

TESTING OF HYPOTHESIS

BASIC CONCEPTS AND FORMULA

Basic Concepts

1. Testing Of Hypothesis Or Test Of Significance

It is a statistical procedure to assess the significance of

- (i) Difference between a statistic and corresponding population parameter.
- (ii) Difference between two independent statistics, known as test of significance.

2. Null Hypothesis (H_0)

It asserts that there is no real difference between the sample statistic and sample parameter or between two independent sample statistics.

3. Alternative Hypothesis (H_1)

Any hypothesis complementary to null hypothesis.

4. Possible Errors in Test of Significance

Four possible errors in test of significance:

Type of error	Actual	Decision from sample	Probability of error
1	H_0 is true	Reject H_0	α
2	H_0 is false	Accept H_0	β

5. One-tailed Test

A hypothesis test in which rejection of the null hypothesis occurs for values of test statistic in one tail of the sampling distribution.

6. Two-way Test

A hypothesis test in which rejection of the null hypothesis occurs for values for test statistic in either tail of its sampling distribution.

7. Critical Value

A value that is compared with the test statistic to determine whether H_0 should be rejected.

8. Procedure for Large Sample Test (t-test)

Step 1: Set up Null hypothesis H_0 and alternative hypothesis H_1 .

Step 2: Compute $Z = \frac{t - E(t)}{SE(t)}$

Step 3: Testing significance at desired level, usually 5% & 1%

	At 1% Level	At 5% Level	
Significant values of $ Z $	2.58	1.96	Two tailed test
Significant values of $ Z $	2.33	1.645	One tailed test

9. Analysis of Variance (ANOVA): Test Analysis of variance can be used for testing equality of k population means.

$H_0: \mu_1 = \mu_2 = \dots = \mu_k$

H_1 : Not all population means are equal.

Where m_j = mean of j^{th} population

Let x_{ij} = value of observation i for treatment j

n_j = No. of observation for treatment j

$\bar{x}_j$ = sample mean for treatment j

s_j^2 = sample variance for treatment j

$\bar{x}$ = overall sample mean n_t = Total Sample Size

Sum of Square due to treatment

$$SSTR = \sum_{j=1}^k n_j (\bar{x}_j - \bar{x})^2$$

Mean Square due to treatment

$$MSTR = \frac{SSTR}{k - 1}$$

Sum of Square due to error

$$SSE = \sum_{j=1}^k (n_j - 1) s_j^2$$

$$\text{Mean of square due to error MSE} = \frac{\text{SSE}}{n_t - k}$$

Test Statistic for equality of k population mean

$$F = \frac{\text{MSTR}}{\text{MSE}}$$

ANOVA Table

Source of Variation	Sum of Squarely	Degree of freedom	Mean Square
Treatment	SSTR	$k - 1$	$\text{MSTR} = \frac{\text{SSTR}}{k-1}$
Error	SSE	$n_t - k$	$\text{MSE} = \frac{\text{SSE}}{n_T - k}$
Total	SST	$n_T - 1$	$F = \frac{\text{MSTR}}{\text{MSE}}$

9.1 ANOVA For Randomized Block Design (2- ways classification)

k = No. of treatments

b = No. of blocks

n_T = Total sample size = kb

r = replications

x_{ij} = Value of observation responding to treatment j in block j

$\bar{x}_j$ = sample mean of j^{th} treatment

$\bar{x}_i$ = sample means of i^{th} stock

$\bar{\bar{x}}$ = overall sample mean

Total Sum of Square

$$\text{SST} = \sum_{i=1}^b \sum_{j=1}^k (x_{ij} - \bar{\bar{x}})^2$$

Sum of Square due to treatments

$$\text{SSTR} = b \sum_{j=1}^k (\bar{x}_j - \bar{\bar{x}})^2$$

Advanced Management Accounting

Sum of Square due to blocks

$$SSBL = k \sum_{i=1}^b (\bar{x}_i - \bar{x})^2$$

Sum of Square due to error SSE = SST – SSTR – SSBL

ANOVA TABLE

Source of Variation	Sum of Square	Degree of freedom	Mean Square	F
Treatment (TR)	SSTR	K – 1	$MSTR = \frac{SSTR}{K - 1}$	$\frac{MSTR}{MSE}$
Block (BL)	SSBL	b - 1	$MSBL = \frac{SSBL}{b - 1}$	$\frac{MSBL}{MSE}$
TR x BL	SSTB	(k-1) (b-1)	$\frac{MSTB}{\frac{SSTB}{(k-1)(b-1)}} = \frac{MSTB}{MSE}$	
Error	SSE	Kb (r – 1)	$MSE = \frac{SSE}{kb (r - 1)}$	
Total	SST	n _T - 1		

Basic Formulas

1. Test Statistic for Hypothesis Test, about a Population Mean is known

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}}$$

μ = population mean
n = sample size

2. Test Statistic for Hypothesis Test, about a population Mean; is unknown

$$t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

s = sample mean

3. Test Statistic for Hypothesis Tests about a Population Proportion

$$Z = \frac{p - p_0}{\sqrt{\frac{p_0 (1 - p_0)}{n}}}$$

Question 1

Write a short note on the procedure in hypothesis testing.

Answer

Procedure in Hypothesis Testing: Following procedure is followed in hypothesis testing:

1. Formulate the hypotheses: Set up a null hypothesis stating, for e.g. $H_0: \theta = \theta_0$ and an alternative hypothesis H_1 , which contradicts H_0 . H_0 and H_1 cannot be done simultaneously. If one is true, the other is false.
2. Choose a level of significance, i.e. degree of confidence. This determines the acceptance rejection region. For example, $Z_{.05}$ in a 2 tailed 'Z' test is.
3. Select test statistic: For $n > 30$, Z statistic is used, implying normal distribution for large samples. For small samples, we use t_1 , F_1 and χ^2 distribution.
4. Compute the sample values according to the test statistic.
5. Compare with the table value of the statistic and conclude.

Question 2

A factory manager contends that the mean operating life of light bulbs of his factory is 4,200 hours. A customer disagrees and says it is less. The mean operating life for a random sample of 9 bulbs is 4,000 hours, with a sample standard deviation of 201 hours. Test the hypothesis of the factory manager, given that the critical value of the test statistic as per the table is (-) 2.896.

Answer

$$\begin{aligned}\text{Manager's Hypothesis} \quad H_0 \quad \mu_0 &= 4,200 \\ H_1 \quad \mu &< 4,200 \text{ (Left Tail test)} \\ t &= \frac{\bar{X} - \mu_0}{\sigma},\end{aligned}$$

$$\text{where } \sigma = \frac{s}{\sqrt{n}} = \frac{201}{\sqrt{9}} = \frac{201}{3} = 67$$

$$t = \frac{4,000 - 4,200}{67} = \frac{-200}{67} = -2.985$$

Calculated $t = 2.985$, < table value of $t_{.01} \text{ (sdf)}$ which is -2.896

Hence reject the null hypothesis H_0 . i.e. Accept H_1

The customer's claim is correct.

Question 3

Advanced Management Accounting

In the past, a machine has produced pipes of diameter 50 mm. To determine whether the machine is in proper working order, a sample of 10 pipes is chosen, for which mean diameter is 53 mm and the standard deviation is 3 mm. Test the hypothesis that the machine is in proper working order, given that the critical value of the test statistic from the table is 2.26.

Answer

Null Hypothesis $H_0 : \mu = 50$ mm i.e. the M/c works properly.

$H_1 : \mu \neq 50$ mm. i.e. the M/c does not work properly

Sample Size = 10, small.

use 't' statistic

$$t = \frac{\bar{x} - \mu}{S / \sqrt{n-1}}$$

$$\bar{x} = 53$$

$$\mu = 50$$

$$n = 10; \sqrt{n-1} = \sqrt{9} = 3$$

$$S = \text{std dev} = 3$$

$$T = \frac{53-50}{3/3} = \frac{3}{1} = 3$$

Table Value = 2.26

Calculated t > table value

Reject H_0

i.e. The M/c is not working properly.

Question 4

A manufacturer claimed that at least 95% of the equipment which he supplied to a factory conformed to specifications. An examination of a sample of 200 pieces of equipment revealed that 18 were faulty. Test this claim at a significance level of (i) 0.05 (ii) 0.01.

Answer

In the usual notations, we are given $n = 200$. x = No. of pieces conforming to specifications in the sample = $200 - 18 = 182$.

$$\therefore P = \text{Proportion of pieces conforming to specifications in the sample} = \frac{182}{200} = 0.91.$$

Null hypothesis. $H_0 : P \geq 0.95$, i.e., the proportion of pieces conforming to specifications in the lot is at least 95%.

Alternative Hypothesis. H_1 ; $P < 0.95$ (Left-tailed alternative).

It will suffice to test H_0 : $P = 0.95 \Rightarrow Q = 1 - P = 0.05$

Level of significance (i) $\alpha = 0.05$, (ii) $\alpha = 0.01$

Test statistic. Under H_0 , the test statistic is $Z = \frac{P - E(P)}{SE(P)} = \frac{P - P}{\sqrt{PQ/n}} \sim N(0,1)$,

Since sample is large

$$= \frac{0.91 - 0.95}{\sqrt{0.95 \times 0.05/200}} = \frac{-0.04}{\sqrt{0.00237}} = \frac{-0.04}{0.0154} = -2.6.$$

(i) Significance at 5% level of significance.

Since the alternative hypothesis is one-sided (left-tailed), we shall apply left-tailed test for testing significance of Z. The significant value of Z at 5% level significance for left-tail test is -1.645 .

Since computed value of $Z = -2.6$ is less than -1.645 (or since $|z| > 1.645$), we say Z is significant (as it lies in the critical region) and we reject the null hypothesis at 5% level of significances. Hence, the manufacturer's claim is rejected at 5% level of significance.

(ii) Significance at 1% level of significance. The critical value of Z at 1% level of significance for single-tailed (left-tailed) test is -2.33 . Since the computed value $Z = -2.6$ is less than -2.33 (is $|z| > 2.33$),

$\therefore H_0$ is rejected at 1% level of significance also.

Question 5

For the following data representing the number of units of production per day turned out by five workers using from machines, set-up the ANOVA table (Assumed Origin at 20).

Workers	Machine Type			
	A	B	C	D
1.	4	-2	7	-4
2.	6	0	12	3
3.	-6	-4	4	-8
4.	3	-2	6	-7
5.	-2	2	9	-1

Advanced Management Accounting

Answer

Null Hypothesis

- (a) The machines are homogenous

$$\text{i.e., } \mu_A = \mu_B = \mu_C = \mu_D$$

- (b) The workers are homogeneous

$$\text{i.e., } \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$$

Alternative Hypothesis

- (a) At least two of the machines differ significantly

- (b) At least two of the workers differ significantly

In the usual notation, we have:

$$K = 5, H = 4, N = KH = 5 \times 4 = 20$$

$$G = \sum \sum X_{ij} = 20;$$

Calculation for Various S.S

Workers	Machine Type				Total
	A	B	C	D	
I	4	-2	7	-4	$R_1 = 5$
II	6	0	12	3	$R_2 = 21$
III	-6	-4	4	-8	$R_3 = -14$
IV	3	-2	6	-7	$R_4 = 0$
V	-2	2	9	-1	$R_5 = 8$
Total	$C_1 = 5$	$C_2 = -6$	$C_3 = 38$	$C_4 = -17$	$G = 20$

$$\text{Corrector Factor (CF)} = \frac{G^2}{n} = \frac{20^2}{20} = 20$$

$$\begin{aligned} \text{Raw S.S (RSS)} &= \sum \sum X_{ij}^2 \\ &= [(16+4+49+16)+(36+0+144+9) + (36+16+16+64) + (9+4+36+49) \\ &\quad + (4+4+81+1)] \\ &= 594 \end{aligned}$$

$$\text{Total S.S} = \text{RSS} - \text{CF} = 594 - 20 = 574$$

$$\begin{aligned}\text{S.S Rows (Workers)} &= \frac{R_1^2 + R_2^2 + R_3^2 + R_4^2 + R_5^2}{4} - CF \\ &= \frac{5^2 + 21^2 + (-14)^2 + 0 + 8^2}{4} - 20 \\ &= \frac{25 + 441 + 196 + 64 + 80}{4} = \frac{646}{4} = 161.5\end{aligned}$$

$$\begin{aligned}\text{S.S Columns (Machine Type)} &= \frac{C_1^2 + C_2^2 + C_3^2 + C_4^2}{5} - CF \\ &= \frac{5^2 + (-6)^2 + 38^2 + (-17)^2}{5} - 20 \\ &= \frac{25 + 36 + 1,444 + 289 + 100}{5} \\ &= \frac{1,694}{5} = 338.8\end{aligned}$$

$$\therefore \text{SSE} = \text{Error S.S} = \text{TSS} - \text{SSR} - \text{SSC}$$

$$= 574 - 161.5 - 338.8$$

$$= 73.7$$

Since the various sum of the squares are not affected by change of origin, the ANOVA table for the original data and the given data obtained on changing the origin to 20 will be same and is given in following table.

Degrees of Freedom for various S.S

$$\text{d.f for TSS} = n - 1 = 20 - 1 = 19$$

$$\text{d.f for Rows (Workers)} = 5 - 1 = 4$$

$$\text{d.f for Column (Machines)} = 4 - 1 = 3$$

$$\text{d.f for SSE} = 19 - (4 + 3) = 12$$

$$\text{OR d.f for SSE} = (\text{d.f for Rows}) \times (\text{d.f for columns})$$

$$= (3 \times 4) = 12$$

ANOVA TABLE

Advanced Management Accounting

Sources of variation	d.f	S.S	$MSS = \frac{S.S}{d.f}$	Variance Ratio (F)
Rows (Workmen)	4	161.5	40.38	$\frac{40.38}{6.14} = 6.58 \sim F(4,12)$
Columns (Machine)	3	33.8	112.93	$\frac{112.93}{6.14} = 18.39 \sim F(3,12)$
Errors	12	73.7	6.14	
Total	19	574		

Question 6

Given below in the contingency table for production is three shifts and the number of defective good turn out- Find the value of C. It is possible that the number defective goods depends on the shifts then by them, No of Shifts:

Shift	I Week	II Week	III Week	Total
I	15	5	20	40
II	20	10	20	50
III	25	15	20	60
	60	30	60	150

Answer

Let H_0 : Defective is good does not depend upon the shift run by the factory the first Expected value is

$$= E = \frac{40 \times 60}{150} = 16$$

O	E	O-E	$(O-E)^2$	$(O-E)^2/E$
15	16	-1	1	0.063
20	20	0	0	0
25	24	1	1	0.042
5	8	-3	9	1.125
10	10	0	0	0
15	12	3	9	0.750
20	16	4	16	1.0

Testing of Hypothesis

20	20	0	0	0
20	24	-4	16	0.667
				3.647

$$D: F = V = (r - 1)(c - 1) = (3 - 1)(3 - 1) = 4$$

$$: \Psi^2(4, 0.05) = 9.488$$

Here, the calculated value of Ψ^2 is less than of table value.

Hence, the hypothesis is accepted.

i.e., the number of defective does not depend on shift run by the factory.

EXERCISE

Question 1

The contingency table below summarize the results obtained in a study conducted by a research organization with respect to the performance of four competing brands of tooth paste among the users

	Brand A	Brand B	Brand C	Brand D	Total
No. of Cavities	9	13	17	11	50
One of five	63	70	85	82	300
More than five	28	37	48	37	150
Total	100	120	150	130	500

Test the hypothesis that incidence of cavities is independent of the brand of the tooth paste used.
Use level of significance 1% and 5%.

Answer

Incidence of cavities is independent of the brand of the tooth paste used.

Question 2

Below are given the yield (in kg.) per acre for 5 trial plots of 4 varieties of treatment. Carry out an analysis of variance and state conclusion

	Treatment			
Plot no.	1	2	3	4
1	42	48	68	80
2	50	66	52	94
3	62	68	76	78
4	34	78	64	82
5	52	70	70	66

Answer

∴ The null hypotheses is rejected

∴ The treatment does not have same effect.

Question 3

The sales data of an item in six shops before and after a special promotional campaign are as under

Testing of Hypothesis

Shops	A	B	C	D	E	F
Before Campaign	53	28	31	48	50	42
After Campaign	58	29	30	55	56	45

Can the campaign be judged to be a success?

Test at 5% level of significance using t-test.

Answer

H_0 is rejected at 5% level of significance and we conclude that the special promotional campaign has been effective in increasing the sales.