**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?**

**Ans**

## Null Hypothesis H0: No relation between Gender and education

## Alternate Hypothesis H1: There is a relation between Gender and education

## The Expected frequency under the null hypothesis is given

E = row total \* column total / sample size

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

Therefore, we have expected frequencies as follows:

Gender High-School Bachelors Masters Ph.d.  
Female 50.886 49.868 50.377 49.868   
Male 49.114 48.132 48.623 48.132

Therefore, Χ2 = 1.632 + 0.342 + 0.38 + 1.577 + 1.691 + 0.355 + 0.394 + 1.634 = 8.006

The critical value of χ2 with 3 degree of freedom is 7.815.

Since 8.006 > 7.815, we reject the null hypothesis and conclude that gender and education are dependent on a5% 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]**

**Ans**

Null Hypothesis H0 : Means are equal for Group1,Group2 and Group 3

Alternate Hypothesis H1: Means are different

Sample means for Group1=48.2, Group 2 = 35.4, Group 3 = 69.8

Variance for Group 1 = 153.2, Group 2 = 128.8, Group 3 = 183.2

Mean Squared Error = SUM(Variance for three groups)/3 = 155.07

dferror = 15 – 3 = 12

dfgroups = 3 -1 = 2

If we compute the F statistic and its critical value

F- Statistic = MeanSqaurebetween/MeanSquared Error

= 1511.45/155.07 = 9.75

Fcritical(2,12) = 3.89

We got 9.75 > 3.89, so we reject null hypothesis

We conclude that means are different for each of these groups

**ANOVA table**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **source** | **SS** | **df** | **MS** | **F** |
| group | 3023 | 2 | 1511.5 | 9.75 |
| error | 1861 | 12 | 155.07 |  |
| total | 4884 |  |  |  |

**Effect size**

**n2 = 3023/4884 = 0.62**

**APA writeup**

*Fcrit*(2, 12)=9.75, *p* <0.05, **n2** **=0.62**

**Problem Statement 3:**

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

**For 10, 20, 30, 40, 50:**

**Ans**

We must compute Variance of these two groups first and divide to get F Test statistic

F Test = Variance for first group/Variance of second group

**For 10,20,30,40,50**

Mean = 30

Variance = Sum(x1-xm)2/N-1

= (10-30)2+…..(50-30)2/4 = 250

**For 5,10,15,20,25**

Mean = 15

Variance = Sum(x1-xm)2/N-1

**=** (5-15)2**+…..**(25-15)2/4 **= 62.5**

**F- Test = Variance of first group/Variance of second group**

**= 250/62.5**

**= 4**