0.2. The Probability Space

Note. In this section, we set up the environment in which probability theory is performed. We consider the axiomatic approach of Andrey Kolmogorov and give an intuitive motivation for it.

Note/Definition. Suppose a given random experiment is repeated n times and that event A occurs $f_n(A)$ times. Then the *relative frequency* of occurrences of A is $r_n(A) = f_n(A)/n$. The *stabilization of relative frequencies* "property" claims that:

$$r_n(A) \text{ converges to some real number as } n \to \infty.$$

This idea is more formally addressed in the Law of Large Numbers (see Section 6.5. The Law of Large Numbers and the Central Limit Theorem). The "frequency interpretation" is the starting point for an informal development of probability.

Note. Andrey Kolmogorov presented an axiomatization of the theory of probability in his 1933 *Grundbegriffe der Wahrscheinlichkeitsrechnung* ("Basic Concepts of Probability Theory"). To motive Kolmogorov's axioms, consider the following intuitive observations:

(a) Since $0 \le f_n(A) \le n$ for any event A, then $0 \le r_n(A) \le 1$. So the probability of an event is a real number in the interval $[0, 1]$.

(b) If $A = \varnothing$, then $f_n(\varnothing) = 0$ and hence $r_n(\varnothing) = 0$. The probability of $\varnothing$ is then 0. If A is the whole space of possible outcomes, which we denote Ω, then $f_n(\Omega) = n$ for all $n \in \mathbb{N}$ and hence $r_n(\Omega) = 1$. The probability of Ω is then 1.

(c) Let B be the complement of A (that is, $B = \Omega \setminus A$). Since in each performance of the experiment, either A or B occurs and never both simultaneously, then $f_n(A) + f_n(B) = n$, and hence $r_n(A) + r_n(B) = 1$. The sum of the probability of an event and the probability of its complement is 1.

(d) Suppose that the event A is contained in event B (that is, $A \subseteq B$). Then $f_n(A) \leq f_n(B)$ and hence $r_n(A) \leq r_n(B)$. The probability of A is less than or equal to the probability of B.

(e) Suppose events A and B are disjoint, and let C be their union (denoted $C = A \mathbin{\dot\cup} B$). Then $f_n(C) = f_n(A) + f_n(B)$ and hence $r_n(C) = r_n(A) + r_n(B)$. The probability of the union of two disjoint events equals the sum of their individual probabilities. This is called *finite additivity*.

(f) If A and B are not disjoint, then we have $f_n(C) \leq f_n(A) + f_n(B)$ and hence $r_n(C) \leq r_n(A) + r_n(B)$. The probability of the union of two events is less than or equal to the sum of the individual probabilities.

(g) More specifically, if A and B are not disjoint, then $f_n(C) = f_n(A) + f_n(B) - f_n(D)$ where D equals the intersection of A and B (i.e., $D = A \cap B$). Hence $r_n(C) = r_n(A) + r_n(B) - r_n(D)$. The probability of the union of two events equals the sum of the individual probabilities minus the probability of their intersection.

Motivated by these observations, we now start a formal statement of Kolmogorov's approach.

Definition. Consider nonempty set Ω, called the *sample space*. An *elementary event* ω is an element of the sample space, $\omega \in \Omega$. Let $\mathcal{F}$ be a collection of subsets of Ω which is (1) closed under complements (i.e., if $A \in \mathcal{F}$ then the complement $A^c = \Omega \setminus A \in \mathcal{F}$), (2) closed under countable unions (i.e., if $A_1, A_2, A_3, \ldots \in \mathcal{F}$ then $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$), and (3) closed under countable intersections (i.e., if $A_1, A_2, A_3, \ldots \in \mathcal{F}$ then $\cap_{i=1}^{\infty} A_i \in \mathcal{F}$). The elements of $\mathcal{F}$ are called *events* (assuming they are "measurable"; see the following note).

Note. The three conditions imposed on $\mathcal{F}$ imply that it is a σ-algebra (also called a σ-field; see my online notes on Real Analysis 1 [MATH 5210] on Section 1.4. Open Sets, Closed Sets, and Borel Sets of Real Numbers, and my online notes for Mathematical Statistics 1 [STAT 4047/5047] on Section 1.3. The Probability Set Function). By de Morgan's Laws, conditions (1) and (2) imply condition (3) (and conditions (1) and (3) imply condition (2)), so the previous definition could be slightly simplified by eliminating condition (2) or condition (3). Since we will commonly deal with integrals over events in $\mathcal{F}$, is is necessary that we assume all events are *Lebesgue measurable*. We largely gloss over this detail in these notes, but it is an essential part of a measure theoretic approach to probability theory (such as that given in my online notes on Measure Theory Based Probability).

Definition. A *probability space* $(\Omega, \mathcal{F}, P)$ is a nonempty set Ω, a σ-algebra of Lebesgue measurable sets $\mathcal{F}$, and a function $P : \mathcal{F} \to \mathbb{R}$ such that

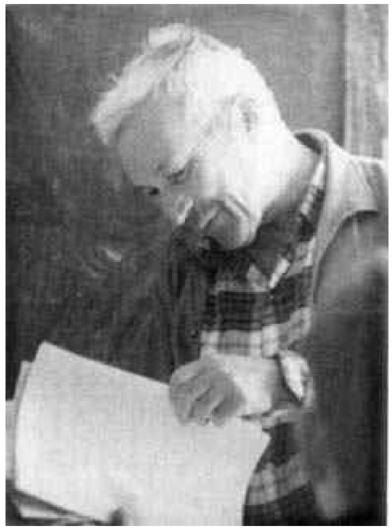
Axiom 1. For any $A \in \mathcal{F}$, we have $P(A) \geq 0$.

Axiom 2. $P(\Omega) = 1$.

Axiom 3. Let $\{A_n \mid n \in \mathbb{N}\}$ be a countable collection of pairwise disjoint events, and let A be their union. Then $P(A) = \sum_{n=1}^{\infty} P(A_n)$.

The value $P(A)$ is the *probability* of event $A \in \mathcal{F}$. Axiom 3 is the property of *countable additivity*, and includes the property of *finite additivity* as a special case: For $A_1, A_2, \ldots, A_n \in \mathcal{F}$, we have $P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A_2) + \cdots + P(A_n)$.

Note. Andrey Kolmogorov entered Moscow State University in 1920. While still an undergraduate, he published papers on set theory and function theory. He graduated in 1925 and in that same year published eight papers, including his first paper on probability. He completed his doctorate in 1929 (by which time he had 18 publications). He started as a professor at Moscow University in 1931 and published *Grundbegriffe der Wahrscheinlichkeitsrechnung* ("Basic Concepts of Probability Theory") in 1933, as mentioned above. In this, he gave his axioms of probability theory. In 1938-39, Moscow University created the Department of Probability and Statistics with Kolmogorov as the chair. Other areas to which he contributed include the theory of Markov random processes, dynamical systems (in connection with planetary motion), and topology (where he introduced cohomology groups; the American mathematician James Alexander independently studied these ideas at about the same time). Kolmogorov was recognized by mathematical and scientific organizations around the world, including in the USSR, Romania, the UK, the USA, Netherlands, and France. This historical information and the following figure are from MacTutor History of Mathematics Archive Kolmogorov webpage.



Andrey N. Kolmogorov

April 25, 1903–October 20, 1987

Revised: 10/8/2022