Conjugate Priors

Andy Shen, Devin Francom

Statistical Sciences Group (CCS-6) - Los Alamos National Laboratory

Problem 1. Binomial

```
p1 <- read.table("https://madison.byu.edu/bayes/binomial2.dat", header = FALSE)
```

```
sum_ni <- sum(p1$V2); sum_ni
```

```
## [1] 476
```

```
sum_yi <- sum(p1$V3); sum_yi
```

```
## [1] 73
```

```
sum_ni_yi <- sum(p1$V2 - p1$V3); sum_ni_yi
```

```
## [1] 403
```

We have that $\sum_{i=1}^{153} n_i = 476$, $\sum_{i=1}^{153} y_i = 73$, and $\sum_{i=1}^{139} n_i - y_i = 403$.

The conjugate prior distribution is a **Beta** distribution with parameters a and b. $\theta \sim Beta(a, b)$.

It follows that

$$P(\theta) \propto \theta^{a-1}(1-\theta)^{b-1}$$

The likelihood $L(y \mid \theta)$ of $P(y \mid \theta)$ is given by

$$L(y \mid \theta) = \prod_{i=1}^{159} \binom{n_i}{y_i} \theta^{y_i}(1-\theta)^{n_i - y_i} \propto \theta^{73}(1-\theta)^{403}$$

Multiplying the likelihood by the prior, we get that

$$P(\theta \mid y) \propto \theta^{73+a-1}(1-\theta)^{403+b-1}$$

which follows a Beta distribution with parameters $73 + a$ and $403 + b$.

We select $a = 1$ and $b = 249$ so that our posterior follows a Beta distribution with parameters 74 and 652. This way, our expected value is around 0.10, which is the record proportion of at-bats that are home runs.

The prior distribution is $Beta(1, 652)$ such that

$$P(\theta \mid y) \propto \theta^{74-1}(1-\theta)^{652-1}$$

Deriving the maximum likelihood estimate of our likelihood $L(y \mid \theta) = \theta^{73}(1-\theta)^{403}$

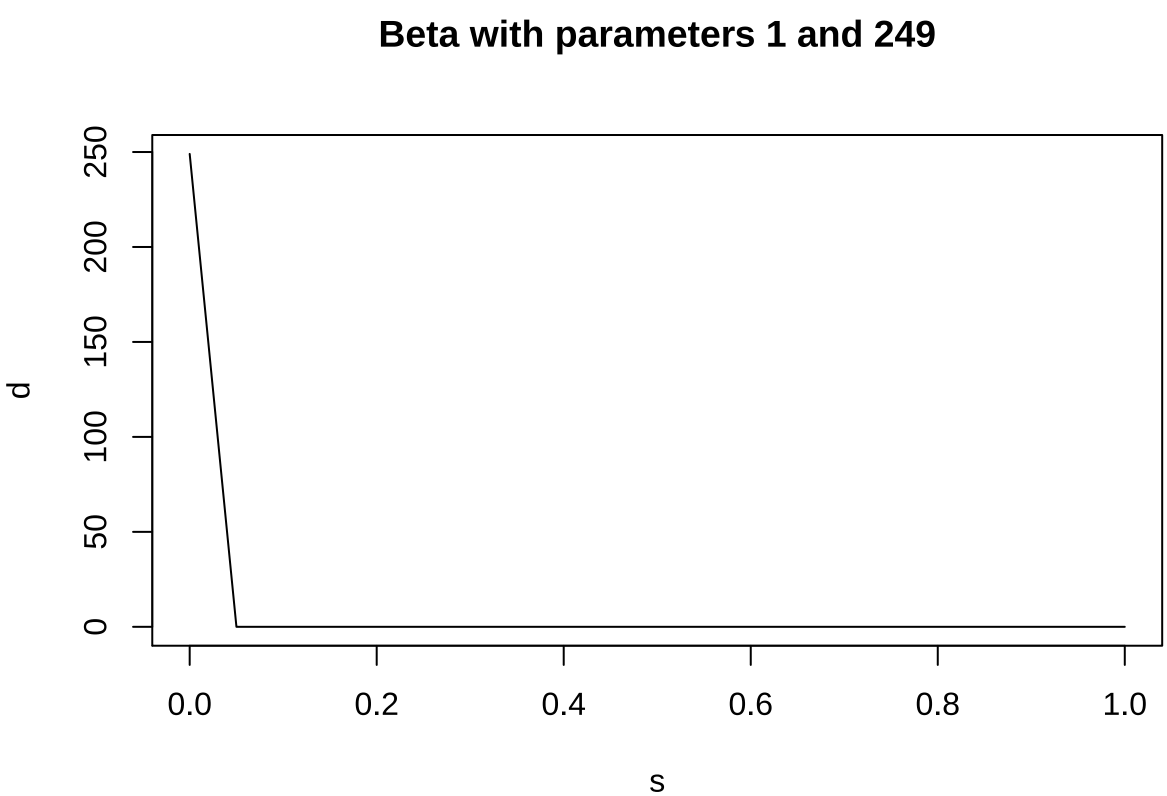
$$\log L(y \mid \theta) = 73\log(\theta) + 403\log(1-\theta)$$
$$\frac{\partial \log L(y \mid \theta)}{\partial \theta} = \frac{\partial}{\partial \theta}(73\log(\theta) + 403\log(1-\theta))$$
$$0 = \frac{73}{\theta} - \frac{403}{1-\theta}$$

$$\frac{73}{\theta} = \frac{403}{1-\theta}$$

$$\theta = \frac{73}{476} \approx 0.1533613$$

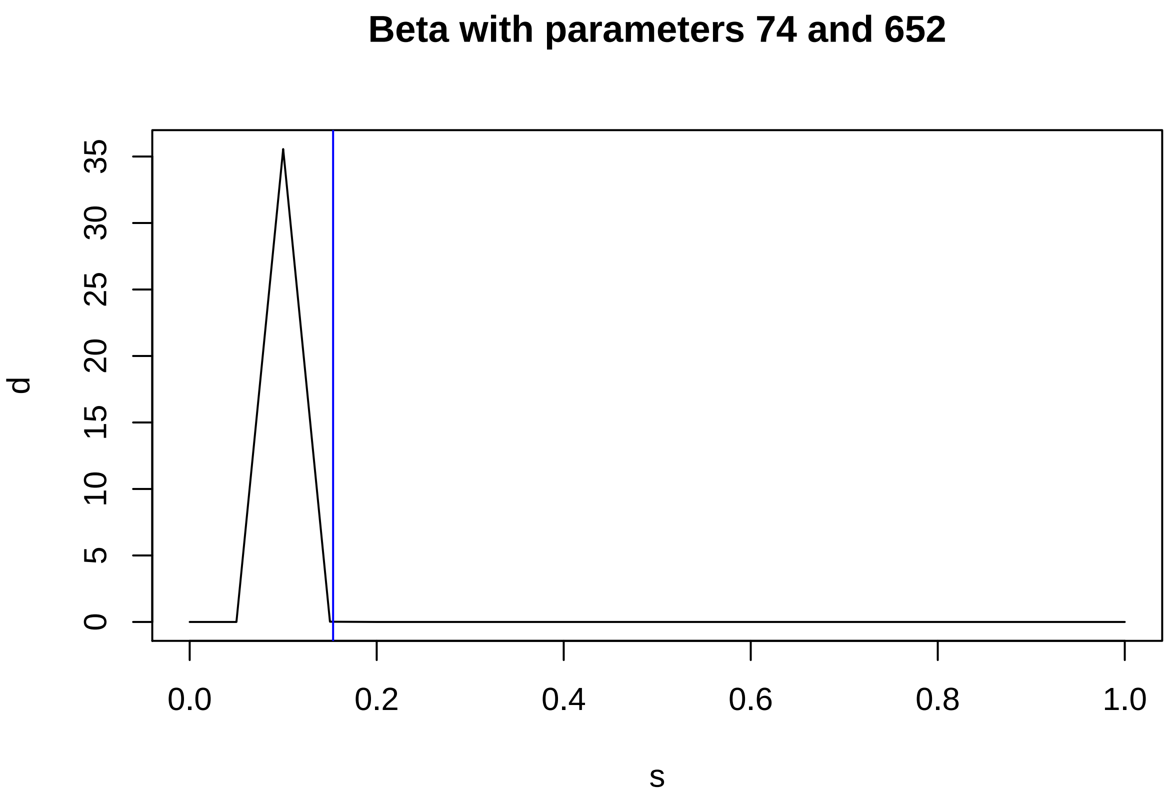
Plotting the prior distribution, we get:

```
a <- 1
b <- 249
s <- seq(0, 1, by = 0.05)
d <- dbeta(s, a, b)
title <- paste0("Beta with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
```



Plotting the posterior distribution, we get:

```r
mle <- 73/476
a <- 74
b <- 652
s <- seq(0, 1, by = 0.05)
d <- dbeta(s, a, b)
title <- paste0("Beta with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
abline(v = mle, col = "blue")
```



Posterior Mean

The expected value of a Beta Distribution is

$$E(X) = \frac{a}{a+b} = \frac{74}{726} \approx 0.102$$

Posterior Median

```r
n <- rgamma(10e6, 74, 652)
median(n)
```

```
## [1] 0.1129851
```

Posterior Mode

$$\frac{a-1}{a+b-2} = \frac{73}{724} \approx 0.101$$

Posterior Variance

$$Var(X) = \frac{ab}{(a+b)^2(a+b+1)} = 0.0001259133$$

Posterior Standard Deviation

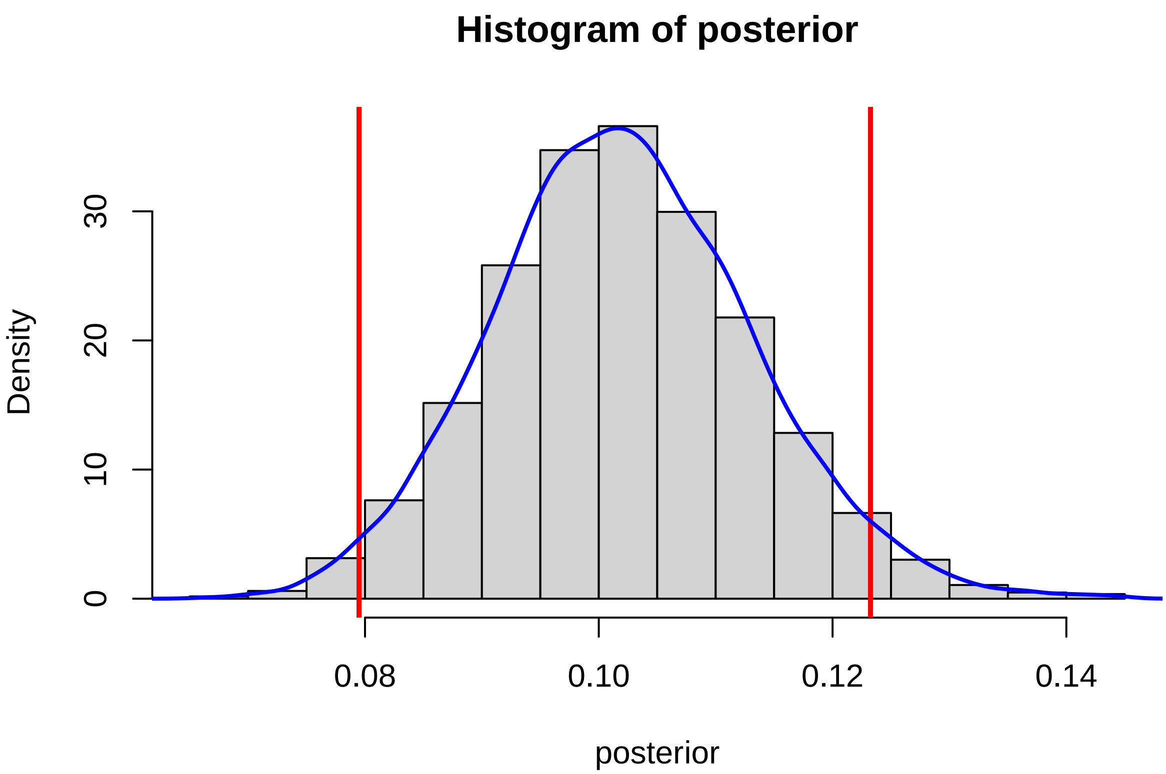
```
sd <- sqrt(0.0001259133); sd
```

```
## [1] 0.01122111
```

95% Credible Interval

Acknowledgments: https://cran.r-project.org/web/packages/bayestestR/vignettes/credible_interval.html#what-is-a-credible-interval

```
posterior <- rbeta(10000, 74, 652)
ci_hdi <- ci(posterior, method = "HDI", ci = 0.95)
low <- ci_hdi$CI_low
high <- ci_hdi$CI_high
ci <- c(low, high)
hist(posterior, prob = TRUE)
abline(v = ci, col = "red", lwd = 2.5)
lines(density(posterior), lwd = 2,col = "blue")
```



Problem 2. Poisson

```
p2 <- read.table("https://madison.byu.edu/bayes/poisson.dat")
p2 <- as.vector(p2[,1])
p2
```

```
##  [1] 3 4 3 1 1 4 1 4 2 1 1 3 1 1 4
```

```
sum(p2)
```

```
## [1] 34
```

The conjugate prior of a Poisson distribution is a Gamma distribution with parameters a and b.

$$\theta \sim \Gamma(a, b)$$

$$L = \prod_{i=1}^{15} \frac{\lambda^{y_i} exp(-\lambda)}{y_i!}$$

$$\sum_{i=1}^{15} y_i = 34$$

$$L(y \mid \theta) \propto \lambda^{34} exp(-15\lambda)$$

$$P(\theta) = \frac{\lambda^{a-1} exp(\frac{-\lambda}{b})}{b^a \Gamma(a)} \propto \lambda^{a-1} exp(-\frac{\lambda}{b})$$

It then follows that

$$P(\theta \mid y) \propto \lambda^{34+a-1} exp(-\lambda(15 + \frac{1}{b}))$$

We see that $P(\theta \mid y)$ is proportional to a gamma distribution with parameters $34 + a - 1$ and $1 + \frac{1}{b}$, $\Gamma(34 + a, \frac{1}{15+\frac{1}{b}})$.

We let $a = b = 1$ so that

$$P(\theta \mid y) \propto \lambda^{35-1} exp(-\frac{\lambda}{16})$$

$$P(\theta \mid y) \sim \Gamma(35, \frac{1}{16})$$

We compute the MLE:

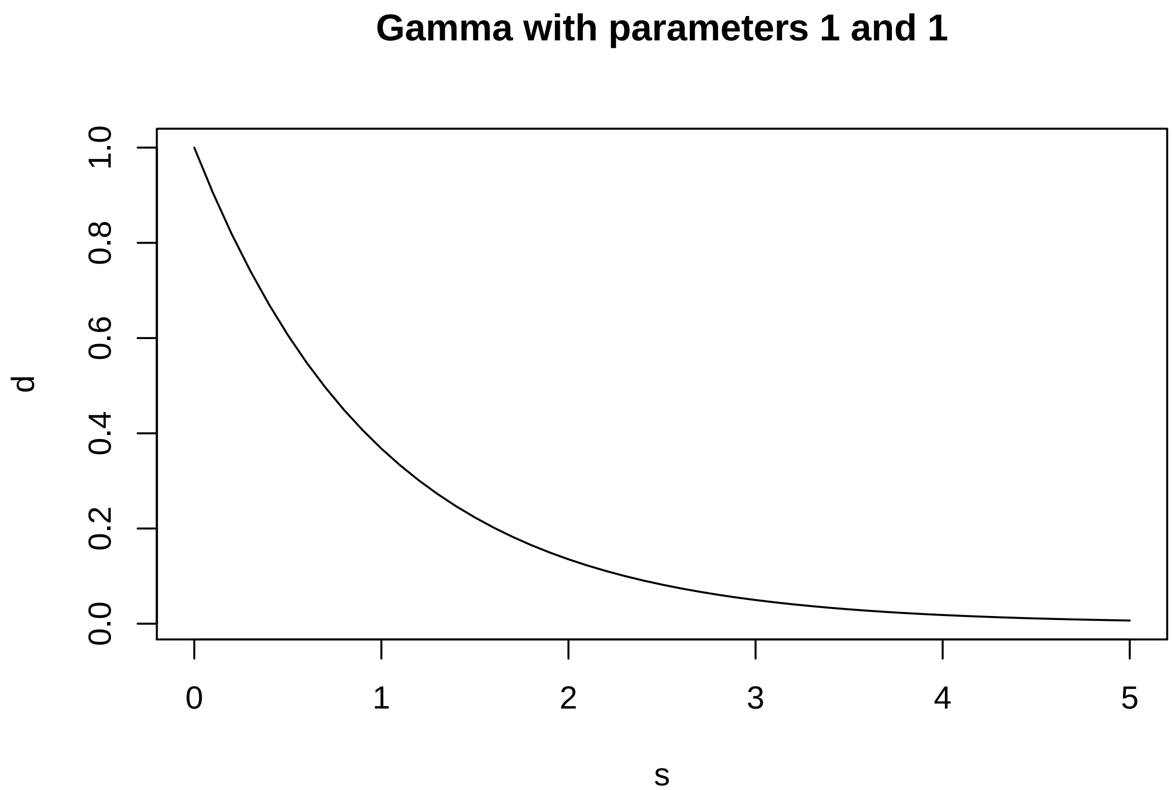
$$L(y \mid \theta) = \lambda^{34} exp(-15\lambda)$$

$$\log L(y \mid \theta) = 34 \log(\lambda) - \frac{\lambda}{15}$$

$$\frac{\partial \log L(y \mid \theta)}{\partial \lambda} = \frac{34}{\lambda} - \frac{1}{15} = 0$$
$$\hat{\lambda} = 34(15) = 510$$

Plotting the prior distribution, we get

```
a <- 1
b <- 1
s <- seq(0,5, by = 0.1)
d <- dgamma(s, a, b)
title <- paste0("Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
```

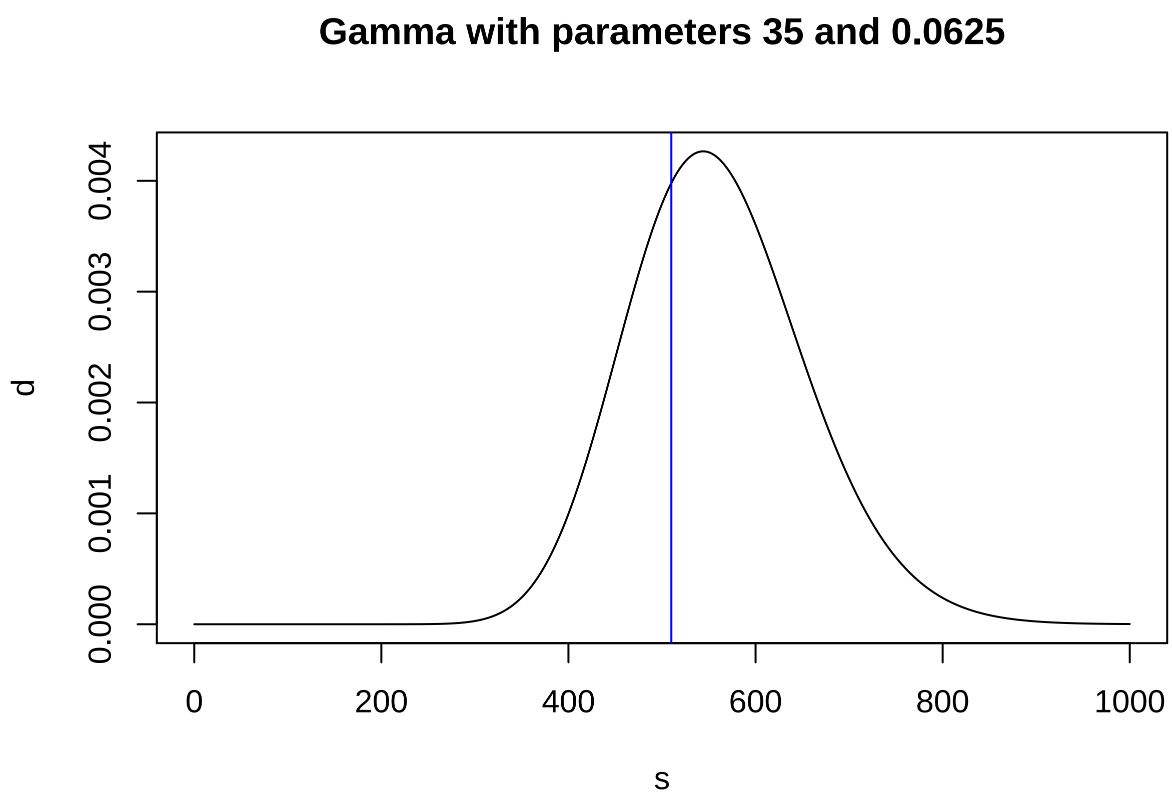


Our posterior is

$$\Gamma(35, \frac{1}{16})$$

Plotting the posterior distribution, we get

```
a <- 35
b <- 1/16
s <- seq(0, 1000, by = 0.1)
d <- dgamma(s, a, b)
title <- paste0("Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title); abline(v = 510, col = "blue") #MLE
```



Posterior Mean

$$E(X) = ab = (35)16 = 560$$

Posterior Median

```r
n <- rgamma(10e6, 35, 1/16)
median(n)
```

```
## [1] 554.6399
```

Posterior Mode

$$(a-1)b = 544$$

Posterior Variance

$$Var(X) = ab^2 = 8960$$

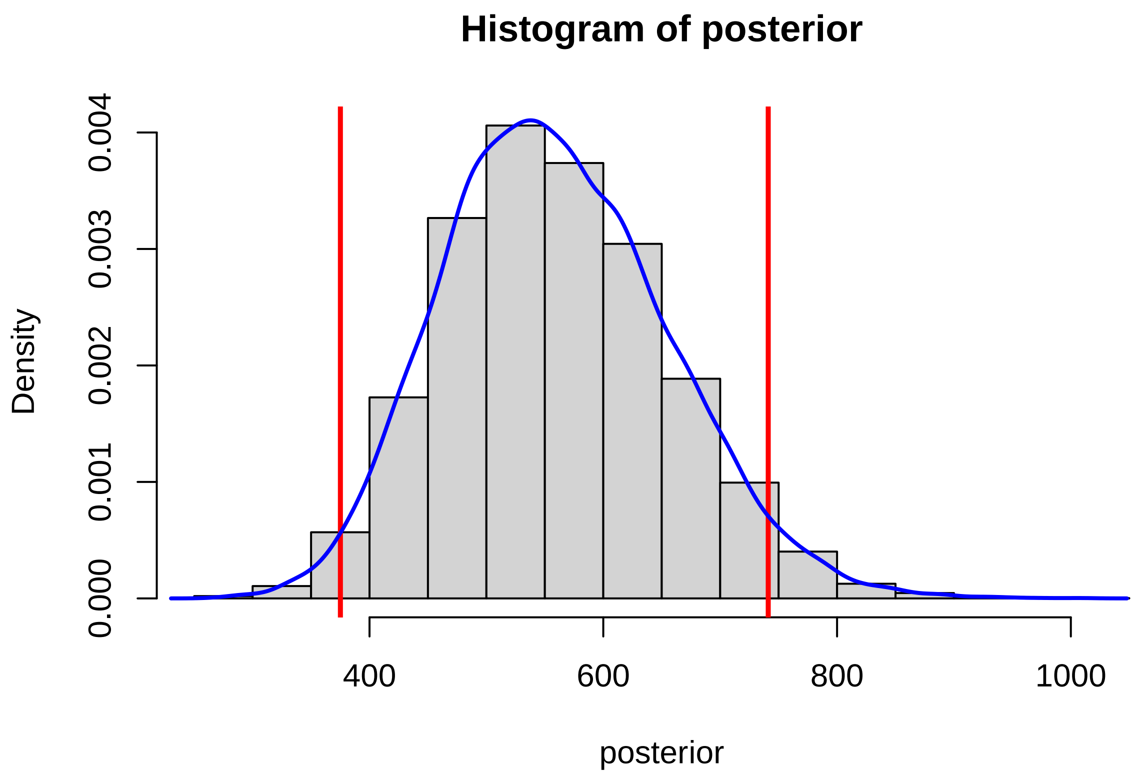
Posterior Standard Deviation

```r
sqrt(8960)
```

```
## [1] 94.65728
```

95% Credible Interval

```
posterior <- rgamma(10000, 35, 1/16)
ci_hdi <- ci(posterior, method = "HDI", ci = 0.95)
low <- ci_hdi$CI_low
high <- ci_hdi$CI_high
ci <- c(low, high)
hist(posterior, prob = TRUE)
abline(v = ci, col = "red", lwd = 2.5)
lines(density(posterior),lwd=2,col="blue")
```



Problem 3. Exponential

```r
p3 <- read.table("http://madison.byu.edu/bayes/exponential.dat")
p3 <- as.vector(p3[,1])
```

$$P(y_i \mid \lambda) = \frac{1}{\lambda} exp(-\frac{y_i}{\lambda})$$

```r
sum(p3)
```

```
## [1] 83357
```

$$\sum_{i=1}^{141} y_i = 83357$$

Then,

$$L(y \mid \lambda) = \prod_{i=1}^{141} \lambda^{-1} exp(-\lambda y_i) = \lambda^{-141} exp(-\frac{83357}{\lambda})$$

The conjugate prior of an Exponential distribution is an Inverse Gamma Distribution with parameters a and b.

Our prior:

$$P(\lambda) \propto \lambda^{-a-1} \exp\{-b/\lambda\}$$

Our posterior:

$$P(\lambda|y) \propto L(y|\lambda)P(\lambda)$$

$$\propto \lambda^{-141-a-1} exp\{-\frac{1}{\lambda}(83357 + b)\}$$
$$\propto \Gamma^{-1}(141 + a, 83357 + b)$$

Computing the MLE we get:

$$L(y \mid \lambda) = \prod_{i=1}^{141} \lambda^{-1} exp(-\lambda y_i) = \lambda^{-141} exp(-\frac{83357}{\lambda})$$

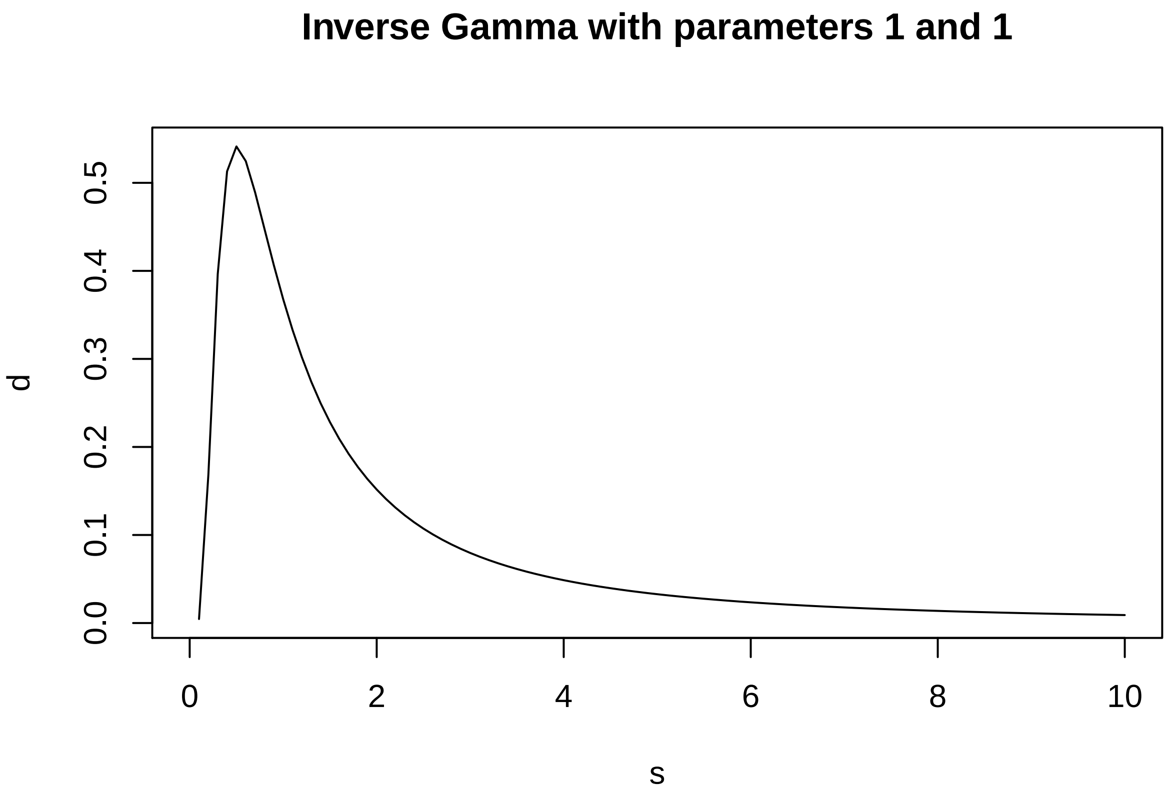
$$\log L(y \mid \lambda) = -141 \log \lambda - \frac{83357}{\lambda}$$

$$\frac{\partial L(y \mid \lambda)}{d\lambda} = -\frac{141}{\lambda} + \frac{83357}{\lambda^2} = 0$$
$$-141 + \frac{83357}{\lambda} = 0$$

$$\hat{\lambda} = \frac{83357}{141} = 591.1844$$

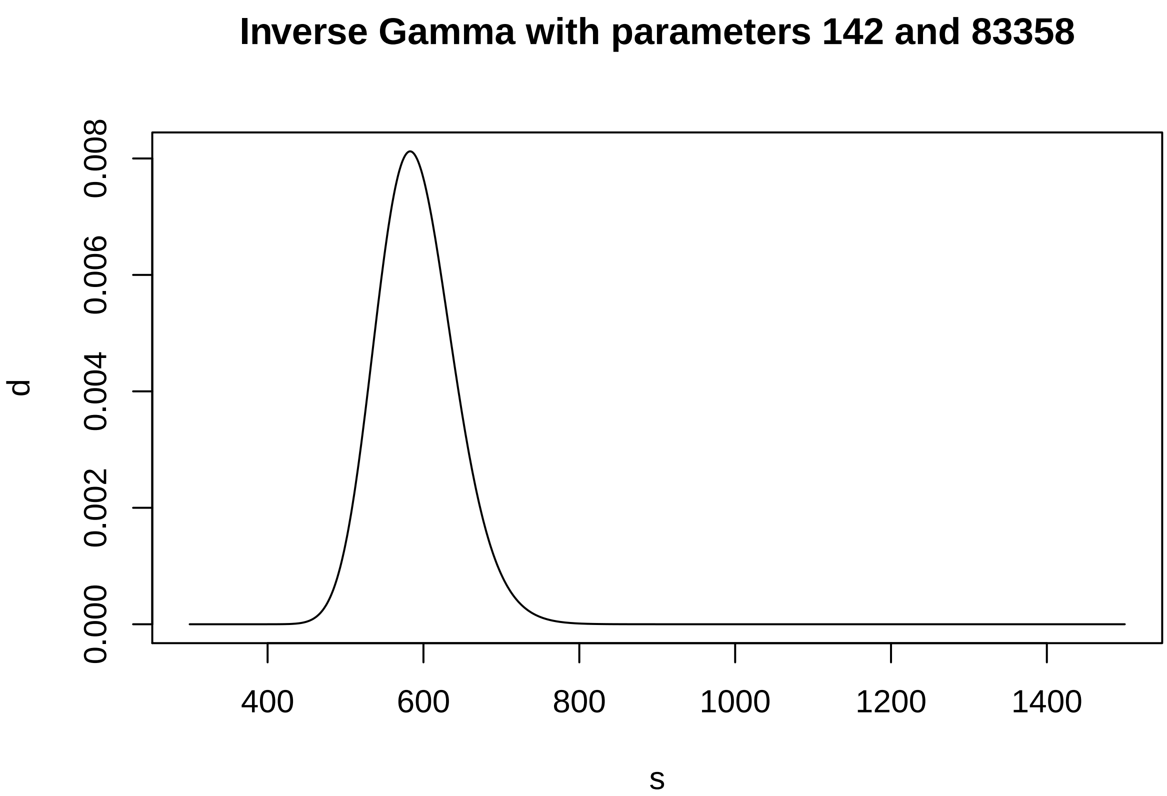
We let $a = 100$ and $b = 1$. Plotting the prior distribution, we get

```
library(invgamma)
a <- 1
b <- 1
s <- seq(0, 10, by = 0.1)
d <- dinvgamma(s, a, b)
title <- paste0("Inverse Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
```



Plotting the posterior distribution, we get

```
a <- 142
b <- 83358
s <- seq(300, 1500)
d <- dinvgamma(s, a, b)
title <- paste0("Inverse Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
abline(v = 0.121644, col = "blue")
```



$$a = 142 b = 83358$$

Posterior Mean

$$E(X) = \frac{b}{a-1} = 591.1915$$

Posterior Median

```r
n <- rinvgamma(1e6, 142, 83358)
median(n)
```

```
## [1] 588.3826
```

Posterior Mode

$$\frac{b}{a+1} = 582.9231$$

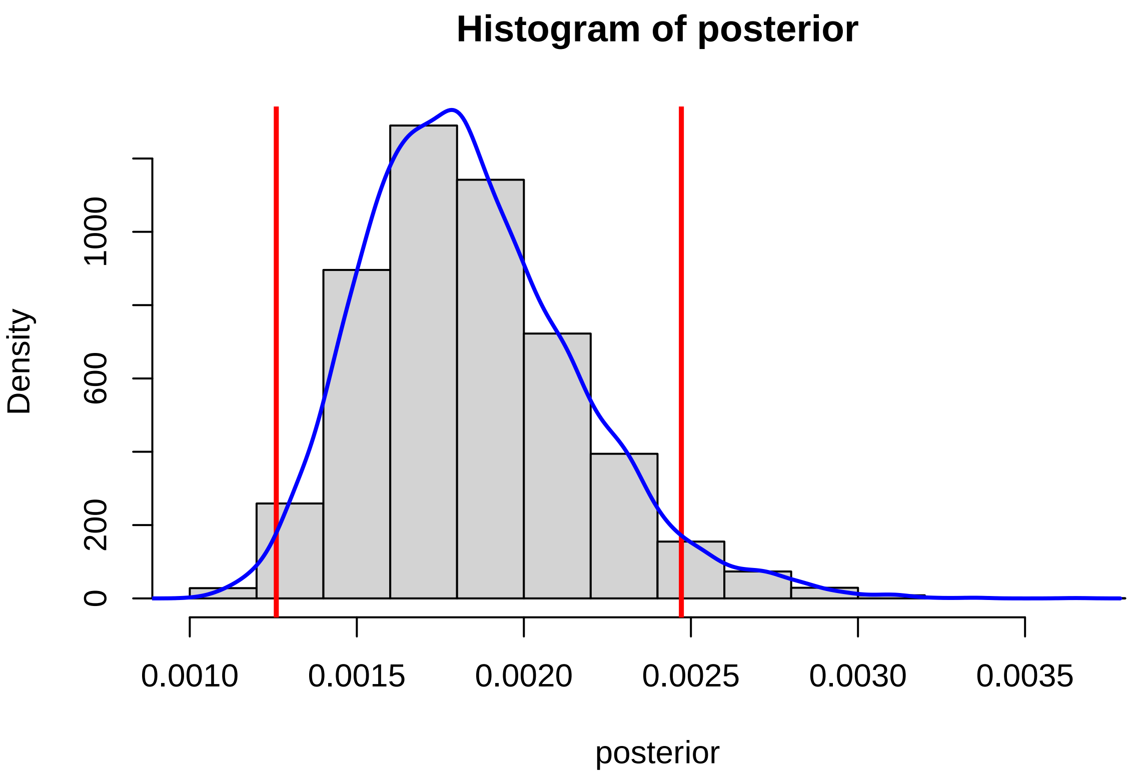
Posterior Variance and SD

$$Var(X) = \frac{b^2}{(a-1)^2(a-2)} = 2496.481$$

$$sd(X) = 49.9648$$

95% Credible Interval

```
posterior <- rinvgamma(10000, 35, 1/16)
ci_hdi <- ci(posterior, method = "HDI", ci = 0.95)
low <- ci_hdi$CI_low
high <- ci_hdi$CI_high
ci <- c(low, high)
hist(posterior, prob = TRUE)
abline(v = ci, col = "red", lwd = 2.5)
lines(density(posterior), lwd = 2, col = "blue")
```



Problem 4. Normal with Constant Variance

```
p4 <- read.table("http://madison.byu.edu/bayes/normalmean.dat")
p4 <- as.vector(p4$V1)
p4
```

```
##  [1]  98  92  89  77  87  84  75  73  95  86  67  86  86 100 100  92 100  97  95
## [20]  77  87  87  95  84  84  74  86  84  94
```

```
sum(p4)
```

```
## [1] 2531
```

$$\sum_{i=1}^{29} y_i = 2531$$

$$P(y \mid \theta) = \frac{1}{9\sqrt{2\pi}} e^{\frac{-(y_i-\theta)^2}{162}}$$

$$L(y \mid \theta) \propto e^{\frac{-\sum(y_i-\theta)^2}{162}}$$

$$P(\theta) = \frac{1}{\sigma_0\sqrt{2\pi}} e^{\frac{-(\theta-\mu_0)^2}{2\sigma_0^2}}$$

Then, we have that our posterior likelihood, $P(\theta \mid y)$ is proportional to

$$e^{(\frac{-\sum(y_i-\theta)^2}{162} - \frac{-(\theta-\mu_0)^2}{2\sigma_0^2})} (1)$$

We know our posterior is proportional to

$$e^{\frac{-(\theta-\mu_1)^2}{2\sigma_1^2}} (2)$$

For some μ_1 and σ_1

We then expand (1) and set it equal to (2) to get the result that

$$\theta \sim N(\frac{1}{\frac{n}{81} + \frac{1}{\sigma_0^2}}(\frac{2531}{81} + \frac{\mu_0}{\sigma_0^2}), \sqrt{(\frac{n}{81} + \frac{1}{\sigma_0^2})^{-1}})$$

We let $\mu_0 = 85$ and $\sigma_0^2 = 12$.

Then, we have that

$$P(\theta \mid y) \sim N(86.84615, 1.505236)$$

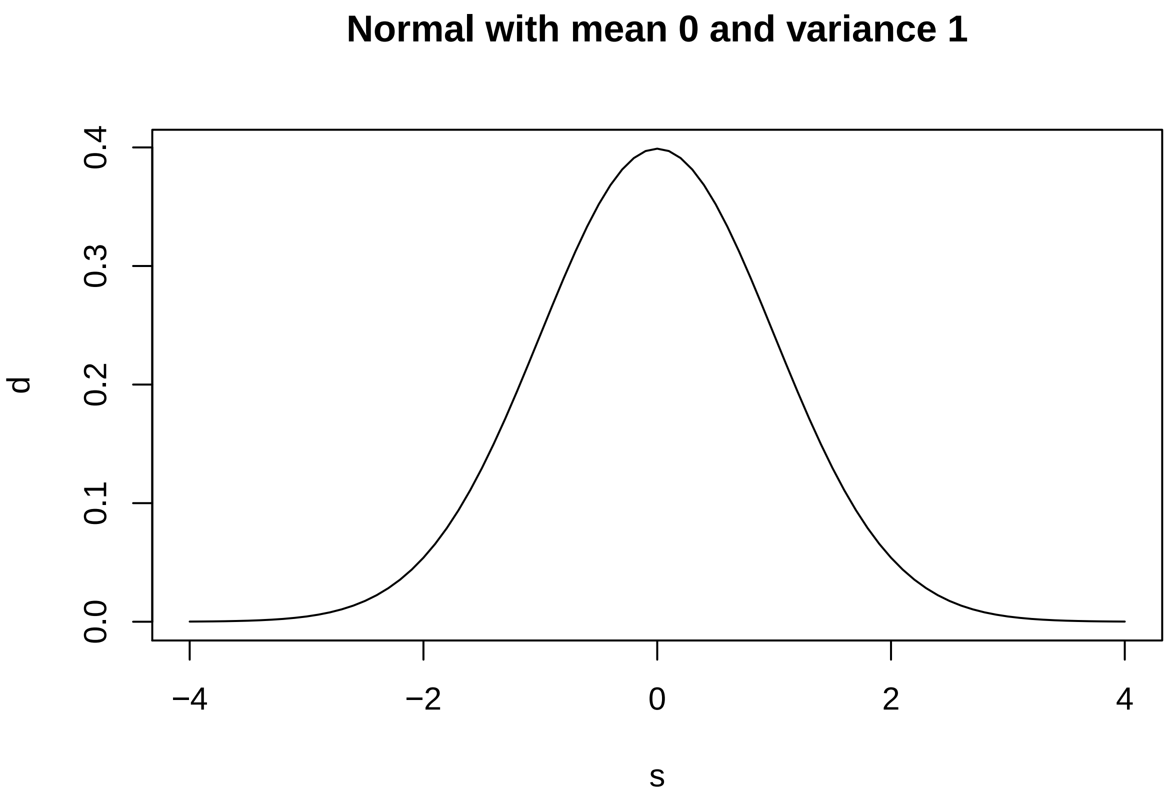
We calculate the MLE:

$$L(y \mid \theta) = (2\pi\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\sum(y_i-\mu)^2}$$

$$\log L = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum(y_i - \mu)^2$$
$$\hat{\mu} = \bar{y} = 86.84615$$
$$\hat{\sigma}^2 = 2.265734$$

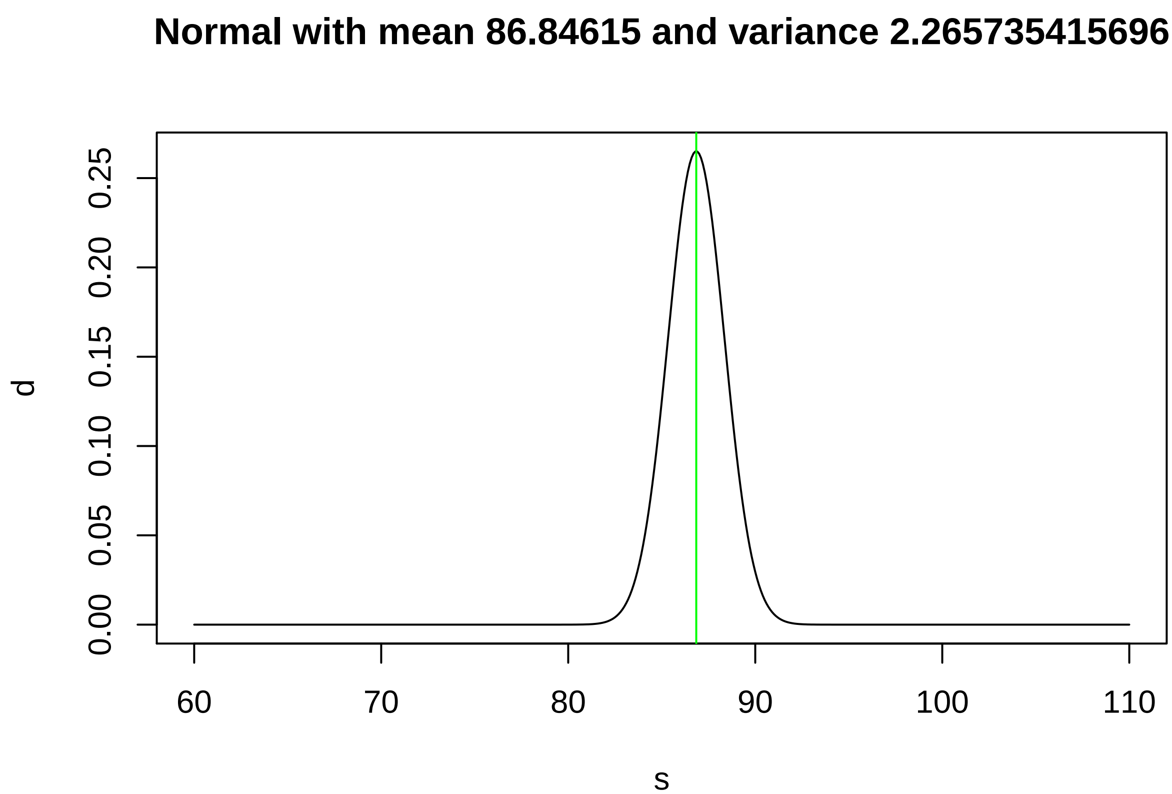
Plotting our prior distribution, we get:

```r
mu <- 0
sd <- 1 # Already the defaults for dnorm() I think
s <- seq(-4, 4, by = 0.1)
d <- dnorm(s)
title <- paste0("Normal with mean ", mu, " and variance ", sd^2, "\n")
plot(s, d, type = "l", main = title)
```



Plotting the posterior, we get

```r
mu <- 86.84615
sd <- 1.505236
s <- seq(60, 110, by = 0.1)
d <- dnorm(s, mu, sd)
title <- paste0("Normal with mean ", mu, " and variance ", sd^2, "\n")
plot(s, d, type = "l", main = title)
abline(v = mu, col = "green")
```



Posterior Mean

```
mu
```

```
## [1] 86.84615
```

Posterior Median

```
mu
```

```
## [1] 86.84615
```

```
n <- rnorm(10e6, mu, sd)
median(n) #interesting finding
```

```
## [1] 86.8465
```

Posterior Mode

```
mu
```

```
## [1] 86.84615
```

Poster Variance and Standard Deviation

```
var <- sd^2; var
```

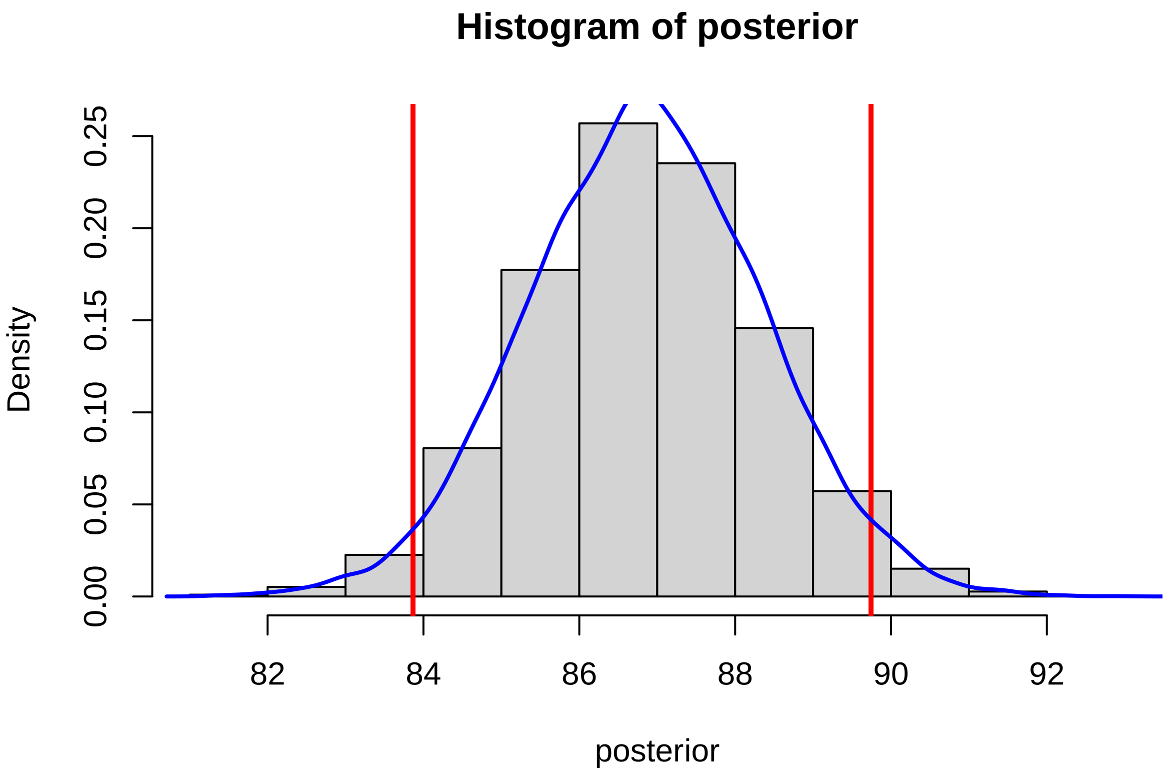
```
## [1] 2.265735
```

```
sd
```

```
## [1] 1.505236
```

95% Credible Interval

```
posterior <- rnorm(10000, mu, sd)
ci_hdi <- ci(posterior, method = "HDI", ci = 0.95)
low <- ci_hdi$CI_low
high <- ci_hdi$CI_high
ci <- c(low, high)
hist(posterior, prob = TRUE)
abline(v = ci, col = "red", lwd = 2.5)
lines(density(posterior), lwd = 2, col = "blue")
```



Problem 5. Normal with Constant Mean

```r
p5 <- read.table("http://madison.byu.edu/bayes/normalvariance.dat")
p5 <- as.vector(p5$V1)
p5
```

```
##  [1]  98  92  89  77  87  84  75  73  95  86  67  86  86 100 100  92 100  97  95
## [20]  77  87  87  95  84  84  74  86  84  94
```

```r
sum(p5)
```

```
## [1] 2531
```

$$\sum_{i=1}^{29} y_i = 2531$$

We are given that $\mu = 85$.

We use the **Inverse Gamma Distribution** as our conjugate prior distribution.

$$P(y \mid \theta) = \frac{1}{\theta\sqrt{2\pi}} e^{\frac{-(y_i-85)^2}{2\theta^2}}$$

Then, we have that

$$L(y \mid \theta) \propto (\sigma^2)^{-\frac{n}{2}} e^{\frac{\sum_{-}(y_i-85)^2}{2\sigma^2}}$$

$$P(\theta) \propto (\sigma^2)^{-a-1} e^{(-\frac{2b}{2\sigma^2})}$$

Since our mean μ is fixed our parameter θ is σ^2.

Then,

$$P(\sigma^2 \mid y) \propto L(y \mid \sigma^2)P(\sigma^2)$$

$$\propto (\sigma^2)^{-\frac{n}{2}-a-1} e^{-\frac{2b+\sum_{(} y_i-85)^2}{2\sigma^2}}$$

$$= (\sigma^2)^{-(\frac{n}{2}+a)-1} e^{-\frac{b+0.5\sum_{(} y_i-85)^2}{\sigma^2}} \propto (\sigma^2)^{-(\alpha_1)-1} e^{-\frac{\beta_1}{\sigma^2}}$$

We see that our variance σ^2 follows an Inverse Gamma distribution with parameters $a+\frac{n}{2}$ and $b+\frac{1}{2}\sum(y_i-85)^2$. Therefore,

$$\sigma^2 \sim \Gamma^{-1}(a + \frac{n}{2}, b + \frac{1}{2}\sum(y_i - 85)^2)$$

$$\sigma^2 \sim \Gamma^{-1}(a + \frac{29}{2}, b + \frac{1}{2}\sum(y_i - 85)^2)$$

Expanding the sum $\sum(y_i - 85)^2)$ gives us:

$$\sum(y_i^2 - 170y_i + 85^2)$$

$$= \sum y_i^2 - 170\sum y_i + (29)85^2$$

$$= 223065 - 430270 + 209525 = 2320$$

Therefore, our posterior σ^2 is distributed as follows:

$$\sigma^2 \sim \Gamma^{-1}(a + \frac{29}{2}, b + \frac{1}{2}\sum(y_i - 85)^2)$$

$$\sigma^2 \sim \Gamma^{-1}(a + \frac{29}{2}, b + 1160)$$

If we choose $a = \frac{5}{2}$ and $b = 120$, we get that

$$\sigma^2 \sim \Gamma^{-1}(17, 1280)$$

So our expected value is

$$E(X) = \frac{b}{a-1} = 80$$

Computing the MLE, we get that

$$L(y \mid \sigma^2) \propto (\sigma^2)^{-\frac{n}{2}} e^{\frac{\sum_{-}(y_i-85)^2}{2\sigma^2}}$$

$$\log L(y \mid \sigma^2) = -\frac{n}{2}\log(\sigma^2) - \frac{2320}{2\sigma^2}$$

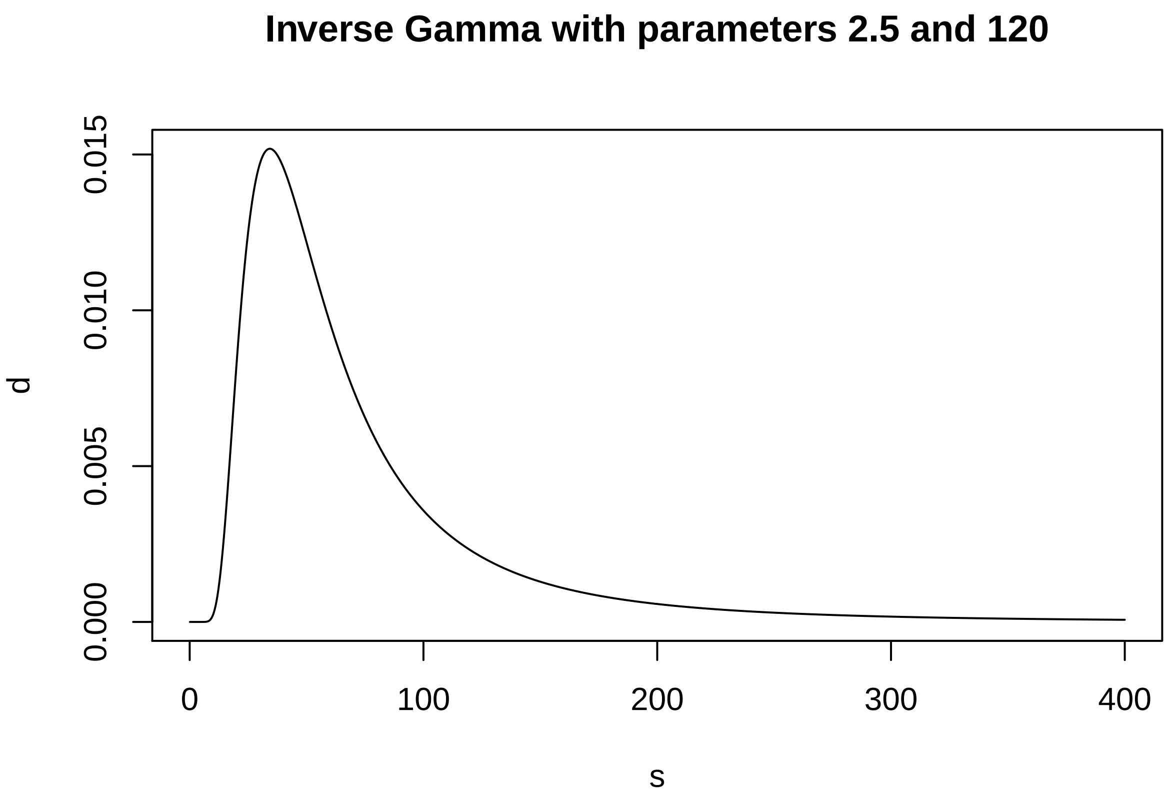
$$\frac{\partial L(y \mid \sigma^2)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{2320}{2\sigma^4} = 0$$

$$n = 29 = \frac{2320}{\sigma^2}$$

$$\sigma^2 = \frac{2320}{29} = 80$$

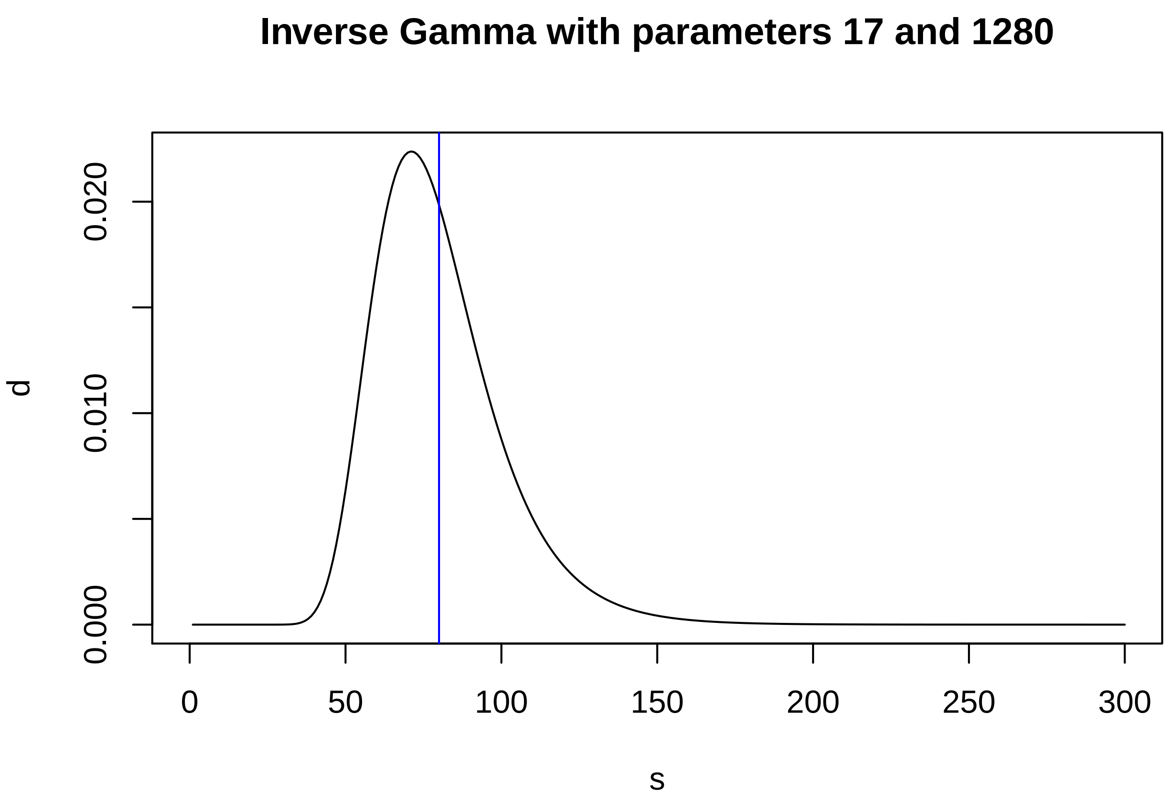
Plotting the prior, we get that

```
library(invgamma)
a <- 2.5
b <- 120
s <- seq(0, 400, by = 0.1)
d <- dinvgamma(s, a, b)
title <- paste0("Inverse Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
```



Plotting the posterior distribution, we get

```
a <- 17
b <- 1280
s <- seq(0, 300)
d <- dinvgamma(s, a, b)
title <- paste0("Inverse Gamma with parameters ", a, " and ", b, "\n")
plot(s, d, type = "l", main = title)
abline(v = 80, col = "blue")
```



Posterior Mean

$$E(X) = \frac{b}{a-1} = 80$$

Posterior Median

```r
n <- rinvgamma(10e6, 17, 1280)
median(n)
```

```
## [1] 76.79556
```

Posterior Mode

$$\frac{b}{a+1} = 71.111$$

Posterior Variance and SD

$$Var(X) = \frac{b^2}{(a-1)^2(a-2)} = 426.667$$

$$SD = 20.656$$

95% Credible Interval

```
posterior <- rinvgamma(10000, 17, 1280)
ci_hdi <- ci(posterior, method = "HDI", ci = 0.95)
low <- ci_hdi$CI_low
high <- ci_hdi$CI_high
ci <- c(low, high)
hist(posterior, prob = TRUE)
abline(v = ci, col = "red", lwd = 2.5)
lines(density(posterior), lwd = 2, col = "blue")
```