STAT 638: Solution for Homework #5

6.1

a) The joint density of θ and γ is

$$p(\theta, \gamma) \propto \theta^{a-1} e^{-b\theta} \gamma^{c-1} e^{-d\gamma}.$$

Now make the change of variable $\theta_A = \theta$ and $\theta_B = \theta\gamma$ to see what the joint density of θ_A and θ_B is. We have

$$\pi(\theta_A, \theta_B) = \theta_A^{a-1} e^{-b\theta_A} \left(\frac{\theta_B}{\theta_A}\right)^{c-1} e^{-d\theta_B/\theta_A} \times \frac{1}{\theta_A},$$

where θ_A^{-1} is the Jacobian of the transformation. So,

$$\pi(\theta_A, \theta_B) = \theta_A^{a-c-1} e^{-b\theta_A}\, \theta_B^{c-1} e^{-d\theta_B/\theta_A},$$

and since this isn't a product of a function of θ_A only times a function of θ_B only, θ_A and θ_B are not independent. This prior would be reasonable when what you know about θ_A is independent of what you know about the ratio of the two means.

b) The posterior is proportional to

$$\theta^{54} e^{-58\theta} (\theta\gamma)^{305} e^{-218\theta\gamma} \theta^{a-1} e^{-b\theta} \gamma^{c-1} e^{-d\gamma}.$$

Therefore, the full conditional of θ is proportional to

$$\theta^{359+a-1} \exp\left[-\theta(58 + 218\gamma + b)\right],$$

and therefore the full conditional of θ is gamma$(359 + a, 58 + 218\gamma + b)$.

c) From the previous part, the full conditional of γ is gamma$(305 + c, 218\theta + d)$.

d) The following results were obtained on the basis of 10,000 Gibbs replications for each combination of $(a, b) = (2, 1)$ and $(c, d) = (c, c)$:

c	*Estimate of* $E[\theta_B - \theta_A \mid \boldsymbol{y}_A, \boldsymbol{y}_B]$	*Credible interval for* $\theta_B - \theta_A$
8	0.3775	(0.0981,0.6457)
16	0.3292	(0.0683,0.5753)
32	0.2655	(0.0177,0.4931)
64	0.1966	(-0.0140,0.3934)
128	0.1302	(-0.0475,0.2980)

As c gets larger the estimate of $\theta_B - \theta_A$ moves closer to 0. The credible interval also moves toward 0 and becomes more narrow. This makes sense because the prior for γ has mean 1 and standard deviation $1/\sqrt{c}$. Therefore, as c increases the prior says that it is more and more likely that the two Poisson means are very close to each other, meaning that their difference is probably close to 0.

7.1 (a) Since the integral of the Jeffreys prior with respect to $\boldsymbol{\theta}$ is infinite, p_J cannot be a probability density for $(\boldsymbol{\theta}, \Sigma)$.

(b)

$$
\begin{aligned}
p_J(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, ...\boldsymbol{y}_n) &\propto p_J(\boldsymbol{\theta}, \Sigma) p(\boldsymbol{y}_1, ..., \boldsymbol{y}_n | \boldsymbol{\theta}, \Sigma) \\
&\propto |\Sigma|^{-(p+2)/2} |\Sigma|^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})^T \Sigma^{-1} (\boldsymbol{y}_i - \boldsymbol{\theta})\}
\end{aligned}
$$

$$
\begin{aligned}
p_J(\boldsymbol{\theta} | \Sigma, \boldsymbol{y}_1, ..., \boldsymbol{y}_n) &\propto \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})^T \Sigma^{-1} (\boldsymbol{y}_i - \boldsymbol{\theta})\} \\
&= \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}} + \bar{\boldsymbol{y}} - \boldsymbol{\theta})^T \Sigma^{-1} (\boldsymbol{y} - \bar{\boldsymbol{y}} + \bar{\boldsymbol{y}} - \boldsymbol{\theta})\} \\
&= \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})\} \exp\{-\frac{n}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\} \\
&\propto \exp\{-\frac{n}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\}
\end{aligned}
$$

Hence, $\boldsymbol{\theta} | \Sigma, \boldsymbol{y}_1, ..\boldsymbol{y}_n$ is distributed as multivariate normal$(\bar{\boldsymbol{y}}, \Sigma/n)$.

$$
\begin{aligned}
p_J(\Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n) &= \int p_J(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n) d\boldsymbol{\theta} \\
&= \int |\Sigma|^{-(p+2+n)/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{y})^T \Sigma^{-1} (\boldsymbol{y}_i - \bar{y})\} \exp\{-\frac{n}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\} d\boldsymbol{\theta} \\
&\propto ||\Sigma|^{-(p+2+n)/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})\} |\Sigma|^{1/2} \\
&= |\Sigma|^{-(p+1+n)/2} \exp\{-\frac{1}{2} tr[\sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T (\boldsymbol{y}_i - \bar{\boldsymbol{y}}) \Sigma^{-1}]\}
\end{aligned}
$$

Hence, $\Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n$ is distributed as inverse Wishart$(n, \boldsymbol{S}^{-1})$ where $\boldsymbol{S} = \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T$. Since $p_J(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n) \propto p_J(\boldsymbol{\theta} | \Sigma, \boldsymbol{y}_1, ..., \boldsymbol{y}_n) p_J(\Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n)$, the posterior distribution is normal-inverse Wishart.

7.2 (a)

$$
p(\boldsymbol{y}_1, .., \boldsymbol{y}_n | \boldsymbol{\theta}, \Psi) = (2\pi)^{-p/2} |\Psi|^{1/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})^T \Psi (\boldsymbol{y}_i - \boldsymbol{\theta})\}
$$

Thus, log likelihood of $\boldsymbol{\theta}, \Psi$ is

$$
\begin{aligned}
l(\boldsymbol{\theta}, \Psi | \boldsymbol{y}) &\propto \frac{n}{2} \log |\Psi| - \frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})^T \Psi (\boldsymbol{y}_i - \boldsymbol{\theta}) \\
&= \frac{n}{2} \log |\Psi| - \frac{1}{2} tr[\sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})(\boldsymbol{y}_i - \boldsymbol{\theta})^T \Psi]
\end{aligned}
$$

Thus,

$$
\log p(\boldsymbol{\theta}, \Psi) \propto \frac{1}{2} \log |\Psi| - \frac{1}{2n} tr[\sum_{i=1}^{n} (\boldsymbol{y}_i - \boldsymbol{\theta})(\boldsymbol{y}_i - \boldsymbol{\theta})^T \Psi]
$$

$$
\begin{aligned}
&= \frac{1}{2} log|\Psi| - \frac{1}{2n} tr[\sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Psi + n(\bar{\boldsymbol{y}} - \boldsymbol{\theta})(\bar{\boldsymbol{y}} - \boldsymbol{\theta})^T \Psi] \\
&= \frac{1}{2} \log|\Psi| - \frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Psi (\boldsymbol{y}_i - \bar{\boldsymbol{y}}) - \frac{1}{2} (\bar{\boldsymbol{y}} - \boldsymbol{\theta})^T \Psi (\bar{\boldsymbol{y}} - \boldsymbol{\theta})
\end{aligned}
$$

Then,

$$
\begin{aligned}
p_U(\boldsymbol{\theta}, \Psi) &\propto |\Psi|^{1/2} \exp\{-\frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Psi (\boldsymbol{y}_i - \bar{\boldsymbol{y}})\} \exp\{-\frac{1}{2} (\bar{\boldsymbol{y}} - \boldsymbol{\theta})^T \Psi (\bar{\boldsymbol{y}} - \boldsymbol{\theta})\} \\
&= |\Psi|^{1/2} \exp\{-\frac{1}{2} (\bar{\boldsymbol{y}} - \boldsymbol{\theta})^T \Psi (\bar{\boldsymbol{y}} - \boldsymbol{\theta})\} |\Psi|^{(p+1-p-1)/2} \exp\{-\frac{1}{2n} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Psi (\boldsymbol{y}_i - \bar{\boldsymbol{y}})\} \\
&= p_U(\boldsymbol{\theta}|\Psi) p_U(\Psi)
\end{aligned}
$$

Hence, $\boldsymbol{\theta}|\Psi$ is multivariate normal($\bar{\boldsymbol{y}}, \Psi^{-1}$), and Ψ is Wishart(p+1, $\boldsymbol{S}$) where $\boldsymbol{S} = \frac{\sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})(\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T}{n}$.

(b) Since Ψ follows Wishart $(p+1, \boldsymbol{S})$, Σ follows inverse-Wishart$(p+1, \boldsymbol{S}^{-1})$.
Therefore,

$$
\begin{aligned}
p_U(\boldsymbol{\theta}, \Sigma | \boldsymbol{y}_1, ..., \boldsymbol{y}_n) &\propto p_U(\boldsymbol{\theta}|\Sigma) p_U(\Sigma) p(\boldsymbol{y}_1, ..., \boldsymbol{y}_n | \boldsymbol{\theta}, \Sigma) \\
&= |\Sigma|^{-1/2} \exp\{-\frac{1}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\} |\Sigma|^{-(2p+2)/2} \exp\{-\frac{1}{2n} tr[\boldsymbol{S}\Sigma^{-1}]\} \\
&\times |\Sigma|^{-n/2} \exp\{-\frac{1}{2} \sum_{i=1}^{n} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{y}_i - \bar{\boldsymbol{y}})\} \exp\{-\frac{n}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\} \\
&= |\Sigma|^{-1/2} \exp\{-\frac{n+1}{2} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})^T \Sigma^{-1} (\boldsymbol{\theta} - \bar{\boldsymbol{y}})\} |\Sigma|^{-(2p+n+2)/2} \exp\{-\frac{1}{2} tr[(n+1)\boldsymbol{S}\Sigma^{-1}]\}
\end{aligned}
$$

This is a multivariate normal-inverse Wishart distribution, and can be considered as a posterior distribution for $\boldsymbol{\theta}$ and Σ.