NAG Toolbox for MATLAB

g12za

1 Purpose

g12za creates the risk sets associated with the Cox proportional hazards model for fixed covariates.

2 Syntax

```
[num, ixs, nxs, x, id, nd, tp, irs, ifail] = g12za(ns, z, isz, ip, t, ic, isi, mxn, 'n', n, 'm', m)
```

3 Description

The Cox proportional hazards model (see Cox 1972b) relates the time to an event, usually death or failure, to a number of explanatory variables known as covariates. Some of the observations may be right-censored, that is, the exact time to failure is not known, only that it is greater than a known time.

Let t_i, for $i = 1, 2, \ldots, n$, be the failure time or censored time for the ith observation with the vector of p covariates z_i. It is assumed that censoring and failure mechanisms are independent. The hazard function, $\lambda(t, z)$, is the probability that an individual with covariates z fails at time t given that the individual survived up to time t. In the Cox proportional hazards model, $\lambda(t, z)$ is of the form

$$\lambda(t, z) = \lambda_0(t) \exp\left(z^{\mathrm{T}}\beta\right),$$

where λ_0 is the base-line hazard function, an unspecified function of time, and β is a vector of unknown parameters. As λ_0 is unknown, the parameters β are estimated using the conditional or marginal likelihood. This involves considering the covariate values of all subjects that are at risk at the time when a failure occurs. The probability that the subject that failed had their observed set of covariate values is computed.

The risk set at a failure time consists of those subjects that fail or are censored at that time and those who survive beyond that time. As risk sets are computed for every distinct failure time, it should be noted that the combined risk sets may be considerably larger than the original data. If the data can be considered as coming from different strata such that λ_0 varies from strata to strata but β remains constant, then g12za will return a factor that indicates to which risk set/strata each member of the risk sets belongs rather than just to which risk set.

Given the risk sets the Cox proportional hazards model can then be fitted using a Poisson generalized linear model (g02gc with g04ea to compute dummy variables) using Breslow's approximation for ties (see Breslow 1974). This will give the same fit as g12ba. If the exact treatment of ties in discrete time is required, as given by Cox 1972b, then the model is fitted as a conditional logistic model using g11ca.

4 References

Breslow N E 1974 Covariate analysis of censored survival data *Biometrics* **30** 89–99

Cox D R 1972b Regression models in life tables (with discussion) *J. Roy. Statist. Soc. Ser. B* **34** 187–220

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

5 Parameters

5.1 Compulsory Input Parameters

1: **ns – int32 scalar**

The number of strata. If $\mathbf{ns} > 0$ then the stratum for each observation must be supplied in **isi**.

Constraint: $\mathbf{ns} \geq 0$.

2: **z(ldz,m) – double array**

ldz, the first dimension of the array, must be at least **n**.

The ith row must contain the covariates which are associated with the ith failure time given in **t**.

3: **isz(m) – int32 array**

Indicates which subset of covariates are to be included in the model.

$\mathbf{isz}(j) \geq 1$

The jth covariate is included in the model.

$\mathbf{isz}(j) = 0$

The jth covariate is excluded from the model and not referenced.

Constraint: $\mathbf{isz}(j) \geq 0$ and at least one value must be nonzero

4: **ip – int32 scalar**

p, the number of covariates included in the model as indicated by **isz**.

Constraint: **ip** = the number of nonzero values of **isz**

5: **t(n) – double array**

The vector of n failure censoring times.

6: **ic(n) – int32 array**

The status of the individual at time t given in **t**.

$\mathbf{ic}(i) = 0$

Indicates that the ith individual has failed at time $\mathbf{t}(i)$.

$\mathbf{ic}(i) = 1$

Indicates that the ith individual has been censored at time $\mathbf{t}(i)$.

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

7: **isi(*) – int32 array**

Note: the dimension of the array **isi** must be at least **n** if $\mathbf{ns} > 0$, and at least 1 otherwise.

If $\mathbf{ns} > 0$, the stratum indicators which also allow data points to be excluded from the analysis.

If $\mathbf{ns} = 0$, **isi** is not referenced.

$\mathbf{isi}(i) = k$

Indicates that the ith data point is in the kth stratum, where $k = 1, 2, \ldots, \mathbf{ns}$.

$\mathbf{isi}(i) = 0$

Indicates that the ith data point is omitted from the analysis.

Constraint: if $\mathbf{ns} > 0$, $0 \leq \mathbf{isi}(i) \leq \mathbf{ns}$, for $i = 1, 2, \ldots, \mathbf{n}$.

8: **mxn – int32 scalar**

and

Constraint: **mxn** must be sufficiently large for the arrays to contain the expanded risk sets. The size will depend on the pattern of failures times and censored times. The minimum value will be returned in **num** unless the function exits with **ifail** $= 1$ or 2

5.2 Optional Input Parameters

1: **n – int32 scalar**

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

n, the number of data points.

Constraint: $\mathbf{n} \geq 2$.

2: **m – int32 scalar**

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

the number of covariates in array **z**.

Constraint: $\mathbf{m} \geq 1$.

5.3 Input Parameters Omitted from the MATLAB Interface

ldz

5.4 Output Parameters

1: **num – int32 scalar**

The number of values in the combined risk sets.

2: **ixs(mxn) – int32 array**

The factor giving the risk sets/strata for the data in **x** and **id**.

If $\mathbf{ns} = 0$ or 1, $\mathbf{ixs}(i) = l$ for members of the lth risk set.

If $\mathbf{ns} > 1$, $\mathbf{ixs}(i) = (j-1) \times \mathbf{nd} + l$ for the observations in the lth risk set for the jth strata.

3: **nxs – int32 scalar**

The number of levels for the risk sets/strata factor given in **ixs**.

4: **x(mxn,ip) – double array**

The first **num** rows contain the values of the covariates for the members of the risk sets.

5: **id(mxn) – int32 array**

Indicates if the member of the risk set given in **x** failed.

$\mathbf{id}(i) = 1$ if the member of the risk set failed at the time defining the risk set and $\mathbf{id}(i) = 0$ otherwise.

6: **nd – int32 scalar**

The number of distinct failure times, i.e., the number of risk sets.

7: **tp(n) – double array**

$\mathbf{tp}(i)$ contains the ith distinct failure time, for $i = 1, 2, \ldots, \mathbf{nd}$.

8: **irs(n) – int32 array**

Indicates rows in **x** and elements in **ixs** and **id** corresponding to the risk sets. The first risk set corresponding to failure time **tp**(1) is given by rows 1 to **irs**(1). The lth risk set is given by rows **id**$(l-1)+1$ to **id**(l), for $l = 1, 2, \ldots,$ **nd**.

9: **ifail – int32 scalar**

0 unless the function detects an error (see Section 6).

6 Error Indicators and Warnings

Errors or warnings detected by the function:

ifail $= 1$

On entry, **m** < 1,
or **n** < 2,
or **ns** < 0,
or **ldz** $<$ **n**.

ifail $= 2$

On entry, **isz**$(i) < 0$ for some i,
or the value of **ip** is incompatible with **isz**,
or **ic**$(i) \neq 1$ or 0.
or **ns** > 0 and **isi**$(i) < 0$,
or **ns** > 1 and **isi**$(i) >$ **ns**.

ifail $= 3$

mxn is too small, the minimum value is returned in **num**.

7 Accuracy

Not applicable.

8 Further Comments

When there are strata present, i.e., **ns** > 1, not all the **nxs** groups may be present.

9 Example

```
ns = int32(0);
z = [0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
     0;
```

```
     0;
     0;
     0;
     0;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1;
     1];
isz = [int32(1)];
ip = int32(1);
t = [1;
     1;
     2;
     2;
     3;
     4;
     4;
     5;
     5;
     8;
     8;
     8;
     8;
     11;
     11;
     12;
     12;
     15;
     17;
     22;
     23;
     6;
     6;
     6;
     7;
     10;
     13;
     16;
     22;
     23;
     6;
     9;
     10;
     11;
     17;
     19;
     20;
     25;
     32;
     32;
     34;
     35];
ic = [int32(0);
```

```
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(0);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1);
     int32(1)];
isi = [int32(0)];
mxn = int32(1000);
[num, ixs, nxs, x, id, nd, tp, irs, ifail] = g12za(ns, z, isz, ip, t, ic,
isi, mxn)
```

```
num =
         418
ixs =
     array elided
nxs =
          17
x =
     array elided
id =
     array elided
nd =
          17
tp =
     1
     2
     3
     4
     5
     6
     7
     8
    10
    11
    12
```

```
    13
    15
    16
    17
    22
    23
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
     0
irs =
          42
          82
         120
         157
         192
         225
         254
         282
         305
         326
         344
         360
         375
         389
         402
         411
         418
          13
          32
          26
          33
          14
          15
          34
          16
          17
          27
          18
          28
          19
          35
          36
          37
          20
          29
          21
          30
          38
```

```
          39
          40
          41
          42
ifail =
           0
```