Conjugate Priors for Normal Data

Hoff Chapter 5

October 1, 2010

Normal Model

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

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

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

Likelihood

$$\begin{aligned} L(\mu, \phi | Y) &\propto \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \phi^{1/2} \exp\{-\frac{1}{2}\phi(Y_i - \mu)^2\} \\ &\propto \phi^{n/2} \exp\{-\frac{1}{2}\phi \sum_i (Y_i - \mu)^2\} \end{aligned}$$

Likelihood Factorization

$$
\begin{aligned}
L(\mu, \phi | Y) &\propto \phi^{n/2} \exp\{-\frac{1}{2}\phi \sum_i (Y_i - \mu)^2\} \\
&\propto \phi^{n/2} \exp\{-\frac{1}{2}\phi \sum_i \left[(Y_i - \bar{Y}) - (\mu - \bar{Y})\right]^2\} \\
&\propto \phi^{n/2} \exp\{-\frac{1}{2}\phi \left[\sum_i (Y_i - \bar{Y})^2 + n(\mu - \bar{Y})^2\right]\} \\
&\propto \phi^{n/2} \exp\{-\frac{1}{2}\phi s^2 (n-1)\} \exp\{-\frac{1}{2}\phi n(\mu - \bar{Y})^2\}
\end{aligned}
$$

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

Likelihood for Normal Model

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

Sufficient statistics

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

Conjugate Prior Distribution

for (μ, ϕ) is Normal-Gamma.

$$\begin{aligned} \mu|\phi &\sim \mathsf{N}(m_0, 1/(p_0\phi)) \\ \phi &\sim \mathsf{G}(v_0/2, \mathsf{SS}_0/2) \end{aligned}$$

$$p(\mu, \phi) \propto \phi^{v_0/2-1} \exp\{-\phi \frac{\mathsf{SS}_0}{2}\} \phi^{1/2} \exp\{-\phi \frac{p_0}{2}(\mu - m_0)^2\}$$

$\mu, \phi \sim \mathsf{NG}(m_0, p_0, v_0/2, \mathsf{SS}_0/2)$ Normal-Gamma family

$\mu, \phi \mid Y \sim \mathsf{NG}(m_n, p_n, v_n/2, \mathsf{SS}_n/2)$ Posterior is Normal-Gamma

Updating the Posterior Parameters

Under the Normal-Gamma prior distribution:

$$\mu \mid \phi, Y \sim \mathrm{N}\left(m_n, \frac{1}{p_n \phi}\right)$$
$$\phi \mid Y \sim \mathrm{G}(\frac{v_n}{2}, \frac{\mathrm{SS}_n}{2})$$

where

$$p_n = p_0 + n$$
$$m_n = \frac{n\bar{y} + p_0 m_0}{p_n}$$
$$v_n = v_0 + n$$
$$\mathrm{SS}_n = \mathrm{SS}_0 + \mathrm{SS} + \frac{n p_0}{p_n}(\bar{y} - m_0)^2$$

Interpretation

- p_n precision for estimating μ after n observations
- m_n expected value for μ 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

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

- $\mathrm{SS}_n = \mathrm{SS}_0 + \mathrm{SS} + \frac{np_0}{p_n}(\bar{y} - m_0)^2$ posterior variation:
  - prior variation,
  - observed variation (sum of squares),
  - variation between prior mean and sample mean

Derivation

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

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

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

Derivation

Take second line and complete the square:

$$
\begin{aligned}
p(\mu|\phi, Y) &\propto \phi^{1/2} e^{\{-\phi \frac{p_0}{2}(\mu - m_0)^2\}} \times e^{\{-\phi \frac{n}{2}(\mu - \bar{y})^2\}} \\
&= \phi^{1/2} e^{\{-\frac{\phi}{2}\left(p_0\mu^2 - 2p_0 m_0 \mu + p_0 m_0^2 + n\mu^2 - 2n\bar{y}\mu + n\bar{y}^2\right)\}} \\
&= \phi^{1/2} e^{\{-\frac{\phi}{2}\left((p_0 + n)\mu^2 - 2(p_0 m_0 + n\bar{y})\mu + p_0 m_0^2 + n\bar{y}^2\right)\}} \\
&= \phi^{1/2} e^{(-\frac{\phi}{2}\left[(p_0+n)\mu^2 - 2p_n \frac{(p_0 m_0 + n\bar{y})}{p_n}\mu + p_n m_n^2\right])} \times \\
&\qquad e^{(-\frac{\phi}{2}\left(p_0 m_0^2 + n\bar{y}^2 - p_n m_n^2\right))}
\end{aligned}
$$

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

Derivation

Factor and combine terms from earlier

$$p(\mu, \phi | Y) \propto \phi^{1/2} e^{\{-\phi \frac{p_n}{2}(\mu - m_n)^2\}} \times$$
$$\phi^{\frac{v_0+n}{2}-1} e^{\{-\phi\left(\frac{\mathrm{SS}_0+\mathrm{SS}}{2}\right)\}} e^{\{-\frac{\phi}{2}\left(p_0 m_0^2 + n\bar{y}^2 - p_n m_n^2\right)\}}$$

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

$$p(\phi \mid Y) \propto \phi^{\frac{v_0+n}{2}-1} \exp\{-\phi \left(\frac{\mathrm{SS}_0 + \mathrm{SS} + \frac{p_0 n}{p_n}(\bar{y} - m_0)^2}{2}\right)\}$$

Marginal Distribution for $\mu \mid Y$

$$
\begin{aligned}
p(\mu \mid Y) &\propto \int p(\mu, \phi \mid Y) d\phi \\
&= \int \phi^{\frac{v_n+1}{2}-1} \exp[-\phi \left\{ \frac{\text{SS}_n + p_n(\mu - m_n)^2}{2} \right\}] d\phi
\end{aligned}
$$

This has the form of a Gamma integral with $a = (v + 1)/2$ and b equal to the mess multiplying ϕ in the exponential term, so that the result is $\propto b^{-a}$ (at least that is all that matters)

$$
p(\mu|Y) \propto \left\{ \text{SS}_n + p_n(\mu - m_n)^2 \right\}^{\frac{-(v_n+1)}{2}}
$$

Student t Distribution

X has a Student t distribution with location ℓ, scale $\mathcal{S}$ and degrees of freedom δ:

$$X \sim t_\delta(\ell, \mathcal{S}) \Rightarrow p(t; \delta, \ell, \mathcal{S}) \propto \left(\delta + \frac{(x-\ell)^2}{\mathcal{S}}\right)^{(\delta+1)/2}$$

Rearrange posterior distribution:

$$\begin{aligned} p(\mu|Y) &\propto \left\{\mathrm{SS}_n + p_n(\mu - m_n)^2\right\}^{\frac{-(v_n+1)}{2}} \\ &\propto \left(v_n + \frac{(\mu - m_n)^2}{\frac{1}{p_n}\frac{\mathrm{SS}_n}{v_n}}\right)^{-(v_n+1)/2} \end{aligned}$$

Student $t_{v_n}(m_n, s_n^2)$ location m_n, df $= v_n$, scale $s_n^2 = \frac{1}{p_n}\frac{\mathrm{SS}_n}{v_n}$

Standard Student t

Standardize $X \sim t_\delta(\ell, \mathcal{S})$ by subtracting location and dividing by square root of the scale:

$$\frac{X - \ell}{\sqrt{\mathcal{S}}} \sim t_\delta(0, 1)$$

(new location 0 and scale 1)

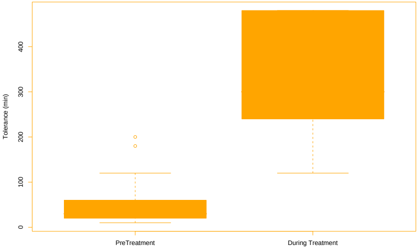
$$\Rightarrow \frac{\mu - m_n}{s_n} \sim t_{v_n}(0, 1)$$

$$\mu \stackrel{D}{=} m_n + t_{v_n} s_n$$

Use `rt`, `qt`, `pt`, `dt` in `R`

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

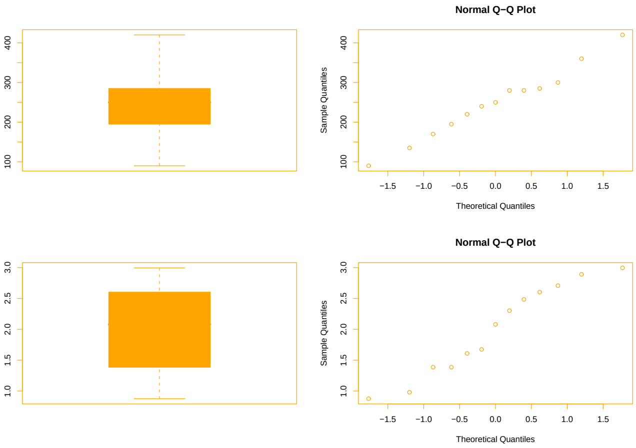
A paired design may be more powerful than two sample design because of patient to patient variability, particularly if there is positive correlation

$$V(d) = V(Y_T) + V(Y_B) - 2\text{Cov}(Y_B, Y_T)$$

- 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


Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
100 200 300 400
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5
Normal Q−Q Plot
Sample Quantiles
Theoretical Quantiles
1.0 1.5 2.0 2.5 3.0
−1.5 −1.0 −0.5 0.0 0.5 1.0 1.5

Model for SPF

- Model `Y = log(TRT) - log(CONTROL)` as $N(\mu, 1/\phi)$
- $\mathrm{E}(\log(\mathrm{TRT}/\mathrm{CONTROL})) = \mu = \log(\mathrm{SPF})$
- Want distribution of $\exp \mu \equiv$SPF
- Summary statistics
  - `ybar = 1.998`
  - `s2 = 0.525`
  - `n = 13`

Informative Prior for SPF

Construct an informative prior distribution for μ:

- 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:
$m_0 = \log(16)$, $p_0 = 25$, $v_0 = p_0 - 1$

$$P(\mu < \log(64)) = 0.99 \text{ where } \frac{\mu - m_0}{\sqrt{SS_0/(v_0 p_0)}} \sim t_{v_0}$$

$\Rightarrow SS_0 = 185.7$

Posterior Distribution

Summary statistics

- `ybar = 1.998`
- `SS = 6.297`
- `n = 13`

Posterior hyperparameters:

$$
\begin{aligned}
p_n &= 25 + 13 = 38 \\
m_n &= (25 * 2.773 + 13 * 1.998)/38 = 2.508 \\
v_n &= 24 + 13 = 37 \\
\mathrm{SS}_n &= 185.7 + 6.297 + (1.998 - 2.773)^2 * (13 * 25)/38 = 197.134
\end{aligned}
$$

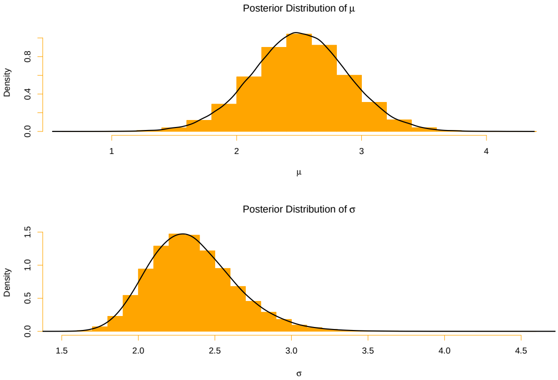
$$\mu \mid Y \sim t(37, 2.508, 197.134/(38 * 37))$$

Samples from the Posterior

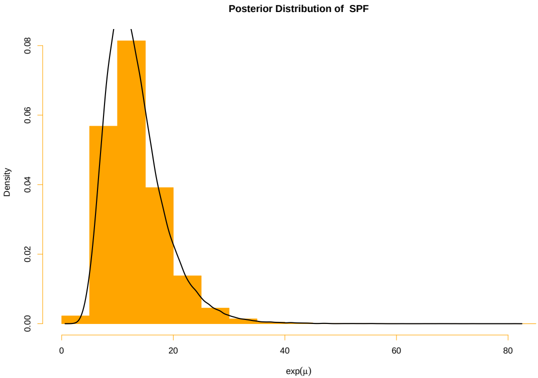
To draw samples of SPF from the posterior distribution:

- Draw $\phi | Y$
  `phi = rgamma(10000, vn/2, rate=SSn/2)`
- Draw $\mu | \phi, Y$
  `mu = rnorm(10000, mn, 1/sqrt(phi*pn))`
- –or– Draw $\mu | Y$ directly
  `mu = rt(10000,vn)*sqrt(SSn/(pn*vn))+ mn`
- transform `exp(mu)`
- `HPDinterval(exp(mu))`

Distributions



Distribution for SPF



95% HPD Interval 4.54 to 23.758

Reference 95% HPD Interval 4.47 to 10.89