Snedecor's F-test

The test statistic is

$$F = \frac{S_1^2}{S_2^2}, \text{ if } S_1^2 > S_2^2$$

or

$$= \frac{S_2^2}{S_1^2} \text{ if } S_2^2 > S_1^2$$

where $S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}$ $\quad S_2^2 = \frac{n_2 s_2^2}{n_2 - 1}$

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

df is $(n_1 - 1, n_2 - 1)$ or $(n_2 - 1, n_1 - 1)$

Application

1) Testing significant difference between population variance

Partc 2) Testing whether the samples are drawn from the same population.

(X.)

Test 1: Testing difference between population variances

Q) In a sample of ~~tako~~ 10 observations the sum of squares of deviation of the sample value from the mean is 120. And then another sample of 12 observations was 314. Test whether there is a significant diff. b/w the samples at 5% LOS.

$n_1 = 10$, ~~$n_2 = 12$~~ $n_2 = 12$

$\Sigma(x_1 - \bar{x}_1)^2 = 120$ $\Sigma(x_2 - \bar{x}_2)^2 = 314$

$$S_1^2 = \frac{n_1 s_1^2}{n_1 - 1}$$

$$S_1^2 = \frac{\Sigma(x_1 - \bar{x}_1)^2}{n_1} = \frac{120}{10} = 12$$

$$S_2^2 = \frac{\Sigma(x_2 - \bar{x}_2)^2}{n_2} = \frac{314}{12} = 26.16$$

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

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

Test statistic

$$F = \frac{S_1^2}{S_2^2} \text{ or } \frac{S_2^2}{S_1^2}$$

$$= \frac{26.16}{12} = 2.18$$

LOS: 5%

df: $(n_2 - 1, n_1 - 1)$ = ~~(11, 12, 1)~~ (11, 9)

$F_{cal} = 2.18$

At 5% LOS, (11, 9) df $F_{tab} = 3.15$

$F_{cal} < F_{tab}$

H_0 is accepted

$\Rightarrow \sigma_1^2 = \sigma_2^2$

Q) Two random samples from a normal population gives (X) the following results

Sample I : 20 16 26 27 23 22 —

Sample II : 27 33 42 35 32 34 38

Test whether the samples are drawn from same population.

To test for sample came for ~~some~~ same population, we have to apply

i) t - test

ii) F - test

NH H_0 : The samples are drawn from same population.

i.e $\mu_1 = \mu_2$

$\sigma_1^2 = \sigma_2^2$

AH H_1 : $\mu_1 \neq \mu_2$, $\sigma_1^2 \neq \sigma_2^2$

t - test

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

x_1	x_2	$x_1 - \bar{x}_1$ $x_1 - 22.3$	$x_2 - \bar{x}_2$ $x_2 -$	$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$
20	27	-2.33	-7.43	5.42	55.20
16	33	-6.33	-1.43	40.06	2.044
26	42	3.67	7.57	13.96	57.30
27	35	4.67	0.57	21.8	0.32
23	32	0.67	-2.43	0.44	5.90
22	34	-0.33	-0.43	0.1	0.18
—	38	—	4.43	—	19.6
134	241			81.504	133.68

$$\bar{x}_1 = \frac{\Sigma x_1}{n_1} = \frac{134}{6} = 22.3$$

$$\bar{x}_2 = \frac{241}{7} = 34.43$$

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

$$= \frac{81.504 + 133.68}{6 + 7 - 2}$$

$$= 19.56$$

$$S = 4.42$$

$$t = \frac{22.3 - 34.43}{4.42\sqrt{\frac{1}{6}+\frac{1}{7}}}$$

$$\boxed{t = 4.95}$$

At 5% LOS, 11 df, $t_{tab} = 2.2$

$t_{cal} = 4.95$

$t_{cal} > t_{tab}$

F-test:

$$S_1^2 = \frac{\Sigma (x_1 - \bar{x}_1)^2}{n_1}, \quad S_2^2 = \frac{\Sigma (x_2 - \bar{x}_2)^2}{n_2}$$

$$= 19.11$$

$$= 13.384$$

$$F = \frac{S_2^2}{S_1^2} = \frac{19.11}{13.384} = 1.388$$

$F_{cal} = 1.388$

At (6, 5) df, $F_{tab} = 4.95$, $F_{cal} < F_{tab} \Rightarrow \sigma_1^2 = \sigma_2^2$

Since $\mu_1 \neq \mu_2$, the sample are not drawn from same population.

CT-2, 12 mark

Q) The nicotin contained in two random samples of tobacco are given below:

Sample I	21	24	25	26	27	-
Sample II	22	27	28	30	31	36

Can you say that the two samples came from same population?

$n_1 = 5, n_2 = 6$

$\bar{X}_1 = \frac{\Sigma x_1}{n_1}$ $\quad \bar{X}_2 = \frac{\Sigma x_2}{n_2}$

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

$S_1^2 = \frac{\Sigma (x_1 - \bar{x}_1)^2}{n_1}$ $\quad S_2^2 = \frac{\Sigma (x_2 - \bar{x}_2)^2}{n_2}$

x_1	x_2	$x_1 - \bar{x}_1$	$x_2 - \bar{x}_2$	$(x_1 - \bar{x}_1)^2$	$(x_2 - \bar{x}_2)^2$
21	22	-3.6	-7	12.96	49
24	27	-0.6	-2	0.36	4
25	28	0.4	-1	0.16	1
26	30	1.4	1	1.96	1
27	31	2.4	2	5.76	4
-	36	-	7	-	49
123	174			23.6	108

$\bar{X}_1 = \frac{123}{5} = 24.6$

$\bar{X}_2 = \frac{174}{6} = 29$ ~~83~~

$s^2 = \frac{\Sigma (x_1 - \bar{x}_1)^2 + \Sigma (x_2 - \bar{x}_2)^2}{} = 23.6 + 108$

$S_1^2 = \frac{\Sigma(x_1-\bar{x}_1)^2}{n_1} = \frac{23.6}{5} = 4.7$

$S_2^2 = \frac{\Sigma(x_2-\bar{x}_2)^2}{n_2} = \frac{108}{6} = 18$

NH H_0: The samples are drawn from same population

$\mu_1 = \mu_2$, $\sigma_1^2 = \sigma_2^2$

AH H_1: $\mu_1 \neq \mu_2$, $\sigma_1^2 \neq \sigma_2^2$

t-test

The test statistic is

$$t = \frac{\bar{x}_1 - \bar{x}_2}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

$$= \frac{24.6 - 29}{3.82\sqrt{\frac{1}{5}+\frac{1}{6}}} = -1.91$$

$|t| = 1.91$

$t_{cal} = 1.91$

At 5% LOS, 9 df,

$t_{tab} = 3.26$

$t_{cal} < t_{tab}$

$\Rightarrow \mu_1 = \mu_2$

F-test

The test statistic is

$$F = \frac{S_2^2}{S_1^2} = \frac{18}{4.7} = 3.67$$

$F_{cal} = 3.67$

For $(n_2-1, n_1-1) = (5, 4)$ df

At 5% LOS,

$F_{tab} = 6.26$

$F_{cal} < F_{tab}$

$\Rightarrow \sigma_1^2 = \sigma_2^2$

$\Rightarrow H_0$ is accepted

Both sample came from same population.

Chi - Square test (χ^2 - test)

The test statistic is

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

where O_i - observed frequency

E_i - Expected (or) Theoretical frequency.

Application:

1) Testing difference between observed and expected frequencies.
2) Testing goodness of fit
3) Testing independence of Attributes.

Condition for χ^2-test

1) $4 \le n \le 16$, n is no. of observation
2) Individual observed frequency should be ≥ 10
3) If any $O_i < 10$, the frequency can be combined with neighbouring frequency.

Test I: Testing diff b/w observed & expected frequencies

Q) The following table gives the no. of accidents that occur during the various days of the week. Test whether the accidents are uniformly distributed over the week.

Day	Mon	Tue	Wed	Thurs	Fri	Satur
No. of Accidents	15	19	13	12	16	15

Expected frequency $= \dfrac{15+19+13+12+16+15}{6}$

$= \dfrac{90}{6} = 15$

NH H_0: The number of accidents are distributed uniformly over a week.

$O_i = E_i$

AH H_1: $O_i \neq E_i$

Test statistic: $\chi^2 = \sum_i \dfrac{(O_i - E_i)^2}{E_i}$

L.O.S = 5%

df = n − 1 = 6 − 1 = 5

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\dfrac{(O_i-E_i)^2}{E_i}$
15	15	0	0	0
19	15	4	16	1.06
13	15	−2	4	0.26
12	15	−3	9	0.6
16	15	1	1	0.06
15	15	0	0	0

$\chi^2 = 1.98$

$\chi^2_{cal} = 1.98$

At 5% LOS, 5df, $\chi^2_{tab} = 11.07$

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

H_0 is accepted

⇒ No. of accidents are distributed uniformly over a week.

Q) Theory predicts that the proportion of beans in 4 groups ~~is~~ A, B, C, D are in the ratio 9:3:3:1. In an experiment amount 1600 ~~days~~ beans, the numbers in the 4 groups where 882, ~~330~~ 313, 287 & 118. ~~Thus~~ Does the experiment support the theory?

Expected frequencies

$$E(A) = \frac{9}{16} \times 1600 = 900$$

$$E(B) = \frac{3}{16} \times 1600 = 300$$

$$E(C) = \frac{3}{16} \times 1600 = 300$$

$$E(D) = \frac{1}{16} \times 1600 = 100$$

Null Hypothesis H_0: The experiment results support the efficiency.

$$O_i = E_i$$

A.H H_1: There is a significant difference b/w expected and observed frequencies.

Test statistic: $\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i}$, $\sum_i O_i = \Sigma E_i$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i-E_i)^2}{E_i}$
882	900	-18	324	0.36
313	300	13	169	0.56
287	300	-13	169	0.56
118	100	18	324	3.24
1600	1600			$\chi^2 = 4.72$

$$\chi^2_{cal} = 4.72$$

$$df = n-1 = 4-1 = 3$$

At 5% LOS, 3df, $\chi^2_{tab} = 7.82$

conclusion: $\chi^2_{cal} < \chi^2_{tab}$

$\Rightarrow H_0$ is accepted.

$\Rightarrow$ The experimental results support the theory.

Test 2 Testing goodness of fit

(*) 12mark

Q) Fit a binomial distribution for the following data.

x :	0	1	2	3	4	5	6	Total
f :	5	15	28	12	7	8	5	80

Test the goodness of fit.

The Pmf of Binomial distribution is

$P(X=x) = nC_x \, p^x q^{n-x}, \quad x = 0, 1, 2 \dots n$

$n = 6$ Mean $= np$, Variance $= npq$

Mean $= np = \frac{\Sigma fx}{\Sigma f} = \frac{205}{80} = 2.563$

x	f	fx
0	5	0
1	15	15
2	28	56
3	12	36
4	7	28
5	8	40
6	5	30
	80	205

$np = 2.563$

$6p = 2.563$

$p = \frac{2.563}{6} = 0.427$

$q = 1 - p = 1 - 0.427$

$= 0.573$

$n = 6, \quad p = 0.427, \quad q = 0.573$

$P(X = x) = 6C_x (0.427)^x (0.573)^{6-x}, \quad x = 0, 1, 2, 3, 4, 5, 6$

Expected frequencies:

$f(x) = N \cdot P(x) \qquad N = \Sigma f$

$= 80 P(x)$

x	$P(X=x) = 6C_x (0.427)^x (0.573)^{6-x}$	Expected frequencies $N \cdot P(x)$
0	$6C_0 (0.427)^0 (0.573)^6 = 0.035$	2.8
1	$6C_1 (0.427)^1 (0.573)^5 = 0.158$	12.64
2	$6C_2 (0.427)^2 (0.573)^4 = 0.294$	23.5
3	$6C_3 (0.427)^3 (0.573)^3 = 0.29$	23.4
4	$6C_4 (0.427)^4 (0.573)^2 = 0.16$	13.09
5	$6C_5 (0.427)^5 (0.573)^1 = 0.048$	3.9
6	$6C_6 (0.427)^6 (0.57)^0 = 0.00606$	0.48
		$\Sigma E_i = 79.8 \simeq 80$

NH. H_0: The BD is the best fit

$O_i = E_i$

AH H_1: $O_i \neq E_i$

Test Statistic:

$$\chi^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}, \quad \Sigma O_i = \Sigma E_i$$

LOS 5%

Df for B.D = $n - 1$

χ^2 Table

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
20	15.44	4.56	20.79	1.346
28	23.5	4.5	20.25	0.867
12	23.4	-11.4	129.9	5.55
20	17.47	2.53	6.40	0.366
			χ^2	8.139

$\chi^2_{cal} = 8.139$

At 5% L.O.S, $n-1 = 4-1 = 3$ df.

$\chi^2_{tab} = 7.82$

Conclusion:

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

H_0 is rejected.

$\Rightarrow$ B.D is not the best fit.

Q) Fit a Poisson distribution to the following data & Test the goodness of fit.

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

Note: df for Bionimal distribution = $n-1$
df for Poisson distribution = $n-2$
df for Normal distribution = $n-3$

The Pmf of P.D

$$P(X=x) = \frac{e^{-\lambda}\lambda^x}{x!}, \quad x = 0, 1, 2 \ldots$$

$$\text{Mean } \lambda = \frac{\Sigma fx}{\Sigma f} =$$

x:	0	1	2	3	4	5	
f:	142	156	69	27	5	1	400
fx:	0	156	138	81	20	5	400

$$\lambda = 1$$

$$P(X = x) = \frac{e^{-\lambda} \lambda^x}{x!} = \frac{e^{-1}}{x!}, \quad x = 0, 1, 2 \ldots$$

Fitting P.D

x	$P(X = x)$	$N \cdot P(x)$
0	$e^{-1}/0! = 0.3678$	$147.12 \simeq 147$
1	$e^{-1} = 0.3678$	$147.12 \simeq 147$
2	$e^{-1}/2 = 0.183$	$73.2 \simeq 73$
3	$e^{-1}/6 = 0.06$	$24 \simeq 24$
4	$e^{-1}/24 = 0.015$	$6 \simeq 6$
5	$e^{-1}/100 = 0.003$	$1.2 \simeq 1$
		$\Sigma E_i = 398.64$

χ^2 Table

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
142	147	-5	25	0.170
156	147	9	81	0.55
69	73	-4	16	0.219
33	31	2	4	0.12
				$\chi^2 = 1.05$

NH. H_0: The P.D is the best fit. i.e $O_i = E_i$

AH H_1: $O_i \neq E_i$ ~~Fet~~

Test statistic: $\chi^2 = \frac{\Sigma (O_i - E_i)^2}{E_i}$

$$\chi^2_{cal} = 1.05$$

for P.D, $df = n-2 = 4-2 = 2$

At 5% LOS, for 2df,

$$\chi^2_{tab} = 5.99$$

Conclusion:

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

H_0 is accepted.

$\Rightarrow$ P.D is the best fit

Test 3:

Testing independence of Attributes

Let A & B be two characteristics (attributes) then

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

i.e 2 × 2 contingency Table

X\Y	A_1	A_2	Total
B_1	a	b	a+b
B_2	c	d	c+d
Total	a+c	b+d	$N = a+b+c+d$

For 2×2 contingency table,

$$\chi^2 = \frac{N(ad-bc)^2}{(a+c)(b+d)(a+b)(c+d)}$$

Expected frequencies

$$E(a) = \frac{(a+b)(a+c)}{N}$$

$$E(b) = \frac{(a+b)(b+d)}{N}$$

$$E(c) = \frac{(a+b)(a+c)}{N} \qquad \frac{(c+d)(a+c)}{N}$$

$$E(d) = \frac{(b+d)(c+d)}{N}$$

Scanned with CamScanner

Problems on Independence of attributes

Q) The following data are collection on two characters

Attributes	Smokers	Non-Smokers	Total
Literates	83	57	140
Illiterates	45	68	113
Total	128	125	253

Based on this, can you say that there is a relation b/w smoking & literacy?

NH H_0: There is no relation b/w smoking & literacy.

AH H_1: There is a relation b/w smoking & literacy.

LOS: 5%

Df: $(r-1)(s-1)$, r - no. of rows, s - no. of columns

$r=2, s=2$

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

Test statistics

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

Expected frequency

$$E(83) = \frac{140 \times 128}{253} = 70.83 \simeq 71$$

$$E(57) = \frac{140 \times 125}{253} = 69.17 \simeq 69$$

$$E(45) = \frac{113 \times 128}{253} = 57.16 \simeq 57$$

$$E(68) = \frac{113 \times 125}{253} = 55.83 \simeq 56$$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
83	71	12	144	2.02
57	69	-12	144	2.08
45	57	-12	144	2.52
68	56	12	144	2.57
253	253			$\chi^2 = 9.19$

$\chi^2_{cal} = 9.19$

$\chi^2_{tab} = 3.84$ at 5% LOS for 1 df

$\chi^2_{cal} > \chi^2_{tab} \Rightarrow H_0$ is rejected.

conclusion: There is a solution smoking & literacy.

(or)

2nd only method for 2x2 table

$$\chi^2 = \frac{N(ad-bc)^2}{(a+b)(a+c)(b+c)(b+d)}$$

$$= \frac{253((83)(68)-(45)(57))^2}{(140)(113)(128)(125)}$$

$= 9.47$

) The company has to choose 3 pension plans and the opinions of random sample of 500 employees are showed in the following ways.

Job classification	Pension Plans 1	2	3
Salared workers	160	140	40
Hourly workers	40	60	60

Determine whether the preferences of plan is independent of Job classification.

				Total
	160	140	40	340
	40	60	60	160
Total	200	200	100	500

Expected frequency

$$E(160) = \frac{340 \times 200}{500} = \cancel{680}\ 136$$

$$E(140) = \frac{200 \times 340}{500} = 136$$

$$E(40) = \frac{340 \times 100}{500} = 68$$

$$E(40) = \frac{200 \times 160}{500} = 64$$

$$E(60) = \frac{160 \times 200}{500} = 64$$

$$E(60) = \frac{160 \times 100}{500} = 32$$

NH H_0: There is no preference of plan over job classification.

AH H_1: There is a preference of plan over job classification.

Test statistic : $\chi^2 = \sum_i \frac{(O_i - E_i)^2}{E_i}$

LOS : 5%.

Df : $(r-1)(s-1)$

$r=2,\ s=3 \Rightarrow (r-1)(s-1) = (2-1)(3-1) = 2$

$df = 2$

O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
160	136	24	576	4.23
140	136	4	16	0.117
40	68	-28	784	11.52
40	64	-24	576	9
60	64	-4	16	0.25
60	32	28	784	24.5
				49.617

$\chi^2_{cal} = 5.991$

At 5% LOS, 2df, $\chi^2_{tab} = 3.43$

$\chi^2_{cal} > \chi^2_{tab} \Rightarrow H_0$ i.e. ~~accepo~~ rejected.

Large samples (n > 30)

1. Testing for single mean
2. Testing difference b/w mean
3. Testing for single proportion
4. Testing for difference b/w ~~th~~ two ~~pa~~ proportion

Test 1, (single mean)

Test statistic:

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

$\bar{x}$ - sample mean
μ - population mean
σ - population S.D
n - sample size

Test-2 (Difference b/w mean)

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{(or)}$$

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Critical values for Large Samples (z- test)

Nature of test	5% LOS	1% LOS
Two tailed test) $(\mu_1 \neq \mu_2)$	1.96	2.58
One tailed test (Right tailed $\mu_1 > \mu_2$ or Left tailed $\mu_1 < \mu_2$)	1.65	2.33

Problems on Test 1 (single mean)

Q) A sample of 100 students is taken from a large population with mean ~~high~~ height of 160 cm. Can it be reasonably regarded that the population the mean height is 165 cm with S.D = 10 cm

Since $n = 100 > 30$, the given sample is large sample.

NH H_0: $\mu = 165$ cm

AH H_1: $\mu \neq 165$ cm

Test Statistic:

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

$\bar{x} = 160$, $\mu = 165$, $\sigma = 10$, $n = 100$

$$z = \frac{160 - 165}{10/\sqrt{100}} = \frac{-5}{10/10} = -5$$

$|z| = 5$

$Z_{cal} = 5$

LOS: 5%

$Z_{tab} = 1.96$

$Z_{cal} > Z_{tab} \Rightarrow$ Reject H_0

$\Rightarrow \mu \neq 165$ cm

Q) A random sample of 200 tins of Coconut oil, gave an average weight 4.95 kg, The S.D of 0.21 kg. Due you accept the hypothesis of new weight, 5 kg per tin at 5% LOS.

Since $n = 200 > 30$, the given is large sample

NH H_0 : $\mu = 5$ kg.

AH H_1 : $\mu \neq 5$ kg.

Test statistic:

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

$\bar{x} = 4.95$, $\mu = 5$

$\sigma = 0.21$ $n = 200$

$$z = \frac{4.95 - 5}{0.21/\sqrt{200}}$$

$= -3.367$

$|z| = 3.367$

$Z_{cal} = 3.367$

At 5% LOS, $Z_{tab} = 1.96$

$Z_{cal} > Z_{tab}$

Reject H_0

$\mu \neq 5$ kg //

Test-2 (Difference of Mean)

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{where } \sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2}$$

Q) Test the significance of difference between the mean of samples drawn from two normal dis. with same S.D from the following data

	Size	Mean	S.D
Sample I	100	61	4
Sample II	200	63	6

$n_1 = 100 \qquad n_2 = 200$

$\bar{x}_1 = 61 \qquad \bar{x}_2 = 63$

$s_1 = 4 \qquad s_2 = 6$

$$\sigma^2 = \frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2} = \frac{(100)(16) + (200)(36)}{100 + 200}$$

$$= \frac{1600 + 7200}{300} = \frac{88\cancel{00}}{3\cancel{00}}$$

$$\sigma^2 = 29.3$$

$$\sigma = 5.42$$

NH H_0: $\mu_1 = \mu_2$

There is no significant difference between population mean.

AH H_1: $\mu_1 \neq \mu_2$

Test statistic:

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{61 - 63}{5.42\sqrt{\frac{1}{100} + \frac{1}{200}}}$$

$$= -2.13$$

$|Z| = 2.13 \qquad Z_{cal} = 2.13$

At 5% LOS, $Z_{tab} = 1.96$

$Z_{cal} > Z_{tab}$

H_0 is rejected.

$\Rightarrow \mu_1 \neq \mu_2$

At 1% LOS, $Z_{tab} = 2.58$

$Z_{cal} < Z_{tab}$, Accept H_0

$\mu_1 = \mu_2$

Q) A simple sample of 6400 english men have a mean height 170 cm. with S.D = 6.4 cm. While a simple sample of 1600 americans has a mean height 172 cm with S.D = 6.3 cm. Do the data indicate that the americans are taller than the english men.

NH H_0: $\mu_1 = \mu_2$

There is no significant difference between mean height of english men & american.

AH H_0: $\mu_1 < \mu_2$ (left tailed test)

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\bar{x}_1 = 170$ $\quad \bar{x}_2 = 172$

$n_1 = 6400$ $\quad n_2 = 1600$

$S_1 = 6.4$ $\quad S_2 = 6.3$

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

$$= \frac{(6400)(40.96) + (1600)(39.69)}{8000}$$

$$= \frac{262144 + 63504}{8000} = 40.706$$

$\sigma = 6.4$

$$z = \frac{170 - 172}{6.4\sqrt{\frac{1}{6400} + \frac{1}{1600}}} = -11.2$$

$|z|_{cal} = 11.2$

At 5% LOS, for one tailed test, $Z_{tab} = 1.65$

$Z_{cal} > Z_{tab} \Rightarrow$ Reject $H_0 \Rightarrow$ Accept H_1

$\mu_1 < \mu_2$

$\Rightarrow$ Americans are taller than Englishmen//

Test 3: Test for single proportion.

Test statistic is

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

where p - sample proportion

P - population proportion

$Q = 1 - P$, $q = 1 - p$

NOTE:

1. The confidence (or) Fiducial limits (95%) for sample proportion at 5% L.O.S is

$$P \pm z_{0.05} \sqrt{\frac{PQ}{n}}$$

99% confidence limits are

$$P \pm z_{0.01} \sqrt{\frac{PQ}{n}}$$

2. Standard error is the standard deviation of respective test statistic.

Q) A coin was tossed 400 times and the head appeared 216 times. The Test the hypothesis that the coil is unbiased.

NH H_0: The coin is unbiased

$$P = \frac{1}{2} = 0.5$$

AH H_1: The coin is biased

$$P \neq 0.5$$

Test Statistic

$$z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$P = 0.5,\ Q = 1 - P = 0.5$

$$p = \frac{216}{400} = 0.54$$

$$Z = \frac{0.54 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{400}}} = 1.6$$

$$|Z| = 1.6$$

At 5% LOS, $Z_{tab} = 1.96$

$Z_{cal} < Z_{tab}$ => H_0 is accepted

=> The coin is unbiased.

Q) A producer confesses that 22% of items ~~that~~ manufactured by the company will be defective. To test his claim, random sample of 80 items are selected and known to contain 13 defectives. Test the validity of producers claim at 1% LOS

NH H_0: $P = \frac{22}{100} = 0.22$

AH H_1: $P \neq 0.22$

Test statistic:

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}}$$

$$p = \frac{13}{80} = 0.1625$$

$P = 0.22,\ Q = 1 - P = 0.78,\ n = 80$

Test statistic

$$Z = \frac{p - P}{\sqrt{\frac{PQ}{n}}} = \frac{0.16 - 0.22}{\sqrt{\frac{(0.22)(0.78)}{80}}} = -1.3$$

$$|Z| = 1.3$$

At 1% LOS, $Z_{tab} = 2.33$

$Z_{cal} < Z_{tab}$ => H_0 is accepted

=> $P = 0.22$

=> Producer's claim is valid for 1% LOS

Test 4: Testing for difference of proportion

Test statistic

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{when}$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}, \quad Q = 1 - P$$

(Population proportion are not given)

(or)

when population proportions are given then

$$z = \frac{p_1 - p_2}{\sqrt{\frac{P_1 Q_1}{n_1} + \frac{P_2 Q_2}{n_2}}} \qquad Q_1 = 1 - P_1 \qquad Q_2 = 1 - P_2$$

Q) Random samples of 400 men and 600 women were asked whether, they would like to have a flyover near the residence. 200 men and 325 women were in favour of proposal. Test the hypothesis that proportions of men and women in favour ofs proposal ~~of~~ are same at i) 5% LOS ii) at 1% LOS

$n_1 = 400 \qquad n_2 = 600$

$p_1 = \frac{200}{400} = 0.5 \qquad p_2 = \frac{325}{600} = 0.541$

NH H_0: The men and women were equal in favour of proposal.

$$P_1 = P_2$$

AH H_1: $P_1 \neq P_2$

Test statistic:

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{when}$$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}, \quad Q = 1 - P$$

$$= \frac{(400)(0.5) + (0.541)(600)}{400 + 600}$$

$$= \frac{200 + 325}{1000} = \frac{525}{1000} = 0.525$$

$$Q = 1 - 0.525 = 0.475$$

$$z = \frac{0.5 - 0.541}{\sqrt{(0.525)(0.475)\left(\frac{1}{400} + \frac{1}{600}\right)}}$$

$$= 0.625$$

$$|z|_{cal} = 0.625$$

$z_{0.05} = 1.96$, $z_{0.01} = 2.58$

At 5% and 1% LOS, $z_{cal} < z_{tab}$, Accept H_0.

$$\Rightarrow P_1 = P_2$$

Q) In a large city A, 20% of random sample of 900 school boys had a ~~slide~~ slight physical defect. In other large city B 18.5% of random sample of 1600 school boys had the same defect. Is the difference between the proportion significant.

NH H_0:

$n_1 = 900$ $\quad$ $n_2 = 1600$

$p_1 = 0.2$ $\quad$ $p_2 = 0.185$

$$P = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$$

$$= \frac{(900)(0.2) + (1600)(0.185)}{900 + 1600}$$

$$= \frac{180 + 296}{2500} = 0.1904$$

$$Q = 1 - P = 0.81$$

$$z = \frac{p_1 - p_2}{\sqrt{PQ\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = 0.915$$

At 5% LOS, $z_{tab} = 1.96$, $z_{cal} < z_{tab}$, Accept H_0 //