John Gilmer
SDS 383 D - Statistical Modeling 2
Homework

Exercise 1.1

$X_1, \dots X_n \sim Bernoulli(p), \quad h(p) \sim Beta(\alpha, \beta)$

$$\begin{aligned} p|X_1, \dots X_N &\propto f(x_1, \dots x_n|p) \cdot h(p) \\ &\propto p^{\sum x_i}(1-p)^{n-\sum x_i} \cdot p^{\alpha-1}(1-p)^{\beta-1} \\ &\propto p^{\sum x_i+\alpha-1}(1-p)^{\beta+n-\sum x_i-1} \end{aligned}$$

$$p|X_1, \dots X_N \sim Beta(\sum x_i + \alpha, \beta + n - \sum x_i)$$

Exercise 1.2

$X_1, \dots X_N \overset{iid}{\sim} Cat(p), \quad p = (p_1, \dots p_K), \quad f(x|p) = \prod_{i=1}^n \prod_{k=1}^K p_k^{I(x_i=k)}$
$p \sim Dirichlet(\alpha_1, \dots \alpha_K)$

$$\pi(p) = \frac{\Gamma(\sum_k \alpha_k)}{\prod_K \Gamma(\alpha_k)} \prod_{k=1}^K p^{\alpha_k-1} I(\sum_{k=1}^{K-1} p_k \leq 1, p_k \geq 0)$$

$$\begin{aligned} p|X_{1:n} &\propto f(x_{1:n}|p) \cdot \pi(p) \\ &\propto \prod_{i=1}^n \prod_{k=1}^K p_k^{I(x_i=k)} \cdot \prod_{k=1}^K p_k^{\alpha_k-1} \\ &\propto \prod_{k=1}^K p_K^{n_k+\alpha_k-1}, \quad \text{where } n_k = \sum_{i=1}^N I(x_i = k) \\ p|X_{1:n} &\sim Dir(n_1+\alpha_1, \dots n_k+\alpha_k) \end{aligned}$$

Exercise 1.3

$X \sim Gam(\alpha, b)$ and $Y \sim Gam(\beta, b)$ and X is independent of Y.

1.3.a

Let $W = X+Y$ and $Z = X/(X+Y)$. Show that W and Z are independent with $W \sim Gam(\alpha+\beta, b)$ and $Z \sim Beta(\alpha, \beta)$.

This is a bivariate random variable transformation:

Change of variables equation: $f_{w,z} = f_{x,y}(h_1(w,z), h_2(w,z))|J|$

$g_1(X,Y) = W = X+Y$

$g_2(X,Y) = Z = X/(X+Y)$

$h_1(Z,W) = X = ZW$

$h_2(Z,W) = Y = W(1-Z)$

$$|J| = abs(\begin{vmatrix} \frac{\partial X}{\partial W} & \frac{\partial X}{\partial Z} \\ \frac{\partial Y}{\partial W} & \frac{\partial Y}{\partial Z} \end{vmatrix}) = \begin{vmatrix} Z & W \\ 1-Z & -W \end{vmatrix} = abs(-w) = w$$

$$f_{w,z} = f_x(zw)f_y(w-zw)w$$

$$f_{w,z} = \frac{b^{\alpha}(zw)^{\alpha-1}e^{-bzw} \cdot b^{\beta}(w(1-z))^{\beta-1}e^{-bw(1-z)}w}{\Gamma(\alpha)\Gamma(\beta)} \cdot I(zw \geq 0) \cdot I(w - zw \geq 0)$$

$$f_{w,z} = \frac{w^{\alpha+\beta-1}b^{\alpha+\beta}e^{-bw} \cdot z^{\alpha-1}(1-z)^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \cdot I(0 \leq z \leq 1) \cdot I(w \geq 0)$$

$$f_{w,z} = \frac{b^{\alpha+\beta}w^{\alpha+\beta-1}e^{-bw}I(w \geq 0)}{\Gamma(\alpha+\beta)} \cdot \frac{\Gamma(\alpha+\beta)z^{\alpha-1}(1-z)^{\beta-1}I(0 \leq z \leq 1)}{\Gamma(\alpha)\Gamma(\beta)}$$

So W and Z are independent by factorization theorem and $W \sim Gam(\alpha+\beta, b)$ and $Z \sim Beta(\alpha, \beta)$

1.3.b

Explain how to use a random gamma sampler to sample from a $Beta(\alpha, \beta)$ random variable using R:

We can use the R function (rgamma(n, shape, rate)) to sample two gamma random variables with shape parameters $\alpha = 10, \beta = 2$ and rate parameters equal $b = 5, b = 5$ and then use the transformation from the previous step to get a sample from a $Beta(\alpha = 10, \beta = 2)$ distribution.

```
x <- rgamma(1, 10, 5)
y <- rgamma(1, 2, 5)
z = x / (x + y)
```

Exercise 1.4

Suppose $X_1 \dots X_n \overset{iid}{\sim} \text{Normal}(\theta, \sigma_0^2)$ where σ_0^2 is known. Suppose $\theta \sim \text{Normal}(m, v)$, find posterior of $\theta | X_{1:N}$

$$\begin{aligned}
\theta|X_{1:N} &\propto \exp\{\frac{-1}{2\sigma_0^2}\sum_{i=1}^{N}(x_i-\theta)^2\} \cdot \ \exp\{\frac{-1}{2v^2}(\theta-m)^2\} \\
&\propto \exp\{\frac{-1}{2\sigma_0^2}\big(\sum_{i=1}^{N}(x_i^2+\theta^2-2\theta_i)\big) + \frac{-1}{2v^2}\big(\theta^2+m^2-2\theta m\big)\} \\
&\propto \exp\{\frac{-\theta^2}{2}\cdot(\frac{1}{v^2}+\frac{n}{\sigma_0^2}) + \theta(\frac{m}{v^2}+\frac{\sum x_i}{\sigma_0^2}) - (\frac{m^2}{2v^2}+\frac{\sum x_i^2}{2\sigma_0^2})\} \\
&\propto \exp\{\frac{-\theta^2}{2}\cdot(\frac{1}{v^2}+\frac{n}{\sigma_0^2}) + \theta(\frac{m}{v^2}+\frac{\sum x_i}{\sigma_0^2})\} \\
&\propto \exp\{\frac{-1}{2}\big(\theta^2(\frac{1}{v^2}+\frac{n}{\sigma_0^2}) - 2\theta(\frac{m}{v^2}+\frac{\sum x_i}{\sigma_0^2})\big)
\end{aligned}$$

Completing the square:

$$\begin{aligned}
a &= \frac{1}{v^2}+\frac{n}{\sigma_0^2} \\
b &= \frac{m}{v^2}+\frac{\sum x_i}{\sigma_0^2} \\
a\theta^2 - 2\theta b &= a(\theta^2 - \frac{2\theta b}{a}) = a((\theta-\frac{b}{a})^2 - \frac{b^2}{a^2})
\end{aligned}$$

$$\begin{aligned}
\theta|X_{1:N} &\propto \exp\{\frac{-1}{2}a(\theta-\frac{b}{a})^2\} \\
&\propto \exp\{\frac{-1}{2}\frac{\left(\theta-\frac{b}{a}\right)^2}{\frac{1}{\sqrt{a}}}\} \\
\theta|X_{1:N} \sim &\text{Normal}(m=\frac{b}{a}, v=a^{-1}) \\
m &= \frac{b}{a} = \frac{\frac{m}{v^2}+\frac{\sum x_i}{\sigma_0^2}}{\frac{1}{v^2}+\frac{n}{\sigma_0^2}}, \quad v = (\frac{1}{v^2}+\frac{n}{\sigma_0^2})^{-1}
\end{aligned}$$

Exercise 1.5

Suppose $X_1, \dots X_N \overset{iid}{\sim}$ Normal(θ, σ^2) with θ known but σ^2 unknown. Suppose that $\omega = \sigma^{-2}$ has a Gam(α, β) prior. Derive posterior of $\omega | X_1, \dots X_n$

$$\begin{aligned}
h(\omega|X_{1:n}) &\propto f(x_{1:n}|\omega)h(\omega) \\
&\propto \omega^{n/2}\text{exp}\{-w/2\sum(x_i-\theta)^2\} \cdot \omega^{\alpha-1}\text{exp}\{-\beta\omega\} \\
&\propto \text{exp}\{-\omega(\beta + \frac{\sum(x_i-\theta)^2}{2})\omega^{\alpha+n/2-1} \\
\omega|X_{1:n} &\sim \text{Gamma}(\alpha + \frac{n}{2}, \beta + \frac{\sum(x_i-\theta)^2}{2})
\end{aligned}$$

Exercise 1.6

Show that for squared error loss $R(F, a) = Var_F(a) + Bias_F(a)^2$. Show that $R(F, a)$ is minimized at $a \equiv \mu(F)$.

$$R(F,a) = \int (\mu(F) - a)^2 dF$$

$$R(F,a) = \int (\mu(F)^2 - 2\mu(F)a + a^2) dF$$

$$R(F,a) = \mu(F)^2 - 2\mu(F)\int adF + \int a^2 dF$$

$$\text{Var}_F(a) = E_F(a^2) - E_F(a)^2 = \int a^2 dF - (\int adF)^2$$

$$\text{Bias}_F(a)^2 = E_F(a - \mu(F))^2 = (\int (a - \mu(F))dF)^2$$

$$\text{Bias}_F(a)^2 = (\int adF - \mu(F))^2 = (\int adF)^2 - 2\mu(F)\int adF + \mu(F)^2$$

$$\text{Var}_F(a) + \text{Bias}_F(a)^2 = \mu(F)^2 - 2\mu(F)\int adF + \int a^2 dF = R(F,a)$$

Minimizing R(F, a)

$$R(F,a) = \mu(F)^2 - 2\mu(F)\int adF + \int a^2 dF$$

$$\frac{d}{da}(R(F,a)) = 2\int adF - 2\mu(F) = 0$$

$$\mu(F) = \int adF \quad \text{R(F,a) is minimized at } a = \mu(F)$$

Exercise 1.7 Suppose that $L(F,a)$ is the squared-error loss $(\mu(F)-a)^2$. Show that the Bayes action is given by the posterior mean.

$$\tilde{a} = \min_a \int L(f,a)\Pi(dF) = \min_a \int (\mu(F)-a)^2\Pi(dF)$$
$$\tilde{a} = \min_a \int (\mu(F)^2 - 2a\mu(F) + a^2)\Pi(dF)$$
$$\tilde{a} = \min_a \int \mu(F)^2\Pi(dF) - 2a\int \mu(F)\Pi(dF) + a^2$$
$$\frac{d}{da} = 2a - 2E[\mu(F)|Z=z] = 0$$
$$\tilde{a} = \tilde{\mu} = E[\mu(F)|Z=z]$$

Exercise 1.8 Refer to exercises 1.4 and 1.1. Find the Bayes estimators for these problems and compare them to the maximum likelihood estimates.

From exercise 1.1:

$$X_1,\ldots X_n \sim Bernoulli(p) \quad p|X_1,\ldots X_N \sim Beta(\alpha + \sum x_i, \beta + n - \sum x_i)$$
$$\tilde{p} = \frac{a + \sum x_i}{\alpha+\beta+n} \quad \hat{p} = \frac{\sum x_i}{n}$$
$$\tilde{p} = \frac{\alpha}{\alpha+\beta+n} + \frac{n\hat{\theta}}{\alpha+\beta+n}$$
$$\tilde{p} = \frac{\alpha/\alpha+\beta}{\frac{\alpha+\beta+n}{\alpha+\beta}} + \frac{n\hat{\theta}}{\alpha+\beta+n}$$
$$\tilde{p} = \frac{\alpha+\beta}{\alpha+\beta+n}\cdot\frac{\alpha}{\alpha+\beta} + \hat{\theta}\cdot\frac{n}{\alpha+\beta+n}$$

From exercise 1.4:

$$X_1 \ dots X_N \sim \text{Normal}(\theta, \sigma_0^2) \quad \theta \sim Normal(m, v)$$

$$\theta | X_1, \dots X_N \sim \text{Normal}(\frac{1}{\frac{1}{v^2} + \frac{n}{\sigma_0^2}} \cdot (\frac{m}{v} + \frac{\sum x_i}{\sigma_0^2}), \frac{1}{\frac{1}{v^2} + \frac{n}{\sigma_0^2}})$$

$$\tilde{\theta} = \frac{1}{\frac{1}{v^2} + \frac{n}{\sigma_0^2}} \cdot (\frac{m}{v} + \frac{\sum x_i}{\sigma_0^2})$$

$$\hat{\theta} = \bar{x}$$

$$\tilde{\theta} = \frac{m/v^2}{1/v^2 + n/\sigma_0^2} + \frac{n\hat{\theta}/\sigma_0^2}{1/v^2 + n/\sigma_0^2}$$

$$\tilde{\theta} = \frac{1/v^2 \cdot m}{1/v^2 + n/\sigma_0^2} + \frac{n/\sigma_0^2 \cdot \hat{\theta}}{1/v^2 + n/\sigma_0^2}$$