Azmayen Fayek Sabil

ID: 190042122

Assignment ( QUIZ-04 )

## chi-square test:

Example 4:

We want to test $H_0$: Accidents are uniformly distributed over the week

$H_1$: Accidents are not uniformly distributed over the week

We know,

$$\chi^2 = \Sigma \frac{(O-E)^2}{E} \quad \text{with } d.f = n-1$$

here, O = observed value

E = expected value

As,

E and $\bar{x}$ (arithmetic mean) are same,

$$\therefore E = \frac{\text{total number of accidents}}{\text{total days}} = \frac{14+16+8+20+11+9+14}{7} = \frac{92}{7} = 13.14$$

Now,

| Day | O | E | $(O-E)^2$ | $(O-E)^2/E$ |
|---|---|---|---|---|
| Sun | 14 | 13.14 | 0.7396 | — |
| Mon | 16 | 13.14 | 8.1796 | — |
| Tue | 8 | 13.14 | 26.4196 | — |
| Wed | 20 | 13.14 | 47.0596 | — |
| Thu | 11 | 13.14 | 4.5796 | — |
| Fri | 9 | 13.14 | 17.1396 | — |
| Sat | 14 | 13.14 | 0.7396 | — |

So, we get, $\Sigma \frac{(O-E)^2}{E} = 7.98$ and $df = n-1 = 7-1 = 6$

The calculated value of $\chi^2$ is $\chi^2_{cal} = 7.98$ and the table value of $\chi^2$,

for d.f. 6 at 5% level of significance is $\chi^2_{tab} = 12.59$

since $\chi^2_{cal} < \chi^2_{tab}$ So the null hypothesis is accepted.

## Example 5 :

We want to test ;

Ho : The populations are homogeneous with respect ----

We know,

$$\chi^2 = \sum \frac{(O-E)^2}{E} \text{, with d.f.} = (r-1)(c-1)$$

here,

O = observed frequency

$$E = \frac{\text{row total} \times \text{col total}}{\text{total number of observation}}$$

| Age group | Drama | Talk show | Cinema | Total |
|---|---|---|---|---|
| Under 30 | 120 | 30 | 50 | 200 |
| 30-40 | 10 | 75 | 15 | 100 |
| 45 and above | 10 | 30 | 60 | 100 |
| TOTAL | 140 | 135 | 125 | 400 |

| O | E | $(O-E)^2$ | $(O-E)^2/E$ |
|---|---|---|---|
| 120 | 70 | 2500. | – |
| 10 | 35 | 625 | – |
| 10 | 35 | – | – |
| 30 | 67.5 | – | – |
| 75 | 33.75 | – | – |
| 30 | 33.75 | – | – |
| 50 | 62.5 | – | – |
| 15 | 31.25 | – | – |
| 60 | 31.25 | – | – |

$\chi^2 = \Sigma \frac{(O-E)^2}{E} = 180.5$ d.f. $= (3-1)(3-1) = 4$

The calculated value of $\chi^2$ is $\chi^2_{cal} = 180.5$ and the table value of $\chi^2$ for df 4

at 5% is $\chi^2_{tab} = 9.4877$

$\chi^2_{cal} > \chi^2_{tab}$. So the hypo is rejected.

Ex: Example 19.6 :

1. $H_0: M = 98.5$

2. $H_1: M \neq 98.5$

3. $\alpha = 0.05$

4. Reject the null if $T \leq T_{0.05}$

5. Subtracting 98.5 from each value and ranking the differences without regard to their sign,

| Observation X | X−98.5 | \|X−98.5\| | Rank |
|---|---|---|---|
| 97.5 | −1 | 1 | 4 |
| 95.2 | −3.3 | 3.3 | 12 |
| 97.3 | −1.2 | 1.2 | 6 |
| 96.0 | −2.5 | 2.5 | 10 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 94.9 | −3.6 | 3.6 | 13 |

So that $T^- = 4+12+6+\cdots+13 = 95$

$T^+ = 8+2=10$ and $T=10$

From the table 10, we find that $T_{0.05} = 21$ for $n=14$

$T_{obs} < T_{0.05}$ . Null rejected .

Ex. Example 19.7:

(i) $H_0: M_1 = M_2$

(ii) $H_1: M_1 \neq M_2$

(iii) $\alpha = 0.05$

(iv) reject null hyp if $T \leq T_{0.05}$

(v)

| $X$ | $Y$ | $D = (X-Y)$ | $\lvert D \rvert$ | rank | $R(+)$ | $R(-)$ |
|---|---|---|---|---|---|---|
| 16 | 4 | 12 | 12 | 9.5 | ✓ | |
| 12 | 18 | −6 | 6 | 5 | | ✓ |
| 22 | 10 | 12 | 12 | 9.5 | ✓ | |
| 16 | 14 | 2 | 2 | 1.5 | ✓ | |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | | |
| 22 | 12 | 10 | 10 | 7.5 | ✓ | |

$T^- = 8$, $T^+ = 47$

$T = \min(8, 47) = 8$

$T_{0.05} = 8$

$T \leq T_{0.05}$ . So null must be rejected

Hypo example :

Given that,

$H_0 : \mu = 138$

$H_1 : \mu = 142$

$\alpha = 0.05$

standard deviation = 2

we know,

$$z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}}$$

Critical value for 0.05 significance level is 1.645 (greater than)

Now,

$$\bar{x} = \mu_0 + z_{0.05} \sigma_{\bar{x}}$$

$= 138 + 1.645 * 2$

$= 141.30$

The corresponding z value ;

$$z = \frac{141.30 - 142}{2}$$

$= -0.35$

Accepted $H_0$ .

Ex: Cov problem;

Given that, $f(x,y) = \begin{cases} 6/5\,(x^2+2xy) & ; \ 0 \le x \le 1, \quad 0 \le y \le 1 \\ 0 & ; \ \text{otherwise} \end{cases}$

The marginal density function of X is,

$$g(x) = \frac{6}{5}\left[x^2 y + x y^2\right]_0^1$$

$$= \frac{6}{5}(x^2 + x)$$

$$= \frac{6}{5}x(x+1)$$

$$h(y) = \frac{6}{5}\left[x^3/3 + x^2 y\right]_0^1$$

$$= \frac{6}{5}\left[\frac{1}{3} + y\right]$$

$$= \frac{2}{5}(1+3y)$$

$$\mu_x = \int x \cdot \frac{6}{5}x(x+1)\,dx$$

$$= \frac{6}{5}\int (x^3+x^2)\,dx = \frac{6}{5}\left[\frac{x^4}{4} + \frac{x^3}{3}\right]_0^1$$

$$= \frac{6}{5}\left(\frac{1}{4}+\frac{1}{3}\right) = \frac{3+4}{12} = \frac{7}{10}$$

$$\mu_y = \int y \cdot \frac{2}{5}(1+3y)\,dy = \frac{3}{5}$$

$$E[XY] = \frac{6}{5}\int\int xy(x^2y + 2xy)\,dx\,dy$$

$$= \frac{6}{5}\int\int (x^3y + 2x^2y^2)\,dx\,dy$$

$$= 5/12$$

$$\text{cov}(X,Y) = 5/12 - (7/10 \times 3/5)$$

$$= -1/300 \quad (\text{Ans})$$

---

$H_0: \mu = 3.65$ $\quad\sigma = 0.16$

$H_1: \mu \neq 3.65$ $\quad n = 36$

$\bar{x} = 3.45$

$\alpha = 0.05 \longrightarrow Z < -1.96$

$Z > 1.96$

$$Z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.45 - 3.65}{0.16/\sqrt{36}} = -7.5$$

$H_0 \longrightarrow \times$

$n = 40$, $\bar{x} = 2.8$, $\sigma = 0.35$, $\alpha = 0.01$, $z < 2.57$, $z > 2.57$

$H_0: \mu = 2.4$

$H_1: \mu \neq 2.4$

$$z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} = 7.22$$

$$\mu_0 \bar{x} = \bar{x} \pm z \times \frac{\sigma}{\sqrt{n}}$$

$$= 2.8 \pm 2.57 \times \frac{0.15}{\sqrt{40}}$$

$$= (2.26, 2.54)$$

| x | y | xy | x² |
|---|---|---|---|
| 10 | 5 | 50 | 100 |
| 12 | 6 | 72 | 144 |
| 13 | 7 | 91 | 169 |
| 17 | 9 | 153 | 289 |
| 18 | 13 | 234 | 324 |
| 70 | 40 | 600 | 1026 |

$$\frac{600 - \frac{70 \times 40}{5}}{1026 - \frac{70^2}{5}}$$

0.87

## Formulas

৫। ANOVA

(i) hypothesis, $H_0, H_1$ লেখব; $\alpha$ লেখব

(ii) group এর $x_1, x_2 \cdots x_n$ আবার দেখব and $x_1^2, x_2^2 + i \cdots$ বের করব

with "$\Sigma$" of each column

| $x_1$ | $x_2$ | $\cdots$ | $x_1^2$ | $x_2^2$ | $\cdots$ |
|---|---|---|---|---|---|
| | | | | | |
| $T_i \to \Sigma$ | | | | | |

n হলো total element

(iii) $T = \Sigma T_i$

$$CT = \frac{T^2}{n}$$

(iv) $SST = \Sigma T_i^2 - CT$

$$SSB = \left(\Sigma \frac{T_i^2}{n_i}\right) - CT$$

$$SSW = SST - SSB$$

(v) degree of freedom বের করব,

n = total element
k = group

$df_{SSB} = k - 1$

$df_{SSW} = n - k$

$df_{SST} = n - 1$

mean squares

$MSB = SSB / k-1 = \checkmark$

$MSW = SSW / n-k = \checkmark$

$MSB / MSW$ = f-ratio

compare f-ratio with f-critical from table

it ke accepted or rejected Ho

৪২

# Nonparametric test

\# $Z = \dfrac{S - np}{\sqrt{np(1-p)}} = \dfrac{S - 0.5n}{0.5\sqrt{n}}$ here $p = 0.5$

(i) math section added

$n$ = total number of ...

$S$ = number of plus signs

# wilcoxon signed rank test;

| X | X − mean | \|X − mean\| | rank |
|---|---|---|---|
| − | | | |
| − | | | |
| − | | | |

rank; 1, 2, 4, 6, 8, 9, 9, 10, 11, 11

① ② ③ ④ ⑤ 6.5 6.5 8 9.5 9.5

$6+7 = 13 = 13/2 = 6.5$

$9+10 = 19 = 9.5$

then $T^+ = \Sigma$ (rank of all positives)

$T^- = \Sigma$ (rank of all negatives)

$T = T^+$

table থেকে $T_{table}$ = ৫ then compare

( $T$ with $T_{table}$ )

reject if

| | |
|---|---|
| $H_1: M \neq M_0$ | $T < T_\alpha$ |
| $H_1: M > M_0$ | $T^- < T_{2\alpha}$ |
| $H_1: M < M_0$ | $T^+ < T_{2\alpha}$ |

# spearmanh rank correlation coefficient :

$$r_s = 1 - \frac{6 \Sigma d^2}{n(n^2-1)} \qquad d = R_1 - R_2$$

# Chi- square test ✓

# Goodness of fit

$$\chi^2 = \Sigma \frac{(O-E)^2}{E}$$

O = observed value

E = expected value = mean

df = n-1

# test of homogeneity

$$\chi^2 = \Sigma \frac{(Ob-E)^2}{E}$$

df = (r-1) (c-1)

O = observed value

$$E = \frac{\text{row total} \times \text{col total}}{\text{total no. of observation}}$$

Ex Hypo

$$z = \frac{\bar{x} - \mu_0}{\sigma_{\bar{x}}} \qquad \bar{x} = \mu_0 + z \cdot \sigma_{\bar{x}}$$

or

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \qquad \bar{x} = \mu_0 \pm z \frac{\sigma}{\sqrt{n}} \quad \text{(twotailed)}$$

if rejected,

then $\mu_0 = \bar{x} \pm z \frac{\sigma}{\sqrt{n}}$

$H_0$ ~~: μ = 880~~ $\quad$ $\mu_0 =$

~~$H_a$ μ ≠ 880~~ $\quad$ $\alpha$ ~~= 0.05~~

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \times$$

$$z = \frac{\bar{x}_1 - x_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

## Covariance - correlation

$$\text{cov}(X,Y) = E[XY] - \mu_X \mu_Y$$

correlation coeffient $\rho(X,Y) = \dfrac{\text{Cov}(X,Y)}{\sqrt{V(X)\,V(Y)}}$

$$\mu_x = E[X] = x \int g(x)$$

$$\mu_y = E[Y] = y \int h(y)$$

$$V[X] = E[X^2] - (E[X])^2$$

or,

$$r = \frac{\sum xy - \dfrac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \dfrac{(\sum x)^2}{n}\right)\cdot\left(\sum y^2 - \dfrac{(\sum y)^2}{n}\right)}}$$

probable error, $P.E(r) = \dfrac{1-r^2}{\sqrt{n}} \times 0.6745$ → আসল n থেকে error calculate

$6 \times P.E(r) < r$ হলে good value

Rank correlation;

$$R = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

$d = (R_1 - R_2)$ → individual rank থেকে আসবে

Regression analysis

$$y = a + bx$$

$$a = \bar{y} - b\bar{x} \qquad b = \frac{\Sigma xy - \frac{\Sigma x \Sigma y}{n}}{\Sigma x^2 - \frac{(\Sigma x)^2}{n}}$$

$$\hat{y} - \bar{y} = b(x - \bar{x})$$

$r = \sqrt{b \times d}$ → $b \to x$, $d \to y$ regression coefficient

correlation coefficient

coefficient of determination, $r^2 = \dfrac{\Sigma(\hat{y}-\bar{y})^2}{\Sigma(y-\bar{y})^2}$

ANOVA

$$SST = SSB + SSW$$

SSB, dof = $k-1$ ② mean squares of SSB = $\dfrac{SSB}{k-1}$ = MSB

SSW, dof = $n-k$ ③ " " " SSW = $\dfrac{SSW}{n-k}$ = MSW

SST, dof = $n-1$ ①

F-ratio = MSB/MSW