Hume's Fallacy: Miracles, Probability, and Frequency

Paul Mayer

Abstract

Frequency-based arguments against rational belief in a miracle occurring have been present for centuries, the most notable being from David Hume. In this essay, I will show Hume's argument rests on an equivocation of probability, with him using the term interchangeably to refer to two different and incompatible perspectives: Bayesianism and Frequentism. Additionally, I will show that any frequentist arguments against miracles relies on a view of probability that is only dubiously linked to rationality. In other words, the frequentist cannot have it both ways: if probability is indeed frequency, then miracles are indeed improbable but not necessarily irrational to believe. On the other hand, if probability is an agent's confidence in a given belief, than under certain assumptions miracles are indeed highly probable and thus rational to believe despite their rarity. As a result, regardless of which view of probability one takes, it does not follow that believing in a miracle is irrational simply because miracles are rare.

1 Introduction

Miracles are defined as events outside of the usual course of nature. As Mackie (2013) writes,

> The laws of nature. . . describe the ways in which the world—including, of course, human beings—works when left to itself, when not interfered with. A miracle occurs when the world is not left to itself, when something distinct from the natural order as a whole intrudes into it.

If we assume most observations are the result of nature when left to itself, than miracles are (relatively) rare. With this in mind, it is perhaps unsurprising that some would argue the rarity of miracles is a reason to disbelieve them.

The most famous and influential argument against rational belief in miracles comes from Section X of David Hume's Enquiry Concerning Human Understanding (Hume and Steinberg, 1993). While there are differing interpretations on what point Hume intended to make[1], my reading of Hume is that he is making an argument as follows:

- P1: One should only believe miracle testimony T_M if the probability that T_M is true is larger than the probability that T_M is false.[2]
- P2: The probability that miracle testimony T_M is true is *always* smaller than the probability that T_M is false.
- C: Therefore, one should *never* believe miracle testimony T_M is true

The controversial premise here is P2, and many have taken aim at the suggestion that miracle claims are always less probable than the testimony being false (Earman, 2000). For instance, Charles Babbage (often considered the "Father of the Computer") wrote in the Ninth Bridgewater Treatise

> [I]f independent witnesses can be found, who speak truth more frequently than falsehood, it is ALWAYS possible to assign a number of independent witnesses, the improbability of the falsehood of whose concurring testimony shall be greater than the improbability of the alleged miracle (Babbage, 1841).

Of course, Babbage's suggestion is only true if the prior probability of a miracle is nonzero. Hume argues that the prior probability of a miracle is zero because the frequency of miracles is zero. He writes:

> A miracle is a violation of the laws of nature; and as a firm and unalterable experience has established these laws, the proof against a miracle, from the very nature of the fact, is as entire as any argument from experience can possibly be imagined[3]...It is no miracle that a man, seemingly in good health, should die on a sudden: because such a kind of death, though more unusual than any other, has yet been frequently observed to happen. But it is a miracle, that a dead

[1] Some believe Hume was arguing against the *possibility* of miracles, however under this interpretation, Hume has been criticized as begging the question (Lewis, 2009) and proving too much (Geisler, 2014). For more interpretations and discussion of Hume, see Hájek (2008) and Earman (2000) for a critical examination.

[2] This premise is based on the following from Hume's Enquiry:

> No testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish...When anyone tells me, that he saw a dead man restored to life, I immediately consider with myself, whether it be more probable, that this person should either deceive or be deceived, or that the fact, which he relates, should really have happened...If the falsehood of his testimony would be more miraculous, than the event which he relates; then, and not till then, can he pretend to command my belief or opinion (Hume and Steinberg, 1993).

[3] Others, such as Lewis (2009) and Campbell (1824), have argued Hume begs the question against miracles by assuming natural laws are based on firm and unalterable experience. While this is also an important critique, my target is the frequency-based nature of Hume's argument.

> man should come to life; because that has never been observed in any age or country (Hume and Steinberg, 1993).

Hume is only able to make this move because he equivocates probability: he using the same the term to refer to both frequency (the frequentist view of probability, described in Section 2.2) *and* as an epistemic concept related to rational credences (the Bayesian view of probability, described in Section 2.3). It is this equivocation that exposes the flaw in Hume, and indeed any frequency-based argument, against miracles. My argument is that Hume cannot have it both ways; probability cannot be both primarily be frequency and be an epistemic concept due to the incompatibility between these interpretations.

Indeed we see Hume himself switching between these interpretations of probability within his Enquiry. In Section VI, *Of Probability* Hume writes, "Though there be no such thing as *Chance* in the world; our ignorance of the real cause of any event has the same influence on the understanding, and begets a like species of belief or opinion (Hume and Steinberg, 1993)". This Hume adopting a Bayesian view, where probability is an epistemic concept referring to one's degree of belief[4]. Yet in Section 10, Of Miracles, we see frequentist Hume using probability interchangeably with frequency, where he writes (emphasis added):

> A wise man, therefore, proportions his belief to the evidence. In such conclusions as are founded on an infallible experience, he expects the event with the last degree of assurance, and regards his past experience as a full *proof* of the future existence of that event. In other cases, he proceeds with more caution: He weighs the opposite experiments: **He considers which side is supported by the greater number of experiments**: to that side he inclines, with doubt and hesitation; and when at last he fixes his judgment, the evidence exceeds not what we properly call *probability* (Hume and Steinberg, 1993).

If the Bayesian view of probability is correct, than assuming miracles have a zero prior probability is begging the question against them, since this amounts to claiming one *a priori* knows all current and future miracle claims are false before investigation. If the frequentist view of probability is correct [5], probability and rational belief are no longer (necessarily) to the same concept, and the frequentist needs to show what this necessary link is to argue belief in miracles is irrational.

[4] Italian statistician Bruno DeFinetti, a Bayesian, opened his two volume treatise on probability by declaring (similar to Bayesian Hume), "Probability does not Exist! (De Finetti et al., 2017)"

[5] Frequentism has come under much recent criticism, with many philosophers of statistics now arguing it is an untenable interpretation of probability (Hubert, 2021; Hájek, 2009).

2 What is Probability

Statisticians typically agree with Andrey Kolmogorov's axiomatization of probability theory into a three-element tuple (Ω, E, P), consisting of a sample space Ω (a set of outcomes), an event space E (a set of sets of outcomes, and a $\sigma-$algebra over Ω), and a probability measure P that assigns each event $e \in E$ to a probability between 0 and 1 inclusive (Kolmogorov and Bharucha-Reid, 2018). P must obey a set of probability axioms such as finite additivity for (Ω, E, P) to be considered a legitimate probability space.

What is less clear is what the meaning or interpretation of these terms are. For instance, American mathematician and statistician L.J. Savage once remarked "It is unanimously agreed that statistics depends somehow on probability. But, as to what probability is and how it is connected with statistics, there has seldom been such complete disagreement and breakdown of communication since the Tower of Babel (Savage, 2012)." Furthermore, Bertrand Russell once quipped in a lecture, "Probability is the most important concept in modern science, especially as nobody has the slightest notion what it means (Bell, 2012)."

2.1 The Classical View of Probability

Probability developed by applying mathematics to games of chance, with figures such as Girolamo Cardano, Blaise Pascal, Pierre de Fermat, and Christian Huygens laying the groundwork for modern statistics. The view that emerged from these thinkers is called the classical view of probability, where the probability of an event is the number of possible ways an event can occur divided by the total number of possibilities (Hájek, 2023). This was famously put by French mathematician Abraham de Moivre in a 1718 book on probability that was apparently highly prized by gamblers, "[I]f we constitute a fraction whereof the numerator be the number of chances whereby an event may happen, and the denominator the number of all the chances whereby it may either happen or fail, that fraction will be a proper designation of the probability of happening (Moivre, 1738)." Formally, under the classical view, the probability of an event $e \in E$ is defined as $P(e) \equiv |e|/|\Omega|$, where $|\cdot|$ represents the cardinality of a set.

The classical view of probability relies on what is called the principle of insufficient reason, or as economist John Maynard Keynes famously called it, the principle of indifference (Keynes, 2013), which says that when we have no reason to believe that one outcome will occur preferentially to another, we assign them equal probability (Weisstein). This assumption, while reasonable in many cases, leads to two issues with the classical view.

The first is that the classical view has no way of *updating* probabilities if the data shows uniformity is not a good assumption. The classical view only deals with possibility, and has no way of incorporating frequencies or taking into account direct observations. Consider a loaded die, weighted in such a way so it lands on one number more frequently than others: after rolling such a the die, it seems there is now a reason to prefer some numbers over others, yet the classical view has no way of overcoming the principle of indifference and adapting to these observations.

The second issue with the classical view involves what are called Bertrand's Paradoxes, named after French mathematician Joseph Bertrand (Shackel, 2007). These are demonstrated mathematically in Appendix A; in essence Bertrand's paradoxes show that the principle of indifference can provide inconsistent results when dealing with a sample space that is uncountably infinite. While the statistician E.T. Jaynes provides a response to this in terms of the principle of maximum entropy (Jaynes, 1968), the main point is the classical view alone does not appear to provide a way of arbitrating between different answers that come from equivalent reformulations.

2.2 The Frequentist View of Probability

Like the classical view, the frequentist view uses counts to determine probabilities, dividing the number of outcomes that belong to a certain event by the total number of trials. However unlike the classical view, the frequentist uses actual observed counts rather than possibilities to determine the probability (Hájek, 2023). This view was famously stated by John Venn, who wrote "probability *is* nothing but that proportion (Venn, 1866)." Formally, let $o = [\omega_1, \omega_2, \ldots, \omega_n]$ be a sequence of n outcomes observed from Ω belonging to a probability space (Ω, E, P). The probability of an event $e \in E$ is thus defined as $P(e) \equiv \frac{|\{\omega_i \in o \text{ s.t. } \omega_i \in e\}|}{|o|}$.

Unlike the classical view of probability, this change to actual counts allows the frequentist view of probability to capture intuitions loaded die and unfair coins. This makes the frequentist view more data-driven than the classical view. However, one immediate problem is that different sequences of outcomes lead to different probabilities: a coin that is never flipped has an undefined probability, while a coin flipped only once has a probability of either 0 or 1 depending on what the result of this flip is. Another issue has to do with unrepeatable events, such as historical events; often called the "problem of the single case." There is no way to assign frequencies to unrepeatable events other than to give them a value of 0 or 1. These problems are related to what is called *finite frequentism,* where $|o| \in \mathbb{Z}^+$, and issues with this view are explored in (Hájek, 1996).

One promising solution to the inconsistencies of different (finite) samples is to let probability be the *limit* of the aforementioned ratio for a sequence of outcomes. This is a view called *hypothetical frequentism,* with the word hypothetical used since it is impossible to perform an infinite number of trials. Formally, hypothetical frequentism defines probability as $P(e) \equiv \lim_{n\to\infty} \frac{|\{\omega_i \in o \text{ s.t. } \omega_i \in e\}|}{|o|}$, for any $n-$length sequence of outcomes o. Indeed these ratios often do converge as $n \to \infty$, so it seems the hypothetical frequentist is able to avoid inconsistencies resulting from different samples that plagued the finite frequentist.

However the hypothetical frequentist has given up the data-driven benefit of frequentism by invoking counterfactuals (Starr, 2022). Indeed to make a claim such as "the democrats will *probably* win the next election" is to imagine an infinite collection of hypothetical repeated elections. If these hypothetical elections differ from one another or from the actual course things take, than the argument is implicitly invoking counterfactuals. Despite possible appearing more data-driven (due to the demand for actual observations), the hypothetical frequentist view of probability is arguably just as (if not more) subjective as the classical view since it is tied to speculative metaphysical judgments about an infinite number of counterfactuals.

2.3 The Bayesian View of Probability

The Bayesian view of probability sees probability as an epistemic notion that quantifies uncertainty. The Bayesian View of probability was famously stated by British mathematician and logician Augustus DeMorgan, who wrote, "By degree of probability, we really mean, or ought to mean, degree of belief (Morgan, 1847)." It is this view that allows probability theory to be an extension of logic that allows one to reason when they lack complete knowledge and update their beliefs in light of new evidence. According to classical logic, (Shapiro and Kouri Kissel, 2024), propositions are either true or false, and due to Aristotle's notions such as the law of the excluded middle (LEM) and law of non-contradiction (LNC), must be one or the other.

George Boole, in his famous work on mathematical logic, connected Aristotle's logic to algebra by letting a 0 represent false and a 1 represent true, providing operations for combining and relating propositions (Boole, 1911; Davis, 2000). It is in this context that a probability, as a measure between 0 and 1, represents an agent's *confidence* that a given proposition is true. Physicist Richard Cox famously made this connection in his 1948 paper Probability, Frequency and Reasonable Expectation (Cox, 1946), later expanding upon it in his 1961 book The Algebra of Probable Inference (Cox, 1961). Fellow physicist and statistician E.T. Jaynes also shared this view in his book Probability Theory: The Logic of Science (Jaynes, 2003). Recent popular-level and academic work has argued for the Bayesian interpretation of probability; see (Clayton, 2021; Hoang, 2020).

At the heart of Bayesian probability is Bayes' Theorem, which describes how a rational agent should update their belief in a model M being true when presented with new evidence E.

$$P(M|E) = \frac{P(E|M)P(M)}{P(E)} \tag{1}$$

Here, $P(M|E)$ is the posterior, referring to an agent's belief that a model M is true after seeing evidence E. $P(E|M)$ is called the likelihood function, and refers to an agent's belief that they would see the evidence E if M is in fact true. This is multiplied by $P(M)$, called the prior, which is one's confidence that M is true before observing evidence E. Lastly, $P(E)$ in the denominator is called the evidence, and refers to one's confidence in seeing E under any model. We will come back to this term later.

The Bayesian view provides a way of overcoming the shortcomings of both the classical and frequentist perspectives. The Bayesian view is able to adapt to observed data, meaning it can reason about loaded die and unfair coins, unlike the classical view. The Bayesian view also avoids Bertrand's paradoxes: in addition to providing a mechanism of updating beliefs via the posterior, the Bayesian view shows how the choice of a prior can be subjective[6]. As Kendal and Moran wrote in 1964 (referring to Bertrand's Paradoxes), "Since the ascription of a measure to such elements is not quite an obvious procedure, a number of 'paradoxes' can be produced by failure to distinguish the reference set (Kendall and Moran, 2012)." In other words, the Bayesian gets to choose how they assign a probability measure to the prior $P(M)$, and if they are using the principle of indifference, they must to select a specific measure to use as a prior.

Unlike the frequentist, who either needs to posit dubious counterfactual scenarios or collect actual data, the Bayesian can still reason about possibility in the absence of data. Of course, once data is observed, these observations can be used to revise one's belief, but observations are not strictly necessary for the Bayesian to reason about probability. In summary, the Bayesian view allows one to reason about possibility when one lacks data, and incorporate data when one has it.

3 Probability and Rationality

One's interpretation of probability has a direct bearing on the relationship between probability and rationality. To the frequentist, probability *is* frequency, and any attempt to link frequency to rational belief is non-obvious and subject to the famous Problem of Induction. To the Bayesian, probability involves belief, with no necessary link to frequencies: bayesian epistemology argues a rational agent will update their belief via Bayesian conditioning (Arnborg and Sjödin, 2001), and as Ramsey (2016) famously showed, using Bayesian probabilities is the only way to avoid getting duped by a "dutch book" style betting argument.

We come dilemma for those advancing a Hume-style argument against belief in miracles. She must choose to either

1. Adopt a frequentist view of probability (described in Section 3.1), or
2. Adopt a Bayesian view of probability

If she chooses the first horn, then probability *is* frequency. On this view, miracles are indeed improbable (because they have low frequency) but are not necessarily irrational to believe. This is because under the frequentist view, the link between probability (frequency) and rationality is non-obvious and subject to the Problem of Induction. This is detailed in Section 3.1.

[6]The subjectivity of assigning credence does not mean Bayesian inference is irrational; indeed one could argue epistemology is subjective since it involves belief and what information a given person has access to.

On the other hand, if she chooses the second horn, then probability and rationality are indeed linked via Bayesian epistemology, but frequency has less to do with probability than Bayes' theorem does. As a result, we show it is possible for miracles to have a high probability and thus be rational to believe, regardless of their low frequency. This is detailed in Section 3.2.

Thus, either way, belief in miracles is not irrational. It is only by equivocating between these two incompatible views of probability that one can definitively argue miracles are irrational because of their low frequency.

3.1 Horn 1: Frequentism and the Problem of Induction

Hume is also famous for positing his famous problem of induction elsewhere in the Enquiry (Henderson, 2022), which (ironically) is also the most famous argument *against* trusting frequencies. Hume writes "it implies no contradiction that the course of nature may change, and that an object seemingly like those which we have experienced, may be attended with different or contrary effects."

On one hand, Hume claims that "If I ask, why you believe any particular matter of fact, which you relate, you must tell me some reason," while on the other, he writes "tis impossible for us to satisfy ourselves by our reason, why we shou'd extend [our] experience beyond those particular instances, which have fallen under our observation (Hume and Steinberg, 1993)." By thus concluding that one must take it on blind faith that the future will resemble the past (combined with his evidentalist epistemology), Hume has provided an undercutting defeater for his own frequentist reasoning about miracles. As C.S. Lewis famously remarked, "no man knew this better than Hume. His Essay on Miracles is quite inconsistent with the more radical, and honourable, scepticism of his main work. (Lewis, 2009)."

My argument is that this is not a problem unique to Hume, rather it is a problem for frequentism in general: if probability is indeed frequency, then one may ask what *reason* we have to trust frequencies, calling into question the link between probability and rationality. Furthermore, to say the past is a good guide to the future because the past has predicted the present well is to beg the question in favor of a link between frequentism and rationality, since this involves using the fact that the past is good at predicting the past to predict the future.

Here's an example adopted from Russell (2001): imagine a turkey living comfortably on a farm for years. Suppose that the turkey believes this provides a reason for believing he will wake up the next day, which happens to be Thanksgiving. Unbeknownst to the turkey, the very fact he was raised and fed was for the purpose of being Thanksgiving dinner. The key insight here is the farmer has justified knowledge that the turkey will be killed and eaten on Thanksgiving, despite the fact that the number of days that the turkey will be killed and eaten (1 day) is small compared to the total number of days the turkey has lived on the farm wihtout being killed (hundreds of days). Under frequentism, the probability that the turkey will be killed and eaten is low[7] for both the farmer and the turkey *despite* the fact that the farmer has justified knowledge of the turkey's eventual death and consumption. Thus, under frequentism, probability and rationality are separate concepts.

3.2 Horn 2: Bayesianism and Frequencies

To the Bayesian, probability is not necessarily frequency, and rather is one's degree of belief in a given proposition being true. Returning to the turkey in Section 3.1, the turkey would perhaps assign high probability to model where he wakes up and is fed as usual on Thanksgiving and future days, while the farmer would not. The difference is the farmer has additional knowledge about the turkey, that if the turkey had, would cause a (rational) turkey to revise his belief.

A Bayesian update for the Turkey would work as follows: let T be the event where the Turkey becomes thanksgiving dinner, and T^C be it's complement (i.e. the turkey wakes up fine on Thanksgiving). Let F be the farmer's testimony that he raised and fed the turkey for the past 500 days so it would be Thanksgiving dinner. Applying Bayes' Theorem (Equation 1), we have $P(T|F) = \frac{P(F|T)P(T)}{P(F)}$. Now the turkey may ascribe low initial confidence to his being thanksgiving dinner, corresponding to a low prior $P(T)$. However if the turkey's confidence that the farmer would give this testimony if indeed he would end up as thanksgiving dinner is high enough, then this can overcome the initial improbability the turkey assigned to $P(T)$. What is key here is the evidence term, $P(F)$, which represents the confidence the turkey has that the farmer would claim the turkey would be thanksgiving dinner under *any* model. Using the law of total probability, $P(F) = P(F|T)P(T) + P(F|T^C)P(T^C)$. Thus if $P(F|T^C) \approx 0$, meaning the turkey doesn't think the farmer would lie, then the posterior $P(T|F) \approx 1$, meaning the turkey has updated his belief and is now fairly confident he *will* be thanksgiving dinner. This is despite the fact that the frequency of days the turkey is killed and eaten is indeed low and despite the turkey's initial low prior. Again, the turkey can be justified believing in a low frequency event if the observed evidence is far better explained by the low frequency event than a higher frequency alternative.

In a similar way, let M a miracle. Even if one assigns a low prior $P(M)$ to miracles on the basis of their low frequency, as long as this prior is nonzero,

[7] If the turkey is only killed and eaten on one day out of hundreds, than the probability that the turkey will be killed on a given day (such as Thanksgiving) is thus low, regardless of the farmer's knowledge or intention. A sly frequentist could argue that what should really be considered is the frequency of Thanksgivings, not days in general. Which items are to be used in a probability computation is called the reference class problem, coined by Reichenbach (1971). While the reference class problem is present in any substantive view of probability (Hájek, 2007), it is particularly an issue for the frequentist because the frequentist needs to specify *what* frequency they are talking about. Of course, this is how a frequentist can still argue for the rationality of miracles; that there is no analogous reference class for a certain religious figure and thus the reference class of naturalistic explanations or everyday poeple is inappropriate and does not count against the frequency.

there can always be testimony T_M such that $P(M|T_M) > P(M^C|T_M)$, meaning believing a miracle occurred is more rational than its denial.

4 Alternative Models and Uncertainty

The denominator of Bayes' Theorem, or the evidence $P(E)$, implicitly takes into account other possible models could explain the evidence.

References

J. L. Mackie, *The Miracle of Theism: Arguments for and Against the Existence of God*, repr ed. Oxford: Clarendon Press, 2013.

D. Hume and E. Steinberg, *An Enquiry Concerning Human Understanding*, 2nd ed. Indianapolis: Hackett, 1993.

C. S. Lewis, *Miracles.* Harper Collins, Jun. 2009, google-Books-ID: tH8di3qQDhEC.

N. L. Geisler, "4. Miracles & The Modern Mind," in *In Defense of Miracles: A Comprehensive Case for God's Action in History.* InterVarsity Press, Apr. 2014, google-Books-ID: 4W4uDwAAQBAJ.

A. Hájek, "Are Miracles Chimerical?" *Oxford Studies in Philosophy of Religion*, vol. 1, pp. 82–104, 2008, publisher: Oxford University Press.

J. Earman, *Hume's Abject Failure: The Argument Against Miracles.* Oxford University Press, Nov. 2000, google-Books-ID: 2_DQCwAAQBAJ.

C. Babbage, *The Ninth Bridgewater Treatise. A fragment ... Second edition.* Lea & Blanchard, 1841, google-Books-ID: XyRbAAAAcAAJ.

G. Campbell, *A Dissertation on Miracles: containing an examination of the principles advanced by David Hume, Esq; in an Essay on miracles.* Thomas Tegg, 1824.

B. De Finetti, A. Machì, A. F. M. Smith, and B. De Finetti, *Theory of Probability: A Critical Introductory Treatment.* Chichester, UK ; Hoboken, NJ: John Wiley & Sons, 2017.

M. Hubert, "Reviving frequentism," *Synthese*, vol. 199, no. 1, pp. 5255–5284, Dec. 2021. [Online]. Available: https://doi.org/10.1007/s11229-021-03024-8

A. Hájek, "Fifteen Arguments Against Hypothetical Frequentism," *Erkenntnis*, vol. 70, no. 2, pp. 211–235, Mar. 2009, 49 citations (Crossref/DOI) [2024-12-07]. [Online]. Available: https://doi.org/10.1007/s10670-009-9154-1

A. N. Kolmogorov and A. T. Bharucha-Reid, *Foundations of the Theory of Probability: Second English Edition.* Courier Dover Publications, Apr. 2018.

L. J. Savage, *The Foundations of Statistics.* Courier Corporation, Aug. 2012.

E. T. Bell, *The Development of Mathematics.* Courier Corporation, Sep. 2012.

A. Hájek, “Interpretations of Probability,” in *The Stanford Encyclopedia of Philosophy*, winter 2023 ed., E. N. Zalta and U. Nodelman, Eds. Metaphysics Research Lab, Stanford University, 2023. [Online]. Available: https://plato.stanford.edu/archives/win2023/entries/probability-interpret/

A. d. Moivre, *The Doctrine of Chances, Or, A Method of Calculating the Probabilites of Events in Play ...* author, 1738, google-Books-ID: iYhj8z6UNwkC.

J. M. Keynes, *A Treatise on Probability.* Courier Corporation, Sep. 2013, google-Books-ID: rVCoAAAAQBAJ.

E. W. Weisstein, “Principle of Insufficient Reason,” publisher: Wolfram Research, Inc. [Online]. Available: https://mathworld.wolfram.com/PrincipleofInsufficientReason.html

N. Shackel, “Bertrand’s Paradox and the Principle of Indifference,” *Philosophy of Science*, vol. 74, no. 2, pp. 150–175, Apr. 2007. [Online]. Available: https://www.cambridge.org/core/journals/philosophy-of-science/article/abs/bertrands-paradox-and-the-principle-of-indifference/DC735A7B90AD19EB0572A5EA9C5B07BB

E. T. Jaynes, “Prior Probabilities,” *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 3, pp. 227–241, Sep. 1968, conference Name: IEEE Transactions on Systems Science and Cybernetics. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/4082152?casa_token=SovUWUGjLIgAAAAA:0lkG_M8fdDjmkAeMrhUKpfg2o_QxQaHoUFR7TrM2kuEis9vZILwp_XGVhAG8LE8R-Rv-6nyO

J. Venn, *The Logic of Chance: An Essay on the Foundations and Province of the Theory of Probability, with Especial Reference to Its Application to Moral and Social Science.* Macmillan, 1866.

A. Hájek, ““Mises redux” — Redux: Fifteen arguments against finite frequentism,” *Erkenntnis*, vol. 45, no. 2, pp. 209–227, Nov. 1996. [Online]. Available: https://doi.org/10.1007/BF00276791

W. Starr, “Counterfactuals,” in *The Stanford Encyclopedia of Philosophy*, winter 2022 ed., E. N. Zalta and U. Nodelman, Eds. Metaphysics Research Lab, Stanford University, 2022. [Online]. Available: https://plato.stanford.edu/archives/win2022/entries/counterfactuals/

A. D. Morgan, *Formal Logic: Or, The Calculus of Inference, Necessary and Probable.* Taylor and Walton, 1847.

S. Shapiro and T. Kouri Kissel, “Classical Logic,” in *The Stanford Encyclopedia of Philosophy*, spring 2024 ed., E. N. Zalta and U. Nodelman, Eds. Metaphysics Research Lab, Stanford University, 2024. [Online]. Available: https://plato.stanford.edu/archives/spr2024/entries/logic-classical/

G. Boole, *The Laws of Thought (1854)*. Open Court Publishing Company, 1911.

M. Davis, *The Universal Computer: The Road from Leibniz to Turing*, first edition ed. New York, NY London: W. W. Norton & Company, Oct. 2000.

R. T. Cox, “Probability, Frequency and Reasonable Expectation,” *American Journal of Physics*, vol. 14, no. 1, pp. 1–13, Jan. 1946, 781 citations (Crossref/DOI) [2024-11-15]. [Online]. Available: https://doi.org/10.1119/1.1990764

——, *The Algebra of Probable Inference*. Baltimore: Johns Hopkins Press, 1961.

E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge University Press, Apr. 2003, google-Books-ID: tTN4HuUNXjgC.

A. Clayton, *Bernoulli's Fallacy: Statistical Illogic and the Crisis of Modern Science*. Columbia University Press, Aug. 2021, google-Books-ID: HzQGEAAAQBAJ.

L. N. Hoang, *The equation of knowledge: from Bayes' rule to a unified philosophy of science*, first edition ed. Boca Raton, FL: CRC, 2020.

M. G. Kendall and P. A. P. Moran, *Geometrical Probability*. Ulan Press, Sep. 2012.

S. Arnborg and G. Sjödin, “On the foundations of Bayesianism,” *AIP Conference Proceedings*, vol. 568, no. 1, pp. 61–71, May 2001. [Online]. Available: https://doi.org/10.1063/1.1381871

F. P. Ramsey, “Truth and Probability,” in *Readings in Formal Epistemology: Sourcebook*, H. Arló-Costa, V. F. Hendricks, and J. van Benthem, Eds. Cham: Springer International Publishing, 2016, pp. 21–45. [Online]. Available: https://doi.org/10.1007/978-3-319-20451-2_3

L. Henderson, “The Problem of Induction,” in *The Stanford Encyclopedia of Philosophy*, winter 2022 ed., E. N. Zalta and U. Nodelman, Eds. Metaphysics Research Lab, Stanford University, 2022. [Online]. Available: https://plato.stanford.edu/archives/win2022/entries/induction-problem/

B. Russell, *The Problems of Philosophy*. OUP Oxford, Mar. 2001, google-Books-ID: y8loAgAAQBAJ.

H. Reichenbach, *The Theory of Probability: An Inquiry into the Logical and Mathematical Foundations of the Calculus of Probability (English Translsyion by Ernest H. Hutten and Maria Reichenbach)*, repr. 2. ed ed. Berkeley: Univ. of California Pr, 1971, original Year: 1949.

A. Hájek, "The reference class problem is your problem too," *Synthese*, vol. 156, no. 3, pp. 563–585, Jun. 2007. [Online]. Available: https://doi.org/10.1007/s11229-006-9138-5

B. C. v. Fraassen, *Laws and Symmetry*. Clarendon Press, Nov. 1989, google-Books-ID: 6QdREAAAQBAJ.

G. Casella and R. L. Berger, *Statistical Inference*, 2nd ed. Pacific Grove, Calif: Duxbury, 2002.

A Bertrand's Paradox: An Illustration

Adopting an example from Fraassen (1989), consider a factory that produces cubes with a side length from 0 to 1. Suppose one wants to know the expected or average side length, $\mathbb{E}[s]$. Using the principle of indifference on the side length, one assumes a uniform distribution over side lengths, where $s \sim U[0,1]$, and we calculate $\mathbb{E}[s] = \int_s P(s) s ds$ to get the following:

$$\mathbb{E}[s] = \int_0^1 s ds = \frac{s^2}{2}\Big|_{s'=0}^{1} = \frac{1}{2}$$

So the average side length from a factory producing cubes with a side length that ranges from 0 to 1 is $\frac{1}{2}$.

However, one can equivalently restate the problem as the same factory producing cubes with a *face area* that ranges from 0 to 1. Notice how this is still the same factory; if the side lengths of the cube range from 0 to 1, then the face area will range from $0^2 = 0$ to $1^2 = 1$. However under this reformulation, the principle of indifference says to assume a uniform distribution over the face area, or $s^2 \sim U[0,1]$. To calculate the expected side length, one can invoke the Law of the Unconscious Statistician (Casella and Berger, 2002), which states $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$, where $f_X(x)$ represents the probability density function of a random variable X[8]. The expected side length is thus

$$\mathbb{E}\left[\sqrt{x^2}\right] = \int_{s^2} \sqrt{s^2} P(s^2) ds^2 = \int_0^1 z^{1/2} dz = \frac{2z^{3/2}}{3}\Big|_{z=0}^{1} = \frac{2}{3}.$$

By restating the same problem a different way, one gets a different, and indeed inconsistent, answer. One can do this a third time as well by restating

[8] This is a special, continuous case of the Lebesgue–Stieltjes integral over $F_X(x)$, the cumulative density function of X In greater generality, the Law of the Unconscious Statistician states that $\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x) dF_X(x)$.

the problem again as the factory producing cubes with a *volume* that ranges from 0 to 1. Since $0^3 = 0$ and $1^3 = 1$, this is the same factory as the one that produces cubes with a side that ranges from 0 to 1. However if, as the principle of indifference suggests, one assumes $s^3 \sim U[0, 1]$, then the expected side length is now

$$\mathbb{E}\left[\sqrt[3]{x^3}\right] = \int_{s^3} \sqrt[3]{s^2} P(s^3) ds^3 = \int_0^1 z^{1/3} dz = \frac{3z^{4/3}}{4}\Big|_{z=0}^1 = \frac{3}{4}.$$

These three inconsistent answers show the principle of indifference alone lacks the specificity needed to avoid "dutch book" style arguments, such as the ones suggested by Ramsey (2016).