BIOMATHEMATICS TRAINING PROGRAM

AN ALGORITHM FOR THE APPROXIMATE NULL DISTRIBUTION OF HOTELLING'S GENERALIZED T_o^2

by

Ronald W. Helms

and

Elizabeth DeLong

Department of Biostatistics
University of North Carolina
Chapel Hill, NC 27514

Institute of Statistics Mimeo Series No. 970
July, 1975

Keywords: T_0^2, Trace Criterion, Multivariate Analysis

LANGUAGE

ANSI Fortran

Summary

An algorithm is presented for computing an approximation to the cumulative distribution function of Hotelling's generalized T_0^2 statistic.

DESCRIPTION AND PURPOSE

Hotelling's Generalized T_0^2 statistic is defined as

$$T_0^2 = n_2 \text{ Trace } (\underset{\sim}{H}\, \underset{\sim}{E}^{-1})$$

where $\underset{\sim}{H}$ and $\underset{\sim}{E}$ are independent p x p matrices;

$$\underset{\sim}{H} \sim W_p\, (n_1, \underset{\sim}{\Sigma}\, , \underset{\sim}{\Delta}), \text{ and is positive semidefinite;}$$

$$\underset{\sim}{E} \sim W_p\, (n_2, \underset{\sim}{\Sigma}\, , \underset{\sim}{0}), \text{ and is positive definite;}$$

and where $W_p\, (n, \underset{\sim}{\Sigma}, \underset{\sim}{\Delta})$ denotes the Wishart distribution for p x p matrices, with n degrees of freedom, covariance parameter matrix $\underset{\sim}{\Sigma}$, and noncentrality parameter matrix The purpose of the algorithm is to compute an approximation to the cumulative distribution function of the *central* distribution (i.e., $\underset{\sim}{\Delta} = \underset{\sim}{0}$) of T_0^2:

$$\text{CDF} = \Pr\, [T_0^2 \leq T \mid \underset{\sim}{\Delta} = \underset{\sim}{0}]. \tag{1}$$

The central distribution is also called the "null" distribution, in part because $\underset{\sim}{\Delta}$ is "null" and in part because this distribution is usually the relevant one under the "null hypothesis" of certain multivariate tests.

METHOD

This algorithm evaluates the cumulative distribution fuction of the distribution of Hotelling's T_0^2 for certain special cases; for combinations of the parameters which do not satisfy the special cases the algorthim evaluates the cdf of a distribution which approximates the T_0^2 distribution. The following notation will be helpful; let

$$U^{(s)} = \text{Trace}\ (\underset{\sim}{H}\ \underset{\sim}{E}^{-1})$$

where $\underset{\sim}{H}$ and $\underset{\sim}{E}$ are as described above and

$$s = \min\ \{n_1,\ p\}$$

is the number of non-zero eigenvalues of $\underset{\sim}{H}\ \underset{\sim}{E}^{-1}$. The case $s = n_1 < p$ leads to difficulties which can be bypassed: the distribution of $U^{(s)}$ can be derived from the distribution of $U^{(p)}$ $[p < n_1]$ by mapping

$$(n_1,\ n_2,\ p) \rightarrow (p,\ n_1 + n_2 - p,\ n_1). \qquad (2)$$

Thus only the case $n_1 \geq p$ need be considered. The algorithm automatically handles this mapping in the case $n_1 < p$.

<u>Special cases</u>. The algorithm produces exact results (to the accuracy of the incomplete beta function algorithm) in the following special cases.

In the case p = 1 the T_0^2 is proportional to an F-statistic:

$$T_0^2 / n_1 \sim F(n_1, n_2),$$

hence,

$$[T_0^2 / (n_2 + T_0^2)] \sim \text{Beta}\left(\frac{n_1}{2}, \frac{n_2}{2}\right),$$

and the algorithm evaluates:

$$\Pr[T_0^2 \leq T] = I_w\left(\frac{n_1}{2}, \frac{n_2}{2}\right)$$

where

$$w = T/(T+n_2) = U/(U+1)$$

$$U = T/n_2.$$

In his original paper Hotelling (1951) derived the exact distribution of T_0^2 for the special case p = 2:

$$\Pr[T_0^2 \leq T] = I_w(n_1-1, n_2) + \sqrt{\pi}\,\frac{\Gamma[(n_1+n_2-1)/2]}{\Gamma(n_1/2)\,\Gamma(n_2/2)}\left(\frac{1-w}{1+w}\right)^{(n_2-1)/2} I_{w^2}\left(\frac{n_1-1}{2}, \frac{n_2+1}{2}\right)$$

where

$$w = T/(2n_2+T) = U/(U+2).$$

This result is used by the algorithm when p=2. It is interesting that Johnson and Kotz (1972, page 199, equation (41)) give an incorrect version of this result.

Note that because of the mapping (2) one of these special cases will be invoked if $n_1 = 1$ or $n_1 = 2$ and $n_1 < p$. Rao(1965, Section 8b(xii)) treats the special case $n_1 = 1$ separately and shows that the cdf of the T_0^2 distribution is equivalent to the incomplete beta function, but because of the mapping (2) no special coding is needed for this case.

Moment-Based Approximations. Pillai and Young(1971) developed the technique of approximating the density of $U^{(p)}$ by a density similar to the density of a central F distribution. The approximation is based upon finding the parameters of the F-type distribution which make the first three moments of the distribution of $U^{(p)}$ equal to the first three moments of the F-type distribution. The first three central moments of the central distribution of $U^{(p)}$ are given by Pillai(1960), Pillai and Samson(1959), and again in Pillai and Young(1971) as:

$$\mu_1 = p\ (2m + p + 1)\ /\ (2n)$$

$$\mu_2 = [p\ (2m + p + 1)\ (2m + 2n + p + 1)\ (2n + p)]/\ [4n^2(n - 1)\ (2n + 1)]$$

$$\mu_3 = \frac{p\ (2m + n + p + 1)\ (2m + p + 1)\ (2m + 2n + p + 1)\ (n + p)\ (2n + p)}{2n^3\ (n - 1)\ (n - 2)\ (n + 1)\ (2n + 1)}$$

where

$$m = (n_1 - p - 1)\ /\ 2$$

$$n = (n_2 - p - 1)\ /\ 2.$$

The F-type density function

$$f(x;\ a,b,K) = x^a/\{\beta\ (a+1,\ b-a-1)\ K^{a+1}\ (1+x/K)^b\} \qquad 0 < x < \infty$$

has the following first three central moments

$$\mu_{F1} = K\ (a+1)\ /\ (b-a-2)$$

$$\mu_{F2} = [K^2\ (a+1)\ (b-1)]\ /\ [\ (b-a-2)^2\ (b-a-3)]$$

$$\mu_{F3} = [2\ K^3\ (a+1)\ (b-1)\ (a+b)]\ /\ (b-a-2)^3\ (b-a-3)\ (b-a-4)].$$

This F-type distribution is identical to the standard central F distribution with k_1 and k_2 degrees of freedom if one imposes the restriction

$$K = k_2\ /\ k_1$$

and uses the relations

$$a = (k_1-2)\ /\ 2$$

$$b = (k_1+k_2)\ /\ 2.$$

However, without the restriction on K the F-type distribution has three parameters: a,b, and K.

Setting the first three moments of the distribution of $U^{(p)}$ and the F-type distribution equal and solving for a,b, and K yields

$$a = (2\mu_1^3\mu_2 + 3\mu_1^2\mu_3 - 6\mu_1\mu_2^2 - \mu_2\mu_3) \,/\, (\mu_2\mu_3 + 4\mu_1\mu_2^2 - \mu_1^2\mu_3)$$

$$b = [\ (a+1)\ (a+3) - \mu_1^2 \,/\, \mu_2] \,/\, [\ (a+1) - \mu_1^2 \,/\, \mu_2]$$

$$K = \mu_1\ (b-a-2) \,/\, (a+1)$$

Note that in Pillai and Young (1971) there is a typographical error in the expression for a; the last minus (-) sign in the expression above is erroneously given as a plus (+) sign in that paper.

The first three moments of these distributions do not exist for all possible values of the parameters. The following is a summary of necessary and sufficient conditions for the existence of the indicated moments:

Distribution

$U^{(p)}$ Moment	$U^{(p)}$ Condition for existence	F-Type Moment	F-Type Condition for existence
μ_1	$n_2 > p+1$	μ_{F1}	$b-a > 2$
μ_2	$n_2 > p+3$	μ_{F2}	$b-a > 3$
μ_3	$n_2 > p+5$	μ_{F3}	$b-a > 4$

In each of these cases it is assumed the mapping (2) has been performed if $n_1 < p$ and the statements and expressions apply to the transformed parameters.

If the combination of parameters is such that one of the third-order moments (μ_3 or μ_{F3}) does not exist, one can obtain a two-moment approximation by equating the first two moments of the two distributions. The following expressions for a, b, and K will generate a two-moment approximation:

$$K = p$$

$$a = [\mu_2 (\mu_1 - p) + \mu_1^2 (\mu_1 + p)]/(p\mu_2)$$

$$b = [\mu_1 (\mu_1 + p)^2 + \mu_1\mu_2 + 2p\mu_2]/(p\mu_2)$$

Similarly, a one-moment approximation is generated by the following expressions:

$$K = p$$

$$a = p (2m + p + 1) /2 - 1$$

$$b = p(2m + 2n + p + 1) /2 + 1.$$

Operational procedure. The algorithm first checks the input parameter values to determine whether they have proper values (e.g., $T \geq 0$, etc.). The algorithm then checks for the special case $n_1 < p$; if so, the mapping (2) is performed. The upper limit, T, is converted to a $U^{(p)}$: $U = T/n_2$. The algorithm next checks for one of the special cases p = 1 or p = 2. If a special case applies, the exact cdf is evaluated. If neither of the special cases apply, the algorithm computes m and n (defined above) and determines whether the third moments of both distributions exist. If both third moments exist the three-moment approximation is used. Otherwise a two-moment approximation is attempted. If one of the second-order moments does not exist, a one-moment approximation is attempted. If one of the first-order moments does not exist the algorithm sets the failure indicator (IFAULT) and returns control to the calling program. If the three-, two-, or one-moment approximation is used, the algorithm uses the incomplete beta function subprogram to evaluate:

$$CDF = I_w(a+1,\ b-a-1)$$
$$\cong Pr[\ T_o^2 \leq T\ ]$$

where

$$w = T/(T+Kn_2) = U/(U + K)$$

and where a, b, and K are evaluated from the expressions of the highest-moment approximation applicable.

STRUCTURE

SUBROUTINE HOTELL (T, N1, N2, IP, CDF, IFAULT, IAPRX)

Formal Parameters:

T	Real	input:	The algorithm computes CDF $\cong \Pr[T_0^2 \leq T]$
N1	Integer	input:	Number of degrees of freedom of the distribution of $\underset{\sim}{H}$ in $T_0^2 = n_2 \mathrm{Tr}(\underset{\sim}{H}\,\underset{\sim}{E}^{-1})$
N2	Integer	input:	Number of degrees of freedom of the distribution of $\underset{\sim}{E}$.
IP	Integer	input:	The matrices H and E are IP x IP; IP corresponds to the parameter p in the discussion.
CDF	Real	output:	CDF $\cong \Pr[T_0^2 \leq T]$.
IFAULT	Integer	output:	Failure indicator.
IAPRX	Integer	output:	Indicates the type of approximation used.

IAPRX=-3:	Exact computation for T=0
IAPRX=-2:	Exact computation for special case p=2
IAPRX=-1:	Exact computation for special case p=1
IAPRX=0:	Algorithm failure; approximation technique not applicable
IAPRX=1:	One-moment approximation used
IAPRX=2:	Two-moment approxmiation used
IAPRX=3:	Three-moment approximation used

Failure Indicators:

IFAULT = 0	indicates no errors were detected.
IFAULT = -1	indicates one of the input parameters has an improper value [i.e., T<0, N1 < 1, N2 < 1, IP < 1]. CDF is returned with a value of -1E75.
IFAULT = 1	indicates the combination of parameters T, N1, N2, and IP is such that the first moment of the approximating distribution is not finite, i.e., the technique is not applicable. CDF is returned with a value of -1E75 and IAPRX = 0.
IFAULT = 2	indicates W≤0 or W ≥ 1 where W = T/(T + K * N2), which should never happen, i.e., IFAULT = 2 indicates a programming error. CDF is returned with a value of -1E75.

RESTRICTIONS

The parameters n_1, n_2, and p must satisfy the restrictions given above for the existence of moments for a particular approximation to apply; i.e., the minimal restriction is $n_2 > p+1$, which applies to both the original parameters and the transformed parameters if the mapping (2) is invoked. The restrictions shown for a and b must also be satisfied, but there is no simple statement of these restrictions in terms of n_1, n_2, and p. These restrictions guarantee that the algorithm will operate as described but do not guarantee that the approximation used will be accurate. Additional restrictions are described in the section discussing accuracy.

AUXILLARY ALGORITHMS

The present version of HOTELL uses BETAIN, Algorithm AS 63, "The Incomplete Beta Integral ," by Majumder and Bhattacharjee (1973). BETAIN requires the prior computation of the complete beta function, $B(p,q) = \Gamma(p)\Gamma(q)/\Gamma(p+q)$. HOTELL assumes the availability of a single precision function subprogram, ALGAMA (X), which evaluates $\log_e \Gamma(X)$. Virtually all computer manufacturers provide an efficient, accurate subprogram for $\log \Gamma(X)$ and such a routine may be easily substituted for ALGAMA in HOTELL. If such a routine is not available the algorithm of Pike and Hill (1960) may be implemented in Fortran under the name ALGAMA.

ACCURACY

Errors in this approximation have three components: (a) differences between the cdf's of the F-type distribution and the distribution of $U^{(p)}$, (b) errors in approximating the incomplete beta function, and (c) calculations performed in the routine.

Pillai and Young (1971) included a comparison of the approximations to the exact distribution for $n = (n_2-p-1)/2 = 5, 10, 15, 20, 30, 40, 50, 60, 80$, and 100, and for $(p,m) = (p, (n_1-p-1)/2) = (3,0), (3,3), (4,0)$, and $(4,2)$. They concluded that this approximation "...provides about three significant digits accuracy in the *percentage points* for $n \geq 10$. In some cases $n \geq 5$ is sufficient for this accuracy" and that "...the distribution function for [this approximation] provides a good approximation [about three decimal places] to the exact distribution of $U^{(p)}$ for $n \geq 10$ and for the whole range of $U^{(p)}$." Note that the first quote refers to the percentage points, i.e., the inverse of the cdf, and the second quote refers to the cdf itself. (Emphasis added.)

In our tests of this routine we reproduced the relevant results of Table II in Pillai and Young (1971) to the number of digits given there.

Errors of types (b) and (c) above are more within the control of the user of the algorithm. If HOTELL is implemented on an IBM 360, 370, or other computer with a short ($\leq$ 24 bits) single precision floating point mantissa, double precision is absolutely essential for accurate computation, especially for the special case $p = 2$. A double precision version is easily produced; one need change only the initial declarations, including the declaration of the arithmetic statement function CBF (complete beta function), the names of the functions EXP, ALGAMA, and ALOG in statement 9, and function BETAIN. In a double precision version of HOTELL (and BETAIN), or a single precision version on machines with long floating point mantissas, the computational errors in

BETAIN and HOTELL will typically be negligible compared to the error of the approximating technique (type (a), above).

REFERENCES

Hotelling, H. (1951) A Generalized T-test and Measure of Multivariate Dispersion, Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability, 23-41.

Johnson, N. L. and S. Kotz (1972). Distributions in Statistics: Continuous Multivariate Distributions. John Wiley & Sons, Inc., New York, 1972.

Majumder, K. L. and G. P. Bhattacharjee (1973). Algorithm AS63. The Incomplete Beta Integral. Applied Statistics, **22**, 409-411.

Pike, M.C. and I. D. Hill (1966). Algorithm 291. Logarithm of the Gamma Function. Commun. Ass. Comput.Mach., **9**, 684.

Pillai, K.C.S. (1960). Statistical Tables for Tests of Multivariate Hypotheses. The Statistical Center, University of the Philippines, Manila.

Pillai, K.C.S. and P. Samson, Jr. (1959). On Hotelling's Generalization of T^2. Biometrika, **46**, 160-165.

Pillai, K.C.S. and D. L. Young (1971). On the Exact Distribution of Hotelling's Generalized T_0^2. Journal of Multivariate Analysis, **1**, 90-107.

Rao, C. R. (1965). Linear Statistical Inference and Its Applications. John Wiley & Sons, Inc., New York, 1965.

SUBROUTINE HOTELL LISTING

```
C***
C
C    SUBROUTINE HOTELL(T,N1,N2,IP,CDF,IFAULT,IAPRX)
C
C
C    DESCRIPTION AND PURPOSE
C
C      HOTELLING'S GENERALIZED T(0)**2 STATISTIC IS DEFINED AS
C
C         T(0)**2 = N2*TRACE(H*E**-1)
C
C      WHERE H AND E ARE INDEPENDENT P X P MATRICES;
C
C         H HAS A W(P;N1,S,D) DISTRIBUTION AND IS POSITIVE SEMIDEFINITE;
C
C         E HAS A W(P;N2,S,0) DISTRIBUTION AND IS POSITIVE DEFINITE;
C
C      AND WHERE W(P;N,S,D) DENOTES THE WISHART DISTRIBUTION FOR P X P
C      MATRICES, WITH N DEGREES OF FREEDOM, COVARIANCE PARAMETER MATRIX
C      S, AND NONCENTRALITY PARAMETER MATRIX D.  THIS ALGORITHM COMPUTES
C      AN APPROXIMATION OF THE CUMULATIVE DISTRIBUTION FUNCTION OF THE
C      CENTRAL DISTRIBUTION (I.E., D=0) OF T(0)**2;  THAT IS, THIS
C      ALGORITHM COMPUTES AN APPROXIMATION TO
C
C         CDF = PR(T(0)**2 < T | D=0)
C
C      THE CENTRAL DISTRIBUTION IS ALSO CALLED THE "NULL" DISTRIBUTION,
C      IN PART BECAUSE D IS "NULL" AND IN PART BECAUSE THIS DISTRIBUTION
C      IS USUALLY THE RELEVANT ONE UNDER THE "NULL HYPOTHESIS" OF CERTAIN
C      MULTIVARIATE TESTS.
C
C
C    METHOD
C
C      THE ALGORITHM PRODUCES EXACT RESULTS (TO THE ACCURACY OF THE
C      INCOMPLETE BETA FUNCTION ALGORITHM) IN THE SPECIAL CASES IP=1,
C      IP=2, AND T=0.
C
C      IF NONE OF THE SPECIAL CASES APPLY, THEN THE ALGORITHM
C      IS AN IMPLEMENTATION OF AN APPROXIMATION DEVELOPED BY PILLAI
C      AND YOUNG (1971).  THE APPROXIMATION IS BASED UPON SELECTING AN
C      F-TYPE DISTRIBUTION WHICH HAS THE SAME FIRST THREE MOMENTS AS
C      THE DISTRIBUTION OF
C
C         U = (T(0)**2)/N2
C
C      THEN   PR(T(0)**2 < T)   IS APPROXIMATED BY   PR(F < T/N2).  AFTER
C      A TRANSFORMATION,   PR(F < T/N2)    IS EVALUATED AS
C      PR(BETA < W = T/(T+K*N2)) = INCOMPLETE BETA FUNCTION EVALUATED AT
C      W, WHERE   BETA   HAS THE APPROPRIATE BETA DISTRIBUTION.
C
C      IF THE THIRD MOMENT OF THE F-TYPE DISTRIBUTION DOES NOT EXIST,
C      BUT THE SECOND MOMENT DOES EXIST, THE APPROXIMATION IS BASED ON
C      THE FIRST TWO MOMENTS.  LIKEWISE, IF BOTH THE THIRD AND SECOND
C      MOMENTS DO NOT EXIST, BUT THE FIRST MOMENT DOES EXIST, THE
C      APPROXIMATION IS BASED ON THE FIRST MOMENT.  IF NONE OF THE
C      THREE MOMENTS EXISTS, THE APPROXIMATION IS NOT APPLICABLE AND
C      THE ALGORITHM DOES NOT ATTEMPT AN APPROXIMATION.
C
C
C    USAGE
C
```

SUBROUTINE HOTELL LISTING

```
C     CALL HOTELL(T,N1,N2,IP,CDF,IFAULT,IAPRX)
C
C
C   DESCRIPTION OF PARAMETERS
C
C
C     T       -REAL INPUT:  THE ALGORITHM COMPUTES
C                           CDF = PR(T(0)**2 < T)   APPROXIMATELY
C
C     N1      -INTEGER INPUT:  NUMBER OF DEGREES OF FREEDOM OF THE
C                              DISTRIBUTION OF H  IN T(0)**2=TR(H*E**-1)
C
C     N2      -INTEGER INPUT:  NUMBER OF DEGREES OF FREEDOM OF THE
C                              DISTRIBUTION OF E
C
C     IP      -INTEGER INPUT:  THE MATRICES H AND E ARE IP X IP; IP
C                              CORRESPONDS TO THE PARAMETER P IN THE
C                              DISCUSSION
C
C     CDF     -REAL OUTPUT:  CDF = PR(T(0)**2 < T)   APPROXIMATELY
C
C
C     IFAULT  -INTEGER FAILURE INDICATOR:
C
C                  IFAULT=0    INDICATES NO ERRORS WERE DETECTED
C
C                  IFAULT=-1   INDICATES ONE OF THE INPUT PARAMETERS
C                              HAS AN IMPROPER VALUE, (I.E. T<0, N1<1,
C                              N2<1, IP<1).  CDF IS RETURNED WITH A
C                              VALUE OF -1E75.
C
C                  IFAULT=1    INDICATES THE COMBINATION OF PARAMETERS
C                              T, N1, N2, AND IP IS SUCH THAT THE
C                              FIRST MOMENT OF THE APPROXIMATING
C                              DISTRIBUTION IS NOT FINITE, I.E., THE
C                              TECHNIQUE IS NOT APPLICABLE.  CDF IS
C                              RETURNED WITH A VALUE OF -1E75 AND
C                              IAPRX=0.
C
C                  IFAULT=2    INDICATES W .LE. 0 OR W .GE. 1 WHERE
C                              W=T/(T+K*N2), WHICH SHOULD NEVER
C                              HAPPEN, I.E., IFAULT=2 INDICATES A
C                              PROGRAMMING ERROR.  CDF IS RETURNED
C                              WITH A VALUE OF -1E75.
C
C
C     IAPRX   -INTEGER INDICATOR OF TYPE OF APPROXIMATION USED
C
C                  IAPRX=-3    EXACT COMPUTATION FOR T=0
C
C                  IAPRX=-2    EXACT COMPUTATION FOR SPECIAL CASE P=2
C
C                  IAPRX=-1    EXACT COMPUTATION FOR SPECIAL CASE P=1
C
C                  IAPRX=0     ALGORITHM FAILURE; APPROXIMATION
C                              TECHNIQUE NOT APPLICABLE
C
C                  IAPRX=1     ONE-MOMENT APPROXIMATION USED
C
C                  IAPRX=2     TWO-MOMENT APPROXIMATION USED
C
C                  IAPRX=3     THREE-MOMENT APPROXIMATION USED
C
C
C   SUBROUTINES REQUIRED
C
```

SUBROUTINE HOTELL LISTING

```
C        BETAIN     (APPLIED STATISTICS ALGORITHM AS63.  SEE FIRST
C                   REFERENCE BELOW)
C
C
C   REFERENCES
C
C    MAJUMDER, K.L. AND G.P. BHATTACHARJEE (1973).  ALGORITHM AS63.
C           THE INCOMPLETE BETA INTEGRAL.  APPLIED STATISTICS, 22,
C           409-411.
C
C     PILLAI, K.C.S. AND D.L. YOUNG (1971).  ON THE EXACT DISTRIBUTION
C           OF HOTELLING'S GENERALIZED T(0)**2.  JOURNAL OF
C           MULTIVARIATE ANALYSIS, 1, 90-107.
C
C
C
C
C   SUBROUTINE HOTELL(T,N1,N2,IP,CDF,IFAULT,IAPRX)
C
C     THIS ALGORITHM COMPUTES AN APPROXIMATION TO THE CUMULATIVE
C     DISTRIBUTION FUNCTION OF THE CENTRAL DISTRIBUTION OF HOTELLING-S
C     GENERALIZED   T(0)**2   STATISTIC, I.E.,
C
C        CDF = PR(T(0)**2 .LT. T), APPROXIMATELY.
C
C
C        VERSION DATE  14 JULY 1975
C
      SUBROUTINE HOTELL(T,N1,N2,IP,CDF,IFAULT,IAPRX)
C
C   DECLARATIONS
C
C        ARGUMENTS
C
      REAL    CDF,T
      INTEGER N1,N2,IP,IFAULT,IAPRX
C
C        FUNCTION SUBPROGRAMS
C
      REAL    BETAIN,CBF
      CBF(A,B)= EXP(ALGAMA(A)+ALGAMA(B)-ALGAMA(A+B))
C
C        INTERNAL VARIABLES
C
      REAL    A,A1,A2,B,B1,BETA,TADD,TEMP1,TERM1,TERM2,TK,TM,TMU1,TMU2
      REAL    TMU3,TMU1SQ,TMU2SQ,TN,TN1,TN2,TP,TWOTN,U,W
C
C        INTERNAL CONSTANTS
C
      REAL    ZERO/0E0/,ONE/1E0/,TWO/2E0/,THREE/3E0/,FOUR/4E0/,SIX/6E0/
      REAL SQRTPI/1.7724539/
C
C   BEGIN EXECUTION
C
      IAPRX = 0
      IFAULT = -1
      CDF = -1E75
C
C   CHECK FOR T .LE. ZERO  OR INVALID VALUES OF N1, N2, AND IP
C
      IF(N1 .LT. 1  .OR.  N2 .LT. 1  .OR.  IP .LT. 1) GO TO 20
      IF(T)20, 2, 3
    2 CDF=ZERO
      IAPRX = -3
      IFAULT = 0
      GO TO 20
```

SUBROUTINE HOTELL LISTING

```
C
C   IF N1 .LT. IP,
C
C     (N1,N2,IP) MAPS TO (IP,N1+N2-P,N1)
C
    3 IF (N1 .GE. IP)  GO TO 4
      TN1 = IP
      TN2 = N1+N2-IP
      TP = N1
      GO TO 5
    4 TN1 = N1
      TN2 = N2
      TP = IP
    5 U = T/N2
C
C   CHECK FOR SPECIAL CASE P = 1
C
      IF (TP. GT. ONE) GO TO 6
      IAPPX = -1
      A1=TN1/TWO
      B1=TN2/TWO
      W=U/(U+ONE)
      GO TO 15
C
C   CHECK FOR SPECIAL CASE P=2
C
    6 IF (TP .GT. TWO) GO TO 10
      IAPPX=-2
C
C      NOTE THAT TN1 .GT. 1  BECAUSE 2 = TP .LE. TN1
C
      A1=TN1-ONE
      W=U/(U+TWO)
      BETA=CBF(A1,TN2)
      TERM1=BETAIN(W,A1,TN2,BETA,IFAULT)
      IF(IFAULT .NE. 0) GO TO 20
      A1=A1/TWO
      A2=(TN2+ONE)/TWO
      TEMP1=W*W
      BETA=CBF(A1,A2)
      TERM2=BETAIN(TEMP1,A1,A2,BETA,IFAULT)
      IF(IFAULT .NE. 0) GO TO 20
    9 TERM2=TERM2*SQRTPI* EXP( ALGAMA((TN1+TN2-ONE)/TWO)
     1  -ALGAMA(TN1/TWO) -ALGAMA(TN2/TWO)
     2 +((TN2-ONE)/TWO)*ALOG((ONE-W)/(ONE+W)))
      CDF=TERM1-TERM2
      GO TO 20
C
C   COMPUTE M (IN TM) AND N (IN TN)
C
   10 TM = (TN1-TP-ONE)/TWO
      TN = (TN2-TP-ONE)/TWO
C
C   IF TN .LE. 0 THE FIRST MOMENT (TMU1) OF THE DISTRIBUTION OF T(0)**2
C   DOES NOT EXIST AND THE APPROXIMATION TECHNIQUE IS INVALID.
C
      IF(TN .LE. ZERO) GO TO 13
      TWOTN = TWO*TN
      TADD = TWO*TM+TP+ONE
      TMU1 = TP*TADD/TWOTN
C
C   IF TN .LE. 1 THE SECOND MOMENT (TMU2) DOES NOT EXIST
C
      IF(TN .LE. ONE)GO TO 12
      TMU2 = TMU1*(TWOTN+TADD)*(TWOTN+TP)/(TWOTN*(TN-ONE)*(TWOTN+ONE))
      TMU1SQ= TMU1*TMU1
```

SUBROUTINE HOTELL LISTING

```
C
C   IF TN .LE. 2 THE THIRD MOMENT (TMU3) DOES NOT EXIST
C
      IF(TN .LE. TWO)GO TO 11
      TMU3 = (TWO*TMU2*(TN+TADD)*(TN+TP))/(TN*(TN-TWO)*(TN+ONE))
      TMU2SQ = TMU2*TMU2
C
C   COMPUTE A AND B FOR THE THREE-MOMENT APPROXIMATION
C
      A = (TMU1*(-SIX*TMU2SQ+TMU1*(THREE*TMU3+TWO*TMU1*TMU2))-TMU2*TMU3)
     1   /(TMU2*TMU3+FOUR*TMU1*TMU2SQ-TMU1SQ*TMU3)
      B = ((A+ONE)*(A+THREE)-TMU1SQ/TMU2)/((A+ONE)-TMU1SQ/TMU2)
C
C   CHECK FOR (B-A) .LE. 4
C
      IF (B-A .LE. FOUR) GO TO 11
      IAPRX = 3
      TK = TMU1*(B-A-TWO)/(A+ONE)
      W=U/(U+TK)
C
C   W  IS THE UPPER LIMIT FOR USE WITH THE CUMULATIVE DISTRIBUTION
C   FUNCTION OF THE   BETA(A1,B1)   DISTRIBUTION;
C
C     PR(T(0)**2 .LT. T) = PR(BETA .LT. W), APPROXIMATELY
C
      GO TO 14
C
C   COMPUTE A AND B FOR THE TWO-MOMENT APPROXIMATION
C
   11 TEMP1 = TP*TMU2
      A = (TMU2*(TMU1-TP)+TMU1SQ*(TMU1+TP))/TEMP1
      B = (TMU1*(TMU1+TP)**2+TMU1*TMU2+TWO*TEMP1)/TEMP1
C
C   CHECK FOR B-A .LE. 3
C
      IF (B-A .LE. THREE) GO TO 12
      IAPRX = 2
      W=U/(U+TP)
      GO TO 14
C
C   COMPUTE A AND B FOR THE ONE-MOMENT APPROXIMATION
C
   12 A = TP*TADD/TWO-ONE
      B = TP*(TADD+TWOTN)/TWO+ONE
C
C   CHECK FOR B-A .LE. 2
C
      IF (B-A .LE. TWO) GO TO 13
      IAPRX = 1
      W=U/(U+TP)
      GO TO 14
C
C   WHEN B-A .LE. 2 THE APPROXIMATION TECHNIQUE IS NOT APPLICABLE.
C
   13 IAPRX = 0
      IFAULT = 1
      GO TO 20
   14 A1=A+ONE
      B1=B-A-ONE
C
C   A CHECK IS MADE TO INSURE THAT THE PARAMETERS  W, A1, B1,  ARE
C   WITHIN THE RANGE OF THE INCOMPLETE BETA FUNCTION.
C
   15 IF (W .LT. ZERO) GO TO 20
      IF (A1 .LE. ZERO .OR. B1 .LE. ZERO) GO TO 20
      BETA=CBF(A1,B1)
```

SUBROUTINE HOTELL LISTING

```
      CDF=BETAIN(W,A1,B1,BETA,IFAULT)                                    HOTL3260
      IF (IFAULT .NE. 0) CDF = -1E75                                     HOTL3270
   20 RETURN                                                             HOTL3280
      END                                                                HOTL3290
```

SUBROUTINE BETAIN LISTING

```
C***
      FUNCTION BETAIN(X,P,Q,BETA,IFAULT)
C
C        ALGORITHM AS 63  APPL STATIST. (1973), VOL.22, NO.3
C
C        COMPUTES INCOMPLETE BETA INTEGRAL FOR ARGUMENTS
C        X BETWEEN ZERO AND ONE, P AND Q POSITIVE.
C        COMPLETE BETA FUNCTION IS ASSUMED TO BE KNOWN.
C
      LOGICAL INDEX
C
C        DEFINE ACCURACY AND INITIALISE
C
      DATA ACU /0.1E-7/
      BETAIN = X
C
C        TEST FOR ADMISSIBILITY OF ARGUMENTS
C
      IFAULT = 1
      IF (P.LE.0.0.OR.Q.LE.0.0) RETURN
      IFAULT = 2
      IF (X.LT.0.0.OR.X.GT.1.0) RETURN
      IFAULT = 0
      IF (X.EQ.0.0.OR.X.EQ.1.0) RETURN
C
C        CHANGE TAIL IF NECESSARY AND DETERMINE S
C
      PSQ = P + Q
      CX = 1.0 - X
      IF (P.GE.PSQ*X) GOTO 1
      XX = CX
      CX = X
      PP = Q
      QQ = P
      INDEX = .TRUE.
      GOTO 2
    1 XX = X
      PP = P
      QQ = Q
      INDEX = .FALSE.
    2 TERM = 1.0
      AI = 1.0
      BETAIN = 1.0
      NS = QQ + CX * PSQ
C
C        USE SOPER-S REDUCTION FORMULAE.
C
      RX = XX/CX
    3 TEMP = QQ - AI
      IF (NS.EQ.0) RX = XX
    4 TERM = TERM * TEMP * RX / (PP + AI)
      BETAIN = BETAIN + TERM
      TEMP = ABS(TERM)
      IF ( TEMP.LE.ACU.AND.TEMP.LE.ACU*BETAIN) GOTO 5
      AI = AI + 1.0
      NS = NS - 1
      IF (NS.GE.0) GOTO 3
      TEMP = PSQ
      PSQ = PSQ + 1.0
      GOTO 4
C
C        CALCULATE RESULT
C
```

SUBROUTINE BETAIN LISTING

```
5 BETAIN = BETAIN * EXP(PP*ALOG(XX)+(QQ-1.0)*ALOG(CX))/(PP*BETA)
  IF (INDEX) BETAIN = 1.0 - BETAIN
  RETURN
  END
```

THE FOLLOWING RESULTS OBTAINED FROM SUBROUTINE HOTELL AGREE WITH THOSE OF TABLE II IN PILLAI AND YOUNG (1971)

N1	N2	P	T/N2	CDF
4	14	3	2.5064	.949997
4	14	3	3.6951	.990000
4	24	3	1.1562	.949993
4	24	3	1.5623	.990000
4	34	3	.74777	.950001
4	34	3	.98252	.990000
4	44	3	.55207	.950001
4	44	3	.71559	.990000
4	64	3	.36218	.949999
4	64	3	.46326	.990000
4	84	3	.26943	.949997
4	84	3	.34239	.990000
4	104	3	.21449	.950002
4	104	3	.27151	.989998
4	124	3	.17815	.949996
4	124	3	.22494	.989999
4	164	3	.13306	.949992
4	164	3	.16748	.990001
4	204	3	.10619	.950009
4	204	3	.13340	.990002
10	14	3	5.4723	.949998
10	14	3	7.6114	.990000
10	24	3	2.4700	.950005
10	24	3	3.1258	.990000
10	34	3	1.5845	.950006
10	34	3	.19465	.989999
10	44	3	1.1649	.950006
10	44	3	1.4107	.990003
10	64	3	.76090	.949997
10	64	3	.90866	.990000
10	84	3	.56480	.950003
10	84	3	.66987	.990000
10	104	3	.44901	.949997
10	104	3	.53039	.990000
10	124	3	.37261	.950000
10	124	3	.43896	.990000
10	164	3	.27799	.950004
10	164	3	.32640	.990001
10	204	3	.22168	.949991
10	204	3	.25977	.989998
5	15	4	3.8146	.950001
5	15	4	5.3980	.990000
5	25	4	1.7419	.950004
5	25	4	2.2532	.990001
5	35	4	1.1230	.950008
5	35	4	1.4127	.989999
5	45	4	.82786	.949999
5	45	4	1.0277	.990003
5	65	4	.54237	.950001
5	65	4	.66464	.990000

N1	N2	P	T/N2	CDF
5	85	4	.40321	.949996
5	85	4	.49103	.990000
5	105	4	.32087	.950002
5	105	4	.38930	.989999
5	125	4	.26645	.950005
5	125	4	.32248	.989999
5	165	4	.19895	.949998
5	165	4	.24007	.990001
5	205	4	.15874	.950005
5	205	4	.19120	.990001
9	15	4	6.3433	.950001
9	15	4	8.6636	.990000
9	25	4	2.8631	.950001
9	25	4	3.5626	.990001
9	35	4	1.8384	.950004
9	35	4	2.2236	.990002
9	45	4	1.3525	.949994
9	45	4	1.6140	.990002
9	65	4	.88428	.949999
9	65	4	1.0416	.990001
9	85	4	.65673	.950001
9	85	4	.76869	.990000
9	105	4	.52228	.949997
9	105	4	.60905	.990000
9	125	4	.43352	.950002
9	125	4	.50430	.990001
9	165	4	.32353	.949997
9	165	4	.37521	.990000
9	205	4	.25806	.950006
9	205	4	.29874	.990002