Unit-3.

1

(1a) Define Alternative Hypothesis ?

Sol:- Any hypothesis which contradicts the null hypothesis is called an Alternative Hypothesis. It is denoted by H_1. The two hypothesis H_0 & H_1 are so that if one is true, the other is false and vice versa. For example if we want to test the null hypothesis thus the population has a special mean μ_0 (say) i.e, $H_0 = \mu = \mu_0$, then alternative hypothesis would be

i) $H_1 : \mu \neq \mu_0$ (Two-tailed alternative hypothesis).

ii) $H_1 : \mu > \mu_0$ (Right-tailed alternative hypothesis)

iii) $H_1 : \mu < \mu_0$ (left-tailed alternative hypothesis)

b) Define Critical Region?

Sol:- A region corresponding to a satistics 't' in the sample space S which leads to the rejection of H_0 is called critical region (or) Rejection region. Those region which lead to the acceptance or H_0 give us a region called acceptance region.

c). Define type I and II errors ?

i) Type-I error :- Reject H_0 when it is true

If the null hypothesis H_0 is true but it rejected by test procedure then the error made is called Type-I error or "α" error.

ii) Type-II error :- Accept H_0 when it is wrong that is accepted by test then error ~~committed is called~~ Type H_0 when H_1 is true.

If Null hypothesis H_0 is false but it is accepted by test then error committed is called Type-II error or β error

(d) Write the four important tests to test the significance under large sample tests. (2)

Sol:- 1. Testing of significance for single proportion.

2. Testing of significance for difference of proportions.

3. Testing of significance for single mean.

4. Testing of significance for difference of means.

(e) Derive critical values of z for both two tailed and single tailed test at 1%, 5% and 10% level of significance.

Sol:-

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 tailed test.	$Z_\alpha = 2.33$	$Z_\alpha = 1.64$	$Z_\alpha = 1.28$
Critical values for left tailed test.	$Z_\alpha = -2.33$	$Z_\alpha = -1.645$	$Z_\alpha = -1.28$

(2a)

(3). A sample of 64 students has a mean weight of 70kg. Can this be regarded as a sample from a population which mean weight 56 kgs and standard deviation 25 kgs.

Sol:- Given that,

The mean of the population $\mu = 56$ kgs

and standard deviation $\sigma = 25$kg.

The mean of sample $\bar{x} = 70$ kgs

and n = sample size = 64.

NH H_0 : A sample of 64 students with mean weight of 70kgs can be regarded as a sample from a population with mean weight 56 kgs.

(3)

Alternative Hypothesis H_1: Sample cannot be regarded as one coming from the population.

level of significance : 0.05 (assumption)

The test statistics is $z = \dfrac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \dfrac{70-56}{\left[\frac{25}{\sqrt{64}}\right]} = 4.48$

The null hypothesis H_0 is rejected, since $|z| > 1.645$

Note:- The null hypothesis can be rejected even at 1% level of significance

~~Short notes on procedure for testing of Hypothesis?~~

~~Sol:- Procedure for testing hypothesis~~ :-

~~Step 1~~:-

(2b) A Die is tossed 960 times and it falls with 5 upwards 184 times Is the die unbiased at a level of significance of 0.01.

Sol:- Given that.

$n = 960$.

The probability of throwing 5 with one die $p = \frac{1}{6}$

Then $q = 1 - p \Rightarrow 1 - \frac{1}{6} = \frac{5}{6}$

Mean $\mu = np \Rightarrow \overset{160}{\cancel{960}} \times \frac{1}{\cancel{6}} \Rightarrow 160$

~~Mean~~ The variance is $\sigma^2 = npq \Rightarrow 960 \times \frac{1}{6} \times \frac{5}{6} = 133.33$.

The standard deviation is $\sigma = \sqrt{133.33} = 11.5$.

$\therefore x =$ no. of success $= 184$

NH H_0 : The die is unbiased ($\mu = \mu_0$)

NH H_1 : The die is biased ($\mu \neq \mu_0$)

level of significance : $\alpha = 0.01$ (1%)

The test statistic is $z = \frac{x-\mu}{\sigma} = \frac{184-160}{11.54} \Rightarrow \frac{24}{11.54} = 2.07$

The z value at 1% level of significance is 2.58 has $|z| < 2.58$, the null hypothesis H_0 is accepted at 1% level of significance

$\therefore$ The die is unbiased at 1% level of significance

2C. Among 900 people in a state 90 are found to be chapatti eaters construct 99% of confidence interval for the proportion.

Sol:- Given $x = 90$, $n = 900$.

$\therefore P = \frac{x}{n} = \frac{90}{900} = \frac{1}{10} = 0.1$ and $Q = 1-P = 0.9$

Now $\sqrt{\frac{PQ}{n}} = \sqrt{\frac{(0.1)(0.9)}{900}} = 0.01$

confidence interval is

$$\left[P - \frac{z_\alpha}{2}\sqrt{\frac{PQ}{n}}, P + \frac{z_\alpha}{2}\cdot\sqrt{\frac{PQ}{n}}\right] \text{ or } \left[P - 3\sqrt{\frac{PQ}{n}}, P + 3 \cdots \sqrt{\frac{PQ}{n}}\right]$$

i.e, $(0.1 - 0.03, 0.1 + 0.03)$

i.e, $(0.07, 0.13)$.

(b) 2a. In a big city 325 men out of 600 men were found to be smokers. Does this information support the conclusion that the majority of men in this city are smokers?

Sol:- Given $n = 600$

No. of smokers $= 325$

p = sample proportion of smokers $= \frac{325}{600} = 0.5417$

P = population proportion of smokers in the city $= \frac{1}{2} = 0.5$

$Q = 1-P = 1-0.5 = 0.5$.

1. NH H_0: The no. of smokers and non-smoker are equal in the city.

AH: $P > 0.5$ (Right tailed)

The test statistics is $z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0.5417-0.5}{\sqrt{\frac{0.5 \times 0.5}{600}}} = 2.04.$

$\therefore$ calculated value of $z = 2.04$.

Tabulated value of z at 5% level of significance for right tail test is 1.645.

Since calculated value of z > tabulated value of z, we reject the NH and conclude that the majority of men in the city are smokers.

Q2. A manfucterer claimed that atleast 95% of the equipment which the supplied to a factory conformed to specification. An examination of a sample of 200 pieces of equipment revealed that 18 were fault. Test his claim at 5% level of significance.

Sol:- Given sample size, $n = 200$

No. of pieces confirming to specification $= 200 - 18 = 182$.

$\therefore p$ = Proportion of pieces confirming to specifications

$$= \frac{182}{200} = 0.91$$

P = population proportion $= \frac{95}{100} = 0.95$

1) NH H_0: The proportion of pieces confirming to specifications

i.e., $P = 95\%$.

2) AH H_1: $P < 0.95$ (left-tail test)

3) The test statistic is $$z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0.91-0.95}{\sqrt{\frac{0.95 \times 0.05}{200}}} = \frac{-0.04}{0.0154} = -2.59$$

Since alternative hypothesis is left tailed, the tabulated value of z at 5% level of significance is 1.645

Since calculated value of $|z| = 2.6$ is greater than 1.645, we reject the null hypothesis H_0 at 5% level of significance and conclude that the manufactures claim is rejected.

3a) write a procedure for testing Hypothesis?

Sol: procedure for testing hypothesis:

Step 1: Null hypothesis: Define (or) setup a NH H_0 taking into consideration the nature of the problem and data involved.

Step 2: Alternative hypothesis: set up AH, H_1 so that we could decide whether we should use one failed (or) two failed test.

Step 3: Level of significance: Select the appropriate level of sig (α) depending on reliability of estimates and permissible risk. That is suitable α is selected in advance. R is not given in the problem (usually we choose 5% level of sig)

Step 4: Test statistic: compute the test statistic $z = \frac{t - E(t)}{S.E \text{ of } t}$ under the NH.

* Hence t is sample statistic & S.E is standard error of t.

Step 5: Conclusion: we compare the computed value of the test statistic z with the critical value z_α at given level of sig α

If $|z| < z_\alpha$ (i.e absolute value of computed value of z_α the critical vaule z_α at given level of sig (α))

then we conclude that it is not significant. Hence we accept NH if $|z| > z_\alpha$ then the difference is significant & hence the NH is rejected at level (sig α)

b) The mean life of a sample of 10 electric light bulbs was found be 1456 hours with standard deviation of 423 hours. A second sample of 17 bulbs chosen from a different batch showed a mean life of 1280 hours with standard deviation of 398 hours. Is there significant difference between the means of two batches?

Sol: Given that

$n_1 =$ sample size of first batch $= 10$

$n_2 =$ sample size of second batch $= 17$

$\bar{x}_1 =$ mean life of first batch $= 1456$

$\bar{x}_2 =$ mean life of second batch $= 1280$

$\sigma_1 =$ standard deviation of first batch $= 423$

$\sigma_2 =$ standard deviation of second batch $= 398$

1. NH $H_0: \mu_1 = \mu_2$

2. AH $H_1: \mu_1 \neq \mu_2$

3. Level of significance, $\alpha = 0.05$

4. The test statistic is $Z = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}} = \dfrac{1456 - 1280}{\sqrt{\dfrac{(423)^2}{10} + \dfrac{(398)^2}{17}}}$

$= \dfrac{176}{\sqrt{17892.9 + 9317.88}} = \dfrac{176}{164.96} = 1.067$

Since $Z < Z_\alpha = 1.96$, we accept the null hypothesis H_0 i.e no, there is no difference, between the mean life of electric bulbs of two batches.

3C). A sample of 400 items is taken from a population whose standard deviation 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.

Sol: Given $n = 400$, $\bar{x} = 40$, $\mu = 38$ and $\sigma = 10$

1. NH $H_0: \mu = 38$

2. AH $H_1: \mu \neq 38$

3. level of significance, $\alpha = 0.05$

4. The test statistic is, $Z = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{40 - 38}{10/\sqrt{400}} = 4.892 = 4 > 1.96$

$\therefore$ We reject the NH H_0.

i.e the sample is not from the population whose mean is 38.

95% confidence interval is $\left[\bar{x} - 1.96\frac{\sigma}{\sqrt{n}}, \bar{x} + 1.96\frac{\sigma}{\sqrt{n}}\right]$

i.e $\left[40 - \frac{1.96(10)}{\sqrt{400}}, 40 + \frac{1.96(10)}{\sqrt{400}}\right]$ (or) $\left[40 - \frac{1.96}{20}, 40 + \frac{1.96}{20}\right]$

(or) $(40 - 0.98, 40 + 0.98)$ i.e, $(39.02, 40.98)$

3d) 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 variable of 16 minutes Test the claim at 0.05 level significance.

Sol: Given $n = 36$, $\bar{x} = 11$, $\mu = 10$ and $\sigma = \sqrt{16} = 4$

1. NH $H_0 = \mu = 10$
2. AH $H_1 : \mu < 10$
3. level of significance, $\alpha = 0.05$
4. The test statistic is, $Z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{11 - 10}{4/\sqrt{36}} = \frac{6}{4} = 1.5$

Tabulated value of z at 5% level of significance is 1.645. Hence calculated $z <$ tabulated z

$\therefore$ we accept the null hypothesis H_0.

i.e The sample has been drawn from the population with mean $\mu = 3.25$ 95%. Confidence limits are given by

$$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} = 3.4 \pm 1.96 \times \frac{2.61}{\sqrt{900}} = 3.4 \pm 0.1705 \quad 10$$

i.e. 3.57 and 3.2295

ii, Given $n = 1000$

p = sample proportion of rice eaters $= \frac{540}{1000}$

$= 0.54$

P = Population. proportion of rice eaters $= \frac{1}{2} = 0.5$

$\therefore Q = 0.5$

~~Null hypothesis~~

Continuation in next page

P = population proportion of rice eaters $= \frac{1}{2} = 0.5$

(11)

$\therefore Q = 0.5$

NH : Ho : Both rice and wheat are equally popular in the state

AH H1 : $P \neq 0.5$ (two tailed alternative)

Test statistics is $z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.54 - 0.5}{\sqrt{\frac{0.5 \times 0.5}{1000}}} = 2.532$

The calculated value of $z = 2.532$

The tabulated value of z at 1% level of significance ~~to~~ for two to test is 2.58.

Since calculated $z <$ tabulated z. we accept the NH Ho at 1% level of significance and conclude that both rice and wheat. equally popular in the state.

4 b) .i) If two large populations, that are 30% and 25% respectively fair haired people. Is this difference likely to be hidden in same of 1200 and 900 respectively from the two populations.

ii) In a sample of 500 from a village in Rajasthan, 280 are from to be wheat eaters and the rice eaters can we assume of the both articles are equally popular.

1) Given $n_1 = 1200$, $n_2 = 900$

P_1 = proportion of fair haired people in the first population $= \frac{3}{100} = 0.3$

P_2 = proportion of fair haired people in the second population $= \frac{25}{100} = 0.25$

NH H_0: Assume that the sample proportions are equal i.e., the difference in populations is likely to be hidden in sampling

i.e., $H: P_1 = Q_2$

AH $H_1: P_1 \neq P_2$

The test statistic is $Z = \dfrac{P_1 - P_2}{\sqrt{\dfrac{P_1 Q_1}{n_1} + \dfrac{P_2 Q_2}{n_2}}}$

where $Q_1 = 1 - P_1 = 1 - 0.3 = 0.7$

$Q_2 = 1 - P_2 = 1 - 0.25 = 0.75$

$$\therefore Z = \frac{0.3 - 0.25}{\sqrt{\dfrac{0.3 \times 0.7}{1200} + \dfrac{0.25 \times 0.75}{900}}} = \frac{0.05}{0.0195} = 2.55$$

i.e. $Z = 2.5$

Since $Z > 1.96$, therefore we reject the null hypothesis H_0 at 5% level of sig (Two-tailed test) i.e., the sample proportions are not equal. Thus we conclude that the difference in population proportions is unlikely that the real difference will be hidden.

Given $n = 500$

p = sample proportion of rice eaters = $\frac{280}{500}$ = 0.56

P = population proportion of rice eaters = $\frac{1}{2}$ = 0.5

$\therefore Q = 1 - P = 1 - 0.5 = 0.5$

2. AH $H_1: P \neq 0.5$

3. level of significance, $\alpha = 0.01$

4. The test statistic is $z = \dfrac{p-P}{\sqrt{PQ/n}} = \dfrac{0.56-0.5}{\sqrt{(0.5)(0.5)/500}} = \dfrac{0.06}{0.022} = 2.68$

$\therefore$ The calculated value $z = 2.68$

The tabulated value of z at 1% level of significance for two taled test is 2.58, since calculated value of $z >$ tabulated z, we reject the null hypothesis H_0. i.e., Both rice and wheat are not equally popular in Andhra pradesh at 1% level of significa-nce.

(c)
(i) Experience had shown that 20% of a manufactured product is of the top quality. In one day's production of 400 articles only 50 are of top quality. Test the hypothesis at 0.05 level.

ii) 20 people were affected by a disease and only 18 survived will you reject the hypothesis that the survival rate affected by this disease is 85% in favour of the hypothesis that is more at 5% level.

Sol: we have $n = 400$, $x = 50$ and $p = \dfrac{x}{n} = \dfrac{50}{400} = 0.125$
c(i).

1. NH $H_0: P = 0.2$

2. AH $H_1: P \neq 0.2$

3. level of significance, $\alpha = 0.05$

The test statistic is $z = \frac{p-P}{\sqrt{\frac{PQ}{n}}}$

(14)

i.e., $z = \frac{0.125-0.2}{\sqrt{\frac{(0.2)(0.8)}{400}}} = \frac{-0.075}{0.02} = -3.75$

since $|z| = 3.75 > 1.96$, we reject the Null hypothesis H_0 at 5% level of significance.

i.e., $P = 20\%$ is not correct.

4(iii) Given n = sample size = 20

x = no. of survived people = 18

p = proportion of survived people = $\frac{x}{n} = \frac{18}{20} = 0.9$

$P = 0.85$

$\therefore Q = 1-P = 1-0.85 = 0.15$

NH $H_0 : P = 0.85$

AH $H_1 : P > 0.85$ (Right tailed test)

Level of significance, $\alpha = 0.5$

The test statistic is

$$Z = \frac{p-P}{\sqrt{\frac{PQ}{n}}} = \frac{0.9-0.85}{\sqrt{\frac{0.85 \times 0.15}{20}}} = \frac{0.5}{0.8} = 0.625$$

$\therefore$ Calculated $z = 0.625$

Tabulated z at 5% level of significance, $z_\alpha = 1.645$

Since calculated $z <$ tabulated z, we accept the Null hypothesis H_0

i.e., The proportion of the survived people is 0.85.

(3e) In a random sample of 1000 persons from town A. 400 are found to be consumers of wheat. In a sample of 800 from town B, 400 are found to be consumers of wheat. Do these data reveal a significant difference between town A and town B, so far as the proportion of wheat consumers is concerned.

Sol:- we have n_1 = Sample size of town A = 1000.

n_2 = " " B = 800.

x_1 = No. of consumers of wheat from town A = 400.

x_2 = " " " B = 400.

P_1 = proportion of consumers of wheat in town A

$= \frac{x_1}{n_1} = \frac{400}{1000} = 0.4$

P_2 = " " " B $= \frac{x_2}{n_2} = \frac{400}{800} = 0.5$

$$\therefore P = \frac{x_1+x_2}{n_1+n_2} = \frac{400+400}{1000+800} = \frac{800}{1800} = \frac{8}{18} = \frac{4}{9}.$$

$$q = 1-P = \frac{5}{9}.$$

① Null hypothesis ⇒ $H_0: P_1 = P_2$ i.e there is no difference.

② Alternative hypothesis ⇒ $H_1: P_1 \neq P_2$ i.e there is a difference

③ LOS: $\alpha = 0.05$ ④ $$Z = \frac{P_1-P_2}{\sqrt{pq\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}} = \frac{0.4-0.5}{\sqrt{\frac{4}{9}\cdot\frac{5}{9}\left(\frac{1}{1000}+\frac{1}{800}\right)}} = -4.242$$

$\therefore |Z| = 4.2424$. also $Z_\alpha = 1.96$

$\therefore$ we reject the null hypothesis H_0 at 5% level of significance.

$\therefore$ There is significant difference between town A and town B as the proportion of wheat consumers