Simulation of the exponential distribution

Ariane

Saturday, September 20, 2014

We simulate the exponential distribution. The rate parameter λ is set to 0.2. The mean of the exponential distribution is then $1/\lambda = 5$ and the standard deviation of the exponential distribution is also $1/\lambda = 5$. We investigate the distribution of the mean of 40 exponentials. To carry out this analysis, we need to do a thousand simulated data for each exponential distribution.

```
# Data
lambda <- 0.2 # rate parameter
NS <- 1000 # number of exponentials
ND <- 40 # number of data for each exponential distribution
NB <- 10 # number of breaks to draw the histograms
#
# Simulations of the 40 exponentials
data <- matrix(nr=ND, nc=NS)
for (i in 1:NS){ data[,i] <- rexp(ND, lambda) }
#
# Computation of the average
m <- rep(0, NS)
for (i in 1:NS){ m[i] <- mean(data[,i]) }
```

The theoretical center of the distribution of the average of 40 exponentials is $1/\lambda = 5$. The theoretical variance of the distribution is $(1/\lambda)^2/ND$ where ND = 40 is the number of simulations for each exponential. The theoretical variance is equal to 0.625.

```
mth <- 1/lambda
vth <- (1/lambda)^2 / ND
pth <- 0.95
```

We compute the empirical value of the mean, the variance and the 95% confidence interval of the average of the 40 exponentials.

```
ms <- mean(m)
vs <- var(m)
ps <- length(m[m >= mth - 1.96 * sqrt(vth) & m <= mth + 1.96 * sqrt(vth)]) / NS
```

The empirical value of the mean of the average of 40 exponential distributions is 4.9916 and is close to the theoretical value 5. The empirical value of the variance is 0.6048 and is close to the theoretical value 0.625. The empirical value of the confidence interval is 0.957 and is close to the theoretical value 0.95.

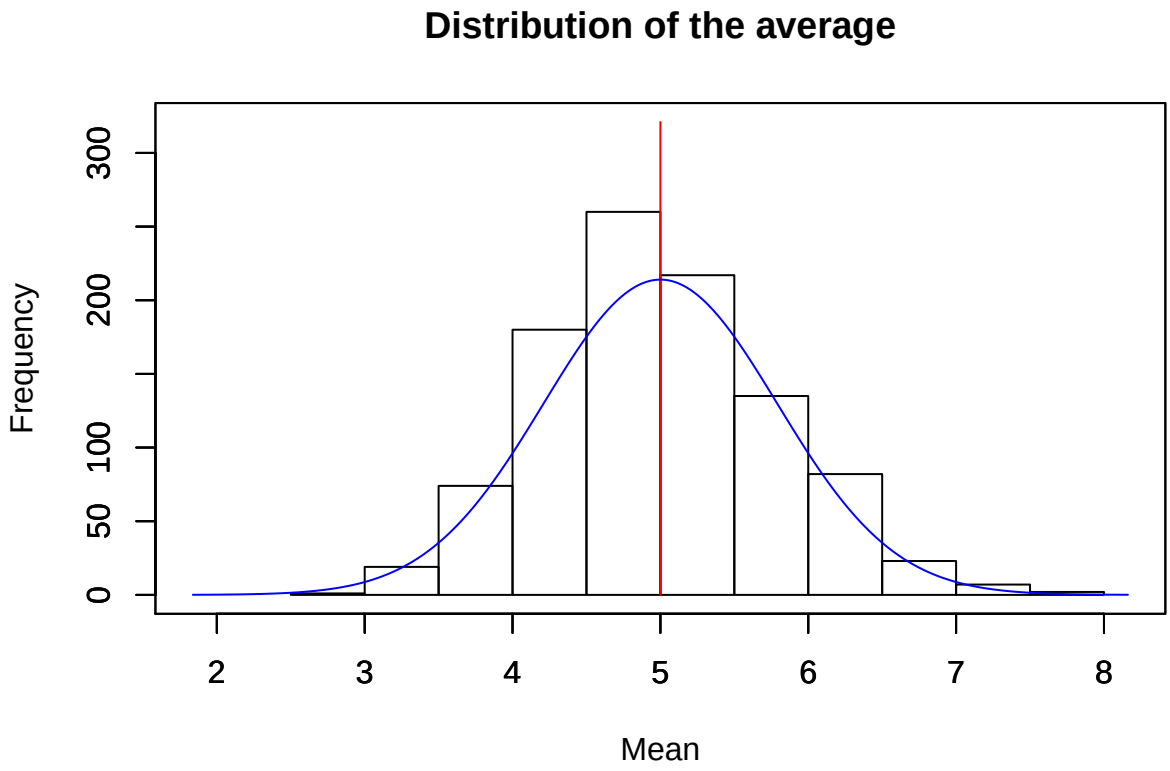
To compare the distribution of the average of the 40 exponentials to a normal distribution, we compute the values of a normal distribution with mean equal to the theoretical mean of the average of the 40 exponentials and variance equal to the theoretical variance of the average.

```
sample <- sqrt(vth) * 0.01 * c(-400:400) + mth
dsample <- dnorm(0.01 * c(-400:400))
```

We set the scale of the normal distribution at the same scale of the histogram of the average of the 40 exponentials.

```
norm_hist <- hist(rnorm(NS), breaks=NB, plot=FALSE)
max_density <- max(norm_hist$counts) / max(dsample)
```

We draw an histogram of the average of the 40 exponentials.

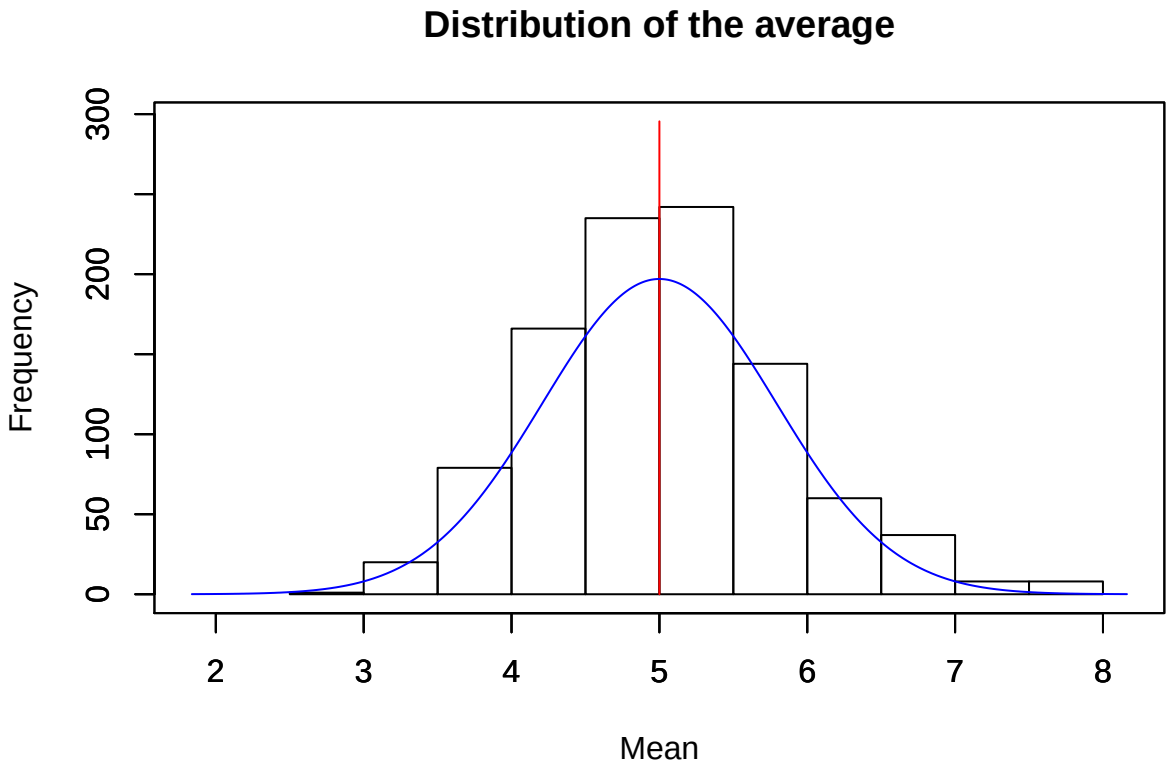


The red bar is the theoretical mean of the average of the 40 exponentials. We can see that the histogram of the average is nearly equally distributed on both sides of the theoretical mean.

The blue line is the normal distribution. We can see that the histogram of the average of the 40 exponentials is close to the normal distribution.

Appendix

We now do the same analysis with the average of 400 exponential distributions and draw the new histogram.



The new empirical value of the mean is 5.0283, the new empirical value of the variance is 0.6858 and the new empirical value of the confidence interval is 0.934. These new values are close to the theoretical values. The histogram of the average of the 400 exponentials is also close to the normal distribution.