NAG Fortran Library Routine Document

G12AAF

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

G12AAF computes the Kaplan–Meier, (or product-limit), estimates of survival probabilities for a sample of failure times.

2 Specification

```
SUBROUTINE G12AAF(N, T, IC, FREQ, IFREQ, ND, TP, P, PSIG, IWK, IFAIL)
INTEGER           N, IC(N), IFREQ(*), ND, IWK(N), IFAIL
real              T(N), TP(N), P(N), PSIG(N)
CHARACTER*1       FREQ
```

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.

G12AAF computes the Kaplan–Meier estimates and the corresponding estimates of the variances, $\hat{\mathbf{var}}(\hat{S}(t))$, using Greenwood's formula,

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

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

1: N – INTEGER *Input*

On entry: the number of failure and censored times given in T.

Constraint: N ≥ 2.

2: T(N) – ***real*** array *Input*

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

3: IC(N) – INTEGER array *Input*

On entry: IC(i) contains the censoring code of the ith observation, for $i = 1, 2, \ldots, \mathrm{N}$.

If IC(i) $= 0$ the ith observation is a failure time.

If IC(i) $= 1$ the ith observation is right-censored.

Constraint: IC(i) $= 0$ or 1 for $i = 1, 2, \ldots, \mathrm{N}$.

4: FREQ – CHARACTER*1 *Input*

On entry: indicates whether frequencies are provided for each time point.

If FREQ = 'F', then frequencies are provided for each failure and censored time.

If FREQ = 'S', then the failure and censored times are considered as single observations, i.e., a frequency of 1 is assumed.

Constraint: FREQ = 'F' or 'S'.

5: IFREQ(*) – INTEGER array *Input*

Note: the dimension of the array IFREQ must be at least N if FREQ = 'F' and 1 if FREQ = 'S'.

On entry: if FREQ = 'F', then IFREQ(i) must contain the frequency of the ith observation. If IFREQ = 'S' then a frequency of 1 is assumed and IFREQ is not referenced.

Constraint: if FREQ = 'F', IFREQ(i) ≥ 0, for $i = 1, 2, \ldots, \mathrm{N}$.

6: ND – INTEGER *Output*

On exit: the number of distinct failure times, n_d.

7: TP(N) – ***real*** array *Output*

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

8: P(N) – ***real*** array *Output*

On exit: P(i) contains the Kaplan–Meier estimate of the survival probability, $\hat{S}(t)$, for time TP(i), for $i = 1, 2, \ldots, n_d$.

9: PSIG(N) – ***real*** array *Output*

On exit: PSIG(i) contains an estimate of the standard deviation of P(i), for $i = 1, 2, \ldots, n_d$.

10: IWK(N) – INTEGER array *Workspace*

11: IFAIL – INTEGER *Input/Output*

On entry: IFAIL must be set to 0, -1 or 1. Users who are unfamiliar with this parameter should refer to Chapter P01 for details.

On exit: IFAIL $= 0$ unless the routine detects an error (see Section 6).

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, for users 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.**

6 Error Indicators and Warnings

If on entry 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:

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.

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 leukemia 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

Note: the listing of the example program presented below uses ***bold italicised*** terms to denote precision-dependent details. Please read the Users' Note for your implementation to check the interpretation of these terms. As explained in the Essential Introduction to this manual, the results produced may not be identical for all implementations.

```
*     G12AAF Example Program Text
*     Mark 15 Release. NAG Copyright 1991.
*     .. Parameters ..
      INTEGER          NIN, NOUT
      PARAMETER        (NIN=5,NOUT=6)
      INTEGER          NMAX
      PARAMETER        (NMAX=18)
*     .. Local Scalars ..
      INTEGER          I, IFAIL, N, ND
*     .. Local Arrays ..
      real             P(NMAX), PSIG(NMAX), T(NMAX), TP(NMAX)
      INTEGER          IC(NMAX), IFREQ(NMAX), IWK(NMAX)
*     .. External Subroutines ..
      EXTERNAL         G12AAF
*     .. Executable Statements ..
      WRITE (NOUT,*) 'G12AAF Example Program Results'
*     Skip heading in data file
      READ (NIN,*)
      READ (NIN,*) N
      IF (N.LE.NMAX) THEN
         READ (NIN,*) (T(I),IC(I),IFREQ(I),I=1,N)
         IFAIL = 0
```

```
*
         CALL G12AAF(N,T,IC,'Frequencies',IFREQ,ND,TP,P,PSIG,IWK,IFAIL)
*
         WRITE (NOUT,*)
         WRITE (NOUT,*) '  Time   Survival    Standard'
         WRITE (NOUT,*) '        probability  deviation'
         WRITE (NOUT,*)
         DO 20 I = 1, ND
            WRITE (NOUT,99999) TP(I), P(I), PSIG(I)
   20    CONTINUE
      END IF
      STOP
*
99999 FORMAT (1X,F6.1,F10.3,2X,F10.3)
      END
```

9.2 Program Data

```
G12AAF 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
```

9.3 Program Results

```
 G12AAF 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
```