

Keith K. Niall

# Johannes von Kries: Principles of the Probability Calculus

A Logical Investigation

# **Studies in History and Philosophy of Science**

## **Volume 59**

### **Series Editor**

Catherine Abou-Nemeh, History Programme, Victoria University of Wellington, Wellington, New Zealand

### **Advisory Editors**

Rachel A. Ankeny, University of Adelaide, Adelaide, SA, Australia

Peter Anstey, School of Philosophical & Hist Inquiry, University of Sydney, Sydney, NSW, Australia

Steven French, Department of Philosophy, University of Leeds, Leeds, UK

Ofer Gal, Unit for History and Philosophy of Science, University of Sydney, Sydney, Australia

Clemency Montelle, School of Mathematics & Statistics, University of Canterbury, Christchurch, New Zealand

Nicholas Rasmussen, UNSW Sydney, Kensington, Australia

John Schuster, University of New South Wales Sydney, Kensington, Australia

Richard Yeo, Griffith University, Brisbane, Australia

*Studies in History and Philosophy of Science* is a peer-reviewed book series, dedicated to the history of science and historically informed philosophy of science. The series publishes original scholarship in various related areas, including new directions in epistemology and the history of knowledge within global and colonial contexts. It includes monographs, edited collections, and translations of primary sources in the English language. These cover a broad temporal spectrum, from antiquity to modernity, and all regions of the world.

Keith K. Niall

# Johannes von Kries: Principles of the Probability Calculus

A Logical Investigation



Springer

Keith K. Niall  
Toronto, ON, Canada

ISSN 1871-7381                    ISSN 2215-1958 (electronic)

Studies in History and Philosophy of Science

ISBN 978-3-031-36505-8            ISBN 978-3-031-36506-5 (eBook)

<https://doi.org/10.1007/978-3-031-36506-5>

Translation from the German language edition: “Die Principien der Wahrscheinlichkeitsrechnung” by Johannes (deceased) von Kries, © J.C.B. Mohr 1886. Published by J.C.B. Mohr. All Rights Reserved.

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Switzerland AG 2023

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG  
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

***Dedication by the Translator***

*To the friends of my youth who died of AIDS:  
I returned, and saw under the sun, that the race  
is not to the swift, nor the battle to the strong,  
neither yet bread to the wise, nor yet riches to  
men of understanding, nor yet favour to men  
of skill; but time and chance happeneth to  
them all.*

*Ecclesiastes 9:11.*

# Translator's Preface

Der ist nicht klug, der vieles wagt,  
Geringen Vortheil zu erwischen.  
Dieß heisset, wie August gesagt,  
Mit einem güldnen Angel zu fischen.  
(von Hagedorn (1764). ‘An einen Freund’, p. 100.)

Researchers in psychology and researchers in medicine don't often stop to ponder the foundations of probability: they have much else to do. Their neglect of foundations is not a product of disdain; they rest assured that mathematicians have taken care of the important questions already. Probability is anything which satisfies **Kolmogorov**'s (1933/1950) axioms – and if one collects sufficient data, it makes sense to compute probabilities on those observations, correct? The idea that probabilities are not to be computed everywhere may seem novel. The idea that some probabilities are not to be expressed in numbers at all may seem fanciful. When they apply statistics, researchers often skip over a few notions inherited from study of the foundations of probability. Three stand out: *comparability*, *originality*, and *independence*. Rather than being syntactic or formal properties of the mathematical study of probability, they are semantic properties (they are defined in the text). Sometimes even those responsible for teaching statistics in psychology and medicine will skip over such semantic notions in their curriculum. They skip over them, in a haste to introduce canonical correlation or multivariate analysis of variance to their students in the space of a semester or two. Then to turn a proverb: act in haste, repent of replicability at leisure. When the computation of probability is so easy, what point can there be to any discussion of conditions under which probability may not be expressed in numbers?

The problem does not lie with mathematics; there is nothing unsound in the proliferation of statistical techniques in mathematics. For a long time, the notion of probability has been misapplied to observations in psychology and medicine (among other social sciences). Observations in these disciplines have been interpreted as the results of games of chance – which is to say as results from a domain where there is reason to believe the notion of probability can be applied. Though the problem may

seem viciously circular, the conceptual issue is that we still do not understand other disciplines well enough to say where the notion of probability applies, and where it does not. The issue is not that great masses of data are lacking. Instead, it is that some basic characteristics requisite to probability have not been revealed or specified in those disciplines, or else adequate conventions are lacking. The characteristics in question occur but rarely – they are far from omnipresent. Masses of mathematics will not help the situation, as much as will the cautious use of *exploratory* analysis in specific domains of psychology and medicine. There is no guarantee such efforts will prove to be productive; such patient reflections on assumptions about data have been applied before in mathematical psychology, particularly in the study known as psychophysics. Yet the effort to establish psychophysics has failed consistently for many of the same reasons as basic notions of probability fail to apply in other domains of psychology. However grandiose a proposition it may sound, the present work offers an overview of the nature of probability, and the right way the notion of probability ought to be applied in these empirical disciplines. It does so with extended narrative and examples, and without many formulas.

Johannes **von Kries**'s theory of probability was first published in 1886 and reprinted in 1927 as *Die Principien der Wahrscheinlichkeitsrechnung*. Chapters 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10 of the present volume are a translation of that text. The two Prefaces of those volumes also appear here in English translation. Two chapters have been added. Chapter 11 translates most of Chapter 26 of von Kries (1916), which is his book on logic. That chapter is meant to summarize the original 1886 work, and to reply to subsequent criticisms (notably those by **von Stumpf**). Here Chap. 12 translates an 1882 article by von Kries, on the idea of a psychophysical scale. The material in Chap. 12 had appeared in translation as **Niall** (1995); it is reprinted here with permission. Von Kries claims that often probabilities may be mischaracterized as numeric when they are not, for many of the reasons he gives in the 1882 article for the incoherence of **Fechner**'s notion of a psychophysical scale. The use that von Kries makes of the notion of convention in measurement is also made clear there. Notes which appeared as footnotes in the original have been translated and collated after Chap. 12. Passages which first appeared in French or in Latin have been translated into English in the main body of the text. The original untranslated passages appear together with von Kries's Notes. The positions of Notes in the text are indicated by successive Arabic numerals for each chapter. I find the translation of von Kries's work significantly more difficult than, say, translation of works by his contemporary Erwin **Schrödinger** (cf. Niall 2017). I do not mean to say that von Kries is unclear; instead, his use of sentence structure in German can be rich and intricate.

Von Kries made a large contribution to the theory of probability: his remains a viable theory – even though it is expressed largely in narrative form. (The first full-blooded equation occurs in Chap. 5, Section 3.) Von Kries gives some pointers *how* to read the book, at the end of his Preface to the first printing. His theory emerged in the nineteenth century – it doesn't offer techniques new to us – but it does speak to our understanding, especially since it anticipates formal developments in probability

theory, such as the method of arbitrary functions. The goal of his theory is *propae-deutic*; it provides a general orientation to ideas which guide the application of probability toward reasonable understanding. Here is a sort of introductory manual which tells us when to apply techniques derived from probability theory, and *when not*. Probabilistic reasoning is used in the exploration of data and in their analysis, in the twenty-first century as in the nineteenth. Probabilistic reasoning can go awry because people misunderstand basic concepts of probability. Such profound misunderstandings have occurred throughout the development of the mathematical theory of probability (see Chap. 9, Sections 13 and 14 on **Poisson**, and the whole of Chap. 10 on the history of probability theory). One is better prepared to apply probabilistic reasoning by avoiding these confusions, having understood von Kries's account. "Probability is still used uncritically, ignoring the assumptions underlying its main interpretations. For that reason, investigation of the foundations of probability, its meaning, and the assumptions made by each of its interpretations still looks like a useful exercise" (Galavotti 2017, p.11).

The force of von Kries's theory of probability is that it guides conceptual change, rather than developing new computational formulae; he seeks to change our way of thinking about probability. One contemporary example he emphasized comes from **Boltzmann**'s (1880, 1881a, 1881b) statistical theory of gases. Von Kries upends general notions which emerge from that theory – for example, that "more organized states transition to less organized ones." It has been mentioned how von Kries anticipates the method of arbitrary functions (see Hopf 1934, 1936 for a systematic account). "About all this von Kries is quite clear, although he lacks some modern terms like 'density function' or 'bounded variation'" (Rosenthal 2016, p. 154). He reinterprets some central notions in probability theory, examples being the law of large numbers and the interpretation of **Bayes**'s principle: notions which are by no means unimportant today. For the latter, it is important to distinguish the legitimate use of prior distributions summarizing genuine knowledge, and the use of uniform priors to summarize unknowing (Zabell 2022, p. 1). Von Kries develops another important idea which is not fully appreciated: that is the relation between the independence of observations and the dispersion of distributions. There is an intimate relation between von Kries's philosophy of probability and developments in the philosophy of induction. Similar to von Kries, **Kneale** (1949) has a range theory of probability which approaches issues in the philosophy of induction. As he puts it (op. cit. p. 211): "It is now time to discuss the fundamental question whether the probability which attaches to the conclusion of an induction can be brought within the theory of chances. I wish to maintain that it cannot. If this is the right answer to the question, all such attempts to justify induction... must be mistaken exercises in ingenuity." The *Principien* is not the only work by von Kries which bears on the theory of probability. Other articles of note are: von Kries (1888, 1892, 1899, 1914, 1916, 1919); an overview of his publications may be found in **Buldt** (2016). As **Zabell** (2016, p. 148) puts the matter: "we are now – at last – coming to gradually recognize the philosophical riches to be found in von Kries's neglected masterpiece."

Certainly, the book has been praised. **Kamlah** (1987, p. 110) says: “We are now able to see why his *Principien*... was the most intelligent and sophisticated book on probability in Germany before World War I.” And von Kries has been praised as “the most cogent and influential of the German philosophers who discussed probability in the nineteenth century” (**Shafer & Vovk** 2018, p. 11). It is safe to say that his book has been underappreciated, even when its ideas have been borrowed (such as by Max **Weber**: see **Heidelberger** 2010, p. 254 and **Treibler** 2015). Zabell (2016, p. 132) makes a strong case for its influence: “von Kries remains today overall one of the least known and most underestimated contributors to the foundations of probability.” The book’s one time obscurity may be blamed on its style of narrative exposition, or on the unavailability of fitting mathematical tools at the time (Rosenthal 2012, p. 228). Still “von Kries’s anticipation of **Poincaré**’s method of arbitrary functions cannot fail to impress even the mathematician” (Zabell 2016, p. 147; **Poincaré** 1912). Perhaps von Kries’s ideas have been too subtle or too controversial in the way they bridge the logical and physical aspects of probability, or perhaps no one knew how to enact them (**Shafer & Vovk** 2018, P. 11). So Kamlah (op. cit. p. 110) claims that “only today in light of some later statistical mechanics is it easy to understand what von Kries was talking about.” Part of the book’s obscurity may also arise from its position outside the English tradition of the philosophy of probability. At least one English author studied von Kries’s *Principien*: John Maynard **Keynes** (see **Fioretti** 1998, 2003). He acknowledges von Kries’s rejection of the “Principle of insufficient reason” – a principle which Keynes attributes to James **Bernoulli**. Keynes (1921, p. 44) gave the principle the name by which it is now familiar: the “Principle of indifference.” “The Principle of indifference asserts that if there is no known *reason* for predication of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an *equal probability*” (op. cit. p. 45). Keynes also says “in order that numerical measurement may be possible, we must be given a number of equally probable alternatives” (p. 44), and the Principle of indifference is a means to that end. Keynes saw that the principle is faulty, based on criticisms advanced by von Kries in the *Principien*. Von Kries supposes that the Principle of indifference (or of insufficient reason) is a poor fiction which has nothing to do with the actual course of ordinary thought, and which has little place in the theory of probability. Zabell (2005, p. 129) claims “one of the most effective parts of Keynes’s *Treatise* is its criticism of the principle of indifference.” The criticism illustrates the futility of applying operations like congruence to assignments of value on such premises in actual situations without further consideration. It is futile to assign probabilities to empirical premises, and to expect that comparisons or combinations of those estimates may be made by rules of mathematics, unless *for other reasons* one knows independently that such rules of comparison or combination do hold.

There are two related morals in von Kries’s account of probability. One is logical: that in general the Principle of indifference does not apply (even in the guise of uniform priors for Bayesian analysis). Its habitual application arises from a misunderstanding of foundational concepts of probability. The other moral is physical or ontological: there are no bare particulars. (**Hacking** 1971, p. 341 says that the very

meaning of “equally possible” is an equivocation between logical possibility and physical possibility.) The perennial examples used to introduce statistical concepts (the urn, the roulette wheel) are more impoverished than ideal. Without invoking natural kinds, it may happen that after more sophisticated analysis, things we have counted as similar items of a series are simply not similar. We can assume that we make observations of things which are similar, but often enough they turn out to be things which are neither similar nor independent. It is unsuitable to cast them as a series of bare particulars. The assumption is faulty unless explicitly justified: it can become a questionable means to smooth a path to numeric measurement. The “independence” of cases may mean an equal possibility of combination, but “it is always and only reasonable to treat a series of similar cases as independent in light of a definite bit of knowledge” (from Chap. 11). The explicit justification of our assumptions may be instrumental in saying why the Principle of indifference has seemed at all compelling. Those intuitions may be explained in terms of ranges: “The principle of indifference may be seen as the consequence of putting the following constraint on the assignment of probabilities to a set of mutually exclusive outcomes: *Symmetries in the probability distribution ought to mirror symmetries in our knowledge concerning the outcomes.*” (Strevens 1998, pp. 7–8). As Zabell (2005, p. 30) says: “Symmetry arguments lie at the heart of probability. But they are tools, not axioms, always to be applied with care to specific instances rather than general propositions.”

The determination of probability by ranges may occur anywhere that our knowledge is incomplete. More than a simple theory of physical probability, von Kries’s range theory of probability incorporates human competence; von Kries indicates that the theory supersedes Laplace only where it is necessary to ask about such competence. The determination of probability by ranges may occur anywhere our knowledge is incomplete. It is in this sense that von Kries bridges the subjective and the objective: we have competence to judge the bounds of what we know, and to gauge ratios of ranges, or their parts. “To put it very briefly, the probability of an event is the proportion of those initial states within the initial state-space attached to the underlying random experiment which led to the event in question, provided that the space has a certain structure” (Rosenthal 2010, p. 72). Yet we decide what counts as the same for a purpose, or what may count as equal in range. Our expectations may *develop*: the assessments we make of probability are dynamic rather than static, even for enduring scenarios. There may be no single concept of probability that we must apply in advance, where any is applicable (cf. Cohen 1989, p. 113).

Kamlah (1987, p. 110) claims that “von Kries was led to the study of probability by his wish to defend the thesis that psychological quantities are not measurable.” In the theory of probability, the Principle of indifference is a means to establishing equal possibility or equal probability. Von Kries raises the objection that such equality is a matter of convention. Without the explicit or public establishment of such a convention, probability is simply not amenable to measurement. “As far as the choice of the measure is a matter of convention, as far as the initial-state space has no ‘naturally-built-in’ measure, or class of measures that are equivalent regarding the proportions of the outcomes, the probabilities we get out of the range

approach are merely conventional as well" (Rosenthal 2010, p. 90; see also Roberts 2016). That is the connection between von Kries's rejection of the Principle of indifference and his repudiation of Ernst **Weber** and Gustav **Fechner**'s "psychophysics" for the psychology of the senses (von Kries 1882; Chapter 12). In psychophysics, the "just-noticeable difference" (or *jnd*) indicates the minimum change in a physical stimulus (say, the intensity of a light) which gives rise to an impression of difference (in brightness). Equal *jnds* along a scale of stimulus intensity are meant to define a scale of sensation. "Although generations of psychologists have managed to convince themselves that the equal-*jnd* assumption is plausible, if not obvious, it is not and never has been compelling; and in this respect, an equal-ratio assumption is not much different" (Luce 1963, p. 70). Von Kries insists this equality is conventional: the equality of *jnds* along a scale is at best conventional. "Likewise, we can only speak of the measurement of sensations once we have established an arbitrary convention that determines what we shall consider as equal" (von Kries 1882, p. 294; Niall 1995, p. 301). In the psychological domain, an unwarranted notion of equality has been applied in an attempt to codify a system of measurement where none may be reasonable. Von Kries does not consider the social sciences (or research on the efficacy of outcomes in medicine) to be amenable to treatment in terms of probability. Much the same holds for econometrics when he says (Chap. 7, Section 10). "There is no definite and universally valid principle according to which we might compare the value of goods in our possession to the value of a prospect or the hope of obtaining such a thing."

A central question of the book is: what is subjective and what is objective in probability? The distinction between subjective and objective is often drawn simplistically. Though "subjective" is correlative to "objective," both have multiple, often tacit meanings. "Subjective" can be reduced to a term of abuse, along the lines of "perhaps unreliable" or "too much a function of the particular but unascertainable psychological conditions" (Anscombe 1981, p. 48). There may be an aspersion that what is subjective is incorrigible and useless, meaning "there is no room for correctness." The term "objectivity" provides no refuge: at one time "if something were called an object that would have raised the question 'object of what?'" (Anscombe 1981, p. 3). The term "object" is now more often used as a *prosortal*. Prosortals are placeholders for sortal terms, rather than being extremely general sortal terms. (Sortal terms in grammar classify entities by kinds.) They are empty in their abstraction. As such they avoid ontological commitments – they enable a retreat into obscurity. In his interpretation of von Kries's theory, Rosenthal (2010, p. 76) abandons the subjective altogether, calling it "a proposal for an objective interpretation of probability statements, i.e. as providing truth conditions for such statements that do not depend on our state of mind or our state of information and are therefore called 'objective'." In a different twist, Pulte (2016, p. 124) calls the elements of von Kries's configuration spaces (configuration – Spielräume) *objective* but not *realized*. Zabell (2016, p. 147) puts forward a less controversial opposition of logical and physical, in place of the less palatable opposition of subjective and objective. Even that may be controversial: what counts as purely physical? "Subjective and objective" and "logical and physical" capture something of the opposition put forward by

von Kries in the *Principien*. I suggest that the poles of that concept could also be expressed as “possible and actual.” That recognizes legitimacy in what counts as possible, and it leaves room to define what counts as actual. As **Goodman** (1955, p. 56) puts the matter in his analysis of counterfactual conditionals, “What we mistake for the actual world is one particular description of it. And what we mistake for possible worlds are just equally true descriptions in other terms. We have come to think of the actual as one among many possible worlds. We need to repaint that picture. All possible worlds lie within the actual one.” A probability theorist may have a cut-and-dried description in mind for a problem set (though it may change and develop). Randomness can be just what falls outside the description we have ready to hand. Goodman (1955, p. 87) makes an accessory point when he says: “while confirmation is indeed a relation between evidence and hypotheses, this does not mean that our definition of this relation must refer to nothing other than such evidence and hypotheses.”

Given Kolmogorov’s axioms, the study of probability seems to have been fully arithmetized, so that a central question then becomes: “Why apply the *Lebesgue measure* to a state space arising from a *linear mapping* of *standard physical quantities* onto mathematical space in order to determine the outcome probabilities?” (Rosenthal 2016, p. 165). That question is not superfluous, nor is it wrong, but it may be secondary in precedence. Many interesting and basic questions are left behind. Hacking (1965, p. 8) points out how “Text-books persistently repeat the idea that the formal theory conveyed by Kolmogoroff’s axioms is a model for frequency in the long run. If this does not conceal an actual confusion, it at least makes it easy to ignore problems about long run frequency.” Later he adds (op. cit. p. 27): Kolmogoroff’s axioms, or some equivalent... do not even determine which of a pair of hypotheses about a given distribution is better supported by given data.” Von Kries emphasizes that “Mathematicians – who are mainly occupied with the computational methods of our discipline – have largely neglected the philosophical establishment or basis of the discipline, or else they have often explicitly refused the discussion” (cf. the Preface to the second printing). For him, “the specification of probabilities cannot be a general aim of research in any sense: instead, there are only very special objective conditions which enable that task” (Chap. 9, Section 14).

The psychologist Edwin **Boring** appreciated von Kries’s account of probability. He appreciated its rejection of the Principle of indifference, in saying: “Knowledge simply does not come out of ignorance” (Boring 1920, p. 30). Not only did Boring recommend the theory: he urged psychologists to study it despite their other concerns, saying: “No amount of practically successful ‘mental measurement’ in laboratories, school-systems, factories or the army can relieve us, if we do not wish to waste time, of the necessity of stopping, every so often, to take account of first principles.” Von Kries’s foundational theory of probability is well worth the time and effort that psychologists (or others) can afford.

## References

- Anscombe, G.E.M. 1981. *Metaphysics and the philosophy of mind. Collected papers, vol. 2.* Oxford: Basil Blackwell.
- Boltzmann, L. 1880. Zur Theorie der Gas-Reibung I. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 81(2): 117–158.
- . 1881a. Zur Theorie der Gas-Reibung II. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 84(2): 40–135.
- . 1881b. Zur Theorie der Gas-Reibung III. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 84(2): 1230–1263.
- Boring, E.G. 1920. The logic of the normal law of error in mental measurement. *The American Journal of Psychology* 31(1): 1–33.
- Buldt, B. 2016. Johannes von Kries: A bio-biography. *Journal for General Philosophy of Science* 47(1): 1–19.
- Cohen, L.J. 1989. *The philosophy of induction and probability.* Oxford: Clarendon Press.
- Fioretti, G. 1998. John Maynard Keynes and Johannes von Kries. *History of Economic Ideas* 6(3): 52–80.
- . 2003. No faith, no conversion: The evolution of Keynes's ideas on uncertainty under the influence of Johannes von Kries. Chapter 10 of: J. Runde & S. Mizuhara. *The philosophy of Keynes's economics: Probability, uncertainty and convention.* London: Routledge.
- Galavotti, M.C. 2017. The interpretation of probability: Still an open issue? *Philosophies* 2 (20): 13.
- Hacking, I. 1965. *Logic of statistical inference.* Cambridge: Cambridge University Press.
- . 1971. Equipossibility theories of probability. *The British Journal for the Philosophy of Science* 22(4): 339–355.
- Hagedorn, F. von 1764. *Sämmtliche poetische Werke. Erster Theil.* Hamburg: Johann Carl Bohn.
- Heidelberger, M. 2010. From Mill via von Kries to Max Weber: Causality, explanation, and understanding. Chapter 13 in: U. Feest, Ed. *Historical perspectives on Erklären und Verstehen.* (New Studies in the History of Science and Technology: Archimedes, vol. 21). New York: Springer, 241–265.
- Hopf, E. 1934. On causality, statistics and probability. *Journal of Mathematics and Physics* 13: 51–102.
- . 1936. Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 46: 176–195.
- Kamlah, A. 1987. The decline of the Laplacian theory of probability: A study of Stumpf, von Kries, and Meinong. Chapter 5 of: L. Krüger, L.J. Daston, and M. Heidelberger, Ed. *The probabilistic revolution, vol. 1: Ideas in history.* Cambridge: MIT Press, pp. 91–116.
- Kneale, W. 1949. *Probability and induction.* Oxford at the Clarendon Press.
- Kries, J. von 1882. Ueber die Messung intensiver Größen und über das sogenannte psychophysische Gesetz. *Vierteljahrsschrift für wissenschaftliche Philosophie* 4(3): 257–294.
- . 1886. *Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung.* Freiburg i.B: J.C.B. Mohr (Paul Siebeck). (Reprinted by Mohr in 1927).
- . 1888. Ueber den Begriff der objectiven Möglichkeit und einige Anwendungen desselben. *Vierteljahrsschrift für wissenschaftliche Philosophie* 12(2) erster Artikel, 179–240; 12(3), zweiter Artikel, 287–323; 12(4), dritter Artikel (Schluss), 393–428.
- . 1892. Ueber Real- und Beziehungs – Urtheile. *Vierteljahrsschrift für wissenschaftliche Philosophie* 16: 253–288.
- . 1899. Zur Psychologie der Urteile. *Vierteljahrsschrift für wissenschaftliche Philosophie* 23(1): 1–48.
- . 1914. *Immanuel Kant und seine Bedeutung für die Naturforschung der Gegenwart.* Berlin: Springer.

- \_\_\_\_\_. 1916. *Logik: Grundzüge einer kritischen und formalen Urteilslehre*. Tübingen: J.C.B. Mohr (Paul Siebeck). Chapter 26: pp. 595–636.
- \_\_\_\_\_. 1919. Ueber Wahrscheinlichkeitsrechnung und ihre Anwendung in der Physik. *Die Naturwissenschaften* 7(1–2): 2–7.
- Kolmogorov, A.N. 1950. *Foundations of the theory of probability*. New York: Chelsea Publishing Company. (originally published in 1933 as *Grundbegriffe der Wahrscheinlichkeitsrechnung*).
- Luce, R.D. 1963. On the possible psychophysical laws. In *Readings in mathematical psychology*, ed. R.D. Luce, R.R. Bush, and E. Galanter, vol. 1, 69–83. New York: Wiley.
- Niall, K.K. 1995. Conventions of measurement in psychophysics: von Kries on the so-called psychophysical law. *Spatial Vision* 9 (3): 1–30. The journal *Spatial Vision* continues as *Multisensory Research*.
- \_\_\_\_\_. 2017. *Erwin Schrödinger's color theory, translated with modern commentary*. New York: Springer.
- Pulte, H. 2016. Johannes von Kries's objective probability as a semiclassical concept. Prehistory, preconditions and problems of a progressive idea. *Journal for General Philosophy of Science* 47: 109–129.
- Rosenthal, J. 2010. The natural-range conception of probability. In *Time, chance and reduction. Philosophical aspects of statistical mechanics*, ed. G. Ernst and A. Hüttemann, 71–91. Cambridge: Cambridge University Press.
- \_\_\_\_\_. 2012. Probabilities as ratios of ranges in initial-state spaces. *Journal of Logic, Language and Information* 21: 217–236.
- \_\_\_\_\_. 2016. Johannes von Kries's range conception, the method of arbitrary functions, and related modern approaches to probability. *Journal for General Philosophy of Science* 47: 151–170.
- Shafer, G., and V. Vovk. 2018. The origins and legacy of Kolmogorov's Grundbegriffe. Working Paper #4, The Game-Theoretic and Finance Project, Rutgers School of Business. <https://doi.org/10.18550/arXiv.1802/06071>.
- Strevens, M. 1998. Inferring probabilities from symmetries. *Noûs* 32: 231–246.
- Treibler, H. 2015. Max Weber, Johannes von Kries and the kinetic theory of gases. *Max Weber Studies* 15(1): 47–68.
- Zabell, S.L. 2005. *Symmetry and its discontents: Essays on the history of inductive probability*. New York: Cambridge University Press.
- \_\_\_\_\_. 2016. Johannes von Kries's *Principien*: A brief guide for the perplexed. *Journal for General Philosophy of Science* 47: 131–150.
- \_\_\_\_\_. 2022. Fisher, Bayes, and predictive inference. *Mathematics* 10: 1634, 16 pp.

# Preface to the First Printing

In publishing the investigations which follow, I had a pair of goals in mind. For one, I wished to draw the attention of philosophers to certain concepts: concepts whose logical significance appeared to me noteworthy. They turn out simply to be basic concepts of the probability calculus. First among them is the concept of a *Spielraum* (range), followed by a number of other notions which prove requisite to the intellectual significance of these *ranges*. These concepts have been wanting until now: if no one has been able to derive a satisfactory account of the probability calculus – including its methods and findings – from abstract logic, *that* is the central reason. In considering the undoubtedly large importance given to the calculation of probability in many disciplines, it may be deemed a desirable thing for logic to pay attention to this characteristic principle, as it will be applied here. This desire would appear even better motivated, should it turn out that the probability calculus reveals itself as only one special case of a very general principle. A clear grasp of that principle provides us a not inconsiderable insight to the logical bases of knowledge. Understandably the lack of a satisfactory logical foundation has made itself felt already, since it has been very troublesome in application of the probability calculus. In fact until now, systematic expositions of the probability calculus have contained a diverse mix of true and false propositions, as of valuable and worthless ones. A delineation of these from those has often been thought necessary and has repeatedly been attempted. I intended my *second* goal to be at least the rough sketch of such a distinction. This can be achieved without difficulty, once one arrives at a satisfactory understanding of the principles. Essential clues are offered which guide a proper and accurate application of the probability calculus to various domains of investigation.

The nature of the subject is such, that several passages of the investigation run along conceptual paths that will be easily followed only by readers readily acquainted with mathematical formalisms. Several of these sections (Chaps. 5 and 8) can be skipped without affecting one's understanding of the main issues. That is not the case for the exposition given on pages 38 through 74 [here, roughly Chap. 2, Section 4 through Chap. 3, Section 8], whose content is more or less essential to all that follows. Since some readers may find this very section hard to understand, the

following observation may be apt: the most important results can be assimilated even if one skims over the strict exposition contained in these sections, and one settles for – should I say – a more popular reading. Something similar happens with almost any concept defined by mathematical terms in scientific theory. Why and in what sense two pressures or two forces can be compared quantitatively, or how potential energy can be compared with kinetic energy, is not entirely comprehensible without strict definitions for these concepts. Still a certain understanding of their meaning and use may very well be achieved nonetheless. If someone wishes to achieve such a limited perspective on the probability calculus, the above-mentioned sections can be replaced by the following brief comments. One should be aware of the direct link to be made with the results set out on pages 36 and 37 [here the latter part of Chap. 2, Section 3].

The constraint which emerges in the numeric representation of probabilities is, first of all, that different possible premises extend over comparable ranges. For certain relations, this constraint is satisfied. Consider the universal set of configurations of constraining circumstances which lead to the throw of a die in a game of dice. Consider those which lead to a throw of one; likewise consider all those which produce a throw of two, or three, etc. It may be claimed these six complexes all have the same magnitude. Further it may be claimed that this relation may be understood in full generality, if we focus on conditions that obtain immediately before the die comes to rest, or if we focus on an earlier state of the constraining circumstances at an arbitrary point in time. Our expectations will still be fixed by the equipotency of circumstances if we pose this question: What previous behavior of the die led to any of the current arrangements? (Their relative magnitude is “original.”) Finally it may be shown that our expectations of one outcome over another are drawn from the scope of circumstances in each respective complex. They are not to be gauged by any other type of logical relation (the ranges are “indifferent”). Then the probability of each of the six possible throws should really be called equal in a strict sense. Something similar obtains for games of chance generally, as it does in many other situations. The most important constraint here is that even very small changes in constraining conditions produce a change in outcome. So when one considers a wider range of variation in these circumstances, arrangements which represent various outcomes will occur in regular succession. Any one of them can be produced by many different arrangements of circumstances. – I believe that with this help, and without deeper justification, one will still accept the idea that determinate ratios of magnitude can be given under certain specific assumptions for the scope of different conditions of arrangement or grouping. This ratio lends measure to our expectations; our assignment of numbers to probabilities expresses nothing else. Attainment of this perspective is enough for the reader to follow the derivation of the most important results, particularly the accounts given in Chaps. 4, 6, and 7.

Freiburg im Breisgau, Germany  
February 1886

Johannes von Kries

## Preface to the Second Printing

The first printing of this version of the work appeared 41 years ago. A reprint seemed indicated, since there was continued demand for the book – naturally not a very large demand, given its nature. After 1920 that demand could no longer be met. There was no question but that this reprint was to be in the form of a photomechanical reproduction. As a consequence, any addenda or changes in the text were ruled out ahead of time. I had to restrict myself to introducing the new imprint with a short preface by way of a few short comments. This might have been enough, all the more since I have held onto the main concepts – basic concepts which were developed in the text – without any noteworthy changes. Yet it seemed advisable to say something here about the set form of the text. It seemed advisable, in part to refer to several publications of my own where I have developed the concepts further, and where I have pushed them in new directions. In part it also seemed advisable to address briefly several positions taken by other authors, some of whom affirm and some of whom dispute these notions.

The first of the aforementioned basic concepts runs as follows. One can characterize two intuitions – which have long competed – shortly, to the point: the propositions of the probability calculus have been given either a *subjective* or else an *objective* sense. Both are justified in particular circumstances, and the two senses have a certain connection. Naturally the relative probability of two expectations or two premises is always something subjective. Yet a numeric evaluation is admissible only if there is specific knowledge of an objective referent, that is to say of the *Ranges* of constraining circumstances by which one expectation or another is shaped and made manifest. This means that the two ranges stand in a definite proportion to which numbers can be assigned, and that this ratio of magnitude is known to us. This holds for games of chance – for ideal games of chance it holds strictly. The quantifiable property of games of chance can be recognized most easily in the game of roulette. There the circumstance which determines the outcome – the force of the initial impulse – is unknown ahead of time, within broad bounds. But then extremely small changes in the impulse are enough to change the outcome, to put red where there was black. The arrangement of the colored pockets ensures that

substantial changes in initial propulsion represent a regular and periodic alternation of the two outcomes: red and black. This depends on the condition that the range of forces can be called “equal” which lead to one outcome rather than another, in an unbiased way across equal extents of red and black stripes. In an ideal game of chance, this would hold strictly for stripes of infinitely small width.

It is correct to call the proportion of ranges of the class of circumstances which determine one outcome over another – the ratio of their *objective possibilities*. Then the calculation of probabilities is possible only where objective possibilities of different outcomes stand in some numerically expressible ratio, and where that ratio of possibilities is known to us.

I have also examined these notions in the framework of a more comprehensive work (my *Logic*,<sup>1</sup> which is concerned with the theory of probability from the perspective of logic (*Logik*, Chapter 19, p. 422 ff.; Chapter 26, p. 595 ff.) [Most of the latter chapter is translated as Chap. 11 here.] In that work there was opportunity to present divergent accounts – especially one propounded by von **Stumpf**,<sup>2</sup> which proceeds from disjunctive judgment, and which is essentially built on the principle of insufficient reason (the Principle of indifference). Contrary to that intuition, it may be shown that the necessary foundation for application of the probability calculus is only ever given if certain relations obtain between abstract representations of the elements of a disjunction. For the rest, I must forego a review of that side of the literature, since it would be impossible in the space available here to consider the matter productively.<sup>3</sup> Still, the following comments may be in order. If logicians have shown themselves to be very reticent about the Range theory, perhaps this is the reason: they have not thoroughly understood either the general epistemological foundations of mathematics, or measurement theory in general, as I believe those subjects must be understood. Actually I believe that only those who understand all of mathematics in a way that basically fits with **Kant**’s approach will understand this. They will succeed in seeing a larger context for the circumstances to which the numeric assessment of probability is tied; perhaps only then the assessment of probability can be fully understood. I developed these intuitions about mathematics in my volume on logic (*Logik*, Chapter 1, p. 15 ff.). Here I should allude to the fact

<sup>1</sup> von **Kries**, J. (1916). *Logik: Grundzüge einer kritischen und formalen Urteilslehre*. Tübingen: J.C.B. Mohr (Paul Siebeck).

<sup>2</sup> von **Stumpf**, C. (1892). Ueber den Begriff der mathematischen Wahrscheinlichkeit. *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften, philosophisch-philologische Classe*, 37–120.; von **Stumpf**, C. (1892a). Ueber die Anwendung des mathematischen Wahrscheinlichkeitsbegriffes auf Teile eines Continuums. *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften zu München, philosophisch-philologische Classe*, 681–691. (p. 691)

<sup>3</sup> I regret this all the more, because the most significant work of its kind in the literature (**Meinong**, A., 1915, *Über Möglichkeit und Wahrscheinlichkeit: Beiträge zur Gegenstandstheorie und Erkenntnistheorie*. Leipzig: Johann Ambrosius Barth) was published just before my book on logic, so that it is not considered there either.

that I have written a short summary of these intuitions in my treatise on Kant.<sup>4</sup> They culminate in the statement that “*the object of mathematics consists of the internal relations of those contents of consciousness which represent a concatenation of elements, all exactly the same in kind.*”<sup>5</sup> What is meant by this concatenation – and in what sense elements are to be called the same in kind – escapes general explanation, and may only be clarified by illustration with individual examples (representations of number, time, and space). The work of Lourié<sup>6</sup> has already received a brief mention in my book on logic; it may be considered one attempt to approach problems of the computation of probability through purely logical considerations.

Mathematicians – who are mainly occupied with the computational methods of our discipline – have largely neglected the philosophical establishment or basis of the discipline, or they have even explicitly refused the discussion. **Von Mises** is a prime example.<sup>7</sup> However, if that author hopes “to have made the greater number of such discussions superfluous by the unequivocal establishment of the probability calculus *as a discipline of mathematics*” and “to evade those interminable wide-ranging discussions which turn on the definition of ‘equal possibility’,” then I confess I am no longer able to follow him in the matter. After all, purely mathematical concepts are not the object of the probability calculus: real events and processes are its object. For **von Mises**, it follows he also believes the probability calculus to be “a science with the same standing as geometry or theoretical mechanics.” He proceeds from initial assertions of the existence of a “threshold value” as if this were “a fact of experience supported by innumerable observations.” Yet what the probability calculus needs by way of foundation<sup>8</sup> is not individual facts which have actually been experienced but rather inductive *generalization* of them. So here we stand before a “philosophical question,” and thus we arrive at the crux of an old

<sup>4</sup> **von Kries**, J. (1914). *Immanuel Kant und seine Bedeutung für die Naturforschung der Gegenwart*. Berlin: Springer.

<sup>5</sup> Op. cit. p. 35.

<sup>6</sup> **Lourié**, S. (1910). *Die Prinzipien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung des disjunktiven Urteils*. Tübingen: Mohr (Paul Siebeck).

<sup>7</sup> **von Mises**, R. (1919). Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 5, 52–99. (p. 52) I think I can restrict myself to citation of this one author, since his exposition of the probability calculus has received especially wide recognition from among mathematicians. Often it seems to be taken as a model of its kind. To delve into the treatment of our subject by **Czuber**, by **Polya**, by **Bruns** as well as many others, would lead well outside the bounds of this investigation (cf. Czuber, 1884, 1902, 1914).

<sup>8</sup> Admittedly, mathematicians can restrict themselves – here as elsewhere – to the formal or mathematical development of implications that may be produced from specific premises which have been introduced as postulates. Except for this: if any real meaning is to be claimed for the probability calculus, then the attempt to prove the *cogency* of those premises will in some sense still be indispensable. Of course mathematicians can dispense with this task for themselves, and that might even be very useful to a clean division of labor for such tasks. But the investigation must be pursued somewhere at least, be that in physics, be that in logic, or on whatever shore a proof may wash up. We should not kid ourselves, of course, that this treatment by mathematics may not signify the postponement of a task rather than its resolution.

antinomy. Should we consider such a regularity – that makes direct reference to collective phenomena and which only appears through them – as final in the same sense as the lawfulness of nature? Should we regard our intellectual requirements as having been satisfied insofar as they can be satisfied, or do we feel an irrepressible demand that we ought to look for another foundation, another explanation? **Von Mises**'s basis for the probability calculus has cogency only if one answers the *former* question in the affirmative. Only then will his foundation seem at least necessary and sufficient. **Von Mises** has neither excluded nor circumvented the philosophical question, then. Rather he has answered the question in a quite definite way, if a tacit and perhaps unconscious way. At least he has taken a stance on the philosophic question. – Yet the close contemplation of what we may assume about the lawfulness of nature, and what we can establish of its essential framework, has consistently led a great many thinkers to answer the question in the *latter* way (as an irrepressible demand) in the affirmative. What kind of law excludes strict and exact formulation, or admits any old exception as possible just when exceptions seem very improbable? That cannot be the lawfulness of reality, which we seek to comprehend as a final goal. And it has always been thought necessary to base the relative frequency of many observations on some circumstance that can be demonstrated in every case; approximately equal frequencies occur only when an equal possibility is given to one outcome or other in an individual case. We have not been in a position to declare what “equally probable” means. Thus even along the path that **von Mises** has blazed, naturally we encounter an old stumbling block since this concerns such hard questions as the conceptual imperative for a valid form of inference to be drawn. Here we stand before an unbridgeable chasm – before a confrontation of outlooks, perhaps one which will never be spanned completely. In my opinion, this difficulty may only be resolved by an idea, one which is fundamentally different from the lawfulness of nature but which has equal standing. And so we come to the other basic idea to be introduced here. It begins with the fact that lawful regularity does not cover the real world in all its entirety and in all its particulars. Rather, the conduct (the behavior) of things in the real world is not subject to lawful regularity in every respect; some conduct is discernable only in concrete. We need to distinguish *nomological* from *ontological* determinants of the behavior of things in the world. It may happen that an event is produced by the ontological behavior of real things, which make for a range of a majority *fraction approaching one* of behaviors which are possible in some respect, but which are not covered by the totality of laws, while it follows as a consequence that the realization of a contrary event is assigned a vanishingly small fraction. If we can demonstrate this point, then the regular occurrence of the former event even seems comprehensible. Then the intellectual imperative we feel in the face of an observable regularity will be satisfied: we should recognize that the regularity may be warranted not only as a *lawful regularity* of nature, but we should also recognize that it may follow if as a consequence the regularity represents a *far greater part of the manifest range*. At the same time, the corresponding expectation for future cases will also have adequate warrant. The warrant holds even if it is not authorized by the general lawfulness of nature as in most cases, but must be authorized by appeal to another principle. This holds for all

those regularities which come to light on a macroscopic scale as collective phenomena. That is because a mathematical theorem can be transposed directly to apply to the magnitudes of ranges: that is the mathematical theorem we are given to call “the Law of large numbers” in discussion of probability. If the outcomes *red* and *black* represent *equal* ranges of constraining conditions for every spin of the roulette wheel, then an approximately equal frequency of the two outcomes is represented by a fraction of almost one in the range which consists of *all* combinations of these outcomes. It is not a lawful regularity of nature that explains why many spins of the wheel produce about the same frequency of red and black in roulette. Such an explanation is unnecessary, because this result is brought about by the prevailing range of real configurations. Smoluchowski<sup>9</sup> came to the probability calculus from the study of *games of chance*, as did I. By a close examination of games of chance he arrived at results remarkably close to mine, without knowing my work. On many points, his results correspond exactly with mine. Perhaps it lay outside the scope of his thought to develop a complete theory of mathematical probability from these results; at least that was not his intention.

An important application of the probability calculus is found in *theoretical physics*. From the scientific literature, it may not be said with any certainty how often my views have found acceptance or rejection among academic representatives of that discipline. The discipline has not yet undertaken a treatment of the one question which seems fundamental to me in this respect. Namely, that is the question I alluded to earlier in a general way: whether observable regularities of the diffusion of heat – including all that is involved in the kinetic theory of gases – find an adequate foundation in a general formulation of laws of nature (which are indispensable in any case), or whether as in the regularities we observe in ordinary games of chance, we need to augment them by invoking logical principles of quite a different nature – the Range principle. The accounts given by von Planck and Nernst (among others) do not take a clear position on this question, as has been mentioned. Nonetheless Planck inclines to draw a strict division that contrasts “statistical laws” with other laws of nature. That division approaches my present intuition, and – by deeper consideration of the analogy to games of chance – it could lead (so far as I can see) inevitably to the Range principle.

I treat these relations in some detail, in an article that deals with the application of the probability calculus to theoretical physics.<sup>10</sup> Several points are represented there (the derivation of Boltzmann’s Law concerning the distribution of speeds of molecules in a volume of gas; the demarcation of certain problems yet to be solved

---

<sup>9</sup>Smoluchowski (1918) Ueber den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik, in dem Herrn M. Planck gewidmeten Festheft der Naturwissenschaften. April, p. 253. [Smoluchowski, M.V. (1918). Über den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik. *Die Naturwissenschaften*, 17(7), (April, 6th year), 23 pp. Max Planck zur Feier seines 60. Geburtstages, pp. 195–264.]

<sup>10</sup>von Kries, J. (1919). Ueber Wahrscheinlichkeitsrechnung und ihre Anwendung in der Physik. *Die Naturwissenschaften*, 7(1–2), pp. 2 & 17.

by the mathematical theory), and the logic of the argument is perhaps both clearer and fuller than here in the *Principles*.

At this juncture, I should mention two articles by **Reichenbach**<sup>11</sup> as an investigation which proceeds from theoretical physics, and which attempts to develop a general theory of probability. I have not yet had the opportunity to consider his comprehensive account of the subject; this would not be the place to do so.

In *Principles of the probability calculus*, I have shown that the formal relations which hold for games of chance are evident only in part, and with significant exceptions, for the *collective phenomena of human society*. For this reason, application of the probability calculus can lead to large errors or blunders if not constrained by the careful exercise of critical thought. I have the impression that such wanted care in application of the probability calculus has been exercised more and more in past decades, at least in the field of medical statistics and its accessory applications. Perhaps the reason lay with previous error-laden methods, whose barrenness and misleading nature was immediately apparent in the results derived from them. Insofar as theoretical justification may be wanted in this regard, it may be given in the form of a statement that commonly-applied concepts were taken from games of chance (bounding conditions, atomic vs. molar collective phenomena, etc.). Perhaps that may even be enough. The elevated probabilities which count toward a therapeutic outcome may be entirely subject to masking illusions in face of the slightest doubt, whether any of the states which may be called general conditions of the course of disease have been constantly present, or whether states have undergone unrelated changes (apart from the treatment which was introduced). Later, in my *Logik* (Chapter 19, p. 420 ff.), I set out the logical and formal relations of collective phenomena, which are described only incompletely as those for games of chance. There I set out a description which may be more satisfying than that found in the *Principles*. Namely in the *Logic* it is shown how laws of observed collective phenomena become comprehensible only if one brings to bear *both* principles: that of the general lawfulness of nature, but also the Range principle.

My ideas have received an unexpected amount of attention from the domain of *jurisprudence*. That is to say the concept of the range and of its place in probability calculus have given rise to a series of concepts which are important to the legal doctrine of causation. This has given secure support and accessible meaning to the notion of *abetting circumstances*. In that context, certain distinctions prove to have been drawn correctly, which had been devised in jurisprudence but which had been provided no satisfactory explanation: as a result, they had often been dismissed as

---

<sup>11</sup> **Reichenbach** Die physikalischen Voraussetzungen der Wahrscheinlichkeitsrechnung. *Naturwissenschaften* 1919, pp. 2 & 17. [Reichenbach, H. (1920). Die physikalischen Voraussetzungen der Wahrscheinlichkeitsrechnung. *Die Naturwissenschaften*, 8(3), 46–55.] – Derselbe, Philosophische Kritik der Wahrscheinlichkeitsrechnung. Ebenda, 1920. p. 146. [Reichenbach, H. (1920a). Philosophische Kritik der Wahrscheinlichkeitsrechnung. *Die Naturwissenschaften*, 8(8), 146–153.]

baseless. As I showed in the précis devoted to this topic,<sup>12</sup> it is of legal import when an outcome is brought about by a blameworthy act (i.e., an act without which it would not have occurred) if the act is generally likely to bring about such an outcome; if the act represents an abetting circumstance for the outcome, increasing its probability; or else if that is not the case. I proposed that we speak of an *adequate* cause in the first place, and of an *accidental* cause in the latter. This is a distinction to which jurisprudence has given close examination.<sup>13</sup>

At that time I did not think I had expressed some new idea by establishing (initially in the *Principles*) a sharp distinction between nomological and ontological determinants of reality. Rather I thought this expressed something that is often assumed implicitly, and especially often in mathematical physics. Often there, for example, a law appears in the form of a differential equation that describes a motion, and ontological constraints appear as an integration constant. That idea is close to the distinction which **Windelband** drew later on, between nomothetic sciences and idiographic sciences. Afterward, it was promulgated by **Rickert** as the distinction – applied so extensively and so productively – between the natural sciences and humanities. I returned to that initial starting point later myself under more of a formal, logical perspective (*Logik*, Chapter 1, p. 53). There I showed that the contrast between the nomological and the ontological can only be applied with caution in characterizing whole areas of knowledge: it cannot be drawn in a schematic way (Chapter 23, p. 508 ff.).

A review of the literature that has appeared since 1886 provides a still more impressive picture than I could have sketched at the time, of the widely ramifying significance of the circle of concepts surrounding the probability calculus. So far as I can see, the ideas that I developed to apply to all those many diverse problems have found approbation only in part, and with limited scope. It is unclear whether this will change in future. One thing speaks in favor of the expectation: until now there has been no other way we have found to deal with this whole complex of problems in any really satisfying way. But even if this expectation should not be sustained (and who would wager that wholesale recognition of any particular idea will become necessary – who would ever wager on a prediction of that certainty...), nonetheless it will appear a useful step forward, to have pointed out the issue which is the single unifying foundation for a great many topics that seem so widely disparate.

Freiburg, Germany  
May 1927

Johannes von Kries

---

<sup>12</sup>von Kries, J. (1888). Ueber den Begriff der objectiven Möglichkeit und einige Anwendungen desselben. Separatabdruck aus der *Vierteljahrsschrift für wissenschaftliche Philosophie*, **12**(2-4), 179–240; 287–323; 393–428. (Fues' Verlag, R. Reisland)

<sup>13</sup>From the uncommonly wide-ranging literature on this topic, I would cite M. **Rümelin**, *Civil Archiv* **90**, p. 171 f.; Träger, L. (1904). *Der Kausalzusammenhang im Straf- und Zivilrecht* Marburg: Elwert.

## Acknowledgments

I would like to thank Elizabeth Wener, Foreign Rights Manager, Rechte und Lizenzen, Mohr Siebeck Verlag, for her clarification of copyright issues (message of August 30, 2018). I would like to thank the publisher Brill for permission to re-use Brill content (Brian Johnson, rights@brill.com, message of July 4, 2022), as reprinted in Chap. 12 from Niall, K.K. (1995). Conventions of measurement in psychophysics: von Kries on the so-called psychophysical law. *Spatial Vision*, 9(3), 1–30. All translations from Latin are the work of Drs. Sarah H. and Elliot J. Rossiter, who I would like to thank for their generosity in allowing the free use of their translations. Drs. David Nicolas and Gail Edwards provided advice and encouragement in the process of writing.

# Contents

|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>The Meaning of Probability Statements . . . . .</b>   | <b>1</b>  |
| 1.1      | Difficulty About the Interpretation Given to Reports<br>of Numeric Probability . . . . .   | 1         |
| 1.2      | Their Psychological Interpretation . . . . .   | 2         |
| 1.3      | Relation of Propositions About Probability to Propositions<br>of Practical Reason . . . . .  | 4         |
| 1.4      | The Logical Interpretation. Principle of Insufficient Reason<br>(Principle of Indifference). Insufficiency of the Latter . . . . .                                 | 4         |
| 1.5      | Empirical Interpretation. Partial Analogies . . . . .  | 10        |
| 1.6      | Concept of Reports of Probability as Statements of Empirically<br>Given Regularities. Distinguished from Descriptions Which<br>Express Lawful Connection . . . . . | 12        |
| <b>2</b> | <b>The Convention of Equally Warranted Premises . . . . .</b>  | <b>15</b> |
| 2.1      | Connection of Expectations to Relations Among Ranges.<br>Indifference of Ranges and the Free Formation<br>of Expectation . . . . .                                 | 15        |
| 2.2      | Non-numeric Probabilities . . . . .  | 16        |
| 2.3      | Comparability and Originality. Formulation of the Range<br>Principle . . . . .   | 19        |
| 2.4      | Probability in the Context of the Temporal Coincidence<br>of Independent Processes . . . . .   | 23        |
| 2.5      | Configurations of Ranges . . . . .   | 24        |
| 2.6      | Originality of Ranges with Reference to the Lawful<br>Connections Between Conditions and Consequences . . . . .  | 27        |
| <b>3</b> | <b>The Theory of Games of Chance . . . . .</b>   | <b>31</b> |
| 3.1      | Objective Ratios of Numbers and Extents in Games<br>of Chance . . . . .  | 31        |

|          |   |           |
|----------|---|-----------|
| 3.2      | The Bowling – Game. Dependence of Assignments<br>of Probability on the Assumption of Constant Probability<br>for the Effects of Arbitrary Motions . . . . . | 32        |
| 3.3      | Extension of the Same Intuition to all Games of Chance . . . . .  | 35        |
| 3.4      | The Interpretation of These Propositions About Probability<br>in Terms of the Range Principle . . . . .   | 38        |
| 3.5      | Indifference . . . . .  | 39        |
| 3.6      | Comparability Within the Smallest Ranges . . . . .  | 39        |
| 3.7      | Justification of Premises About Constant Probabilities . . . . .  | 43        |
| 3.8      | Formulation of the Conditions for Numeric Reports<br>of Probability . . . . .   | 45        |
| <b>4</b> | <b>The Special Theory of Probability . . . . .</b>  | <b>47</b> |
| 4.1      | The Objective and the Subjective Meaning of Propositions<br>About Probability . . . . .   | 47        |
| 4.2      | Numeric Probability as a Case of a Logical Constraint Which<br>is Only Approximately Realized . . . . .   | 49        |
| 4.3      | Independence of Various Cases . . . . .   | 51        |
| 4.4      | Difference Between Nomological and Ontological<br>Determinants. Ontological Ranges of Behavior.<br>Concept of Objective Possibility . . . . .               | 53        |
| 4.5      | The Law of Large Numbers . . . . .  | 55        |
| 4.6      | The Concept of Likelihood . . . . .   | 56        |
| 4.7      | Generally Valid Probabilities. The Notion of Chance.<br>Determination of Probability . . . . .  | 58        |
| 4.8      | Randomness. General Conditions and Their Random<br>Arrangement . . . . .  | 59        |
| 4.9      | Dispersion. The Relation of the Connection Between<br>Individual Cases as a Condition of Normal, Hypernormal,<br>or Hyponormal Dispersion . . . . .         | 63        |
| 4.10     | The Meaning of the Totality of Possibilities. Equality<br>of Chance for Individual Cases . . . . .  | 66        |
| <b>5</b> | <b>Varieties of Numeric Probability . . . . .</b>   | <b>69</b> |
| 5.1      | Relation of the Probability of Premises to the Probability<br>of a Conclusion . . . . .   | 69        |
| 5.2      | The Probability of Connected Premises . . . . .   | 70        |
| 5.3      | <b>Bayes's Principle . . . . .</b>  | <b>72</b> |
| 5.4      | Partially Admissible Ranges . . . . .   | 73        |
| 5.5      | Formation of Numeric Probability by the Characteristic<br>of the Condition of Admissibility . . . . .   | 75        |
| 5.6      | Logical Equivalence in Form of All Numeric Probabilities . . . . .  | 77        |

|   |            |
|---|------------|
| <b>6 Establishing and Justifying Probability Statements . . . . .</b>   | <b>81</b>  |
| 6.1 The Justification of Statements About Probability.  |            |
| The Difference Between This and Mere <i>A Posteriori</i>  |            |
| Determination of a Numeric Value . . . . .  | 81         |
| 6.2 Deductive Justification in Detail . . . . .   | 83         |
| 6.3 Summary Deductive Justification . . . . .   | 84         |
| 6.4 Empirical Proof of the Constancy of General Conditions . . . . .  | 86         |
| 6.5 Empirical Establishment of the Dependence or Independence<br>of Individual Cases, from the Dispersion of Ratios . . . . .                                 | 88         |
| 6.6 Empirical Investigation of Individual Cases and Their Equal<br>Chances of Occurrence . . . . .  | 88         |
| 6.7 The Orthodox Procedure . . . . .  | 89         |
| 6.8 Tracing Phenomena Back to Their General Circumstances . . . . .   | 90         |
| 6.9 Judgments About Individual Cases . . . . .  | 93         |
| 6.10 Expectations for Many Cases of the Same Kind . . . . .   | 94         |
| <b>7 On the Significance of the Range Principle, and the Probability<br/>Calculus . . . . .</b>   | <b>97</b>  |
| 7.1 Logical Significance of Ranges in the Absence<br>of Comparability in Numbers . . . . .  | 97         |
| 7.2 The Range Principle as a Definitive Principle of Expectation.<br>Its Relation to the Principle of the Lawfulness of Events . . . . .                      | 98         |
| 7.3 Explanation of Regularities Which Occur in Collective<br>Phenomena . . . . .  | 102        |
| 7.4 Non-empirical Nature of the Range Principle . . . . .   | 104        |
| 7.5 Its Indispensability. Certainty of the Expectations<br>Based on it . . . . .  | 105        |
| 7.6 The Interpretation of Computations Based on Ranges.<br>The Exploratory Use of the Probability Calculus . . . . .  | 106        |
| 7.7 Relation of the Probability Calculus to Conclusions Drawn<br>by Analogy . . . . .   | 108        |
| 7.8 The Interpretation of Probability Computations Based<br>on Inexact or Uncertain Premises . . . . .  | 109        |
| 7.9 The Estimation of Arbitrary Probabilities . . . . .   | 110        |
| 7.10 Relation of Numeric Probability to Pragmatic Rules . . . . .   | 112        |
| <b>8 Application of the Probability Calculus to Theoretical Physics . . . . .</b>   | <b>117</b> |
| 8.1 Application of the Probability Calculus . . . . .   | 117        |
| 8.2 Task of the Probability Calculus in Physics. Relation<br>of Physical Theory to Logical Principles . . . . .   | 118        |
| 8.3 Basis for Enlisting the Probability Calculus in Explanation<br>of Certain Phenomena . . . . .   | 119        |
| 8.4 The Putative Transition from Less Probable to More<br>Probable States. Originality of Ranges and the Constancy<br>of Assignments of Probability . . . . . | 121        |

|           |  |     |
|-----------|--|-----|
| 8.5       | The Real Meaning of <b>Maxwell's Law</b> . . . . .   | 123 |
| 8.6       | Conclusion to the Theory of An Ideal Case. Compelling Derivation of Assignments of Probability . . . . .   | 127 |
| 8.7       | Task of Extending the Theory . . . . .   | 129 |
| 8.8       | Results . . . . .  | 132 |
| <b>9</b>  | <b>More Applications of the Probability Calculus</b> . . . . .   | 133 |
| 9.1       | The Theory of Errors in Observation. Composition of Total Error From Many Elementary Errors . . . . .  | 133 |
| 9.2       | General Conditions for Individual Elementary Errors. Empirical Investigations Concerning Their Constancy . . . . .   | 135 |
| 9.3       | Independence and the Equal Chances of Successive Observations . . . . .  | 137 |
| 9.4       | <b>Gauss's Law of Error</b> . . . . .  | 138 |
| 9.5       | Results Concerning the Theory of Error . . . . .   | 140 |
| 9.6       | The Collective Phenomena of Human Society. Meaning of the Overall Ratios . . . . .   | 142 |
| 9.7       | Empirical Investigations Concerning the Constancies of the Latter. Typical Series and Problematic Series. Numeric Determination of the Certainty of Inferences to Prior Situations . . . . .   | 143 |
| 9.8       | Incorrect Application of the Probability Calculus in This Domain . . . . .   | 145 |
| 9.9       | The Exploratory Application of the Probability Calculus to Statistical Results . . . . .   | 148 |
| 9.10      | Medical Statistics. The Importance of the Probability Calculus in the Justification of Statements About Therapeutic Value . . . . .  | 149 |
| 9.11      | Views of <b>Fick</b> and <b>Liebermeister</b> on the Subject . . . . .   | 151 |
| 9.12      | Jury Verdicts. Conforming and Dissenting Decisions. Impossibility of Determining Their Relative Frequency . . . . .  | 153 |
| 9.13      | <b>Poisson's</b> Investigations on the Probability of Judgments . . . . .  | 157 |
| 9.14      | General Characterization of the Subjects Which Call for Application of the Probability Calculus . . . . .  | 158 |
| <b>10</b> | <b>On the History of Probability Theory</b> . . . . .  | 161 |
| 10.1      | Beginnings of the Probability Calculus with <b>Pascal</b> and <b>Fermat</b> . Probability Defined by ‘values’ . . . . .  | 161 |
| 10.2      | Initial Calculation of the Value of Annuities. <b>Halley</b> . . . . .   | 162 |
| 10.3      | Development of the Probability Calculus by the Mathematical Academy. Jacob and Daniel <b>Bernoulli</b> . The Article ‘ <i>Probabilité</i> ’ Appears in the <i>Encyclopédie</i> . <b>Laplace</b> . The Principle of Insufficient Reason (Principle of Indifference) . . . . . | 163 |
| 10.4      | Characterization of the Orthodox Method as the Treatment of all Series of Phenomena by the Schema Applied to Games of Chance . . . . .   | 167 |

|       |   |     |
|-------|---|-----|
| 10.5  | Opponents of the Academic Approach of Mathematics,<br>in the Last Century. <b>d'Alembert. Prévost. Béguelin</b> . . . . .   | 170 |
| 10.6  | The Requirement to Distinguish Different Kinds of Probability.<br><b>Fries. Cournot</b> . . . . .   | 172 |
| 10.7  | <b>Fick's Account of the Objective Meaning of Statements<br/>About Probability</b> . . . . .  | 175 |
| 10.8  | The Investigations Carried Out by <b>Lexis</b> . . . . .  | 177 |
| 10.9  | The Perspective of Modern Logic. The Force of the Subjective<br>Interpretation of Propositions About Probability. <b>Lotze.</b><br><b>Sigwart. Windelband.</b> The English Logicians. Starting Points<br>of My Own Investigations . . . . . | 177 |
| 10.10 | Sketch of a Form to Represent the Probability Calculus,<br>Suitable for the Classroom . . . . .   | 182 |
| 11    | <b>On Probability Theory</b> . . . . .  | 185 |
| 12    | <b>Conventions of Measurement in Psychophysics</b> . . . . .  | 219 |
|       | References . . . . .  | 240 |
|       | <b>Bibliography</b> . . . . .   | 243 |
|       | <b>Index</b> . . . . .  | 249 |

# Chapter 1

## The Meaning of Probability Statements



**Abstract** Though the computational aspects of probability theory are well-developed, the foundations of probability theory have not been well articulated. It is unclear what sense is to be made of numeric probability. Every report of numeric probability hangs on a notion of ‘equal possibility’. Probability statements represent some knowledge of objective relations, rather than knowledge of lawful connection. There are notions which do not capture what we mean by probability in numbers: inductive conclusion by analogy is one. There is no measure of probability as psychological expectation, just as there is no measure of the intensity of sensation. Also, numeric probability cannot be based on the Principle of indifference (here called the Principle of insufficient reason). Two examples are given to reinforce these claims: the probability of a meteor striking the earth’s surface, and the probability that terrestrial elements exist on a distant star.

**Keywords** Discrete probability · Logic of probability · Conclusion by analogy · Equal possibility · Abetting conditions · Principle of indifference

1. The principles of the probability calculus are the subject of a not inconsiderable number of investigations, both of recent and of early date. Despite that, renewed treatment of these many well-considered and often-discussed issues does not seem supernumerary at the outset, because previous efforts have not really arrived at any conclusive and fully satisfactory result. To be specific, perhaps it ought to count as generally accepted that the widest circle of literature which touches on the computation of probability offers only very unsatisfactory answers to the important questions. Certainly one finds that the computational aspect of probability theory has been developed to a high degree, especially in articles written by mathematicians. By contrast the foundations of the whole subject have mostly been handled in a way which staggers through obscurity, and which engenders all manner of doubt. Even the efforts of philosophers have not relieved this deficiency. At least none of them have offered a theory which might have been up to the task of furnishing a popular preamble. Rather, as a glance at the literature tells, even in the most recent publications a host of contradictory intuitions are represented. So for example, the concept of objective probability or that of physical possibility figures in numerous articles –

especially those which deal with the application of computational methods to restricted domains. Quite often the same concept is designated expressly as being both important and indispensable, adding that it is essentially distinct from and opposed to the concept of subjective probability. Logicians take a different position, and say that possibility in general is a subjective business about how a thing behaves one way or another, as is any probability. Following their argument, it seems inadmissible and wrongheaded to lend these concepts objective meaning. When as in a recent textbook<sup>1</sup> types of probability are introduced as subjective, objective, and *a posteriori*, one has to concede that the mere introduction of such categories raises a great many questions without answering a single one. One does not even learn if by this the three types should be indicated by distinct statements about probability, or if there are only three methods indicated for different cases, by which one may formulate propositions which are anyhow equivalent in meaning.

Doubtless the main source for all this tentativeness lies in this: we are not sufficiently clear what meaning may be attached to a numeric account of probability. Surely we ought to be able to declare – in an entirely satisfying and exhaustive way, without fuss – what may be meant if we assign a probability of  $\frac{1}{6}$  to the probability of one throw of a cubic die coming up 4. Yet so long as doubt remains on this point, of course nothing certain can be made of several central issues, such as the conditions to which application of the probability calculus is tied. As a consequence we select the starting point of our investigation so that we may best ascertain the meaning which can be given to probability statements expressed in terms of numbers. At least initially we will restrict ourselves in this examination to those statements which can undoubtedly hold true, and which ought to count as typical examples for the probability calculus. In other words we will restrict ourselves to statements which have some relation to so-called games of chance. We should like to proceed as follows: to run through a whole series of concepts that we encounter. We examine a number of concepts to determine of each one, if it proves to be sound, and to what extent it may be called sound.

**2.** If we look to discover the sense of a proposition about probability that has been described, then our first thought is that it ought to express something about an actually existent mental state. That is, it ought to express the subjective degree of certainty with which a particular individual believes something or expects something at a particular moment. Let us call this its psychological interpretation. This much is sure: we can hold the widest range of subjective certainty or ‘conviction’ about the content of any arbitrary judgment. That ranges from zero, when the correctness or incorrectness of a premise is all the same to us. An innumerable series of graded steps leads from there to the unshakeable conviction with which we assert our own existence, or with which we expect the sun to rise tomorrow.

---

<sup>1</sup>A. Meyer, *Vorlesungen über Wahrscheinlichkeits-Rechnung*, translated by Czuber from the French, Leipzig, 1875. [Meyer, A. (1879). *Vorlesungen über Wahrscheinlichkeitsrechnung*. Leipzig: B.G. Teuber (E. Czuber, Trans.)]

Yet on first consideration this subjective certainty is not at all a topic of measurement. Rather in this case, all the objections pertain which I have made elsewhere<sup>2</sup> against the psychological measurement of intensive magnitudes. [See Chap. 12.] Some differences and comparisons of degree may be admitted for those magnitudes, but an assignment of numbers is ruled out, because they do not represent the concatenation of uniform elements of a single kind. Then we can say that *A* may be more probable, even much more probable than *B*. Still no reasonable sense is to be attached to the claim that one of them may be doubly or ten times as probable as something else. It is perverse to look for this kind of equivalence in numeric terms. This can be seen clearly once the certainties of things are in question, which are not subject to description another way in terms of probability. So it might be asked: how probable is it for me that there will be war between England and Russia next year? The actual political situation might be completely ignored in response to this question: it could simply be a matter of my present state of belief and expectation as it stands at this moment, without reference to its justification. Just as there is no measure of the strength of a sensation, so a measure of the intensity of expectation does not exist – there is no measure for this degree of conviction.

Let us put aside the question of a theoretical rationale for such a process of measurement. Then it is still easy to show that ordinary numeric statements of probability cannot be interpreted in this psychological fashion. It follows that a lengthy series of factors remains neglected in the evaluation of probability. Those factors do exert a significant influence in psychological terms on one's actual degree of expectation. So the hope with which we expect a good thing – or the apprehension with which we expect a bad thing – are both dependent in large measure on individual temperament, on fleeting mood, or on other similar things. No notice is taken of these factors in the calculation of probability. Everyone acknowledges, e.g., that the probability of throwing a 12 with a pair of dice is 1/36. One can specify this without consideration of special conditions; one is used to thinking of this as wholly and generally valid. Now suppose we indicate that there is a large prize for a throw of 12, and that for many people the certainty with which they expect a throw of 12 is augmented by their associated hope. Then our unconditional assessment would be that this statement of probability has nothing to do with the measurement of psychological states as such. It is in no way the task of the probability calculus to assess in what degree the certainty of hopeful optimists may overestimate their rosy chances as they anticipate good fortune, or in what degree the frightened may overestimate their darker chances as they anticipate some misfortune. Instead, the subject of a convention can only be this: there is a certainty with which “something may reasonably be expected”.

---

<sup>2</sup>von Kries, J. (1882). Ueber die Messung intensiver Größen und über das sogenannte psychophysische Gesetz. *Vierteljahrsschrift für wissenschaftliche Philosophie*, 4(3), 257–294. [Translated in: Niall, K.K. (1995). Conventions of measurement in psychophysics: von Kries on the so-called psychophysical law. *Spatial Vision*, 9(3), 1–30. The journal *Spatial Vision* continues as *Multisensory Research*. DOI: <https://doi.org/10.1163/156856895X00016>]

If we concentrate on the negative result which has been gained here, at least we could say that probability statements contain no expression of the psychological state present at a given time for a given individual.

**3.** Let us examine what may be a second interpretation: that sense which was just mentioned by way of contrast to the psychological sense, which imputed information about certainty to the probability statement, and which might be expected in any relatively reasonable account. This too proves untenable, at least in the sense to be described below. It may be considered very contentious, if we are not thought to expect an event with greater certainty – *ceteris paribus* – in that we are better prepared for the event when it is undesirable than if it is favorable. Far-ramified factors seem to demand similar consideration: as in the precautionary steps we may take to prevent or produce those events, or the preparatory measures that the occurrence or absence of those events may call for, and so on. If we continue with such speculation, we find we are forced to declare that the ‘reasonable account’ which was just taken up in explanation of probability statements, cannot be meant in the practical, ethical sense which calls upon the necessary use of conditions which are neither sharply definable nor easily judged. Pragmatic rules are much too subjective and arbitrary, that one might imagine their representation to be the business of the probability calculus. And this account would be excluded just as the psychological account was, particularly by its neglect of all the same kinds of factors which are not of equal validity in our actions, and for the general validity which we confer exclusively in games of chance on our understanding of propositions about probability.

**4.** At this point one can attempt to grasp the rules – those which guide expectations and which seemed to be represented in probability statements – in a logical sense rather than a pragmatic sense. Consequently the numeric values which appear in those statements would no longer be considered the measure of an existent psychological state. Neither would they be something demanded by pragmatic rules; rather they would be considered a logical relation. And the story how this determination in numbers might be understood is not a difficult story to tell. Every report of numeric probability either consists in, or at least may be traced to, the fact that a number of distinct cases are declared ‘equally possible’. In addition it is established that such-and-such a large portion of all equally possible cases is characterized by a certain property, let us say by the satisfaction of some outcome or another. In fact the whole enterprise of the computation of probability may be seen as the establishment and enumeration of such equally possible cases. Then if the logical operation of our understanding may be seen in the presentation of a number of equally possible cases, an explanation follows easily: that two or more cases are to be considered equally possible, if in the present state of our knowledge there is no reason to call one of them more probable than another. – If as appears at first, any incomplete sort of knowledge licenses such arrangements and their decomposition, then there arises the opportunity that we might characterize the probability of this supposition or that expectation, et cetera, by the entirely definite application of a numeric measure.

Let us give this interpretation a short name: it is the logical interpretation. The principle it employs as a basis for the probability calculus is called the Principle of insufficient reason [a.k.a. Principle of indifference]. At first this appears altogether satisfactory. It may also be the principle which has found greatest acceptance among accounts of the probability calculus. That is the reason we need to deal with this principle in some detail.

The former two interpretations founded on a single concern; here by comparison that same consideration is dismissed easily. Of course there is no getting away from the fact that the subjective coloration and the significant uncertainty of the Principle do contrast markedly with the apparent generality and the precision of probability statements that pertain to games of chance. If the balance of expectation at issue here is somehow modulated by dependence on a momentary state of knowledge, then it may well happen that  $a$  and  $b$  are equally possible for  $X$ , while  $Y$  – whose information may be more complete – is in possession of a reason to declare that  $a$  is more probable than  $b$ . And suppose for example, as we read in the work of an English logician: “If  $A$  and  $C$  are wholly unknown things, we have no reason to believe that  $A$  is  $C$  rather than that it is not  $C$ ; the antecedent probability is then  $\frac{1}{2}$ .” Then only with difficulty are we able to refrain from the conjecture that the probability spoken of in games of chance derives its measure from some other basis than this. It may be thought valid that human knowledge and human ignorance should always be counted essentially the same, just considering the great complexity of conditions in games of chance. They could also be essentially the same, in that the logical constraint on expectations of the various outcomes is always the same. Despite the factors mentioned above – that are different from case to case, and which underlie human judgment – it is entirely characteristic that they influence strength of expectation according to laws of psychology. Perhaps they might also influence it according to rules of practical reason. Yet these factors do not stand in logical connection to the thing which is expected. In fact on closer examination, one is forced to admit – for example – that when dice are thrown, under no circumstance does anyone have a reason in the logical sense to expect one outcome over another.

Certainly there can be no basis for doubt that the definition just posed is a necessary one for equally possible cases in the probability calculus. When we assume equally possible cases, surely there can be no basis given to justify calling one case more probable than another. As illuminating as this may be, on further consideration another insight forces itself on us as indispensable: the insight that this explanation is not sufficient. In no way does the simple absence of a basis for making a distinction suffice to constitute equally possible cases in the sense required for the probability calculus.

Our knowledge of the conditions of any event can be incomplete in many different ways, and to many different extents. If we know that equally many black and white balls are contained in an urn, then we assign an equal probability to obtaining a white or a black ball on a single draw: both probabilities are counted as  $\frac{1}{2}$ . Then this assignment of probability might also be justified under our principle, if we knew only that there are black and white balls in the urn, while their numeric ratio is entirely unknown to us. Yet only the former case represents the ordinary and

completely unobjectionable general procedure which is recognized in the probability calculus. It is also easy to see that the intellectual moves which are made in these two cases are essentially different. In the former case in particular, the convention of an equal probability of drawing a white ball or a black ball finds an altogether positive justification in a ratio which has been established objectively, and which is known to us. In the latter case, that is not demonstrable in the same measure. When we investigate examples of the latter sort more closely, soon we find that the assignment of equally possible cases quite often runs into peculiar difficulties. To settle on a particular example, let us consider an event which by its nature can happen at any place on earth, like a meteor impact. A question is posed about the expectation for the probability of the event of such an impact on some part of the earth. We want to presume quite a thorough state of ignorance about the proportions in question. Let there be someone who is in utter ignorance of anything about meteor impacts. Further let the person's knowledge of the earth's surface be that the earth is divided into a certain number of parts  $n$ , perhaps by counting the names of countries and seas: he would know nothing about their extents. If the subject person is asked: which part of the earth's surface will be struck by a meteor?, he has no reason to expect that it will be greater for any of these  $n$  parts than for any other – for Spain more than for Denmark, or for Mexico more than for the North Sea. He would have to assign the probability of impact to be equal for each of the  $n$  parts, in other words to be  $1/n$ . One may think what one wants about the justification for assigning a specific certainty to such a solution with this expectation: a very basic objection can be raised against this whole approach – an objection which on closer examination renders the approach completely illusory. As long as the measurement of probability is guided by such complete ignorance as we have postulated, the identification of equally possible cases is simply arbitrary.

We can assemble larger or smaller numbers of parts of the earth's surface into more general categories than were considered initially. Then we can achieve any division at all, let us say one for the five continents and the named oceans. The initial convention (of countries and seas) will produce a certain probability of meteor impact for each of the regions. This might be recast as the number of parts of the former kind which each region contains, because the case of impact on such a larger region is more general, and includes several specific cases. The probability of the general case could be set equal to the sum of probabilities of all the individual cases it contains. Then we would obtain very different probabilities of impact for Europe, for Asia, etc. It seems clear that the new division can be chosen as the starting point for our investigation with just the same warrant as the earlier division. When we say there is no reason to expect a meteor impact in one region over another, that holds for continents just as well as for individual countries. By that reasoning we would have to assign equal probabilities to each of the continents, but unequal probabilities to countries which belong to the various continents. Initially it cannot be claimed that one way of considering the matter has any advantage over the other. – Similar situations are encountered very frequently; I restrict myself to introducing another example which is framed somewhat differently.

Let the question be posed as follows: does some celestial body (say Sirius) contain a certain terrestrial substance (say iron) or not? We are completely ignorant of the composition of this star, — we can imagine ourselves back in history when there were no methods of spectral analysis — and we have no reason to prefer the conjecture that Sirius has iron over the conjecture that Sirius lacks iron. Then we assign the probability of each of these two premises to be =  $\frac{1}{2}$ . Subsequently we inquire about the presence of gold, and the same thing applies. One should also note that both questions — and all successive questions which make reference to different elements — are quite independent of one another. The probability of the presence of any other substance can always be assigned =  $\frac{1}{2}$  in turn, even if the presence or the absence of one or more other substances may be determined already. According to familiar rules, we establish the probability that Sirius lacks both iron and gold as =  $\frac{1}{4}$ . If we proceed along these lines, then we obtain a truly minimal probability that all 68 terrestrial elements are not present in Sirius, or else a truly huge probability approaching complete certainty that the star contains at least one terrestrial substance. The unfevered imagination strains not to concede the force of such an argument, I believe. But the value of this argument is also invalidated right away, by comparison with the simple assessment of alternatives we are able to make in advance and without further ado: that either Sirius contains terrestrial elements or it does not. We may even consider these two premises as equally warranted as well, and hence their probability seems to be =  $\frac{1}{2}$ , while previously their probability seemed to be near one, falling short of a value of one only by a vanishingly small fraction. Now in a case like this, there are surely all kinds of reasons to declare that the former assignment of probability is valid, or the latter is valid, or that quite different assignments of probability could be valid. What remains unquestionable, however, is that this whole procedure of dividing up cases does not seem at all applicable: there is no definite cue to indicate what must be construed as equally possible.

Accordingly it is now evident that a very essential requirement has been left out of the general formulation of the logical principle given above for the probability calculus, that the disposition of equally possible cases has to be established in a compelling manner, not in any arbitrary manner. This means for example that the probability of a general case must always be concordant, whether we evaluate it by direct comparison with coordinate propositions, or if — beginning with the observation of particular cases — we represent it as the sum of probabilities of all the separate cases it covers. The satisfaction of this requirement is far from assured in any general sense if we assign equally possible cases by the Principle of insufficient reason. That is because situations can very easily be construed so that there is no reason to assign different probabilities to cases  $a$ ,  $b$ ,  $c$ ,  $d$ ,  $e$ , and  $f$ , while at the same time we may assemble four cases  $a$ ,  $b$ ,  $c$ , and  $d$  under the rubric of category  $A$ , and contrast that with  $B$ , which contains only  $e$  and  $f$ . Again there is no reason to expect  $A$  over  $B$ . Under these circumstances there is not even any immediately evident and reliable rule for choosing between different assignments of probability. Then it is not an empirical matter that everything which comes to our consideration will always admit some decomposition into a number of definite premises which are indisputably and

evidently equal in value. Rather this is only the case (1) if a certain definite criterion does exist for declaring possibilities equal, (2) if this criterion appears to be met consistently, and (3) if once we have said things are equal, a declaration of their difference counts as false. Only in such cases can we speak of a really valuable representation of logical relations in terms of numeric probabilities. It may happen that there is no reason to hold one thing to be more probable than another – that is, while that may occur in very many ways and with the most varied of materials – still only very specific states of knowledge and ignorance will have the requisite properties which are associated with the assignment of numeric probabilities in these ways. There are only very particular states under which a strict and definite account of equally possible cases may be produced.

A specific conjecture suggests itself here, to show how this epistemological situation may be considered. We are close to making an assumption that we may discern the criterion just postulated for the measurement of probability, in some objective relation of magnitude. In the first example that was introduced – which concerned a partition of the earth's surface – any uncertainty might in fact be dispelled, should we have been able to restrict our judgment of probability to the surface areas of each part as correctly and objectively specified. Any arbitrariness would have been eliminated if we had set the probability contributed by any section of the earth's surface to be proportional to its surface area. In a similar way, the insight – that the determination of probabilities must always be connected to ratios of objective magnitudes – coincides with another notion. That is a notion I suspect would have occurred at least to some readers right at the beginning of the example. That is to say: no useful determination of probability at all may be sustained, when complete ignorance of spatial extents is presumed. Rather this demands a knowledge of spatial extents, by necessity.

We encounter a second difficulty worthy of discussion if we seek to explain certain methods by the Principle of insufficient reason which are considered plainly correct in the theory of games of chance. In dice games it is well known that we assign an equal probability of  $1/6$  to each of the results 1, 2, 3, 4, 5, or 6 for any throw of a single die. And if we calculate probabilities for a sequence of several throws, we begin with a notion connected with arbitrarily many outcomes of previous throws. The outcome of a subsequent throw, whether 1, 2, 3, 4, 5, or 6, remains equally probable for any arbitrary outcome of previous throws. From that it follows that over many throws any arbitrary sequence of outcomes is to be considered equiprobable to any other. The admissibility of this assumption is not at all evident, so long as its justification is taken from the Principle of insufficient reason alone. Let us take the perspective of someone completely ignorant of the relevant conditions which determine the outcomes of dice throws. Then we could not deny that an ordinary rule of analogy would have to be used to evaluate the preponderance of our expectations: that that which has happened once or else many times will be more probable in future than that which has happened seldom or not at all. Yet then a higher probability of a 6 is to be expected on the next throw after a 6 has occurred, as opposed to all other outcomes. And then it is also more probable in advance that the same number will occur twice in a row than two different numbers in succession. Successions of 1 1,

2 2, etc. would be more probable than say, 3 4 or 6 5, etc. In general it has to be that the rule of analogy makes it more probable for us that series which repeat themselves in some way are more probable than series which do not, if nothing is known about the conduct of the throws.

Such a way of considering the matter really does take hold once we fail to ensure certain positive conditions may be assumed, or known outright, which in ordinary situations we are correct to assume. Usually in dice games we proceed on the assumption that the centre of gravity of the die in play coincides at least to a coarse approximation with its geometric centre. If under special conditions an appreciably eccentric location for the centre of gravity is known, or if that location even comes into question, then the justification of the usual convention about probability ceases to hold immediately. Of course if we do not know which face is now nearer to the centre of gravity, the probability of any of the 6 possible outcomes will be the same on the first throw. The situation is different for the outcome of several throws, by contrast. The outcome of the first throw will enable us to make conjectures about the location of the centre of gravity from then on, and expectations of subsequent throws will be modified in that regard. Accordingly, probabilities can be assigned in advance too. For example the probability of the sequence 4 4 4 would not be the same as that for the sequence 4 2 3; instead the former can at least be assumed to be the larger probability.

One is used to reporting the distinction between these cases and ordinary ones by saying that for these cases the existence of abetting conditions for one outcome or another must be taken into account. Ordinarily it can be considered certain that such abetting conditions as favor one of the 6 possible results of throws are in no way present. What does this mean, then, and how are we so sure of that? One will not find that this question can be answered on its own. Only this much is sure: that it is not enough merely to declare that the conditions may be entirely unknown, meaning the conditions on which the outcomes of the throws depend. Rather it is indisputable that there has to be some positive sense attached to the objective relations relevant to games of chance, together with that claim about abetting conditions. For example, in a game of dice at least *inter alia* it must hold that the centre of gravity for each die is not positioned eccentrically. At least in this way it is generally comprehensible that the ordinary rule of analogy finds no application in judging a sequence of several cases. There are simply certain items of knowledge which are capable of swiveling our expectations around to new vantage points. Which objective relation may be said to constitute or not to constitute an abetting condition: there is a question that is joined quite inseparably to our main problem. Then in completely general terms it can only be said that abetting circumstances are those which make the occurrence of an event more probable than the occurrence of another event. We are led to this question by the difficulty just touched upon: what knowledge of objective significance counts as a premise for the assignment of probability in games of chance?

In general, the logical interpretation guides us to the result we found. Of course the explanation is correct in those terms: that two cases are equally possible just when we have no reason to hold one as more probable than the other. That explanation cannot satisfy us, however, because it does not give us a characteristic

property of numeric probabilities. Instead what is to be added is that a reliable criterion must always be present in each instance, of a compellingly definite assignment (and not an arbitrary one) for the equality of possibility. It may be conjectured that such a criterion can only be established by taking objective ratios of magnitude into account. Moreover the justification for conventional assignments of probability in games of chance is quite definite, but it is not necessarily precise and deliverable knowledge of objective import, other things being equal. That would exclude application of the ordinary rule of analogy with respect to expectations over several cases. Subsequent investigation will set us the task of discovering which specific complexes of known facts admit the definite establishment of equally possible cases, as well as the task of discovering the role played by facts which count as objective, and of course whether objective ratios of magnitude actually do play a role. Before we ready ourselves to approach these questions, it is necessary to consider still other attempts at explanation, ones which proceed from very different concepts than the one just discussed.

**5.** One can try to make sense of probability statements, in that they might be considered the direct expression of particular empirical knowledge; in other words, those propositions express what is known to us about the actual course or progress of cases of a certain type observed already. In fact it is easy to see that if we put the question whether a specific outcome will succeed or fail under certain circumstances, then the expectation to be formed must at least be influenced as follows. Say that a number  $x$  of similar cases are known to us. Of those,  $y$  cases have led to the outcome of interest, while the remaining  $x - y$  have not. One might then conjecture that this is the meaning – as nothing else is – of the statement that a probability of  $\frac{y}{x}$  may be attributed to the occurrence of successful outcomes, under those circumstances. The actual meaning of the number attached to the probability was based on past events, and those events known to us.

Anyways, an important concordance is revealed between the logical relation which occurs here and the one characteristic of the probability calculus, as definite numeric values are considered for the strength of our expectation. Just because of this similarity the concordance has a certain interest.

If we have observed the occurrence of a certain outcome several times under some circumstances, if subsequently we expect the same course of events under circumstances that we find similar, and if we appeal directly to the authority of that experience, then we have drawn a conclusion by analogy. Since this is based upon cases which have been observed already, which are similar to a present case, then it is immediately clear that its certainty grows with the number of those cases. In addition it is clear this depends on the strength and the type of the correspondence which these cases bear to one another, and to the case at hand. How that dependence may be stated is no simple business of the strength of the correspondence and its type. It must also be emphasized that it happens in a very special way that the outcome in question occurred for all cases subsumed by the analogy. It is entirely and generally conceivable that the outcome occurs in some arbitrary percentage of those cases, while being excluded in the remainder. Such a situation can also

determine our expectations, though in another manner altogether. Since it is necessary to have short names for these different types of argument by analogy, let us speak of the formation of expectation by complete analogy, or by partial analogy, respectively. The latter is always characterized as a proper fraction. The fraction tells us how large a portion of previously observed cases has resulted in the outcome in question over the course of events. This proper fraction may be called the figure of merit for analogy. If it were  $= \frac{1}{2}$ , then previous experience would have made the outcome and the failure of the outcome equally probable. Rising values would allow us to conjecture success of the outcome with ever-increasing probability, while sinking values would let us conjecture an ever-increasing probability of failure. Finally, maximum certainty would be achieved if the fraction equaled 1 or 0: that is, if partial analogy became total analogy. Our current attempt at explanation takes a fraction from the probability calculus simply as such a figure of merit for analogy. In that much, the unsatisfactory nature of this view is easily shown by a few very important distinctions which separate this view from that of the probability calculus. First: with real values of probability, evidence of experience with a number of cases of the same kind is never demonstrated, as a rule. Often enough we enter into the most varied types of games of chance with full confidence in the calculation of probabilities, even if those games are brand new, and if evidence from an investigation of them is still pending. But yet in another respect, statements of probability contain much more than would be the case, if their content were given as the simple result of experience in the way put forward. For example, let us consider an ordinary game of dice once more, this time as the result of actual observation. Suppose that each and every number were to land approximately equally often over the course of many throws. Then the same notion becomes an impossibility, in the face of an unbounded set of consequences involved in this general proposition: that the probability of the result of every single throw is independent of the outcomes of all the rest. As a consequence of this proposition we also maintain that it would be equally probable to throw the following sequence in 6 successive turns, for example: 1 2 3 4 5 6 or 6 5 4 3 2 1. No one denies that this claim is correct, but there are no grounds for doubt that it cannot be taken in the sense in question here: as the empirical basis which would need to be demonstrated. Not only does there not exist an empirical basis here – as would need to be demonstrated under this conception – but any attempt to establish it would founder under the enormous number of throws necessary for its establishment.

The difference between the fraction for probability and the figure of merit for analogy is brought to light in a still more decisive way, in that the very sense which we indicate to be the primary meaning of the former does not pertain at all to the latter. About that much there can be no doubt: that is just its logical meaning. So long as one is concerned with the probability calculus at all, one thinks of its numbers as follows: that a value of one should indicate complete certainty, while any fraction should be the warrant of a definite shade of expectation. The whole meaning which has always been placed on methods of the probability calculus has rested on this formalism, too. Yet as can be seen at a glance, this intuition is quite inadmissible for the figure of merit for analogy. A value of one – given to total analogy – does not

represent complete certainty here at all. The fact that all previously observed cases of a certain category follow such-and-such a course of events is never grounds for complete certainty. Rather, it is grounds to expect very different probabilities for a new instance of that same course of events, depending on other conditions. And so formal values which are intrinsically variable are given to the figure of merit for analogy; those values are not reportable as fixed once and for all. This holds for partial analogies too, just as for total analogy. So long as we want to proceed to interpret statements about probability on the assumption that these propositions actually have the same properties which one has always attributed to them as being absolutely essential, we will not be able to think of identifying them with figures of merit for analogy. Probability statements must contain something else than merely a historical record of past events.

**6.** We have arrived at the last attempt at interpretation which we will discuss here. One is given to say that the ratio of frequencies – with which an outcome does or does not occur in previously observed cases of a particular kind – may not however be regarded as the probability of that outcome without further consideration. It may happen, though, that the outcome may have occurred this or that many times in a large number of cases. Then the conclusion may be drawn that the relative frequency of the same outcome will be just as large – at least approximately as large – in a sufficiently large number of cases to come. This conclusion, borne of experience, which would draw initially on the totality of results for a large number of prospective cases, would form the content of statements about probability. In particular, this would be the real meaning of numeric values which figure in such statements.

Here we should set aside one circumstance: that under this view – as under the one just discussed – the claim enlists an empirical basis for probabilities in numbers which is hardly ever present in general. It will not be necessary to return to this point; rather we will confine ourselves to such factors as are characteristic of the interpretation under review. This much is correct: that statements about probability, say for games of chance, give rise to the very strong expectation that a particular outcome will occur approximately so or so often over a large number of cases. Yet we cannot claim to possess a satisfactory understanding of the content and the foundation of such statements, if we are merely able to report that they are derived from experience. Rather, at a minimum shall want to form quite a general idea of that which experience actually teaches us. We shall want to have an idea what objectively valid knowledge serves as the basis of our expectations about future courses of events. In the face of such a desideratum we are now in a position to indicate that for all such inferences, we assume that a thoroughgoing lawfulness holds for every event. Given the fact of the occurrence of an event, this lawfulness is always to be considered as having been brought about by specific conditions. Accordingly the occurrence of any event is also bound by necessity to the manifesting of certain circumstances. Experience can reveal such connections to us; any general conclusion drawn from experience would represent more or less fully-formed insight into such connections. In them we come to recognize those very objective relations which are known to us, on which our expectations are based. The assumption that such connections are present everywhere also appears to be an assumption necessary to the logical warrant

of every such conclusion. Then it will constitute no important difficulty to understanding those conclusions, if they should admit the interpretation that a specific event must occur under definite circumstances as a result of lawful connections. But then if we attempt to interpret some expectations in this way – those which produce statements of probability for games of chance, given a totality of many trials – then we are easily convinced of the inadequacy of such an interpretation. We are quite unable to form an idea of a lawful connection that supports this – say that for example red and black have to land approximately an equal number of times in a thousand spins of the roulette wheel. There is absolutely nothing in the constraining circumstances as described – a thousand spins of the wheel – to make that approximate equality necessary in an intelligible way. Indeed it is the case – and this point is decisive – that such necessity is never even assumed. That would always need to be put forward in the form that the relative frequencies of spin outcomes may not deviate from equality by more than a given value. But something of that kind only happens when we stipulate that individual trials are dependent on one another in some way, which departs from the ordinary arrangements for games of chance. Let us draw a number of balls from an urn which contains 500 white balls and 500 black balls. The balls are not replaced in the urn after they are drawn. Then the outcome of every draw modifies the probability of the next event. So in fact it happens by necessity that, for example, at least 200 white balls are drawn in 700 trials, simply because there are no more than 500 black balls. Here circumstances dictate a necessity that is both specifiable and entirely specific. Yet on the contrary, it is absolutely essential to our ordinary notion of the procedures of games of chance that the outcome of an individual trial is entirely independent of other trials. Therefore the attempt to produce a similar explanation will fail for our expectations which concern the totality of outcomes. Even the usual assignment of probability itself gives rise to the greatest deviations from the expected ratio, as in this example: that it is not impossible in roulette that the outcome is red a thousand times in succession. That deviation does not represent an impossibility of course, but rather an extremely small probability, though one which can be expressed as a finite numeric value. By extending our calculations it may be shown that for a sufficiently large number of series – of 1000 spins apiece – the occurrence of so-and-so relatively many series in which one lands on red for all thousand spins, may be expected with maximum certainty. Even the most minimal probability is fundamentally distinct from an impossibility, and we cannot bridge this chasm even if we allow the numbers to increase yet much further. Still it remains a characteristic feature of probability statements, that they report a larger or smaller warrant for an expectation, but also that for each, failure seems possible. Then there can be no satisfaction in solving the issue about the signal meaning of probability statements – the issue of an immediately evident necessity, the issue of a lawful connection – if we avert our attention from individual cases and turn instead to the totality of outcomes over many cases. As before, we hold these propositions to have the same type of content, and it is a change of entirely secondary importance if in the effort we have replaced moderately high numeric probabilities with very high ones, that is, with probabilities whose values come very close to unity.

Let us also admit something which may perhaps not allow misunderstanding in many cases: that statements about probability can express a result of experience gotten by inductive means. Yet we must deny that the very meaning of those statements is rendered intelligible by this observation. Rather we can only conjecture that, as other inductively obtained knowledge may give expression to lawful connection, so probability statements express another subject of experience – meaning something which counts as an objectively valid subject. Surely that is something fit to regulate our expectations in a definite way. It appears not to be expressible what this content may be, at least not offhand, or how it may become the measure of our expectations.

And so we arrive at essentially the same point to which the discussion of logical interpretation had led us earlier. Proceeding on the assumption that an item of knowledge gained inductively from experience – hence knowledge about objective relations – is represented in probability statements, we find that this is not to be considered knowledge of certain reportable and specific lawful connections. Rather we find that an objective situation – not to be labelled for the moment – appears to be the basis of characteristic rules that guide our expectations. Earlier we had the institution of equally warranted premises. They framed the real task of the probability calculus as capturing a certain sort of formation of expectation. We were then led to the conjecture that such a convention would have to operate (at least for the most part) by depending on a certain knowledge of objective import.

Since at this juncture we have exhausted readily available attempts at interpretation, we can address questions which rebounded to us as we pursued the Principle of insufficient reason. Two different paths lie open to our investigation here. On one hand, we can set ourselves the task of addressing the question in a purely logical and theoretical manner: which forms of knowledge produce compelling premises which are both equally warranted and also not arbitrary? We will dedicate ourselves to this task at first. On the other hand, we could try to ascertain the objective properties of relations in games of chance which are suited to serve as the basis for conventional assignments of probability – and which are generally held to be valid. This line of inquiry will also be pursued, but the way will only be open to practical inquiry once we have found a guiding thread in the answer to that first question.

# Chapter 2

## The Convention of Equally Warranted Premises



**Abstract** Incomplete knowledge about the state of a system gives rise to statements about probability. There is more than one kind of probability; most are not to be reckoned in numbers. The notion of equal possibility can be applied to ranges of possible events (or possible things). A range theory of probability is developed; a complex of ranges is known as a range configuration. Statements about the equality of ranges may be compared. The comparability of ranges is requisite to numeric probability. Ranges must be indifferent to be compared, meaning they have parts assigned to be equal in value. Those ranges should not be reducible to other ranges, meaning they are original. Numeric probability only makes sense where probability statements span original ranges which are indifferent and comparable in magnitude, and where the statements cover all possibilities. An example of numeric probability is given for two series of processes known not to influence one another.

**Keywords** Numeric probability · Independent observations · Logic of induction · Spielraum theory · Range principle · Indifference of ranges · Comparable ranges · Originality of ranges

1. Suppose we pose ourselves the question: under which circumstances does there exist a specific, non-arbitrary convention about a value of probability ascribed to a premise? We can try to generalize, and draw the consequences of a conjecture alluded to earlier. In the formation of expectations about a meteor strike, it seemed reasonable to assign a probability proportional to the area of a section of the earth's surface as the probability that the meteor strikes that section of the earth. From there it was a small step to the idea that there might always be something similar, should our premises find a subject for which – so far as we know – an operational range seems possible which is subject to measurement and which can be divided into parts. So for example, if we knew that the specific weight of a substance was greater than 5.0 and smaller than 6.0, then for this value a range (Spielraum) between 5.0 and 6.0 would be in question. The particular premises we might develop under these circumstances, say that the specific weight lay between 5.4 and 5.5, or that it lay between 5.7 and 5.8, would be held to be equally warranted in a direct and evident

way when and only when parts of this range (the Teil-Spielräume) are equal to one another in what they encompass. With a number of equal-valued premises, we can exhaust the totality of all extant possibilities. Then it would appear strictly warranted, say, to attribute a probability of  $\frac{1}{100}$  to the premise that said value may lie between 5.34 and 5.35.

Right at this point however, it is necessary to make an essential codicil. The rule which we tried to institute just now should of course only count as admissible if our knowledge is of this sort: it positively excludes values below 5 and above 6, but also it provides no reason at all to claim that any value between 5.0 and 6.0 is more probable than any other value in the range. Only under this assumption will the magnitudes in the range exclusively determine the logical warrant of individual premises. Rather than trading in reasons in any way, we only have to trade in weighing premises guided by ratios of magnitudes. Call this the unrestricted formation of expectation, and let us call operational ranges indifferent if there is no logical precedence of one range of behavior over another in the way just outlined. To apply this to the example above, that could consist of two premises that the specific weight lay between 5.34 and 5.35, and that it lay between 5.68 and 5.69, which premises encompass equal and indifferent ranges. The two appear to have equal warrant for this reason. It would be good to introduce an afterthought one might have here, which stands opposed to this view. One might ask how such a specific form of knowledge may be supported, so that we might be in a position to state that a value may definitely lie between two bounds, or that real behavior may be bracketed by a specific range, even though there may exist no reason to draw distinctions between different parts of the possible range. Instead, shouldn't inexact knowledge always be represented in another form: that a specific behavior is maximally probable, and that the probability of other arrangements becomes less and less as they deviate from it more and more – without a sharp distinction to be made anywhere, or without establishing the complete equality in value that we call indifference? – For the moment I must ask the reader to suppress the afterthought, even though it be completely justified. Since we are in the midst of developing our logical theory, this ought not to seem untoward. So it ought still to be left undecided if logical relations are realized as we have cast them, and under which particular circumstances that may occur. Still I wish to note in advance that the difficulties raised here do in fact disappear for games of chance as in many other domains, as a consequence of some special characteristics.

**2.** As we turn to closer examination of the conjecture we have just raised, at first we can easily confirm that in contrast to the unrestricted formation of expectation – in which probabilities are based on relations of magnitude among indifferent ranges – there also exist relations of logical connection. Insofar as some items count as certain, others count as having larger or smaller probabilities though there is no numeric measure for them. In fact only the completely rigid connection which binds the conclusion of a deductive argument to its premises will do to transpose the certainty of the latter to the former, unchanged and without remainder. For other kinds of logical relation the situation is different – for conclusions by analogy, to

begin with. Suppose we have seen one case or several cases proceed in a definite manner, and we expect a similar case to proceed in the same way. Then of course this expectation does not share the same certainty as the premises on which it is based. If those premises are taken as certain, the expectation is always more or less probable, and no more. The logical relation which obtains here is nothing that can be represented in numeric terms. The probability of the previously observed outcomes climbs with the number of cases that are revealed. That probability also depends on the degree and type of similarity which individual cases have to one another, especially the similarity that currently judged cases have to formerly presented cases. Any clue to an assignment of equally-valued premises is simply lacking. The same holds for that logical relation which we had contemplated under the rubric of partial analogy. Famously it remains a widespread notion that reliably definite and easily ascertained numeric probabilities can be ascribed to conclusions by analogy. Since this opinion rests on certain mistaken ideas which are not unimportant in themselves, we ought not to let them go unconsidered here.<sup>1</sup> Under the procedure according to which numeric probabilities are obtained from the kind of logical relations which concern us here – especially for partial analogies – one proceeds on the following assumption: that one ought to attribute a specific probability to each case of the relevant kind in the process under consideration; that this probability may be found to be approximately so-and-so large on the basis of previous experience (here it is unnecessary to elaborate how one tries to determine the certainty of this result in numbers); and that the value so obtained for the expectation of the relevant outcome should be considered the measure of a fresh case. In the first place, the unproven assumption made here is that if one abstracts a general concept from a number of similar cases or cases that are somehow the same kind, then given a new case is subsumed under the concept, a definite and ostensibly numeric probability must be implicated for any instance of the outcome. Yet it may always seem questionable that, for example, the general report that someone has typhus really provides the basis for a definite probability of the expectation of a fatal outcome. It may seem at least questionable that the value – which one is simply trying to establish – exists at all in that sense. If we wish to admit this is valid at all, it ought to be so only in the sense that the prospectively derived value of probability with which the fatal outcome is expected – to stay with our example – is just the probability if we know nothing else about the case except that it is ‘a case of typhus’. Yet this probability is not to be identified with that which seems to be warranted in a single case, since in the present example an agglomeration of specific and relevant bits of knowledge is necessarily available for this case. If the process under discussion is considered apart from that circumstance, the tacit and prior assumption has been made that only one thing is important to the expectation of the relevant outcome in each case, in which it corresponds to all previous cases. Every single case must count solely as an example of the general category in mind. Insofar as one adopts this

---

<sup>1</sup> Compare the more thorough critique of this procedure in Chap. 6, Section 5 ff., which is only touched on here.

framework, one sees that the logical value of conclusion by analogy is not so much determined in numbers – instead, conclusion by analogy has been completely pushed aside to be replaced by a wholly different concept. Namely the judgment passed on a new case seems to be that an account of real significance can be given, which is to say that it belongs to a specific category with certainty, and that this fact supports a particular expectation in line with a general rule derived from previous experience. But in fact that is no longer a conclusion by analogy. That type of conclusion is excluded as a matter of course, once it is assumed that the accumulated cases all lend themselves in the same way (univocally and measurably) to determining the expectation of the outcome. We will see later that this does occur, and how we should consider the content of a general rule of expectation. There can be no lingering doubt that this conception of the nature and function of the way we adjust expectations to previous experience cannot hold in general. We cannot eradicate analogy in the sense given here from the logic of empirical knowledge, because individual cases falling under a general concept are almost always discriminated from one another by an expectation of this or that course of events. A multitude of particular properties seems essential to each case. As long as we do not merely rely on general rules to determine our expectations – which rules are surely tied to the realization of conditions present in each case – and instead we seek to guide expectations by previous experience of similar cases in a way that admits no further analysis, then we are forced to deal with the logical relation of analogy. This might well happen extraordinarily often. Under most conditions there are many cases that are the subject of experience which have a multitude of distinctions unknown to us. Diversity in outcomes is not to be excluded entirely in this context. Rather a characteristic property may align more frequently now with one course of events and now with another. A fresh case is the same in broad terms of course, but in many respects it resembles now these cases more closely, and now those other cases. A full and exact correspondence is never achieved, either in whole or in part. There is only ever greater or lesser similarity. All these many factors that completely evade exact description are to be considered in the expectation associated with a process. Maybe after the aforementioned mistaken conception is swept aside, it may be admitted generally that a numeric measure of a logical relation does not obtain under these circumstances. We cannot characterize similarities by the enumeration of factors which correspond or do not. Even if one didn't recoil from the assumption of such a possibility, still it would be inexplicable how a definite measure of expectation for an existing case would be derived from the number and similarity of previous cases.

Exactly the same thing as holds for analogy also holds for the logic of the process of induction. By that we should understand a procedure by which we infer propositions of abstract content from a more or less extensive knowledge of experience. Especially if such a proposition has manifold consequences and it finds widespread application, and therefore can be supported by many different kinds of empirical results, then there is no denying that a numeric measure of its support or its justification by experience does not exist. It would be an entirely wrongheaded venture to look for a numeric datum for the certainty of, say, the law of inertia or

the principle of conservation of energy. The same holds for other less well-founded theorems in other domains. For any proposition of the kind, an array of factors represents the range and precision of its empirical tests, its domain of effect and the applications it will generate, no less than the objections which may disconfirm it, and which themselves may be thwarted by fresh assumptions. This array of factors defies determination in numeric terms, as a matter of principle.

Without going deeper into the matter explicitly, perhaps we ought to state that even the logical relation that a general proposition has to the individual facts which support it or affirm it, is a relation that lends a larger or smaller probability to the general proposition, but does not admit characterization in numbers.

**3.** If relations of logical connection do exclude the numeric determination of probability values proper to them, then one sees it can only be the unrestricted formation (i.e., unconstrained by other reasons) of expectation which can be represented as a probability given in numbers. At this point we will concern ourselves with that in more detail. Even when the unrestricted formation of expectation does take place, and when the available range is measurable, the establishment of equal-valued premises often runs into difficulties. Already, in close examination of our previous example – where the specific weight of a substance was in question – we are compelled to make ever more specific assumptions about the measurability of ranges. Let  $\delta$  stand for specific weight, and corresponding to what was said earlier, let the probability that its value lies within a small region  $d\delta_1$  be proportional to that magnitude  $d\delta_1$ . The specific weight is the weight of a unit volume of the substance at hand; in its place we can also set the volume of a unit of weight – its specific volume, which we may call  $\omega$ . The premise that the actual value of  $\delta$  lies in the small range  $d\delta_1$  is identical with one that the value of  $\omega$  lies in a certain small range  $d\omega_1$ , and should therefore by the same principle also be set to be proportional to the value  $d\omega_1$ . Yet  $\delta$  and  $\omega$  are inverse values; elements of the one and elements of the other which correspond to one another, do not stand in constant ratio. For example the probability that a specific weight may lie between 1.0 and 2.0 or else between 10.0 and 11.0, appears equal in terms of values of  $\delta$ . By contrast, in terms of values of  $\omega$ , the range of the former premise would be represented as from 1.0 to 0.5, while the range of the latter would be 0.1 to 0.09. The former would seem 50 times more probable than the latter, accordingly.

This approach to probability is therefore a vague one. Examples of a similar type are easily multiplied: the same uncertainty arises for the relations of the properties of a pendulum, since we can take into consideration either length, or frequency, or cycle duration.

Consideration of other sorts of relations may lead us to something fundamentally new. If it is known that an urn contains some black balls and some white balls among a thousand balls, but there is no information how many of the balls are black and how many are white, then here is a general and pragmatic approach to probability: that it should be considered equiprobable that none or else one or else two or else three etc. of the balls are black, and that correspondingly all the rest may be white. This seems quite thought-provoking as a simple example of the unrestricted formation of

expectation, for a measurable range. Yet then let us consider the following example, which reflects an application of the same procedure. Two playing cards lie on a table; we turn one over and find it is from a black suit. How would the probability be affected that the other card is either black or red? I believe a naïve reader would say the card could as likely be red as it could be black, in other words the probability of either premise is equal, and a value of  $\frac{1}{2}$  is to be assigned. But for example, in **Poisson**'s treatment of this case, the probability of the card turning up black once more is  $2/3$ , and the probability of its turning up red is only  $1/3$ .<sup>2</sup> This is supported by the previously-mentioned approach to probability, in a way that will be discussed later on.<sup>3</sup> In fact it enters into an easily-demonstrated contradiction with the same premise that appeared so central to us a moment ago, that either the first or the second card might as well be red as it might be black. By his reasoning we would assume in advance that the probability both cards are black is  $1/3$ ; the probability one is black and the other is red is  $1/3$ ; and the probability both are red is  $1/3$ . Following the other way of seeing the matter, we find by contrast that the probability both cards are black is initially presumed to be  $\frac{1}{4}$ , and the probability both are red is  $\frac{1}{4}$ . The probability that one is red and one black is  $\frac{1}{2}$ , since the latter case covers two possibilities: namely that the first card may be red and the second black, and also the converse.

Now is it correct to say that any card might equally well be red as black, or is it correct to say that with two cards, either none or else one or else two cards can be assumed to be black with the same probability? When presented this dilemma, we may have to admit that one alternative is as correct as the other: the assignment of probability is entirely arbitrary. There seems to be a difficulty similar to the earlier one which we encountered with specific weight, but here we may go still further. That is, it will have to be considered how the state of affairs currently under examination may be thought to have arisen. Consequently a follow-on question may present itself: whether both cards were taken from one and the same deck, or else from different decks. Given either of these assumptions, it will hardly be likely to seem any more intelligible to declare that these are numerically tractable possibilities. Yet right at this point here the modest beginnings of a set of possibilities present themselves which develop into an unbounded class. If we use the first assumption as a starting point, then under the further assumption that both cards are not chosen with foresight, but that they were arbitrarily drawn, admittedly some numeric probability will result for each of the cases considered. But that result will differ depending on the deck in play: a game of whist with 52 cards, or a game of piquet with 32. Consequently we are driven to consider another question in turn: what probability is to be assigned if our cards belonged to one kind of card game, or if they belonged to another? What probability is assigned that they were chosen

<sup>2</sup>**Poisson** uses this example, which underlines the arbitrariness of the ordinary procedure especially clearly, without a second thought to its explanation. (*Recherches sur la probabilité des jugements*, p. 96) [Poisson, S.-D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris: Bachelier, Imprimeur-Libraire.]

<sup>3</sup>In the context of the so-called Bayesian principle.

randomly, and what probability that they were selected intentionally? Once we begin to go down this path to its end, it becomes evident that it is quite impossible to develop a totality of all the premises which present themselves, or to tally a determinate value of probability for each, and so it is impossible to arrive at a definitive result. Then no doubt can arise that of the two assignments of probability – definite propositions to begin with, so it seemed – neither one nor the other can be considered the ‘correct’ one. Rather this presents us a case in which the result is not a definite numeric probability. Now let us return once again to the example of an urn filled with some black balls and with some white balls. At first here too there seems to be the idea that each individual ball could just as well be black as white, and that idea is only as warranted as in the case mentioned above. There is an assumption that any arbitrary number of balls may be black while the rest are white, and that premise lends a distinct and equal value of probability to any number of balls which may be chosen. Under that premise, the predominate probability would be that for any number drawn, an approximately equal number of black balls and white balls would be expected. But here too, further consideration of the matter leads to very different assignments.

Certainly we might also say that initially we are most likely to assume an urn filled with balls of only one type, or else to assume a random mixture of both types. Subsequently if a thousand balls are present, the premise that a thousand are black or five hundred or none at all, must be accorded a larger probability than the premise that, say, 873 of them may be black. The attempt to produce a complete partition of all possibilities would force us to wander lost in an endless labyrinth; necessarily it ends in no result at all.

We may easily draw further contingencies from the examples under discussion, which must be satisfied for numeric probabilities to result from the unrestricted formation of expectation. Even if the framing of an example appears on first inspection to offer a number of equally valued premises, then the importance of this partition may freshly be placed into question, if the behavior of the process under examination is considered the necessary result of some other process. Because as soon as that is the case, premises which may be taken at first to be equal in value may also represent consequences of other premises which may no longer seem equally warranted. Instead they may give rise to completely different assignments of probability. It is then evident we should only consider two possibilities as really the same in value, if that equality of value also holds in the same sense for the conditions on which we have to think those probabilities depend. Then as soon as any actual state of affairs comes under close observation, it is an eminently clarifying step not to rest content with the starting point that framed the question, but to appeal to constraining circumstances. The probability of a present state of affairs – or a future state – must be judged by the probability of those prior types of behavior which are apt to produce them. If those change, as they do regularly when laws in the domain under which the processes play out are known only roughly, then investigation may be split into conjectures about the form of these laws, and conjectures about actual behaviors. If those behaviors must be traced to ever earlier and earlier stages of prior conditions,

then the prospect of an exhaustive partition of possibilities runs into a complication. The complication makes it seem questionable in general that a specific endstate can be attained under any circumstances whatsoever.

Nevertheless we will not be able to dispense with the strict character of this logical requirement. This may be phrased to say that the ranges to which the formation of expectation is directed, are not reducible to other ranges. In short we should like to say they are original. This requirement is satisfied in a particularly simple way if our view of probability incorporates such things as we have occasion to regard as invariant constants of nature. Indeed this ought to occur, if as in the example given earlier, we ask about a substance's specific weight. Yet we shall soon see that this is by no means the most important role of original ranges.

A second contingency presents itself merely as a somewhat more exact expression of the concept of measurability. If different physical magnitudes depend on one another (specific weight and specific volume; the length and period of a pendulum, etc.) and the ranges of values in which they correspond represent one and the same range of actual behavior, then the range is measurable, though in different ways. This does not lead directly to the institution of equally warranted premises, since then entirely different relations of magnitude may result, each according to the kind of measure applied to its parts. There exists no unique and compellingly definite comparability in magnitude for the parts of such a range. In many cases the arbitrary nature of comparison between disparate parts of a large range can be immediately evident. For example, suppose that we make conjectures about the temperature of the sun. Suppose too that within a very broad range, its temperature is unknown to us. Then the range between 5000 °C and 6000 °C may not seem comparable to that between 500,000 °C and 600,000 °C. Then a definite ratio of probabilities could not be reported for the two premises, even if the temperature were considered to have been fixed for all eternity. In fact for each of these ranges very different relations of magnitude would be obtained under different types of measurement. Therefore it is necessary for us to change our earlier account, and to say in its place that a condition of numeric probability is not whether a range is subject to measurement; rather it is whether the comparability of the ranges spanned by the different premises has been established.

Then as the overall result of our logical investigation, we have the proposition that premises stand in a numerically expressible ratio of probability: if they span original ranges that are both indifferent and comparable in magnitude, and if definite values of probability are thereby produced, for which the totality of possibilities can be exhausted by enumeration of such premises.

Under the conceptual viewpoint we have attained, above all else the characteristic idea is that the ability to represent something in numbers is seen as only a property of one kind of probability. Strictly speaking, our investigation has taught us something more – and something more important – than the answer to the question with which we began. It has acquainted us with this special kind of probability, and has taught us how to distinguish it. We have found that on occasion the probabilities of various premises may essentially be measured in terms of the ranges each of them spans – or,

as one might want to express the matter – by their scope, or by their generality. We can express this result in the form of a principle of probability, and now say something to the effect that: two premises are equally probable if they span equal, indifferent, and original ranges. The importance of this principle becomes clearer, however, if we propose that a very high degree of probability be assigned to such premises as span the greatest part of a range considered as a whole. Under this account, attention is drawn to the exceptionally great certainty of expectations which are based on this principle under some conditions. As may be said right here and now, this Range principle is of very general importance – it has significance which extends far beyond the probability calculus. This point will be sketched briefly in Chap. 7, insofar as is possible in the present work. As far as our proper subject of numeric probability is concerned, at least we have come one important step closer to our goal with the establishment of the Range principle. Still, we cannot claim to have arrived at our goal. First, in addition to the notion of equally possible cases, there are a whole series of other concepts in the theory of probability that seem to demand investigation and elucidation. Above all it will be desirable to determine which real proportions give rise to numeric probabilities, especially if we should not find we have spun some pure fiction which has no application of any kind. The nature of numeric probability will only be revealed to us by consideration of the specific circumstances under which the purported logical relation is actually established. Specific circumstances tell us a more complete story about numeric probability, in a much more satisfying way than the abstract formulation given above allows.

Now at first, abstract logic offers us no foothold on such an investigation. But still it seems that the conjecture is justified *prima facie*, that at least in a subject domain proper to the probability calculus – that is, for games of chance – its application might prove warranted even in the face of that theory's strict requirements. Accordingly we may hope that the investigation already laid out – which concerns objective proportions in games of chance – will provide us the further explanation we desire. Then it serves our purpose to precede this investigation by discussion of a few particularly simple examples of a different type.

**4.** To begin with, let us suppose there are two series; these series are the results of processes which are known not to influence one another. Neither is there anything common to the range of effects of one series and the effects of the other. In short, each process is wholly independent of the other; each proceeds on its own. One of them exhibits a steady alternation of two states *a* and *b*: either *a* appears or else *b* does, but the two never appear simultaneously. For example, we might designate *a* as the state that the barometric pressure is 760 mm or more, and *b* as the state that the barometric pressure is less than 760 mm. The series from the other process includes an event that occurs only once for a short time: the event  $\alpha$ . Thus if we knew ahead of time only that *a* and *b* regularly alternate with one another, and that they always last for the same amount of time, then – in case nothing is known about the temporal relations of the series from the two processes – we would believe ourselves justified in saying that event  $\alpha$  might coincide with *a*, just as well as with *b*. On the

other hand, if it were known that  $a$  always lasted a thousand times longer than  $b$ , then we might hold the concurrence of event  $\alpha$  with event  $a$  to be much more probable than its occurrence with  $b$ . The expectation that is warranted here would need to consider the relative extent in time of  $a$  and  $b$  to be authoritative in the absence of other information. If we should want to formalize a general procedure by which equally warranted premises are established here, then it would needs be said that: the temporal coincidence of a certain point from one series of a process appears equally possible given equally long intervals of time from another series, if nothing is known about the temporal relations of the two series, but it ought to be assumed certain that the two series are independent of one another in the sense specified above.

So we have here an elementary example of the unrestricted formation of expectation in which there are comparable ranges, and – as is plainly evident – in which the ranges are original. That is because the temporal relation of two completely independent series generated by separate processes is in fact “original” in the sense mentioned above. Insofar as we think of one series transposed in time with respect to the other, we consider the relations of coincidence as being altered in the same measure for all their parts. Consequently it is all the same if we direct our attention to this or that part of each series, to the event  $\alpha$ , or to conditions prior to it. The temporal relation of the two series is just something objectively given as such and such. There is no corner given to explanatory reference by any other elements.

By contrast, if a causal connection of any kind whatsoever were to have obtained between the two series, then the coincidence of event  $\alpha$  with  $a$  or  $b$  would be traceable to some difference in previous circumstances which were decisive, though that might not be immediately discernable. The assignment of probability would lose its justification, since the ranges to be compared could not be considered original. So for example, one should not really set the probability of dying in the daytime and dying at nighttime to be the same *ceteris paribus*, because physiological states have a specific connection to the alternation of day and night.

**5.** In many respects the example just sketched proves to be unadulterated fiction. To begin with, it would never come to pass that processes for two series could run completely independently, without any causal link. Instead it may be surmised that some connection exists between any two events, even if that connection is mediated and very distant.

We stand to gain a qualitatively new understanding of probability, similar in some respects to the one we have been acquainted with here, but one which is far more productive and consequential for further investigation. We gain that kind of understanding if we choose this question to begin our inquiry: how everything actually behaves at a specific instant – in that initially we disregard the changes or the events in themselves.

Here too we will have to deal with fictitious examples for the moment; that is, we want to restrict our considerations to elementary relations, meaning purely mechanical relations of a system of point – masses distributed in space. Then we consider the question to have been raised, what may be anticipated as the state of such a system at a given point in time. We might have to explore how – and if – this question may be

posed in terms of a number of equally warranted premises. In order to frame this case appropriately and align it with the Range principle, it would also need to be assumed that complete familiarity with the totality of varieties of behavior in question would somehow be available. That would enable a specific range of behavior to seem possible, but within that we would not even have a rough clue whether to expect one arrangement over another. Under these circumstances only those considerations would be applied which were almost entirely similar to the considerations which resulted from temporal coincidence before. The simplest case of this type might just consist of the following: that for a point-mass, it would be known that the point may be found within a certain space  $V$ , but any knowledge of its location within that space would be lacking. Then two premises – that the point may lie either in this or that portion of  $V$  – may then and only then be considered to be equally warranted, if both portions are of the same size. We can also say that the probability that the point is located within a certain portion  $v$  of the entire space  $V$  can be measured by the fraction  $\frac{v}{V}$ . Something similar holds for the direction of motion that the point-mass follows at the moment under consideration, as holds for position. It is simply unnecessary to assume prior knowledge of constraints here because the values of all possible directions are constrained already. The expectation that the direction of motion may lie within a certain solid angle  $\omega$  would be measurable as the ratio of that magnitude to the subtense of  $\Omega$  – the entire solid angle in question – meaning the fraction  $\frac{\omega}{\Omega}$ . Should there have been no known prior constraints, a value of  $4\pi$  would take the place of  $\Omega$  here, as the maximum value of all conceivable directions.

In the same way there might be incomplete knowledge about the state of a system which is composed of very many point-masses. It may be that certain ranges which are defined in the way given above could be given for all the spatial properties under consideration – that is, for the locations and the directions of motion of all points. The totality of system states which might seem possible to us under these conditions, would be determined in that varieties of behavior could form entirely arbitrary combinations with one another in any relation whatsoever. If for example, two volumes of space  $V_1$  and  $V_2$  seem to be possible locations for a point-mass, and two volumes  $W_1$  and  $W_2$  are the possible location of another, then overall the result for these possibilities would be:

1. the first within  $V_1$ , the second within  $W_1$ .
2. " "  $V_2$ , " "  $W_1$ .
3. " "  $V_1$ , " "  $W_2$ .
4. " "  $V_2$ , " "  $W_2$ .

The differences in location and direction of motion of all the point – masses may enter into arbitrary combinations in just this way. We may also stipulate that even these combinations all seem equally warranted insofar as there is no reason to hold that one is less probable than any other, to our knowledge. The totality of the said spatial determinants of the system we may call its configuration, for short. Then the term range configuration indicates all the configurations which seem possible, given

the incompleteness of our knowledge. Here it is easily provable that such a range configuration can be separated into parts just as a range for the location of a single point-mass may be. Those parts can definitely be represented as equal in value. From this a large number of equally warranted premises emerge for our expectation of the way the system actually behaves. To make this clear, let us give the labels  $S_1$ ,  $S_2$ , and  $S_3 \dots$  to the ranges of the individual spatial determinants. We consider the first of these,  $S_1$ , to be divided into a very large number of small equal parts or particles: they are to be called  $dS_1$ , and there are  $n_1$  many of them. Similarly we divide  $S_2$  into  $n_2$  many equal particles  $dS_2$ , and so on. Then let there be a range configuration which is defined such that the first of its determinants is confined to a particular part  $dS_1$ , the second is confined to a particular small part  $dS_2$ , the third to a small part  $dS_3$ , and so forth, each constrained to represent an element of the entire range configuration in question. All such elements are completely equal in value to one another, since they span equal parts  $dS_1$  as well as equal parts  $dS_2$  and equal parts  $dS_3$  and so forth. A number of equally warranted premises result for expectations about the actually manifest configuration of the system. Meanwhile it will be useful to describe the measurement of probability which results in a somewhat more surveyable form. That is to say, to any arbitrary part of a range configuration, we can assign a definite size which may be compared to the size of the whole. Thereby we obtain the probability that the actual state of the system may be contained within this part, in the same way as before: as measured by the fraction expressing this ratio of size. The number of equal-valued elements which span the whole range configuration as described, is then given by the total number of possible combinations among the  $dS_1$ ,  $dS_2$ ,  $dS_3$  etc., which is to say by the product  $n_1 \cdot n_2 \cdot n_3 \cdot n_4 \dots$

Now let us ask about the probability that the configuration of the system may occur in some narrower range which also subsumes individual ranges  $s_1$ ,  $s_2$ ,  $s_3 \dots$  just as the whole range configuration subsumes  $S_1$ ,  $S_2$ ,  $S_3 \dots$ . Then it is easy to see that it will only be necessary to describe how many of the previously-described equal-valued elements are contained in this narrower range. That number may be compared to their total number.

Let  $s_1$  of the  $n_1$  particles  $dS_1$  into which  $S_1$  had been divided, contain the number  $m_1$ . Similarly let  $s_2$  be composed of  $m_2$  particles of  $dS_2$ , etc. Then our partial range which spans a number of elements  $m_1 \cdot m_2 \cdot m_3 \dots$  will span those elements which form the number  $n_1 \cdot n_2 \cdot n_3 \dots$  of the whole range. Hence a probability is to be measured as the fraction

$$\frac{m_1 \cdot m_2 \cdot m_3 \dots}{n_1 \cdot n_2 \cdot n_3 \dots}$$

or, as we could also write, as the value

$$\frac{s_1 \cdot s_2 \cdot s_3 \dots}{S_1 \cdot S_2 \cdot S_3 \dots}$$

We can then regard the product  $s_1 \cdot s_2 \cdot s_3 \dots$  as the measure of the size of a range configuration.<sup>4</sup> We can say: the probability that the state of the system lies within a certain range configuration is to be measured as the ratio of the size of this range configuration to the size of the whole range configuration, taking into account our inexact knowledge of the range configuration in question.<sup>5</sup>

**6.** Until now we had restricted our considerations to states which happened at a specific instant in time, abstracting entirely from changes which might occur by laws of inertia or the mutual interaction of elements. If now we also bring these into the fold of our conjectures, it is immediately evident that those magnitudes which we have just come to know as range configurations cannot determine our expectations on their own. While certainly they may be comparable, and while keeping with the assumption that they may also be indifferent, yet in general they might not be original ranges. Rather, every current state is now to be considered as the necessary outcome of certain previous varieties of behavior, which leads to an insight: that even a behavior which spans only a very tiny range, and which therefore at first glance might appear to be very improbable, must still be regarded as very probable if the laws governing events are such that they facilitate the behavior, or that they lead to it by necessity. When, for example, many atoms unite in a close packing to form a relatively small body, a structure might be formed which only spans a minor and relatively negligible range. We might well expect such a structure if we know that there exist attractive forces among the atoms, by which they are drawn together. It is only under a very special assumption that despite this complication, the case under discussion would give rise to a specific institution of equally warranted premises. In order to achieve that, next we assume that the laws which govern changes in the state of our system are laws which are exact and fully known. Insofar as that is not the case, conjectures about this or that aspect of the said laws must be folded into the whole corpus of our deliberations, of course. Yet here we ought to ignore such conjectures as those for which there is no numeric account of probability. If there does exist complete knowledge of those laws, then we should be able to make a claim about the whole complex of configurations which seem possible at a given

<sup>4</sup>It is easily seen that if different varieties of behavior cannot be combined with one another completely and freely, then the product  $s_1 \cdot s_2 \cdot s_3 \dots$  is to be replaced by the multiple integral  $\iiint \dots ds_1 \cdot ds_2 \cdot ds_3 \dots$ , which spans the totality of all admissible combinations. Things are somewhat similar, should we consider the rectangular coordinates XYZ for the location of an individual point, instead of parts of space. The probability that that point would be found in a certain space, would be measured by the size of the space, and would have the value  $\iiint dx dy dz$ . It would be measured by the product  $x \cdot y \cdot z$  instead only if the various values of the three coordinates were entirely and freely combinable with one another, that is, if the whole range were described as a die or hexahedron. It is unnecessary at this point to expound further on this generalization.

<sup>5</sup>One may find it easier to understand this measurement of magnitude – just as one may find the later expositions easier – by tying it to spatial representations. One may consider the range configuration as a manifold similar to a space which may have very many dimensions, and whose independent coordinates could represent various values which stand for the arrangement defining the system. Under the analogy, the determination of distributed magnitudes takes the place of measurement of spatial elements.

moment. We may designate that complex as the range  $X_0$ ; that has as its necessary precursor a range  $X_t$  which existed at some arbitrarily early time  $t$ . In turn range  $X_0$  corresponds to the necessary outcome of the latter  $X_t$ . Similarly an arbitrary part of  $X_0$  that we may call  $x_0$  will be simply homologous to some part of  $X_t$  which we may call  $x_t$ . At the very moment the whole range  $X_0$  comes into play, the question arises whether a configuration might be realized which belongs to part  $x_0$ . The answer to that question may be referred to the prior question: whether the state manifest at time  $t$  might have been contained within range  $x_t$ . But then the probability of that, which we assume to be the ratio  $\frac{x_t}{X_t}$ , can as little be definitively assessed as it was by the ratio  $\frac{x_0}{X_0}$ . At this juncture it seems that a definitive result may only be obtained under the assumption that in tracing back events as precursors, we may come to a conclusive end; the aim is that we may consider the lawful course of events to hold once and for all from a given point in time. This point may be considered the starting point of all changes in state, which is to say it may be considered the origin of the timeline. We are not resigned to this fiction, however. Rather things will emerge just as they have here: a ratio of range configurations will again lend its definitive measure to our premises if, as we appeal to states further and further back in time, the ratio remains unchanged throughout. With the help of the expression we introduced previously, in that case we are able to say in short that there would be a certain ratio of size between those original range configurations. They represent any current varieties of behavior whatever by prior conditions. Then such current or future states as are consequences of equal and original range configurations will in fact have to be attributed the same probability. Premises that we assign as equally warranted have the property that, even if we should hark back to the premises from which they are thought to originate, we always obtain premises which are just the same themselves equally warranted. They should therefore, in a definitive way and in the sharpest sense, be considered as really equal in value. Anything which concerns a large manifold of original configurations will be expected to have negligible probability. Particular types of contemporary configurations which develop from comparatively large manifolds of original behavior will be considered likely; in that light they may be expected more often, and so on. Still, definite ratios of size among the original range configurations would be the measure of all these expectations.

As we have seen already, in the course of our investigations we will need to turn our attention repeatedly to the subject of the ‘original’ character of ratios between ranges. That is an uncommonly important idea for the theory of numeric probability. And so it is reasonable to dwell on this notion a little longer, given the specific way that it is applied here. As long as we fix our sights on questions of the balance of our expectations which involve only the very next stage in behavior, these considerations remain almost worthless, even if we can tie them to specific quantitative relations between an entire range and its separate parts. We will always need to demand proof that the thing which initially seems equal-valued is also to be considered equal-valued with respect to the laws which govern changes in states. We will always have to ask if this thing might not have emerged from a larger complex of previous states, or if that thing might not have emerged from a smaller

complex of previous states. Only where this question is firmly denied can such proof really be offered. Only then can it seem both conceivable and justified that such ratios of size may count as the measure of our expectations, unconditionally and unbiasedly. In fact it is a cardinal question for the whole theory of numeric probability how an assignment of probability can seem to be definitively valid and rigidly determined, despite the circumstance that every state is subject to lawful change, and that therefore various processes past and present are connected. This is a problem which simply seems to be shoved aside in the cases mentioned earlier, where physical constants are involved, or else where temporal relations of two entirely independent series of events are in question. Here we see that something very similar may occur, even when the varieties of behavior to be compared implicate differences in a current state and the totality of antecedents which constrain it.

If we speak of operational ranges in this succinct way, they just represent something no longer traceable to something else – i.e. something original. The argument we have just set forward serves to show – at least in some circumstances – that a certain comparison of magnitude may be carried out. We have gained a new concept of what a ratio of sizes (which is original in the sense given) between two operational ranges represents. Of course this new concept may only be applied to very specific situations, though where it may be applied, it is in fact well-suited to produce a conclusive and unbiased assignment of probability.

We may consider two things to be essential outcomes of this preliminary investigation. First we have garnered some idea how an operational range – as a concatenation of very many separate determinants – can be considered to have measurable magnitude under certain circumstances. Second, we have become acquainted with a procedure that is obviously very important: the means by which ratios of such ranges are represented as original. Both will prove essential to our understanding of games of chance, and both find application there.

# Chapter 3

## The Theory of Games of Chance



**Abstract** The range theory accounts for our conventions about probability in games of chance. Generally, movements in these games are the composite effect of many factors acting in concert. Expectations are formed freely by the comparison of different ranges. Such games of chance exhibit the properties which allow numbers to be assigned to probabilities – including the comparability of smallest neighbouring parts of ranges. Examples of dice games and urn problems are given. The notion of a probability function is developed from numeric ratios that are expressible for the objective conditions in the examples. The function should not exhibit the periodicity apparent in those objective conditions. The notion of continuity for a probability function is discussed in this context. An extended example called the bowling-game is used to illustrate what is currently known as the method of arbitrary functions.

**Keywords** Games of chance · Measurability · Periodicity · Probability function · Method of arbitrary functions · Comparability of regions

1. Let us consider on the one hand the objective conditions present in any so-called game of chance, and on the other hand the convention about probability which counts as admissible and valid in terms of the game. Right away we notice that in the objective conditions there are always certain numerically expressible ratios which are demonstrable, and to which assignments of probability are connected, so that the numbers which appear are nothing but the values of those objective ratios. Perhaps this is shown most clearly by the case which is favoured in exposition of the theory: selections drawn from an urn filled with black balls and white balls. One assigns the proportion of probability that a black ball may be drawn as opposed to a white ball: that is just the ratio of the number of black balls and the number of white balls contained in the urn. Other types of games like roulette behave similarly. There the probability that the ball comes to rest on the pocket of the wheel corresponding to this number or that number depends on the sizes of those pockets. Yet the situation is basically no different than in a coin-toss or in dice games. As we know, that is because the usual assignments of probability will only count as correct under the assumption that the objects in play are regular, both in geometric terms and in

physical terms. That means among other things that their centre of mass lies in their middle. Thus an objective equivalence must also obtain for certain relations between two sides of a coin, and among the six sides of a single die. Then this equivalence is reflected in assignments of probability, just as other arbitrary numeric ratios figure in other games of chance.

2. Then too, we find that different games of chance coincide in this: numerical statements about probability are made with reference to certain results or outcomes which we produce by the motion of certain objects in play – by the toss of a coin, the throw of a die, the spin of a roulette wheel, or the drawing of a ball from an urn, etc. As one is given to say, these motions are carried out freely and arbitrarily. Now we should like to subject a very simple game to investigation: the simplest game possible of its kind, and one which is easy to imagine. For short, let us call it the bowling game. It consists of the following: we imagine a long, straight, and horizontal alley that has a level floor and is bounded by vertical walls. At the beginning a ball is set down, and rolled by hand with a force (not too weak a force) that sets it in motion along the alley. The floor of the alley is painted, too, with stripes perpendicular to its length. In the path of its motion, the ball passes regularly over black and white stripes which alternate regularly. The width of each stripe is 1 mm, measured in the direction of travel. After moving a longer way or a shorter way, the ball comes to rest; naturally it rests either on a black stripe or a white stripe. One can tell that this game is basically very similar to roulette; one will agree uncritically to the claim that the ball would be just as likely to come to rest on a black patch as on a white patch, in other words the probability of either case will be equally large. This assignment of probability would be warranted anyways by another assumption about the nature of the game, which we would introduce because it serves to simplify the conditions further. We assume that the conditions of air resistance and of surface friction in the alley are known, and that they remain the same for repeated turns of the game. Under these circumstances the outcome depends only on one factor unknown to us: that is, on the force imparted to the ball, better stated as the strength of the component of that impulse in a direction lengthwise along the alley.

If we seek to justify this convention about probability, we note that it may not be based anyways on the notion that the probability of the ball coming to rest on an arbitrary stripe is the same as for all the rest. Rather it can seem wholly doubtful that this probability can be assumed to be the same throughout, say for the 500th as opposed to the 10,000th stripe, which is to say assumed to coincide for a ball bowled with medium force and one bowled with very significant force.

Further, we may consider an attempt to weigh the probability of two such disparate cases against one another exactly. If we must abandon the attempt as one which is destined in advance not to deliver a result, then there remains only one notion about our initial assignment of probability – though it is a notion both accessible and profound – which is not endangered by that renunciation. That is to say, one can hardly doubt that we can assume the difference is quite unappreciable between the probability that the ball comes to rest on the possibly-black 10,000th stripe and the probability that the ball comes to rest on the possibly-white 10,001st stripe. In other words, almost the same probability can be attributed readily throughout to any two stripes that are not separated by much; in that way the sum total of

probability for black and for white must prove to be equal (to an excellent approximation). If we wish to express the premise posed here in somewhat more formal terms, then we may say something like this: for the purpose of computation, given a small element  $\Delta l$  of alley length and a value  $\varphi$  which we call the probability function, then a probability may be assigned which is equal to the product  $\varphi \Delta l$ . The probability function is certainly not the same for all points along the path – all values of  $l$  – but it would very nearly be equal for nearby points separated from one another by only a few stripes. We shall call such an assignment of probabilities a continuous one.<sup>1</sup>

We have grounded the convention about probability with which we began on a very simple assumption. That is, we have assumed *a priori* something that was unknown for a specific real behavior: that within broad bounds, probability is continuous. In a moment we will assess the importance of this step, and we will need to test that general assumption further. However before we do so, we must finish our rough conceptual picture in what concerns the tracing of specific assignments of probability back to general premises.

If the meaning of continuity for the probability function is assumed to be as specified a moment ago, then the claim that the total probability of the outcome is equal for black and for white is a claim which rests on yet another assumption. Though that is in some sense pretty well self-evident, still it must be rendered explicit. It says that the function  $\varphi$  exhibits no periodicity which correlates with the width of the stripes. Actually, it seems at least conceivable that the function could be somewhat larger in a regular manner at those values which stand for white stripes than at those values which stand for black stripes, without the function becoming discontinuous at any value of  $l$ . If that is the case, the total value of all the white stripes can be noticeably greater than that for all the black stripes. Then in the original convention about probability, it is also postulated that such periodicity is absent, quite apart from the question of continuity. It is immediately evident we can show these two assumptions suffice to produce the right and relevant assignment of probability – to a close approximation. It is not difficult to illustrate the way in which

<sup>1</sup>It must be noted that something different – something even more – is meant here than might be expressed by the common mathematical use of the word ‘continuity’. Here mathematical continuity would only imply that  $\varphi$ , considered as a function of  $l$ , exhibits no discontinuities, i.e., that not infinitely few different values of  $l$  correspond to finitely many different values of  $\varphi$ . Instead here, the meaning of continuity is that few different values of the argument pertain to few different values of the function. This type of continuity by no means represents a sharply defined concept, then; there is a more-or-less to it, and it can be present in larger or smaller degree. Then the assumption made here is basically an estimate of that degree, leading to the point that the relevant correspondence may be considered sufficiently continuous to produce a best approximation to equality for the total probability for black and for white. Since the entire discussion of this postulate is about a certain estimate rather than about the affirmation or the negation of a specific behavior, here it emerges very clearly why it seems necessary for us (in justification of the convention about probability) to consider a certain smallness of the period with which black and white alternate. At any rate it seems necessary to be able to think that a specific value for its period, say here a millimetre of stripe width, ought to be sufficiently small. We will have more to say later about continuity and correspondence, in the same vein.

these overall equivalences develop. For example, let us imagine a rubber band, painted so as to divide it into very narrow stripes which are all of equal width, alternating black and white. Then the total length given over to black and to white would be very nearly equal, no matter how we might stretch or expand this band, presuming that the degree of expansion changes only gradually from one place to another, and that it does not change periodically in such a way that at any time the white stripes would be subject to somewhat stronger or somewhat weaker expansion overall than the black stripes.

The possibility of assigning a specific numeric probability to the event that the ball may come to rest on a white portion – or a black portion, say – depends on the particular arrangement in which the black and white stripes repeat. It depends on their arrangement along the entire length of the path; of course it depends on their arrangement where it seems likely the ball might land, which arrangement may include many alternations of stripes, and which arrangement assumes a constant ratio of extent. It is easy to see how essential this arrangement is. If instead of the previously postulated width of a millimetre, we were to give the stripes another width – say of 10 m for example – then the whole range which at first might seem likely to include the endpoint of the traversed path would itself span only a few stripes, perhaps three or four stripes anyways. Then we would not really be able to judge the game by the principles of games of chance without further information. Rather, in any more detailed investigation, the assignment of probability might prove to be uncertain and even no longer representable in numeric terms: then many kinds of consideration might seem possible, which are very different in the probabilities assigned to neighbouring stripes. As a consequence the total probability of the results for black and white might even no longer seem to be equal. It will not be necessary to investigate such fine distinctions in possibility. Without further discussion we have an overview which shows they do not render numeric probabilities for us. We can then rest satisfied to have established one thing about an oft-repeated alternation which is to be compared for its effect on probability: that it is truly an essential property along the whole range in question, at least for this form of a game of chance.

The assignment of probability – already much altered, and which is now revealed to be the core of the originally established convention – will now need to be the object of our further examination. At this point it must be taken into consideration – as we saw previously – that the length of the path traversed by the ball is essentially determined by the force of the impulse by which it is pushed forward along the alley. With respect to this contingency it is easy for us to add something else: the value of the path length must change with the force of the impulse, in a manner that we can call continuous in the sense given above, i.e., likewise there must be ranges of values of force which represent very small, adjacent, and congruent ranges of alley length, and which themselves are very small, adjacent, and approximately congruent. From this it follows that the convention about probability which we make about path length may likewise be considered as a very similar convention about the probability of force, and this convention about force may be substituted for that of path length. For this too, we must assume that probability is continuous and that periodicity is not evident in the probability function either: at least that any periodicity has been excluded which corresponds to the periodic alternation between – to twist a phrase – black-producing

values and white-producing values of force. And once these assumptions have been set, it follows immediately that the sum of probability values must be equal for outcomes of black as for outcomes of white, to a best approximation.

The more detailed examination of our bowling – game then leads us to the insight that for the assignment of probability offered here, one thing is essential to which we can hardly object: a particular arrangement of elements in the game, meaning the steady alternation of black and white stripes of equal width. Besides that, the game also rests on a convention about the continuity of probability, one that is made ahead of time about the result of an arbitrary motion, that is, with respect to a result whose scope is not known to any considerable extent beforehand.<sup>2</sup>

**3.** Things behave quite similarly even in more complex cases, as in dice games. The result produced by one throw, that is, which side finally faces up, is determined by a whole series of movements. We shake the die up and down, left and right, forwards and backwards; we let it rotate about several axes, first in one sense and then in the other. As it leaves our grasp at last, we impart certain speeds of rotation and forward motion to it. Before the die is cast, we know that a large number of such movements will be undertaken; within broad limits it seems that the number, the sequence, and the magnitude of each is unknown. Then one sees that the totality of all possibilities which arise in this way encompass a vast count of repeated instances of each of the six possible end states: *a priori*, innumerable chains of such movements can be considered which would lead to a 6 being cast; innumerable ones which would lead to a 5 being cast, and so on. That which was gleaned from the previous example of alternating black and white stripes along an alley then stands, and it is established in equal measure here, since over a wide range of variation in motion the six possibilities are the sole and repeated outcomes. Let us suppose then that for each movement in question, that a convention about continuous probability has been established. Then here too equality of probability is the immediate result for each of the end states 1–6. As before the equal width of black and white stripes necessarily ensured a specific and coherent complex of possible movements, so here the geometric and physical regularity of the die ensures that one connected complex of possible movements produces, say, an end state of 6. For every relation there are coherent alternative complexes of almost the same span that are just a little different, which would have as their effect the end states of 1, 2, 3, 4, and 5. Each of them

---

<sup>2</sup>We must emphasize that a continuous assignment of probability with respect to one magnitude always entails a like assignment with respect to any other which is its continuous effect. That relation follows analytically; we have already made use of this a moment ago. It is also implied directly here that for the claim – that a continuous probability may be assigned to the magnitude of an objective quantity – no difficulty arises from the circumstances which were cited earlier. One and the same range of actual behavior may be considered the range of values for this or that magnitude in many ways, and thus its measurement can be conducted by different means. Whether we pay particular attention to the specific weight or the specific volume of a substance, whether to the length or the period of a pendulum, or else to refractive index or optical density, it is all the same, because the continuous probability assigned to the amount of a quantity always represents a continuous probability of the other quantity at the same time. Here likewise nothing is changed in our deliberations by the way we consider force to have been measured either: it is irrelevant whether we take the speed imparted to the ball as our measure, or its kinetic energy.

returning in regular permutation, these six kinds of movement completely fill the entire range of possible movements. In analogy to our previous example, the result which emerges is that the sum totals of ranges for movements which produce the outcomes 1, 2, 3, 4, 5, and 6 must all be equal, to a best approximation. And from this it follows that each of these outcomes must also have the same probability, under the same assumptions as were made in that example.

The present case is distinguished from the earlier one only in a secondary matter; that is to say, here we are concerned with a combination of several effects while there a single effect of movement is considered – the effect of force. Accordingly we must frame the postulate on which our assignment of probability depends here, in a slightly different way. Let us call the various effects of movement in the game  $b_1, b_2, b_3 \dots$ . Such a partition could be performed in manifold ways, yet their governing procedure counts as just the same. – Then the probability that the first of them may be found within the small range  $\Delta b_1$ , the second within the small range  $\Delta b_2$ , and so forth, is set to be proportional to the others. That is, the probability  $\Delta b_1$  is to be proportional to the probabilities of  $\Delta b_2, \Delta b_3$ , and so on. The probability would then have to be expressed by the product  $\varphi \cdot \Delta b_1 \cdot \Delta b_2 \cdot \Delta b_3 \dots$ , where  $\varphi$  is once again a continuous function (in the earlier sense) of all the  $b$ . Shortly put, we would have to assume a continuous probability for the combined ranges of values. Similarly the postulate would have to be modified for the absence of any periodicity in the probability function which corresponds to alternations in the results; that postulate is also as essential as the one for continuity, as in the earlier example of the bowling – game. But here the aim is to exclude not only simple periodicity, but multivariate periodicity. That is a point whose more exact expression can be dispensed with, both for the relatively trivial importance attached to it, and for the difficulty of the formal mathematical treatment to which it leads.

We can provide a rather more general account of the characteristic of the game, on which we base the same convention about probability that we are counting on here. We can say that all the ranges of all that which seems possible will exhibit continuous permutation and a constant ratio of extension for the portions representing various outcomes. We can give the sense which really attaches to this arrangement clearer prominence here, in that we simulate certain transformations in the game. As we imagine stripe width to have been altered for the bowling – game, so here it is simple for us to imagine certain other rules of this game to have been changed, so that some characteristic of the game would no longer be realized. Perhaps the game could be played in such a manner that the die – we might think of it as being large and heavy – would be given an impetus that budges it only a small way, so that only a few throws may be cast. Under these circumstances new considerations seem possible, such as that if a 6 is cast once, the outcome of the next throw might not at all be that every position of the die would be equally probable. Experience might well tell us that certain end states of throws never succeed one another, or else rarely. So for example, immediate repetition of the same result would be excluded entirely if the force of the throw fell below a threshold, so that the die turned over once at least but never more than three times. In fact the usual assignment of probability would be inadmissible under those circumstances.

It is not especially difficult to convince oneself of the applicability of analogous behaviors to other forms of games of chance. Next I will mention those in which there is some business of mixing many of the same objects: cards, or balls, etc. Here one is used to assuming that if such things are shuffled many times, then be the initial order whatever it may, the resulting endstate can be expected equally well – with equal probability – to be in any arbitrary order whatsoever. Perhaps the more common assumption in the popular imagination – the assumption that for two different types of objects, as in black balls and white balls, an ‘even mixture’ may be expected – is a result of this assumption as a consequence of simple mathematical relations which are unnecessary to elaborate here.

Each movement which is undertaken with the aim of shuffling or mixing creates a specific change in the arrangement of the relevant structure. Each movement is subject to very similar considerations as were established just now concerning the motions of dice. With the support of the assumptions which have been made explicit previously, we should expect such arrangements as are the effect of only one mixing (or several movements) to be little different from one another, and we should expect such arrangements to have probabilities that are nearly equal. As we highlighted the equal width of the stripes in the bowling – game; as in the dice game we highlighted the regularity of the form of a die; here the congruence of individual objects is highlighted. The equivalence of ranges is directly implied by them: by way of example, they can even be ranges of movements which lead to the grasping of one specific ball over a neighbouring one, and that would transport the grasped ball to this place or a nearby place, and so forth. From this it follows directly that the validity of the stated assignment of probabilities is conditional on sortings being undertaken that are sufficiently numerous, and repeated sufficiently often. We can easily express this condition more precisely. The possibility of any particular arrangement must surely be subsumed once we assume that any individual sorting-movement is possible only over a very small operational range. Or what comes to the same thing, with respect to any sorting movement, *a priori* there must be a much-elevated uncertainty of any arbitrary arrangement in the endstate, since any arbitrary arrangement may possibly arise in many different ways. These conditions replicate exactly what we find in the bowling – game. There too, small variations in movement are sufficient to produce an outcome of white in place of an outcome of black. From the whole uncertainty which actually surrounds the outcome of a movement, the occurrence of any of these results seems to be possible in very many ways. – During the sorting, the number of cases which become equally possible is already huge for a middling number of objects to be sorted. Accordingly, only a very prolonged mixing or sorting will also serve as warrant for the assignment of probability at issue. If such sorting does not take place to a sufficient extent, then it should not be ignored that the probability of that arrangement which existed before the sorting took place must also be taken into account. Then no definite result has been produced, whether that prior arrangement is itself known or unknown. In particular it must be emphasized that the simple absence of definite knowledge about the pre-existing arrangement does not warrant any definite assignment of initial probability for us, since under these circumstances the most diverse kinds of questions press upon us – for example, whether the prior arrangement was established intentionally, or for this or that purpose, and so forth. As we have seen already, a definite result with numeric measure does not emerge in the face of such considerations.

Finally let us turn to the investigation of those kinds of games in which emphasis is placed on the selection of many differently colored balls from an urn. This presents no difficulty to the understanding, either, once we make some further determinations about proportions, determinations which are entirely necessary to ensure the validity of the usual assignments of probability as well. The well-known conventions about probability are these: that if the urn contains  $n$  black and  $m$  white balls, the probabilities of drawing a black ball or a white ball are expressed by the fractions  $\frac{n}{n+m}$  and  $\frac{m}{n+m}$ , respectively. Among the general assumptions which were derived in discussion of the previous example, it is easy to spot one: that the selection of adjoining balls is almost equally probable – though if, as one might perhaps like to assume in justifying that convention about probability, any arbitrary ball is expected to be drawn with equal probability – for example, one lying at the centre and one found at the bottom of the urn – then equiprobability cannot be claimed anyhow. Even by the psychology of common experience, that is most certainly not the case.

It also follows that the usually-accepted convention about probability can only count as warranted, either if that arrangement of the balls ought to be considered equally probable to others, meaning that one has sorted them thoroughly and carefully, or else if it is known that black balls and white balls are found to be contained in the same  $n$  – to –  $m$  numeric ratio either exactly or approximately in every small region of the whole urn. Yet if one or the other of these postulates does hold, then the validity of our assignment of probability is self-evident – better said, the fact is self-evident that it can be traced back to the general premises developed from other common games of chance.<sup>3</sup>

4. It will not be necessary to pursue fine-grained examples any further. Rather at this point we must approach this question: if the conventions about probability to which we have been led are really admissible, in other words if the notion is acceptable which was to have vindicated them in the discussion given by the previous chapter. Before we attempt to answer the question, it is necessary for us to be perfectly clear what it is we are aiming at. Let us proceed on the assumption that assignments of probability are strictly correct with regard to games of chance. Then we would have a motive to ask for a general and formal proof that they really express proportions of size among indifferent and original operational ranges. Now it stands to reason that no absolutely compelling proof, as in a mathematical or a purely logical proof, can ever be given for this, just as no such proof may be given for any fact of real importance. Then if we should declare the relevant probability statements to be warranted, that can only mean they are based on well-founded concepts. It does not seem at all possible to claim a definite kind or a definite degree of certainty for these concepts in advance. Soon enough we will have to delve further into this situation,<sup>4</sup> which is so exceptionally important to the whole theory of probability: that statements about probability always depend on postulates with

<sup>3</sup> Since the drawing of balls from an urn is employed so often as an illustration in what follows, let me say here that in the examples, the game is always presumed to have this kind of structure.

<sup>4</sup> Compare Chap. 4, Section 2; and Chap. 6.

objective meaning. As a consequence they rest only on relative certainty. Currently we see that it cannot be considered to be the all-consuming goal of our investigation to prove the probability statements in question, beginning with the Range principle. Rather the goal is only to render them understandable in light of that principle. However, it may already be conjectured that the notions on which they are based will prove to be exceptionally reasonable and well-founded.

5. Front and centre, it is now completely evident that a formation of expectation must actually occur, if we should ask whether this or that will occur in games of chance. The sum of all that we know about the behavior of real things is insufficient to predict the outcome in advance. The reason is not simply due to a gap in our deliberations, or due to an unexploited pocket of our knowledge. Rather it is a real absence of positive knowledge which allows the outcome to seem uncertain. And of course we will have to weigh expectations against one another: the totality of constraining circumstances may behave so that now it stands for the occurrence of one result, and now it corresponds to the occurrence of another. Yet we can also understand that this formation of expectation is unfettered. Everything that we know, and everything that we are warranted in presuming according to the rules of logical inference – all that provides us no lever to tell the difference between the operational range for one outcome and the range for another, or to tell that one range which spans those premises is indifferent to another range spanning those premises. This must be the case once the parts of the range to be compared (within a range which seems as broad as possible) are recombined in steady permutation, maintaining a specific ratio of size. I note that the idea which was held out before has now come to fruition here – against the general objection raised to the premise of indifferent ranges. In fact as was rightly shown earlier, it hardly ever happens that all portions of any whole range are to be considered equal in value. That is certainly not the case here, either. In the example of the ball rolled along the length of an alley, we suppose with maximum probability certain middling alley distances to be traversed, while we suppose other smaller or larger distances to be traversed with comparatively negligible probability. But these probabilities, which are vaguely characterized and wholly indeterminate, drop out of consideration. The formation of expectation is unfettered, since our example deals with the comparison of total values for these stripes of equal width which alternate steadily between black and white. Of course the relation of indifference can never obtain between any two arbitrary portions of the range, but it may obtain between all those portions which constrain a result of black and all those which constrain a result of white. Then this also means that all the considerations by which we could make this or that distinguishing characteristic of the constraining circumstances somehow probable – either a hard or a weak throw, or a cast of the die of this type or that – still seem completely irrelevant to the expectation of one outcome or another.

6. If then there really is a forming of expectation which depends essentially on the comparison of different ranges, then secondarily it will need to be asked if the ranges to be judged stand in a definite and expressible ratio of size. That means we ask if in the sense of the term introduced earlier, they are comparable. This question requires special treatment, too, since even if those values – which specify contiguities in time or spatial arrangement – do seem (easily and in general) to be commensurate, then still as we have seen, this thoroughgoing comparability is not to be found just anywhere. Rather there also exist real behaviors of different kinds, among which

distinct ranges of values can be compared to one another in many different ways, which means they are not comparable in the strict sense of the word. An example of magnitudes of the latter type has been given already, by the speeds of a system of point-masses. Even there, there is no golden rule to evaluate the balance between the expectation that the speed lies between 1 and 2 millimetres per second, and the expectation that the speed lies between 1 and 2 kilometres per second. Then for all those things which determine the outcome in any game of chance, there is no doubt that our vague knowledge of objectively manifest behavior remains fragmentary, over a great number of the most varied types of relations. It is quite the unsurveyable class of circumstances, therefore, which exercise influence over the regularity of our voluntary behavior, especially if they should be conducted ‘in an entirely arbitrary manner’. All such things are known to us only vaguely at the time, or not at all. Here sensory impressions come into play which occur in a snap, and which depend on an entire manifold of environmental states: the internal physiological conditions of the brain, the nerves, and the muscles are all to be considered. This in turn points to a widely ramified dependency: psychological states count as well, and as moods and streams of thought, etc. they are influenced by a host of prior conditions. However various they may be, one can always imagine all these things as partitioned into a very large number of separate categories of behavior. For each of them a certain range seems possible, even if the range is not demarcated at all sharply because of the imprecise scope of our knowledge. Obviously one can think of this partition in very different ways: one could pay attention to differences in temporal coincidence, or to differences in simultaneously occurring behavior; the latter may be illustrated in the most varied of ways, too. However this partition may be carried out, after all a complex totality of possibilities does arise by the combination of all those individual differences which seem possible. Right away one can catch a glimpse of a general principle, how two different categories of behavior may span wholly incommensurate ranges. Yet the particular characteristic of games of chance carries this moral with it: that assignments of probability which pertain to them do not demand such comparability in general. Instead one always depends on the comparison of ranges which span only a very few categories of behavior in any respect. That is because small variations in varieties of behavior (which are necessary to substitute the results of any one turn for that of another) can also be produced by small variations in the whole class of circumstances by which the movement is brought about. Actually it is very easy to say that the question whether someone will throw a 5 or a 6 in dice does not require us to decide between assumptions about two quite different varieties of behavior, nor do we call on an arbitrary state of pre-existing circumstances which determined these events. Rather it concerns only quite negligible differences which are very much smaller than the total scope of what seems possible. So if we think of the totality of all varieties of behavior as partitioned into determinate individual portions then at least if not all, by far most of them must be of the same type, in that a certain range seems possible for them. Within that range there is a comparability of smallest neighbouring parts. This must hold for all proportions which can be represented by a determination of extensive magnitudes or intensive magnitudes. Think of a temperature, or think of a speed: certainly, distantly separate ranges of

values are not commensurable, since magnitudes of this kind can be measured in different ways. Nevertheless it ought to be claimed, for example, that a range of temperature from  $5.65^\circ$  to  $5.66^\circ$  and one from  $5.66^\circ$  to  $5.67^\circ$  are to be considered very approximately equal, since by any means of measurement whatever this arrangement would be found to be equally comparable in extent. So in any determination of a continuous (in the mathematical sense of the term) variable, it can be claimed with complete generality that comparability exists between elements belonging to the same tiny range. For infinitely small regions this holds with absolute precision; within small finite ranges it holds to a good approximation that is nearly exact.<sup>5</sup> At this point I believe I can propose that this idea of comparability within local neighbourhoods may be assumed to have been adequately illustrated to serve as a postulate. It is evident that an exhaustive proof cannot be given outside a complete formal account of measurement theory. That would be beyond the bounds we must hold to here, in keeping with our specific task.

One is easily convinced that this property of comparability within local regions really does hold: the property of continuous variation among nearly all the points over which we can consider actual behavior to vary. If we should think of spatial arrangements or temporal relations; if we should think of the magnitude or decomposition of forces with kinetic energies; or if we should think of the amounts of substances which may ever be present in a mix – on first inspection the same characteristic is shown by all these examples. Such ratios can be attributed no less correctly to such conditions as admit no immediate representation in terms of spatial magnitudes, temporal magnitudes, or in terms of units of mass. The qualitative or quantitative differences between psychological states may provide an example, since for them too it is sufficient to have some functional dependence on material processes to derive a conventional measure – in the sense of an arbitrary representation in mathematical terms – under which smallest adjacent elements must be unequivocally comparable.

It follows directly that small neighbouring elements within a compositely determined range can be considered comparable in magnitude, even when the range contains very many items of the same kind in combination.

Likewise it seems understandable that as a consequence of the mixing of parts, that for the total values of ranges which represent various results, some expression for their relative magnitude may be obtained. It can be obtained once it is presumed that the ratio for regularly alternating elements is approximately the same everywhere. Here we have a not-uncommon finding in mathematics, that a fraction has an expressible value – that is, there is a fixed ratio of size between numerator and denominator – though neither numerator nor denominator represents a neatly expressible magnitude on its own. Other types of examples are simple to give, too, by which we may elaborate this relation. So we may think about someone whose worldly possessions consist of various elements composited together. Under such

---

<sup>5</sup>We will return later to discuss the circumstance that in games of chance, the comparability of ranges cannot be considered to be absolutely precise in the mathematical sense.

fictional conditions, money, real estate, cattle, etc. are items which could prove altogether incomparable in value. All the same it can be said the property owned by person *A* is double that of person *B*, if *A* has 20 silver thalers while *B* has 10; if *A* has two acres of land and *B* has one; if *A* has 100 sheep and *B* has 50, etc.

In a similar way, operational ranges which determine throws of 5 and of 6 are to be called equal, though no definite expression for them in terms of a definite unit of stateable magnitude is attached to either separately.<sup>6</sup>

On first consideration a slight difficulty seems to attend the notion that numbers attached to probabilities should always indicate the relative magnitudes of small neighbouring elements of a range. That difficulty arises from the very large number of possibilities which may be compared for probability under certain circumstances, especially if the results of many individual cases are taken into account. So for example, if a die is thrown 6 times,  $6^6$  equally possible series of throws may be obtained. Therefore it must be taken into consideration that the greater the number of

<sup>6</sup>I would like to try and draw a more intuitive picture of the ways that those parts which represent the various results of a game alternate with one another in a densely connected operational range of constraining circumstances. I will confine myself to the case of only two results, though – here we may think of the bowling game – which should be called black and white for short. Accordingly, the entire operational range of constraining circumstances might be divided into portions which are black-producing and white-producing. As far as its formation and its arrangement are concerned, it would be a mistake to presume that each of the constraining categories of behavior must have the following property: that if we think of their variation (and they are the only things varying) then this variation could only represent an oft-repeated alternation between black and white. Let us imagine that which we call the operational range – as was proposed earlier, on Chapter 2, footnote 5 – as a space of high dimensionality. Then the form and the placement of the black-producing and the white-producing parts need not be of the same type at all; they need not be such that any arbitrary straight line drawn through the black-producing and the white-producing parts will alternate through them many times. If all the differences of the operational range were encompassed by only a single relation – in other words if this could be represented as a straight line – then indeed the overall equality of the black-producing and the white-producing parts could only take on such a form that the whole line was partitioned into small pieces of one kind or the other. If instead two categories of behavior are considered, they may be represented as two coordinates of a plane. That result might occur, then, if there were multiple alternations in both directions, but also if they occurred in only one direction. In the former case the plane divides into small black-producing and white-producing regions, like a checkerboard. In the latter the plane divides instead into tiny and more-or-less stretched-out stripes. Given three determinants, which we may represent as the three coordinates of the imagined space of our example, the elements may be small in all three directions, or in two, or only in one. The space may then be partitioned accordingly, either into small cubes, or into narrow columns, or into layers. All these cases would be completely equivalent for the result. The greater the number of categories of behavior which are in play, the greater number of possibilities there are for these types of partition, too. Of course it is even possible that one type of partition obtains for one portion of the entire operational range, and that another type obtains for another. Then if we should claim that every black-producing element may always stand alongside a neighbouring white-producing element that is nearly equal, then more precisely it should be stated that: both elements are distinguished only slightly by their numbers of relations, be that a smaller number or a larger number, and they not be distinguished at all by others. Also, in a few or in many dimensions they may be smaller, broader, or narrower, but they share the same boundaries (or come close to sharing them).

determinate things in our purview (whose differences constitute the operational range), so too the greater the number of elements contiguous to a single element, hence which are in its neighbourhood. While differences in a single determinant can be represented by the points of a single straight line, so that it is bounded by two other elements; yet in a plane divided by a quadratic curve, whose coordinates we may represent by two determinants, an element is adjoined to 8 other elements. In a triply determined space, each element divides into cubes which adjoin 26 other cubes. Since our operational ranges always span very many individual determinants, it is understandable that so many neighbouring elements are indeed comparable.

7. From what has been said, it follows that the indifference of ranges should be considered to be realized – as their comparability is – from assumptions relative to the outcomes of games of chance; that is, if an approximately equal ratio is found everywhere between small neighbouring elements of the range which represent various results. Our present interest is focused narrowly on one question: if this can be assumed to be really general in nature, that is, if it can be assumed for categories of behavior which seem possible at an arbitrary stage of the constraining conditions. According to the earlier argument, this question coincides with another: if small and equal neighbouring ranges of equal value for any effect of arbitrary motions belong to small and equal neighbouring elements of that operational range. Of course it can never be proven strictly that this really is the case. Yet it does seem to be a pretty obvious and comprehensible assumption. Cursory deliberation tells us that this assumption is really remarkably plausible; it is hardly even imaginable that it is incorrect. Then we may notice that the steady dependence of one objective behavior on any other represents dependence that is at least not miraculous. Instead it is a kind of dependence which is often realized, of course. The content of the assumption in question is nothing but such continuous dependence. Further it may be understood that the manner in which the arbitrary movements in question arise here is very special, and not without importance for justification of that assumption. That is to say: those movements arise as the composite effect of a large number of factors acting in concert. The manner in which variations in each of these factors influence the result is independent of the contemporaneous behavior of all the rest, within certain bounds. So we can tell that the value on which the bowling-game depends – the strength of the component force along the direction of the alley – is determined on the one hand by the overall force of the initial throw, and on the other hand by its variable direction (again within certain bounds). But variations in direction influence the effect in a manner not changed abruptly by changes in force; rather they are close to being independent, whether the force is somewhat larger or smaller. Moreover we may remember that if the initial throw had been given as a certain type of motion, then the posture of the ball at that moment could still be different; so the force and direction of the impulse which moves the ball might vary. Finally it may be brought to bear that each movement is composed of the simultaneous contraction or contractions in quick succession of a number of muscles. Each may be considered to vary independently of the others, at least within narrow bounds. So actually a great number of individual processes act in concert towards the end of movement (which we finally consider) that is the component of

magnitude of the throw along the alley. In gross terms with such a type of convergence, there must always emerge a continuous probability by the various contributions of the effects. Given a factor for which probability is discontinuous, an entirely determinate value for the force of the throw may be considered to be especially probable – as a consequence of the continuous probability which then obtains for the direction of the throw. In nearly the same way that these contributions influence causally effective components when there is but little difference in force, then the discontinuity already seems to be compensated in some measure in the probability which arises with respect to one factor. Then for an effect which arises in the way described above – as a joint but independent effect of many individual factors – the continuous probability exhibited by a few of those factors is enough to make the probability of the total effect continuous, too. As these individual factors become greater in number and as they are more varied in kind, continuity can be ascribed to the probability of the effect with that much greater certainty, and in that much higher a degree.

We can express this overview of behavior in a stricter form as well, once we have recourse to the elements of circumstances which constrain an operational range. Given the interdependence of very many varieties of behavior, it will always be assumed there are at least a few of them which – within certain bounds – serve to alter the relevant effect of movement continuously, by their own continuous changes. – This is enough to establish an equivalence in magnitude of those neighbouring elements represented by equal ranges of values for the effect.

With that we are well-oriented to the most important points of the whole investigation: apart from several remarks, there is little more to tell.

We can resolve the question of the originality of these ratios of magnitude quite simply. First it may be shown that the considerations just established are entirely general. They also ought to count as established for earlier stages of constraining circumstances arbitrarily far back in time. Even from that much we know we are dealing with ratios of original operational ranges; this matter has already been settled, insofar as those observations are at all correct. If we should attempt a yet more exact demonstration of our notions in this regard, first we must clarify one matter for ourselves. When someone casts dice, of course the final result is determined not only by a state of affairs just past, meaning the state of all that which existed at the previous moment. The result is just as well necessarily determined by states of affairs which took place a year ago, or a 100 years ago. Yet the further back that we consider the conditions of a current event – which conditions pertain only to a relatively negligible portion of everything at hand – as they are traced back to earlier and earlier moments, the more numerous are the factors that might be considered, and at the same time the more negligible the differences which might have sufficed to condition appreciable differences among those events. Then that kind of dependency becomes more prominent which we claim to have established as occurring for ratios between ranges. Just this type of originality is found, which we had elaborated in our fictional story about range configurations. We do not reach any endpoint in our deliberations, to which we must trace everything backwards, beyond which it would be impossible to reach. The ratio that is relevant is definitive and

valid still, insofar as it holds necessarily in the same way at any arbitrary moment no matter how far back in time.

Finally, in very short order we can deal with the last of the postulates on which we found propositions about probability to be based. This is the postulate that the probability assigned to diverse effects of motion is to be represented by functions which have periodicity coinciding with fluctuations in the results of the game.

No substantive doubt can undermine the notion that this condition is fulfilled here. The composition of effects from many separate factors counts for as much here as it did for continuity of probability before. Even for a single factor of motion, such an idiographic contingency scarcely seems tenable: that regularly alternating ranges of value would ever represent elements that are now larger, now smaller in the operational range assumed to exist under the circumstances – and that these periods would match up exactly, in that their fluctuations would match the alternating results of the game. If such periodicity were realized, nevertheless it would be blurred by its joint effect with other factors, and it would no longer exist in the final effect. The more numerous and diverse the factors are which act in concert, the more certainly the development of such periodicities are excluded, as could make operational ranges unequal which represent various results of the game.

**8.** The harvest of this investigation can be summarized as follows: our conception of conventions about probability for games of chance is at least admissible in the sense given by the Range principle. In keeping with that interpretation, no significant underlying doubt can be accorded to the assumptions on which those propositions about probability are based.

On the one hand, with this we have a meaningful test of the principle. We have also come to the conviction that the logical relations postulated to hold for numeric probabilities are not mere fictions, either. Rather they really do emerge with specific subject matters. On the other hand we have achieved a deeper and more fundamental understanding of probability statements. It would astonish me if people did not find the interpretation we have come to give these propositions to be satisfactory, in the sense of being largely and immediately satisfactory. The claim that the probability of casting a one with a single throw of a die is equal to  $\frac{1}{6}$  proves to be objectively meaningful in so many respects. Once one has settled on this interpretation, it is hardly to be doubted that the whole complex of possibilities which we imagine for these constraining circumstances is somehow measurable. Accordingly the sixth part of this complex is ascribed the property of producing a throw of one. This intuition follows directly from the Range principle, given the presumption that we could gather that the probability statements are indeed correct. Nevertheless I could not have foregone the detailed exposition of games of chance given above. That was natural, since it concerned just how this business is to be understood: how the capacity to measure such a complex may be represented, and how we provide a foundation for claims about that measurement. After that task is completed, it is easy to express the conditions to which the establishment of numeric probability seems to be tied, in a way that is more general and easily understood. Thereby we garner valuable first cues towards a decision: whether there are other domains which exhibit anything similar.

In games of chance, events which are compared for probability in numeric terms – the results of the game – occur as manifold outcomes of possible varieties of behavior (which seem possible, given the constraining circumstances). It may be considered an essential property of games of chance that these events represent small portions of an operational range, in their regular combinations and in nearly constant relation of extent. Often we are able to form something of an idea to tell if something similar happens for any other events, on the basis of general knowledge of the way those events come about. It requires only a slight change in the formulation given above, to lend expression to this not altogether foreign way of looking at things. That is: the parts of the whole operational range which represent the occurrence of an event and the parts which represent its absence are folded into one another many times. That means nothing else than what one usually expresses by saying that the occurrence and absence seem to be possible in very many ways. This would be an initial condition necessary to assigning numbers in the measurement of probability. To that is added a second condition: that in general a very nearly constant proportion of size can be assumed to hold between the parts which occasion the occurrence of the relevant event and the parts which occasion its absence. This can also be expressed in a form more familiar to us. That is just to say our familiarity with constraining conditions is such that there are certain differences which are neither precise nor determinable in numbers. Some are most probable though they represent ill-defined and imprecise varieties of behavior, while others depart from the norm significantly and are less probable. One condition is that the relative sizes of these parts of the operational range which are not equally valued – the ratio of those parts which engender our outcome and those which hinder it – should be the same everywhere. Obviously this corresponds to another condition: that conjectures we can form with respect to gross differences in behavior seem irrelevant to expectation of the outcomes at hand. This holds for the bowling – game because black stripes and white stripes are the same in width everywhere, but it will be irrelevant to the expectation of results for black and white should we have a reason to suppose there will be either a forceful throw or a weak one. If their proportion were unequal – say if black dominated at the beginning and the end of the alley, while white dominated in the middle – then conjectures about the greater or lesser intensity of the throw to be expected would not count as indifferent for the expectation of black (or for the expectation of white).

In both these points there lie insights about the most important characteristics for ascribing numeric probabilities to games of chance. But against that, it may be overlooked that we permit ourselves to claim simple continuity of probability for any old objective behavior. Certainly we do so exceptionally often, without noticing that we cannot elevate those characteristics to be the numeric mark of probabilities in general.

Let us move on, to a demonstration of a theory of numeric probability based on the Range principle.

# Chapter 4

## The Special Theory of Probability



**Abstract** In the range theory, knowledge of objective import is the basis for statements of probability. The range theory provides a satisfactory explanation of the law of large numbers, and it clarifies the notion of likelihood. Statements about probability are associated with essential premises; pride of place is given to the assumption of independence among cases. Two types of factors determine probability: nomological and ontological. Nomological statements concern generally valid laws; ontological statements have singular meaning. It is our ignorance of ontological relations which makes outcomes seem uncertain. An event is called random, if more precise ontological determinants (withheld from our knowledge) are constraints on which an event depends. The notion of a normal distribution is introduced, by the example of series of draws from an urn. Different procedures for the draws are introduced to illustrate hypernormal and hyponormal dispersion for distributions. Hyponormal dispersion occurs when successive cases are not independent.

**Keywords** Law of large numbers · Randomness · Normal dispersion · Hyponormal dispersion · Hypernormal dispersion · Objective probability · Subjective probability · Nomological · Ontological · Span of possibility

1. If we say that a certain event is expected with a definite probability assigned a number, then as a consequence we must consider that a positive claim is implied: that the relevant event is necessarily tied to the nth part of an operational range shaped by the imprecision of our knowledge. As conjectured earlier, knowledge of objective import really is the basis for propositions about probability. At this point we should turn our attention to the objective content expressed by statements of probability. With a little effort we can recognize some essentially distinctive components. On one hand there is some knowledge of actual behavior. Its measure is instrumental to defining the range of possibilities essential to the determination of probabilities. That must be inexact – at least in part – in the sense outlined earlier. Yet for other reasons exact knowledge may also be required, for example knowledge about the ratios which determine probabilities. So in roulette, if we expect that the ball will land on

red (or on black) with a probability of  $\frac{1}{2}$ , this depends on familiarity that the red and the black pockets of the wheel are equal in width. It depends on knowledge that someone sets the roulette wheel spinning by a certain application of force, that the wheel revolves through a larger or smaller number of turns, and so on. If it is stated that the range consisting of original varieties of behavior is of such-and-such a size, and this represents some variation on the expected event, then this claim necessarily supposes familiarity with causal laws which hold in this domain. – Without entering into closer investigation of the contents of all the knowledge under consideration (we will devote ourselves to a closer investigation of that soon enough), we can still turn our attention to the way an old controversy is settled, about the subjective or objective nature of probability statements. Every proposition about probability implies an immediate and evident roster of cases that seem equally possible given one's current state of knowledge, and so the proposition has subjective meaning. Yet the establishment of such a roster is only possible with reference to definite data with objective meaning; the data are expressed factively by probability statements. Hence one can say they also possesses objective meaning *implicite*. Often this objective sense is frequently far more important. If we stick to the simplest of cases, then the familiar claim – that when drawing a black ball (or a white ball) from a given urn, we should suppose the draw to have a probability of  $\frac{1}{2}$  – means implicitly that equally many black balls as white balls are contained in the urn. Now it is evident that objective meaning, as an implicit factor in the probability statement, can be discharged or expressed fully without alluding to the subjective import of probability at all. Yet neither does it hold that this may be achieved until one has learned how to elucidate the objective relations which determine probability. Only a few items of objective content – such as the numeric ratio of black versus white balls or the well-balanced location of centre of mass in a die, etc. – can be named explicitly under ordinary circumstances. Still for other objective content – such as for relations of likelihood and independence soon-to-be-specified – one is in no position to characterize them, other than by the manner they affect our expectations. On the one hand it might be said: “The probability of throwing a 5 or a 6 with a single die is really equal – that much is objectively correct. Neither one nor the other is more likely, or objectively more probable. In a similar sense, a well-defined value that is objectively correct can also be assigned, i.e., for a probability when an observation is conducted on a result that is to some extent inaccurate.” By this expression, one urges that there is an objective meaning which is contained tacitly in these probability statements; one would not like to unshackle this from its basically non-essential link to probability. On the other hand it may justly be repeated that all probability is subjective, as the expression and consequence of our imprecise or incomplete knowledge. By contrast, the notion of ‘objective probability’ is nonsense, a *contradictio in adjecto*. Since the former conception seemed every bit as indispensable as the latter was incontrovertible, their contradiction remained unresolved. That contradiction has made itself felt more or less clearly, and extraordinarily often. It has been a primary motivation for ever-renewed attempts to reshape the principles of the probability calculus in a strict and satisfactory manner.

2. The investigation of games of chance in the previous chapter taught us that there really are circumstances under which numeric probabilities may legitimately be assigned. Closer examination forces us to modify this result just a little. We notice that the uncertainty we encounter in our dealings with the results of games of chance do coincide with the state of knowledge denoted by the Range principle. That is to say they coincide to a very close approximation, but not completely. It is not difficult to prove that such accuracy can never be attained in an absolute sense. Thus numeric probability represents an instance of ideal logical operations that can be accurate to an extraordinary degree, but which never achieve absolute precision. In assignments of numeric probability, if distinct cases are deemed equally possible by appeal to a state of knowledge which is objectively valid, then expression has been given to a mental state for which the content of that very knowledge counts as certain. If we know that an urn contains equally many black and white balls, and the probability is deemed the same that a draw currently taking place will deliver a black ball or a white ball, then there is no room in that arrangement for consideration that the number of black and white balls may be unequal, at least no more than that some unknown effect of the black or white color acts on our sense of touch, and lends an advantage to the drawing of one kind of ball.

It is easy to see that these kinds of objective postulates are essential not only to the cases we examined earlier – rather, they are always essential to propositions about probability. Every time a claim is invoked that the relative magnitude of ranges is original, as a rule this involves a not insignificant bit of objective knowledge. One might hope to have transcended such an assumption, if the formation of expectation were to involve some magnitude that can be regarded as a constant of physics. Yet if a probability of  $\frac{1}{10}$  were proposed for the state of affairs that a body has a specific gravity that has the numeral 3 in its third decimal place, that would seem unsupported by any objective hypothesis. But here too a more thorough examination cannot escape the notion that propositions about probability can only be considered an exact formal expression of our mental activity. That holds especially if we should deny certain assumptions to be objectively correct – if the assumptions do not correspond to reality, though they can well be imagined – meaning that their correctness depends once again on a certain amount of knowledge. We may think that specific weight might not be a physical constant after all, but that it may represent a property subject to many kinds of variation. And one might think that if one were in the right position and had the right opportunity, one could attribute different specific weights to bodies for one purpose or another – at least under some assumptions whose consequences are left unconsidered, since it appears sure enough they are incorrect.

Doubtless just any knowledge of the kind which has objective significance does not provide absolute evidence: strictly speaking, it is all subject to some uncertainty, be that ever so slight. Then no numeric probability really expresses the whole range of things which may be imagined, in all completeness. Forecasts are always made under specific assumptions which may be so enormously improbable that our habit of neglecting them no longer seems unjustified. The fact that we may arbitrarily

trivialize, but never quite do away with our awareness of this incongruity is interesting in a couple of ways. The certainty we attribute to objective postulates in reporting numeric probabilities is always based on logical arguments which produce merely non-numeric probabilities – particularly with conclusions from analogy and inductive inference, as we showed earlier. If numeric assignments of probability are to be made, such arguments would actually have to be left out of consideration altogether. That is how we arrived at the expression that such determinate assignments seem confined to the free formation of expectation. With respect to questions of real significance, there can be no mental state for which those non-numeric probabilities really have been eliminated entirely. Still the special means by which we transcend consideration of non-numeric probabilities is that they embody a certainty that excludes any doubt worthy of mention. – Obviously a minimal uncertainty in objective premises is irrelevant to many cases in practical terms. Nevertheless it cannot be ignored that the larger the numeric probabilities based upon them, the more careful consideration is called for. One must always keep in mind (and this is a rule of exceptional importance for the probability calculus) that probability statements may simply be false – not merely as a consequence of a logical error or cognitive bias, but as the consequence of factual error. If a probability statement tells us that some event is very improbable, then it ought to be examined whether some mistake or other has not been excluded carefully enough from the premises on which that proposition rests. If that is not the case, then the assignment of numerals to probabilities has no claim to be meaningful.

If very many balls are drawn from an urn that contains a known number of black balls and white balls, we obtain huge probabilities for certain ratios in the total of all the drawings. The higher these probabilities are, the more often we should be asked how the number of balls has been determined, and then if error is excluded in that respect, should we want to use the probabilities for a logical or practical end.

The notion that has been developed here, of the relation between propositions about probability and their associated premises, will seem paradoxical to many on first impression. That will be the impression of those who are accustomed to read an immediate and conclusive expression of knowledge into probability statements. Nevertheless I believe that this notion is incontrovertible. How one comes to pronounce common assignments of probability – for example, about repeated throws of dice made in succession – is entirely incomprehensible if one does not regard it as completely certain objectively that no biasing conditions hold for any throw in particular, but yet one does not hold this assumption itself in doubt. As soon as one has clarified this issue, one has to admit that assignments of probability always rest upon assumptions in a similar way. The certainty of those premises is indeterminate, though often they can have an enormously high degree of certainty.

In still another regard, the representation of probability in numbers represents an ideal case, which is only ever realized approximately. That is to say, this determination in numbers cannot ever be mathematically precise. Rather it is always characterized as an approximation, one valid to an exceptional extent. Let us return to our example of the bowling – game introduced in the last chapter. However certain it appears that the probabilities estimated for two adjacent stripes of the alley could hardly be distinguished in any significant way, it can only be concluded that

the total value for black and for white must prove to be equal to a fine approximation. Exact equality cannot be claimed. Further, the explanations by which we sought to render the notions of comparability and originality comprehensible for operational ranges are themselves of the sort that they do not bulwark statements about absolutely precise magnitudes – of course they hold roughly. The comparison of ranges still remains arbitrary to some extent, though that be minimal. Even their relative magnitude has to fluctuate as constraining circumstances are traced back in time. It is easy to see that this indeterminacy is no exception that places the value of those numeric proportions in question. Rather it is a property attaching to almost all numeric values in the same way – those which have real meaning, rather than purely mathematical or logical meaning. Then it needs no explanation that the true value of a real magnitude cannot be reported with absolute precision, because of the uncertainty inherent in any measurement. A few simple considerations show that similar conceptual indeterminacies attend any such magnitude. This holds for every concrete magnitude, in the most pronounced way. When one tries to determine the surface area of a country, or a person's height, or the weight of a stone block – these are not purely and sharply defined in conceptual terms. But even for values of more general significance – say values assessed by physics or chemistry – as a conceptual matter one is conditioned not to press much beyond the precision of measurement which that determination may achieve. In the definition of specific weight, for example, it is assumed in each instance that the unit of measure should be given for water at  $4^{\circ}$  of temperature. One neglects to add the current barometer reading, as well as the fixed mass and form of the container in which the water is held, though those circumstances change its density to a minimal extent.<sup>1</sup>

3. Now since we are generally acquainted with the relations between propositions about probability and their objective premises, let us turn to discussion of some real relations which constrain important forms of assignments of probability. Pride of place is earned by a property we are given to call the independence which several cases have to one another. First we would like to explicitly call the characteristic a property of assignments of probability, and then we can be acquainted with its basis in fact. That is just the characteristic we had earlier declared (Chap. 1, Section 3) not to be explicable by the Principle of insufficient reason. To formalize the property more generally, it consists in this: that if  $a_1 \ a_2 \ a_3 \dots$  stands for a series of equally possible cases, as does  $b_1 \ b_2 \ b_3 \dots$ , as does  $c_1 \ c_2 \ c_3 \dots$  and so on, then the manifestation of any combination  $a \ b \ c$  whatsoever is to be considered equally probable. If  $a$  stands for the outcome of the initial throw of a die, and  $b, c$  etc. for the outcomes of the second, third, and subsequent throws, then with that general property, any arbitrary series of results would also be declared equally probable. To put the matter shortly, the ‘independence’ that we attribute to various cases, just means equal possibility of combination.

---

<sup>1</sup>A fully precise report of numeric probability might be possible in ideal cases, as in a game of chance in which various results alternate in an infinitely small period, such as for a bowling-game where alternating black and white stripes are infinitely thin.

If we try to specify the real behavior this presumes, we note that the ‘independence’ of single cases should not be taken in the strict sense, but may be considered an improperly shortened form of expression. If we wanted to base a relevant assignment of probability on the lack of connection among various isolated cases, it is not just that such should be excluded when an actual process directly influences possibilities which arise afterwards (as when balls from an urn are not replaced each time over many draws). Rather there ought to be no factor at all that complexes of constraints have in common for several cases. Individual cases have to be fully independent in the sense that – as on an earlier occasion – we proposed a counterfactual and total lack of connection between two series from different processes. Under this assumption, the correctness of that assignment of probability would be immediately evident. Yet it stands to reason this kind of connectedlessness never obtains over many cases at all. Specifically it never obtains over many cases of the same kind. Only then would it be correct to say that the assignment of probability is made as if the cases were entirely independent of one another.

Nonetheless the objective behavior which comes into question here is not difficult to specify more closely. Evidently this consists initially of the fact that the outcome of each case is conditioned at least in part by other factors. Even if we should fixate on an arbitrary antecedent stage of constraining circumstances, there must always be a multitude of constraints that influence the outcome of just one case. Accordingly, certain combinations of antecedent forms of constraint represent combinations of outcomes. There are always a rich number of factors which are simply not understood, or not fully understood in each case. Those stand in contrast to other factors of general significance, that is, which coincide in many cases or in all, and which also act to determine outcomes. For example, when we are throwing dice, those factors include: the size and weight of the dice, and the elasticity and roughness of the table on which the dice are cast, together with general laws of motion, and so forth. Then if we declare arbitrary combinations of individual outcomes to be equally possible, this means simply – something that is both evident and indubitable in games of chance – that innocence about individual varieties of behavior (idiothetic to each case) results in a definite ratio of probability for the various results of the game, without regard to those general determinants. If we knew exactly all that may be held consistently in most or all the cases, expectations of outcome would still remain just the same. Everything which has more than individual significance must either be known with certainty in advance – as was the location of the centre of mass for the die, or the numeric ratio of black and white balls in the urn – or else it is completely irrelevant to expectation of the outcome. The question raised earlier – why a rule of analogy is not applied under these circumstances (why, for instance, after repeated throws of 6 with a single die, the same result of 6 is not expected with a larger probability than otherwise) is a question rather easily resolved. Essentially positive knowledge of the prevailing state of affairs supersedes inference by analogy. The circumstances of any single case are known to us with better certainty, since they are based on a definite expectation of various results which adhere to a certain rule: to the Range principle. We know that the actual courses of events already observed give expression to certain types of real behavior, but they certainly do not contain valid determinants for

subsequent cases, which are meaningful for expectation. It really does hold that the actual course of one case (or many) does not affect the expectation of subsequent cases in any way. Since it is desirable to have a short expression for the said characteristic, let us stick with the usual one, and speak of “independence in the sense of the probability calculus” – or else where there can be no misunderstanding, we can simply declare several cases “independent of one another”.

4. It becomes especially meaningful that we can combine probability statements for many individual cases – to combine them with one another in the way just described – when cases of the same kind are in question. The singular meaning that other factors have does rest on another characteristic, one which is grasped only with effort. I do not think I can forego discussion of this subject, partly because it bears a certain interest for the present discussion, and partly for its importance to discussion of the notions of chance and randomness later on. To be clear about what is at stake, first we must hone a distinction we have alluded to many times. We distinguish two essentially different parts of the business of knowing about reality. On one side we investigate laws by which individual objects change their state, or persist: laws by which they affect one another, hence laws by which the whole panoply of events proceeds. This can also be characterized as an exploration of the properties of existent objects. If our knowledge were complete in that regard, still it would not do us any good, if we did not have additional information about the way things behave. That information provides a point of departure from which we may imagine changes flow as determined by those laws. Acquaintance with the law of gravity tells us nothing about the actual motion of the planets. In order to gauge those motions, we also have to know which masses are present, and in which states of motion and spatial distribution they have existed.

I would like to lend the adjectives nomological and ontological to these kinds of determining factors, since I do not find sufficiently short and meaningful expressions for them in the literature. We are used to assuming – as an indispensable principle of all our knowledge of nature – that reality presents us with repeated instances of things of the same kind. Laws governing events can be established as valid for many events. An experience which we may have had, can also be applied to certain other events. Thus nomological propositions will always be taken to be generally valid, i.e., assumed to hold for a whole category of things that have an indefinitely large class of exemplars. Ontological determinants have always and only a singular meaning, by contrast. Since they represent determinants that may change in every particular over the course of time, then the conditions which have occurred for a series of things do not immediately produce an expectation – even for things of just the same kind. Ontological determinants cover that which is purely factual, namely that which cannot be traced back to what is general and necessary. Suppose that the presence or absence of an outcome is determined by the merely individual constitution of a case, and that despite any correspondence of properties for the objects in question, yet their evolution can still vary from case to case. Then it is clear that it is the ontological determinants as merely factual determinants of single cases on which the outcomes depend. It is our ignorance of ontological relations which makes outcomes appear uncertain to us in each case. However directly illuminating our

data might be about historical connections, causal laws, and nomological propositions, uncertainty may not be brushed aside whether a ball lands on black or red on the current spin of the roulette wheel. Instead our information is incomplete. We do not know how constraining circumstances would behave in new ways from case to case, since they are variable by their nature. This ontological knowledge is so incomplete, that even with the best information about all the laws governing these events, the probabilities of outcomes would be represented just the same way.

The distinction between nomological and ontological finds its nicest expression in the often-employed and highly meaningful notion of objective (or physical) possibility. One cannot say baldly of any event that its presence or its absence is objectively possible. Either one or the other is necessary, or else one or the other is impossible, given the totality of bounding conditions. For a single event, possibility is always an expression of insufficient information, just as probability is. Things are different when we are dealing with an event under a general denotation. That an event might be present or absent under certain circumstances, or that both are possible objectively: that is a conjecture with a completely tenable and comprehensible meaning, given that its characterization of the bounding circumstances is general and imprecise. That characterization covers a number of different varieties of behavior, or a range. Such ‘generality’ of determining conditions occurs very often, although by no means exclusively, in this way: only a portion of the circumstances necessary to the outcome are determined, while another portion is left undetermined. The description of these circumstances is only ‘abstract’ in the sense it is partial. It covers all varieties of behavior which could arise under circumstances left indeterminate. – Now under this assumption, any conjecture about objective possibility is less than meaningless. This has nothing to do with the other assumption: that because of some lack of information, we are in no position to predict the presence or absence of an event. If one calls it objectively possible, that in dice a 6 is cast ten times in succession, one wants to emphasize that nothing is covered by the bounding circumstances – which are set out in wholly indeterminate fashion – such that when the die is cast ten times, the result would be excluded. Under certain imprecisely specified conditions, we call an outcome objectively possible if more thoroughly determined circumstances are imaginable which would effect the outcome. However, this always presumes the circumstances are represented as indeterminate in ontological status: that ontological determinants are imaginable, suited to lead to the outcome under valid nomological conditions (which could actually obtain). Given a body that really floats on water, we call its sinking objectively impossible, since that would be effected only by altering nomological determinants (its specific weight) – by changing things that are actually inalterable. Should we not know what the measure of a body’s specific weight is, and should that body’s sinking seem possible as a consequence, perhaps one ought to call such a possibility ‘merely subjective’.

In claiming that an outcome may be possible under generally denoted circumstances, by implication what is said expresses knowledge of nomological content: some information about laws governing events. We have only to take a step further to grasp another idea which will prove important. A determinate value in numbers

can be given for a possibility represented by certain general conditions (in the sense given above) of the presence or absence of an outcome under some circumstances. It is not difficult to see how this is to be understood. If we think of ontological determinants as being varied within a frame traced out by general conditions, a definite proportion of those ranges represents the presence and the absence of the outcome. Then we could say that the general condition of casting a die represents the possibility  $\frac{1}{6}$  for each of the outcomes 1 2 3 4 5 6. This notion of a possibility reportable in numbers, one which stands for any general condition of the outcome, is one to which we will return often. It is essential to this notion that the ranges which are its foundation and which it expresses should always and only represent variation in ontological relations, of bounding circumstances. Only presuming very particular nomological relations can a definite ratio be claimed for original ranges representing various outcomes. If then one speaks of a possibility which certain circumstances represent for an outcome, this is always to be considered as supported by quite definite laws of events.

**5.** Perhaps the most important impetus to logical understanding given us by familiarity with the Range principle, is the satisfactory explanation it gives for the so-called **Law of large numbers**. If a certain issue is expected with probability  $\frac{1}{n}$  for one case of a specific kind, consequently it is assumed with the greatest certainty that for very large numbers of cases of the same kind, approximately the  $n^{\text{th}}$  part of them will prove to have that outcome. One calls this the Law of large numbers, wishing to draw attention to a regularity: that despite indeterminacy of expectation in each case, still a result of near-absolute certainty may be conjectured for very many cases. Once we know what it means for probability is to be assigned to one case, it may be understood how – for an independent series of the same kind – expectation is based on the Range principle for many cases, as for single cases. From all of this a theorem of pure mathematics may be derived. Fundamentally the proposition in question is a theorem of combinatorics. As a formula of abstract mathematics, it can be expressed like this: If one forms all possible series of elements  $a_1 a_2 a_3 \dots a_n$ , and these series contain a large number  $p$  of elements overall, then the number of series which contain a particular element  $a_k$  approximately  $\frac{p}{n}$  times is preponderantly large. The ratio of the number of series which contain  $a_k$  more often than  $\frac{p}{n} - \delta$  and less often than  $\frac{p}{n} + \delta$  (when their limit is defined) to the total number of all series, is a ratio which approaches a value of one as the value of  $p$  increases without bound.

If one considers  $a$  equal possibilities which obtain for cases of a certain type, then each series containing a number  $p$  of those elements in any order, represents equal possibilities of that issue over a series of  $p$  such cases. The larger number of those which contain the single element  $a_k$  approximately  $\frac{p}{n}$  times, results in the very large probability that for the types of outcome designated as  $a_k$  – which have a probability of  $\frac{1}{n}$  individually – the outcome occurs approximately  $\frac{p}{n}$  times in  $p$  cases.

Thus there is an expectation – to introduce a simple example – that in 100 spins of the roulette wheel, the ball lands about 50 times on red and 50 times on black. This expectation is based on the circumstance that every series of events which could take place in 100 spins is equally probable. Among those series by far the most frequent

are those in which the ball shows up about 50 times on red and 50 times on black. It is not difficult to recognize a basis for this in the large-scale permutability of series. 50 elements of  $a$  and 50 of  $b$  may be arranged in  $\frac{1 \cdot 2 \cdot 3 \cdots 100}{1 \cdot 2 \cdot 3 \cdots 50 \cdot 1 \cdot 2 \cdot 3 \cdots 50}$  different ways, while 90  $a$  and 10  $b$  may be arranged in  $\frac{1 \cdot 2 \cdot 3 \cdots 100}{1 \cdot 2 \cdot 3 \cdots 90 \cdot 1 \cdot 2 \cdot 3 \cdots 10}$  ways. The ratio of the latter number to the former is something like 1 to 5800 billion. This result is important, because it shows us how expectations about the totality of results over a large number of independent cases of the same kind, are always based on the Range principle. Consequently probability statements all have the same meaning, though they make reference to a single case or to many uniform and independent cases. If I expect to cast approximately 500 heads when I toss a coin 1000 times; or if I expect not to find a white ball in one drawing from an urn containing 10,000 black balls and a single white ball, but expect to find one of the black balls: in both scenarios the expectation is of the same kind, based on the same principle. Of all varieties of behavior which seem possible, it is by far the largest portion of them which leads to the expected outcome. In neither scenario is it known what may have excluded the unexpected outcome. But it is a very improbable outcome under either scenario in the same sense.<sup>2</sup>

**6.** We have already had occasion to allude to the concept of likelihood. Still we must pursue the notion more fully than we could have earlier. It is not always true that if an outcome of a pair is more probable than another, that we call the former outcome the likely one. Although it is more probable one throws a combination of 7 with a pair of dice than it is to throw 12, one does not usually call these likely conditions for a throw of 7. Instead one restricts use of this expression to situations where two probabilities prove to be unequal on closer investigation, which might have had some justification to be equal. This occurs when two varieties of behavior span equal ranges but unequal probability, because the ranges are not original. Between two varieties of behavior which span the same contemporaneous ranges, the one called ‘likely’ corresponds to a larger original range. If a die is geometrically regular but not regular in physical terms, so that its centre of mass is positioned eccentrically to a significant extent, then the form of the die does not provide us a standard for the probability of various throws, as it would normally. Despite the geometrical symmetry of the sides, different throws are not equally probable: they represent unequal ranges of the possibilities which determine the throw. So there are some throws we can call more ‘likely’ than others. In an important objective sense – in the way one speaks of likelihood – they do not exhibit the proportions which

---

<sup>2</sup>From this it can be seen that the label “the Law of large numbers” is really not very suitable. That is because there is basically nothing here except a proposition of mathematics that exhibits certain peculiarities which follow from the size ratios of ranges and consequently from the expectations based on them. It is true that the Law of large numbers has often been taken for a proposition of far greater import. That is, one thinks it legitimate to claim that in the phenomena in any area there must be found a repetition of cases of the same kind, so that for a great number of such repetitions, an arbitrary characteristic may always be observed which occurs with approximately the same relative frequency. Of course that is something altogether different than a mere mathematical relation; that is a claim about a very specific regularity of nature. Yet as we shall see later on, this proposition is not at all valid in general, and the general proofs which have been sought for it are based on fallacies.

belong to games of chance. For example, with any process there are possibilities which could come about at any time of day. If observation tells us that a large number of cases are not distributed almost equally over the 24 hours into which our day is divided, but instead they are unevenly distributed in a significant way, we conclude that something is at work which makes the exclusion of those varieties of behavior likely at certain hours – or, as one might say – which makes their omission more probable objectively. So we know that the final labour of pregnancy may happen at any time of the day or night, but there is a greater likelihood of it in the evening hours or the night. That simply means the operational ranges are not the same as those given by occurrence of the event at any hour of the day or night. Instead they correspond to unequal and original ranges of behavior.

Very similarly we find that a somewhat different idiom for likelihood can be explained as well, as when one speaks of a circumstance making an event more likely. One might say of the example just discussed, that an eccentric location of the centre of mass makes a certain result likely to be cast – say a result of 4. In most cases it would not be sufficient to read that the relevant circumstance leads to the event, when that event would not have occurred in its absence. The reason is that the eccentric location of the centre of mass produces a result of 4 under circumstances which would have produced 1 2 3 5 or 6 otherwise. Still it is not less effective in producing 1 2 3 5 or 6 under circumstances which would have produced 4 with a centrally-positioned centre of mass. So it is essential to add that the former cases are more numerous – correctly stated, they are more probable. Basically what is meant here, is that with addition of that determinant (an eccentric position of the centre of mass) the original ranges which produce 4 are relatively larger than when they are not present.<sup>3</sup> It is essential to the correct evaluation of expressions like these, that there always must be found some tacit premise about other conditions – conditions to be added to the likening circumstance – which might be supposed to substitute for the lack. This is not always evident. Often enough, it may not be asserted or denied – in an absolute sense – that a specific circumstance makes an event more likely. For example, whether travel lengthens one's lifespan is just too vague a question to admit any answer in general.

As one can see, often it is just these very important dependencies which one habitually takes to be characterized by their importance for expectations, though one is very likely conscious of their objective significance in general. Without calling on the notion of ranges, though, it is impossible to express that objective significance directly.

For this reason it seems to me eminently clear that these forms of representation are indispensable to the notion of likelihood. At the same time, we have seen how essential it is to the principle of expectation that ranges should be determined to be original. The fact that the delivery of a baby might start between noon and one in the afternoon, or the fact that it might start between 8 and 9 in the evening: these are initially varieties of behavior which span equal ranges. Yet they do not represent

---

<sup>3</sup>This is the same condition that **Poisson** characterizes when he calls that circumstance a cause “in the larger sense” of the event in question.

equal and original ranges: it is just this which is signaled by the unequal probabilities attributed to them.

7. In our treatment of probability statements, until now we have completely avoided one question: Who should be considered the subject of probability – for whom should probabilities count? Until now it was unnecessary to introduce a specific premise to this effect, since if a probability statement emerges as the expression of a specific state of knowledge, naturally the task of logical investigation is just to establish which state of knowledge this may be, not in whom it may reside. Had propositions about probability been so restricted in scope that they represented only mental states of individuals, one might have had very little opportunity to engage in calculations of probability. When outcomes have more extended validity, or very general validity, quite a different interest must be attached to those methods. This is obviously the case as soon as the framework of knowledge (under which those methods are valid) is manifest in many individuals, or in everyone. If imprecision of knowledge lies at the heart of probability not as an individual matter, but as a general matter; if the acquisition of exact knowledge on which the sure prediction of outcomes could be based is something which transcends human capabilities altogether, or at least which goes beyond our means to date; then the relative sizes of ranges for the presence or absence of an outcome will provide the only general measure of expectation for outcomes. Let us recall the special circumstances which have always been attached to determining probabilities in numbers. It is too easy for us to say that general validity is not the exception, but rather an ordinary thing which holds for the majority of propositions about probability. That is because constraints on numeric probability – how an outcome may be influenced by small changes across a great number of circumstances – evidently implies that any exact specification of circumstances must be a task of really enormous difficulty to transcend the unrestricted formation of expectation. Add to this the additional circumstance that if such similar cases present themselves frequently and repeatedly, in the inexact inquiry of which we are capable, the same expectations are based on the same ratios. From this, reports of probability are possible which are generally valid in two senses: either they are valid for all cases of a specific category, or they are valid for anyone who wants to form some expectation of the arc of such a case. In that sense, it is generally valid that the probability of throwing a 4 with one die =  $\frac{1}{6}$ . This insight serves in the explanation of two concepts which play a large role in probability theory. The first is that of “chance”, or as one may say with emphasis, “objective chance”.<sup>4</sup> Ordinarily what one indicates by this, in reality is nothing other than a probability generally valid in the sense given above. If one then contrasts chance with other probabilities, they are probabilities of merely individual significance,

---

<sup>4</sup>Given the appreciable vagueness in terminology which attends literature on the probability calculus, it is clear that the word ‘chance’ is used often enough purely to mean probability, and not used in the particular sense that has been given here. Nevertheless it is always ‘chance’ that is spoken of where the general validity of a probability is to be emphasized. Then since a short expression is altogether necessary for the concept which has been elaborated above, I believe the term ‘chance’ should be applied to it.

which may be changed by the growth of knowledge. Then if we say that the probability of throwing a combination of 12 with a pair of dice is counted as  $\frac{1}{36}$ , and that this is the objective chance, that ought to mean: “as soon as a pair of dice are cast, occurrence of that result is expected with probability  $\frac{1}{36}$ . This is settled once and for all. No investigation of the specific conditions of a single case changes anything about it.” Accordingly when an event is reported as chance, that means specific considerations have been dropped as superfluous. They do not impend on the issue whether the relevant outcome could occur in a particular case.

If chance represents a kind of probability valid once and for all, holding for all cases, then it has already been declared that observation is pointless. Whatever is actually done in a number of such cases, they do not serve to modify expectations about future cases. From this it is to be supposed that chance always represents proportions of ontological ranges of behavior. That is why they stand in close relation to randomness, as we shall see shortly. – This same circumstance (the general validity of certain probabilities) renders comprehensible how it can be said that the probability an event has under specific circumstances may be ascertained. It may be established by experience. One ought to attribute a definite value of probability in an absolute sense, if a tacit premise has been made about the state of knowledge it ought to represent. Such is the case here: if the basis of a very specific expectation for a process is what can be ascertained at all in any single case, then surely that expectation represents knowledge beyond which we cannot go. It is the object to be determined. One may describe this state of knowledge narrowly by saying it covers everything which obtains for all individual cases under the circumstances.

**8.** We should not neglect to review the origins of the concept of randomness in approaching a theory of the probability calculus. As much as it might seem advisable that one avoid this equivocal word entirely, on the other hand it is altogether imperative to call common ideas by common names, to ease one’s understanding and uptake of theoretical considerations by tying them to popular currents of thought (insofar as that is even possible). Accordingly the concept of randomness needs to be discussed, since it is essential to our subject. An elaborate explanation may be waived, especially since a full and excellent exposition has appeared only recently.<sup>5</sup>

Among the many qualifications expressed by the word ‘random’, one which most likely holds steady is that it negates some necessity or lawfulness. Proceeding on that premise, it is understandable if an event is called ‘random’ relative to some class of circumstances. What is meant, is that it is not these circumstances (but some others) which are considered necessarily tied to the event – that is, to which it can be traced as if to a cause. One might say that changes in weather are random relative to phases of the moon; or that the spread of Phylloxera is random relative to existing freight transport routes, and so on. If one inquires into this idea of relative randomness more closely, a stable meaning presents itself only when it is claimed there is no

---

<sup>5</sup> Windelband, W. (1870). *Die Lehren vom Zufall*. Berlin: F. Henschel, 80 pp. (Diss. Univ. Göttingen).

connection at all between two series of phenomena. But this is strictly never realized, and the conditions to which the term is applied, are simply those we had discussed earlier in the context of likelihood. The event would be random with respect to certain processes, if they do not have spatial or temporal relations which are ‘likely’. That event would be random with respect to a specific circumstance if presence of the circumstance represented neither its likely presence nor its likely omission.

Yet this sense is not foremost when one connects the concept of randomness to the probability calculus. Strictly speaking another sense develops from incorrect use of the word. In what we are used to calling games of chance – perhaps for many other kinds of phenomena – one observation is paramount: that this outcome or that is not bound by necessity to any appointed property of the constraining circumstances. If one expresses this by saying in an absolute sense one deems the outcome to be random, one loses sight of the fact that it is also necessary with respect to all the constraining circumstances. With this insight, in keeping with the original concept, one can say nothing at all may be absolutely or purely random. The term has long been used in that much to indicate the characteristics of dice games or other similar games. It is used in that way more generally and frequently than in its simplest, original sense. One might say it has acquired a modified sense. Currently it may be applied in this new sense to characterize important provisions under which some events might arise. The demystification of this sense is the task now open to our interest.

If some phenomena are attributed randomness absolutely, rather than with respect to a distinguished category of circumstances, this can often be understood in the sense that the totality of conditions – insofar as they are open to our knowledge – does not involve necessity, either for the presence or for the absence of the event in question. Necessity would not be asserted (it would be negated) for more or less vaguely conceived conditions, compared to the exact standard of knowledge we can attain. If we claim this holds in general for an entire category of indefinitely frequent repetitions of the same kind of cases, then a further view has been invoked: that unknown to us, the individual determinants of single cases always determine relevant outcomes. On that view, it follows that an event ought to be called random, if more precise ontological determinants (withheld from our knowledge) are the constraining circumstances on which presence or absence of the event depends.

Suppose experience had taught us that outcomes of 6, 5, 4 etc. in a game of dice are tied to particular circumstances open to observation. Then we would be in a position to foretell outcomes, and it might be said that this was not really a game of chance in the sense given earlier. One can confirm the validity of the explanation given, in that for us, there generally seems to be a logical characteristic bound up with the randomness of a particular outcome with it, which does not permit the progress of one case or several cases to sanction a subsequent modification of expectation for the next. In fact we are used to brushing aside this kind of conclusion by analogy, with the phrase that the whole course of events is based on randomness.

If one takes the word ‘random’ in this sense, this much is sure at least: foremost, the probability calculus must be concerned with domains of randomness. – We have noted that of all the numeric probabilities, those which represent chance are by far those which receive greatest interest. We have noted the reason for such interest, too. If we compare the explanation of this early concept of chance to the definition of randomness just developed, then we have a clear overview. Chance always represents the probability of events whose presence or absence depends on randomness. The often-heard saying that probability calculus is an accounting of randomness appears justified in a very general sense at least: that in terms of calculation, it deals with domains subject to randomness. Firstly, on this subject it may be emphasized that there can emerge other numeric probabilities – we will become acquainted with some – which do not bear the same relation to randomness. Secondly, there is no point in establishing chances in numeric terms generally when we speak of randomness. In other words, not every sort of randomness is the subject of calculation.

Naturally one might have wished to attribute a particularly high level of achievement to the probability calculus as a rule, since it has been described as an accounting of randomness. Understandably this was meant to say more than that the subject of its study stands in some relation to randomness. Confirmation by observation – as found in the end products of the probability calculus by the average results of many cases of the same kind – seemed to be a special triumph, leaving aside the incalculability of the randomness of individual cases. One might even see it as a triumph for which much is owed to the probability calculus, since laws could be established to which “even randomness must be subjugated”. That idea can only be recognized as correct in a fundamentally different way. Certainties which we attain in domains of interest have only limited significance: even when outcomes turn on randomness, it is our familiarity with ratios of ranges which often permits our expectations to be very highly certain. What is characteristic of methods by which we become masters of randomness – so to speak – consists entirely and simply of the systematic formation of expectations, based on the Range principle. The probability calculus is nothing else but a systematic accounting of ratios of ranges, strictly speaking. Surely it is a very misleading expression, then, if one should call the probability calculus an accounting of randomness. Strictly speaking, one can claim it only tallies our expectations about random events. On occasion, it permits the formation of expectations to which great certainty is attached.

Not infrequently ‘random’ is used with an essentially different meaning. Substituting this meaning for the one just mentioned, the definition of the probability calculus becomes more apt. Often the attribution of ‘random’ to an event does not just deny any lawful connection to recognizable conditions – which would exclude the possibility of expectations which are entirely certain. Rather, something is expressed about a degree to which expectation may be warranted. It is in that sense we call an event ‘random’ under vaguely conceived circumstances. To be more expressive, it is an “especially” or “remarkably” random event if a really negligible probability is expected. Then it would be a “random event” if someone

casts a die, and it comes up 6 three times in a row. By contrast, should one result come up significantly more often than usual in many throws, for us it would not be a ‘random’ event if the die’s centre of mass is located outside its geometric centre.

Following this interpretation, one really can call the probability calculus an accounting of randomness, since randomness is graded, and a figure of merit for probability tells us how seriously an event should be considered “random”.

This application of the word ‘random’ then leads us to another change in meaning. This must also be discussed, since we cannot avoid using the word with yet another new meaning. If we describe a future event as contingent on randomness, as a rule the meaning of this claim is the one entertained until now: that its presence or absence may depend on particular determinants or circumstances that lie outside the scope of our understanding. Yet if we describe an event which has already happened as random, as a rule this is less entwined with the interpretation that the same idiosyncrasies of causation could not have been ascertained. Instead it bears the interpretation that the same circumstances are not simply expressed by any coherent and unitary concept. Randomness in that sense is the direct opposite of intentionality – that which is brought about methodically by human will. In this sense, randomness can be attributed to the statement that the sweeping out of equal areas to the sun by the motions of the planets is not random. Outcomes we deal with in so-called games of chance are attributed randomness in this sense, as well as in the sense originally put forward: that there is no unitary and reportable cause for throwing a 6 once in dice. In probability theory the word ‘random’ is used frequently in this sense, too. Among the circumstances on which an event depends, on one hand one distinguishes general determinants. On the other hand one distinguishes particular arrangements that may occur within the frame of general determinants. As should be stressed, initially they are called ‘general’ for the sole reason that they subsume different varieties of behavior. So for example, if “someone tosses a coin”, that counts as a general determinant, because a varied and unsurveyable class of happenings fall under the concept. One is used to saying that for the presence or absence of such an event, what sets the standard is those general conditions on one hand, and their specific ‘random’ arrangements on the other. The notion that these particular arrangements may be beyond the grasp of our discovery is superfluous in all this. Instead, this is a matter of the enormous class of varieties of behavior left more or less undetermined by the general concept. – That which is accessible to our understanding is not always understood to be the epitome of what counts as a general condition. Instead where there is a series of cases of the same kind, one ascribes their general conditions to be all that which as a rule holds for them in equal measure. In this instance it seems natural to take the description of some circumstances to be ‘general’ in this sense. Yet this use of the word should be expressly avoided in what follows. We will be involved in repeated investigations to see if certain general determinants (in the former sense) hold for all cases of a certain kind – to say if they are general in the latter sense. Since it is irrational to use words promiscuously in both senses, we should like to label the latter characteristic – where we consider several cases to be successive, as a rule – as constancy. Certain overall conditions

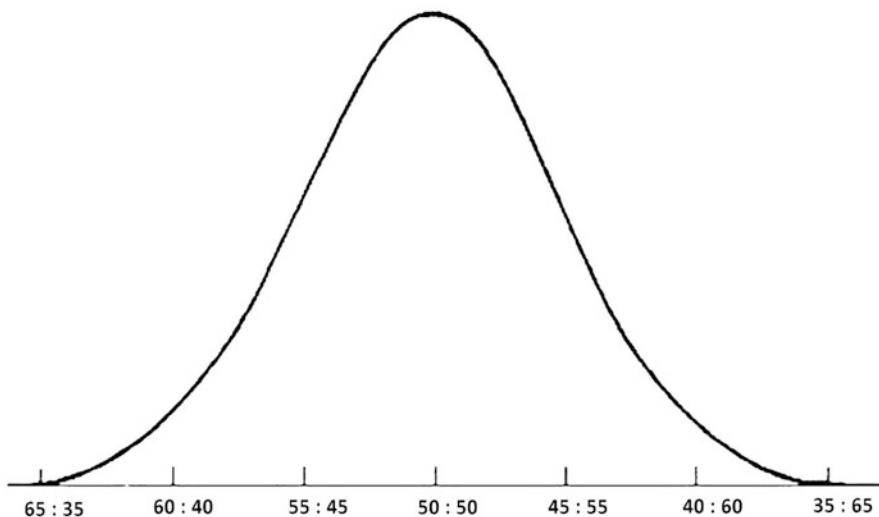
may be constant, or else they may not be constant. As has been stated, in many instances it is customary to take all that which stays constant to constitute general conditions. Yet if (as happens not infrequently) this becomes inapplicable because in some essential respects constancy is not found, then it will be more or less arbitrary, what are to be called general conditions and what is called randomness. That is an issue we will examine more thoroughly at a later opportunity.<sup>6</sup>

**9.** Let us turn to discussion of certain proportions which arise for series of cases of the same kind. Here ‘similarity in kind’ should be understood to mean that all cases are covered by one and the same concept. All offer the possibility of continuing along the same lines. It should also be presumed that certain persistent and constant conditions exist for series that can be continued indefinitely. Those conditions hold for each and every case. Next let us consider an ordinary game of chance, played repeatedly and very often. Say there is an urn that contains an equal number of black and white balls. It is planned that many balls will be drawn from the urn. The balls are chosen with replacement: after each draw the chosen ball is replaced and mixed with the others. It is equally probable that a black ball or a white ball will be drawn on each turn. According to the Law of large numbers, it is to be expected this equality will come to be expressed in the compendium of results over very many draws from the urn. – At this point we would like to consider the game as partitioned into separate series of 100 draws apiece. The most probable result overall for such a tranche is that a white ball will be drawn 50 times. It is somewhat less probable this would occur 49 or 51 times; still less probable that it occurs 48 or 52 times, and so on. These probabilities have entirely settled values, since under the proposed arrangement, all individual draws are considered independent of one another. Then if the game continues at length, so that a large number of such series of 100 draws each are formed, gradually these probabilities emerge. Some series that exhibit certain overall outcomes are most frequent, and other series occur less frequently. While to an ever-increasing approximation the average outcome of all series approaches an equal frequency of black balls and white balls drawn, at the same time a specific ratio of frequencies emerges – as is to be expected with as great a certainty. Series exhibit larger or smaller deviations from the overall average. The relative frequency of larger or smaller deviations from the overall average – which emerges in the results of many individual series – is what one may call the dispersion of the series outcomes.<sup>7</sup> One can represent this most fittingly by a curve. The points of a horizontal line under the curve mark different proportions of white and black balls drawn:  $0/100$ ,  $1/99$ ,  $2/98 \dots 50/50 \dots 98/2$ ,  $99/1$ , and  $100/0$ .

---

<sup>6</sup>Cf. Chap. 6, Section 6 on this topic.

<sup>7</sup>Here I adopt the term used by Wilhelm Lexis (1877, *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*), who has introduced some exceptionally important applications of the concept of dispersion, as will be discussed later. [Lexis, W. (1877). *Zur Theorie der Massenerscheinung in der menschlichen Gesellschaft*. Freiburg i. B.: Fr. Wagner.]



The height of the curve above each point of the line gives the measure of probability, or the frequency with which proportions are expected over very many series. One obtains a curve like that in the accompanying Figure. The curve has its maximum in the centre – representing the results when 50 black and 50 white balls are drawn – and it declines to either side in a particular manner. Let us call this dispersion a normal dispersion. A normal dispersion for an ordinary game of chance that lasts a sufficiently long while is to be expected with maximum probability. That is because of the independence of individual cases. – Now we may easily contrive games also consisting of series of draws, each of which delivers a black ball or a white ball. These games may equally well be continued indefinitely, following fixed rules. Yet these games may depart significantly from the previous type of behavior nonetheless. For example, there could be many urns filled with black and white balls in various proportions. We can imagine that it should be decided by lottery which urn is to be used, but that the arrangement is of a kind that generally a large number of successive draws are taken from each urn. This would be the case if after every draw, a lottery were used to decide if the same urn should be used over again. This could be arranged so that there is a significant probability in favor of a decision of using the same urn. One can see that oftentimes there will be more, and oftentimes fewer draws made from the same urn, though in general a larger number. As soon as it was decided by lot that one was to switch from one urn to another, it would be decided – once again, by lot – which new urn was to be used. So on each turn each urn could have the same chance of being used. Then let us think of this game as played continuously for a very long time. It is to be expected with maximum certainty that all the urns will be picked equally often by lot. It is also to be expected that equally many draws will be taken from each urn on average. Here too, the ratio of draws which deliver black balls to the total number of draws will gradually approach a definite value more and more closely. Let  $n$  be the number of urns, and let  $\alpha_1, \alpha_2, \alpha_3, \dots$  be the ratio of black balls to the sum of black and white balls in the

first, second, third, etc. urn. That definite value is obviously just the arithmetic mean:  $\frac{\alpha_1 + \alpha_2 + \alpha_3 + \dots + \alpha_n}{n}$ . This may be described as the possibility representing the whole series of general determinants, meaning the ratios of fullness among urns, and the rules of play for drawing a black ball. Call these values the span of possibilities encompassed by a course of events, since they depend on all the determinants of the whole series. The span of possibilities should be the same for drawing a black ball and drawing a white ball. Therefore it should be expected that an equality will gradually come to be expressed in equal frequencies of the two outcomes. It is important to state this will not happen in general with the same speed and certainty evident for chances in the type of game we spoke of earlier. Even in a long series it is easy for quite significant deviations to occur from that ratio. If we aggregate each 100 successive draws together in a series once again, large deviations from the average will be much more probable than before for the result of a series. A definite probability may be reported for the values of this deviation; a definite dispersion is to be expected with maximum certainty here too. This is called a hypernormal dispersion. If we present it in graphic form in the same way as normal dispersion, the curve which represents it has a lower peak and higher side lobes than normal. One can see that an expectation of hypernormal dispersion represents assignments of probability for which successive cases should not be considered independent. With the rules of the game in play, it may be expected with great probability that two of the same balls – rather than two different ones – may be drawn in succession. Shortly put, the essential characteristic of the game is that circumstances which are subject to randomness are present, which influence the chance of a greater number of successive draws in the same way. Hypernormal dispersion can be quite different, depending on the rules of the game. Differences in the ratio of black and white balls across separate urns are of great influence on the distribution. – Now let us turn to consider another arrangement of the game, which leads to an expectation of hyponormal dispersion. The expectation for a series which contains a number of individual draws is different: large departures from the average ratio are expected with smaller probability, while small departures are expected with larger probability than for normal dispersion. In graphic representation, the hyponormal dispersion curve has a higher peak and lower side lobes than the normal dispersion curve. Thus it is distinguished from normal in the opposite manner from hypernormal dispersion. One can see that hyponormal dispersion is expected once there is a greater probability two successive draws have unequal results than when they have the same result. This behavior can also be realized by suitably arranging a game, too. For example, let us erect two urns filled in this way: the black balls of one outnumber white in the same ratio as the white outnumber black in the other urn. We proceed to draw balls: each time a white ball is drawn, the next draw is taken from the first urn. After each time a black ball is drawn, the next draw is taken from the second urn. The more we let black balls dominate in one urn and the more we let white balls dominate in the other, evidently the smaller the dispersion to be expected. If any urn held balls of exclusively one color, then a black ball and a white ball would need to be drawn in steady alternation, and the dispersion of even-numbered series would equal zero.

Another procedure under those conditions – by which we may imagine a series of cases of the same kind can be regulated – is sketched here to foreshadow later applications.

While in the institution of this game – by which we have just become acquainted with hypernormal dispersion – it depends on randomness which urn is to be drawn from, we can also imagine that there is a regular order fixed in advance by which a certain number of draws should be taken from a first urn, then a certain number from a second urn, and so forth. Bounded series are formed in that way, where necessarily each corresponds to all the others in the frequency of draws taken from each urn. An easily described span of possibility also arises under these circumstances for drawing a black ball or a white ball. If  $p_1 p_2 p_3 \dots$  are the numbers of draws to be taken from the first, second, and third etc. of the urns, and if  $\alpha_1 \alpha_2 \alpha_3 \dots$  are the chances of drawing a black ball in each, then this span of possibility has the value:

$$\frac{p_1\alpha_1 + p_2\alpha_2 + p_3\alpha_3 + \dots}{p_1 + p_2 + p_3 + \dots}$$

which may be designated by  $\bar{\alpha}$ . Over many series, this gradually becomes evident in the ratio of frequencies with which black balls are drawn. But the results of individual series will exhibit a very different dispersion than if all draws had been taken from one urn: of course this will always be a hyponormal dispersion. One sees this right away, if one considers a span of possibility consisting of only two separate options, one almost 0 and the other almost 1. The series is then assembled from individual cases in a specific ratio in which one process or the other is almost certain. Already under these circumstances, moderate deviations from an average result are very improbable, and expected dispersion is quite negligible. The latter would be equal to zero, if all individual chances were equal to one or equal to zero.

From this it may be understood that for series of phenomena which (1) correspond in the sense that they consist of repetitions of cases of the same kind, and which (2) progress continuously under identical general conditions, and finally which (3) seem to be governed by randomness for the course of individual cases – those series may still behave very differently from one another. The basis of those differences can be described briefly: the content of the general conditions which constantly regulate the processes is essentially different. As a consequence the roles prove to be quite unequal which had been played by randomness in events otherwise (meaning the particular arrangements possible under those conditions). In series for which hypernormal dispersion is to be expected, randomness – one might well say – has greater influence than in series where individual cases are quite independent. Under such conditions as produce hyponormal dispersion, this significance may be deemed still more negligible.<sup>8</sup>

**10.** In the sense just given, series of the same kind can be distinguished by another relation which is quite different but no less basic. Whatever general conditions may

---

<sup>8</sup> The description of definite dispersion, be that of an expected dispersion or be it of an observed dispersion, always presumes a determination in advance of the principles by which individual cases are aggregated into series. It may not be too much to stress this explicitly once again. In the first of our examples, the succession of draws gives us this principle; in the latter example, the rules of the game give us definite bounds for the series. And if one wanted to form series by the random aggregation of single cases, then naturally one would always obtain a normal dispersion.

be, as the series is extended sufficiently they always produce spans of possibility which may be asserted to be ratios of frequencies for one outcome over another. Let us suppose we know this span of possibility for a series of cases of the same kind. Often it is a very important question whether or not this should be considered a measure of the expectation for any single case. Obviously this is so for the ordinary rules of games of chance: for example, if all cases are drawn from the same urn, and the balls which are drawn are always replaced each time to be mixed with the others. But by contrast, for the complex relations which have been considered, in this respect extreme differences may arise.

If it is decided by lot on each turn which urn is to be used, then the expectation for an upcoming draw could depart from the span of possibilities. That would occur if we knew the urn from which the ball is to be drawn, and if we knew the chosen urn contains a particular ratio of black and white balls.

Of course probability will be determined differently if such ratios are not communicated – say, communicated to someone who might only hear the end result each time, whether white or black has been drawn. It is also entirely conceivable that for certain types of observation of this game, non-numeric probabilities are involved. This or that circumstance could give rise to conjectures about the urn to be the site of the next draw. Clearly it is a characteristic of ordinary games of chance – each of whose cases is entirely uniform – that the span of possibility immediately gives the probability relevant to each case. Otherwise this characteristic is contingent: it is not a property that is necessarily realized at all. Let us call this property the definitive meaning of the span of possibility, otherwise the equality of chances for individual cases. Where complex general conditions exist, if idiosyncratic random effects occur in the game, and they make a valid serial contribution, then generally failure of this equality of chance is conditioned. That is because our most recent expectation has been established after part of that decision has already been taken – after its outcome has become known to us. To draw a simple example which speaks to this point, let us imagine both mild and serious cases of an illness have been observed; it is a purely random matter which is observed. Then let there be certain persistent and general conditions which offer concrete possibilities for such and such a case to emerge. Further, let a definite chance of fatal outcome be attributed to each of these possible characters of illness. In that way a well-defined ratio of the span of possibility exists for a case of illness ending in death, and another exists for a case of illness ending in recovery. Awareness of this ratio would give us reason to say something with great certainty before the fact. Over a large number of cases, it allows us to say how large average mortality will prove to be for the illness in question. Yet that is not at all the probability which we have to assume in individual cases for the occurrence of a fatal outcome. Observation tells us, or it allows us to make the conjecture, whether we have a case of this kind or that. It tells us how the first decision (which depended on randomness) has actually turned out. The span of possibility has no definitive meaning here for the expectation of a fatal outcome, then. It should be noted that anytime individual cases are arranged in series by some principle, and those series exhibit non-normal dispersion, equality of chance has already been excluded. It has been excluded because expectation is always developed on the actual course of

previously observed events which belong to the same series. That is how expectation is developed for a new case, and that is how it departs either more or less from the average value.

As we shall see, substantive applications can be made of the concepts expressed in this section. They apply to judgments made about series of phenomena presented to us in the world. The importance of these concepts is shown even more clearly through those applications.

# Chapter 5

## Varieties of Numeric Probability



**Abstract** Every probability must have the definite content of a judgment as its object: a propositional content which is either true or false. The pursuit of a method to transpose or compare probabilities by way of logical argument is largely futile. Probability is not transposed by simple inference: the same probability must be attributed to two premises, only if neither may be established in absence of the other. However, the probability derived from two ranges which are very small and nearly adjacent may be represented in numeric terms by Bayes's rule. Suppose there are two ranges which stand in definite ratio; then say there are gaps in the ranges, so that parts of the ranges remain. The ratio of the residual part-ranges may be shown to follow Bayes's rule. Yet Bayes's rule is applied frequently in inappropriate situations, where probabilities are combined more or less arbitrarily. That can lead to meaningless or incorrect statements of probability.

**Keywords** Types of probability · Combination of probabilities · Bayesian rule · Bayes's formula · Loaded dice · Uniform priors · Individual probability

1. In Chap. 3, our investigation showed us something about numeric probability. It showed that numeric probabilities which we may report for the occurrence of this or that outcome in games of chance, are to be explained by an exceptional circumstance. The exceptional circumstance is that parts of a range which represent different outcomes can always be exchanged in a rule-governed way within a large operational range. They also stand in constant ratio. At first one may be inclined to consider this the only means by which probability can be represented in numbers. As we had intimated already, that is because only there have we found an explanation for the indifference and comparability of ranges spanned by separate premises. And at first we also knew these premises to be the crucial ones for evaluation of probability in numbers. Nevertheless when the subject is pursued further, there exist still other forms of numeric probability. They diverge significantly from the forms we have considered previously. Admittedly in comparison these are not anything fundamentally new, but they can be considered to be specific and intelligible modifications of previous forms. One is led to these special kinds of numeric probability naturally, by consideration of some common and well-grounded

procedures. That is how I will proceed here, since an explanation of these procedures is valuable in itself. – A question may be posed, if the definite probability we ascribe to one premise must also hold for others which are considered its consequences. That is, if a numeric probability estimated for one behavior should not be transposed to others, by way of inference. So one might think that when one premise of a deductive argument is certain, while another premise is more or less in doubt, then it follows that the conclusion of the argument shares the probability of the latter premise. In fact this opinion has frequently been put forward. It has sometimes been proposed explicitly, but sometimes it has been based on tacit transposition of numeric probabilities to premises, when those probabilities are not valid for those premises in their own right. To begin with, it is evident that the circumstance that one proposition shares the probability of another, is a circumstance not adequately described in such a simplistic way. Indeed, the same probability may be ascribed to two propositions  $a$  and  $b$  with necessity, only when they mutually imply one another: if  $a$  implies  $b$ , but also that  $b$  implies  $a$ . In deductive inference, the probability of premise  $P_1$  can only be transposed to a conclusion  $C$  – assuming the other premise  $P_2$  is certain – when the converse is valid. That is, from the validity of  $C$  and  $P_2$  the validity of  $P_1$  is implied in turn. Under these circumstances the assumed propositions  $P_1$  and  $C$  are in fact fully equivalent, given the certain truth of  $P_2$ . The simplest of examples clearly illustrates the inadmissibility of transposition of probability when implication is not mutual, meaning that the converse of the argument fails to hold. The premise  $P_1$  “This stone is a diamond” may be attributed only a very small probability, when the second premise  $P_2$  “Diamond cuts glass” is certain. The conclusion “This stone cuts glass” does not necessarily share the negligible probability of the first premise. It is entirely possible a stone may cut glass without being a diamond. The probabilities of  $P_1$  and  $C$  stand in relation to one another: the one cannot be smaller than the other, though it may be larger. The probability that the stone in question has the property of cutting glass, is at least as large as the probability that it may be a diamond. Consideration of this form of argument leads to a very simple theorem about probability: the same probability must be attributed to two premises, if neither may be established in absence of the other.

**2.** If then we attempt to use this proposition to transpose numeric probabilities, we run into further difficulty. To tie this to another specific example, let us consider that numeric probability  $p$  had been reported in the course of observing some phenomena. Probability  $p$  is the probability that the anticipated result does not depart from a true value to be assessed, by more than a certain amount  $\pm\alpha$ . Observations are conducted, and their outcome is the value  $x$ . Then it seems correct we should say that it can be assumed with probability  $p$  that the true value does not deviate from  $x$  more than  $\pm\alpha$ . It is easy to recognize this transposition of probability as the relation given above. We know with great certainty that value  $x$  has been obtained by observation. Thus both propositions “The error of observation is less than  $\pm\alpha$ ” and “The true value lies between  $x - \alpha$  and  $x + \alpha$ ” actually have the same meaning. The probability  $p$  that holds for the first proposition is also recognized as valid for the second. Yet in this way of looking at things, one important circumstance has been overlooked. We must consider that the probability  $p$  is one we had expected before settling the

observation of an error to be less than  $\pm\alpha$ . It is not at all evident this will have retained its validity unchanged after observations have been conducted – after they have produced a specific value. Rather, and just as well, the premise that the true value lies within the bounds  $x \pm \alpha$ , may itself have had a specific (either large or small) probability before observations were conducted, even if such a probability could not generally have been assessed in numbers. Then after outcome  $x$  has been obtained, the final probabilities of the two premises are linked. Both probabilities must be considered: that previously attributed to one proposition, and the other probability on its own. The joint probability attributed to the two must be produced from both these elements somehow. Differently than we presumed at first, we can think of this as being realized in that before observations were conducted, some relevant magnitudes would have been known, maybe as the result of otherwise more precise and more stable methods. Then our probability changes, because the value  $x$  which has been found is not the usual probability of its having this or that value. Instead it is merely a probability for the magnitude of the observational error in question. How large that was has been established directly. But if there is no definite knowledge, when only loose conjectures about the true values of the quantities have been made, then of course wouldn't the ensuing probability be modified? In the context of the said connections of two premises, it is completely inadmissible simply to transpose the probability which is thought to hold for one alone, to the other – before their connection becomes known. In such cases there needs to be an entirely fresh assessment of probability. That assessment must be carried out at least for a combination of the two probabilities, which combination had been thought at first to be of two independent premises. This pursuit of an attempt to transpose probabilities by way of logical argument, then brings us to the insight that this cannot be done at all. At the same time this is an occasion for us to explore particular assumptions which must be made about probability, once two premises of different content are connected to one another by necessity.<sup>1</sup>

We may represent the connection formed between two premises  $a$  and  $b$  as a constraint on the combinations of two possibilities for different objects. Propositions  $a$  and  $\sim a$  [not  $a$ ] might seem to be permissible in one respect, and  $b$  and  $\sim b$  in another, while really only  $a$  and  $b$  are allowed in combination, as only  $\sim a$  and  $\sim b$  in another. Certain combinations of all available possibilities are permissible, while others count as excluded. We will want to make use of this idea, to discover the

---

<sup>1</sup> The result just derived – that there is no transposition of probability to be made – needs a certain qualification. In the whole exposition just set out, we proceeded on the assumption that a definite probability could be attributed to either of the two premises separately. Clearly this is not the case if for example, one of the premises cannot be advanced as either probable or improbable at all, independently of the other. So all our conjectures about future events are exclusively constrained by what we can say about present or past behavior in the world. Such is the nature of things that there cannot be a probability with which this or that future event can be expected, independently of the probability of those present varieties of behavior with which we are forced to think those future events must be connected. Admittedly in such a case a transposition of probability is made, if one really wants to call it that.

circumstances under which even such combined probabilities may be represented in numbers.

**3.** However we should also indicate in advance that analysis of another procedure which plays an outsized role in the probability calculus, leads to the very same situation in which possibilities cannot be combined in unrestricted fashion. Let us begin with a specific example. Say six dice correspond completely in geometric and in physical terms. They are distinguished only in the way their sides are marked. That is to say, they can be marked on one or two or three or four or five or six sides by a cross, +. All remaining sides are always marked with a zero, 0. We shake the dice thoroughly in a dicecup, and draw one of them out. Without inspecting the die, we cast it several times. We always observe just the symbol inscribed on the uppermost side. If we continue this way for some time, of course we may glean with some certainty which of the six dice we have. We glean this from the number of times + is thrown and the number of times 0 is thrown. Let us assume the die is cast only three times. The series has come up as +, 0, and then +. We ask if a numeric probability (and perhaps even which numeric probability) may be given that we have the first die, or the second, or another one. What one sees of methods by which this question and similar ones are treated in textbooks – methods which lead to entirely correct outcomes – runs as follows. Initially it is equally possible any die may be drawn. There are initially six equally possible premises. After a few throws have been observed, the probability attributed to one premise must be assigned to be proportional to that probability which produces the state of affairs assumed for the outcome. This is the so-called Bayesian rule. The Bayesian rule has justifiably been considered one of the most important propositions in probability theory. The outcome under the premise that the first die would have been drawn – the outcome observed in our example – would have probability:

$$\frac{1}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{216};$$

and in order for the other dice, the values would be:

$$\frac{2}{6} \cdot \frac{4}{6} \cdot \frac{2}{6} = \frac{16}{216} \text{ for the second die,}$$

$$\frac{3}{6} \cdot \frac{3}{6} \cdot \frac{3}{6} = \frac{27}{216} \text{ for the third,}$$

$$\frac{4}{6} \cdot \frac{2}{6} \cdot \frac{4}{6} = \frac{32}{216} \text{ for the fourth,}$$

$$\frac{5}{6} \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{216} \text{ for the fifth, and}$$

$$\frac{6}{6} \cdot 0 \cdot \frac{6}{6} = 0 \text{ for the sixth die.}$$

It follows that the probability of having the first, or the second, etc. of the dice, should be proportional to the numbers 5, 16, 27, 32, 25, and 0. Consequently the probabilities are expressed by the fractions  $\frac{5}{105}$ ,  $\frac{16}{105}$ ,  $\frac{27}{105}$ , etc. Now let us clarify what kinds of considerations we might use as a foundation for these assignments of probability.

Initially, a certain operational range needs to be assumed for mixing the dice and for drawing one of the dice. One says the occurrence of any die is set equal in probability to that of any other. This holds just as surely when the draw has been made as when it about to be made – but only so long as we are certain the real result of the draw has in no way become salient in the larger context of events which have come to our attention. Of course considerations of probability never relate essentially and directly to the future. They always depend on present or past behavior. Under that assumption, it is moot whether a die has already been drawn. – Motions of the die give rise to a certain operational range. By its measure we consider six possible end states to be equally probable for each of the dice, on each of three throws. A long series of possibilities arises from the combination of operational ranges. After three throws have been carried out, and the results have been observed, observation tells us that a certain class of results has not been realized: namely all of those not leading to the series + 0 +. Observation forces us, so to speak, into choosing among cases which were originally posited as equally likely. We eliminate those which contradict experience, and we assess probability by the ratio of those consonant with experience. Here too, possibilities cannot be combined unrestrictedly between possibilities for mixing and possibilities for drawing, and possibilities for movements of the dice. Rather some premises are derived from what has been observed in one respect, and other premises are derived from what has been observed in another.

**4.** Then let us direct our attention to ask how a numeric representation of probability emerges for combined premises under these circumstances. This can be described in simple terms, at least for the example. If we consider the cases initially set to be equal, still to be equal after observation of three throws has been conducted, and we eliminate only those which have not been realized, it is easy to see that the same ratio results as given by the rule previously introduced.

In the example, drawing one die among six and three throws with that die, provided us with  $6 \cdot 216 = 1296$  cases of equal value. Only 105 are consonant with that which was observed, which excludes the other 1191 cases. Acts of drawing single dice, represented by numbers cited earlier of 5, 16, 27, and so forth are contained among the remaining 105. This is the ordinary derivation of the rule to ascertain the probability of “causes of observed events”, according to **Bayes**. The mathematical form of the rule, which is not to be reviewed here, is that if an actual event can be produced by several different circumstances, the probability of each is given by the fraction

$$\frac{p_1\alpha_1}{\sum p\alpha};$$

Here  $p$  indicates the probability of each circumstance by which this might be supposed to occur without regard to the outcome actually observed;  $\alpha$  indicates the probability with which the event is to be expected given the circumstance;  $p_1$  and  $\alpha_1$  represent values for any one of the circumstances, and  $\Sigma$  indicates the sum of products  $p\alpha$  for all the circumstances which enter into consideration as causes.

We may provide a better foundation for the assumption about probability expressed here. Of course this is a direct reference to operational ranges, as represented after the event has been observed. We may not overlook the fact that the derivation of Bayes's Rule, as we have just set it out and as it is usually given, is not exactly irreproachable. That is: we set out a number of equal-valued premises; by experience a portion of them prove to be incorrect; we dispose of those, and assess probability by enumeration of those which remain. Should they still be considered equal in value? Perhaps this newly-obtained knowledge might upset the equilibrium that had existed before? Those questions are easily resolved, once we consider what a range actually looks like, whose parts are compared in terms of assignments of probability based on Bayes's Rule. Then we can see that the whole range – formed by the exhaustive combination of all types of behavior – initially contains such parts which are thoroughly mixed in constant ratio. There are parts that represent actually observed results, and other parts that do not – in other words, such parts as we consider, and others we must ignore. For short, let us call all the former type the ‘admissible’ portion of the entire range. One might say this represents a fragmentary range, beset with gaps. In the example, all the admissible parts form  $\frac{105}{296}$  of the whole. These represent acts of drawing each of the dice; they stand in fixed ratio to one another. All the properties on which numeric representation of probability depends (the indifference, comparability, and originality of ranges) are realized in the same way as for the usual case of a range without gaps. What sets this apart is the restriction on possibility imposed by experience. One could say it decimates the collective range. It is also easy to understand how the numeric values – which one commonly takes as probabilities which would have arisen under certain postulates – in fact have another meaning. Their meaning can be expressed simply, even for a specific state of knowledge. For instance, take the probability expected under the postulate that die #3 has been drawn. That the actual result + 0 + had been expected means nothing but that a portion of a certain range counts as ‘admissible’, which range represents the draw in question. Suppose there are two ranges which stand in the ratio  $\frac{p_1}{p_2}$ . If the residual whose measure is the fraction  $\alpha_1$  is a part of the first, and if the residual whose measure is the fraction  $\alpha_2$  is a part of the second, then evidently the ratio of these parts is  $\frac{\alpha_1 p_1}{\alpha_2 p_2}$ . That is simply the content of Bayes's formula.

One should not call the coefficients  $\alpha$  ‘probabilities’ either, strictly speaking, since they are not considered as such. Rather, they are possibilities which represent all circumstances in question for the results actually observed. The form of the range with gaps – that is, the range admissible only in part – represents the first turning-point in our understanding of numeric probability which was to have been indicated. This subsumes Bayes's rule, of a “probability of the causes of observed events” (‘cause’ is better expressed as the manner of their coming into being), but it includes

the rule only given an essential premise: that all values entering into the calculation are correct, and that all are numeric probabilities or possibilities based on the Range principle.

5. Actually, Bayes's Rule is used far more frequently in cases where the said premise does not hold. Instead, a number of probabilities are combined more or less arbitrarily. Indeed they are often found to be decidedly incorrect. If results are to be obtained here in the end – results which are to be claimed essentially correct – that depends on the probabilities in play being representable in numbers (again, in one of the ways that have been mentioned). Let us connect this to the earlier example. Imagine that in the course of making some observations, a numeric probability is reported that an error of a certain size is associated with a result. Clearly the probability is expressed as a function of the error: if we let  $z$  be the size of the error, then the probability that the error has a value between  $z$  and  $z + dz$  is  $=\varphi_{(z)} dz$ , where  $\varphi$  is some function. About the form of this function, we need only know it has a maximum value at  $z = 0$ , that it decreases with increasing  $z$ , and that it lies close to zero for certain values of  $z$ . Generally small errors are the most probable, while it seems almost ruled out that the error exceeds certain large values. When the observation is conducted, and a definite result  $x_0$  has been delivered, Bayes's rule provides a ready procedure by which a numeric probability seems to represent that the true value lies within certain bounds. That is at least more rigorous than the simple ‘transposition’ which had been shown to be invalid previously. To apply Bayes's rule to cases of this kind, one must be able to establish, independently of observation, the probability that a given value of estimate would be expected to have. In its usual application, let us call this stipulation the “*a priori*” assignment of probability. One is used to making the stipulation this way: one assumes that before observation has been conducted, every relevant value may be considered equally probable. Or better, the probability that the value lies within an arbitrary range  $\Delta x$  is set to be proportional to  $\Delta x$ . Initially a numeric ratio is produced for the probability of two premises following Bayes's rule, on the basis of this premise. Suppose that one premise is that the true value lies between  $x_1$  and  $x_1 + dx$ , and the other premise is that it is bounded by  $x_2$  and  $x_2 + dx$ . Then the probability of both premises behaves just as those probabilities expected on the basis of the outcome that actually occurs: namely the observed event  $x_0$ . That is, they stand in the relation of  $\varphi_{(x_1 - x_0)}$  to  $\varphi_{(x_2 - x_0)}$ . From this the form of Bayes's rule can be derived easily. It is the probability that the true value lies between bounds  $a$  and  $b$ , as given by the fraction:

$$\frac{\int_a^b \varphi_{(x - x_0)} dx}{\int_{-\infty}^{+\infty} \varphi_{(x - x_0)} dx}$$

This formula may be thought to follow from the formula  $\frac{p_1 \alpha_1}{\sum p \alpha}$  given above, as summed over a number of elements, which number becomes infinitely large. – Now there can be no reason to doubt that among the premises underlying this calculation, one may not be considered strictly correct: namely that included in the formulation of the *a priori* assignment of probability. First, it hardly ever happens that determinations of magnitude are made without some supposition having been made about their true value. Even if those are still very uncertain, they have to modify the *a priori* assessment of probability to some extent. However, if we abstract from such conjectures altogether, the *a priori* assignment of probability represents the case already described – that a substantial range exists for the unrestricted formation of expectation, though different parts of that range may not be comparable to one another, which means the measurement of probability becomes arbitrary. Consequently, once we bring to light what that *a priori* assignment of probability means for a specific example, it cannot be ignored that such a thing cannot be sustained. For example, observations may involve specific weight. Before relevant observations have been conducted, the claim that it is equally probable a specific weight may lie between 0.5 and 0.6, and that it may lie between 100.1 and 100.2, proves to be inadmissible. It is inadmissible with respect to both points just indicated. The characteristic mark of this case and similar ones rests on the particular form of the function  $\varphi$ . According to that, values which depart significantly from the observational result  $x_0$  are so exceptionally improbable, that one may consider them to have been wholly ruled out. And with that, whatever disparate forms we may lend to the *a priori* assignment of probability, all of them deliver the same result (to a very close approximation). Observation may produce a specific gravity of 3.45. If that is exact to some extent, then premises which attribute values under 3.44 or over 3.46 are no longer to be considered, because of the enormous improbability of such large errors in observation. The more exact form of the *a priori* assignment of probability becomes quite irrelevant. If numeric probabilities obtained in this way must be deemed correct after all, we can recognize the reason they are. The reason is that comparisons of *a priori* probability are not at all necessary for any range of values in general, but only for parts of tiny ranges which include the result of observation. Small parts are comparable without arbitrariness worthy of mention, within such narrow bounds. But there would be no objection to the validity of this procedure, insofar as the conjectures which have been made about prospective values are constrained to lie within a small range within which there are no significant differences – that is, should current observations far exceed previous knowledge or conjecture about the object, in certainty and precision. – If we try to make these cases comprehensible in terms of operational ranges, in the main they do not differ from the ordinary account of numeric probability. Instead they represent a new and interesting modification of it.

At first it is easy to understand that the probability of two equal ranges – ranges which are very small and nearly adjacent – may be represented in numeric terms by Bayes's Rule. The *a priori* probability of these ranges would be unreflectively established as equal. After observations had been conducted, their probability

would be assessed by familiar calculations to show how large parts of these equal ranges are to be considered “admissible” on the basis of experience. Ordinarily a restriction of comparisons to small adjacent ranges leads only to determinations of probability which are of less significance, namely relative determinations of probability. Even if we should claim that a certain ratio exists between the probability that a value lies between 5.40 and 5.41, and the probability that the value lies between 5.41 and 5.42, then that does not do too much for us, if we cannot compare these probabilities with others for which entirely different magnitudes are broached. Though now we have an account in numeric terms of the ratio of probabilities with two premises, we have not made all possibilities comparable that way. What is really and solely valuable in the end, can only be attained in a couple of situations where comparability is confined to the smallest domains. The first of these is the ordinary rule with which we began, where parts to be compared completely fill the entire range, by regular combination and in constant proportion. But the other case is the one with which we are now concerned. Initially that case rests on the fact that judgment of probability is merely a business of comparing “admissible parts” of a whole range. Then the ratio of admissible portions to excluded portions is not equal across the entire range at hand. It only represents a fraction of appreciable size, at some definite point and in its proximate neighbourhood. That ratio lessens quickly with distance from that point, and outside certain narrow bounds it approaches zero. Once more all possibilities are given exhaustively by premises which have comparable values of probability. Other premises do come into play here, premises with values which are not comparable strictly speaking. Yet their probability is so tiny that it ought to be ignored. While the exhaustion of possibilities is usually represented by saying that the portions to be compared fill out the entire range in constant proportion, that happens here by rapid diminution of the proportion of admissibility instead.

If we survey the types of numeric probability that have been considered, we notice immediately that the comparability and the indifference of very small and neighbouring elements of ranges always form the basis of these probabilities. It is evident that the two types of probability which have just been contrasted are the only ones in which an exhaustion of all possibilities may take place with premises that have comparable values of probability. If we abstract away from the determination of probabilities which are relative, and which deal only with comparison of a restricted number of premises, we should be able to claim that we have characterized the proportions which fully enable statements about numeric probabilities, in a way that is at once general and complete.

**6.** If we compare the types of probability with which we are now familiar, then we may have the facile impression they might be distinguished in other ways. We assign the probability that a total of 18 will be thrown with three dice, to be equal to  $\frac{1}{216}$ . Is it not something basically different, if we say that it can be assumed with such-and-such a probability, that some event has actually occurred, or that some real magnitude has a certain value?

Difficulties seem to arise once again about a question we thought we had done away with earlier. Does the first-named of those assignments of probability not have objective meaning, while the latter has only subjective meaning? Closer examination shows that differences in the comparison of these distinct propositions about probability – which one thinks we perceive so directly – are differences which are actually present in part, and which may be understood and exactly gauged through our earlier investigation. Yet in part these differences are illusory, in that they emerge from certain confusions which are not at all easily alleviated.

To begin with this last point, if we pay attention we see that the outcome of a throw in progress is fully and precisely determined by existing proportions. As evident as that may seem if we think of it in isolation, yet it is easily overlooked by ordinary superficial intuition. Surely this touches on the issue that we are inclined to distinguish the probability of a prospective event from the probability of things which have happened already, or which have continued for some time. It seems natural to contrast this – the thing as really still undecided – with that – the thing as objectively established already. Let us be clear that this idea is most decidedly in error, and let us be clear that even the question whether a 5 or a 6 will presently be cast is merely an expression of our subjective uncertainty in the face of behavior that is already fixed. Then another point remains, by which there seems to arise a distinction between one kind of probability and another. Recall that at an earlier point we had opportunity to remark that the importance attached to probability statements in the realm of games of chance, is principally tied to the general validity of those propositions. We count the probability of throwing a 6 with a single die as equal to  $\frac{1}{6}$ . With that we have expressed something valid for everyone, under any conditions. By contrast, if we report the probability of an event which has occurred already, then generally that is not valid in the same sense, since it does not refer to a whole category of cases, only to a single one. One might say it is a singular case. Further, it dispenses with generality in the sense that attainment of a bit of knowledge which generates very different probabilities seems in no way excluded. Once an event has occurred, we tell ourselves it is entirely possible someone might have observed it; that at any moment we can perceive this or that mark of its progress; and that the probability with which we assume something might have happened, always seems to be probability only in a restricted and individually valid sense. Naturally something similar holds for the probability of behavior that has continued for some time, and that has been observable all along. The basic distinction to which we are led, is one mentioned earlier, between probability which is generally valid on the one hand, and that which is singular and individual on the other. Obviously this makes no difference to the logical character of the relevant propositions. Rather it is a completely accidental feature. Whether the mental state they represent is manifest in one person or many; whether it is found frequently and repeatedly in just the same way or not: evidently all that is quite irrelevant. Here we are only concerned to establish what this mental state may be. I would like to lend special emphasis to this point, since many have endorsed the opinion that any numeric probability has to be entirely general, and that numeric probability bears no relation at all to single,

isolated events. Against that opinion, I have to claim that every probability is singular, strictly speaking. All make reference to single and determinate behaviors.

Indeed every probability must have the definite content of a judgment as its object: a propositional content which is either true or false. By necessity, every probability is a probability that this or that behaves in such-and-such a way. I can assume with a certain probability that Cajus or Sempronius or Titus will die over the course of a year. If we speak of the probability of a general case instead, which covers an indefinite number of individual varieties of behavior, say the probability “that a man who is forty years old today will live another twenty years”, then it is clear this word must be taken in some habitual sense, and we are dealing with some abbreviated expression. In fact, such a proposition – if it should have anything to do with probability at all – only has the meaning that it places some general rubric on an arbitrary number of individual probabilities. As we had shown earlier, general probability statements are basically propositions with objective content, by whose use our expectations are determined over a large or a small number of cases – possibly over an unbounded number. In conjunction with specific knowledge about an individual case, these propositions produce numeric probabilities for each case. For example, if I know that “the probability of throwing a 6 with one die” is equal to  $\frac{1}{6}$ , then no numerically determinate expectation attends us. Only when I know that someone casts the die, do I expect the outcome of 6 with a certain probability on the basis of that general proposition.

Once this point has been clarified, I think this leads to fuller insight: that there is no fundamental distinction between the real numeric probabilities which pertain to the two cases compared previously. The probability – in the game of roulette which I observe – that the ball will land on black in the next ten spins at least once, and the probability that the true value of a real magnitude lies within certain bounds: both are my conjectures about actual behavior. They are either objectively true or objectively false, and they can be represented in numbers for exactly the same reason, following exactly the same principle.

# Chapter 6

## Establishing and Justifying Probability Statements



**Abstract** There is not the least justification for indiscriminate use of the word ‘probability’ for any statistically estimated frequency of behavior. The task of a general theory of probability can only consist of making important methods known; the establishment of generally applicable methods is nothing to strive for. We should take into account not only frequencies, but also the groupings which can be made of cases. One may try to tie a series of phenomena to the behavior of a game of chance, but so often this leads to unsatisfactory results – especially if circumstances of observation change over time. If observation of a sufficiently large number of series produces a non-normal dispersion, we should conclude that the phenomena do not behave on analogy to games of chance. That is to say, some questions about the enduring or constant circumstances of phenomena can be answered empirically.

**Keywords** Range principle · Imprecise knowledge · Laplacian probability · Non-normal dispersion · Mortality rates · Collective behavior · Constancy of conditions · *a posteriori* probability

1. According to the Range principle, some knowledge of objective import seems to be necessary to every application of the probability calculus. Then the investigation which tells us how such items of knowledge about different objects are acquired – which items are the basis for familiarity with statements about probability – is a particularly important part of the whole theory of probability. In this very respect the results of intuition depart markedly from what is contained in textbooks and other systematic presentations of the probability calculus. A review and a revision of some representative modern expositions on application of the probability calculus seems to be called for, all the more because – as we shall see – current treatments are largely insufficient. Often they have to be termed decidedly misleading.

Even in cases commonly trotted out as typical examples of the “assessment of probability”, we can reveal a basic difference in concept. Say we draw balls many times in succession from an urn filled with black and white balls. Then say we replace each one every time and mix it with the others. If over a large number of draws we obtain a black ball  $n$  times and a white ball  $m$  times, then a conclusion seems justified: that the ratio of black and white balls in the urn does not vary

substantially from  $\frac{n}{m}$ . Then for a subsequent draw we assume that the probability of obtaining a black ball ought to be very close to  $\frac{n}{n+m}$ , and the probability of obtaining a white ball is very close to  $\frac{m}{n+m}$ . This procedure is known as the a posteriori determination of probability. Let us pose ourselves the question, how generally this should be considered to be applicable. We have to pay close attention, because in the example given above, an unquestioned warrant drags in certain bits of knowledge, which conventional expositions clearly fail to make explicit. Already we know with certainty that when balls are drawn from the urn, the probability of obtaining a white ball or a black ball is solely determined by the ratio in which the balls are present. We also know each draw can be assigned a definite chance independently of the outcome of other draws, and we know that the number of balls does not change between draws. In a word, we know the entire phenomenon to be judged represents a game of chance, and constant chances are set out for a sequence of events that are separate and independent in a game of chance. What we determine empirically is just the quantitative value of that chance. It follows that this procedure is in no way applicable in general. Rather it is tied to some essential premises. Following the usual prescription, suppose that for one series of similar cases which may run partly one way and partly another, we ask purely and simply what “the probability of a certain course of events in such a case” may be. Initially (by a point which was raised earlier) it may be questioned if there is any value to which such meaning can be attributed. We are acquainted with the circumstances on which this depends. So we may say with certainty that anyhow this will not be the case generally. The very task before us cannot just consist of describing a method for the numeric assessment of any arbitrary probability whatsoever. In keeping with some important and objective postulates expressed as statements about probability, we need to provide some basis for these probabilities. That covers much more and far different things than just the determination of numeric value. We might convince ourselves that such a basis is to be sought largely through concepts of a very general nature. Those concepts might draw on the ways the relevant phenomena come about, and the behaviors which occasion them. But there are simply no completely general rules which can be established for the acquisition of such concepts, or their demonstration. It can seem an open question at first, whether observation is to be made of a large number of cases, or if the relative frequency of this trend or that ought to be turned to use. Closer study shows us how this can be done. In most cases, and for nearly all propositions of real significance, the objective basis of any proposition about probability is tractable: it is capable of being examined and tested in various ways. That may be done without establishing which proof is to be considered definitive, or what weight should be given to one proof or another. Under these circumstances, the task of a general theory can only consist in making the most important methods known. They are the methods by which propositions can be proven, or by which their validity may be proven. By contrast, no decision may be taken about when or under what auspices each of these methods finds a home. The establishment of a generally applicable method is nothing we ought to strive for at all in these investigations.

2. The first case – the simplest theoretically – would be that in which we are in a position to tell what is known in a basic and completely precise way, which means what is to be considered certain about the real behavior of some objects, but whose range makes allowance for our inexact knowledge. We would also have to be in a position to know the laws by which the course of events unfolds in the domain in question. Finally, we would have to be in a position to know how large the original operational ranges are, the ranges which represent this or that course of events.

This method of grounding probability statements can be called a deductive method. We should call it a detailed deductive method, to distinguish it from a method soon to be discussed. One can see this requires extensive and exact knowledge of the domain in question. That knowledge must have been garnered elsewhere somehow, and it has to count as wholly certain. A complete and detailed reduction of numeric probability to the Range principle would then be possible. We could then elaborate all varieties of behavior considered equally probable, describing them exactly; we could fully specify the ratios on which their equal value is based. Yet I wouldn't know how to specify such a domain, in which numeric probabilities could be represented and grounded this way. It is conceivable such a domain does not exist: we never possess sure items of knowledge about laws of events to serve as a foundation. At best one might believe that certain areas of physics – say, the kinetic theory of gases – permit us the kind of perspective on probability which has just been sketched. Still closer examination shows that this is not strictly the case either.

Without foreshadowing comprehensive discussion of this interesting application of the probability calculus, its most important characteristic can be introduced here briefly. In the kinetic theory of gases, we consider a volume of gas to be a large number of separate molecules engaged in brisk movements hither and yon. They bounce off one another, or off the solid or fluid bodies which contain them. We can now ground the considerations which are to be established here in a very direct way. We say that it is only the collective effect of very many molecules which we observe. Their particular arrangement escapes our knowledge entirely. We may form certain postulates about causal laws – for example, we assume that molecules bounce off one another in the same way elastic objects behave, or else we assume that a repulsive force acts between them that falls off especially quickly with their separation from one another – With that much we can show that if a defined unit of gas is left to itself, enclosed in a space for a long time, certain characteristics are to be expected with maximum probability. That is how the kinetic theory of gases seeks to derive the equilibrium of pressure and temperature established in a gas left to itself, isolated from external influences. Similarly the transport of heat from a region of higher gas temperature to a colder region is derived. Considerations of probability have a marked transparency which is enabled by a process which allows a simply described, very perspicuous comparison of magnitude. Undoubtedly here we have a correspondence in foundational principles with the principles we called detailed and deductive. This correspondence is still only superficial; the logical situation is fundamentally different. That is because we do not have anything like an independent knowledge of gas molecules and their causal laws, which can be considered certain. The whole theory should be ascribed only a certain degree of intrinsic

probability, insofar as some simple and maybe rational assumptions are made about the properties of molecules. The whole theory is not to be elaborated any more closely. Yet a part of the justification which is just as important – or which may be even more important – is this: that all the phenomena just alluded to have actually been observed, such as the attainment of an equilibrium of temperature and of pressure (and so on). There is knowledge of those very outcomes on which considerations of probability may be established.

From this it may be seen that even in the kinetic theory of gases there is a conglomerate foundation for assignments of probability, one which is difficult to characterize. Nevertheless this approach is not superficial. At least it tells us how we might think about the ideal case of a theory with a detailed deductive basis, and that in an especially clear way.

**3.** Games of chance behave in an essentially different manner in the justification of probability statements. There is no question whatsoever of an exact, complete, and certain knowledge of causal laws for the phenomena. There is no attempt even in a hypothetical sense to establish such laws.

Nonetheless a basis for those probability statements can be found by a deductive method, in connection with some notions about the general conditions of any game of chance. They are somewhat specific to the case in question. We pursued similar considerations in Chap. 3: there we learned that even without a specific and detailed knowledge of things and their causal laws, certain findings may still be established as probable. Their probability may be established with the highest degree of certainty by the relative sizes of original ranges of behavior.

We may call such an inference summarily deductive. Its characteristic is this: it is connected with certain and precise knowledge in some respects, but it proceeds only from extremely general assumptions in others. One assumption could be the constancy of a certain correlation. If as before we make a short list of the premises on which we have relied, then we may consider an objective numeric ratio by which we are given to measure probability as known with certainty (in the usual sense of the phrase). Those are such premises as the number of black and white balls contained in an urn, or the regular form of a die or a coin to be tossed, or the equal widths of black and red pockets on a roulette wheel, etc.

If the assessment of probability occurs *a posteriori* in this way, it will simply be presumed certain the relevant ratio remains unchanged over all cases of a lengthy series.

It can count as certain that properties which distinguish different outcomes which are compared in probability do not directly involve a likelihood of one over another. The black color or the white color of a ball has no effect on its being chosen. Marks on the sides of dice have no effect on the way they land. If for example, a black ball was chosen under specific circumstances, there is no doubt that in the event everything else behaves the same way, and that a white ball would have been found in the same place, then the white ball would have been drawn. Certain probabilities pertaining to the results of the game are connected to probabilities relating to the implementation of arbitrary motions during the game. It counts as certain that every case is influenced by a large number of circumstances, and there

are always new circumstances to determine fresh cases. Contrary to the remarks by which we sought to render assignments of probability comprehensible in games of chance, these observations are of a characteristic sort. They convert the whole inference to a summarily deductive inference. From the way in which relevant processes arise, it may be shown almost convincingly that for certain processes, unchanging assignments of probability can be set – that small neighbourhoods of values over a range represent nearly equal original ranges. This deduction of probability statements is not at all limited to games of chance. We can very easily imagine other questions for which numeric probabilities may be given that are warranted on very similar grounds. This is particularly important because here too, exceptionally safe expectations can be formed for the totality of outcomes over very many cases. This is important because the reliability we ascribe to them allows us to see the logical value of summarily deductive inference. So if we apply considerations about probability to the height of an indicator on a barometer, then we may develop the odd conviction – just as we derived probability statements for games of chance – that at any point in time, values are equally probable which contain the numeral 4 or the numeral 5 in their second decimal place. We should be able to demonstrate that it ought to be expected with maximal certainty that over a long period, that values of the one kind and values of the other occur for equal periods of time overall. Perhaps it might seem justified to regard this as almost certain before the fact. One might even consider it a pointless and superfluous enterprise to furnish an empirical proof.

Nevertheless it is self-evident – as should be emphasized once again – that summary deduction never bestows absolute certainty on probability statements. After all, it remains conceivable (even if extremely improbable) that a certain type of movement could predominate in the motions of a die, for instance. That – given that a certain side has landed face-up – makes one specific result more probable on the next turn than the others, and so on.

One should also consider it important under these circumstances that empirical observation of outcomes has always led to the establishment of probabilities which correspond closely to those found by deduction. For that reason, one should speak of a direct empirical justification for such statements of probability. Such conditions have never been acknowledged until now for the throw of a regular die which would have accorded some additional likelihood to observing a specific throw, or a specific series of throws. The other requisites (to be discussed presently) for an empirical justification of probability statements have always been sufficient for games of chance, insofar as one may have attempted to define such things.<sup>1</sup> What stock one may put in the deductive method, and what stock one may put in empirical justification – that balance cannot be described in a way that is generally valid. To me it seems unnecessary to make pronouncements on the matter. There can be no

---

<sup>1</sup> Compare, for example, the experiments on coin-tosses reported by **Jevons** (*Principles of Science*, p. 208). [Jevons, W.S. (1874). *The principles of science: A treatise on logic and scientific method*. London: Macmillan and Co.]

doubt that probability statements which have a bearing on games of chance and similar subjects ought to count as valid at any rate, with maximum certainty and to a best approximation.

4. In contrast to the deductive inferences which have been considered, we would like to see how probability statements are formed in a more narrowly empirical way, that is, how the observation of relevant cases on its own may yield propositions about probability.<sup>2</sup> The first condition that one can hope to have a numeric probability at all, is obviously that a number of cases exhibit a certain similarity. Some take one course while others take another course; for these the result is this outcome while for those the result is that outcome. Already there a certain connection to the succession of individual cases in a game of chance. This correspondence is even fuller if our knowledge of each case is incomplete, so that the outcome seems uncertain every time as in a game of chance. So say we observe a large number of cases of typhus over a long period of time. Some of the sick die, and another group recovers to becomes healthy. It is not to be said with certainty before the fact in each case, whether its state will eventually be one or the other. It seems at least conceivable under such circumstances that the series of phenomena behaves like an ordinary game of chance – that our knowledge of individual cases, the knowledge by which we construe this to be a case of typhus, leaves open a range of behavior. Within these comparable, indifferent, and original ranges are represented both fatalities and recoveries in a definite proportion. For each case of typhus there could arise a certain chance of one or the other, and it may seem a random matter whether this of that course of the disease occurs.

Now it may be asked: how can we gain any conviction that this is how things really behave ? First I would like to say there may be no domain where a test of this assumption is to be found entirely and exclusively in the empirical method considered here. One must always try to connect this with what is known or what may be conjectured about the origins of relevant phenomena, as well as about their mutual dependence. In that way – with the aid of some deduction – very likely one may understand the basis of the probability statements which were to be established. The better and the more completely this succeeds, and the more that the premises (which the probability statements might involve) can plausibly be made in other contexts too, then the whole idea will seem to be so much better justified. But all this should be ignored for the moment: what we consider here is just the observation of cases themselves and their progress.

The first, most essential question is evidently: if there are any general conditions which really do remain constant for the proportions in question, knowledge of which would justify the assignment of particular chances. We had alluded to the certainty of similar postulates in games of chance: that the die cast a large number of times always bears the same marks on its sides; that the urn contains the same number of

---

<sup>2</sup>On this point, see Lexis, *Zur Theorie der Massenerscheinung in der menschlichen Gesellschaft*. [Lexis, W. (1877). *Zur Theorie der Massenerscheinung in der menschlichen Gesellschaft*. Freiburg i. B.: Fr. Wagner.]

black balls and white balls before each draw, and so on – there is no doubt about those things. Given an entirely arbitrary series of phenomena of like kind, analogous behavior should not be assumed in advance as certain. To put the matter shortly, that needs to be put to the test, to show if the series behaves at all according to the same rules as does an indefinitely prolonged game of chance.

Most often this effort proves successful, though it can be reported only roughly and to a degree (if not quite precisely) how the general conditions are to be represented, by which we consider those chances to be determined. When we speak of the probability of an observational error of such-and-such a size, general determining factors will be understood to include an observer's attention, limits on the capacities of his sense-organs, the condition of his instruments, and so on. In what concerns the probability of a lethal course of a particular disease, similarly general conditions of hygiene, the constitution of the infectious agent, etc. would be considered. As soon as we bestow meaning on "general determinants" this way, even if only to lend them an approximate and imprecise meaning, it is immediately clear how questionable enduring constancy is in general for those factors. It would be useless to object by saying one might always abstract some overarching and general determinants from any series of observed cases whatsoever, which would involve definite chances of this or that course of events in general. Even granted that this could be imagined in general, one would always be laden with doubt whether the same things would hold just for the next case, or else continue for many cases to follow. Clearly the conviction that there will exist general conditions which remain constant, which determine the chances of some course of events over a series of similar cases, may be based upon a simple fact that over a long period of observation, the relative frequency of various courses of events remains the same (at least to a good approximation). One should not confuse this with what is ordinarily called the determination of probability *a posteriori*. Even with a surfeit of certainty and precision, this is no business of the derivation of a numeric value which might occur generally, over some series of a few hundred cases in general. Rather this is about investigating whether such a numeric value with enduring constancy is even available. For that, it is always necessary to observe an ever-greater number of such series.

Let us suppose a sufficiently large amount of observational material were available. Let us suppose these observations establish the constancy of a value in a most satisfactory manner. Nevertheless one thing should not be forgotten. For future cases as for any other, an overall assumption of identical circumstances can be attributed no more than the moderate certainty of a conclusion by analogy. We may be led to the assumption that certain conditions existed in the same measure over long periods; consequently we expect this will also hold in future. Still we are unable to pin any definite necessity for this constancy on present behavior which is known with certainty.

Of course we see that the initial question about the existence of overall constant circumstances cannot be answered with full surety. Yet it can be answered empirically (with greater or lesser probability) just in case there exists a sufficient amount of material for observation.

5. Let us think of this question as having been answered in the affirmative. Additional questions follow, which are no less fundamental to the establishment of specific numeric probabilities. After it has been shown that the relevant phenomena follow the same rules as a game of chance, played persistently over an extended period, then it is necessary to discover what kind of rules these are. Above all it is important to know whether the phenomena ought to be considered on analogy to an ordinary game of chance. It is important to know those general conditions produce the same probabilities for every course of events in every case. It is also important to know that they are all independent of one another, in the sense given by the probability calculus. We have seen before how these relations may become evident. They become evident in a certain way under sufficiently sustained observation. That is, it is determined what dispersion results as we assemble larger or smaller numbers of cases into distinct series according to some rule or other; we compare all the results for these series to one another.

Previously we had outlined circumstances on which dispersion depends. Here we can rely on those earlier results. They tell us that when observation of a sufficiently large number of series produces a non-normal dispersion, then at least we ought to conclude that the phenomena do not behave by analogy to ordinary games of chance. Instead a hypernormal dispersion leads us to conclude something else for most cases which are grouped together: that circumstances exist which modify the ratio of various possibilities in the same direction. If dispersion proved to be normal every time, only then would we infer that the aggregated cases really are independent of one another, according to the rule by which we had grouped cases into series, and that the same ratio of various possibilities obtains for all.

It becomes immediately clear how important the determination of dispersion is. It becomes clear once we consider that one of the most basic applications of the probability calculus consists in this: the formation of expectations of appreciable certainty about a substantial series of cases which belong together. Those cases may be successive. This is impossible without knowledge of the conditions which determine the dispersion, as we see right away. These may even be derived empirically. Since in the process we also make use of observations of cases themselves in order to draw conclusions about the means by which they arise, we succeed in establishing propositions about probability. What is characteristic of this evaluation of observational material is this: we take account not only of the frequencies of courses of events, but we also take account of the groupings of cases which go one way or another.

6. Given a particular grouping of materials with large scope, at last we can try to answer a question: whether in the end possibilities within the relevant range are to be considered as defining probabilities which hold for single cases. The question is whether there exist “equal chances” for each and every case.

Let us consider the totality of possibilities for various courses of events to be established with some certainty by observation of a very large number of cases. Then it could be proven that its sense shows through as we partition the totality of materials, according to some criteria which hold for some cases and not for others. We would have to choose criteria by way of example as we are inclined to trust had

some connection to the course of events, of course. We might – taking up the example cited earlier – categorize observed cases of typhus. Perhaps we could arrange patients by age, or contrast patients found to be well-fed with others found to be emaciated, or something similar. Suppose it happened that whenever we divided groups by a characteristic in any way at all, that the relative frequency of a fatality remained the same across all groups. Gradually we might be convinced that no particular characteristic of cases we can observe will make any course of events more probable. Then the notion of a totality of possibilities makes definitive sense: there exists complete equality of chances for all cases. Then we could, for example – for some game of chance where it is really unnecessary – furnish empirical proof that arbitrarily applied procedures do not affect the relative frequency of outcomes. Approximately the sixth part of all throws of a die will always come up 1, whether we toss the die with our left hand or our right, whether we toss it in the morning or afternoon, whether we toss it during the moon's waning phase or its waxing, etc. But in the earlier examples of cases of illness, as soon as sufficient observational material is available, it is easy to show an essential dependence of the course of events on various characteristics. So it happens that there the passage of phenomena differs from that of ordinary games of chance.

7. While the establishment of statements about probability appears to be tied to investigations which are far from simple, complete disregard of such a test of relevant series is a characteristic mark of methods which the academy of mathematicians has developed for applications of the probability calculus. For them, everything becomes subordinate to the calculation of probability: meaning any series of cases of the same kind where one part charts one course and another part another course. For them it is as if many things were thoroughly established already: the constancy of conditions, the property of independence, and the equality of chances for single cases. As a consequence we find quite simple and definite rules from which “the probability of the same thing in a similar kind of case” is derived from the frequency of outcomes for cases already observed. Even probabilities for future cases of the same kind are to be derived from them.<sup>3</sup> These rules are established without restriction, as when we read – to point out only one instance – in Laplace:<sup>4</sup> “Most simple events have a probability that is unknown; in considering the probability *a priori*, all values included between zero and one seem admissible, but if one

---

<sup>3</sup> Compare the rules for determining *a posteriori* probability in **Laplace**, *Théorie analytique des probabilités*, Book 2, Chapters 1 and 6; **Lacroix**, *Traité élémentaire du calcul des probabilités*, p. 134; **Meyer**, *Vorlesungen über Wahrscheinlichkeits-Rechnung*, pp. 176 and 179, etc. Even **Poisson**'s much more involved theory only appears to be distinct from the simpler common theories. Basically, for them the chances assigned to individual cases are basically composed of several random judgments. The constancy of general conditions and the independence of individual cases are still assumed to be self-evident. [Laplace, P.-S. (1812). *Théorie analytique des probabilités*. Paris: Courcier.; Lacroix, S.-F. (1816). *Traité élémentaire du calcul des probabilités*. Paris: Courcier.; Meyer, A. (1879). *Vorlesungen über Wahrscheinlichkeitsrechnung*. Leipzig: B.G. Teuber (E. Czuber, Trans.)].

<sup>4</sup> Op. cit. Chapter 6.

has observed a result composed of many of these events, the manner in which they enter in makes some of these values more probable than others. Then inasmuch as the observed result is composed by the development of simple events, their actual possibility is known more and more, and it becomes more and more probable that it falls within bounds which – continually narrowing – would finally coincide if the number of simple events were to become infinite. In order to determine the laws according to which this possibility may be discovered . . . etc.”<sup>5</sup>

Now one might believe this general formula is just some vague expression. One might believe that certain tacit assumptions will have been made about the cases to which the indicated method is applied. But this is not so at all. In the first place, an entirely general application of the probability calculus is repeatedly emphasized. It is emphasized on the authority of these rules to the point of insistence, though really the method is applied even to multitudes of things for which the necessary postulates fail to hold at all. A collection of not irreproachable results is obtained by such means. Some are simply illusory. Examples occur in actuarial accounts of human society, particularly with mortality statistics. Other examples occur in witness testimony and judicial decisions, and even in the findings of medical statistics. We have yet to concern ourselves in particular with subjects where attempts have been made to apply the probability calculus. Still at this point we may proceed to elucidate the generally unsatisfactory nature of results to which one is led when one judges these phenomena as ordinary games of chance. That is to say: when one applies a schema that does not represent what is at hand. It is clear that this academic procedure – call it the orthodox procedure, for short – produces probabilities which one would have to estimate for an expectation if certain postulates were to count as certain. In other words, if these conditions counted as certain: constancy of general conditions, independence of cases, and equality of chances for single cases. It is clear these procedures may lead to entirely different results than for probabilities where we recognize those assumptions hold. The results could be different if we consider that these postulates are in doubt, or even if we know for certain that they do not hold. The difference between results of the orthodox procedure from what could be considered a valid expression of knowledge, is something we now turn to examine more closely with respect to various procedures.

**8.** The first of these consists in drawing a conclusion about the conditions under which some cases arise, on the basis of observed cases. The probability calculus teaches us that if  $n$  cases follow one course and  $m$  cases follow another, then overall conditions existed from which the possibility of one course of events and the

<sup>5</sup> « La probabilité de la plupart des événements simples est inconnue; en la considérant a priori elle nous paraît susceptible de toutes les valeurs comprises entre zéro et l’unité; mais si l’on a observé un résultat composé de plusieurs des ces événements, la manière, dont ils y entrent, rend quelques unes de ces valeurs plus probables que les autres. Ainsi à mesure que le résultat observé se compose par le développement des événements simples, leur vraie possibilité se fait de plus en plus connaître et il devient de plus en plus probable qu’elle tombe dans des limites qui, se resserrant sans cesse, finiraient par coincider, si le nombre des événements simples devenait infini. Pour déterminer les lois, suivant lesquelles cette possibilité se découvre etc. »

possibility of the other are given by  $\frac{n}{n+m}$  and  $\frac{m}{n+m}$ , approximately. Beyond that it teaches us – if we are to derive a definite probability – that these values do not differ by more than some measure from the real extent of those possibilities. Here we may append the remark that such a consideration has meaning only if the cases actually did follow different courses, and therefore neither  $n$  nor  $m$  is equal to zero. Such a consideration has meaning if we make it clear to ourselves in some way, how conditions can even be imagined which produce the possibility of one course of events or another. Suppose **Laplace** wants to derive how large the real probability of sunrise may be on some day, from the multitude of days on which sunrise and sunset have been observed to occur. It is easy to the point of being facile to say what he seeks here is not something real. – It is a make-believe value sought under an entirely arbitrary fiction which is doubtless false. Suppose in that way somebody did decide by lottery if the sun would rise or not each morning. We might stop to ask how large the chances of that event (or its contrary) might be on each occasion. Yet since we have to conceive this situation differently, we do not lend any meaning at all to such an investigation. – When we compare a series of cases to a game of chance, then to some approximation the possibility can really be deduced from a large number of cases – given some general conditions for each course of events. That all depends on the courses of individual events as determined in part by certain general conditions, and in part by the random arrangement of conditions on each occasion. No doubt undermines the validity of this procedure in general. For that reason it is all the more important that when we use it we are not bound to the premise that general conditions are precisely constant. Even conclusions drawn from the phenomena themselves may be used to establish how circumstances may change gradually – phenomena from which the general conditions have only just been evaluated. That is a question foreign to the orthodox way of thinking.<sup>6</sup> However, it is one that must be raised. Here there lies a danger: the danger of being led to false results by the ordinary way of applying the probability calculus. One must always be suspicious of this, once one no longer restricts the investigation to reporting definite values which are approximately correct for possibilities under general conditions. Instead sometimes one wants to characterize certainty and precision in numbers. The question how much certainty may be inferred from observed phenomena about this or that behavior under existing general conditions, is of a piece with another question: which possibilities may represent definite general conditions for various results of observation. No definite claim may be made, if we do not know what general conditions in question may be significant, hence what dispersion is expected for the results of those series when the relevant conditions are constant. For a certain number of observed cases, it is easy to say much can be concluded about the conditions from which they arise. Conclusions can be made with greater confidence the more evidently the general conditions are impressed upon each series – in other words, the smaller the dispersion that is to be expected. Suppose the ordinary

---

<sup>6</sup>Following **Lexis**, it is this way of thinking according to which certain series are called 'problematic'.

procedure consistently represents the connection of overall conditions to the results of the series. Then given a constancy of conditions we should expect a normal dispersion of results. But in that way the certainty of conclusions which proceed from observations to general conditions will be wrong in general. Often they will be overestimates; often they will be underestimates. An example should make this crystal-clear.

Beginning with the number of crimes against property committed annually in a state, we are concerned to draw conclusions about the constancy or change in general conditions which constrain that number. We can imagine that for any individual, there is a smaller or larger yet definite possibility the individual could commit such a crime over the course of a year. This involves factors of his moral constitution as well as factors of his living conditions. The general state of society might then be characterized by the average value of this possibility. It might be characterized by the number of individuals for whom this value lies between 0.0 and 0.1, between 0.1 and 0.2, and so on. This possibility would be very large (nearly equal to 1) for a small number of people, while for most it would be comparatively tiny (nearly equal to 0). By contrast we do not imagine – at least not very distinctly – general conditions that the same possibility is generated for every single individual. That which may interest us about the moral condition of the society can surely only be represented in the former way, not the latter. It follows that the general conditions are expressed with greater fidelity as the numbers of crimes actually committed, than could be expressed if the possibility were fixed as the same for all individuals. Randomness plays but a minor role here. Say the general conditions in question remain constant over a long period. Then if each specific chance always arose for a definite and unchanging number of people, a very strongly hyponormal dispersion of the annual number of crimes would be expected. Following the procedure detailed on pages 108 and 109 [here the latter part of Chap. 4, Section 9], cases of very disparate probability are always collated in series which have a very specific numeric ratio. And given this, there is no doubt we cannot endorse a conclusion from the number of crimes committed to the ‘moral state of society’ by attributing a numeric rating to its certainty and precision. At any rate we underestimate both certainty and precision if we treat these phenomena in the same way as ordinary games of chance.

This example alone draws attention to a situation which only gains in importance for many other examples. That is the vagueness of the idea of conditions which apply to any case whatsoever. As was mentioned, we must be conscious that often this idea seems to be bulwarked only by the fact that certain ratios remain constant for a time. For example, by the overall conditions of a game of chance, we never understand anything else than all those determining factors which obtain for the game’s whole duration, as in each case. As soon as we are concerned with an inference from observed phenomena to the conditions which hold for them, and once we set out to discover the constancy or variability of those conditions, then that definition no longer holds. Often under these circumstances it becomes quite arbitrary what one is willing to call the general conditions of a series of phenomena. Whether one has this idea or that in mind, the certainty of the inference will be entirely different, just as expected dispersion is unequal, whether we think this or that general condition will

be fixed over a long period. An example that shows this very clearly is found in the tally of people who die of an epidemic disease. We may catch a glimpse of the general conditions which underlie a definite possibility of death by typhus for each individual. On one hand that consists in the extant amount of the typhus agent, ignoring such things as health and lifestyle factors, etc. Under that assumption we should represent the prospective situation as having no significant error, as if a certain level of danger awaited each individual. It is as if the survival or succumbing of each individual represents a series of independent cases. Then if a conclusion is to be made about these general conditions from incidence numbers, the usual normal dispersion seems applicable in the basic formula by which certainty and precision are calculated. On the other hand we can also include those relations among the general conditions which produce definite possibilities for the emergence of larger or smaller epidemics, or else pronounced or less-symptomatic epidemics. Let us consider conditions of the latter sort as enduring and fixed. Clearly there are circumstances that randomly modify the possibility of dying of typhus in a large number of cases grouped in spatial or temporal terms, but they modify the possibility in the same direction. A strongly hypernormal dispersion would be expected as a consequence. Correspondingly, we can then make an inference – often with great certainty – from variations in typhus mortality to changes in ‘general conditions’ in the former sense. At the same time there is no sufficient basis to assume any variation in proportions under the latter interpretation of ‘general conditions’. So clearly in fact only sometimes can it be concluded with an appreciable degree of certainty and precision from a series of observed cases to some general conditions (or other) which determine them. That happens only when it may be one knows for sure what dispersion would be expected among the outcomes of many similar series, given that the constancy of such general conditions is expected. If such constancy does not occur, then of course the desired value cannot be derived empirically. Then one depends on making conjectures on the basis of what is known about the creation of the relevant phenomena, or their connection with one another, and so forth. Here it is necessary above all that one’s ideas are as rigidly precise as possible – one’s ideas about what can really be concluded, and about what are understood to be general conditions. But here we gain a new insight: this is not the way to a definite result in numbers which expresses the certainty of the conclusions at hand.

9. Let us turn to the second task that was mentioned. That is the task of reporting a probability of this or that course of events for a new case, on the basis of a large number of observed cases. Here I confine myself to one circumstance in order to demonstrate the inadmissibility of consequences to which orthodox procedures lead us in judging a single case. Naturally one can ask if the general conditions under which previous cases arose hold equally well for a fresh case. That is relatively unimportant, compared to another question: if the empirically determined totality of possibilities for one course of events or other is to be considered as setting a definitive chance which is valid in every case. There is no doubt this does not hold at all generally. There is not the least justification for indiscriminate use of the word ‘probability’ for a statistically estimated frequency of any behavior, meaning a probability it is expected to have in every single case. Often this is so glaringly

obvious, that one shies away from making such claims. At least these days no one talks about determining the chances – valid for anyone – of committing a crime, or of marrying, and so on. Yet this objection undermines – perhaps negligibly – what one is used to calling the probability of dying within a year for a single individual, as found by average mortality for the appropriate tranche of age. This counts as distinct only if any particular circumstances which might be considered are circumstances of the case entirely unknown to us. For example that might happen if we need to form some expectation about a person who is selected by lot from a whole population with a known average. One must be aware that facile numeric probabilities reported under such circumstances are of very limited validity. They represent only someone's opinion who is totally ignorant of the person in question. Consequently such probabilities should in no way be declared "to represent the actual chances". – Sound human understanding tells us incisively that probabilities which are judged to hold for individuals should be applied with care if they are derived by the usual methods. Then the trap of being significantly deceived on this point may have only been encountered seldom, or to small extent. One sees very easily that the results of observations of collective phenomena cannot constitute the standard of measure by which we judge individual cases. We see this once our familiarity with those individuals covers anything that motivates a different expectation. Yet if one restricts this type of judgment to conditions where knowledge of individual circumstances contributes very little – only what is insignificant to prediction – of course one may apply those methods repeatedly. At least one will not be exposed to gross deception, though one's methods prove not to be correct and will not withstand closer examination.

**10.** The situation is very different for the task which will now be discussed. We are concerned to report probabilities for the outcomes of a large number of prospective cases. It is easy to see that here the tacit postulate of a constancy of general conditions occasionally guides orthodox procedures to results which completely diverge from a really adequate expression of knowledge. Suppose that the probability of some course of events is shown to have such-and-such a size on the basis of a large number of earlier observations; the same probability is credited to future cases. Then for a middling number of those cases we obtain a very high probability that the events in question occur approximately so-and-so many times. – More precisely, we obtain a high probability that their relative frequency lies within specific bounds.

Naturally one considers expressions of this kind – which produce high probabilities for collective behavior over many cases – as the most valuable results of the probability calculus. Almost every application of the probability calculus by these methods boils down to this: drawing out large probabilities for a behavior of interest. It is of the utmost importance to recognize that the very way these very results are obtained places their accuracy deep in doubt.

To introduce an example, not infrequently we encounter such ascriptions in the application of the probability calculus to the evaluation of therapeutic success. On the basis of previous experience, it may be reported what probability is assigned to a fatal outcome for a case of typhus treated by a given method. This is followed by a statement of probability that we ought to expect so-and-so many fatalities among

20 or 30 new cases treated by the same method. These kinds of numbers are nothing but an illusion. One can tell this immediately once one remembers they report probability under the postulate that the general conditions for the genesis and progress of the disease are identical to that in previous cases. High probabilities are unfounded where there is the slightest doubt that general conditions may have altered in some way. The numbers may give us a coarse gauge of a very high probability – a probability that there is no particular grouping which stands out among the same general conditions, nor that any noteworthy random event might lead to dramatically different results than results found earlier. Those numbers give no expression to the (perhaps) quite negligible certainty with which we are justified in assuming such general conditions have been secured. Therefore they are also far removed from appropriately reporting the certainty with which we should expect the result. The situation is just the same with some high values for probabilities one can compute: for example, that the fatalities which occur in one year in a large community may fall within certain bounds. As with any probability, such numbers express only that given certain general conditions, one needs an exceptional arrangement or disposition to give rise to significant deviations from the expected result. The numbers indicate how large an operational range might be represented by this or that outcome. Only if it counts as free from doubt that all conditions which determine mortality, such as general health, lifestyle, other harmful propensities, etc. – only if those conditions persist as they have been revealed by previous experience – can one then hope to express the certainty of our expectations justified by a number merely for the possibility of arranging these conditions in groups. – Add to this, that in such situations even the (similarly presumed) independence of single cases proves to be very doubtful as a rule – given some thought. Or else it proves not to be present at all. If any general conditions can actually be assumed to persist, then for the total results of a series to specify the probability of a future series adequately, above all it is necessary to know that a hypernormal distribution is not to be expected from the series. Some slight deliberation tells us it is most probable that chances will vary in the same random direction for a good number of cases which stick together, with quite extraordinary frequency (that includes the example just presented). Under those circumstances even the numbers we mentioned will be incorrect, if one presumes ahead of time that general conditions we think are constant really are constant. Later we will introduce another common example where an unjustified postulate of independence among single cases also leads to entirely erroneous results.

## Chapter 7

# On the Significance of the Range Principle, and the Probability Calculus



**Abstract** Representation in numbers is by no means an essential mark of probability. A special position is occupied by probabilities which can be represented in numbers. When probabilities are not numeric, no methodology alleviates the lack, though one may still talk about communicating a psychological degree of surety. Acquaintance with the range principle – of probabilities as ratios of ranges – enriches our knowledge, and resolves certain issues in explanation. The range principle is not empirical, in that it is neither proven nor refuted by experience. It does offer a satisfactory explanation of the law of large numbers. The theory of probability has been extended to doctrines of rational methods for preference and choice. Yet there is no general principle by which the value of goods in hand may be compared to the value of a prospect of possessing those goods, as emphasized by a presentation of the St. Petersburg problem.

**Keywords** Range principle · Lawfulness of events · St. Petersburg problem · Logic of probability · Probability and value · Conventional assignment · Reasonable wager · Psychological value

1. The theory of numeric probability – which I was at pains to frame and develop – hews close to some considerations and questions of more general import, which I believe should not be suppressed here. Though they do not belong to the subject of our study in the strictest sense, they throw new light on our study, which may relieve some afterthoughts that remain.

The basis of our theory may be considered the proposition that premises are equally probable which span equal, original, and indifferent ranges. From this it follows that premises which span either a very large or a very small fraction of a range which seems as complete as possible – a fraction that approaches one or approaches zero – are to be considered very probable or very improbable, respectively. In the first place it may be noticed that for this probability, a numeric representation is by no means essential. It is no constant mark of such probability. The measurement of probability in terms of ranges happens particularly often, even when several premises span ranges which are not quite indifferent, and not strictly comparable. It follows that the Range principle has an importance which reaches far

beyond numeric probability. Say for example, we were to focus on an inference which led to the conclusion that the specific weight of a newly-discovered body lay just between 4.25 and 4.26. In the event that that appeared significantly uncertain in advance, we would be consistent in saying it was “not at all expected” that the specific weight had just this value. As a consequence we might attach a very negligible probability to the premise that it was. That is not to be represented numerically, for reasons we have already rehearsed. At any rate we can say only that the premise at issue spans a very small part of the whole possible range. Simple deliberation shows us many ways the unconscious logic of everyday cognition brings this principle to bear in an entirely straightforward way in application. Any time we consider something improbable because its realization would be “a remarkably chance event” or “a particular coincidence”, we may be clear this deals with the formation of expectation based on the Range principle. In most cases we see that this occurs without it being known to us what specific ratio exists between different ranges – as a rule, without there being such a property at all. Here as in so many cases, we must be satisfied with vague expressions like many or few, and large or small.

In this extended sense the Range principle should find its place in a general theory of probability beside the logical relations of analogy and induction as characteristic sources of probability. (I leave aside the question whether there may be other logical relations of interest.) Since it is not my present intention to construct a theory, it is enough to indicate that the logic of our knowledge and conjecture usually offers us the most varied types of probability in ramified combinations and condensations. The ability to specify probability in numbers for a premise seems connected to two characteristics of logical relations. The first is this: probabilities are not considered which do not rely on the Range principle. The second is: that indifferent ranges must stand in an original and expressible ratio if they are to be compared, in fulfillment of the first characteristic. Accordingly it is clear what a special position is occupied by probabilities which can be represented in numbers, among all kinds of probability. This is how a natural error which was implicit in earlier theoretical accounts has been avoided. The error is that an arithmetization of all probabilities must be the aim and end result of a thoroughgoing logical examination of our knowledge. It is not just a gap in our combinatorial powers or deductive powers if we should not arrive at this result.

**2.** We may consider the Range principle from yet another perspective, and find a certain parallel to the principle of the lawfulness of events.

If we form expectations about future events on the basis of something which has already been explained, then we are prompted to ask: in the final analysis, what is the basis for such expectations, by which they can be justified? As alluded to earlier (p. 19 in the original and the passage following: Chap. 1, Section 6), often this question finds a satisfactory resolution by appeal to an assumption of the thorough-going lawfulness of events. Indeed we can perceive certain conditions in the world directly, with immediate certainty that precludes all doubt. That explains our insistence on always expecting anything we can think of as lawfully connected with these conditions.

By contrast, if we have noted the frequency with which a particular kind of outcome has occurred for many cases, and if over many additional cases we expect various outcomes to have the same relative frequency on average, then as we saw, this procedure is not rendered legitimate by a principle of lawfulness alone. Justification by such a principle does not seem intelligible on its own. The course of our investigation has shown that such expectations are in fact justified, but they may always find their definitive basis in a quite different principle: the Range principle.

Suppose we expect that for the oft-repeated tossing of a coin, heads and tails turn up about equally often. Through the Range principle on one hand we develop an expectation about constraining circumstances with which we are vaguely acquainted. On the other hand as a necessary consequence we count ourselves justified in attributing various possible combinations of specific outcomes to the tosses. The principle of lawfulness (on the one hand) and the Range principle (on the other) are similar in logical importance. They have this similarity despite their differences in kind. Clearly their similarity consists in this: both are definitive principles, they admit no further justification, and they are capable of no further explanation. We develop expectations about actual behavior from them both. Despite the circumstance by which we ascribe objective meaning to the principle of lawfulness, but merely subjective meaning to the Range principle, here both are considered the same way. That is because the former principle is a principle that fixes subjective certainty at the same time. It has a wholly similar place to the Range principle in that much. Some things we can say very little about: one is what right we have to assume that certain circumstances necessarily lead to certain consequences – accordingly, what right we have to expect that realization of such circumstances really leads to their lawful consequences. That is inchoate, just as we might not say what right we have to hold the realization of a distinctive form of behavior as very improbable within a large range of possibilities. – Even from this perspective, the meaning of our principle proves far more general. It is far more general than if it were restricted solely to cases of probability which really are numeric. The most salient feature of the principle was found to be a satisfactory explanation of the so-called Law of large numbers. This is no less valid for cases in which appreciable numeric probabilities cannot be claimed. It is no less valid in situations where only approximate equivalence may be expected with any certainty for average proportions. We know when completely similar cases recur very often, it can be expected with great certainty (from ratios of ranges) that the average frequency of a phenomenon may not deviate significantly from a certain value. That average frequency may be predicted given only very general and coarse-grained familiarity. From this it follows that where in broad terms there is very frequent repetition of similar cases in approximately equal proportions, more or less approximate constancy ought to be assumed for the average frequency of an event. That represents an overwhelmingly large range of combinatorial possibilities. Generally such expectations are justified, if it can be presumed that certain general characteristics which mark each case will remain nearly constant, or change only slowly. Simple deliberation tells us that it is often necessary to base such expectations about the average behavior of many cases by appealing to the Range principle. Often it is not difficult to come to the

recognition that adherence to the expected regularity is not to be considered necessarily connected to what is known to be manifest. – Such a guarantee cannot be ensured as generally necessary by some connection among individual cases. Once we can say that much, no doubt the relevant expectation has to be based in part on the Range principle. Through this insight we dispel difficulties which would otherwise proceed from such considerations. For example, in almanacs we find that of all the letters sent by post, only a small and almost unchanging percentage ever bear an incomplete address. There is no apparent basis for doubt much the same will occur next year. It seems quite unimaginable there should be a lawful regularity which produces this result. It is impossible that persons *X*, *Y*, and *Z* give incomplete addresses to their letters because a limited number of such cases are available, and hence other circumstances exert an influence. That would mean persons *A*, *B*, and *C* do not commit the same lapse in addressing their letters because the number of lapses is limited. There is no connection among separate cases to guarantee their compliance with a specific frequency of such omissions. Determinants of current behavior are known to us so seldom, of the kind where we can imagine a putative result to arise from them by necessity. Consequently we tell ourselves (at least at first) only an approximate constancy may be expected for any overall condition. We tell ourselves that a given frequency of such omissions is just to be considered the most probable outcome, because it represents the largest portion of the range which seems possible.

To justify the certainty with which we consistently expect certain regularities for some events – say the sprouting of a seed under favourable conditions – we call upon the lawfulness of events in general, of course without being able to articulate specific laws in play on the occasion. In the same way in other cases, we are able only incompletely (if at all) to describe overall conditions from whose particular structure we develop expectations by the Range principle. The ways and means by which we tie our expectations to principles which underlie them, are often even less well-determined. If we develop expectations for average outcomes over very many cases of the same kind, often enough a complete independence of individual cases cannot be claimed. Then it remains *in suspenso* how much and in which aspects our expectations are based on an assumption of lawfulness, and just how much they are based on the Range principle. For example, if we state that in Goethe's prose writings the letter *E* generally occurs 160 times on average for every 1000 letters, we ought to expect with great certainty that each time we tabulate letters of a printed page in Goethe's *Collected Works*, we will find about the same proportion. Yet obviously it would be unsound to imagine that for each position in a long string of these letters, there is an independent and equal probability that that position will be taken by the letter *E*. For one thing, the character of the language excludes the possibility that three or more letters *E* follow one after another. Forms of declension and conjugation demand that very long sentences cannot be written while avoiding all use of *E*. Even significant reduction in the usual proportion might not be contrived without unusual duress, or without damage in aesthetic terms. So a certain minimal and maximal frequency of the letter *E* is necessarily given by the character of the language, as well as by the nature of what is to be expressed. Besides, we may also

expect that random fluctuations which may always be considered possible, are found to be balanced over a large number of sentences. Thus it does not appear to be expressible how much that expectation is based on lawful connections with conditions known to be present, and how much is based on the assumption of a balance of random events – in other words, on the Range principle.

Similarly we might tell ourselves that the expectation of a certain yearly average of temperature at some location must be based in part on understandable natural necessity, but also in part on the estimation of ranges. The warmth the earth receives as radiation from the sun, and how much it reflects back to nearby space, the available volume of water, and the extent of cloud cover, etc. are factors whose exact or very approximate constancy we may predict with certainty. We may think it ruled out by known conditions that a location which had a yearly average of +8° or +9° degrees previously, will exhibit an average of -4° next year. Perhaps we would not hold a value of +6° to be entirely impossible, but we might consider that to represent a noteworthy but random exception anyways. The certain expectation of a value between 8° and 9° owes some justification to both principles of expectation, but not in a manner precisely determinable.

If the Range principle renders expectations of this kind comprehensible generally, there is one essential advantage of this explanation. The advantage is that the danger is removed that we may be deceived by some misconception about those expectations.

One knows what errors lurk here. Consequently one knows what a very controversial thing the “Law of large numbers” has been. So long as one knows no other principle of expectation than necessity of causal connection, one is always subject to the temptation of construing regularities of collective phenomena exclusively in this sense. That is, one is subject to the temptation of conceiving all cases as connected, so that some approximation to definite average results must necessarily emerge given greater and greater numbers. The confusion is uncommonly widespread with almost all collective phenomena, even for games of chance. **Laplace** even warned us about such illusions in the passage « Illusions dans l'estimation des probabilités ». He introduces a relevant example: if a coin toss turns up heads nine times in a row, most people would assume with heightened certainty that the next toss will finally turn up tails. This expectation would only be well-founded if individual trials are not independent of one another, but cohere. Results which were not likely at one time would make the opposite outcome likely for a number of cases to follow. Such a connection – which would make itself known by a hyponormal dispersion for the average results of series – most certainly does not occur. Even in the case just mentioned, on the tenth toss heads and tails would be expected with the same probability once again. Some departure from this pattern would occur only if we were unsure of the physical regularity of the coin in play, but then it needs might happen in the sense that a result observed nine times in a row might be considered more probable on the next toss. **Laplace** continues: “I have seen men, who fervently desire to have a son, disappointed to learn how many boys are born in the month they expect to be fathers. Imagining that the proportion of these births to those of girls must be the same at the end of each month, they thought that newborn boys render

the forthcoming births of girls more probable".<sup>1</sup> **Laplace** assumed correctly that an illusion is at work. – More importantly than in this arena, similar illusions about regularities appear in statistics of ethical concern. There the same misbegotten grasp of the concept – mainly advanced by **Quetelet** – is held dear, as if the actions of individuals are determined not just by their conditions, but in some incomprehensible way by the states of all the others too. It appears as if “the budget of the scaffold and the prisons” had always forced its own logical fulfillment. Maybe if a certain number of crimes were still lacking, someone would be compelled to commit them by an inexplicable influence. If one characterizes this idea only as a very ephemeral error, still one must admit there actually was a lacuna in understanding here which called for a solution. The inadverturous notion which is currently widespread cannot claim to have overcome the difficulty. For example, **Lexis** – who has stressed the independence of individual cases especially clearly and sharply for most collective phenomena of human society – restricts himself to citing other examples as corroboration. He cites games of chance, and the relative frequency of letters in spoken or written language, and the like – there, even without any connection between cases, all the expected and regular features of a totality of outcomes may appear. The Range principle grants us more satisfactory understanding, even if only in a quite general way at first. It teaches us what the definitive basis of those expectations is. Even for phenomena of human society, we seek prematurely and in vain to demarcate what may be attributed to one principle of expectation, and what is to be attributed to the other. We are capable of telling ourselves that physiological laws must guarantee some similarity of nature between one generation and the ones to follow, of course, or that habits and needs are necessarily maintained with a certain steadiness, or else that governmental and social institutions only support rapid and substantial change by exception. At the same time it is indisputable that only some really odd groupings of these familiar conditions could even have led to changes or fluctuations we do not expect. The constancy of collective phenomena is far too large for what we ought to consider necessary from conditions generally.

3. Concerning ranges, interpretation of the regularities which do occur for collective phenomena seems to prove insufficient on one very important point. One may be inclined to assert that some fundamental difference exists between the principle of the lawfulness of events, and the Range principle. One may assert that the former has objective meaning while the latter has merely subjective meaning. Then if the latter seems suited only to justification of a certain type of expectation, one does not need to explain why these expectations are confirmed by experience. To draw this out into an explanation, would only be to repeat an attempt which logicians have so often proven wanting – that is, finding a way of basing explanation of real behavior on a principle of purely subjective import. It remains as

---

<sup>1</sup> « J'ai vu des hommes, désirant ardemment d'avoir un fils, n'apprendre qu'avec peine les naissances des garçons dans le mois, où ils allaient devenir pères. S'imaginant que le rapport de ces naissances à celle de filles devait être le même à la fin de chaque mois, ils jugeaient que les garçons déjà nés rendaient plus probables les naissances prochaines de filles. »

incomprehensible as before, where those regularities come from. – Here the difficulty is alleviated by very simple considerations; they are actually necessary to gain a right understanding how these regularities may be explained.

To begin, I hope that to this point the whole presentation of the merely subjective character given to the Range principle as a principle of expectation, has not aroused any doubt. So if we should speak at all of the explanation of certain regularities which really have been observed, the explanation is not given by the rule of expectation expressed by this principle. Perhaps something can even be found in those objective facts, by which this rule determines our expectations, namely by the ratios of certain ranges. If a phenomenon always appears under some generally described circumstances, then an “explanation” of it that satisfies our intellectual needs will be found, if we are in a position to realize that the phenomenon is connected not to all the specific arrangements which span this general characteristic, but rather it is connected to the large majority of them, by some necessity we can understand. Of course one ordinarily calls such an insight into the connections between events an ‘explanation’. It would be as useless as it is pointless to debate whether or not it should be correct to call this ‘explanation’. The taking-into-account of ranges tells us that there are other explanations based exclusively on lawful necessity. They are distinguished by the sense in which they constitute explanations. They tell us further that explanations which contain only statements of lawful connection cannot be given for those regularities. Therefore it is wrong-headed to demand such explanations from the outset, since they depend on an erroneous premise. Nevertheless we can imagine a state of knowledge that represents acquaintance with every detail, under which every particular event could be understood as necessary. Yet for general propositions which express the regularities at hand, in this sense there is no explanation at all. There is no real need to demand such an explanation. That merely represents an intellectual compulsion borne of delusion. Inasmuch as one demands an explanation that a phenomenon occurs regularly under certain conditions, one is already tempted to proceed on the premise that an explanation may be given by describing specific lawful connections. Yet the demand for general explanation in this sense – as confirmed by the very safe expectations given by the probability calculus – is inadmissible. What has been demanded simply does not exist. Consider the question how it happens that in roulette the ball always lands equally often on red as on black (to a close approximation). There is only one answer (leaving aside any indication of the preponderant size of the ranges which led to this result): that is the way the real situation is constituted. It is useful to clarify this by appeal to analogous conditions in which ratios of ranges are directly open to inspection. We would find it absurd, if someone were to remark about the celestial sphere, that: (1) a large number of bright stars never appears distributed around the same great circle with even angular separation, and (2) there must be some reason that such an arrangement may be impossible, and so (3) an investigation must be undertaken to determine how this situation has been prevented. It just seems to us entirely satisfactory that this quite peculiar arrangement of stars simply goes unrealized. In the same way we always find our intellectual curiosity satisfied, if we have proof that certain phenomena we have never observed, are only found for a

very special arrangement of their constituents. Acquaintance with ratios of ranges represents an enrichment of our knowledge, one which resolves certain issues in explanation. If one wishes to cling to the idea that only necessary statements can be called ‘explanations’, one should recognize that an investigation which offers any kind of explanation in this sense is already important. Our meditation on ranges teaches us that this is frequently not the case, even where one is inclined ahead of time to believe it is. – The afterthought with which we began – that the Range principle might still not explain actual regularities – is then resolved as follows. Acquaintance with ratios of ranges does explain those regularities in a definite way. This leads directly to the insight that it is neither possible nor necessary to explain those regularities with some other interpretation.

Naturally the situation is different when definite knowledge of general circumstances and the ranges of some processes are still works in progress. – Then we can bring an observed regularity into correlation with the Range principle in the very general way sketched previously. How does it happen that the same number of marriages are celebrated each year? We cannot yet render this completely comprehensible by referring to certain ratios of ranges, as we may when a roulette ball lands on red or black with equal frequency. Still, it is clear this is not the task of a logical principle. Rather it is an investigation involving real conditions. Even in the face of such phenomena, the explanation which is positively offered by consideration of ranges consists in this: a range is the token of a general form for the results of such an investigation. It is a token by which we can represent results to ourselves *a priori*, by which they morph into a comprehensible problem. Without this form they must have appeared an incomprehensible riddle.

Here too one can speak about explanation in a useful way, which may hardly be entirely objectionable. There exists a difference which is simple to understand between the type of explanation given here, and the kind of explanation spoken about earlier. There we draw connections between phenomena to be explained and facts which are otherwise safely known. Here on the other hand, we draw connections to premises which we deem acceptable or plausible. As far as I can see everything is set for explanation of these problems by ordinary logic once this result has been achieved. Ordinary logic is capable of explaining them, and its duty is to explain them.

**4.** Obviously the Range principle is not empirical, very like the law of cause and effect. Like that law, it is neither proven nor refuted by experience. The confidence with which we apply it may increase according to laws of psychology, of course, should expectations based on it often be confirmed by experience. Nonetheless one cannot speak of empirical proof, simply because this principle has no content which could be derived, as content is inferred from empirical propositions. Just as for the law of cause and effect, failure to confirm expectations would have the immediate consequence that specific assumptions made about empirical subject matter are incorrect, to the extent that they have brought the principle to bear. In other words, the ratios of some operational ranges are different from what we had believed. If what we consider very improbable stubbornly intrudes itself beyond what we consider probable, initially we assume errors have been committed in determining

the relevant probability. If our expectations should be disconfirmed even after the validity of those determinations were beyond doubt, we would declare this a “very remarkable coincidence”. Nothing would be produced from this which could hinder further application of the principle, or which might replace it with something different. – At the same time I would hesitate to speak of propositions about probability as synthetic *a priori* propositions, as **Fick** does. His phrase pays insufficient attention to the composite nature of the subject matter expressed. The determination of magnitudes is a matter of real significance, and so the content of propositions is called empirical, even if one attributes *a priori* evidence to the comprehensive but purely mathematical relations they have. One principle which really is non-empirical (“Premises are equally probable when they span equal, indifferent, and original ranges.”) ought to be called a synthetic judgment *a priori*, if one would have it count as synthetic at all. So far as I can see, that much is arbitrary. It appears to me that the understanding of other synthetic judgments *a priori* – such as that 2 times 2 equals 4 – will not be advanced significantly if such a heterogeneous proposition as this principle of probability is listed in the same category.

5. Clearly one essential difference between the Range principle and the principle of lawful regularity is this: the former does not represent a necessary and indispensable element of all empirical knowledge, in the way the latter principle does. We can imagine a mental state that represented such exact and detailed familiarity with reality that it would end talk of seemingly-possible ranges, or the unrestricted formation of expectation. If by the nature of things even such quality of knowledge could not be considered absolutely certain, then the probability we attribute to it – given the strength of its certain correspondence with experience – could be such that the Range principle would no longer be considered. Given that kind of knowledge, familiarity with ranges would no longer serve any purpose in evaluating expectations, nor would it elicit any other interest.

The unattainability of such a propositional attitude seems to me more important than its imaginability. Perhaps it is not too much to take a moment to clarify the large role our principle plays in contemporary fields of knowledge. One may say it is always called to play this role. Often there are cases where exact determination of behavior seems completely unfeasible, as would be necessary in helping to predict future events. In part that is because a huge number of inquiries must be made. In part it is because an extraordinary precision would be demanded. If under these circumstances our knowledge of the world remains incomplete, then both the assessment of ranges and the formation of expectation find their place. Just as one is inclined to believe at first, knowledge gained in this way may be called satisfactory in many respects. First it is to be emphasized that knowledge may need be incomplete only in an ontological sense, while explanation of the lawful connections of events can be complete and accurate in terms of nomological conditions. For example, in a kinetic theory of gases which assigned definite laws of action to molecules, and which satisfactorily explained all phenomena on the basis of such premises, we would have knowledge to fulfil the most important and the most pressing of our intellectual needs. By comparison, questions about ontological

determination would seem of more isolated importance. For example, it seems of subordinate importance how all the molecules are arranged and how they move in a certain gas at a given time. It is particularly important that the Range principle justifies expectations of such a degree of certainty that we should harbour no misgivings about regulating our actions accordingly. – At least that happens under some circumstances, certainly very often under circumstances which are actually realized, as if our expectations stood for absolute certainties. We really are used to relying on probabilities of this kind. Our confidence can be shown clearly with the help of many examples. I will not say that we consider the transmission of heat from warmer bodies to colder ones as one of the most certain facts of experience at this point, since it may be doubted this represents a proposition about probability. But we are used to considering other expectations as not in the least uncertain, though on closer consideration they prove capable of being based solely on propositions about probability. For example, no one takes objection to an instruction to mix two powders uniformly by stirring them prolongedly and shaking them thoroughly. That simply means to mix them so that within each spatial part of moderate size, both powders are present as a fixed proportion of the total. We have not the least doubt that such a result can be achieved even when it is not apparent to the eye, as with powders which look the same. Yet if our ideas about the physics of solid bodies are not wholly incorrect, there is no necessity that a thorough mixture will result. Rather, it is only that the outcome is expected with very high probability.

6. If these deliberations show that the Range principle has a definite importance, and not an insignificant one – even in cases which do not result in numerically expressible probabilities – then a question may be raised. The question is how often the calculation of range ratios finds purchase, in other words when computational methods of the probability calculus need to be applied. In this respect, what is required emerges without difficulty once we turn our attention to the many tasks performed in scientific research.

First it is clear such calculations must always be brought to bear, when we investigate whether certain phenomena are contingent upon some set of general conditions. It is just a simplified case of this type of investigation where – in the way the matter is put for the **Bayesian** principle – numerical probabilities can be reported that a certain phenomenon may have been produced. We might have shown that if two circumstances are considered equally probable, apart from some certain (*‘a priori’*) fact, then their probabilities behave like the possibilities they stand for. Something similar occurs, if for two premises there exists no numerically determined proportion whatsoever of *a priori* probabilities. Let us imagine we have had two reports about an urn: one says that the urn contains 50 black balls and 50 white balls; the other says that the numbers are 30 and 70. One numerically expressible fraction of probability does not fit both premises on offer. That becomes especially clear if one report seems more trustworthy. If then we draw a large number of balls from the urn, and if black balls were drawn 295 times while white balls were drawn 705 times, there would be no basis for any significant doubt that the latter report was indeed correct. Yet specification of this probability in numbers would be impossible, just as before. To explain how this conclusion is justified, we must be clear that according to

the probability calculus, the possibility that the urn is filled with 50 black and 50 white balls is a much smaller possibility for the observed outcome than the possibility that the urn is filled with 30 black and 70 white balls. These conditions are still very simple, insofar as this is only about comparing two premises which do not differ in any point of general significance, that is, they are not distinguished by any nomological relation. Instead there are two different types of ontologically distinguished processes, each supported by a class of circumstances. They represent known nomological conditions of different possibilities for the observed outcome. But any derivation of general propositions from experience rests on a much deeper application of a principle which can be expressed in full generality: nomological premises are more probable, the larger the range of ontological varieties of behavior which permit them to seem possible in light of empirical observation. Basically there is nothing else which gives us reason to treat a regularly observed succession as necessary, and to decline the notion that the succession may have only been random. Similarly, we consistently regard ideas about laws of action as well-founded, if they allow certain regularly observed conjunctions of events to appear – well, certainly not as necessary in general – but perhaps as representing a majority of possibilities in which the events can be arranged. Any derivation of general propositions from collective phenomena is based on that much. For example, if we find fewer pneumonia patients treated with cold baths die on average than those treated indifferently, and from that we conclude that cold baths influence the course of the disease positively, then the justification of this conclusion is based at least in part on this: that actually observed events represent a far larger range of real behavior for the basis of this premise than for the contrary premise which would explain the phenomenon by chance. As a rule, we recommend a premise with nomological content when it represents actually observed behavior by encompassing a very large representative sample of original varieties of behavior. Clearly we regard the discovery of such premises as a logical postulate anywhere that that regularity is a guide to understanding what is observed. Every regularity calls for explanation in this sense.

The calculation of range proportions is indispensable throughout investigations of this kind, as it is easy to see. But it is no less clear that results gained in this way do not always represent definite values of probability. That is because it is not possible in general to enumerate all the kinds of representation in question completely, still less to apportion an *a priori* probability value to each. Most often it is not even feasible to express in numbers those possibilities which would represent actual observations.

If we observe regularity in certain average results, and we correlate this with ratios of ranges for certain constant and general determinants, yet the probabilities of some very different premises might not be comparable at all. Those premises might perhaps try to explain a regularity by an overriding necessary connection between cases. So there really are cases where the psychological value of an entire class of interpretations cannot be characterized exactly. They may be judged in altogether different ways. The part of theoretical physics based on considerations of probability belongs to this category. It seems not to be ruled out entirely ahead of time that regularly observed connections may also prove to be necessary under some

representation of a completely different nature. There is no established standard to weigh the logical value of such a theory against a theory of probability.<sup>2</sup>

For all investigations of this kind, their business is not at all the calculation of probabilities, strictly speaking. They are concerned with the calculation of ratios of ranges. Certainly those have great significance, but not one simply described in psychological terms. If one wishes to characterize such calculations in a conventional way – as the computation of probability – then it would be appropriate to set them apart, as something like the “exploratory application” of methods, distinct from the actual computation of probabilities. At any rate it would be more accurate to speak simply of a calculation of ranges here, or else a calculation of possibilities.

7. If we consider that the second task of probability calculus is to present us with expectations of great certainty for collective outcomes of very many cases of the same kind, it is easy to see these methods cannot be enlisted should there exist no sufficiently reliable and precise postulates on which to proceed. Suppose we have no clear and reliable notion of the constancy of general conditions, and none of the independence of successive individual cases, or – what is worse – suppose we are confined to just the observations made of such cases and to conclusions by analogy drawn from them. Then even expectations developed for all outcomes over many cases seem just to be conclusions by analogy. Then the significance of the probability calculus is confined to rendering the justification of such conclusions comprehensible in a general way. Under those circumstances it is warranted to expect nearly equal frequencies for phenomena which were previously observed to have nearly equal frequencies. Yet it is impossible to form a similar expectation for any process whose frequency has not actually been observed. Application of the probability calculus seems superfluous from that perspective. It teaches nothing which is not directly evident from conclusion by analogy. It has none but a general justification of inference by analogy, so it cannot lead us farther. One catches a glimpse of a very basic application of the probability calculus, when as in such contexts as games of chance, reliable expectations are allowed even when specific experience is not available in advance. In recognizing this, one finds that application of the probability calculus seems extremely limited. One can claim that where reliable ideas are not to be obtained from the laws governing events, even the probability calculus does not enable us to transcend what is “purely empirical” – to rise above conclusion by analogy. We are still confined to expectations about what has been observed to date. However, notice that even for phenomena which are unlikely in general from our perspective, there are still some specific and isolated situations, where a few premises of general import can be formed with sufficient certainty, that we may succeed in forming reliable expectations about conditions for which particular observations have not been made. Not infrequently we are justified in assuming with great certainty that several possibilities are independent. If we should consider

---

<sup>2</sup>Here we affirm from another perspective the proposition that was expressed earlier (p. 29 ff. in original; Chap. 2, Section 2): that the inductive bases of general propositions cannot be represented in numeric terms.

possibilities  $a$  and  $b$  which arise as independent of possibilities  $\alpha$  and  $\beta$  which arise in some other respect, then a slight acquaintance with the frequencies expected for  $a$  and  $b$  on the one hand and for  $\alpha$  and  $\beta$  on the other suffice to judge what the frequencies will be of their various combinations  $a\alpha$ ,  $a\beta$ ,  $b\alpha$ , and  $b\beta$ . That works even if previous experience had not included any enumeration of those combinations. If we know how many men in Germany marry each year, and how many men are blond and how many dark-haired, it could be reported with some assurance how many blond men and how many dark-haired men marry each year. We need not take a finer reckoning of statistics in the matter. Investigations of this sort count as methods of the probability calculus, because they concern frequencies of combinations. It is clearly evident those same methods call for very careful examination of their assumptions. Later we will allude to calculations (about the lengths of marriages, widowhood status, and the like) which fall into these categories, but which are far from impeccable.

**8.** Finally we pose this question: how much logical significance in the narrow sense is to be given the probability calculus? In other words, how is a strict account to be given of the logical justification of assumptions in domains which do not admit representation by precise probabilities in numeric terms, but which still call for application of the Range principle? Under these circumstances it is only possible to calculate a numeric probability that an expectation would have under some postulates. Either those postulates may be uncertain, or else we know with confidence that they do not hold precisely. We must form a schema to obtain any kind of numeric determination in some respects. We must presume general conditions to be constant, even if we know they wobble more or less. Else we have to consider individual cases as independent, even if we know they are not completely independent. That much is unavoidable: we cannot simply pin numbers on whatever deviations from these schema reality may offer. Under these circumstances it is not the calculations which present questions of real interest for us. Rather, calculations concern schematized examples (so to speak) which can be of use in illustrating how we deal with real problems. It is also quite impossible to go beyond these means in exploiting the probability calculus, by the nature of such cases. There is no prescribed way of drawing a boundary to show where those means are of service, of course. Simple deliberations can provide us a reasonable orientation. Under these circumstances the probability we attribute to some premise can be assessed from a couple of viewpoints. First, we have to know how large its numeric value would be under the schema (as we may say for short). Second, we have to extract the degree of correspondence between the schema and the case in question. Our interest in numeric evaluation is lost to the extent that this correlation becomes negligible. Of course the measure of that is simply described as a ratio. If the relevant premise seems especially unreliable because of a lack of correspondence between actual ratios and the schema, but the numeric probability this produces is large, then precise evaluation of probability is of little or no interest. When several dubious assumptions stand behind a premise, it must be of primary importance to orient oneself to the degree of certainty attributed to the one least certain. Analogously, in order to be aware of the precision of an experimental procedure – as in chemical analysis – it is essential to consider the sources of error prone to produce the largest errors.

Whether application of the probability calculus is of significant interest in those cases, depends essentially on doubt. It depends whether the relevant assumption is placed into question, either by a negligible measure of probability produced by the schema, or else by a lack of correlation between actual ratios and the schema. In a general sense it is important to calculate those numeric values, once we see that only moderate values of probability are produced. It will be superfluous if we can tell even in advance that numeric probabilities are very large, though the assumptions of interest remain relatively uncertain. In such situations it is absolutely necessary to have an overview of the gross magnitude of the pertinent numeric probabilities. It can be superfluous to perform each and every tabulation and calculation for all probabilities – even misleading. The proportion of randomness that can be accounted for may often be the sole source of error in such investigations, just as errors in the weighing of substances may be in chemical analysis. If we were told to use a very faulty scale for weight, some acquaintance with the precision of the scale would be required to say how exact an analysis would be. But say we knew the scale's precision was wholly sufficient to the purpose. – Say we knew that the error which might arise by imprecise weighing could be neglected entirely, compared to errors involved in analytic procedures. – Then it would be equivocal whether the probable error of weighing was somewhat larger or smaller. Although it is obvious we have to be familiar with the precision of our procedures in weighing, it would be wrong-headed to demand that probable error in weight must be appended to the result of every analysis in order to assess its reliability. Accordingly it may be expected – we will have opportunity to confirm this later – that very many domains of phenomena rely on a schematic treatment of this kind. It may be that in performing calculations of probability for assumptions which have yet to be established, or conclusions which have yet to be drawn, such calculations are of little use to anyone. Suppose they are still demanded or attempted. Then very often such an effort may be found to rely on the salient mistakes which have been reviewed already, by which all series of phenomena of the same kind are considered to represent the same kind of thing as ordinary games of chance. The numeric values which are calculated would be mistakenly held out as probabilities of real importance, probabilities which might actually be assigned to premises or to conclusions.

**9.** If our investigations seem to allow the numeric representation of probability only in an extremely restricted kind of case, this provides us an opportunity to make a few additional remarks. First, what one may call logical methods of the probability calculus represent procedures of entirely universal application. Those methods include the partition of possible phenomena into numbers of separate cases, as well as the establishment and exhaustive enumeration of possibilities – in short, the systematic development of disjunctive judgments. This logical form leads to numeric outcomes only by the conventional assignment of equally possible cases. Consequently this is the material basis of any result produced by the probability calculus. Such a result is only produced in an immediate and reliably illuminating way for cases we have described as characteristic. So these domains are fenced off as the only ones in which we can assign numbers to probabilities.

The question may be raised, if there cannot also be numeric representation of quite arbitrary probabilities, ones that may be compared with probabilities which are authentically numeric in nature. It seems conceivable one could estimate such a probability, by reporting a numeric probability that seems to have the same degree of certainty. Of course there is not much to argue against those rough kinds of assessment, but it may be necessary to pay attention to their meaning, and to the difficulties that underlie them. If we estimate the value in numbers of a conclusion by analogy, and we report a probability of  $\frac{5}{6}$ , then relations between that analogy and the unrestricted formation of expectation for ranges in a  $\frac{1}{5}$  ratio, show completely heterogeneous things in logical terms. By their nature such heterological relations cannot be compared. The point on which these two may be compared is purely psychological. What is compared is the psychological surely found in one case and in the other. That is the only thing common to both.

So this appraisal seeks to unite two logical relations: to compare them by psychological effect. Such appraisals become more difficult, just because such psychological matters are irregular and variable. As far as I can see, they may prove helpful in only one situation. If a complex argument leads me to a conjecture which is more or less certain, and I want to communicate the gist of the result to someone unfamiliar with what is involved in the argument, sometimes it will not be too much to express the certainty of the conjecture, and to express it by an estimate in numbers. A doctor might try state his hope or concern for a patient's condition, by giving numbers that are more precise than otherwise possible. Say he estimates the probability of recovery to stand at  $\frac{3}{4}$ , while his estimate of fatality stands at  $\frac{1}{4}$ . Then despite the indeterminacy which necessarily pervades the estimate, it would be more valuable than a few turns of phrase about risk and hope, and so on. But it is very clear that this is a matter of communicating a psychological state, not a generally valid characterization of some logical relation. The estimate dispenses with the advantage one expects to be the most important property of an assignment of numbers to probabilities, which is to avoid idiosyncratic and arbitrary license in judging proportions. When probabilities are not numeric, obviously no methodology alleviates the lack. Suppose clinician X treats 20 cases of rheumatism with salicylic acid, and he believes he can draw a reliable conclusion about its positive effects. At the same time clinician Y thinks this conclusion is risky. Their disagreement in interpretation may rest on this: one clinician believes the relevant cases may be assumed to correspond in other respects to earlier cases, while the other thinks this dubious. Nothing is gained if the first clinician estimates this probability to be  $\frac{19}{20}$ , while the other estimates the same value to be  $\frac{2}{3}$ . Another no less important use of a numeric evaluation of probability may be to dispatch an estimate. Then it can be reported with certainty which probability is to be considered larger and which smaller. This can be important when the business at hand is deciding matters from probabilities, or choosing between alternatives etc. Sometimes an accurate numeric determination of probability can regulate decisions in a way that excludes misgivings. But when it seems questionable which of two probabilities is larger, uncertainty cannot be sidestepped by estimating both. Instead those estimates must be presented in their raw form, in all their uncertainty.

Detailed investigation shows us why such estimates cannot lay claim to significant value. The arbitrariness of assigning what should count as equally possible, was the reason why a numeric treatment of probability is restricted to a few cases. That arbitrariness springs from the undetermined and indeterminate nature of certain logical relations. It is also shown in practice: it is always seems unreportable what degree of psychological surety we are justified in lending any such relation. Consequently a basic postulate also fails for the estimate: that a definite and clearly representable degree of subjective certainty really does correspond to some logical relation. Even with numeric probabilities, there is no definite connection to psychological certainty. At least no clear mental images are left to memory of the different intensities of such a connection. Who would claim he is able to represent facts with probabilities of  $\frac{9}{10}$ ,  $\frac{99}{100}$ , and  $\frac{999}{1000}$  by well-formed mental states? It may be doubted that the arbitrary assignment of numerals to specific premises can have any value for more complex conjectures. I believe this all the less, because in general it is not permissible to estimate the probability of single premises, and then to perform operations on the results by combinatorial methods of the probability calculus. Instead the probability of each combination of premises always has to be assessed anew each time.

It is not to be feared that such attempts at estimation – when conducted rationally – involve the risk of falling into a conceptual trap. That would only happen if one considered such numbers for probabilities to have real meaning – as ratios of ranges. It is not entirely incorrect if one reverts to illusions which have occasionally been entertained when applying the probability calculus – if one reverts to some confluence of the subjective and objective meanings of statements about probability. But the reason for the mistake is not that one places a psychological value on an estimated probability which does not belong. Instead the reason is that one proceeds as if these numbers possessed a real meaning they do not. It is not the estimate which is suspicious, but rather a misapplied judgment about real proportions, and a sketchy parsing of possibilities by which one accedes to statements about probability – statements which are then unsustainable as valid estimates, on stricter examination.

**10.** In the interest of completeness and in the interest of historical significance, it may be justified if I append some remarks on the relation of the theory of probability to rules of practical reason – in other words, to doctrines of rational methods for preference and choice.

In order to answer questions which have been raised in this context, and answer them in an entirely satisfactory manner, just one general insight is required. The insight is that no definite and universally valid principle exists by which we might compare the value of goods we possess to the value of a prospect or the hope of possessing them. If we consider goods with cash value, then for prospects of this kind, an amount of profit may be considered – call it  $A$  – as well as a probability with which we may expect to obtain that profit. Let us think of the last as definite in numbers – and call it  $\alpha$ . It is well-known that a basic postulate of probability theory is that the value of such a hope can be measured as the product  $\alpha A$ . Since  $\alpha = 1$  for whatever we possess with certainty, the product  $\alpha A$  represents the secure possession

which would be equivalent to expected profit  $A$  with probability  $\alpha$ . Accordingly it seems reasonable to pick one of two prospects which offers the larger product of its chance of success and its profit. Now when this rule is applied to specific cases, it does not always lead to acceptable outcomes. Very frequently it leads to results which deviate considerably from what is reasonable. The establishment and maintenance of this rule is facilitated by several circumstances that are likely to foster illusions. Foremost is ignorance of the specific circumstances to which the possibility of measurement is always tied. Another is the standard impression that both the value of a possession and its prospect have to be expressible in numbers. It even may have appeared self-evident that any prospect can be deemed equivalent to some possession whose worth can be safely tallied and precisely stated. If that were true the arithmetic product would be the simplest and easiest determination of value.

A further circumstance may be added, whose explanation is our most crucial task here. Let someone be invited to enter into a game of chance, in which for a stake  $C$  he obtains the prospect of making a profit  $A$  with the chance of success given by  $\alpha$ . Admission to the game means trading the secure possession of  $C$ , against an uncertain but expected possession of  $A$  with probability  $\alpha$ . If such a game is repeated an indefinite number of times, it is clear that entering the game will be advantageous whenever  $A\alpha$  is larger than  $C$ . Given an increasing number of plays, the probability grows indefinitely larger that the total amount of profit exceeds the sum of the stakes. This holds once  $A\alpha$  is not much larger than  $C$ . Once accepted there is no room for doubt that under these premises, it can always be stated with assurance whether the game is to the player's advantage or not – at least, with an unbounded number of plays. Only when the relation between stakes, profit, and chance of success is standardized in a quite specific way – so that  $A\alpha = C$  – can it be claimed reasonable for the bank or for the player to enter the game. Any deviation from this balance disadvantages one or the other. Here the estimate – taken as the product of the actually equally valued stake  $C$  and the profit  $A$  to be expected with probability  $\alpha$  – is confirmed if  $C = A\alpha$ . This seems to make sense – but just here it is important to note the proximity of a trap – that if entering a game often is reasonable and advantageous, then entering the game once must be too. This does not take into account that an oft-repeated swap which is not advantageous once might finally become advantageous, or vice versa. It is advantageous to play a game one million times, when I obtain a  $\frac{1}{1000}$  chance of success of winning 1001 thaler coins for a stake of one thaler coin each time. It seems to follow directly that on a single play which has the prospect of winning 1001, which counts as having a probability of  $\frac{1}{1000}$ , that one play is worth more than possessing one thaler. But this line of argument proceeds from a completely false premise. It is assumed for two identical values, that their repetition an arbitrary number of times, or their division into equal parts, must produce equal values in turn. That does not take into account the subjective value whether a choice is to be called reasonable or not. Over many trials that subjective value is modified in various ways by the effects that individual outcomes have on one another. The subjective value of a composite object can never be found accurately by adding the values which its parts have in isolation. So it may appear advantageous for me to buy

a house with a garden for 30,000 marks, when otherwise I might decline an offer to purchase the house alone for 20,000 and another to purchase the garden for 10,000, as not being in my interest. Subjective value is determined by context. Similarly it can be advantageous for me to buy a bouquet of roses for 3 marks, while there is no doubt that 30,000 marks represents a higher value for me than 10,000 bouquets. It is easy to say these values are modified over many repetitions. Once I give a bunch or a few bunches as presents, or else have some placed in my room, I might not find any use for the rest. So repeated trials are basic to the subjective value of chances of success too. The possession of a multitude of prospects, where each has a definite and modest probability and where each is independent of the rest, represents a prospect approaching certainty of a unit of profit. The subjective value of many chances of success and that of a single one cannot be compared in the simple way presumed. – The value of the whole is equal to the sum of values of the parts only if each part has value in the sense of a market value. That way an arbitrarily large number of identical examples can be bought or sold for a price without difficulty and without default. But then without knowing about the whole class, it is always advantageous to buy below market price, and disadvantageous to buy above market price. Subjective value would no longer be considered, because of the completely unfettered possibility of sales turnover. If every chance of an arbitrary profit had a market price in this sense, that chance could be nothing but that given by the product of profit and chance of success. Any deviation would afford the possibility of unbounded increase in wealth, either insofar as a player might pay for chances of success, or insofar as the banker sets the rules to sell chances of success. In practise these chances have as little a free market as anything else does, since subjective considerations always decide whether purchase or sale is advantageous or reasonable. Under these circumstances, proportions of value do not remain fixed as objects of value are multiplied. Many examples show how this holds for values assigned to chances of success. Whoever insures his property against fire damage always makes a bad deal, judged in schematic terms – though it is generally agreed this is reasonable nonetheless. That example shows clearly how different assessments are given for many cases than a collection of singletons. If someone – for example a railway company – owns very many houses, it is generally agreed to be more reasonable for the company not to insure them. – Further, assume someone has a fortune of 100,000 marks, on which he lives. Assume he is offered a wager where he stands to lose all his assets, perhaps a bet with great probability, say  $\frac{99}{100}$ . Common sense says he would be foolish to take the wager. It would be the same if, by the rules of the wager, the profit to be expected had a probability of  $\frac{1}{100}$ , whether the profit were 10 million marks, or ten, or a hundred, or a thousand times as large. By contrast, suppose there is the possibility of repeating the wager under identical conditions, say 100,000 times in rapid succession. Then it just might be acceptable if the prospect of a profit of a little more than 10 million were dangled.

The difficulties to which estimates of value as the product of profit and chances of success can lead, are difficulties which find colourful expression in the so-called St. Petersburg problem. This famous *crux* of probability theory can be discussed

briefly. Here is the situation. Peter tosses a coin, and he repeats this as often as it takes for the coin to turn up heads. That ends one round of the game. Peter engages to give his opponent Paul:

1. ducat if the coin turns up heads on the first toss
2. ducats if it turns up heads on the second toss

In case the coin turns up heads only on the third, the fourth. . . or on the  $n^{\text{th}}$  toss, Peter gives Paul 4, 8, . . . or  $2^{n-1}$  ducats, respectively.

The question is: what stake should Paul wager – how much should he pay for a prospect of profit in the game Peter sets? Now the probability heads will turn up on the  $n^{\text{th}}$  toss – that is, that tails turn up  $n-1$  times before heads – is  $\frac{1}{2^n}$ . What is offered to Paul is an infinite series of possibilities of winning, where profits are ever larger and the probabilities of obtaining them increasingly small. If we draw out the series, initially the chances of winning an amount of 1 are  $\frac{1}{2}$ . Following the rule, this represents a value of  $\frac{1}{2}$ . Later the chances are  $\frac{1}{4}$  of winning 2 ducats, which produces a value of  $\frac{1}{2}$  once more. In general, the chances of winning an amount of  $2^{n-1}$  are equal to  $\frac{1}{2^n}$ , which results in a value of  $\frac{1}{2}$  again.

Each member of the infinite series represents the value of half a ducat. Following the rule, it would be reasonable to pay any large sum whatever for a stake in the game. Unbridled consideration of the problem tells us the opposite. For a middling number of tosses, one tells oneself there is the highest probability heads will still appear once. Assuming this might occur on the sixth toss, the profit would amount to 32 ducats. That means with a stake of 33 or more, even this somewhat likely case would result in a loss. Suppose one believes that the coin might turn up tails nine times in a row. That is already an unheard-of stroke of luck, but with a stake of more than 512 ducats, Paul would still suffer losses. With a stake of a few thousand, one might say a loss could safely be predicted. On close consideration, the behavior of value is the same here as was discussed earlier. There is no error in the way the chances of winning are determined. We are easily convinced that if we think the game is repeated indefinitely often in finite time, any finite stake set by Peter is to Paul's detriment. What is characteristic of this game is an enormous sum of profit and the correspondingly minimal chance of success. The possibility of profit increases and the chance of success decreases in geometric progression with progression of the series over the course of the game. Values estimated as the product of the two always remain the same. When initial stakes are high, repetition is required to make the game into one that is somewhat likely to be advantageous: numbers of repetitions which can exceed imagining. Naturally our considerations are restricted to one instance, or a number of instances which may be called small in comparison to a number that offers any significant certainty. Then if the game seems to be decidedly unreasonable for any high stakes from Paul's perspective, then this shows the minimal chance of obtaining an enormous profit cannot be equated to the certainty of a middling fortune – by the measure of the product of profit and chance of success. It becomes less and less true they can be equated, the smaller the chances of profit and with that the more certain a loss.

Undoubtedly there is a contradiction between what is reasonable and what is estimated by the product of terms in these examples. It is just as clear otherwise that this means of estimation – applied to a sufficiently large number of equal prospects of profit – should always deliver accurate determinations of value. Consequently we might be able to deduce the general proposition from which we began, in a completely rigorous way. Since it shows – to put the point very generally – that if  $U$  stands for an unknown possession, and  $S$  stands for a secure possession, and  $U$  and  $S$  do not have the same value in general, nevertheless the simultaneous possession of many  $U$  can represent the same value as the same number of  $S$ . It follows that the value of prospective profits cannot be counted the same as the certain possession of things, in order that the equivalence of  $U$  and  $S$  is valid for any arbitrary combination. After we have attained this perspective, we no longer require proof to show that the logical significance which numbers have when assigned to probabilities, is not to be compared to a similarly definite pragmatic significance. Perhaps one will admit that no directly evident and precise general rule can be established for choices made between what is certain and what is uncertain. It does not seem fitting to pursue this point further, or to delve into a deeper rationale for this intuition.

# Chapter 8

## Application of the Probability Calculus to Theoretical Physics



**Abstract** The results of earlier investigations of the range theory of probability are displayed in a new context: the kinetic theory of gases in physics. The Second Law of thermodynamics is considered as a statement of probability. Boltzmann's mathematical investigations are used as a starting point in the theory of gases. The presuppositions of this probabilistic theory are outlined: probabilities are already measured in terms of ranges. This application of the theory of probability is not a new or immature hypothesis in physics. The present account contradicts the notion that states which are less probable will transition to states which are more probable. The most probable state of a gas at rest is really an extraordinary number of different states. This account is consonant with Maxwell's distribution for the speeds of molecules in a gas at rest (in an adiabatic flask).

**Keywords** Second law of thermodynamics · Kinetic theory of gases · Carnot's theorem · Maxwell's distribution · Boltzmann theory · Normal state of a gas · Entropy and randomness · Equilibria

1. Although in the general arguments to this point, there has been ample opportunity to take a position on many specific applications of the probability calculus, still it ought to be useful to display the results of our theoretical investigations in context. At least it ought to be useful with respect to the most important domains in which such applications have been attempted. It may be necessary to repeat some points in the process. I would ask the reader to forgive this on the grounds that explication of the general theory by citation of examples should not be neglected. The complex relations inherent in many of these domains might not have found adequate treatment before, either.

We may pass over examples drawn from games of chance, since we have investigated them already for the purpose of developing the theory. Instead what is to be discussed are applications of the probability calculus to mathematical physics, to the theory of errors in observation, and to doctrines of the collective phenomena of human society. We will also discuss applications to decisions about therapeutic efficacy based on medical statistics, and applications to judicial decisions. Of course we will not have begun to exhaust the subjects which people have attempted to

handle by methods of the probability calculus. It should suffice to present results in these principal domains, since on that basis it will not be difficult to gain perspective on other domains as well.

**2.** It is well known that certain areas of physics have a close connection to the probability calculus. Namely those areas are the kinetic theory of gases, and the part of mechanical thermodynamics which concerns the so-called Second Law: **Carnot's** theorem. I should like to review these subjects – subjects which have not usually been presented in systematic expositions of the probability calculus. I should like to do so all the more since (as I believe) an excellent illustration of our theoretical and logical notions can be made in these somewhat simple domains, and the relevant physical theories are even illuminated by such consideration. Thus they are brought closer to fully satisfactory implementation. It may be useful to foreshadow the result which will be reached. A theory which seeks to tie regularly-observed phenomena to considerations of probability, bears the task of proving that the relevant phenomena are a majority of possibilities, that is, they represent the largest part of original possibilities of behavior. The theory should demonstrate that those phenomena are always expected with maximum certainty under identical general conditions, in the absence of detailed knowledge about the situation. The theory has to make certain regularities of events intelligible in a similar way, as the explanation we gave in Chap. 3. That explanation cast light on regularities for many cases over total results in games of chance. That investigation has an advantage, insofar as it can attempt to provide more than a summary account. It can provide a detailed deductive treatment of the subject. Yet it has the disadvantage that it must proceed on more-or-less hypothetical premises. Consequently it can only provide an explanation of phenomena in the sense that it traces them to fairly reasonable assumptions.<sup>1</sup> Accordingly the core of this physical theory lies in the calculation of ratios of original ranges of behavior. This concept – particularly the fundamental criterion of originality – has not been described clearly in the theory to this point. It is clear that the mathematical investigations which **Boltzmann** has offered, contain nothing but a part (still, the most important part) of the proof demanded by a logical theory of probability. Those investigations must be interpreted in a way which departs somewhat from their literal wording. As I would like to emphasize, I do not presume to judge how this aligns with the authors' actual opinions or by how much. It seems to me its alignment is much more easily imaginable since here we have the case alluded to earlier: in absence of a sufficient theory for the probability of objective behavior as it should be expressed, no completely adequate description may be found. Those proofs are not quite complete, and so the theory should not count as finished. Indeed this view corresponds to one expressed many times as a desideratum by **Boltzmann**. We will need to formulate his view in a slightly different way. Overall it follows from our examination that as a rule, the standards we have set for the probability calculus in this domain are standards maintained strictly and exactly – without having a sufficiently precise formulation. This situation is understandable, because when

---

<sup>1</sup>Cf. Chap. 7, Section 3 on the meaning of such explanations.

one wants to begin to study a new subject and to apply arbitrary conventions about probability, soon enough one may arrive at results whose worthlessness may be recognized at a glance. For example, let us imagine the problem set by the kinetic theory of gases is altered, so that repulsive forces are attributed to molecules, which act at a finite distance. If we were then to say that the probability is equal that any molecule is contained in an arbitrary and equal part of the whole space enclosed, that would be an inadequately grounded and entirely perverse convention. It can hardly be overlooked that every convention about probability stands in need of a particular foundation, which makes reference to postulates about laws of cause and effect. As a matter of principle, one may consider that such a foundation will almost have to coincide with that which we found necessary, whatever point of view one takes.

**3.** Here and elsewhere, instigation for linking some phenomena to considerations of probability – apart from an exceptionless regularity which they have exhibited in previous experience – is that a necessity of some kind seems inconceivable. It must be emphasized that this inconceivability, and with it a compulsion to categorize relevant propositions as propositions about probability, is based on some general assumptions. Those are assumptions to which one is inclined to attribute a large degree of inherent probability. Firstly, this is the assumption that all material bodies are composed of individual material particles. Especially for gases, that means the framework is adopted which one usually calls the kinetic theory. Secondly, one assumes that forces which material particles exert on one another at any moment, always and only depend on their position at the time, but not on the speed or direction of their motion. In other words, forces are to be considered as functions of coordinates of the positions of particles. They are not also to be considered in terms of the differential quotients of those coordinates over time.<sup>2</sup> Under this postulate (whose inherent probability one may dispute) freedom from necessity follows for certain connections of events which are observed to be completely regular. – If we leave a gas to itself in a space bounded by solid walls,<sup>3</sup> we observe that after the passage of a certain interval of time, gas fills the whole space in equal measure. That means unit volumes everywhere inside it contain equal quanta of gas, and that the same pressure and the same temperature obtain everywhere. Consider a perceptible quantity of a gas for which there is always an enormous number of individual molecules in motion which collide with one another and with the walls of the container. Then this fact could be reported: for gases left to themselves over a long period, equal numbers of molecules with the same average kinetic energy are

---

<sup>2</sup>Incidentally, it may be shown that modern attempts to base a theory of matter solely on contact forces (and in place of the atomic theory, to base it on an assumption about paths of motion through a frictionless fluid which permeates space evenly) lead to the same consequences.

<sup>3</sup>Here I note that strictly speaking, this premise involves a bit of fiction; we cannot “leave the gas to itself”, that is, we cannot isolate it from all external influence. Rather, the actual situation is that it is in an equilibrium of temperature with its container. However, it is expedient to cling to that fiction temporarily, since the theory of this ideal case must be treated on its own merits, and so we may consider the gas to be contained in an adiabatic flask against which molecules may rebound with no loss of force.

contained within the same volume of space everywhere – to a good approximation. In apparent contradiction to the regularity with which this occurs, it may be shown that given the postulates cited earlier – under which any arbitrary distribution of molecules in space is possible at any time, and any arbitrary distribution of total kinetic energy is possible too – all manner of arrangements are possible for their speeds, in both magnitude and direction. None of those arrangements is excluded by the condition that the gas remains without external influence for a long period. The proof is quite extraordinarily simple. Suppose that a series of such states of the gas is indicated by  $a b c d \dots n$ , as they develop by the laws of cause and effect, so that  $a$  transitions necessarily to  $b$ , and  $b$  to  $c$ , and so on. From each of these states we can imagine another corresponding state arising – the ‘reverse’ as one may say for short. For the reverse state all speeds are conferred in opposite direction, while all else remains unchanged – namely all locations of particles and all magnitudes of speed. The new series of states obtained – call it  $\alpha \beta \gamma \delta \dots \nu$  – represents a development which follows the laws of cause and effect (though in the contrary direction) according to the postulate about force cited previously. Then  $\nu$  transitions to  $\mu$ ,  $\mu$  to  $\lambda \dots \delta$  to  $\gamma$ ,  $\gamma$  to  $\beta$ , and from there to  $\alpha$ . For a gas left to itself, any arbitrary state must transition to any other state after some period of time. According to the correspondence just established, for any state which exists at a given instant, there exists another with the property that it transitions to its reverse after a definite period of time. If we ask which state must have been present at time  $t_0$  in order to produce state  $\alpha$  at a later time  $t_1$ , then we may think of state  $\alpha$  in the opposite sense. In this reversal,  $a$  transitions to state  $n$  over a temporal interval equal to  $t_0 - t_1$ . The desired state is the reversal of the earlier state, which is just  $\nu$ . From this it follows that any arbitrary state could be effected at any time. The state of the gas at the moment one left it to itself would be entirely determinate. Therefore the realization of specific states is not conditioned necessarily by lengthy isolation (even an infinitely long one) from external influence.

Then one has to rely on assuming that only very special initial conditions can lead to a non-uniform distribution of mass and nonuniformity of pressure and of temperature at a later point in time. One has to assume the predominant and largest range of states is that which gives rise to the phenomena which are always observed.

In a similar way we may – again under the previous postulate – show that even the Second Law of Thermodynamics<sup>4</sup> is to be considered a statement of probability. In

---

<sup>4</sup>For those unfamiliar with the subject, I believe that the import of this statement is most readily rendered intelligible if I allude to its specific consequences: for example, that with contact between a warmer body and a colder one, heat always transfers from the former to the latter, but never the reverse. That is, it never happens that the warmer body becomes even warmer and the colder one colder. Further, it never happens that the kinetic energy – which is disarrayed and takes the form of heat – can be transformed into some directed motion in the form of translation or rotation (for bodies of finite dimensionality), and so forth.

that respect I would refer to the argument that **Boltzmann** has made in response to an article written by **Loschmidt**.<sup>5</sup>

These sorts of considerations show that we have reason not to hold some connections between phenomena to be necessary, despite their regularity. The familiarity which we can gain with actual behavior in these subjects is only ever general, since details of the arrangements and movements of molecules lie outside our capacity to perceive them. That holds for gases, as for anything that concerns the exchange of heat between bodies in general. So we can say these considerations of probability are entirely in order.

**4.** Physicists in general, but **Boltzmann** in particular, have added a special twist to investigation of the kinetic theory of gases. They do not seek to make the previously cited phenomena (equilibria of pressure and of temperature) interpretable mainly through consideration of probability, since they do not seem to require any deeper justification than is given by a few very general remarks. They proceed to conduct investigations of something that may not be controlled experimentally in a direct way: something called the distribution of speeds for a gas at rest. As may be evident, various molecules have different speeds; let us take it as specified how many have speeds which lie between 0 and  $\delta$ ,  $\delta$  and  $2\delta$ , or  $2\delta$  and  $3\delta$  etc. at a given time, where  $\delta$  stands for a small measure of speed. That is the task of finding their distribution of speed at some time. According to **Maxwell**'s proposition, one whose central importance does not need to be explained, just as a uniform distribution of mass in space and an equilibrium of pressure and temperature is achieved, so a very specific distribution of speed is achieved over some time in a gas left to itself. This is a distribution where the number of molecules whose speed lies between  $c + dc$  is proportional to  $c^2 e^{-hc^2} dc$ . Here  $e$  is the basis of natural logarithms, and  $h$  is a particular constant determined by the total kinetic energy contained in the gas, and the number of molecules (that is, by average kinetic energy over all molecules). Proving this distributional law has been the principal task of a large number of articles. Its content is what we are concerned with here, since statements which are made about equilibria of pressure and temperature are connected to the same dispositions which do double duty as initial premises for the proof.

Since we are concerned to examine the content of this physical theory, it is advisable we begin by getting rid of an apparent contradiction in the context of our discussions about logic. We have established the proposition that the probability of any behavior is necessarily identical with the probability of the prior conditions under which the behavior is bound to occur (p. 34 in original; Chap. 2, Section 3). Then in order to determine probability, it follows that ranges must be original in nature.

We can express this by saying: the probability we assume the present existence of a certain behavior  $X$  to have, must always be equal to the probability which we

---

<sup>5</sup> Boltzmann, L. (1877). Bemerkungen über einige Probleme der mechanischen Wärme-Theorie. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe*, 75(2), p. 67.

expect any subsequent behavior to have, if the present occurrence of  $X$  is the necessary and sufficient condition for that subsequent behavior to be realized. This assertion – which is the evident basis of any logic of probability – seems to stand in contradiction to the proposition current in physical theory: that as a rule, what is less probable transitions to a more probable state. The contradiction is easily resolved, once we examine what this proposition really means. Let us continue with an example: a closed space is filled with a certain gas so that one half of the space contains more molecules in motion, or more fast-moving molecules, than the other half. When one attempts to show that these differences tend to equalize over time, one always proceeds in the same way. Speculations about probability are made about the distribution of molecules and their speeds within each half, and also about collisions and changes of state. This results in an approximation to a uniform distribution being the distribution considered most probable. Strictly speaking, given the indeterminacy which is stipulated even for the initial state, arrangements that have the largest probability are just those constituted to transition to a uniform distribution. If the physicist characterizes this as a transition from an ‘improbable’ state to a ‘more probable’ one, it is clear he does not have in mind an expectation justified by some epistemological state. Rather he takes the word ‘probable’ in a different sense: an absolute one. That can be stated simply: it means that probability which arises given complete ignorance of the distribution of the gas. What comes to the same thing, it is the most probable state among all those which represent the largest class of original varieties of possible behavior. It is not very useful to speak simply of “the most probable state” of a gas enclosed in a space, since a basic postulate is not made explicit. The basic postulate is that the meaning attached to this expression is just that which it has in ordinary speech. Hence in what follows, I call a state which (under certain general conditions) counts as most probable in this sense as the normal state. – So it would be postulated in advance, that for a gas contained in an adiabatic flask, the normal state consists of a uniform distribution of mass, temperature, and pressure. In the same sense, equality of temperature could be considered the normal state of two such flasks that come into contact. This may be expressed by saying that even with indefinitely more knowledge about the behavior of a gas, the most probable arrangements are always those which transition to a normal state. That is true at least insofar as our knowledge is indeterminate rather than completely exact.<sup>6</sup> What is yet to be noted is that the normal state – what physicists call the most probable state – is not one well-defined state, but rather the embodiment of an extraordinary number of different states. – Thus the proposition we have reached departs from the one used in physics – the one from which we had

---

<sup>6</sup>If one bears in mind that within certain general conditions the normal state is that state which is most probable given the total absence of all exact knowledge, and that there really always is such a lack with respect to the determinants of the tiniest particles, then the confident expectation of a general and enduring approximation to the normal state may be expressed as follows: that everywhere within a finite number of dimensions, just that variety of behavior is to be expected, which we assume for the current states of the smallest particles. In fact, this proposition represents the final result of all investigations involving probability in the field of molecular physics.

derived the proposition at hand by a slight change in expression. It still departs from the usual proposition insofar as it does not posit that the transition from an atypical state to the normal state is necessary, but only the most probable behavior under all circumstances. That point will be amplified in what follows.

The principal requirement to which we may cling without coming into contradiction with the laws of physics is easily formulated in mathematical terms. It provides an essential condition which must be adequately addressed by any assignment of probability. Let us designate all the elements which determine the momentary state of a gas as  $r_1 r_2 r_3 \dots$ , meaning the coordinates of all the particles, and the vector components of their speeds. Then the probability that the gas exists in a particular state at time  $t$ , a state is determined by the values  $r_1 r_2 r_3 \dots$  with ranges  $\Delta_{r_1} \Delta_{r_2} \Delta_{r_3} \dots$ ,<sup>7</sup> would be expressed as the function

$$f(r_1 r_2 r_3 \dots \Delta_{r_1} \Delta_{r_2} \Delta_{r_3} \dots t).$$

This function must have the property that  $\frac{df}{dt}$ , the overall differential quotient of the function over time will vanish. That quotient is equal to

$$\frac{\partial f}{\partial r_1} \cdot \frac{dr_1}{dt} + \frac{\partial f}{\partial r_2} \cdot \frac{dr_2}{dt} + \dots + \frac{\partial f}{\partial \Delta_{r_1}} \cdot \frac{d\Delta_{r_1}}{dt} + \frac{\partial f}{\partial \Delta_{r_2}} \cdot \frac{d\Delta_{r_2}}{dt} + \dots \frac{\partial f}{\partial t}$$

on the condition that values are substituted for the following terms:

$$\frac{dr_1}{dt}, \frac{dr_2}{dt} \dots \frac{d\Delta_{r_1}}{dt}, \frac{d\Delta_{r_2}}{dt} \dots$$

Those values are produced by the relevant laws of cause and effect, when the state of the system is known.

**5.** After these preliminary remarks, let us turn to study the normal state of a gas, as found in **Boltzmann's** work. We draw attention to the stated goal of his investigations. In general terms, his goal is to demonstrate that a certain state is ‘possible’, namely the state characterized by **Maxwell's** proposition. But this expression is interpreted further, to say that such a state – once present at some time – remains unaltered by the collisions of molecules, and by the progress of their motion.<sup>8</sup> What is meant by a ‘possible’ type of distribution is always one possible in the long term: a

<sup>7</sup>I count a state as determined “by the value  $r_1$  with range  $\Delta_{r_1}$ ” if the value in question falls between  $r_1$  and  $r_1 + \Delta_{r_1}$ .

<sup>8</sup>So for example, in the *Wiener Sitzungsberichte* etc. vol. **63**(2), p. 400: “a value of function  $f$ , which is not further altered by the collisions of molecules, and consequently represents at least a possible type of distribution for states of molecules ...”; also vol. **78**(2), p. 41 and in many other places. [Boltzmann, L. (1871). Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe*, **63**(2), 397–418. See also Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen, 2. Band*. Leipzig: Barth (F. Hasenöhrl). Reissued 1969 – New York: Chelsea.]

stable distribution. This should not be considered as if a change in the state in question – meaning a transition to another state not corresponding to Maxwell's proposition – should be impossible. Rather, it is irrefutable that just as deviant states may transition to the normal state, so a distribution characterized as normal may also transition to a deviant one. This illuminates the conditions of reversal described earlier, when one adds that if a state represents the constraints of the normal state, then of course its ‘reverse’ always does as well. In the mathematical development of the subject, if there is the appearance that as soon as a normal state arises then any other state is impossible, then it is certain that some misinterpreted expression has led to a conceptual mistake.

It seems to me that a vague expression of this kind really is present. It is always present at a very particular point in the investigation. The mark of a state is assigned to be the number of those molecules for which the coordinates and the vector components of velocity take on certain values ( $x_1 x_2 x_3 \dots, u_1 u_2 u_3 \dots$ ) that have very small ranges ( $dx_1 dx_2 dx_3 \dots, du_1 du_2 du_3 \dots$ ). Let us consider this number to be given by the function

$$f(x_1 x_2 x_3 \dots u_1 u_2 u_3 \dots) dx_1 dx_2 dx_3 \dots du_1 du_2 du_3 \dots$$

Evidently the stability of the state is given by the condition that this value remains constant. It is constant if individual coordinates and speeds are altered in ways which respect the laws of cause and effect. Suppose  $a$  and  $b$  are two states characterized by different values of  $x$  and  $u$  and by their ranges. Then stability should be thought of in this way: at any time just as many molecules transition from state  $a$  to state  $b$ , as the reverse. Then if we take the explanation of function  $f$  to be literal, it is obvious there is no specification which satisfies the requirement. There is really no way to shift another molecule into place just as one molecule vacates its position. A function which characterizes the state of a gas precisely – as this does – can never remain constant. Construed in this way, the function would be discontinuous, inasmuch as a molecule would either be present in this state or not. The function would always have a value of 1 or a value of 0 for specific coordinates and speeds. Once we consider that the function  $f$  is discontinuous, in the literal wording of the explanation, obviously it would be impossible to determine the function in a way that satisfies the requirement of constancy. The proof can actually be given that there is a specific form of the function  $f$  which possesses this property. That depends on the function being treated as continuous, but this stands in contradiction to the interpretation it is given, so it becomes necessary to modify the interpretation. Clearly this needs to happen as follows: we consider function  $f$  as the product of probabilities over the total number of molecules, for which we assume (for every molecule) that a molecule's state falls within the range determined by coordinates  $x_1 x_2 \dots$ , speeds  $u_1 u_2 u_3 \dots$ , and their derivatives  $dx_1 dx_2 dx_3 \dots du_1 du_2 du_3 \dots$ . At once it is clear that this fresh interpretation almost coincides with the original, for large numbers of molecules and for such ranges as span both finite portions of the enclosing space and finite ranges of speed. Suppose that for a very large number  $N$  of molecules, it is

assumed with probability  $\frac{f}{N}$  that the state of each molecule falls within a certain range. By the Law of large numbers it is expected with maximum certainty that the number of molecules for which this holds is very nearly equal to  $f$ . This number is given by the integral of the probability function  $f$  over such a range.<sup>9</sup> Further it may be seen that under this reading of function  $f$ , the discovery of a form for  $f$  which satisfies the condition of constancy, can be given a very important interpretation, one which can be described precisely. What can be found is an assignment of probability which satisfies the principal requirement above: one which ascribes equal probability to varieties of behavior which are necessarily connected to one another as antecedent and consequent. Since here probabilities are measured in terms of ranges, this may also be expressed as follows: that a means may be found to determine the sizes of ranges, which has the property of deriving original ratios. We arrive at an assignment of probability which is authorized under the laws of cause and effect attributed to molecules. It is important that this is independent of time. For any arbitrary point in time, an unbounded duration may be brought to bear. As a consequence we ought to expect certain characteristics of behavior with maximal probability at any point in time. These may be inferred directly from the form which the function  $f$  is found to have. These characteristics consist of a uniform distribution for mass and for kinetic energy on the one hand, since the function is coordinate-independent. On the other hand it consists of **Maxwell's** distribution of speed, as the function of speed which corresponds to those characteristics.

The condition that the function  $f$  remains constant, leads in an accepted way to establishment of this finding: in general it is sufficient to determine that the probability of a certain type of behavior for two molecules is unchanging if one mode transitions to another by collision. There is not the slightest difficulty in understanding this condition under the present interpretation. For example, if we hold to the simplest, most primitive account of the situation (one **Boltzmann** gave in volume

<sup>9</sup>The probabilities which are to be assigned to individual molecules would not be independent of one another, just insofar as the location of each one excluded the possibility that another would occupy a position within an exceedingly small element of volume. That would not affect the result established here concerning the postulate about the size of the range. More importantly, this independence of the state of each individual molecule does not hold in connection with the constancy of the total kinetic energy present. This circumstance implies that the representation which we obtain for the probability of any state is in fact not strict in a formal sense, since the probability  $f$  has been assigned to each molecule, and all those probabilities are treated as independent. At most, these probabilities would be valid given infinitely many molecules. Hence it is easy to see that for a finite number of molecules, and considering the convention about probability for the conservation of energy (as given for example by **Boltzmann** in the second part of his article contained in volume 58 of the *Wiener Sitzungsberichten*), this is derived in just such a way as necessitated by the interpretation given here. It is known that the same result obtains once the number of molecules becomes very large, and those results do not depart significantly from the results one obtains by neglecting their dependence. For this reason, I believed I ought to restrict myself to the much simpler formulation given above. [Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe*, **58**, 517–560.]

58 of the *Wiener Sitzungsberichte*), one can adopt a wholesale accounting of the collision of two molecules. One can formulate the conditions to be fulfilled, by saying that the same probability is attributed to the state following collision as is to be attributed to that preceding collision. So this all comes down to a difference in words, if against **Boltzmann**, we require that probabilities (or ranges) or that numbers of molecules remain constant. The results are the same in mathematical terms.<sup>10</sup> This situation also holds in the case described in the second part of the same article, which takes into account the contingency or non-independence introduced by a constant amount of total energy, given a finite number of molecules.

---

<sup>10</sup>One can make the meaning of this calculation more intuitive in the following way. If a point-mass continues straightforward in uniform motion, we can gauge the probability that at a certain point in time it may be found in a unit of space  $v$ , and its direction of motion may be found within the solid angle  $\Delta$ , by the integral  $\iint dv d\Delta$ . Then this integral remains constant, if we concentrate on the totality of states to which it may transition at any possible time. Then this represents the size of an original range. One sees this most easily for infinitely small ranges which are determined by two plane sections  $dF_1$  and  $dF_2$  which are perpendicular to the direction of motion, and by such a range of directions of motion where it would be possible to determine the direction of each point of the element  $dF_1$  by each point of  $dF_2$ . As one sees at first glance, at the beginning of the time period in question, the operational range is given by  $dF_1$  multiplied by the solid angle which  $dF_2$  appears to have from the perspective of  $dF_1$ . At the end of the time period, it is given by  $dF_2$  multiplied by the solid angle which  $dF_1$  appears to have from  $dF_2$ ; these two products are equal to one another. Without the influence of external forces, the range measured in this way is not changed by progressive motion. In the same way, by the ideal law of elastic impact, it is not changed by reflection against a solid wall (i.e., equality of angle of incidence and angle of reflection, and unaltered speed). If on the other hand, two bodies of comparable masses collide, then the operational range does change, since the distribution of speed does too.



Then one may no longer make a satisfactory assignment of probability without considering speed; rather, the operational range must also span a certain range of speed. At first one might obtain values like  $\iiint dv d\Delta dc$ , where  $dc$  would be a derivative of speed. Let us call the ranges of two molecules specified in this way  $s_a$  and  $s'_a$  before the collision. After the collision let them be called  $s_p$  und  $s'_p$ . Then further investigation shows that  $s_a \cdot s'_a$  is not equal to  $s_p \cdot s'_p$ ; rather these products vary in inverse proportion to the products of the kinetic energies before the collision,  $c_a^2 \cdot c_a'^2$  and after the collision,  $c_p^2 \cdot c_p'^2$ . One of the requirements of the measurement of ranges – to respect originality – is then maintained in that one does not set the ranges of speeds to be proportional to the derivative  $dc$ , but rather proportional to that multiplied by a function  $f$  of speed, which is determined by the equation:

**6.** With that, the theory is still far from complete. A convention about probability has been established, and its admissibility in a specific circumstance has been demonstrated. To argue it really does determine our expectations, some proof is necessary (as the logical theory demands) that what is called equally probable here, also represents equal, indifferent, and original ranges. If we are to be adequately prepared for the possibility of such a proof, it is necessary to specify the problem in more detail. Initially we should focus on the ideal case, as if we had to deal with a gas contained in an adiabatic flask for some long but finite period, about whose present or past states nothing is known apart from its mass and the total value of its kinetic energy.

The link to logical theory is quite uncommonly simple, since the probabilities are already measured in terms of ranges. We saw that the investigations just put forward tell us nothing but that definite measurement of ranges produces *original*, enduring, and unchanged ratios for them. If further, following the postulate, nothing is known but that the gas is enclosed in a defined and unchanging space, the ranges of the various behaviors are also *indifferent*. Under those circumstances, only one test is lacking: that is, to determine if they are also strictly *comparable*. That means the measure introduced may not be arbitrary (meaning the measure by which we call two ranges equally large, and by which we call two kinds of behavior equally probable). Such a misgiving may seem to haunt the measurement of ranges of speed by the value of  $c^2 e^{-hc^2}$ . Still it has not actually been ruled out as a measure, since the principles of originality and indifference really do hold for it. The same internal consistency of an approach to probability would obtain, if we claimed as much for a behavior unchanging over time. If one were to succeed in partitioning the range of such a behavior into indifferent parts, then their ratios could be established in a completely arbitrary way. Each would have the property that we could use it as a measure of probability, without contradicting the requirement of originality. That works simply because no transformations have been effected. Any one of them might be entirely arbitrary, nonetheless. Something similar might happen for our

$$s_a \cdot s'_a \cdot f(c_a) \cdot f(c'_a) = s_p \cdot s'_p \cdot f(c_p) \cdot f(c'_p)$$

or

$$\frac{f(c_a) \cdot f(c'_a)}{c_a^2 \cdot c_a^2} = \frac{f(c_p) \cdot f(c'_p)}{c_p^2 \cdot c_p^2}.$$

If we should call some operational ranges  $\varphi$  – which are ranges gauged by the magnitudes  $\iiint d\nu d\Delta f_{(c)} dc$  – then it would always be that:

$$\varphi_{(a)} \cdot \varphi'_{(a)} = \varphi_{(p)} \cdot \varphi'_{(p)}.$$

gas, if we could assume several quite different varieties of behavior  $A$ ,  $B$ , and  $C$ , each of which had to be maintained over time. Then if  $A$  were present, only a specific category of states could be swapped for one another; and if  $B$  were present only quite another category ... etc.; but the presence of  $A$  would persistently exclude the introduction of  $B$  or  $C$ . Under such circumstances it would be conceivable that for measurement of the probability that  $A$ ,  $B$ , or  $C$  was present, many different assignments might be made with equal justification. Under this way of considering the matter, an assignment of probability might prove to be compelling, under similar circumstances as we have seen for games of chance. The following considerations convince me this does happen. Even here it is not the case – because of the specific manner that states change following collision – that in comparing any two varieties of behavior, we might have two wholly different and hence incommensurable things before us. Quite minimal differences in behavior at one time lead to arbitrarily large differences at a somewhat later time. Very small variations in the coordinates and directions of motion of two molecules before collision give rise to very significant differences in the ways collision may occur – consequently to significant differences in behavior afterwards.

Large differences in behavior at one time may be traced to minimally different states at an earlier time, and vice versa. And from this the comparability of very different behaviors emerges, insofar as the behavior might be traced to very closely correlated behaviors at any time under previously existing conditions. If we proceed on this basis, then the determination of size given by the theory follows compellingly. For example, if  $S_1$  and  $S_2$  are two very different ranges to be compared, and if both are very small, it would be enough to have a compelling way to discover a ratio for the probability of  $dS_1$  (an element of  $S_1$ ) and  $dS_2$  (an element of  $S_2$ ). This does happen if we can claim that at least two infinitely small elements  $dS_1$  and  $dS_2$  can be found, which are states caused by processes  $d\sigma$  and  $d\sigma^1$  at some earlier time, and which both belong to the same very small range  $\sigma$ . The  $d\sigma$  and  $d\sigma^1$  are directly comparable. In this way the ratio of probabilities for  $dS_1$  and  $dS_2$  also emerges without further ado. Since the magnitudes  $d\sigma$  and  $d\sigma^1$  are equal if  $dS_1$  and  $dS_2$  stand in equal ratio, by which the theory attributes them the same probability, then it follows that the assignment of probability is uniquely determined.

The postulate which is the basis of these considerations could be formulated more strictly, as follows: If  $S_t$  and  $S'_t$  are arbitrary states which are well-defined at time  $t$ , at any earlier point in time there may always be found two states  $\sigma$  and  $\sigma^1$  such that  $\sigma$  and  $\sigma^1$  differ exceptionally little in the elements that determine them. They are such that  $\sigma$  would lead to a state  $S_t$  at time  $t$  which is only very little different in the elements which determine it, from state  $S'_t$  – the state at time  $t$  to which  $\sigma^1$  would give rise.

This truth seems free from doubt, because of the laws which govern the collisions of minuscule bodies. If one does think them questionable, it is still conceivable that different varieties of behavior can be specified, for which values of probability may not be compared. Yet if one accepts that much, there remains no doubt that for a gas contained in an adiabatic flask for a long period, whose state is unknown to us, the notion of probability given by the theory is not only admissible, but unavoidable. It may even be the only notion possible.

Therefore the uniform distribution of mass and kinetic energy at any given time, as well as Maxwell's distribution of speed, are to be assumed to hold with maximum probability at any given time.

There can be no concern that "no exact proof has been offered, that as a consequence of the equilibrium in temperature, the average kinetic energy in the neighbourhood of a molecule might be influenced by the state of the molecule itself."<sup>11</sup> The only proof which can be provided, is to the effect that arrangements under which such a relation obtains, represent neither larger nor smaller original ranges than others. That proof has been rendered in full.<sup>12</sup>

Similarly, as the result of investigations of the Second Law – unnecessary to pursue here in more detail – for the moment it may be reported that equality of temperature should be considered the more probable state for two closed bodies which have been in long contact, but which are free from any other influence, and about which nothing else is known.

**7.** Before we leave this ideal case behind and pass on to see what may be required of the theory for other problems closer to reality – we should inform ourselves how **Boltzmann** formulates other strictures in his theory.

According to **Boltzmann**, the proof has yet to be delivered that the state characterized by function  $f$  is the only one possible. To that end, it would not only be necessary that the function  $f$  should be proven to be the only one satisfying the condition of constancy, but (as **Boltzmann** says in response to remarks by

<sup>11</sup> Boltzmann, L. (1878). Weitere Bemerkungen etc. *Wiener Sitzungs-Berichte*, **78**(2), p. 44.

<sup>12</sup> By the way, the premise that was just advanced also leads to the conclusion that there could not be two different and continuous assignments of probability which satisfy the condition of constancy. That is, let us consider two infinitely small ranges  $s_1$  and  $s_2$ , which represent correspondingly infinitely small and very closely neighbouring ranges,  $\sigma_1$  and  $\sigma_2$ , at an earlier time. Then any assignment of probability must fulfill the condition that the probabilities attributed to both states  $s_1$  and  $s_2$  – let those probabilities be called  $\varphi(s_1)$  and  $\varphi(s_2)$  – must have the same ratio as  $\sigma_1$  to  $\sigma_2$ . That ratio is uniquely determined, since  $\sigma_1$  and  $\sigma_2$  are extremely close as neighbours. Then it holds that:  $\varphi(s_1) : \varphi(s_2) = \sigma_1 : \sigma_2$ . And if  $\psi$  were to be a second probability function, then immediately the result follows that

$$\varphi(s_1) : \varphi(s_2) = \psi(s_1) : \psi(s_2)$$

or

$$\frac{\varphi(s_1)}{\psi(s_1)} = \frac{\varphi(s_2)}{\psi(s_2)},$$

which means the two functions could only be distinguished by a constant factor. As is easy to see, whether continuity is here taken in its mathematical sense, or taken in the sense previously described (p. 51 in original; Chap. 3, Section 2), obviously will depend on the assumption one wants to make about the traceability of finitely different contemporary states to a few different earlier varieties of behavior (or for infinitely long periods of time, to infinitely few different earlier varieties). Here I restrict the postulate to the sense given previously, and hence once again I construe continuity in the non-mathematical sense.

**Loschmidt**<sup>13</sup>) “the proof that this is the only possible distribution of states becomes significantly more difficult, since one would also need to show that it is impossible in the stationary state – where the function which determines the distribution of states is still dependent on time – that a small rise or fall in temperature may not occur: now here, now there, or altogether irregularly throughout the gas. That might arise from the motions of the molecules themselves, just as it could vanish on the same basis.”

Here too it seems clear that fully determinate arrangements may be imaginable, which do not lead to irregular distributions of molecules and of kinetic energy. Instead they may lead to periodic and uneven distributions which cycle over an indefinite period. Then we would be dealing with a case analogous to that which **Boltzmann** calls ‘labile equilibrium’. Even this requirement must be made more precise, to say that the convention about probability characterized by function  $f$  is derivable from any arbitrary initial indeterminacy whatsoever, be that ever so small. If one could prove that much, it seems to me irrefutable that the theory would satisfy any requirement one could place on a theory of probability. With that the computational part of the theory would be entirely complete.

But simple considerations show that this proof cannot be developed, at least not without significant modification, as long as we hold fast to the fiction of a gas enclosed in an adiabatic flask: a container whose walls reflect molecules without loss of force. Every constraint on initial conditions which is known to us is by the same token a constraint on the state which occurs an arbitrary time afterwards. If we were to know that an equilibrium of temperature existed at time 0, then all states would be excluded at time  $t$  which would have emerged from an unequal distribution of temperature at time 0. To gain this insight, we need only recall that states at time 0 and states at time  $t$  have a unique correspondence. For any constraint on the initial state, a convention about a probability function would be inadmissible which allowed any state to seem possible, since that would contradict what we know of the initial state. It may be concluded that such a convention about probability is not merely unnecessary, moreover it is not even strictly accurate under these circumstances.

But at this point it is necessary to consider that this difficulty arises only as long as we are content to rest with the ideal case of a gas in an adiabatic flask. It is this fiction which bounds possibility and lands us in difficulty. The fiction does not represent reality. Our entire perspective is altered if we consider that for the state of a gas at time  $t$ , not only is its state to be taken into account at time 0, but also that of an enormous number of other bodies, since the molecular motion of every point-mass exerts an ever-widening influence.

The whole problem is changed by consideration of this circumstance. If we regard a body as being affected by an almost unlimited number of other bodies over a finite period, it follows we are no longer able to conclude that there is an equalization of

---

<sup>13</sup> **Boltzmann**, L. (1878). Weitere Bemerkungen etc. *Wiener Sitzungs-Berichte*, **78**(2), p. 45. [See also Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*, 2. Band. Leipzig: Barth (F. Hasenöhrl). Reissued 1969 – New York: Chelsea.]

pressure and temperature in the gas. We are even unable to conclude that an equilibrium of heat is the most probable outcome in general, since this is not the result which actually occurs overall. Yet this constraint on what the theory can achieve represents the scope of what we are entitled to demand of it. It is sufficient if the theory shows that the most probable result under all circumstances – for a great number of bodies in close contact – is that the temperature, (and at least for gases) pressure and flow pattern may be represented as continuous functions of time and spatial coordinates. In this way specific relations hold between these values and their differential quotients over coordinates and over time. One can call that the achievement of an equilibrium of states for the smallest parts of space. This assumption can and must receive more precise specification by further specific acquaintance with actual conditions at all times, especially if equilibrium is to be assumed. In reality we should only assume a state of equilibrium with certainty for a gas once we know relative temperatures in the environment, at least approximately. Inasmuch as we make claims about such a state, we proceed from a postulate. The postulate is that when considering the conduction of heat in the flask which contains the gas, differences in environmental temperature ought to be ignored.

The global import of the theory is modified inasmuch as the constraining circumstances are changed, because the circumstances connect specific phenomena, as their most probable consequences. Those circumstances include not only some spatial arrangements, such as the containment of a gas in a defined space; they involve details of pressure, temperature, and motion too. It is not to be proven that a gas attains equilibrium – which is not generally the case anyways – but only that if a gas exhibits the same pressure and temperature everywhere at one time, and does not exhibit any currents, then Maxwell's distribution of speed does apply. Noticeable differences in temperature and flow will never occur in its interior spontaneously. That does not show that two bodies in contact necessarily attain the same temperature – likewise that holds only under some postulates. It does show that heat flows in the direction of descending temperature; that if body *a* loses heat to body *b*, and body *b* loses heat to *c*, then if one brings *a* and *c* into contact, heat is also lost from the former to the latter, etc.

If we consider this problem to be solved, then the result could be a much more general convention about probability which subsumes equilibrium as a special case. For a separate and arbitrary specification of values, it represents a case of flow, or the conduction of heat, etc. Thus it would lead to the derivation of constants for friction and for thermal conduction. It would not be fitting to express what is still only conjecture: whether the development of a closed form for the theory is possible in this sense. But we can predict the effort would be laden with tremendous mathematical difficulty. There are investigations of this kind – based on this principle – which Boltzmann<sup>14</sup> has developed to derive constants of friction for gases. They are still developed under a particularly simple postulate of flow which is stationary or time-

---

<sup>14</sup> Boltzmann, L. (1880, 1881, 1881a). Zur Theorie der Gas-Reibung. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften*, **81**(2), p. 117 and **84**(2), pp. 40 & 1230.

independent. Since calculations become especially complex here, initial prospects for solving the more general problem seem not very bright.

8. One should conclude from this overview that in physics the theory of probability represents no immature hypothesis. Rather, it is a hypothesis which has been developed logically, in a thoroughly impeccable way. – In an unmistakable way it has focused on what we called the originality of operational ranges. This, and the affirmation that a theory of probability arises from it based on the Range principle, is what I have given greatest weight. But since it assumes the essential correctness of the physical theory, the present discussion is not intended to materially reform or to round out that theory to any significant extent. I believe the exposition of this theory, and the recognition of other issues, must profit by clearer apprehension of the logical principles which are involved. I would hardly know how to offer a simpler and clearer exposition of the investigations sketched here, than by beginning with the notion of original ranges. In the equation found in **Maxwell's Law**, one sees proof of the originality of certain ratios of ranges. In the accounts which have been offered to date, one cannot fail to recognize what a barrier to understanding that uncertainty about these principles has posed. (I would remind the reader of **Meyer**'s 1877 *Kinetic theory of gases*, p. 259 onwards).

# Chapter 9

## More Applications of the Probability Calculus



**Abstract** There are many different types of error; one can describe where the probability calculus best describes error. Calculable random error is only one type. Bessel describes the many sources of error when using an astronomical instrument. The probability calculus may be applied when total random error is compounded from many sources, and where each makes a small independent contribution. The situation is well described by Gauss's law of error, but that law is not absolute. There are major theoretical differences between errors in observation and games of chance, for example. In general one cannot draw a strict parallel. The probability calculus is not well applied in the social sciences, and in medical statistics. The statistical treatment of majority decisions is also discussed, as in decisions made by juries. A critique is offered of the account of them given by Laplace and by Poisson.

**Keywords** Least-squares methods · Errors in observation · Gaussian error · Sources of error · Constant error · Mortality tables · Life expectancy · Majority decisions · Judicial decisions · Medical statistics

**1.** At first sight it seems the theory of errors in observation must be framed in a very similar way to the theory of games of chance. Just as for games of chance, outcomes are involved which are influenced by a very great number of varied circumstances. Nevertheless a closer inspection reveals a great number of significant differences.

We begin with a question: whether it is proper to report a numeric probability that an observation conducted under specific circumstances departs from a correct result by a definite amount.<sup>1</sup> To answer this question, we have to proceed under terms we established earlier (p. 134 and ff. in original; Chap. 6, Section 2). First we must try to form a clear idea of the general determinants presumed to be known *a priori* about the circumstances of observation. Among them are ranges of behavior which have caused errors of a certain extent, and which stand in definite and original ratios to one

---

(Errors in observation. General and medical statistics. Polling data.)

<sup>1</sup>Here as in what follows, it is understood that this is always a definite amount with a small and definite range. Strictly speaking this concerns the probability that the error is larger than a specific value  $x$ , but smaller than  $x + \Delta x$ .

another. Second, we will have to see what experience tells us about the occurrence and frequency of larger or smaller errors.

Initially one will find it plausible that for observations being conducted, that certain general statements may be made from earlier observations of the same type. From a series of observations, it seems possible to gain some impression of the capacity of the sense-organs of an observer, as well as some impression of his skill and concentration, of the quality of the apparatus, and so forth. There are a great number of factors here whose near constancy is not improbable, at least. By these means we would like to gain a clearer picture of the overall conditions, and their operational ranges. – From the many efforts dedicated to our purpose, initially we can garner a proposition that is easily verifiable: even with the simplest of methods, the error of an observational result always consists of a convergence of large numbers of distinct errors which may be called elementary.

In order to lend some perspective on the way this may be considered, I cite the sources of error given by **Bessel**, which must be considered “when a difference in the polar position of a fixed star is observed about a circular meridian constructed by **Reichenbach**’s method”.<sup>2</sup>

**Bessel** says: “First, the instrument must be positioned towards the star, and this adjustment may be in error for several reasons, namely: 1) because the resolution limit of the telescope has been reached, beyond which its orientation is haphazard; 2) because the point-image of the star which one intends to bring to the line of the reticle can be random within certain bounds, and this of course may diverge more for large bright stars than for smaller stars which are less bright, and from that it may happen that different points may be chosen at night than by day, or with a brighter sky as opposed to one less bright; 3) because the star never appears fixed or seldom does, but rather exhibits a shaky motion which comes from a lack of stillness in the atmosphere, and thus a choice must be made within the outermost bounds of this motion. These are associated with sources of error which are quite independent of the settings of the apparatus, for example: 4) the effect of the elasticity of the metal of the apparatus, which produces a reading now of one value and then another, randomly as the consequence of external circumstances. This may also have the consequence that the direction of the telescope does not stay the same at the instant the reading is taken, as when the adjustment was made; 5) an uncertain reading along the circle, which arises from small inequalities of separation between its tickmarks and the tickmarks on the nonius line, which is expressed as a variable error, since ordinarily different tickmarks are aligned to coincide for each repetition of an observation; 6) the uncertainty represented by the limited resolution of the optical aid by which the reading is achieved; 7) that error which arises from the circumstance that the estimate of the nonius line reading can only be gauged to half the smallest interval which is indicated, say of  $2''$ . Because of this, observations taken

---

<sup>2</sup>**Bessel**, *Gesammelte Abhandlungen*, vol. 2, p. 388. [Bessel, F.W. (1876). *Gesammelte Abhandlungen, Band 1.: Theorie der Instrumente, Stellarastronomie, Mathematik*. Leipzig: W. Engelmann.]

from all four nonius lines of this instrument may be judged as a full second, or as a quarter or a half or three quarters of a second, but not by any other fraction. In addition, external circumstances come into play such as 8) the influence of the observer's body heat on the circle or on other parts of the apparatus; 9) the influence of a temperature difference which is generally present between the lower edge and the upper edge of the circle – which produces stresses in the metal of the circle, and changes its shape. There is also 10) the presumption that the spirit level of the alidade (turning board) may not reach an undisturbed equilibrium on each reading, which represents a random error; 11) and another such error may proceed from the assumption that the instrument will remain in just the same state for two observations which are to be compared, though it is not infrequent that changes in state are noticed over a shorter or a longer period of time. The so-called error of observation is compounded with other biases, such as 12) the influence that the mistaken assumption may have, that the state of the atmosphere – as indicated by barometer and thermometer – is the sole and precise influence on the measure of the refractive index, and 13) the influence of small gaps in the reduced data of the observations. Presumably I will have overlooked several things in this enumeration of causes which converge to effect a putative error of observation, just insofar as I have not wanted to allude to random moments of distraction in conducting certain observational procedures, or to mottled or else unfavorable illumination of the threads and the chords of the circle, or to any influence which cold may exert on the apparatus, etc. However, our goal has been reached by this enumeration of sources of error, which is to draw attention to the fact that even this simple kind of observation must exhibit a total error which is compounded of many sources, each of which contributes to the total independently of the rest."

2. The insight that any overall error (to which the ready result of observation is subject) comes about as the joint effect of many elementary errors, will not get us significantly further. Evidently, that is because a numeric probability for the magnitude of overall error results only if a numeric probability obtains for each of the elementary errors separately. Yet we find that writers are inclined to say in general that a certain "probability holds" for every such elementary error, meaning that the error attains such-and-such a magnitude. If we may repeat the general consideration (which we had posited initially for overall error) with the same validity for any elementary error, then tracing that overall error to these elementary errors may be interpreted merely as a preparatory step. As we shall see, this is both useful and important, primarily because it allows us to postulate a certain 'law of error' *a priori*. That is a definite function which gives the probability of any amount of error, or at least it allows us to make a conjecture in advance. On the other hand, if we are to demonstrate that reporting numeric probabilities is at all comprehensible, we should pay attention to more detailed description of the constituents which determine the probability of elementary errors. We should look to gain a clear and distinct idea of their proportions. We encounter several categories of such general determinants. First, there are many factors we may consider to be constant in the best sense of the word. Among them are certain structures in the sense-organs of an observer, and certain functions of an apparatus. To this easily understood class of determinants we

add a second: that a circumscribed scope of possibility exists for certain varieties of behavior. As we came to know for games of chance (particularly for dice games), continual repetition of the same instances represents a very broad range of constraining circumstances. So if a reading on a scale is to be taken, there is a series of possibilities available to position a marker along the scale's gradations. Each position of the marker determines a reading error which is definite, once all other conditions are well determined, as when we scan the whole scale, or else when we estimate the position of the marker between two points across a tenth of the distance. Under these circumstances we encounter general determinants disposed to produce numeric probabilities for certain sources of error.

We come to a third type of general determinant which is far more difficult to specify clearly. For instance, consider disturbances in observation brought about by variations in temperature or vibration. We see that the numeric probability of the extent of error depends in turn on the probability that these disturbing influences are present with this or that intensity at some time. Should we need to trace a longer or shorter chain of contingencies, at least we can say what kind of conditions we must be led to consider. For any varieties of behavior which appear over many repetitions of the same kind, in the final analysis the general determinants must represent a tally of the entire set under which they must arise, though their temporal and spatial distributions remain undetermined.

So to cite a simple example, prevailing flows of traffic imply that a definite, nearly determinate number of larger or smaller vehicles drive down a street in a given stretch of time. They disturb our observations by shaking the ground, but the distribution of these disturbances over time falls within certain bounds which depend by chance on conditions unknown to us. Apart from the postulates mentioned earlier – that some factors remain constant, and that there exists a particular definite and bounded scope of possibilities – a numeric probability would have to be traced to the idea that for a large collection of things, the total amount they contribute to the phenomena must almost be considered to be determinable in advance.

Something similar would have to be taken as valid in many other conditions, as was established here for perturbing influences in the narrow sense of the term, especially for random fluctuations in attention. In general, numeric probability is tied to the constraint that phenomena depend on the random coincidence of very many individual varieties of behavior. Each of them must necessarily occur in many repetitions of the same type, in an aggregate which may be determined roughly in advance. This result makes clear to us how conditions for the comparability of ranges are satisfied. That is to say it is based on the wholesale repetition of behaviors of the same kind, which may be taken to be effective disturbances. Just as in games of chance – constant repetition of the same outcomes corresponds to extensive variation in constraining circumstances. Satisfaction of this condition is even more evident if we proceed from the other formulation given earlier (p. 73 of original; Chap. 3, Section 8). One is disposed to assume that among the possibilities left open to our inexact and incomplete knowledge, observation may be faulty to some extent, and in very many different ways. One is no less apt to assume that everything is indifferent to the expectation of a better or a worse result which we can establish,

say, about the particular circumstances of particular observations (apart from the sum of all their general determinants). Then if certain statements can be made about the sum of perturbing circumstances, within certain bounds their distribution is still a matter of the free formation of expectation. Hence a numeric probability may be given for various extents of error.

When we are faced with the question, whether and with what certainty such a summation statement can actually be made, it should at least be confirmed that our knowledge is insufficient to characterize all the factors to be considered, and insufficient to make definite statements about them directly by perception. Consider the fact that a series of observations – which we attempt to say are as much alike in kind as possible – will always count as having the same general determinants. That is something we conjecture, but it cannot be claimed with certainty. It is of the utmost importance that we examine this assumption in a purely empirical way, from long series of observations. Such tests – we will have more to say about them in a moment – are made very often. In fact they show the nearly complete validity of that assumption. As a rule they are not conducted with the intention of responding to questions about the constancy of general conditions. Rather they are conducted with a different aim: to establish a fixed rule for the relative frequency with which observations occur that are faulty to some extent. For that reason it is evident this might not succeed, if general circumstances did not remain very nearly constant over a whole series of one type. One should see to it that a definite regularity of the said type is established, and that the assumption of this constancy really is confirmed. To recapitulate: overall it ought to be concluded that it is meaningful to report numeric probabilities for errors of observation, that the postulate of such an *a posteriori* determination may be held true, and that the constancy of general conditions that underlies probability may be assumed – not assumed with complete certainty or exactly, but rather as an approximate rule of high probability.

3. Once this most important postulate has been settled, we may pass relatively quickly over two other postulates about statements of numeric probability which are considered next. The nature of most sources of error makes something immediately clear: observations are to be considered as a series of cases independent of one another. An empirical proof of independence is hardly necessary in this respect: it is unlikely that the requisite enormous number of observations of a single type could be provided, which are indispensable to such a proof. There may also be sources of error that can be considered to contribute equally along a whole series of observations. Theoretically we distinguish such ‘constant’ errors from ‘random’ errors. Since constant errors cannot be the subject of the probability calculus, we explicitly confine our investigation to random errors. The theory becomes simpler this way, but it becomes correspondingly narrow in application. Some additional conjectures are still necessary when we wish to judge the actual precision of a result derived from many distinct observations. One conjecture is that the methods (common to all) include errors which are constant, not random. Others are conjectures for which there is no way to assign a general rule. It must be noted that the usual distinction made in the theory of error – between constant errors and random errors – is by no means exhaustive. There can very well be other sources of error. Some sources of error vary

randomly but influence a fair number of consecutive observations the same way. Those sources would be neither completely constant nor completely random in the sense described here. It is not even too much to consider the presence of such sources of error, when we have deduced the precision of a long series of observations from some series, and then we apply the result to assess the precision of another small number of successive observations. Under these circumstances it is necessary to examine another possibility closely – that beyond constant errors, there are also errors which vary randomly in the same direction over a number of cases.

As far as the equality of chances for single cases is concerned (which is the last point to be made here), there is no doubt that the circumstances that constrain errors lie so completely outside our detailed knowledge, that we should judge every single case by the average chances of its category. Still the remark is not superfluous that the ratio which existed previously may change during the course of observation, and may not exist after an observation has been conducted. It is not at all uncommon during the course of observation that we are convinced we have particularly auspicious circumstances, and that we have garnered a result of exemplary precision. Perhaps we may have the opposite impression. This is of major importance, since the probability of a fine or a coarse exactness of observation is not considered in forming our expectations. It is not considered at first, but it may be used to judge the exactness of results after observations have been conducted. One sees that the introduction of a probability presumed in advance, in order to form conclusions about some results, is a business which depends on the postulate that conducting an observation gives us no purchase in attributing it either a finer or a coarser exactness.

**4.** From what has been said, it follows that a numeric probability for the extent of error can be reported to a good approximation most often, when the relative frequency of larger and smaller errors is determined by a sufficiently extended series of the most similar observations possible. This is simplified to a great extent – indeed application of the probability calculus to the theory of error only becomes possible to a great extent – if one assumes that the relative frequency of various amounts of error is regulated in a definite way governed by formulae of mathematics. Evaluation at each point in time has nothing to do with the type of functional relation which obtains between probability and the extent of error. That only puts a numeric value to a relation which is everywhere functionally equivalent. In other words it is assumed that with observations conducted under certain general circumstances, the probability of an error between  $x$  and  $x + dx$  is proportional in extent to the value  $e^{-\frac{h^2}{2}x^2}$ . This is the assumption **Legendre** and **Gauss** make, on which the so-called method of least squares has been based. Initially this assumption suffices to find the value in any series of observations of the same kind, which must be considered most probable. It also allows us to report the probability of different extents of error, by ascertaining the value  $h$  which enters into the formula, either for individual observations, or for the combination of several results. The value  $h$  is called the measure of precision for a certain type of observation.

Though many attempts have been made to prove this “law of error” in all generality, without assumptions, it is evident that such a proof may be constructed

only in connection with specific conditions manifest in such attempts. In the previously quoted work, **Bessel** states how well and under what postulates the law of errors may be considered valid. He sets that out in a complete and incisive manner. From his work it follows that **Gauss**'s assumption is more than a somewhat plausible premise which is recommended by its simplicity of mathematical form. The probability of different extents of total error always takes on a very good approximation to the function  $e^{-h^2x^2}$  which **Gauss** has assumed. That is true if total error arises as the independent and joint effect of very many elementary errors, without consideration of what smaller or larger extent may be credited to any one of these errors. Only two postulates must be made for this to be true. First, for any elementary error, equal positive and negative contributions are equally probable. Second, the average value of any error – or of the minority of errors – does not significantly exceed all others in size. One sees that the latter postulate is just a finer restatement of the main assumption: that the error is the joint effect of many separate errors. One would not be able to claim this, if overall error were overwhelmingly biased by one error or a minority of them, in the face of which all others could be disregarded. – It is not an unjustified assumption that such conditions are often realized. Considerations have already been offered which have acquainted us with the large number of independent errors that enter into an observation. As **Bessel** certainly points out correctly, complex observations are always orchestrated as much as possible. They are orchestrated so that individual procedures lend about equal precision to the final result, for the reason that it would be pointless to lend fine precision to one part of a procedure, if the final result is not improved because of the uncertainty of other parts. It emerges from all this that **Gauss**'s law of error should be considered a convenient, approximately correct assumption for many kinds of observation. This receives much-desired confirmation from the results of long series of observations, in which larger and smaller deviations appear with approximately the same frequencies as would be expected on **Gauss**'s assumption. The following Table (op. cit. **Bessel**) which applies to 470 observations made by **Bradley**, may serve as evidence. The third column of the Table lists actual frequencies in percent; the fourth column lists the percent frequency expected from those observational results; while their errors (more precisely, their deviations from the most probable values) appear within the fifth column. One can recognize from the fifth column, which lists the differences between theoretical values and observed values, just how negligible these differences are – how good the correspondence is.

| From | To  | Observed | Theory | Difference |
|------|-----|----------|--------|------------|
| 0.0  | 0.1 | 20.0     | 19.6   | -0.4       |
| 0.1  | 0.2 | 18.7     | 18.4   | -0.3       |
| 0.2  | 0.3 | 16.6     | 16.3   | -0.3       |
| 0.3  | 0.4 | 12.4     | 13.6   | +1.2       |
| 0.4  | 0.5 | 10.8     | 10.6   | -0.2       |
| 0.5  | 0.6 | 7.7      | 7.8    | +0.1       |
| 0.6  | 0.7 | 5.5      | 5.4    | -0.1       |

(continued)

| From | To         | Observed | Theory | Difference |
|------|------------|----------|--------|------------|
| 0.7  | 0.8        | 3.0      | 3.6    | +0.6       |
| 0.8  | 0.9        | 2.1      | 2.1    | 0.0        |
| 0.9  | 1.0        | 1.5      | 1.3    | -0.2       |
| 1.0  | etc. . . . | 1.7      | 1.3    | -0.4       |

But then **Bessel** shows that circumstances occur where the random influence of one source of error comes to determine the value of observational results. Quite different relations of probabilities of larger and smaller errors then take the place of errors represented by **Gauss**'s formula. The usual assumptions do not hold in such cases, and application of the method of least squares leads to false results.

As far as I can see, everything is given in **Bessel**'s investigation that is necessary to evaluate the sense to be made of the law of error, as well as everything which needs to be brought to bear on the issue generally. Any further conjectures must emerge from specific conditions of the methods of observation at the time.

**5.** From the points made so far, an approximate (not exact) and a frequent (not thoroughgoing) validity holds for the assumptions on which the mathematical treatment of the theory of error is based. In other respects, something similar may be noticeable, insofar as procedures are almost always subject to arbitrariness in small measure, and often even in large measure. The former sort of arbitrariness was found when we (p. 123 of original; Chap. 5, Section 5) described an *a priori* convention about probability, by which we hoped to lend certainty and precision to some observational results, lending precision to numeric computation. We showed that such *a priori* assignments of probability are always arbitrary. As a rule they are made in a way that is outright incorrect, but this is without practical significance, because the influence of all this arbitrariness on the end result is exceedingly tiny.

We can connect this to other situations, in which a certain arbitrariness in method is more immediately salient. For example, this is the case if it is necessary to combine observations of different types, and one has reason to suppose that one set of observations is more precise than the other, without definite indication of the measure of precision having been produced from the series of observations itself. Then it is necessary to strike an arbitrary convention for ‘weights’ to be applied to various observations. One might count the results of one observer to have two or three times the weight in a calculation than results from another observer.<sup>3</sup> Thus the arbitrary nature of this assessment can be seen.

Now if only approximate – not universal – validity can be bestowed on all the postulates of the theory of error, we should neither overestimate nor underestimate the scope of our result. First it should be considered that in nearly every case it is

---

<sup>3</sup>Cf. for example, **Hagen**, *Grundzüge der Wahrscheinlichkeits-Rechnung*, p. 54. [Hagen, G. (1867). *Grundzüge der Wahrscheinlichkeitsrechnung*. 2d. ed. Berlin: Ernst & Korn.; Gauss, C.F. (1821). *Theoria combinationis observationum erroribus minimus obnoxiae, pars prior*. Göttingen: Apud Henricum Dieterich, 6–7. Also in: Gauss, C.F. (1880). *Werke* 4. Göttingen: Königlichen Gesellschaft der Wissenschaften.]

enough if one can attribute approximate validity to one's results. In general it is not crucial that the probable error of a specific datum is assumed to be somewhat larger or smaller. Which value is taken as most probable on the basis of previous observation – that would be changed to an extremely small extent by petty modifications of the postulates. So it might be unimportant in practical terms if we should wish to replace **Gauss**'s law of error by some other rule that departs from it only a little: for example, a rule under which very large errors are deemed impossible. In the same way, it is immaterial for practical purposes that there is a certain arbitrariness to *a priori* assignments of probability. It is necessary to make the point that what is achieved here is – as a rule – the only thing that can be achieved. There can be no entirely general method for the assessment of observations which is guaranteed to deliver results that have to be considered maximally probable, excluding any doubt or arbitrariness. By the nature of the beast, there can be no such universal method. In particular cases the logic of the situation is such that a tidy numeric representation is simply out of order. Attempts to achieve such a representation cannot ever be satisfied. No effort of detail about assumptions or about the complexity of calculation will ever be able to do away with residual uncertainty and arbitrariness. It must be emphasized that the achievement of a solid methodology for combining many observations already represents large and important progress. There simply must exist such a means of combination. At least in some specific cases, a definite rule of combination does away with uncertainty, arbitrariness, and the possibility that one's procedures are just an uncontrolled motley. This has value even if it must be admitted that standardization is itself to some extent simply arbitrary, and that the assumptions are not quite right. – In many cases it seems both permissible and indicated to declare oneself satisfied with least-squares methods, and to let them be applied simply because there is nothing better to take their place. Yet one has to avoid something not uncommonly encountered: overemphasis of these techniques, and of their basic assumptions. I would like to draw attention to two main points. First, it is perverse to consider **Gauss**'s law of error as an absolute law which is strictly valid in any case whatsoever. Even if it has no bearing on the results of interest whether one considers very large errors to be improbable or impossible, still it is in the interest of clarity and accuracy of our ideas to ensure that untrue claims on this subject do not find their way into textbooks.

Then we should not approve, when **Hagen**<sup>4</sup> attributes the fact that very large errors of observation never arise, to the colossal improbability given them by **Gauss**'s formula. The postulate on which that formula is based – that of infinitely many elementary errors – is simply not a postulate which holds in all strictness. If one admits (as **Hagen** proposes) that in tracing errors to ever-more distal causes, causes become ever larger in number, still it must be taken into account that their influence becomes correspondingly less and less important. But apart from that, given a certain accumulation of errors, factors of a completely fresh sort are always in play. They make it impossible that the error of the result overflows certain bounds. The bounded

---

<sup>4</sup>Op. cit. p. 31.

nature of our yardsticks provides an example. It is utterly impossible that one will find an interval which actually covers ten centimeters to be equal to one kilometer, or find it equal to one tenth of a millimeter. It is quite impossible to arrange that this may happen by some random accident, with any constellation of ordinary circumstances. The law of error has meaning for small errors and for moderately large errors. To extend it over very large errors which never occur in fact, and to attribute to them probabilities of the said order of magnitude – however harmless that may be for practical application of the method, it is entirely ruled out in a general exposition of the theory.

I have to stress a second point that is more important. It is a far too closely-held idea, and an impression which is facilitated by common opinion about the importance of the probability calculus. The idea is that whenever any observation is evaluated by the provisions of the theory, a definite result is issued which is most probable, and its probable error is a definite one. The result is thought to be set by definition, not subject to any objection that the relevant result may have been misjudged, or any question about what degree of certainty should be attributed to it. Against this it cannot be emphasized strongly enough: there is no theory of error and no probability calculus which achieves that. All of them proceed from postulates which must be tested with every fresh case. Whether constant errors enter into observation, whether the law of error holds or whether there are sources of error which follow some other law of error and which prove to be preponderant, whether certain observations have been conducted under the same general circumstances as others: all these are questions which must be posed explicitly. The value of our results depends crucially on answers to these questions. What are always in play are conditions that may completely exclude determination in numbers – just as they completely exclude provisions of general validity. Ordinary expositions of the theory of error do not adequately consider this point. That is deeply rooted in the mistake of thinking that numerals for probabilities always serve to express a relevant state of knowledge, and do so in an exhaustive way. I have no doubt that most scientists who have to deal with applications of the theory of error, have no illusion on this point. A physicist who becomes aware of a new observation – if he has any reservations about method – will allow himself to be impressed very little that the probable error of the result is reported to be very small. He should immediately intuit that such a report is of pretty dubious value initially. That is especially the case when a constant error is to be considered, or if the precision of methods has been established first by a series of preliminary trials and then only reputed to hold for definitive observations made later.

**6.** Let us turn to the broad field of collective phenomena of human society. As statistics tell us, almost all those forms of human events and actions we have the opportunity to witness (as very frequent recurrences of the same kind) give the uniform appearance of an enumeration under which they occur over an extended length of time in a particular community. This enumeration has some regularity which is more or less exact. Either the regularity consists in approximate constancy to an absolute value, or it consists in an approximately constant ratio to a number, say the total number of individuals who belong to the community in question.

If we turn our attention to numbers of births, marriages, and deaths; if we establish numbers of crimes committed or numbers of trials held; or if we inquire how many letters sent to the postal service bear an incomplete address: from year to year we do not find the same proportions the same way, exactly. We do find proportions which vary between narrow bounds. Then in each of these cases the expectation seems justified that in the year to come, just about the same will come to pass once again.

Imagine we pose the same conjectures about these things as we had posed about errors of observation before. Then it is easy to convince ourselves that according to the Range principle we ought to imagine that for an approximate constancy of average results, there are ratios on which those expectations are based. We can imagine overall ratios which subsist because things remain constant in some way for a long time. The frequent repetition of other things too may be conjectured given a set size which is more or less fixed, though any detail of the rules governing their spatial and temporal distribution escape our knowledge. So it may appear comprehensible that the conditions under which someone lives have some general description, and that knowledge of this description – together with the age and the state of health of the individual in question – gives us an approximate chance that the individual will die over the course of a year. We might also have to consider that a series of hazardous conditions, meteorological factors, agents of contagion (and so forth) would have to occur to given extents. It might depend on the details of the assignment of these conditions to sample groups whether one of these factors would prove fatal to the person in question. Even in circumstances which govern people's origins and their life changes, we could imagine general determinants which represent a certain relative probability – say for individuals of a given age or a particular constitution – which represent their probability of dying in terms of their basic physiological constitution. Suppose too we consider that in the vast majority of similar cases it is always to be expected with maximum probability (according to repeatedly-cited theorems of mathematics) that the average frequency of a phenomenon represents just about its possibility in any single case. Then we would be in a position to set the expectation of an approximate constancy for such an average result. The expectation is set by the reasonable assumption of an approximate constancy of certain general conditions in transparently-known proportion. We could clarify those general conditions with respect to other varieties of behavior (such as marriages or the like). Yet without reference to the Range principle, it will not be easy in general to describe strictly where constancy should be supposed.

7. Now it is quite a different question whether specific numeric probabilities may be reported for any expectation in this domain of inquiry. Since there can be no question of direct determination of the overall ratios under consideration, then of course we are consigned to empirical inquiry. With the help of such inquiry we may easily find the sum of possibilities which general conditions represent for individual cases which go this way or that. But to arrive at numeric probability from this assessment, what is necessary are the postulates listed earlier: equality of chance for individual cases, constancy of general conditions, and independence of individual cases. Those are postulates which may be shown in part by simple deliberation, and in part empirically. Only a few words need be said about the first point: consider the

proportion of possibilities which constitute the overall ratio that a man may die between the ages of 40 and 41 years, and the overall ratio that a man may die at an age beyond 41 years. That ratio does not measure the probability over the next twelve months with which we have to expect death for one man who is now 40 years old. I will not delve into this argument or similar ones again, since they have been covered and sufficiently referenced in the text.

So far as the totality of results for many prospective cases is concerned, here the possibility of attributing values of probability to our expectations, depends above all on whether a real constancy of general proportions ought to be assumed. As we saw, that is a question which may be answered empirically with some certainty. In that light, we are now in a position to do something with the results of **Lexis**'s extended investigations.<sup>5</sup> **Lexis** distinguishes two kinds of series for phenomena. In the series he calls typical, a constant numeric ratio counts as lasting, and this constancy is expressed in the series. Such series indicate the continuing presence of general proportions; they correspond to a game of chance that continues indefinitely according to overarching rules. The role of randomness can still be quite varied in general – the rules of those games to which such a series can be compared may still be different. Thus typical probabilities may occur with normal or with hypernormal dispersion. The other kind of series is called problematic. It is characterized by a more or less variable state of human society. For **Lexis**, “in only one case has it happened that an undoubtedly typical degree of probability has been found with normal dispersion; namely, the probability of the birth of a boy or a girl”. It may be contended that “the collective phenomena of society overwhelmingly form problematic series.”<sup>6</sup> A numeric probability for an average result over many cases of the same kind can hardly proceed from a different postulate than from the constancy of general conditions and the independence of individual cases – in other words from the assumption that this is a typical probability with normal dispersion. Then it follows that apart from the ratio of genders in newborns, a numeric probability cannot be given in connection with this kind of event. **Lexis** appears to reach the same opinion in his investigations. He says, for example: “The raw annual figures for suicides in developed countries form a decidedly problematic descriptive series, which runs parallel to a certain social evolution. If one divides these numbers by contemporary population numbers, then one might perhaps consider these ratios as empirical values for composite sums of possibilities, in the sense given above. Since one would have to assume that the abstract probabilities would increase overall by a variable amount from year to year, then a profitable theoretical treatment of those ratios would still not be possible under the rules of the probability calculus. This has nothing to do with random variations in a typical value, nor with any magnitude which progresses according to some norm by a definite rule over time.”<sup>7</sup> **Lexis** expresses something similar at another point: “Of course one may formally consider

---

<sup>5</sup> **Lexis**, W. (1877). *Zur Theorie der Massenerscheinungen in der menschlichen Gesellschaft*. Freiburg i. B.: Fr. Wagner.

<sup>6</sup> Loc. cit. pp. 64 and 91.

<sup>7</sup> Loc. cit. p. 83.

every single value of a problematic series as a value which approximates an abstract degree of probability. Then one must assume further, that the probability on which this is based itself changes from year to year, or changes from one stretch of time to the other in a manner unknown to us, so that there is little to be gained by the introduction of such a notion of probability.”<sup>8</sup>

Under these circumstances the establishment of numeric probabilities in an extended sense remains conceivable in only one respect, that is, in relation to inferences made from statistical findings on putative changes in overall circumstances. We have already shown that the certainty of such an inference can be quite varied, depending on the conditions we have in mind. There is little doubt that for most phenomena, one will not succeed in defining general circumstances from whose constancy a normal dispersion of the results of a series would be expected. Take the general circumstances which determine mortality: we can think of them in the simplest of ways, by a report of how many individuals die in a year, that is, how their chances fall between probabilities of 0 and 0.1, between 0.1 and 0.2, 0.2 and 0.3, and so forth. As we have shown, there is reason to believe circumstances specified in this way should have a hyponormal dispersion if they persist. That is because in each tranche, cases with very different chances are brought together in one numeric proportion. Proportions of older and younger, healthy and sickly, wealthy and poverty-stricken individuals are lumped together under general circumstances. Necessarily, these find stronger expression in the observed numbers of fatalities; their expression is less dependent on chance than conjectured under the assumption of normal dispersion. The same will not hold for more distal causes, whose constancy in proportion still admits some measure of random variation. It hardly seems possible to characterize such distal causes in a fixed way. Neither ought one to claim independence for cases in general under any such conditions. Most of them, particularly those we are interested in knowing something about, are those which admit random variation, variation which holds in the same direction for the larger number of individuals who belong together. – So we arrive at the following result: if one wishes to draw conclusions about variation under certain general conditions from variations which are exhibited year to year in the total numbers of some social phenomena, generally we have no misgiving. However, if one wishes to vindicate this conclusion with a definite number that stands for certainty and precision, then that enterprise calls for more exact specification of what is meant by such general circumstances. Yet as soon as this is sought, it proves to be impossible to find.

**8.** If we ascend to the notion that for the collective phenomena of human society, reasonable reports of numeric probability can almost never be made, then we should like to define the consequences of this idea in detail. We would like to know how often we contradict intuitions which are current and well-recognized. One is used to thinking of the importance of the probability calculus as so marked with just these kinds of phenomena. The uses of the calculus are thought to have been tested by so

---

<sup>8</sup>Loc. cit. p. 91.

many different experiments over time, that this general dismissal itself seems questionable. But more detailed examination shows that anywhere pragmatic interest is solely involved, some near constancy of conditions can reasonably be expected – on average. To provide a basis for this expectation, it is simply necessary – as was discussed earlier – to draw on the Range principle in a general way. The numbers attached to any particular probability are not in question here. Here is an example: in a life-insurance company, it is extremely important to know that about this-or-that many deaths occur for a particular tranche of age over the course of a year. The company does not have the slightest interest in knowing whether a probability of death can be calculated for an individual in this tranche. Numeric assessments are very seldom wanted, or turned to account for a probability that expresses which departures of larger or smaller size are expected from the mean value of the results to date. No insurance company can or will adjust their reserve holdings for considerations formulated in purely mathematical terms. What is more here, the assurance which seems necessary in the face of large variations in average conditions, is measured by the quality of that reconciliation which those funds have shown to date, without proceeding on the basis of some other representation by a mathematical formula.

So far as I can tell, the situation is similar for all the domains which belong to this cluster. Then by the demonstration just given, I believe that what experience has shown to be a useful and indispensable methodology, need not be done away with. Still it is necessary more than ever to emphasize the falsity of results which may be obtained by ordinary procedures, where the constancy of general conditions and the independence of individual cases is assumed in advance to be evident. As an example, I would mention the method **Laplace** proposed to determine the size of a country's population, and the calculation he offers for the precision of his method. Of course this is a subject which is no longer of any practical significance. It is still one suited to a very clear illustration of the fallibility of his procedure – one which still appears in textbooks, by the way. **Laplace** began by saying that the number of births over the course of a year may be determined more easily and certainly than the total number of individuals living at that time. His method was meant to have been this: that the population is enumerated carefully in several smaller districts, while the number of births over the course of a year is also elicited for those districts. A quotient of these two numbers is found. If the annual number of births is known for the entire country, then total population may be found by multiplying that number by the quotient. In order to model this method on the probability calculus, the following approach is taken. The number  $N$  of individuals in the population is compared to the total number of draws made from an urn filled with black balls and white balls. Under the analogy, the number  $M$  of births is compared to draws that deliver a white ball. If analogous values  $n$  and  $m$  are produced by surveys of values for individual districts, it may be concluded that the ratio of white balls to the sum of black and white balls,  $\frac{w}{w+s}$ , should not deviate significantly from the value of  $\frac{m}{n}$ , or  $\frac{M}{N}$ . From this it is found that  $N = M \frac{n}{m}$ . Established methods allow one to determine the probability that  $\frac{m}{n}$  on the one hand, and  $\frac{M}{N}$  on the other, may differ by no more than a given amount from the quotient  $\frac{w}{w+s}$ . From this a certain datum ought to be evident for the precision to be attributed to the method. However, more careful examination reveals how badly actual ratios are represented by those for games of chance, which

have been called on in the comparison. Mainly, the reliability of the method is tied to the postulate that general conditions which determine the probability of births, are the same for all  $n$  parts of the country. That is very doubtful, and indeed unlikely ever to be the case. If we want to make this illustrative example work, we have to add that it is completely unknown whether draws of  $n$  and  $m$  made on the one hand and  $M$  draws made on the other, were taken from the same urn. They might well have been taken from two distinct urns that contain white balls and black balls in unequal proportions.

Let us assume what is actually the case: the relevant proportions are different across various districts of a large country. Then no judgment whatsoever may be delivered about the precision of results, without some knowledge of these differences. One might just as well claim an analogous finding, say, based on numbers of crimes against property rather than numbers of births. The error bounds would appear to be just the same at first (ignoring the influence of the smaller numbers involved). Nevertheless, no one would doubt that this would provide an unreliable reckoning, because local conditions obviously prove to have pronounced differences.

**Laplace**'s account of the errors to be guarded against in his method, is subject to a specific factor in this calculation, which allows residual possibilities of error seem to be the ones which confer most weight, to the neglect of all others. I will forebear from introducing further examples of this kind. Instead I will cite a case of a rather different sort, where the tacit and unjustified supplanting of a basic postulate leads to very contentious results about a relationship that is not unimportant in practical terms. I have in mind the postulate of independence for two types of behavior which have a specific relation to one another. As previously outlined (p. 178 in original; Chap. 7, Section 8), this leads us to derive the frequency of combined events by a statistic which rightfully pertains to only one of the elements.

A procedure of this kind is used to deduce results about the duration of marriages, widowhood, and so forth, from ratios of life expectancy. Those ratios are found separately for men and for women.<sup>9</sup> A simple schema for the solution to the problem runs as follows: Suppose there are two people  $A$  and  $B$ ,  $m$  and  $n$  years of age, respectively. Let  $p$  stand for the probability that  $A$  continues to live after  $t$  years have passed. Let  $q$  stand for the probability that  $A$  is no longer alive after that. Analogous values for person  $B$  are given by  $p'$  and  $q'$ . The composite probability that both

---

<sup>9</sup> Among others, cf. **Drobisch** (1880). Ueber die nach der Wahrscheinlichkeits-Rechnung zu erwartende Dauer der Ehen. *Berichte über die Verhandlungen der königlich – sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematische – physische Klasse*, 32, 1–21. ‘Let us indicate the number of men who survive at the end of each year  $m, m + 1, m + 2 \dots$  as being  $l_m, l_{m+1}, l_{m+2} \dots$ ; these are numbers to be taken from mortality tables. Then the probabilities that a man of  $m$  years of age will still be alive after 1, 2, 3 … etc. years will be:  $\frac{l_{m+1}}{l_m}, \frac{l_{m+2}}{l_m}, \frac{l_{m+3}}{l_m}$ , etc.. Now if  $l_w, l_{w+1}, l_{w+2} \dots$  signify the same numbers for women, then similarly the probabilities that a woman of  $w$  years of age will still be alive after 1, 2, 3 … etc. years will be:  $\frac{l_{w+1}}{l_w}, \frac{l_{w+2}}{l_w}, \frac{l_{w+3}}{l_w}$ , etc.. The product of these probabilities:  $\frac{l_{m+1}}{l_m} \cdot \frac{l_{w+1}}{l_w}, \frac{l_{m+2} \cdot l_{w+2}}{l_m \cdot l_w}, \frac{l_{m+3} \cdot l_{w+3}}{l_m \cdot l_w}$ , etc. are the probabilities that this man and this woman may live together or jointly after 1, 2, or 3 years, and the sum of these probabilities gives the number of years they may hope to live together, to the end. Then in that much, this number may be termed the duration of the relationship, or if the relation is marital, as the length of the marriage.’

people are alive after  $t$  years is  $p \cdot p'$ . The composite probability that  $A$  lives on, but  $B$  does not, is equal to  $p \cdot q'$ , and so on. These assignments of probability invoke the postulate that longer or shorter life expectancies for  $A$  and  $B$  are independent, in the sense given by the probability calculus. This is surely false as soon as we do not mean  $A$  and  $B$  to be two arbitrarily chosen people, but we mean that  $A$  is a married man and  $B$  is his wife. It ought to be considered that mortality tables – from which the values of  $p$  and  $p'$  are likely taken – cover all sorts of living conditions. For a married couple a goodly number of them necessarily coincide, such as the conditions of hygiene in their home, their level of wealth, and so on. Also, the death of one of the pair is always a factor which significantly affects the subsequent life of the other. Hence that death makes a contribution to the life expectancy of the other member of the pair. We are in no position to evaluate the degree of this mutual dependence in advance. We can only ascertain solid facts about the frequency of longer-lasting or shorter-lasting marriages or widowhoods, by direct investigation. If that does not seem achievable on practical grounds, admittedly the results of this method may be the best that can be done. I am far from declaring such results worthless; they could serve as an orientation to the topic, since the influences which are unconsidered may really not be significant. Still in the interest of methodological accuracy and in the interest of a correct account of the results, it is indispensable to bring to light the postulates that enter into an investigation, if those postulates are not strictly accurate. Should that be left undone, once again we have to face the consequences of this method, which has already and repeatedly been characterized in general terms.

9. From what has been said, the probability calculus is only to be applied to collective phenomena of human society in broad strokes, for the purpose of investigating which general conditions must be considered to underlie social phenomena. In other words they are methods which we characterized as ‘exploratory’ methods earlier (p. 176 in original; Chap. 7, Section 7). In **Lexis**’s investigations, normal dispersion is exhibited by relative frequencies of male births and female births, in the aggregation of cases by time and location. The important conclusion may be drawn that no circumstance alters this probability ratio over times and districts. Another outcome of **Lexis**’s investigations may be even more interesting. Statistics show a definite functional relation between the frequency of deaths which occur in the  $n^{\text{th}}$  year of life, and the deviation of that age from another age – a ‘normal’ age of about 70 years. (The number is not exactly the same for men and women, or the same across countries.) This functional relation represents the frequency with which observed values depart by differing amounts from the correct value, as expressed by **Gauss**’s formula.

At first, this makes it seem highly probable that a simple physiological meaning can be assigned to this ‘normal’ age. One might say “just as the human organic constitution determines a certain normal body size, so it also determines a certain normal longevity”, and that this is reflected in the numbers. It may be conjectured that each larger or smaller deviation from this normal longevity is determined by the independent and joint action of a large number of circumstances, which remain valid over a whole range of general determinants through time. The direction in which and the amount to which this deviation holds, depend on the form of the specific

constellation of determinants at any time, which is to say: they depend on chance. How such “plots of specific longevity” and the chance circumstances which modify them are to be considered, and how such questions are to be addressed, is of course a matter of biology. – More detailed examination of this interesting topic is thwarted here. Such an examination would have to take into account the fact that the reliability of the conclusions is impaired, in that the functional relation does not leap directly out of mortality tables. Instead it is obscured in part by so-called premature deaths. At least the present exposition suffices to show a curious way the probability calculus may be applied in this instance. In doing so, it may make clear the benefit of such investigations.

**10.** For many reasons it is suitable to have a separate section to discuss phenomena that are the subject of medical statistics. Strictly speaking they should be covered under the supervenient notion of collective social phenomena. This approach recommends itself, mainly because the goal has been especially explicit in the application of the probability calculus to clinical experiments. The aim has been to establish general conclusions from the evaluation of observations. That is, the aim has been to establish maximally strict methods free from arbitrary influence, and to establish a precise figure for probability – meaning characterization in numbers. Propositions which assert that some circumstance (i.e., a specific treatment) exerts an influence on the course of a disease should be assessed exactly. The empirical basis of these effects should be fully demonstrated, and the certainty of such results should be assumed to be rigorously determined. Then we ought to ask the question, whether it is possible to assign numeric probabilities to propositions of this kind. Let us begin with the following deliberation: say we treat a large number of cases of a certain illness by a new method, and we observe an unusually favourable result – such as a very low incidence of mortality. The probability that we may have to consider this new therapy a success depends fundamentally on the probability we ought to expect the same result by the original treatment. The smaller the latter, the larger the former must prove to be, in general. Accordingly, the initial investigation will be directed towards a question: what probability might the result (just observed) have under the previous treatment ? We can only hope to answer this question by reporting a numeric figure for that probability, when a sufficiently large number of similar cases have come to light, which had been treated by the previous method. In order to report a numeric probability from such an observation – to say that a number of new cases result in a specific total mortality – further knowledge would be required. One needs to know that the general conditions under which the new cases occur are the same as those already observed, and that these general conditions allow the series to have a normal dispersion. We never know so well the general conditions on which the manner of the incidence and the course of a disease depend, that we could ever establish the validity of those assumptions directly. We would have to rely on the method outlined in Chap. 6, to investigate whether among those conditions which do not differ markedly to us, there is a constancy of mortality and a normal dispersion of results from the series. Maybe in this way those propositions we are convinced of can actually be realized – propositions on whose foundation numeric probabilities are to be based. Maybe one could claim this never happens. First, treatment methods

cannot be controlled in experiment in the way that would be necessary to serve such a purpose. Steadily richer data from experiment; constant improvement in general conditions of sanitation (in homes as well as in hospitals); and finally the duty to introduce rough treatments without delay, if they offer some prospect of improved success: these are all important factors that prevent us from controlling even those conditions we can control.

As a consequence, one ought seldom to have any success in proving, for example, that an incidence of mortality remains constant over a long period. Nevertheless, such examples exist: I am reminded of the tabulation which **Fismer**<sup>10</sup> made of mortality rates for cases of pneumonia treated between 1838 and 1866 at a hospital in Basel. Namely from 1839 to 1848, 24.7% died; from 1849 to 1857, 24.9% died; and from 1858 to 1866, 25.9%. In this result we may find the expression of a very approximate constancy for the incidence and the course of general conditions governing these lung infections. Though the principles of treatment did not remain exactly the same over this entire period, whatever changes there may have been, they had no meaningful effect. But on the basis of such empirical results, it cannot be considered certain that in the succeeding years general conditions remained unchanged, with exception of revised therapeutic methods introduced in the year 1867. We can lend that premise a somewhat, but not exactly-defined degree of probability. If we recall what a multitude of things have changed in living conditions and sanitary conditions, or in the provision of hospital supplies, etc., then we ought not to esteem the probability of this premise as being all that high, even in this exceptional circumstance.

The question is no less important, whether we should consider separate and successive cases to be independent, and therefore if we should expect normal dispersion – that is, under general conditions which may remain constant, or nearly constant. The certainty of the expectation given to the average result for a set number of cases will be modified greatly, if circumstances are present to which many cases observed in succession are equally susceptible, either in one direction or another. On that score, we know that some diseases vary in incidence over time without directly perceptible cause. Hence one speaks of small outbreaks and severe epidemics of measles, or scarlet fever. Then if one treats 50 or 100 cases of scarlet fever by a new method and achieves especially favourable results, that is but a trifling proof of the efficacy of the treatment, compared to the treatment of an illness with constant uniform incidence.

Of course it is unknown how other diseases may behave in this respect. It does not seem downright impossible that very often they exhibit a hypernormal dispersion. It gives us pause, that if we assume a variable character of the agent of contagion in our explanation of small outbreaks and severe epidemics, a similar factor can be found for any disease which depends on an external noxious agent. They may have

---

<sup>10</sup>**Fismer**, F.H. (1873). Die Resultate der Kaltwasserbehandlung bei der acuten croupösen Pneumonie im Basel Spitäle von Mitte 1867 bis Mitte 1871. *Deutsches Archiv für klinische Medizin*. **11**, pp. 395 & 396.

unpredictable changes by nature, though they act on many people at the same time in the same manner. Other injurious effects that have some connection to meteorological phenomena might be like this, no less than any number of factors of general hygiene and social conditions which might vary randomly. Any of them would lead us to expect a hypernormal dispersion for many diseases.

At this juncture we have another point to add. Everything which was said until now dealt with the probability with which we could expect to have specific results from a process when methods of treatment are applied in the same manner. We found that this probability was not specifiable in numbers, not strictly. It must also be considered that this *a priori* probability does not directly or exclusively determine judgments of therapeutic efficacy. The cases in which successes were achieved, are just a matter of observation. Then the problem becomes: would these actually observed cases have produced a specific mortality if they had been treated by the previous method? Whenever<sup>11</sup> the statistics of mortality covers cases of an illness which present themselves differently to investigating physicians, and hence to which a different prognosis is assigned from the outset, observation of the cases at hand provides us with clues to judge how these cases would have progressed under the previous treatment. If the relevant cases have been treated by the new method, knowledge of the types of those cases may give us more solid conviction about the efficacy of this method, than is achievable without any consideration of these circumstances. Occasionally a prudent and experienced clinician makes relative judgments which are very convincing and assured, from treatment of a very few cases, especially if they are severe cases which had a positive outcome. If the disease admits fairly definite prognoses, then this may well happen. It follows that a numeric probability can never be reported, that a therapeutic intervention influences the course of a disease in this way or that. In the absence of further considerations, it is not permissible to treat successive cases of a disease like turns in an ordinary game of chance; it is not permissible to draw the parallel between treatment modifications and changes to the odds in a game.

**11.** Even with this result, it seems worthwhile to indicate how much this may contradict, or how it may be consonant with the views of other authors. In recent times application of the probability calculus to this domain has been heartily recommended. It has been deemed quite essential. Some authors who recommend it have not simply found themselves committing the error represented by a schematic treatment along the lines of an analogy to games of chance; neither have they appreciated well enough the distinctions which have been drawn here. **Fick** even cites the postulate explicitly, that apart from the change of therapies, “all other circumstances had remained the same”<sup>12</sup>. He leaves it unconsidered that that cannot

---

<sup>11</sup>In the face of these circumstances, if one believes one must require that conclusions should be based solely on the observation and the comparison of cases which are really entirely of the same kind, then I think that means one requires the subject of the investigation to be very different than it is in reality.

<sup>12</sup>**Fick**, *Medizinische Physik*, 3d. ed., p. 430. [Fick, A. (1885). *Die medicinische Physik*. 3d. ed. Braunschweig: Friedrich Vieweg und Sohn.]

be established with certainty. Purely and simply he describes the numbers which are calculated as the probability with which the method of treatment exerts a positive effect. It is most intimately connected with **Fick's** theoretical ideas (which will be examined in the last chapter; here Chap. 10) that this point – the concordance of auxiliary conditions – goes unemphasized. His principal emphasis is on the point that the “concept of disease” is understood the same way for two series to be compared. – **Liebermeister**<sup>13</sup> has done far better justice to the situation in question. In his work we find it expressed with all the precision one could desire, that one must always think of changes in general conditions: changes we do not control. He says: “the probability calculus tells us in all strictness, what level of probability is to be assumed, that a difference has occurred in the constant conditions (this may be supposed to mean ‘general conditions’) on which a successful outcome depends; it is completely incapable of saying anything about the nature of that difference. The investigation of causes is a matter of clinical analysis, as is the question whether this is to be sought as a difference in treatment, or as the difference of other constant conditions.” Although this is appropriately expressed<sup>14</sup> and repeatedly urged, still he does not draw the consequences.<sup>15</sup> In fact, here we have the general case that was examined earlier (p. 179 in original; Chap. 7, Section 8), that concerned the judgment of probabilities for some schematized examples, where the probability of a single premise of interest was to have been represented numerically. We saw the conditions which determine whether or not significant interest is owed to specification in numbers. If we proceed from the principles which were established there, there is hardly any getting away from the fact that a schematic procedure for the calculation of probability is superfluous in most instances for this subject. Even if we support our conclusions with great masses of numbers, there is the danger of being deceived by a specific grouping of certain general conditions. The calculable risk of being deceived is relatively trivial in comparison to the chance that the comparison series are distinguished by general conditions yet unknown, distinct from treatment methods. Hence the probability of the conclusion is essentially determined by the certainty we have that this was not the case. This can be demonstrated by some simple considerations. So it is heavily weighted, if results are seen to improve after a novel therapeutic method is introduced not just at one place, but rather when the same observation has been made at several different locations. The probability of a positive effect of the novel treatment is much greater, than if all cases had belonged to the same hospital or even just the same location. It does not seem inconceivable that within a restricted region, some unknown change might have been introduced almost simultaneously with introduction of the revised therapy.

---

<sup>13</sup> **Liebermeister**, C. (1877). Ueber Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik, *Sammlung klinischer Vorträge*. (Innere Medicin No. 31 – 64), 110, 935–962.

<sup>14</sup> Here I glean from this circumstance that still, only such general conditions are considered whose constancy would determine a normal dispersion.

<sup>15</sup> Op. cit. p. 18.

Positive results might be attributed to that unknown change. But it would be an extremely improbable assumption that this could have occurred in the same measure at several places.

The more extensive a series of observations at one's disposal, the smaller the probability of being deceived by random effects – and therefore more weight should be given to examining other postulates. The usefulness of numeric evaluation is lessened accordingly. If the observations subsume only a few cases, then close consideration of individual cases becomes all the more possible. Here too the numbers which may be calculated should not be seen to lead the investigation. Deliberations of this nature, to whose number we need not add, show that performing calculations of probability in making decisions about therapeutic outcome is at most necessary as a very limited exercise, with limited usefulness. Calculable random error is one source of error among many. As a rule it is not the most important one. Some acquaintance with specific numeric proportions is then necessary, just as in chemical analyses it is indispensable to know how to measure weight precisely. In most scenarios, reports of probable error at each step of the weighing procedure would be equally pointless. It gives us no idea of the certainty of the analytic results. If one begins with the perspective that observation of a mere fourteen cases can prove nothing, even when the course of a disease is altered by the application of a therapy fourteen times in succession, and when this altered course is seldom observed otherwise, then one is completely in the dark – in complete confusion – about what can be attributed to randomness. Such confusion may – must – be alleviated by the practise of calculating probabilities that have meaning only as illustrative examples. If an event appears three times in 10,000 similar cases, and later it appears five times in a further 10,000 cases, anyone who believes this is unlikely to be random because of the large number of observations is subject to a crude error.

I cherish the hope that in the exposition given above, that the place and significance the probability calculus can achieve, or must achieve in medical science, is properly valued. Under the crucial presumption that one does not surrender to illusion about the face value of numbers given to probabilities, recourse to the probability calculus can only be useful and praiseworthy. Some knowledge of these methods may even be called indispensable. Nevertheless it is conceivable, as it is justified in practise, that these computations play no important role in the literature. One may suspect this will be the same in future.

**12.** Perhaps the most noteworthy of all the applications made of the probability calculus, is the assessment of decisions reached by majority vote. Majority votes occur in elections, judicial decisions, and ballots of various kinds. For our purposes it is enough to discuss one example in more detail. For this I choose verdicts made by jurors in a trial. They have been handled on the model given by Condorcet,<sup>16</sup> first by

---

<sup>16</sup>de Condorcet, M.J.A.N. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris, de l'Imprimerie Royale.

**Laplace**,<sup>17</sup> and especially by **Poisson**.<sup>18</sup> I wish to show that both **Laplace** and **Poisson**, proceeding on incorrect premises, arrive at results which do not bear the significance attributed to them. I hope to achieve this by a simpler means than a detailed critique of **Poisson's Works**. Such a critique would lead us far afield from the bounds of the present investigation. Instead I hope to give a concise and independent treatment of the subject, occasionally accompanied by examples from those authors.

In order to set the stage with the simplest conditions, let us assume that in reporting a verdict, a jury is composed of twelve people who are chosen by lottery from men whose names appear in eligibility rolls. In general not all jurors will judge the case in the same way. Hence the verdict is not unanimous, but delivered with a larger or smaller majority. This diversity of opinion leads to the possibility that someone may be declared innocent, who would have been judged guilty by a jury of another composition, or vice versa. Then to some extent, the result of the judicial proceedings depends on the outcome of a lottery: it depends on randomness. Obviously the significance of this is greater, the closer that probabilities of opposing decisions are to being equal; it is smaller the more one decision outweighs the other. There is a shortcoming in this arrangement because of the ‘randomness’ of decisions. That explains the desire to assess the role of chance in the outcome of verdicts more closely, to learn how that role may be mitigated. To approach both these tasks more closely, upfront we should consider that by the nature of the case, and by the psychological makeup of people whose names appear on eligibility rolls, that the way each person will vote in the end may be thought to be determined ahead of time, just in case that person becomes a jury member. Then before the fact there are a certain number  $x$  of people who would develop their views one way, and another number  $y$  of people who are likely to form the opposite opinion. Let  $x$  be the larger number. We may call a decision with  $x$  a conforming decision, and a decision with  $y$  a dissenting opinion. Then a view is either conforming or dissenting, depending how it agrees with or fails to agree with the judgment of the majority. According to the presumption just made, the judgment of the majority is considered an ideal judgment. If we knew ahead of time how large the probability is of obtaining a conforming verdict, or else a dissenting verdict – in other words how large the relative frequency was of dissenting verdicts, over very many cases – then we might have a measure to say how often “randomness finds a home” in the existing arrangement. The requirement posed earlier is to formulate – insofar as possible – when the probability of conforming verdicts outweighs the probability of dissenting verdicts. If one cannot prevent false decisions from being made – it is in the nature of many cases that they lead to a factually incorrect judgment – still it is a revealing stipulation that the resulting judgment is not determined by lottery, but should turn out (as much as possible) in favour of the overall majority.

---

<sup>17</sup> Laplace (1812). *Théorie analytique des probabilités*. 3d. ed., pp. XXXVI & 460. [Laplace, P.-S. (1812). *Théorie analytique des probabilités*. Paris: M<sup>me</sup> V<sup>e</sup> Courcier.]

<sup>18</sup> Poisson, S.-D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris: Bachelier, Imprimeur-Libraire.

If we try to weigh the probability of conforming verdicts versus dissenting verdicts, then we see that the more jurors who form a jury at any time, the larger the probability of one and the smaller the probability of the other. If we draw balls from an urn which contains more white balls than black, then we ought to expect a greater certainty of obtaining white balls than black, the larger the number of times we undertake to draw one. This is very important – insofar as, if it turned out that the influence of randomness is significant for a specific proceedings, one would be in a position to mitigate the influence of randomness merely by composing a jury of more people. In order to be able to determine that probability, knowledge of one factor is necessary first of all, one which cannot be foretold. One would have to know how numbers of conformists stack up in proportion to numbers of dissenters in general. If for a specific case the former were 90% and the latter 10%, then evidently the risk of a dissenting judgment is much smaller than if their proportion were 60 to 40. If their ratio were only 50/50, then even with the strongest jury, it would depend only on randomness which verdict would emerge. The probability of a verdict of innocence and the probability of a guilty verdict would be equally large at the outset, simply on the nature of the case, and the temperaments and intellectual qualifications of all jurors listed on the eligibility rolls. Now it was an inspired idea on Laplace's<sup>19</sup> part, initially a correct one, that one may have opportunity to discern this ratio along the lines of the empirical determination of probability, once one had at one's disposal extensive statistics which would bring to light the relative frequency of verdicts delivered with some majority (or other). Here the circumstance is not counted as a significant difficulty that the numeric ratio of conformists and dissenters may often vary for different cases. As one might assume for a large number of cases, even for the relative frequency of those which are judged with greater or lesser divergence in opinion, one will be somewhat authoritative. So for the latter, which we may consider to be measured by the fraction  $\frac{y}{x+y}$ , there will be a certain average value. One could assess this value, and that would also be sufficient to provide a measure of the relative frequency with which conforming sentences and dissenting sentences are pronounced. The court of the *Administration de la justice criminelle* delivered a statistical report of the desired sort for the years 1825–1833. For several reasons, the relative frequency of judgments pronounced with the smallest majority (7 versus 5) was then established. Judgment was given against 61 of the 100 accused, on average, and of those 7 were judged by that smallest of majorities.<sup>20</sup>

That in discovering such a ratio there seems to be means to become familiar with “the average degree of divergence in opinion” as one may say for short – and that in this way one can hope to determine the risk of dissenting verdicts: that much is understandable straightforward. The act of drawing twelve random names from a list with some conforming and some dissenting voices, is an act that represents a series of twelve independent cases, for each of which there is a certain chance of one outcome or another. If we select series of twelve draws each from an urn filled with

---

<sup>19</sup> Op. cit. Laplace, p. 460.

<sup>20</sup> Op. cit. Poisson, p. 14.

black and white balls, then from the relative frequency of series that contain seven balls of one kind and five of the other, we may deduce the numeric ratio of black and white balls in the urn without difficulty. The problem about verdicts might be handled in the same way, too, by familiar methods of the probability calculus.

However, it is necessary to examine the suitability of the simplified premise with which we began. Every person whose name stands on the list may be considered ahead of time as conforming or dissenting in a given case. Thus the random draw of twelve names in forming a jury represents a series of twelve independent cases. A definite chance of one outcome or the other may be attributed to each of the twelve. As we may say after some thought, this premise is altogether incorrect. The individual members of a jury do not form their decisions independently. They always have a more or less intense exchange of opinion. The vote that an individual delivers in the end, depends fundamentally on who else sits on the jury. Under these circumstances we have to modify the meaning of randomness, as being the relative frequency of dissenting judgments. That is: at present we may only consider the verdict determined in advance, which a jury delivers after it has been composed of any twelve people from the list. If the number of people on the list runs to 500, from whom members of the jury are chosen randomly, there are a large number of possible  $\frac{500 \cdot 499 \cdots 489}{1 \cdot 2 \cdot 3 \cdots 12}$  distinct combinations of people on the jury. All are equally probable in advance, of course. The judgment which the majority would finally deliver can be called conforming, and the contrary judgment may be called dissenting, as before. Once again we have to ask about the relative frequency of dissenting judgments. But since under the current premise (one that is closer to reality), individual lottery draws are no longer independent with respect to the delivery of a conforming or a dissenting voice, there does not seem to be a simple ratio which gives that frequency – as the percentage of judgments pronounced with a larger or smaller majority. One gets the point right away: the more that opinions of individual jurors influence one another *in camera*, and the more that jurors find conciliation, the greater will be the majorities will be by which verdicts are delivered. Yet this result tells us nothing about the risk of dissenting judgments. Let us suppose that there might be a significant number of individuals who have a proclivity to let their opinion be swayed by the way prospective views of the majority lean in consultation. By this means, a much stronger predominance of large majorities will be achieved than would be observed otherwise. If we are indulging in fiction, we might even imagine that reconciliation of opinions went so far as always to produce – or almost always to produce – unanimous decisions. The ‘randomness’ associated with results could still be very large. In this multifaceted dependence of individual votes, we are dealing with a factor that eludes exact assessment or empirical determination.<sup>21</sup> This insight makes it clear that there is no story of a suitable investigation here, which tells us how much verdicts depend on randomness.

---

<sup>21</sup>For the purpose of an empirical assessment it would be necessary to present particular cases in succession to different juries for their decision. Of course there are good reasons why this should not happen; one is likely to disregard how the material for such an investigation should be provided – material from which one might gain some picture of the “degree of randomness” attached to verdicts.

Shortly put, the actual problem differs from the abstract schema because individual cases are not independent. Therefore results gotten from the schema lose any real importance.

**13.** Up to this point, I have confined myself to simple and comprehensible questions. I have shown that answers to these questions are impossible to obtain by the probability calculus, for a very specific reason. Now let us cast a glance over some investigations conducted by mathematicians, particularly those conducted by **Poisson**. It may be said at the outset that an erroneous premise of the complete independence of individual cases underlies all their results. There is talk of the chance that the judgment of a juror randomly chosen from the rolls may be correct or incorrect. This is connected to a certain probability of a true verdict or a false verdict on one hand, and to a larger or smaller majority decision on the other. This is the clearest error in the entire investigation. If one proceeds on the assumption – which seems less than clever – that the mutual influence of jurors is always such that it increases a majority, then one can claim that the influence of randomness on verdicts is greater than (it may be much more significant than) it seems from **Poisson**’s figures. – Even this incisive critique of those investigations would be unable to reveal the most important flaw in this mistaken premise. There is a far more significant flaw in the notion of what is taken to be the subject of investigation. The subject of interest is truth and falsity; it is not at all the probability or relative frequency of conforming and dissenting judgments. What may be meant by a correct judgment here, is not a judgment that declares those who are guilty to be guilty, and those who are not guilty to be innocent. A frequency of true judgments and false judgments in this sense cannot be established, which is illuminating in itself. Yet in **Poisson**’s writings truth or falsity of judgment has a relation to something objective: the “actual probability” standing for the guilt or innocence of the accused. **Poisson** proceeds on the premise that one may declare an accused to be guilty, if one assumes his guilt has a probability significantly beyond  $\frac{1}{2}$ .<sup>22</sup> He ends up with a conceptually-based criterion (loc. cit. 1837, p. 389), as when we read: “For the potential jurors listed by any criminal court, and for both kinds of crime which we have identified, we should consider that there is a certain probability  $z$  which is judged necessary and sufficient for conviction. That being the case, the chance  $u$  that a juror (picked at random from the eligibility rolls of this district) will not be mistaken in his vote, is the probability that he will judge the probability of guilt of the accused to be equal to or greater than  $z$  if it is, effectively. Otherwise he will judge the probability to be less than  $z$  if it does not effectively surpass this limit.”<sup>23</sup> It is incredible to imagine that the

---

<sup>22</sup> Also see a similar exposition given by **Laplace** in his *Théorie analytique . . .*, p. LXXXVI of his Introduction. [Laplace, P.-S. (1812). *Théorie analytique des probabilités*. Paris: M<sup>me</sup> V<sup>e</sup> Courcier.]

<sup>23</sup> « Pour les jurés du ressort de chaque cour d’assises et pour chacun des deux genres de crimes que nous avons distingués on doit donc concevoir qu’il y a une certaine probabilité  $z$ , jugée suffisante et nécessaire pour la condamnation. Cela étant, la chance  $u$ , qu’un juré pris au hazard sur la liste de ce département ne se trompera pas dans son vote, est la probabilité qu’il jugera celle (la probabilité) de la culpabilité de l’accusé égale ou supérieure à  $z$  si elle l’est effectivement, ou bien, inférieure à  $z$ , si en effet, elle n’atteint pas cette limite. »

probability of guilt for the accused should actually attain or exceed a value  $z$ . Several other of **Poisson**'s concepts are subject to similar concerns. Perhaps we should leave off trying to follow the threads of his investigation. There could hardly be a publication which shows us so instructively, how the most ingenious investigations can be led so far astray by lack of theoretical clarity. Once accustomed to associating a precise notion with the word 'probability', one is rendered unable to read the expository parts of **Poisson**'s book (as opposed to the computational parts). One is unable to read them without being beset with anxiety that none of it succeeds in making clear just what he is talking about.

**14.** An overview of the domains to which the probability calculus has been applied, brings us to another question. The question is how reasonable it might be to describe the conditions on which the usefulness of those methods depend. In the present work, general conditions that are conducive to probability in numbers have been described in logical terms already. Whether such logical conditions hold for attempts to gain knowledge in fresh subjects – that must depend principally on the objective character of the subject, together with the nature of tools we have for its investigation. From this perspective it may not be too much to characterize the domains to which the probability calculus may be applied. A few remarks may be in order in that regard.

Initially the Range principle can only play a significant role where there is some reason that exact and detailed knowledge is unattainable. The numeric formulation of probability is connected to the situation that certain factors tending to the outcome alternate, representing substantial variation in behaviors constrained by overall circumstances. If we survey the ways and means by which this characteristic type of dependence is brought to pass, different cases may be distinguished.

The simplest is when something may only transition over a very specific and limited class of states; a large manifold of influences only brings about repetitions of these same varieties of behavior. For games of dice and coin tosses, this circumscribed nature of the states is given by constraints on the orientations which are possible. In the final analysis they are given by the constitution of space itself. However many turns we take with a die or with a coin, we only have repetitions of the same definitive possibilities of final position. There are many other ways to achieve something similar. If a gas molecule is enclosed in a space bounded by rigid walls, in a very similar way we depend on the following for an assignment of probability which we may make about its locus in this or that region of space. That is, a great number of different types of effects to which the molecule may be subject by collision, can only lead to a definite and limited class of behaviors. If we think of all the variations in changing influences, behaviors exhibit continual repetition as a consequence.

While here the individual character seems to be tied to bounds on whatever is possible, in other cases we are led to think quite another way. Let us consider the bowling – game which has been described. We see its specifiability in numbers depends on a specific objective arrangement: that within a defined spatial extent there are real objects of exactly the same kind (the black and white stripes of the alley) in continual alternation. We find the same factor again when we deal with the

drawing of balls from an urn, insofar as it can be presumed in advance that black balls and white balls are present in approximately equal numeric ratio in small regions of the urn. We saw too, that when we are dealing with the chance of errors in observation, or with the risk of contracting a disease, that the frequently repeated presence of the same kinds of things, or the occurrence of the same kinds of events, is considered a reason or a condition for probabilities to be represented in numeric terms. We see very clearly what a very special arrangement in the real world is necessary for this to occur. Anywhere that this is not realized, or where it is realized only incompletely, our expectations cannot be specified in numbers, or not exactly. Sometimes things are not just of the same kind *de facto*, but instead more or less similar things are just repeated over time or across space. Sometimes the relative extent in which their presence and their absence alternate with one another is no definite ratio but rather a more or less fluctuating one. Then by the nature of this enterprise, there is no definite number to be given as a precise and accurate expression of probability. Very often it may mean chasing a will-o-the-wisp, if one sets oneself the task of deriving specific values for probability in numbers. Very often general determinants that are accessible or known to us, are not at all conducive to precisely reportable probabilities of outcomes. We would do well to maintain that when we speak (for example) of a specific attentional state or the skill of an observer, or else of a certain general state of health, the kinds of objective processes which are indicated do not actually admit an exact classification and comparison of ranges. By the nature of this enterprise, any probability based on them cannot be given precise specification in numbers. Then the specification of probabilities cannot be a general aim of research. Rather, there are only very special objective conditions which ever enable that to be accomplished.

Whether the examination of ranges proves to be of central importance, depends on other circumstances. This must be the case chiefly where it offers high probabilities, and for such premises as have not been disconfirmed in an offhand way, but for premises connected to issues of major interest. The Range principle always produces a large probability for varieties of behavior that are very general (they span a large range). An essential condition for useful application of the principle seems to be that fundamental and univocal importance is associated with the general determinants. This is shown clearly where a sum or an average over very many cases of the same kind is in question. For example, the expectation that for 1000 throws of a roulette wheel, the ball lands on red 500 times, covers a number of different possibilities. The possibilities are represented by various series of 500 throws of one sort and 500 throws of another sort. All are subsumed under the general concept. For deep-seated reasons, these collections of results are of essential and independent interest in such conditions. The probability calculus owes to those circumstances the prominence and importance given to it. The principle becomes even more prominent, if total outcomes are the only thing of which we have any knowledge. Then our interest may be entirely and exclusively focused on them. Such is the case in molecular physics. According to everything we know, any uniform body of perceptible size contains an immense number of molecules of just the same kind. Their exact arrangement is indeterminable for us, and the imprecision of our knowledge leaves

room for an operational range of this kind. So from a limited number of determinants, extremely large probabilities are produced for generally characterized factors of overall behavior – as for example, the attainment of equilibrium in temperature in small regions of space. Even for future events, we can only form expectations of great probability that are expressed in general terms, such as that the temperature of a body will rise or fall, not conjecturing anything in detail about the interactions of molecules. Once again, those total phenomena are the only things we can observe. Hence they are the only things we have pronounced interest in making conjectures about.

These remarks make it clear that even from the perspective which has been presented, there are only specific subjects where application of the probability calculus has marked importance. Other exceptional subjects (as have been of interest here) offer this in a more or less approximate way, if at all. It is in the nature of the subjects themselves that a deep or a superficial exploitation of the probability calculus finds its justification.

# Chapter 10

## On the History of Probability Theory



**Abstract** The history of probability begins with Pascal and Fermat on games. This history of probability delineates changes in our understanding of basic concepts of probability, rather than tracing general advances in methods. Pragmatic value is central at first, and probability is soon applied to the calculation of annuities. The significance of the probability calculus is thought to be very broad, as a ‘general logic of uncertainty’. The greatest of hopes is placed on its extended application. Yet the literature of the eighteenth century is full of critiques of orthodox accounts. Mistakes in the orthodox approach have emerged by acceptance of the principle of indifference and the habit of interpreting all series of phenomena in the same way as games of chance, as outlined in the work of Cournot, Mill, Lotze, and Sigwart. The task of establishing a formal discipline on theoretical principles has not been given much notice by mathematicians, in general.

**Keywords** History of probability · Annuities · Equal possibility · Mathematical probability · Chance and probability · Unsaturated judgment · Disjunctive judgments · Principle of insufficient reason

1. From the perspective of our investigation, a glance over the history of the probability calculus is not without interest, since – disregarding general advances in mathematical methods – changes may be traced in the ideas that determine our understanding of the probability calculus. I believe a few remarks on historical development will be appropriate here, since that aspect has been given only cursory attention in **Todhunter**<sup>1</sup>’s comprehensive work.<sup>1</sup>

The investigations conducted by Pascal and Fermat are usually considered the bare beginnings of probability calculus. They are contained in the exchange of letters

---

<sup>1</sup> **Todhunter**, I. (1865). *History of the mathematical theory of probability from the time of Pascal to that of Laplace*. Cambridge and London: Macmillan and Co. (Reprinted in 1949 by New York: Chelsea Publishing Company).

between these two learned authors.<sup>2</sup> As so frequently later, already primary emphasis is placed on the solution of specific problems and hence on mathematical methods. By contrast, any expression of an interpretation given to the whole procedure is missing entirely. That much might seem sufficiently clear in restriction to games of chance, without discussion of detail. However, if we observe the verbal form in which problems and solutions are usually cast, then we find – what is of more than a little interest – that value in a pragmatic sense is most often held to be the thing to be determined. Every chance appears to be the chance of profit, and its determination is nothing but the ascertaining of a concrete possession equivalent to it. The questions treated by **Pascal** and **Fermat** relate primarily to the so-called *Problème des parties*, that is, to the division to be made of stakes in a game when two players agree to break off the game at some arbitrary point. Then it is asked: what division is ‘just’, or what ‘value’ the momentary state of the game represents to each player. Very occasionally a turn of phrase is found there: « *le hazard est égal* »,<sup>3</sup> which one may consider the first appearance of the concept of equally possible cases. The same formulation tied to contexts of pragmatic importance is also commonly found in the work of authors who succeeded them, as in: “the expected value will be  $\frac{a+b}{2}$ ” [expectatio dicenda est valere  $\frac{a+b}{2}$ ], or “if we wish to divide the money exactly equally” [si pecuniam prout aequum est partiri velimus]<sup>4</sup> and the like. It is noteworthy that very much later **Bayes** – who gives an entirely general definition of probability – begins with the same perspective. His somewhat cumbersome explanation reads: “The probability of any event is the ratio between the value at which an expectation, depending on the happening of the event, ought to be computed, and the value of the thing expected on its happening.”<sup>5</sup>

2. In the year 1654 only a short time after the works by **Pascal** and **Fermat** (just cited), the probability calculus got a boost from quite a different quarter. With that, questions were once again initially posed about value in a pragmatic sense. People began to look for a definite and rational measure for values of annuities. This was the question to be answered: given that an individual is now of a certain age, which allowance should be given in an economical way or paid in a rational manner towards a liability to pay the person some amount annually as long as that person lives? It is clear that in addressing this question one runs into another: the probability with which an individual who is now of a certain age may be expected to have a

<sup>2</sup>**Pascal**, B. (1779). *Œuvres, tome 4*. La Haye: chez Detune, pp. 412–443, and Todhunter (1865, Chapter 2).

<sup>3</sup>Op. cit. p. 414.

<sup>4</sup>**Christiaan Huyghens**, *Tractatus de ratiociniis in ludo aleæ*, proposition I.IV. The same thing stands in part I of **Jacob Bernoulli**’s *Ars conjectandi*. [Hugenii, C. (1657). *De ratiociniis in ludo aleæ. Exercitationum Mathematicarum, liber 5*. Leyden: Francisci à Schooten, pp. 521–534. See volume 14 of his *Works* (1655–1666/1920); Bernoulli, J. (1713). *Ars conjectandi* (opus posthumum). Basilieæ: Thurnisiorum, Fratrum].

<sup>5</sup>**Bayes**, 1763. An Essay towards solving a problem in the Doctrine of Chances. By the late Mr. Bayes F.R.S., communicated by Mr. Price, in a Letter to John Canton, A.M.F.R.S. *Philosophical Transactions of the Royal Society of London*, 53, 370–418 (see also Bayes 1764).

longer or a shorter lifespan. As **Todhunter** reports, the first attempts to answer this question in a rational way were made by **John Graunt** (1662)<sup>6</sup> and by **Jan de Witt** (1671).<sup>7</sup> There is a work dated 1693 by **Halley**<sup>8</sup> which handles this subject. In the latter, “the Chances of Mortality at all ages and likewise how to make a certain Estimate of the value of Annuities for Lives etc.” are ascertained from statistical data on the city of Breslau. In that work the expression used most often is either “the odds” (for example, “the odds that any person does not dye before he attain any proposed age”) or else “Chances”. By contrast, “probability” is still not found anywhere.

**3.** In short, application of the probability calculus in such heterogeneous domains as games of chance and mortality tables was enough to make its significance seem exceptionally broad in nature. Certainly we see how in only a few short steps a theoretical interpretation is reached, to which a great number of writers – principally mathematicians – have been converted up to the present day in essentially the same way. At the same time we see that methodology constantly proliferates, and a solid tradition forms which is authoritative – at least for this circle of authors. Perhaps its internal consistency permits us to call its theories and procedures the ‘orthodox’ tradition, as we have done. **Laplace** can be called the principal representative of this theoretical push; his *Théorie analytique des probabilités* is still the most comprehensive and important work in this category. Next let us examine this orthodox treatment of the theory of probability.

If one wants to characterize briefly the principal stance which was maintained, one cannot but concede that it was a logical stance. In fact as early as **Jacob Bernoulli**’s *Ars conjectandi*,<sup>9</sup> the view is set out fully and clearly that the probability calculus is a general logic of uncertainty. Its applicability is deemed expressly to be universal. This view can also be attributed to the somewhat exuberant way **Bernoulli** appraises the importance and value of the subject in his General Introduction. Let us cite a passage of *Ars conjectandi* as being characteristic, where we read:<sup>10,11</sup> “Although so far this difficult subject would [seem to] be a mathematical consideration (since it is determined

<sup>6</sup> **John Graunt** (1662). *Natural and political observations made upon the bills of Mortality*. London: Thomas Roycroft.

<sup>7</sup> **Jan de Witt**, *De vardye van de lifrenten*. [de Witt, J. (1671). *Waardije van Lyf-renten naer Proporcie van Los-Renten* (The value of life annuities compared to redemption bonds). In’s Graven-Hage: Jacobus Scheltus.]

<sup>8</sup> **Halley** (1693). An Estimate of the Degrees of the Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslaw; with an Attempt to ascertain the Price of Annuities upon Lives. *Philosophical Transactions of the Royal Society of London*, **17**(196). 596–610.

<sup>9</sup> This appeared in 1713, after the author’s death.

<sup>10</sup> Quamquam hoc negotii eatenus sit considerationis Mathematicae, quatenus in subducendo calculo terminatur, si tamen usum et necessitatem speces, universale prorsus est et ita comparatum, ut sine illo nec sapientia Philosophi nec Historici exactitudo, nec Medici dexteritas aut Politici prudentia consistere queat.

<sup>11</sup> In the proemium to the second part.

by a calculation), nevertheless, if we consider [its] use and necessity, it is entirely universal and thus shared; for without it we could establish neither the wisdom of the philosopher, nor the accuracy of the historian; neither the skill of the doctor, nor the prudence of the politician.” Naturally such an evaluation necessarily involves the premise that the procedure is not considered a determination of values in the pragmatic sense as before, but rather a method of essentially psychological significance. The fourth part of *Ars conjectandi* contains a more detailed exposition of the theory. There light is cast on the subjective meaning of statements about probability (which **Pascal** and **Fermat** do not mention at all).<sup>12</sup> There it says: “The things seen by us in succession do not all have the same certainty, but rather, this certainty varies widely by a greater or lesser magnitude.” Further: “Probability . . . is a *degree* of certainty, and differs from certainty as a part differs from the whole. Indeed, if there is complete and absolute certainty — which we designate with the letter A, or the unit 1 — when for example, five should stand for the probabilities or parts (i.e., possibilities) from which three possibilities entail a particular future event, while the rest do not, then that event is said to have  $\frac{3}{5}$  A, i.e.,  $\frac{3}{5}$  of an absolute certainty.” . . . “To make a conjecture about anything is to measure its probability; therefore, the art of conjecture, or guessing, we define as the art of measuring the probabilities of things as exactly as possible. We do this in order always to be able to choose and follow, in our judgments and actions, that which is determined to be better, more satisfactory, more prudent, and better-advised. The wisdom of every philosopher and the prudence of every politician should be founded on this alone.”

It is important to note the manner in which the measurement of probability is explained, that is, objective meaning displaces subjective meaning in a way that goes quite unremarked.<sup>13</sup> “Probabilities are estimated from both the number and the weight of arguments . . . By ‘weight’ I mean the strength of the proof . . .”. The latter, the strength of proof [*vis probandi*], is then explained as follows: “One can weigh the strength of proof by considering many cases, where it is indicated that a thing may or may not occur – or even where the contrary of the thing is indicated.” Right afterwards it is said: “Moreover I posit that all possible cases may exist or occur with equal facility. Indeed, caution is to be

<sup>12</sup>Certitudo rerum . . . spectata in ordine ad nos, non omnium eadem est, sed multipliciter variat secundum magis et minus. . . . Probabilitas . . . est gradus certitudinis et ab hac differt ut pars a toto. Nimurum si certitudo integra et absoluta, quam litera a vel unitate 1 designamus, quinque verb. gr. probabilitatibus seu partibus constare supponatur, quarum tres militent pro existentia aut futuritione alicujus eventus, reliquae contra, eventus ille dicetur habere  $\frac{3}{5}$  a, seu  $\frac{3}{5}$  certitudinis . . . Conjicere rem aliquam est metiri ejus probabilitatem; ideoque Ars conjectandi sive stochasticis nobis definitur ars metiendi quam fieri potest exactissime probabilitates rerum, eo fine ut in judiciis et actionibus nostris semper eligere vel sequi possimus id, quod melius, satius, tutius aut consultius fuerit deprehensum; in quo solo omnis Philosophi sapientia et Politici prudentia versatur.

<sup>13</sup>Probabilitates aestimantur ex numero simul et pondere argumentorum . . . Per pondus autem intelligo vim probandi. The latter, the *vis probandi*, is then explained as follows: vim probandi pendere a multitudine casuum, quibus illud existere vel non existere, indicare vel non indicare aut etiam contrarium rei indicare potest. . . . Pono autem, omnes casus aequae possibles esse seu pari facilitate evenire posse; alias enim moderatio est adhibenda et pro quovis casu faciliori tot alii casus numerandi sunt, quoties is ceteris facilius evenit: ex. gr. pro casu triplo faciliori numero tres casus qui pari cum caeteris facilitate contingere possint.

applied otherwise, and whenever some cases occur more easily than others, many additional cases are to be counted in the same manner as the earlier case, so long as that case occurs more easily than all the others: for example, three cases which occur with equal facility as some triplet of the other cases.”

Finally here, quite likely one notes that in this whole formulation, the numeric ratio of the various equally possible cases is basically considered to be an objective ratio. Perhaps this emerges even more clearly and prominently, in that the empirical determination of probability – which was familiar to **Bernoulli** with regard to ratios of mortality – is considered the ascertainment of those sorts of numeric ratios of equally possible cases.<sup>14</sup> “Because it may not be taken from the *a priori*, we may take it from the *a posteriori* – that is, from observing similar examples of the event many times – because with so many cases, we ought to presume that the event will occur or not occur in later cases just as often as it has been observed to occur or not to occur previously in a similar situation.”

**Bernoulli** finds no problem with the idea that right here the observation of a large number of cases puts us in the same epistemological situation which we may find ourselves in with games of chance. This shows just how the ‘equipossibility’ of cases is taken in an objective sense. The problem with describing this interpretation in a satisfying way, is that it forces one to put the subjective sense of probabilities in the foreground. An even sharper turn in this direction is clearly exhibited in the concept as described by subsequent authors.

In fact what was needed to motivate this was only the hint – which has been repeated so often against the mistaken intuition offered by naïve and untutored thought – that outcomes are objectively and completely determined by the totality of constraining circumstances; that is the reason one case is certain while all others are ruled out as impossible. Perhaps the earliest authors on the subject of the probability calculus should not perhaps have been in the dark on this point.<sup>15</sup> But given this insight, the relation of what is equally possible to an individual’s knowledge follows directly. This is clearly formulated in what **Daniel Bernoulli**<sup>16</sup> says, where sure enough he speaks of an hypothesis:<sup>17</sup> “The whole proof of this proposition . . . may be seen in the initial hypothesis, since where there is no reason to expect one thing over another, each of the parts would have to be judged to be equal.” – After that, the article “*Probabilité*” in the Encyclopédie (**Diderot** and

<sup>14</sup>Quod a priori elicere non datur, saltem a posteriori, hoc est, ex eventu in similibus exemplis multoties observato eruere licebit; quandoquidem praesumi debet, tot casibus unumquodque posthac contingere et non contingere posse, quoties id antebac in simili rerum statu contigisse et non contigisse fuerit deprehensum.

<sup>15</sup>Todhunter (1865, p. 4) cites an interesting passage where **Kepler** expresses this interpretation with all the clarity and precision one could want.

<sup>16</sup>Daniel Bernoulli, *Specimen theoriae novae de mensura sortis*. *Comment. Acad. Imp. Petrop.*, 1730 and 1731. [Bernoulli, D. (1738). *Specimen theoriae novae de mensura sortis*. *Commentarii Academiae Scientiarum Imperialis Petropolitanae (In Classe Mathematica)*, 5, 175–192. (contributions submitted in 1730 and 1731)]

<sup>17</sup>Demonstraciones hujus propositionis . . . omnes videbis hac inniti hypothesi, quod cum nulla sit ratio cur expectanti plus tribui debeat uni quam alteri, unicuique aequae sint adjudicandae partes.

**d'Alembert**, 1751–1766) emphasizes the subjective meaning of probability much more sharply than **Jacob Bernoulli** does. Certainly the article begins by proceeding from the notion of ‘value’ in a pragmatic sense. “Let us suppose that I have just been told I have won a jackpot of ten thousand livres in a lottery. I doubt the truth of this announcement. Someone nearby asks me what amount I am willing to give him to insure the prize. I offer him half, meaning that I consider the probability of the announcement as only half-certain.”<sup>18</sup> – Subsequently however this is ignored, and the probability calculus is established simply as a logical method of universal application. “This means of determining by probability the relation of the causes which give rise to an event to those which undermine it – or more generally the proportion of reasons or conditions which establish the truth of a proposition to those which give rise to its contrary – can be applied to anything which may happen or may not happen, to anything which could be or could not be.”<sup>19</sup> And immediately afterwards one finds the precise explanation – in a subjective sense – of equally possible cases. “I have said this principle would be employed if we assumed the various cases to be equally possible. And in effect, it is only by assumption relative to our bounded knowledge that we say, for example, that any side of the die may equally well turn up; it is only when it is rolled in its dicebox that it may have the disposition for what must happen when, together with the dispositions of the dicebox, the felt of the table, or the force and manner in which the die is thrown: all that makes it certain what happens; but since all that is wholly unknown to us, we have no reason to prefer one side over another. As a consequence we assume it is equally easy any of them should turn up.”<sup>20</sup>

And with that the stance which **Laplace** has taken has essentially been reached, too. In his writings we find the following short explanation: “equally possible cases, which is to say those for which we may be equally undecided about their existence.”<sup>21</sup> That is a notion which coincides with what we have called the Principle of insufficient reason [i.e., the Principle of indifference].

<sup>18</sup> « Supposons que l'on vienne me dire, que j'ai eu à un loterie un lot de dix mille livres; je doute de la vérité de cette nouvelle. Quelqu'un qui est présent me demande, quelle somme je voudrais donner pour qu'il me l'assurât. Je lui offre la moitié, ce qui veut dire que je ne regarde la probabilité de cette nouvelle que comme une demi-certitude. »

<sup>19</sup> « Cette manière de déterminer probablement le rapport des causes qui font naître un évènement à celles qui le font manquer, ou plus généralement la proportion des raisons ou conditions qui établissent la vérité d'une proposition avec celles qui donnent le contraire, s'applique à tout ce qui peut arriver ou ne pas arriver, à tout ce qui peut être ou ne pas être. »

<sup>20</sup> « J'ai dit que ce principe s'employait quand nous supposions les divers cas également possibles. Et en effet ce n'est que par supposition relative à nos connaissances bornées que nous disons, par exemple, que tous les points d'un dez peuvent également venir; ce n'est pas que quand ils roulent dans le cornet celui qui doit se présenter n'ait déjà la disposition qui, combinée avec celle du cornet, du tapis, ou de la force et de la manière avec laquelle on jette le dez le doit faire sûrement arriver; mais tout cela nous étant entièrement inconnu nous n'avons pas de raison de préférer un point à un autre; nous les supposons donc tous également faciles à arriver. »

<sup>21</sup> « cas également possibles, c'est à dire tels que nous soyons également indécis sur leur existence, »

4. This is not the place to delve into much detail to say how well this paralleled the spirit of that previous century, in placing the greatest of hopes on extended application of the probability calculus – to expect of it what might be termed a thoroughly rational classification of human thought. It is only that broad current of thought and culture which explains the extraordinarily intensive engagement with our subject which existed at that time. Some words of Condorcet's<sup>22</sup> are particularly characteristic of this, as in the following: "There is no one who has not observed in himself, that he has changed his mind about certain objects, according to his age, circumstances, or other events, without however being able to say that this alteration has been based on fresh intentions, and without being able to assign another cause than the more or less salient impression made by those same objects. But if instead of judging by that impression which multiplies or exaggerates a part of the objects, while it attenuates others or hides them from view, we could count or evaluate them by calculation, our thinking would cease to be the slave of our impressions."<sup>23</sup>

As these extremely general formulations of the principle and of the task of the probability calculus emerge, at the same time they are also applied to areas of investigation which really are new. The most important of them may be alluded to here: errors of observation, which Simpson<sup>24</sup> and Laplace<sup>25</sup> were first to subject to the probability calculus; decisions based on a majority of votes, which Condorcet<sup>26</sup> was first to treat by methods of the probability calculus; and finally the statistics of disease cures, which Black<sup>27</sup> may have been the first to try to evaluate in this

<sup>22</sup> Condorcet, *Essai sur l'application de l'analyse à la probabilité des décisions*. Discours prélim. CLXXXV. [de Condorcet, N. (1785). *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. À Paris, de l'imprimerie royale.]

<sup>23</sup> « Il n'y a personne qui n'ait observé sur lui-même, qu'il a changé d'opinion sur certains objets, suivant l'âge, les circonstances, les évènements, sans pouvoir dire cependant, que ce changement ait été fondé sur de nouveaux motifs, sans pouvoir y assigner d'autre cause que l'impression plus ou moins forte des mêmes objets. Or si au lieu de juger par cette impression qui multiplie ou exagère une partie des objets, tandis qu'elle atténue ou empêche de voir les autres, on pouvait les compter ou les évaluer par le calcul, notre raison cesserait d'être l'esclave de nos impressions. »

<sup>24</sup> Simpson, T. (1755). A letter to the Right Honourable Earl of Macclesfield, President of the Royal Society, on the Advantage of taking the Mean of a number of Observations in practical Astronomy. *Philosophical Transactions of the Royal Society of London*, **49**, 82–93.

<sup>25</sup> Laplace, P.-S. (1774). Déterminer le milieu que l'on doit prendre entre trois observations données d'un même phénomène. *Mémoires des savans étrangers*, **4**. [Cf. Laplace, P.-S. (1904). *Oeuvres complètes*. Paris: Gauthier-Villars.]

<sup>26</sup> Condorcet, *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. 1785.

<sup>27</sup> Black, W. (1789). *Analyse arithmétique et médicale des maladies et de la mortalité de l'espèce humaine 2<sup>me</sup> éd.* (1789, *An arithmetical and medical analysis of the diseases and mortality of the human species*. London: J. Johnson.); I find this work – which is not known to me – cited in Lacroix's (1816) *Traité élémentaire* . . . , p. 200. [Lacroix, S.-F. (1816). *Traité élémentaire du calcul des probabilités*. Paris: M<sup>me</sup> V<sup>e</sup> Courcier.] Among the works which are now better known, the oldest which deals with these subjects is Gavarret, J. (1840). *Principes généraux de statistique médicale, ou Développement des règles qui doivent présider à son emploi*. Paris: Bechet jeune & Labé.

manner. The collective phenomena of society are not to be mentioned separately once again, since application of the probability calculus to problems of longevity was mentioned already as one of its earliest applications. Still, it should be remarked that the scope of this kind of application soon expanded to a huge extent, as investigation by similar means spread from mortality rates, to births, marriages, the commission of crimes, and so on.

It is easy to say that the generality of the principle was not an insignificant factor in this expansion of the use of the probability calculus: there could hardly be a domain of investigation, to which the Principle of insufficient reason could not assign an example of equally possible cases.

Yet meanwhile one must admit that the sort of application of the probability calculus, as developed in the orthodox tradition of the academy, does not find a substantive explanation in this central idea. In fact the Principle of insufficient reason even leads necessarily – once correctly and fully analyzed – to the realization that in many cases the establishment of numeric probabilities cannot succeed, if only because of the enormous complexity of possible assumptions. Then the principle can only be held responsible in a negative way for the legion of mistaken applications which would have been useful, just insofar as it allowed constraints on their correct application to go unrecognized.

Furthermore one can claim that even in orthodox applications of the probability calculus, the consequences of the principle were never effectively elaborated. Often this made itself felt directly, in that there lay a significant difficulty in determining what is ‘equally possible’. So Laplace says: “If they are not (if the cases are not equally possible), next one will determine their respective possibilities, the accurate assessment of which is one of the finer points of the theory of chance.”<sup>28</sup> In particular, perhaps the notion of a probability which can be attributed to certain objective ratios in themselves has always been received as valid. In a strict sense, this is shown by all those investigations which seek to establish *a posteriori* what the ‘actual probability’ of an event may be. That is because this way of putting the question is of course easily and satisfactorily construed if the objective factor which determines probability can be immediately and transparently cited, as for example the numeric ratio of black and white balls contained in an urn. By contrast it makes no sense to say that there is a clearly expressible probability – following the Principle of insufficient reason – that is a probability of death for a particular individual, or else to speak of ‘determining’ that probability. Admittedly, superficial deliberation could have led to a glimpse of that sort of probability, which represents a very singular state of knowledge for us – a state at the limit of our knowledge. However, explanation of the relevant probability in this sense is not forthcoming, so far as I know. Yet it may not be uncommon for a fitting explanation to be given for these probabilities (or other similar ones) in terms of the relevant events themselves without any reference to our state of knowledge – in blatant contradiction to the general principle.

---

<sup>28</sup> « s’ils ne le sont pas ‘(si les cas ne sont pas également possibles)’ on déterminera d’abord leurs possibilités respectives dont la juste appréciation est un des points les plus délicats de la théorie des hazards. »

The following passages from **Poisson**'s *Recherches* serve as evidence: there the distinction between ‘chance’ and ‘probabilité’ is discussed in what one could call a very whimsical way.<sup>29</sup>

In ordinary language, the words ‘chance’ and ‘probability’ are almost synonymous. Most often we will use one for the other indifferently; but as it becomes necessary to draw a distinction between their meanings, in the present work we will use the word ‘chance’ for events in themselves, independently of the knowledge we have of them, and we will retain the earlier definition for the word ‘probability’. Hence by its nature an event will have a chance which is larger or smaller, known or unknown; while its probability will be relative to our knowledge of it.<sup>30</sup>

Many similar passages can be found, and it is conceivable that the fundamentally objective meaning which was to be attributed to certain statements of probability, made itself felt even more strongly, the more often they were paired with others which dispensed with this meaning. Some sense of this distinction could not have been lost entirely on those authors who count as being in the mathematical camp.

Now the characteristic manner in which one was given to apply the probability calculus, is to be attributed to a very large extent to the habit of interpreting all possible series of phenomena in the same way as in games of chance – at any rate to a much larger extent than any clearly-conceived and well-considered theoretical framework is ever mimicked ordinarily. There is nothing astonishing about this phenomenon. How often do we see that procedures which are applied correctly and usefully in one domain, are transposed to others where they do not fit, and where they lend themselves to outright self-deception! All kinds of circumstances came together to facilitate this common habit. Above all it should be mentioned here, that as statistical investigations flourished to greater and greater extent, an approximate constancy of average results always seemed to be produced ever more fully and generally. They seemed to be produced whenever sufficiently large numbers of observations could be made available. Naturally it was very important that even the theoretical understanding of games of chance was itself incomplete, and consequently no adequate reply could be given to the question whether any given series of phenomena would be suited to logical analysis in the same way – whether an adequate convention existed. And yes – one must even say this question never was neglected, but the wrong answer was almost always given to it. That is, one just believed that that logical treatment had to be introduced anywhere that numberless

<sup>29</sup> **Poisson**, *Recherches sur la probabilité des jugements*, p. 31. [Poisson, S.-D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris: Bachelier, Imprimeur-Libraire.]

<sup>30</sup> « Dans le langage ordinaire les mots chance et probabilité sont à peu près synonymes. Le plus souvent nous emploierons indifféremment l'un et l'autre; mais lorsqu'il sera nécessaire de mettre une différence entre leurs acceptations, on rapportera, dans cet ouvrage, le mot chance aux évènements en eux-mêmes et indépendamment de la connaissance que nous en avons, et l'on conservera au mot probabilité sa définition précédente. Ainsi, un évènement aura, par sa nature, une chance plus ou moins grande, connue ou inconnue; et sa probabilité sera relative à nos connaissances en ce qui le concerne »

unknown details coalesce to produce an outcome. In the process one neglected to take account of the most important factor to which any calculation about a game of chance is subject, which is implicit in the enduring constancy of certain quite definite relations of magnitude. Consequently an only more or less distant similarity which held between games of chance and many other phenomena could then be mistaken for an exact correspondence. Attempts to give a theoretical basis for this parallelism are made but rarely, and then they are only sought along narrow paths, namely with respect to the constancy of general conditions. Undeniably it is nothing but a fantasy exercise to prove the existence of such a thing in some completely general and purely conceptual way. Nevertheless we encounter many arguments which may be considered attempts to construct such a proof in wholly general terms.<sup>31</sup> By contrast, the properties of independence and equality of chances for individual cases – which had been presumed for reasons which were just as generally unfounded – cannot appear to be necessary marks of any series of similar cases, from any theoretical perspective. The unjustified assumption of such properties is merely the glimpse of an illusion – an illusion fostered by the similarity to games of chance.

If we find this reduction of all series of phenomena to the schema of an ordinary game of chance to be the most prominent identifying characteristic of the methods of mathematical orthodoxy, then in no sense should it be said that those theoretical ideas never determined particular features of procedure, at least occasionally. Notably, one could occasionally appeal to the statement that even ‘objective’ probabilities might be nothing else than the expression of our inexact knowledge, and from that it might seem justified to consider any probability founded on the Principle of insufficient reason to be equivalent to them, and then to combine them lavishly. Among the latter probabilities, the most important ones, at least the most commonly applied, would be those which were characterized as *a priori* probabilities founded on the Principle of insufficient reason. Following the rule established by **Bayes**, one became used to assuming that before an observation is given, any value of an (objective) probability is equally likely. Then once this rule was established, a method was garnered which could always be applied without any further thought, once a large number of similar cases had been observed. If  $n$  cases came to light, of which  $m$  had taken a specific course, then one could report a definite probability for them: the probability that the course of events in question would lie between  $\frac{n}{m} + \delta$  and  $\frac{n}{m} - \delta$ . For substantial values of  $n$  this probability becomes large, even for fairly small values of  $\delta$ , and one can “determine that probability” with fair accuracy. Then that is enough to report a probability for a number of new cases, that in turn the relative frequency of the relevant course of events may lie within some bounds. It has been explained above, how the nature of this calculation and similar calculations serves to compound a number of rank illusions, so it is unnecessary to repeat the refrain.

**5.** Of course only a small number of the many kinds of considerations and objections can be contemplated here, which have been brought against the views

---

<sup>31</sup> For example **Poisson**, *Recherches* etc. pp. 132 & 144.

of the orthodox school. Even the literature of the eighteenth century is replete with work that falls under this heading. It may be permissible to bring a few of them forward, in order to characterize the theoretical controversies of that period.

What is of keenest interest is the well-known opposition posed by **d'Alembert** to representatives of the orthodox theory. In turn the most remarkable point of his opposition is his view that, for certain phenomena whose occurrence is never observed, they must not be declared to be very improbable (as probability theory does), but instead they must be declared actually impossible. If the theory of probability construed it to be possible that heads might turn up a thousand times consecutively in a coin toss, still this is lacking an empirical basis and as such is unsustainable. – “Yet speaking in physical terms, heads cannot come up consecutively more than a fixed number of times. What is that number? I wouldn’t undertake to determine it. Still I would go further, and ask why heads couldn’t come up an infinite number of times in succession, speaking in physical terms. One can only give the following reason: it is no part of nature that an effect may remain constant, forever the same, just as it is no part of nature that all men look alike, or that all trees look alike.”<sup>32,33</sup> It is noteworthy that in his attempt to justify his view, **d'Alembert** himself was led to lend no other explanation to the absence of such very frequent repetitions, but to say that things just do not behave that way. If nevertheless the basis of his whole line of argument is still the opinion that the relevant constellations of events may actually be considered to be excluded in some way – then here we are basically only dealing with the very common and incorrect conception of the Law of large numbers, which takes the improbability of varieties of behavior as lawful exclusion: that is, it takes expectations based on the Law of large numbers to be certainties.

Such theories must be considered to be public declarations of intellectual bankruptcy, according to which the probability calculus proceeds on the basis of arbitrary fictions. For example, according to **Prévost**<sup>34</sup> every assignment of probability is founded on the tacit hypothesis that various series of cases considered equally possible, must actually occur in serial order. That is an intuition which soon proves

<sup>32</sup> « Donc physiquement parlant croix ne peut arriver de suite qu'un nombre fixé de fois. Quel est ce nombre ? C'est que je n'entreprends point de déterminer. Mais je vais plus loin et je demande par quelle raison croix ne saurait arriver une infinité de fois de suite, physiquement parlant. On ne peut en donner que la raison suivante: c'est qu'il n'est pas dans la nature qu'un effet soit toujours et constamment le même, comme il n'est pas dans la nature que tous les hommes et tous les arbres se ressemblent »

<sup>33</sup> **d'Alembert**, J. le R. (1768). *Doutes et questions sur le calcul des probabilités. Mélanges de littérature, d'histoire et de philosophie. Tome 5.* Amsterdam, aux dépens de la compagnie.

<sup>34</sup> **Prévost**, P. and **l'Huilier**, S. (1796). *Mémoire sur l'art d'estimer la probabilité des causes par les effets. Mémoires de l'Académie Royal des Sciences et des Belles-Lettres de Berlin*, 6, 3–24. “When by virtue of a certain specification of causes, many events appear equally possible to us, we pretend that all these events occur successively in turn, without repetition.” [« Lorsqu'en vertu d'une certaine détermination des causes plusieurs événements nous paraissent également possibles nous feignons que tous ces événements ont lieu successivement tour-à-tour et sans répétition. »]

fundamentally inadmissible, since then how is it that such an actually nonexistent fiction can claim to be the basis of any calculation of value?

Béguin's<sup>35</sup> intuitions on this matter bear a certain similarity: "The probability calculus strikes a middle position between arbitrary fortune and physical necessity; it decides what event will be, not insofar as it is steered by chance and not insofar as it is determined by mechanical causes. Rather it supposes the event to be prescribed through laws of fairness by equitable and impartial judgment. – It is not so much that we calculate what chance will do, but rather what it ought to do, if it sowed its favours with exact impartiality."<sup>36</sup>

It is no miracle that these attempts and many similar ones exerted no influence worthy of mention on the traditions that were developing among mathematicians.

**6.** Far greater significance may be attributed to a number of works contemporary to our time. Two works may be cited first off; each contains a richness of important and valuable insights. In addition, they are even more interesting in that they exhibit a most conspicuous correspondence to one another in many aspects, although no doubt the ideas were arrived at independently, because of the timing of their publications, and also because the premises from which the two start are the most disparate one can imagine. One of the authors is the German philosopher **Fries**, whose 'Versuch einer Kritik der Principien der Wahrscheinlichkeits-Rechnung' appeared in 1842; the other is the French mathematician **Cournot** who published an *Exposition de la théorie des chances et des probabilités* (1843, Paris). Undoubtedly the strength of both these works lies in a rejection, that is, in their sharp opposition to the schematic procedures of mathematicians. In that much – in this most interesting position – both authors are in almost complete concord. Both explain that (for example) **Poisson**'s investigation of probability in judicial decisions, in the application of the probability calculus to witness testimonies, etc. and also to a large extent in their application to collective social phenomena, are either completely incorrect or else in need of significant modification. It is clearly evident that this completely accurate insight into the central issues, which both authors have in equal measure, has been conditioned by the fact that both had taken very clear notice of the essentially objective meaning attributed to propositions about probability in some applications of the methods, but which was neglected in others. As a consequence, **Fries** and **Cournot** also agree in that they distinguish different types of probability. They explain that certain probabilities are not amenable to measurement at all; and they emphasize as entirely fundamental the question whether the reporting of probability in numbers can have objective meaning, or else merely subjective meaning. Both declare that complete disregard of these differences is the principal

<sup>35</sup> de Béguin (1767). Sur l'usage du principe de la raison suffisante dans le calcul des probabilités. *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin*, 23, 382–412.

<sup>36</sup> « Le calcul des probabilités prend donc un milieu entre l'arbitraire fortuit et la nécessité physique; il décide quel sera l'évènement, non en tant qu'il est dirigé par le hazard, non en tant qu'il est déterminé par les causes mécaniques, mais en le supposant prescrit par les lois de la convenance, par l'équité d'un juge impartial ..... C'est qu'on ne calcule pas ce que le hazard fera, mais ce qu'il devrait faire, s'il distribuait ses faveurs avec une exacte impartialité. »

error in the procedures put forward by **Laplace** and **Poisson**. The agreement between these two authors goes as far as these points; we have to endorse these points in that much as well. But we find their stance about which probabilities may actually be considered to be measurable quite different. **Fries** describes them – he calls them ‘mathematical probabilities’ in contrast to ‘philosophical’ probabilities – as follows: “Conclusions in mathematical probability are of an entirely different nature. Here we must be cognizant of a certain range of knowledge, and of the most general of laws; yet the determination of individual cases still depends on other influences, which do not follow from a single rule, and for whose variations we know no law. So we know that we might have no fixed rule to predict a single event that we should like to determine in advance. However, within the scope of our fallible knowledge, we look to secure an overview of all possible cases which remain, and from that we attempt to derive a claim about the individual case.” ... And further on: “The conditions under which a probability may be held to be subject to calculation, are indicated in what has just been discussed. Consider a question about an event, or a series of events, about which we know or we postulate that they exist in proportions which generally remain the same, so that a certain scope of knowledge is encompassed. However, other variable conditions must be added to distinguish individual cases. Of course I am in no position to determine the individual cases, but in general I am in a position to enumerate how many kinds they can only belong to, and how often one kind may occur in proportion to the others, in general.<sup>37</sup> – No special proof is required to show that this explanation leaves us without any satisfactory idea what should be meant in an objective sense by numbers which stand for probabilities. An attempt is made to describe and to portray (as much as is possible) the oddity which is clearly sensed in certain cases of probability, but the attempt does not succeed in rising above some altogether unformed and general notions. How far this lies from a solution to the problem, is shown most clearly by the turn of phrase given in conclusion to the passage just cited: I cannot ‘enumerate’ how often one thing or another (the variable conditions) will occur; instead, that can only be concluded somehow. There the task would be to discover on just what basis or on just what principle this conclusion can be drawn.

**Cournot**’s theoretical intuitions are perhaps even more remarkable, inasmuch as the various possibilities of concurrence, the « *concours des causes indépendantes* » play an essential role. “The notion of chance is that of a concurrence of independent causes towards the production of a determinate event. The combinations of various independent causes which contribute equally to the production of the same event are what we should understand as the chances of this event occurring.”<sup>38</sup> Then connected to that notion of chance, there follows a definition of mathematical

<sup>37</sup> Op. cit. **Fries**, p. 17ff. [Fries, J.F. (1842). *Versuch einer Kritik der Principien der Wahrscheinlichkeitsrechnung*. Braunschweig: Friedrich Vieweg und Sohn.]

<sup>38</sup> « L’idée du hazard est celle du concours des causes indépendantes pour la production d’un événement déterminé. Les combinaisons des diverses causes indépendantes qui donnent également lieu à la production d’un même événement sont ce qu’on doit entendre par les chances de cet événement. »

probability as: “the ratio between the number of chances which favor the event, and the total number of chances”.<sup>39</sup> In addition it is said: “that this can be considered to measure the possibility of the event, or the ease with which it is produced. Similarly in this sense, mathematical probability expresses a ratio which stands outside the mind that conceives of it – a law to which phenomena are subject, and whose existence does not depend on the range or the bounds of our knowledge about the circumstances of their production.”<sup>40,41</sup>

Then in contrast to this mathematical probability it is said of subjective probability that: “If, in the imperfect state of our knowledge, we have no reason to suppose that one combination will arise rather than another . . . and if we understand by the probability of an event, the ratio between the number of combinations which favor it, and the total number of combinations which we place in the same series, then given nothing better, this probability could yet serve to determine the conditions of a wager, or of some sort of lottery scheme, but that would no longer represent a ratio which exists actually or objectively between things. It would acquire a purely subjective character, and would be susceptible to differences from one individual to another, according to the measure of their knowledge. Nothing is more important here than to distinguish carefully between the two senses of the term probability, which now is taken in an objective sense, and now is taken in a subjective sense – that is, if we want to avoid error and confusion in our explanation of the theory as well as in the applications we make of it.”<sup>42</sup>

Obviously the weakness of even this theoretical formulation lies in the determination of mathematical probability, since one is in no position to study how chances may be enumerated as combinations of causes, or what should be meant in this sense by a numeric ratio of chances in favor of an event to the total number of all chances.

<sup>39</sup> « rapport entre le nombre des chances favorable à l'évènement et le nombre total des chances »

<sup>40</sup> « qu'elle peut être considérée comme mesurant la possibilité de l'évènement, ou la facilité avec laquelle il se produit. En ce sens pareillement la probabilité mathématique exprime un rapport subsistant hors de l'esprit qui le conçoit, une loi à laquelle les phénomènes sont assujettis, et dont l'existence ne dépend pas de l'extension ou de la restriction de nos connaissances sur les circonstances de leur production »

<sup>41</sup> Op cit. Cournot, pp. 437 & 438. [Cournot, M.A.A. (1843). *Exposition de la théorie des chances et des probabilités*. Paris: L. Hachette.]

<sup>42</sup> « Si dans l'état d'imperfection de nos connaissances nous n'avons aucune raison de supposer qu'une combinaison arrive plutôt qu'un autre . . . et si nous entendons par probabilité d'un événement le rapport entre le nombre des combinaisons qui lui sont favorables et le nombre total des combinaisons mises par nous sur la même ligne, cette probabilité pourra encore servir, faute de mieux, à fixer les conditions d'un pari, d'un marché aléatoire quelconque; mais elle cessera d'exprimer un rapport subsistant réellement et objectivement entre les choses; elle prendra un caractère purement subjectif, et sera susceptible de variations d'un individu à un autre, selon la mesure de ses connaissances. Rien n'est plus important que de distinguer soigneusement la double acceptation du terme de probabilité, pris tantôt dans un sens objectif, et tantôt dans un sens subjectif, si l'on veut éviter la confusion et l'erreur, aussi bien dans l'exposition de la théorie que dans les applications qu'on en fait. »

Very similar phrasings occur elsewhere, and not infrequently; namely those which purport to represent the fraction of probability directly as the numeric ratio of those ‘causes’ which favour (or enable) an outcome and those which hinder it. For instance, this definition can be found in the work of John Stuart Mill: “If the actuary of an insurance office infers from his tables that among a hundred persons now living of a particular age, five on the average will attain the age of seventy, his inference is legitimate, not for the simple reason that this is the proportion who have lived until seventy in times past, but because the fact of their having had so lived shows that this is the proportion existing at that place and time, between the causes which prolong life to an age of seventy and those tending to bring it to a close.”<sup>43</sup> But what meaning may one give to the proposition that the numbers (or the magnitudes?) of causes which prolong life and those which shorten it stand in a ratio of 95 : 5? How can one imagine counting effective causes in a complex collection of the most disparate things? How should one classify things with respect to a specific effect, into those which favor and those which hinder? Then it is still wholly obscure what such a ratio may be, and how it has an effect. Having said that, despite its vagueness, this interpretation of propositions about probability comes at least very close to expressing a decidedly incorrect interpretation of the Law of large numbers. One is tempted to regard it as naturally evident that over the course of an extended period – or for the totality of outcomes over many cases – the actual ratio of constraining factors will emerge, and that small deviations from that same definite value may not surpass it if they are small with respect to the total number. Then it seems that the expectation for many cases is shown to be a certain one, and to reveal a characteristic feature of propositions about probability.

7. Anyways, we have A. Fick<sup>44</sup> to thank for a much more valuable interpretation of propositions about probability. That interpretation should be appended here, since in a very explicit way it also represents the meaning of those propositions as objective. According to Fick, every proposition about probability contains the expression of both a condition and a consequence. The condition is formulated so that it covers a certain ‘domain’. “The probability of an incompletely-expressed hypothetical judgment is represented by a proper fraction of the whole domain of the condition, and it is to the realization of that condition that the outcome which is expressed as a postscript is necessarily tied.” Initially it is noteworthy here – particularly in comparison to the Principle of insufficient reason – that as one asks about the probability of an outcome, one’s attention turns away from the outcome towards the constraining circumstances. If we should claim that the entire ‘sphere’ of

---

<sup>43</sup> John Stuart Mill, *System der deductiven und induktiven Logik*, in Schiel’s German edition (4th edition, II, p. 79). [Trans.: As it stands, the quotation is taken directly from Mill’s 8<sup>th</sup> edition, rather than being translated back from the German edition. I see no significant difference. Mill, J.S. (1882). *A System of Logic, Ratiocinative and Inductive, Being a Connected View of the Principles of Evidence, and the Methods of Scientific Investigation*. NY: Harper & Brothers, Publishers. Chapter 18: Of the Calculation of Causes, pp. 668–9.]

<sup>44</sup> Fick, A. (1883). *Philosophischer Versuch über die Wahrscheinlichkeiten*. Würzburg: Stahel’schen Univers.- Buch.- & Kunsthändlung.

a generally formulated condition produced half a certain outcome, then at any rate we put forward a proposition which has objective meaning and which is not at all identical to the proposition that we have no reason to expect the exclusion of the outcome to be more or less likely than its inclusion. Of course this presumes that the parts into which we imagine that sphere to be divided, ought to be described as objectively equal – which is a point we will return to in a moment. The emphasis on this interpretation of propositions about probability – which is objective, at least in **Fick**'s view – goes so far as to neglect ever to say what relation this has to our assumptions and expectations. Perhaps one may also need to add that, insofar as the sphere of generally formulated conditions is divided into a part which leads to a certain outcome, and another which permits its exclusion, then the ratio of the sizes of these parts should be the (logically justified) measure of the expectation of the relevant outcome, once constraining circumstances are admitted under that general concept. It seems necessary to me to add this stipulation, since if we should wish to omit it, certainly one would have numbers of some objective meaning for those ratios, but one would not have any clear idea or justification why they should be called probabilities. It also seems to me in **Fick**'s account that the whole motivation for a methodological application of the probability calculus is only comprehensible given this postulate. So far as I can tell, that is the only essential progress shown by **Fick**'s intuition: that the objective ratio to which the rule governing expectation must be tied, is couched in such a way that is certainly more widely applicable than before, even if ‘the domain’ or ‘the sphere’ of a general condition still begins as a very vague idea. This notion is specified in an unsatisfactory way, and that is the reason why – despite his basically very different perspective on the subject – **Fick** ends up with an application of the probability calculus which is not significantly different from that of the orthodox mathematical tradition. Just as in that tradition, this application simply must be described as decidedly incorrect on several points. To round out **Fick**'s theory in a satisfactory way, first the question needs to be asked, if and how the ‘domain’ of a generally formulated condition may be considered to be partitioned into pieces of comparable size. **Fick** does not deal with this question at all; rather, it seems to him that in general a number can always be given for probability in the sense he gives. However, the way things are in implementation of the theory, it happens at least occasionally that the congruence of parts is established merely on subjective grounds: two pieces are deemed to be equal because there is no reason to hold one or the other to be predominant. Yet in that much one has reverted to the Principle of insufficient reason, despite the apparently objective sense that the primary formulation lends to propositions about probability. Therefore an appeal to outcome over a sphere of contingencies is basically a useless appeal. But further, even when we consider such a number as measures the ratio of two parts of a sphere in a way both explicitly known and reportable, still this would not have the same meaning in general as a numeric probability. Our knowledge of constraining circumstances is never exhausted by the problem which is set: that the case to be judged was to be subsumed under the concept of such a sphere, as abstracted from a number of previous cases. It may be asked further, when and why knowledge of the particular determinants of a case seems so irrelevant, when significant psychological

weight is bestowed on the measure of that numeric ratio. To put the matter shortly, **Fick**'s theory contains an entirely accurate statement of a concept which is similar to that of a range, but it neglects to elaborate the epistemological consequence of that concept in detail. In particular, it fails to pose questions which are fundamental to numeric probability: "When are probabilities determined exclusively by ranges?" and "When are ranges commensurate?". Accordingly, **Fick**'s theory stands in the same relation to the theory which I hold to be true, as does the Principle of insufficient reason. It is correct to say that if two cases are equally possible, we have no reason to expect one over the other; it is also correct that each throw of a die (1, 2, . . . etc.) represents  $\frac{1}{6}$  of the whole domain encompassed by the contingencies of the toss. The more important questions only begin with such a stipulation, however. Insofar as one fails to pose these questions, the universal supposition of behavior which is realized only in exceptional cases will lead to the same kind of errors in application of the methods.

**8.** Likewise, **Lexis**'s investigations (mentioned in Chaps. 6 and 9)<sup>45</sup> show large departures from the methods given by mathematicians. Of course **Lexis** is establishing specific norms for application of the probability calculus in a particular area, that is, for the collective phenomena of society. Hence we do not find anything there which one might call an exposition of the logical principles of probability calculus. Despite that, **Lexis**'s effort must be called one of the most important in the entire literature on the probability calculus, just for his regard of those principles. Its essential contribution can be seen to be that – for the first time – investigation is required to determine whether and how much any series of phenomena do follow the analogy to a game of chance. Further, methods are given for the empirical solution of this question. Finally a large variety of conditions is placed in suitable perspective, which conditions may exist even given the sustained constancy of any general conditions whatsoever, especially differences in dispersion. It does not detract from this contribution, that a definitive explanation is not attempted for the particular properties which are characteristic of games of chance. Rather, games of chance are defined merely by their empirical manifestation, and that remains as unexplained as the notion of 'physical possibility', as emphasized by **Cournot**. I should like to claim that of all the writers engaged in applying the probability calculus to phenomena of social life, **Lexis** is the first to have recognized in an entirely accurate way what rules are to be followed in this pursuit. And in this connection, he transposes the methods to answer questions which are quite different from those everyone was used to raising before. As a consequence his investigations are not only far more accurate, but they also represent a much more clever and useful application of the probability calculus than can be found in earlier literature on the same topic.

**9.** Insofar as it touches on the calculation of probability, modern logic has generally confined itself to accentuating the subjective meaning of propositions about probability. If thereby it neglects to pay sufficient attention to the objective

---

<sup>45</sup> **Lexis** (1877). *Zur Theorie der Massenerscheinung in der menschlichen Gesellschaft*. Freiburg i. B.: Fr. Wagner.

content which those propositions actually express, then that is easily understood. To explore that objective content, it is crucial to find an unimpeded opportunity to apply the probability calculus to specific real objects, where one has opportunity to notice that objectively valid data are expressed in the form of propositions about probability. By contrast, general theory provides no immediate opportunity for such an investigation. In **Lotze**'s and in **Sigwart**'s<sup>46</sup> writings on systems of logic, we find the ‘merely subjective’ nature of probability calculus to be emphasized in quite parallel ways. So **Lotze**<sup>47</sup> very deftly explodes the common confusion which is often committed that elevated probabilities for total outcomes over many cases represent ‘conceptual necessities’. “If this expectation is not confirmed in a significant number of cases, then constant circumstances can be responsible, but it can also be that an unregulated combination of variables is responsible; as often as this may be confirmed, still it is a fact which cannot surprise us just because it was improbable ahead of time. Its appearance may be as little proved to be necessary as the occurrence of any other mathematical probability.” In the same vein, **Lotze** says that concerning the “experiment with the tumbler and balls”, he cannot convince himself “that the gradually emerging stability of the ratio between balls of different colors may really be explained, if more than just probability is to be understood by that expression”. Just as these observations are undoubtedly correct, still on the other hand it appears doubtless to me they only ever teach us what considerations of probability cannot achieve. They leave us in the dark about what they can achieve. And the reason is simply that their dependence on ordinary games of chance – for which a statement of probability appears in a very elementary way to be the result of our lack of knowledge – means that the valuable objective knowledge which actually underlies the proposition goes unnoticed. Yes, it is correct that a state of knowledge for which we make some assignment of probability, is always different from one in which we may expect something with complete certainty, and for which we may deem the occurrence of the thing expected an understandable necessity. Yet it is no less correct that the propositional attitude that permits us to make estimates of probability which are accurate and well-founded (i.e., assignments grounded by familiarity with ranges), represents an exceptionally important advance over the other attitude, where we cannot develop any definite expectations given our complete ignorance of proportions. And in just the same way, that propositional attitude by which we know that an observed fact neither necessitates nor permits further general explanation than we are able to provide, is distinguished from another attitude, under which any understanding at all is lacking. Basically, it is only an insufficient consideration of this distinction – which figures so prominently in untutored thought – which **Lotze** has offered up as an ‘explanation’ of the regularities in question, as opposed to habits of ordinary language. It cannot be swallowed that, if approximately equally

---

<sup>46</sup> **Lotze**, H. (1843). *Logik*. Leipzig: Weidmann'sche Buchhandlung, pp. 414–434.; **Sigwart**, C. (1878). *Logik*, 2. Band: *Die Methodenlehre*. Tübingen: H. Laupp'schen Buchhandlung, p. 270 ff.

<sup>47</sup> **Lotze**, op. cit. p. 436.

many black and white balls are always drawn from an urn, that an ‘explanation’ of this phenomenon is to be found in the assumption, or in the fact that the urn contains equally many black balls as white balls. And if this counts as an example of an explanation, basically it stands as an example of the complete lack of understanding we have in facing a similar phenomenon, should it occur under other circumstances.

Secondarily, another circumstance may be held responsible for inaccurate construal of the logic of the probability calculus. It was not known (or at least it had been inadequately considered) that any specification in numbers – of whatever kind – is tied to the availability of elements which are unequivocal, and whose equivalence may be compared in non-arbitrary ways. That is, it is tied to a constraint which is in no way general: rather, it is realized only under special conditions. This omission is even more strongly palpable in **Lotze**’s work, because he has taken pains to codify precise rules for establishing equally possible cases.

According to **Lotze**, those cases should be considered equally possible “which are single cases that are coordinate in being of equally-valued kinds within the scope of the general case”<sup>48</sup> – which is an accurate but surely inadequate specification. The omission of the important point is felt here still more clearly as we read:<sup>49</sup> “The simplest determination of any magnitude of probability presupposes that a disjunction of all possible cases is given, that each of these is identical to itself and not equal to another, and finally that any one may exclude all others.” And, as if we needed to add, the partition must not only deliver coordinate cases, but rather must also deliver elements which can be made equal or else compared in magnitude.

Seen from this perspective, it is absolutely consistent if **Lotze** fails to concern himself with the dependence of propositions about probability on objective knowledge: “a finer evaluation of those probabilities, which are based on the more-or-less well-known internal associations of given states of affairs, is removed from the general remittance of logic and should be left to a discipline-specific familiarity with the particular case at that time.”<sup>50</sup>

In many respects, **Sigwart**’s treatment of the probability calculus is similar. According to him,<sup>51</sup> “the purely subjective nature of probability is frequently obscured, in that the examples chosen as illustrations contain further information, which is not contained in what is expressed by disjunctive judgment. For instance, in the example of the toss of a single die, we know – perhaps from the makeup of causal factors in particular cases, or perhaps from our experience that individual outcomes occur almost equally often with a single die in the great majority of cases. In other words the real causes which lead to a particular outcome of a throw will alternate, in such a way that they do not give an advantage to one outcome over another. Then the equal possibility of items in the disjunction is not merely given, inasmuch as the

<sup>48</sup>Loc. cit. p. 414.

<sup>49</sup>Loc. cit. p. 429.

<sup>50</sup>Loc. cit. p. 414.

<sup>51</sup>**Sigwart**, *Logik II*, p. 273. [Sigwart, C. (1878). *Logik, 2. Band: Die Methodenlehre*. Tübingen: H. Laupp’schen Buchhandlung.]

person making the judgment has no reason to deny one and affirm the other. It is given insofar as the actual circumstances are such that they realize the various possibilities one after another, should they only have the time to ring changes over all combinations . . . Small wonder, then, that the probability calculus finds a solid foundation in such cases as secure a particular meaning for their outcomes. Such postulates are not necessarily encompassed by the nature of the process. Disjunctive judgments and fractions of probability derived from them say nothing at first – or in themselves – about the ratio of frequencies of the actual occurrence of mutually exclusive cases; . . . rather they only express an authoritative measure of subjective expectation, that we should entertain for want of better information.” In this passage too, I believe that one can recognize the unresolved difficulty clearly. Principally, this seems an inadequate expression of that objective knowledge which, occasionally touching on the probability calculus, itself should deserve more prominent significance. It cannot be enough to say we would know that every single throw would ‘appear’ with approximately equal frequency over a great number of cases. Clearly that is comprehensible only under the inadequate interpretation that they have really appeared equally frequently in the course of previous experience, or else it is comprehensible under the incorrect interpretation that they must appear approximately equally often, as a matter of necessity. Neither interpretation gives us a real sense there is any matter of objective knowledge present in such cases. As soon as one corrects this interpretation, one sees too that this is not an incidental interpretation to be added to propositions about probability. Rather it is wholly indispensable and basic to those propositions.

In the same way we find the subjective character of probability calculus emphasized by **Windelband**, though the relation to actual events is cast in a fundamentally different way. He says:<sup>52</sup> “Probability is no property of the events which are expected; rather it is only a proportion according to which we gauge the strength of our expectation . . . If probability is just a numeric ratio of possibilities, then all determinations made in the probability calculus would only count as being about possibilities, and not about reality. If probability is no property of *facta*, but rather a degree of strength of our expectation for them, then every law of the probability calculus becomes, not a law of facts, but just a law about our expectation of them.”

But later, where it reads:<sup>53</sup> “It is part and parcel of the concept of equally possible cases, that for a sufficiently large number of cases, every possibility is offered an equally large set of opportunities for its realization”, then this expression for the Law of large numbers – which was found earlier in **Fries**’s work – must be characterized as incorrect, just as the principal distinction is incorrect which is drawn between the particular case – for which the computation of probability “makes absolutely no sense, in its very concept”<sup>54</sup> – and a series of many cases. If equal possibility is

<sup>52</sup> **Windelband**, W. (1870). *Die Lehren vom Zufall*. Berlin: F. Henschel, p. 32 [Habilitationsschrift, Universität Göttingen].

<sup>53</sup> Loc. cit. p. 34.

<sup>54</sup> Loc. cit. p. 33.

determined only subjectively, then it can be no part of its concept that the same possibility appears approximately equally many times in a series of many independent cases. Actually nothing further can be wrung from that concept, than the very large probability with which we may expect the approximately equal frequency of one event or another – but not the concrete realization of that expectation. Deduction, and certainly mathematical proof, ties certain events to a general rule of expectation. In turn events represent expectations. But proof cannot be wrung from such heterogeneous content.

The facts which have been assembled here should serve to emphasize what I wished to show: how in many ways necessity has made it worthwhile to gain clear insight into the objective foundations of the probability calculus. That insight attained on this basis has a different meaning than interpretation made without it, is likely never misunderstood. Moreover it has been understood and expressly emphasized that losing such insight always leads to illusion. But since the attempt to specify this objective foundation did not produce any satisfying result before, and since neither did the effort prove successful to render its necessity understandable in some way, then finally the “Principle of insufficient reason” was left standing, and no guiding thread could be found for application of these methods.

The theoretical difficulty of the problem could then have remained unnoticed, if only one considered it permissible without further comment to conclude that regularities simply emerged from large numbers of cases of the same kind, that the same regularities were to be expected for future cases, and if one dispensed entirely with the question what the justification for such conclusions might actually be. This is essentially the point of view taken by several English logicians. To them, it seems sufficient to define the probability of an event as its relative frequency over a large number of cases.<sup>55</sup> It is unnecessary to delve into their intuitions in more detail here. They play no role in the history of our present problem, since they do not even raise it.

The pressing and urgent conviction that a truly valuable application of the probability calculus would necessarily postulate objective items of knowledge was the spark for the present investigations. The question what these items might be, gave a first impetus to the investigations. A second point of departure lay in the general theoretical notion that every numeric determination of magnitude is tied to a specific condition, that is, to the concatenation of objects to be measured from among a number of elements which are in all respects identical, and hence comparable in magnitude. This in turn led to another question: what are the circumstances under which what is (subjectively) possible can be partitioned into a number of premises which are assured to have the same value? I trust that I have done justice to both these questions. Our effort has been successful in showing that the reporting of numeric probabilities presupposes one has specific information about objective content. It is no less true to say that proof was furnished that only very specific

---

<sup>55</sup> Compare for example **Venn**, J. (1866) *The logic of chance: An essay on the foundations and province of the theory of probability*. London: Macmillan and Co., p. 145.

mental phenomena admit any characterization in numbers: in fact this property is completely tied to such objective knowledge.

**10.** Let us conclude this overview with a forward-looking commentary. From the logician's point of view, the task of establishing a mathematical discipline on theoretical principles is an undertaking to which even mathematicians have not given much notice, in general. If they notice the products of such an investigation at all, then as a rule they find the concepts given to them too complex and unwieldy for them to be able to proceed with much comfort. Doubtless they are generally right in that much. For example, who would want to have the usual expositions of differential calculus and of mechanics burdened by profound considerations of logical issues? Things are essentially different in a theory of the probability calculus. Here the issue is not the same as in those other disciplines, where the issue is to find the most rigorous and the most logically satisfying formulation from a host of theories which are firmly established. Instead a wide-ranging doctrine patched together from error-laden components – or at least parts which are clearly badly represented – must be purified, and a form found for them. That form has at least to satisfy definite requirements of logical correctness and of comprehensibility. The desire to contribute to the unearthing and the exegesis of such a formal account of the probability calculus in these investigations, bound me to an obligation at least to sketch how I found that such an account might be possible – with some consideration to be given to the (certainly justified) demand that the account should not be unduly freighted with abstract demonstrations in logic. So far as I can tell, the mathematical theory will find its way forward most securely and easily, if it clings strictly to the notion that it must deal with actual magnitudes which have a meaning simply given in objective terms. In order to do that, the theory must begin with the concept of ranges. That concept, and the ways it determines our expectations in the simplest of cases, might first be explained, say, by relations of coincidence in time between two independent series of behaviors. And then the determination of measure for a composite range could be explained, perhaps most usefully by the possibilities of arrangement for an isolated system of point-masses enclosed within a definite space. Then that would serve to express the primary sense given to calculation by the methods of combinatorics.

After it has been clarified in this way what the size of a range should be understood to be, that is enough of an idea on which to hang all the basic concepts of the probability calculus. The first of these is the concept of originality. It can be seen without further elaboration that two varieties of behavior are really to be considered equally probable, not just if they encompass equal ranges, but only if they also arise from equally large complexes of previous varieties of behavior. Moreover, the central meaning of originality may be illustrated through simply-contrived examples of likely varieties of behavior.

After such a development of the principle, initially it would be very practical to treat an idealized game of chance as an example application of the principle – perhaps the bowling game could be used, or else a game of roulette with infinitely small pockets on the wheel could be used. From there, the originality of the range ratios in question could be derived from the property of continuity of all the

functions, and so a convention about probability could then be shown to have a formal basis. Naturally the Law of large numbers can then be derived as a proposition in the discipline of combinatorics. The simple and cogent sense that the Law of large numbers lends to assignments of probability may be made clear, just at first, by an idealized game of chance.

By contrast, the application of propositions about probability to actual games of chance should expressly not be introduced as the immediate logical outcome of our uncertainty. Rather it should be introduced as a proposition with actual meaning, that is, as an hypothesis. It ought to be shown that the hypothesis might be assumed to be well-founded in part because of a rough correspondence to an idealized game of chance, and in part because of the results of experience. How much emphasis one wishes to put on the former factor and how much on the latter, will always be a matter of individual taste (one that is basically irrelevant). Taking that hypothesis as a basis, later the remaining elements of the general theory may be explained in terms of games of chance – those remaining elements being **Bayes**' principle, the determination of probability *a posteriori*, and the relations of independence or connectedness of individual cases. In all essential points the theory would not have much to distinguish it from the form which has been usual until now, with the exception of one point: that the subject of the calculations would not appear to be subjective probability, but rather the sizes of ranges. That much is entirely sufficient for one to reach the correct perspective in a snap about application of the probability calculus to what is a very special domain: namely, the theory of error. It would have to be shown that the assumptions of objective content which underlie **Gauss**'s theory of error, may be demonstrated both initially by deduction proceeding from not unreasonable premises, and also by being confirmed empirically in many different ways. Thus the content of those premises might then seem to be well-founded, but they would need to be tested carefully in each specific situation. Then the right assessment of results which are produced by least-squares methods would appear *sui generis*, which is to say methods of probable error would be evident in themselves.

Along the way it might be recommended that one show how other phenomena differ from games of chance, such as those which are treated by medical statistics, or else witness statements and judicial decisions. They do not permit application of the probability calculus under the same muster as ordinary games of chance, because of the lack of constancy in their general conditions, or else because of an indeterminate degree of connection among individual cases. – In this way the whole exposition would be simplified in comparison to that which is now common, basically because the probability calculus is stripped of the pretense that it is some general logic of the unknown. As far as I can tell, much more complex investigations will need to be passed over in the teaching of mathematics, which should not do any harm, because in general mathematicians occupy themselves with subject domains in which these two premises hold (that is, more complex investigations of the conditions under which ranges are solely responsible for determining our expectations; and just when two ranges may have a relative size which can be expressed at all). Some indication such as a footnote would then be enough to say that this is not always the case.

If I do not deceive myself, an exposition of the theory which follows the design just sketched would also be conducive to better understanding by students than the exposition which is usual nowadays. At the same time, it easily satisfies the usual standard for rigor and accuracy, without a single quantum of purely logical explanation being involved.

# Chapter 11

## On Probability Theory



**Abstract** A response is given to criticisms of the range theory of probability. That theory stands opposed to general application of the principle of indifference. Much hinges on the notion of equally probable cases, but this equality is contrasted with equivalence relations in mathematics. The range theory takes account both of objective ratios, and of what we do and do not know. Everywhere that our knowledge is incomplete or imprecise, the theory contributes to the determination of probability. Often this is a matter of unsaturated judgments that encompass a range, concerning extents or behaviors represented by continuous magnitudes. Stumpf has objected to earlier examples given in support of the range theory. His position is shown to be misleading and extreme; it depends on a logical connection to disjunctive judgment that is easily overemphasized. Derivations made by Maxwell and Boltzmann in the kinetic theory of gases are used in an extended example to support the range theory.

**Keywords** Comparison judgments · Symmetry of errors · Equivalence relations · Partition of ranges · Propositional attitudes · Incomplete knowledge · Probabilistic fallacy · Heuristic significance · Nomological admissibility · Maxwellian distribution

The theory of probability I developed in 1886, which is given here with no change worth mentioning, has met with widespread agreement to many of its components over the course of time. To other components the theory encountered many opposing views and objections. So I would like to discuss a few points in greater detail, which could not be reviewed (or reviewed thoroughly enough) in the context of the main argument, in the interest of full transparency and for removal of many a misunderstanding. This may be achieved in part by a direct response to certain objections raised against the Range theory.

---

*Logic: Outline of a critical and formal theory of judgment*, 1916, Chapter 26

One of the starting points of my investigation – not the only one – was tied to the numeric evaluation of probabilities, and had raised some discussion. All of this hinges on the notion of equally probable cases. Evidently the question proceeds from this intuitive idea: that we can call premises or expectations equally probable, if they hold the balance between reasons which speak for one thing or another, and if deliberations provide us no definite cue for holding one more probable. One might say the probability that the war in Europe will come to an end in six months, may be just as large as the probability that it will not be the case. The view that the probability calculus finds its characteristic and only foundation in establishing equally probable cases, is what I have called the Principle of insufficient reason. No doubt we can speak of equally probable cases. In this sense the judgment we express when we call two expectations or two premises ‘equally probable’, is evidently a comparison-judgment in the sense previously discussed, similar to one we make when we call two sensations equally strong, or the difference between sensations equally large or equally clear, etc. It has been emphasized that we must distinguish the equality of such propositions from the strict and definitive notion found in mathematics. The operations of congruence expressed by propositions of mathematics imply that the relations of two objects to be compared may not be characterized as ‘more’ or ‘less’. The broad scope over which such comparisons may be carried out, depends on this exceedingly broad connotation, but also on a correspondingly indefinite sense of ‘more’ or ‘less’, ‘stronger’ or ‘weaker’, and so on. We can think of these comparisons as an expression of the relation of incidence for those vague concepts (a connection of membership under the general concept). Earlier we had said that for more or less vague general concepts, the relation of incidence is atypical. It seems there are many ways to doubt that an individual falls under the general concept. We emphasized that this uncertainty – which emerges from the indeterminacy of the concept – should not be confused with some question of truth or falsity in an objective sense. Whether a sensation we have – say one that falls between yellow and red – is still to be called red; whether a noise we hear is to be called loud: where these things seem altogether doubtful, they elude meaningful discussion. They are undecidable by the nature of the question. Such comparisons are possible over the broadest of scopes, but for many reasons they are exceptionally uncertain. It is important to maintain that when we are still in doubt, we are not to ask for definite outcomes, as we would validly ask or seek to have otherwise. We can call the difference between red and green greater than the difference between a tone of 400 Hz and one of 405 Hz. Nevertheless comparisons between differences in optical sensations and differences in acoustical sensations are entirely uncertain. One would be deluded if one set out to ascertain, nay, that one must ascertain a difference in color equal to an augmented third in music.<sup>1</sup> – This is what happens with relations here. Often we can compare the probability of two premises by their specific

---

<sup>1</sup>On this whole subject, including comparison judgments, the psychological notion of equality, and the atypicality of relations of incidence, etc., see the earlier passages from my *Logik* (1916) on pages 12, 31, and 475.

outcome, to determine if one outcome is larger or smaller than another. But as a rule, such comparisons are not feasible; their results seem both uncertain and arbitrary. This has been illustrated earlier with simple examples. We can compare the probability of an inductive generalization with a probability from a game of chance. We might say rashly that it is greater than the probability with which we expect one side to turn up in the toss of a die, but smaller than the probability which we assume that a roulette ball will land on red at least once in the next 100 spins. It does not seem feasible to report a numeric probability to which it is strictly equal. The same holds in general for probabilities that concern ratios of ranges, and probabilities measured in those terms.<sup>2</sup>

Just like other comparison judgments, as a rule evaluations of probabilities by the Principle of insufficient reason lead to wobbly results. Yes, given the complex nature of what is to be compared, uncertainty and arbitrariness will be particularly evident. One may notice that one's perspective changes, when instead of paying attention to processes, one asks about the probability of constraining circumstances which lead to some outcome. On the basis of some consideration which seems appropriate under the rules of probability calculus, we might calculate a very high probability for a behavior. Then we become aware that for another admissible way of considering the problem, a negligible probability is obtained for the same behavior. If we find in games of chance that assignments of probability seem fixed and free from arbitrariness; if expectations with enormous probability values are formed from the results of very many cases without some objection being raised; it may be assumed some special circumstances give rise to such measurability. Then we may ask: under what constellation of psychological conditions could a firm and independent evaluation of probabilities be made, free from the arbitrary influence of other considerations? The Range theory provides an answer. In those very cases a ratio of magnitudes is known to us: a ratio for ranges of constraints that lead to different outcomes. It is just these ratios which lend measure to expectations. So one can say that assignments of probability are based on specific knowledge with objective reference.

Philosophical objections have been raised to this very point. In particular, **Stumpf**<sup>3</sup> has interceded for **Laplace**'s old formulation, against this position. He takes pains to show that **Laplace**'s view suffices to achieve definite estimates of probability free from arbitrariness. – It may be remarked that the way in which **Stumpf** defines equally probable cases – hence the way conditions for assigning numeric values are laid down – is not exactly equivalent to the definition I had given for the Principle of insufficient reason, which I had sought to prove insufficient. In that lies the danger of a certain misunderstanding, connected with a mismatched

---

<sup>2</sup>Compare my *Logik* (1916), pp. 415 and 432.

<sup>3</sup>von **Stumpf**, C. (1892). Ueber den Begriff der mathematischen Wahrscheinlichkeit. *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften zu München, philosophisch-philologischen Classe*, pp. 37–120; von **Stumpf**, C. (1892a). Ueber die Anwendung der mathematischen Wahrscheinlichkeitsbegriffes auf Teile eines Continuums, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften zu München, philosophisch-philologischen Classe*, pp. 681–691.

interpretation of **Laplace**'s so-called classical definition. In **Stumpf**'s text a special emphasis is placed on a condition I never referred to: complete unfamiliarity. **Stumpf** says (op. cit. p. 41): “**Laplace**’s view is completely expressed if we say: cases are equally possible, if we find ourselves in the same state of ignorance with regard to them. And since the measure of ignorance can only be set equal, then if we know absolutely nothing to tell which distinctive case will occur, then we can definitely substitute this statement for the other.” This condition of complete lack of knowledge is stressed in other places (op. cit. p. 67), in the form that our knowledge is exhausted by the contents of disjunctive judgment – that it must contain nothing beyond that.<sup>4</sup> It is still a large difference, whether we know everything possible (with certainty or with little certainty) about things to be compared – and our thoughts about the comparison lead us to no decisive preference – or whether we really find ourselves in total ignorance faced with those cases. Then one may recognize that the numeric specification of probabilities is not completely valid. One might agree that the reasons I gave for this are justified. One might also admit that the probability of decisions formed by analogy or inductive generalization do not admit numeric measure. One might draw the conclusion that probabilities are said to be equal, just when such considerations are completely eliminated for the members of a disjunction, by the force of a propositional attitude. And one might be of the opinion that this happens just when such evaluations may not be conducted, out of an absolute lack of knowledge. That intuition represents an extreme refinement of the Principle of insufficient reason: let it be called the Principle of absolute ignorance.

Prolonged pursuit of this idea – which may be argued for in many ways – leads directly to a condition which is hardly ever realized. It is not even realized in cases where it might seem to be realized initially. Some knowledge of relevance is represented by the very meaning of those concepts (in which the members of the disjunction are named). There must always be relations among the several members of a disjunction. Those relations will be important to the evaluation of probability, by the measure of this or that item of knowledge. Whether this is contradiction by

---

<sup>4</sup> Of course it is inconsequential whether **Laplace**'s view is interpreted completely accurately here, or if it can be more cogently identified with the Principle of insufficient reason, as I gave it. Still, it may be remarked that when **Laplace** defines: “Cases which are equally possible, which means those for which we may be equally undecided about their existence” [« *Cas également possibles, c'est à dire tels que nous soyons également indécis sur leur existence* »], the expression « *également indécis* » is not strictly to be translated as equally unknown, but rather as equally uncertain or equally undecided, which should accord better with the latter interpretation. In a similar argument, **Stumpf** also repeats **Sigwart**'s view, which advocates taking complete lack of knowledge about the members of a disjunction (considered in isolation) to be an additional premise. However, I doubt this represents **Sigwart**'s view entirely. In the passage which has been selected, it reads “always presuming that our knowledge is restricted to that which is expressed by the disjunctive judgment, and that there is no other reason which allows us to expect the one over the other.” Obviously this codicil points to a weighing of various reasons, which would always presume a certain measure of knowledge. In that much, one can simply say that this stands in a certain contradiction to the earlier premise.

antithesis, or a tale of notions that are somehow concordant, the content of these notions gives rise to a series of questions in any situation whatever. For some, it is impossible for them to be met by total lack of knowledge. They offer a series of footholds to the understanding. The same thing is easy to show for games of chance, if one manages to avoid confusion. As we are used to saying for games of dice, it is utterly unknown to us what makes us expect one side of a die to turn up rather than another. We ought not to be led astray by this idiom, to say that the absolute ignorance demanded by the Principle is present. Let us set aside all the circumstances which coincide for the cases, which would be excluded anyways. There is always a series of distinctions we know among the cases to be compared, to which conjectures about their probability may be tied. About a single die we know its sides are marked with one dot or two or three; those consist of indentations or flecks of color or something similar. We know that a throw of 6 often has a certain psychological significance as the highest of the numbers, and so forth. We can pose a series of questions directly: whether the centre of gravity may not have been displaced a little by application of the marks; whether attention paid to the side marked 6 may not influence the outcome; whether the red or black color of the pockets of the roulette wheel may not exercise some attraction on the moving ball, or if one color might offer more friction to that movement, and so forth. We do know this is not the case, or that such things make a minimal contribution. If we say that we “know nothing about what would prompt us to weigh several probabilities unequally”, that really means we know positively that there is no reason to give for a series of conditions. Besides theoretical points, we should also remember we have experience with games of chance. It is actually very likely known to us from roulette that the balls lands on red and on black nearly the same number of times, over a large number of spins.

Specific proportions really do arise in games of chance, and the usual assignments of probability emerge as being cogent and well-founded by acquaintance with them. It is not impossible we might arrive at the same result even without appealing to that knowledge. Might it not be at least acceptable to recognize another way of considering the matter, which abstracts from such knowledge entirely? Will the Principle of absolute ignorance prove a sufficient foundation for the evaluation of probability that way? It seems enticing to pursue such a modified perspective on games of chance. Perhaps the widespread intuition is – and if I am not mistaken, this is the direction taken by **Stumpf** – that one neither needs to admit nor argue against the characterization of games of chance I argue for (that individual results reflect equal ranges of constraining conditions). It may be useless to bang one’s head on the wall about this, since correct and useful assignments of probability are established without consideration of such knowledge. Those assignments are convincing, and they may be established independently of arbitrary changes in perspective. – Simple reflection tells us that the ordinary (and correct) treatment of games of chance is not based on mere ignorance. Suppose one takes the mental stance of someone who really does know less about the game than we do, or else nothing about it. Suppose that is someone who has never taken a die in their hand, or who has never seen a game played. He is in no position to reckon the influence that markings on the die – or the unshakeable idiosyncrasies of the game, or else the situation of the die

preceding the throw – may have on the outcome. Plainly he will have to deal with the possibility that only one number can come up with any real die. With more than one throw of the die, he may be forced to reckon that repetition of the same result is higher in probability than an arbitrary sequence of several different results: that the probability of ‘six and six’ or ‘three and three’ is higher than the probability of ‘four and five’, etc. He has no solid foundation to evaluate or determine just how often this may happen. The treatment of individual cases as independent therefore stands in need of a certain basis of knowledge. It is legitimized by that knowledge. Even the most important assessments of probability, which approach certain expectations for collections of results over a multitude of cases, would go unproven without positive knowledge, and would come to nothing.

Things are different in each case. One might say that if we knew nothing about dice with a displacement of the centre of gravity, and the influence it has on outcomes, then we would know nothing about the side to which such a change leans, in terms of our postulates. The probability of all six outcomes would be the same. Even if we only pursue expectations about the outcome of one toss, conditions never fall in just the way that would support the Principle of absolute ignorance. Mainly this concerns a physical object which is geometrically regular, whose displacements in space form a mathematical group. Among the arrangements which are possible mathematically, those possible in real terms (those which represent some stable equilibrium) represent a definite fraction which is itself a definable mathematical entity. The considerations employed here find their simplest and most perspicuous analog in the spatial orientation of a real line. The class of geometrically available directions form a particular set, which we may measure on the surface of a unit sphere. We assign the probability that the directions lie within a certain solid angle, to be proportional to the measure of that angle – equal to that value divided by  $4\pi$ . Here probability is measured in terms of a range. We are led to behaviors not considered under our general description, only insofar as directions or angular extents exhibit idiothetic proportions. Mainly this is when they form a complete and closed totality, which may not be extended infinitely in a positive or negative sense as lengths or temporal intervals can be. There is no great variety of possible ways to measure them, as there are for physical magnitudes. Complete lack of knowledge about them seems imaginable at least, since their parts can all have equal values. – Something very similar happens for the throw of a die. What we call a specific outcome (e.g., that the die comes up six) is characterized by a whole system of spatial arrangements. In part this happens because we think of the die coming to rest on arbitrary places on the board, and in part because we think of the same position being rotated about an axis perpendicular to the board. In turn, positions which stand for particular outcomes represent strictly defined mathematical systems, and which are of the same size for each outcome. Right here we have the decisive reason why we reject alternate conventions about probability as false. We reject the contrast between positions that the side which shows six may be oriented vertically to the board, or oriented horizontally to the board.

If probabilities are measured in terms of ranges – even for this game of dice – initially they concern the resting state of the die. We can also proceed to modify the

description, given the constellation of processes which lead to one outcome or another. That description takes no account of what is known about the situation. To simplify relations further, let us focus on a coin toss (heads or tails) instead of using a game of dice. To render this in simplest form, assume what is tossed is a circular disk on which one side has been marked. If the meaninglessness of such a marking to the outcome were not known, it could be conjectured that the mark could favor one possible outcome. It is known that to any premise in this situation one can counter what may be called a symmetric premise. To the premise that the symbol favors one outcome in some measure, one can counter with the premise that it favors the opposite in equal proportion. Then despite the ignorance of nomological relations assumed here, a probability of  $\frac{1}{2}$  is to be assigned to each of the outcomes. Even in this isolated case, a specific evaluation of probability depends on equivalence relations between different members of a disjunction. The premises treated as equally probable, are once more those which encompass equal ranges. The case before us is distinguished from those discussed earlier, only in that the unbiased stipulation of equality between two ranges depends on complete symmetry (represented as positive values and negative values), and not some composition of tiny parts that alternate regularly. For premises with nomological content, we may assume as complete an ignorance – without entering into the realm of pure fiction – as is necessary to render the ranges suitable for the measurement of probability. It may even be conceded that under this fictional perspective – which ignores what we know about games of chance – there is a definite basis for assignment of probability. But we must combat the notion that the Principle of absolute ignorance finds justification here. The assignment of equal probabilities is not in any way based merely on the notion that in the absence of knowledge, all conjectures come to nothing. Even here they are grounded in objective ratios known to us. Even here the last word is given to the Range principle.

Ratios of ranges are not shown to advantage here, as they have been shown important to our theory of games of chance. I do not see any contradiction in this, but at most an elaboration which is evident to some extent. It should not be claimed that when those particular conditions are not fulfilled the evaluation of probability becomes arbitrary: that 0.1 may be assigned just as well as 0.9. It is only claimed that the assignment of probability is uncertain within broad bounds; there is no specific value which is strictly correct. Suppose we find certain cases where a definite numeric value for probability is strict and unbiased. That does not exclude that some clue may be given that permits some evaluation to appear more obvious and appropriate, even if the necessary conditions are not present for them to be both unbiased and compelling. Still one will find such clues from things we know, rather than out of complete ignorance.

Sometimes consideration of fitting examples leads us to cases where – first appearances to the contrary – positive clues point to an assignment of probability. More frequently it leads to cases where such knowledge as we have makes a specific assignment seem appropriate, while on closer examination its justification seems doubtful. As an example of the former, there is the hackneyed example whether a

child will be of male or female gender. One may be totally in the dark about lawful connections about that; it may be utterly unknown how gender is determined through procreation and development. We may assume that nothing is known about the relative frequency of male and female births, as found by statistical inquiry. But that the two genders are present in at least almost equal number, is a fact known unconditionally: such knowledge is hardly to be wished away. Even ignoring that, what remains is still given by the notion of male and female genders: one of one kind and one of the other contributes to procreation. In this question, the assignment of 0.5 to each proves to be both obvious and appropriate, but in no way is that the result of absolute ignorance. Rather it is due to knowledge given by the assignment of members to a disjunction. – The importance of these circumstances is apparent as we think of similar examples, where either such clues are not present, or they are given in another form. Thus we come to cases where the assignment of probability seems appropriate at first, but proves controversial on closer examination. So it might be asked whether a certain star we see in the night sky is a planet or a fixed star. It is imaginable (here we should presume) that the questioner has had no instruction about the relative frequency of planets and fixed stars in the night sky. Out of absolute ignorance, one could form the conviction that the probability in both cases should be  $\frac{1}{2}$ . Only someone who knows about planets (that they belong to our solar system) and fixed stars (that they are celestial bodies outside the solar system) will be inclined to make the correct assessment of a much higher probability of fixed stars.

The consideration of a few more examples seems instructive in many respects. If a disk is marked before it is tossed, then we may consider it equally probable (because of the disk's exact symmetry) that the marking gives an advantage to one outcome. Numeric values expressed by nomological propositions hardly ever exhibit such complete symmetry of positive and negative values as is the case here. Although the relations just mentioned are similar, still conditions for the numeric evaluation of probability do not have to be present. One might investigate how mortality for some disease is affected by medication. One could suspect that an influence in positive or negative terms, or that decrease or increase in mortality may be balanced symmetrically, or that the probability of either condition could prove to be 0.5 in the absence of other knowledge. But even if mortality should reach 20 percent, then in no way does an increase in mortality to 100 percent or a decrease to 0 percent represent symmetry. An increase to 100 – meaning certain death – seems imaginable (for any effective poison) but by contrast a decrease to zero – the avoidance of any fatality – can hardly be expected. So if we know nothing about the treatment, it seems doubtful that we must assign equal probabilities to favorable effects and to deleterious effects. One might well imagine evaluating the probability of a deleterious influence highly at first. – So it seems to me not an unbiased argument, by which **Stumpf** tries to show that when we draw a ball from an urn – about which we know only that it contains black and white balls in some unknown ratio – we have to assign the probability of drawing black or white to be 0.5, uniformly for each. Even the symmetry between black and white present in some

games of chance, is in no sense absolute. So presumably there are many more materials which are white in their original state (wood, ivory, bone, and celluloid, among others) than black. Of course materials may be dyed, as when black materials are required, or when equal numbers of black and white balls are called for. In the absence of such an occasion, those materials may be expected to be white materials, and enter into commerce as such. Then it is justified to conjecture that there might be more white balls than black. But suppose that out of total lack of knowledge, all such conjectures vanish; it might still be asked whether the concept of white and that of black extend to equal ranges of surface properties, or to equal ranges of sensations. Or one might assert that similar modifications of surface properties as change white objects into colored objects may be imagined for black objects, but that these properties of black objects are unnoticeable, so they always remain black. Consequently it may be conjectured that black may stand for a larger manifold of surface properties than white. As a result (given the presumed lack of accessory knowledge) the probability of a black object may be assigned to be higher than that of a white object. The justification of such considerations is clear given the opponency of many color terms which explicitly have a relatively narrow meaning – such as white and crimson red.

The general consideration of a point raised in the context of games of chance leads us to the most important circumstance of this kind. There it was emphasized that it is always and only reasonable to treat a series of similar cases as independent in light of a definite bit of knowledge. If such well-defined and positive facts are not known as would allow us to treat the cases as independent, the probability of a recurrence is always to be estimated greater than the probability of different outcomes. Just as we contrast members of a disjunction – which subsumes a series of similar cases where one always represents the same outcome while another represents an alternating outcome (as for example BBBB and BRRB for roulette) – the contribution of each member of a disjunction represents a bit of knowledge which is not to be ignored or imagined away. It is a bit of knowledge which is important to the evaluation of cases. Without any knowledge of the way outcomes arise or the conditions on which they depend, the probability of recurrence would have to be estimated as distinctly higher than the probability of alternation. Only well-defined and positive knowledge justifies us in declaring all the innumerable series of possible events to be equally probable. It is only that which leads us to expect that various outcomes occur in approximately equal numbers over long series. It is in this crucial case – meaning the judgment of long series of similar cases – that the Principle of absolute ignorance never finds application, because of the meaning attributed to individual members of the disjunction.<sup>5</sup>

---

<sup>5</sup> Considerations of the kind that has been discussed here also enter into a case that I had mentioned previously, and which had then been discussed by **Stumpf** as well. Let it be known to us that an urn contains 1000 balls, some of which are black and some of which are white. Yet it is unknown how many of them are black and how many are white. We can then ask about the probability of each possible case (that 1000 are white; that one is black and 999 are white; that two are black and 998 are white, and so on). Then one might think that a definite assignment of probability may result,

In these arguments, it has been confirmed that there is no easy arrangement of the formal conditions for the evaluation of probability to be found in the Principle of absolute ignorance. Where one is inclined to think of applying that Principle, there is no such arrangement. There may be some interest in asking if there remain any conditions, where that Principle is given adequately and completely. This might be the case, but just given the fictional nature of the conditions to which we are then driven, we should pay close attention to this as fiction. Absolute ignorance with respect to the members of a disjunction might be realized by a symbolic expression of indeterminate reference. We may not arrive at such a state of knowledge by direct experience, but it may be imparted as information. Someone might say that within some period of time, one of two outcomes occurred, or will occur. One is called  $x$  and the other  $y$ . Now if we experience nothing else, there really will be no reason to hold  $x$  more or less probable than  $y$ . But even this case represents an assignment of probability familiar to us from games of chance – though incompletely. Yet this case is distinguished from games of chance: with games of chance all departures from ordinary convention seem positively false. In the tale of maximum ignorance, doubtless one can say that as a foundational matter, basically any assignment is equally justified as equivalence across the board.<sup>6</sup> It may be asked whether an unbiased person would not decline to evaluate probability as intractable in the absence of a point of reference. It may be conceded that when one is forced to pick numbers, that one may settle on a half-and-half ratio – at least I won't argue the point. But I cannot rid myself of the impression that there is a somewhat suspicious sense in which the whole approach follows in **Buridan**'s tracks. It does when we are concerned with pure fictions that play no role in the actual course of ordinary thought.

Suppose we can devise cases which are realized as two or more members of a disjunction, for which any point of reference is lacking. If we concede that we arrive at an equal assignment of probabilities under these circumstances, it becomes clear these are cases tailored to fit the theory, which lack practical significance. We cannot

if we assume that nothing at all is known to us about the way the urn has been filled. But this postulate cannot be realized, in the sense that with it all considerations were imagined away, never to be realized. Still we know very well that for all sorts of reasons and for all sorts of purposes, huge numbers of balls will coincide in one color, or else possibly in equal numbers of black balls and white balls, or that balls of one color may be clumped together in large numbers, and so on. Consequently I cannot then recognize that the result to which **Stumpf** arrives is appropriate (op. cit. p. 66), that for any single ball the probability may be independently counted as equal, either that the ball is white or that the ball is black. Then and only then one might be able to reckon that any ball could be drawn randomly in a manner that gave an equal chance to black and white, as say from another urn which contained an equal number of white and black balls, and as if we knew all this. If by contrast we really know nothing – in the sense given by ordinary language – about the manner in which the urn has been filled, then the considerations just mentioned will find a foothold, and it will seem to be ruled out entirely that one may evaluate the probability of all white balls or all black balls to be so very tiny, as the results of **Stumpf**'s calculations would show.

<sup>6</sup> As one might say figuratively, the formal conditions of the two cases are different, in analogy to the difference between a stable equilibrium and an indifferent equilibrium in mechanics.

conceive this as being a basis for pinning numbers to probabilities in actual cases, meaning any basis for the Principle of absolute ignorance in games of chance. Yet even in the Principle of absolute ignorance there is surely a grain of truth. This is implicit in the notion that in games of chance we are always led to the notion of an absolute lack of knowledge. Suppose everyday thought asserts that it may be completely ‘random’ whether one outcome or another occurs. A sound opinion on which that might be justified is that the conditions which lend measure to the situation lie completely outside our capacity to recognize them. That is the sense in which we may speak of an absolute lack of knowledge (without losing ourselves in meaningless fiction). In a certain respect that is established when the occurrence of one outcome or another is determined by indefinitely small variations in constraining circumstances, when the variations lie utterly outside our powers of observation. This has great significance for the quantitative evaluation of probabilities. This becomes clear if, drawing on our initial formulation, we require that all conjecture should be completely dispensed with, which leads to obscurity and to avoidance of numeric specificity. In practise if we ask about the probability of this event or that, the large majority of our current concerns about comparison and evaluation are dispensed with – when changes in outcome are tied to indefinitely small variations in circumstances entirely inaccessible to our observation. But the conditions for evaluation in numbers are never given by this alone – particularly not the assignment of some uniform equivalence of probabilities. Under those circumstances everything depends on ranges of conditions which produce one outcome or the other. Even these considerations – connected to a multitude of nomological relations – generally lead us to domains which do not permit any specification of numeric values. Even those vanish if we know the ranges to be equal. So having taken another path, here we arrive back at the Range theory. Indeed, we can say emphatically that the condition of absolute lack of knowledge (in whatever sense that can be fulfilled) is just what I have called the ‘indifference’ of ranges, (*Principien der Wahrscheinlichkeitsrechnung*, p. 24 f.; here Chap. 2, Section 1) once we consider the matter further.

There can be no absolute lack of knowledge so complete (leaving aside speculative fiction) that it supplies the sole and sufficient condition for assigning equivalent probabilities. Perhaps in certain specific scenarios, complete lack of knowledge is essential to an assignment of equivalence, even if it is not the sole and sufficient condition.

It has been demonstrated (*Logik*, Chapter 19, p. 432) why a completely general application of numeric probability is unfeasible. It is unnecessary to rehearse those conditions once more. In *Principles of the probability calculus* I had tried to show this with a few examples. I would like to return to two of them, adding a few words in response to criticisms which have been raised.<sup>7</sup> One concerned the situation where some event, say the fall of a meteor, contacts an arbitrary location on the earth’s

---

<sup>7</sup> Several other examples that I had mentioned are also touched upon in the previous section (*Logik*, 1916, p. 606 and p. 607 ff.; here the text just prior to note 119).

surface. I sought to show that we may report a numeric probability that the meteor may strike a certain part of the earth, only if the size of that part and its relation to the entire surface of the earth is known. I have not been able to convince myself that **Stumpf**'s objections to my statements are cogent ones. I have to agree with **Stumpf** that one must pay attention to the nature and significance of different partitions, and that one may distinguish 'real' from 'fictive' partitions. In itself this distinction is insufficient to lead to an unbiased assignment of probability. Let us assume there may only be one partition of the earth's surface which is recognized and useful, say a division into parts *A*, *B*, *C*, *D*, and *E*. Then it may seem that our assignment of probability must be based on this division, and one may deny that it is permissible to swap this for some 'fictive' scheme. An example of a fictive scheme might combine parts *A*, *B*, and *C*, or it might distinguish northern and southern halves of *A*, and so form two parts of *A* – and so forth. Modifications of this kind are unjustified only if we have some reason to assume that the original partition must be valid to some extent. Should this not be the case (even if we know nothing further about that partition, except that it is useful), one is right to say that any modification by consolidation or by finer division has as much claim as the original partition to form the basis of our considerations about probability. It is quite unnecessary to introduce such arbitrary postulates. It may very well be the case that two or more such partitions exist alongside one another, and have uses known to us. This holds for the earth's surface, which we divide into continents, and which we also divide into countries with political boundaries. We may arrive at entirely different assignments of probability, depending on the partition we begin with. And we have no fixed point of reference to prefer one as criterial or to exploit both in some combination. A single numeric assessment is impossible, and one we might devise somehow will seem arbitrary. This should be admitted under **Stumpf**'s theoretical stance, as a matter of consistency. The absolute ignorance he postulates – the restriction of our knowledge to a matter of disjunctive judgment – simply does not happen when several partitions are given, about which their connections one to another are known. Apart from exceptional cases, we proceed directly from this to the insight that such a state of absolute ignorance cannot be realized.

**Stumpf** makes an assertion which is to be denied unconditionally: that there is an absurdity in the way the problem is initially posed, hidden in the infinite divisibility of surfaces and the infinite number of possible cases. Is infinite divisibility not also present where proportions of size are known, where we set the probability for a plane fragment to be proportional to its size? It may as little be claimed that we always have to have recourse to a least (the smallest) known part. We cannot establish an equivalence between the smallest parts of loci known to us, and other larger regions alien to us about which we know nothing, and hence whose partitions are unknown. Such a principle cannot be considered in the absence of a characterization of the nature of the partitions in question, for which a notion of a 'real' division is at least insufficient.

If one considers things in a less fictional light, then there is no doubt that if we need to calculate the probability of a meteor strike on any known portion of the earth's surface (the Fiji islands or the Australian continent; the city of Paris or the

lakes of Romania) then we will attempt to produce a graph of relevant surface areas and their relation to the whole surface of the earth (even if only to a coarse approximation). There is no doubt that the results may vary within wide bounds, given insufficient information.

Another example which I drew – and which **Stumpf** had criticized – concerned the probability that terrestrial elements are present on an extraterrestrial body. There I began with what I considered the simplest of conditions: that (in absence of all positive knowledge of the subject) one might be inclined to assign as 0.5 the probability that the star Sirius contains iron – and the probability it does not. I can agree completely with the objections that **Stumpf** raises against this. Yet in no sense did I propose that this assignment was correct – I only wanted to show that we have no adequate point of reference to make a definite evaluation. One might have been more justified in posing an objection to my demonstration at that time, that it may be quixotic to treat questions of the presence or absence of all the elements independently, since when one terrestrial element is present (or when a few are), the probability is at least increased thereby that others are present too. Just these considerations, as well as **Stumpf**'s considerations which hang on atomic weight, serve to teach us only that questions of this kind do not lead to definite assignments of probability. And finally, when **Stumpf** obtains the result that it may be impossible to give a definite figure for the number of equally possible cases, there is not even a convoluted notion of probability which will save his argument (*op. cit.* p. 74), and so his result coincides completely with what I had hoped to demonstrate.

If the treatment that logicians give to the probability calculus has often led directly to results which I cannot say are unsatisfactory, then I believe the reason lies principally in the essentially formal approach which motivates them to attribute a far greater significance to disjunctive judgments than those judgments really have – which is surely and evidently characteristic of judgments in the problems set out here. The existence of a close connection between the measurement of probability and the disjunctive form in logic is self-evident, of course. As with any measurement, so here we stand in need of a unit which is cited in reports of measurement. Only with that should we consider it completely certain. On the other hand, it is understandable that logical contingencies will always lead us directly and immediately to attribute some proportion to the probabilities of two or more premises which are uncertain. It follows that evaluations in absolute terms result only when, for a number of probabilities, their ratio is fixed, their total is fixed, and their sum is assigned a value of 1. That is the case, if we are allowed to attribute a definite ratio to the probability of several mutually exclusive expectations, and if we know that one of the expectations is realized. Then of course it follows that any assignment of probability susceptible to numeric measure must be represented as the ratio of probability between two (or among several) members of a disjunctive judgment. Still, it follows no less that this discovery does little for us in itself – even less so, because for example any judgment can be placed in contradiction with its denial in the context of a disjunction. When is such an evaluation possible, and how might it ensue? – The form of disjunctive judgment is mute on this at first. There is no doubt that whenever the connection between the probability calculus and disjunctive

judgment is emphasized, this proceeds tacitly from specific postulates about its nature. At first in connection with the conventional conception of disjunctive judgment, the argument might run that members of the disjunction refer to formally distinct types of behavior, but that they are not demarcated by some sort of boundary drawn across a continuous variation in behavior. However, against this remark we must pit the claim that it is very often just under the conditions when a monotonic gradation seems possible for some behavior, and when individual behaviors are demarcated by the drawing of some arbitrary boundary, that evaluations of probability may very well be possible. Here the connection to disjunctive judgment only leads to the neglect of a case of primary importance – one might even say the very most important kind of case. – Let us consider here that case which authors may have dangled at first, in connection with successes in games of chance: that members of a disjunction refer to varieties of behavior which are formally distinct, and which are not unified by transitional types. Then here too we find an emphasis on disjunctive form to be less conducive than it may seem to be at first glance. Of course in turn it is obvious that, if we can assemble a large number of equally probable cases, the probability of the separate cases is given by their total number, and that the probability of any behavior given under a general rubric and which covers several cases, is a probability determined by the ratio of those cases which are covered by it to those which are not – i.e., the ratio of the ‘propitious’ cases to the ‘unpropitious’ ones. But this is still just the means of calculation, by which we derive novel results (ones which are essentially only formally different) from a given basis. Despite that, what really matters here is that the initial assignment of equally probable cases is simply assumed as given. What is in question, how this is to be achieved, and under what conditions this is possible: the point of reference that is the disjunctive form offers us no further clue to any of that. It is reasonable to focus on many other requirements which might satisfy the nature of disjunctive judgment in leading to equally probable cases. Foremost among them is the requirement that different varieties of behavior must pick out ‘coordinate cases’, or in **Sigwart**’s more noteworthy expression, this is about “equally differentiated speciations of general behavior”. Surely one must concede that this is more about setting a problem than it is a solution to a problem, since what is to be understood by all this would at least require more incisive investigation. Nevertheless these are the ideas whose pursuit leads to the point of primary interest here. On the other hand, it seems to me that if one lays aside those requirements, and stresses only the exhaustion of knowledge in disjunctive judgment, in other words if one pursues the intuition we have called the Principle of absolute ignorance, then one overestimates the significance of the disjunctive form in the most outright way, and one overlooks that which is most important.

An investigation which begins with the form of disjunctive judgment (connected to the idea of coordinate cases or that of ‘equally differentiated speciations’) will always have to emphasize this as its central notion: that between the conceptual referents of the members of the disjunction there must always be some relation or other, or (as one may say more generally) some relations of equivalence must always be there. Of course a more detailed pursuit of this idea requires us to pay close

attention to the conceptual content of our notion of reality. And strictly speaking, that tells us this requirement is not fulfilled by a disjunctive judgment with a limited number of members, but rather it may be fulfilled by more unsaturated judgments which encompass a certain range, and which concern behavior that can be represented by continuous magnitudes.<sup>8</sup> – And so as promised, we have indicated the point which naturally comes into the foreground in any account which begins with disjunctive judgment. As a secondary matter, it may then be proposed there is another requirement: that nothing in our stock of knowledge may postpone an evaluation of probabilities as ratios of magnitudes. That requirement boils down to this: that in certain very specific respects there is absolute ignorance on the subject whether certain types or collections of considerations ought to be set completely aside. And in certain frames of mind, this is the case. Then we are not led into unrealizable fictions when we try to base the equivalence of members of the disjunction on the lack of any knowledge, just in case we have no basis given by any relations of magnitude. – And in addition it is clear that we cannot answer the decisive question (about equally probable cases) with any finality, without giving a full account of the most basic foundations of any probability, and without accounting for the psychological nature of comparison judgments, making a particular distinction between the psychological meaning of equality and the mathematical concept of equality. Having done that much, then we run into considerations that were discussed previously (*Logik*, Chapter 19, pp. 399–435). They tell us that probabilities of our assumptions or expectations which concern any real proportions do not in admit precise numeric evaluation in general. These considerations also beckon us to characterize in an exhaustive way those particular propositional attitudes which are connected with the formal establishment of equivalence for probabilities.

It is unnecessary to return to these explanations once more *in extenso*. Yet this may be the right place to shine light on a few specific points. The disregard of these points, or their misunderstanding, or errant judgments made about them, may have been most responsible in forestalling recognition of the Range theory. The first is that concepts which we apply to the behavior of real things are largely of a kind that permits continuous variation. And this is closely connected to the notion that our uncertain knowledge is generally represented in the form of unsaturated judgments – as we saw, a disjunctive judgment (with a finite number of members) represents a special case of an unsaturated judgment connected with exceptional conditions or singletons. It may be of some interest to show how this holds even for games of chance. For example in roulette, within broad limits the force of the impetus given to the ball is unknown to us. But since red pockets and black pockets alternate regularly with one another, the ball must either come to rest on red or come to rest on black.

---

<sup>8</sup>One may note that these reflections (developed from the notion of a disjunctive judgment) do not simply lead to the Range theory on their own because we consider the number of members of the disjunction to be infinitely large. Rather the decisive point is still this: that there is a relation of equivalence which holds among the infinite number of members of the disjunction. Though that relation of equivalence encompassed by any single range may become infinitely small, still there must be an equivalence.

When a coin is tossed, or if a die is thrown, the circumstances (the orientation, or the linear and rotational speed of the object thrown) which determine the outcome are unknown to us. However, the form of the object which hits the plane of support determines that it can come to rest in one of two orientations (or one of six, respectively). Now of course it may be admitted that such peculiarities of causal connection often occur, by which continuous variations in constraining circumstances correspond to an appreciable number of outcomes that are sharply distinguished. At the same time it is important to keep in mind that in our everyday knowledge of real things, the most common case by far is not represented by a finite and exhaustive number of members of a disjunction, but rather it is represented by judgment most often given to us by our knowledge of real things, that is, by comprehensive judgment in a case that is unsaturated with respect to an overarching concept. Even in games of chance, we are led right into an inexact state of knowledge in this sense as soon as we ask – not about evident outcomes, but rather about the constraining circumstances which are prone to bring them about. Then the role of disjunctive judgment has been usurped by the far more common form of an indeterminate proposition, which represents a certain open range of behavior. In the theory of probability, too, we see that this case must be dealt with initially, rather than being treated as a paradoxical case.

The other point to be mentioned here concerns the notion of comparison judgments. We set out to show that the notion of equivalence which they imply must be distinguished from the mathematical notion of equivalence. The former notion arises as an expression of incidence, in other words by inclusion under a concept which is more or less indeterminate. The shifting foundation and the inchoate nature of such judgments are connected to the peculiarities of such incidence relations. One must keep an eye on all such relations, just to be able to accurately evaluate those judgments which predict a balance between a reason for one thing, and a reason for its opposite. Whoever is incapable of affirming this fundamental notion wholesale, but who is instead of the opinion that equality might mean something which holds in a strict sense between the most varied of objects, and who then attempts to discover how to express this in only one way that is valid, that person will declare that a search for specific conditions for the reliable evaluation of probability is uncalled-for as a consequence. That person must defend the position that, insofar as certain propositions serve as the reliable basis of our empirical knowledge, it may be taken to be only a matter of the computation and the solution of a combinatorial problem to assign a definite mathematical value to any premise, which mathematical value is its probability. By contrast whoever takes on our basic intuition as his own will find nothing difficult and nothing peculiar in the idea that the mathematical evaluation of probability is tied to certain propositional attitudes or mental states, and that it is specific to just those states.

In addition we may cite a misunderstanding which may threaten this notion of the measurement of probability as confined to certain behaviors, and its connection to specific knowledge. Of course we ought not to confuse knowledge of objective content, by which reliable assignments of probability are produced, with the conventions themselves. The judgment that all throws of the dice arise from equally

large ranges of conditions, is not identical to the judgment that we expect any throw of the die with equal probability – rather it provides a basis for the latter. Then it would be a misunderstanding if one wished to claim that the Range theory may swoon into the same error which infects many older theories: that probability is purely and simply something given by real ratios, without regard to anything which may be contributed by knowledge (however it is formed) or by its lack. A remark made by **Stumpf** seems to me to indicate a misunderstanding of this kind, when he says that people who subscribe to **Laplace**'s definition will not consider cases to be equally possible just because physically equivalent magnitudes are in question. Rather, they will consider them equivalent because we find ourselves to be in the same state of ignorance about all of them (op. cit. p. 684). It is not only proponents of **Laplace**'s theory who would consider matters this way: proponents of the Range theory will also, by the same token. If a proponent of the Range theory proceeds on the basis that we know that any throw arises from an equally large range of constraining circumstances, then in that he recognizes only what allows us to expect all outcomes to have the same probability. And if one wishes to ask more particularly what the actual basis for the equivalence of probabilities might be, then under this conception one should find – one must find – that it lies in the strict uniformity of the reasons which speak for one proposition or the other, in the complete absence of any factor which might favor one premise over another.<sup>9</sup> Thus the Range theory does not stand in contradiction to the fundamental ideas of **Laplace**'s theory. The Range theory only supersedes it insofar as it is found necessary to ask about the states of propositional attitude under which such a perfect uniformity may be given. If one considers the conditions in this context, one can make no objection to the claim that a certain positive knowledge underlies the numeric evaluation of probability, and that such knowledge is expressed in the evaluation. It is basically self-evident that probabilities are always the result of propositional attitudes: in other words, they must be conditioned by what we know and what we do not know.

Finally in this context, we may touch upon another objection which has been raised against the Range theory and which equally well may rest upon a misunderstanding. Even in my earlier writings, I had given particular weight to the notion that this theory resolves a contradiction, between positions that have been sustained in theories of probability for a long time. These positions may be described shortly as an understanding of the concept of probability in a subjective sense, and its understanding in an objective sense. Among the former theories I understood some very cogent claims emphasized by logicians: that probability means nothing at all concerning that which we can express about objective conditions alone – rather, it is a logical relation which depends of course on some mental state and at least some incomplete (that is, inexact or unsaturated) knowledge. In apparent contradiction to this account ran an account developed from games of chance, under which

---

<sup>9</sup>This is consonant with **Lourié**'s account, where he says that all the positive knowledge offered by the Range theory only serves “to smooth the road to complete lack of knowledge”. (1910, *Prinzipien der Wahrscheinlichkeitsrechnung*. Tübingen: J.C.B. Mohr, p. 146.)

probabilities may be defined in a very specific way from objective proportions. That is an intuition which carried considerable weight when the distinction was stressed between the situation when the ‘true and accurate probability’ of an event is known, and when it is not. If at first this objective interpretation of the notion of probability should appear unfeasible and inconsistent with the one mentioned earlier, this contradiction is eliminated once one begins with the notion that these are probabilities which represent a very specific propositional attitude, and hence the whole investigation is limited to domains whose idiosyncratic nature it is to permit such a very specific propositional attitude to be assayed. This does happen under some conditions, and the classical example is found in games of chance. When outcomes depend on infinitely small variations of constraining circumstances which are wholly outside our powers of observation, then here on one hand an impassable boundary is set as a limit to our understanding. On the other hand, there is the possibility that we may come to know with certainty that there is a ratio between the ranges which bring about each of those outcomes – a ratio which is equal for large numbers of similar cases, as a rule. Then it happens that there is a definite way of thinking tacitly assumed by the said limit, but which in the second respect has already attained everything accessible to us. In this regard one may say that here there is a probability *par excellence*: this is what is specified in a determinate way by objective proportions. Therefore while it would be nonsensical to ask how large the probability *realiter* may be that it will rain today, still one is entirely in the right to examine or to report how large the ‘true probability’ may be of throwing a six with a single die which has a measurably eccentric centre of mass (etc.). For certain ranges we can speak of probabilities determined by ratios, and in that much we may speak of objective values of probability, but that is by no means the situation generally. We may speak of such objective values without being distracted from the fact – or without contradiction – that here too, probability refers to a logical relation conditioned by a propositional attitude: it should be taken to have subjective meaning in just that much.

If this is taken as a completely transparent and unbiased inquiry, then some misunderstanding is indicated in **Lourié**’s statement about my work,<sup>10</sup> that “what is most characteristic is an unparalleled collision of subjective and objective viewpoints”. Such a misunderstanding becomes even clearer immediately afterwards, when he asks: “Why is there such effort to establish an objective reference point, when its whole value for the result is destroyed at a stroke, and one stands before the abyss of arbitrary subjectivity ?”. Here **Lourié** seems to identify the expression ‘subjective’ with ‘arbitrary’. This interpretation is completely alien to me – as it always has been – and it is not in that sense I emphasized the subjective nature of probability, even in the passage he cites. Only in the sense that has just been mentioned – in the sense properly emphasized by logicians (in contrast to older theories which completely ignore this aspect of the real meaning of the concept of probability) – should the connection of the notion of probability to incomplete

---

<sup>10</sup> Op. cit. p. 176.

knowledge and to propositional attitudes be accentuated, and only in that sense should its subjective meaning be emphasized. By contrast it is the fundamental idea on which the entire Range theory rests, that under certain propositional attitudes, probability values are found which are entirely unbiased and compelling. Therefore they are in no sense arbitrary, but they are subjective if the term is to be taken in that way. The idea is that the establishment of such a convention is not guaranteed in general, but it may work under particular conditions. And with that we are set a task, which is to find those particular conditions, and to say how it happens that they lead to a stable and compelling evaluation of probability.

Concerning the formulation of conditions expressed by the Range principle for an evaluation of probability in numbers, originally I was led to this by a circumstance that seemed to be present just in games of chance, but which on further examination led to the conditions established here. Quite apart from the question whether or not those conditions are indispensable to a numeric characterization of probability values, naturally it may be asked if such conditions really are fulfilled in games of chance. Here it should be mentioned first, that extended experience has acquainted us with the almost-equal frequencies in which red and black come up in roulette, and in which each of the six sides of a die come up over a long series of throws. A strict examination of the logical relations involved produces a justification for deriving the conclusion that the various outcomes are produced by ranges of constraining circumstances which are of approximately equal size. And there are a great many subjects where we can only ground our intuitions about ratios of ranges in this way. However, these conjectures proceed from logical principles which are not universally recognized, or which at least are not in common use everywhere (even if, strictly speaking, they stand in no need of proof). In this regard it is worthwhile for us to demonstrate the equality between ranges of constraints – right there in real games of chance – which constraints lead to one outcome among others, and it is important we can demonstrate this directly with the help of purely physical considerations. This needs to be achieved in a way that can be called deductive, in contrast to one supported by the results of collective phenomena and which may be called empirical. Such a proof is the subject of considerations which I have already developed in *Principles of the probability calculus*, and which it may not be superfluous to return to briefly. First let us consider an ideal situation, like the bowling game<sup>11</sup> (A ball is given a push by which it moves along a straight and flat alley, whose length is divided crosswise by alternating red and black stripes of equal width.) Here we can see right away that, if we consider the strength of impetus to vary continuously, then the ranges of such forces as give rise to an end result of red and those which give rise to an end result of black, must alternate constantly. And if we assume that the stripes are infinitely small in width, then the ranges which give rise to red and those which give rise to black become exactly equal, however we imagine force to be measured. Nothing about this changes, if we take into account circumstances which occurred earlier in time. Of course it is evident that the most varied of circumstances will have

---

<sup>11</sup> Cf. my *Principien der Wahrscheinlichkeitsrechnung*, p. 49 [here Chap. 3, Section 1].

some influence in determining whether a stronger or a weaker impulse is imparted to the ball in each case. But since – at least in a great number of respects – the connection between constraining circumstances and their consequences is of the type that continuous variation in the one represents continuous variation in the other, then there is no doubt that even those more distal ranges of constraints which lead to outcomes of red or black will still be equal in magnitude. And so, according to the measure of what we know on the one hand about widths of stripes and about general laws of the succession of events on the other, we really may claim with certainty that the ranges of constraints which lead to one outcome or the other are in fact of equal size. Then here we have an example of an ideal game of chance, which in a way absolutely satisfies the logical requirements that have been set out. – How do things behave, if we do not have a game of chance in our sights that is perfected in an ideal way, but instead a game of chance which is really in front of us? Here let us consider the more complex case of dice games. Simple considerations of mechanics tell us that the behavior of a single die, in the instant at which it leaves the hand, is fixed by nine values which determine its position and its motion exhaustively.<sup>12</sup> If we think of these as varying, then we obtain a manifold determined by nine variables, and a precise knowledge of physical laws would allow us to tell which parts of the manifold lead to each of the six possible outcomes. If we compare the conditions given here to those of an ideal game of chance, then we note that here too, the elements of that manifold which determine each throw will succeed each other in a regularly alternating fashion. And certainly, at least for a part of them (namely for rotational speeds) the situation is that even very small changes will induce a shift in the outcome. Within the entire range – which seems possible in a subjective way when we reckon it by the measure of our imprecise knowledge – any outcome of a throw may be produced in altogether different ways. Next let us suppose that the game may be performed with an ideal die. Thus we would have to claim that each of the ranges of conditions which lead to each of the throws 1 through 6 (and which are to be considered parts of that manifold determined by nine variables) are equal, to a fine degree of approximation. Of course in supposing this, we run into some difference from the ideal case. Since this concerns the comparison of small but not infinitely small ranges of conditions, it is conceivable that the entire range which leads to each throw does not turn out to be equal to the others in a strictly mathematical sense. Instead they may depart from it by a minute amount, should we consider that the parts which produce the various outcomes are mediated by general laws governing events. It is also conceivable that things will turn out somewhat differently, depending on our choice of variables by which we determine those conditions, and the way we think of measuring them. That will hold true for the outcome given directly by the nine variables alluded to earlier, but just as well for

---

<sup>12</sup>It is known that the position and the speed of a rigid body are exhaustively determined by six values apiece. However, these twelve values are reduced to nine for the result under consideration – the throw of a single die – since the outcome is unchanged if the point of origin is translated horizontally, or if the whole system of the die (its position and speed) is rotated about a normal axis.

conditions which are farther back in time, to which we may also resort. Further, we must always reckon with the possibility of very small geometrical eccentricities in the centre of mass, or suchlike. If such possibilities given by the general conditions of individual throws are not known with absolute precision (because of these circumstances or others like them), then the totality of outcomes which is to be expected with maximum probability over many cases cannot be reported exactly, but only approximately.<sup>13</sup> The value of those considerations is not affected as a consequence by the merely approximate realizability of certain ideal conditions. That is because even an understanding of real games of chance – and a correct assessment of their logical contingencies – is not possible without revealing the involvement of the Range principle, even where it does not come to be known in a completely pure and unadulterated way. Yes, one may say that that which is given in real games of chance, and which is worthy of attention, can only be described most simply and cogently as a thoroughgoing approximation to that ideal case. And of course it has to be noted too, that the deductive stance we take to ratios of ranges does round out our understanding in a valuable way. Doubtless the importance of the Range principle is enhanced to a large extent by the ease with which it may be grasped and intuited, just when we can specify and describe the ranges in question in a very immediate way. And this importance is undiminished when the manifestation of the ideal case is only approximate. In this respect things are no different than for the theoretical and logical significance attributed to games of chance in general. If one wants to describe the characteristic relation which obtains between these particular ranges of real varieties of behavior and general principles of logic, one must always remember that the Range principle has entirely general purchase – as has been stressed repeatedly. It applies under the most varied of circumstances. Everywhere that our knowledge is imprecise, it contributes to the determination of probability. Nevertheless, under particular conditions it leads to pointedly conspicuous logical manifestations. Firstly, it is that probabilities can be evaluated in numeric terms; secondly that for certain varieties of behavior, meaning for the totality of results over a multitude of cases, enormously large probabilities are produced which approximate absolute certainty; thirdly that these probabilities may be deemed to be general in application. In games of chance, all these conditions are given to such close approximation that they must have seemed remarkable just from an everyday perspective, and even more remarkable in the practise of science. Thus games of chance represent a domain of actual processes for which certain pointedly conspicuous and very particular features of the Range principle do hold to a very close approximation. As a consequence this domain is especially well-suited to investigation of the principle on one hand, and on the other hand to investigation of conditions on which those particular features of its application depend. Then one may attribute an outstanding heuristic significance to all the logical conditions of games of chance considered here. The circumstance that the ratios of interest to us are realized only

---

<sup>13</sup>Cf. my *Principien der Wahrscheinlichkeitsrechnung*, p. 77 ff. [here Chap. 4, Section 2], on the difference between ideal games of chance and games actually played.

approximately in actual games of chance is as little an impediment, as that probabilities may occasionally be deemed to be equivalent under other conditions even if their equivalence cannot be given a similarly strict basis, though it may still be recognized as justified.

To some extent, a similar heuristic significance may be attributed to very many domains. Earlier we emphasized that the most diverse array of collective phenomena can only be understood fully if we trace the collective phenomena back to logical principles developed from games of chance. All of them represent analogous problems in that much. It is simply that the solution is made so much easier with games of chance, in that they approximate (to a great extent) an ideal case which can be characterized precisely. In this respect the deductive treatment just mentioned is especially important, in that there is a possibility of creating an immediately intuitive model in physical terms for the ranges in question, as well as for their ratios. In that respect it is a question of great interest – though a question perhaps more of mathematical interest than of logical interest – whether a similar treatment should also be possible in other domains. It is clear that such an interpretation is ruled out for the collective phenomena of society. And perhaps here we can make it clear generally that the arrangement of each individual case is conditioned by a series of distinctive ratios of spatial and temporal coincidences – that is, insofar as relations of ranges may be considered at all. Yet of course we cannot speak of any strict mathematical treatment here. Doubtless there are many other domains which may offer some foothold for consideration by analogy. One may be reminded of the theory of error, and of mixtures of rigid bodies, and so on. We have to refrain from pursuing this somewhat tempting investigation at present, however. Yet our attention may be called to the possibility that the formal comparability of ranges – and with that a certain basis for the evaluation of probability – may also be brought about by other principles than those of regular and periodic permutation, as hold in most games of chance. One may also wonder if, besides the bowling – game, there are other ‘ideal games of chance’ which rely on another principle altogether. This conjecture becomes particularly suggestive in application of the probability calculus to theoretical physics. We should delve into this matter in more detail, since the subject of this investigation is of great interest to several different studies in logic too. In doing so, we will find once again that we encounter the question just raised.

It is known that several areas of physics have a close connection to the probability calculus, namely thermodynamics and the molecular theory of gases. From a logical perspective, it is also of great interest to ask about the investigations physicists have made in these domains, and to examine the way in which the probability calculus is applied. It is of special interest to alleviate certain contradictions which at least seem to hold between these investigations and our own ideas about the theory of probability. In the kinetic theory of gases (to which we will restrict our remarks here), one begins with the premise that certain laws of mechanics hold true for the collision of molecules which fly by one another at random – the laws of elastic collision – and that the movements and arrangement of individual molecules are not otherwise bound by any generally stated rule, which is to say that they are independent of one another in the sense given by the probability calculus. If one proceeds on that

basis, then it follows that certain phenomena which we observe routinely, such as the transfer of heat from a warmer body to a colder body, may not be derived by necessity from the laws which govern the transitions of events. The contrary eventuality – that for two volumes of gas brought into contact, the warmer one would become still warmer and the colder one still colder – is conceivable without contradicting such laws. Such an eventuality may be called possible in that sense. But that could only occur if just the more slowly travelling molecules of the warm gas were to collide ‘randomly’ with the more quickly travelling molecules of the cold gas – shortly put, given a very special arrangement of their positions and velocities, in other words under conditions which are enormously improbable. This quite abstract idea is due to **Boltzmann**, who based it on a series of considerations, some more formally defined and some drawing on details of the phenomena.

Next, a couple of peculiar features of the whole subject will be indicated, which stand out as being worth of attention in this regard, where considerations of measurement which make sense of the notion of probability find their ultimate foundation. The first runs as follows. Let there be two molecules which have such positions and velocities that they collide with one another, and a very great change in their velocity is effected by their collision. Then let us suppose that the behavior which occurs before collision is changed but a little, enough to allow them to fly by one another rather than colliding. Still smaller variations would suffice to make their impact more or less eccentric, so that then the outcome would also be greatly changed. So minimal differences in present behavior represent coarse differences in behavior at later points in time. Then the prediction of events becomes impossible as soon as our knowledge of present or antecedent behavior is imbued with even the slightest measure of uncertainty (as is always the case). The course of events depends on details which are entirely outside our recognition. This is a matter of randomness in an entirely similar way – and in an even more strict sense – to the games which have been cited. Then let us add a second circumstance: the same pattern is exhibited a huge number of times under our conditions of observation, and the same processes repeat themselves. Some of those also act on our sense organs in ways which are primitive and cannot be distinguished further – they are elements of such dimensions, that the number of molecules within them is still overwhelmingly large. Then the varieties of behavior which are available to our observation acquire the significance of a totality of outcomes for very many cases covered by the same general conditions. Similarly to the totality of very many cases in a game of chance, it is clear that for such phenomena – meaning those which are observable to us – definite rules may be established with approximate certainty about what may be expected. Even in this somewhat condensed description one may find it reasonable that the most probable result by far is that there will be a distribution of molecules, for which there are approximately equal numbers of molecules and equal amounts of energy in spatial regions of equal size.

The task set out here should count as having been resolved in full, once we are able to achieve a goal similar to that which is achievable (and which has been achieved) for games of chance – specifically for ideal games of chance. Mainly it would have to be shown – meaning that it would have to be derived from general

assumptions which have been made – that that which we see to occur regularly (though it may not be derived from laws governing the succession of events) represents a predominant range in the arrangements of the constraining circumstances.<sup>14</sup> These regularities would be explained in this way, just as they have been set out for games of chance, in the sense that this is feasible at all. At the same time even our assumptions about those same regularities would become legitimate for our expectations of future cases. Then evaluations of operational ranges would have to be based on a compelling comparison of magnitudes that is both unbiased and non-arbitrary. But of course the requirement would also have to be satisfied, that the probability of a behavior cannot be established in isolation from the probabilities of other behaviors, to which it is connected univocally and necessarily according to laws governing the succession of events. That conjectures about probability, and the explanations which are based on them, etc. are consonant with laws governing events in this sense (a point which will prove to be especially important in a moment) – may be called their nomological admissibility.<sup>15</sup>

In that sense, a conclusive resolution of the problems about probability which are posed in the kinetic theory of gases is not to be derived just from the investigations which have been made up to this point. And it is even less fitting for us to attempt such a thing here.<sup>16</sup> We ought to confine ourselves to examining and explaining some of the investigations which physicists have conducted – and the interpretations which have taken root there – from the perspective offered by our conception of the task.

A first subject that I would like to discuss in this light, is the so-called Maxwellian law of the distribution of velocity: that is to say the manner in which this has been

<sup>14</sup> Of course in a certain sense, one might still rest satisfied with another point of view. One might say that it is justified to conclude from the regular occurrence of certain phenomena, that they represent a predominant range of constraining circumstances. Then that deductive basis for assignments of probability would be lacking, which basis was usefully rounded out by enlisting our knowledge about games of chance. We would be in a similar position, as with games of chance whose general conditions were unknown to us initially, and which were instead only concluded from the totality of outcomes over a large number of cases. In that much it seems comprehensible that this intellectual stance cannot be entirely satisfying here, since the postulates for a deductive treatment are given entirely by premises about the physical situation. If nonetheless the approach does not succeed, that implies we cannot prove the purely mathematical relationship of whose existence we are convinced.

<sup>15</sup> This is a matter of the same requirement which had been described in *Principien der Wahrscheinlichkeitsrechnung*: that for the size ratios of ranges to determine a measure of expectation, they must be original. It has already been mentioned above (*Logik*, Chapter 20, p. 451) that the assignments of probability in games of chance do satisfy this requirement.

<sup>16</sup> It may be enough to indicate that the investigations which are required here would have to resort to another mathematical principle than one which is considered for most games of chance. In those, the regular and periodic repetition of observable outcomes under continuous variation in constraining conditions offers a very simple basis for consideration – a basis which is evidently lacking here. Much will depend whether the existence of fixed and definite size ratios may be demonstrated on other bases than this. (It may well be that this question is connected with an issue of subsequent general mathematical interest.)

derived and justified in **Boltzmann**'s work. I believe his derivation must be considered in this sense: it fixes and integrates our theoretical requirements in a particularly interesting way. If a gas is contained in an adiabatic flask over time, of course the presence of irregularly occurring collisions means that some molecules move with much higher velocity, while others move with more negligible velocity. But just as we see that local inequalities of temperature and density come to equilibrium, one is justified in assuming that a stable condition emerges in that situation, and therefore a nearly fixed distribution of velocity emerges gradually and is steadily maintained afterwards. That would need to be represented by an expression for the function  $\varphi$ , which function assigns a value proportional to  $\varphi(c) dc$  to the great number of those molecules whose velocity falls between  $c$  and  $c + dc$ . If one proceeds on that premise, the question arises what form the function  $\varphi$  may have. **Maxwell** assumed the provisional form  $c^2 e^{-hc^2}$  for it, where  $e$  is the base of natural logarithms, and  $h$  stands for a constant fixed by general and lasting conditions. Later **Boltzmann** provided a formal justification for this expression. – Now if one examines **Boltzmann**'s derivation, it is apparent that it depends simply on the principle of nomological admissibility. What it demonstrates is just the nomological admissibility of a convention about probability which produces the **Maxwellian** distribution, or a corresponding assessment in terms of ranges. If, without taking the course of events into account, we focus only on behavior at a single instant, then a definite probability cannot be given that the velocity of a molecule may have some one value, or that it falls within certain bounds. Rather, any assignment would be completely arbitrary. Certainly, the probability that the velocity lies between  $c$  and  $c + dc$  will always be proportional to the value  $dc$ ; but if it is set equal to  $\varphi(c) dc$ , then  $\varphi$  can be almost any function at all. – Then it becomes comprehensible that this function emerges from laws which determine the succession of events, as one pays closer attention to the process of collision. When two molecules collide, it is fixed by laws governing the impact that all kinetic energy remains the same after impact as before impact, meaning that the sum of squared velocities remains the same. On the other hand, the distribution of energy may change. From a state in which approximately half the energy is lost from each molecule, meaning that both velocities are about the same, there may follow a state in which the velocity of one is very large while the other is practically equal to zero, or vice versa. It depends on the particular features of the collision whether one state or another will occur. Those features are the result of the spatial arrangement of the molecules, the direction of their motion, and so forth. The probability of larger or smaller velocities – and consequently even the most probable distribution of velocities – will then (as one may put it somewhat shortly) have to conform to how easy or how difficult it is to produce approximately corresponding or else very different speeds for the collision – that is, by the sizes of ranges of antecedent circumstances. Then there is a difference here which distinguishes this from games of chance, insofar as these same conditions – that is, the velocities we are interested in as outcomes – also enter into antecedent relations. And one might think that as a consequence, the derivation of determinate probabilities might become impossible. However, from the postulates, there ensues a

consequence quite specific to the case under consideration. If we should ask about a distribution of velocities which is to be expected eventually under any conditions – similar to an equilibrium of pressure and temperature – we proceed on the premise that the varieties of behavior which may be given by some initial state and which may be known to us, bear no importance for the formation of more distal states. That is, the influence of the initial state gradually disappears. And if everything which may be known to us about an initial state is of no relevance to judgments about more distal prospective varieties of behavior, finally this must depend solely on determinants which produce persistent and constant results: the total amount of gas, and the total kinetic energy. It is reasonable that with such a perspective (quite similar in a way to the ideal case in which nothing at all would be known to us but the said circumstances of duration) the probability of any behavior at one point in time cannot be assigned differently from that at any other. We may call such a convention about probability a definitive assignment. Here the requirement of nomological admissibility can be construed to be especially simple in form. Of course the probability of any behavior (stated without reference to a point in time) must correspond to the probability of those to which its lawful outcome is connected, meaning all which can be called consequent states of that behavior. A definitive assignment of probability must have the property that the probability of any range of behavior remains unchanged despite a transformation induced by laws of the succession of events.

In order to arrive at a definition of the desired distribution of velocities from this, we must focus on the collision of a pair of molecules. Evidently we consider the state of both molecules at a specific time before the collision to be a little – infinitely little – changed by their location, their velocity, and their direction of motion. By the laws governing impact, the infinitely small range of behavior which results will represent a range of behavior which is infinitely small in turn at a specific point in time after the collision. In any case, the probability of the behavior which pertains to such a range is then proportional to the product of the six infinitely small elements (of space, direction of motion, and velocity for each of the pair of molecules), though its expression may also contain coefficients which depend on location, or direction, or velocity. Then the purely mechanical deliberation which stands at the core of **Boltzmann**'s derivation tells us that the products of those six differential terms are not equal in general for an earlier range and a later range. Rather, they vary inversely as the product of the squared velocities. In order for the probabilities of both ranges – the earlier and the later – to turn out equal, no factors which depend on location or direction may be appended to these products of differentials, though those which are a function of velocity may be. That requirement is even sufficient to specify the function at hand: this just proves to be the cited Maxwellian function. For the sake of completeness it may be added that in this assessment the probability of an earlier operational range and the probability of the corresponding later operational range are found to correspond even for the case in which a molecule flies away, or even if it bounces off the wall of the adiabatic container without change in velocity – this is left as an exercise for the reader, which one can demonstrate for oneself.

The considerations which have been sketched briefly here tell us that if every position and every direction of motion is presumed to be equally probable for each molecule, and moreover the probability of the various velocities follows the Maxwellian function, then this convention has the property required for a definitive stance on probability. To repeat, it consists of this: that if we assign probability at one point in time, then because of causal laws, the same assignment of probability results at any subsequent point in time. – We can also add an immediate consequence which may be more readily understood. If the probability of a certain behavior for each molecule is assigned according to that rule, then for a huge number of molecules it should be expected with near certainty that the number of them belonging to a particular range of behavior is simply proportional to that probability. The property of the said convention about probability even acquires a perspicuous meaning from this argument. It consists of this: if we expect something with certainty from this rule at a point in time, that does not require us to expect something else necessarily with respect to the laws governing events as a consequence – that is, we should not expect some departure with certainty at a subsequent point in time. Rather we are committed to just the said expectation at any subsequent point in time.<sup>17</sup>

In fact the requirement is that our stance on probability (in the theoretical sense outlined) must be consonant with laws governing events, and that is the requirement of nomological admissibility, which by deduction leads to the Maxwellian function. At the same time, the investigation tells us that an expectation of overall equal temperature and density is admissible in the same sense. – On the other hand it is reasonable to say that the whole task which has been posed here (in the sense given earlier) has not yet been given conclusive resolution by this proof. That is because initially we had used a prior assumption as a basis, the premise that the initial conditions known to us are unimportant to an evaluation of probability at subsequent times: that such an evaluation has to be definitive. But then initially it is shown that the convention about probability which interests us is admissible as definitive only

---

<sup>17</sup>On the other hand, it would be inappropriate to characterize the Maxwellian distribution of velocities as a stable distribution in the sense that it would necessarily have to remain the same over time once it has been realized at one specific point in time. This arises just from the consideration that the laws of force permit the reversal of all processes, which is a familiar point. Then just as well as any arbitrary state can transition to the state of equilibrium represented by this rule (even a notably uneven distribution of temperature, or a distribution of velocity which departs from the Maxwellian rule), then just as well it is possible (that is, it is compatible with the laws governing events) that a state of this latter type transitions to one which deviates from it. The requirement on which the deduction is based is not the necessary stability of a state (that is, one which results from laws governing events), but rather the requirement is the nomological correctness of an assessment of probability: that is just what defines the meaning of the result. – I have already shown in *Principien der Wahrscheinlichkeitsrechnung* (p. 192 and ff. in original; here Chap. 8) that **Boltzmann**'s argument must be interpreted in the way set out here, though his derivation gives no prominence to the point. The deduction carried out there corresponds to what has been said here, in what is most important. Nevertheless it seems to me now that the earlier presentation stands in need of some improvement on a few points.

under a very specific perspective, not however that it is the only possible view and therefore compellingly given. Then whatever may be known about the initial state, if we expect an equilibrium of temperature and pressure as well as a distribution of velocity which satisfies Maxwell's formula at a subsequent time – then the accompanying justification may indeed be doubted in this point.<sup>18</sup> With respect to the former point, it is clear that we can depend on the exceptionless results of observation – but the mathematical treatment which would be required to render these regularities comprehensible in the manner set out earlier, would have yet another loophole. It would lie beyond the goal we are pursuing to give further details of the proof which is wanted.<sup>19</sup>

On the other hand we should not neglect to mention some other formulations here, and to examine how the regularities of concern to us are usually expressed in the physics literature. There are formulas which may appear at least conspicuous on first inspection, from our perspective. First we encounter the proposition that it is always the case that states which are less probable transition to more probable states. If one considers that any evaluation of probability holds true under a specific propositional attitude, and must refer to it, one could ask what the propositional attitude may be that is presumed by the probabilities noted here. In order to answer this question, we may turn our attention to a specific instance of pertinent processes, such as the equilibrium of temperature just mentioned. For a gas contained in an adiabatic flask, we deem the state of equilibrium in temperature to be more probable than temperatures which fluctuate across locations within the flask. That is indeed the case, should we not know which conditions have remained unchanged over time. And it is just this very specific state of our knowledge which is tacitly assumed in advance by that description. So if we say that the less probable state transitions to a more probable one, then what is meant is probability under specifically imagined conditions, which basically only goes to show that we deem a state of a nonuniform distribution of temperature to be improbable, although we may know its initial constitution with certainty.

---

<sup>18</sup>Similarly it may be asked in the ideal case about which nothing is known to us apart from the conditions of continuing constancy, whether there may not be several types of behavior whose probabilities may not be compared with one another in a non-arbitrary way.

<sup>19</sup>It may be sufficient to adduce that the proof alluded to in the second place may be furnished by some elementary considerations, in a manner I believe to be satisfactory. Shortly put, these considerations tell us that there can be no more than a single assignment of probability which is nomologically admissible in the sense given above (cf. *Principien der Wahrscheinlichkeitsrechnung*, p. 209 f.; here Chap. 8, Section 6). As far as the proof proposed in the first place is concerned, it would be a matter of showing that any familiarity with initial conditions is insignificant to the evaluation of probability for subsequent states (insofar as those initial conditions exhibit some imprecision). It is just at this point that the fundamental difference from most games of chance (which was mentioned earlier) comes into focus. What is missing here is the great simplification given there by the regular and periodic alternation of outcomes. As far as I can tell, the proof which is indicated here has not ever been conducted. It may well encounter some particularly high hurdles in mathematics.

It is important to understand that we do not mean the aforementioned propositional attitude to be some wholly arbitrary state of knowledge which is insufficient or incomplete. Rather we mean that recognizable and lawfully specified conditions are assumed to be known in advance. Then what are in question here are probabilities which can be characterized as being correctly measured, as we usually find for games of chance. That means such probabilities as may be produced when their “general conditions” are known to us, and to which a more general significance is accorded. Here we may speak of probabilities which are firmly characterized by objective relations in a still more formal sense than for games of chance. For one thing, this is because everything recognizable to us – or which we presume to know – is sharply and adequately defined; for another, everything that falls outside our powers of recognition is also sharply and adequately defined.

In yet another respect, the proposal that states which we see emerge gradually are the most probable states is a proposal that deserves some additional explanation. It is valid in a restricted sense, as when we want to say about a game of chance that what happens over very many cases is always in close correspondence to what we expect is most probable. Obviously such a formula only holds in a restricted sense. It is valid only so long as – and insofar as – we restrict our consideration of what has occurred to the totality of results: so for example in roulette, how often the ball lands on black or on red. By contrast, it is inappropriate as soon as we have actual events in mind, in all their determinate individuality. We have no warrant to expect – as could have been the case – that the ball will land first on red for the first two spins of the wheel, once on black, once again on red and then three times on black, and so on. And in that sense the course of events – that which actually occurs – always represents something which is exceedingly improbable to consideration of the situation in advance. This is completely representative of what holds for the behaviors we are concerned with here, too. That equal numbers of molecules and equal amounts of energy are contained in observable slices of space: that is something we should describe as maximally probable. And insofar as this is realized in any final state, we can say that the gas has transitioned to a maximally probable state. However, let us concentrate on the state which is actually realized, in all its detail, meaning the state consisting of the arrangement that each molecule is found in a quite specific location, and moves in a quite specific trajectory. Then obviously – as at an earlier point in time – we may only call this arrangement a totally improbable one at a later point in time. Therefore a strict formulation will not blindly attribute a higher probability to the later state. It is only that concept which covers a larger manifold of states by which we characterize what is observable in all this. It is the recognizable varieties of behavior of successive states to which we ought to ascribe progressively larger probability.

The correct interpretation of these conditions is threatened by a particular delusion connected with another way of describing the rule given above. One may also be used to saying that more organized states transition to less organized ones. And together with the other formula about increasing probability, one can easily tie this to the idea that organized types of behavior are purely and simply less probable – that disordered states are more probable. This is a fallacy which is completely familiar in

an analogous way within the domain of games of chance – one which **Laplace** had opposed emphatically. One may be inclined to think that in a game of dice, an orderly sequence of throws such as 1 2 3 4 or 3 3 3 3 is less probable than, for example, a sequence like 3 1 6 2 which offers no such recognizable order. Really the same probability is attributed to every distinct sequence. It is only that the count of the unordered sequences is far greater than the count of those sequences which are in some sense orderly. The corresponding mistake is also observed here. A disordered behavior, such as a state of equilibrium in temperature, is not to be called probable for all the specific minutiae by which it is actually realized. Nor is it any more probable than any behavior in which some coarsely observable order may be recognized. It is only the far more widely encompassing manifold of those states which are lacking observable order which justifies us in calling the presence of any disordered behavior a probable one.

The considerations given above suffice to show that important investigations conducted in physics – which have already proved their worth by their generation of empirical results – are satisfying and reasonable from our perspective. They may be incorporated in the framework of our theory of probability. That is, our deliberations show that there is no logical objection if we should trace certain actually observed regularities back to probabilities – or better, to ratios of ranges – rather than to laws in the conventional sense. There is no objection if we content ourselves with explanation in this sense, and we justify our expectations in this way. In fact, the conditions which hold here may in fact parallel the well-known and fully transparent conditions offered by idealized games of chance in every way. However, one still ought to say that some uncertainty – or at least some incompleteness – may be recognized in the whole way the concept of probability is treated in the physics literature. In the face of this uncertainty, a recourse to fundamentals fully explained in logical terms seems a somewhat worthy goal, a goal which would not be without its advantages. Doubtless there is something a little disconcerting at first, if in the explanation of certain regularities observed as exceptionless, it is said that which always occurs is just what is most probable for us. This impression vanishes once we make it clear that what is in question here are probabilities which are determined by very specific objective proportions, very likely even proportions which can be described directly. And we may compare this formulation – that that which actually occurs as a rule is just that which arises from a predominant range of combinations of constraining circumstances – as surely more reasonable and insightful than others.

On the other hand it will of course be necessary to emphasize that the expectations which we form here depend on a special principle, one which does not enter into consideration in the better part of our knowledge about the world. And at this point we have to agree completely with **Boltzmann** if he accentuates the fundamental difference between propositions about probability, and laws of nature in the properly narrow sense. I would like to emphasize the legitimacy of this contrast all the more emphatically, because one can find statements by leading authorities, which by contrast seem to point to a different interpretation – in my opinion an inappropriate interpretation. So for example, **Planck** says that if **Boltzmann** claims that certain processes which are never observed – such as the transfer of warmth from a

colder to a warmer body – are surely hugely improbable but not to be counted as impossible, then “we do not need to follow him in that much”. “Nature in which such things occur as the flow of heat towards warm bodies, or the spontaneous unmixing of two gases, would simply no longer be nature as we know it.”<sup>20</sup> As insightful as such a consideration may seem at first glance, nevertheless it is still deceptive in that it leaves the decisive point untouched. Certainly there is no doubt we do not have to deal with processes of this kind: they come into play only conceptually for us, rather than being realized. But that just raises the question, how and in what measure such an event is to be distinguished from actual events: that is, whether it deviates from actual events in some respects that we can codify in a general formula, or whether it is distinguished from them only in that we find what is realized to be constituted in the most complex way, one not subject to any laws we imagine. And when in the same place **Planck** says that: “just with the assumption that the lawful regularities which are defined here might not exist, the necessity of all natural events is restored by the principle of basic disorder, because fulfillment of this principle entails the increase of entropy as a direct consequence of the laws of the probability calculus.”, then he leaves it unstated (if not misunderstood) what the subject of the probability calculus is, and what one may call laws of the probability calculus.<sup>21</sup> Right there a distinction of fundamental importance has been blurred. So long as we assume that laws in the strict sense are only the laws of force which determine the collisions of molecules, but otherwise we proceed on the basis that their arrangements are not subject to any expression of lawfulness, and therefore that they are to be judged only by axioms of the probability calculus, then of course to that extent we can even grant warrant to our expectations about such processes. Yet if we want to be clear about the logical fundamentals themselves, we must keep in mind the basic distinction between something we ought to call maximally probable, and something we ought to expect to be lawfully determined – the distinction between that which is overwhelmingly improbable and that which is impossible in the sense that it stands in contradiction to natural law – the distinction between these two is the distinction between the Range principle and the principle of uniformity. Together they determine the certainty we have in our ideas about reality.

Finally as may be emphasized here, of course it is an undecidable question – and the topic of indisputably subjective judgment – if we ought to be satisfied or should be satisfied with this whole notion based on considerations of probability, or if here too the way seems open to us to trace the regularities in question to laws in the strict sense, in other words to trace them to theoretically definable types of congruences.

---

<sup>20</sup> Planck (1909). *Die Einheit des physikalischen Weltbildes*. Leipzig: S. Hirzel, p. 24. (Vortrag gehalten am 9. Dezember 1908 in der naturwissenschaftlichen Fakultät des Studentenkorps an der Universität Leiden.)

<sup>21</sup> However, in a more recent lecture ('Dynamic and static lawfulness', Leipzig, 1914) **Planck**, like **Boltzmann**, has emphasized the distinction between natural laws in the strict sense and propositions about probability, as an unbridgeable divide. And it almost seems, then, as if he may have changed his conception of those relations to some extent. Therefore I gladly leave open the possibility that I may have read more into his earlier exposition than his actual opinion represents.

Earlier it was indicated that even in games of chance there are difficult and obscure conditions (at least to the popular mind), and those have given rise to repeated misgivings. Should we find that for large numbers of spins of a roulette wheel, the ball does land on red nearly as often as on black, then this does not seem to be adequately explained by the circumstance that we are to expect such an outcome in advance with maximum certainty. And in just the same way, one is inclined to ask what then justifies us in claiming that it is possible for the ball to land on red a thousand times in a row, if nothing of the kind has ever happened before. We need not return to review the discussion which was developed previously, by which these questions were resolved. For games of chance where we may assume with confidence that we have exhaustive acquaintance with the laws in play, doubtless we should be content. But the situation is different here, where we are dealing with theoretical concepts which are still uncertain and undeveloped in many respects. If anyone feels that his intellectual curiosity is unsatisfied that we should consider an event which is certain to occur as one that is enormously probable given our vague knowledge – if that person feels the lack when we designate processes as possible (though enormously improbable) which have never actually been observed, then there is no room for debate. And one should not call the attempt unjustified – perhaps one should not even call it superfluous – to search for other concepts and other ways of thinking about reality which are free from that affliction; which allow us to understand what occurs regularly as not merely that which is maximally possible; and in that much to consider that such regularities stand in no need and in no want of an explanation – rather they are rendered comprehensible because they are established by a general law. What prospects such attempts may have, or what avenues they may open, is something that cannot be elaborated here. Still, here one may be reminded of the idea that the notion of temperature may not be reduced to relations of motion – that is, it may not be defined in kinetic terms – meaning that temperature might be treated as an independent and definitive notion. But of course it may also be thought that our premises about the laws of motion must undergo some modification or some expansion. These laws might have been represented as basic laws which were to underlie elastic collision in an approximate way, but might not have been represented altogether fully or strictly. And lastly a complete account would still have to mention that apart from the laws of force which govern the collision of molecules, there might be other kinds of expressions which may exist that might equally well be characterized as laws in the strict sense. The very effort to develop this idea – which idea may be both evident and completely impeccable in logical terms – will surely encounter some exceptional difficulties. Let us assume that the course of events in a gas with a given determinate state is given uniquely by laws of force. Then there must be some law in addition to those, which must be represented as uniquely constraining behavior at any instant – a law of simultaneity – if those laws are not to interact with one another in a way that is ruled out methodologically. Then a specific difficulty arises in the following circumstance. If our assumptions about the laws governing events are correct, the course of processes would be reversed entirely once we consider the whole arrangement which exists at a point in time to have been transformed into its temporal mirror-

image. That is, the arrangement would be transformed so that all molecules would retain their positions, and their magnitudes of velocities would remain the same, but their directions of motion would be reversed. A law covering simultaneity in arrangement, which is called for here, would have to be of the form that it allows one arrangement, but can exclude its temporal mirror-image. It is very difficult at least to have any idea of the formal nature of such a law.

# Chapter 12

## Conventions of Measurement in Psychophysics



**Abstract** Intensive magnitudes are not measurable in themselves, because the establishment of an equivalence between different steps in a scale of intensity does not make any sense without further clarification. Where intensive magnitudes are determined in a domain of the natural sciences, von Kries claims this is only a matter of counting, and of the measurement of temporal and spatial magnitudes. Every measurement of intensity should then be reduced to these operations by explicit conventions. Likewise, we can only speak of the measurement of sensations once we have established an arbitrary convention that determines what we will consider equal. The debate whether sensation varies with the logarithm of stimulus intensity, is then not a difference over matters of fact. Instead it is an empty dispute over words that is rooted in misunderstanding.

**Keywords** Psychophysics · Measurement of sensation · Intensive magnitudes · Conventions of measurement · Just-noticeable difference · Psychological scale · Multimodal scaling · Conceptual confusion

Any two stimuli that impinge upon our sense organs must differ either qualitatively or quantitatively by a finite value for their difference to be noticeable; as is generally recognized, E. H. **Weber** (e.g. Weber, 1834) both discovered this fact and explored it in some detail. Weber succeeded in establishing a definite rule for the necessary magnitude of this difference in a number of cases. Together with similar facts he discovered on his own, **Fechner** made use of these facts as a foundation for his ‘psychophysical law’, which asserts a definite relation between the magnitudes of stimulus and sensation: that sensation increases in proportion to the logarithm of the stimulus. Since this psychophysical law was first proposed, the nature of the functional relation between stimulus and sensation has been much debated. In

---

1882. Ueber die Messung intensiver Größen und über das sogenannte psychophysische Gesetz. *Vierteljahrsschrift für wissenschaftliche Philosophie*, 4(3), 257–294.

particular, critics have maintained it necessary that sensation increase in direct proportion to stimulus intensity.

I do not wish to enter into the arguments that have been brought to bear for one view or the other; rather I intend to show that the entire debate rests upon an assumption held generally and tacitly, but which upon closer analysis nonetheless proves to be an error (albeit a natural error).

A distinction can be made between the *theoretical* and the *practical measurability* of any thing. In general, measurement depends on an operation that brings distinct things into a relation of equality. For theoretical measurement, the issue is whether or not any such equality holds in a given domain: if the equation of two distinct parcels (if I may use this term for just a moment) is permissible in the domain and makes sense. If that is the case, the subsequent question of practical measurability may be framed. This means ascertaining whether a given concrete parcel is the same as (or different than) another concrete parcel. In general this is most likely to involve *realia* of identical extent at different loci of a system to be measured; or it involves processes at different times that have just the *same duration*; or it involves bodies at different places that occupy just the *same space*. These realia (processes or bodies) are our *standards of measure*. Their constitution is of paramount importance to practical measurability, though it is irrelevant to theoretical measurability. We can assert that certain categories of things are not practically measurable, without this having anything to do with their theoretical measurability. For example we can state that the distance of a fixed star from the earth becomes practically unmeasurable once it exceeds a certain value. Still, we do not doubt that the distance of any such star has a determinate value, even if we should never be capable of determining it. Obviously every attempt to carry out determinate measurements within any domain (to make it measurable *in practise*) presupposes some analysis of the theoretical measurability of the domain.

The following discussion will consider the theoretical measurability of intensive magnitudes; this will be elucidated first and foremost in the domains to which it is unquestionably applicable: the various domains of physics.

To better analyze the conditions for the establishment of any quantitative relation whatsoever, first we will consider a simpler case than that of the intensive magnitudes: that of extensive magnitudes. And here we focus first on the simplest of cases: the quantitative comparison in length of two line segments. Clearly such a comparison is predicated on an assumption: the assumption that there is a clear intuitive sense to the claim that two distinct straight-line segments may be *equal to* (or different from) one another. What really may be thought to follow from such a claim of equality (or inequality) is not a question likely to receive unanimous response from mathematicians or philosophers. It is not my intention nor is it part of my present task to enter into such a discussion here, since I am not concerned to expound the theoretical measurability of extensive magnitudes, but merely wish to use this example. However one views the matter, it is clear that the sense of geometry as a whole depends on the sense of such a claim, and that the whole of geometry would be nonsense if somehow that claim proved to be nonsense. The same holds true for comparisons of planar regions or of volumes; perhaps there it is even more

difficult to specify what might be meant by equality or congruence, but there must be assumed to be some sense to the notion if geometry is to be no delusion. If, however, these assumptions are fulfilled, they lend unambiguous meaning to all quantitative geometric relations. Then we can say that two line segments are equal, or that one is some multiple of another. After all, ‘many times as much’ just indicates the composition of a number of *equal* parts. And if ‘equal’ means something definite, it is also clear what is meant by ‘many times as much’. It would lead us far afield from our current discussion to show how the unbounded divisibility of spatial quantities provides an application for irrational numbers. It suffices to understand the main point of all this: the centrality of meaning of an equality between distinct elements. In the same way the same assumption holds for *time*, and assures the theoretical measurability of time as well.

Since we will be dealing with the measurement of physical magnitudes, we should familiarize ourselves with an essential measurement of a quite different kind, namely the measurement of *mass*. Here we encounter an entirely new principle for the first time: that of an arbitrary *convention*. This is as effective a means of comparison as the direct comparison just mentioned for elements of space and time. Now the requirement that distinct parcels be comparable is only satisfied unequivocally for one and the same substance. What we mean when we say that quantity A of gold and quantity B of gold are the same, needs no explanation. It signifies the equally frequent iteration of individual portions that match one other *in every respect*. In fact this business is just one of *counting*. The complete homology of individual portions makes this possible. But we might mean if we declare that the mass of lump A of gold is the same as that of lump B of copper, is not evident at all. Rather its meaning is bestowed by a convention that a unit of mass of a given substance shall be regarded as the same quantum, which has the same *weight* as a designated quantum of a certain substance (say one cubic centimetre of water at maximum density). We recognize two bodies as having the same weight if they produce the same alterations when they are placed on a compressible pad, or if they remain in equilibrium on a balance whose beam is of equal length on either side. Previously we referred to the convention for equality of mass as an arbitrary convention. In fact there are no logical grounds against adopting some other convention, for example that those quanta of any substance should be considered as equal, whose temperature is raised from 0 °C to 1 °C by application of the same amount of heat. This does not exclude the normal convention from being by far the simplest and most expedient. Actually the normal convention is simple and expedient only because, so far as we know at present, certain empirical laws hold strictly and without exception. It would lose this property if these laws proved to be inaccurate. The first law concerns the proportionality between increments in weight and increments in mass. If we have ten pieces of silver ( $S_1 S_2 \dots S_{10}$ ) each paired with a piece of copper ( $C_1 C_2 \dots C_{10}$ ) and the pieces of each pair are equal in weight, then we can be sure that the weight of all  $S$ s is also equal to the sum of the weight of  $C$ s. In other words the relation of equality of weight does not hold just between the quanta identified as units of mass, but also between any multiple of these. Imagine for a moment this were not the case, but that in different substances weight increased at

different rates with increments in mass. Then in order to compare masses it would be necessary to specify the quantity (say one gram) at which equality of mass was given by equality of weight. Then one gram of any substance would be equal to a gram of any other substance, both in mass and in weight. A hundred grams of one substance would also be equal to a hundred grams of another in mass, that is, it would still be a hundred times that unit of mass – but it might not be equal in weight. Should this account strike you as too difficult, perhaps the following will be easier. In general we assume that the gravitational attraction of masses is independent of their quality as substances, which means two quanta of different substances that have the same weight at any location on the earth's surface also bear this relation to each other at *any other* location. Things could well be different. If very different bodies were present in a very different distribution at separate locations on the earth's surface, and if the masses of those bodies exerted gravitational forces that were unequal for different substances, then it is conceivable that for a given pair of masses, one could weigh more at one location, and the other could weigh more at another location. If such were the case, the current convention in force about units of mass would be invalidated, and it would have to be replaced by another. This would also be the case if the weight of a body changed with alterations in its situation (through rotation), for example. So in evaluating mass we deal with a *convention* that is adopted for maximum utility in accord with certain empirically verifiable facts. Following the adoption of this convention we compare only masses of the same stuff – by counting. Indeed without this convention, it would make no sense to assert that a cubic centimetre of gold may be equal in mass to a certain volume of silver.

Next we wish to consider how other kinds of physical magnitudes can be measured that are neither magnitudes of mass, nor of space, nor of time. In the process we find that *all* of these physical magnitudes are formed as combinations of magnitudes of space, time, and mass (that is how these magnitudes are determined). A very simple example is that of *speed*. We measure this as the magnitude of a linear extent of space divided by a magnitude of time. Since such a division cannot be carried out, it can only serve as a symbol with some other meaning. This is indeed the case. What do we mean when we say that some speed is n times as great as another? This is immediately clear if we take the equation:

$$L \cdot T^{-1} = n \cdot L_1 \cdot T_1^{-1}$$

(where  $L$  and  $L_1$  indicate lengths;  $T$  and  $T_1$ , durations) and rewrite it in another form:

$$L \cdot L_1^{-1} = n \cdot T \cdot T_1^{-1}.$$

What we have then is a ratio of two durations (a scalar value) that is n times as great as the ratio of the distances travelled in those times (another scalar value). Generalizing, it is clear that it would be an excessive, even harmful pedantry to attempt to avoid using units like these, constructed from the units of space, time, and mass. (Let us call these *composite scientific units*.) Based on what we have just said, it makes

sense to fix a unit of speed and to compare other speeds by assigning them distinct numbers. Other than spatial and temporal magnitudes, nothing else is assumed to be measurable. Every composite unit admits of numerical evaluation in this same way: each unit is compared with itself, and only with itself. And each is in some way or other a composite of space, time, and mass; nothing is ever measured in terms other than space, time, or mass. The great advantage of these composite units lies in this: that despite their rigour they facilitate logical reasoning. And so we may observe how concepts of this kind develop immediately with the inauguration of each new branch of science. To the oldest of these units – speed, force, pressure, weight, momentum, kinetic energy – gradually new ones have been added: thermal units, amperage, voltage, electromotive force, electrical resistance, magnetic flux, etc. The unit of each of these physical magnitudes is composed of units of space, time, and mass. Let  $L$  indicate length,  $T$  time, and  $M$  mass. Then we have, to list some of the most important:

|                          |   |                               |
|--------------------------|---|-------------------------------|
| Speed =                  | $\text{length} \times (\text{time}^{-1}) =$         | $L \cdot T^{-1}$              |
| Acceleration =           | $\text{speed} \times (\text{time}^{-1}) =$          | $L \cdot T^{-2}$              |
| Force =                  | $\text{mass} \times \text{acceleration} =$          | $M \cdot L \cdot T^{-2}$      |
| Weight =                 | $\text{force} =$                                    | $M \cdot L \cdot T^{-2}$      |
| Pressure (hydrostatic) = | $\text{weight} \times (\text{surface area}^{-1}) =$ | $M \cdot L^{-1} \cdot T^{-2}$ |
| Kinetic energy =         | $\text{mass} \times \text{speed squared} =$         | $M \cdot L^2 \cdot T^{-2}$    |
| Work =                   | $\text{force} \times \text{displacement} =$<br>etc. | $M \cdot L^2 \cdot T^{-2}$    |

Thus every physical equation is of the form:

$$l^p t^q m^r = n \cdot (l_1^p t_1^q m_1^r).$$

where  $n$  is a scalar,  $l$  and  $l_1$  are linear extents,  $t$  and  $t_1$  durations,  $m$  and  $m_1$  masses, and where the exponents  $p$ ,  $q$ , and  $r$  are positive or negative, integers or fractions.  $L^p T^q M^r$  is called the ‘dimension’ of that magnitude which is evaluated in the equation. Again (just as in the equation given above for speed) it is evident that this equation is best expressed when cast in the form:

$$(l \cdot l_1^{-1})^p (t \cdot t_1^{-1})^q (m \cdot m_1^{-1})^r = n.$$

What we have here is still an equation between scalar terms. Furthermore we can see that any equation is absurd whose left and right sides bear different units. Such an imbalance would mean equating a length to a time, a mass to a scalar value, or something of the kind.

Neither the unit nor the dimension of any physical magnitude arises spontaneously; rather both require the establishment of an arbitrary convention, based initially on experience and on *utility*. For example, we define the *unit of force* as that which imparts a unit of acceleration to a unit of mass, or that which produces a

unit of momentum in unit time. That is how the unit of force is specified, and its dimension may be derived after a little reflection.

On this point we should recall that the speed which is induced by a constant force increases in proportion to the time during which this force is applied. Accordingly, since a given force must be equal to that by which speed increases  $n$  times in time interval  $n$ , the expression for force must remain unaltered if we multiply both speed and time by  $n$ . That is, the expression must be the quotient (speed  $\times$  (time $^{-1}$ )). We also know that when forces (which we have reason to regard as equal) act upon unequal masses in equal intervals of time, they generate speeds that are inversely proportional to mass. We arrange our expression accordingly, so that it remains unaltered if we multiply the mass by  $n$  while dividing the speed by  $n$ , that is, so that mass appears as a factor of speed. Finally then the unit of motive force takes the form ( $M \cdot V \cdot T^{-1}$ ) where we can substitute the equivalent ( $L \cdot T^{-1}$ ) for  $V$  to obtain ( $M \cdot L \cdot T^{-2}$ ).

The unit of electricity in electrostatic systems is derived in an analogous way; it is defined as that quantity of electricity which acts on a corresponding quantity with unit force at a unit of distance. Since these forces are directly proportional to the product of the electrical quantities (charge elements) and inversely proportional to the square of their distance, then by analogy we can derive the dimension of the unit of electricity (for electrostatic systems) as ( $M^{1/2} \cdot L^{3/2} \cdot T^{-1}$ ).

Now we wish to focus on a particularly instructive example, namely that of determining the intensity of a galvanic current. It turns out that this intensity is subject to several different systems of measure. Within an *electrostatic system* the unit of current is defined when the electrostatic unit of electricity (specified above) flows through the cross-sectional area of a conducting material in unit time. Since that unit of electricity had dimension ( $M^{1/2} \cdot L^{3/2} \cdot T^{-1}$ ), so current intensity as measured for an electrostatic system has dimension ( $M^{1/2} \cdot L^{3/2} \cdot T^{-2}$ ).

The unit of magnetic field strength in an electromagnetic system is defined just as we have defined a unit for electrostatics above; so the dimension of magnetic field strength is ( $M^{1/2} \cdot L^{3/2} \cdot T^{-1}$ ) because of the same relationship of force and separation distance. Further, the unit of current intensity is defined as the current flowing in a conductor of length which exercises a unit motive force at a unit separation from a magnetic particle of a unit of magnetic field strength. Since this force is proportional to the inverse square of the separation, to the length of the conductor traversed by the current, and to the strength of the magnetic field, the dimension is:

$$(Force \times (\text{Separation})^2) \times (\text{Length} \times \text{Field Strength})^{-1} = (M^{1/2} \cdot L^{1/2} \cdot T^{-1})$$

Thus the dimension is different from that of current intensity as measured in the electrostatic system. Their relation is still straightforward to the extent that two currents which have a ratio of  $1 : X$  in one scale of measure have the same ratio in the other system as well. This because the current which carries  $X$  times the quantity of electricity (in the electrostatic system) per unit time also exerts  $X$  times the effect upon a magnetic particle. The difference between the two systems of measure shows

itself in the following. Were we to change the units of length, mass, and time with which we began – for example, were we to introduce feet and minutes instead of metres and seconds – total current intensities would need to be multiplied by a constant: 20 498 for the electrostatic system, and 107.10 for the electromagnetic system.

Things are somewhat different again in the *determination of temperature*. First, we define two temperatures as equal if two bodies of those temperatures neither give nor receive heat from one another on contact. But this still does not determine what we mean when we say that a change in temperature from  $T$  to  $T^1$  may be equal to a change from  $t$  to  $t^1$ . We establish this by the convention that the expansion of a gas maintained at constant pressure is criterial for such changes. We could easily determine temperature using the expansion of mercury too, or that of platinum, as is often done for high temperatures. Measurements may differ substantially between one system and another. Then one can say (somewhat simplistically but quite comprehensibly) : the gradations of a mercury thermometer are not equal by the standard of an air thermometer. Instead those gradations become comparatively smaller as temperature increases. In other words: the volume of mercury does not increase in constant proportion to that of air. Thus two differences in temperature can be the same on one scale (or system of measure), and different from one another on another scale. Clearly it would make no sense to debate whether mercury or platinum or air expands in proportion to The Temperature – unless one has first determined in some other way how temperature ought to be measured. If one considers the expansion of gases as the normative criterion for temperature, as is usual nowadays, then one ought to say that mercury does not expand in constant proportion to this norm.

The preceding discussion is sufficient to show how we proceed in physics to make intensities tractable to quantitative measure. We can say that *in the final analysis it is always and only values of length, time, and mass that are compared one to another. The reduction of all other magnitudes to these values is mediated by a convention that involves an appropriate and pragmatic consideration of empirical relations*. Often enough we cannot establish such conventions out of hand because of a lack of empirical data; accordingly often there will be many things we are unable to measure. One example is coercive force in the magnetization of iron and steel. Currently if one were to ask how much greater is the coercive force of steel than that of iron, there would be no point to the question, nor to any number that could be given as an answer. That is inasmuch as we have no law for the attenuation of magnetism. If on the other hand we knew that magnetism attenuates according to a certain law (after other magnetic forces no longer have an effect); say it were that  $M = M_0 e^{-\alpha t}$ , where  $M_0$  denotes the magnetism present at an arbitrary initial moment of time,  $e$  the basis of natural logarithms,  $t$  time, and  $\alpha$  a constant that is substance-dependent. Then we might define coercive force as  $\alpha^{-1}$  and say that coercive force is this or that much greater for steel than for soft iron.

Likewise what we call the stability of a chemical compound is not a measurable quantity. We can say that some compounds are more stable than others (e.g. water is more stable than hydrogen peroxide), but it makes no sense to ask how much more

stable one is than another until we have decided how this stability is to be measured. Suppose that we succeeded in establishing that any body decomposes if the kinetic energy (that which is held by molecules as heat) of its molecules attains a certain value. And suppose that the stability of compounds conformed to this value in general. Then one could adopt these values as a standard of measure, and so obtain numerical expressions for the stability of chemical compounds.

While relatively simple empirical laws were a guide to the establishment of conventions governing units and dimensions in the cases we have considered, let us now consider some examples where our principles of measurement encounter greater difficulties. I have chosen the measurement of *luminous intensity* as an example. Patently there is no difficulty here so long as the lights to be compared have the same wavelength. Thus what may be meant when we say that a light *A* is *n* times the intensity of light *B*, is clear for the case where *A* and *B* have the same wavelength. The situation is different if one wishes to compare two lights of different spectral composition, or a pure spectral light with a light composed of different wavelengths. Theoretically, the simple and correct procedure is to compare the kinetic energies of the two lights, which can be determined by the heat that is generated when the light is completely absorbed (being converted exclusively to heat, and not to chemical work). Then one can say of light whose wavelength lies between  $\lambda_1$  and  $\lambda_2$  in sunlight, that its intensity is *n* times the intensity of light whose wavelength lies between  $\lambda_3$  and  $\lambda_4$  in sunlight. But frequently, especially in physiological or psychological investigation, one determines quantities of light quite differently. Two luminous intensities may be called equal if they appear equally bright to our eyes, for example. Such a convention or criterion must be *unequivocal* to be useful, in any event. That is the single requirement that must absolutely be fulfilled. Obviously a convention (like the one just introduced) fulfils this requirement only under certain conditions which must be verified in experiment before one adopts the convention. Here these conditions are easy to specify. First, two lights which appear equal in brightness to a third must also appear equal in brightness to one another. That they do so appear, may be admitted without further consideration. But moreover it is assumed in the abstract statement given above, that if this property holds for two lights *A* and *B*, then it holds as well if their kinetic energies are multiplied by a factor *n*. Now we know this is not the case. Thus if this calibration is to be useful, either brightnesses can be designated to be equal in appearance by a convention about units for a single intensity only, or if brightnesses are required to be equal across the range of intensities, then increments in intensity will be held proportional to the kinetic energy of a single wavelength only, not for other wavelengths. In the latter case the intensity of a sample light would be defined by the kinetic energy of a light of an arbitrarily chosen standard wavelength which appears to have the same brightness as the sample.

The application of a principle of measurement becomes even more difficult when it concerns processes that are more complex or by their nature unknown, such as the measurement of a *stimulus* or an *excitatory process*. If we dwell on the measurement of stimulus strength, we find that quite different stimuli can affect the same motor nerves. If we wish not to make a physical measurement of the material operations as

such by which we stimulate the nerves (which would be easy enough except that stimuli of different kinds would then be incommensurate), but if we wish rather to measure *stimulus strength*, then first we must set the requirement that two stimuli count as equal if they effect the same excitation.

If we confine our inquiry to electrical stimuli for a moment, then we must ask how the intensity of an electrical stimulus is to be measured. One answer could be: as the current intensity, determined by physical means. But then the success of our result will depend upon the time course of current intensity, primarily on its rate of change. This raises the possibility of an infinitely manifold diversity of electrical stimuli. How then should we compare, e.g. an induction pulse with the stimulus value of a simple onset or cessation of current? One might regard all those values as equal which produce a minimum stimulus when applied to the same place on the same nerve, that is, which are capable of eliciting a minimal muscle twitch. This would produce a unit for any sort of time course, and one could then find an equivalence between a given induction pulse and the cessation of a given current. But, employing these units, if one wished to go further and proceeded to call an induction pulse ten units strong (as determined by physical means) the same as the cessation of a current ten units strong, then one would find that these currents have completely different excitatory value. So they should not be considered the same stimuli. I do not need to produce explicit proof of this point here; I would only remind those of my readers who are familiar with the area, of the phenomenon of intervals for rising pulses of induced current, and of the distinctive behavior of the cessation of rising or decaying currents, and so forth. However, it may not be too bold to speculate that we will succeed in establishing a measure of ‘stimulus intensity’ just when we are able to formulate an empirical law of excitation. This would require us to identify a function of current intensity, the coefficients of its differential equation involving time and intensity, and the propagation of those intensities along nerves (as well as the current density, direction, length of neural paths, etc.), which function has the property of increasing monotonically with excitation. This function would then be useful as a measure of stimulus strength.

With more complex processes such as excitatory processes in nerves, the situation is again a little different; here it is quite conceivable that even with complete empirical knowledge of the process, it remains arbitrary how a measure of intensity should be chosen. As a matter of fact, for any process in which electromotive forces, galvanic currents, chemical conversions, temperature changes, and so forth play a role, it is pretty doubtful that a simple measure of intensity is to be found so easily. One may suppose some process to be a function of a *single* variable (with respect to excitation, as we may wish to suppose for simplicity’s sake). Even if this process lends itself to a *single* scale only, still the way in which one may choose to define its intensity will remain altogether arbitrary. If the requirements of the standard are satisfied by a system of measure in which a given intensity is dubbed  $i$ , then those requirements will also be satisfied by any arbitrary function of  $i$  which has the property of increasing monotonically with  $i$ , such as  $i^2$  or  $e^i$ . Admittedly one is not tempted to introduce just any such function as a standard; one retains that method of measuring which is physically practical to use. But once a process is found to be

concatenated of many linked events, it becomes arbitrary which part we employ in our standard. Consider for example a device that consists of a platinum wire that is brought to incandescence by a galvanic current. The radiant heat and light of this device produce some effect. Then for these concatenated processes we can establish a measure of intensity in various ways: either by current intensity, or by the amount of heat radiated from the platinum wire, or by the final effects of the radiation, assuming we can measure them. All these measures increase monotonically one with another, but not uniformly.

Naturally I do not maintain that the excitatory processes of nerves act like this, by any means. But it seems to me important to indicate this possibility, in face of the confident assurance with which we often tacitly suppose that the intensity of an excitatory process is something that can be measured free of any theoretical complication.

From the preceding it can be inferred that where the measurement of intensive magnitudes is current and has proven to be of use in the natural sciences (in the narrow sense), this measurement is concerned exclusively with the comparison of spatial or temporal magnitudes, or else with counting operations.<sup>1</sup> Thus we turn our hand to a completely new investigation when we attack the problem of measurement of *intensive magnitudes in the realm of psychology*, such as in the measurement of sensations. Straight off we see that a reduction of terms to spatial and temporal magnitudes is excluded. I would like, to simplify matters in the following section, to make a provisional (and factually indifferent) assumption: that the entire sensory system is not dynamic, meaning that a given (objective) stimulus corresponds to the same sensation throughout. Let  $R_1R_2R_3\dots$  be a series of sensations which correspond to the stimuli  $S_1S_2S_3\dots$ , and let  $S_1S_2\dots$  be a series of increasing intensities. Any attempt at magnitude estimation is here premised on the outcome of one question: whether it makes sense (and what sense it does make) to say that a change in sensation from  $R_1$  to  $R_2$  may be equal to a change in sensation from  $R_K$  to  $R_L$  – or what comes to the same thing, that sensation  $R_M$  is this or that much greater than sensation  $R_N$ . Impartial reflection leads, in my view, to the inevitable conclusion that no sense at all can be attached to this expression. The uniformity of elements that marks our conceptions of time and space is simply lacking in intensive series of sensations. If a location on the skin is subjected first to a two-pound and then a three-pound load, and subsequently to a ten-pound and then a fifteen-pound load, the latter two sensations of pressure occupy quite a different place in the whole series of sensations than the first two. Thus the one increment is something quite different from the other; at first they admit of no comparison. The claim that they may be equal makes absolutely no sense. As a matter of fact this is no different than the claim that a change in sound source and the movement of a light may be equal. In both cases the

---

<sup>1</sup>Counting operations are not restricted to the determination of similar masses only, as when a lump of gold is designated to be  $x$  times another lump of gold. We can use counting operations on various kinds of things to establish composite units, and we can find valid methods to establish a basis for measuring these magnitudes. It is easy to see that population density is one intensive magnitude (among others) that is entirely measurable; it is composed of a number (of people) and a planar area.

claim acquires a sense only when we have established some arbitrary convention about what is to be considered congruent – about what it is that we mean by ‘equal’. It may be regrettable, but the nature of this topic is such that what I affirm here cannot really be proven. I can only call into question the bare possibility that there is any evident sense to the relevant judgment of equality. One can dispute the nature of such a basic element of our mental capacity for judgment in so little as one can describe the primitive elements of sensation (such as redness or sweetness). Even if some people were to set themselves the task of determining what magnitude of length is equal to a second of time, we would only be able to inform them that for us it makes no sense to establish any congruence at all between spatial magnitudes and temporal magnitudes. And to the confident assertion that though the task is difficult, surely it could be resolved in principle, we could only respond by appeal to the immediate data of intuition, which teach us just the contrary. The situation is exactly the same when people seek to establish the congruence of two different increments in sensation.

Considering the confidence with which we expect uniformity in the basic organization of human intellect (and likely we are justified in general), I do not doubt other people will arrive at the same results by impartial examination. In my opinion the whole programme of measuring intensive magnitudes of mental life has consisted of no more than a careless and unjustified importation of the conventions about intensive magnitudes accepted in physics. It passes unnoticed that these latter are concerned solely with measurements of time and space – and this might easily have remained unnoticed, since at one time physics did not concentrate on ‘positive’ standards of measure as it does today (those standards which require the reduction of all quantities to space, time, and mass).

Later we shall elaborate the kind of delusion which can easily be engendered by a general tendency to apply objective values to sensations (e.g. values measurable by an objective standard). If one rules out this source of error, leaving aside (inasmuch as is possible) the objective process that serves as the stimulus, then one must concede that there exists no quantitative relation between the different steps of a series of intensities. This is clearest in cases where we are not bent on objectification, such as with pain. What it means to say that one pain may be exactly ten times as strong as another, is simply unfathomable.

In general we can say that intensive magnitudes are (theoretically) unmeasurable because an operation of congruence between different increments (from *a* to *b* and from *p* to *q*) has no meaning. This becomes all the more obvious when one considers that in most cases a greater intensity does not admit any fractionation, such that we can perceive the lesser intensities of which it is composed. A loud tone does not conceal within itself this or that many faint tones, in the same sense that a foot contains twelve inches or a minute contains sixty seconds. Accordingly the measurement of intensive magnitudes becomes possible in theoretical terms only when a convention is established arbitrarily, to determine what ‘equality’ should mean. In physics this is achieved by the definition of each and every intensity in terms of magnitudes of space, time, and mass. Obviously in psychology it is impossible to proceed in that way.

From these considerations it follows that the whole debate over laws of correlation between sensations and stimuli is a totally senseless debate. We can *fix a convention* such that all just-noticeable increments in sensation from a single series of intensities are to be considered equal. Having done so, we can represent a number of observed facts in that we attribute to sensation an increment proportional to the logarithm of the stimuli. But this law means nothing at all *without* that convention, as we have seen. And *with the convention* this law means nothing but the observed facts. We can just as well establish the convention that those increments in sensation should be considered equal which correspond to equal increments in stimulation. Then we can represent the same set of facts by saying that the stronger a sensation, the greater the just-noticeable difference. One schema is no more accurate than the other. The only difference between them is a question of usefulness. It is blatantly obvious how this puts our position in stark contrast to the received reading. For **Fechner** and for everyone else<sup>2</sup> who has worked in his tradition since, the problem has been one of discovering '*the correct standard of measure*' for the intensity of sensation. What does that mean? The very *task* of measuring spatial magnitudes presupposes that one has a clear idea what it should mean that spatial magnitudes may be congruent. Yet the execution of *specific* operations of measurement is tied to the existence of bodies which have the property that they can move freely in space without change of form. At least these operations depend on the property in this much: they are enormously facilitated by it, and perhaps without such bodies these operations would be impossible. Let us imagine there were only a very few such bodies. Clearly then our execution of operations of measurement in space would depend on our ability to recognize and use those few bodies. Whoever had discovered the first of these bodies would be justified in saying: until now we could not have measured spatial extent, for want of an appropriate standard; now we have found the standard which permits us to do so. This is the sense in which **Fechner** thought he had revealed the just-noticeable difference as the 'true' standard of measure for the intensity of sensation. In contrast, discussion has shown that we are really dealing with an arbitrary convention about relations of magnitudes – which can be expedient or inexpedient, but which cannot be right or wrong.

Since this fundamental concept is crucial to the question at hand, it would seem to warrant a little more attention. I am apprised of the objection that the very notion of

---

<sup>2</sup>**Zeller** 's (1881) recent article is also directed against the performance of actual measurements. It leaves unmentioned what I must take to be the core of the whole matter. When **Zeller** (1881, p. 9) says 'that psychological processes are not measurable in the same sense as magnitudes of space and time, because they have no invariant standard of measure...', then it is evident he considers a situation similar to that which would hold for space if there were no rigid bodies. To this one might justifiably counter that in the domain of sensations we could produce at least close or approximate values that might serve as such 'invariant standards of measure'. Nor am I ready to concede that psychological processes may be unmeasurable because appearances in consciousness can only ever be compared with appearances in consciousness. No form of measurement does anything else, but to establish the relations of like with like. So we measure space only by space, and time only by time. Compare **Wundt** 's (1883, pp. 253–254) comments on the subject, which I endorse wholeheartedly.

an intensive magnitude presupposes that the magnitude estimation of increments is at least conceivable, even if it is not always practicable. Although it is not difficult to demonstrate that this opinion is false, it would be useless to argue the point, since it comes down to a matter of semantics. Even if one wished to define the notion of an intensive magnitude in that way, the notion would still be inapplicable to sensations (it would be inapplicable to anything, as we shall indicate presently). We will need to be clear what we actually mean when we speak of *more or less, stronger or weaker* with regard to sensations, as we will need say clearly how we come to label some sorts of changes as *increases* or *decreases* in contrast to other changes that appear to us as qualitative. As soon as one poses the question in this way one sees that a criterion of equality for distinct increments of magnitude is just not required. In fact no such criterion is available. Instead we see all that matters here is the existence of a continuous variable which has zero at one end of its range. Then clearly two tasks are left to psychological investigation: first, to discover how we achieve representation of a continuous series of changes – especially how we manage to differentiate changes in one direction (or a certain sense) from changes in another direction; and second, to determine how it is we can interpret the terminus of such a series as the origin of a scale, as an absence of sensation. These questions of psychology are moot for our purposes; it suffices to say that these elements of representation (an origin and a continuous change in one direction) are something given empirically. They constitute necessary and sufficient conditions for the emergence of a representation of intensity. And (by contrast) nowhere is there implicated any composition of equal parts. That psychological magnitudes do have such characteristics admits of no further proof, as has been intimated. Yet it can still be shown that fundamentally the same thing holds in the natural world for what we call ‘intensities’. Let us consider a string fixed at both ends, which can be made to swing in harmonic motion; it is instantly clear what we mean if we say that its oscillations are stronger or weaker.<sup>3</sup> This is an example of continuous variation in a state of affairs, beginning from rest (that is, from an origin). We can speak unequivocally of intensity here. Still this intensity is not yet determinable in quantitative terms. It becomes determinate only if we arbitrarily identify this intensity with some function of the magnitudes of mass, space, and time which are observable in the process. This identification is indeed arbitrary, as will become clear even to those to whom it is not immediately intuitive, once it is remembered how very differently this problem can be framed. In our present example one can assign the range of motion of a given point as a measure though one might just as well assign the kinetic energy that the string has at the moment all of its parts pass by its initial position at equilibrium. The latter measure, though apparently less simple, has been the more useful. Intensities measured in this way do not increase with the range of motion, but with the square of the range. Hence it is clear that while we can well speak of an intensity of oscillation, it is still completely undetermined whether one oscillation should be considered

---

<sup>3</sup>It is assumed these oscillations occur in a plane, without the formation of nodes, so that the process can be described as a function of a single variable.

twice or four times another – in other words, an assignment of numbers is not implied by the concept of intensity. If we ascribe a value of four units of intensity to an oscillation of doubled amplitude, then one sees that there is no intelligible sense in which the larger intensity can be represented as four simple and equal components. A similar situation is observed anywhere we do not yet possess a definite physical convention for what should be understood by ‘intensity’, or where this convention has not become part of our nature, as it were. Let me cite the so-called ‘current strength’ of a river as an example. One can, at will, understand by this either the average speed of water particles or this quantity multiplied by the cross-sectional area of the water (that gives the flow per unit interval).

Doubtless though the significance of any measurement of sensation is greatly depreciated by calling attention to the *arbitrarily* conventional meaning of equivalence of sensation, nevertheless study of the constraints on such arbitrary conventions of measurement is not without interest – namely to determine the extent to which ‘just noticeable differences’ satisfy these requirements.

We noted earlier that the only requirement that absolutely must be fulfilled is the requirement for an unequivocal criterion. If one wishes to measure nothing more than the strengths of sensations that accompany a merely intensive ordering of stimuli, then just-noticeable differences will satisfy this requirement in general, at least to a good approximation. Yet we should not forget that the constitution of our sense organs renders the requirements for a suitable standard very complicated. A first complication is introduced for touch and for haptics in general by the extended nature of some sensory surfaces. This affords a new possibility: that a stimulus can act on different parts of a sensory apparatus. We are unquestionably capable of a definite (if not very precise) comparison of sensations from separate locations. We are in a position to say: this spot of light, which falls on the periphery of my retina, is as bright as (is brighter than, or darker than) that one which falls on the centre of my retina. Similarly for the pressure sense of the skin. Let us leave aside consideration of what the real significance, or what the origin of our ability to make these comparisons may be; we need only say beyond a doubt that we possess these abilities. Then one requirement for a system of measure may be stipulated: that two sensations associated with different sites on a sensory surface and considered by us as equal must be represented as the same in our system of measure. In other words it would be desirable to choose a system of measure so that it would be equally valid for all locations on the given sensory apparatus. Should the question arise if just-noticeable differences satisfy this requirement, then the answer is no. We know that the centre and the periphery of the retina act in such a way that a given light will not appear much less bright when presented indirectly than when presented centrally,<sup>4</sup> despite the fact that central difference thresholds for brightness are substantially smaller, and the absolute thresholds are virtually identical between the centre and periphery.

---

<sup>4</sup>I rely here on **Aubert**’s report (Aubert & Foerster, 1857). The comparison of a bright spot in central vision with one in peripheral vision appears to me a rather shaky business, though I do believe **Aubert**’s report to be correct.

Accordingly two differences in brightness would have to be considered equal though the one perceived centrally would be composed of many more just-noticeable differences than the one perceived peripherally.

We encounter quite a similar phenomenon if, instead of different locations along a sense organ, we introduce different qualities of stimuli as another complication in our task. Here again we are capable of applying rather imperfect criteria for the evaluation e.g. of the brightnesses of different lights. The difference threshold for changes in brightness varies greatly with hue, according to studies by **Lamanski** and **Dobrowolski** (e.g. Dobrowolsky, 1872). Consider a red  $R_1$  and a blue  $B_1$  which an observer deems to be equally bright. If we multiply each light by some amount, so that they are increased by an equal number of just-noticeable differences, one might expect that the subsequent values would again appear equally bright. But this is not at all the case. As two hues of apparently equal brightness increase uniformly in intensity, the less refrangible light will appear to be more salient, according to **Helmholtz**. By contrast, the difference threshold for refrangible lights is greater than that for weakly refrangible lights.

These facts (whose list may be expanded) reveal that the measurement of sensations by **Fechner**'s standard does not show any of the generativity that one would be justified in expecting, if that standard of measure had any objective validity. Instead these findings provide a vivid documentation of the arbitrariness of introducing **Fechner**'s standard of measure.

The same objection is raised even more forcibly in the measurement of qualities. We can report quite exactly whether a loud tone and a faint tone are the same in pitch or not. Thus a difference of a fifth is the same for a loud tone and a faint tone. Yet this difference subsumes a vastly greater number of just-noticeable differences for tones of medium loudness than for faint tones, in which the recognition of small intervals is more difficult. Likewise difference thresholds for pitch are subject to great variation between individuals. Practice will heighten sensitivity substantially for any individual, but this does not allow one to say that a fifth will appear to be as great a difference as an octave did previously.

Why did people seem so compelled, following **Fechner**, to overrate the just-noticeable difference? The answer becomes clear if we imagine what might have been the achievements (under the best of circumstances) of a numerical reckoning of psychological magnitudes. We may imagine that we could have had an arbitrary convention that fixed intensities in the way that has been mentioned; i.e. it would only have had to satisfy the requirement that the vanishing (the indistinguishability) of a sensation, as well as its increase, decrease, and constancy (that is, its approximation to an asymptotic value) were expressed in a definite way. It is conceivable that these definitions would have allowed us to establish laws for psychological phenomena – laws of simple and closed form. Here the results would still have been bound by an arbitrary standard; hence they would be uninterpretable without help of the definition. But despite this the results could still have been useful, insofar as they might have produced suitable expressions for the vanishing, the increase and decrease, and the constancy of sensation (namely for the type identity of a state).

Let us illustrate this by an example. A psychological state whose values can be arranged in a monotonic order  $f$  may arise by the joint action of a number of other states  $a, b, c \dots$ . Let us consider  $f$  as well as  $a, b, c \dots$  to be determined quantitatively along an arbitrary scale. Then it is possible a law could be established describing this process, perhaps of the form  $f = \varphi(a b c \dots)$ , where  $\varphi$  is some function. Now if  $f$  also arises by the joint action of  $\alpha \beta \gamma$ , and these admit measurement along a certain conventional scale, then it is possible to write  $f = \psi(\alpha \beta \gamma \dots)$ , where  $\psi$  is another functional relation.

Neither of these propositions makes sense in the absence of a definition of intensity. Yet taking the two propositions together, under certain conditions it follows that:

$$\varphi(a b c \dots) = \psi(\alpha \beta \gamma \dots).$$

This has a definite and specifiable meaning, namely that this kind of psychological state is generated by  $a \alpha \beta \gamma \dots$  in just the same way as it is generated by  $a_1 b_1 c_1 \dots$ <sup>5</sup>

Whether such functional relations can in fact be established will depend on the existence of a suitable method for the measurement of all these quantities. Should the attempt to find such a scale be successful, then while this still amounts to measurement by convention, still it would be a valuable convention indeed. A system of psychological laws erected on this basis might teach us a great many important things, even though any report of magnitudes would derive its final interpretation from consideration of some defining equation.

**Fechner** was of the opinion (if I am not mistaken) that consequently something like ‘the external validity of the scale’ would have to be acknowledged. Now one can see plainly that just-noticeable differences were selected precisely because a form or kind of psychological process may be simply expressed in those terms. This is the way in which the perception of difference is tied to changes in sensation, whose simple statement in a formula is used as a convention about a standard of measure for sensation. Thus we frame our definition in terms of a causal mechanism within psychology if we say: we will consider those increments in sensation as equal, which induce a perception of difference.

If one proceeds (as **Fechner** did) from the notion that there is an objectively valid scale of measure for sensation, it will seem reasonable to assume this scale will become evident just insofar as simple laws of psychological events can be formulated in those terms. That is, it will seem reasonable to assume that those increments of sensation are equal which cause the same perceptions of difference (namely the just noticeable ones). On the other hand, once we recognize that there is no question

---

<sup>5</sup>I hardly need point out that one of these two propositions  $f = \varphi(a b c \dots)$  could itself be the defining equation. And under certain circumstances, it can remain arbitrary which of several propositions is regarded as the defining equation and which as a law. Thus it is arbitrary whether one regards the electrostatic, the electrodynamic, or the electromagnetic determination of current intensity as definitive (by analytic judgment). In any case the remaining two will then be subjects of synthetic judgment.

of an ‘objective validity’, we are forced to question if great importance should still be attached to the just-noticeable difference in establishing laws of psychology. Naturally these laws will still become evident in *some way or other*, as in study of the basic processes on which they are modelled. This business has been settled in our previous inquiry, where we asked how well just-noticeable differences satisfy the requirements of an arbitrary convention for a scale of measure. We saw they were not satisfactory, even as we attempted a first generalization. We were unable to declare just-noticeable increases in brightness at the region of sharpest vision to be equal to just-noticeable increments in the periphery of the visual field. We find this over and over again, and as a result we can say that so far the just-noticeable difference has in no way matched the hopes that were placed upon it.

It remains conceivable that future investigation into the laws of psychology will make a numerical account of intensive magnitudes in psychological activity attractive in some way. Future investigation may even characterize certain conventions about the meaning of such an account as almost self-evidently useful, as is now the case for intensive magnitudes in physics. Still this does not alter the logical situation one bit. Let us say there had been a dispute before we came to understand the relationship between kinetic energies and the temperatures of gases: a dispute whether air or mercury expands in proportion to temperature. We could now say only that the dispute had been futile, since at that time no one knew what should be understood by ‘temperature’ at all.

The moral for further research is that it is not our task to ‘discover’ the relation between intensities of sensations and stimuli, or between sensations and excitatory processes. Instead it is our task to derive psychological laws – laws whose formulation incorporates arbitrary convention (where such conventions serve as an aid to discourse) about those things that should be considered equal in magnitude.

The same holds for the measurement of qualitative differences as holds for the measurement of intensities. *That two differences of some kind are equal in size or unequal in size* is an entirely arbitrary assertion: it has no fixed or intelligible meaning at all. Here again only an arbitrary convention as to what should be considered equally great will serve to give content to such propositions. Here again there is more or less only as soon as sensations are ranged in a qualitative series. The difference between yellow and red is greater than that between orange and red, and the difference between musical notes *d* and *c* is smaller than that between *e* and *c*. Then is the difference between red and yellow larger or smaller than that between *c* and *d*? No one would judge that any reply is admissible to such a question.

Nevertheless what we encounter (especially but not exclusively) in the area of qualitative scales are often rigid notions about the relative magnitudes of any two differences whatsoever. So one may have occasion to hear the statement that the difference between *c* and *C* is equal to that between *d* and *D*, and so forth.

But once a thoroughgoing analysis has taught us that there is no question of really measuring differences in sensation, be these either qualitative or intensive differences, still we will not be able to consider our work entirely finished. Rather we must examine what we mean by our use of language in everyday life, that is, what is meant

by comparisons of magnitude when they are laid down in ordinary language. It will be necessary to see how these uses originated, and what legitimizes them. I believe several different cases can be distinguished here.

Firstly we attribute to sensation (without due consideration) that which holds primarily for the objective relations that we estimate by means of those sensations. This is certainly the case when (as may happen) we speak of ‘an extensive magnitude of sensations’, but mean by that the magnitude which the physical object seen or the physical object felt appears to have. This ‘extensive magnitude of sensations’ is not a happy turn of phrase by any means. Still should one persist in using it, one may say that (extensive) magnitudes of sensation generally increase in proportion to the objective extensions of physical things. But this is to say nothing more than that we are usually capable of making accurate estimates of extensive magnitudes in general. That fact becomes significant only in juxtaposition to a wholly false assumption: that the extensive magnitude of sensation (i.e. the apparent magnitude of physical objects) can be accounted for by numbers of just-noticeable differences. **Weber** ’s (1852) famous experiment – in which two fixed compass points appear to approach one another as we move them from the finger up over the hand and then along the arm – merely demonstrates that estimates of objective relations no longer parallel those objective relations *once we approach the limits of our capacity to discriminate*. This is a phenomenon to which we will return shortly, for discussion at greater length. As long as this sort of thing is excluded, no *systematic* error of any real significance occurs in observers’ estimates. Nothing could be further mistaken than the supposition that an equal number of just-noticeable differences in spatial extent appear equal to one another, e.g. in the central and in the peripheral visual field. It requires great attentiveness to convince oneself of the petty illusions of estimation by eye that occur in indirect vision. If on the other hand, the just-noticeable distance were our standard of measure according to which we estimated magnitudes at any place, then every object that passed from direct to indirect view would appear to dwindle to little more than nothing, because of the extreme dullness of spatial sense in the periphery of our vision by comparison to central vision. Similarly for touch: the distance from the nape of the neck to the small of the back would appear only as large to us as, say, a stretch of a few centimetres across the hand. We know that with experience we estimate extensive magnitudes accurately in general, quite independently of just-noticeable differences. This does not exclude the possibility that these latter phenomena could be the source of specific illusions under special circumstances (see below).

The same thing holds for sensations of pressure. If we speak of the intensity of a sensation of pressure, or should we want to evaluate such a sensation, then either we attach no clear meaning to the intensity, or we give it the meaning of an assessment of the weight causing the pressure. This again is something objective: something which admits of an objective measure. **Hering** (1876) has shown explicitly that we estimate weights accurately in general, and not according to their logarithm. Here where there is such an obvious connection to physical measurement of the stimulus, people will find it hard to interpret anything else by ‘intensity of sensation’ than just the intensity of the stimulus, judged by means of sensation. Consequently one can derive a law of proportionality here too, only one whose significance does not extend beyond specific facts.

Were we equally accustomed to estimate intensities of sounds, then presumably we would also call that sound sensation ‘double’ which arose from double the intensity of sound source.

Observers in **Delboeuf**’s (1873) experiments seem to use visual sensations in a different technique of estimation. I find this of particular interest since his results again show how estimates of large differences approximate the values of objective properties. **Delboeuf**’s experiment involved the following: three distinct patches of brightness (an inner circle, a middle annulus, and an outer annulus) are presented to the eye for comparison. Each of these patches is a white sector whose arc length can be adjusted; these white sectors rotate in front of an almost completely black background. In this way each of the three patches can be adjusted continuously in brightness. If the observer’s task is to adjust the brightness of the middle ring so that its value is midway between that of the outer ring and the inner circle, this problem is solved in a consistent manner. The (adjusted) objective brightness is set to be very close to the geometric mean of the other two brightnesses.<sup>6</sup> This experimental result has an immediate and evident connection to **Fechner**’s result (reported some time ago, 1859) that the objective brightness of different magnitude classes of stars corresponds to a geometric progression in brightness. It does not seem difficult to explain how this should happen. As we attend to the appearance of various bodies, we want to apprehend a property characteristic to them, and commit this to memory. What is important to us is not the incident light intensity, but rather the surface properties of bodies, which modulate but are not the sole influence on light intensity. The brightness under which we see an object is not constant but is quite variable, and this makes it very difficult to identify the key property in the situation. If we study closely the manner in which apparent brightness varies, we find a definite principle at work here. For terrestrial objects which are not self-luminous, brightness (that is, the amount of light reflected by a unit surface area) depends chiefly on the brightness of incident light (such as the brightness of daylight), which by its nature varies immensely. And certainly the brightness of an object in this sense is simply proportional to incident brightness. Accordingly the brightnesses of two distinct and unchanging bodies may be  $a$  and  $b$  at one time,  $n \cdot a$  and  $n \cdot b$  at another time, and  $m \cdot a$  and  $m \cdot b$  on a third occasion, where  $n$  and  $m$  represent numeric values. The brightness we see an object to have depends further on pupil diameter; the larger the pupil, the larger the cone of light that leads from each point of the visible object into the eye. This has the same effect as do variations in the brightness of incident light, so the brightness of the retinal image might be  $\alpha \cdot a$  and  $\alpha \cdot b$  in one instance, and  $\gamma \cdot a$  and  $\gamma \cdot b$  in another, where  $\alpha$  and  $\gamma$  stand for pupil diameters. Fatigue of the eye has a similar subsequent effect; we know that the approximate effect of a state of retinal fatigue is as if all objective intensities of light were reduced in fixed proportion. And finally, this is the same effect that variable absorption of light by the atmosphere has upon starlight. In a similar way the brightnesses of two stars are  $a$  and  $b$  at one time, and  $p \cdot a$  and  $p \cdot b$  at another. So if we perceive that one and the same objective

<sup>6</sup>Here I disregard any consideration of the small correction introduced for the natural light of the retina. [Trans.: for spontaneous phosphenes].

proportion corresponds now to brightnesses  $a$  and  $b$ , then to  $n \cdot a$  and  $n \cdot b$ , then to  $\gamma \cdot a$  and  $\gamma \cdot b$ , and then to  $p \cdot a$  and  $p \cdot b$ , it becomes understandable how we are used to judging light intensities by ratios, thus not in an arithmetic series but in a geometric series. If (in **Delboeuf**'s experiment described above) the brightness of the inner circle was 4 and that of the outer ring 100, then we would adjust the midpoint to 20; we would set the difference between 4 and 20 to be equal to the difference between 20 and 100. This is because of the way the brightnesses 20 and 100 would look under various conditions. Either under a fifth of the intensity of daylight, or with a pupil five times smaller, or in a greater state of retinal fatigue, this pair would look exactly the same as the other pair (of 4 and 20) looked originally. By this account **Delboeuf**'s results may easily be ascribed to the particular way that we make use of visual sensations to form judgments about objective proportions.

Proportions of a quite specific nature must be considered in regard to the estimation of larger qualitative differences across tonal series. Here it seems intuitive and natural for judgment to consider *equal intervals* (equal ratios of frequencies) as equal changes in pitch. Although it is indisputable that one cannot simply *assert* an equivalence between a change from  $C$  to  $G$  and a change from  $c$  to  $g$ , yet such an equivalence does seem to have some title as a convention. One thing is sure: this has nothing to do with just-noticeable differences, inasmuch as the same interval represents very different numbers of just-noticeable differences for tones of different intensities, and for tones along different parts of the scale. What we mean by an equivalence of distinct intervals across any part of the scale, is something that may be established without much difficulty. Clearly this is all about relations of consonance and dissonance, which are independent of absolute levels of pitch and which repeat themselves across the entire scale, depending only upon numerical ratios of frequencies. This fact, which is elaborated in **Helmholtz**'s theory of timbre, has the consequence that any melody can be reproduced faithfully in a different register, so to speak, as a specific multiple of all its frequencies. So as is well known, most people do not even notice if a particular piece is performed at one time in one key and at another time in another (as when a singer transposes a song), while the same people would sense any deviation from the correct intervals as a mistake. Thus the significance of these intervals (ratios of frequencies) does not consist of a simple equivalence of qualitative differences in sensation for this case either, but consists of an equivalence relation based upon harmony.

Now there are many cases (and there are perhaps just as many again as those just discussed) where we perform a certain kind of magnitude estimation by a quite different means than by reference to relations of objects, or similar principles. One could perhaps describe these as cases of estimation by *psychological effect*. We will see presently that in such cases just-noticeable differences have a certain role to play – though we will discount the possibility that they produce an illusion, by generalization.

What this means is best illustrated by some examples. Although we have no means of comparing the strengths of visual and auditory sensations, people will say instinctively that a blinding light is more intense than a sound just above the detection threshold for hearing. We take the difference of two sensations that we discern only

with great concentration to be a smaller difference than one which is so large it cannot pass unnoticed. Differences as well as simple intensities are judged very often (in ordinary unbiased thought) according to the role they play in the development of representations. Does it follow, then, that numerical comparison can be made with a standard of measure like this? Absolutely not. This is obvious as soon as one realizes that even in the context of such a mental *effect* (that renders the comparison of heterogeneous things possible), a gradation of intensity is involved, but there is no possibility of making a comparison in numbers between different steps.

Inasmuch as we estimate by effects, for a moment we may be inclined to call any two sensations or any two differences equal which stand just at the threshold of perceptibility. In a similar way we might call two completely heterogeneous sensations equally strong at just the points where they begin to be painful. So we can speak of the ease (or the difficulty) of tasks in very different contexts and we can compare these one to another, since again we have points of comparison: namely the point of barely noticeable expenditure of energy, the point at which the tasks cause us significant discomfort, and the point beyond which they cannot be carried out at all. Under this or other similar schemes, we could then speak of *strong* and *weak*, *much* and *little*, *large* and *small*, in almost any area of psychological activity without ever expecting or demanding a numerical reckoning. People will concede that the beauty of **Raphael's** Sixtina is greater than that of a genre painting by **Knaus**, but it should not occur to anyone that it is possible (I do not mean practically, but *conceptually*) to determine a numeric expression for this relationship. In the same sense that we speak of the similarity or difference of two sensations, we can also speak of the similarity or difference of two landscapes, two architectural styles, or two languages. We can speak of *much* or *little* in these contexts, too. No one would find fault, or find it incomprehensible were we to say that the similarity between Italian and Latin is greater than that between German and Sanskrit. On the other hand, one should not consider it sensible to stipulate that the former is 225 times greater than the latter.<sup>7</sup> Notwithstanding, should we still want to measure similarities, our attention is drawn more clearly to the arbitrary nature of convention than it is when we study sensation. One might, for example, define the similarity of

---

<sup>7</sup> **Wundt** (1883, p. 254) has objected to an apparently similar argument by **Zeller** (1881) as follows: [Naturally psychology must forever renounce the hope of determining by measure or in numbers the boringness of a conversation as greater or less. Psychology should be all the readier to do so, because such a determination would be as useless as it is hopeless. Why didn't **Zeller** also reproach physicists for being in no position to count the number of water droplets in the ocean, or to predict the course that a storm-tossed boat will travel on the high seas?]. This objection finds its target in **Zeller's** perspective on the subject, but does not impinge at all on the explanation given above. Here we are not concerned whether or not a measurement can be made practically in a specific case; rather we are concerned that such reports as may be made offer anything meaningful, quite apart from their material truth or falsity. The proposition that the amount of liquid water on the face of the earth amounts to exactly 100 kilograms at a given time is simply false. It is false only because it has a fully determinate sense, like the proposition that that boat will move at a uniform speed along a cubic curve. Yet the proposition that one conversation is ten times as boring as another is neither true nor false, but is simply a string of words to which no sense may be attached.

vocabularies according to the absolute number of common roots, or according to their proportion to a total. This numerical datum would then indicate something definite.

If we do not strike a convention then numerical data have no meaning, though one may still talk of much or little. In any particular case what is responsible for our attribution can usually be specified without too much effort. For, from the aggregate psychological effect of the causal conditions to be compared, one will always be able to show that a certain state can appear now as intense and then as weaker, now frequently and then seldomly, now as long-lasting and then as quickly vanishing. Then we can readily understand how it is possible that we come to apply ‘more’ or ‘less’ even to collective terms or rubrics for many different kinds of effects like these (as in the examples above: the similarity of languages, or two architectural styles, or the difficulty of a language, etc.).

In conclusion, the essential results of our present inquiry can be summarized in a few brief sentences.

Intensive magnitudes are not measurable in themselves, because the establishment of an equivalence between different steps in a scale of intensity does not make any sense without further clarification. Where intensive magnitudes are determined in a domain of the so-called exact sciences, this is really only a matter of counting, and of the measurement of temporal and spatial magnitudes. All measurement of intensity is derived from these operations by specific conventions. Likewise, we can only speak of the measurement of sensations once we have established an arbitrary convention that determines what we shall consider as equal. In sum, the debate whether sensation varies with the logarithm of stimulus intensity, or in direct proportion to stimulus intensity, is not a difference over matters of fact. Instead it is an empty dispute over words that is rooted in misunderstanding.

J. von Kries

*Freiburg (Breisgau)*

## References

- Aubert, H., and R. Foerster. 1857. Untersuchungen über den Raumsinn der Retina. *Archiv für Ophthalmologie (Berlin)* 3 (2): 1–37.
- Delboeuf, J.R.L. 1873. Étude psychophysique: Recherches théoriques et expérimentales sur la mesure des sensations et spécialement des sensations de lumière et de fatigue. *Mémoires couronnés et autres mémoires* (de l’Académie Royale des Sciences, des Lettres, et des Beaux-arts de Belgique), 34: 50–101.
- Dobrowolsky, W. 1872. Ueber die Empfindlichkeit des Auges gegen die Lichtintensität verschiedener Spectralfarben. *Archiv für Ophthalmologie (Berlin)* 18 (1): 74–92.
- Fechner, G.T. 1859. Ueber ein psychophysisches Grundgesetz und dessen Beziehung zur Schätzung der Sterngrößen. *Berichte über die Verhandlungen der königlich – sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physische Classe* 4: 457–532.

- Hering, E. 1876. Zur Lehre von der Beziehung zwischen Leib und Seele I: Über Fechner's psychophysisches Gesetz. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe. Abt. 3*, 72, 310–348.
- Weber, E.H. 1834. *De pulsu, resorptione, auditu et tactu*. Leipzig: C.F. Koehler.
- Wundt, W. 1883. Ueber die Messung psychischer Vorgänge. *Philosophische Studien* 1: 251–260.
- Zeller, H. 1881. Ueber die Messung psychischer Vorgänge. *Abhandlungen der Akademie der Wissenschaften zu Berlin, philosophisch-historische Klasse* 3: 1–16.

# Bibliography

- Anscombe, G.E.M. 1981. *Metaphysics and the philosophy of mind. Collected papers, vol. 2.* Oxford: Basil Blackwell.
- Aubert, H., and R. Foerster. 1857. Untersuchungen über den Raumsinn der Retina. *Archiv für Ophthalmologie (Berlin)* 3 (2): 1–37.
- Bayes, T. 1763. An Essay towards solving a problem in the Doctrine of Chances. By the late Mr. Bayes F.R.S., communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S. *Philosophical Transactions of the Royal Society of London* 53: 370–418.
- . 1764. A demonstration of the second rule in the Essay towards the solution of a problem in the demonstration of chances. *Philosophical Transactions of the Royal Society* 54: 296–325.
- Béguelin, N. de 1767. Sur l'usage du principe de la raison suffisante dans le calcul des probabilités. *Histoire de l'Académie Royale des Sciences et des Belles-Lettres de Berlin* 23: 382–412.
- Bernoulli, J. 1713. *Ars conjectandi* (opus posthumum). Basileæ: Thurnisiorum, Fratrum.
- Bernoulli, D. 1738. Specimen theoriae novae de mensura sortis. *Commentarii Academiae Scientiarum Imperialis Petropolitanae (In Classe Mathematica)* 5: 175–192. (contributions submitted in 1730 and 1731).
- Bessel, F.W. 1876. *Gesammelte Abhandlungen, Band 1.: Theorie der Instrumente, Stellarastronomie, Mathematik*. Leipzig: W. Engelmann.
- Black, W. 1789. *An arithmetical and medical analysis of the diseases and mortality of the human species*. London: J. Johnson.
- Boltzmann, L. 1868. Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe* 58: 517–560.
- . 1871. Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe* 63 (2): 397–418.
- . 1877. Bemerkungen über einige Probleme der mechanischen Wärme-Theorie. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe* 75 (2): 62–100.
- . 1878. Weitere Bemerkungen über einige Probleme der mechanischen Wärmetheorie. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 78 (2): 7–46.
- . 1880. Zur Theorie der Gas-Reibung I. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 81 (2): 117–158.
- . 1881a. Zur Theorie der Gas-Reibung II. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 84 (2): 40–135.

- \_\_\_\_\_. 1881b. Zur Theorie der Gas-Reibung III. *Wiener Sitzungsberichte der kaiserlichen Akademie der Wissenschaften* 84 (2): 1230–1263.
- \_\_\_\_\_. 1909. *Wissenschaftliche Abhandlungen, 2. Band*. Leipzig: Barth (F. Hasenöhrl). Reissued 1969 – New York: Chelsea.
- Boring, E.G. 1920. The logic of the normal law of error in mental measurement. *The American Journal of Psychology* 31 (1): 1–33.
- Buldt, B. 2016. Johannes von Kries: A bio-biography. *Journal for General Philosophy of Science* 47 (1): 1–19.
- Cohen, L.J. 1989. *The philosophy of induction and probability*. Oxford at the Clarendon Press.
- Cournot, M.A.A. 1843. *Exposition de la théorie des chances et des probabilités*. Paris: L. Hachette.
- Czuber, E. 1884. *Geometrische Wahrscheinlichkeiten und Mittelwerte*. Leipzig: B.G. Teubner.
- \_\_\_\_\_. 1902. *Probabilités et moyennes géométriques*. Paris: A. Hermann.
- \_\_\_\_\_. 1914. *Wahrscheinlichkeitsrechnung und ihre Anwendung auf Fehlerausgleichung, Statistik, und Lebensversicherung*. Leipzig & Berlin: B.G. Teubner.
- d'Alembert, J. le R. 1768. Doutes et questions sur le calcul des probabilités. *Mélanges de littérature, d'histoire et de philosophie. Tome 5*. Amsterdam, aux dépens de la compagnie.
- Condorcet, N. de 1785. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. À Paris, de l'imprimerie royale.
- Delboeuf, J.R.L. 1873. Étude psychophysique: Recherches théoriques et expérimentales sur la mesure des sensations et spécialement des sensations de lumière et de fatigue. *Mémoires couronnés et autres mémoires* (de l'Académie Royale des Sciences, des Lettres, et des Beaux-arts de Belgique), 34: 50–101.
- Diderot, D., and J. le R. d'Alembert. 1751–1766. *Encyclopédie, ou dictionnaire raisonné des sciences, des arts et des métiers*. Paris: Chez Briasson.
- Dobrowsky, W. 1872. Ueber die Empfindlichkeit des Auges gegen die Lichtintensität verschiedener Spectralfarben. *Archiv für Ophthalmologie (Berlin)* 18 (1): 74–92.
- Drobisch, M.W. 1880. Ueber die nach der Wahrscheinlichkeits-Rechnung zu erwartende Dauer der Ehen. *Berichte über die Verhandlungen der königlich – sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematische – physische Klasse* 32: 1–21.
- Fechner, G.T. 1859. Ueber ein psychophysisches Grundgesetz und dessen Beziehung zur Schätzung der Sterngrößen. *Berichte über die Verhandlungen der königlich – sächsischen Gesellschaft der Wissenschaften zu Leipzig, mathematisch-physische Classe* 4: 457–532.
- Fick, A. 1883. *Philosophischer Versuch über die Wahrscheinlichkeiten*. Würzburg: Stahel'schen Univers.- Buch.- & Kunstdrucklung.
- \_\_\_\_\_. 1885. *Die medicinische Physik*. 3rd. ed. Braunschweig: Friedrich Vieweg und Sohn.
- Fioretti, G. 1998. John Maynard Keynes and Johannes von Kries. *History of Economic Ideas* 6 (3): 52–80.
- \_\_\_\_\_. 2003. No faith, no conversion: The evolution of Keynes's ideas on uncertainty under the influence of Johannes von Kries. Chapter 10 of: J. Runde & S. Mizuhara. *The philosophy of Keynes's economics: Probability, uncertainty and convention*. London: Routledge.
- Fismer, F.H. 1873. *Die Resultate der Kaltwasserbehandlung bei der acuten croupösen Pneumonie im Basel Spitäle von Mitte 1867 bis Mitte 1871*. (Inaugural-Dissertation der Hohen Medicinischen Facultät der Universität Basel). Leipzig: J.B. Hirschfeld. (Separat Abdruck, *Deutsches Archiv für klinische Medizin*. 11).
- Fries, J.F. 1842. *Versuch einer Kritik der Principien der Wahrscheinlichkeitsrechnung*. Braunschweig: Friedrich Vieweg und Sohn.
- Galavotti, M.C. 2017. The interpretation of probability: Still an open issue? *Philosophies* 2 (20): 13 pp.
- Gauss, C.F. 1821. *Theoria combinationis observationum erroribus minimus obnoxiae, pars prior*. Göttingen: Apud Henricum Dieterich.
- \_\_\_\_\_. 1880. *Werke IV*. Göttingen: Königlichen Gesellschaft der Wissenschaften.
- Gavarret, J. 1840. *Principes généraux de statistique médicale, ou Développement des règles qui doivent présider à son emploi*. Paris: Bechet jeune & Labé.

- Goodman, N. 1955. *Fact, fiction, & forecast*. Cambridge, MA: Harvard University Press.
- Graunt, J. 1662. *Natural and political observations made upon the bills of Mortality*. London: Thomas Roycroft.
- Hacking, I. 1965. *Logic of statistical inference*. Cambridge: Cambridge University Press.
- . 1971. Equipossibility theories of probability. *The British Journal for the Philosophy of Science* 22 (4): 339–355.
- Hagedorn, F. von 1764. *Sämmliche poetische Werke. Erster Theil*. Hamburg: Johann Carl Bohn.
- Hagen, G. 1867. *Grundzüge der Wahrscheinlichkeitsrechnung*. 2nd ed. Berlin: Ernst & Korn.
- Halley, E. 1693. An estimate of the degrees of the mortality of mankind, drawn from curious tables of the births and funerals at the City of Breslaw; with an attempt to ascertain the Price of Annuities upon Lives. *Philosophical Transactions of the Royal Society of London* 17 (196): 596–610.
- Heidelberger, M. 2010. From Mill via von Kries to Max Weber: Causality, explanation, and understanding. Chapter 13 in: U. Feest, Ed. *Historical perspectives on Erklären und Verstehen*. (New Studies in the History of Science and Technology: Archimedes vol. 21) New York: Springer, 241–265.
- Hering, E. 1876. Zur Lehre von der Beziehung zwischen Leib und Seele I: Über Fechner's psychophysisches Gesetz. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften, mathematisch-naturwissenschaftliche Classe. Abt. 3*, 72, 310–348.
- Hopf, E. 1934. On causality, statistics and probability. *Journal of Mathematics and Physics* 13: 51–102.
- . 1936. Über die Bedeutung der willkürlichen Funktionen für die Wahrscheinlichkeitstheorie. *Jahresbericht der Deutschen Mathematiker-Vereinigung* 46: 176–195.
- Hugenii, C. (Christiaan Huyghens), 1657. De ratiociniis in ludo aleæ. *Exercitationum Mathematicarum, liber 5*. Leyden: Francisci à Schooten, pp. 521–534.
- Huygens, C. 1655–1666/1920. Calcul des probabilités, travaux de mathématiques pures. Tome 14e des *Œuvres complètes de Christiaan Huygens*. La Haye: Martinus Nijhoff.
- Jeavons, W.S. 1874. *The principles of science: A treatise on logic and scientific method*. London: Macmillan and Co.
- Kamlah, A. 1987. The decline of the Laplacian theory of probability: A study of Stumpf, von Kries, and Meinong. Chapter 5 of: L. Krüger, L.J. Daston, and M. Heidelberger, Ed. *The probabilistic revolution*, vol. 1: *Ideas in history*. Cambridge: MIT Press, pp. 91–116.
- Keynes, J.M. 1921. *A treatise on probability*. London: Macmillan and Co.
- Kneale, W. 1949. *Probability and induction*. Oxford at the Clarendon Press.
- Kolmogorov, A.N. 1950. *Foundations of the theory of probability*. New York: Chelsea Publishing Company. (originally published in 1933 as *Grundbegriffe der Wahrscheinlichkeitsrechnung*).
- Kries, J. von 1882. Ueber die Messung intensiver Größen und über das sogenannte psychophysische Gesetz. *Vierteljahrsschrift für wissenschaftliche Philosophie* 4 (3): 257–294.
- . 1886. *Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung*. Freiburg i.B: J.C.B. Mohr (Paul Siebeck). (Reprinted by Mohr in 1927).
- . 1888. Ueber den Begriff der objectiven Möglichkeit und einige Anwendungen desselben. *Vierteljahrsschrift für wissenschaftliche Philosophie* 12 (2) erster Artikel, 179–240; 12(3), zweiter Artikel, 287–323; 12(4), dritter Artikel (Schluss), 393–428.
- . 1892. Ueber Real- und Beziehungs – Urtheile. *Vierteljahrsschrift für wissenschaftliche Philosophie* 16: 253–288.
- . 1899. Zur Psychologie der Urteile. *Vierteljahrsschrift für wissenschaftliche Philosophie* 23 (1): 1–48.
- . 1914. *Immanuel Kant und seine Bedeutung für die Naturforschung der Gegenwart*. Berlin: Springer.
- . 1916. *Logik: Grundzüge einer kritischen und formalen Urteilslehre*. Tübingen: J.C.B. Mohr (Paul Siebeck). Chapter 26: pp. 595–636.

- \_\_\_\_\_. 1919. Ueber Wahrscheinlichkeitsrechnung und ihre Anwendung in der Physik. *Die Naturwissenschaften* 7 (1–2): 2–7.
- Lacroix, S.-F. 1816. *Traité élémentaire du calcul des probabilités*. Paris: M<sup>me</sup> V<sup>e</sup> Courcier.
- Laplace, P.-S. de 1812. *Théorie analytique des probabilités*. Paris: M<sup>me</sup> V<sup>e</sup> Courcier.
- \_\_\_\_\_. 1904. *Oeuvres complètes*. Paris: Gauthier-Villars.
- Lexis, W. 1877. *Zur Theorie der Massenerscheinung in der menschlichen Gesellschaft*. Freiburg i. B: Fr. Wagner.
- Liebermeister, C. 1877. Ueber Wahrscheinlichkeitsrechnung in Anwendung auf therapeutische Statistik, *Sammlung Klinischer Vorträge*. (Innere Medicin No. 31–64), 110, 935–962.
- Lotze, H. 1843. *Logik*, 414–434. Leipzig: Weidmann'sche Buchhandlung.
- Lourié, S. 1910. *Die Prinzipien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung des disjunktiven Urteils*. Tübingen: J.C.B. Mohr (Paul Siebeck).
- Luce, R.D. 1963. On the possible psychophysical laws. In *Readings in mathematical psychology*, ed. R.D. Luce, R.R. Bush, and E. Galanter, vol. 1, 69–83. New York: Wiley.
- Meinong, A. 1915. *Über Möglichkeit und Wahrscheinlichkeit: Beiträge zur Gegenstandstheorie und Erkenntnistheorie*. Leipzig: Johann Ambrosius Barth.
- Meyer, O.E. 1877. *Die kinetische Theorie der Gase: in elementarer Darstellung mit mathematischen Zusätzen*. Breslau: Maruschke & Berendt.
- Meyer, A. 1879. *Vorlesungen über Wahrscheinlichkeitsrechnung*. Leipzig: B.G. Teuber (E. Czuber, Trans.)
- Mill, J.S. 1882. *A system of logic, ratiocinative and inductive, being a connected view of the principles of evidence, and the methods of scientific investigation*. Manhattan: Harper & Brothers, Publishers.
- Mises, R. von 1919. Grundlagen der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift* 5: 52–99.
- Niall, K.K. 1995. Conventions of measurement in psychophysics: von Kries on the so-called psychophysical law. *Spatial Vision* 9 (3): 1–30. The journal *Spatial Vision* continues as *Multisensory Research*.
- \_\_\_\_\_, ed. 2017. *Erwin Schrödinger's color theory, translated with modern commentary*. New York: Springer.
- Pascal, B. 1779. *Oeuvres, tome 4*. La Haye: chez Detune.
- Planck, M. 1909. *Die Einheit des physikalischen Weltbildes*, 24. Leipzig: S. Hirzel. (Vortrag gehalten am 9. Dezember 1908 in der naturwissenschaftlichen Fakultät des Studentenkorps an der Universität Leiden.).
- Poincaré, H. 1912. *Calcul des probabilités*. 2nd ed. Paris: Gauthier-Villars.
- Poisson, S.-D. 1837. *Recherches sur la probabilité des jugements en matière criminelle et en matière civile*. Paris: Bachelier, Imprimeur-Libraire.
- Prévest, P., and S.A.J. l'Huilier. 1796. Mémoire sur l'art d'estimer la probabilité des causes par les effets. *Mémoires de l'Académie Royal des Sciences et des Belles-Lettres (Berlin)* 6: 3–24.
- Pulte, H. 2016. Johannes von Kries's objective probability as a semiclassical concept. Prehistory, preconditions and problems of a progressive idea. *Journal for General Philosophy of Science* 47: 109–129.
- Reichenbach, H. 1920a. Die physikalischen Voraussetzungen der Wahrscheinlichkeitsrechnung. *Die Naturwissenschaften* 8 (3): 46–55.
- \_\_\_\_\_. 1920b. Philosophische Kritik der Wahrscheinlichkeitsrechnung. *Die Naturwissenschaften* 8 (8): 146–153.
- Roberts, J.T. 2016. The range conception of probability and the input problem. *Journal for General Philosophy of Science / Zeitschrift für allgemeine Wissenschaftstheorie* 47 (1): 171–188.
- Rosenthal, J. 2010. The natural-range conception of probability. In *Time, chance and reduction. Philosophical aspects of statistical mechanics*, ed. G. Ernst and A. Hüttemann, 71–91. Cambridge: Cambridge University Press.
- \_\_\_\_\_. 2012. Probabilities as ratios of ranges in initial-state spaces. *Journal of Logic, Language and Information* 21: 217–236.

- . 2016. Johannes von Kries's range conception, the method of arbitrary functions, and related modern approaches to probability. *Journal for General Philosophy of Science* 47: 151–170.
- Shafer, G., and V. Vovk. 2018. The origins and legacy of Kolmogorov's Grundbegriffe. Working Paper #4, The Game-Theoretic and Finance Project, Rutgers School of Business. <https://doi.org/10.18550/arXiv.1802/06071>.
- Sigwart, C. 1878. *Logik, 2. Band: Die Methodenlehre*. Tübingen: H. Laupp'schen Buchhandlung.
- Simpson, T. 1755. A letter to the Right Honourable Earl of Macclesfield, president of the Royal Society, on the Advantage of taking the Mean of a number of Observations in practical Astronomy. *Philosophical Transactions of the Royal Society of London* 49: 82–93.
- Smoluchowski, M.V. 1918. Über den Begriff des Zufalls und den Ursprung der Wahrscheinlichkeitsgesetze in der Physik. *Die Naturwissenschaften*, 17(7), April, 6th year, 23 pp. Max Planck zur Feier seines 60. Geburtstages, pp. 195–264.
- Strevens, M. 1998. Inferring probabilities from symmetries. *Noûs* 32: 231–246.
- Stumpf, C. von 1892a. Ueber den Begriff der mathematischen Wahrscheinlichkeit. *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften zu München, philosophisch-philologischen Classe*, pp. 37–120.
- . 1892b. Ueber die Anwendung der mathematischen Wahrscheinlichkeitsbegriffes auf Teile eines Continuums, *Sitzungsberichte der königlich bayerischen Akademie der Wissenschaften zu München, philosophisch-philologischen Classe*, pp. 681–691.
- Todhunter, I. 1865. *A history of the mathematical theory of probability, from the time of Pascal to that of Laplace*. Cambridge and London: Macmillan and Co.. (Reprinted in 1949, Chelsea Publishing Company).
- Träger, L. 1904. *Der Kausalzusammenhang im Straf- und Zivilrecht*. Marburg: Elwert.
- Treibер, H. 2015. Max Weber, Johannes von Kries and the kinetic theory of gases. *Max Weber Studies* 15 (1): 47–68.
- Venn, J. 1866. *The logic of chance: An essay on the foundations and province of the theory of probability*. London: Macmillan and Co.
- Weber, E.H. 1834. *De pulsu, resorptione, auditu et tactu*. Leipzig: C.F. Koehler.
- Windelband, W. 1870. *Die Lehren vom Zufall*. Berlin: F. Henschel. [Dissertation, Universität Göttingen].
- Witt, J. de 1671. *Waardije van Lyf-renten naer Proportie van Los-Renten* (The value of life annuities compared to redemption bonds). In's Graven-Hage: Jacobus Scheltus.
- Wundt, W. 1883. Ueber die Messung psychischer Vorgänge. *Philosophische Studien* 1: 251–260.
- Zabell, S.L. 2005. *Symmetry and its discontents: Essays on the history of inductive probability*. New York: Cambridge University Press.
- . 2016. Johannes von Kries's *Principien*: A brief guide for the perplexed. *Journal for General Philosophy of Science* 47: 131–150.
- . 2022. Fisher, Bayes, and predictive inference. *Mathematics* 10: 1634, 16 pp.
- Zeller, H. 1881. Ueber die Messung psychischer Vorgänge. *Abhandlungen der Akademie der Wissenschaften zu Berlin, philosophisch-historische Klasse* 3: 1–16.

# Index

## A

d'Alembert, J. le R., 166, 171  
Anscombe, G.E.M., xiv  
Aubert, H., 232

## B

Bayes, T., xi, 73–76, 162, 170, 183  
Béguelin, N. de, 172  
Bernoulli, D., 165  
Bernoulli, J., xii, 162, 163  
Bessel, F.W., 134, 139, 140  
Black, W., 167  
Boltzmann, L., xi, xxix, 118, 121, 123, 125,  
    126, 129–131, 207, 209–211, 214, 215  
Boring, E.G., xv  
Buldt, B., xi

## C

Cohen, L.J., xiii  
Condorcet, N. de, 153, 167  
Cournot, M.A.A., 172, 173, 177  
Czuber, E., xxi

## D

Delboeuf, J.R.L., 237, 238  
Diderot, D., 165  
Dobrowolsky, W., 233  
Drobisch, M.W., 147

## F

Fechner, G.T., x, xiv, 219, 230, 233, 234, 237  
Fick, A., 105, 151, 175–177  
Fioretti, G., xii  
Fismer, F.H., 150  
Foerster, R., 232  
Fries, J.F., 172, 173, 180

## G

Galavotti, M.C., xi  
Gauss, C.F., 138–141, 148, 183  
Gavarret, J., 167  
Goodman, N., xv  
Graunt, J., 163

## H

Hacking, I., xii, xv  
von Hagedorn, F., ix  
Hagen, G., 141  
Halley, E., 163  
Heidelberger, M., xii  
Hering, E., 236  
Hopf, E., xi  
l'Huilier, S.A.J., 171  
Huyghens, C., 162

## J

Jevons, W.S., 85

**K**

- Kamlah, A., xi–xiii  
 Keynes, J.M., xii  
 Kneale, W., xi  
 Kolmogorov, A.N., ix, xv  
 von Kries, J., x–xv

**L**

- Lacroix, S.-F., 89  
 Laplace, P.-S. de, 89, 91, 101, 146, 147, 154,  
     155, 163, 166–168, 173, 187, 188,  
     201, 214  
 Lexis, W., 102, 144, 148, 177  
 Liebermeister, C., 152  
 Lotze, H., 178, 179  
 Lourié, S., xxv, 201, 202  
 Luce, R.D., xiv

**M**

- Meinong, A., xx  
 Meyer, A., 89  
 Meyer, O.E., 132  
 Mill, J.S., 175  
 von Mises, R., xxi, xxv, xxvii

**N**

- Niall, K.K., x, xiv

**P**

- Pascal, B., 161, 162, 164  
 Planck, M., xxix, 214, 215  
 Poincaré, H., xii  
 Poisson, S.-D., xi, 20, 57, 89, 154, 157, 158,  
     169, 172, 173

Prévost, P., 171

Pulte, H., xiv

**R**

- Reichenbach, H., xxix, 134  
 Roberts, J.T., xii  
 Rosenthal, J., xi–xv

**S**

- Sigwart, C., 178, 179, 198  
 Simpson, T., 167  
 Smoluchowski, M.V., xxvii  
 Strevens, M., xiii  
 von Stumpf, C., x, xxiv

**T**

- Todhunter, I., 161, 163  
 Träger, L., xxv  
 Treiber, H., xii

**V**

- Venn, J., 181

**W**

- Weber, E.H., 219, 236  
 Weber, M., xii, xiv  
 Windelband, W., 59, 180  
 de Witt, J., 163  
 Wundt, W., 230, 239

**Z**

- Zabell, S.L., xi–xiv  
 Zeller, H., 230, 239