Assignment 3

Jeff Gould

3/4/2021

Theoretical Exercises

1)

$$Y_i \overset{iid}{\sim} Gamma(\alpha, \beta)$$
$$P(\alpha, \beta) \propto \frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} p^{\alpha - 1} e^{-\beta q}$$

$$\mathcal{L}(\alpha, \beta | y_i, \dots y_n) = \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} y_i^{\alpha-1} e^{-\beta y_i} = \frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \left(\prod_{i=1}^{n} y_i \right)^{\alpha-1} e^{-\beta \sum y_i}$$

$$\mathcal{L}(y|\alpha, \beta) p(\alpha, \beta) \propto \left(\frac{\beta^{n\alpha}}{\Gamma(\alpha)^n} \left(\prod_{i=1}^{n} y_i \right)^{\alpha-1} e^{-\beta \sum y_i} \right) \left(\frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} p^{\alpha-1} e^{-\beta q} \right) =$$
$$\frac{\beta^{\alpha(n+s)}}{\Gamma(\alpha)^{n+r}} \left(p \prod_{i=1}^{n} y_i \right)^{\alpha-1} e^{-\beta(\sum y_i + q)}$$

Let: $p' = p \prod y_i$, $q' = q + \sum y_i$, $s' = s + n$, and $r' = n + r$

$$\frac{\beta^{\alpha(n+s)}}{\Gamma(\alpha)^{n+r}} \left(p \prod_{i=1}^{n} y_i \right)^{\alpha-1} e^{-\beta(\sum y_i + q)} = \frac{\beta^{\alpha s'}}{\Gamma(\alpha)^{r'}} (p')^{\alpha-1} e^{-\beta q'}$$

This takes the same form as the joint prior, therefore the prior is a conjugate prior for a Gamma distribution with unknown parameters α and β

2)

$y \sim MVN(X\beta, \lambda^{-1} I_{n \times n)}$

$\pi(\beta, \lambda) \propto \lambda^{-1}$

$$\mathcal{L}(Y|X, \beta, \lambda^{-1}) \propto |\lambda^{-1} I_{n\times n}|^{-1/2} \exp\left[-\frac{1}{2} (Y - X\beta)' (\lambda^[- 1] I)^{-1} (Y - X\beta) \right]$$
$$= \lambda^{n/2} \exp\left[-\frac{\lambda}{2} (Y - X\beta)'(Y - X\beta) \right]$$

$$
\begin{aligned}
P(\beta,\lambda) &\propto \pi(\beta,\lambda)\mathcal{L}(Y|X,\beta,\lambda^{-1}) \\
&\propto (\lambda^{-1})\lambda^{n/2}\exp\left[-\frac{\lambda}{2}(Y-X\beta)'(Y-X\beta)\right] \\
&= (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}(Y-X\beta)'(Y-X\beta)\right] \\
&= (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)+(\hat\beta-\beta)'X'X(\hat\beta-\beta)\}\right] \\
&= (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]\exp\left[-\frac{\lambda}{2}\{(\hat\beta-\beta)'X'X(\hat\beta-\beta)\}\right]
\end{aligned}
$$

So for the marginal distribution of β:

$P(\beta|\lambda,Y,X)\propto\exp\left[-\frac{1}{2\lambda^{-1}}\{(\hat\beta-\beta)'X'X(\hat\beta-\beta)\}\right]$

Which is the kernel for a multivariate normal distribution with mean $=\hat\beta$ and a Covariance matrix of $\Sigma=\lambda^{-1}(X'X)^{-1}$, and with $\hat\beta=(X'X)^{-1}X'Y$

$$
\begin{aligned}
P(\lambda|X,Y) &= \int(\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]\exp\left[-\frac{\lambda}{2}(\hat\beta-\beta)'X'X(\hat\beta-\beta)\right]d\beta \\
&\propto (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]\int\exp\left[-\frac{\lambda}{2}(\hat\beta-\beta)'X'X(\hat\beta-\beta)\right]d\beta \\
&\propto (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]\lambda^{k/2}\int\exp\lambda^{-k/2}\left[-\frac{\lambda}{2}(\hat\beta-\beta)'X'X(\hat\beta-\beta)\right]d\beta \\
&= (\lambda)^{n/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]\lambda^{k/2} \\
&= (\lambda)^{(n+k)/2-1}\exp\left[-\frac{\lambda}{2}\{(Y-X\hat\beta)'(Y-X\hat\beta)\}\right]
\end{aligned}
$$

Which we recognize as the kernel for a variable with a gamma distribution with parameters: $\alpha=\frac{n+k}{2}$ and $\beta=\frac{(Y-X\hat\beta)'(Y-X\hat\beta)}{2}$

3)

$P(\theta|W)\sim\mathcal{N}(\hat\theta,[I(\hat\theta)]^{-1})$

$I[\theta]=-\frac{d^2}{d\theta^2}\log[P(\theta|W)]$

$$
\begin{aligned}
P(\theta|W) &\propto \pi(\theta)\mathcal{L}(W|\theta) \\
&\propto (\tau^2)^{-1}\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\tau^2}}\exp[-\frac{1}{2}\frac{(w_i-\mu)^2}{\tau^2}] \\
&\propto (\tau^2)^{-(n/2+1)}\exp\left[-\frac{1}{2\tau^2}\sum(w_i-\mu)^2\right]
\end{aligned}
$$

$$\log P(\theta|W) = -(n/2+1)\log(\tau^2) - \frac{1}{2\tau^2}\sum(w_i-\mu)^2$$
$$\frac{\partial}{\partial\mu}\log P = \frac{1}{\tau^2}\sum(w_i-\mu) = 0 \rightarrow$$
$$\sum(w_i-\mu) = 0 \rightarrow$$
$$\hat{\mu} = \frac{1}{n}\sum w_i = \bar{w}$$

$$\frac{\partial^2}{\partial\mu^2}\log P = \frac{\partial}{\partial\mu}\frac{1}{\tau^2}\sum(w_i-\mu) = \frac{-n}{\tau^2}$$

$$\frac{\partial}{\partial\tau^2}\log P = \frac{-(n/2+1)}{\tau^2} + \frac{1}{2(\tau^2)^2\sum(w_i-\mu)^2} = 0 \rightarrow$$
$$\frac{1}{2(\tau^2)^2}\sum(w_i-\bar{w})^2 = \frac{(n/2+1)}{\tau^2} \rightarrow \frac{1}{2}\sum(w_i-\bar{w})^2 = \tau^2(n/2+1) \rightarrow$$
$$\hat{\tau}^2 = \frac{1}{n+2}\sum(w_i-\bar{w})^2$$

$$\frac{\partial^2}{\partial(\tau^2)^2}\log P = \frac{\partial^2}{\partial(\tau^2)}\left[\frac{-(n/2+1)}{\tau^2} + \frac{1}{2(\tau^2)^2\sum(w_i-\mu)^2}\right]$$
$$= \frac{(n/2+1)}{(\tau^2)^2} - (\tau^2)^{-3}\sum(w_i-\mu)^2$$

$$\frac{\partial^2}{\partial\mu\partial(\tau^2)}\log P = \frac{\partial}{\partial\tau^2}\left[\frac{1}{\tau^2}\sum(w_i-\mu)\right]$$
$$= -(\tau^2)^{-2}\sum(w_i-\mu)$$

$$I = -\begin{bmatrix} \frac{\partial^2}{\partial\mu^2} & \frac{\partial^2}{\partial\mu\partial\tau^2} \\ \frac{\partial^2}{\partial\mu\partial\tau^2} & \frac{\partial^2}{\partial(\tau^2)^2} \end{bmatrix}$$
$$= -\begin{bmatrix} \frac{-n}{\tau^2} & -(\tau^2)^{-2}\sum(w_i-\mu) \\ -(\tau^2)^{-2}\sum(w_i-\mu) & \frac{(n/2+1)}{(\tau^2)^2} - (\tau^2)^{-3}\sum(w_i-\mu)^2 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{n}{\tau^2} & (\tau^2)^{-2}\sum(w_i-\mu) \\ (\tau^2)^{-2}\sum(w_i-\mu) & -\frac{(n/2+1)}{(\tau^2)^2} + (\tau^2)^{-3}\sum(w_i-\mu)^2 \end{bmatrix}$$

\

$$I(\theta)|_{\theta=\hat{\theta}} = \begin{bmatrix} \frac{n}{n+2}\sum(w_i-\bar{w})^2 & (\frac{\sum(w_i-\bar{w})^2}{n+2})^{-2}\sum(w_i-\bar{w}) \\ (\frac{\sum(w_i-\bar{w})^2}{n+2})^{-2}\sum(w_i-\bar{w}) & -\frac{n/2+1}{(\frac{1}{n+2}\sum(w_i-\bar{w})^2)^2} + (\frac{\sum(w_i-\bar{w})^2}{n+2})^{-3}\sum(w_i-\bar{w})^2 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{n}{n+2}\sum(w_i-\bar{w})^2 & 0 \\ 0 & -\frac{n/2+1}{(\frac{1}{n+2}\sum(w_i-\bar{w})^2)^2} + (\frac{\sum(w_i-\bar{w})^2}{n+2})^{-3}\sum(w_i-\bar{w})^2 \end{bmatrix}$$

And so, $(\mu, \tau^2) \sim \mathcal{MVN}(\hat{\theta}, I(\hat{\theta})^{-1})$, where $\hat{\theta}$ and $I(\hat{\theta})$ are as defined above

Analysis Exercises

1)

$$f_X(x|\lambda, k) = \frac{1}{(k-1)!}\lambda^k x^{k-1} e^{-\lambda x} f_X(x|\lambda, k = 22) \qquad = \frac{1}{(22-1)!}\lambda^{22} x^{22-1} e^{-\lambda x}$$

$\pi(\lambda) \propto [J(\lambda)]^{1/2}$
$J(\lambda) = -E\left[\frac{d^2 l(x|\theta)}{d\theta^2}|\theta\right]$

$\mathcal{L}(X|\lambda) = \frac{1}{(22-1)!}\lambda^{22}x^{22-1}e^{-\lambda x} \to \log\mathcal{L} \propto 22\log(\lambda) + 21\log(x) - \lambda x \to \frac{dl}{d\lambda} = \frac{22}{\lambda} - x \to \frac{d^2 l}{d\lambda^2} = -\frac{22}{\lambda^2}$

$-E[-\frac{22}{\lambda^2}|\lambda] = \frac{22}{\lambda^2}$

So Jeffrey's Prior: $\pi(\lambda) = \frac{\sqrt{22}}{\lambda} \propto \frac{1}{\lambda}$

To get the normal approximation:

$\pi\mathcal{L}(x) = \frac{1}{\lambda}\prod_{i=1}^n \frac{1}{(22-1)!}\lambda^{22}x_i^{22-1}e^{-\lambda x_i} \propto \lambda^{22n-1}\prod_{i=1}^n x_i^{22-1}e^{-\lambda x_i} \to \log\mathcal{L} = (22n-1)\log\lambda + \sum\log(x_i) + \lambda\sum(x_i)$

$\frac{\partial}{\partial\lambda}\log\mathcal{L} = \frac{22n-1}{\lambda} - \sum x_i = 0 \to \hat{\lambda} = \frac{22n-1}{\sum x_i}$

$\frac{\partial^2}{\partial\lambda^2}\log\mathcal{L} = -\frac{22n-1}{\lambda^2} \to I(\hat{\lambda}) = \frac{22n-1}{(\frac{22n-1}{\sum x_i})^2} = \frac{(\sum x_i)^2}{22n-1}$

$\lambda \sim \mathcal{N}\left(\frac{22n-1}{\sum x_i}, \frac{22n-1}{(\sum x_i)^2}\right)$

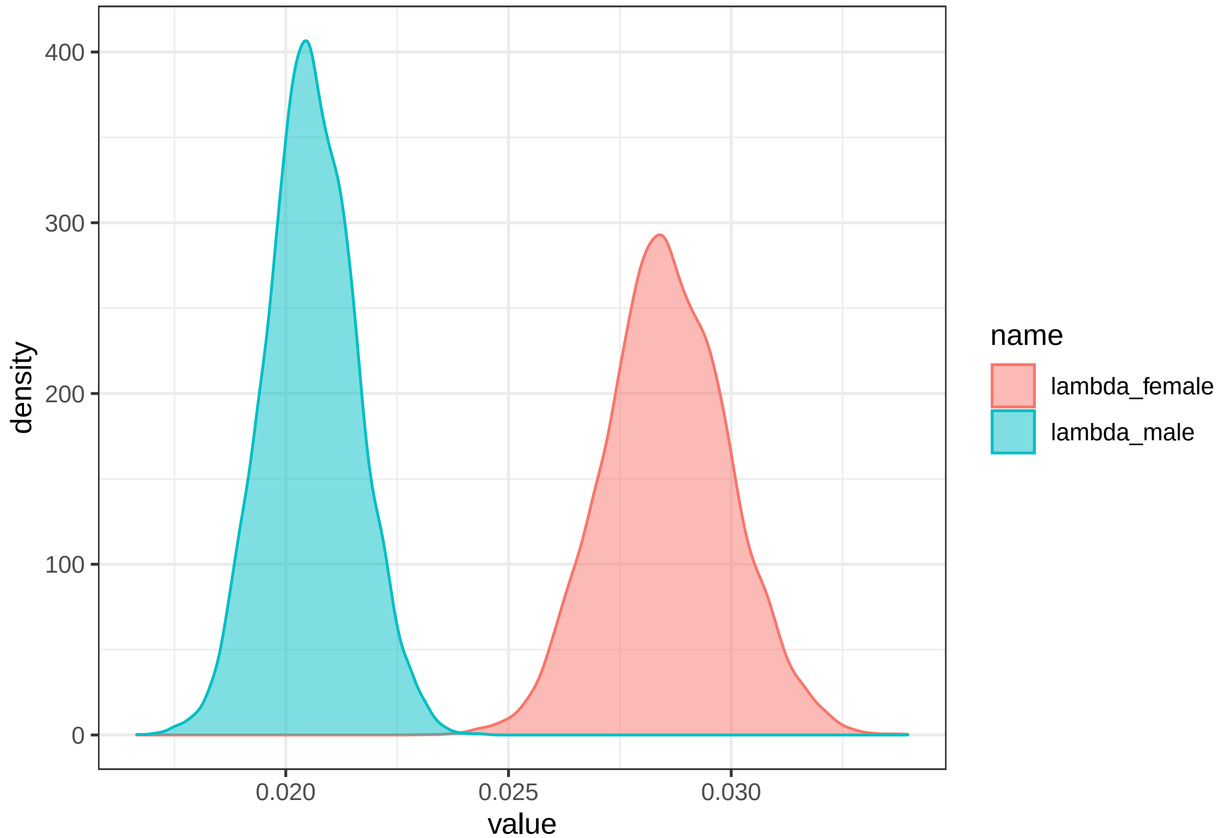
```
library(tidyverse)
incidence <- read.delim(file = "incidenceUK.txt")
k <- 22
n <- nrow(incidence)
x_i_male <- sum(incidence$male)
x_i_female <- sum(incidence$female)
B <- 10000

I_male <- sum(x_i_male)^2 / (n * k - 1)
I_female <- sum(x_i_female)^2 / (n * k - 1)
Lambda <- data.frame(lambda_male = rep(NA, B),
                     lambda_female = rep(NA,B))


set.seed(2020)
Lambda$lambda_male <- rnorm(n = B, mean = (k*n-1)/x_i_male, sd = sqrt(1/I_male))
set.seed(2020)
Lambda$lambda_female <- rnorm(n = B, mean = (k*n-1)/x_i_female, sd = sqrt(1/I_female))

Lambda %>%
  pivot_longer(cols = lambda_male:lambda_female) %>%
ggplot() +
```

```
  geom_density(aes(x = value, color = name, fill = name), alpha = 0.5) +
  theme_bw()
```



We see that the distribution for λ for males has a lower mean and variance than it does for females. Since $1/\lambda$ is the time it takes to develop cancer, this suggests that cancer develops more slowly in males than it does for females

2)

```
coup <- read.delim("coup1980.txt", sep = " ")
X <- as.matrix(coup[,3:5])
Y <- coup[,2]
B <- 10000
library(mvtnorm)
library(MCMCpack)

n <- length(Y)
k <- ncol(X)

bhat    <- c(solve(t(X)%*%X)%*%(t(X)%*%Y))
SSY <- t(Y - X%*%bhat)%*%(Y - X%*%bhat)
XtXi    <- solve(t(X)%*%X)

rbeta   <- matrix(0, nrow = B, ncol = k)
```

```r
shape <- (n + k)/2
rate <- (1/2) * t(Y - X %*% bhat) %*% (Y - X %*% bhat)

set.seed(1980)
rsig <- rgamma(B, shape, rate)

for(i in 1:B){
  CovX      <- (rsig[i]^(-1))*XtXi
  rbeta[i,] <- c(rmvnorm(1, mean = bhat, sigma = CovX))
}

rbMat      <- apply(rbeta, 2, quantile, probs = c(0.5, 0.025, 0.975))
postp    <- apply(rbeta > 0, 2, mean)
outMat   <- cbind(t(rbMat), postp)
rownames(outMat) <- colnames(X)


outMat
```

```
##                   50%         2.5%         97.5%  postp
## democracy -1.812849019 -2.428042758 -1.1976958430 0.0000
## age       -0.099995313 -0.108753172 -0.0909868069 0.0000
## tenure    -0.003715921 -0.006946732 -0.0005631526 0.0101
```

All three variables are statisitically significant in regards to their impact on a coup Whether or not the country is a democracy, on average, changed the log-Probabilty of a coup by 1.81 points Each year of age of the leader decreased the log-Probabilty of a coup by 0.1 points, on average. Lastly, the longer the leader was in office, the lower the probabilty of a coup as well.