Statistical Inference Project Part 1

JN Gabra

3/31/2021

Overview

This report consists of a simulation exercise as well as basic inferential data analysis. An exponential distribution is simulated, and the distribution of the averages of 40 exponentials is observed and compared to a normal distribution.

Simulations

For this exercise we will run 1000 simulations of 40 exponential distributions (each). The rate of the exponentials (i.e. lambda) is set to 0.2.

```
lambda<-0.2

mns = {NULL}
vrs = {NULL}
for (i in 1:1000) {
      mns = c(mns,mean(rexp(40,lambda)))
      vrs = c(vrs,var(rexp(40,lambda)))
}
```

Sample Mean versus Theoretical Mean

For the exponential distributions with a rate = lambda = 0.2 the theoretical mean is 1/lambda.

```
mean_theor<-1/lambda
mean_sample<-mean(mns)
mean_sample
```

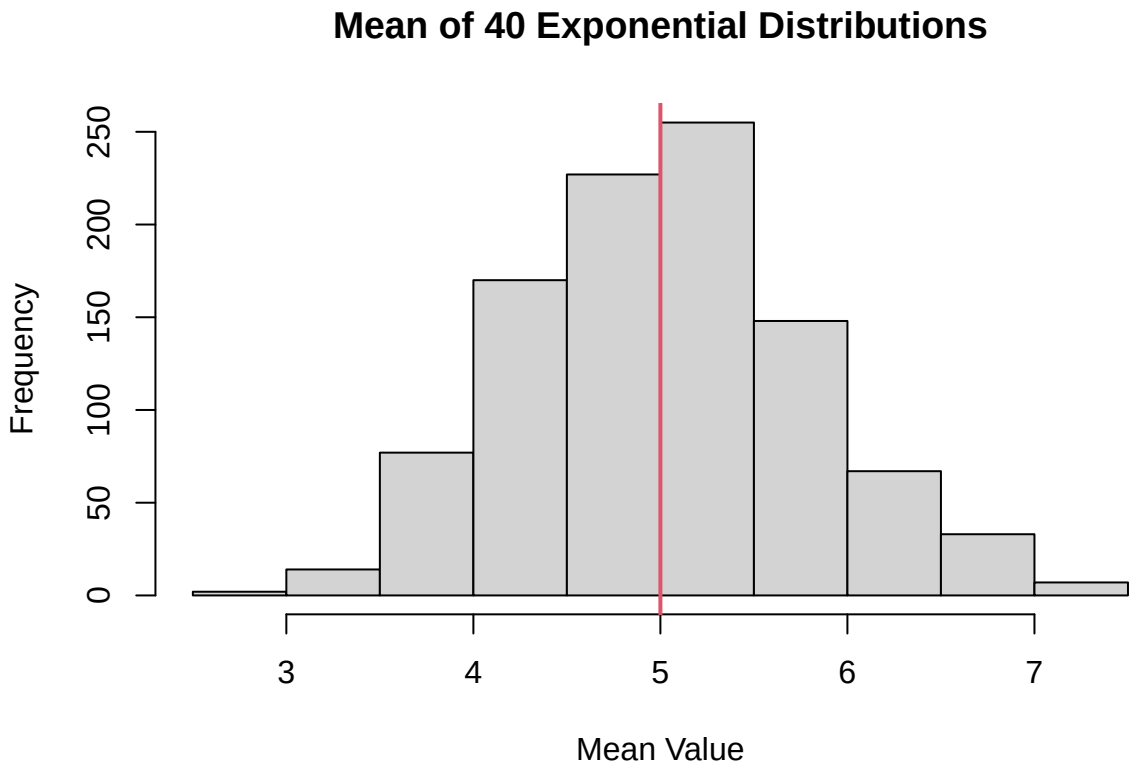
```
## [1] 5.024153
```

```
mean_theor
```

```
## [1] 5
```

We can see that the the sample mean is pretty close to the theoretical mean.

```
hist(mns,main="Mean of 40 Exponential Distributions",xlab="Mean Value")
abline(v=mean_theor,col=2,lwd=2)
```



In the plot above, we see a histogram of the 1000 means of the 40 exponential distributions. The theoretical mean is displayed as a vertical red line. We can see that the most common mean value of 40 exponentials is approximately 5.

Sample Variance versus Theoretical Variance

For the exponential distributions with a rate = lambda = 0.2 the theoretical standard deviation is 1/lambda. The variance is equivalent to the square of the standard deviation. Thus the theoretical variance is equal to 25 = 5^2.

```
sd_theor<-1/lambda
var_sample<-mean(vrs)
var_theor<-sd_theor^2
var_sample
```

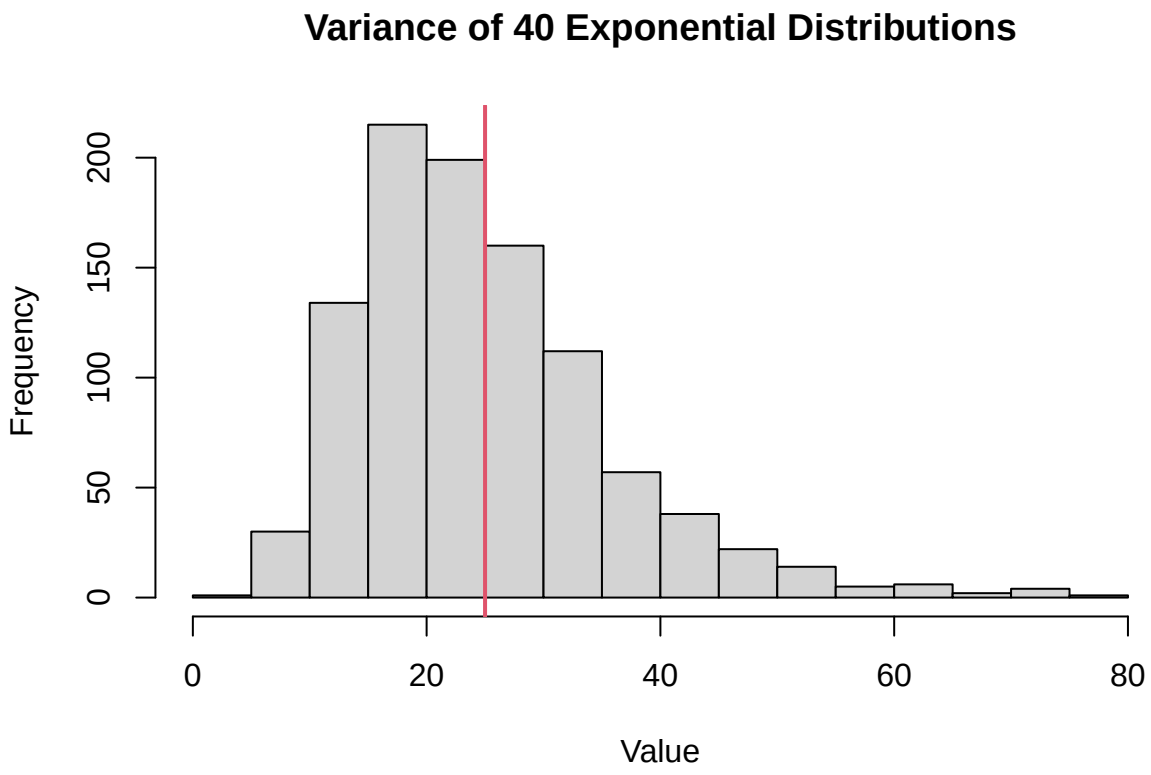
```
## [1] 24.82346
```

```
var_theor
```

```
## [1] 25
```

We can see that the sample variance is in close agreement to the theoretical variance of 25.

```
hist(vrs,main="Variance of 40 Exponential Distributions",xlab="Value")
abline(v=var_theor,col=2,lwd=2)
```

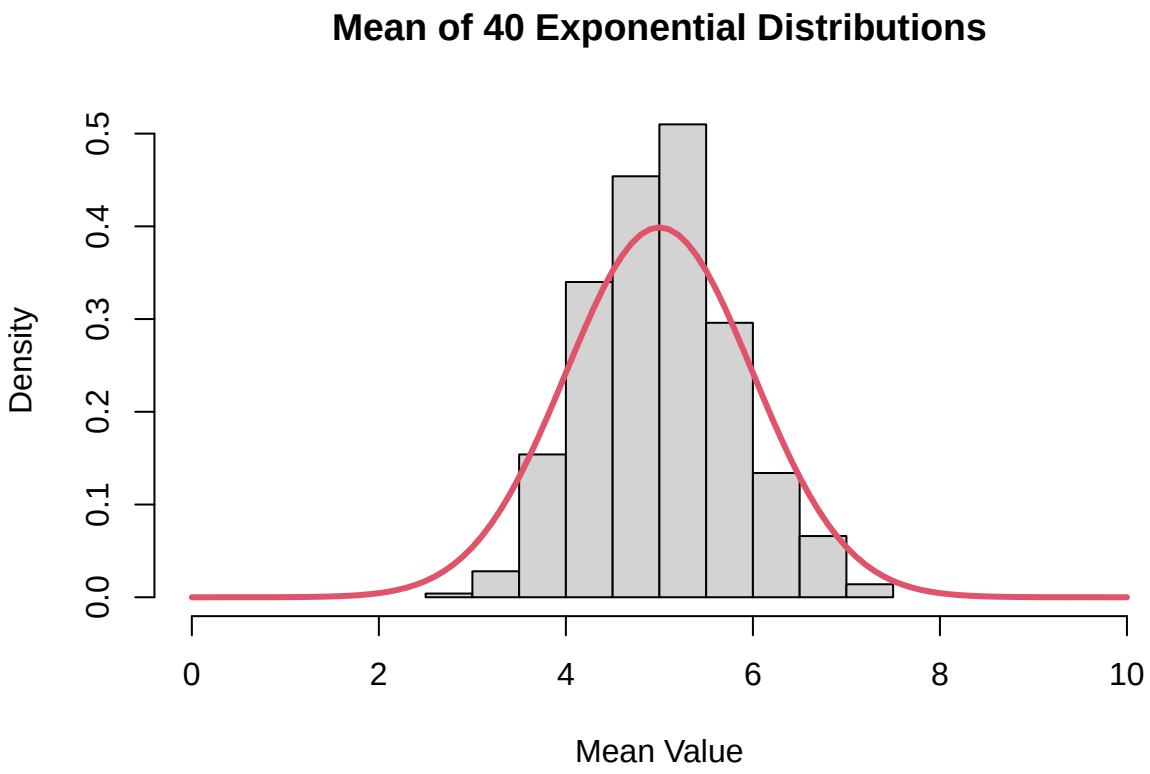


In the plot above, we see a histogram of the 1000 variances of the 40 exponential distributions. The theoretical variance is displayed as a vertical red line. We can see that the most common variance value of 40 exponentials is approximately 25.

Distribution

Below is the density histogram of the 1000 means of 40 exponentials. Overlayed in red is a normal distribution curve.

```
hist(mns,freq = FALSE,xlim=c(0,10),main="Mean of 40 Exponential Distributions",xlab="Mean Value")
curve(dnorm(x,mean=mean_theor,sd=1),add=TRUE,col=2,lwd=3)
```



We can see that the distribution of the means is approximately normal. The histogram of the mean values as approximately the same shape as a normal distribution curve centered at the theoretical mean.