

NAG Toolbox for MATLAB

g01db

1 Purpose

g01db calculates an approximation to the set of Normal Scores, i.e., the expected values of an ordered set of independent observations from a Normal distribution with mean 0.0 and standard deviation 1.0.

2 Syntax

```
[pp, ifail] = g01db(n)
```

3 Description

g01db is an adaptation of the Applied Statistics Algorithm AS 177.3, see Royston 1982. If you are particularly concerned with the accuracy with which g01db computes the expected values of the order statistics (see Section 7), then g01da which is more accurate should be used instead at a cost of increased storage and computing time.

Let $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ be the order statistics from a random sample of size n from the standard Normal distribution. Defining

$$P_{r,n} = \Phi(-E(x_{(r)}))$$

and

$$Q_{r,n} = \frac{r - \epsilon}{n + \gamma}, \quad r = 1, 2, \dots, n,$$

where $E(x_{(r)})$ is the expected value of $x_{(r)}$, the current function approximates the Normal upper tail area corresponding to $E(x_{(r)})$ as,

$$\tilde{P}_{r,n} = Q_{r,n} + \frac{\delta_1}{n} Q_{r,n}^\lambda + \frac{\delta_2}{n} Q_{r,n}^{2\lambda} - C_{r,n}.$$

Estimates of ϵ , γ , δ_1 , δ_2 and λ are obtained for $r = 1, 2, 3$, and $r \geq 4$. A small correction $C_{r,n}$ to $\tilde{P}_{r,n}$ is necessary when $r \leq 7$ and $n \leq 20$.

The approximation to $E(X_{(r)})$ is thus given by

$$E(x_{(r)}) = -\Phi^{-1}(\tilde{P}_{r,n}), \quad r = 1, 2, \dots, n.$$

Values of the inverse Normal probability integral Φ^{-1} are obtained from g01fa.

4 References

Royston J P 1982 Algorithm AS 177: Expected normal order statistics (exact and approximate) *Appl. Statist.* **31** 161–165

5 Parameters

5.1 Compulsory Input Parameters

1: **n** – int32 scalar

n , the size of the sample.

Constraint: $n \geq 1$.

5.2 Optional Input Parameters

None.

5.3 Input Parameters Omitted from the MATLAB Interface

None.

5.4 Output Parameters

1: **pp(n)** – double array

The Normal scores. **pp**(i) contains the value $E(x_{(i)})$, for $i = 1, 2, \dots, n$.

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, **n** < 1.

7 Accuracy

For $n \leq 2000$, the maximum error is 0.0001, but g01db is usually accurate to 5 or 6 decimal places. For n up to 5000, comparison with the exact scores calculated by g01da shows that the maximum error is 0.001.

8 Further Comments

The time taken by g01db is proportional to n .

9 Example

```
n = int32(10);  
[pp, ifail] = g01db(n)
```

```
pp =  
-1.5388  
-1.0014  
-0.6561  
-0.3757  
-0.1227  
0.1227  
0.3757  
0.6561  
1.0014  
1.5388  
ifail =  
0
```