NAG Toolbox

nag_stat_prob_kolmogorov1 (g01ey)

1 Purpose

nag_stat_prob_kolmogorov1 (g01ey) returns the upper tail probability associated with the one sample Kolmogorov–Smirnov distribution.

2 Syntax

```
[result, ifail] = nag_stat_prob_kolmogorov1(n, d)
[result, ifail] = g01ey(n, d)
```

3 Description

Let $S_n(x)$ be the sample cumulative distribution function and $F_0(x)$ the hypothesised theoretical distribution function.

nag_stat_prob_kolmogorov1 (g01ey) returns the upper tail probability, p, associated with the one-sided Kolmogorov–Smirnov test statistic D_n^+ or D_n^-, where these one-sided statistics are defined as follows;

$$D_n^+ = \sup_x [S_n(x) - F_0(x)],$$

$$D_n^- = \sup_x [F_0(x) - S_n(x)[.$$

If $n \leq 100$ an exact method is used; for the details see Conover (1980). Otherwise a large sample approximation derived by Smirnov is used; see Feller (1948), Kendall and Stuart (1973) or Smirnov (1948).

4 References

Conover W J (1980) *Practical Nonparametric Statistics* Wiley

Feller W (1948) On the Kolmogorov–Smirnov limit theorems for empirical distributions *Ann. Math. Statist.* **19** 179–181

Kendall M G and Stuart A (1973) *The Advanced Theory of Statistics (Volume 2)* (3rd Edition) Griffin

Siegel S (1956) *Non-parametric Statistics for the Behavioral Sciences* McGraw–Hill

Smirnov N (1948) Table for estimating the goodness of fit of empirical distributions *Ann. Math. Statist.* **19** 279–281

5 Parameters

5.1 Compulsory Input Parameters

1: **n** – INTEGER

n, the number of observations in the sample.

Constraint: **n** ≥ 1.

2: **d** – REAL (KIND=nag_wp)

Contains the test statistic, D_n^+ or D_n^-.

Constraint: $0.0 \leq$ **d** ≤ 1.0.

5.2 Optional Input Parameters

None.

5.3 Output Parameters

1: **result**

The result of the function.

2: **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** < 1.

ifail $= 2$

On entry, **d** < 0.0,
or **d** > 1.0.

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

The large sample distribution used as an approximation to the exact distribution should have a relative error of less than 2.5% for most cases.

8 Further Comments

The upper tail probability for the two-sided statistic, $D_n = \max\left(D_n^+, D_n^-\right)$, can be approximated by twice the probability returned via nag_stat_prob_kolmogorov1 (g01ey), that is $2p$. (Note that if the probability from nag_stat_prob_kolmogorov1 (g01ey) is greater than 0.5 then the two-sided probability should be truncated to 1.0). This approximation to the tail probability for D_n is good for small probabilities, (e.g., $p \leq 0.10$) but becomes very poor for larger probabilities.

The time taken by the function increases with n, until $n > 100$. At this point the approximation is used and the time decreases significantly. The time then increases again modestly with n.

9 Example

The following example reads in 10 different sample sizes and values for the test statistic D_n. The upper tail probability is computed and printed for each case.

9.1 Program Text

```
     function g01ey_example

fprintf('g01ey example results\n\n');

% Upper tail probabilities for 1-sample Kolmogorov--Smirnov distribution.
n = zeros(10,1,nag_int_name);
n(1: 5) = 10;
n(6:10) = 400;
d = [0.323; 0.369; 0.409; 0.457; 0.489; 0.0535; 0.061; 0.068; 0.076; 0.0815];

fprintf('     d       n    one-sided probability\n');
for j = 1:numel(d)

  [p, ifail] = g01ey( ...
                       n(j), d(j));

  fprintf('%8.4f%6d%17.4f\n', d(j), n(j), p);
end
```

9.2 Program Results

```
     g01ey example results

     d       n    one-sided probability
  0.3230    10           0.0994
  0.3690    10           0.0497
  0.4090    10           0.0251
  0.4570    10           0.0099
  0.4890    10           0.0050
  0.0535   400           0.1001
  0.0610   400           0.0502
  0.0680   400           0.0243
  0.0760   400           0.0096
  0.0815   400           0.0048
```