Fakenet & Chi-square Exercise

Background: An international humanitarian organization reported unfair labor conditions at Fakenet corporation, an electronics manufacturing company known for making Apple circuit boards and hard drives for iPhone, iMac, and Macbook computers. Activists cited reports that the number of workplace suicides were through the roof. An investigation was launched into Fakenet’s labor practices, which confirmed the initial reports that employees attempted suicide by jumping from the rooftops at Fakenet.

You’re all going to receive some data with numbers of suicides either attempted or succeeded. Our job is to analyze these numbers for an association between gender and suicide outcomes. Note: if an employee attempted suicide, that individual was transferred to a factory in a one-story building so it would not happen again. We assume each attempt was independent. That is, a person attempting suicide is only in one column and did not, say, attempt twice and succeed on the third try. Each person listed had only one opportunity to arrive at a suicide outcome.

Predictor variable: Gender (male vs. female)

Outcome variable: Suicide outcome (attempt vs. success)

Sample size: 200

Research question: Is there an association between gender and suicide outcome?

Null hypothesis: There is no association; gender is independent of suicide outcome.

Alternative hypothesis: Reject the null hypothesis, finding support that there may be association between variables.

Department 1: Apple iPhone Circuit Boards *(Contingency table)*

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Suicide Attempt | Suicide Success | Total |
| Male | 9 | 14 | 23 |
| Female |  | 28 | 30 |
| Total | 11 | 42 | 53 |

Department 2: Asus televisions *(Contingency table)*

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Suicide Attempt | Suicide Success | Total |
| Male | 5 | 10 | 15 |
| Female |  | 28 | 31 |
| Total | 8 | 38 | 46 |

Department 3: Beats Headphones *(Contingency table)*

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Suicide Attempt | Suicide Success | Total |
| Male | 8 |  | 16 |
| Female | 4 | 29 | 33 |
| Total | 12 | 37 | 49 |

Department 4: Generic Hard Drives *(Contingency table)*

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Suicide Attempt | Suicide Success | Total |
| Male | 6 | 16 | 22 |
| Female | 1 | 29 | 30 |
| Total |  | 45 | 52 |

Contingency table showing gender and suicide outcome across Fakenet departments.

|  |  |  |  |
| --- | --- | --- | --- |
| Gender | Suicide Attempt | Suicide Success | Total |
| Male | 28 | 48 | 76 |
| Female | 10 | 114 | 124 |
| Total | 38 | 162 | 200 |

(Review variables)

With contingency table, what do we do? Calculate Chi-square

We would usually calculate general fit of the model with total error:

Total error = ∑ni=1 (observedi – modelij)2

But we have categorical data so we use chi-square instead:

Chi-square = ∑ (observedij – modelij)2 / modelij

I = rows in contingency table, j = columns in contingency table

Observed data are the frequencies in the contingency table. We can’t calculate it yet because we do not have the model values. We calculate model values ourselves. If it was ANOVA, we would use group means. But we have two categorical variables so we can’t use group means. We compare “expected frequencies” with observed frequencies to serve as our model.

Modelij = Eij = Row totali X column totalj / *n*

Our *n* is number of observations. N = 200 Fakenet worker suicide outcomes.

Let’s calculate the model values so we can see how well our model fits.

modelAttempt, Male = RowTotalMale X ColumnTotalAttempt / *n* = 76 X 38/200 = 14.44

modelAttempt. Female = RowTotalFemale X ColumnTotalAttempt / *n* = 124 X 38/200 = 23.56

modelSuccess, Male = RowTotalMale X ColumnTotalSuccess / *n* = \_\_\_\_\_\_\_\_\_ X 162/200 = 61.56

modelSuccess. Female = RowTotalFemale X ColumnTotalSuccess / *n* = \_\_\_\_\_\_\_\_\_ X 162/200 = 100.44

Now that we have model values, we can calculate Chi-square to see if the model fit is significant:

Chi-square statistic = ∑ (observedij – modelij)2 / modelij

=(28–14.44)2/\_\_\_\_\_ + (10–23.56)2/23.56 + (\_\_\_–61.56)2/61.56 + (114–100.44)2/\_\_\_\_\_\_

= 12.73 + 7.80 + 2.99 + \_\_\_\_\_\_\_\_\_\_

= 25.\_\_

(Our chi-square statistic is \_\_\_\_\_\_\_\_\_\_)

We need to calculate *degrees of freedom* before we find out if our model fit is significant:

Df = (r-1)(c-1)

= (2-\_\_\_)(2-1)

= (1)(\_\_\_\_\_\_)

= 1

Now we look at either SPSS output or the Appendix in the Field book with critical values. Based on the degrees of freedom, we know what minimum chi-square test statistic we need to be able to say our model is significant at either p = .05 of p = .01.

Here, we look at Field and see critical value Df of 1 = 3.84 at p - .05; and 6.63 at p = .01.

Our observed chi-square is (higher/lower) \_\_\_\_\_\_\_\_\_\_\_ so we conclude it is (significant/not significant) \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ at p < .01.

The CEO (will/will not) \_\_\_\_\_\_\_\_\_\_\_\_\_ be pleased! We have done what is asked by company. Our modesty will be rewarded.

Analysis

Does it meet our assumptions?

#1 We need Expected Frequencies above 5 for each box in the contingency table. (Ours is fine.)

#2 We need independent observations, which we know to be true because employees who attempted suicide were transferred to that one-story factory so they could not attempt it again.

We can conclude that we meet our assumptions. Keep this document nearby. Now you will write it into SPSS to compare outputs. We will send these over to the executive suite after we make sure they look okay. We would bring shame to company by having errors in our calculations.

SPSS data set: FakenetWeight.sav