NAG Toolbox

nag_stat_prob_durbin_watson (g01ep)

1 Purpose

nag_stat_prob_durbin_watson (g01ep) calculates upper and lower bounds for the significance of a Durbin–Watson statistic.

2 Syntax

```
[pdl, pdu, ifail] = nag_stat_prob_durbin_watson(n, ip, d)
[pdl, pdu, ifail] = g01ep(n, ip, d)
```

3 Description

Let $r = (r_1, r_2, \ldots, r_n)^{\mathrm{T}}$ be the residuals from a linear regression of y on p independent variables, including the mean, where the y values $y_1, y_2, \ldots, y_n$ can be considered as a time series. The Durbin–Watson test (see Durbin and Watson (1950), Durbin and Watson (1951) and Durbin and Watson (1971)) can be used to test for serial correlation in the error term in the regression.

The Durbin–Watson test statistic is:

$$d = \frac{\sum_{i=1}^{n-1} (r_{i+1} - r_i)^2}{\sum_{i=1}^{n} r_i^2},$$

which can be written as

$$d = \frac{r^{\mathrm{T}} A r}{r^{\mathrm{T}} r},$$

where the n by n matrix A is given by

$$A = \begin{bmatrix} 1 & -1 & 0 & \ldots & : \\ -1 & 2 & -1 & \ldots & : \\ 0 & -1 & 2 & \ldots & : \\ : & 0 & -1 & \ldots & : \\ : & : & : & \ldots & : \\ : & : & : & \ldots & -1 \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

with the nonzero eigenvalues of the matrix A being $\lambda_j = (1 - \cos(\pi j / n))$, for $j = 1, 2, \ldots, n-1$.

Durbin and Watson show that the exact distribution of d depends on the eigenvalues of a matrix HA, where H is the hat matrix of independent variables, i.e., the matrix such that the vector of fitted values, $\hat{y}$, can be written as $\hat{y} = Hy$. However, bounds on the distribution can be obtained, the lower bound being

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

and the upper bound being

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

where u_i are independent standard Normal variables.

Two algorithms are used to compute the lower tail (significance level) probabilities, p_l and p_u, associated with d_l and d_u. If $n \leq 60$ the procedure due to Pan (1964) is used, see Farebrother (1980), otherwise Imhof's method (see Imhof (1961)) is used.

The bounds are for the usual test of positive correlation; if a test of negative correlation is required the value of d should be replaced by $4 - d$.

4 References

Durbin J and Watson G S (1950) Testing for serial correlation in least squares regression. I *Biometrika* **37** 409–428

Durbin J and Watson G S (1951) Testing for serial correlation in least squares regression. II *Biometrika* **38** 159–178

Durbin J and Watson G S (1971) Testing for serial correlation in least squares regression. III *Biometrika* **58** 1–19

Farebrother R W (1980) Algorithm AS 153. Pan's procedure for the tail probabilities of the Durbin–Watson statistic *Appl. Statist.* **29** 224–227

Imhof J P (1961) Computing the distribution of quadratic forms in Normal variables *Biometrika* **48** 419–426

Newbold P (1988) *Statistics for Business and Economics* Prentice–Hall

Pan Jie–Jian (1964) Distributions of the noncircular serial correlation coefficients *Shuxue Jinzhan* **7** 328–337

5 Parameters

5.1 Compulsory Input Parameters

1: **n** – INTEGER

n, the number of observations used in calculating the Durbin–Watson statistic.

Constraint: **n** $>$ **ip**.

2: **ip** – INTEGER

p, the number of independent variables in the regression model, including the mean.

Constraint: **ip** ≥ 1.

3: **d** – REAL (KIND=nag_wp)

d, the Durbin–Watson statistic.

Constraint: **d** ≥ 0.0.

5.2 Optional Input Parameters

None.

5.3 Output Parameters

1: **pdl** – REAL (KIND=nag_wp)

Lower bound for the significance of the Durbin–Watson statistic, p_l.

2: **pdu** – REAL (KIND=nag_wp)

Upper bound for the significance of the Durbin–Watson statistic, p_u.

3: **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, **n** $\leq$ **ip**,
or **ip** < 1.

ifail $= 2$

On entry, **d** < 0.0.

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

On successful exit at least 4 decimal places of accuracy are achieved.

8 Further Comments

If the exact probabilities are required, then the first $n - p$ eigenvalues of HA can be computed and nag_stat_prob_chisq_lincomb (g01jd) used to compute the required probabilities with **c** set to 0.0 and **d** to the Durbin–Watson statistic.

9 Example

The values of n, p and the Durbin–Watson statistic d are input and the bounds for the significance level calculated and printed.

9.1 Program Text

```
     function g01ep_example

fprintf('g01ep example results\n\n');

% bounds for the significance of a Durbin--Watson statistic.
n  = nag_int(10);
ip = nag_int(2);
```

```
d  = 0.9238;

[pdl, pdu, ifail] = g01ep(n, ip, d);

fprintf('d_l(n=%4d,p=%4d,d=%7.4f) = %7.4f\n',n,ip,d,pdl);
fprintf('d_u(n=%4d,p=%4d,d=%7.4f) = %7.4f\n',n,ip,d,pdu);
```

9.2 Program Results

```
     g01ep example results

d_l(n=  10,p=   2,d= 0.9238) =  0.0610
d_u(n=  10,p=   2,d= 0.9238) =  0.0060
```