

APPLIED STATISTICS INSTRUCTION SHEET

THE ANALYSIS OF CONTINGENCY TABLES IN SPSS AND PSPP

Instructions

Contingency tables (also known as crosstabulations) are used when there are two categorical variables. The analysis has three components: the table itself, an analysis of the goodness of fit of the table, and measures of association describing the effect size and, if present, the direction of the relationship. The path to the procedure in both software packages is Analyze>Descriptive Statistics>Crosstabs.

Instructions

- Toggle the row and column variable into their respective boxes. It is customary to make the rows the dependent variable and the columns the independent, though this can be reversed if the number of columns expands beyond the size of the page.
- Under the Cells button, select the statistics for each cell. If the columns are the independent variable, it is customary to include the column percents. If the rows are the independent variable, then include the row percents.
- Under the Statistics button, select Chi-square and at least one measure of association depending the type of data. Select Kendall's tau-b (BTau in PSPP) if both variables are ordinal, and Phi and Cramer's V if at least one is nominal. (Cramer's V has to be calculated by hand if using PSPP.) Binary variables customarily are viewed as nominal unless there is a clear hierarchy (e.g., values like high/low).

Key Statistics

- The Pearson's chi-square measures the goodness of fit of the table. The null hypothesis is that the number of observations in each cell is randomly distributed in a pattern based on the marginals of the table. Use the asymptotic significance as the p-value, and if the p-value is less than .05, the observations are not randomly distributed and the null hypothesis is rejected.
- The chi-square does not tell the direction of the relationship if the variables are ordinal. Even if a direction is an acceptable result, the chi-square does not identify where the non-randomness in the table is, and the pattern might not be systematic but a single cell.
- Kendall's tau-b provides direction and effect size if the data is ordinal. Phi and Cramer's V provide only an effect size.
- The effect size is the absolute value of the specified coefficient. The interpretation of the effect size is based on the following ranges: weak relationships correlation between .01 and .29, medium or moderate relationship are between .03 and .49, and strong relationships equal or exceed .50.

Written Interpretation

Written comments highlight the significance of the chi-square statistic and the effect size and, where appropriate, the direction of statistically significant results. Statistics can be included in

the text parenthetically when accompanied by a statement that conveys the key result in plain English. If there is no significant effect, this fact should be stated in the text.