Can Miracle Scepticism be Unscientific?[1]

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

Abstract
It is well-known that Hume's argument against miracles (1748) anticipates Bayes Rule. The rule also demonstrates that religious 'fundamentalists' (who do not doubt miracle claims at all) and non-religious sceptics (who accord them zero prior probability) both regard evidence as irrelevant. We show that Hume's argument also anticipates a hypothesis test where the status quo belief (the null hypothesis) is that a miracle has not occurred. Proponents of extreme miracle scepticism thus advocate vanishingly small hypothesis test sizes, in contrast to the scientific community's common range of 1 per cent to 10 per cent.

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 anticipates Bayes Rule (Dawid and Gillies, 1989) in the sense that what is sometimes referred to as his 'balancing principle' can be derived from the rule. Hume's principle says, in effect, that when one is faced with a miracle claim the key judgment to make is whether it is less likely that the testimony is made up, or, that the miracle occurred.

'… 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)

In this paper we will reiterate the connection between Bayes rule and these two probabilities, and then go on to show that under Bayes rule extreme priors are replicated irrespective of evidence. This point is well understood with respect to religious 'fundamentalists' where that term is understood as a cavalier attitude to evidence but we will show it can be equally relevant for anti-religious sceptics.

We also show that Hume's balancing principle anticipates a hypothesis test where the null is that a miracle has not occurred, and the Bayesian prior is equivalent to a test size. To our knowledge this is a new result.

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

Both a Bayesian prior and a test size are subjective numbers, and it is helpful to be aware of their equivalence in this context. Mainstream science has opted for the convention of defaulting to 1 per cent, 10 per cent, or, most commonly, 5 per cent for a hypothesis test (Larson, 1982). The 5 per cent test size creates discipline in the use of evidence both ways, being neither excessively credulous nor excessively sceptical about a status quo belief. It has also earned respect within the scientific community as a fruitful approach in many inference situations.

There have been attempts to justify extremely low Bayesian priors in the case of miracles, in a debate that stretches all the way back to Hume's original essay. We do not discuss these attempts in detail, but Hajek (2008) provides an appraisal of them and rejects what he considers to be Hume's inadequate account of probability. The point of this paper is simply that if these or any other attempts to justify extremely low priors are not compelling, then those who adopt them for miracles – who are, as we have shown, actually adopting test sizes far lower than the standard 5 per cent, should recognize their methodology as a departure from a disciplined use of evidence.

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, P(A), and updating it with the evidence, to arrive at a new probability P(A|E) that takes the evidence into account.[2] The usual formulation of Bayes Rule is

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

where $\bar{A}$ denotes the complement of A, that is, the event of A not occurring and P(E|A) is the probability of E occurring, given that if A does occur. $P(E|\bar{A})$ is defined analogously.

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 $P(\bar{A}) = 0$ 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

[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.

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 $P(\bar{A}) = 0.9994$) 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 $P(E|\bar{A}) = 0.03$. 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'. We will follow that line of thought in what follows.

2.2 A Brief Digression

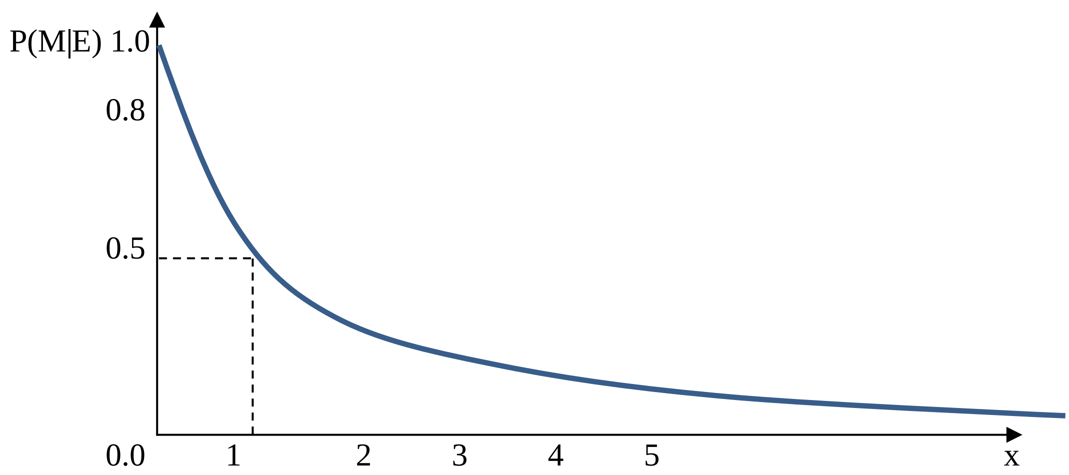
If we return to equation (1) and divide numerator and denominator by P(E|A)P(A), we obtain Bayes Rule in the form

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

where

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

Figure 1: Updated Probability



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

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

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

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

2.2 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 miracles, the quantity x defined in (5), simplifies to

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

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 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. $P(E|\overline{A})=p>>0$.

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 her performance. Presumably these phenomena are extraordinary enough for people to pay to see her, and not to be expected in the ordinary course of things. However, familiarity with the techniques of her 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).[4] 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).

In the two uninteresting cases above p is large relative to α, and so (from (6)) x will be large relative to unity and $P(M|E)=1/x$ 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 $P(M|E)$.

[4] 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.

The ratio x in (6) reflects the balancing principle outlined by Hume in the quote at the beginning of this paper. If the 'miracle' of false testimony $E|\bar{A}$ 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 believing in something that is better than even odds is straightforward enough, and is used in Bayesian hypothesis testing (Berger, 1980).[5] We will return to a discussion of x at the end of the paper.

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 $P(E|\bar{M})=p\approx 2\times 10^{-7}$ 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 they do not appear to have understood that since their event is evidential, extraordinary and uniquely plausible, all that matters is p/α. Naturally, a miracle believer should be grateful for any gifts, but this one is a token gesture.[6] 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.

Thus there are two points to take from the use of (6) as it applies to their argument.

Once (6) is relevant, as is the case for all interesting claims about extraordinary events, then detailed discussions, or concessions for the sake of argument, about P(M|E) are neither here nor there.

[5] 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.

[6] Doubtless the mistake was unintentional, for otherwise their request for thanks at the top of page 63 'Note that this is much higher than in our analysis of the lottery case' would be the height of insensitivity!

Second, Hume was right to emphasise the importance of x. But in this context we would note that this probability might be hard to judge. Thus Dawid and Gillies intuition that it would seem to be unreasonable to take p much lower than 10^{-3} is evidently not self-evident, and a certain amount of scepticism might be warranted for their prior probability too, a point to which we will return at the end of the paper. In the meantime, we note that if the prior probability in this particular case were to be, say, 5 per cent, the probability in (3) is approximately unity, even with a p of 10^{-3}.

2.5 Extreme Religious and Nonreligious Priors Make Evidence Irrelevant

We now return to the point made earlier in regard to equation (2). This point applies whether or not A satisfies the conditions for a miracle M. However, because we are interested in the case of 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)

One wonders if Dawkins, in this quote, is being a little too kind to himself, because if it is the case that he believes in a closed universe, which automatically rules M out, he would be untouched (indeed untouchable) by any evidence for M. Any sceptic who claims scientific education and insight to justify *ruling M out as impossible* is implicitly setting $\alpha=0$ and they are violating the fundamental principle of science that all theories can in principle be refuted by evidence.

Perhaps, though, it is we who are being unkind. Maybe a sceptic would respond that if the evidence were strong enough they would, in fact, believe in a miracle. The question to which we now turn is then the nub of the matter "What is the threshold of evidence-strength that is needed for a scientist to believe in a miracle?"

3 Hypothesis Testing

3.1 Bayesian Priors and Test Sizes

Although we concede that 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 Hume, try saying aloud ‘all the best evidence considered, I think M is more likely than not, but I’m going to believe $\overline{M}$’ and see how it sounds.

If we can accept Hume’s balancing principle (or, more correctly, a balancing principle in the spirit of Hume) as our standard of proof for a miracle then this amounts to believing in a miracle if and only if $x<1$.

We now make on an observation which, to our knowledge, is new. The condition $x<1$ for believing in a miracle (and disbelieving it if $x\geq 1$) can, because of our mathematical description of a miracle, call upon the approximation in equation (6) to deliver the following 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.

Anyone trained in classical statistics will recognize a comparison between the probability of obtaining evidence under a given belief with a small cut-off chosen by a researcher.

This is rather like a hypothesis test.

The insight obtained from this realization is that a vanishingly small prior probability such as 10^{-6} in Dawid and Gillies (1989) is rather like a vanishingly small test size in a hypothesis test. A practicing scientist might have to give a careful account of himself if he eschewed a standard 1, 5 or 10 per cent test size in favour of a test size of 10^{-4} per cent (i.e. a probability of 10^{-6}). So we do not think it unreasonable that a philosopher be scrutinized for a prior probability that small as well.

But we can make a stronger claim that this.

The probability $p=P(E|\overline{M})$ is *formally* equivalent to a p-value in a hypothesis test, rather than just being a metaphor for it. And so if we reinterpret the prior probability α as a test size, the rule given is (7) is *exactly* equivalent to a hypothesis test. In other words the following hypothesis test is equivalent to (7):

H_0: $\overline{M}$
H_1: M
Decision Rule: reject H_0 if p-value$=p=P(E|\overline{M})<\alpha$.[7] (8)

[7] We are aware that the equivalence between Bayesian and classical hypothesis testing in the case of miracles is not exact. To the extent that $(1-\alpha)/\alpha$ differs from $1/\alpha$ in (5) they will differ. But for small αs, such as 0.1, 0.05 and 0.01,

Thus, any decision rule for believing in a miracle that uses Hume's balancing principle can be written either as a Bayesian hypothesis test with prior α on $\overline{M}$, or, as a classical hypothesis test where that same α is the test size. In either case, we disbelieve $\overline{M}$ and believe M if $p<\alpha$.

3.2 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 a whole raft of scientific practice to guide us in our choice of Bayesian priors in miracle debates, which turn out to have been test sizes all along.

Thus a scientist who spends her days at the lab conducting experiments with a 5 per cent test size might, when reading about miracles in the evening, be justified in raising the prior probability to something like 5 per cent. We earlier noted that a value of α in Dawid and Gillies (1989) of 5 per cent (or even 1 per cent) would see us believe in their miracle.

The levels of test sizes used by scientists are context specific, so one can't be too dogmatic about out other test sizes, and by implication other reasonable priors for miracle debates.[8] Nevertheless, the most popular choice, 5 per cent, has proven to be fruitful, on average, and so it could arguable function as a default value.

> "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." Effron (2005 pg. 5)

In science, discipline goes both ways with regards to evidence – it restrains the credulous, but it can also challenge the sceptical. We often teach undergraduate students why test sizes are so small, but 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, but not so reluctant that evidence can be ignored. That is why these values settle at 'a few per cent', that is 1, 5 or 10 per cent. We speculate that this even-handed discipline may have been lacking in miracle debates.

3.3 Can Miracle Scepticism be Unscientific?

Thus the correspondence between a Bayesian prior and a test size presses upon us the importance of the arguments made, by Hume and others since, that sceptics about evidence are 'off the hook' because of the alleged special features of miracles. The final chapter of Dawkins (2011) echoes Hume in this regard.

---

and *a fortiori* for hyper-sceptical values of α like 10^{-6}, the range of p-values for which the two differ is negligible and can be practically ignored.

[8] There are situations where scientists have adopted lower test sizes than 1 per cent (Franklin, 2013), so Effron's comments which follow are a generalization.

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 the argument on Hume's core idea about probability – that we assign probabilities to things based on analogies with past events – and yet he rejects Hume's account. 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)

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 certainly 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 priors are equivalent to vanishingly low test sizes in hypothesis tests, so the issue has re-appeared in another guise.

To put this another way, there are no discontinuities with scepticism. It may be true that a sceptic can give logical lip service to the benefits of evidence by saying that in principle they could be persuaded of a miracle, if the evidence were good enough. But microscopically small priors, shown in this paper to be microscopic test sizes, blow their epistemological cover and make them just like religious 'fundamentalists' whom they despise.

References

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

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.

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/>.