

Nonparametric test

Introduction

Choosing the right test to compare measurements is a bit tricky, as you must choose between two families of tests: parametric and nonparametric. Many statistical tests are based upon the assumption that the population from which the samples are drawn are normally distributed. These tests are referred to as parametric tests. The z test, t test, F test, ANOVA is some examples of parametric tests.

Non parametric tests are used when data is not normally distributed. Nonparametric methods were developed to be used in cases when the researcher knows nothing about the parameters of the variable of interest in the population. Hence the name *nonparametric* is used. Nonparametric methods do not rely on the estimation of parameters describing the distribution of the variable of interest in the population. Therefore, these methods are also sometimes called *parameter free* methods or *distribution free* methods.

Advantages of non parametric test

1. Non Parametric tests are readily comprehensible, easy to apply and do not require complicated sample theory.
2. No assumption is made about the population from which sampling is done.(i.e. they do not incorporate all the restrictive assumptions characteristics of parametric tests)
3. No parametric will apply to the data which are mere classifications, while non parametric methods are readily applicable.
4. Since the socio-economic data are not, in general, normally distributed, non parametric test have found application in various social sciences like — Psychometry, Sociology and Educational Statistics etc.
5. Non Parametric tests are available to deal with data which are given in ranks or grades. Use of ranks makes non parametric techniques less sensitive to measurement error.

Disadvantages of Non- Parametric Test

1. Non Parametric test can be used only if the measurement are nominal or ordinal. Even in that cases, if parametric test exists they are more powerful than non parametric test.
2. If the null hypothesis is false, the non parametric test would require a larger sample to provide sufficient evidence to reject it.
3. Non Parametric tests are designed to test statistical hypothesis only and not for estimating parameter.
4. The hypothesis tested by nonparametric techniques tends to be less specific than tested by traditional methods. Because they rely on ranks rather than on the actual values of the observations, non parametric tests do not use everything that is known about a distribution.
5. So far, no nonparametric test is available for testing interactions in ANOVA model unless specific assumptions are made about the additivity of the the model.

Types of non parametric tests

A large number of non parametric tests exist, but in this chapter we will discuss some non parametric tests which are widely used. Some of them are:

The sign test, Median test, The run test, Wilcoxon test, Mann-Whitney U test, Kruskal-Wallis tests, Friedman's test, Fisher exact test, McNemar test, Chisquare test

Basically, there is at least one nonparametric equivalent for each parametric general type of test. In general, these tests fall into the following categories:

- Tests of differences between groups (independent samples);
- Tests of differences between variables (dependent samples);
- Tests of relationships between variables.

| Parametric tests | Non parametric tests |
|--|---|
| 1 sample t test | Sign Test/Wilcoxon Signed Rank test |
| Paired t test | Sign Test/Wilcoxon Signed Rank test |
| 2 sample t test | Mann Whitney U test Wilcoxon Sum Rank test Kolmogorov-Smirnov two sample test |
| ANOVA | Kruskal Wallis test/ Median test |
| Qualitative data | |
| For independent samples | Chi square test Fisher's exact test |
| For matched case-control samples | McNemar test |
| Bivariate correlation (quantitative data) | |
| Normality assumption satisfied | Pearson's correlation |
| Normality assumption not satisfied or ordinal qualitative data | Spearman's correlation |

Chi square test

Definition

A measure of discrepancy between the observed and expected frequencies is known as χ^2 statistic and is defined by,

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

$$= \frac{(O_1 - E_1)^2}{E_1} + \frac{(O_2 - E_2)^2}{E_2} + \dots + \frac{(O_n - E_n)^2}{E_n}$$

Where O_i 's are the observed (or actual) frequencies and E_i 's are the corresponding expected (or theoretical) frequencies.

Degree of freedom

If the data are presented as a series of variate values in row or column, the degree of freedom will be the number of items in the series less than one. i.e. $\nu = n - 1$. If the number of frequencies are presented in cells in a contingency table (a table with two or more than two rows and columns), the degree of freedom will be $\nu = (r - 1)(c - 1)$ where r and c denotes the number of rows and columns respectively.

Conditions for applying χ^2 test

- i. The total number of observations must be large i.e. χ^2 test is appropriate if $n \geq 50$.
- ii. χ^2 cannot be used for estimating the value of the population parameter.
- iii. No theoretical cell frequency should be less than 5. If it is less than 5, we have to pool the succeeding or preceding cell frequency so that the frequency must be 5 or more than 5 and the degree of freedom is adjusted accordingly. For the adjustment of degree of freedom, the number of items which are pooled are subtracted in the degree of freedom.

Applications of χ^2 test

When the given data are available in the form of frequencies χ^2 test is applied. χ^2 test has a wide application in testing of hypothesis. Some applications of this test are:

1. Test of goodness of fit
2. Test of independence of attributes
3. Test of homogeneity
4. Test of population variance

Test of goodness of fit

The steps in testing are as follows:

1. Formulate null and alternative hypothesis

H_0 : there is no significant difference between experimental and theoretical frequency or the given model or theory fits the data

H_1 : there is a significant difference between experimental and theoretical frequency or the given model or theory does not fits the data

2. Compute the test statistic

Under H_0 , the test statistic is,

$$\chi^2 = \sum \frac{(O - E)^2}{E} \sim \chi_{n-1}^2$$

Where, O = observed frequency, E = expected frequency = $\frac{\sum O}{n}$

3. **Critical value:** Obtain the tabulated value of χ^2 at α % level of significance and $\nu = n - 1$ degree of freedom. In this case, there is only one variable and therefore only one row or column and the d.f. is $(n - 1)$ only. The critical value of χ^2 is obtained from the table.
4. **Decision:** If $\chi_{cal}^2 \leq \chi_{tab}^2$, H_0 is accepted otherwise rejected.

Test of independence of attributes

The steps in testing are as follows:

1. Formulate null and alternative hypothesis

H_0 : the attributes are independent to each other or the treatment has no effect.

H_1 : the attributes are not independent to each other.

2. Compute the expected cell frequency by using the relation,

$$\text{Expected frequency (E)} = \frac{\text{Row total} \times \text{Column total}}{\text{Grand total}}$$

3. Compute the test statistic

Under H_0 , the test statistic is,

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

4. **Critical value:** Obtain the tabulated value of χ^2 at α % level of significance and $(r - 1)(c - 1)$ degree of freedom

5. **Decision:** If $\chi_{cal}^2 \leq \chi_{tab}^2$, H_0 is accepted otherwise rejected.

Remark: For 2×2 contingency table as shown below,

| Attributes | Attributes | | Total |
|----------------|----------------|----------------|-------------------|
| | B ₁ | B ₂ | |
| A ₁ | a | b | a + b |
| A ₂ | c | d | c + d |
| Total | a + c | b + d | a + b + c + d = n |

The value of χ^2 can be directly computed by the formula,

$$\chi^2 = \frac{n(ad - bc)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Yates correction

If any cell frequency (observed frequency or expected frequency) for 2×2 contingency table is less than 5, then we use the technique of pooling which consists in adding the frequencies which are less than 5 with the preceding or succeeding frequency to get sum 5 or more.

In 2×2 contingency table, the number of degree of freedom is,

$$(r - 1)(c - 1) = (2 - 1)(2 - 1) = 1$$

But 1 d.f is lost due to pooling, so d.f after pooling equals to 0, which is meaning less.

Since χ^2 must be at least 1 d.f. In this case, we apply a correction given by F- Yates which is usually known as '*Yates correction for continuity*'. The working rule for the application of the correction is to add 0.5 to the cell frequency which is less than 5 and adjust the remaining frequencies accordingly by fixing the row and column totals. Then apply χ^2 test without pooling.

Lets us suppose 2×2 contingency table,

| Attributes | Attributes | | Total |
|----------------|----------------|----------------|-------------------|
| | B ₁ | B ₂ | |
| A ₁ | a | b | a + b |
| A ₂ | c | d | c + d |
| Total | a + c | b + d | a + b + c + d = n |

Then the value of χ^2 after Yates correction is,

$$\chi^2 = \frac{n \left(\left| ad - bc \right| - \frac{n}{2} \right)^2}{(a + b)(c + d)(a + c)(b + d)}$$

Exercise

1. Random samples of weekly child births at the local maternity hospital showed that average number of births on each day of the week as follows:

| Day | Sun | Mon | Tue | Wed | Thurs | Fri | Sat |
|---------------|-----|-----|-----|-----|-------|-----|-----|
| No. of births | 65 | 53 | 35 | 38 | 62 | 54 | 43 |

At 1% level of significance, can we conclude that the proportion of births on some days is higher than on other days?

2. Genetic theory states that children having one parent of blood type M and another blood type N will always be one of the three types M, MN, N and the proportion of three types will be on average be as 1:2:1. A report states that out of 300 children having one M parent and one N parent 30% were found to be type M, 45% type MN and remainder type N. Test the hypothesis by χ^2 test.
3. Among 64 offspring of a certain cross between guinea pigs, 34 were red, 10 were black and 20 were white. According to the genetic model these numbers should be in the ratio 9:3:4. Are the data consistent with the model at 5% level of significance?

(1.44)

4. 192 families (for each of which the probability of an albino child being born is otherwise established) has the following distribution of albinos among the first children.

| No. of albinos | 0 | 1 | 2 | 3 | Total |
|-----------------|----|----|----|---|-------|
| No. of families | 77 | 90 | 20 | 5 | 192 |

Find the expected frequencies on the basis of the theoretical probability 0.25 of a child being born an albino and test the goodness of fit. (2.031)

5. Records taken of the number of male and female births in 800 families having four children are given below:

| | | | | | |
|----------------------|----|-----|-----|-----|----|
| No. of female births | 4 | 3 | 2 | 1 | 0 |
| No. of male births | 0 | 1 | 2 | 3 | 4 |
| Frequency | 32 | 178 | 290 | 236 | 64 |

Test whether the data is consistent with the hypothesis that the binomial distribution holds and the chance of male births is equal to that of female births.

6. On the basis of information given below about the treatment of 200 patients suffering from a disease, state whether the new treatment is comparatively superior to the conventional treatment.

| Treatment | No. of patients | |
|--------------|--------------------|-------------|
| | Favorable response | No response |
| New | 60 | 20 |
| Conventional | 70 | 50 |

For drawing your inference use the value of χ^2 for 1 d.f at 5% level of significance. ($\chi^2 = 5.861$)

7. From the following data use χ^2 test and conclude whether inoculation is effective in preventing tuberculosis.

| | Attacked | Not attacked | Total |
|----------------|----------|--------------|--------------------------------------|
| Inoculation | 31 | 469 | 500 |
| Not inoculated | 185 | 1315 | 1500 |
| Total | 216 | 1784 | 2000 ($\chi^2=14.64$, significant) |

8. A sample of 500 college students participated in a study designed to evaluate the level of college students' knowledge of a certain group of common diseases. The following table shows the students classified by major field of study and level of knowledge of the group of diseases:

| Major | Knowledge of disease | | Total |
|------------|----------------------|------|-------|
| | Good | Poor | |
| Premedical | 31 | 91 | 122 |
| Other | 19 | 359 | 378 |
| Total | 50 | 450 | 500 |

Do these data suggest that there is a relationship between knowledge of the group of diseases and major field of study of the college students from which the present sample was drawn? Let $\alpha = 0.01$

9. The following data correspond to number of patients who were given treatment for tuberculosis under DOTS strategy and under non DOTS strategy with corresponding number of cure patients.

| Outcome | Treatment | | Total |
|-----------|-----------|----------|-------|
| | DOTS | Non DOTS | |
| Cured | 140 | 90 | 230 |
| Not cured | 20 | 60 | 80 |
| Total | 160 | 150 | 310 |

Is there a significant difference in the cure rates in patients with TB who were given treatment under DOTS and non DOTS strategy?

10. Associations between blood groups and severity is presented in the contingency table below:

| Condition | Blood group | | | | Total |
|-----------|-------------|-----|-----|-----|-------|
| | O | A | B | AB | |
| Severe | 51 | 40 | 10 | 9 | 110 |
| Moderate | 105 | 103 | 25 | 17 | 250 |
| Mild | 384 | 527 | 125 | 104 | 1140 |
| Total | 540 | 670 | 160 | 130 | 1500 |

Can we conclude that the data are compatible with the hypothesis that severity of the condition and blood type are associated? ($\chi^2 = 13.35$)

11. In an epidemic of a certain disease 92 children contracted the disease. Of these 41 received no treatment and of these 10 showed after effects, of the remainder who did receive treatment, 17 showed after effects. Test the hypothesis that treatment was not effective. (0.85)
12. Following table provides data with regard to stature of the fathers and their sons at the age of 25 years.

| | | Stature of fathers | | Total |
|-----------------|-------|--------------------|-------|-------|
| | | Tall | Short | |
| Stature of sons | Tall | 8 | 2 | 10 |
| | Short | 4 | 6 | 10 |
| Total | | 12 | 8 | 20 |

Test that the stature of sons is independent of the stature of the fathers.

$$(\chi^2 = 1.875, H_0 \text{ accept})$$

13. In an experiment on the immunization of goats from anthrax, the following results were obtained. Derive your inference on the efficiency of the vaccine.

| | Died | Survived | Total |
|----------------|------|----------|-------|
| Inoculated | 2 | 10 | 12 |
| Not inoculated | 6 | 6 | 12 |
| Total | 8 | 16 | 24 |

(1.6875)

Fisher exact test

Fisher exact probability test for two independent samples is essentially a restricted chi square test with 2×2 contingency table and is extremely useful non parametric test for analyzing discrete data (either nominal or ordinal). It is designed to test equality of proportions in two classes.

Lets us suppose 2×2 contingency table,

| Attributes | Attributes | | Total |
|----------------|----------------|----------------|-------------|
| | B ₁ | B ₂ | |
| A ₁ | a | b | a + b |
| A ₂ | c | d | c + d |
| Total | a + c | b + d | a+b+c+d = n |

Null and alternative hypothesis

H₀: $P_1 = P_2$, the two samples do not differ significantly in the proportion of attributes to them.

H₁: $P_1 \neq P_2$, the two samples differ significantly in the proportion of attributes to them.

Mechanism:

- Cast the observed frequencies in a 2×2 contingency table.
- Compute the marginal totals and marginal total sums to n.
- Compute the exact probability of the configurations of the contingency table of order 2×2 such as from smallest observed frequency up to zero by reducing each time the smallest cell frequency by unity keeping the marginal totals fixed.
- For a 2×2 contingency table with fixed marginal totals, the exact probability of a particular set of frequencies is given by hypergeometric distribution.

$$p_k = \frac{\binom{a+c}{a} \binom{b+d}{b}}{\binom{n}{a+b}} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!}$$

- Finally compute $p = p_k + p_{k-1} + \dots + p_1 + p_0$ which is the probability corresponding to one tail of the distribution.

Test statistic

The test statistic is provided by the p- value given by,

$$p = p_k + p_{k-1} + \dots + p_1 + p_0$$

Critical region

Fix a level of significance α

Decision

Reject H_0 at $\alpha\%$ level of significance if $p \leq \alpha$ otherwise accept.

Example: A researcher wishing to find the popularity of new curriculum and old curriculum among students revealed the information as follows:

| Sex | Old curriculum | New curriculum | Total |
|-------|----------------|----------------|-------|
| Boys | 2 | 5 | 7 |
| Girls | 3 | 2 | 5 |
| Total | 5 | 7 | 12 |

Do the information provide sufficient evidence to test the hypothesis that greater proportion of boys are in favor of new curriculum. Use 1% level of significance.

Solution:

Contingency table I

| Sex | Old curriculum | New curriculum | Total |
|-------|----------------|----------------|-------|
| Boys | 2 | 5 | 7 |
| Girls | 3 | 2 | 5 |
| Total | 5 | 7 | 12 |

Given, $a=2, b=5, c=3, d=2$ and $n = 12$.

H_0 : Proportion of enrollment of students in both new and old curriculum are same.

H_1 : Greater proportion of boys is in favor of new curriculum as compared to girls.

Test statistic is,

$$p_k = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!} = \frac{7!5!7!}{12!2!5!3!2!} = 0.26515$$

Contingency table II

| Sex | Old curriculum | New curriculum | Total |
|-------|----------------|----------------|-------|
| Boys | 1 | 6 | 7 |
| Girls | 4 | 1 | 5 |
| Total | 5 | 7 | 12 |

Here $a = 1, b = 6, c = 4, d = 1, n = 12$

$$p_{k-1} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!} = \frac{7!5!7!}{12!!6!4!!} = 0.04399$$

Contingency table III

| Sex | Old curriculum | New curriculum | Total |
|-------|----------------|----------------|-------|
| Boys | 0 | 7 | 7 |
| Girls | 5 | 0 | 5 |
| Total | 5 | 7 | 12 |

Here $a = 0, b = 7, c = 5, d = 0, n = 12$

$$p_{k-2} = \frac{(a+b)!(c+d)!(a+c)!(b+d)!}{n!a!b!c!d!} = \frac{7!5!7!}{12!0!7!5!0!} = 0.00126$$

$$\begin{aligned} \therefore p &= p_k + p_{k-1} + p_{k-2} \\ &= 0.26515 + 0.04399 + 0.00126 = 0.31040 \end{aligned}$$

Here, $\alpha = 0.01$

Decision:

Since $p = 0.31040 > 0.01$, we accept H_0 . Which means proportion of enrolment of students in both new and old curriculum is same.

Exercise

1. In a comparative study of two new drugs A and B, 9 patients were treated with drug A and 6 patients treated with drug B and the following results were obtained.

| Drug | Cured | Not cured | Total |
|-------|-------|-----------|-------|
| A | 1 | 8 | 9 |
| B | 6 | 0 | 6 |
| Total | 7 | 8 | 15 |

Do the results demonstrate a significantly higher cure rate with drug B than drug A? Test at 5% level of significance. (0.00139, Reject H_0)

2. From a universe of 12 females using oral contraceptives and 8 females using other contraceptive devices were randomly selected and the numbers of hypertensive cases from both groups were recorded as below:

| Types of contraceptives | Hypertensive | Normal | Total |
|-------------------------|--------------|--------|-------|
| Oral | 11 | 1 | 12 |
| Other | 2 | 6 | 8 |
| Total | 13 | 7 | 20 |

Test the hypothesis that the two groups differ in the proportion of patients with hypertension and normal. (0.0050, Reject H_0)

3. Interview with 12 persons engaged in a stressful occupation revealed the following classification.

| Mental status | Alcoholic | Non alcoholic | Total |
|---------------|-----------|---------------|-------|
| Depressed | 1 | 6 | 7 |
| Not depressed | 4 | 1 | 5 |
| Total | 5 | 7 | 12 |

Do the data provide sufficient evidence to conclude an association between alcoholism and mental depression? (0.0905, Accept H_0)

4. Following table gives the number of births according to their sex and condition at the time of birth.

| Sex | Normal | Abnormal | Total |
|--------|--------|----------|-------|
| Male | 2 | 8 | 10 |
| Female | 5 | 10 | 15 |
| Total | 7 | 18 | 25 |

Test at $\alpha = 0.05$, whether the condition at birth depends in the sex of the child.

5. A sociologist studying rehabilitation of paraplegics examines cases of marriage after suffering the physical disability. The idea is that financial status of the crippled individual may have some relation with whether the spouse is physically handicapped or normal. The data are given below.

| Spouse | Financial status of patients | | Total |
|------------------------|------------------------------|------|-------|
| | Rich | Poor | |
| Normal | 12 | 3 | 15 |
| Physically handicapped | 2 | 2 | 4 |
| Total | 14 | 5 | 19 |

(0.272, accept H_0)

McNemar Test

McNemar test is designed to see the effect of some treatment, training or advertisement which brings about the change in attitude of individuals such as: change on voters preferences before and after TV debate; change in smoking habit before and after meditation etc.

Layouts

- a) Consider an independent bivariate random variable (x_i, y_i) ; $i = 1, 2, \dots, n$ whose frequencies of type (+) and (-) signs are classified into 2×2 contingency table as shown below.

| Before meditation | After meditation | | Total |
|-------------------|------------------|------------|-------|
| | Non smoker (-) | Smoker (+) | |
| Smoker (+) | A | b | a + b |
| Non smoker (-) | C | d | c + d |
| Total | a + c | b + d | N |

- b) Obviously, responses changed from (+) to (-) and (-) to (+) are included in the cell 'a' and 'd' respectively. Responses of no changes are enclosed in the cells 'b' and 'c'.

Assumptions

- The pairs (x_i, y_i) mutually independent.
- The measurement scale is either nominal or ordinal.

Mechanism

- Count the total number of frequencies changed from (+) to (-) and (-) to (+) i.e (a + d).
- Effective sample size = (a + d)
- Compute the expected frequencies to cells 'a' and 'd' as $\frac{a+d}{2}$.

Null and alternative hypothesis

H_0 : $\Pr(a) = \Pr(d)$; meditation has no effect.

$H_1: \Pr(a) \neq \Pr(d)$, meditation has effect.

Test statistic

$$\chi^2 = \frac{(|a - d| - 1)^2}{a + d} \text{ with d.f} = 1$$

Critical region

Fix a level of significance α , and obtain the tabulated value of test statistics from table (Chi square table)

Give decision

If χ^2 calculated $\leq \chi^2$ tabulated, null hypothesis is accepted otherwise rejected.

Exercise

1. A medical researcher wishes to determine if a pill has the undesirable side effect of reducing the blood pressure of the user. The study involves recording the initial blood pressures 30 college age women. After they use the pill regularly for six months, their blood pressures are again recorded. The researchers wish to draw inferences about the effect of the pill on blood pressure from the observations given below.

| B. P Before | B.P After | | Total |
|-------------|-----------|-----|-------|
| | (-) | (+) | |
| (+) | 12 | 5 | 17 |
| (-) | 5 | 8 | 13 |
| Total | 17 | 13 | 30 |

Do the data substantiate the claim that use of the pill reduces blood pressure? Test at $\alpha = 0.01$ (0.8, accept H_0)

2. A newly invented drug is administered to a sample of 150 patients in a random manner and hypertensive cases, a common symptom among females using oral contraceptive, were classified into high and low level.

| Level before administration | Level after administration | | Total |
|-----------------------------|----------------------------|------|-------|
| | Low | High | |
| High | 20 | 60 | 80 |
| Low | 30 | 40 | 70 |
| Total | 50 | 100 | 150 |

Is the drug an effective means of controlling hypertensive level? Test at 5% level of significance. (6.02, reject H_0)

3. Leprosy is a disease with considerable prejudice and stigma attached to it. Education is one means by which this situation may be improved. A film prepared for a mass education program us to be tested for efficacy. The following experiment is conducted. A sample of 100 individuals is asked to choose between agree and disagree regarding the statement. Leprosy is just like any other disease and can be cured. After the options are recorded, the film is shown to the group and again the same question is asked. Thus

two responses are recorded on each individual, one before viewing the film and one afterwards. The data are given in the following 2×2 table, which gives the reaction to the statement before and after viewing the film.

| Before | After | | Total |
|----------|----------|-------|-------|
| | Disagree | Agree | |
| Agree | 5 | 33 | 38 |
| Disagree | 40 | 22 | 62 |
| Total | 45 | 55 | 100 |

Does the information provide sufficient evidence to conclude that the film has the effect of making people more favorable to the statement? (9.481, Reject H_0)