TUTORIAL - IV

U19CS012

SAMPLING OF VARIABLES

BHAGYA VINOD RANA

U19CS012

1.) A machine which produces mica insulating washers for use in electric device to turn out washers having a thickness of 10 mm. A sample of 10 washers has an average thinkness of 9.52 mm with a standard deviation of 0.6 mm. Find out t.

1.) μ_0 = Population mean = 10 mm $s = 0.6$ mm

$\bar{x}$ = Sample mean = 9.52 mm $n = 10$

$$t = \frac{\bar{x} - \mu_0}{(s/\sqrt{n})} = \frac{(9.52-10)\sqrt{10}}{(0.6)} = -0.8 \times \sqrt{10} = \boxed{-2.5298}$$

2.) Ten individuals are chosen at random from a population and their heights are found to be in inches 63, 63, 64, 65, 66, 69, 70, 70, 71. Discuss the suggestion that the mean heights of universe is '65'.
[For 9 degree of freedom, t at 5% level of significance = 2.262]

2.) N=9 (ie <30, so it is small sample)

STEP 1: Null Hypothesis (H0) : $\mu_0 = 65$

Alternate Hypothesis (H_0) : $\mu \neq 65$ (Two tailed test)

STEP 2: LOS = 5% (two tailed test)

Degree of Freedom = n-1 = 9-1 = 8

Critical Value (t_α) = 2.262

$$\bar{X} = \Sigma x/n = 601/9 = \boxed{66.77}$$

$$S = \sqrt{\frac{\Sigma(x-\bar{x})^2}{(n-1)}} = \boxed{3.2821}$$

$$t_{(o)} = \frac{\bar{x} - \mu}{S/\sqrt{n}} = \frac{(66.77-65)\sqrt{9}}{(3.2821)}$$

$$= \boxed{1.6178}$$

X	$X-\bar{X}$	$(X-\bar{X})^2$
63	-3.77	14.2129
63	-3.77	14.2129
64	-2.77	7.6729
65	-1.77	3.1329
66	-0.77	0.5929
69	2.33	5.4289
70	3.33	11.0889
70	3.33	11.0889
71	4.33	18.7489
$\Sigma = 601$		Σ 26.1801

U19CS012

Since $|t_{(a)}| < t_{x}$, μ_0 is accepted

$\therefore$ Mean height of Universe is 65 inches.

3.) A random sample of size 16 values from a normal population showed a mean of 53 and a sum of squares of deviations from the mean equals to 150. Can this sample be regarded as taken from the population having 56 as mean? Obtain 95% and 99% confidence limits of the mean of the population.

$\gamma = 15$, $\alpha = 0.05$, $t = 2.131$, $\alpha = 0.01$, $t = 2.947$

3.) STEP 1: Setup Null and alternative hypothesis H_0 and H_1,

$H_0: \bar{x} = \mu$

$H_1: \bar{x} \neq \mu$

STEP 2: Let the level of significance be 5% and 1% with $n-1 = 16-1 =$ 15 degree of freedom

$t_{0.05}$ for $(16-1)$ i.e., 15 dof $= 2.131$

$t_{0.0001}$ for $(16-1)$ i.e., 15 dof $= 2.947$

STEP 3: Test statistics $t = \dfrac{(\bar{x} - \mu)}{S/\sqrt{n}}$, where $\bar{x} = \dfrac{1}{n}\sum_{i=1}^{n} x_i$

We are given, $n = 16$, $\bar{x} = 53$, $\mu = 56$, $\sum_{i=1}^{n}(x_i - \bar{x}) = 150$ & $s^2 = \dfrac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2$

$s^2 = ?$

$$s^2 = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i-\bar{x})^2 = \frac{1}{(16-1)} 150 = \boxed{10} \qquad \boxed{S = \sqrt{10}}$$

$$t = \frac{(\bar{x}-\mu)}{S/\sqrt{n}} = \frac{(53-56)}{\sqrt{10}/\sqrt{16}} = \frac{(-12)}{\sqrt{10}} = -3.7947$$

STEP 4: Calculated value of t: -3.7947

Table value of t: 2.131, 2.947 } REJECTED!

Hence, we conclude that the assumption of mean of 56, for the population is Not reasonable.

U19CS012

95% confidence Limit $= \bar{x} \pm t_{0.005}\left(\frac{s}{\sqrt{n}}\right)$

$= 53 \pm 2.131 \frac{\sqrt{10}}{\sqrt{16}}$

$= 53 \pm 1.6847$

$\boxed{\Rightarrow (51.3152,\ 54.6847)}$

99% confidence Limit $= \bar{x} \pm t_{0.01}\left(\frac{s}{\sqrt{n}}\right)$

$= 53 \pm (2.947) \frac{\sqrt{10}}{\sqrt{16}}$

$= 53 \pm 2.3298$

$\boxed{\Rightarrow (50.67,\ 55.3298)}$

4> Two independent samples of 8 and 7 items resp. had the following values of variable (weight in ounces)

Sample 1 : 9 11 13 11 15 9 12 14

Sample 2 : 10 12 10 14 9 8 10

Is The difference between the means of the samples significant?

[Given for $V = 13$, $t_{0.05} = 2.16$]

4.> $n_1 = 8$ and $n_2 = 7$ (< 30, it is small Sample)

STEP1 : Null Hypothesis $(H_0) = (\mu_1 = \mu_2)$ (ie difference between Sample is not significant)

Alternate Hypothesis $(H_1) = \mu_1 > \mu_2$ | $\mu_2 > \mu_1$ (ie difference between sample is significant) [Two tailed Test]

STEP2 : LOS = 5% (Two tailed test)

Degree of Freedom $= n_1 + n_2 - 2 = 2 + 7 - 2 = 13$

$\therefore$ Critical Value $(t_\alpha) = 2.16$

U19CS012

Sample 1	$x_{1t}-\bar{x}_1$	$(x_{1t}-\bar{x}_1)^2$	Sample 2	$x_{2t}-\bar{x}_2$	$(x_{2t}-\bar{x}_2)^2$
9	-2.75	7.5625	10	-0.42	0.1764
11	-0.75	0.5625	12	1.58	2.4964
13	1.25	1.5625	10	-0.42	0.1764
11	-0.75	0.5625	14	3.58	12.8164
15	3.25	10.5625	9	-1.42	2.0164
9	-2.75	7.5625	8	-2.42	5.8564
12	0.25	0.0625	10	-0.42	0.1764
14	2.25	5.0625			

$\Sigma = 34$ $\qquad\qquad \Sigma = 74$

$\bar{x}_1 = \dfrac{\Sigma x_{1t}}{n_1} = \dfrac{94}{8} = \boxed{11.75}$ $\qquad \bar{x}_2 = \dfrac{\Sigma x_{2t}}{n_2} = \dfrac{73}{7} = \boxed{10.42}$

$$Sp = \sqrt{\frac{\Sigma(x_{1t}-\bar{x}_1)^2 + \Sigma(x_{2t}-\bar{x}_2)^2}{n_1+n_2-2}}$$

$$= \sqrt{\frac{33.5 + 23.7148}{13}} = \sqrt{\frac{57.2148}{13}} = \boxed{2.0978}$$

$$S.E = SP\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 2.0978\sqrt{\frac{1}{8}+\frac{1}{7}}$$

$$= 2.0978\sqrt{\frac{15}{56}}$$

$$= \boxed{1.0857}$$

STEP4: Test Statistics

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{S.E.} = \frac{11.75-10.42}{1.0857} = \frac{1.33}{1.0857} = \boxed{1.2250}$$

STEP 5:

Since $|t_{cal}| < t_\alpha$, Ho is accepted.

ANS: Difference between sample mean is not significant.

VISION

1119CS012

5.) Memory capacity of 9 students was tested before and after the course of meditation for a month. State ~~wet~~ whether course was effective or not from the data below (in some units)

Before	10	15	9	3	7	12	16	17	4
After	12	17	8	5	6	11	18	20	3

[t_{tab} at 0.05 at 8 d.o.f is 2.31]

5.) $n_1 = 9$ and $n_2 = 9$ (< 30, small sample)

STEP 1:

Null Hypothesis (H_0) = $\mu_1 = \mu_2$ (difference insignificant) (course is not effective)

Alternate Hypothesis (H_0) = $\mu_1 > \mu_2$ | $\mu_2 > \mu_1$ (course is EFFECTIVE) (difference significant)

[Two tailed test]

STEP 2:

LOS = 5% (two tailed test)

Degree of freedom = $n_1 + n_2 - 2 = 9+9-2 = 16$

∴ Critical Value $t_2 = [2.31]$

STEP 3:

Before	$x_{1t} - \bar{x}_1$	$(x_{1t} - \bar{x}_1)^2$	After	$x_{2t} - \bar{x}_2$	$(x_{2t} - \bar{x}_2)^2$
10	-0.33	0.1089	12	0.89	0.7921
15	4.67	21.8089	17	5.89	34.6921
9	-1.33	1.7689	8	-3.11	9.6721
3	-7.33	53.7289	5	-6.11	37.3321
7	-3.33	11.0889	6	-5.11	26.1121
12	1.67	2.7889	11	-0.11	0.0121
16	5.67	32.1989	18	6.89	47.4721
17	6.67	44.4889	20	8.89	79.0321
4	-6.33	40.0689	3	-8.11	65.7721

$\Sigma = 93$ $\qquad\qquad$ $\Sigma = 100$

$\bar{x}_1 = \frac{\Sigma x_1 t}{n_1} = \frac{93}{9} = \boxed{10.33}$ $\qquad$ $\bar{x}_2 = \frac{\Sigma x_2 t}{n_2} = \frac{100}{9} = \boxed{11.11}$

$$SP = \sqrt{\frac{\Sigma(x_{1t} - \bar{x}_1)^2 + \Sigma(x_{2t} - \bar{x}_2)^2}{n_1 + n_2 - 2}} = \sqrt{\frac{208.001 + 300.8899}{16}} = \boxed{5.6396}$$

U19CS012

$$S.E. = S_P \sqrt{\frac{1}{9} + \frac{1}{9}}$$

$$= \frac{5.6396 \sqrt{2}}{3}$$

$$\boxed{S.E. = 2.6581}$$

STEP4 : Test statistics

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{S.E} = \frac{10.33 - 11.11}{2.6581} = -0.2934$$

$$|t_{cal}| = 0.2934$$

STEP 5 :

Since $|t_{cal}| < t_\alpha$, H_0 is accepted

Course was not effective in improving performance.

6.) A sample of 20 items has mean 42 units and S.D 5 units. Test the hypothesis that it is a random sample from Normal population with mean 45 units. [t_{tab} at 5% LOS for 19 d.o.f = 2.09]

6.) STEP 1 : Set up Null and Alternative hypothesis (H_0 & H_1)

H_0 : $\bar{x} = \mu$

H_1 : $\bar{x} \neq \mu$

STEP 2 : Let the Level of Significance be 5%

with $n-1 = 20-1 = 19$ degree of freedom

$$\boxed{t_{0.05} = 2.09}$$

STEP 3 : Test statistics

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} = \frac{42-45}{5/\sqrt{20}} = \boxed{-2.68}$$

$$|t| = 2.68$$

STEP 4 : $t_{cal} = 2.68$ Hypothesis $|t_{cal}| > t_\alpha$

$t_\alpha = 2.09$ (Rejected)

ANS : Hence, we conclude that the assumption of mean of 45 for population is not reasonable.

U19CS012

7.) Two samples of Sodium Vapour bulbs were tested for length of life and the following results were got

	Size	Sample Mean	Sample S.D.
Type I	8	1234 hrs	36 hrs
Type II	7	1036 hrs	40 hrs

Is the difference in means significant to generate that type I is superior to type II regarding length of life ? [$t_{0.05}$ at 13 d.o.f = 1.77]

7.) $n_1 = 8$ and $n_2 = 7$ (< 30 small sample)

STEP 1 :

Null Hypothesis (H_0) : $\mu_1 = \mu_2$ (difference insignificant)

Alternate Hypothesis (H_1) : $\mu_1 < \mu_2$ (difference significant)

STEP 2 :

LOS = 5% (two tailed test)

Degree of freedom $= n_1 + n_2 - 2 = 8+7-2 = 13$

critical value $(t_\alpha) = 1.77$

STEP 3 :

$\bar{x}_1 = 1234$ hrs $\qquad \bar{x}_2 = 1036$ hrs

$$SP = \sqrt{\frac{s_1^2 n_1 + s_2^2 n_2}{n_1 + n_2 - 2}} \qquad SE = SP\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= \sqrt{\frac{36^2 \times 8 + 40^2 \times 7}{13}} \qquad = 40.7317\sqrt{\frac{1}{8} + \frac{1}{7}}$$

$$= \sqrt{\frac{21569}{13}} \qquad = 40.7317\sqrt{\frac{15}{56}}$$

$$= \boxed{21.0806}$$

$$= \boxed{40.7317}$$

STEP 4 : Test Statistics

$$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{S.E.} = \frac{1234 - 1036}{21.0806} = \boxed{9.3925}$$

STEP 5 : $|t_{cal}| > t_\alpha$ H_0 is rejected

H_1 is accepted ✓

ANS : Type I is definately superior to Type II

VISION

U19CS012

8.) The following table is given:

		Eye color in Sons		
		Not Light	Light	Total
Father's Eye color	Not Light	230	148	378
	Light	251	471	622
	Total	381	619	1000

Test whether the colour of son's eyes is associated with that of the fathers.

[Given: Value of X^2 is 3.84 for 1 d.o.f]

8.) Expected count for each cell

Cell 1: $(NL_F) \times (NL_S)$ $\frac{378 \times 381}{1000} = 144.018$ Cell 3: $(L_F \times NL_S)$ $\frac{381 \times 622}{1000} = 231.982$

Cell 2: $(NL_F) \times (L_S)$ $\frac{378 \times 619}{1000} = 233.982$ Cell 4: $\frac{622 \times 619}{1000} = 385.018$

$$X^2_{cal} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} = \frac{(230 - 144.018)^2}{144.018} + \frac{(148 - 233.982)^2}{233.982} + \frac{(251 - 236.982)^2}{236.982} + \frac{(471 - 385.018)^2}{385.018}$$

$$X^2 = 51.33 + 31.59 + 0.82 + 19.20$$

$$X^2 = 102.94$$

Ans: $X^2_{cal} > X^2_{tab}$

Ho Rejected, H_1 accepted [Color of Son's eye Not Associated with color of father's eye]

9.) From the following table, showing the number of plants having certain characters, test the hypothesis that the flower colour is independent of flatness of leaf.

[X^2_{tab} = 0.0158 at 0.1 < 0.5 for 1. D.O.F.]

VISION

U19CS012

	Flat Leaves	Curved Leaves	Total
White Flowers	99	36	135
Red Flowers	20	5	25
Total	119	41	160

Expected count for each cell

Cell 1 (WF) : $\frac{119 \times 135}{160} = 100.40$ Cell 3 (RF) : $\frac{25 \times 119}{160} = 18.59$

Cell 2 (WC) : $\frac{135 \times 41}{160} = 34.59$ Cell 4 (RC) : $\frac{(25 \times 41)}{160} = 6.40$

$$X^2_{cal} = \sum \frac{(\text{observed} - \text{expected})}{\text{expected}} = \frac{(99-100.40)^2}{100.40} + \frac{(36-34.59)^2}{34.59} + \frac{(20-18.59)^2}{18.59} + \frac{(5-6.40)^2}{6.40}$$

$$= 0.0195 + 0.0574 + 0.1069 + 0.3062$$

$$= 0.4900$$

ANS : $X^2_{cal} > X^2_{tab}$, Ho is rejected

[Flower colour is not independent of flatness]

10.) A set of five similar coins is tossed 320 times and the result is :

No. of Heads	0	1	2	3	4	5
Frequency	6	27	72	112	71	32

Test the hypothesis that the data follow a binomial distribution.
[$X^2_{tab} = 11.07$ for 5 d.o.f]

10.) Null Hypothesis

Ho : The Given data fits the Binomial distribution

$p = q = 1/2$ $n = 5$ $N = 320$

PTO →

U19CS012

No. of heads	$P(X=x) = {}^5C_x\, p^x q^{5-x}$	Expected Frequency $N.P(X=x)$
0	$1/32$	10
1	$5/32$	50
2	$10/32$	100
3	$10/32$	100
4	$5/32$	50
5	$1/32$	10
		320

Computation of Chi square values

Observed Frequency	Expected Frequency	$(O-E)^2$	$\frac{(O-E)^2}{E}$
6	10	16	1.6
27	50	529	10.58
72	100	784	7.84
112	100	144	1.44
71	50	441	8.82
32	10	484	48.4
			78.68

$\chi^2_{cal} = 78.68$

$\boxed{\chi^2_{cal} > \chi^2_{tab}\text{ , Ho is rejected}}$

∴ The Given data ~~fits~~ does not fits the binomial distribution

11.) Fit a Poisson distribution to the following data and test the goodness of fit.

x	0	1	2	3	4	5	6
f	275	72	30	7	5	2	1

[χ^2_{tab} at 5% L.O.S for 2 d.o.f = 5.94]

U19CS012

11) Null Hypothesis : The Given data fits the Poisson distribution
Level of significance $\alpha = 0.05$

Computated of expected frequencies $= \frac{(6\times1 + 5\times2 + 5\times4 + 7\times3 + 30\times2 + 72\times1 + 275\times0)}{275 + 72 + 30 + 7 + 5 + 2 + 1}$

$= \frac{189}{392} = \boxed{0.482}$

$P(0) = \frac{e^{-0.482} \times (0.482)^0}{0!} = 0.617$

$f(0) = N.P(0) = 392 \times 0.617 = \boxed{241.864}$

The other expected frequencies will be obtained by using recurrence formula.

$f(x+1) = \frac{m}{(x+1)} \times f(x)$

Putting $x = 0, 1, 2$, we obtain the following frequency

$f(1) = 0.482 \times 241.864 = 116.578$

$f(2) = \left(\frac{0.482}{2}\right) \times 116.578 = 28.095$

$f(3) = \left(\frac{0.482}{3}\right) \times 28.095 = 4.513$

$f(4) = \left(\frac{0.482}{4}\right) \times 4.513 = 0.543$

$f(5) = \left(\frac{0.482}{5}\right) \times 0.543 = 0.052$

$f(6) = \left(\frac{0.482}{6}\right) \times 0.052 = 0.004$

x	f
0	241
1	116
2	28
3	7
4	1
5	0
6	0
	$\Sigma = 392$

P. T. O. →

U19CS012

Observed Frequency	Expected Frequency	$(O-E)^2$	$\frac{(O-E)^2}{E}$
275	241	1156	4.996
72	116	1936	16.689
30	28	4	0.142
7	4	9	2.25
5	1	16	16
2	1	1	1
1	1	0	0

$\Sigma = 40.877$

$\chi^2_{cal} = 40.877$

$\chi^2_{cal} > \chi^2_{tab}$

ANS. Poisson Distribution (P.D.) is Not Good Fit.

SUBMITTED BY :

U19CS012

BHAGYA VINOD RANA

CSE (IInd Yr)