Competing Risk Survival Models

Dan N

2024-02-28

Example Cause-Specific Hazard Model in SAS

The hazard function is the instantaneous rate of the occurrence of the event of interest in subjects who are currently at risk of the event (or for whom the event has not yet occurred). Cox proportional hazards regression model is used to estimate the relative effect of covariates on the hazard function. In survival analysis, a competing risk is an event whose occurrence precludes the occurrence of the primary event of interest.The cause-specific hazard function for a given event type is the instantaneous rate of occurrence of the given type of event in subjects who are currently event-free (Austin, et al, 2016).

Data Prep

```
time variable      = wl_months
censor variable    = status
indicator variable = dz (slk=0 lt=1)
covariates         = gender (m=0, f=1)
                   = init_age (continuous)

data a; set sample_data;
keep gender dayswait_chron init_age status dz wl_months _status _gender _dz ;

if gender = 'M' then _gender = 0;
  else     /*F*/     _gender = 1;

if dz = 'SLK' then _dz = 1;
  else  /*LT*/     _dz = 0;

if status          = "Waiting"   then        _status = 0;
  else if status   = "TX"        then        _status = 1;
    else if status = "Drop"      then        _status = 2;
      else          /*Died*/                _status = 3;

wl_months=dayswait_chron/30.5;
wl_months=ROUND(wl_months,1);
run;
```

Assess Proportional Hazard Assumption

Need to assess for each covariate whether this assumption of proportional hazards is reasonable. If non-proportional hazards are present:

- Use separate relative risks for early and late
- Stratified model

```
ods graphics on;
ods pdf file="c:/Assess PH.pdf" startpage=no;
proc phreg data=a;
model wl_months*_status(0,1,2) = _dz init_age _gender dz_time age_time gender_time;
dz_time      = _dz*log(wl_months);
age_time     = init_age*log(wl_months);
gender_time  = _gender*log(wl_months);
title 'Assess PH Assumption with Time-Dependent Method';
run;
ods pdf close;
ods graphics off;
```

Assess PH Assumption with Time-Dependent Method

Analysis of Maximum Likelihood Estimates							
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	Label
_dz	1	0.66998	0.15874	17.8127	<.0001	1.954	
init_age	1	0.01808	0.01596	1.2837	0.2572	1.018	CALCULATED AGE AT LISTING
_gender	1	0.55967	0.11256	24.7217	<.0001	1.750	
dz_time	1	0.09429	0.10608	0.7902	0.3740	1.099	
age_time	1	-0.00393	0.01060	0.1376	0.7107	0.996	
gender_time	1	-0.14205	0.07482	3.6045	0.0576	0.868	

Here i took a time-dependent method approach to test proportionality. This involves generating the time dependent covariates by creating interactions of the predictors and a function of survival time and included these interactions in the model.
If any of the interactions are significant, then those predictors are not proportional. Since none of the interactions are significant, we can be confident the proportional hazard assumption does indeed hold.

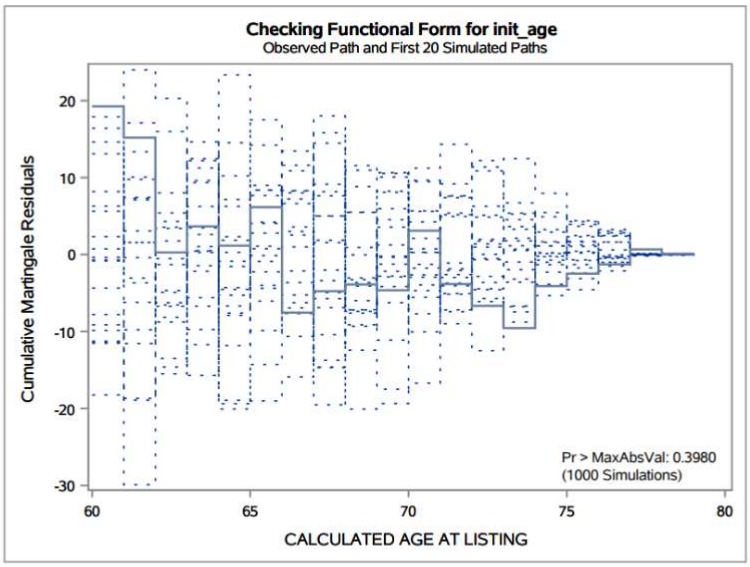
Check Functional Form of a Continuous Covariate

Checking the functional form of a continuous covariate. We can plot the cumulative Martingale residuals against all levels of the covariate.

- Unusually large values suggest a problem with function form.
- This plot is a randomly generated residual process (1000 times)

```
ods graphics on;
ods pdf file="C:/Check Functional Form.pdf" startpage=no;
proc phreg data=a;
model wl_months*_status(0,1,2) = init_age/ rl;
assess var=(init_age)/resample;
title 'Check Functional Form of Continuous Variables';
run;
ods pdf close;
ods graphics off;
```

The PHREG Procedure



Supremum Test for Functional Form

Variable	Maximum Absolute Value	Replications	Seed	Pr > MaxAbsVal
init_age	19.2539	1000	1827489653	0.3980

The observed process (solid line) does not appear to be atypical from any of the simulated processes. We can conclude that age has no problems with linearity, and no transformation is needed.

Model Building

By treating all competing events (“TX”=1, “Drop”=2) as censored we can ensure the risk set at each event time holds only those (“Died”=0) who have not failed from a competing event. Note that this risk set will also contain those who are truly censored (“Waiting”=0).

Note I omitted simple modeling for brevity.

Build a Stepwise Model

```
ods graphics on;
ods pdf file="C:/Stepwise Model.pdf" startpage=no;
proc phreg data=a concordance=harrell;
class dz(ref='LT') gender(ref='F');
model wl_months*_status(0,1,2) = dz gender init_age /rl selection=stepwise;
title 'Stepwise Selection Model';
run;
ods pdf close;
ods graphics off;
```

Step 1. Effect DZ is entered. The model contains the following effects:

DZ

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	23292.887	23222.798
AIC	23292.887	23224.798
SBC	23292.887	23229.986

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	70.0890	1	<.0001
Score	88.0354	1	<.0001
Wald	84.1515	1	<.0001

Step 2. Effect gender is entered. The model contains the following effects:

DZ gender

Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics		
Criterion	Without Covariates	With Covariates
-2 LOG L	23292.887	23176.595
AIC	23292.887	23180.595
SBC	23292.887	23190.972

Testing Global Null Hypothesis: BETA=0			
Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	116.2922	2	<.0001
Score	136.2905	2	<.0001
Wald	131.6507	2	<.0001

Note: No (additional) effects met the 0.05 level for entry into the model.

Note that age (int_age) did not meet criteria for inclusion into the model.

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
DZ	1	71.4746	<.0001
gender	1	47.2717	<.0001

Analysis of Maximum Likelihood Estimates										
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Label
DZ	SLK	1	0.69010	0.08163	71.4746	<.0001	1.994	1.699	2.340	DZ SLK
gender	M	1	-0.38350	0.05578	47.2717	<.0001	0.681	0.611	0.760	TCR RECIPIENT GENDER M

Summary of Stepwise Selection								
Step	Effect Entered	Effect Removed	DF	Number In	Score Chi-Square	Wald Chi-Square	Pr > ChiSq	Effect Label
1	DZ		1	1	88.0354		<.0001	
2	gender		1	2	47.8351		<.0001	TCR RECIPIENT GENDER

Harrell's Concordance Statistic					
		Comparable Pairs			
Source	Estimate	Concordance	Discordance	Tied in Predictor	Tied in Time
Model	0.5782	3797009	2171593	4421343	87707

Type 3 Tests which I use to assess overall effect of a covariate, it is more informative if you have more than 2 levels of a categorical covariate. If that were the case, this test will tell us if there are any differences in event rates across any of the levels of the covariate. Since DZ group and gender only have 2 levels, we can move on to the analysis of MLEs to get hazard ratios and CIs. DZ group ‘SLK’ has HR = 1.99, indicating the instantaneous hazard of death is 99% higher compared to ‘LT’. Gender ‘M’ has a lower hazard compared to ‘F’. In the output for Harrell’s C, we can use this as a measure of model fit, where larger value of the estimate indicates our model can correctly predict survival, when in a pair, the subject with the better observed survival time will also have the better predicted survival time. We can compare to a multiple model and see if this estimate improves.

Build a Multiple Model and Compare to Stepwise

```
ods graphics on;
ods pdf file="C:/multiplemod.pdf" startpage=no;
proc phreg data=a concordance=harrell;
class dz(ref='LT') gender(ref='F');
model wl_months*_status(0,1,2) = dz gender init_age /rl;
title 'Compare Full Model to Stepwise Model';
run;
ods pdf close;
ods graphics off;
```

The PHREG Procedure

Type 3 Tests			
Effect	DF	Wald Chi-Square	Pr > ChiSq
DZ	1	71.6726	<.0001
gender	1	46.7966	<.0001
init_age	1	0.6583	0.4172

Analysis of Maximum Likelihood Estimates									
Parameter		DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits		Label
DZ	SLK	1	0.69115	0.08164	71.6726	<.0001	1.996	1.701	2.342	DZ SLK
gender	M	1	-0.38182	0.05582	46.7966	<.0001	0.683	0.612	0.762	TCR RECIPIENT GENDER M
init_age		1	0.00646	0.00796	0.6583	0.4172	1.006	0.991	1.022	CALCULATED AGE AT LISTING

Harrell's Concordance Statistic					
		Comparable Pairs			
Source	Estimate	Concordance	Discordance	Tied in Predictor	Tied in Time
Model	0.5796	5833199	4179531	377215	87707

The output from the multiple model does not indicate better model performance compared to the stepwise model. Including age (init_age), which is not significant (95%CI includes null value 1), did not improve Harrell's C.