

UNIVERSITI PUTRA MALAYSIA

A SIMULATION STUDY ON COMPETING RISKS WITH CENSORED DATA USING COX MODEL

IING LUKMAN

FSAS 1999 3



A SIMULATION STUDY ON COMPETING RISKS WITH CENSORED DATA USING COX MODEL

By

IING LUKMAN

Thesis Submitted in Fulfilment of the Requirements for the Degree of Master of Science in the Faculty of Science and Environmental Studies
Universiti Putra Malaysia

December 1999



2

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

A SIMULATION STUDY ON COMPETING RISKS WITH CENSORED DATA USING COX MODEL

By

IING LUKMAN

December 1999

Chairman: Noor Akma Ibrahim, Ph.D.

Faculty: Science and Environmental Studies

A simulation study was performed to compare two regression methods for competing risks with censored data. The first method was the conventional Cox's proportional hazard regression model (Cox model). The second method was based on Cox model using a duplicated data technique of Lunn and McNeil (or the modified Lunn-McNeil). Samples with various sizes and censoring percentages were generated and fitted using both methods. This study was conducted by comparing the inference of both methods, using Root Mean Square Error (RMSE), the power tests, and the Schoenfeld residuals analysis. The power tests used in this study were likelihood ratio test, Rao-score test, and Wald statistics. The Schoenfeld residuals analysis was conducted to check the proportionality of the model through its covariates. The estimated parameters were computed for cause-specific hazards. Results showed the RMSE were



generally smaller for the model of the modified Lunn-McNeil method than that of the ordinary Cox method. The power tests of the likelihood ratio statistics and Rao-score test were only powerful for the unstratified Cox model, so that, it could be concluded that the model had more advantages than the modified Lunn-McNeil one. However, results from the analysis of Schoenfeld residuals indicated that the modified Lunn-McNeil was better than the ordinary Cox in complying with the proportional hazards model assumption with respect to certain covariates.



4

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

KAJIAN SIMULASI TERHADAP RISIKO BERSAING DENGAN DATA TERTAPIS MENGGUNAKAN MODEL COX

Oleh

IING LUKMAN

Disember 1999

Pengerusi:

Noor Akma Ibrahim, Ph.D.

Fakulti:

Sains dan Pengajian Alam Sekitar

Kajian simulasi dijalankan untuk membandingkan dua kaedah regresi bagi risiko bersaing dengan data tertapis. Kaedah pertama ialah model regresi kadaran bahaya Cox biasa (model Cox). Kaedah kedua ialah model Cox yang berlandaskan pada penggunaan teknik data yang sama dari Lunn dan McNeil (atau kaedah Lunn-McNeil terubahsuai). Beberapa sampel dengan saiz berbeza dan peratusan tapisan berbeza dijana dan dianalisis menggunakan kedua-dua kaedah tersebut. Kajian ini dijalankan dengan membandingkan inferens dari kedua-dua kaedah tersebut, menggunakan Punca Kuasadua Min Ralat (PKMR), ujian kuasa, dan analisis reja Schoenfeld. Ujian kuasa yang digunakan ialah ujian nisbah kebolehjadian, ujian skor-Rao, dan statistik Wald. Analisis reja Schoenfeld dijalankan untuk meneliti keseimbangan model menerusi kovariatnya. Anggaran parameter dihitung bagi punca bahaya tertentu.



Keputusan menunjukkan PKMR secara amnya lebih kecil bagi model Lunn-McNeil terubahsuai berbanding dengan kaedah Cox biasa. Ujian kuasa dari statistik nisbah kebolehjadian dan ujian skor-Rao adalah hanya berkuasa bagi model Cox takberstrata, jadi dapat disimpulkan bahawa model ini memiliki kelebihan ke atas kaedah Lunn-McNeil terubahsuai. Bagaimanapun, keputusan daripada analisis reja Schoenfeld menunjukkan bahawa Lunn-McNeil terubahsuai adalah lebih baik berbanding Cox biasa kerana ia mematuhi andaian model kadaran bahaya meskipun untuk beberapa kovariat tertentu sahaja.



ACKNOWLEDGMENTS

Praise be to Alloh Subhanahu Wata'ala who hath given me the permission to write this thesis, and peace be upon His Messenger, Muhammad Sallallahu 'Alaihi Wasallam (SAW) and his followers. Alloh Subhanahu Wata'ala senth Muhammad SAW for the benefit of the entire universe (Holy Koran epistle 21, verse 107) until the ages ending.

My great thanks to Dr. Noor Akma Ibrahim, the chairperson of the supervisory committee, because of her guidance, patience, and critics have given me the understanding of the whole work written in this thesis. My great thanks also goes to Associate Prof. Dr. Isa bin Daud, the committee member, for giving me a better understanding about Schoenfeld residuals. My great thanks is also to Puan Fauziah bt. Maarof, the committee member, for checking on the hypothesis I used in the power tests and her permission to let me used the computer in the computer laboratory.

An another great thanks goes to the authorities of IRPA project with code number 51119 which is led by Associate Prof. Dr. Isa bin Daud for giving me the financial assistance through the Graduate Assistantship scheme.



My thanks also goes to Associate Prof. Dr. Jambari Hj. Ali of Biology Department (UPM), and Associate Prof. Dr. Abdurrauf Rambe, Mst., of Statistics Department, Institut Pertanian Bogor for giving me the letters of recommendation at the time I applied to study at UPM. To the whole staffs of the UPM Graduate School Office, I express my thanks.

To my beloved wife, Maria Viva Rini, and my beloved children, Tanukh, Haifa, Puput, and Aida, I send them this special thanks. Also a special thank you to my parents, brothers and sisters in Bandung for their supports and prayer. A lot of thanks to all of my Indonesian friends, especially Ida Yuhana, Frediantony Nasdean, Lalang Buana, Iman Rahayu, and Syamsuddin Toaha. Also thanks to my Sudanese friends such as Faiz Elfaki, Yusuf, and Anwar. May Alloh Subhanahu Wata'ala give a lot of rewards to those I say thanks.



TABLE OF CONTENTS

		Page
ABSTRACT	,	2
ABSTRAK	,	4
ACKNOWLE	EDGMENTS	6
APPROVAL	SHEETS	8
DECLARAT	ION FORM	10
LISTS OF TA		13
LISTS OF FI		17
CHAPTER		
I	INTRODUCTION	25
	Background	25
	Aim and Objective of the Research	29
	The Scope of the Thesis	30
II	COX PROPORTIONAL HAZARDS MODEL	
	(COX MODEL)	31
	Censoring	37
	Estimation of the Regression Parameters	38
	Hypothesis Testing	42
	Several Types of Failure	44
	Example of Parameter Estimation	45
	Five Points Data	45
III	COMPETING RISKS	50
	Application of Cox Regression Model to the Competing	
	Risks	54
	Lunn-McNeil Method	55
	Simplification or Modification of Lunn-McNeil	
	Method	60
	Results from Lunn-McNeil and its Modifications	64
	Sample Size and Power	70
	Distribution Used in Generating Data	73
	Binomial Distribution	74
	Exponential Distribution	74



	A Simple Test of the Validity of the Proportional	
	Hazards Model	76
	Schoenfeld Residuals	77
	Power Test	80
IV	SIMULATION STUDY	81
	Results from Simulation Study for the RMSE	84
	Discussion	105
	Statistical Tests on Regression Coefficients	107
	Results from Simulation study for the Power Tests	113
	Discussion	118
	Schoenfeld Residuals Computation	119
	Results from Simulation Study for Schoenfeld	
	Residuals	120
	Discussion	150
V	CONCLUSION	154
	Summary	154
	Conclusion	155
	Suggestion for Further Research	157
BIBLIOGRA	APHY	159
APPENDIX		
A	Programming in SAS/PROC IML	166
В	Stanford Heart Transplant Data	169
VITA		173



LIST OF TABLES

Tables		Page
1	Cox Model for Two-group Treatment in Clinical Studies	35
2	Parameter Estimate Calculation from SAS/PROC IML	49
3	Data Handling of Lunn-McNeil Method	56
4	Data Handling of Modified Lunn-McNeil Method	61
5	Result of the First Model of Lunn-McNeil (1995)	66
6	Result of the First Model of Modified Lunn-McNeil	67
7	Result of the Second Model of Lunn-McNeil (1995)	68
8	Result of the Second Model of Modified Lunn-McNeil	69
9	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=15 and cp=25% in the ordinary Cox Method	84
10	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=15 and cp=25% in the modified Lunn-McNeil Method	85
11	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=15 and cp=50% in the ordinary Cox Method	86
12	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=15 and cp=50% in the modified Lunn-McNeil	86
13	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=45 and cp=25% in the ordinary Cox Method	87



14	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=45 and cp=25% in the modified Lunn-McNeil Method	88
15	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=45 and cp=50% in the ordinary Cox Method	88
16	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=45 and cp=50% in the modified Lunn-McNeil Method	89
17	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=25% in the ordinary Cox Method	90
18	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=25% in the modified Lunn-McNeil Method	91
19	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=50% in the ordinary Cox Method	92
20	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=50% in the modified Lunn-McNeil Method	92
21	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=75% in the ordinary Cox Method	93
22	Cause-Specific Proportional Hazards Model Unstratified (1 st Model) Over One Thousand Simulated Samples for n=80 and cp=75% in the modified Lunn-McNeil Method	94
23	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=15 and cp=25% in the ordinary Cox Method	95



24	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=15 and cp=25% in the modified Lunn-McNeil Method	95
25	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=15 and cp=50% in the ordinary Cox Method	96
26	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=15 and cp=50% in the modified Lunn-McNeil Method	97
27	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=45 and cp=25% in the ordinary Cox Method	98
28	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=45 and cp=25% in the modified Lunn-McNeil Method	98
29	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=45 and cp=50% in the ordinary Cox Method	99
30	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=45 and cp=50% in the modified Lunn-McNeil Method	100
31	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=25% in the ordinary Cox Method	101
32	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=25% in the modified Lunn-McNeil Method	101
33	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=50% in the ordinary Cox Method	102



34	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=50% in the modified Lunn-McNeil Method	103
35	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=75% in the ordinary Cox Method	104
36	Cause-Specific Proportional Hazards Model Stratified (2nd Model) Over One Thousand Simulated Samples for n=80 and cp=75% in the modified Lunn-McNeil Method	104
37	Results of the Power Tests of the Cause-Specific Hazards in Competing Risks with Censored Data for Unstratified Model per 1000 simulated Data set	104
38	Results of the Power Tests of the Cause-Specific Hazards in Competing Risks with Censored Data for Stratified Model per 1000 Simulated Data set	116



LIST OF FIGURES

Figures		Page
1	Plot of the Hazard functions for Cox Model with One Binary Covariate for Two-group in the Table 1	36
2	Plot of the Logarithm of the Hazard Functions for Cox Model with One Binary Covariate	36
3	Cause-specific Curves of Heart Transplant Survival	65
4	Probability Density Function of Exponential Distribution	75
5	Schoenfeld Residuals on Carcinogenesis Data	79
6	Root Mean Square Error for n=15 and cp=25% in the Unstratified Model for Both Methods	85
7	Root Mean Square Error for n=15 and cp=50% in the Unstratified Model for Both Methods	87
8	Root Mean Square Error for n=45 and cp=25% in the Unstratified Model for Both Methods	88
9	Root Mean Square Error for n=45 and cp=50% in the Unstratified Model for Both Methods	90
10	Root Mean Square Error for n=80 and cp=25% in the Unstratified Model for Both Methods	91
11	Root Mean Square Error for n=80 and cp=50% in the Unstratified Model for Both Methods	93
12	Root Mean Square Error for n=80 and cp=75% in the Unstratified Model for Both Methods	94
13	Root Mean Square Error for n=15 and cp=25% in the Stratified Model for Both Methods	96



14	Root Mean Square Error for n=15 and cp=50% in the Stratified Model for Both Methods	97
15	Root Mean Square Error for n=45 and cp=25% in the Stratified Model for Both Methods	99
16	Root Mean Square Error for n=45 and cp=50% in the Stratified Model for Both Methods	100
17	Root Mean Square Error for n=80 and cp=25% in the Stratified Model for Both Methods	102
18	Root Mean Square Error for n=80 and cp=50% in the Stratified Model for Both Methods	103
19	Root Mean Square Error for n=80 and cp=75% in the Stratified Model for Both Methods	105
20	Power Tests of the Likelihood Ratio, Rao-score and Wald Statistics per Sample Size and Its Censoring Percentage in the Unstratified ordinary Cox	115
21	Power Tests of the Likelihood Ratio, Rao-score and Wald Statistics per Sample Size and Its Censoring Percentage in the Unstratified modified Lunn-McNeil	115
22	Power Tests of the Likelihood Ratio, Rao-score and Wald Statistics per Sample Size and Its Censoring Percentage in the Stratified ordinary Cox	117
23	Power Tests of the Likelihood Ratio, Rao-score and Wald Statistics per Sample Size and Its Censoring Percentage in the Stratified modified Lunn-McNeil	117
24	Schoenfeld Residuals for Age with n=15 and cp=25% of the Cause-Specific Hazards of the ordinary Cox in Competing Risks	122



25	Schoenfeld Residuals for Age with n=15 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	122
26	Schoenfeld Residuals for Mismatch Score with n=15 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	123
27	Schoenfeld Residuals for Mismatch Score with n=15 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	123
28	Schoenfeld Residuals for ffage with n=15 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	124
29	Schoenfeld Residuals for ffage with n=15 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	124
30	Schoenfeld Residuals for ffm with n=15 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	125
31	Schoenfeld Residuals for ffm with n=15 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	125
32	Schoenfeld Residuals for Age with n=15 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	126



33	Schoenfeld Residuals for Age with n=15 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in CompetingRisks	126
34	Schoenfeld Residuals for Mismatch Score with n=15 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	127
35	Schoenfeld Residuals for Mismatch Score with n=15 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	127
36	Schoenfeld Residuals for ffage with n=15 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	128
37	Schoenfeld Residuals for ffage with n=15 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	128
38	Schoenfeld Residuals for ffm with n=15 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	129
39	Schoenfeld Residuals for ffm with n=15 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	129
40	Schoenfeld Residuals for Age with n=45 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	130
41	Schoenfeld Residuals for Age with n=45 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	130



42	Schoenfeld Residuals for Mismatch Score with n=45 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	131
43	Schoenfeld Residuals for Mismatch Score with n=45 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	131
44	Schoenfeld Residuals for ffage with n=45 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	132
45	Schoenfeld Residuals for ffage with n=45 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	132
46	Schoenfeld Residuals for ffm with n=45 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	133
47	Schoenfeld Residuals for ffm with n=45 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	133
48	Schoenfeld Residuals for Age with n=45 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	134
49	Schoenfeld Residuals for Age with n=45 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	134
50	Schoenfeld Residuals for Mismatch Score with n=45 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	
51	Schoenfeld Residuals for Mismatch Score with n=45 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	135



52	Schoenfeld Residuals for ffage with n=45 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	136
53	Schoenfeld Residuals for ffage with n=45 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	136
54	Schoenfeld Residuals for ffm with n=45 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	137
55	Schoenfeld Residuals for ffm with n=45 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	137
56	Schoenfeld Residuals for Age with n=80 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	138
57	Schoenfeld Residuals for Age with n=80 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	138
58	Schoenfeld Residuals for Mismatch Score with n=80 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	139
59	Schoenfeld Residuals for Mismatch Score with n=80 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	139
60	Schoenfeld Residuals for ffage with n=80 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	140
61	Schoenfeld Residuals for ffage with n=80 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	140



62	Schoenfeld Residuals for ffm with n=80 and cp=25% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	141
63	Schoenfeld Residuals for ffm with n=80 and cp=25% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	141
64	Schoenfeld Residuals for Age with n=80 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	142
65	Schoenfeld Residuals for Age with n=80 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	142
66	Schoenfeld Residuals for Mismatch Score with n=80 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	143
67	Schoenfeld Residuals for Mismatch Score with n=80 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	143
68	Schoenfeld Residuals for ffage with n=80 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	144
69	Schoenfeld Residuals for ffage with n=80 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	144
70	Schoenfeld Residuals for ffm with n=80 and cp=50% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	145
71	Schoenfeld Residuals for ffm with n=80 and cp=50% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	145



72	Schoenfeld Residuals for Age with n=80 and cp=75% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	146
73	Schoenfeld Residuals for Age with n=80 and cp=75% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	146
74	Schoenfeld Residuals for Mismatch Score with n=80 and cp=75% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	147
75	Schoenfeld Residuals for Mismatch Score with n=80 and cp=75% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	147
76	Schoenfeld Residuals for ffage with n=80 and cp=75% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	148
77	Schoenfeld Residuals for ffage with n=80 and cp=75% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	148
78	Schoenfeld Residuals for ffm with n=80 and cp=75% of the Cause-Specific Hazards of the ordinary Cox Method in Competing Risks	149
79	Schoenfeld Residuals for ffm with n=80 and cp=75% of the Cause-Specific Hazards of the modified Lunn-McNeil Method in Competing Risks	149

