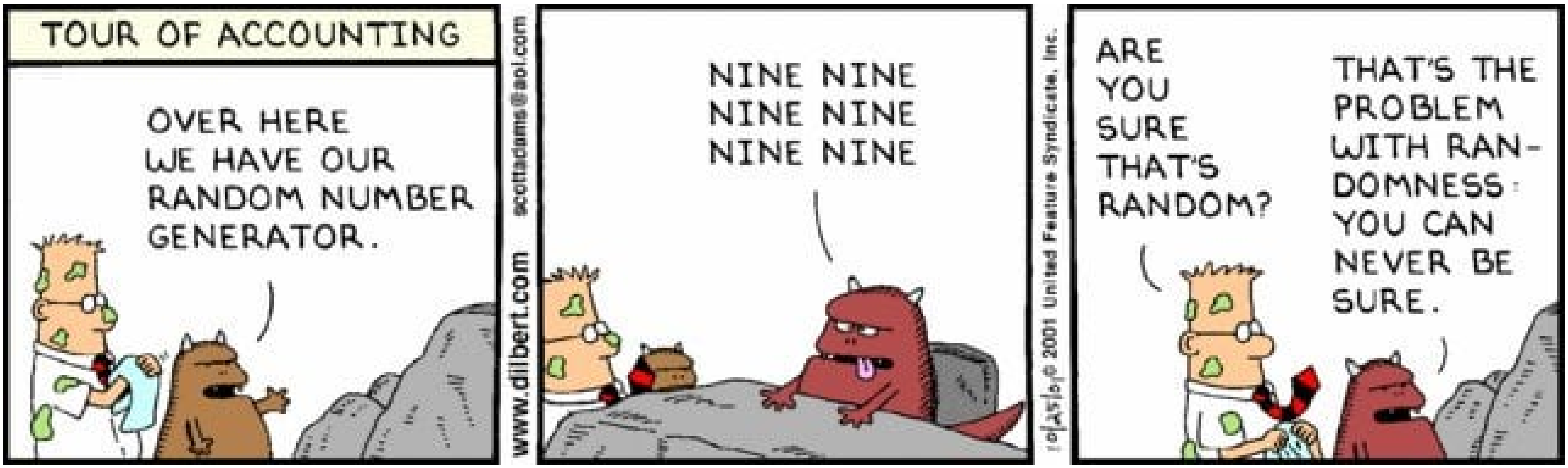
Richard von Mises' Approach to Probability

Ilya Pavlyukevich
Institute for Mathematics
Friedrich Schiller University Jena

Workshop at the Virtual Center for Structure, Methods, and Theory
Friedrich Schiller University Jena
1 April 2026

Randomness


TOUR OF ACCOUNTING
OVER HERE WE HAVE OUR RANDOM NUMBER GENERATOR.
scottadams@aol.com
www.dilbert.com
NINE NINE NINE NINE NINE NINE
© 2001 United Feature Syndicate, Inc.
ARE YOU SURE THAT'S RANDOM?
THAT'S THE PROBLEM WITH RANDOMNESS: YOU CAN NEVER BE SURE.

Randomness and probability

Toss a fair coin 20 times:

$$10010010001111010001$$

or

$$01100011001101110011$$

or

$$00000000000000000000$$

$$01010101010101010101$$

All these sequences are equally probable, $p = (\frac{1}{2})^{20}$

What is random and what is probability?

For infinite sequences like

$$1001000000111101000101100011001101110011011111110000001111101\ldots$$

$$1010101010101010101010101010101010101010101010101010101010\ldots$$

the probability to observe a concrete sequence is zero, $p = 0$

Historically: Probability $\approx$ chances

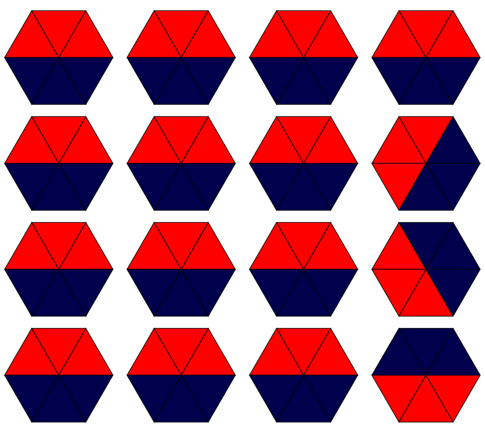
Gerolamo Cardano (1501–1576), Christiaan Huygens (1629–1695) etc

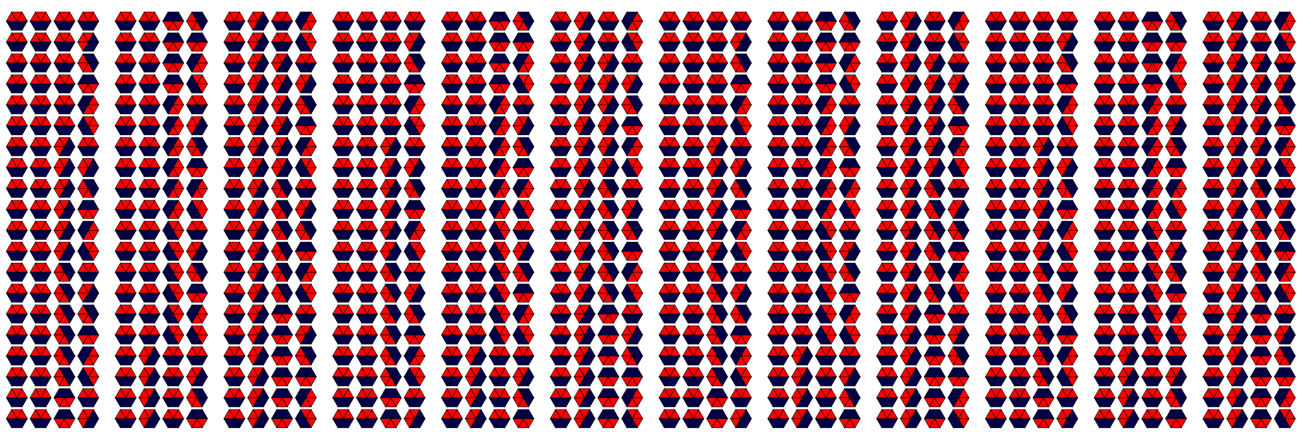
$$\text{Probability} \approx \text{chances} = \frac{\text{number of favorable outcomes}}{\text{total number of equally possible outcomes}}$$

Calculations related to “Games of chance”, combinatorial methods.

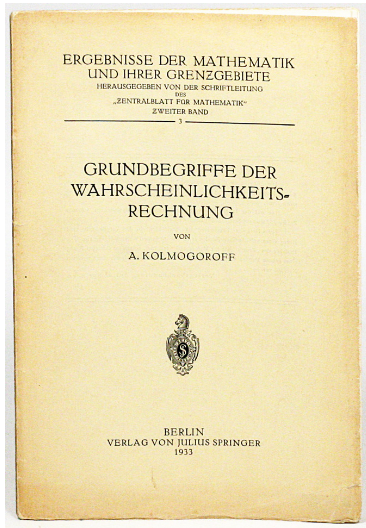
Jacob Bernoulli (1654–1705), “Ars Conjectandi”, 1713, for the first considered *potentially infinite* sequences of repeated trials and posed the question of the limiting behavior of the *frequencies* of occurrence of certain events in these trials → **Law of Large Numbers**.

Arguments were limited to elementary arithmetic and combinatorial methods.





$\cdots \to$ **Kolmogorov, 1933**



Andrey N. Kolmogorov (1903–1987)
$*$Tambov–Moscow†
Since 1931: professor at the Lomonosov Moscow State University

Hilbert's sixth problem (1900): to axiomatize those branches of physics (including probability) in which mathematics is prevalent.

$(\Omega, \mathscr{F}, \mathbf{P})$ is a probability space:

$\Omega \neq \varnothing$ is a space of elementary samples

$\mathscr{F}$ is a σ-algebra of "measurable" subsets of Ω or "random" events

Probability $\mathbf{P}$ is a **measure**: σ-additive set function,

$$\mathbf{P}\colon \mathscr{F} \to [0,1], \quad \mathbf{P}(\Omega) = 1,$$

$$A \mapsto \mathbf{P}(A) \quad \text{— probability of } A$$

von Mises, 1919 $\leftarrow \cdots$

Richard von Mises (1883–1953) $*$Lemberg–Vienna–Brünn–Strasbourg–Dresden–Berlin–Istanbul–Harvard†
Applied mathematician, far-famed for his works in mechanics, esp. in hydrodynamics and theory of flight.
Since 1919: director and full professor at the Institute of Applied Mathematics, University of Berlin
Since 1939: in Harvard, Professor of Aerodynamics and Applied Mathematics (1944).

Goal: to define probability theory as a mathematical discipline.

"We state here explicitly: The rational concept of probability, which is the only basis of probability calculus, applies only to problems in which either the same event repeats itself again and again, or a great number of uniform elements are involved at the same time. Using the language of physics, we may say that in order to apply the theory of probability we must have a practically unlimited sequence of uniform observations."

Then, the main object is the **Kollektiv**

Kollektiv

Let $\omega = (\omega_i)$ be an infinite sequence of elements. Each element ω_i takes values in a space M (“Merkmalraum”, e.g., $M = \{0, 1\}$). Such a sequence is called *Kollektiv* if it satisfies axioms I and II.

I. Existence of a limit. Let A be an arbitrary specified subset of the space M, and define

$$\nu_n(A; (\omega_i)_{i \leqslant n}) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{I}_A(\omega_i)$$

as the frequency of occurrence of the “event” A in n trials $(\omega_1, \ldots, \omega_n)$. Then there exists a limit

$$\lim_{n \to \infty} \nu_n(A; (\omega_i)_{i \leqslant n}) =: W(A; \omega)$$

II. Irregularity of the assignment. For any subsequence $(\omega'_i)_{i \geqslant 1} = (\omega'_1, \omega'_2, \ldots)$ obtained from the sequence $\omega = (\omega_1, \omega_2, \ldots)$ by any “admissible” selection of elements $\omega'_i = \omega_{n_i}$, where $n_1 < n_2 < \cdots$ there exists the limit

$$\lim_{n \to \infty} \nu_n\big(A; (\omega'_i)_{i \leqslant n}\big) = W(A; \omega') = W(A; \omega).$$

“Erst das Kollektiv, dann die Wahrscheinlichkeit”

The problem of place selection

What is “any ‘admissible’ selection of elements”?

Examples of simple methods for selecting subsequences:

a) Choose n if n is prime.

b) Choose n if it is divisible by 6 (arithmetic rules)

c) Choose n if the $(n-9)$-th, $(n-8)$-th,..., $(n-1)$-th terms of ω are all equal to 1 (a gambling strategy).

d) Take a second coin (Kollektiv ω'), supposed to be independent of the first. Choose n if the outcome $\omega'_n = 1$.

etc.

Critisism of the Kollektiv approach

No rigorous definition of the admissible place selection.

The properties of $W(A;\omega)$ are not clear, for example, is $W(A,\omega)$ σ-additive?

Do Kollektivs exist?

To justify the very existence of collective von Mises referred to the existence of gambling houses and to impossibility of construction of winning strategies against the “random” sequences produced in casinos.

What could be a Kollektiv?

Consider an infinite coin tossing in Kolmogorov's axiomatics.

$\Omega = \{0,1\} \times \{0,1\} \times \cdots = \{0,1\}^{\mathbb{N}}$.

$\mathscr{F}$ is the σ-algebra generated by "finite" sequences $(\omega_1, \ldots, \omega_n)$ (cylindrical σ-algebra)

Then, there exists a probability $\mathbf{P}$ (a product measure)

$$\mathbf{P} = \bigotimes_{n=1}^{\infty} \left(\frac{1}{2}\delta_0 + \frac{1}{2}\delta_1 \right)$$

In particular, for each n

$$\mathbf{P}(\{(\omega_1, \ldots, \omega_n)\}) = \frac{1}{2^n}$$

There are limit theorems about sequences of independent coin tosses (random walks) that hold almost surely (with probability 1)

The strong law of large numbers

With probability 1

$$\frac{\omega_1 + \cdots + \omega_n}{n} \to \frac{1}{2}, \quad n \to \infty.$$

That is, there is a set $\Omega_0 \in \mathscr{F}$ with $\mathbf{P}(\Omega_0) = 1$ such that for each $\omega \in \Omega_0$

$$\frac{\omega_1 + \cdots + \omega_n}{n} \to \frac{1}{2}, \quad n \to \infty.$$

Ω_0 can be a candidate for the set of Kollektivs. For instance,

$$\omega = (1,1,1,1,1,1,1,1,1,1,1,1,\ldots) \notin \Omega_0$$
$$\omega = (1,0,0,1,0,0,1,0,0,1,0,0,\ldots) \notin \Omega_0$$

but

$$\omega = (1,0,1,0,1,0,1,0,1,0,1,0,1,\ldots) \in \Omega_0$$
$$\omega = (1,1,0,0,1,1,0,0,1,1,0,0,1,\ldots) \in \Omega_0$$

Law of the iterated logarithm

With probability 1

$$\limsup_{n\to\infty} \frac{(2\omega_1 - 1) + \cdots + (2\omega_n - 1)}{\sqrt{2n \ln\ln n}} = 1,$$
$$\liminf_{n\to\infty} \frac{(2\omega_1 - 1) + \cdots + (2\omega_n - 1)}{\sqrt{2n \ln\ln n}} = -1,$$

That is, there is a set $\Omega_1 \in \mathscr{F}$ with $\mathbf{P}(\Omega_1) = 1$ such that for each $\omega \in \Omega_1$ these limits hold.

Then, for example,

$$\omega = (1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, \dots) \notin \Omega_1$$

In any case, it is not clear how to describe the null-sets $\bar{\Omega}_0$, $\bar{\Omega}_1$.

Very roughly speaking, from the viewpoint of Kolmogorov's axiomatics null-sets are not of practical importance.

Wald, 1937

Back to Kollektivs.

How to choose a subsequence?

Obviously, one has to impose restrictions on the place selection rules.

Wald: The place selection in von Mises' sense is should be made with the help of a function $\mathbf{W}$ (actually, a family of functions) taking values 0 or 1.

Denote

$$\omega^{(n)} = (\omega_1, \ldots, \omega_n), \quad \omega^{(0)} = \varnothing.$$

We define a subsequence ω'

$$\omega' = (\omega_{n_1}, \omega_{n_2}, \ldots)$$

where

$$n_1 := \min\{k > 0 \colon \mathbf{W}(\omega^{(k-1)}) = 1\},$$

$$n_2 := \min\{k > n_1 \colon \mathbf{W}(\omega^{(k-1)}) = 1\},$$

$$\ldots$$

Existence of Kollektivs

The basic result of Wald is a proof of **existence** of Kollektivs.

For any countable family of selection rules $\{\mathbf{W}_i\}$ *there exists a sequence* ω *that has the property of frequency stability with respect to all the rules in this family.*

However:

1) one cannot guarantee that the limiting set functions $W(A, \omega)$ are σ-additive...

2) the "probability" $W(A, \omega)$ depends on the chosen countable family of selection rules $\{\mathbf{W}_i\}$...

Thus, the notion of a Kollektiv becomes non-empty (Kollectivs do exist) if, from all possible selection rules, one selects some countable family.

Can we choose a particular countable family of selection rules?

Church, 1940

One should take **computable functions** (exactly countably many):

→ von Mises–Wald–Church Kollectivs

But: Ville's example

There is a sequence ω such that it is Kollektiv in the sense of von Mises–Wald–Church,

$$\frac{\omega_1 + \cdots + \omega_n}{n} \to \frac{1}{2}, \quad n \to \infty$$

but for all $n \in \mathbb{N}$

$$\frac{\omega_1 + \cdots + \omega_n}{n} \geqslant \frac{1}{2}$$

which violates the law of the iterated logarithm!

...

However: these questions open the door to new fields like algorithmic complexity, theory of chaos, etc.

How to calculate probabilities with help of Kollektivs?

“Probability” is a primitive notion of the theory comparable to those of “energy” or “mass” in other physical theories.

Whereas energy or mass exist in fields or material objects, probabilities exist the mathematical idealization of Kollektivs (random sequences).

How to determine probability of some complex event?

One has to construct some Kollektivs, and then, derive, according to certain rules, new Kollektivs from given ones and calculate the distributions of these new Kollektivs.

von Mises vs. Kolmogorov

This formalization [Kolmogorov (1933)], measure-theoretic probability, has served and still serves as the standard foundation of probability theory; **virtually all current mathematical work on probability uses the measure-theoretic approach**. [Vovk and Schaefer 2001]

However,

“I have already expressed the view [. . .] that the basis for the applicability of the results of the mathematical theory of probability to real ’random phenomena’ must depend on some form of **the frequency concept of probability, the unavoidable nature of which has been established by von Mises in a spirited manner**.” [Kolmogorov, 1963]

References

N. Vereshchagin, V. A. Uspensky, A. Shen, Kolmogorovstaya slozhnost' i algoritmicheskaya sluchainos' (in Russian), https://hal-lirmm.ccsd.cnrs.fr/lirmm-00786255v1

R. von Mises, Probability, Statistics and Truth, Dover Pblications, Inc.

Selected Papers of R. von Mises, Vol. 2, Probability and Statistics, General, 1964, AMS

M. Schield and T. V. V. Burnham, Von Mises' Frequentist Approach to Probability Section on Statistical Education — JSM 2008.

A. N. Shiryaev, On the Evolution of the von Mises' notion of randomness, https://www.ambitprocesses.au.dk/fileadmin/pdfs/ambit/Shiryaev.pdf

M. van Lambalgen, Randomness And Foundations Of Probability: Von Mises' Axiomatisation Of Random Sequences, Statistics, Probability and Game Theory, IMS Lecture Notes — Monograph Series (1996) Volume 30.

Photos: internet