Exploring Variances of the Exponential Distribution

ww44ss

October 2014

Synopsis

This analysis shows how the mean and standard deviation of a finite random sample of a uniform distribution converge to population mean with increasing sample size.

Here is a sample of the data.

```
##create data for analysis

        ##Set seed to make results reproducible
        set.seed(8675309)

        ##makes sample sizes variable
        rowsofsamples <- 42
        columnsofdistributions <- 1003

        ##set lambda
        lambda <- .2

        ##create data frame using slightly different method than in class
        ##c rows is the sample index and column indexes the distribution.
        rdata <- replicate(columnsofdistributions, rexp(rowsofsamples, lambda))

        ##look at structure of data
        rdata[1:10, 1:11]
```

```
##         [,1]    [,2]    [,3]    [,4]   [,5]    [,6]    [,7]    [,8]
##  [1,] 8.3108 2.21911  5.7452  4.9706 2.3460  9.1790 14.8307  9.9862
##  [2,] 6.1163 0.00644  8.4450  0.8119 6.4402  3.9995  6.2526  1.9988
##  [3,] 3.8367 6.65512  6.6883  1.8342 3.7670  0.7838  0.4101  0.6106
##  [4,] 1.7305 2.43901  2.8348  7.4282 2.0314  0.6605  2.4097  3.5832
##  [5,] 1.5428 0.25482 25.2812  7.0943 4.5731  0.1878  2.1739 10.2474
##  [6,] 3.3823 1.30645  4.1566  1.0877 0.2332  5.2007 21.5939  2.0287
##  [7,] 0.8332 6.50322  9.0648 17.1492 8.1698  2.7127  0.3139  1.7888
##  [8,] 0.1095 6.03537  2.6453  4.4358 5.9621 27.4551  0.2918  3.4522
##  [9,] 3.6691 4.37594  0.6022  5.3980 6.1676  5.1189 12.5936  2.1047
## [10,] 2.4949 5.19336  3.4753  1.4396 0.4709 10.0943  0.3898  3.4128
##          [,9]  [,10]   [,11]
##  [1,]  1.5576 3.2057 10.7935
##  [2,]  0.4156 6.1352  2.1928
##  [3,]  2.6618 1.6070 10.4130
##  [4,]  5.1680 0.4647  1.1694
##  [5,]  7.3741 1.4501  0.8454
##  [6,] 13.7490 3.1437  4.0568
##  [7,]  3.4455 3.7691  8.5890
##  [8,]  7.7237 1.7594  0.6331
##  [9,]  6.7943 7.9235 18.5087
## [10,]  2.3633 1.5877  2.9955
```

This analysis looks at a set of 1003 distributions of 42 samples.

Question1: Show where the distribution is centered and compare it to the theoretical center of the distribution.

```
##Take means of columns and plot histrogram with mean and theoretical mean

        ##Compute Column means
        expmeans <- colMeans(rdata)

        ##Calculate "Mean of Means"
        meanofdistmeans<-mean(expmeans)
```

The mean of the distribution 5.0068 differs by 0.1364% from the theoretical population mean of 5.0

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

```
        ## Calculate the Variance
        varofdistmeans = var(expmeans)
```

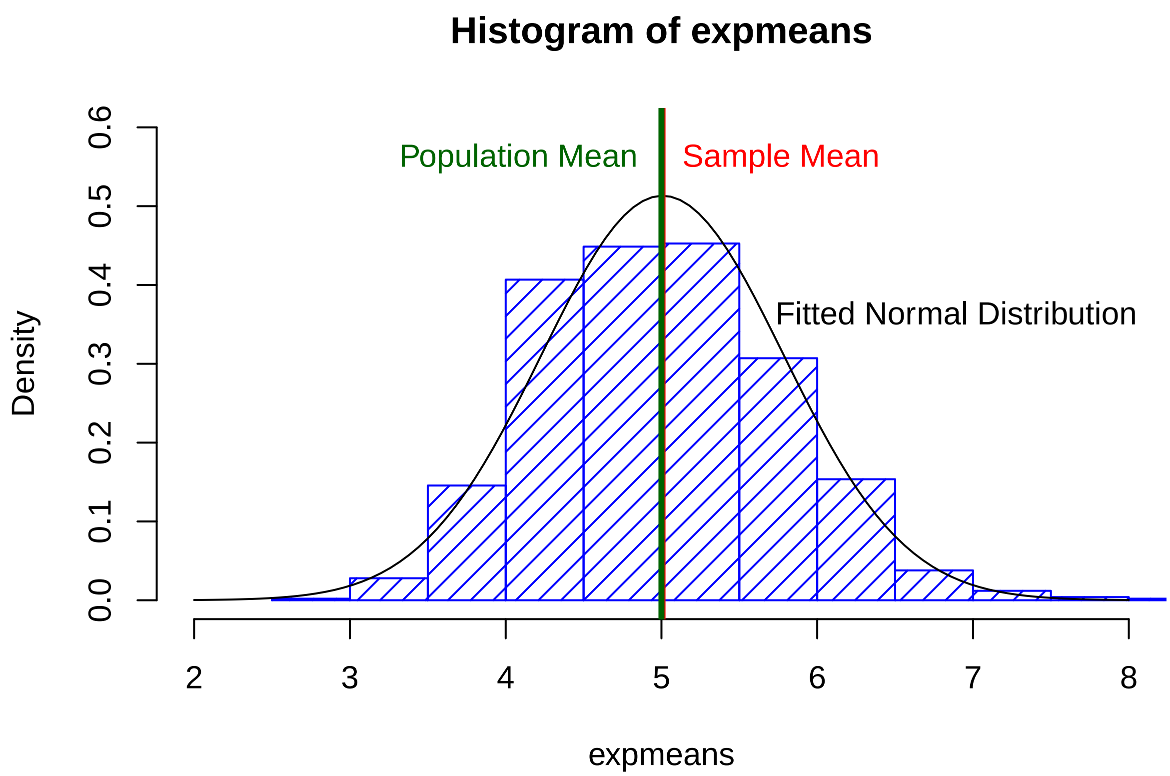
The theoretical variance of the distribution is 0.2 versus the variance of the data 0.6042, which is 3.0208 times larger.

Distribution: Show that the distribution is approximately normal.

```
        ##Generate Historgram
        hist(expmeans, prob=TRUE, density=12, angle=45, xlim=c(2,8), ylim=c(0,0.6),col="blue")
        curve(dnorm(x, mean=mean(expmeans), sd=sd(expmeans)), add=TRUE)


        ##Draw lines at mean and theoretical mean = 5
        abline(v=meanofdistmeans, lwd=3, col="red")
        abline(v=5, lwd=3, col="darkgreen")

        ##Label means on graph
        text(meanofdistmeans,0.55, "Sample Mean", col="red", adj=c(-0.1,0))
        text(5,0.55, "Population Mean", col="darkgreen", adj=c(1.10,0))
        text(5.5,0.35, "Fitted Normal Distribution", col="black", adj=c(-.10,0))
```




```
###Draw Std Deviation Lines
```

Above is a plot of a histogram of the exponentials. The sample mean of the data, shown in the graph by the red vertical line at 5.0068 is barely distinguishable from the popultion mean at 5.000. The fit of the normal distribution, from the sample mean and standard deviation 0.7773 follows the data well.

Convergence: Evaluate the coverage of the confidence interval for 1/lambda

```
a<-1.0
```