

BDA - Assignment 3

3/10/2021

Contents

1	1
a)	1
b)	2
2	4
b)	6
3	11
b)	13

1

In the exercise we will be using a uninformative prior $p(\mu, \sigma^2) \propto \frac{1}{\sigma}$. The likelihood is the normal distribution $p(y|\mu, \sigma^2) = \mathcal{N}(\mu, \sigma^2)$. We get the posterior by the product $p(\mu, \sigma^2)p(y|\mu, \sigma^2) = \frac{\mathcal{N}(\mu, \sigma^2)}{\sigma}$.

The marginalized posterior density for μ follows Student's t distribution $t_{n-1}(\bar{y}, s^2/n)$ where.

a)

We will calculate a point estimate as the mean. The mean of the posterior can simply be calculated as:

```
mu_point_est = function(data){
  return(mean(data))
}

mu_interval = function(data, prob){
  mu = mu_point_est(data)
  n = length(data)
  freedom = n-1
  scale = sqrt(var(data)/n)
  low = (1-prob)/2
  high= 1-low
  int = c(low,high)
  qt_scl <- qtnew(int,freedom)
  res = c(mu + qt_scl[1]*scale, mu + qt_scl[2]*scale)
  return(res)
}
```

Test data:

Real data:

```

mu = mu_point_est(windshieldy1)
interval = mu_interval(data = windshieldy1, prob = 0.95)
mu

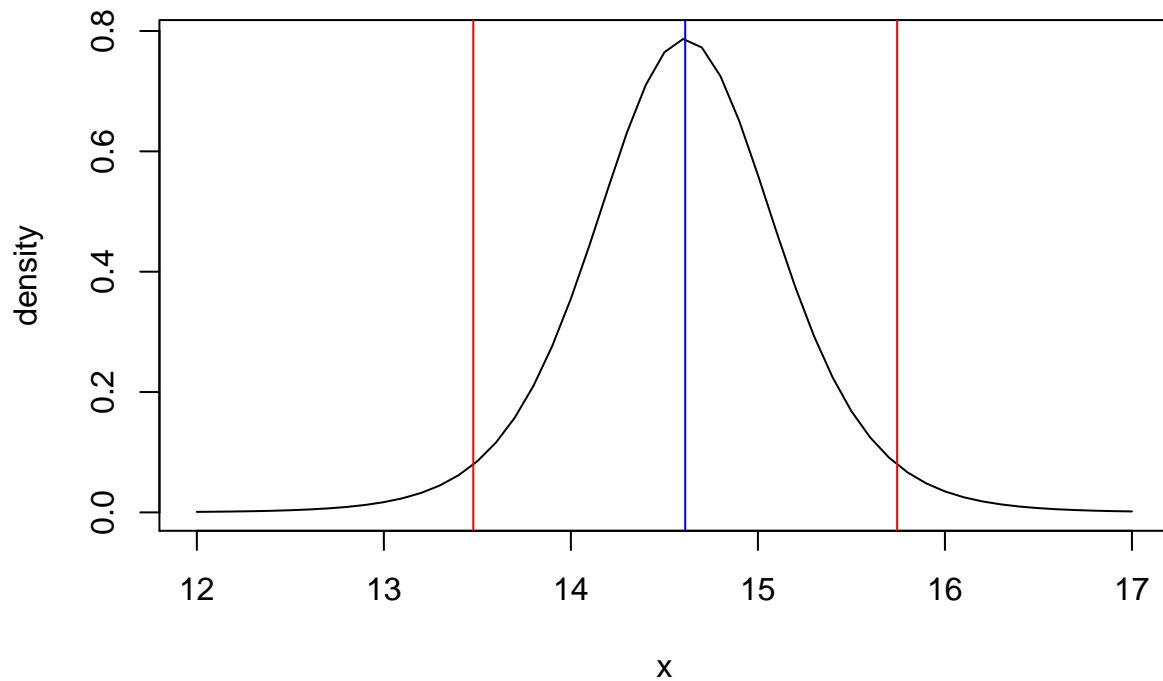
## [1] 14.61122
interval

## [1] 13.47808 15.74436

n = length(windshieldy1)
x = seq(from=12,to=17, by=0.1)
freedom = n-1
density = dtnew(x, freedom, mean=mu, scale=sqrt(var(windshieldy1)/n) )
plot(x,density,
      xlab = 'x',
      type='l')

abline(v = mu, col = "blue")
abline(v = interval[1], col = "red")
abline(v = interval[2], col = "red")

```



b)

In order to predict a future value we need the future posterior predictive distribution. The distribution is a student-t distribution with location \bar{y} , scale $(1 + \frac{1}{n})^{1/2}s$ and $n-1$ degrees of freedom.

```

mu_pred_point_est = function(data){
  return(mean(data))
}

mu_pred_interval = function(data, prob){
  mu = mu_point_est(data)
  n = length(data)
  freedom = n-1
  scale = sqrt(var(data)/n)
  low = (1-prob)/2
  high= prob+low
  int = c(low,high)
  res <- qtnew(int,freedom, mean=mu, scale = scale)
  return(res)
}

mu = mu_pred_point_est(data = windshields1)
interval = mu_pred_interval(data = windshields1, prob = 0.95)
mu

## [1] 14.61122
interval

## [1] 13.47808 15.74436

```

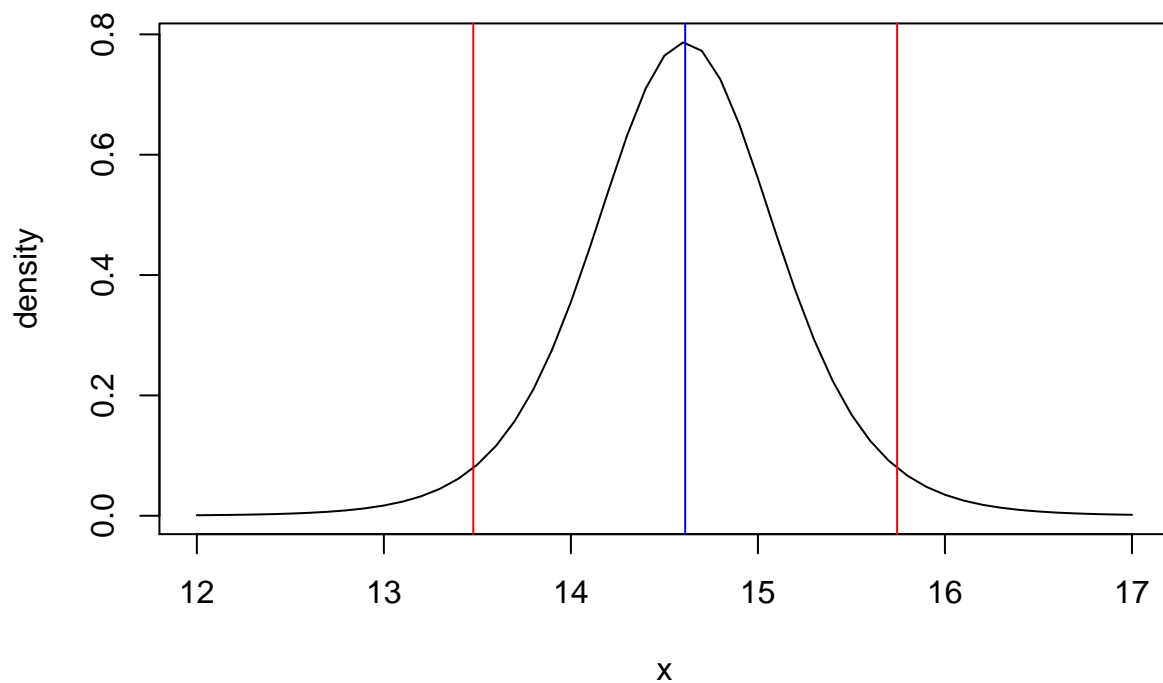
The mean, used for the point estimate is 14.6 and the intervals lower bound 13.5 and upper bound 15.7. Next we will plot the density function with the and the interval.

```

n = length(windshields1)
x = seq(from=12,to=17, by=0.1)
freedom = n-1
density = dtnew(x, freedom, mean=mu, scale=sqrt(var(windshields1)/n) )
plot(x,density,
     xlab = 'x',
     type='l')

abline(v = mu, col = "blue")
abline(v = interval[1], col = "red")
abline(v = interval[2], col = "red")

```



2

The likelihood $p(y|\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$, the prior $p(\pi) = \text{Beta}(\alpha, \beta)$ and thus the likelihood becomes $p(y|\pi) = \binom{n}{y}\pi^y(1-\pi)^{n-y}$. Using a uniform Beta, $\text{Beta}(1, 1)$ as our weakly informative prior we get that the posterior is $p(\pi|y) = \text{Beta}(\pi|1+y, 1+n-y)$

```
posterior_odds_ratio_point_est = function(p0, p1){
  ratio = (p1/(1 - p1))/(p0/(1 - p0))
  return(mean(ratio))
}

posterior_odds_ratio_interval = function(p0, p1, prob){
  data = (p1/(1 - p1))/(p0/(1 - p0))
  lower = (1-prob)/2
  upper = prob+lower
  interval = c(lower,upper)
  res = quantile(data,interval)
  return(res)
}
```

```
n0 = 674
y0 = 39
n1 = 680
y1 = 22

prior_alpha = 1
```

```

prior_beta = 1

post_alpha0 = prior_alpha + n0 - y0
post_beta0 = prior_beta + y0

post_alpha1 = prior_alpha + n1 - y1
post_beta1 = prior_beta + y1

p0 <- rbeta(100000, post_alpha0, post_beta0)
p1 <- rbeta(100000, post_alpha1, post_beta1)

point_est = posterior_odds_ratio_point_est(p0 = p0, p1 = p1)
interval = posterior_odds_ratio_interval(p0 = p0, p1 = p1, prob = 0.95)
paste0("The point estimate is: ", round(point_est,digits=2))

## [1] "The point estimate is: 1.89"

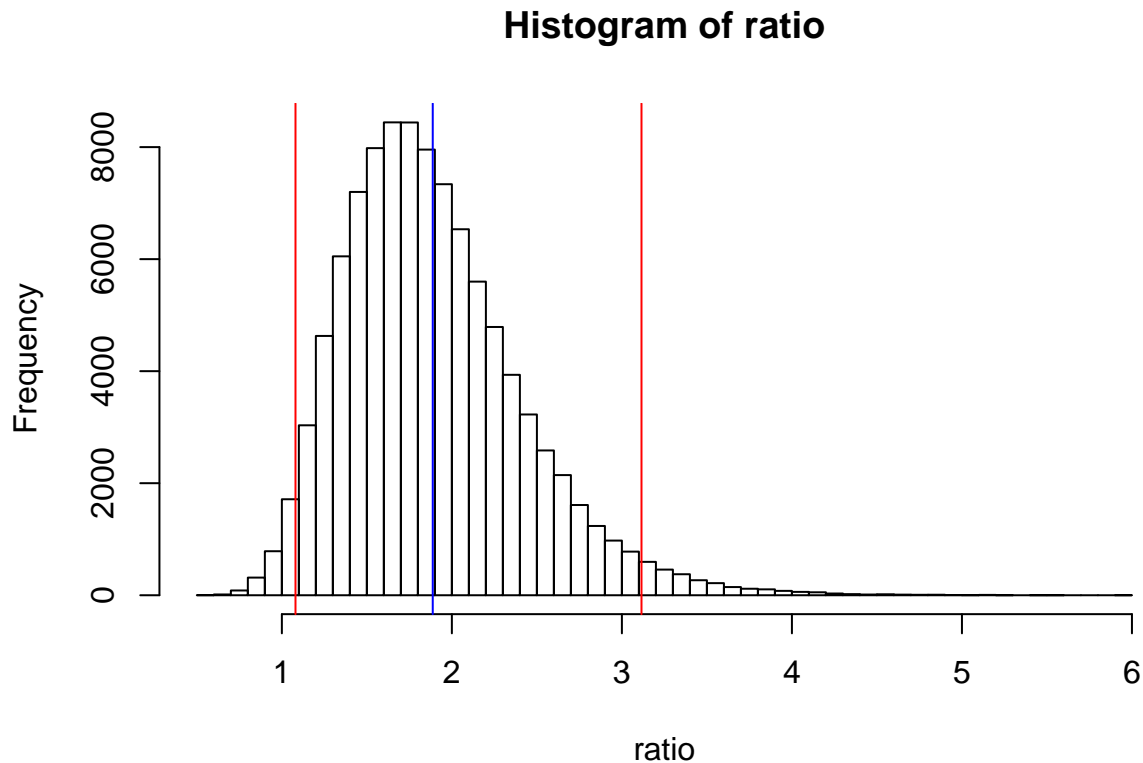
paste0("The interval is: [", round(interval[1],digits=2), ",", round(interval[2],digits=2),"]")

## [1] "The interval is: [1.08,3.12]"

The interval is [1.08, 3.12]

ratio = (p1/(1 - p1))/(p0/(1 - p0))
hist(ratio, breaks=50)
abline(v = interval[1], col='red')
abline(v = interval[2], col='red')
abline(v = point_est, col='blue')

```



b)

By using a uniform prior, $\alpha = \beta = 1$, we have decided to “let the data speak for itself”. The sum $\alpha + \beta = 2$ essentially tells us how much prior data we have used. We will now use different values for α and β and plot our results.

```
prior_alpha = c(0.1,0.5,1,5,100)
prior_beta = prior_alpha

post_alpha0 = numeric(length(prior_alpha))
post_beta0 = numeric(length(prior_alpha))
post_alpha1 = numeric(length(prior_alpha))
post_beta1 = numeric(length(prior_alpha))
p0 = numeric(length(prior_alpha))
p1 = numeric(length(prior_alpha))
point_est = numeric(length(prior_alpha))
interval = numeric(length(prior_alpha)*2)

post_alpha0 = prior_alpha + n0 - y0
post_beta0 = prior_beta + y0

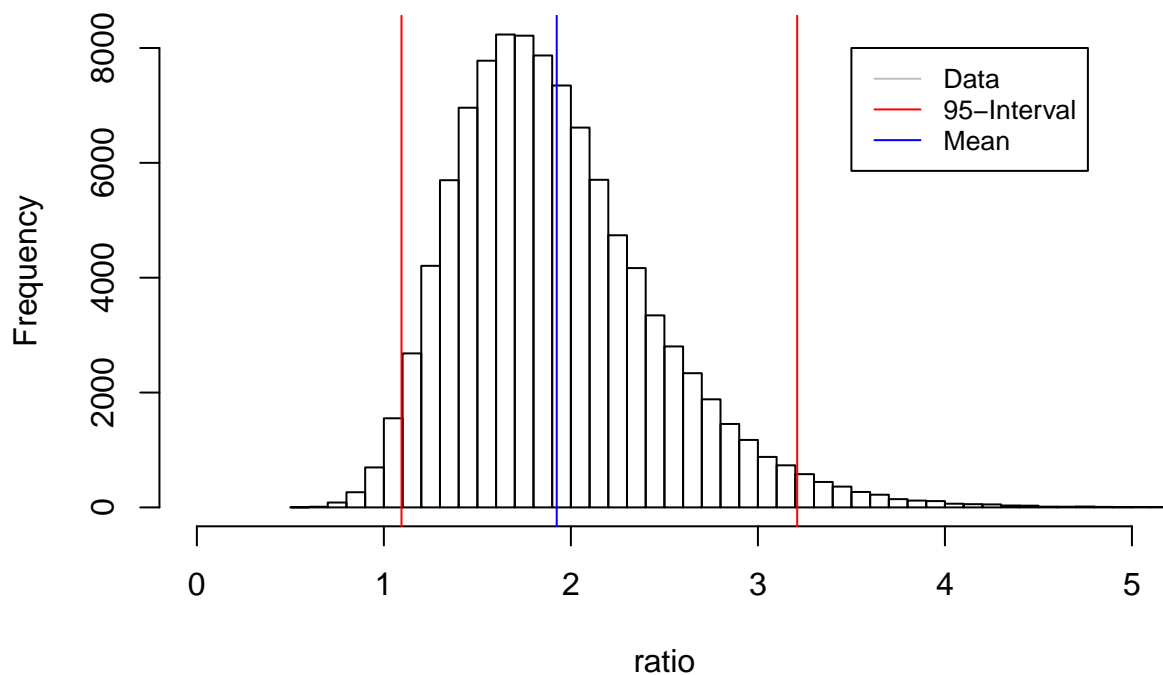
post_alpha1 = prior_alpha + n1 - y1
post_beta1 = prior_beta + y1
```

```

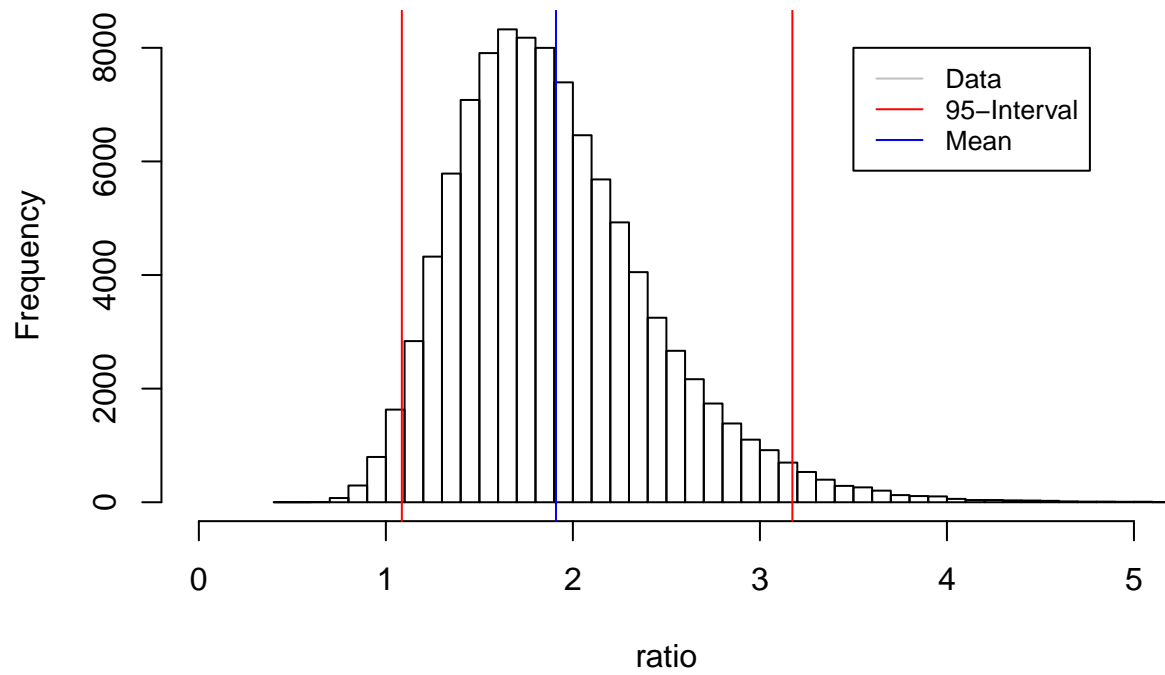
for(i in 1:length(prior_alpha)){
  p0 <- rbeta(100000, post_alpha0[i], post_beta0[i])
  p1 <- rbeta(100000, post_alpha1[i], post_beta1[i])
  point_est = posterior_odds_ratio_point_est(p0 = p0, p1 = p1)
  interval = posterior_odds_ratio_interval(p0 = p0, p1 = p1, prob = 0.95)
  ratio = (p1/(1 - p1))/(p0/(1 - p0))
  hist(ratio,
       breaks=50,
       main=paste0("Histogram using prior alpha = ", prior_alpha[i], " and beta = ", prior_beta[i]),
       xlim=c(0,5))
  abline(v = interval[1], col='red')
  abline(v = interval[2], col='red')
  abline(v = point_est, col='blue')
  legend(3.5,8000, legend=c("Data", "95-Interval", "Mean"),col=c("grey", "red", "blue"), lty=1, cex=0.8)
}

```

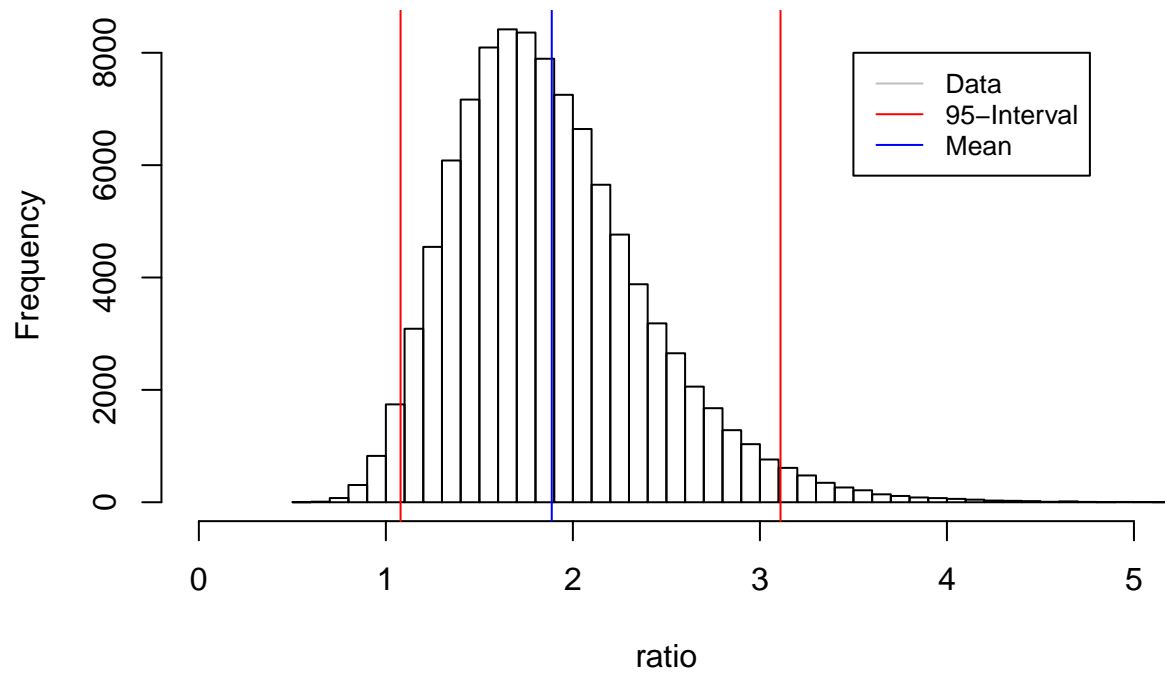
Histogram using prior alpha = 0.1 and beta = 0.1



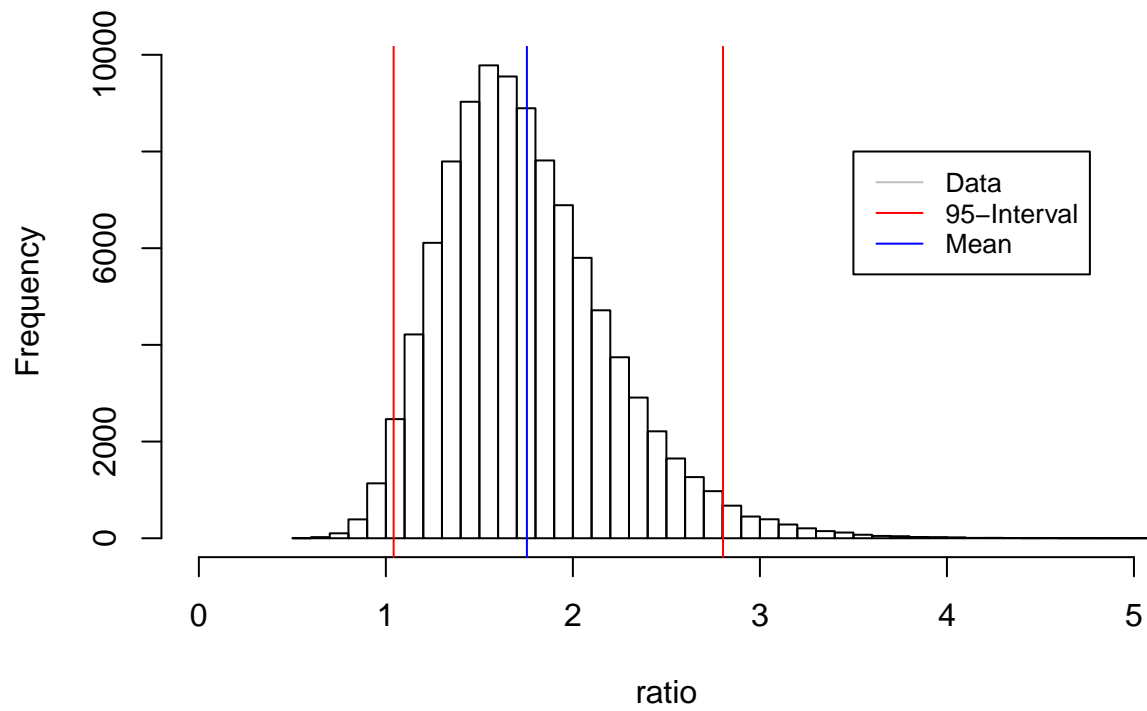
Histogram using prior $\alpha = 0.5$ and $\beta = 0.5$



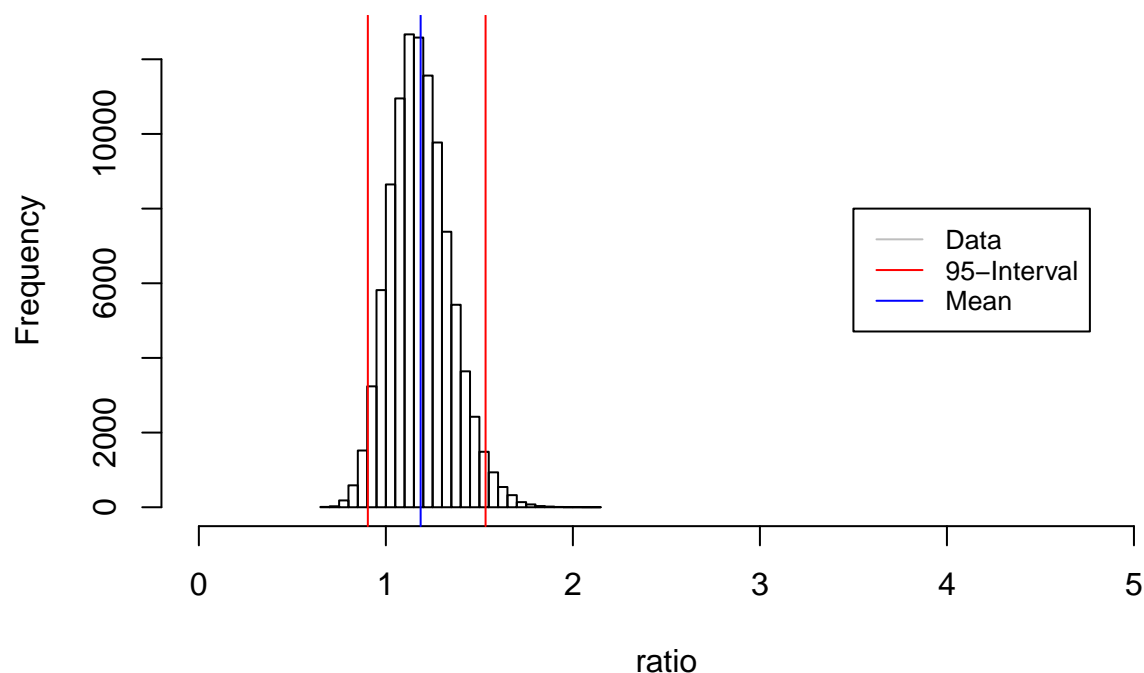
Histogram using prior $\alpha = 1$ and $\beta = 1$



Histogram using prior alpha = 5 and beta = 5



Histogram using prior alpha = 100 and beta = 100



As we can see from the plots, small values for our parameters barely change the point estimate nor the interval. Thus, the posterior is not very sensitive to changes in the prior as long as they are small. Using large numbers has a large affect on the posterior distribution, but that might not be considered noninformative or weakly informative.

3

In the exercise we will be using a uninformative prior $p(\mu, \sigma^2) \propto \frac{1}{\sigma}$. The likelihood is the normal distribution $p(y|\mu, \sigma^2) = \mathcal{N}(\mu, \sigma^2)$. We get the posterior by the product $p(\mu, \sigma^2)p(y|\mu, \sigma^2) = \frac{\mathcal{N}(\mu, \sigma^2)}{\sigma}$.

The marginalized posterior density for μ follows Student's t distribution $t_{n-1}(\bar{y}, s^2/n)$ is the same as in part 1.

```
#rm(list = ls())
data("windshieldsy1")
data("windshieldsy2")

mu_point_est = function(data){
  return(mean(data))
}

n1 = length(windshieldsy1)
df1 = n1 - 1
mu1 = mu_point_est(windshieldsy1)
scale1 = sqrt(var(windshieldsy1)/n1)
```

```

n2 = length(windshieldy2)
df2 = n2 - 1
mu2 = mu_point_est(windshieldy2)
scale2 = sqrt(var(windshieldy2)/n2)

n_samples <- 100000
samples1 <- rtnew(n_samples, df=df1, mean = mu1, scale = scale1)
samples2 <- rtnew(n_samples, df=df2, mean = mu2, scale = scale2)

mu_d <- mu1 - mu2
#mu_d = mean(samples1-samples2)
paste0("Difference in the means are: ", mu_d)

## [1] "Difference in the means are: -1.2098547008547"

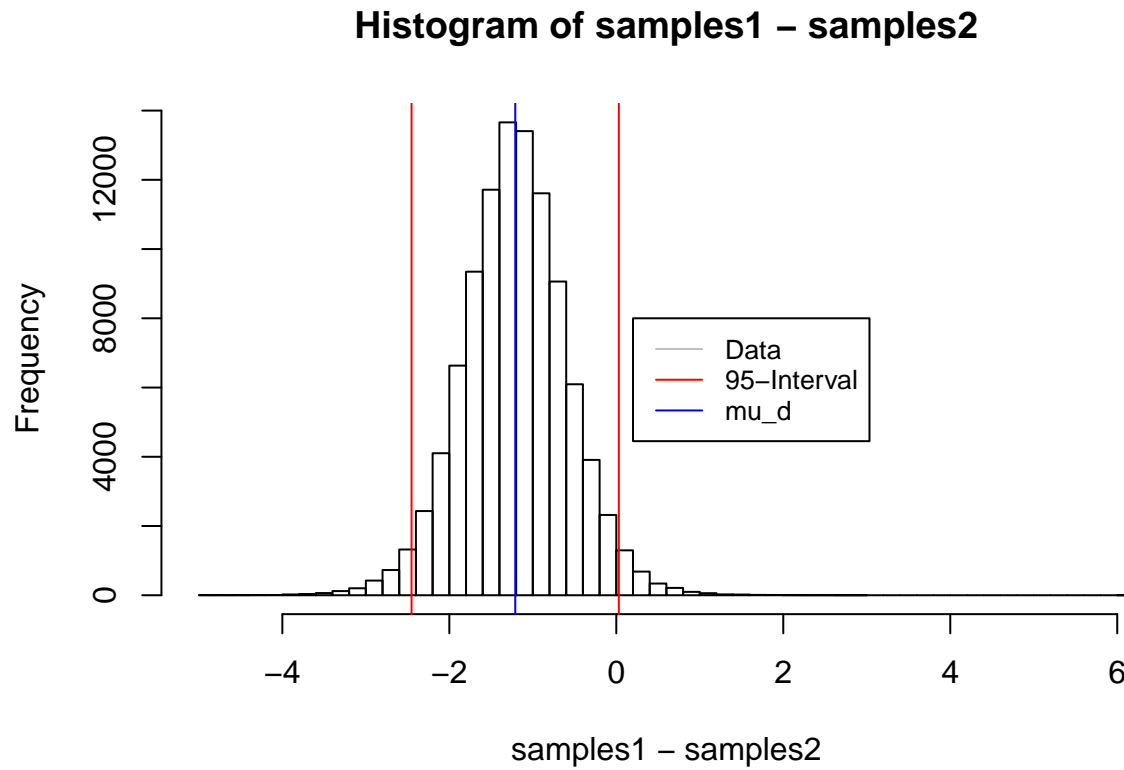
As we can see, the differences in the means are approximately -1.21.

mu_d_interval = function(samples1, samples2, prob){
  low = (1-prob)/2
  high = prob+low
  int = c(low,high)
  sample = samples1-samples2
  res = quantile(sample, int)
}
interval = mu_d_interval(samples1 = samples1, samples2 = samples2, 0.95)
paste0("Interval low: ", round(interval[1],digits=2), " and high: ", round(interval[2],digits=2))

## [1] "Interval low: -2.45 and high: 0.03"

hist(samples1-samples2,
      breaks=50)
abline(v = interval[1], col='red')
abline(v = interval[2], col='red')
abline(v = mu_d, col='blue')
legend(0.2 ,8000, legend=c("Data", "95-Interval", "mu_d"),col=c("grey", "red", "blue"), lty=1, cex=0.8)

```



b)

For any continuous variable, the probability to take on any exact value, i.e. the point density value, is 0. Hence is $P(\mu_d = 0) = 0$. Based on the credible intervals we can see that 0 is within the 95-interval, but quite close to the edge. This means that it's not very likely.