Chapter 1

Wednesday, January 23, 2019 4:31 PM

Problem 1

A) $x_1, \dots, x_n \sim$ bernulli (ω), Suppose $\omega \sim$ Beta(a,b)

$$p(\omega) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \omega^{a-1}(1-\omega)^{b-1}$$

Derive the posterior $p(\omega \mid x_{1:n})$

$p(\omega \mid x_{1:n}) \propto p(\omega) p(x_{1:n} \mid \omega)$

$f(\omega) \propto g(\omega)$
$f(\omega) = c\, g(\omega)$ where c is a constant

For Bayes rule, $c = \frac{1}{p(x_{1:n})} = \frac{1}{\int p(x \mid \omega) p(\omega) d\omega}$

$p(x_{1:n} \mid \omega) = \prod_{i=1}^{n} \omega^{x_i}(1-\omega)^{1-x_i}$

$p(\omega \mid x_{1:n}) \propto \omega^{a-1}(1-\omega)^{b-1} \omega^{s}(1-\omega)^{n-s}$ $\quad s = \sum_{i=1}^{n} x_i$

$= \omega^{a+s-1}(1-\omega)^{n-s+b-1}$

$\sim$ beta$(a+s, n-s+b)$

b) The PDF of a gamma RV $x \sim GA(a,b)$ is

$$p(x) = \frac{b^a}{\Gamma(a)} x^{a-1} \exp(-bx)$$

Suppose $x_1 \sim Ga(a_1, 1)$ $\quad x_2 \sim Ga(a_2, 1)$. Define

$$y_1 = \frac{x_1}{x_1 + x_2} \qquad y_2 = x_1 + x_2.$$

Find the joint density for (y_1, y_2) using a direct PDF transformation (& its jacobian). Use this method to find $p(y_1)$, $p(y_2)$ & propose a method for simulating beta RVs assuming you've got gamma RVs.

$x_1 \sim ga(a_1, 1)$ $\quad x_2 \sim ga(a_2, 1)$

To get gamma RVs, add two $\exp(\lambda_1) + \exp(\lambda_2)$

$y_1 = \frac{x_1}{x_1 + x_2}$ $\quad y_2 = x_1 + x_2$

To get exponential, use uniforms w/ inverse logit function

$A: \{(x_1, x_2) : x_1 \in [0, \infty) \times x_2 \in [0, \infty)\}$

$B: \{(y_1, y_2) : y_1 \in [0, 1] \times y_2 \in [0, \infty)\}$

one to one & onto.

$x_1 = h_1(y_1, y_2) = y_1 y_2$ $\quad x_2 = h_2(y_1, y_2) = y_2 - y_1 y_2$

$$J = \begin{vmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ -y_2 & 1 - y_1 \end{vmatrix} = y_2(1-y_1) + y_1 y_2 = y_2 - y_1 y_2 + y_1 y_2$$

$$= y_2$$

$f_{y_1 y_2}(y_1, y_2) = f_{x_1 x_2}(h_1, h_2)|J|$

$f_{x_1 x_2}(x_1, x_2) = \frac{1^{a_1} 1^{a_2}}{\Gamma(a_1)\Gamma(a_2)} x_1^{a_1 - 1} \exp(-x_1) x_2^{a_2 - 1} \exp(-x_2)$

$f_{y_1 y_2}(y_1, y_2) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} (y_1 y_2)^{a_1 - 1} \exp(-y_1 y_2)(y_2 - y_1 y_2)^{a_2 - 1} \exp(-y_2 + y_1 y_2)(y_2)$

$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1 - 1} y_2^{a_1} \cancel{\exp(-y_1 y_2)} (y_2)^{a_2 - 1} (1-y_1)^{a_2 - 1} \exp(-y_2) \cancel{\exp(y_1 y_2)}$

$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1 - 1}(1-y_1)^{a_2 - 1} y_2^{a_1 + a_2 - 1} \exp(-y_2)$

$p(y_1) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} \int_0^\infty y_1^{a_1 - 1}(1-y_1)^{a_2 - 1} y_2^{a_1 + a_2 - 1} \exp(-y_2)\, dy_2$

$= \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1 - 1}(1-y_1)^{a_2 - 1} \Gamma(a_1 + a_2)$

$= \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1 - 1}(1-y)^{a_2 - 1} \sim$ beta(a_1, a_2)

$$= \frac{}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1}(1-y_1)^{a_2-1} \quad \Gamma(a_1+a_2)$$

$$= \frac{\Gamma(a_1+a_2)}{\Gamma(a_1)\Gamma(a_2)} y_1^{a_1-1}(1-y)^{a_2-1} \sim \text{beta}(a_1, a_2)$$

$$p(y_2) = \frac{1}{\Gamma(a_1)\Gamma(a_2)} y_2^{a_1+a_2-1} \exp(-y_2) \int_0^1 y_1^{a_1-1}(1-y_1)^{a_2-1} dy_1$$

$$= \frac{1}{\cancel{\Gamma(a_1)\Gamma(a_2)}} y_2^{a_1+a_2-1} \exp(-y_2) \frac{\cancel{\Gamma(a_1)\Gamma(a_2)}}{\Gamma(a_1+a_2)}$$

$$= \frac{1}{\Gamma(a_1+a_2)} y_2^{a_1+a_2-1} \exp(-y_2) \sim \text{gamma}(a_1+a_2, 1)$$

To generate a beta (α, β) transform $ga(\alpha,1)$, $ga(\beta,1)$

C) Suppose that we take independent $x_{1:n}$ from a normal sampling model w/ unknown mean θ and known variance σ^2 $x_i \sim N(\theta, \sigma^2)$. Suppose that θ is given a normal prior dist'n with mean m and variance v. Derive $p(\theta | x_{1:N})$

- normals are nice b/c precisions add together
- posterior mean is convex combination (linear combo where weights add to 1) of data mean and prior mean.

$$x_{1:n} \sim N(\theta, \sigma^2)$$

$$p(\theta) \sim N(m, v) = \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{1}{2v}(\theta - m)^2\right)$$

$$p(x_{1:n} | \theta, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{1}{2\sigma^2}(x_i - \theta)^2\right)$$

$$= \frac{1}{(2\pi\sigma)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right)$$

$$p(\theta | x_{1:n}) \propto \exp\left(-\frac{1}{2v}(\theta - m)^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \theta)^2\right)$$

$$= \exp\left(-\frac{1}{2v}(\theta^2 - 2\theta m + m^2) - \frac{1}{2\sigma^2}(\Sigma x_i^2 - 2x_i\theta + \theta^2)\right)$$

$$\propto \exp\left\{-\left(\frac{\theta^2 - 2\theta m}{2v} + \frac{n\theta^2 - 2n\bar{x}\theta}{2\sigma^2}\right)\right\}$$

$$= \exp\left\{-\left(\frac{\sigma^2(\theta^2 - 2\theta m) + v(n\theta^2 - 2n\bar{x}\theta)}{2v\sigma^2}\right)\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}(\sigma^2\theta^2 - 2\sigma^2\theta m + vn\theta^2 - 2nv\bar{x}\theta)\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}\left(\theta^2(\sigma^2 + nv) - 2\theta(\sigma^2 m - vn\bar{x})\right)\right\}$$

$$= \exp\left\{-\frac{1}{2v\sigma^2}(\sigma^2 + nv)\left(\theta^2 - 2\theta\left(\frac{\sigma^2 m - v\bar{x}n}{\sigma^2 + nv}\right) + \left(\frac{\sigma^2\mu - vn\bar{x}}{\sigma^2 + nv}\right)^2 - \left(\frac{\sigma^2\mu - vn\bar{x}}{\sigma^2 + nv}\right)^2\right)\right\}$$

$$\propto \exp\left\{-\frac{\sigma^2 + nv}{2v\sigma^2}\left(\theta - \left(\frac{\sigma^2 m - vn\bar{x}}{\sigma^2 + nv}\right)^2\right)\right\}$$

$$\sim N\left(\frac{\sigma^2 m - vn\bar{x}}{\sigma^2 + nv}, \frac{v\sigma^2}{\sigma^2 + nv}\right)$$

D) $x_{1:n} \overset{iid}{\sim} N(\underset{\text{known}}{\theta}, \underset{\text{unknown}}{\sigma^2})$, $\quad w = 1/\sigma^2$

$$p(x_i | \theta, w) = \left(\frac{w}{2\pi}\right)^{1/2} \exp\left\{-\frac{w}{2}(x_i - \theta)^2\right\}$$

w has a gamma prior w/ hyper parameters a and b. Derive the posterior $p(w | x_{1:n})$. Reexpress this as a posterior for σ^2, the variance.

$$p(w) = \frac{b^a}{\Gamma(a)} w^{a-1} \exp(-bw) \qquad p(x_{1:n} | w, \theta) = \left(\frac{w}{2\pi}\right)^{n/2} \exp\left\{-\frac{w}{2}\Sigma(x_i - \theta)^2\right\}$$

$$p(w | x_{1:n}) \propto \frac{b^a}{\Gamma(a)} w^{a-1} \exp(-bw)\, w^{n/2} \exp\left(-\frac{w}{2}\Sigma(x_i - \theta)^2\right)$$

$$\propto w^{a+\frac{n}{2}-1} \exp\left(-w\left(b + \frac{1}{2}\Sigma(x_i - \theta)^2\right)\right)$$

$$p(\omega|x_{1:n}) \propto \frac{b^a}{\Gamma(a)} \omega^{a-1} \exp(-b\omega)\, \omega^{n/2} \exp\left(-\frac{\omega}{2} \sum (x_i - \theta)^2\right)$$

$$\propto \omega^{a+\frac{n}{2}-1} \exp\left(-\omega\left(b + \frac{1}{2}\sum(x_i-\theta)^2\right)\right)$$

$$p(\omega|x_{1:n}) \sim \text{gamma}\left(a+\frac{n}{2},\ b+\frac{1}{2}\sum(x_i-\theta)^2\right)$$

Thus $p(\sigma^2|x_{1:n}) \sim \text{inv gamma}\left(a+\frac{n}{2},\ b+\frac{1}{2}\sum(x_i-\theta)^2\right)$

E) $x_{1:n} \sim N(\underset{\text{unknown}}{\theta}, \underset{\text{known}}{\sigma_i^2}) \qquad \theta \sim N(m, v)$

Derive $p(\theta|x_1, \dots, x_N)$. Express the posterior mean in a form that is clearly interpretable as a weighted avg of the observations and the prior mean

$$p(x_{1:n}|\theta, \sigma_i^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2\sigma_i^2}(x_i-\theta)^2\right)$$

$$= \left(\frac{1}{2\pi}\right)^{n/2} \prod_{i=1}^n \left(\frac{1}{\sigma_i}\right) \exp\left(-\frac{1}{2}\sum_{i=1}^n \frac{1}{\sigma_i^2}(x_i-\theta)^2\right)$$

$$\propto \exp\left(-\frac{1}{2}\sum \frac{1}{\sigma_i^2}(x_i-\theta)^2\right)$$

$$p(\theta|x_{1:n}) \propto \exp\left(-\frac{1}{2v}(\theta-m)^2 + -\frac{1}{2}\sum_{i=1}^n \frac{1}{\sigma_i^2}(x_i^2 - 2x_i\theta + \theta^2)\right)$$

$$= \exp\left\{-1\left(\frac{\theta^2 - 2\theta m + m^2}{2v} + \frac{\sum \frac{1}{\sigma_i^2}(x_i^2 - 2x_i\theta + \theta^2)}{2}\right)\right\}$$

$$\propto \exp\left\{-\left(\frac{\theta^2 - 2m\theta + v\sum\frac{1}{\sigma_i^2}\theta^2 - v\sum\frac{1}{\sigma_i^2}2x_i\theta}{2v}\right)\right\}$$

$$= \exp\left\{-\frac{1}{2v}\left(1+v\sum\frac{1}{\sigma_i^2}\right)\left(\theta^2 - 2\theta\left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}}\right) + \left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}}\right)^2 - \left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}}\right)^2\right)\right\}$$

$$\propto \exp\left\{-\frac{\left(1+v\sum\frac{1}{\sigma_i^2}\right)}{2v}\left(\theta - \left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}}\right)^2\right)\right.$$

$$\sim N\left(\frac{m+v\sum\frac{1}{\sigma_i^2}x_i}{1+v\sum\frac{1}{\sigma_i^2}},\ \frac{v}{1+v\sum\frac{1}{\sigma_i^2}}\right)$$

F) Suppose $(x|\omega) \sim N(m, \omega^{-1})$ where $\omega \sim Ga\left(\frac{a}{2}, \frac{b}{2}\right)$ prior. Show that the marginal dist'n of x is student's t w/ d degrees of freedom center m, and scale parameter $(b/a)^{1/2}$

$$p(\omega) = \frac{(b/2)^{a/2}}{\Gamma(a/2)} \omega^{a/2-1} \exp(-b\omega/2)$$

$$p(x|\omega) = \left(\frac{\omega}{2\pi}\right)^{1/2} \exp\left\{-\frac{\omega}{2}(x-m)^2\right\}$$

$$p(x,\omega) = p(x|\omega)p(\omega)$$

$$p(x) = \int_0^\infty p(x,\omega)\, d\omega$$

$$= \int_0^\infty \frac{(b/2)^{a/2}}{\Gamma(a/2)} \omega^{a/2-1} \exp(-b\omega/2) \left(\frac{\omega}{2\pi}\right)^{1/2} \exp\left\{-\frac{\omega}{2}(x-m)^2\right\} d\omega$$

$$= \frac{(b/2)^{a/2}}{\Gamma(a/2)(2\pi)^{1/2}} \underbrace{\int_0^\infty \omega^{a/2-1+\frac{1}{2}} \exp\left\{-\omega\left(\frac{b}{2}+\frac{1}{2}(x-m)^2\right)\right\} d\omega}_{\text{kernel of } Ga\left(\frac{a+1}{2},\ \left(\frac{b}{2}+\frac{1}{2}(x-m)^2\right)\right)}$$

$$= \frac{(b/2)^{a/2}}{\Gamma(a/2)(2\pi)^{1/2}} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\left(\frac{b}{2} + \frac{(x-m)^2}{2}\right)^{\frac{a+1}{2}}}$$

Kernel of $Ga\left(\frac{a+1}{2}, \left(\frac{b}{2} + \frac{1}{2}(x-m)^2\right)\right)$

$$= \frac{(b/2)^{a/2}}{\Gamma(a/2)(2\pi)^{1/2}} \frac{\Gamma\left(\frac{a+1}{2}\right)}{\left(\frac{b}{2} + \frac{(x-m)^2}{2}\right)^{\frac{a+1}{2}}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(2\pi)^{1/2}} \qquad \frac{\left(\frac{b}{2}\right)^{a/2} \Big/ \left(\frac{b}{2}\right)^{\frac{a+1}{2}}}{\left(\frac{b}{2} + \frac{(x-m)^2}{2}\right)^{a+1/2} \Big/ \left(\frac{b}{2}\right)^{\frac{a+1}{2}}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(2\pi)^{1/2}\left(\frac{b}{2}\right)^{1/2}} \left(1 + \frac{\frac{(x-m)^2}{2}}{\frac{b}{2}}\right)^{\frac{a+1}{2}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)(b\pi)^{1/2}\left(\frac{\sqrt{a}}{\sqrt{a}}\right)} \left(\frac{a}{a}\left(1 + \frac{(x-m)^2}{b}\right)\right)^{\frac{a+1}{2}}$$

$$= \frac{\Gamma\left(\frac{a+1}{2}\right)}{\Gamma\left(\frac{a}{2}\right)\left(\frac{b}{a}\right)^{1/2}\sqrt{\pi a}} \left(\frac{a + \left(\frac{x-m}{b/a}\right)^2}{a}\right)^{\frac{a+1}{2}}$$