Monty Hall, Doomsday and confirmation

DARREN BRADLEY & BRANDEN FITELSON

Imagine you are on a game show. You are faced with three doors (1, 2 and 3), behind one of which is a prize and behind the other two is no prize. In the first stage of the game, you tentatively select a door (this is your initial guess as to where the prize is). To fix our ideas, let's say you tentatively choose door 3. Then the host, Monty Hall, who knows where the prize is, opens one of the two remaining doors. Monty Hall can never open either the door that has the prize or the door that you tentatively choose – he must open one remaining door that does not contain the prize. Now you learn that Monty Hall has opened door 1. The standard question asked about this set-up is: should you now change your (tentative) choice from door 3 to door 2? This is typically seen as being equivalent to the following question: is the posterior probability that the prize is behind door 2 greater than the posterior probability that the prize is behind door 3? If various 'lottery' assumptions are made about the prior probabilities and the likelihoods in

ANALYSIS 63.1, January 2003, pp. 23–31.

this game, then (perhaps somewhat surprisingly) the answer to this question is 'yes'. But, the 'lottery' assumptions required for this conclusion about the posteriors are non-trivial, and they have been the source of great controversy about this game and its proper probabilistic analysis (see vos Savant 1995 for an entertaining discussion of the controversies involved).

In the present paper, we propose an alternative, confirmation-theoretic analysis of the Monty Hall problem that leads to a much more robust and less controversial argument. Here, we borrow from analogous confirmation-theoretic analyses of the Doomsday Argument. In §1, we begin with a discussion of the Doomsday Argument. We show that the Doomsday Argument – when reconstructed confirmation-theoretically – is quite robust, and does not require very strong 'lottery' assumptions about either priors or likelihoods to get off the ground. Then, in §2, we show how an analogous analysis of the Monty Hall problem leads to an *even more robust* argument that requires *no lottery assumptions whatsoever*.

1. *Confirmation-theoretic analysis of the Doomsday Argument*

Imagine there are three possibilities for how many people there are, and will ever be, in the entire universe. Either (H_1) there will be one person – called number 1, or (H_2) there will be two people – number 1 and number 2, or (H_3) there will be three people – number 1, number 2 and number 3. These people are always created in order. That is, there cannot be number 2 without there first being number 1, and there cannot be number 3 without there first being both number 1 and number 2. Now, you learn your birth rank (i.e. you learn that you were the i^{th} person born in the universe: E_i). To fix our ideas, assume you discover that you are number 2 (E_2). At this point, one might ask: is the posterior probability that the total population of the universe is 2 greater than the posterior probability that the total population of the universe is 3? In other words, is 'doom sooner' more probable a posteriori (i.e. conditional upon your birth-rank) than 'doom later'? In order to answer this question precisely, we would need to make some rather strong assumptions about the priors $\Pr(H_j)$ and the likelihoods $\Pr(E_i \mid H_j)$. In 'lottery' versions of the Doomsday Argument (e.g. Bartha & Hitchcock 1999: S342–5), it is typically assumed that the likelihoods satisfy the following constraint: $\Pr(E_i \mid H_j) = 1/j$, for all $i \leq j$. But, in order to derive an inequality between the posteriors $\Pr(H_2 \mid E_2)$ and $\Pr(H_3 \mid E_2)$, we would *also* need some strong assumptions about the *priors* $\Pr(H_j)$. The most natural 'lottery' assumption would be to make the H_j equiprobable, a priori. Given these two 'lottery' assumptions, Bayes's theorem shows that the answer to the comparative question about the posteriors is 'yes.' We present the argument formally now. First, some notation and terminology:

H_j = The total population of the universe is j
E_i = Your birth rank is i
n = The largest possible population (assumed, for analogy with Monty Hall, to be 3 here)
$\Pr(H_j)$ = The prior probability of H_j
$\Pr(H_j \mid E_i)$ = The posterior probability of H_j, given E_i
$\Pr(E_i \mid H_j)$ = The likelihood of H_j (on E_i)

We now formally deduce that $\Pr(H_2 \mid E_2) > \Pr(H_3 \mid E_2)$, given our two 'lottery' assumptions.[1]

(1) For all j, $\Pr(H_j) = 1/n = 1/3$ (and the H_j are exclusive and exhaustive)
(2) For all $i \leq j$, $\Pr(E_i \mid H_j) = 1/j$

$$\therefore (3) \quad \Pr(E_2) = \Pr(H_1) \times \Pr(E_2 \mid H_1) + \Pr(H_2) \times \Pr(E_2 \mid H_2) + \Pr(H_3) \times \Pr(E_2 \mid H_3) = \left(\frac{1}{3} \times 0\right) + \left(\frac{1}{3} \times \frac{1}{2}\right) + \left(\frac{1}{3} \times \frac{1}{3}\right) = \frac{5}{18}$$

$$\therefore (4) \quad \Pr(H_3 \mid E_2) = \frac{\Pr(H_3) \times \Pr(E_2 \mid H_3)}{\Pr(E_2)} = \frac{\frac{1}{3} \times \frac{1}{3}}{\frac{5}{18}} = \frac{2}{5}$$

$$\therefore (5) \quad \Pr(H_2 \mid E_2) = \frac{\Pr(H_2) \times \Pr(E_2 \mid H_3)}{\Pr(E_2)} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{5}{18}} = \frac{3}{5}$$

$$\therefore (6) \quad \Pr(H_2 \mid E_2) > \Pr(H_3 \mid E_2)$$

So, given our two 'lottery' assumptions, it is more probable a posteriori (i.e. given that your birth rank is 2) that the total population of the universe is 2 than it is that the total population of the universe is 3. This argument is fully general. That is, it will go through for *any n*. So long as n is finite, the 'lottery' assumptions (1) and (2) will suffice to show that 'doom sooner' has a greater posterior probability than 'doom later'.[2]

[1] There are also some *logical* constraints imposed on the likelihoods by the formulation of the Doomsday set-up (e.g. that $\Pr(E_i \mid H_j) = 0$ if $i > j$). We take such logical constraints for granted throughout, without comment.

[2] That is, if $i \leq j < k \leq n$ and n is finite, then the lottery assumptions (1) and (2) above will suffice to ensure that $\Pr(H_j \mid E_i) > \Pr(H_k \mid E_i)$. For simplicity, and for the purposes of analogy with the Monty Hall problem, we have assumed that n is finite (and known). This assumption can be relaxed in a confirmation-theoretic rendition of the argument (see Bartha & Hitchcock 1999 for a confirmation-theoretic rendition that allows n to be infinite). This is another advantage of thinking about Doomsday confirmation-theoretically rather than posterior-probabilistically.

Interestingly, this is *not* how the Doomsday Argument is typically formulated (see, for instance, Bartha & Hitchcock 1999; Bostrom 2002; Korb & Oliver 1999; Leslie 1997; Sober 2002). The most sophisticated versions of the argument begin with (something tantamount to) the following *different* question about the Doomsday set-up.

(Q) Does E_2 confirm H_2 more strongly than E_2 confirms H_3?

This is because satisfactorily answering the question about posteriors requires some strong and controversial assumptions about the priors of the H_j (like (1)). As it turns out, answering the *confirmation-theoretic* question (Q) does *not* require such strong and controversial assumptions. The confirmation-theoretic treatment is much more robust (and less controversial) than the posterior-probabilistic analysis, as we will now see.

Following many contemporary authors, including Horwich (1982), Howson & Urbach (1994), Milne (1995, 1996), and Schlesinger (1991, 1995), we will assume that the degree to which E confirms H is properly measured by the ratio $\Pr(H \mid E)/\Pr(H)$ of the posterior to the prior probability of H. Given this assumption about how to measure degree of confirmation, our question (Q) becomes:

$$(Q^*)\quad \text{Is it the case that } \frac{\Pr(H_2 \mid E_2)}{\Pr(H_2)} > \frac{\Pr(H_3 \mid E_2)}{\Pr(H_3)}?$$

An application of Bayes's theorem simplifies (Q^*) to the following *logically equivalent* question.[3]

$$(Q^*)\quad \text{Is it the case that } \Pr(E_2 \mid H_2) > \Pr(E_2 \mid H_3)?$$

What's neat about (Q^*) is that it can be answered in the affirmative without *assuming anything about the prior probabilities of* H_2 *and* H_3.[4] So, we can answer (Q^*) affirmatively without appeal to assumption (1) *or any other significant assumption about the priors of the* H_j. Moreover, we don't need as strong an assumption as (2) concerning the likelihoods of the H_j to get

[3] It is often seen as a distinguishing *virtue* of the ratio measure of degree of confirmation that whether E favours H_1 over H_2 depends *only* on the *likelihoods* of H_1 and H_2, and *not* their priors. That is, the ratio measure is distinguished because it satisfies the *Law of Likelihood* (Hacking 1965). A wide variety of philosophers and statisticians (both Bayesian and non-Bayesian) have defended the Law of Likelihood (see, for instance, Royall 1997 and Sober 1994). Other measures of confirmation that have been proposed in the literature *violate* this principle of comparative support (see Milne 1996 and Schlesinger 1991). For a recent reconstruction of the Doomsday Argument based *directly* on the Law of Likelihood, see Sober 2002.

[4] Except that $\Pr(H_2) \neq 0$. We will assume throughout that extreme probabilities are *only* assigned in cases where *logical* constraints apply (i.e. we will assume that Pr is *strictly coherent* in the sense of Shimony 1955).

an affirmative answer to (Q^*). All we need is the following weaker assumption about the likelihoods of the H_j:

(7) If $i \leq j < k$, then $\Pr(E_i \mid H_j) > \Pr(E_i \mid H_k)$.

All (7) requires is that the likelihood $\Pr(E_i \mid H_j)$ is a *strictly decreasing function of* j, for all $i \leq j < k$. This is weaker than the 'lottery' assumption (2), which requires *equi-likelihood*. Assumption (7) also seems to us more plausible than assumption (2) in the context of Doomsday. Here, (7) only requires that (in the absence of any other information) the probability of having a particular birth rank i in a universe of size j gets smaller as j gets larger. This does not require us to accept any 'principle of indifference (or insufficient reason)' concerning birth ranks and universe sizes.[5] Thus, a confirmation-theoretic rendition of the Doomsday Argument is bound to be substantially more robust than a posterior-probabilistic one. Next, we show how an analogous confirmation-theoretic treatment of the Monty Hall problem leads to an *even more robust* argument.

2. *Confirmation-theoretic analysis of the Monty Hall problem*

We begin with a brief review of the standard probabilistic analysis of the Monty Hall problem. Imagine you are on a game show. There are three doors in front of you (1, 2 and 3). You know that behind just one of them is a prize (let H_j be the hypothesis that the prize is behind door j). You get to make an initial guess. Let's say you guess door 3 (i.e. you guess H_3). Then the host, Monty Hall, who knows where the prize is, opens one of the two other doors (let E_i be the observation that Monty opens door i). He must open a remaining door that does not contain the prize. Say Monty Hall opens door 1 (i.e. you observe E_1). Typically, one is now asked the following question: is the posterior probability that the prize is behind door 2, $\Pr(H_2 \mid E_1)$, greater than the posterior probability that the prize is behind door 3, $\Pr(H_3 \mid E_1)$? As was the case with Doomsday, in order to answer this question precisely, we need to make some rather strong assumptions about the priors $\Pr(H_j)$ and the likelihoods $\Pr(E_i \mid H_j)$. In the standard treatments, the following two 'lottery' assumptions are made about the Monty Hall set-up.[6]

[5] Even this weaker assumption is controversial. Elliott Sober (2002) argues that assumption (7) – in the context of the Doomsday Argument – has implausible *empirical* consequences.

[6] The following two *logical* constraints on the likelihoods are also implicit in the formulation of the Monty Hall problem: (*i*) For all i and j, if $i \neq j$ and $j \neq 3$, then $\Pr(E_i \mid H_j) = 1$, and (*ii*) For all i and j, if $i = j$, then $\Pr(E_i \mid H_j) = 0$. As in the Doomsday case, we take such logical constraints for granted throughout, without comment.

(8) For all j, $\Pr(H_j) = 1/n = 1/3$ (and the H_j are mutually exclusive and exhaustive)

(9) For all i and j, if $i \neq j$ and $j = 3$ then $\Pr(E_i \mid H_j) = 1/(n-1) = 1/2$

We present the argument formally now. First, some notation.[7]

H_j = The prize is behind door j
E_i = Monty Hall opens door i
n = The total number of doors (typically, n is 3) = the # of the door you tentatively choose.

We now formally deduce that $\Pr(H_2 \mid E_1) > \Pr(H_3 \mid E_1)$, given our two 'lottery' assumptions.[8]

$$\therefore (10)\quad \Pr(E_1) = \Pr(H_1) \times \Pr(E_1 \mid H_1) + \Pr(H_2) \times \Pr(E_1 \mid H_2) + \Pr(H_3) \times \Pr(E_1 \mid H_3) = \left(\frac{1}{3} \times 0\right) + \left(\frac{1}{3} \times 1\right) + \left(\frac{1}{3} \times \frac{1}{2}\right) = \frac{1}{2}$$

$$\therefore (11)\quad \Pr(H_3 \mid E_1) = \frac{\Pr(H_3) \times \Pr(E_1 \mid H_3)}{\Pr(E_1)} = \frac{\frac{1}{3} \times \frac{1}{2}}{\frac{1}{2}} = \frac{1}{3}$$

$$\therefore (12)\quad \Pr(H_2 \mid E_1) = \frac{\Pr(H_2) \times \Pr(E_1 \mid H_2)}{\Pr(E_1)} = \frac{\frac{1}{3} \times 1}{\frac{1}{2}} = \frac{2}{3}$$

$$\therefore (13)\quad \Pr(H_2 \mid E_1) > \Pr(H_3 \mid E_1)$$

So, given the standard 'lottery' assumptions, it is more probable a posteriori (i.e. given that Monty Hall opens door 1) that the prize is behind door 2, and the player should revise the tentative choice of door 3 to a choice of door 2. A parallel argument can be made to show that a switch should also be made if Monty Hall opens door 2. So, given the symmetries of the

[7] To make the analogy to the Doomsday Argument clear, the hypothesis that there is just 1 person in the universe is like the first door containing the prize, the hypothesis that there are 2 people is like the second door containing the prize and the hypothesis that there are 3 people is like the third door containing the prize. The H_j's are strictly analogous between ($n = 3$) Doomsday and Monty Hall. Moreover, learning that you are number 2 is analogous to Monty Hall opening door 1. In both cases, H_1 is eliminated as a possibility. More generally, E_i in the Monty Hall problem corresponds (roughly) to E_{i+1} in the ($n = 3$) Doomsday Argument.

[8] See Cross 2000 for a canonical layout of the argument behind the standard ($n = 3$) Monty Hall problem. Chun (1999) shows how to generalize the argument (in various ways) to $n > 3$ doors.

problem, the player should *always* switch doors once Monty Hall opens a door – *no matter which door is tentatively chosen and no matter which door is opened*! Many people find this result counter-intuitive. It is often thought that we should be indifferent between the two remaining doors (and *not* be motivated to switch doors). There has been much written about this issue (see, for instance, Chun 1999; Cross 2000; vos Savant 1995).

We will not rehearse the various debates surrounding Monty Hall here. For our present purposes, it will suffice to point out that the conclusion (13) of this standard argument depends on two substantive assumptions about the agent's degrees of belief. The first assumption is (8), that the H_j should be *equiprobable*, a priori. The second assumption is (9), that the likelihoods of the H_j should be split equally between the two remaining possible door eliminations (provided these likelihoods are non-extreme).[9] Next, we will show that a confirmation-theoretic analysis of the Monty Hall problem obviates the need to make either of these two (potentially controversial) assumptions about the Monty Hall agent's degrees of belief. (8) can be (effectively) disposed of, and (9) can be *substantially* weakened.

As was the case with the Doomsday Argument, we may ask the following (simplified, analogous) *confirmation-theoretic* question about the Monty Hall problem:

(Q') Is it the case that $\Pr(E_1 \mid H_2) > \Pr(E_1 \mid H_3)$? (i.e. Does E_1 *favour* H_2 over H_3?)

And, as was the case with Doomsday, we may give an affirmative answer to (Q') without making *any* assumption about the prior probabilities of the H_j (except that they are non-zero), and without making as strong an assumption as (9) about the likelihoods $\Pr(E_i \mid H_j)$. All we need for an affirmative answer to (Q') is assumption (7), which is *substantially* weaker than (9). In the Monty Hall case, (7) only requires that $\Pr(E_1 \mid H_3) < 1$.[10] But, that is *nearly trivial*, since all it requires is that *Monty Hall might not open door 1 if the prize is behind door 3*. So, even if you object to the standard posterior-probabilistic analysis of the Monty Hall problem, it seems you must agree (given the symmetries of the problem) that – *no matter which*

[9] In other words, (9) says that the probabilities of Monty Hall eliminating doors 1 or 2 (conditional on the location of the prize) are the same (provided that these probabilities are non-extreme).

[10] There is an interesting corollary to this result. If one *grants* the uniform prior distribution assumption (8), then (13) is *guaranteed*, *unless* one assigns $\Pr(E_1 \mid H_3) = 1$. This is another sense in which the Monty Hall argument is more robust than the Doomsday Argument. Moreover, since there is no *logical* constraint imposed by H_3 on E_1 in the Monty Hall set-up, $\Pr(E_1 \mid H_3) < 1$ follows from the mere *strict coherence* (Shimony 1955) of Pr *alone*. As such, our confirmation-theoretic rendition of the Monty Hall problem involves *no lottery assumptions whatsoever*.

door Monty Hall opens and no matter which door you tentatively chose – Monty Hall's door-opening provides *better evidence* for the hypothesis that the prize is behind the door you did *not* tentatively choose than it does for the hypothesis that your tentative choice was correct.[11]

Stanford University
Stanford, CA 94305–2155, USA
dbradley@stanford.edu

San José State University
San José, CA 95192–0096, USA
branden@fitelson.org

References

Bartha, P. and C. Hitchcock 1999. No one knows the date of the hour: an unorthodox application of Rev. Bayes's theorem. *Philosophy of Science* 66 (Proceedings): S329–53.

Bostrom, N. 2002. *Anthropic Bias: Observation Selection Effects in Science and Philosophy*. New York: Routledge.

Chun, Y. 1999. On the information economics approach to the generalized game show problem. *American Statistician* 53: 43–51.

Cross, C. 2000. A characterization of imaging in terms of Popper functions. *Philosophy of Science* 67: 316–39.

Hacking, I. 1965. *Logic of Statistical Inference*. Cambridge: Cambridge University Press.

Horwich, P. 1982. *Probability and Evidence*. Cambridge: Cambridge University Press.

Howson, C. and P. Urbach 1994. *Scientific Reasoning: The Bayesian Approach*. La Salle: Open Court.

Korb, K. and J. Oliver 1999. A refutation of the doomsday argument. *Mind* 107: 403–10.

Leslie, J. 1997. Observer-relative chances and the doomsday argument. *Inquiry* 40: 427–36.

Milne, P. 1995. A Bayesian defence of Popperian science? *Analysis* 55: 213–15.

Milne, P. 1996. $log(p(h|eb)/p(h|b))$ is the one true measure of confirmation. *Philosophy of Science* 63: 21–26.

Royall, R. 1997. *Statistical Evidence: A Likelihood Paradigm*. London: Chapman & Hall.

Schlesinger, G. 1992. *The Sweep of Probability*. South Bend: Notre Dame University Press.

Schlesinger, G. 1995. Measuring degrees of confirmation. *Analysis* 55: 208–12.

Shimony, A. 1955. Coherence and the axioms of confirmation. *Journal of Symbolic Logic* 20: 1–28.

[11] The authors wish to thank Nick Bostrom, Ben Bradley, Otávio Bueno, Andrew Celotto, Mark Colyvan, Ben Escoto, Tim German, Alan Hájek, Chris Hitchcock, Mehul Mathlkia, Elliott Sober, Stephen Street, Michael Strevens and Daniel Wallis for useful discussions and comments.

Sober, E. 1994. Contrastive empiricism. In his *From a Biological Point of View*. Cambridge: Cambridge University Press.

Sober, E. 2002. An empirical critique of two versions of the doomsday argument – Gott's line and Leslie's wedge. *Synthese* 133.

vos Savant, M. 1995. *The World's Most Famous Math Problem. The Proof of Fermat's Last Theorem and Other Mathematical Mysteries*. New York: St. Martin's Press.