

Lindley's Paradox

by Glenn Shafer

gshafer@andromeda.rutgers.edu

3 August 1996

A sharp null hypothesis may be strongly rejected by a standard sampling-theory test of significance and yet be awarded high odds by a Bayesian analysis based on a small prior probability for the null hypothesis and a diffuse distribution of one's remaining probability over the alternative hypothesis. This disagreement between sampling-theory and Bayesian methods was first studied by Harold Jeffreys (1939), and it was first called a paradox by Dennis Lindley (1957).

The paradox can be exhibited in the simple case where we are testing $\theta = 0$ using a single observation Y from a normal distribution with variance one and mean θ . If we observe a large value y for Y ($y = 3$, for example), then standard sampling theory allows us to confidently reject the null hypothesis. But the Bayesian approach advocated by Jeffreys can give quite a different result. Jeffreys advised that we assign a non-zero prior probability π_0 to the null hypothesis and distribute the rest of our probability over the real line according to a fairly flat probability density $\pi_1(\theta)$. If the range of possible values for θ is very wide, then the set of values within a few units of y will be very unlikely under $\pi_1(\theta)$, and consequently the overall likelihood of the alternative hypothesis,

$$L_1 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \exp(-(y-\theta)^2/2) \pi_1(\theta) d\theta ,$$

will be very small. It may even be so much smaller than the likelihood of the null hypothesis,

$$L_0 = \frac{1}{\sqrt{2\pi}} \exp(-y^2/2),$$

that the odds in favor of the null hypothesis,

$$\frac{P(\theta=0 | Y=y)}{P(\theta \neq 0 | Y=y)} = \frac{\pi_0 L_0}{1-\pi_0 L_1} , \quad (1)$$

are substantial.

We can think of (1) as a way of balancing arguments for and against the null hypothesis. Against the null hypothesis is its small initial probability (small π_0) and the unlikeliness of the observation under the null hypothesis (small L_0). For the null hypothesis is the unlikeliness of alternative values of θ near y (small $\pi_1(\theta)$, leading to small L_1). There is no strong constraint between the arguments for and against. No matter how small π_0 and L_0 are, a sufficiently diffuse $\pi_1(\theta)$ can make L_1 small enough to counterbalance them.

If we are confident of the specified prior distribution—if, for example, we are working with a series of problems involving θ s that are zero about π_0 of the time and distributed roughly according to $\pi_1(\theta)$ the rest of the time—then the Bayesian analysis is unassailable, and hence we must reject the standard sampling theory. An observation three standard deviations from the null hypothesis is not adequate to reject the null hypothesis if that observation is even more unlikely under the alternative hypothesis. This has led many authors to suggest that we make tests increasingly stringent as measurements become more precise relative to the range of possible values

for what is being measured. We should, for example, lower the significance level as the sample size grows. More sophisticated suggestions are made by Berger and Delampady (1987).

On the other hand, if diffuseness of $\pi_1(\theta)$ reflects merely a wide uncertainty about θ rather than a positive prior confidence that values of θ near y are likely to occur, the conflict seems to constitute a criticism of the Bayesian analysis. If we have no idea how θ arises, then our mere ignorance cannot justify a skepticism about values close to y so strong as to outweigh real evidence against the value of zero. This has motivated non-Bayesian approaches to weighing the evidence, such as the belief-function approach discussed by Shafer (1982).

References

- Berger, James, and Mohan Delampady (1987). *Statistical Science* **2**, 317-352.
- Jeffreys, Harold (1939). *Theory of Probability*. Oxford University Press.
- Lindley, Dennis V. (1957). *Biometrika* **44**, 187-192.
- Shafer, Glenn (1982). *Journal of the American Statistical Association* **77**, 325-351.