WLW: A SAS Macro for Combining Evidence on Multiple Events

Suppose that we are interested in K types of events, such as improvement by one category, "clinical improvement" (i.e., improvement by two categories or hospital discharge), improvement by three categories, deterioration by one category, deterioration by two categories, critical illness, and death. For $k = 1, \ldots, K$, let T_k denote the time from randomization to the kth type of event, and X_k and Z_k be the corresponding treatment indicator and prognostic factors (e.g., baseline severity rating, age), respectively. We relate T_k to X_k and Z_k through proportional hazards models, such that the hazard function of T_k takes the form

$$\lambda_k(t|X_k, Z_k) = \lambda_{k0}(t)e^{\beta_k X_k + \gamma_k^{\mathrm{T}} Z_k} \qquad (k = 1, \ldots, K),$$

where $\lambda_{k0}(\cdot)$ is an arbitrary baseline hazard function, β_k is the log hazard ratio for treatment vs control, and γ_k is the set of log hazard ratios for the prognostic factors. We code X_k in such a way that a positive value of β_k means that treatment is beneficial. Thus, X_k is coded as 1 for treatment and 0 for control if the kth event pertains to improvement of clinical status (e.g, recovery, clinical improvement), and X_k is coded as 0 for treatment and 1 for placebo if the kth event pertains to deterioration of clinical status (e.g., critical illness, death). Using the method of Wei, Lin and Weissfeld (1989) (WLW), we obtain the estimates $(\widehat{\beta}_1, \ldots, \widehat{\beta}_K)$, together with their (estimated) covariance matrix $\widehat{\Psi}$. This is available in SAS; see the section on "Marginal Cox Models for Multiple Events Data" in the PHREG documentation. The coding of the treatment indicators may differ from the example in the PHREG documentation, depending on what kinds of events we are combining.

To test the null hypothesis that $\beta_1 = \beta_2 = \ldots = \beta_K = 0$, we consider the test statistic

$$T = \frac{\sum_{k=1}^{K} w_k \widehat{\beta}_k}{\left(\sum_{k=1}^{K}\sum_{l=1}^{K} w_k w_l \widehat{\psi}_{kl}\right)^{1/2}},$$

where $\widehat{\psi}_{kl}$ is the (k,l)th element of $\widehat{\Psi}$. This test statistic is asymptotically standard normal under the null hypothesis.

WLW proposed the following weighting scheme: $(w_1, \ldots, w_K)^{\mathrm{T}} = (e^{\mathrm{T}}\widehat{\Psi}^{-1}e)^{-1}\widehat{\Psi}^{-1}e$, where $e = (1, \ldots, 1)^{\mathrm{T}}$. This choice of weights minimizes the variance of the linear combination of $\widehat{\beta}_1, \ldots, \widehat{\beta}_K$. The AVERAGE option in SAS PHREG computes the optimal weights and performs a one degree of freedom test. This joint test will be more powerful than the separate test of the treatment effect on each type of event if $\beta_1, \ldots, \beta_K$ are similar.

It is more intuitive to combine the p-values or the Z-scores. If we set $w_k = 1/\widehat{\psi}_{kk}^{1/2}$, then

$$\sum_{k=1}^{K} w_k \widehat{\beta}_k = \sum_{k=1}^{K} \widehat{\beta}_k / \widehat{\psi}_{kk}^{1/2},$$

which is the sum of the Z-scores. We can calculate the corresponding test statistic T by using the SAS PRREG output for $\widehat{\beta}_1, \ldots, \widehat{\beta}_K$ and $\widehat{\Psi}$.

We have written a SAS macro to implement the above two versions of the WLW method, to be referred to as "Optimal weights" and "Combined Z-scores", respectively. We generate the test statistics and p-values for testing the null hypothesis that $\beta_1 = \beta_2 = \ldots = \beta_K = 0$. We also output the log hazard ratio estimates and the standard error estimates.

Synopsis

This macro is defined with the name "**WLW**":

%macro **WLW**(**Data** = , **ID** = , **Enum** = , **Time** = , **Status** = , **Treatment** = , **Covariates** = %str(NONE), **Test** = %str(one.sided));

.....
%mend **WLW**;

Options

Data Name of the input data file.

ID Name of the identification variable such that all observations from a subject have the same ID.

Enum Name of the variable to index the event type (e.g., Enum=1 for event 1, Enum=2 for event 2). This variable can be numeric or character.

Time Name of the variable to represent the observation time (event time or censoring time).

Status Name of the variable to indicate whether the observation time is the event time or the censoring time, with the value that indicates censoring enclosed in parentheses. This is the same as the MODEL statement in SAS PHREG, where Status=1 is the default for censoring, with no parentheses needed. If status=0 indicates censoring, then Status=status(0) should be used.

Treatment Name of the treatment indicator. This variable should be coded in such a way that a positive value of log hazard ratio means that treatment is beneficial.

Covariates Names of baseline prognostic factors. Leave this argument blank if there are no prognostic factors. Must be space delimited if there is more than one factor (e.g., Covariates=Var1 Var2 Var3). Prognostic factors must be numeric.

Test This option must be "one.sided" (default) or "two.sided", corresponding to the one-sided and two-sided tests, respectively.

Example

The file "simdata" shown below is a simulated data set with 500 subjects and 4 types of events. The variable "treatment" is coded as 1 for treatment and 0 for control. There are 7 categories of severity rating, and the baseline severity rating, denoted by "basecat", can be 3, 4, or 5. The four events are:

- Event 1: clinical improvement (i.e., decline of severity by two categories or reaching Category 2);
- Event 2: improvement of clinical status by one category;

- Event 3: clinical recovery (i.e., reaching Category 2);
- Event 4: death (i.e., reaching Category 7).

The variable "time$k$" is the observation time for the kth event, which is either censored (status$k$=0) or observed (status$k$=1), for $k = 1, 2, 3, 4$.

id	treatment	basecat	time1	status1	time2	status2	time3	status3	time4	status4
1	0	4	9	1	2	1	9	1	28	0
2	1	4	10	1	4	1	10	1	28	0
3	1	3	26	1	26	1	26	1	28	0
4	1	4	20	1	6	1	20	1	28	0
5	1	3	28	0	28	0	28	0	28	0
6	0	5	9	1	1	1	28	0	28	0
7	1	3	1	1	1	1	1	1	28	0
8	1	4	10	1	9	1	10	1	28	0
9	0	4	28	0	6	1	28	0	28	0
10	0	3	2	1	2	1	2	1	28	0
11	0	3	4	1	4	1	4	1	28	0
12	1	3	10	1	10	1	10	1	28	0
13	0	4	28	0	19	1	28	0	28	0
14	1	4	28	0	9	1	28	0	28	0
15	0	4	28	0	28	0	28	0	28	0
16	0	4	7	1	1	1	7	1	28	0
17	1	4	28	0	28	0	28	0	9	1
18	0	4	6	1	4	1	6	1	28	0
19	1	3	1	1	1	1	1	1	28	0
20	0	4	3	1	0	1	3	1	28	0
21	0	3	28	0	28	0	28	0	28	0
22	0	4	6	1	5	1	6	1	28	0
23	0	3	7	1	7	1	7	1	28	0
24	0	5	28	0	28	0	28	0	5	1
25	0	5	28	0	4	1	28	0	10	1
26	0	3	4	1	4	1	4	1	28	0
27	0	4	28	0	23	1	28	0	28	0
28	1	3	11	1	11	1	11	1	28	0
29	1	3	6	1	6	1	6	1	28	0
30	1	3	6	1	6	1	6	1	28	0

A dataset named "simdata"

Suppose that we wish to combine the evidence of treatment effects on events 1 and 2 (i.e., clinical improvement and improvement by one category) and also wish to combine the evidence of treatment effects on events 3 and 4 (i.e., clinical recovery and death). We convert "simdata" into two separate files named "example1" and "example2", which can be passed to the SAS macro **WLW**. The file "example1" contains the data on events 1 and 2, and the file "example2" contains the data on events 3 and 4. In both files, we create two dummy variables named "basecat4" and "basecat5" for baseline Categories 4 and 5, respectively, with Category 3 as the reference. For events 1, 2, and 3, "treatment" is coded as 1 for treatment and 0 for control. For event 4, "treatment" is coded as 0 for treatment and 1 for control. We use "event" to denote the event type and "time" and "status" to denote the observation time and the censoring status (0 for censoring), respectively. A snapshot of the two files is provided below.

id	treatment	basecat4	basecat5	event	time	status
1	0	1	0	1	9	1
1	0	1	0	2	2	1
2	1	1	0	1	10	1
2	1	1	0	2	4	1
3	1	0	0	1	26	1
3	1	0	0	2	26	1
4	1	1	0	1	20	1
4	1	1	0	2	6	1
5	1	0	0	1	28	0
5	1	0	0	2	28	0
6	0	0	1	1	9	1
6	0	0	1	2	1	1
7	1	0	0	1	1	1
7	1	0	0	2	1	1
8	1	1	0	1	10	1
8	1	1	0	2	9	1
9	0	1	0	1	28	0
9	0	1	0	2	6	1
10	0	0	0	1	2	1
10	0	0	0	2	2	1
11	0	0	0	1	4	1
11	0	0	0	2	4	1
12	1	0	0	1	10	1
12	1	0	0	2	10	1
13	0	1	0	1	28	0
13	0	1	0	2	19	1
14	1	1	0	1	28	0
14	1	1	0	2	9	1
15	0	1	0	1	28	0
15	0	1	0	2	28	0

id	treatment	basecat4	basecat5	event	time	status
1	0	1	0	3	9	1
1	1	1	0	4	28	0
2	1	1	0	3	10	1
2	0	1	0	4	28	0
3	1	0	0	3	26	1
3	0	0	0	4	28	0
4	1	1	0	3	20	1
4	0	1	0	4	28	0
5	1	0	0	3	28	0
5	0	0	0	4	28	0
6	0	0	1	3	28	0
6	1	0	1	4	28	0
7	1	0	0	3	1	1
7	0	0	0	4	28	0
8	1	1	0	3	10	1
8	0	1	0	4	28	0
9	0	1	0	3	28	0
9	1	1	0	4	28	0
10	0	0	0	3	2	1
10	1	0	0	4	28	0
11	0	0	0	3	4	1
11	1	0	0	4	28	0
12	1	0	0	3	10	1
12	0	0	0	4	28	0
13	0	1	0	3	28	0
13	1	1	0	4	28	0
14	1	1	0	3	28	0
14	0	1	0	4	28	0
15	0	1	0	3	28	0
15	1	1	0	4	28	0

Input files "example1" (left) and "example2" (right)

We run the following SAS statement to combine the evidence of treatment effects on events 1 and 2:

```
%WLW(Data=example1, ID=id, Enum=event, Time=time, Status=status(0),
Treatment=treatment, Covariates=basecat4 basecat5, Test=one.sided);
```

The output is given below. In the first table, "treatment1", "basecat41", and "basecat51" pertain to event 1, and "treatment2", "basecat42", and "basecat52" pertain to event 2. The optimal weights are shown in the second table. The test statistics and p-values for the two versions of the WLW method are presented in the third table. Both versions of the WLW method provide stronger evidence about the treatment benefit than the separate analysis on event 1 or 2.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	StdErr	ChiSq	ProbChiSq
treatment1	1	0.20418	0.10579	3.7247	0.0536
treatment2	1	0.18443	0.09316	3.9195	0.0477
basecat41	1	-0.88025	0.11042	63.5475	<.0001
basecat42	1	-0.12393	0.09631	1.6556	0.1982
basecat51	1	-1.00091	0.19561	26.1828	<.0001
basecat52	1	-0.41937	0.18881	4.9331	0.0263

Optimal Weights for Test Treatment	
Parameter	Optimal weight
treatment1	0.16264
treatment2	0.83736

One-Sided Test Results of the WLW Method		
Method	Test statistic	P-value
Optimal weights	2.025808161	0.0213922237
Combined Z-scores	2.0498116009	0.0201914096

Output for example1

Likewise, we run the following SAS statement to combine the evidence of treatment effects on events 3 and 4:

%WLW(Data=example2, ID=id, Enum=event, Time=time, Status=status(0),
Treatment=treatment, Covariates=basecat4 basecat5, Test=one.sided);

The output is given below. In the first table, "treatment3", "basecat43", and "basecat53" pertain to event 3, and "treatment4", "basecat44", and "basecat54" pertain to event 4. Both versions of the WLW method provide stronger evidence about the treatment benefit than the separate analysis on event 3 or 4.

Analysis of Maximum Likelihood Estimates					
Parameter	DF	Estimate	StdErr	ChiSq	ProbChiSq
treatment3	1	0.18977	0.10765	3.1075	0.0779
treatment4	1	0.41867	0.23146	3.2718	0.0705
basecat43	1	-0.88263	0.11156	62.5954	<.0001
basecat44	1	0.87703	0.29709	8.7148	0.0032
basecat53	1	-1.53572	0.21785	49.6952	<.0001
basecat54	1	2.01610	0.33803	35.5731	<.0001

Optimal Weights for Test Treatment	
Parameter	Optimal weight
treatment3	0.99499
treatment4	0.00501

One-Sided Test Results of the WLW Method		
Method	Test statistic	P-value
Optimal weights	1.7735374676	0.0380698445
Combined Z-scores	2.092598573	0.018192502

Output for example2

References

Wei, L. J., Lin, D. Y., and Weissfeld. L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association*, **84**, 1065–1073.