NAG Fortran Library Routine Document

G08AHF

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

G08AHF performs the Mann–Whitney U test on two independent samples of possibly unequal size.

2 Specification

```
      SUBROUTINE G08AHF(N1, X, N2, Y, TAIL, U, UNOR, P, TIES, RANKS, WRK,
     1                  IFAIL)
      INTEGER           N1, N2, IFAIL
      real              X(N1), Y(N2), U, UNOR, P, RANKS(N1+N2), WRK(N1+N2)
      LOGICAL           TIES
      CHARACTER*1       TAIL
```

3 Description

The Mann–Whitney U test investigates the difference between two populations defined by the distribution functions $F(x)$ and $G(y)$ respectively. The data consist of two independent samples of size n_1 and n_2, denoted by $x_1, x_2, \ldots, x_{n_1}$ and $y_1, y_2, \ldots, y_{n_2}$, taken from the two populations.

The hypothesis under test, H_0, often called the null hypothesis, is that the two distributions are the same, that is $F(x) = G(x)$, and this is to be tested against an alternative hypothesis H_1 which is

H_1 : $F(x) \neq G(y)$; or

H_1 : $F(x) < G(y)$, i.e., the x's tend to be greater than the y's; or

H_1 : $F(x) > G(y)$, i.e., the x's tend to be less than the y's,

using a two-tailed, upper-tailed or lower-tailed probability respectively. The user selects the alternative hypothesis by choosing the appropriate tail probability to be computed (see the description of argument TAIL in Section 5).

Note that when using this test to test for differences in the distributions one is primarily detecting differences in the location of the two distributions. That is to say, if we reject the null hypothesis H_0 in favour of the alternative hypothesis H_1: $F(x) > G(y)$ we have evidence to suggest that the location, of the distribution defined by $F(x)$, is less than the location, of the distribution defined by $G(y)$.

The Mann–Whitney U test differs from the Median test (see G08ACF) in that the ranking of the individual scores within the pooled sample is taken into account, rather than simply the position of a score relative to the median of the pooled sample. It is therefore a more powerful test if score differences are meaningful.

The test procedure involves ranking the pooled sample, average ranks being used for ties. Let r_{1i} be the rank assigned to x_i, $i = 1, 2, \ldots, n_1$ and r_{2j} the rank assigned to y_j, $j = 1, 2, \ldots, n_2$. Then the test statistic U is defined as follows;

$$U = \sum_{i=1}^{n_1} r_{1i} - \frac{n_1(n_1+1)}{2}$$

U is also the number of times a score in the second sample precedes a score in the first sample (where we only count a half if a score in the second sample actually equals a score in the first sample).

G08AHF returns:

(a) The test statistic U.

(b) The approximate Normal test statistic,

$$z = \frac{U - \text{mean}(U) \pm \frac{1}{2}}{\sqrt{\text{var}(U)}}$$

where

$$\text{mean}(U) = \frac{n_1 n_2}{2}$$

and

$$\text{var}(U) = \frac{n_1 n_2 (n_1 + n_2 + 1)}{12} - \frac{n_1 n_2}{(n_1 + n_2)(n_1 + n_2 - 1)} \times TS$$

where

$$TS = \sum_{j=1}^{\tau} \frac{(t_j)(t_j - 1)(t_j + 1)}{12}$$

τ is the number of groups of ties in the sample and t_j is the number of ties in the jth group.

Note that if no ties are present the variance of U reduces to $\frac{n_1 n_2}{12}(n_1 + n_2 + 1)$.

(c) An indicator as to whether ties were present in the pooled sample or not.

(d) The tail probability, p, corresponding to U (adjusted to allow the complement to be used in an upper one-tailed or a two-tailed test), depending on the choice of TAIL, i.e., the choice of alternative hypothesis, H_1. The tail probability returned is an approximation of p is based on an approximate Normal statistic corrected for continuity according to the tail specified. If n_1 and n_2 are not very large an exact probability may be desired. For the calculation of the exact probability see G08AJF (no ties in the pooled sample) or G08AKF (ties in the pooled sample).

The value of p can be used to perform a significance test on the null hypothesis H_0 against the alternative hypothesis H_1. Let α be the size of the significance test (that is, α is the probability of rejecting H_0 when H_0 is true). If $p < \alpha$ then the null hypothesis is rejected. Typically α might be 0.05 or 0.01.

4 References

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

Neumann N (1988) Some procedures for calculating the distributions of elementary nonparametric teststatistics *Statistical Software Newsletter* **14 (3)** 120–126

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

5 Parameters

1: N1 – INTEGER *Input*

On entry: the size of the first sample, n_1.

Constraint: N1 ≥ 1.

2: X(N1) – ***real*** array *Input*

On entry: the first vector of observations, $x_1, x_2, \ldots, x_{n_1}$.

3: N2 – INTEGER *Input*

On entry: the size of the second sample, n_2.

Constraint: N2 ≥ 1.

4: Y(N2) – ***real*** array *Input*

On entry: the second vector of observations. $y_1, y_2, \ldots, y_{n_2}$.

5: TAIL – CHARACTER*1 *Input*

On entry: indicates the choice of tail probability, and hence the alternative hypothesis.

If TAIL = 'T', then a two-tailed probability is calculated and the alternative hypothesis is $H_1 : F(x) \neq G(y)$.

If TAIL = 'U', then an upper-tailed probability is calculated and the alternative hypothesis $H_1 : F(x) < G(y)$, i.e., the x's tend to be greater than the y's.

If TAIL = 'L', then a lower-tailed probability is calculated and the alternative hypothesis $H_1 : F(x) > G(y)$, i.e., the x's tend to be less than the y's.

Constraint: TAIL = 'T', 'U' or 'L'.

6: U – ***real*** *Output*

On exit: the Mann–Whitney rank sum statistic, U.

7: UNOR – ***real*** *Output*

On exit: the approximate Normal test statistic, z, as described in Section 3.

8: P – ***real*** *Output*

On exit: the tail probability, p, as specified by the parameter TAIL.

9: TIES – LOGICAL *Output*

On exit: indicates whether the pooled sample contained ties or not. This will be useful in checking which routine to use should one wish to calculate an exact tail probability.

TIES = .FALSE., no ties were present (use G08AJF for an exact probability).

TIES = .TRUE., ties were present (use G08AKF for an exact probability).

10: RANKS(N1+N2) – ***real*** array *Output*

On exit: contains the ranks of the pooled sample. The ranks of the first sample are contained in the first N1 elements and those of the second sample are contained in the next N2 elements.

11: WRK(N1+N2) – ***real*** array *Workspace*

12: 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, N1 < 1,
or N2 < 1.

IFAIL $= 2$

On entry, TAIL $\neq$ 'T', 'U' or 'L'.

IFAIL $= 3$

The pooled sample values are all the same, that is the variance of U $= 0.0$.

7 Accuracy

The approximate tail probaility, p, returned by G08AHF is a good approximation to the exact probability for cases where $\max(n_1, n_2) \geq 30$ and $(n_1 + n_2) \geq 40$. The relative error of the approximation should be less than 10 percent, for most cases falling in this range.

8 Further Comments

The time taken by the routine increases with n_1 and n_2.

9 Example

The example program performs the Mann–Whitney test on two independent samples of sizes 16 and 23 respectively. This is used to test the null hypothesis that the distributions of the two populations from which the samples were taken are the same against the alternative hypothesis that the distributions are different. The test statistic, the approximate Normal statistic and the approximate two-tail probability are printed. An exact tail probability is also calculated and printed depending on whether ties were found in the pooled sample or not.

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.

```
*     G08AHF Example Program Text
*     Mark 14 Release.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          MAXN1, MAXN2, MAXLW, MAXIW
      PARAMETER        (MAXN1=25,MAXN2=25,MAXLW=8000,MAXIW=100)
*     .. Local Scalars ..
      real             P, PEXACT, U, UNOR
      INTEGER          I, IFAIL, LWRK, N, N1, N2, NSUM
      LOGICAL          TIES
*     .. Local Arrays ..
      real             RANKS(MAXN1+MAXN2), WRK(MAXLW), X(MAXN1),
     +                 Y(MAXN2)
      INTEGER          IWRK(MAXIW)
*     .. External Subroutines ..
      EXTERNAL         G08AHF, G08AJF, G08AKF
*     .. Intrinsic Functions ..
      INTRINSIC        INT, MIN
*     .. Executable Statements ..
```

```
      WRITE (NOUT,*) 'G08AHF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N1, N2
      WRITE (NOUT,*)
      WRITE (NOUT,99999) 'Sample size of group 1 = ', N1
      WRITE (NOUT,99999) 'Sample size of group 2 = ', N2
      WRITE (NOUT,*)
      IF (N1.LE.MAXN1 .AND. N2.LE.MAXN2) THEN
         READ (NIN,*) (X(I),I=1,N1)
         WRITE (NOUT,*) 'Mann-Whitney U test'
         WRITE (NOUT,*)
         WRITE (NOUT,*) 'Data values'
         WRITE (NOUT,*)
         WRITE (NOUT,99998) '    Group 1  ', (X(I),I=1,N1)
         READ (NIN,*) (Y(I),I=1,N2)
         WRITE (NOUT,*)
         WRITE (NOUT,99998) '    Group 2  ', (Y(I),I=1,N2)
         IFAIL = 0
*
         CALL G08AHF(N1,X,N2,Y,'Lower-tail',U,UNOR,P,TIES,RANKS,WRK,
     +               IFAIL)
*
         WRITE (NOUT,*)
         WRITE (NOUT,99997) 'Test statistic           = ', U
         WRITE (NOUT,99997) 'Normal Statistic         = ', UNOR
         WRITE (NOUT,99997) 'Approx. tail probability = ', P
         WRITE (NOUT,*)
         IF ( .NOT. TIES) THEN
            WRITE (NOUT,*) 'There are no ties in the pooled sample'
            LWRK = INT(N1*N2/2) + 1
*
            CALL G08AJF(N1,N2,'Lower-tail',U,PEXACT,WRK,LWRK,IFAIL)
*
         ELSE
            WRITE (NOUT,*) 'There are ties in the pooled sample'
            N = MIN(N1,N2)
            NSUM = N1 + N2
            LWRK = N + N*(N+1)*NSUM - N*(N+1)*(2*N+1)/3 + 1
*
            CALL G08AKF(N1,N2,'Lower-tail',RANKS,U,PEXACT,WRK,LWRK,IWRK,
     +                  IFAIL)
*
         END IF
         WRITE (NOUT,*)
         WRITE (NOUT,99997) 'Exact tail probability   = ', PEXACT
      ELSE
         WRITE (NOUT,*) 'Either N1 or N2 is out of range :'
         WRITE (NOUT,99996) 'N1 = ', N1, ' and N2 = ', N2
      END IF
      STOP
*
99999 FORMAT (1X,A,I5)
99998 FORMAT (1X,A,8F5.1,2(/14X,8F5.1))
99997 FORMAT (1X,A,F10.4)
99996 FORMAT (1X,A,I16,A,I16)
      END
```

9.2 Program Data

```
G08AHF Example Program Data
 16 23
 13.0  6.0 12.0  7.0 12.0  7.0 10.0  7.0
 10.0  7.0 16.0  7.0 10.0  8.0  9.0  8.0
 17.0  6.0 10.0  8.0 15.0  8.0 15.0 10.0 15.0 10.0 14.0 10.0
 14.0 11.0 14.0 11.0 13.0 12.0 13.0 12.0 13.0 12.0 12.0
```

9.3 Program Results

```
G08AHF Example Program Results

Sample size of group 1 =    16
Sample size of group 2 =    23

Mann-Whitney U test

Data values

    Group 1   13.0  6.0 12.0  7.0 12.0  7.0 10.0  7.0
              10.0  7.0 16.0  7.0 10.0  8.0  9.0  8.0


    Group 2   17.0  6.0 10.0  8.0 15.0  8.0 15.0 10.0
              15.0 10.0 14.0 10.0 14.0 11.0 14.0 11.0
              13.0 12.0 13.0 12.0 13.0 12.0 12.0

Test statistic           =    86.0000
Normal Statistic         =    -2.8039
Approx. tail probability =     0.0025

There are ties in the pooled sample

Exact tail probability   =     0.0020
```