

Chapter 11 ANOVA exercise

1. Please conduct ANOVA analysis of the following data

Here is some hypothetical data with a small enough N to keep your calculations fairly simple. Imagine a sample of 12 teenagers from three high schools—Washington, Adams, and Jefferson. Here are the school each teenager attends and the number of school days missed during a school year.

<u>School Attended</u>	<u>Number of Days Missed</u>
Washington	6
Washington	5
Washington	7
Adams	5
Adams	4
Adams	6
Adams	5
Jefferson	2
Jefferson	1
Jefferson	3
Jefferson	1
Jefferson	3

1) null hypothesis: school districts do not affect the days of missing

2) calculating all the means

$$\bar{X}_T = \frac{\sum X_i}{N} = \frac{6 + 5 + 7 + 5 + 4 + 6 + 5 + 2 + 1 + 3 + 1 + 3}{12} = 4$$

$$\bar{X}_{Washington} = \frac{6 + 5 + 7}{3} = 6$$

$$\bar{X}_{Adams} = \frac{5 + 4 + 6 + 5}{4} = 5$$

$$\bar{X}_{Jefferson} = \frac{2 + 1 + 3 + 1 + 3}{5} = 2$$

3) calculating SS_{Total} and $SS_{Between}$

$$(SS_{Total}) = \sum (X_i - \bar{X}_T)^2 = 44$$

$$(SS_{Total}) = (6 - 4)^2 + (5 - 4)^2 + (7 - 4)^2 + (5 - 4)^2 + (4 - 4)^2 + (6 - 4)^2 + (5 - 4)^2 \\ + (2 - 4)^2 + (1 - 4)^2 + (3 - 4)^2 + (1 - 4)^2 + (3 - 4)^2$$

$$(SS_{Between}) = \sum (\bar{X}_G - \bar{X}_T)^2 \times N_G$$

$$(SS_{Between}) = (6 - 4)^2 \times 3 + (5 - 4)^2 \times 4 + (2 - 4)^2 \times 5 = 36$$

4) calculating SS_{within}

$$\sum (X_i - \bar{X}_G)^2 = \sum (X_i - \bar{X}_T)^2 - \sum (\bar{X}_G - \bar{X}_T)^2 \times N_G$$

$$SS_{within} = SS_{Total} - SS_{Between} \Rightarrow SS_{within} = 44 - 36 = 8$$

5) calculating the dfs

$$df_{between} = k - 1 \quad k: \text{the total number of groups in the independent variable}$$

$$df_{within} = N - k \quad N: \text{the total number of cases in the sample}$$

$$df_{between} = 3 - 1 = 2$$

$$df_{within} = 12 - 3 = 9$$

6) calculating the mean sum of square MSS

$$\text{MSS for between: } MSS_{between} = \frac{SS_{between}}{(k-1)} = \frac{36}{2} = 18$$

$$\text{MSS for within: } MSS_{within} = \frac{SS_{within}}{(N-k)} = \frac{8}{9} = .89$$

7) calculating the f ratio

$$F_{df_{between}, df_{within}} = \frac{MSS_{between}}{MSS_{within}} = \frac{18}{.89} = 20.2$$

8) determining the p level

$$P < .01$$

9) decision regarding the null hypothesis, type of error committed

Reject the null hypothesis, committing type I error.

10) calculating eta-square (E^2)

$$E^2 = \frac{SS_{between}}{SS_{total}} = \frac{36}{44} = 81.8\%$$

11) interpret eta-square

Eta-square is PRE: knowing independent variable reduces errors in estimating the value of dependent variable by X%,

knowing school districts reduces errors in estimating the school missing days by 81.8%.

ANOVA Table

	Sum of square	df	MSS	f	p	Eta-square
Between	36	2	18	20.2	P<.01	81.8%
Within	$44 - 36 = 8$	9	.89			
total	44					