PSTAT 220A HW #1

Jaxon Stuhr

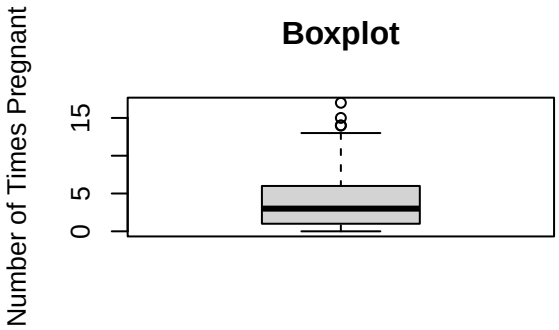
2022-09-27

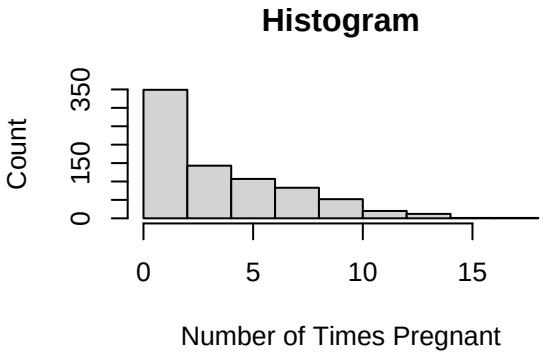
Question 1

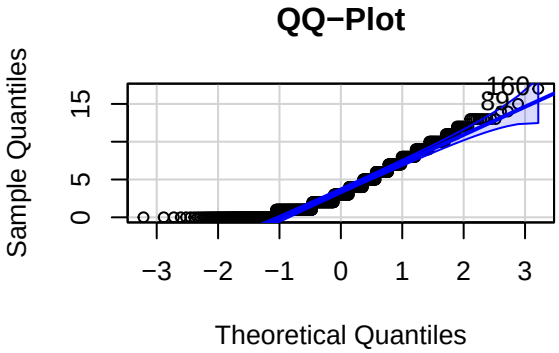
```
# this function returns the standard deviation, mean absolute deviation, median absolute deviation, and
dispersion_measures = function(dataset_x) {
  sd = sd(dataset_x) # standard deviation
  mean_ad = mad(dataset_x, center = mean(dataset_x)) # mean absolute deviation
  median_ad = mad(dataset_x, center = median(dataset_x)) # median absolute deviation
  iqr = IQR(dataset_x) # inter-quartile range
  return(c(sd, mean_ad, median_ad, iqr))
}
```

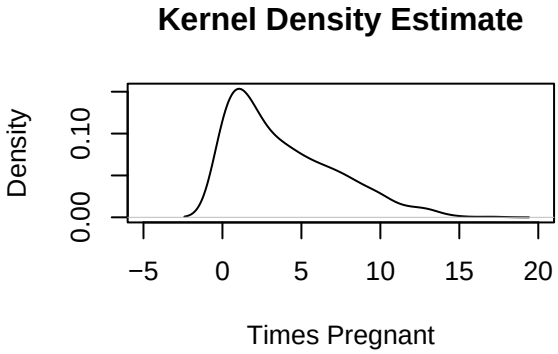
Question 2

```
## [1] 160  89
```

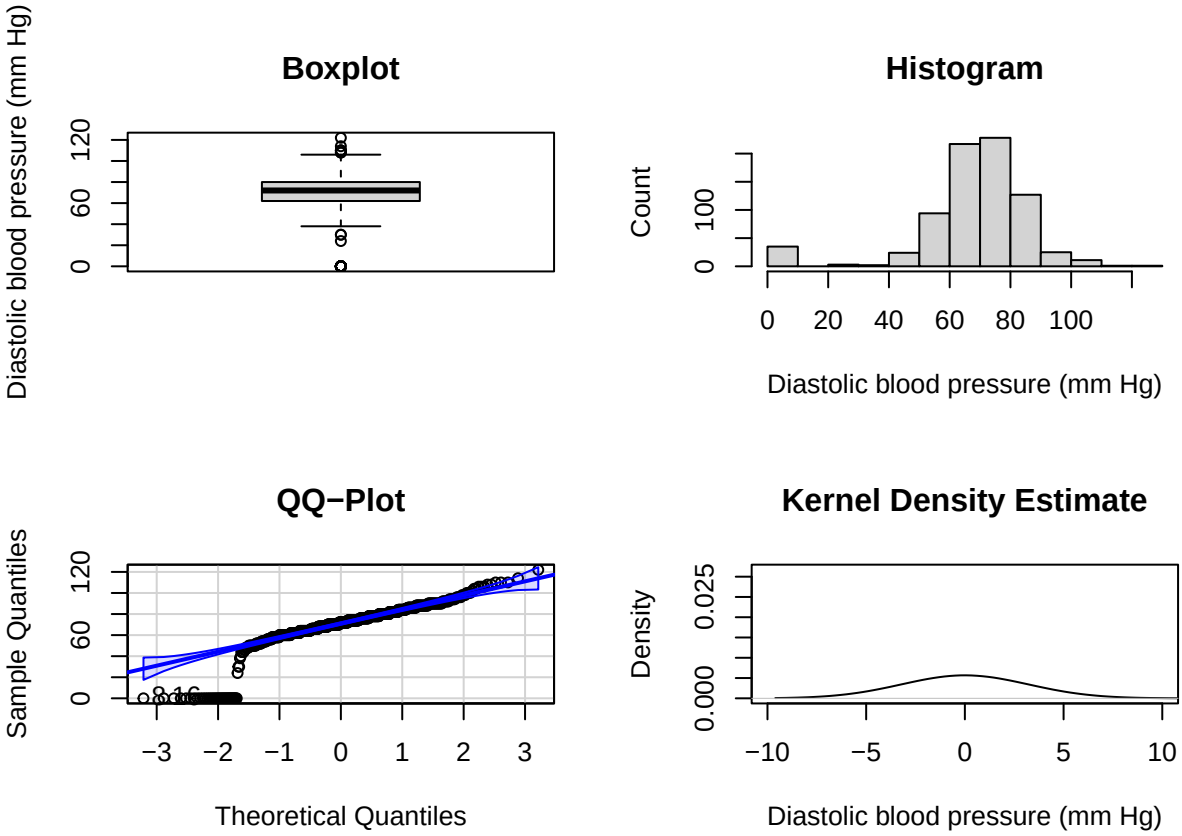

Boxplot
Number of Times Pregnant
0
5
15


Histogram
Count
0
150
350
0
5
10
15
Number of Times Pregnant

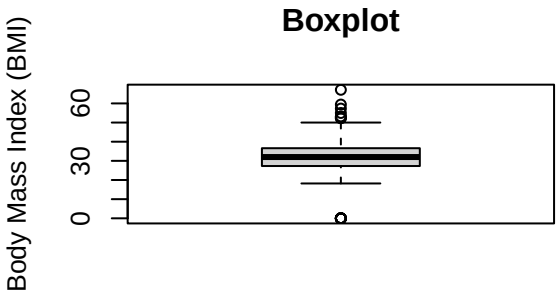

QQ−Plot
Sample Quantiles
0
5
15
160
−3
−2
−1
0
1
2
3
Theoretical Quantiles

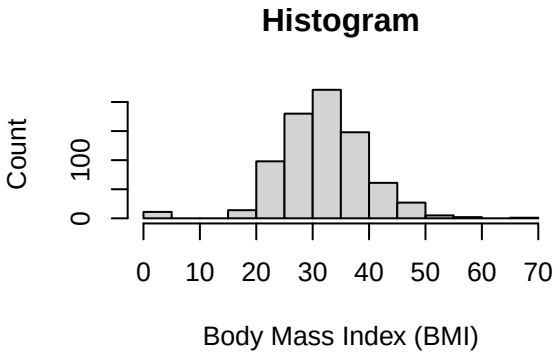

Kernel Density Estimate
Density
0.00
0.10
−5
0
5
10
15
20
Times Pregnant

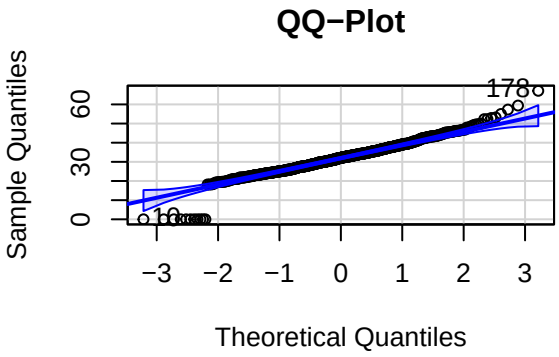
```
## [1]  8 16
```

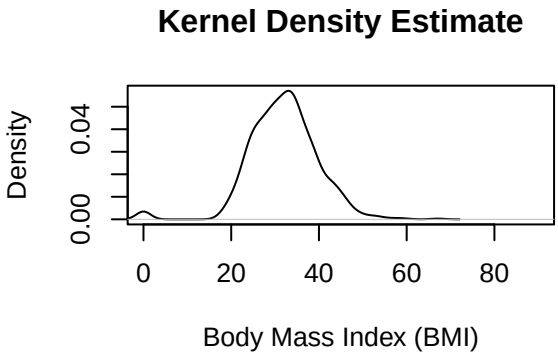

Boxplot
Diastolic blood pressure (mm Hg)
0
60
120
Histogram
Count
0
100
0
20
40
60
80
100
Diastolic blood pressure (mm Hg)
QQ−Plot
Sample Quantiles
0
60
120
−3
−2
−1
0
1
2
3
Theoretical Quantiles
Kernel Density Estimate
Density
0.000
0.025
−10
−5
0
5
10
Diastolic blood pressure (mm Hg)

```
## [1] 178  10
```

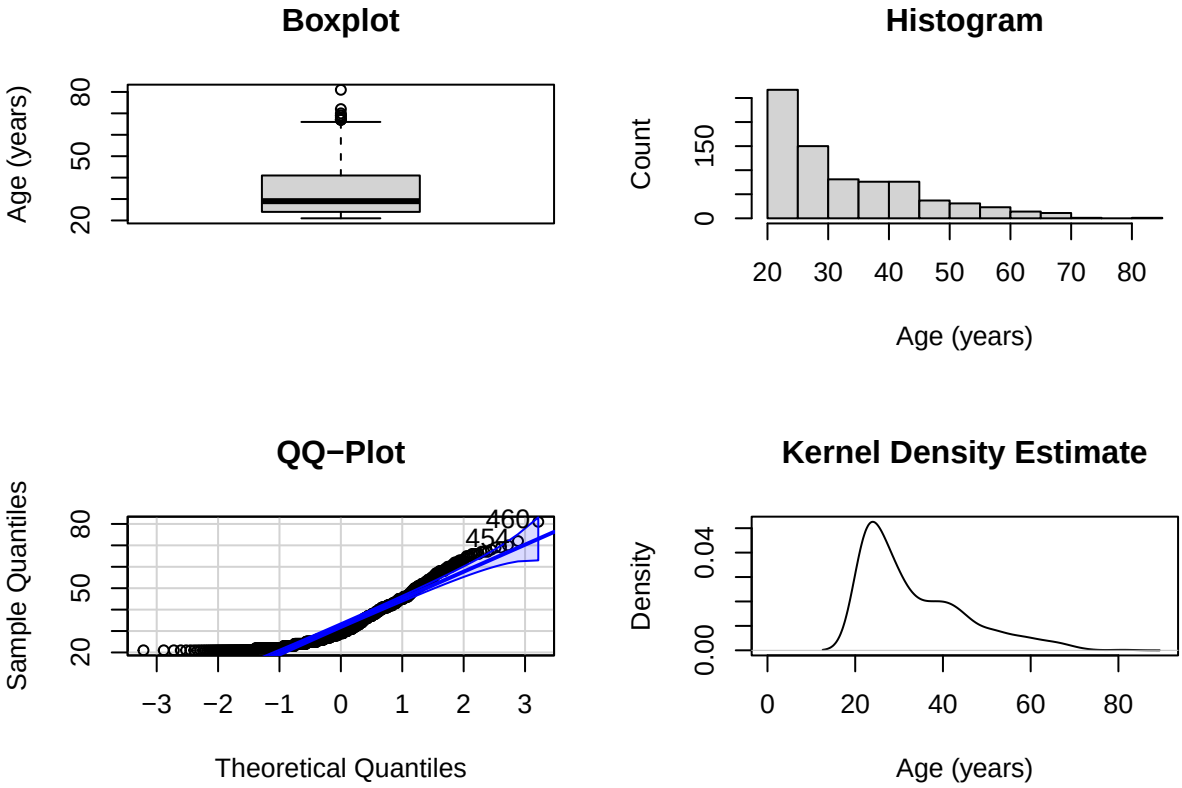

Boxplot
Body Mass Index (BMI)
0
30
60


Histogram
Count
0
100
0
10
20
30
40
50
60
70
Body Mass Index (BMI)


QQ−Plot
Sample Quantiles
0
30
60
178
−3
−2
−1
0
1
2
3
Theoretical Quantiles


Kernel Density Estimate
Density
0.00
0.04
0
20
40
60
80
Body Mass Index (BMI)

```
## [1] 460 454
```



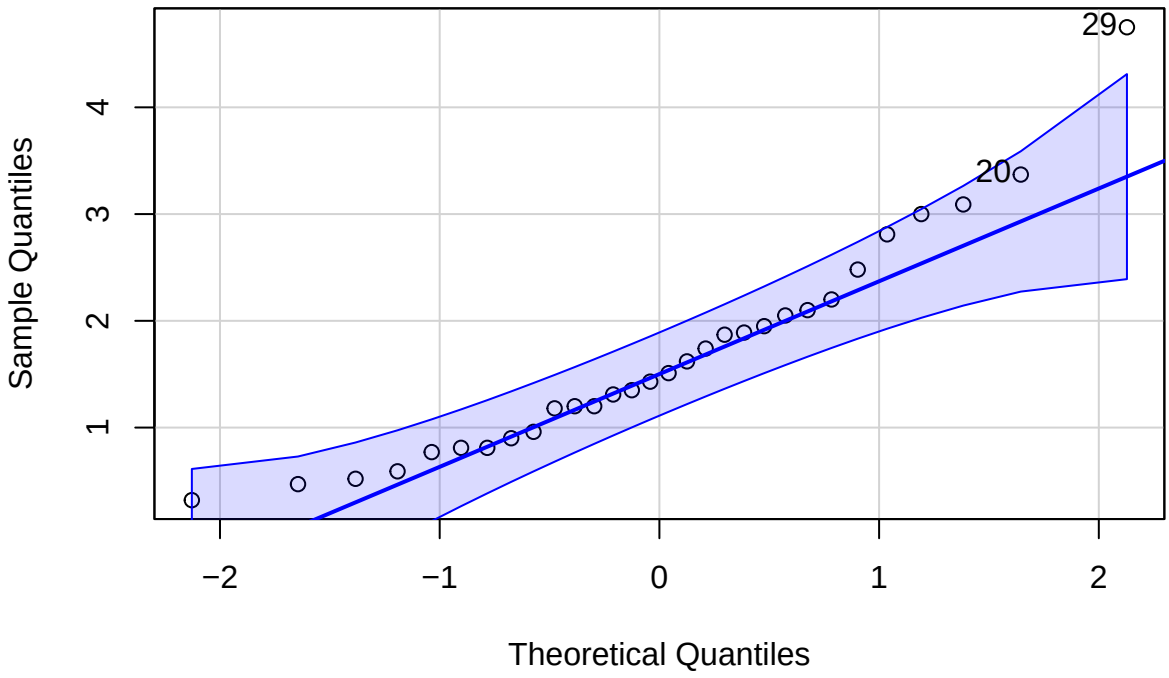
Pregnant: a few very high outliers, distribution appears to be Gamma

Blood Pressure: distribution looks normal except for outliers with extremely low bloodpressure

BMI: distribution also looks relatively normal, but again there are extreme outliers with very low BMI that fall outside the distribution

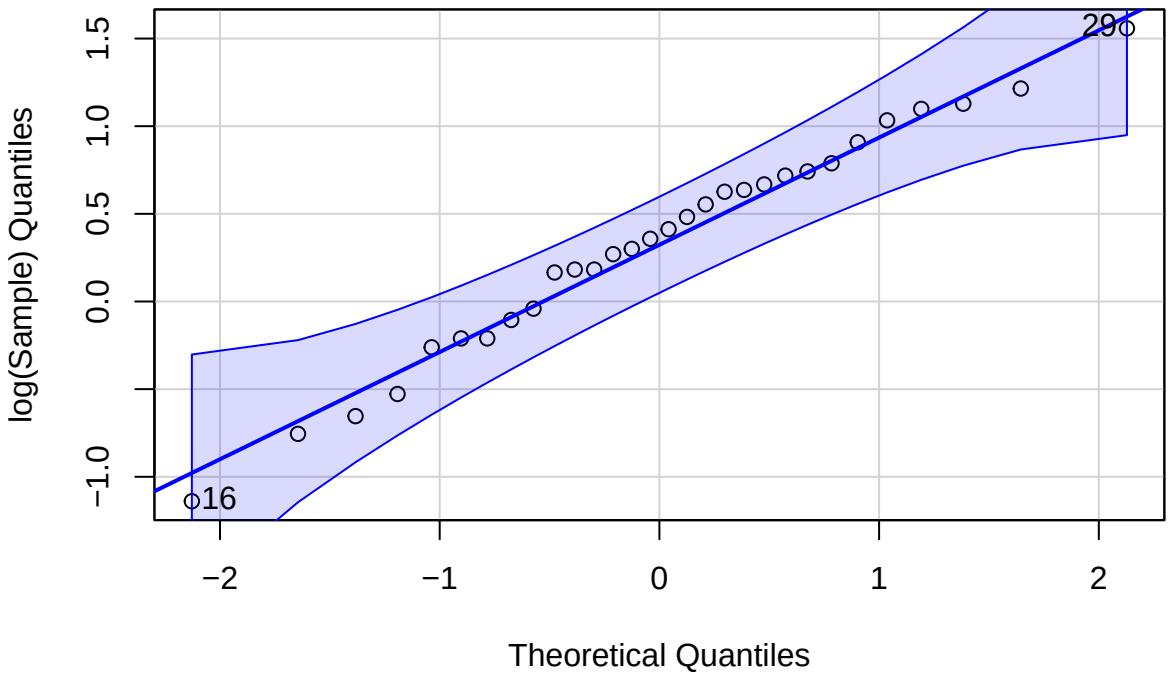
Age: looks like a gamma distribution, or exponential after year 18... Definitely not normal due to QQ-plot being highly non-linear

Question 3




```
## [1] 29 20
```

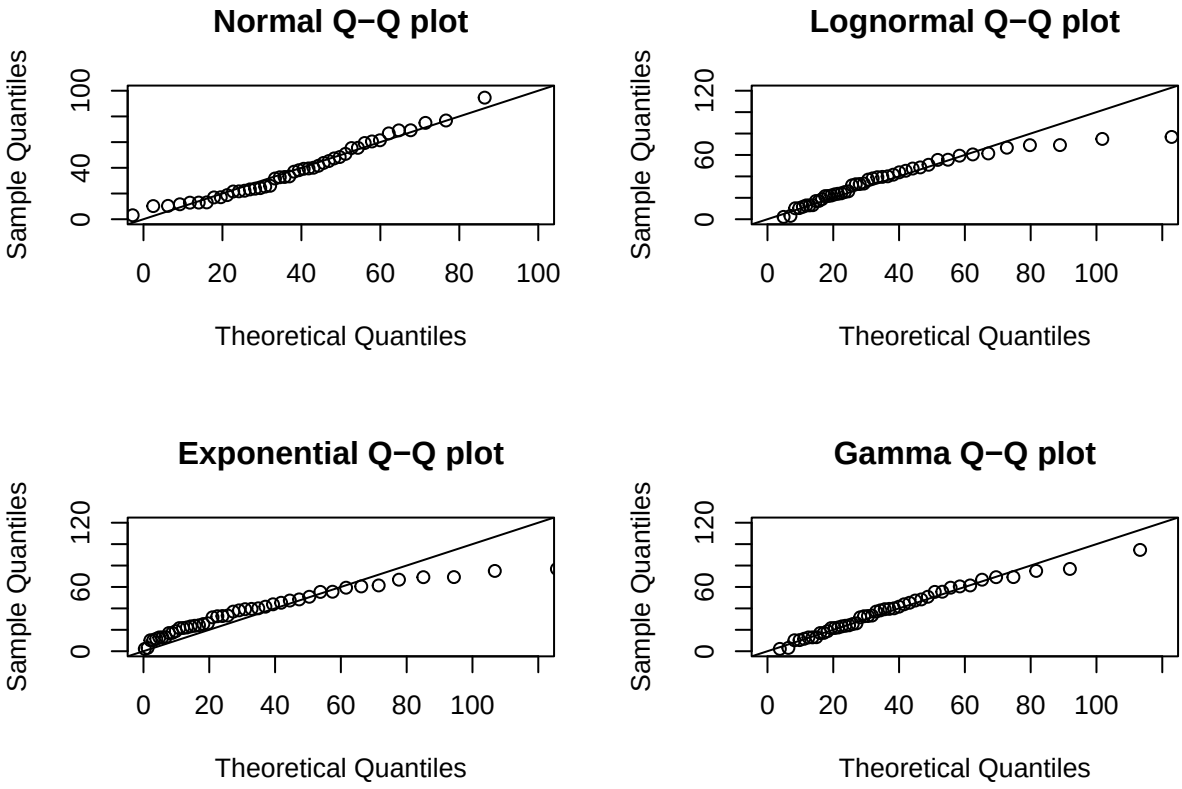
Looks farily normal, but tails are not linear. Try log(precip)...



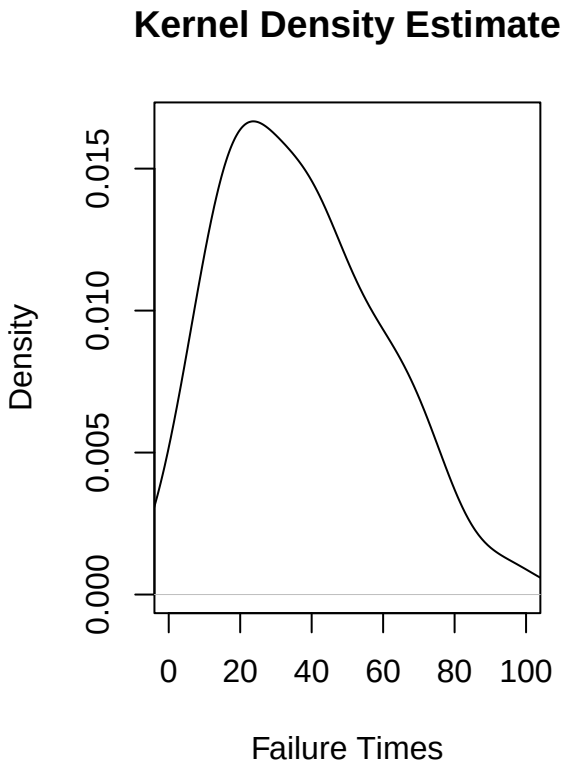

```
## [1] 16 29
```

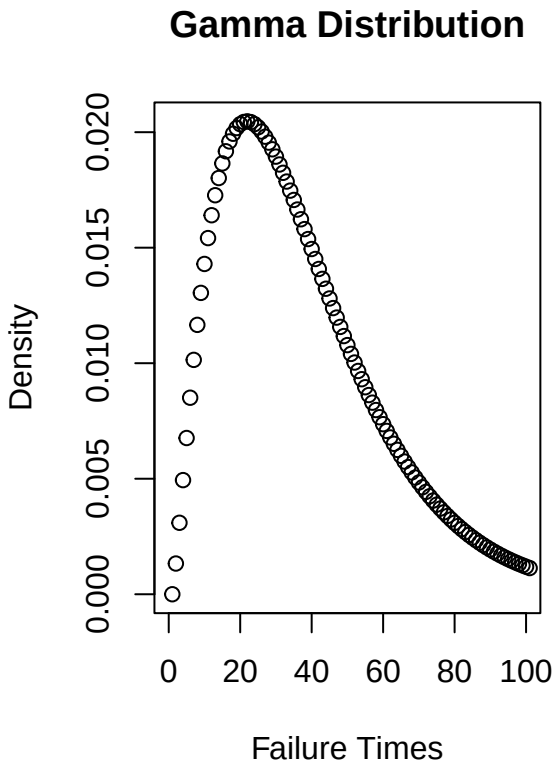
That looks like a normal distribution!

Question 4



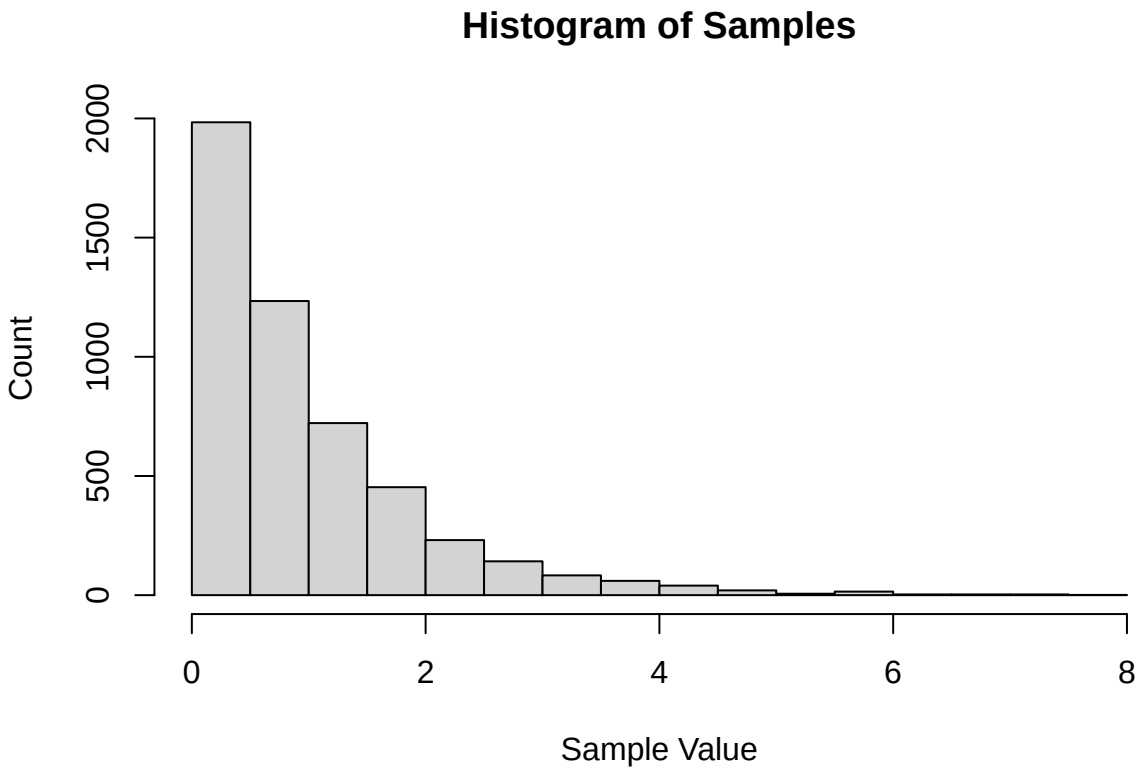
Based on the above plots, the Gamma distribution appears to fit best as the points closely follow the line, except for the last two outliers. The Normal distribution fits relatively well, and the exponential and lognormal are very poor fits.



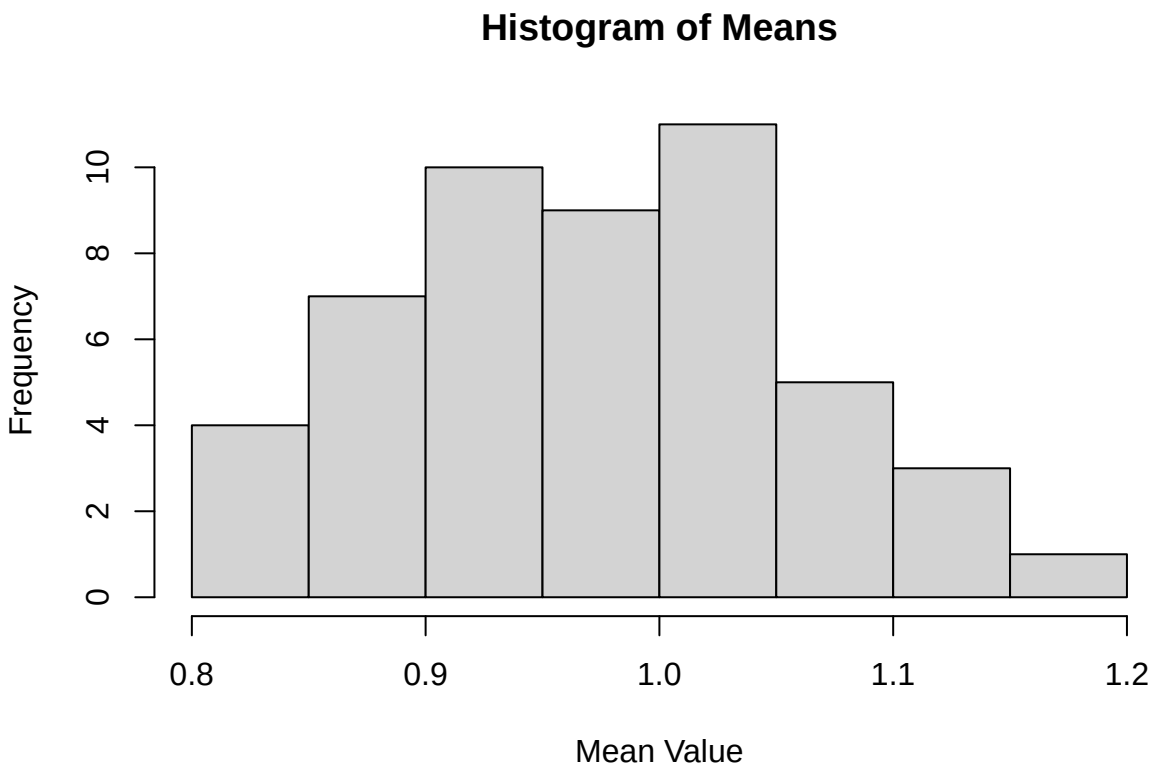


Gamma distribution with given paramters aligns quite well with KDE!

Question 5

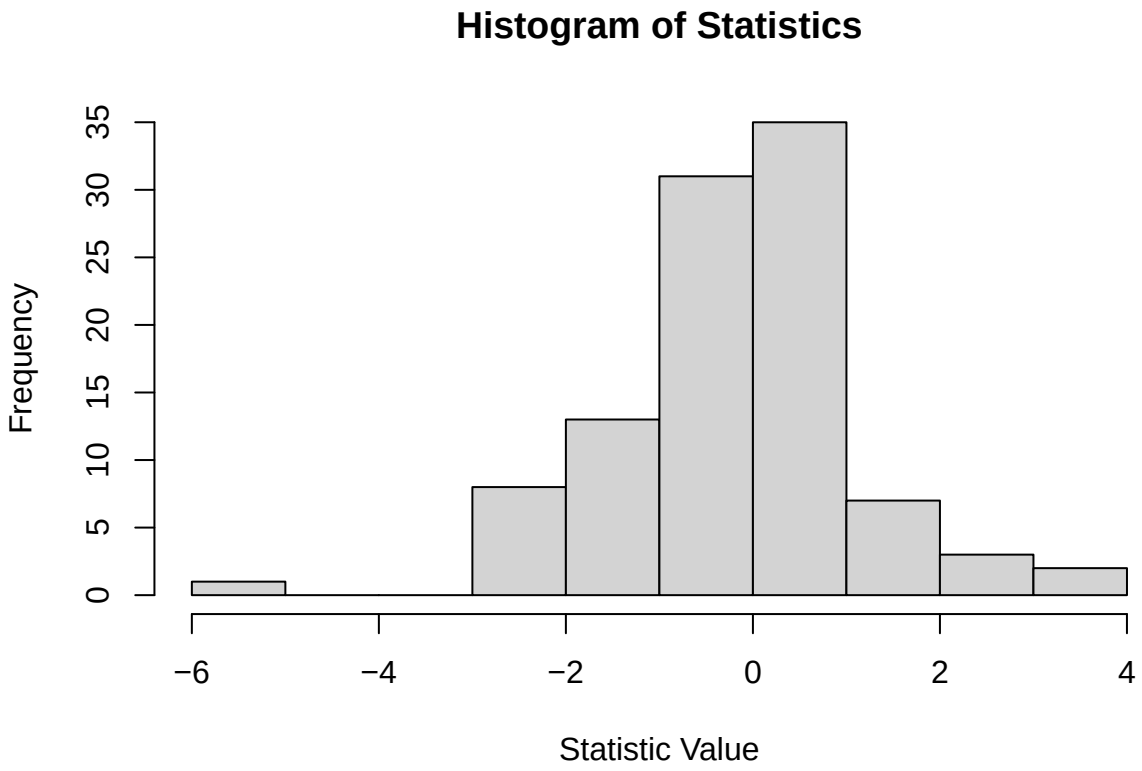


The shape drops of as x increases but has a long tail, as we would expect because the exponential function goes as $f(x) = e^{-x}$



They don't have the same shape, because the larger values that make up the tail of the exponential histogram will be averaged with the higher density of lower values, leaving a more normal distribution around 1.

Question 6



We would expect this statistic to be normally distributed about zero, which it is.