Assignment 2 Question 2

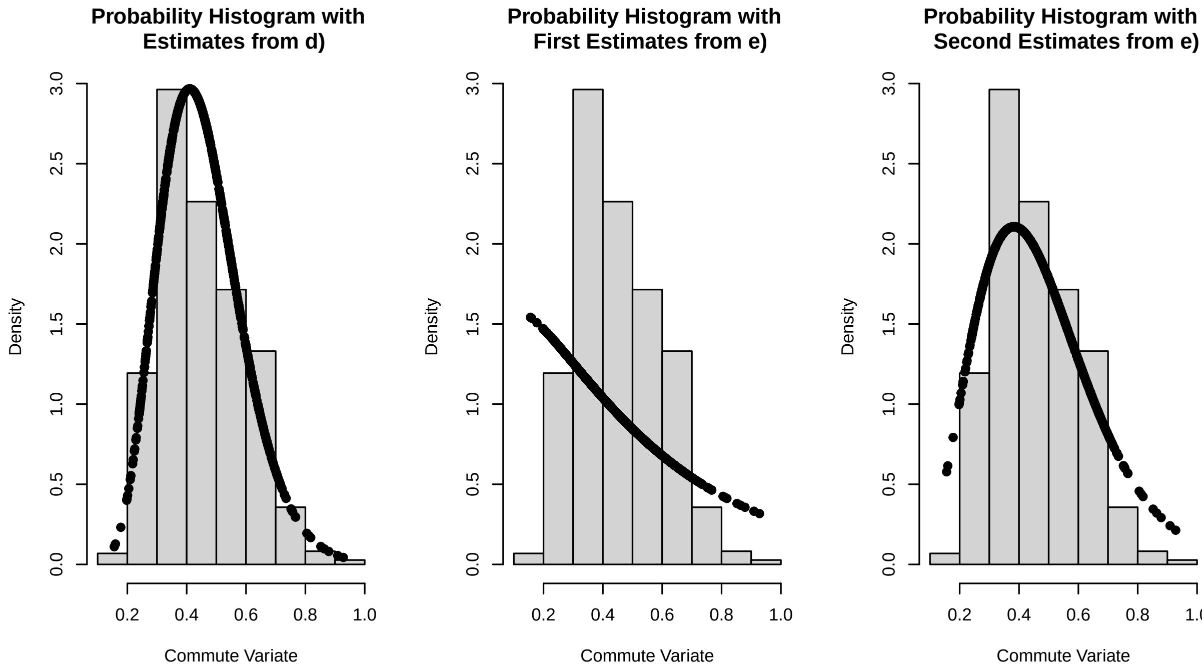
Sheen Thusoo

Part g)

```
data <- read.csv("EconomicMobility.csv")
```

```
pdf <- function(alpha, beta) {
  y <- data$Commute
  result <- array(NA, dim = length(y))
  for (i in 1:length(y)) {
    result[i] <- ( (beta^alpha * y[i]^(alpha - 1)) / gamma(alpha) ) * exp(-y[i] * beta)
  }
  return(result)
}
```

```
# values from d)
alpha1 <- 10.47992
beta1 <- 23.09889

# values from first part of e)
alpha2 <- 1.278860
beta2 <- 2.692789

# values from second part of e)
alpha3 <- 5.244547
beta3 <- 11.094237

par(mfrow = c(1, 3))
hist(data$Commute, probability=TRUE,
     main='Probability Histogram with \n Estimates from d)', xlab='Commute Variate',
     ylab='Density')
points(data$Commute, pdf(alpha1, beta1), pch = 19)

hist(data$Commute, probability=TRUE,
     main='Probability Histogram with \n First Estimates from e)', xlab='Commute Variate',
     ylab='Density')
points(data$Commute, pdf(alpha2, beta2), pch = 19)

hist(data$Commute, probability=TRUE,
     main='Probability Histogram with \n Second Estimates from e)', xlab='Commute Variate',
     ylab='Density')
points(data$Commute, pdf(alpha3, beta3), pch = 19)
```



What do you observe? Justify your observation.
The pdf of the Gamma distribution with the estimates from d) superimposed onto the probability histogram of the *Commute* variate fits the shape of the histogram quite closely. These estimates do the best job of the three estimates in matching the histogram. This is because we used the following sequence for lambda: $\lambda = 0, 0.01, 0.02, ..., 5$. Thus, our $lambaMax = 5$ and $lambdaStepSize = 0.01$. This combination helped us converge to the correct values of alpha and beta since *lambda* is the step size taken by the gradient ascent function. This step size was small enough and the maximum lambda value was large enough that the gradient ascent function could find the optimal values of alpha and beta to converge to.

The pdf of the Gamma distribution with the first estimates from e) has a negative slope and seems to be decreasing as the *Commute* variate increases. These estimates do not do a good job of matching the histogram. This is because we used the following sequence for lambda: $\lambda = 0, 1, 2, ..., 5$. Thus, our $lambdaMax = 5$ and $lambdaStepSize = 1$. In this case, the step size is much too large so the solution converges to the wrong value. A large step size takes larger steps towards the solution, so it can plausibly jump over the the maxima we are trying to reach. Thus, our solutions do not converge to the correct values of alpha and beta since we overshoot.

The pdf of the Gamma distribution with the second estimates from e) has the same general shape as the probability histogram of the *Commute* variate; however, it does not exactly match the peak of the histogram. These estimates do not do as good of a job as the estimates from d). This is because we used the following sequence for lambda: $\lambda = 0, 0.0001, 0.0002, ..., 5$. Thus, our $lambdaMax = 0.0001$ and $lambdaStepSize = 0.01$. Thus, our $lambdaMax = 0.01$ and $lambdaStepSize = 0.0001$. In this case, the step size is much too small so the gradient ascent function does not reach the maximum solution. A very small step size will cause the gradient ascent function to be very slow - as we saw in e) it reaches the max iterations and does not converge. Thus, this causes the solution to diverge and not reach the correct solution.