**Chi Square Example (Handcalculations) Spring 2017**

Question: Does a person's major have any signficant influence over the region of the country in which they work? (in other words--are the 2 variables “DEPENDENT”?)

MAJOR freq expected for each cell:

observed I |cell 1 = (42)(75)/138 = 22.826

expected I |cell 2 = (42)(63)/138 = 19.174

Col % I NonCJ CJ ROW |cell 3 = (32)(75)/138 = 17.391

chi sq obtI 0 I 1 I TOTAL |cell 4 = (32)(63)/138 = 14.609

REGION ‑‑‑‑‑‑‑‑+‑‑‑‑‑‑‑‑+‑‑‑‑‑‑‑‑+ |cell 5 = (64)(75)/138 = 34.783

1 I 26 I 16 I 42 |cell 6 = (64)(63)/138 = 29.217

Midwest I22.826 I19.174 I |

I 34.7% I 25.4% I |chi square for each cell:

I I I |cell 1 = (26 - 22.826)2 / 22.826 = 10.0743 / 22.826 = .4413

+‑‑‑‑‑‑‑‑+‑‑‑‑‑‑‑‑+ |cell 2 = (16 - 19.174)2 / 19.174 = 10.0743 / 19.174 = .5254

2 I 16 I 16 I 32 |cell 3 = (16 - 17.391)2 / 17.391 = 1.9349 / 17.391 = .1113

South I17.391 I14.609 I |cell 4 = (16 - 14.609)2 / 14.609 = 1.9349 / 14.609 = .1324

I 21.3% I 25.4% I |cell 5 = (33 - 34.783)2 / 34.783 = 3.1791 / 34.783 = .0914

I I I |cell 6 = (31 - 29.217)2 / 29.217 = 3.1791 / 29.217 = .1088

+‑‑‑‑‑‑‑‑+‑‑‑‑‑‑‑‑+ |chi square obtd =(.4413+.5254+.1113+.1324+.0914+.1088) =

3 I 33 I 31 I 64 | = 1.4106

Other I34.783 I29.217 I |

I 44.0% I 49.2% I |

I I I |

+‑‑‑‑‑‑‑‑+‑‑‑‑‑‑‑‑+ |

COLUMN 75 63 138 |

TOTAL

CHI‑SQUARE VALUE DF SIGNIFICANCE

‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑‑ ‑‑‑‑‑‑‑‑‑‑‑ ‑‑‑‑ ‑‑‑‑‑‑‑‑‑‑‑‑

# PEARSON 1.41064 2 .49395

LIKELIHOOD RATIO 1.42172 2 .49122

MANTEL‑HAENSZEL .95909 1 .32742

MINIMUM EXPECTED FREQUENCY ‑ 14.609

NUMBER OF MISSING OBSERVATIONS: 215

5 Step Model

1. Assumptions: independent random samples, nominal or short ordinal variables involved

2. Hypotheses: Null hypo says that the 2 variables, REGION and MAJOR, are independent of one another

Research hypo says that the 2 variables are dependent on one another

3. Picture: use chi square sampling distribution with an alpha of .05, degrees of freedom for chi square,

df = (# rows - 1)(# columns - 1) = (2)(1) = 2, therefore chi square critical = 5.991

4. Find chi square obtained or the chi square test statistic, first find frequency expected for each cell,

freq exp for a cell = (row marginal)(column marginal)/# of cases overall, N

chi square obtained = E (freq obs - freq exp)(freq obs - freq exp)/(freq exp)

5. Make a decision: (compare chi square critical to chi square obtained) Because chi square obtained is less than the chi square critical we conclude in saying that REGION and MAJOR are independent of each other. In other words, being a nonCJ or CJ major does not have any statistically significant influence on where a person works. The F obtained value is 1.41 which keeps us in the “independent” zone (within the 5.991 steps). Looking at the p-value or significance that corresponds to the 1.41, we see that 49% of the time, the connection between REGION and MAJOR is due to chance which is well beyond the 5% of the time we’re willing to allow our calculations to have in order to say that they are dependent on each other.