A SIMULATION STUDY OF ESTIMATORS IN STRATIFIED PROPORTIONAL HAZARDS MODELS

Hsi-Wen Liao
Pharmaceutical Research Associates, Inc. , Shrewsbury, NJ, U.S.A.

Abstract: It is common for large population-based surveys to select a sample from a population using a complex design. A simulation study was conducted to compare the estimates from the stratified proportional hazards model with the weighted estimates of the Binder method, when a stratified random sample of the population is used. The SAS PHREG procedure performs regression analysis of survival data based on the proportional hazards model while the SUDAAN SURVIVAL procedure estimates the parameters using the Binder method. Both procedures were used in this study to give estimates of regression parameters based on the two different methods of interest. This study shows that the estimates form the stratified proportional hazard model perform well under the assumptions for stratified proportional hazard models. The estimates from the Binder method show a tendency to underestimate the true value and gave the estimates from the unstratified finite population.

1. INTRODUCTION

The analysis of survival data requires special techniques because it involves incomplete observation of the failure times. Through a modeling approach to the analysis of survival data, we can explore the relationship between the survival time of an individual and the explanatory variables. One frequently used model for survival data is the proportional hazards model, which was proposed by Cox(1972) and is widely known as the Cox regression model. One of the reasons that the model is so popular is because the unknown parameter β can be estimated by partial likelihood without putting a structure on baseline hazard. This model requires that the hazards between any two individuals are proportional across time. To use the partial likelihood for inference, the failure times are required to be independent.

In a survey study, the design parameters for the survey are often related to the hazard function but are not fitted in the model. On some other occasions the independence assumption might be violated. Sometimes correlations exist within each level of nesting. These could cause biases and affect variances of the parameter estimates. Care must be exercised when trying to make inference for the proportional hazards model when the assumptions are violated. Lin&Wei(1989) proposed a method for estimating the covariance matrix of the estimated parameters, when the model was misspecified. Binder(1992) extended the method of Lin&Wei(1989) by considering unequal probability sampling and correlations among sampling units. This method provides a weighted estimator for the β coefficient in the Cox model which is appropriate for use in complex sampling designs. It is common for large population-based surveys to select a sample from a population using a complex design. In this paper we conduct a simulation study to compare the estimates from stratified proportional hazards model with the weighted estimates of Binder method, when a stratified random sample of the population is obtained.

2. DESCRIPTION OF MODELS

2.1 Proportional Hazards Models

Let T be a nonnegative random variable representing the failure time of an individual in the population. The survival distribution function of T is expressed as

$$S(t) = \Pr(T \geq t).$$

An alternative way of expressing the distribution of the failure time T is through its hazard function $\lambda(t)$ which specifies the instantaneous failure rate at t. If T is a continuous random variable, $\lambda(t)$ is expressed as

$$\lambda(t) = \lim_{\Delta t \to 0^+} \frac{\Pr(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)}$$

where f(t) is the probability density function of T.

The proportional hazards model assumes no particular form of probability distribution for the survival times with hazard function given by

$$\lambda(t) = \lambda_0(t) \exp(X\beta) \qquad (2.1.1)$$

where X is the vector of explanatory variables, and $\lambda_0(t)$ is the hazard function at X=0 which is an unspecified function of time t. β can be estimated by the partial likelihood method. Let the observed follow up time of the ith individual be t_i with corresponding covariates X_i (i = 1, ….., K, K is the number of individuals), the conditional probability for ith individual failing at t_i given that the individual is from the risk set $R(t_i)$ (i.e., $R(t_i)$ = { j: $t_j >= t_i$ } is :

$$\frac{\lambda_0(t_i)\exp(X_i\beta)}{\sum_{l\in R(t_i)}\lambda_0(t_i)\exp(X_l\beta)} \qquad (2.1.2)$$

Assume there are S failures, the partial likelihood function is then:

$$\prod_{i=1}^{S}\frac{\exp(X_i\beta)}{\sum_{l\in R(t_i)}\exp(X_l\beta)} \qquad (2.1.3)$$

Let δ_i denote the individual failing at time t_i, i.e. $\delta_i=1$ if the ith unit is an observed failure, and $\delta_i=0$ if the ith unit is censored; The above partial likelihood function (2.1.3) can be expressed as:

$$L(\beta)=\prod_{i=1}^{K}\left[\frac{\exp(X_i\beta)}{\sum_{j=1}^{K}Y_j(t)\exp(X_j\beta)}\right]^{\delta_i} \qquad (2.1.4)$$

where $Y_j(t)$= 1 when $t \leq t_j$, otherwise $Y_j(t)$=0

For a sample of size K, the log partial likelihood for expression (2.1.4) is

$$\text{logL}(\beta) = \sum_{i=1}^{K}\delta_i\left[\frac{\exp(X_i\beta)}{\sum_{j=1}^{K}Y_j(t)\exp(X_j\beta)}\right] \qquad (2.1.5)$$

The maximum partial likelihood estimate of β can be obtained as a solution to the equation

$$\frac{\partial LogL(\beta)}{\partial\beta}=0$$

$$\Rightarrow\sum_{i=1}^{K}\delta_i\left[X_i-\frac{\sum_{j=1}^{K}Y_j(t)X_j\exp(X_j\beta)}{\sum_{j=1}^{K}Y_j(t)\exp(X_j\beta)}\right]=0 \qquad (2.1.6)$$

2.2 Finite Population Model

For a finite population of size N, the partial likelihood function is

$$L(\beta)=\prod_{i=1}^{N}\left[\frac{\exp(X_i\beta)}{\sum_{j=1}^{N}Y_j(t)\exp(X_j\beta)}\right]^{\delta_i} \qquad (2.2.1)$$

The maximum partial likelihood estimator (MPLE) of β could be obtained by the solution of

$$\frac{\partial LogL(\beta)}{\partial\beta}=0$$

$$\Rightarrow\sum_{i=1}^{N}\delta_i\left[X_i-\frac{\sum_{j=1}^{N}Y_j(t)X_j\exp(X_j\beta)}{\sum_{j=1}^{N}Y_j(t)\exp(X_j\beta)}\right]=0 \qquad (2.2.2)$$

The solution to the equation is the finite population value of β and we wish to estimate it from a sample size n.

$$\sum_{i=1}^{n}\delta_i\left[x_i-\frac{\sum_{j=1}^{n}Y_j(t)X_j\exp(X_j\hat{\beta})}{\sum_{j=1}^{n}Y_j(t)\exp(X_j\hat{\beta})}\right]=0 \qquad (2.2.3)$$

If the design is a complex design, a weighted version of (2.2.3) is required. The weights are constructed so that the weighted estimator β is unbiased for the solution to (2.2.2). The estimating equation for β as given by Binder(1992) is:

$$\sum_{i=1}^{n} \omega_i \delta_i \left[x_i - \frac{\sum_{j=1}^{n} \omega_i Y_j(t) X_j \exp(X_j \hat{\beta})}{\sum_{j=1}^{n} \omega_i Y_j(t) \exp(X_j \hat{\beta})} \right] = 0 , \quad (2.2.4)$$

where ω_i is the weight, reflecting the sampling design. The estimator β which is the solution of the expression(2.2.4) is referred to as the pseudo-maximum likelihood estimator (Skinner, 1989)

2.3 Inclusion of Strata

The proportional hazards model requires that for two covariate sets X_1 and X_2, the relative hazard of the two individuals are constant over time. Sometimes the different levels of some important factors produce hazard functions which differ markedly from proportionality. To accommodate such a study, Kalbfleisch and Prentice (1980) extended the model(2.1.1) by including the stratum effect into the baseline hazards λ_{0j}, j= 1, …..s. The baseline hazards are allowed to be arbitrary and unrelated for s strata. The hazard function for an individual in the jth stratum is

$$\lambda_j(t) = \lambda_{0j}(t) \exp (X\beta) \quad (2.3.1)$$

The stratum effects are expressed in the baseline hazards instead of the regression coefficient and the effects of X are proportional within strata. In this study, we will use a common regression coefficient through all strata. The partial likelihood of β is the product of terms as in (2.1.4) for each stratum. The Newton-Raphson technique is used to estimate β and the maximization of the likelihood is easily accomplished on a computer.

We are considering a stratified random sample and the weight (ω_i) for the ith individual in the jth stratum for the Binder model is

$$\omega_{ij} = \frac{N_j}{n_j}$$

where N_j is the number of individuals in the jth stratum of the finite population and n_j is the number of individuals in the jth stratum of the random sample.

2.4 Software for Implementation of The Method

The SAS PHREG procedure performs regression analysis of survival data based on the proportional hazards model. The SUDAAN SURVIVAL procedure estimates the parameters using Binder method. Both of them assume a parametric form for the effects of the explanatory variables, but allow an unspecified form for the underlying survival function. The two procedures were used in this study to give estimates of regression parameters β based on the two different methods in which we are interested.

3. SIMULATION

3.1 Description of Methods

Our interest in this study is to compare the estimates from stratified proportional hazards model with the weighted estimates of the Binder method. We performed a simulation study based on the stratified proportional hazards assumption. The survival time of the ith person in the jth strata of a population is assumed to follow its own hazard function, $\lambda_{ih}(t)$ which expressed as :

$$\lambda_{ij}(t_{ij}) = \lambda_{0j}(t_{ij}) \exp (x_{ij}\beta)$$

where $\lambda_{0j}(t)$ is the baseline hazard function for the jth strata, X_{ij} is the explanatory variable for the ith person in the jth strata, and β is the unknown regression parameter. Assume the effect of X on the hazard is constant within each stratum. (i.e., the hazard ratio for one unit change of x is $\lambda_{0j}\exp(\beta)$ in jth stratum) X_{ij} is obtained from independent uniform (0,1) distribution. The survival distribution function(also known as survivor function) is expressed as

$$S(t) = \Pr(T_{ij} \geq t)$$

We can specify the distribution of failure time for the ith person in the jth stratum T_{ij} through its own hazard function. The failure time T_{ij} for this study were obtained from an exponential distribution ($\lambda_j\exp(X_{ij}\beta)$). Two strata were used in this study.

In the first part of the simulation study, the total sample size is fixed and chosen to be 1000, with 750 in stratum1 and 250 in stratum2. The failuretime of the ith person in the 1st stratum was generated from an exponential distribution ($\exp(X_{i1}\beta)$) and the ith person in the 2nd stratum was generated from an exponential distribution ($2\exp(X_{i2}\beta)$). We fitted 3 models with 3 different β coefficients (β = 0, 1, ln(2)). We did 500 replications for each model using two methods: the Cox proportional hazards model (performed by SAS PHREG procedure) and the Binder model (performed by SUDAAN SURVIVAL procedure). In this part of simulation, we did not fix the original finite population. We assumed that equal stratum size with 1000 individuals composed the finite population. We also generated the failuretime with 500 in stratum1 and 500 in stratum2 from the two independent distribution so it's more representative for the original population we assumed earlier.

Then we assumed the baseline hazard ratio did not change with different strata and generated the failure time with 750 in stratum1 and 250 in stratum2 from exponential distribution ($\exp(X_{ij}\beta)$). We then did 500 times of simulation in each of the two models with $\beta = 0$ and 1.

In the second part, we generated the finite population with the parameter fixed and then only drew samples from the population to fit the models. The goal is to see how Cox & Binder methods perform in estimating the β coefficient in the population. The size of the finite population is 10000 with 5000 in each stratum. The total sample size was fixed to be 1000 and three sampling strategies were used. We used the equal allocation with 500 in each stratum and unequal allocation with 750 in stratum1 and 250 in stratum2 and sample selection with 250 in stratum1 and 750 in stratum2. The baseline hazard ratio is chosen to be one for the first stratum and two for the second stratum. So the failure time for the ith individual in the 1^{st} stratum is obtained from the exponential distribution ($\exp(X_{i1}\beta)$) and for the ith individual in the 2^{nd} stratum is obtained from the exponential distribution ($2\exp(X_{i2}\beta)$). We considered when under the null hypothesis that the hazard ratio doesn't change with the covariates (X_{ij}) and when $\beta = 1$.

Table1. Expected value, standard deviation and MSE of parameter estimates based on 500 simulations. Sample size (stratum1=750, stratum2=250)

True	Models			
	Cox		Binder	
β	$\hat{\beta}$ (sd)	MSE	$\hat{\beta}$ (sd)	MSE
0	0.002(0.007)	0.026	-0.001(0.007)	0.027
1	1.013(0.010)	0.054	0.843(0.011)	0.056
Ln2(=0.6931)	0.708(0.009)	0.040	0.589(0.010)	0.040

* $\lambda_{01} = 1$, $\lambda_{02} = 2$

Table2. Expected value, standard deviation and MSE of parameter estimates based on 500 simulations. Sample size (stratum1=750, stratum2=250)

True	Models			
	Cox		Binder	
β	$\hat{\beta}$ (sd)	MSE	$\hat{\beta}$ (sd)	MSE
0	0.001(0.007)	0.024	-0.001(0.008)	0.035
1	1.008(0.010)	0.045	1.008(0.011)	0.063

* $\lambda_{01} = 1$, $\lambda_{02} = 1$

Table3. Expected value, standard deviation and MSE of parameter estimates based on 500 simulations. Sample size (stratum1=500, stratum2=500)

True	Models			
	Cox		Binder	
β	$\hat{\beta}$ (sd)	MSE	$\hat{\beta}$ (sd)	MSE
1	1.024(0.012)	0.072	0.851(0.012)	0.073

* $\lambda_{01} = 1$, $\lambda_{02} = 2$

Table4. Compare the β estimates of finite population(stratified-, unstratified-) with the 3 sampling selection strategies used Cox and Binder method based on the average of 500 simulations

β	$\hat{\beta}$ for the finite population		Sample selection (1st stratum, 2nd stratum)					
			1.(750,250)		2.(250,750)		3.(500,500)	
	Stratified	unstratified	Cox	Binder	Cox	Binder	Cox	Binder
1	0.929	0.766	0.919	0.778	0.995	0.785	0.948	0.780
0	-0.035	-0.040	-0.032	-0.032	-0.008	-0.014	-0.015	-0.019

The strata statement in SAS PHREG was used to perform a stratified analysis. We assumed simple random sample selection and did not use the weighted method in PHREG. The Breslow likelihood function and Newton-Raphson Method were applied to give β estimate. A stratified random sampling with replacement was chosen to be the sample design with weighted statement in SUDAAN. The partial likelihood function on the failure and censoring time is weighted by the sampling weight. And the β estimated by SUDAAN is referred to as the pseudo-maximum likelihood estimator. We fixed the censored time as .75.

3.2 Results for The Simulation

In Table 1, we give the average values of the parameter estimates over the 500 simulations. The standard deviation of the sampling distribution of the parameter estimator is also given. Table 1 shows that under the situation when β=0, the β estimates from stratified proportional hazards model and Binder methods both performed well and give the estimates close to the true value. When β is not 0, Binder method is less likely to give the correct estimates. The stratified proportional hazards model seems to give consistent and good estimates under the 3 different conditions.

In Table 2, we chose the same baseline hazard function in both strata and both methods performed well under null hypothesis and alternative hypothesis. Unequal allocation was used in both Table 1 and Table 2 with 750 in stratum1 and 250 in stratum2. We then used equal allocation with 500 in each strata. Two different baseline hazard functions were chosen with true β =1. Again, the Binder method gave a lower estimates when stratified proportional hazards model still gave us estimates close to the true value. In this part of the simulation, generally the Binder method gives lower estimates and only performs well when the baseline hazards are the same in both strata or when β=0. The β estimates from stratified proportional hazards model are always close to the true value.

We also used the index of mean squared error (MSE) to evaluate the true and estimated hazard ratios, among 500 samples.

$$\sum_{i=1}^{500} \frac{(\hat{\beta}-\beta)^2}{500}$$

The MSE in Tables 1, 2 and 3 show that the Binder method gave larger deviations from the true values. The mean squared errors of estimated hazard ratios indicated that β estimator from the stratified proportional hazards model performed better than the one from the Binder method.

In the second part of this simulation study, individual survival function estimates from stratified proportional hazards model and Binder method are compared with the true survival function to show how well the estimates are. Table 4 gives the estimated parameter for the finite population with stratified and unstratified option. When β=0, both methods gave good estimates. We can see that the estimate from the stratified proportional hazards model is closer to population parameter. In contrast, the estimates from Binder method give the estimates closer to the population parameter when we ignore the stratum effect.

3.3 Interpretation

Due to the results of this simulation study, we suggest that the β estimate from the Binder method is inferior to the one from stratified proportional hazards model if the stratified proportional hazards model is appropriate. The estimates from the Binder method exhibit larger biases in this case. Generally, the Binder method gives a smaller estimate than the true value while the stratified proportional hazards model gives almost the same estimate as the true value. The Binder method perform as well as stratified proportional hazards model only when we assume the same baseline hazard function in both strata or under the situation when β = 0.

For the finite population, the Binder method estimates the β for the finite population without considering the stratum effect while the stratified proportional hazards model gives the estimate assuming stratification.

Overall, based on this simulation study, the stratified proportional hazards model shows a better estimation on β coefficient under several different situations. The Binder method has larger bias of β estimates. This is possibly due to the model we used for Binder method ignored the stratum effect.

4. EXAMPLE

The Lipid Research Clinics (LRC) Program was created in 1971. This program include two major sets of studies, the first one is the Population studies and the second one is the Coronary Primary Prevention Trial. The Prevalence Study is one of the Population Studies for LRC program. The Prevalence Study consisted of two screens designed as Visit 1 and Visit 2. A sample for further study at Visit 2 was selected from all attendees of Visit 1. The Visit 2 sample was drawn from those individuals whose cholesterol and triglyceride levels were properly measured at Visit 1. All participants were divided into three strata using age-specific cutpoints for cholesterol and triglyceride. The variable LIPSTRAT is used to indicate the strata. If both cholesterol and triglyceride were in the normal range, the value of LIPSTRAT was set to 1. Extreme values of cholesterol were set to 2, and of triglyceride to 3.

We used the dataset from the LRC Prevalence Study to perform an example and the results are shown in Table 5. The dataset was stratified by the variable LIPSTRAT and has three strata. The covariates we chose in the first model are baseline age, cholesterol reading in Visit 1, and smoking status. For the second model, we chosen the baseline age, cholesterol reading in Visit 2, and smoking status. Failure times from both models were censored by mortality. The estimated risk ratios of baseline age(RR≈ 1.10) and Cholesterol Reading in V1&V2(RR≈ 1.04) shows weak association with mortality when the smoking status(RR≈ 1.96) is significantly associated(P<0.001) with mortality.

In the two models, the parameter estimates are similar from both methods. This might be due to that LIPSTRAT is not related to the hazards. Thus even when we considered the stratum effect in the stratified proportional hazards model, it wouldn't give different estimates from Binder method without considering stratum effect.

Table5. Compare β estimates and its standard deviation of the LRC Prevalence Study dataset using Cox and Binder method.

		β estimates (sd)	
Group	Variables	Cox	Binder
1	Baseline Age	0.097(0.003)	0.096(0.003)
	Cholesterol Reading at V1	0.001(0.001)	0.002(0.001)
	Smoking Status	0.677(0.068)	0.687(0.069)
2	Baseline Age	0.097(0.003)	0.096(0.003)
	Cholesterol Reading at V2	0.001(0.001)	0.002(0.001)
	Smoking Status	0.675(0.068)	0.687(0.069)

5. CONCLUSION

The main purpose of these simulations is to provide guidelines on the performance of the β estimates from two different methods under the assumptions for stratified proportional hazard model. This study has shown that the stratified proportional hazards model approach to estimate β coefficient performs well in a stratified proportional hazard model. The different allocations of sample did not affect the result. The β estimates from Binder model showed a tendency to underestimate the true value and gave the estimates for the unstratified finite population.

If the sample has been drawn from a population using a complex design with different hazards in different stratum, the Binder method would give us biased estimates, if the underlying model is the stratified proportional hazards model. Even if there is no stratum effect, the estimate from Binder method is still less efficient compared to the estimate from stratified proportional hazards model. The Binder method could be extended to adjust stratum effect and give unbiased β estimates. Further simulations could be conducted to compare the efficiency of β estimates from the stratified proportional hazards model with the Binder method with strata.

REFERENCES

Binder, D.A. (1992). Fitting Cox's Proportional Hazards Models from Survey Data, *Biometrika*, 79, 1, pp. 139-147.

Cox, D.R. (1972). Regression Models and Life Tables. *Journal of the Royal Statistical Society*, Series B34, 187-220.

Collet, D. (1994). Modeling Survival Data in Medical Research. *Chapman & Hall*.

Kalbfleisch, J.D. and Prentice, R.L. (1980). The Statistical Analysis of Failure Time Data, New York: *John Wiley & Sons*, *Inc*.

Lin, D.Y., and L.J. Wei (1989). The Robust Inference for the Cox Proportional Hazards Model. *Journal of American Statistician Association*, 84, pp. 1074-1078.

National Institutes of Health (1980). The Lipid Research Clinics Population Studies Data Book Volume 1 The Prevalence Study. *NIH Publication* No. 80-1527.

Research Triangle Institute (1996). SUDAAN User's Manual, Version 6.4, Second Edition. Research Triangle Park, NC: *Research Triangle Institute*.

SAS Institute Inc. (1992). SAS Technical Report p-229. Cary, NC: *SAS Institute Inc*.