Exponential Inference Simulation

Author: Byron.Bian

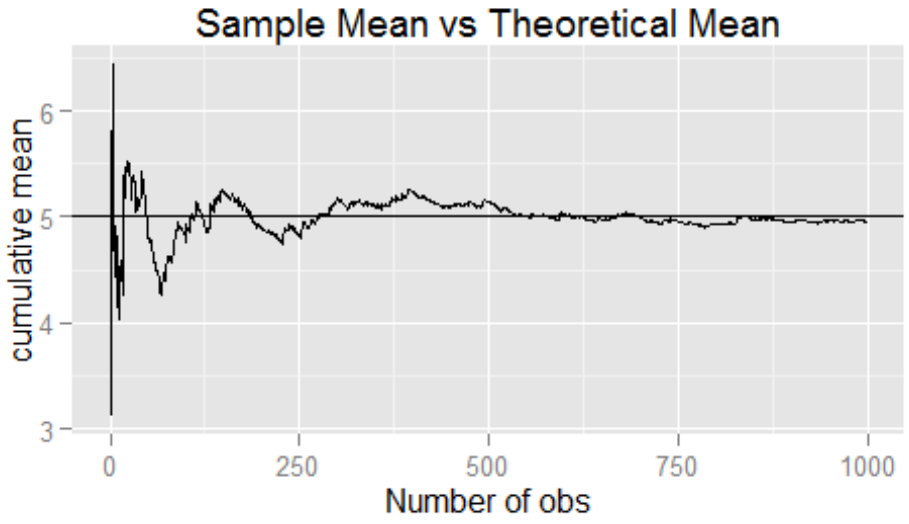
Date: Sunday, May 17, 2015

Overview:
In this report, I will verify the theorem of CLT by researching three simulation test for exponential inference. The first one is about the sample mean versus theoretical mean, second one is for sample variance (standard error) versus theoretical variance and the last one is for asymptotoic distribution of sample mean.

Simulation I: Sample Mean versus Theoretical Mean

In this simulation, i will compare the sample mean with theoretical mean (1/lambda) along with the accruement of sample size from 1 to 1000.

```
#simulation sample size is 1000, lambda is 0.2
n<-1000
lambda<-0.2
library(ggplot2)
samp.mean<-cumsum(rexp(n,lambda))/(1:n)
g<-ggplot(data.frame(x=1:n,y=samp.mean),aes(x=x,y=y),ylim=c(1,7))
g<-g+geom_hline(yintercept=1/lambda)+geom_line(size=.5)
g<-g+labs(x="Number of obs",y="cumulative mean",title="Sample Mean v
s Theoretical Mean")
print(g)
```

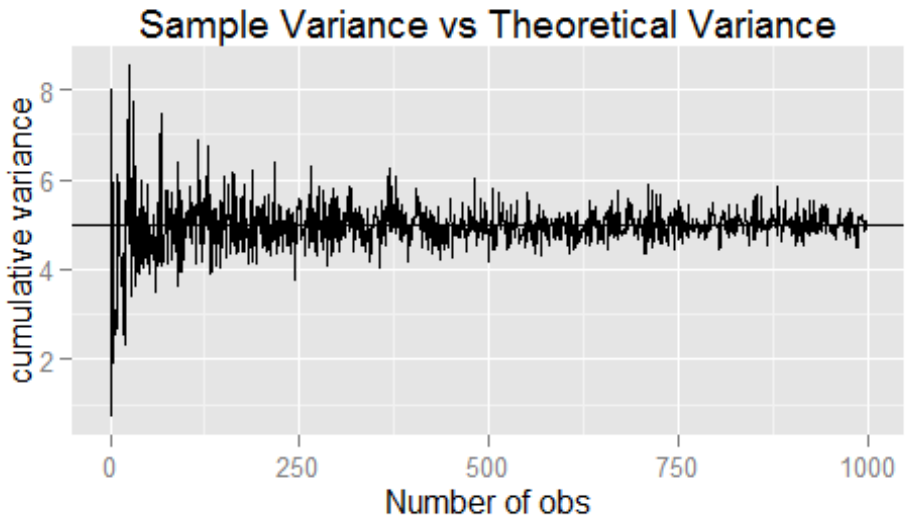


Conclusion: Along with the accruement of sample size, the sample mean is becoming the asymptotic value of theoretical mean (1/lamba)

Simulation II: Sample Variance versus Theoretical Variance

In this simulation, i will compare the sample variance with theoretical variance (1/lambda) along with the accruement of sample size from 1 to 1000.

```
#simulation sample size is 1000, lambda is 0.2
n<-1000
lambda<-0.2
library(ggplot2)
set.seed(100)
samp.var<-NULL
for (i in 2:(n+1)) samp.var<-c(samp.var,sd(rexp(i,lambda)))
g<-ggplot(data.frame(x=1:n,y=samp.var),aes(x=x,y=y))
g<-g+geom_hline(yintercept=1/lambda)+geom_line(size=.5)
g<-g+labs(x="Number of obs",y="cumulative variance",title="Sample Va
riance vs Theoretical Variance")
print(g)
```



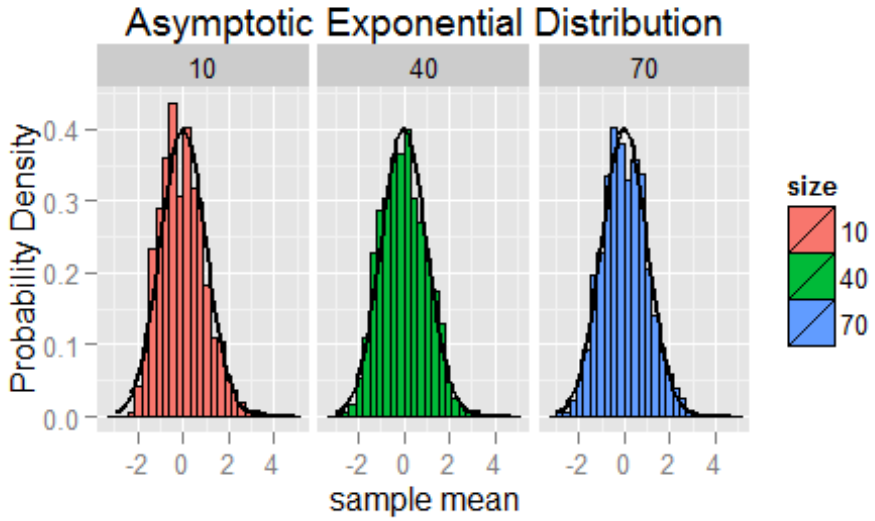
Conclusion: Along with the accruement of sample size, the sample variance (standard error) is becoming the asymptotic value of theoretical variance (1/lamba)

Simulation III: Asymptotic Distribution

In this simulation, i will research the asymptotic distribution of exponennt, the sample size is 40 and for convenient comparision with other sample sizes, I choose alternative sample sizes of 20 and 30. The simulation will be tried 1000 times

```
nosim<-1000
lambda<-0.2
cfunc <- function(x, n) (mean(x)-(1/lambda))/((1/lambda)/sqrt(n))
dat <- data.frame(
```

```
  x = c(apply(matrix(rexp(nosim * 10, rate=lambda), nosim), 1, cfunc,
 10),
        apply(matrix(rexp(nosim * 40, rate=lambda), nosim), 1, cfunc,
 40),
        apply(matrix(rexp(nosim * 70, rate=lambda), nosim), 1, cfunc,
 70)),
      size = factor(rep(c(10, 40, 70), rep(nosim, 3))))
g <- ggplot(dat, aes(x = x, fill = size)) +
     geom_histogram(binwidth=.3, colour = "black", aes(y = ..densit
y..)) +
     labs(x="sample mean",y="Probability Density",title="Asymptotic
Exponential Distribution")
g <- g + stat_function(fun = dnorm, size = 1)
g + facet_grid(. ~ size)
```



Conclusion: Along with the accruement of sample size (10->40->70), the distributions of sample mean are approximately normal more and more.