Distribution Agreement

Signature:

_________________________________ _______________

Huaying Dong Date

Handling Tied Events in Cox Proportional Hazard Regression Modeling

By

Huaying Dong

Master of Public Health

Department of Biostatistics and Bioinformatics

____

Jose Binongo, Ph.D.

(Thesis Advisor)

____

George Cotsonis, MA.

(Thesis Co-Advisor)

Handling Tied Events in Cox Proportional Hazard Regression Modeling

By

Huaying Dong

B.S.

Shandong University

2017

Thesis Committee Chair: Jose Binongo, Ph.D.

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University
in partial fulfillment of the requirements for the degree of
Master of Public Health
in Biostatistics
2019

Abstracts

Handling Tied Events in Cox Proportional Hazard Regression Modeling

By Huaying Dong

Background: Tied survival times are quite common in real-life survival analysis as survival time is usually measured discretely. One of the assumptions of the Cox proportional hazard model is that there are no tied survival times. However, the violation of the assumption does not mean the Cox proportional hazard model should be discarded.

Application: There are four methods that have been developed to handle ties. The exact and discrete methods provide the gold standard, but they are computationally intensive especially when the percentage of ties is high. Early recognition of these computational difficulties led to the development of Breslow's and Efron's procedures. Breslow's approximation performs well when the percentage of tied observations is not too high; Efron's approximation, on the other hand, almost always gives results very close to the exact method, and it performs well even when the percentage of tied observations is high.

Discussion: In general, Breslow's approximation is recommended when the number of ties is not extensive. When the percentage of ties is high, Efron's approximation could be used as a good substitute of the exact method or discrete method. When computation time is not a concern and accuracy is required, the exact method or the discrete method is appropriate.

Key words: Cox proportional model, partial likelihood function, tied survival times

Handling Tied Events in Cox Proportional Hazard Regression Modeling

By

Huaying Dong

B.S.

Shandong University

2017

Thesis Committee Chair: Jose Binongo, Ph.D.

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Public Health

in Biostatistics

2019

Table of Contents

Chapter 1 Background

1.1 Partial Likelihood

In survival analysis, one of the topics of interest is the quantification of the heterogeneity that may exist among groups when the outcome is time to an event of interest. For example, researchers may be interested in the difference in survival rate between female and male smokers, or the difference in divorce rate between people who have at least a college degree and people who did not finish college. Cox [1972] proposed a proportional hazard model to address this problem and gave the hazard function below:

$$h(t|z_i) = h_0(t)\exp(z_i\beta) \tag{1}$$

where

z_i is the vector of covariates associated with the individual whose failure time is t_i,

$h(t|z_i)$ is the hazard rate at time t of an individual with covariate z_i,

$h_0(t)$ is the baseline hazard function which only depend on time t,

$\exp(z_i\beta)$ is the covariate-related function.

This model is called "proportional" since the hazard ratio is proportional:

$$\frac{h(t|z_i)}{h(t|z_j)} = \exp\left(z_i - z_j\right)\beta \tag{2}$$

and independent of time t. To estimate the covariates, instead of the full likelihood function, Cox derived an alternative expression, which he called "partial likelihood". Anderson [1993] showed that the estimation of the covariates in the partial likelihood function is the same as that in the full likelihood function. The expression of the partial likelihood is:

$$L(\beta) = \prod_{i=1}^{D} L_i = \prod_{i=1}^{D} \frac{\exp(z_i\beta)}{\sum_{j\in R_i}\exp(z_j\beta)} \tag{3}$$

where

$t_1 < t_2 < \cdots < t_D$ denote the ordered event times,

z_i is the vector of covariates associated with the individual whose failure time is t_i,

R_i is the risk set at time t_i.

It can be seen from the partial likelihood function that only the uncensored observations are taken into account; the censored observations do not affect the denominator.

1.2 Tied Event Times

In the presence of tied event times, an alternate formula for partial likelihood is needed for two reasons.

First, one of the assumptions that Cox made for the partial likelihood is that there are no tied event times among the observed survival times. In (3) it is clear that the ordering of events matters. If individuals i, j and k have the same survival time, it is not possible to tell which one happens at first, which one is second or third. For example, if i happens first, the risk set should include j and k. Thus the risk set depends on which one happens at first. This will in turn change the denominator of the partial likelihood.

The second reason is that the occurrence of ties will make the formula more complicated and will increase the computing time significantly. With the development of an integral representation of the likelihood, it is much easier to evaluate numerically (DeLong, Guirguis, and So, 1994). However, computation of the exact likelihood when large numbers of events occur at the same time can take a lot of computing time (Allison, 2010). Early recognition of these computational difficulties led to the development of approximate methods.

1.3 Approaches dealing with tied event times

The exact method (Kalbfleisch and Prentice, 2002; DeLong, Guirguis, and So 1994) and discrete method (Cox, 1972), regarding time as continuous and discrete, respectively, provide the best results, but these methods are computationally intensive, requiring more computing time. Breslow's approximation [1974] usually generates better results when the percentage of ties is not small. Efron's approximation [1977] generates better results even when the percentage of ties is high.

1.4 Software Options

1.4.1 SAS

When there are no ties, all the four methods result in the same likelihood and yield identical estimates. In SAS, the default, TIES=BRESLOW, is the most efficient method when there are no ties. When ties are not extensive, EFRON and BRESLOW methods provide satisfactory approximations to the EXACT method for the continuous time-scale model [Allison, 2010]. The EXACT method can take a considerable amount of computer resources. In general, Efron's approximation gives results that are much closer to the EXACT method results than Breslow's approximation does. If the time scale is genuinely discrete, the DISCRETE method should be used. The DISCRETE method is also required in the analysis of case-control studies when there is more than one case in a matched set [J. Mandrekar, 2004].

1.4.2 R

In coxph function, there are three possible choices for handling tied event times. Although almost all computer routines use the BRESLOW approximation as the default as it is the easiest to program, in R EFRON is the default option [Therneau, 2000]. Using the "exact partial likelihood" approach the Cox partial likelihood is equivalent to that for matched logistic regression. The clogit function uses the coxph code to do the fit [Therneau, 2000]. It is technically appropriate when the time scale is discrete and has only a few unique values, and some packages refer to this as the DISCRETE option. There is also an "exact marginal likelihood" due to Prentice [1973] which is not implemented here. The calculation of the exact partial likelihood is numerically intense. For example, if there are 300 subjects at risk at a certain timepoint, of which 30 had an event, then the code needs to compute sums over all C_{30}^{300} different possible subsets of size 30. Although there is an efficient recursive algorithm for this problem, the computation can be extremely long. With (start, stop) data it is much worse since the recursion needs to start a new for each unique start time [Therneau, 2000].

1.4.3 STATA

There are four options in STATA for handling tied event times in calculating the Cox partial likelihood: BRESLOW, EFRON, EXACTM, and EXACTP. If there are no ties in the datasets, the results are identical, no matter which option is selected [StataCorp, 2013]. The default option is the Breslow method since it is fast when there are not so many ties. But when there are so many ties in the dataset, Breslow's approximation will not be accurate since there are too many observations in the risk sets [StataCorp, 2013]. The Breslow method is an approximation of the exact marginal likelihood. Efron's approximation is a more accurate approximation of the exact marginal likelihood than Breslow's but the computational time is longer [StataCorp, 2013]. For exact methods we can use the EXACTP option (for the exact partial likelihood) or the EXACTM option (for the exact marginal likelihood) in the stcox or cox command. The exact methods are slower than Efron's approximation when the sample size is small, but the difference in the computation time diminishes when samples become larger [StataCorp, 2013].

1.5 Objectives and Significance of the Problem

In this thesis, we compare the performance of these four methods of handling ties, based on two real datasets, with different percentages of ties, in terms of the estimates of coefficients, standard error of coefficients, fit statistics and computation time. SAS Version 9.4 is used in all the analyses. By comparing the existing four methods, a better understanding of handling tied event times may be achieved, and appropriate choices may be made when facing different types of datasets. The goal is to improve accuracy and efficiency of the data analysis.

Chapter 2 Exact & Discrete Method

The mainly difference between exact and discrete method is the way time is treated – whether as continuous or discrete.

Cox [1972] presumed that the events really occurs at exactly the same time (time is really discrete). For example, when two or more events appear to happen at the same time, it is presumed that there is no underlying ordering. The exact method takes each event into account, thus it fits the model very well but it takes a rather long time as a result of permutations.

Kalbfleisch and Prentice's [2002] presume that time is truly continuous; that is, there is a true but unknown ordering for the tied event times. The assumption is that ties are merely the result of the imprecise measurement of time). This is a generalization of the discrete model that Cox came up. The partial likelihood function for discrete method is:

$$\mathrm{L}(\beta) = \prod_{i=1}^{D} \frac{\exp[\left(\sum_{j \in D_i} z_j\right)\beta]}{\sum_{q \in Q_i} \exp(s_q^* \beta)} \tag{4}$$

Where

Q_i denotes the set of all subsets of d_i individuals who could be selected from the risk set R_i. Q_i includes all the subsets of d_i failures at time $\mathrm{t_i}$.

$s_q^* = \sum_{j=1}^{d_i} z_{qj}$, which denotes the sum of covariate values for subset q.

Chapter 3 Approximation

3.1 Breslow's Approach

Breslow [1975] came up with an approximation of the partial likelihood, and it is commonly used as default in major statistical software, such as SAS and Stata. Breslow assumed that the underlying survival distribution is continuous, with the hazard h_i being constant in each interval (t_{i-1}, t_i). Moreover, he regarded the withdrawals or censored observations that lies within the interval (t_{i-1}, t_i) as occurring at timepoint t_i. In this way, he obtained the estimation of β and $h_0(t)$ simultaneously. This method multiplies the summation of all the discrete events at each time point and use it as the denominator rather than that in the previous partial likelihood formula. The partial likelihood function is:

$$\mathrm{L}(\beta) = \prod_{i=1}^{D} \frac{\exp[\left(\sum_{j \in D_i} z_j\right)\beta]}{[\sum_{j \in R_i} \exp(z_j \beta)]^{d_i}} \tag{5}$$

Where

z_j is the covariate for the jth individual,

d_i is the number of failures at t_i,

D_i is the set of all individuals who fails at time t_i.

Breslow states that when there are no ties, his likelihood function generates the same results for β, compared to those of Cox [1972] and Kalbfleisch and Prentice [1973], and when the ties exist, his method will be an approximation of those discrete likelihoods.

3.2 Efron's Approach

Efron [1977] puts forward an approximation which is more precise than Breslow's. He points out that real censored datasets are often discrete, where have failures lie between intervals rather than their exact times.

He also assumes that the hazard ratio is constant within each time interval and no changes other than those due to failures in risk set $R(t)$ occur within such interval. The partial likelihood function is:

$$\mathrm{L}(\beta) = \prod_{i=1}^{D} \frac{\exp[\left(\sum_{j\in D_i} z_j\right)\beta]}{\prod_{l=1}^{d_i}[\sum_{j\in R_i}\exp(z_j\beta) - \frac{l-1}{d_i}\sum_{j\in D_i}\exp(z_j\beta)]} \tag{6}$$

Where

z_j is the covariate for the jth individual,

d_i is the number of failures at t_i,

D_i is the set of all individuals who fails at time t_i.

Efron’s approximation is closer to the correct partial likelihood based on a discrete hazard model than Breslow’s likelihood. When the number of ties is small, Efron’s and Breslow’s likelihoods are quite close (K. Dietz, M. Gail, K. Krickeberg, J. Samet, A. Tsiatis). In many applied settings, there will be little or no practical difference between the estimators from the two approximations, so Breslow’s approximation is more commonly used (Hosmer, D. W., Lemeshow, S., May, S., 2008).

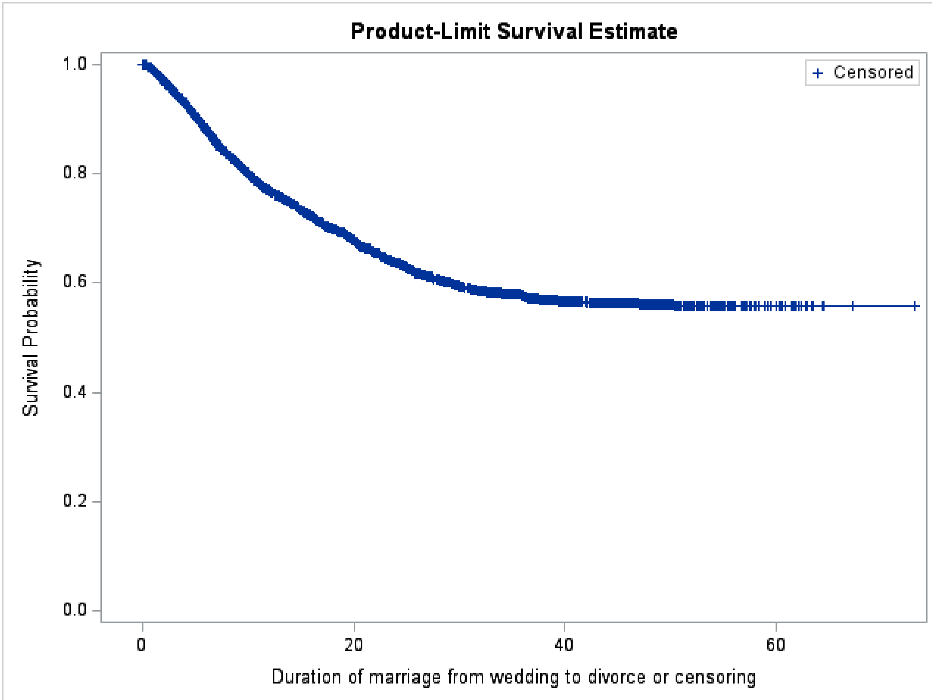
Chapter 4 Applications & Discussion

4.1 Divorce Dataset

The divorce dataset is based on a longitudinal survey conducted in the U.S.. The unit of observation is the couple and the event of interest is divorce, with interview and widowhood treated as censored events. There are three covariates: education of the husband and two indicators of the couple's ethnicity: whether the husband is black and whether the couple is mixed. The variables are: id (a couple of number), heduc (education of the husband, coded 0 when less than 12 years, 1 when 12 to 15 years and 2 when 16 or more years), heblack (coded 1 if the husband is black and 0 otherwise), mixed (coded 1 if the husband and wife have different ethnicity (defined as black or other), 0 otherwise), years (duration of marriage, from the date of wedding to divorce or censoring (due to widowhood or interview)), div (the failure indicator, coded 1 for divorce and 0 for censoring).

The dataset has 3771 couples, 1032 of which are failures (divorce). Of the event times, 24.5% are ties. The Kaplan-Meier curve looks well (Fig 1).

Fig 1 Kaplan Meier curve for the original divorce dataset



We apply the four methods using PROC PHREG in SAS. Table 1 shows that all of the four methods generate almost the same results. The percentage of ties is 24.5%, which is relatively small. Efron's approximation has the same results with the exact method, while Breslow's approximation seems to underestimate the results of the exact method. The discrete method tends to provide higher parameter estimates and standard errors than those of the exact method. The discrepancy is due largely to the fact that completely different models are being estimated. The exact method uses Cox proportional hazard model, while the discrete method uses the logit model. The logit coefficients will usually be larger [Allison 2010].

Table 2 shows that the hazard ratios and confidence intervals are all the same among the four methods for all of the covariates. Since all of the confidence intervals don't include 1, they are significant.

Table 3 shows the fit statistics of the four methods. Since Akaike's Information Criterion (AIC) is based on the Kullback-Leibler information measure of discrepancy between the true distribution of the response variable and the distribution specified by the model, better models are identified with smaller AIC. The Schwarz's Bayesian Criterion (SBC), also known as the Bayesian Information Criterion (BIC), is an increasing function of the model's residual sum of squares and the number of effects. Unexplained variations in the response variable and the number of effects increase the value of the SBC. As a result, a lower SBC implies either fewer explanatory variables, better fit, or both. SBC penalizes free parameters more strongly than AIC. Table 3 shows that the exact and discrete methods provide the better results than Breslow's and Efron's approximations when the percentage of ties is not high.

Table 4 shows the computation time for the four methods. Real Time is the actual, real world, time that the step takes to run and will be the same as if you timed it with a stopwatch. CPU Time is the amount of time the step utilises CPU resources. and we can see that Breslow's and discrete methods are faster than the exact and Efron's methods.

Table 1 Parameter estimate, standard error and p-value of covariates for the original divorce dataset

Method	Parameter Estimate for heduc	SE for heduc	p-value for heduc	Parameter Estimate for heblack	SE for heblack	p-value for heblack	Parameter Estimate for mixed	SE for mixed	p-value for mixed
Exact	0.09427	0.04718	0.0457	0.18376	0.07974	0.0212	0.22948	0.07929	0.0038
Breslow	0.09424	0.04718	0.0458	0.18373	0.07974	0.0212	0.22945	0.07929	0.0038
Efron	0.09427	0.04718	0.0457	0.18376	0.07974	0.0212	0.22948	0.07929	0.0038
Discrete	0.09425	0.04718	0.0457	0.18377	0.07975	0.0212	0.22948	0.07929	0.0038

Table 2 Hazard ratio for the original divorce dataset

Method	Exact	Breslow	Efron	Discrete
Hazard Ratio for heduc	1.099 (1.002, 1.205)	1.099 (1.002, 1.205)	1.099 (1.002, 1.205)	1.099 (1.002, 1.205)
Hazard Ratio for heblack	1.202 (1.028, 1.405)	1.202 (1.028, 1.405)	1.202 (1.028, 1.405)	1.202 (1.028, 1.405)
Hazard Ratio for mixed	1.258 (1.077, 1.469)	1.258 (1.077, 1.469)	1.258 (1.077, 1.469)	1.258 (1.077, 1.469)

Table 3 Fit statistics for the original divorce dataset

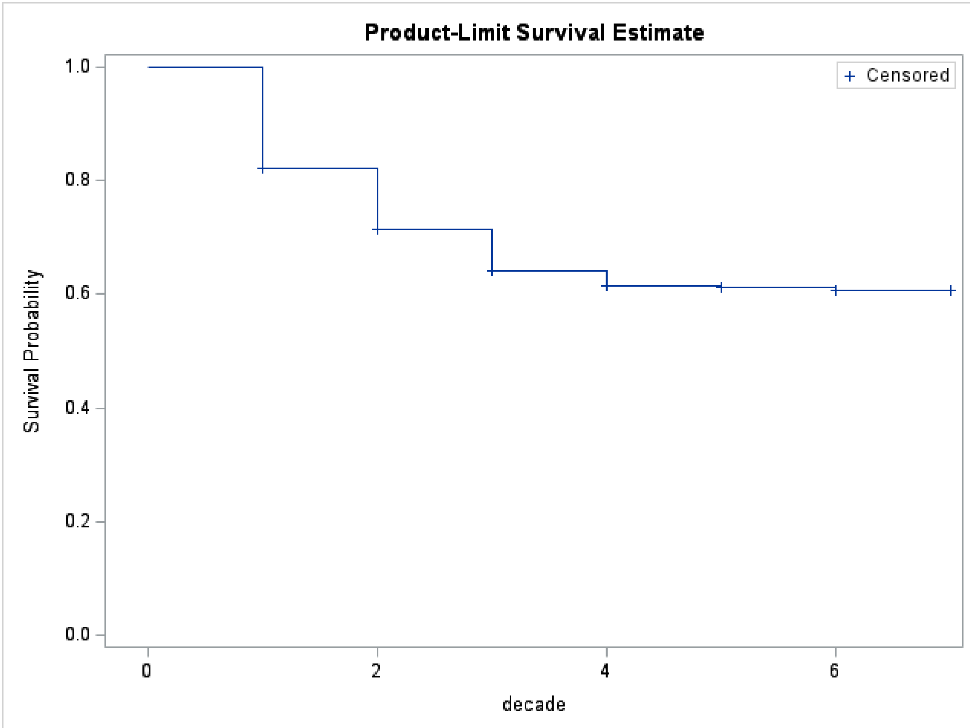
Method	Exact	Breslow	Efron	Discrete
-2LOG L	15428.346	15669.975	15669.803	15428.347
AIC	15434.346	15675.975	15675.803	15434.347
SBC	15449.164	15690.793	15690.621	15449.165

Table 4 Computation time for the original divorce dataset

Method	Exact	Breslow	Efron	Discrete
Real time	0.45	0.43	0.48	0.43
CPU	0.23	0.15	0.14	0.14

Since there are few ties in the divorce dataset, we recode survival time to increase the number of ties. Considering the range of the variable ‘years’ is 0 to 75, we divide the years into seven intervals and create a new variable ‘decade’ as survival time: 0-10, 11-20, 22-30, 31-40, 41-50, 51-60, 61 or more. Using ‘decade’ as survival time leads to 99.9% of ties. Fig 2 shows the Kaplan-Meier curve using the recoded survival time. A less smooth (i.e., step) function results, which is expected.

Fig 2 Kaplan Meier curve for the recoded divorce dataset



Comparing the results in the following tables, we can find the performance of Efron's approximation is better than Breslow's approximation and the discrete method. As is shown in Table 5, the parameter estimates and standard errors of the covariates for Efron's approximation and the exact method are very close, and those for the discrete method are a little bit larger. But statistics generated from Breslow's approximation are not very close to those from other methods.

Table 6 also shares the same performance for those methods, with Breslow's hazard ratio being a little bit lower than the others and Efron's and the exact method are almost the same.

For the fit statistics in Table 7, the exact method and discrete method have the lowest values, while Breslow's and Efron's are much higher. And Efron's is still lower than Breslow's, which demonstrates the advantage of Efron's approximation when there are many ties.

Table 8 shows the computation time. It appears that Breslow's approximation uses the least time and the discrete method requires a lot of time. All of these tables demonstrate the superiority of Efron's approximation to Breslow's approximation when the percentage of ties is high.

Table 5 Parameter estimate, standard error and p-value of covariates for the recoded divorce dataset

Method	Parameter Estimate for heduc	SE for heduc	p-value for heduc	Parameter Estimate for heblack	SE for heblack	p-value for heblack	Parameter Estimate for mixed	SE for mixed	p-value for mixed
Exact	0.05272	0.04745	0.2666	0.08199	0.07925	0.3009	0.24457	0.07893	0.0019
Breslow	0.04793	0.04754	0.3133	0.07675	0.07913	0.3320	0.22357	0.07880	0.0046
Efron	0.05254	0.04739	0.2676	0.08172	0.07913	0.3018	0.24380	0.07881	0.0020
Discrete	0.05680	0.05164	0.2713	0.09178	0.08664	0.2895	0.26708	0.08660	0.0020

Table 6 Hazard ratio and confidence interval for the recoded divorce dataset

Method	Exact	Breslow	Efron	Discrete
Hazard Ratio for heduc	1.054 (0.961, 1.157)	1.049 (0.956, 1.152)	1.054 (0.961, 1.157)	1.058 (0.957, 1.171)
Hazard Ratio for heblack	1.085 (0.929, 1.268)	1.080 (0.925, 1.261)	1.085 (0.929, 1.267)	1.096 (0.925, 1.299)
Hazard Ratio for mixed	1.277 (1.094, 1.491)	1.251 (1.072, 1.459)	1.276 (1.093, 1.489)	1.306 (1.102, 1.548)

Table 7 Fit statistics for the recoded divorce dataset

Method	Exact	Breslow	Efron	Discrete
-2LOG L	5831.605	16131.515	15964.364	5831.447
AIC	5837.605	16137.515	15970.364	5837.447
SBC	5852.422	16152.332	15985.181	5852.264

Table 8 Computation time for the recoded divorce dataset

Method	Exact	Breslow	Efron	Discrete
Real time	0.39	0.26	0.46	2.01
CPU	0.06	0.03	0.17	1.68

4.2 Readmission Dataset

The readmission data set consists of patients who underwent coronary artery bypass graft surgery at a single U.S multi-hospital institution from July 2014 to May 2017. The outcome of interest was time to hospital readmission after discharge (days). The exposure of interest was sex (i.e., female or male).

The dataset has 5710 observations, 527 of which are failures (readmission) and 97.5% of ties. Figure 3 shows the Kaplan-Meier curve.

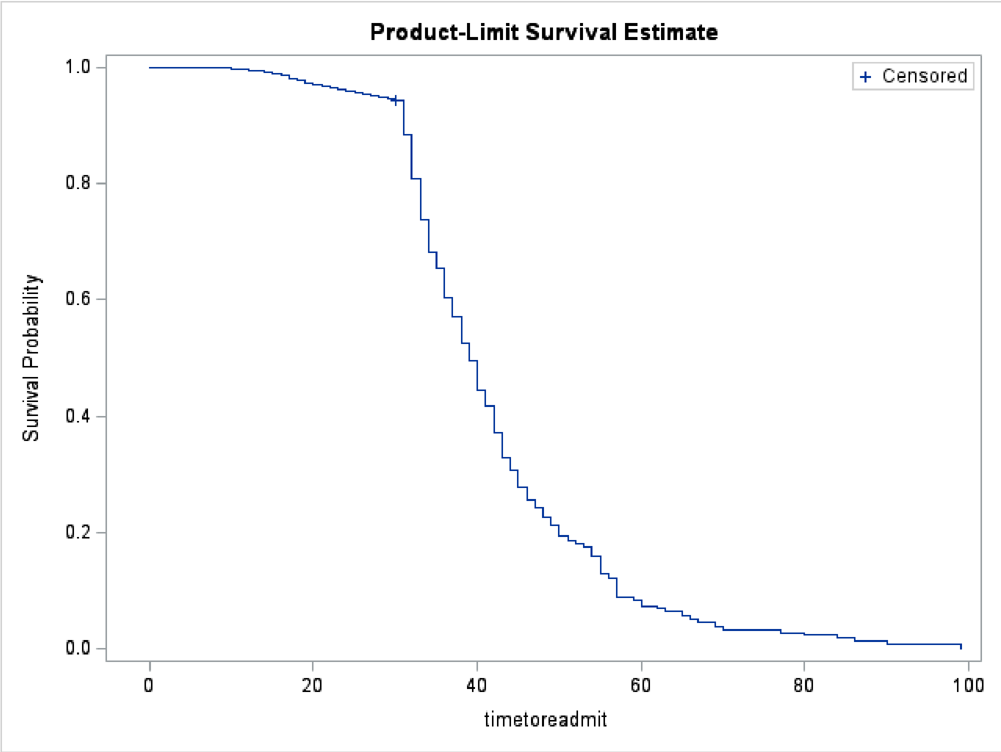


Fig 3 Kaplan Meier curve for readmission dataset

We examine this data set in the same manner that we did previously. One point we need to specify here is that since the event time is regarded as continuous in this dataset, the exact method is the gold standard here.

Table 9 shows the parameter estimates and standard errors for the covariate timetoreadmit and we can see the Efron's approximation for the estimation of the covariate is the closest to the exact method when the percentage of ties is high enough, and also the standard error is the smallest among all the methods. In contrast, Breslow's method performs less euqually well.

In Table 10 we can see that hazard ratio calculated using Efron's procedure is the same as the exact method, while the other methods produce hazard ratios that a little bit different from the exact method. Efron's method also provides the narrowest confidence interval, which indicates good precision.

For the fit statistics in Table 11, we can see that the exact method still has the best performance of goodness-of-fit, and the discrete method also performs well.

Table 12 shows that Efron's method has the shortest computation time and Breslow's method is also much faster than the exact and discrete methods.

Table 9 Parameter estimate and standard error of covariates for readmission dataset

Method	Parameter Estimate for timetoreadmit	Standard Error for timetoreadmit	p-value
Exact	-0.02050	0.09639	0.8316
Breslow	-0.01717	0.09612	0.8582
Efron	-0.02039	0.09613	0.8320
Discrete	-0.01785	0.09802	0.8555

Table 10 Hazard ratio and 95% Wald Confidence Interval for readmission dataset

Method	Exact	Breslow	Efron	Discrete
Hazard Ratio	0.980	0.983	0.980	0.982
95% Wald Confidence Interval	(0.811, 1.183)	(0.814, 1.187)	(0.812, 1.183)	(0.811, 1.190)

Table 11 Fit statistics for readmission dataset

Method	Exact	Breslow	Efron	Discrete
-2LOG L	5606.026	7362.298	7344.467	5606.038
AIC	5608.026	7364.298	7346.467	5608.038
BIC	5612.293	7368.565	7351.034	5612.305

Table 12 Computation time for readmission dataset

Method	Exact	Breslow	Efron	Discrete
Real time	0.50	0.40	0.42	0.46
CPU	0.21	0.15	0.15	0.25

Chapter 5 Conclusion & Recommendation

The divorce dataset and the readmission dataset demonstrate difference in performance among the different methods for handling tied survival times, when the percentage of ties varies. For the original divorce dataset, the four methods generate similar results. The Exact method can take a considerable amount of computer resources. Breslow's approximation performs well when ties are not extensive, in terms of the accuracy and efficiency of the estimates of β and computation time. Efron's approximation also provides satisfactory approximations for the continuous time-scale model. When there are extensively percentage of ties, such as the recoded divorce dataset and the recoded readmission dataset, the performance of Breslow's approximation deteriorates while Efron's continues to do better, as can be seen in the parameter estimates, fit statistics and calculation time.

In general, Efron's approximation gives results that are much closer to the exact method results than Breslow's approximation does. If the time scale is genuinely discrete, we should use the discrete method. If there are no ties, all four methods result in the same likelihood and yield identical estimates. The default, TIES=BRESLOW, is the most efficient method when there are no ties.

Bibliography

Therneau, T., Grambsch, P., *Modeling Survival Data: Extending the Cox Model*. Springer-Verlag, 2000.

Andersen, P., Borgan, Ø., Gill, R.D., Keiding, N., 1993, *Statistical Models Based on Counting Processes*, Springer-Verlag, New York.

Breslow, N., 1975, *Covariance Analysis of Censored Survival Data*, *Biometrics*, vol. 30, no. 1, s. 89–99.

Cox, D.R., 1972, *Regression Models and Life-Tables*, Journal of the Royal Statistical Society. Series B (Methodological), vol. 34, no. 2, s. 187–220.

Efron, B., 1977, *The Efficiency of Cox's Likelihood Function for Censored Data*, Journal of the American Statistical Association, vol. 72, no. 359, s. 557–565.

Allison, P. 2010. *Survival Analysis Using SAS: A Practical Guide*. Cary, NC, USA: SAS Institute Inc.

DeLong, D. M., Guirguis, G. H., and So, Y. C., 1994, *Efficient Computation of Subset Selection Probabilities with Application to Cox Regression*. Biometrika, 81, 607-611.

Kalbfleisch, J.D., Prentice, R.L., 2002, *The Statistical Analysis of Failure Time Data*, John Wiley & Sons, New York.

Kalbfleisch, J.D., Prentice, R.L., 1973, *Marginal Likelihoods Based on Cox's Regression and Life Model*, Biometrika, Vol. 60, No. 2, pp. 267-278

K. Dietz, M. Gail, K. Krickeberg, J. Samet, A. Tsiatis. *Statistics for Biology and Health*, Springer-Verlag New York

Hosmer, D. W., Lemeshow, S., May, S., 2008, *Applied Survival Analysis: Regression Modeling of Time to Event Data, 2nd Edition*, John Wiley & Sons, New York.

SAS Institute Inc. 2010. SAS/GRAPH 9.2 Reference, Second Edition. Cary, NC: SAS Institute Inc.

StataCorp. 2013. Stata: Release 13. Statistical Software. College Station, TX: StataCorp LP.

J. Mandrekar, S. Mandrekar, 2004, *An Introduction to Matching and its Application using SAS*, Statistics and Data Analysis, SUGI 29, Paper 208-29