NAG Toolbox for MATLAB

g01dc

1 Purpose

g01dc computes an approximation to the variance-covariance matrix of an ordered set of independent observations from a Normal distribution with mean 0.0 and standard deviation 1.0.

2 Syntax

```
[vec, ifail] = g01dc(n, exp1, exp2, sumssq)
```

3 Description

g01dc is an adaptation of the Applied Statistics Algorithm AS 128, see Davis and Stephens 1978. An approximation to the variance-covariance matrix, V, using a Taylor series expansion of the Normal distribution function is discussed in David and Johnson 1954.

However, convergence is slow for extreme variances and covariances. The present function uses the David–Johnson approximation to provide an initial approximation and improves upon it by use of the following identities for the matrix.

For a sample of size n, let m_i be the expected value of the ith largest order statistic, then:

(a) for any $i = 1, 2, \ldots, n$, $\sum_{j=1}^{n} V_{ij} = 1$

(b) $V_{12} = V_{11} + m_n^2 - m_n m_{n-1} - 1$

(c) the trace of V is $tr(V) = n - \sum_{i=1}^{n} m_i^2$

(d) $V_{ij} = V_{ji} = V_{rs} = V_{sr}$ where $r = n + 1 - i$, $s = n + 1 - j$ and $i, j = 1, 2, \ldots, n$. Note that only the upper triangle of the matrix is calculated and returned column-wise in vector form.

4 References

David F N and Johnson N L 1954 Statistical treatment of censored data, Part 1. Fundamental formulae *Biometrika* **41** 228–240

Davis C S and Stephens M A 1978 Algorithm AS 128: Approximating the covariance matrix of Normal order statistics *Appl. Statist.* **27** 206–212

5 Parameters

5.1 Compulsory Input Parameters

1: **n – int32 scalar**

n, the sample size.

Constraint: **n** ≥ 1.

2: **exp1 – double scalar**

The expected value of the largest Normal order statistic, m_n, from a sample of size n.

3: **exp2 – double scalar**

The expected value of the second largest Normal order statistic, m_{n-1}, from a sample of size n.

4: **sumssq – double scalar**

The sum of squares of the expected values of the Normal order statistics from a sample of size n.

5.2 Optional Input Parameters

None.

5.3 Input Parameters Omitted from the MATLAB Interface

None.

5.4 Output Parameters

1: **vec**$(\mathbf{n} \times (\mathbf{n}+\mathbf{1})/\mathbf{2})$ **– double array**

The upper triangle of the n by n variance-covariance matrix packed by column. Thus element V_{ij} is stored in **vec**$(i+j \times (j-1)/2)$, for $1 \leq i \leq j \leq n$.

2: **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, $\mathbf{n} < 1$.

7 Accuracy

For $n \leq 20$, where comparison with the exact values can be made, the maximum error is less than 0.0001.

8 Further Comments

The time taken by g01dc is approximately proportional to n^2.

The arguments **exp1** $(= m_n)$, **exp2** $(= m_{n-1})$ and **sumssq** $(= \sum_{j=1}^{n} m_j^2)$ may be found from the expected values of the Normal order statistics obtained from g01da (exact) or g01db (approximate).

9 Example

```
n = int32(6);
exp1 = 1.267206361712849;
exp2 = 0.641755038918563;
sumssq = 4.116565238504731;
[vec, ifail] = g01dc(n, exp1, exp2, sumssq)
```

```
vec =
    0.4159
    0.2085
    0.2796
    0.1394
    0.1889
    0.2462
    0.1025
    0.1397
```

```
    0.1834
    0.2462
    0.0774
    0.1060
    0.1397
    0.1889
    0.2796
    0.0563
    0.0774
    0.1025
    0.1394
    0.2085
    0.4159
ifail =
           0
```