Is the Pope the Pope?

SIR — Beck-Bornholdt and Dubben (*Nature* **381**, 730; 1996) describe a common misinterpretation of the *P*-value of a classical statistical test using the following example. The chance that a randomly chosen human being is the Pope is about 1 in 6 billion. John Paul II is the Pope. What are the chances that John Paul II is human? By analogy to syllogistic reasoning, Beck-Bornholdt and Dubben suggest 1 in 6 billion but point out that this is "obviously not sensible".

It certainly isn't. The probability of data given a hypothesis, $P(D|H)$, is not the same as the probability of the hypothesis given the data, $P(H|D)$. This is an elementary error regardless of one's preferred statistical approach. Bayesian statistical inference, which includes syllogistic reasoning as a special case, is particularly well-suited for avoiding this sort of mistake (H. Jeffreys, *Theory of Probability*, Oxford University Press, 1939).

A similar pitfall is the infamous 'prosecutor's fallacy', in which a probability that a DNA fingerprint match would occur in someone other than the true criminal — $P(\text{match}|\text{innocent})$ — is used incorrectly as the probability that a suspect is innocent — $P(\text{innocent}|\text{match})$. In a city of ten million people, a one-in-a-million DNA fingerprint match will give ten other people the same fingerprint as the true criminal. In the absence of other evidence, the odds that the suspect is innocent are better than 90%, not one in a million.

Let H represent the class (hypothesis) of humanness, A represent the class (hypothesis) of alienness, and J represent the observation (data) that a randomly chosen individual is Pope John Paul II. Bayes's theorem tells us how to infer the probability that the Pope is human, $P(H|J)$:

$$P(H|J)=\frac{P(J|H)P(H)}{P(J|H)P(H)+P(J|A)P(A)}$$

Thus, to infer the probability that the Pope is human, $P(H|J)$, we have to have two more numbers in addition to the probability $P(J|H)$ of drawing the Pope at random from the class of humans:

(1) $P(J|A)$, the probability of choosing the Pope at random from the class of aliens.

(2) $P(A)$, the *a priori* probability that a randomly chosen individual is an alien instead of a human.

$P(H)$ is just $1-P(A)$, if only these two hypotheses are considered.

Presumably the probability $P(J|A)$ of choosing the Pope from the class of aliens is infinitesimal. The prior probability of choosing an alien as opposed to a human, $P(A)$, is also expected to be quite small, except perhaps near secret US Air Force bases. As either $P(J|A)$ or $P(A)$ approaches zero, the probability that the Pope is human approaches one.

It is a shame that Bayesian methods are not part of all introductory statistics classes. In this case, they quickly reassure us that the Pope is (probably) not an alien.

Sean R. Eddy
Department of Genetics,
Washington University School of Medicine,
St Louis, Missouri 63110, USA
David J. C. MacKay
Cavendish Laboratory,
University of Cambridge,
Madingley Road, Cambridge CB3 0HE, UK

SIR — In their syllogistic reasoning to present a papal paradox, surely Beck-Bornholdt and Dubben's sequence is incorrect. It should be as follows: (1) If an individual is human s/he is probably not the Pope (premise); (2) John Paul II is human (premise); (3) therefore John Paul II is probably not the Pope (conclusion).

In fact, of course, he is the Pope, but probability allows for this possibility (it was just extremely unlikely) and the logical progression holds. Their sequence made the classic mistake of assuming that "if A, then not B" implies "if B, then not A", which can be disproved by many examples more trivial than the papal paradox. Indeed, if applied to their first example, it would imply that all mortals are human, a hypothesis easily disproved by considering any other living being.

If the logical sequence they describe for statistical hypothesis testing is actually the one in use, then it too makes the same mistake, and is therefore in error, but not for the reason they suggest. I think the conclusion in the sequence should be "(3) therefore the null hypothesis is PROBABLY wrong", with probability theory providing the degree of uncertainty, leading to conclusions such as "it is extremely unlikely (5% probability) that this result arose by chance" — but I'm probably wrong!

Stephen P. Gosden
Chaussée de Wavre,
B-1160 Brussels, Belgium
e-mail: stephen.gosden@dg12.cec.be

SIR — The argument of Beck-Bornholdt and Dubben can be paraphrased as follows. We have a human-looking organism before us and wish to test the hypothesis that it is in fact human. We know that if it *is* human, it is very unlikely to be the Pope. We find out that it is indeed John Paul II, and have to conclude that it is unlikely to be human. In claiming that this conclusion (and with it, statistical hypothesis testing) is absurd, Beck-Bornholdt and Dubben not only contradict one of their own assumptions but also describe bad scientific practice.

We use statistics to test hypotheses because we know that the same set of data could arise whether or not a given hypothesis is true. In making the comparison with statistical hypothesis testing, Beck-Bornholdt and Dubben assume at the outset that a non-human could be the Pope. They cannot therefore claim that the conclusion of their argument is ridiculous, merely that it is factually wrong.

Furthermore, this incorrect conclusion arises only if the randomly picked human-looking organism happens to be the Pope. On the many other occasions when we perform the experiment, we will quite reasonably fail to reject the hypothesis that our subject is human. The same applies to statistical reasoning: we will occasionally observe highly improbable data, and incorrectly reject the null hypothesis.

In their statement of their argument, Beck-Bornholdt and Dubben have implicitly allowed themselves the luxury of observing subjects until they find one that is the Pope. This procedure is equivalent to a scientist, anxious to reject a null hypothesis, repeating an experiment until a suitable result turns up.

B. J. Craven
Department of Psychology,
University of Stirling,
Stirling FK9 4LA, UK
L. S. Craven
49 Broompark East, Menstrie,
Clackmannanshire FK11 7AN, UK

SIR — Beck-Bornholdt and Dubben have shown at most that the Pope is not a random sample. Their 'syllogism':

Premise: "If... we randomly pick a human being, the probability that it is the Holy Father is extremely low..."

Conclusion: "Therefore, if an individual is human, it is probably not the Pope."

A syllogism states that an attribute shared by all members of a class is possessed by each member. Their conclusion ignores the fact that the attribute of probable nonidentity with the Pope was limited explicitly to the class of randomly chosen human beings. If the sampling method is not part of the class specification, why not just give as the premise "The Pope is, with high probability, not the Holy Father," thus suggesting an even wittier title for the published letter?

Statistical inference about membership in a hypothesized distribution applies to data randomly sampled from the population of inference. The failure of inference here results from no philosophical paradox about probability, but only from verbal shiftiness about the population and the sampling method.

James C. Nelson
Department of Plant Breeding and Biometry,
Cornell University,
Ithaca, New York 14853, USA
e-mail: jcn5@cornell.edu