

**SDS 383D Ex 02:**  
**Bayes and the Gaussian Linear Model**

January 18, 2016

**Jennifer Starling**

## A Simple Gaussian Location Model

### Part A

The marginal prior  $p(\theta)$  is a gamma mixture of normals. Show it takes the form of a centered, scaled t distribution:  $p(\theta) \propto \left(1 + \frac{1}{\nu} \cdot \frac{(x-m)^2}{s^2}\right)^{-\frac{\nu+1}{2}}$ .

$$(\theta|\omega) \sim N(\mu, \omega^{-1}k^{-1}) \text{ and } \omega \sim \text{gamma}\left(\frac{d}{2}, \frac{\eta}{2}\right)$$

Joint prior for  $p(\theta, \omega)$  has form

$$p(\theta, \omega) = p(\theta|\omega)p(\omega) \propto \omega^{\left(\frac{(d+1)}{2}-1\right)} \exp\left[-\omega \cdot \frac{k(\theta-\mu)^2}{2}\right] \exp\left[-\omega \frac{\eta}{2}\right]$$

Marginal prior for  $\theta$  is calculated as follows:

$$\begin{aligned} p(\theta) &= \int_0^\infty p(\theta, \omega) \delta\omega \\ &= \int_0^\infty \omega^{\left(\frac{(d+1)}{2}-1\right)} \exp\left[-\omega \cdot \frac{k(\theta-\mu)^2}{2}\right] \exp\left[-\omega \frac{\eta}{2}\right] \delta\omega \\ &= \int_0^\infty \omega^{\left(\frac{(d+1)}{2}-1\right)} e^{-\omega\left(\frac{k(\theta-\mu)^2}{2} + \frac{\eta}{2}\right)} \delta\omega \end{aligned}$$

This is the kernel of a  $\text{gamma}\left(\frac{(d+1)}{2}, \frac{k(\theta-\mu)^2}{2} + \frac{\eta}{2}\right)$  distribution.

The integral will therefore equal  $\frac{1}{c}$ , where  $c$  is the constant of proportionality for this gamma density,  $\frac{1}{\beta^\alpha/\Gamma(\alpha)} = \frac{\Gamma(\alpha)}{\beta^\alpha}$ .

$$\begin{aligned} &= \Gamma\left(\frac{d+1}{2}\right) \left[\frac{k(\theta-\mu)^2}{2} + \frac{\eta}{2}\right]^{-\frac{d+1}{2}} \\ &\propto \left[\left(\frac{\eta}{2}\right) \left(1 + \frac{k(\theta-\mu)^2}{\eta}\right)\right]^{-\frac{d+1}{2}} \\ &= \left(\frac{\eta}{2}\right)^{-\frac{d+1}{2}} \left[1 + \frac{k(\theta-\mu)^2}{\eta}\right]^{-\frac{d+1}{2}} \end{aligned}$$

The left term cancels, as it is constant wrt  $\theta$ .

$$\propto \left[1 + \frac{k(\theta-\mu)^2}{\eta}\right]^{-\frac{d+1}{2}}$$

Multiply numerator and denominator by  $d$ .

$$\begin{aligned} &= \left[1 + \frac{dk(\theta-\mu)^2}{d\eta}\right]^{-\frac{d+1}{2}} \\ &= \left[1 + \left(\frac{1}{d}\right) \frac{(\theta-\mu)^2}{\frac{\eta}{dk}}\right]^{-\frac{d+1}{2}} \end{aligned}$$

This has the desired central-scaled t form, with mean  $m = \mu$ , df  $\nu = d$ , and scale  $s^2 = \frac{\eta}{dk}$ . ■

## Part B

Assume the sampling model  $(y_i|\theta, \sigma^2) \sim N(\theta, \sigma^2)$  for  $i = 1 \dots n$ , so  $y = (y_1, \dots, y_n)^T$ . Work with precision  $\omega = 1/\sigma^2$ .

Assume the same normal-gamma prior has form

$$p(\theta, \omega) \propto \omega^{\left(\frac{d+1}{2}-1\right)} \exp\left[-\omega \cdot \frac{k(\theta - \mu)^2}{2}\right] \exp\left[-\omega \frac{\eta}{2}\right]$$

Calculate the joint posterior up to a constant of proportionality (factors not dependent on  $\theta$  or  $\omega$ .) Show the joint posterior is also normal-gamma.

The posterior  $\propto$  sampling model \* prior, so

$$p(\theta, \omega|y) \propto p(y|\theta, \omega)p(\theta, \omega)$$

Sampling model (from the normal likelihood), using  $S_y = \sum_{i=1}^n (y_i - \bar{y})^2$ , is

$$p(y|\theta, \omega) \propto \omega^{\frac{n}{2}} \exp\left[-\frac{\omega}{2} \left(S_y + n(\bar{y} - \theta)^2\right)\right]$$

Therefore, the joint posterior is

$$\begin{aligned} p(\theta, \omega|y) &\propto \omega^{\left(\frac{d+1}{2}-1\right)} \exp\left[-\omega \cdot \frac{k(\theta - \mu)^2}{2}\right] \exp\left[-\omega \frac{\eta}{2}\right] \cdot \omega^{\frac{n}{2}} \exp\left[-\frac{\omega}{2} \left(S_y + n(\bar{y} - \theta)^2\right)\right] \\ &= \omega^{\left(\frac{n+d+1}{2}-1\right)} \cdot \exp\left[-\frac{\omega}{2} \left(S_y + n(\bar{y} - \theta)^2 + k(\theta - \mu)^2 + \eta\right)\right] \end{aligned}$$

Simplify the exponential term multiplied by  $-\frac{\omega}{2}$ .

$$\begin{aligned} &\left(S_y + n(\bar{y} - \theta)^2 + k(\theta - \mu)^2 + \eta\right) \\ &= S_y + n(\bar{y} - \theta)(\bar{y} - \theta) + k(\theta - \mu)(\theta - \mu) + \eta \\ &= S_y + n\bar{y}^2 - 2n\bar{y}\theta + n\theta^2 + k\theta^2 - 2k\theta\mu + k\mu^2 + \eta \\ &= (n+k)\theta^2 - 2(n\bar{y} + k\mu)\theta + \left(S_y + n\bar{y}^2 + k\mu^2 + \eta\right) \end{aligned}$$

This has the form  $ax^2 - 2bx + c$ . We can complete the square as follows:

$$\begin{aligned} &ax^2 - 2bx + c \\ &= a \left[ x^2 - 2\frac{b}{a}x + \frac{c}{a} \right] \\ &= a \left[ x^2 - 2\frac{b}{a}x + \left(\frac{b}{a}\right)^2 - \left(\frac{b}{a}\right)^2 + \frac{c}{a} \right] \\ &= a \left[ \left(x - \frac{b}{a}\right)^2 - \left(\frac{b}{a}\right)^2 + \frac{c}{a} \right] \\ &= a \left(x - \frac{b}{a}\right)^2 - \frac{b^2}{a} + c \end{aligned}$$

Plugging in the appropriate a, b and c terms yields:

$$\begin{aligned}
 &= (n+k) \left[ \theta - \frac{n\bar{y} + k\mu}{n+k} \right]^2 - \frac{(n\bar{y} + k\mu)^2}{n+k} + S_y + n\bar{y}^2 + k\mu^2 + \eta \\
 &= (n+k) \left[ \theta - \frac{n\bar{y} + k\mu}{n+k} \right]^2 + \frac{nk(\bar{y} - \mu)^2}{(n+k)} + S_y + \eta
 \end{aligned}$$

Plug this rearranged term back into the entire posterior:

$$\begin{aligned}
 p(\theta, \omega | y) &\propto \omega^{\left(\frac{n+d+1}{2}-1\right)} \cdot \exp \left[ -\frac{\omega}{2} \left( (n+k) \left[ \theta - \frac{n\bar{y} + k\mu}{n+k} \right]^2 + \frac{nk(\bar{y} - \mu)^2}{(n+k)} + S_y + \eta \right) \right] \\
 p(\theta, \omega | y) &\propto \omega^{\left(\frac{n+d+1}{2}-1\right)} \cdot \exp \left[ -\frac{\omega}{2} \left( (n+k) \left[ \theta - \frac{n\bar{y} + k\mu}{n+k} \right]^2 \right) \right] \cdot \exp \left[ -\frac{\omega}{2} \left( \frac{nk(\bar{y} - \mu)^2}{(n+k)} + S_y + \eta \right) \right]
 \end{aligned}$$

This is the desired form of the normal-gamma distribution, with posterior parameters as follows.

$$d^* = (n + d) \tag{1}$$

$$k^* = (n + k) \tag{2}$$

$$\mu^* = \frac{n\bar{y} + k\mu}{(n+k)} \tag{3}$$

$$\eta^* = \frac{nk(\bar{y} - \mu)^2}{(n+k)} + S_y + \eta \tag{4}$$

## Part C

The conditional posterior of  $p(\theta | y, \omega)$  is

$$p(\theta | y, \omega) \propto \exp \left[ -\frac{\omega}{2} \cdot (n+k) \left( \theta - \frac{n\bar{y} + k\mu}{(n+k)} \right)^2 \right]$$

This is the Normal distribution form, parameterized with precision.

$$p(\theta | y, \omega) \sim N \left( \frac{n\bar{y} + k\mu}{(n+k)}, -\omega(n+k) \right)$$

**Part D**

From the joint posterior in Part A, what is the marginal posterior  $p(\omega|y)$ ?

Note: Will indicate posterior parameters with a \*, for notational simplicity, and will substitute values of posterior parameters at end.

$$\begin{aligned} p(\omega|y) &= \int_{-\infty}^{\infty} p(\theta, \omega|y) \delta\theta \\ &\propto \int_{-\infty}^{\infty} \omega^{\left(\frac{(d^*+1)}{2}-1\right)} \exp\left[-\frac{\omega}{2}k^*(\theta - \mu^*)^2\right] \exp\left[-\omega\frac{\eta^*}{2}\right] \delta\theta \\ &= \omega^{\left(\frac{(d^*+1)}{2}-1\right)} e^{\left(-\omega\frac{\eta^*}{2}\right)} \int_{-\infty}^{\infty} \exp\left[-\frac{\omega}{2}k^*(\theta - \mu^*)^2\right] \delta\theta \end{aligned}$$

The integral is a normal kernel:  $N(\mu^*, \omega k^*)$ . It integrates to  $\frac{1}{c}$ , where c is the constant of proportionality for the normal density. The integral =  $\frac{1}{\omega^{1/2}} = \omega^{-\frac{1}{2}}$ .

$$\begin{aligned} &= \omega^{\left(\frac{(d^*+1)}{2}-1\right)} e^{\left(-\omega\frac{\eta^*}{2}\right)} \omega^{-\frac{1}{2}} \\ &= \omega^{\left(\frac{d^*}{2}-1\right)} e^{\left(-\omega\frac{\eta^*}{2}\right)} \end{aligned}$$

This is the kernel for  $\text{gamma}\left(\frac{d^*}{2}, \frac{\eta^*}{2}\right)$ , where

$$\begin{aligned} d^* &= (n + d) \\ \eta^* &= \frac{nk(\bar{y} - \mu)^2}{(n + k)} + S_y + \eta \end{aligned}$$

**Part E**

From C and D, we know the marginal posterior  $p(\theta|y)$  is a gamma mixture of normals. Show this takes the form of a centered, scaled t distribution of form  $p(\theta) \propto \left(1 + \frac{1}{v} \cdot \frac{(x-m)^2}{s^2}\right)^{-\frac{v+1}{2}}$  and state what m, s and v are.

Joint posterior  $p(\theta, \omega|y)$ :

$$p(\theta, \omega|y) \propto \omega^{\left(\frac{n+d+1}{2}-1\right)} \cdot \exp \left[ -\frac{\omega}{2} \left( (n+k) \left[ \theta - \frac{n\bar{y} + k\mu}{n+k} \right]^2 \right) \right] \cdot \exp \left[ -\frac{\omega}{2} \left( \frac{nk(\bar{y} - \mu)^2}{(n+k)} + S_y + \eta \right) \right]$$

Marginal posterior  $p(\theta|y)$ :

$$\begin{aligned} p(\theta|y) &= \int_0^\infty p(\theta, \omega|y) \delta\omega \\ &= \int_0^\infty \omega^{\frac{d^*+1}{2}-1} \cdot \exp \left[ -\frac{\omega}{2} \left( k^*(\theta - \mu^*)^2 + \frac{\eta^*}{2} \right) \right] \delta\omega \end{aligned}$$

This is the kernel of a  $\text{gamma} \left( \frac{(d^*+1)}{2}, \frac{k^*(\theta - \mu^*)^2 + \eta^*}{2} \right)$  distribution.

The integral will therefore equal  $\frac{1}{c}$ , where c is the constant of proportionality for this gamma density,  $\frac{1}{\beta^\alpha / \Gamma(\alpha)} = \frac{\Gamma(\alpha)}{\beta^\alpha}$ .

$$\begin{aligned} &= \Gamma\left(\frac{d^*+1}{2}\right) \left[ \frac{k^*(\theta - \mu^*)^2}{2} + \frac{\eta^*}{2} \right]^{-\frac{d^*+1}{2}} \\ &\propto \left[ \left( \frac{\eta^*}{2} \right) \left( 1 + \frac{k^*(\theta - \mu^*)^2}{\eta^*} \right) \right]^{-\frac{d^*+1}{2}} \\ &= \left( \frac{\eta^*}{2} \right)^{-\frac{d^*+1}{2}} \left[ \left( 1 + \frac{k^*(\theta - \mu^*)^2}{\eta^*} \right) \right]^{-\frac{d^*+1}{2}} \end{aligned}$$

The left term cancels, as it is constant wrt  $\theta$ .

$$\propto \left[ \left( 1 + \frac{k^*(\theta - \mu^*)^2}{\eta^*} \right) \right]^{-\frac{d^*+1}{2}}$$

Multiply numerator and denominator by  $d^*$ .

$$\begin{aligned} &= \left[ \left( 1 + \frac{d^*k^*(\theta - \mu^*)^2}{d\eta^*} \right) \right]^{-\frac{d^*+1}{2}} \\ &= \left[ \left( 1 + \left( \frac{1}{d^*} \right) \frac{(\theta - \mu^*)^2}{\frac{\eta^*}{d^*k^*}} \right) \right]^{-\frac{d^*+1}{2}} \end{aligned}$$

This has the desired central-scaled t form, with mean  $m = \mu^*$ , df  $\nu = d^*$ , and scale  $s^2 = \frac{\eta^*}{d^*k^*}$ . ■

**Part G**

True or False? In the limit, as prior parameters  $k, d, \eta \rightarrow 0$ , priors  $p(\theta)$  and  $p(\omega)$  are valid probability distributions. (Must integrate to 1, or something finite so can be normalized over their domains.)

$p(\theta)$ :

The prior  $p(\theta)$  is a gamma mixture of normals, which can arrange as a scaled, centered t.

$$\lim_{d,k,\eta \rightarrow 0} \Gamma\left(\frac{d+1}{2}\right) \left[ \frac{k(\theta - \mu)^2}{2} + \frac{\eta}{2} \right]^{-\frac{d+1}{2}}$$

Goes to:

$$\Gamma\left(\frac{1}{2}\right) \left[ \frac{0(\theta - \mu)^2}{2} + \frac{0}{2} \right]^{-\frac{1}{2}} \rightarrow 0$$

This is not a valid probability distribution. It is a point mass at 0. It is an improper prior when the prior parameters go to zero.

$p(\omega)$ :

The prior  $p(\omega)$  is  $gamma(\frac{d}{2}, \frac{\eta}{2})$ .

$$p(\omega) = \frac{(\frac{\eta}{2})^{(\frac{d}{2})}}{\Gamma(\frac{d}{2})} \omega^{\frac{d}{2}-1} e^{-\omega \frac{\eta}{2}}$$

$$\lim_{d,k,\eta \rightarrow 0} \frac{(\frac{\eta}{2})^{(\frac{d}{2})}}{\Gamma(\frac{d}{2})} \omega^{\frac{d}{2}-1} e^{-\omega \frac{\eta}{2}}$$

Goes to:

$$(1)(\omega^{-1})(1) = \omega^{-1} = \omega^{0-1} e^{-0\omega}$$

This is the form of  $gamma(0,0)$ , which is not a valid probability distribution, since gamma limits both parameters to values greater than zero. This is also an improper prior.

The answer is False for both prior distributions; neither are proper probability distributions.

**Part H**

True or False? In the limit, as prior parameters  $k, d, \eta \rightarrow 0$ , priors  $p(\theta|y)$  and  $p(\omega|y)$  are valid probability distributions. (Must integrate to 1, or something finite so can be normalized over their domains.)

As the hyperparameters go to zero, the posterior parameters are as follows.

$$d^* = (n + d) \rightarrow d^* = n \quad (5)$$

$$k^* = (n + k) \rightarrow k^* = n \quad (6)$$

$$\mu^* = \frac{n\bar{y} + k\mu}{(n + k)} \rightarrow \mu^* = \bar{y} \quad (7)$$

$$\eta^* = \frac{nk(\bar{y} - \mu)^2}{(n + k)} + S_y + \eta \rightarrow \eta^* = S_y \quad (8)$$

$p(\theta|y)$ :

The marginal posterior  $p(\theta|y)$  is also a gamma mixture of normals, which can be written as a scaled, centered t. Plug the updated posterior parameters into the central t form.

$$\left[ \left( 1 + \left( \frac{1}{d^*} \right) \frac{(\theta - \mu^*)^2}{\frac{\eta^*}{d^* k^*}} \right) \right]^{-\frac{d^*+1}{2}} \rightarrow \left[ \left( 1 + \left( \frac{1}{n} \right) \frac{(\theta - \bar{y})^2}{\frac{S_y}{n^2}} \right) \right]^{-\frac{n+1}{2}}$$

This is a valid probability distribution. It has form of the centered, scaled t distribution, with parameters

$$\begin{aligned} m &= \bar{y} \\ s^2 &= \frac{S_y}{n} = \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n^2} \\ \nu &= n \end{aligned}$$

$p(\omega|y)$ :

The marginal posterior  $p(\omega|y)$  is  $\text{gamma}\left(\frac{d^*}{2}, \frac{\eta^*}{2}\right)$ . Plug the updated posterior parameters into the gamma density.

$$p(\omega|y) = \frac{\left(\frac{\eta^*}{2}\right)^{\left(\frac{d^*}{2}\right)}}{\Gamma\left(\frac{d^*}{2}\right)} \omega^{\frac{d^*}{2}-1} e^{-\omega \frac{\eta^*}{2}} \rightarrow p(\omega|y) = \frac{\left(\frac{S_y}{2}\right)^{\left(\frac{n}{2}\right)}}{\Gamma\left(\frac{n}{2}\right)} \omega^{\frac{n}{2}-1} e^{-\omega \frac{S_y}{2}}$$

This is a valid probability distribution. It is  $\text{Gamma}\left(\frac{n}{2}, \frac{S_y}{2}\right)$ .

The answer is True for both posterior marginal distributions. Both are valid probability distributions. Moreover, their parameters are intuitive based on priors assuming zero information.



**Part I**

The Bayesian Credible Interval has the form  $\theta \in m \pm t^* \cdot s$  where  $m$  and  $s$  are the posterior mean and scale, and  $t^*$  is the appropriate  $t$  critical value using the posterior degrees of freedom:  $t^* = t_{\nu^*, 1 - \frac{\alpha}{2}}$ .

True or False? In the limit, as prior parameters  $k, d, \eta \rightarrow 0$ , the Bayes Credible Interval equals the frequentist confidence interval for  $\theta$ .

This is true. We can use the marginal posterior of  $\theta|y$  from Part E, which is centered and scaled  $t$ .

$$p(\theta|y) \sim t\left(m = \mu^*, \nu = d^*, s^2 = \frac{\eta^*}{d^*k^*}\right)$$

where, in the limit, the posterior parameters go to

$$d^* = (n + d) \rightarrow d^* = n$$

$$k^* = (n + k) \rightarrow k^* = n$$

$$\mu^* = \frac{n\bar{y} + k\mu}{(n + k)} \rightarrow \mu^* = \bar{y}$$

$$\eta^* = \frac{nk(\bar{y} - \mu)^2}{(n + k)} + S_y + \eta \rightarrow \eta^* = S_y$$

The Bayes Credible Interval goes to

$$m \pm t^* \cdot s \quad \rightarrow \quad \bar{y} \pm t^* \cdot \sqrt{\frac{S_y}{n^2}}$$

THIS IS CLOSE - but there is an extra  $n$  hanging out in denominator. Look into this.

## **Appendix: R Code**