PSYC 60 – Statistics
Spring 2012
Notes #10

Chi-square - One Variable

Book: Chapter 19 - Test for Qualitative (Nominal) Data
Sections 19.1 - 19.5

Data

The professor predicts that his class will be made up of 10% freshman, 20% sophomores, 40% juniors and 30% seniors.

Example roster:

Name	Year
Tristan	Senior
Teresa	Sophomore
Ieva	Freshman
Richard	Senior

Roster totals:

Year	Expected	Observed in data:
Freshman		9
Sophomore		16
Junior		40
Senior		19

Were the actual class proportions significantly different than the professor's prediction?

Terms

Frequency =

Proportion =

Expected frequency =

Observed frequency =

Formula

X^2 =

Values for X^2

Exactly as expected

Small deviations from expected

Large deviations from expected

X^2 Distribution and Table

Building the distribution -

For a population with known proportions, take a sample and calculate the X^2 value. Repeat this many time. Build a distribution of X^2 values that occur _____.

Shape - Zero = _____ deviations, _____ by chance

Small values = _____ deviations, _____ by chance

Large values = _____ deviations, _____ by chance

For each distribution we have encountered, we've needed a table to judge the likelihood

of getting different values _____. In particular, we need to find a _____ using alpha.

Distributions/Tables that we use:

Statistical test

1. Hypotheses

2. Alpha
3. Critical value

Degrees of freedom (df) for one variable test =

1 vs 2 tailed

4. Test statistic

5. Decision

Chi-square - Two Variable

Book: Chapter 19 - Test for Qualitative (Nominal) Data
Sections 19.6 - 19.10

Data

Example roster:

Name	Gender	Major
Tristan	Male	Bio
Teresa	Female	Bio
Ieva	Female	Psych
Richard	Male	Psych

Roster totals:
11 male-biology
9 female-biology
38 male-psychology
52 female-psychology

Option 1: The professor predicts that the class will be 20% male-biology majors, 10% female-biology majors, 40% female-psychology majors and 30% male-psychology majors:
Were the actual class proportions significantly different than the professor's prediction?
Option 2: *Are the factors of gender and major indepedent?*

Expected frequencies for two indepedent factors

You commute to school 20 days a month. The weather at UCSD is sunny 80% of the days and cloudy the other 20%. Your parking budget limits you to driving 25% of the days and taking the shuttle the other 75%. If the weather did not determine how you got to school, how many days would you be driving in sunny weather?

	Sunny	Cloudy
Drive		
Shuttle		

Procedure:
1. Draw out the table
2. Find total for each column, each row, and grand total
3. Determine the % for each column/row
4. Determine the % for each cell by mulitplying _____ x _____
5. Determine the count for each cell by _____ x _____

If the weather did not determine how you commute, then these factors are _____

If you always tried to drive on cloudy days, what would be the new percentages for each cell? Is the weather still independent of how you commute?

	Sunny	Cloudy
Drive		
Shuttle		

Dependent

the value of one factor _____ or _____ another factor

Independent

Percentage for each cell is based only on percentage for _____ and _____

Two variables, but more than two values per variable
Weather: Cloudy, Rainy, Sun, Hot
Commute: Car, Shuttle, Bike

Degrees of freedom (df) for two variable test =

Test for independence
1. Hypothesis

2. Alpha
3. Critical value
4. Test statistic
5. Decision

Null Hypotheses
The expected values for the _____ hypothesis that we "test against" can be determine using several different conditions:

Expected values set by:

A. Individually specified
Example: Do students in PSYC 60 have the same proportion of majors as UCSD overall, which is 30% psyc, 20% bio, 15% communications, 15% engineering, 20% other.

B. All the same
Example: Are all UCSD colleges represented equally in PSYC 60?

C. Independence
Example: Are gender and pass/fail independent in PSYC 60?