تاریخچه‌ای از نظریه احتمال

گردآورنده: سهند حشمتی افشار

اردیبهشت 1392

Gerolamo Cardano(1501-1576)

ax^3 +bx+c=0

Galileo(1564-1642)

Galileo(1564-1642)

Galileo Paradox: which shows that there are as many perfect squares as there are whole numbers, even though most numbers are not perfect squares.

Simplicio: Here a difficulty presents itself which appears to me insoluble. Since it is clear that we may have one line greater than another, each containing an infinite number of points, we are forced to admit that, within one and the same class, we may have something greater than infinity, because the infinity of points in the long line is greater than the infinity of points in the short line. This assigning to an infinite quantity a value greater than infinity is quite beyond my comprehension.

Salviati: This is one of the difficulties which arise when we attempt, **with our finite minds, to discuss the infinite**, assigning to it those properties which we give to the finite and limited; but this I think is wrong, **for we cannot speak of infinite quantities as being the one greater or less than or equal to another**. To prove this I have in mind an argument which, for the sake of clearness, I shall put in the form of questions to Simplicio who raised this difficulty.
I take it for granted that you know which of the numbers are squares and which are not.

Simplicio: I am quite aware that a squared number is one which results from the multiplication of another number by itself; thus 4, 9, etc., are squared numbers which come from multiplying 2, 3, etc., by themselves.

Salviati: Very well; and you also know that just as the products are called squares so the factors are called sides or roots; while on the other hand those numbers which do not consist of two equal factors are not squares. Therefore if I assert that all numbers, including both squares and non-squares, are more than the squares alone, I shall speak the truth, shall I not?

Simplicio: Most certainly.

Salviati: If I should ask further how many squares there are one might reply truly that there are as many as the corresponding number of roots, since every square has its own root and every root its own square, while no square has more than one root and no root more than one square.

Simplicio: Precisely so.

Salviati: But if I inquire how many roots there are, it cannot be denied that there are as many as the numbers because every number is the root of some square. This being granted, we must say that there are as many squares as there are numbers because they are just as numerous as their roots, and all the numbers are roots. Yet at the outset we said that there are many more numbers than squares, since the larger portion of them are not squares. Not only so, but the proportionate number of squares diminishes as we pass to larger numbers, Thus up to 100 we have 10 squares, that is, the squares constitute 1/10 part of all the numbers; up to 10000, we find only 1/100 part to be squares; and up to a million only 1/1000 part; on the other hand in an infinite number, if one could conceive of such a thing, he would be forced to admit that there are as many squares as there are numbers taken all together.

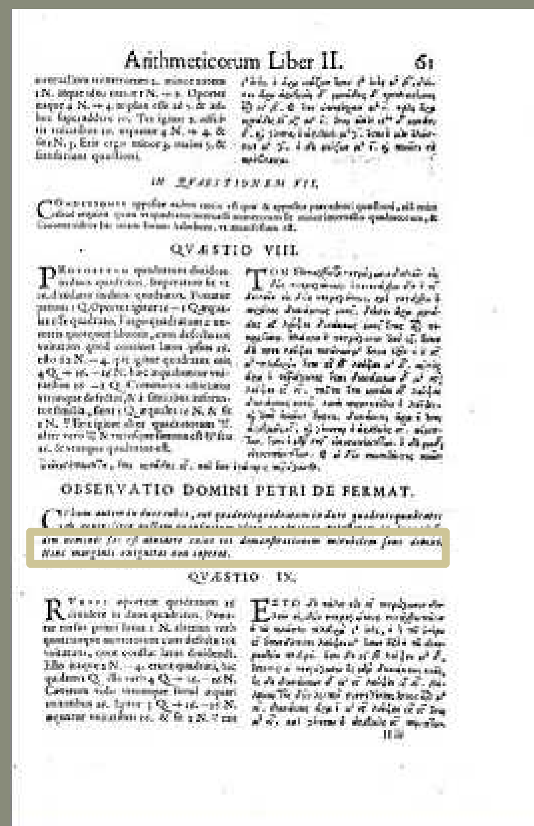
Sagredo: What then must one conclude under these circumstances?

Salviati: So far as I see we can only infer that the totality of all numbers is infinite, that the number of squares is infinite, and that the number of their roots is infinite; neither is the number of squares less than the totality of all the numbers, nor the latter greater than the former; **and finally the attributes "equal," "greater," and "less," are not applicable to infinite, but only to finite, quantities**. When therefore Simplicio introduces several lines of different lengths and asks me how it is possible that the longer ones do not contain more points than the shorter, I answer him that one line does not contain more or less or just as many points as another, but that each line contains an infinite number.— *Galileo, Two New Sciences*

Pascal (1623-1662)

Fermat (1601 or 1607-1665)

$a^n + b^n = c^n$



Arithmeticorum Liber II. 61

IN QVAESTIONEM VII.

QVAESTIO VIII.

OBSERVATIO DOMINI PETRI DE FERMAT.

QVAESTIO IX.

"I have found for this truly wonderful proof, but the margin is too small to hold it."

Probability Becomes Science: Correspondence between Fermat and Pascal

- Problem of Points:

Two equally skilled players are interrupted while playing a game of chance for a certain amount of money. Given the score of the game at that point, how should the stakes be divided?

Huygens(1629 – 1695)

Probability in Print

Leibniz(1646–1716German)

Dissertatio de Arte Combinatoria: $C(n,r)=\frac{n!}{r!(n-r)!}$

Leibniz was the first to use the term *analysis situs*([illegible])

John Wallis(1616–1703)

He is credited with introducin the symbol for infinity.

$$P(n,r)=\frac{n!}{(n-r)!}$$

Jacob Bernoulli (1655 -1705, Switzerland)

$$y' = p(x)y + q(x)y^n.$$

Bernoulli discovered the constant *e*:

$$\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

The number of ways in which a total of *m* points can be obtained by throwing *n* dice at once is equal to the coefficient of x^m in:

$$(x + x^2 + x^3 + x^4 + x^5 + x^6)^n$$

Daniel Bernoulli(1700–1782)

A casino offers a game of chance for a single player in which a fair coin is tossed at each stage. The pot starts at 1 dollar and is doubled every time a head appears. The first time a tail appears, the game ends and the player wins whatever is in the pot. Thus the player wins 1 dollar if a tail appears on the first toss, 2 dollars if a head appears on the first toss and a tail on the second, 4 dollars if a head appears on the first two tosses and a tail on the third, 8 dollars if a head appears on the first three tosses and a tail on the fourth, and so on. In short, the player wins 2^{k-1} dollars if the coin is tossed k times until the first tail appears.

What would be a fair price to pay the casino for entering the game? To answer this we need to consider what would be the average payout: With probability 1/2, the player wins 1 dollar; with probability 1/4 the player wins 2 dollars; with probability 1/8 the player wins 4 dollars, and so on. The expected value is thus

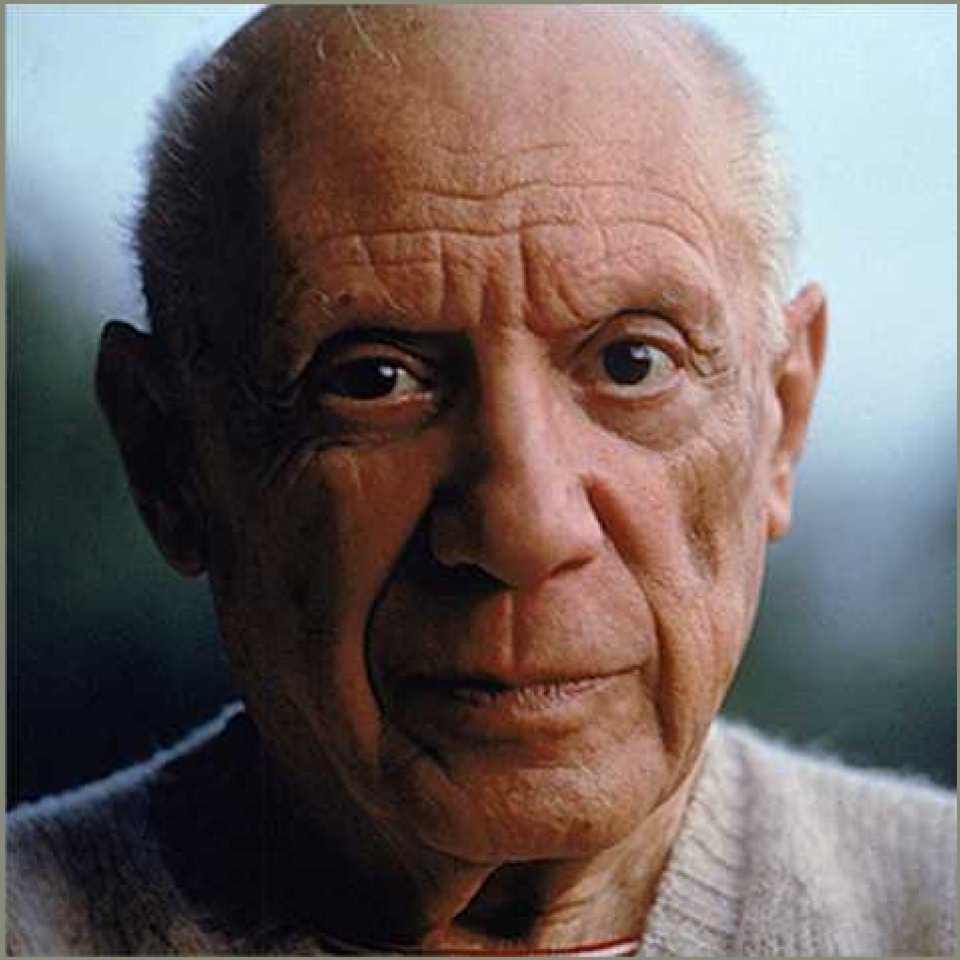
$$\begin{aligned} E &= \frac{1}{2}\cdot 1 + \frac{1}{4}\cdot 2 + \frac{1}{8}\cdot 4 + \frac{1}{16}\cdot 8 + \cdots \\ &= \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots \\ &= \sum_{k=1}^{\infty} \frac{1}{2} = \infty . \end{aligned}$$

Assuming the game can continue as long as the coin toss results in heads, in particular that the casino has unlimited resources, this sum grows without bound, and so the expected win for the player, at least in this idealized form, is an infinite amount of money. Considering nothing but the expectation value of the net change in one's monetary wealth, one should therefore play the game at any price if offered the opportunity. Yet, in published descriptions of the game, many people expressed disbelief in the result. Martin quotes Ian Hacking as saying "few of us would pay even $25 to enter such a game" and says most commentators would agree. The paradox is the discrepancy between what people seem willing to pay to enter the game and the infinite expected value.

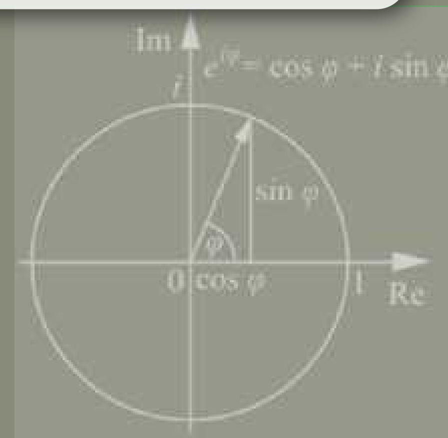
De Moivre(1667-1754)

If p is the probability of the success of an event, the probabilty of exactly r successes followed by n-r failures is:

$$C(n,r)p^r(1-p)^{n-r}$$



Leonhard Euler(1707–1783, German)



$$\gamma = \lim_{n\to\infty}\left(1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\cdots+\frac{1}{n}-\ln(n)\right).$$

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \lim_{n\to\infty}\left(\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots+\frac{1}{n^2}\right) = \frac{\pi^2}{6}.$$

$\neq, i$

Buffon (1707–1788)

Introduced differential and integral calculus to probability

Buffon's Needle

The problem in more mathematical terms is: Given a needle of length dropped on a plane ruled with parallel lines t units apart, what is the probability that the needle will cross a line?

Let x be the distance from the center of the needle to the closest line, let θ be the acute angle between the needle and the lines.

The uniform probability density function of x between 0 and t /2 is

$$\begin{cases} \frac{2}{t} & : 0 \le x \le \frac{t}{2} \\ 0 & : \text{elsewhere.} \end{cases}$$

The uniform probability density function of θ between 0 and π/2 is

$$\begin{cases} \frac{2}{\pi} & : 0 \le \theta \le \frac{\pi}{2} \\ 0 & : \text{elsewhere.} \end{cases}$$

The two random variables, x and θ, are independent, so the joint probability density function is the product

$$\begin{cases} \frac{4}{t\pi} & : 0 \le x \le \frac{t}{2},\ 0 \le \theta \le \frac{\pi}{2} \\ 0 & : \text{elsewhere.} \end{cases}$$

The needle crosses a line if

$$x \le \frac{l}{2} \sin\theta.$$

This GIF image describes the solution of Buffon's Needle Problem for the "short needle" case

Suppose . $t \ge l$

Integrating the joint probability density function gives the probability that the needle will cross a line:

$$P = \int_{\theta=0}^{\frac{\pi}{2}} \int_{x=0}^{(l/2)\sin\theta} \frac{4}{t\pi}\, dx\, d\theta = \frac{2l}{t\pi}.$$

A particularly nice argument for this result can alternatively be given using "Buffon's noodle".

Estimating π

In the first, simpler case above, the formula obtained for the probability can be rearranged to:

$$\pi = \frac{2l}{tP}$$

Thus, if we conduct an experiment to estimate P, we will also have an estimate for π.

Suppose we drop *n* needles and find that *h* of those needles are crossing lines, so is approximated by the fraction . This leads to the formula:

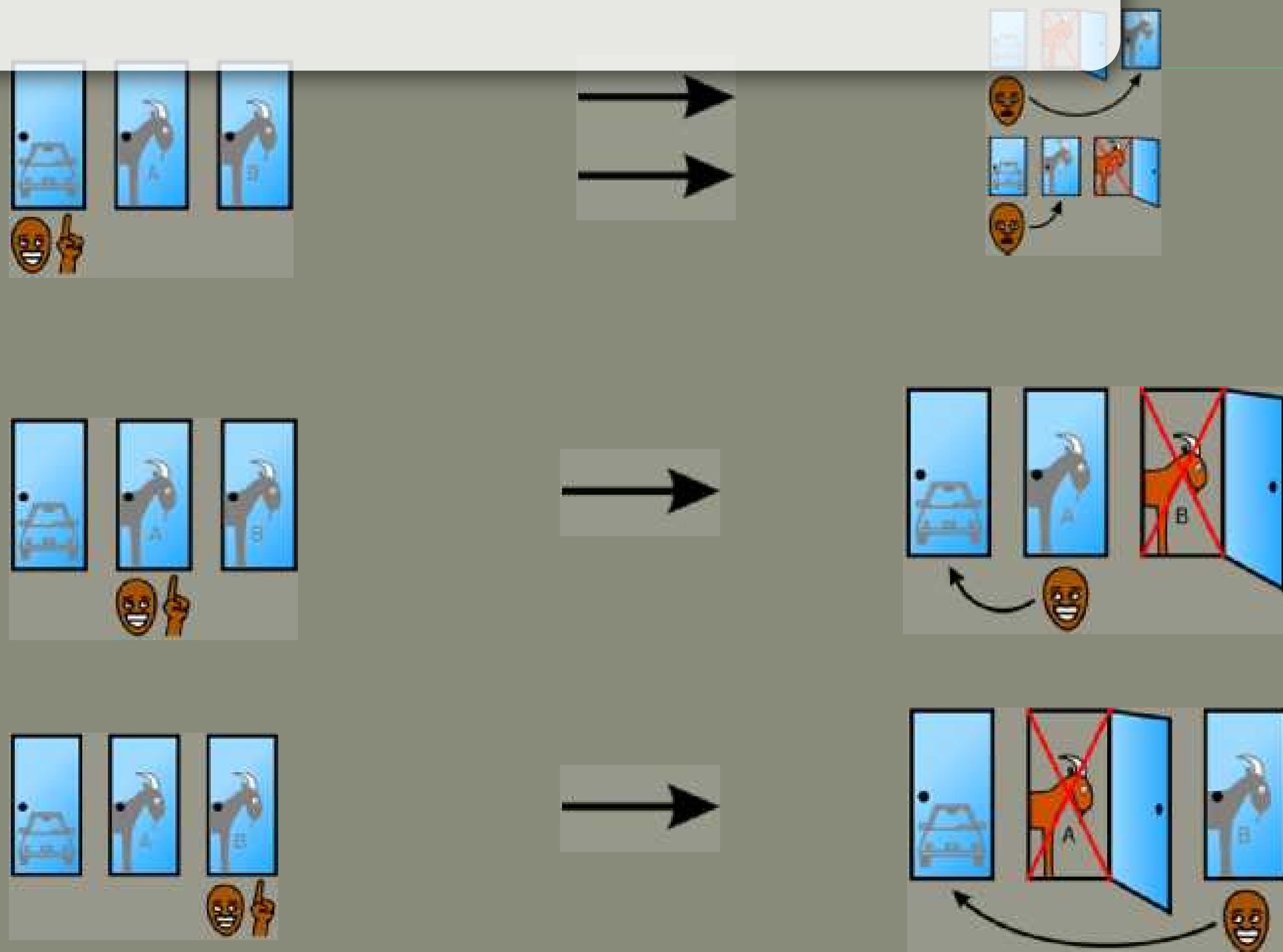
$$\pi \approx \frac{2l \cdot n}{th}.$$

In 1901, Italian mathematician Mario Lazzarini performed the Buffon's needle experiment. Tossing a needle 3408 times, he obtained the well-known estimate 355/113 for π, which is a very accurate value, differing from π by no more than 3×10^{-7}

Thomas Bayes(1702-1761)

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}.$$

-  Bertrand's Box Paradox
- -Monty Hall Problem(Steve Selvin)
- -Three Prisoner Problem(Martin Gardner) 



Laplace (1749–1827)

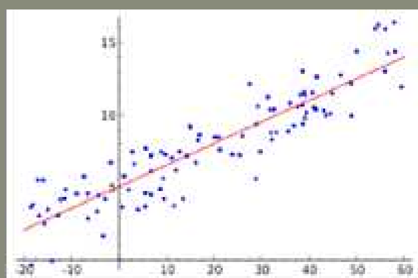
"I had no need of such a hypothesis."

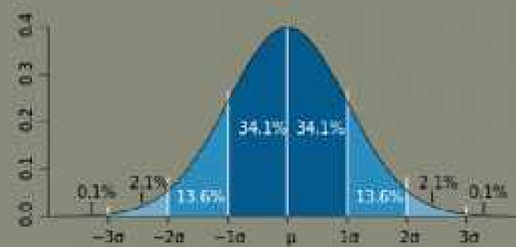
Classic definition of probability:

1. Probability is the ratio of the "favored events" to the total possible events.
2. The first principle assumes equal probabilities for all events. When this is not true, we must first determine the probabilities of each event. Then, the probability is the sum of the probabilities of all possible favored events.
3. For independent events, the probability of the occurrence of all is the probability of each multiplied together.
4. For events not independent, the probability of event B following event A (or event A causing B) is the probability of A multiplied by the probability that A and B both occur.
5. The probability that *A* will occur, given that B has occurred, is the probability of *A* and *B* occurring divided by the probability of *B*.
6. Three corollaries are given for the sixth principle, which amount to Bayesian probability. Where event $A_i \in \{A_1, A_2, \ldots A_n\}$ exhausts the list of possible causes for event B, $\Pr(B) = \Pr(A_1, A_2, \ldots A_n)$. Then

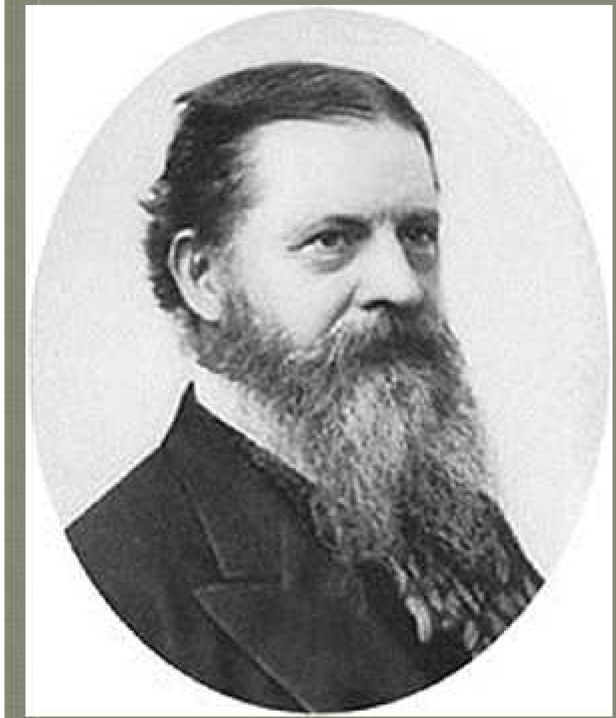
$$\Pr(A_i|B) = \Pr(A_i)\frac{\Pr(B|A_i)}{\sum_j \Pr(A_j)\Pr(B|A_j)}.$$

Gauss(1777 –1855)





Charles Peirce



Pragmatism

In 1860 he suggested a cardinal arithmetic for infinite numbers, years before any work by Georg Cantor (who completed his dissertation in 1867) and without access to Bernard Bolzano's 1851 (posthumous) *Paradoxien des Unendlichen*.

In 1881 he set out the axiomatization of natural number arithmetic, a few years before Richard Dedekind and Giuseppe Peano.

In 1885 he distinguished between first-order and second-order quantification. In the same paper he set out what can be read as the first (primitive) axiomatic set theory, anticipating Zermelo by about two decades (Brady 2000, pp. 132–3).

In 1886 he saw that Boolean calculations could be carried out via electrical switches, the same idea that was used decades later to produce digital computers.

Charles S. Peirce, 1839 to 1914, was one of America's most outstanding intellects. Philosopher, mathematician, and scientist, he wrote profusely, the equivalent of almost 100,000 printed pages in all. He was able to publish only about one-seventh of this, and most of his writings in mathematics and logic were never published during his lifetime.

The Collected papers of Charles Sanders Peirce appeared from 1931 to 1958; volumes 1 to 6 were edited by Charles Hartshorne and Paul Weiss (1931–1935), and volumes 7 and 8 by the present writer (1958). While these volumes included some previously unpublished papers in mathematical logic, by design they included almost none of Peirce's other papers in mathematics, nor his drafts of textbooks.

Carolyn Eisele, Professor Emeritus of Mathematics at Hunter College, has now filled this gap. She has edited about 2500 pages of the unpublished manuscripts, encompassing pure mathematics, numerous applications, and some rather ingenious textbook materials. *The new elements of mathematics* includes Peirce's papers on linear algebra and matrices, Euclidean and non-Euclidean geometry, topology and Listing numbers, graphs, and the four-color problem; also, his mathematical applications to economics, map projections, engineering, and the theory of errors. In addition, there are writings on the logic of relatives, Boolean algebra, and the nature of continuity; on probability, inductive logic, and applications of induction to historical inquiry. Finally, Professor Eisele provides most of Peirce's drafts of textbooks on arithmetic, geometry, and trigonometry.

Charles Peirce was the son of Benjamin Peirce (1809–1880), America's first original mathematician, whose *Linear associative algebra* appeared in 1870. Charles derived many results from his father's algebras, and he demonstrated their connection to relations (matrices).

Charles Peirce also proved a theorem about the rotation of bodies in four-dimensional space. But his most important mathematical results were in symbolic logic, a subject not generally accepted by mathematicians in Peirce's time. He developed the formalism of the propositional calculus and the general logic of quantifiers, independently of, though a little later than, Gottlob Frege. Independently of Dedekind, Peirce defined a finite set as one that cannot be put in one-one correspondence with a proper subset of itself.

Keynes (1883–1946)

Richard von Mises(1883-1953)

The law of stability of statistical frequencies:

$$\lim_{n \to \infty} \frac{m(A)}{n}$$

The law of excluded gambling system:

$$\lim_{n \to \infty} \frac{m(A)}{a(n)}$$

Andrey Kolmogorov (1903–1987)

First axiom

The probability of an event is a non-negative real number:

$$P(E) \in \mathbb{R}, P(E) \geq 0 \qquad \forall E \in F$$

where F is the event space and E is any event in F. In particular, $P(E)$ is always finite, in contrast with more general measure theory. Theories which assign negative probability relax the first axiom.

Second axiom

This is the assumption of **unit measure**: that the probability that some elementary event in the entire sample space will occur is 1. More specifically, there are no elementary events outside the sample space.

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

This is often overlooked in some mistaken probability calculations; if you cannot precisely define the whole sample space, then the probability of any subset cannot be defined either.

Third axiom

This is the assumption of σ-additivity:

Any countable sequence of disjoint (synonymous with *mutually exclusive*) $E_1, E_2, \ldots$ events satisfies

$$P(E_1 \cup E_2 \cup \cdots) = \sum_{i=1}^{\infty} P(E_i).$$

Some authors consider merely finitely additive probability spaces, in which case one just needs an algebra of sets, rather than a σ-algebra.Quasiprobability distributions in general relax the third axiom

Birthday Problem

- Pigeonhole theorem vs. Probability Theory

n	P(n)
20	41.1%
23	50.7%
30	70.6%
50	97%
57	99%
100	99.99997%
200	99.9999999999999999999999999998%

Refrences

- Pathways to Probability, Amy C. King/Cecil Read, 1963.
- An Objective Theory of Probability, Gillies, 1973.
- wikipedia.com