UNIVERSITI PUTRA MALAYSIA

MEASURES OF INFLUENCE AND WEIGHTED PARTIAL LIKELIHOOD ESTIMATION FOR COX PROPORTIONAL HAZARDS REGRESSION

REBECCA LOO TING JIIN

FS 2016 53

MEASURES OF INFLUENCE AND WEIGHTED PARTIAL LIKELIHOOD ESTIMATION FOR COX PROPORTIONAL HAZARDS REGRESSION

By

REBECCA LOO TING JIIN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

February 2016

DEDICATIONS

Mum
Dad
. . .

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

MEASURES OF INFLUENCE AND WEIGHTED PARTIAL LIKELIHOOD ESTIMATION FOR COX PROPORTIONAL HAZARDS REGRESSION

By

REBECCA LOO TING JIIN

February 2016

Chairman : Anwar Fitrianto, PhD
Faculty : Science

In this study, we consider the development of influential diagnostics to assess case influence for the Cox proportional hazards model and stratified Cox proportional hazards regression model. We examine various residuals previously proposed for these models and develop a diagnostics method using the case-deletion technique. However, existing diagnostics methods are affected by masking effect. This effect may cause diagnostics methods to fail to correctly detect influential cases. Therefore, we propose an influential diagnostics method that has lower masking effect as compared to other methods. The proposed influential diagnostics method is approximately Chi-square distribution with p degress of freedom.

The simulation study is implemented to evaluate the performance of the proposed influential diagnostics method via comparison with existing diagnostics method. Then, the diagnostics methods are applied into the real data such as kidney catheter data, Worcester Heart Attack study and also Stanford Heart Transplant study. The performance of the proposed influential detection method is better than that of the existing influential detection method. The partial likelihood estimation for the Cox regression model is biased when there are measurement errors in the covariate. Therefore, a weighted partial likelihood estimation for Cox regression model is proposed when there is violation of underlying assumptions due to measurement error in the covariates. In the simulation study, the proposed weighted partial likelihood estimations for parameter coefficients have smaller bias, root mean square errors, and ratio of bias over standard error than the existing parameter estimators, both with and without contamination of the covariates. The demonstrated performance of the proposed influential methods and weighted partial likelihood estimators are superior to existing influential detection methods and parameter estimators.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

UKURAN PENGARUH DAN WAJARAN SEPARA PENGANGGARAN KEBOLEHJADIAN UNTUK REGRESI COX BAHAYA BERKADARAN

Oleh

REBECCA LOO TING JIIN

Februari 2016

Pengerusi : Anwar Fitrianto, PhD
Fakulti : Sains

Dalam kajian ini, kami menyelidik pembangunan diagnostik berpengaruh bagi menilai pengaruh kes untuk model bahaya berkadaran Cox dan model regresi bahaya berkadaran Cox berstrata. Kami memeriksa pelbagai reja yang telah diusulkan untuk semua model ini, dan mengusulkan kaedah diagnostik dengan menggunakan teknik penghapusan kes. Namun begitu, kaedah diagnostik sedia ada telah dipengaruhi kesan penyelubungan. Kesan ini akan menyebabkan kaedah diagnostik itu tidak dapat mengesan kes berpengaruh dengan betul. Oleh itu, kami cuba untuk mengusulkan kaedah diagnostik berpengaruh yang mempunyai kesan penyelubungan yang lebih rendah berbanding kaedah lain. Kaedah diagnostik berpengaruh yang diusulkan adalah anggaran taburan Chi-square dengan darjah kebebasan *p*.

Kajian simulasi dilaksanakan bagi menilai prestasi kaedah diagnostik berpengaruh yang diusulkan dengan membandingkannya dengan kaedah diagnostik sedia ada. Kemudian, kaedah diagnostik diaplikasikan pada data sebenar seperti data kateter buah pinggan, kajian serangan jantung Worcester, dan juga kajian pemindahan jantung Stanford. Prestasi kaedah pengesanan berpengaruh yang diusulkan adalah lebih baik berbanding kaedah pengesanan berpengaruh sedia ada. Anggaran kebolehjadian separa untuk model regresi Cox terpincang apabila ada ralat ukuran dalam kovariat. Oleh itu, anggaran kebolehjadian wajaran separa untuk model regresi Cox diusulkan apabila terdapat pelanggaran andaian dasar iaitu ralat ukuran dalam kovariat. Daripada kajian simulasi, anggaran kebolehjadian wajaran separa yang diusulkan untuk pekali parameter mempunyai pincang, ralat punca min persegi, dan nisbah pincang melawan ralat piawai yang lebih kecil berbanding penganggar parameter sedia ada dengan dan tanpa pencemaran dalam kovariat. Oleh itu, prestasi kaedah berpengaruh dan juga penganggar kebolehjadian wajaran separa yang diusulkan adalah lebih baik berbanding kaedah pengesanan berpengaruh dan penganggar parameter sedia ada.

ACKNOWLEDGEMENTS

It would be impossible for me to say thanks to all the nice people who have helped me in numerous ways during my master study at Universiti Putra Malaysia. However, I must take the time to express my gratitude to the main contributors to this work.

First and foremost, I wish to thank all my supervisory committee members, Dr. Anwar Fitrianto, Associate Prof. Dr Jayanthi Arasan, and Prof. Dr. Habshah Midi, for the guidance during these years. I consider myself extremely fortunate to have them as my supervisory committee members for my master study.

I also wish to express my gratitude to my family, whose support was always there when I needed it. Thank you for loving me so much. I owe everything to them.

Finally, I also thank my friends for their selfless help throughout my time at UPM and for providing me with some memorable experiences.

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science.

The members of the Supervisory Committee were as follows:

Anwar Fitrianto, PhD
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

Jayanthi Arasan, PhD
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

Habshah Binti Midi, PhD
Professor
Faculty of Science
Universiti Putra Malaysia
(Member)

_____
BUJANG KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citation have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Rebecca Loo Ting Jiin , GS37708

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: _____
Name of Chairman of Supervisory Committee: Dr. Anwar Fitrianto

Signature: _____
Name of Member of Supervisory Committee: Associate Professor Dr. Jayanthi Arasan

Signature: _____
Name of Member of Supervisory Committee: Professor Dr. Habshah Binti Midi

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BMI	Body Mass Index
cdf	Cumulative Distribution function
cp	Censoring Proportion
d.f.	Degree of Freedom
DRLD	Diagnostics Robust Likelihood Displacement
DRLDs	Stratified Diagnostics Robust Likelihood Displacement
id	Identification Number
pdf	Probability Density Function
RMSE	Root Mean Square Error
SAS	Statistical Analysis Software
WHAS 500	Worcester Heart Attack Study

CHAPTER 1

INTRODUCTION

Statistical modelling is an essential part in data analysis and has an important role as it can lead to understanding the effect of explanatory variables on the survival times. Covariates are the explanatory variables that affect the time-to-event. However, these models can only be presented correctly if the data fulfils certain assumptions of the model. If there is a minor modification is seriously affects the key results of the analysis, then the model or data set should be re-examined to investigate the cause of the problems.

Cox (1972) proposed a statistical model that can be used to assess the relationship between a set of covariates and time-to-event with censored data. This model is known as Cox proportional hazards regression model.

Influential diagnostics are generally considered to be useful to identify unusual observation. A few general methods are available and developed in the context other than normal linear or generalized linear regression such as Cook's distance, likelihood displacement and so on. However, the models that are considered by these authors do not include censored data. In the latter, authors like Schoenfeld (1982), Cain and Lange (1984), Pettitt and Bin Daud (1989), Nardi and Schemper (1999) had developed methods for model adequacy assessment in the presence of censored data in Cox (1972) proportional hazards model.

The work in this thesis is mainly extending influential diagnostics method due to modest perturbation which included censored data. Perturbation in single and multi-covariates model is investigated. The outcome of the analysis due to abnormal subjects existed is investigated through the case-deletion technique and has a cut-off point with its approximating distribution.

The perturbation schemes that are considered in most of the works here are based on contaminating the data which modified subjects to extreme values. The proposed influential diagnostics method in this thesis is basically from the idea of the single case deletion method, which was extended to the Cox regression model and stratified Cox regression model from the methods that were introduced by Cain and Lange (1984) and Pettitt and Bin Daud (1989). This case deletion technique provides the basis to detect the influential cases in the data set to contribute influence effect on the parameter estimation. An observation might be judged as influential observation if the estimated parameter coefficients are altered substantially in the model by the observation.

The influential observation is observation that give dramatic influence to parameter estimates. The estimated parameters that are affected by the observation that is judged as influential will not present information correctly from the fitted model. In short, the estimated coefficients of the Cox regression model is biased under violation of underlying assumptions such as measurement error or contamination in the covariates (Bednarski, 1993). Therefore, few parameter estimation methods in Cox model that are less bias in case there are existing influential observations, were developed by Sasieni (1993a), Sasieni (1993b), Bednarski (1993), Viviani and Farcomeni (2010) and Farcomeni and Viviani (2011).

Methods for robust estimation that were discussed by the researchers are either based on weighting or trimming the suspected observations that give more influence to the estimation. Downweighting even just one subject in the Cox regression model is directly corresponds to downweighting few risk sets. Let $x_{(i)}$ is an observation with survival time, $t_{(i)}$ that was sorted in ascending order of the rank of time. The longer the survival time, the more risk sets of the observations get involved.

Before the scope of this thesis is outlined, we shall brief overview of some important keywords definition used in the thesis.

1.1 Censored Data

The main difficulty in the analysis of time-to-event data is that it involved some individuals that may not be observed for the full follow up time until it experience the event such as failure. Not all subjects have failed at the end of the study, and for the subjects that have not failed before end of the study will have longer survival or failure time. Such an incomplete observation is known as censoring. The data sets that contain these kinds of incomplete observations are called censored data.

Three main types of censoring are right censoring, left censoring and interval censoring. The right censoring of a subject is that the subject does not experience the event when the study ended. The right censoring can be separated into three main types of censoring which are type I, type II and random censoring.

In type I censoring, let c_f be a fixed censoring time. If the observed lifetime is less than c_f then the lifetime is uncensored and the rest are censored at c_f. While in type II right censoring, the study ends when a fixed number of subjects have failed and the rest is censored. At last, random censoring could be described as each subject has potential censoring time and a potential lifetime, both are assumed to be independent random variables. The lifetime is only observed if the potential lifetime is less than

the censoring time.

For left censoring, the interested event has already occurred before the observed lifetime but the exact lifetime is unknown and only know that it is less than the censoring time.

The exact lifetime in interval censoring is usually unknown because the subject's lifetime is failing within a certain interval $[l_i, r_i]$ where l_i and r_i are left and right censoring times, respectively. Left and right censoring are a special case of interval censoring.

The lifetime data that is truncated due to some of the conditions in the study design is known as truncation. Let say u_i and v_i are left and right truncation times, the individual's time-to-event do not lie between the interval $(u_i, v_i]$ will not be included in the study.

Ignoring the censored observations may lead to bias in inferences so that those observations have to be encountered and should not be ignored in the analysis. Special statistical procedures are adopted to treat the censored data (Collett, 2003).

1.2 Semiparametric Models for Survival Data

A model that specifies hazard function of time-to-event conditional on a set of covariates is a product of baseline hazard function and exponential of covariates with an effect of hazards. Cox (1972) introduced an estimator by using maximum likelihood estimation which does not involve all the observations, which is known as partial likelihood. The popular of Cox regression model is that it does not need underlying assumption for baseline hazards function so that, it is known as semiparametric model. Proportional hazards assumption is needed in the Cox regression model but it is often violated. Some other semiparametric models were introduced as extension of the Cox model when the proportional hazards assumption is violated such as stratified Cox model.

1.2.1 Cox Proportional Hazards Regression Model

Cox proportional hazards model was first proposed by Cox (1972) which is concerning the analysis of the effect of vector of variables with censored failure times by

using regression model. The values of a set of explanatory variables, x are provided by all individual who involves in the study and a follow up period is observed until it experienced the event or censored. The hazard function, which is defined as the instantaneous failure rate is formulated as a function of these covariates with the unknown regression coefficients, β.

Cox proportional hazards model is a popular method to study time-to-event data in various fields, e.g. credit scoring, clinical trials, marital instability, food security and so on. One of the interesting properties of the Cox model is that it does not require the full specification of the distribution of the follow up duration of the study. Besides that, not all the observations that enrolled in the study do experienced the event, such data is called censored data.

A more detailed discussion about Cox proportional hazards regression model and literature of the previous works are given in Chapter 2.

1.2.2 Stratified Cox Proportional Hazards Model

In stratified Cox regression, the assumption that subjects have different baseline hazard if they are from different stratum. The baseline hazards are allowed to differ by stratum, but the coefficient values, β are the same.

To estimate the parameter coefficients, the subjects are categorized according to each stratum and the risk set in likelihood function are separated by stratum. The strata in stratified Cox model are similar as block, a nuisance factor that will affect the estimation. Then, the effect of the strata is not interested to be estimated in the stratified regression model.

1.2.3 Risk Set

Risk set is a set of observation has risk to experience the event at time, t_i. Suppose observation j fails before i in which $j < i$ then observation i is said to be in j's risk set. This is the set of observations that are at risk to experience the event or fail when j is failing at time, t_j. The observations that survived longer in the study contribute in more risk sets in which they are included. The more risk sets that the observations are included, the more contribution of the observations to the partial likelihood to estimate parameter coefficients in the Cox proportional hazards model.

1.3 Influential Observation

In a set of data, it might appear an observation to be inconsistent with the majority of the data. This observation might cause suspicion with different mechanism. A suspected observation will show up large gap between "abnormal" and "normal" observations, and the deviation between "unusual" and the group of "normal" observations is slightly larger than the majority. These such unusual observations should be foreign to the main population because of their nature, they may cause some difficulties in the attempt to represent the population. Also, they might come from different population with different nature. In short, influential observations are the subject that can change grossly the estimate parameter coefficients in some model for the population.

There are few ways to give rise to unusual subjects. One is misspecification in the collection or recording the observations. And, there might exist some observations that display different information as compared to the majority in the data set. For example, most patients who experienced certain diseases and weak healthy condition might have shorter life, but there still might exist some patients who lived in very long period. From this influential assessment, we could be able to identify which observations displays differently in the model and should be investigated. Besides that, throughout the influential diagnostics, we could detect these kind of "influential" patients and then analyse differently since they might give useful information in the analysis.

Influential observation does not mean an outlier in the data set. Influential observation gives some influences to the parameter estimation or inferences and presented differently to give higher effect to change the inferences substantially than other majority subjects. From the influential observation, we might get more useful information and could investigate the reason of why this observation is "influential".

1.4 Problem Statements

Influential diagnostics is essential for identifying the influential observations. These observations should be detected because they will affect the parameter estimation in the regression model. But, not all the influential diagnostics methods are effective in detecting all the influential subjects since there will be a risk to have a masking effect. Masking effect is an effect that causes the influential diagnostics method fail to detect the influential observation correctly. This effect is the main problem to the diagnostics methods (Pettitt and Bin Daud, 1989). When the number of influential observations is large, the diagnostics methods might not able to detect those observations as influential.

The covariate values with contamination or measurement error might cause significant effect to change the estimates parameter (Wang et al., 2006, Desmarais and Harden, 2012). Therefore, these influential observations should be identified and the masking effect should be decreased to avoid the diagnostics method to detect those influential observations as non-influential.

Cox regression coefficients will be affected by the influential observation and will have a dramatic change when the underlying assumption is violated such as measurement error in the covariate (Heritier et al., 2009, Desmarais and Harden, 2012). Therefore, the parameter coefficient will be biased. When the parameter in the regression model is biased, the effect of the covariate will not present the results correctly.

1.5 Objectives of the Study

The objectives of the research are,

1. to develop a new influence measure method to identify influential observations that can avoid masking effect when there is measurement error or contamination in the covariate in Cox proportional hazards regression model.
2. to develop a new method of influence measure to identify influential observations that avoid masking effect in the presence of measurement error or contamination in the covariate for stratified Cox proportional hazards regression model.
3. to propose parameter estimator that has lower bias but higher efficiency in the presence of measurement error or contamination in the covariate for Cox proportional hazards regression model.

1.6 Scope of the Thesis

This thesis consists of 6 chapters. Chapter 1 is the general introduction for Cox proportional hazards regression model and stratified Cox proportional hazards regression model as well as influential diagnostics.

Chapter 2 presents some background of the study that are related to the current work. It reviews literatures on some basis of Cox and stratified Cox models. Background of residual analysis and influential diagnostics are also discussed here. Finally, the literatures on robust parameter estimation on censored modeling is reviewed.

Chapter 3 concentrates on influential measures with the violation of the assumptions such as measurement error or contamination in the covariate in the Cox proportional hazards model. A proposed method for identifying influential observation is applied to the kidney catheter data and Worcester Heart Attack study. Simulation studies are implemented to evaluate the performance of the proposed influential diagnostics method.

Chapter 4 focuses on influential diagnostics in stratified Cox proportional hazards regression model. The proposed influential detection method for the stratified Cox model is similar to the proposed method in Chapter 3 which is applied to the Stanford Heart Transplant study. A simulation study is conducted to obtain the model parameters and a few covariate values were contaminated randomly. Most of the contaminated subjects are influential subjects but not all the influential observations are contaminated observations. Comparison between the proposed and existing influential observations detection method is implemented by using simulation study and real data analysis for the stratified Cox model.

Chapter 5 explores the robust parameter estimation in Cox proportional hazards model in the presence of measurement error or contamination in the covariate. A simulation study is conducted and the performance is evaluated based on bias and root mean square error (RMSE). The simulation study is used to evaluate the performance of the proposed partial likelihood estimator with the ordinary Cox partial likelihood estimator and other estimators which are known as robust.

Last but not least, Chapter 6 is the last chapter in this thesis. It presents the conclusion of our research which is included influential diagnostics and also parameter estimation that illustrated lower bias and higher efficiency when there is contamination in the covariate.

BIBLIOGRAPHY

Aggarwal, C. C. 2013. *Outlier Analysis*. Springer.

Austin, P. C. 2012. Generating survival times to simulate Cox proportional hazards models with time-varying covariates. *Stat Med.* 31 (29): 3946–3958.

Bednarski, T. 1993. Robust estimation in Cox's regression model. *Scandinavian Journal of Statistics* 20 (3): 213–225.

Bednarski, T. and Nowak, M. 2003. Robustness and efficiency of Sasieni-type estimators in the Cox model. *Journal of Statistical Planning and Inference* 115: 261–272.

Belsley, D. A., Kuh, E. and Welsch, R. E. 1980. *Regression Diagnostics Identifying Influential Data and Sources of Collinearity*. John Wiley and Sons.

Bender, R., Augustin, T. and Blettner, M. 2005. Generating survival times to simulate Cox proportional hazards models. *Statistics in Medicine* 24 (11): 1713–1723.

Burton, A., Altman, D. G., Royston, P. and Holder, R. L. 2006. The design of simulation studies in medical statistics. *Statistics in Medicine* 25: 4279–4292.

Cain, K. C. and Lange, N. T. 1984. Approximate case influence for the proportional hazards regression model with censored data. *Biometrics* 40: 493–499.

Cheung, M. W. L. 2009. Constructing approximate confidence intervals for parameters with structural equation models. *Structural Equation Modeling* 16: 297–294.

Chiang, J. T. 2007. The Masking and Swamping Effects Using the Planted Mean-Shift Outliers Models. *Int. J. Contemp. Math. Sciences* 2 (7): 297–307.

Cochran, W. G. 1977. *Sampling Techniques*. John Wiley & Sons, Incorporated, New York.

Collett, D. 2003. *Modelling Survival Data in Medical Research*. 2nd edn. CRC Press.

Cook, R. D. 1986. Assessment of local influence. *J. R. Statist. Soc. B* 48 (2): 133–169.

Cox, D. R. 1972. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological)* 34 (2): 187–220.

Cox, D. R. 1975. Partial likelihood. *Biometrika* 2 (2): 269–276.

Cox, D. R. and Oakes, D. 1984. *Analysis of Survival Data*. Chapman and Hall.

Cox, D. R. and Snell, E. J. 1968. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)* 30 (2): 248–275.

Desmarais, B. A. and Harden, J. 2012. Comparing partial likelihood and robust estimation methods for the Cox regression model. *Political Analysis* 20: 113–135.

Efron, B. 1977. The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association* 72: 557–565.

Escobar, L. A., Meeker, W. Q. and Jr. 1992. Assessing influence in regression analysis with censored data. *Biometrics* 48 (2): 507–528.

Farcomeni, A. and Ventura, L. 2012. An overview of robust methods in medical research. *Statistical Methods in Medical Research* 21 (2): 111–133.

Farcomeni, A. and Viviani, S. 2011. Robust estimation for the Cox regression model based on trimming. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 53: 956–973.

Grambsch, P. and Therneau, T. M. 1994. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* 81 (3): 515–526.

Hampel, F. R., Ronchetti, E. M., Peter, J. R. and Stahel, W. A. 1986. *Robust Statistics*. John Wiley & Sons.

Hebel, J. R. and McCarter, R. J. 1990. *A Study Guide to Epidemiology and Biostatistics*. 7th edn. Jones & Bartlett Learning.

Heritier, S., Cantoni, E., Copt, S. and Victoria-Feser, M. 2009. Robust Methods in Biostatistics .

Hosmer, D., Lemeshow, S. and May, S. 2008. *Applied Survival Analysis*. 2nd edn. Wiley Series in Probability and Statistics.

Lin, D. Y. and Wei, L. J. 1989. The Robust Inference for the Cox Proportional Hazards Model. *Journal of the american statistical association* 84 (408): 1074–1078.

McGilchrist, C. A. and Aisbett, C. W. 1991. Regression with frailty in survival analysis. *Biometrics* 47 (2): 461–466.

Minder, C. E. and Bednarski, T. 1996. A robust method for proportional hazards regression. *Stat. Med.* 15 (10): 1033–1047.

Nardi, A. and Schemper, M. 1999. New residuals for Cox regression and their application to outlier screening. *Biometrics* 55 (2): 523–529.

Norazan, M. R., Midi, H. and Imon, A. H. M. R. 2009. Estimating Regression Coefficients Using Weighted Bootstrap with Probability. *WSEAS Transactions on Mathematics* 8 (7): 362–371.

O'Quigley, J. 2008. *Proportional Hazards Regression*. Springer.

Pettitt, A. N. and Bin Daud, I. 1989. Case-weighted measures of influence for proportional hazards regression. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 38 (1): 51–67.

Rahman, G., Mubeen, S. and Rehman, A. 2015. Generalization of Chi-square distribution. *Journal of Statistics Applications & Probability* 4 (1): 119–126.

Sasieni, D. 1993a. Maximum weighted partial-likelihood estimators for the Cox model. *Journal of the American Statistical Association* 88 (421): 144–152.

Sasieni, P. 1993b. Some new estimators for Cox regression. *The Annals. of Statistics* 21 (4): 1721–1759.

Schoenfeld, D. 1982. Partial residuals for the proportional hazards regression model. *Biometrika* 69 (1): 239–241.

Selvin, S. 2008. *Survival Analysis for Epidemiologic and Medical Research*. Cambridge University Press.

Smith, P. J. 2002. *Analysis of Failure and Survival Data*. Chapman & Hall/CRC.

Therneau, T. M., Grambsch, P. M. and Fleming, T. R. 1990. Martingale-based residuals for survival models. *Biometrika* 77 (1): 147–160.

Velleman, P. F. and Welsch, R. E. 1981. Efficient computing of regression diagnostics. *The American Statistician* 35 (4): 234–242.

Viviani, S. and Farcomeni, A. 2010. Trimmed Cox regression for robust estimation in survival studies. *Technical Report 7. Department of Statistics. Sapienza-University of Rome* .

Wang, H. M., Jones, M. P. and Storer, B. E. 2006. Comparison of case-deletion diagnostic methods for Cox regression. *Statistics in Medicine* 25: 669–683.

Weissfeld, L. A. 1990. Influence diagnostics for the proportional hazards model. *Statistics & Probability Letters 10* 10 (5): 411–417.