Lecture 4 Probability Axioms

BIO210 Biostatistics

Xi Chen

Fall, 2025

School of Life Sciences
Southern University of Science and Technology

南方科技大学生命科学学院
SUSTech · SCHOOL OF
LIFE SCIENCES

Probability

Probability theory is nothing but common sense reduced to calculation.

Laplace

Notations

Set

A set is a well-defined collection of distinct objects.

S = { list or description of the objects in the set }

Definitions

Sample space (Ω)

Set of all possible outcomes

Outcomes: **mutually exclusive** and **collectively exhaustive**

Sample space example 1

Example 1: flipping a coin four times

Sample space $\boldsymbol{\Omega}$ $= \{$ HHHH, HHHT, HHTH, HTHH,
THHH, HHTT, HTHT, THHT,
THTH, TTHH, HTTH, TTTH,
TTHT, THTT, HTTT, TTTT $\}$

Sample space example 2

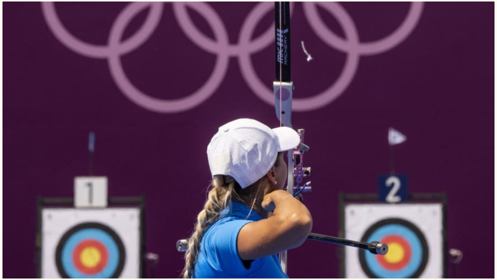
Example 2: an exam contained ten questions; each has 10 points; what is the total points you may get ?

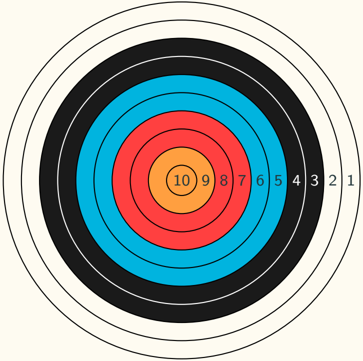
Sample space $\boldsymbol{\Omega} = \{0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$

Alternative sample space $\boldsymbol{\Omega} = \{$ 0, 10, 20, 30, 40, 50, 60, 70, 80, 90,
100 and you are using your lucky pen,
100 and you are not using your lucky pen $\}$

Sample space example 3

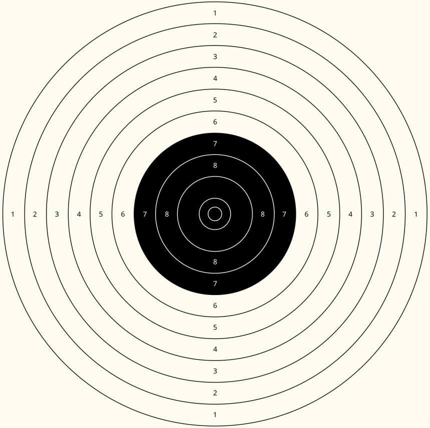
Example 3: shooting on a circular target (archery)





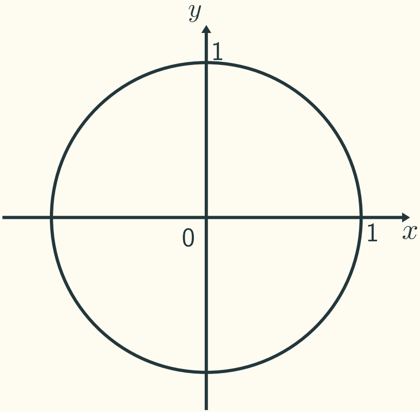
Sample space example 3

Example 3: shooting on a circular target (ISSF 10 Metre Air Pistol)



Sample space example 3

Example 3: positions on a target



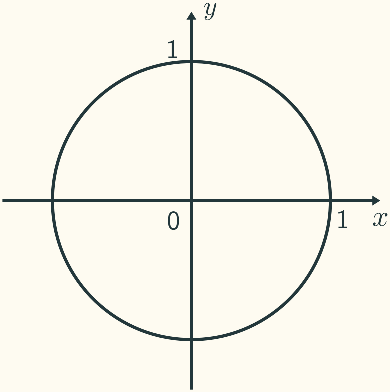
Sample space $\boldsymbol{\Omega}$ $= \{\ (x,\ y) \mid x^2 + y^2 \leqslant 1\ \}$

Assign probability to outcomes ... ?

Now, we can assign probability to individual outcomes ...

Not exactly!

What is the probability of hitting (0, 0)?



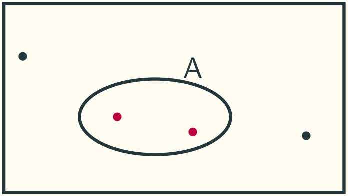
Event

Event

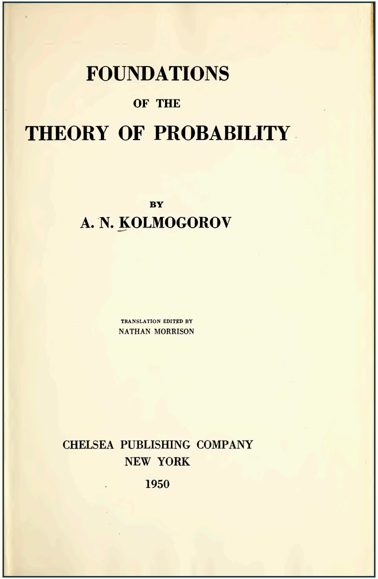
An event (A, B, C, D, *etc.*): a subset of the sample space Ω

- Probabilities are assigned to events. The probability represents our belief on how likely we think an event will occur.
- Event A has occurred. $\leftarrow$ what does this mean?

Sample space Ω



Probability axioms


FOUNDATIONS

OF THE

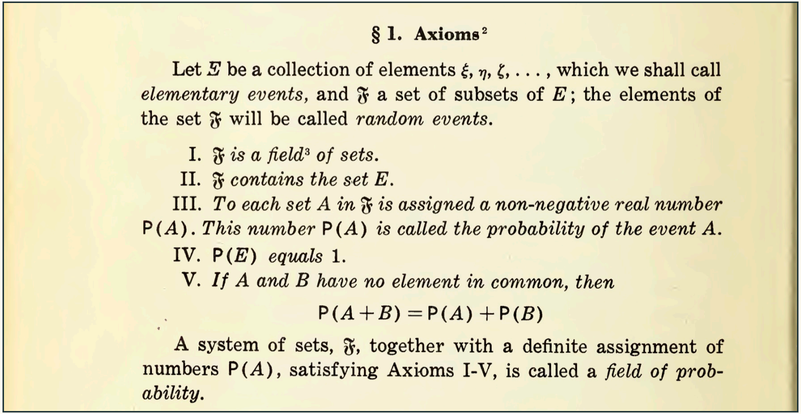
THEORY OF PROBABILITY

BY

A. N. KOLMOGOROV

TRANSLATION EDITED BY
NATHAN MORRISON

CHELSEA PUBLISHING COMPANY
NEW YORK
1950


§ 1. Axioms[2]

Let E be a collection of elements $\xi, \eta, \zeta, \ldots$, which we shall call *elementary events*, and $\mathfrak{F}$ a set of subsets of E; the elements of the set $\mathfrak{F}$ will be called *random events*.

I. *$\mathfrak{F}$ is a field*[3] *of sets.*

II. *$\mathfrak{F}$ contains the set E.*

III. *To each set A in $\mathfrak{F}$ is assigned a non-negative real number $\mathsf{P}(A)$. This number $\mathsf{P}(A)$ is called the probability of the event A.*

IV. *$\mathsf{P}(E)$ equals* 1.

V. *If A and B have no element in common, then*

$$\mathsf{P}(A+B) = \mathsf{P}(A) + \mathsf{P}(B)$$

A system of sets, $\mathfrak{F}$, together with a definite assignment of numbers $\mathsf{P}(A)$, satisfying Axioms I-V, is called a *field of probability*.

Probability axioms

The Kolmogorov Axioms

1. Nonnegativity: $\mathbb{P}(A) \geqslant 0$
2. Normalisation: $\mathbb{P}(\boldsymbol{\Omega}) = 1$
3. Additivity: if A and B are distjoint $(A \cap B = \emptyset)$, then $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$

Nice properties

- The probability of any event is always between 0 and 1.
- If $A_1,\ A_2,\ A_3,\ \cdots,\ A_n$ are disjoint, then

$$\mathbb{P}\left(A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_n\right) = \mathbb{P}\left(A_1\right) + \mathbb{P}\left(A_2\right) + \mathbb{P}\left(A_3\right) + \cdots + \mathbb{P}\left(A_n\right)$$

- $s_1,\ s_2,\ s_3,\ \cdots,\ s_k$ are individual outcomes from the sample space, then

$$\begin{aligned} \mathbb{P}\left(\{s_1,\ s_2,\ s_3,\ \cdots,\ s_k\}\right) &= \mathbb{P}\left(\{s_1\}\right) + \mathbb{P}\left(\{s_2\}\right) + \cdots + \mathbb{P}\left(\{s_k\}\right) \\ &= \mathbb{P}\left(s_1\right) + \mathbb{P}\left(s_2\right) + \cdots + \mathbb{P}\left(s_k\right) \leftarrow \text{abuse notation} \end{aligned}$$

Frequentist interpretation

Probabilities as long-term relative frequencies

If an experiment is repeated n times under essentially the identical conditions, and if the event A occurs m times, then as n grows large, the ratio $\frac{m}{n}$ approaches a fixed limit that is the probability of A:

$$\mathbb{P}\left(A\right) = \frac{m}{n}\text{, where } n \text{ is large.}$$

Measure of Belief

Probabilities as a measure of belief

- The probability that my current manuscript about RGCs will get published without revision is 1%.
- The probability that you will get a full score in BIO210 is 5%.
- The probability that it rains tomorrow is 80%.

Assigning probability

Experiment 1: flipping a fair coin four times

Sample space Ω $= \{$ HHHH, HHHT, HHTH, HTHH,
THHH, HHTT, HTHT, THHT,
THTH, TTHH, HTTH, TTTH,
TTHT, THTT, HTTT, TTTT $\}$

All possible outcomes are **equally likely**, so we can let every possible outcome have a probability of $1/16$.

Calculate the probabilities of the following events:

$A = \{$all heads or tails$\}$
$B = \{$exactly two head$\}$
$C = \{$at least two tails$\}$

Discrete uniform law

Discrete Uniform Law

Let all outcomes be equally likely, then

$$\mathbb{P}(A) = \frac{\text{number of elements of } A}{\text{total number of sample points}} = \frac{|A|}{|\Omega|}$$

Computing probability is essentially just counting!

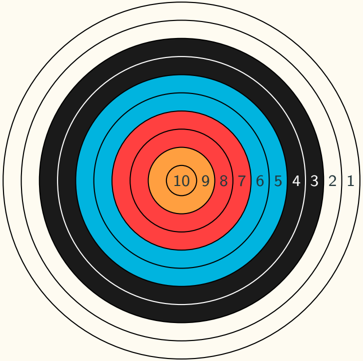
Continuous uniform law

Experiment 2: archery

Sample space $\boldsymbol{\Omega} = \{\ (x,\ y) \mid x^2 + y^2 \leqslant 1\ \}$

All possible outcomes are equally likely, Then probability = the ratio of areas.

$A = \{\text{hitting the red area}\}$
$B = \{(x,\ y) \mid x + y \leqslant 1\}$
$C = \{(0,\ 1), (1,\ 0), (0,\ -1), (-1,\ 0)\}$



Countable additive axiom

Experiment 3: keep flipping a fair coin until you obtain a head for the first time and stop.

Sample space $\boldsymbol{\Omega} = \{$ H, TH, TTH, TTTH, TTTTH, $\cdots\}$

Let n be the number of flips, $\mathbb{P}(n) = \dfrac{1}{2^n}$, $n = 1, 2, 3, 4, \cdots$

$A = \{\ n$ is an even number $\}$, $\mathbb{P}(A) = ?$

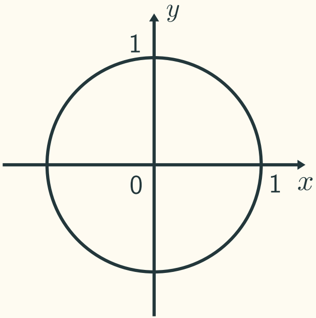
Countable additivity axiom

Countable Additivity Axiom

If a sequence of events A_1, A_2, A_3, $\cdots$ are disjoint, then

$$\mathbb{P}\left(A_1 \cup A_2 \cup A_3 \cup \cdots\right) = \mathbb{P}\left(A_1\right) + \mathbb{P}\left(A_2\right) + \mathbb{P}\left(A_3\right) + \cdots$$

Misuse of the countable additivity axiom



Sample space $\boldsymbol{\Omega} = \{\ (x,\ y) \mid x^2 + y^2 \leqslant 1\ \}$

Paradox 1??

$$1 = \mathbb{P}(\boldsymbol{\Omega}) = \mathbb{P}\left(\bigcup\{(x,y)\}\right) = \sum_{x,y} \mathbb{P}(\{(x,y)\}) = \sum_{x,y} 0 = 0$$

Take-home message: $\{(x,\ y)\}$ is uncountable: it is not possible to list every single one of $(x,\ y)$.

Paradox 2??

An experiment is performed, and the outcome is $\left(\frac{1}{2}, \frac{1}{2}\right)$

Take-home message: probability of 0 does NOT mean impossible.