

ARTICLE

Have Bayesians Solved the Paradox of the Ravens?

Amit Karmon 

University of Oxford, UK
Email: amit.karmon@reuben.ox.ac.uk

(Received 21 May 2025; revised 04 March 2026; accepted 07 March 2026)

Abstract

The standard Bayesian solution to the paradox of the ravens maintains that the degree of confirmation provided by seeing a nonblack nonraven is positive but negligible compared to that provided by seeing a black raven. I show that, unless we impose severe and unmotivated restrictions on the subject's priors, this has the consequence that the cumulative confirmation provided by all the nonblack nonravens the subject expects to see is nonnegligible compared to the cumulative confirmation provided by all the black ravens the subject expects to see. If this is so, however, then the paradox retains its full force.

1. Introduction

There is, I think, a widely (though not unanimously) held view among philosophers that Hempel's (1945) paradox of the ravens is no more than a historical anecdote because the paradox has already been solved by the Bayesian approach.¹ We can find this view expressed in some encyclopedic entries on confirmation (Huber 2024) and promoted in slides shown to undergraduate students in epistemology or philosophy of science classes. The view that a Bayesian approach can solve the paradox of the ravens, I hope to show, is far from the truth. The Bayesian approaches to the paradox don't even begin to solve it.

Let's present the paradox. Take the following seemingly plausible assumptions. First, special cases of a general hypothesis confirm the hypothesis. For example, a black raven is a special case of the hypothesis that all ravens are black, so if we see an

¹ Examples of Bayesian solutions to the paradox can be found as early as Hosiasson-Lindenbaum (1940), Alexander (1958), Mackie (1963), and others. More recent Bayesian treatments of the paradox include Horwich (1982) and Howson and Urbach (1993). Vranas (2004) presents a version of a Bayesian solution that is sometimes referred to in the literature as the *Standard* or *Canonical* Bayesian Solution. Fitelson and Hawthorn (2010a, 2010b, and 2010c, which is an addendum to 2010b) discuss the paradox and the Bayesian approach to it in detail. Rinard (2014) criticizes the canonical solution and promotes adopting what she takes to be an alternative Bayesian solution (but see fn. 26).

object that is a black raven, this fact confirms the hypothesis that all ravens are black. This assumption is referred to as *Nicod's Condition (NC)*. Second, if evidence E confirms a hypothesis H , and H is logically equivalent to another² hypothesis H' , then E confirms H' too, and to the same degree. This assumption is referred to as the *Equivalence Condition (EC)*. From these seemingly plausible assumptions, we can derive the *Paradoxical Conclusion (PC)* that the fact we saw, for example, a white shoe confirms the hypothesis that all ravens are black.

The derivation of *PC* is quite simple. Take the hypothesis H' stating that all nonblack things are nonravens. A white shoe is a special case of H' , so according to *NC*, seeing a white shoe confirms H' . But H' is logically equivalent to H , the hypothesis that all ravens are black. Thus, according to *EC*, seeing a white shoe also confirms H .

This seems paradoxical because white shoes (and brown tables and green apples) seem completely irrelevant to the color of ravens and learning about the color of shoes doesn't seem to provide any information about the color of ravens. As Goodman (1983, 70) quipped, this can seem to give rise to a new field of indoor ornithology. We are left with the dilemma of whether to accept this seemingly paradoxical conclusion or reject one of the premises. If we choose to accept the conclusion, we had better also be able to explain why it seems paradoxical at first sight.

By "the Bayesians," I shall refer for now to those who choose to accept that both black ravens and nonblack nonravens confirm the hypothesis that all ravens are black, and who therefore accept the seemingly paradoxical conclusion. However, they emphasize that confirmation is a matter of degree, and from the fact that both black ravens and nonblack nonravens confirm the hypothesis that all ravens are black, it doesn't follow that a nonblack nonraven confirms the hypothesis that all ravens are black to the same degree as a black raven does. In fact, with further assumptions about our prior probabilities of sampling objects from our universe, we can derive the conclusion that the degree of confirmation provided by a nonblack nonraven is negligible compared to the degree of confirmation provided by a black raven. The conclusion seemed paradoxical to us because it did not include further information that the degree of confirmation provided by a nonblack nonraven is negligible. People take nonblack nonravens to be irrelevant to the hypothesis that all ravens are black because they confuse a minute degree of confirmation with zero confirmation. It doesn't change a great deal in practice.

Thus, in what follows, to qualify as a "Bayesian" solution to the paradox, the two desiderata below need to follow from it:

D₁: Both sampling a black raven and sampling a nonblack nonraven confirm the hypothesis that all ravens are black.

D₂: The degree of confirmation provided to the hypothesis that all ravens are black by sampling a nonblack nonraven is negligible compared to the degree of confirmation provided by sampling a black raven.

² I individuate hypotheses as linguistic entities, such that H and H' can be seen as different hypotheses even though they are logically equivalent. Alternatively, we can also say that H and H' are two guises of the same hypothesis and let *EC* to state that if evidence E confirms a hypothesis H under one guise, then it confirms H to the same extent under every guise.

Some might wonder whether this is too restrictive. It should suffice for a solution to be qualified as “Bayesian” if it uses statistical and probabilistic considerations to solve the paradox, whether they satisfy D_1 and D_2 in the preceding text or not. While indeed the most famous Bayesian solutions satisfy D_1 and D_2 , maybe other solutions that can be qualified as “Bayesian” don’t. I will say something about this issue in the conclusion.

I hope to show that any solution that satisfies D_1 and D_2 suffers from a fatal flaw. The central idea of this article is a rather simple one. A satisfying solution to the Paradox of the Ravens need not only show that the support provided to the hypothesis that all ravens are black by a *single* nonblack nonraven is negligible compared to the support provided by a *single* black raven. It also needs to show that the support provided by *all* the nonblack nonravens we see is negligible compared to the support provided by *all* the black ravens we see. Otherwise, even if the support provided by a single nonblack nonraven is negligible compared to the support provided by a single black raven, it might be that we see many more nonblack nonravens than black ravens and the cumulative support provided by all the nonblack nonravens we see can be nonnegligible compared to the cumulative support provided by all the black ravens we see. This way, the prospects of indoor ornithology are on the rise, and we are left again in a paradoxical position.

On most Bayesian solutions, *what makes it the case* that a nonblack nonraven contributes negligibly compared to a black raven is exactly the (obviously true) assumption that it is overwhelmingly more likely when we sample an object at random to sample a nonblack object than to sample a raven. As we shall see in the next section, this assumption implies that it is also overwhelmingly more likely to sample a nonblack nonraven than to sample a black raven. Thus, the attention to the cumulative support provided by all nonblack nonravens we see and all black ravens we see is a natural worry.

If the cumulative support provided by all nonblack nonravens we see is nonnegligible, then the paradox retains its full force. We can still do indoor ornithology and learn about the color of ravens without ever seeing any ravens. To get a feel for the kind of results that will follow, suppose that the cumulative support provided by all the nonblack nonravens I see is equal to the cumulative support provided by all the black ravens I see. Suppose further that my late grandmother, who had similar priors as I have, saw roughly similar things in her life as I did, except that she lived all her life never seeing any raven (or getting any testimony about their color, for that matter). It follows that my grandmother, who was more than twice my age when she passed away, had stronger evidence than I have for the hypothesis that all ravens are black, even though she had never seen any ravens. She probably saw more than double the amount of nonblack nonravens that I saw at the time of her death, and they more than compensate for all the ravens I have seen, and she hadn’t. If the cumulative contribution of all the nonblack nonravens I see is (say) half that of the black ravens, we only need to go back in time to a moment when my grandmother’s age tripled mine (I was old enough to know the color of ravens by then).³

³ An anonymous reviewer raised the point that it seems that there are things we can learn about ravens by just seeing nonravens. For instance, a person who lived in Beijing all their life and have never seen any ravens can have strong evidence that ravens don’t live in Beijing. For considerations that go

Thus, what seems to matter more is not whether a single nonblack nonraven provides a nonnegligible degree of confirmation, but rather whether the *cumulative* support that all the nonblack nonravens we saw give to the hypothesis that all ravens are black is still negligible compared to the support provided by all the black ravens we saw. The paradox stemmed from the paradoxical flavor of the conclusion that we can learn the color of ravens by looking only at nonblack nonravens. The Bayesians allegedly solved the paradox by showing that under some plausible assumptions, a single nonblack nonraven contributes only negligibly. But it is barely a consolation if, from these very assumptions, it follows that the tons of nonblack nonravens we see will together let us gain substantive confirmation for the claim that ravens are black, without us ever seeing any ravens.

In my view, the need to show that the cumulative support of all the nonblack nonravens is negligible in normal circumstances is part of the difficulty posed by the original paradox, surprisingly unnoticed as it is. Hence the title and my choice of words in the article. Some might contend that this is not part of the original puzzle but rather a separate revenge puzzle. I don't want to pay too much attention to this matter because nothing crucial hangs on that. It's open for the readers to defend the Bayesians' honor and insist that the Bayesians *did* solve *the* Paradox of the Ravens, even if I'm right to say that they lack the means to solve an equally disturbing revenge puzzle that stems naturally from the first. The bottom line will remain the same.

To sum up, I take the Paradox of the Ravens to pose another desideratum that must follow from every plausible solution to it:

D₃: The totality of nonblack nonravens we see contributes negligibly to the confirmation of the hypothesis that all ravens are black compared to the totality of black ravens we see.

To solve the paradox, one needs to derive *both* D_2 and D_3 from plausible assumptions. The Bayesians want to establish D_1 and D_2 , and they (like everyone else) usually ignore D_3 . I aim to show that, except under some extremely narrow (and usually implausible) assumptions, D_3 is incompatible with D_1 and D_2 . Therefore, to solve the Paradox of the Ravens, that is, to establish D_2 and D_3 , we need to reject D_1 , that is, we need it to be the case that a nonblack nonraven does *not* positively support the hypothesis that all ravens are black.

beyond the scope of this article, I don't fully agree with that. However, in any case, this seems like a general point about biased sampling procedures. If your sampling procedure is biased not to include objects with property P , then merely failing to sample objects of type T is some evidence that objects of type T have the property P . When a fisherman repeatedly casts a net with a mesh size of 5 cm in a lake and catches nothing, this is evidence that all the fish in the lake are smaller than 5 cm in size. Similarly, if for some reason your sampling procedure cannot sample black objects, and you don't see ravens, this is evidence in favor of the hypothesis that all ravens are black. However, this is somewhat orthogonal to the discussion of the Paradox of the Ravens, which is not about the evidential import of biases in the sampling procedure. The paradox is formulated without assuming a biased sampling procedure. It is not clear how actual biases in our sampling procedures can tell us why the Paradoxical Conclusion is not really paradoxical or why one of the premises of the paradox is false. I thank Bernhard Salow for suggesting this line of response to me.

Showing that inevitably requires some undergraduate-level calculus. I tried to make it as accessible as possible for readers who are not fluent with these topics, and I have put some of the more technical details in the footnotes and the appendix. In a sense, there is nothing particularly important in the specific mathematical details. The philosophical value of the mathematical details comes from the conclusion they support.

The rest of the article is structured as follows. In the next section, I present the most famous version of a Bayesian solution, sometimes referred to in the literature as the *standard*, or the *canonical*, Bayesian solution. This version is adapted from Vranas (2004) and is the one that, to my impression, is commonly shown to undergraduates in their introductory classes.⁴ In section 3, I show why the assumptions in Vranas's version imply that (except for extreme conditions) the cumulative contribution of all the nonblack nonravens is not negligible compared to that of the black ravens. In section 4, I explain why that problem is not specific to Vranas's version. The desiderata D_1 - D_3 can hold together only under implausible conditions. Because the argument for that is more technical, the details are in the appendix. Section 5 is a conclusion.

2. The standard Bayesian solution

The standard Bayesian solution aims to establish D_1 and D_2 from our priors regarding the probabilities of sampling black ravens, nonblack ravens, black nonravens, and nonblack nonravens and from whether and how these probabilities change if the hypothesis that all ravens are black is true. It is common to model the situation as follows: We sample objects from the universe, like drawing balls from an urn. Because we are going to discuss multiple draws in the following text, we will assume we replace the ball we draw back into the urn.⁵ According to the objects we sample (the balls we draw), we try to estimate what the rest of the objects in the universe are (what all the balls in the urn are like). Let's denote the hypothesis that all ravens are black by H and denote the propositions that the object in question that we have sampled is a black raven, or nonblack raven, or black nonraven, or nonblack nonraven, by br , or nbr , or bnr , or $nbnr$, accordingly.

We can get all the relevant information regarding the prior probability distribution by specifying the probability of sampling an object from each category conditional on H and conditional on $\sim H$, and by specifying the prior probability of H . That is, we need to specify the values of all the probabilities in Table 1 and Table 2.

⁴ Vranas doesn't necessarily adopt it as a solution. His paper is a critical but undecided evaluation of whether the solution is successful.

⁵ Modeling the situation with replacement simplifies the calculation but is not crucial. In section (ii) of the appendix, I show that under the assumption that (a) according to the subject's priors the number of objects depends only negligibly on whether H is true, and (b) the number of black ravens the subject sees is small compared to the total number of black ravens in the universe, the divergence between modeling the situation with or without replacement is small. Thus, solely modeling the situation without replacement is unlikely to change the overall verdict. I thank an anonymous reviewer for encouraging me to clarify that.

With probability $P(H)$:

Table 1. Prior probabilities conditional on H

	r	$\sim r$
b	$P(br H)$	$P(bnr H)$
$\sim b$	$P(nbr H)$	$P(nbnr H)$

With probability $P(\sim H)$:

Table 2. Prior probabilities conditional on $\sim H$

	r	$\sim r$
b	$P(br \sim H)$	$P(bnr \sim H)$
$\sim b$	$P(nbr \sim H)$	$P(nbnr \sim H)$

That is, to fully specify our priors, we need the value of the following ten variables: $P(br|H)$, $P(nbr|H)$, $P(bnr|H)$, $P(nbnr|H)$, $P(br|\sim H)$, $P(nbr|\sim H)$, $P(bnr|\sim H)$, $P(nbnr|\sim H)$, $P(H)$, and $P(\sim H)$.

Not all ten variables are independent. First, of course, $P(\sim H) = 1 - P(H)$. Moreover, $P(br|H) + P(nbr|H) + P(bnr|H) + P(nbnr|H)$ should sum up to one as together the four variables cover the whole universe (conditional on H). Ditto for $P(br|\sim H) + P(nbr|\sim H) + P(bnr|\sim H) + P(nbnr|\sim H)$. That is, one in each of these sets of four variables is dependent on the other three. Because $P(bnr|H)$ and $P(bnr|\sim H)$ will not play any role in the calculations that follow, we will treat them as the dependent variables. Furthermore, due to the specific meaning of H (i.e., all ravens are black), $P(nbr|H)$ must be zero. Thus, we are left with six independent variables:

$$P(br|H), P(nbnr|H), P(br|\sim H), P(nbr|\sim H), P(nbnr|\sim H), \text{ and } P(H).$$

By “independent,” of course, I don’t mean that they can get any value. We are talking about probabilities, so each must range between 0 and 1, and as I said, the probability that an object we sampled is from the whole universe needs to be 1. To avoid triviality and divisions by zero, let’s adopt the common practice and assume that all these six variables are *strictly* between 0 and 1. The Bayesians aim to solve the paradox for all plausible cases anyway, so surely there is no harm in assuming that.

Note that the probability of sampling a black raven, say, regardless of whether H is true is (according to the law of total probability) simply $P(br) = P(H)P(br|H) + P(\sim H)P(br|\sim H)$, and ditto for sampling a nonblack raven, black nonraven, or nonblack nonraven. Thus, whenever we see a probability of sampling one of the four categories unconditional on H or $\sim H$, we can treat it as an abbreviation of an expression that uses only the preceding six variables.

We are now ready to present the standard Bayesian solution. It is common to present it as follows. Our background knowledge, it is claimed, seems to agree with the following three assumptions (nb stands for the proposition that the object in question we have sampled is nonblack, and r stands for it being a raven):⁶

$$A_1) P(nb) \gg P(r)$$

$$A_2) P(r|H) = P(r)$$

$$A_3) P(nb|H) = P(nb)$$

A_1 states that the probability of sampling a nonblack object at random is overwhelmingly higher than the probability of sampling a raven. A_2 states that the probability of sampling a raven is independent of whether the hypothesis that all ravens are black is true, that is, the probability of sampling a raven conditional on the hypothesis that all ravens are black is simply the probability of sampling a raven in general. A_3 states that the probability of sampling a nonblack object is also independent of the hypothesis that all ravens are black. Note that it's a direct result of A_2 and A_3 that $P(r|H) = P(r|\sim H)$ and $P(nb|H) = P(nb|\sim H)$.⁷

These three assumptions suffice to satisfy D_1 and D_2 . There are various ways to show that. It will be useful for what follows to present the case in terms of the preceding six independent variables, so let's see what each of A_1 - A_3 tells us about them. Let's denote the ratio $\frac{P(nb)}{P(r)}$ by m . A_1 tells us that m is a huge number. Note that because both nb and r are probabilistically independent of H (as we learn from A_2 and A_3), m is independent of whether H is true or not, that is, $m = \frac{P(nb)}{P(r)} = \frac{P(nb|H)}{P(r|H)} = \frac{P(nb|\sim H)}{P(r|\sim H)}$. It follows⁸ that:

$$P(nbnr|H) = mP(br|H) \tag{2.1}$$

And therefore:

$$P(nbnr|H) \gg P(br|H) \tag{2.2}$$

Moreover, it follows⁹ from A_2 and A_3 that:

$$P(nbnr|\sim H) = mP(br|\sim H) + (m - 1)P(nbr|\sim H) \tag{2.3}$$

Therefore, $P(nbnr|\sim H)$ is at least m times greater than $P(br|\sim H)$ and at least $(m-1)$ times greater than $P(nbr|\sim H)$. If $P(br|\sim H)$ and $P(nbr|\sim H)$ are not too small compared to one another, then $P(nbnr|\sim H)$ is even more than m times greater than both $P(br|\sim H)$ and $P(nbr|\sim H)$. That is, we can derive from A_1 both:

⁶ Usually, the assumptions are presented with the probabilities conditionalized on our background knowledge K . E.g., A_1 is presented as $P(nb|K) \gg P(r|K)$, and ditto for A_2 and A_3 . This is because, as Good (1967) has famously shown, according to some background knowledge black ravens can *disconfirm* H , so it's a good practice to always conditionalize on our background knowledge K . However, because K will not contribute anything to the following calculations, I chose to omit it.

⁷ It follows from the law of total probability. For instance, in the case of A_2 :

$$P(r) = P(H)P(r|H) + P(\sim H)P(r|\sim H) = P(H)P(r) + P(\sim H)P(r|\sim H).$$

Therefore, $(1 - P(H))P(r) = P(\sim H)P(r) = P(\sim H)P(r|\sim H)$. Therefore, $P(r) = P(r|\sim H)$.

⁸ Because $P(nbr|H) = 0$, we have $P(nbnr|H) = P(nb|H) = P(nb) = mP(r) = mP(r|H) = mP(br|H)$.

⁹ $P(nbnr|\sim H) + P(nbr|\sim H) = P(nb|\sim H) = P(nb) = mP(r) = mP(r|\sim H) = mP(br|\sim H) + mP(nbr|\sim H)$. Thus, $P(nbnr|\sim H) = mP(br|\sim H) + (m - 1)P(nbr|\sim H)$.

$$P(nbr|\sim H) \gg P(br|\sim H) \quad (2.4)$$

$$P(nbr|\sim H) \gg P(r|\sim H) \quad (2.5)$$

The \gg sign is interpreted as at most slightly weaker than the \gg sign in A_1 and in many (perhaps most) cases as even stronger.

From A_2 we infer that $P(r|H)=P(r|\sim H)$ and therefore we can derive¹⁰ that:

$$P(br|H) = P(br|\sim H) + P(nbr|\sim H) \quad (2.6)$$

From A_3 we know that $P(nb|H)=P(nb|\sim H)$ and therefore we can similarly derive¹¹ that:

$$P(nbr|H) = P(nbr|\sim H) + P(nbr|\sim H) \quad (2.7)$$

Now, there is no single agreed-upon way to measure the degree of confirmation a piece of evidence E confers on a hypothesis H . In the literature (Fitelson and Hawthorne 2010a), there are three major relevance measures:

The Difference: $d(H, E) = P(H|E) - P(H)$

The Log-Ratio: $r(H, E) = \log\left(\frac{P(H|E)}{P(H)}\right)$

The Log-Likelihood Ratio: $l(H, E) = \log\left(\frac{P(E|H)}{P(E|\sim H)}\right)$

Among these three, the third one, l , is the most suitable for quantitative comparisons, and there are independent reasons to prefer it (Fitelson 2001). For this reason, I'll use this measure in this article.¹²

Measured by the log-likelihood ratio, the degrees of confirmation provided to the hypothesis that all ravens are black by a single black raven and by a single nonblack nonraven are:¹³

$$l_{br} = \log\left(1 + \frac{P(nbr|\sim H)}{P(br|\sim H)}\right) \quad (2.8)$$

$$l_{nbr} = \log\left(1 + \frac{P(nbr|\sim H)}{P(nbr|\sim H)}\right) \quad (2.9)$$

¹⁰ $P(r|H) = P(br|H) + P(nbr|H) = p(br|H) + 0 = P(br|H)$. Now, because $P(r|H) = P(r|\sim H)$ we get that: $P(br|H) = P(r|\sim H) = P(br|\sim H) + P(nbr|\sim H)$.

¹¹ $P(nb|H) = P(nbr|H) + P(nbr|H) = P(nbr|H) + 0 = P(nbr|H)$. Because $P(nb|H) = P(nb|\sim H)$, we get: $P(nbr|H) = P(nb|\sim H) = P(nbr|\sim H) + P(nbr|\sim H)$.

¹² Nothing crucial seems to change in the case of the other measures. The same results we get using the log-likelihood ratio seem to hold using the other measures as well. The math in these cases is much more complicated so it's harder to prove the results with the same rigor. Suffice it to say that if for some reason one is reluctant to use the log-likelihood ratio as a measure, one carries the burden of proving that the other measures don't lead to the same bottom line.

¹³ According to (2.6) and (2.7),

$$l_{br} = \log\left(\frac{P(br|H)}{P(br|\sim H)}\right) = \log\left(\frac{P(br|\sim H) + P(nbr|\sim H)}{P(br|\sim H)}\right) = \log\left(1 + \frac{P(nbr|\sim H)}{P(br|\sim H)}\right)$$

$$l_{nbr} = \log\left(\frac{P(nbr|H)}{P(nbr|\sim H)}\right) = \log\left(\frac{P(nbr|\sim H) + P(nbr|\sim H)}{P(nbr|\sim H)}\right) = \log\left(1 + \frac{P(nbr|\sim H)}{P(nbr|\sim H)}\right).$$

Because all the preceding terms are positive, we see immediately that the confirmation provided by sampling a black raven and by sampling a nonblack nonraven are both positive and therefore both the sampling of a black raven and the sampling of a nonblack nonraven confirm the hypothesis that all ravens are black. Thus, D_1 is satisfied. Furthermore, because $P(nbnr|\sim H) \gg P(br|\sim H)$ we get that $\frac{P(nbnr|\sim H)}{P(br|\sim H)} \gg \frac{P(nbr|\sim H)}{P(nbnr|\sim H)}$. Therefore, the contribution of a nonblack nonraven is negligible compared to that of a black raven, and D_2 is also satisfied. However, we shall now see that D_3 , in general, cannot be satisfied, except under extreme and implausible assumptions regarding our priors.

3. The nonblack nonravens strike back

To satisfy D_3 , we need to compare the support provided to H by all the black ravens we see and all the nonblack nonravens we see. Because we model each sampling of a black raven or a nonblack nonraven as independent and the log-likelihood ratio is additive, we can compare instead the support H receives from seeing a single black raven, and the support H receives from seeing the number of nonblack nonravens we see per black raven. What is this number?

The expected value of the number of nonblack nonravens we see for each black raven is given by $\frac{P(nbnr)}{P(br)}$, which under A_1 - A_3 can be rewritten as $\frac{P(nbnr|\sim H)+P(H)P(nbr|\sim H)}{P(br|\sim H)+P(H)P(nbr|\sim H)}$. One can argue that $\frac{P(nbnr)}{P(br)}$ is not the right value to use. We can narrow the discussion to the confirmation provided by all the nonblack nonravens only when it is in fact the case that all ravens are black. In such a scenario, the expected value of the number of nonblack nonravens we see for each black raven is given by $\frac{P(nbnr|H)}{P(br|H)}$ that under A_1 - A_3 can be rewritten as $\frac{P(nbnr|\sim H)+P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)}$, which is smaller than $\frac{P(nbnr)}{P(br)}$. Taking $\frac{P(nbnr|H)}{P(br|H)}$ as the expected number simplifies the calculations, and the resulting expected cumulative support of all the nonblack nonravens we see per raven is smaller than if we had taken $\frac{P(nbnr)}{P(br)}$ instead as the relevant magnitude. Because I want to be generous toward the Bayesian solution and simplify the calculation, I'll indeed use $\frac{P(nbnr|H)}{P(br|H)}$ as the relevant magnitude (let's call it choice 1). I'll sometimes say (in the footnotes, without the full calculation) what is added by taking $\frac{P(nbnr)}{P(br)}$ instead (let's call it choice 2).

In section (i) of the appendix, I show that under the assumption that each sampling of a nonblack nonraven is independent,¹⁴ the contribution of this number of nonblack nonravens can be approximated by:¹⁵

$$l_{nbnr's \text{ per } br} \approx \frac{1}{\frac{P(br|\sim H)}{P(nbr|\sim H)} + 1} \tag{3.1}$$

The ratio $\frac{P(br|\sim H)}{P(nbr|\sim H)}$ appeared (in its inverse form) in l_{br} as well (see 2.8). This ratio tells us what we take in advance the proportion between black ravens and nonblack ravens

¹⁴ See section (ii) of the appendix for why the assumption of independence is not crucial.

¹⁵ If we take choice 2, we get instead $l_{nbnr's \text{ per } br} \approx \frac{1}{\frac{P(br|\sim H)}{P(nbr|\sim H)} + P(H)}$.

to be, in case not all ravens are black. For instance, if we think in advance that if not all ravens are black, then half of them are, then the value of this ratio is 1. If we think in advance that if not all ravens are black, then 10 percent of them are, then the value of this ratio is $\frac{1}{9}$. If we think in advance that if not all ravens are black, then either (with a probability of 50 percent) half of them are, or (with a probability of 50 percent) none of them are, then the value of this ratio is $\frac{0.5 \cdot 0.5 + 0.5 \cdot 0}{0.5 \cdot 0.5 + 0.5 \cdot 1} = \frac{1}{3}$. In general, I think, we expect this ratio not to be extremely big or extremely small.¹⁶ At the very least, it seems *legitimate* to take it in advance to be moderate.

For simplicity, let's denote this ratio by x , that is, $\frac{P(br|\sim H)}{P(nbr|\sim H)} = x$. We can then rewrite l_{br} and $l_{nbr's\ per\ br}$ as $l_{br} = \log\left(1 + \frac{1}{x}\right)$ and $l_{nbr's\ per\ br} = \frac{1}{1+x}$.¹⁷ To satisfy D_3 , we need $\frac{1}{1+x}$ to be *negligible* compared to $\log\left(1 + \frac{1}{x}\right)$. One way to formalize this requirement is with the following inequality:

$$\log\left(1 + \frac{1}{x}\right) > K\left(\frac{1}{1+x}\right) \quad (3.2)$$

Here K is a big *negligibility constant*. For instance, if to be negligible the value of $l_{nbr's\ per\ br}$ needs to be (say) less than 1 percent of the value of l_{br} , then $K = 99$. I tend to think that K should be significantly greater than 99, but we don't need to decide upon it now to proceed. (3.2) is equivalent¹⁸ to:

$$\log\left(\left(1 + \frac{1}{x}\right)^{1+x}\right) > K \quad (3.3)$$

Exponentiating both sides of the inequality yields.¹⁹

$$\left(1 + \frac{1}{x}\right)^{1+x} > e^K \quad (3.4)$$

The function $f(x) = \left(1 + \frac{1}{x}\right)^{1+x}$ has been studied in great detail. For our purpose, though, we need only note two things. First, $f(x)$ is descending for every $x > 0$. Second, (provided that we take K to be greater than 2, say),²⁰ the inequality holds for $x = 1/e^k$, but fails to hold for $x = 2/e^k$. Therefore, for the inequality in (3.4) to hold, x must be smaller than a number that lies somewhere between $1/e^k$ and $2/e^k$. Therefore, to satisfy D_3 , the ratio $P(br|\sim H)/P(nbr|\sim H)$ must be smaller than $2/e^k$.

Because K is supposed to be a pretty big number, e^K should be gigantic.²¹ This means that for the standard Bayesian solution to satisfy D_3 , we must think in advance

¹⁶ We can think that it's a possibility that if not all ravens are black then none of them are. However, as long as we think that there is a nonnegligible probability that if not all ravens are black then still a significant portion of them are, we expect $\frac{P(br|\sim H)}{P(nbr|\sim H)}$ not to be extremely small.

¹⁷ If we take choice 2, we can rewrite: $l_{nbr's\ per\ br} = \frac{1}{P(H)+x}$.

¹⁸ Because $(1+x)\log\left(1 + \frac{1}{x}\right) = \log\left(\left(1 + \frac{1}{x}\right)^{1+x}\right)$.

¹⁹ If we take choice 2, we get instead that $\left(1 + \frac{1}{x}\right)^{P(H)+x} > e^K$.

²⁰ Or not smaller than about 1.92 to be exact.

²¹ Note that we get from (3.4) that while the cumulative support of the nonblack nonravens is not negligible compared to the cumulative support of the black ravens it still must be smaller. To see that, simply take K to be 1. However, if we had taken choice 2, we could have easily come up with examples of priors according to which the cumulative support of all the nonblack nonravens is greater than that of the black ravens. E.g., take x to be 1 and $P(H)$ to be $1/3$.

Table 3. How small x should be for some chosen values of K

Chosen value of K	The nonblack nonravens contribute at most . . .	x should be at least smaller than . . .	Suppose that $\sim H$ and that we have sampled a raven. This raven being black must be less likely than . . .
19	5%	$\frac{2}{e^{19}} \approx 10^{-8}$	Winning the lottery ²²
99	1%	$\frac{2}{e^{99}} \approx 10^{-44}$	Winning the lottery six times in a row
999	0.1%	$\frac{2}{e^{999}} \approx 10^{-434}$	Randomly sampling the same specific elementary particle from the whole universe five times in a row

that if not all ravens are black, and we sample a raven, then the probability that the raven we sample is black should be extremely tiny. To get a feel of how tiny, look at Table 3. The table states how small x should be for some chosen values of K , that is, how likely it is that a given raven we sample is black, conditional on not all ravens being black. To help the reader get a grasp of the magnitudes, in each row, I gave an example of a roughly equiprobable event with which we are more familiar.

Because ravens are discrete objects and there aren't that many of them, we get that for the standard Bayesian solution to satisfy D_3 , it must be the case that we think in advance that if not *all* ravens are black, then it's overwhelmingly likely that none of them are.²³ However, there is no reason to think that if not all ravens are black, then none of them are.

Furthermore, remember that the color of ravens is just an example. The paradox seemingly applies to many other properties F and G , and there is even less reason to assume in advance that for all these properties F and G , either all the F 's are G 's or, in probability close to certainty, none or only an unimaginably tiny minority of them are. For example, it seems unreasonable to think we are committed not only to thinking in advance that if not all ravens are black then none of them are but also to thinking in advance that if not all parrots are green then none of them are, if not all chairs have three legs then none of them do, if not all oranges are orange then none of them are, and so forth. I conclude that the standard Bayesian solution cannot plausibly satisfy all three desiderata, and if I'm right in saying that both D_2 and D_3 are crucial to answering the paradox, it follows that the standard Bayesian solution fails to address the paradox of the ravens.²⁴

²² The chance of winning the lottery changes from country to country. I take it to be in the order of magnitude of 10^{-7} .

²³ Taking very small values for K like 19 allows one very special raven to be black even if not all of them are, but taking K to have such a small value seems implausible (at least to me).

²⁴ It is known at least since Horwich (1982) that the standard Bayesian solution has the consequence that seeing black nonravens *disconfirms* H (see Rinard 2014). I believe people tend ignore it because the contribution of a single black nonraven is negligible. The cumulative negative contribution of all the black nonravens we see is nonnegligible, though. I thank an anonymous reviewer clarifying the need to mention this.

4. The prospects of any alternative Bayesian solution

At this point, the reader might suspect that the challenge I presented applies only to a specific variant of the Bayesian solution, namely, the one presented by Vranas, and while indeed this is probably the most famous version of a Bayesian solution, it needn't be the only one. What if instead of assumptions A_1 - A_3 , we adopt different assumptions? Can't we then save a Bayesian approach to the paradox of the ravens? My answer is that the prospects for that are not promising, at least when "Bayesian approach" is understood as the strategy of establishing both D_1 and D_2 (i.e., acknowledging that both seeing a black raven and seeing a nonblack nonraven confirm H but a black raven provides stronger support). Making this generalized argument is significantly more complicated than making the preceding argument (though still requires roughly the same level of mathematical proficiency), so I have put the mathematical details in section (iii) of the appendix. Here, I present the general outline of the argument, but I urge the suspicious reader to examine the appendix as well.

The idea of the argument is to examine what follows from adopting the three desiderata D_1 - D_3 , without making any assumptions regarding our background knowledge. That is, recall, we examine what follows from assuming that both seeing a black raven and seeing a nonblack nonraven support the hypothesis that all ravens are black (D_1), but the contribution of both a single nonblack nonraven compared to a single black raven (D_2), and the aggregate contribution of all the nonblack nonravens we are expected to see compared to the aggregate contribution of all the black ravens we are expected to see (D_3) are negligible. We derive that for these three desiderata to hold together, at least one of three conditions must hold, but it is implausible to impose any of these conditions on our priors.

The first condition that allows D_1 - D_3 to hold together states that the ratio $\frac{P(br|H)}{P(br|\sim H)}$ is extremely big. That is, it is overwhelmingly more likely to sample a black raven if all ravens are black than to sample a black raven if not all of them are, or practically, that if not all ravens are black, then none of them are. Again, there is no reason that our priors regarding ravens must be tuned this way, all the more so our priors regarding any properties F and G for which the paradox applies. Note that this is a generalization of the result we got in the previous section. Earlier, we got from A_1 - A_3 that to satisfy D_1 - D_3 , the ratio $\frac{P(br|\sim H)}{P(nbr|\sim H)}$ needs to be extremely small, or equivalently, that $\frac{P(nbr|\sim H)}{P(br|\sim H)}$ needs to be extremely big. It follows from A_1 - A_3 that $\frac{P(br|H)}{P(br|\sim H)} > \frac{P(nbr|\sim H)}{P(br|\sim H)}$ so if $\frac{P(nbr|\sim H)}{P(br|\sim H)}$ is extremely big, then so is $\frac{P(br|H)}{P(br|\sim H)}$.

The second condition that allows D_1 - D_3 to hold together states that $P(br|H) \geq P(nbr|H)$, that is, that if all ravens are black, then sampling a raven is at least as likely as sampling a nonblack object. An implausible condition indeed. This condition comes from the fact that, in the argument presented in the appendix, I assume for convenience that $\frac{P(nbr|H)}{P(br|H)} > 1$, so negating this assumption seemingly lets us evade the conclusion of the argument presented there. In fact, the argument can work with even smaller values than 1,²⁵ but it is not crucial. A solution that requires our priors to be such that $P(br|H) \geq P(nbr|H)$ is a nonstarter.

²⁵ Technically, $\frac{P(nbr|H)}{P(br|H)}$ just needs to be big compared to $\frac{1}{k}$.

The third condition that allows D_1 - D_3 to hold is somewhat more complicated, and results from the conclusion of the argument in the appendix (see proposition iii.16). It states, roughly, that for the ratios $l = \frac{P(br|H)}{P(br|\sim H)}$, $m = \frac{P(nbr|H)}{P(br|H)}$ and $n = \frac{P(nbr|\sim H)}{P(br|\sim H)}$, the following bounds on n must hold:

$$ml - \frac{l}{K} \log(l) < n < ml \quad (4.1)$$

When K is the negligibility constant.

While not easy to see at first glance, this condition is completely implausible. When you look more closely at the inequality in iii.16, you see that the length of the interval of permissible values for n is dependent only on the values of l and K and independent of the value of m . Suppose, like in the appendix, that (say) $l = 10$, $K = 100$. It follows that the length of the interval of permissible values for n is about 0.23. If $m = 1000$, that is, if conditional on all ravens being black, it is 1,000 times more likely to sample a nonblack object than to sample a raven, then n must be in the interval between 10,000 and 10,000 minus 0.23. If $m = 1$ billion then n must lie in the interval between 10 billion and 10 billion minus 0.23. This is too exact to be taken seriously. The fact that what remains invariant for various values of l , m , and n is the absolute length of the interval rather than its proportional length makes it even more bizarre. D_1 - D_3 severely restrict the set of permissible priors, and in a way that is seemingly hopeless to motivate. At the very least, I can't see any way to motivate it. Anyone who argues for a solution that maintains D_1 - D_3 needs to provide us with an argument for why our priors must be so fine-tuned.

I conclude that it is not the specific three assumptions A_1 - A_3 of the standard Bayesian solution that get the Bayesians into trouble. It is the very strategy of taking a nonblack nonraven to support the hypothesis that all ravens are black, but only negligibly, that does.

5. Conclusion

The standard Bayesian solution fails because according to it, the aggregate contribution of all the nonblack nonravens we see must be nonnegligible unless we severely and implausibly restrict the set of permissible priors. However, it is not only the standard Bayesian solution that fails. Any solution that wishes to maintain all of the desiderata D_1 - D_3 needs to severely restrict the set of permissible priors in an unmotivated way. Because D_2 - D_3 are desiderata for any solution to the paradox of the ravens, it follows that for a solution to work, D_1 needs to be false. A nonblack nonraven contributes zero or negatively to the confirmation of the hypothesis that all ravens are black. Because there is no reason to think that a nonblack nonraven should disconfirm this hypothesis, the conclusion is that the contribution of a nonblack nonraven to the hypothesis needs to be zero. For some reason, nonblack nonravens (as such) are irrelevant to the confirmation of the hypothesis that all ravens are black.

It is time to discuss the accusation that I have been too restrictive with my use of the term "the Bayesians." Offering a Bayesian solution to the Paradox of the Ravens, one might protest, does not commit one to adopt D_1 . Rather, it just commits one to solving the paradox using considerations regarding our prior probabilities for sampling black ravens and nonblack nonravens in every scenario.

However, we just saw that any plausible solution to the paradox must take the support provided by a nonblack nonraven to be zero, and only a very specific set of possible prior probability distributions (namely, those in which $P(nbnr|H) = P(nbnr|\sim H)$) give that. There doesn't seem to be any reason our priors must be like that, a fortiori why they should be like that for every F and G to which the paradox applies. There must be something else in play here beyond our prior statistical beliefs that makes it the case that nonblack nonravens are irrelevant (and that non- F 's non- G 's are irrelevant for F 's and G 's to which the paradox applies). The paradox will not be solved by statistical considerations alone.²⁶ If what I said here is correct, the conclusion seems to be that Bayesian considerations alone will not solve the paradox. We need a different strategy.

Acknowledgments. This article was written under the supervision of Tim Williamson. I thank Tim for his invaluable comments on many versions of the manuscript and for connecting me with Branden Fitelson at a stage when the project was a little more than a vague intuition. I am grateful to Branden for all the effort he put into providing me with very useful background and references, and for his comments on an earlier version. I also thank him for prompting me to generalize the argument I originally had and write section 3 of the appendix. Bernhard Salow became my second supervisor at a later stage of the project, and I thank him for his comments and his help at the revision stage. I thank Ittay Nissan-Rosen for a helpful discussion in the very early stages of the project, and Yotam Dikstein for his endless patience and willingness to help whenever I have a question in mathematics, in this article and in general. I am grateful to Paolo Faglia and Dan Grimmer for their comments on earlier drafts and especially for their meticulous review of the mathematical appendices in search of errors. I thank two anonymous reviewers for this journal for their comments and for verifying the mathematical derivations. Of course, I'm solely responsible for any mathematical error in the text, should there be one. I wish to thank the Reuben Foundation and Reuben College at Oxford for their financial support during the time the article was written. My late grandmother, who features in one of the examples in this article, passed away just days after I wrote down this example. I also wish to use the acknowledgments to commemorate this special woman, who, while focusing on human feelings rather than ornithology, was able to make brilliant inferences from very indirect evidence.

Funding Statement. None to declare.

Declarations. None to declare.

References

- Alexander, H. G. 1958. "The Paradoxes of Confirmation." *The British Journal for the Philosophy of Science* 9 (35):227–33. <https://doi.org/10.1093/bjps/IX.35.227>.
- Fitelson, Branden. 2001. "A Bayesian Account of Independent Evidence with Applications." *Philosophy of Science* 68 (S3):S123–40. <https://doi.org/10.1086/392903>.
- Fitelson, Branden, and James Hawthorne. 2010a. "How Bayesian Confirmation Theory Handles the Paradox of the Ravens." In *The Place of Probability in Science*, edited by Ellery Eells and James H. Fetzer, 247–75. *Boston Studies in the Philosophy of Science*, vol. 284. Dordrecht: Springer. https://doi.org/10.1007/978-90-481-3615-5_11.
- Fitelson, Branden, and James Hawthorne. 2010b. "The Wason Task(s) and the Paradox of Confirmation." *Philosophical Perspectives* 24 (1):207–41. <https://doi.org/10.1111/j.1520-8583.2010.00191.x>.
- Fitelson, Branden, and James Hawthorne. 2010c. "An Even Better Solution to the Paradox of the Ravens." http://fitelson.org/wason_addendum.pdf.

²⁶ Rinard (2014) discusses why our priors are and should be such that a nonblack nonraven provides zero support to H . Although she calls her paper "A New Bayesian Solution to the Paradox of the Ravens," she demonstrates my point. There is nothing particularly Bayesian in her solution. What does the work in her solution are metaphysical considerations regarding the hierarchy of natural kinds, rather than statistical considerations. I will not go into the details of her solution here.

Good, I. J. 1967. "The White Shoe Is a Red Herring." *The British Journal for the Philosophy of Science* 17 (4):322–32. <https://doi.org/10.1093/bjps/17.4.322>.

Goodman, Nelson. 1983. *Fact, Fiction, and Forecast*. 4th ed. Cambridge, MA: Harvard University Press.

Hempel, Carl G. 1945. "Studies in the Logic of Confirmation (I)." *Mind* 54 (213):1–26. <https://doi.org/10.1093/mind/LIV.213.1>.

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

Hosiasson-Lindenbaum, Janina. 1940. "On Confirmation." *Journal of Symbolic Logic* 5 (4): 133–48. <https://doi.org/10.2307/2268173>.

Howson, Colin, and Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach*. 2nd ed. Chicago: Open Court.

Huber, Franz. 2024. "Confirmation and Induction." *The Internet Encyclopedia of Philosophy*. ISSN 2161-0002. <https://iep.utm.edu/confirmation-and-induction/#SH6c>.

Mackie, J. L. 1963. "The Paradox of Confirmation." *The British Journal for the Philosophy of Science* 13 (52):265–76. <https://doi.org/10.1093/bjps/XIII.52.265>.

Rinard, Susanna. 2014. "A New Bayesian Solution to the Paradox of the Ravens." *Philosophy of Science* 81 (1):81–100. <https://doi.org/10.1086/674202>.

Vranas, Peter B. 2004. "Hempel's Raven Paradox: A Lacuna in the Standard Bayesian Solution." *The British Journal for the Philosophy of Science* 55 (3):545–60. <https://doi.org/10.1093/bjps/55.3.545>.

Appendix

(i) Calculating the cumulative contribution of many *nbnr*'s

In this section of the appendix, we calculate the cumulative support of seeing $\frac{P(nbnr|H)}{P(br|H)}$ nonblack nonravens. $\frac{P(nbnr|H)}{P(br|H)}$ is not necessarily an integer, but is assumed to be very big, so taking the nearest integer instead should provide practically the same result. We can calculate the degree of confirmation provided by $\frac{P(nbnr|H)}{P(br|H)}$ nonblack nonravens as follows (see 2.6 and 2.7):

$$\begin{aligned} I_{nbnr's \text{ per } br} &= \log \left(\frac{P(nbnr|H) \frac{P(nbnr|H)}{P(br|H)}}{P(nbnr| \sim H) \frac{P(nbnr|H)}{P(br|H)}} \right) = \log \left(\left(\frac{P(nbnr|H)}{P(nbnr| \sim H)} \right)^{\frac{P(nbnr|H)}{P(br|H)}} \right) \\ &= \log \left(\left(\frac{1 + P(nbr| \sim H)}{P(nbnr| \sim H)} \right)^{\frac{P(nbnr| \sim H) + P(nbr| \sim H)}{P(br| \sim H) + P(nbr| \sim H)}} \right) \end{aligned} \tag{i.1}$$

This expression resembles another, very famous expression. It is well known that $\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$, and more generally that:

$$\lim_{x \rightarrow \infty} \left(1 + a \frac{1}{x}\right)^x = e^a \tag{i.2}$$

We know (from 2.3) that $\frac{P(nbnr| \sim H)}{P(br| \sim H) + P(nbr| \sim H)}$ is very large and thus $\frac{P(nbr| \sim H)}{P(br| \sim H) + P(nbr| \sim H)}$ only adds a negligible contribution to the term. The expression in (i.1) can be made very similar to the one in (i.2) if in the expression $\frac{P(nbr| \sim H)}{P(nbnr| \sim H)}$ in (i.1) we multiply both the numerator and the denominator by the factor $P(br| \sim H) + P(nbr| \sim H)$:

$$\begin{aligned}
& \log \left(1 + \frac{P(nbr|\sim H)}{P(nbr|\sim H)} \frac{P(nbr|\sim H)+P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)} \right) \\
&= \log \left(1 + \frac{P(nbr|\sim H)}{P(br|\sim H) + P(nbr|\sim H)} \frac{P(br|\sim H) + P(nbr|\sim H)}{P(nbr|\sim H)} \frac{P(nbr|\sim H)+P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)} \right) \quad (i.3) \\
&\approx \log \left(1 + \frac{P(nbr|\sim H)}{P(br|\sim H) + P(nbr|\sim H)} \frac{P(br|\sim H) + P(nbr|\sim H)}{P(nbr|\sim H)} \frac{P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)} \right)
\end{aligned}$$

Because $\frac{P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)}$ is much smaller than $\frac{P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)}$, we can now use (i.2) to approximate the last expression in (i.3) as:

$$\log \left(e^{\frac{P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)}} \right) \quad (i.4)$$

And this last term in (i.4) is equal to:

$$\log \left(e^{\frac{P(nbr|\sim H)}{P(br|\sim H)+P(nbr|\sim H)}} \right) = \frac{P(nbr|\sim H)}{P(br|\sim H) + P(nbr|\sim H)} = \frac{1}{\frac{P(br|\sim H)}{P(nbr|\sim H)} + 1} \quad (i.5)$$

In the end, we get that:

$$l_{nbr's \text{ per } br} \approx \frac{1}{\frac{P(br|\sim H)}{P(nbr|\sim H)} + 1} \quad (i.6)$$

Which is what we wanted to demonstrate.

(ii) Sampling without replacement

In this section of the appendix, I show that given two reasonable assumptions, modeling the situation as sampling without replacement does not cause significant divergence from modeling the situation as sampling with replacement. I present only the case for the cumulative contribution of all black ravens the subject sees. The case for the cumulative contribution of all nonblack nonravens the subject sees is analogous.

The first assumption is that the subject believes that the number of objects in the universe depends only negligibly on whether all ravens are black or not. That is, if N_H stands for the believed number of objects in the universe conditional on H , and $N_{\sim H}$ is the believed number of objects in the universe conditional on $\sim H$, then for all practical purposes, we can assume that $N_H = N_{\sim H} = N$.²⁷ Of course, the subject might assign prior probabilities to various values of N , but we assume that for each specific value of N , the probability for this value of N is, for all practical purposes, independent of H . Thus, we can treat N as a specific number, and the argument generalizes for a probability distribution for values of N .

Suppose then that the subject's prior belief is that there are N objects in the universe. Let K_H stand for the number of black ravens the subject believes there are in the universe in case H is true, and let $K_{\sim H}$ stand for the number of black ravens the subject believes there are in the universe in case H is false (note, K_H and $K_{\sim H}$ have nothing to do with the negligibility constant K introduced previously). We can express

²⁷ This assumption significantly simplifies the calculation but can be weakened. See the next footnote.

the subject prior conditional probability for sampling a black raven on the first draw as $P(br|H) = \frac{K_H}{N}$, $P(br|\sim H) = \frac{K_{\sim H}}{N}$. The log-likelihood ratio before the first draw (i.e., after 0 draws) is:

$$l_{br}(0) = \log\left(\frac{K_H/N}{K_{\sim H}/N}\right) = \log\left(\frac{K_H}{K_{\sim H}}\right) \tag{ii.1}$$

Note that according to D_1 , the contribution of seeing a black raven should be positive, so K_H should be greater than $K_{\sim H}$.²⁸ Let n stand for the number of objects sampled so far, and k stand for the number of black ravens sampled so far. In general, the log likelihood ratio of seeing one black raven after seeing n objects and k black ravens without replacement is $\log\frac{(K_H-k)/(N-n)}{(K_{\sim H}-k)/(N-n)}$. The $N-n$ cancels out, so the contribution of seeing each new black raven is a function of k alone. That is, the contribution of seeing a black raven after sampling k black ravens (i.e., the contribution of the $(k+1)$ th black raven) is:

$$l_{br}(k) = \log\frac{K_H - k}{K_{\sim H} - k} \tag{ii.2}$$

Note that for each k , $l_{br}(k)$ can be decomposed to:²⁹

$$l_{br}(k) = \log\frac{K_H}{K_{\sim H}} + \log\frac{\left(1 - \frac{k}{K_H}\right)}{\left(1 - \frac{k}{K_{\sim H}}\right)} \tag{ii.3}$$

The left-hand side is just $l_{br}(0)$ and the right-hand side is a correction term that is a function of k . Let's call it $Corr(k)$.

$$Corr(k) = \log\frac{\left(1 - \frac{k}{K_H}\right)}{\left(1 - \frac{k}{K_{\sim H}}\right)} \tag{ii.4}$$

That is, $l_{br}(k) = l_{br}(0) + Corr(k)$. The cumulative contribution of seeing k black ravens *with* replacement is just $kl_{br}(0)$. *Without* replacement, the cumulative contribution of seeing k black ravens is:³⁰

²⁸ Instead of assuming that $N_H = N_{\sim H} = N$, we can alternatively assume that the difference between N_H and $N_{\sim H}$, i.e., $|N_H - N_{\sim H}|$ is negligible compared to both N_H and $N_{\sim H}$, and that $\log\frac{N_H}{N_{\sim H}}$ is negligible compared to $\log\frac{K_H}{K_{\sim H}}$. The last assumption says in effect that the support provided to H by seeing a black raven is not due to considerations involving the difference in the total number of objects in the universe given H and given $\sim H$. The calculations lead to the same result and don't reveal anything deep as far as I can tell.

²⁹ $l_{br}(k) = \log\frac{K_H-k}{K_{\sim H}-k} = \log\frac{K_H\left(1-\frac{k}{K_H}\right)}{K_{\sim H}\left(1-\frac{k}{K_{\sim H}}\right)} = \log\frac{K_H}{K_{\sim H}} + \log\frac{\left(1-\frac{k}{K_H}\right)}{\left(1-\frac{k}{K_{\sim H}}\right)}$.

³⁰ Note that if we assume that $N_H = N_{\sim H}$, it follows from $K_H > K_{\sim H}$ that there believed to be more objects that are not black ravens conditional on $\sim H$ than conditional on H . Thus, if all we know about an object is that it is not a black raven, sampling it is evidence against H , and sampling a great deal of them disconfirm H nonnegligibly. I thank an anonymous reviewer for making me clarify that.

$$\sum_0^{k-1} l_{br}(0) + \text{Corr}(k) = kl_{br}(0) + \sum_0^{k-1} \text{Corr}(k) \quad (\text{ii.5})$$

The second assumption I'm making is that the total number of black ravens the subject sees, call it k_0 , is small compared to the total number of black ravens, whether all ravens are black or not. That is, $k_0 \ll K_{\sim H} (< K_H)$. If the subject sees k_0 black ravens, the divergence between their cumulative contribution with and without replacement is $\text{Div}(k_0) = \sum_{k=0}^{k_0-1} \text{Corr}(k)$. If k_0 is small compared to $K_{\sim H}$, then for every $k < k_0$, $\text{Corr}(k)$ can be approximated by:³¹

$$\text{Corr}(k) \approx \frac{k}{K_{\sim H}} - \frac{k}{K_H} \quad (\text{ii.6})$$

Therefore, the total divergence after seeing k_0 black ravens is:³²

$$\text{Div}(k_0) \approx \left(\frac{k_0(k_0 - 1)}{2} \right) \left(\frac{1}{K_{\sim H}} - \frac{1}{K_H} \right) \quad (\text{ii.7})$$

To understand how big the divergence is, we need to compare $\text{Div}(k_0)$ with the cumulative contribution *with* replacement, $k_0 l_{br}(0)$. To do that, note that $l_{br}(0) = \log \frac{K_H}{K_{\sim H}}$ can be rewritten as $l_{br}(0) = \log \left(1 + \frac{K_H - K_{\sim H}}{K_{\sim H}} \right)$. Define: $\frac{K_H - K_{\sim H}}{K_{\sim H}} = a$. Note that $K_H = (a + 1)K_{\sim H}$. We can then rewrite $k_0 l_{br}(0)$ as $k_0 \log(1 + a)$ and rewrite $\text{Div}(k_0)$ as $\text{Div}(k_0) \approx \left(\frac{k_0(k_0 - 1)}{2} \right) \left(\frac{a}{(a + 1)K_{\sim H}} \right)$. We can then express the ratio between $\text{Div}(k_0)$ and $k_0 l_{br}(0)$ as:

$$\frac{\text{Div}(k_0)}{k_0 l_{br}(0)} \approx \frac{a}{(1 + a) \log(1 + a)} \frac{k_0 - 1}{2K_{\sim H}} \quad (\text{ii.8})$$

Because $0 < \frac{a}{(1+a) \log(1+a)} < 1$ for every $a > 0$ and $k_0 \ll K_{\sim H}$, we get that $\text{Div}(k_0)$ is small compared to $k_0 l_{br}(0)$. That is, given the assumptions that the subject takes the total number of objects in the universe to be independent from the hypothesis that all ravens are black and that the subject only sees a small subset of all black ravens, the divergence between modeling the situation as sampling with replacement and sampling without replacement is small. Therefore, modeling the situation as sampling without replacement is unlikely to change the overall verdict.

(iii) D_I - D_3 without A_I - A_3

Suppose we don't adopt any assumptions about our background knowledge, but we want all the desiderata D_I - D_3 to hold. These three desiderata, it turns out, are only consistent with a narrow set of possible priors, and I don't think there is any independent argument for why we need to have these specific priors. Let's see why.

³¹ $\text{Corr}(k) = \log \left(\frac{1 - \frac{k}{K_H}}{1 - \frac{k}{K_{\sim H}}} \right) = \log \left(1 - \frac{k}{K_H} \right) - \log \left(1 - \frac{k}{K_{\sim H}} \right) \approx \frac{k}{K_{\sim H}} - \frac{k}{K_H}$. Note that for every $x > 0$, $x > \log(1 + x)$, so the approximation leads to overestimation of $\text{Corr}(k)$.

³² $\sum_{k=0}^{k_0-1} \text{Corr}(k) \approx \sum_{k=0}^{k_0-1} \frac{k}{K_{\sim H}} - \frac{k}{K_H} = \frac{1}{K_{\sim H}} \sum_{k=0}^{k_0-1} k - \frac{1}{K_H} \sum_{k=0}^{k_0-1} k = \frac{k_0(k_0-1)}{2} \left(\frac{1}{K_{\sim H}} - \frac{1}{K_H} \right)$.

As we saw previously, the degree of confirmation provided by a single black raven is $\log\left(\frac{P(br|H)}{P(br|\sim H)}\right)$ and the degree of confirmation provided by a single nonblack nonraven is $\log\left(\frac{P(nbnr|H)}{P(nbnr|\sim H)}\right)$. The first and the second desiderata imply together that:³³

$$\log\left(\frac{P(br|H)}{P(br|\sim H)}\right) > \log\left(\frac{P(nbnr|H)}{P(nbnr|\sim H)}\right) > 0 \tag{iii.1}$$

Or, equivalently, that:

$$\frac{P(br|H)}{P(br|\sim H)} > \frac{P(nbnr|H)}{P(nbnr|\sim H)} > 1 \tag{iii.2}$$

To simplify, let's say (as in section 4) that $\frac{P(nbnr|H)}{P(br|H)} = m$ and that $\frac{P(nbnr|\sim H)}{P(br|\sim H)} = n$. We can then infer from (iii.2) that:

$$n = \frac{P(nbnr|\sim H)}{P(br|\sim H)} > \frac{P(nbnr|H)}{P(br|H)} = m \tag{iii.3}$$

And that:

$$\frac{P(br|H)}{P(br|\sim H)} > \frac{mP(br|H)}{nP(br|\sim H)} > 1 \tag{iii.4}$$

We tend to think that both m and n need to be very big, but we won't need this assumption for what follows. All we will need is that m won't be very small. It is useful though to think of m as at least greater than 1, that is, that according to our priors, given that all ravens are black, it is at least as likely to sample a nonblack nonraven as it is to sample a black raven (which, conditional on all ravens being black, is equivalent to the claim that it is at least as likely to sample a nonblack object as it is to sample a raven). What I say in the text that follows can work with values of m that are even smaller than 1, but in any case, any solution to the paradox of the ravens should be compatible with m being more than 1. The ratio $\frac{P(br|H)}{P(br|\sim H)}$ will be used many times in what follows. For convenience, let's substitute it with the single letter l (for "likelihood"), that is, $\frac{P(br|H)}{P(br|\sim H)} = l$.³⁴ Note that it follows from (4) that $\frac{m}{n}l > 1$, or, equivalently, that

$$\frac{n}{m} < l \tag{iii.5}$$

Note that (iii.5) is in effect just another way to say that $P(nbnr|\sim H) < P(nbnr|H)$, or that $\frac{P(nbnr|H)}{P(nbnr|\sim H)} > 1$, as we already know from (iii.2).

Let's turn now to D_3 . If H is true, we expect to see m nonblack nonravens for each black raven. If H is false, we expect to see n nonblack nonravens for each black raven. The expected value for the number of nonblack nonravens we see per each black raven is $mP(H) + nP(\sim H)$, which is greater than m , because $n > m$. Therefore, taking

³³ Actually, D_2 tells us that $\log\left(\frac{P(nbnr|H)}{P(nbnr|\sim H)}\right)$ is not only smaller than $\log\left(\frac{P(br|H)}{P(br|\sim H)}\right)$ but also that it is negligible compared to it, but it won't be important for what follows.

³⁴ Note that earlier we denoted $\log\left(\frac{P(br|H)}{P(br|\sim H)}\right)$ by l_{br} . Therefore, in our new notation we get that $l_{br} = \log l$.

m as the expected value of the number of nonblack nonravens we see per each black raven is again generous toward the Bayesians. To satisfy D_3 , we need that:

$$\log\left(\frac{P(br|H)}{P(br|\sim H)}\right) > K \log\left(\frac{P(nbnr|H)}{P(nbnr|\sim H)}\right)^m \quad (\text{iii.6})$$

Or, equivalently:

$$\log(l) > K \log\left(\frac{m}{n}\right)^m = mK \log\left(\frac{m}{n}\right) = mK \left(\log\left(\frac{m}{n}\right) + \log(l)\right) \quad (\text{iii.7})$$

Where K is again a negligibility constant that is supposed to be pretty big. Rearranging (iii.7) gives us:

$$(mK - 1) \log(l) < -mK \log\left(\frac{m}{n}\right) = mK \log\left(\frac{n}{m}\right) \quad (\text{iii.8})$$

Which in turn gives:

$$\frac{mK - 1}{mK} \log(l) = \left(1 - \frac{1}{mK}\right) \log(l) = \log(l^{1 - \frac{1}{mK}}) < \log\left(\frac{n}{m}\right) \quad (\text{iii.9})$$

Exponentiating (iii.9) gives:

$$\frac{n}{m} > (l)^{1 - \frac{1}{mK}} \quad (\text{iii.10})$$

(iii.5) and (iii.10) together yield:

$$l^{1 - \frac{1}{mK}} < \frac{n}{m} < l \quad (\text{iii.11})$$

To be more systematic, we can separate n and rewrite (iii.11) as:

$$m^{1 - \frac{1}{mK}} < n < ml \quad (\text{iii.12})$$

That is, to answer desiderata D_1 - D_3 , our prior value for n (i.e., recall, our prior value for the ratio $\frac{P(nbnr|\sim H)}{P(br|\sim H)}$) needs to be bounded within an interval, which in cases in which K is big, and m is big compared to $\frac{1}{K}$, is a very narrow one. If m is significantly bigger than $\frac{1}{K}$ then the difference between l and $l^{1 - \frac{1}{mK}}$ is tiny. To get a feel of how narrow the interval between them should be, suppose (say) that $l = 10$, $K = 100$, and $m = 1000$. In such a case, we get that $9,999.77 \approx 1000 \cdot 10^{\frac{99,999}{100,000}} < n < 1000 \cdot 10 = 10,000$.

The difference between the expression on the right-hand side of (iii.12) and the expression on the left-hand side of (iii.12) is $ml - m^{1 - \frac{1}{mK}} = m(l - l^{1 - \frac{1}{mK}})$. If m is significantly bigger than $\frac{1}{K}$, then $\frac{1}{mK}$ is a pretty small number. Therefore, $l - l^{1 - \frac{1}{mK}}$ can be approximated using the Taylor expansion near zero of l^{1-x} where we take $x = \frac{1}{mK}$ as our variable. (iii.12) presents the Taylor expansion of l^{1-x} near $x = 0$:

$$\begin{aligned} l^{1-x} &= e^{\log(l)(1-x)} \\ &= e^{\log(l)(1-0)} - \log(l)e^{\log(l)(1-0)}x + \frac{1}{2}\log^2(l)e^{\log(l)(1-0)}x^2 + \dots \\ &= l - \log(l) \cdot lx + \frac{1}{2}\log^2(l) \cdot lx^2 + \dots \approx l - \log(l) \cdot l \cdot \frac{1}{mK} \end{aligned} \quad (\text{iii.13})$$

The approximation at the end of (iii.13) is derived from the fact that because $x = \frac{1}{mK}$ is close to zero, the higher powers of x are very small compared to x (and we assume that l is not gigantic enough for the coefficients to compensate). It follows from the approximation in (iii.12) that:

$$l - l^{1-\frac{1}{mK}} \approx \log(l) \cdot l \cdot \frac{1}{mK} \tag{iii.14}$$

Therefore,

$$m(l - l^{1-\frac{1}{mK}}) \approx m\log(l) \cdot l \cdot \frac{1}{mK} = \frac{l}{K}\log(l) \tag{iii.15}$$

That is, (iii.12) can be rewritten approximately as:

$$ml - \frac{l}{K}\log(l) < n < ml \tag{iii.16}$$

What can be learned from this? First, we see that the length of the interval of permissible values for n is approximately independent of m . In the preceding case, where $l = 10$, $K = 100$, it always has a length of about 0.23. That is, if (say) $m = 2$, then n must lie in the interval between 20 (i.e., ml) and 20 minus 0.23. If $m = 1000$ (as we saw in the preceding text), n must lie in the interval between 10,000 and 10,000 minus 0.23. If $m = 1$ billion then n must lie in the interval between 10 billion and 10 billion minus 0.23.

Second, it's reasonable to assume that K is determined independently of the specifics of the setting of the paradox. K , after all, is a negligibility constant, that is, it should determine only when the contribution of a piece of evidence e is negligible. It should have nothing to do specifically with our priors regarding ravens. Therefore, it seems that l is the only flexible value in $\frac{l}{K}\log(l)$, and we can see that $\frac{l}{K}\log(l)$ becomes bigger the bigger l is. This is in accord with what we have found about the canonical Bayesian solution. l , recall, is the ratio $\frac{P(br|H)}{P(br|\sim H)}$, and if l is very big, it implies that the probability of sampling a black raven in case H is false is tiny compared to the probability of sampling a black raven in case H is true. However, for reasonable values of l , we get that the interval of possible values for n is very narrow and independent of the probabilities of sampling a nonblack nonraven or a black raven.

Another point to note is that when we unpack the notation in the inequality (iii.16) back to our original variables, we see that it can be rewritten as:

$$P(nbnr|H) - \frac{P(br|H)}{K} \log\left(\frac{P(br|H)}{P(br|\sim H)}\right) < P(nbnr|\sim H) < P(nbnr|H) \tag{iii.17}$$

That is, while greater l indeed implies a greater length of the interval of possible values for $(nbnr|\sim H)$, because in any case $P(br|H) < 1$, and therefore in any case $\frac{P(br|H)}{K} < \frac{1}{K}$, the growth in l contributes almost entirely *logarithmically* to the growth of the interval of possible values for $P(nbnr|\sim H)$. That is, to widen the interval of possible values for $P(nbnr|\sim H)$ just by a tiny bit, l will need to grow by a great deal. If l is not too big, then as I explain in section 4, this result poses an implausible and somewhat bizarre restriction on the set of permissible priors.