NAG Toolbox for MATLAB

g01ep

1 Purpose

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

2 Syntax

```
[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 – int32 scalar**

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

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

2: **ip – int32 scalar**

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

Constraint: **ip** ≥ 1.

3: **d – double scalar**

d, the Durbin–Watson statistic.

Constraint: **d** ≥ 0.0.

5.2 Optional Input Parameters

None.

5.3 Input Parameters Omitted from the MATLAB Interface

work

5.4 Output Parameters

1: **pdl – double scalar**

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

2: **pdu – double scalar**

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

3: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 1$

On entry, $\mathbf{n} \leq \mathbf{ip}$,
or $\mathbf{ip} < 1$.

ifail $= 2$

On entry, $\mathbf{d} < 0.0$.

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 g01jd used to compute the required probabilities with **c** set to 0.0 and **d** to the Durbin–Watson statistic.

9 Example

```
n = int32(10);
ip = int32(2);
d = 0.9238;
[pdl, pdu, ifail] = g01ep(n, ip, d)
```

```
pdl =
    0.0610
pdu =
    0.0060
ifail =
           0
```