**Central Limit Theorem (CLT)**

**Define**

**Proper statement**

**A random sample form Any dist f(x)**

**The dist of Sum(x) approaches normal dist (mu, var)**

**if *n* tends to infinity**

**A random sample form normal dist**

**The dist of Sum(x) follows normal dist (mu, var)**

**n>+1( What ever n) (there is no condition on n)**

**Linear function of normal varaibles also follows normal dist**

**Draw r sample n from any fx**

**Calculate**

**Z[i]=(xbar-mu)/(sigma/sqrt(n)**

**Repeat a large number of times**

**Hist (z)**

**Save**

**Then change n**

**Z**

**Hist (z)**

**Save**

**As n increases, the shape of z will like normal, bell shaped**

**######################### Sampling dist of a statisitc**

**n=50; #sample size**

**mn=c(); # storage vector**

**for(i in 1:100000) # to Replicate 100000**

**{**

**x=rexp(n,1) # random number of size n**

**mn[i]=mean(x) # statistic**

**}**

**mean(mn) # mean of 100000 statistic stored above**

**var(mn)**

**hist(mn)**

**Form Binomial**

**n <- 4 # Number of trials (population size)**

**p<-0.05**

**s <- 2000 # Number of simulations**

**m <- c(20, 100, 500, 1000)**

**EX <- n\*p**

**VarX <- n\*p\*(1-p)**

**Z\_score <- matrix(NA, nrow = s, ncol = length(m))**

**for (i in 1:s){**

 **for (j in 1:length(m)){ # loop over sample size**

 **samp <- rbinom(n = m[j], size = n, prob = 0.05)**

 **sample\_mean <- mean(samp) # sample mean**

 **# Calculate Z score for mean of each sample size**

 **Z\_score[i,j] <- (sample\_mean-EX)/sqrt(VarX/m[j])**

 **}**

**}**

**# Display distribution of means**

**par(mfrow=c(4,1))**

**for (j in 1:4){**

 **hist(Z\_score[,j], xlim=c(-5,5),**

 **freq=FALSE, ylim=c(0, 0.5),**

 **ylab="Probability", xlab="",**

 **main=paste("Sample Size =", m[j]))**

 **# Density curve**

**x <- seq(-4, 4, by=0.01)**

 **y <- dnorm(x)**

 **lines(x, y, col="blue")**

**}**

**For ur learning**

**Just understand the codes**

**For exponential and binomial are being provided in the group.**

**What you have to do?**

**Take poison distn**

**Apply central limit theorem by changing n (5,10,25,50,100,200,500, 1000)**

**And store histogram**

**Then interpret the result by looking at histogram**

**Do this for**

**Log normal and**

**Laplace dist**