

NAG Library Routine Document

G13ABF

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 (KIND=nag_wp) X(NX), XM, XV, R(NK), STAT

3 Description

The data consists of n observations x_i , for $i = 1, 2, \dots, n$ 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 $k = 1, 2, \dots, K$, 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 of Box and Jenkins (1976) 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_K^2 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 Box and Jenkins (1976) provides further details of the use of STAT.

4 References

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

5 Parameters

- 1: X(NX) – REAL (KIND=nag_wp) array *Input*
On entry: the time series, x_i , for $i = 1, 2, \dots, n$.
- 2: NX – INTEGER *Input*
On entry: n , the number of values in the time series.
Constraint: $NX > 1$.
- 3: NK – INTEGER *Input*
On entry: K , the number of lags 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 (KIND=nag_wp) *Output*
On exit: the sample mean of the input time series.
- 5: XV – REAL (KIND=nag_wp) *Output*
On exit: the sample variance of the input time series.
- 6: R(NK) – REAL (KIND=nag_wp) array *Output*
On exit: the sample autocorrelation coefficient relating to lag k , for $k = 1, 2, \dots, K$.
- 7: STAT – REAL (KIND=nag_wp) *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. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.
 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, if you are 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.**
On exit: IFAIL = 0 unless the routine detects an error or a warning has been flagged (see Section 6).

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, $NX \leq NK$,
or $NX \leq 1$,
or $NK \leq 0$.

$IFAIL = 2$

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

If $n < 100$, or $K < 10 \log(n)$ then the autocorrelations are calculated directly and the time taken by G13ABF is approximately proportional to nK , otherwise the autocorrelations are calculated by utilizing fast fourier transforms (FFTs) and the time taken is approximately proportional to $n \log(n)$. If FFTs are used then G13ABF internally allocates approximately $4n$ real elements.

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

```

Program g13abfe

!      G13ABF Example Program Text

!      Mark 24 Release. NAG Copyright 2012.

!      .. Use Statements ..
Use nag_library, Only: g13abf, nag_wp
!      .. Implicit None Statement ..
Implicit None
!      .. Parameters ..
Integer, Parameter          :: nin = 5, nout = 6
!      .. Local Scalars ..
Real (Kind=nag_wp)         :: stat, xm, xv
Integer                     :: i, ifail, nk, nx
!      .. Local Arrays ..
Real (Kind=nag_wp), Allocatable :: r(:), x(:)
!      .. Executable Statements ..
Write (nout,*) 'G13ABF Example Program Results'
Write (nout,*)

!      Skip heading in data file
Read (nin,*)

!      Read in the problem size
Read (nin,*) nx, nk

```

```

        Allocate (x(nx),r(nk))

!       Read in data
        Read (nin,*) x(1:nx)

!       Compute autocorrelation
        ifail = 0
        Call g13abf(x,nx,nk,xm,xv,r,stat,ifail)

!       Display results
        Write (nout,99999) 'The first ', nk, ' coefficients are required'
        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,nk)
        Write (nout,*)
        Write (nout,99998) 'The value of STAT is ', stat

99999 Format (1X,A,I2,A)
99998 Format (1X,A,F12.4)
99997 Format (1X,I6,F10.4,I8,F10.4)
        End Program g13abfe

```

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

```
