5

The Evaluation of Probabilities

5.1 How should Probabilities be Evaluated?

In order to say something about this subject without running the risk of being misunderstood, it is first of all necessary to rule out the extreme dilemma that a mathematical treatment often poses: that of either saying everything, or of saying nothing. As far as the evaluation of probabilities is concerned, one would be unable to avoid the dilemma of either imposing an unequivocal criterion, or, in the absence of such a criterion, of admitting that nothing really makes sense because everything is completely arbitrary.

Our approach, in what follows, is entirely different. We shall present certain of the kinds of considerations that do often assist people in the evaluation of their probabilities, and might frequently be of use to You as well. On occasion, these lead to evaluations that are generally accepted: You will then be in a position to weigh up the reasons behind this and to decide whether they appear to You as applicable, to a greater or lesser extent, to the cases which You have in mind, and more or less acceptable as bases for your own opinions. On other occasions, they will be vaguer in character, but nonetheless instructive. However, You may want to choose your own evaluations. You are completely *free* in this respect and it is entirely your own *responsibility*; but You should beware of superficiality. The danger is twofold: on the one hand, You may think that the choice, being subjective, and therefore arbitrary, does not require too much of an effort in pinpointing one particular value rather than a different one; on the other hand, it might be thought that no mental effort is required, as it can be avoided by the mechanical application of some standardized procedure.

5.2 Bets and Odds

5.2.1. One activity which frequently involves the numerical evaluation of probabilities is that of betting. The motivation behind this latter activity is not usually very serious-minded or praiseworthy, but this is no concern of ours here. We should mention, however, that such motivations (love of gambling, the impulse to bet on the desired outcome, etc.) may to some extent distort the evaluations. On the other hand, motives of a different kind lead to similar effects in the case of insurance, where the first objection does not apply.

However, with all due reservation, it is worthwhile starting off with the case of betting, since it leads to simple and useful insights.

Theory of Probability: A Critical Introductory Treatment, First Edition. Bruno de Finetti.

5.2.2. An important aspect of the question (one to which we shall frequently return) is the necessity of 'getting a feeling' for numerical values. Many people if asked how long it takes to get to some given place would either reply 'five minutes' or 'an hour', depending on whether the place is relatively near, or relatively far away: intermediate values are ignored. Another example arises when people are unfamiliar with a given numerical scale: a doctor, although able to judge whether a sick man has a high temperature or not, simply by touching him, would be in trouble if he had to express that temperature on a scale not familiar to him (Fahrenheit when he is used to Centigrade, or vice versa). Likewise, in probability judgments, there are also those who ignore intermediate possibilities and pronounce 'almost impossible' to everything that to them does not appear 'almost certain'. If neither YES nor NO appears sufficiently certain to them, they simply add 'fifty-fifty' or some similar expression. In order to get rid of such gaps in our mental processes it is necessary to be fully aware of this and to get accustomed to an alternative way of thinking.

In this respect, betting certainly provides useful experience. In order to state the conditions for a bet, which have to be precise, it is necessary to have a sufficiently sensitive feeling for the correspondence between a 'numerical evaluation' and 'awareness' of a *degree of belief*. In becoming familiar with judging whether it is fair to pay 10, 45, 64 or 97 lire in order to receive 100 lire if a given event occurs, You will acquire a 'feeling' for what 10%, 45%, 64% or 97% probabilities are. Together with this comes an ability to estimate small differences and a sharpening of that 'feeling for numerical values', which must be improved for the purpose, of course, of analysing actual situations.

5.2.3. These two aspects come together in the particularly delicate question of evaluating very small probabilities (and, complementarily, those very close to 1). Approximations that are adequate (according to the circumstances and purposes involved) in the vicinity of $p = \frac{1}{2}$ (e.g. 50% ±5%, ± 1%, ±0.1%) are different from those required in the case of very small probabilities: here, the problem concerns the order of magnitude (whether, for example, a small probability is of the order of 10^{-3}, or 10^{-7}, or 10^{-20},...). In this connection, it is convenient to recall Borel's suggestion of calling 'practically impossible', with reference to '*human*, *earthly*, *cosmic* and *universal* scales', respectively, events whose probabilities have the orders of magnitude of 10^{-6}, 10^{-15}, 10^{-50} and 10^{-1000}. This is instructive if one wishes to give an idea of how small such numbers (and therefore such probabilities) are, provided that no confusion (in words or, worse, in concepts) with 'impossibility' arises.[1]

5.2.4. *On the use of 'odds'*. In the jargon used by gamblers, the usual way of expressing numerical evaluations is somewhat different, although, of course, equivalent. Instead of referring to the *probability p*, which (in the sense we have given) is the amount of a bet, we refer to the *odds*,

1 Borel himself, and other capable writers, fail to avoid this misrepresentation when they give the status of a principle – 'Cournot's principle' – to the confusion (or the attempt at a forced identification) between 'small probabilities', which, by convention, could be termed 'almost impossibility', and 'impossibility' in the true sense. What is overlooked here is that 'prevision' is not 'prediction'. The topic is dealt with in E. Borel, *Valeur pratique et philosophie des probabilités* (p. 4 and note IV), part of the great *Traité du Calcul des Probabilités* which he edited; Gauthier–Villars, Paris (1924) (and subsequent editions).

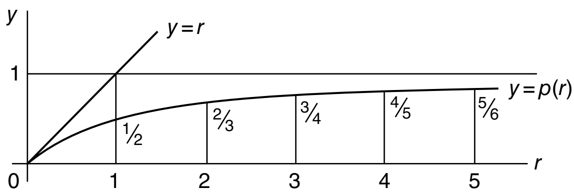

Figure 5.1 The relation between probability (p) and odds (r): $r = p/(1-p)$.

$$r = p/(1-p) = p/\tilde{p}.$$

These are usually expressed as a fraction or ratio, $r = h/k = h{:}k$ (h and k integers, preferably small), by saying that the odds are '*h to k on*' the event, or '*k to h against*' the event. Of course, given r, that is the odds, or, as we shall say, the *probability ratio*, the probability can immediately be obtained by

$$p = r/(r+1), \quad \text{i.e.} \quad (\text{if } r \text{ is written as } h/k) \quad p = h/(h+k). \tag{5.1}$$

A few examples of the correspondence between probabilities and probability ratios, and vice versa, are shown below and illustrated in Figure 5.1:

p	$p/\tilde{p}$	$= r$	$= h/k$	in words	(check) $h/(h+k) = p$
0.20	20/80	=0.25	=1/4	'4 to 1 against'	$1/(1+4) = 0.20$
$2/7 = 0.286$	28.6/71.4	=0.40	=2/5	'5 to 2 against'	$2/(2+5) = 0.286$
0.50	50/50	=1	=1/1	'evens'	$1/(1+1) = 0.50$
0.75	75/25	=3	=3/1	'3 to 1 on'	$3(3+1) = 0.75$

Observe that to the complementary probability, $\tilde{p} = 1-p$, there corresponds the reciprocal ratio, $\tilde{p}/p = 1/(p/\tilde{p}) = 1/r$ (i.e. to 'h to k on' there corresponds the symmetrical phrase 'k to h on').[2]

5.2.5. *Extensions*. Probability is preferable by far as a numerical measure (additivity is an invaluable property for any quantity to possess!).[3] However, there are cases in which it is advisable to employ the probability ratio (especially in cases involving likelihood – Chapter 4 – which are often considered in the form of 'Likelihood Ratio') and it

2 It would perhaps be better to introduce a notation to indicate that we are passing from probability to 'odds'; similar to that used for 'complementation' ($\tilde{P} = 1-p$). An analogous approach would be to take $\underset{\sim}{p} = p/\tilde{p}$ (and if $p = \mathbf{P}(E)$ to use therefore $\underset{\sim}{\mathbf{P}}(E) = \mathbf{P}(E)/\mathrm{P}(\tilde{E}) = \mathbf{P}(E)/\!\sim \mathbf{P}(E)$). We prefer merely to draw attention to the possibility without introducing and experimenting with more new ideas than prove to be absolutely necessary. To avoid any difficulties, or risks of confusion in notation, we denote the odds more clearly by writing $O(x) = x/(1-x)$, $O[\mathbf{P}(E)] = \mathbf{P}(E)/\mathbf{P}(\tilde{E})$.

3 A newspaper, in considering three candidates for the American presidential election, attributed odds of 2 to 1 on, 3 to 1 against and 5 to 1 against; these are equivalent to probabilities of $\frac{2}{3}, \frac{1}{4}, \frac{1}{6}$, with sum $(8+3+2)/12 = 13/12 > 1$. It is difficult for a slip of this nature to pass unnoticed when expressed in terms of probabilities; using percentages especially, it would certainly not escape notice that $67\% + 25\% + 17\% = 109\%$ was inadmissible.

is useful to indicate, at this juncture, the way in which we shall generalize its use (or, in a certain sense, substitute for it) in cases where the need arises.

In accordance with, and in addition to, the conventions introduced in Chapter 3, Section 3.5, concerning the use of the symbol **P**, we can denote that $r = h/k$ by writing

$$\begin{aligned}\mathbf{P}\left(E,\ \tilde{E}\right) &= \left(h/\left(h+k\right),\ k/\left(h+k\right)\right) \\ &= \left(h, k\right)/\left(h+k\right) = K\left(h, k\right) = \left(h:k\right) = \mathbf{P}\left(E:\tilde{E}\right),\end{aligned} \tag{5.2}$$

where we have successively and implicitly made the following conventions:

a common factor, such as $1/(h+k)$, can be taken outside the parentheses; that is $m(a, b) = (ma, mb)$;

such a factor may be taken as understood, denoting it by K, to simply mean that proportionality holds;

the same thing may be indicated by simply using the 'colon' (:) as the dividing sign, rather than the comma. This means that two n-tuples of numbers $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$, not all zero, are said to be proportional if $b_i = Ka_i$ where K is a nonzero constant. Proportionality is sometimes denoted by the sign $\propto$ (which is not very good), and can also be expressed by $= K$. We make the convention – once and for all – that K denotes a *generic* coefficient of proportionality, whose value is not necessarily the same, not even for the duration of a given calculation: we can write, for example, $(2,1,3) = K(4,2,6) = K(6,3,9)$. The equals sign is sufficient on its own if the n-tuple with ':' in place of ',' is interpreted as 'up to a coefficient of proportionality' (like homogeneous coordinates); that is as a multi-ratio. Hence, for example, $(2:1:3) = (4:2:6) = (6:3:9)$.

Sometimes the omission of the proportionality factor is irrelevant because it is determined by normalization: for example, if it is known that $E_1 \ldots E_n$ constitute a partition, and we write

$$\mathbf{P}\left(E_1 : E_2 : \ldots : E_n\right) = \left(m_1 : m_2 : \ldots : m_n\right), \tag{5.3}$$

it is clear that $\mathbf{P}(E_i) = m_i/m$, $m = m_1 + m_2 + \ldots + m_n$, because the sum must equal 1. In other cases (for any E_i whatsoever, even if they are compatible), one can make the common divisor m enter in explicitly, for example by adding in 1 = the certain event:

$$\mathbf{P}\left(E_1 : E_2 : \ldots : E_n : 1\right) = \left(m_1 : m_2 : \ldots : m_n : m\right). \tag{5.4}$$

The resulting convenience is most obvious when the m_i are small integers. For example, if A, B, C form a partition ($A + B + C = 1$), by writing $\mathbf{P}(A{:}B{:}C) = (1{:}5{:}2)$(even without the refinement $\mathbf{P}(A{:}B{:}C{:}1) = (1{:}5{:}2{:}8)$) it becomes obvious that

$$\mathbf{P}\left(A\right) = 1/8 = 12\cdot 5\%, \quad \mathbf{P}\left(B\right) = 5/8 = 62\cdot 5\%, \ \mathbf{P}\left(C\right) = 2/8 = 25\%.$$

At this point we shall also introduce the operation of the *term-by-term* product of multiratios, denoting it by *:

$$\left(a_1 : a_2 : \ldots : a_n\right) * \left(b_1 : b_2 : \ldots : b_n\right) = \left(a_1 b_1 : a_2 b_2 : \ldots : a_n b_n\right). \tag{5.5}$$

This frequently provides some advantage in handling small numbers or simple expressions in a long series of calculations, and will turn out to be particularly useful for the applications to likelihood which we mentioned above.

The time has now come to end this digression concerning methods of numerically denoting probabilities and to return to questions of substance.

5.3 How to Think about Things

5.3.1. In discussing the central features of the analysis which must underlie each evaluation, it will be necessary to go over many things which, although obvious, cannot be left out, and to add a few other points concerning the calculus of probability.

The following recommendations are obvious, but not superfluous:

- to think about every aspect of the problem;
- to try to imagine how things might go, or, if it is a question of the past, how they might have gone (one must not be content with a single possibility, however plausible and well thought out, since this would involve us in a *prediction*: instead, one should encompass all conceivable possibilities, and also take into account that some might have escaped attention);
- to identify those elements which, compared with others, might clarify or obscure certain issues;
- to enlarge one's view by comparing a given situation with others, of a more or less similar nature, already encountered;
- to attempt to discover the possible reasons lying behind those evaluations of other people with which, to a greater or lesser extent, we are familiar, and then to decide whether or not to take them into account. And so on.

In particular, in those cases where bets are made in public (e.g. horse races, boxing matches – in some countries even presidential elections) some sort of 'average public opinion' is known by virtue of the existing odds. More precisely, this 'average opinion' is that which establishes a certain 'marginal balance' in the demand for bets on the various alternatives. This might be taken into consideration in order to judge, after due consideration, whether we wish to adopt it, or to depart from it, and if so in which direction and by how much.

5.3.2. In order to provide something by way of an example, let us consider a tennis match between two champions, A and B.[4] You will cast your mind back to previous matches between them (if any); or You will recall matches they had with common opponents (either recently, or a long time ago, under similar or different conditions); You will consider their respective qualities (accuracy, speed, skill, strength, fighting spirit, temper, nerves, style, etc.) and the variation in these since the last occasion of direct or indirect comparison; You will compare their state of health and present form, and so on; You will try to imagine how each quality of the one might affect, favourably or otherwise, his opponent's capacity to settle into the game, to fight back when behind, to avoid losing heart, and so on. For instance, you may think that B, although on the whole a better player, will lose, because he will soon become demoralized as a result of A's deadly service. However, it would be naïve to stop after this first and lone supposition: it would mean to aspire to making a prediction rather than a prevision. You will go on next to think of what might happen if this initial difficulty for B does not materialize,

4 This example has already been discussed by Borel and again by Darmois (see p. 93 of the Borel work mentioned previously, and again on p. 165 Darmois' note VI). As is clear from this and other examples (like his discussion, again on p. 93, of the evaluation of a weight – similar to our example in Chapter 3, Section 3.9.7), Borel seems to be inspired – in the greater part of his writings – by the subjectivistic concept of probability: he can thus be regarded as one of the great pioneers, although incompatible statements and interpretations crop up here and there, as was pointed out in the footnote to Section 5.2.3.

or is overcome, and little by little You will obtain a summary view – but not a one-sided or unbalanced one – of the situation as a whole. Your ideas about the values to attribute to the winning probabilities for A and B will in this way become more precise. You may have the opportunity to compare your ideas and previsions with those of other people (in whose competence and information You have a greater or lesser confidence, and whom You may possibly judge to be more or less optimistic about their favourite). In the light of all this, You might think over your own point of view and possibly modify it.

5.3.3. Our additional remarks concerning the calculus of probability consist in pointing out that the conditions of coherence, even if they impose no limits on the freedom of evaluation of any probability, do in practice very much limit the possibility of 'extreme' evaluations. More precisely, an *isolated* eccentric evaluation turns out to be impossible (the same thing happens, for instance, to a liar, who, in order to back up a lie, has to make up a whole series of them; or to a planner, who must modify his entire plan if one element is altered).

It is easy to say 'in my opinion, the probability of E is, roughly speaking, twice what the others think it is'. However, if You say this, I might ask 'what then do You consider the probabilities of A, B, C to be?', and, after You answer, I may say 'so do You think the probability of H is as small as this; $\frac{1}{10}$ of what is generally accepted?', and so on. If You remain secure in your coherent view, You will have a complete and coherent opinion that others may consider 'eccentric' (with as much justification as You would have in calling the common view 'eccentric') but will not otherwise find defective. However, it will more often happen that as soon as You face the problem squarely, in all its complexity and interconnections, You come to find yourself in disagreement *not only with the others but also with yourself*, by virtue of your eccentric initial evaluation.

We have been talking in terms of bets and the evaluations of probability, and not of previsions of random quantities, although they are the same thing in our approach. This was simply a question of the convenience of fixing ideas in the case where the probabilistic aspect is most easily isolated: however, one should note that the same considerations could in fact be extended to the general situation.

5.4 The Approach Through Losses

The betting set-up is related to the 'first criterion' of Chapter 3, Section 3.3; the scheme we are now going to discuss is based on the 'second criterion': it is this latter – as we remarked previously, and will shortly see – which turns out to be the more suitable.

First of all, we shall find it convenient to present this scheme right from the beginning again, referring ourselves now to the case of events. Because this is the simplest case, and because we are treading an already familiar path – which we shall illustrate clearly with diagrams – everything should appear both more straightforward and of wider application.

5.4.1. Instead of some general random quantity X, You must now think in terms of an event E, such that You are free to choose a value x, bearing in mind that You face a loss

$$L = L_x = (E - x)^2. \tag{5.6}$$

Expanding this (remembering that $E^2 = E$), one obtains the following alternative expressions (in the last one, p is any number whatsoever):

$$\begin{aligned}
&(\mathrm{a}) \quad L_x = x^2 + (1-2x)E, \\
&(\mathrm{b}) \quad = x^2\tilde{E} + (1-x)^2 E, \qquad (5.7)\\
&(\mathrm{c}) \quad = E(1-p) + (p-x)^2 + (E-p)(p-2x).
\end{aligned}$$

They all reveal (5.7b most explicitly) that L_x equals x^2 or $(1-x)^2$ according to whether $E = 0 =$ false or $E = 1 =$ true.

Since we have already used the criterion as a definition – and hence already know what the probability $p = \mathbf{P}(E)$ of E is – we can, 'being wise after the event', examine how the criterion behaves by looking at $\mathbf{P}(L_x)$, considered as a function of a value x and of a probability p, assumed to be arbitrary (so we adopt the notation $L_x(p)$). Putting $E = p$ in equations 5.7a, 5.7b, and 5.c (which are linear in E) we obtain:

$$\begin{aligned}
&(\mathrm{a}) \quad L_x(p) = x^2 + (1-2x)p \\
&(\mathrm{b}) \quad = x^2\tilde{p} + \tilde{x}^2 p, \qquad (5.8)\\
&(\mathrm{c}) \quad = p(1-p) + (p-x)^2 = p\tilde{p} + (p-x)^2.
\end{aligned}$$

5.4.2. We now examine the variation in $L_x(p)$ as p varies, x being an arbitrary fixed value. As might have been expected (5.8a shows this up most clearly), L_x varies linearly from $L_x(0) = x^2$ to $L_x(1) = \tilde{x}^2$ (which are the two possible values for L_x, depending on the occurrence of either $\tilde{E}(p = 0)$ or $E(p = 1)$). The straight lines in Figure 5.2, connecting these extreme values, give a visual impression of how they go together: that is of how, in order to reduce the penalty resulting in one case, one must increase it in the other.[5]

The figure also shows, in an indirect way, the variation of $L_x(p)$ for varying x, with p fixed. Geometrically, one can see (and equation 5.8c) presents it explicitly) that the straight lines are the tangents to the parabola $y = p(1-p) = p\tilde{p}$, and that none of them can go beneath their envelope (this is within the interval [0, 1]: the others would correspond to values $x < 0$ and $x > 1$; see footnote 5). Given p, the best one can do is to take the tangent at p, obtained (as we already know!) by choosing $x = p$: this gives $L_x(p)$ its minimum value (as x varies), $L_x(p) = p\tilde{p}$. Choosing a different x gives rise, in prevision, to an additional loss $(x-p)^2$; that is the square of the distance from x to p: equation 5.8c shows this explicitly, by splitting the linear function $L_x(p)$ into the sum of $p(1-p)$ (the parabola) and $(x-p)^2$ (the deviation from the parabola of the tangent at $p = x$). We observe also, and this confirms what has been said already, that this deviation is the same for all the tangents (starting, of course, from their respective points of contact).

The maximum loss is 1, and this is achieved by attributing probability zero to the case that actually occurs: the minimum loss is 0, and is achieved when a probability of 1 (or 100%) is attributed to this case. For any given x, the loss varies between x^2 and $\tilde{x}^2$ (as we have seen already). For a given p, we already know that the minimum is $p\tilde{p}$ (for $x = p$), and it is readily seen that the minimum is $p \vee \tilde{p}$: more precisely, if $p \leqslant \frac{1}{2}$ it is $1-p$, obtained by choosing $x = 1$; if $p \geqslant \frac{1}{2}$ it is p, obtained by choosing $x = 0$. If $p = \frac{1}{2}$, we have the maximum of

5 This is for $0 \leqslant x \leqslant 1$: we already know, and can also see, that, in every case, $x < 0$ or $x > 1$ is worse than $x = 0$ or $x = 1$, respectively, and is thus automatically ruled out (without need of any convention).

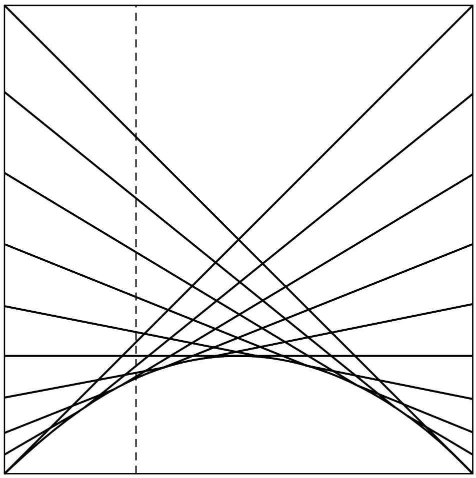

Figure 5.2 The straight lines correspond to the combinations of penalties among which the method allows a choice (the penalty can be reduced in one of the two cases at the expense of increasing it in the other: lowering the ordinate at one end raises it at the other). The ordinate of a particular straight line at the point *p* is the prevision of the loss for the person who chooses that line and attributes probability *p* to the event under consideration. In this case, the minimum value that can be attained is given by the ordinate of the parabola (no straight line passes beneath it!), and the optimal choice of straight line is the tangent to the parabola at the point with abscissa *p*.

the minimum ($p\tilde{p} = \frac{1}{4}$), and the minimum of the maximum ($p \vee \tilde{p} = \frac{1}{2}$), and hence the largest discrepancy ($\max - \min = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$); in general, the discrepancy is $x^2 \vee \tilde{x}^2$, that is the maximum of x^2 and $(1-x)^2$, and attains its maximum (=1) for $x = 0$ and $x = 1$.[6]

5.4.3. *The case of many alternatives.* We can deal with the case of many alternatives (of a multi-event, of a partition), and the more general case of any number of arbitrary but not incompatible events, by applying the previous scheme to each event separately. In this way, things reduce to the treatment given in Chapter 3, and to the geometric representation which was there illustrated. Here, we simply wish to review the approach in the spirit of the above considerations, and then to look at a few modifications.

It will suffice to consider a partition into three events (such as E_1, E_2 and E_3 of Chapter 3, 3.9.2). We shall call them A, B and C ($A + B + C = 1$) and represent them as points, $A = (1, 0, 0)$, $B = (0, 1, 0)$ and $C = (0, 0, 1)$, in an orthogonal Cartesian system. For the time being, we shall distinguish the probabilities, $p = \mathbf{P}(A)$, $q = \mathbf{P}(B)$ and $r = \mathbf{P}(C)$, attributed to them, from the values x, y and z chosen in accordance with the second criterion (we know they must coincide but we want to investigate what happens if we choose them to be different, either through whim, oversight or ignorance).

6 Among other decision criteria that are employed (inspired by points of views which differ from ours) is one which is called the 'minimax' criterion: it consists in taking that decision that minimizes the maximum possible loss. Observe that, in the above situation, this criterion would have us always choose $x = \frac{1}{2}$ (then, in fact, the loss would $= \frac{1}{4}$, with certainty, whereas every other choice would give a smaller loss in one of the two cases, although greater in the other). Since it is incoherent to attribute probability $\frac{1}{2}$ to all events, such a criterion is absurd (in this kind of application; not so, however, in the theory of games – see Chapter 12, Section 12.7.4 – where it provides a solution in situations of a different kind, nor even in this situation under an hypothesis of an extremely convex utility function where it would no longer lead to the choice of $p = \frac{1}{2}$).

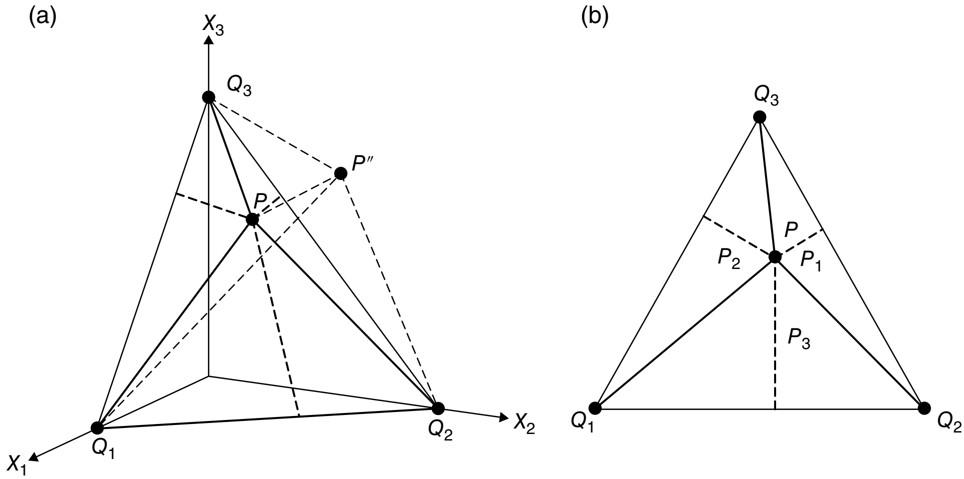

Figure 5.3 The triangles of points such that $x+y+z=1$ (x, y, z non-negative) seen in (a) space, and (b) in the plane. It is clear from geometrical considerations that the choice of a loss rule corresponding to the point (x, y, z) is inadmissible (in the case of three incompatible events) if it is not within the given triangle. Moreover, if one attributes the probabilities (p, q, r) to the three events, it pays then to choose $x=p$, $y=q$, and $z=r$. In other words, the method rewards truthfulness in expressing one's own evaluations.

We shall denote by P the prevision-point $P=(p, q, r)$; the decision-point will be denoted by P'', $P''=(x, y, z)$ (Figure 5.3).

The total loss will then be

$$L=(A-x)^2+(B-y)^2+(C-z)^2 \tag{5.9}$$

and

$$\mathbf{P}(L)=\left[p\tilde{p}+q\tilde{q}+r\tilde{r}\right]+\left[(p-x)^2+(q-y)^2+(r-z)^2\right]; \tag{5.10}$$

in other words,

$$\mathbf{P}(L)=\left(\text{first term involving only the prevision}-\text{point } P\right)+\left(P''-P\right)^2,$$

the latter being the square of the distance between P'' and P. Hence, in order to avoid an extra loss, whose prevision is equal to the square of the distance between P'' and P, the point P'' must be made to coincide with P.

The argument given previously (Chapter 3, 3.9.2) was saying the same thing, but without reference to a preselected prevision $\mathbf{P}$. Given a point $P''=(x, y, z)$, outside the plane of A, B and C (i.e. with $x+y+z\neq 1$), its orthogonal projection P' onto this plane has distance less than P'' from A, B and C; if P' falls outside the triangle ABC, the above-mentioned distances decrease if one moves from P' to the nearest point P on the boundary. This shows that only the points of the triangles are admissible (in the sense of Pareto optimality; there are no other points giving better results in all cases). The present argument is less fundamental, but more conclusive, because – assuming the notion of probability to be known in some way (e.g. on the basis of the first criterion) – it shows how and why the evaluations x, y, z of the second criterion must be chosen to coincide with the probabilities p, q, r of A, B and C.

5.4.4. We have here dealt with the most formally immediate case, that of applying to the different events (A, B, C) one and the same scheme with the same maximum loss, namely unity. We know, however (see Chapter 3, 3.3.6), that, so far as the evaluation of probabilities is concerned, and this is what interests us, no modifications would be required were we to use different coefficients: for instance, if one were to take

$$L = a^2\left(A-x\right)^2 + b^2\left(B-y\right)^2 + c^2\left(C-z\right)^2$$

with arbitrary a, b, c. Geometrically, the three orthogonal unit vectors, $A-O$, $B-O$, $C-O$, must now be taken to have lengths a, b and c. This implies – and it is this aspect which may be of interest to us – that the loss, which always equals the square of the distance, is given by $(A-B)^2 = a^2 + b^2$, if in prevision all the probability is concentrated on A, and B actually occurs (and conversely): similarly for $(A-C)^2 = a^2 + c^2$ and $(B-C)^2 = b^2 + c^2$. In the plane of A, B and C, the triangle ABC can be any acute-angled triangle (in the limit, if one of the coefficients is zero, it can be right-angled): in fact, $a^2 = (B-A)\times(C-A) = \overline{AB}\ .\ \overline{AC}.\ \cos\widehat{BAC}$, $\cos\widehat{BAC} > 0$, and so on. In any case, the scheme would work in the same way even if ABC were taken to be any triangle whatsoever, although if it were obtuse-angled, we could not obtain it as we just did in orthogonal coordinates (merely by changing the three scales). This is obvious by virtue of the affine properties, a point we have made repeatedly. In the general case, the only condition imposed on the three losses $\overline{AB}^2, \overline{AC}^2, \overline{BC}^2$ is the triangle inequality for $\overline{AB}, \overline{AC}, \overline{BC}$.

5.4.5. Why are we bothering about the possibility of modifying the shape of the triangle: that is the ratios of the losses in the different cases? After all, this is irrelevant from the point of view of evaluating probabilities. Despite this, it may sometimes be appropriate to draw a distinction between the more serious 'mistakes', and the less 'serious' (the former to be punished by greater losses), in those cases in which the losses could also serve as a useful means of comparison when considering how things turn out for different individuals (as we shall see shortly).

A good example, and one to which we shall subsequently return, is that of a football match (or some similar game) in which the following three results are possible: A = victory, B = draw, C = defeat. In the most usual case (triangle ABC equilateral), one considers it 'equally bad' if either a draw or defeat results when one has attributed 100% probability to victory. If, on the other hand, the distance between victory and defeat is considered greater than the distance between each of these and a draw, we could take an isosceles triangle with the angle B greater than 60°; if we take this angle <90°, we have a combination of three losses for the three results, and the loss for victory–defeat will be less than twice the loss for draw–defeat (or for draw–victory). For a right angle, this ratio will be exactly double (the ratio of the sides = $\sqrt{2}$) and the scheme will only be applicable to the events victory and defeat (a draw is only taken into account as complementary to the other two). For angles between 90 and 180°, the interpretation as combinations of losses for the three events no longer holds; for the case of a draw, the loss would have to be *negative* in order for things to proceed smoothly! The 180° case means that we are effectively considering prevision in terms of 'points' (0 for a defeat, 1 for a draw, 2 for a victory), in the sense that previsions like (0,1,0) and $(\frac{1}{2}, 0, \frac{1}{2})$ – that is of being certain of a draw, or of equal probabilities for victory and defeat, excluding a draw – are considered identical.

5.5 Applications of the Loss Approach

5.5.1. The employment of this method (or something similar) by various people for evaluating probabilities should be given great emphasis and, for many, many reasons, deserves wide publicity.

Sometimes, one is interested in knowing the opinion of a given individual, or of various individuals, concerning the probabilities of certain events under consideration. Sometimes, in order to make some kind of psychological analysis, one is interested in knowing how the various individuals react to information, or other new factors. In certain other cases, it might be interesting to be able to judge, in a more precise fashion, the extent of the 'partial knowledge' of individuals under examination: for instance, one might discourage them from 'guessing'.[7] And so on.

In all these cases, one should take into account the no less important value of repeated experiences of this kind. They greatly aid one in acquiring the 'feeling for numerical values' with which one expresses 'degrees of belief', and hence they contribute to building up a keen and accurate understanding of the problems of prevision, and of the spirit – not cut-and-dried – in which probability theory must approach them.

5.5.2. With this aim in mind, we must now supply all the details of the method. It must be understood that it is preferable to express one's own evaluations sincerely and accurately, and that otherwise one suffers a loss, equal, *in prevision* (in one's own evaluation), to the square of the distance between one's own true evaluation and the one expressed. In addition, there is a definite advantage in obeying the conditions of coherence (in our example; $x, y, z \geqslant 0$, $x+y+z=1$): to do otherwise is to arrange to suffer one part of the loss *with certainty*. If, instead, one wishes to check – in a decision-theoretic sense – the ability of a given individual to do the right thing without having a systematic knowledge of the situation and of the theory, the characteristic features of the method should not be revealed (except for mentioning what losses are). This is a different problem, however; a far cry from those for which we have introduced the method under present discussion (and it seems unlikely, anyway, that anyone could come to sensible decisions without knowing and applying – with great care! – the theory of probability).[8]

Let us now proceed to some concrete examples of various types of applications.

5.5.3. *The opinions of experts*. It often happens that one turns to the experts for information. This is, in actual fact, nothing other than an evaluation of probability. One is

---

7 By 'guessing', we mean 'guessing at random'. This should not be confused with the usage conveyed in Pólya's 'Let us teach guessing', where it means to make useful conjectures (first guess, then prove!).

8 Experiments of this kind, which are made in order to check the extent to which actual behaviour conforms to the norms derived from the theory of probability, are often considered as 'proving' or 'disproving' the validity of probability theory (or of the related theory of decision making under conditions of uncertainty). This would be so if such theories were to be regarded as empirical–psychological theories of actual behaviour, but, in fact, it is completely at odds with what we are considering here: a *normative* theory for *coherent* behaviour.

Many criticisms derive from this confusion (or from the refusal to accept that a subjectivistic theory can distinguish incoherent and coherent behaviour, rather than just being an acritical, empirical observation of actual behaviour as it happens to be). This kind of empirical evidence is also of interest from our standpoint, but in the same way as a mathematician might find the mistakes of laymen, students, or even other mathematicians, interesting. He does not modify mathematics by incorporating these 'mistakes', as though, simply because someone has enunciated them, they 'should' be included by virtue of their being part of some psychological truth, or of the indiscriminate collection of mathematical statements made in the course of history.

not always in a position to weigh-up for oneself all the probabilities relevant to a given situation; this then is the time to behave like the Prince, who, according to Machiavelli, 'sometimes understands things by himself, sometimes through the understanding of others: while the former is excellent, the latter is also very good.'

An example, one of thousands, is given by the case of a geologist who is asked to give an opinion as to whether it is worth drilling a hole at a particular site during an oil search. This is a useful example to consider, since it has, to some extent, been treated by Grayson,[9] and so the interested reader can delve deeper into those aspects which we shall not discuss. The geologist himself does not have any say in the final decision of whether or not to drill: this decision must be taken (by the 'decision maker') after consideration of all the various pieces of information, of which that of the geologist is just one. He, for his part, cannot state categorically that oil is present or not present (thus making a prediction rather than a prevision), nor can he sin in the opposite direction and merely list the information about the geology of the area (reliable, but analytical), leaving to others the task of synthesis and drawing some conclusions. The synthesizing and the conclusions about the probable outcome of the drilling – given from a geological standpoint – are precisely what his expertise is called upon to provide.

In actual fact, the geologist's report does provide this answer, but usually couched in extremely vague adjectives or phrases (such as: fairly good prospects, or good, favourable, uncertain, promising, etc.; sometimes preceded by little words like 'very', 'not very', 'quite', 'rather', and followed cautiously by 'unless anything unexpected happens', 'perhaps', 'it's difficult to say', 'in my humble opinion',..., 'God only knows'). The only solution worthy of serious consideration is to have the geologist express the probabilities numerically, and some companies actually do this. The objection could be raised (and often is) that the knowledge of the geologist is too vague to be represented numerically. It would certainly be unwise and overzealous to assert that the probability of striking oil at a given site is 0.1307594, but to state that the probability is 0.131, or 0.13, or even simply 10–15%, is always preferable to a string of adjectives whose vagueness depends upon the nature of the opinion itself, on the inadequacy of language, and, perhaps, on a desire to state the conclusions in the least compromising way – that is essentially ambiguous, but not appearing to be.[10]

5.5.4. There remains the problem, however: *how can we interest the expert* – in our case the geologist – *in giving an honest answer; in expressing accurately his deep-felt belief*? This problem was examined by Grayson in the light of the 'first criterion', without any satisfactory solution being obtained. The method we suggest here – that of the 'second criterion' – would seem to give a perfectly satisfactory solution, and is precisely what Grayson requires; 'a system to discourage falsification'. For the practical application at present under consideration, it would be sufficient to agree that some part of the agreed fee (neither insignificant, nor excessive; say, 5–10%) be held back until the eventual outcome was known, and then the loss deducted (up to a maximum of the amount held back) before payment. In certain cases, however, like those of experts who are

9 C.J. Grayson, *Decisions Under Uncertainty*: *Drilling Decisions by Oil and Gas Operators*, Harvard Business School (1960).

10 Someone made the acute observation that often the ability to make accurate predictions consists in expressing them in a sufficiently imprecise fashion (this principle is mentioned on p. 213 of Good's anthology – see footnote 12 – and also in a review article of mine; see *Civilta delle macchine*, No. 1 (1963), 71–72). On the other hand, the limit-case of Sibylline predictions ('*Ibis, redibis,...*') is well known.

consulted regularly, or who hold positions within the firm, one might also add up the losses – expressed as 'scores' – in order to make global comparisons (of the 'goodness' of the previsions of two individuals based on comparisons of the cases examined by both of them). These comparisons could be made separately according to 'type' of problem, time period and so on, and could be taken into account when considering the merits of someone in connection with appointments, promotion and so on.

The following discussion is useful both in real-life and as an example.

5.5.5. *Forecasting sports results.* We consider sports results, football in particular, because they give plenty of scope for experiments of this kind: they can be observed regularly (e.g. every weekend) and sufficiently often; the outcomes are clear-cut (in football, the home team either wins, loses or draws), officially ratified, and the situation is well known to most people. In addition, there is considerable background information and comment in the newspapers. However, leaving aside the convenience (for the reasons given above) of sports results, we could consider forecasting in any area (e.g. politics, economics, meteorology, everyday affairs, culture, judicial or sanitary matters, personal or business affairs, etc.).

There are, as is well known, various organized pools for betting on football and horse racing. These, however, are motivated by the concept of 'prediction', in that they reward those who *guess* all (or almost all) of the *results*. Moreover, the sensitivity of the system is completely distorted by the practice of sharing out the available prize money among the winners. Indeed, the net result of all this is as follows: those who write down ridiculous forecasts, that by chance turn out to be correct, receive fantastic prizes; whereas those who write down forecasts which could reasonably be thought probable receive, if they win, only very small amounts, since the prize, in this case, will presumably have to be shared with many others.[11] Consequently, the 'most reasonable' way to gamble would not be to bet on the result for which the probability of occurrence is highest but, instead, to consider the probability multiplied by the prevision of the reciprocal of the number of people betting on it, and to bet on the result for which this is highest.

The betting approach that we discussed previously, illustrating its merits and demerits, is in line with the notion of prevision (as opposed to prediction). The scheme we are now going to present is intended to build on the merits and eliminate the demerits. It should, therefore, permit us to achieve those goals that we have already mentioned: to develop a feeling for what a prevision (not a prediction) is, and a feeling for the numerical scale on which it is to be expressed; to teach one how to take into account the relevant circumstances, bearing in mind one's own level of competence. Moreover, all this is achieved within the agreeable format of a competition, there being the additional opportunity to reflect and to compare, after the event, one's own previsions with those of others, and with the results themselves. It will be necessary to consider rather carefully the latter point; that is '*being wise after the event*'. We shall do so in Sections 5.9 and 5.10 of this chapter, and will come back to it on several occasions later in the book.

11 This brings to mind a rather tragic story: a man died, overwhelmed with joy, on learning that he had guessed correctly all the 13 football results on the Italian football pools. In fact, he was lucky, because otherwise he would have died of disappointment the next day on learning that his winnings were so small (about 3000 Lire), owing to the predictability of the results, which were therefore foreseen by many others besides himself.

5.5.6. One could organize a competition more or less along the following lines (this has already been tried, although on a small scale).[12] The participants have to hand in, each week, previsions for the forthcoming matches, giving, for each match, the probabilities (expressed in percentages) of the three possible outcomes (in the order: win, draw, defeat); writing, for instance, 50–30–20, 82–13–05, 32–36–32 and so on. Given the results, one can evaluate, game by game, the losses and the total losses for the day (and, possibly, a prize for the day), as well as the cumulative sum needed for the final classification. This final classification must be seen as the primary objective. If there are prizes, the largest should be reserved for the final placements, and, in order to conform to the spirit of the competition, the prizes must complement losses; that is they should depend on them in a *linearly* decreasing fashion.[13]

The lessons of experience tell us much about the necessity of avoiding the mentality of prediction when making previsions. It is true that total success – that is no penalty – is achieved if and only if the whole probability, 100%, is attributed to the case which actually occurs. For this reason, many find it tempting, especially at first, to attempt to get the result spot-on, with evaluations which ignore the possibility of uncertainty (i.e. 100–00–00, 00–100–00 or 00–00–100, which are equivalent, in the notation of the football pools, to the 'predictions' '1', 'X', '2'). However, these participants come to realize very quickly that they have fallen behind – this happens on the individual days, but shows up most in the final classification – relative to those who distribute probability in a sensible way: they soon modify their approach.

We shall come back to this example later.

5.5.7. *Replies to multiple-choice questions*. One is often required in a 'quiz', or even in an examination (especially in America), to choose from among a few given answers the one which one believes to be correct. The exact details may differ somewhat: one may either have to tick one and only one answer; or be allowed to choose none; or to choose a subset within which the correct answer is thought to lie (and, in this case, there are two variants, according to whether one indicates an order of preference or not). In any case, there must be an agreed method of scoring according to the way in which the answers given compare with the correct answers. A problem arises from the necessity of discouraging people from 'guessing'; this is often dealt with by estimating statistically what the effect of the assumed presence of 'guessing' would be, in a mass of people.

12 It was tried twice, in 1960–1961 and 1961–1962, in the Economics Faculty of Rome University. There were about 30 participants (students and a few teachers) on each occasion, and the study centred on the nine football matches played every week in the first division of the Italian league. Some discussion of this can be found in B. de Finetti, 'Does it make sense to speak of "good probability appraisers"?' in the volume entitled *The Scientist Speculates: An Anthology of Partly-baked Ideas* (edited by I.J. Good) Heinemann, London (1962). The experiment was repeated again in Rome (Faculty of Science) from 1966 on, and experiments of this kind have recently been made in the United States.

13 If, for example, no prize is to be awarded to those who come last (by whatever ruling is proposed), not only do the tail-enders have no motivation to exercise care in their evaluations but, on the contrary, they have a vested interest in trying outlandish evaluations, which they presume to be different from those of individuals in a better position. This is their only hope of overtaking them and getting a prize. If the first prize were extremely large, the temptation to behave in this way would be greatest for those in second place on the next to last day. In any case, such a distortion of interest occurs whenever *linearity* is abandoned.

This latter problem is completely resolved if one applies the method under consideration.[14] Observe that, in this context, there is no question of events which could be considered 'uncertain' in some 'objective' sense. For example: it is clear that if we ask which of A = Antonio, B = Brutus, C = Caesar said the famous line 'Alea jacta est', we are not asking for any sort of testimony or opinion concerning the fact that some great man uttered the phrase in the course of his life; we simply wish to check whether the examinee knows that the phrase relates to Caesar and the crossing of the Rubicon. In the same way, if we ask whether $\log x + \log y$ equals

$$A = \left(\log(x+y)\right), \quad B = \left(\log\ xy\right) \quad \text{or} \quad C = \left(\log\left(e^x + e^y\right)\right);$$

or whether $\sqrt{26}$ is A = rational, B = algebraic or C = transcendental; or whether, at the battle of Waterloo, Napoleon A = won, B = lost or C = drew; or whether the city of Bahia is in A = Argentina, B = Brazil or C = Chile; and so on; in all these cases, the probability, the doubt to be measured, comes solely from the ignorance, uncertain knowledge or bad memory of the person questioned.

In every other respect, on the other hand, the situation is identical to that of the football pools: for the person who judges, for the person concerned with his state of doubt, there is no difference. It is sufficient to realize that a person could forecast the football results on a Sunday evening, when the facts are part of the past and are known to everybody (provided they are not known to him), or even a year later, provided that he then recollects them with something less than certainty.

5.5.8. The adoption of the proposed system in the case of multiple-choice questions would turn out to be instructive, in addition to the reasons that hold generally (i.e. learning how to express one's own opinion by translating it into numerical values), for the 'lesson' which would show how it is also *advantageous* (where sensible rather than stupid rules are in use) to strive for the greatest honesty and accuracy in expressing one's own doubts or lacunae. Conversely, stupid rules (like stupid laws) encourage dishonesty and reticence, encourage that complex of underhand and stupid actions which are euphemistically described by the phrase 'trying to be clever'; in our case, they encourage 'guessing'.

For the examiners too, it would be extremely useful to have precise information about those who 'know' (e.g. those who write down Antonio 00%, Brutus 00%, Caesar 100%), with the suspicion of 'guessing' now removed, and even more to be able to make a detailed analysis, on the basis of precise and meaningful data, of the frequency, intensity and nature of the doubts (possibly with a view to investigating their origin and suggesting ways of dealing with inadequacies in the teaching). In addition, they would be able to examine the degree of accuracy with which the evaluations are made (e.g. not simply using 50%–50% if there is uncertainty between two alternatives). In the case under consideration, there could, of course, be any number of alternatives whatsoever; in the examples above, we considered three for convenience, and in order to be able to retain the analogy with football results, and the possibility of imagining the situation as always representable in terms of Figure 5.3.

14 The betting approach, on the other hand, could not be used. Anyone in a state of some doubt would certainly lose against an opponent (e.g. the examiner) who knows the right answer.

5.5.9. *Applications in economics*. In the field of economics, the importance of probability is, in certain respects, greater than in any other field. Not only is uncertainty a dominant feature but the course of events is itself largely dependent on people's behaviour, which is itself determined, in a more or less unconscious and confused fashion, by evaluations and arguments of a probabilistic nature. It is, therefore, probability theory, in the broadest and most natural sense, that best aids understanding in this area (and not those fragments of the theory which never progress beyond the drawing of 'equally likely' balls from an urn, or 'stable' frequencies).

This point of view was presented in a clear and authoritative manner by T. Haavelmo in a celebrated critical speech delivered as president of the Econometric Society,[15] where he stated that previsions and evaluations of subjective probabilities '*are realities in the minds of people*' and that it was to be hoped that '*ways and means can and will he found to obtain actual measurements of such data*'

Another point, of particular importance for applications in operational research, is the possibility of making use of those evaluations of probability which represent a decision-maker's own opinions. For example, only the decision maker himself can say what probabilities he attributes to the different reactions of his most direct competitors to possible decisions of his. How, though, are we to interrogate him? Indirect approaches are necessary; questions about his preferences under some hypothetical sets of conditions should be posed in such a way as to provide, in turn, both a complete picture and a check of consistency. These are, however, expedients to make up for the lack of training in expressing oneself in terms of probabilities; the difficulty would not exist if such training became general practice.

Finally (in order not to dwell on too many other aspects[16]), there are important applications to the more theoretical field of econometric models. As E. Malinvaud says, in his treatise on statistical methods in econometrics,[17] the justification of the introduction of random models into econometrics rests, in his view, on an appeal to subjective probabilities, so that 'l'établissement d'une statistique subjectiviste qui reposerait sur le principe de Bayes' would be desirable (even though, in his opinion, research in this direction is, as yet, not sufficiently advanced to make a systematic application possible: on the other hand, there are those, for example A. Zellner,[18] who are attempting to do this).

5.6 Subsidiary Criteria for Evaluating Probabilities

Having analysed the meaning and the method of evaluating probabilities that a person might be led or compelled to make in order to sort out his ideas about what might occur, and to choose wisely any decision that has to be made, we are now in a position,

15 Trygve Haavelmo, *The rôle of the econometrician in the advancement of economic theory*, Presidential Address, Meeting of the Econometric Society, Philadelphia, 29 Dec. 1957; see *Econometrica*, **26** (1958), 351–357.
16 I have recently provided a wide ranging discussion of these topics (with a fairly mathematical treatment) in 'L'incertezza nell'economia', part I of: B. de Finetti and F. Emanuelli, *Economia delle assicurazioni*, Vol. XVI of *Trattato italiano di economia* (Edited by C. Arena and G. Del Vecchio), Utet, Torino (1967).
17 Edmond Malinvaud, *Méthodes statistiques dans l'économétrie*, Dunod, Paris (1964).
18 Arnold Zellner, *An Introduction to Bayesian Inference in Econometrics*, John Wiley & Sons (1970). It should be noted, however, that, although the treatment is Bayesian, the interpretation is not subjectivistic. The choice of the initial distribution does not derive from a case-by-case consideration of the factual circumstances, but from adopting once and for all a mathematically convenient form for each type of problem.

and are in fact obliged, to return to the essence of the problem of evaluation. We wish to discover whether the task of translating more or less vague impressions and opinions into numerical form could be facilitated by using some suitable subsidiary criterion. Fortunately, this turns out to be the case.

This happy circumstance derives, in general, from the observation that in many cases in the calculus of probability, under restrictions that are often very natural, certain probabilities, which are calculated on the basis of certain others, vary very little as one's evaluations of these other probabilities are varied. Consequently, even if the latter seem, to a given individual, rather vague, the former may very well appear to him capable of being evaluated with sufficient precision and confidence. As a brief aside on the question of interpersonal comparisons, we note that this explains why individuals often make practically identical judgments of prevision, even though they start off with very different opinions.

These general considerations will become clearer as we proceed further. For the time being, we restrict ourselves to illustrating the two subsidiary criteria which are of the greatest and most immediate interest: the first one we shall deal with in a reasonably detailed manner; the second, which, from a logical point of view, is based on material we shall meet much later on, is dealt with in a necessarily superficial way.

5.7 Partitions into Equally Probable Events

5.7.1. Every quantitative measurement is made both easier and more precise when it is possible to reduce it to a qualitative comparison. For example; it is much easier to say that A.N. Other has eaten $\frac{2}{9}$ (i.e. about $22 \cdot 2$ %) of a cake knowing that it was divided into 18 pieces, which could be taken as equal, and that he has eaten four of them, than to directly estimate that his portion was $22 \cdot 2$ % of the whole, undivided cake. In precisely the same way, it is obvious that if I judge n events of a partition to be equally likely, I cannot avoid attributing probability $p = 1/n$ to each of them (because the sum of the n terms, each equal to p, must be 1). Judgments of this kind arise rather frequently: it is sufficient that, given the present state of information, one finds oneself in a situation of *symmetry*. This will often, although not necessarily always, reduce to a state of *symmetry* regarding certain physical, or at any rate external, circumstances, which we regard as essential and relevant elements of our state of information.

When tossing a coin, we usually attribute the same probability $\frac{1}{2}$ to both faces and, similarly, probability $\frac{1}{6}$ to each of the six faces of a die. If we have n balls in an urn, we again, in general, attribute the same probability $1/n$ to any particular one of them being drawn: in this case, if we also know that m of the balls are white, we have no choice but to attribute probability m/n to the drawing of a white ball. This judgment of equiprobability (relative to a single toss, throw or drawing – this is not the place to consider more complicated cases) reflects a symmetric situation which is often made objectively precise by stating that the balls must be identical, the coin and the die perfect (physically symmetric) and so on. However, the criterion remains essentially subjective, because the choice, of a more or less arbitrary character, of those more or less objective requirements which are to be included in this concept of 'identical,' reflects the subjective distinction drawn by each individual of what is, and what is not, a circumstance that influences his opinion. It was necessary to point this out, in order to avoid giving the impression that in problems of this kind we are dealing with a different kind of probability; objective rather than subjective. It is true, however, that in this context opinions generally do coincide (although the agreement is less strong and unconditional than one would tend to think). Independently of all this, we can always talk about the case of equiprobability, provided we state (or take it as implicit) that this simply means that You (or the individual concerned) attribute the same values to the probabilities in question.

5.7.2. Returning to our examples, we observe that by means of these kinds of set-ups – it might be sufficient just to consider drawings from an urn – we can easily obtain a representation of events of any given probability (to be more precise, any rational m/n). For example, if one wants to get an idea of the magnitude of a probability expressed to two or three decimal places, for example expressed in percentage terms like 13% or 13·2%, it is sufficient to think of an urn with 100 balls, 13 of which are white (or 1000 balls, 130 or 132 of which are white). One can avoid talking about colours, and changing the percentage of white balls, by simply thinking of the balls as numbered consecutively (from 1 to 100 or 1 to 1000): this enables one to say – albeit in less suggestive language – that 13% is the probability of drawing a number not exceeding 13 (out of 100; or 130 out of 1000) and so on.

Using the 'representations' of this 'scale' one can – if the method seems easier – reduce the evaluation of any probability to comparison with cases of this kind, and forget all about both the betting approach and that in terms of losses. In order to translate into figures the probability – according to You – of striking oil by drilling at a given spot, it is sufficient that You decide how many balls, out of 1000, should be white, in order to obtain the same probability of drawing a white ball; if You think the number should be 131, this implies that You think the probability of striking oil is 13·1%.

It is convenient to express all this formally:

Theorem. *If the n events of a partition are considered as equally probable, the probability of each of them is 1/n, and the probability of an event which is the sum of m of them is m/n.*

The classical statement is that, under these conditions, *the probability is given by the ratio of the 'number of favourable cases'(m) to the 'total number of possible cases' (n).*

5.7.3. *Criterion of comparison* (or 'third criterion' – following the two in Chapter 3, Section 3.3). Having at one's disposal a model of a partition into n events, which are judged equally probable (e.g. an urn), *the probability of any event E can be evaluated, by comparison with events composed of sums of events of the partition, with an error of less than* $1/n$. In fact, if E_m and E_{m+1} are sums of m and $m+1$ events, each of probability $1/n$, and if one judges $\mathbf{P}(E_m) \leqslant \mathbf{P}(E) \leqslant \mathbf{P}(E_{m+1})$, then $m/n \leqslant \mathbf{P}(E) \leqslant (m+1)/n$. In order to make the comparison operational, it is sufficient to express it by saying that You would rather receive one lira if E occurs than one lira if E_m occurs, but vice versa if the comparison is made with E_{m+1}. In this way, its subjective nature is clear; it remains somewhat in the shade when we speak of 'comparison' in the abstract, with no precise meaning.

There are many points, both historical and critical, that one could raise at this juncture, but they would require overlong, and in part untimely, digressions: they will be considered instead at the end of the Appendix.

Let us just say something, however, in order to make the above a little more precise, at least in its essential features. Evaluations made on the grounds of symmetry are generally accepted as a basis for problems concerning games, drawings from an urn, lotteries, dice and so on, and one often regards as 'equally probable cases' certain outcomes which are 'combined' (like the 6^{10} possible sequences obtained by tossing a die 10 times, or the 90 !/85! possible sets of five numbers on the Lottery, or the 90! permutations in a drawing of all the 90 numbers at Bingo etc.), rather than elementary (like the score obtained at the next throw of a given die, or the number 'drawn first' at a given Lottery wheel next Saturday). Recall the remarks made in Chapter 4, 4.10.3, which are relevant to this procedure.

5.7.4. We note, however, that it is not just in examples of games of chance that considerations of symmetry can act as a guide but, in fact, in any practical problem whatsoever. For example: if we consider the maximum annual temperature (at a given location) in three consecutive years, then it can either:

- increase (type 1–2–3, where 1, 2 and 3 schematically denote the three temperatures in increasing order),
- decrease (3–2–1),
- be maximal in the middle year (types 1–3–2 and 2–3–1), or be minimal in the middle year (types 2–1–3 and 3–1–2).

Now, whatever one's evaluations of the probabilities of more or less high summer temperatures might be, under certain conditions it may very well be natural for us to attribute the same probability ($\frac{1}{6}$) to each of the possible cases.

Example A. *Increases and decreases in agricultural production*. This is a (true!) example of a fallacious analysis, based on the observation that, by comparing agricultural production in successive years, the numbers of *inversions of trend* (i.e. the number of times in which an increase was followed by a decrease, or vice versa) was about twice the number of *permanences* (i.e. repetitions of an increase or of a decrease). An agricultural expert argued that rich and poor crops alternate, and it required a statistician to point out the mistake (the numbers are in agreement with what we have just seen above).

Example B. *Breaking an existing record*. In connection with temperatures, agricultural production, or even the results in an annual competition, for example the winning throw in the national discus championships (assuming the given hypotheses continue to hold: i.e. there exists no reason to expect an improvement due to better training, more participants etc.), one can pose the following sorts of problems: what is the probability that in the nth year (of the competition, of keeping temperature records etc.) a new record is set up? (Ans. $1/n$); that the record (set in the first year) be broken for the first time? (Ans. $1/n(n-1)$); that the previous record had stood for h years ($h=n-1$, $n-2,\ldots, 3, 2, 1$)? (Ans. $1/(n-1)$ for any h); what is the prevision of the number of times the record was broken in the first n years? (Ans. $\sum(1/h)(1 \leqslant h \leqslant n) '' \log n$); and what is the prevision of the number of years that the record lasts until the next improvement? (Ans. $+\infty$). As an exercise, verify these answers and pose yourself some further problems (these are easy to find, although not always easy to solve).

5.8 The Prevision of a Frequency

5.8.1. When considering events $E_1, E_2,\ldots, E_n$, it may happen that we know with certainty what the number of successes $Y=E_1+E_2+\ldots+E_n$ (or, equivalently, the frequency Y/n) must be: $Y=y$, say; that is $Y/n=y/n$. Clearly (see Chapter 3, 3.10.3), the sum of the

$p_i = \mathbf{P}(E_i)$ must be equal to y (i.e. their arithmetic mean must be equal to y/n); in particular, if the E_i are judged to be equally probable, $p_i = p$, then we must have $p = y/n$ (the probability equal to the known frequency: for $y = 1$ we have the case of a partition, as considered previously). However, even if the frequency is not known with certainty, the relation still holds if we substitute the prevision of the frequency: *the sum of the probabilities must equal the prevision of the number of successes*. In other words, dividing by n, we have

Theorem. *The arithmetic mean of the probabilities must equal the prevision of the frequency:*

$$\left(p_1 + p_2 + \ldots + p_n\right)/n = \mathbf{P}\left(Y/n\right) = \mathbf{P}\left(Y\right)/n. \tag{5.11}$$

In particular, *if the E_i are judged equally probable, $p_i = p$, we have $p = \mathbf{P}(Y/n) = \mathbf{P}(Y)/n$: the probability* (*common to all the events*) *is equal to the prevision of the frequency.*

5.8.2. In order that correct use be made of this theorem, we must make very clear that it is essentially trivial: otherwise, we run the risk of goodness knows what being read into it. Observe first of all that the E_i can be any events whatsoever, however diverse, so long as the number of successes is given by addition: for example success in an examination, a victory for one's favourite football team, finding a traffic-light green, throwing a double six at dice, and anything else, however dissimilar. The 'theorem' is an identity: it imposes no restrictions, apart from informing us that the same thing, written in two ways, remains one and the same thing (rather like the sum of a double-entry table, which can be taken either over rows or over columns).

Well then: *it is in this very thing – and in nothing else – that the value of any theorem in the calculus of probability lies, and it cannot be otherwise. It is to tell us whether, in making the same evaluation in two different ways, we arrive at different conclusions, and, in this case, to invite us to think again and to rectify the situation by modifying one or the other.*

There is no *unique way* of doing this: we do not begin with one side already fixed and the other to be 'deduced'. Instead, we have on both sides evaluations that should agree, and which must be modified if they do not. How should this be done? Generally speaking, one of the evaluations usually seems to be more immediate, so one is inclined to look for a modification of the other; however, one should be open-minded about it, since appearance might well be only appearance.

5.8.3. Turning now to our particular case, You might find that the probabilities which You have evaluated, when added together give, for instance, a value which is greater than the number of successes, $\mathbf{P}(Y)$, which, in prevision, seems to You reasonable. You must then ask yourself: 'have I given the p_i values which are on average too large, or are the values which I thought of for the number of successes Y (or the frequency Y/n) too low?' It is fairly difficult to answer this if the events are rather disparate, but when they are more alike, and especially if we know the frequency of other similar events, which have already been observed, it often happens that one places greater confidence in prevision of the future frequency (under the assumption that it will remain close to that previously observed).

Why is this so? The answer to this cannot be given at present (see Chapter 11) but, even without going into the whys and wherefores, the idea that there is a degree of stability in the frequency of occurrence of events usually grouped together as 'similar' is one which seems quite intuitive to most people. At the present time, this phenomenon may even be somewhat exaggerated as a result of overly simple and rigid formulations current among many statisticians. However, it rests on a very real foundation, since this is how things appear, even to the naïve layman (who, for example, is really surprised if in a given period certain phenomena re-occur with an unusual frequency). Let us accept things as they are.

As a particular case, suppose the events under consideration are so similar that one judges them equally probable: it will turn out that their probability p will be evaluated on the basis of a frequency f observed among similar events in the past, and that p will be close to f. Notice that in this case the evaluation is based not only on the prevision of a frequency, *but also requires a judgment of equal probabilities*.[19]

5.8.4. *Some examples*. Statistics show that the percentage (or frequency) of males among live births is always about 51·7% (hence, a few more males than females); that, according to the Italian tabulations for 1950–1953 and 1954–1957, respectively, the percentages of deaths in the first year were 6·75% and 5·49% for males, 5·88% and 4·67% for females; that the overall annual percentage of deaths in Italy in 1960 attributable to cancer was 1·51%, but broken down into age groups it was

Age:	0–5	5–25	25–55	55–75	over 75
%	0·013	0·009	0·078	0·524	1·131

and into regions (not distinguishing age groups) it varied from 0·220% in Liguria, 0·210% in Tuscany, to 0·089 % in Puglia and 0·073% in Basilicata and Calabria. To change the subject completely, statistics also show that the results of championship football matches are distributed (in terms of home fixtures) as 50% wins, 30% draws, 20% defeats.

Thinking of such frequencies as stable, we could adopt them universally as probabilities for any similar events, or future cases; or, at least, we could evaluate the probabilities of individual cases in such a way as to make them compatible, in arithmetic mean, to these frequencies. However

5.8.5. *The need for realism*. Even though we have expressed our previous considerations with a certain amount of caution (which itself might appear overdone and unnecessary to anyone accustomed to a different approach), it is necessary, in fact, to go still further and provide additional warnings in emphasis of that caution. We seek to reduce everything to three questions (and in answering these we shall delve deeper).

5.8.6. *The first question*: are we justified, in real applications, in attributing the same probability to all the events of a given type? This question is equally relevant to both of the subsidiary criteria; that is symmetric partitions and frequencies. However, we must first point out that it is meaningless unless we bear in mind that the probability is not an *external fact* relating to the event, but, instead, relating to your state of information regarding the event, and the previsions which You derive from this state of information. If You know the innate qualities, the past records and the degree of preparedness of

19 This is often overlooked: if, for example, one speaks of 'the probability of a newly born baby being a boy', it is not made explicit that one is dealing with one unspecified event out of infinitely many ill-defined events, each of which is understood to be equally probable.

every student, your evaluation of the probability of passing an examination will vary from student to student. Even with all this background information, however, if You only know the students by sight (i.e. are ignorant of the name of any given student) and are asked name by name to give the probabilities, then your evaluations will all be equal (the same would be true if knowledge by sight or by name were the other way around). In much the same way, your probabilities for the results of different football matches on a given day will be different if You know the merits of the respective teams, and are in a position to express a prevision for each match. However, if You had to fill in a pools coupon knowing what the matches were, but not the order in which they were listed, You could only assign the same probabilities to them all (the averages of those for the individual matches). For example, it might be 40–20–40 if in about half of the matches the away teams are first-rate and favourite to win, or, if You had to fill it in without even knowing what matches were being played that day, You might adopt a standard average probability like 50–30–20. Even in the near legendary case of drawings from an urn, for instance drawing one from among 90 identical balls (numbered from 1 to 90), equality would not necessarily hold if one knew the position of each ball in the urn at the instant before the drawing took place (You might know, or believe, that the person drawing the ball has a habit of drawing from the top, or from the left-hand side, and so on, and taking this into account might lead You to judge the probabilities to be different).

5.8.7. *The second question*: if I wish to make use of a frequency, which one should I base my opinions on? Given an event E in which You are interested, there are usually several classes of events already observed, which are, in different ways, more or less directly similar to yours, but with each class providing a different frequency: the choice is largely arbitrary.

Let us consider, for example, the problem of life insurance for a certain individual (for simplicity, suppose it is a question of a capital sum being provided if he dies within a year). How shall we determine the 'premium'; that is the probability of his death within the year (not taking into account any 'extras' – for expenses, etc.). We could check the statistics of the deaths of individuals in the same country (or region, county, city, district, etc.), of the same age (sex, class, etc.), having the same profession (income, degree, etc.), of similar constitution (height, weight, etc.), same name or initial of surname, or house number, or born in the same month, and so on: or we could group together some number (large or small) of these sort of characteristics, or any others. Each grouping will yield a different frequency, and this forces us to adopt a reasoned evaluation rather than a mechanical one; one which takes into account those classifications which it appears most reasonable to assume related to the phenomenon (for instance, age), and not the others (like the person's name). What is 'reasonable' depends not only on whether and *to what extent* this or that circumstance influences the phenomenon, but also on *how* it has an influence. If, for example, it appears reasonable (on general grounds, and on the basis of corroborative evidence) to think that the death rate increases with age (once we pass childhood), one would be inclined to stick to this when evaluating the probabilities of death in the immediate future, even for those countries for which the most up-to-date statistics would show oscillations from year to year. One would appeal to some sort of *smoothing* procedure, in an attempt to preserve the general outline, which is considered significant, and to eliminate what are thought of as misleading perturbations.

Finally, one is always faced with the aspect we have already spoken of; that of individual differences (which the insurance companies take into account through the results of the medical examination).

This is a general situation and examples are easy to come by. We shall consider just one other, which shows how meaningful variations in frequency, for appropriately chosen subdivisions, can occur, even in those situations where it appears to be more correct to view the probability as invariant with respect to any of the background circumstances. Given that the frequency of males among new-born babies is almost completely invariant over time, races, or countries, there would seem to be no possibility of differentiating probabilities on the basis of frequency statistics selected according to some factor or other. On the contrary, the research of Gini (using Geissler's data on Saxony, 1876–1885) brought to light a differentiation on the basis of families: there were too many families with an excess of either males or females for it to be 'attributed to chance'.[20] Presumably, one could always find some differentiation if one could succeed in finding appropriate factors on which to base a classification. On the other hand, clearly, as a kind of converse, for those for whom every attempt at picking out significant factors is unsuccessful, every combination of cases automatically appears uniform (even if this is not so for those who do succeed in picking out such factors).

5.8.8. *The third question*: are we justified in expecting frequencies to be stable? The remarks concerning the second question have already led us to consider the differences in frequencies when we refer to subgroups (e.g. in questions concerning people, age groups, regional groupings, etc.), not to mention individual differences (as discussed in the first question). The stability of all these frequencies is an hypothesis, incompatible with the variability exhibited by the overall composition in terms of subgroupings (e.g. dividing the population according to age, region, etc.). In actual fact, in practice, we can usually assume that the overall composition changes rather slowly and, therefore, that the incompatibility is not obvious over a short time period: from a logical point of view, however (and in some cases from a practical one too), the objection is completely valid. On the other hand, even if we leave all this out of consideration, there may be – and there usually are – causes of variation resulting from the evolution of the situation itself. For example, if we consider mortality, there has been great progress in sanitation, medicine, general living standards and so on, as a result of which mortality has progressively declined significantly (this can be seen even from the snippets of data we reported above, relating to very close time periods like 1950–1953, 1954–1957). It might, therefore, appear reasonable, in evaluating a future probability, to extrapolate the rate of improvement rather than base oneself on the hypothesis of the preservation of the present level.[21] In any case, the force of the 'stability of frequencies' as a probabilistic or statistical principle is completely illusory, and without solid foundation.

Similar considerations apply, of course, in other fields. We could add the obvious examples of frequencies of car accidents, and similar matters in connection with

---

20 C. Gini, *Il sesso dal punto di vista statistico*, Sandron (1908), Ch. X, 'La variabilità individuate nella tendenza a produrre i due sessi' (pp. 371–393). I do not know whether there has been any more recent research confirming these results: in any case, it is the argument which is of interest here rather than the facts.

21 Questions of this nature have been discussed with particular reference to the actuarial field; see R.D. Clarke, 'The concept of probability', *J. Inst. Actuaries*, 1954.

technical or economic development. In the case of football, the changing character of systems of play, tactics, and many other things, may alter the influence of playing at home or away, and therefore the probabilities of the three results. In addition, even without changes of this kind, frequencies will be altered if the imbalance between 'top' teams and 'bottom' teams is altered.[22]

5.9 Frequency and 'Wisdom after the Event'

5.9.1. Let us repeat an earlier remark, whose function is to prevent a certain confusion; one which we have already warned against, but to which we are particularly vulnerable in the case of previsions of frequencies.

Previsions are not predictions, so there is no point in comparing the previsions with the results in order to discuss whether the former have been 'confirmed' or 'contradicted', as if it made sense, being 'wise after the event', to ask whether they were 'right' or 'wrong'. For frequencies, as for everything else, it is a question of prevision not prediction. It is a question of previsions made in the light of a given state of information; these cannot be judged in the light of one's 'wisdom after the event', when the state of information is a different one (indeed, for the given prevision, the latter information is complete: the uncertainty, the evaluation of which was the subject under discussion no longer exists). Only if one came to realize that there were inadequacies in the analysis and use of the original state of information, which one should have been aware of at that time (like errors in calculation, oversights which one noticed soon after, etc.), would it be permissible to talk of 'mistakes' in making a prevision.

Any reluctance one feels in accepting these obvious explanations is possibly accounted for by their seeming to preclude any possibility of taking past experiences into account when thinking about the future. This is not so, however: the latter is rather different from 'correcting previous evaluations'. One must emphasize that this phrase is wrong, even though it may only be a confused way of expressing an actual need. It is not, however, a harmless inaccuracy: in actual fact, it distorts the basic question, and generates a tangle of confusions and obscurities.

This should be made absolutely clear. If, on the basis of observations, and, in particular, observed frequencies, one formulates new and different previsions for future events, or for events whose outcome is unknown, it *is not* a question of a *correction*. It is simply a question of a new evaluation, *cohering with the previous one*, and making use – by means of Bayes's theorem – of the new results which enrich one's state of information, drawing out of this the evaluations *corresponding to this new state of information*. For the person making them (You, me, some other individual), these evaluations are as correct *now*, as were, and are, the preceding ones, thought of *then*. There is no contradiction in saying that my watch is correct because it now says 10.05 p.m., and that it was also correct four hours ago, although it then said 6.05 p.m.

5.9.2. Discussions and refinements of this kind, which might seem rather pointless when made in the abstract and reduced to mere phrases, are not only of genuine

---

22 If, for example, one half of the teams were so much stronger than the others that they beat them with certainty, then about half the matches would have the assigned result; if the frequencies 50–30–20 were retained for the other half, we would have, overall, the frequencies 50–15–35 (the averages of 50–00–50 and 50–30–20).

relevance to the conceptual and mathematical construction of the theory of probability, but they also have implications which demand the attention of everyone; even those not interested in topics of this nature.

The meaning which attaches to statements about 'being wise after the event' does not seem to correspond in a unique way to attitudes either for or against the considerations just made. It is often both different and opposite. This happens when the sentence is uttered as a reproach to someone who belatedly admits that 'he was wrong' – as if to tell him '*tu l'as voulu…*' ['it is your own fault…']. It is conceivable that in some situations this reproach is justified: one often makes mistakes through lack of concentration, or because one was unable to resist the temptation, although fully aware of being in the wrong.

However, the reproach is often made when there is no fault – apart from that of failing to be a prophet. Judgment *by results*, the notion that someone's merit should be measured in terms of his successes, is often passed off as 'realism': to dwell upon the *ifs* and *buts* is considered meaningless. Of course, it is meaningless as far as the facts are concerned; no-one doubts that these cannot be reversed or modified by any *ifs* and *buts*. The facts themselves are not open to question, but when we turn to *judgments based on those facts, evaluations of personal responsibility, appreciation or criticism of someone's actions*, it is a different matter. In these matters, it is by no means true that the facts provide any definite answers; in fact, they provide no answers at all. Their only value might be in helping one towards a better understanding of the range of ifs and buts. It is precisely these which allow one to judge someone's actions in the one way that makes any sense: that is taking into account, moment by moment, *the context, the situation and the state of information in which the actions took place*.

It would perhaps be overstating the case to suggest, for these reasons, the removal of any distinction between – let us say – being found guilty of murder and of attempted murder. It could happen that 'missing' killing someone was evidence of a lesser intention of doing so; but if everything hinges on a miraculous piece of surgery, how is the offence in any way less serious, or the culprit more deserving of leniency? Anyway, since legal matters are somewhat of a mystery to me, I do not wish to pursue the question.

Something that can be criticized with more certainty is what seems to me the deplorable habit of picking on someone as a scapegoat when something goes wrong. Apart from being unfair, the practice encourages people to avoid taking on responsibility, so that one gets the worst of all worlds. Those who acted loyally, in a sensible manner, cannot be reproached if, by chance, the outcome was unfavourable; those who blundered (in an honest fashion) are advised to learn from the experience and take more care in the future. In contrast, those who had not done everything possible, in terms of organization, control and efficiency, to reduce the risk of unfavourable outcomes are punished – whether or not anyone was responsible.

To set against such stupidity, there is an alternative practice, which can be taken as an example of the beneficial effect of a mode of thinking based on operational research. It was brought to my attention by Pasquale Saraceno,[23] and is established practice in the industrial group of which he is one of the leading figures. When examining the actions of the various companies, and especially those with unfavourable outcomes, the analysis is based on drawing a distinction between that which could and should have been foreseen, on the

23 *Translators' note.* Italian economist; former head of the I.R.I. (the state controlled Institute for Industrial Reconstruction).

basis of the information at hand, and that which could not possibly have been foreseen. This sort of calm criticism and self-examination is undoubtedly what is required in order to encourage a sense of responsibility in a climate of honesty and mutual confidence.

5.9.3. The remarks above were made in order to underline the importance of breaking away from these destructive hangovers of the confusion between prediction and prevision: this is important from a general – one might even say *moral* – point of view. Let us turn to a technical aspect of the problem, which should help to remove such confusion. I say 'should', because I know only too well that such errors (these, it seems, more than most) are difficult to eliminate; like the Hydra with a thousand heads. Were it not for this, I would have simply said, as it seems to me, that each objection raised is decisive in itself, and should be sufficient.

In order to combat the idea that the influence of the facts, or, to be more precise, of information regarding the facts, on prevision should be interpreted as a mechanism for refutation and correction (and also to point out the inadequacy and awkwardness of language which gives this impression), we observed that the 'new' opinion, far from being new, was already contained in the 'old', which, far from being refuted, was used when we took over as the 'new' the opinion it had already provided as appropriate for such an eventuality (as for any other possible outcome).

Let us note at this point that such an 'opinion implicitly contained' in the initial one, and already provided for such a contingency, is integrated with it to such an extent that it practically lends itself to being used *without even the occurrence of the facts under consideration*.

5.9.4. The '*device of imaginary observations*', put forward, in particular, by Good (1950), is a method of evaluating probabilities, and, as such, deserves mention in the present chapter. It is a device that is particularly useful for evaluating very small probabilities, and which is more accurate in this context than the direct approach. A simple example will suffice to make the notion clear.

A person claims that he is able to guess in which hand you are concealing a certain object; You do not believe him. If You are invited to be more precise and say what probability p You attach to the possibility that he really can do what he says, You reply 'very small', but are not really in a position to sort out the different implications of saying 10^{-2}, or 10^{-10}, or some other value. Then, according to Good, one can do better by reformulating the question in the following way. Imagine that You put him to the test, and that he guesses correctly three times in a row, or ten times, or fifty times, …; after how many consecutive correct guesses would You consider as equally likely (probabilities $\frac{1}{2}$ and $\frac{1}{2}$) the two possibilities that either his claims are justified, or that he has guessed correctly by chance?

It is easily seen that at each trial where he guesses correctly the probability ratio in favour of his claims doubles (likelihood ratio $1:\frac{1}{2} = 2:1$); after n such trials it is 2^n. If, after n trials, the ratio of the probabilities has become $\frac{1}{2}:\frac{1}{2} = 1$, it must mean that initially it was given by $p:\tilde{p} = 2^{-n}$; in other words, we have approximately, $p = (\frac{1}{2})^n = 10^{-n\log 2} \simeq 10^{-(0.3)n}$. For example: if $n = 10$, we have $p = 10^{-3} = 0.1\%$; if $n = 30$, $p = 10^{-9}$; if $n = 50$, $p = 10^{-15}$.

There is no doubt that, with this interpretation available, a comparison between the meaning of answers such as $p = 10^{-3}$,or $p = 10^{-100}$, is no longer unattainable (although a certain vagueness or unfamiliarity is inherent in questions of this kind, and cannot be removed altogether; any method or device of this kind is intended as an *aid*, not a panacea:

once one gets to a certain stage, there is nothing to do but try to sharpen the feeling for numerical values of probabilities, including the very small ones).

The conclusion regarding the principle of this method seems to derive further support, psychologically speaking, from a consideration of the paradoxical – I would even say grotesque – position that a contrary point of view leads one into. Its formulation would have to be along the following lines (any attempt at spicing it up in order to increase its paradoxical and mind-bending flavour would only spoil it):

'*My initial evaluation was* $p' = (\frac{1}{2})^n$;

'*it was based on consideration of a hypothetical possibility; that of a succession of n experiments, in which the person claiming to be able to guess obtains successes on every trial, and on my reaction to such a hypothetical result, consisting precisely in the fact that my final evaluation would then have been* $p'' = \frac{1}{2}\tilde{p}$;

'*now, the eventuality considered as the hypothesis has actually occurred, and my reaction has been precisely the presupposed one, therefore…*

'*the initial evaluation, which was, and still is, a logical consequence of these assumptions* (*actual or hypothetical*)*… WAS FALSE*!'.

5.10 Some Warnings

5.10.1. It is necessary to point out a number of pitfalls. Although it is premature to talk about the dangers before we understand their causes, some pointers must be given in order to guard against the doubts and distortions that might get mixed up with what we have said concerning the evaluation of probabilities, giving rise to confused and contradictory notions.

The following remarks should, in one sense, be unnecessary. All the dangers have already been mentioned and the details already given at the appropriate time would be sufficient to render these additional comments superfluous, if – and this is the difficulty – they remained firmly implanted in one's mind, together with all their ramifications, and with such clarity that any dangers reappearing, in whatever disguise, could be dealt with just as effectively as when they were first encountered. This not being the case, it is preferable, and perhaps necessary, to repeat ourselves; to go over the details mentioned above, in their different variants and versions, pointing out the many forms the dangers may assume. (There are such a number of them that perhaps some, even important ones, will be overlooked; hopefully, though, the pattern-book of objections and counter-objections will be sufficiently representative for the reader to be able to answer, by analogy, possible objections not covered, by means of suitable counter-objections.)

5.10.2. It might be argued that the kind of problems we have considered in this treatment, and for which we have discussed the appropriate methods of evaluating probabilities, are outside the 'true' ambit of the calculus of probability, or, at most, they constitute a small and specialized part of it.

The arguments put forward will be, by and large, the standard ones; however, if they are given with reference to physics, for instance, they may appear novel, or at least more substantial and difficult to refute.

There are cases where the probabilities in physics are given by combinatorial arguments, in accordance with the 'classical' idea of 'equally likely cases'; that is, they are given by the Maxwell–Boltzmann, Bose–Einstein, Fermi–Dirac 'statistics' (to use the jargon of physicists): for further details, see Chapter 10, Section 10.3. Who can argue in this case that we are dealing with a probability whose value is objectively determined by '*a priori*' considerations? It is precisely this example (as Feller observed, vol. I, pp. 5 and 21) which shows how fallacious any *a priori* conclusion would be: nobody could have foreseen that the computation of 'equally likely cases' had to be carried out using completely different methods in problems where different 'statistics' apply (and the explanation only came later, through the distinction between particles with integer or semi-integer spin).

5.10.3. Everyone will probably agree, therefore, that it makes no sense to be willing to deduce properties of phenomena, or previsions regarding their outcomes, basing oneself solely on superficial, preconceived ideas. The confirmation of experience is required, and this, certainly, leads on to an objective conclusion. One might well say that, for the physicist, probability coincides with frequency.

And this statement is, in a certain sense, true. However, this form of expression is completely wrong from a conceptual point of view, even if at first sight it presents no difficulties.

Let us swiftly demolish, one by one, the main arguments put forward with the intention of transforming probability from being subjective to being objective, by means of more or less overt confusion or connection of the notion with that of frequency.

5.10.4. Firstly, we present an objection frequently raised against the notion of the *probability of an individual event*: either this event occurs or it does not, and therefore it either has probability *one* or *zero*; it makes no sense to attribute to it an intermediate probability p. I accept this argument completely, in that it refers to an objective probability p: but I observe that the same argument holds even in cases where my opponent forgets that it does – when he says that in n 'individual cases' there is an objective meaning to p because np of them will occur. This is not true: either *zero*, or *one*, or *two*,..., or *all* n of them occurs, and the objective probability (if one prefers to use this term as a useless and misleading synonym for frequency) is one of the $n+1$ values $0, 1/n, 2/n, \ldots, h/n, \ldots, (n-1)/n, 1$, although it is not known which one.

It is only in a subjective sense that it makes sense to speak of p, as the arithmetic mean of these $n+1$ possible values, *taking as weights the subjective probabilities of the single frequencies* (*still* '*individual cases*'!).

5.10.5. It might be objected that in many cases (those to which an opponent would limit himself) the probability is concentrated near a certain frequency p, which could be defined as objective probability. But here, and in every case in which something 'very probable' is said to be 'practically certain' (or even 'certain', for the sake of brevity), and, symmetrically, something 'very improbable' is said to be 'practically impossible' (or even 'impossible'), an *either–or* must be clearly established. In fact, such sentences can either say something obvious, with which one has no choice but to agree, or, alternatively, they can completely falsify the meaning of things. The field of probability and statistics is then transformed into a Tower of Babel, in which only the most naïve amateur claims to understand what he says and hears, and this because, in a language devoid of convention, the fundamental distinctions between what is certain and what is not, and between what is impossible and what is not, are abolished. Certainty and impossibility then become confused with high or low degrees of a subjective probability, which is itself denied precisely by this falsification of the language.

On the contrary, the preservation of a clear, terse distinction between certainty and uncertainty, impossibility and possibility, is the unique and essential precondition for making *meaningful* statements (which could be either right or wrong), whereas the alternative transforms every sentence into a nonsense.

5.10.6. We have already made abstract reference to this confusion (Section 5.2.3), so let us confine ourselves here to an illustration in the context of physics (with the warning that we are anticipating things to come for the purpose of preventive therapy; later, Chapter 7, we will be concerned with the true meaning of 'laws of large numbers' and suchlike).

It cannot be denied that two different explanations of the same phenomenon may turn out to be indistinguishable in practice; particularly when one explanation is deterministic and the other probabilistic. One thinks immediately of the diffusion of heat, or any other similar phenomenon, which can either be considered in terms of a differential equation, describing the continuous development of the phenomenon in a manner governed precisely by deterministic laws, or as a random process in which elementary phenomena occur in a nondeterministic way, but such that there is a high probability of the phenomenon developing at a macroscopic level in a manner practically identical to that indicated by the deterministic theory.

However, this in no way implies that the two explanations are similar, and even less that they are the same, or substitutable. On the contrary, they are exact opposites; diametrically opposed and absolutely incompatible. The deterministic explanation makes certain assumptions which preclude any departure from predetermined behaviour. Any similar explanation, albeit less rigid, which laid down that some conclusion was compulsory and certain, would, at the very least, require some sort of self-regulatory mechanism, some sort of '*feedback*'. The probabilistic explanation makes no assumptions of this kind: it states nothing other than that everything is possible. If it *appears* to state something more, it is only because such a statement, which may seem quite precise, corresponds to a property common to 'almost all' the possible cases.

A probabilistic explanation of the diffusion of heat *must* take into account the fact that heat could accidentally move from a cold body to a warmer one, making the former even colder, the latter even warmer (in Jeans' example: water being frozen rather than boiled when put on the stove). That this is very improbable is merely due to the fact that the 'unordered' possibilities (heat equally diffused) are far more numerous than the 'ordered' possibilities (all the heat in one direction), and not because the former enjoy some special status.

To rule out the possibility of those cases which seem 'exceptional', in no way *improves* the probabilistic explanation, by somehow making it simpler, or more scientific: on the contrary, it negates it. Acceptance of the probabilistic explanation has the following implications: it means that what we state about the phenomenon must not be regarded as necessary, but, instead, must be attributed to 'chance', and, hence, regarded as only approximate and probable. It means that one must regard it as essential to deny the existence of certain and exact laws which are obeyed only apparently; it means that one must consider as necessary the possibility of studying departures from any rigid law, fluctuations, the effects of discontinuities (the *shot effect*), and all that a cursory identification with a different form of explanation would sweep away without a thought.

5.10.7. What we just said is itself open to misinterpretation. It would be a mistake to infer that an explanation based on a 'tendency to disorder' takes care of every application of probabilistic concepts and not merely the particular example given above.

'Chance' (if we can adopt this convenient terminology as a summary of complicated and uncertain factors without its being taken too seriously) plays a no less important rôle in biological and social processes, where the outcome depends on highly ordered and organized structures, like chromosomes, cells and human beings (and also in physics, in processes like crystallization).

The following needs to be said in order to disprove the thesis which considers a levelling down into a debased chaos (entropic death) to be an inevitable consequence of the validity of this or that 'law' of probability. *The calculus of probability can say absolutely nothing about reality; in the same way as reality, and all sciences concerned with it, can say nothing about the calculus of probability*. The latter is valid whatever use one makes of it, no matter how, no matter where. One can express in terms of it any opinion whatsoever, no matter how 'reasonable' or otherwise, and the consequences will be reasonable, or not, for me, for You, or anyone, according to the reasonableness of the original opinions of the individual using the calculus. As with the logic of certainty, the logic of the probable adds nothing of its own: it merely helps one to see the implications contained in what has gone before (either in terms of having accepted certain facts, or having evaluated degrees of belief in them, respectively).

Physics can make greater or lesser use of the calculus of probability, but the relationship between the two is simply the relationship between a certain field of research, which remains itself, no matter what tools it uses, and a logical tool, unconditionally valid, which remains itself, whatever use is made of it, in whatever field.

5.10.8. Let us return to the necessity of avoiding the dangers implicit in attempts to confuse certainty with 'high probability'. We have to stress this point because these attempts assume many forms and are always dangerous. In one sentence: to make a mistake of this kind leaves one inevitably faced with all sorts of fallacious arguments and contradictions whenever an attempt is made to state, on the basis of probabilistic considerations, that something must occur, or that its occurrence confirms or disproves some probabilistic assumptions.

From such a point of view, the calculus of probability seems to be regarded, more or less explicitly, as a nothingness; saying nothing when the probabilities in question are intermediate in value, but capable of miraculous transformation into a warrantor of absolute truths when the probabilities are very large or very small, since, in these cases, the difference can be ignored and one can simply say that something is true or false. One thus has a mechanism that is considered to be useless when it says that which it is capable of saying, and wishes to say, but is blindly trusted when the things one wants to make it say are not the things it does say or could say.

5.10.9. We present three examples of this form of observation.

First example. The statement that '*an event of small probability does not occur*' is sometimes made, under the heading of 'Cournot's principle' (Section 5.2.3). A kind of corollary or special case of this is referred to as the '*empirical law of chance*' (meaning that frequency and probability actually behave in many cases according to the 'law of large numbers').

Second example. In accordance with the identification of small probability with impossibility, Neyman finds a contradiction in the behaviour of an individual who travels by aeroplane and at the same time takes out insurance. If he considers it possible to have an accident, why does he travel? If he does not, why does he insure himself? The paradox here specifically relates to 'decision theory', which, in the restricted sense to which it is often reduced by 'objectivist' statisticians, considers only the question 'what decision is appropriate given the *accepted hypothesis*' and not 'what decision is appropriate in the *given state of uncertainty*'.

Third example. Again, in this context (of 'objectivist statisticians'), one aspires to '*accept*' or '*reject*' an hypothesis on the basis of an experiment, instead of considering how its outcome modifies the initial probabilities (which *one wants to do without*!) in order to give the final probabilities (which therefore cannot be obtained!). Here the absurdity reaches new heights, because it cannot even be claimed that 'accept' and 'reject' correspond to the minimal requirement of the probabilities being large or small. The use of these two words is a meaningless convention; an apparent attempt to answer a question by disregarding everything that makes it a meaningful question in the first place.

It is as if, in comparing two weights, we were to decide upon which was the heavier by choosing the one which tilted the balance to its side, without taking into account, and, indeed, refusing to consider it legitimate to take into account, any difference between the arms of the balance, even knowing that the difference could be considerable.

5.11 Determinism, Indeterminism, and other 'Isms'

5.11.1. Continuing with the same theme, there is a clear philosophical point to be made. It derives from the strange fact that precisely the same disposition to accept an objective probability is often justified in two completely opposite ways.

For some people, the ideal instrument for producing an objective probability with value p would be a totally invariable device, working under strictly unchangeable conditions, and for which the tendency to produce successes with frequency p would be a 'built-in property', or, more specifically, a '*dispositional property*' (following, for example, Hacking). Any perturbations would result in a deviation from the desired result; that is, from the realization of a frequency close to p.

For others, whole-hearted determinists, any such device could but yield the same result; always successes or always failures. The fact that both successes and failures occur implies that there exists something causing perturbations. In general, it is assumed that there are a great number of small, accidental, causal factors, which are largely unknown. The fact that the frequency is expected to be around p would be an effect of the combined and random actions of these causal factors (following, for example, Paul Lévy).

So far as the subjectivistic conception is concerned, it has the advantage and, indeed, the preoccupation of remaining outside of disputes of this kind. The thing that really matters, and which justifies, in fact requires, our arguing on the basis of probabilistic logic, is the impossible nature of the situation in which we find ourselves when we attempt to foresee a given outcome with certainty. This is so whatever the reason: whether it be ignorance of certain deterministic laws; or the nonexistence of such laws; or an inability to perform the requisite calculations even though we know the laws; or an inability to obtain precise data (or the impossibility of doing so). At any given time, it does not matter. It is only with respect to the prospects for the advance of science in the future that it matters, and, even here, only in a minor way, since reference to such rigid and preconceived positions seems rather unnecessary anyway.

From time to time, as scientific prospects change, this or that particular mental attitude may be useful, in that it facilitates the formulation of theories which – for the moment – give better agreement with this or that point of view. However, nothing remains for ever unchanged; nothing is absolute. The particular mould in which one sets is not so important: what matters far more is not to set too firmly in any one pattern. To set fast is to no longer be alive.

5.11.2. These same remarks need repeating more generally, in connection with all those ways devised to saddle probability with an objective something (meaning, interpretation, justification, definition or whatever). In the first place, it is a fact that these attempts are not successful, and cannot be so, since, having the resolve to express matters relating to uncertainty in terms of the logic of certainty, they force themselves, *ab initio*, into a vicious circle, with no means of escape – 'per la contradizion che no'l consente'.[24] It is as if someone were to wish to hoist himself up by his bootlaces. Logic only permits the exposure of a tautology on the basis of what is taken as known; a prevision, however, is not simply a tautological consequence of what is already known. To be thus, would be to constitute something implicitly known, and would not involve uncertainty and, therefore, would not give rise to prevision.

However, even if we were to consider someone's arguments to constitute an acceptable basis for an objective meaning of probability (and, in general, such arguments will be different and concerned with special and different types of event, according to the different points of view), our thesis consists in believing that these arguments would be irrelevant anyway. All such conceptions, all the 'isms' they reduce to, are rejected here, but *not* in support of yet another 'ism' (as might be thought; e.g. 'subjectivism' or 'solipsism'), which one wants to put forward and contrapose to the others. The latter are rejected because, whatever the explanation of the uncertainty might be (attributing it to 'chance', 'fate', 'hidden laws', 'Providence', or 'statistical regularities', or to something else – or ... words(?)...), the sole concrete fact which is beyond dispute is that someone (me, You, somebody else) feels himself in a state of uncertainty, and has to decide on and adopt some point of view as a basis for previsions and related decisions.

5.11.3. This subjective meaning is an objective and unquestionable fact: all the rest (even if there were no dispute about it) is, in any case, something of an extra, which, at best, serves to help fix one's ideas. It is analogous to a vivid piece of writing that succeeds in forming something like an idea in our minds, although its meaning is not clear, and an analysis of the sentence in fact shows it to be inconsistent.

It is the case, however, that this view of the logic of uncertainty, complete and clear as it is, is far from achieving general acceptance. Why is this? Perhaps it is only the state of being certain which appears to most people as worthy of consideration and fit to be part of the edifice of science (which, according to the prevailing view, appears to express or aspire to omniscience – notwithstanding the fact that all progress, pushing back as it does the frontier of what is known, makes the horizon of what can be seen as unknown even broader). Perhaps the unknown and the uncertain disturb and annoy us, and

---

24 *Translators' note*. Ruled out by the principle of contradiction. (The line is from Dante's Inferno: canto XXVII, line 120.)

provoke those who are most upset to attempt to suppress them, or at least to make them disappear. There is not much point in philosophical arguments or speculations of this kind: they do provide, however, a possible explanation of why these different attitudes exist (we had to mention them, or, at least, to give some indication of how a person who finds our point of view natural could try to understand its lack of acceptance).

These different attitudes are, essentially, only variations on the same theme: the attempt to avoid the problems of uncertainty by simply pretending to overcome it; restricting the treatment to cases in which it can be presented in such a watered down way that it looks like something else.

The classical variant limits itself to cases like those of games of chance (where probability should acquire an objective meaning by virtue of the 'definition' based upon 'equally likely cases'). In the view of the most rigid supporters of this position, every application of the theory of probability outside this field would only be a questionable transposition by analogy.

The position that is at present most widely accepted restricts itself to cases of a certain statistical type (where probability should acquire an objective meaning by virtue of the 'definition' based upon 'frequency'). According to its most rigid adherents, the term 'probability', when used outside this context, has no more in common with its 'scientific' meaning than the 'energy' of a team leader has with the same term as applied to physical motion.[25]

Other approaches, which, having the aim of acting as guides in decision making, follow less rigid notions, attempt nevertheless to avoid those components of the argument which many find unpalatable (like the 'initial probabilities' required for Bayesian induction).[26]

Others adopt an eclectic attitude, accepting that one can base one's thinking on 'that probability which we evaluate for previsions and decisions' (i.e. the one corresponding to the conception of the present author), but, on the other hand, asserting that 'there is also another type of probability, the one with which statistics is concerned' (or, alternatively, 'the type valid in games of chance', or both).[27]

We should point out, here and now, that the mathematical treatment is unaffected (or, at most, very little affected) by these disagreements. In this sense, we can give a reassurance that everything we shall say mathematically is independent of questions of this kind, and should be acceptable to everyone. However, the interpretation is often different; there are certain nuances which, when looked into closely, completely change the spirit in which a given statement (perhaps expressible, in the same words, in the imprecise manner of everyday language) is to be understood.

So far as our own attitude is concerned, we wish to make clear that it is not utterly opposed to the attitude we have termed 'eclectic', even though it differs from it in a very real sense.

It is not utterly opposed because we recognize the importance of the problems, concepts and criteria that are the object of the various practical theories, even though we

25 The phrases given here, in characterizing the two attitudes, are due to Castelnuovo (transposition by 'analogy') and von Mises ('energy' and energy), respectively.

26 The followers of 'objectivistic statistics' in its various schools, including that of A. Wald (who we particularly have in mind here, as the nearest in approach to the Bayesian school).

27 The quotations are from V. Castellano. Typical examples of the eclectic attitude are provided by R. Carnap (who differentiates between 'probability$_1$' *logical*, and 'probability$_2$', *statistical*) and I.J. Good who admits the possible value of distinguishing many 'kinds of probability' (although in the context of a conception which is essentially subjectivistic).

study them within the framework of the general theory. Only by renouncing their alleged autonomy is it possible to compensate for those deficiencies in the foundations of the particularistic theories which render their conclusions meaningless, and the interpretation of them arbitrary.

It differs from it because we do not accept the existence of probabilities of different kinds, nor the autonomous validity of theories which set out to consider them, leaving aside some of the assumptions of the general theory, all of which are at all times essential.

All this has been summed up in an expressive manner by L.J. Savage (in a rather more specialized context): it is as though one wished to make a probabilistic omelette without breaking probabilistic eggs. There are two possible outcomes: either the result is not an omelette; or the eggs have in fact been used, either surreptitiously or inadvertently. All comments that we shall have occasion to make concerning 'other points of view' will essentially be continuations of the above analogy.