Keynes and the Weight of Evidence

Lok C. Chan

1   Introduction

In this chapter I shall examine the key claims about the weight of evidence
made by Keynes. My main purpose is to argue against Keynes' idea that the
weight of evidence for a belief is the absolute amount of relevant evidence we
have for that belief. Keynes defines *relevance* in terms of the difference between
prior and posterior probabilities - an account still accepted by many Bayesians.
Nevertheless, he notices that sometimes intuitively relevant evidence could
change the probability of the hypothesis in question through conditionalization.
He then proposes a "strict" definition to capture those cases. As we shall see, the
second definition has the problematic implication of regarding any proposition
P not implied by another proposition H and our background knowledge as
being relevant to H. After examining and criticizing his notion of evidential
weight as the amount of relevant evidence, I turn to an important remark
by Keynes that, I shall suggest, opens the door to the core epistemological
problem the very idea of evidential weight is supposed to address. I call this

the problem of epistemic urgency.

2 Weight and Relevance

Probability, in Keynes's view, is defined as a probability relation between two propositions. These are objective and logical relationship that are independent of our subjective state, even though it can only be known through intuition. In order to render the numeric assignment of *some* probabilities possible, Keynes proposes a number of basic judgments about probability. We make *judgments* of indifference, for instance, when we judge that probabilities are uniformly distributed for a set of events.

More pertinent to us is the *judgment of irrelevance.* Keynes' observation is that we often can judge whether one proposition E counts as being relevant to another proposition H by considering whether the probability of H would change on the supposition that E is true. Keynes's example is that, in a typical urn example with some black and white balls, if we want to know the probability of a white ball being randomly chosen, the *color* of the ball would not change its probability of being chosen, so the idea is that a ball's probability of being chosen conditional on being (say) white is the same as the probability of the ball being chosen in general.[1] So, Keynes proposes that evidence E is irrelevant to the proposition H if and only if:

---

[1] John Maynard Keynes, *A Treatise on Probability* (Macmillan; Co., Limited., 1921), 59.

$$P(H|E) \neq P(H)$$

Now that the notion of relevance has been introduced, we come to Keynes' idea of the weight of evidence. He first brings our attention to the fact that when we consider the conditional probability of the hypothesis in question under all relevant evidence, the resultant number constitutes the *balance* between favorable and unfavorable evidence.[2] For instance, we may say that when the probability is above 0.5, then the overall evidence is somewhat in favor of the hypothesis. Of course, the balance changes as we gather more relevant evidence, and it might go from favorable from unfavorable depending on the nature of the new evidence.

However, as Keynes points out, this is not the only way in which we think of how evidence can render a hypothesis acceptable or unacceptable, for we not only care about how much the current evidence favors the hypothesis, but we also concern ourselves with the *amount* of evidence involved in calculating the balance of the evidence, and Keynes calls this measure the *weight* of evidence. But, unlike the balance of the evidence, which can go either direction, the weight of evidence can only go up as we gather more relevant evidence. In Keynes' words, "New evidence will sometimes decrease the probability of an argument, but it will always increase its 'weight.' "[3]

A critical point Keynes makes in the just mentioned quote is that evidential

---

[2]Keynes, 78.

[3]Keynes, 78.

weight forms a linear relationship with the amount of *relevant* evidence: bringing in new evidence *always* increases the weight of the evidence. As a matter of fact, Keynes suggests that as a matter of definition, to introduce new relevant evidence is simply another way to say that the weight of evidence has increased.[4] This is the claim that I shall endeavor to examine in this chapter. In the next section, we will review a criticism made by Popper against this particular notion of evidential weight.

3　The Paradox of Ideal Evidence

Popper has challenged Keynes' definition, which is adopted by Carnap and Bayesian epistemologists, by using what he calls the *paradox of ideal evidence.*[5] He asks us to consider a certain coin with an unknown bias: let N be the proposition "the next toss of the penny will yield heads". Now, what should $P(N)$ be? He suggests, either by appealing to intuition or the Principle of Indifference, it ought to be clear that $P(N) = 0.5$.[6]

Now let I be what he calls *the ideal statistical evidence* in favor of the idea that the coin in question is a fair one. Popper's example is to let I be the proposition that 'in a million tosses, the coin landed on heads roughly half a million times.' The exact number is not important, as long as the number of heads and tails would make it practically certain that the coin is fair - the

---

[4]Keynes, 78–79.

[5]Karl Popper, *The Logic of Scientific Discovery* (Routledge, 2002), 425.

[6]It should be noted that Popper is not attacking the principle of indifference in this context. That is, for this argument he is willing to grant that Bayesians have some way of arriving at $P(N)$ - it could be by indifference, through elicitation, etc.

same point could be made using 10 millions instead of a million. Now, given we have ideal evidence I, what is the probability of N? Popper claims that it would have to be 1/2. So

$$P(N|I) = P(N) = \frac{1}{2}$$

However, as discussed earlier, evidence E is irrelevant to the hypothesis H if and only if

$$P(N|I) = P(N)$$

If $P(N|I) = P(N) = 1/2$, this means that the ideal evidence is also irrelevant evidence. Popper then concludes:

> Now this is a little startling; for it means, more explicitly, that our so-called 'degree of rational belief' in the hypothesis, $[N]$, ought to be completely unaffected by the accumulated evidential knowledge, $[I]$; that the absence of any statistical evidence concerning [the hypothesis that the coin is fair] justifies precisely the same 'degree of rational belief' as the weighty evidence of millions of observations which, prima facie, support or confirm or strengthen our belief.[7]

What is 'startling' about this? This returns to Keynes' initial observation: *something* is changed by the ideal evidence, but the probability fails to reflect this, so this serves as a counterexample to Keynes' claim that the increase of

[7]Popper, 426.

relevant evidence always increase its weight. The intuition is that our attitude toward N must have shifted by I somehow, so I must be relevant evidence, in a sense not captured by Keynes' definition. The weight of evidence is precisely what is missing.

Nevertheless, Popper's argument contains a sleight of hand that shifts between two ways of thinking about N's probability. The argument begins by asking, rather innocuously, for your prior for N, but the ideal evidence I Popper immediately introduced is not for N but for the hypothesis *about N*, so the distinction missing here is the one between having ideal evidence for N and having ideal evidence for $P(N) = 1/2$. We clearly ought to be careful about the distinction between the two and the different kinds of evidence they require. The ideal evidence for N would presumably be evidence that makes it practically certain that N is true, so, the ideal evidence for N would point the balance of the evidence toward being overwhelmingly in favor of the coin having a bias toward heads. But this is clearly not what I is; instead, what I is supporting is $P(N)$ being a certain value. So the ideal evidence is not directly relevant to the outcome of the next toss.

Popper is explicit about *that*, but what he is not explicit about is *this*: he has convinced us that I is both evidentially ideal and relevant to $P(N) = 0.5$. That, however, is a misdirection, because he immediately starts talking the conditional probability on I, *not* of $P(N) = 0.5$, but of N. Now, Popper speaks as if this is the only way Bayesian can incorporate the ideal evidence by thinking of the situation in terms of $P(N|I)$, but that is not the only way.

A more statistically minded Bayesian may say that the probability of the toss landing on heads is non-sensical is begin with, unless it is conditional on a hypothesis about the bias of the coin, θ. In other words, instead of saying $P(N) = 0.5$, we should really think of the situation as $P(N|\theta) = 0.5$ So the prior here would not be on N, but on θ, which in standard Bayesian statistics would just be modeled as a Beta Distribution. In this context, it is clear that I is not relevant to N, but to θ, so according to this line of thought, there is nothing startling about I being irrelevant to $P(N)$.

Even in the way I just proposed, relevance and weight still come apart; because, now there is no straightforward way to talk about how I has increased our weight of evidence for θ in the Keynesian sense. First, to help Keynes' account of weight, we need to show:

$$P(\theta = 0.5|I) \neq P(\theta = 0.5)$$

But since Beta is a uniform distribution, this is clearly impossible.

Second, I will narrow the spread the distribution for θ, so we can calculate use the cumulative distributions to say that

$$P(0.49 < \theta < 0.51) \neq P(0.49 < \theta < 0.51|I)$$

But the problem is that the modes for both distributions will remain the same, and they will have the same expected value, 0.5. This means that when

push comes to shove, the posterior distribution will make no practical difference from the prior distribution. I shall return to this in greater detail in chapter 3, when we consider some of the formal details regarding evidential weight.

4   The Strict Definition of Relevance

It must be pointed out that Keynes is explicitly ambivalent about the notion of weight he proposes. He in fact anticipates counterexamples such as the paradox of ideal evidence. He points out that sometimes complex evidence E might be irrelevant to H, but its logically simpler components E_1 and E_2 could be relevant to H.

What Keynes has in mind seems to be this: suppose you have reasons to believe that the urn in front of you contains 50 white balls and 50 black balls, so your probability for H, drawing a black ball, would be 0.5. Further suppose E_x says 'a ball with the color x is drawn', where $x = w$ for white and $x = b$ for black. Let $E = E_b \wedge E_w$. At this point, an argument could be made that your degree of belief for H should stay at 0.5, since sample mean is exactly that. In fact, this is a situation in which the use of Laplace's "rule of succession" is justified..[8] According to this rule, $P(H|E_b \wedge E_w)$ is

$$\frac{1 + 50}{2 + 50 + 50} = \frac{51}{102} = \frac{1}{2}$$

[8]'... Thus one finds that when an event has happened any number of times running, the probability that it will happen again next time is equal to this number increased by 1, divided by the same number increased by 2.', Pierre-Simon Laplace, *Philosophical Essay on Probabilities* (Spring, 1995), 11

But suppose E_b and E_w are not gathered at the same time: say E_b was received first - this would sensibly nudge your opinion slightly toward the opinion that the proportion of black balls is higher than 50%. This could be verified by the rule of succession: $P(H|E_b) = \frac{51}{101} \neq P(H)$, so E_b is relevant to H. *Mutatis mutandis* , the same argument could be made for E_w. So we have a situation where getting E_w and E_b together does not increase the weight of the evidence, but getting them sequentially would lead to an increase. Though he, too, has to speculate Keynes' precise meaning, Carnap[9] gives a similar but formal example.

To avoid this type of problems, Keynes proposes a 'strict' definition of irrelevance. While Keynes calls this the 'strict' definition, the purpose of the definition is to relax the requirement for relevance, so that evidence that is irrelevant "as a whole" may still be relevant due to its components. The strict definition states: E is irrelevant to H on background knowledge K if and only if there is no proposition ϕ, such that

1. ϕ is inferable from $(E \wedge K)$.
2. ϕ is not inferable from K.
3. $P(H|\phi \wedge K) \neq P(H|K)$

This roughly states that E is irrelevant to H on K if and only if there is no proposition implied by E and K together, but not K by itself, such that it would change the probability of H. So in the case where $E = E_b \wedge E_w$, E is *not* irrelevant to H, because E_b, which is inferable from $E \wedge K$, does raise

[9]Rudolf Carnap, *Logical Foundations of Probability* (Chicago]University of Chicago Press, 1950), 419.

the probability H. Also, this definition also means that any proposition that satisfies the three conditions is relevant. In other words, a piece of evidence is relevant so long as a part of it is relevant.

So it would appear that the paradox of ideal evidence is subject to a similar analysis. That half a million out of a million tosses landed on head implies that the coin handed on heads for half a million times, and conditional on the tosses that landed on heads alone would seem to change the probability of next toss landing a heads.

Nevertheless, there is something peculiar about this way of explicating evidence. It seems that we are distorting the evidence instead of analyzing it. If I was the result of a randomized experiment, the integrity of the inference X. In the case of the ideal paradox, the definition works, because it matches our intuition about I's relevance, but a deeper worry is whether the definition is too accommodating that even intuitively irrelevant evidence could be analyzed as being relevant. In the next section, we will see that the situation is actually worse. Carnap has proven that Keynes' strict definition has the consequence that almost any proposition can be interpret as relevant evidence for any hypothesis.

5 The Problematic Implication of the Strict Definition

What Carnap has shown is as follows: Keynes' restricted definition has the consequence that any proposition E not implied by background knowledge K would be relevant to hypothesis H conditional on K.[10] The idea is that Keynes's new definition has essentially counted anything not implied by background knowledge as being relevant, because the problem is that any proposition that is not implied by K will always have implied components that are relevant to H, however weakly. The argument is roughly that, for any H on background knowledge K, if J is not implied by $H \wedge K$ and R is relevant evidence to H, J can always be shown to be relevant by drawing trivial inferences that that involves R, such as $J \vee \neg R$.

For a formal argument, first consider these two assumptions about relevance.

1. If E is relevant, then $\neg E$ is also relevant.
2. If a hypothesis is not implied by our background knowledge, then there is some proposition such that it is relevant to the hypothesis.

Note that these assumptions are explicitly accepted by Keynes, so we do not have to find them plausible; we just have to accept them for the sake of this particular argument. In addition, Carnap has also proven that Keynes' definition of irrelevance is equivalent to the following: E is irrelevant to H on

---

[10]Carnap, 420.

K if and only if $P(H|K) = P(H|K \wedge E)$ *or* $E \wedge K$ is a contradiction. This alternative definition will be useful in the proof.

Carnap's proof is roughly as follows.[11] Suppose both evidence J and hypothesis H are contingent on background knowledge K. Also suppose a different piece of evidence R that is relevant to H on K - R must exist because of assumption 2 above. For a *reductio*, suppose J is irrelevant to H. Now let E be $(R \wedge \neg J)$. E is either relevant or irrelevant to H, and from this we can make an argument by cases.

Suppose E is relevant: this means that E is relevant to H by virtue of Keynes' stricter definition; because, E contains R, which is relevant to H on K. By assumption 1, since E is relevant to H, $\neg E$ is also relevant to H. But this means J is relevant to H as well: $\neg E$ is equivalent to $(\neg R \vee J)$, which is inferable from J. Since an implication of J is relevant to H given K, this means J must be relevant to H.

Now suppose E is irrelevant. Since R is relevant to H on K, the only way for E to be irrelevant is for $(\neg J \wedge K \wedge R)$ to be a contradiction (by Carnap's alternative definition), so $\neg(\neg J \wedge K \wedge R)$ has to be true. That is logically equivalent to $(\neg J \wedge K) \rightarrow \neg R$. Now, by assumption 1, $\neg R$ is relevant, because R is relevant. The above formula means that from $\neg J$ plus K, we can infer $\neg R$, so this means $\neg J$ is relevant, which entails J is relevant.

So, we see that both possibilities implies that J is relevant. By *reductio*, then, J is relevant. Thus, if J is not inferable from K, then J is relevant to H.

---

[11]Carnap, p.420. My presentation also partly follows Peter Gärdenfors, "On the Logic of Relevance," *Synthese* 37, no. 3 (1978): 351–67

In addition, if J is relevant to H on K, then J is not inferable from K, since one of the conditions of the strict definition of relevance is that relevant evidence cannot be a proposition that is a logical consequence of our background knowledge. So, this means that J is relevant to H on K if and only if J is not inferable from K.

In other words, any proposition not implied by our background knowledge is by definition relevant, so Carnap's result shows that the new definition is unhelpful: if we and his suggestion is that we should simply stick with the simpler definition. Modern Bayesian have followed Carnap in accepting Keynes' basic definition of relevance as a minimal requirement for any confirmation measure. I think at this point, it is clear that Keynes' goal of forging the link between the amount of relevant evidence and the weight of evidence has failed, and we should forgo the idea that the increase of relevant evidence always increase its weight.

6 Epistemic Urgency and the Value of Evidence

At this point, we must ask: why should we care about the weight of our evidence? We ought to make this clear before getting lost between the formalism about probability and the metaphor of balance and weight. Here's one proposal: Probability, as the saying goes, is the very guide of life, so evidential weight is a meaningful concept only insofar as it captures aspects of rational decision

making not captured by the balance of evidence. Keynes clearly realizes this:

> For in deciding on a course of action, it seems plausible to suppose that we ought to take account of the weight as well as the probability of different expectations.[12]

Keynes remains non-committal, but I think he puts his finger on the crux of the problem. He points out that probability theorists, such as Bernoulli, often assumes what Carnap later calls "the requirement of total evidence": when making judgments about probability, we ought to consider all available evidence.[13][14]. Now Keynes asks: does this principle also suggest that we ought to find as much evidence as possible? But Keynes suggests that this cannot be the case:

> There clearly comes a point when it is no longer worth while to spend trouble, before acting, in the acquisition of further information, and there is no evident principle by which to determine how far we ought to carry our maxim of strengthening the weight of our argument.[15]

This line of inquiry, which is unfortunately not further developed by Keynes, points to the general direction of the problem about evidential weight I intend to examine in this dissertation. Let us call this *the problem of epistemic urgency*, that is, the problem of determining whether a willingness, or the lack thereof, to seek further evidence is rational or justified. This is what evidential

---

[12]Keynes, *A Treatise on Probability*, 84.
[13]Keynes, 84.
[14]Carnap, *Logical Foundations of Probability*, 211.
[15]Keynes, *A Treatise on Probability*, 84–85.

weight should tell us: a belief with weighty evidence is one where our decision to stop seeking further evidence is justifiable.

A. J. Ayer[16] puts the matter in this follow way: should "total evidence" include relevant evidence that I do not yet have in possession? The answer seems to be obviously affirmative: if finding the truth value of some proposition P would sufficiently sway the balance of my evidence, then I should definitely acquire it. Thus the principle of total evidence seems to suggest that I am also rationally compelled to consider some evidence I do not yet have, but Ayer, as does Keynes, points out that this cannot be the whole picture: taken as a rule of rationality, this means we should never stop acquiring unless we are certain that we have acquired all relevant evidence.

This theme is picked up by Ramsey, perhaps the most famous critic of Keynes' theory of probability. Interestingly, however, he has proven essentially the opposite conclusion reached by Ayer and Keynes: his argument shows that getting more evidence is always better. How can this be? If Ramsey's result is accepted without qualification, then the problem of epistemic urgency would be essentially a pointless one; therefore, we must examine what exactly Ramsey's proof entails.

Ramsey's argument is roughly that, *if* we assume A to be a perfect Bayesian agent and that new information does not cost anything, then A will be no worse off or better off getting new evidence.[17] A perfect Bayesian agent

---

[16] *Probability and Evidence* (Macmillan, 1972).

[17] F. P. Ramsey, "Weight or the Value of Knowledge," *British Journal for the Philosophy of Science* 41, no. 1 (1990): 1–4, also see I. J. Good, "On the Principle of Total Evidence," *British Journal for the Philosophy of Science* 17, no. 4 (1966): 319–21.

is someone who studiously updates her opinions based on Bayes' rule and then act by choosing the action that maximize her expected utility. An intuitive example is perhaps more helpful than reproducing the proof.[18] Suppose we have three hypotheses about the content of an urn in front of us:

1. H_b: 90 black balls and 10 white balls

2. H_w: 10 white balls and 90 black balls

3. H_n: 50 white balls and 50 black balls.

We then start by assuming $P(H_b) = P(H_w) = P(H_n) = 1/3$. Suppose we win \$1 by picking the correct hypothesis. Our expected payoff for choosing each hypothesis would be the same at $1/3$. Nevertheless, we are allow to sample with replacement as many times as we wish. Should we get more evidence? To begin, at this point, the probability of getting a black ball is the same as getting a white ball. Let E_b be "a black ball is drawn" and E_w for white balls. So:

$$P(E_b) = P(H_b)P(E_b|H_b) + P(H_w)P(E_b|H_w) + P(H_n)P(E_b|H_n)$$

$$1/3(0.9) + 1/3(0.1) + 1/3(0.5) = 0.5$$

And $P(E_w) = 1 - P(E_b) = 0.5$. So, in the event of drawing a black ball from the urn, we would update our belief like so:

[18]This example is adapted from Isaac Levi, "The Weight of Argument," in *Fundamental Uncertainty: Rationality and Plausible Reasoning*, ed. Silva Marzetti Dall'Aste Brandolini and Roberto Sczzieri (Palgrave MacMilan, 2011), 39–58

$$P(H_b|E_b) = \frac{P(H_b)P(E_b|H_b)}{P(E_b)} = \frac{1/3(0.9)}{0.5} = 0.6$$

Similarly, applying the calculation on the other hypotheses, we get:

$$P(H_w|E_b) = 0.067$$

$$P(H_n|E_b) = 0.333$$

Similar argument can be made assuming E_w, that is, a white ball is chosen. In that case $P(H_w|E_w) = 0.6$. If we are an ideal Bayesian agent, we should pick H_b if E_b, and pick H_w if E_w. So our expected payoff after getting one piece of evidence is $0.5(0.6) + 0.5(0.6) = 0.6$, which is better than what it would have been had we decided not to get any evidence. The net gain in payoff, $0.6 - 0.33 = 0.27$, is referred to as *the sample value of information* in the decision theory literature.[19] Ramsey calls this the *value of evidence* in his note.

It turns out that we would be even better off if we were to draw from the urn again. Suppose the first draw yields a black ball. Let us refer to our state of belief after the first draw as $H_b', E_b', ..$ and so on. One notable change is that $P(E_b') = 0.7132$ and $P(E_w') = 0.2868$. If we draw again and get a black ball, this means:

$$P(H_b'|E_b') = 0.757$$

[19]Howard Raiffa and Robert Schlaifer, *Applied Statistical Decision Theory* (Harvard, 1964), 89–90. For a more digestible presentation see Robert Winkler, *An Introduction to Bayesian Inference and Decision* (Probabilistic Publishing, 2010) sec.6.3.

$$P(H'_w|E'_b) = 0.009$$

$$P(H'_n|E'_b) = 0.233$$

If a white ball were to be drawn:

$$P(H'_b|E'_w) = 0.21$$

$$P(H'_w|E'_w) = 0.21$$

$$P(H'_n|E'_w) = 0.58$$

So our expected payoff if we were to draw from the urn again is: $0.7132(0.757) + 0.2867(0.58) = 0.706$, which is an improvement over just drawing once. Ramsey's proof shows that this will go on forever and we will never be worse off.

What should we make of Ramsey's proof? As Ramsey seems to have refuted Keynes' intuition that sometimes more evidence is not worth our time, the crucial assumptions he makes need to be addressed: if Ramsey's assumptions hold, then it's always rational to look for more evidence for beliefs that are not necessarily true, because we are never worse off by doing it. Further, if it's always rational to look for more evidence for beliefs that are not necessarily true, then epistemic urgency is not a problem. So, if Ramsey's assumptions hold, then epistemic urgency is not a problem.

My point is that if we accept Ramsey's result and accept the underlying

conditions, then we should never refuse looking for new evidence, since we always end up with the same or better utility. In other words, more evidence is almost always better. If this is true, it naturally follows that there is no such a thing as the problem of epistemic urgency, since the very idea of it is to ask how it is possible that sometimes it is rational to look for new evidence. *But* I take it that we would agree with Keynes that more evidence is not always worth the effort, so it must mean that Ramsey's assumptions cannot be all true. What I wish to consider are the reasons for which these assumptions do not hold, and what the implications are.

Perhaps the most crucial assumption here is that new information has no cost. Without this assumption, it becomes evident that Ramsey's argument is limited in applicability. Suppose it costs us 25 cents for each sample. This means that we would be gaining only $0.27 - 0.25 = 0.02$ in expected payoff for the first draw, and the second draw would definitely not be worth the additional 25 cents. Cost might also enter into consideration in different forms, e.g., computational cost or memory. Of course, there was no indication that Ramsey thought that information was free. That in itself is trivial point. What is interesting is to consider an epistemology that takes into account the cost and value of evidence. In the next chapter, I shall reconstruct a Peircean account for such a view.

Ramsey makes another assumption that Keynes does not accept, that is, the ideal Bayesian agent can express all of her beliefs in precise numerical values. Keynes famously denies this: he holds that some probabilities can at

best be expressed as intervals, and some are totally indeterminate.[20] Ramsey's assumption has come to known as the "Dogma of Precision." In contrast, some have accepted the existence of "imprecise probabilities".

In the context of evidential weight, the question of whether probabilities should be sharp is a significant one, for one view of evidential weight is that it is a measure of the precision of a probability. We already have a relevant case in the context of the paradox of ideal evidence. We saw that what the ideal evidence manifest itself not in terms of the expected value, but in the width of the posterior credible interval. An intuition would be to say that we should make our decisions based on the interval and not the point estimate.

Roughly speaking, the idea is that before observing the ideal evidence regarding the bias of the coin, our degree of belief is rather vague - it could be as vague as $[0, 1]$ - we know that it's within this interval but we have no reason to think that it's one particular value or another. As we take into account the ideal evidence, we then narrow down our estimate to 0.5 to a high degree of precision. Theories of imprecise probabilities formalize this intuition, by holding that belief should be represented by a set of distribution, instead of one. Some theorists, such as Walley[21], thinks that evidential weight just is the difference between the upper and lower bounds of this interval. We shall return to this issue in our development of a Peircean view of evidential weight in the next chapter.

---

[20]Keynes, *A Treatise on Probability*, ch. III.
[21]Peter Walley, *Statistical Reasoning with Imprecise Probabilities*, 1991

Reference

Ayer, A. J. *Probability and Evidence.* Macmillan, 1972.

Carnap, Rudolf. *Logical Foundations of Probability.* Chicago]University of Chicago Press, 1950.

Gärdenfors, Peter. "On the Logic of Relevance." *Synthese* 37, no. 3 (1978): 351–67.

Good, I. J. "On the Principle of Total Evidence." *British Journal for the Philosophy of Science* 17, no. 4 (1966): 319–21.

Keynes, John Maynard. *A Treatise on Probability.* Macmillan; Co., Limited., 1921.

Laplace, Pierre-Simon. *Philosophical Essay on Probabilities.* Spring, 1995.

Levi, Isaac. "The Weight of Argument." In *Fundamental Uncertainty: Rationality and Plausible Reasoning*, edited by Silva Marzetti Dall'Aste Brandolini and Roberto Sczzieri, 39–58. Palgrave MacMilan, 2011.

Popper, Karl. *The Logic of Scientific Discovery.* Routledge, 2002.

Raiffa, Howard, and Robert Schlaifer. *Applied Statistical Decision Theory.* Harvard, 1964.

Ramsey, F. P. "Weight or the Value of Knowledge." *British Journal for the Philosophy of Science* 41, no. 1 (1990): 1–4.

Walley, Peter. *Statistical Reasoning with Imprecise Probabilities*, 1991.

Winkler, Robert. *An Introduction to Bayesian Inference and Decision.* Probabilistic Publishing, 2010.