Non-Parametric Tests and Multiple Regression

- Non-Parametric tests are called distribution free test and are used when population distributions are not known. This happens when samples are very small and distribution cannot be generated accurately.
- The non-parametric tests make less restrictive assumptions than parametric tests. They are easier to carry out and understand and are applicable in a wide range of situations. These tests are based on ranked data or nominal data (as with nominally scaled data) or on ranks (sometimes even ordering is not required).
- However, in these distribution-free tests there is a certain amount of ignoring of information (when ranks and not numerical data are used). Therefore, they are weak and less efficient than the relevant standard parametric tests
- Nonparametric does not meant that model lack parameters but that the number and nature of the parameters are flexible

Chi-Square test:

The Chi-square test is commonly used when the measurements pertain to assigning observations to distinct categories, as is the case with nominal scale data. This test is particularly useful when a researcher aims to compare different categories. As described by Williams (1968), the Chi-square test is best understood as a discrepancy statistic. It measures the discrepancies between the frequencies observed in the data and the frequencies expected based on a theoretical distribution.

Chi-square is an important non-parametric test and as such no rigid assumptions are necessary in respect of the type of population. We require only the degrees of freedom (implicitly of course the size of the sample) for using this test. As a non-parametric test, chi-square can be used (i) as a test of goodness of fit and (ii) as a test of independence

The following conditions should be satisfied before $\chi 2$ test can be applied:

(i) Observations recorded and used are collected on a random basis.

(ii) All the itmes in the sample must be independent.

(iii) No group should contain very few items, say less than 10. In case where the frequencies are less than 10, regrouping is done by combining the frequencies of adjoining groups so that the new frequencies become greater than 10. Some statisticians take this number as 5, but 10 is regarded as better by most of the statisticians.

(iv) The overall number of items must also be reasonably large. It should normally be at least 50, howsoever small the number of groups may be.

(v) The constraints must be linear. Constraints which involve linear equations in the cell frequencies of a contingency table (i.e., equations containing no squares or higher powers of the frequencies) are known are know as linear constraints

The various steps involved are as follows:

(i) First of all calculate the expected frequencies on the basis of given hypothesis or on the basis of null hypothesis. Usually in case of a 2 × 2 or any contingency table, the expected frequency for any given cell is worked out as under:

$$\text{Expected frequency of any cell} = \left[\frac{(\text{Row total for the row of that cell}) \times (\text{Column total for the column of that cell})}{(\text{Grand total})}\right]$$

(ii) Obtain the difference between observed and expected frequencies and find out the squares of such differences i.e., calculate $(O_{ij} - E_{ij})^2$

(iii) Divide the quantity $(O_{ij} - E_{ij})^2$ obtained as stated above by the corresponding expected frequency to get $(O_{ij} - E_{ij})^2/E_{ij}$ and this should be done for all the cell frequencies or the group frequencies.

(iv) Find the summation of $(O_{ij} - E_{ij})^2/E_{ij}$ values or what we call This is the required $\chi 2$ value.

$$\sum \frac{\left(O_{ij} - E_{ij}\right)^2}{E_{ij}}$$

The $\chi 2$ value obtained as such should be compared with relevant table value of $\chi 2$ and then inference be drawn as stated above.

We now give few examples to illustrate the use of $\chi 2$ test.

1. A die is thrown 132 times with following results:

Number turned up	1	2	3	4	5	6
Frequency	16	20	25	14	29	28

Is the die unbiased?

Solution: Let us take the hypothesis that the die is unbiased. If that is so, the probability of obtaining any one of the six numbers is 1/6 and as such the expected frequency of any one number coming upward is 132 ×1/6 = 22. Now we can write the observed frequencies along with expected frequencies and work out the value of χ2 as follows:

No. turned up	*Observed frequency O_i*	*Expected frequency E_i*	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
1	16	22	–6	36	36/22
2	20	22	–2	4	4/22
3	25	22	3	9	9/22
4	14	22	–8	64	64/22
5	29	22	7	49	49/22
6	28	22	6	36	36/22

∑ [(Oi – Ei) ^{2} /Ei] = 9.

Hence, the calculated value of χ2 = 9.

Degrees of freedom in the given problem is (n – 1) = (6 – 1) = 5.

The table value* of χ2 for 5 degrees of freedom at 5 per cent level of significance is 11.071. Comparing calculated and table values of χ2 , we find that calculated value is less than the table value and as such could have arisen due to fluctuations of sampling. The result, thus, supports the hypothesis and it can be concluded that the die is unbiased.

2.

Find the value of χ^2 for the following information:

Class	*A*	*B*	*C*	*D*	*E*
Observed frequency	8	29	44	15	4
Theoretical (or expected) frequency	7	24	38	24	7

Solution: Since some of the frequencies less than 10, we shall first re-group the given data as follows and then will work out the value of χ^2 :

Table 10.3

Class	*Observed frequency O_i*	*Expected frequency E_i*	$O_i - E_i$	$(O_i - E_i)^2/E_i$
A and *B*	(8+29)=37	(7+24)=31	6	36/31
C	44	38	6	36/38
D and *E*	(15+4)=19	(24+7)=31	–12	144/31

$$\therefore \quad \chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i} = 6.76 \text{ app.}$$

3.Genetic theory states that children having one parent of blood type A and the other of blood type B will always be of one of three types, A, AB, B and that the proportion of three types will on an average be as 1 : 2 : 1. A report states that out of 300 children having one A parent and B parent, 30 per cent were found to be types A, 45 per cent per cent type AB and remainder type B. Test the hypothesis by χ2 test.

Solution: The observed frequencies of type A, AB and B is given in the question are 90, 135 and 75 respectively.

The expected frequencies of type A, AB and B (as per the genetic theory) should have been 75, 150 and 75 respectively. We now calculate the value of χ2 as follows:

Type	Observed frequency O_i	Expected frequency E_i	$(O_i - E_i)$	$(O_i - E_i)^2$	$(O_i - E_i)^2/E_i$
A	90	75	15	225	225/75 = 3
AB	135	150	–15	225	225/150 = 1.5
B	75	75	0	0	0/75 = 0

$$\therefore \quad \chi^2 = \Sigma\frac{(O_i - E_i)^2}{E_i} = 3 + 1.5 + 0 = 4.5$$

$$\because \quad \text{d.f.} = (n - 1) = (3 - 1) = 2.$$

Table value of χ2 for 2 d.f. at 5 per cent level of significance is 5.991.

The calculated value of χ2 is 4.5 which is less than the table value and hence can be ascribed to have taken place because of chance. This supports the theoretical hypothesis of the genetic theory that on an average type A, AB and B stand in the proportion of 1 : 2 : 1.

4. The table given below shows the data obtained during outbreak of smallpox

	Attacked	Not attacked	Total
Vaccinated	31	469	500
Not vaccinated	185	1315	1500
Total	216	1784	2000

Test the effectiveness of vaccination in preventing the attack from smallpox. Test your result with the help of χ2 at 5 per cent level of significance.

Solution: Let us take the hypothesis that vaccination is not effective in preventing the attack from smallpox i.e., vaccination and attack are independent. On the basis of this hypothesis, the expected frequency corresponding to the number of persons vaccinated and attacked would be: Expectation of (AB) = ((A)x (B))/N

where A represents vaccination and B represents attack.

(A) = 500, (B) = 216, N = 2000

Expectation of (AB)=(500X216)/2000

=54

Now using the expectation of (AB), we can write the table of expected values as follows:

	Attacked: B	*Not attacked: b*	*Total*
Vaccinated: *A*	$(AB) = 54$	$(Ab) = 446$	500
Not vaccinated: *a*	$(aB) = 162$	$(ab) = 1338$	1500
Total	216	1784	2000

Table 10.5: Calculation of Chi Square

Group	*Observed frequency* O_{ij}	*Expected frequency* E_{ij}	$(O_{ij} - E_{ij})$	$(O_{ij} - E_{ij})^2$	$(O_{ij} - E_{ij})^2/E_{ij}$
AB	31	54	–23	529	529/54 = 9.796
Ab	469	446	+23	529	529/44 = 1.186
aB	158	162	+23	529	529/162 = 3.265
ab	1315	1338	–23	529	529/1338 = 0.395

$$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 14.642$$

Degrees of freedom in this case = (r – 1) (c – 1) = (2 – 1) (2 – 1) = 1. The table value of $\chi 2$ for 1 degree of freedom at 5 per cent level of significance is 3.841. The calculated value of $\chi 2$ is much higher than this table value and hence the result of the experiment does not support the hypothesis. We can, thus, conclude that vaccination is effective in preventing the attack from smallpox.

5. Two research workers classified some people in income groups on the basis of sampling studies. Their results are as follows:

Investigators	*Income groups*			*Total*
	Poor	*Middle*	*Rich*	
A	160	30	10	200
B	140	120	40	300
Total	300	150	50	500

Show that the sampling technique of at least one research worker is defective.

Solution: Let us take the hypothesis that the sampling techniques adopted by research workers are similar (i.e., there is no difference between the techniques adopted by research workers). This being so, the expectation of A investigator classifying the people in

(i) Poor income group $= \frac{200 \times 300}{500} = 120$

(ii) Middle income group $= \frac{200 \times 150}{500} = 60$

(iii) Rich income group $= \frac{200 \times 50}{500} = 20$

Similarly the expectation of *B* investigator classifying the people in

(i) Poor income group $= \frac{300 \times 300}{500} = 180$

(ii) Middle income group $= \frac{300 \times 150}{500} = 90$

(iii) Rich income group $= \frac{300 \times 50}{500} = 30$

Groups	*Observed frequency* O_{ij}	*Expected frequency* E_{ij}	$O_{ij} - E_{ij}$	$(O_{ij} - E_{ij})^2 E_{ij}$
Investigator A				
classifies people as poor	160	120	40	1600/120 = 13.33
classifies people as middle class people	30	60	–30	900/60 = 15.00
classifies people as rich	10	20	–10	100/20 = 5.00
Investigator B				
classifies people as poor	140	180	–40	1600/180 = 8.88
classifies people as middle class people	120	90	30	900/90 = 10.00
classifies people as rich	40	30	10	100/30 = 3.33

Hence, $$\chi^2 = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = 55.54$$

Degrees of freedom = (c – 1) (r – 1)

= (3 – 1) (2 – 1) = 2.

The table value of χ2 for two degrees of freedom at 5 per cent level of significance is 5.991. The calculated value of χ2 is much higher than this table value which means that the calculated value cannot be said to have arisen just because of chance. It is significant. Hence, the hypothesis does not hold good. This means that the sampling techniques adopted by two investigators differ and are not similar. Naturally, then the technique of one must be superior than that of the other

IMPORTANT CHARACTERISTICS OF χ2 TEST

(i) This test (as a non-parametric test) is based on frequencies and not on the parameters like mean and standard deviation.

(ii) The test is used for testing the hypothesis and is not useful for estimation.

(iii) This test possesses the additive property as has already been explained.

(iv) This test can also be applied to a complex contingency table with several classes and as such is a very useful test in research work.

(v) This test is an important non-parametric test as no rigid assumptions are necessary in regard to the type of population, no need of parameter values and relatively less mathematical details are involved.

One sample sign test: It is based on the direction or the plus or minus signs of observation in a sample and not on their numerical magnitudes.

Problem: 1. The compressive strength of insulating blocks used in the construction of new houses is tested by a civil engineer. The engineer needs to be certain at the 5% level of significance that the median compressive strength is at least 1000 psi. Twenty randomly selected blocks give the following results:

Observation	Compressive Strength	Observation	Compressive Strength	Observation	Compressive Strength	Observation	Compressive Strength
1	1128.7	6	718.4	11	1167.1	16	1153.6
2	679.1	7	787.4	12	1387.5	17	1423.3
3	1317.2	8	1562.3	13	679.9	18	1122.6
4	1001.3	9	1356.9	14	1323.2	19	1644.3
5	1107.6	10	1153.2	15	788.4	20	737.4

Test (at the 5% level of significance) the null hypothesis that the median compressive strength of the insulting blocks is 1000 psi against the alternative that it is greater

Solution: The hypotheses are $H0 : \theta = 1000$ and $H1 : \theta > 1000$

Comp. Strength	Sign	Comp. Strength	Sign	Comp. Strength	Sign	Comp. Strength	Sign
1128.7	+	718.4	−	1167.1	+	1153.6	+
679.1	−	787.4	−	1387.5	+	1423.3	+
1317.2	+	1562.3	+	679.9	−	1122.6	+
1001.3	+	1356.9	+	1323.2	+	1644.3	+
1107.6	+	1153.2	+	788.4	−	737.4	−

We have 14 plus signs and the required probability value is calculated directly from the binomial formula as (for binomial distribution p =1/2

$$P(X=r)=\binom{n}{r}q^{n-r}p^{r}=\binom{n}{r}(1-p)^{n-r}p^{r}$$

$$P(X\geq 14) = \sum_{r=14}^{20}\binom{20}{r}\left(\frac{1}{2}\right)^{20-r}\left(\frac{1}{2}\right)^{r}$$

$$=\frac{20.19.18.17.16.15}{1.2.3.4.5.6}\left(\frac{1}{2}\right)^{20}+\frac{20.19.18.17.16}{1.2.3.4.5}\left(\frac{1}{2}\right)^{20}+\frac{20.19.18.17}{1.2.3.4}\left(\frac{1}{2}\right)^{20}+\frac{20.19.18}{1.2.3}\left(\frac{1}{2}\right)^{20}+\frac{20.19}{1.2}\left(\frac{1}{2}\right)^{20}+\frac{20}{1}\left(\frac{1}{2}\right)^{20}+\left(\frac{1}{20}\right)^{20}$$

$$=\left(\frac{1}{2}\right)^{20}(38760+15504+4845+1140+190+20+1)$$

$$=0.05766$$

Since we are performing a one-tailed test, we must compare the calculated value with the value 0.05. Since 0.05 < 0.05766 we conclude that we cannot reject the null hypothesis and that on the basis of the available evidence, we cannot conclude that the median compressive strength of the insulating blocks is greater than 1000 psi.

2.A certain type of solid rocket fuel is manufactured by binding an igniter with a propellant. In order that the fuel burns smoothly and does not suffer either “flame-out” or become unstable it is essential that the material bonding the two components of the fuel has a shear strength of 2000 psi. The results arising from tests performed on 10 randomly selected samples of fuel are as follows

Using the 5% level of significance, test the null hypothesis that the median shear strength is 2000 psi.

Observation	Shear Strength	Observation	Shear Strength
1	2128.7	6	1718.4
2	1679.1	7	1787.4
3	2317.2	8	2562.3
4	2001.3	9	2356.9
5	2107.6	10	2153.2

Answer The hypotheses are H0 : θ = 2000 H1 : θ ≠ 2000

Shear Strength	Sign	Shear Strength	Sign
2128.7	+	1718.4	−
1679.1	−	1787.4	−
2317.2	+	2562.3	+
2001.3	+	2356.9	+
2107.6	+	2153.2	+

We have 7 plus signs and the required probability value is calculated directly from the binomial formula as

$$P(X \geq 7) = \sum_{r=7}^{10} \binom{10}{r} \left(\frac{1}{2}\right)^{10-r} \left(\frac{1}{2}\right)^{r} = \frac{10.9.8}{1.2.3}\left(\frac{1}{2}\right)^{10} + \frac{10.9}{1.2}\left(\frac{1}{2}\right)^{10} + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10}$$

$$= \frac{10.9.8}{1.2.3}\left(\frac{1}{2}\right)^{10} + \frac{10.9}{1.2}\left(\frac{1}{2}\right)^{10} + \frac{10}{1}\left(\frac{1}{2}\right)^{10} + \left(\frac{1}{2}\right)^{10} = \left(\frac{1}{2}\right)^{10}(120 + 45 + 10 + 1) \simeq 0.172$$

Since we are performing a two-tailed test, we must compare the calculate value with the value 0.025. Since 0.025 < 0.172 we cannot reject the null hypothesis on the basis of the available evidence and we cannot conclude that the median shear strength is different to 2000 psi.

-> Wilcoxon-Mann-Whitney test (or U-test):

This is a very popular test amongst the rank sum tests. This test is used to determine whether two independent samples have been drawn from the same population. This test applies under very general conditions and requires only that the populations sampled are continuous.

To perform this test, we first of all rank the data jointly, taking them as belonging to a single sample in either an increasing or decreasing order of magnitude. We usually adopt low to high ranking process which means we assign rank 1 to an item with lowest value, rank 2 to the next higher item and so on.

In case there are ties, then we would assign each of the tied observation the mean of the ranks which they jointly occupy. For example, if sixth, seventh and eighth values are identical, we would assign each the rank (6 + 7 + 8)/3 = 7. After this we find the sum of the ranks assigned to the values of the first sample (and call it R1) and also the sum of the ranks assigned to the values of the second sample (and call it R2). Then we work out the test statistic i.e., U, which is a measurement of the difference between the ranked observations of the two samples as under:

$$U = n_1 \cdot n_2 + \frac{n_1(n_1 + 1)}{2} - R_1$$

where n1 , and n2 are the sample sizes and R1 is the sum of ranks assigned to the values of the first sample.

If the null hypothesis that the n1 + n2 observations came from identical populations is true, the said 'U' statistic has a sampling distribution with

$$\text{Mean} = \mu_U = \frac{n_1 \cdot n_2}{2}$$

and Standard deviation (or the standard error)

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

1. The values in one sample are 53, 38, 69, 57, 46, 39, 73, 48, 73, 74, 60 and 78. In another sample they are 44, 40, 61, 52, 32, 44, 70, 41, 67, 72, 53 and 72. Test at the 10% level the hypothesis that they come from populations with the same mean. Apply U-test.

Solution: First of all we assign ranks to all observations, adopting low to high ranking process on the presumption that all given items belong to a single sample. By doing so we get the following:

Size of sample item in ascending order	Rank	Name of related sample: [A for sample one and B for sample two]
32	1	B
38	2	A
39	3	A
40	4	B
41	5	B
44	6.5	B
44	6.5	B
46	8	A
48	9	A
52	10	B
53	11.5	B
53	11.5	A
57	13	A
60	14	A
61	15	B
67	16	B
69	17	A
70	18	B
72	19.5	B
72	19.5	B
73	21.5	A
73	21.5	A
74	23	A
78	24	A

From the above we find that the sum of the ranks assigned to sample one items or R1 = 2 + 3 + 8 + 9 + 11.5 + 13 + 14 + 17 + 21.5 + 21.5 + 23 + 24 = 167.5

and similarly we find that the sum of ranks assigned to sample two items or R2 = 1 + 4 + 5 + 6.5 + 6.5 + 10 + 11.5 + 15 + 16 + 18 + 19.5 + 19.5 = 132.5 and we have n1 = 12 and n2 = 12

$$\text{Hence, test statistic } U = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$= (12)(12) + \frac{12(12+1)}{2} - 167.5$$

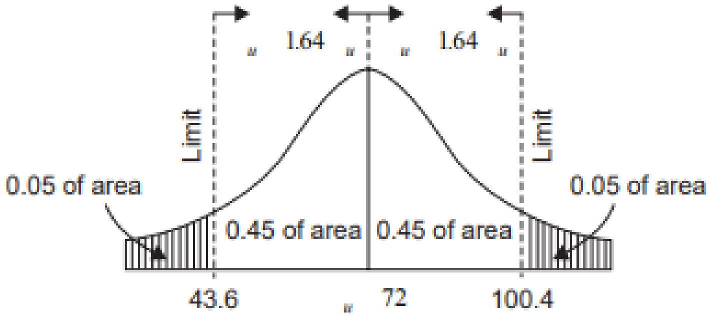
$$= 144 + 78 - 167.5 = 54.5$$

Since in the given problem n1 and n2 both are greater than 8, so the sampling distribution of U approximates closely with normal curve. Keeping this in view, we work out the mean and standard deviation taking the null hypothesis that the two samples come from identical populations as under:

$$\mu_U = \frac{n_1 \times n_2}{2} = \frac{(12)(12)}{2} = 72$$

$$\sigma_U = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} = \sqrt{\frac{(12)(12)(12+12+1)}{12}}$$
$$= 17.32$$

As the alternative hypothesis is that the means of the two populations are not equal, a two-tailed test is appropriate. Accordingly the limits of acceptance region, keeping in view 10% level of significance as given, can be worked out as under:



(Shaded portion indicates rejection regions)

As the z value for 0.45 of the area under the normal curve is 1.64, we have the following limits of acceptance region:

Upper limit = μ_U +1.64 σ_U =72 +1.64 (17.32)= 100.40

Lower limit = μ_U – 1.64σ_U = 72 -1.64(17.32) =43.60

As the observed value of U is 54.5 which is in the acceptance region, we accept the null hypothesis and conclude that the two samples come from identical populations (or that the two populations have the same mean) at 10% level. We can as well calculate the U statistic as under using R2 value:

$$U = n_1 \cdot n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

$$= (12)(12) + \frac{12(12+1)}{2} - 132.5$$

$$= 144 + 78 - 132.5 = 89.5$$

The value of U also lies in the acceptance region and as such our conclusion remains the same, even if we adopt this alternative way of finding U.

2. Two samples with values 90, 94, 36 and 44 in one case and the other with values 53, 39, 6, 24, and 33 are given. Test applying Wilcoxon test whether the two samples come from populations with the same mean at 10% level against the alternative hypothesis that these samples come from populations with different means.

Solution: Let us first assign ranks as stated earlier and we get

Size of sample item in ascending order	*Rank*	*Name of related sample (Sample one as A Sample two as B)*
6	1	*B*
24	2	*B*
33	3	*B*
36	4	*A*
39	5	*B*
44	6	*A*
53	7	*B*
90	8	*A*
94	9	*A*

Sum of ranks assigned to items of sample one = 4 + 6 + 8 + 9 = 27

No. of items in this sample = 4

Sum of ranks assigned to items of sample two = 1 + 2 + 3 + 5 + 7 = 18

No. of items in this sample = 5

As the number of items in the two samples is less than 8, we cannot use the normal curve approximation technique as stated above and shall use the table giving values of Wilcoxon's distribution.

To use this table, we denote 'Ws ' as the smaller of the two sums and 'Wl ' the larger. Also, let 's' be the number of items in the sample with smaller sum and let 'l' be the number of items in the sample with the larger sum. Taking these notations we have for our question the following values:

Ws = 18; s = 5; Wl = 27; l = 4

The value of Ws is 18 for sample two which has five items and as such s = 5.

We now find the difference between Ws and the minimum value it might have taken, given the value of s. The minimum value that Ws could have taken, given that s = 5, is the sum of ranks 1 through 5 and this comes as equal to 1 + 2 + 3 + 4 + 5 = 15.

Thus, (Ws – Minimum Ws) = 18 – 15 = 3.

To determine the probability that a result as extreme as this or more so would occur, we find the cell of the table which is in the column headed by the number 3 and in the row for s = 5 and l = 4.

The entry in this cell is 0.056 which is the required probability of getting a value as small as or smaller than 3 and now we should compare it with the significance level of 10%.

Since the alternative hypothesis is that the two samples come from populations with different means, a two-tailed test is appropriate and accordingly 10% significance level will mean 5% in the left tail and 5% in the right tail. In other words, we should compare the calculated probability with the probability of 0.05, given the null hypothesis and the significance level. If the calculated probability happens to be greater than 0.05 (which actually is so in the given case as 0.056 > 0.05), then we should accept the null hypothesis. Hence, in the given problem, we must conclude that the two samples come from populations with the same mean.

- The Kruskal-Wallis test (or H test): This test is conducted in a way similar to the U test described above. This test is used to test the null hypothesis that 'k' independent random samples come from identical universes against the alternative hypothesis that the means of these universes are not equal.

  This test is analogous to the one-way analysis of variance, but unlike the latter it does not require the assumption that the samples come from approximately normal populations or the universes having the same standard deviation. In this test, like the U test, the data are ranked jointly from low to high or high to low as if they constituted a single sample. The test statistic is H for this test which is worked out as under:

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

where $n = n_1 + n_2 + \ldots + n_k$ and R_i being the sum of the ranks assigned to n_i observations in the ith sample.

If the null hypothesis is true that there is no difference between the sample means and each sample has at least five items* , then the sampling distribution of H can be approximated with a chisquare distribution with (k – 1) degrees of freedom. As such we can reject the null hypothesis at a given level of significance if H value calculated, as stated above, exceeds the concerned table value of chi-square. Let us take an example to explain the operation of this test:

- Use the Kruskal-Wallis test at 5% level of significance to test the null hypothesis that a professional bowler performs equally well with the four bowling balls, given the following results

With Ball No. *A*	271	282	257	248	262
With Ball No. *B*	252	275	302	268	276
With Ball No. *C*	260	255	239	246	266
With Ball No. *D*	279	242	297	270	258

Solution: To apply the H test or the Kruskal-Wallis test to this problem, we begin by ranking all the given figures from the highest to the lowest, indicating besides each the name of the ball as under

Bowling results	*Rank*	*Name of the ball associated*
302	1	*B*
297	2	*D*
282	3	*A*
279	4	*D*
276	5	*B*
275	6	*B*
271	7	*A*
270	8	*D*
268	9	*B*
266	10	*C*
262	11	*A*
260	12	*C*
258	13	*D*
257	14	*A*
255	15	*C*
252	16	*B*
248	17	*A*
246	18	*C*
242	19	*D*
239	20	*C*

For finding the values of R_i, we arrange the above table as under:

Ball A	Rank	Ball B	Rank	Ball C	Rank	Ball D	Rank
271	7	252	16	260	12	279	4
282	3	275	6	255	15	242	19
257	14	302	1	239	20	297	2
248	17	268	9	246	18	270	8
262	11	276	5	266	10	158	13
$n_1 = 5$	$R_1 = 52$	$n_2 = 5$	$R_2 = 37$	$n_3 = 5$	$R_3 = 75$	$n_4 = 5$	$R_4 = 46$

Now we calculate *H* statistic as under:

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

$$= \frac{12}{20(20+1)} \left\{ \frac{52^2}{5} + \frac{37^2}{5} + \frac{75^2}{5} + \frac{46^2}{5} \right\} - 3(20+1)$$

$$= (0.02857)\,(2362.8) - 63 = 67.51 - 63 = 4.51$$

As the four samples have five items each, the sampling distribution of H approximates closely with χ2 distribution. Now taking the null hypothesis that the bowler performs equally well with the four balls, we have the value of

χ2 = 7.815 for (k – 1) or 4 – 1 = 3 degrees of freedom at 5% level of significance. Since the calculated value of H is only 4.51 and does not exceed the χ2 value of 7.815, so we accept the null hypothesis and conclude that bowler performs equally well with the four bowling balls.

Multiple Regression:

Multiple regression is a logical and mathematical extension of simple linear bivariate regression. It examines the relationship between two or more intervally scaled predictor variables and one intervally scaled criterion variable. Ordinal data that are near 'interval', such as semantic differential scale data can also generally be used (Darlington, 1968). Assume that Y is a dependent variable that depends on X1, X2,..., Xn, and the relationship between Y and Xi 's is linear, then a mathematical relationship between Y and Xi 's can represented by the following equation,

$$Y = b_0 + b_1 X_1 + b_2 X_2 + . \ldots + b_n X_n$$

Where

B_0 = constant derived from the analysis (parameter)

B_i = coefficients of association with the predictor variable (parameters)

X_i = predictor or explanatory variables that influence the criterion variable Y.

Often, Y is referred to as the 'regressand' and X1, X2,..., Xn as the 'regressors'. The above equation is an exact relationship explaining the variations in Y as fully attributed to the changes in Xi 's. However, in reality, the dependent variable is influenced by a large number of other factors. Hence, errors are likely to occur while estimating the parameters.

The sources of errors are,

(i) omission of certain variables from the function,

(ii) errors of aggregation, and

(iii) errors of measurement (refer Koutsoyiannis, 1977).

In order to account for all these sources of errors, one more variable u (error term) has to be introduced into the above equation. This is termed as the error component or the random component. Then the above equation becomes,

$$Y = (b_0 + b_1 X_1 + b_2 X_2 + ... + b_n X_n) + u$$

Problem: A Distributor of frozen desert pies want to evaluate factors thought to influence demand.

Week	Pie Sales	Price ($)	Advertising ($100s)
1	350	5.50	3.3
2	460	7.50	3.3
3	350	8.00	3.0
4	430	8.00	4.5
5	350	6.80	3.0
6	380	7.50	4.0
7	430	4.50	3.0
8	470	6.40	3.7
9	450	7.00	3.5
10	490	5.00	4.0
11	340	7.20	3.5
12	300	7.90	3.2
13	440	5.90	4.0
14	450	5.00	3.5
15	300	7.00	2.7

Depend Variable: Pie Sales (units per week)

Independent Variables: Price Advertising

Data are collected for 15 weeks

Solution

Sales = $b_o + b_1$ (price) + b_2 (Advertising)

	y	x_1	x_2	x_1y	x_2y	x_1x_2	x_1^2	x_2^2
1	350	5.5	3.3	1925	1155	18.15	30.25	10.89
2	460	7.5	3.3	3450	1518	24.75	56.25	10.89
3	350	8	3	2800	1050	24	64	9
4	430	8	4.5	3440	1935	36	64	20.25
5	350	6.8	3	2380	1050	20.4	46.24	9
6	380	7.5	4	2850	1520	30	56.25	16
7	430	4.5	3	1935	1290	13.5	20.25	9
8	470	6.4	3.7	3008	1739	23.68	40.96	13.69
9	450	7	3.5	3150	1575	24.5	49	12.25
10	490	5	4	2450	1960	20	25	16
11	340	7.2	3.5	2448	1190	25.2	51.84	12.25
12	300	7.9	3.2	2370	960	25.28	62.41	10.24
13	440	5.9	4	2596	1760	23.6	34.81	16
14	450	5	3.5	2250	1575	17.5	25	12.25
15	300	7	2.7	2100	810	18.9	49	7.29
Σ		99.2	52.2	39152	21087	345.46	675.26	185

$$b_1 = \frac{\sum x_1 y \sum x_2^2 - \sum x_1 x_2 \sum x_2 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2},$$

$$b_2 = \frac{\sum x_2 y \sum x_1^2 - \sum x_1 x_2 \sum x_1 y}{\sum x_1^2 \sum x_2^2 - (\sum x_1 x_2)^2},$$

$(\Sigma x_1 x_2)^2 = 26814169$

$b_1 =- 24.97509$

$b_2 = 74.131$

$b_0 = \frac{\sum y}{n} - \frac{\sum x_1}{n} - \frac{\sum x_2}{n}$

$B_0 = 306.526$

$$\text{Sales} = 306.526 - 24.975(\text{Price}) + 74.131(\text{Advertising})$$

An Overview of Multiple Regression Analysis Procedure

The measures generally derived in multiple regression analysis can be summarised as follows:

1. A least squares function fitting the data, the regressor equation is obtained.

2. An F test is made for checking the significance of the overall regression model.

3. R2 , the coefficient of multiple determination (both sample based and population ad�justed) is obtained.

4. Standard errors of regression coefficients are computed.

5. Partial hypotheses regarding the significance of the regression coefficients bj (param�eters) are tested using the t-test.

6. Selection of a subset of regressors is made if required (using the backward elimination procedure).

Variants of Regression Analysis

A few of the variants of the basic multiple regression analysis are now discussed.

- Multiple Regression with nominal variables:

In some situations researchers may have to employ nominally scaled predictor variables, such as marital status, gender, or occupational category, in a multiple regression analysis.

Maddala (1972) suggests the use of dummy or indicator variables as long as there are relatively few such variables. While natural dichotomies can be coded as 0 or 1, polynomous data (multiple categories), such as occupation and caste, in each category serve as variables. Dummy variables can also be used,

(i) to measure the shift of a regression function over time,

(ii) to measure the change of parameters over time,

(iii) as proxies for the dependent variable, and

(iv) for seasonal adjustment of time series. If, however, the dependent variable is an indicator variable, the shape of the response function will frequently be curvilinear. In these cases, transformations of non-linear function into log-linear (logit) or probit analysis is recommended.

- Multiple Regression with rank order or rank transformed variables:

There are situations when the dependent and independent variables are rank orders and are best transformed into such ranking. To estimate the association between such sets of data, rank correlation techniques can be used.

- Multiple Regression with lagged variables:

Often, instead of selecting different variables as the explanatory variables, it may so happen that the dependent variable is expected to be related to the lagged values of the dependent variables among the set of explanatory variables.

- Non-linear Multiple Regression:

Often Y may be related to a function f (x), which is not necessarily linear. Quite a few procedures are available to obtain the estimates of the parameters.

Applications

The multiple regression analysis has various applications such as:

(i) Measuring the determinants of say, market structure, and demand;

(ii) Forecasting sales;

(iii) Determining the effect of a particular independent variable on the dependent variable, while the other predictor vari�ables are held constant;

(iv) Determining whether predictor variables, other than those being considered, are related to the criterion variable, adjusting data obtained from experiments for factors not controlled and believed not to be randomly distributed;

(v) Estimating values for the missing data (item non-response) in surveys; and

(vi) Comparing data obtained on the same data items (but from different respondents) by different interviewers, to determine if any bias exists.

Interpretation of Co-efficients:

When it is desirable to compare the relative effects of predic tor variables, care must be taken to code these using the same measurement units. If it is not possible, standardised* regression scores must be developed to compare the relative changes in variables measured in different units.

Causation

One of the most ‘naive’ but highly resorted to abuse of multiple regression analysis is the assumption that the level of predictor variables causes the level of criterion variable. However, all they indicate is association between the variables. Association is evidence of causation, but not a proof of it.

Order of importance of predictor variables:

Most analysts have an urge to use measures such as squared partial correlations or squared beta coefficients to rank predictors in order of importance in accounting variation in the criterion. However, if the predictors are correlated, this urge should be tempered. In case of correlated predictors, there is no unambiguous measure of relative importance of predictor variables.

Hypothesis Testing And Parametirc Tests:

Hypothesis Testing:

- The process of hypothesis testing in research is analogous to the judicial processes by which an accused individual is judged in a court of law.
- The person brought before the judge is considered to be innocent.
- The responsibility of proving him guilty rests with the prosecution.
- Let us assume the person accused is innocent by *RI* and that the person is guilty by *RG*.
- If the judge or the jury finds the evidence provided by the prosecution is inconsistent with *RI* they reject *RI* and accept *RG* that is, that the person is guilty.
- The possibilities of the judgement situation are:
    1. The defendant is innocent (*RI* is true) and the judge/jury judges that he is innocent (retains *RI*). Therefore, the decision is correct.
    2. The defendant is guilty (*RI* is false) and judge/jury judges that he is innocent (retains *RI*).Therefore, the decision is wrong.
    3. The defendant is innocent (*RI* is true) and the judge/jury judges he is guilty (rejects *RI*).Therefore, the decision is wrong.
    4. The defendant is guilty (*RI* is false) and the judge/jury decides that he is guilty (reject *RI*). Therefore, the decision is correct.
- In hypothesis testing, much like in a judicial trial, the focus is on evaluating whether to retain or reject a preconceived assumption, known as the null hypothesis (H_0).
- Just as the legal system emphasizes avoiding the wrongful conviction of an innocent person, hypothesis testing aims to minimize the risk of incorrectly rejecting a true null hypothesis (Type I error).
- This caution reflects a scientific preference for missing a true discovery over accepting a false one.
- In classical hypothesis testing, a hypothesis about a population parameter is proposed and tested using sample data, with decisions made based on statistical methods.
- Alternatively, the Bayesian approach incorporates subjective probabilities based on prior beliefs and expert opinions, which are updated as new data is collected, allowing for a dynamic reassessment of the hypothesis.

LOGIC OF HYPOTHESIS TESTING:

- In hypothesis testing, two kinds of hypothesis are involved.
    1. Null Hypothesis.
    2. Research Hypothesis.

NULL HYPOTHESIS:

- The first type is the null hypothesis, which can be assessed using probabilities derived from the sample statistics.
- This is a statement that the statistical differences or relationships have occurred for no other reason than laws of chance operating in an unrestricted manner.
- Suppose we are testing the difference between the means μ1 and μ2 of two populations. The null hypothesis *H*o is stated as,

$$H0 : \mu1 = \mu2$$

RESEARCH HYPOTHESIS:

- The second kind is research hypothesis, which is intended to test the research prediction.
- The null hypothesis is the logical opposite of the research hypothesis.
- Thus, if the null hypothesis is rejected, then the research hypothesis is considered acceptable.
- This is a statement of differences or relationships among phenomena, the acceptance or rejection of which is based upon resolving the null hypothesis, which is its logical alternative. It is stated as,

$$H0 : \mu1 \neq \mu2$$

CHOICE OF PROBABILITY LEVEL:

- A probability level (α) of 0.05 is typically used as the threshold for rejecting the null hypothesis.
- If the calculated probability (p-value) is lower than 0.05, the null hypothesis is rejected, known as the rejection or significance level.

- The significance level is the probability threshold set by the researcher for rejecting the null hypothesis, represented by p.
- If the p-value is higher than the significance level, the test is considered inconclusive, meaning the null hypothesis cannot be rejected.

Substantive significance:

- **Statistical significance** refers to the acceptable level of error in a test, indicating whether the observed results are likely due to chance or not.
- **Substantive significance** focuses on the practical importance or strength of the relationship, which is of greater interest to the researcher in real-world applications. A relationship can be statistically significant but not necessarily meaningful in practice.

Strong and weak testing:

- **Strong vs. Weak Data:** Data is considered strong when it can be modeled using a distribution and is measured on an interval or ratio scale (metric). Weak data is non-metric, measured on an ordinal or nominal scale.
- **Strong vs. Weak Statistical Techniques:** Strong statistical techniques, known as parametric tests, can extract more information from data and are dependent on distribution assumptions. Weak statistical techniques, called non-parametric or distribution-free tests, extract less information and do not rely on distribution assumptions.
- **Application Flexibility:** Traditionally, strong techniques are applied to strong data and weak techniques to weak data. However, researchers can choose to apply strong techniques to weak data or weak techniques to strong data, offering a wider range of testing options.

Degrees of freedom:

- **Definition of Degrees of Freedom (df):** Degrees of freedom refer to the number of independent pieces of information available for estimating statistical parameters.
- **Calculation:** Degrees of freedom are calculated as df=n−kdf = n - kdf=n−k, where nnn is the total number of data points, and kkk is the number of parameters or constraints applied to the data.

- **Application in Hypothesis Testing:** Degrees of freedom are crucial in hypothesis testing techniques, such as χ2\chi^2χ2-tests, t-tests, and F-tests, affecting the sensitivity and accuracy of the results.
- **Significance:** The degrees of freedom indicate how much independent information is available to make inferences about the population, impacting the reliability of the statistical technique.

Errors in Hypothesis Testing:

- **Type I error:** In a statistical test, this is the error committed while rejecting null hypotheses that are true. The probability of making this error is represented by α.
- **Type II error:** In a statistical test this error is committed in accepting null hypotheses that are false. Probability of type II error is represented by β.

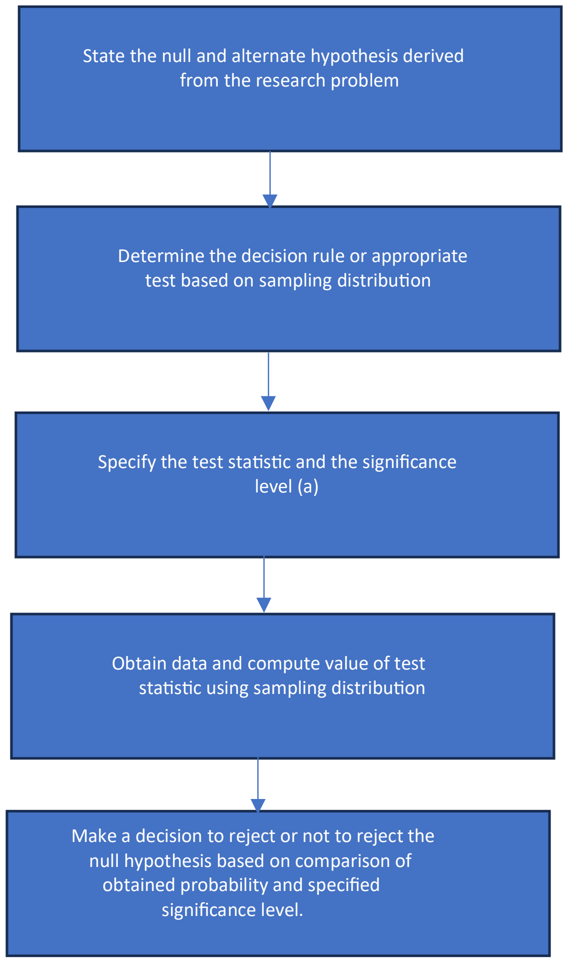
- **Goodness of a Statistical Test:** The effectiveness of a statistical test is measured by two probabilities: α\alphaα (the probability of a Type I error) and β\betaβ (the probability of a Type II error). Increasing the rejection region increases α\alphaα but decreases β\betaβ, and vice versa.
- **Power of the Test:** The power of a statistical test, defined as 1−β1 - \beta1−β, is the probability of correctly rejecting a false null hypothesis. This power increases as the difference between the true value and the hypothesized value of the population parameter becomes larger.
- **Practical Implications:** While α\alphaα is known beforehand, β\betaβ is often unknown. Therefore, when the null hypothesis falls within the acceptance region, it is safer to withhold judgment rather than accept it outright, unless β\betaβ is known. This cautious approach is why researchers support their hypothesis by rejecting the null hypothesis rather than directly testing the research hypothesis.

Reducing type II error:

- The probability of acceptance of *H*o, which is false (type II error),cannot be determined, as the real population parameter is never known.
- There are several means of reducing this probability.
- They are:

1. Increase the level of significance used as the standard for rejecting the null hypothesis (use a = 0.01 in place of 0.05 or even lower).
2. Use a one tailed test rather than a two tailed test if there is a basis for it.
3. Decrease the standard error, that is, take a bigger sample.

Steps in hypothesis testing procedure:



- Five steps are to be performed for a typical testing of a hypothesis, as shown in the flowchart above.
- Generally, in any research, the first three steps are completed before collecting sample data.
- In the statistical analysis of research data, the first and primary task of the researcher is to choose the appropriate test or method of analysis.

- There are a large number of well developed tools and techniques of statistics available to the researcher for this purpose.
- In order to make an appropriate choice of the technique in the second step, the following considerations should be kept in mind.

   a. **Type of analysis:** Descriptive analysis summarizes and describes data within a specific group, using measures like mean and standard deviation. Inferential analysis, however, uses sample data to make generalizations or predictions about a larger population.
   b. **Type of data:** Nominal data categorize objects without involving numerical operations, making mean and median irrelevant, with mode being the only valid measure of central tendency. For ordinal or interval data, the decision between using parametric or non-parametric tests depends on the data's characteristics, such as the distribution and scale. Parametric tests are generally preferred for interval data with a normal distribution, while non-parametric tests are more suitable for ordinal data or when the assumptions for parametric tests are not met.
   c. **The size of the sample:** When interval scaled data is being analysed for inferential purposes, the size of the sample, whether large (> = 30) or small (< 30), determines if a ztest or t-test should be used. In certain cases of multivariate analysis (for example, factor analysis) the minimum size of the sample is dictated by the number of variables being considered.
   d. **Number of variables:** Depending upon whether a single independent variable is analysed (uni-variate) or two or more variables (multi-variate) are to be analysed, both in hypothesis testing and in determining relationships, the techniques adopted may be different.
   e. **Assumptions underlying samples:** When more than one sample is considered for purposes of comparison and testing (whether from the same population or different populations), it is important to know whether the samples are independent or dependent. The following questions are raised: Does selection of one sample limit the selection of the other ? Are the populations normally distributed? Are the variances of the populations same?

f. **Nature of relationship:** In the analysis, whether the researcher is seeking a casual relationship or only an associative relationship will determine the type of technique to be used.

g. **Number of groups:** Techniques for analysing a single group are different from those used for analysing multiple groups. While a single group may be explored for natural variables (for example, factor analysis), comparative analysis and relational analysis may be required for multiple groups.

h. **Levels or differences:** While testing hypothesis, the level of a variable value may be ofinterest (as in a univariate analysis) or the significance of difference between levels of a variable in two or more groups or samples may be of interest as in multivariate analysis.

- In most research, the choice of significance level is determined by the researcher at the time of selecting the test.
- A significance level refers to the probabilitya ofmaking type I error.
- Based on the significance level the critical value(s) of a test static is determined.
- The actual value of the test statistic is compared with the critical value and a decision is made based on the hypothesis (obtained by meaning probability verses a).

IDENTIFICATION OF AN APPROPRIATE TEST FOR HYPOTHESIS TESTING:

- The user can identify the appropriate test for his research (univariate hypothesis testing) by providing the following information.
    1. Whether the scale of measurement is nominal/ordinal or interval.
    2. Whether the test is a parametric test or distribution free test.
    3. Whether the sample involved is a single sample or multiple samples.
    4. Whether the samples are dependent or independent when two or more samples are drawn.

Parametric Tests:

- **Definition:** Parametric tests that make specific assumptions about the parameters (the mean and variance) of the population distribution from which the data is drawn.
- **Normality**: Many parametric tests assume that the data follows a normal distribution (bell-shaped curve)
- **Homogeneity of Variance**: Some parametric tests assume that different groups have similar variances ($\sigma 2$)

Applications :

- Used for continuous variables (y = f(x)).
- Used when data are measured on appropriate interval of ratio scale of measurement .
- Data should follow normal distribution.

Z-Test:

Logic of Z-test: When samples are large (*n* greater than 30), tests for hypothesis concerning the levels of population parameters μp (for a single population), or differences in the population parameters, (μ 1 – μ2) or (*p*1 – *p*2) [two populations] are all based on normally distributed test statistics and may be regarded as essentially the same test and is termed as the Z-Test.

Given that H0 : μ = μ0

When sampling distribution mean = μ0

Two situations may be met with:

- *H*a : is the research hypothesis. The acceptable regions and rejection given are shown in Fig. 14.2(a)—a two tail test that is divided into two equal areas as shown. Test statistic is Zo = (μ – μ0)/(σ0)

  Reject H0 if Zo > Za/2 or Zo < -Za/2
- $\left.\begin{array}{l}\text{If Ha : } \mu > \mu 0 \text{ (1)} \\ \text{If Ha : } \mu < \mu 0 \text{ (2)}\end{array}\right\}$ is the research hypothesis
- Test statistic is Zo = (μ – μ0)/(σ0)
  - Reject H0 if Zo > Za ,Therefore, accept Ha : μ > μ0,
  - Reject Ho if Zo < Za ,Therefore, accept *H*a : μ < μ0.

t-Test:

- It is Used to compare the means of two groups to determine if they are statistically significantly different from each other.
- Independent t-test (comparing two separate groups) .
- Paired t-test (for comparing two related groups) .
- One-Sample t-Test (To compare the mean of a single group to a known value or population mean) .

Logic of the t-test:

When the samples are small (n less than 30), the central limit theorem does not hold good and the test statistic as given in Z-test.

That is, $\frac{\bar{x} - \mu}{S/\sqrt{n}}$ does not follow a normal distribution, but follows a student's *t*-distribution.

Student's *t*-distribution gives greater variances for smaller values of *n* and approaches the normal distribution values (that is, Z) when *n* is very large. However, in t-tests also the population is assumed to be normally distributed. In testing the hypothesis with small samples *t*-distribution values are used (instead of Z values) from the tables.

$$\text{Test statistic } t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$$

when differences between means are tested,

$$t = \frac{(\bar{x}_1 - \bar{x}_2)}{\sigma_{differences}}$$

The numerator is almost always the same. If null hypothesis is true

$$(\bar{x}_1 - \mu_1)(\bar{x}_2 - \mu_2), \text{ or } \{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2) \text{ becomes } (\bar{x}_1 - \bar{x}_2)\}$$

- The difference arises mostly in the denominator, that is, while using *t*-values, it is necessary to know the degrees of freedom that reflect the sample's size.

- If there is a single sample involved, degrees of freedom is (n–1), if two samples are involved, (as finding the difference in means between two populations) the degree of freedom is (n1 + n2 – 2) [(n1 – 1) + (n2 – 1)].
- The degree of freedom associated with any statistic, refers to the number of components that are free to vary. (n–1) degrees of freedom associated with a single sample of n is meant to reflect that if we know (n–1) values, the nth one can be determined easily, hence (n–1) values are free to vary.
- It reflects the number of squared deviations available for estimating variance. As in Ztests, two tail and single tail tests can be used in t-tests too.

Problems:

1. The average shelf life of a deteriorable product is claimed by the manufacturers as 30 months. A random sample of 81 units had a mean of 28.7 months and a standard deviation of 8 months. Perform a hypothesis test using a significance level of 0.05.

Solution:

A one tail Z-test is used (n > 30)

H0 : μ = 30

Ha : μ < 30

a = 0.05 $\bar{x}$=28.7

Test statistic $t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$

= (27.7-30)/(8/$\sqrt{81}$)

=-1.46

for $a = 0.05$, $Z_c = -1.65$.
Reject H_0 if $Z_c < -1.65$
Do not reject H_0 if test is inconclusive and Z > -1.65

2. A worker demands an average time of 15 minutes for an operation. The industrial engineer feels the operation takes much less time than 15 minutes. He observes 16 randomly selected repetitions of the operation and calculates the average as 12.4 minutes with a standard deviation of 1.3 minutes. For a significance level of 0.01, how good is the case for the industrialengineer?

Solution:

H0 : μ = 15

Ha : μ < 15

$a = 0.01$ $\bar{x}$=13.4 n=16 S=1.3

Test statistic $t = \dfrac{\bar{x} - \mu}{S/\sqrt{n}}$

= $(13.4-15)/(1.3/\sqrt{16})$

= -4.92

Degree of Freedom=n-1=16-1=15

Reject H0 if and only if t_c <-2.602

Therefore,Reject H_0 .

3. A customer wants to find out whether the cooking gas supplied by two suppliers is the same with regard to performance in stoves. Eight stoves were selected randomly. One stove was used on the cooking gas supplied by A and another on that supplied by B. The time the stoves worked on the same quantity of gas as supplied by A. Nine stoves were randomly selected.
Time for A—19.7, 18.1, 20.4, 22.6, 18.8, 19.2, 19.3, 21.5
Time for B—18.9, 23.2, 19.8, 22.4, 21.3, 22.2, 21.6, 20.3, 24.7
Are the two cooking gases equivalent?

Solution: Since both the samples are small, *t*–statistic is the appropriate test statistic. The conditions for *t*–statistic to be used are (i) populations are normally distributed, (ii) population variances are same, and (iii) the samples are selected independently.

H0 : μ1 = μ2 $t_c = \pm 1.753$

Ha : μ1≠μ2

(n1,n2<30)

Test Statistic is $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1-1)(n_2-1)}}} \times \dfrac{1}{\sqrt{\dfrac{1}{n_1^2} + \dfrac{1}{n_2^2}}}$

$\bar{x}1$=(19.7+18.1+…….+19.3+21.5)/8 =19.95

$\bar{x}2$=(18.9+23.2+19.8+…..+24.7)/9 =21.6

S_1^2 = $((18.1-19.95)^2 +(18.1-18.19.95)^2 +…..(21.5-19.95)^2)/(8-1)$ =2.2029

Similarly S_2^2 =3.185

T_0 =(19.95-21.6)/((8-1)2.209+(9-1)3.185) *1/($\sqrt{\frac{1}{8}+\frac{1}{9}}$) = -2.056

$T_0<t_c$

Therefore Reject H_0 .

Thus, the performance of the two suppliers, A and B, are not the same.