## 22-09-2015

## Experiment-7

**AIM:** To test both the samples follow the same distribution using Kolmogorov-Smirnov test.

**EXPERIMENT:** A class is randomly divided into two groups. One group is taught using a particular method 1 and second group is taught using method 2. At the end of the session each student is given a test for which the scores obtained are as follows:

Group 1: 227 176 252 149 16 55 234 294 247 92 184 147 88 161 171

Group 2: 202 14 165 171 292 271 151 235 147 99 63 284 53 228 271

Test whether the scores of the two groups follow the same distribution or not using Kolmogorov-Smirnov Test.

**THEORY:** In statistics, the Kolmogorov–Smirnov test (K–S test or KS test) is a nonparametric test of the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution (one-sample K–S test), or to compare two samples (two-sample K–S test).

H0: Both samples follow same distribution.

H1: They do not follow the same distribution.

**Test-Statistic:**

Where, sup x is the supremum of the set of distances

Fn1(x) = Empirical distribution function of sample 1

Fn2(x) = Empirical distribution function of sample 2

Fn(xi) = l/n , where l= number of values in the sample x(i) and n is the sample size. Here x(i) is the ordered statistic of the combined distinct observations.

**Test-Criteria:** Reject H0 if otherwise we accept the hypothesis. [ C(α)=1.36 and α=0.05]

**ALGORITHM:** We use two files pract7.txt and output7.txt for input and output and we construct two functions namely “sort” function to arrange the series in ascending order and “fun” function to compute Fn(xi) and then we follow these steps:

* We input the two random samples and merge them.
* Then we sort the combined sample in ascending order and print the distinct observations.
* Compute the functional value for the observations of both the samples and compute the difference between them.
* Sort the series of differences to get the supremum value which is the required test statistic.
* Then we check the value of test statistic with the threshold value and print the result.
* Output is obtained and the files are closed.

**RESULT:** We have the calculated value of Test Statistic Dn as 0.2 which is less than the threshold value, i.e., 0.667.

**CONCLUSION:** Since the value of test statistic is less than threshold value, we fail to reject the Null Hypothesis, which implies that both the samples follow a common/same distribution.