Bayes class 12 – Name:

Sean van der Merwe

# Bayesian Hypothesis Testing

In the classical hypothesis testing framework:

We assume that the world is boring. There is no change, no difference, no impact, all coefficients are zero, all models are the same. We observe something interesting. We ask ourselves, “how likely is it to observe something at least this interesting under the assumption that the world is boring?”

If the answer is large then we don’t conclude anything and continue to assume our gloomy outlook.

If the answer is small then we take it as evidence against the null hypothesis of boringness and conclude that there is at least one interesting thing going on.

Problems:

Evidence against the null hypothesis is not necessarily evidence for the alternative. If the alternative is at least as unlikely to have produced the interesting result as the null then we are bound to make false conclusions.

It does not take practical significance into account at all. Statistical significance is established first and then practical significance is considered as an entirely separate step.

It does not account for any prior knowledge or similar tests already constructed. This can create bias in both directions very easily.

Can you think of a case where prior tests would favour the null and where it would favour the alternative?

Bayesian approaches:

Two popular Bayesian alternatives are:

1. Regions of posterior equivalence. This is where we explicitly bring in a measure of practical significance and calculate a posterior probability of being within or outside the region, or of posterior regions overlapping an interesting amount.
2. Model comparison. This is where we build models that fit each (of possibly many) reasonable hypotheses and see which is best supported by the data.

How would you test equality of means and equality of variances at the same time when your data is grouped?