**Martingale Residual Plot and Optimal Cut Point**

**Macro Name:** FINDCUT

**Created Date/Author** Feb. 2012/Yuan Liu

**Last Update Date/Person** Dec. 9, 2013/Dana Nickleach

**Current Version**: V17

**Working Environment:** SAS 9.3 English version

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

**Purpose:** To search for an optimal cut point for a continuous variable that predicts time to event outcome and produce a martingale residual plot for the variable, which will further help us to detect any nonlinear relationship.

**Notes:** The major part of this macro was from <http://www2.sas.com/proceedings/sugi28/261-28.pdf>.

Please cite properly. The log-rank rank statistic, not log-rank chi-square statistic, is used to choose the optimal cut point. The cut point is chosen based on the data being dichotomized into groups X< cut point vs. X>= cut point.

The macro MPLOT V11 or later and the macro NO\_OUT V4 or later is required if OUTLIER=T or B. Besides the RTF document a SAS dataset containing the optimal cut points will be created.

**Parameters:**

|  |  |
| --- | --- |
| **Macro variable** | **Description** |
| DS | The name of the data set to be analyzed. |
| TIME | Time to event outcome variables separated by spaces. |
| STAT | The variable name for censoring indicators separated by spaces. The order of this list should correspond to the order of the TIME parameter. **It is required that 1 is used for the event and 0 for censored cases.** |
| CUTVAR | The numerical variables to be plotted separated by spaces. The variable name cannot be more than 26 characters. |
| SM | Smoothing parameter. The default value is 0.65. |
| DOC | Set to T to create a RTF file containing the output or F to suppress creation of the RTF file (optional). The default value is T. |
| OUTPATH | Path for output table to be stored |
| FNAME | File name for output table. |
| MINOUT | Request minimal output (optional). Set to T to minimize output, which will only include the martingale residual plot, percentiles, and optimal cut point. The default value if F. |
| SUBTIT | Text to be added to the title to appear after, but on the same line as the first title (optional). Text should not be enclosed in quotes. By default no additional text will be printed. |
| PLOT | Set to F to suppress the martingale residual plots (optional). The default value is F. |
| HREF | Set to T for a horizontal reference line at zero on the plot (optional). The default value is F. |
| VREF | Set to T for vertical reference lines at the median and optimal cut point on the plot (optional). The default value is F. |
| CHOSEN | Value of the chosen cut point to display with a vertical reference line in red on the plot (optional). This is not recommended if processing more than one probe or data set since the same point will be used for all plots. |
| DEC | Fixed number of decimal places to display in the percentile table (optional). Note that the maximum format width is 7. So the number of decimal places can't be more than 5, but depending on the size of the numbers may be less. The best format is used by default. |
| OUTLIER | Set to T, F, or B (optional). T will exclude extreme outliers from the analysis. F will not exclude any points. B will perform the analysis both ways, including all points and excluding outliers. Note that if no outliers are found the second analysis will not be reported. The default value is F. |
| OUT | The name of the SAS dataset that contains the optimal cut point results. By default it will be called Cutpoint. |
| 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:**

**DATA** analysis;

input 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 M 40 44 20 No B

2 1 F 45 46 16 Yes A

3 1 F 40 32 20 No B

4 1 F 47 32 23 No B

5 0 M 41 25 22 No B

6 1 M 54 35 13 No B

7 1 M 48 50 9 Yes A

8 1 M 36 33 12 Yes B

9 0 F 49 51 8 Yes A

10 1 M 49 52 10 Yes A

11 1 M 44 35 12 No A

12 1 M 49 50 8 Yes A

13 1 M 44 44 14 Yes A

14 1 M 50 31 10 Yes A

15 1 M 53 40 15 No B

16 0 M 52 29 20 Yes B

17 1 F 46 45 5 Yes A

18 1 F 37 44 11 Yes A

19 1 M 49 46 13 No B

20 1 M 42 31 11 No A

;

TITLE 'Figure 1 Martingale Residual Plot and Optimal Cut Point';

%FINDCUT(ds = analysis,

time = os,

stat = os\_censor,

cutvar = duration,

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

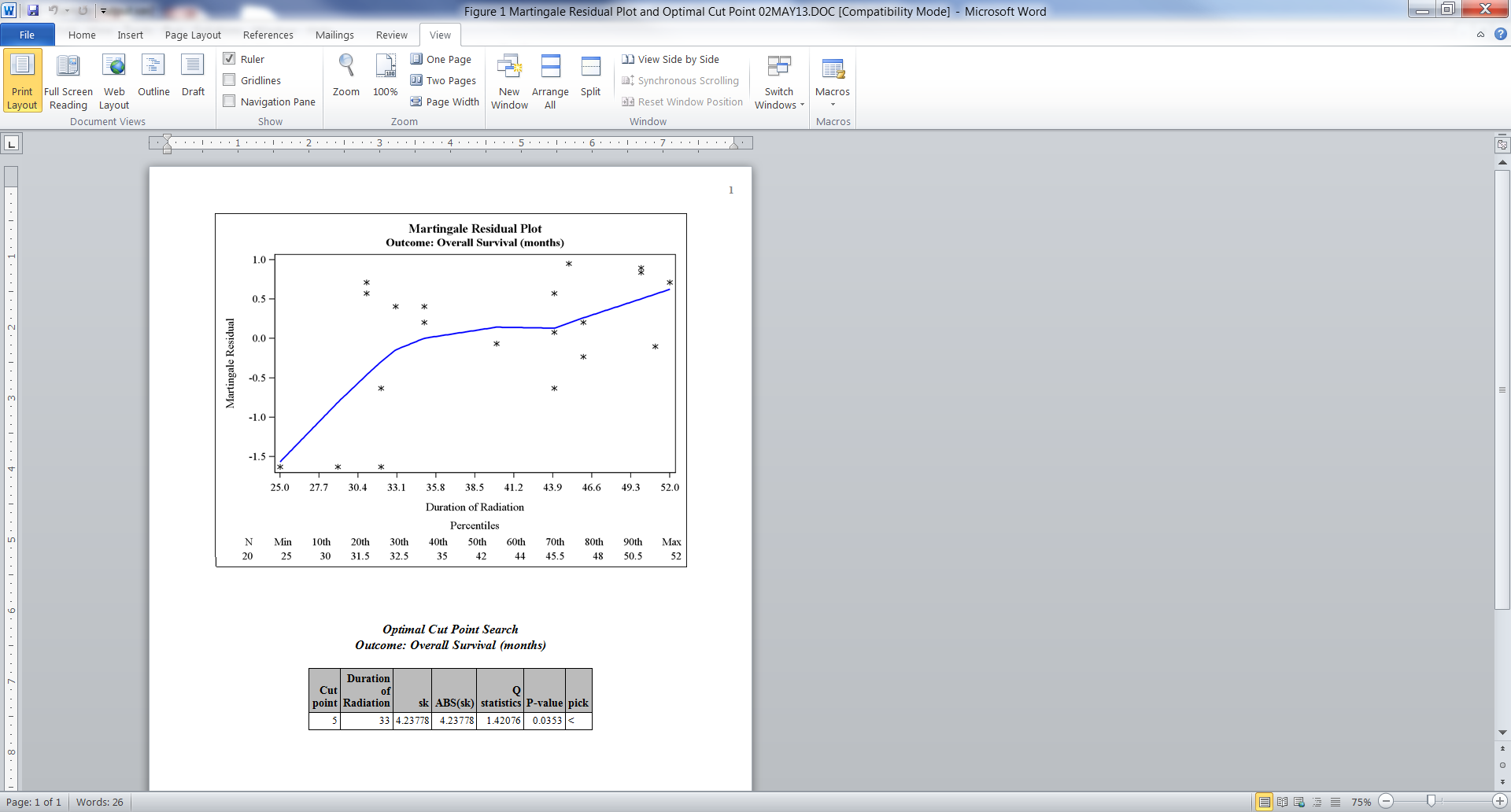
fname = Figure 1 Martingale Residual Plot and Optimal Cut Point,

minout = T,

subtit = );

TITLE;

**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** |
| 7/19/2012 | Dana Nickleach  (dnickle@emory.edu) | Prevented deletion of temporary data sets not created by the macro and added restoration of original options at the end of the macro. Removed obsno from a keep statement to avoid a warning. Removed creation of nvar data set as it was unnecessary and was causing a merge issue. Added minout and subtit parameter to macro. Trimmed extra space in title. Changed “<===” to “<” due to problems in subsetting. | V2 |
| 7/24/2012 | Dana Nickleach (dnickle@emory.edu) | Added debug parameter and slightly increased horizontal axis to avoid points being cut off. | V3 |
| 9/13/2012 | Dana Nickleach (dnickle@emory.edu) | Added vref option and dropped unnecessary variables to decrease processing time. | V4 |
| 10/22/12 | Dana Nickleach (dnickle@emory.edu) | Added check for cutvar that is too many characters and prevented case sensitivity for T/F parameters. | V5 |
| 10/26/12 | Sungjin Kim  (skim61@emory.edu) | Got rid of leading space in RTF file name. | V6 |
| 11/7/12 | Dana Nickleach (dnickle@emory.edu) | Added bodytitle option to RTF file, added N to output, labeled percentiles in output, added outlier parameter, changed proc means to use output statement instead of ODS to prevent output of means if running with minout=T. | V7 |
| 11/12/12 | Dana Nickleach (dnickle@emory.edu) | Modified format so that small p-values will be printed as <.0001. | V8 |
| 12/21/12 | Dana Nickleach (dnickle@emory.edu) | Added noobs option to print of optimal cut point and adjusted so that the printed cutvar would not lose its original label | V9 |
| 2/8/13 | Dana Nickleach (dnickle@emory.edu) | Replaced creation of martingale residual plot with a call to MPLOT so that code updates to not have to be made in two different places, added outlier option B. | V10 |
| 2/12/13 | Dana Nickleach (dnickle@emory.edu) | Added capability to handle multiple outcomes and multiple cutvars, modified deleting of data sets, added out parameter and creation of final data set that holds all cutpoints. | V11 |
| 2/27/13 | Dana Nickleach (dnickle@emory.edu) | Removed variable name from plot title and redefined subtit parameter so that additional text can be added to the main title, changed VREF parameter to be called HREF, added VREF and DEC parameters, and defined macro vars as local to avoid problems when calling from other macros. | V12 |
| 4/10/13 | Dana Nickleach (dnickle@emory.edu) | Added CHOSEN parameter. | V13 |
| 7/16/13 | Dana Nickleach (dnickle@emory.edu) | Made some changes to the output data set in order to be consistent with clinical workflow macros. | V14 |
| 10/10/13 | Dana Nickleach (dnickle@emory.edu) | Fixed so that commas in variable labels will not cause errors. | V15 |
| 10/28/13 | Dana Nickleach (dnickle@emory.edu) | Added PLOT parameter and modified so that outlier variable always appears in output SAS data set. | V16 |
| 12/9/13 | Dana Nickleach (dnickle@emory.edu) | Made some minor changes to clean up code and decrease runtime. | V17 |