Can Miracle Scepticism be Unscientific?[1]

Dr G. Menzies
University of Technology, Sydney
Dr H. E. Doran,
University of Queensland

Abstract
Bayes Rule can be used to discuss the strength of belief concerning miracles. The rule demonstrates that religious 'fundamentalists' who do not doubt miracle claims at all are not impacted by evidence. However, our first result is that non-religious sceptics who accord miracles zero prior probability also are not impacted by evidence, which is unscientific. Our second result is that extremely small yet non-zero priors for miracles are equivalent to very small probability thresholds in classical hypothesis testing (what are called test sizes or levels of significance), far smaller than the conventional 1%, 5% or 10%. Miracle sceptics who use very low priors thus depart from the scientific community's common practice, and are open to the charge of being unscientific is this second sense.

Keywords: *Miracles, Bayes Rule*

---

1. Introduction

Hume's essay 'Of Miracles' (1748) concerns itself with the truth of miracle claims, and famously rejects them. His analysis motivated the development of Bayes Rule, which has subsequently been used to discuss miracle claims. Bayes rule is a natural formula for these questions since it makes allowance for a starting strength of belief, and then modifies the strength of belief according to the evidence presented. Although it arose historically in the context of miracle debates, its validity in no way depends on whether one accepts Hume's argument. Following writers such as Dawid and Gillies (1989), Bayes rule will be our preferred framework for our analysis of miracles.

In this paper we will show that under Bayes rule extreme priors are replicated irrespective of evidence. That is to say, someone who believes a miracle claim without a doubt – a prior probability of unity – will end up with the same posterior probability, namely unity. This point is well understood with respect to religious 'fundamentalists' where that term is understood as a complete disregard to evidence (Dawkins, 2007).

However, it turns out that insensitivity to evidence cuts both ways. Someone who disbelieves a miracle claim without a doubt – a prior probability of zero – will end up with the same posterior probability, namely zero, and this is our first result.

[1] We thank Peter Docherty and Jack Gray for useful discussions.

Of course, not many miracle sceptics will claim or admit that evidence is completely unimportant to them, since that would be tantamount to a declaration of contempt for science. We therefore consider what happens when prior probabilities for miracles are extremely low, yet above zero.

Our second result is that extremely low priors for miracles are equivalent to extremely low probability thresholds in classical hypothesis testing (what are called test sizes or levels of significance), far smaller than the conventional 1%, 5% or 10%. Miracle sceptics who use very low priors thus depart from the scientific community's common practice, and are open to the charge of being unscientific is this second sense.

In stating clearly what our paper achieves, we need to be clear about what our paper does not achieve.

First, we make no claim to be Hume scholars. Our logic is built upon Bayes rule rather than Hume's argument. Having said that, we will briefly mention one critique of his famous essay 'Of Miracles' by Hajek (2008) since Hume's argument could, if valid, provide a justification for test sizes far lower than the scientific community's common practice.

Second, we do not substitute any actual probabilities into Bayes rule in our analysis, apart from the extreme cases of zero and unity. Nor do we arrive at our best guess for the probability of any specific miracle. Our argument is stronger for being connected to the underlying mathematicological structure, rather than guessed values for probabilities. We venture later in the paper that sceptics can be inconsistent in their scepticism, not that they get a particular number wrong. Although the former error can lead to the latter, the nub of our argument does not reside there.

2. Bayes Rule

2.1 A Statement of Bayes Rule

If there is an event A for which evidence (or 'testimony') E is offered, Bayes rule provides a means of taking an initial (or 'prior') probability of A, and updating it with the evidence, to arrive at a new probability that takes the evidence into account.[2] We use standard terminology where a vertical line | means 'given', P means 'probability' or 'strength of belief' and a bar over an event refers to the complement of an event. Thus P(A) is the probability of event A, P(A|E) is the probability of A given that evidence E is observed, and P($\bar{A}$) is the probability that A does not occur. Using this notation, the usual formulation of Bayes Rule is

$$P(A|E) = P(E|A)\, P(A) \,/\, [P(E|A)\, P(A) + P(E|\bar{A})\, P(\bar{A})] . \qquad (1)$$

[2] If 'testimony' is interpreted broadly, most of our beliefs do indeed travel through the conduit of the testimony of others: in the press, in academic journals or from the observations of friends, colleagues and family.

We note here an immediate implication of (1). If P(A) = 0, then clearly P(A|E) is also zero. Also, if P(A) = 1, then P($\overline{A}$) = 0 and (1) shows that P(A|E) = 1. Thus, if P(A) is either 0 or 1, it follows that

$$P(A|E) = P(A), \qquad (2)$$

implying that evidence E will never change the original P(A), or alternatively, the two 'extreme' values of P(A) render evidence irrelevant. We will return to this point later.

There are many examples of the usefulness of (1). To give just one, suppose a woman tests positive to a breast cancer test. Here we define A to be the event she has cancer and E is a positive result. The woman will be intensely interested in an answer to the question "Given I have a positive mammogram result what is the probability that I actually have cancer?" Using our notation, the question becomes "What is the value of P(A|E)?" We can use Bayes rule to update the probability she actually has cancer.[3] Suppose the probability from the general female population is P(A)=.0006, (implying P($\overline{A}$) = 0.9994) and that the probability of testing positive if she really has cancer is P(E|A)=0.84. Finally, suppose the probability of testing positive when she doesn't have cancer, a 'false positive' is P(E| $\overline{A}$) = 0.03. Substituting these numbers into (1), we obtain P(A|E) = 0.016. Thus the positive test raises the probability of cancer from 6 in 10,000 to just under 2 in 100. Bayes rule has registered an important increase in the probability, even though it is still rather unlikely she has cancer.

In this example, all the probabilities used are presumably measured using realized frequencies and experimental data, but Bayes rule is not always used that way. It is common to let the starting or prior probability P(A) in (1) be a subjective 'strength of belief' as we flagged earlier, and this is the most natural sense when discussing miracles or other unrepeatable historical events.

2.2 Representing P(A|E) graphically

If we return to equation (1) and divide numerator and denominator by P(E|A)P(A), we obtain Bayes Rule in the form

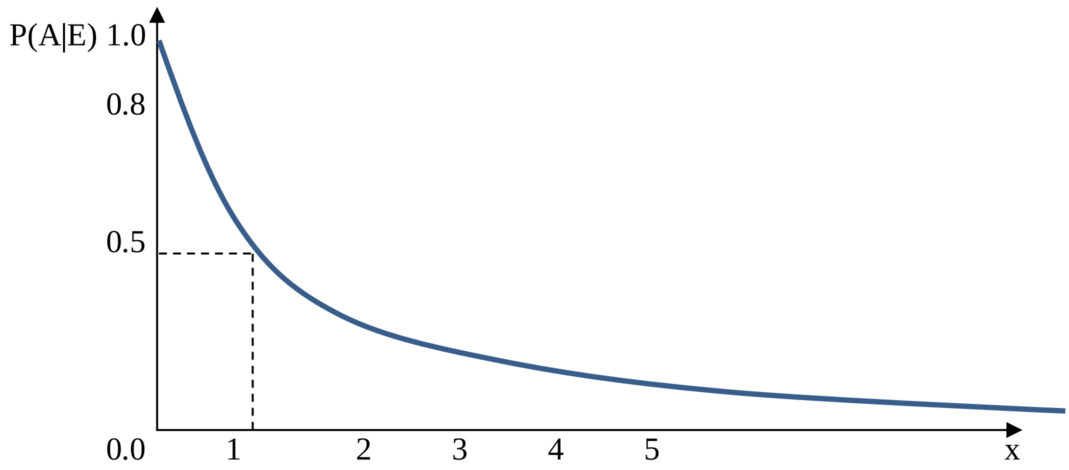
$$P(A|E) = 1 / (1 + x) \qquad (3)$$

where

$$x = P(E|\overline{A})\, P(\overline{A}) \,/\, P(E|A)\, P(A) \qquad (4)$$

[3] See http://www.ncbi.nlm.nih.gov/pubmed/11002452 and related links for probabilities like the ones in the text.

Figure 1: Updated Probability



The graph of the function 1/ (1 + x) shown above enables us to see the main features of Bayes Rule, as follows:

(i) If x = 0, P(A|E) = 1
(ii) Small x (relative to 1) implies high P(A|E)
(iii) x<1 implies P(A|E)>0.5
(iv) Large values of x result in small values of P(A|E)

In addition, from (3), if x is much larger than 1, then P(A|E) is, for all practical purposes equal to 1/x. For example, if x = 1000, then P(A|E), calculated from (3) is almost identical to 0.001.

Finally, for convenience in what follows, we introduce the following notation:

$P(A) = \alpha$

and

$P(E|\bar{A}) = p$.

The quantity x is now given as

$$x = p(1 - \alpha)/(\alpha P(E|A)) \quad (5)$$

and from now on, we will use Bayes Rule in the form of equations (3) and (5).

2.3 Of Miracles

We will now turn our attention to a special kind of event which describes a miracle claim. We have two relevant criteria.

An event A is *evidential* if $P(E|A) \approx 1$; viz. if A is true the evidence advanced for it is to be expected. Despite appearances of no common ground among protagonists in miracle debates, there is often agreement by everyone that this probability is in fact close to one (McGrew, 2013).

The event A is *extraordinary* if $P(A)=\alpha\approx 0$; viz. A is not to be expected in the ordinary course of events. Again, everyone agrees that miracles are not common. If these two conditions are fulfilled we will, for the purposes of this paper, describe A as a miracle M.

It follows that, when we are using Bayes Rule to investigate evidential and extraordinary events, aka miracles, the quantity x defined in (5) simplifies to a ratio p/α in (6). This ratio is well-known as the origin of Hume's so called balancing principle.[4]

$$x = p(1-\alpha)/(\alpha P(E|A)) \approx p(1-\alpha)/\alpha \quad \text{for evidential events}$$

$$= p/[\alpha/(1-\alpha)] \approx p/(\alpha+\alpha^2) \quad \text{for extraordinary events}$$

$$\approx p/\alpha \qquad (6)$$

Furthermore, in addition to these two conditions, the event we are now calling M may also be *uniquely plausible* which is to say $P(E|\overline{M})=p\approx 0$. This condition says that if M is not true the evidence E is very hard to account for. Much of the debate about miracles really revolves around this last condition and whether or not it is fulfilled. That is, we seek to answer the question "Is there a plausible, non-miraculous alternative explanation for the evidence at hand?" In order to understand why so much hinges on this last condition, we will consider two uninteresting miracle claims which are evidential, extraordinary, but not uniquely plausible.

A religious example of an uninteresting miracle claim would be the suggestion that all geological features around us were created miraculously by a deity 10,000 years ago with the appearance of age. If this miracle occurred, then it explains the evidence, including the evidence that the earth appears to be older. Furthermore, on the face of it this seems to be an extraordinary event. However, if the miracle did not occur, then there are credible theories commanding near-unanimous support of the relevant scholars which explain what we see, so most people would not entertain this notion when there are perfectly good alternative explanations; i.e. $P(E|\overline{A})=p>>0$.

[4] '.. no testimony is sufficient to establish a miracle, unless the testimony be of such a kind, that its falsehood would be more miraculous, than the fact, which it endeavours to establish ..' Hume (1748, last paragraph in Section I). The 'fact' under question is the miracle so the associated probability is α. Hume has in mind the evidence as testimony so false testimony is the event $E|\overline{A}$ with associated probability p. So in terms or our notation Hume is saying a miracle might be believed if $p<\alpha$ or equivalently if $x<1$. We will come to this point presently.

A non-religious example of an uninteresting miracle claim would be a stage magician who claimed paranormal powers which, if actually present, would explain various phenomena during his performance. Presumably these phenomena are extraordinary enough for people to pay to see him, and not to be expected in the ordinary course of things. However, familiarity with the techniques of his art suggests that the miraculous claims could easily be explained in the absence of special powers, so that again $P(E|\overline{A})=p>>0$. The ease of alternative explanations means that there are few people who seriously believe in these paranormal powers.

A potential candidate for an interesting miracle claim is the alleged bodily return to life of Jesus after his death (the resurrection).[5] Historians who do not accept this miracle must account for narratives which describe: an empty tomb; Jesus' post-resurrection appearances; and, the psychological transformation of the disciples which allowed them to promote their new faith against strong opposition, all of which could reasonably be expected if the miracle had occurred. Furthermore, the search for a naturalistic explanation for the narratives has been marked by the proposing of various mutually exclusive theories, none of which commands broad support from the relevant scholars (unlike the first example of the geological miracle). Some have used this lack of consensus to reject the viability of a naturalistic hypothesis (see the debate between Habermas and Flew, 2003, and a historical case mounted by Bruce, 1981).[6]

In the two uninteresting cases above p is large relative to α, and so (from (6)) x will be large relative to unity and $P(M|E)\approx 1/x$ will be correspondingly low. However, if one accepts, or at least entertains, that there could be uniquely plausible miracles then that would imply (6) could produce a small value of x, and hence a high probability P(M|E).

The ratio x in (6) has a crucial value at unity. If a naturalistic explanation $E|\overline{A}$ is more improbable than the prior belief in the miracle itself, x is less than unity and (3) will be greater than one half. Thus, in Figure 1, x=1 is a plausible Bayesian cutoff between believing and not believing that M happened, leading to a rule: believe M if $x<1$. We leave it to Hume scholars to discern whether Hume himself could have seriously entertained that possibility, but the logic of

[5] The analysis below is not tied to a particular event or miracle claim, so we ask that a reader sceptical of the resurrection forebear with us until the structure of the argument is laid out. Its general nature will then be apparent.
[6] There has also been a dramatic though gradual change in the scholarly dating of the relevant documents. Historians in the late 1800s dated key New Testament documents in the second century A.D., around one hundred years after the life of Jesus. The situation is different now. Through the twentieth century, a very wide consensus emerged that most of the documents belong to the first century. This is not to say that late datings are now unheard of - rather, it is just to say that they are just far less frequent among the experts. This tendency to date the documents earlier has not been confined to scholars who are conservative Christians. John Robinson, famous for his sceptical book entitled 'Honest to God' (Robinson, 1963), wrote a new book entitled 'Redating the New Testament' (Robinson, 1976) where he argued that all the New Testament documents were written prior to 70 A.D.

believing in something that is better than even odds is straightforward enough, and is used in Bayesian hypothesis testing (Berger, 1980).[7] We will return to a discussion of x subsequently.

2.4 What Really Matters in Miracles Debates

Equation (3) and the approximation of x for miracles in (6) makes the calculations in Dawid and Gillies (1989) very simple indeed, and leads to insights about their methodology. They discuss the report of a particular person winning a lottery where the event has a near-zero prior probability ($\alpha=10^{-6}$). Their probability that the event is misreported is $P(E|\overline{M})=p\approx 2\times 10^{-7}$ and the probability of a correct reporting is close to unity with P(E|M)=0.8. With all our requirements for an interesting miracle satisfied we can use the approximation for x in (6). In this case x is 0.2, leading to P(M|E) ≈0.8.

When it comes to what they call 'the miracles case', they allow the prior to be the same for comparison purposes. They then claim that by making P(E|M)=0.99 they are making 'the case as favourable as possible to the establishment of a miracle' (pg. 62) but this is not what our algebra shows. Their event is evidential, extraordinary and uniquely plausible, so all that matters is p/α. Naturally, a miracle believer should be grateful for any gifts, but this one is a token gesture.[8] Whether P(E|M) is 0.9, 0.99, or even 1.0 is completely irrelevant.

They draw on Hume's analysis of the pitfalls of testimony (fakery, deception, awe at the marvellous and hallucinations) and declare a floor on p of 10^{-3} with an 'it would seem unreasonable' to go lower, as their summary argument. Thus their value of $x=p/\alpha$ is $10^{-3}/10^{-6}=1000$ and the miracle probability (approx. 1/1000) is essentially zero.

Hume was right to emphasise the importance of x in his 'balancing principle' shown in equation (6), rather than P(E|M). In contrast, Dawid and Gillies, by their arbitrary choice of p as 10^{-3} have multiplied x by 5000, from 0.2 in the lottery to 1000 in the miracle, as they move across examples. As Hume foresaw, it is only this value of x that matters in miracle debates, and it is their choice of x, rather than any supposed concessions about P(E|M), that ensures miracles remain implausible.

2.5 Extreme Religious and Nonreligious Priors Make Evidence Irrelevant

We now return to the point about extreme priors made earlier in regard to equation (2), and the conditions under which P(A|E) = P(A). The discussion there applied whether or not A satisfies the conditions for a miracle M. However, because we are interested in the case of

---

[7] As it happens, Hajek (2008) does not think this is a plausible reading of Hume. Matters are of course more complicated if one is allowed to be agnostic, but as an intellectual disciplining device we are not considering the merits or the meaningfulness of that position in this paper.

[8] They do not appear to have understood the unimportance of their concession since, at the top of page 63, they make something of it again: 'Note that this is much higher than in our analysis of the lottery case'.

miracles, we will use the symbol M. We have seen that when $\alpha = 1$ or 0, evidence becomes irrelevant.

The first case, $\alpha = 1$ characterises the 'fundamentalist' – no matter what evidence is produced against M, their view that M occurred is never changed. If M is a 'religious' miracle (eg, Jesus rising from the dead), they are derided as being wilfully ignorant, behind-the-times and anti-scientific.

> "Passion for passion, an evangelical Christian and I may be evenly matched. But we are not equally fundamentalist. The true scientist, however passionately he may "believe", in evolution for example, knows exactly what would change his mind: evidence! The fundamentalist knows that nothing will." Dawkins (2007)

It is unsettling that Prof Dawkins entertains moving from one scientific theory, namely evolution, to another one more consonant with yet-to-be-discovered evidence. It would have been more reassuring that his α about theism is not actually zero if he had conducted a thought experiment about moving between atheism and theism, perhaps by saying something like this:

"The true scientist, however passionately he may "believe" in atheism for example, knows exactly what would change his mind: miracle evidence!"

We are agnostic about whether Prof Dawkins would be prepared to say this, but surely some atheists would not. Anyone who believes in a closed universe which automatically rules M out would be untouched (indeed untouchable) by any evidence for M. But this comes at a cost. Any sceptic who claims scientific education and insight to justify *ruling M out as impossible* is implicitly undercutting their claim. By setting $\alpha=0$ they are violating a principle of science that all theories can be effectively interrogated by evidence.[9]

In what follows, we therefore confine our attention to scientific atheists who would admit a role for evidence in miracle debates. We take it that they would have non-zero, albeit very low, values of the prior probability, either because they think chaotic and miraculous-looking events are possible in a tiny fraction of instances, or because they harbour some small degree of doubt about their atheism and closed-universe hypotheses.

[9] The dream of straightforward falsificationism has foundered on significant philosophical difficulties. As a result the so called demarcation problem has been identified in philosophy whereby it is not a straightforward matter to draw a clean line between science and pseudo-science (Laudan, 1983). Nonetheless, the principle that evidence is important and should be used in whatever validation of theories is possible remains a key element of the scientific approach.

3 Hypothesis Testing

3.1 The Bayesian Decision Rule Can Be Expressed in Terms of x, p and α

Although many atheists, and perhaps Hume himself, might not have been happy with the idea of believing in a miracle that has better than even odds, we noted above that this is consistent with so called Bayesian hypothesis testing (Berger, 1980). With due respect to some readings of Hume, try saying aloud 'all the best evidence considered, I think M is more likely than not, but I'm still going to believe $\overline{M}$ ' and see how it sounds.

Thus the Bayesian decision rule expressed in terms of P(M|E) requires us to believe in the miracle if P(M|E) >0.5. According to Figure 1 (applied to miracles so A becomes M) this amounts to believing in a miracle if and only if $x<1$.

We now make the observation that the condition $x<1$ for believing in a miracle (and disbelieving it if $x\geq1$) can, because of our mathematical description of a miracle, call upon the approximation in equation (6) to deliver the following simple Bayesian decision rule:

Believe in M if and only if $p<\alpha$ (7)

It will be recalled that $p=P(E|\overline{M})$, and this is the probability of obtaining the actual evidence we see, if the miracle did *not* occur. According to (7), this is to be compared to a prior probability of M, which is a subjective number chosen by a researcher.

3.2 This Rule Looks Like a Hypothesis Test

Anyone trained in classical statistics will be struck that this Bayesian decision rule in (7) compares the probability of obtaining evidence under a given belief with a small cut-off chosen by a researcher. This is close in spirit to a hypothesis test, where a researcher asks if the evidence seen is a 'rare event' under the 'null hypothesis', called H_0, and rejects the null if it is.[10]

For example, imagine Person A shows Person B an urn with 1,000,000 hidden balls and then announces that 999,999 balls are white and 1 ball is black. Person B could be forgiven for doubting the announcement if Person B is given the opportunity to draw a single ball, only to find that it is black.

Person B's doubts can be formalized using a hypothesis test. The so-called null hypothesis H_0 says that the urn contains 999,999 white balls and 1 black ball. The probability of obtaining the

[10] In the language of hypothesis testing a hypothesis can be 'accepted' or 'rejected' which is close enough to 'believe' and 'disbelieve' for us to use the latter terms. A technicality here is that it is not common to say that a null hypothesis is accepted, whereas it is conventional to say an alternative is accepted (should the evidence point that way). The point being made by the different terminology has to do with the different probabilities of making a mistake, but need not detain us.

evidence of one black ball if H_0 is true is called the p-value: $p=P(\{1 \text{ black draw}\}|H_0)=.000001$. This is so unlikely, i.e. 'rare', as to cast overwhelming doubt on H_0.

More generally, the basic procedure of a hypothesis test is to reject H_0 if p falls below a small cut-off value α. We reject H_0 if $p < \alpha$, where in conventional scientific research α takes the values of either 0.01, 0.05, or 0.10. It is reasonable to accept that improbable events happen to a point (events as unlikely as, say 0.01), but once the probability p of observing the evidence we see gets smaller than that, it is reasonable to suspect that hypotheses other than H_0 are responsible for what we observe. The cut-off value α is also called the significance level, or the test size, and the most common choice is $\alpha=0.05$ because it has provided scientists with good overall guidance in the past.

> "The .05 significance cutoff has been used literally millions of times since Fisher proposed areas of science. I don't think that .05 could stand up to such intense use if it wasn't producing basically correct scientific inferences most of the time." Efron (2005 pg. 5)

It is important to note that the smaller α is, the more we are protecting H_0 from being rejected by evidence. Therefore α can be too small – if it is zero one will never reject the null hypothesis no matter how compelling the evidence because $p<\alpha$ can never be true. This mirrors the case discussed in section 2.5 when a Bayesian prior of zero was likewise so small that evidence was not allowed to speak.

What are we to make of the strong similarities between the Bayesian approach and a hypothesis test, both in general tenor and in the irrelevance of evidence if the prior or the test size are zero?

3.3 The Bayesian Decision Rule can be Turned into a Hypothesis Test

We now make an observation which, to our knowledge, is new. We can, by a judicious choice of a null hypothesis, make our Bayesian decision rule "Believe in the miracle if $p<\alpha$" entirely equivalent to a hypothesis test, as long as we are prepared to reinterpret α.

We first note that our Bayesian decision rule bids us to believe in the miracle if $p<\alpha$ and we then remind ourselves that $p=P(E|\overline{M})$. It follows that this p is definitionally the same as the p-value in a hypothesis test if we choose H_0 to be the belief that the miracle did NOT occur, namely belief in $\overline{M}$.[11] In other words the following hypothesis test will give exactly the same outcomes, in terms of whether H_0 is believed or not, as the Bayesian decision rule in (7):

H_0: $\overline{M}$ the miracle did not happen
H_1: M the miracle did happen

[11] If the null is rejected in a hypothesis test, the alternative belief H_1 is believed instead. It is easy to stipulate H_1 in this case – it is simply that the miracle did happen.

Decision Rule: reject H_0 $p<\alpha$.[12] (8)

Thus, looking across the two methods for evaluating the plausibility of miracles (events that are both evidential $P(E|M) \approx 1$ and extraordinary $P(M)=\alpha\approx 0$) we have for the first time demonstrated the complete equivalence of two widely used frameworks:

1. Reject $\overline{M}$ in favour of M if $p<\alpha$ is a Bayesian hypothesis test where $p=P(E|\overline{M})$ and α is the prior probability on M, and
2. Reject $\overline{M}$ in favour of M if $p<\alpha$ is a classical hypothesis test where $p=P(E|\overline{M})$ and α is the level of significance.

The pronumeral p is the same across the two frameworks, but what has not been understood in the past is that the significance level for a hypothesis test and the Bayesian prior, which are generally chosen according to different criteria by researchers, are actually identical for the miracle case.[13]

3.4 Should we Use Hypothesis-Testing Conventions?

The importance of seeing the equivalence between Bayesian and classical hypothesis testing for miracles is that we now have available to us the scientific conventions for the latter in our execution of the former. This raises an intriguing possibility. Perhaps it is possible to break the deadlock that has characterized this area of philosophy of religion, or at least provide a scientific basis for the choice of a prior which then can be a benchmark - a starting value to which the proposed prior probabilities can be compared. Thus a scientist who spends their days at the lab conducting experiments with a 5 per cent test size might, when reading about miracles in the evening, be justified in starting with 5 per cent in that realm of inference too.

We are aware that there are circumstances where it is defensible to put a low prior probability on an event and therefore, by the above result, use a low test size (consequently protecting the null from rejection). We actually think that Dawid and Gillies (1989) are correct in choosing the 10^{-6} prior for a lottery application (with 6 numbers), and there are situations where scientists have adopted lower test sizes than 1 per cent in particle physics (Franklin, 2013).

But there is good sense in the conventional 0.05 too. The modelling discipline associated with classical test sizes goes both ways with regards to evidence – it restrains the credulous, but it also challenges the sceptical. We often teach undergraduate students why test sizes are so small, *but*

---

[12] We are aware that the equivalence between Bayesian and classical hypothesis testing in the case of miracles is subject to a second-order small error, to the extent that $(1-\alpha)/[\alpha\ P(E|M)]$ differs from $1/\alpha$ in (5). But for small αs, such as 0.1, 0.05 and 0.01, and for $P(E|M) = 1-\delta$ close to unity, we approximate $(1-\alpha)/[\alpha\ P(E|M)]$ as $1/[\alpha\ (1+\alpha)(1-\delta)] \approx 1/[\alpha+\alpha^2-\alpha\delta]$ and the denominator approaches α very rapidly. For example, if $\alpha=0.05$ and $\delta=0.1$ we have $\alpha+\alpha^2 - \alpha\delta = 0.0475$.

[13] The test size is formally the probability of a false rejection of a null hypothesis, and a Bayesian prior has a range of justifications that permit more latitude (see Berger, 1980). It is a surprise that they are identical.

we could gainfully teach them why they are as large as they are. A balance must be struck at a value that makes one reluctant to change one's mind (α not too large), but not so reluctant that evidence can be ignored (α not too small). That is why conventions about these values have settled at 'a few per cent', that is 1, 5 or 10 per cent.

The attraction for using standard test sizes for miracles as a benchmark is thus twofold. There is the argument just made that standard test sizes strike a plausible balance necessary for any use of evidence. Second, and more tellingly, it is in any case impossible to assign objective prior probabilities to miracles in a way that will satisfy everyone, since any assignment will involve a whole swath of worldview assumptions.

3.5 Can Miracle Scepticism be Unscientific?

Hume and other early writers in science (such as Descartes (1984)) are famous for doubting past certitudes, and helping to launch the scientific enterprise. Yet this paper has argued that there is a potential conflict between Humean scepticism and the scientific valuing of evidence. This is obviously true when the prior (and therefore the test size) for a miracle is zero (as we showed in section 2), but in this last section we have shown that vanishingly low miracle priors are equivalent to vanishingly low test sizes in hypothesis tests, so this possibility has re-appeared in another guise. Arguably, the onus is on sceptics to justify what makes miracle different when there is a 0.05 scientific standard widely adopted across many different disciplines. These applications represent a vast array of environments and of actual priors that any particular scientist brings to their experiment.

This puts the spotlight firmly back onto the philosophical arguments made, by Hume and others since, that there are good reasons for having extremely small priors.[14] It should now be clear from our analysis that one must either find a 'special case' argument against miracles which justifies the small test sizes, such as Hume attempted to do, or be open to a charge of inconsistency in the use of evidence.

As we said in the introduction, we make no claims to be Hume scholars, but given the stakes of finding a special case argument against miracles we cannot avoid saying something about him here.

There is a controversy about whether Hume himself believed miracles are unbelievable in principle, or in practice. The former claim asserts that $\alpha=0$ and implies that the examination of evidence is irrelevant.[15] The latter is a claim that what we have called p is really quite high, due perhaps to fakery, deception, awe at the marvellous or hallucinations. However, such a claim poses no particular problem to our call for sceptics to return to mainstream scientific practice.

[14] The final chapter of Dawkins (2011) echoes Hume in this regard.
[15] Since Hume (1748) spends a good deal of time looking at evidence, Hajek (2008) questions that this is a correct reading of Hume, though he acknowledges the contrary view advanced by Antony Flew.

Presumably the values of p are high enough to not go below a standard test threshold of 1, 5 or 10 per cent and miracles would be rejected without the need for implausibly small test sizes. Whatever Hume meant – zero α or high p – the end result that $p>\alpha$ was, to his way of thinking, inescapable.

Concerning the in principle argument that α is zero or vanishingly close to it, Hajek (2008) discusses some readings of Hume, and the reasons he gave for miracles to be discarded on philosophical grounds. In the end, Hajek locates the crux of an in principle argument on Hume's core idea about probability – that we assign probabilities to things based on analogies with past events. Miracles are so extremely unlikely on this basis, says this reading of Hume, that they are not be entertained. Yet in the end Hajek rejects Hume's account of probability. The death knell, he suggests, lies in some of the discoveries of modern science:

> "… if strength-of-analogy is such a crucial determinant of a reasonable person's probability function, then that person should also be a skeptic [sic] about all spectacular scientific discoveries. And that is absurd." (page 27)

If one accepts that Hume's argument against miracles is ultimately unconvincing, let alone absurd, there is clearly a danger that low priors and hence low test sizes can function as an immunizing strategy for a sceptic. This, it will be recalled, is an argument or procedure employed in support of a belief system which makes it more or less invulnerable to empirical evidence.[16] As we have noted before, the smaller α is, the more we are protecting H_0 from being rejected by evidence, and herein lies the potential for the misuse of α.

In our view a miracle sceptic should advance a solid philosophical reason why miracles are special, one that carries more weight than Hume's argument, or adopt the discipline of a scientific test size of 'a few per cent', Otherwise, an unthinking sceptic can have a free hand to give lip service to the scientific ideal of letting evidence speak, without any real willingness to do so. Any such departure from scientific discipline is just as deserving of the 'unscientific' epitaph as sceptics who more straightforwardly announce a zero prior probability of a miracle.

References

Berger, J. O. (1980), *Statistical Decision Theory and Bayesian Analysis*, Springer, New York.

Boundry, M. (2011), 'Immunizing Strategies and Epistemic Defense Mechanisms', Philosophia, March, 39 (1), 145-161.

[16] Boudry (2011) gives a number of examples of making theories invulnerable to evidence. For horoscopes, predictions are framed in such a way that anything that happens can be interpreted as being consistent with the forecast. For Jehovah's Witnesses, repeated failures to forecast the second coming of Christ have lent themselves to increasingly non-physical interpretations of what 'coming' means. For psychoanalysis or classic Marxism, questioning the validity of theory can be dismissed as motivated by psychologically induced 'resistance' in the former case and 'bourgeois class consciousness' in the latter.

Bruce, F. F. (1981), *The New Testament Documents: Are They Reliable?* Downers Grove, IL: InterVarsity Press.

Dawid, P. and D. Gillies (1989), 'A Bayesian Analysis of Hume's Argument Concerning Miracles', *The Philosophical Quarterly,* 39(154), 57-65.

Dawkins, R. (2007), Letter to *the Times*, 'How dare you call me a fundamentalist', May 12.

Dawkins, R. (2011), *The Magic of Reality: How we Know What is Really True*, Bantam, UK.

Descartes, R. (1984), *The Philosophical Writings of Descartes*, trans. John Cottingham, Robert Stoothoff, Dugald Murdoch and Anthony Kenny, Cambridge: Cambridge University Press, 3 vols., 1984-1991.

Efron, Bradley (2005), 'Modern Science and the Bayesian-frequentist Controversy', *Technical Report*, Stanford University, Dept of Statistics.

Franklin, Allan (2013), 'Prologue: The rise of the sigmas' in *Shifting Standards: Experiments in Particle Physics in the Twentieth Century,* (1st ed.), Pittsburgh, University of Pittsburgh Press, Ii–Iii.

Habermas, G. R. and Flew, A. G. N. (2003), *Did Jesus Rise from the Dead? The Resurrection Debate*, ed. Terry L. Miethe, Wipf and Stock.

Hajek, A. (2008), 'Are Miracles Chimerical', *Oxford Studies in Philosophy and Religion*, Volume 1, OUP 82-104

Hume D. (1748),'Of Miracles', *An Enquiry concerning Human Understanding* X, i, 86. As revealed in Hume's private correspondence, various versions of this work were originally circulated without this section to avoid offending religious people.

Laudan, L. (1983), 'The Demise of the Demarcation Problem', *Physics, Philosophy and Psychoanalysis: Essays in Honour of Adolf Grunbaum*, R. S. Choen and L. Laudan. Dordrecht, D. Reidel: 111-128.

Larson, H. J. (1982), *Introduction to Probability Theory and Statistical Inference*, 3rd ed. Wiley, New York.

McGrew, T. (2013), 'Miracles', *The Stanford Encyclopedia of Philosophy* (Spring 2013 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2013/entries/miracles/>.

Robinson, J. (1963), *Honest to God,* SCM Press, Westminster Press Philadelphia.

Robinson, J. (1976), *Redating the New Testament*, Wipf and Stock, Eugene Oregon.