STAT 578: Advanced Bayesian Modeling

Week 2 – Lesson 1

Mean-Only Normal Sample

Fall 2019

Flat Prior Analysis

Review

Have sample of n observations from $\mathrm{N}(\mu, \sigma^2)$, σ^2 known.

With $\mathrm{N}(\mu_0, \tau_0^2)$ prior for μ, posterior is

$$\mu \mid y \;\sim\; \mathrm{N}(\mu_n, \tau_n^2)$$

$$\mu_n = \frac{\frac{1}{\tau_0^2}\mu_0 + \frac{n}{\sigma^2}\bar{y}}{\frac{1}{\tau_0^2} + \frac{n}{\sigma^2}} \qquad\qquad \frac{1}{\tau_n^2} = \frac{1}{\tau_0^2} + \frac{n}{\sigma^2}$$

What happens as $\tau_0^2 \to \infty$?

As $\tau_0^2 \to \infty$,

$$\mu \mid y \quad \longrightarrow_d \quad \mathrm{N}(\bar{y}, \sigma^2/n)$$

Note: Limiting distribution does not depend on prior mean μ_0.

Larger τ_0^2 implies prior $\mathrm{N}(\mu_0, \tau_0^2)$ is wider, “flatter,” less precise, less concentrated, more diffuse, less *informative*.

Leads to posterior inference that is driven by the data, not by prior knowledge

Flat Prior

Is there a prior density that produces $\mathrm{N}(\bar{y}, \sigma^2/n)$ as the posterior?

Yes, but it is not a **proper** density – it doesn't represent an actual distribution.

Consider

$$p(\mu) \;\propto\; 1 \qquad\qquad -\infty < \mu < \infty$$

Since the integral over $p(\mu)$ is infinite, it cannot be normalized to a true density – it is **improper**.

This is a **flat** prior. It is also considered **noninformative**.

Is it reasonable to use an improper prior?

Yes, provided Bayes' rule formally produces a proper posterior.

In this case,

$$
\begin{aligned}
p(\mu \mid y) \quad &\propto \quad p(\mu)\, p(y \mid \mu) \quad \propto \quad 1 \cdot p(y \mid \mu) \\
&\propto \quad \exp\left(-\frac{n}{2\sigma^2}(\mu - \bar{y})^2\right) \qquad\qquad -\infty < \mu < \infty
\end{aligned}
$$

This is proportional (in μ) to a $\mathrm{N}(\bar{y}, \sigma^2/n)$ density, so is proper.

Remarks:

- Even though the *posterior* converges as $\tau_0^2 \to \infty$, the $\mathrm{N}(\mu_0, \tau_0^2)$ *prior* does not converge.

  In particular, the prior density (of $\mathrm{N}(\mu_0, \tau_0^2)$) converges to zero – it does not converge to a flat prior proportional to 1.

- An improper prior does not define a true distribution, so sampling from it is impossible.

  This makes forward sampling impossible, which will have implications later.

Example: Flint Data

$n = 271$ log-scale observations (first-draw) of lead concentration (log-ppb):

$$\bar{y} \;\approx\; 1.40 \qquad\qquad \text{and suppose} \quad \sigma^2 \;=\; s^2 \;\approx\; 1.684$$

Using flat prior for log-scale mean μ,

$$\text{95\% posterior interval:} \qquad \bar{y} \;\pm\; 1.96 \cdot \sqrt{\sigma^2/n} \quad\approx\quad (1.25, 1.56)$$

(same as a classical 95% confidence interval)