→ the candidate is more popular among female candidate.

Chi square test of popⁿ variance $\sigma^2$

Function of the test

This test is used to determine wheather the popⁿ variance (or standard deviation) of a variable is equal to some specified value or not.

Test Assumption

1. Variable of interest is normally distributed in the population
2. Variable is numerical.
3. Sample is drawn randomly from the popⁿ so that sample is representative.

Hypothesis to test

Let $\sigma^2$ be popⁿ variance of the process.

Null Hypothesis

$H_0: \sigma^2 = \sigma_0^2$

$(H_0: \sigma = \sigma_0)$

Alternative Hypothesis

$H_1: \sigma^2 \neq \sigma_0^2$

or, $H_1: \sigma^2 < \sigma_0^2$

or, $H_1: \sigma^2 > \sigma_0^2$

## Test Statistic

The appropriate test statistic is given by,

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

where,

$n$ = Sample size

$s^2$ - Sample S.D

$\sigma_0^2$ = Hypothesized value of $\sigma^2$ (Test value)

The test statistic follow $\chi^2$ distribution with $(n-1)$ degree of freedom.

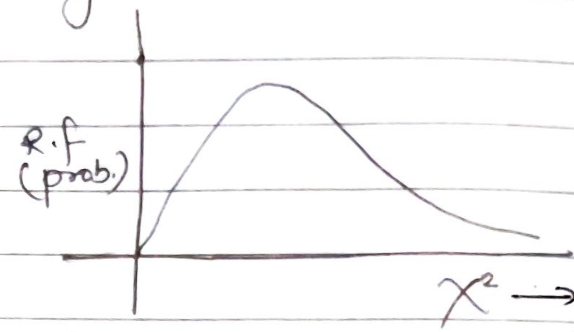
### Decision Rule

| Hypothesis | Decision Rule |
|---|---|
| 1. $H_0: \sigma^2 = \sigma_0^2$ <br> $H_1: \sigma^2 \neq \sigma_0^2$ | 1. Accept $H_0$ if $\chi^2_{\alpha/2}(n-1) < \chi^2 < \chi^2_{1-\alpha/2}(n-1)$ otherwise reject |
| 2. $H_0: \sigma^2 \geq \sigma_0^2$ <br> $H_1: \sigma^2 < \sigma_0^2$ | 2. Reject $H_0$ if cal $\chi^2 \geq \chi^2_{1-\alpha}(n-1)$ |
| 3. $H_0: \sigma^2 \leq \sigma_0^2$ <br> $H_1: \sigma^2 > \sigma_0^2$ | 3. Reject $H_0$ if cal $\chi^2 \leq \chi^2_{\alpha}(n-1)$ |

Characterstic:

1. The characterstic distribⁿ is exact distribⁿ. There is a family chi-square curve for each sample size.
2. Chi-square distribⁿ is right skewed (Non-symmetric)
3. Chi-square value lies betⁿ 0 and ∞

$0 \leq \chi^2 \leq \infty$




# The variability of weights in 2 lbs. Packets of guaranteed is expressed by a S.D of 0.05 oz. To test this, a sample of 25 packet was picked and weighed giving following results in (in oz.)

| | | | | |
|---|---|---|---|---|
| 32.11 | 31.97 | 32.18 | 32.05 | 32.16 |
| 32.25 | 32.07 | 32.07 | 32.15 | 32.03 |
| 32.05 | 32.14 | 32.19 | 31.96 | 31.98 |
| 32.07 | 31.99 | 32.09 | 32.08 | 32.03 |
| 32.16 | 32:09 | 32.18 | 32.64 | 31.98 |

Perform a chi-square to test the claim at $\alpha = 0.05$ level.

Soln:-

Step 1: Null and Alternative Hypothesis

Here, X = Weight of packets of guaranteed food (oz).

Let $\sigma^2$ be the variability of weight of packets

<u>Null Hypothesis</u>

$H_0 : \sigma^2 \leq (0.05)^2 \leq 0.0025$ (Production running smoothly) (confirmation to quality)

<u>Alternative Hypothesis</u>

$H_1 : \sigma^2 > 0.0025$

<u>Step 2</u> choice of $\alpha$

<u>Step 3</u> : Test statistic

The appropriate test statistic is given by,

$$\chi^2 = \frac{(n-1)\, s^2}{\sigma_0^{\,2}}$$

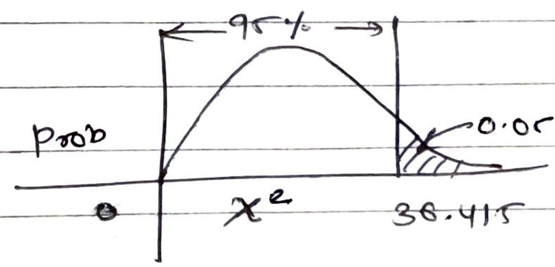
<u>Step 4</u> : Critical $\chi^2$ (Tabulated $\chi^2$)

The test is right sided at $\alpha = 0.05$

$D.f = n - 1$

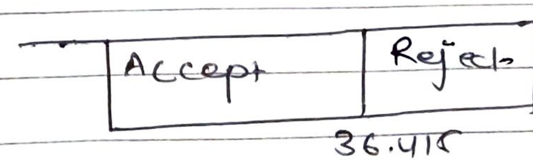
$= 25 - 1$

$= 24$




<u>From table of $\chi^2$ distribⁿ,</u>

$\chi^2_{0.05}(24) = 36.415$




$AR = \chi^2 < 36.415$

$RR : \chi^2 \geq 36.415$

Step 5 : Calculation of $\chi^2$

Here. $\Sigma x = 801.95$

$\Sigma x = 25725.0951$

$\bar{x} = 32.078$

$s = 0.007719$

Now,

$$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$

$$= \frac{(25-1) \times (0.07719)}{(0.05)^2} = 57.1996$$

Step 6 : Statistical Decision

Since, Cal $\chi^2 = 57.1996 >$ Tab $\chi^2 = 36.415$, we reject Ho at 5% level of significance.

Step 7: Conclusion

The S.D of weight of packets is actually higher than 0.05 lbs so that there is non-confiomation to the quality.

F - test of difference of two population variances

Function of the test :

The function of the test is used to test wheather the variances of two popⁿs. are equal or not.

This test may be used prior to t test for diffⁿ of two popⁿ means to know wheather variances are homogeneous or not.

Test Assumptions

1. Variable of interest is normally distributed in both popⁿs or groups. F-test is more sensitive to deviations from normality than t-test.
2. Samples are drawn randomly from each popⁿ or group
3. Samples are independent of each other.

Hypothesis to test

Let $\sigma_1^2$ and $\sigma_2^2$ be the variance of popⁿ 1 and popⁿ 2 respectively. The null and alternative hypothesis of the test is given by,

| Two tailed test | Left tailed test | Right tailed test |
|---|---|---|
| $H_0: \sigma_1^2 = \sigma_2^2$ | $H_0: \sigma_1^2 \geq \sigma_2^2$ | $H_0: \sigma_1^2 \leq \sigma_2^2$ |
| $H_1: \sigma_1^2 \neq \sigma_2^2$ | $H_1: \sigma_1^2 < \sigma_2^2$ | $H_1: \sigma_1^2 > \sigma_2^2$ |

Test Statistics

The appropriate test statistics for this test is given by,

$$F = \frac{S^2_{larger}}{S^2_{smaller}} = \frac{\text{Larger of two variances}}{\text{Smaller of two variances}}$$

Putting larger variances always on the numerator makes the ratio greater than one and one sided test always becomes right sided.

# Decision Rule

| Hypothesis | Decision Rule |
|---|---|
| 1. Two tailed test<br>$H_0 : \sigma_1^2 = \sigma_2^2$<br>$H_1 : \sigma_1^2 \neq \sigma_2^2$ | 1. Reject $H_0$ if calculated F falls outside the range<br>$(F_{\alpha/2}$ or $F_L$, $F_{1-\alpha/2}$ or $F_U)$<br>Note: $F_L = \frac{1}{F_U}$ |
| 2. One tailed test | 2. Reject $H_0$ if calculated $F \geq F_\alpha(n_1 - 1, n_2 - 1)$ |

## Problem

1. Two types of instruments for measuring the amount of Sulphur monoxide in the atmosphere are being compared in an air pollution experiment. It is desired to determine wheather the two types of instruments yield measurements having the same variability. The following readings were recorded for the two instruments.

| Instrument A | Instrument B |
|---|---|
| 0.86 0.82 0.75 0.61 | 0.87 0.74 0.63 0.55 0.76 |
| 0.89 0.64 0.81 0.68 | 0.70 0.69 0.57 0.53 |
| 0.65 | |

Assuming the population of measurements to be approximately normally distributed test the hypothesis that $\sigma_A = \sigma_B$ against the alternative that $\sigma_A \neq \sigma_B$

2. In most colleges, it is desirable for the variances of exam grades to be nearly the same among instructors. College administration is interested in whether or not there is any variation in the way they grade math exams. They each grade the same set of 10 exams. The first instructors' grades have a variance of 52.3. The second instructor's grades have a variance of 89.9. Test the claim that the first instructor's variance is smaller using level of significance of 5%.

1. Soln:-

Step 1:- Setting up Null and Alternative Hypothesis

Let $X_1$ = measurement of sulphur dioxide in the ($X_A$) ~~absorption~~ atmosphere using instrument A

$X_2$ = " " " " " " " " B

$\sigma_1^2$ = Variance of measurements of sulphur dioxide in the atmosphere using instrument A

$\sigma_2^2$ = " " " " " B.

$H_0: \sigma_1^2 = \sigma_2^2$ (Two instruments are equally precise)

$H_1: \sigma_1^2 \neq \sigma_2^2$ (Precision of two instrument are different)

$\left( H_1 : \frac{\sigma_1^2}{\sigma_2^2} \neq 1 \right)$

Step 2: Choice of $\alpha$ for the test

$\alpha$ = level of significance employed in the test

= Prob {Type I error}

= 0.05

Step 3: Test Statistic

The appropriate test statistic is given by,

$$F = \frac{S^2_{larger}}{S^2_{smaller}} = \frac{\text{Larger variance}}{\text{Smaller variance}}$$

The test statistic has F distribution with $n_1 - 1$ diffⁿ in numerator and $n_1 - 1$ diffⁿ in denominator.

$n_1$ = Size of the sample which gives larger variance

$n_2$ = Size of " " smaller variance

Step 4: Calculated F (observed F)

| Instrument A | Instrument B |
|---|---|
| $\Sigma x_1 = 6.71$ | $\Sigma x_2 = 6.04$ |
| $\Sigma x_1^2 = 5.0893$ | $\Sigma x_2^2 = 4.1534$ |
| $\overline{x_1} = 0.7456$ | $\overline{x_2} = 0.6711$ |
| $S_1^2 = 1.1267$ | $S_2^2 = 0.01249$ |
| (Smaller variance) | (Larger variance) |

$\therefore$ cal F $= \dfrac{S^2 \text{ larger}}{S^2 \text{ smaller}} = \dfrac{0.01249}{0.01083} = 1.15$

Step 5 : Critical f

The test is two-sided and $\alpha = 5\%$.

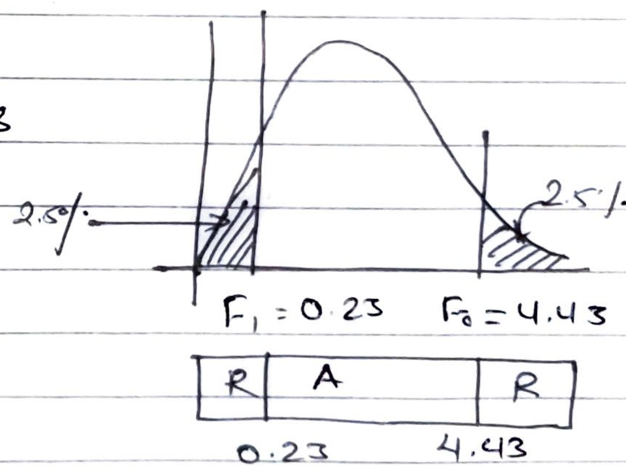
Numerator d.f $= 9-1 = 8$

Denominator d.f $= 9-1 = 8$

From the table,

$F_{0.05}(8,8) = F_0$

$= 4.43$




Step 6 : Statistical Decision

Since, cal $F = 1.15$ lies in the acceptance region AR $(0.23 < F < 4.43)$, we donot reject $H_0$ at 5% level of significance.

Step 7: Conclusion

Two instruments are equally precise in measuring the amount of sulphur monoxide in the atmosphere.

## Non - Parametric Hypothesis Testing

Defn:-

Non-parametric tests are distribution free tests because they dont assume that your data follow specific distribution. One of the fundamental assumption of parametric test is that data is approximated normally distributed. Further, Non-parametric procedure can be used to test hypotheses that do not involve specific popn parameters such as $\mu$, $\sigma$ and $\pi$ etc. Parametric test involves specific probability distributions (eg., the normal distribution) & the tests involve estimation of the key parameters of that distribution (eg:- the mean) from the sample data.

Advantages of Non-parametric method

- They can be used to test popn parameters when the variable is not normally distributed.
- They can be used even when the data are categorical which are measured either on nominal or ordinal scale.
- They can be used to test hypotheses that donot involve popn parameters.

- In most cases, the computations are easier than those for the corresponding parametric methods.
- They are easy to understand and apply.

## Disadvantages of NP Method

- Non-parametric tests tend to use less information than the parametric tests because exact numerical data are often reduced to a qualitative form for example ranks, categories such as signs (+/-), low/moderate/high etc. When we do that we waste information.

- They are less sensitive than their parametric counterparts when the assumptions of the parametric methods are met. Therefore, larger differences are needed ~~to~~ before the null hypothesis can be rejected.

- They are less efficient than their parametric counterparts when the assumptions of the parametric methods are met. That is, larger sample sizes are needed to overcome the loss of information. For example, the non-parametric sign test is about 60% as efficient as its parametric counterpart, the z-test. Thus, a sample size of 100 is needed for the use of the sign test, compared with a sample size of 60 for use of the z test to obtain the same results.

# Differences Between Parametric and Non-Parametric Methods

* Scale of Measurement :- Most parametric methods are relev for data measured on a quantitative (interval or ratio) scale. Nonparametric tests are primarily concerned with nominal or ordinal data taken from a (typically) continuous popⁿ distribution.

* Distribution Assumption :- Fundamental assumption of all parametric tests is that data comes from the normally distributed popⁿ. Non-parametric tests do not require that the probability distribution of some popⁿ characteristics assume any functional form.

* Sample Size :- PT are used only when sample size requirement is met and we have confidence that data is normally distributed. NPT is suitable for making inferences about popⁿs with relatively large sample size.

* Degree of information in decision :- Conclusions obtained from parametric hypothesis tests have a high degree of information content.

* Statistical Power :- In instances where both parametric and non-parametric techniques apply (i.e. the classical assumptions hold), the power of a parametric test

(is greater than that of its non parametric counterpart given $n$, $\alpha$ and the true situation).

## Ranking

Many non-parametric tests involve the ranking of data, that is, the positioning of a data value in a data array according to some rating scale. A rank is a number assigned to an individual sample item according to its order in the sorted list. The first item is assigned to rank of 1, the second item is assigned a rank of 2, and so on. Numerical data when converted to rank become ordinal data. In doing so, we lose information.

## Order Statistics

Order statistics and rank statistic are among the most fundamental tools used in the non-parametric statistic and inference. The order statistic of a random sample $X_1, X_2, \ldots\ldots X_n$ are the sample values placed in ascending order. They are denoted by $X_{(1)}, X_{(2)}, \ldots\ldots, X_{(n)}$.

$X_{(1)}$ = First order statistics = Min $X_i$, $i = 1, 2, \ldots, n$

$X_{(2)}$ = Second order statistics

....

$X_{(n)}$ = $n^{th}$ order statistic = Max $X_i$

The $k^{th}$ order statistic of a sample is equal to its $k^{th}$ smallest value. The order statistic are random variable themselves and satisfies the following relationship:-

$$X_{(1)} \leq X_{(2)} \leq \ldots\ldots \leq X_{(n)}$$

The following are some statistics that are easily defined in terms of the order statistics :-

(1) Sample Range

Sample range(R) = $X(n) - X(1)$

(2) Sample median

Median(md) = $X_{n+1/2}$ if n is odd (here n is order statistic)

$= \dfrac{X_{n/2} + X_{n/2+1}}{2}$ if n is even (here n is a fxn of two order statistic)

List of Non-Parametric Tests

| Non-parametric Test | Parametric test (Equivalent test) |
|---|---|
| Run test for randomness | Not available |
| Sign test (single sample/paired sample) | t test for mean single sample /paired t test |
| Wilcoxon signed rank test (single sample / paired sample) | " " " |
| Wilcoxon Rank Sum Test (Mann-Whitney Test) | t test for diffⁿ of means, two ind. samples |
| Krustikal Wallis Test | One way ANOVA test |
| Friedman test | Repeated measure ANOVA test |
| Cochran's Q test | " " " " |
| Median test | t test for mean (2 samples case), One way ANOVA ($\geq 3$ samples case) |
| Chi-square test of goodness of fit | Not available |
| Chi-square test of independence of categorical values | Not available |
| Chi-square test of homogeneity of popⁿ wrt some attribute | Not available |
| Kolmogorov-Smirov Test (one-sample case / Two sample case) | |

# Run test of randomness

## Function of the test

One of the most important assumption of all types of statistical test is that ~~is~~ ~~the~~ the sample must be representative of the population as possible. It is because the sample result used in decision regarding the population from which the sample is drawn. This requires that the sample is random. The randomness of the data can be tested by simple run test.

A run test can be used in two cases:-

1. To test the randomness of nominal scale categorical data.
2. To test the randomness of numerical data.

## What is run?

→ A run is defined as a sequence of identical symbols. It is a sequence of like elements bounded on both sides by either unlike elements or no elements.

$R$ = Population no. of runs

$r$ = sample no. of runs.

## Example 1

A coin is tossed 20 times and sequence of head and tail is given as follows:-

$\underline{H}\ \underline{TT}\ \underline{HH}\ \underline{TTT}\ \underline{H}\ \underline{T}\ \underline{H}\ \underline{TT}\ \underline{HH}\ \underline{T}\ \underline{H}\ \underline{T}\ \underline{HH}$

1 2 3 4 5 6 7 8 9 10 11 12 13

Find the no. of runs.

Soln:-

No. of runs $(r) = 13$

Few runs → not random

Too many run → not random

r distribution

HHH TTT HHH TTT

HTHT HTHTHT HT

## Example 2

Marks obtained by students

92 32 48 69 77 13 52

Find the no. of runs.

Soln:-

Use '−' sign if data $\leq$ Md

Use '+' sign if data $>$ Md

The given sequence is,

92 32 48 69 77 13 52

Array:- 13 32 48 52 69 77 92

$$\text{Median} = X\left(\frac{n+1}{2}\right) = X(4) = 52$$

| Data | Sign | |
|---|---|---|
| 90 | + | 1 |
| 32 | − | 2 |
| 48 | − | |
| 69 | + | 3 |
| 77 | + | |
| 13 | − | 4 |
| 52 | − | |

Hypothesis

$H_0$ : Data is in random order
$H_1$ : Data is not in random order } Two tailed test

$H_0$ : Data is in random order
$H_1$ : Too few runs } Left tailed test.

$H_0$ : Data is in random order
$H_1$ : Too many runs } Right tailed test.

Test Statistics

(i) Small sample case : ($n_1 < 20$ and $n_2 < 20$)

In this case, we first find the $n_1$ and $n_2$, where $n_1$ = no. of elements of first type and $n_1$ = no. of elements of second type. If both $n_1$ and $n_2$ are smaller than 20 or one of them is smaller than 20, we consider it as a small sample case. In this case, we use exact distribution of r.

Test statistic = No. of runs in the sequence (r)

(ii) Large sample case: ($n_1 \geq 20$ and $n_2 \geq 20$)

When $n_1$ and $n_2$ increase in size, the sampling distribution of r approaches to a normal probability distribⁿ with,

$$\text{mean}(\mu_r) = \frac{2n_1 n_2}{n} + 1$$

$$\& \text{ standard deviation } (\sigma_r) = \sqrt{\frac{(2n_1n_2)(2n_1n_2 - n)}{n^2(n-1)}}$$

The z- transformation of r is given by,

$$z_r = \frac{r - \mu_r}{\sigma_r}$$

The statistics z follow standard normal distribⁿ with mean 0 and SD of 1.

Test Statistic = Standardized r-value

Decision Rule

Small sample case:

| Hypothesis | Decision Rule |
| --- | --- |
| Case I: Two-tailed test | Accept $H_0$ if $r_l < r < r_u$ otherwise reject $H_0$ |
| Case II: Left tailed test | Accept $H_0$ if $r > r_l$ otherwise reject $H_0$ |
| Case III: Right tailed test | Accept $H_0$ if $r < r_u$ otherwise reject $H_0$ |

## Large sample case

For large sample case, we use Z approximation for r and use the following decision rule.

| Hypothesis | Decision Rule |
| --- | --- |
| Case I | Accept $H_0$ if $-Z_{\alpha/2} < Z_r < +Z_{\alpha/2}$ otherwise reject $H_0$ |
| Case II | Reject $H_0$ if $Z_r \leq -Z_\alpha$ |
| Case III | Reject $H_0$ if $Z_r \geq +Z_\alpha$ |

Problems

1. What is a run test? At a small soda factory, the amount of soda put into each 12-ounce bottle by the bottling machine varies slightly for each filling. The plant manager suspects that the machine has a random pattern of overfilling and underfilling the bottles. The following are the results of filling 18 bottles, where O denotes 12 ounces or more of soda in a bottle and U denotes less than 12 ounces of soda.

U U U O O O O U U O O
O U O U U U U

Using the runs test at 5% significance level, can you conclude that there is a non-random pattern of overfilling and underfilling such bottles?

2. The students in a statistics class were asked if they could be a good random number generator. Each student was asked to write down a single digit from 0 through 9. The data were collected starting at the front left of the class, moving row by row, to the back right of the class. The sequence of digits was as follows:-

7 4 3 6 9 5 4 4 4 3 6 3 3 7 7 7 6 3 6 7 6 9 6 7

3 7 7 3 4 6

Do these data show a randomness about the median value at 5% level of significance?

3. The following arrangement of men (M) and women (W) lined up to purchase tickets for a rock concert.

M W M W M M M W M W M M M W W M M M M W W

M W M M M W M M M W W W M W M M M

W M W M M M M W W M M W

Test for randomness of the arrangement at $\alpha = 0.05$ level of significance.

Soln:-

~~Step 1: Setting up Null and Alternative Hypothesis~~

~~Let $X_1$ = Measurements of $SO_2$ in the ($X_A$) atmosphere using instrument A~~

~~$X_2$ = " " " " ($X_B$) atmosphere using " " B~~

~~$\sigma_1^2$ = Variance of measurement of $SO_2$ in the atmosphere using instrument A~~

~~$\sigma_2^2$ = " " " " " " " B~~

## Run test for randomness

If $n_1 < 20$, $n_2 < 20$ (Small sample case)

$n_1 \geq 20$, $n_1 \geq 20$ (Large sample case)

1. Soln:

Step 1: Null and Alternative Hypothesis

X = Underfilled (U) / overfilled (O)

$H_0$: The sequence of underfill and overfill is in random order

$H_1$: The sequence of underfill and overfill isn't in random order

Step 2: Level of significance

Here, $\alpha = 5\% = 0.05$

Step 3: Test statistic

The given sequence is,

UUU OOOO UU OOO U O UUUU

$n_1$ = No. of units in first category = 10

$n_2$ = No. of units in second category = 8

∴ Test statistic = No. of runs in the sequence.

Step 4: Critical r

Here, $n_1 = 10$

$n_2 = 8$

$\alpha = 0.05$

From the table of r- distribution

$r_L$ = lower critical $r = 5$

$r_U$ = Upper critical $r = 15$

AR: $6 \leq r \leq 14$

RR: $r \geq 15$ or $r \leq 5$

Step 5: Observed r (Calculated r)

No. of runs in sequence $(r) = 7$

Step 6 : Statistical Decision

Since, cal $r = 7$ false in AR( $6 \leq r \leq 14$ )

We accept $H_0$ at 5 % level of significance.

Step 7 : Conclusion

The underfilling and overfilling of machine is in random order.

# Sign Test

## Function of the test

The sign test is the location test. It is used in two hypothesis testing situations.

(1) Sign test for single sample
It is used to determine wheather the popⁿ from which the sample is drawn has some specific median value $M_0$ or not.

$H_0: M = M_0$

(2) Paired sample case
It is used to determine wheather two dependent samples or paired sample have same median value or not.

## Test Assumptions

(1) The sample is drawn randomly from the popⁿ
(2) Variable is at least ordinal

## Hypothesis to test

(i) Single sample case

$H_0: M = M_d$ (Popⁿ has specific median value)

Against

$H_1: M \neq M_d$

OR,

$H_1: M < M_d$

OR,

$H_1 : M > M_d$

(ii) Paired sample case

$H_0 : M_1 = M_2$

Against,

$H_1 : M_1 \neq M_2$

OR,

$H_1 : M_1 < M_2$

OR,

$H_1 : M_1 > M_2$

Test Statistic

(1) Small sample case ($n < 25$)
(Exact test)

(2) Large sample case ($n \geq 25$)
($Z$ distribution)

Small sample case

Step 1. Find difference $D_i = X_i - M_d$ (single sample case and $D_i = X_{i1} - X_{i2}$ (Paired sample case)

Step 2. Find the no. of '+' sign and '−' sign.

$S^+$ = No. of '+' sign

$S^-$ = No. of '−' sign

If there are tied observation, omit the data and reduce the sample size accordingly.

Step 3 :- Test Statistic

Small sample ~~test~~ case

Test statistic $S = \text{Min}\{S^+, S^-\}$ Two-tailed test

Test statistic $S = S^+$ left sided test

Test statistic $S = S^-$ Right sided test

$H_1 : M < M_d$

Large sample case

Convert S into Z

We know the distribution S is Binomial.

If n is large distribution of S converges to Z distribution.

$\mu_S$ = Mean of $S = n \times p = n \times \frac{1}{2} = \frac{n}{2} = 0.5n$

$\sigma_S = \sqrt{n \times p \times q} = \sqrt{n \times \frac{1}{2} \times \frac{1}{2}} = \frac{\sqrt{n}}{2} = 0.5\sqrt{n}$

Then, Z transformation of S,

$$Z = \frac{S - \mu_S}{\sigma_S} = \frac{S - 0.5n}{0.5\sqrt{n}}$$

Decision Rule (small)

Case I

$H_0 : M = M_0$ vs $H_1 : M \neq M_0$

$H_0 : M_1 = M_2$ Vs $H_1 : M_1 \neq M_2$

Reject $H_0$ if cal $S = \text{Min}\{S^+, S^-\} < S_\alpha(n)$

$S_\alpha(2)(n)$

Case II

Reject $H_0$ if cal $S = S^+ < S_\alpha(n)$

$S_\alpha(1)(n)$

Case III

Reject $H_0$ if cal $S = S^- < S_\alpha(n)$

Date
Page

## Decision Rule (Large)

Case I   Reject Ho if cal z lie outside the interval $-z_{\alpha/2} < z < +z_{\alpha/2}$

Case II   Reject Ho if cal $z < -z_{\alpha}$

Case III   Reject Ho if cal $z > +z_{\alpha}$

## Numerical

The following data in (tons) are the amt. of $SO_2$ emitted by a large industrial plant in 40 days.

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 24 | 15 | 20 | 29 | 19 | 18 | 22 | 25 | 27 | 9 |
| 17 | 20 | 17 | 6 | 24 | 14 | 15 | 23 | 24 | 26 |
| 19 | 23 | 28 | 19 | 16 | 22 | 24 | 17 | 20 | 13 |
| 19 | 10 | 23 | 18 | 31 | 13 | 20 | 17 | 24 | 14 |

Use a sign test to test the hypothesis that median $SO_2$ level is 21 against more at 5% level of significance.

Soln:-

Step 1 : Null and Alternative Hypothesis

Here, X = Amount of $SO_2$ emitted by the large industry

M = Median amount of $SO_2$ emitted by the large industry

$H_0 : M \le 21$

$H_1 : M > 21$

Step2: Level of significance

$\alpha = 5\%$ (given)

Step 3: <u>Test Statistic</u>

Since $n \geq 25$; we consider large sample case.

$$\text{Test statistic } (z) = \frac{S - \mu_s}{\sigma_s}$$

where, $S$ = Smaller no. of plus or minus sign

$\mu_s =$

$\sigma_s =$

Step 4: <u>Observed z</u>

| S.N | Data ($X_i$) | $D_i = X_i - 21$ | Sign | |
|---|---|---|---|---|
| 1. | 24 | +3 | + | 1 |
| 2. | 15 | −6 | − | 2 |
| 3. | 20 | −1 | − | |
| 4. | 29 | 8 | + | 3 |
| 5. | 19 | −2 | − | 4 |
| 6. | 18 | −3 | − | |
| 7. | 22 | 1 | + | 5 |
| 8. | 25 | 4 | + | |
| 9. | 27 | 6 | + | |
| 10. | 9 | −12 | − | |
| 11. | 17 | −4 | − | 6 |
| 12. | 20 | −1 | − | |
| 13. | 17 | −4 | − | |
| 14. | 6 | −15 | − | |
| 15. | 24 | 3 | + | 7 |
| 16. | 14 | −7 | − | 8 |

| | | | | |
|---|---|---|---|---|
| 17 | 15 | -6 | - | |
| 18 | 23 | 2 | + | 9 |
| 19 | 24 | 3 | + | |
| 20 | 26 | 5 | + | |
| 21 | 19 | -2 | - | 10 |
| 22 | 23 | 2 | + | 11 |
| 23 | 28 | 7 | + | |
| 24 | 19 | -2 | - | 12 |
| 25 | 16 | -5 | - | |
| 26 | 22 | 1 | + | 13 |
| 27 | 24 | 3 | + | |
| 28 | 17 | -4 | - | |
| 29 | 20 | -1 | - | |
| 30 | 18 | -3 | - | 14 |
| 31 | 19 | -2 | - | |
| 32 | 10 | -11 | - | |
| 33 | 23 | 2 | + | 15 |
| 34 | 18 | -3 | - | 16 |
| 35 | 31 | 10 | + | 17 |
| 36 | 13 | -8 | - | 18 |
| 37 | 20 | -1 | - | |
| 38 | 17 | -4 | - | |
| 39 | 24 | 3 | + | 19 |
| 40 | 14 | -7 | - | 20 |

classmate
Date
Page

No. of plus sign $(S^+) = 16$

No. of minus sign $(S^-) = 24$

Mean of the sign $(\mu_s) = 0.5\,n = 0.5 \times 40 = 20$

S.D of the sign $(\sigma_s) = 0.5 \times \sqrt{40} = 3.1623$

Now, the z-value is given by,

$$z = \frac{S - 0.5n}{0.5\sqrt{n}} = \frac{16-20}{3.1623} = -1.26$$

Step 5:- Tabulated z (Critical z)

The test is right sided and $\alpha = 0.05$

From the table of critical value of z distribution,

critical $z = +1.65$

AR: $z \leq +1.65$

RR: $z > +1.65$

Step 6. Statistical Decision

Since cal $z = -1.26$ is less than critical $z = +1.65$, we donot reject $H_0$ at 5% level of significance.

Step 7:- Conclusion

The amount of $SO_2$ produced by a large factory is 21 or less.

2 Soln:—

Given.

7 4 3 6 9 5 4 4 4 3 6 3 3 7 7 7 6 3 6 7 6 9

6 7 3 7 7 3 4 6

Step 1: Null and Alternative Hypothesis

$H_0$: The sequence of no. is in random order

$H_1$: The " " " isn't " " "

Step 2:

$\alpha = 0.05$

Step 3: Test statistic

The given sequence is: 7 4 3 6 9 5 4 4 4 3 6 3 3

7 7 7 6 3 6 1 6 9 6 7 3 7 7 3 4 6

Data Array:- 3 3 3 3 3 3 3 4 4 4 4 4 5 6 6 6 6

6 6 6 7 7 7 7 7 7 7 7 9 9

Use '−' sign if $x_i \le M_d$

Use '+' sign if $x_i > M_d$

$M_d = X\left(\frac{n+1}{2}\right) = X(15.5) = \frac{X_{(15)} + X_{(16)}}{2} = \frac{6+6}{2} = 6$

Sign:—

+ − − − − + − − − − − − − − − − + + + − − −

+ − + − + − + + − − −

$n_1$ = No. of '−' sign = 20

$n_2$ = No. of '+' sign = 10

∴ Test statistic = No. of runs(r)

Step 4 Critical r

$n_1 = 20$ From table,

$n_2 = 10$ $r_L = 3$

$\alpha = 0.05$ $r_U = 20$

AR : $10 \leq r \leq 19$

RR : $r \geq 20$ OR $r \leq 9$

| Reject | Accept | Reject |
|---|---|---|
| 10 | 14 | 19 |

Step 5 : Observed r

Observed no. of runs $(r) = 14$

Step 6 :- Statistical Decision

Accept $H_0$

Step 7 :- Conclusion

3. Soln :—

Step 1 :- Null and Alternative Hypothesis

$H_0$ = The gender of a person who lined up to purchase the tickets for a rock concert is in random order. (Sequence of men and women is in random order)

$H_1$ = The gender of a person who lined up to purchase tickets for a rock concert (Sequence is not random order)

Step 2 :- Choice of $\alpha$ for the test

$\alpha$ = Level of significance of the test

= Probability {Type I error}

= 0.05 (given)

**Step 3:- Test Statistic**

$n_1$ = No. of men in the line = 31

$n_2$ = No. of women in the line = 19

We use large sample approximation,

The test statistic is given by, $z = \dfrac{r - \mu_r}{\sigma_r}$

where, $r$ = No. of runs

$\mu_r$ = Mean of no. of runs $= \dfrac{2n_1 \times n_2}{n} + 1$

$\sigma_r$ = S.D of runs $= \sqrt{\dfrac{(2n_1 n_2)(2n_1 n_2 - n)}{n^2 (n-1)}}$

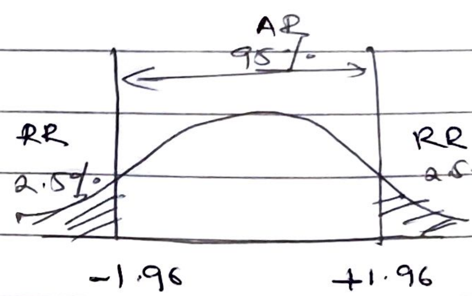
Test statistic follows SND with mean 0 & SD of 1.

**Step 4:- Tabulated z (Critical z)**

Test is two sided & $\alpha = 5\%$.

AR $= -1.96 < z < +1.96$

RR $= z \geq +1.96$ or $z \leq -1.96$




**Step 5:- Calculated z**

From the given sequence, no. of runs ($r$) = 28

$$\mu_r = \frac{2n_1 n_2}{n} + 1$$

$$= \frac{2 \times 31 \times 19}{50} + 1 = 24.5$$

$$\sigma_r = \sqrt{\frac{(2n_1 n_2)(2n_1 n_2 - n)}{n^2 (n-1)}}$$

$$= \sqrt{\frac{(2 \times 31 \times 19)(2 \times 31 \times 19 - 50)}{50^2 (50-1)}}$$

$= 3.293$

Again,

$$z = \frac{r - \mu_r}{\sigma_r} = \frac{28 - 24.5}{3.293} = 1.04$$

Step 6:- Statistical Decision

Since cal $z = 1.04$ which falls in the accepted region at 5% level of significance

Step 7:- Conclusion

Hence, we can conclude that the sequence of men and women is in random order.

## Wilcoxon Signed rank test

Function of the test:-

The Wilcoxon signed rank test is the non-parametric alternative to t-test for single sample or paired t test (two related samples)

This test has same function as sign test but this test is better than its since this test considers both sign of the difference and the magnitude of the difference.

Test Assumptions:-

(1) The sample is drawn randomly from the popⁿ

(2) The measurement scale is at least ordinal. (But scale data is most preferred).

$M_1$ = Score before condition / treatment
$M_2$ = " after " "

## Hypothesis to test

**One sample case:-**

$H_0 : M = M_0$

Against,

$H_1 : M \neq M_0$

or $H_1 : M < M_0$

or $H_1 : M > M_0$

**Paired sample case**

$H_0 : M_1 = M_2$

Against,

$H_1 : M_1 \neq M_2$

or $H_1 : M_1 < M_2$

or, $H_1 : M_1 > M_2$

Test Statistic :- (Small sample case $n < 25$)

Steps:-

Step 1: Compute the diff $D_i = X_i - M_0$ (single sample case)

or, $D_i = X_{1i} - X_{2i}$ (Paired sample case)

Step 2: find the absolute diff of $D_i$ i.e. $|D_i|$
(If there are ties in the difference i.e $D_i = 0$, reduce the sample size accordingly)

Step 3: Rank the absolute diff from lowest to highest
In case of ties, give the average rank)

Step 4: Give the ranks '+' sign or '−' sign according to sign in the original difference.

Step 5: Find the sum of positive ranks and sum of negative ranks.

Let $W^+$ = sum of positive ranks
$W^-$ = Sum of negative ranks

## Decision Rule :-

| Hypothesis | Decision Rule |
| --- | --- |
| Case I (Two-tailed test) | Reject Ho if cal $w = \text{Min}\{w^+, \lvert w^- \rvert\} \leq w_{\alpha(2)(n)}$ |
| Case II (Left tailed test) | Reject Ho if $w^+ \leq w_{\alpha(1)(n)}$ |
| Case III (Right tailed test) | Reject Ho if $\lvert w^- \rvert \leq w_{\alpha(2)(n)}$ |

Q. A food inspector examines 16 jars of a certain brand of jam to determine the % of foreign impurities. The following data were recorded :-

2.4 2.3 3.1 2.~~8~~ 2.3 1.2 1.0 2.4 1.7 1.1

4.2 1.9 1.7 3.6 1.6 2.3

Perform a Wilcoxon signed rank test at $\alpha = 5\%$ level of significance to test the hypothesis that the mdn % of the impurities in this brand of jam is 2.5% against ~~not~~. nut.

Soln:-

Step 1 :- Null and Alternative Hypothesis

X = Percentage of impurities in the given brand of the jam.

Let M = Median percentage " " "
" " "

$H_0 : M = 2.5\%$

$H_1 : M \neq 2.5\%$

Step 2:- Level of significance

Here, $\alpha = 5\%$ (given)

Step 3: Test Statistic

Since, $n < 25$, we consider a ~~sm~~ small sample case and test is,

$$W = \text{Min}\{W^+, |W^-|\}$$

Step 4: Observed W (Calculated W)

| S.N. | $X_i$ (Data) | $D_i = X_i - 2.5$ | $\lvert D_i \rvert$ | Rank | Signed Rank |
|---|---|---|---|---|---|
| 1. | 2.4 | −0.1 | 0.1 | 1.5 | −1.5 |
| 2. | 2.3 | −0.2 | 0.2 | 4 | −4 |
| 3. | 3.1 | 0.6 | 0.6 | 7.5 | 7.5 |
| 4. | 2.2 | −0.3 | 0.3 | 6 | −6 |
| 5. | 2.3 | −0.2 | 0.2 | 4 | −4 |
| 6. | 1.2 | −1.3 | 1.3 | 13 | −13 |
| 7. | 1.0 | −1.5 | 1.5 | 15 | −15 |
| 8. | 2.4 | −0.1 | 0.1 | 1.5 | −1.5 |
| 9. | 1.7 | −0.8 | 0.8 | 9.5 | −9.5 |
| 10. | 1.1 | −1.4 | 1.4 | 14 | −14 |
| 11. | 4.2 | 1.7 ~~0.3~~ | 1.7 ~~0.3~~ | 16 | 16 |
| 12. | 1.9 | −0.6 | 0.6 | 7.5 | −7.5 |
| 13. | 1.7 | −0.8 | 0.8 | 9.5 | −9.5 |
| 14. | 3.6 | 1.1 | 1.1 | 12 | 12 |
| 15. | 1.6 | −0.9 | 0.9 | 11 | −11 |
| 16. | 2.3 | −0.2 | 0.2 | 4 | −4 |

$W^+ =$ Sum of positive ranks $= 35.5$
$W^- =$ Sum of negative ranks $= -100.5$

$|W^-| = 100.5$

Calculated Test Statistic

$W = \text{Min}\{W^+, |W^-|\}$
$= \text{Min}\{35.5, 100.5\}$
$= 35.5$

Step 5:- Tabulated W (Critical W)
Test is two sided and $\alpha = 5\%$.
From the table,
$W_{\alpha(2)}(16) = 29$

Step 6: Statistical Decision
Since, cal $W = 35.5 >$ critical $W = 29$, we donot reject $H_0$ at $\alpha = 5\%$

Step 7: Conclusion
Median impulsion of certain brand of jam is not different from $2.5\%$.

# Wilcoxon Signed Rank Test
## (large sample case, $n \geq 25$)

When $n$ is large, the distribution of $W$ approaches to Z ie, distribution of $W$ can be approximated by Z distribution.

The Z transformation of $W$ is given by,

$$Z = \frac{W - \mu_W}{\sigma_W}$$

where,

$W$ = smaller of two sums $W^+$ or $|W^-|$

$\mu_W$ = mean of $W$

$$= \frac{n(n+1)}{4}$$

$\sigma_W$ = S.D of $W$

$$= \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Q. Two nodes of machines are under consideration for purchase. An organization has one of each type for trial. Each operator, out of the team of 25 operator uses each machine for a fixed length of time. There outputs are as follows

| Operation No. | Machine I | Machine II |
|---|---|---|
| 1 | 82 | 80 |
| 2 | 68 | 71 |
| 3 | 53 | 46 |
| 4 | 75 | 58 |
| 5 | 78 | 60 |
| 6 | 86 | 72 |
| 7 | 64 | 38 |
| 8 | 54 | 60 |
| 9 | 62 | 65 |
| 10 | 70 | 64 |
| 11 | 51 | 38 |
| 12 | 80 | 79 |
| 13 | 64 | 37 |
| 14 | 65 | 60 |
| 15 | 70 | 73 |
| 16 | 55 | 48 |
| 17 | 75 | 58 |
| 18 | 64 | 60 |
| 19 | 72 | 76 |
| 20 | 55 | 60 |
| 21 | 70 | 50 |
| 22 | 45 | 30 |
| 23 | 64 | 30 |
| 24 | 58 | 55 |
| 25 | 65 | 60 |

Is there any significant difference betw output capacities of two machines? Use Wilcoxon signed rank test. Take $\alpha = 5\%$.

Soln:-

Step 1: Null and Alternative Hypothesis

Let $X_1$ = No. of outputs produced by machine I

$X_2$ = No. of output produced by machine II

Let $M_1$ and $M_2$ are the median no. of outputs produced by machine I and machine II respectively.

$H_0: M_1 = M_2$ (Output capacities of two machines are

$H_1: M_1 \neq M_2$ ( " " " " " are

Step 2: Level of significance

$\alpha = 5\%$ (Given)

Step 3: Test Statistics

Here, $n = 25$, we consider the large sample case

The test statistic is given by,

$$Z = \frac{W - \mu_W}{\sigma_W}$$

where,

$W$ = smaller of two sums $W^+$ or $|W^-|$

$\mu_W$ = Mean of $W$

$$= \frac{n(n+1)}{4}$$

$$\sigma_W = \text{S.D of } W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Step 4: Calculated $w$ (Observed $w$)

| Operator No. | $X_1$ | $X_2$ | $D_i = X_i - X_2$ | $\lvert D \rvert$ | Rank of $\lvert D \rvert$ | Signed Rank |
|---|---|---|---|---|---|---|
| 1. | 82 | 80 | 2 | 2 | 2 | 2 |
| 2. | 68 | 71 | -3 | 3 | 4.5 | -4.5 |
| 3. | 53 | 46 | 7 | 7 | 14.5 | 14.5 |
| 4 | 75 | 58 | 17 | 17 | 19.5 | 19.5 |
| 5. | 78 | 60 | 18 | 18 | 21 | 21 |
| 6. | 86 | 72 | 14 | 14 | 17 | 17 |
| 7 | 64 | 38 | 26 | 26 | 23 | 23 |
| 8. | 54 | 60 | -6 | 6 | 12.5 | -12.5 |
| 9. | 62 | 65 | -3 | 3 | 4.5 | -4.5 |
| chine are same) 10 | 70 | 64 | 6 | 6 | 12.5 | 12.5 |
| are not same) 11. | 51 | 38 | 13 | 13 | 16 | 16 |
| 12. | 80 | 79 | 1 | 1 | 1 | 1 |
| 13. | 64 | 37 | 27 | 27 | 24 | 24 |
| 14. | 65 | 60 | 5 | 5 | 10 | 10 |
| 15. | 70 | 73 | -3 | 3 | 4.5 | -4.5 |
| 16. | 55 | 48 | 7 | 7 | 14.5 | 14.5 |
| 17. | 75 | 58 | 17 | 17 | 19.5 | 19.5 |
| 18. | 64 | 60 | 4 | 4 | 7.5 | 7.5 |
| 19. | 72 | 76 | -4 | 4 | 7.5 | -7.5 |
| 20. | 55 | 60 | -5 | 5 | 10 | 10 |
| 21. | 70 | 50 | 20 | 20 | 22 | 22 |
| 22. | 45 | 30 | 15 | 15 | 18 | 18 |
| 23. | 64 | 30 | 34 | 34 | 25 | 25 |
| 24. | 58 | 55 | 3 | 3 | 4.5 | 4.5 |
| 25. | 65 | 60 | 5 | 5 | 10 | 10 |

Here,

$w^+$ = Sum of positive ranks = 273

$w^-$ = " " negative ranks = −52

$|w^-| = 52$

$w$ = Smaller of $w^+$ and $|w^-| = 52$

$$\mu_w = \frac{n(n+1)}{4} = \frac{25 \times 26}{4} = 162.5$$

$$\sigma_w = \sqrt{\frac{n(n+1)(2n+1)}{24}} = \sqrt{\frac{25 \times 26 \times 51}{24}}$$

$$= 37.1651$$

Now,

$$z = \frac{w - \mu_w}{\sigma_w} = \frac{52 - 162.5}{37.1651} = -2.97$$

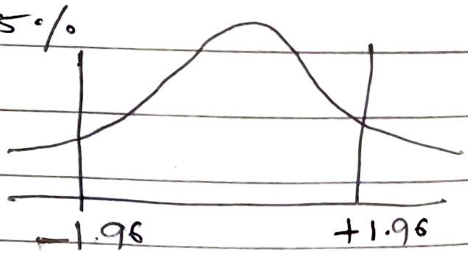
Tabulated z (Critical z)

The test is 2 sided and $\alpha = 5\%$

Critical $z = 1.96$

AR: $-1.96 < z < +1.96$

RR: $z \geq 1.96$ or $z \leq 1.96$




Statistical Decision

Since, cal z = −2.97 falls in the lower critical region ($z \geq -1.96$), we reject Ho in favour of $H_1$ at $\alpha = 5\%$ level of significance.

Conclusion

The output capacities of two machines are different.

# Wilcoxon Rank Sum Test

## Function of the test

Wilcoxon rank sum test is the non parametric alternative to the independent sample t test. This test determines whether two independent samples are from the population with some median or not.

The alternative non-parametric test of Wilcoxon rank sum test are:-

(1) Mann-Whitney U test

(2) Median test

## Test Assumptions:-

1. Two samples of sizes $n_1$ and $n_2$ respectively have been drawn randomly and independently from the respective pop$^{ns}$.

2. The measurement of scale is at least ordinal but the scale data is most preferred.

3. Two pop$^n$ have some basic shape.

## Hypothesis to Test

$H_0: M_1 = M_2$ (Two popns are identical wrt median)

against,

$H_1: M_1 \neq M_2$ ( " " are not " " " )

OR,

$H_1: M_1 < M_2$ (Pop$^n$ I has significantly lower median)

OR,

$H_1: M_1 > M_2$ (Pop$^n$ II has " " ")

## Test Statistic

Steps are:-

Step 1: Combine two samples into a large sample (Pooling of sample)

Pooled sample size $(n) = n_1 + n_2$

Step 2: Rank the n obs. of pooled sample from lowest to highest. In case of tied obs. give average rank

Step 3: Find the sum of ranks of sample 1 and sample 2 separately

Let $W_1$ = Sum of ranks of sample I / group 1

$W_2$ = Sum of ranks of sample 2 / group 2

Test Statistic $W_1 = \text{Min}\{W_1, W_2\}$

$$W_1 = \frac{n(n+1)}{2} - W_2$$

## Decision Rule:-

1 Two – tailed test

Reject $H_0$ if : $W_s \leq W_L$ or $W_s \geq W_U$

| Reject | Accept | Reject |
|---|---|---|

$W_L$ $\quad\quad$ $W_U$

2. Left Tailed test

If $W_s = W_1$ Reject $H_0$ if $W_s \leq W_L$

If $W_s = W_2$ Reject $H_0$ if $W_s \geq W_U$

3. Right tailed test

If $W_s = W_1$ Reject $H_0$ if $W_s \geq W_U$

If $W_s = W_2$ Reject $H_0$ if $W_s \leq W_L$

# Numerical

The following data represents the no. of hours the two different types scientific calculators operate before a recharge is required.

Calculator A

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 5.5 | 5.6 | 6.3 | 4.6 | 5.3 | 5.0 | 6.2 | 5.8 | 5.1 |

Calculator B

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 3.8 | 4.8 | 4.3 | 4.2 | 4.0 | 4.9 | 4.5 | 5.2 | 4.5 |

Use Wilcoxon Rank Sum test to determine if Calculator A operates more than calculator B on full recharge & use $\alpha = 5\%$.

Soln:-

Step 1: Null and Alternative Hypothesis

Let $X_1$ = Length of time a calculator of Brand A operates after full recharge

$X_2$ = Length " " " " Brand B
" " " "

Let $M_1$ and $M_2$ be the median length of time the calculation of brand A and brand B operates after full recharge.

$H_0 : M_1 = M_2$ (The brand of calculator operates for same duration)

Step 2: Level of significance

$\alpha = 5\%$

Step 3: Test Statistic

Steps are:

Step 1: Combine two samples into large sample (Pooling of sample)

Pooled sample size $(n) = n_1 + n_2$

Step 2: Rank the n obs. of pooled sample from lowest to highest. In case of tied obs. give average rank.

Step 3: Find the sum of ranks of sample 1 and sample 2 separately

Let $W_1$ = Sum of ranks of sample 1 / group 1

$W_2$ = Sum of ranks of sample 2 / group 2

Test Statistic $W_1 = \text{Min}\{W_1, W_2\}$

$$W_1 = \frac{n(n+1)}{2} - W_2$$

Step 4: Calculated W

Pooled sample

| A | A | A | A | A | A | A | A | A |
|---|---|---|---|---|---|---|---|---|
| 5.5 | 5.6 | 6.3 | 4.3 | 5.3 | 5.0 | 6.2 | 5.8 | 5.1 |

| B | B | B | B | B | B | B | B | B |
|---|---|---|---|---|---|---|---|---|
| 3.8 | 4.8 | 4.3 | 4.2 | 4.0 | 4.9 | 4.5 | 5.2 | 4.5 |

Sorted Data

3.8, 4.0, 4.2, 4.3, 4.5, 4.5, 4.6, 4.8, 4.9, 5.0, 5.1, 5.2, 5.3, 5.5, 5.6, 5.8, 6.2, 6.3

| Data | Group | Rank |
|---|---|---|
| 3.8 | B | 1 |
| 4.0 | B | 2 |
| 4.2 | B | 3 |
| 4.3 | B | 4 |
| 4.5 | B | 5.5 |
| 4.5 | B | 5.5 |
| 4.6 | B | 7 |
| 4.8 | B | 8 |
| 4.9 | B | 9 |
| 5.0 | A | 10 |
| 5.1 | ~~B~~ A | 11 |
| 5.2 | B | 12 |
| 5.3 | A | 13 |
| 5.5 | A | ~~13~~ 14 |
| 5.6 | A | ~~14~~ 15 |
| 5.8 | A | 16 |
| 6.2 | A | 17 |
| 6.3 | A | 18 |

$W_1$ = Sum of ranks of sample 1

= 121

$W_2$ = Sum of ranks of sample 2

= 50

$W_s = \text{Min}\{W_1, W_2\}$

$= \text{Min}\{50, 121\}$

$= 50$

## Critical $W$

Test is right sided,

$n_1 = 9$

$n_2 = 9$

$\alpha = 5\%$

Critical $W_L = 66$

Critical $W_U = 105$

## Statistical Decision

$W_S = 1 \quad = 50$

Since, $W_S = 50 < W_L = 66$, Reject $H_0$

## Conclusion

Calculator A operates longer in average than B on full recharge.

# Wilcoxon Rank Sum test
## (Large sample case)
## ($n_1 > 10$ , $n_2 > 10$)

If $n_1 > 10$, $n_2 > 10$, the distribution of $W$ can be approximated by $z$-distribution.

The $z$ transformation of $W$ is given by,

$$Z = \frac{W_s - \mu_W}{\sigma_W}$$

where,

$W_s$ = Min$\{W_1, W_2\}$

$\mu_W$ = Mean of the distribution of $W$

$$= \frac{n_s(n_s + n_l + 1)}{2}$$

$\sigma_W$ = Standard deviation of $W$

$$= \sqrt{\frac{n_s n_l (n_s + n_l + 1)}{12}}$$

$n_s$ = No. of observation in the sample that gives smaller $W$

$n_l$ = No. of obs. in the sample that gives larger $W$

$n_1 = 11$ $\quad$ $W_2 = 75$ $\quad$ $n_s = 15$

$n_2 = 15$ $\quad$ $W_1 = 50$ $\quad$ $n_l = 11$

Mean - Whitney U test

Function of the test

It is the non-parametric method which is used to determine whether two independent samples have been drawn from the popⁿ with same median or not.

This test has same function as the Wilcoxon rank sum test, but this test is modified test.

Test Statistics

(i) Small sample case ($n_1 \le 8$, $n_2 \le 8$)

(ii) Large sample case ($n_1 > 8$, $n_2 > 8$)

Small sample case:-

$n_1$ = No. of obs. in sample 1

$n_2$ = No. of obs. in sample 2

$W_1$ = Sum of ranks of sample 1

$W_2$ = " " " " 2

Now, U statistic is given by,

$$U_1 = W_1 - \frac{n_1(n_1+1)}{2}$$

$$U_2 = W_2 - \frac{n_2(n_2+1)}{2}$$

$$\boxed{U_1 = n_1 n_2 - U_2}$$

## Decision Rule:-

Case I : Two tailed test

Reject $H_0$ if $U = \text{Min}\{U_1, U_2\} \le U_{\alpha(2)}(n_1, n_2)$

Case II : Left tailed test

Reject $H_0$ if $U_1 \le U_{\alpha(1)}(n_1, n_2)$

Case III : Right tailed test

Reject $H_0$ if $U_2 \le U_{\alpha(1)}(n_1, n_2)$

Q# Previous Question

Soln:-

Test statistic

Calculated U

$n_1 = 9$ $\quad W_1 = 121$

$n_2 = 9$ $\quad W_2 = 50$

$$U_1 = W_1 - \frac{n_1(n_1+1)}{2}$$

$$= 121 - \frac{9 \times 10}{2}$$

$$= 76$$

Critical U

$U_{\alpha(1)}(9, 9) = 60$

$$U_2 = W_2 - \frac{n_2(n_2+1)}{2}$$

$$= 50 - \frac{9 \times 10}{2}$$

$$= 5$$

Statistical Decision

Since $U_2 = 5 < U_{\alpha(1)}(9,9) = 60$, reject $H_0$

Conclusion:-

Mann-Whitney U test
(Large sample case)

If $n_1 > 8$, $n_2 > 8$, the distribution of U is approximated by z distribution.

The z transformation of U is given by,

$$z = \frac{U - \mu_U}{\sigma_U}$$

$U = \text{Min}\{U_1, U_2\}$ for two tailed test

$= U_1$ for left tailed test

$= U_2$ for right tailed test

$\mu_U$ = Mean of U distribution

$$= \frac{n_1 n_2}{2}$$

$\sigma_U$ = S.D of U distribution

$$= \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

$U_2 = 5$

$\mu_U = \frac{9 \times 9}{2} = 40.5$

$\sigma_U = \sqrt{\frac{9 \times 9 \times 9}{12}} = 11.3247$

$Z = \frac{5 - 4.05}{11.3247}$

$= -3.1347$

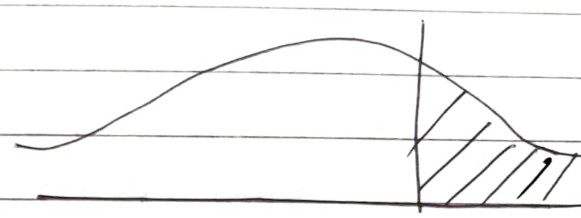
$\therefore$ Cal $Z = -3.1347$

Critical $Z$

$= 1.65$

Accept $H_0$ | Reject $H_0$

$+1.65$



$+1.65$

$Z = \frac{76 - 40.5}{11.3247} = +3.1347$

$|\text{cal } z| \geq$ Tab $z$, Reject $H_0$

Here, $|\text{cal } z| = 3.1347 >$ critical $z$

Reject $H_0$.

## Kruskal - Wallis H Test

### Function of the test

This test is developed by W.H Kruskal and W.A Wallis and it is also known as Kruskal - Wallis H test. It is a generalization of Wilcoxon rank som test or Mann– Whitney U test to the case of $K > 2$ samples, where $k$ = no. of independent samples. The test is non- parametric alternative of one-way analysis of variance.

## Test Assumptions

1. Samples are drawn randomly and independently from their respective populations.
2. Dependent variable should be measured at the ordinal or continuous level.
3. Independent variable should consist two or more categorical, independent groups.
4. Distribution in each group (i.e. the distribution of scores for each group of independent variable) have the same basic shape (which also means the same variability)

## Hypothesis

Null hypothesis assumes that the samples are from identical populations and alternative hypothesis assumes that the samples come from diffⁿ popⁿs.

$H_0 : M_1 = M_2 = \ldots = M_k$ (medians of all groups are equal or k distributions are identical)

$H_1$: At least one popⁿ median of one group is different from the popⁿ median of at least one other group (At least one distribution is different)

## Test Statistic

If one of the sample has size which is less than 5, we consider small. The procedure for finding test statistic is as follows:-

Let $n_i$ be the no. of observations in the $i^{th}$ sample, $i = 1, 2, 3, \ldots, k$

Step 1: Combine all data from k samples into a single series. The no. of observations in combined sample or pooled sample will be $n = n_1 + n_2 + \ldots + n_k$. Arrange all observations in ascending order (or descending order)

Step 2 :- Assign rank to sorted observations from 1 to n. The smallest observation getting a rank of 1 and highest observation getting a rank of n. In case of a repeated value, or a tie, assign ranks to them by averaging their rank position.

Step 3 :- Sum up the different ranks for each of the diff$^n$ groups.

$R_i$ = Sum of ranks for $i^{th}$ sample or group ($i = 1, 2, \ldots, n$)

Then, the test statistic is given by,

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

where, H = Kruskal - Wallis test statistic

N = Total no. of observations in combined samples

$R_i$ = Sum of ranks in ith sample

If each sample size is at least 5, the sampling distrib$^n$ of H statistic is a chi-square distribution with k-1 degrees of freedom where k stands for number of independent samples.

classmate
Date
Page

## Correction for ties

The corrected H for ties is given by,

$$H_c = \frac{H}{C}$$

where, $H_c$ = Corrected value of H

$H$ = Uncorrected value of H

$C$ = Correction for ties $= 1 - \frac{\sum_{i=1}^{k} (t_i^3 - t_j)}{n^3 - n}$

where, $t_j$ is the number of ties in each tie group

If there are few ties, then there will be small diffⁿ betⁿ $H_c$ and $H$.

## Decision Rule

For small sample case, the critical value of H is obtained from the exact distribⁿ table of H statistic. Let $H_\alpha(n_i)$ be the critical value of H from the table of H distribⁿ. We will use following decision rule:

Reject $H_0$ if cal $H \geq H_\alpha(n_i)$

~~For~~ For large sample case, we compute the critical value of Chi-square for k-1 degrees of freedom with upper tail probability of $\alpha$. Let $\chi^2_\alpha(k-1)$ be the critical value of Chi-square distribⁿ, then we will use following decision rule:-

Reject $H_0$ if cal $H \geq \chi^2_\alpha(k-1)$

## Numerical

The following are the final examination of marks of three groups of students who were taught C programming by three diffⁿ methods.

$n_i < 5$ H statistics

$n_i \geq 5$ $\chi^2$ statistics

classmate

Date

Page

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| First method | 94 | 88 | 91 | 74 | 87 | 97 | |
| Second method | 85 | 82 | 79 | 84 | 61 | 72 | 80 |
| Third method | 89 | 67 | 72 | 76 | 69 | | |

Use Kruskal Wallis H test at 5 % level of significance to test the hypothesis that the three methods are equally effective.

Soln:—

Step 1:- Null and Alternative Hypothesis

$X_1$ = Scores obtained by students who were taught (programming by method A)

$X_2$ = ----- B

$X_3$ = ----- C

Let $M_1$, $M_2$, $M_3$ be the median score obtained by students who were taught c program.

$H_0: M = M_1 = M_2$ ; (Three methods are equally efficient)

$H_1$ : At least one median is different from at least one other median

Step 2: Level of significance

$\alpha = 5\%$

Step 3: Test Statistics

Here, $n_1 = 6$, $n_2 = 7$ and $n_3 = 3$

We consider large sample,

Test statistics,

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(n+1)$$

$$= \frac{12}{n(n+1)} \left\{ \frac{R_1^2}{n_1} + \cdots + \frac{R_k^2}{n_k} \right\} - 3(n+1)$$

$n$ = Total sample size $= n_1 + n_2 + n_3$.

The test statistic follows $\chi^2$ distribution with $k-1$ d.f

Step 4: Calculated H

Combined sample

94, 88, 91, 74 87, 97 A

85, 82, 79 84, 61, 72, 80 B

89, 67, 72, 76, 69 C

| Data | Group | Rank |
|---|---|---|
| 61 | B | 1 |
| 67 | C | 2 |
| 69 | C | 3 |
| 72 | B | 4.5 |
| 72 | C | 4.5 |
| 74 | A | 6 |
| 76 | C | 7 |
| 79 | B | 8 |
| 80 | B | 9 |
| 82 | B | 10 |
| 84 | B | 11 |
| 85 | B | 12 |
| 87 | A | 13 |
| 88 | A | 14 |
| 89 | C | 15 |
| 91 | A | 16 |
| 94 | A | 17 |
| 97 | A | 18 |

$n = n_1 + n_2 + n_3$

$= 6 + 7 + 5$

$= 18$

$R_1$ = sum of sample of ranks of sample 1
= 84

$R_2$ = ...... sample 2
= 55.5

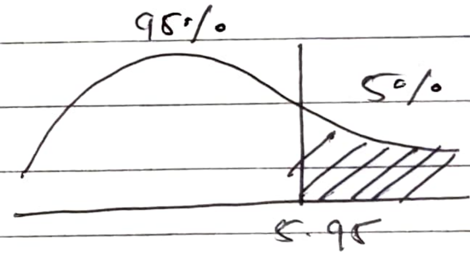
$R_3$ = ...... sample 3
= 31.5

Now, $H = \frac{12}{18(18+1)} \left\{ \frac{84^2}{6} + \frac{55.5^2}{7} + \frac{31.5^2}{5} \right\} - 3(18+1)$

= 6.67

Tabulated $\chi^2$

Test is right sided




D.f = k − 1
= 3 − 1 = 2

From the table of $\chi^2$ distribⁿ

$\chi^2_{0.05}(2) = 5.99$

| Accept | Reject |
| --- | --- |

5.99

Statistical Decision

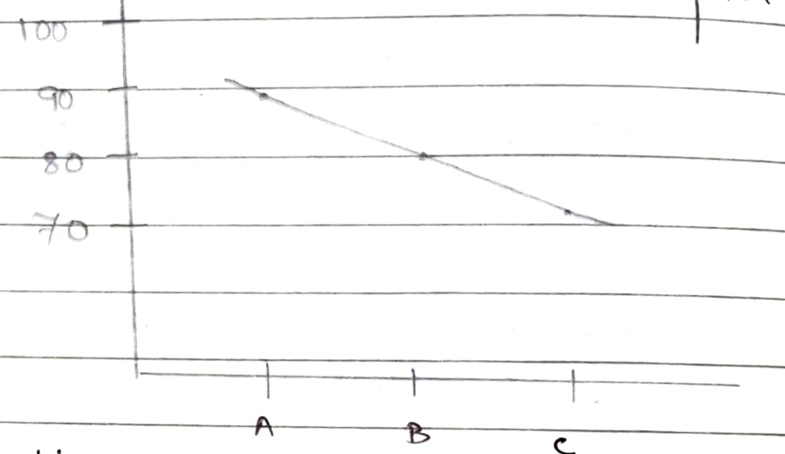
Since cal H = 6.67 > Tab $\chi^2$ = 5.99, we reject $H_0$.

## Conclusion

Three different method of teaching have diff$^n$ effectivenes

Median plot

| Method | Median |
|---|---|
| A | 89.5 |
| B | 80 |
| C | 72 |




Pairwise comparision

$$^nC_2 = {}^3C_2 = 3$$

## Friedman test

### Function of the test

~~Fun~~ Friedman test is used to determine whether several dependent or related samples have same median or not. The test requires that dependent variable is at least ordinal. It can also be used in the case of scale data that are not symmetric. Independent variable is a categorical having three or more groups. Friedman test is the non-parametric alternative to the one way analysis of variance with repeated measures.

Suppose a software company wants to compare the relative effectiveness of three different modes of contacting its registered customers: direct mail, telephone and postal mail. The company conducts a randomized block design experiment. For selected customers, the software company used all 3 modes during a 1-year period and recorded the percentage response to each type of mode. Company can test the hypothesis that percentage of response is same for all modes of contact.

### Data Structure for Friedman test

| Subject (Block) | Conditions / Treatments | | | |
|---|---|---|---|---|
| | 1 | 2 | .... | k |
| 1 | $x_{11}$ | $x_{12}$ | | $x_{1k}$ |
| 2 | $x_{21}$ | $x_{22}$ | | $x_{2k}$ |
| ... | ... | ... | | ... |
| n | $x_{n1}$ | $x_{n2}$ | ... | $x_{nk}$ |

Each row gives the score of one subject under each of the k conditions. So, we have k related samples. The score $x_{ij}$ represent the response for ith subject under jth condition ; $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$

## Test Assumptions

1. Each subject is measured on three or more occasion
2. Minimum allowable data within each block is at ordinal scale of measurement. But scale data is the preferred choice which are not normally distributed.
3. The n- blocks are independent so that the values in one block have no influence on the values in any other block. In other words, the scores obtained by one subject do not influence the scores obtained by other subjects.
4. There is no interaction between the n blocks and k treatment levels.
5. The k popns have the same variability.
6. The k popns have the same basic shape.

## Hypothesis

Ho : the treatments / conditions have identical effects (median effects are same for all treatment or condition) (Distribn of k treatments are same)

$H_1$: at least one treatment produces diff$^n$ effect than at least one of other treatments (Not all the medians are equal) (K distrib$^n$s. differs in location)

→ Test Statistic

Step 1: We first rank the scores in each row (block). We assign a rank of 1 to the lowest score, 2 to next lowest value and so on. In case of tied scores, we assign the avg. rank to each tied score. If the ordinal data are used for example use of internet could be low, intermediate or high, so we give rank of 1 to low, 2 to intermediate, and 3 to high.

Step 2: Find the sum of ranks for each column
Let $R_i$ = Sum of ranks for ith column ($i = 1, 2, \ldots, k$)

The appropriate test statistic is given by,

$$F_r = \frac{12}{n.k(k+1)} \sum_{i=1}^{k} R_i^2 - 3n(k+1)$$

, where,

$n$ = no. of rows
$k$ = no. of columns
$R_i$ = Sum of ranks for the ith treatment / condition

If either $k>6$ or $n>20$, we consider it as large sample, and the distrib$^n$ of $F^r$ is approximated by Chi-square distrib$^n$ with $k-1$ degrees of freedom.

## Decision Rule

If cal $F_r \geq$ tabulated $F_r(\alpha)(n,k)$ which is obtained from exact distribⁿ of $F_r$, we reject $H_0$ in favour of $H_1$.

For n or k is large, in particular, if $k > 6$ or $n > 20$, we reject null hypothesis $H_0$ if the computed value of $F_r$ is greater than or equals to the upper critical value of Chi-square distribⁿ having $k-1$ d.f, leaving $\alpha$ % area on the upper tail.

Reject $H_0$ if calculated $F_r \geq \chi^2_\alpha(k-1)$

Note :-

| No. of samples | Parametric Method | Non-parametric method |
|---|---|---|
| Single sample | t test for mean | - Sign test<br>- Wilcoxon signed rank test |
| Two independent samples | t test for diffⁿ of mean | - Wilcoxon Rank Sum test<br>- Mann Whitney U test<br>- Median test |
| Two dependent samples | Paired t test | - Sign test<br>- Wilcoxon signed rank test |
| More than 2 independent samples | One way ANOVA | - Kruskal Wallis H test<br>- Median test |
| More than 2 dependent samples | Repeated measure ANOVA<br>(Two way ANOVA) | - Friedman Test |

## Exercise

1. Nine experts rated four brands of coffee in a taste - testing experiments. A rating on a 7-point scale (1 = extremely unpleasing, 7 = extremely pleasing) is given for each of the four characterstics : taste, aroma, richness and acidity. The following table displays the summated ratings accumulated over all four characteristics.

| Expert | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| Brand A | 24 | 27 | 19 | 24 | 22 | 26 | 27 | 25 | 22 |
| Brand B | 26 | 27 | 22 | 27 | 25 | 27 | 26 | 27 | 23 |
| Brand C | 25 | 26 | 20 | 25 | 22 | 24 | 22 | 24 | 20 |
| Brand D | 22 | 24 | 16 | 23 | 21 | 24 | 23 | 21 | 19 |

At the 5% level of significance, is there evidence of a difference in the median summated ratings of the four brands of coffee?

Soln:-

Step 1: Null and Alternative Hypothesis (Summated rating)

Let $X_1$ = Cummulated score of brand A coffee (Rating)

$X_2$ = " " B "

$X_3$ = " " C " "

$X_4$ = " " D " "

Let $M_1$, $M_2$, $M_3$ and $M_4$ be the median summated rating of coffee of brand A, B, C and D respectively.

$H_0$ : $M_1 = M_2 = M_3 = M_4$ (Median summated rating of four brands of coffee are same)

$H_1$ : At least median summated rating of one brand of coffee is diffⁿ from at least one other brand of coffee.

Step 2: level of significance

$\alpha = 5\%$

Step 3: Test Statistic

$$F_r = \frac{12}{n \cdot k(k+1)} \sum_{i=1}^{k} R_i^2 - 3n(k+1)$$

$\chi^2_{(k-1)}$ $F_r(n,k)$

$k > 6$, $n > 20$ } We consider large sample

condition / treatment

k

Step 4: Calculated $F_r$

$k = 4$
$n = 9$

| Expert | Brand A | B | C | D |
|---|---|---|---|---|
| A | 24 (2) | 26 (4) | 25 (3) | 22 (1) |
| B | 27 (3.5) | 27 (3.5) | 26 (2) | 24 (1) |
| C | 19 (2) | 22 (4) | 20 (3) | 16 (1) |

| | | | | |
|---|---|---|---|---|
| D | 24 | 27 | 25 | 23 |
| | (2) | (4) | (3) | (1) |
| E | 22 | 25 | 22 | 21 |
| | (2.5) | (3) | (2.5) | (1) |
| F | 26 | 27 | 24 | 24 |
| | (2) | (3) | (1.5) | (1.5) |
| G | 27 | 26 | 22 | 23 |
| | (3) | (4) | (2) | (1) |
| H | 25 | 27 | 24 | 21 |
| | (3) | (4) | (2) | (1) |
| I | 22 | 25 | 20 | 19 |
| | (3) | (4) | (2) | (1) |

Rank sum $R_1 = 25$ $\quad R_2 = 34.5$ $\quad R_3 = 20$ $\quad R_4 = 10.5$

Avg. Rank $\overline{R_1} =$ $\quad \overline{R_2} =$ $\quad \overline{R_3} =$ $\quad \overline{R_4} =$

$n = 9, \ k = 4$

Now,

$$F_r = \frac{12}{n.k(k+1)} \sum_{i=1}^{k} R_i^2 - 3n(k+1)$$

$$= \frac{12}{9 \times 4 \times 5} \left\{ 25^2 + 34.5^2 + 20^2 + 10.5^2 \right\} - 3 \times 9 \times 5$$

$= 20.033$

Step 5: Tabulated $F_r$

Test is right sided & $\alpha = 5\%$.
From table of distribution of $F_r$,
$F_{0.05}(9, 4) = 7.667$

$$\chi^2_{0.05}(3) = 7.815$$

Step 6: Statistical Decision:-

Since, cal $F_r = 20.333 \geq$ Tab $F_r = 7.667$,
we strongly reject $H_0$ at 5% level of significance.

Step 7: Conclusion:-
Median summated rating of 4 brands of coffee are not same.

## Cochran's Q Test

Function of the test
Cochran's Q test is used for testing the significance of two or more matched set of frequencies, where a binary response (eg. 0 or 1) is recorded from each condition with each subject. Cochran's Q test is considered as a special case of Friedman test, which is used to detect differences in multiple matched sets with numeric response. When the responses are binary, the Friedman test becomes Cochran's Q test.

Examples of binary responses are:- True / Fail, Present / Absent, For / Against, Positive reaction / Negative reaction, Pass/Fail in exam etc. For such variable, we assign only two values 1 for success and 0 for failure.

Example:- A school teacher may examine whether pass rates increased as students had more time to study. To check this hypothesis, the teacher may plan a study, taking a random sample of ~~subjects~~ students. He may undertake three exams: first one is a 'surprise exam' to test their current knowledge. They are given a "mock exam" two weeks later before they took a "final exam" a further two weeks later. Students' performance in the exams are assessed in terms of "pass" or "fail".

Data are arranged in a two-way table consisting of n rows and k columns.

| Subject/Block | Conditions / Treatments (k) | | | |
|---|---|---|---|---|
| | 1 | 2 | .... | k |
| 1 | $x_{11}$ | $x_{12}$ | .... | $x_{1k}$ |
| 2 | $x_{21}$ | $x_{22}$ | .... | $x_{2k}$ |
| .... | ... | .... | .... | .... |
| n | $x_{n1}$ | $x_{n2}$ | .... | $x_{nk}$ |

Test Assumptions

1. Each k treatment/condition is independently applied to n subjects. Each subject responds to k different conditions so that there is k matched pairs.
2. Responses are binary. Each column $X_{ij} (i = 1, 2, \ldots n, j = 1, 2, \ldots, k)$ is scored as 0 (failure) or scored as 1 (success)

$X_{ij}$ = 1 if category of interest is present
0 if category of interest is not present

Note :- Remove those blocks having either all 0's or 1's.
This process doesnot affect the value of Q statistic.
The symbol 'n' denotes the no. of blocks after removing all those blocks containing either all 0's or all 1's.

3. The cases (participants) are selected randomly from the popⁿ of all possible cases. (For large sample approximation, cases must be large)

Hypothesis

Let the proportions $\pi_1, \pi_2, \dots \pi_k$ represent the proportion of successes in each of the k groups.

$H_0$ : Propⁿ of responses of a particular kind is the same in each column / condition

$(H_0 : \pi_1 = \pi_2 = \dots = \pi_k)$

$H_1$ : Propⁿ in at least one group is different from at least one another block

($H_1 : \pi_a \neq \pi_b$ for at least one pair $\pi_a, \pi_b$ with $a \neq b$ and $1 \leq a, b \leq k$)

Test Statistic

The test statistics is given by,

$$Q = \frac{(k-1)(kC - T^2)}{(kT - R)}$$

, where k = no. of groups
n = no. of subjects
$G_i$ = Total no. of successes in $i$th column
$B_j$ : Total no. of successes in $j$th column

$$C = \sum_{i=1}^{k} \left( \sum_{i=1}^{n} X_{ij} \right)^2 = G_1^2 + G_2^2 + \dots + G_k^2$$

$$R = \sum_{i=1}^{n} \left( \sum_{i=1}^{k} X_{ij} \right)^2 = B_1^2 + B_2^2 + \dots + B_n^2$$

$$T = \sum_{i=1}^{n} \sum_{i=1}^{k} X_{ij} = G_1 + G_2 + \dots + G_k = B_1 + B_2 + \dots + B_n$$

The distribution of statistic Q is chi-square with k-1 d.f. The condition required for the Chi-square approximation is that $k \geq 4$ and $nk \geq 24$

1 Decision Rule:-

Reject $H_0$ if calculated $Q$ is greater than or equal to critical value of Chi-square distribⁿ with $k-1$ d.f i.e $Q \geq \chi^2(k-1)$

Example:-

A text book distributer wishes to asses potential acceptance of four statistics text books. He asked 15 statistics professor to examine the books and to which ones they would seriously consider for their courses. Positive response (yes) is recorded as 1 and a negative response (no) as 0.

| Professor / text books | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| 2 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 1 |
| 3 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 1 | 0 |
| 4 | 0 | 1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |

Soln:-

Step 1: Null and Alternative Hypothesis

Let $\pi_1, \pi_2, \pi_3$ & $\pi_4$ be the population of positive responses (favourable reviews) to the text book of 1, 2, 3 and 4 respectively.

$H_0: \pi_1 = \pi_2 = \pi_3 = \pi_4$ (Four text books are equally popular among professors)

$H_1$: At least population of positive responses of one book is different from proportion of positive responses of at least one other book

Step 2: Level of significance

$\alpha = 5\%$

Step 3: Test Statistic

$$Q = \frac{(k-1)(kC - T^2)}{kT - R}$$

Q has chi-square dist$^n$ with $k-1$ df

$Q \sim \chi^2_\alpha (k-1)$

Step 4: Calculated Q

Book

| Professor | 1 | 2 | 3 | 4 | $B_i$ | $B_j^2$ |
|---|---|---|---|---|---|---|
| 1 | 1 | 1 | 0 | 0 | 2 | 4 |
| 2 | 1 | 1 | 0 | 1 | 3 | 9 |
| 3 | 1 | 0 | 0 | 0 | 1 | 1 |
| 4 | 1 | 1 | 1 | 1 | 4 | 16 |
| 5 | 1 | 1 | 0 | 1 | 3 | 9 |
| 6 | 0 | 1 | 0 | 0 | 1 | 1 |
| 7 | 0 | 1 | 1 | 0 | 2 | 4 |
| 8 | 1 | 1 | 1 | 0 | 3 | 9 |
| 9 | 0 | 0 | 1 | 0 | 1 | 1 |
| 10 | 1 | 0 | 1 | 0 | 2 | 4 |
| 11 | 0 | 0 | 0 | 0 | 0 | 0 |
| 12 | 1 | 1 | 0 | 0 | 2 | 4 |
| 13 | 1 | 0 | 0 | 1 | 2 | 4 |
| 14 | 0 | 1 | 1 | 0 | 2 | 4 |
| 15 | 1 | 1 | 0 | 0 | 2 | 4 |
| $G_i$ | 10 | 10 | 6 | 4 | | |
| $G_i^2$ | 100 | 100 | 36 | 16 | | |

Now,

$k = 4$

$C = G_1^2 + G_2^2 + G_3^2 + G_4^2 = 100 + 100 + 36 + 16 = 252$

$R = B_1^2 + B_2^2 + \dots + B_{15}^2$

$= 74$

$T = G_1 + G_2 + G_3 + G_4$

$= 10 + 10 + 6 + 4$

$= 30 \; (\Sigma B_j)$

$$Q = \frac{(k-1)(kC - T^2)}{kT - R}$$

$$= \frac{(4-1)(4 \times 252 - 30^2)}{4 \times 30 - 74}$$

$= 7.04347$

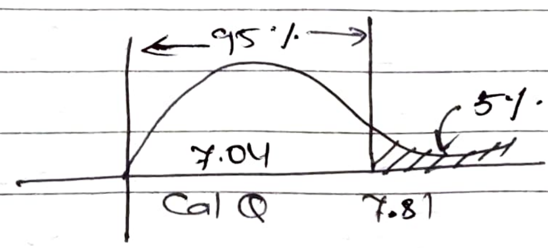
Critical $\chi^2$

The test is right sided and $\alpha = 5\%$

D.f $= k - 1 = 4 - 1 = 3$

From table of chi-square,

$\chi^2_{0.05}(3) = 7.82$




AR : $\chi^2 < 7.81$

RR : $\chi^2 \geq 7.81$

| Accept | Reject |
|---|---|

7.81

classmate
Date
Page

Statistical Decision

Since, cal Q = 7.0465 < critical $\chi^2$ = 7.81, we don't reject Ho at 5% level of significance.

Conclusion

Four text books are equally popular among professors. All books are equally likely for adoption.

## Chi- Square test of independence

Function of the test

→ This test is used to determine whether two nominal scale categorical variables are independent or not. For this test, the data must be presented in R×C contingency table, where R is the row and C is the column. The R×C contingency table is given by,

| Row variable | Column variable | | | | Total |
|---|---|---|---|---|---|
| | 1 | 2 | ... | C | |
| 1 | $O_{11}$ | $O_{12}$ | ... | $O_{1C}$ | $R_1$ |
| 2 | $O_{21}$ | $O_{22}$ | ... | $O_{2C}$ | $R_2$ |
| ... | ... | ... | ... | ... | ... |
| R | $O_{R1}$ | $O_{R2}$ | ... | $O_{RC}$ | $R_R$ |
| Total | $C_1$ | $C_2$ | ... | $C_C$ | n |

Here,

$O_{ij}$ = Observed cell frequency of ith row and jth column
($i = 1, 2, \ldots, R$ ; $j = 1, 2, \ldots, c$)

$n$ = The total size of the sample which is fixed in advance in independence test. The row and column totals are determined by chance.

Hypothesis to test

$H_0$: row categorical variable is independent of column categorical variable

$H_1$: row categorical variable is not independent of column categorical variable

If null hypothesis is true, then the relative frequencies are same for each row categories or relative frequencies are same for each column categories.

Test Statistics

The appropriate test statistic is,

$$\chi^2 = \sum_{i=1}^{R} \sum_{j=1}^{c} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

The test statistic has chi-square distribn with $(R-1)(C-1)$ degrees of freedom.

Here

$E_{ij} = \frac{R_i . C_j}{n}$ = Expected frequency for the cell $(i,j)$ i.e. expected frequency for the ith category of row variable and jth category of the column variable

Classmate

Date

Page

$R_i$ = $i^{th}$ row total

$C_j$ = $j^{th}$ column total &

n = sample size

Yates' continuity correction improves the approximation of the discrete sample chi-square statistic to a continuous chi-square distribⁿ.

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(|O_{ij} - E_{ij}| - 0.5)^2}{E_{ij}}$$

General rule for chi-square test of independence:

1. Expected frequencies should be relatively high.
2. When the expected frequencies are between 5 and 10, Yates' corrected formula should be used.
3. If the expected cell frequencies are high, the corrected and uncorrected formula gives the same value.
4. The value of the cell expected frequencies should be 5 or more in at least 80% of the cells. In other words, no more than 20% of the expected frequencies should be less than 5.
5. For contingency tables with more than 1 d.f, a minimum expectation of 1 is allowable if no more than 20% of the cells have expected frequencies of less than 5.

## 2×2 contingency table

The smallest possible contingency table is 2×2 table. It is most frequently encountered in biology.

| Variable B | Variable A | | Total |
|---|---|---|---|
| | Category 1 | Category 2 | |
| Category 1 | $O_{11}$ | $O_{12}$ | $R_1 = O_{11} + O_{12}$ |
| Category 2 | $O_{21}$ | $O_{22}$ | $R_2 = O_{21} + O_{22}$ |
| Total | $C_1 = O_{11} + O_{21}$ | $C_2 = O_{12} + O_{22}$ | $n$ |

### Test Statistic

Beside general formula to calculate the chi-square given for chi-square by following formula:-

$$\chi^2 = \frac{n\,(O_{11}\cdot O_{12} - O_{12}\cdot O_{21})^2}{R_1\cdot R_2\cdot C_1\cdot C_2}$$

In 2×2 contingency table, $\chi^2$ has 1 d.f. In this case, Yates' corrected formula is used which is given by,

$$\chi_c^2 = \frac{n\left\{\left|O_{11}\cdot O_{12} - O_{12}\cdot O_{21}\right| - \frac{n}{2}\right\}^2}{R_1\cdot R_2\cdot C_1\cdot C_2}$$

Expected frequencies are at least 5. If the expected frequencies are less than 5, the Fisher – Irving exact test should be used.

## Exercise

1. A book publisher wishes to determine whether there is difference in the type of book selected by males and females for recreational reading. A random sample provides the data given below:-

| Gender | Type of Book: Mystery | Romance | Self-help |
|---|---|---|---|
| Male | 243 | 201 | 191 |
| Female | 135 | 149 | 202 |

At $\alpha = 0.05$, test the claim that the type of book selected is independent of the gender of the individual.

2. Random samples of people of ages 15-24 and 25-34 were asked about their preferred method of communication with friends. The respondents were asked to select one of the methods from the following list: cell phone, instant message, email and other.

| Age | Preferred Communication Method: Cell phone | Instant message | Email | Other | Total |
|---|---|---|---|---|---|
| 15-24 | 48 | 40 | 5 | 7 | 100 |
| 25-34 | 41 | 30 | 15 | 14 | 100 |
| Total | 89 | 70 | 20 | 21 | 200 |

Test whether the two popⁿ share the same propⁿs of preferences for each type of communication method. Use 5% level of significance.

3. To determine whether types of professional jobs held in the computing industry independent of the number of years of person has worked in the industry, a random sample of 246 workers was interviewed. The following results were obtained.

| Years of experience | Professional Position: Manager | Programmer | Operator | System Analyst | Total |
|---|---|---|---|---|---|
| 0 – 3 | 6 | 37 | 11 | 13 | 67 |
| 4 – 8 | 28 | 16 | 23 | 24 | 91 |
| More than 8 | 47 | 10 | 12 | 19 | 88 |
| Total | 81 | 63 | 46 | 56 | n = 246 |

Use chi-square test of independence to determine whether type of professional job held in the computer industry is independent of years worked in the in the industry. Test at $\alpha = 0.05$

Chi-square test of homogeneity

This test is used to determine whether two or more pop$^{n}$s. are same wrt to some factor or criterion.

In this situation, samples are drawn from several diff$^{n}$ popns. independently and the researcher is interested in determining whether samples are drawn from pop$^{n}$s that are homogeneous wrt some criterion of classification i.e. prop$^{n}$ of elements is same over several categories for sampled pop$^{ns}$.

Hypothesis to test

$H_0$: Pop$^{ns}$ are homogeneous wrt some criterion variable

$H_1$: Pop$^{ns}$ are not homogeneous wrt some criterion variable

Data structure for this type of study is as follows:-

| Criterion Variable | Sample 1 | 2 | .... | k |
|---|---|---|---|---|
| 1 | $O_{11}$ | $O_{12}$ | .... | $O_{1k}$ |
| 2 | $O_{21}$ | $O_{22}$ | ... | $O_{2k}$ |
| ... | ... | ... | ... | ... |
| n | $O_{n1}$ | $O_{n2}$ | ... | $O_{nk}$ |
| Total | $n_1$ | $n_2$ | ... | $n_k$ |

$O_{ij}$ = observed frequency for ith category of criterion variable and for jth subpopulation, $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, k$

Test procedure is exactly same as chi-square test of independence.

## Exercise

1. A random sample 200 voters of party X, 150 voters of party Y and 150 voters of the party Z of the certain district were surveyed to know whether they are for a proposed abortion law, against it or undecided. The observed responses are given below:-

| Attitude towards abortion law | Political Affiliation X | Y | Z | Total |
|---|---|---|---|---|
| For | 82 | 70 | 62 | 214 |
| Against | 93 | 62 | 67 | 222 |
| Undecided | 25 | 18 | 21 | 64 |
| Total | $n_1 = 200$ | $n_2 = 150$ | $n_3 = 150$ | $n = 500$ |

Test the hypothesis that the popn propn within each row are same. Use $\alpha = 0.05$