Journal of the Royal Statistical Society Series A: Statistics in Society, 2025, **188**, 181–187 https://doi.org/10.1093/jrsssa/qnae044 Advance access publication 17 May 2024 **Original Article** 



# Bayesian issues in the 1950s: an episode involving Karl Popper and Jimmie Savage

Stephen M. Stigler

Department of Statistics, University of Chicago, Chicago, IL, USA

Address for correspondence: Stephen M. Stigler, Department of Statistics, University of Chicago, 5747 S. Ellis Ave., Chicago, IL 60637, USA. Email: stigler@uchicago.edu

## Abstract

In the 1950s, the growing importance of Bayesian inference attracted both supporters and critics. In 1958, the philosopher Karl Popper published what he called a paradox that purported to show the foolishness of subjectivist Bayesian inference. In correspondence, Jimmie Savage pointed out what he considered an error in his reasoning, but Popper was unmoved and did not change his example when he reprinted it. It is argued that Popper did not acknowledge that the new subjectivists reasoned with distributions, not simply expected values. Excerpts from Jimmie Savage's correspondence at the time with Popper and with Allen Wallis highlight the issues involved at this time when Bayesian statisticians were seeking general acceptance. **Keywords:** Bayesian inference, falsification, Jimmie Savage, Karl Popper, R. A. Fisher, William Feller

In the 1950s, the theory of frequentist statistics was undergoing rapid development, and Bayesian statistics was suffering from what might be called benign neglect. Early enthusiasm by Harold Jeffreys and by Bruno de Finetti in the 1930s had not excited many followers, at least until Jimmie Savage's 1954 book (Savage, 1954), and for much of the 1950s Savage's remained a lonely voice. The occasional mentions by others tended to be critical, such as those by William Feller and R. A. Fisher. In his 1950 text, Feller wrote: 'Unfortunately Bayes' rule has been somewhat discredited by metaphysical applications ... The modern method of statistical tests and estimation is less intuitive but more realistic. It may be not only defended but also applied.' (Feller, 1950, pp. 124-125; Gelman & Robert, 2013; Stigler, 2013). In his 1956 book, Fisher wrote: 'Certainly cases can be found, or constructed, in which valid probabilities a priori exist, and can be deduced from the data. More frequently, however, and especially when the probabilities of contrasted scientific theories are in question, a candid examination of the data at the disposal of the scientist shows that nothing of the kind can be claimed.' (Fisher, 1956, p. 17). Feller, Fisher, and many others felt that the lack of either an experimental basis for the choice of the prior, or a model-based theory for the prior, should rule out the application of Bayes theorem. If the case for the prior was firmly based, Feller or Fisher would agree that Bayes was the best way to go. The same is not the case for an objection levelled by Karl Popper in 1958.

# **Karl Popper**

Probability figures prominently in the work of the philosopher Karl Popper, at least superficially. He is perhaps most famous in logic for his influential ideas about falsification as the test of scientific theories, often tied to Neyman-Pearson testing, and he had strong opinions about the basis of probability (he could be described as a frequentist). He used symbols such as C(h, e), intended to be the strength of confirmation that evidence *e* provides for a claim *h* about the value of a probability of, say, a Head in a single toss of a coin, and he introduced a measure of relative corroboration for

Received: October 4, 2023. Revised: April 2, 2024. Accepted: April 20, 2024

© The Royal Statistical Society 2024. All rights reserved. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

Stigler

competing theories, based on the likelihood function when considering coin tosses. His view on weight of evidence was closely tied to his ideas on falsification: Single number measures of agreement with null hypotheses. For him, evidence seemed to have been a logical concept, useful only in judging the falsification of hypotheses.

In the 1950s, he wrote three short papers related to the question, and in the third of these, published in February 1958, he offered what he considered to be a paradox that destroyed the subjective approach. He claimed to be accepting the premises of subjective Bayesians, with specific reference to the weight of evidence, citing John Maynard Keynes (1921) and Jack Good (1950), stating:

'Considerations of the "weight of evidence" lead, within the subjective theory of probability, to paradoxes which, in my opinion, are insoluble within the framework of this theory.... By the subjective theory of probability, or the subjective interpretation of the calculus of probability, I mean a theory that interprets probability as a measure of our ignorance, or of our partial knowledge, or of the rationality of our beliefs, in the light of the evidence available to us.' (Popper, 1958)

Here is how he presented his 'paradox':

'In order to save space, I shall explain the problem of the weight of evidence merely by way of presenting one of the paradoxes to which I referred above. It may be called the "paradox of ideal evidence."

'Let z be a certain penny, and let a be the statement "the *n*th (as yet unobserved) toss of z will yield heads." Within the subjective theory, it may be assumed that the absolute (or prior) probability of the statement a is equal to 1/2, that is to say,

$$P(a) = 1/2 \tag{1}$$

'Now let *e* be some *statistical evidence*; that is to say, a *statistical report*, based upon the observation of thousands or perhaps millions of tosses of *z*, and let this evidence *e* be *ideally favourable* to the hypothesis that *z* is strictly symmetrical—that it is a "good" penny, with equidistribution.<sup>1</sup> We then have no other option concerning P(a, e) than to assume that

$$P(a, e) = 1/2$$
 (2)

'This means that the probability of tossing heads remains unchanged, in the light of the evidence *e*; for we now have

$$P(a) = P(a, e) \tag{3}$$

'But this formula has to be interpreted as asserting that *e* is, on the whole, (absolutely) *irrelevant* information with respect to *a*.

'Now this is a little startling; for it means, more explicitly, that our so-called "*degree of rational belief*" *in the hypothesis, a, ought to be completely unaffected by the accumulated evidential knowledge, e*; that the absence of any statistical evidence concerning *z* justifies precisely the same "degree of rational belief" as the weighty evidence of millions of observations which, corroborate or *prima facie*, support or confirm or strengthen our belief.' (Popper, 1958, pp. 295–296; italics in original)

Let me restate this: consider tossing a penny *n* times. The penny shows no obvious physical reason to favour one side over the other, and the subjective view would say that a priori the probability of Heads in a single trial is 1/2, whether (a) we take it as a known fact that P(Heads) is exactly 1/2, or that (b) we simply have no idea what P(Heads) is but take it to be subjectively plausible that it is distributed over the unit interval with mean 1/2. Then, in either case (a) or (b), if asked to state a value for the probability of Heads on the first toss, the subjectivist would assess it as 1/2, Popper's (1). Now, he says, suppose we have observed a large number of tosses (even millions) with an

182

<sup>&</sup>lt;sup>1</sup> At this point Popper inserted a parenthetical statement to the effect that e did not constitute the entire detailed record, rather a summary such as "in a million observed tosses, heads occurred in 500,000 +/– 20 cases."

essentially even balance of Heads and Tails; then, he says, the subjectivists will still say 1/2, his (2). That is, the subjectivist says 1/2 with or without that particular huge amount of evidence, his (3).

The value  $\frac{1}{2}$  is indeed the same as the initial assessment, but to infer there is no change in the weight of evidence is nonsense. If that subjectivist (b) was asked to state the probability of Heads a priori, the number may well be 1/2, but the true a priori assessment would not be a number but rather a prior distribution, perhaps in this case a uniform distribution over [0,1], with mean 1/2. And, after gathering the evidence, his assessment would be an *a posteriori* distribution, perhaps with mean 1/2, but now very tightly concentrated about 1/2. The posterior view is greatly changed by the evidence.

How could Popper have come to such a conclusion? The answer is his limited view of what the Bayesians were doing. He stated that view clearly: 'The *fundamental postulate of the subjective theory* is the postulate that degrees of rationality of beliefs in the light of evidence exhibit a *linear order*: that they can be measured, like degrees of temperature, on a one-dimensional scale.' (Popper, 1958, p. 296). But that denied the approach of de Finetti, Good, Savage, and in fact Bayes and Price, (Stigler, 2018), a view that reasoned from prior distribution to posterior distribution. Popper's limited view of a Bayesian approach was one that looked only at the posterior mean; he was correct in assessing that limited goal, but that was not the modern Bayesians' view. He addressed a straw Bayesian, not a Bayesian scientist uninterested simply in falsification of null hypotheses.

To be specific, suppose you have a large jar of coins. One possibility is that they are all perfectly fair, with P(Heads) = 1/2. Another possibility is that they are all biased, but that the mean of all the P(Heads) is 1/2. Popper understood correctly that these two possibilities, while very different in fact, are statistically indistinguishable if you choose a single coin at random and toss it, recording the outcome. They are also indistinguishable if you repeat the tosses a very large number of times with the same coin, with or without replacement, and ignore all outcomes except the last.

What Popper missed was that the only interesting version of this, from the standpoint of induction, is where you make many tosses with the same coin that is chosen for that first toss and consider a summary of the long list of outcomes as your evidence *e*. If the biases truly vary, the probability of a sequence of outcomes is quite different from the case where all the coins are fair. Laplace made exactly this point in one of his earliest papers (Laplace, 1774, Section VI; Stigler, 1986 at p. 361 and 375–8). And most important, the posterior distribution reflects the evidence even in the unlikely case that Heads and Tails are in balance. Yes, the posterior mean is 1/2 in that case too, but the confidence in that value, the true weight of evidence, is justly captured by the posterior distribution's concentration.

Even most frequentists, including Feller, would agree that if the assumptions of the subjectivist are granted, Bayes theorem is not only a valid description of the learning process, but also the best such model. Popper accepted those assumptions but denied the conclusion. He was correct in saying the a priori expectation of the probability was, given the assumptions, equal to ½, and he was correct in saying that, with equal numbers of Heads and Tails, the *a posteriori* expectation (sometimes referred to as Laplace's Rule of Succession, a term Popper was familiar with) was also ½. But he was wrong in assuming that was all the subjectivist was interested in. Quite to the contrary, it was the full posterior distribution that reflects the learning which is captured beautifully by Bayes theorem in this case (see Appendix).

Popper's erroneous conclusion has not gone unnoticed. Jeffrey (1965, pp. 183–184) and Skyrms (1977) have pointed out the error without specifically tying it to lack of understanding of the subjectivists' use of Bayes Theorem. But Popper had stated that the point of the critique was precisely the understanding of those very subjectivists of the role of Bayes theorem in assessing weight of evidence.

# **Popper and Savage**

Zappia (2019) has uncovered in Yale's archives correspondence between Popper and Jimmie Savage from 1958, subsequent to the appearance of this 'Third Note.' Savage's book on the subjective approach had appeared in 1954 (Savage called it 'personal probability' not subjective probability), and Jack Good had published a book on the weighting of evidence from a Bayesian perspective in 1950 that Popper had cited (Popper did not cite Savage). Good was unimpressed by Popper's article, but he wrote to Savage, calling his attention to it, and Savage wrote to

Popper on 25 March, saying that he saw no paradox. He gently described what a real subjectivist would say, avoiding mathematical statements, perhaps because he sensed Popper's grasp on mathematics to be weak:

'What you call a paradox in section 3 does not seem paradoxical to me. It is perfectly true that I may attach probability 1/2 to a coin toss in very different circumstances. For example, I may do so if I feel no reason to stake a prize on heads in preference to tails, though I have little or no explicit information about this particular coin, and also I may do it if a tremendous amount of evidence indicates that there is no reason to prefer staking a prize on heads or tails. This situation does not seem to me startling in itself, and if introduction of the term "degree of rational belief" makes it sound startling, that seems to me another fault of the term. It might help, though, if I were to point out one great difference between the two situations for a subjectivist. The evidence that you call e guarantees not only that the probability of heads on the next toss is 1/2, but that the probability of heads on the next two tosses is 1/4, and so on, at least if certain side conditions that I think you meant to imply are fulfilled. (It must, for example, be ruled out at the start that the coin has any tendency to alternate heads with tails, which would be consistent with the data in the elliptical form in which you described it.) On the other hand, in those cases where I attach probability 1/2 to the next throw of a coin without having intimate experience with it, I do not attach probability 1/4 to two consecutive heads, 1/8 to three consecutive heads, etc. Indeed, a person who does that is one whose a posteriori opinions are not affected by evidence. This rigidity is approximately justified in the case of a person who knows e, because he bases his belief on an overwhelming amount of evidence, but it is not justified where there is what we are sometimes inclined to call no evidence.' (Letter from L. J. Savage to K. Popper, 25 March 1958).

Popper replied in a 2 April letter, but not to this criticism. He gave ground on the term 'degree of rational belief', writing, 'I do not think that the use of the term degree of rational belief, or any similar term, is essential to the subjective interpretation of probability: I agree here with your letter. But for precisely this reason, I do not think that, in avoiding this or a similar term, the subjective theory can be rescued from my criticism. My criticism is, fundamentally, that the occurrence of statistical effects in physics has nothing to do with our knowledge, or belief, or behaviour. Therefore, we need an objective theory of probability.' Popper was busy; 'I cannot read your book for at least the next five months.' Savage's mathematical criticism had apparently gone over his head. Savage replied politely 28 April, seeing that mathematical issues were not useful, and effectively signed off, saying, 'We cannot profitably discuss these delicate questions by mail, especially as busy as we both are now.'

Popper's rejection of the subjective view was not based on cogent philosophical grounds, but on a mathematical misunderstanding of the position that he was attacking, in particular as developed by Jack Good and by Jimmie Savage, but going back to Thomas Bayes and Richard Price and John Maynard Keynes and Frank Ramsey and Harold Jeffreys and Bruno de Finetti. Popper was unmoved; subsequent to the correspondence with Savage, he reprinted the Third Note in 1959, essentially unchanged.

Popper's ignorance of the full Bayesian argument seems remarkable. He had earlier put forth a set of axioms for probability (Popper, 1938) that left no impression on the world, but did give evidence of the serious lacunae in his knowledge. The only works on probability that short piece cited were Carnap (1937) on terminology, Mazurkiewicz (1932) (who gave another axiom system), Keynes (1921) *Treatise on Probability*, and an Oxford text by Levy & Roth (1936) that does include Bayes theorem, but not in a way that would have been helpful here. Of these, only Keynes might have keyed him in that he was missing something, but it is clear that Popper looked no further than to see how those authors defined probability, not how it was used. Even if he was only interested in axiomatic treatments in 1938, at that time, he conspicuously missed Kolmogorov's pathbreaking 1933 tract that, for the first time, showed how conditional probability could be dealt with successfully at the axiomatic level.

Popper's view of probability was restricted to work within mathematical logic, work epitomized by that of Carnap (1950) that was fully rigorous but of little relevance to interesting active problems in scientific induction. Popper gets little more than a footnote in modern works such as von Plato (1994). He made no reference to Galton's pathbreaking insights of the 1880s, developed

further by Pearson, Neyman and Fisher, which may not have been philosophically rigorous, but they were scientifically useful. Savage's book was an attempt to repair that.

#### Savage and Fisher

An interesting artefact from October 1957 testifies that Savage was looking deeply into arguments about Bayes at that time. Savage and Allen Wallis were both at Chicago in the 1950s and quite close at that time, even when at a distance. Wallis had arranged for both of them to have an early dictation machine called the Soundscriber, where they could record and exchange messages via small green discs that could be played on ordinary record players. A number of those discs survive with Wallis's papers at the University of Rochester. The following is part of a transcription from one disc from Jimmie Savage to Allen:

'Hello, Allen, this is Saturday, October 20 [1957]. I'm beginning this disc just now because of something in my hand that I didn't want to forget to mention to you. ...

'The trifle in my hand that I wanted to mention is that you may remember, that in the dittoed draft of my book [Savage, 1954], one of the earliest ditto drafts, I attributed to R. A. Fisher the expression of the idea that since the a priori distribution washes out in a large sample, that there ought to be some intrinsic way of analyzing the data in itself without ever postulating a prior distribution at all. I don't remember whether I criticized that argument on the spot, but it's not valid, of course, because the prior distribution does wash out, does so only exponentially, and the rate at which it washes out does depend considerably on what prior distribution it is. Thus for example, since I'm firmly convinced that extrasensory perception does not exist, it would take tremendous amounts of data, of relevant opposing data, to bring me to the opposite point of view. Well, the thing was, we couldn't find this passage anywhere in Fisher and, when I wrote him, he said it was ridiculous, he never could've said any such thing, but Bob Schlaifer has found the reference for me, and it's in Paper 24 of Fisher's collected papers, it's the passage that straddles pages 286 and 287 and I just thought you might like to look at it for yourself.'

Here is the relevant paragraph from Fisher (1934):

'As an axiom this supposition [a uniform prior distribution] of Bayes fails, since the truth of an axiom should be manifest to all who clearly apprehend its meaning, and to many writers, including, it would seem, Bayes himself, the truth of the supposed axiom has not been apparent. It has, however, been frequently pointed out that, even if our assumed form for f(x)dx be somewhat inaccurate, our conclusions, if based on a considerable sample of observations, will not greatly be affected; and, indeed, subject to certain restrictions as to the true form of f(x)dx, it may be shown that our errors from this cause will tend to zero as the sample of observations is increased indefinitely. The conclusions drawn will depend more and more entirely on the facts observed, and less and less upon the supposed knowledge a priori introduced into the argument. This property of increasingly large samples has been sometimes put forward as a reason for accepting the postulate of knowledge a priori. It appears, however, more natural to infer from it that it should be possible to draw valid conclusions from the data alone, and without a priori assumptions.—If the justification for any particular form of f(x) is merely that it makes no difference whether the form is right or wrong, we may well ask what the expression is doing in our reasoning at all, and whether, if it were altogether omitted, we could not without its aid draw whatever inferences may, with validity, be inferred from the data. In particular we may question whether the whole difficulty has not arisen in an attempt to express in terms of the single concept of mathematical probability, a form of reasoning which requires for its exact statement different though equally well-defined concepts.' (Fisher, 1934, pp. 286–287)

We can now see that Fisher was not intending to argue for the validity of Bayesian methods; rather he saw a way to support an approach without priors, specifically the fiducial method he had first considered in 1930.

## Conclusion

The remarkable fact is not simply that Popper would address issues involving inductive reasoning in 1958 without a grasp of the new role of conditional probability distributions in induction; his approach to the philosophy of science did not need that as far as he could judge. But, it was an error to do so in criticizing the approach taken by subjective Bayesians, an approach that was specifically grounded in the manipulation of such distributions. This led to him attacking them as not learning from evidence, when they were using a method that was designed to do exactly that, if their assumptions were accepted as Popper granted for the argument, a method that was and is now widely seen as the perfect tool for that purpose. He had not read Savage and even though he cited Good, his acquaintance there could not have been deep. Popper was not fully familiar with the subjective probability of the modern Bayesians, but he was confident that he did not like it. In any event, his critique fell on deaf ears.

In truth, the tools of probability are more difficult to deeply comprehend than they appear at first blush. As Augustus De Morgan wrote, 'Everyone makes errors in probabilities, at times, and big ones.' (Graves, 1889, p. 459) Popper would have been further emboldened because the paradox he thought he had found supported his preconceived ideas. This would not be the last time he made an error of this type (e.g. Popper & Miller, 1983). Even Isaac Newton erred in answering a question in probability, presenting a simple argument that, while it happened to give the right answer, was unsound in principle (Stigler, 2006).

Conflicts of interest: None declared.

# Appendix

### Bayes theorem for binomial

While well known to all statisticians, reactions from some philosophers to a circulated version of the paper suggest that a brief review may be helpful. Consider the simple situation considered by Popper, inference about the probability  $\theta$  of a success from data for *n* Bernoulli trials:  $P(X = k|\theta) \propto \theta^k (1-\theta)^{n-k}$  for  $0 \le \theta \le 1$  and  $0 \le k \le n$ . In any Bayesian analysis,  $\theta$  is taken as having a prior (to observing data) distribution  $p(\theta)$  over the interval [0,1]. For Laplace and several people since, who have invoked some 'principle' with a name like the 'principle of insufficient reason,' the choice has been to take  $p(\theta)$  as the uniform distribution over [0,1], a density with expected value  $\frac{1}{2}$ . It is a wonderful mathematical fact that if n = 1, the marginal distribution of X depends only on the expected value of the prior distribution,  $\frac{1}{2}$  here, and in that case, the distribution of X is indistinguishable from the case where  $\theta$  is known to be  $\frac{1}{2}$  with certainty. It is also true that this fact fails magnificently when n > 1. Both of these facts were well known to Laplace.

A Bayesian analysis here would move to the joint distribution of X and  $\theta$ , namely  $p(k, \theta) = P(X = k|\theta)p(\theta)$ , and then to the posterior distribution  $p(\theta|X = k) = p(k, \theta)/P(X = k)$ , where P(X = k) is the marginal distribution of X, which for the uniform prior distribution was already known by Thomas Bayes to be P(X = k) = 1/(n + 1) for k = 0, 1, ..., n. In that case (the only one considered by Popper), the posterior distribution of  $\theta$  given X = k is proportional to the density  $\theta^k(1-\theta)^{n-k}$ , known as a beta density over [0,1]. This would give the posterior expectation as (k + 1)/(n + 2) and variance  $(k + 1)(n-k + 1)/[(n + 2)^2(n + 3)]$ . The posterior expectation has come to be known as Laplace's Rule of Succession.

Popper had taken the n = 1 fact as indicating  $\theta = \frac{1}{2}$  rather than that there was a total lack of evidence, and then adopted the posterior expectation with millions of cases equally split as saying the same. He was oblivious to the posterior variance and was hence led to his stating 'that our so-called *degree of rational belief* in the hypothesis, *a*, ought to be completely unaffected by the accumulated evidential knowledge, *e*.' The idea that the Bayesian inference he critiqued neatly incorporates the evidence had escaped him.

# Data availability

No new data were generated or analysed in support of this research.

#### 186

#### References

- The correspondence between Savage and Popper is courtesy of the Yale University Library. The recording of Savage is courtesy of the University of Rochester Library.
- Carnap, R. (1937). The logical syntax of language. Harcourt, Brace.
- Carnap, R. (1950). Logical foundations of probability. University of Chicago Press.
- Feller, W. (1950). An introduction to probability theory and its applications. Wiley.
- Fisher, R. A. (1934). Two new properties of mathematical likelihood. Proceedings of the Royal Society, Series A, 144(852), 285–307. https://doi.org/10.1098/rspa.1934.0050
- Fisher, R. A. (1956). Statistical methods and scientific inference. Oliver & Boyd.
- Gelman, A., & Robert, C. (2013). Not only defended but also applied: The perceived absurdity of Bayesian inference. *The American Statistician*, 67(1), 1–5. https://doi.org/10.1080/00031305.2013.760987
- Good, I. J. (1950). Probability and the weighing of evidence. Charles Griffin.
- Graves, R. P. (1889). Life of Sir William Rowan Hamilton. (Vol. 3). Hodges Figgis.
- Jeffrey, R. (1965). The logic of decision. McGraw-Hill.
- Keynes, J. M. (1921). A treatise on probability. Macmillan.
- Laplace, P. S. (1774). Mémoire sur la probabilité des causes par les évènemens. Mémoires de mathématique et de physique, presentés à l'Académie Royale des Sciences, par divers savans, & lû dans ses assemblées, 6, 621–656. Translated in Stigler (1986).
- Levy, H., & Roth, L. (1936). Elements of probability. Clarendon Press.
- Mazurkiewicz, S. (1932). Zur Axiomatik der Wahrscheinlichkeitsrechnung. Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie Classe III, 25, 1–4.
- Popper, K. (1938). A set of independent axioms for probability. Mind, 47(186), 275–277. https://doi.org/10. 1093/mind/XLVII.186.275
- Popper, K. (1958). A third note on degree of corroboration or confirmation. The British Journal for the Philosophy of Science, 8(32), 294–302. Reprinted as an Appendix in his 1959 The Logic of Scientific Discovery, pp. 406ff. https://doi.org/10.1093/bjps/VIII.32.294
- Popper, K., & Miller, D. (1983). A proof of the impossibility of inductive probability. *Nature*, 302(5910), 687–688. Refutation by Isaac Levi, Nature 310: 433. https://doi.org/10.1038/302687a0
- Savage, L. J. (1954). The foundation of statistics. Wiley.
- Skyrms, B. (1977). Resiliency, propensities, and causal necessity. Journal of Philosophy, 74(11), 704–713. https:// doi.org/10.2307/2025774
- Stigler, S. M. (1986). Laplace's 1774 memoir on inverse probability. Statistical Science, 1(3), 359–378. https://doi. org/10.1214/ss/1177013620
- Stigler, S. M. (2006). Isaac Newton as a probabilist. *Statistical Science*, 21(3), 400–403. https://doi.org/10.1214/ 088342306000000312
- Stigler, S. M. (2013). Bayesian inference: The Rodney Dangerfield of statistics? *The American Statistician*, 67(1), 6–7. Comment on Gelman and Robert (2013). https://doi.org/10.1080/00031305.2012.747448
- Stigler, S. M. (2018). Richard Price, the first Bayesian. Statistical Science, 33(1), 117–125. https://doi.org/10. 1214/17-STS635
- von Plato, J. (1994). Creating modern probability. Cambridge University Press.
- Zappia, C. (2019). Paradox? What paradox? On a brief correspondence between Leonard Savage and Karl Popper. https://papers.ssrn.com/sol3/papers.cfm?abstract\_id=3478165. To appear in Research in the History of Economic Thought and Methodology.