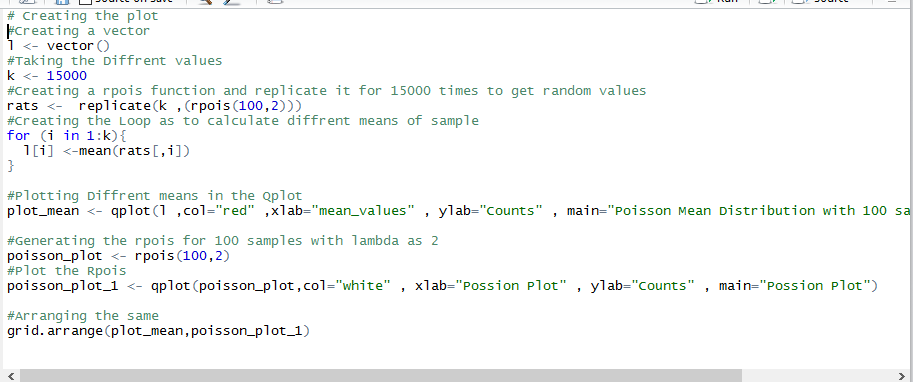
ISQS 5347 – Advanced Statistical Methods Midterm Exam 1.

The Central Limit Theorem says that given sufficiently large sample size, the sampling distribution of sample means will be approximately normal regardless of the shape of the distribution from which the samples are taken. Create an example to illustrate the Central Limit Theorem. Assume you are working with samples of size 100 drawn from a Poisson distribution with a lambda of 2. Your response should show graphs of both the Poisson distribution and your sampling distribution (based on a large number of simulated samples). You should also explain your example to someone who does not have extensive statistical knowledge. (12 points)

Central limit Theorem: - CLT states that if we take the different samples from the population and takes the mean of the samples and plot the samples it will plot like normal distribution



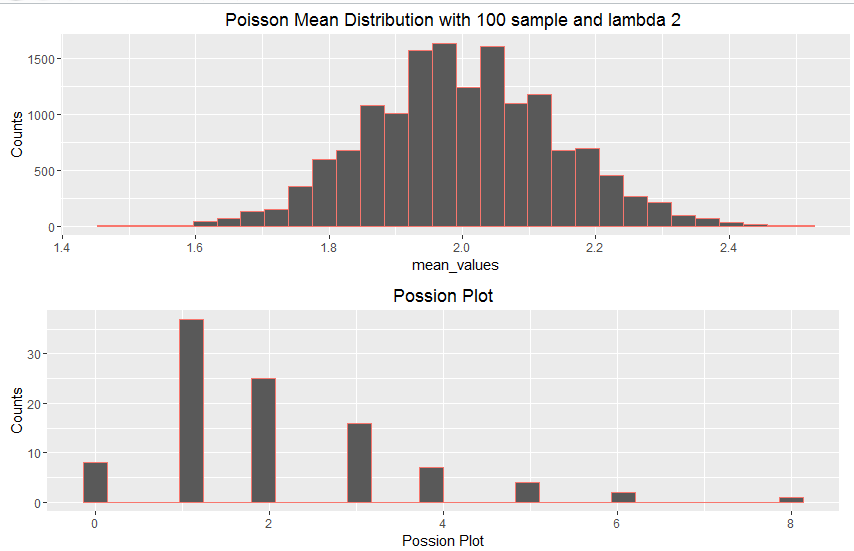
In the above question same scenario is there we have taken the population and taken different sample from the population, in the above question generating the random population with the Poisson Function and replicate it to 15000 times so that we can get the large sample size.

After taking the large population size we will take the small chunks of the sample and takes the mean of the different samples, the sample will be taken from population, sample consist of the 100 values and taking the mean of the different samples and store the samples in the vector.

Also taking the Poisson distribution of 100 samples with lambda of 2 and generate the values of the sample.

Plotting both the graph

1. Mean values generated by the concept of CLT (central limit theorem), is normally distributed and can be shown by the graph where peak values lie in the mean of the distribution.
2. Graph generated by the Poisson Distribution with n =100 and lambda = 2.



Conclusion: -

Hence CLT (Central Limit Theorem) can be proved from the above experiment.

2. How long are resale houses on the market? The answer to this question is likely to depend on local market conditions within an area. The data in the accompanying Excel file report days on the market for houses listed for resale in Houston and Chicago. Assume that the number of days a resale house is on the market is normally distributed and that population variances in the two cities are equal. At α=0.05, test whether the number of days on the market differs between the cities. (10 points)

a. Identify the statistical test you will use and justify your choice.

As we have given the Information that the population is normally distributed and the variance of the population is nearly equal, and Standard deviation of population is unknown therefore in such case we will use t-test.

b. State hypotheses, a decision rule, show computations (including a p-value), and state a decision and managerial conclusion.

H0: - The mean of the resale of the houses in both the cities are nearly equal.

Ha: - The mean of the resale of the houses in both the cities are not equal.

H0: µ(Houston) = µ(Chicago)

H1: µ(Houston) ≠ µ(Chicago)

**Desired level of significance and sample size:**

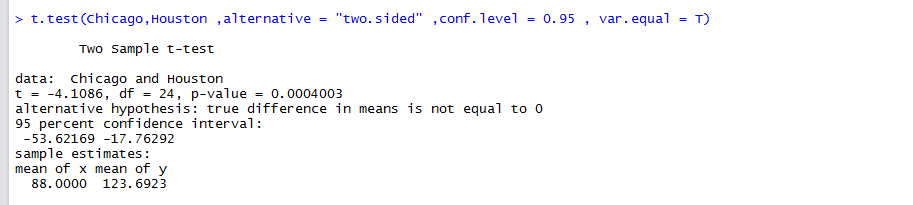
α = 0.05 n = 13

**critical values of t and state rejection rule:**



Reject H0 if tSTAT < -2.07 or if tSTAT > 2.07.

**Collect the data and compute the test statistic:**



**Evaluate test statistic, make decision, state managerial conclusion:**

Reject H0. At the .05 level of significance, there is not sufficient evidence to conclude that mean of resale is different.

**Alternately, test hypothesis using p-value:**

For a two tailed test, is the probability of getting a test statistic at least as extreme (in either direction) as the one found, if the null hypothesis is true.



These computations confirm the p-value reported in R t.test output. The p-value is less than the level of significance, so we reject H0.

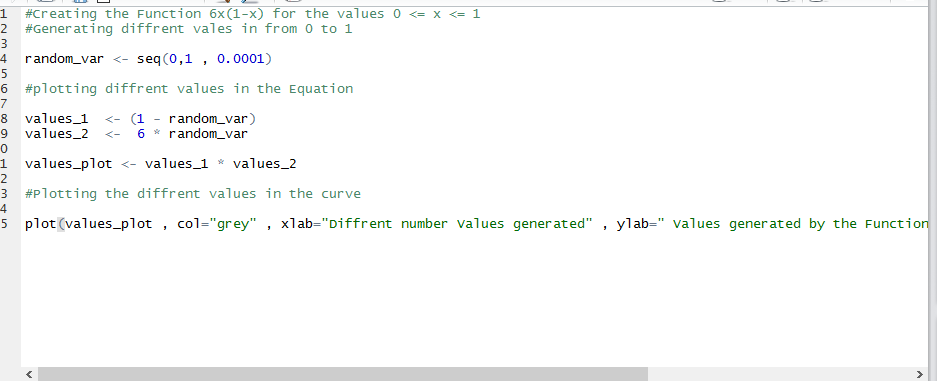
**c. Compute the 95% confidence interval for this data.**

Reject H0 if the hypothesized population mean is outside the confidence interval.

95% confidence interval, -53.62169 < µ < -17.76292

This confidence interval contains the hypothesized population mean, so we reject H0.

3. Given the following function: 𝑓(𝑥) = 6𝑥(1 − 𝑥)𝑓𝑜𝑟 0 ≤ 𝑥 ≤ 1 Does this function meet the criteria for a continuous probability density function? Justify your response. (8 points)



𝑓(𝑥) = 6𝑥(1 − 𝑥)𝑓𝑜𝑟 0 ≤ 𝑥 ≤ 1

Continuous Probability Density Function: -

A continuous random variable takes on an infinite number of possible values. For a discrete random variable X that takes the infinite number of discrete values for all possible values of X and can generate the different values.

Probability Density Function can be represented as: -

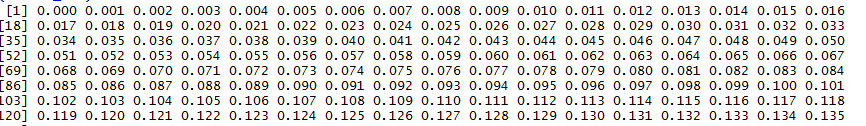
The random variable Y is a function of X; that is, y = f(x).

The value of y is greater than or equal to zero for all values of x.

The total area under the curve of the function is equal to 1.

In the above Equation: -

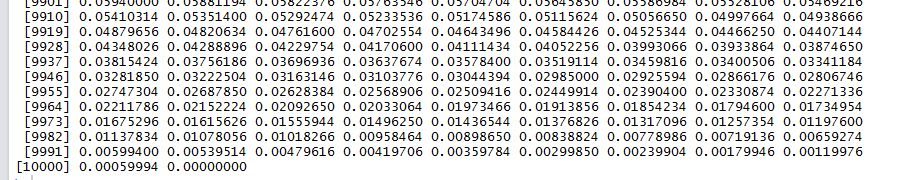
Getting the different values by the equation “(random\_var <- seq (0,1, 0.001)) “



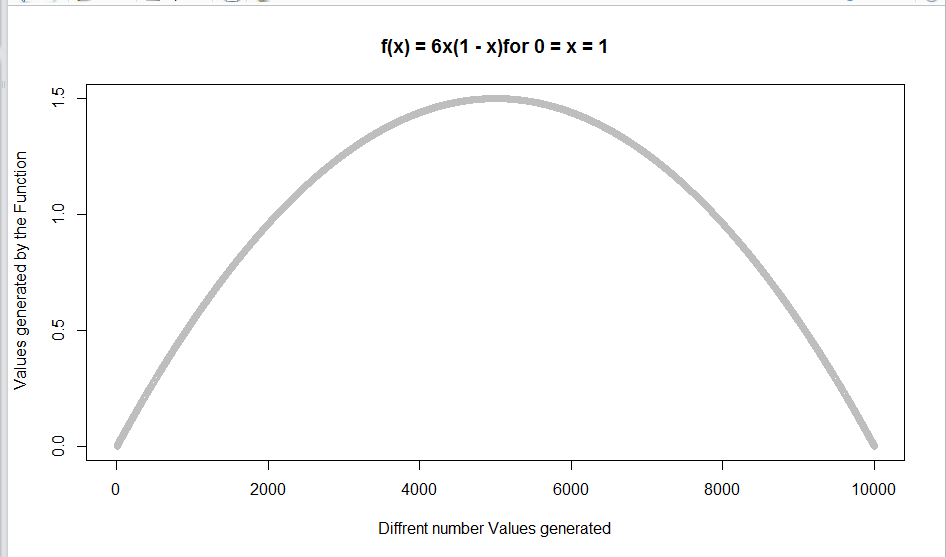
Plotting the Different values in the Function.

𝑓(𝑥) = 6𝑥(1 − 𝑥)𝑓𝑜𝑟 0 ≤ 𝑥 ≤ 1

We get the various randomly generated values.



Plotting the Different Values in the Graph :-



Mathematically represented by: -

P(0 ≤ 𝑥 ≤ 1) =

=

=6[1/2 -1/3]

=6[1/6] =1

Hence area under the curve is equal to 1.