



**UNIVERSITI PUTRA MALAYSIA**

**DIAGNOSTIK DAN PENGARUH BAGI DATA MANDIRIAN  
DALAM RISIKO BERSAING**

**NOOR AKMA IBRAHIM**

**FSAS 1994 5**

**DIAGNOSTIK DAN PENGARUH BAGI DATA MANDIRIAN  
DALAM RISIKO BERSAING**

**NOOR AKMA IBRAHIM**

**DOKTOR FALSAFAH**

**UNIVERSITI PERTANIAN MALAYSIA**

**1994**



**DIAGNOSTIK DAN PENGARUH BAGI DATA MANDIRIAN  
DALAM RISIKO BERSAING**

**Oleh**

**NOOR AKMA IBRAHIM**

**Disertasi Yang Dikemukakan Sebagai Memenuhi Syarat Untuk  
Mendapatkan Ijazah Doktor Falsafah  
di Fakulti Sains dan Pengajian Alam Sekitar**

**Universiti Pertanian Malaysia**

**November 1994**



## PENGHARGAAN

Saya ingin mengambil kesempatan di ruangan ini untuk merakamkan jutaan terima kasih kepada Dr. Isa bin Daud selaku Pengerusi yang dengan sabarnya telah banyak memberikan kata nasihat, bimbingan dan panduan di dalam penyelidikan dan proses menyediakan tesis ini.

Setinggi-tinggi penghargaan juga ditujukan kepada Professor Madya Dr. Mat Yusoff bin Abdullah dan Dr. Kassim bin Haron selaku ahli Jawatankuasa Penyeliaan atas nasihat yang telah diberikan.

Ribuan terima kasih juga disampaikan kepada Professor Madya Dr. Harun bin Budin dan Professor Madya Dr. Ismail bin Mohammad di atas bantuan dan kata nasihat mereka.

Penghargaan juga dirakamkan kepada pegawai Jabatan Perangkaan Malaysia, Bahagian Demografi di Wisma Pahlawan di atas bantuan membekalkan data untuk tujuan analisis.

Saya juga ingin mengucapkan ribuan terima kasih kepada Universiti Pertanian Malaysia yang telah memberikan saya lepasan cuti belajar dan bantuan kewangan.

Tidak lupa juga ucapan terima kasih buat kedua ibu bapa saya serta keluarga tersayang di atas segala-galanya.

## KANDUNGAN

	Mukasurat
<b>PENGHARGAAN</b> .....	ii
<b>SENARAI JADUAL</b> .....	vi
<b>SENARAI RAJAH</b> .....	vii
<b>ABSTRAK</b> .....	xiv
<b>ABSTRACT</b> .....	xvii
 <b>BAB</b>	
<b>I</b>	
<b>Pengenalan</b> .....	1
Latar Belakang .....	1
Ringkasan Keputusan .....	11
<b>II</b>	
<b>Analisis Mandirian</b> .....	16
Data Mandirian dan Masa Kegagalan .....	18
Tapisan .....	18
Kovariat di dalam Analisis Mandirian .....	20
Beberapa Jenis Masa Kegagalan .....	21
Taburan Masa Kegagalan .....	22
Taburan Eksponen .....	25
Taburan Weibull .....	26
Model Regresi .....	27
Model Regresi Eksponen dan Weibull .....	27
Model Kadar Bahaya .....	31
Model Masa Kegagalan Terpecut .....	33
Model Regresi Mudah bagi Data Tertapis .....	35
Model dan Andaian .....	38
Anggaran Parameter .....	40
Komputeran - Penggunaan Perisian SAS .....	41
Contoh - Data Motor Kecil .....	43

<b>III</b>	<b>RISIKO BERSAING</b> .....	50
	Tata-tanda dan Formulasi .....	52
	Pendekatan Berparameter .....	58
	Kes Semua Hayat dan Risiko Diketahui .....	59
	Kes Hayat Tertapis .....	62
	Kes Hayat Dikumpul ke dalam Selang .....	63
	Taburan Tertentu - Risiko Merdeka .....	65
	Taburan Eksponen .....	65
	Taburan Weibull .....	69
	Risiko Bersaing Dengan Kovariat .....	70
	Kaedah Berparameter dengan Tapisan sebagai Risiko .....	71
	Anggaran dan Pentakbiran .....	72
	Pendekatan Kompromi Marquardt Terubahsuai ...	73
	Kebarangkalian Intrinsik .....	78
	Taburan Eksponen dengan Data Terkumpul .....	81
	Taburan Bagi Cerapan Tertapis .....	90
	Contoh .....	90
	Data Barah Oropharinks .....	91
	Data Pemindahan Jantung Stanford .....	99
	Data Kemortalan Bayi .....	102
<b>IV</b>	<b>DIAGNOSTIK DAN PENGARUH</b> .....	115
	Ukuran Pengaruh .....	117
	Jarak Cook .....	119
	Jarak Kebolehjadian .....	121
	Kaedah Satu Langkah .....	122
	Lelaran Satu Langkah .....	123
	Pengahmipan Satu Langkah .....	124
	Contoh .....	124
	Data Motor Kecil .....	125
	Data Barah Oropharinks .....	130
	Data Pemindahan Jantung Stanford .....	134
	Data Kemortalan Bayi .....	137
<b>V</b>	<b>KAJIAN SIMULASI</b> .....	142
	Parameter Untuk Simulasi .....	142



Menjana Kes .....	143
Rangka Analisis .....	144
Keputusan Analisis .....	146
<b>VI KOVARIAT BERSANDARAN KEPADA MASA ..</b>	<b>186</b>
Analisis Reja .....	188
Reja Bagi Risiko Bersaing .....	189
Contoh .....	193
Data Barah Oropharinks .....	193
Data Pemindahan Jantung Stanford .....	197
Data Simulasi .....	199
<b>VII KESIMPULAN .....</b>	<b>208</b>
Ringkasan .....	208
Kesimpulan .....	212
Cadangan Untuk Kajian Masa Hadapan .....	216
<b>BIBLIOGRAFI .....</b>	<b>219</b>
<b>LAMPIRAN .....</b>	<b>228</b>
A .....	229
B .....	235
<b>LATAR DIRI .....</b>	<b>238</b>

## SENARAI JADUAL

Jadual		Mukasurat
1	Anggaran Kuasa Dua Terkecil bagi Contoh Hayat Penebatan Motor Kecil (dengan pembedaan bagi cerapan tertapis) .....	47
2	Anggaran Parameter pada Lelaran Terpilih .....	48
3	Anggaran Parameter dan Ralat Piawai bagi Risiko Pertama Data Barah Oropharynx .....	94
4	Anggaran Parameter dan Ralat Piawai bagi Risiko Kedua Data Barah Oropharynx .....	95
5	Anggaran Parameter, Ralat Piawai dan nilai $P$ ....	96
6	Anggaran Kebarangkalian Kasar, $\hat{Q}_i$ .....	97
7	Anggaran Parameter, Ralat Piawai dan Nilai $P$ ...	98
8	Anggaran Parameter dan Ralat Piawai bagi Risiko Pertama Data Pemindahan Jantung Stanford .....	100
9	Anggaran Parameter dan Ralat Piawai bagi Risiko Kedua Data Pemindahan Jantung Stanford .....	100
10	Anggaran Parameter, Ralat Piawai dan Nilai $P$ ....	102
11	Anggaran Parameter, Ralat Piawai dan Nilai $P$ bagi Data Kemortalan Bayi .....	107
12	Anggaran Kebarangkalian Intrinsik, $\hat{Q}_{1\alpha}$ , $\hat{q}_{1\alpha}$ , $\hat{Q}_{1\alpha(2)}$ , $\hat{Q}_{2\alpha(1)}$ bagi Data Kemortalan Bayi .....	112
13	Nilai Jarak Cook dan Reja bagi Data Motor Kecil ...	128





## SENARAI RAJAH

Rajah		Mukasurat
1	Histogram Kebarangkalian Kegagalan bagi Kovariat Jantina Data Kemortalan Bayi .....	108
2	Histogram Kebarangkalian Kegagalan bagi Kovariat Bangsa Data Kemortalan Bayi - Risiko Pertama .....	109
3	Histogram Kebarangkalian Kegagalan bagi Kovariat Bangsa Data Kemortalan Bayi - Risiko Kedua .....	109
4	Histogram Kebarangkalian Kegagalan bagi Kovariat Kediaman Data Kemortalan Bayi - Risiko Pertama .....	110
5	Histogram Kebarangkalian Kegagalan bagi Kovariat Kediaman Data Kemortalan Bayi - Risiko Kedua .....	110
6	Jarak Cook bagi Data Motor Kecil .....	129
7	Plot Reja bagi Data Motor Kecil .....	129
8	Jarak Cook bagi Data Barah Oropharinks (risiko pertama) .....	131
9	Jarak Kebolehjadian bagi Data Barah Oropharinks (risiko pertama) .....	132
10	Jarak Cook bagi Data Barah Oropharinks (risiko kedua) .....	132
11	Jarak Kebolehjadian bagi Data Barah Oropharinks (risiko kedua) .....	133
12	Jarak Cook bagi Data Stanford (risiko pertama) .....	135
13	Jarak Kebolehjadian bagi Data Stanford (risiko pertama) .....	136
14	Jarak Cook bagi Data Stanford (risiko kedua) .....	136
15	Jarak Kebolehjadian bagi Data Stanford (risiko kedua) .....	137



16	Jarak Cook bagi Data Kemortalan Bayi (risiko pertama) .....	139
17	Jarak Cook bagi Data Kemortalan Bayi (risiko kedua) .....	139
18	Jarak Cook bagi 20 cerapan ke 5 (eksponen) .....	147
19	Jarak Cook bagi 20 cerapan ke 10 (eksponen) .....	148
20	Jarak Cook bagi 20 cerapan ke 15 (eksponen) .....	148
21	Jarak Cook bagi 20 cerapan ke 20 (eksponen) .....	148
22	Jarak Cook bagi 20 cerapan ke 25 (eksponen) .....	149
23	Jarak Cook bagi 20 cerapan ke 30 (eksponen) .....	149
24	Jarak Cook bagi 30 cerapan ke 4 (eksponen) .....	149
25	Jarak Cook bagi 30 cerapan ke 9 (eksponen) .....	150
26	Jarak Cook bagi 30 cerapan ke 14 (eksponen) .....	150
27	Jarak Cook bagi 30 cerapan ke 19 (eksponen) .....	150
28	Jarak Cook bagi 30 cerapan ke 24 (eksponen) .....	151
29	Jarak Cook bagi 30 cerapan ke 29 (eksponen) .....	151
30	Jarak Cook bagi 50 cerapan ke 3 (eksponen) .....	151
31	Jarak Cook bagi 50 cerapan ke 8 (eksponen) .....	152
32	Jarak Cook bagi 50 cerapan ke 13 (eksponen) .....	152
33	Jarak Cook bagi 50 cerapan ke 18 (eksponen) .....	152
34	Jarak Cook bagi 50 cerapan ke 23 (eksponen) .....	153
35	Jarak Cook bagi 50 cerapan ke 28 (eksponen) .....	153
36	Jarak Cook bagi 100 cerapan ke 2 (eksponen) .....	153



37	Jarak Cook bagi 100 cerapan ke 7 (eksponen)	.....	154
38	Jarak Cook bagi 100 cerapan ke 12 (eksponen)	.....	154
39	Jarak Cook bagi 100 cerapan ke 17 (eksponen)	.....	154
40	Jarak Cook bagi 100 cerapan ke 22 (eksponen)	.....	155
41	Jarak Cook bagi 100 cerapan ke 27 (eksponen)	.....	155
42	Jarak Cook bagi 20 cerapan pertama (Weibull)	.....	158
43	Jarak Cook bagi 20 cerapan ke 2 (Weibull)	.....	158
44	Jarak Cook bagi 20 cerapan ke 3 (Weibull)	.....	158
45	Jarak Cook bagi 20 cerapan ke 4 (Weibull)	.....	159
46	Jarak Cook bagi 20 cerapan ke 5 (Weibull)	.....	159
47	Jarak Cook bagi 20 cerapan ke 6 (Weibull)	.....	159
48	Jarak Cook bagi 20 cerapan ke 7 (Weibull)	.....	160
49	Jarak Cook bagi 20 cerapan ke 8 (Weibull)	.....	160
50	Jarak Cook bagi 20 cerapan ke 9 (Weibull)	.....	160
51	Jarak Cook bagi 20 cerapan ke 10 (Weibull)	.....	161
52	Jarak Cook bagi 20 cerapan ke 11 (Weibull)	.....	161
53	Jarak Cook bagi 20 cerapan ke 12 (Weibull)	.....	161
54	Jarak Cook bagi 20 cerapan ke 13 (Weibull)	.....	162
55	Jarak Cook bagi 20 cerapan ke 14 (Weibull)	.....	162
56	Jarak Cook bagi 20 cerapan ke 15 (Weibull)	.....	162
57	Jarak Cook bagi 20 cerapan ke 16 (Weibull)	.....	163
58	Jarak Cook bagi 20 cerapan ke 17 (Weibull)	.....	163



59	Jarak Cook bagi 20 cerapan ke 18 (Weibull)	.....	163
60	Jarak Cook bagi 20 cerapan ke 19 (Weibull)	.....	164
61	Jarak Cook bagi 20 cerapan ke 20 (Weibull)	.....	164
62	Jarak Cook bagi 30 cerapan pertama (Weibull)	.....	164
63	Jarak Cook bagi 30 cerapan ke 2 (Weibull)	.....	165
64	Jarak Cook bagi 30 cerapan ke 3 (Weibull)	.....	165
65	Jarak Cook bagi 30 cerapan ke 4 (Weibull)	.....	165
66	Jarak Cook bagi 30 cerapan ke 5 (Weibull)	.....	166
67	Jarak Cook bagi 30 cerapan ke 6 (Weibull)	.....	166
68	Jarak Cook bagi 30 cerapan ke 7 (Weibull)	.....	166
69	Jarak Cook bagi 30 cerapan ke 8 (Weibull)	.....	167
70	Jarak Cook bagi 30 cerapan ke 9 (Weibull)	.....	167
71	Jarak Cook bagi 30 cerapan ke 10 (Weibull)	.....	167
72	Jarak Cook bagi 30 cerapan ke 11 (Weibull)	.....	167
73	Jarak Cook bagi 30 cerapan ke 12 (Weibull)	.....	168
74	Jarak Cook bagi 30 cerapan ke 13 (Weibull)	.....	168
75	Jarak Cook bagi 30 cerapan ke 14 (Weibull)	.....	169
76	Jarak Cook bagi 30 cerapan ke 15 (Weibull)	.....	169
77	Jarak Cook bagi 30 cerapan ke 16 (Weibull)	.....	169
78	Jarak Cook bagi 30 cerapan ke 17 (Weibull)	.....	170
79	Jarak Cook bagi 30 cerapan ke 18 (Weibull)	.....	170
80	Jarak Cook bagi 30 cerapan ke 19 (Weibull)	.....	170



81	Jarak Cook bagi 30 cerapan ke 20 (Weibull) .....	171
82	Jarak Cook bagi 50 cerapan pertama (Weibull) .....	171
83	Jarak Cook bagi 50 cerapan ke 2 (Weibull) .....	171
84	Jarak Cook bagi 50 cerapan ke 3 (Weibull) .....	172
85	Jarak Cook bagi 50 cerapan ke 4 (Weibull) .....	172
86	Jarak Cook bagi 50 cerapan ke 5 (Weibull) .....	172
87	Jarak Cook bagi 50 cerapan ke 6 (Weibull) .....	173
88	Jarak Cook bagi 50 cerapan ke 7 (Weibull) .....	173
89	Jarak Cook bagi 50 cerapan ke 8 (Weibull) .....	173
90	Jarak Cook bagi 50 cerapan ke 9 (Weibull) .....	174
91	Jarak Cook bagi 50 cerapan ke 10 (Weibull) .....	174
92	Jarak Cook bagi 50 cerapan ke 11 (Weibull) .....	174
93	Jarak Cook bagi 50 cerapan ke 12 (Weibull) .....	175
94	Jarak Cook bagi 50 cerapan ke 13 (Weibull) .....	175
95	Jarak Cook bagi 50 cerapan ke 14 (Weibull) .....	175
96	Jarak Cook bagi 50 cerapan ke 15 (Weibull) .....	176
97	Jarak Cook bagi 50 cerapan ke 16 (Weibull) .....	176
98	Jarak Cook bagi 50 cerapan ke 17 (Weibull) .....	176
99	Jarak Cook bagi 50 cerapan ke 18 (Weibull) .....	177
100	Jarak Cook bagi 50 cerapan ke 19 (Weibull) .....	177
101	Jarak Cook bagi 50 cerapan ke 20 (Weibull) .....	177
102	Jarak Cook bagi 100 cerapan pertama (Weibull) .....	178



103	Jarak Cook bagi 100 cerapan ke 2 (Weibull)	.....	178
104	Jarak Cook bagi 100 cerapan ke 3 (Weibull)	.....	178
105	Jarak Cook bagi 100 cerapan ke 4 (Weibull)	.....	179
106	Jarak Cook bagi 100 cerapan ke 5 (Weibull)	.....	179
107	Jarak Cook bagi 100 cerapan ke 6 (Weibull)	.....	179
108	Jarak Cook bagi 100 cerapan ke 7 (Weibull)	.....	180
109	Jarak Cook bagi 100 cerapan ke 8 (Weibull)	.....	180
110	Jarak Cook bagi 100 cerapan ke 9 (Weibull)	.....	180
111	Jarak Cook bagi 100 cerapan ke 10 (Weibull)	.....	181
112	Jarak Cook bagi 100 cerapan ke 11 (Weibull)	.....	181
113	Jarak Cook bagi 100 cerapan ke 12 (Weibull)	.....	181
114	Jarak Cook bagi 100 cerapan ke 13 (Weibull)	.....	182
115	Jarak Cook bagi 100 cerapan ke 14 (Weibull)	.....	182
116	Jarak Cook bagi 100 cerapan ke 15 (Weibull)	.....	182
117	Jarak Cook bagi 100 cerapan ke 16 (Weibull)	.....	183
118	Jarak Cook bagi 100 cerapan ke 17 (Weibull)	.....	183
119	Jarak Cook bagi 100 cerapan ke 18 (Weibull)	.....	183
120	Jarak Cook bagi 100 cerapan ke 19 (Weibull)	.....	184
121	Jarak Cook bagi 100 cerapan ke 20 (Weibull)	.....	184
122	Plot Reja bagi Kovariat Umur	.....	195
123	Plot Reja bagi Keadaan Pesakit	.....	196
124	Plot Reja bagi Kovariat Peringkat 1	.....	196



125	Plot Reja bagi Kovariat Umur	.....	198
126	Plot Reja bagi Kovariat Pemindahan	.....	199
127	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	200
128	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	201
129	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	202
130	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	202
131	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	203
132	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	203
133	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	204
134	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	204
135	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	205
136	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	205
137	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	206
138	Plot $\hat{r}_{i,j}$ terhadap pangkat $t_{i,j}$	.....	206



Abstrak disertasi yang dikemukakan kepada Senat Universiti Pertanian Malaysia bagi memenuhi keperluan untuk Ijazah Doktor Falsafah.

**DIAGNOSTIK DAN PENGARUH BAGI DATA MANDIRIAN  
DALAM RISIKO BERSAING**

Oleh:

**NOOR AKMA IBRAHIM**

**NOVEMBER 1994**

Pengerusi : Dr. Isa bin Daud

Fakulti : Fakulti Sains dan Pengajian Alam Sekitar

Kajian di dalam tesis ini adalah mengenai perkembangan dan lanjutan bagi teknik penilaian diagnostik dan pengaruh dalam data mandirian yang merangkumi cerapan tertapis. Model mandirian yang dianalisis bermula dengan model hayat terpecut dengan risiko tunggal dikaji secara ringkas. Tumpuan analisis ialah ke atas model risiko bersaing termasuk data terkumpul. Diagnostik pengaruh dilakukan berdasarkan kaedah penghapusan kes manakala ukuran skalar untuk menilai pengaruh berpandukan ukuran jarak Cook (1972). Beberapa bentuk ukuran Cook telah dipertimbangkan untuk menilai pengaruh dalam model hayat terpecut yang mana model ini lazim digunakan dengan meluas untuk menganalisis data kebolehppercayaan. Melalui bantuan pakej komputer yang sesuai didapati kesemua ukuran ini memberikan persetujuan ke atas keputusan yang melibatkan cerapan yang dipercayai berpengaruh.

Model risiko bersaing yang dikaji secara amnya merupakan suatu model mandirian yang mana penilaian ke atas sesuatu risiko tertentu diselidiki di dalam





situasi yang kompleks dengan kehadiran  $m$  risiko yang lain. Walau bagaimanapun penekanan di sini adalah ke atas model risiko bersaing dengan kehadiran dua risiko ( $m = 2$ ) dengan tumpuan kepada kes tapisan sebagai salah satu daripada risiko ini. Di samping itu analisis kepada data terkumpul telah dilaksanakan berdasarkan risiko bersaing dengan  $m = 2$ . Proses transformasi telah digunakan bagi menangani masalah penyuaian model risiko bersaing terutamanya bagi data terkumpul dengan kriteria penumpuan tercapai melalui pendekatan lelaran Marquardt.

Penilaian pengaruh dalam model linear telah dilanjutkan bagi menampung keperluan model risiko bersaing. Ukuran jarak Cook dan jarak kebolehdjian dikaji bagi mengesan kes-kes pengaruh bagi data di dalam model ini. Oleh kerana pengiraan ukuran pengaruh melibatkan suatu proses yang agak panjang, kami sarankan kegunaan teknik lelaran satu-langkah dan ianya diperkembangkan menggunakan pendekatan kompromi Marquardt terubahsuai. Melalui diagnostik penilaian pengaruh dengan penghapusan dan lelaran satu-langkah ini beberapa cerapan berpengaruh telah dapat dikenal pasti.

Penilaian kaedah diagnostik dan pengaruh bagi model risiko ini diselidiki dengan melakukan kajian simulasi ke atas data yang dijana. Dua kajian simulasi dijalankan bagi menentusahkan kesahihan model yang digunakan. Di samping itu kaedah diagnostik yang dibincang di atas dilaksanakan. Gambarajah yang diperolehi daripada keputusan menunjukkan diagnostik pengaruh ini mampu mengesan kes-kes pengaruh lantas memperkuat lagi kaedah.

Kepentingan kovariat yang bersekutu dengan masa kegagalan yang diperkenalkan di dalam model risiko bersaing melalui fungsi bahaya dikaji selan-

jutnya dengan meneliti salah satu daripada kebarangkalian kegagalan intrinsik risiko bersaing. Secara kasar kebarangkalian ini dapat memperteguhkan keputusan ke atas kovariat terhadap kesignifikanannya di dalam model.

Kovariat yang disyaki bersandaran kepada masa diselidiki menggunakan anggaran reja yang dibina dan disesuaikan mengikut keperluan risiko bersaing berlandaskan reja Schoenfeld. Dengan kaedah ini kesahan dan kesesuaian model boleh dikesan.

Abstract of dissertation submitted to the Senate of Universiti Pertanian Malaysia in fulfilment of the requirements for the degree of Doctor of Philosophy.

**DIAGNOSTICS AND INFLUENCE FOR SURVIVAL DATA  
IN COMPETING RISKS**

**By:**

**NOOR AKMA IBRAHIM**

**NOVEMBER 1994**

Chairman : Dr. Isa Bin Daud

Faculty : Faculty of Science and Environmental Studies

The work in this thesis is concerned with the development and extension of techniques for the assessment of diagnostics and influence in survival data encompassing censored observations. The survival models analysed are the accelerated life-time model with one risk which is done in brief, while emphasis is on competing risks models that include grouped data. The influence diagnostics is based on the case deletion method while the scalar measurement to assess influence is based on Cook (1972) distance measure. Various forms of the Cook's measurements are considered for the accelerated life-time model which has been used quite extensively in the analysis of reliability data. Through the aid of a suitable computer package it is found that all measures are in agreement with one another with respect to the observations suspected as influential.

The competing risks models being studied in general are concerned with the assessment of a specific risk in the complicating presence of  $m$  other risks.



However, the emphasis here is on models in the presence of two risks ( $m = 2$ ), treating censoring as one of the risks. In addition, the analysis of grouped data is performed based on competing risks with  $m = 2$ . Transformations have been used to overcome the problem in fitting the competing risks models especially for the grouped data with the convergence criteria met by the Marquardt iteration approach.

The assessment of influence in linear model is extended to accommodate competing risks models with the measurement of influence based on the Cook's distance and the likelihood distance. Due to the long computation process of the influence measures, a one-step iteration technique is proposed and developed using the modified Marquardt compromise approach. Through this influence diagnostics assessment with deletion and the one-step iteration several observations are identified as influential.

Assessment of this diagnostics and influence method is investigated by performing simulation study on generated data. Two simulation studies are carried out to confirm the suitability of the models. Diagnostic methods as discussed above are then implemented. Diagrams obtained from the results indicate that this influence diagnostics is able to detect influential cases, thus substantiating the method.

The importance of the covariates associated with the failure times introduced in the model via the hazard function is further investigated by examining one of the intrinsic probability of failures in competing risks. These probabilities can consolidate the decision made on these covariates with regards of their significance in the model.



The time dependence covariates are investigated utilizing the residual estimates formulated and modified according to the suitability in competing risks based on Schoenfeld's residuals. With this, the validity of the model can be detected.



## **BAB 1**

### **Pengenalan**

#### **Latar Belakang**

Tahun 1960-an memperlihatkan perkembangan regresi dengan kebanyakan penyelidikan tertumpu kepada peranan reja untuk mengesan kekurangan, kelemahan, atau ketidakseimbangan sesuatu model. Reja juga berperanan untuk memperlihatkan titik-titik data yang ganjil. Plot seperti plot reja terhadap nilai ramalan dan plot reja terhadap pembolehubah ramalan disarankan. Ujian terhadap reja telah menjadi sebahagian diterap di dalam kebanyakan analisis statistik dengan adanya kemudahan program komputer. Sehingga tahun 1970-an makalah-makalah yang diterbitkan adalah berlandaskan tema ini. Walau bagaimanapun masalah masih timbul di mana reja gagal memenuhi andaian kenormalan. Masalah ini menggerakkan kegunaan anggaran teguh di dalam pemasalahan regresi. Prosedur teguh ini dapat mengurangkan penonjolan cerapan atau kes yang dikesan sebagai titik terpencil (cerapan yang menghasilkan nilai reja yang besar) di dalam plot reja.

Di penghujung tahun 1970-an, usaha adalah ke arah mengesan titik terpencil dan kes yang ekstrim. Prosedur ini dikenali sebagai diagnostik regresi dan ia berperanan mengesan kes-kes yang berkemungkinan mempunyai pengaruh ke atas anggaran regresor. Ia juga merupakan suatu statistik direka-bentuk untuk membantu para penyelidik membuat keputusan sama ada mereka dapat mempertahankan andaian yang diandaikan didalam proses penyuaian



model. Teks oleh Cook (1977,1979), Andrews dan Pregibon (1978), Cook dan Weisberg (1980), Belsley *et al.* (1980) dan Cook dan Weisberg (1982) memperkenalkan beberapa ukuran diagnostik untuk mengesan dan mengenal pasti kes secara individu atau secara kumpulan. Kes ini dianggap berpengaruh jika ianya berbeza daripada sebahagian besar data.

Di dalam analisis kuasa dua terkecil bagi data berdasarkan model linear berpangkat penuh, suatu cerapan boleh di kira berpengaruh jika penghapuannya daripada analisis menjejaskan ciri penting analisis tersebut. Penyelidikan lepas menunjukkan yang pengaruh sesuatu cerapan walaupun tidak secara menyeluruh dapat diterangkan melalui reja dan varians reja yang berkenaan.

Pertimbangkan model

$$\underline{Y} = X\underline{\beta} + \underline{\epsilon} \quad [1.1]$$

dengan  $\underline{Y}$  vektor cerapan  $n \times 1$ ,  $X$  matriks berpangkat penuh  $n \times s$  bagi pemalar yang diketahui,  $\underline{\beta}$  vektor parameter  $s \times 1$  yang tak diketahui, dan  $\underline{\epsilon}$  vektor pembolehubah rawak merdeka dengan setiap satu mempunyai min sifar dan varians  $\sigma^2$ . Anggaran kuasa dua terkecil bagi  $\underline{\beta}$  adalah

$$\hat{\underline{\beta}} = (X'X)^{-1} X'Y$$

Vektor reja yang sepadan adalah

$$\underline{R} = \underline{Y} - \hat{\underline{Y}} = \underline{Y} - X\hat{\underline{\beta}}$$

Cook (1977) mencadangkan yang pengaruh titik data ke  $j$  boleh diuji menggunakan ukuran jarak

$$D_j = [(\hat{\beta} - \hat{\beta}_{(j)})' X' X (\hat{\beta} - \hat{\beta}_{(j)})] / (s\sigma_*^2) \quad [1.2]$$

$$j = 1, \dots, n$$

$\hat{\beta}$  menandakan anggaran bagi  $\beta$  tanpa menghapuskan titik data ke  $j$  manakala  $\hat{\beta}_{(j)}$  menandakan anggaran bagi  $\beta$  dengan penghapusan titik data ke  $j$  dan  $\sigma_*^2 = R' R / (n - s)$ . Nilai  $D_j$  besar menunjukkan yang titik ke  $j$  yang bersekutu dengannya mempunyai pengaruh yang kuat ke atas anggaran bagi  $\beta$ . Magnitud bagi jarak di antara  $\hat{\beta}$  dan  $\hat{\beta}_{(j)}$  boleh dinilai dengan membandingkan  $D_j$  dengan taburan memusat  $F$  bersama darjah kebebasan  $s$  dan  $n - s$ . Persamaan [1.2] ini menjadi asas kepada kebanyakan ukuran jarak untuk mengesan pengaruh sesuatu cerapan atau kes. Sekiranya ukuran bagi kes yang dihapuskan memperlihatkan beza apabila dibandingkan dengan ukuran yang lain dipercayai kes ini mempunyai pengaruh terhadap model. Penghapusan kes demi kes ini merupakan suatu gangguan terhadap formulasi sesuatu model. Jika gangguan kecil ke atas penganggar mempengaruhi keputusan analisis, ini seharusnya dikhuatiri tetapi jika sebaliknya berlaku, sampel dianggap teguh terhadap gangguan yang disengajakan dan kejahilan kita memilih model yang tepat tidak memudaratkan. Umumnya kajian ke atas keputusan dan pentakbiran yang bergantung kepada pelbagai aspek masalah formulasi dipanggil kajian terhadap pengaruh. Keupayaan untuk mengenal pasti kes pengaruh memberikan manfaat dari segi kebolehpercayaan keputusan ke atas model yang diandaikan.

Dari bahagian sebelum ini jelas bahawa ukuran diagnostik diformulasikan berdasarkan model linear dengan kaedah kuasa dua terkecil. Kebanyakan penyelidikan dan penulisan berkenaan diagnostik pengaruh yang terhasil adalah berdasarkan model regresi linear ini. Cook (1977) memperkenalkan suatu



statistik bagi mengukur pengaruh sesuatu cerapan terhadap model yang tertentu dan ia juga menunjukkan yang statistik ini mengandungi maklumat sama ada cerapan tersebut terpencil atau sebaliknya. Andrews dan Pregibon (1978) menyarankan suatu statistik untuk mengenal pasti satu atau lebih cerapan sebagai titik terpencil, berpengaruh atau kedua-duanya sekali. Draper dan John (1981) menunjukkan yang statistik Cook (1977) jika digabungkan dengan komponen di dalam statistik Andrews dan Pregibon (1978) dapat memberikan maklumat bukan sahaja ke atas titik terpencil dan cerapan berpengaruh tetapi juga ke atas kejauhan (remoteness) cerapan di dalam ruang faktor. Pregibon (1981) membincangkan kaedah mengenal pasti cerapan berpengaruh melalui penyuaian kebolehdajadian maksimum ke atas model regresi logistik. Weisberg (1983) membincangkan beberapa prinsip bagi diagnostik regresi dan analisis pengaruh. Cook dan Wang (1983) mengkaji perubahan dalam parameter transformasi bagi pembolehubah sambutan regresi linear sementara Atkinson (1983) memberi perhatian terhadap analisis regresi diagnostik dengan transformasi kuasa ke atas model. Perlu dinyatakan di sini bahawa sesuatu cerapan yang muncul sebagai titik terpencil setelah model disesuaikan kepada data tidak bererti cerapan ini mempunyai pengaruh terhadap model. Ini bererti jika cerapan ini dikeluarkan daripada analisis, penyuaian persamaan tidak akan terjejas (Andrews dan Pregibon, 1978).

Diagnostik pengaruh yang telah menjadi popular dari segi perlaksanaannya ialah Cook's D, DFBETAS dan DFFITS (lihat Belsley et al. 1980; Cook dan Weisberg, 1982). Ukuran-ukuran jarak ini dibentuk melalui reja yang dibahagikan dengan ralat piawai (studentized residuals) dan matriks diagonal bagi cerapan ke  $j$  dari matriks Hessian ( $H = X(X'X)^{-1}X'$ ). Diagnostik pengaruh