

Foundations of probability theory

Danny Nygård Hansen

15th November 2021

1 • Introduction

In the usual measure-theoretical formulation of probability theory, the following result is a corollary of the Law of Large Numbers:

THEOREM 1.1: *The frequency interpretation of probability*

Let X, X_1, X_2, \dots be i.i.d. real-valued random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For every $B \in \mathcal{B}(\mathbb{R})$ we have

$$\mathbb{P}(X \in B) = \lim_{n \rightarrow \infty} \frac{|\{j \in \{1, \dots, n\} \mid X_j(\omega) \in B\}|}{n}$$

for \mathbb{P} -almost all $\omega \in \Omega$.

PROOF. Thorbjørnsen Korollar 13.6.2. [Maybe reproduce the proof given LLN for completeness?] \square

That is, given a sequence $(X_n)_{n \in \mathbb{N}}$, and one extra X , of i.i.d. random variables, the probability that X lies in some Borel set B can be thought of as the proportion of the X_n that lie in B , as n tends to infinity. In other words, probability is a measure of the *frequency* with which an outcome of a random experiment obtains, if we repeat the experiment many times.

Whether or not this is the correct interpretation of probability as it occurs in the natural world we will not discuss here. Nonetheless the above result is an uncontroversial consequence of the theory, and it certainly aligns with our intuitive understanding of probability.

In this note we turn this result on its head and attempt to use it to motivate the formalisation of probability theory in terms of measure spaces. As we shall see, this is not entirely successful and will require some leaps that are not entirely justified by our conceptual grasp of probability.

2 • Preliminaries

2.1. Boolean algebras

We begin by reviewing some of the purely algebraic properties of Boolean algebras.

DEFINITION 2.1: Boolean algebras

A *Boolean algebra* is a structure $\langle B; \vee, \wedge, ', 0, 1 \rangle$ such that

- (i) $\langle B; \vee, \wedge \rangle$ is a distributive lattice,
- (ii) 0 and 1 are elements of B such that $x \vee 0 = x$ and $x \wedge 1 = x$ for all $x \in B$, and
- (iii) $'$ is a unary operation such that $x \vee x' = 1$ and $x \wedge x' = 0$ for all $x \in B$.

The binary operations \vee and \wedge are called *join* and *meet*, respectively. For $x \in B$ the element x' is called the *complement* of x . In a general bounded lattice L , an element $y \in L$ such that $x \vee y = 1$ and $x \wedge y = 0$ is called a complement of $x \in L$. If L is distributive, complements are unique. Recall also that the lattice structure on B induces a partial order \leq such that $x \leq y$ if and only if $x \vee y = y$ for $x, y \in B$.

Let B be a Boolean algebra. For $x, y \in B$ we define the *symmetric difference* between x and y by

$$x \triangle y = (x \wedge y') \vee (y \wedge x').$$

If $x \triangle y = 0$, then it is easy to show that $x = y$.

Before proceeding we note the following technical result that we shall need later:

LEMMA 2.2

Let $\langle B; \vee, \wedge, ', 0, 1 \rangle$ be a Boolean algebra. Let $(x_i)_{i \in I}$ be a collection of elements in B such that $x_i \wedge x_j = 0$ when $i \neq j$. If $\bigvee_{i \in I} x_i \in B$, then $\bigvee_{i \in J} x_i \in B$ for any cofinite¹ $J \subseteq I$.

PROOF. Let $(x_i)_{i \in I}$ be such a collection of elements, and let $J \subseteq I$ be cofinite. It suffices to prove the lemma in the case $I \setminus J = \{i_0\}$, since the general case then follows by induction. We claim that

$$\bigvee_{i \in J} x_i = x'_{i_0} \wedge \bigvee_{i \in I} x_i.$$

Let $j \in J$ and notice that, since $x_{i_0} \wedge x_j = 0$,

$$x'_{i_0} \wedge x_j = (x'_{i_0} \wedge x_j) \vee (x_{i_0} \wedge x_j) = (x'_{i_0} \vee x_{i_0}) \wedge x_j = 1 \wedge x_j = x_j.$$

¹ Recall that a subset J of a set I is called *cofinite* if the complement $I \setminus J$ is finite.

Now because $x_j \leq \bigvee_{i \in I} x_i$ we get

$$x_j = x'_{i_0} \wedge x_j \leq x'_{i_0} \wedge \bigvee_{i \in I} x_i.$$

Conversely, suppose that $x_j \leq s$ for all $j \in J$. Then $x_i \leq x_{i_0} \vee s$ for all $i \in I$, so

$$\bigvee_{i \in I} x_i \leq x_{i_0} \vee s.$$

It follows that

$$x'_{i_0} \wedge \bigvee_{i \in I} x_i \leq x'_{i_0} \wedge (x_{i_0} \vee s) = 0 \vee (x_{i_0} \wedge s) \leq s,$$

as desired. \square

2.2. Abstract measure spaces

DEFINITION 2.3: Generalised abstract measure spaces

A *measure* on a Boolean algebra B is a map $\mu: B \rightarrow [0, \infty)$ such that $x \wedge y = 0$ implies

$$\mu(x \vee y) = \mu(x) + \mu(y) \quad (2.1)$$

for all $x, y \in B$. If μ is a measure on a Boolean algebra B , then we call the pair (B, μ) a *generalised abstract measure space*. If $x \neq 0$ implies that $\mu(x) > 0$, then μ is called *positive definite*. If $\mu(1) = 1$, then we call μ a *probability measure*.

It is clear that $\mu(\emptyset) = 0$ and that μ is increasing. It follows that $\mu(x) \leq \mu(1)$ for all $x \in B$. Notice that we require that μ is finite, but this is no restriction since we are ultimately interested in the case where μ is a probability measure.

The property (2.1) is called *(finite) additivity* of μ , since an easy induction argument extends it to all finite joins. We will later define more restrictive structures and measures upon them, hence the adjective ‘generalised’.

PROPOSITION 2.4: Boole’s inequality

Let (B, μ) be a generalised abstract measure space B . Then for any $x, y \in B$ we have

$$\mu(x \vee y) \leq \mu(x) + \mu(y).$$

Similar to (2.1), Boole’s inequality may be extended to all finite joins by induction.

PROOF. Notice that

$$(x \wedge y') \wedge y = x \wedge (y' \wedge y) = 0,$$

and that

$$(x \wedge y') \vee y = (x \vee y) \wedge (y' \vee y) = x \vee y.$$

It follows by additivity of μ that

$$\mu(x \vee y) = \mu((x \wedge y') \vee y) = \mu(x \wedge y') + \mu(y) \leq \mu(x) + \mu(y),$$

as desired. \square

DEFINITION 2.5: Metric Boolean algebras

A *pseudometric Boolean algebra* is a tuple (B, ρ) , where B is a Boolean algebra and ρ is a pseudometric on B such that the maps $x \mapsto x'$, $(x, y) \mapsto x \vee y$, and $(x, y) \mapsto x \wedge y$ are continuous.

If ρ is a metric, then (B, ρ) is called a *metric Boolean algebra*.

Next we equip generalised abstract measure spaces with a canonical pseudometric. If (B, μ) is a generalised abstract measure space, define a map $\rho_\mu: B \times B \rightarrow [0, \infty)$ by

$$\rho_\mu(x, y) = \mu(x \triangle y) \tag{2.2}$$

for $x, y \in B$. The next proposition shows that ρ_μ is in fact a pseudometric. We will always equip a generalised abstract measure space with this pseudometric.

PROPOSITION 2.6

Given a generalised abstract measure space (B, μ) , the map ρ_μ defined in (2.2) makes (B, ρ_μ) into a pseudometric Boolean algebra. Furthermore, ρ_μ is a metric if and only if μ is positive definite.

PROOF. *ρ_μ is a pseudometric:* We only need to prove the triangle inequality. To this end, let $x, y, z \in B$ and notice that

$$\begin{aligned} x \wedge z' &= (x \wedge z) \wedge (y' \vee y) \\ &= (x \wedge z' \wedge y') \vee (x \wedge z' \wedge y) \\ &\leq (x \wedge y') \vee (y \wedge z'). \end{aligned}$$

Similarly we have $z \wedge x' \leq (z \wedge y') \vee (y \wedge x')$. It follows that

$$\begin{aligned} x \triangle z &= (x \wedge z') \vee (z \wedge x') \\ &\leq (x \wedge y') \vee (y \wedge x') \vee (y \wedge z') \vee (z \wedge y') \\ &= (x \triangle y) \vee (y \triangle z). \end{aligned}$$

Now Boole's inequality implies that

$$\rho_\mu(x, z) = \mu(x \triangle z) \leq \mu(x \triangle y) + \mu(y \triangle z) = \rho_\mu(x, y) + \rho_\mu(y, z),$$

as desired.

Continuity of lattice operations: Let $x \in B$, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in B that converges to x . Notice that $x'_n \triangle x' = x_n \triangle x$, so $\rho_\mu(x'_n, x') = \rho_\mu(x_n, x)$. Hence the complementation map $x \mapsto x'$ is continuous.

Let further $(y_n)_{n \in \mathbb{N}}$ be a sequence converging to a point $y \in B$. A short calculation shows that

$$\begin{aligned} (x_n \vee y_n) \triangle (x \vee y) &= (x_n \wedge x' \wedge y') \vee (y_n \wedge x' \wedge y') \vee (x \wedge x'_n \wedge y'_n) \vee (y \wedge x'_n \wedge y'_n) \\ &\leq (x_n \wedge x') \vee (x \wedge x'_n) \vee (y_n \wedge y') \vee (y \wedge y'_n) \\ &= (x_n \triangle x) \vee (y_n \triangle y). \end{aligned}$$

Thus Boole's inequality shows that

$$\rho_\mu(x_n \vee y_n, x \vee y) \leq \rho_\mu(x_n, x) + \rho_\mu(y_n, y), \quad (2.3)$$

which implies continuity of the join map $(x, y) \mapsto x \vee y$.

Finally, continuity of the meet map $(x, y) \mapsto x \wedge y$ follows since

$$x \wedge y = (x' \vee y')',$$

so it is a composition of continuous functions.

Positive definiteness: The last claim follows directly from the fact that $x \triangle y = 0$ if and only if $x = y$ for all $x, y \in B$. \square

REMARK 2.7. Notice that the measure μ can be written in terms of ρ_μ , since $\mu(x) = \rho_\mu(x, 0)$. Furthermore, since (pseudo)metrics are continuous, it follows that $\mu(x_n) \rightarrow \mu(x)$ whenever $x_n \rightarrow x$ in B . \lrcorner

It is well-known that any (pseudo)metric space has a completion, i.e. can be isometrically embedded as a dense subset of a complete (pseudo)metric space. See for instance Corollary 24.5 in Willard (1970). A natural question is then: If (B, ρ) is a (pseudo)metric Boolean algebra with metric completion $(\bar{B}, \bar{\rho})$, does \bar{B} also carry the structure of a Boolean algebra?

This is indeed the case, and we sketch the construction: Let $x, y \in \bar{B}$, and let (x_n) and (y_n) be sequences in B that converge to x and y , respectively. Then these are Cauchy sequences in B , and the calculation leading to (2.3) show that $(x_n \vee y_n)$ is also a Cauchy sequence. Thus it converges to some element of \bar{B} . Denote it $x \vee y$. We define x' and $x \wedge y$ similarly. It is easy to check that these operations satisfy the conditions in Definition 2.1. Furthermore, the completion $\bar{\rho}$ of the pseudometric ρ makes $(\bar{B}, \bar{\rho})$ into a pseudometric Boolean algebra.

This takes care of the metric structure. The next proposition shows that we can also extend the measure on a generalised abstract measure space to its completion.

PROPOSITION 2.8

Let (B, μ) be a generalised abstract measure space, and let $(\bar{B}, \bar{\rho}_\mu)$ be the completion of (B, ρ_μ) . Define a map $\bar{\mu}: \bar{B} \rightarrow [0, \infty)$ by

$$\bar{\mu}(x) = \lim_{n \rightarrow \infty} \mu(x_n), \quad (2.4)$$

where $(x_n)_{n \in \mathbb{N}}$ is any sequence in B that converges to x . Then $\bar{\mu}$ is a well-defined measure on \bar{B} . The generalised abstract measure space $(\bar{B}, \bar{\mu})$ is called the completion of (B, μ) .

PROOF. First notice that for any $x \in \bar{B}$ there does in fact exist a sequence in B converging to x . If $(x_n)_{n \in \mathbb{N}}$ is such a sequence, it is a Cauchy sequence in B , and the reverse triangle inequality shows that $(\mu(x_n))$ is a Cauchy sequence in \mathbb{R} , hence convergent. Thus the limit on the right-hand side of (2.4) exists.

Now let (y_n) be another sequence in B that approximates x . Another application of the reverse triangle inequality then shows that

$$|\mu(x_n) - \mu(y_n)| \leq \rho_\mu(x_n, y_n) \leq \rho_\mu(x_n, x) + \rho_\mu(y_n, x) \rightarrow 0.$$

Hence $\mu(x_n)$ and $\mu(y_n)$ converge to the same value, and thus $\bar{\mu}$ is well-defined.

Next we show that $\bar{\mu}$ is finitely additive. Let $x, y \in \bar{B}$ with $x \wedge y = 0$ and choose approximating sequences (x_n) and (y_n) in B . Then

$$(x_n \vee y_n) \wedge (x_n \wedge y_n)' = (x_n \wedge (x_n \wedge y_n)') \vee (y_n \wedge (x_n \wedge y_n)')$$

is the join of disjoint elements of B , so

$$\mu((x_n \vee y_n) \wedge (x_n \wedge y_n)') = \mu(x_n \wedge (x_n \wedge y_n)') + \mu(y_n \wedge (x_n \wedge y_n)').$$

By continuity of the lattice operations we have $(x_n \wedge y_n)' \rightarrow 1$, so the three elements given as arguments to μ above are elements in approximating sequences for $x \vee y$, x and y respectively. By definition of $\bar{\mu}$ it follows that

$$\bar{\mu}(x \vee y) = \bar{\mu}(x) + \bar{\mu}(y)$$

as desired. □

LEMMA 2.9

Let (B, μ) be a generalised abstract measure space. Every monotonic sequence in B is a Cauchy sequence.

PROOF. Let $(x_n)_{n \in \mathbb{N}}$ be an increasing sequence in B . Then since $x_n \leq 1$ we also have $\mu(x_n) \leq \mu(1) < \infty$ for all $n \in \mathbb{N}$. Thus the sequence $(\mu(x_n))$ is a bounded

increasing sequence in \mathbb{R} , hence it converges to some $\alpha \geq 0$. For $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that $\mu(x_n) \in (\alpha - \varepsilon, \alpha]$ for all $n \geq N$.

If then $m, n \geq N$ with $m \leq n$, then $x_m \leq x_n$ and so

$$x_m \wedge x'_n \leq x_n \wedge x'_n = 0.$$

Hence $x_m \triangle x_n = x_n \wedge x'_m$, so it follows that

$$\rho_\mu(x_m, x_n) = \mu(x_n \wedge x'_m) = \mu(x_n) - \mu(x_m) < \varepsilon.$$

Thus (x_n) is indeed a Cauchy sequence. The case where (x_n) is decreasing is similar. \square

PROPOSITION 2.10

Let (B, ρ) be a metric Boolean algebra with completion $(\bar{B}, \bar{\rho})$. Any sequence $(x_n)_{n \in \mathbb{N}}$ in B has a join in \bar{B} , and

$$\bigvee_{n \in \mathbb{N}} x_n = \lim_{n \rightarrow \infty} \bigvee_{i \leq n} x_i.$$

Similarly for meets.

As far as I know, it is not possible to generalise this result to the case where ρ is only a pseudometric. [We also need the measure, don't we? For the lemma? Kolmogorov only looks at this when there is a positive definite measure.]

PROOF. The sequence $(\bigvee_{i \leq n} x_i)_{n \in \mathbb{N}}$ is increasing, so by [Lemma 2.9](#) it has a limit $s \in \bar{B}$. For $k \in \mathbb{N}$ and $n \geq k$ we have $x_k \leq \bigvee_{i \leq n} x_i$, i.e.

$$x_k \vee \bigvee_{i \leq n} x_i = \bigvee_{i \leq n} x_i.$$

Taking the limit as $n \rightarrow \infty$, continuity of (binary) joins implies that $x_k \vee s = s$, or $x_k \leq s$. Thus s is an upper bound of the sequence (x_n) .

On the other hand, if $t \in \bar{B}$ is an upper bound of (x_n) , then $x_k \leq t$. We have just seen that taking limits preserves inequalities, so this implies that $s \leq t$ as desired.

The corresponding result for meets follows similarly, or from the fact that complementation is continuous. \square

2.3. Abstract σ -algebras and continuity

DEFINITION 2.11: Axiom of continuity

A generalised abstract measure space (B, μ) is said to satisfy the *axiom of continuity* if it has the following property: If $(x_n)_{n \in \mathbb{N}}$ is a decreasing sequence

of elements in B such that $\bigwedge_{n \in \mathbb{N}} x_n$ exists and equals 0, then $\lim_{n \rightarrow \infty} \mu(x_n) = 0$.

This is an analogue of the continuity (from above) of ordinary countably additive measures on concrete σ -algebras.

LEMMA 2.12

Let (B, μ) be a generalised abstract measure space with μ positive definite. Then (B, μ) satisfies the axiom of continuity.

PROOF. Let $(x_n)_{n \in \mathbb{N}}$ be a decreasing sequence in B such that $\bigwedge_{n \in \mathbb{N}} x_n = 0$. Since it is decreasing, $x_n = \bigwedge_{i \leq n} x_i$, so because μ is positive definite, [Proposition 2.10](#) implies that the sequence converges to its meet, i.e. it converges to 0. By [Remark 2.7](#), $\mu(x_n)$ converges to $\mu(0) = 0$, proving the claim. \square

PROPOSITION 2.13: The generalised addition theorem

Let (B, μ) be a generalised abstract measure space satisfying the axiom of continuity. If $(x_n)_{n \in \mathbb{N}}$ is a sequence of pairwise disjoint elements in B such that $x = \bigvee_{n \in \mathbb{N}} x_n \in B$, then

$$\mu(x) = \sum_{n=1}^{\infty} \mu(x_n).$$

PROOF. By [Lemma 2.2](#), $r_n = \bigvee_{i > n} x_i$ exists in B , and we claim that $\bigwedge_{n \in \mathbb{N}} r_n = 0$. Let $t \in \mathcal{F}$ be a lower bound of r_n for $n \in \mathbb{N}$. Then for $n \in \mathbb{N}$ we have

$$t \vee \bigvee_{i > n} x_i = \bigvee_{i > n} x_i,$$

and taking the meet of each side with x'_n yields $t \wedge x'_n = 0$. Hence $\bigvee_{n \in \mathbb{N}} t \wedge x_n = 0$. Now notice that $\bigwedge_{n \in \mathbb{N}} x_n = 0$, since $x_n \wedge x_m = 0$ when $n \neq m$. It follows by taking complements that $\bigvee_{n \in \mathbb{N}} x'_n = 1$, and so

$$t = t \wedge \bigvee_{n \in \mathbb{N}} x'_n = \bigvee_{n \in \mathbb{N}} t \wedge x'_n = 0.$$

Thus $\bigwedge_{n \in \mathbb{N}} r_n = 0$ as claimed. It now follows from finite additivity of μ and the axiom of continuity that

$$\mu(x) = \sum_{i=1}^n \mu(x_i) + \mu(r_n) \rightarrow \sum_{i=1}^{\infty} \mu(x_i)$$

as $n \rightarrow \infty$ as desired. \square

PROPOSITION 2.14: *Boole's inequality*

Let (B, μ) be a generalised abstract measure space, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in B . If $x = \bigvee_{n \in \mathbb{N}} x_n \in B$, then

$$\mu(x) \leq \sum_{n=1}^{\infty} \mu(x_n).$$

PROOF. First notice that, by the finite Boole equality,

$$\mu\left(\bigvee_{i \leq n} x_i\right) \leq \sum_{i=1}^n \mu(x_i) \leq \sum_{i=1}^{\infty} \mu(x_i)$$

for all $n \in \mathbb{N}$. By [Proposition 2.10](#), $\bigvee_{i \leq n} x_i \rightarrow x$ as $n \rightarrow \infty$, so $\mu(\bigvee_{i \leq n} x_i) \rightarrow \mu(x)$. The claim follows \square

Of course, the join of a sequence of elements may not exist. In the case where we are ensured the existence of countable joins we use the following terminology:

DEFINITION 2.15: *Abstract σ -algebra*

An *abstract σ -algebra* is a Boolean algebra B with countable joins. That is, if $(x_n)_{n \in \mathbb{N}}$ is a sequence of elements in B , then their join $\bigvee_{n \in \mathbb{N}} x_n$ exists.

If the join $\bigvee_{n \in \mathbb{N}} x_n$ exists, then it follows by taking complements that the meet $\bigwedge_{n \in \mathbb{N}} x'_n$ also exists. Hence an abstract σ -algebra also has countable meets. In the context of abstract measure spaces we obtain the following:

DEFINITION 2.16: *Abstract measure spaces*

An *abstract measure space* is a generalised abstract measure space (B, μ) that satisfies the axiom of continuity, and where B is an abstract σ -algebra.

LEMMA 2.17

If (B, μ) is a generalised abstract measure space with μ positive definite, then the completion $(\bar{B}, \bar{\mu})$ is an abstract measure space.

PROOF. Since the completion of a metric space is a metric space, $\bar{\mu}$ is positive definite, so [Lemma 2.12](#) implies that $(\bar{B}, \bar{\mu})$ satisfies the axiom of continuity.

On the other hand, [Proposition 2.10](#) implies that every sequence in \bar{B} has a join, so \bar{B} is an abstract σ -algebra. \square

3 • The algebra of probability spaces

3.1. Motivation

If the probability of an event is supposed to be a measure of how often this event occurs on repetitions of the random experiment in question, then it seems reasonable to assume that we are, at least in principle, able to distinguish when the event does and does not obtain. For example, after rolling a six-sided die the state of affairs ‘the result of the die roll is three’ is an event, since we can determine the outcome of the roll just by looking at the die. To take another example, after throwing a ball the state of affairs ‘the ball was thrown more than 50 metres’ is also an event: That is, we can determine whether or not the length of the throw was strictly greater than 50 metres.

One might take a different view: Say that one grants that it is possible to *affirm* that the length of the throw, measured in metres, lies in the interval $(50, \infty)$. If the length L in metres does in fact lie in the above interval, we can simply use a ruler whose subdivisions are smaller than $L - 50$ in metres. Still one might disagree that it is possible to *refute* that $L \in (50, \infty)$. For if L is exactly 50 metres, then since any measurement of L carries some error, it is in practice impossible to determine whether L is 50 (or slightly smaller), or whether it is slightly larger than 50. We will not pursue this line further but refer the reader to Vickers (1989) for more on this *logic of affirmative assertions*.

To be precise, after performing the relevant random experiment, we will assume that we are always able to decide whether or not the event has occurred or not. In particular, if E is an event, then the state of affairs ‘ E does not obtain’ is also an event, denoted \bar{E} : If E obtains, then \bar{E} does not. And conversely, if E does not obtain, then \bar{E} does obtain. We call \bar{E} the *complement of* or the *complementary event to* E .² Evidently, the complement of a complement is just the event we started with.

Next consider two events E_1 and E_2 . Since we are able to decide whether each of them have obtained, the same is true for the event ‘both E_1 and E_2 have obtained’ and the event ‘at least one of E_1 and E_2 has obtained’. The first is called the *conjunction* of E_1 and E_2 and is denoted $E_1 \wedge E_2$, and the second is the *disjunction* $E_1 \vee E_2$ of E_1 and E_2 .

Finally it seems natural to allow an ‘impossible event’ 0 which never occurs, as well as a ‘sure event’ 1 that always occurs. Clearly 0 and 1 are each other’s complements. If E_1 and E_2 are events with $E_1 \wedge E_2 = 0$, then this manifestly means that E_1 and E_2 cannot obtain simultaneously: The two events are *incompatible*.

We collect all the relevant events in a set \mathcal{F} and postulate that the structure

² In contrast, in the logic of affirmative assertions we do not allow complementation (i.e. negation). Hence it may not be surprising that this logic ends up being closely tied to topology, the relevant analogy being that the complement of an open set need not be open.

$\langle \mathcal{F}; \vee, \wedge, \neg, 0, 1 \rangle$ is a Boolean algebra. We leave it to the reader to reflect on the reasonability of this assumption. This leads naturally to the following definition:

DEFINITION 3.1: *Generalised abstract probability spaces*

A *generalised abstract probability space* is a generalised abstract measure space $(\mathcal{F}, \mathbb{P})$, where \mathbb{P} is a positive definite probability measure.

The partial order \leq induced by the lattice structure then has the interpretation that if $E \leq F$, then E implies F . For recall that this means that $E \vee F = F$, i.e. if F has already occurred then no information is gained by observing that E has also occurred. Conversely, if F has *not* occurred, then E is impossible.

We have justified every part of this definition except for the positive definiteness of \mathbb{P} . In the usual measure theoretical formulation of probability theory, there may (and often do) exist events that are not empty but still have probability zero. Kolmogorov had the following to say in critique of this approach:

[W]e are forced to give up the principle, formulated in numerous classical works in probability theory, according to which an event of probability zero is absolutely impossible. More precisely, one must allow that an event of positive probability can be decomposed into a (possible continuous) infinity of variants of which each has probability zero. (Kolmogorov and Jeffrey 1995)

Furthermore, this also has the technical benefit that the pseudometric $\rho_{\mathbb{P}}$ induced by \mathbb{P} is in fact a metric. We will return to the consequences of this choice in the next section.

3.2. Continuity of probability measures

Of course, the map \mathbb{P} above is supposed to be analogous to a probability measure on a (concrete) σ -algebra. But ordinary measures are *countably* additive, not just finitely so. It is however difficult to justify extending the finite additivity to sequences of disjoint events purely on conceptual or operational grounds. In fact, according to Kolmogorov himself:

Since the new axiom [countable additivity] is essential for infinite fields of probability only, it is almost impossible to elucidate its empirical meaning. (...) For, in describing any observable random process we can obtain only finite fields of probability. Infinite fields of probability occur only as idealized models of real random processes. *We limit ourselves, arbitrarily, to only those models which satisfy Axiom VI* [countable additivity]. This limitation

has been found expedient in researches of the most diverse sort.
(Kolmogorov 1956)

And furthermore:

[S]omewhat more complicated problems require, if the theory is to be simple and tractable, that probability be subject to the *axiom of denumerable additivity*. However, the justification of that axiom remains purely empirical, in that we have not yet encountered any interesting problem for which we have not been able to construct a probability field conforming to the axiom in question.
(Kolmogorov and Jeffrey 1995)

The axiom of continuity is, as we saw in [Proposition 2.13](#), equivalent to countable additivity in our formalism. While countable additivity is usually preferred in the definition of (probability) measures today, Kolmogorov (1956) instead assumed the axiom of continuity and then proved the generalised addition theorem, as we have done above.

Kolmogorov could apparently only find justification for assuming the axiom in the success of the theory, and not in its conceptual underpinnings. Luckily, in the present context we can avoid taking a stance: We have already heard Kolmogorov argue that the absense of non-empty events of probability zero is conceptually problematic; hence we have disallowed them by assuming our probability measures to be positive definite. And according to [Lemma 2.12](#), in this setting we get the axiom of continuity for free.

We are still not assured that our probability spaces are closed under countable joins. However, [Lemma 2.17](#) tells us that taking the completion of a generalised abstract probability space solves the problem. And according to Kolmogorov, this can be done without issue:

The use of denumerable additivity, with its great concomitant freedom, is not generally possible except in *complete* metric Boolean algebras. It is therefore natural, in probability theory, always to assume that the algebra of events is *complete* by adjunction of ideal elements. As has been pointed out, this situation is always realizable.

So, the use of denumerably additive probabilities is seen to be legitimate and to impose no additional restrictions on the nature of the problems which fall under the scope of the general theory.
(Kolmogorov and Jeffrey 1995)

Notice that he calls the adjoined elements ‘ideal’. Thus he seems to think of these new events as not (necessarily) corresponding to a real, observable event as described above. Instead, he seems only to be concerned with how

their adjunction can improve the flexibility of the theory without restricting which classes of problems the theory can address. And indeed, upon taking the completion of a probability space we always have an isometric copy of the original space inside the new one.

We summarise the result of the above discussion in the following definition, which is the one we will be concerned with going forward:

DEFINITION 3.2: Abstract probability spaces

An *abstract probability space* is a generalised abstract probability space $(\mathcal{F}, \mathbb{P})$, where \mathcal{F} is an abstract σ -algebra.

4 • Set-theoretical probability theory

4.1. Basic definitions and properties

We begin by recalling the standard definitions in order to compare them to the abstract versions we have considered above:

DEFINITION 4.1: Measurable spaces

A (concrete) σ -algebra in a set Ω is a collection \mathcal{F} of subsets of Ω such that

- (i) $\Omega \in \mathcal{F}$,
- (ii) $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$, and
- (iii) if $(A_n)_{n \in \mathbb{N}}$ is a sequence in \mathcal{F} , then $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$.

The sets in \mathcal{F} are called \mathcal{F} -measurable, and the pair (Ω, \mathcal{F}) is called a *measurable space*.

Note that a σ -algebra by this definition is indeed an *abstract* σ -algebra as defined in [Definition 2.15](#). Also contrast the definition of a σ -algebra with that of a *set algebra*: in the latter case also require it to contain Ω and be closed under complementation, but we only require it to be closed under *finite* intersections. In particular a set algebra is a Boolean algebra.

DEFINITION 4.2: Measure spaces and probability spaces

Let (Ω, \mathcal{F}) be a measurable space. A *measure* on (Ω, \mathcal{F}) is a map $\mu: \mathcal{F} \rightarrow [0, \infty]$ such that

- (i) $\mu(\emptyset) = 0$, and
- (ii) μ is countably additive, i.e. for every sequence $(A_n)_{n \in \mathbb{N}}$ of pairwise dis-

joint sets in \mathcal{F} we have

$$\mu\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n).$$

The triple $(\Omega, \mathcal{F}, \mu)$ is called a (*concrete*) *measure space*. If $\mu(\Omega) = 1$ then μ is called a *probability measure* and the triple $(\Omega, \mathcal{F}, \mu)$ a (*concrete*) *probability space*.

Again, if $(\Omega, \mathcal{F}, \mu)$ is a measure space thus defined, then (\mathcal{F}, μ) is also an *abstract* measure space of the type considered in [Definition 2.16](#). However, even if μ is a probability measure, (\mathcal{F}, μ) is not necessarily an *abstract* probability space as defined in [Definition 3.2](#).

The trouble is that $\mu(A) = 0$ does not necessarily imply that $A = \emptyset$, i.e. μ need not be positive definite. In an abstract measure space (perhaps of the generalised kind) (\mathcal{F}, μ) the only information we have about the relationship between two elements $x, y \in \mathcal{F}$ is whether they are comparable and, if so, which one is greater.³ Or in other terms for events E and F , whether one of them implies the other or not.

By contrast, elements of a concrete σ -algebra are subsets of an underlying set. This has several important implications: First of all, this vastly increases the number of objects we have to contend with, namely *every* subset of Ω , not just those lying in \mathcal{F} . This may have benefits of the technical nature; certainly it would have simplified many of the arguments in the previous section on Boolean algebras.

Secondly, it forces us to take the very elements of the *sample space* Ω seriously. Usually these are referred to as *outcomes* or, e.g. by Kolmogorov, *elementary events*. But this may be conceptually problematic:

[T]he notion of an elementary event is an artificial superstructure imposed on the concrete notion of an event. In reality, events are not composed of elementary events, but elementary events originate in the dismemberment of composite events. (Kolmogorov and Jeffrey 1995)

In other words, arriving at the idea of an elementary event requires *analysis* of the collection of events that is given in some random experiment.

A simple example may help elucidate this point: Consider the random experiment consisting of rolling a six-sided die. Say this die is loaded in such a way that it is impossible to roll a 6, and that it is very likely to roll a 1 or a 2.

³ Recall that the ordering on a lattice completely determines the lattice structure.

We construct an abstract probability space $(\mathcal{F}, \mathbb{P})$ to model the outcome of the die roll. The event space might be

$$\mathcal{F} = \{0, A_1, \dots, A_5, E, F, 1\},$$

where we interpret the event A_i as ‘the result was i ’ for $i = 1, \dots, 5$, E as ‘the result was even’ and F as ‘the result was odd’. There is no event A_6 since it is impossible to roll a 6. Certainly we would require that $A_1, A_3, A_5 \leq F$, and that $A_2, A_4 \leq E$. Assigning probabilities to each event we might find that $\mathbb{P}(A_1) = \mathbb{P}(A_2) = 0.35$, and that $\mathbb{P}(A_3) = \mathbb{P}(A_4) = \mathbb{P}(A_5) = 0.1$. Additivity of \mathbb{P} would then imply that $\mathbb{P}(E) = 0.45$ and $\mathbb{P}(F) = 0.55$.

It is tempting to *identify* each of E and F with the (possible) events that imply them, i.e. identify E with the set $\{A_2, A_4\}$ and F with the set $\{A_1, A_3, A_5\}$ of events. Maybe we are even tempted to include a hypothetical event A_6 in the former set. But notice that this is an *analysis* of the events E and F . It would be quite possible to grasp the meaning of the event E without immediately enumerating each of the ‘elementary’ events it consists of. Indeed events in even set-theoretic probability theory (and objects in mathematics as a whole, for that matter) are often defined in terms of their properties, not of their constituents. Or to put it in other terms: by their *intension* and not their *extension*.

On the other hand, the set-theoretic approach and the possibility of non-empty null sets offer certain conceptual advantages as well. Perhaps we would like there to be an event A_6 even if it is impossible. Certainly our model of the die seems incomplete without it; the die is, after all, six-sided! Admittedly this picture seems rather artificial since the probability of rolling a 6 using a physical die is surely positive, if very small. So imagine that this die appears in a video game or in a fantasy novel where this might be a more easily digestible proposition.

Furthermore, if we actually do want to use our theory of probability to *model* physical phenomena, then it might be perfectly reasonable to include some event in the model on conceptual grounds, even if this event happens to be assigned probability zero in the model. Perhaps this fact is due to numerical approximation, missing information or a state of affairs that comes about after each event under consideration has been identified.

4.2. Extending premeasures to σ -algebras

Recall that we in [Lemma 2.17](#) proved that the completion of a generalised abstract measure space (B, μ) is an abstract measure space under the assumption that μ is positive definite. Since this is no longer the case we need another way of extending a finitely additive probability measure on a set algebra to a

σ -algebra. Furthermore, in [Lemma 2.12](#) we showed that (B, μ) automatically satisfies the axiom of

References

- Billingsley, Patrick (1995). *Probability and Measure*. 3rd ed. Wiley. 593 pp. ISBN: 0-471-00710-2.
- Davey, B. A. and H. A. Priestley (2002). *Introduction to Lattices and Order*. 2nd ed. Cambridge University Press. 298 pp. ISBN: 978-0-521-78451-1.
- Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability*. 2nd ed. Chelsea Publishing Company. 84 pp.
- Kolmogorov, A. N. and Richard Jeffrey (1995). 'Complete Metric Boolean Algebras'. In: *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* 77.1, pp. 57–66. ISSN: 00318116, 15730883. URL: <http://www.jstor.org/stable/4320553>.
- Vickers, Steven (1989). *Topology via Logic*. 1st ed. Cambridge University Press. 200 pp. ISBN: 0-521-36062-5.
- Willard, Stephen (1970). *General Topology*. Addison-Wesley Publishing Company, Inc. 369 pp.