Stats.

Q.1.

| Group of People | Baloperidol | | |
|---|---|---|---|
| | X | Y | Z |
| A | 25 | 7 | 13 |
| | 27 | 8 | 18 |
| B | 21 | 16 | 19 |
| | 24 | 11 | 14 |
| C | 29 | 19 | 30 |
| | 31 | 21 | 27 |

→

Step 1 :

total no. of obs. (N) = 18

grand sum $(\Sigma x) = 360$

grand mean $(\bar{x}) = \frac{360}{18} = 20$

Correction Factor (CF) = $(\Sigma x^2)/N = (360)^2/18 = 7200$

$\therefore$ CF = 7200.

Sum of squares of all obs. $(\Sigma x^2)$ = $25^2 + 27^2 + 7^2 + 8^2 + 13^2 + 18^2 + 21^2 + 24^2 + 16^2 + 11^2 + 19^2 + 14^2 + 29^2 + 31^2 + 19^2 + 21^2 + 30^2 + 27^2$

$\therefore \Sigma x^2 = 8144$

Step 2 :

(1) total sum of squares (SST) = $\Sigma x^2 - CF$

$= 8144 - 7200$

$\therefore$ SST = $\cancel{8}944$

(2) Sum of squares between rows (SSR)

$\left\{ \therefore SSR = \frac{\Sigma R_i^2}{b \times r} - CF \right\}$

$R^2_A = (25+27+7+8+13+18)^2 = (98)^2$

$\therefore R^2A = 9604$ —— (1)

$R^2_B = (21+24+16+11+19+14)^2 = (104)^2$

$= 11025$ —— (2)

$R^2_c = (29+31+19+21+30+27)^2 = (157)^2$

$= 24649$ —— (3)

Now, from (1) (2) & (3)

$\Sigma Ri^2 = R^2_A + R^2_B + R^2_c$

$= 9604 + 11025 + 24648$

$= 45278$

$SSR = \frac{45278}{3 \times 2} - CF = \frac{45278}{6} - 7200$

$= 346.33$

(3) Sum of squares btw columns (SSC)

$\left\{ SSC = \frac{\Sigma Ci^2}{a \times r} - CF \right\}$

$Cx^2 = (25+27+21+24+29+31)^2 = (157)^2$

$= 24649$ —— (i)

$Cy^2 = (7+8+16+11+19+21)^2 = (82)^2$

$= 6724$ —— (ii)

$Cz^2 = (13+18+19+14+30+27)^2 = (121)^2$

$= 14641$ —— (iii)

Now., $\Sigma C_i^2 = C_x^2 + C_y^2 + C_z^2$

$= 24649 + 6724 + 14641$

$= 46014.$

$SSC = \frac{46014}{3 \times 2} - CF = \frac{46014}{6} - 7200$

$\therefore SSC = 469.$

(4) Sum of square within cells (SSE)

first compute sum of squares betw^n cells (SSB-cells).

1st cell = 25 + 27 = 52
2nd cell = 7 + 8 = 15
3rd cell = 13 + 18 = 31
4th cell = 21 + 24 = 45
5th cell = 16 + 11 = 27
6th cell = 19 + 14 = 33
7th cell = 29 + 31 = 60
8th cell = 19 + 21 = 40
9th cell = 30 + 27 = 57

$\therefore$ Squares of cells.

$= (52)^2, (15)^2, (31)^2, (45)^2, (27)^2, (33)^2, (60)^2, (40)^2, (57)^2.$

$\Rightarrow$ 2704, 225, 961, 2025, 729, 1089, 3600, 1600, 3249

$\therefore$ Sum = 2704 + 225 + 961 + 2025 + 729 + 1089 + 3600 + 1600 + 3249

= 16182.

$SSB_{cell} = \frac{16182}{r} - CF$

$= \frac{16182}{2} - 7200$

$= 891.$

Now., calculating SSE

$\therefore SSE = SSI - SSB\ cell$

$= 944 - 891$

$= 53.$

5. Sum of square for interaction (SSI).

SSI = SSB cell − SSR − SSC
= 891 − 346.33 − 469
= 75.67.

Step 3: [ANOVA Table.]

| Source of variation. | SS | df | MS. | F ratio. |
|---|---|---|---|---|
| Rows. | 346.33 | (r−1)<br>(3−1) = 2 | 346.33/2<br>= 173.165 | $F_r$ = 173.165/4.417<br>= 39.20 |
| Columns. | 469 | 2 | 469/2<br>= 234.5 | $F_c$ = 234.5/4.417<br>= 53.09 |
| Interactions | 75.67 | 4 | 75.67/4<br>= 18.917 | $F_t$ = 18.917/4.417<br>= 4.28. |
| Error | 53 | 9 | 4.417 | – |
| total. | 944 | 17 | – | – |

(i) Do the baloperidol acts differently?
Yes (F = 53.09, P < 0.05)

$F_\alpha$
F(2,9)
= 4.29.

(ii) Are the differently group of people affect differently?
Yes (F = 39.20, P < 0.05)

F(4,9)
= 3.63

(iii) Is the interaction term significant?
Yes (F = 4.28, P < 0.05)

Q.2. You are conducting, study to compare the performance of three different type of enhancers [enhancer A, enhancer B, and enhancer C] in promoting the growth of a specific plant species. You have collected data on the height of the plants in each group after six weeks. The dataset continue the following measurement

(a) Enhancer A : [18, 20, 22, 25, 21]
(b) Enhancer B : [15, 16, 14, 17, 18]
(c) Enhancer C : [20, 24, 22, 21, 25]

Do the baloperidal set differently?
are the different groups of people affected differently?
is the interaction term significant?

Step 1.
(N) total no. of obs = 18
ground sum ($\Sigma n$) = 360
ground mean ($\bar{x}$) = 360/18 = 20
Correction factor (CF) = $(\Sigma x^2)/N = \frac{(360)^2}{18} = 72.00.$

Sum of squares of all obs. ($\Sigma x^2$) $= 25^2 + 27^2 + 7^2 + 8^2 + 13^2 + 18^2 + 21^2 + 24^2$
$+ 16^2 + 11^2 + 19^2 + 14^2 + 29^2 + 31^2 + 19^2 + 21^2$
$+ 30^2 + 27^2$

NA

a.) Explain the basic concept of the Kruskal-wallis test and when it is appropriate to use it.

→ The Kruskal-wallis test is a non-parametric method used to compare three or more independent groups when the dependent variable is either ordinal or continous, but not normally distributed.

- It is essentially an extension of the mann-whitney test to more than two group.
- the test ranks all the data from the group together then, compares the avg. ranks btwn groups. If the same groups are from same population, their ranks should be similarly distributed.
- A significant result inside indicates that atleast one group tends to have larger or smaller values than the others.

b.) Calculate the Kruskal-Wallis statistic for the given datasets.

→ data :- enhancer (A) : 18, 20, 22, 25, 21

enhancer (B) : 15, 16, 14, 17, 18

enhancer (C) : 20, 24, 22, 21, 25

Step 1 : combine and rank all data.

| values | group | ranks. |
| --- | --- | --- |
| 14 | B | 1 |
| 15 | B | 2 |
| 16 | B | 3 |
| 17 | B | 4 |
| 18 | A | 5.5 |
| 18 | B | 5.5 |
| 20 | A | 7.5 |
| 20 | C | 7.5 |
| 21 | A | 9.5 |
| 21 | C | 9.5 |
| 22 | A | 11.5 |
| 22 | C | 11.5 |
| 24 | C | 13 |
| 25 | A | 14.5 |
| 25 | C | 14.5 |

Step 2 : Sum of ranks for each group.

- enhancer A : 5.5 + 7.5 + 9.5 + 11.5 + 14.5 = 48.5
- enhancer B : 1 + 2 + 3 + 4 + 5.5 = 15.5
- enhancer C : 7.5 + 9.5 + 11.5 + 13 + 14.5 = 56

Step 3: Kruskal-Wallis formula.

$$\left\{ H = \frac{12}{N(N+1)} \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3(N+1) \right\}$$

where ->

$N = 15$     $R_1 = 48.5$

$k = 3$     $R_2 = 15.5$

$n_i = 5$ each     $R_3 = 56$

$$\frac{\Sigma R_i^2}{n_i} = \frac{(48.5)^2}{5} + \frac{(15.5)^2}{5} + \frac{(56)^2}{5}$$

$$= 1145.7$$

$$\therefore H = \frac{12}{15 \times 16} \times 1145.7 - 3 \times 16$$

$$= 9.285$$

Step 4 : Adjust for ties.

18 (2 time) - t = 2
20 (2 time) - t = 2
21 (2 time) - t = 2
22 (2 time) - t = 2
25 (2 time) - t = 2

$$C = 1 - \frac{\Sigma(t^3 - t)}{N^3 - N}$$

$$\Sigma(t^3 - t) = 5 \times (8 - 2) = 5 \times 6 = 30.$$

$$\therefore C = 1 - \frac{30}{3375 - 15} = 0.99107$$

$$\therefore H_{adj} = \frac{H}{C} = \frac{9.285}{0.99107} = 9.37$$

$$\therefore H_{adj} = 9.37$$

c.) State the null and alternative hypothesis for the Kruskal-Wallis test in this context

→ Null hypothesis ($H_0$).
- the population median of plant heights for all the three enhancers are the same.

Alternative hypothesis ($H_1$).
- at least one enhancer tends to yield different plant heights than the others.

d.) Interpret the results of the Kruskals-Wallis test and provide a conclusion regarding the effect of different enhancecers on plant height.

→ from chi-square table,

for $df = k-1 = 2$ ., critical value at $\alpha = 0.05$ is about $5.991$ Our $H \approx 9.37 > 5.991$,

So we reject $H_0$.

Conclusion :

there is a statistically significant difference in plant heights amoung at least one pair of the three enhancers at the 5% significance level.