Lab 3

Senan Hogan-H.

8 March 2018

Question 1.

a. Define $s^2 = \frac{1}{n}\sum(x_i - \bar{x})^2$.

$$
\begin{gathered}
\sum_{i=1}^{n}(x_i-\mu)^2 = \\
\sum_{i=1}^{n}(x_i^2 - 2\mu x_i + \mu^2) = \\
\sum_{i=1}^{n}x_i^2 - 2\mu\sum_{i=1}^{n}x_i + n\mu^2 = \\
\sum_{i=1}^{n}x_i^2 - 2n\bar{x}^2 + n\bar{x}^2 + n\bar{x}^2 - 2\mu\sum_{i=1}^{n}x_i + n\mu^2 = \\
\sum_{i=1}^{n}x_i^2 - 2\bar{x}\sum_{i=1}^{n}x_i + n\bar{x}^2 + n\bar{x}^2 - 2\mu\sum_{i=1}^{n}x_i + n\mu^2 = \\
\sum_{i=1}^{n}(x_i^2 - 2\bar{x}x_i + \bar{x}^2) + n(\bar{x}^2 - 2\mu\bar{x} + \mu^2) = \\
\sum_{i=1}^{n}(x_i-\bar{x})^2 + n(\bar{x}-\mu)^2 = \\
ns^2 + n(\bar{x}-\mu)^2
\end{gathered}
\tag{1}
$$

b. Start with a normal prior, $f(\mu) = \mathcal{N}(\mu_0, \sigma_0^2)$.

c. Start with a normal prior, $f(\mu) = \mathcal{N}(\mu_0, \sigma_0^2)$, where $\mu_0 = \sigma_0 = 1$. Suppose that the real distribution has a mean and standard deviation of 3, so that $\bar{x} = 3$.

```
x_bar <- sd <- 3
mu_0 <- 1
sd_0 <- 1

prob <- seq(0, 6, .01)
# First prior
x_0 <- dnorm(prob, mu_0, sd_0)

# Update with data sample of 10
n <- 10

sd_10 <- ((n/(sd^2) + 1/(sd_0^2)))^(-0.5)
mu_10 <- sd_10^2*((mu_0/(sd_0^2)) + (n*x_bar)/sd^2)
```

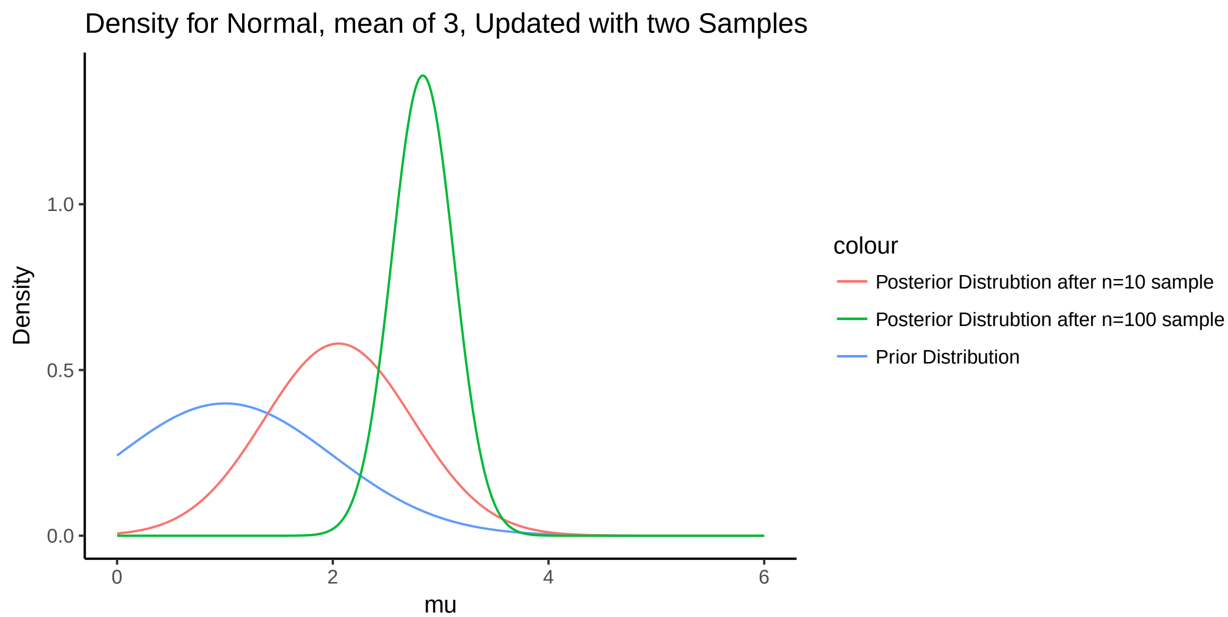
```r
x_10 <- dnorm(prob, mu_10, sd_10)

# Update with another data sample of 100
n <- 100

sd_100 <- ((n/(sd^2) + 1/(sd_0^2)))^(-0.5)
mu_100 <- sd_100^2*((mu_0/(sd_0^2)) + (n*x_bar)/sd^2)

x_100 <- dnorm(prob, mu_100, sd_100)

data.frame(x_0, x_10, x_100, prob) %>% ggplot(aes(x= prob)) +
  geom_line(aes(y= x_0, colour = 'Prior Distribution')) +
  geom_line(aes(y= x_10, colour = 'Posterior Distrubtion after n=10 sample')) +
  geom_line(aes(y= x_100, colour = 'Posterior Distrubtion after n=100 sample')) +
  labs(title = "Density for Normal, mean of 3, Updated with two Samples",
       x = "mu", y = "Density") + theme_classic()
```

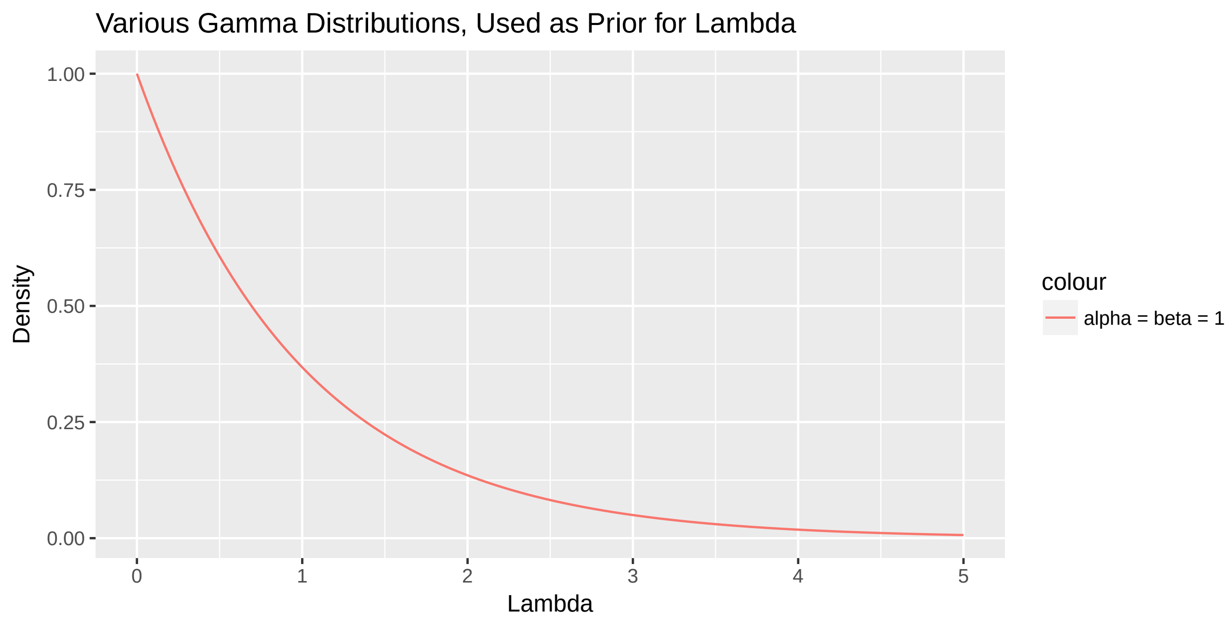


d. Use a prior of the form $\mathcal{G}(\alpha_0, \beta_0)$ where $\alpha_0, \beta_0 > 0$ to estimate λ.

```r
prob <- seq(0, 5, .01)

alpha_0 <- 1
beta_0 <- 1
x_0 <- dgamma(prob, alpha_0, beta_0)

data.frame(x_0, prob) %>%
  ggplot(aes(x = prob)) +
  geom_line(aes(y = x_0, colour = 'alpha = beta = 1')) +
  labs(title = "Various Gamma Distributions, Used as Prior for Lambda",
       x = "Lambda", y = "Density")
```



e. THIS QUESTION IS ACTUALLY NOT POSSIBLE WITH GIVEN PRIOR. CHANGE PROBABILITY TO 0.95.

For the prior, want $\sigma^2 = 4$, $P(\sigma^2 > 2) = 0.95$.

So that $\frac{\alpha_0 - 1}{\beta_0} = 4$, $pbeta(2, \alpha_0, \beta_0) = 1 - 0.95 = 0.05$

```
x <- 2

# start with 5 going to 1 to ensure alpha_0, beta_0 > 0
#vector <- rev(seq(1, 5, by = 0.001))
vector <- seq(1, 3, by = 0.00001)
i <- 1

# (alpha_0 - 1)/beta_0 = 4, so beta_0 = (alpha_0 - 1)/4
alpha_0 <- vector[i]
beta_0 <- (alpha_0 - 1)/4

epsilon <- 0.0001
while (abs(pgamma(x, alpha_0, beta_0) - 0.05) > epsilon){
  #print(i)
  i <- i + 1
  alpha_0 <- vector[i]
  beta_0 <- (alpha_0 - 1)/4
}
rm(vector, i)
alpha_0
```

```
## [1] 1.174
```

```
beta_0
```

```
## [1] 0.04341
```

```
pgamma(x, alpha_0, beta_0) # roughly = 0.05
```

```
## [1] 0.0499
```

f.

```r
# test data
x <- rnorm(10000, 0, 1)

# test hyperparameters
alpha_0 <- 1
beta_0 <- 1


# function that will be defined.
Normal_estimator <- function (x, alpha_0, beta_0){

  return(mean(x))
}
```

g.

Consulted the following resource in working on this exercise: https://www.cs.ubc.ca/~murphyk/Papers/bayesGauss.pdf