Bayesian Inference in a Normal Population

September 22, 2005

Casella & Berger Chapter 7, Gelman, Carlin, Stern, Rubin Sec 2.6, 2.8, Chapter 3.
Normal Model

IID observations $Y = (Y_1, Y_2, \ldots, Y_n)$

$$Y_i \sim N(\mu, \sigma^2)$$

unknown parameters $\mu$ and $\sigma^2$. From a Bayesian perspective, it is easier to work with the precision, $\phi$, where $\phi = 1/\sigma^2$.

Likelihood

$$L(\mu, \phi|Y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \phi^{1/2} \exp\left\{ -\frac{1}{2} \phi (Y_i - \mu)^2 \right\}$$

$$\propto \phi^{n/2} \exp\left\{ -\frac{1}{2} \phi \sum_{i} (Y_i - \mu)^2 \right\}$$
Likelihood

\[ L(\mu, \phi|Y) \propto \phi^{n/2} \exp\left\{ -\frac{1}{2} \phi \sum_i (Y_i - \mu)^2 \right\} \]

\[ \propto \phi^{n/2} \exp\left\{ -\frac{1}{2} \phi \sum_i [(Y_i - \bar{Y}) - (\mu - \bar{Y})^2] \right\} \]

\[ \propto \phi^{n/2} \exp\left\{ -\frac{1}{2} \phi \left[ \sum_i (Y_i - \bar{Y})^2 + n(\mu - \bar{Y})^2 \right] \right\} \]

\[ \propto \phi^{n/2} \exp\left\{ -\frac{1}{2} \phi s^2 (n - 1) \right\} \exp\left\{ -\frac{1}{2} \phi n(\mu - \bar{Y})^2 \right\} \]

where \( s^2 = \sum_i (Y_i - \bar{Y})^2 / (n - 1) \) is the usual sample variance.
Prior Distributions

Conjugate prior distribution for \((\mu, \phi)\) is Normal-Gamma.

\[
\begin{align*}
\mu | \phi & \sim N(\mu_0, 1/(n_0 \phi)) \\
\phi & \sim \text{Gamma}(\nu_0/2, (\nu_0 s_0)/2)
\end{align*}
\]

\[
p(\phi) \propto \phi^{\nu_0/2-1} \exp\{-\phi \nu_0 s_0/2\}
\]

Non-informative prior distribution (improper) Reference Distribution

\[
p(\mu, \phi) = 1/\phi
\]

\(\mu\) is uniform on the real line, \(\log(\phi)\) is uniform on real line: invariance to scale and location changes of the data.
Reference Posterior Distribution

\[ p(\mu, \phi | Y) \propto L(\mu, \phi)p(\mu, \phi) \]

\[ = \phi^{n/2} \exp\left\{-\frac{1}{2}\phi s^2(n - 1)\right\} \exp\left\{-\frac{1}{2}\phi n(\mu - \bar{Y})^2 \right\} \phi^{-1} \]

\[ = \left\{ \phi^{\frac{n-1}{2}} e^{-\frac{1}{2}\phi s^2(n-1)} \right\} \left\{ \phi^{1/2} e^{-\frac{1}{2}\phi n(\mu-\bar{Y})^2} \right\} \]

\[ \propto \text{Gamma} \left( \frac{n-1}{2}, (n-1)\frac{s^2}{2} \right) N \left( \bar{Y}, \frac{1}{\phi n} \right) \]

\[ = p(\phi | Y)p(\mu | \phi, Y) \]
Marginal Distribution for $\mu | Y$

Obtain the marginal distribution for $\mu$ by integrating out $\phi$ from the joint posterior distribution, and recognizing the kernel of the distribution!

\[
p(\mu | Y) \propto \int p(\mu, \phi | Y) d\phi
\]

\[
= \int \phi^{n/2-1} \exp\left\{ -\frac{1}{2} \phi s^2 (n - 1) + n(\mu - \bar{Y})^2 \right\} d\phi
\]
Continued

\[ p(\mu|Y) \propto \int \phi^{n/2-1} \exp\{-\frac{1}{2} \phi s^2 (n - 1) + n(\mu - \bar{Y})^2\} d\phi \]

This has the form of a Gamma integral with \( \alpha = n - 1 \) and \( \beta \) equal to the mess multiplying \( \phi \),

\[ p(\mu|Y) \propto (s^2 (n - 1) + n(\mu - \bar{Y})^2)^{\frac{n-1+1}{2}} \]

\[ \propto \left(1 + \frac{1}{n - 1} \frac{(\mu - \bar{Y})^2}{s^2/n}\right)^{\frac{n-1+1}{2}} \]

Student-t\(_{n-1}(\bar{Y}, s^2/n)\) location \( \bar{Y} \), df = n-1, scale \( s^2/n \)
Proper Conjugate Priors

Under the Normal-Gamma Prior (proper prior distributions):

- Find the (conditional) posterior distribution $\mu | \phi$
- Find the (marginal) posterior distribution of $\phi$
- Find the (marginal) posterior distribution of $\mu$

Hint: Expand quadratics in $\mu$ to read off the posterior precision $p_n$ and mean $m_n$ then complete the square and factor

$$-.5(p_n\mu^2 - 2p_nm_n\mu + p_nm_n^2) = -.5p_n(\mu - m_n)^2$$
Example: SPF

A Sunlight Protection Factor (SPF) of 5 means an individual that can tolerate \( X \) minutes of sunlight without any sunscreen can tolerate \( 5X \) minutes with sunscreen.
Pairing

A paired design may be more powerful than two sample design because of patient to patient variability.

- Analysis should take into account pairing which induces dependence between observations
- Use differences
  - Use ratios or log(ratios) difference in logs

Ratios make more sense given the goals: how much longer can a person be exposed to the sun relative to their baseline.
Data

Differences

Normal Q–Q Plot

log(Ratio)

Normal Q–Q Plot
Model for SPF

- Model $Y = \log(\text{TRT}) - \log(\text{CONTROL})$ as $N(\mu, 1/\phi)$
- $E(\log(\text{TRT}/\text{CONTROL})) = \mu = \log(\text{SPF})$
- Want distribution of $\exp \mu \equiv \text{SPF}$

Summary statistics
- $\bar{y} = 1.998$
- $s^2 = 0.525$
- $n = 13$
Samples from the Posterior

To draw samples of SPF from the posterior distribution:

- Draw $\phi|Y$
  
  $\phi = \text{rgamma}(10000, (n-1)/2, \text{rate}=(n-1) \ast s2/2)$

- Draw $\mu|\phi, Y$
  
  $\mu = \text{rnorm}(10000, ybar, 1/\sqrt{\phi \ast n})$

- summarize, transform $(\sqrt{1/\phi})$, etc

- quantile($\exp(\mu)$, c(.025,.5,.975))
Distributions

Posterior Distribution of $\mu$

Posterior Distribution of $\sigma$
Distribution for SPF

95% Probability Interval 5 to 12 (equal tail area)