G13ABF – NAG Fortran Library Routine Document

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

G13ABF computes the sample autocorrelation function of a time series. It also computes the sample mean, the sample variance and a statistic which may be used to test the hypothesis that the true autocorrelation function is zero.

2 Specification

```
SUBROUTINE G13ABF(X, NX, NK, XM, XV, R, STAT, IFAIL)
INTEGER           NX, NK, IFAIL
real              X(NX), XM, XV, R(NK), STAT
```

3 Description

The data consist of n observations x_i, for $i = 1, 2, \ldots, n$ from a time series.

The quantities calculated are

(a) The sample mean

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

(b) The sample variance (for $n \geq 2$)

$$s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{(n-1)}$$

(c) The sample autocorrelation coefficients of lags $k = 1, 2, \ldots, K$, where K is a user-specified maximum lag, and $K < n$, $n > 1$.

The coefficient of lag k is defined as

$$r_k = \frac{\sum_{i=1}^{n-k}(x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

See page 496 *et seq.* of Box and Jenkins [1] for further details.

(d) A test statistic defined as

$$\text{STAT} = n\sum_{k=1}^{K} r_k^2,$$

which can be used to test the hypothesis that the true autocorrelation function is identically zero.

If n is large and K is much smaller than n, STAT has a χ^2_K distribution under the hypothesis of a zero autocorrelation function. Values of STAT in the upper tail of the distribution provide evidence against the hypothesis; G01ECF can be used to compute the tail probability.

Section 8.2.2 of [1] provides further details of the use of STAT.

4 References

[1] Box G E P and Jenkins G M (1976) *Time Series Analysis: Forecasting and Control* Holden–Day (Revised Edition)

5 Parameters

1: X(NX) — ***real*** array *Input*

On entry: the time series, x_i, for $i = 1, 2, \ldots, n$.

2: NX — INTEGER *Input*

On entry: the number of values, n, in the time series.

Constraint: NX > 1.

3: NK — INTEGER *Input*

On entry: the number of lags, K, for which the autocorrelations are required. The lags range from 1 to K and do not include zero.

Constraint: $0 <$ NK $<$ NX.

4: XM — ***real*** *Output*

On exit: the sample mean of the input time series.

5: XV — ***real*** *Output*

On exit: the sample variance of the input time series.

6: R(NK) — ***real*** array *Output*

On exit: the sample autocorrelation coefficient relating to lag k, for $k = 1, 2, \ldots, K$.

7: STAT — ***real*** *Output*

On exit: the statistic used to test the hypothesis that the true autocorrelation function of the time series is identically zero.

8: IFAIL — INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. For users not familiar with this parameter (described in Chapter P01) the recommended value is 0.

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

6 Error Indicators and Warnings

Errors detected by the routine:

IFAIL $= 1$

On entry, NX $\leq$ NK,
or NX ≤ 1,
or NK ≤ 0.

IFAIL $= 2$

On entry, all values of X are practically identical, giving zero variance. In this case R and STAT are undefined on exit.

7 Accuracy

The computations are believed to be stable.

8 Further Comments

The time taken by the routine is approximately proportional to NX $\times$ NK.

If the input series for G13ABF was generated by differencing using G13AAF, ensure that only the differenced values are input to G13ABF, and not the reconstituting information.

9 Example

In the example below, a set of 50 values of sunspot counts is used as input. The first 10 autocorrelations are computed.

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.

```
*     G13ABF Example Program Text
*     Mark 14 Revised.  NAG Copyright 1989.
*     .. Parameters ..
      INTEGER          NXMAX, NKMAX
      PARAMETER        (NXMAX=50,NKMAX=10)
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
*     .. Local Scalars ..
      real             STAT, XM, XV
      INTEGER          I, IFAIL, NK, NX
*     .. Local Arrays ..
      real             R(NKMAX), X(NXMAX)
*     .. External Subroutines ..
      EXTERNAL         G13ABF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G13ABF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) NX, NK
      WRITE (NOUT,*)
      IF (NK.GT.0 .AND. NK.LE.NKMAX .AND. NX.GT.0 .AND. NX.LE.NXMAX)
     +    THEN
         READ (NIN,*) (X(I),I=1,NX)
         WRITE (NOUT,99999) 'The first ', NK,
     +     ' coefficients are required'
         IFAIL = 0
*
         CALL G13ABF(X,NX,NK,XM,XV,R,STAT,IFAIL)
*
         WRITE (NOUT,99998) 'The input array has sample mean ', XM
         WRITE (NOUT,99998) 'The input array has sample variance ', XV
         WRITE (NOUT,*) 'The sample autocorrelation coefficients are'
         WRITE (NOUT,*)
         WRITE (NOUT,*) '   Lag    Coeff      Lag    Coeff'
         WRITE (NOUT,99997) (I,R(I),I=1,10)
         WRITE (NOUT,*)
         WRITE (NOUT,99998) 'The value of STAT is ', STAT
      END IF
      STOP
*
99999 FORMAT (1X,A,I2,A)
99998 FORMAT (1X,A,F12.4)
99997 FORMAT (1X,I6,F10.4,I8,F10.4)
      END
```

9.2 Program Data

```
G13ABF Example Program Data
 50 10
   5.0  11.0  16.0  23.0  36.0
  58.0  29.0  20.0  10.0   8.0
   3.0   0.0   0.0   2.0  11.0
  27.0  47.0  63.0  60.0  39.0
  28.0  26.0  22.0  11.0  21.0
  40.0  78.0 122.0 103.0  73.0
  47.0  35.0  11.0   5.0  16.0
  34.0  70.0  81.0 111.0 101.0
  73.0  40.0  20.0  16.0   5.0
  11.0  22.0  40.0  60.0  80.9
```

9.3 Program Results

```
G13ABF Example Program Results

The first 10 coefficients are required
The input array has sample mean      37.4180
The input array has sample variance    1002.0301
The sample autocorrelation coefficients are

   Lag    Coeff      Lag    Coeff
     1    0.8004       2    0.4355
     3    0.0328       4   -0.2835
     5   -0.4505       6   -0.4242
     7   -0.2419       8    0.0550
     9    0.3783      10    0.5857

The value of STAT is      92.1231
```