

Chi-square tests in DataDesk

Test for goodness-of-fit

Note: This only works when the hypothesis to be tested is that all categories occur in equal proportion. My DataDesk documentation claims that one gets P-values from this procedure, but I only get the value of the chi-square statistic. With this, one would have to use another technology to get a P-value or look up a critical value in the table (degrees of freedom equals $n - 1$, where n is the number of categories. Please let me know if your version of DataDesk gives a P-value.

1. If you have your data in summary form already, you need to covert it to raw form. If your variables are Category and Count, select Category as Y, Count as X, and choose Manip > Replicate Y by X. DataDesk will create a single variable called Category:Count that contains raw data.
2. Click on the variable that contains raw data and select Calc > Frequency Breakdowns. A frequency table will open.
3. Click on the HyperView menu (little triangle in the upper left corner of the frequency table) and choose Frequency Options.
4. Check off the boxes corresponding to the information you'd like (standardized residual and chi-square, perhaps percent and expected as well) and click OK. The information you selected will appear on the frequency table.

Test for homogeneity

1. If you have your data in summary form already, you need to covert it to raw form. If your variables are Group, Category, and Count, select Group and Category as Y, Count as X, and choose Manip > Replicate Y by X. DataDesk will create a two variables called Group:Count and Category:Count that contain raw data.
2. Click on the two variables that contain raw data and select Calc > Contingency Tables. A contingency table will open.
3. Click on the HyperView menu (little triangle in the upper left corner of the contingency table) and choose Table Options.
4. Check off the boxes corresponding to the information you'd like (standardized residual and chi-square, perhaps percent of row or column total and expected value as well) and click OK. The information you selected, including the number of degrees of freedom and a P-value, will appear on the contingency table.

Test for independence

1. If you have your data in summary form already, you need to convert it to raw form. If your variables are First Category, Second Category, and Count, select First Category and Second Category as Y, Count as X, and choose Manip > Replicate Y by X. DataDesk will create a two variables called First Category:Count and Second Category:Count that contain raw data.
2. Click on the two variables that contain raw data and select Calc > Contingency Tables. A contingency table will open.
3. Click on the HyperView menu (little triangle in the upper left corner of the contingency table) and choose Table Options.
4. Check off the boxes corresponding to the information you'd like (standardized residual and chi-square, perhaps percent of row or column total and expected value as well) and click OK. The information you selected, including the number of degrees of freedom and a P-value, will appear on the contingency table.