Module 5 ~~IX~~

Testing

1) Testing the mean of a normal population

Case I : σ^2 is known

Suppose we want to test $H_0 : \mu = \mu_0$ against $H_1 : \mu < \mu_1$ or $H_1 : \mu > \mu_1$ or $H_1 . \mu \neq \mu_0$.

For the significance of level α, the best critical region (BCR) are respectively ; said to,

$Z < -Z_\alpha$ or $Z > Z_\alpha$ or $|Z| \geq Z_{\alpha/2}$

The test statistic is

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

Using this sample, calculate the value of Z. If the it lies in the critical region (CR), reject H_0 otherwise accept H_1.

Note : Z_α or $Z_{\alpha/2}$ is obtained from the standard normal table using probability of $P(Z < -Z_\alpha)$ or $P(Z > Z_\alpha)$ or $P(|Z| \geq \alpha)$

Scanned by CamScanner

Case II : σ^2 is unknown, n is large ($n \geq 30$).

Follow the case I by replacing $\frac{\sigma}{\sqrt{n}}$ by $\frac{s}{\sqrt{n}}$

σ by s in Z.

Case III : σ^2 is unknown, n is small (1 to 29) {t-test}

Suppose we want to test $H_0 : M = M_0$ against one of the alternatives H_1 $M < M_0$ or $M \neq M_0$.

The test statistic is,

$$t = \frac{\bar{x} - M_0}{s/\sqrt{n-1}}$$

where, $s = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$

for the significant level α, the best critical region are as respectively, $t < t_\alpha$ or $t > t_\alpha$ or $|t| \geq t_{\alpha/2}$.

where, t_α or $t_{\alpha/2}$ is obtained from the t-table for $n-1$ degree of freedom using $P(t > t_\alpha) = \alpha$ and $t_{\alpha/2}$ is obtained from t-table for n-1 degree of freedom using

$P(t > t_{\alpha/2}) = \alpha$.

Calculate the value of t using these sample data and if it lies in critical region ~~otherwise~~ reject H_0 otherwise accept H_0.

1. A sample of 25 items were taken from a population with σ SD is 10 and sample mean is found to be 65. Can it be regarded as a sample from a normal population with $M = 60$

Ans: Given,

$n = 25$

$\sigma = 10$

$\bar{x} = 65$,

$M_0 = 60$

we have to test $H_0 : M = 60$ against,

$H_1 : M_1 \neq 60$

Let $\alpha = 0.05$

the critical region is $|z| \geq z_{\alpha/2}$ = $0.05/2 = 0.025$

$= 0.025$

test statistic is

$$z = \frac{\bar{x}-\mu}{\sigma/\sqrt{n}}$$

$$= \frac{65-60}{10/\sqrt{25}} = \frac{65-60}{10} \times \sqrt{25}$$

$$= \frac{5}{10/5} \quad 6 \cdot \frac{5}{10} \times 5$$

$$= 5/2 = \cancel{2.5} \quad 25/10$$

$$= 2.5$$

From standard normal table,

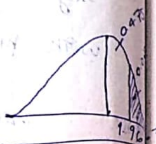
$z_{\alpha/2} = 1.96$.

that is $\alpha/2 = 0.025$

then $= 0.5 - 0.025$

$= 0.475$

$\therefore z_{\alpha/2} = 1.96$



Ans:- follow Case II

$$\boxed{P(z > z_{\alpha/2}) = \alpha/2 = 0.05/2 = 0.025}$$

$|z| = |2.5| = 2.5$

$z_{\alpha/2} = 1.96$

$\therefore 2.5 > 1.96$

$\boxed{z > z_{\alpha/2}}$

z lies in critical region, Ho is rejected.

2. The mean life time of a sample of 100 fluorescent bulb light bulbs produced by a company is computed to be 1570 hours, with a SD of 120 hours. Test the hypothesis that the mean life of all the bulbs manufactured by the company is 1600 hours against the alternative that it is not, using a level of significance of 0.05.

Given,

$n = 100$

$SD = 120$ hrs $\quad S = 120$ hrs

$\bar{x} = 1570$

we have to test $H_0 : \mu_0 = 1600$ hrs against

$H_1 : \mu_1 \neq 1600$ hrs

Let $\alpha = 0.05$

the critical region is $|z| \geq z_{\alpha/2} = 0.05/2$

$= 0.025$

(60)

test statistic is,

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

$$= \frac{1570-1600}{120/\sqrt{100}}$$

$$= \frac{-300}{120} \times 10$$

$$= \frac{-3000}{120} = -2.5$$

from standard normal table,

$P(z > z_{\alpha/2}) = 0.025$

$P(0 < z < z_{\alpha/2}) = 0.5 - 0.025$

$= 0.475$

$z_{\alpha/2} = 1.96$

$|z| = |-2.5| = 2.5$

$z_{\alpha/2} = 1.96$

$z > z_{\alpha/2}$

$2.5 > 1.96$

z lies in the critical region. Then Ho is rejected.

The mean life time of all the bulbs manufactured by the company is not 1600.

3. A sample of 900 members is found to have a mean of 3.4 cm and SD 2.61. can it be regarded as a sample from a large population whose mean is 3.25 cm. Use the two tailed test & $\alpha = 0.01$

Given,

$n = 900$

$\bar{x} = 3.4$

$s = 2.61$

We have to test $H_0: \mu_0 = 3.25$ against

$H_1: \mu_1 \neq 3.25$

Given $\alpha = 0.01$

The critical region is $|z| = z_{\alpha/2} = \frac{0.01}{2}$

$= 0.005$

Test statistic is,

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

(61)

$$= \frac{3.4 - 3.25}{2.61/\sqrt{900}}$$

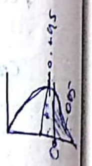
$$= \frac{0.15}{2.61} \times 30$$

$\alpha = \frac{.01}{2} =$

$$\doteq \frac{4.5}{2.61}$$

$$= 1.72$$

from the standard normal table,



$P(z > z_{\alpha/2}) = 0.005$

$P(0 < z < z_{\alpha/2}) = \cancel{0.005 - 0.5}$

$= 0.5 - 0.005$

$= 0.495$

$z_{\alpha/2} = \cancel{2.35}\ 2.57$

$|z| = |1.72| = 1.72$

$z > z_{\alpha/2}$

$\cancel{z > z_{\alpha}}\ 1.72 > \cancel{2.35}\ 2.57$

$1.72 < 2.57$

z does not lies in the critical region. Then H_1 is accepted

4\. A sample of 10 observation means a mean = 38 and SD = 4. Can be conclude that the population mean is 40.

Ans:- Given,

$n = 10$

$\bar{x} = 38$

$s = 4$

$\mu = 40$

We have to test $H_0 : \mu_0 = 40$ against

$H_1 : \mu_1 \neq 40$

Let $\alpha = 0.05$

Therefore, critical region is, $|t| \geq t_{\alpha/2}$

The test static is,

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n-1}}$$

$$= \frac{38 - 40}{4/\sqrt{10-1}}$$

$$= \frac{-2}{4} \times 3$$

$$= -6/4 = \frac{-3}{2} = -1.5$$

$$|t| = |-1.5| = 1.5$$

From the 't' table, $t_{\alpha/2} = \cancel{2.262}\ 2.262$

for $\alpha = 0.05$ & decrease of freedom (df) = 9. $(n-1=10-1)$

$t > t_{\alpha/2}$

$1.5 < 2.262$

t does not lies in critical region. Then H_0 is accepted.

② **Testing equality of two population means** (normal populations)

Case I : σ_1^2 σ_2^2 known. (z-test normal)

suppose we want to test the null hypothesis, $H_0 : M_1 - M_2 = 0$ ie,

$H_0 : M_1 = M_2$ against one of the alternative $H_1 : M_1 - M_2 < 0$ or $H_1 : M_1 - M_2 > 0$

or $H_1 : M_1 - M_2 \neq 0$.

Based on independent random samples of size n_1 and n_2 from two populations, having the means M_1 and M_2.

For the significance level α, the critical regions are respectively $Z < -Z_\alpha$ or $Z > Z_\alpha$ or $|Z| \geq Z_{\alpha/2}$

The test statistic is

$$Z = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad (\because M_1 - M_2 = 0)$$

Calculate the value Z using the sample information and if it lies in the CR reject H_0, otherwise accept it.

Case II : σ_1^2, σ_2^2 Unknown,

n_1, n_2 large

Here $H_0 : M_1 - M_2 = 0$

CR : $Z < -Z_\alpha$ or $Z > Z_\alpha$ or $|Z| \geq Z_{\alpha/2}$

against $H_1 : M_1 - M_2 < 0$ or $H_1 : M_1 - M_2 > 0$ or

$H_1 : M_1 - M_2 \neq 0$

63

test statistic is,

$$Z_1 = \frac{(\bar{x}_1 - \bar{x}_2) - 0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Case III : σ_1^2, σ_2^2 unknown, n_1, n_2 small

$\sigma_1^2 = \sigma_2^2$

Here we are testing

$H_0 : \mu_1 - \mu_2 = 0$ against one of the

alternatives $H_1 : \mu_1 - \mu_2 < 0$ or $H_1 : \mu_1 - \mu_2 > 0$

or $H_1 : \mu_1 - \mu_2 \neq 0$.

The test statistic is,

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where, s_1^2, s_2^2 - are the sample variances,

The CRs are respectively

$t < -t_\alpha$ or $t > t_\alpha$ or $|t| \geq t_{\alpha/2}$

where, t_α or $t_{\alpha/2}$ are obtained from t-table for $n_1 + n_2 - 2$ df using $P(t > t_\alpha) = \alpha$

or $P(|t| \geq t_{\alpha/2}) = \alpha$.

Calculate the value of t and if it lies in CR reject H_0 otherwise accept H_0.

1. A random sample of size 16 has 53 as mean and the sum of squares of the deviation taken from the mean is 150. Can the sample be regarded as arisen from the population with mean 56.

2- A sample of size 8 from a normal population 6, 8, 11, 5, 9, 11, 10, 12. Can such a sample be regarded as drawn from a population with mean 7 at 2% level of significance

3- The mean life of a sample of 10 electric bulbs was observed to be 1309 hours with SD of 420 hrs. A second sample of 16 bulbs of a different batch showed a mean life of 1205 hrs with a SD of 340 hrs. Test whether there is significant difference b/w the means. (use α = 0.0...)

4. Suppose that 64 senior girls from college A and 81 senior girls from college B had mean statures of 68.2" and 67.3" respectively. If the SD for statures of all senior girls is 2.43, is the difference b/w the two groups significant?

5. A random sample of 1000 workers from factory A shows that the mean wages were Rs 47 per week with a SD of Rs. 23. A random sample of 1500 workers from factory B gives a mean wage of Rs. 49 per week with a SD of Rs. 30. Is there any significant difference b/w their mean level of wages.

Answers

1. Given, $n = 16$

$\bar{x} = 53$

~~SD = 150~~

$M = 56$

We have to test $H_0 : M_0 = 56$ against $H_1 : M_1 \neq 56$.

Let $\alpha = 0.05$

Therefore CR is $|t| \geq t_{\alpha/2}$

from the ~~table~~ 't' table, $t_{\alpha/2} = 2.131$ for $\alpha = 0.05$ & $df = 15$.

~~The test statistic~~,

$$t = \frac{\bar{x} - M_0}{s/\sqrt{n-1}}$$

$$= \frac{53 - 56}{150/\sqrt{15}}$$

$$= \frac{-3}{150} \times 3.87$$

$$= \frac{-11.619}{150}$$

$\cancel{t = -0.077}$

$|t| = \cancel{|-0.077| = 0.077}$

$t > t_{\alpha/2}$

$\cancel{0.077 < 2.131}$

~~t does not lies in CR. Then H0 is accepted~~

Here given,

$$\sum_{i=1}^{n} (x - \bar{x})^2 = 150$$

$$\frac{\sum (x - \bar{x})^2}{n} = s^2$$

$$\frac{150}{16} = 9.375$$

$$s = \sqrt{9.375} = 3.0619$$

The test statistic is,

$$t = \frac{\bar{x} - M_0}{s/\sqrt{n-1}}$$

$$= \frac{53 - 56}{3.06/\sqrt{15}}$$

$$= \frac{-3 \times 3.87}{3.06}$$

$$= \frac{-11.61}{3.06}$$

$$= -3.79$$

$|t| = |-3.79| = 3.79$

$t > t_{\alpha/2}$

$3.79 > 2.131$

t lies in CR. The H_0 is rejected.

2- Here, Given, $n = 8$

$\alpha = 0.02$

$M = 7$

x	$x - \bar{x}$	$(x - \bar{x})^2$
6	-3	9
8	-1	1
11	2	4
5	-4	16
9	0	0
10	2	4
10	1	1
12	3	9

$\bar{x} = \frac{\Sigma x}{n}$

$= \frac{62}{8} = 9$

$S^2 = \sqrt{\frac{\Sigma(x-\bar{x})^2}{n}} = \sqrt{\frac{44}{8}}$

$= \sqrt{5.5} = 2.345$

$S = 2.345$

We have to test $H_0 : M_0 = 7$ against

$H_1 : M_1 \neq 7$

From the 't' table,

$t_{\alpha/2} = 2.998$

The test statistic is,

$t = \frac{\bar{x} - M_0}{S/\sqrt{n-1}}$

$= \frac{9-7}{2.34} \times \sqrt{7}$

$= \frac{2 \times 2.645}{2.34}$

$= \frac{5.29}{2.34}$

$= 2.261$

$t > t_{\alpha/2}$

$2.26 < 2.998$

t does not lies in CR. H_0 is accepted.

3. Here given,

$t = 0.697$

$\alpha = 0.05$

2.064

$n_1 = 10$

$\bar{x}_1 = 1309$

$S_1 = 420$

$n_2 = 16$

$\bar{x}_2 = 1205$

$S_2 = 390$

We have to test $H_0 : M_1 = M_2 = 0$ against

$H_1 : M_1 - M_2 \neq 0$.

The test statistic is,

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$s_1^2 = 420^2 = 176400$

$s_2^2 = 390^2 = 152100$

$$= \frac{1309 - 1205}{\sqrt{\frac{(10 \times 176400) + (16 \times 152100)}{10 + 16 - 2}\left(\frac{1}{10} + \frac{1}{16}\right)}}$$

$$= \frac{104}{\sqrt{\frac{1764000 + 2433600}{24}\left(\frac{16+10}{160}\right)}}$$

$$= \frac{104}{\sqrt{\frac{4197600}{24}(0.1625)}}$$

$$= \frac{104}{\sqrt{174900 \times 0.1625}}$$

$$= \frac{104}{\sqrt{28421.25}} = \frac{104}{168.59}$$

$t = 0.617$

Let $\alpha = 0.05$

~~$t = 0.617$~~

Therefore CR is $|t| \geq t_{\alpha/2}$

From the 't' table,

$t_{\alpha/2} = 2.064$

for, $\alpha = 0.05$ and $df = 10 + 16 - 2 = 24$

~~0.617~~

$\therefore |t| \geq t_{\alpha/2}$

$0.617 < 2.064$.

t does not lies in CR. H_0 is accepted.

Testing of equality (difference) of two population proportion.

Here we are testing the eqlty of 2 population proposition or the significant difference b/w 2 sample propositions.

Let P_1: Proportion of success of the 1st popul

P_2: Proportion " " 2nd population

x_1 : no. of success in the 1st sample

x_2 : no. of success in the 2^{nd} sample

n_1 : 1^{st} sample size

n_2 : second sample size

p_1' = Proportion of success of the 1st sample $\left(=\frac{x_1}{n_1}\right)$

p_2' = Proportion of success of the 2^{nd} sample $\left(\frac{x_2}{n_2}\right)$

Suppose we want to test the null hypothesis

$H_0 : P_1 - P_2 = 0$ against one of the

alternatives ; $H_1 : P_1 - P_2 < 0$ or

$H_1 : P_1 - P_2 > 0$ or

$H_1 : P_1 - P_2 \neq 0$

based on the 2 independent large sample of size n_1 and n_2 with proportion of success P_1' and P_2' respectively.

For the significance level α, the CRs are respectively $Z < -Z_\alpha$, $Z > Z_\alpha$ or $|Z| > Z_{\alpha/2}$

The test statistic is,

$$\boxed{Z = \frac{P_1' - P_2'}{\sqrt{p^* q^* \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}}$$

where,

$$p^* = \frac{n_1 P_1' + n_2 P_2'}{n_1 + n_2}$$

$$q^* = 1 - p^*$$

Calculate Z and if it lies in the critical region reject it, otherwise accept it.

6. In a ~~saving~~ survey 70 business firms (64) H.W as found that 45 are planning to expand their capacities next years. Does the sample information contradict the hypothesis that 70% of the firms are planning to expand next years.

7. In a die throwing experiment, the throw of 3 or 6 is reckoned as a success. Suppose 9000 times three the die was thrown resulting 3240 success. Do you have reasons to believe that the dies is an unbiased one.

8. Before an increase in excise duty on tea 800 persons out of a sample 1000 persons were found to be tea drinkers. After an increase in duty 800 people were tea drinkers in a sample of 1200 people. Test whether there is significant decrease in the consumption of tea after the increase in duty.

Answers

4. equality of 2 population means

Given, $n_1 = 64$

$\bar{x}_1 = 68.2$

$\sigma_1^2 = 243$

$n_2 = 81$

$\bar{x}_2 = 67.3$

$\sigma_2^2 = 243$

We want to test $H_0 = M_1 = M_2 = 0$ against

$H_1: M_1 - M_2 \neq 0$

Let $\alpha = 0.05$

the test statistic is,

$$Z = \frac{(\bar{x}_1 - \bar{x}_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_1^2}{n_2}}}$$

$$= \frac{68.2 - 67.3}{\sqrt{\frac{(2.43)^2}{64} + \frac{(2.43)^2}{81}}}$$

$$= \frac{0.9}{\sqrt{\cancel{3.796 + 3}}}$$

$$= \frac{0.9}{\sqrt{\cancel{6.797}}} \quad \frac{0.9}{\sqrt{\frac{5904.9}{64} + \frac{5904.9}{81}}}$$

$$= \frac{0.9}{\sqrt{0.0922.64 + \cancel{0.0729}}}$$

$$= \frac{0.9}{\sqrt{2165.1.\cancel{64}}} = \frac{0.9}{\cancel{40.64}} \quad \frac{0.9}{\cancel{5.4\times5}\ 0.4063}$$

$$= \cancel{0.022}\ 2.215$$

The critical region is

$$|Z| = Z_{\alpha/2} = \frac{0.05}{2} = 0.25$$

$$Z_{\alpha/2} = 1.96$$

$$Z > Z_{\alpha/2}$$

$$\cancel{0.022 < 1.96}$$

$$2.215 > 1.96$$

Z is lies in CR. Ho is rejected.

5. Here,

Given, $\bar{x}_1 = 47$

$n_1 = 1000$

$s_1 = 23$

$n_2 = 1500$

$\bar{x}_1 = 49$

$s_2 = 30$

We have to test, $H_0 : M_1 = M_2 = 0$ against

$H_0 : m_1 \neq m_2$.

Let $d = 0.02$

The test statistic is,

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Scanned by CamScanner

$$= \frac{47-49}{\sqrt{\frac{23^2}{1000}+\frac{30^2}{1500}}}$$

$$= \frac{-2}{\sqrt{\frac{529}{1000}+\frac{9000}{1500}}}$$

$$= \frac{-2}{\sqrt{0.529+.6}}$$

$$= \frac{-2}{\sqrt{\cancel{6.592}}} = \frac{\cancel{-2}}{}$$

$$= \frac{-2}{\sqrt{1.129}} = \frac{-2}{1.0625}$$

$$= \underline{-1.882}$$

$$|z| = |-1.882| = \underline{1.882}$$

$$z_{\alpha/2} = \frac{0.02}{2} = 0.01$$

$$P(z > z_{\alpha/2}) = 0.01$$

$$P(0 < z < z_{\alpha/2}) = 0.5 - 0.01$$

$$= 0.49$$

$$= \underline{2.33}$$

$$z > z_{\alpha/2}$$

$$1.882 < 2.33$$

z donot lies in CR. Ho is accepted.

α	$\alpha/2$
0.01	2.58
0.02	2.33
0.05	1.96

(3) Test of ~~Population~~ Proportion of Success of a Population

By testing population proportion of success we mean the testing of the significant difference b/w population proposition of success and the sample proportion of success.

Let P: Population proportion of success (unknown)

P_0: the assumed value of P (given)

p': $\frac{x}{n}$; the proportion of success of a sample

x: the no. of success

n: sample size

Suppose we want to test the null hypothesis

$H_0: P = P_0$ Vs $H_1: P < P_0$ or $H_1: P > P_0$ or

$H_1 : P \neq P_0$ base on a large sample of size n whose proportion of success p'.

For the significance level α, critical regions are $Z < -Z_\alpha$, $Z > Z_\alpha$ or $|Z| \geq Z_{\alpha/2}$ respectively.

The test statistic is,

$$\boxed{Z = \frac{p' - P_0}{\sqrt{\frac{P_0 q_0}{n}}}}$$

calculate the value of Z and if it lies in the CR, reject H_0, otherwise accept H_0.

(3) Answers

Testing of population of proportion of discuss.

population :-

6. when test is proportion test,

$n = 70$

$x = 45$

$P_0 = \frac{70}{100} = 0.7$

We have to test, $H_0 : P = .70$ against $H_1 : P \neq 0.7$

Let $\alpha = 0.05$

The CR is $|Z| \geq 1.96$

where, $Z = \frac{P' - P_0}{\sqrt{\frac{P_0 q_0}{n}}}$

$p' = \frac{x}{n} = \frac{45}{70} = 0.64$

we have, $P_0 = 0.70$

$q_0 = 1 - 0.7$

$= 0.3$

$$Z = \frac{0.64 - 0.7}{\sqrt{\frac{0.7 \times 0.3}{70}}}$$

$$= \frac{-0.06}{\sqrt{\frac{0.21}{70}}} = \frac{-0.06}{\sqrt{0.003}}$$

$$= \frac{-0.06}{0.0548}$$

$$= -1.0948$$

$|Z| = |-1.0949| = 1.0949$

$z > z_{\alpha/2}$

$1.01049 < 1.96$

z donot lies in the CR. H_0 is accepted.

7. Given, test is proportion test,

$n = 9000$

$P_0 = 2/6 = 1/3$

$x = 3240$

we have to test : $H_0: P = 1/3$ against

$H_1: P \neq 1/3$.

Let $\alpha = 0.05$

The CR. is $|z| \geq z_{\alpha/2} = 1.96$

where $z = \frac{P' - P_0}{\sqrt{\frac{P_0 q_0}{n}}}$

$p' = \frac{x}{n} = \frac{3240}{9000} = 0.36$

$P_0 = 1/3$

$q_0 = 1 - 1/3 = \frac{3-1}{3} = 2/3$

$\therefore z = \frac{0.36 - 0.33}{\sqrt{\frac{0.33 \times 0.66}{9000}}}$

$= \frac{0.03}{\sqrt{\frac{0.2178}{9000}}} = \frac{0.03}{\sqrt{0.0000242}}$

$= \frac{0.03}{0.0049} =$

$= 6.12$

$|z| = |6.12| = 6.12$

$z > z_{\alpha/2}$

$6.12 > 1.96$

z lies in the CR H_0 is rejected.

Answers

8. Given, test is equality proportion test,

$n_1 = 1000$

$x_1 = 800$

$n_2 = 1200$

P_2

$x_2 = 800$

we have, $p_1' = \frac{x_1}{n_1}$

$= \frac{800}{1000} = 0.8$

$p_2' = \frac{800}{1200} = 0.47$

we have to test,

$H_0 : P_1 - P_2 = 0$ vs $P_1 - P > 0$

Let $\alpha = 0.05$, $0.5 - 0.05 = 0.45$

The CR is $z > z_\alpha$ ~~$|z| \geq z_{\alpha/2} = 1.96$~~

$z_\alpha = 1.65$

The test stastic is,

$$z = \frac{p_1' - p_2'}{\sqrt{p^* q^* \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where, $p^* = \frac{n_1 p_1' + n_2 p_2'}{n_1 + n_2}$ | $\frac{x_1 + x_2}{n_1 + n_2}$

$$= \frac{1000 \times 0.8 + 1200 \times 0.667}{1000 + 1200}$$

$$= \frac{800 + 804}{2200} \qquad \frac{1604}{2200}$$

$= 0.72$

$q^* = 1 - p^*$

$= 1 - 0.72 = 0.28$

$$z = \frac{0.8 - 0.67}{\sqrt{0.72 \times 0.28 \left(\frac{1}{1000} + \frac{1}{1200}\right)}}$$

$$= \frac{0.13}{\sqrt{0.2016 \left(0.001 + 0.00084\right)}}$$

$$= \frac{0.13}{\sqrt{0.2016 \times 0.00184}}$$

$$= \frac{0.13}{\sqrt{0.00037}} = \frac{0.13}{0.0192}$$

$= 6.7708$

$|z| = z > z_\alpha$

$= 6.77 > 1.65$

H_0 is rejected. $\therefore P_1 \geq P_2$

there is significant decrease in consumption of tea after raise in excised

9. Given, $n_1 = 600$

$x_1 = 450$

$n_2 = 900$

$x_2 = 450$

We have, $p_1' = \frac{x_1}{n_1} = \frac{450}{600} = 0.75$

$p_2' = \frac{450}{900} = 0.5$

We have to test,

$H_0 : P_1 - P_2 = 0$ vs $P_1 - P_2 \neq 0$

Let $\alpha = 0.05$ $\quad 0.5 - 0.05 = 0.495$

The CR is $= z \geq z_{\alpha/2} = 1.96$

The test statistic is,

$$z = \frac{p_1' - p_2'}{\sqrt{p^* q^* \left(1/n_1 + 1/n_2\right)}}$$

$$p^* = \frac{x_1 + x_2}{n_1 + n_2} = \cancel{600+900}$$

$$= \frac{450+450}{600+900}$$

$$= \frac{900}{1500} = 0.6$$

$$q^* = 1 - 0.6 = 0.4$$

$$z = \frac{0.75 - 0.5}{\sqrt{0.6 \times 0.4 \left(1/600 + 1/400\right)}}$$

$$= \frac{0.25}{\sqrt{0.24\left(\cancel{0.01+0.0067+1.12}\ 0.00167 + 0.0012\right)}}$$

$$= \frac{0.25}{\sqrt{0.24 \times \cancel{2.71}\ 0.00287}} = \frac{0.25}{\sqrt{\cancel{0.06696}}}$$

$$= \frac{0.25}{\cancel{0.259}} \quad \frac{0.25}{\sqrt{0.00068}}$$

$$= \cancel{0.965}$$

$$= \frac{0.25}{0.0262}$$

$$= 9.54$$

$\therefore |z| = \cancel{9.5}\ z > z_{\alpha/2}$

$= 9.54 > 1.96$

Therefore H_0 is rejected.

6) Paired t-test

Consider a data obtained by observing the values of some attribute of different elements of the population, observed 'before & after'. For instance, measure a certain reaction of human subjects before (x_i) and after some treatment. has an effect or not.

Here we have to two samples, $x_1, x_2 \ldots x_n$ (values "before") and $y_1, y_2 \ldots y_n$ (values "after"). But they are dependent samples.

Under some assumptions we use one sample test we imagine that the observations are such that the difference $d_i = x_i - y_i$; $i = 1, 2, 3 \ldots n$ have the same normal distribution with mean μ & variance of σ^2

Note: The values x_i or y_i, separately donot need to have normal distribution.

We may wish to test the null hypothesis ~~H_0~~: $H_0: \mu = \mu_0$ v/s one or two sided alternative $H_1: \mu > \mu_0$ or $H_1: \mu \neq \mu_0$.

In most typical applications we set up hypothesis as $H_0: \mu = 0$ Vs $H_1: \mu \neq 0$

(ie we assume that the treatment has no effect against the alternative that the ~~parameter~~ treatment has effect).

The test statistic is,

$$\boxed{t = \frac{\bar{d}}{S_d / \sqrt{n-1}}}$$

where, $\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$,

$S_d = \sqrt{\frac{1}{n}\sum d_i^2 - \bar{d}^2}$

$d_i = x_i - y_i$, $i = 1, 2, \ldots n$

For the significance level α, the BCRs is $|t| > t_{\alpha/2}$. Here $t_{\alpha/2}$ is obtained by referring the t-table for $n-1$ degrees of freedom and α. Calculate the value of the test statistic if it lies in the CR, reject H_0. otherwise accept it.

Ex 1. Ten soldiers visit a rifle range for two consecutive weeks. For the 1st week, their scores are 67, 24, 57, 55, 63, 54, 56, 68, 33, 43 and during the second week, they score in the same order 70, 38, 58, 58 (A) 56, 56, 67, 68 (71), 72, 43, 38. Examine if there is significant difference in their programme?

Ans:

x_i	67	24	57	55	63	54	56	68	33	43	Total
y_i	70	38	58	58	56	67	68	72	42	38	
d_i (difference)	-3	-14	-1	-3	7	-13	-12	-4	-9	5	-47
d_i^2	9	196	1	9	49	169	144	16	81	25	699

$$\bar{d} = \frac{\Sigma d_i}{n} = \frac{-47}{10} = -4.7$$

$$S_d = \sqrt{\frac{\Sigma d_i^2}{n} - \bar{d}^2}$$

$$= \sqrt{\frac{699}{10} - (-4.7^2)}$$

$$= \sqrt{69.9 - 22.09} = \sqrt{47.81}$$

$$= 6.91$$

H_0 : There is no difference, $M = 0$

H_1 : There are different (performance) $M \neq 0$

The test statistic is,

$$t = \frac{\bar{d}}{S_d / \sqrt{n-1}}$$

$$= \frac{-4.7}{6.91 / \sqrt{9}} = \frac{-4.7}{20.7}$$

$$= -2.04$$

Let $\alpha = 0.05$

$df = n-1 = 10-1 = 9$

Table value (t-table) = 2.262

$|t| = |-2.04| = 2.04$.

$t > t_{\alpha/2}$

$2.04 < 2.262$

H_0 is accepted. That is, there is no difference in their performance.

2. A ~~criterian~~ test stimulus administered to each of the 12 patients resulted in the following increase of blood pressure. 5, 2, 8, -1, 3, 0, -2, 1, 5, 0, 4, 6. Can it be concluded that the stimulus will, in general be accompanied by an increase in blood pressure?

Ans:

d_i	5	2	8	-1	3	0	-2	1	5	0	4	6
d_i^2	25	4	164	1	9	0	4	1	25	0	16	36

$\Sigma d_i = 31$

$\Sigma d_i^2 = 185$

$\bar{d} = \frac{\Sigma d_i}{n} = \frac{31}{12} = 2.58$

$S_d = \sqrt{\frac{\Sigma d_i^2}{n} - \bar{d}^2}$

$= \sqrt{\frac{185}{12} - (2.58)^2}$

$= 2.96$

H_0 : There is no difference (no increase)

H_1 : There is increase in blood pressure

$df = n-1 = 12-1 = 11$

$t = \frac{\bar{d}}{S_d / \sqrt{n-1}} = \frac{2.58}{2.96 / \sqrt{11}} = 2.9$

$t_{\alpha, 12} = 1.796$

$t = 2.9 > 1.796$

H_0 is rejected.

There is increase in blood pressure

ANOVA

1. As head of a department of a consumer's research organisation you have responsibility for testing and comparing life times of light bulbs for four brands of bulbs. Suppose you test the lifetime of three bulbs of each of the four brands. Your test data are as shown below, each entry representing the lifetime of a bulb, measured in hundred of hours.

Brand

A	B	C	D
20	25	24	23
19	23	20	20
21	21	22	20

Can we infer that the mean lifetime of the four brands are equal.

Ans: The null hypothesis is that the avg lifetime of the four brands of the bulbs are equal.

ie, $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$

Scanned by CamScanner

Let, $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_3$, $\bar{x}_4$ denote mean life time of brand A, B, C and D respectively and $\bar{\bar{x}}$ be the overall grand mean.

Then,

	x_1	x_2	x_3	x_4
	20	25	24	23
	19	23	20	20
	21	21	22	20
Total	60	69	66	63

$\bar{x}_1 = 20$, $\frac{\Sigma x_1}{n} = \frac{60}{3} = 20$

$\bar{x}_2 = 23$, $\bar{x}_3 = 22$, $\bar{x}_4 = 21$

Grand mean $= \bar{\bar{x}} = \frac{\bar{x}_1 + \bar{x}_2 + \bar{x}_3 + \bar{x}_4}{n}$

$= \frac{20 + 23 + 22 + 21}{4}$

$= \frac{86}{4} = 21.50$

The variance b/w samples can be computed as follows,

$\bar{x}$	$\bar{\bar{x}}$	$\bar{x} - \bar{\bar{x}}$	$(\bar{x} - \bar{\bar{x}})^2$
20	21.50	-1.5	2.25
23	21.50	1.5	2.25
22	21.50	0.5	0.25
21	21.50	-0.5	0.25
		total	5.00

$\Sigma(\bar{x} - \bar{\bar{x}})^2 = 5.00$

$s_{\bar{x}}^2 = \frac{\Sigma(\bar{x} - \bar{\bar{x}})^2}{n-1} = \frac{5.00}{3} = 1.667$

n = no. of samples

m = sample size

$\sigma^2 = m \times s_{\bar{x}}^2 = 3.16667$

$= 5$

∴ the first estimate of the population variance based on the variance b/w samples, $= 5$, say, s^2

The variance within samples can be computed as follows:-

A		B		C		D	
x	$(x-\bar{x})^2$	x	$(x-\bar{x})^2$	x	$(x-\bar{x})^2$	x	$(x-\bar{x})^2$
20	0	25	4	24	4	23	4
19	1	23	0	20	4	20	1
21	1	21	4	22	0	20	1

for brand A,

$\bar{x} = 20, \quad S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{2}{3-1} = 2/2 = 1$

for brand B,

$\bar{x} = 23, \quad S_2^2 = \frac{8}{2} = 4$

for brand C,

$\bar{x} = 22, \quad S_3^2 = \frac{8}{2} = 4$

for brand D,

$\bar{x} = 21, \quad S_4^2 = \frac{6}{2} = 3$

$\therefore$ the pooled estimate s^2 is given by,

$$s^2 = \frac{S_1^2 + S_2^2 + S_3^2 + S_4^2}{4}$$

$$= \frac{1+4+4+3}{4} = 12/4 = 3$$

The estimate of the population variance based on 'within sample' = 3 say, S_2^2

$$\boxed{\therefore F = \frac{\text{variance b/w the samples}}{\text{variance within samples}}}$$

$$= \frac{S_1^2}{S_2^2} = \frac{5}{3} = 1.67$$

From the 'F'-table, the table value of F for (3, 8) df at 5% level of significance is 4.07

$\therefore F_{cal} = 1.67 < F\text{-table} = 4.07$

ie, $F_{cal} < F\text{-table}$ we accept H_0.

ie, we can infer that the avg life time of different brands of bulbs are equal.

~~χ^2-test~~

χ^2-test for goodness of fit

If we have a set of frequencies of a distribution obtained by an experiment and if we are interested knowing whether these frequencies are consistent with those which may be obtained based on some theory. Then we can use χ^2-test, for goodness of fit for this purpose.

Steps -

1) H_0: There is goodness of fit b/w observed & expected frequencies.

Scanned by CamScanner

2) Compute the test statistic, $\boxed{\chi^2 = \sum \left\{ \frac{(O_i - E_i)^2}{E_i} \right\}}$

where O_i stands for observed frequencies

and E_i stands for expected frequencies.

Note: Observed frequencies are available in a given problem. But expected frequencies are to be computed.

3) Degree of freedom $= n - r - 1$,

where, 'r' is the no. of independent constraints to be satisfied by the frequencies.

Note: For a frequency distribution 'r' is the no. of parameters computed from the data.

4) Obtain the table value of χ^2 for d.f and l.o.s.

5) If the calculated value of χ^2 is "less than" the 'table value', we conclude that there is goodness for fit (concept).

1. 8 coins were tossed 256 times. The results obtained are given below: Test whether the coins are unbiased.

no. of x : 0, 1, 2, 3, 4, 5, 6, 7, 8

frequency : 2, 10, 25, 50, 35, 86, 21, 9, 6

Ans:

$P = P(\text{getting head in toss}) = 1/2$

$n = 8$, $q = 1 - 1/2 = 1/2$

Here, 'no. of heads' follows binomial distribution.

Therefore,

$$P(x) = \binom{8}{x} (1/2)^x (1/2)^{8-x}$$

$x = 0, 1, 2, 3, 4, 5, 6, 7, 8.$

H_0 : There is goodness of fit b/w Observ f & exp fre.

x	$P(x)$	$E_i = N \times P(x)$
0	$1/256$	1
1	$8c_1(1/2)^1(1/2)^{8-1} = 8/256$	8
2	$8c_2(1/2)^2(1/2)^6 = 28/256$	28
3	$8c_3(1/2)^3(1/2)^5 = 56/256$	56
4	$8c_4(1/2)^4(1/2)^4 = \frac{70}{256}$	70
5	$8c_5(1/2)^5(1/2)^3 = \frac{56}{256}$	56
6	$8c_6(1/2)^6(1/2)^2 = 28/256$	28
7	$8c_7(1/2)^7(1/2)^1 = 8/256$	8
8	$8c_8(1/2)^8(1/2)^0 = 1/256$	1

O_i	E_i	$O_i - E_i$	$(O_i-E_i)^2$	$(O_i-E_i)^2/E_i$
12(2+10)	9(1+8)	3	9	1
25	28	-3	9	$9/28 = 0.321$
50	56	-6	36	$36/56 = 0.643$
75	70	5	25	0.357
58	56	2	4	0.071
21	28	-7	49	1.75
15(9+6)	9(8+1)	6	36	4
				8.14

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= 8.14$$

$df = 7 - 1 = 6$

$\alpha = 0.05$

From the χ^2-table,

level of significance, is, 12.592

$\chi^2 < \alpha$

$8.14 < 12.592$

$\chi^2_{cal} = 8.14 < \chi^2$ table $= 12.592$

H_0 accepted. That there is goodness of fit b/w observed & expected frequency. Coins are ~~un~~ unbiased.

χ^2

Testing independence of 2 attributes

1). H_0 : 2 attributes are independent (they are not associated)

2) Calculate the test statistic by the formula,

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

where, O_i is the observed frequency

E_i is the expected frequency

3) Degree of freedom, $df = (r-1)(c-1)$

4) Obtain the table value for the df & desired level of significance α..

(83)

5) If the calculated value of χ^2 less than the table value of χ^2, accept it. Otherwise reject it.

1. From the following data we use χ^2-test and conclude whether innoculation is effective in preventing tuberculosis.

	Attacked	Not attacked
Inoculated	31	469
Non-innoculated	185	1315

Ans: H_0: the two attributes namely, attack & innoculation are independent ie, is not effective

The observed frequencies (Oi) are: 31, 469, 185, 1315

~~observed~~ Expected frequencies are (Ei):

Observed frequency Table

	Attacked	Non-attacked	Total
innoculated	31	469	500
Non-oculated	185	1315	1500
Total	216	1784	2000

Expected frequency Table

$\frac{500 \times 216}{2000} = 54$	$446 = \frac{500 \times 1784}{2000}$
$\frac{1500 \times 216}{2000} = 162$	$\frac{1500 \times 1784}{2000} = 1338$

Oi	Ei	Oi − Ei	$(O_i - E_i)^2$	$(O_i - E_i)^2 / E_i$
31	54	−23	529	9.796
469	446	23	529	1.186
185	162	23	529	3.265
1315	1338.	−23	529	0. 395

(64)

$$\chi^2 = \sum \frac{(O_i - E_i)^2}{E_i}$$

$$= 14.642$$

$$df = (r-1)(c-1)$$

$$= (2-1)(2-1)$$

$$= 1 \times 1 = 1$$

$\alpha = 0.05$

From the χ^2-table, value, 3.841

$\chi^2_{cal} = 14.64 > \chi^2$-table value $= 3.841$

Ho : is rejected.

~~It is ineffective~~

$\therefore$ Attack & innoculation are not independent

The innoculation is effective.