

# NAG Library Function Document

## nag\_tsa\_auto\_corr (g13abc)

### 1 Purpose

nag\_tsa\_auto\_corr (g13abc) 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

```
#include <nag.h>
#include <nagg13.h>
void nag_tsa_auto_corr (const double x[], Integer nx, Integer nk,
    double *mean, double *var, double r[], double *stat, NagError *fail)
```

### 3 Description

The data consist 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  $K < n$ ,  $n > 1$ .

- (d) The coefficient of lag  $k$  is defined as

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

- (e) See page 496 *et seq.* of Box and Jenkins (1976) for further details.

- (f) A test statistic defined as

$$\mathbf{stat} = n \sum_{k=1}^K r_k^2,$$

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  $\chi_{2K}^2$  distribution under the hypothesis of a zero autocorrelation function. Values of **stat** in the upper tail of the distribution provide evidence against the hypothesis.

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 Arguments

- |    |   |                     |
|----|---|---------------------|
| 1: | <b>x</b> [ <b>nx</b> ] – const double   | <i>Input</i>        |
|    | <i>On entry:</i> the time series, $x_i$ , for $i = 1, 2, \dots, n$ .  |                     |
| 2: | <b>nx</b> – Integer   | <i>Input</i>        |
|    | <i>On entry:</i> the number of values, $n$ , in the time series.  |                     |
|    | <i>Constraint:</i> <b>nx</b> > 1.   |                     |
| 3: | <b>nk</b> – Integer   | <i>Input</i>        |
|    | <i>On entry:</i> the number of lags, $K$ , for which the autocorrelations are required. The lags range from 1 to $K$ and do not include zero. |                     |
|    | <i>Constraint:</i> $0 < \mathbf{nk} < \mathbf{nx}$ .  |                     |
| 4: | <b>mean</b> – double *  | <i>Output</i>       |
|    | <i>On exit:</i> the sample mean of the input time series.   |                     |
| 5: | <b>var</b> – double *   | <i>Output</i>       |
|    | <i>On exit:</i> the sample variance of the input time series.   |                     |
| 6: | <b>r</b> [ <b>nk</b> ] – double   | <i>Output</i>       |
|    | <i>On exit:</i> the sample autocorrelation coefficient relating to lag $k$ , for $k = 1, 2, \dots, K$ .                                       |                     |
| 7: | <b>stat</b> – double *  | <i>Output</i>       |
|    | <i>On exit:</i> the statistic used to test the hypothesis that the true autocorrelation function of the time series is identically zero.      |                     |
| 8: | <b>fail</b> – NagError *  | <i>Input/Output</i> |
|    | The NAG error argument (see Section 3.6 in the Essential Introduction).   |                     |

## 6 Error Indicators and Warnings

### NE\_2\_INT\_ARG\_LE

On entry, **nx** =  $\langle value \rangle$  while **nk** =  $\langle value \rangle$ . These arguments must satisfy **nx** > **nk**.

### NE\_INT\_ARG\_LE

On entry, **nk** =  $\langle value \rangle$ .

Constraint: **nk** > 0.

On entry, **nx** =  $\langle value \rangle$ .

Constraint: **nx** > 1.

### NE\_TIME\_SERIES\_IDEN

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 Parallelism and Performance

Not applicable.

## 9 Further Comments

The time taken by `nag_tsa_auto_corr` (g13abc) is approximately proportional to  $\mathbf{nx} \times \mathbf{nk}$ .

## 10 Example

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

### 10.1 Program Text

```

/* nag_tsa_auto_corr (g13abc) Example Program.
 *
 * Copyright 1991 Numerical Algorithms Group.
 *
 * Mark 2, 1991.
 * Mark 8 revised, 2004.
 *
 */

#include <nag.h>
#include <stdio.h>
#include <nag_stdlib.h>
#include <naggl3.h>

int main(void)
{
    Integer    exit_status = 0, i, nk, nx;
    NagError  fail;
    double     mean, *r = 0, stat, *x = 0, xv;

    INIT_FAIL(fail);

    printf("nag_tsa_auto_corr (g13abc) Example Program Results\n");
    /* Skip heading in data file */
    scanf("%*[\n]");
    scanf("%ld %ld", &nx, &nk);

    if (nk > 0 && nx > 1 && nk < nx)
    {
        if (!(r = NAG_ALLOC(nk, double)) ||
            !(x = NAG_ALLOC(nx, double)))
        {
            printf("Allocation failure\n");
            exit_status = -1;
            goto END;
        }
    }
    else
    {
        printf("Invalid nx or nk.\n");
        exit_status = 1;
        return exit_status;
    }
    for (i = 0; i < nx; ++i)
        scanf("%lf", &x[i]);
    printf("\nThe first %2ld coefficients are required\n", nk);

```

```

/* nag_tsa_auto_corr (g13abc).
 * Sample autocorrelation function
 */
nag_tsa_auto_corr(x, nx, nk, &mean, &xv, r, &stat, &fail);
if (fail.code != NE_NOERROR)
{
    printf("Error from nag_tsa_auto_corr (g13abc).\n%s\n",
           fail.message);
    exit_status = 1;
    goto END;
}

printf("The input array has sample mean %12.4f\n", mean);
printf("The input array has sample variance %12.4f\n", xv);
printf("The sample autocorrelation coefficients are\n\n");
printf("  Lag    Coeff\n");
for (i = 0; i < 10; ++i)
    printf("%6ld%10.4f\n", i+1, r[i]);
printf("\nThe value of stat is %12.4f\n", stat);
END:
NAG_FREE(r);
NAG_FREE(x);
return exit_status;
}

```

## 10.2 Program Data

nag\_tsa\_auto\_corr (g13abc) 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

```

## 10.3 Program Results

nag\_tsa\_auto\_corr (g13abc) 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
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

---