Program Z_INDEP.EXE

1. Introduction and notation.

The program performs the one (two)-tailed asymptotic test for testing the null hypothesis of independence (*H*) in a 2×2 table like Table I; in addition, the program determines the validity conditions for these tests. All the details of the following may be seen in [1].

Table I
Presentation of results in the form of a 2×2 table

		Characteristic A YES	Characteristic A NO	Totals
Characteristic B	YES	x_1	y_1	n_1
	NO	x_2	y_2	n_2
Totals		a_1	a_2	n

The results in the above table can have been obtained using 3 types of sampling, which results in 3 types of statistical models and 3 types of test:

i) *Model III* (the row marginal totals and the column marginal totals are fixed): The values of a_i and n_i (i=1, 2) are previously fixed, the sole random variable (r.v.) in the problem is x_1 (for example) –a generalized hypergeometric r.v.– and its distribution under *H* is the hypergeometric one. The exact test for *H* is performed using a conditional test (the known Fisher's exact test) [2], which gives rise to a p-value of $P_{III(E)}$.

ii) *Model II* (only the row marginal totals are fixed): The values of n_i (i=1, 2) are previously fixed, the r.v. for the problem are x_1 and x_2 (for example) –two independent binomial r.v. with parameters p_1 and p_2– and its joint distribution under *H* (p_1=p_2=p) is a double-binomial based on the nuisance parameter *p*. The exact test for *H* is carried out using an unconditional test [3], which yields a p-value of $P_{II(E)}$.

iii) *Model I* (no fixed marginal): Only the value of *n* is previously fixed, the r.v. for the problem are x_1, y_1 and x_2 (for example) –a multinomial distribution with parameters p_{11}, p_{12}, p_{21} and p_{22}– and its distribution under *H* ($p_{ij}=p_{i\bullet}\times p_{\bullet j}$) is a multinomial based on the two nuisance parameters $p_{\bullet 1}$ and $p_{1\bullet}$. The exact test for *H* is carried out using an unconditional test [3], which yields a p-value of $P_{I(E)}$.

The exact p-values $P_{III(E)}$, $P_{II(E)}$ and $P_{I(E)}$ are not unique (because they depend on which order statistic is used to obtain them), generally $P_{III(E)}\geq P_{II(E)}\geq P_{I(E)}$ (although not always) and they are progressively more difficult to calculate (in that order). In particular, it is not possible at present to determine the last two (especially $P_{I(E)}$) when the marginal are moderately large. The value of $P_{III(E)}$ is given in many statistical packages. The value of $P_{II(E)}$ is given by the StatXact package. Free programs for obtaining the values $P_{I(E)}$ and $P_{II(E)}$ may be copied at http://www.ugr.es/local/bioest/SMP.EXE and www.ugr.es/local/bioest/TMP.EXE respectively. If the p-value under a model for the generally most powerful statistic cannot be determined because of computational problems, the p-value under the same model, but for a different less powerful statistic may be determined. If it is still impossible, then the p-value under the higher model may be obtained. The details may be consulted in [4].

For teaching purposes, more ease or the impossibility of calculation, it is usual to resolve the above problem in an approximate manner using the chi-square test with the most appropriate continuity correction c (c.c.) [5]. But, given that the chi-square test is an asymptotic test, it will be subject to certain validity conditions (v.c.). It is usual to require that the minimum expected quantity

$$E = \frac{\min\ (n_1, n_2) \times \min\ (a_1, a_2)}{n}$$

be sufficiently large ($E > E$*****), where E***** is a number which, as will be seen, does not have to be fixed nor have the classic value of 5.

For the following, one must bear in mind that:

1) It is understood that if P_E is the p-value obtained by the exact test and P_A is the p-value obtained by the chi-square asymptotic test, then the condition $E>E^*$ must guarantee that:

$$|P_A - P_E| \leq \delta P_E \quad \text{where} \quad \delta = \begin{cases} 1 & \text{if} \quad P_E \leq 1\% \\ 1.15 - 15 \times P_E & \text{if} \quad 1\% < P_E \leq 5\% \\ 0.5 - 2 \times P_E & \text{if} \quad 5\% < P_E \leq 10\% \\ 0.3 & \text{if} \quad P_E > 10\% \end{cases}$$

In particular, this means that any value $3\% \leq P_A \leq 7\%$ is an acceptable estimation of P_E=5%. The values of E^* and P_A are obtained as indicated below. The values of P_E are obtained for the models *I*, *II* and *III* using the unconditional, unconditional and conditional exact tests, respectively, based on the orders of Yates's chi-square, Barnard [3] and Yates's chi-square respectively. The reasons for this may be seen in [1].

2) In the one-tailed test there are two possible alternatives *K*:

KR (one right tail): "The association is positive";

KL (one left tail): "The association is negative".

Although the program determines the p-value in both cases, the description which follows only considers those results and hypotheses which yield a p-value <0.5 (in order to facilitate the explanation). This means that the one-tailed tests which follow refer to the alternative that is compatible with the results (*KR* if N>0 or *KL* if N<0, where $N = x_1y_2 - x_2y_1$) and are valid only when the results verify $-c < N < +c$: in the other cases the p-value is ≥0.5.

3) In the <u>two-tailed test</u> (alternative K: "There is positive or negative association) the chi-square quantity depends on the term $|N|-c$. In every case it is understood that $|N|-c=0$ when $-c \le N \le +c$.

4) If $F(\bullet)$ refers to the distribution function of a typical normal r.v., then the p-value of the statistic $\chi \ge 0$ for the one- or two-tailed test is $1-F(\chi)$ or $2\times\{1-F(\chi)\}$, respectively.

2. Calculating the asymptotic p-value under the Model *III* (both marginal totals are fixed).

2.1. One-tailed test.

The <u>classic statistic</u> is the chi-square statistic with <u>Yates</u>'s c.c. Thus the p-value is given by $P_Y = 1-F(\chi_Y)$, where:

$$\chi_Y^2 = \frac{\left\{|x_1y_2 - x_2y_1| - n/2\right\}^2}{a_1a_2n_1n_2}n \qquad (1)$$

The <u>optimal statistic</u> [6] is the chi-square statistic with <u>Conover</u>'s c.c. [7]. Thus the p-value is given by $P_{III} = 1-\mathrm{F}(\chi_{III})$, where:

$$\chi_{III}^2 = \chi_Y^2 + \frac{n^3}{4a_1a_2n_1n_2} \qquad (2)$$

For the above tests to be valid, it is necessary that:

$$E > E^* = \frac{1}{K+1}\left[\frac{n}{2}\left\{1-\sqrt{\frac{C}{C+4y}}\right\}\right]^- \text{ where } \begin{cases} C = \dfrac{(n-2)^2}{n-1} \\ y = \begin{cases} 7.785 + 20.627x \text{ for Yates's case } (\chi_Y^2) \\ 6.486 + 19.111x \text{ for Conover's case } (\chi_{III}^2) \end{cases} \\ x = \dfrac{(K-1)^2}{K} \\ K = \begin{cases} \dfrac{\max(n_1,n_2)}{\min(n_1,n_2)} & \text{if } |n_2-n_1| \le |a_2-a_1| \\ \dfrac{\max(a_1,a_2)}{\min(a_1,a_2)} & \text{if } |n_2-n_1| > |a_2-a_1| \end{cases} \\ [z]^- = \text{ the integral part of } z \end{cases} \qquad (3)$$

<u>A universal</u> (and conservative) <u>condition</u> is $E^*=19.2$ (20.7) for the case P_{III} (P_Y).

<u>The universal condition when $n \le 500$</u> is reduced to $E^*=8.1$ for the case P_{III}.

2.2. Two-tailed test.

The classic statistic is the chi-square statistic with the Yates c.c. So the p-value is given by $P_Y = 2\times\{1-F(\chi_Y)\}$. But this gives rise to a very conservative test [8].

The optimal statistic [9] is that of Yates but using the Mantel precaution [8].The p-value is given by $P_{III} = \{1-F(\chi_Y)\}+\{1-F(\chi_{Y'})\}$, where $\chi^2_{Y'}$ refers to the value of χ^2_Y in the table with $x'_1=[2E_{11}-x_1]$, where $E_{11}=n_1a_1/n$ and $[x]$ refers to the rounding of x in the sense of moving away from the value of E_{11} ($[x]=E_{11}$ if $x=E_{11}$; $[x]$ = "the integral part of x" if $x<E_{11}$; $[x]$ = "the smallest integer not less than x" if $x>E_{11}$). When $x'_1 > s=\min(n_1, a_1)$ or $x'_1 < r=\max(0, a_1-n_2)$, then $1-F(\chi_{Y'})=0$. This test is always valid (since $E^*=0$).

3. Calculating the asymptotic p-value under the Model *II* (only the row marginal totals are fixed).

The classic statistic is the chi-square statistic with the Yates c.c. So the p-value is given by $P_Y = 1-F(\chi_Y)$ for the one-tailed test and $P_Y = 2\times\{1-F(\chi_Y)\}$ for the two-tailed test. But this produces a very conservative test [10].

The optimal statistic is Pearson's chi-square statistic with the c.c. of Martín *et al.* [10]:

$$\chi^2_{II}=\frac{\{|x_1y_2-x_2y_1|-c\}^2}{a_1a_2n_1n_2}(n-1) \text{ where } c=\begin{cases}2 & \text{if } n_1=n_2\\ 1 & \text{if } n_1\neq n_2\end{cases} \tag{4}$$

The p-values for the one- and two-tailed tests then are $P_{II} = 1-F(\chi_{II})$ and $P_{II} = 2\times\{1-F(\chi_{II})\}$ respectively.

For the test to be valid, it is necessary that:

$$E>E^*=\left[\frac{1}{K+1}\left[\frac{n}{2}\left\{1-\sqrt{\frac{C}{C+4y}}\right\}\right]\right]^- \text{ where } \begin{cases} C=\dfrac{n^2}{n-1}\\ y=\begin{cases}9.743 + 14.786x & \text{for the one-tailed test}\\ 16.315 + 14.889x & \text{for the two-tailed test}\end{cases}\\ x=\dfrac{(K-1)^2}{K}\\ K=\dfrac{\max(n_1,n_2)}{\min(n_1,n_2)}\\ [z]^- = \text{the integral part of } z\end{cases} \tag{5}$$

A universal (and conservative) condition is $E^*=14.8$ (14.9) for the case of one (two) tails.

The universal condition when $n\le 500$ is reduced to $E^*=7.2$ (7.7) for the case of one (two) tails.

4. Calculating the asymptotic p-value under the Model *I* (no fixed marginal).

The classic statistic is the chi-square statistic with the Yates c.c. Hence the p-value is given by $P_Y = 1 - F(\chi_Y)$ for the one-tailed test and $P_Y = 2\times\{1 - F(\chi_Y)\}$ for the two-tailed test. But this gives rise to a very conservative test [10].

The optimal statistic [11] is the Pearson's chi-square statistic with the c.c. of Pirie and Hamdan [12]:

$$\chi_I^2 = \frac{\left\{\left|x_1 y_2 - x_2 y_1\right| - 0.5\right\}^2}{a_1 a_2 n_I n_2}(n-1) \tag{6}$$

The p-values for the one- and two-tailed tests then are $P_{II} = 1 - F(\chi_I)$ y $P_{II} = 2\times\{1 - F(\chi_I)\}$ respectively.

For the test to be valid it is necessary that:

$$E > E^* = \frac{1}{K+1}\left[\frac{n}{2}\left\{1 - \sqrt{\frac{C}{C+4y}}\right\}\right]^- \text{ where } \begin{cases} C = \dfrac{n^4}{(n-2)^2(n-1)} \\ y = \begin{cases} 7.318 + 3.899x & \text{for the one-tailed test} \\ 7.612 + 6.120x & \text{for the two-tailed test} \end{cases} \\ x = \dfrac{(K-1)^2}{K} \\ K = \begin{cases} \dfrac{\max(n_1, n_2)}{\min(n_1, n_2)} & \text{if } |n_2 - n_I| \le |a_2 - a_I| \\ \dfrac{\max(a_1, a_2)}{\min(a_1, a_2)} & \text{if } |n_2 - n_I| > |a_2 - a_I| \end{cases} \\ [z]^- = \text{the integral part of } z \end{cases} \tag{7}$$

A universal (and conservative) condition is $E^*=3.9$ (6.2) for the case of one (two) tails.
The universal condition when $n \le 500$ is reduced to $E^*=3.0$ (3.9) for the case of one (two) tails.

REFERENCES

1. Martín Andrés, A.; Sánchez Quevedo, M. J.; Tapia García, J.M. and Silva Mato, A. (2005). On the validity condition of the chi-squared test in 2×2 tables. *Test* 14 (1), 1-30.
2. Fisher, R.A. (1935). The logic of inductive inference. *Journal of the Royal Statistical Society A 98,* 39-54.
3. Barnard, G.A. (1947). Significance tests for 2×2 tables. *Biometrika 34*, 123-138.
4. Martín Andrés, A.; Silva Mato, A.; Tapia García, J.M. and Sánchez Quevedo, M. J. (2004). Comparing the asymptotic power of exact tests in 2×2 tables. *Computational Statistics and Data Analysis* 47 (4), 745-756.

5. Greenwood, P.E. and Nikulin, M.S. (1996). *A Guide to Chi-Squared Testing*, 1d ed., ed.Wiley Series in Probability and Statistics. New York, 1996.
6. Martín Andrés, A. and Herranz Tejedor, I. (1997). On condition for validity of the approximations to Fisher's exact test. *Biometrical Journal 39*, 935-54.
7. Conover, W.J. (1974). Some reasons for not using the Yates' continuity corrections on 2×2 contingency tables. *Journal of the American Stat. Assoc. 69*, 374-376.
8. Mantel, N. (1974). Comment and a Suggestion. *Journal of the American Stat. Assoc. 69*, 378-380.
9. Martín Andrés, A., Herranz Tejedor, I. and Luna del Castillo, J.D. (1992). Optimal correction for continuity in the chi-squared test in 2×2 tables (conditioned method). *Commun. Statist. - Simula. and Comp. 21 (4)*, 1077-1101.
10. Martín Andrés, A.; Sánchez Quevedo, M. J. and Silva Mato, A. (2002). Asymptotical tests in 2×2 comparative trials (unconditional approach). *Computational Statistics and Data Analysis* 40 (2), 339-354.
11. Martín Andrés, A. and Tapia Garcia, J. M. (2004). Optimal unconditional asymptotic test in 2×2 multinomial trials. *Commun. Statist. -Simula. and Comp. 33 (1)*, 83-97.
12. Pirie, W.R. and Hamdan, M.A. (1972). Some revised continuity corrections for discrete distributions. *Biometrics 28*, 693-701.