PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

Semester : 1   Subject : Statistics for AI & DS   Academic Year: 20 23 2024

TWO-WAY ANOVA:

The following data represents the number of units of tablet production [in thousands] per day by five different technicians by using four different type of machines.

Workers	A	B	C	D
P	54	48	57	46
Q	56	50	62	53
R	44	46	54	42
S	53	48	56	44
T	48	52	59	48

(a) Test whether the mean productivity of the different machines are same?

(b) Test whether the 5 technicians differ with respect to the mean productivity?

Solution :

Given data : Calculation of Grand Total & Correction factor.

Workers	A	B	C	D	~~Total~~
P	54	48	57	46	
Q	56	50	62	53	
R	44	46	54	42	
S	53	48	56	44	
T	48	52	59	48	
~~Total~~					

PARSHWANATH CHARITABLE TRUST'S
A.P. SHAH INSTITUTE OF TECHNOLOGY
Department of Computer Science and Engineering
Data Science

CSE DATA SCIENCE

Semester : __1__   Subject : Statistics for AI&DS   Academic Year: 20 23 2024

Step 1 : Calculation of Grand Total and Correction factor (Can take any value from the range given. For easy calculation we consider 50)

	A	B	C	D	Total
P	4	−2	+7	− 4	+5
Q	6	0	+12	+3	+21
R	−6	−4	+4	−8	−14
S	+3	−2	+6	−6	+1
T	−2	+2	+9	−2	+7
Total	5	−6	38	−17	20 → Grand Total

Correction factor $= \dfrac{T^2}{N}$

$$= \dfrac{20^2}{(4 \times 5)} \quad \boxed{= 20}$$

Source of Variation	Sum of Squares	Degree of Freedom	Mean sum of Squares	Ratio of F
Between the Coloumns.	SSC = 338.8	$v = (C-1)$ = 4−1 = 3 now	MSC = SSC/(C−1) = 338.8/3 = 112.93	MSC/MSE = 20.16 $\dfrac{=112.93}{5.6.}$
Between the Rows.	SSR = 158	$v = (r-1)$ = 5−1 = 4 now	MSR = SSR/(r−1) = 158/4 = 39.5	MSR/MSE = 7.05 $\dfrac{=39.5}{5.6}$
Residual or Error.	SSE SSC−SSR = 67.2	$v = (C-1)(r-1)$ (4−1)(5−1) = 12 colour	MSE = SSE/((C−1)(r−1)) = 67.2/12 = 5.6	
	SST = 564	done		

Semester: $\underline{\text{I}}$ Subject: Statistics for AI& DS Academic Year: 2023-2024.

Step 2: Calculation of SSC

$$SSC = \frac{A^2}{n_A} + \frac{B^2}{n_B} + \frac{C^2}{n_C} + \frac{D^2}{n_D} - \frac{T^2}{N}$$

$$SSC = \frac{5^2}{5} + \frac{(-6)^2}{5} + \frac{(38)^2}{5} + \frac{(-17)^2}{5} - 20$$

$$= \frac{25}{5} + \frac{36}{5} + \frac{1444}{5} + \frac{289}{5} - 20$$

$$= 5 + 7.2 + 388.8 + 57.8 - 20 \boxed{= 338.8}$$

Step 3: Calculation of SSR.

$$SSR = \frac{P^2}{n_P} + \frac{Q^2}{n_Q} + \frac{R^2}{n_S} + \frac{T^2}{n_T} - \frac{T^2}{N}$$

$$SSR = \frac{(+5)^2}{4} + \frac{(21)^2}{4} + \frac{(-14)^2}{4} + \frac{(+1)^2}{4} + \frac{(7)^2}{4} - 20$$

$$= \frac{25}{4} + \frac{441}{4} + \frac{196}{4} + \frac{1}{4} + \frac{49}{4} - 20$$

$$= 6.25 + 110.25 + 49 + 0.25 + 12.25 - 20 \boxed{= 158}$$

Step 4: Calculation of SST.

$$SST = (4)^2 + (6)^2 + (-6)^2 + (3)^2 + (-2)^2 + \cancel{(6)^2} +$$
$$(-2)^2 + (0)^2 + (-4)^2 + (-2)^2 + (2)^2 + \cancel{(-6)^2} +$$
$$(+7)^2 + (12)^2 + (4)^2 + (6)^2 + (9)^2 +$$
$$(-4)^2 + (3)^2 + (-8)^2 + (-6)^2 + (-2)^2 +$$
$$\cancel{(5)^2} + \cancel{(21)^2} + \cancel{(-14)^2} + \cancel{(1)^2} + \cancel{(7)^2}$$

$$SST = 16 + 36 + 36 + 9 + 4 + 4 + 0 + 16 + 4 + 4 + 49 + 144 +$$
$$16 + 36 + 81 + 16 + 9 + 64 + 36 + 4 \cancel{+251}$$

Semester : Ⅶ

Subject : Statistics for AI&DS   Academic Year: 2023-20 24

$$SST = 564$$

Table value :

$\nu_1 = 12, \quad \nu_2 = 3, \quad F_{0.05} = 3.49$

$\nu_1 = 12, \quad \nu_2 = 4 \quad F_{0.05} = 3.26.$

$F = 20.16 > F_{0.05} = 3.49.$

Hence the productivity of different machines are different.

$F = 7.05 > F_{0.05} = 3.26.$

Hence the 5 technicians differ with respect to the mean productivity.

Semester : ___V___   Subject : Statistics for AI&DS   Academic Year: 2023-2024.

Example2 :- (Two-way ANOVA).

To study the performance of three detergents and three different water temperature, the following whiteness readinggs were obtained with specially designed equipment.

Water Temperature	A	B	C
Cold Water	47	45	50
Warm Water	39	42	52
Hot Water.	44	36	48

Perform a two way ANOVA, using 5% level of significance.

Solution:

Detergent Water Temperature	A	B	C
Cold Water (C)	47	45	50
Warm Water (W)	39	42	52
Hot Water (H)	44	36	48.

(1) Calculation of GrandTotal and Correction factor :

Data is coded by subtracting any guessed mid value (i.e. 40) for easy calculation.

Semester : I Subject : Slatistics for AI&DS   Academic Year: 20 23- 2024 .

	A	B	C	Total
C	7	+5	+10	+22
W	-1	+2	+12	+13
H	+4	-4	+8	+8
T.	+10	3	30	43

$T \rightarrow$ Grand Total = 43.

Correction factor $= \dfrac{T^2}{N}$

$= \dfrac{(43)^2}{9} = \dfrac{1849}{9} = \boxed{205.44}$.

2 Way Anova — Table for Calculation :-

Source of Variation	Sum of Squares	Degree of Freedom	Mean sum of squares	Ratio of F
B/w the Coloumns.	SSC = 130.89	$v = (C-1)$ = 2.	$MSC = SSC/v$ = $130.89/2 = 65.45$	MSC/MSE = $65.45/12.28 = 5.32$
B/w the rows	SSR = 33.55	$v = (r-1)$ = 2	$MSR = SSR/v$ = $\dfrac{33.55}{2} = 16.78$	MSR/MSE = $16.78/12.28 = 1.37$
Residual Error.	SSE = 49.12	$v = (C-1)(r-1)$ = $(2)(2) = 4$	$MSE = SSE/v$ = $49.12/4 = 12.28$.	

(2) Calculation of SSC.

$$SSC = \dfrac{A^2}{n_A} + \dfrac{B^2}{n_B} + \dfrac{C^2}{n_c} - \dfrac{T^2}{N}$$

$$= \dfrac{(10)^2}{3} + \dfrac{(3)^2}{3} + \dfrac{(30)^2}{8} - 205.44 .$$

$$= 100/3 + 9/3 + 900/8 - 205.44$$

$$= 33.33 + 3 + 300 - 205.44$$

$$\boxed{SSC = 130.86}$$

Semester : _____I_____   Subject : Statistics for AI&DS   Academic Year: 20 23 2024 .

(3) Calculation of SSR :-

$$SSR = \frac{C^2}{n_c} + \frac{W^2}{n_w} + \frac{H^2}{n_H} - \frac{T^2}{N}$$

$$SSR = \frac{(22)^2}{3} + \frac{(13)^2}{3} + \frac{8^2}{8} - 205.44$$

$$SSR = \frac{484}{3} + \frac{169}{3} + \frac{64}{8} - 205.44$$

$$= 161.3 + 56.33 + 21.33 - 205.44$$

$$\boxed{SSR = 33.55}$$

(4) Calculation of SST.

$$SST = (7)^2 + (-1)^2 + (4)^2 + (5)^2 + (2)^2 + (-4)^2 + (10)^2 + (12)^2 +$$

$$(8)^2 - 205.44$$

$$= 49 + 1 + 16 + 25 + 4 + 16 + 100 + 144 + 64 - 205.44.$$

$$= 419 - 205.44.$$

$$\boxed{SST = 213.56.}$$

(5) Calculation of SSE.

$$SSE = SST - (SSC + SSR).$$

$$= 213.56 - (130.86 + 33.55).$$

$$\boxed{SSE = 49.12}$$

Tabulated F value , $\nu_1 = 4, \nu_2 = 2$, $F_{0.05} = 6.94$.

5.32 < $F_{0.05}$ = 6.44 ∴. There is no significance difference between the different detergents

Tabulated F value , $\nu_1 = 4, F_{0.05} = 6.94$.

1.87 < $F_{0.05}$ = 6.44 ∴ There is no significance difference b/w the diff water Temperature.