Non-Parametric Test

Md. Aminul Hoque, PhD

Non-Parametric Method

18·7. NON-PARAMETRIC METHODS

Most of the statistical tests that we have discussed so far had the following two features in common.

(i) The form of the frequency function of the parent population from which the samples have been drawn is assumed to be known, and

(ii) They were concerned with testing statistical hypothesis about the parameters of this frequency function or estimating its parameters.

For example, almost all the exact (small) sample tests of significance are based on the fundamental assumption that the parent population is normal and are concerned with testing or estimating the means and variances of these populations. Such tests, which deal with the parameters of the population are known as *Parametric Tests*. Thus, a parametric statistical test is a test whose model specifies certain conditions about the parameters of the population from which the samples are drawn.

On the other hand, a *Non-parametric (N.P.) Test* is a test that does not depend on the particular form of the basic frequency function from which the samples are drawn. In other words, non-parametric test does not make any assumption regarding the form of the population.

However, certain assumptions associated with N.P. tests are :

(i) Sample observations are independent.

(ii) The variable under study is continuous.

(iii) p.d.f. is continuous.

(iv) Lower order moments exist.

Obviously these assumptions are fewer and much weaker than those associated with parametric tests.

Definition

Nonparametric statistics is the branch of statistics that is not based solely on parametrized families of probability distributions. Nonparametric statistics is based on either being distribution-free or having a specified distribution but with the distribution's parameters unspecified.

Non-Parametric Test

Nonparametric tests

- Also known as distribution-free tests because they are based on fewer assumptions (e.g., they do not assume that the outcome is approximately normally distributed).
- Parametric tests involve specific probability distributions (e.g., the normal distribution) and the tests involve estimation of the key parameters of that distribution (e.g., the mean or difference in means) from the sample data.
- There are some situations when it is clear that the outcome does not follow a normal distribution. These include:
  - when the outcome is an ordinal variable or a rank,
  - when there are definite outliers or
  - when the outcome has clear limits of detection.

Parametric Test vs Non-Parametric Test

Comparison Basis	Parametric Test	Non-Parametric Test
Definition	A parametric test is one which has complete information about the population parameter or that can make assumptions about the parameters (defining properties) of the population distribution from which the samples are drawn.	A nonparametric test is one where the researcher has no knowledge about the population parameter, neither he can make specific assumptions, but still it is required to test the hypothesis of the population.
Information about population	Completely known	Not known/Unavailable
Basis of test statistic	Uses a normal probabilistic distribution	The distribution is non-normal/arbitrary
Measurement level	Applied when scale of measurement is a metric scale i.e. Interval or Ratio	Applied for Nominal or Ordinal scale
Measure of central tendency	Mean	Median
Applicability	Variables	Variables and attributes
Correlation test	Pearson	Spearman

Advantages and Disadvantages

Advantages

1. N.P. methods are readily comprehensible, very simple and easy to apply and do not require complicated sample theory.
2. No assumption is made about the form of the frequency function of the parent population from which sampling is done.
3. No parametric technique will apply to the data which are mere classification (*i.e.*, which are measured in nominal scale), while N.P. methods exist to deal with such data.
4. Since the socio-economic data are not, in general, normally distributed, N.P. tests have found applications in Psychometry, Sociology and Educational Statistics.

Drawbacks

1. N.P. tests can be used only if the measurements are nominal or ordinal. Even in that case, if a parametric test exists it is more powerful than the N.P. test. In other words, if all the assumptions of a statistical model are satisfied by the data and if the measurements are of required strength, then the N.P. tests are wasteful of time and data.
2. So far, no N.P. methods exist for testing interactions in 'Analysis of Variance' model unless special assumptions about the additivity of the model are made.

Cont…

5. N.P. tests are available to deal with the data which are given in ranks or whose seemingly numerical scores have the strength of ranks. For instance, no parametric test can be applied if the scores are given in grades such as A^+, A^-, B, A, B^+, etc.

3. N.P. tests are designed to test statistical hypothesis only and not for estimating the parameters.

Remarks:

1. Since no assumption is made about the parent distribution, the N.P. methods are sometimes referred to as *Distribution Free* methods. These tests are based on the 'Order Statistic' theory. In these tests we shall be using median, range, quartil, inter-quartile range, etc., for which an ordered sample is desirable. By saying that $x_1, x_2, \ldots, x_n$ is an ordered sample we mean $x_1 \leq x_2 \leq \ldots \leq x_n$.
2. The whole structure of the N.P. methods rests on a simple but fundamental property of order statistic, *viz. "The distribution of the area under the density function between any two ordered observations is independent of the form of the density function", which we shall now prove.*

18.7.2 Basic Distribution

Non-Parametric Methods

- Rank Test
- Randomness Test
- Run Test
- Sign Test
- Median Test
- Mann-Whitney-Wilcoxon U-Test
- Kruskall-Walis Test
- Wilcoxon Signed-rank Test

Randomness Test

18·7·4. Test for Randomness

Another application of the 'run' theory is in testing the randomness of a given set of observations. Let $x_1, x_2, \ldots, x_n$ be the set of observations arranged in the order in which they occur, *i.e.*, x_i is the *ith* observation in the outcome of an experiment. Then, for each of the observations, we see if it is

above or below the value of the median of the observations and write A if the observation is above and B if it is below, the median value. Thus, we get a sequence of A's and B's of the type, (say),

$$A\ B\ B\ A\ A\ A\ B\ A\ B\ B \qquad \ldots (*)$$

Under the *null hypothesis* H_0 *that the set of observations is random,* the number of runs U in (*) is a *r.v.* with

$$E(U) = \frac{n+2}{2} \quad \text{and} \quad \text{Var}(U) = \frac{n}{4}\left(\frac{n-2}{n-1}\right) \qquad \ldots (18{\cdot}106)$$

For large n (say, > 25), U may be regarded as asymptotically normal and we may use the normal test.

18·7·5. Median Test

...procedure for testing if two independent ordered samples differ in

Median Test

$$\mathrm{Var}(U) = \frac{n}{4}\left(\frac{n-2}{n-1}\right)$$

For large n (say, > 25), U may be regarded as asymptotically normal and we may use the normal test. ... (18·106)

18·7·5. 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 $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$ be two independent ordered samples from the populations with p.d.f.'s $f_1(.)$ and $f_2(.)$ respectively. The measurements 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 number of x's and m_2 the number of y's exceeding the median value M, (say), of the combined sample.

Then, under the null hypothesis that the samples come from the same population or from different populations with the same median, *i.e.*, under $H_0 : f_1(.) = f_2(.)$, the joint distribution of m_1 and m_2 is the hypergeometric distribution with probability function :

$$p(m_1, m_2) = \frac{\binom{n_1}{m_1}\binom{n_2}{m_2}}{\binom{n_1+n_2}{m_1+m_2}} \qquad \ldots (18{\cdot}107)$$

If $m_1 < n_1/2$, then the critical region corresponding to the size of type 1 error α, is given by $m_1 < m_1'$ where m_1' is computed from the equation :

$$\sum_{m_1=1}^{m_1'} p(m_1, m_2) = \alpha \qquad \ldots (18{\cdot}108)$$

The distribution of m_1 under H_0 is also hyper-geometric with

$$E(m_1) = \begin{cases} \dfrac{n_1}{2}, \text{ if } N = (n_1 + n_2) \text{ is even} \\ \dfrac{n_1}{2}\cdot\left(\dfrac{N-1}{N}\right), \text{ if } N \text{ is odd} \end{cases}$$

$$\text{and} \quad \mathrm{Var}(m_1) = \begin{cases} \dfrac{n_1 n_2}{4(N-1)}, & \text{if } N \text{ is even} \\ \dfrac{n_1 n_2 (N+1)}{4N^2}, & \text{if } N \text{ is odd} \end{cases} \qquad \ldots (18{\cdot}109)$$

This distribution is most of the times quite inconvenient to use. However, for large samples, we may regard m_1 to be asymptotically normal and use normal test, *viz.*,

$$Z = \frac{m_1 - E(m_1)}{\sqrt{Var(m_1)}} \sim N(0, 1), \text{ asymptotically.} \qquad \ldots (18{\cdot}110)$$

Mann-Whitney-Wilcoxon test

18·7·7. Mann-Whitney-Wilcoxon U-Test

This non-parametric test for two samples was described by Wilcoxon and studied by Mann and Whitney. It is the most widely used test as an alternative to the t-test when we do not make the t-test assumptions about the parent population.

Let x_i $(i = 1, 2, \ldots, n_1)$ and y_j $(j = 1, 2, \ldots, n_2)$ be independent ordered samples of size n_1 and n_2 from the populations with *p.d.f.* $f_1(.)$ and $f_2(.)$ respectively. We want to test the null hypothesis $H_1 : f_1(.) = f_2(.)$. Like the run test, Mann-Whitney test is based on the pattern of the x's and y's in the combined ordered sample. Let T *denote the sum of ranks of the y's in the combined ordered sample*. For example, for the pattern (18·99) on page 18·41 of combined ordered sample the ranks of y observations are respectively 3, 4, 5, 7, 10 etc. and $T = 3 + 4 + 5 + 7 + 10 + \ldots$ The test statistic U is then defined in terms of T as follows :

$$U = n_1 n_2 + \frac{n_2(n_2+1)}{2} - T \qquad \ldots (18{\cdot}113)$$

If T is significantly large or small then $H_0 : f_1(.) = f_2(.)$ is rejected. The problem is to find the distribution of T under H_0. Unfortunately, it is very troublesome to obtain the distribution of T under H_0. However, Mann and Whitney have obtained the distribution of T for small n_1 and n_2; have found the moments of T in general and shown that T is asymptotically normal. It has been established that under H_0, U is asymptotically normally distributed as $N(\mu, \sigma^2)$, where

$$\mu = E(U) = \frac{n_1 n_2}{2} \quad \text{and} \quad \sigma^2 = \text{Var}(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} \qquad \ldots (18{\cdot}114)$$

Hence,

$$Z = \frac{U - \mu}{\sigma} \sim N(0,1), \text{ asymptotically,} \qquad \ldots (18{\cdot}114a)$$

and normal test can be used. The approximation is fairly good if both n_1 and n_2 are greater than 8.

Remark: The asymptotic relative efficiency (*ARE*) of Mann-Whitney's U-test relative to two samples t-test is greater than or equal to 0·864. For a normal population, this $ARE = 3/\pi = 0{\cdot}955$. Accordingly, Mann-

Sign Test

... not as decisively as the t-test. ... with a little more assurance than does the sign ...

18·7·6. Sign Test

Consider a situation where it is desired to compare two things or materials under various sets of conditions. An experiment is thus conducted under the following circumstances :

(i) When there are pairs of observations on two things being compared.

(ii) For any given pair, each of the two observations is made under similar extraneous conditions.

(iii) Different pairs are observed under different conditions.

Condition (iii) implies that the differences $d_i = x_i - y_i$; $i = 1, 2, \ldots, n$ have different variances and thus renders the paired t-test (Chapter 16) invalid, which would have otherwise been used unless there was obvious non-normality. So, in such a case we use the '*Sign Test*', named so since it is based on the signs (plus or minus) of the deviations $d_i = x_i - y_i$. No assumptions are made regarding the parent population. The only assumptions are :

(i) Measurements are such that the deviations $d_i = x_i - y_i$, can be expressed in terms of positive or negative signs.

(ii) Variables have continuous distribution.

(iii) d_i's are independent.

Different pairs (x_i, y_i) may be from different populations (say *w.r.t.* age, weight, stature, education, etc.). The only requirement is that within each pair, there is matching *w.r.t* relevant extraneous factors.

Procedure. Let (x_i, y_i), $i = 1, 2, \ldots, n$ be n paired sample observations drawn from the two populations with *p.d.f.'s* $f_1(.)$ and $f_2(.)$. We want to test the null hypothesis $H_0 : f_1(.) = f_2(.)$. To test H_0, consider $d_i = x_i - y_i$, $(i = 1, 2, \ldots, n)$. When H_0 is true, x_i and y_i constitute a random sample of size 2 from the same population. Since the probability that the first of the two sample observations exceeds the second is same as the probability that the second exceeds the first and since hypothetically the probability of a tie is zero, H_0 may be restated as:

$$H_o : \quad P[X - Y > 0] = \frac{1}{2} \quad \text{and} \quad P[X - Y < 0] = \frac{1}{2}$$

Let us define:

$$u_i = \begin{cases} 1, & \text{if } x_i - y_i > 0 \\ 0, & \text{if } x_i - y_i < 0 \end{cases}$$

Cont...

U_i is a Bernoulli variate with $p = P(x_i - y_i > 0) = \frac{1}{2}$. Since U_i's, $i = 1, 2, \ldots, n$ are independent, $U = \sum_{i=1}^{n} U_i$, the total number of positive deviations, is a Binomial variate with parameters n and $p\left(=\frac{1}{2}\right)$. Let the number of positive deviations be k. Then

$$P(U \leq k) = \sum_{r=0}^{k} \binom{n}{r} p^r q^{n-r}, \left(p = q = \frac{1}{2} \text{ under } H_0\right)$$

$$= \left(\frac{1}{2}\right)^n \sum_{r=0}^{k} \binom{n}{r} = p', \text{(say)} \qquad \ldots (18{\cdot}111)$$

If $p' \leq 0{\cdot}05$, we reject H_0 at 5% level of significance and if $p' > 0{\cdot}05$, we conclude that the data do not provide any evidence against the null hypothesis, which may therefore, be accepted.

For large samples, $(n \geq 30)$, (under H_0), we may regard U to be asymptotically normal with,

$$E(U) = np = n/2 \quad \text{and} \quad \text{Var}(U) = npq = n/4$$

$$\therefore \quad Z = \frac{U - E(U)}{\sqrt{\text{Var}(U)}} = \frac{U - n/2}{\sqrt{(n/4)}}, \text{ is asymptotically } N(0, 1), \qquad \ldots (18{\cdot}112)$$

and we may use normal test.

18·7·7. Mann-Whitney-Wilcoxon U-Test

... was described by Wilcoxon and studied by Mann and ... do not make https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/sign-test/

Sign Test

The **Sign test** is a non-parametric test that is used to test whether or not two groups are equally sized. The sign test is used when dependent samples are ordered in pairs, where the bivariate random variables are mutually independent It is based on the direction of the plus and minus sign of the observation, and not on their numerical magnitude. It is also called the binominal sign test, with p = .5.. The sign test is considered a weaker test, because it tests the pair value below or above the median and it does not measure the pair difference. The sign test is available in SPSS: click “menu,” select “analysis,” then click on “nonparametric,” and choose “two related sample” and “sign test.

Questions Answered:

Which product of soda (Pepsi vs. Coke) is preferred among a group of 10 consumers?

Assumptions:

•**Data distribution:** The Sign test is a non–parametric (distribution free) test, so we do not assume that the data is normally distributed.

•**Two sample:** Data should be from two samples. The population may differ for the two samples.

•**Dependent sample:** Dependent samples should be a paired sample or matched. Also known as ‘before–after’ sample.

Types of sign test:

1.One sample: We set up the hypothesis so that + and – signs are the values of random variables having equal size.

2.Paired sample: This test is also called an alternative to the paired t-test. This test uses the + and – signs in paired sample tests or in before-after study. In this test, null hypothesis is set up so that the sign of + and – are of equal size, or the population means are equal to the sample mean.

Procedure:

1.Calculate the + and – sign for the given distribution.
Put a + sign for a value greater than the mean value, and put a – sign for a value less than the mean value. Put 0 as the value is equal to the mean value; pairs with 0 as the mean value are considered ties.

2.Denote the total number of signs by ‘n’ (ignore the zero sign)

3. and the number of less frequent signs by ‘S.’

3.Obtain the critical value (K) at .05 of the significance level

4. by using the following formula in case of small samples:

$$K = \frac{n-1}{2} - 0.98\sqrt{n}$$

$${}^{n}C_{x}\, q^{n-x} \cdot p^{x}$$

Sign test in case of large sample:

$$Z = \frac{S - np}{\sqrt{np(1-p)}}$$

For Binomial distribution the formula may be ${}^{n}C_{x} p^{x} q^{n-x}$ with p=0.5

Compare the value of ‘S’ with the critical value (K). If the value of S is greater than the value of K, then the null hypothesis is accepted.

If the value of the S is less than the critical value of K, then the null hypothesis is accepted. In the case of large samples, S is compared with the Z value.

- The **Wilcoxon signed-rank test** is a non-parametric statistical hypothesis test 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.[1] The one-sample version serves a purpose similar to that of the one-sample Student's *t*-test.[2] For two matched samples, it is a paired difference test like the paired Student's *t*-test (also known as the "*t*-test for matched pairs" or "*t*-test for dependent samples"). The Wilcoxon test can be a good alternative to the *t*-test when population means are not of interest; for example, when one wishes to test whether a population's median is nonzero, or whether there is a better than 50% chance that a sample from one population is greater than a

- https://en.wikipedia.org/wiki/Wilcoxon_signed-rank_test