CIND 123
Lab 9

Sampling distribution

- Create a vector containing the same population as the lecture c(5, 10, 15).
- Install the package "gtools" in R.
- Use the premutations function in gtools to list all possible outcomes when 2 numbers are selected with replacement. Hint: Set the argument. repeats.allowed =TRUE in the premutations function
- Use the rowMeans() function to calculate the mean of each possible pair of outcomes.
- Is the mean of the original population == the mean of the row means?
- Use the hist() function to draw a histogram of the pairs' means.

Central Limit Theorem Simulation[1]

Step 1:
Open a new script in R and enter the following function:

```
sdm.sim <- function(n,src.dist=NULL,param1=NULL,param2=NULL) {
 r <- 10000  # Number of replications/samples
 # This produces a matrix of observations with
 # n columns and r rows. Each row is one sample:
 my.samples <- switch(src.dist,
             "N" = matrix(rnorm(n*r,param1,param2),r),
             "P" = matrix(rpois(n*r,param1),r),)
 all.sample.sums <- apply(my.samples,1,sum)
 all.sample.means <- apply(my.samples,1,mean)
 all.sample.vars <- apply(my.samples,1,var)
 par(mfrow=c(2,2))
 par(mar=c(1,1,1,1))
 opar=par(ps=6) # Make text 18 point
 hist(my.samples[1,],col="gray",main="Distribution of One Sample")
 hist(all.sample.sums,col="gray",main="Sampling Distribution of
        the Sum",ps=10)
 hist(all.sample.means,col="gray",main="Sampling Distribution of the Mean",ps=10)
 hist(all.sample.vars,col="gray",main="Sampling Distribution of
        the Variance",cex=.8)
}
```

[1] http://www.r-bloggers.com/sampling-distributions-and-central-limit-theorem-in-r/

Examine the function in the script, most of the functions we have already covered and encountered.

Run the function and make sure you can see it in the memory.

Step 2 (Normal Distribution):

1- On the R console: call the function as follows:

```
>sdm.sim(5, src.dist="N", param1=20, param2=3)
```

2- Change the sample size from 5 -10 and call the function.
3- Change the sample size from 10 -20 and call the function.
4- Continue incrementing the sample size by 10 size till you reach 80.
Record any key findings.

Step 3 (Poisson distribution):

Study the function arguments and call the function for the Poisson distribution instead of the normal distribution. You need to change the src.dist argument for Poisson.

Try out the following scenarios:

Test #1
Start with sample mean (lambda) set at 5 and sample size set 10.
Keep sample size at 10 and increase the lambda in increments of 5 till you reach lambda = 40.
Record your findings.

Test #2
Now keep lambda at 5 and start a test with a sample size equal to 2
Increment the sample size by 1 till you reach a value of 15 and record your findings.

Convergence of Poisson and Binomial Distribution to Normal Distribution

Prove the following statements using R

1. When the sample set is large, Poisson distribution with lambda n converges to Normal distribution with mean n and variance n.
2. When the sample set is large, Binomial distribution with n trials and a probability with p converges to Normal distribution with mean n*p and variance n*p*(1-p).