Homework 4

Jon Campbell

```r
library(tidyverse)
```

Question 1

```r
####### Strategy STAY
S <- 10000
stay_strat <- numeric(S)
doors <- 1:3
for (i in 1:S) {
  prize <- sample(x = 1:3, size = 1
                  , prob = c(1/3, 1/3, 1/3))
  my_choice <- sample(x = 1:3, size = 1
                  , prob = c(1/3, 1/3, 1/3))

  if (my_choice == prize) {
    stay_strat[i] <- 1
  } else {
    stay_strat[i] <- 0
  }
}
stay_prob <- sum(stay_strat)/S

#####Strategy SWITCH

S <- 10000
switch_strat <- numeric(S)
doors <- 1:3
for (i in 1:S) {
  prize <- sample(x = 1:3, size = 1
```

```r
                  , prob = c(1/3, 1/3, 1/3))
  my_choice <- sample(x = 1:3, size = 1
                  , prob = c(1/3, 1/3, 1/3))

  rev <- subset(doors, doors != my_choice & doors != prize)

  if (length(rev)==2) {
    revealed <- sample(rev, size = 1, prob = c(1/2,1/2))
  } else {
    revealed <- rev
  }

  switch <- subset(doors, doors != my_choice & doors != revealed)

  if (switch == prize) {
    switch_strat[i] <- 1
  } else {
    switch_strat[i] <- 0
  }
}
switch_prob <- sum(switch_strat)/S


cat("Stay Strategy Success Prob: ",stay_prob,"\n",
    "Switch Strategy Success Prob: ", switch_prob)
```

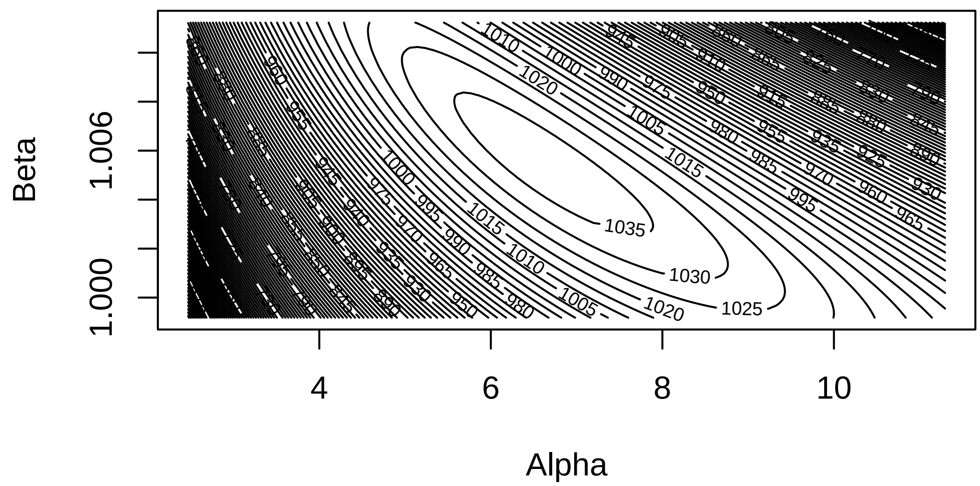
```
Stay Strategy Success Prob:  0.3369
 Switch Strategy Success Prob:  0.6639
```

Question 3

b)

```r
loglikelihood <- function(a, b) {
  logL <- 877*log(a/b)+47370*log(b)-(a/b)*sum(b^(0:99))
  logL
}

a <- 7.06
b <- 1.0042
```

```r
a_seq <- seq(a * 0.35, a * 1.6, length.out = 100)
b_seq <- seq(b * 0.995, b * 1.007, length.out = 100)
logL_matrix <- outer(a_seq, b_seq, Vectorize(loglikelihood))


contour(a_seq, b_seq, logL_matrix, xlab = "Alpha", ylab = "Beta"
        ,nlevels = 100)
```




```r
a_seq <- seq(0.085, 80, length.out = 1000)
b_seq <- seq(0.94, 1.07, length.out = 1000)
logL_matrix <- outer(a_seq, b_seq, Vectorize(loglikelihood))


threshold <- loglikelihood(7.06,1.0042)*0.15

indices <- which(logL_matrix >= threshold, arr.ind = TRUE)

a_range <- range(a_seq[indices[, 1]])
b_range <- range(b_seq[indices[, 2]])
cat("Alpha range: ",a_range,"\n","Beta range: ",b_range)
```

```
Alpha range:  0.164995 63.76102
 Beta range:  0.9544444 1.062843
```

c)

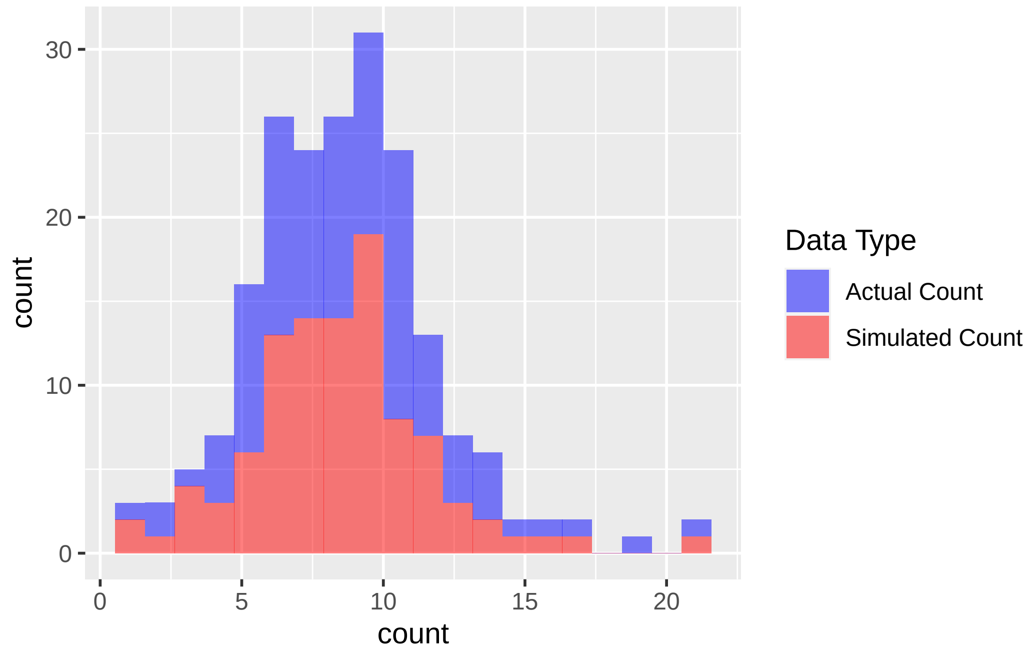
```r
data = read.csv("tropical-cyclones-1901-2000.csv")
```

```r
a <- 7.06
b <- 1.0042

S <- 1000
sim_data <- matrix(NA,nrow = S, ncol = 100)
mses <- numeric(S)

for (i in 1:S) {
  for (j in 1:100) {
    sim_data[i,j] <- rpois(1,a*b^(j-1))
  }
  mses[i] <- sum((data$count - sim_data[i,])^2)
}

cat("Average MSE: ", mean(mses), "\n",
    "Simulated Sample Mean: ",mean(apply(sim_data, MARGIN = 1, FUN = mean)), "\n",
    "Simulated Sample Variance: ", mean(apply(sim_data, MARGIN = 1, FUN = var)), "\n",
    "Observed Sample Mean: ",mean(data$count), "\n",
    "Observed Sample Variance: ",var(data$count), "\n")
```

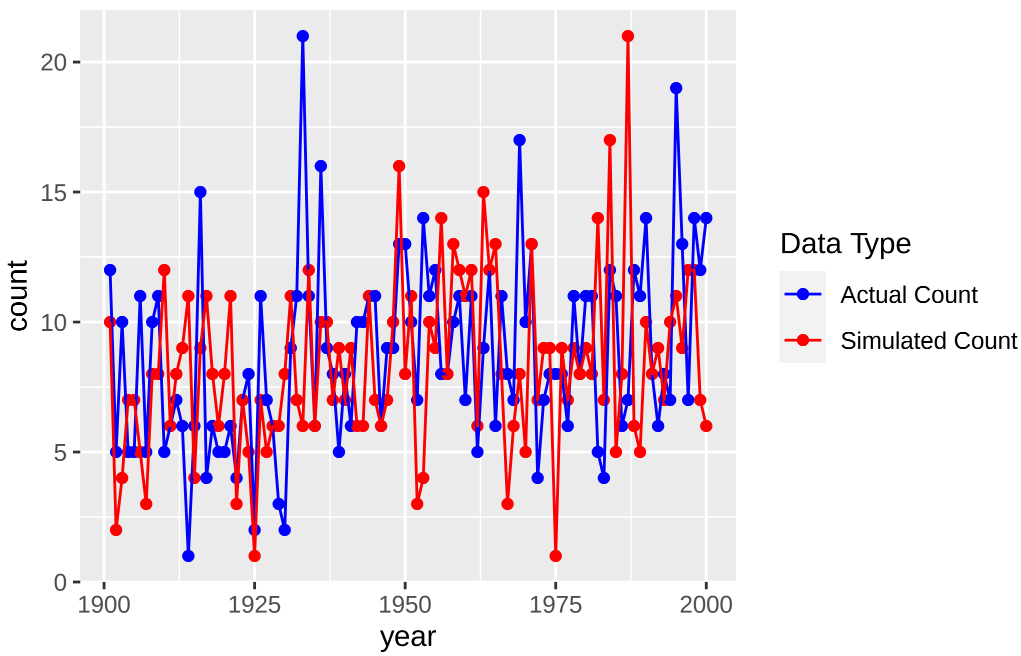
```
Average MSE:  2067.919
 Simulated Sample Mean:  8.75979
 Simulated Sample Variance:  9.95164
 Observed Sample Mean:  8.77
 Observed Sample Variance:  13.16879
```

```r
avg_sim_counts <- apply(sim_data, MARGIN = 2, FUN = mean)
data |>
  mutate(sim = sim_data[1,]) |>
  pivot_longer(cols = c(count, sim), names_to = "type", values_to = "count") |>
  ggplot(aes(x = count, fill = type)) +
  geom_histogram(alpha = 0.3, bins = 20) +
```

```
  geom_histogram(alpha = 0.3, bins = 20) +
  scale_fill_manual(values = c("blue", "red"), name = "Data Type",
                    labels = c("Actual Count", "Simulated Count"))
```




```
data |>
  mutate(sim = sim_data[1,]) |>
  pivot_longer(cols = c(count, sim), names_to = "type", values_to = "count") |>
  ggplot(aes(x = year, , y = count, color = type)) +
  geom_point() +
  geom_line() +
  scale_color_manual(values = c("blue", "red"), name = "Data Type",
                    labels = c("Actual Count", "Simulated Count"))
```


count
year
0
5
10
15
20
1900
1925
1950
1975
2000
Data Type
Actual Count
Simulated Count