# Problem Statement 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:

As the goal is to find if there is a relationship between gender and education level, we can solve this with **Chi Squared** value at 5% significance level.

H0 (Null Hypothesis): Gender and Education level are independent

H1 (Alternate Hypothesis): Gender and Education level are dependent

We know, Chi Squared

χ2 = ∑(fo – fe)2/fe

Where,

Fo = Observed value

Fe = Expected value

Fe = Column total \* Row total / Grand Total

We need to find Fe for each report.

Example:

* There are 60 females observed in High School. Hence Fo = 60
* Fe = 201\*100/395 = 50.88608

The table can be rewritten as,

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **High** | **School** | **Bachelors** | **Masters** | **Ph.d.** | **Total** |
| **Female** | 60 50.88608 | 54 49.86835 | 46 50.37722 | 41 49.86835 | 201 |
| **Male** | 40 49.11392 | 44 48.13165 | 53 48.62279 | 57 48.13165 | 194 |
| **Total** | 100 | 98 | 99 | 98 | 395 |

Chi Squared = (60-50.88608)2/50.88608 + (54-49.86835)2/49.86835 + (46-50.37722)2/50.37722 +

(41-49.86835)2/49.86835 + (40-49.11392)2/49.11392 + (44-48.13165)2/48.13165 +

(53-48.62279)2/48.62279 + (57-48.13165)2/48.13165

**Chi Squared = 8.006061**

The Chi Squared value at 5% significance level (found from the table) is **7.815**

**Conclusion:**

As the obtained Chi Squared value is greater than the critical Chi Squared value, we can reject the null and can conclude that gender is **dependent** on the education level at 5% significance level.

# Problem Statement 2

Using the following data, perform a oneway analysis of variance using α=.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:

F = Mean Square (Between Groups) / Mean Square (Within Groups)

Mean Square (Between) = Sum of Squares (Between) / Degrees of freedom (Between)

Mean Square (Within) = Sum of Squares (Within) / Degrees of freedom (Within)

### Anova Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sum of Squares** | **df** | **Mean Square** | **F** |
| **Between Groups** | 3022.933333 | 2 | 1511.466667 | 9.747206 |
| **Within Groups** | 1860.8 | 12 | 155.0666667 |  |
| **Total** | 4883.733333 | 14 |  |  |

Proportion of total variation, Cohen’s d (η2) = Sum of Squares(Between Groups)/Sum of Squares(Total)

=3022.933333/4883.733333

=0.61898

### APA (American Psychological Association) Style:

F (2, 12) = 9.75, p <0.05, η2 = 0.61898

# Problem Statement 3:

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

## Solution:

F = Mean Square (Between Groups) / Mean Square (Within Groups)

Mean Square (Between) = Sum of Squares (Between) / Degrees of freedom (Between)

Mean Square (Within) = Sum of Squares (Within) / Degrees of freedom (Within)

### ANOVA table:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sum of Squares** | **df** | **Mean Square** | **F** |
| **Between Groups** | 562.5 | 1 | 562.5 | 3.6 |
| **Within Groups** | 1250 | 8 | 156.25 |  |
| **Total** | 1812.5 | 14 |  |  |

**F (1, 8) = 3.6**