# Assignment 2

Jeff Gould

2/21/2021

## Theoretical Exercises

### 1)

$\mathcal{L}(y|\mu_0) = \prod_{i=1}^n (2\sigma^2\pi)^{-1/2} \exp\left[-1/2\sigma^2(y_i-\theta)^2\right] = (2\sigma^2\pi)^{-n/2} \exp\left[-1/2\sum(y_i-\theta)^2\right]$
$\propto \exp\left[-1/2\sum(y_i-\theta)^2\right]$
$\pi(\theta) = (2\pi\tau_0^2)^{-1/2}\exp[-1/2\tau_0^2(\theta-\mu_0)^2] \propto \exp[-1/2\tau_0^2(\theta-\mu_0)^2]$
$P(\theta|y) \propto \exp\left[\frac{-1}{2\sigma^2}\sum(y_i-\theta)^2\right]\exp\left[\frac{-1}{2\tau_0^2}(\theta-\mu_0)^2\right] = \exp\left[\frac{-1}{2}(\sigma^{-2}\sum(y_i-\theta)^2+{\tau_0}^{-2}(\theta-\mu_0^2)^2)\right] =$

$\exp\left[\frac{-1}{2}(\sigma^{-2}(\sum(y_i^2-2\theta y_i+\theta^2))+{\tau_0}^{-2}(\theta^2-2\theta\mu_0+\mu_0^2))\right] = \exp\left[\frac{-1}{2}(\frac{n\theta^2}{\sigma^2}\sigma^{-2}(\sum(y_i^2-2\theta y_i))+\frac{\theta^2}{\tau_0^2}+{\tau_0}^{-2}(-2\theta\mu_0+\mu_0^2))\right] =$

$\exp\left[\frac{-1}{2}\left(n\theta^2(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})-2\theta\{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}\}+\frac{\mu_0^2}{\tau_0^2}+\frac{\sum y_i^2}{\sigma^2}\right)\right] =$

$\exp\left[\frac{-1}{2}\left(n\theta^2(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})-2\theta\{\frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\}+\frac{\frac{\mu_0^2}{\tau_0^2}+\frac{\sum y_i^2}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right)\right] =$

$\exp\left[\frac{-1}{2}(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})\left(n\theta^2-2\theta\{\frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\}+\frac{\frac{\mu_0^2}{\tau_0^2}+\frac{\sum y_i^2}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right)\right] \propto$

$\exp\left[\frac{-1}{2}(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})\left(\left\{\theta-\frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right\}^2+\frac{\frac{\mu_0^2}{\tau_0^2}+\frac{\sum y_i^2}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}-\left\{\frac{\frac{\mu_o}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right\}^2\right)\right] =$

$\exp\left[\frac{-1}{2}(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})\left\{\theta-\frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right\}^2\right]\exp\left[\frac{-1}{2}(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})\left(\frac{\frac{\mu_0^2}{\tau_0^2}+\frac{\sum y_i^2}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}-\left\{\frac{\frac{\mu_o}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right\}^2\right)\right] \propto$

$\exp\left[\frac{-1}{2}(\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2})\left\{\theta-\frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}\right\}^2\right]$

Which we recognize as the kernel for a normal distribution with:

$$\mu_1 = \frac{\frac{\mu_0}{\tau_0^2}+\frac{\sum y_i}{\sigma^2}}{\frac{1}{\sigma^2}+\frac{1}{n\tau_0^2}}$$

$$\frac{1}{\tau_1^2} = \frac{1}{\sigma^2} + \frac{1}{n\tau_0^2}$$

### 2) $X_i \sim exponential(\lambda)$

$\lambda e^{-\lambda x}$

$\mathcal{L}(x_i|\lambda) = \prod_{i=1}^n \lambda e^{-\lambda x_i} = \lambda^n e^{-\lambda \sum x_i}$

Define:

$C(\theta)^n = \lambda^n$

$T(X) = \sum t(x_i)$

$t(x_i) = x_i$

$w(\theta) = -\lambda$

Then our conjugate prior takes the form:

$\lambda^\eta \exp[-\lambda\nu]$

Which is the kernel for a Gamma distribution

### 3) Let $Z$ be a geometric random variable with probabilty of success $\theta$. Find Jeffrey's Prior

$\mathcal{L}(Z|\theta) = (1-\theta)^Z\theta$

Jeffrey's Prior: $[J(\theta)]^{1/2}$

$J(\theta) = -E\left[\frac{d^2\updownarrow(z|\theta)}{d\theta^2}|\theta\right]$

$\updownarrow(z|\theta) = z\log(1-\theta) + \log(\theta)$

$\frac{d\updownarrow}{d\theta} = -z(1-\theta)^{-1} + \theta^{-1}$

$\frac{d^2\updownarrow}{d\theta^2} = -z(1-\theta)^{-2} - \theta^{-2}$

For a rv $x$ that is geometrically distributed with probability of success $p$: $E[x] = \frac{1-p}{p}$

$-E\left[\frac{d^2\updownarrow(z|\theta)}{d\theta^2}|\theta\right] = -E[-z(1-\theta)^{-2} - \theta^{-2}] = -(1-\theta)^{-2}E[-z] + \theta^{-2} = (-\frac{1}{(1-\theta)^2})(-\frac{1-\theta}{\theta}) + \frac{1}{\theta^2} =$

$\frac{1}{\theta(1-\theta)} + \frac{1}{\theta^2} = \frac{\theta}{\theta^2(1-\theta)} + \frac{(1-\theta)}{\theta^2(1-\theta)} \propto \frac{1}{\theta^2(1-\theta)} = \theta^{-2}(1-\theta)^{-1}$

$[J(\theta)]^{1/2} = (\theta^{-2}(1-\theta)^{-1})^{1/2} = \theta^{-1}(1-\theta)^{-1/2} = \theta^{0-1}(1-\theta)^{1/2-1}$

Which looks like the kernel for a beta distribution, except it is improper as $\alpha = 0$ in this kernel, when it should be strictly greater than 0.

## Analysis Exercises

### 1)

#### a)

$Unif(0,1) \iff beta(1,1)$

Then $\alpha_1 = \alpha_0 + x$, and $\beta_1 = \beta_0 + N - x$, where $x$ is the number of "successes" and $N$ is the sample size

So with the Huff post data, we have $x = 380$ and $N = 1000$

Then: $\alpha_1 = 1 + 380 = 381$ and $\beta_1 = 1 + 1000 - 380 = 621$

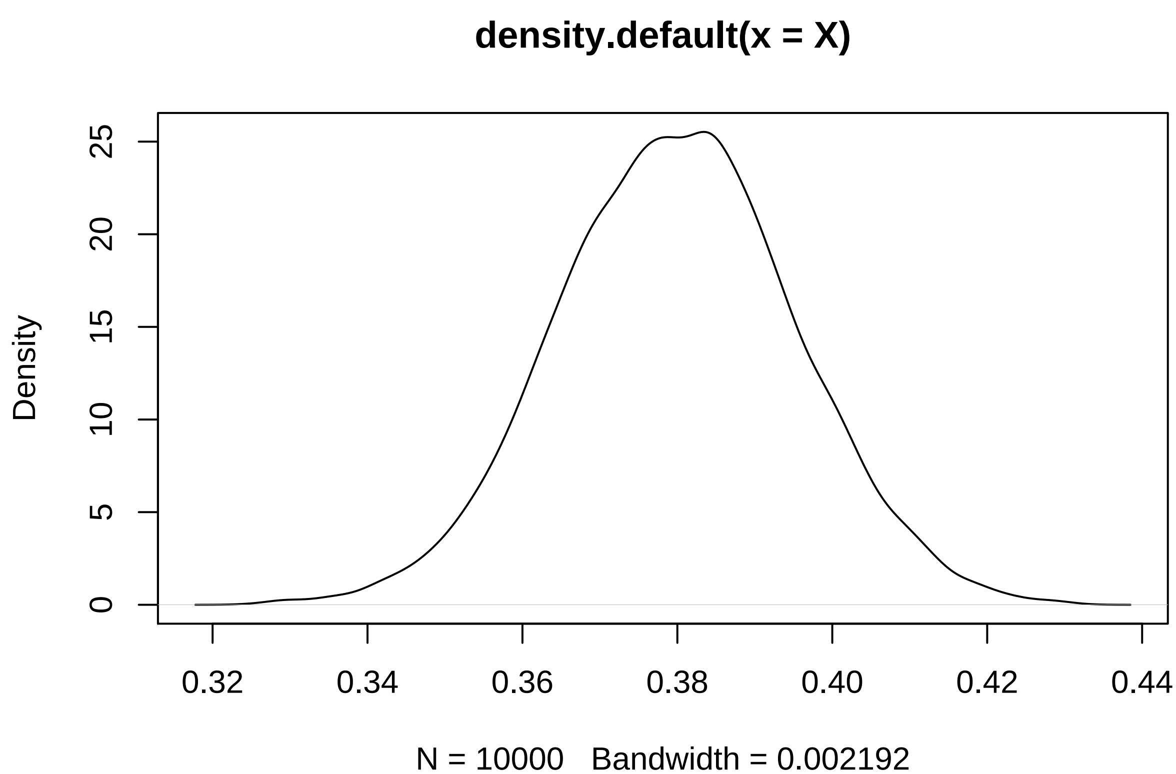
```
alpha = 381
beta = 621

set.seed(2019)
X <- rbeta(n = 10000, shape1 = alpha, shape2 = beta)
```

We find that an estimated 38% of the population is very concerned about the coronavirus, and we have about a 6-percentage point window in which we estimate the true mean lies - (0.3496, 0.4101)

Density Plot:

```
plot(density(X))
```




Median:

```
median(X)
```

```
## [1] 0.3799
```

95% Confidence Interval:

```r
quantile(X, probs = c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 0.3496 0.4101
```

**b)**

$\alpha_2 = \alpha_1 + x_2 = 381 + 450 = 831$

$\beta_2 = \beta_1 + N_2 - x_2 = 621 + 1000 - 450 = 1171$

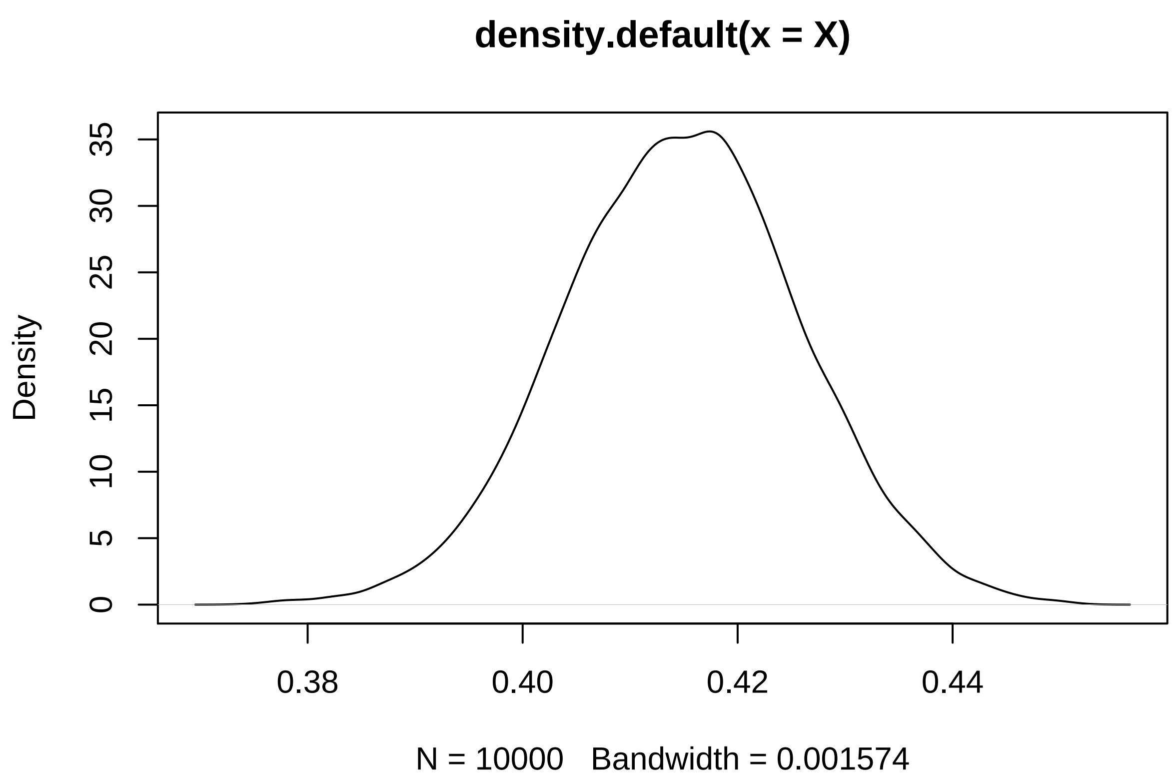
```r
alpha = 831
beta = 1171

set.seed(2019)
X <- rbeta(n = 10000, shape1 = alpha, shape2 = beta)
```

Updating our prior, we now estimate that about 41.5% of the population is very concerned with the coronavirus, and the range of the distribution/confidence interval has shrunk as we have increased our sample size. The window of the confidence interval is now just 4-percentage points - (0.3931, 0.4364)

Density Plot:

```r
plot(density(X))
```




Median:

```r
median(X)
```

```
## [1] 0.4149
```

95% Confidence Interval:

```r
quantile(X, probs = c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 0.3931 0.4364
```

### c)

Using standard uniform prior $\iff$ beta(1,1):

$\alpha_3 = \alpha_0 + x_3 = 1 + 737 = 738$

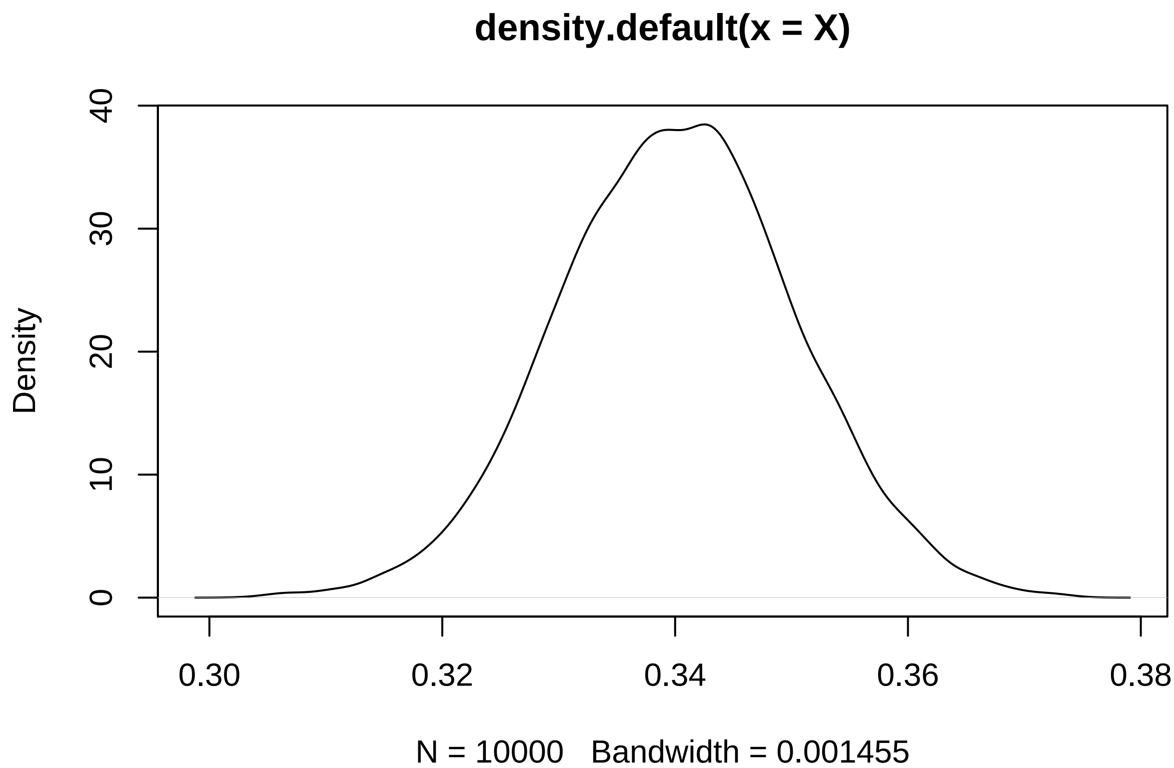
$\beta_3 = \beta_0 + N_3 - x_3 = 1 + 2166 - 737 = 1430$

```r
alpha = 738
beta = 1430

set.seed(2019)
X <- rbeta(n = 10000, shape1 = alpha, shape2 = beta)
```

Density Plot:

```r
plot(density(X))
```




Median:

```r
median(X)
```

```
## [1] 0.3402
```

95% Confidence Interval:

```r
quantile(X, probs = c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 0.3201 0.3602
```

Now using the posterior from b) as our prior:

$\alpha_3 = \alpha_2 + x_3 = 831 + 737 = 1568$

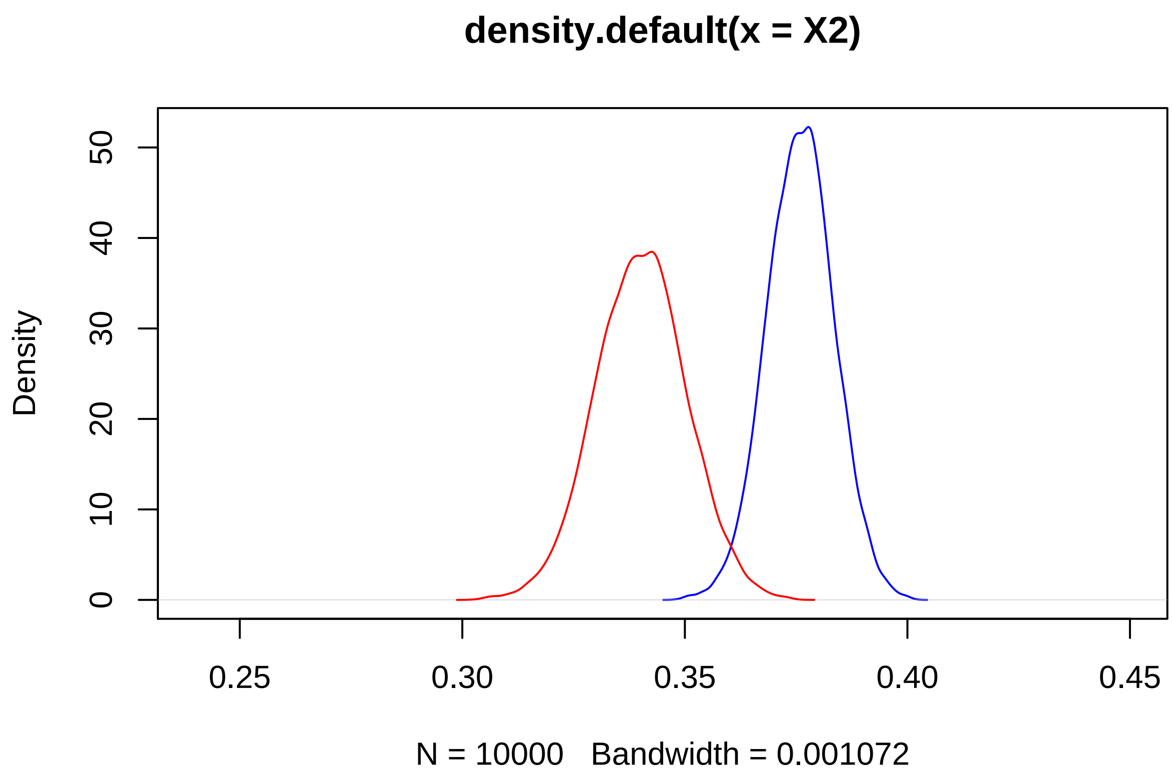
$\beta_3 = \beta_2 + N_3 - x_3 = 1171 + 2166 - 737 = 2600$

```r
alpha = 1568
beta = 2600

set.seed(2019)
X2 <- rbeta(n = 10000, shape1 = alpha, shape2 = beta)
```

Density Plot: We see that the distribution formed using the posterior prior from b) as our conjugate prior has a higher mean and tighter distribution than the posterior distribution formed using the standard uniform as a prior.

```r
plot(density(X2), col = "blue", xlim = c(0.24,0.45))
lines(density(X), col = "red")
```




Median:

```r
median(X2)
```

```
## [1] 0.376
```

95% Confidence Interval: The confidence interval formed using the posterior from b) as the prior has a higher range, and also only a three percentage point window, compared to approximately a 4 percentage point window when we used the standard uniform as our prior

```r
quantile(X2, probs = c(0.025, 0.975))
```

```
##   2.5%  97.5%
## 0.3612 0.3908
```

## 2

$\mathcal{L}(x|\lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \rightarrow \updownarrow(x|\theta) = x\log(\lambda) - \lambda - \log(x!)$

$\frac{d\updownarrow}{d\lambda} = \frac{x}{\lambda} - 1$

$\frac{d^2\updownarrow}{d\lambda^2} = \frac{-x}{\lambda^2}$

$-E\left[\frac{-x}{\lambda^2}\right] = \frac{1}{\lambda}$

$\pi(\lambda) \propto \frac{1}{\lambda^{1/2}}$

$\mathcal{L}(x|\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} \propto e^{-n\lambda}\lambda^{\sum x_i}$

$\mathcal{L}(x|\lambda)\pi(\lambda) \propto e^{-n\lambda}\lambda^{\sum x_i}\frac{1}{\lambda^{1/2}} = \lambda^{\sum x_i + \frac{1}{2} - 1}e^{-n\lambda}$,

Which is the kernal for a gamma distribution:

$P(\lambda|x_1, \ldots x_n) \sim Gamma(\frac{1}{2} + \sum x_i, n)$

```
skin <- read.delim("skin.txt", header = T, sep = " ")
sumX <- sum(skin$numsc)
n <- sum(!is.na(skin$numsc))
```

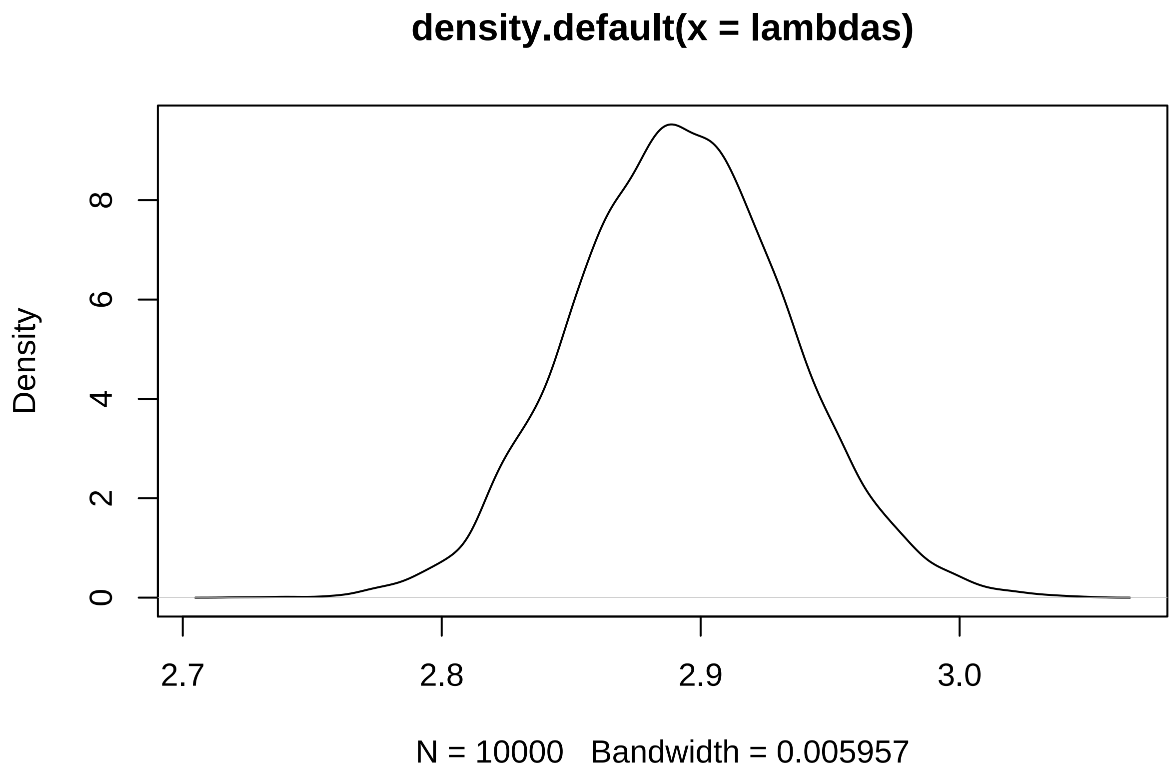
So our posterior for $\lambda$ is

$P(\lambda|x_1, \ldots x_n) \sim Gamma(\frac{1}{2} + 4867, 1683)$

```
alpha <- 1/2 + sumX
beta <- n

set.seed(11)
lambdas <- rgamma(n = 10000, shape = alpha, rate = beta)

plot(density(lambdas))
```




```r
median(lambdas)
```

```
## [1] 2.892
```

```r
quantile(lambdas, probs = c(0.025, 0.975))
```

```
##  2.5% 97.5%
## 2.814 2.976
```