

Foundations of probability theory

Danny Nygård Hansen

13th 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 • Event spaces as Boolean algebras

If the probability of an event is supposed to be a measure of how often this event occurs, then it seems reasonable to assume that we are, in principle, able to distinguish when this event occurs. For example, 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, 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: One might agree that it is possible to *affirm* that the length of the throw, measured in metres, lies in the interval $(50, \infty)$. If the length is L , we can simply take a ruler whose subdivisions are smaller than $L - 50$ in metres. However, 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 E' : If E obtains, then E' does not. And conversely, if E does not obtain, then E' does obtain. We call E' the *complement of* or the *complementary event to* E .¹

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 ‘empty event’ 0 which never occurs, as well as a ‘universal event’ 1 that always occurs.

With these operations in place, we claim that this logic of events constitute a *Boolean algebra*. We recall the definition of such a structure:

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,

¹ 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.

(ii) $a \vee 0 = a$ and $a \wedge 1$ for all $a \in B$, and

(iii) $a \vee a' = 1$ and $a \wedge a' = 0$ for all $a \in B$.

We leave it to the reader to reflect on whether or not Boolean algebras offer a reasonable way to formalise our intuition of events (in the opinion of the author, they do). This leads naturally to the following definition:

DEFINITION 2.2: *Generalised abstract probability spaces*

A *generalised abstract probability space* is a pair $(\mathcal{F}, \mathbb{P})$, where \mathcal{F} is a set and $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$, such that

(i) $\langle \mathcal{F}; \vee, \wedge, ', 0, 1 \rangle$ is a Boolean algebra,

(ii) $\mathbb{P}(1) = 1$, and

(iii) $E \wedge F = 0$ implies $\mathbb{P}(E \vee F) = \mathbb{P}(E) + \mathbb{P}(F)$ for all $E, F \in \mathcal{F}$.

In [ref] we consider more restrictive abstract probability spaces, hence the adjective ‘generalised’. We refer to the property (iii) as *(finite) additivity* of \mathbb{P} .

We define an ordering on events in the usual way: If E and F are events such that $E \vee F = F$, then we write $E \leq F$ and say that E *implies* F . The intuition is that, if F obtains just when *either* E or F obtains, then there is no way for E to obtain without F also doing so.

LEMMA 2.3

Let $(\mathcal{F}, \mathbb{P})$ be a generalised abstract probability space.

(i) $\mathbb{P}(0) = 0$.

(ii) \mathbb{P} is increasing, i.e. $E \leq F$ implies $\mathbb{P}(E) \leq \mathbb{P}(F)$ for all $E, F \in \mathcal{F}$.

PROOF. The first claim follows since $0 \vee 0 = 0 \wedge 0 = 0$, so by additivity of \mathbb{P} ,

$$\mathbb{P}(0) = \mathbb{P}(0 \vee 0) = \mathbb{P}(0) + \mathbb{P}(0),$$

hence $\mathbb{P}(0) = 0$.

To prove the second claim, first notice that if $E \leq F$ then $F = E \vee (F \wedge E')$. But $E \wedge (F \wedge E') = 0$, so

$$\mathbb{P}(F) = \mathbb{P}(E \vee (F \wedge E')) = \mathbb{P}(E) + \mathbb{P}(F \wedge E') \geq \mathbb{P}(E),$$

as claimed. □

References

- Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability*. 2nd ed. Chelsea Publishing Company. 84 pp.
- Vickers, Steven (1989). *Topology via Logic*. 1st ed. Cambridge University Press. 200 pp. ISBN: 0-521-36062-5.