Michael Leibert

Math 640

Homework 3

Theoretical Exercises

1. Multi-parameter distributions often lack convenient conjugate priors (if they have one at all). One such case is when Y_i are iid $Gamma(\alpha, \beta)$ where *both* α and β are unknown. The conjugate prior, while proper, is not a named density. Show that the joint prior

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

is actually a conjugate prior for the Gamma distribution with unknown α and β. That is, show that when this joint prior is used, the resulting posterior has the same parametric form. Be sure to determine the parameters. (Hint: this prior is parameterized by p, q, r, and s thus the posterior should have four parameters as well.)

$$\begin{aligned}
\mathcal{L}(Y_i|\alpha, \beta) &= \prod_{i=1}^{n} \frac{\beta^{\alpha}}{\Gamma(\alpha)} \, Y_i^{\alpha-1} \, \exp(-\beta Y_i) \\
&= \frac{\beta^{\alpha n}}{\Gamma(\alpha)^n} \left[\prod_{i=1}^{n} Y_i\right]^{\alpha-1} \exp\left(-\beta \sum_{i=1}^{n} Y_i\right)
\end{aligned}$$

Let $\boldsymbol{\theta} = (\alpha, \beta, p, q, r, s)$

$$\begin{aligned}
P(\boldsymbol{\theta}|Y_i) &\propto \frac{\beta^{\alpha n}}{\Gamma(\alpha)^n} \left[\prod_{i=1}^{n} Y_i\right]^{\alpha-1} \exp\left(-\beta \sum_{i=1}^{n} Y_i\right) \frac{\beta^{\alpha s}}{\Gamma(\alpha)^r} \, p^{\alpha-1} e^{-\beta q} \\
&= \frac{\beta^{\alpha(n+s)}}{\Gamma(\alpha)^{n+r}} \left[p \prod_{i=1}^{n} Y_i\right]^{\alpha-1} \exp\left[-\beta \left(q + \sum_{i=1}^{n} Y_i\right)\right] \\
&= \frac{\beta^{\alpha s^*}}{\Gamma(\alpha)^{r^*}} \, {p^*}^{\alpha-1} \exp\Big(-\beta q^*\Big)
\end{aligned}$$

$s^* = n + s \qquad r^* = n + r \qquad p^* = p \prod_{i=1}^{n} Y_i \qquad q^* = q + \sum_{i=1}^{n} Y_i$

2. Let $\boldsymbol{y} = [\; y_1 \quad y_2 \quad \cdots \quad y_n \;]'$ be an $n \times 1$ vector of regression outcomes. Further let X denote an $n \times p$ matrix of covariates and $\boldsymbol{\beta}$ be a $p \times 1$ vector of coefficients. Assume $\boldsymbol{y}$ is normally distributed of the form

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

That is, the standard regression assumption where we've parametrized the model in terms of the precision, λ. Using the joint prior $\pi(\boldsymbol{\beta}, \lambda) \propto \lambda^{-1}$, find the marginal distribution posterior of $\lambda|\boldsymbol{y}, X$ and the conditional posterior distribution of $\boldsymbol{\beta}|\lambda, \boldsymbol{y}, X$.

$$
\begin{aligned}
\mathcal{L}(\boldsymbol{y}|X,\boldsymbol{\beta},\lambda) &\propto \left|\lambda^{-1} I_n\right|^{\frac{1}{2}} \ \exp\left[-\frac{1}{2}(\boldsymbol{y}-X\boldsymbol{\beta})^T \left(\lambda^{-1} I_n\right)^{-1} (\boldsymbol{y}-X\boldsymbol{\beta})\right] \\
&\propto \left(\lambda^{-1}\right)^{-\frac{n}{2}} \ \exp\left[-\frac{\lambda}{2}(\boldsymbol{y}-X\boldsymbol{\beta})^T (\boldsymbol{y}-X\boldsymbol{\beta})\right]
\end{aligned}
$$

$$
\begin{aligned}
P(\boldsymbol{\beta},\lambda|\boldsymbol{y},X) &\propto \lambda^{\frac{n}{2}-1} \ \ \exp\left[-\frac{\lambda}{2}(\boldsymbol{y}-X\boldsymbol{\beta})^T (\boldsymbol{y}-X\boldsymbol{\beta})\right] \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left[-\frac{\lambda}{2}\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}+X\hat{\boldsymbol{\beta}}-X\boldsymbol{\beta}\right)^T \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}+X\hat{\boldsymbol{\beta}}-X\boldsymbol{\beta}\right)\right] \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left(-\frac{\lambda}{2}\left[\boldsymbol{y}-X\hat{\boldsymbol{\beta}}+X\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right]^T \left[\boldsymbol{y}-X\hat{\boldsymbol{\beta}}+X\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right]\right) \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left(-\frac{\lambda}{2}\left[\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)+2\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)^T X^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)+\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)^T X^T X\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right]\right) \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left(-\frac{\lambda}{2}\left[\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)+\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)^T X^T X\left(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\right)\right]\right)
\end{aligned}
$$

$$
P(\boldsymbol{\beta}|\lambda,\boldsymbol{y},X) \propto \exp\left[-\frac{\lambda}{2}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)^T X^T X\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]
$$

$$
\boldsymbol{\beta}|\lambda,\boldsymbol{y},X \sim N\left[\hat{\boldsymbol{\beta}},\ \lambda^{-1}\left(X^T X\right)^{-1}\right]
$$

$$
\begin{aligned}
P(\lambda|X,\boldsymbol{y}) &\propto \int \lambda^{\frac{n}{2}-1} \ \exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)\right]\exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)^T X^T X\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]\mathrm{d}\boldsymbol{\beta} \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)\right] \ \int \frac{(2\pi)^{\frac{p}{2}}\ \left|\lambda^{-1} I_p\left(X^T X\right)^{-1}\right|^{\frac{1}{2}}}{(2\pi)^{\frac{p}{2}}\ \left|\lambda^{-1} I_p\left(X^T X\right)^{-1}\right|^{\frac{1}{2}}} \ \exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)^T X^T X\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right]\mathrm{d}\boldsymbol{\beta} \\
&\propto \lambda^{\frac{n}{2}-1} \ \exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)\right] \ \left|\lambda^{-1} I_p\right|^{\frac{1}{2}} \ \left|\left(X^T X\right)^{-1}\right|^{\frac{1}{2}} \\
&\propto \lambda^{\frac{n-p}{2}-1} \ \exp\left[-\frac{\lambda}{2}\ \left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)\right]
\end{aligned}
$$

$$
\lambda|X,\boldsymbol{y} \sim Gamma\left[\frac{n-p}{2},\ \frac{1}{2}\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)^T\left(\boldsymbol{y}-X\hat{\boldsymbol{\beta}}\right)\right]
$$

3. Let $W_i \sim N(\mu, \tau^2)$ for $i = 1, \ldots, n$ where both μ and τ^2 are unknown. Determine the form of normal approximation to the joint posterior of μ and τ^2 when using the non-informative joint prior, i.e. $\pi\left(\mu, \tau^2\right) \propto \left(\tau^2\right)^{-1}$. (Hint: this will require find the posterior modes for both μ and τ^2 as well as the information matrix, i.e. the negative of the Hessian matrix.)

$$\mathcal{L}(W|\mu, \tau^2) \propto \left(\tau^2\right)^{-\frac{n}{2}} \exp\left[-\frac{1}{2\tau^2}\sum_{i=1}^{n}\left(W_i - \mu\right)^2\right]$$

$$P(\mu, \tau^2|W) \propto \left(\tau^2\right)^{-\frac{(n+2)}{2}} \exp\left[-\frac{1}{2\tau^2}\sum_{i=1}^{n}\left(W_i - \mu\right)^2\right]$$

$$\log\left[P(\mu, \tau^2|W)\right] \propto -\frac{(n+2)}{2} \log\left(\tau^2\right) - \frac{1}{2\tau^2}\sum_{i=1}^{n}\left(W_i - \mu\right)^2$$

$$\begin{aligned}
\frac{\partial \log(P)}{\partial \mu} &\propto \frac{1}{\tau^2}\sum_{i=1}^{n}\left(W_i - \mu\right) \\
0 &= \frac{1}{\tau^2}\sum_{i=1}^{n}\left(W_i - \mu\right) \\
\sum_{i=1}^{n}\mu &= \sum_{i=1}^{n}W_i \\
n\mu &= n\overline{W} \\
\hat{\mu} &= \overline{W}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \log(P)}{\partial \tau^2} &\propto -\frac{(n+2)}{2\tau^2} + \frac{1}{2}\left(\tau^2\right)^{-2} \sum_{i=1}^{n}\left(W_i - \mu\right)^2 \\
0 &= -\frac{(n+2)}{2\tau^2} + \frac{1}{2}\left(\tau^2\right)^{-2} \sum_{i=1}^{n}\left(W_i - \mu\right)^2 \\
\frac{(n+2)}{2\tau^2} &= \frac{1}{2}\left(\tau^2\right)^{-2} \sum_{i=1}^{n}\left(W_i - \mu\right)^2 \\
\tilde{\tau}^2 &= \frac{1}{(n+2)} \sum_{i=1}^{n}\left(W_i - \mu\right)^2 \\
\hat{\tau}^2 = \tilde{\tau}^2 \Big|_{\mu=\hat{\mu}} &= \frac{1}{(n+2)} \sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \log(P)}{\partial \mu^2} &\propto -\frac{n}{\tau^2} \\
\frac{\partial^2 \log(P)}{\partial \mu^2}\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &\propto -n\left(\hat{\tau}^2\right)^{-1} \\
&= -n\left[\frac{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2}{n+2}\right]^{-1} \\
&= -\frac{n(n+2)}{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2}
\end{aligned}$$

$$\frac{\partial^2 \log(P)}{\partial \left(\tau^2\right)^2} \propto \frac{(n+2)}{2\left(\tau^2\right)^2} - \left(\tau^2\right)^{-3} \sum_{i=1}^{n} \left(W_i - \overline{W}\right)^2$$

$$\begin{aligned}
\left.\frac{\partial^2 \log(P)}{\partial \left(\tau^2\right)^2}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &\propto \frac{n+2}{2}\left[\frac{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2}{n+2}\right]^{-2} - \left[\frac{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2}{n+2}\right]^{-3} \sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2 \\
&= \frac{(n+2)^3}{2\left[\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2\right]^2} - \frac{(n+2)^3}{\left[\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2\right]^2} \\
&= -\frac{(n+2)^3}{2\left[\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2\right]^2}
\end{aligned}$$

The mean and the mode are equal for a Gaussian Distribution.

$$\frac{\partial^2 \log(P)}{\partial \tau^2 \, \partial \mu} = \frac{\partial^2 \log(P)}{\partial \mu \, \partial \tau^2} \propto -\frac{1}{\left(\tau^2\right)^2} \sum_{i=1}^{n} \left(W_i - \mu\right)$$

$$\begin{aligned}
\left.\frac{\partial^2 \log(P)}{\partial \tau^2 \, \partial \mu}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = \left.\frac{\partial^2 \log(P)}{\partial \mu \, \partial \tau^2}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} &\propto -\frac{1}{\left(\hat{\tau}^2\right)^2} \sum_{i=1}^{n} \left(W_i - \hat{\mu}\right) \\
&= -\frac{1}{\left(\hat{\tau}^2\right)^2} \sum_{i=1}^{n} \left(W_i - \overline{W}\right) \\
&= 0
\end{aligned}$$

$$\hat{\boldsymbol{\theta}} = \left(\overline{W}, \ \frac{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2}{(n+2)} \right) \qquad I\left(\hat{\boldsymbol{\theta}}\right) = \begin{pmatrix} \dfrac{n(n+2)}{\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2} & 0 \\ 0 & \dfrac{(n+2)^3}{2\left[\sum_{i=1}^{n}\left(W_i - \overline{W}\right)^2\right]^2} \end{pmatrix}$$

$$\mu, \tau^2 | W \sim N\left[\hat{\boldsymbol{\theta}}, \ I\left(\hat{\boldsymbol{\theta}}\right)^{-1}\right]$$

Analysis Exercises

1. The age distribution of the incidence of cancer can be modeled using the Erlang distribution which has as PDF

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

where $x \in [0,\infty)$, $k \in \mathbb{Z}^+$, and $\lambda \in (0,\infty)$. Here the parameter k can be interpreted as the number of carcinogenic events needed for a cancer to develop while $1/\lambda$ is the average time to developing cancer. The data file `incidenceUK.txt` contains age specific incidence of all cancers in both males and females in the United Kingdom for the years 2013 to 2015. Using an Erlang distribution with $k = 22$[1], fixed, find the posterior distribution of the average time to developing cancer in males and females, separately, using the normal approximation to the posterior density. Use Jeffreys' prior for λ. Generate posterior summaries and compare between males and females. Draw a conclusion in context. Use $B = 10000$ samples for each model and set the seed to 2020.

$$\begin{aligned}
\mathcal{L}\left(x|k,\lambda\right) &= \prod_{i=1}^{n} \Big[(k-1)!\Big]^{-1} \lambda^k \; x_i^{k-1} \; \exp\left(-\lambda x_i\right) \\
&= \Big[(k-1)!\Big]^{-n} \lambda^{nk} \left[\prod_{i=1}^{n} x_i\right]^{k-1} \exp\left(-\lambda \sum_{i=1}^{n} x_i\right) \\
\ell\left(x|k,\lambda\right) &= -n\log\Big[(k-1)!\Big] + kn\log(\lambda) + (k-1)\log\left(\sum_{i=1}^{n} x_i\right) - \lambda\sum_{i=1}^{n} x_i
\end{aligned}$$

$$\frac{\partial \ell}{\partial \lambda} = \frac{nk}{\lambda} - \sum_{i=1}^{n} x_i \qquad\qquad J(\lambda) = -E\left[\frac{\partial^2 \ell}{\partial \lambda^2}\right] = -E\left[\frac{-\,nk}{\lambda^2}\right] = \frac{nk}{\lambda^2}$$

$$\frac{\partial^2 \ell}{\partial \lambda^2} = \frac{-\,nk}{\lambda^2} \qquad\qquad \Big[J(\lambda)\Big]^{\frac{1}{2}} \propto \lambda^{-1}$$

$$P(\lambda|k,x) \propto \frac{\lambda^{nk-1}}{\Big[(k-1)!\Big]^n} \left[\prod_{i=1}^{n} x_i\right]^{k-1} \exp\Big(-\,n\bar{x}\lambda\Big)$$

$$\lambda|k,x \sim Gamma(nk, n\bar{x})$$

[1] Note: 22 is roughly the average number of carcinogenic events needed from the 20 most common cancers.

$$\log\Big[P(\lambda|k,x)\Big] \propto (nk-1)\log(\lambda) - n\bar{x}\lambda$$

$$\frac{\partial}{\partial\lambda}\log\Big[P(\lambda|k,x)\Big] = \frac{nk-1}{\lambda} - n\bar{x}$$

$$0 = \frac{nk-1}{\lambda} - n\bar{x}$$

$$n\bar{x} = \frac{nk-1}{\lambda}$$

$$\hat{\lambda} = \frac{nk-1}{n\bar{x}}$$

$$\frac{\partial^2}{\partial\lambda^2}\log\Big[P(\lambda|k,x)\Big] = -\frac{nk-1}{\lambda^2}$$

$$I(\lambda)\Big|_{\lambda=\hat{\lambda}} = \frac{nk-1}{\hat{\lambda}^2}$$

$$= \cancel{(nk-1)}\,\frac{(n\bar{x})^2}{(nk-1)^{\cancel{2}}}$$

$$= \frac{(n\bar{x})^2}{nk-1}$$

$$\lambda|k,x \sim N\left(\frac{nk-1}{n\bar{x}},\ \frac{nk-1}{(n\bar{x})^2}\right)$$

```
n = nrow(dat); k = 22
xybar <- mean( dat$male )
xxbar <- mean( dat$female )
Nsim <- 10000

set.seed(2020) #male
xy <- rnorm( Nsim , (n*k-1) / (n * xybar ) , sqrt(n*k-1) / (n * xybar ) )
set.seed(2020) #female
xx <- rnorm( Nsim , (n*k-1) / (n * xxbar ) , sqrt(n*k-1) / (n * xxbar ) )

#Male
quantile(  xy , probs = c(.5,.025,0.975) ); mean( xy)
```

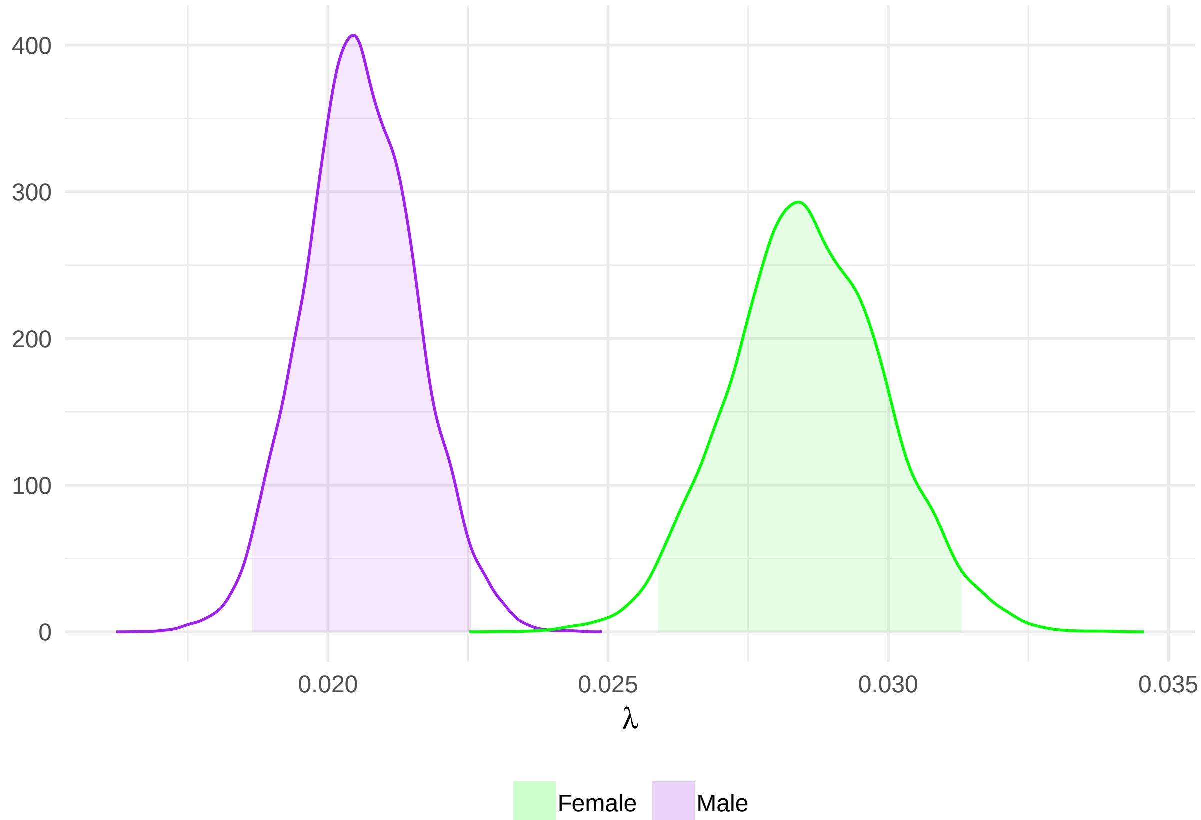
```
##        50%       2.5%      97.5%
## 0.02054315 0.01863656 0.02256389
```

```
## [1] 0.02056537
```

```
#Female
quantile(  xx  , probs = c(.5,.025,0.975) ); mean( xx)
```

```
##        50%       2.5%      97.5%
## 0.02852165 0.02587458 0.03132719
```

```
## [1] 0.02855249
```




```
#Male
quantile( 1/xy , probs = c(.5,.025,0.975) ); mean(1/xy)
```

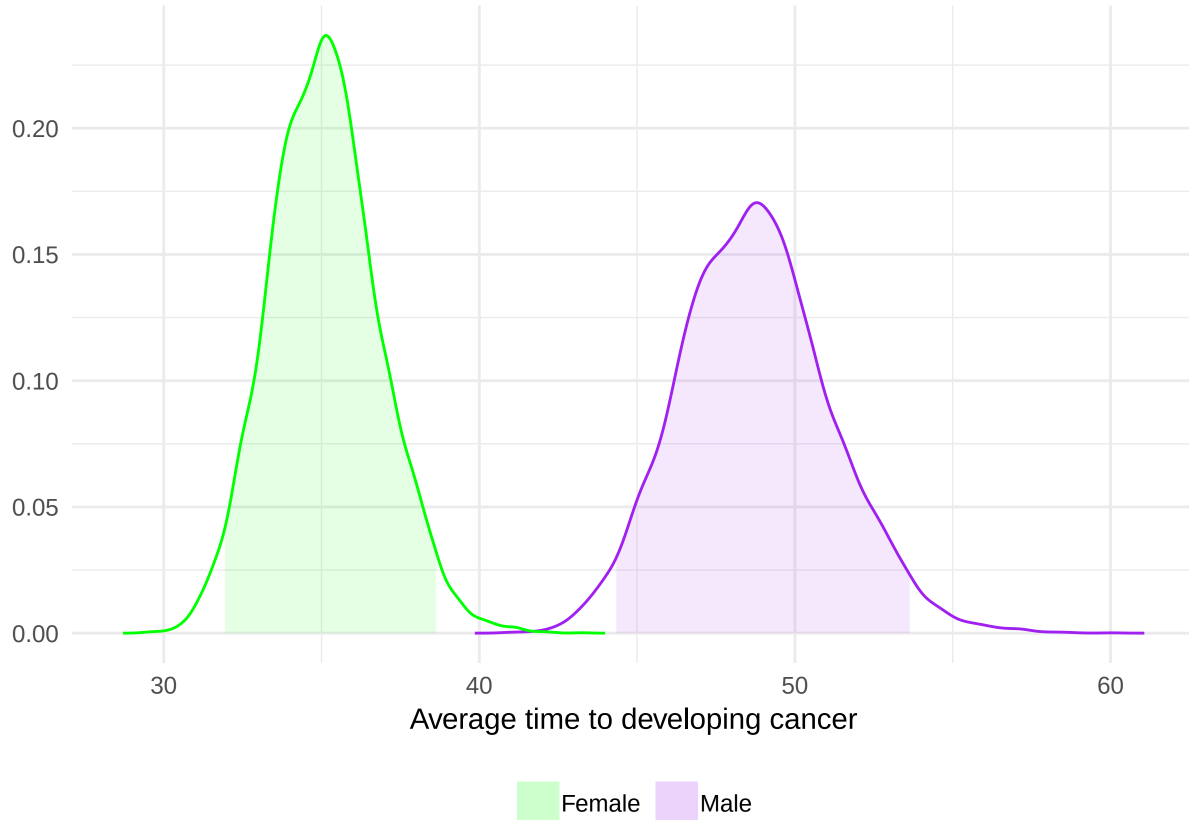
```
##      50%     2.5%    97.5%
## 48.67802 44.31860 53.65796
```

```
## [1] 48.74202
```

```
#Female
quantile( 1/xx  , probs = c(.5,.025,0.975) ); mean(1/xx)
```

```
##      50%     2.5%    97.5%
## 35.06108 31.92115 38.64797
```

```
## [1] 35.10719
```



We see that women have a shorter average time to developing cancer vs men, 35 vs 48.6. It is also interesting to note, that women have a narrower credible interval than men. A pratical implication from these data suggest that women should get tested for cancer at a younger age than men, but when men approach the median time they should get tested more often because their interval of incidence is wider.

2. The dataset `coup1980.txt` contains the coup risk in the month of June from 1980 for 166 different countries. Using your result from Theoretical Exercise 2, build a linear regression model to predict `logCoup` risk using the covariates `democracy` (1 = yes, 0 = no), `age` (the leader's age in years), and `tenure` (the leader's tenure in months). Conduct relevant inference to determine significant predictors and describe how each variable impacted coup risk during June of 1980. Use $B = 10000$ samples and set the seed to 1980. (Hint: your description of the impact can be an interpretation, in context, of the coefficients.)

```
n = nrow(dat); Nsim = 10000
y <- dat[,2]
X <- as.matrix( dat[, - c(1,2 ) ] )
X <- cbind(1,X) ;   colnames(X)[1] <- "(Intercept)"
p <- ncol(X)
BHat <- solve( t(X)%*%X ) %*% t(X) %*% y

set.seed(1980)
lambda <- rgamma(Nsim, (n-p)/ 2 , .5*t( y- X %*% BHat )%*% (y- X %*% BHat ))

set.seed(1980)
Beta <- matrix(NA, Nsim , p )
```

```r
for( i in 1:Nsim){ Beta[i,] <- mvrnorm( 1, BHat, (  1/lambda[i] ) *     solve( t(X)%*% X )    ) }

#GLM results
summary(glm( logCoup~.    , data = dat[,-1 ] ))
```

```
##
## Call:
## glm(formula = logCoup ~ ., data = dat[, -1])
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -3.5459  -0.9093   0.1790   1.0216   3.8848
##
## Coefficients:
##              Estimate Std. Error t value         Pr(>|t|)
## (Intercept) -4.200169   0.567141  -7.406 0.00000000000676 ***
## democracy   -1.826276   0.277603  -6.579 0.00000000062489 ***
## age         -0.024529   0.010930  -2.244         0.026172 *
## tenure      -0.005347   0.001422  -3.760         0.000237 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 2.318831)
##
##     Null deviance: 530.38  on 165  degrees of freedom
## Residual deviance: 375.65  on 162  degrees of freedom
## AIC: 616.66
##
## Number of Fisher Scoring iterations: 2
```

```r
colnames(Beta) <- colnames(X)
round( t( apply(Beta, 2, quantile, probs = c(0.5, 0.025, 0.975)) ) , 4 )
```

```
##                  50%    2.5%   97.5%
## (Intercept) -4.1971 -5.2988 -3.0918
## democracy   -1.8273 -2.3822 -1.2764
## age         -0.0245 -0.0461 -0.0032
## tenure      -0.0053 -0.0081 -0.0025
```

When looking at an x variable, we hold the other x variables fixed. If the country is a democracy, it shifts the regression line downward with a new intercept of about -6.02. Holding other variables constant, for a one-unit increase in age `logCoup` decreases by .02. And holding other variables constant, for a one-unit increase in tenure `logCoup` decreases by .005. We can consider all the coefficients significant because 0 does not appear in any of the credible intervals.