Ann. Acad. Rom. Sci.
Ser. Math. Appl.
ISSN 2066-6594 Vol. 17, No. 3/2025

ON BERTRAND'S PARADOX AND A BRIDGE BETWEEN THE PRINCIPLE OF INDIFFERENCE AND THE FREQUENTIST PROBABILITY*

Luigia Caputo† Aniello Buonocore‡

Dedicated to Prof. Liliana Restuccia on the occasion of her 70th anniversary

DOI 10.56082/annalsarscimath.2025.3.369

Abstract

The "Bertrand's paradox" arises from a geometric probability problem for which J. Bertrand provides three solutions with different values of the probability sought; from here the "paradox" arises, a term that several authors attribute to J. Poincaré. The interest in this issue has lasted for over a hundred years and even today it is possible to find articles in which Bertrand's paradox, in one way or another, is take into consideration. Indeed, it as well as in the philosophical debate concerning the opposition between the frequentist and Bayesian approaches to Inferential Statistics, it has also used to highlight (hypothetical) logical inconsistencies of the "principle of indifference" in problems in which the total number of cases is not countable.

The Principle of Indifference, whose origin is, often, attributed to P. S. Laplace, is a fundamental concept in the field of probability but also in decision-making under uncertainty. It offers a rational starting point for assigning probabilities in the absence of any information. However, it is essential to be aware of the importance of updating our beliefs as added information about the issue becomes available.

*Accepted for publication on November 26, 2025

†`luigia.caputo@unina.it`, Università di Napoli Federico II–Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"

‡`aniello.buonocore@unina.it`, Università di Napoli Federico II–Dipartimento di Matematica e Applicazioni "Renato Caccioppoli"

In this article, in order to resolve Bertrand's paradox, we highlight the role, often overlooked, of the random experiment (and also of the random device for its implementation) that generates the support of a probability space. Afterwards, from the problem itself it is possible to trace the class of generating events and a pre-measure on this class to complete a probability space consistent with the problem. This way of proceeding is illustrated for each of the three solutions identified by Bertrand and for another recently proposed solution. We conclude that all solutions are equally valid because, once the appropriate probability space has specified, the only correct solution will emerge in a logical and formal way.

Keywords: historical definitions and paradoxs, geometric probability, combinatorial probability.

MSC: 60-00, 60-03, 60C05.

1 Introduction

This article starts from the criticism formulated by J.M. Keynes in [9] with respect to "reckless" application in probability of the *Principle of Indifference* (PI),[1] and from the ever-current interest (see, for example, [17], [8], [4], [14]) with regard to a geometric probability problem formulated by Bertrand [2] and known as *Bertrand's problem*. The criticism and the problem are linked by the fact that the latter, in Keynes' opinion, would provide reasonable evidence to refute the (indiscriminate) adoption of the PI in the evaluation of the probability of an event.

According the line depicted by Keynes in the PI asserts that "if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability. Thus equal probabilities must be assigned to each of several arguments, if there is an absence of positive ground for assigning unequal ones".

With regard to *Bertrand's paradox* (BP),[2] the question is simple: Bertrand introduces an event B related to the "*random choice*" of a chord of a given circle: "*the random chord is less than of the side of the inscribed equilateral triangle*" in order to evaluate its probability. Bertrand finds three possible solutions that lead to different values of the required probability: hence the "paradox". Such a paradox, at first, is ascribed to the fact that the

[1] Keynes states that drew many of the arguments in support of his criticism from J. von Kries' volume [18].

[2] This term, in the opinion of many authors, is due to J. Poincaré.

chords admit cardinality greater than the countable and, in these situations, the PI could not be applied. However, in the following, further aspects have been considered and analyzed.

In [6] E. T. Jaynes wonders what the correct solution is, and among the authors he cites E. Borel as the only one who expresses a clear preference for one of those proposed by Bertrand (see [3]) and L. von Mises as the one who takes a strong position by declaring that this kind problems (including the similar problem of the "Buffon's Needle", briefly described below) are not within the competence of probability theory. The others, including Bertrand, are in an intermediate position asserting that, due to the presence of "at random" wording in its statement, the problem is "ill-posed", so that it cannot have a well-defined solution. On this argument, Jaynes says:

> *"In works on probability theory this state of affairs has been interpreted, almost universally, as showing that the principle of indifference must be totally rejected. Usually, there is the further conclusion that the only valid basis for assigning probabilities is frequency in some random experiment. It would appear, then, that the only way of answering Bertrand's question is to perform the experiment."*[3]

Incidentally, Jaynes' suggestion was taken into account in [5]; here the authors propose to adopt a physical experiment, known as "Buffon's Needle", as described by G.-L. Leclerc de Buffon in his 1777 *Essai D'Arithmétique Morale*.[4] The experiment consists of throwing a straw of length δ on a floor on which parallel lines are drawn at a distance d from each other. The aims is to calculate the probability that the straw will not intersect any of the parallel lines. In [5], the authors replace the parallel lines with a circle Γ of radius r and use the symbol L instead of the symbol δ.[5] Then the probability that the straw will find a chord of length greater than the side of the equilateral triangle inscribed in Γ, given that it intersects the circumference at two points, is calculated as a function of the ratio L/r.[6] In

[3]Hence, there WAS a clear preference of frequencies over the PI.

[4]This experiment was later used to obtain an estimate of $1/\pi$ and was thus the forerunner of the "Monte Carlo Method". Buffon's intent was to propose betting games to be played on a table and, for this purpose, he replaces the straw with a sewing needle or a headless pin; hence the name of the experiment.

[5]M.J. Stoka in [16] was among the first to use variants of the experiment of the "Buffon's Needl" with regard to problems of *geometric probability.*

[6]Note that the probability required by Bertrand's problem is the 1's complement of this probability.

our opinion, to fall within the scope of Bertrand's problem it is necessary to get rid of the additional parameter L and, given the radius r of the circle, this can only be achieved by sending L to infinity. In [5] it can be seen that this asymptotic probability is equal to $1/2$ (that coincides with probability obtained in the second Bertrand's solution): probably due their choice to consider only parallel chords.

However, Jaynes states that the correct solution to the Bertrand's problem must satisfy three invariance properties ("rotational", "positional" and "translational").[7] In this way, Bertrand's problem is well-posed since there is only one of the three solutions for which all the invariance properties hold, thus saving the application of the PI. In [11] L. Marinoff proposes a resolution of BP, by taking as a fundamental element the procedures that generate the random chords: thus, he makes the three solutions, as others proposed by himself, equally valid; furthermore he concludes his analysis by stating that the PI can be applied to infinite sets as long as the problems are formulated in an unambiguous way.

Among the authors of the most recent articles in which there's a refutation of the PI, we cite N. Shachel, who in [15] considers the strategy of making Bertrand's problem well posed unsuccessful, while those who argue to have positively resolved BP are (i) N. Wang and R. Jachson that in [19] believe that they have rigorously demonstrated that two of the three solutions are not valid due to the presence of false hypotheses concerning them and that only one solution is correct – which, incidentally, is the same as that identified by Jaynes – (ii) D. Aerts and M. Sassoli de Bianchi that in [1] identify two Bertrand's problems, one of which is "easy" and the other is "difficult", so that once the easy problem has been solved, even the difficult one becomes solvable under an appropriate condition – (iii) R. A. Chechile [4] finds a solution to Bertrand's problem which does not coincide with any of the three proposed by Bertrand (and others subsequently proposed), which are the result of not perfectly random chords.

Furthermore, in [8] another solution to Bertrand's problem is proposed while in [17] a new procedure for generating a random chords is presented, the solution of which coincides with that of the third proposed by Bertrand. Finally, in [14] the paradox is analyzed in detail and with an emphasis on the philosophy of probability that the author considers largely neglected.

The article is organized as follows. In section 2 the PI criticism of Keynes is presented. Then, we describe: the three solutions offered by Bertrand;

[7] Another widely used earlier proposal by Jaynes to try to transform an ill-posed problem into a well-posed one is based on the "principle of maximum entropy" (see [7]).

possible random devices for generating random chords for each of the three solutions;[8] a simple resolution of BP (based on the axiomatic theory of Kolmogorov) with the conclusion that the PI does not suffer from any logical difficulty and can be applied in both the discrete and continuous cases. In section 3, some counting problems related to an n-dimensional sample space and their solutions are recalled. In section 4, we prove that starting from a uniform probability distribution with finite support, S,[9] by virtue of Bayes' formula and a random sample from S, we obtain the frequency distribution of the sample. Finally, a short summary section is presented.

2 On the principle of indifference and the fundamental role of the random experiment

Pierre Simon Laplace considered, without any label, the PI intuitively obvious and in [10] (pag. 7) he wrote:[10]

> *"The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all the cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible."*

The term *Principle of Indifference* appears for the first time in [9] where chapter IV is entirely dedicated to this subject. Here the author, J.M. Keynes, first provides the reasons

> *"The discovery of a rule, by which equiprobability could be established, was, therefore, essential. A rule, adequate to the purpose, introduced by James Bernoulli,*[11] *who was the real founder of*

[8] A random device is any artifact that can produce more than one outcome; for example, a coin, a die, a deck of cardstocks, a true random number generator (TRNG).

[9] That is, applying the PI to the elements of S.

[10] The translation here quoted is from the English edition (page 6) by F. W. Truscott and F. L. Emory and published in 1902 by John Wiley & Sons and Chapman & Hall.

[11] Also remembered as Jacques, Jakob or Giacomo.

> *mathematical probability, has been widely adopted, generally under the title of* The Principle of Non-Sufficient Reason*, down to the present time. This description is clumsy and unsatisfactory, and, if it is justifiable to break away from tradition, I prefer to call it* The Principle of Indifference*."*

then he enunciates the PI

> *"The Principle of Indifference asserts that if there is no known reason for predicating of our subject one rather than another of several alternatives, then relatively to such knowledge the assertions of each of these alternatives have an equal probability. Thus equal probabilities must be assigned to each of several arguments, if there is an absence of positive ground for assigning unequal ones."*

Continuing in his exposition, Keynes treats some examples (i.e. the *paradox of color book*, the problem of the specific volume of a given substance, the Bertand's paradox), which in his opinion profoundly undermine the effective application of the PI.

Fortunately, in 1933 A. N. Kolmogorov, taking inspiration from Lebesgue's works on the theory of measure and integration, published the monograph "*Grundbegriffe der Wahrscheinlichkeitrechnung*" ("Foundations of the Theory of Probability"), definitively marking the birth of the Probability theory; in fact, in the introduction Kolmogorov writes:

> *"The purpose of this monograph is to give an axiomatic foundation for the theory of probability. The author set himself the task of putting in their natural place, among the general notions of modern mathematics, the basic concepts of probability theory – concepts which until recently were considered to be quite peculiar."*

However, in our opinion, it is still necessary to highlight an aspect that can give rise to erroneous interpretations. In fact, there is a substantial difference between measure and probability theory: in the first, the support set, S, of the measure space is assigned and does not need further attention, while in the second, the probability space support is the set Ω of possible outcomes of a random experiment $\mathcal{E}$ and the events are some of its subsets. Ultimately, $\mathcal{E}$ generates Ω and, once atomic events have been identified, the σ-algebra $\mathcal{F}$

of the events must also be specified: therefore Ω must be specified in detail. Thus, the description of the random device $\mathcal{D}$ that performs $\mathcal{E}$ becomes even more important.

Consider now the paradox of color book, the first example proposed by Keynes in his criticism of the Principle of Indifference:

> *"If, for instance, having no evidence relevant to the colour of this book, we could conclude that 1/2 is the probability of "This book is red" we could conclude equally that the probability of each of the propositions 'This book is black' and 'This book is blue' is also 1/2. So that we are faced with the impossible case of three exclusive alternatives all as likely as not."*

The context illustrated by Keynes, right from the incipit "*If, for instance, having no evidence relevant to the colour of this book*", does not fall within the perimeter established by the axiomatic Kolmogorov theory because, in it, all the possible colors for the cover of the book must be known: otherwise it is not possible to establish who is Ω, and then, who is $\mathcal{F}$. Furthermore, not only it is necessary to know Ω and, in particular, that $|\Omega| = m \in \mathbb{N}$, but also in what random way the typographer establishes the element of Ω for the cover of the newly composed book.

Currently, the paradox of color book is resolved by applying the PI when the $\mathcal{D}$ device for the choice of color consists, for example, of a deck of m cardstocks, well mixed on a table, on each of which is shown on the front one of the colors of Ω, from which the typographer takes the one on top. Thus, if red is a color of Ω the probability that the cover is red is $1/m$, while its negation has probability $(m-1)/m$.

Regarding to Bertrand's problem,[12]

> *"...:if a chord in a circle is drawn at random, what is the probability that it will be less than the side of the inscribed equilateral triangle."*

it can be observed that in its formulation there is the characteristic of being a "ill-posed" problem. Keynes himself, quoting Bertrand, recognizes this characteristic writing

> *"(a) It is indifferent at what point one end of the chord lies. If we suppose this end fixed, the direction is then chosen at random.*

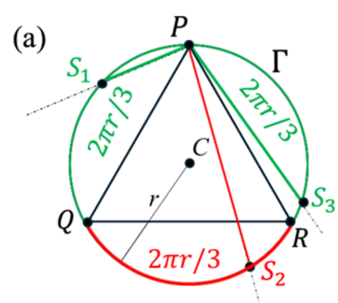
[12] For which it is not possible to use random devices such as cards, coins, dice, etc.

> *In this case the answer is easily shown to be* 2/3*."*
>
> *"(b) It is indifferent in what direction we suppose the chord to lie. Beginning with this apparently not less justifiable assumption, we find that the answer is* 1/2*."*
>
> *"(c) To choose a chord at random, one must choose its middle point at random. If the chord is to be less than the side of the inscribed equilateral triangle, the middle point must be at a greater distance from the centre than half the radius. But the area at a greater distance than this is* 3/4 *of the whole. Hence our answer is* 3/4*."*

but he ascribes the different evaluations of the probability of the considered event to the pernicious application of the PI, while, in our opinion, it only depends on the random device with which the chord is obtained in a random way.

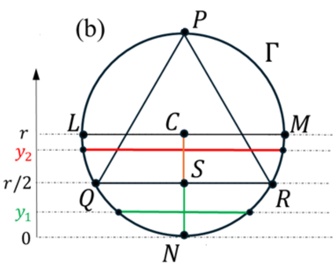
In fact, denoted by Γ the circumference with radius r and by γ the circle enclosed in it, for the situation described in the previous list, the result is:[13]

(a) The random device $\mathcal{D}_{(a)}$ consists of a straw of length greater than 2π and with one of its endpoints located on a prefixed point of Γ; the straw can rotate around this point (e.g., P in the figure on the right). Thus the chord is determined by the point of intersection of Γ with the straw when it stops after it has been vigorously operated: for example, S_1, S_2 and S_3 in the figure on the right. Note that point S_2, as all the points on the arc from Q to R (included), is not favorable to the event B specified in Bertand's problem.

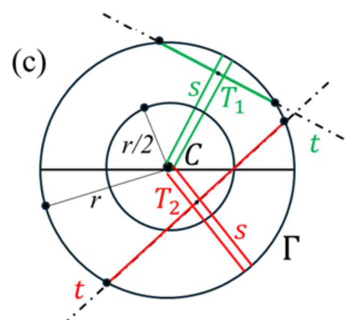


[13]Each of the following three experiments must be interpreted as a "conceptual experiment" (Gedankenexperiment), in the same way as Hilbert's Hotel with infinite rooms or the "Brownian pawl" (or "Feynman-Smoluchowski pawl") device that was analyzed for the first time by M. Smoluchowski and made popular by R. Feynman. Furthermore, in the following figures, the green (red) chords are less (not less) than the side of the equilateral triangle inscribed in Γ.

(b) The $\mathcal{D}_{(\mathrm{b})}$ random device consists of a straw of length greater than $2r$, parallel to one side of the inscribed equilateral triangle (for example, the side QR in the figure on the right); the straw can freely move, maintaining the same direction, from its point of contact with Γ (the point N in the figure on the right) until it overlaps the corresponding diameter (the segment LM in the figure on the right) and vice versa. So the chord is determined by the ordinate of the point of intersection of the straw with the orthogonal radius (the segment CN in the figure on the right) when it stops after being vigorously operated. Note that $y_2 \in (0, r]$ is an ordinate that is not favorable to the event B.



(c) The $\mathcal{D}_{(\mathrm{c})}$ random device consists of a straw, s, of length r pivoted on the center C of Γ, and of another straw, t, of length greather than $2r$ centered on s (see figure on the right): while s is rotated, t moves along s. When the device stop after being vigorously operated, the midpoint of s provides the desired random point of γ. Note that $T_2 \in \gamma$ is a point not favorable to the event B.



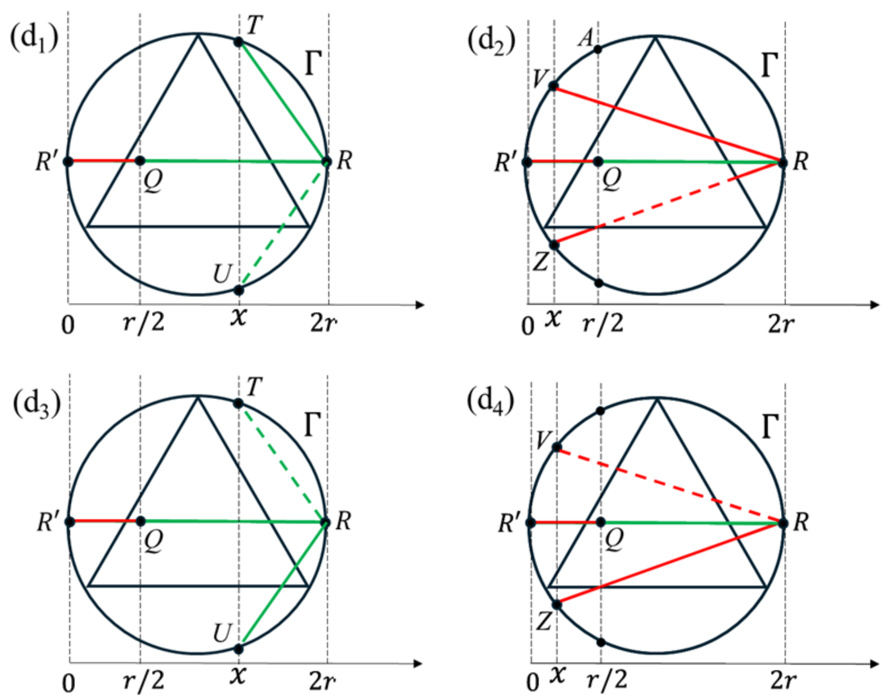
Consequently, in our opinion, BP is resolved by the observation that each of the three solutions needs a appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$:

(a) the support $\Omega_{(\mathrm{a})}$ coincides with Γ, the σ-algebra of the events $\mathcal{F}_{(\mathrm{a})}$ is generated by the arcs of Γ (right end point included) and the probability measure $\mathbb{P}_{(\mathrm{a})}$ is the normalized extension on $\mathcal{F}_{(\mathrm{a})}$ of the (uniform) pre-measure on the class of the arc of Γ (the random choice of a chord means that any two of these arcs, having equal amplitude l, have the same pre-measure value which is proportional to l);

(b) the support $\Omega_{(\mathrm{b})}$ coincides with the interval $(0, r]$, the σ-algebra of the events $\mathcal{F}_{(\mathrm{b})}$ is the Borel σ-algebra of $(0, r]$ and the probability measure $\mathbb{P}_{(\mathrm{b})}$ is the normalized Lebesgue measure on $\mathcal{F}_{(\mathrm{b})}$ (the random choice of a chord means that any two subintervals of $(0, r]$, having equal amplitude h, have the same probability which is proportional to h);

(c) the support $\Omega_{(\mathrm{c})}$ coincides with γ, the σ-algebra of the events $\mathcal{F}_{(\mathrm{c})}$ is the Borel σ-algebra of γ and the probability measure $\mathbb{P}_{(\mathrm{c})}$ is the normalized Lebesgue measure on $\mathcal{F}_{(\mathrm{c})}$ (the random choice of a chord means that any two circle inside γ, having equal radius ρ, have the same probability which is proportional to ρ).

To accompany the exposition with another example, we consider also the aforementioned solution proposed in [8] (solution (d) in the sequel); see the following figure, in which r represents the radius of Γ.



The R point is one end of the chord. The other end point is determined by randomly choosing a real x in the interval $(0, 2r]$ (this can be achieved with a similar mechanism described in the solution (b); in this case a vertical straw moves from R to R' and vice versa) and considering the two points of intersection of Γ with the perpendicular to the diameter $R'R$ passing through the point of abscissa x: points T and U in Figures (d_1,d_3) and points V and

Z in Figures (d_2,d_4). With another independent random device $\widehat{\mathcal{D}}$ — a fair coin, a fair die, etc. — one establishes which of these two points should be considered; for example, by virtue of $\hat{D}$ in Figure (d_1) the chord RT has been chosen, while, in Figure (d_3) the chord RU has been chosen.[14] It is quite easy to verify with elementary geometric reasoning that for values of x belonging to the interval $(r/2, 2r]$ the length of the chord is less than the side of the inscribed equilateral triangle. Therefore, the event B has probability $\frac{2r-r/2}{2r} = 3/4$. This fact is interesting because it shows that two completely different solutions of Bertrand's problem, solutions (c) and (d), determine the same probability value for event B.

After this, one has:

(d) the support $\Omega_{(\mathrm{d})}$ coincides with the interval $(0, 2r]$, the σ-algebra of the events $\mathcal{F}_{(\mathrm{d})}$ is the Borel σ-algebra of $(0, 2r]$ and the probability measure $\mathbb{P}_{(\mathrm{d})}$ is the normalized Lebesgue measure on $(0, 2r]$ (the random choice of a chord means that any two subintervals of $(0, 2r]$, with equal amplitude δ, have also the same probability which is proportional to δ).

We explicitely emphasize that the proposal described above only aims to formalize an idea already introduced by various authors in articles and books. For example, in [3] (pag. 110), Borel writes:[15]

> *Should we think that these three solutions are equally good and, consequently, equally bad? Not at all,* it is simply a matter of specifying the method by which the experimental verification will be carried out*, that is to say, how we can draw a chord* at random *in a circle: if we force this chords to pass through a fixed point of the circle, or if we fix its midpoint* at random*, it is the second or the third solution that is the good one; but it is easy to see that most of the natural processes that we can imagine lead to the first one.*[16]

3 Some positions and preliminary calculations

Let n, k be two positive integers and let $\mathcal{E}$ be the random experiment consisting of rolling a k-face die for which the manufacturing house guarantees

[14] This further choice is necessary only in order not to neglect the chords that lie below the considered diameter $R'R$.

[15] The English translation is performed by the present authors.

[16] The solution that Borel refers to corresponds to the solution at point b).

perfect balance.[17] Thus, applying the PI, we have:

$$j \in \Omega := \{1, 2, \dots, k\}, \quad p_j := \mathbb{P}(j) = \frac{1}{|\Omega|} = \frac{1}{k} =: p.$$

On the Ω space of the results of $\mathcal{E}$ consider the number X uniformly distributed over the (finite) support $S_X = \{\xi_1, \xi_2, \dots, \xi_k\} \subset \mathbb{R}$:

$$j \in \Omega, \quad X(j) = \xi_j \quad \text{and} \quad \mathbb{P}(X = \xi_j) = \mathbb{P}(j) = \frac{1}{k}.$$

For the support of X it results,

$$|S_X| = k. \tag{1}$$

Repeating the $\mathcal{E}$ experiment n times under the same conditions yields a simple random sample (i.e., with independent and identically distributed observations):

$$\underline{X} = (X_1, X_2, \dots, X_n), \quad \text{with } X_1 \sim X.$$ [18]

For the (finite) support $S_{\underline{X}}$ of $\underline{X}$ it results,

$$S_{\underline{X}} = S_X^n \subset \mathbb{R}^n \quad \text{and} \quad |S_{\underline{X}}| = D_{k,n}^{(\mathrm{r})} = k^n, \tag{2}$$

because, in order to construct the generic element of $S_{\underline{X}}$ we must choose (in sequence and with repetition)) any of the k elements of S_X for each observation of $\underline{X}$.[19]

For our purposes it is also useful to consider the random vector of relative frequencies

$$\underline{N} = (N_1, N_2, \dots, N_k),$$

where with $j \in \Omega$, $N_j \sim \mathrm{Bin}(n, p_j)$, $N_1 + N_2 + \cdots + N_k = n$ and $\underline{N}$ has a multinomial law.
For the support of $\underline{N}$ it results,

$$S_{\underline{N}} = \left\{ \underline{n} = (n_1, n_2, \dots, n_k) \in \mathbb{N}_0^k : n_1 + n_2 + \cdots + n_k = n \right\},$$

[17] Then, under this guarantee $\mathcal{E}$ is a random experiment.

[18] The symbol $\sim$ enclosed between two random numbers indicates that they have the same law; the same meaning applies in the case where a law of probability appears instead of the second random number.

[19] With $D_{k,n}^{(\mathrm{r})}$ has been denoted the number of n-arrangements with repetition consisting of the elements of a set with cardinality k.

and

$$|S_{\underline{N}}| = \binom{k+n-1}{n} = C^{(\mathrm{r})}_{k,n}, \tag{3}$$

because here the counting problem is equivalent to that of determining how many ways indistinguishable objects can be placed in k drawers, each of which can contain from 0 to n.[20]

Finally, the $C^{(\mathrm{r})}_{k,n}$ points of $S_{\underline{N}}$ partition $S_{\underline{X}}$ into classes defined as follows:

$$\begin{aligned} &\underline{n} = (n_1, n_2, \dots n_k) \in \underline{N}, \\ &\qquad C_{\underline{n}} := \{\underline{x} \in S_{\underline{X}} : j \in \Omega,\ \ \xi_j \text{ occurs } n_j \text{ times}\}. \end{aligned} \tag{4}$$

Obviously, it turns out:[21]

$$|C_{\underline{n}}| = P^{(\mathrm{r})}_{n;n_1,n_2,\dots,n_k} = \frac{n!}{n_1!n_2!\ \cdots\ n_k!}. \tag{5}$$

It is explicitly noted that the sum of the cardinalities of all $C_{\underline{n}}$ coincides with the cardinality of $S_{\underline{X}}$:

$$\sum_{\substack{n_1=0 \\ n_1+n_2+\cdot+n_k=n}}^{n} \sum_{n_2=0}^{n} \cdots \sum_{n_k=0}^{n} \frac{n!}{n_1!n_2!\ \cdots\ n_k!} = k^n. \tag{6}$$

In fact, the (6) is true for $k=2$:[22]

$$\begin{aligned} \sum_{\substack{n_1=0 \\ n_1+n_2=n}}^{n} \sum_{n_2=0}^{n} \frac{n!}{n_1!n_2!} &= \sum_{n_1=0}^{n} \frac{n!}{n_1!(n-n_1)!} \\ &= \sum_{n_1=0}^{n} \binom{n}{n_1} \\ &= \sum_{n_1=0}^{n} \binom{n}{n_1} 1^{n-n_1} \cdot 1^{n_1} = (1+1)^n = 2^n; \end{aligned}$$

[20]With $C^{(\mathrm{r})}_{k,n}$ has been denoted the number of n-combinations with repetition consisting of the elements of a set with cardinality k.

[21]With $P^{(\mathrm{r})}_{n;n_1,n_2,\dots,n_k}$ has been denoted the number of n-permutations with repetition having n_1 symbols equal to ξ_1, n_2 symbols equal to $\xi_2, \dots$, n_k symbols equal to ξ_k.

[22]Eq. (6) is also true for $k=1$ as the first and second members coincide with 1.

furthermore, assuming that the (6) is true in the case where S_X contains $k-1$ elements, we have:

$$\sum_{\substack{n_1=0\\ n_1+n_2+\ldots+n_k=n}}^{n} \sum_{n_2=0}^{n} \cdots \sum_{n_k=0}^{n} \frac{n!}{n_1!n_2!\ \cdots n_k!} =$$

$$= \sum_{n_1=0}^{n} \frac{n(n-1)\cdots(n-n_1+1)}{n_1!} \sum_{\substack{n_2=0\\ n_2+\ldots+n_k=n-n_1}}^{n-n_1} \cdots \sum_{n_k=0}^{n-n_1} \frac{(n-n_1)!}{n_2!\ \cdots n_k!}$$

$$= \sum_{n_1=0}^{n} \binom{n}{n_1} (k-1)^{n-n_1} \cdot 1^{n_1} = [(k-1)+1]^n$$

$$= k^n.$$

Hence, the result (6) follows from the principle of induction.

4 From the principle of indifference to frequencies

Taking into account the (2), the (5) and the fact that all points of $S_{\underline{X}}$ have the same probability of occurring, by virtue of the PI we have:

$$\underline{n} = (n_1, n_2, \ldots n_k) \in S_{\underline{N}},$$
$$\mathbb{P}\left(\underline{N} = \underline{n}\right) = \frac{n!/(n_1!n_2!\ \cdots\ n_k!)}{k^n}. \tag{7}$$

Note that the following formula is also true

$$j \in \Omega, \quad \underline{n} = (n_1, n_2, \ldots n_k) \in S_{\underline{N}},$$
$$\mathbb{P}\left(\underline{N} = \underline{n} | X_n = \xi_j\right) = \frac{(n-1)!/[n_1!\ \cdots\ n_{j-1}!(n_j-1)!\ \cdots\ n_k!]}{k^{n-1}}, \tag{8}$$

because we can neglect the last component of $\underline{X}$, due to the fact that $X_n = \xi_j$, and the previous reasoning is repeated.

Then, taking into account the (7) and(8), by virtue Bayes' formula, we have:

$$
\begin{aligned}
&j \in \Omega, \quad \underline{n} = (n_1, n_2, \dots n_k) \in S_{\underline{N}} : n_j > 0, \\
\mathbb{P}\left(X = \xi_j | \underline{N} = \underline{n}\right) &= \mathbb{P}\left(X_n = \xi_j | \underline{N} = \underline{n}\right) \\
&= \frac{\mathbb{P}\left(\underline{N} = \underline{n} | X_n = \xi_j\right)}{\mathbb{P}\left(\underline{N} = \underline{n}\right)} \mathbb{P}(X_n = \xi_j) \\
&= \frac{1}{k} \frac{\mathbb{P}\left(\underline{N} = \underline{n} | X_n = \xi_j\right)}{\mathbb{P}\left(\underline{N} = \underline{n}\right)} \\
&= \frac{1}{k} \frac{\{(n-1)!/\left[n_1! \ \cdots \ n_{j-1}!(n_j - 1)! \ \cdots \ n_k!\right]\}/k^{n-1}}{\left[n!/(n_1! n_2! \ \cdots \ n_k!)\right]/k^n} \\
&= \frac{n_j}{n}.
\end{aligned}
\tag{9}
$$

It is explicity note that Eq. (9) also holds in the case $n_j = 0$; indeed under $\{\underline{N} = \underline{n}\}$ the event $\{X = \xi_j\}$ is impossible.

From a probabilistic point of view, it can be said that the PI used by Laplace for his definition of the probability of an event is the first example of the application of Bayesian reasoning. The uniform law for the number X is the one that in the Bayesian context is called "a priori": it photographs the current state of knowledge as the k-face die is guaranteed to be perfectly balanced by the manufacturer. The realization $\underline{n}$ of (absolute) frequencies provides an additional information for which it is necessary to update the law of X:

$$
\begin{aligned}
&\underline{n} = (n_1, n_2, \dots n_k) \in S_{\underline{N}}, \\
&X | \{\underline{N} = \underline{n}\} \sim \begin{pmatrix} \xi_1, & \xi_2, & \dots, & \xi_k \\ \frac{n_1}{n}, & \frac{n_2}{n}, & \dots, & \frac{n_k}{n} \end{pmatrix},
\end{aligned}
\tag{10}
$$

which is the so-called "a posteriori" law of X.

In summary:

$$
\begin{aligned}
&j \in \Omega, \quad \underline{n} = (n_1, n_2, \dots n_k) \in S_{\underline{N}}, \\
&X(j) = \xi_j \quad \text{and} \quad \mathbb{P}(X = \xi_j) = \frac{1}{k}, \\
&\mathbb{P}\left(X = \xi_j | \underline{N} = \underline{n}\right) = \frac{n_j}{n}.
\end{aligned}
\tag{11}
$$

5 Conclusions

The main purpose of the present article was to give a contribution to the resolution of BP. The idea, illustrated in detail in the second section, is the following: in order to proceed with the calculation of the probability of event B considered by Bertrand (a random chord is less than the side of the equilateral triangle inscribed in a given circle), an appropriate probability space $(\Omega, \mathcal{F}, \mathbb{P})$ must be considered. This is achieved by first introducing a suitable random device $\mathcal{D}$, whether physical or conceptual, with which the experiment $\mathcal{E}$ of obtaining a random chord is carried out. After that,

1 Ω is the set of all possible outcomes of $\mathcal{E}$;

2 $\mathcal{F}$ is the σ-algebra generated by a family $\mathcal{A} \subset \Omega$ chosen in such a way that: (i) B actually belongs to $\mathcal{A}$ and (ii) on $\mathcal{A}$ it is possible to identify an pre-measure $\mathbb{P}_{\mathcal{A}}$ consistent with the request of the random choice of the chord;

3 $\mathbb{P}$ is the probability measure that is obtained by normalizing the extension on $\mathcal{F}$ of the pre-measure $\mathbb{P}_{\mathcal{A}}$.

In conclusion, there may be numerous solutions to Bertrand's problem since its modelling (deriving from the axiomatic approach of Kolmogorov) can be completely different and this justifies the different values of the probability already found. Of course, there is nothing to prevent two different probability spaces from leading to the same value of the probability.

Another notable point that has been addressed in this article has been to highlight the intimate connection between the application of the PI and the use of relative frequencies when one wants to evaluate the probability of an event.[23] This connection is obtained by using Bayes' formula: it surprisingly enough establishes that the distribution of relative frequencies is the "a posteriori" distribution of a uniform law assigned by applying the PI. In other words, the frequency distribution is an upgrade of a uniform law when the information obtained from a simple random sample is translated into a second-level representation (the frequency distribution, in fact).

Acknowledgement. Paper written with financial support of project PRIN2022-MUR 2022XZSAFN.

References

[1] D. Aerts and M. Sassoli de Bianchi, Solving the hard problem of Bertrand's paradox, *J. Math. Phys.* 55 (2014), 083503.

[23] It is explicitly noted that while the PI is applicable, under the conditions deemed appropriate, in all random experiments, to use the frequencies it is necessary to be in the presence of a random experiment that can be repeated indefinitely.

[2] J. Bertrand, *Calcus des Probabilités*, Gauthier-Villars et Fils (Paris), 1907.

[3] E. Borel, *Éléments de la Théorie des Probabilités*, Librairie Scientifiques A. Hermann & Fils (Paris), 1909.

[4] R.A. Chechile, Bertrand's paradox resolution and its implications for the Bing-Fisher problem, *Mathematics* 11 (2023), 083503.

[5] P. Di Porto, E. Crosignani, A. Ciattoni and H.C. Liu, Bertrand's paradox: a physical way out along the lines of Buffon's needle throwing experiment, *Eur. J. Phys.* 32 (2011), 819-825.

[6] E.T. Jaynes, The well-posed problem, *Found. Phys.* 33 (1973), 477-493.

[7] E.T. Jaynes, Where do we stand on maximum entropy? In: *The Maximum Entropy Formalism*, R.D. Levine and M. Tribus (eds), MIT Press, Cambridge MA, 1979.

[8] P. Kaushik, A new solution of Bertrand's paradox, *Theory Probab. Appl.* 67 (2022), 199-202.

[9] J.M. Keynes, *A Treatise of Probability*, Mc Millan and Co (London), 1921.

[10] P.S. Laplace, *Essai Philosophique sur les Probabilités*, Courcier (Paris), 1814.

[11] L. Marinoff, A resolution of Bertrand's paradox, *Philos. Sci.* 61 (1994), 1-24.

[12] J. Poincaré, *Calculus des Probabilités*, Gauthier-Villars (Paris), 1912.

[13] R. Rosa, *Lezioni sul Metodo di Monte Carlo*, Editrice Lo Scarabeo (Bologna), 1992.

[14] N. Shachel, *Bertrand's Paradox and the Principle of Indifference*, Routledge, Taylor & Francis Group (New York and London), 2024.

[15] N. Shachel, Bertrand's paradox and the principle of indifference, *Philos. Sci.* 74 (2007), 150-175.

[16] M.I. Stoka, Probabilités géométriques de type "Buffon" dans le plan euclidien, *Atti della Accademia delle Scienze di Torino* 110 (1976), 53-59.

[17] Z. Vidović, Random chord in a circle and Bertrand's paradox: new generation method, extreme behaviour and length moments, *Bull. Korean Math. Soc.* 58 (2021), 433-444.

[18] J. von Kries, *Die Principien der Wahrscheinlichkeitsrechnung. Eine logische Untersuchung*, J.C.B. Mohr (Freiburg), 1886.

[19] N. Wang and R. Jackson, Resolving Bertrand's probability paradox, *Int. J. Open Problems Compt. Math.* 4 (2011), 72-103.