Frequently Asked Questions

1. What do you mean by Non parametric tests?

   Answer:

   Nonparametric methods provide an alternative series of statistical methods that require no or very limited assumptions to be made about the data. There is a wide range of methods that can be used in different circumstances, but some of the more commonly used are the nonparametric alternatives to the t-tests

   If the test does not require the knowledge of the parent population or in other words if the test does not require any such specification of the parameters it is known as Non Parametric tests.

   For example : Chi-square test, Sign test, Run test, Mann Whitney U test etc.

2. State the assumptions for the Non parametric tests

Answer:

Nonparametric Assumptions

1.Observations are independent

2.Variable under study has underlying continuity

3. The proabability distribution function is continuous

4. The lower order moments like mean and variance exists

3. What do you mean by a sign test ?

Answer:

The sign test is the simplest of the nonparametric tests, and is similar to testing if a two-sided coin is fair. Its name comes from the fact that it is based on the direction (or signs for pluses and minuses) of a pair of observations and not on their numerical magnitude. Count the number of positive values (larger than hypothesized median), the number of negative values (smaller than the hypothesized median), and test whether there are significantly more positives (or negatives) than expected. The One Sample

Sign Test is a nonparametric equivalent to the parametric One Sample t-Test.

4. Write a note on the type of the hypothesis used for a one sample sign test

Answer:

The one-sample sign test is used to test the null hypothesis that the probability of a random value from the population being above the specified value is equal to the probability of a random value being below the specified value.

Nonparametric tests do not make assumptions that the population is from a specific distribution. Therefore its results are more robust than a parametric test when such assumptions are violated.

Test	**Null Hypothesis, H_0**	**Alternate Hypothesis, H_1**
One-tailed	$M = M_0$	$M < M_0$ or $M > M_0$
Two-tailed	$M = M_0$	$M \neq M_0$

where M is the population median and M_0 is the hypothesized population median.

5. What are the assumptions behind the sign test?

Answer:

Data is non-normally distributed, even after log transforming.

This test makes no assumption about the shape of the population distribution, therefore this test can handle a data set that is non-symmetric, that is skewed either to the left or the right.

6. What are the concepts to be familiar with to analyze and interpret results of the *one-sample sign test*

Answer:

To properly analyze and interpret results of the *one-sample sign test*, you should be familiar with the following terms and concepts:

- One sample problem
- Independent samples
- Violation of test assumptions
- Distribution free tests
- Rank tests

7. What is the principle of sign test?

Answer:

Under the hypothesis that the sample median (m) is equal to some hypothesized value (Mo , so H0: M = Mo), then you would expect half the data set S of sample size n to be greater than the hypothesized value Mo. If S > 0.5n then M > Mo, and if S < 0.5n then M< Mo. The SIGN TEST simply computes whether there is a significant deviation from this assumption, and gives you a p value based on a binomial distribution.

8. Discuss the advantages of one sample sign test

Answer:

Sign tests require no or very limited assumptions to be made about the format of the data, and they may therefore be preferable when the assumptions required for t tests are not valid.

Sign test can be useful for dealing with unexpected, outlying observations that might be problematic with a parametric approach.

Sign tests are intuitive and are simple to carry out by hand, for small samples at least.

Sign tests are often useful in the analysis of ordered categorical data in which assignation of scores to individual categories may be inappropriate.

Probability statements obtained from most nonparametric statistics are exact probabilities, regardless of the shape of the population distribution from which the random sample was drawn. The 1-Sample Sign is a nonparametric test of population location (median) and also calculates the corresponding point estimate and confidence interval. Its parametric counterpart is the 1-sample Z and 1-sample t tests.

- If sample sizes as small as N=6 are used, the simplest procedure available is sign test
- Treat samples made up of observations from several different populations.
- Can treat data which are inherently in ranks as well as data whose seemingly numerical scores have the strength in ranks
- They are available to treat data which are classificatory
- Easier to learn and apply than parametric tests

9. What are the disadvantages of one sample sign tests?

Answer:

Sign test may lack power as compared with more traditional approach one sample t-test. This is a particular concern if the sample size is small or if the assumptions for the corresponding parametric method (e.g. Normality of the data) hold.

Sign tests are geared toward hypothesis testing rather than estimation of effects. It is often possible to obtain nonparametric estimates and associated confidence intervals, but this is not generally straightforward.

Tied values can be problematic when these are common, and adjustments to the test statistic may be necessary.

Appropriate computer software for one sample sign test can be limited, although the situation is improving. In addition, how a software package deals with tied values or how it obtains appropriate p values may not always be obvious.

The main drawback of sign test is that they are not as efficient and powerful as t test that are based on a known underlying distribution

10. State the concept behind the one sample sign test?

Answer:

If you want to test whether the hypothesized value is not equal to the sample median test uses both the tails of the distribution.

If you are only interested in whether the hypothesized value is greater or lesser than the sample median (H0: M > or < Mo), the test uses the corresponding upper or lower tail of the distribution.

1Make null hypothesis about true median

2. Let S = number of values greater than median

3.Each sampled item is independent

4. If null hypothesis is true, S should have binomial distribution with success probability .5

11. Write a note on procedure of sign test?

Answer:

The procedure is very simple. Let a random sample x1,x2,…,xn of size n be drawn from a population with distribution function F(x) where F(x) is assumed to be continuous in close vicinity of a n average (say Median). Suppose the Median of F(x) is m then P[X=M]=0

To test H0:M=M0 against the alternative hypothesis H1: M≠M0 where M0 is the given value of the population median. We know that P[X>M0]=P[X<M0]=0.5. Hence a null hypothesis under test is equivalent to H0:P[X>M0]=P[X<M0] against H1: P[X>M0]≠P[X<M0]

To perform the sign test we take the differences xi – M0 for i=1,2,…,n and consider the signs. Let the number of positive signs be U and negative signs (n-U) .

For the test we consider only the number of positive signs. In this way the data have been dichotomized which consists of number of positive signs and negative signs. The distribution of U given n is a Binomial distribution with p=P[X>M0]. Thus the null hypothesis H0 changes to H0:p=0.5 . So now we test H0:p=0.5 against H0:p ≠0.5.

Compute

$P(X \leq U) = \sum^{U} n_{c_x} (\frac{1}{2})^x (\frac{1}{2})^{n-x}$ and this value is compared with α/2, the level of significance. If the calculated value of P(X≤U) is less than the predicted α/2, the null hypothesis is rejected

12. How do you carry on with the sign test when the sample size is large?

Answer:

If n≥25, the value of Z is computed and the normal test is applied to decide about H0 where

$$Z = \frac{(U + 0.5) - np_0}{\sqrt{np_0(1 - p_0)}}$$ when $U<np_0$

$$Z = \frac{(U - 0.5) - np_0}{\sqrt{np_0(1 - p_0)}}$$ when $U>np_0$ where $p_0 = 0.5$

If $Z \geq Z_{\alpha/2}$ or $Z \leq Z_{1-\alpha/2}$. a null hypothesis is rejected which implies $|Z| \geq Z_{\alpha/2}$ a null hypothesis is rejected

13. Explain one sided test procedure for one sample sign test.

Answer:

To test H0:p=0.5 against H0:p >0.5. Compute

the value of Z and the normal test is applied to decide about H0 where

$$Z = \frac{(U + 0.5) - np_0}{\sqrt{np_0(1 - p_0)}}$$ when $U<np_0$

$Z = \frac{(U-0.5)-np_0}{\sqrt{np_0(1-p_0)}}$ when U>np_0 where p_0 =0.5

If Z > Z_α , a null hypothesis is rejected

To test H0:p=0.5 against H0:p <0.5. Compute

the value of Z as above and the normal test is applied to decide about H0

If Z <- Z_α , a null hypothesis is rejected

14. The inventory ordering policy of a particular shoe store is partly based on the belief that the median foot size of teenage boys is 10.25 inches. To test this hypothesis the foot size of each of the random sample of 50 boys was determined. Suppose that 36 boys had sizes in excess of 10.25 inches . Does this disprove the hypothesis that the median size is 10.25?

Answer:

- Let X be a Binomial random variable with parameters 50 and ½. Since 36 is larger than 50X ½ = 25, we see that the p value is
- P value = 2P{X≥36]
- We can now see the Normal approximation
- E(X)= 50X1/2=25 V(X)= 12.5
- P value = 2P{X≥36]= 2P{X≥35.5](the continuity correction)

$$pvalue = 2P\left\{\frac{X-25}{\sqrt{12.5}} \geq \frac{35.5-25}{\sqrt{12.5}}\right\} = 2P[Z \geq 2.97] = 0.0030$$

Thus the belief that the median shoe size is 10.25 inches is rejected even at the 1% level of significance. There appears to be a strong evidence that the median shoe size is greater than 10.25

15. The PQR campany claims that the life time of a type of battery that it manufactures is more than 250 hours. A consumer advocate wishing to determine whether the claim is justified measures the life times of 24 of the company's batteries. Assuming the sample to be random, determine whether the company's claim is justified at the 0.05 significance level. Observations are :

271, 230, 198, 275, 282, 225, 284, 219

253, 216, 262, 288, 236, 291, 253, 224

264, 295, 211, 252, 294, 243, 272, 268

Answer:

Let H0 be the hypothesis that the company's batteries have a life time equal to 250 hours and let H1 be the hypothesis that they have a life time greater than 250 hours. To test H0 against H1 we can use sign test.

To do this we substract 250 from each entry of the above table and record the signs of differences as shown in the table below. We see that there are 15 plus signs and 9 minus signs

+	-	-	+	+	-	+	-
+	-	+	+	-	+	+	-
+	+	-	+	+	-	+	+

Using a one tailed test at te 0.05 significance level , we would reject H0 if the Z score were greater than 1.645. Since the Z score using a correction for continuity is (Since U > np_0 that is 15 > (24)(0.5))

$$Z = \frac{(U-0.5)-np_0}{\sqrt{np_0(1-p_0)}} = \frac{(15-0.5)-(24)(0.5)}{\sqrt{(24)(0.5)(0.5)}} = 1.02$$

Since Z is less than Z_α at 5% level of significance we cannot accept the alternative hypothesis. The company's claim cannot be justified at the 0.05 level