NAG Fortran Library Routine Document

G08ALF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of ***bold italicised*** terms and other implementation-dependent details.

1 Purpose

G08ALF performs the Cochran Q-test on cross-classified binary data.

2 Specification

```
SUBROUTINE G08ALF(N, K, X, LDX, Q, PROB, IFAIL)
INTEGER            N, K, LDX, IFAIL
real               X(LDX,K), Q, PROB
```

3 Description

Cochran's Q-test may be used to test for differences between k treatments applied independently to n individuals or blocks (k related samples of equal size n), where the observed response can take only one of two possible values; for example a treatment may result in a 'success' or 'failure'. The data is recorded as either 1 or 0 to represent this dichotomization.

The use of this 'randomized block design' allows the effect of differences between the blocks to be separated from the differences between the treatments. The test assumes that the blocks were randomly selected from all possible blocks and that the result may be one of two possible outcomes common to all treatments within blocks.

The null and alternative hypotheses to be tested may be stated as follows.

H_0 : the treatments are equally effective, that is the probability of obtaining a 1 within a block is the same for each treatment.

H_1 : there is a difference between the treatments, that is the probability of obtaining a 1 is not the same for different treatments within blocks.

The data is often represented in the form of a table with the n rows representing the blocks and the k columns the treatments. Let R_i represent the row totals, for $i = 1, 2, \ldots, n$, and C_j represent the column totals, for $j = 1, 2, \ldots, k$. Let x_{ij} represent the response or result where $x_{ij} = 0$ or 1.

	Treatments				
Blocks	1	2		k	Row Totals
1	x_{11}	x_{12}	$\cdots$	x_{1k}	R_1
2	x_{21}	x_{22}	$\cdots$	x_{2k}	R_2
$\vdots$			$\vdots$		$\vdots$
n	x_{n1}	x_{n2}	$\cdots$	x_{nk}	R_n
Column Totals	C_1	C_2		C_k	$N =$ Grand Total

If $p_{ij} = \Pr(x_{ij} = 1)$, for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$, then the hypotheses may be restated as follows

H_0 : $p_{i1} = p_{i2} = \ldots = p_{ik}$, for each $i = 1, 2, \ldots, n$.

H_1: $p_{ij} \neq p_{ik}$, for some j and k, and for some i.

The test statistic is defined as

$$Q = \frac{k(k-1)\sum_{j=1}^{k}\left(C_j - \frac{N}{k}\right)^2}{\sum_{i=1}^{n} R_i(k - R_i)}.$$

When the number of blocks, n, is large relative to the number of treatments, k, Q has an approximate χ^2 distribution with $k-1$ degrees of freedom. This is used to find the probability, p, of obtaining a statistic greater than or equal to the computed value of Q. Thus p is the upper-tail probability associated with the computed value of Q, where the χ^2 distribution is used to approximate the true distribution of Q.

4 References

Conover W J (1980) *Practical Nonparametric Statistics* Wiley

Siegel S (1956) *Nonparametric Statistics for the Behavioral Sciences* McGraw-Hill

5 Parameters

1: N – INTEGER *Input*

On entry: the number of blocks, n.

Constraint: $N \geq 2$.

2: K – INTEGER *Input*

On entry: the number of treatments, k.

Constraint: $K \geq 2$.

3: X(LDX,K) – ***real*** array *Input*

On entry: the matrix of observed zero-one data. $X(i,j)$ must contain the value x_{ij}, for $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

Constraint: $X(i,j) = 0.0$ or 1.0, for all $i = 1, 2, \ldots, n$; $j = 1, 2, \ldots, k$.

4: LDX – INTEGER *Input*

On entry: the first dimension of the array X as declared in the (sub)program from which G08ALF is called.

Constraint: $LDX \geq N$.

5: Q – ***real*** *Output*

On exit: the value of the Cochran Q-test statistic.

6: PROB – ***real*** *Output*

On exit: the upper tail probability, p, associated with the Cochran Q-test statistic, that is the probability of obtaining a value greater than or equal to the observed value (the output value of Q).

7: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, for users not familiar with this parameter the recommended value is 0. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**

6 Error Indicators and Warnings

If on entry IFAIL $= 0$ or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

IFAIL $= 1$

On entry, N < 2,
or K < 2,
or LDX $<$ N.

IFAIL $= 2$

On entry, X$(i,j) \neq 0.0$ or 1.0 for some i and j, $i = 1,2,\ldots,n$; $j = 1,2,\ldots,k$.

IFAIL $= 3$

The approximation process used to calculate the tail probability has failed to converge. The result returned in PROB may still be a reasonable approximation.

7 Accuracy

The use of the χ^2 distribution as an approximation to the true distribution of the Cochran Q-test statistic improves as k increases and as n increases relative to k. This approximation should be a reasonable one when the total number of observations left, after omitting those rows containing all 0 or 1, is greater than about 25 and the number of rows left is larger than 5.

8 Further Comments

None.

9 Example

The following example is taken from page 201 of Conover (1980). The data represents the success of three basketball enthusiasts in predicting the outcome of 12 collegiate basketball games, selected at random, using 1 for successful prediction of the outcome and 0 for unsuccessful prediction. This data is read in and the Cochran Q-test statistic and its corresponding upper-tail probability are computed and printed.

9.1 Program Text

Note: the listing of the example program presented below uses ***bold italicised*** terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*     G08ALF Example Program Text
*     Mark 15 Release. NAG Copyright 1991.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          LDX, KMAX
      PARAMETER        (LDX=20,KMAX=3)
*     .. Local Scalars ..
      real             DF, P, Q
      INTEGER          I, IFAIL, J, K, N
*     .. Local Arrays ..
      real             X(LDX,KMAX)
*     .. External Subroutines ..
```

```
      EXTERNAL         G08ALF
*     .. Intrinsic Functions ..
      INTRINSIC        real
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G08ALF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N, K
      IF (N.LE.LDX .AND. K.LE.KMAX) THEN
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Cochrans Q test'
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Data matrix'
         DO 20 I = 1, N
            READ (NIN,*) (X(I,J),J=1,K)
            WRITE (NOUT,99999) (X(I,J),J=1,K)
   20    CONTINUE
         IFAIL = 0
*
         CALL G08ALF(N,K,X,LDX,Q,P,IFAIL)
*
         DF = real(K-1)
         WRITE (NOUT,*)
         WRITE (NOUT,99998) 'Cochrans Q test statistic = ', Q
         WRITE (NOUT,99997) 'Degrees of freedom = ', DF
         WRITE (NOUT,99998) 'Upper-tail probability = ', P
      END IF
      STOP
*
99999 FORMAT (1X,3F6.1)
99998 FORMAT (1X,A,F12.4)
99997 FORMAT (1X,A,F6.1)
      END
```

9.2 Program Data

```
G08ALF Example Program Data
 12  3
 1 1 1
 1 1 1
 0 1 0
 1 1 0
 0 0 0
 1 1 1
 1 1 1
 1 1 0
 0 0 1
 0 1 0
 1 1 1
 1 1 1
```

9.3 Program Results

```
 G08ALF Example Program Results

 Cochrans Q test

 Data matrix
     1.0   1.0   1.0
     1.0   1.0   1.0
     0.0   1.0   0.0
     1.0   1.0   0.0
     0.0   0.0   0.0
     1.0   1.0   1.0
     1.0   1.0   1.0
     1.0   1.0   0.0
     0.0   0.0   1.0
     0.0   1.0   0.0
     1.0   1.0   1.0
     1.0   1.0   1.0

 Cochrans Q test statistic =       2.8000
 Degrees of freedom =    2.0
 Upper-tail probability =      0.2466
```