4. Sampling Distribution

Suppose we have different samples of size n drawn from the population and for each and every sample of size n compute the means. Obiously these means will not be same. If we group these charecteristics according to its frequency the frequency distribution so generated is called sampling distribulton of means.

Note : 1) The sampling distribution of the large samples is assumed to be normal distribution.

2) The standard diviation of sampling distribution is also called standard error.

case 1 : Sampling with replacement.

Here the items drawn are put back to the population before the next draw. If N is the size of the population and n is size of the sample then we have N^n number of samples. And $\mu_{\bar{x}}$ of the frequency distribution of the sample mean will be equal to the population mean μ. $(\mu_{\bar{x}} = \mu)$

The variance $\sigma_{\bar{x}}^2$ of the frequency distribution of the sample means will bo equal to σ^2/n where σ^2 is variace of population

i.e $\sigma_{\bar{x}}^2 = \dfrac{\sigma^2}{n}$.

case 2 : Sampling without replacement.

Here the items are drawn and are not put back to the population before the next draw. In this case $^N C_n$ samples can be collected

Mean : $\mu_{\bar{x}} = \mu$

Variance : $\sigma_{\bar{x}}^2 = \left(\dfrac{N-n}{N-1}\right) \dfrac{\sigma^2}{n}$

correcting factor $\left(\dfrac{N-n}{N-1}\right) \approx 1$ for large samples

(~~with replacement~~)

① A population consists of 5 numbers 2, 3, 6, 8, 11. Consider all possible samples of size 2 which can be drawn with replacement from this population. F

(i) Find mean and standard deviation of population

(ii) The mean and standard deviation of sampling distribution of means.

(iii) Considering samples without replacements find mean and standard deviation of sampling distribution of means.

Solⁿ : $N = 5$

$n = 2$

(i) population mean $(\mu) = 6$ $\quad [2+3+6+8+11/5 = 6]$

$S.D\ (\sigma) = 3.28$ $\quad \left[\sqrt{\dfrac{\Sigma (x_i - \mu)^2}{n}}\right]$

(ii) Sample with replacement are

(2,2) (2,3) (2,6) (2,8) (2,11)

(3,2) (3,3) (3,6) (3,8) (3,11)

(6,2) (6,3) (6,6) (6,8) (6,11)

(8,2) (8,3) (8,6) (8,8) (8,11)

(11,2) (11,3) (11,6) (11,8) (11,11)

Now individual means of each samples

2	2.5	4	5	6.5
2.5	3	4.5	5.5	7
4	4.5	6	7	8.5
5	5.5	7	8	9.5
6.5	7	8.5	9.5	11

$\bar{x}$:	2	2.5	3	4	4.5	5	5.5	6	6.5	7	8	8.5	9.5
f :	2	2	1	2	2	2	2	1	2	4	1	2	2

$$\mu_{\bar{x}} = \frac{\Sigma \bar{x} f}{\Sigma f} = 6$$

standard deviation $= \sigma_{\bar{x}} = 2.32$

$$\left\{ \sigma_{\bar{x}}^2 = \frac{\sigma_x^2}{2} \Rightarrow 5.38 = 5.38 \right\}$$

iii) without replacement samples are

$$N_{C_n} = 5_{C_2} = 10 \text{ samples}$$

(2, 3) (2, 6) (2, 8) (2, 11) (3, 6) (3, 8) (3, 11) (6, 8) (6, 11) (8, 11)

Now individual means are

2.5 ~~4~~ 5 6.5 4.5 5.5 7 7 8.5 9.5

$$\mu_{\bar{x}} = \mu = 6$$

$$\sigma^2_{\bar{x}} = \left(\frac{N-n}{N-1}\right)\frac{\sigma^2}{n} = \left(\frac{5-2}{5-1}\right)\left(\frac{10.76}{2}\right) = 4.035$$

(2) Certain tubes manufactured by a company have mean life time of 800 hrs and standard deviation 60 hrs. Find the probability that a sample of 16 tubes taken from the group will have a mean lifetime

(i) b/w 790 hrs to 810 hrs

(ii) Less than 785 hrs

(iii) more than 820 hrs

Given: for population $\mu_x = 800$ $\sigma_x = 60$ $n = 16$

converting population data to sample

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \qquad \left\{ \begin{array}{l} \mu_{\bar{x}} = \mu \\ \sigma^2_{\bar{x}} = \sigma^2/n \end{array} \right.$$

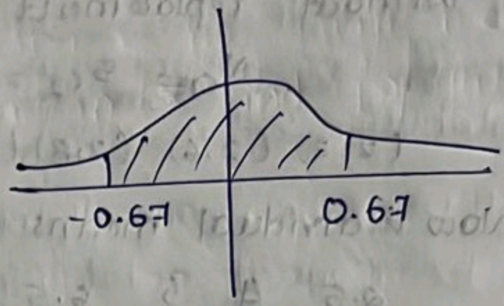
$$= \frac{x - \mu}{\left(\frac{\sigma}{\sqrt{n}}\right)}$$

$$z = \frac{\bar{x} - 800}{60/4} = \frac{\bar{x} - 800}{15}$$

(i) $P(790 < \bar{x} < 810)$

$$= P\left(\frac{790-800}{15} < \frac{\bar{x}-800}{15} < \frac{810-800}{16}\right)$$

$$= P(-0.67 < z < 0.67)$$



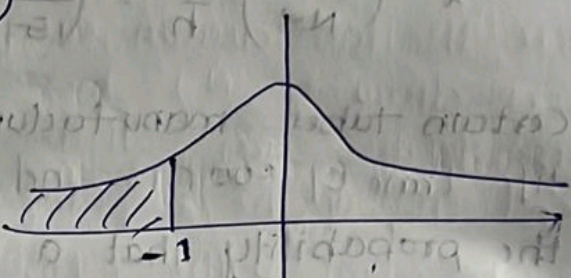
$$= 2A(0.67)$$

$$= 0.497$$

(ii) $P(\bar{x} < 785)$

$$= P\left(\frac{\bar{x}-800}{15} < \frac{785-800}{15}\right)$$

$$= P(z < -1)$$



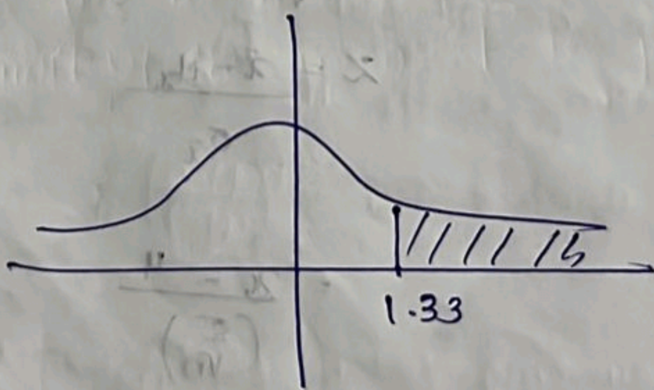
$$= 1/2 - A(-1)$$

$$= 0.158$$

(iii) $P(\bar{x} > 820)$

$$= P\left(\frac{\bar{x}-800}{15} > \frac{820-800}{15}\right)$$

$$= P(z > 1.33)$$



$$= 1/2 - A(1.33)$$

$$= 0.091$$

(5) If the mean population is 575 with standard deviation 8.3. How large a sample must be in order that there will be one chance in 100 that the mean sample is less than 572.

Population : $\mu = 575$ $\sigma = 8.3$ $n = ?$

$\bar{x} < 572$

For sample

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} \quad \left\{ \begin{array}{l} \mu_{\bar{x}} = \mu \\ \sigma_{\bar{x}} = \sigma/\sqrt{n} \end{array} \right.$$

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{\bar{x} - 575}{8.3/\sqrt{n}}$$

For $\bar{x} = 572$

$$z = \frac{572 - 575}{8.3/\sqrt{n}} = -0.36(\sqrt{n})$$

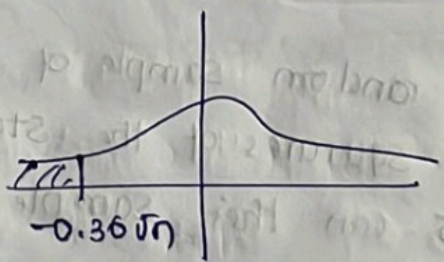
Given, $P(\bar{x} < 572) = 0.01$

$P(z < -0.36\sqrt{n}) = 0.01$

$1/2 - A(0.36\sqrt{n}) = 0.01$

$0.49 = A(0.36\sqrt{n})$

$A(2.35) = A(0.36\sqrt{n})$



By data $A(2.35) = 0.49$

$2.35 = 0.36\sqrt{n}$

$n = 42.6$

$$\boxed{n = 43}$$

5

* **Test of significance of small samples** 26/07/23

Student t-distribution (t-Test)

1) t-test of significance of mean of a random sample

Under the null hypothesis that there is no significance difference between the sample mean $\bar{x}$ and population mean μ.

The test statistic is given by

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

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

With degree of freedom $n-1$ (D.f)

Here the formula is used when the population standard deviation is not known

The confidence limit of population mean is given by

$$\bar{x} - t_\alpha (S/\sqrt{n}) < \mu < \bar{x} + t_\alpha (S/\sqrt{n})$$

1) A random sample of size 16 has 53 as mean. The sum of the squrares of the standard deviation from mean is 135. can this sample be regarded as taken from the population having 56 as the mean. Also obtain 95% and 99% confidence limit of the mean of the population.

$n = 16$

$\bar{x} = 53$

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

$\mu = 56$

Null hypothesis $H_0: \mu = 56$

alternate hypothesis $H_1 : \mu \neq 56$ (two tailed test)

test statistics

$$t \text{ or } z = \frac{\bar{x} - \mu}{(s/\sqrt{n})}$$

$$s^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{135}{15} = 9$$

$$\boxed{s = 3}$$

$$t \text{ or } z = \left(\frac{53-56}{3/\sqrt{16}}\right)$$

$$z = -4$$

$|z| = |-4| = 4 > 2.13$ at 5% level of significance and $n-1$ i-e 15 as degree to freedom

(check under 0.05 at 15 DoF freedom in given table)

$|z| = 4 > 2.13$ so H_0 is rejected. at 5% level of significance.

~~$|z| = 4 > 2.95$ at 1% level of significance and 15 as degree of freedom.~~

~~so H0 is reject at 1% significant level~~

The confidence limit at 95%.

$$\bar{x} - t_\alpha(s/\sqrt{n}) < \mu < \bar{x} + t_\alpha(s/\sqrt{n})$$

at 95% : ie $\bar{x} \pm t_{0.05}(s/\sqrt{n})$

$$53 \pm 2.13\left(3/4\right)$$

$$53 \pm 1.5975$$

$$\Rightarrow [54.59 \quad 54.25 \quad 51.4025]$$

at 99% : $\bar{x} \pm t_{0.01}(s/\sqrt{n})$

$\Rightarrow \quad 53 \pm 2.95\,(3/4)$

$\Rightarrow \quad 53 \pm 2.2125$

$[55.2125 \quad \cancel{2.2125} \quad 50.7875]$

2) 9 items of the sample have following values 45, 47, 50, 52, 48, 47, 49, 53, 51. Does the mean of these value differ significantly from assumed mean of 47.5.

$n = 9 \quad \mu = 47.5 \quad d.f = 9-1 = 8 \quad \bar{x} = 49.11$ (calculator)

$H_0 : \mu = 47.5$

$H_1 : \mu \neq 47.5$ (two tailed test)

test - statistics

t or $z = \dfrac{\bar{x} - \mu}{(s/\sqrt{n})}$

$s^2 = \dfrac{\Sigma(x-\bar{x})^2}{n-1}$

x	45	47	50	52	48	47	49	53	51
$(x-\bar{x})$	-4.1	-2.1	0.9	2.9	-1.1	-2.1	-0.1	3.9	1.9
$(x-\bar{x})^2$	16.81	4.41	0.81	8.41	1.21	4.41	0.01	15.21	3.61

$\Sigma(x-\bar{x})^2 = 54.89$

$s^2 = \dfrac{54.89}{8} = 6.861$

$\boxed{s = 2.619}$

assumed mean is nothing but population mean

$$t \text{ or } z = \frac{\bar{x}-\mu}{(s/\sqrt{n})} = \frac{49.1-47.5}{(2.619/\sqrt{8})}$$

$z = 1.832 < 2.31$ at 5% level of significance at d.f = 8 as degree of freedom.

$\therefore$ Ho is accepted.

3) A sample of 20 items has mean 42 units and standard deviation 5 units. Test the hypothesis that it is random sample from the normal population with mean 45 units.

$n = 20$ ~~s=5~~ S.D = 5 $H_0 : \mu = 45$

$\bar{x} = 42$ D.of.f = 19 $H_1 = \mu \neq 45$ (two tailed test)

$\mu = 45$

t- test

$t = \frac{\bar{x}-\mu}{s/\sqrt{n-1}}$ ⟶ (Here in formula we have to take instead of 'S' we take 's' because S.D is mentioned)

$$= \frac{42-45}{5/\sqrt{20-1}}$$

$t = -2.6$

$|t| = 2.6 \neq 2.09$ at 5% level of significance as 19 as degree of freedom.

Ho is rejected.

* **t-test for diff of means of two samples (Small)**

This test is used to check whether the two samples $x_1, x_2 \dots x_{n_1}$ and $y_1, y_2 \dots y_{n_2}$ which are of the sizes n_1 and n_2 respectively. and are drawn from two normal population with mean μ_1 and μ_2 under the assumption that population variance are equal (i-e $\sigma_1^2 = \sigma_2^2 = \sigma^2$).

Under the null hypothesis that the samples are drawn from the normal population with mean μ_2 and μ_2

$$t = \frac{\bar{x} - \bar{y}}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where $\bar{x}$ and $\bar{y}$ are means of the samples and degree of freedom is '$n_1 + n_2 - 2$.'

Note: 1) If the two samples standard deviation s_1 and s_2 are given $S^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$

2) If S_1 and S_2 are not given

$$S^2 = \frac{\sum (x-\bar{x})^2 + \sum (y-\bar{y})^2}{n_1 + n_2 - 2}$$

1) Two samples of sodium vapour bulbs were tested for length of life and the following results were obtained.

	Size	Sample mean	Sample S.D
Type 1	8	1234 hrs	36 hrs
Type 2	7	1036 hrs	40 hrs

is the difference in the mean significant to generelise that type 1 is supirior to type 2. regarding length of life.

$n_1 = 8 \quad n_2 = 7$

$\bar{x} = 1234 \quad \bar{y} = 1036$

$S_1 = 36 \quad S_2 = 40$

$H_0 : \mu_1 = \mu_2$

$H_1 : \mu_1 > \mu_2$ (one tailed test)

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{1/n_1 + 1/n_2}}$$

$$s^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}$$

$s = 40.73$

$t = 9.392$

do not double it.

$t = 2 \times 9.39$

$t = 18.79$ (Here we are considering single tailed test and formula is on two tailed test so we have to double it)

$t = 18.79 > 1.77$ at 2α i.e $2 \times 0.05 = 0.1$ significance level at df 13.

$\therefore$ Ho is rejected.

2) Two horses A and B were tested according to time to run a particular race with the following results

Horse A	28	30	32	33	33	29	34
Horse B	29	30	30	24	27	29	

Test whether you can discriminate b/w two horses.

$u_1 = 7$ $\quad \bar{x} = 31.28$

$u_2 = 6$ $\quad \bar{y} = 28.16$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$ (two tailed test)

$$t = \frac{\bar{x} - \bar{y}}{s\sqrt{1/n_1 + 1/n_2}}$$

$$s^2 = \frac{\Sigma (x-\bar{x})^2 + \Sigma (y-\bar{y})^2}{n_1 + n_2 - 2}$$

x	28	30	32	33	33	29	34
$(x-\bar{x})$	-3.28	-1.28	0.72	1.72	1.72	-2.28	2.72
$(x-\bar{x})^2$	10.75	1.63	0.51	2.95	2.95	5.19	7.39

$$\Sigma (x-\bar{x})^2 = 31.37$$

y	29	30	30	24	27	29
$(y-\bar{y})$	0.84	1.84	1.84	-4.16	-1.16	0.84
$(y-\bar{y})^2$	0.70	3.38	3.38	17.3	1.34	0.70

$$\Sigma (y-\bar{y})^2 = 26.8$$

$$s^2 = \frac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1+n_2-2}$$

$$= \frac{31.37 + 26.8}{11}$$

$$s = 2.29$$

$$t = \frac{\bar{x}-\bar{y}}{s\sqrt{1/n_1 + 1/n_2}} = \frac{31.28 - 28.16}{2.29\sqrt{1/7 + 1/6}}$$

$t = 2.44 > 2.20$ at 5% significance level of degree of freedom 11

$\therefore$ Ho is rejected.

Chi-square Test

If O_i for $(i = 1, 2, 3 \ldots n)$ is a set of experimental frequencies or observed and e_i $(i = 1, 2, 3 \ldots n)$ is the corresponding set of expected frequencies or theoretical then chi-square is defined as (χ^2)

$$\chi^2 = \sum_{i=1}^{n} \left[\frac{(O_i - E_i)^2}{E_i} \right]$$

where $\Sigma O_i = \Sigma E_i = N$ $\quad (N = \Sigma f_i)$

and the degree of freedom (d.f) is ~~N-1~~ $n-1$

Ex ①: The theory predicts that the proportion of beans in four groups G_1, G_2, G_3 and G_4 should be in the ratio 9:3:3:1. In experiment with 1600 beans the numbers in 4 groups were 882, 313, 287, 118. Does the experimental result support the theory?

Soln:

Experimental data

O_i : 882 313 287 118

Theoretical data

→ We have ratio of data

i.e 9:3:3:1

$$E(G_1) = \frac{9}{16} \times 1600 = 900 \quad \left[\frac{\text{no. of part taken}}{\text{total parts}} \times \text{total items} \right]$$

$E(G_2) = 300$

$E(G_3) = 300$

$E(G_3) = 100$

E_i : 900 300 300 100

Ho: There is no significance diff b/n observed and theoretical values

H_1: There is diff b/n O_i and E_i

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

$$= \left[\frac{(882-900)^2}{900} + \frac{(313-300)^2}{300} + \frac{(287-300)^2}{300} + \frac{(118-100)^2}{100}\right]$$

$$= \frac{(-18)^2}{900} + \frac{13^2}{300} + \frac{(-13)^2}{300} + \frac{18^2}{100}$$

$\chi^2 = 4.7266 < 7.820$ $[\chi^2_{0.05} = 7.820]$ for $n-1$ ie $4-1 = 3$

d.f = 3 at 5% level of significance

Ho is accepted.

② A die is thrown 276 times and the results of these are given below.

No. appeared on die	1	2	3	4	5	6
frequency	40	32	29	59	57	59

Test whether the die is biased or not.

$$E_i = \frac{276}{6} = 46 \left[\frac{\text{total no. of trials}}{\text{no. of options}}\right]$$

Each E_i has value 46.

O_i :	40	32	29	59	57	59
E_i =	46	46	46	46	46	46

H_0 : Die is biased

H_1 : Dis is unbiased

$$\chi^2 = \Sigma\left[\frac{(O_i - E_i)^2}{E_1}\right]$$

$$= \frac{(40-46)^2 + (32-46)^2 + (29-46)^2 + (59-46)^2 + (57-46)^2 + (59-46)^2}{46}$$

$$= \frac{36 + (-14)^2 + (-17)^2 + (13)^2 + 11^2 + 13^2}{46}$$

$$= 980/46$$

$\chi^2 = 21.30 > 11.07$ at 5% significance level at degree of freedom 5.

$\chi^2_{0.05} = 11.07$

H_0 is rejected

③ Records taken of the no. of male and female births in 800 families having 4 childern are as fallows

no. of male birth	0	1	2	3	4
no. of female birth	4	3	2	1	0
No. of families	32	178	290	236	64

Test whether the data are consistant with the hypothesis that the binomial law hols and the chance of male birth is equal to that of female birth.

$n=4 \quad N=800 \quad p=1/2 \quad q=1/2$ [p and q are probabilities of having equal male and female childs]

$p(x) = {}^nC_x\, p^x q^{n-x}$ (binomial distribution)

$E(x) = N\,P(x)$ [fitting to binomial distribution]

$E(x=0) = 800\left[{}^4C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{4-0}\right] = 50$

$E(x=1) = 800\left[{}^4C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3\right] = 200$

$E(x=2) = 800\left[{}^4C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2\right] = 300$

$E(x=3) = 800\left[{}^4C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1\right] = 200$

$E(x=4) = 800\left[{}^4C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0\right] = \underline{50}$

800

E_i : 50 200 300 200 50

O_i = 32 178 290 236 64

H_0 : male & female births are equal

H_1 : births are not equal

$$\chi^2 = \sum_{i=1}^{n}\left[\frac{\cancel{E_i}-(O_i - E_i)^2}{E_i}\right]$$

$$= \frac{(50-32)^2}{32} + \frac{(200-178)^2}{178} + \frac{(300-290)^2}{290}$$

$$+ \frac{(200-236)^2}{236} + \frac{(50-64)^2}{64}$$

$= 19.63 > 9.49$ at 5% level significance level at degree of freedom

H_0 is rejected

4) Fit a possion distribution to the following data and test the goodness of fit.

x :	0	1	2	3	4
f :	109	65	22	3	1

$$P(x) = \frac{e^{-m} m^x}{x!} \qquad m = 0.61$$

$E(x) = N P(x)$ [fitting into poissons distribution]

$$= N\left[\frac{e^{-m} m^x}{x!}\right]$$

$$E(0) = 200\left[\frac{e^{-0.61}(0.61)^0}{0!}\right] = 108.67 = 109$$

$$E(01) = 200\left[\frac{e^{-0.61}(0.61)^1}{1!}\right] = 66.28 = 66$$

$$E(02) = 200\left[\frac{e^{-0.61}(0.61)^2}{2!}\right] = 20.21 = 20$$

$$E(3) = 200\left[\frac{e^{-0.61}(0.61)^3}{3!}\right] = 4.11 = 4$$

$$E(4) = 200\left[\frac{e^{-0.61}(0.61)^4}{4!}\right] = 0.62 = 1$$

200

x	O_i	E_i	$(O_i - E_i)^2/E_i$
0	109	109	0
1	65	66	0.015
2	22	20	0.2
3	4 $\left\{3\right.$	5 $\left\{4\right.$	0.2
4	1	1	

$\Sigma (O_i - E_i)^2/E_i = 0.415$

write hypothesis

whenever the values of frequency are less than 5 we are going to club it up.

$$\chi^2 = \Sigma \frac{(Oi - Ei)^2}{Ei} = 0.415 < 5.990$$ at 5% level of significance at degree of freedom 2. (n-2)

$$\chi^2_{0.05} = 5.990$$

Here 3 degrees are going off one is because of defination, one for clubing up, and one is because we have taken mean from given data.

F - Test (Fisher Test)

To test whether the samples are drawn from the same population or from two population with the same variance. (σ^2). The null hypothesis set up here is Ho is

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$

The test for significance of diff of variance is

$$F = \frac{S_1^2}{S_2^2} \quad [S_1^2 > S_2^2] \text{ if this is not happening the swap.}$$

S_1

If the calculated value of F for $n_1 - 1$ and $n_2 - 1$ degree of freedom we can conclude that the ratio is significant at 0.05 level of significance.

① In two independent samples of sizes 8 and 10. The sum of squares deviations of sample values from the respective sample mean were 84.4 and 102.6. Test whether the diff of variance of population is significant or not.

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma^2$

$H_1: \sigma_1^2 \neq \sigma_2^2$

$n_1 = 8 \qquad n_2 = 10$

$\Sigma(x - \bar{x})^2 = 84.4$

$\Sigma(y - \bar{y})^2 = 102.6$

$$S_1^2 = \frac{\Sigma(x - \bar{x})^2}{n_1 - 1} = 12.05$$

$$S_2^2 = \frac{\Sigma(y - \bar{y})^2}{n_2 - 1} = 11.4$$

Now F distribution

$$F = \frac{S_1^2}{S_2^2} = \frac{12.05}{11.4} = 1.057$$

$F = 1.057 < 3.29$ at 5% level of significance at 7 and 9 are degree of freedom.

denominator goes in rows
numerator goes in colums

H_0 is accepted.

2) Two independent samples of sizes 7 and 6 had the following values

Sample A	28	30	32	33	31	29	34
Sample B	29	30	30	24	27	28	

Examine whether the samples are drawn from normal populations having same variance.

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$\bar{x} = 31$

$\bar{y} = 28$

$F = \frac{S_1^2}{S_2^2} \quad (S_1^2 > S_2^2)$

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

$x-\bar{x}$	-3	-1	1	2	0	-2	3
$(x-\bar{x})^2$	9	1	1	4	0	4	9

$\Sigma(x-\bar{x})^2 = 28$

$y-\bar{y}$	+1	+2	+2	-4	-1	0
$(y-\bar{y})^2$	1	4	4	16	1	0

$\Sigma(y-\bar{y})^2 = 26$

$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{19}{7-1} = \frac{28}{6} = 4.66$

$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n-1} = \frac{26}{5} = 5.2$

$$F = \frac{s_1^2}{s_2^2} \quad (s_1^2 < s_2^2)$$

so swap values $\because$ s_1^2 should be $> s_2^2$

$$F = \frac{s_2^2}{s_1^2} = \frac{5.2}{4.66} = 1.11$$

$F = 1.11 < 4.39$ at 5% ~~degree~~ significance level

at degree of freedoms 6 and 5 respectively

Here we have swapped values so denominator takes degree of freedom 6 and numerator 5.

3) Two independent samples are drawn from two normal populations are as follows

A	17	27	18	25	27	29	13	17
B	16	16	20	27	26	25	21	

Test whether the samples are from the same normal population.

In this problem we have 2 tests

(i) Equality of population variance by applying F-test.

(ii) Equality of means by applying t-test.

$H_0 : \sigma_1^2 = \sigma_2^2$

$H_1 : \sigma_1^2 \neq \sigma_2^2$

$$F = \frac{s_1^2}{s_2^2}$$

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

$\bar{x} = 21.625$ $\bar{y} = 21.571$ n_1

$x-\bar{x}$	-4.625	5.375	-3.625	3.375	5.375	7.375	-8.625
$(x-\bar{x})^2$	21.39	28.89	13.14	11.39	28.89	54.39	74.39

-4.625

21.39

$$\Sigma (x-\bar{x})^2 = 253.87$$

$$S_1^2 = \frac{\Sigma(x-\bar{x})^2}{n-1} = \frac{253.87}{7} = 36.26$$

$y-\bar{y}$	-5.571	-5.571	-1.571	5.429	4.429	3.429	-0.571
$(y-\bar{y})^2$	31.03	31.03	2.46	29.47	19.61	11.75	0.32

$$\Sigma (y-\bar{y})^2 = 125.67$$

$$S_2^2 = \frac{\Sigma(y-\bar{y})^2}{n-1} = \frac{125.67}{6} = 20.945$$

$$F = \frac{S_1^2}{S_2^2} \quad (S_1^2 > S_2^2)$$

$$= \frac{36.26}{20.945}$$

$F = 1.73 < 4.21$ at 5 % level of significance at degree of freedom 7 and 6 respectively

$\therefore$ Ho is accepted

Both variance are same.

To check whether data is from same population we have to do t-test to check whether mean are equal.

$n_1 = 8 \qquad n_2 = 7$

$$S^2 = \frac{\Sigma(x-\bar{x})^2 + \Sigma(y-\bar{y})^2}{n_1+n_2-2} = \frac{379.54}{13}$$

$S^2 = 29.19$

$S = 5.4$

$$t = \frac{\bar{x}-\bar{y}}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} = \frac{21.625 - 21.571}{5.4\sqrt{1/8 + 1/7}}$$

$= 0.0193 < 2.16$ at 5% level of significance at degree of freedom n_1+n_2-2 i.e 13

$\therefore$ samples are from same population

Practic t test

* Two types of batteries are tested for theire length of life and the following results were obtained

Battery A $\quad n_1 = 10 \quad \bar{x}_1 = 500$ hrs $\quad S_1^2 = 100$

Battery B $\quad n_2 = 10 \quad \bar{x}_2 = 500$ hrs $\quad S_2^2 = 121$

Compute students t-test and test whether there is significance difference in the two means.

$H_0 = H_1 = H_2$

$H_1 = H_1 \neq H_2$

$$t = \frac{\bar{x}-\bar{y}}{S\sqrt{1/n_1+1/n_2}} = \frac{\bar{x}-x_2^2}{S\sqrt{1/n_1+1/n_2}}$$

$$S^2 = \frac{n_1S_1^2 + n_2S_2^2}{n_1+n_2-2} = \frac{10(100)+10(121)}{18} = 122.77$$

$S = 11.08$

$$t = \frac{560-500}{122.77\sqrt{1/10 + 1/10}}$$

t = ~~1.091~~ $12.108 > 2.10$ at 5% level. of significance at degree of freedom n_1+n_2-2 i.e 18.

Ho is rejected.

* A machine is expected to produce nales of length 3 inches. A random sample of 25 nales gave an averange length of 3.1 inches with standard deviation 0.3. Can it be said that machine is producing nales as per specification. (Given $t(0.05) = 2.064$ for 24 the degree freedom)

$n = 25$ $\mu = 3$

$\bar{x} = 3.1$ $s = 0.3$

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n-1}} = \frac{3.1-3}{0.3/\sqrt{25-1}} = \frac{0.1}{1.5} = 0.068 < 2.064$$ at 5% level of significance at degree of freedom 24.

Because here we have considered ~~popul~~ Standard deviation so instead of n we have to take n-1

If they do not mention population mean (μ) assume it to be zero.