Ch 2 Bayes and the Gaussian linear model

Wednesday, February 13, 2019 1:02 PM

A) The marginal prior distribution $p(\theta)$ is a gamma mixture of normals. Show that this takes the form of a centered, scaled t distribution

$$p(\theta) \propto \left(1+\frac{1}{v}\cdot\frac{(x-m)^2}{s^2}\right)^{-\frac{v+1}{2}}$$

with center m, scale s, and degrees of freedom v, where you fill in the blank for m, s^2, and v in terms of the four parameters of the normal gamma family.

$$p(\theta) = \int p(\theta,\omega)\,d\omega$$

$$= \int_0^\infty \frac{\sqrt{\omega k}}{\sqrt{2\pi}} \exp\left(-\frac{\omega k}{2}(\theta-\mu)^2\right)\frac{(n/2)^{d/2}}{\Gamma(d/2)}\,\omega^{d/2-1}\exp\left(\frac{-\omega n}{2}\right)d\omega$$

$$= \frac{\sqrt{k}}{\sqrt{2\pi}}\frac{(n/2)^{d/2}}{\Gamma(d/2)}\int_0^\infty \underbrace{\omega^{\frac{(d+1)}{2}-1}\exp\left(-\frac{\omega}{2}\left(k(\theta-\mu)^2+n\right)\right)}_{\text{kernel of gamma}\left(\frac{d+1}{2},\ \frac{k(\theta-\mu)^2+n}{2}\right)}d\omega$$

$$= \frac{\sqrt{k}}{\sqrt{2\pi}}\frac{(n/2)^{d/2}}{\Gamma(d/2)}\frac{\Gamma\left(\frac{d+1}{2}\right)}{\left(\frac{k(\theta-\mu)^2+n}{2}\right)^{\frac{d+1}{2}}}$$

$$\propto \left(\frac{n}{2}\right)^{d/2}\left(\frac{k(\theta-\mu)^2+n}{2}\right)^{-\left(\frac{d+1}{2}\right)}$$

$$= \frac{\left(\frac{n}{2}\right)^{d/2}\Big/\left(\frac{n}{2}\right)^{\frac{d+1}{2}}}{\left(\frac{n}{2}+\frac{k(\theta-\mu)^2}{2}\right)^{\frac{d+1}{2}}\Big/\left(\frac{n}{2}\right)^{\frac{d+1}{2}}}$$

$$= \left(\frac{n}{2}\right)^{-1/2}\left(1+\frac{\frac{k(\theta-\mu)^2}{2}}{\frac{n}{2}}\right)^{-\left(\frac{d+1}{2}\right)}$$

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

$$= \left(1+\frac{1}{d}\frac{(\theta-\mu)^2}{\left(\sqrt{\frac{n}{kd}}\right)^2}\right)^{-\left(\frac{d+1}{2}\right)}$$

B) Assume a normal sampling model and the normal gamma prior (dependent). Calculate the joint posterior density $p(\theta,\omega \mid y)$, up to constant factors not depending on ω or θ. Show that this is also a normal/gamma prior in the same form as above:

$$p(\theta,\omega\mid y) \propto \omega^{(d^*+1)/2-1}\exp\left\{-\omega\cdot\frac{k^*(\theta-\mu^*)^2}{2}\right\}\exp\left\{-\omega\frac{n^*}{2}\right\}$$

Let $S_y = \sum_{i=1}^{n}(y_i-\bar{y})^2$

Let $(y_i \mid \theta,\sigma^2) \sim N(\theta,\sigma^2),\ i=1,\dots,n$

Priors: $(\theta\mid\omega) \sim N(\mu, (\omega k)^{-1})$

Let $(y_i \mid \theta, \omega) \sim N(\theta, \omega^{-1}),\ i = 1, \dots, n$

Priors: $(\theta \mid \omega) \sim N(\mu, (\omega K)^{-1})$

$$\omega \sim \text{Gamma}\left(\frac{d}{2}, \frac{n}{2}\right)$$

$$p(\theta, \omega) \propto \omega^{\frac{(d+1)}{2}-1} \exp\left\{-\omega \frac{K(\theta-\mu)^2}{2}\right\} \exp\left\{-\omega \frac{n}{2}\right\}$$

Let the likelihood of $y_{1:n}$ be

$$p(y_{1:n} \mid \theta, \omega) \propto \omega^{n/2} \exp\left\{-\omega\left(\frac{S_y + n(\bar{y}-\theta)^2}{2}\right)\right\} \quad \text{where } S_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Then the posterior is

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

$$\propto \omega^{n/2} \exp\left(-\omega\left(\frac{S_y + n(\bar{y}-\theta)^2}{2}\right)\right) \omega^{\frac{(d+1)}{2}-1} \exp\left\{-\omega \frac{K(\theta-\mu)^2}{2}\right\} \exp\left\{-\omega \frac{n}{2}\right\}$$

$$(1) \quad = \omega^{\frac{n+d+1}{2}-1} \exp\left(-\frac{\omega}{2}\left(\underbrace{n(\bar{y}-\theta)^2 + K(\theta-\mu)^2}_{*} \right)\right) \exp\left(-\omega \frac{n+S_y}{2}\right)$$

Let us first complete the square inside $*$:

$$n(\bar{y}-\theta)^2 + K(\theta-\mu)^2 = n\bar{y}^2 - 2n\bar{y}\theta + n\theta^2 + K\theta^2 - 2K\theta\mu + K\mu^2$$

$$= (n+K)\theta^2 - 2(n\bar{y} + K\mu)\theta + n\bar{y}^2 + K\mu^2$$

$$= (n+K)\left(\theta^2 - 2\frac{n\bar{y}+K\mu}{n+K}\theta + \left(\frac{n\bar{y}+K\mu}{n+K}\right)^2 - \left(\frac{n\bar{y}+K\mu}{n+K}\right)^2\right) + n\bar{y}^2 + K\mu^2$$

$$= (n+K)\left(\theta - \left(\frac{n\bar{y}+K\mu}{n+K}\right)\right)^2 - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2$$

plugging $*$ back into (1):

$$= \omega^{\frac{n+d+1}{2}-1} \exp\left[-\frac{\omega}{2}\left((n+K)\left(\theta - \left(\frac{n\bar{y}+K\mu}{n+K}\right)\right)^2 - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2\right)\right] \exp\left(-\omega \frac{n+S_y}{2}\right)$$

$$= \omega^{\frac{n+d+1}{2}-1} \exp\left(-\frac{\omega(n+K)}{2}\left(\theta - \left(\frac{n\bar{y}+K\mu}{n+K}\right)\right)^2\right) \exp\left(-\frac{\omega}{2}\left(n + S_y - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2\right)\right)$$

Thus, we can state the new parameters. Again, the form of the posterio is

$$p(\theta, \omega \mid y) \propto \omega^{(d^*+1)/2-1} \exp\left\{-\omega \cdot \frac{K^*(\theta-\mu^*)^2}{2}\right\} \exp\left\{-\omega \frac{n^*}{2}\right\}$$

$$\mu^* = \frac{n\bar{y}+K\mu}{n+K}$$

$$K^* = n + K$$

$$d^* = n + d$$

$$d^* = n + d$$

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

C) From the joint posterior of b), what is the conditional posterior distribution $p(\theta \mid y, \omega)$?

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

given ω, the conditional posterior is normal. We can simply ignore any piece that does not contain θ. Therefore,

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

$$\sim N\left(\frac{n\bar{y} + k\mu}{n+k}, \frac{1}{\omega(n+k)}\right)$$

D) From the joint posterior, what is the marginal posterior distribution of $p(\omega \mid y)$?

$$p(\omega \mid y) = \int_{-\infty}^{\infty} p(\theta, \omega \mid y)\, d\theta$$

$$\propto \int_{-\infty}^{\infty} \omega^{\frac{n+d+1}{2} - 1} \exp\left(-\frac{\omega(n+k)}{2}\left(\theta - \left(\frac{n\bar{y} + k\mu}{n+k}\right)\right)^2\right) \exp\left(-\frac{\omega}{2}\left(\eta + S_y - \frac{(n\bar{y} + k\mu)^2}{n+k} + n\bar{y}^2 + k\mu^2\right)\right) d\theta$$

$$= \omega^{\frac{n+d+1}{2} - 1} \exp\left(-\frac{\omega}{2}\left(\eta + S_y - \frac{(n\bar{y} + k\mu)^2}{n+k} + n\bar{y}^2 + k\mu^2\right)\right) \underbrace{\int_{-\infty}^{\infty} \exp\left(-\frac{\omega(n+k)}{2}\left(\theta - \left(\frac{n\bar{y} + k\mu}{n+k}\right)\right)^2\right) d\theta}_{\text{This is a } N\left(\frac{n\bar{y} - k\mu}{n+k}, \frac{1}{\omega(n+k)}\right) \text{ kernel}}$$

$$= \omega^{\frac{n+d+1}{2} - 1} \exp\left(-\frac{\omega}{2}\left(\eta + S_y - \frac{(n\bar{y} + k\mu)^2}{n+k} + n\bar{y}^2 + k\mu^2\right)\right) \frac{\sqrt{2\pi}}{\sqrt{\omega(n+k)}}$$

$$\propto \omega^{\frac{n+d}{2} - 1} \exp\left(-\frac{\omega}{2}\left(\eta + S_y - \frac{(n\bar{y} + k\mu)^2}{n+k} + n\bar{y}^2 + k\mu^2\right)\right)$$

$$\sim \text{Gamma}\left(\frac{n+d}{2}, \frac{1}{2}\left(\eta + S_y - \frac{(n\bar{y} + k\mu)^2}{n+k} + n\bar{y}^2 + k\mu^2\right)\right)$$

E) Show that the marginal posterior $p(\theta \mid y)$ takes the form of a centered scaled t distribution and express the parameters of this t distribution in terms of the four parameters of the normal gamma posterior for (θ, ω).

In part A, we showed that the normal gamma prior yields a centered, scaled t distribution

$$p(\theta) \propto \left(1 + \frac{1}{d} \frac{(\theta - \mu)^2}{\left(\sqrt{\frac{\eta}{kd}}\right)^2}\right)^{-\left(\frac{d+1}{2}\right)}$$

Since the posterior is also a normal gamma distribution, albeit with more complicated parameters, the marginal distribution of $p(\theta \mid y)$ will be

$$p(\theta \mid y) \propto \left(1 + \frac{1}{d^*} \frac{(\theta - \mu^*)^2}{\left(\sqrt{\eta^* \ldots}\right)^2}\right)^{-\left(\frac{d^*+1}{2}\right)} \qquad \mu^* = \frac{n\bar{y} + k\mu}{n+k}$$

$$p(\theta|y) \propto \left(1+\frac{1}{d^*}\frac{(\theta-\mu^*)^2}{\left(\sqrt{\frac{\eta^*}{k^* d^*}}\right)^2}\right)^{-\left(\frac{d^*+1}{2}\right)} \quad \text{where} \quad \mu^* = \frac{n\bar{y}+K\mu}{n+K}$$

$$K^* = n+K$$

$$d^* = n+d$$

$$\eta^* = \left(\eta + S_y - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + \mu^2\right)$$

F) True or False: in the limit as the prior parameters K, d, and η approach zero, the priors $p(\theta)$ and $p(\omega)$ are valid probability distributions.

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

We already showed in part A that the marginal of θ is

$$p(\theta) \propto \left(1+\frac{1}{d}\frac{(\theta-\mu)^2}{\left(\sqrt{\frac{\eta}{Kd}}\right)^2}\right)^{-\left(\frac{d+1}{2}\right)}$$

The normalizing constant for this kernel is:

$$\frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\frac{\eta}{Kd}}\sqrt{d\pi}\,\Gamma\left(\frac{d}{2}\right)} \quad \text{as found on the internet}$$

Thus, the full pdf is:

$$p(\theta) = \frac{\Gamma\left(\frac{d+1}{2}\right)}{\sqrt{\frac{\eta \cancel{d}\pi}{K\cancel{d}}}\,\Gamma\left(\frac{d}{2}\right)} \left(1+\frac{1}{d}\frac{(\theta-\mu)^2}{\left(\sqrt{\frac{\eta}{Kd}}\right)^2}\right)^{-\left(\frac{d+1}{2}\right)}$$

The limit $\lim_{(d,K,\eta)\to 0} P(\theta)$ does not exist since the variables going to zero are in the denominator. Depending on the direction the limit approaches zero from, the limit would take different values so the limit doesn't exist. Therefore, $p(\theta)$ is not a not a proper density.

$$p(\omega) \propto \int_{-\infty}^{\infty} \omega^{\frac{(d+1)}{2}-1} \exp\left\{-\omega \frac{K(\theta-\mu)^2}{2}\right\} \exp\left\{-\omega \frac{\eta}{2}\right\} d\theta$$

$$\propto \frac{\left(\frac{\eta}{2}\right)^{\frac{d+1}{2}}}{\Gamma\left(\frac{d+1}{2}\right)} \omega^{\left(\frac{d+1}{2}\right)-1} \exp\left(-\frac{\eta\omega}{2}\right) \sqrt{\frac{2\pi}{\omega K}} \sim \text{Gamma}\left(\frac{d}{2}, \frac{\eta}{2}\right)$$

As d and η go to 0, we end up with a gamma(0,0) distribution which is not a proper density.

G) True or False: in the limit as the prior parameters K, d, η approach zero, the posteriors $p(\theta|y)$ and $p(\omega|y)$ are valid probability distributions.

In the case of the posterior, we have that the kernel has the following form,

$$p(\theta,\omega|y) \propto \omega^{(d^*+1)/2-1} \exp\left\{-\omega \cdot \frac{K^*(\theta-\mu^*)^2}{2}\right\} \exp\left\{-\omega \frac{\eta^*}{2}\right\}$$

$$\mu^* = \frac{n\bar{y}+K\mu}{n+K}$$

$$K^* = n+K$$

$$d^* = n+d$$

$$\eta^* = \left(\eta + S_y - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2\right)$$

We can plug in the updated parameters into the t and gamma marginals for $p(\theta|y)$ and $p(\omega|y)$ to obtain:

$$p(\theta|y) \propto \left(1 + \frac{1}{d^*}\frac{(\theta-\mu^*)^2}{\left(\sqrt{\frac{\eta^*}{K^* d^*}}\right)^2}\right)^{-\left(\frac{d^*+1}{2}\right)} \quad \text{where} \quad \mu^* = \frac{n\bar{y}+K\mu}{n+K}$$

$$K^* = n+K$$

$p(\theta|y) \sim t(\mu = \bar{y}, v = n, s^2 = \frac{S_y}{n^2})$

$$d^* = n+d$$

$p(\omega|y) \sim \text{gamma}(\frac{n}{2}, \frac{S_y}{2})$

$$\eta^* = \left(\eta + S_y - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2\right)$$

Again, as $d, K, \eta \to 0$, $K^* = n$, $d^* = n$ and $\eta^* = S_y$. Thus, $p(\theta|y)$ is a legitimate density and will integrate to 1

For $p(\omega|y) \sim \text{Gamma}\left(\frac{n+d}{2}, \frac{1}{2}\left(\eta + S_y - \frac{(n\bar{y}+K\mu)^2}{n+K} + n\bar{y}^2 + K\mu^2\right)\right)$, as $\eta, d, K \to 0$, we get a $\text{Gamma}(\frac{n}{2}, \frac{1}{2}S_y + n\bar{y}^2)$ distribution. This is a proper distribution and will integrate to 1.

The take away is that even if the prior is improper, you can still end up with a proper posterior.

H) Your result in E) implies that a Bayesian credible interval for θ takes the form

$$\theta \in m \pm t^* \cdot s,$$

where m and s are the posterior center and scale parameters from E), and t^* is the appropriate critical value of the t distribution for your coverage level and degrees of freedom. True or false: In the limit as the prior parameters K, d, η approach zero, the Bayesian credible level for θ becomes identical to the classical frequentist confidence interval for θ at the same confidence level.

Statement: True or false. In the limit as the prior parameters η, d, κ approach zero, the Bayesian credible level for θ becomes identical to the classical frequentist confidence interval for θ at the same confidence level.

Judging by the marginal posterior distribution of θ,

$$p(\theta|y) \propto \left(1 + \frac{1}{d^*} \frac{(\theta-\mu^*)^2}{\left(\sqrt{\frac{\eta^*}{\kappa^* d^*}}\right)^2}\right)^{-\left(\frac{d^*+1}{2}\right)} \quad \text{where} \quad \mu^* = \frac{n\bar{y} + \kappa\mu}{n+\kappa}$$

$$\kappa^* = n + \kappa$$

$$d^* = n + d$$

$$\eta^* = \left(\eta + S_y - \frac{(n\bar{y} + \kappa\mu)^2}{n+\kappa} + n\bar{y}^2 + \kappa\mu^2\right)$$

We can see that as $\eta, d, \kappa \to 0$, the center and scale parameters of $p(\theta|y)$ become

$$\mu^* = \frac{n\bar{y}}{n} = \bar{y}$$

$$s^2 = \frac{\eta^*}{\kappa^* d^*} = \frac{S_y}{n^2} \quad \text{where } S_y = \sum_{i=1}^{n} (y_i - \bar{y})^2$$

Thus, the bayesian credible interval for θ becomes

$$\theta \in \bar{y} \pm \frac{t^*}{\sqrt{n}} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{y})^2} = \bar{y} \pm \frac{t^*}{\sqrt{n}} s_n$$

This is the same as the confidence interval generated for the frequentist estimation of μ. In the frequentist interpretation, when we do not know the population variance, we estimate it using $\frac{s_n}{\sqrt{n}}$. The sampling distribution is a t-distribution centered at $\bar{y}$.