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(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 $i = 1, 2, \ldots, n$ 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 (1976) for further details.

(d) A test statistic defined as

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

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.

Section 8.2.2 of Box and Jenkins (1976) provides further details of the use of **stat**.

4. Parameters

x[nx]
Input: the time series, x_i, for $i = 1, 2, \ldots, n$.

nx
Input: the number of values, n, in the time series.
Constraint: **nx** > 1.

nk
Input: 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**.

mean
Output: the sample mean of the input time series.

var
Output: the sample variance of the input time series.

r[nk]
Output: the sample autocorrelation coefficient relating to lag k, for $k = 1, 2, \ldots, K$.

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

fail
The NAG error parameter, see the Essential Introduction to the NAG C Library.

5. Error Indications and Warnings

NE_INT_ARG_LE
On entry, **nx** must not be less than or equal to 1: **nx** = ⟨*value*⟩.
On entry, **nk** must not be less than or equal to 0: **nk** = ⟨*value*⟩.

NE_2_INT_ARG_LE
On entry **nx** = ⟨*value*⟩ while **nk** = ⟨*value*⟩. These parameters must satisfy **nx** $>$ **nk**.

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.

6. Further Comments

The time taken by the routine is approximately proportional to **nx** $\times$ **nk**.

6.1. Accuracy

The computations are believed to be stable.

6.2. References

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

7. See Also

nag_tsa_auto_corr_part (g13acc)

8. Example

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

8.1. Program Text

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

#include <nag.h>
#include <stdio.h>
#include <nag_stdlib.h>
#include <nagg13.h>

#define NXMAX 50
#define NKMAX 10

main()
{

  double mean, xv, stat;
  Integer i, nk, nx;
  double r[NKMAX], x[NXMAX];

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

  if (nk>0 && nk<=NKMAX && nx>0 && nx<=NXMAX)
    {
      for (i=0; i<nx; ++i)
        Vscanf("%lf", &x[i]);
      Vprintf("\nThe first %2ld coefficients are required\n", nk);

      g13abc(x, nx, nk, &mean, &xv, r, &stat, NAGERR_DEFAULT);

      Vprintf("The input array has sample mean %12.4f\n", mean);
      Vprintf("The input array has sample variance %12.4f\n", xv);
      Vprintf("The sample autocorrelation coefficients are\n\n");
      Vprintf("   Lag    Coeff\n");
      for (i=0; i<10; ++i)
        Vprintf("%6ld%10.4f\n", i+1, r[i]);
      Vprintf("\nThe value of stat is %12.4f\n", stat);
    }
  else
    {
      Vfprintf(stderr, "One or both of nk and nx are out of range:\
 nk = %-3ld while  nx = %-3ld\n", nk, nx);
      exit(EXIT_FAILURE);
    }
  exit(EXIT_SUCCESS);
}
```

8.2. Program Data

```
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
```

8.3. Program Results

```
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
```