Appendix Course Project - Statistical Inference

Candice MH

Sunday, November 09, 2014

This appendix documents the code required to reproduce the analysis for the course project.

## Simulation Code

lambda<-0.2  
samplesize<-40  
nreps<-1000  
samplemeans<-rep(0,nreps)  
  
for (i in 1:nreps) {  
 tempsample<-rexp(samplesize,lambda)  
 samplemeans[i]<-mean(tempsample)  
}

## Distribution centrality

# Calculate

#sample (mean)  
mean(samplemeans)  
#theory (mean)  
theo\_mean<-1/lambda  
theo\_mean

# Plot densities

hist(samplemeans,breaks=20,col="red",freq=FALSE,xlab="Sample means",main="Density of sample means (red),\n with theoretical density (blue)")  
xfit<-seq(min(samplemeans),max(samplemeans),length=samplesize)   
yfit<-dnorm(xfit,mean=1/lambda,sd=1/(lambda\*sqrt(samplesize)))   
lines(xfit, yfit, col="blue", lwd=2)

## Distribution variability

# calculate

#sample (variability about mean)  
var(samplemeans)  
#theory (variability about mean)  
theo\_var\_about\_mean<-1/(samplesize\*lambda^2)  
theo\_var\_about\_mean

## Distribution normality

# Plot qq plot

qqnorm(samplemeans,ylab="Sample means from simulation"); qqline(samplemeans, col = 2)

# Plot densities

sample2<-rexp(nreps,lambda)  
y<-c(samplemeans,sample2)  
x<-c(rep("Mean of sample",nreps),rep("Single sample",nreps))  
xy\_df<-data.frame(x,y)  
library(lattice)  
histogram(~ y | as.factor(x),xlab="Sample value",main="Density of means of many samples (left)\n and a single sample (right)")