STA 365: Applied Bayesian Statistics

Boris Babic
Assistant Professor, University of Toronto

Week 5B: Normal Models

Normal Model: Unknown Mean and Variance

Let $y_1, \ldots, y_n \overset{\text{iid}}{\sim} \mathrm{N}(\mu, \sigma^2)$.

For this model we will consider two different prior settings:

- Improper, noninformative prior
- Conjugate prior

- Note that this is a univariate normal model with two unknown parameters, the scale μ, and the variance, σ^2.
- This is a multiparameter model, but a univariate model. We will look at multivariate models later.

Improper, Uninformative Prior

Non-informative prior

$$\pi(\mu, \sigma^{-2}) \propto (\sigma^{-2})^{-1}.$$

- This, recall, is the joint Jeffreys' prior assuming independence of the scale and location parameters.
- Hence, the Jeffreys' prior for the precision is proportional to the variance
- And conversely, the Jeffreys' prior for variance is proportional to the precision.
- $\pi(\mu|\sigma^2) \propto 1$ and $\pi(\sigma^{-2}) \propto (\sigma^{-2})^{-1} = \sigma^2$
- Hence, $\pi(\mu, \sigma^{-1}) = \pi(\mu|\sigma^{-2}) \times \pi(\sigma^{-2}) = 1 \times (\sigma^{-2})^{-1} = (\sigma^{-2})^{-1}$
- Another way to think about this: $\pi(\sigma) \propto 1/\sigma$ and $\pi(\log(\sigma)) \propto 1$
- Hence, we will sometimes adopt a uniform prior on $(\mu, \log \sigma)$.

Joint Posterior

The joint posterior distribution,

$$\pi(\mu, \sigma^{-2} \mid y) \propto (\sigma^{-2})^{n/2-1} \exp\left(-\frac{\sigma^{-2}}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2]\right),$$

where $\bar{y} = n^{-1}\sum_{i=1}^{n} y_i$ and $s^2 = (n-1)^{-1}\sum_{i=1}^{n}(y_i - \bar{y})^2$

- The joint posterior is proportional to the normal likelihood multiplied by the factor σ^2.

Student's t Distribution

- Suppose $Z \sim \mathrm{N}(0,1)$ and $V \sim \mathrm{G}(n/2, 1/2)$ or $\chi^2(n)$. And Z and V are independent. Let
$$T = \frac{Z}{\sqrt{V/n}}.$$
- Then T has probability density function
$$f(t) = \frac{\Gamma\{(n+1)/2\}}{\sqrt{n\pi}\Gamma(n/2)} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$
This is the student's t distribution with n degrees of freedom.
- Let $X = \mu + \sigma T$. Then X follows a student's t distribution with mean μ, scale parameter σ^2 and n degrees of freedom, denoted
$$X \sim t_n(\mu, \sigma^2).$$
What is the density of X: $f(x)$?
$$f(x) = \frac{\Gamma\{(n+1)/2\}}{\sqrt{n\pi\sigma^2}\Gamma(n/2)} \left(1 + \frac{(x-\mu)^2}{n\sigma^2}\right)^{-\frac{n+1}{2}}.$$

Scale Mixture of Normal Distributions

Let

$$(\theta \mid \mu, \sigma^2, k) \sim \mathrm{N}(\mu, \sigma^2/k),$$
$$k \sim \mathrm{G}(\nu/2, \nu/2),$$

Then

$$\theta \sim t_\nu(\mu, \sigma^2).$$

How to show this?

$$\begin{aligned}
&\pi(\theta \mid \mu, \sigma^2) \\
&= \int_0^\infty \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} k^{\nu/2-1} \exp\left[-(\nu/2)k\right] (2\pi\sigma^2/k)^{-1/2} \exp\left[-0.5k(\theta-\mu)^2/\sigma^2\right] \\
&= \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)\sqrt{2\pi\sigma^2}} \times \int_0^\infty k^{(v+1)/2-1} \exp\left[-0.5(\nu + (\theta-\mu)^2/\sigma^2)\right] dk \\
&= \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)\sqrt{2\pi\sigma^2}} \times \frac{\Gamma((\nu+1)/2)}{(\nu/2 + ((\theta-\mu)/\sigma)^2/2)^{(\nu+1)/2}}.
\end{aligned}$$

The Marginal Distributions

$$
\begin{aligned}
\pi(\mu \mid y) &\propto \int_0^\infty (\sigma^{-2})^{n/2-1} \exp\left(-\frac{\sigma^{-2}}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2]\right) d\sigma^{-2} \\
&= \Gamma(n/2)\left(\frac{(n-1)s^2 + n(\bar{y}-\mu)^2}{2}\right)^{-n/2} \\
&\propto \left[1 + \frac{1}{n-1}\left(\frac{\bar{y}-\mu}{s/\sqrt{n}}\right)\right]^{-n/2}.
\end{aligned}
$$

This implies that

$$\mu \mid y \sim t_{n-1}(\mu, s/\sqrt{n}).$$

And

$$
\begin{aligned}
\pi(\sigma^{-2} \mid y) &\propto \int_0^\infty (\sigma^{-2})^{n/2-1} \exp\left(-\frac{\sigma^{-2}}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2]\right) d\mu \\
&\propto (\sigma^{-2})^{(n-1)/2-1} \exp\left\{-\frac{\sigma^{-2}}{2}(n-1)s^2\right\}.
\end{aligned}
$$

Thus,

$$\sigma^{-2} \mid y \sim \mathrm{G}\{(n-1)/2, (n-1)s^2/2\}.$$

Posterior Predictive Distributions

Let

$$\tilde{y} \sim \mathrm{N}(\mu, \sigma^2).$$

The posterior predictive distribution of $\tilde{y}$ is given by

$$\tilde{y} \mid y \sim t_{n-1}\left\{\bar{y}, \left(1 + \frac{1}{n}\right) s^2\right\}.$$

How to show this? Let $k = s^2/\sigma^2$.

$$\tilde{y} \mid k, y \sim \mathrm{N}\left\{\bar{y}, \left(1 + \frac{1}{n}\right) s^2/k\right\}.$$

$$k \mid y \sim \mathrm{G}\{(n-1)/2, (n-1)/2\}.$$

Then the results are followed by the property of the scale mixture of normal distributions

Conjugate Prior

To derive the conjugate priors for μ, σ^{-2}, we consider the functional form of the likelihood

$$\pi(y \mid \mu, \sigma^{-2}) \propto (\sigma^{-2})^{n/2} \exp\left(-\frac{\sigma^{-2}}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2]\right)$$

how about

$$\pi(y \mid \sigma^{-2}) \propto (\sigma^{-2})^{n/2} \exp\left(-\frac{\sigma^{-2}}{2}[(n-1)s^2]\right)$$

This implies that the conjugate priors should be

$$\sigma^{-2} \sim \mathrm{G}\left(\nu/2, \nu\tau^2/2\right).$$

and

$$\mu \mid \sigma^{-2} \sim \mathrm{N}(\theta, \sigma^2/k).$$

Joint Posterior

$$
\begin{aligned}
\pi(\mu, \sigma^{-2} \mid y) \quad &\propto \quad (\sigma^{-2})^{\nu/2-1} \exp\left(\frac{\nu\tau^2}{2}\sigma^{-2}\right) (\sigma^{-2})^{1/2} \exp\left\{-\frac{k\sigma^{-2}}{2}(\mu-\theta)^2\right\} \\
&\quad \times (\sigma^{-2})^{n/2} \exp\left(-\frac{\sigma^{-2}}{2}\sum_i (y_i-\mu)^2\right) \\
&= \quad (\sigma^{-2})^{\frac{v+n+1}{2}-1} \exp\left(-\frac{\sigma^{-2}}{2}[\nu\tau^2 + k(\mu-\theta)^2]\right) \\
&\quad \times \exp\left\{-\frac{\sigma^{-2}}{2}[(n-1)s^2 + n(\bar{y}-\mu)^2]\right\}.
\end{aligned}
$$

Marginal Posterior of σ^{-2}

$$
\begin{aligned}
\pi(\sigma^{-2} \mid y) &= \int_{-\infty}^{\infty} \pi(\mu, \sigma^{-2} \mid y) d\mu \\
&\propto \underbrace{(\sigma^{-2})^{\frac{\nu+n+1}{2}-1} \exp\left\{-\frac{\sigma^{-2}}{2}[\nu\tau^2 + (n-1)s^2]\right\}}_{A} \\
&\quad \times \int_{-\infty}^{\infty} \exp\left\{-\frac{\sigma^{-2}}{2}[k(\mu-\theta)^2 + n(\bar{y}-\mu)^2]\right\} d\mu \\
&\propto A \exp\left[-\frac{\sigma^2}{2}(k\theta^2 + n\bar{y}^2)\right] \int_{-\infty}^{\infty} \exp\left\{-\frac{(k+n)\sigma^{-2}}{2}\left[\mu^2 - 2\left(\frac{k\theta+n\bar{y}}{k+n}\right)\mu\right]\right\} \\
&= A \exp\left[-\frac{\sigma^2}{2}(k\theta^2 + n\bar{y}^2)\right] \exp\left[\frac{\sigma^{-2}}{2}(k+n)\left(\frac{k\theta+n\bar{y}}{k+n}\right)^2\right] \\
&\quad \times \int_{-\infty}^{\infty} \exp\left[-\frac{(k+n)\sigma^{-2}}{2}\left(\mu - \frac{k\theta+n\bar{y}}{k+n}\right)^2\right] d\mu \\
&= A \exp\left\{-\frac{\sigma^2}{2}\left[k\theta^2 + n\bar{y}^2 - \frac{(k\theta+n\bar{y})^2}{k+n}\right]\right\} (2\pi)^{1/2}[(k+n)\sigma^{-2}]^{-1/2} \\
&\propto (\sigma^{-2})^{\frac{\nu+n}{2}-1} \exp\left\{-\frac{\sigma^2}{2}\left[\nu\tau^2 + (n-1)s^2 + \frac{kn}{k+n}(\bar{y}-\theta)^2\right]\right\}.
\end{aligned}
$$

Marginal Posterior of σ^{-2}

This is a kernel of a gamma distribution. Thus,

$$[\sigma^{-2} \mid y] \sim G\left(\frac{\nu+n}{2}, \frac{\nu\tau^2+(n-1)s^2+\frac{kn}{k+n}(\bar{y}-\theta)^2}{2}\right).$$

Marginal Posterior of μ

The full conditional of μ is given by

$$\begin{aligned}
\pi(\mu \mid \sigma^2, y) &\propto \exp\left[-\frac{k\sigma^{-2}}{2}(\mu-\theta)^2 - \frac{\sigma^{-2}}{2}\sum_i (y_i-\mu)^2\right] \\
&\propto \exp\left[-\frac{k\sigma^{-2}}{2}(\mu^2-2\theta\mu) - \frac{\sigma^{-2}}{2}(-2n\bar{y}\mu + n\mu^2)\right] \\
&= \exp\left\{-\frac{\sigma^{-2}}{2}(k+n)\left[\mu^2 - 2\left(\frac{k\theta+n\bar{y}}{k+n}\right)\mu\right]\right\}.
\end{aligned}$$

This implies that

$$[\mu \mid \sigma^{-2}, y] \sim N\left(\frac{k\theta+n\bar{y}}{k+n}, \frac{\sigma^2}{k+n}\right).$$

The marginal posterior distribution of μ is

$$[\mu \mid y] \sim t_{\nu+k}\left(\frac{k\theta+n\bar{y}}{k+n}, \frac{\nu\tau^2+(n-1)s^2+\frac{kn}{k+n}(\bar{y}-\theta)^2}{(\nu+n)(k+n)}\right).$$