18/04/2022

STATISTICS EXAM

Q1 ANOVA - Analysis of Var.

$\Sigma V_1 = 500$, $\Sigma V_2 = 530$, $\Sigma V_3 = 440$, $n = 30$

$\Sigma V_1^2 = 26742$, $\Sigma V_2^2 = 28586$, $\Sigma V_3^2 = 20422$

Correction term : $C_x = \frac{\Sigma(x)^2}{n} = \frac{(500+530+440)^2}{30} = \frac{(1470)^2}{30}$

$$\Rightarrow \frac{2160900}{30} = 72030$$

a. Sum of squares total - SST

$$SST = \Sigma x^2 - cx = (26742 + 28586 + 20422) - 72030$$

$$\Rightarrow 75750 - 72030$$

$$SST \Rightarrow 3720$$

b. Sum of squares among grp - SSA

$$SSA = \left(\frac{\Sigma x^2}{n}\right) - C_x = \left(\frac{560^2}{10} + \frac{536^2}{10} + \frac{446^2}{10}\right) - 72030$$

$$\Rightarrow \left(\frac{250060}{10} + \frac{280900}{10} + \frac{193600}{10}\right) - 72030$$

$$\Rightarrow (25000 + 28096 + 19360) - 72030$$

$$\Rightarrow 72450 - 72030$$

$$SSA \Rightarrow 420$$

(c) Sum of square within the group.

$$SS_W = SST - SSA$$
$$= 3720 - 420$$
$$= 3300$$

(d) mean of Sum of Squ. among group.

$$\mu_{SS_A} = \frac{SS_A}{k-1} = \frac{420}{3-1} = \frac{420}{2}$$

k - Category of table

$$= 210$$

(e) mean of Sum of Squ. within group

$$\mu_{SS_W} = \frac{SS_W}{N-k} = \frac{3300}{30-3} \Rightarrow \frac{3300}{27} = 122.22$$

$$F\text{-Ratio} = \frac{\mu_{SS_A}}{\mu_{SS_W}} = \frac{210}{122.22} = 1.718$$

Comparison of F-Ratio Vs F-value @ 0.05 Sign. level

Source of Variance	df	SS	μ_{SS}	F-Ratio	F-Value
Among grp	2	420	210	1.718	3.354
within grp	27	3300	122.22		

$\Rightarrow 1.718 < 3.354 \Rightarrow$ F.Ration < F.value

We accept the null Hypothesis.

Q12) Determining ANOVA with STOCK MARKET EXAMPLE.

Financial(x), Energy (4), Utilities(z)

$\Sigma x = 92.68$, $\Sigma 4 = 110.77$, $\Sigma z = 71.3$, $n = 24$

$\Sigma x^2 = 1257.46$, $\Sigma 4^2 = 1699.56$, $\Sigma z^2 = 779.13$

Correction term: $C_x = \frac{\Sigma(x)^2}{n} = \frac{(92.68 + 110.77 + 71.3)^2}{24}$

$= \frac{(274.75)^2}{24} = \frac{75487.56}{24} = 3145.31$

Sum of squ. total = $SS_T = \Sigma x^2 - C_x$

$= (1257.46 + 1699.56 + 779.13) - 3145.31$

$= 3736.16 - 3145.31$

$SS_T = 590.85$

Sum of squ. among grp - $SS_A = \frac{(\Sigma x^2)}{n} - C_x$

$= \frac{(92.68^2 + 110.77^2 + 71.3^2)}{24} - 3145.31$

$= \frac{(8589.58 + 12269.99 + 5083.69)}{8} - 3145.31$

$= 337.89 + 54.24 + 211.82 - 3145.31$

$= 1088.96 - 3145.31$

$= -2056.35$

$= (1073.69) + 1533.74 + 635.46) - 3145.31$

$= 3242.90 - 3145.31$

$SS_A = 97.59$

Sum of Square within the group:-

$SS_W = SS_T - SS_A$

$= 590.85 - 97.59$

$= 493.26$

Sum of Squ. among group - $= \frac{SS_A}{K-1}$

$\mu SS_A = \frac{97.59}{3-1} = \frac{97.59}{2} = 48.79.$

Mean of Sum of Squ. within group.

$\mu_{SS_W} = \frac{SS_W}{N-K} = \frac{493.26}{24-3} = \frac{493.26}{21} = 23.48$

F-Ratio $= \mu SS_A / \mu SS_W = 48.79 / 23.48 = 2.077$

Compare.

Source of Variance	df	MSS	SS	F. Ration	F. Value (0.05)
Among grp	2	48.79	97.59	2.077	3.466
Within grp	21	23.48	493.26		

F. Ratio $\neq$ F. value.

So there is a difference in rate of return on of Industrie

(Q3) Hyp. test for 2 Sample Variables.

H_0 = There is no significant diff. between Machine 1 and 2 on thickness.

H_1 = There is significant diff. b/w machine 1 and 2 on thickness.

$\alpha = 0.05$, Two Tail

F. dist = $F = \frac{S_1^2}{S_2^2} = \frac{\text{est. } \sigma_1^2}{\text{est. } \sigma_2^2}$

$\frac{S_1^2}{S_2^2} = \frac{0.797}{0.688} =$

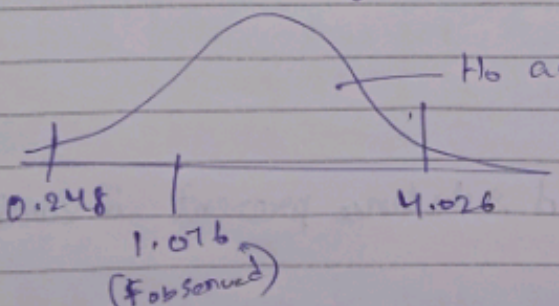
$\frac{\sigma_1}{\sigma_2} = \frac{0.893}{0.829} = 1.076 = F_{observed}$

F-table with df 4 Samples, $\alpha = 0.025$

$F_{0.25,9,9} = 4.026$

For Two-tail.

$F_{0.975} = \frac{1}{F_{.025}} = \frac{1}{4.026} = 0.248$



Since Fobserved is falling under Critical Region Ho is accepted.

So there is no significant difference on thickness between machine 1 and machine 2.

Q4. t-dist. and C.I for Can Speed.

$t_{\frac{\alpha}{2}} = \bar{x} \pm t \times S/\sqrt{n}$, $t = \frac{\bar{x} - \mu}{S/\sqrt{n}}$

$\Rightarrow \bar{x} = 100.83$

$\Rightarrow \mu = 100$

$\Rightarrow S = 5.55$

$\Rightarrow \sqrt{n} = 3.16$

$\Rightarrow t = \frac{100.83 - 100}{5.55(3.16)}$

$= \frac{0.83}{17.57} = 0.0472$

Sample err. $= 5.55/3.16$

$S/\sqrt{n} = 1.75$

$\Rightarrow t \times S/\sqrt{n} = 0.0472 \times 1.75$

margin of Error $= 0.083$

$t_{dist} = 100.83 + 0.083 < \mu < 100.83 - 0.083$

$= 100.913 < \mu < 100.747$

So the interval of the true mean will fall under above Confidence Interval.

Q5) T-test of book Sales.

H_0 = There is no Significant diff. in book Sales after the book enhancement.

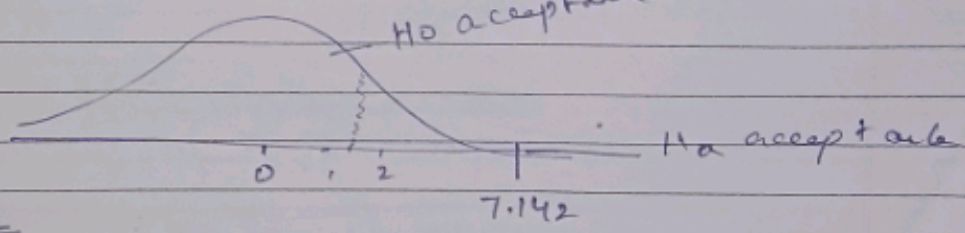
H_a = There is Significant diff. in book Sales after book enhancement.

$T\text{-test} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ $\Rightarrow \mu = 105,\ \bar{x} = 125,\ s = 14,\ n = 25$

$\alpha = 0.05$, One-tailed.

$= \frac{125 - 105}{14/\sqrt{25}} = \frac{20}{2.8} = 7.142$

$t_{0.05, 24} = 1.711$



Reject H_0

Since the value of 7.142 falls in Rejection area the, it is highly likely that mean Sale is greater and the book enhancement is probably a success.

(6) Z-test for Mean distance.

H_0 = The mean distance of U.S. workers is equal to the mean distance of workers in New York

H_0 = The mean distance is not same.

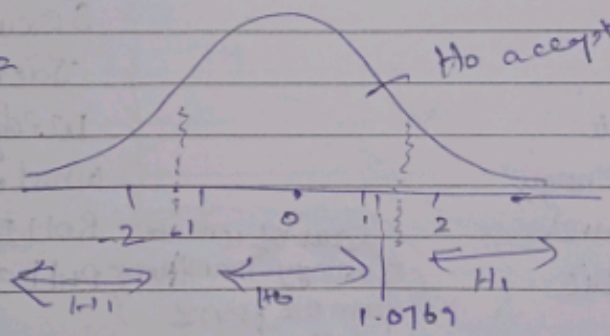
$$Z\text{-test} = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \Rightarrow \bar{x} = 16,\ \mu = 15,\ \sigma = 14,\ n = 169.$$

$\alpha = 0.10$, Two tailed.

$$= \frac{16 - 15}{14/13} = \frac{1}{0.928} = 1.0769.$$

$Z\ Score(\alpha/2) \Rightarrow z_{0.05} = 1.645$

=



Since the Z score is falling in acceptance region H_0 is accepted.

Interpretation:- The national average distribution does describe the mean commuting distance of all workers in New York area.