ARTICLE

Likelihoodism and Guidance for Belief

Tamaz Tokhadze[1] 📛

Abstract
Likelihoodism is the view that the degree of evidential support should be analysed and measured in terms of likelihoods alone. The paper considers and responds to a popular criticism that a likelihoodist framework is too restrictive to guide belief. First, I show that the most detailed and rigorous version of this criticism, as put forward by Gandenberger (2016), is unsuccessful. Second, I provide a positive argument that a broadly likelihoodist framework can accommodate guidance for comparative belief, even when objectively well-grounded prior probabilities are not available. As I show, the shift from non-relational to comparative probabilities opens up a new space for addressing the belief guidance problem for likelihoodism.

Keywords Likelihoodism · Belief guidance · Non-relational and relational probabilities · Comparative belief · The problem of priors · The ratio form of Bayes' Theorem

1 Introduction

Contemporary statistics is home to a couple of competing paradigms for interpreting scientific data as evidence. For the past 60 years or so, the two most popular approaches to statistical inference have been the *frequentist* paradigm and the *Bayesian* paradigm.

A central procedure of frequentist statistics is the so-called Null Hypothesis Significance Testing (NHST). A significance test starts with a hypothesis, called the "null hypothesis", which is examined against some relevant outcome or data. Simply put, NHST says that if a null hypothesis renders certain outcomes as highly improbable and if such an improbable outcome occurs, then the null hypothesis should be rejected.

While the guiding idea behind NHST seems plausible, many have found the method to be fundamentally defective.[1] Moreover, in certain fields of science, primarily in the social,

---

[1] An immediate problem with NHST is that it embodies a defective form of inductive reasoning; even if a hypothesis renders a certain observation as unlikely, the observation might still support the hypothesis; this is so because the observation might be *even more improbable* if the hypothesis is false. To demonstrate this, suppose two individuals share a copy of a rare allele; only 1 in 10 000 have it. As siblings share half of their alleles on average, $P(rare\ allele|siblings) = 0.5 * 0.0001$, which is a very small number. However, if the two individuals are unrelated the probability is much lower: $P(rare\ allele|unrelated) = 0.0001 * 0.0001$. Hence, the data supports the sibling hypothesis, even if the hypothesis renders the data quite unlikely. See Sober (2008, 48–58) for an accessible discussion.

✉ Tamaz Tokhadze
tamaz.tokhadze@gmail.com

[1] University of Sussex, Brighton BN1 9RH, UK

behavioural, and biomedical sciences, some of the important results that relied on NHST failed to be replicated.[2] This new evidence showing the important lack of replication in these fields puts an additional strain on frequentism, so much so that there is an increasing call for some kind of statistical reform.

Many critics of frequentism see Bayesianism as providing superior methods of data-analysis (Dienes 2011; Wetzels et al. 2011; Kruschke 2013). The key characteristic of Bayesianism is the use of the so-called *prior probabilities*. Unlike frequentism, Bayesian theory requires a probability distribution over both the sample space and statistical hypotheses.[3] A probability distribution over statistical hypotheses is called *prior distribution*. A prior distribution encodes how likely the competing hypotheses are before the relevant evidence comes in.

Certainly, the indispensability of priors in data analysis is the Achilles heel of Bayesian methods. The problem is that, in many contexts, there seems to be no objective, uncontentious way to fix priors. And due to this unmistakably subjectivist component in Bayesian methods, many are quite reluctant to give up on the traditional frequentist methods.[4] While some Bayesians have proposed various theories for grounding "objective" prior probabilities, none of these proposals has been generally accepted.[5] Hence the problem of the subjective priors continues to haunt Bayesian methods.

Likelihoodism can be seen as an attempt to overcome the frequentism-Bayesian controversy by paving the way between "the illogic of the frequentists and the subjectivity of the Bayesians" (Royall 1997, XIV). The central principle of likelihoodism is the so-called *Law of Likelihood* according to which an outcome, E, is evidence for a hypothesis, A, over its competitor, B, when E is more likely if A is true than if B is true. Like frequentism (and unlike Bayesianism), likelihoodism requires only a probability distribution over the sample space and not over hypotheses themselves. And like Bayesianism (and unlike frequentism), likelihoodism holds that the impact of evidence on any two hypotheses is wholly determined by the likelihoods of these hypotheses.[6] Hence, likelihoodism endorses some of the true-and-tried principles from both frequentism and Bayesianism, without relying on controversial NHST or subjective priors.[7]

Certainly, there are several problems associated with likelihoodism. The problem that will preoccupy us in this paper is that, unlike frequentist and Bayesian approaches, likelihoodism remains silent on matters of belief. According to an orthodox likelihoodist position, strong evidential support does not licence either a categorical belief or a degree of belief in a proposition. This point can be illustrated by the following quick example.

---

[2] For a general philosophical discussion of the replication crisis see Romero (2019). For an explanation/diagnosis of the crisis see Bird (2021).

[3] NHST only considers probability distribution over sample space (i.e., how likely an outcome is on the supposition of a statistical hypothesis) and not over statistical hypotheses themselves (i.e., how likely the relevant statistical hypotheses are).

[4] There have been some attempts to marrying the two paradigms together, in a unified frequentist-Bayesian theory. But so far, the most serious disputes between the two approaches are still raging. See Mayo (2018) for a lengthy discussion.

[5] For a positive, systematic account of objective or impermissive Bayesianism see Williamson (2007; 2010). For a critical discussion, see Meacham (2014).

[6] This view on the impact of evidence is called the *Likelihood Principle*. Frequentism is in tension with the principle as it allows various non-likelihood related factors to influence the impact of evidence. For a detailed discussion of the Likelihood Principle see Berger and Wolpert (1988) and Gandenberger (2015).

[7] For an influential statement of the likelihoodist program see Royall (1997). For a more philosophically rich discussion, see Sober (2008) and Bandyopadhyay et al. (2016).

Consider a TB test which is 99% reliable: the test would correctly report the presence of TB in 99% of cases and falsely report the presence of TB in 1%. Now suppose that the test indicates the presence of TB in a randomly chosen individual in the UK. According to the Law of Likelihood, the test result is strong evidence that the individual has TB. However, it is still highly unlikely that the individual has TB, as the incidence rate of TB is extremely low in the UK. It is overwhelmingly more likely that the test has given a false positive rather than true positive report. So, even if there is a strong piece of evidence that the person has TB, this does not license the conclusion that the person probably has TB.

In the above example, the assumption about the prior probability of the disease was unproblematic. The incidence rate or frequency data of the disease fixes the priors in an objective, uncontentious manner. But, in many scientific settings, a prior distribution cannot be fixed in the same way. Sometimes the relevant frequency data is unknown. And, in some cases, it seems incoherent to suppose that frequency data can provide a basis for assigning prior probabilities. For instance, what type of frequency information can ground a prior probability distribution to, say, the general theory of relativity or the anthropomorphic climate change hypothesis (more on this in Sect. 2.2)?

Now, likelihoodists eschew the use of prior probabilities when priors are not supported by empirical evidence (Edwards 1972; Royall 1997; Sober 2008). And given that empirically grounded priors are unavailable in many, and arguably, in most settings, likelihoodist methods seem to be practically useless for science (Gandenberger 2016, 12).

While the lack of belief guidance has been identified as a problem for likelihoodism for some time, more recently, Gandenberger (2016) has articulated the worry in a detailed and rigorous way. As he has concluded, due to the lack of belief guidance, likelihoodism is not a viable alternative to either frequentism or Bayesianism.

This paper argues that, contrary to the received view, a likelihoodist framework can accommodate belief guidance. Like Gandenberger, I focus on *comparative beliefs*; that is, beliefs of the form "*A* is more probable than *B*".[8] As I shall argue, when there is no objective basis for assigning prior probabilities to hypotheses, rational comparative beliefs can still be formed, without invoking Bayesian subjective priors. Following Salmon (1990), my main argumentative strategy is to move from *non-relational probabilities* of individual hypotheses to *comparative evaluations* of competing hypotheses. A non-relational probability of a hypothesis, *A*, is commonly represented by a point-valued probability: e.g. when *A* is assigned a probability of, say, 0.6 (I also allow non-relational probabilities to be represented by *intervals* or *sets* of probability distributions). By contrast, comparative evaluations of competing hypotheses *A* and *B* do not require us to assign any non-relational probabilities to them. For instance, scientists may not have an objective basis for assigning

---

[8] Also, like Gandenberger, I am concerned with belief guidance for *simple* rather than *composite* (or catchall) hypotheses. Composite hypotheses are disjunctions of mutually exclusive simple hypotheses. For instance, the hypothesis H_1: "the coin is fair" is simple while the hypothesis $\neg H_1$: "the coin is *not* fair"—is composite: as $\neg H_1$ is the disjunction of all the specific alternatives to H_1.

The likelihoods of composite hypotheses are sensitive to prior probabilities. For this reason, such likelihoods raise several issues for likelihoodism that go beyond the scope of this paper. See Bandyopadhyay et al. (2016, Appendix to Chapter 2) for a discussion and argument that the likelihoodist account of evidence can deal with composite hypotheses.

non-relational probabilities to *A* and *B*, but they can still rationally judge that *A* and *B* are roughly equally plausible.

So, contrary to Gandenberger's criticism, I will conclude that, even when there is no objective basis for assigning prior probabilities, it is feasible to guide comparative belief without collapsing into subjective Bayesianism.

The paper runs as follows. Section 2.1 gives a general, broad-brush overview of the main aspects of likelihoodism, and Sect. 2.2 gives a precise statement of the distinct problem that likelihoodism faces concerning belief guidance. Section 3 provides detailed analyses and criticism of Gandenberger's (2016) anti-likelihoodist argument. In Sect. 4, I put forward a positive, likelihoodist account of guidance for comparative beliefs. This account utilises the so-called ratio form of Bayes' Theorem. I will illustrate both the applicability and limits of likelihoodist belief guidance by analysing two examples: one from cognitive neuroscience and one from philosophy. I conclude in Sect. 5 that likelihoodism can provide substantive guidance for comparative belief.

2 Setting the Stage

2.1 Two Tenets of Likelihoodism

The core of likelihoodism consists of (*i*) a *comparative*, *relational* conception of evidential support and (*ii*) the likelihood ratio measure of the degree of relational support. The first is qualitative and the second is a quantitative aspect of likelihoodism. In what follows, I will characterise and explicate each of these aspects, starting with the likelihoodist view of (evidential) support.

To explain the likelihoodist view of support, it is useful to contrast it with a more orthodox, non-relational view. A theory of support is non-relational when it defines support for an individual hypothesis, without contrasting the hypothesis to its alternative, competitor hypothesis. For instance, consider the standard Bayesian view which I call *Support-IP* (IP for Increase in Probability):

> Support-IP: For any hypothesis *H*, evidence *E*, and personal probability function *P*, *E* supports *H* relative to *P* iff $P(H|E) > P(H)$.

Support-IP defines support in terms of the increase-in-probability relation (or confirmation). And Support-IP is a non-relational view because support for a hypothesis is defined without appealing to any *competitor hypothesis*.

By contrast, the likelihoodist view of support is *inherently relational*, as it requires two competitor hypotheses to define the relation of evidential support. This view is expressed by the so-called *Law of Likelihood* (LL), which roughly says that for any two competitor hypotheses *A* and *B*, *E* supports *A* more strongly than *B* iff *E* is more likely on the supposition that *A* than on the supposition that *B*. More precisely:

> LL: For any two competitor hypotheses *A* and *B*, *E* supports *A* over *B* iff *A* confers greater probability on *E* than *B* does: $P(E|A) > P(E|B)$.

Why accept LL over its Bayesian competitors? The main strength of LL—according to its supporters—is that it provides an objective, inter-personally justifiable criterion for evidential support. LL defines support in terms of two likelihoods; i.e., the probabilities of the following form: $P(Evidence|Hypothesis)$. A likelihood encodes the empirical content

of a hypothesis; that is, what the hypothesis says about evidence. For instance, let $h =$ "25% of philosophy undergraduates are introverts", and let $e =$ "randomly chosen philosophy undergraduate is an introvert". There is a certain logico-conceptual relationship between h and e that is articulated by likelihood $P(e|h)$. And even if we have no clue about the prior probability of h and e, the likelihood of $P(e|h)$ is still objectively given: $P(e|h) = 0.25$.

The example illustrates what Hawthorne (2005, 278) has called the *publicness of likelihoods*. So, even if two agents disagree about the prior probability of h, they can still agree on the value of the likelihood, $P(e|h)$.[9]

Fixing likelihoods is not always as easy as in the above example. But even the so-called *subjective* Bayesians—that is, Bayesians who allow the multitude of coherent prior distributions as rationally permissible—grant that likelihoods can be objectively well-grounded in many scientific contexts (Edwards et al. 1963).

In addition to LL, likelihoodists provide a measure of (comparative) evidential support, which quantifies the basic idea behind LL; so that the degree of evidential support between A and B is defined as the ratio of their respective likelihoods.

Relational Measure of Support: The degree to which evidence E supports a hypothesis A over its competitor B equals the ratio of their respective likelihoods. In symbols:

$$\frac{P(E|A)}{P(E|B)}$$

Ratios of likelihoods (R_L, for short) has useful mathematical properties. Whenever the data is more likely on A than on B, the R_L is always greater than 1.[10] And the better E fits A over B, the greater the ratio. Following Royall (1997), it is common to postulate an arbitrary cut-off point for characterising weak and strong evidence. For instance, we can say that if $1 < R_L < 8$, then E provides *weak* evidence for A. And if $R_L \geq 8$, then E provides *strong* evidence for A.[11]

The combination of LL and the measure of relational support, R_L, comprises the core of likelihoodism.

Many (e.g. Fitelson 2007; 2011; Mayo 1996; 2018) have criticised these core principles on various grounds. This paper will not address any potential difficulties with either LL or the likelihoodist measure of support. Rather, the focus is on the applicability of these principles to the question of belief guidance.

The next section gives a detailed statement of the problem of belief guidance for likelihoodism.

---

[9] By the standard definition of conditional probability, likelihoods are still mathematically related to priors: as $P(E|H) = P(E \text{ and } H)/P(H)$. But this mathematical connection between likelihoods and priors does not imply that we cannot make an independent sense of $P(E|H)$, without appealing to the prior probability of H. For one thing, there is an important logical asymmetry between $P(E|H)$ and $P(E \text{ and } H)$ and $P(H)$. Knowing the values of $P(E \text{ and } H)$ and $P(H)$ fixes the value of $P(E|H)$. But not the other way around. So we can make an independent sense of $P(E|H)$ without assuming that the prior probabilities are known or well-defined. For a more detailed discussion see Sober (2008, 38–41).

[10] Except when $P(E|B) = 0$.

[11] The reader should not attach too much significance to the cut-off point 8. Certainly, whether evidence E provides strong evidence for A over B is a context-sensitive matter and depends on the evidence and hypotheses in question (Bandyopadhyay et al. 2016, 24).

2.2 The Problem of Belief Guidance

It has long been recognised that likelihoodist methods for interpreting data as evidence do not, by itself, determine what one ought to believe. The point is well-illustrated by Royall (1997, 2–4) by distinguishing three types of questions regarding the analysis of evidence:

(Q1)   What does the present evidence support?
(Q2)   What should you believe in light of your present evidence?
(Q3)   What should you do in light of your present evidence?[12]

The core of likelihoodism only applies to the first question. By contrast, answering questions (Q2) and (Q3) require more than the information about the likelihoods. To illustrate this, consider a physician, "*you*", who investigates whether a patient, "*Eve*", has a skin disease. Eve has taken a test, and the result came up positive. The probability of a true-positive is quite high, 95%, and the probability of a false-positive is quite low, 5%. From this information, we can already answer Royall's first question by using the Law of Likelihood (*LL*): the test result strongly supports the hypothesis that Eve has the skin condition over the hypothesis that she does not have it.[13]

But this information is insufficient to answer either question (Q2) or (Q3). To answer (Q2), you also need to know the prior probability of the disease. If the disease is quite rare and only 1 in 10 000 people have it, then the test result does not license the belief that Eve has the disease. So, to answer (Q2), you need to know both the relevant likelihoods and prior probabilities.

Regarding (Q3): whether you should give any medication to Eve, based on the test result, depends not only on your probabilities but on relevant *utilities*. If the common medication against the disease is harmless, then you can reasonably prescribe it to Eve, even without knowing the exact prior probability of the disease. And, if the medication can be harmful to a healthy person, you would prescribe it only when you are quite certain that Eve has the disease.

To sum up then: even if *E* strongly supports *A* over *B* this will not imply that *E* justifies either categorical belief in *A* or comparative belief in *A* over *B* (more on this in the next section). Hence, the core of likelihoodism only applies to question (1) and not to questions (2) or (3).

But how does a likelihoodist answer the belief question? As one of the motivating ideas of likelihoodism is to avoid the subjectivity of Bayesianism, likelihoodists cannot rely on *subjective priors* to guide beliefs. So, instead, they should rely on *objectively well-grounded priors*.

Generally speaking, there are two broad strategies for grounding objective priors: by appealing to (*i*) empirical information about frequencies or (*ii*) some a priori principle (e.g. the so-called *Principle of Indifference*). However, as I discuss next, both strategies are problematic for likelihoodists (Gandenberger 2016).

---

[12] For a discussion about a broader significance of Royall's three questions for the philosophy of statistics see Bandyopadhyay and Forster (2011, Sect. 2).

[13] $P(+$ result | Eve has the disease$) = 0.95$ and $P(+$ result | Eve does not have the disease$) = 0.05$. So, the likelihood ratio is $0.95/0.05 = 19$. Hence, the positive test result provides strong evidence that Eve has the disease.

Regarding the first strategy: it is widely accepted that frequency data can provide objective justification for fixing prior probabilities. Frequency information is often *out there*, independent of our knowledge, as when a certain disease has some objective incidence rate in the population. For instance, the incidence rate of TB in England is approximately 9.2 per 100,000. And we can estimate the prior probability of a randomly selected individual in England to have TB, based on this frequency data.

Using empirically informed priors to guide belief seems to meet the likelihoodist standard of objectivity. But what if such priors are unavailable? One popular likelihoodist position is that, when empirically well-grounded priors are unavailable, the only rational doxastic response is a suspension of judgment. Such a view about rational belief is well-summarised and endorsed by Sober (2008, 32):

> When prior probabilities can be defended empirically, … you should be a Bayesian. When priors and likelihoods do not have this feature, you should *change the subject*. In terms of Royall's three questions …, you should shift from question (2), which concerns what your degree of belief should be, to question (1), which asks what the evidence says.

Unfortunately, though, Sober's proposal is unsatisfactory. Sober himself points out that in many cases, empirically informed priors are simply unavailable. As he (2008, 26) articulates the point:

> There is a world of difference between this quotidian case of medical diagnosis and the use of Bayes' theorem in testing a deep and general scientific theory, such as Darwin's theory of evolution or Einstein's general theory of relativity. … When we assign prior probabilities to these theories, what evidence can we appeal to in justification? We have no frequency data as we do with respect to the question of whether *S* has tuberculosis. If God chose which theories to make true by drawing balls from an urn (each ball having a different theory written on it), the composition of the urn would provide an objective basis for assigning prior probabilities, if only we knew how the urn was composed. But we do not, and, in any event, no one thinks that these theories are made true or false by a process of this kind.

Sober's view about the scarcity of frequency data is the majority view in the philosophy of science and statistics; as most would agree that a prior probability assignment cannot be defended empirically in many cases of interest. Hence, the proposal that the talk of belief is inappropriate in the absence of frequency data seems to lead to a sceptical view of science, where scientific theories and models are rarely useful for guiding belief.

So, can likelihoodists pursue an alternative strategy and appeal to some a priori principle(s) to ground objective prior probabilities? This strategy is also problematic for likelihoodists, as they are generally sceptical about the prospects of grounding priors on a priori principles.[14]

To illustrate this, let us consider the most prominent a priori rule for fixing priors, the so-called *Principle of Indifference* (PoI). PoI roughly says that if you have no reason to favour a proposition over its competitor, then you should assign equal probabilities to them. More generally and precisely:

---

[14] See Sober (2008, 27–28).

PoI: Let U be a finite set of all mutually exclusive and exhaustive hypotheses; if an agent has no evidence that favours any member of U over any other, then for all x in U, $P(x) = \frac{1}{|U|}$, where $|U|$ is the cardinality of U.

Likelihoodists are sceptical towards PoI for two separate but interconnected reasons. First, it is a common practice for scientists to consider a handful of competitor hypotheses, at any given time. In most cases, no one thinks that the considered hypotheses exhaust the space of all serious possibilities. And scientists rarely know all members of the set of realistic hypotheses. But the application of PoI depends on such a set, whose members and cardinality are explicitly known. For instance, consider the contemporary theories of quantum gravity. There are just a couple of well-articulated theories of gravity, and no working physicist would think that these hypotheses exhaust the space of all possible realistic hypotheses. Of course, one can negate the disjunction of the competing hypotheses, and hence fill the space of all possibilities. But this manoeuvre leads to a second problem. The problem is that there is more than one way to carve this logical space. For instance, assume that scientists focus on only three *specific* competitor theories of quantum gravity: Q_1, Q_2, and Q_3. So, in total, they must consider four competitor hypotheses: $\{Q_1, Q_2, Q_3$, and $\neg(Q_1 \vee Q_2, \vee Q_3)\}$. Assuming that one is indifferent between these four hypotheses, PoI mandates to assign the probability of $1/4$ to each. But it is possible to carve the space of possibilities in a more coarse-grained or fine-grained manner (for instance, by introducing another specific theory of quantum gravity). Different carvings would have licensed different priors. PoI, in itself, does not settle which carvings should be favoured.[15]

All such a priori rules for deriving priors are relative to the set of competitor hypotheses; hence the two problems I have mentioned are not restricted to PoI and apply to other a priori rules for deriving priors.

To wrap up the above: According to the standard likelihoodist position, frequency data, essentially, is the only admissible evidence for grounding priors for scientific hypotheses. And as such data is unavailable for most scientific hypotheses, the likelihoodist methods seem practically useless for science.

Of course, one can simply deny that the lack of belief guidance is a problem. To paraphrase Sober, when empirically grounded priors are unavailable, one must simply change the subject and answer the evidence question instead of the belief question. This paper will not argue that such a response is illegitimate. But I do not expect that this response would convince the critics. Hence, I shall proceed by presupposing that belief guidance is a genuine problem for likelihoodism.

In the remaining sections of the paper, I shall argue that this received view on the inapplicability of likelihoodist methods to the belief question is incorrect.

Before I defend my positive proposal, first, I need to address a general worry against the very possibility of likelihood-based guidance for belief. The worry has been articulated in a detailed, rigorous manner by Gandenberger (2016). The next section provides a detailed analysis and critique of Gandenberger's argument.

---

[15] While likelihoodists think that PoI is problematic even with *discrete* cases, there is also a well-known Bertrand's paradox that poses problems for PoI with respect to *continuous* probabilities (where one cannot straightforwardly appeal to the "finest" partition of the space of possibilities). Some Bayesians (e.g. Williamson 2007; 2010) have provided novel, nuanced defences of PoI. It is beyond the scope of the paper to evaluate these defences as I am solely concerned with why likelihoodists think that PoI is wrong.

3 An Argument Against Likelihood-based Belief Guidance

Gandenberger (2016) has articulated an argument against the possibility of deriving belief guidance from a likelihoodist framework. The argument identifies a principle that, as he claims, all likelihoodists should endorse. He calls this principle "*minimal comparative proportionalism*" (MCP, for short). To quote Gandenberger (2016, 7):

> This principle [*MCP*] says that there is a real number $r > 1$ such that for any pair of hypotheses A and B, a rational agent believes A over B either in an absolute sense or at least to some degree, if its total evidence favours A over B to degree r or greater.

Now, accepting MCP, as he demonstrates, leads to various epistemic paradoxes. Hence, he concludes that one cannot derive rules for belief from the likelihood framework alone.

A simple, but representative counterexample, similar to the one that Gandenberger puts forward, is as follows:

> *Example*: There is an urn colnsisting of 10 tickets, labelled $T_0, T_1, \ldots T_9$, and a machine that selects tickets from the urn, without replacement. For each ticket, the machine will either select the ticket or not: so, it could select all 10 tickets, only some tickets, or no tickets at all. You want to know whether the machine selects the tickets randomly or deterministically. The machine may be selecting the tickets by following a random process, where each ticket has a 50% chance of being drawn. Alternatively, the machine may be following some deterministic rule and select the same set of tickets in each experiment. You do not know which process underlies the selection.

> You have decided to switch the machine on and see which tickets it would select. In the first round, the machine has selected tickets 0, 4, 6, and 8; let us denote the data as d_{0468}.
> Now let h_{random} be the hypothesis that *the machine selects tickets randomly* and let h_{0468} be the hypothesis that *the tickets 0, 4, 6, and 8* were bound to be selected.

> Should you believe h_{random} over h_{0468}?

Now, the likelihoods of the observed data, d_{0468}, on each competing hypothesis are as follows: $P(d_{0468}|h_{random}) = 1/1024$; $P(d_{0468}|h_{0468}) = 1$.[16] Thus, the degree of evidential support of h_{0468} over h_{random} is 1024. The data seems to support h_{0468} quite strongly. So, if we let the threshold value, r, in MCP to be less than 1024, then you ought to believe h_{0468} over h_{random}.

But this is clearly absurd. The data does not make h_{0468} more believable than h_{random}. We know from the outset that, for some deterministic hypothesis h_x, the first trial would inevitably favour h_x over h_{random}. Therefore, the first experiment cannot be interpreted as making any deterministic hypothesis more believable than h_{random}.

Notice that the above-identified problem for MCP would remain intact if we had chosen a higher threshold value than 1024. For any finite value of r, a similar counterexample can

---

[16] On h_{random}, each ticket is equally likely to be selected; hence each possible outcome is equally probable. As each of the 10 tickets is either selected or not, there are 2^{10} or 1024 possible outcomes, the probability of d_{0468}, on the supposition of h_{random}, is $1/1024$. And, on the supposition of the deterministic hypothesis, h_{0468}, the probability of d_{0468} is 1.

easily be devised (by increasing the number of tickets in the urn). Therefore, it is tempting to conclude that there are no reasonable likelihood-based rules for belief.[17]

As I show shortly, the above conclusion is premature.

One can tease out two readings from the original MCP, depending on the position of the "there is a real number r" quantifier in relation to the universally quantified sentence "over pair of hypotheses A and B". These two readings are as follows:

> MCP_{weak}: For any pair of hypotheses A and B, there is a real number r such that a rational agent believes A over B if her total evidence favours A over B to degree r or greater.
>
> MCP_{strong}: There is a real number r, such that for any pair of hypotheses A and B, a rational agent believes A over B if her total evidence favours A over B to degree r or greater.

Any likelihoodist account that endorses MCP_{strong} is susceptible to a type of counterexample identified by Gandenberger. But notice that accepting MCP_{weak} alone does not give rise to the same problem. This is so because MCP_{weak} allows threshold r to vary across *contexts of reasoning*. For instance, if we set threshold r to be equal to 1025 (instead of, say, 1023), then the first experiment would not settle the question of which hypothesis should be believed. Of course, the second experiment can go either of the two ways: (*i*) the machine can select the same set of tickets as in the first experiment or (*ii*) it can select a different set of tickets. If the second possibility is actualised, then the data would conclusively settle the issue in favour of h_{random}. On the other hand, if the machine selects the same set of tickets, then this would provide overwhelming evidence for the deterministic selection process. The probability that the machine selected the same set of tickets in two trials, on the supposition of a random process is $\frac{1}{1024} * \frac{1}{1024} = \left(\frac{1}{2}\right)^{20}$. Therefore, if we set the threshold value in Example to be greater than 1024 and less then 2^{20}, we would have avoided the problem.[18]

I will discuss at the end of the next section, which aspects of an agent's context determine the value of threshold r. But even at this point of argumentation, we have reached an important conclusion: once we dissect MCP into two principles, MCP_{strong} and MCP_{weak}, it becomes evident that Example is only problematic for MCP_{strong}. Hence, Gandenberger overall argument is inapplicable to MCP_{weak}.

Now, it is fairly uncontroversial that likelihoodists should accept MCP_{weak}. After all, if there are normative principles that relate likelihood functions with belief, then there should be *some* value for the ratio of likelihoods that would make A more probable than B.[19] But

---

[17] An anonymous referee has pointed out that Gandenberger's example has some similarities with Howson's (2013) *Santa* example against the Law of Likelihood (LL). See Bandyopadhyay et al. (2017) and Howson (2017) for a debate on the Santa example (and related issues).

I must emphasise that Gandenberger's example is designed to be a counterexample against MCP and not against LL.

[18] My strategy for blocking Gandenberger's objection is similar to a recent defence of the so-called Lockean thesis by Leitgeb (2017, Chapter 3), who allows the Lockean threshold to vary across contexts of reasoning.

[19] More than that, assuming that A and B are mutually exclusive and have non-zero probabilities, it is a consequence of Bayes' Theorem that there is some value for the ratio of likelihoods, $P(E|A)/P(E|B)$, that would make A more probable than B. This is evident from the following theorem of probability calculus:

For any mutually exclusive hypotheses A and B:

$$\frac{P(A|E)}{P(B|E)} = \frac{P(E|A)}{P(E|B)} * \frac{P(A)}{P(B)}$$

MCP_{weak} does not entail MCP_{strong}. And it is not clear why likelihoodists should accept MCP_{strong}. After all, why believe that there is one unique threshold value that should fix beliefs in all reasoning contexts? Even bracketing its paradoxical consequences, the existence of such a unique threshold is rather implausible on its own; and I do not see how likelihoodists can be forced to accept such a principle. Hence the argument against a likelihood-based account of belief is wanting.

Of course, my response here is solely negative, as MCP_{weak}, on its own, is insufficient to derive any belief guidance. It remains to be seen whether a broadly likelihoodist account of belief guidance is tenable.

4 The Case for Likelihood-based Belief Guidance

As MCP_{weak}, by itself, cannot guide belief, some other principle(s) is needed to connect likelihood functions with beliefs. Like Gandenberger, I also focus on belief guidance for comparative belief; that is, beliefs of the following form: "A is more probable than B", where A and B are any two competitor propositions (hypotheses/theories).

There are two additional reasons for focusing on comparative belief. Firstly, as we have seen, likelihoodism endorses a comparative conception of evidential support. Hence, it is to be expected that likelihoodist methods would be better suited to accommodate comparative belief rather than categorical belief.

Secondly, comparative judgements and evaluations are *indispensable* in science. Typical testing in science is contrastive, where rival hypotheses are assessed against relevant evidence. And scientists often do interpret comparative testing in doxastic terms; as, when biologists conclude that the change in allele frequencies in a population is *probably* due to genetic drift rather than due to selection. Such comparative judgements in science seem to be less problematic, from the epistemic point of view, than categorical or non-relational probabilistic judgements.

So, can likelihoodists provide guidance for belief without lapsing into subjective Bayesianism? To answer this question, we need to be more clear about what "lapsing into subjective Bayesianism" means. From Gandenberger remarks, it is clear that by "lapsing into subjective Bayesianism" he means accepting this core subjective Bayesian principle, which I call *Subjectivity*:

> Subjectivity: When objective, empirically grounded priors are unavailable, scientists can rationally assign prior probabilities to hypotheses that reflect their *subjective degrees of belief* in the hypotheses.

Now, from Gandenberger remarks, it is clear that by "prior probabilities" he means precise or point-valued prior probabilities. But, to make Subjectivity more appealing, I do not assume that probabilities are always point-valued. Instead, in some cases, prior probabilities may be represented with *ranges* or sets of probability functions. So, Subjectivity is

---

Footnote 19 (continued)

So, for any given value for the ratio of priors, there is *some* value for the ratio of likelihoods that would make $P(A|E) > P(B|E)$. Hence, MCP_{weak} is not something that a Bayesian—or anyone who accepts the standard definition of conditional probability—can reject.

The above theorem will play a crucial role for deriving guidance for comparative belief in the next section.

assumed to be consistent with situations where scientists represent the probability of *H* by some range, say [0.1, 0.6].

Now, independent from whether we represent priors with points or ranges, the core of Subjectivity is the view that it is rational for scientists to assign priors to *H* based on their subjective degree of belief in *H*.

In what follows, I show how likelihoodists can accommodate belief guidance without accepting Subjectivity. By arguing this, I grant the main premise of Gandenberger's criticism: that scientists cannot appeal to objective, empirically grounded priors in many relevant cases. However, even granted this, we can make sense of rational comparative belief. Let me explain how.

It has been pointed out by Wesley C. Salmon (1990), among others, that when the information about prior probabilities is unavailable, rational comparative belief can be guided via the so-called *ratio form of Bayes' Theorem*[20]:

$$\frac{P(A|E)}{P(B|E)} = \frac{P(E|A)}{P(E|B)} * \frac{P(A)}{P(B)}$$

If we let R_{Post} be the ratio of *posteriors*, R_L the ratio of *likelihoods*, and R_{Prior} the ratio of *priors*, then the theorem can be summarised succinctly as:

$$R_{Post} = R_L * R_{Prior}$$

Now, the ratio form of Bayes' Theorem frees us from the need of knowing the exact, or even approximate prior probability of either *A* or *B* to determine whether *A* is more probable than *B*. All we need to know is the value (or approximate value) of the *ratio* of priors, R_{Prior}, and not the value of priors themselves. And fixing the approximate value of R_{Prior} requires strictly less information than fixing the approximate value of priors. To explain this, we need to differentiate *non-relational* priors from *relational priors*.

Non-relational priors are priors of an individual hypothesis (or a set of hypotheses); when, for instance, we assign a prior of 0.6 to *A*, or a range of [0.1, 0.6] to *A*, we attribute a non-relational probability to *A*. By contrast, relational priors have to do with the relationship between competing hypotheses, *A* and *B*. And we may be rational in believing that *A* and *B* do not differ significantly in their probabilities without knowing their non-relational probabilities. All we need to know is that the ratio of their priors is approximately 1, $P(A)/P(B) \approx 1$. This ratio can be approximated for many competing theories by appealing to such non-subjective characteristics as their overall predictive accuracy, simplicity, explanatory scope, fruitfulness, etc. So, we may have a good objective basis for concluding that hypotheses *A* and *B* are roughly equal in prior plausibility, without knowing their non-relational probabilities. Again, I emphasise that such relational judgments do not require the assignment of non-relational probabilities to the hypotheses in question. Therefore, even when *non-relational priors* cannot be objectively well-grounded, we can still guide comparative belief in a way that meets the likelihoodists standard of objectivity.

Let us illustrate this with an example from cognitive neuroscience. It involves the famous Trolley Problem, which essentially is about whether it is morally permissible/required to sacrifice one innocent life to save several.

---

[20] It is interesting to note that Earman (1992, Chapter 7, Sect. 3) has criticised Salmon's strategy as too restrictive for Bayesianism, for reasons similar to Gandenberger's criticism of likelihoodism. This fact has been brought to my attention by an anonymous referee.

First, we need to distinguish two types of Trolley cases: the *impersonal* cases (otherwise known as the *bystander case)*, where one needs to hit a switch which diverts a runaway trolley that kills one person but saves five; and the *personal* cases (otherwise known as the *footbridge case)*, where one needs to push someone from a bridge to stop a runaway trolley. It is well-known that people respond differently to the impersonal and personal versions of the Trolley Problem. When presented with the bystander case, people tend to answer that one should hit a switch and save five. By contrast, when presented with the footbridge case, most object to pushing someone to divert the trolley.

Greene et al. (2001) used brain scanning techniques to study which brain regions were "activated" when people engaged with the impersonal and personal Trolley problems. They were primarily concerned with the following two hypotheses (I borrow the formulation of these hypotheses from Machery 2014, 258):

H1 People respond differently to moral-personal and moral-impersonal dilemmas because the former elicit more emotional processing than the latter.

H2 People respond differently to moral-personal and moral-impersonal dilemmas because the single moral rule that is applied to both kinds of dilemmas (for example, the doctrine of double effect) yields different permissibility judgments.

Now, Greene et al. (2001) found that the personal cases elicited relatively greater activation of brain regions associated with automatic emotional responses; while the impersonal cases elicited relatively greater activation of brain regions associated with conscious reasoning. Let us denote this neuroimaging evidence as e_{new}.

How should we interpret this neuroimaging evidence? One relatively uncontroversial inference is that the likelihood of e_{new} is higher on the supposition of H_1 than on the supposition of H_2. This is so because H_2, unlike H_1, cannot account for why the two cases elicit the activation of brain regions associated with two very different psychological processes. But, it is unclear whether we can make an informed estimate of the posterior probability of H_1 (or H_2) on this evidence. First, it is unclear how we should estimate the prior probabilities for these hypotheses. And even if priors can be fixed in some non-arbitrary way, there may well be some alternative hypothesis that predicts the evidence far better than H_1. As Machery (2014, 256) puts it:

> …in many cases, cognitive neuroscientists have no sense of the probability of obtaining a particular pattern of brain activation if psychological process p is not recruited by experimental tasks and, as a result, they do not know whether the observed pattern of activation gives them a reason to conclude that the psychological process of interest was involved during the task under consideration.

However, notice that even if we cannot estimate the posterior probability of H_1 and H_2, we can still rationally conclude the evidence renders H_1 more probable than H_2. By the ratio form of Bayes' Theorem, to make this comparative inference, the only required information is that the ratio of likelihoods, $P(e_{new}|H_1)/P(e_{new}|H_2)$, is *greater than the reciprocal of* the ratio of priors, $P(H_1)/P(H_2)$. And it is reasonable to think that the available evidence licenses us to accept this inequality. Therefore, even if we have no clue about the non-relational prior and posterior probabilities of H_1 and H_2, we can still conclude that the former is more probable than the latter, on the relevant evidence.

Of course, I grant that there are many cases where the ratio of priors cannot be estimated in a non-subjective manner. In such cases, judging that say, hypothesis *A* is more probable than *B* may be problematically sensitive to some subjective factors. As an example, consider a hotly debated topic of cosmological fine-tuning. Some background would be required to explain this.

According to contemporary physics, the fact that life exists in our universe depends on the very precise values that the so-called *fundamental constants* of physics take. For instance, if the mass of proton had been slightly different from its actual value, then the complex and stable structures we find in the universe, like galaxies, stars and planets, would not have existed; hence, life would not have existed. So, given the laws that most contemporary physicists accept, the existence of stable structures and, specifically, the existence of life is very improbable. But life does exist in our universe. How can we account for this puzzling evidence?

Some (e.g. Leslie 1989; Hawthorne and Isaacs 2018) think that the likelihood that our universe is fine-tuned for life (denoted as *F*) is roughly the same relative to these two very different hypotheses:

> *G*: The cosmological constants of the universe have been consciously designed by the God of traditional theism (as, if such a God exists, she would create the universe that can support life).
>
> *M*: There exist very many (maybe infinitely many) universes. And most (maybe all) possible values of cosmological constants are actualised in some universe(s). Therefore, it is to be expected that some universe(s) among this vast ensemble of universes is fine-tuned for life; and we inhabit such fine-tuned universe.

Let us suppose that the ratio of their likelihood with respect to the fine-tuning evidence, *F*, is around 1: $P(F|G)/P(F|M) \approx 1$. Now, on this supposition, it is not clear whether there is a relatively unbiased or uncontentious way to evaluate the relative plausibilities of *G* and *M*, given evidence *F*.[21] Some philosophers (e.g. Hawthorne and Isaacs 2018, Sect. 7.7.3) suggest that the prior of *G* should be greater than *M*, because "… [it is] quite strange indeed to suppose that we are living in a multiverse" (2018, 160). Certainly, many reject this. For instance, one may argue that most non-theists should assign a far greater subjective probability to the multiverse hypotheses than to the God hypothesis. Because, for most non-theists, the universe with God in it is more "strange" than the universe without God.[22]

Therefore, at least at the first blush, there does not seem to be a non-subjective way of assessing the relative plausibilities of *G* and *M*. And cases like these are abundant in philosophy and science.

Now, it should be clear that the existence of such cases does not conflict with the main argument of this paper. Likelihoodists are not committed to the claim that comparative beliefs can be formed in *all* evidential situations. By contrast, all we needed to show is that, in *many* cases, comparative beliefs can be freed from subjective priors. And this is exactly what I have argued here: comparative belief can be objectively well-grounded even when the empirical information about non-relational priors are unavailable.

---

[21] If we use the imprecise probability framework, we may say that on some rationally permissible probability distributions, *G* is more likely than *M*, but for some other permissible distributions—*M* is more likely than *G*. Hence, on this framework, it seems that the evidence supports *suspending judgement* on whether *G* is more probable than *M*.

[22] I have developed this type of response to the fine-tuning argument in detail elsewhere (Tokhadze 2022).

Before concluding, let me briefly discuss the connection between the ratio form of Bayes' Theorem and MCP_{weak}. To remind the reader, MCP_{weak} is the thesis that:

> For any pair of hypotheses A and B, there is a real number r such that a rational agent believes A over B if her total evidence favours A over B to degree r or greater.

As I have already noted in the previous section, threshold r in MCP_{weak} may be fixed differently in different contexts of reasoning. But I have not elaborated on what this context is and how it fixes the relevant threshold. For our purposes, the most salient factor of an agent's reasoning context, with respect to competing hypotheses A and B, is her *comparative prior probability function* which provides an estimate of the value of the ratio of priors (for A and B). So, if an agent estimates that A and B's prior ratio is some number c, then threshold r should be greater than the *reciprocal* of c. As before, the rationale behind this is provided by the ratio form of Bayes' Theorem. So, the threshold r would be sensitive to an agent's estimate for the relevant ratio of priors.[23]

In some cases, like in the considered example from cognitive neuroscience, the relevant value of threshold r can be estimated in a relatively uncontentious manner. However, in the fine-tuning example, different agents may have different estimates for threshold r. And as I have already discussed, this is perfectly consistent with the main argument of this paper.

This concludes the positive argument of this paper.

5 Conclusion

From its inception, the main objective of the likelihoodist program has been to provide an objective (i.e., non-Bayesian) and logically coherent (i.e., non-frequentist) account of scientific evidence. Many have criticised the likelihoodist program on theoretical grounds. But some have levelled a more practical objection against likelihoodism; as these critics have argued, likelihoodist methods are too restrictive to guide belief.

This paper has called the above, received view into question. As I have argued, a broadly likelihoodist framework can accommodate belief guidance without appealing to Bayesian subjective prior probabilities. My main argumentative strategy has been the move from the non-relational to relational or comparative probabilities: probabilities of the form "A is more probable than B". This shift towards comparative probabilities, I believe, opens up a new space for addressing the various issues concerning belief guidance for likelihoodists and non-likelihoodists alike.

Declarations

Conflict of interest I declare that I have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

---

[23] My proposal follows Leitgeb's view (2017, Chapter 3), who takes an agent's degree of belief function P to be part of her reasoning context.

References

Bandyopadhyay, P.S., and M.R. Forster. 2011. Philosophy of statistics: An introduction. In *Philosophy of statistics*, pp. 1–50. Amsterdam: North-Holland.

Bandyopadhyay, P.S., G.G. Brittan, and M.L. Taper. 2017. Non-Bayesian accounts of evidence: Howson's counterexample countered. *International Studies in the Philosophy of Science* 30 (3): 291–298.

Bandyopadhyay, Prasanta S., Gordon G. Brittan, and Mark L. Taper. 2016. *Belief, evidence, and uncertainty: problems of epistemic inference*. Basel: Springer International Publishing.

Berger, J., & Wolpert, R. 1988. *The likelihood principle.* Beachwood: Institute of Mathematical Statistics, 2nd edition.

Bird, Alexander. 2021. Understanding the replication crisis as a base rate fallacy. *The British Journal for the Philosophy of Science* 72 (4): 965–993. https://doi.org/10.1093/bjps/axy051.

Dienes, Z. 2011. Bayesian versus orthodox statistics: Which side are you on? *Perspectives on Psychological Science* 6 (3): 274–290.

Earman, J. 1992. *Bayes or bust? A critical examination of bayesian confirmation theory*. Cambridge, MA: MIT Press.

Edwards, A.W. 1972. *Likelihood*. Cambridge: Cambridge University Press.

Edwards, W., H. Lindman, and L.J. Savage. 1963. Bayesian statistical inference for psychological research. *Psychological Review* 70 (3): 193–242.

Fitelson, B. 2007. Likelihoodism, bayesianism, and relational confirmation. *Synthese* 156: 473–489.

Fitelson, B. 2011. Favoring, likelihoodism, and bayesianism. *Philosophy and Phenomenological Research* 83 (3): 666–672.

Gandenberger, G. 2015. A new proof of the likelihood principle. *The British Journal for the Philosophy of Science* 66: 475–503.

Gandenberger, G. 2016. Why I am not a likelihoodist. *Philosophers' Imprint* 16 (7): 1–22.

Greene, J.D. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science* 293 (5537): 2105–2108.

Hawthorne, J. 2005. Degree-of-belief and degree-of-support: Why Bayesians need both notions. *Mind* 114 (454): 277–320.

Hawthorne, J., and Y. Isaacs. 2018. Fine-tuning fine-tuning. In *Knowledge, belief, and god: New insights in religious epistemology*, ed. M.A. Benton, J. Hawthorne, and D. Rabinowitz, 136–168. Oxford: Oxford University Press.

Howson, C. 2013. Exhuming the no-miracles argument. *Analysis* 73 (2): 205–211.

Howson, C. 2017. How pseudo-hypotheses defeat a non-Bayesian theory of evidence: Reply to Bandyopadhyay, Taper, and Brittan. *International Studies in the Philosophy of Science* 30 (3): 299–306.

Kruschke, J.K. 2013. Bayesian estimation supersedes the t test. *Journal of Experimental Psychology: General* 142 (2): 573–603.

Leitgeb, H. 2017. *The stability of belief: How rational belief coheres with probability*. Oxford: Oxford University Press.

Leslie, J. 1989. *Universes, London*. London: Routledge.

Machery, E. 2014. In defense of reverse inference. *The British Journal for the Philosophy of Science* 65 (2): 251–267.

Mayo, D. 1996. *Error and the growth of experimental knowledge*. Chicago: University of Chicago Press.

Mayo, D. 2018. *Statistical inference as severe testing: How to get beyond the statistics wars*. Cambridge: Cambridge University Press.

Meacham, C. 2014. Impermissive Bayesianism. *Erkenntnis* 79: 1185–1217.

Romero, F. 2019. Philosophy of science and the replicability crisis. *Philosophy Compass*. https://doi.org/10.1111/phc3.12633.

Royall, R. 1997. *Scientific evidence: A likelihood paradigm*. London: Chapman and Hall.

Salmon, W. 1990. Rationality and objectivity in science or Tom Kuhn meets Tom Bayes. *Minnesota Studies in the Philosophy of Science* 14: 175–204.

Sober, Elliott. 2008. *Evidence and evolution: The logic behind the science*. Cambridge: Cambridge University Press.

Tokhadze, Tamaz. 2022. Fine-tuning, weird sorts of atheism and evidential favouring. *Analytic Philosophy* 63 (3): 192–203. https://doi.org/10.1111/phib.12232.

Wetzels, R., D. Matzke, M.D. Lee, J.N. Rouder, G.J. Iverson, and E.J. Wagenmakers. 2011. Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science* 6 (3): 291–298.

Williamson, J. 2010. *In defence of objective Bayesianism*. Oxford: Oxford University Press.

Williamson, Jon. 2007. Motivating objective Bayesianism: from empirical constraints to objective probabilities. In W. E. Harper, & G. R. Wheeler (Eds.), Probability and Inference: Essays in Honour of Henry E. Kyburg, Jr. Amsterdam: Elsevier.