MARUDHAR KESARI JAIN COLLEGE FOR WOMEN, VANIYAMBADI

DEPARTMENT OF STATISTICS

SUBJECT NAME : TESTING STATISTICAL HYPOTHESIS

SUBJECT CODE : CST61

CLASS : III B.Sc Statistics

SYLLABUS

UNIT IV

Non-parametric tests - Run, Median, sign and Mann Whitney tests (one sample and two sample) problems. Wilcoxon Signed rank test, test sum test, Kolmogorov's Smirnov one sample test and Kruskal Wallis test.

Testing of Statistical Hypothesis

Non-Parametric test :

Non-Parametric tests are the mathematical methods used in statistical hypothesis testing, which do not make assumptions about the frequency distribution of variables that are to be evaluated

The non-parametric experiment is used when there are skewed data, and it comprises techniques that do not depend on data pertaining to any particular distribution.

The word non-parametric does not mean that these models do not have any parameters.

The fact is, the characteristics and no of parameters are pretty flexible and not predefined. Therefore, these models are called distribution free models.

Kruskal Wallis Test :

kruskal wallis test is used to compare the continuous outcome in greater than two independent samples.

Null Hypothesis (Ho):

Population medians are equal.

Test Statistic :

If N is the total sample size, K is the num

of comparison groups. R_j is the sum of the ranks in the j^{th} group and n_j is the sample size in the j^{th} group, then the test statistic, H is given by

$$H = \left(\frac{12}{N(N+1)} \sum_{j=1}^{k} \frac{R_j^2}{n_j}\right) - 3(N+1)$$

Decission Rules

Reject the Null Hypothesis H_0,

if $H \geq$ critical value.

Median test:

Median Test is a statistical procedure for testing, if two independent ordered samples differ in their central tendencies. In other words, it gives information if two independent samples are likely to have been drawn from the populations with the same median.

As in run test, let $x_1, x_2 \ldots x_n$ and y_1, y_2 be two independent ordered samples from the populat with pdf's $f_1(\cdot)$ and $f_2(\cdot)$ respectively. The measurem must be at least ordinal.

Let $z_1, z_2 \ldots z_{n_1+n_2}$ be the combined ordered sample. Let m_1 be the no. of x's and m_2 the no. of y exceeding the median value M (say). & the combine sample.

Then, under the null hypothesis that the sa comes from the same population or from different.

Mann-Whitney U Test:

Mann-Whitney U test also called the Mann-Whitney Wilcoxon) is a non-Parametric test of the null hypothesis that, for randomly selected values x and y from two populations, the probability of x being greater than y is equal to the probability of y being greater than x.

Non-Parametric test used on two dependent samples are the sign test and the Wilcoxon signed-rank test.

Assumptions:

1) All the observations from both groups are independent of each other.

2) The responses are at least ordinal (ie one can atleast say, of any two observations, which is greater).

3) Under the null Hypothesis H_0, the assumptions of both populations are identical.

4) The alternative Hypothesis H_1 is that the distribution are not identical.

The maximum value of U is the product of the sample sizes for the two samples (ie) $U = n_1 n_2$

In such a case, the other U would be 0.

Wilcoxon Signed rank test :

The wilcoxon signed-rank test is a non-parametric rank test for statistical hypothesis testing used either to test the location of a population based on a sample of data, or to compare the locations of two populations using two matched samples.

The one-sample version serves a purpose similar to that of the one-sample Student's t-test.

For two matched samples, it is a paired difference test like the paired student test also known as t-test for matched pairs or t-test for dependent samples.

Wilcoxon test can be good alternative to the t-test when population means are not of interest, for example, when one wishes to test whether a populations median is non-zero or whether there is a better than 50% chance that a sample from one population is greater than a sample from another population.

Kolmogorov–Smirnov Test :

In statistics, the Kolmogorov-Smirnov test is a non-parametric test of the equality of continuous, one dimentional probability distributions that can be used to test whether a sample

Came from a given reference probability distribution or to test whether two samples came from the same distribution.

Intuitively, the test provides a method of qualitatively answer the question "How likely is it that we would see a collection of samples like this if they were drawn from that probability distribution?" or in the second case

The Kolmogorov-Smirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function of the reference distribution or between the empirical distribution functions of two samples.

The null distribution of this statistic is calculated under the Null Hypothesis that the sample is drawn from the reference distribution or that the samples are drawn from the same distribution.

Kruskal Wallis-Test

The Kruskal wallis test by ranks, Kruskal-wallis H test or one-way ANOVA on ranks is a non-parametric method for testing whether samples originate from the same distribution.

It is used for comparing two or more independent samples of equal or different sample sizes. It extends the Mann-Whitney U test, which is used for comparing only two groups.

The parametric equivalent of the Kruskal-Wallis test is the one-way analysis of variance (ANOVA).

A significant Kruskal-Wallis test indicates that at least one sample stochastically dominates one other sample.

The test does not identify where this stochastic dominance occurs or for how many pairs of groups stochastic dominance obtains.