**Conduct Backward Selection in Multivariable Competing Risk Analysis by Fine & Gray Model**

**Macro:** %FINEGRAY\_SEL

**Created Date/Author:** Sep 23, 2016/Chao Zhang

**Last Update Date/Person**: Oct 2016/Yuan Liu

**Current Version**: V2

**Working Environment:** SAS 9.4 English version

**Contact**: Dr. Yuan Liu [yliu31@emory.edu](mailto:yliu31@emory.edu)

**Purpose:** To conduct multivariable analysis for competing risk model. The proportional subdistribution hazards model as proposed by Fine and Gray (1999) was used followed by backward elimination.

**Notes:** The model runs using PROC PHREG. The final list of variables selected is saved in two macro variables: &\_finalcvar and &\_finalvar, and is also written into the log. The macro “MUTLIPLE\_PHREG V21” or later is also required.

**Reference**: Fine, J. P. and Gray, R. J. (1999), “A Proportional Hazards Model for the Subdistribution of a Competing Risk,” Journal of the American Statistical Association, 94, 496–509.

**Parameters:**

|  |  |
| --- | --- |
| **Macro variable** | **Description** |
| DSN | The name of the data set to be analyzed. |
| EVENT | Name of time to event outcome variable. |
| CENSOR | Name of censoring indicator variable. Values of 0 indicate censored. |
| EVENT\_CODE | The value in CENSOR that indicate event of interest, and this value will appear EVENTCODE= option. |
| VAR | The list of variables on interest in the initial model that would be eliminated during the backward selection procedure separated by spaces. The order of variables in this list will be preserved in the final report. |
| CVAR | The list of categorical variables that are in VAR. If need to change the reference group, you can follow each variable name by (DESC) or by (ref = “Ref level in formatted value”) where needed and separate terms by \*. See code example. |
| ALPHA | The significance level for removing variables from the model (optional). The default value is .2. |
| ID | Variable to be used in the ID statement in PHREG (optional). Refer to SAS Help and Documentation for proper use of this option. ID and COVSAGG will be used together in the model to compute the robust sandwich covariance matrix estimate for cluster data. The ID statement identifies the variable that represents the clusters. If observations have more than one record in the data file, the number of observations reported in the table footer will be the number of records, not unique observations. |
| CLNUM | Set to T if you want to see the number of observations for each level of covariates. The default is T. |
| TYPE3 | Set to F to suppress type III p-values from being reported in the table (optional). The default value is T. This only has an effect if REPORT = T. |
| REPORT | Set it to T if a results summary table is desired. Otherwise check Log for variable selected by the backward elimination. |
| ORIENTATION | Orientation of the output Word table. Default is portrait, can be changed to landscape. |
| FILENAME | File name for output table. This is necessary if report=T. |
| OUTPATH | File path for output table to be stored. This is necessary if report=T. |
| DEBUG | Set to T if running in debug mode (optional). Work datasets will not be deleted in debug mode. This is useful if you are editing the code or want to further manipulate the resulting data sets. The default value is F. |

**Usage Example:**

**proc** **format**;

value DiseaseGroup **1**='ALL' **2**='AML-Low Risk' **3**='AML-High Risk';

**data** Bmt;

input Disease T Status @@;

label T='Disease-Free Survival in Days';

format Disease DiseaseGroup.;

datalines;

1 2081 0 1 1602 0 1 1496 0 1 1462 0 1 1433 0

1 1377 0 1 1330 0 1 996 0 1 226 0 1 1199 0

1 1111 0 1 530 0 1 1182 0 1 1167 0 1 418 2

1 383 1 1 276 2 1 104 1 1 609 1 1 172 2

1 487 2 1 662 1 1 194 2 1 230 1 1 526 2

1 122 2 1 129 1 1 74 1 1 122 1 1 86 2

1 466 2 1 192 1 1 109 1 1 55 1 1 1 2

1 107 2 1 110 1 1 332 2 2 2569 0 2 2506 0

2 2409 0 2 2218 0 2 1857 0 2 1829 0 2 1562 0

2 1470 0 2 1363 0 2 1030 0 2 860 0 2 1258 0

2 2246 0 2 1870 0 2 1799 0 2 1709 0 2 1674 0

2 1568 0 2 1527 0 2 1324 0 2 957 0 2 932 0

2 847 0 2 848 0 2 1850 0 2 1843 0 2 1535 0

2 1447 0 2 1384 0 2 414 2 2 2204 2 2 1063 2

2 481 2 2 105 2 2 641 2 2 390 2 2 288 2

2 421 1 2 79 2 2 748 1 2 486 1 2 48 2

2 272 1 2 1074 2 2 381 1 2 10 2 2 53 2

2 80 2 2 35 2 2 248 1 2 704 2 2 211 1

2 219 1 2 606 1 3 2640 0 3 2430 0 3 2252 0

3 2140 0 3 2133 0 3 1238 0 3 1631 0 3 2024 0

3 1345 0 3 1136 0 3 845 0 3 422 1 3 162 2

3 84 1 3 100 1 3 2 2 3 47 1 3 242 1

3 456 1 3 268 1 3 318 2 3 32 1 3 467 1

3 47 1 3 390 1 3 183 2 3 105 2 3 115 1

3 164 2 3 93 1 3 120 1 3 80 2 3 677 2

3 64 1 3 168 2 3 74 2 3 16 2 3 157 1

3 625 1 3 48 1 3 273 1 3 63 2 3 76 1

3 113 1 3 363 2

;

**data** bmt;

set bmt;

call streaminit(**123**);

u = rand("Uniform");

if u>=**0.7** then Sex='Female'; else if u<**0.5** then Sex='Male';

if u>=**0.4** then Race='white'; else Race='AA';

call streaminit(**456**);

Age = rand("Uniform")\***90**;

Drop u;

**run**;

option mprint;

%let dir = C:\;

title 'Table 3 Multivariable Survival Analysis by Fine-Gray Model';

%***finegray\_sel*** (dsn=bmt, event=T, censor=Status, event\_code=**1**,

var= sex race age disease,

inc= 1,

cvar=race(ref="AA")\*Disease(desc)\*sex(ref="Male"),

alpha=**0.2**, type3=T,

outpath = &dir,

filename = Table **3** Multivariable Survival Analysis by Fine-Gray Model);

title;

**Summary Table Example:**

Table 3 Multivariable Survival Analysis by Fine-Gray Model

|  | | | **Disease-Free Survival in Days** | | |
| --- | --- | --- | --- | --- | --- |
|  | | | **----------------------------------------** | | |
| **Covariate** | **Level** | **N** | **Hazard Ratio (95% CI)** | **HR P-value** | **Type3 P-value** |
| Sex | Female | 54 | 1.37 (0.69-2.73) | 0.374 | 0.374 |
| Male | 57 | - | - |
|  | | | | | |
| Disease | AML-High Risk | 35 | 2.16 (0.96-4.84) | 0.062 | **0.001** |
| AML-Low Risk | 46 | 0.44 (0.17-1.16) | 0.097 |
| ALL | 30 | - | - |
|  | | | | | |
| \*  Number of observations in the original data set = 137. Number of observations used = 111. \*\* Backward selection with an alpha level of removal of 0.2 was used. The following variables were forced in the model: sex. The following variables were removed from the model: Age, and Race. | | | | | |

**Permission:**

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:   
  
The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.  
  
THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

**Log of Updates:**

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **By** | **Description** | **Version** |
| Sep, 2016 | Chao Zhang | Initial Creation | V1 |
| Oct, 2016 | Yuan Liu | Simplify variable elimination step; allow categorical variable to set desired reference level. The output order of variable is the same as input order | V2 |
| Oct, 2016 | Yuan Liu | Remove FORCEINVAR, and add INC parameter, so that this macro is consistent with %PHREG\_SEL | V3 |
| March, 2017 | Chao Zhang | Remove the parameter, i.e.,COVSAGG | V4 |