nag_prod_limit_surviv_fn (g12aac)

1. Purpose

nag_prod_limit_surviv_fn (g12aac) computes the Kaplan-Meier, (or product-limit), estimates of survival probabilities for a sample of failure times.

2. Specification

```
#include <nag.h>
#include <nagg12.h>

void nag_prod_limit_surviv_fn(Integer n, double t[], Integer ic[],
            Integer freq[], Integer *nd, double tp[], double p[],
            double psig[], NagError *fail)
```

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_prod_limit_surviv_fn computes the Kaplan-Meier estimates and the corresponding estimates of the variances, $\hat{\mathrm{var}}(\hat{S}(t))$, using Greenwood's formula,

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

4. Parameters

n

Input: the number of failure and censored times given in **t**.
Constraint: $\mathbf{n} \geq 2$.

t[n]

Input: the failure and censored times; these need not be ordered.

ic[n]

Input: **ic**$[i-1]$ contains the censoring code of the ith observation, for $i = 1, 2, \ldots,$**n**.

If **ic**$[i-1] = 0$ the ith observation is a failure time.

If **ic**$[i-1] = 1$ the ith observation is right-censored.
Constraint: **ic**$[i-1] = 0$ or 1 for $i = 1, 2, \ldots,$**n**.

freq[n]

Input: indicates whether frequencies are provided for each failure and censored time point. If frequencies are provided then **freq** must be dimensioned at least **n**.

If the failure and censored times are to be considered as single observations, i.e., a frequency of 1 is to be assumed then **freq** must be set to NULL.

Constraints: Either **freq** = (Integer $*$)0 or **freq**$[i-1] \geq 0$, for $i = 1, 2, \ldots,$**n**.

nd

Output: the number of distinct failure times, n_d.

tp[n]

Output: **tp**$[i-1]$ contains the ith ordered distinct failure time, t_i, for $i = 1, 2, \ldots, n_d$.

p[n]

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

psig[n]

Output: **psig**$[i-1]$ contains an estimate of the standard deviation of **p**$[i-1]$, for $i = 1, 2, \ldots, n_d$.

fail

The NAG error parameter, see the Essential Introduction to the NAG C Library.

5. Error Indications and Warnings

NE_INT_ARG_LT

On entry, **n** must not be less than 2: **n**=⟨*value*⟩.

NE_INVALID_CENSOR_CODE

On entry, **ic**[⟨*value*⟩]=⟨*value*⟩. The censor code for an observation must be either 0 or 1.

NE_INVALID_FREQ

On entry, **freq**[⟨*value*⟩]=⟨*value*⟩. The value of frequency for an observation must be ≥ 0.

NE_ALLOC_FAIL

Memory allocation failed.

NE_INTERNAL_ERROR

An internal error has occurred in this function. Check the function call and any array sizes. If the call is correct then please consult NAG for assistance.

6. 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.

6.1. Accuracy

The computations are believed to be stable.

6.2. 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.

7. See Also

None

8. Example

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

8.1. Program Text

```
/* nag_prod_limit_surviv_fn(g12aac) Example Program.
 *
 * Copyright 1996 Numerical Algorithms Group.
 *
 * Mark 4, 1996.
 *
 */

#include <nag.h>
#include <stdio.h>
#include <nag_stdlib.h>
#include <nagg12.h>

#define NMAX 18

main()
{

  double psig[NMAX];
  double p[NMAX];
  double t[NMAX];
  double tp[NMAX];

  Integer i, n;
  Integer ifreq[NMAX], ic[NMAX], nd;

  Vprintf("g12aac Example Program Results\n");

  /*     Skip heading in data file */
  Vscanf("%*[^\n] ");

  Vscanf("%ld ", &n);
  if (n <= NMAX)
    {
      for (i = 0; i < n; ++i)
        Vscanf("%lf %ld %ld ", &t[i], &ic[i], &ifreq[i]);

      g12aac(n, t, ic, ifreq, &nd, tp, p, psig, NAGERR_DEFAULT);

      Vprintf("\n   Time   Survival    Standard\n");
      Vprintf("         probability  deviation\n\n");
      for (i = 0; i < nd; ++i)
        Vprintf(" %6.1f%10.3f  %10.3f\n", tp[i], p[i], psig[i]);
    }
  exit(EXIT_SUCCESS);

}
```

8.2. Program Data

```
g12aac Example Program Data
18
6.0 1 1 6.0 0 3 7.0 0 1 9.0 1 1 10.0 0 1 10.0 1 1
11.0 1 1 13.0 0 1 16.0 0 1 17.0 1 1 19.0 1 1 20.0 1 1
22.0 0 1 23.0 0 1 25.0 1 1 32.0 1 2 34.0 1 1 35.0 1 1
```

8.3. Program Results

```
g12aac Example Program 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
```