Assignment

Wednesday, May 12, 2021 6:48 PM

3. The number of students arriving late at five high schools on different days of week are given the following table:

RBD:

Days of week	High School			
	A	B	C	D
Monday	5	4	5	7
Tuesday	4	5	3	2
Wednesday	4	3	4	5
Thursday	4	4	3	2

Is there any significant difference in the number of late arrivals among different days of the week at the significance level $\alpha = 1\%$?

- Why Two-Way ?
  + Goal : Difference between different days

- Remember we are really interested in SSC or column/group variance (cities) as a part of SST, but we need to account for variation between the secret shoppers.
- Anything that is not SSC is error; SSE.
- Introducing BLOCKS, allows us to potentially reduce SSE even more by splitting the original SSE (as in One-Way) into SSB and a smaller SSE (Two-Way).
- If SSC is much larger than minimized SSE, then there are likely group/column differences.
- Comparing SSC to the *minimized* SSE makes the test more powerful and more sensitive; SSC relative to SSE.

Note:

$SSC = SS_{Treatment}$

$SST = SS_T$

$SSE = SS_E$

$SSB = SS_{Blocks}$

- Randomize Block Design (RBD) ?
  + Blocks are independent

- R. Complete. B. D (RCBD)?
  + All Blocks has a value

- With or Without ?

→ Without : Only 1 result per method/block per column

With Replication

Industries	Districts			
	1	2	3	4
Refrigeration	2.5, 2.7, 2.0, 3.0	13.1, 3.5, 2.7	2.0, 2.4	5.0, 5.4
Construction Materials	0.6, 10.4	15.0	9.5, 9.3, 9.1	19.5, 17.5
Computer services	1.2, 1.0, 9.8, 1.8	2.0, 2.2, 1.8	1.2, 1.3, 1.2	5.0, 4.8, 5.2

Column (School)

Block (Day)

Days of week	High School			
	A	B	C	D
Monday	5	4	5	7
Tuesday	4	5	3	2
Wednesday	4	3	4	5
Thursday	4	4	3	2

Days of week	A	B	C	D	Total $y_{i\cdot}$	Ave. $\bar{y}_{i\cdot}$
Monday	5	4	5	7	21	5,25
Tuesday	4	5	3	2	14	3,5
Wednesday	4	3	4	5	16	4
Thursday	4	4	3	2	13	3,25
Total : $y_{\cdot j}$	17	16	15	16		
Average : $\bar{y}_{\cdot j}$	4,25	4	3,75	4		

$\mu = 4$ (overall mean)

$\Sigma = 4.16 = 64$

Assume

$H_1 : \tau_i \neq 0$

$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = 0$ (different days)

13-4.1 Design and Statistical Analysis

The appropriate linear statistical model:

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, b \end{cases}$$

We assume

- treatments and blocks are initially fixed effects
- blocks do not interact
- $\sum_{i=1}^{a} \tau_i = 0$ and $\sum_{j=1}^{b} \beta_j = 0$

Independent

$a = 4 \quad b = 4$

$SS_T = 5^2 + 4^2 + 5^2 + \ldots + 2^2 - \frac{64^2}{16} = 24$

$SS_{Treatment} = \frac{21^2 + 14^2 + 16^2 + 13^2}{\ldots} - \frac{64^2}{\ldots} = 9,5$

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - \frac{y_{\cdot\cdot}^2}{ab} \quad (13\text{-}18)$$

$$SS_{Treatments} = \frac{1}{b} \sum_{i=1}^{a} y_{i\cdot}^2 - \frac{y_{\cdot\cdot}^2}{\ldots} \quad (13\text{-}19)$$

16

$$SS_{Treatment} = \frac{21^2 + 14^2 + 16^2 + 13^2}{4} - \frac{64^2}{16} = 9{,}5$$

$$SS_{Blocks} = \frac{17^2 + 16^2 + 15^2 + 16^2}{4} - \frac{64^2}{16} = 0{,}5$$

$$SS_E = 24 - 9{,}5 - 0{,}5 = 14$$

$$SS_T = \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij}^2 - \frac{y_{..}^2}{ab} \quad (13\text{-}18)$$

$$SS_{\text{Treatments}} = \frac{1}{b} \sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{ab} \quad (13\text{-}19)$$

$$SS_{\text{Blocks}} = \frac{1}{a} \sum_{j=1}^{b} y_{.j}^2 - \frac{y_{..}^2}{ab} \quad (13\text{-}20)$$

and

$$SS_E = SS_T - SS_{\text{Treatments}} - SS_{\text{Blocks}} \quad (13\text{-}21)$$

$$\Rightarrow MS_{Treatment} = \frac{9{,}5}{3}$$

$$MS_{Blocks} = \frac{0{,}5}{3}$$

$$MS_E = \frac{14}{9}$$

The mean squares are:

$$MS_{\text{Treatments}} = \frac{SS_{\text{Treatments}}}{a-1}$$

$$MS_{\text{Blocks}} = \frac{SS_{\text{Blocks}}}{b-1}$$

$$MS_E = \frac{SS_E}{(a-1)(b-1)}$$

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0
Treatments	$SS_{\text{Treatments}}$	$a - 1$	$\frac{SS_{\text{Treatments}}}{a-1}$	$\frac{MS_{\text{Treatments}}}{MS_E}$
Blocks	SS_{Blocks}	$b - 1$	$\frac{SS_{\text{Blocks}}}{b-1}$	
Error	SS_E (by subtraction)	$(a-1)(b-1)$	$\frac{SS_E}{(a-1)(b-1)}$	
Total	SS_T	$ab - 1$		

$$\Rightarrow F_0 = 2{,}0357$$

is true. We would reject the null hypothesis at the α-level of significance if the computed value of the test statistic in Equation 13.23 is $f_0 > f_{\alpha,a-1,(a-1)(b-1)}$.

⟶ Book 1371

$\alpha = 1\%$, $a = 4$, $b = 4$

$\Rightarrow f_{\alpha,3,9} = 6{,}99 > F_0$

$\Rightarrow$ We fail to reject H_0

$\Rightarrow$ There's no difference

Is there any significant difference in the number of late arrivals among different days of the week at the significance level $\alpha = 1\%$?