Choosing the Correct Statistical Test
Chi-Square Analysis

February 20, 2006

## Choosing the Correct Statistical Test

- Knowing which statistical test to use in order to test the relationship between your independent and dependent variables depends on the 'type' of data that you have.

## Scales of Measurement

- A Variable is anything we measure
- Remember NOIR

## Continuous vs. Discreet Variables

- **Continuous Variables** – a variable that can represent any value on the scale used to measure the variable (typically interval or ratio variables)
  - Number of Cokes drink per week (1, 1.5, 2)
  - "how much" --- amount
- **Discreet Variables** – a variable that can represent only a few specified values on a designated scale (typically nominal, ordinal variables)
  - Male (1) or female (2)
  - "how many" --- type

## Measurement vs. Categorical Variables

- **Measurement Variables** – things to which we can assign an 'actual' number. Something we can measure, quantitative.
  - Age, height, weight
- **Categorical Variables** – things to which we may assign a number to, but the number represents type rather than amount, thus qualitative.
  - Race, gender
  - Dichotomous categorical variables – only two possible outcomes (yes or no; male or female)

CHOOSING THE RIGHT STATISTICAL TEST

| | | **Dependent Variable** | |
|---|---|---|---|
| | | **Measurement Continuous (Quantitative)** | **Categorical Discreet Qualitative** |
| **Independent Variable** | **Measurement Continuous (Quantitative)** | Regression<br>Correlation | Logistic Regression<br>Point-Biserial Correlation |
| | **Categorical Discreet Qualitative** | T-test<br>ANOVA<br>Regression<br>Point-Biserial Correlation | Chi-Square<br>Logistic Regression<br>Phi & Cramer's V |

So, what test do we use to test our primary hypothesis?

- Our independent variable is ….
- Our dependent variable is…

## Chi Square Analysis

- Chi square Analysis ($\chi^2$) provides probabilities based on frequencies.
- Thus, data consists of frequency counts of your variable categories.
- Chi square analysis compares the observed frequency counts to the frequency counts that would be expected if the null hypothesis was true (thus, chance)
- This is the statistical analysis to statistically tests for statistical significance.

## Chi Square Analysis

- **To calculate $X^2$, us this formula**
  $X^2 = \Sigma [(O - E)^2 / E]$
  - where O = observed frequency, or the number of events in each category
  - where E = expected frequency, use '**test of independence**' or 'goodness of fit' calculations
- Degrees of freedom in a chi square analysis
  - *df* = (R-1)(C-1)

## Chi Square Analysis – Finding the Expected Score

1. Test of Independence
   - Tests the null hypothesis that the two variables are independent, or that there is NO relationship to one another.
   - The most common test used for chi square analyses

## Chi square analyses

**Hypothetical data on attitudes toward war**

- 140 participants
- Male or Female – variable is…
- Acceptable or Unacceptable – variable is…

***Keep in mind in chi square, participants can only be in ONE of the dependent categories**

| | | War | |
|---|---|---|---|
| | Acceptable | Unacceptable | Σ |
| Men | 59 | 29 | **88** |
| Women | 15 | 37 | **52** |
| Σ | **74** | **66** | **140** |

## Chi-Square – Expected Frequencies

| | | War | |
|---|---|---|---|
| | Acceptable | Unacceptable | Σ |
| Men | 59 (46.51) | 29 (41.49) | **88** |
| Women | 15 (27.49) | 37 (24.51) | **52** |
| Σ | **74** | **66** | **140** |

**HOW DO WE GET THESE EXPECTED FREQUENCIES…FOLLOW THE MATH…**

## How do we get to the $X^2$ Statistic?

$X2 = \Sigma [(O - E)2 / E]$

| O | E | O-E | $(O-E)^2$ | $(O-E)^2/E$ |
|---|---|---|---|---|
| 59 | 46.51 | | | |
| 15 | 27.49 | | | |
| 29 | 41.49 | | | |
| 37 | 24.51 | | | |
| | | | | $X^2$ = |

## Chi Square Analysis

- Don't forget degrees of freedom
  - *df* = (R-1)(C-1)
- Critical Values for Chi Square analysis
  - Critical values table on page 360 (Table C.2) in your Cozby textbook
  - In general, $X^2_{.05}$ (1*df*) = 3.84, then you may reject your null hypothesis.

## Chi Square Analysis

- What can we conclude?

## Chi square analysis

- Effect Size
  - How MUCH difference is there?
  - We may know that the differences are statistically significant…but HOW different are the two groups…this is effect size (the **degree** of the relationship)
  - For chi square, we look at **Phi coefficient** or Cramer's V (if extends beyond a 2 x 2)

## Chi Square Analysis - Effect Size

- **Phi coefficient** (Φ) – used in 2 x 2 tables

  $\Phi = \sqrt{\chi^2/N}$

  Thus, our Phi coefficient in our example would be…
- What is a good effect size?
  - .10 – small effect
  - .30 – medium effect
  - .50 – large effect

## Limitations of Chi square

- Responses must be independent and mutually exclusive and exhaustive
  - Each case from a sample MUST fit into one and only one cell in our matrix