Simulation Course Project Part 1

Libardo Lopez

Thursday, August 21, 2014

Illustrate via simulation and associated explanatory text the properties of the exponential distribution.

The mean of exponential distribution is 1/lambda and the standard deviation is also also 1/lambda.

Set lambda = 0.2 for all of the simulations.

If you like to see de R code, please Go to Github (https://github.com/Libardo1/Statistical_Inference_Course_Project)

Initialization:

```
nosim   <- 1000      # number of simulations
n      <- 40         # sample size of 40 as requested
lambda <- 0.2
mu     <- 1/lambda  # mu theoretical population mean = 5
s      <- mu         # s theoretical population standard deviation = 5
SE     <- s/sqrt(n) # SE is the theoretical standard error
set.seed(7890)
```

The simulated samples are stored in a matrix, where each row contains one sample of 40 random exponential variables.
The mean and standard-deviation were calculated for each sample, and stored as vectors - *meanx*, and *sdx* :

```
x<-matrix(rexp(n*nosim,lambda),nosim)
meanx<-apply(x, 1, mean)             #means of all samples
sdx<-apply(x, 1, sd)                 #std-deviations of all samples
```

1. Showing where the distribution is centered at, and comparing it to the theoretical center of the distribution:

The **average mean** of all the samples simulated is 4.997 compared with the theoretical mean of the exponential distribution, which is $\mu = 1/\lambda = 5$

2. Show how variable it is and compare it to the theoretical variance of the distribution.

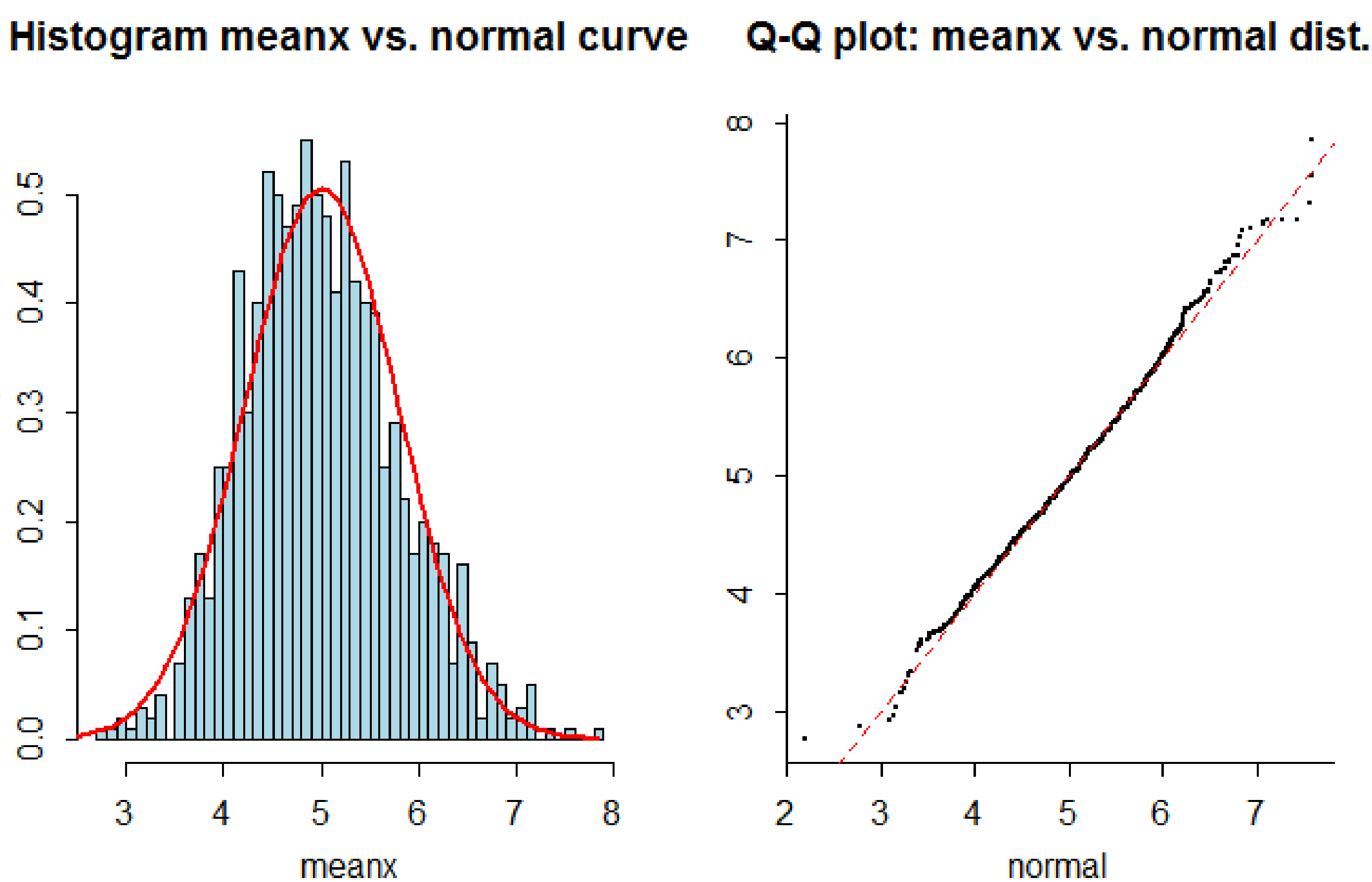
The **variance** of all the samples simulated means is 0.6176 compared with the theoretical variance of the exponential distribution, for sample-size of *n*, which is $s^2 = \sigma^2/n = (1/\lambda)^2/n = 5^2/40 = 0.625$

3. Show that the distribution is approximately normal.

With the histogram of the means simulated, compared to the density function curve of normal distribution of $\mu = 1/\lambda$ and $\sigma = 1/\lambda/\sqrt{n}$ and the Q-Q plot of the sample means compared to random normal sample with the same expected parameters, we confirm the **normalized** distribution.

```
par (mfrow = c(1,2), mar=c(5,2,4,2), mgp=c(2,.7,0), bty="l", tck=-.02)
hist(meanx, breaks=40, col="lightblue", freq=F, main="Histogram meanx vs. normal curve")
xaxis<-seq(0,max(meanx),length=100) #prepare the x-axis
lines(xaxis,
      dnorm(xaxis,1/lambda,sd=1/lambda/sqrt(n)),
      type="l", col="red", lwd=2)  #add ref normal curve

qqplot(rnorm(1000,1/lambda,sd=1/lambda/sqrt(n)),meanx, pch=20, cex = .5, col="black", main="Q-Q
 plot: meanx vs. normal dist.", xlab="normal")
abline(0,1, col = "red",lwd=1,lty=2)
```



4. Evaluation of the coverage of the confidence interval for $1/\lambda$ using $\overline{X} \pm 1.96 * s/\sqrt{n}$:

First at all, figure the interval's lower and upper limit for each sample:

```
ul<-meanx+sdx*1.96/sqrt(n) #upper limit vector
ll<-meanx-sdx*1.96/sqrt(n) #lower limit vector
```

And evaluate the **good intervals**, which contain the *mean=5*, with my settings, **the coverage** is 92.5%.