

## NAG Toolbox

### nag\_tsa\_uni\_autocorr (g13ab)

## 1 Purpose

nag\_tsa\_uni\_autocorr (g13ab) 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 Syntax

```
[xm, xv, r, stat, ifail] = nag_tsa_uni_autocorr(x, nk, 'nx', nx)
[xm, xv, r, stat, ifail] = g13ab(x, nk, 'nx', nx)
```

## 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  $\chi_K^2$  distribution under the hypothesis of a zero autocorrelation function. Values of **stat** in the upper tail of the distribution provide evidence against the hypothesis; nag\_stat\_prob\_chisq (g01ec) 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

### 5.1 Compulsory Input Parameters

1: **x(nx)** – REAL (KIND=nag\_wp) array

The time series,  $x_i$ , for  $i = 1, 2, \dots, n$ .

2: **nk** – INTEGER

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

*Constraint:*  $0 < \mathbf{nk} < \mathbf{nx}$ .

### 5.2 Optional Input Parameters

1: **nx** – INTEGER

*Default:* the dimension of the array **x**.

$n$ , the number of values in the time series.

*Constraint:*  $\mathbf{nx} > 1$ .

### 5.3 Output Parameters

1: **xm** – REAL (KIND=nag\_wp)

The sample mean of the input time series.

2: **xv** – REAL (KIND=nag\_wp)

The sample variance of the input time series.

3: **r(nk)** – REAL (KIND=nag\_wp) array

The sample autocorrelation coefficient relating to lag  $k$ , for  $k = 1, 2, \dots, K$ .

4: **stat** – REAL (KIND=nag\_wp)

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

5: **ifail** – INTEGER

**ifail** = 0 unless the function detects an error (see Section 5).

## 6 Error Indicators and Warnings

Errors or warnings detected by the function:

**ifail** = 1

On entry,  $\mathbf{nx} \leq \mathbf{nk}$ ,  
 or  $\mathbf{nx} \leq 1$ ,  
 or  $\mathbf{nk} \leq 0$ .

**ifail = 2 (warning)**

On entry, all values of **x** are practically identical, giving zero variance. In this case **r** and **stat** are undefined on exit.

**ifail = -99**

An unexpected error has been triggered by this routine. Please contact NAG.

**ifail = -399**

Your licence key may have expired or may not have been installed correctly.

**ifail = -999**

Dynamic memory allocation failed.

## 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 nag\_tsa\_uni\_autocorr (g13ab) 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 nag\_tsa\_uni\_autocorr (g13ab) internally allocates approximately  $4n$  real elements.

If the input series for nag\_tsa\_uni\_autocorr (g13ab) was generated by differencing using nag\_tsa\_uni\_diff (g13aa), ensure that only the differenced values are input to nag\_tsa\_uni\_autocorr (g13ab), 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

```
function g13ab_example

fprintf('g13ab example results\n\n');

x = [ 5; 11; 16; 23; 36; 58; 29; 20; 10; 8;
      3; 0; 0; 2; 11; 27; 47; 63; 60; 39;
     28; 26; 22; 11; 21; 40; 78; 122; 103; 73;
     47; 35; 11; 5; 16; 34; 70; 81; 111; 101;
     73; 40; 20; 16; 5; 11; 22; 40; 60; 80.9];

nk = nag_int(10);
% Compute autocorrelation
[xm, xv, r, stat, ifail] = g13ab( ...
    x, nk);

% Display results
fprintf('The first %4d coefficients are required\n',nk);
fprintf('The input array has sample mean      %12.4f\n', xm);
fprintf('The input array has sample variance %12.4f\n', xv);
fprintf('The sample autocorrelation coefficients are\n\n');
fprintf('   Lag   Coeff   Lag   Coeff\n');
ivar = double([1:nk]);
fprintf('%6d%10.4f%8d%10.4f\n',[ivar; r(1:nk)]);
fprintf('\nThe value of stat is %12.4f\n', stat);
```

```
% For plotting use all possible lags
nk = nag_int(numel(x)-1);
[xm, xv, r, stat, ifail] = g13ab( ...
x, nk);

fig1 = figure;
refline = 1.96/sqrt(numel(x));
hold on
h = bar(r,0.1);
h.FaceColor = [1 0 0];
h.EdgeColor = [1 0 0];
h.ShowBaseLine = 'off';
plot([0 nk+1],[refline refline],'green');
plot([0 nk+1],[-refline -refline],'green');
axis([0 50 -0.6 1]);
ax = gca;
ax.YTick = [-0.6:0.2:1];
ax.XTick = [0:10:50];
xlabel('Lag');
ylabel('ACF');
title('Sample autocorrelation coefficients');
hold off
```

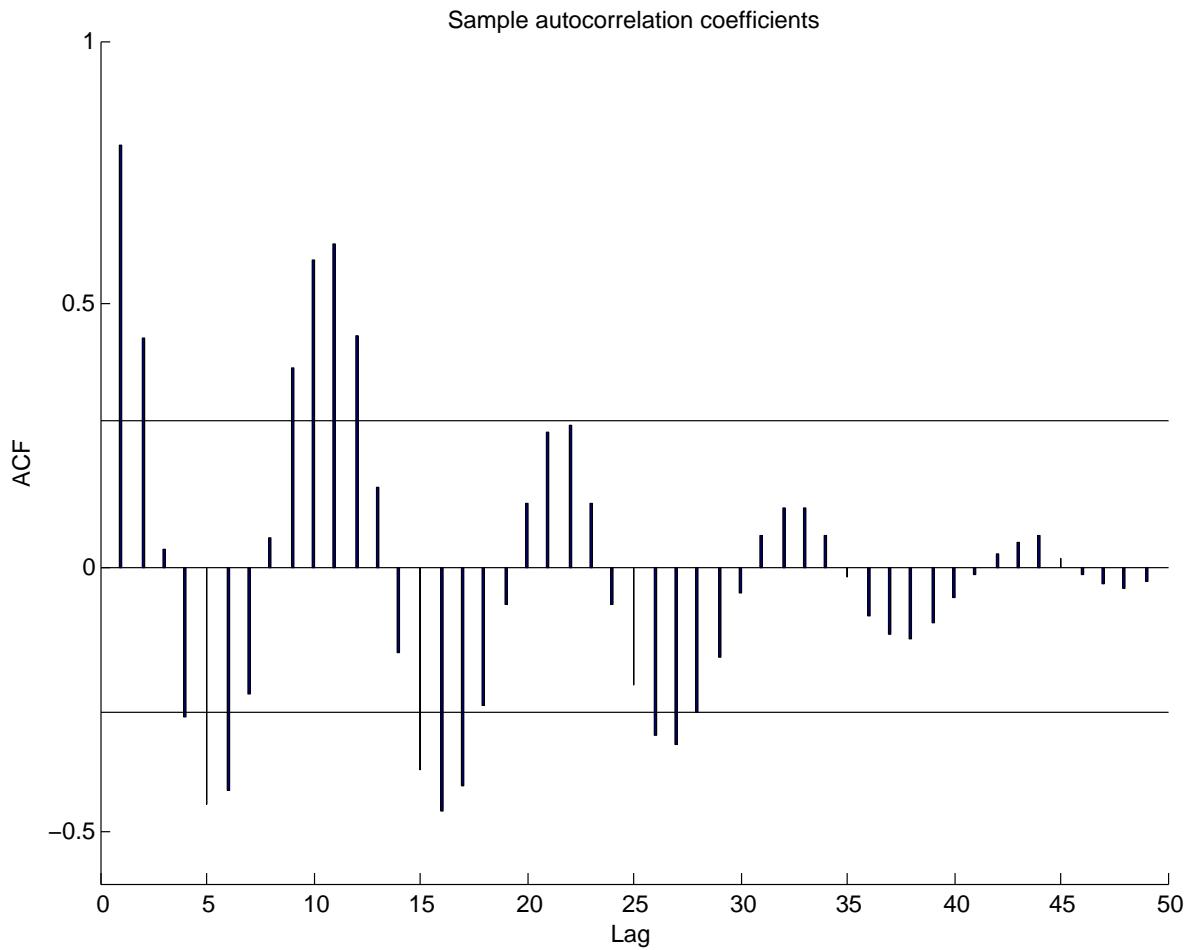
## 9.2 Program Results

g13ab example 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



This plot shows the autocorrelations for all possible lag values. Reference lines are given at  $\pm z_{0.975}/\sqrt{n}$ .

---