Assignment-1

Name: Keshav Gambhir Roll Number:2019249

Question-1

We need to find the maximum likelihood function for the gamma distribution. The following is the calculation for the same

Ans-1

a. Given $x_i \overset{iid}{\sim}$ Gamma($\lambda = 2, r = 5$)

to find: log maximum likelyhood function

$$f(x) = \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$$

for calculating the likelyhood f^n we would take product of $f(x_i)\ \forall i$ from $1 \dots n$

$$L(\theta) = \prod_{i=1}^{n} f(x) = \prod_{i=1}^{n} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}$$

for calculating the log likelyhood fx^n we would take natural log on both side

$$\ln(L(\theta)) = \ln\left[\prod_{i=1}^{n} f(x)\right] = \ln\left[\prod_{i=1}^{n} \frac{\lambda^r}{\Gamma(r)} x^{r-1} e^{-\lambda x}\right]$$

$$= nr\ln(\lambda) - \lambda \sum x_i - \ln(\Gamma(r)^n) + (r-1)\ln\left(\sum_{i=1} x_i\right)$$

$$\boxed{\ln(L(\theta)) = nr\ln(\lambda) - \lambda \sum x_i + (r-1)\sum_{i=1} \ln(x_i) - n\ln(\Gamma(r))}$$

The derived formula has been coded using R as follows

```r
# Question-1 Part A
maxLikelyhoodFunction.gamma <- function(params) {
  return (-1*n*params[1]*log(params[2]) + n*lgamma(params[1]) + sum(data)*params[2] - (params[1]-1)*sum(log(data)))
}
```

The function **maxLikelyhoodFunction.gamma** returns the log of maximum likelihood function as derived in the above derivation.

In part B of question 1 we need to generate a 1000 random sample of gamma distribution. The following are the generated data points and the code for generating the data

```r
n <- 1000
r <- 5
lambda <- 2
data <- round(rgamma(n,shape = r,rate = lambda),2)
print(data)
```

```
2.91 3.07 2.14 2.27 5.38 2.63 1.96 4.36 1.80 2.66 2.84 1.52 2.74 3.60 2.58 3.15 4.06 1.62 0.80 4.57 1.77 2.25 2.08 1.62
  [25] 4.95 4.31 2.53 2.80 1.92 2.58 2.69 1.29 5.96 2.26 0.91 1.00 2.29 3.02 1.98 4.79 2.31 2.64 1.59 1.95 3.21 5.10 5.00 3.23
  [49] 1.40 0.73 2.72 3.63 1.40 2.31 1.49 4.47 2.24 3.36 3.21 1.47 2.07 2.99 3.49 1.94 1.36 4.67 3.02 1.94 3.99 3.37 2.02 4.70
  [73] 1.85 1.31 1.75 0.40 1.57 3.76 2.29 3.08 0.94 1.72 1.62 1.80 1.35 1.06 2.88 2.08 1.75 2.83 2.29 3.97 1.94 3.87 2.60 4.17
  [97] 1.83 4.47 1.62 6.63 3.95 1.02 5.95 1.28 5.84 1.97 1.82 3.19 4.42 2.86 0.91 1.68 2.40 0.86 2.69 2.77 1.55 1.99 1.51 2.32
 [121] 1.61 2.75 2.00 1.77 1.53 2.97 1.13 2.95 3.68 2.82 3.42 2.61 4.25 3.04 2.59 2.63 2.94 2.46 3.17 1.35 1.84 1.60 2.33 4.06
 [145] 2.66 2.07 2.14 1.75 3.54 1.41 3.18 3.24 1.80 1.94 0.98 1.54 3.51 3.22 2.36 0.78 3.10 1.96 3.43 1.43 3.45 1.75 2.60 3.00
 [169] 4.52 3.27 2.78 3.84 0.32 3.02 2.03 2.83 1.61 2.33 1.50 1.98 3.29 1.78 2.65 1.76 2.52 1.57 2.34 2.66 3.58 3.81 3.30 2.50
 [193] 1.81 3.12 1.47 4.13 3.54 2.63 3.01 1.92 1.79 1.58 3.07 3.80 2.10 2.45 3.28 1.88 2.79 1.17 2.01 1.57 2.35 2.59 3.08 2.10
 [217] 1.93 1.04 0.93 2.78 1.63 2.99 2.65 2.36 1.59 1.55 1.82 1.83 2.69 1.33 1.15 1.64 4.79 3.15 0.47 1.81 3.52 2.17 3.70 2.63
 [241] 1.91 1.49 2.16 4.74 2.38 1.86 1.68 2.07 1.51 2.27 1.77 2.78 0.53 1.80 2.68 1.54 1.27 2.79 2.64 3.25 2.70 2.66 0.93 1.44
 [265] 2.77 4.99 1.94 2.74 3.70 2.84 3.90 0.88 0.34 2.23 2.16 2.92 3.14 1.85 2.64 2.28 2.78 1.38 5.66 2.99 2.59 4.37 3.02 1.97
 [289] 0.94 2.78 2.91 3.01 1.50 2.13 3.14 5.29 1.77 2.34 4.01 3.29 1.42 3.31 5.26 1.16 3.35 1.93 2.38 1.88 1.80 2.98 2.59 2.21
 [313] 1.99 2.25 2.56 1.31 1.68 2.16 3.05 4.16 2.35 2.56 3.01 1.24 1.41 2.66 2.19 1.21 2.39 1.67 1.69 1.34 1.82 1.75 3.75 2.94
 [337] 1.46 3.27 2.71 0.96 2.80 5.25 3.30 3.18 4.77 2.69 2.97 2.36 3.45 3.72 2.81 7.78 1.75 3.07 3.47 1.69 2.41 1.35 4.12 2.09
 [361] 1.83 1.09 2.38 1.67 2.12 2.66 3.64 2.55 1.87 3.84 2.22 3.93 2.43 2.41 1.83 1.74 1.36 2.22 2.72 2.28 2.02 2.48 3.12 2.40
 [385] 2.42 2.50 1.19 4.22 1.88 2.01 1.64 1.44 1.52 6.31 1.60 7.17 2.08 1.84 3.62 1.30 3.24 2.20 2.30 2.82 1.49 2.09 2.93 1.55
 [409] 3.09 1.83 2.26 4.74 2.65 1.05 0.77 4.93 1.92 1.89 4.94 0.71 3.68 2.88 2.19 3.08 1.39 4.73 4.30 5.02 1.58 2.64 1.98 1.78
 [433] 2.07 3.65 3.98 5.06 1.24 3.32 1.08 2.37 2.32 1.63 1.43 1.59 1.70 1.29 3.38 1.00 2.85 1.15 2.22 3.75 2.78 0.91 2.11 1.05
 [457] 1.38 3.09 2.47 0.89 2.08 3.53 5.44 3.44 1.81 4.39 2.02 3.18 1.43 0.85 2.48 4.46 1.51 1.07 2.56 4.62 2.83 1.33 2.50 1.01
 [481] 1.59 3.18 2.19 2.51 1.57 4.05 2.47 2.19 0.97 2.18 2.18 2.56 3.33 1.28 2.45 2.58 1.26 1.65 2.13 2.12 1.78 3.40 3.55 1.88
 [505] 0.70 2.09 2.72 1.80 1.47 2.00 2.30 2.57 5.43 2.13 4.69 1.03 4.32 2.25 1.95 2.55 1.93 2.25 3.39 2.95 1.88 3.75 2.79 1.57
 [529] 2.28 1.59 1.50 1.24 3.77 3.72 3.39 3.65 2.67 2.09 0.70 1.81 4.05 5.23 1.78 2.88 5.36 4.90 2.19 3.49 1.85 3.34 1.88 2.12
 [553] 2.24 3.30 3.43 1.19 1.52 2.22 0.81 2.00 1.72 5.23 3.06 3.28 1.31 2.09 2.18 2.68 2.63 0.69 2.70 2.03 1.35 1.66 1.83 2.77
 [577] 4.38 2.32 0.89 1.67 3.49 1.53 1.53 5.48 5.42 1.62 4.64 1.37 2.28 2.60 2.99 1.14 3.26 3.40 2.95 5.68 3.64 4.68 2.01 0.90
 [601] 2.80 3.22 3.34 3.36 1.70 2.05 1.31 1.53 2.32 3.44 1.87 5.10 3.48 3.27 1.63 2.70 4.04 2.07 2.82 5.06 1.47 1.74 3.01 2.80
 [625] 3.54 3.63 0.86 0.45 1.53 1.78 3.37 1.60 2.62 1.73 3.27 2.40 3.36 1.60 0.69 2.22 2.26 1.41 1.55 2.39 3.16 1.88 3.24 2.95
 [649] 0.64 2.33 2.41 1.72 3.29 1.23 2.08 4.42 3.27 4.35 2.23 3.29 1.72 2.40 3.03 1.69 2.95 2.26 2.82 4.35 4.43 1.84 7.68 2.72
 [673] 3.52 1.54 3.35 2.92 4.12 0.79 2.83 3.31 2.38 3.36 2.83 1.70 1.61 1.67 2.42 3.03 2.05 2.38 4.09 5.76 1.97 2.70 1.71 4.49
 [697] 3.12 2.12 2.38 3.55 1.81 1.93 3.07 1.55 2.93 3.03 2.26 2.04 1.23 4.13 3.18 5.08 3.47 3.74 2.04 4.11 1.91 1.15 2.03 3.43
 [721] 3.81 2.34 2.28 2.59 2.17 1.53 3.25 2.03 1.33 3.68 2.04 0.90 2.16 3.41 1.92 1.95 2.05 1.12 1.70 1.71 1.50 1.98 4.24 2.11
 [745] 1.35 1.24 2.01 1.89 2.49 1.75 3.55 3.22 1.95 3.45 3.08 1.37 4.66 1.97 5.10 1.14 1.93 4.26 2.01 2.77 4.40 2.82 3.57 2.54
 [769] 2.45 2.22 4.38 2.34 1.43 1.93 0.69 1.28 3.02 1.88 3.27 4.95 2.83 3.10 2.85 4.95 2.17 3.14 2.16 2.69 2.20 2.14 1.38 2.16
 [793] 1.18 1.74 4.01 2.81 2.56 1.55 1.67 2.68 1.23 3.26 3.76 1.71 1.58 1.24 0.98 3.07 2.66 1.63 7.21 1.61 1.45 2.49 3.33 3.21
 [817] 2.03 2.04 1.14 1.16 3.62 4.34 3.12 3.50 1.41 3.53 1.25 2.65 1.04 2.80 1.37 3.30 1.08 2.24 5.69 1.77 2.84 3.65 2.06 3.71
 [841] 2.22 2.00 2.38 3.63 2.76 4.91 1.57 2.90 2.97 4.34 3.24 3.97 2.53 1.12 2.16 2.91 2.39 2.95 1.88 2.90 1.85 1.82 2.49 3.32
 [865] 4.20 3.80 3.13 3.00 3.00 4.20 2.71 2.28 0.44 4.05 3.54 1.14 3.37 5.66 3.24 0.89 4.02 3.56 2.12 2.18 1.93 5.46 2.30 2.45
 [889] 1.75 0.84 1.78 1.49 3.17 3.50 2.77 4.23 3.69 3.04 1.71 5.88 1.15 2.10 3.75 2.01 3.75 0.92 2.53 1.01 2.64 2.38 0.74 3.03
 [913] 1.18 1.19 2.46 1.49 2.43 2.36 1.17 1.48 1.75 4.37 1.60 2.52 3.63 3.21 2.63 1.86 3.79 1.64 2.71 3.07 3.70 2.44 3.59 4.75
 [937] 2.02 0.87 6.54 4.97 3.15 2.06 0.84 1.29 1.93 0.81 3.23 2.31 2.62 0.99 3.57 1.11 2.68 1.78 2.90 0.97 0.91 1.64 3.06 3.93
 [961] 2.41 2.29 3.91 1.88 2.22 3.59 2.48 2.91 1.58 2.31 3.01 5.56 3.48 3.79 0.57 3.41 2.46 0.87 3.07 0.74 2.02 2.56 3.48 5.34
 [985] 2.19 1.85 1.82 3.14 3.00 1.69 3.01 2.65 1.76 1.38 0.58 2.13 0.83 2.61 1.58 1.97
```

The next step is to calculate the MLE of the distribution using the method of moments and using the random values

To calculate the output the following code snippet has been used

```
# Question-1 Part B
#Computing using method of moments
# moment alpha of gamma data is mean square/variance
momentAlpha <- mean(data)^2/var(data)
# moment beta of gamma data is mean/variance
momentBeta <- mean(data)/var(data)
output <- nlm(maxLikelyhoodFunction.gamma,c(momentAlpha,momentBeta),hessian = T)
print(output)

#Computing using random initial value-1
output2 <- nlm(maxLikelyhoodFunction.gamma,c(8,0.8),hessian = T)
print(output2)

#Computing using random initial value-3
output3 <- nlm(maxLikelyhoodFunction.gamma,c(3,3),hessian = T)
print(output3)
```

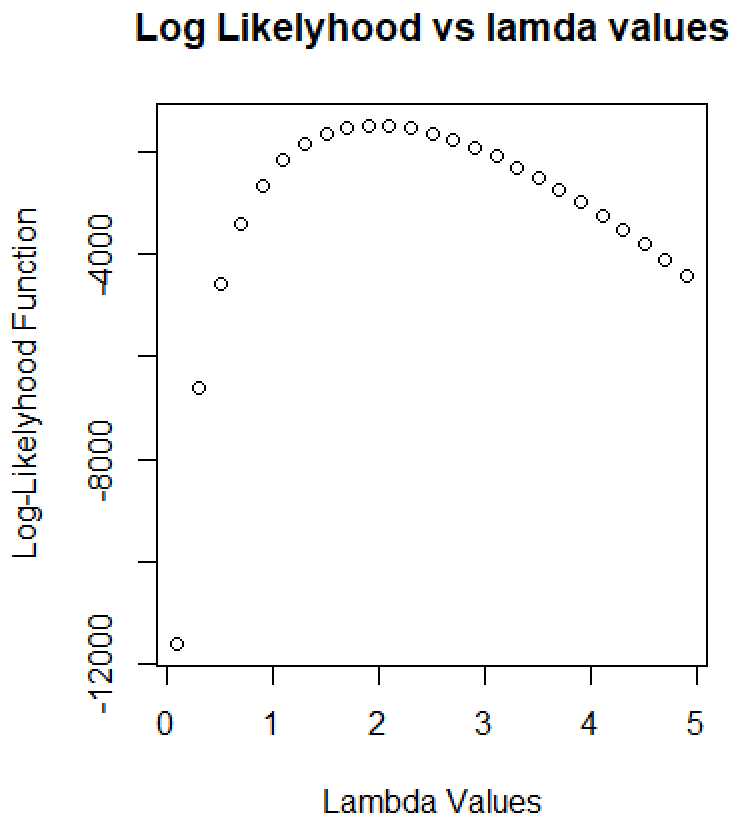
The output for the MLE for unknown parameters are:

```
$estimate
[1] 4.958907 1.941341
```

In the third part we need to generate different values of lambda and plot the graph of Log-MLE against the various values of lambda

The plot obtained is as follows:

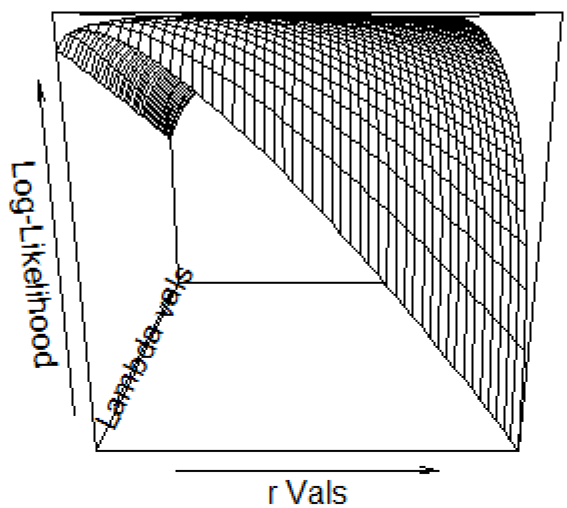
```
# Question-1 Part C
lambdaVals = seq(0.1,5,by = 0.2)
mleOutput <- list()
i = 1
for (lVal in lambdaVals) {
  output = -1*maxLikelyhoodFunction.gamma(c(5,lVal))
  mleOutput[i] <- output
  i <- i +1
}
plot(lambdaVals,mleOutput,xlab = "Lambda Values",ylab = "Log-Likelyhood Function", main = "Log Likelyhood vs lamda values")
```



In the last part we have to repeat the same with different values of r and lambda.

```
# Question-1 Part D
rVals = seq(0.1, 10, by=0.2)
lambdaVals = seq(0.1,4,by = 0.1)
maxLikelyhoodFunction.gammaNew <- function(alpha,beta) {
  tmp <- -1*n*alpha*log(beta) + n*lgamma(alpha) + sum(data)*beta - (alpha-1)*sum(log(data))
  return (tmp)
}
mleOutput2 = -1*outer(rVals,lambdaVals,maxLikelyhoodFunction.gammaNew)
print(mleOutput2)
persp(rVals,lambdaVals,mleOutput2)
```

Log Likelyhood vs r and lambda



Question-2

In the first part of the question we have to first take the csv as an input and find the mean and the variance of the data. This is done as follows

```
data = unlist((read.csv(file = "C:\\Users\\Keshav Gambhir\\Desktop\\Assignment-1 SI\\data_Q2.csv"))['x'])
dataMean = mean(data)
dataVar = var(data)
n = 1000
```

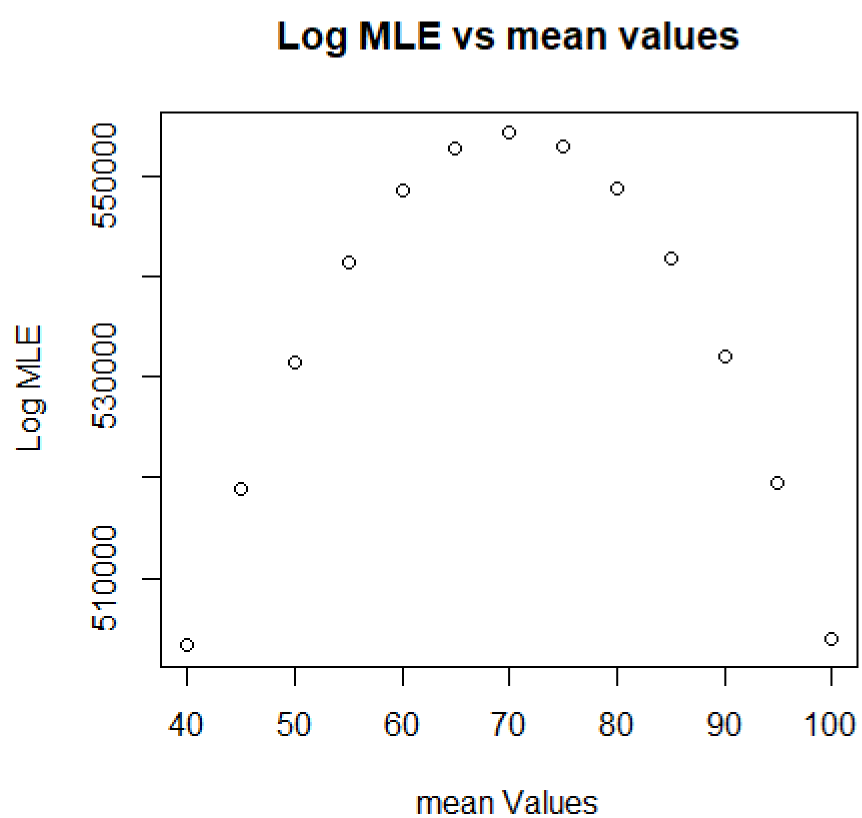
```
mle.norm<-function(params)
{
  loglikelihood.norm <- ((n/2)*log(2*pi*params[2])) - (-1/(2*params[2]))*(sum(data^2) + 2*params[1]*sum(data) - n*(params[1]^2))
  return (-1*loglikelihood.norm)
}

output <- nlm(mle.norm,c(dataMean,dataVar),hessian = T)
output
```

The output estimator for mean and variance is:70.10977 and 61930.00757

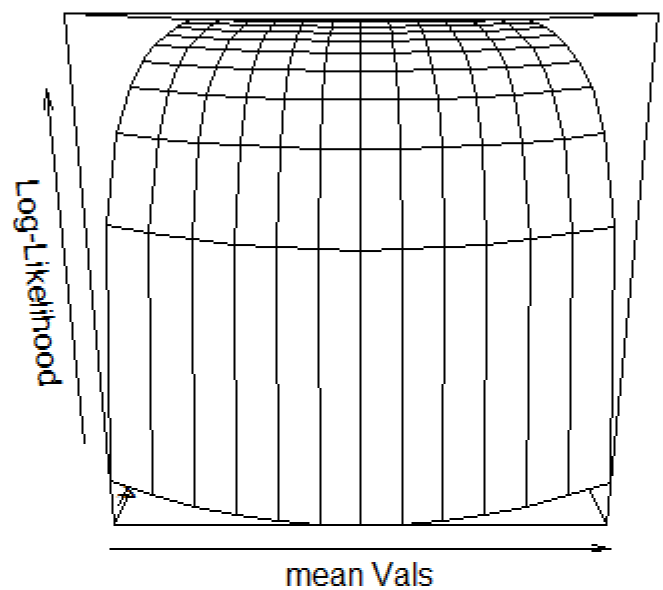
```
$estimate
[1]    70.11377 61930.00757
```

In the second part of the question we have to show that the maxima of the graph of likelihood function is attained at above estimate.

So plotting the graph between various mean values and the log MLEs we can clearly see the max is around 71



Log Likelyhood vs mu and sigma square



Derivation of the Log MLE of the normal distribution

Given: $x_i \overset{iid}{\sim} N(\mu, \sigma^2)$

To find: MLE

$$PDF = \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)2}$$

$\because$ the x_i are IID we can take the product of IIDs to find $L(\theta)$

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}$$

to find log of $L(\theta)$ we take natural log on both sides

$$\ln(L(\theta)) = \ln\left(\prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-1/2\left(\frac{x-\mu}{\sigma}\right)^2}\right)$$

$$= \log_e\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \log_e\left(-\frac{1}{2}\left(\frac{x_1-\mu}{\sigma}\right)^2\right) + \log\left(\frac{1}{\sigma\sqrt{2\pi}}\right) + \log_e\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)\right.$$

$$- - - -$$

$$\Rightarrow n \log_e\left(\frac{1}{\sigma\sqrt{2\pi}}\right) - \frac{1}{2}\left\{\left(\frac{x_1-\mu}{\sigma}\right)^2 + \left(\frac{x_2-\mu}{\sigma}\right)^2 - - - \left(\frac{x_n-\mu}{\sigma}\right)^2\right\}$$

$$\Rightarrow -n\left[\frac{1}{2}\log_e(2\pi) + \log_e(\sigma)\right] - \frac{1}{2}\Sigma\left(\frac{x_i-\mu}{\sigma}\right)^2$$

$$\text{Log MLE} = -\frac{n}{2}\log_e(2\pi) - n\log_e(\sigma) - \frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-\mu}{\sigma}\right)^2$$