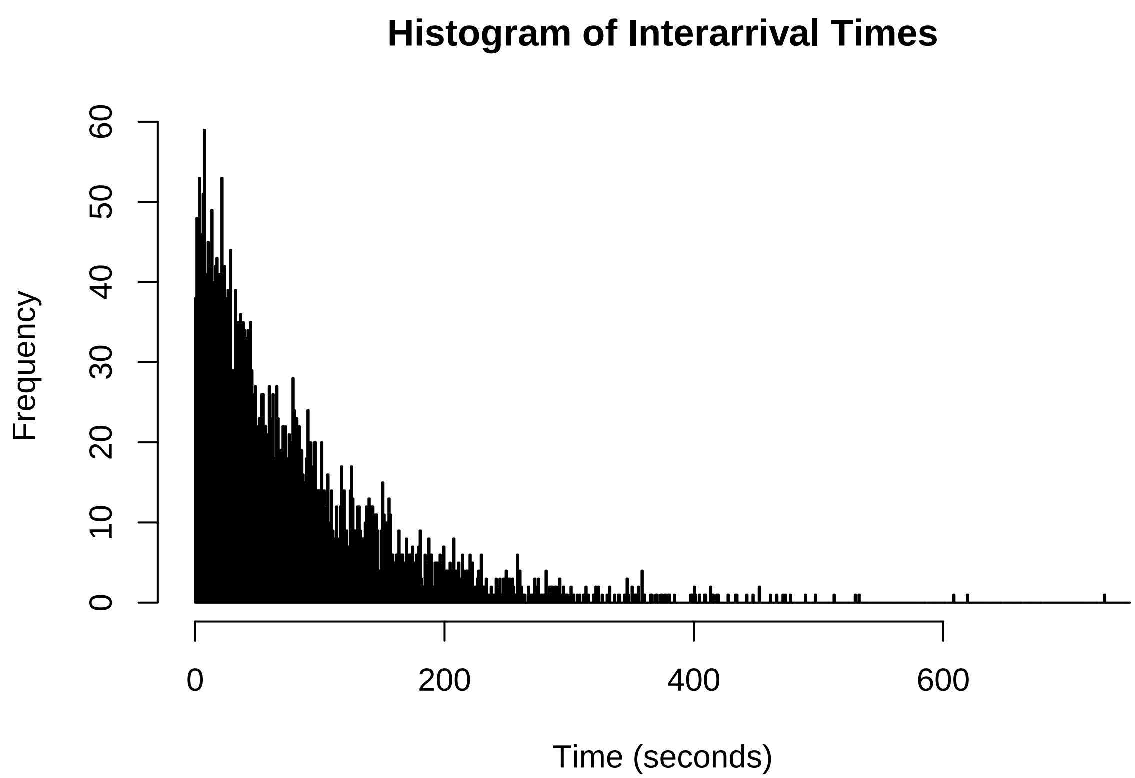
# Problem 43

## Gamma Ray Interarrival Times

### Initial Histogram

The data consists of 3935 observations of interarrival times for gamma rays, with units in seconds.

```
hist(data, breaks = seq(from = 0, to = 750, by = 1), main = 'Histogram of Interarrival Times', xlab = '
```




The interarrival times look roughly exponential, or at least gamma, since the exponential distribution is a special case of the gamma distribution. For the moment, we will assume that interarrival times are distributed via a gamma distribution.

### Parameter Estimates

#### Method of Moments

The method of moments estimators for $\alpha$ and $\gamma$ are

$$\hat{\alpha}_{MoM} = \frac{\bar{X}^2}{\bar{\sigma}^2}$$

$$\hat{\lambda}_{MoM} = \frac{\bar{X}}{\bar{\sigma}^2}$$

```
#Computes the MoM estimates of alpha and lambda for a gamma distribution, respectively
gammaMoM <- function(data) {
  output <- list()
  output$alpha <- (mean(data)^2) / var(data)
  output$lambda <- mean(data) / var(data)
  return(output)
}
foo <- gammaMoM(data)
aMoM <- foo$alpha
lMoM <- foo$lambda
rm(foo)
```

Thus the method of moments estimates are $\hat{\alpha}_{MoM} = 1.0120949$ and $\hat{\lambda}_{MoM} = 0.0126614$.

### Maximum Likelihood

The maximum likelihood estimator for $\alpha$ satisfies the following equation

$$nln(\hat{\alpha}) - nln(\bar{X}) + \sum_{i=1}^{n} ln(X_i) - n\psi(\hat{\alpha}) = 0$$

Where $\psi(x)$ is the digamma function, $\frac{\Gamma'(x)}{\Gamma(x)}$. Then an estimator for $\lambda$ is

$$\hat{\lambda}_{MLE} = \frac{\hat{a}}{\bar{X}}$$

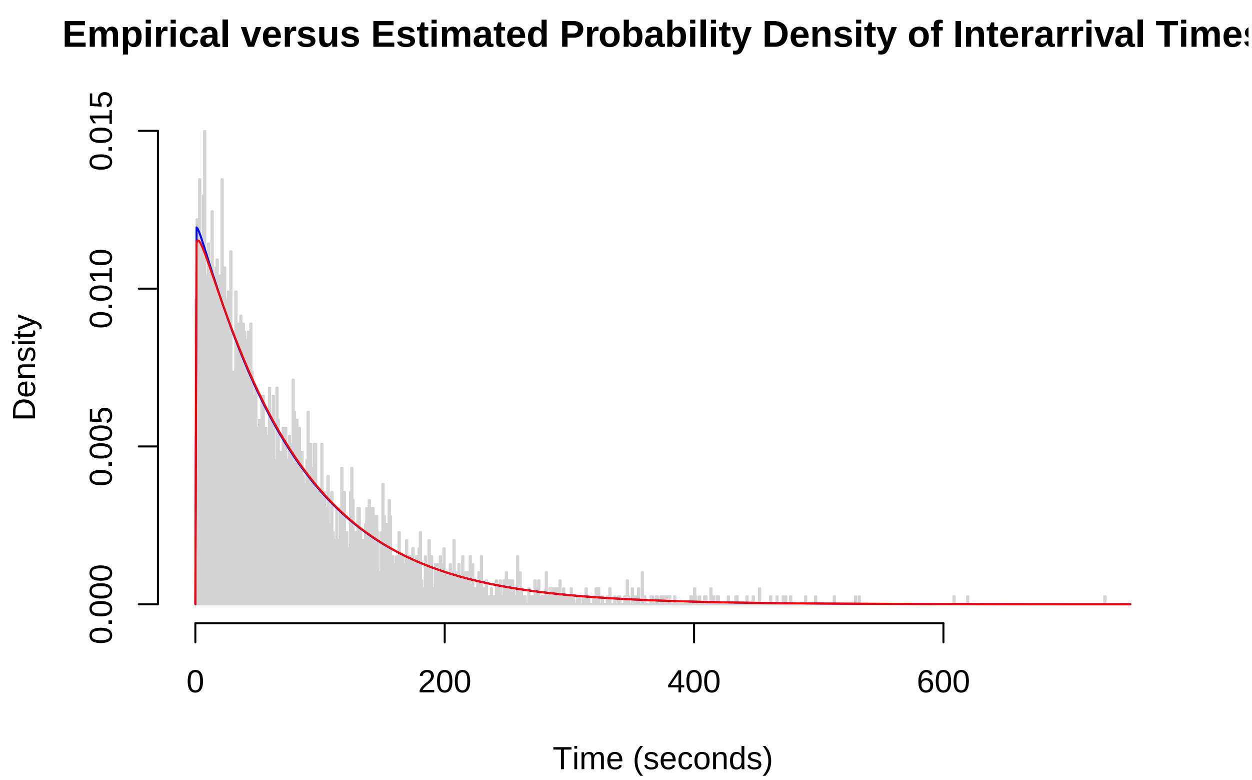
We will use the method of moments estimator as an initial guess to solve for $\hat{\alpha}_M LE$.

```
#Computes the MLE estimates of alpha and lambda for a gamma distribution, respectively
gammaMLE <- function(data) {
  aMoM <- gammaMoM(data)$alpha
  #Initialize variables
  n <- length(data)
  lnXBar <- log(mean(data))
  sumLnX <- sum(log(data))
  aMLEFun <- function(a) {
    n*log(a) - n*lnXBar + sumLnX - n*digamma(a)
  }

  #diagnostic plots
  #x <- seq(from = aMoM - .5, to = aMoM + .5, by = 0.01)
  #plot(x, aMLEFun(x))
  output = list()
  output$alpha <- uniroot(aMLEFun, interval = c(aMoM - .5, aMoM + .5))$root
  output$lambda <- output$alpha / mean(data)
  return(output)
```

```
}

foo <- gammaMLE(data)
aMLE <- foo$alpha
lMLE <- foo$lambda
rm(foo)
```

Thus the maximum likelihood estimates are $\hat{\alpha}_{MLE} = 1.026308$ and $\hat{\lambda}_{MLE} = 0.0128392$. We see that the two are nearly identical. Looking at the estimated distributions versus the histogram shows that the fit is good, and that the two estimated densities are nearly identical.

```
hist(data, freq = FALSE, breaks = seq(from = 0, to = 750, by = 1), density = 1, main = 'Empirical versus
xSeq <- seq(from = 0, to = 750, by = 1)
#MoM fit
lines(xSeq, dgamma(xSeq, shape = aMoM, rate = lMoM), col = 'blue', lwd = 1)
#MLE fit
lines(xSeq, dgamma(xSeq, shape = aMLE, rate = lMLE), col = 'red', lwd = 1)
```




### Estimating Standard Errors via Bootstrap

```
#First step in a bootstrap. Generates simulated gamma random variables.
#Creates dataSize iid draws from a gamma distribution, and estimate alpha and lambda hat using calcFun.
#Returns a list with alphaHats, lambdaHats being the respective estimates
```

```
gammaSim <- function(dataSize, reps, alpha, lambda, calcFun) {
  output <- list()
  output$alphaHats <- vector(mode = 'numeric', length = reps)
  output$lambdaHats <- vector(mode = 'numeric', length = reps)

  for (i in 1:reps) {
    data <- rgamma(dataSize, shape = alpha, rate = lambda)
    ests <- calcFun(data)
    output$alphaHats[i] <- ests$alpha
    output$lambdaHats[i] <- ests$lambda
  }

  return(output)
}


#Bootstrap estimation of standard errors of alpha and lambda of a gamma distribution calculated uses ca
gammaSE <- function(dataSize, reps, alpha, lambda, calcFun) {
  ests <- gammaSim(dataSize, reps, alpha, lambda, calcFun)

  #Return the estimated standard errors
  output <- list()
  output$alphaSE <- sd(ests$alphaHats)
  output$lambdaSE <- sd(ests$lambdaHats)
  return(output)
}
```

```
set.seed(1000)
MoMSE <- gammaSE(dataSize = length(data), reps = 1000, alpha = aMoM, lambda = lMoM, calcFun = gammaMoM)
MLESE <- gammaSE(dataSize = length(data), reps = 1000, alpha = aMLE, lambda = lMLE, calcFun = gammaMLE)
```

The estimates for $s_{\hat{\alpha}_{MoM}}$ and $s_{\hat{\lambda}_{MoM}}$ are thus 0.0311579 and $4.3464725 \times 10^{-4}$. The corresponding estimates for $s_{\hat{\alpha}_{MLE}}$ and $s_{\hat{\lambda}_{MLE}}$ are thus 0.0208755 and $3.2373483 \times 10^{-4}$. The estimated standard errors for the maximum likelihood estimates are significantly lower than the corresponding estimates for the method of moments.

## Estimating Confidence Intervals via Bootstrap

```
#Computes the bootstrapped 100(1-confidenceAlpha) confidence intervals of gamma distribution parameters
#Note the confidenceAlpha should be a decimal, not a percent.
#Returns a list with alphaCI and lambdaCI being the confidence intervals.
gammaCI <- function(dataSize, reps, alpha, lambda, calcFun, confidenceAlpha) {
  #Convert confidenceAlpha to a decimal if it was accidentally inputted as a percent
  if (confidenceAlpha > 1) {
    confidenceAlpha = confidenceAlpha / 100
  }
  ests <- gammaSim(dataSize, reps, alpha, lambda, calcFun)

  #Order the lists
  sortedAlpha <- sort(ests$alphaHats)
  sortedLambda <- sort(ests$lambdaHats)
```

```r
  low <- floor((confidenceAlpha / 2) * reps)
  high <- reps - low

  output <- list()
  output$alphaCI <- c(sortedAlpha[low], sortedAlpha[high])
  output$lambdaCI <- c(sortedLambda[low], sortedLambda[high])
  return(output)
}
```

```r
set.seed(2000)
MoMCI <- gammaCI(dataSize = length(data), reps = 1000, alpha = aMoM, lambda = lMoM, calcFun = gammaMoM,
MLECI <- gammaCI(dataSize = length(data), reps = 1000, alpha = aMLE, lambda = lMLE, calcFun = gammaMLE,
```

The estimated 95% confidence intervals for $\hat{\alpha}_{MoM}$ and $\hat{\lambda}_{MoM}$ are (0.9507651, 1.0763429) and (0.0118158, 0.0135985), respectively. The corresponding estimates for $\hat{\alpha}_{MLE}$ and $\hat{\lambda}_{MLE}$ are (0.9864015, 1.06545) and (0.0122045, 0.0135627), respectively. As we can see, the parameter estimates estimated via method of moments are much less dispersed than the corresponding estimates from method of moments.

The interarrival times are consistent with a Poisson arrival process. In a Poisson arrival process, interarrival times are exponentially distributed, which corresponds to a gamma distribution with $\alpha = 1$. As we can see by the 95% confidence intervals, we can not reject the null hypothesis that $\alpha = 1$ for neither the method of momentsn nor maximum likelihood estimates.