NAG Library Routine Document

G01DBF

Note: before using this routine, please read the Users' Note for your implementation to check the interpretation of ***bold italicised*** terms and other implementation-dependent details.

1 Purpose

G01DBF 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 Specification

```
SUBROUTINE G01DBF (N, PP, IFAIL)

INTEGER            N, IFAIL
REAL (KIND=nag_wp) PP(N)
```

3 Description

G01DBF is an adaptation of the Applied Statistics Algorithm AS 177.3, see Royston (1982). If you are particularly concerned with the accuracy with which G01DBF computes the expected values of the order statistics (see Section 7), then G01DAF which is more accurate should be used instead at a cost of increased storage and computing time.

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

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

and

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

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

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

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

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

$$E\left(x_{(r)}\right) = -\Phi^{-1}\left(\tilde{P}_{r,n}\right), \qquad r = 1, 2, \ldots, n.$$

Values of the inverse Normal probability integral Φ^{-1} are obtained from G01FAF.

4 References

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

5 Parameters

1: N – INTEGER *Input*

On entry: n, the size of the sample.

Constraint: $\mathrm{N} \geq 1$.

2: PP(N) – REAL (KIND=nag_wp) array *Output*

On exit: the Normal scores. $\mathrm{PP}(i)$ contains the value $E\big(x_{(i)}\big)$, for $i = 1, 2, \ldots, n$.

3: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. If you are unfamiliar with this parameter you should refer to Section 3.3 in the Essential Introduction for details.

For environments where it might be inappropriate to halt program execution when an error is detected, the value -1 or 1 is recommended. If the output of error messages is undesirable, then the value 1 is recommended. Otherwise, if you are not familiar with this parameter, the recommended value is 0. **When the value -1 or 1 is used it is essential to test the value of IFAIL on exit.**

On exit: $\mathrm{IFAIL} = 0$ unless the routine detects an error or a warning has been flagged (see Section 6).

6 Error Indicators and Warnings

If on entry $\mathrm{IFAIL} = 0$ or -1, explanatory error messages are output on the current error message unit (as defined by X04AAF).

Errors or warnings detected by the routine:

$\mathrm{IFAIL} = 1$

On entry, $\mathrm{N} < 1$.

7 Accuracy

For $n \leq 2000$, the maximum error is 0.0001, but G01DBF is usually accurate to 5 or 6 decimal places. For n up to 5000, comparison with the exact scores calculated by G01DAF shows that the maximum error is 0.001.

8 Further Comments

The time taken by G01DBF is proportional to n.

9 Example

A program to calculate the expected values of the order statistics for a sample of size 10.

9.1 Program Text

```
    Program g01dbfe

!     G01DBF Example Program Text

!     Mark 24 Release. NAG Copyright 2012.

!     .. Use Statements ..
      Use nag_library, Only: g01dbf, nag_wp
!     .. Implicit None Statement ..
      Implicit None
!     .. Parameters ..
```

```
      Integer, Parameter               :: nout = 6
!     .. Local Scalars ..
      Integer                          :: ifail, n
!     .. Local Arrays ..
      Real (Kind=nag_wp), Allocatable  :: pp(:)
!     .. Executable Statements ..
      Write (nout,*) 'G01DBF Example Program Results'
      Write (nout,*)

!     Set the problem size
      n = 10

      Allocate (pp(n))

!     Calculate the normal scores
      ifail = 0
      Call g01dbf(n,pp,ifail)

!     Display results
      Write (nout,99999) 'Sample size = ', n
      Write (nout,*) 'Normal scores'
      Write (nout,99998) pp(1:n)

99999 Format (1X,A,I2)
99998 Format (10X,5F12.4)
    End Program g01dbfe
```

9.2 Program Data

```
G01DBF Example Program Data
```

9.3 Program Results

```
 G01DBF Example Program Results

 Sample size = 10
 Normal scores
              -1.5388     -1.0014     -0.6561     -0.3757     -0.1227
               0.1227      0.3757      0.6561      1.0014      1.5388
```