UNIT-V

TEST OF SIGNIFICANCE

Part-A Large samples

Model No 5.1: Test of Significance of a single mean

Model No 5.2: Test of Significance for difference of means

Model No 5.3: Test of Significance for Single Proportion

Model No 5.4: Test for equality of two Proportions (or)

Test of significance of difference between two sample Proportions

Test of hypothesis: In many circumstances, to arrive at decisions about the population on the basis of sample information, we make assumptions (or guesses) about the population parameters involved. Such an assumption (or statement) is called a statistical hypothesis which may or may not be true. The procedure which enables us to decide on the basis of sample results whether a hypothesis is true or not, is called Test of Hypothesis or Test of significance.

Null Hypothesis: A null hypothesis is the hypothesis which asserts there is no significant difference between statistic and the population parameter and whatever observed difference is there, is merely due to fluctuations in sampling from the population.

For applying test of significance, we first set up a hypothesis- a definite statement about population parameter. Such a hypothesis is usually a hypothesis of no difference, is called null hypothesis.

Null hypothesis is a statement of no differences.

Alternative Hypothesis (H_1): It is denoted by H_1, is the opposite statement of null hypothesis.

Type I and Type II errors

Type –I error: Reject null hypothesis when it is true

Type –I error: Accept null hypothesis when it is false.

Level of significance: The selection of the level of significance depends on the choice of the researcher. Generally level of significance is taken to be 5% or 1%. It is denoted by α, is the probability of committing type I error. Thus L.O.S. measures the amount of risk or error associated in taking decisions. L.O.S. is expressed in percentage. Thus L.O.S. $\alpha = 5\%$ means that there are 5 chances in 100 that null hypothesis is rejected when it is true.

Critical Region (C.R.): In any test of hypothesis, a test statistic S^*, calculated from the sample data is used to accept or reject the null hypothesis. Consider the area under the probability curve of the sampling distribution of the test statistic S^*. This area under the probability curve is divided into two regions, namely the region of rejection where N.H. is rejected and the region of acceptance where N.H. is accepted. Thus critical region is the region of rejection of N.H. The area of the critical region equals to the level of significance α. Note that C.R. always lies on the tail of the distribution.

Right tailed test: When the alternative hypothesis H_1 is of the greater than type *i.e.,* $H_1 : \mu > \mu_0$ *or* $H_1 : \sigma_1^2 > \sigma_2^2$ *etc*. Then the entire critical region of area α lies on the right side of the curve. In such case the test of hypothesis is known as right tailed test.

Left tailed test: When the alternative hypothesis H_1 is of the less than type *i.e.,* $H_1 : \mu < \mu_0$ *or* $H_1 : \sigma_1^2 < \sigma_2^2$. *etc*. Then the entire critical region of area α lies on the left side of the curve. In such case the test of hypothesis is known as left tailed test.

Two tailed test: When the alternative hypothesis H_1 is of the not equals type *i.e.,* $H_1 : \mu \neq \mu_0$ *or* $H_1 : \sigma_1^2 \neq \sigma_2^2$ *etc*. Then the entire critical region of area α lies on the both sides of the curve. In such case the test of hypothesis is known as two tailed test.

PROCEDURE FOR TESTING OF HYPOTHESIS:

(i) Null Hypothesis (H_0)**:** Define a Null Hypothesis H_0 taking into consideration the nature of the problem and data involved.

(ii) Alternative Hypothesis (H_1)**:** Define an Alternative Hypothesis H_1 so that we could decide whether we should use one-tailed or two-tailed test.

(iii) Level of Significance (α)**:** Select the appropriate level of significance α depending on the reliability of the estimates and permissible risk.

(iv) Test Statistic: Compute the test statistic $z = \dfrac{t - E(t)}{S.E.\ of\ t}$

(v) Conclusion: We compare the computed value of the test statistic Z with the critical value Z_α at a given level of significance α

(i) If $|z| < z_\alpha$ we accept the Null Hypothesis H_0

(ii) If $|z| > z_\alpha$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1

Critical Values of z			
Level of Significance α	**1%**	**5%**	**10%**
Critical values for two-tailed test	$\lvert Z_\alpha \rvert = 2.58$	$\lvert Z_\alpha \rvert = 1.96$	$\lvert Z_\alpha \rvert = 1.645$
Critical values for right-tailes test	$Z_\alpha = 2.33$	$Z_\alpha = 1.645$	$Z_\alpha = 1.28$
Critical values for left-tailed test	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

Formulas for large samples

Large Samples: $n \geq 30$		Test Statistic	Identification
1	**Test of Significance of a single mean**		
a.	Direct	$z = \dfrac{\bar{x} - \mu}{\left(\sigma / \sqrt{n}\right)}$	One sample mean One population mean One population S. D One sample size
b.	Population S. D. (σ) is not known. *Sample S. D. (S) is known	$z = \dfrac{\bar{x} - \mu}{\left(s / \sqrt{n}\right)}$	One sample mean One population mean One sample S. D One sample size
2	**Test of Significance for difference of means**		
a.	Direct	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ or $z = \dfrac{\left(\bar{x}_1 - \bar{x}_2\right) - \left(\mu_1 - \mu_2\right)}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}$ $\mu_1 - \mu_2 = 0$ $\qquad$ $\mu_1 - \mu_2 \neq 0$	Two sample means Two population S. D.s Two sample sizes
b.	When the samples are taken from the same population. $\sigma_1 = \sigma_2 = \sigma$	$z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}$	Two sample means One Population S. D.s Two sample sizes
c.	The sample variances s_1 & s_2 are given	$z = \dfrac{\left(\bar{x}_1 - \bar{x}_2\right)}{\sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}}}$	$s_1^2 = \dfrac{\sum\left(x_1 - \bar{x}_1\right)^2}{n_1 - 1}$ $s_2^2 = \dfrac{\sum\left(x_2 - \bar{x}_2\right)^2}{n_2 - 1}$
3	**Test of Significance for Single Proportion**		
a.	Direct	$z = \dfrac{p - P}{\sqrt{\dfrac{PQ}{n}}}$	No sample mean No S. D.s One sample size Observations are given (Probability)
4	**Test for equality of two Proportions (or)** **Test of significance of difference b/w two sample Proportions**		
a	Direct	$z = \dfrac{P_1 - P_2}{\sqrt{\dfrac{P_1 Q_1}{n_1} + \dfrac{P_2 Q_2}{n_2}}}$	No sample mean No S. D.s Two sample sizes Two Observations are given (Probability)
b	**Method of pooling:** Two Sample proportions p_1 and p_2 into a single proportion p	$z = \dfrac{p_1 - p_2}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $\quad$ $z = \dfrac{\left(p_1 - p_2\right) - \left(P_1 - P_2\right)}{\sqrt{pq\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$ $P_1 - P_2 = 0$ $\qquad$ $P_1 - P_2 \neq 0$	$p = \dfrac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \dfrac{x_1 + x_2}{n_1 + n_2}$ $q = 1 - p$

TEST OF SIGNIFICANCE FOR LARGE SAMPLES

Model No 5.1: Test of significance for single mean:

(i) Null Hypothesis (H_0): $\bar{x} = \mu$ i.e., "there is no significance difference between the sample mean and population mean" or "the sample has been drawn from the population"

(ii) Alternative Hypothesis (H_1): (i) $\bar{x} \neq \mu$ or (ii) $\bar{x} < \mu$ or (iii) $\bar{x} > \mu$

(iii) Level of Significance (α): Set a level of significance

(iv) Test Statistic:

Case(i): When the S.D. of the population σ is known, The test statistic $z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$

Case(ii): When the S.D. of the population is not known. In this case, we take S.D. of the sample

The test statistic $z = \frac{\bar{x} - \mu}{s / \sqrt{n}}$

(v) Conclusion: (i) If $|z| < z_\alpha$ we accept the Null Hypothesis H_0

(ii) If $|z| > z_\alpha$ we reject the Null Hypothesis H_0 i.e., we accept the Alternative Hypothesis H_1

Problem 1: According to the norms established for a mechanical aptitude test, persons who are 18 years old have an average height of 73.2 with a S.D. of 8.6. If ~~4~~ 31 randomly selected persons of that age averaged 76.7, test the hypothesis $\mu = 73.2$ against the alternative hypothesis $\mu > 73.2$ at the 0.01 level of significance.

Solution: Sample size $n = 31$, sample Mean $\bar{x} = 76.7$, Population Mean $\mu = 73.2$, Population S.D $\sigma = 8.6$

There is No Significant difference between Population Mean and sample Mean. The Sample is drawn from Population

(i) Null Hypothesis (H_0): $\mu = 73.2$, $\bar{x} = \mu$

(ii) Alternative Hypothesis (H_1): $\mu > 73.2 \Rightarrow$ RIGHT-TAILED TEST.

(iii) Level of Significance (α): $\alpha = 0.01$ Single Tailed Test $\Rightarrow 0.5 - \alpha = 0.5 - 0.01 = 0.49$

(iv) Test Statistic: The test statistic

$$z = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{76.7 - 73.2}{\left(\frac{8.6}{\sqrt{31}}\right)} = \cancel{0.8139}\ 2.2659$$

(v) Conclusion: Tabulated value of $Z_{tab} = 2.33$

Calculated value of $Z_{cal} = \cancel{0.8139}\ 2.2659$

Calculated value of	Tabulated value of
$\lvert Z_{cal} \rvert$ <	Z_{tab}

Hence, The Null Hypothesis is Accepted i.e. $\bar{x} = \mu$ (or) There is No Significant difference b/w Sample Mean & Population Mean (or) The Sample is drawn from Same Population.

Problem 2: A sample of 64 students have a mean weight of 70kgs. Can this be regarded as a sample from a population with mean weight 56kgs and standard deviation 25kgs .

Solution: Given Sample size $n=64$, Sample Mean $\bar{x}=70$, Population Mean $\mu=56$
$\sigma=25$ Population S.D

(i) Null Hypothesis (H_0): $\bar{x}=\mu$, $\mu=56$

(ii)Alternative Hypothesis (H_1): $\mu\neq 56$ Two-Tailed Test

(iii) Level of Significance (α): $\alpha=5\%=0.05 \Rightarrow 0.5-\frac{\alpha}{2}=0.5-\frac{0.05}{2}=0.4475$ $\boxed{Z_{tab}=1.96}$

(iv) Test Statistic: The test statistic

$$Z_{cal}\left(\frac{\bar{x}-\mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}\right)=\frac{70-56}{\left(\frac{25}{\sqrt{64}}\right)}=4.48$$

(v) Conclusion: Tabulated value of $Z_{tab}=1.96$

Calculated value of $Z_{cal}=4.48$

Calculated value of Tabulated value of

$\boxed{Z_{cal} > Z_{tab}}$ Null Hypothesis is Rejected

Problem 3: A sample of 900 members has a mean of 3.4 cms and S.D. 2.61 cms. Is this sample has been from a large population of mean 3.25 cm and S.D. 2.61 cms. If the population is normal and its mean is unknown find the 95% confidence limits of true mean.

Solution: Sample : Size $n=900$, Mean $\bar{x}=3.4$, S.D $S=2.61$

Population: Mean $\mu=3.25$, SD $\sigma=2.61$

(i) Null Hypothesis (H_0): $\bar{x}=\mu$, $\mu=3.25$

(ii)Alternative Hypothesis (H_1): $\mu\neq 3.25$ Two-Tailed Test

(iii) Level of Significance (α): $\alpha=5\%=0.05 \Rightarrow 0.5-\alpha/2=0.475$ $Z_{tab}=1.96$

(iv) Test Statistic: The test statistic

$$Z_{cal}=\left(\frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}\right)=\frac{3.4-3.25}{\left(\frac{2.61}{\sqrt{900}}\right)}=1.724$$

(v) Conclusion: Tabulated value of $Z_{tab}=1.96$

Calculated value of $Z_{cal}=1.72$

Calculated value of Tabulated value of

$\boxed{Z_{cal} < Z_{tab}}$ Null Hypothesis is Accepted

The confidence limits are

$$=\bar{x}\pm Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$

$$=3.4\pm 1.96\left(\frac{2.61}{\sqrt{900}}\right)$$

$$=[3.2294, 3.5705]$$

Problem 4: A sample of 400 items is taken from a population whose S.D. is 10. The mean of the sample is 40. Test whether the sample has come from a population with mean 38. Also calculate 95% confidence interval for the population.

Solution: Sample: Size $n = 400$, Mean $\bar{x} = 40$,

Population: Mean $\mu = 38$, S.D $\sigma = 10$

(i) Null Hypothesis (H_0): $\bar{x} = \mu$, $\mu = 38$

(ii) Alternative Hypothesis (H_1): $\mu \neq 38$ Two Tailed Test

(iii) Level of Significance (α): $\alpha = 5\% = 0.05 \Rightarrow 0.5 - \alpha/2 = 0.475$ $\boxed{Z_{tab} = 1.96}$

(iv) Test Statistic: The test statistic

$$Z_{cal} = \left(\frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}\right) = \frac{40 - 38}{\left(\frac{10}{\sqrt{400}}\right)} = 4$$

(v) Conclusion: Tabulated value of $Z_{tab} = 1.96$

Calculated value of $Z_{cal} = 4$

Calculated value of Tabulated value of

$\boxed{Z_{cal} > Z_{tab}}$ Null Hypothesis is Rejected

The confidence limits are

$$\bar{x} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right) = 40 \pm 1.96\left(\frac{10}{\sqrt{400}}\right) = [39.02, 40.98]$$

Problem 5: An ambulance service claims that it takes on the average less than 10 minutes to reach its destination in emergency calls. A sample of 36 calls has a mean of 11 minutes and the variance of 16 minutes. Test the claim at 0.05 level of significance.

Solution: Sample: size $n = 36$, Mean $\bar{x} = 11$, Variance $s^2 = 16$, S.D $s = 4$

Population: Mean $= \mu = 10$

(i) Null Hypothesis (H_0): $\bar{x} = \mu = \mu = 10$

(ii) Alternative Hypothesis (H_1): $\mu < 10$ left Tailed Test

(iii) Level of Significance (α): $\alpha = 0.05 \Rightarrow \alpha = 0.05 \Rightarrow 0.5 - \alpha = 0.5 - 0.05 = 0.45$

(iv) Test Statistic: The test statistic

$\boxed{Z_{tab} = 1.645}$

$$Z_{cal} = \frac{\bar{x} - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)} = \frac{11 - 10}{\left(\frac{4}{\sqrt{36}}\right)} = 1.5$$

(v) Conclusion: Tabulated value of $Z_{tab} = 1.645$

Calculated value of $Z_{cal} = 1.5$

Calculated value of Tabulated value of $\boxed{Z_{cal} < Z_{tab}}$ H_0 Accepted

Problem 6: In a random sample of 60 workers, the average time taken to set to work is 33.8 minutes with a standard deviation of 6.1 minutes. Can we reject the null hypothesis $\mu = 32.6$ minutes in favor of alternative null hypothesis $\mu > 32.6$ at ~~$\alpha = 0.025$ level of significance~~ $\alpha = 5\% = 0.05$

Solution: Sample: Size $n = 60$, Mean $\bar{x} = 33.8$, S.D $s = 6.1$

Population: Mean $\mu = 32.6$

(i) Null Hypothesis (H_0): $\mu = \bar{x} \Rightarrow \mu = 32.6$

(ii) Alternative Hypothesis (H_1): $\mu > 32.6$ Right Tailed Test

(iii) Level of Significance (α): $\alpha = 5\% = 0.05 \Rightarrow 0.5 - \alpha = 0.5 - 0.05 = 0.45$ $\boxed{Z_{tab} = 1.645}$

(iv) Test Statistic: The test statistic

$$Z_{cal} = \frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)} = \frac{33.8 - 32.6}{\left(\frac{6.1}{\sqrt{60}}\right)} = 1.523$$

(v) Conclusion: Tabulated value of $Z_{tab} = 1.645$

Calculated value of $Z_{cal} = 1.523$

Calculated value of Tabulated value of $\boxed{Z_{cal} < Z_{tab}}$ H_0 Accepted

Problem 7: An insurance agent has claimed that the average age of policy holders who issue through him is less than the average for all agents which is 30.5 years. A random sample of 100 policy holders who had issued through him gave the following age distribution.

Age	16-20	21-25	26-30	31-35	36-40
No. of persons	12	22	20	30	16

Solution:

(i) Null Hypothesis (H_0): $\mu = \bar{x} \Rightarrow \mu = 30.5$

Problem-7: Now, we have to find Sample Mean & Sample S.D, since population S.D is Not Given.

Population Mean $\mu = 30.5$

x	true 'x'	mid 'x'	f	d=x-A	fd	d^2	fd^2
16-20	15.5-20.5	18	12	-10	-120	100	1200
21-25	20.5-25.5	23	22	-5	-110	25	550
26-30	25.5-30.5	28 A	20	0	0	0	0
31-35	30.5-35.5	33	30	5	150	5	750
36-40	35.5-40.5	38	16	10	160	100	1600
			$\Sigma fi = 100$		$\Sigma fd = 80$		$\Sigma fd^2 = 4100$

$$\bar{x} = A + \frac{\Sigma fd}{\Sigma f} = 28 + \frac{80}{100} = 28.8$$

$$S.D = \sqrt{\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2} = \sqrt{\frac{4100}{100} - \left(\frac{80}{100}\right)^2} = \boxed{S = 6}$$

$$\boxed{S = 6.352}$$

(ii)Alternative Hypothesis (H_1): $\mu < 30.5$ Left Tailed Test

(iii) Level of Significance (α): $\alpha = 0.05 \Rightarrow 0.5 - \alpha = 0.45$ $\boxed{Z_{tab} = 1.645}$

(iv) Test Statistic: The test statistic

$$Z_{cal} = \left(\frac{\bar{x} - \mu}{\left(\frac{s}{\sqrt{n}}\right)}\right) = \left(\frac{28.8 - 30.5}{\left(\frac{6.352}{\sqrt{100}}\right)}\right) = -2.6763$$

(v) Conclusion: Tabulated value of $Z_{tab} = 1.645$

Calculated value of $|Z_{cal}| = 2.6763$

Calculated value of Tabulated value of

$\boxed{|Z_{cal}| > Z_{tab}}$ $\therefore$ Ho is Rejected.