belief is true: but why is knowing this needed for having a method that works? In the absence of such a requirement, a supporter of Nozick can easily accommodate Garrett's examples as genuine cases of knowledge.[3]

6238 Orange Street,
Los Angeles, California 90048, U.S.A.

[3] I am grateful to Professor Nozick for many discussions about his theory of knowledge. He should not, however, be held responsible for anything in this article.

PROBABILITY: SUBJECTIVE AND MATHEMATICAL

By P. J. R. MILLICAN

IN 'Is Epistemic Preferability Transitive?' (ANALYSIS 41.3, June 1981), Roy Sorensen attempted to give a counterexample to show that the 'more likely to be correct than' relation is not always transitive. My refutation of this argument ('On the Transitivity of Epistemic Preferability', ANALYSIS 42.2, March 1982) focused on Sorensen's erroneous assumption that iterated probabilities may be collapsed, that if p implies *probably q*, and q implies *probably r*, so that p implies *probably (probably r)*, then p must imply *probably r tout court*. Nowhere do I suggest that such an assumption might be justified even under some alternative, non-mathematical interpretation of 'probable', and yet in his reply ('Subjective Probability and Indifference', ANALYSIS, 43.1, January 1983) Sorensen appears to take for granted that this is the case. He writes: 'Peter Millican has pointed out that my argument ... is unsound *if* epistemic preferability is given a probabilistic interpretation', and he then goes on to 'provide *another* counterexample ... that *also* shows that epistemic preferability cannot be given a probabilistic interpretation' (my italics). He thus appears to enlist me to the ranks of those who *reject* the probabilistic view, despite the fact that I use probabilistic reasoning precisely to highlight 'the flaw in Sorensen's argument'! This misunderstanding I should like to rectify.

Let us suppose that a company director wishes to recruit two new employees but, because of his difficulty in choosing amongst the ten equally qualified applicants, decides to eliminate them one by one until he is left with two to fill the posts. As one of these ten, I am likely (epistemically) to survive the first elimination (since only one of the ten will fail to do so), and having become one of the remaining nine, I am also likely (epistemically) to survive the second. Similar reasoning would seem to show that I am likely to survive each elimination in turn, and if iterated likelihoods are allowed to collapse, this has the consequence that I am likely to be appointed. When there are five times as many applicants as jobs, such a result is surely absurd under any interpretation of 'likely', epistemic or otherwise!

Without his implausible collapsing of iterated probabilities, Sorensen's purported counterexample to the transitivity of epistemic preferability has absolutely no force whatever. In his later article, however, he presents a second apparent counterexample (strictly, to the transitivity of epistemic *indifference*), which we must now consider:

> Suppose there are 21 buckets of water $b_1, b_2, \ldots, b_{21}$ such that bucket b_i contains water at i degrees centigrade. Jim makes pairwise comparisons between b_i and b_{i+1} where $1 \leqslant i < 21$ by putting his hands in the appropriate buckets. Since Jim cannot discern a temperature difference of 1 degree by this method, he is epistemically indifferent between propositions p_i and p_{i+1} where 'p_i' reads 'The bucket containing water of the highest temperature of all 21 buckets is bucket i'. So if the epistemic indifference relation is transitive, Jim should be indifferent between p_1 and p_{21}. However, Jim can discern a temperature difference of 20 degrees and so he is not indifferent between p_1 and p_{21}. Therefore, epistemic indifference is not transitive.

This argument contains a subtle but devastating flaw, which is disguised by Sorensen's use of the present tense to describe all of Jim's actions and judgements. For transitivity has been undermined only if Jim is epistemically indifferent between p_i and p_{i+1} *at the same time* as he epistemically prefers p_{21} to p_1, and once the story is filled out a little, it becomes far from clear that this could be the case. Presumably Jim's preference for p_{21} over p_1 comes either from an explicit comparison, or else from a gradual realization that the temperature of the buckets has increased. Either way, this new information provides him with additional data which might require a revision of his previous judgements of epistemic indifference between p_i and p_{i+1}. Sorensen's case is obviously incomplete until he has dismissed such a possibility, and this he cannot do, for it can in fact be shown that once Jim knows b_{21} to be hotter than b_1, he does indeed have reason to prefer p_{i+1} to p_i in every case ($1 \leqslant i < 21$).

Let us suppose that Jim is able to detect a difference in temperature if and only if that difference is at least D degrees (such a sensory threshold will not of course be so well-defined in practice, but this makes no substantial difference to the argument). Thus he is faced with a series of indiscernible differences $d_1, d_2, d_3, \ldots, d_{20}$ (where d_i is the increase in temperature between b_i and b_{i+1}), all of them between $-D$ and $+D$, which together add up to an overall rise in temperature of D or more. Now *before comparing* b_1 *with* b_{21}, and in the absence of any further information about these indiscernible differences, Jim is quite epistemically indifferent between the possible values of, say, d_1 in the range $-D$ to $+D$: he knows that d_1 lies somewhere within this range, but he does not know where. The same is true of d_2, d_3 etc.; so as far as he is concerned, for all i, the value of d_i has a uniform probability density function over the range $-D$ to $+D$. Under these circumstances the probability that $d_i = 0$ is negligible, d_i having a 0.5 probability of being positive and 0.5 of being negative. It should be obvious that the same comment applies to the sum $S = d_1 + d_2 + d_3 + \ldots + d_{20}$, although this probability density function, over the range $-20D$ to $+20D$, is not uniform.[1] One important corollary is that, *before comparing* b_1 *with* b_{21}, the probability that $S \geqslant +D$ is less than 0.5 (the actual figure is approximately 0.349). In other words, it is unlikely that the difference in temperature between b_1 and b_{21} will be both positive and discernible.

There are 21 possibilities $h_0, h_1, h_2, \ldots, h_{20}$ concerning the direction of the 20 indiscernible differences $d_1 \ldots d_{20}$, where h_n represents the hypothesis that n of them are positive (representing a rise in temperature) and $(20-n)$ negative (representing a fall). Since each of these differences is independent of the others and equally as likely to be positive as negative, it follows that, *before comparing* b_1 *with* b_{21}, the probability that h_n is true yields a (binomial) distribution function in which h_{10} is the most likely hypothesis, while h_9 and h_{11}, h_8 and h_{12}, h_7 and $h_{13}, \ldots, h_0$ and h_{20}, are pairs of equally likely hypotheses, the pairs being themselves in order of decreasing probability. That h_n and h_{20-n} are indeed equally likely can be seen easily when the equal probability of positive and negative differences is combined with the observation that h_n postulates n positive and $(20-n)$ negative differences, while h_{20-n} postulates n negative and $(20-n)$ positive differences.

If h_n is true, then n of the 20 differences are positive and $(20-n)$ negative. So in the absence of any information to distinguish amongst them, any one of them stands an $n/20$ chance of being positive, and a $(20-n)/20$ chance of being negative. It should therefore be clear that *if, after comparing* b_1 *with* b_{21}, h_n *can be shown to be more probable for high values of* n *than for low values, then it will follow that any particular difference* d_i *is more likely to be positive than to be negative.* In this case, Sorensen's argument against the transitivity of epistemic indifference would be completely undermined, since p_{i+1} would be in all cases epistemically preferable to p_i, whereas Sorensen claims them to be epistemically indifferent. We now proceed to demonstrate, by an inverse probability argument using Bayes' Theorem, that *after comparing* b_1

[1] By the Central Limit Theorem, the distribution will be approximately normal, with mean zero and variance $20D^2/3$. In fact, the argument in the text in no way depends upon a *uniform* distribution function even for d_i: the important result is simply that Jim has no reason to advance to suppose b_i to be hotter than b_{i+1} – something that Sorensen's argument itself presupposes.

with b_{21}, Jim has reason to prefer h_{11} to h_9, h_{12} to h_8, h_{13} to h_7, and so on, thus fulfilling the antecedent of the italicized conditional.

Let us start by assuming that either h_9 or h_{11} is correct: we choose this pair because, as shown previously, the probabilities of h_9 and h_{11} 'initially' (i.e. before the comparison of b_1 and b_{21}) are exactly the same. Both hypotheses postulate nine positive indiscernible differences and nine negative – they differ in that according to h_9 the remaining two differences are negative, while according to h_{11} they are positive. It should be manifestly clear that the sum S of all twenty differences is more likely to exceed $+D$ if h_{11} is true than if h_9 is true.[2]

Bayes' Theorem states that the probability of an hypothesis H after observing some evidence E is equal to the initial probability of (H and E) divided by the initial probability of E. The initial probability of (H and E) is equal to the initial probability of H multiplied by the conditional probability that E is true given that H is true. Putting these together:

$$\text{Probability (H given E)}$$
$$= \frac{\text{Initial Probability of H} \times \text{Probability (E given H)}}{\text{Initial Probability of E}},$$

which is proportional to the conditional probability of E given H. Thus where we are given a single piece of evidence for two rival hypotheses which have the same initial probability, we should prefer whichever hypothesis makes that evidence more likely to have occurred, since this will give the greater value for the conditional probability of E.

In the case under discussion, we are comparing the hypotheses h_9 and h_{11} in the light of the 'evidence' that $S \geqslant +D$ (since the difference between b_1 and b_{21} is positive and discernible). The two hypotheses have an equal initial probability, but h_{11} makes our evidence more likely than does h_9: that evidence therefore reciprocates by making h_{11} more probable than h_9, and similar reasoning will show h_{12} to be more probable than h_8, h_{13} than h_7, h_{14} than h_6, and so on. We have seen that this is sufficient to prove p_{i+1} to be more probable than p_i (the probability that b_{i+1} is hotter than b_i is approximately 0.605)[3], so if Jim is rational, although his epistemic preferences amongst p_1, p_2, p_3 etc. will change with the new evidence, there is nothing to suggest that they will cease to be transitive. Sorensen's second example is therefore no more successful than his first: he has given no reason to question the transitivity of epistemic preferability, and hence no reason to reject an interpretation of that notion which conforms to the calculus of probabilities.

[2] To find these probabilities mathematically, we again use the Central Limit Theorem to show that h_n yields an approximately normal distribution function for S, with mean $(n-10)D$ and variance $5D^2/3$. h_9 and h_{11} thus give figures of 0.061 and 0.5 respectively for the initial probability that $S \geqslant +D$.

It might perhaps be thought that the above argument itself relies on probabilistic assumptions such as Bayes' Theorem, and thus presupposes a probabilistic (and therefore transitive) conception of epistemic preferability. This objection, however, would be misguided, since the point of the argument is not to establish that epistemic preferability *is* transitive, but only to refute Sorensen's attempt to prove that it is *not*. His 'counterexample' can be accommodated perfectly well within a probabilistic interpretation of epistemic preferability: it therefore fails to overthrow that interpretation.

One final comment. Sorensen hopes to find cases where contradictions arise from the supposition that epistemic preferability is probabilistic. Unless probability theory is actually inconsistent, however, it is difficult to see how he can possibly succeed unless he can show that the relevant basic methods of acquiring information (e.g. by putting hands in buckets of water) themselves yield results which resist probabilistic interpretation. To take Sorensen's first example: as soon as it is agreed that the probability of a research group's success is proportional to its resources, it requires only elementary mathematics to show that the transitivity of epistemic preferability has not been undermined. Similarly with his second example: if the assumptions about initial probabilities given earlier are accepted (and Sorensen has given no reason whatever why they should not be), then we can be sure that probability theory will be able to deal with the apparent contradiction, even if our ignorance of the Central Limit Theorem and of Bayes' Theorem prevents us from performing the necessary calculations ourselves. Thus if Sorensen wishes to show that epistemic preferability is not transitive, he must prove that at least some basic forms of evidence are irreducibly non-probabilistic. How he is to go about doing so is another question!

Lincoln College,
Oxford OX1 3DR

[3] The complete calculation for $P(d_i > 0)$ turns out like this, ignoring in the intermediate stages values of n which make a negligible difference. Three decimal places are given, but more accurate figures are used in the calculation.

a Initial $P(S \geqslant +D) = 0.349$

b	n:	8	9	10	11	12	13	14	15	16	17
c	Initial $P(h_n)$:	0.120	0.160	0.176	0.160	0.120	0.074	0.037	0.015	0.005	0.001
d	$P(S \geqslant +D)$, given h_n):	0.010	0.061	0.219	0.500	0.781	0.939	0.990	0.999	1	1
e	Final $P(h_n)$ $= c \times d/a$:	0.003	0.028	0.111	0.229	0.269	0.199	0.105	0.042	0.013	0.003
f	$P(d_i > 0$, given $h_n)$ $= n/20$:	0.400	0.450	0.500	0.550	0.600	0.650	0.700	0.750	0.800	0.850
g	Final $P(h_n$ and $d_i > 0) = e \times f$:	0.001	0.013	0.055	0.126	0.161	0.129	0.073	0.032	0.011	0.003

h Final $P(d_i > 0)$ = sum of row g = 0.605