Survival Estimation to Cox Proportional Hazard Regression Models with Time-varying Coefficients

Abstract

Cox proportional hazard model is one of the most used statistical methods in survival analysis, and is highly relied on the proportional hazards (PH) assumption - the hazard ratios should be constant. However, the proportional-hazards assumption of constant hazard ratios is frequently violated in the data of clinical trials. An extension to the Cox PH model with time-varying covariates was adopted when proportionality assumption are violated. After checking the assumption violation to Cox PH model, our practice here is to offer estimation to time-varying coefficient Cox PH model and have them implemented in SAS.

Keywords: time-varying covariates, Cox PH model, survival estimation, SAS

1. Introduction

Survival analysis, or time-to-event analysis, can be conducted to explore the occurrence of events to be interested, such as death, or progression of disease, etc, since the intervention to population of subjects. Survival analysis considers not only about events but also time that events happen. Especially in tumor analysis, if the observation time is long enough, all patients will tend to have the same event—death. Only by considering both events and time can reveal the truth. Survival analysis includes the description of the survival progress, the comparison of survival progress and analysis of the survival time influencing factors.

Cox proportional hazard model (Cox, 1972) is one of the most widely applied statistical analysis methods to explore the relationship between survival explanatory variables and outcomes. The assumption to Cox PH model is the hazard ration of a group to a baseline group is assumed constant through the time. The assumption may be violated and will get misleading effect estimates and even wrong conclusion if time-varying covariates are included into the model without appropriate modeling. Furthermore, the hazard of the event from any group is a constant multiple of the hazard in any other. The hazards for groups should be proportional and cannot cross or diverge. Strong violations of the proportional hazards assumption can have detrimental effects on the validity and efficiency of the partial likelihood inference. Therefore, checking to the proportionality of the hazard should be conducted prior to the survival analysis by a Cox PH model.

The paper is organized as below. In the section 2 will introduce the Cox PH model with time-varying covariates. Section 3 will present the methods in SAS used to validate the assumptions to Cox PH model. In the last section, based on the example data Stanford heart transplant program (Crowley, J. and Hu, M., 1977), the survival estimation to Cox PH model will be described.

2. Cox PH model with time varying coefficients

In Cox PH model

$$\mathrm{h}_1(\mathrm{t}) = \mathrm{h}_0(\mathrm{t})e^{\beta Z} \tag{2.1}$$

the hazard function $h(t)$ is dependent on these *p* covariates $x_1, x_2, \ldots, x_p$. And the interpretation of β can be explained by the log hazard ratio, which means the hazard ratio for two treatment group of the variable Z from studies is

$$HR_1 = \frac{\mathrm{h}_0(\mathrm{t})\mathrm{e}^{\beta Z_1}}{\mathrm{h}_0(\mathrm{t})\mathrm{e}^{\beta Z_2}} = e^{\beta(Z_1 - Z_2)} \tag{2.2}$$

Since survival data occur over time, some important covariates considered in the model may also change over time. And this will violate the assumptions of that the hazards for groups should be proportional and cannot cross or diverge (Allison, 2010) .

To address such issue, time dependent covariates are introduced into the model, which is

$$\mathrm{h}_2(\mathrm{t}) = \mathrm{h}_0(\mathrm{t})e^{\beta_1 Z_1 + \beta_2 X_1(t)} \tag{2.1'}$$

and the corresponding hazard ration is updated as below

$$HR_2 = \frac{\mathrm{h}_0(\mathrm{t})e^{\beta_1 Z_1 + \beta_2 X_1(t)}}{\mathrm{h}_0(\mathrm{t})e^{\beta_1 Z_1 + \beta_2 X_2(t)}} = e^{\beta_1(Z_1 - Z_2) + \beta_2(X_1(t) - X_2(t))} \tag{2.2}$$

3. Assessment of proportionality assumptions

Some methods (Xue X.N., Xie, X.H., et al, 2013) (Xue Y, , Schifano ED., 2017) have been proposed to assess the violation to the assumption of Cox PH model, such as graphical methods: Log Cumulative Hazard plot, Schoenfeld Partial Residuals plot, cumulative martingale residual plot and Standardized Sco-re Process plot, and some other statistical testing methods as below.

3.1 Likelihood ratio test

Cox has proposed to have time-dependent covariates included into the Cox PH model (Cox, 1972), and likelihood ratio test is used to test if the time-dependent covariates will contribute to the model. The corresponding test statistics are:

$$LR = -2\log\big(h_1(t)\big) - \big(-2\log\big(h_2(t)\big)\big)$$

which is follow a chi-square distribution with freedom of $p = df_{h_1(t)} - df_{h_2(t)}$. And the method can be implemented with procedure of PHREG to fit the null and alternative models.

3.2 Schoenfeld residuals test and plot

Schoenfeld residual (Grambsch, P. M. , Therneau, T.M., 1994) to test time-dependent variables was introduced to test the time-dependent covariates. This method can be implemented with code as below.

```
proc phreg data = stan;
    class treatment strata_factors subseq_att;
    **** covariate to sub sequent therapies is to test for time dependent;
    model overall_survival*cnsr(1) = treatment strata_factors subseq_att;
    output out = resid ressch = sch_subseq_att;
run;

proc rank data = resid out = resrank ties = mean;
       var overall_survival;
      ranks overall_survival_rank;
run;

proc reg data = resrank ;
       model overall_survival_rank = sch_subseq_att;
run;
```

With information from Output window below, the PH assumption is violated at the significant level of$p = 0.01$.

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	Intercept	1	42.54662	2.94244	14.46	<.0001
schtrans	Schoenfeld Residual for trans	1	19.07863	7.19523	2.65	0.0098

There is another method with scaled Schoenfeld residual (Lin DY, Wei LJ, and Ying Z, 1993)to test PH assumption, which can be implemented with SAS program below.

```
proc phreg data = stan;
    class treatment strata_factors subseq_att;
    **** covariate to sub sequent therapies is to test for time dependent;
    model overall_survival*cnsr(1) = treatment strata_factors subseq_att;
    assess ph/resample;
run;
```

4. Estimation to time-varying Cox PH model

Time-dependent variables can be used to model the effects of subjects transferring from one treatment group to another. Example data for analysis is the Stanford heart transplant program (Crowley, J. and Hu, M., 1977). In the data, 103 patients are accepted if physicians judge them suitable for heart transplant, and 89 patients have received transplant. In the study, when a donor becomes available, physicians choose transplant recipients according to various medical criteria. A patient's status can be changed during the study from waiting for a transplant to being a transplant recipient. Transplant status can be defined by the time-dependent covariate function $z = z(t)$ as

$$z(t) = \begin{cases} 0, if\ the\ patient\ has\ not\ received\ trasplant\ at\ the\ time\ t \\ 1\ , if\ the\ patient\ has\ received\ trasplant\ at\ the\ time\ t \end{cases}$$

In the data, time-dependent variable transplant status takes value 1 or 0 at time (measured from the date of acceptance), depending on whether or not the patient has received a transplant at that time.

The following statements fit this model:

```
proc phreg data= Heart;
   model Days*Status(0)= XStatus Acc_Age / ties = EFRON;
   if (WaitDays = . or Days < WaitDays) then XStatus=0;
   else  XStatus= 1;
run;
```

With analysis output to this model, transplantation, which is assumed to be associated to the decrease of risk, is not significant (p = .9893), and the age to accept transplant is significant in the model (p=.0339).

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
XStatus	1	-0.00418	0.31208	0.0002	0.9893	0.996
ageaccpt	1	0.03074	0.01450	4.4975	0.0339	1.031

If we have the time dependent variable LOG(Age accept transplant + survival time) included into the model, the corresponding program is as below.

```
proc phreg data= Heart;
   model Days*Status(0)= XStatus Acc_Age LogAge / ties = EFRON;
   if (WaitDays = . or Days < WaitDays) then XStatus=0;
   else  XStatus= 1;
   LogAge = LOG(Acc_Age + Days);
run;
```

And the output is as below.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
XStatus	1	0.04010	0.29365	0.0186	0.8914	1.041
surg	1	-0.75871	0.35939	4.4569	0.0348	0.468
LogAge	1	1.57967	1.02483	2.3759	0.1232	4.853

In the example data, there are some other time-dependent covariates - waiting time until transplant, three measures of tissue matching and days since transplant, which is changed at unpredictable times. We will consider the case of that the time-dependent covariates change in time intervals.

```
proc phreg data= Heart;
    model Days * Status(0) = Surg Acc_Age XStatus
                             mltd m2td m3td waittd dottd /
```

```
                                ties = EFRON;
    if wait > Days OR wait =. then XStatus=0;
    else XStatus = 1;
    if XStatus = 1 then do;
       mltd = m1;
       m2td = m2;
       m3td = m3;
       waittd = wait;
       dottd = dot;
    end;
    else do;
       mltd = 0;
       m2td = 0;
       m3td = 0;
       waittd = 0;
       dottd = 0;
    end;
run;
```

the corresponding result for each period is as below.

Analysis of Maximum Likelihood Estimates

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio
surg	1	-0.17369	0.39218	0.1961	0.6579	0.841
ageaccpt	1	0.03859	0.01531	6.3554	0.0117	1.039
XStatus	1	0.48582	1.83505	0.0701	0.7912	1.626
mltd	0	0	.	.	.	.
m2td	1	1.48677	1.77539	0.7013	0.4023	4.423
m3td	1	-0.21451	0.19436	1.2181	0.2697	0.807
waittd	1	-1.48506	1.77537	0.6997	0.4029	0.226
dottd	1	-0.00455	0.0007896	33.1524	<.0001	0.995

5. Conclusion

In the paper, we have introduced the Cox PH model with time-dependent covariates. For the Cox PH model, the PH assumption is the basis for the conducting of such model. But in some case, the covariates may be time-dependent. The methods to testing the assumption to Cox PH model is introduced, and the corresponding SAS programs are also exhibited for some example. In the model, methods to estimate the model for the case of that the covariates are time-dependent.

Bibliography

Allison, P. (2010). *Survival Analysis Using SAS: A Practical Guide* (2nd edition ed.). Cary.: SAS Publishing.

Cox, D. (1972). Regression Models and Life Tables. *Journal of the Royal Statistical Society B, 34*(2), 187-220.

Crowley, J. and Hu, M. (1977). Covariance Analysis of Heart Transplant Survival Dat.

Journal of the American Statistical Association, 72, 27–36.

Grambsch, P. M. , Therneau, T.M. (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika, 81*, 515-526.

Lin DY, Wei LJ, and Ying Z. (1993). Checking the Cox model with cumulative sums of martingalebased. *Biometrika, 80*, 557-572.

Xue X.N., Xie, X.H., et al. (2013). Testing the proportional hazards assumption in case-cohort analysis. *BMC Medical Research Methodology, 13*, 88.

Xue Y, , Schifano ED. (2017). Diagnostics for the Cox model. *Communications for Statistical Applications and Methods, 24*, 583-604.