A program for exact power of the 2×2 chi-square

WILLIAM P. DUNLAP

Tulane University, New Orleans, Louisiana

The most widely available power tables for 2×2 contingency tables are those by Cohen (1988), which are based on a normal approximation following the use of the arc sine transformation. Dunlap (1982) showed that Cohen's (1988) algorithm and two somewhat more accurate algorithms by Zelen and Severo (1965) and Fleiss (1973) are particularly inaccurate when the proportions being compared are quite small (or quite large). This was shown when the power estimated by the three algorithms was compared with the actual power computed directly by the present program.

The program, described below, is limited to the 2×2 case, which may appear to be a substantial shortcoming in terms of experimental planning, where larger contingency tables may often be anticipated. However, if one concentrates on the two groups expected to differ the most, and if one is willing to collapse multicategory response variables into only two categories, the program can still be used to give accurate power estimates for quite complex experiments. A second limitation to the present program is that the exact power calculations only pertain to what Pearson (1947) called "Case 2 data," which consists of an experimental design in which two groups (presumably an experimental and control group) are compared in terms of the observed proportions of a binomial event in the two groups. This case is different from "Case 1 data," where two binomial variables are measured and the hypothesis of independence is tested. Although chi-square treats these two cases identically, the underlying distributions differ, such that the multinomial distribution is appropriate for the latter case, as opposed to the binomial for the former. Therefore, with Case 1 (correlational) 2×2 data, the power computations of this program would be only approximate, and not exact. The mathematical steps in computing power by the present program are described below.

Mathematical development. With a 2×2 contingency table, the power of the chi-square test can be computed directly, if somewhat laboriously, by examining all pos-

sible 2×2 tables, given the sample sizes in the two groups. If there are n_1 cases in the first group, there are $(n_1 + 1)$ possible outcomes for that group (0, 1, 2, ..., n_1). Similarly, there are $(n_1 + 1)(n_2 + 1)$ possible 2×2 tables. The program generates each possible table and computes chi-square for each. If the chi-square exceeds its 0.05 critical value (3.84146), the program computes the probability of that particular table occurring from the product of two binomial probabilities, as shown in Equation 1 below. Pearson (1947) gives the actual probability of each table with Case 2 data as

$$\frac{n_1!}{s_1! (n_1 - s_1)!} P_1^{s_1} (1 - P_1)^{(n_1 - s_1)} \times \frac{n_2!}{s_2! (n_2 - s_2)!} P_2^{s_2} (1 - P_2)^{(n_2 - s_2)}, \quad (1)$$

where n_1 and n_2 are the sample sizes, s_1 and s_2 the numbers of events, and P_1 and P_2 the population proportions of events for Groups 1 and 2, respectively. The actual probabilities of each table that results in a significant chi-square are summed, and the result is the power, if P_1 differs from P_2 , or the Type I error rate, if $P_1 = P_2$.

The program (see Appendix) is written in single-precision FORTRAN IV and runs on an IBM 3081 KX. For the IBM, Unit 9 defaults to the CRT screen and keyboard. By replacing Unit 9 in the READ and WRITE statements with Unit *, the program will compile and run on an IBM PC or clone. The user is warned, however, that because the program generates and examines all possible contingency tables, problems with large sample sizes require very long run times on microcomputers.

Program availability. A listing of the program may be obtained free of charge from William P. Dunlap.

REFERENCES

- COHEN, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- DUNLAP, W. P. (1982). An interactive FORTRAN IV program for calculating aspects of power with dichotomous data. *Behavior Research Methods & Instrumentation*, **14**, 422-424.
- FLEISS, J. L. (1973). *Statistical methods for rates and proportions*. New York: Wiley.
- PEARSON, E. S. (1947). The choice of statistical tests illustrated on the interpretation of data classed in a 2×2 table. *Biometrika*, **34**, 139-167.
- ZELEN, M., & SEVERO, N. C. (1965). Probability functions. In M. Abramowitz & I. A. Stegun (Eds.), *Handbook of mathematical functions* (p. 942). New York: Dover.

Correspondence may be addressed to William P. Dunlap, Department of Psychology, Tulane University, New Orleans, LA 70118.

APPENDIX

```
C -- EXACT POWER COMPUTATION FOR THE 2x2 CHI SQUARE, CASE 2
C -- CVCS IS THE CRITICAL VALUE TO BE USED -- SET BELOW FOR
C -- ALPHA = 0.05 AND A TWO-TAILED TEST
      CVCS=3.84146
```

APPENDIX (Continued)

```

1  WRITE(9,2)
2  FORMAT(1X,'ENTER P1,P2,N1,N2 (0,0,0,0 TO STOP)')
   READ(9,*)P1,P2,N1,N2
   IF(N1+N2.EQ.0)STOP
   XN1=N1
   XN2=N2
   N1PO=N1+1
   N2PO=N2+1
   GN=N1+N2
   PC=0.0
   DO 4 I=1,N1PO
     A=I-1
     PG1=BP(XN1,A,P1)
     IF(PG1.EQ.0.0)GOTO 4
     C=N1PO-I
     DO 3 J=1,N2PO
       B=J-1
       D=N2PO-J
       CS=0.0
       BOT=(A+B)*(C+D)*(A+C)*(B+D)
       IF(BOT.GT.0.0)CS=GN*(A*D-B*C)**2/BOT
       IF(CS.LT.CVCS)GOTO 3
       PC=PC+PG1*BP(XN2,B,P2)
3    CONTINUE
4    CONTINUE
   WRITE(9,5)PC
5   FORMAT(' BINOMIAL PROB =',F12.6)
   GOTO 1
   END

C
   FUNCTION BP(X,A,P)
C  COMPUTES THE BINOMIAL PROBABILITY OF A OUT OF X EVENTS
C  WITH PROBABILITY P --  $X!/(A!X-A!)P^A Q^{(X-A)}$ 
   B=X-A
   Q=1.0-P
   TEST=P*Q
   IF(TEST.GE.0.0)GOTO 2
   WRITE(9,1)P
1  FORMAT(' P =',F10.6,' IS OUT OF BOUNDS')
   STOP
2  IF(TEST.GT.0.0)GOTO 3
   BP=0.0
   IF(A.EQ.0.0.AND.P.EQ.0.0)BP=1.0
   IF(A.EQ.X.AND.P.EQ.1.0)BP=1.0
   RETURN
3  K=A
   IF(B.LT.A)K=B
   CON=0.0
   DO 4 I=1,K
     Y=I
     CON=CON-ALOG(Y)+ALOG(X-Y+1.0)
4  CONTINUE
   BPL=CON+ALOG(P)*A+ALOG(Q)*B
   BP=0.0
   IF(BPL.GT.-25.0)BP=EXP(BPL)
   RETURN
   END

```