How likely is the null hypothesis?

Andy J. Wills

September 30, 2019

In our example, we estimated that the probability the null hypothesis is true, prior to data collection, was .95. So $P(H_0) = .95$. Our t-test gave $p = .049$, so $P(data|H_0) = .049$. To work out $P(H_0|data)$, we use Bayes' theorem. Bayes' is most typically written as:

$$P(H_0|data) = \frac{P(data|H_0)P(H_0)}{P(data)} \tag{1}$$

but can equivalently be expressed as:

$$P(H_0|data) = \frac{P(data|H_0)P(H_0)}{P(data|H_0)P(H_0) + P(data|\overline{H_0})P(\overline{H_0})} \tag{2}$$

where $P(\overline{H_0})$ means the probability the null hypothesis is false. Equation 2 is more useful in this case, as $P(\overline{H_0}) = 1 - P(H_0)$. The null hypothesis is either false or true, so the sum of these two probabilities must be one.

This leaves us to work out $P(data|\overline{H_0})$. It is tempting to further assume that $P(data|\overline{H_0}) = 1 - P(data|H_0)$. However, recall that $P(data|H_0)$ here is shorthand for "probability of a mean difference at least as extreme as the one observed, given the null hypothesis is true". So, $P(data|\overline{H_0})$ is the "probability of a score at least as extreme as the one observed, given the null hypothesis is false". This is not knowable unless one makes some assumptions about the distribution of scores when the null hypothesis is false, which in turn depends on how large you think the effect would be if it existed.

What we can say, however, is that $P(data|\overline{H_0})$ can't be greater than 1, so the lowest $P(H_0|data)$ can be is:

$$P(H_0|data) = \frac{0.049 \times .95}{.049 \times .95 + 1 \times .05} = 0.48 \tag{3}$$

So, actually it is an optimisitic simplification to say that the probability of the null hypothesis after this significant t-test is close to 50 : 50. It won't be lower than that, but it could be much higher.