

# *The Thomas-Nolt Dispute: Some Lessons about Induction*

DAVID HITCHCOCK

*McMaster University*

**Abstract:** I resolve an apparently unresolved dispute about how probable uniform experience makes an extrapolation from it, and draw some general lessons about such enumerative induction. Uniform experience does not necessarily confer a high probability on an extrapolation of or generalization from that experience. Rational extrapolation or generalization typically involves a lot of specific background information, though not necessarily a general assumption that nature is uniform or that the future will resemble the past. And new evidence which is highly likely on one hypothesis but highly unlikely on any of its competitors does not necessarily make the former hypothesis highly probable.

**Résumé:** Je résous une controverse apparemment irrésolue concernant la probabilité des extrapolations fondées sur l'uniformité de l'expérience, et je tire quelques conclusions générales sur l'induction énumérative. L'uniformité des événements n'accorde pas nécessairement une haute probabilité à une extrapolation ou à une généralisation basée sur cette uniformité. Une extrapolation ou une généralisation rationnelle s'étaye typiquement sur beaucoup d'information spécifique de fond, mais ne s'appuie pas nécessairement sur une supposition que la nature est uniforme ou que l'avenir ressemblera au passé. Et lorsqu'une hypothèse rend très probables des nouvelles preuves et que des hypothèses qui lui font concurrence rendent peu probable ces mêmes preuves, celle-là n'est pas nécessairement très probable.

**Keywords:** Induction, enumerative induction, inductive extrapolation, inductive extrapolation, inductive generalization, Bayes' theorem, epistemic probability, Stephen Thomas, John Nolt.

Consider balls drawn from an urn at random. Suppose there are fifty balls in the urn, and that the first forty-nine, drawn at random, all have been blue. . . . If one calculates the probability that the remaining ball is blue given that the first forty-nine drawn at random have been blue, the probability is well in excess of 80%. (Thomas 1984: 32)

This claim [that the probability is well in excess of 80%] seems plausible, but in fact there is no way to calculate such a probability from the information Thomas gives. . . . if we do not know the number of blue balls in the urn initially and we make only the assumptions that Thomas gives, then no calculation will yield the probability that the remaining ball is blue. We can, without violating any mathematical law, assign that proposition any probability we like. (Nolt 1985: 56)

This dispute apparently remains unresolved: Thomas in the latest edition of his textbook (1997: 131) continues to claim a high conditional probability for a similar case. Though the example is artificial, it raises a fundamental question about the legitimacy of all extrapolation and generalization from uniform experience. I propose therefore to resolve the dispute, and to draw from this resolution some general lessons about inductive reasoning.

### 1. Analysis of the example

We can set out the example in the form of an argument:

There were fifty balls in the urn.  
 The first forty-nine, drawn at random, all have been blue.  
 Therefore, probably, the remaining ball is blue.

The dispute concerns how probable the premisses of this argument make its conclusion.

The probability in question cannot be construed, at least in any obvious way, as any sort of relative frequency, either actual or hypothetical, since it does not concern the probability that an arbitrarily chosen member of a class will have a certain property. It rather concerns the probability, relative to specified information, that a definite proposition is true, that the remaining ball in the urn is blue. This is an epistemic probability, the degree of confidence that it is reasonable to attach to a proposition given certain information. An initial approximation to understanding this kind of probability is to think of it in terms of the fair betting odds that the proposition is true. Given that the first 49 balls selected from the urn are blue, what would be fair odds if one person bet another person a dollar that the remaining ball will also be blue? If it is fair to accept such a bet at even odds, then the rational degree of confidence, given this information, that the remaining ball is blue is .5, or 1/2. In general, if the fair odds, given background information *K*, on a bet that a proposition is true are  $x:y$ , then the rational degree of confidence in the proposition is roughly  $x \div (x + y)$ . To make the rational degree of confidence in a proposition an exact function of the fair odds in a bet that it is true, however, we have to make some patently absurd assumptions: that the fairness of the odds is not influenced by the potential bettor's enjoyment of or aversion to gambling, that the only thing of value is the currency of the potential bet, that the value of that currency increases exactly proportionately with the amount, and so on. The absurdity of such assumptions vitiates the usual Dutch book argument that the degrees of confidence it is rational to assign to propositions must conform to the axioms of the classical probability calculus. For more sophisticated arguments, see Kaplan (1996), Ramsey (1990/1925-29) and Savage (1972/1954). On the basis of these more sophisticated arguments, I shall assume in what follows that rational degree of confidence is a probability function in the sense of the classical probability calculus. This assumption is compatible with various theories about rational degree of confidence, e.g., that it is a logical probability (Carnap 1962/

1950), a rationally constrained subjective degree of confidence (Howson and Urbach 1989, Kaplan 1996), or a mixed physical and epistemic probability (Pollock 1990).

In form, the argument based on Thomas' example is what we might call an *inductive extrapolation*; it extrapolates a property found in all observed members of a class to an unobserved member of that class. Russell (1948: 401) called such reasoning "particular induction by simple enumeration". In substance, however, since there is only one unobserved member left, the argument is an *inductive generalization*; it amounts to generalizing from the presence of a property in all observed members of a class to its presence in all members of the class. Russell (1948: 401) called such reasoning "general induction by simple enumeration".

To prove that the conditional probability at issue is the same regardless of whether we express the conclusion as a singular statement or a universal generalization, we can appeal to the definition of conditional probability: the probability that a hypothesis  $H$  is true given certain evidence  $E$ , written " $p(H|E)$ ", is the result of dividing the probability that both the hypothesis and the evidence are true (" $p(H \& E)$ ") by the probability that the evidence obtains (" $p(E)$ "), provided that this latter probability is not 0; if  $p(E) = 0$ , then the conditional probability is undefined.<sup>1</sup> In what follows, I shall assume, since we are given that  $E$  actually occurs, that  $p(E) \neq 0$ . That is, it was not rational to have absolute confidence in advance that  $E$  would not occur, given that it actually did occur; a rational thinker would never think it fair, I suppose, to accept infinite odds against a proposition which later turns out to be true.

If we let " $E$ " stand for the premisses of our example, then the question is whether  $p(\text{remaining ball is blue}|E) = p(\text{all 50 balls blue}|E)$ . Applying the definition of conditional probability, this amounts to the question whether  $p(\text{remaining ball is blue} \& E) \div p(E) = p(\text{all 50 balls blue} \& E) \div p(E)$ . Since  $p(E) \neq 0$ , our question is whether  $p(\text{remaining ball is blue} \& E) = p(\text{all 50 balls blue} \& E)$ . But "remaining ball is blue &  $E$ " entails "all 50 balls blue &  $E$ ", and vice versa; the two conjunctions are logically equivalent. Assuming that logically equivalent propositions have the same probability, we can conclude that the conditional probability at issue is the same regardless of whether we express the conclusion as a singular statement or a universal generalization. In what follows I shall take the conclusion to be a universal generalization.

## 2. Preliminary solution

Bayes' theorem allows us to calculate the epistemic probability that a hypothesis  $H$  is true given certain new evidence  $E$ , a probability generally referred to as the *posterior probability*, provided we are given three other epistemic probabilities, each construed as a rational degree of confidence in a proposition. First, we need the *prior probability* of the hypothesis, that is, the probability that the hypothesis is true, given our background information independently of the new evidence. (I shall call this " $p(H|K)$ ", where  $p$  is the probability function,  $H$  is the hypothesis and

K is our background information apart from the new evidence. The posterior probability we are looking for is thus given by “ $p(H|E \ \& \ K)$ ”, where E is the new evidence.) Second, we need the *posterior likelihood*, the likelihood of the evidence on the assumption that the hypothesis is true, again assuming the same background information which we have independently of the new evidence. (I shall call this “ $p(E|H \ \& \ K)$ ”, where E is the new evidence.) Third, we need the *prior likelihood* of the evidence, the likelihood that the evidence is true on the assumption of our background information, without assuming the truth of the hypothesis under investigation. (I shall call this “ $p(E|K)$ ”.) Bayes’ theorem tells us that, if the prior likelihood is not zero, the posterior probability of a hypothesis on new evidence is its prior probability multiplied by the ratio of the posterior likelihood of the evidence to its prior likelihood:

$$p(H|E \ \& \ K) = p(H|K) \times p(E|H \ \& \ K) \div p(E|K).$$

The proof of the theorem rests on the above-mentioned definition of a conditional probability  $p(A|B)$  as the result of dividing the probability that both A and B obtain by the probability that B obtains, provided that this latter probability is not zero. If one replaces the conditional probabilities in Bayes’ theorem according to this definition, one sees that the theorem is correct, provided that the prior likelihood of the evidence [ $p(E|K)$ ] is not zero.

Our hypothesis H is that all the balls in the urn are blue. The new evidence E is that there were 50 balls in the urn and the first 49 drawn at random from the urn are blue. The relevant background information K is difficult to specify completely, but it would include knowledge about the general properties of balls (that they retain their individuality over time, neither evaporating like mothballs nor merging like drops of water; that they do not spontaneously change colour as chameleons do; etc.), as well as assumptions about the particular situation which are not explicit in Thomas’ description of it (that the balls drawn from the urn were not put back in it, that nobody else put balls in the urn or took them out after the drawing of the 49 blue balls began, that nobody came along after the 49 balls were drawn and repainted the remaining one, etc.). It is important to realize that such background information is always present when we extrapolate or generalize from instances; philosophical discussions of the problem of justifying induction distort the practice when they treat it as if it were just a matter of extrapolating or generalizing from observed instances, without any other information.

Since the evidence of the first 49 balls drawn being blue is a logical consequence of the hypothesis that all 50 balls are blue (given implicit background assumptions such as those mentioned in the preceding paragraph), the posterior likelihood of the evidence is 1.<sup>2</sup> Hence, in this case the posterior probability of the hypothesis will simply be the prior probability of the hypothesis divided by the prior likelihood of the evidence.

$$1. p(\text{all 50 balls blue} | \text{first 49 blue \& K}) = p(\text{all 50 balls blue} | \text{K}) \div p(\text{first 49 blue} | \text{K})$$

Of course, we are not given information which would allow us to assign definite numbers to this prior probability and prior likelihood. What can we do?

One strategy is to notice that, once 49 blue balls have been drawn from the urn, there is only one alternative hypothesis about the colour of the 50 balls which has not been refuted. That is the hypothesis that 49 balls in the urn were blue and the other one was not blue. The posterior likelihood of the evidence, given this hypothesis and background information, is 1/50. For, given the fact that the 49 balls were drawn at random, and assuming that exactly 49 of the 50 balls originally in the urn were blue, there was a probability of 49/50 that the first ball drawn would be blue, of 48/49 that the second ball would be blue given that the first one was blue, and so on, up to a probability of 2 that the 49th ball would be blue given that the first 48 drawn were blue. These 49 probabilities are probabilities of independent events: that the first ball drawn is blue is for example independent of the event that the second ball drawn is blue given that the first one is blue. Hence, by the multiplication rule of the classical probability calculus for independent events, the probability of the first 49 balls being drawn at random being blue is the product of these 49 probabilities, or 1/50. This is of course a frequency probability, but it can be used to justify an epistemic probability of the same magnitude in a step of direct inference (Pollock 1990, Bacchus 1990). Hence the posterior probability that 49 of the 50 balls are blue, given relevant background information along with the evidence that the first 49 balls drawn at random were blue, is 1/50 of the prior probability of the hypothesis divided by the prior likelihood of the evidence. That is,

$$2. p(49 \text{ balls blue} | \text{first 49 blue \& K}) = p(49 \text{ balls blue} | \text{K}) \div 50p(\text{first 49 blue} | \text{K}).$$

But the evidence, along with the background information, entails that either all 50 balls in the urn were blue or 49 of them were. Hence, by the reasoning of note 2, we can conclude that:

$$p(\text{all 50 balls blue or 49 balls blue} | \text{first 49 blue \& K}) = 1.$$

The classical probability calculus tells us that the probability of a disjunction of two mutually exclusive propositions is the sum of the probabilities of the two disjuncts. And the propositions that all 50 balls in the urn are blue and that 49 of them are blue are mutually exclusive. Hence:

$$3. p(\text{all 50 balls blue} | \text{first 49 blue \& K}) + p(49 \text{ balls blue} | \text{first 49 blue \& K}) = 1.$$

Applying some algebraic transformations to the above three equations, we can express the posterior probability as a function of the ratio  $r$  of the prior probability that all 50 balls are blue to the prior probability that 49 balls are blue:

$$4. p(\text{all 50 balls blue} | \text{first 49 blue \& K}) = 50r \div (50r + 1), \text{ provided that } p(49 \text{ blue balls blue} | \text{K}) \neq 0.^3$$

Alternatively, by reasoning parallel to that of note 3, we can express the posterior probability of the 50 blue balls hypothesis as a function of the inverse ratio  $r^1$  of the prior probability that 49 balls are blue to the prior probability that all 50 balls are blue:

5.  $p(\text{50 blue balls}|\text{first 49 balls blue \& K}) = 50 \div (50 + r^1)$ , provided that  $p(\text{all 50 balls blue}|\text{K}) \neq 0$ .

### 3. A logical approach

Of course, we are not given the ratio of the prior probabilities of our two remaining hypotheses, let alone absolute values of each prior probability. We might try, however, to calculate the ratio on the basis of an a priori weighting of the logical possibilities. Prior to any inspection of the 50 balls in the urn, the information given by Thomas, along with our background information, leaves open all the logical possibilities about their colours. Each ball may be any possible colour, or any combination of those possible colours. Putting together such a possibility for each of the 50 balls gives rise to what Carnap (1962/1950: 70-72) called a “state description”, which in this case would be a conjunction of 50 singular propositions postulating a particular colour or combination of colours for each of the 50 balls. (For simplicity, we ignore the other components of a complete state description, on the ground that they are irrelevant to the present problem.) Carnap proposed to assign logically prior probabilities not to those state descriptions but to structural descriptions which each covered all the structurally isomorphic state descriptions; for example the structural description “49 balls are blue and one ball is red” covers 50 state descriptions, each of which makes a different one of the 50 balls the red one. The hypothesis that 49 balls are blue and one is non-blue is in fact a conjunction of structural descriptions, one for each of the colours other than blue. How many colours other than blue are there? Since there are many shades of blue, the logical parallels to blue are general colours like red, yellow, orange, green, purple, brown, black, grey, white, transparent. Let us take this list as the complete list of alternatives to blue, while recognizing some arbitrariness in where the lines are drawn between one colour and another and how many colours to put on the list. This leaves 10 structural descriptions in which 49 balls are blue and the other is a single colour other than blue. But notice that our background information does not exclude the possibility that a ball in the urn is multi-coloured; any ball may have two, three, four, up to 11 colours. Adding in these possibilities gives us 2,036 additional structural descriptions in which 49 balls are blue and the other is not blue, for a total of 2,046.<sup>4</sup> If we assign an equal prior probability to all the structural descriptions not ruled out by the first 49 drawings, we find that the ratio  $r$  is  $1/2,046$ . Putting this ratio into our formula for the posterior probability of the 50 blue balls hypothesis, we get a value of  $50/2,046 \div 2,096/2,046$ , i.e.,  $50/2,096$ , or .024. On this calculation the premisses of the argument make it highly probable (.976) that the 50th ball is not blue.

Of course, we are not obliged to give each structural description the same logically prior probability as any other one. Carnap's attempt (1962/1950) to base an "inductive logic" on this sort of probability foundered on the inability to single out a unique measure function which would assign absolute prior probabilities to each set of structurally isomorphic state descriptions (Carnap and Jeffrey 1971, Jeffrey 1980). We might, for some reason, suppose that the absolute prior probability of an  $n$ -coloured ball decreases as  $n$  increases. Thus the logical approach does not give us a definite answer to our problem.

#### 4. Some hypothetical answers

We can however use equations 4 and 5 above to draw some hypothetical conclusions about the posterior probability of the 50 blue balls hypothesis.

First, the assertion by Stephen Thomas that this probability is "well in excess of 80%" is logically equivalent to the claim that the ratio  $r$  is well in excess of .08.<sup>5</sup> In other words, his claim assumes that, before any balls are drawn from the urn, the hypothesis that all 50 are blue is at least one-tenth as probable as the hypothesis that 49 are blue and one is non-blue.<sup>6</sup> This seems like a rather conservative assumption, but nothing in the information Thomas provides entitles him to make it.

Second, Nolt's claim that "we can, without violating any mathematical law, assign that proposition [*sc.* the 50 blue balls hypothesis—DH] any probability we like" assumes that no mathematical law prevents us from assuming any ratio we like of the prior probabilities, or in the limiting cases setting one or the other of them at 0. If we assign the 50 blue balls hypothesis a posterior probability of 0, we are assuming that our background information rules out that hypothesis from the very beginning, even before any balls are drawn from the urn; in other words,  $r = 0$ . For example, we may know in advance that the "balls" are a set of Moslem prayer beads, which (I am told) always have one bead of a different colour than the others, symbolizing that nothing on earth is perfect.<sup>7</sup> If at the opposite extreme we assign the 50 blue balls hypothesis a posterior probability of 1, we are assuming that our background information rules out the 49 blue balls hypothesis from the very beginning; in other words  $r^{-1} = 0$ . For example, we may know in advance that the balls are all of the same colour, a specific version of a "uniformity of nature" assumption. Notice, however, that we can calculate a value for the posterior probability without making any general assumption that nature is uniform, contrary to the claims of some philosophers unduly influenced by Hume. To assign some intermediate value for the posterior probability of the 50 blue balls hypothesis, we simply calculate the required assumed ratio of the prior probabilities; a posterior probability of .5, for example, requires a ratio of the prior probabilities of .02; in other words, in advance of drawing any balls from the urn, the prior probability that 49 are blue is 50 times the prior probability that all 50 are blue. We might suppose, for example, that the "balls" are a set of Moslem prayer beads dawn at random from the assembly line of a factory which by mistake produces a

set of 50 beads of the same colour once for every 50 times it correctly produces a set with 49 beads of one colour and the 50th of a different colour. By varying the defect rate in this imaginary factory, we can produce any posterior probability we like. None of these scenarios appears to involve any violation of a mathematical law. Thus, as far as one can see, Nolt's claim is correct.

In the absence of specific information and in default of a well-grounded inductive logic, we can at best make qualitative judgements about the prior degree of confidence it is rational to place in each of the subsequently unrefuted hypotheses. It is not rational to be absolutely certain in advance that all 50 balls are blue, or absolutely certain in advance that not all 50 balls are blue; the background information and new evidence are consistent both with this hypothesis and with its negation, and it is not rational to place absolute confidence in a proposition which later turns out to be false, as either of these might. For the same reason, it is not rational to be absolutely certain in advance that 49 balls are blue, or that the number of blue balls is not 49. That is, it would not be rational, given the information supplied and the background information available to us, to assume that either relevant prior probability was 1, or to assume that either relevant prior probability was 0. Beyond that, it seems difficult to go. How do we express the fact that our background information, before drawing any balls from the urn, neither makes it absolutely certain that 50 balls are blue, nor makes it absolutely certain that not all 50 balls are blue? Such vague qualitative judgements can be expressed by indicating that the rational degree of confidence lies in a range. If we let  $j$  be  $p(50 \text{ blue balls} | K)$ , the prior rational degree of confidence in the 50 blue balls hypothesis, then we can say that  $0 < j < 1$ . Similarly, if we let  $m$  be  $p(49 \text{ blue balls} | K)$ , the prior rational degree of confidence in the 49 blue balls hypothesis, then we can say that  $0 < m < 1$ . Since  $r = j \div m$ , and the posterior probability  $p(50 \text{ blue balls} | \text{first 49 drawn are blue} \ \& \ K) = 50r \div (50r + 1)$ , we can conclude only that  $0 < p(50 \text{ blue balls} | \text{first 49 drawn are blue} \ \& \ K) < 1$ .<sup>8</sup>

### 5. An invalid argument schema

A tempting approach to our example is to note that the evidence is much more likely to occur on the hypothesis that all 50 balls are blue than on the hypothesis that 49 balls are blue and one is not blue; the likelihoods are respectively 1 and .02, as indicated in section 2 above. And the evidence rules out all other competing hypotheses. Since the result we observed was inevitable on the first hypothesis but highly unlikely on the only alternative hypothesis consistent with our evidence, does this result not make it highly probable that the first hypothesis is true, and thus that the last ball in the urn is blue?

The considerations advanced above show that this method of reasoning is invalid; that is, its premisses do not necessarily confer a high probability on the conclusion of the argument. A Bayesian explanation of why it is invalid is that it does not take into account the prior probability of the two hypotheses. If our



background information gives the 50 blue balls hypothesis a very much lower prior probability than the 49 blue balls hypothesis, the fact that the evidence is exactly what we would expect on the basis of the 50 blue balls hypothesis, but highly unlikely given the 49 blue balls hypothesis, is not enough to make the 49 blue balls hypothesis highly probable.<sup>9</sup> Thus, the following argument schema, though plausible, must be rejected as invalid:

The observed results rule out all but  $n$  mutually exclusive hypotheses.  
On one of these hypotheses, the observed results were bound to occur.  
On any of the others, the observed results were highly unlikely.  
Therefore, probably the first hypothesis is correct.

## 6. Conclusion: general lessons about induction

In the dispute which occasioned this article, Nolt was right and Thomas was wrong. From the resolution of the dispute, we can draw the following general lessons about induction:

1. Uniform experience does not necessarily confer a high epistemic probability on an extrapolation or generalization from that experience. We must always keep in mind Russell's chicken, whose experience of getting its daily ration of grain from its keeper gradually increased its confidence that it would get the grain the next day, until one day the keeper cut its head off. (Russell 1959/1912, 63) Analogues of Russell's chicken can be found in the daily newspaper making theoretically ungrounded extrapolations from the past performance of stock markets and other trading forums.
2. Rational extrapolation or generalization from uniform experience typically involves a great deal of background information.
3. Rational extrapolation or generalization from uniform experience does not require a general assumption that nature is uniform or that the future will resemble the past.
4. In many situations it is rational to make only a vague qualitative judgement about the probability of a hypothesis given our background information. Such a vague qualitative judgement can be interpreted quantitatively in terms of a range of values.
5. The fact that new evidence at our disposal is just what we would expect on one hypotheses, but highly unlikely on the basis of any of its unrefuted competitors, does not necessarily confer high probability on the former hypothesis. The posterior probability of the unrefuted hypotheses depends also on their epistemic probabilities prior to learning of the new evidence.<sup>10</sup>

## Notes

<sup>1</sup>This definition is standard; see for example Hymans (1967: 51-52) and Mosteller, Rourke and Thomas (1961: 88-89). It is more easily understood in terms of probabilities which are frequencies; think for example of the probability that a card in a deck is a queen, given that it is a face

card. The idea behind interpreting conditional rational degrees of confidence in the same way is that the rational degree of confidence in a hypothesis given new evidence is the rational degree of confidence, given background information (other than the new evidence) that the hypothesis and the evidence are true, divided by the rational degree of confidence, given the same information, that the evidence obtains. This idea seems intuitively plausible.

<sup>2</sup> If  $(H \& K)$  entails  $E$ , then  $(E \& H \& K)$  is logically equivalent to  $(H \& K)$ , so that  $p(E \& H \& K) = p(H \& K)$ : logically equivalent propositions have the same probability. Hence, since  $p(E|H \& K) = p(E \& H \& K) \div p(H \& K)$ ,  $p(E|H \& K) = 1$ . In general, the rational degree of confidence in a proposition, given certain other information from which that proposition follows necessarily, is 1. Roughly speaking, this amounts to saying that it would be fair to give as high odds as you like on a bet that the other information entails the proposition in question.

<sup>3</sup> The algebraic calculation goes as follows. Let  $x = p(\text{all 50 balls blue}|\text{first 49 blue} \& K)$ ,  $y = p(49 \text{ balls blue}|\text{first 49 blue} \& K)$ ,  $j = p(\text{all 50 balls blue}|K)$ ,  $m = p(49 \text{ balls blue}|K)$ , and  $n = p(\text{first 49 balls blue}|K)$ . Then the three numbered equations above become:

1.  $x = j/n$
2.  $y = m/50n$
3.  $x + y = 1$ .

From 2, assuming that  $m \neq 0$  and thus  $y \neq 0$ , we get  $n = m/50y$ . By substitution of this value for "n" in 1, we get  $x = 50jy/m$ . Now  $r$ , as defined above, is  $j/m$ . Hence  $x = 50ry$ . By substitution of this value for "x" in 3, we get  $50ry + y = 1$ , i.e.,  $y(50r + 1) = 1$ . Since  $50r + 1 \neq 0$ , we can divide both sides of this equation by  $50r + 1$ , to get  $y = 1/(50r + 1)$ . By substitution of this value for "y" in 3, we get  $x + 1/(50r + 1) = 1$ . Subtracting  $1/(50r + 1)$  from both sides, we get  $x = 1 - 1/(50r + 1)$ . Simplifying the right-hand side, we get  $x = 50r/(50r + 1)$ .

<sup>4</sup> The number of ways that a ball can have  $n$  of the 11 possible colours is  $11!/n!(11 - n)!$  For example, the number of ways that a ball can have 6 of the 11 possible colours is  $11!/6!5!$ , or  $(11 \times 10 \times 9 \times 8 \times 7)/(5 \times 4 \times 3 \times 2)$ , or 462. Adding the number of combinations for two-coloured, three-coloured, and so on up to 10-coloured balls gives a total of 2,036 combinations. In taking all these combinations to be part of a structure description in which 49 balls are blue and one is not blue, I assume that a ball which is partly blue and partly one or more other colours is not blue. The opposite assumption would reduce the number of combinations to 1,012 ways that a ball can have from two to 10 of the 10 colours other than blue.

<sup>5</sup> Since the posterior probability  $x = 50r/(50r + 1)$ , the claim that  $x$  is well in excess of 80% is the claim that  $50r/(50r + 1) \gg .8$ . That is,  $50r \gg .8(50r + 1)$ . That is,  $50r \gg 40r + .8$ . Hence  $50r - 40r \gg .8$ . That is,  $10r \gg .8$ . Hence  $r \gg .08$ .

<sup>6</sup> I assume here that a ratio well in excess of .08 is at least .1.

<sup>7</sup> Francisca Snoeck Henkemans suggested this example.

<sup>8</sup> A posterior probability of 0 is ruled out by the fact that  $j \neq 0$ , which implies that  $r \neq 0$ , which implies that  $50r \div (50r + 1) \neq 0$ . A posterior probability of 1 is ruled out by the fact that  $m \neq 0$ . Any intermediate value of the posterior probability can be accounted for by solving for  $r$  in the equation for the posterior probability, setting the larger of the two prior probabilities at some arbitrarily chosen number between 0 and 1, then solving for the other prior probability. For example, if the posterior probability is .00059, then  $50r \div (50r + 1) = .00059$ , and  $r \approx .000012$ . Since  $j < m$ , we first set  $m$  at some arbitrarily chosen positive number less than 1, say .0045. Then  $j = r \times m \approx .000012 \times .0045 = .0000054$ .

<sup>9</sup> One may appreciate this fact more readily if one expands the prior likelihood  $p(E|K)$  in Bayes' theorem, using the probability calculus, to get:  $[p(E|H \& K) \times p(H|K)] + [p(E|\sim H \& K) \times (1 - p(H|K))]$ . Suppose the evidence  $E$  is highly likely if the hypothesis is true but highly unlikely if the hypothesis is false. For example, let  $p(E|H \& K) = 1$  and  $p(E|\sim H \& K) = .0008$ . Now

suppose that the prior probability of the hypothesis is very low; for example, let  $p(H|K) = .0001$ . Then  $p(H|E \& K) = p(H|K) \times p(E|H \& K)$ ,  $\{[p(E|H \& K) \times p(H|K)] + [p(E|\sim H \& K) \times (1 - p(H|K))]\} = .0001 \times 1 \div \{[1 \times .0001] + [.0008 \times (1 - .0001)]\} = .0001 \div \{.0001 + [.0008 \times .9999]\} = .0001 \div \{.0001 + .0008\} = .0001 \div .0009 = 1/9 = .1111$ . So, even though the evidence is bound to occur if the hypothesis is true and highly unlikely if the hypothesis is false, the posterior probability of the hypothesis, given the evidence, is only .1111.

<sup>10</sup> Ancestors of this paper were read at the Ontario Society for the Study of Argumentation's conference on argumentation and rhetoric in May 1997, at the University of Windsor in March 1998, and at the 4th International Conference on Argumentation in June 1998. The present paper is substantially different, and I hope better, than those ancestors. I would like to thank those who commented on previous versions, including especially Roderic Girle, Robert C. Pinto, Robert H. Ennis, Sally Jackson, and two anonymous referees for the Canadian Philosophical Association. They are of course not responsible for any faults in the present paper, with which indeed some of them would vehemently disagree.

## References

- Bacchus, Fahiem. 1990. *Representing and Reasoning with Probabilistic Knowledge: A Logical Approach to Probabilities*. Cambridge, MA, and London: MIT Press.
- Carnap, Rudolf. 1962/1950. *Logical Foundations of Probability*, Second Edition. Chicago: University of Chicago Press. First edition published in 1950.
- Carnap, Rudolf, and Jeffrey, Richard C., editors. 1971. *Studies in Inductive Logic and Probability*, volume I. Berkeley, Los Angeles and London: University of California Press.
- Howson, Colin, & Urbach, Peter. 1989. *Scientific Reasoning: the Bayesian approach*. La Salle, IL: Open Court.
- Hymans, Saul H. 1967. *Probability Theory, with Applications to Econometrics and Decision-Making*. Englewood Cliffs, NJ: Prentice-Hall.
- Jeffrey, Richard C., editor. 1980. *Studies in Inductive Logic and Probability*, volume II. Berkeley, Los Angeles and London: University of California Press.
- Kaplan, Mark. 1996. *Decision Theory as Philosophy*. Cambridge: Cambridge University Press.
- Mosteller, Frederick, Rourke, Robert E. K., and Thomas, George B. Jr. 1961. *Probability and Statistics*. Reading, MA: Addison-Wesley.
- Nolt, John. 1985. More on induction and possible worlds: replies to Thomas and Kahane. *Informal Logic* 7, 1: 55-60.
- Pollock, John L. 1990. *Nomic Probability and the Foundations of Induction*. New York: Oxford University Press.
- Ramsey, F. P. 1990. *Philosophical Papers*, edited by D. H. Mellor. Cambridge: Cambridge University Press. Papers written between 1925 and 1929.
- Russell, Bertrand. 1959/1912. *The Problems of Philosophy*. New York: Oxford University Press.
- Russell, Bertrand. 1948. *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster.
- Savage, Leonard J. 1972/1954. *The Foundations of Statistics*, second revised edition. New York. Dover. First edition published in 1954.

- Thomas, Stephen N. 1984. Degrees of validity and ratios of conceivable worlds. *Informal Logic* 6, 3: 31-34.
- Thomas, Stephen N. 1997. *Practical Reasoning in Natural Language*, 4th edition. Upper Saddle River, NJ: Prentice-Hall. (Previous editions published in 1973, 1981 and 1986.)

*David Hitchcock, Department of Philosophy  
McMaster University, Hamilton, Ontario, Canada L8S 4K1  
hitchckd@ssevox.cis.mcmaster.ca*