**Tetrachoric Correlation[[1]](#footnote-1)©**

If you have dichotomous data on two variables but are willing to assume that the underlying variables are normally distributed, you may use the tetrachoric correlation to estimate the size of the Pearson ρ between the underlying variables.

Obtain from my [StatData Page](http://core.ecu.edu/psyc/wuenschk/StatData/StatData.htm) the Potthoff.dat file. From my [SAS Programs](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Programs.htm) page, obtain Tetrachoric.sas. After being sure that the program file correctly points to the location of the data file on your computer, submit the program file. Look at the program and the output as I explain both.

Suppose you are interested in the relationship between [misanthropy](http://personal.ecu.edu/wuenschk/Misanthropy.htm) and attitudes about animal rights. You have data on 91 nonidealistic people. For each you know whether e is a misanthrope (1) or not (0) and whether e supports animal rights (1) or not. You are willing to assume that the underlying variables are both normally distributed.

I have created such a situation by doing a “median split” on continuous measures of misanthropy (misanth) and support for animal rights (ar). While this is OK for pedagogical purposes, it is otherwise very poor practice. There rarely, if ever, is a good reason for converting a continuous variable to a dichotomous variable.

Proc Freq has been employed to conduct a contingency table analysis with the tetrachoric correlation reported. It is the PLCORR option that results in the tetrachoric correlation being reported. The tetrachoric correlation is a special case of the polychoric correlation, the case where both measurement variables are dichotomous. In the more general case the measurement variables are ordinal groups.

From the contingency table we can see that 62% of the misanthropes support animal rights but only 37% of the not misanthropic participants support animal rights. If our variables represented genuine dichotomies, we could use the phi coefficient (.24) as a measure of the strength of association between misanthropy and support of animal rights. I prefer odds ratios. In this case, misanthropes are  times more likely to support animal rights than are the not misanthropic. Notice that SAS gives a confidence interval for this odds ratio, ranging from 1.15 to 6.82.

The Pearson χ2 shows us that the association between misanthropy and support for animal rights is statistically significant at *p* = .02.

The tetrachoric correlation, .38, estimates the Pearson *ρ* between the two underlying variables.

Since we have continuous measurement of the two variables, we can directly compute the sample *r*. It is .36. In this case the tetrachoric correlation and the Pearson *r* are very close to one another. Notice that both are considerably larger than is φ. Also notice that the sample *r* is significantly greater than zero, *p* < .001. Notice that the *p* is much smaller than it was with the contingency table – why? Well, when you degrade continuous variables by dichotomizing them you generally lose power.

Also see [Biserial and Polychoric Correlation Coefficients](http://core.ecu.edu/psyc/wuenschk/SAS/Biserial-Polychoric.htm)

[Return to Wuensch’s SAS Lessons Page](http://core.ecu.edu/psyc/wuenschk/SAS/SAS-Lessons.htm)

Copyright 2011, Karl L. Wuensch - All rights reserved.

1. © Copyright 2011, Karl L. Wuensch - All rights reserved. [↑](#footnote-ref-1)