"The prob of the value of the variate falling in critical region is level of significance"

29 4 2012

Sampling tests

process of selecting samples

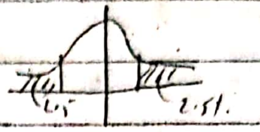
On the basis of sample information we make certain decisions on the popln. On taking such decision we make certain assumption — assumption statistical hypothesis

Testing of hypothesis :-

SMALL SAMPLES → subset of popln

collecting individuals or attributes of popln

(n < 30) ← t - distribution



$z = -1.96$ $z = 1.96$

Part I :-

Testing the significance of the difference betwn mean of sample ($\bar{x}$) and mean of population (μ)

A random sample of size 16 has mean of 54. The sum of the squares of deviations of the observations from their mean is 135. Can this sample be regarded as drawn from population having mean of 56?

For population, $\mu = 56$

For sample, $\bar{x} = 54$

$\Sigma (x - \bar{x})^2 = 135$

$n = 16$

H_0 : Hypothesis of no diff. — no diff betn observed & expected value

H_0 : There is no difference betwn mean of sample and mean of population (i.e. $\mu = 56$ is valid)

H_a : $\mu \neq 56$ [for 2-tail test]

Now $s = \sqrt{\dfrac{\Sigma (x - \bar{x})^2}{n - 1}}$ ← [For small sample]

$= \sqrt{\dfrac{135}{16 - 1}} = \sqrt{9} = 3$

$\therefore$ standard Error $= \frac{s}{\sqrt{n}}$ [$\because$ s is calcul...

$= \frac{3}{\sqrt{16}}$

$= 0.75$

then $t = \frac{|\bar{x} - \mu|}{S.E.} = \frac{|54-56|}{0.75} = 2.\overline{6}$

Also $\vartheta = n - 1 = 16 - 1 = 15$

$\therefore t_{0.05} = 2.131$ [for 2 tail test]

thus $t > t_{0.05}$

$\Rightarrow$ Difference between mean of sample mean of population is significant

Hence the sample cannot be as drawn from population having 56 at 5% level of significance.

Note! - For 1% L.O.S. we have $t_{0.01} = 2.947$ [for 2-tail t...]

Thus $t < t_{0.01}$

$\Rightarrow$ Difference between mean of sample and mean of population is not significant.

Hence the sample can be regarded as drawn from population having mean of 56 at 1% level of significance. [L.O.S]

Scanned with CamScanner

2) The weekly sales of a powder in a super market was 146.3 kg. After a special advertising campaign the mean weekly sales in 22 of its branches increased to 153.7 kg. with S.D. of 17.2 kg. Was the advertising campaign successful? → i.e. mean of popl^n also increase

Ans:- For population, $\mu = 146.3$

For sample, $\bar{x} = 153.7$

$s = 17.2$

$n = 22$

If advertising is successful there will be diff. between the mean of sample & mean of popl^n

Ho: There is no difference between mean of sample and mean of population (i.e $\mu = 146.3$ is valid) i.e (advertising campaign wasn't successful)

// i.e even if sample mean increase mean of the popul. don't increase

Ha: $\mu > 146.3$ [Right-tail test]

Now, $S.E = \dfrac{s}{\sqrt{n-1}}$ [$\because$ s is given]

$= \dfrac{17.2}{\sqrt{21}}$

$= 3.75$

Then $t = \dfrac{|\bar{x} - \mu|}{S.E} = \dfrac{|153.7 - 146.3|}{3.75} = 1.97$

Also $\nu = n - 1 = 22 - 1 = 21$

$\therefore t_{0.05} = 1.721$ [for 1-tail test]

Thus $t > t_{0.05}$

⇒ Difference betw^n mean of sample & mean of population is significant.

Hence the advertising campaign is successful.

3.) The weights of 10 persons of a locality are found to be 70, 67, 62, 68, 61, 68, 70, 64, 64, 66, kg. At 1% level of significance is it reasonable to believe that the average weight of the people of that locality is (i) 64 kg (ii) > 64 kg (iii) < 64 kg. Also find 99% confidence limits for population mean is (i) above.

Ans:-

Note:- In the above problem either (i) or (ii) or (iii) can be asked at a time (not all together).

Soln:- For population, $\mu = 64$

For sample,

x	$(x-\bar{x})^2$
70	16
67	1
62	16
68	4
61	25
68	4
70	16
64	4
64	4
66	0
$\Sigma x = 660$	$\Sigma(x-\bar{x})^2 = 90$

Now, $\bar{x} = \frac{660}{10} =$

&

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

$$= \sqrt{\frac{90}{9}}$$

$$= 3.162$$

Scanned with CamScanner

$t < t_{0.05}$ ∗ not significant means 'H_0' is true.
$t > t_{0.05}$ ∗ significant means 'H_a' is true.

$\therefore$ $S.E. = \frac{s}{\sqrt{n}} = \frac{3.162}{\sqrt{10}} = 1$ [$\because$ s is calculated]

Then, $t = \frac{|\bar{x} - \mu|}{S.E.}$

$= \frac{|66 - 64|}{1}$

$= 2$

Also $\nu = n - 1 = 10 - 1 = 9$

H_0: There is no difference ~~dif~~ between mean of sample and mean of population (i.e $\mu = 64$ is valid).

i) H_a: $\mu \neq 64$ [2 - tail test]
$\therefore$ $t_{0.01} = 3.250$
Thus $t < t_{0.01}$
$\Rightarrow$ Difference between mean of sample and mean of population is not significant.
Hence it is reasonable to believe that the average weight of the people of that locality is 64 kg.

ii) H_a: $\mu > 64$ [right tail test]
$\therefore$ $t_{0.01} = 2.821$ [for 1-tail test]
Thus $t < t_{0.01}$

$\Rightarrow$ Difference between mean of sample and mean of population is not significant.

Hence it is not reasonable to believe th[at] the average weight of the people of that locality is greater than 64 kg.

iii) Ha: $\mu < 64$ [left tail test]

$\therefore t_{0.01} = 2.821$ [for 1-tail test].

Thus $t < t_{0.01}$

$\Rightarrow$ Difference between mean of sample an[d] mean of population is not significant.

Hence it is not reasonable to beli[eve] that the average weight of the people of that locality is less than 64 kg.

The 99% confidence limits for the popula[tion] mean in (i) above are given by:

$\bar{x} \mp t_{0.05} \times S.E.$

i.e. $66 \mp 3.250 \times 1$

i.e $\boxed{62.75 \text{ and } 69.25}$ kg.

Part II :- Testing the significance of the difference between means ($\bar{x}_1$ and $\bar{x}_2$) of 2 samples.

A) If samples are independent :-

1) The mean life of a sample of 10 electric bulbs is 1456 hrs. with S.D. of 423 hrs. A second sample of 17 bulbs from different batch has a mean of 1280 hrs with S.D. of 398 hrs. Is there a significant difference between means of two samples?

→ Sample - I

$\bar{x}_1 = 1456$

$s_1 = 423$

$n_1 = 10$

Sample - II

$\bar{x}_2 = 1280$

$s_2 = 398$

$n_2 = 17$

Ho :- There is no difference between means of two samples [i.e. $\mu_1 = \mu_2$]

Ha : $\mu_1 \neq \mu_2$ [2-tail test]

Now $\sigma = \sqrt{\dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$

$= \sqrt{\dfrac{10(423)^2 + 17(398)^2}{10 + 17 - 2}}$

$= 423.42$

S.E. →

$$S.E. = \sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}} = 423.42\sqrt{\frac{1}{10}+\frac{1}{17}} = 168$$

then $$t = \frac{|\overline{x}_1 - \overline{x}_2|}{S.E.} = \frac{|1456-1280|}{168.74} = 1.04$$

Also $\nu = n_1 + n_2 - 2 = 10 + 17 - 2 = 25$

$\therefore t_{0.05} = 2.060$ [from table]

Thus $t < t_{0.05}$

$\Rightarrow$ Difference between the means of the two samples is not significant.

Hence the two samples can be regarded as drawn from the same population or from two different populations with same means.

2) For a random sample of 10 persons fed on diet A, the increase in weights for a certain period were 10, 6, 16, 17, 13, 12, 8, 14, 15, 9 pounds.

For another sample of 12 persons fed on diet B, the increase in weight for the same period were 7, 13, 22, 17, 15, 12, 14, 18, 8, 21, 23, 10. pounds.

(i) test ~~whether the~~ two diets, differ significantly as regards to increase in weight?

(ii) Do you agree with the claim that diet B is superior to A as regards to increase in weight.

Note:- In the above prob: either (i) or (ii) will be asked simultaneously not together.

Solution:-

Diet A		Diet B	
x_1	$(x_1-\bar{x}_1)^2$	x_2	$(x_2-\bar{x}_2)^2$
10	4	7	64
6	36	13	4
16	16	22	49
17	25	17	4
13	1	15	0
12	0	12	9
8	16	14	1
14	4	18	9
15	9	8	49
9	9	21	36
$\Sigma x_1=120$	$\Sigma(x_1-\bar{x}_1)^2=120$	23	64
		10	25
		$\Sigma x_2=180$	$\Sigma(x_2-\bar{x}_2)^2=314$

Now $\bar{x}_1=\frac{\Sigma x_1}{n_1}=\frac{120}{10}$
$=12$.

$\bar{X}_2=\frac{\Sigma x_2}{n_2}=\frac{180}{12}$
$=15$

Now $\sigma=\sqrt{\frac{\Sigma(x_1-\bar{x}_1)^2+\Sigma(x_2-\bar{x}_2)^2}{n_1+n_2-2}}$

$=\sqrt{\frac{120+314}{10+12-2}}=4.66$

$\therefore S.E.=\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}=4.66\sqrt{\frac{1}{10}+\frac{1}{12}}=1.99$

Then, $t=\frac{|\bar{x}_1-\bar{x}_2|}{S.E}=\frac{|12-15|}{1.99}=1.5$

Also $\nu = n_1 + n_2 - 2 = 10 + 12 - 2 = 20$

Ho: There is no diff betⁿ means of two samp (i.e $\mu_1 = \mu_2$ is valid)

i) Ha: $\mu_1 \neq \mu_2$ [2 tail-test]

∴ $t_{0.05} = 2.086$ [for 2-sided test]

Thus $t < t_{0.05}$

⇒ Diff. betwⁿ means of two samples is not significant. Hence there is no difference be the two diets as regards in increase in

ii) Ha: $\mu_1 < \mu_2$ [left tail test]

∴ $t_{0.05} = 1.725$ (for 1-sided test]

Thus $t < t_{0.05}$

⇒ Diff. betwⁿ means of two samples is not significant. Hence the claim that diet B is superior to diet A is not superior.

B) If samples are dependent.

1) To varify whether the a course is mathematic improved performance or not, two tests were given to 12 participants [1 before & other after the course]. Marks of test before the course were 44, 40, 61, 52, 32, 44, 70, 41, 67, 72, 53, 72.
Corresponding marks of test after course were 53, 38, 69, 57, 46, 39, 73, 48, 73, 74, 60, 7
Determine whether the couse was usef or not.

Sol^n :- H_0: There is no difference bet^n mean marks of two tests (i.e $\mu_1 = \mu_2$ is valid)

H_a: $\mu_1 < \mu_2$ [i.e left tail test]

x_1	x_2	$d = x_2 - x_1$	$(d-\bar{d})^2$
44	53	9	16
40	38	-2	49
61	69	8	9
52	57	5	0
32	46	14	81
44	39	-5	100
70	73	3	4
41	48	7	4
67	73	6	1
72	74	2	9
53	60	7	4
72	78	6	1
		$\Sigma d = 60$	$\Sigma (d-\bar{d})^2 = 278$

Now $\bar{d} = \frac{\Sigma d}{n} = \frac{60}{12} = 5$

& $s = \sqrt{\frac{\Sigma (d-\bar{d})^2}{n-1}} = \sqrt{\frac{278}{11}} = 5.027$

$\therefore$ S.E. $= \frac{s}{\sqrt{n}} = \frac{5.027}{\sqrt{12}} = 1.45$

then $t = \frac{|\bar{d}|}{S.E} = \frac{5}{1.45} = 3.45$

Also $\nu = n-1 = 12-1 = 11$

$\therefore$ $t_{0.05} = 1.796$ [from one sided test (table)]

Thus $t > t_{0.05}$

$\bar{d} = \bar{x}_2 - \bar{x}_1$ or $\bar{x}_1 - \bar{x}_2$

$\Rightarrow$ Diff betn the means of two samples is signifi
Hence the cause is useful.

2.) A drug is given to 10 patients and the inc
in their blood pressures were recorded
3, 6, −2, 4, −3, 4, 6, 0, 0, 2.

i) It is reasonable to believe that the dr
has no effect on change in B.P.
ii) Do you agree that the B.P. increases wi
intake of drug.
Note :- either (i) or (ii) will be asked not both

Ans:-

d	$(d-\bar{d})^2$
3	1
6	16
−2	16
4	4
−3	25
4	4
6	16
0	4
0	4
2	0
$\Sigma d = 20$	$\Sigma(d-\bar{d})^2 = 90$

Now $\bar{d} = \frac{\Sigma d}{n} = \frac{2}{1}$

$= 2$

$s = \sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}}$

$= \sqrt{\frac{90}{9}} = 3.162$

$\therefore S.E = \frac{s}{\sqrt{n}} = \frac{3.162}{\sqrt{10}}$

Then $t = \frac{|\bar{d}|}{S.E.} = \frac{2}{1} = 2$

Also $\nu = n-1 = 10-1 = 9$

$H_0: \mu_1 = \mu_2$ (i.e there is no change in B.P. to d

i) $H_a: \mu_1 \neq \mu_2$ [2-tail test]

$\therefore t_{0.05} = 2.262$

Thus $t < t_{0.05}$

$\Rightarrow$ Difference betⁿ the means of the two samples is not significant.

Hence the drug has no significant effect on change in B.P.

ii) $H_a: \mu_1 < \mu_2$ [1-tail test]

$\therefore t_{0.05} = 1.833$

Thus $t > t_{0.05}$

$\Rightarrow$ Difference between means of the two samples is significant. Hence the B.P. increase with intake of drug.

5-12

B] Large samples :-

($n > 30$) - normal distribution

level	2-tail	1-tail
5%	$z_{0.05} = 1.96$	$z_{0.05} = 1.65$
1%	$z_{0.01} = 2.58$	$z_{0.01} = 2.33$

I) Testing the significance of the difference betⁿ mean of sample ($\bar{x}$) and mean of population (μ);

1) A Sample of 400 student is taken from a large population. There mean height is ~~171.38 cm with std~~. deviation of 3.3 cm. Can it reasonably be concluded that this ~~height~~ sample is drawn from population with mean ~~limits~~ ht of 171.17 cm. Also find 99% fiducial limits for population mean.

$\Rightarrow$ for population, $\mu = 171.17$

$n = 400$

For sample, $\bar{x} = 171.38$, $s = 3.3$,

H_0: There is no diff. betn mean of sample & mean of populatn (i.e $\mu = 171.17$ is valid)

H_a: $\mu \neq 171.17$ (2 sided test).

Now $S.E = \frac{\sigma_p}{\sqrt{n}}$ whr σ_p = S.D of population

If σ_p is not given then,

$SE = \frac{s}{\sqrt{n}}$ where s = S.D of sample.

$\therefore S.E = \frac{s}{\sqrt{n}} = \frac{3.3}{\sqrt{400}} = 0.165$

$\therefore z = \frac{|\bar{x} - \mu|}{SE} = \frac{|171.38 - 171.17|}{0.165} = 1.27$

But $z_{0.05} = 1.96$ {for 2-sided test}

Thus $z < z_{0.05}$

$\Rightarrow$ Diffn betn mean of sample and mean of popul is not significant.

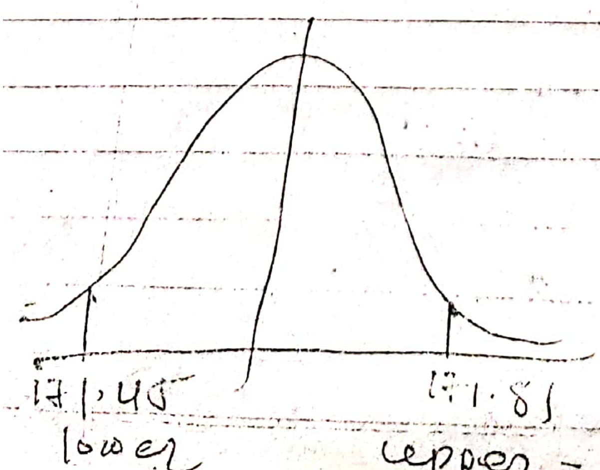
Hence the sample can be regarded as drawn from the population having mean height of 171.17

The 99% fiducial limits for population mean are given by,

$$\bar{x} \mp z_{0.01} \times S.E$$

i.e $171.38 \mp 2.58 \times 0.165$

i.e $\boxed{170.95 \text{ \& } 171.81 \text{ cm}}$



Ans. $\sigma = 2.5$, $n_1 = 1000$

$n_2 = 2000$

$\bar{x}_1 = 67.5$

$= 0.15$

$z = 3$

cal. value > table value

rejected

II) Testing the significance of the difference between means ($\bar{x}_1$ & $\bar{x}_2$) of two samples

2.) In a certain factory there are 2 independent processs to manufacture the same product. the avg. wt of a sample of 250 item from the 1st process is found to be 120g with S.D. of 12gm. The corresponding figures for the sample of 400 items from another process are 124gm & 14gm respectively. Is there any difference between the two process. Also find 95% confidence limits for the diff in aveg wt of the prod manufactured by 2 processes.

Declir

The means of two samples, ~~of sizes 1000 and 2000~~ respectively are 67.50 & 68 inches can ~~the~~ sample be regarded as drawn from same population of SD = 2.5 inches

Process I	Process II
$\bar{x}_1 = 120$	$\bar{x}_2 = 124$
$S_1 = 12$	$S_2 = 14$
$n_1 = 250$	$n_2 = 400$

Ho: There is no diff bet^n the two processes [i.e. $\mu_1 = \mu_2$]

Ha: $\mu_1 \neq \mu_2$ [i.e. 2 sided test]

Now $S.E. = \sigma_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ whr, σ_p = S.D. of population

If σ_p is not given then,

$S.E. = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$ whr, S_1 & S_2 = S.D's of samples

$\therefore S.E. = \sqrt{\frac{(12)^2}{250} + \frac{(14)^2}{400}} = 1.03$

Then $z = \frac{|\bar{x}_1 - \bar{x}_2|}{S.E.} = \frac{|120 - 124|}{1.03} = 3.87$

But $Z_{0.05} = 1.96$ (for 2 sided test)

Thus $z > Z_{0.05}$

$\Rightarrow$ Diff bet^n means of the two samples is significant.

Hence there is diff bet^n the two processes.

* If large diff. betn the value then it is good fit data.

The 95% confidence limits for diff in the avg w/s the products manufactured by the two processes are given by, $|\bar{x}_1 - \bar{x}_2| \mp Z_{0.05} \times S.E.$

i.e $4 \mp 1.96 \times 1.03$, i.e $\boxed{1.982 \quad 6.018}$ lcg

new chapter ⟹ CHI - SQUARE (χ^2) TEST

A] For testing the goodness of fit.

1) The fig. given below are the observed freq of a variable X^2: 1, 5, 20, 28, 42, 22, 15, 5, 2. On fitting a normal distribution to data, the corresponding freq worked out as 1, 6, 18, 25, 40, 25, 18, 8, 1. Is it reasonable to believe that the normal distribution is a good fit to the data.

Soln:- Ho :- The normal distribution is good fit to data.

Observed freq (O)	Expected freq (E)	$\frac{(O-E)^2}{E}$
1, 5 } 6*	1, 6 } 7*	0.143
20	18	0.222
28	25	0.360
42	40	0.100
22	25	0.360
15	18	0.500
5, 2 } 7*	8, 1 } 9*	0.444
		$\boxed{2.129}$

* If individual freq is less than 5 then group two or more freq together so that the sum is at least 5.

Now $\chi^2 = \sum \frac{(O-E)^2}{E} = 2.129$,

Also $v = n - k = 7 - 3 = 4$ [$\because k = 3$ for N.D.]

$\tau = 1$ for B.D, $k = 2$ for P.D then $k = 3$ for N.D

Note: $\tau = 1$ otherwise.

(34) 50

$\therefore \chi^2_{0.05} = 9.49$ [from table], Thus, $\chi^2 < \chi^2_{0.05}$

$\Rightarrow$ Diff. between the two freq. is not significant.

Hence the normal distribution is a good fit to the data.

Ex.:-

2.) A set of 5 coins is tossed 3200 times & no. of heads appearing each time are noted.

no. of heads:	0	1	2	3	4	5
freq:-	80	570	1100	900	500	50

Test the hypothesis that the coins are biased.

Sol:- Binomial distribution:

Success: Getting a head

If the coins are unbiased then, $p = \frac{1}{2}$, $q = 1 - \frac{1}{2} = \frac{1}{2}$

Also $n = 5$, $N = 3200$

By B.D we have,

$$P(X = x) = P(x) = {}^nC_x\, p^x\, q^{n-x}$$
$$= {}^5C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{n-x} = {}^5C_x \left(\frac{1}{2}\right)^5$$

Then the theoretical freqs. are given by,

$N \cdot P(X = x) = 3200 \times {}^5C_x \left(\frac{1}{2}\right)^5 = 100 \times {}^5C_x$ where $x = 0$ to 5

Hence the theoretical freq work out as

100, 500, 1000, ~~1500~~ 1000, 500, 100

χ^2- test:

Q - H_0:- The coins are unbiased.

The following table gives the no. of accidents in a city during a week. Find ~~whether the accidents~~ are uniformly distributed over a week using χ^2 test.

day	S	M	T	W	Th	F	Sat	Total
No. of accidents	13	15	09	11	12	10	14	84

Observed freq (O)	expected freq. (E)	$\frac{(O-E)^2}{E}$
80	100	4.0
570	500	9.8
1100	1000	10.0
900	1000	10.0
500	500	0.0
50	100	25.0
		58.8

Now $\chi^2 = \sum \frac{(O-E)^2}{E} = 58.8$

Also $\nu = n-k$ ($\because k=1$ for B.D.) [$n \to$ no. of observation i.e 6]

$= 6-1 = 5$

$\therefore \chi^2_{0.05} = 11.07$, Then $\chi^2 > \chi^2_{0.05}$

$\Rightarrow$ Difference betn obsewed & expected freq. is significant.

Hence the coins are biased.

In an experiment on pea breeding, Mendal obtained the following freq of seeds: 315 round seeds of yellow colour, 101 wrinkled seeds of Yellow colour,

108 round & " green "

32 wrinkled & " " ".

According to his theory of heredity, these nos should be in the proportion of 9:3:3:1. Is their any evidence to doubt his theory of heredity at 1% level of significance.

Ho: There is no difference betn obsewed & expected frequencies of the different types of seeds.

Now total no. of seeds $= 315+101+108+32 = 556$

According to the theory of heredity, the nos. of different types of seeds should be in the proportion of 9:3:3:1
Hence their expected frequencies are,

$$E_1 = \frac{9}{16} \times 556 = 312.75$$

$$E_2 = \frac{3}{16} \times 556 = 104.25$$

$$E_3 = \frac{3}{16} \times 556 = 104.25$$

$$E_4 = \frac{1}{16} \times 556 = 34.75$$

Observed freq 'O'	Expected freq 'E'	$\frac{(O-E)^2}{E}$
315	312.75	0.016
101	104.25	0.101
108	104.25	0.134
32	34.75	0.217
		0.468

Now, $\chi^2 = \sum \frac{(O-E)^2}{E} = 0.468$

Also $\nu = n - k = 4 - 1 = 3$

$\therefore \chi^2_{0.01} = 11.34$ [from table]

Then $\chi^2 < \chi^2_{0.01}$

$\Rightarrow$ Difference betn observed & expected frequencies is not significant
Hence there is no evidence to doubt mendal's theory of heredity.

Attribute = Qualitative characteristics.

Testing the association between two attributes:—

In an experiment on ~~immunization~~ immunization of ~~cattle~~ cattle for T.B. The following results were obtained.

Type of cattle	Affected	Not affected
Inoculated	12	28
Not Inoculated	18	4

use chi-square test to discuss the effectiveness of vaccine in controlling T.B.

H_0: There is no association betn inoculation and disease.

After applying Yate's correction the observed frequencies are tabulated as follows-

$O_{11} = 12.5$ $\quad O_{12} = 27.5$ $\quad R_1 = 40$

$O_{21} = 17.5$ $\quad O_{22} = 4.5$* $\quad R_2 = 22$

$C_1 = 30$ $\quad C_2 = 32$ $\quad N = 62$

The corresponding expected frequencies are calculated as follows:

$E_{11} = \frac{R_1 \times C_1}{N} = 19.35$, $\quad E_{12} = \frac{R_1 \times C_2}{N} = 20.65$

$E_{21} = \frac{R_2 \times C_1}{N} = 10.65$; $\quad E_{22} = \frac{R_2 \times C_2}{N} = 11.35$

* whenever cell freq^n is less than 5, apply yate's correction (i.e +0.5, -0.5)

observed freq 'O'	Expected freq 'E'	$\frac{(O-E)^2}{E}$
12.5	19.35	2.42
27.5	20.85	2.27
17.5	10.65	4.41
9.5	11.35	4.13
		13.23

Now $\chi^2 = \Sigma \frac{(O-E)^2}{E}$ $\cancel{= 13.23}$

Also $V = (r-1)(c-1) = (2-1)(2-1) = 1$

$\therefore \chi^2_{0.05} = 3.84$ (from table)

Then $\chi^2 > \chi^2_{0.05}$

$\Rightarrow$ Difference betn observed & expected frequencies is significant.

Hence the vaccine is effective in controlling T.B.

2) A random sample of students of a particular university was selected & they were asked about their opinion on autonomy to colleges.

Class of students	In favour of autonomy	Not in favour of autonomy
B.Sc	120	80
B.Com	130	70
B.A.	70	30
P.Grad	80	20

Test the hypothesis that the opinions of the students are independent of their class groupings.

H_0: There is no association betn opinions of students & their class groupings.
The observed freqn are tabulated as follows.

$O_{11} = 120$ $\quad O_{12} = 80$ $\quad R_1 = 200$

$O_{21} = 130$ $\quad O_{22} = 70$ $\quad R_2 = 200$

$O_{31} = 70$ $\quad O_{32} = 30$ $\quad R_3 = 100$

$O_{41} = 80$ $\quad O_{42} = 20$ $\quad R_4 = 100$

$C_1 = 400$ $\quad C_2 = 200$ $\quad N = 600$

The corresponding expected frequencies are calculated as follows.

$E_{11} = \frac{R_1 \times C_1}{N} = 133.33$ $\qquad E_{12} = \frac{R_1 \times C_2}{N} = 66.67$

$E_{21} = \frac{R_2 \times C_1}{N} = 133.33$ $\qquad E_{22} = \frac{R_2 \times C_2}{N} = 66.67$

$E_{31} = \frac{R_3 \times C_1}{N} = 66.67$ $\qquad E_{32} = \frac{R_3 \times C_2}{N} = 33.33$

$E_{41} = \frac{R_4 \times C_1}{N} = 66.67$ $\qquad E_{42} = \frac{R_4 \times C_2}{N} = 33.33$

observed freqn 'o'	Expected freqn 'E'	$\frac{(O-E)^2}{E}$
120	133.33	
80	66.67	
130	133.33	
70	66.67	
70	66.67	
30	33.33	
80	66.67	
20	33.33	
		12.75

Now, $\chi^2 = \dfrac{\Sigma (O-E)^2}{E} = 12.75$

$|t| = 2.68$

Also, $v = (r-1)(c-1)$

$= (4-1)(2-1)$

$= 3$

$v = 9+7-2 = 14$

$t_{(0.05)} = 2.145$

$t > t_2$ H0 is rejected

$\therefore \chi^2_{0.05} = 7.82$ (from table)

The samples can't be considered to have been drawn from the same population

Thus $\chi^2 > \chi^2_{0.05}$

$\Rightarrow$ Difference bet^n observed and expected frequencies is significant.

Hence there is association between opinions of students and their class groupings.