Behavior Research Methods & Instrumentation
1981, Vol. 13 (3), 363-366

FORTRAN IV functions for calculating probabilities associated with Dunnett's test

WILLIAM P. DUNLAP, MARCIA S. MARX,
and GREGORY J. AGAMY
Tulane University, New Orleans, Louisiana 70118

Dunnett's (1955, 1964) test was designed specifically to permit the researcher to compare a control group mean with several experimental group means while holding the probability of one or more Type I errors for the entire set of comparisons, the so-called experimentwise error rate, at a specified value. Dunnett's procedure, in fact, provides a test that corresponds well to what a researcher very often wants to test, concentrating power for only comparisons between treatment groups and the control or standard, rather than diluting power to test all possible pairwise differences between group means, as do tests based on the studentized range distribution, such as Tukey's test (see Dunlap, Powell, & Konnerth, 1977, for computing probabilities of these tests). Perhaps the reason that Dunnett's test is not more widely used involves the availability of adequate tables. The potential user is most often referred to tables by Dunnett (1964), but these tables give values that are appropriate only when the correlation between comparisons equals .5, the case when all group sample sizes are equal. Dunnett pointed out that his test is actually more powerful when the control group is allocated a larger sample size than experimental groups; however, in this case, the correlations between contrasts decrease and the tabled values must be modified with provided correction factors, or the user may refer to somewhat more extensive tables by Hahn and Hendrickson (1971), in which the correlation between contrasts is a parameter. Given the potential usefulness of Dunnett's test, Miller (1966, p. 75) has made clear the need for more extensive tables.

The present program calculates the probabilities associated with Dunnett's test directly, thus providing an alternative to more extensive tabulation of critical values. Although comparing control and experimental group means is the most common use of this distribution, Hahn and Hendrickson (1971) outline a number of other applications (e.g., confidence intervals for regression equations, interactions, and future observations or means). Shaffer (1977) describes procedures for extending Dunnett's test to provide conservative tests for other pairwise and more complex contrasts between group means.

The program described in the present paper computes the probabilities of Dunnett's distribution in the following form:

$$1 - F(t/df;k;r) =$$

$$\int_0^{\infty} \left\{ \int_{-\infty}^{\infty} \left[E\left(\frac{r^{1/2}x+ts}{(1-r)^{1/2}}\right) - E\left(\frac{r^{1/2}x-ts}{(1-r)^{1/2}}\right) \right]^k z(x)dx \right\} g(s;df)ds,$$

where

$$z(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

$$E(u) = \int_{-\infty}^{\infty} z(u)du,$$

and

$$g(s;df) = \frac{df^{df/2}}{\Gamma(df/2)2^{df/2-1}} s^{df-1} e^{-df\, s^2/2}.$$

The function $z(x)$ is the equation for ordinates of the nominal curve. The integral, $E(t)$, is the cumulative probability of a standard normal deviate that is solved by the probability function provided in Dunlap and Duffy (1975). Since extreme normal deviates are encountered, this function is augmented by another approximation for extreme scores provided by Zelen and Severo (1965, p. 933, Function 26.2.24).

The inner integral of the top expression is solved numerically with about 20 points using Simpson's rule. Since the distribution integrated is symmetric, the program calculates the area from 0 to infinity and then doubles this amount. The solution to this inner integral gives probability values of Dunnett's test when the degrees of freedom, df, are infinite. In practice, the program solves this inner integral only whenever the degrees of freedom exceed 2,000.

The outer integral studentizes Dunnett's test; that is, it adjusts the distribution for finite degrees of freedom associated with the error term. The function $g(s;df)$ is the distribution of the ratio of the sample to the population standard deviation, and it is not symmetric; therefore, the integration is done in two parts, from the modal value 1 to infinity and then from 1 to 0. This integration involves from 24 to 30 points fitted with Simpson's rule.

The first input parameter for the function is the Student t statistic, calculated in the usual manner:

$$t = (\bar{X}_c - \bar{X}_E)/(S\sqrt{1/N_c + 1/N_E}),$$

where $\bar{X}_c$ and $\bar{X}_e$ are the control and experimental means, N_c and N_E are the respective sample sizes, and S

0005-7878/81/030363-04$00.65/0

is the estimated common standard deviation. Usually, the square root of the mean square error from an analysis of variance involving all groups is used to estimate S. The second parameter, df, is the degrees of freedom associated with S (equal to the df error from the analysis of variance). The third input parameter, K, is the number of experimental group means, excluding the control mean. The last parameter, r, is the correlation between contrasts, calculated as

$$r = N_E/(N_E + N_c),$$

which clearly equals .5 for equal sample sizes, but which will be less than .5 for proportionally larger N_c. For a more complete description and a worked example of Dunnett's test, see Edwards (1968).

Accuracy. The probabilities produced by the function agree to as many places of accuracy as are available in any of the published tables of Dunnett's test. Thus the calculations are accurate to at least the fourth decimal place. The inaccuracy introduced by switching to the asymptotic distribution for degrees of freedom greater than 2,000 is in the fourth decimal place and so should not present much problem to the practical user.

Time Requirements. Because two nested integrals are solved numerically, with approximately 600 (20 X 30) points, the time required to find each probability value is rather long; however, run time does not vary greatly as a function of the parameters used. On a DEC-20 computer, the average time for each probability calculation is in the neighborhood of .70-.85 sec.

Program Availability. The working part of the program is written as a series of four function subprograms in standard single-precision FORTRAN IV and runs on a DEC-20 computer. Since the language is standard, these functions should run without change on any machine that has a FORTRAN IV compiler. A nine-statement main routine to operate the functions is also included at the front and contains some FORTRAN statements peculiar to the DEC system; however, the users of these functions will probably wish to write their own main routines, depending upon the particular use they have in mind.

There are no special peripheral requirements; memory size need only be sufficient to load the function and a main routine to call it. A listing of the function may be obtained at no cost from William P. Dunlap, Department of Psychology, Tulane University, New Orleans, Louisiana 70118.

REFERENCES

Dunlap, W. P., & Duffy, J. A. FORTRAN IV functions for calculating exact probabilities associated with z, chi-square, t, and F values. *Behavior Research Methods & Instrumentation,* 1975, **7**, 59-60.

Dunlap, W. P., Powell, R. S., & Konnerth, T. K. A FORTRAN IV function for calculating probabilities associated with the studentized range statistic. *Behavior Research Methods & Instrumentation,* 1977, **9**, 373-375.

Dunnett, C. W. A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association,* 1955, **50**, 1096-1121.

Dunnett, C. W. New tables for multiple comparisons with a control. *Biometrics,* 1964, **20**, 482-491.

Edwards, A. L. *Experimental design in psychological research* (3rd ed.). New York: Holt, Rinehart, & Winston, 1968.

Hahn, G. J., & Hendrickson, R. W. A table of percentage points of the distribution of the largest absolute value of k Student t variates and its applications. *Biometrika,* 1971, **58**, 323-332.

Miller, R. G. *Simultaneous statistical inference.* New York: McGraw-Hill, 1966.

Shaffer, J. P. Multiple comparisons emphasizing selected contrasts: An extension and generalization of Dunnett's procedure. *Biometrics,* 1977, **33**, 293-303.

Zelen, M., & Severo, N. C. Probability functions. In M. Abramowitz & I. A. Stegun (Eds.), *Handbook of mathematical functions.* New York: Dover, 1965.

Appendix

```
C
C  MAIN ROUTINE FOR EQUAL N DUNNETT PROBABILITIES
C
1       TYPE 2
2       FORMAT (' ENTER Q,DF,K ',$)
        ACCEPT *,Q,DF,K
        R=.5
        P=PDUN(Q,DF,K,R)
        TYPE 3,P
3       FORMAT(' PROBABILITY = ',F12.7)
        GOTO 1
        END
C
        FUNCTION PDUN(T,D,K,R)
C
C  COMPUTES THE PROBABILITY FOR DUNNETTS TEST
C     T=STUDENTS T BETWEEN A CONTROL & EXPERIMENTAL GROUP
C     D=DEGREES OF FREEDOM ASSOCIATED WITH T
C     K=NUMBER OF EXPERIMENTAL GROUPS (NOT COUNTING CONTROL)
C     R=CORRELATION BETWEEN T'S -- 0.5 FOR EQUAL N
C        R=N(EXP)/(N(EXP)+N(CONT)) FOR UNEQUAL N
C  PDUN RETURNS WITH TWO TAILED PROBABILITY OF T
```

```
C  FOR DF > 2000 DF IS SET EQUAL TO INFINITY
C
         COMMON GAML,DF,BT,TP
         Q=ABS(T)
         DF=D
         BT=SQRT(1.-R)
         TP=SQRT(R)
         IF(DF.GT.2000.)GOTO 7
C
C  FIND THE LOG OF GAMMA(DF/2.)
C
         X=DF/2.
         GAML=ALOG(SQRT(3.14159))
         N=X+.5
         IF(X-FLOAT(N).EQ.0.)GAML=0.
         IF(X.LE.1.)GO TO 2
         DO 1 I=2,N
         Y=I
1        GAML=GAML+ALOG(X-Y+1.)
C
C  OUTER INTEGRAL FROM 1 TO INFINITY
C
2        A1=0.0
         S=.14/SQRT(DF)
         X0=1.
         F0=DUN(Q*X0,R,K)*SD(X0)
3        X1=X0+S
         F1=DUN(Q*X1,R,K)*SD(X1)
         X2=X1+S
         F2=DUN(Q*X2,R,K)*SD(X2)
         SUB=S/3.*(F0+4.*F1+F2)
         A1=A1+SUB
         X0=X2
         F0=F2
         S=S*1.05
         IF(A1/SUB.LT.10E7)GOTO 3
C
C  OUTER INTEGRAL FROM 1 TO 0
C
         A2=0.0
         S=-.14/SQRT(DF)
         XINC=1.05
         IF(DF.GT.12.)GOTO 4
         S=-.03125
         XINC=1.
4        X0=1.
         F0=DUN(Q*X0,R,K)*SD(X0)
         DO 5 KK=1,16
         X1=X0+S
         F1=DUN(Q*X1,R,K)*SD(X1)
         X2=X1+S
         F2=DUN(Q*X2,R,K)*SD(X2)
         SUB=-S/3.*(F0+4.*F1+F2)
         A2=A2+SUB
         IF(A2/SUB.GT.10E7)GOTO 6
         X0=X2
         F0=F2
5        S=S*XINC
6        PDUN=1.0-A1-A2
         IF(PDUN.LT.0.)PDUN=0.
         RETURN
7        PDUN=1.-DUN(Q,R,K)
         END
C
         FUNCTION DUN(Q,R,K)
C
C  SOLVES INNER INTEGRAL USING SIMPSONS RULE
C  WIDTH OF FIT PARABOLAS IS INCREASED EACH STEP
C  INTEGRATION STOPS WHEN SUBAREA/AREA IS LESS THAN 1E-6
C  DISTRIBUTION IS SYMMETRIC SO THE INTEGRAL FROM 0 TO INFINITY
C  IS DOUBLED FOR SOLUTION
C
```

```
      COMMON GAML,DF,BT,TP
      DATA SP/.39894228/
      AREA=0.0
      IF(Q.LE.0.0)RETURN
      S=.07
      X0=0.
      F0=(ZPRB((TP*X0+Q)/BT)-ZPRB((TP*X0-Q)/BT))**K*SP*EXP(-X0*X0/2.)
1     X1=X0+S
      F1=(ZPRB((TP*X1+Q)/BT)-ZPRB((TP*X1-Q)/BT))**K*SP*EXP(-X1*X1/2.)
      X2=X1+S
      F2=(ZPRB((TP*X2+Q)/BT)-ZPRB((TP*X2-Q)/BT))**K*SP*EXP(-X2*X2/2.)
      SUB=S/3.*(F0+4.*F1+F2)
      AREA=AREA+SUB
      X0=X2
      F0=F2
      S=S*1.05
      IF(AREA/SUB.LT.10E7)GOTO 1
      DUN=2.*AREA
      RETURN
      END
C
      FUNCTION SD(S)
C
C ORDINATE OF THE FREQUENCY DISTRIBUTION OF S/SIGMA
C DF IS THE DEGREES OF FREEDOM AND GAML IS THE LOG OF GAMMA(DF/2.)
C
      COMMON GAML,DF,BT,TP
      SD=0.0
      IF(S.LE.0.0)RETURN
      D=DF/2.
      SD=ALOG(DF)*D+ALOG(S)*(DF-1.)-D*S*S-GAML-.6931472*(D-1)
      SD=EXP(SD)
      RETURN
      END
C
      FUNCTION ZPRB(Z)
C
C COMPUTES THE CUMULATIVE PROBABILITY OF NORMAL DEVIATE Z
C (INTEGRAL OF THE NORMAL DISTRIBUTION FROM -INFINITY TO Z)
C
      X=ABS(Z)
      ZPRB=0.
      IF(X.GT.12.)GOTO 2
      Q=.39894228*EXP(-X*X/2.)
      IF(X.GT.3.7) GO TO 1
      T=1./(1.+.2316419*X)
      P=.31938153*T
      P=P-.356563782*T**2
      P=P+1.78147937*T**3
      P=P-1.821255978*T**4
      P=P+1.330274429*T**5
      ZPRB=Q*P
      GO TO 2
1     ZPRB=Q*(SQRT(4.+X*X)-X)/2.
2     IF (Z.GT.0.) ZPRB=1.-ZPRB
      RETURN
      END
```

(Accepted for publication March 26, 1981.)