NAG C Library Function Document

nag_durbin_watson_stat (g02fcc)

1 Purpose
nag_durbin_watson_stat (g02fcc) calculates the Durbin–Watson statistic, for a set of residuals, and the upper and lower bounds for its significance.

2 Specification

```c
void nag_durbin_watson_stat (Integer n, Integer p, const double *res[], double *d, double *pdl, double *pdu, NagError *fail)
```

3 Description

For the general linear regression model

\[ y = X\beta + \epsilon, \]

where \( y \) is a vector of length \( n \) of the dependent variable,

\( X \) is a \( n \) by \( p \) matrix of the independent variables,

\( \beta \) is a vector of length \( p \) of unknown parameters,

and \( \epsilon \) is a vector of length \( n \) of unknown random errors.

The residuals are given by

\[ r = y - \hat{y} = y - X\hat{\beta} \]

and the fitted values, \( \hat{y} = X\hat{\beta} \), can be written as \( Hy \) for a \( n \) by \( n \) matrix \( H \). Note that when a mean term is included in the model the sum of the residuals is zero. If the observations have been taken serially, that is \( y_1, y_2, \ldots, y_n \) can be considered as a time series, the Durbin–Watson test can be used to test for serial correlation in the \( \epsilon_i \), see Durbin and Watson (1950), Durbin and Watson (1951) and Durbin and Watson (1971).

The Durbin–Watson statistic is

\[ d = \frac{\sum_{i=1}^{n-1} (r_{i+1} - r_i)^2}{\sum_{i=1}^{n} r_i^2}. \]

Positive serial correlation in the \( \epsilon_i \) will lead to a small value of \( d \) while for independent errors \( d \) will be close to 2. Durbin and Watson show that the exact distribution of \( d \) depends on the eigenvalues of the matrix \( HA \) where the matrix \( A \) is such that \( d \) can be written as

\[ d = \frac{r^T A r}{r^T r}, \]

and the eigenvalues of the matrix \( A \) are \( \lambda_j = (1 - \cos(\pi j/n)) \), for \( j = 1, 2, \ldots, n - 1 \).

However bounds on the distribution can be obtained, the lower bound being

\[ d_l = \frac{\sum_{i=1}^{n-p} \lambda_i u_i^2}{\sum_{i=1}^{n} u_i^2}, \]

and the upper bound being

\[ d_u = \frac{\sum_{i=1}^{n-p} \lambda_{i-1+p} u_i^2}{\sum_{i=1}^{n} u_i^2}, \]

where the \( u_i \) are independent standard Normal variables. The lower tail probabilities associated with these bounds, \( p_l \) and \( p_u \), are computed by nag_prob_durbin_watson (g01epc). The interpretation of the bounds
is that, for a test of size (significance) $\alpha$, if $p_l \leq \alpha$ the test is significant, if $p_u > \alpha$ the test is not significant, while if $p_l > \alpha$ and $p_u \leq \alpha$ no conclusion can be reached.

The above probabilities are for the usual test of positive auto-correlation. If the alternative of negative auto-correlation is required, then a call to nag_prob_durbin_watson (g01epc) should be made with the parameter $d$ taking the value of $4 - d$; see Newbold (1988).

4 References

5 Parameters

1: n – Integer $\quad$ Input

On entry: the number of residuals, $n$.

Constraint: $n > p$.

2: p – Integer $\quad$ Input

On entry: the number, $p$, of independent variables in the regression model, including the mean.

Constraint: $p \geq 1$.

3: res[n] – const double $\quad$ Input

On entry: the residuals, $r_1, r_2, \ldots, r_n$.

Constraint: the mean of the residuals $\leq \sqrt{\epsilon}$, where $\epsilon = \text{machine precision}$.

4: d – double $\quad$ Output

On exit: the Durbin–Watson statistic, $d$.

5: pdl – double $\quad$ Output

On exit: lower bound for the significance of the Durbin–Watson statistic, $p_l$.

6: pdu – double $\quad$ Output


7: fail – NagError $\quad$ Input/Output

The NAG error parameter (see the Essential Introduction).

6 Error Indicators and Warnings

NE_INT

On entry, $p = (\text{value})$.

Constraint: $p \geq 1$. 
NE_INT_2
    On entry, \( n = \langle \text{value} \rangle \), \( p = \langle \text{value} \rangle \).
    Constraint: \( n > p \).

NE_RESID_IDEN
    On entry, all residuals are identical.

NE_RESID_MEAN
    On entry, The mean of \( \text{res} \) is not approximately 0.0, mean = \( \langle \text{value} \rangle \).

NE_ALLOC_FAIL
    Memory allocation failed.

NE_BAD_PARAM
    On entry, parameter \( \langle \text{value} \rangle \) had an illegal value.

NE_INTERNAL_ERROR
    An internal error has occurred in this function. Check the function call and any array sizes. If the
call is correct then please consult NAG for assistance.

7 Accuracy
The probabilities are computed to an accuracy of at least 4 decimal places.

8 Further Comments
If the exact probabilities are required, then the first \( n - p \) eigenvalues of \( HA \) can be computed and
\nag_prob_lin_chi_sq (g01jdc) used to compute the required probabilities with the parameter \( c \) set to 0.0
and the parameter \( d \) set to the Durbin–Watson statistic \( d \).

9 Example
A set of 10 residuals are read in and the Durbin–Watson statistic along with the probability bounds are
computed and printed.

9.1 Program Text
/* nag_durbin_watson_stat (g02fcc) Example Program. *
 * Copyright 2002 Numerical Algorithms Group. *
 * Mark 7, 2002. */
#include <stdio.h>
#include <nag.h>
#include <nag_stdlib.h>
#include <nag02.h>

int main(void)
{
    /* Scalars */
    double d, pdl, pdu;
    Integer exit_status, i, p, n;
    NagError fail;

    /* Arrays */
    double *res=0;

    /* Scalars */
    double d, pdl, pdu;
    Integer exit_status, i, p, n;
    NagError fail;

    /* Arrays */
    double *res=0;

INIT_FAIL(fail);
exit_status = 0;
Vprintf("g02fcc Example Program Results\n");

/* Skip heading in data file */
Vscanf("%*[\n]");
Vscanf("%ld%*[\n] ", &p);
n = 10;

/* Allocate memory */
if ( !(res = NAG_ALLOC(n, double)) )
{
  Vprintf("Allocation failure\n");
  exit_status = -1;
  goto END;
}
for (i = 1; i <= n; ++i)
  Vscanf("%lf", &res[i - 1]);
Vscanf("%*[\n]");
g02fcc(n, p, res, &d, &pdl, &pdu, &fail);
if (fail.code != NE_NOERROR)
{
  Vprintf("Error from g02fcc.\n%s\n", fail.message);
  exit_status = 1;
  goto END;
}
Vprintf("\n");
Vprintf(" Durbin-Watson statistic %10.4f\n", d);
Vprintf(" Lower and upper bound %10.4f%10.4f\n", pdl, pdu);
END:
  if (res) NAG_FREE(res);
  return exit_status;
}

9.2 Program Data

g02fcc Example Program Data
2
3.735719 0.912755 0.683626 0.416693 1.9902
-0.444816 -1.283088 -3.666035 -0.426357 -1.918697

9.3 Program Results

g02fcc Example Program Results

  Durbin-Watson statistic   0.9238

  Lower and upper bound    0.0610    0.0060