**Contingency Table Analysis in SAS[[1]](#footnote-1)©**

Obtain the following files:

* Howell.dat -- from my [StatData Page](http://core.ecu.edu/psyc/wuenschk/StatData/StatData.htm)
* Chisq.sas, CMH.sas, & Kappa.sas – from my [SAS Programs Page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm)
* The Howell.dat data file is described on pages 686 and 687 of your test (8th Edition of *Statistical Methods for Psychology*).
* After being sure that the program file correctly points to the location of the data file on your computer, submit the Chisq.sas program file. Look at the program and the output as I explain both.

[PROC FORMAT](http://core.ecu.edu/psyc/wuenschk/SAS/Help/Format-SAS.htm) is used to assign alphanumeric labels to the numeric values of a variable.

In the data step named Lotus, we input four scores from each line of the data stream, with the data in the program. I obtained these data from the article "Stairs, Escalators, and Obesity," by Meyers et al. (*Behavior Modification 4*: 355‑359). The researchers observed people using stairs and escalators. For each person observed, the following data were recorded: Whether the person was obese, overweight, or neither; whether the person was going up or going down; and whether the person used the stairs or the escalator.

Each line of data in our program represents one cell in a 3 x 2 x 2 contingency table. The first score represents weight category for the cell, the second score the direction traveled, and the third score the device used to travel. The fourth score is the frequency for that cell.

PROC FREQ is invoked and asked to create TABLES. The first table requested, direct\*device, has direction as its rows and device as its columns. Options requested (after the slash) are CHISQ (Pearson chi square analysis), RELRISK (get odds ratios), NOPERCENT (suppress printing of cell total percentages), and NOCOL (suppress printing of column percentages). The WEIGHT statement identifies the freq variable as the one which contains the cell counts.

Look at your output. Also [look at my annotated output](http://core.ecu.edu/psyc/wuenschk/SAS/ChiSq-Output.docx) We see that a significantly greater number of people use the stairs when doing down (24%) than when going up (6%). The magnitude of this correlation between direction and device, as expressed by the **phi coefficient**, is .26. Another way to express this relationship is by an **odds ratio**. When going down, the odds that the stairs were used was 331/1031 = .3210. Going up, these odds were 114/1741 = .0655. The odds ratio is .3210/.0655 = **4.9**. In other words, the odds of using the stairs were nearly 5 times higher when going down than when going up. Note that SAS reported the ratio of the odds of using the stairs when going up divided by the odds of using the stairs going down, obtaining the inverse of 4.9, which is .204. If you, like most people, prefer odds ratios > 1, just invert them when SAS gives you one < 1. You can do the same with the endpoints of the confidence interval.

I like to report not only odds ratios but also confidence intervals for the odds ratios. Look at the “Estimates of the Relative Risk.” The odds ratio reported there is .2040 with a *CI.95* running from .1626 to .2558. If you invert these you get an odds ratio of 1/.2040 = 4.90 with the confidence interval running from 3.91 to 6.15. Please note that it is possible (but unlikely) to get a significant Pearson or LR Chi-Square but also get a confidence interval for the odds ratio that includes the value of one. The confidence interval for the odds ratio is constructed using the Wald Chi-Square, a statistic that is more conservative than the Pearson and LR Chi-Square statistics.

The second table requested was weight\*direct\*device, which results in a direct\*device table at each level of weight. The relationship between direct and device is statistically significant for each weight group, but the relationship differs in magnitude across the weight groups. When we study log linear analysis of multidimensional contingency tables I shall show you how to determine that this difference among weight groups is significant.

Look at the analysis for the obese people. Phi is only .17, and the odds ratio only .1728/.0488 = 3.5. Very few of the obese people took the stairs (8%). There cannot be much of a correlation between two variables when one of the variables (device in this case) has almost no variance. With the overweight people, phi is much larger, .33, and the odds ratio is .0409/.3844 = 9.4. With the normal (and skinny) people, phi is not so large, .22, and the odds ratio is .3010/.0822 = 3.7. The percentage use of the stairs (14%) was about the same in the overweight people as in the normal people. Perhaps the overweight people realized that they should be getting more exercise, such as using stairs, but using them going up was just too much to ask, thus the strong relationship between direction and device among them.

**Attention Deficit Disorder and Repeating a Grade in School**

Look back at the program, this time at the data step named Sol. The data file is that found in the back of your text book. We have one line of data for each subject. The subjects are students in Vermont. My client wishes to determine if there is a relationship between score on the ADDSC variable and whether or not a student has to repeat a grade. I recommend a *t*-test, comparing mean ADDSC in students who have repeated a grade with mean in students who have not repeated a grade, but my client wants to do a contingency table analysis, with the ADDSC variable converted into a three category variable. Converting a continuous variable into a categorical variable is almost always a very bad idea, but my client is paying me my usual fee ($200 per hour) to do what she wants, so I do it. Please read the document [Dichotomizing Continuous Variables: A Bad Idea](http://core.ecu.edu/psyc/wuenschk/StatHelp/Dichot-Not.pdf).

Proc Format is used to label the groups, and If Then statements are used to recode the ADDSC variable into the three group variable, ADD\_CAT. Note the use of the underscore character to create what looks much like an imbedded blank in the variable name. I first set ADD\_CAT to 1 for all subjects with ADDSC from 0 to 47, then to 2 for subjects having ADDSC above 47 but < 56, then to 3 for subjects with ADDSC > 56.

Proc Freq is used to produce a contingency table involving our two categorical variables. Expected frequencies and a chi‑square analysis are requested, printing of column and total percentages is suppressed, and our groups are labeled with the Formats we previously defined.

The output warns us that we have many low expected frequencies. We know that that is no matter for concern, since we have rejected the null. The row percentages in the “Repeated” column clearly indicate a positive association between ADD and repeating a grade, with 1/3 of the high‑ADD group repeating.

Here is another way to recode a continuous variable into a categorical variable:

**data** Format2;

title 'Another Way to Recode A Continuous Variable to Categorical.'; **run**;

**proc** **format**;

value add low-**47** = 'low add' **48**-**56** = 'medium add' **57**-high = 'high add';

value rep **0** = 'promoted' **1** = 'repeated';

**data** Athena; infile 'C:\Users\Vati\Documents\StatData\howell.dat';

input addsc sex repeat iq engl engg gpa socprob dropout;

format addsc add. repeat rep. ;

**proc** **freq**; tables addsc\*repeat / expected nocol nopercent chisq; **run**;

If you were to run this code, you would get exactly the same 3 x 2 analysis produced earlier.

**Follow-Up Analysis**

The 3 x 2 analysis suggests some additional hypotheses of interest. We could do “pairwise comparisons,” by comparing pairs of groups on percentage of students who repeat a grade. This would be a series of three comparisons, each done with a 2 x 2 analysis – low versus high, low versus medium, and medium versus high. The low versus medium comparison would fall short of statistical significance. Accordingly, I combined the low and medium groups and compared those combined groups with the high group.

**Cohen’s Kappa**

Run the program Kappa.sas to illustrate how to compute kappa with SAS. The data are those on page 166 of our text. For more information on kappa and weighted kappa, and related statistics, I recommend <http://www.john-uebersax.com/stat/kappa.htm> .

The McNemar-Bowker test of symmetry tests the null hypotheses that disagreement between raters does not vary across types of disagreement. As an example of a case where this would not be true, the raters might disagree more about whether the case was internalizing or externalizing than they did about whether the case was internalizing or no problem.

ASE is the standard error used to produce the confidence interval.

The weighted kappa is used when the categories are ordinal, making some disagreements more serious than others. For example, suppose the categories were tastes disgusting, tastes OK, and tastes great. Disagreement between tastes disgusting and tastes great should lower kappa more than disagreement between tastes OK and tastes great.

**Cochran-Mantel-Haenszel Statistic**

Please read this document: Bickel, P. J., Hammel, E. A., & O'Connell, J. W.  (1975).  Sex bias in graduate admissions:  Data from Berkeley.  *Science*, *187*, 398-404. Then run the program CMH.sas. The data are those from Table 6.9 in our text. Please refer to [my output](http://core.ecu.edu/psyc/wuenschk/SAS/CMH-Output.docx), which I have trimmed and annotated.

For the analysis with Department A excluded, the CMH statistic indicates that, controlling for department, there is no significant difference between admission rates for women and admission rates for men, *p* = .72. Notice that the CMH statistic reported in our text book is slightly lower than that reported by SAS. I believe this is due to Howell having used the correction for continuity, which I think is not appropriate.

If you ignore department, the admission rate for women is significantly lower than that for men, odds ratio = 0.69.

In Department A, women are much more likely to be admitted than are men, odds ratio = 2.86.

The odds ratios differ significantly across departments. The CMH statistic assumes that they do not, so the analysis including Departments A through F is invalidated.

**Barnard’s Exact Test**

When sample size is small, there are alternative analyses that may perform better, especially for the 2 x 2 contingency. The most popular is Fisher’s exact test, which although popular, is also controversial: Fisher's exact test, strictly speaking, assumes that the row and columns totals are fixed. In practice, this assumption is not met in many experimental designs and almost all non-experimental ones. An alternative exact test, Barnard's test, has been developed and is discussed in the following paper <http://www.cytel.com/Papers/twobinomials.pdf> . Proponents of Barnard's exact  test suggest that this method is more powerful, particularly in 2 by 2  tables.  However, many authorities suggest that Fisher's exact test can nonetheless be safely used.” ([source](http://spssx-discussion.1045642.n5.nabble.com/Chi-square-and-Fisher-s-Exact-Test-td1075915.html))

[SAS](http://support.sas.com/documentation/cdl/en/statug/65328/HTML/default/viewer.htm#statug_freq_syntax03.htm): Use Exact option with Proc Freq.

[Online Calculator](http://scistatcalc.blogspot.com/2013/11/barnards-test-calculator.html): There are others out there too.

[Karl L. Wuensch](http://core.ecu.edu/psyc/WuenschK/KLW.htm), October, 2014

* [Wuensch’s Stats Lessons](http://core.ecu.edu/psyc/wuenschk/StatsLessons.htm)
* [Fair Use of this Document](http://core.ecu.edu/psyc/wuenschk/Fair-Use.htm)
* [How to get people to use the stairs](http://www.youtube.com/watch?v=2lXh2n0aPyw)

Copyright 2014, Karl L. Wuensch - All rights reserved.

1. © Copyright 2014, Karl L. Wuensch - All rights reserved. [↑](#footnote-ref-1)