Exact unconditional p-values for 2X2 tables

This document (based on a suggestion by Tomonori Ishikawa) was abstracted from comments in Professor Roger Berger's Fortran program XUN2X2 and is used to explain the options available in running the program from his website.

Program Overview

This program computes exact unconditional p-values for analyzing 2X2 tables. The program gives the user the choice of using either a binomial or multinomial model, either a one- or two-sided test, and the choice of three test statistics. It is unusual in considering multinomial models and a variety of test statistics.

Exact unconditional tests are often more powerful than exact conditional tests for these problems, while maintaining the same Type I error level α. See Agresti (2002, p. 91) and Santner & Duffy (1989, p. 217) for discussions of exact unconditional tests.

The program reports a valid unconditional p-value, that is, a value such that if the user rejects the null hypothesis if and only if the p-value is less than or equal to a specified α, the size of the test is at most α.

The Problem

The data in a 2X2 table are usually reported as

A	B	M
C	D	N
R	S	T

A, B, C and D are counts of observations classified according to two dichotomous variables, the row and column variables. In the binomial model, A and C are independent binomial random variables with success probabilities p_A and p_C and fixed sample sizes M and N. In the multinomial model, (A,B,C,D) is a multinomial vector with cell probabilities p_A, p_B, p_C and p_D and sample size T.

The hypotheses that may be tested are:

Binomial Model

One-sided		Two-sided
$H_o:\ p_A \geq p_C$	$H_o:\ p_A \leq p_C$	$H_o:\ p_A = p_C$
$H_A:\ p_A < p_C$	$H_o:\ p_A > p_C$	$H_A:\ p_A \neq p_C$

Multiinomial Model

One-sided		Two-sided
$H_o:\ p_A p_D \geq p_B p_C$	$H_o:\ p_A p_D \leq p_B p_C$	$H_o:\ p_A p_D = p_B p_C$
$H_A:\ p_A p_D < p_B p_C$	$H_o:\ p_A p_D > p_B p_C$	$H_A:\ p_A p_D \neq p_B p_C$

The multinomial hypotheses can be stated in various ways. For example, the one-sided alternative states that the odds ratio is less than one or that the conditional probability of being in the "C" cell, given that the observation is in the second row, is greater than the conditional probability of being in the "A" cell, given that the observation is in the first row. The two-sided test is a test of the hypothesis that the row and column classifications are independent.

Program Usage

The program is interactive; the user must enter data or select options in the following order:

1. DATA – input A, B, C and D
2. MODEL – choose binomial or multinomial
3. HYPOTHESIS – choose one- or two-sided
4. STATISTIC – choose one of three test statistics, Fisher's exact as used in Boschloo (1970), Z-pooled or Z-unpooled from Suissa & Shuster(1985)
5. INTERVAL – choose whether to use a confidence interval as recommended in Berger & Boos(1994)
6. CONFIDENCE COEFFICIENT – If confidence interval method is chosen, input confidence coefficient.

The program finds the maximum of $P(T \leq t_{obs})$ on the boundary of H_o and H_A. t_{obs} is the statistic calculated from input table. All test statistics are defined so that small values give evidence for H_A. This value is reported as the exact, unconditional p-value. (The error probability is added to this maximum for the confidence interval method, as described in Berger & Boos(1994).)

Values of T greater than 100 can take a few seconds. Multinomial tables require more time because a two-dimensional maximization is required. Multinomial tables with T = 100 can require about two minutes. The time increases with T.

Recommendation

Based on the power comparison in Berger (1994), we recommend the use of Fisher's exact test statistic and a .999 confidence interval for binomial tables. For multinomial tables, the use of the Z-pooled statistic, rather than Fisher's exact test statistic, will speed the computations a great deal for T ¿ 100. A .999 confidence interval is recommended for multinomial tables, too.

References

Agresti, A. (2002), Categorical Data Analysis Second Edition , New York:Wiley.

Berger, R. L. (1994), "Power comparison of exact unconditional tests for comparing two binomial proportions," Technical report 2266, Department of Statistics, North Carolina State University.

Berger, R. L. (1996), "More powerful tests from confidence interval p values," The American Statistician, 50, 314-318.

Berger R. L., and Boos, D. D. (1994), "P values maximized over a confidence set for the nuisance parameter," Journal of the American Statistical Association, 89, 1012-1016.

Boschloo, R. D. (1970), "Raised conditional level of significance for the 2X2 table when testing the equality of two probabilities," Statistica Neerlandica, 24, 1-35.

Santner, T. J. and Duffy, D. E. (1989), The Statistical Analysis of Discrete Data, New York:Springer.

Suissa, S., and Shuster, J. J. (1985), “Exact unconditional sample sizes for the 2X2 binomial trial,” Journal of the Royal Statistical Society, Series A, 148, 317-327.