Large Sample Test

① A sample of 100 students is taken from a large population. The mean height of the students in this sample is 160 cm. Can it be reasonably regarded that in the population, the mean height is 165 cm & the S.D. is 10 cm.

Ans Given $\bar{x} = 160$ (mean of sample), $n = 100$, $\mu = 165$ (mean of population) & $\sigma = 10$

Step I) $H_0: \bar{x} = \mu$

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

Step II) Two-tailed test is to be used

Let L.o.s. be 1% $\therefore Z_\alpha = 2.58$

Step III) $Z_{cal} = \dfrac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \dfrac{160 - 165}{10/\sqrt{100}} = -5$

Step IV) $|Z_{cal}| \nless Z_\alpha \Rightarrow H_0$ is rejected

Step V) It is not statistically correct that, the mean height of the population is 165 cm.

② In a random sample of size 500, the mean is found to be 20. In another independent sample of size 400, the mean is 15. Could the samples have been drawn from the same population with S.D. 4?

Ans Given $\bar{x}_1 = 20$ (mean of 1st sample), $n_1 = 500$, $\bar{x}_2 = 15$ (mean of 2nd sample), $n_2 = 400$, $\sigma = 4$

Step I) $H_0: \bar{x}_1 = \bar{x}_2$

$H_1: \bar{x}_1 \neq \bar{x}_2$

Step II) Two tailed test is to be used

Let L.O.S. be 1% $\therefore Z_\alpha = 2.58$

Step III) $Z_{cal} = \dfrac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \dfrac{20 - 15}{4\sqrt{\frac{1}{500} + \frac{1}{400}}} = 18.6$

Step IV $|Z_{cal}| \nless Z_\alpha \Rightarrow H_0$ is rejected

Step V The samples could not have been drawn from the same population.

Eg (3) A simple sample of heights of 6400 English mean has a mean of 170 cm & a S.D. of 6.4 cm, while a simple sample of heights of 1600 Americans has a mean of 172 cm & a S.D. of 6.3 cm. Do the data indicate that Americans are on the average, taller than the English men?

Soln

Given $n_1 = 6400$, $\bar{x}_1 = 170$, $S_1 = 6.4$

$n_2 = 1600$, $\bar{x}_2 = 172$, $S_2 = 6.3$

Step I $H_0: \mu_1 = \mu_2$

$H_1: \mu_1 < \mu_2$

Step II) One tail test to be used.

Let L.O.S. be 1% $\therefore Z_\alpha = +2.33$

Step III) $$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{170-172}{\sqrt{\frac{(6.4)^2}{6400} + \frac{(6.3)^2}{1600}}} = -11.32$$

Step IV $|Z_{cal}| \nless Z_\alpha \Rightarrow H_0$ is rejected

Step V The Americans are, on the average taller than the Englishmen.

Eg (4) The average marks scored by 32 boys is 72 with a S.D of 8 while that for 36 girls is 70 with S.D. of 6. Test at 1% L.O.S whether the boys perform better than girls.

Step I) $H_0: \bar{x}_1 = \bar{x}_2$

$H_1: \bar{x}_1 > \bar{x}_2$

Step II) One tail test to be used, $Z_\alpha = 2.33$

Step III) $$Z_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} = \frac{72-70}{\sqrt{\frac{(8)^2}{32} + \frac{6^2}{36}}} = 1.15$$

Step IV $|Z_{cal}| < Z_\alpha$ $\therefore$ H_0 is accepted

statiscally we can't conclude that the boys perform better than girls

Eg (5) Test the significance of the difference betn the means of the samples, drawn from two normal populations with the same S.D. from the following data.

	Size	Mean	S.D.
sample 1	100	61	4
sample 2	200	63	6

Ans step I H_0: $\bar{x_1} = \bar{x_2}$

H_1: $\bar{x_1} \neq \bar{x_2}$

step II) Let L.O.S be 5% $Z_\alpha = 1.96$

} Populations have same S.D.

Step III $$Z_{cal} = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\frac{S_1^2}{n_2} + \frac{S_2^2}{n_1}}}$$ OR $$\left\{ Z_{cal} = \frac{\bar{x_1} - \bar{x_2}}{\sqrt{\left(\frac{n_1 S_1^2 + n_2 S_2^2}{n_1 + n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \right\}$$

$$Z_{cal} = \frac{61 - 63}{\sqrt{\frac{4^2}{200} + \frac{6^2}{100}}} = -3.02$$

Step IV $|Z_{cal}| \nless Z_\alpha$ $\therefore$ H_0 is rejected.

step V There is the significance of the difference betn the means of the samples.

Small Sample Test

Eg ① Tests made on the breaking strength of 10 pieces of a metal were gave the results: 578, 572, 570, 568, 572, 570, 570, 572, 596 & 584 kg. Test if the mean breaking strength of the wire can be assumed as 577 kg.

Ans Here $n=10$, use small sample test

$$\bar{x} = \frac{\sum_{i=1}^{10} x_i}{10} = 575.2 = E(X) \quad \& \quad E(X^2) = \frac{\sum_{i=1}^{10} x_i^2}{10}$$

SD = S is given as $S^2 = E(X^2) - [E(X)]^2 = 68.16$

i.e. $S = 8.26$

Step I) $H_0 ; \bar{x} = \mu$

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

Step II) Let L.O.S. be 5%, Two tailed test to be used.

Here $\nu = n-1 = 9$, $t_{5\%} = 2.26$ } From t-table.

Step III) $$t_{cal} = \frac{\bar{x} - \mu}{S/\sqrt{n-1}} = \frac{572.2 - 577}{8.26/\sqrt{9}} = -0.65$$

Step IV) $|t_{cal}| < t_\alpha$ $\therefore$ Ho is accepted.

Step V The mean breaking strength of the wire can be assumed as 577 kg at 5% L.O.S.

Eg ② A certain injection administered to each of 12 patients resulted in the following increases of blood pressure:

5, 2, 8, −1, 3, 0, 6, −2, 1, 5, 0, 4

Can it be concluded that the injection will be, in general, accompanied by an increase in B.P.?

Ans. $\bar{x} = \frac{\sum x}{n} = \frac{31}{12} = 2.58$

S.D "s" of the sample is given by $s^2 = E(x^2) - [E(x)]^2$

i.e. $s^2 = \frac{\sum x^2}{n} - \left(\frac{\sum x}{n}\right)^2 = \frac{185}{12} - (2.58)^2 = 8.76$

$\therefore s = 2.96$

Step I) $H_0 : \bar{x} = \mu$

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

{ $\mu = 0$, i.e. the injection will not result in increase in B.P. }

Step II) Let L.o.S. be 5%.

$\nu = n-1 = 12-1 = 11$, one tailed test to be used.

$\underset{\text{1 tailed test}}{t_{5\%, (\nu=11)}} = \underset{\text{2 tailed test}}{t_{10\%, (\nu=11)}} = 1.80$ } from t-table.

Step III) $t_{cal} = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{2.58 - 0}{2.96/\sqrt{11}} = 2.89$

Step IV $|t_{cal}| \nless t_{tab}$ $\therefore$ H_0 is rejected.

Step V We can conclude that the injection is accompanied by an increase in B.P.

Eg. The following data relate to the marks obtained by 11 students in two tests, one held at the begining of a year & the other at the end of the year after intensive coaching. Do the data indicate that the students have benefited by coaching?

Test 1	19	23	16	24	17	18	20	18	21	19	20
Test 2	17	24	20	24	20	22	20	20	18	22	19

Ans Let $d = x_1 - x_2$ $\Rightarrow d = 2, -1, -4, 0, -3, -4, 0, -2, 3, -3, 1$

where x_1 = marks in test 1 }
x_2 = — " — 2. }

Step I) $H_0 : \bar{d} = 0 \quad (\bar{x}_1 = \bar{x}_2)$

$H_1 : \bar{d} < 0 \quad (\bar{x}_1 < \bar{x}_2)$

Step II) Let L.O.S. be 5%. one tailed test to be used.

$\nu = n-1 = 11-1 = 10$

$$t_{5\%,(\nu=10)} = t_{10\%,(\nu=10)} = 1.81$$

one tailed test   2-tailed test

Step III) $t_{cal} = \dfrac{\bar{d}}{s/\sqrt{n-1}} = \dfrac{(-1)}{2.296/\sqrt{10}} = -1.38$

$$\left\{ \begin{array}{l} \text{Here } \bar{d} = \dfrac{\Sigma d_i}{11} = \dfrac{-11}{11} = -1 \\ s^2 = \dfrac{\Sigma d^2}{n} - \left(\dfrac{\Sigma d}{n}\right)^2 = \dfrac{69}{11} - 1 = 5.27 \Rightarrow s = 2.296 \end{array} \right\}$$

Step IV $|t_{cal}| < t_\alpha$ $\therefore$ Ho is accepted.

Step V The students have not benifited by coaching.

eg The mean height & the S.D. height of 8 randomly chosen soliders are 166.9cm & 8.29 cm respectively. The corresponding values of 6 randomly chosen sailors are 170.3 cm & 8.50cm respectively. Based on this data, can we conclude that the soliders are, in general, shorter than sailors?

Ans Given $\bar{x}_1 = 166.9$, $s_1 = 8.29$, $n_1 = 8$ → soliders

$\bar{x}_2 = 170.3$, $s_2 = 8.50$, $n = 6$ → sailors

Step I) $H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 < \bar{x}_2$

Step II) Let L.O.S. be 5%, One tailed test to be used.

Here $\nu = n_1 + n_2 - 2 = 12$

$$t_{5\%,(\nu=12)} = t_{10\%,(\nu=12)} = 1.78$$

Step III $t_{cal} = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\dfrac{n_1 s_1^2 + n_2 s_2^2}{n_1+n_2-2}\right)\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)}} = \dfrac{-3.4}{\sqrt{\left(\dfrac{983.29}{12}\right)\left(\dfrac{1}{8}+\dfrac{1}{6}\right)}} = -0.695$

Step IV $|t_{cal}| < t_\alpha$ $\therefore$ H_0 is accepted.

Step V We can't conclude that soldiers are, in general, shorter than sailors.

eg) Table shows the biological values of protein from cow's milk & buffalo's milk at a certain level. Examine if the average values of protein in the 2 samples significantly diff

Cow's milk	1.82	2.02	1.88	1.61	1.81	1.54
Buffalo's milk	2.00	1.83	1.86	2.03	2.19	1.88

Ans

Step I) $H_0 : \bar{x}_1 = \bar{x}_2$

$H_1 : \bar{x}_1 \neq \bar{x}_2$

Step II) Let L.O.S. be 5%. 2 tailed test be used.

$\nu = n_1 + n_2 - 2 = 10$

$\nu = 2n-2$ $\} n_1 = n_2 = n$

$t_{5\%, (\nu=10)}$ 2 tailed test $= 2.33$

Step III) $t_{cal} = \dfrac{\bar{x}_1 - \bar{x}_2}{\sqrt{\dfrac{s_1^2 + s_2^2}{n-1}}} = \dfrac{1.78 - 1.965}{\sqrt{\dfrac{0.0261 + 0.0154}{5}}} = \dfrac{-0.185}{\sqrt{0.0083}} = -2.03$

Step IV) $|t_{cal}| < t_\alpha$ $\therefore$ H_0 is accepted.

Step V) The difference betn the average values of protein in the 2 samples is not significant.

eg The I.Q.'s (intelligence quotients) of 16 students from one area of a city showed a mean of 107 with a S.D. of 10, while the I.Q.'s of 14 students from another area of city showed a mean of 112 with a S.D. of 8. Is there a significant difference betn the I.Q.'s of 2 group at $\alpha = 5\%$ L.O.S.

Soln

Step I $H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 \neq \mu_2$

Step II Let L.O.S be 5%. $\nu = n_1 + n_2 - 2 = 16 + 14 - 2 = 28$

Two tailed test to be used.

$t_{5\%(\nu = 28)} = 2.05$

Step III) $$t_{cal} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \frac{112 - 107}{\sqrt{\left(\frac{16(10)^2 + 14(8)^2}{16 + 14 - 2}\right)\left(\frac{1}{16} + \frac{1}{14}\right)}}$$

$t_{cal} = 1.45$

Step IV $|t_{cal}| < t_\alpha$ ∴, H_0 is accepted.

Step V There is no significant difference betn the I.Q.'s of two groups at $\alpha = 5\%$ L.O.S.

χ^2-test (Chi-square test)

1) Fit a Poisson distribution for the following distribution & also test the goodness of fit.

x	0	1	2	3	4	5	Total
f	142	156	69	27	5	1	400

Ans <u>Step I)</u> H_0: The fit is good. (The poisson fit for the given dist. is satisfactory)

H_1: The fit is not

<u>Step II)</u> mean $= m = \dfrac{\sum fx}{\sum f} = \dfrac{400}{400} = 1$, p.d.f $f(x) = \dfrac{e^{-m} m^x}{x!}$

Expected frequencies $\cong N \times f(x)$, $0 \le x \le 5$

<u>NOTE</u> ~~If Ei < 10~~

$N \times f(x)$ =	147.15	147.15	73.58	24.53	6.13	1.23
E =	147	147	74	25	6	1

<u>NOTE</u> Expected freq. in every class must be greater than or equal to 10.

$|E_i| \geq 10 \ \forall i$

$\therefore$ The last 3 classes are combined into one.

After regrouping,

O_i	142	156	64	33
E_i	147	147	74	32

<u>Step III)</u> $\chi^2_{cal} = \sum \left[\dfrac{(O_i - E_i)^2}{E_i}\right] = \dfrac{(5)^2}{147} + \dfrac{(9)^2}{147} + \dfrac{5^2}{74} + \dfrac{1^2}{32}$

$= 1.09$

<u>Step IV)</u> Let L.O.S. be 5%, $\nu = n-2 = 4-2 = 2$

$\chi^2_{5\%(\nu=2)} = 5.99$

<u>Step V)</u> $\chi^2_{cal} < \chi^2_{\alpha}$ $\Rightarrow$ Ho is accepted

i.e. The fit is good.

eg. Use the chi-sq. test to determine the goodness of fit of the data.

No. of Heads	0	1	2	3	4	5	Total
No. of Tosses	38	144	342	287	164	25	1000

Ans Step I) H_0 : The fit is good
H_1 : The fit is not good

Step II) (To find Expected frequencies)

$$m = \frac{\Sigma f x}{\Sigma f} = \frac{2470}{1000} = 2.47$$

But $m = np \Rightarrow p = \frac{2.47}{5} = 0.494$, $q = 0.506$.

$f(x) = {}^nC_x \, p^x q^{n-x}$.

$N \times f(x)$	33.2	161.9	316.2	308.7	150.7	29.4
					151	29
(E_i) Exp. freq.	33.	162	316	309		

Step III) $\chi^2_{cal} = \sum \frac{(O_i - E_i)^2}{E_i} = 7.54$

Step IV) $\nu = n - 2 = 6 - 2 = 4$, L.O.S. = 5%

$\chi^2_{5\%(\nu=4)} = 9.488$

Step V) $\chi^2_{cal} < \chi^2_{\alpha}$ $\therefore$ H_0 is accepted.
(i.e the fit is good)

3) The following table shows the observed & expected frequencies in tossing a die 120 times. Test the hypothesis that the die is fair using significance level of 5%.

Face	1	2	3	4	5	6
Observed freq.	25	17	15	23	24	16

Ans ~~Eq~~

Step I) H_0: The die is fair

H_1: The die is not fair

Step II) Expected frequencies $= \frac{120}{6} = 20$

$\therefore$

O_i	25	17	15	23	24	16
E_i	20	20	20	20	20	20

Step III) $\chi^2_{cal} = \sum \frac{(O_i - E_i)^2}{E_i} = 5$

Step IV) $\nu = n - 1 = 6 - 1 = 5$, $\alpha = 5\%$, $\chi^2_{5\%,(\nu=5)} = 11.07$

Step V) $\chi^2_{cal} < \chi^2_{\alpha}$ $\therefore$ Ho is accepted (The die is fair)

4) A random number table of 250 digits showed the following distribution of the digits 0, 1, 2, --- 9. Does the observed distribution differs significantly from the expected distribution?

Digit	0	1	2	3	4	5	6	7	8	9
obs. freq	17	31	29	18	14	20	35	30	20	36

Ans

Step I) Ho: Observed dist. does not differ from Expected dist.

H_1: Observed dist differs from Expected dist.

Step II) Expected frequencies $= \frac{250}{10} = 25$

Step III) $\chi^2_{cal} = \sum \frac{(O-E)^2}{E} = 23.3$

Step IV $\nu = n - 1 = 10 - 1 = 9$, $\chi^2_{5\%,(\nu=9)} = 21.66$

Step V $\chi^2_{cal} > \chi^2_{\alpha}$ $\therefore$ Ho is rejected (H_1 is accepted)

Eg. For the following data, test the hypothesis that the serum helps to cure the disease using L.O.S. of 5%.

	Recover	Do not Recover	Total
Gr A	75	25	100
Gr B	65	35	100
Total	140	60	200

Ans Step I) H_1: The serum helps to cure the disease

H_0: —"— does not help —"— (No effect)

Step II)

To calculate Exp. frequencies.

	Recover	Do not Recover	Total
Gr A	70	30	100
Gr B	70	30	100
Total	140	60	200

Step III). $\chi^2_{cal} = \sum \frac{(O-E)^2}{E} = \frac{(75-70)^2}{70} + \frac{(25-30)^2}{30} + \frac{(65-70)^2}{70} + \frac{(35-30)^2}{30}$

$= 2.38$

Step IV) $\nu = (h-1)(k-1)$ $\Big\}$ h = # of rows, k = # of columns.

$\nu = (2-1)(2-1) = 1$

LOS = 5%. $\chi^2_{5\%(\nu=1)} = 3.84$

Step V) $\chi^2_{cal} < \chi^2_{\alpha}$ $\therefore H_0$ is accepted

i.e. The serum does not help to cure the disease

Eg)

	Mr X	Mr Y	Mr Z	Total
Passed	50	47	56	153
Failed	5	14	8	27
Total	55	61	64	180

The table shows the numbers of students passed & failed by 3 instructors Mr X ...

Test the hypothesis that the proportions of students failed by the 3 instructors are equal.

Ans step I) H_0: The proportions of students failed by 3 instructors are equal

H_1: are not not equal

Step II) To calculate Exp. frequencies.

% of the failed students $= \frac{27 \times 100}{180} = \frac{300}{20} = 15\%$

∴ % of the passed students = 85%.

Exp. freq	Mr X	Mr Y	Mr Z	Total
passed	85% of 55 46.75	85% of 61 51.85	85% of 64 54.40	153
failed	15% of 25 8.25	15% of 61 9.15	15% of 64 9.60	27
Total	55	61	64	180

Step III)

$$\chi^2_{cal} = \sum \frac{(O-E)^2}{E} = \frac{(50-46.75)^2}{50} + \cdots$$

$$= 4.84$$

Step IV) $\nu = (h-1)(k-1) = (2-1)(3-1) = 2$

L.O.S. = 5%, $\chi^2_{5\%, (\nu=2)} = \cancel{5.99}\ 5.99$

Step V) $\chi^2_{cal} < \chi^2_{\alpha}$ ∴ H_0 is accepted

i.e. The proportions of students failed by 3 instructors are eq

29) {NOTE Each freq is less than or equal to 10}

$$\chi^2_{cal} = \sum \frac{[|O-E| - 0.5]^2}{E}$$

Two batches of 12 animals, each are given test of inoculation. One batch was inoculated & the other was not. The numbers of dead & surving animals are given in the following table for both cases. Can the inoculation be regarded as effective against the disease at 5% L.O.S (Using Yates Correction)

Step I) H_0: There is no association betn inoculation & death (No effect of inoculation)

H_1: There is association betn — " —

Step II) To calculate expected frequencies.

Given

	Dead	Surviving	Total
Inoculated	2	10	12
Not inoculated	8	4	12
Total	10	14	24

Exp. freq.	Dead	Surviving	Total
Inoculated	$\frac{10\times12}{24}=5$	$\frac{14\times12}{24}=7$	12
Non-inoculated	$\frac{10\times12}{24}=5$	$\frac{14\times12}{24}=7$	12
Total	10	14	24

Step III) $\chi^2_{cal} = \sum \frac{\{|O-E|-0.5\}^2}{E}$

$$\chi^2_{cal} = \frac{(|2-7|-0.5)^2}{5} + \frac{(|10-7|-0.5)^2}{7} + \frac{(|8-5|-0.5)^2}{5} + \frac{(|4-7|-0.5)^2}{7}$$

$\chi^2_{cal} = 4.29$

Step IV $\nu = (h-1)(k-1) = (2-1)(2-1) = 1$

Let L.O.S. = 5%, $\chi^2_{5\%,(\nu=1)} = 3.841$

Step V $\chi^2_{cal} \nless \chi^2_{table}$ ∴ Ho is rejected

i.e. The inoculation is effective against the disease

eg) In Mendel's experiments, with peas, he observed 315 round & yellow, 108 round & green, 101 wrinkled & yellow, 32 wrinkled & green. According to his theory of heredity the numbers should be in the proportion 9:3:3 Is there any evidence to doubt his theory at L.O.S of 1% & L.O.S. of 5%.

Ans. H_0: The theory & experiment are in agreement
The total no. of peas = 315+108+101+32 = 556

∵ Expected nos are in the proportion 9 : 3 : 3 : 1 (9+3+3+1=16)

∴ We would expect

(i) $\frac{9}{16}(556) = 312.75$ round & yellow

(ii) $\frac{3}{16}(556) = 104.25$ wrinkled & yellow

(iii) $\frac{3}{16}(556) = 104.25$ round & green

(iv) $\frac{1}{16}(556) = 34.75$ wrinkled & green

$$\therefore \chi^2_{cal} = \frac{(315-312.75)^2}{312.75} + \frac{(108-104.25)^2}{104.25} + \frac{(101-104.25)^2}{104.25} + \frac{(32-34.75)^2}{34.75}$$

$\chi^2_{cal} = 0.470$

∵ There are 4 categories $k = 4$

∴ $\nu = k-1 = 3$

$\alpha = 0.01$ $\chi^2_{table} = 11.3$

$\alpha = 0.05$ $\chi^2_{table} = 7.81$

∴ $\chi^2_{cal} < \chi^2_{table}$ for $\alpha = 0.05$ & 0.01

∴ H_0 is accepted.