STAT610 - HWK Solution 4

1.2.15 The posterior $\pi(\theta|N=n)$ is proportional to

$$\begin{aligned}\pi(\theta|N=n) &\propto f(N=n|\theta)\pi(\theta)\\ &\propto \frac{(\sum n_i)!}{\prod n_i!}\prod \theta_i^{n_i}\frac{\Gamma(\sum \alpha_i)}{\prod \Gamma(\alpha_i)}\prod \theta_i^{\alpha_i-1}\\ &\propto \prod \theta_i^{n_i+\alpha_i-1} \sim \mathcal{D}(\alpha+n)\end{aligned}$$

1.4.14 (a) $E(Y|Z)=(1+Z)E[Z_1|Z]=(1+Z)E[Z_1]=\frac{1+Z}{\lambda}$

(b) $E[E(Y|Z)]=E(\frac{1+Z}{\lambda})=\frac{1+E(Z)}{\lambda}=\frac{1}{\lambda}+\frac{1}{\lambda^2}$

(c) $\text{Var}(E(Y|Z))=\text{Var}(\frac{1+Z}{\lambda})=\frac{1}{\lambda^2}\text{Var}(Z)=\frac{1}{\lambda^4}$

(d) $\text{Var}(Y|Z)=\text{Var}((1+Z)Z_1|Z)=(1+Z)^2\text{Var}(Z_1^2)=\frac{(1+Z)^2}{\lambda^2}$

(e) $E[Var(Y|Z)]=E\left[\frac{(1+Z)^2}{\lambda^2}\right]=\frac{1}{\lambda^2}[1+2E(Z)+E(Z^2)]=\frac{1}{\lambda^2}+\frac{2}{\lambda^3}+\frac{2}{\lambda^4}$

(f) Since the best MSPE predictor $E(Y|Z)=\frac{1}{\lambda}(1+Z)$ is linear, the best linear predictor is also $\frac{1}{\lambda}(1+Z)$.

1.4.18 (a) Since $Y=y \Leftrightarrow Y_0=\frac{\sigma y}{\beta}$,

$$\begin{aligned}Z_0|Y=y &\sim Z_0|Y_0=\frac{\sigma y}{\beta}\sim N(\mu+\beta\cdot\frac{\sigma y}{\beta},\sigma^2)=N(\mu+\sigma y,\sigma^2)\\ \Rightarrow Z=\frac{z_0-\mu}{\sigma}|Y=y &\sim N(\frac{\mu+\sigma y-\mu}{\sigma},\sigma^2/\sigma^2)=N(y,1)\end{aligned}$$

(b) By Bayes' Theorem, $\pi(y|z)=\dfrac{f(z|y)\pi(y)}{f(z)}$ where

$$\begin{aligned}f(z|y)\pi(y) &= \frac{1}{\sqrt{2\pi}}e^{-(z-y)^2/2}\cdot\lambda e^{-\lambda y}\\ &= \frac{\lambda}{\sqrt{2\pi}}\exp\left\{-\frac{[y-(z-\lambda)]^2}{2}\right\}\exp\left\{-\frac{z^2-(z-\lambda)^2}{2}\right\} \qquad (y>0,\ z\in\mathbb{R})\\ f(z) &= \int_0^\infty f(z|y)\pi(y)dy\\ &= \lambda\exp\left\{-\frac{z^2-(z-\lambda)^2}{2}\right\}\int_{-(z-\lambda)}^\infty \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx \qquad (x=y-(z-\lambda))\\ &= \lambda\exp\left\{-\frac{z^2-(z-\lambda)^2}{2}\right\}\Phi(z-\lambda)\\ \therefore \pi(y|z) &= \frac{c^{-1}}{\sqrt{2\pi}}\exp\left\{-\frac{[y-(z-\lambda)]^2}{2}\right\} \qquad (y>0, c=\Phi(z-\lambda))\end{aligned}$$

(c) since $\pi(Y=y|Z_0=z_0)=\pi(Y=y|Z=\frac{z_0-\mu}{\sigma})=\dfrac{1}{\sqrt{2\pi}\Phi(\frac{z_0-\mu}{\sigma}-\lambda)}\exp\{-\frac{[y-(\frac{z_0-\mu}{\sigma}-\lambda)]^2}{2}\}$,

$$\begin{aligned}\pi_0(Y_0=y_0|Z_0=z_0) &= \pi(Y=frac\beta y_0\sigma|Z_0=z_0)\cdot|J| \qquad \text{(By Jacobian)}\\ &= \frac{\beta}{\sigma}\frac{1}{\sqrt{2\pi}\Phi(\frac{z_0-\mu}{\sigma}-\lambda)}\exp\left\{-\frac{[\frac{\beta y_0}{\sigma}-(\frac{z_0-\mu}{\sigma}-\lambda)]^2}{2}\right\}\end{aligned}$$

(d) Let $m(z_0)$ be the best predictor which should be the condintional median of Y_0 given Z_0, then

$$\begin{aligned}
&\int_o^{m(z_0)} \pi(y_0|z_0)dy_0 = \frac{1}{2} \\
\Rightarrow &\int_0^{m(z_0)} \frac{\beta}{\sigma}\frac{1}{\sqrt{2\pi}\Phi(\frac{z_0-\mu}{\sigma}-\lambda)} \exp\left\{-\frac{[\frac{\beta y_0}{\sigma}-(\frac{z_0-\mu}{\sigma}-\lambda)]^2}{2}\right\} dy_0 = \frac{1}{2} \\
\Rightarrow &\int_{-(\frac{z_0-\mu}{\sigma}-\lambda)}^{\frac{\beta}{\sigma}m(z_0)-(\frac{z_0-\mu}{\sigma}-\lambda)} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{\Phi(\frac{z_0-\mu}{\sigma}-\lambda)}{2} && \left(x = \frac{\beta}{\sigma}y_0 - (\frac{z_0-\mu}{\sigma}-\lambda)\right) \\
\Rightarrow &\Phi\left(\frac{\beta}{\sigma}m(z_0) - (\frac{z_0-\mu}{\sigma}-\lambda)\right) - \Phi\left(-(\frac{z_0-\mu}{\sigma}-\lambda)\right) = \frac{\Phi(\frac{z_0-\mu}{\sigma}-\lambda)}{2} \\
\Rightarrow &\frac{\beta}{\sigma}m(z_0) - \left(\frac{z_0-\mu}{\sigma}-\lambda\right) = \Phi^{-1}\left(\frac{c}{2}+1-c\right) && \left(c = \Phi(\frac{z_0-\mu}{\sigma}-\lambda)\right) \\
\Rightarrow &m(z_0) = \frac{1}{\beta}(z_0-\mu-\lambda\sigma) + \frac{\sigma}{\beta}\Phi^{-1}\left(1-\frac{c}{2}\right)
\end{aligned}$$

(e) The best MSPE predictor of Y given $Z = z$ can be obtained by

$$\begin{aligned}
E(Y|Z=Z) &= \frac{c^{-1}}{\sqrt{2\pi}}\int_0^\infty ye^{-\frac{[y-(z-\lambda)]^2}{2}}dy && (c = \Phi(z-\lambda)) \\
&= \frac{c^{-1}}{\sqrt{2\pi}}\int_{-(z-\lambda)}^\infty [x+(z-\lambda)]e^{-\frac{x^2}{2}}dx && (x = y-(z-\lambda)) \\
&= \frac{c^{-1}}{\sqrt{2\pi}}e^{-\frac{(z-\lambda)^2}{2}} + c^{-1}(z-\lambda)\Phi(z-\lambda) \\
&= c^{-1}\phi(z-\lambda) - (\lambda - z)
\end{aligned}$$

1.6.11 (b) $\Gamma(p,\lambda) \sim \frac{x^{p-1}}{\Gamma(p)}\exp\{-\lambda x + p\log\lambda\} \Rightarrow T(X) = X, \eta = -\lambda, A(\eta) = -p\log(-\eta)$. So $T(X) = X$ is the sufficient statistic for $\theta = \lambda$, and the moment-generating function of T is given by

$$M(s) = \exp\{A(s+\eta) - A(\eta)\} = \left(\frac{\eta}{\eta+s}\right)^p = \left(\frac{\lambda}{\lambda-s}\right)^p.$$

1.6.18 By example 1.5.5 and 1.6.6, $\theta = (\beta_1, \beta_2, \sigma^2)^T$ is identifiable, and the density

$$p(y,\theta) = (2\pi)^{-\frac{n}{2}}\exp\left\{\frac{-\sum Y_i^2 + 2\beta_1\sum Y_i + 2\beta_2\sum z_iY_i}{2\sigma^2} - \frac{\sum(\beta_1+\beta_2 z_i)^2}{2\sigma^2} - \frac{n}{2}\log\sigma^2\right\}$$

can be expressed in the canonical form with $k = 3$,

$$T(x) = (T_1(x), T_2(X), T_3(X))^T = (\sum Y_i, \sum z_iY_i, \sum Y_i^2)^T,$$

$$\eta = (\eta_1,\eta_2,\eta_3)^T = (\beta_1/\sigma^2, \beta_2/\sigma^2, -1/(2\sigma^2))^T,$$

$$A(\eta) = -\frac{\sum(\eta_1+\eta_2 z_i)^2}{4\eta_3} + \frac{n}{2}\log(\frac{-1}{2\eta_3})$$

and $\varepsilon = \{(\eta_1,\eta_2,\eta_3)^T : \eta_1 \in R, \eta_2 \in R, \eta_3 < 0\}$.

The rank can be directly verified by Thm 1.6.4(ii).

The mean and variance of T can be obtained by Cor 1.6.1

$$
\begin{aligned}
E(T) &= \dot{A}(\eta) = \left(\frac{\partial}{\partial\eta_1}A(\eta), \frac{\partial}{\partial\eta_2}A(\eta), \frac{\partial}{\partial\eta_3}A(\eta)\right)' \\
&= \left(-\frac{\sum(\eta_1+\eta_2 z_i)}{2\eta_3}, -\frac{\sum z_i(\eta_1+\eta_2 z_i)}{2\eta_3}, \frac{\sum(\eta_1+\eta_2 z_i)^2}{4\eta_3^2} - \frac{n}{2\eta_3}\right)' \\
&= \left(\sum \mu_i, \sum z_i\mu_i, \sum \mu_i^2 + n\sigma^2\right)' \qquad \left(\text{let } \mu_i = \beta_1 + \beta_2 z_i = -\frac{\eta_1+\eta_2 z_i}{2\eta_3}\right) \\
\mathrm{Var}(T) &= \ddot{A}(\eta) = \left(\frac{\partial^2}{\partial\eta_i\partial\eta_j}A(\eta)\right)_{i,j} \\
&= \begin{bmatrix} -\frac{n}{2\eta_3} & -\frac{\sum z_i}{2\eta_3} & \frac{\sum(\eta_1+\eta_2 z_i)}{2\eta_3^2} \\ & -\frac{\sum z_i^2}{2\eta_3} & \frac{\sum z_i(\eta_1+\eta_2 z_i)}{2\eta_3^2} \\ & & -\frac{\sum(\eta_1+\eta_2 z_i)^2}{2\eta_3^3} + \frac{n}{2\eta_3^2} \end{bmatrix} = \begin{bmatrix} n\sigma^2 & \sigma^2\sum z_i & 2\sigma^2\sum\mu_i \\ & \sigma^2\sum z_i^2 & 2\sigma^2\sum z_i\mu_i \\ & & 4\sigma^2\sum\mu_i^2 + 2n\sigma^4 \end{bmatrix}
\end{aligned}
$$

or by the model assumption,

$$
\begin{aligned}
E(T_1) &= E\left(\sum Y_i\right) = \sum E(Y_i) = \sum \mu_i \\
E(T_2) &= E\left(\sum z_i Y_i\right) = \sum E(z_i Y_i) = \sum z_i\mu_i \\
E(T_3) &= E\left(\sum Y_i^2\right) = \sum E(Y_i^2) = n\sigma^2 + \sum \mu_i^2 \\
\mathrm{Var}(T_1) &= \mathrm{Var}\left(\sum Y_i\right) = \sum \mathrm{Var}(Y_i) = n\sigma^2 \\
\mathrm{Var}(T_2) &= \mathrm{Var}\left(\sum z_i Y_i\right) = \sum z_i^2 \mathrm{Var}(Y_i) = n\sigma^2 \sum z_i^2 \\
\mathrm{Var}(T_3) &= \mathrm{Var}\left(\sum Y_i^2\right) = \sum \mathrm{Var}\left[(Y_i - \mu_i + \mu_i)^2\right] \\
&= \sum \mathrm{Var}\left[(Y_i-\mu_i)^2 + 2\mu_i(Y_i-\mu_i) + \mu_i^2\right] \\
&= \sigma^4 \sum \mathrm{Var}\left[\left(\frac{Y_i-\mu_i}{\sigma}\right)^2\right] + \sum 4\mu_i^2 \mathrm{Var}(Y_i) \qquad \left(\left(\frac{Y_i-\mu_i}{\sigma}\right)^2 \sim \text{i.i.d. } \chi_1^2\right) \\
&= 2n\sigma^4 + 4\sigma^2 \sum \mu_i^2 \\
\mathrm{Cov}(T_1, T_2) &= \mathrm{Cov}\left(\sum Y_i, \sum z_i Y_i\right) = \sum z_i \mathrm{Var}(Y_i) = \sigma^2 \sum z_i \\
\mathrm{Cov}(T_1, T_3) &= \mathrm{Cov}\left(\sum Y_i, \sum Y_i^2\right) = \sum \mathrm{Cov}(Y_i, Y_i^2) \\
&= \sum E(Y_i^3) - E(Y_i)E(Y_i^2) \\
&= \sum E((Y_i - \mu_i + \mu + i)^3) - \mu_i(\sigma^2 + \mu_i^2) \\
&= \sum 3\mu_i + \mu_i^3 - \mu_i\sigma^2 - \mu_i^3 \\
&= 2\sigma^2 \sum \mu_i \\
\mathrm{Cov}(T_2, T_3) &= \mathrm{Cov}\left(\sum z_i Y_i, \sum Y_i^2\right) = \sum z_i \mathrm{Cov}(Y_i, Y_i^2) \\
&= 2\sigma^2 \sum z_i\mu_i
\end{aligned}
$$

1.6.31 (a) Let $\eta = (\mu, \tau)$ and $\eta_j = (\mu_j, \tau_j)$. Then η follows a multinomial distribution with

$$\Pr(\eta = \eta_j) = \lambda_j$$

The posterior can be obtained by

$$\begin{aligned}\pi(\theta|x) &= \int \pi(\theta, \eta|x) d\eta \\ &= \int \pi(\theta|\eta, x)\pi(\eta|x) d\eta \\ &= \sum_j \pi(\theta|\eta = \eta_j, x) \Pr(\eta = \eta_j|x)\end{aligned}$$

- For $\pi(\theta|\eta = \eta_j, x)$. η can be treated as a constant vector, then θ is simply a normal r.v. with mean μ_j and variance τ_j^2. Therefore,

$$\begin{aligned}\pi(\theta|\eta = \eta_j, x) &\propto f(x|\theta, \eta = \eta_j)\pi(\theta|\eta = \eta_j) \\ &\propto \exp\left\{-\frac{(x-\theta)^2}{2\sigma_0} - \frac{(\theta - \mu_j)^2}{2\tau_j}\right\} \\ &\propto \exp\left\{-\frac{(\theta - \mu_j(x))^2}{2\tau_j(x)}\right\}\end{aligned}$$

where $\tau_j(x) = \sqrt{\left(\frac{1}{\sigma_0^2} + \frac{1}{\tau_j^2}\right)^{-1}}$, $\mu_j(x) = \left(\frac{1}{\sigma_0^2} + \frac{1}{\tau_j^2}\right)^{-1}\left(\frac{x}{\sigma_0^2} + \frac{\mu_j}{\tau_j^2}\right)$.

- For $\Pr(\eta = \eta_j|x)$,

$$X|\eta_j \sim N(\mu_j, \sigma_j^2 = \sigma_0^2 + \tau_j^2) \Rightarrow f(x|\eta_j) = \phi_{\sigma_j}(x - \mu_j)$$

$$\Rightarrow \Pr(\eta = \eta_j|x) = \frac{\Pr(\eta = \eta_j) f(x|\eta = \eta_j)}{f(x)} = \frac{\lambda_j \phi_{\sigma_j}(x - \mu_j)}{\sum_i \lambda_i \phi_{\sigma_i}(x - \mu_i)}$$

Let $\lambda_j(x) = \Pr(\eta = \eta_j|x)$, so we have $\sum_j \lambda_j(x) = \sum_j \Pr(\eta = \eta_j|x) = 1$. Thus,

$$\pi(\theta|x) = \sum_{j=1}^{k} \lambda_j(x)\phi_{\tau_j(x)}(\theta - \mu_j(x))$$

which defines a conjugate prior for the $N(\theta, \sigma_0^2)$ distribution.

(b) By Theorem 1.5.2, $\pi(\theta|x_1, \cdots, x_n) = \pi(\theta|T(x))$, where $T(x)$ is the sufficient statistic for θ. In this case, $T(x) = \bar{x} \sim N(\theta, \sigma_0^2/n)$. Also

$$\begin{aligned}\pi(\theta|x_1, \cdots, x_n) &\propto f(x_1, \cdots, x_n|\theta)\pi(\theta) \\ &\propto \exp\left\{-\frac{1}{2\sigma_0^2}\sum(x_i - \theta)^2\right\} \cdot \pi(\theta) \\ &\propto \exp\left\{-\frac{n}{2\sigma_0^2}(\bar{x} - \theta)^2\right\} \cdot \pi(\theta)\end{aligned}$$

apply same calculation in (a), the posterior is given by

$$\pi(\theta|T(x)) = \sum_j \lambda_j(\bar{x})\phi_{\tau_j(\bar{x})}(\theta - \mu_j(\bar{x}))$$

with σ_0^2 replaced by σ_0^2/n.

1.6.36 (a) Let $\theta = (\mu, \lambda)$. The density function can be expressed as

$$\begin{aligned} f(x;\theta) &= \left(\frac{\lambda}{2\pi}\right)^{1/2} x^{-3/2} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\} \\ &= (2\pi)^{-1/2} x^{-3/2} \exp\left\{-\frac{\lambda}{2\mu^2}x - \frac{\lambda}{2}\frac{1}{x} + \frac{\lambda}{\mu} + \frac{1}{2}\log\lambda\right\} \end{aligned}$$

where $\eta = \left(\frac{\lambda}{\mu^2}, \lambda\right)'$, $T(X) = \left(-\frac{X}{2}, -\frac{1}{2X}\right)'$, $B(\mu,\lambda) = -\left(\frac{\lambda}{\mu} + \frac{1}{2}\log\lambda\right)$.

(b) $\eta_1 = \mu^{-2}\lambda, \eta_2 = \lambda \Rightarrow \mu = \sqrt{\eta_2/\eta_1}, \lambda + \eta_2$. So

$$A(\eta) = -\sqrt{\eta_1\eta_2} - \frac{1}{2}\log\eta_2$$

where $(\eta_1, \eta_2) \in \varepsilon = [0, +\infty) \times (0, +\infty)$.

(c) The MGF is given by

$$M(s) = \exp\left\{-\sqrt{(s_1+\eta_1)(s_2+\eta_2)} - \frac{1}{2}\log(s_2+\eta_2) + \sqrt{\eta_1\eta_2} + \frac{1}{2}\log\eta_2\right\}$$

and

$$\begin{aligned} E(T_1) &= \frac{\partial A}{\partial \eta_1} = -\frac{1}{2}\mu & &\Rightarrow E(X) = \mu \\ E(T_2) &= \frac{\partial A}{\partial \eta_2} = -\frac{1}{2\mu} - \frac{1}{2\lambda} & &\Rightarrow E(1/X) = \frac{1}{\lambda} + \frac{1}{\mu} \\ \mathrm{Var}(T_1) &= \frac{\partial^2 A}{\partial \eta_1^2} = \frac{1}{4}\mu^3\lambda^{-1} & &\Rightarrow \mathrm{Var}(X) = \frac{\mu^3}{\lambda} \\ \mathrm{Var}(T_2) &= \frac{\partial^2 A}{\partial \eta_2^2} = \frac{1}{4}\mu^{-1}\lambda^{-1} + \frac{1}{2}\lambda^{-2} & &\Rightarrow \mathrm{Var}(1/X) = (\lambda\mu)^{-1} + 2\lambda^{-2} \end{aligned}$$

(d) If $\mu = \mu_0$ is known,

$$f(x;\mu_0,\lambda) = (2\pi)^{-1/2} x^{-3/2} \exp\left\{-\lambda\left(\frac{x}{2\mu_0^2} + \frac{1}{2x}\right) + \left(\frac{\lambda}{\mu_0} + \frac{1}{2}\log\lambda\right)\right\}$$

By Prop 1.6.1, the conjugate prior is given by

$$\begin{aligned} \pi(\lambda; t_1, t_2) &= \exp\left\{\lambda t_1 + t_2\left(\frac{\lambda}{\mu_0} + \frac{1}{2}\log\lambda\right) - \log w(t_1,t_2)\right\} \\ &\propto \exp\left\{\lambda t_1 + t_2\left(\frac{\lambda}{\mu_0} + \frac{1}{2}\log\lambda\right)\right\} \\ &= \lambda^{t_2/2}\exp\left\{\lambda(t_1 + t_2/\mu_0)\right\} \sim \Gamma(\alpha,\beta) \end{aligned}$$

where $\alpha = \frac{t_2}{2} + 1$,$\beta = t_1 + \frac{t_2}{\mu_0}$.

(e) If $\lambda = \lambda_0$ is known,

$$f(x;\mu,\lambda_0) = (2\pi)^{-1/2} x^{-3/2} \exp\left\{-\frac{\lambda_0}{2x} + \frac{1}{2}\log\lambda_0\right\} \exp\left\{-\frac{1}{\mu^2}\frac{\lambda_0 x}{2} + \frac{\lambda_0}{\mu}\right\}$$

$$\Rightarrow \pi(\mu; t_1, t_2) = \exp\left\{-\frac{t_1}{\mu^2} + \frac{\lambda_0 t_2}{\mu} - \log w(t_1, t_2)\right\} \propto \exp\left\{-\frac{t_1}{\mu^2} + \frac{\lambda_0 t_2}{\mu}\right\}$$

For any t_1, t_2, $\exp\left(-\frac{t_1}{\mu^2} + \frac{\lambda_0 t_2}{\mu}\right) \to 1$ as $\mu \to \infty$.

$$\Rightarrow \int_0^\infty e^{-\frac{t_1}{\mu^2} + \frac{\lambda_0 t_2}{\mu}} d\mu = \infty$$

Therefore Ω is empty.

(f) The conjugate prior for (λ, μ) is given by

$$\pi(\lambda, \mu; t_1, t_2, t_3) = \exp\left\{-\frac{\lambda t_1}{\mu^2} - \lambda t_2 + t_3(\frac{\lambda}{\mu} + \frac{1}{2}\log\lambda) - \log w(t_1, t_2, t_3)\right\}$$

Since $\exp\left\{-\frac{\lambda t_1}{\mu^2} - \lambda t_2 + t_3(\frac{\lambda}{\mu} + \frac{1}{2}\log\lambda)\right\} \to \exp\left\{-\lambda t_2 + \frac{1}{2} t_3 \log\lambda\right\} > 0$ as $\mu \to \infty$,

$$\int_0^\infty \exp\left\{-\frac{\lambda t_1}{\mu^2} - \lambda t_2 + t_3(\frac{\lambda}{\mu} + \frac{1}{2}\log\lambda)\right\} d\mu = \infty$$

$$\Rightarrow \int\int \exp\left\{-\frac{\lambda t_1}{\mu^2} - \lambda t_2 + t_3(\frac{\lambda}{\mu} + \frac{1}{2}\log\lambda)\right\} d\mu = \infty, \qquad \forall t_1, t_2, t_3.$$