**Conduct Backward Selection in Multivariable Generalized Linear Regression**

**Macro:** GENMOD\_SEL

**Created Date/Author:** Dec. 11, 2013/Dana Nickleach

**Last Update Date/Person/Contact:**

**Current Version:** V1

**Working Environment:** SAS 9.3 English version

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

**Purpose:** To conduct backward selection on a multivariable generalized linear regression model using the maximum possible sample size at each stage of the selection process instead of restricting to the sample size from the first step as SAS does when using their selection methods. The normal, binomial, Poisson, or negative binomial distributions can be used. An identity link will be used with a normal distribution, a logit link with binomial, and a log link with Poisson or negative binomial. Generalized estimating equations (GEE) may also be used. Optionally, a table of the resulting model can be generated.

**Notes:** The model is run using PROC GENMOD. The final list of variables selected will be written to the log. Additionally, two global macro variables, \_finalvar and \_finalcvar will be created containing the list of all variables and categorical variables selected, respectively. If you are requesting a table with the model results then the macro “MUTLIPLE\_GENMOD V7” or later is also required. Interactions can be included if you are not requesting a table (REPORT=F). However, model hierarchy will not be maintained so interactions should only be used with caution. Also, the correct results might not be produced if interactions terms are included without their main effects and the macro cannot produce a report table if the model contains interactions.

**Parameters:**

|  |  |  |
| --- | --- | --- |
| **Macro variable** | **Description** | **Required** |
| DSN | The name of the data set to be analyzed. | Yes |
| OUTCOME | The name of the continuous or binary outcome variable. | Yes |
| DESC | Set to T to reverse the order of a binary outcome. The default value is F. | No |
| VAR | List of variables to include in the model separated by spaces. | Yes |
| CVAR | List of categorical variables to include in the model separated by spaces. These should also appear in the VAR parameter. If you want to change the reference group you can follow each variable name by (DESC) where needed. However, you will need to separate terms with an asterisk instead of a space. | Yes |
| SUBJECT | Subject-effect to be specified in REPEATED statement. Leave blank if not using a GEE model. | Yes |
| WITHINSUB | Within-subject-effect to be specified in REPEATED statement. This is optional if using a GEE model. Leave blank if not using a GEE model. | No |
| DIST | Probability distribution to use for the model. Valid values are NORMAL, BINOMIAL, POISSON, and NEGBIN. The default value is NORMAL. | No |
| TYPE | Correlation structure keyword to be specified in REPEATED statement. Leave blank if not using a GEE model. | Yes |
| INC | Number of variables to include in the model. The first n variables in the VAR parameter will be included in every model. The default value is 0. | No |
| SLSTAY | The significance level for removing variables from the model. The default value is .05. | No |
| TYPE3 | Set to F to suppress type III p-values from being reported in the table. The default value is T. This only has an effect if REPORT = T. | No |
| REPORT | Set this to T if you want a table of the resulting model generated. The default value is F. | No |
| OUTPATH | File path for output table to be stored. | Yes, if REPORT=T |
| FNAME | File name for output table. | Yes, if REPORT=T |
| DEBUG | Set to T to run in debug mode. 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. | No |

**Usage Example:**

DATA analysis;

input group id os\_censor Sex $ Age duration os progress $ trt $;

LABEL os = 'Overall Survival (months)'

progress = 'Progression'

trt = 'Treatment'

duration = 'Duration of Radiation';

DATALINES;

1 1 1 M 40 44 20 No B

1 2 1 F 45 46 16 Yes A

1 3 1 F 40 32 20 No B

1 4 1 F 47 32 23 No B

2 5 0 M 41 25 22 No B

2 6 1 M 54 35 13 No B

2 7 1 M 48 50 9 Yes A

3 8 1 M 36 33 12 Yes B

3 9 0 F 49 51 8 Yes A

3 10 1 M 49 52 10 Yes A

4 11 1 M 44 35 12 No A

4 12 1 M 49 50 8 Yes A

4 13 1 M 44 44 14 Yes A

4 14 1 M 50 31 10 Yes A

5 15 1 M 53 40 15 No B

5 16 0 M 52 29 20 Yes B

6 17 1 F 46 45 5 Yes A

6 18 1 F 37 44 11 Yes A

6 19 1 M 49 46 13 No B

6 20 1 M 42 31 11 No A

;

TITLE 'Table 4 GEE Model';

%genmod\_sel(dsn=analysis,

outcome=duration,

desc=F,

var= age trt sex,

cvar=trt sex,

subject=group,

type=EXCH,

dist=NORMAL,

inc=1,

slstay=.05,

type3=F,

report=T,

outpath= C:\Documents and Settings\User\My Documents\,

fname=Table 4 Linear Regression);

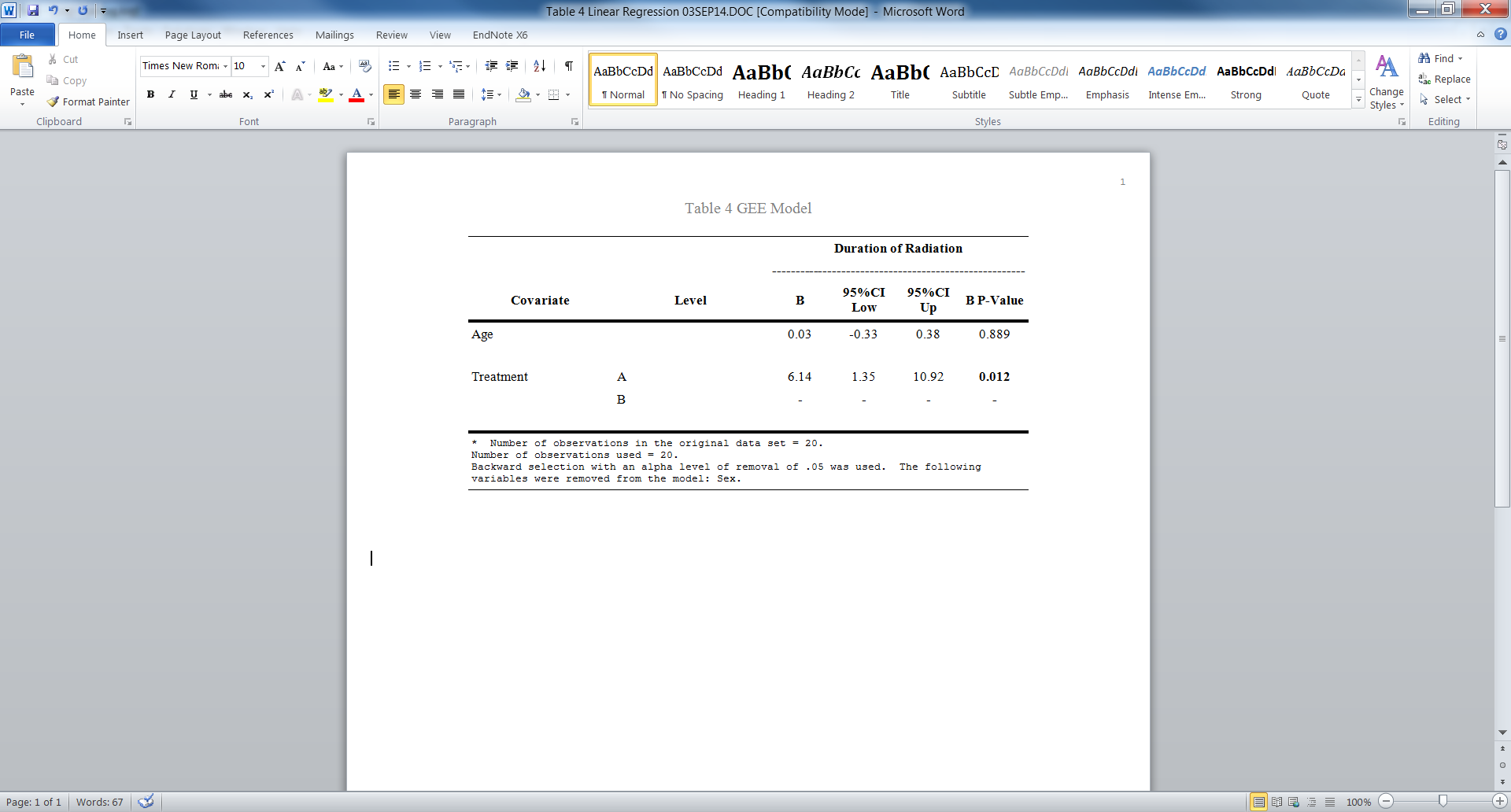
TITLE;

This message will be written in the log:

Categorical variables selected: TRT

All variables selected: AGE TRT

**Summary Table Example:**



**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** |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |