Tests of independence and association in SPSS

Bro. David E. Brown, BYU–Idaho Dept. of Mathematics

February 14, 2012

The **Crosstabs** tool in SPSS can be used for creating crosstabulations (also known as “contingency tables” or “two-way tables”) for categorical data, thereby giving you expected counts and percentages. It can also create clustered bar charts. The following instructions assume your data are categorical. (If not, then a crosstabulation is probably not what you really need, so go back and double-check.)

1. Open your data file or type in your data. If you’d like instructions on entering a two-way table, click here.

2. In the **Variable View**, give your variables names, and make sure the **Type** is correct. Add **Labels** if you want, and take particular care of the level of **Measurement**. For example, if one of your variables is nominal, please make sure the **Measurement** column says **Nominal**; if ordinal, then **Ordinal**.

3. Are your data actual, individual measurements? If so, skip to Step 4. If instead you have data values and their frequencies, you **must “weight” the cases**. This way, SPSS will know how many of each measurement you have, and can correctly calculate statistics and create graphical displays. If you have already weighted the cases, skip to Step 4. Otherwise, to weight the cases,...

   (a) Look on the button bar for a picture of an old-fashioned “pan balance” type of scale (for weighing things), and click it. A dialog appears.

   (b) In the dialog, click the radio button for **Weight cases by**....

   (c) Identify the variable that contains the weights (that is, the frequencies, relative frequencies, or percents), select it, and click the arrow button to move that variable to the **Weight cases by**... box.

   (d) Click **OK**. Nothing will change on your screen. You can check your work by clicking the “weight the cases” button again and seeing whether you’ve used the correct variable, or make sure the cases are actually weighted. You are now done weighting the cases.

4. In the **Analyze** menu, click on **Descriptive Statistics**. A submenu will appear.

5. In the submenu, click **Crosstabs**... The **Crosstabs** dialog will appear.

6. In the **Crosstabs** dialog, select the name of one of your variables and click the arrow that points to the **Column(s)** box. The name of your variable will move into the **Column(s)** box. Select the name of the other variable of interest, and click the arrow that points to the **Row(s)** box. The name of your variable will move from the one box to the other.\(^1\)

7. To select the cell statistics you want, or if you want clustered bar charts,...

   (a) Click the **Cells**... button. The **Crosstabs: Cell Display** dialog will appear.

---

\(^1\) Relative frequencies and percents are OK, too.

\(^2\) Most of the time, it won’t matter which variable is the row variable and which is the column variable. Under some circumstances (not usually found in introductory courses), it will. Be sure to follow your teacher’s instructions on this matter.
(b) Click the check boxes next to the counts and percentages you want. (Examples: If you check Observed, and Column, you’ll get the observed column counts and percentages. If you click Observed, Expected, and Column, you’ll get both the observed column counts and percentages and the expected column counts and percentages.)

(c) Click Continue. SPSS returns you to the Crosstabs dialog.

8. Click OK. The Crosstabs dialog closes, and the Output or Viewer window opens.

9. Reading the output:

(a) The first thing in the output is the Case Processing Summary table. CHECK THIS TABLE to make sure all your data were valid. If not, go repair your data and repeat the above steps.

(b) Next is a table of frequencies, except SPSS calls it the Crosstabulation table. This table has the observed or expected counts or percentages, depending on the choices you made in the Crosstabs: Cell Display dialog. (The names of your variables will be in the title of the table, separated by an asterisk [*]. This is an old-fashioned way of indicating that the two variables are being analyzed together.)

(c) Finally, if you asked for clustered bar charts, they will appear last. There will be a cluster of bars for each value in your row variable. (If you want the bars clustered differently than they are, re-do the crosstabulation, but swap the row and column variables.)

As always, if you have questions, please ask them!