STAT 431 — Applied Bayesian Analysis — Course Notes

Assessment of Priors

Fall 2022

Sensitivity Analysis

How do the results (estimates, posterior probabilities, intervals) vary depending on what prior is used?

Practical Approach:

Choose several reasonable priors (including noninformative and subjective), compute results for each, and assess the results for agreement.

R Example 1.5:

Population Proportion — Sensitivity Analysis

Posterior Predictive Checking

How compatible is a prior with the data?

A strong prior that is misinformative may lead to excessive bias in estimates and misleading conclusions.

Idea:

Compare the observed data to their posterior predictive distribution.

If that distribution is unlikely to produce “data” that look like the observed data, conclude there is incompatibility.

Eg: Pet survey (binomial data) with beta prior

Hypothetically consider repeating the survey with $n^* = n = 70$ participants, and suppose

$$Y^* \quad = \quad \text{number of them having pets}$$

Then Y^* follows the same model as the data Y.

We will compare the observed value of Y, $y = 12$, with the posterior predictive distribution of Y^*.

We first obtain its density $f^*(y^* \mid y = 12)$ ...

For simplicity, consider priors of the kind $\mathrm{Beta}(\alpha, \beta)$.

- If we were initially uncertain about the population proportion θ, we might choose the flat (uniform) prior $\mathrm{Beta}(1, 1)$:

$$\alpha = 1, \ \beta = 1$$

- If instead we were initially quite certain that almost everyone owns pets, we might try, for example

$$\alpha = 100, \ \beta = 1$$

Recall: A $\mathrm{Beta}(\alpha, \beta)$ prior led to the posterior

$$\theta \mid Y = 12 \quad \sim \quad \mathrm{Beta}(12 + \alpha, 58 + \beta)$$

Deriving the posterior predictive density for Y^* (similarly to before) leads to

$$f^*(y^* \mid y = 12) \quad =$$

$$\binom{70}{y^*} \frac{\Gamma(70+\alpha+\beta)}{\Gamma(12+\alpha)\,\Gamma(58+\beta)} \frac{\Gamma(y^*+12+\alpha)\,\Gamma(128-y^*+\beta)}{\Gamma(140+\alpha+\beta)}$$

$$\text{for} \;\; y^* = 0, 1, \ldots 70$$

Plotting this kind of density along with the data value $y = 12$ lets us make the comparisons ...

R Example 1.6:

Population Proportion — Posterior Pred. Checks

Remark:

More generally, both sensitivity analysis and posterior predictive checking can assess the data model, not just the prior. (More later.)