Exam #1

Elliot Smith

10/12/2018

Problem 1

Part i

Sampling Distribution: $p(y_i|\theta) = \frac{1}{\theta}\text{I}\{0 < y_i < \theta\}$

Prior Distribution: $p(\theta) = \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\text{I}\{\theta > \beta\}$

Posterior Distribution:

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

$$\propto \frac{1}{\theta^n}\text{I}\{m < \theta\} \times \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}}\text{I}\{\beta < \theta\}$$

$$\propto \alpha\beta^\alpha \times \frac{1}{\theta^{n+\alpha+1}}\text{I}\{max(\beta,m) < \theta\}$$

$$\propto \frac{1}{\theta^{n+\alpha+1}}\text{I}\{max(\beta,m) < \theta\}$$

$$\sim Pareto(n + \alpha, max(\beta,m))$$

Where we defined: $m = (y_1, ..., y_n)$.

Part ii

Yes, the posterior distribution is a conjugate prior by definition because it is in the same probability distribution family (Pareto) as the prior distribution. The prior parameters, α and β may be interpreted as follows:

- The α parameter is a shape parameter in that it affects the shape of the density in a way separate from a location shift or a scale contraction/expansion. In this distribution it defines how steady (smaller values of α) or drastic (larger values of α) the curviture of the density is as it progresses along the x-axis.
- The β parameter is a scale parameter; the larger the parameter, the more spread out the distribution. In this distribution it defines where the initial "peak" value of the distribution is before it descends.

Part iii

Prior Expected Value:

$$E[\theta] = \int_{\theta=\beta}^{\infty} \theta \times \frac{\alpha\beta^\alpha}{\theta^{\alpha+1}} d\theta$$

$$= \alpha\beta^\alpha \int_{\theta=\beta}^{\infty} \frac{1}{\theta^\alpha} d\theta$$

$$= \frac{\alpha\beta^\alpha}{(\alpha-1)\beta^{\alpha-1}} \int_{\theta=\beta}^{\infty} \frac{(\alpha-1)\beta^{\alpha-1}}{\theta^\alpha}$$

$$= \frac{\alpha\beta}{\alpha-1}$$

Posterior Expectation: $E[\theta|y] = \frac{(n+\alpha)(max(\beta,m))}{(n+\alpha)-1}$

Part iv

For my weakly informative prior I selected $Pareto(1.25, 1)$ where $\alpha = 1.25$ and $\beta = 1$. The reason for selecting these parameters is that from this website, https://www.thezebra.com/insurance-news/848/uber-vs-lyft/#info, I found that the approximate average wait time for an Uber or Lyft is 4.5 minutes. And so, I continued to tweak the parameters until I was confident that a random sample from a Pareto distribution with those parameters could reasonably expect an average wait time of 4.5 minutes.

For my noninformative prior, I selected $Pareto(0.5, 0.5)$ where $\alpha = 0.5$ and $\beta = 0.5$. I select these parameters because I believe it adds no new information and allows the "data to speak for itself".

Weakly Informative Prior

Histogram – Minimum Value   Histogram – Median Value   Histogram – Maximum Value
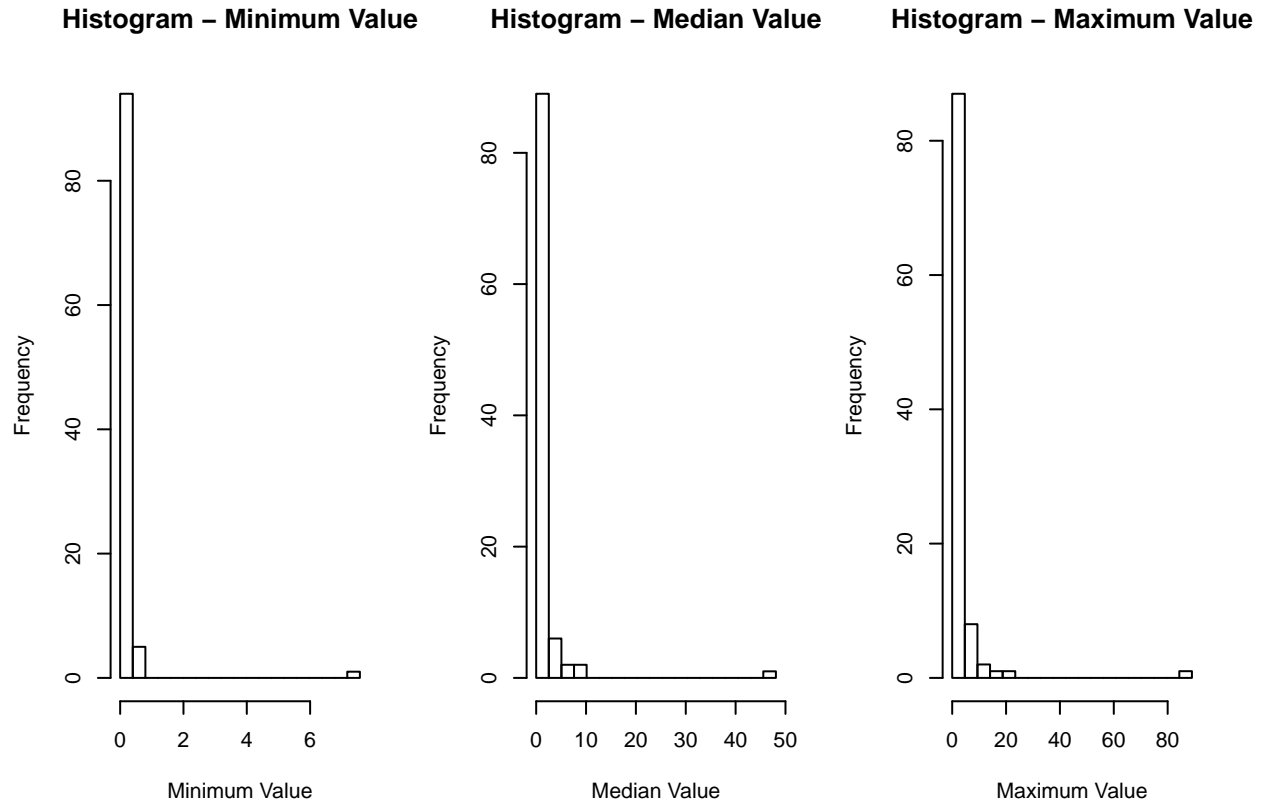
We get what appears to be reasonable results from our weakly informative prior. Our minimum value appears to hover around 0 to 1, our median value hovers around 0 to 10, while our maximum value hovers around 0 to 20. As we can see, a majority of the realized values hover around 0; this is a symptom of trying to make our prior as weakly informative as possible while simultaneously putting pressure on our distribution to pull its values closer to the heavier density areas closer to 0.
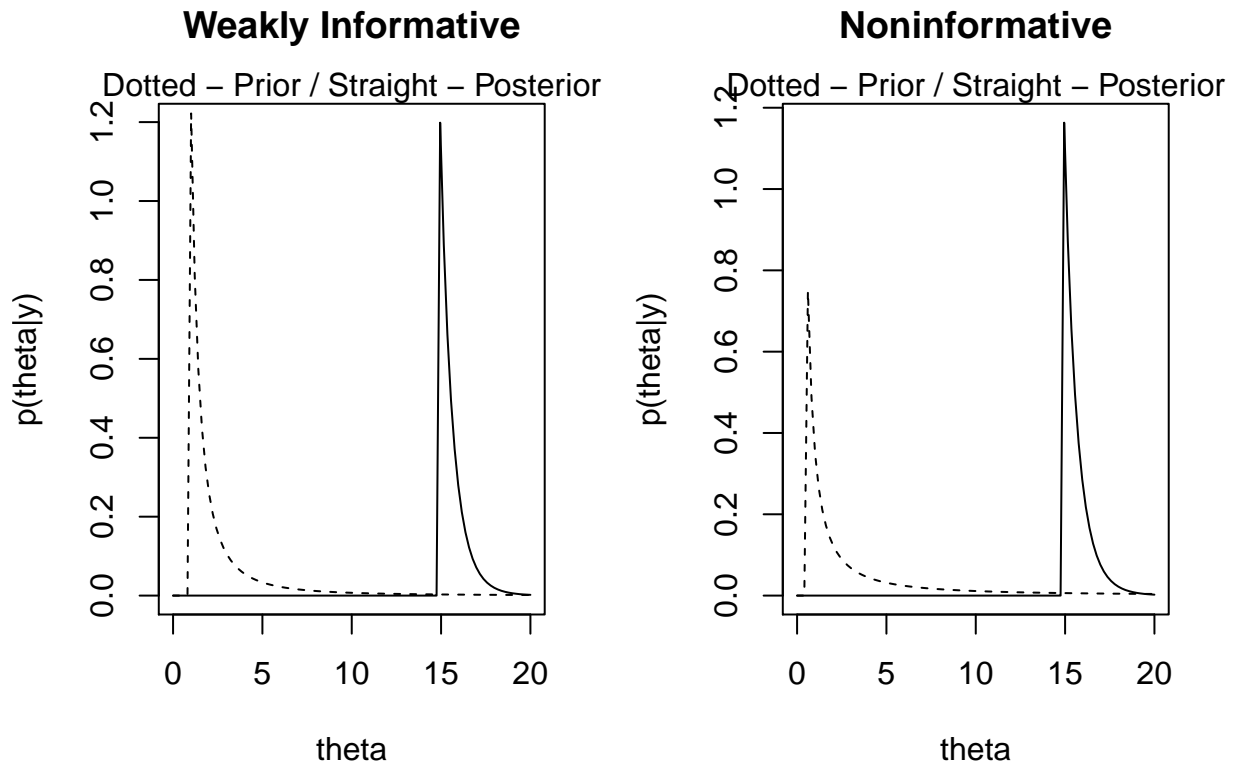
Noninformative Prior

| Histogram – Minimum Value | Histogram – Median Value | Histogram – Maximum Value |
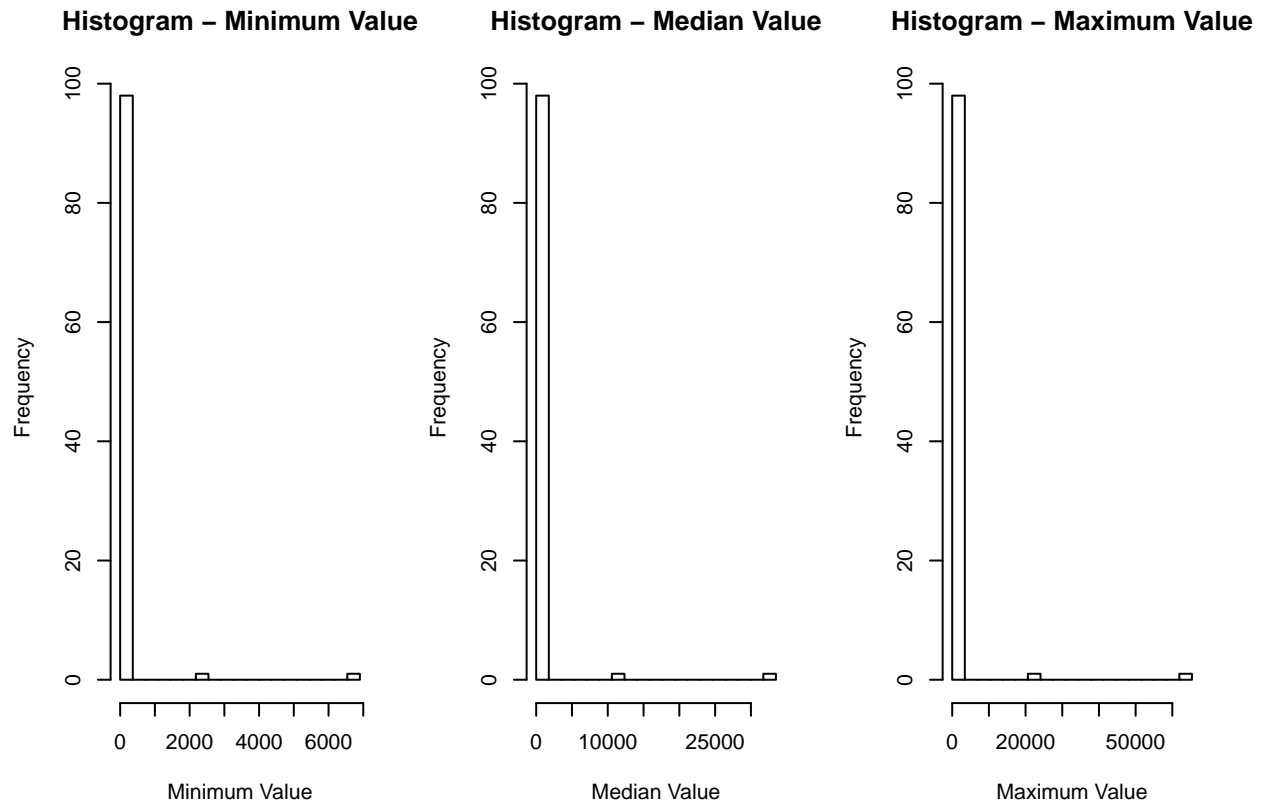
The results from our noninformative prior do not appear to be very reasonable at all. A symptom of allowing the "data to speak for itself" is that we are subject to the construct of our distribution. In this case, we can see for our minimum, median and maximum, the densities are correctly strongest around 0. However, due to the noninformative nature of our prior distribution, we are seeing more extreme values that we would not be very likely to see (given anecdotal understanding of wait times for a Lyft/Uber).

Summary

Of the two Pareto prior distributions selected, I would certainly go with the weakly informative prior as the calculated statistics seem much more reasonable. A weakness of our noninformative approach is that more likely values are not being given a "heavier weight", and as such, we see more extreme values that make our model suspect. A practical next step in my mind would be to try a prior that is more strongly informative; this would give us the ability to pull the density away from 0, as well as away from the extreme values as we currently did with our weakly informative prior.

Part v

Weakly Informative

Dotted – Prior / Straight – Posterior

Noninformative

Dotted – Prior / Straight – Posterior

```
## Posterior Credible Interval, Weakly Informative:  ( 14.83009 , 17.12185 )
```

```
## Posterior Credible Interval, Noninformative:  ( 14.83 , 17.23462 )
```

Both the weakly informative prior and noninformative prior appear to have almost no affect on the posterior distribution. As we can see from the density plots for each, the posterior distributions have no discernable differences across the two prior distributions. This fact is confirmed by the posterior credible intervals which are again almost completely identical! This leads me to believe that both of our priors have almost no affect on the posterior at all.

Part vi

Weakly Informative Prior
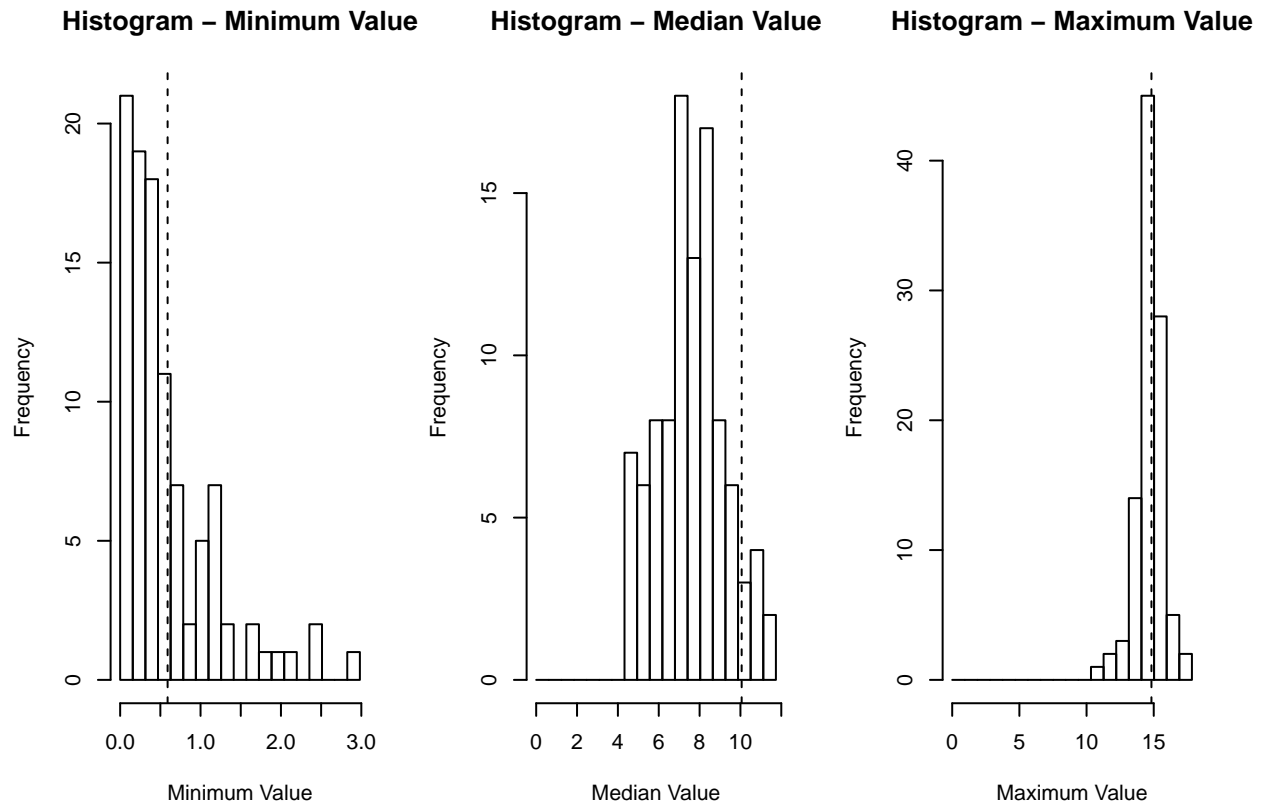
Our results look quite promising for the weakly informative prior. The distributions that represent our minimum value and our max value seem to do a very good job of capturing our observed estimates of these values. Our median value however is not great; our observed value is not well centered on our posterior distribution. In general, I believe this model is a decently good fit, and while there are areas to improve it (perhaps a more informative prior), I am very happy with the result.

Noninformative Prior
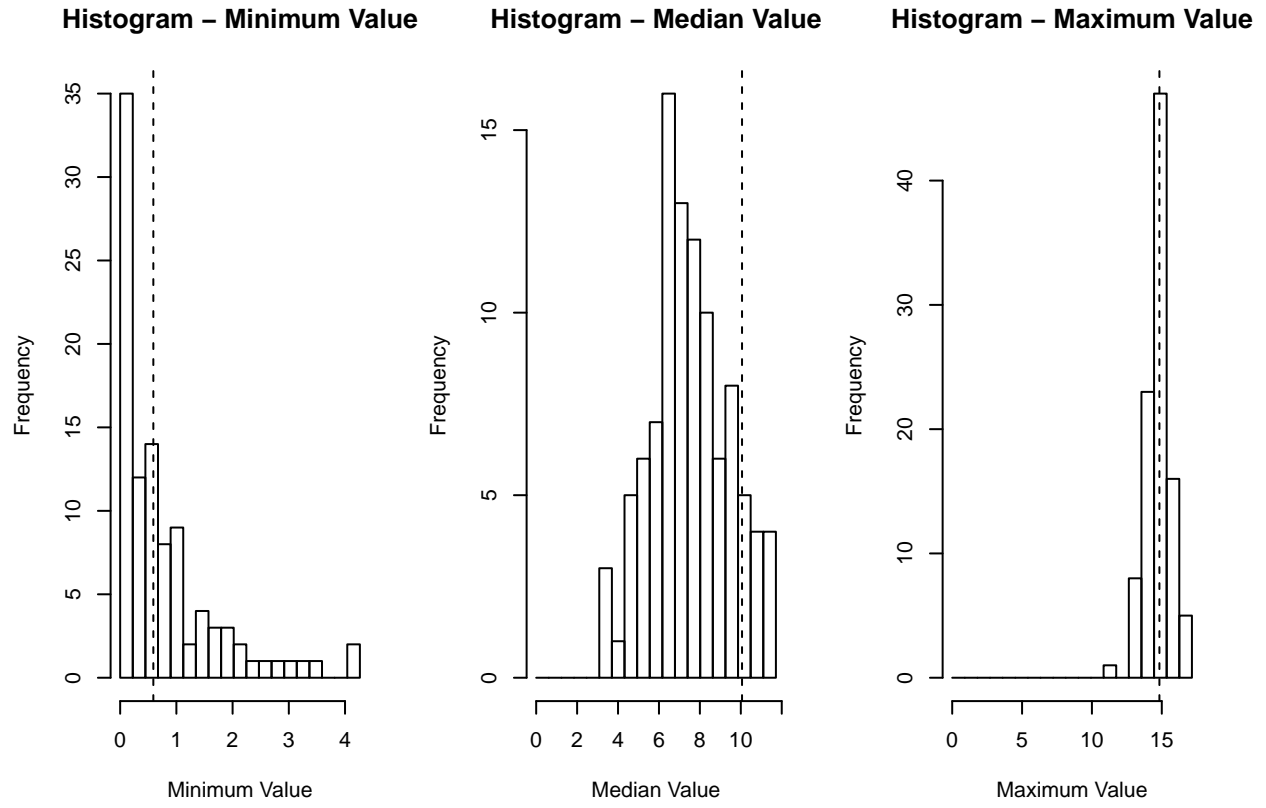
Our results for the noninformative prior are very similar to our results for the weakly informative prior: while our posterior predictive distributions represent our minimum and maximum values quite well, our median value is not well represented. As with the weakly informative prior, I feel the result is satisfactory, while again noting that I would perhaps like to explore a more informative prior and see how well that model may fit.

Problem 2

Part i

I propose the following prior distributions:

- $p(\mu_1) \sim N(0,1)$
- $p(\mu_2) \sim N(0,1)$
- $\tau_1 = \frac{1}{\sigma_1^2} \sim Gamma(2,1)$
- $\tau_2 = \frac{1}{\sigma_2^2} \sim Gamma(2,1)$
- $k \sim Uniform(1,2,...,n)$

Part ii

$p(\mu_1|y, \mu_2, \tau_1, \tau_2, k) \propto p(y|\mu_1, \mu_2, \tau_1, \tau_2, k) \times p(\mu_2|y, \mu_1, \tau_1, \tau_2, k) \times p(\tau_1|y, \mu_1, \mu_2, \tau_2, k) \times p(\tau_2|y, \mu_1, \mu_2, \tau_1, k) \times p(k|y, \mu_1, \mu_2, \tau_1, \tau_2) \times p(\mu_1)$

Posterior for $p(\mu_1|y, \mu_2, \tau_1, \tau_2, k)$:

$$p(\mu_1|y, \mu_2, \tau_1, \tau_2, k) \propto exp\left[-\tau_1 \sum_{t=1}^{k} \frac{(y_i - \mu_1)^2}{2} \right] \times exp\left[\frac{-\mu_1^2}{2} \right]$$

$$\propto exp\left[-\frac{1 + k\tau_1}{2}\left(\mu_1 - \frac{\tau_1 \sum_{t=1}^{k} y_i}{1 + k\tau_1} \right)^2 \right]$$

$$\sim N\left(\frac{\tau_1 \sum_{t=1}^{k} y_i}{1 + k\tau_1}, \frac{1}{1 + k\tau_1} \right)$$

$p(\mu_2|y, \mu_1, \tau_1, \tau_2, k) \propto p(y|\mu_1, \mu_2, \tau_1, \tau_2, k) \times p(\mu_1|y, \mu_2, \tau_1, \tau_2, k) \times p(\tau_1|y, \mu_1, \mu_2, \tau_2, k) \times p(\tau_2|y, \mu_1, \mu_2, \tau_1, k) \times p(k|y, \mu_1, \mu_2, \tau_1, \tau_2) \times p(\mu_2)$

Posterior for $p(\mu_2|y, \mu_1, \tau_1, \tau_2, k)$:

$$p(\mu_2|y, \mu_1, \tau_1, \tau_2, k) \propto exp\left[-\tau_2 \sum_{t=k+1}^{n} \frac{(y_i - \mu_2)^2}{2} \right] \times exp\left[\frac{-\mu_2^2}{2} \right]$$

$$\propto exp\left[-\frac{1 + (n-k)\tau_2}{2}\left(\mu_2 - \frac{\tau_2 \sum_{t=k+1}^{n} y_i}{1 + (n-k)\tau_2} \right)^2 \right]$$

$$\sim N\left(\frac{\tau_2 \sum_{t=k+1}^{n} y_i}{1 + (n-k)\tau_2}, \frac{1}{1 + (n-k)\tau_2} \right)$$

$p(\tau_1|y, \mu_1, \mu_2, \tau_2, k) \propto p(y|\mu_1, \mu_2, \tau_1, \tau_2, k) \times p(\mu_1|y, \mu_2, \tau_1, \tau_2, k) \times p(\mu_2|y, \mu_1, \tau_1, \tau_2, k) \times p(\tau_2|y, \mu_1, \mu_2, \tau_1, k) \times p(k|y, \mu_1, \mu_2, \tau_1, \tau_2) \times p(\tau_1)$

Posterior for $p(\tau_1|y, \mu_1, \mu_2, \tau_2, k)$:

$$p(\tau_1|y, \mu_1, \mu_2, \tau_2, k) \propto \tau_1^{\frac{n}{2}} exp\left[-\tau_1 \sum_{t=1}^{k} \frac{(y_i - \mu_1)^2}{2} \right] \tau_1 exp[-\tau_1]$$

$$\propto \tau_1^{1 + \frac{n}{2}} exp\left[-\tau_1\left(1 + \frac{1}{2}\sum_{t=1}^{k}(y_i - \mu_1)^2 \right) \right] exp\left[-\frac{1}{\tau_1} \right]$$

$$\sim Gamma\left(2 + \frac{k}{2}, 1 + \frac{1}{2}\sum_{t=1}^{k}(y_i - \mu_1)^2 \right)$$

$p(\tau_2|y, \mu_1, \mu_2, \tau_2, k) \propto p(y|\mu_1, \mu_2, \tau_1, \tau_2, k) \times p(\mu_1|y, \mu_2, \tau_1, \tau_2, k) \times p(\mu_2|y, \mu_1, \tau_1, \tau_2, k) \times p(\tau_1|y, \mu_1, \mu_2, \tau_2, k) \times p(k|y, \mu_1, \mu_2, \tau_1, \tau_2) \times p(\tau_2)$

Posterior for $p(\tau_2|y, \mu_1, \mu_2, \tau_1, k)$:

$$p(\tau_2|y,\mu_1,\mu_2,\tau_1,k) \propto \tau_2^{\frac{n}{2}} exp\left[-\tau_2\sum_{t=k+1}^{n}\frac{(y_i-\mu_2)^2}{2}\right]\tau_2 exp[-\tau_2]$$

$$\propto \tau_2^{1+\frac{n}{2}} exp\left[-\tau_2\left(1+\frac{1}{2}\sum_{t=k+1}^{n}(y_i-\mu_2)^2\right)\right]exp\left[-\frac{1}{\tau_2}\right]$$

$$\sim Gamma\left(2+\frac{n-k}{2},1+\frac{1}{2}\sum_{t=k+1}^{n}(y_i-\mu_2)^2\right)$$

$p(k|y,\mu_1,\mu_2,\tau_1,\tau_2) \propto p(y|\mu_1,\mu_2,\tau_1,\tau_2,k) \times p(\mu_1|y,\mu_2,\tau_1,\tau_2,k) \times p(\mu_2|y,\mu_1,\tau_1,\tau_2,k) \times p(\tau_1|y,\mu_1,\mu_2,\tau_2,k) \times p(\tau_2|y,\mu_1,\mu_2,\tau_1,k) \times p(k)$

Posterior for $p(k=j|y,\mu_1,\mu_2,\tau_1,\tau_2)$:

$$p(k=j|y,\mu_1,\mu_2,\tau_1,\tau_2) = \frac{p(y|\mu_1,\mu_2,\tau_1,\tau_2,k=j)}{\sum_{i=1}^{n}p(y|\mu_1,\mu_2,\tau_1,\tau_2,k=i)}$$

$$= p(y|\mu_1,\mu_2,\tau_1,\tau_2,k=j)$$

$$= \prod_{t=1}^{j}\frac{\sqrt{\tau_1}}{\sqrt{2\pi}}exp\left(-\frac{\tau_1}{2}(y_t-\mu_1)^2\right) \times \prod_{t=j+1}^{n}\frac{\sqrt{\tau_2}}{\sqrt{2\pi}}exp\left(-\frac{\tau_2}{2}(y_t-\mu_2)^2\right)$$

$$\propto \left[\prod_{t=1}^{j}\sqrt{\tau_1}exp\left(-\frac{\tau_1}{2}(y_t-\mu_1)^2\right)\right]\left[\prod_{t=j+1}^{n}\sqrt{\tau_2}exp\left(-\frac{\tau_2}{2}(y_t-\mu_2)^2\right)\right]$$

$$\propto \left[\prod_{t=1}^{j}\sqrt{\tau_1}exp\left(-\frac{\tau_1}{2}(y_t-\mu_1)^2\right)\right]\left[\prod_{t=1}^{n}\sqrt{\tau_2}exp\left(-\frac{\tau_2}{2}(y_t-\mu_2)^2\right)\prod_{t=1}^{j}\tau_2^{-\frac{1}{2}}exp\left(\frac{\tau_2}{2}(y_t-\mu_2)^2\right)\right]$$

$$\propto \left[\prod_{t=1}^{n}\sqrt{\tau_2}exp\left(-\frac{\tau_2}{2}(y_t-\mu_2)^2\right)\right]\left[\prod_{t=1}^{j}\frac{\sqrt{\tau_1}}{\sqrt{\tau_2}}exp\left(\frac{\tau_2}{2}(y_t-\mu_2)^2-\frac{\tau_1}{2}(y_t-\mu_1)^2\right)\right]$$

Where:

$f(y,\mu_1,\mu_2,\tau_1,\tau_2) = \left[\prod_{t=1}^{n}\sqrt{\tau_2}exp\left(-\frac{\tau_2}{2}(y_t-\mu_2)^2\right)\right]$

$g(j,y,\mu_1,\mu_2,\tau_1,\tau_2) = \left[\prod_{t=1}^{j}\frac{\sqrt{\tau_1}}{\sqrt{\tau_2}}exp\left(\frac{\tau_2}{2}(y_t-\mu_2)^2-\frac{\tau_1}{2}(y_t-\mu_1)^2\right)\right]$

Part iii

Below is the function in R to compute $p(y|\mu_1,\mu_2,\tau_1,\tau_2,k=j)$:

```
p_full_cond <- function(tau_1, tau_2, mu_1, mu_2, j) {

    sqrt(tau_2) * exp(-(tau_2 / 2) * cumsum((y - mu_2)^2)) *
        (sqrt(tau_1 / tau_2)) * exp((tau_2 / 2) *
            cumsum((y - mu_2)^2) - (tau_1 / 2) * cumsum((y - mu_1)^2))

}
```
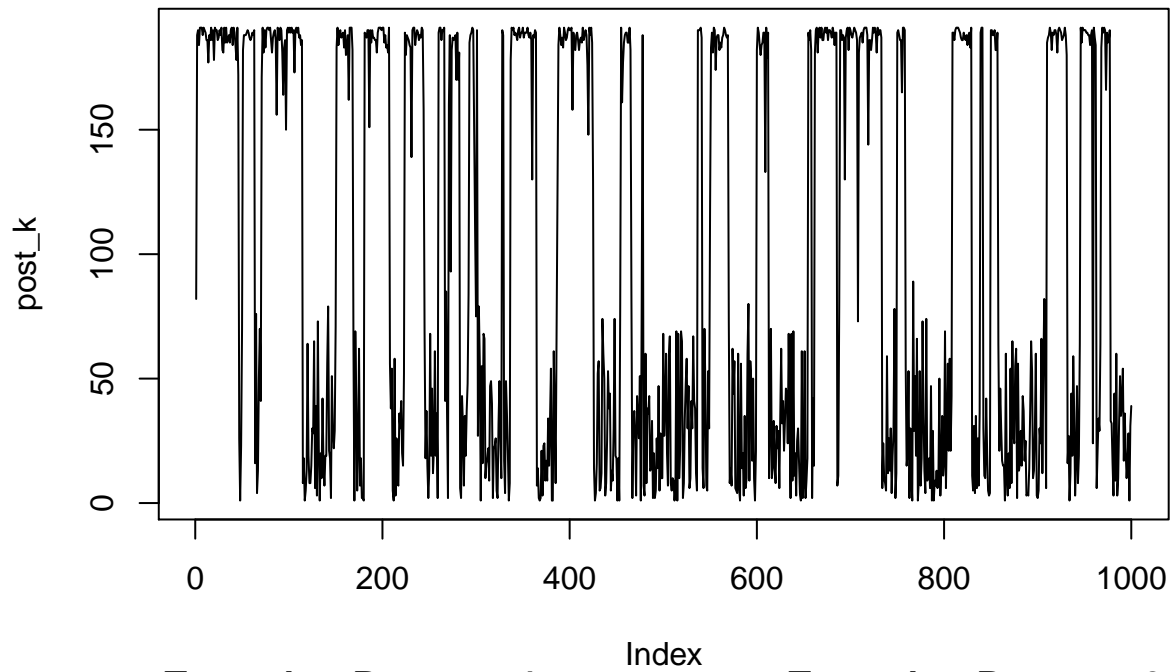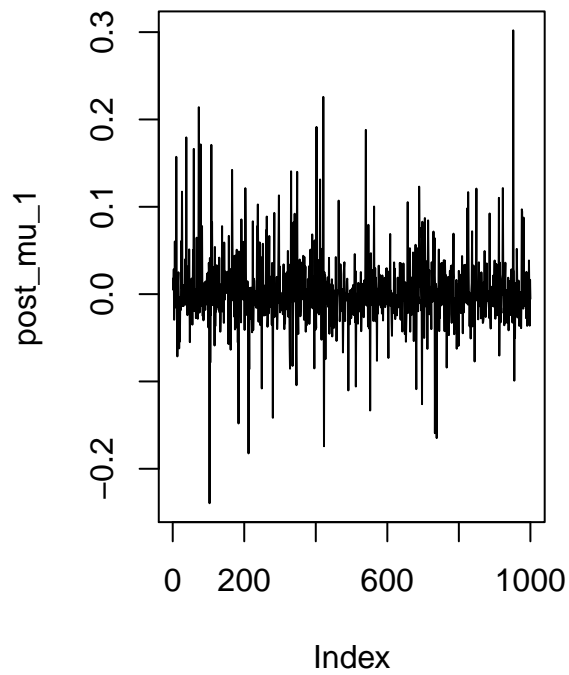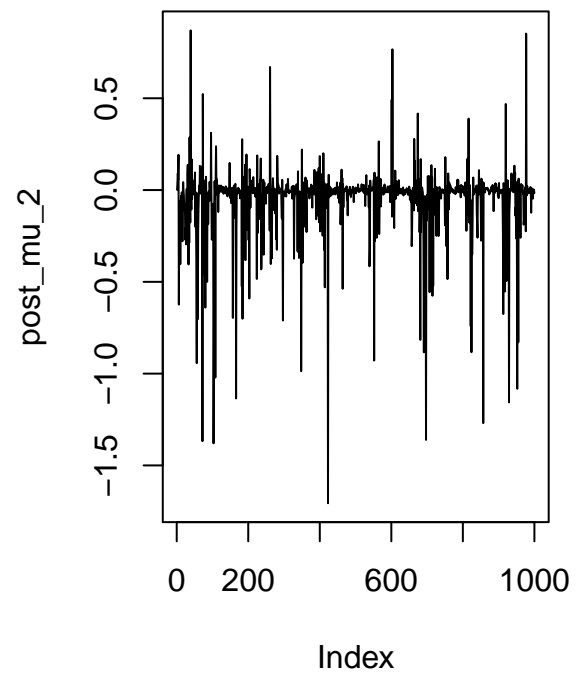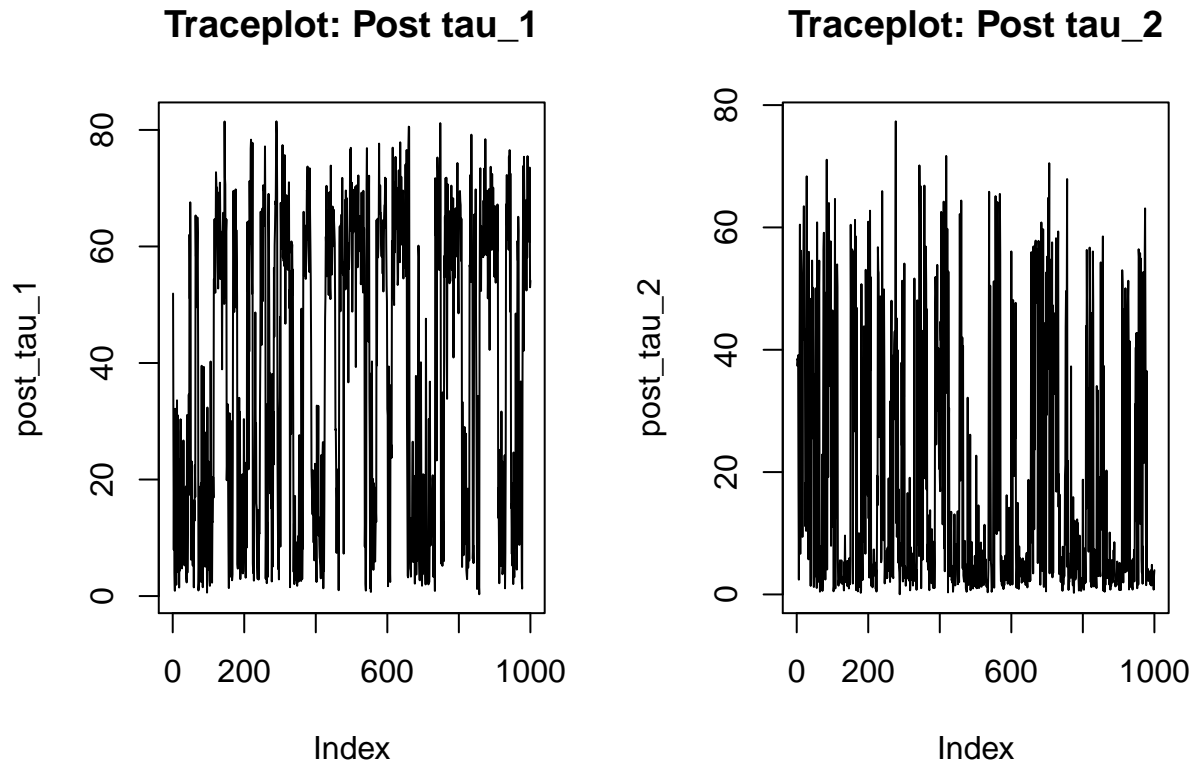
Part iv

Traceplot: Post k

Traceplot: Post mu_1

Traceplot: Post mu_2

Traceplot: Post tau_1

Traceplot: Post tau_2


I think it is very clear that, based on the traceplots, that the chains have not converged.

First looking at the traceplot for post_k, the chain appears to mostly oscillate between extreme values of k and, for small periods of time, lower values of k (around 0 to 50); we clearly have not converged to a value that would identify our change point.

Looking at our traceplots for μ_1 and μ_2 we see that our traceplots tend to hover around 0, which seems to represent the data well enough, but as we can see, we also observe extreme values later on in the chain, which indicates that we have not yet converged.

Finally, the traceplots for τ_1 and τ_2 also indicate a lack of convergence. We again see very large osciallations back and forth between extreme values and are unable to draw any substantive conclusions from our result.

Part v

BAC Return
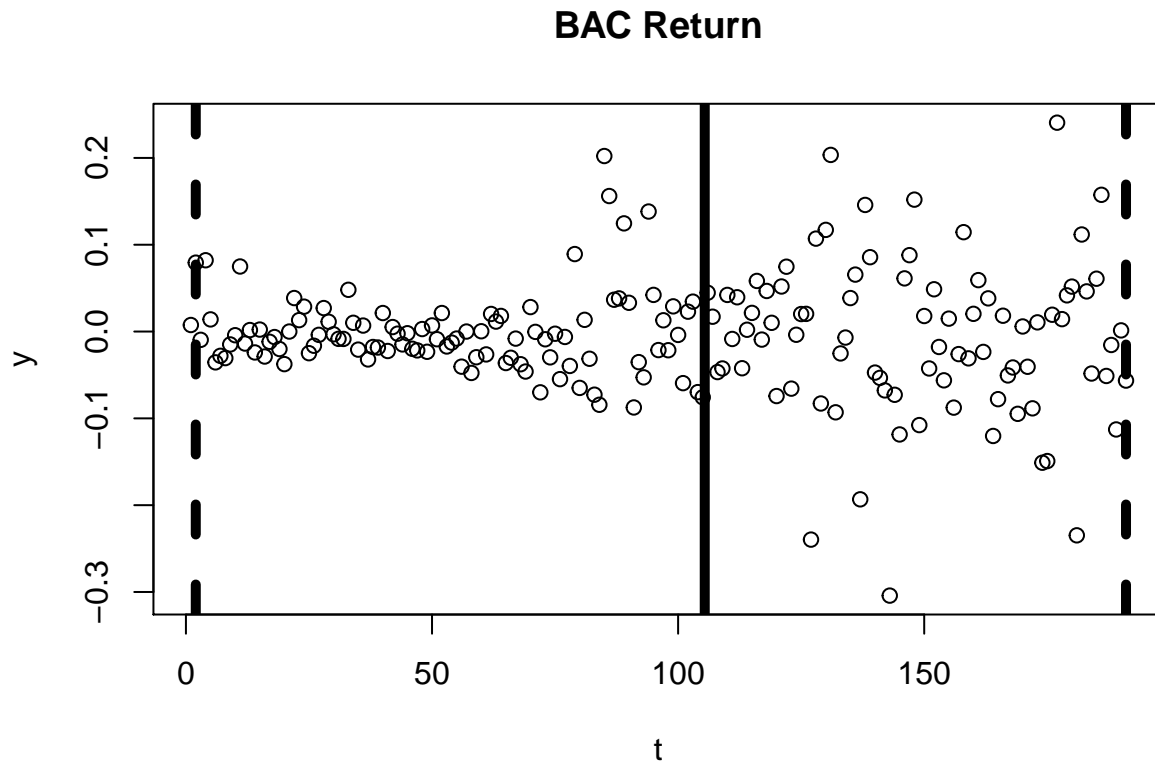
The change point does appear to be reasonable; the posterior credible interval however does not appear very reasonable.

By the "eye test" we can see a clear change in variance at the change point; aside from the cloud of 5 outlying points at t approximately equal to 80 to 95, the variance appears very constant prior to the change point. After the change point there is a noticeable change in variance. In fact I feel strongly, again using the "eye test" as confirmation, that the variance truly begins to grow right at the moment of the change point; this is a great result.

The credible interval does make the result not as nice. Essentially, the posterior credible interval is saying that any t might be the change point; this is clearly not a great result as it increases the uncertainty associated with our change point.

Part vi

```
## Posterior Credible Interval, mu_2 - mu_1:  ( -0.5683991 , 0.2540994 )
```

```
## Posterior Credible Interval, tau_2 - tau_1:  ( -74.78819 , 47.14178 )
```

Both differences, $\tau_2 - \tau_1$ and $\mu_2 - \mu_1$ contain 0, this implies that there is overlap, that the means and variances compared across the change point may be the same. For the mean, this is a positive result; looking at our data it seems like a safe bet to conclude that the mean is a constant 0. For the variance however, this is not a great result. We can clearly see a difference in variance before and after the change point, so the fact that the credible intervals contains 0 is a troubling result for us.

To summarize:

- There is NOT strong evidence for a nonconstant mean
- There is NOT strong evidence for a nonzero mean
- There is NOT strong evidence for a nonconstant variance according to our analysis, but again, according to the plot of the data which appears to show nonconstant variance, this may prove issues with the analysis

Code Appendix

```r
########## Workspace Preparation ##########



## Load in the necessary packages

suppressMessages(
    suppressWarnings(
        library(EnvStats)
    )
)
library(coda)



########## Problem 1 ##########



##### Part i #####



##### Part ii #####



##### Part iii #####


```

```r
##### Part iv #####


## Input the known paramters

S <- 100
n <- 20

shape_weak <- 1.25
location_weak <- 1

shape_non <- 0.5
location_non <- 0.5



### Weakly Informative Prior ###



## Sample from the weakly informative prior distribution

theta_weak_samp <- rpareto(n = S, location = location_weak, shape = shape_weak)

## Construct the container for the replicate data sets for the weakly informative prior

y_rep_weak <- matrix(0, nrow = S, ncol = n)

## Loop through and sample from the weakly informative prior

for(s in 1:S) {

    y_rep_weak[s,] <- runif(n = n, min = 0, max = theta_weak_samp[s])

}

## Compute the required test statistics for each replicate data sets for the weakly informative prior

y_rep_weak_min <- apply(y_rep_weak, 1, min)
y_rep_weak_med <- apply(y_rep_weak, 1, median)
y_rep_weak_max <- apply(y_rep_weak, 1, max)

## Plot the histograms of the statistics

par(mfrow = c(1, 3))

hist(y_rep_weak_min, breaks = seq(0, max(y_rep_weak_min), l = 20),
     main = "Histogram - Minimum Value", xlab = "Minimum Value")
hist(y_rep_weak_med, breaks = seq(0, max(y_rep_weak_med), l = 20),
     main = "Histogram - Median Value", xlab = "Median Value")
hist(y_rep_weak_max, breaks = seq(0, max(y_rep_weak_max), l = 20),
     main = "Histogram - Maximum Value", xlab = "Maximum Value")
```

```r
### Noninformative Prior ###


## Sample from the noninformative prior distribution

theta_non_samp <- rpareto(n = S, location = location_non, shape = shape_non)

## Construct the container for the replicate data sets for the noninformative prior

y_rep_non <- matrix(0, nrow = S, ncol = n)

## Loop through and sample from the noninformative prior

for(s in 1:S) {

    y_rep_non[s,] <- runif(n = n, min = 0, max = theta_non_samp[s])

}

## Compute the required test statistics for each replicate data sets for the weakly informative prior

y_rep_non_min <- apply(y_rep_non, 1, min)
y_rep_non_med <- apply(y_rep_non, 1, median)
y_rep_non_max <- apply(y_rep_non, 1, max)

## Plot the histograms of the statistics

par(mfrow = c(1, 3))

hist(y_rep_non_min, breaks = seq(0, max(y_rep_non_min), l = 20),
     main = "Histogram - Minimum Value", xlab = "Minimum Value")
hist(y_rep_non_med, breaks = seq(0, max(y_rep_non_med), l = 20),
     main = "Histogram - Median Value", xlab = "Median Value")
hist(y_rep_non_max, breaks = seq(0, max(y_rep_non_max), l = 20),
     main = "Histogram - Maximum Value", xlab = "Maximum Value")



##### Part v #####


## Load in the data

y = c(2.21, 14.49, 2.42, 12.37, 11.54, 9.56, 0.59, 5.43, 5.07, 8.77, 13.84,
      0.86, 5.97, 3.38, 14.83, 11.87, 10.56, 12.54, 10.97, 13.28)

## Set our theta

theta <- seq(from = 0, to = 20, length.out = 100)
```

```r
## Set our plots side-by-side

par(mfrow = c(1, 2))

## Generate the plots for the weakly informative prior

plot(x = theta,
     y = dpareto(x = theta, location = max(location_weak, y), shape = shape_weak + n),
     type = "l", ylab = "p(theta|y)", main = "Weakly Informative")
mtext("Dotted - Prior / Straight - Posterior")

lines(x = theta,
      y = dpareto(x = theta, location = location_weak, shape = shape_weak),
      type = "l", lty = 2)

## Generate the plots for the noninformative prior

plot(x = theta,
     y = dpareto(x = theta, location = max(location_non, y), shape = shape_non + n),
     type = "l", ylab = "p(theta|y)", main = "Noninformative")
mtext("Dotted - Prior / Straight - Posterior")

lines(x = theta,
      y = dpareto(x = theta, location = location_non, shape = shape_non),
      type = "l", lty = 2)

## Draw a sample from each posterior distribution

post_weak <- rpareto(n = 10000, location = max(location_weak, y), shape = shape_weak + n)
post_non <- rpareto(n = 10000, location = max(location_non, y), shape = shape_non + n)

## Compute the HPD intervals

hpd_int_weak <- as.vector(HPDinterval(as.mcmc(post_weak), prob = 0.95)[1, 1:2])
hpd_int_non <- as.vector(HPDinterval(as.mcmc(post_non), prob = 0.95)[1, 1:2])

## Output the results

cat("Posterior Credible Interval, Weakly Informative: ", "(", hpd_int_weak[1], ",", hpd_int_weak[2], ")"
cat("Posterior Credible Interval, Noninformative: ", "(", hpd_int_non[1], ",", hpd_int_non[2], ")")



##### Part vi #####



## Compute the real data statistics

y_min <- min(y)
y_med <- median(y)
y_max <- max(y)
```

```r
### Weakly Informative Prior ###


## Sample from the weakly informative posterior distribution

theta_weak_samp_post <- rpareto(n = S, location = max(location_weak, y), shape = shape_weak + n)

## Construct the container for the replicate data sets for the weakly informative posterior

y_rep_weak_post <- matrix(0, nrow = S, ncol = n)

## Loop through and sample from the weakly informative posterior

for(s in 1:S) {

    y_rep_weak_post[s,] <- runif(n = n, min = 0, max = theta_weak_samp_post[s])

}

## Compute the required test statistics for each replicate data sets for the weakly informative posteri

y_rep_weak_post_min <- apply(y_rep_weak_post, 1, min)
y_rep_weak_post_med <- apply(y_rep_weak_post, 1, median)
y_rep_weak_post_max <- apply(y_rep_weak_post, 1, max)

## Plot the histograms of the statistics

par(mfrow = c(1, 3))

hist(y_rep_weak_post_min, breaks = seq(0, max(y_rep_weak_post_min), l = 20),
     main = "Histogram - Minimum Value", xlab = "Minimum Value")
abline(v = y_min, lty = 2)
hist(y_rep_weak_post_med, breaks = seq(0, max(y_rep_weak_post_med), l = 20),
     main = "Histogram - Median Value", xlab = "Median Value")
abline(v = y_med, lty = 2)
hist(y_rep_weak_post_max, breaks = seq(0, max(y_rep_weak_post_max), l = 20),
     main = "Histogram - Maximum Value", xlab = "Maximum Value")
abline(v = y_max, lty = 2)



### Noninformative Prior ###


## Sample from the noninformative posterior distribution

theta_non_samp_post <- rpareto(n = S, location = max(location_non, y), shape = shape_non + n)

## Construct the container for the replicate data sets for the noninformative posterior
```

```r
y_rep_non_post <- matrix(0, nrow = S, ncol = n)

## Loop through and sample from the noninformative posterior

for(s in 1:S) {

    y_rep_non_post[s,] <- runif(n = n, min = 0, max = theta_non_samp_post[s])

}

## Compute the required test statistics for each replicate data sets for the noninformative posterior

y_rep_non_post_min <- apply(y_rep_non_post, 1, min)
y_rep_non_post_med <- apply(y_rep_non_post, 1, median)
y_rep_non_post_max <- apply(y_rep_non_post, 1, max)

## Plot the histograms of the statistics

par(mfrow = c(1, 3))

hist(y_rep_non_post_min, breaks = seq(0, max(y_rep_non_post_min), l = 20),
     main = "Histogram - Minimum Value", xlab = "Minimum Value")
abline(v = y_min, lty = 2)
hist(y_rep_non_post_med, breaks = seq(0, max(y_rep_non_post_med), l = 20),
     main = "Histogram - Median Value", xlab = "Median Value")
abline(v = y_med, lty = 2)
hist(y_rep_non_post_max, breaks = seq(0, max(y_rep_non_post_max), l = 20),
     main = "Histogram - Maximum Value", xlab = "Maximum Value")
abline(v = y_max, lty = 2)




########## Problem 2 ##########



## Load in the data

dat <- read.csv('~/Documents/Rice_University/Fall_2018/STAT525/Exam_1/bac.csv')

## Assign variables

y <- dat$y
x <- dat$X
dates <- as.Date(dat$dates)
n <- length(y)



##### Part i #####
```

```r
##### Part ii #####



##### Part iii #####



## Input the function to compute the y full conditional on k = j

p_full_cond <- function(tau_1, tau_2, mu_1, mu_2, j) {

    sqrt(tau_2) * exp(-(tau_2 / 2) * cumsum((y - mu_2)^2)) *
        (sqrt(tau_1 / tau_2)) *
        exp((tau_2 / 2) * cumsum((y - mu_2)^2) - (tau_1 / 2) * cumsum((y - mu_1)^2))

}



##### Part iv #####



## Number of iterations

S <- 10000

## Select some initial values

k <- ceiling(n / 2)

mu_1 <- mean(y[1:k])
mu_2 <- mean(y[(k + 1):n])
tau_1 <- sqrt(1 / var(y[1:k]))
tau_2 <- sqrt(1 / var(y[(k + 1):n]))

## Set containers for the variables

post_k <- array(0, c(S, 1))
post_mu_1 <- array(0, c(S, 1))
post_mu_2 <- array(0, c(S, 1))
post_tau_1 <- array(0, c(S, 1))
post_tau_2 <- array(0, c(S, 1))

## Loop through to perform the Gibbs sampling

for (s in 1:S) {

    ## Included this to deal with the case where k = n

    if (k == n) {k = (n - 1)}
```

```r
    ## Sample from mu_1

    mu_1 <- rnorm(n = 1,
                  mean = (tau_1 * sum(y[1:k])) / (1 + (k * tau_1)),
                  sd = sqrt(1 / (1 + (k * tau_1)))
                  )

    ## Sample from mu_2

    mu_2 <- rnorm(n = 1,
                  mean = (tau_2 * sum(y[(k + 1):n])) / (1 + (k * tau_2)),
                  sd = sqrt(1 / (1 + (k * tau_2)))
    )

    ## Sample from tau_1

    tau_1 <- rgamma(n = 1,
                    shape = 2 + (k / 2),
                    rate = 1 + (1 / 2) * sum((y[1:k] - mu_1)^2)
                    )

    ## Sample from tau_2

    tau_2 <- rgamma(n = 1,
                    shape = 2 + ((n - k) / 2),
                    rate = 1 + (1 / 2) * sum((y[(k + 1):n] - mu_2)^2)
    )

    ## Sample from k

    k_func <- (sqrt(tau_1 / tau_2)) * exp((tau_2 / 2) *
                  cumsum((y - mu_2)^2) - (tau_1 / 2) * cumsum((y - mu_1)^2))

    k <- sample(x = 1:n, size = 1, prob = k_func)

    ## Store the results

    post_k[s] <- k
    post_mu_1[s] <- mu_1
    post_mu_2[s] <- mu_2
    post_tau_1[s] <- tau_1
    post_tau_2[s] <- tau_2

}

## Plot the traceplot for post k

par(mfrow = c(1, 1))
plot(post_k[seq(1, length(post_k), 10)], type = "l",
     main = "Traceplot: Post k", ylab = "post_k")

## Plot the traceplot for post mu_1 and mu_2
```

```r
par(mfrow = c(1, 2))
plot(post_mu_1[seq(1, length(post_mu_1), 10)], type = "l",
     main = "Traceplot: Post mu_1", ylab = "post_mu_1")
plot(post_mu_2[seq(1, length(post_mu_2), 10)], type = "l",
     main = "Traceplot: Post mu_2", ylab = "post_mu_2")

## Plot the traceplot for post tau_1 and tau_2

par(mfrow = c(1, 2))
plot(post_tau_1[seq(1, length(post_tau_1), 10)], type = "l",
     main = "Traceplot: Post tau_1", ylab = "post_tau_1")
plot(post_tau_2[seq(1, length(post_tau_2), 10)], type = "l",
     main = "Traceplot: Post tau_2", ylab = "post_tau_2")



##### Part v #####



par(mfrow = c(1, 1))

hpd_int <- as.vector(HPDinterval(as.mcmc(post_k)))

plot(x = x, y = y, xlab = 't', main = 'BAC Return')
abline(v = hpd_int, lwd = 5, lty = 2)
abline(v = mean(post_k), lwd = 5, lty = 1)



##### Part vi #####



## Compute the hpd intervals

mu_hpd <- as.vector(HPDinterval(as.mcmc(post_mu_2 - post_mu_1)))
tau_hpd <- as.vector(HPDinterval(as.mcmc(post_tau_2 - post_tau_1)))

## Output the results

cat("Posterior Credible Interval, mu_2 - mu_1: ", "(", mu_hpd[1], ",", mu_hpd[2], ")")
cat("Posterior Credible Interval, tau_2 -tauu_1: ", "(", tau_hpd[1], ",", tau_hpd[2], ")")
```