## 12.2: Test of Independence.

→ In general:

- $H_0$: The Row variable and the column variable are independent.

  $H_1$: The Row variable and the column variable are not independent.

Remark: Not independent mens dependent.

- We need to take a random sample:

$f_{ij}$ : observed freq.   ★ بنتعامل معه انه upper tail بس هو مش هيك

$e_{ij}$ : expected freq.

$n$ : # of Rows.

$m$ : # of columns.

$$\boxed{e_{ij} = \frac{(\text{row } i \text{ total})(\text{column } j \text{ total})}{\text{sample size}}}$$

Note: $\sum_j \sum_i f_{ij} = \sum_j \sum_i e_{ij}$ = sample size.

- Test statistic:

$$\chi^2 = \sum_j \sum_i \frac{(f_{ij} - e_{ij})^2}{e_{ij}} \quad \text{with } df = (n-1)(m-1).$$

assuming $e_{ij} \geq 5 \ \forall i \ \forall j$.

- Rejection Rule:

→ Reject $H_0$ if p-value $\leq \alpha$

→ Reject $H_0$ if $\chi^2 \geq \chi^2_\alpha$.

Q13:

Part Quality

| supplier | $f_{ij}$ Good | $e_{ij}$ | $f_{ij}$ Minor defect | $e_{ij}$ | $f_{ij}$ major Defect | $e_{ij}$ | Total |
|---|---|---|---|---|---|---|---|
| A | 90 | 88.76 | 3 | 6.07 | 7 | 5.17 | 100 |
| B | 170 | 173.09 | 18 | 11.83 | 7 | 10.08 | 195 |
| C | 135 | 133.15 | 6 | 9.10 | 9 | 7.75 | 150 |
| total | 395 | | 27 | | 23 | | 445 |

$\alpha = 0.05$.

→ $H_0$: supplier & Part Quality are independent.

$H_1$: supplier & part quality are not independent.

→ $\chi^2 = \frac{(90-88.76)^2}{88.76} + \dots + \frac{(9-7.75)^2}{7.75} = 7.72$

→ $df = (3-1)(3-1) = (2)(2) = 4$

By chi-square table:

| df | 0.05 |
|---|---|
| 4 | 9.488 |

$\chi^2_\alpha = 9.488$

| df | 0.90 | 0.10 |
|---|---|---|
| 4 | 1.064 | 7.779 |

upper tail area $\in (0.1, 0.9)$

P-value $\in (0.1, 0.9)$

→ • $\chi^2 < \chi^2_\alpha$ (critical approach)

• $(0.1, 0.9) > 0.05$ (P-value approach).

So we don't reject ($\alpha = 0.05$),

supplie and part quality are independent ($\alpha = 0.05$).