Chi-Square Significance Tests

Chi-square is a family of distributions commonly used for significance testing.

Key Concepts and Terms

- **Significance testing**, of which chi-square tests are a type, is also treated in a separate section.
- **Types of Chi-Square**
  - **Pearson's chi-square** is by far the most common type of chi-square significance test. If simply "chi-square" is mentioned, it is probably Pearson's chi-square. This statistic is used to test the hypothesis of no association of columns and rows in tabular data. It can be used even with nominal data. Note that chi square is more likely to establish significance to the extent that (1) the relationship is strong, (2) the sample size is large, and/or (3) the number of values of the two associated variables is large. A chi-square probability of .05 or less is commonly interpreted by social scientists as justification for rejecting the null hypothesis that the row variable is unrelated (that is, only randomly related) to the column variable. Its calculation is discussed below.
    - **Yates' correction** is an arbitrary, conservative adjustment to chi-square when applied to tables with one or more cells with frequencies less than five. It is only applied to 2 by 2 tables. Some authors also apply it to all 2 by 2 tables since the correction gives a better approximation to the binomial distribution. Yates' correction is conservative in the sense of making it more difficult to establish significance. SPSS. Some computer packages label Yates' correction as **continuity corrected chi-square** in their output. Its calculation is also discussed below.
  - **Chi-square goodness-of-fit test**. The goodness-of-fit test is simply a different use of Pearsonian chi-square. It is used to test if an observed distribution conforms to any other distribution, such as one based on theory (ex., if the observed distribution is not significantly different from a normal distribution) or one based on some other known distribution (ex., if the observed distribution is not significantly different from a known national distribution based on Census data). The Kolmogorov-Smirnov goodness-of-fit test is preferred for interval data, for which it is more powerful than chi-square goodness-of-fit.
  - **Likelihood ratio chi-square** is an alternative procedure to test the hypothesis of no association of columns and rows in nominal-level tabular data. It is supported by SPSS output and is based on maximum likelihood estimation. Though computed differently, likelihood ratio chi-square is interpreted the same way. For large samples, likelihood ratio chi-square will be close in results to Pearson chi-square. Even for smaller samples, it rarely leads to different substantive results.
  - **Mantel-Haenszel chi-square**, also called the *Mantel-Haenszel test for linear association* or *linear by linear association chi-square*, unlike ordinary and likelihood ratio chi-square, is an ordinal measure of significance. It is preferred when testing the significance of linear relationship between two ordinal variables because it is more powerful than Pearson chi-square (more likely to establish linear association). Mantel-

Haenzel chi-square is not appropriate for nominal variables. If found significant, the interpretation is that increases in one variable are associated with increases (or decreases for negative relationships) in the other greater than would be expected by chance of random sampling. Like other chi-square statistics, M-H chi-square should not be used with tables with small cell counts.

- **Stratified analysis**, also called blocked analysis and matched analysis, is a form of control variable analysis conducted with the Mantel-Haenszel coefficient. For each of k categories of a control variable (called the *stratification* variable), a 2-by-2 table is created for the independent and dependent variables. The stratification variable need not be ordinal but it is assumed that the row and column marginals be the same for each of the k 2-by-2 tables, a circumstance which occurs mainly in experimental situations. The Mantel-Haenszel chi-square coefficient tests whether the common odds ratio across the k strata is 1.0, indicating no effect of the stratification variable. SPSS provides a macro (mh.sps) for Mantel-Haenszel stratified analysis which outputs M-H chi-square and its significance..

- **SPSS Output**. To obtain chi-square in SPSS: Select Analyze, Descriptive Statistics, Crosstabs; select row and column variables; slick Statistics; select Chi-square.

  - *Chi-square with control variables*: In the context of crosstabulation, use of a control variable creates one subtable (similar to the overall table) for each value of the control variable. Evaluation of the subtable with chi-square is identical to evaluation of the main table. There is a control effect if at least one subtable is non-significant. In SPSS, move the control variable to the Layer 1 box when selecting variables. Use of multiple control variables is possible.

Assumptions

- **Random sample data** are assumed. As with all significance tests, if you have population data, then any table differences are real and therefore significant. If you have non-random sample data, significance cannot be established, though significance tests are nonetheless sometimes utilized as crude "rules of thumb" anyway.

- **A sufficiently large sample size** is assumed, as in all significance tests. Applying chi-square to small samples exposes the researcher to an unacceptable rate of Type II errors. There is no accepted cutoff. Some set the minimum sample size at 50, while others would allow as few as 20. Note chi-square must be calculated on actual count data, not substituting percentages, which would have the effect of pretending the sample size is 100.

- **Adequate cell sizes** are also assumed. Some require 5 or more, some require more than 5, and others require 10 or more. A common rule is 5 or more in all cells of a 2-by-2 table, and 5 or more in 80% of cells in larger tables, but no cells with zero count. When this assumption is not met, Yates' correction is applied.

- **Independence**. Observations must be independent. The same observation can only appear in one cell. This means chi-square cannot be used to test correlated data (ex., before-after, matched pairs, panel data).

- **Similar distribution**. Observations must have the same underlying distribution.

- **Known distribution**. The hypothesized distribution is specified in advance, so that the number of observations that are expected to appear each cell in the table can be calculated without reference to the observed values. Normally this expected value is the crossproduct of the row and column marginals divided by the sample size.

- **Non-directional hypotheses** are assumed. Chi-square tests the hypothesis that two variables are related only by chance. If a significant relationship is found, this is not equivalent to establishing the researcher's hypothesis that A causes B, or that B causes A.

- **Finite values**. Observations must be grouped in categories.

- **Normal distribution of deviations** (observed minus expected values) is assumed. Note chi-square is a *nonparametric test* in the sense that is does not assume the parameter of normal distribution for the data -- only for the deviations.

- **Data level**. No assumption is made about level of data. Nominal, ordinal, or interval data may be used with chi-square tests.

Frequently Asked Questions

- **How is Pearsonian chi-square calculated for tabular data?**
- **How is the chi-square goodness-of-fit test calculated to test if observed data are normally distributed?**
- **My statistics program prints out the chi-square contribution of each cell in the table. Can this be used to establish the significance of each cell?**
- **One of my variables is a multi-response item. What do I do?**

How is Pearsonian chi-square calculated for tabular data?

Calculation of this form of chi-square requires four steps:

1. *Computing the expected frequencies*. For each cell in the table, the expected frequency must be calculated. The expected frequency for a given column is the column total divided by n, the sample size. The expected frequency for a given row is the row total divided by n. The expected frequency for a cell is the column expectation times the row expectation times n. This formula reduces to the expected frequency for a given cell equaling its row total times its column total, divided by n.

2. *Application of the chi-square formula*. Let O be the observed value of each cell in a table. Let E be the expected value calculated in the previous step. For each cell, subtract E from O, then square the result and then divide by E. Do this for every cell and sum all the results. This is the chi-square value for the table. If Yates' correction for continuity is to be applied, due to cell counts below 5, the calculation is the same except for each cell, subtract an additional .5 from the difference of O - E, prior to squaring and then dividing by E. This reduces the size of the calculated chi-square value, making a finding of significance less likely -- a penalty deemed appropriate for tables with low counts in some cells.

3. *Calculate the degrees of freedom*. The chi-square value is not interpretable directly but must be compared to a *table of the chi-square distribution*. The columns of this table are alternative significance levels (.001, .01, .05, etc.) and

the rows are degrees of freedom (df). For a table, df = (r - 1)*(c - 1) , where r is the number of rows and c is the number of columns. That is, if you know the column and row totals, when all cells are filled in except one row and one column, these may be calculated from the information already given.

4. *Using the chi-square table*. A chi-square table, which in effect is built into statistical software packages, gives a *critical value*. The calculated chi-square value must be greater than the critical value to reject the null hypothesis that the row variable is unrelated to the column variable, at the level of significance selected by reading down the appropriate column in the chi-square table (ex., the .05 significance column). In practice, computer programs are used in place of chi-square tables, and computer printout shows the significance level (often labeled p) directly, but the interpretation is the same.

- **My statistics program prints out the chi-square contribution of each cell in the table. Can this be used to establish the significance of each cell?**
  No, a different procedure is needed for that purpose. However, cell contributions are useful in determining which ranges of the two variables depicted in a table are contributing the most to the overall relationship.

- **One of my variables is a multi-response item. What do I do?**
  One method is to compute a new nominal variable in which every combination of responses is a separate value (use the IF statements in SPSS), then you may use chi-square or nominal measures of association. Alternatively, you might consider using the SPSS CATEGORIES module after recoding your multi-response item as a set of separate variables. CATEGORIES then gives you the canonical correlation coefficient between your multiple response set as independents and another variable such as income level as dependent.

Bibliography

- Agresti, Alan (1996). *Introduction to categorical data analysis*. NY: John Wiley and Sons. Agresti discusses Mantel-Haenszel chi-square stratified analysis on pp. 231-236.

- Levin, Irwin P. (1999). *Relating statistics and experimental design*. Thousand Oaks, CA: Sage Publications. Quantitative Applications in the Social Sciences series #125. Elementary introduction covers t-tests and various simple ANOVA designs. Some additional discussion of chi-square, significance tests for correlation and regression. and non-parametric tests such as the runs test, median test, and Mann-Whitney U test.

- Lieberman, Bernhardt, ed. (1971). *Contemporary problems in statistics*. NY: Oxford. Section 5 deals with assumptions of chi-square procedures.

Back
