NAG Toolbox

nag_surviv_kaplanmeier (g12aa)

1 Purpose

nag_surviv_kaplanmeier (g12aa) computes the Kaplan–Meier, (or product-limit), estimates of survival probabilities for a sample of failure times.

2 Syntax

```
[nd, tp, p, psig, ifail] = nag_surviv_kaplanmeier(t, ic, freq, ifreq, 'n', n)
[nd, tp, p, psig, ifail] = g12aa(t, ic, freq, ifreq, 'n', n)
```

3 Description

A survivor function, $S(t)$, is the probability of surviving to at least time t with $S(t) = 1 - F(t)$, where $F(t)$ is the cumulative distribution function of the failure times. The Kaplan–Meier or product limit estimator provides an estimate of $S(t)$, $\hat{S}(t)$, from sample of failure times which may be progressively right-censored.

Let t_i, $i = 1, 2, \ldots, n_d$, be the ordered distinct failure times for the sample of observed failure/censored times, and let the number of observations in the sample that have not failed by time t_i be n_i. If a failure and a loss (censored observation) occur at the same time t_i, then the failure is treated as if it had occurred slightly before time t_i and the loss as if it had occurred slightly after t_i.

The Kaplan–Meier estimate of the survival probabilities is a step function which in the interval t_i to t_{i+1} is given by

$$\hat{S}(t) = \prod_{j=1}^{i} \left(\frac{n_j - d_j}{n_j} \right),$$

where d_j is the number of failures occurring at time t_j.

nag_surviv_kaplanmeier (g12aa) computes the Kaplan–Meier estimates and the corresponding estimates of the variances, $\hat{\text{var}}\left(\hat{S}(t)\right)$, using Greenwood's formula,

$$\hat{\text{var}}\left(\hat{S}(t)\right) = \hat{S}(t)^2 \sum_{j=1}^{i} \frac{d_j}{n_j(n_j - d_j)}.$$

4 References

Gross A J and Clark V A (1975) *Survival Distributions: Reliability Applications in the Biomedical Sciences* Wiley

Kalbfleisch J D and Prentice R L (1980) *The Statistical Analysis of Failure Time Data* Wiley

5 Parameters

5.1 Compulsory Input Parameters

1: **t**(**n**) – REAL (KIND=nag_wp) array

The failure and censored times; these need not be ordered.

2: **ic**(**n**) – INTEGER array

ic(i) contains the censoring code of the ith observation, for $i = 1, 2, \ldots,$ **n**.

ic(i) = 0
The ith observation is a failure time.

ic(i) = 1
The ith observation is right-censored.

Constraint: **ic**(i) = 0 or 1, for $i = 1, 2, \ldots,$ **n**.

3: **freq** – CHARACTER(1)

Indicates whether frequencies are provided for each time point.

freq = 'F'
Frequencies are provided for each failure and censored time.

freq = 'S'
The failure and censored times are considered as single observations, i.e., a frequency of 1 is assumed.

Constraint: **freq** = 'F' or 'S'.

4: **ifreq**(:) – INTEGER array

The dimension of the array **ifreq** must be at least **n** if **freq** = 'F' and at least 1 if **freq** = 'S'

If **freq** = 'F', **ifreq**(i) must contain the frequency of the ith observation.

If **ifreq** = 'S', a frequency of 1 is assumed and **ifreq** is not referenced.

Constraint: if **freq** = 'F', **ifreq**(i) ≥ 0, for $i = 1, 2, \ldots,$ **n**.

5.2 Optional Input Parameters

1: **n** – INTEGER

Default: the dimension of the arrays **ic**, **t**. (An error is raised if these dimensions are not equal.)

The number of failure and censored times given in **t**.

Constraint: **n** ≥ 2.

5.3 Output Parameters

1: **nd** – INTEGER

The number of distinct failure times, n_d.

2: **tp**(**n**) – REAL (KIND=nag_wp) array

tp(i) contains the ith ordered distinct failure time, t_i, for $i = 1, 2, \ldots, n_{\mathrm{d}}$.

3: **p**(**n**) – REAL (KIND=nag_wp) array

p(i) contains the Kaplan–Meier estimate of the survival probability, $\hat{S}(t)$, for time **tp**(i), for $i = 1, 2, \ldots, n_d$.

4: **psig**(**n**) – REAL (KIND=nag_wp) array

psig(i) contains an estimate of the standard deviation of **p**(i), for $i = 1, 2, \ldots, n_d$.

5: **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** < 2.

ifail $= 2$

On entry, **freq** $\neq$ 'F' or 'S'.

ifail $= 3$

On entry, **ic**$(i) \neq 0$ or 1, for some $i = 1, 2, \ldots,$ **n**.

ifail $= 4$

On entry, **freq** $=$ 'F' and **ifreq**$(i) < 0$, for some $i = 1, 2, \ldots,$ **n**.

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 computations are believed to be stable.

8 Further Comments

If there are no censored observations, $\hat{S}(t)$ reduces to the ordinary binomial estimate of the probability of survival at time t.

9 Example

The remission times for a set of 21 leukaemia patients at 18 distinct time points are read in and the Kaplan–Meier estimate computed and printed. For further details see page 242 of Gross and Clark (1975).

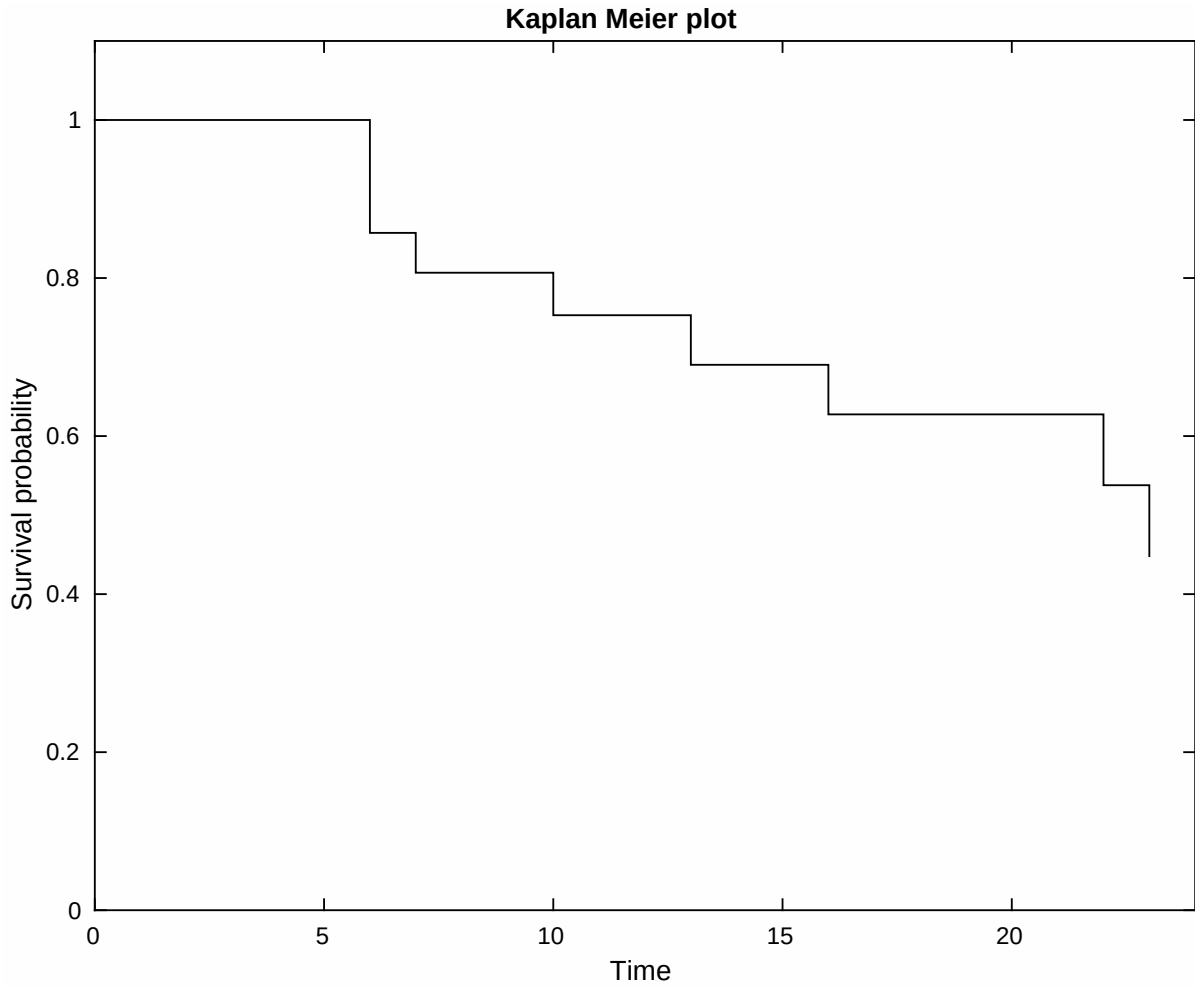
9.1 Program Text

```
function g12aa_example

fprintf('g12aa example results\n\n');

t  = [         6;  6;  7;  9; 10; 10; 11; 13; 16;
              17; 19; 20; 22; 23; 25; 32; 34; 35];
ic = [nag_int(1);  0;  0;  1;  0;  1;  1;  0;  0;
               1;  1;  1;  0;  0;  1;  1;  1;  1];

freq  = 'Frequencies';
ifreq = ones(numel(t),1,nag_int_name);
ifreq(2)  = 3;
ifreq(16) = 2;

% Calculate Kaplan-Meier statistic
```

```
[nd, tp, p, psig, ifail] = g12aa( ...
                                  t, ic, freq, ifreq);

% Display the results
fprintf('  Time   Survival    Standard\n');
fprintf('        probability  deviation\n\n');
fprintf('%6.1f%10.3f%12.3f\n', [tp(1:nd) p(1:nd) psig(1:nd)]');

fig1 = figure;
stp = [0; tp(1:nd)];
sp  = [1; p(1:nd)];
stairs(stp,sp);
xlabel('Time');
ylabel('Survival probability');
title('Kaplan Meier plot');
legend('Off');
axis([0 tp(nd)+1 0 1.1]);
```

9.2 Program Results

```
     g12aa example results

  Time   Survival    Standard
        probability  deviation

   6.0     0.857       0.076
   7.0     0.807       0.087
  10.0     0.753       0.096
  13.0     0.690       0.107
  16.0     0.627       0.114
  22.0     0.538       0.128
  23.0     0.448       0.135
```


Kaplan Meier plot
Survival probability
1
0.8
0.6
0.4
0.2
0
0
5
10
15
20
Time