**IO 519 Statistics**

**Chi-Square Demonstration: Fakenet dataset**

**The Chicago School of Professional Psychology**

*Open* ***FakenetWeight.sav*** *in SPSS.*

Analyze Descriptive Statistics Crosstabs

* Select **Gender** and drag into rows
* Select **Suicide Outcome** into columns
* Select *Exact*
  + Check *Exact*
  + Check *Time limit per test*
  + Enter *5* minutes
* Select *Statistics*
  + Check *Chi-square*
  + Check *Contingency coefficient*
  + Check *Phi and Cramer’s V*
  + Select *Continue*
* Select *Cells*
  + Check *Observed*
  + Check *Expected*
  + Check *Row*
  + Check *Column*
  + Check *Total*
  + Check *Round cell counts*
  + Check *Compare column proportions*
  + Check *Adjusted p-values*
  + Check *Standardized*
  + Select *Continue*
* Select *OK*

How does the Chi-Square Test Statistic on SPSS compare to what we hand calculated?

What about degrees of freedom?

Did we get the significance values right?

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Chi-Square Tests** | | | | | | | |
|  | Value | df | Asymptotic Significance (2-sided) | Exact Sig. (2-sided) | Exact Sig. (1-sided) | Point Probability |
| Pearson Chi-Square | 25.356a | 1 | .000 | .000 | .000 |  |
| Continuity Correctionb | 23.520 | 1 | .000 |  |  |  |
| Likelihood Ratio | 24.932 | 1 | .000 | .000 | .000 |  |
| Fisher's Exact Test |  |  |  | .000 | .000 |  |
| Linear-by-Linear Association | 25.229c | 1 | .000 | .000 | .000 | .000 |
| N of Valid Cases | 200 |  |  |  |  |  |
| a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 14.44. | | | | | | | |
| b. Computed only for a 2x2 table | | | | | | | |
| c. The standardized statistic is 5.023. | | | | | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gender \* Suicide Outcome Crosstabulation** | | | | | |
|  | | | Suicide Outcome | | Total |
| Suicide Attempt | Suicide Success |
| Gender | Male | Count | 28a | 48b | 76 |
| Expected Count | 14.4 | 61.6 | 76.0 |
| % within Gender | 36.8% | 63.2% | 100.0% |
| % within Suicide Outcome | 73.7% | 29.6% | 38.0% |
| % of Total | 14.0% | 24.0% | 38.0% |
| Standardized Residual | 3.6 | -1.7 |  |
| Female | Count | 10a | 114b | 124 |
| Expected Count | 23.6 | 100.4 | 124.0 |
| % within Gender | 8.1% | 91.9% | 100.0% |
| % within Suicide Outcome | 26.3% | 70.4% | 62.0% |
| % of Total | 5.0% | 57.0% | 62.0% |
| Standardized Residual | -2.8 | 1.4 |  |
| Total | | Count | 38 | 162 | 200 |
| Expected Count | 38.0 | 162.0 | 200.0 |
| % within Gender | 19.0% | 81.0% | 100.0% |
| % within Suicide Outcome | 100.0% | 100.0% | 100.0% |
| % of Total | 19.0% | 81.0% | 100.0% |
| Each subscript letter denotes a subset of Suicide Outcome categories whose column proportions do not differ significantly from each other at the .05 level. | | | | | |

1. Has the expected frequencies assumption been met? How do you know? (top of pg 741).

*Yes assumption is met*

*All expected frequencies are greater than 5.*

1. What does the Pearson’s chi-square test examine? (pg 742)

*Whether there is an association between two categorical variables, whether two variables are independent*

1. Is our chi-square test significant?

*Yes, it is significant (p < .001)*

1. Will we accept or reject our hypothesis that the variables are independent? (pg 742)

*(Our null hypothesis was independence. The alternative hypothesis for a chi-square test is that variables are dependent/associated with one another.)*

*We reject the hypothesis meaning that variables are in some way related.*

1. What do our significant results indicate? (pg 742)

*The highly significant result indicates that there is an association between gender and whether they successfully committed suicide.*

1. Using the crosstabulation table, compute and interpret the odds ratio. (pg 744-745)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Gender \* Suicide Outcome Crosstabulation** | | | | | |
| A | | | Suicide Outcome | | Total |
| Suicide Attempt | Suicide Success |
| Gender | Male | Count | 28a | 48b | 76  B |
| Expected Count | 14.4 | 61.6 | 76.0 |
| % within Gender | 36.8% | 63.2% | 100.0% |
| % within Suicide Outcome | 73.7% | 29.6% | 38.0% |
| % of Total | 14.0% | 24.0% | 38.0% |
| Standardized Residual  C | 3.6 | -1.7 | D |
| Female | Count | 10a | 114b | 124 |
| Expected Count | 23.6 | 100.4 | 124.0 |
| % within Gender | 8.1% | 91.9% | 100.0% |
| % within Suicide Outcome | 26.3% | 70.4% | 62.0% |
| % of Total | 5.0% | 57.0% | 62.0% |
| Standardized Residual | -2.8 | 1.4 |  |
| Total | | Count | 38 | 162 | 200 |
| Expected Count | 38.0 | 162.0 | 200.0 |
| % within Gender | 19.0% | 81.0% | 100.0% |
| % within Suicide Outcome | 100.0% | 100.0% | 100.0% |
| % of Total | 19.0% | 81.0% | 100.0% |
| Each subscript letter denotes a subset of Suicide Outcome categories whose column proportions do not differ significantly from each other at the .05 level. | | | | | |

*Now we use the Odds ratio formula as follows*

*Odds ratio*

*= ( a X d ) ÷ ( b X c )*

*= (Suicide Outcome Attempt MALE X Suicide Outcome Success FEMALE) / (Suicide Success MALE X Suicide Attempt FEMALE)*

*= ( 28 X 114 ) / ( 48 x 10 )*

*= (3192) / (480)*

*= 6.65*

*Male workers were 6.65 more likely to unsuccessfully attempt suicide than female workers at Fakenet.*

*What’s our recommendation to company? Do we hire more males or females? Who will be less likely to need replacement?*