**SECTION IV**

1. Hypothesis testing is a statistical procedure to provide a conclusion about a parameter in a population using data measured in a sample is true or false by comparing two data sets i.e., against an ideal data model.

Types of Hypothesis testing:

* F-test – for equality of variance

There is no general method to do this but when both populations are normal, we use F-test.

* Chi-square test – for categorical data and to find the goodness of fit and independent of attributes.
* ANOVA- Analysis Of Variance

To test the mean of more than two populations are equal

* T-test – for a normally distributed population where standard deviation in unknown and sample size is comparatively small.

1. Correlation: technique to find relationship between two continuous variables. For ex: Age and breath rate (in bpm)

a.) Computing methods:

* Pearson’s correlation coefficient – gives linear relation between variables
* Spearman’s rank correlation coefficient – measure of strength of linear relation between two data sets

b.)P-value: The probability of obtaining a sample outcome, given that the value stated in null hypothesis is true when compared with level of significance. It varies between 0 to1 and can never be negative. The p value can be calculated using the Z-score.

In case of standard normal distribution, check if null hypothesis is greater than alternative and z-value greater than test statistics. Then find the corresponding probability from z-table and subtract from 1.

1. a) Mean = 47.7 ,Variance=18.57, SD= 4.31

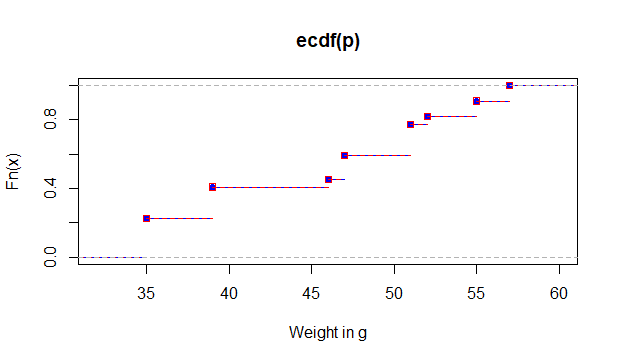
Empirical distributive function is plotted

**Code:**

p<-c(35,35,35,35,35,39,39,39,39,46,47,47,47,51,51,51,51,52,55,55,57,57)

s<-ecdf(p)

plot(s,col="blue",xlab="Weight in g")



b) 25% Quartile = 39

50% Quantile = 47

75% Quantile= 52

Median= 47

4. i) Mean of the population μ = 173.1

Variance =611.09

ii) The 95% confidence interval is between (157 – 189.2) cm