Conjugate Priors for Normal Data

September 22, 2008

Gill Chapter 3. Sections 4, 7-8
Normal Model

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

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

unknown parameters $\mu$ and $\sigma^2$ or precision, $\phi = 1/\sigma^2$.

Likelihood:

$$L(\mu, \phi|Y) \propto \phi^{n/2} \exp\left\{-\frac{1}{2} \phi \text{SS}\right\} \exp\left\{-\frac{1}{2} \phi n(\mu - \bar{y})^2\right\}$$

Sufficient statistics

- sample mean $\bar{y} = \sum_{i=1}^{n} y_i$
- sample sum of squares $\text{SS} = \sum_i(y_i - \bar{y})^2$
Conjugate Prior Distribution

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

\[
\mu | \phi \sim \text{N}(m_0, 1/(p_0\phi)) \\
\phi \sim \text{G}(v_0/2, SS_0/2)
\]

\[
p(\mu, \phi) \propto \phi^{v_0/2-1} \exp\{-\phi \frac{SS_0}{2}\} \phi^{1/2} \exp\{-\phi \frac{p_0}{2}(\mu - m_0)^2\}
\]

\[
\mu, \phi \sim \text{NG}(m_0, p_0, v_0/2, SS_0/2) \text{ Normal-Gamma family}
\]

\[
\mu, \phi \mid Y \sim \text{NG}(m_n, p_n, v_n/2, SS_n/2) \text{ Posterior is Normal-Gamma}
\]
Updating the Posterior Parameters

Under the Normal-Gamma prior distribution:

\[ \mu \mid \phi, Y \sim N \left( m_n, \frac{1}{p_n \phi} \right) \]

\[ \phi \mid Y \sim G \left( \frac{v_n}{2}, \frac{SS_n}{2} \right) \]

where

\[ p_n = p_0 + n \]

\[ m_n = \frac{n \bar{y} + p_0 m_0}{p_n} \]

\[ v_n = v_0 + n \]

\[ SS_n = SS_0 + SS + \frac{np_0}{p_n} (\bar{y} - m_0)^2 \]
Interpretation

- $p_n$ precision for estimating $\mu$ after $n$ observations
- $m_n$ expected value for $\mu$ after $n$ observations

$$m_n = \frac{n}{p_n} \bar{y} + \frac{p_0}{p_n} m_0 \text{ weighted average}$$

- $v_n$ degrees of freedom

$$\phi \sim G(a/2, b/2) \Leftrightarrow \phi b \sim \chi^2_a \text{ with degrees of freedom } a$$

- $SS_n = SS_0 + SS + \frac{np_0}{p_n} (\bar{y} - m_0)^2 \text{ posterior variation:}$
  - prior variation,
  - observed variation (sum of squares),
  - variation between prior mean and sample mean
Derivation

\[ p(\mu, \phi \mid Y) \propto \phi^{v_0/2-1} \exp\left\{-\phi \frac{SS_0}{2}\right\} \times \phi^{n/2} \exp\left\{-\phi \frac{SS}{2}\right\} \times \phi^{1/2} \exp\left\{-\phi \frac{p_0}{2} (\mu - m_0)^2\right\} \times \exp\left\{-\phi \frac{n}{2} (\mu - \bar{y})^2\right\} \]

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

\[-\frac{1}{2} (p_n \mu^2 - 2p_n m_n \mu + p_n m_n^2) = -\frac{1}{2} p_n (\mu - m_n)^2 \]

(Note: page 80 in Gill has errors; need to keep track of terms after completing the square.)
Derivation

Take second line and complete the square:

\[ p(\mu|\phi, Y) \propto \phi^{1/2} e^{-\phi \frac{p_0}{2} (\mu - m_0)^2} \times e^{-\phi \frac{n}{2} (\mu - \bar{y})^2} \]

\[ = \phi^{1/2} e^{-\phi \frac{1}{2} \left( p_0 \mu^2 - 2p_0 m_0 \mu + p_0 m_0^2 + n \mu^2 - 2n \bar{y} \mu + n \bar{y}^2 \right)} \]

\[ = \phi^{1/2} e^{-\phi \frac{1}{2} \left( (p_0 + n) \mu^2 - 2(p_0 m_0 + n \bar{y}) \mu + p_0 m_0^2 + n \bar{y}^2 \right)} \]

\[ = \phi^{1/2} e^{-\phi \frac{1}{2} \left[ (p_0 + n) \mu^2 - 2p_n \left( \frac{p_0 m_0^2 + n \bar{y}^2}{p_n} \right) \mu + p_n m_n^2 \right]} \times \]

\[ e^{-\phi \frac{1}{2} \left( p_0 m_0^2 + n \bar{y}^2 - p_n m_n^2 \right)} \]

Read off posterior mean and precision from terms in \[ [\ ] \].

Conjugate Priors for Normal Data – p.7/17
Derivation

Factor and combine terms from earlier

\[ p(\mu, \phi | Y) \propto \phi^{1/2} e^{-\phi \frac{pn}{2} (\mu - m_n)^2} \times \]
\[ \phi^{\frac{v_0+n}{2} - 1} e^{-\phi \left( \frac{SS_0 + SS}{2} \right)} e^{-\frac{\phi}{2} \left( p_0 m_0^2 + n \bar{y}^2 - p_n m_n^2 \right)} \]

The first line is the (unnormalized) posterior density for \( \mu \) given \( \phi \) and the second line is proportional to the posterior density for \( \phi \). The last term comes from the left-over terms after completing the square. Simplifying

\[ p(\phi | Y) \propto \phi^{\frac{v_0+n}{2} - 1} \exp \left\{ -\phi \left( \frac{SS_0 + SS + \frac{p_0 n}{p_n} (\bar{y} - m_0)^2}{2} \right) \right\} \]
Marginal for $\mu$

If $\mu \mid \phi \sim N \left( m, \frac{1}{p\phi} \right)$

$\phi \sim G(v/2, SS/2)$

then

$\mu \sim t \left( v, m, \frac{SS}{v p} \right)$

$\mu \overset{D}{=} m + t_v \sqrt{\frac{SS}{vp}}$

$\frac{\mu - m}{\sqrt{\frac{SS}{vp}}} \sim t(v, 0, 1)$
Example: SPF

Construct an informative prior distribution for $\mu$:

- Take prior median SPF to be 16
- $P(\mu > 64) = 0.01$
- Information in prior is worth 25 observations

Solve for hyperparameters that are consistent with these quantiles: $m_0 = \log(16)$, $p_0 = 25$, $v_0 = p_0 - 1$

$$P(\mu < \log(64)) = 0.99 \quad \text{where} \quad \frac{\mu - m_0}{\sqrt{SS_0/(v_0p_0)}} \sim t_{v_0}$$

$\Rightarrow SS_0 = 185.7$
Posterior Distribution

Summary statistics

- $\bar{y} = 1.998$
- $SS = 6.297$
- $n = 13$

Posterior hyperparameters:

- $p_n = 25 + 13 = 38$
- $m_n = (25 \times 2.773 + 13 \times 1.998)/38 = 2.508$
- $v_n = 24 + 13 = 37$
- $SS_n = 185.7 + 6.297 + (1.998 - 2.773)^2 \times (13 \times 25)/38 = 197.134$
- $\mu \mid Y \sim t(37, 2.508, 197.134/(38 \times 37))$
Samples from the Posterior

To draw samples of SPF from the posterior distribution:

- **Draw** $\phi | Y$
  \[ \phi = \text{rgamma}(10000, \ vn/2, \ \text{rate}=SSn/2) \]

- **Draw** $\mu | \phi, Y$
  \[ \mu = \text{rnorm}(10000, \ mn, \ 1/\sqrt{\phi \ast pn}) \]

- **-or-** **Draw** $\mu | Y$ **directly**
  \[ \mu = \text{rt}(10000, \vn) \ast \sqrt{SSn/(pn \ast vn)} + \ mn \]

- **transform** $\exp(\mu)$

- **HPDinterval**($\exp(\mu)$)
Distributions

Posterior Distribution of $\mu$

Posterior Distribution of $\sigma$
Distribution for SPF

95% HPD Interval 4.54 to 23.758
Reference 95% HPD Interval 4.47 to 10.89
Predictive Distribution

What is the predictive distribution of a new observation \( Y^* \) given the current data \( Y \)? \( p(Y^* \mid Y) \)?

\[
Y^* \mid \mu, \phi \text{ independent of } Y
\]

\[
p(Y^* \mid Y) = \int \int p(Y \mid \mu, \phi) p(\mu, \phi \mid Y) \, d\mu \, d\phi
\]
Integrals

Use “normal trick” to integrate out $\mu$:
If $X \sim \text{N}(m, s^2)$ then $X$ is equal in distribution to $sZ + m$
where $Z \sim \text{N}(0, 1)$

$$Y^* \overset{D}{=} \mu + Z / \sqrt{\phi}$$

Sum of two independent normals:

$$Y^* | \phi, Y \sim \text{N} \left( m_n, \frac{1}{\phi} \left(1 + \frac{1}{p_n}\right) \right)$$

Use previous result about $t$ distributions

$$Y^* | Y \sim t \left( v_n, m_n, \frac{\text{SS}_n}{v_n} \left(1 + \frac{1}{p_n}\right) \right)$$
Predictive Distribution for New Subject

\[ Y^* = \log(\text{TRT}) - \log(\text{BASELINE}) \overset{D}{=} \mu + Z/\phi \]

\[ \text{TRT/BASELINE} = \exp(Y^*) \mid Y \overset{D}{=} \exp\{t(37, 2.5, 5.32 \ast (1 + 1/38))\} \]

- Distribution is the exponential of a Student \( t \)
- Simulate from predictive distribution
- 50% HPD interval is \((0.0003, 12.4)\) from CODA

Predict that with sunscreen there is a 50% chance that the next subject could be exposed from 0 to 12 times longer than without sunscreen.

Prior influence?