

Problem 1

Is gender independent of education level? A random sample of 395 people were surveyed and each person was asked to report the highest education level they obtained. The data that resulted from the survey is summarized in the following table:

	High School	Bachelors	Masters	Ph.d.	Total
Female	60	54	46	41	201
Male	40	44	53	57	194
Total	100	98	99	98	395

Question: Are gender and education level dependent at 5% level of significance? In other words, given the data collected above, is there a relationship between the gender of an individual and the level of education that they have obtained?

Solution:

Expected values

	High School	Bachelors	Masters	Ph.d.	Total
Female	50.886	49.868	50.377	49.868	201
Male	49.114	48.132	48.623	48.132	194
Total	100	98	99	98	395

So, working this out,

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}} = 1.632 + 1.691 + 0.342 + 0.354 + 0.380 + 0.394 + 1.576 + 1.633$$

$$\chi^2 = 8.006$$

The critical value of with 3 degree of freedom is 7.815. Since $8.006 > 7.815$, we reject the null hypothesis and conclude that the education level depends on gender at a 5% level of significance.

Problem Statement 2:

Using the following data, perform a oneway analysis of variance using $\alpha=.05$. Write up the results in APA format.

[Group1: 51, 45, 33, 45, 67]

[Group2: 23, 43, 23, 43, 45]

[Group3: 56, 76, 74, 87, 56]

Solution:

Group 1

Value	Mean	SD	SD^2
51	48.2	2.8	7.84
45	48.2	-3.2	10.24
33	48.2	-15.2	231.04
45	48.2	-3.2	10.24
67	48.2	18.8	353.44
Sum			612.8

Group 2

Value	Mean	SD	SD^2
23	35.4	-12.4	153.76
43	35.4	7.6	57.76
23	35.4	-12.4	153.76
43	35.4	7.6	57.76
45	35.4	9.6	92.16
Sum			512.2

Group 3

Value	Mean	SD	SD^2
56	69.8	-13.8	190.44
76	69.4	6.2	38.44
74	69.4	4.2	17.64
87	69.4	17.2	295.84
56	69.4	-13.8	190.44
Sum			732.8

$$\text{Var1} = 612.8 / 5 - 1 = 153.2$$

$$\text{Var2} = 512.2 / 5 - 1 = 128.8$$

$$\text{Var3} = 732.8 / 5 - 1 = 183.2$$

$$\text{MS}_{\text{error}} = 153.2 + 128.8 + 183.23 = 155.07$$

$$\text{Df}_{\text{error}} = 15 - 3 = 12$$

$$\text{SS}_{\text{error}} = (155.07)(15 - 3) = 1860.8$$

$$\text{Grand mean } (\bar{x}_{\text{grand}}) = 48.2 + 35.4 + 69.83 = 51.13$$

Group mean	Grand Mean	SD	SD^2
48.2	51.13	-2.93	8.58
35.4	51.13	-15.73	247.43
69.83	51.13	18.67	348.57

Sum of squares (SS_{means}) = 604.58

$Var_{\text{means}} = 604.583 - 1 = 302.29$

$MS_{\text{between}} = (302.29)(5) = 1511.45$

Calculating the remaining between (or group) terms of the ANOVA table:

$Df_{\text{groups}} = 3 - 1 = 2$

$SS_{\text{group}} = (1511.45)(3 - 1) = 3022.9$

Test statistic and critical value

$F = 1511.45 / 155.07 = 9.75$

$F_{\text{critical}}(2, 12) = 3.89$

Reject H_0 null hypothesis.

Source	SS	Df	MS	F
Group	3022.9	2	1511.45	9.75
Error	1860.8	12	155.07	
Total	4883.7			

Effect size

$\eta^2 = 3022.9 / 4883.7 = 0.62$

APA

$F(2, 12) = 9.75, p < 0.05, \eta^2 = 0.62.$

Problem Statement 3:

Calculate F Test for given 10, 20, 30, 40, 50 and 5, 10, 15, 20, 25.

Solution:

Set 1

10, 20, 30, 40, 50

$Mean(x_1) = 10 + 20 + 30 + 40 + 50 / 5 = 150 / 5$

$Mean(x_1) = 30$

$SD_1 = \sqrt{1/(N-1) * ((x_1 - x_m)^2 + (x_2 - x_m)^2 + \dots + (x_n - x_m)^2)}$

$= \sqrt{1/(5-1) * ((10-30)^2 + (20-30)^2 + (30-30)^2 + (40-30)^2 + (50-30)^2)}$

$= \sqrt{1/4 * ((-20)^2 + (-10)^2 + (0)^2 + (10)^2 + (20)^2)}$

$= \sqrt{1/4 * ((400) + (100) + (0) + (100) + (400))}$

$$=\text{sqrt}(250)$$

$$=15.8114$$

$$\text{Variance1}=\text{SD1}^2$$

$$\text{Variance1}=15.8114^2$$

$$\text{Variance1}=250$$

Set 2

For 5, 10, 15, 20, 25:

$$\text{Mean}(x_2) = 5 + 10 + 15 + 20 + 25 / 5$$

$$\text{Mean}(x_2) = 75/5$$

$$\text{Mean}(x_2) = 15$$

$$\text{SD2} = \text{sqrt}(1/(N-1) * ((x_1 - x_m)^2 + (x_2 - x_m)^2 + \dots + (x_n - x_m)^2))$$

$$=\text{sqrt}(1/(5-1)((5-15)^2 + (10-15)^2 + (15-15)^2 + (20-15)^2 + (25-15)^2))$$

$$=\text{sqrt}(1/4((-10)^2 + (-5)^2 + (0)^2 + (5)^2 + (10)^2))$$

$$=\text{sqrt}(1/4((100) + (25) + (0) + (25) + (100)))$$

$$=\text{sqrt}(62.5)$$

$$=7.9057$$

$$\text{Variance2}=\text{SD2}^2$$

$$\text{Variance2}=7.9057^2$$

$$\text{Variance2}=62.5$$

To calculate F Test

$$\text{F Test} = (\text{variance of } 10, 20, 30, 40, 50) / (\text{variance of } 5, 10, 15, 20, 25)$$

$$= 250/62.5$$

$$= 4.$$

The F Test value is 4.