Rajnish Maurya DA24M015 Tutorial - 9

07/11/24

Solution 1:

a) Null Hypothesis (H_0) : Customer satisfaction is independent of the region
Alternate Hypothesis (H_1) : Customer satisfaction is not independent of the region

b) Observed frequencies

Region	Satisfied	Neutral	Dissatisfied	Row Total
North	45	30	25	100
South	40	35	25	100
West	50	25	25	100
Col Total	135	90	75	300

Expected frequencies:-

Region	Satisfied	Neutral	Dissatisfied
North	$\frac{100\times135}{300} = 45$	$\frac{100\times90}{300} = 30$	$\frac{100\times75}{300} = 25$
South	$\frac{135\times100}{300} = 45$	$\frac{100\times90}{300} = 30$	$\frac{100\times75}{300} = 25$
West	$\frac{135\times100}{300} = 45$	$\frac{100\times90}{300} = 30$	$\frac{100\times75}{300} = 25$

c) $\chi^2 = \frac{\Sigma(O_{ij} - E_{ij})^2}{E_{ij}}$

$= \frac{(45-45)^2}{45} + \frac{(30-30)^2}{30} + \frac{(25-25)^2}{25} + \frac{(40-45)^2}{45} + \text{---} + \frac{(25-25)^2}{25}$

$= 0 + 0 + 0 + 0.556 + 0.833 + 0 + 0.556 + 0.833 + 0$

$= 2.778$

d) degrees of freedom $= (r-1)(c-1)$

$df = (3-1)(3-1) = 4$

e) Critical value for $df = 4$ & $\alpha = 0.05 = 9.488$

$\chi^2 = 2.778$ Critical value $= 9.488$

Since $\chi^2 < 9.488$, fail to reject the null Hypothesis.

~~Sol~~ Solution 2:

Null Hypothesis (H_0): The mean strength of asphalt is same across the three levels of air void.

Alternate Hypothesis (H_1): Atleast one of mean strength is different.

$SSB = n \sum_{i=1}^{a} (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2 = 8(87.576 + 1.76 + 64.32)$
$= 1228.92$

$MSB = \frac{1228.92}{3-1} = 614.46$

$SST = \sum_{i=1}^{a} \sum_{j=1}^{a} (y_{ij} - \bar{y}_{\cdot\cdot})^2$
$= 2786$

$MST = \frac{2786}{24-1} = 121.13$

$SSE = 2786 - 1228.92 = 1557.08$

$MSE = \frac{1557.08}{24-3} = 74.15$

F-statistics $= \frac{MSB}{MSE} = \frac{614.41}{74.15} = \boxed{8.28.}$

$df_{between} = 2$ $\quad df_{within} = N - a = 21$, $\alpha = 0.01$

Critical value at $\alpha = 0.07$ for degree of freedom 2 & 21 is 5.78

Since Critical value < F-statistics, we reject the null Hypothesis.

Thus different levels of air voids affect the mean retained strength.

b) p-value for f-statistics $= 8.28$ is 0.0022

c) $CI = \bar{X} \pm t_{\alpha/2, df} * \frac{S}{\sqrt{n}}$
$= 75.5 \pm t_{0.025, 7} \frac{8.228}{\sqrt{8}}$
$= 75.5 \pm 2.365 * \frac{8.228}{\sqrt{8}}$
$= 75.5 \pm 2.83$

The CI is $(72.67, 78.33)$

d) $CI = (\bar{X}_1 - \bar{X}_2) \pm t_{\alpha/2, df} * \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$
$= (92.875 - 75.5) \pm t_{0.025, 14} \sqrt{\frac{(8.55)^2}{8} + \frac{(8.228)^2}{8}}$
$= 17.375 \pm 2.145 * 4.19$
$= 17.375 \pm 8.99$

CI for difference in mean $= (8.385, 26.365)$

Solution 4.

Null Hypothesis (H_0): Both are independent

Alternate Hypothesis (H_1): Breastfeeding duration & autism are not independent

$$E_{ij} = \frac{(\text{Row total}) \times (\text{Column total})}{(\text{Grand total})}$$

Expected frequencies:

	None	Less than 2	2-6 months	more than 6 months
Autism	228.48	195.18	167.38	226.95
No Autism	32.52	27.82	23.62	32.02

$$\chi^2_{test} = \Sigma \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

$= 11.22$

$df = (8-1)(6-1) = (2-1)(4-1) = 3$

for $\alpha = 0.01$ & $df = 3$, critical value $= 11.345$

So we reject the null Hypothesis as $\chi^2_{test} <$ critical value.

Solution 3.

H_0: The mean survival times for all types of cancer are equal

H_1: Atleast two types of cancer have different mean survival times.

$\bar{Y}_{stomach} = 286$, $\bar{Y}_{colon} = 457.41$, $\bar{Y}_{bronchus} = 211.589$

$\bar{Y}_{ovary} = 884.34$, $\bar{Y}_{breast} = 1395.9$, $\bar{\bar{Y}}_{..} = 558.265$

$SSB = 11535760.5$, $df_{SSB} = a-1 = 5-1 = 4$

$MSB = 11535760.5 / 4 = 2883940.13$

$SSE = 26448144.98$, $df_{SSE} = N-a = 64-5 = 59$

$MSE = 448273.64$

$SST = SSE - SSB = 37983905.0$

$df_{SST} = N-1 = 63$

$MST = 602919.127$

$F_{stat} = \frac{MSB}{MSE} = 6.434$

Critical value at $\alpha = 0.01$ for df 4 & 59 ≈ 2.04

Since the F-stats > Critical value at 1% Significance level, we reject the null Hypothesis.

Therefore, atleast two types of cancer have different mean survival times.