

## TASK 2

### Assignment 2.1

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	Phd	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*

**H<sub>0</sub>:** Gender and education independent

**H<sub>1</sub>:** Gender and education dependent

The Expected frequency under the null hypothesis is given

$E = \text{row total} * \text{column total} / \text{sample size}$

Row totals and column totals are given above sample size = 395

Therefore, we have expected frequencies as follows:

High	School	Bachelors	Masters	Phd
Female	50.886	49.868	50.377	49.868
Male	49.114	48.132	48.623	48.132

Therefore,  $X^2 = 1.632 + 0.342 + 0.38 + 1.577 + 1.691 + 0.355 + 0.394 + 1.634 = 8.006$

$X^2_{0.95}$  for degrees of freedom 3 = 7.815.

Since  $8.006 > 7.815$ , we reject the null hypothesis. Therefore, gender and education are dependent

## Assignment 2.2

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

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$H_1$ : Population means are not equal

$$\text{Sum of squares between Treatments} = \sum n_i(\bar{X}_i - \bar{X})^2$$

$$\text{Sum of squares (Error)} = \sum \sum (X - \bar{X}_i)^2$$

$$\text{Test-statistic} = \sum n_i(\bar{X}_i - \bar{X})^2 / (k - 1) / \sum \sum (X - \bar{X}_i)^2 / (N - k) \text{ where } k = 3, N = 15$$

Source of Variation Sums of Squares Degrees of Freedom Mean Squares F between

Treatments: 3022.93 21511.47 9.75

Error or Residual 1860.8 12 155.07

$$F_{0.5}(2,12)=3.885$$

Since,  $9.75 > 3.885$ , We Reject the null hypothesis.

Therefore, the population means are not equal.

In APA format, A one-way between subjects ANOVA was conducted to compare the effect of the group on values in group 1, group 2 and group 3. There was a significant effect of group on values at  $\alpha = 0.05$  level for group 1, group 2, group 3.

## Assignment 2.3

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

## Solution

$$\text{Sample variance } s^2 = \sum (X - \bar{X})^2 / N - 1$$

$$s_1^2 = (400.0 + 100.0 + 0.0 + 100.0 + 400.0) / 4 = 250.0$$

$$s_2^2 = (100.0 + 25.0 + 0.0 + 25.0 + 100.0) / 4 = 62.5$$

$$F\text{-statistic} = s_1^2 / s_2^2 = 250 / 62.5 = 4$$