Experimental tests of Bertrand’s question and the Duhem–Quine problem

Zhenning Liu and Charles S Adams

Joint Quantum Center (JQC) Durham–Newcastle, Department of Physics, Durham University, South Road, Durham, DH1 3LE, United Kingdom

E-mail: zhenning_liu@outlook.com and c.s.adams@durham.ac.uk

Received 10 June 2019, revised 26 July 2019
Accepted for publication 8 August 2019
Published 18 September 2019

Abstract

In this paper we report on an experimental test of Bertrand’s question on the probability to find a random chord drawn inside a unit-radius circle with length greater than \(3\). In an experiment performed by tossing straws onto a circle, we confirm a theoretical prediction that the answer depends on the ratio of the circle diameter, \(2R\), to the straw length, \(L\), and that the special case, which follows from rotational and translation invariance using integral geometry, is only obtained in the experimentally unattainable limit of infinite straw length, \(\bar{d} = \frac{2R}{L} \to 0\). In addition, we observe a systematic discrepancy in the limit, \(\bar{d} = \frac{2R}{L} \to 1\), where a large number of events are rejected. We conclude that the experimental test of Bertrand’s paradox provides a good illustration of the Duhem–Quine problem: that hypothesis testing is always conditional on a bundle of real auxiliary assumptions.

Keywords: Bertrand’s paradox, Duhem–Quine problem, probability, experimental physics

(Some figures may appear in colour only in the online journal)

1. Introduction

In 1889 Bertrand [1] asked the purely mathematical question: if one draws a chord inside a circle at random, then what is the probability for its length to be longer than the side of the inscribed equilateral triangle? He gave three answers, \(1/4\), \(1/3\) and \(1/2\), depending on how we choose to interpret the words at random; see figure 1. In 1973 Jaynes [2] proposed a ‘solution’ to this apparent ‘paradox’ based on invariance to other assumptions—sometimes called ‘the principle of maximum ignorance’. Jaynes’ argued that the absence of additional

1 Author to whom any correspondence should be addressed.
information leads us to Laplace’s principle of indifference—a ‘cornerstone’ of probability theory, with wide-ranging significance across the physical sciences from statistical and quantum physics [3] to econophysics [4], even though in practice, there is often no reason to prefer an indifference probability distribution over any other [5]. In reality, Jaynes’ asked a different question to Bertrand: a long straw is tossed at random onto a circle; given that it falls so that it intersects the circle, what is the probability that the chord thus defined is longer than a side of the inscribed equilateral triangle? This experimental question is distinct from Bertrand’s purely mathematical question, [6, 7] and requires additional assumptions, such as, how ‘long’ does the straw need to be relative to other length scales? The case of finite straw length has been analysed theoretically by Porto et al [8], who showed that the correct answer depends on the ratio of the circle diameter, $2R$, to the straw length, $L$, and only in the mathematical limit, $\bar{d} = 2R/L \to 0$, can we expect to recover Jaynes’ ‘maximum ignorance’ result.

The experimental form of Bertrand’s question is a good example of the so-called Duhem–Quine problem: the impossibility of testing a hypothesis without auxiliary assumptions [9–11]. In practice, the design and analysis of an experiment requires many real world assumptions. In addition to the finite length of real straws, the experimental test of Bertrand’s questions requires choices about how the random events are created, how to deal with failed attempts, and how we choose to record and analyse the data. Each of these assumptions plays a role in determining the assumed distribution and hence the final answer. As Bertrand’s ‘paradox’ is both historically significant in the developed of probability theory [2, 6], and more recently has become important in the context of an interesting class of metamaterials [12], such experimental and interpretative questions are significant. Although, Porto et al [8]
performed a numerical experiment, and there are related experimental studies on random coin toss ing [13], there have been no rigorous experimental tests of Bertrand’s question. As the experiment can be performed using simple equipment, exploit standard techniques in image and data analysis, and raises interesting interpretive question, it follows that systematic experimental investigations are long overdue.

In this letter, we perform an experimental test of Bertrand’s question. We show how the practicalities of the experiment force us to address the additional assumptions in a way that may not arise in theory or simulation. For example, we observe a systematic discrepancy in the limit, $\bar{d} = \frac{2R}{L} \rightarrow 1$, where a large number of events are rejected, which suggests that theoretically assumed distribution is not reproduced experimentally. We conclude that the experimental test of Bertrand’s ‘paradox’ is a useful test case to explore the Duhem–Quine problem that hypothesis testing is typically dependent on a bundle of interdependent assumptions.

2. Experimental test

The principle of the experiment is illustrated in figure 2. We throw spinning straws of length $L$ onto a large sheet of card, figure 2(a), where a circle of radius $R$ is drawn, figure 2(b). Immediately obvious from figure 2(a) is that there are two distinct random aspects of straw throwing. First, the angle of the straw is randomised (isotropy or rotational randomness), and second the point where the straw lands (which we can take as its mid-point) is also random (spatial homogeneity or translational randomness). In an idealised scenario, we could imagine the experiment as an analogue simulator of translational and rotational randomness—randomly choosing a point where one end of the straw touches the floor, and second randomly choosing a direction in which the straw topples. In this respect there is a close analogy with Buffon’s needle where a similar interplay between translational and rotational randomness is found [8].

To record each event, we take a photograph of the straw, figure 3(a), then draw virtual circles with a desired diameter, figure 3(b). The process is repeated 3600 times until all resulting images form the complete dataset for analysis. About 1200 out of 3600 lines obtained from fitting the images are shown in figure 3(c). We repeat the above with different radii to generate data with different parameters. For any particular radius, we make a histogram of the distribution of chord lengths and calculate the probability that the chord length is longer than the side length of an inscribed equilateral triangle. We might expect that the experiment would produce one of the three standard answers to Bertrand’s question shown in figure 1, and that the experiment will tell us which parameters are randomised in a random straw toss. However, the answer is more complex and depends on how we choose to analyse
the images. As both the circle size, $2R$, and position, can be varied in post processing this allows us to generate an infinite dataset from the original finite set of images. As highlighted by Porto et al [8] a key parameter in the experiment is the circle diameter to length ratio, $\tilde{d} = 2R/L$, and the answer to Bertrand’s question turns out to be a continuous function of this ratio [6, 8].

Not all trials produce a straw that completely overlaps with the circle as illustrated in figure 3(b), and it is necessary to post-select events where a chord is defined. This post-selection involves degrees of freedom about how we chose to analyse the data. In practice, we need to impose additional assumptions in order to obtain a result, and then we find that the result obtained depends on these assumptions. In selecting successful trials, there are three possible choices depending on whether we require zero, one or two intersections between the circle and the straw. We refer to these cases as the line, ray and segment viewpoints, respectively. The line viewpoint assumes that wherever the straw falls we can extend its length and if it crosses the circle, it is a valid trial. The ray viewpoint assumes that a straw that crosses the circle once can be extended to cross a second time. The segment viewpoint assumes that the trial is only valid if the straws makes two intersections with the circle. The line, ray and segment viewpoints become equivalent in the mathematically ideal infinite-straw-length limit $\tilde{d} = 2R/L \rightarrow 0$. For the data presented here, we use the segment viewpoint, however, it is possible to show the main qualitative conclusions are not dependent on this choice. As in the experiment the length of straws is fixed, we modify the $R/L$-ratio (the dimensionless parameter $\tilde{d}$) by changing the radius of the circles in the analysis. In practice, for each value of $R$, we draw circles finely distributed on a grid to cover as much area of the plane as possible. We use a near-uniform distribution of circles avoiding any overlap between them because otherwise some regions might be considered for multiple times, creating a potential source of systematic errors. For an individual run, we measure the chord length, $\ell$, and bin the dimensionless length $x = \ell/(2R)$ in bins with width 0.02. The resulting normalised histograms for two value of the parameter $\tilde{d}$ are shown in figure 4. Note that the straws in the experiment were thrown by hand, and the results could be affected by a bias of the researcher. Indeed we observe that our method leads to the unbalanced distribution shown

Figure 3. (a) Camera image of the tossed straw on the paper. (b) The image processing routine fits a line segment and adds tightly arranged circles. The size of circles is varied in the analysis. We show example circles to indicate, two successful trials (the red circles) and unsuccessful trials (black circles) where the straw fails to intersect circles at 2 points. Such cases are omitted from further analysis leading to a post-selection of ‘events’. The length of the line may either match the length of the straw or be extended as desired. (c) The experiment is repeated 3600 times and the paths of 1200 of them are added. In all plots, the coordinates on the axes are pixel counts of the camera.
Although we follow Jaynes’ design of experiment, there could be a better way of sampling chords to minimise human imperfections.

3. Theoretical derivation

The probability distributions shown in figure 4 follow from a simple geometrical argument. Using rotational symmetry we can always rotate the chord axis to be vertical as in figure 5. We consider all possible positions of first touch that give a valid trial in the segment viewpoint. Using axial symmetry, we only need consider the left part of the first-touch area. Let \( p(x) \) be the probability to measure a chord with length between \( 2Rx \) and \( 2R(x + dx) \), where \( x \)
is chord-to-diameter ratio; see figure 5(b). First we derive a probability cumulative function, $F(x) = \int_0^x p(x)\,dx$, by considering the area $A$ where all straws that first touch in $A$ generate a chord length shorter than $2R\ell$. Using geometry, we obtain the following function for $A$:

$$A = LR\sqrt{1 - x^2} - R^2\sin^{-1}(x) + R^2\sqrt{1 - x^2}. \quad (1)$$

The cumulative probability $F(x) = A/(A_1 + A_2)$:

$$F(x) = \frac{4\sqrt{1 - x^2} - 2d\sin^{-1}(x) - 2x\sqrt{1 - x^2}}{4 - \pi d}. \quad (2)$$

The probability density function $p(x)$ is given by the derivative of $F(x)$:

$$p(x)\,dx = \left(1 - \frac{\pi d}{4}\right)^{-1} \left(\frac{x - x^2d}{\sqrt{1 - x^2}}\right)\,dx. \quad (3)$$

This formula is used to plot the line in figure 4 and gives very good agreement with the data. In the limit, $d \to 0$, one obtains Jaynes’ distribution [2]:

$$p(x)\,dx = \frac{x\,dx}{\sqrt{1 - x^2}}. \quad (4)$$

Equation (4) is the limiting case of segment viewpoint but it is possible to show that the same distribution is obtained using the line or ray viewpoints. The only difference is in how we determine the area of first touch. Instead in figure 5, for the ray viewpoint, the first-touch area is enclosed by upper half of the solid circle, the upper half of dotted circle and the two dotted lines. Repeating the analysis proves the equivalence between the ray and line viewpoints, and the limiting case of segment viewpoint in the limit $d \to 0$.

Comparing the experimental data of figure 4 to the predictions of equation (3), we obtain a mean-square weighted deviation ($\chi^2$) [14] of 8.8 and 1.8 for short and long straws, respectively. The experiment becomes a more accurate two-randomiser ‘machine’ in the assumed limit of infinite straw length. The nature of the breakdown in the limit of large post-selection requires further investigation. A formula for the average ratio of chord length to circle diameter, dashed lines in figure 4, using the segment viewpoint, can also be derived by integrating $xp(x)$ using (3) from 0 to 1:

$$\bar{x} = \left(1 - \frac{\pi d}{4}\right)^{-1} \left(\frac{\pi}{4} - \frac{2}{3}d\right). \quad (5)$$

Note that in the limit $d \to 0$, the average chord length becomes $\bar{x} \to \pi/4$.

From the distribution shown in figure 4, we can calculate the probability, $P$, that the chord length is longer that $\sqrt{3}R$ ($x > \sqrt{3}/2 = 0.866$). As apparent by comparing figures 4(a) and (b), $P$ depends on the circle diameter to straw length ratio $d = 2R/L$. The measured probabilities for 18 different value of $d$ are plotted in figure 6. Again we can obtain an analytical result using geometry. If we want the chord to be longer than $\sqrt{3}R$, then the position of first touch must be within the area $A_2$, and the probability of a chord length greater than $\sqrt{3}R$ is given by the ratio $A_2/(A_1 + A_2)$; see figure 5(a). Using geometry,

$$A_2 = \frac{LR}{2} - \frac{\pi R^2}{6} - \frac{\sqrt{3}R^2}{4}, \quad (6)$$

$$A_1 + A_2 = LR - \frac{\pi R^2}{2}. \quad (7)$$
and the probability $A_2/(A_1 + A_2)$ is

$$P(\tilde{d}) = \frac{12 - 2\pi\tilde{d} - 3\sqrt{3}\tilde{d}}{24 - 6\pi\tilde{d}}. \tag{8}$$

In fact, the probability can also be computed by directly integrating equation (4):

$$P(\tilde{d}) = \int_{\pi/2}^{1} p(x)dx = \frac{12 - 2\pi\tilde{d} - 3\sqrt{3}\tilde{d}}{24 - 6\pi\tilde{d}}. \tag{9}$$

We plot this analytic expression together with the measurements in figure 6, and find reasonable agreement with the experimental data. However, there is a clear systematic discrepancy between theory and experiment in the limit $\tilde{d} \to 1$, where the amount of post-selection is maximal—because for large $R = \tilde{d}L/2$ it is less likely for a segment with a given $L$ will intersect the circle at 2 points. The reason for this discrepancy is unknown and will require further investigation. In the limit of infinitely long straws, $\tilde{d} \to 0$, we obtain $P(\tilde{d}) \to 0.5$ corresponding to Jaynes’ result [2] as expected.

It is also worth noting that integral geometry [15] provides a more elegant approach to obtain $P = 1/2$. An important object in integral geometry is the strip, the region bounded by two parallel lines. In a two-dimensional space, the density of strips of a certain breadth can be written as

$$dB = dp \wedge d\phi, \tag{10}$$

where $p$ and $\phi$ are polar coordinates of the foot-point of the strip’s mid-parallel line. The foot-point of a line is the point at which the line is nearest to the origin. It can be shown that rotational and translational invariances are satisfied by equation (10). This equation leads to a useful corollary, the measure $m$ of the set of strips of breadth $a$ that intersect a circle $C$ with radius $r$ is
\[ m(r) = \int_{B \cap C_{=\Omega}} dB = 2\pi r + \pi a. \]  \hspace{1cm} (11)

As is shown in the \( P = 1/4 \) solution, chords that intersect the concentric circle with half the radius are longer than \( \sqrt{3}R \). Therefore we can express the probability as

\[ P = \frac{m(R/2)}{m(R)} = \frac{\pi R + \pi a}{2\pi R + \pi a} = 1/2, \]  \hspace{1cm} (12)

where \( a = 0 \) since we can treat a line as a strip with zero breadth. It is interesting that the integral geometry solution contains no explicit sampling of chords, similar to Jaynes’ ‘transformation groups’ approach. The result is solely derived from the invariances and the line viewpoint.

### 4. Discussion

The main conclusion of figure 6 is that rather than one of these three standard answers to Bertrand’s question (illustrated in figure 1), we find that the answer is a continuous function of the \( \tilde{d} = 2R/L \) ratio as predicted theoretically in [6, 8]. In the limit of infinite straw length (\( \tilde{d} = 2R/L \rightarrow 0 \)) the answer tends towards 0.5. In this limit the line, ray and segment viewpoints merge\(^2\). In the opposite limit, \( \tilde{d} \rightarrow 1 \), the amount of post-selection increases which reduces the probability of finding a chord longer than \( \sqrt{3}R \).

These results can also be interpreted using Jaynes’ principles of invariances. We have shown that, under the segment viewpoint, the chord length distribution depends on \( \tilde{d} \), which clearly violates scale invariance, which requires that the chord distribution is invariant to circle radius. Thus, the probability generally deviates from 1/2. If however we use the line viewpoint, \( \tilde{d} \equiv 0 \), then equation (3) reduces into equation (4), a distribution in compliance with all invariances, corresponding to the \( P = 1/2 \) solution. As shown by Jaynes and our integral geometry approach, any distribution of chords that satisfies principles of invariances must lead to \( P = 1/2 \).

One interpretation of the Duhem–Quine thesis [11] is that we cannot assess which assumptions, e.g. edge effects, that the experiment is a perfect random number generator. or that the straw length is sufficiently large, is responsible for the discrepancy between the measurement and the expected result. In our example by simulating different straw lengths we are able to separate the effect of the straw length from other assumptions. The fact that we observe a systematic discrepancy in the limit, \( \tilde{d} = 2R/L \rightarrow 1 \), suggest a residual bias in the experiment that cannot be eliminated simply by taking more data. The origin of this discrepancy is most likely due to the finite size of the raw images.

### 5. Summary

In summary, we have performed an experimental test of Jaynes’ variation of Bertrand’s paradox which can be stated in the form of a question: a long straw is tossed at random onto a circle; given that it falls so that it intersects the circle, what is the probability that the chord thus defined is longer than a side of the inscribed equilateral triangle?\(^2\) We confirm in a real world practical implementation the predictions of numerical experiments that answer is a continuous function of the ratio of the circle radius to cord straw length \( \tilde{d} = 2R/L \) [6, 8].

\(^2\)Although Jaynes did not specify a viewpoint, in fact, he must have assumed the line viewpoint in order to obtain equation (4) [2].
highlight some experimental differences to the expected result that are due to the nature of post-selection of successful results. The results and analysis illustrate the fundamental nature of assumptions implicit in the scientific method, i.e. that experimentation requires auxiliary conditions and that almost all experiments are designed with a prior probability distribution in mind. The experiment is in effect only a test of the assumed distribution, particularly in the case of post-selection. In the limit of long straw length the experimental test of Bertrand’s paradox selects Laplace’s principle of indifferent, like quantum experiments post-select quantumness, and Black–Scholes hedging selects Gaussian fluctuations [16].

Acknowledgments

We thank Nathan J Cinnamond, Mia West, Robert Heighton for assistance with the experiment, in particular, obtaining the images used in the analysis, graphics and for stimulating discussions. We also thank Tom Lancaster, Matthias C M Troffaes, Yuxun Zhang, Ifan Hughes and Robert Bettles for discussions, and Susan Hilton for technical support. We acknowledge funding from Durham University. CSA acknowledges financial support from Durham University and EPSRC Grant Nos. EP/M014398/1, EP/P012000/1, EP/R002061/1, EP/R035482/1, and EP/S015973/1. The data presented in this paper are available at doi:10.15128/r2s1784k734.

ORCID iDs

Zhenning Liu © https://orcid.org/0000-0002-0794-419X
Charles S Adams © https://orcid.org/0000-0001-5602-2741

References

[1] Bertrand J 1889 Calcul des Probabilites (Paris: Gauthier-Villars)
[2] Jaynes E T 1973 The well-posed problem Found. Phys. 3 477–93
[3] Fuchs C A and Schack R 2013 Quantum-Bayesian coherence Rev. Mod. Phys. 85 1693–715
[4] See e.g. ch 1 in Garibaldi U and Scalas E 2010 Finitary Probabilistic Methods in Econophysics (Cambridge: Cambridge University Press)
[5] Walley P 1996 Inferences from multinomial data: learning about a bag of marbles (with discussion) J. R. Stat. Soc. B 58 3–57
[6] Marinoff L 1994 A resolution of Bertrand’s paradox Phil. Sci. 61 1–24
[7] Arts D and Sassoli de Bianchi M 2014 Solving the hard problem of Bertrand’s paradox J. Math. Phys. 55 083503
[8] Di Porto P, Crosignani B, Ciattoni A and Liu H C 2011 Bertrand’s paradox: a physical way out along the lines of Buffon’s needle throwing experiment Eur. J. Phys. 32 819–25
[9] Duhem P 1991 The Aim and Structure of Physical Theory 2nd edn (Princeton, NJ: Princeton University Press)
[10] Quine W V O 1951 Two dogmas of empiricism Phil. Rev. 60 20–43
[11] Stevens M 2001 The bayesian treatment of auxiliary hypotheses Brit. J. Phil. Sci. 52 515–37
[12] Stevens M 2006 Bayesian approach to philosophy of science Encyclopedia of Philosophy ed D M Borchert 2nd edn (Detroit, MI: Macmillan Reference USA)
[13] Yong E H and Mahadevan L 2011 Probability, geometry, and dynamics in the toss of a thick coin Am. J. Phys. 79 1195–201
[14] Hughes I G and Hase T P A 2010 Measurements and Their Uncertainties: A Practical Guide to Modern Error Analysis (Oxford: Oxford University Press)
[15] Santaló L A 2004 *Integral Geometry and Geometric Probability* (Cambridge: Cambridge University Press) pp 68–71
[16] James J 2017 *Quantitative Finance* (Bristol: IOP Publishing Ltd)