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

H0: Gender and education independent

H1: Gender and education dependent

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 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 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| E = row total \* column total / sample size | | | | |
|  | High School | Bachelors | Masters | Ph.d. |
| Female | 50.88608 | 49.86835 | 50.37722 | 49.86835 |
| Male | 49.11392 | 48.13165 | 48.62278 | 48.13165 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| O | E | O-E | (O — E)2 | (O — E)2/ E |
| 60 | 50.88608 | 9.113924 | 83.06361 | 1.632345 |
| 54 | 49.86835 | 4.131646 | 17.0705 | 0.342311 |
| 46 | 50.37722 | -4.37722 | 19.16001 | 0.380331 |
| 41 | 49.86835 | -8.86835 | 78.64771 | 1.577107 |
|  |  |  |  |  |
| 40 | 49.11392 | -9.11392 | 83.06361 | 1.691244 |
| 44 | 48.13165 | -4.13165 | 17.0705 | 0.354663 |
| 53 | 48.62278 | 4.377215 | 19.16001 | 0.394054 |
| 57 | 48.13165 | 8.868354 | 78.64771 | 1.634012 |
|  | Total | | | 8.006066 |

The degrees of freedom for a test of independence equals the product of the number of categories in each variable minus 1. In this case we have a 2x4 table so df = 1x3 = 3.

Chi Square distribution table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Df | 0.5 | 0.10 | 0.05 | 0.02 | 0.01 | 0.001 |
| 1 | 0.455 | 2.706 | 3.841 | 5.412 | 6.635 | 10.827 |
| 2 | 1.386 | 4.605 | 5.991 | 7.824 | 9.210 | 13.815 |
| 3 | 2.366 | 6.251 | 7.815 | 9.837 | 11.345 | 16.268 |
| 4 | 3.357 | 7.779 | 9.488 | 11.668 | 13.277 | 18.465 |
| 5 | 4.351 | 9.236 | 11.070 | 13.388 | 15.086 | 20.517 |

Critical Value= 7.815

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

**Problem Statement 2:**

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]

# H0:μ1 =μ2 =μ3

# H1: Population means are not equal

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Group1 | Group2 | Group3 | Group1 Var | Group2 Var | Group2 Var | Group1 SD | Group2 SD | Group3 SD |  |
|  | 51 | 23 | 56 | 2.8 | -12 | -14 | 7.8 | 154 | 190 |  |
|  | 45 | 43 | 76 | -3 | 7.6 | 6.2 | 10 | 58 | 38.4 |  |
|  | 33 | 23 | 74 | -15 | -12 | 4.2 | 231 | 154 | 17.6 |  |
|  | 45 | 43 | 87 | -3 | 7.6 | 17.2 | 10 | 58 | 296 |  |
|  | 67 | 45 | 56 | 19 | 9.6 | -14 | 353 | 92 | 190 |  |
|  |  |  |  |  |  |  |  |  |  |  |
|  | 241 | 177 | 349 |  |  |  |  |  |  |  |
| Mean | 48.2 | 35.4 | 70 |  |  | Group(S) SD | 613 | 515 | 733 |  |
|  |  |  |  |  |  | Group(S) Var | 153 | 129 | 183 | 465.2 |
|  |  |  |  |  |  |  |  |  |  | 155.0666667 |

Calculating the remaining error (or within) terms for the ANOVA table:

dferror=15−3=12

SSerror=(155.07)(15−3)=1860.8

Intermediate steps in calculating the variance of the sample means:

Grand mean (x¯grand) = 48.2+35.4+69.83=51.13

|  |  |  |  |
| --- | --- | --- | --- |
| Grp. Mean | grand mean |  |  |
| 48.2 | 51.13 | -2.93 | 8.5849 |
| 35.4 | 51.13 | -15.73 | 247.4329 |
| 69.8 | 51.13 | 18.67 | 348.5689 |
|  |  |  | 604.5867 |

Varmeans=604.58/(3−1)=302.29

MSbetween=(302.29)(5)=1511.45 Note: This method of estimating the variance IS sensitive to group mean differences!

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

dfgroups=3−1=2

SSgroup=(1511.45)(3−1)=3022.9

Test statistic and critical value

F=1511.45155.07=9.75F=1511.45155.07=9.75

Fcritical(2,12)=3.89

Decision: reject H0

**Problem Statement3:**

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

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Group1 | Group2 | Grp1 Var | Grp2 Var | Grp1 SD | Grp2 SD |
|  | 10 | 5 | -20 | -10 | 400 | 100 |
|  | 20 | 10 | -10 | -5 | 100 | 25 |
|  | 30 | 15 | 0 | 0 | 0 | 0 |
|  | 40 | 20 | 10 | 5 | 100 | 25 |
|  | 50 | 25 | 20 | 10 | 400 | 100 |
|  | 150 | 75 |  |  |  |  |
| Mean | 30 | 15 |  | Group(S) SD | 1000 | 250 |
|  |  |  |  | Group(S) Var | 250 | 62.5 |
|  |  |  |  |  |  |  |
|  |  |  |  | F-Stat | 4.00 |  |

**F-statistic is 4**