

國立中央大學

統計研究所

博士論文

加權 Kaplan-Meier 統計量之推廣

指導教授：陳玉英 博士

研究生：李維鈞

中華民國九十年六月



國立中央大學圖書館 碩博士論文授權書

本授權書所授權之論文全文與電子檔，為本人於國立中央大學，撰寫之碩/博士學位論文。(以下請擇一勾選)

- () 同意 (立即開放)
() 同意 (一年後開放)，原因是： _____
() 同意 (二年後開放)，原因是： _____
() 不同意，原因是： _____

授與國立中央大學圖書館，基於推動讀者間「資源共享、互惠合作」之理念，於回饋社會與學術研究之目的，得不限地域、時間與次數，以紙本、光碟、網路或其它各種方法收錄、重製、與發行，或再授權他人以各種方法重製與利用。

研究生簽名： 李維鈞

論文名稱： 加權 Kaplan-Meier 統計量之推廣

指導教授姓名： 陳玉英

系所： 統計 所 博士 碩士 班

學號： 83225009

日期：民國 90 年 6 月 22 日

備註：

1. 本授權書親筆填寫後（電子檔論文可用電腦打字），請影印裝訂於紙本論文書名頁之次頁，未附本授權書，圖書館將不予驗收。
2. 上述同意與不同意之欄位若未勾選，本人同意視同授權立即開放。

誌 謝

衷心感謝指導教授陳玉英博士多年來的悉心教導及指正，使得本篇論文得以順利完成，並感謝所上老師和口試委員們的指導，及學長、同學和學弟妹的相互砥礪，使我對統計有更深一層的認識與了解。更感謝我的家人們給我的支持及鼓勵，讓我得以專心的學習及完成學業。

僅將本文獻給長期支持我的家人：媽媽、妻子昭如、昱霖和佩芹，及所有指導我的師長們。

李維鈞僅識於
國立中央大學統計研究所
中華民國九十年六月

摘 要

本論文推廣應用雙樣本加權 Kaplan-Meier 統計量(Pepe和Fleming, 1989)於具有隨機右設限存活資料的雙樣本和多樣本相關問題。首先修正單尾加權 Kaplan-Meier 檢定為雙尾之檢定。其次，在雙樣本檢定問題中，為考量共變數之影響，並且能夠及早結束動物實驗或臨床試驗，推廣 Li (1999)所提之群序列雙樣本加權 Kaplan-Meier 檢定，建立調整共變數後之群序列雙樣本加權 Kaplan-Meier 檢定。最後，就多組具有右設限的存活資料，檢定多樣本問題中有關存活函數之傘形對立假說。針對傘頂已知或未知情形，推廣雙樣本加權 Kaplan-Meier 統計量建構適當之檢定方法。本論文使用蒙地卡羅方法在具有早期、中期、晚期、等比例或交叉危險函數差異的成段指數分布下，研究本文所提檢定方法在小樣本情形下的水準及檢定力表現。最後，引用適當的數據加以分析。

目 錄

第一章	緒論	1
1.1	研究動機及目的	1
1.2	加權 Kaplan-Meier 統計量	3
1.3	本文結構	8
第二章	Renyi-type 加權 Kaplan-Meier 檢定	9
2.1	Renyi-type 加權對數秩檢定	9
2.2	Renyi-type 加權 Kaplan-Meier 檢定	9
2.3	模擬研究	11
第三章	調整共變數之加權平均存活群序列檢定	20
3.1	簡介及文獻回顧	20
3.1.1	加權平均存活統計量	20
3.1.2	雙樣本加權 Kaplan-Meier 群序列檢定	21
3.1.3	調整共變數後之加權對數秩群序列檢定	23
3.2	方法推導與建立	24
3.3	實例分析	35
第四章	加權 Kaplan-Meier 傘檢定	37
4.1	加權對數秩傘檢定	37

4.2	傘頂已知之傘檢定	39
4.3	傘頂未知之傘檢定	41
4.4	傘頂已知之漸進相對效率	44
4.5	模擬研究	48
4.6	實例分析	52
第五章	結論與未來研究方向	59
5.1	結論	59
5.2	未來研究方向	59
參考文獻	60

圖 表 目 錄

圖 1-1 不同存活函數型態之雙樣本單尾對立假說.....	7
圖 2-1 $n=30$, RWKM和 RLR之模擬抽樣分布	15
圖 2-2 存活函數相交之對立假說	16
圖 4-1 Survival configuration for alternatives	47
圖 4-2 The Kaplan-Meier estimates for the injection sites-transfer data	53
圖 4-3 The cumulative hazard functions for the injection sites-transfer data	53
圖 4-4(a) 前列腺癌資料，癌症等級三之存活函數估計及對數累積危險函數圖	56
圖 4-4(b) 前列腺癌資料，癌症等級四之存活函數估計及對數累積危險函數圖	58
表 1-1 名目水準 $\alpha = 0.05$ ，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之水準估計	5
表 1-2 名目水準 $\alpha = 0.05$ ，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之檢定力估計	6
表 2-1 名目水準 $\alpha = 0.05$ ，存活分布為標準指數分布，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之靴環法之水準估計	12
表 2-2 名目水準 $\alpha = 0.05$ ，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之靴環法之檢定力估 計	13
表 3-1 the prostatic cancer data 之群序列檢定統計量， $\alpha = 0.1$	36
表 4-1 在相同設限分布和相同樣本數下 $z_{\max}(k, \alpha)$	44
表 4-2 在 $k = 5$ ， $p = 5$ ，相同樣本數和共變數為 (q_1, \dots, q_5) 且設限分布為 $U(0, R)$ 之 ARE(WKM,WLR)值	46

表 4-3 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之水準估計	49
表 4-4 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之檢定力估計	50
表 4-5 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之檢定力估計	51
表 4-6 the injection sites-transfer data 之傘檢定統計量	52
表 4-7 the prostatic cancer data 之傘檢定統計量 the cancer stage = 3	55
表 4-8 the prostatic cancer data 之傘檢定統計量 the cancer stage = 4	57

第一章 緒論

1.1 研究動機及目的

自從 1958 年 Kaplan 和 Meier 提出隨機右設限(censoring)資料的存活函數(survival function)估計之後，不完整資料分析方法不斷推陳出新。除單一樣本的分析之外，雙樣本問題中有關危險函數(hazard function)差異的檢定問題，基本上有加權對數秩(weighted logrank, 記為 WLR)檢定(Harrington 和 Fleming, 1982)可資應用，其權數決定於危險函數差異發生的時段與型態。當危險函數交叉時，也有因應的 Renyi-type WLR 雙尾檢定(Gill, 1980)。因為雙樣本加權對數秩統計量衡量的是危險函數之差異，其中權數之決定也與危險函數差異有關，所以基於 WLR 檢定對於存活函數差異的不靈敏缺失，Pepe 和 Fleming (1989, 1991)進一步根據兩組 Kaplan-Meier 存活函數估計差異，提出加權 Kaplan-Meier(weighted Kaplan-Meier, 記為 WKM)統計量用於 $H_0: (S_1(t) = S_2(t) \text{ for all } t)$ 相對於 $H_1: (S_1(t) \leq S_2(t) \text{ for all } t \text{ and } S_1(t) < S_2(t) \text{ for some } t)$ 的檢定，其中權數不因存活或危險函數型態而異。此一 WKM 統計量可視為 Kolmogorov-Smirnov 檢定和 Cramer-von Mises 檢定之延伸。其研究結果指出，在比例危險模式(proportional hazards model)之下，雙樣本 WKM 檢定與對數秩(logrank 記為 LR)檢定具有相當之效率(efficiency)，但對早期即呈現差異或交叉之危險函數而言， WKM 檢定明顯優於對數秩檢定。基於這些原因，所以，下述所提出之統計量均是以雙樣本 WKM 統計量為建構之基礎。

因為 Pepe 和 Fleming (1989, 1991)研究中的比較對象僅為 LR 檢定，然而在 WLR 檢定族中包含針對危險函數差異主要發生在早期或晚期的檢定，例如：針對差異發生在早期時之 Peto-Prentice-Wilcoxon 檢定(Peto 和 Peto, 1972; Prentice, 1978)等。此外，當危險函數交叉時，也有 Renyi-type WLR 檢定可資應用。因此，為求比較之周延，本文的第一個問題將先針對雙樣本單尾檢定統計量，利用模擬的方法比較各檢定統計量的顯著水準(level)與檢定力(power)表現，以便提供使用者選擇檢定方法之參考。因為 Pepe 和 Fleming (1989)所提之加權 Kaplan-Meier 檢定為單尾之雙樣本檢定，針對對立假說為

$H_2 : (S_1(t) \neq S_2(t) \text{ for some } t)$ 之檢定，若存活函數相交時，直接使用 WKM 統計量的平方或絕對值並不恰當，因為會有前後期存活函數差異抵消的情形發生。因此，利用 Renyi-type 的方法修正加權 Kaplan-Meier 檢定，建構 Renyi-type 加權 Kaplan-Meier 檢定。此一方法的建立，可以提供更一般化的雙樣本檢定方法。

近年來，隨機化控制臨床試驗(randomized controlled clinical trial)在醫學研究上對於評價新藥或新治療法的效益已成為一種重要的方法。基於倫理或成本上的考量，週期性地檢驗資料，已成為提早發現醫藥或療法效益的依據。在試驗進行時，若無充分令人信服的證據支持新藥或新療法，則隨機化控制臨床試驗就繼續到預定的時間或預定的樣本數達到為止。若有足夠的證據累積支持或反對新藥或新療法，則可提前停止試驗。因此，群序列檢定(group sequential test)的建立是必要的。若觀測資料為存活時間時，目前已有許多文獻探討 WLR 和 WKM 的群序列檢定，例如：Slud 和 Wei (1982)，Tsiatis (1982)，Slud (1984)，Lin 和 Wei (1989)和 Li (1999)等。然醫學研究中，收集的資料並非僅限存活時間，常有與病人生理狀態和病情相關的背景資料（稱為共變數(covariate)），這些共變數經常提供解釋存活現象的重要訊息。針對調整共變數之群序列檢定已有 Tsiatis et al.(1985)，Lin (1992)和 Gu 和 Ying (1995)等文獻探討。其中，Tsiatis et al.(1985) 為避免共變數效應之混同(confounding)，進一步在 Cox(1972)模式之下，建立調整共變數效應之群序列 LR 檢定，但這些方法僅適用於比例危險函數(proportional hazards function)模式。Shen 和 Fleming (1997)曾針對雙樣本存活資料檢定問題中，當收集之資料含有共變數時，在無比例危險函數模式假說下，建立一調整共變數之 WKM 檢定，稱之為加權平均存活(weighted mean survival，記為 WMS)檢定統計量。為能更廣泛分析各種形態的危險函數模式的群序列資料，本文參考 Li (1999)和 Shen 和 Fleming (1997)之研究，建構一個調整共變數效應的 WKM 群序列檢定。

在劑量反應研究中，若反應值隨著劑量漸增而遞增，但是受制於在高劑量下可能產生之反效果，之後反應值隨劑量漸增而遞減，此時，劑量反應呈先增後減之傘形(umbrella)型態。當資料完整收集時，已有諸多文獻對此種型態之劑量反應提出無母數檢定方法，稱之為傘檢定 (Jonckheere (1954)，Terpstra (1952)，Mack 和 Wolfe (1981)，Simpson 和 Margolin (1986)，Chen 和 Wolfe (1990, 1991)) Liu 等人(1993)針對多組右方

設限存活資料的順序對立假說 $H_{1O} : (S_1(t) \leq \dots \leq S_k(t) \text{ for some } t \text{ with at least one strict inequality})$, 根據雙樣本加權對數秩統計量 , 推廣 Jonckheere (1954)–Terpstra (1952) 檢定方法。Chen 和 Wolfe (2000) 則針對傘形對立假說 $H_{1U} : (S_1(t) \leq \dots \leq S_p(t) \geq \dots \geq S_k(t) \text{ for some } t \text{ with at least one strict inequality})$, 其中 p 為此一傘形之傘頂 (peak of the umbrella) , 進一步根據加權對數秩統計量推廣傘頂已知的 Mack-Wolfe (1981) 檢定和傘頂未知的 Chen-Wolfe (1990) 檢定。然傘形對立假說中檢定的是存活函數之型態 , 而非危險函數之型態。所以 , 本文中最後根據雙樣本 WKM 統計量推廣傘頂已知的 Mack-Wolfe (1981) 檢定和傘頂未知的 Chen-Wolfe (1990) 檢定 , 建立存活函數型態之傘檢定。

1.2 加權 Kaplan-Meier 統計量 :

Pepe 和 Fleming (1989) 基於雙樣本加權對數秩檢定對於存活函數差異的不靈敏缺失 , 為真正反應兩組存活函數之差異 , 根據兩組樣本的 Kaplan-Meier 存活函數估計量的差 , 提出加權 Kaplan-Meier 統計量。

假設存活時間 $T_{i1}, \dots, T_{in_i} \sim iid S_i(t)$, 對應之設限時間 $C_{i1}, \dots, C_{in_i} \sim iid C_i(t)$, $i=1,2$, 其中 $S_i(t)$ 和 $C_i(t)$ 分別為 T_{ij} 和 C_{ij} 之存活函數。又假設 T_{ij} 's 與 C_{ij} 's 為獨立之隨機變數。此時 , 觀察之雙樣本存活資料為 $X_{ij} = \min\{T_{ij}, C_{ij}\}$ 和設限指標 $d_{ij} = I\{T_{ij} < C_{ij}\}$, $j=1, \dots, n_i$, $i=1,2$ 。針對

$$H_0 : (S_1(t) = S_2(t) \text{ for all } t)$$

相對於

$$H_1 : (S_1(t) \leq S_2(t) \text{ for all } t \text{ and } S_1(t) < S_2(t) \text{ for some } t)$$

的雙樣本單尾檢定問題 , Pepe 和 Fleming (1989) 建議採用下述統計量建立檢定方法 :

$$WKM_{12} = \sqrt{\frac{n_1 n_2}{n}} \int_0^{T_c} \hat{w}_{12}(t) \{ \hat{S}_2(t) - \hat{S}_1(t) \} dt , \quad (1.1)$$

其中

$$T_c = \sup\{t : \min\{ \hat{S}_1(t), \hat{S}_2(t), \hat{C}_1(t), \hat{C}_2(t) \} > 0\} , \quad (1.2)$$

$\hat{S}_i(\cdot)$ 為 $S_i(\cdot)$ 的 Kaplan-Meier 估計量 , $i=1,2$,

$\hat{C}_i(\cdot)$ 為 $C_i(\cdot)$ 的 Kaplan-Meier 估計量 , $i=1,2$,

$$\hat{w}_{12}(t) = \frac{\hat{C}_1(t-)\hat{C}_2(t-)}{\hat{p}_1\hat{C}_1(t-)+\hat{p}_2\hat{C}_2(t-)} , \hat{p}_i = \frac{n_i}{n} , i=1,2。$$

此時加權函數 \hat{w}_{12} 的作用是穩定 WKM_{12} 之變異，以便求得 WKM_{12} 之極限分布。當在完整觀測資料下， WKM_{12} 統計量可寫為 $c(\bar{X}_1 - \bar{X}_2)$ ，其中 \bar{X}_i 為第 i 組樣本平均數， $i=1,2$ ，所以雙樣本 WKM 檢定可視為雙樣本 t 檢定的推廣。Pepe 和 Fleming (1989) 亦驗證 WKM_{12} 統計量在 H_0 之下的極限分布為常態分布，並且求得 $\text{Var}(WKM_{12})$ 之一致估計量 $\hat{\sigma}^2$ 後，建議若

$$W\hat{K}M_{12}^* = WKM_{12}/\hat{\sigma} \geq z(\mathbf{a}) , \quad (1.3)$$

則棄卻 H_0 ，結語 H_1 成立，其中 $z(\mathbf{a})$ 為 $N(0,1)$ 之上 100 百分位。

Pepe 和 Fleming (1991) 在不同形態之危險函數差異下，推導出 WKM 統計量的效益(efficacy)，同時計算 WKM 統計量相對於對數秩(logrank，記為 LR) 統計量之漸進相對效率(asymptotic relative efficiency)，記為 $\mathbf{ARE}(WKM, LR)$ 。其研究結果指出，在比例危險模式之下，雙樣本 WKM 檢定與 LR 檢定具有相等之效率，但對早期即呈現差異或交叉之危險函數而言， WKM 檢定明顯優於 LR 檢定。

由於 Pepe 和 Fleming (1989, 1991) 僅研究相對於未加權的 LR 檢定與 WKM 檢定的檢定力表現。然而在 WLR 檢定族中包含針對危險函數差異主要發生在早期或晚期的檢定，例如：針對危險函數差異發生在早期時之 Peto-Prentice- Wilcoxon 檢定(Peto 和 Peto, 1972; Prentice, 1978，記為 PPW)，和針對危險函數差異發生在晚期時之 LRL 檢定。本文採用 Fleming 和 Harrington (1981) 所建議之權數形式 $\{S(t)\}^r \{1-S(t)\}^g$ ，其中 $r \geq 0, g \geq 0$ 。因此考慮 WKM 檢定相對於加權對數秩檢定統計量： LR ($r=g=0$) 統計量， PPW ($r=1, g=0$) 統計量和針對危險函數差異發生在晚期之 LRL ($r=0, g=1$) 統計量的檢定力表現。圖 1-1 為模擬研究中各種不同型態差異之危險函數，表 1-1 和表 1-2 則列出模擬 5000 次在名目水準(nominal level) $\mathbf{a} = 0.05$ ，樣本數相同(顯著水準估計時使用之樣本數分別為 10, 20 和 30，檢定力估計使用之樣本數則分別為 20 和 30)，且設限分布 $U(0, R)$ (討論 $R=2$ 或 3) 相同時， WKM, LR, PPW 和 LRL 之顯著水準及檢定力估計。其中顯著水準模擬之標準差為 0.003 ($\sqrt{(0.95)(0.05)/5000}$)，檢定力模擬之標準差為 0.007 ($\sqrt{(0.5)(0.5)/5000}$)。

由表 1-1 得知除 $n = 10$ 之外，各檢定之顯著水準均在誤差範圍內。表 1-2 顯示在比例危險函數之下， WKM 檢定之檢定力與 LR 檢定之檢定力相當，在危險函數差異發生在早期時， WKM 檢定之檢定力與 PPW 檢定之檢定力相當，但在危險函數差異發生在晚期時， WKM 檢定之檢定力比 LRL 及 LR 檢定之檢定力差。這是因為 WKM 統計量的權數是隨存活時間遞增而遞減，所以差異發生在晚期時， WKM 統計量的權數少因此降低了檢定力。若危險函數交叉，則 WKM 檢定之檢定力比研究中的 WLR 檢定之檢定力高。

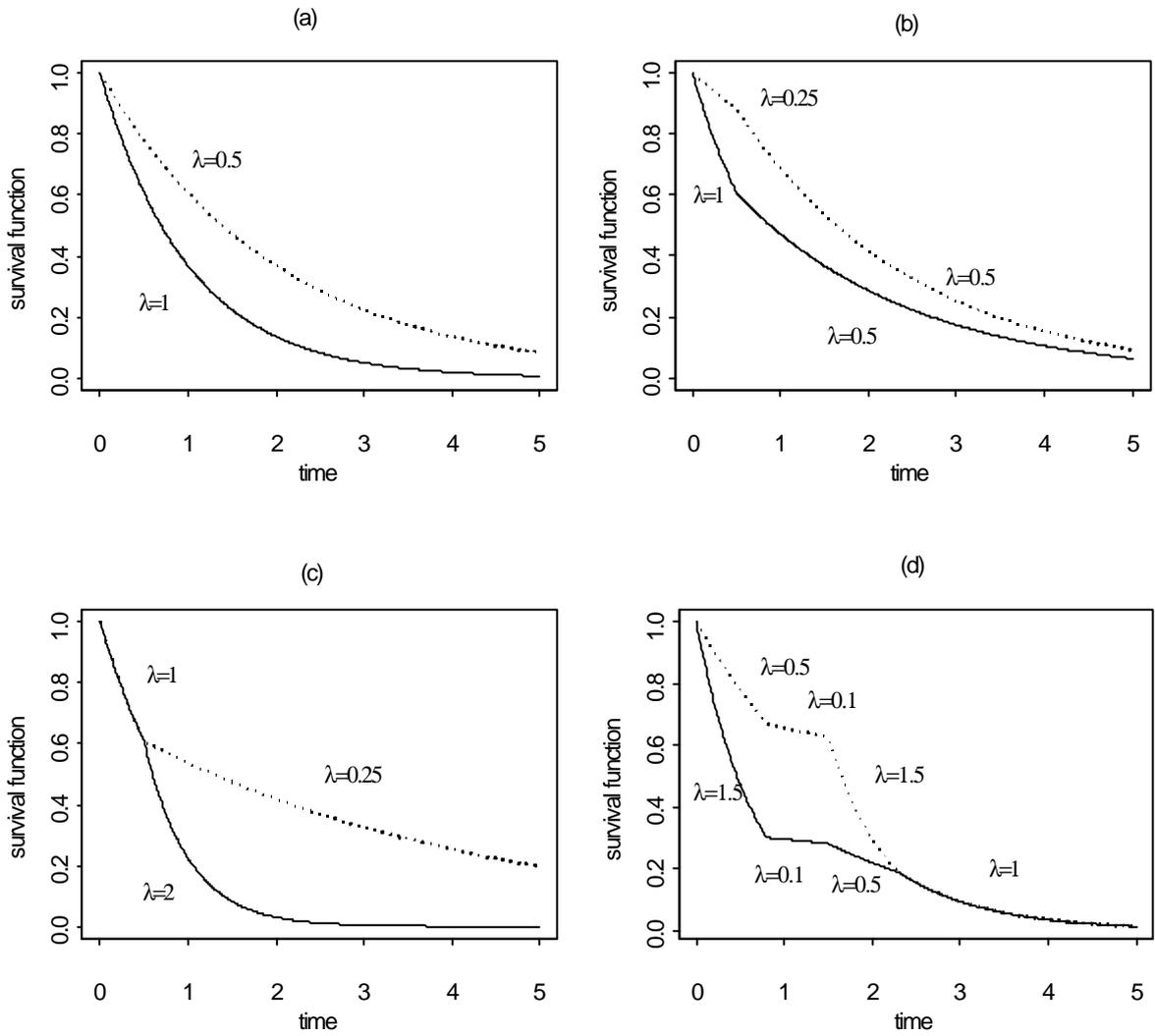
表 1-1 名目水準 $\alpha = 0.05$ ，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之水準估計

R	n	WKM	LR	PPW	LRL
(a) Standard exponential distribution					
2	10	0.057	0.050	0.049	0.059
	20	0.053	0.051	0.046	0.057
	30	0.051	0.051	0.051	0.054
3	10	0.053	0.052	0.046	0.069
	20	0.052	0.051	0.050	0.054
	30	0.050	0.049	0.048	0.051
(b) Piecewise exponential distribution					
$I(t) = I_{[t \leq 0.5]} + 0.5I_{[t > 0.5]}$					
2	10	0.058	0.052	0.047	0.063
	20	0.054	0.051	0.051	0.053
	30	0.052	0.050	0.051	0.053
3	10	0.058	0.053	0.052	0.059
	20	0.050	0.050	0.052	0.055
	30	0.052	0.049	0.053	0.054

表 1-2 名目水準 $\alpha = 0.05$, 設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之檢定力估計

參數		R	n	WKM	LR	PPW	LRL
I_1	I_2	(a) Exponential distribution					
1	0.5	2	20	0.421	0.425	0.402	0.358
			30	0.558	0.567	0.541	0.463
		3	20	0.490	0.489	0.453	0.421
			30	0.633	0.634	0.590	0.537
(b) Early hazard difference ($I(t) = I_i I_{[t \leq 0.5]} + 0.5 I_{[t > 0.5]}$)							
1	0.25	2	20	0.508	0.421	0.497	0.190
			30	0.657	0.563	0.650	0.228
		3	20	0.458	0.398	0.493	0.161
			30	0.599	0.517	0.638	0.180
(c) Late hazard difference ($I(t) = I_{[t \leq 0.5]} + I_i I_{[t > 0.5]}$)							
2	0.25	2	20	0.343	0.408	0.218	0.682
			30	0.432	0.529	0.279	0.835
		3	20	0.552	0.553	0.293	0.845
			30	0.722	0.733	0.405	0.953
(d) Crossing hazard difference ($I(t) = I_{i1} I_{[t \leq 0.8]} + 0.1 I_{[0.8 < t \leq 1.5]} + I_{i2} I_{[1.5 < t \leq 2.3]} + I_{[t > 2.3]}$)							
1.5	0.5	2	20	0.696	0.578	0.670	0.481
0.5	1.5		30	0.847	0.804	0.824	0.581
		3	20	0.650	0.536	0.647	0.261
			30	0.796	0.658	0.789	0.276

圖 1-1 不同存活函數形態之雙樣本單尾對立假說



1.3 本文結構

本文第二、三章討論雙樣本的問題。第二章根據 Renyi-type 加權對數秩檢定統計量的方法，建立 Renyi-type 加權 Kaplan-Meier 檢定統計量，並利用模擬比較 Renyi-type 加權 Kaplan-Meier 檢定統計量與 Renyi-type 加權對數秩檢定統計量。本文第三章推導建立調整共變數之加權 Kaplan-Meier 群序列檢定，並且利用實例說明。

本文第四章討論多樣本問題。推廣雙樣本之加權 Kaplan-Meier 統計量檢定傘形對立假說。首先，討論傘頂已知和未知之加權 Kaplan-Meier 傘檢定，除推導出傘頂已知傘檢定之效率(efficacy)，且利用模擬比較加權 Kaplan-Meier 傘檢定與加權對數秩傘檢定之優劣。最後利用實例說明上述傘檢定之應用。本文最後一章總結，且陳述未來研究的方向。

第二章 Renyi-type 加權 Kaplan-Meier 檢定

針對檢定虛無假說

$$H_0 : (S_1(t) = S_2(t) \text{ for all } t)$$

相對於雙尾對立假說

$$H_2 : (S_1(t) \neq S_2(t) \text{ for some } t) \quad (2.1)$$

的雙樣本問題，Gill (1980)根據雙樣本加權對數秩統計量建立 Renyi-type 加權對數秩統計量，用以檢定兩危險函數不同之雙尾檢定。針對存活函數相交時，為檢定存活函數的差異，本章則根據雙樣本加權 Kaplan-Meier 統計量建立 Renyi-type 雙樣本存活函數差異之雙尾檢定。

2.1 Renyi-type 加權對數秩統計量

Gill (1980)針對危險函數交叉時將加權對數秩(WLR)檢定修正，提出 Renyi-type WLR 檢定，其檢定統計量為

$$Q = \sup\{|WLR(t)|, t \leq t\} / s(t), \quad (2.2)$$

其中 $t = \min\{t_1, t_2\}$ ， t_i 為各組最後的觀測值。

Gill 亦證明 Q 統計量之極限分布為 $\sup\{|W(x)|, 0 \leq x \leq 1\}$ ，其中 W 為標準布朗運動過程(standard Brownian motion process)。又因

$$P(\sup_{0 \leq t \leq 1} |W(t)| > y) = 1 - \frac{4}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \exp\{-\pi^2(2k+1)^2 / 8y^2\}, \quad (2.3)$$

且 Schumacher (1984)提供 $H(y) = P(\sup_{0 \leq t \leq 1} |W(t)| \leq y)$ 值之表，因此，利用此表便可找出檢定之適當之臨界值。

2.2 Renyi-type 加權 Kaplan-Meier 檢定

針對對立假說為(2.1)式之雙樣本檢定問題，若直接使用 $(\hat{WKM})^2$ 建立雙尾檢定並不合適，因為其中存活函數估計之差異會相互抵消。因此，本節參考 Renyi-type WLR 檢

定(Gill, 1980)的方法，建立 Renyi-type WKM 統計量。

假設合併兩樣本之觀測時間點為 $0 < t_1 < t_2 < \dots < t_D = T_c$ ，其中

$$T_c = \sup\{t : \min\{\hat{S}_1(t), \hat{S}_2(t), \hat{G}_1(t), \hat{G}_2(t)\} > 0\}。$$

首先，計算在每一觀測時間點上之 WKM 統計量值，得到 $WKM(t)$ 序列，其中 $0 < t \leq T_c$ 。然後，參考 Renyi-type WLR 檢定的方法，建立 Renyi-type WKM 檢定方法。定義檢定統計量

$$RW = \sup\{|WKM(t)|, t \leq \mathbf{t}\} / \mathbf{s}(\mathbf{t}) , \quad (2.4)$$

其中

$$\mathbf{t} = \sup\{t : \min(S_1(t), S_2(t), G_1(t), G_2(t)) > 0\} ,$$

$$\mathbf{s}^2(\mathbf{t}) = -\int_0^{\mathbf{t}} \frac{(\int_x^{\mathbf{t}} w(u)S(u)du)^2}{S^2(x)} \frac{p_1 C_1(x-) + p_2 C_2(x-)}{C_1(x-)C_2(x-)} dS(x)。$$

定理 2.1

在虛無假說之下， $RW \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|$ ，其中 W 為標準布朗運動過程， \xrightarrow{d} 代表分布收斂(converge in distribution)。

證明：

因為在虛無假設 H_0 之下

$$WKM \xrightarrow{d} Z_0^\infty , \text{ 在 } D[0, \mathbf{t}] ,$$

其中， WKM 定義同(1.1)式，且 Z_0^∞ 是均數為 0 之高斯過程(Gaussian Process)。又對任意 $t \in [0, \mathbf{t}]$

$$\text{Var}(Z_0^\infty(t)) = \mathbf{s}^2(t) = -\int_0^t \frac{(\int_x^t w(u)S(u)du)^2}{S^2(x)} \frac{p_1 C_1(x-) + p_2 C_2(x-)}{C_1(x-)C_2(x-)} dS(x) ,$$

其一致估計式為

$$\hat{\mathbf{s}}^2(t) = -\int_0^t \frac{(\int_x^{T_c} \hat{w}(u) \hat{S}(u) du)^2}{\hat{S}^2(x)} \frac{\hat{p}_1 \hat{C}_1(x-) + \hat{p}_2 \hat{C}_2(x-)}{\hat{C}_1(x-) \hat{C}_2(x-)} d\hat{S}(x). \quad (2.5)$$

且當 $n \rightarrow \infty$ 時，對所有 $t \in [0, t]$ ， $\hat{\mathbf{s}}^2(t) \xrightarrow{p} \mathbf{s}^2(t)$ (Pepe 和 Fleming, 1989, 1991)，其中 \xrightarrow{p} 代表機率收斂 (converge in probability)。又因

$$P\left(\sup_{s \in [0, t]} \frac{|Z_0^\infty(s)|}{\sqrt{\text{Var}(Z_0^\infty(t))}} \leq x\right) = P(\sup_{0 \leq s \leq 1} |W(s)| \leq x)。$$

所以得証

$$RW = \sup\{|WKM(t)|, t \leq t\} / \mathbf{s}(t) \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|。$$

因為 $\mathbf{s}^2(t)$ 是未知的，又 $\mathbf{s}^2(t)$ 之一致估計式為 $\hat{\mathbf{s}}^2(T_c)$ ，定義統計量為如下之 \widehat{RW} ，其極限分布與 RW 相同，亦即

$$\widehat{RW} = \sup\{|WKM(t)|, t \leq T_c\} / \hat{\mathbf{s}}(T_c) \xrightarrow{d} \sup_{0 \leq t \leq 1} |W(t)|，$$

其中 $\hat{\mathbf{s}}^2(T_c)$ 如(2.5)式所列。因此亦可利用(2.3)式求得 Renyi-type WKM 檢定之臨界值。

2.2 模擬研究

利用蒙地卡羅法模擬在名目水準 $\alpha = 0.05$ ，樣本數相同及相同設限分布下，得知 Renyi-type WKM (記為, $RWKM$) 檢定之顯著水準太過保守，僅維持在 0.03 左右。圖 2-1 為在虛無假說下，記錄模擬 5000 次樣本數均為 30 之 Renyi-type WKM 統計量及 Renyi-type WLR (記為, RLR) 統計量之值並描繪其分布所得。由圖 2-1 可知 Renyi-type WKM 檢定統計量之收斂速度較慢，所以使用(2.3)式之臨界值在模擬的表現並不佳。因此建議利用靴環法(bootstrap)求其臨界值。針對存活函數分布為指數分布或成段指數分布及設限分布為均勻分布 $U(0, R)$ ，建構出存活函數交叉在不同時段之對立假說，分別探討在小樣本 ($n_1 = n_2 = n = 10, 20$ 或 30) Renyi-type WKM 檢定及 Renyi-type LR 檢定之表現。在此一模擬研究中，我們首先從存活函數分布及設限分布中，抽取一組資料為

樣本，再從所得樣本中重覆抽取 200 次，建立檢定統計量之靴環分布，並找出其相對應之臨界值，然後與原始抽出之資料之檢定統計量比較，最後重覆此程序 5000 次，並記錄其結果。圖 2-2(a)(b)(c)(d) 為不同形態存活函數相交之對立假說。表 2-1 為在名目水準 $\alpha = 0.05$ ，樣本數均為 10, 20 和 30 下，且相同設限分布為 $U(0,R)$ ($R=3$ 或 4)，利用靴環法模擬顯著水準之結果。表 2-2 為根據圖 2-2(a)(b)(c)(d) 之對立假說，使用名目水準 $\alpha = 0.05$ ，樣本數均為 20 和 30 下，且相同設限分布為 $U(0,R)$ ($R=3$ 或 4)，利用靴環法模擬檢定力之結果。

由表 2-1 得知兩檢定方法在虛無假說下之顯著水準維持在誤差範圍內。由表 2-2 得知在危險函數交叉時，唯交叉的時間點唯 2.5 時 Renyi-type *WKM* 檢定比 Renyi-type *LR* 檢定之檢定力佳，其餘皆是 Renyi-type *LR* 檢定之檢定力較佳。這是因為存活函數之相交時間點較危險函數之相交時間點晚。所以，以存活函數為基礎之 Renyi-type *WKM* 檢定便會比以危險函數為基礎之 Renyi-type *LR* 檢定表現為差。

表 2-1 名目水準 $\alpha = 0.05$ ，存活分布為標準指數分布設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之靴環法之水準估計。

n	R	RWKM	RLR
10	3	0.052	0.051
20		0.050	0.047
30		0.053	0.050
10	4	0.053	0.051
20		0.051	0.049
30		0.050	0.050

表 2-2 名目水準 $\alpha = 0.05$ ，設限分布為 $U(0,R)$ 且 $n_1 = n_2 = n$ 之靴環法之檢定力估計。

參數		R	n	t0	RWKM	RLR			
0.5	1	4	20	1	0.167	0.202			
1	0.5			1.5	0.279	0.312			
				2	0.385	0.352			
				2.5	0.418	0.409			
			30	1	0.216	0.320			
					1.5	0.399	0.442		
					2	0.541	0.541		
					2.5	0.562	0.554		
			3	20	1	0.209	0.253		
						1.5	0.276	0.293	
						2	0.361	0.357	
						2.5	0.383	0.381	
					30	1	0.270	0.330	
						1.5	0.416	0.430	
2	0.491					0.495			
2.5	0.518		0.512						
0.5	1.5	4	20	1	0.494	0.578			
1.5	0.5			1.5	0.683	0.685			
				2	0.768	0.755			
				2.5	0.729	0.724			
			30	1	0.676	0.756			
					1.5	0.871	0.880		
					2	0.904	0.903		
					2.5	0.886	0.882		
			3	20	1	0.428	0.548		
						1.5	0.701	0.735	
						2	0.743	0.742	
						2.5	0.764	0.756	
					30	1	0.592	0.755	
							1.5	0.842	0.882
							2	0.905	0.914
			2.5	0.932			0.927		

續表 2-2

參數		R	n	t0	RWKM	RLR					
1	1	4	20	1	0.161	0.209					
0.5	2			1.5	0.296	0.307					
				2	0.340	0.357					
				2.5	0.397	0.384					
			30	1	0.211	0.271					
					1.5	0.401	0.430				
					2	0.497	0.509				
					2.5	0.570	0.565				
					3	20	1	0.181	0.194		
							1.5	0.322	0.333		
							2	0.338	0.342		
							2.5	0.366	0.366		
							30	1	0.256	0.293	
								1.5	0.417	0.428	
2	0.492							0.503			
2.5	0.490							0.506			
1	1	4	20	1	0.371	0.418					
2	0.5			1.5	0.455	0.438					
				2	0.477	0.458					
				2.5	0.450	0.428					
			30	1	0.435	0.515					
							1.5	0.613	0.617		
							2	0.627	0.610		
							2.5	0.669	0.648		
					3	20	1	0.385	0.400		
							1.5	0.432	0.407		
							2	0.458	0.419		
							2.5	0.434	0.422		
							30	1	0.472	0.530	
									1.5	0.607	0.605
									2	0.640	0.635
			2.5	0.610					0.586		

圖 2-1 $n=30$, RWKM 和 RLR 之模擬抽樣分布。

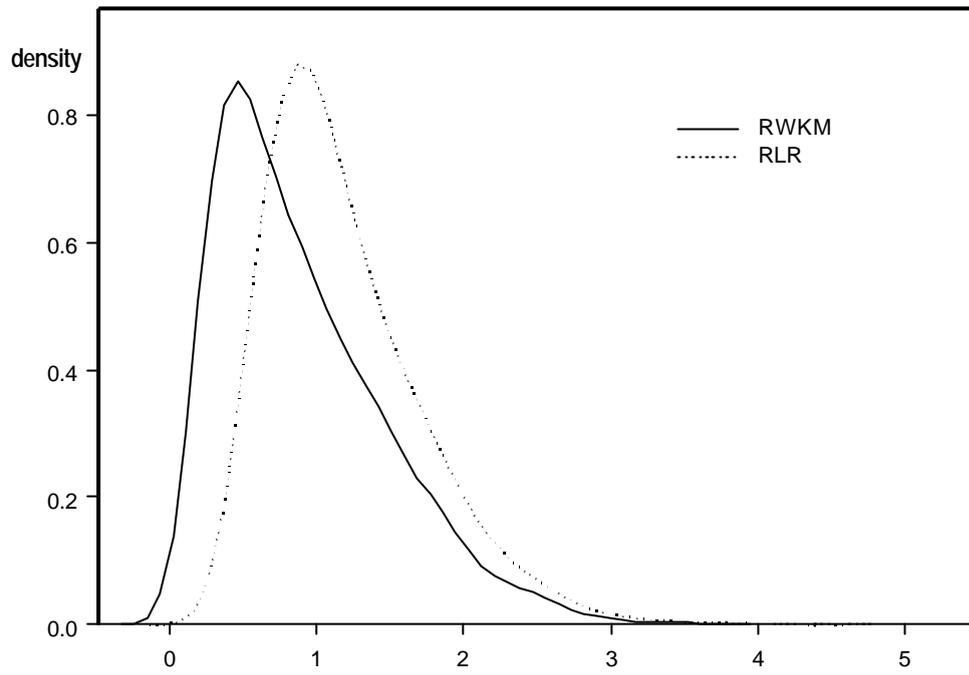
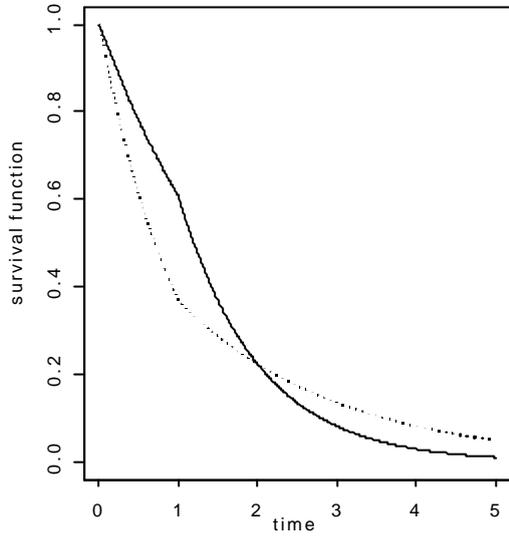


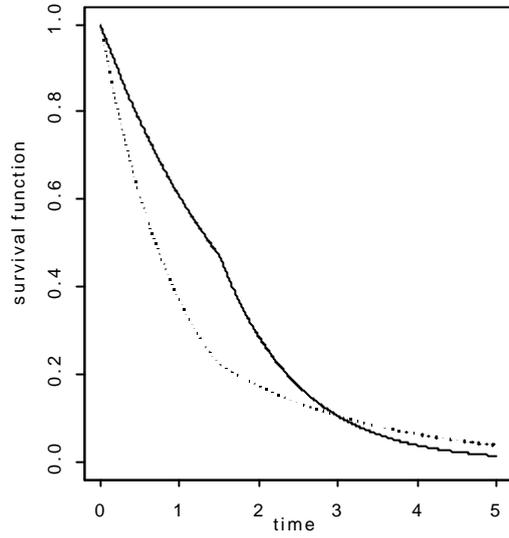
圖 2-2 存活函數相交之對立假說。

(a)



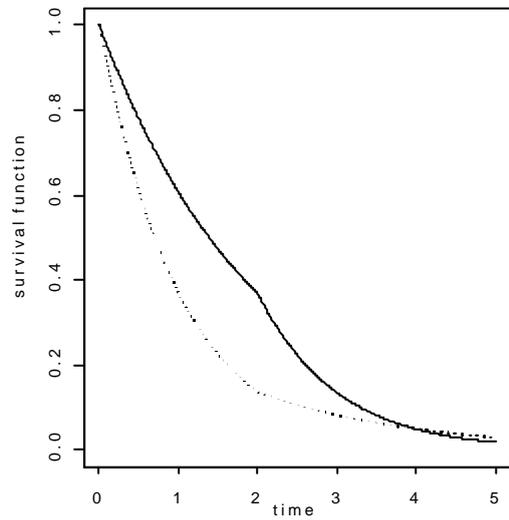
— $I_1(t) = 0.5I[t < 1] + I[t \geq 1]$

..... $I_2(t) = I[t < 1] + 0.5I[t \geq 1]$



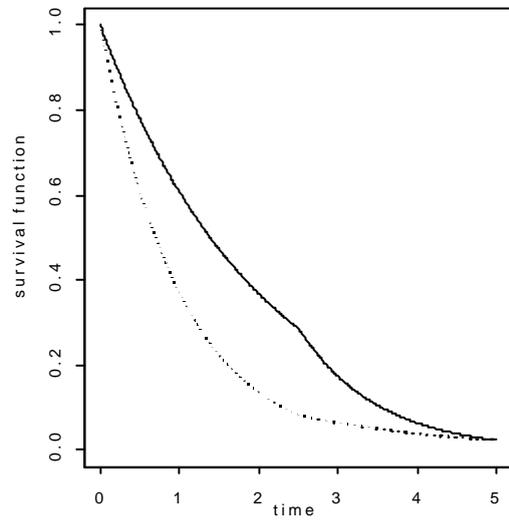
— $I_1(t) = 0.5I[t < 1.5] + I[t \geq 1.5]$

..... $I_2(t) = I[t < 1.5] + 0.5I[t \geq 1.5]$



— $I_1(t) = 0.5I[t < 2] + I[t \geq 2]$

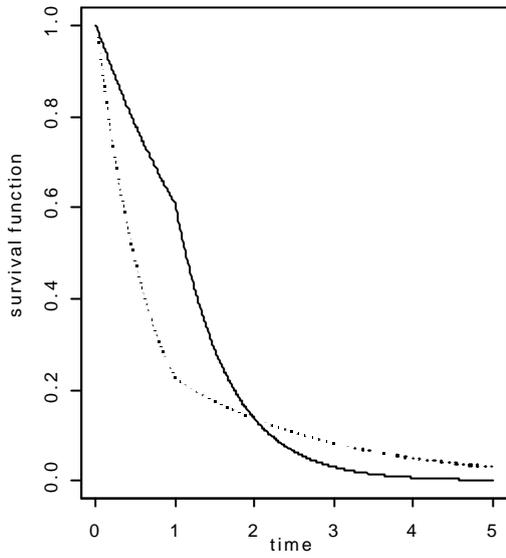
..... $I_2(t) = I[t < 2] + 0.5I[t \geq 2]$



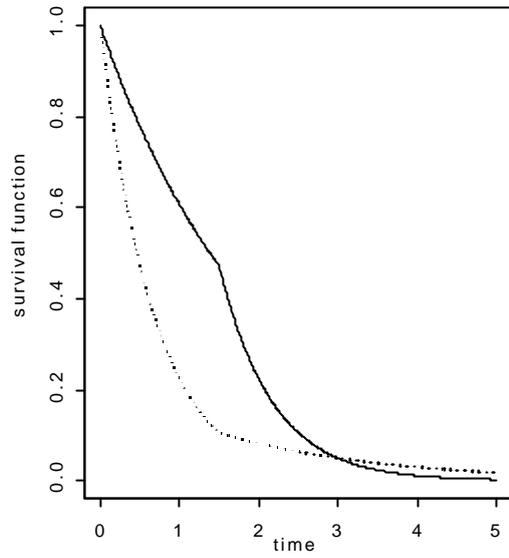
— $I_1(t) = 0.5I[t < 1.5] + I[t \geq 1.5]$

..... $I_2(t) = I[t < 2.5] + 0.5I[t \geq 2.5]$

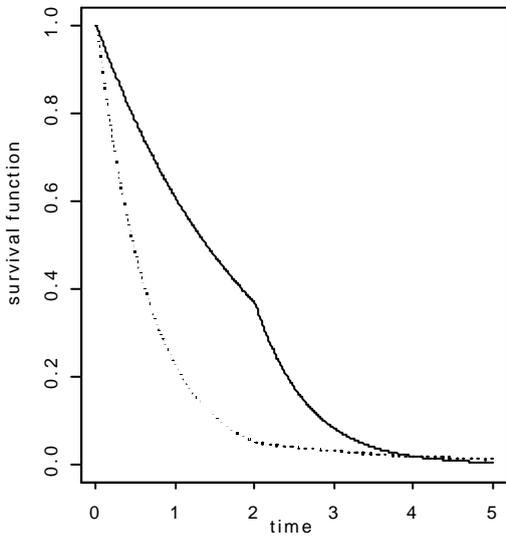
(b)



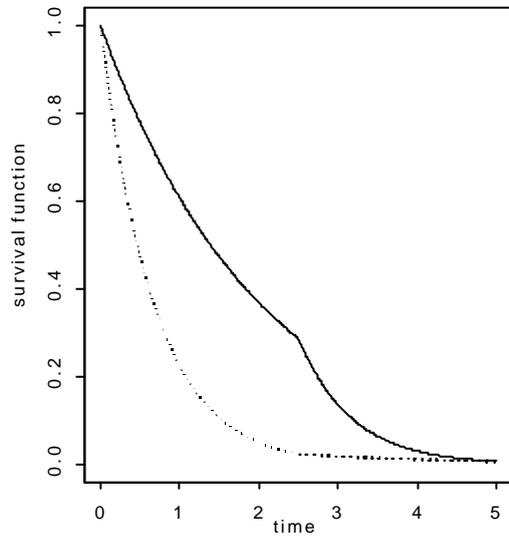
— $I_1(t) = 0.5I[t < 1] + 1.5I[t \geq 1]$
- - - $I_2(t) = 1.5I[t < 1] + 0.5I[t \geq 1]$



— $I_1(t) = 0.5I[t < 1.5] + 1.5I[t \geq 1.5]$
- - - $I_2(t) = 1.5I[t < 1.5] + 0.5I[t \geq 1.5]$

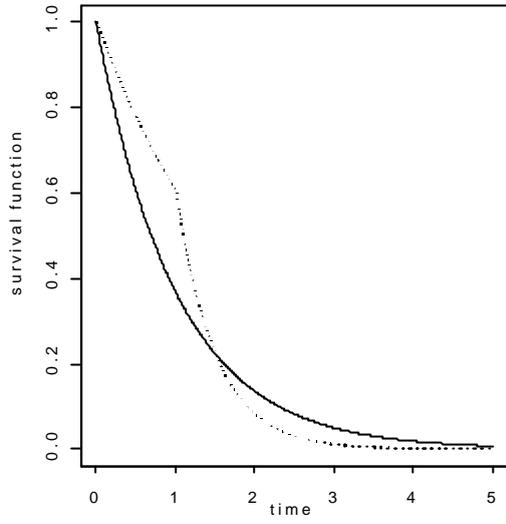


— $I_1(t) = 0.5I[t < 2] + 1.5I[t \geq 2]$
- - - $I_2(t) = 1.5I[t < 2] + 0.5I[t \geq 2]$

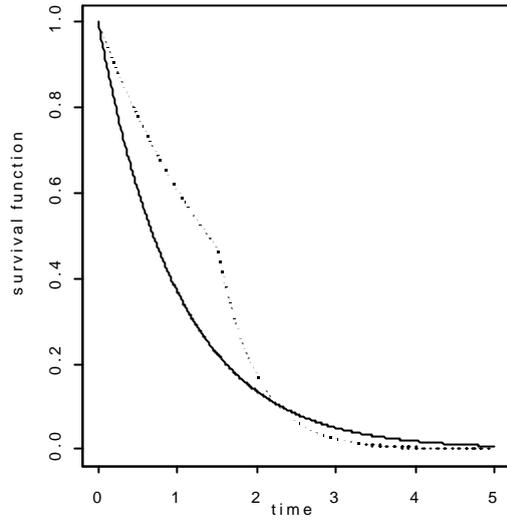


— $I_1(t) = 0.5I[t < 2.5] + 1.5I[t \geq 2.5]$
- - - $I_2(t) = 1.5I[t < 2.5] + 0.5I[t \geq 2.5]$

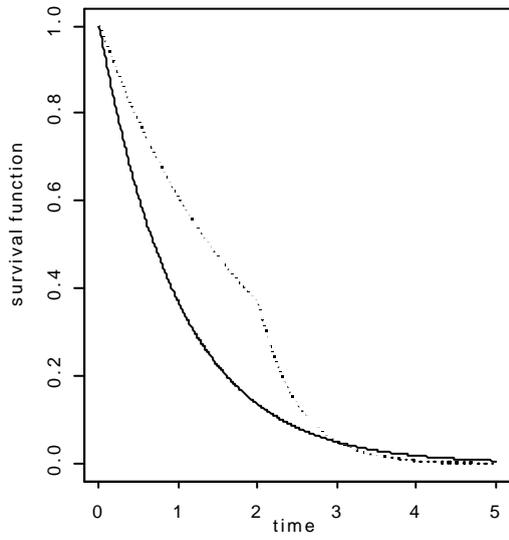
(c)



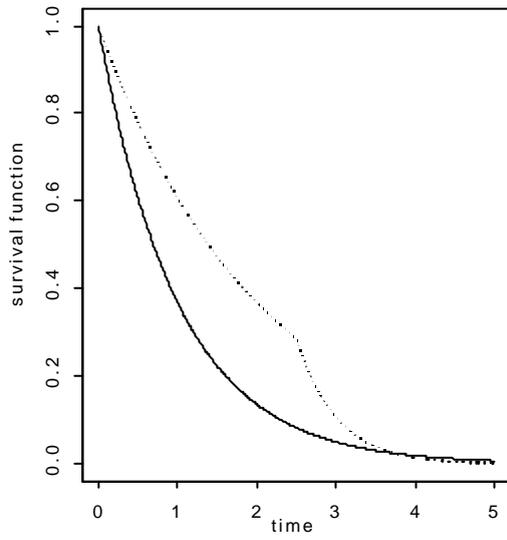
— $I_1(t) = I[t < 1] + I[t \geq 1]$
- - - $I_2(t) = 0.5I[t < 1] + 2I[t \geq 1]$



— $I_1(t) = I[t < 1.5] + I[t \geq 1.5]$
- - - $I_2(t) = 0.5I[t < 1.5] + 2I[t \geq 1.5]$

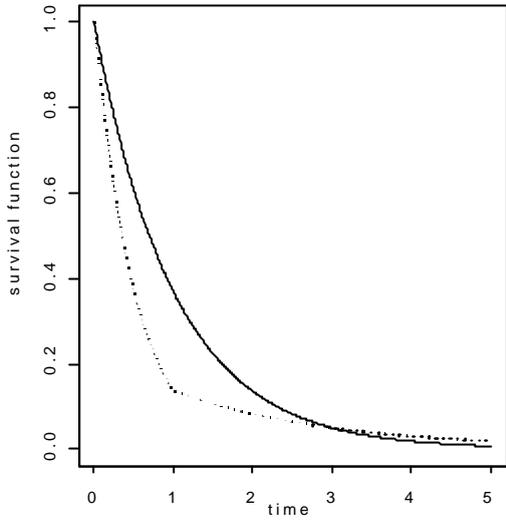


— $I_1(t) = I[t < 2] + I[t \geq 2]$
- - - $I_2(t) = 0.5I[t < 2] + 2I[t \geq 2]$

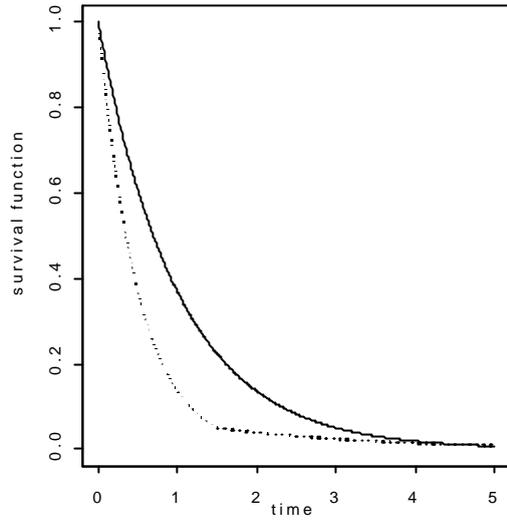


— $I_1(t) = I[t < 1.5] + I[t \geq 1.5]$
- - - $I_2(t) = 0.5I[t < 2.5] + 2I[t \geq 2.5]$

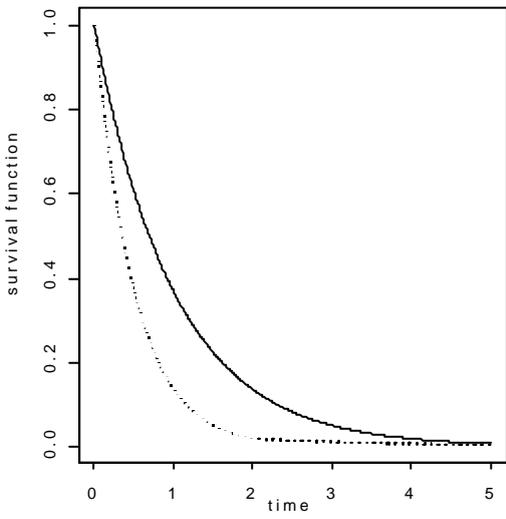
(d)



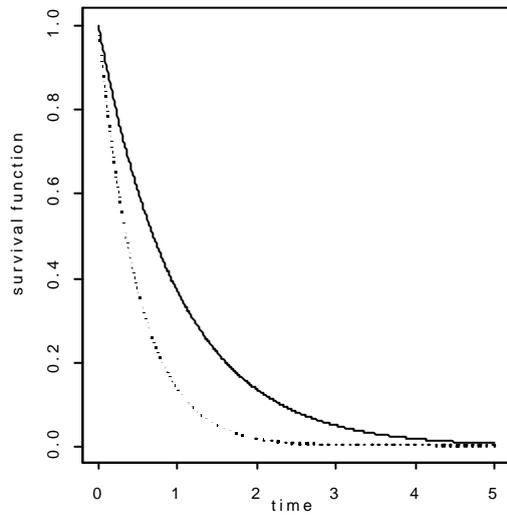
— $I_1(t) = I[t < 1] + I[t \geq 1]$
- - - $I_2(t) = 2I[t < 1] + 0.5I[t \geq 1]$



— $I_1(t) = I[t < 1.5] + I[t \geq 1.5]$
- - - $I_2(t) = 2I[t < 1.5] + 0.5I[t \geq 1.5]$



— $I_1(t) = I[t < 2] + I[t \geq 2]$
- - - $I_2(t) = 2I[t < 2] + 0.5I[t \geq 2]$



— $I_1(t) = I[t < 1.5] + I[t \geq 1.5]$
- - - $I_2(t) = I[t < 2.5] + 0.5I[t \geq 2.5]$

第三章 調整共變數之加權平均存活群序列檢定

3.1 簡介及文獻回顧

隨機化控制臨床試驗在醫學研究上對於評價新藥或新治療法的效益已成為一種重要的方法。基於倫理或成本上的考量，週期性地檢驗資料，已成為提早發現醫藥或療法效益的依據。在試驗進行時，若無充分令人信服的證據支持新藥或新療法，則隨機化控制臨床試驗就繼續到預定的時間或預定的樣本數達到為止。若有足夠的證據累積支持或反對新藥或新療法，則可提前停止試驗。因此，群序列檢定的建立是必要的。然醫學研究中，收集的資料並非僅限存活時間，常有與病人生理狀態和病情相關的背景資料（稱為共變數(covariate)），這些共變數經常提供解釋存活現象的重要訊息。因此建立調整共變數之群序列檢定有其必要。在此先回顧一些重要的文獻。

3.1.1 加權平均存活統計量：

在雙樣本存活資料檢定問題中，當收集的資料中含有共變數時，為避免共變數效應之混同，一般是採用 Cox 迴歸模式(Cox, 1972)，但 Cox 迴歸模式乃植基於比例危險模式之假設。若在非比例危險模式下，Cox 迴歸模式檢定兩樣本存活函數的差異，便不夠靈敏。所以，Shen 和 Fleming (1997)建議在雙樣本加權 Kaplan-Meier 檢定統計量(Pepe 和 Fleming, 1989)中，於各組存活函數引入 Cox 迴歸模式，提出雙樣本加權平均存活檢定統計量。假設在存活時間和設限時間之外，尚測得共變數 Z ，則可配適 Cox 迴歸模式：

$$S_i(t|Z) = \exp\{-\Lambda_{0i}(t)\exp(\mathbf{b}_i'Z)\}, i = 1, 2,$$

其中 $\Lambda_{0i}(t)$ 為累積基準危險函數(cumulative base-line hazard function)， Z 為 $p \times 1$ 的向量， \mathbf{b}_i 為第 i 組 Z 的係數向量， $i = 1, 2$ 。因為 $S_i(t) = E[S_i(t|Z)]$ ，若 Z 之經驗分布為 F_n ，則

$$\hat{S}_i(t) = \int \hat{S}_i(t|z) dF_n(z) ,$$

其中 $\hat{S}_i(t|Z) = \exp\{-\hat{\Lambda}_{0i}(t) \exp(\hat{\mathbf{b}}_i' Z)\}$, $\hat{\Lambda}_{0i}(t)$ 為估計之累積基準危險函數 , $\hat{\mathbf{b}}_i$ 為 \mathbf{b}_i 之最大部分概似函數估計 , 亦即

$$\hat{S}_i(t) = \frac{1}{n_1 + n_2} \sum_{p=1}^2 \sum_{q=1}^{n_p} \hat{S}_i(t|Z_{pq}) , i = 1, 2 .$$

所以為檢定

$$H_0 : (S_1(t) = S_2(t) \text{ for all } t)$$

相對於

$$H_1 : (S_1(t) \leq S_2(t) \text{ for all } t \text{ and } S_1(t) < S_2(t) \text{ for some } t) ,$$

Shen 和 Fleming 建議使用加權平均存活統計量

$$WMS(t) = \sqrt{\frac{n_1 n_2}{n}} \int_0^t \hat{w}(t) [\hat{S}_2(u) - \hat{S}_1(u)] du ,$$

其中 $\hat{w}(t)$ 為權數之估計且與設限時間相關 , 其作用是穩定 WMS 統計量之變異 , 以便求得 WMS 統計量之極限分布。

Shen 和 Fleming 同時指出在比例危險模式 (proportional hazards model) 之下 , Cox 迴歸模式的雙樣本檢定優於 WMS 檢定 , 但是在非比例危險模式或危險函數交叉時 WMS 檢定優於 Cox 迴歸模式之檢定。

3.1.2 雙樣本加權 Kaplan-Meier 群序列檢定 :

針對雙樣本存活資料之群序列分析中不考慮共變數影響時 , Li (1999) 推廣 Pepe 和 Fleming (1989) 的雙樣本加權 Kaplan-Meier 統計量 , 提出加權 Kaplan-Meier 群序列檢定。

假設第 i 組第 j 個病人進入實驗之真正時間為 Y_{ij} , 並且實驗前擬於 t_1, \dots, t_K 時間點進行資料之檢驗。則在日曆時間 (calendar time) t 進行檢驗時 , 所得之雙樣本存活資料為

$$\{ X_{ij}(t), \mathbf{d}_{ij}(t), j=1, \dots, n_i, i=1, 2 \} ,$$

其中

$$X_{ij}(t) = \min\{ T_{ij}, C_{ij}, \max(t - Y_{ij}, 0) \}$$

和設限指標

$$\mathbf{d}_{ij}(t) = I\{T_{ij} < \min(C_{ij}, \max(t - Y_{ij}, 0))\} .$$

對所有 $x \leq t$, 定義

$$N_{ij}(x; t) = I\{X_{ij}(t) \leq x, \mathbf{d}_{ij}(t) = 1\} , i=1, 2, j=1, \dots, n_i ,$$

為第 i 個處理組中第 j 個病人的點過程(counting process) $\bar{N}_i(x; t) = \sum_{j=1}^{n_i} N_{ij}(x; t)$ 為第 i 個處理組在日曆時間 t 進行檢驗時, 存活時間小於或等於 x 的總人數。並且令

$$n_i(x; t) = \sum_{j=1}^{n_i} I\{X_{ij}(t) \geq x\} , i=1, 2 ,$$

為第 i 個處理組在日曆時間 t 進行檢驗且存活時間為 x 時涉險(at risk)的總人數。

所以第 i 個處理組在日曆時間 t 的 Kaplan-Meier 存活函數估計可表示為

$$\hat{S}_i(x; t) = \prod_{s \leq x} \left\{ 1 - \frac{d\bar{N}_i(s; t)}{n_i(s; t)} \right\} , i=1, 2 .$$

針對 H_0 相對於 H_1 的檢定, Li 提出加權 Kaplan-Meier 群序列檢定, 其統計量為

$$WKM(t_i) = \sqrt{\frac{n_1 n_2}{n}} \int_0^{t_i} \hat{w}(s; t_i) \{ \hat{S}_2(s; t_i) - \hat{S}_1(s; t_i) \} ds , i=1, 2, \dots, K$$

其中 $\hat{w}(s; t_i)$ 為權數, 與設限時間及進入時間相關, 主要在穩定 $WKM(t_i)$ 統計量之變異, $\hat{S}_k(s; t_i)$ 為 $S_k(s)$, $k=1, 2$, 在日曆時間 t_i 時的 Kaplan-Meier 存活函數估計。同時利用鞅論證明對所有 $t_1 < \dots < t_K$, $\{WKM(t_1), \dots, WKM(t_K)\}$ 的極限分布為多維常態分布, 並求出其共變異矩陣之一致估計。在顯著水準的配置上, 使用 Slud 和 Wei (1982) 或 Lan 和 DeMets (1983) 的方法配置群序列檢定之顯著水準之 Pocock (1977) 或 O' Brien-Fleming (1979) 邊界。

Li 利用模擬研究得知, 在等比例危險函數形態下, 加權 Kaplan-Meier 群序

列檢定與對數秩群序列檢定的檢定力相當，但在危險函數交叉或早期差異時，加權 Kaplan- Meier 群序列檢定的檢定力高於對數秩群序列檢定的檢定力。

3.1.3 調整共變數後之加權對數秩群序列檢定：

利用群序列檢定比較兩組接受不同處理的存活資料時，若有其他相關共變數時，Tsiatis 等人(1985)建議使用 Cox (1972)迴歸模式，分析兩個處理組之間是否有異。假設隱藏的存活時間為 V ，病人進入實驗的時間為 Y ， W 為研究中感興趣的共變數，例如：處理組的指標函數，假設其均數為 m ，變異數為 s_w^2 。同時有其他與存活時間相關的共變數 Z ，並且 W 與 Z 獨立。所以檢定兩處理組是否有異，亦即檢定存活時間是否與 Z 無關。假設在比例危險模式(proportional hazards model)下，配適 Cox 迴歸模式為：

$$I(x|z, w) = I_0(x) \exp(\mathbf{g}w)h(z, \mathbf{b}) ,$$

其中 $I(x|w, z)$ 為當 $W = w$ 和 $Z = z$ 時的條件危險函數， $I_0(x)$ 為基準危險函數 (base-line hazard function)， $h(z, \mathbf{b})$ 為參數為 \mathbf{b} 的任意特定函數，例如： $h(z, \mathbf{b}) = \exp(\mathbf{b}z)$ 。所以，檢定兩處理組無差異的虛無假說亦可表為 $H_0 : \mathbf{g} = 0$ 。

假設有 n 個病人在時間 Y_1, \dots, Y_n 進入實驗，且 $Y_i \stackrel{iid}{\sim} H(y) = P(Y \leq y)$ ， $i = 1, \dots, n$ 。則所對映第 i 個病人的隨機向量為 (V_i, Y_i, W_i, Z_i) ， $i = 1, \dots, n$ 。若在時間 t 進行檢驗，則觀測時間為 $X_i(t) = \max\{\min(V_i, t - Y_i), 0\}$ ，且設限指標 $\Delta_i(t) = 1$ ，若 $V_i < t - Y_i$ ， $\Delta_i(t) = 0$ ，其他。因為有共變數的存在，Tsiatis 等人根據部分概似 (partial likelihood, Cox, 1975) 檢定虛無假說，在時間 t 之部分概似函數為

$$L(t; \mathbf{g}, \mathbf{b}) = \prod_{i=1}^n \left\{ \frac{\exp(\mathbf{g}W_i)h(Z_i \mathbf{b})}{\sum_{j \in R(t, X_j(t))} \exp(\mathbf{g}W_j)h(Z_j \mathbf{b})} \right\}^{\Delta_i(t)} ,$$

其中 $R(t, x) = \{j : X_j(t) \geq x\}$ 。為檢定虛無假說 $H_0 : \mathbf{g} = 0$ ，得到 score 檢定統計量為 $S_n(t, \hat{\mathbf{b}})$ ，其中

$$\begin{aligned}
S_n(t, \mathbf{b}) &= [\partial \log L(t; \mathbf{g}, \mathbf{b}) / \partial \mathbf{g}]_{\mathbf{g}=\mathbf{0}} \\
&= \sum_{i=1}^n \Delta_i(t) \left\{ Z_i - \frac{\sum_{j \in R\{t, X_i(t)\}} W_j h(Z_j, \mathbf{b})}{\sum_{j \in R\{t, X_i(t)\}} h(Z_j, \mathbf{b})} \right\},
\end{aligned}$$

且 $\hat{\mathbf{b}}$ 為當 $\mathbf{g} = \mathbf{0}$ 時 \mathbf{b} 之最大部分概似估計量。

Tsiatis 等人證明：當在 $t_1 < \dots < t_K$ 進行資料檢驗時， $n^{-\frac{1}{2}} \{S_n(t_1, \hat{\mathbf{b}}), \dots, S_n(t_K, \hat{\mathbf{b}})\}$ 的極限分布為多維常態，均數為零，共變異矩陣為

$$\mathbf{s}(t_i, t_j) = \mathbf{s}_W^2 P(\Delta_i(t) = 1), \quad i \leq j; \quad i, j = 1, \dots, K,$$

並求得共變異矩陣之一致估計。同時利用模擬研究得知，當有共變數存在時，若忽略共變數的影響只使用對數秩統計量檢定時，將損失許多檢定力。所以，當有共變數存在時，共變數的調整是必要的。

3.2 方法推導與建立

假設除存活時間 T 和設限時間 C 外，尚測得共變數 Z ，其中 Z 為 $p \times 1$ 的向量。又第 i 組第 j 個病人進入實驗之真正時間為 Y_{ij} ，則在日曆時間(calendar time) t 進行檢驗時，所得之雙樣本存活資料為

$$\{ X_{ij}(t), \mathbf{d}_{ij}(t), Z_{ij}, j = 1, \dots, n_i, i = 1, 2 \},$$

其中

$$X_{ij}(t) = \min\{ T_{ij}, C_{ij}, \max(t - Y_{ij}, 0) \},$$

和設限指標

$$\mathbf{d}_{ij}(t) = I\{T_{ij} < \min(C_{ij}, \max(t - Y_{ij}, 0))\}.$$

本章擬於 t_1, \dots, t_K 時間點針對虛無假說

$$H_0 : (S_1(t) = S_2(t) \text{ for all } t)$$

相對於對立假說

$$H_1 : (S_1(t) \leq S_2(t) \text{ for all } t \text{ and } S_1(t) < S_2(t) \text{ for some } t),$$

進行群序列檢定(group sequential test)。

本章考慮調整共變數後之加權 Kaplan-Meier 群序列檢定統計量為

$$WMS(t_k) = \sqrt{\frac{n_1 n_2}{n}} \int_0^{t_k} \hat{w}(u; t_k) [\hat{S}_2(u; t_k) - \hat{S}_1(u; t_k)] du, \quad k=1, \dots, K, \quad (3.1)$$

其中

$$\hat{S}_i(u; t_k) = \frac{1}{n_1 + n_2} \sum_{p=1}^2 \sum_{q=1}^{n_p} \hat{S}_i(u; t_k | Z_{pq}), \quad i=1, 2,$$

為 $S_i(u)$ 在 t_k 時之估計式, $\hat{S}_i(u; t_k | Z) = \exp\{-\hat{\Lambda}_{0i}(u; t_k) \exp(\hat{\mathbf{b}}_i^k{}' Z)\}$, $\hat{\Lambda}_{0i}(u; t_k)$ 為基準累積危險函數之估計, $\hat{\mathbf{b}}_i^k$ 為 \mathbf{b}_i 在檢驗時間 t_k 時之最大部分概似估計式。為推廣 Li (1999) 的加權 Kaplan-Meier 群序列檢定, 利用 Shen 和 Fleming (1997) 所提的雙樣本加權平均存活統計量將共變數做調整, 得到加權平均存活群序列檢定 (weighted mean survival group sequential test)。

對所有 $x \leq t$, 定義

$$N_{ij}(x; t) = I\{X_{ij}(t) \leq x, \mathbf{d}_{ij}(t) = 1\} = I[T_{ij} \leq x], \quad i=1, 2, \quad j=1, \dots, n_i,$$

為第 i 個處理組中第 j 個病人的點過程 (counting process), 則 $\bar{N}_i(x; t) = \sum_{j=1}^{n_i} N_{ij}(x; t)$

為第 i 個處理組在日曆時間 t 時, 存活時間小於或等於 x 的總人數。令

$$n_{ij}(x; t) = I\{X_{ij}(t) \geq x\}, \quad i=1, 2, \quad j=1, \dots, n_i,$$

則 $n_i(x; t) = \sum_{j=1}^{n_i} I\{X_{ij}(t) \geq x\}$, $i=1, 2$, 為第 i 個處理組在日曆時間 t 時存活時間至少為 x 的涉險 (at risk) 人數。最後, 定義

$$M_{ij}(x; t) = N_{ij}(x; t) - \int_0^x n_{ij}(u; t) \exp(\mathbf{b}_i' Z_{ij}) d\Lambda_{0i}(u), \quad i=1, 2, \quad j=1, \dots, n_i,$$

為日曆時間 t 時, 均數為 0 之鞅 (martingale), 其中 $\Lambda_{0i}(t)$ 為基準累積危險函數。

若在日曆時間 t 時, 針對各組資料配適 Cox 迴歸模式, 則其對數部分概似函數 (log-partial likelihood function) 為

$$L_i(t; \mathbf{b}_i) = \sum_{j=1}^{n_i} \int_0^t \log \left\{ \frac{\exp(\mathbf{b}_i' Z_{ij})}{\sum_{l=1}^{n_i} n_{il}(u; t) \exp(\mathbf{b}_i' Z_{il})} \right\} dN_{ij}(u; t), \quad i=1, 2.$$

令

$$E_i(\mathbf{b}_i, x; t) = \frac{S_i^{(1)}(\mathbf{b}_i, x; t)}{S_i^{(0)}(\mathbf{b}_i, x; t)},$$

$$S_i^{(r)}(\mathbf{b}_i, x; t) = \frac{1}{n_i} \sum_{j=1}^{n_i} n_{ij}(x; t) \exp(\mathbf{b}_i' Z_{ij}) Z_{ij}^{\otimes r}(x; t),$$

$r = 0, 1, 2$, 其中 $A^{\otimes 0} = 1$, $A^{\otimes 1} = A$ 和 $A^{\otimes 2} = AA'$, 且 $S_i^{(1)}(\mathbf{b}_i, x; t) = \frac{\partial}{\partial \mathbf{b}_i} S_i^{(0)}(\mathbf{b}_i, x; t)$,

$S_i^{(2)}(\mathbf{b}_i, x; t) = \frac{\partial^2}{\partial \mathbf{b}_i^2} S_i^{(0)}(\mathbf{b}_i, x; t)$, $i = 1, 2$ 。則 \mathbf{b}_i 在檢驗時間 t_k 時之最大部分概似估

計式 $\hat{\mathbf{b}}_i^k$, 即為

$$\frac{\partial}{\partial \mathbf{b}_i} L_i(t_k; \mathbf{b}_i) = U_i(\mathbf{b}_i, t_k) = 0$$

之解, 其中

$$U_i(\mathbf{b}_i, t_k) = \sum_{j=1}^{n_i} \int_0^{t_k} \{Z_{ij} - E_i(\mathbf{b}_i, x; t_k)\} dM_{ij}(x; t_k)。$$

且 $\hat{\mathbf{b}}_i^k$ 之共變異矩陣估計式為

$$\hat{\Sigma}_i^k = \int_0^{t_k} \left\{ \frac{S_i^{(2)}(\mathbf{b}_i, x; t_k)}{S_i^{(0)}(\mathbf{b}_i, x; t_k)} - E_i(\mathbf{b}_i, x; t_k)^{\otimes 2} \right\} dN_i(x; t_k)。$$

再令

$$s_i^{(r)}(\mathbf{b}_i, x; t) = E(S_i^{(r)}(\mathbf{b}_i, x; t)), \quad r = 0, 1, 2,$$

$$e_i(\mathbf{b}_i, x; t) = \frac{s_i^{(1)}(\mathbf{b}_i, x; t)}{s_i^{(0)}(\mathbf{b}_i, x; t)}, \quad i = 1, 2。$$

根據 Jennison 和 Turnbull (1997) 結果, 得知在下列二條件之下,

條件一:

假設 \mathbf{b}_i 之真值為 \mathbf{b} , H 為 \mathbf{b}_i 之鄰域(neighborhood), 則對所有的 $k = 1, \dots, K$, 當 $n_i \rightarrow \infty$ 時,

$$\sup\{\|S_i^{(m)}(\mathbf{b}, x; t_k) - s_i^{(m)}(\mathbf{b}, x; t_k)\|; \mathbf{b} \in H; x \in [0, t]\} \xrightarrow{p} 0, \quad m = 0, 1, 2。$$

條件二:

存在 $d > 0$, 使得當 $n_i \rightarrow \infty$ 時,

$$n_i^{-1/2} \sup\{|Z_{ij}| I(\mathbf{b}_i' Z_{ij} > -d | Z_{ij}); j = 1, \dots, n_i\} \xrightarrow{p} 0。$$

下列敘述為真

(1) $U_i(\mathbf{b}_i, t_k) = 0$ 有唯一解 $\hat{\mathbf{b}}_i^k$ 。

(2) 當 $n_i \rightarrow \infty$ 時, $\hat{\mathbf{b}}_i^k \xrightarrow{p} \mathbf{b}_i$ 。

(3) $(\hat{\mathbf{b}}_i^1, \dots, \hat{\mathbf{b}}_i^K)$ 之極限分布為多為常態分布, 且 $n_i^{1/2}(\hat{\mathbf{b}}_i^k - \mathbf{b}_i) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \hat{\mathbf{O}}_i^{k-1})$,
 $\text{Cov}(n_i^{1/2} \hat{\mathbf{b}}_i^{k_1}, n_i^{1/2} \hat{\mathbf{b}}_i^{k_2}) = \text{Var}(n_i^{1/2} \hat{\mathbf{b}}_i^{k_2})$, 對 $1 \leq k_1 \leq k_2 \leq K$ 。當 $n \rightarrow \infty$ 時,
 $\hat{\Sigma}_i^k \xrightarrow{p} \Sigma_i^k$ 。

因為在虛無假說之下, $WMS(t_k)$ 可分解成 $W_2(t_k)$ 與 $W_1(t_k)$ 之差, 其中

$$W_i(t_k) = \sqrt{n_i} \int_0^{t_k} w(x; t_k) \{ \hat{S}_i(x; t_k) - S_i(x) \} dx, \quad i=1, 2.$$

為建立 $WMS(t_k)$, $k=1, \dots, K$, 之極限分布, 本節首先推導 $W_i(t_k)$, $i=1, 2$, 之極限分布。至於 $W_i(t_k)$ 極限分布的推導, 則在 $\hat{S}_i(x; t_k) - S_i(x)$ 減一項 $\bar{S}_i(x)$ 加一項 $\bar{S}_i(x)$ 之後, 根據 $\hat{S}_i(x; t_k) - \bar{S}_i(x)$ 的二維度 Taylor 展開式得到 $\tilde{W}_i(t_k)$, 其中

$$\bar{S}_i(x) \equiv \frac{1}{n} \sum_{p=1}^2 \sum_{q=1}^{n_p} S_i(x | Z_{pq}),$$

$$\begin{aligned} \tilde{W}_i(t_k) = & \int_0^{t_k} w(x; t_k) \mathbf{y}_i'(\mathbf{b}_i, x; t_k) dx \Sigma_i^{k-1} \frac{U_i(\mathbf{b}_i; t_k)}{\sqrt{n_i}} - \int_0^{t_k} h_i(x; t_k) dV_i(x; t_k) \\ & + \sqrt{\frac{n_i}{n}} \frac{1}{\sqrt{n}} \sum_{p=1}^2 \sum_{q=1}^{n_p} \int_0^{t_k} w(x; t_k) \{ S_i(x | Z_{pq}) - E(S_i(x | Z)) \} dx, \end{aligned} \quad (3.2)$$

在(3.2)式中, $V_i(x; t_k)$ 為一個均數為 0 的鞅, 其形式為

$$V_i(x; t_k) = \frac{\sqrt{n_i} \int_0^x d \sum_{j=1}^{n_i} M_{ij}(u; t_k)}{\sum_{j=1}^{n_i} n_{ij}(x; t_k) \exp(\mathbf{b}_i' Z_{ij})},$$

$$h_i(x; t_k) = \int_x^{t_k} w(u; t_k) E[\exp(\mathbf{b}_i' Z) S_i(u | Z)] du,$$

$$\begin{aligned} \mathbf{y}_i(\mathbf{b}_i, x; t_k) = & \int_0^x [e_i(\mathbf{b}_i, u; t_k) E\{\exp(\mathbf{b}_i' Z) S_i(x | Z)\} \\ & - E\{Z \exp(\mathbf{b}_i' Z) S_i(x | Z)\}] d\Lambda_{0i}(u), \end{aligned}$$

並且

$$\Sigma_i^k = \int_0^{t_k} \left\{ \frac{s_i^{(2)}(\mathbf{b}_i, x; t_k)}{s_i^{(0)}(\mathbf{b}_i, x; t_k)} - e_i(\mathbf{b}_i, x; t_k)^{\otimes 2} \right\} s_i^{(0)}(\mathbf{b}_i, x; t_k) d\Lambda_{0i}(x).$$

事實上 $\tilde{W}_i(t_k)$ 為兩個獨立鞅加上一個獨立且同分布隨機變數之和。本章先在引理

3.1 證明 $W_i(t)$ 與 $\tilde{W}_i(t)$ 漸進相等(asymptotically equivalent)。然後於引理 3.2 利用鞅中央極限定理及中央極限定理求得 $\tilde{W}_i(t)$ 之極限分布。最後在定理 3.3 中求證 $WMS(t_k)$ 之極限分布。

引理 3.1

$W_i(t)$ 與 $\tilde{W}_i(t)$ 漸進相等，其中漸進相等為兩統計量之差在 $[0, t_k]$ 間均勻地機率收斂到 0，亦即

$$\sup_{0 \leq t \leq t_k} \|W_i(t) - \tilde{W}_i(t)\| \xrightarrow{p} 0。$$

證明：

首先證明 $\{\hat{S}_i(x; t_k) - \bar{S}_i(x)\} / \sqrt{n_i}$ 與 $g_i(x; t_k)$ 漸進相等。其中

$$\begin{aligned} g_i(x; t_k) &= \mathbf{y}_i(\mathbf{b}_i, x; t_k) \Sigma_i^{k-1} \frac{U_i(\mathbf{b}_i; t_k)}{\sqrt{n_i}} \\ &\quad - E\{\exp(\mathbf{b}_i' Z) S_i(x | Z)\} \sqrt{n_i} \int_0^x \frac{d\bar{M}_i(u; t_k)}{n_i S_i^{(0)}(\mathbf{b}_i, u; t_k)}。 \end{aligned}$$

利用二維度 Taylor 展開式將 $\hat{S}_i(x; t_k | Z_{ij})$ 在 $(\Lambda_{0i}(x), \mathbf{b}_i')$ 展開，所以

$$\begin{aligned} &\sqrt{n_i} \{\hat{S}_i(x; t_k) - \bar{S}_i(x)\} \\ &\approx - \left\{ \frac{1}{n_i} \sum_{j=1}^{n_i} \exp(\mathbf{b}_i' Z_{ij}) S_i(x | Z_{ij}), \Lambda_{0i}(x) \frac{1}{n_i} \sum_{j=1}^{n_i} Z_{ij} \exp(\mathbf{b}_i' Z_{ij}) S_i(x | Z_{ij}) \right\} \\ &\quad \times \sqrt{n_i} \begin{pmatrix} \hat{\Lambda}_{0i}(x; t_k) - \Lambda_{0i}(x) \\ \hat{\mathbf{b}}_i^k - \mathbf{b}_i \end{pmatrix}， \end{aligned}$$

應用 Anderson 和 Gill (1982) 文中定理 .1 可知

$$\sup_{\{\mathbf{b} \in B, t \in [0, t_k]\}} \left\| \frac{1}{n_i} \sum_{j=1}^{n_i} Z_{ij} \exp(\mathbf{b}_i' Z_{ij}) S_i(t | Z_{ij}) - E\{Z \exp(\mathbf{b}_i' Z) S_i(t | Z)\} \right\| \xrightarrow{p} 0$$

和

$$\sup_{\{\mathbf{b} \in B, t \in [0, t_k]\}} \left| \frac{1}{n_i} \sum_{j=1}^{n_i} \exp(\mathbf{b}_i' Z_{ij}) S_i(t | Z_{ij}) - E\{\exp(\mathbf{b}_i' Z) S_i(t | Z)\} \right| \xrightarrow{p} 0。$$

且

$$\sqrt{n_i}(\hat{\mathbf{b}}_i^k - \mathbf{b}_i) \approx \sqrt{n_i} \Sigma_i^{k-1} U_i(\mathbf{b}_i; t_k) ,$$

又

$$\begin{aligned} & \sqrt{n_i}(\hat{\Lambda}_{0i}(t; t_k) - \Lambda_{0i}(t)) \\ & \approx -\sqrt{n_i} \int_0^t e_i(\mathbf{b}_i, x; t_k) d\Lambda_{0i}(x) (\hat{\mathbf{b}}_i^k - \mathbf{b}_i) + \int_0^t \frac{d\bar{M}_i(x; t_k)}{n_i S_i^{(0)}(\mathbf{b}_i, x; t_k)} \circ \end{aligned}$$

因此， $\{\hat{S}_i(x; t_k) - \bar{S}_i(x)\} / \sqrt{n_i} \approx \mathbf{g}_i(x; t_k)$ 。所以，

$$\begin{aligned} \int_0^{t_k} w(x; t_k) \mathbf{g}_i(x; t_k) dx &= \int_0^{t_k} w(x; t_k) \mathbf{y}_i(\mathbf{b}_i, x; t_k) dx \Sigma_i^{k-1} \frac{U_i(\mathbf{b}_i; t_k)}{\sqrt{n_i}} \\ & - \int_0^{t_k} w(x; t_k) E\{\exp(\mathbf{b}_i' Z) S_i(x | z)\} V_i(x; t_k) dx , \end{aligned} \quad (3.3)$$

利用部分積分則

$$\int_0^{t_k} w(x; t_k) E\{\exp(\mathbf{b}_i' Z) S_i(x | z)\} V_i(x; t_k) dx = \int_0^{t_k} h_i(x; t_k) dV_i(x; t_k) ,$$

其中

$$h_i(x; t_k) = \int_x^{t_k} w(x; t_k) E\{\exp(\mathbf{b}_i' Z) S_i(x | Z)\} dx$$

為有界可預測之過程(bounded predictable process)。因此，得證。

由引理 3.1 可知 $W_i(t_k)$ 與 $\tilde{W}_i(t_k)$ 漸進相等，因此，若能求得 $\tilde{W}_i(t_k)$ 之極限分布時，便亦可知道 $W_i(t_k)$ 之極限分布。引理 3.2 將證明 $\tilde{W}_i(t_k)$ 之極限分布為一常態分布。

引理 3.2

若 $n \rightarrow \infty$ 且 $n_i/n \rightarrow \mathbf{r}_i$ ， $i=1, 2$ ，則

$$\tilde{W}_i(t_k) \xrightarrow{d} N(0, \mathbf{s}_i^2(t_k)) ,$$

其中

$$\begin{aligned} \mathbf{s}_i^2(t_k) &= \left\{ \int_0^{t_k} w(x; t_k) \mathbf{y}_i(\mathbf{b}_i, x; t_k) dx \right\} \Sigma_i^{k-1} \left\{ \int_0^{t_k} w(x; t_k) \mathbf{y}_i(\mathbf{b}_i, x; t_k) dx \right\} \\ &\quad + \int_0^{t_k} h_i^2(x; t_k) \frac{d\Lambda_{0i}(x)}{e_i^{(0)}(\mathbf{b}_i, x; t_k)} + \mathbf{r}_i \text{Var} \left\{ \int_0^{t_k} w(x; t_k) S_i(x | Z) dx \right\}. \end{aligned} \quad (3.4)$$

證明：

由引理 3.1 知 $\int_0^{t_k} w(x; t_k) \mathbf{g}_i(x; t_k) dx$ 為一平方可積之鞅 (square integrable martingale)。若 $U_{il}(\mathbf{b}_i; \cdot)$ 為其第 l 個元素, $l = 1, \dots, p$, 則

$$\begin{aligned} &\left\langle U_{il}(\mathbf{b}_i; \cdot), \int_0^\cdot h_i(x; \cdot) dV_i(x; \cdot) \right\rangle(t) \\ &= \int_0^t \left\{ \frac{S_{il}^{(1)}(\mathbf{b}_i, x; t)}{S_i^{(0)}(\mathbf{b}_i, x; t)} - E_{il}(\mathbf{b}_i, x; t) \right\} h_i(x; t) d\Lambda_{0i}(x) = 0. \end{aligned}$$

所以, $U_i(\mathbf{b}_i; \cdot)$ 與 $\int_0^\cdot h_i(x; \cdot) dV_i(x; \cdot)$ 垂直。又因為,

$$\int_0^{t_k} w(x; t_k) \{ \bar{S}_i(x) - S_i(x) \} dx \quad (3.5)$$

可以改寫成獨立且同分布, 均數為 0 和有限變異數之隨機變數之和。所以利用中央極限定理 (central limit theorem), 可得知其極限分布。因為 Shen 和 Fleming (1995) 已證明對任一局部有界且 F_t 可預測之過程 (locally bounded and F_t -predictable process) H , $E\left\{ \int_0^t H(s) dM(s) | F_Z \right\} = 0$, 其中 F_Z 為 Z 之過濾 (filtration)。所以,

$$\begin{aligned} &\text{Cov} \left\{ U_{il}(\mathbf{b}_i, t_k), \int_0^{t_k} w(s; t_k) \{ \bar{S}_i(s) - S_i(s) \} ds \right\} \\ &= E \left(\left[\int_0^{t_k} w(s; t_k) \{ \bar{S}_i(s) - S_i(s) \} ds \right] E \{ U_{il}(\mathbf{b}_i; t_k) | F_Z \} \right) = 0. \end{aligned}$$

同樣的, 亦可知

$$\text{Cov} \left\{ \int_0^{t_k} h_i(s; t_k) dV_i(s; t_k), \int_0^{t_k} w(s; t_k) \{ \bar{S}_i(s) - S_i(s) \} ds \right\} = 0.$$

因此, (3.5) 式與 (3.3) 式垂直, 得知 $\tilde{W}_i(t_k)$ 分布收斂至兩個漸進獨立之高斯過程且變異數為 $\mathbf{s}_i^2(t_k)$ 。

引理 3.1 及 3.2 中說明 $W_i(t_k)$ 與 $\tilde{W}_i(t_k)$ 之極限分布皆為均數為 0，但變異數為 $\mathbf{s}_i^2(t_k)$ 之常態分布，其中變異數 $\mathbf{s}_i^2(t_k)$ 之一致估計式為

$$\begin{aligned} \hat{\mathbf{s}}_i^2(t_k) = & \left\{ \int_0^{t_k} \hat{w}(x; t_k) \mathbf{y}_i(\hat{\mathbf{b}}_i^k, x; t_k) dx \right\} \hat{\Sigma}_i^{k-1} \left\{ \int_0^{t_k} \hat{w}(x; t_k) \mathbf{y}_i(\hat{\mathbf{b}}_i^k, x; t_k) dx \right\} \\ & + \int_0^{t_k} \hat{h}_i^2(x; t_k) \frac{d\hat{\Lambda}_{0i}(x; t_k)}{E_i^{(0)}(\hat{\mathbf{b}}_i^k, x; t_k)} + \frac{n_i}{n} \widehat{Var} \left\{ \int_0^{t_k} w(x; t_k) S_i(x|Z) dx \right\}, \end{aligned} \quad (3.6)$$

$$\hat{h}_i(t; t_k) = \int_t^{t_k} \hat{w}(x; t_k) \frac{1}{n} \sum_{p=1}^2 \sum_{q=1}^{n_p} \exp(\hat{\mathbf{b}}_i^k' Z_{pq}) \hat{S}_i(x|Z_{pq}) dx, \quad ,$$

$$\hat{\Sigma}_i^k = \int_0^{t_k} \left\{ \frac{S_i^{(2)}(\hat{\mathbf{b}}_i^k, x; t_k)}{S_i^{(0)}(\hat{\mathbf{b}}_i^k, x; t_k)} - E_i(\hat{\mathbf{b}}_i^k, x; t_k)^{\otimes 2} \right\} S_i^{(0)}(\hat{\mathbf{b}}_i^k, x; t_k) d\hat{\Lambda}_{0i}(x; t_k).$$

所以，在虛無假說下，檢定統計量 $WMS(t_k)$ 之極限分布便可求得。定理 3.3 將證明 $WMS(t_k)$ 之極限分布亦為常態分布。

定理 3.3

在虛無假說下，若 $n \rightarrow \infty$ 且 $n_i/n \rightarrow \mathbf{r}_i$ ， $i=1, 2$ ，則

$$WMS(t_k) \xrightarrow{d} N(0, \mathbf{s}_W^2(t_k)), \quad ,$$

其中

$$\begin{aligned} \mathbf{s}_W^2(t_k) = & \sum_{i=1}^2 (1 - \mathbf{r}_i) \mathbf{s}_i^2(t_k) \\ & - 2\mathbf{r}_1 \mathbf{r}_2 \text{Cov} \left\{ \int_0^{t_k} w(x; t_k) S_1(x|Z) dx, \int_0^{t_k} w(x; t_k) S_2(x|Z) dx \right\}. \end{aligned}$$

證明：

在虛無假說下，對檢驗時間點 t_k ， $WMS(t_k)$ 可寫成 $\sqrt{\frac{n_1}{n}} W_2(t_k) - \sqrt{\frac{n_2}{n}} W_1(t_k)$ ，

由引理 3.1 及 3.2 知 $\sqrt{\frac{n_1}{n}} W_2(t_k) - \sqrt{\frac{n_2}{n}} W_1(t_k)$ 與 $\sqrt{\frac{n_1}{n}} \tilde{W}_2(t_k) - \sqrt{\frac{n_2}{n}} \tilde{W}_1(t_k)$ 漸進相等，

且具有相同極限分布，其極限分布為均數為 0 變異數為 $\mathbf{s}_w^2(t_k)$ 之常態分布，其中

$$\begin{aligned} \mathbf{s}_w^2(t_k) &= \sum_{i=1}^2 (1 - \mathbf{r}_i) \mathbf{s}_i^2(t_k) \\ &\quad - 2\mathbf{r}_1 \mathbf{r}_2 \text{Cov}\left\{\int_0^{t_k} w(x; t_k) S_1(x|Z) dx, \int_0^{t_k} w(x; t_k) S_2(x|Z) dx\right\}. \end{aligned}$$

定理 3.3 已證明在虛無假說下對任意檢驗點 t_k $WMS(t_k)$ 之極限分布，其漸進變異數 $\mathbf{s}_w^2(t_k)$ 之一致估計式為

$$\begin{aligned} \hat{\mathbf{s}}_w^2(t_k) &= \sum_{i=1}^2 \frac{n_{3-i}}{n} \hat{\mathbf{s}}_i^2(t_k) \\ &\quad - 2 \frac{n_1 n_2}{n^2} \widehat{\text{Cov}}\left\{\int_0^{t_k} w(x; t_k) S_1(x|Z) dx, \int_0^{t_k} w(x; t_k) S_2(x|Z) dx\right\}, \end{aligned} \quad (3.7)$$

其中 $\hat{\mathbf{s}}_i^2(t_k)$ 列在(3.6)式。為能在檢驗時間點 $t_1 < \dots < t_K$ 之下，建構雙樣本加權平均存活群序列檢定，本章在定理 3.4 求證在虛無假說之下 $\{WMS(t_1), \dots, WMS(t_K)\}$ 之極限分布。

定理 3.4

在虛無假說之下，若 $n \rightarrow \infty$ 且 $n_i/n \rightarrow \mathbf{r}_i$ ， $i=1, 2$ ，則

$$\{WMS(t_1), \dots, WMS(t_K)\} \xrightarrow{d} \mathbf{N}(\mathbf{0}, \hat{\mathbf{O}}),$$

其中 $\hat{\mathbf{O}} = (\hat{\phi}_w(t_p, t_q))_{K \times K}$ ，對任意 $t_p < t_q$

$$\begin{aligned} \mathbf{s}_w(t_p, t_q) &= \sum_{i=1}^2 \left\{ \mathbf{r}_{3-i} \left(\int_0^{t_p} w(x; t_p) \mathbf{y}_i(\mathbf{b}_i, x; t_p) dx \right)' \Sigma_i^{q-1} \left(\int_0^{t_q} w(x; t_q) \mathbf{y}_i(\mathbf{b}_i, x; t_q) dx \right) \right. \\ &\quad \left. + \int_0^{t_p} h_i(x; t_p) h_i(x; t_q) \frac{d\Lambda_{0i}(x)}{S_i^{(0)}(\mathbf{b}_i, x; t_q)} \right\} + \mathbf{r}_1 \mathbf{r}_2 \times \\ &\quad \text{Cov}\left(\int_0^{t_p} w(x; t_p) [S_2(x|Z) - S_1(x|Z)] dx, \int_0^{t_q} w(x; t_q) [S_2(x|Z) - S_1(x|Z)] dx\right). \end{aligned} \quad (3.8)$$

證明：

給定任意不完全為 0 之常數 a_1, \dots, a_K ，得到

$$\sum_{i=1}^K a_i WMS(t_i) = \sum_{i=1}^K a_i W_2(t_i) - \sum_{i=1}^K a_i W_1(t_i) ,$$

並利用引理 3.1 及 3.2 知其與

$$\sum_{i=1}^K a_i \tilde{W}_2(t_i) - \sum_{i=1}^K a_i \tilde{W}_1(t_i) \quad (3.9)$$

漸進相等，且 (3.9) 式可化簡為兩個獨立鞅之和再加上一個獨立且同分布隨機變數之和。所以，利用鞅中央極限定理、中央極限定理及 Cramar-Wold device，得知 $\{WMS(t_1), \dots, WMS(t_K)\}$ 之極限分布為 K 維常態分布，其共變異矩陣利用引理 3.2 可化簡成 (3.8) 式。

進一步可求 $\{WMS(t_1), \dots, WMS(t_K)\}$ 極限分布之共變異矩陣 $\hat{\mathbf{O}} = (\hat{\phi}_w(t_p, t_q))_{K \times K}$ 之一致估計式為 $\hat{\mathbf{O}} = (\hat{\phi}_w(t_p, t_q))_{K \times K}$ ，其中

$$\begin{aligned} \hat{\mathbf{S}}_w(t_p, t_q) &= \sum_{i=1}^2 \left\{ \frac{n_{3-i}}{n} \left(\int_0^{t_p} \hat{w}(x; t_p) \mathbf{y}_i(\hat{\mathbf{b}}_i^k, x; t_p) dx \right)' \hat{\Sigma}_i^{q-1} \left(\int_0^{t_q} \hat{w}(x; t_q) \mathbf{y}_i(\hat{\mathbf{b}}_i^k, x; t_q) dx \right) \right. \\ &\quad \left. + \int_0^{t_p} \hat{h}_i(x; t_p) \hat{h}_i(x; t_q) \frac{d\hat{\Lambda}_{0i}(x; t_p)}{S_i^{(0)}(\hat{\mathbf{b}}_i^k, x; t_q)} \right\} + \frac{n_1 n_2}{n^2} \times \\ &\quad \widehat{Cov} \left(\int_0^{t_p} w(x; t_p) [S_2(x|Z) - S_1(x|Z)] dx, \int_0^{t_q} w(x; t_q) [S_2(x|Z) - S_1(x|Z)] dx \right). \end{aligned} \quad (3.10)$$

檢定統計量 $WMS(t)$ 中權數的選擇，是使得 $WMS(t)$ 之變異數穩定。假設設限變數與共變數獨立，則其充分條件為：給定常數 $\Gamma > 0$ ，

$$|w(s; t)| \leq \Gamma \{G_i^*(s; t)\}^{1/2} , \quad (3.11)$$

$i = 1, 2$ 。因此選擇一個簡單而又滿足 (3.11) 式之權數為：

$$w(s; t) = \frac{G_1^*(s; t) G_2^*(s; t)}{(n_1/n) G_1^*(s; t) + (n_2/n) G_2^*(s; t)} ,$$

其中 $G_i^*(s;t) = P(C_{ij} \wedge (t - Y_{ij})^+ \geq s)$ 。若無設限情形，且各組進入時間服從相同分布時，則

$$\hat{w}(s;t) = \hat{G}^*(s;t) = \sum_{i=1}^2 \sum_{j=1}^{n_i} I_{\{Y_{ij} \leq t-s\}} / n$$

為 $w(s;t) = G^*(s;t) = P((t - Y_{ij})^+ \geq s)$ 之估計，亦即在時間 t 時，病人接受處理時間至少是 s 之機率。

令第 k 次檢驗時之檢定統計量為

$$Z(t_k) = WMS(t_k) / \sqrt{\hat{\mathbf{s}}_W^2(t_k)} \quad , k = 1, \dots, K ,$$

其中 $\hat{\mathbf{s}}_W^2(t_k)$ 為 $\mathbf{s}_W^2(t_k)$ 之一致估計量。則在第 k 次檢驗時，群序列分析之棄卻域為：

$$Z(t_k) > c_k \quad , k \leq K .$$

此群序列之臨界值 c_k , $k = 1, \dots, K$, 可由

$$P\{Z(t_1) \leq c_1, \dots, Z(t_{i-1}) \leq c_{i-1}, Z(t_i) > c_i\} = \mathbf{p}_i \quad (3.12)$$

求得，其中 $\{\mathbf{p}_i, i = 1, \dots, K\}$ 滿足 $\sum_{i=1}^K \mathbf{p}_i = \mathbf{a}$ 。又因為 $\{Z(t_1), \dots, Z(t_K)\}$ 之極限分布服從 K 維常態分布，而 $Z(t_i)$ 與 $Z(t_j)$ 之漸進相關係數亦可利用 $\hat{\mathbf{s}}_W(t_i, t_j) / \sqrt{\hat{\mathbf{s}}_W^2(t_i) \hat{\mathbf{s}}_W^2(t_j)}$ 估計之。因此，在給定 \mathbf{p}_i 後(3.12)式中之臨界值便可利用 Drezner (1992) 之程式計算。

至於 \mathbf{p}_i 的選擇有二種方法，第一種方法為 Slud 和 Wei (1982) 所建議，在試驗前先給定之特定值，例如： $K=3$, $\mathbf{a} = 0.05$ 時，給定 \mathbf{p}_i , $i=1,2,3$, 分別為 0.01 , 0.015 和 0.025。所得之邊界與 O' Brien-Fleming (1979) 邊界相似。另一種方法為 Lan 和 DeMets (1983) 所提出，利用給定之消耗函數(use function) 計算出 \mathbf{p}_i 之值。至於分割的方法，一方面需給定特定之消耗函數，另一方面需定義訊息(information) 及訊息分數(information fraction)。消耗函數的形式很多，一般而言，是使用 Pocock (1977) 形式或 O' Brien-Fleming 形式之消耗函數。其中 Pocock 形式之消耗函數為

$$\mathbf{a}_1^*(t) = \mathbf{a} \ln\{1 + (e-1)t\} \quad , t = i/K \quad , i = 1, \dots, K \quad ,$$

且 $\mathbf{a}_1^*(1/2) \approx 0.62\mathbf{a}$ ，表示會有提早結束的傾向。而 O' Brien-Fleming 形式之消耗函數為

$$\mathbf{a}_2^*(t) = \begin{cases} 0 & t = 0 \\ 2 - 2\Phi\left(\frac{Z_{\mathbf{a}/2}}{\sqrt{t}}\right) & 0 < t \leq 1, \quad t = i/K, \quad i = 1, \dots, K. \end{cases}$$

對 $\mathbf{a} = 0.05$ 時， $\mathbf{a}_2^*(1/2) < 0.006 \approx 0.12\mathbf{a}$ ，則適合長期執行之臨床試驗。訊息部份，則可以使用群序列分析中觀察到存活時間之次數替代之，所以第 k 次檢驗之訊息分數為第 k 次檢驗時觀察到存活時間之次數除以預期試驗中觀察到存活時間之總次數。因此，假設消耗函數為 $\mathbf{a}^*(t)$ ，在第 k 次檢驗時所選擇的訊息分數估計為 \hat{v}_k ，所以，在第 k 次檢驗分析時使用之型 I 誤差為 $\mathbf{p}_k = \mathbf{a}^*(\hat{v}_k) - \mathbf{a}^*(\hat{v}_{k-1})$ 。

3.3 實例分析

本章將引用 Byar (1985)之前列腺癌資料。此資料比較四種不同處理：對照組(安慰劑(placebo)或無 Diethylstilbestrol(DES)) 0.2mg、1.0mg 和 5.0mg 之 DES 之隨機化臨床試驗。也就是說，前列腺癌病人進入試驗後，被隨機分配服用零劑量之 DES、0.2mg 之 DES、1.0mg 之 DES 或 5.0mg 之 DES，然後觀測其死亡時間。在第四章中針對不同癌症期數做傘檢定之分析得知，在癌症三期時傘頂位置在第三組，但是在癌症四期時，傘頂位置在第四組。因此，考慮合併第三組和第四組成為處理組，研究處理組是否優於對照組。病人之收集期從 1967 年 4 月至 1969 年 12 月止，其中對照組共收集 128 位病人，1.0mg 和 5.0mg 之 DES 組共收集 253 位病人。預定分別於 1968 年 12 月、1670 年 12 月、1972 年 12 月和 1974 年 12 月對資料進行檢定。預計期望觀察到之存活時間（因前列腺癌死亡）之總個數為 88 人，而在每次檢驗時觀察到之存活時間之個數分別為 15 人、57 人、82 人和 88 人。因此，訊息分數分別為 15/88，57/88 和 82/88。在此取 $\mathbf{a} = 0.05$ ，且利用 Lan 和 DeMets (1983)的消耗函數，建構 Pocock's-type 和 O' Brien-Fleming's-type 之單尾檢定邊界。在此為說明調整共變數之重要性，分別計算

WKM 群序列檢定統計量（未調整）與 WMS 群序列檢定統計量（調整）之值，其結果列於表 3-1。

由表 3-1 可知，WMS 群序列檢定於第三次檢驗時已有充分證據證明兩組存活函數不同，而 WKM 群序列檢定在使用 O' Brien-Fleming' s-type 之邊界時，在第三次檢驗時也有充分證據證明兩組存活函數不同。但是，若是使用 Pocock' s-type 之邊界時，直到結束試驗時證據都還不夠充分拒絕兩組存活函數相同之虛無假說。由於癌症期數確實會影響存活時間，所以，經過調整癌症期數之 WMS 群序列檢定之檢定結果，比未調整之 WKM 群序列檢定結果更具檢定力。因此，由調整癌症期數後之 WMS 群序列檢定結果可知，在 1972 年 12 月檢驗時，已有足夠證據證明服用 1.0mg 以上劑量之 DES 確實可以延長前列腺癌之存活時間。

表 3-1 the prostatic cancer data 之群序列檢定統計量， $\alpha = 0.05$ 。

時間 統計量	Dec., 68	Dec., 70	Dec., 72	Dec., 74
WMS*	1.92	1.08	1.88	1.93
Boundary ¹	2.23	1.82	1.80	1.81
Boundary ²	4.01	2.17	1.73	1.70
相關矩陣	1	0.859	0.739	0.714
		1	0.943	0.923
			1	0.980
WKM*	1.82	1.03	1.75	1.79
Boundary ¹	2.23	1.84	1.84	1.90
Boundary ²	4.01	2.17	1.74	1.71
相關矩陣	1	0.814	0.672	0.641
		1	0.929	0.901
			1	0.980

*:為標準化之統計量

1: Pocock' s-type 單尾檢定邊界

2: O' Brien-Fleming-type 單尾檢定邊界

第四章 加權 Kaplan-Meier 傘檢定

本章討論多樣本存活資料之傘檢定問題。根據雙樣本加權 Kaplan-Meier 檢定統計量建立傘頂已知和未知之加權 Kaplan-Meier 傘檢定。除評估傘頂已知之 *WKM* 傘檢定相對於 *WLR* 傘檢定(Chen 和 Wolfe, 2000)的漸近相對效率，且模擬傘頂未知 *WKM* 傘檢定之表現。最後以實例說明其應用。

4.1 加權對數秩傘檢定：

由於比較的對象為加權對數秩傘檢定(Chen 和 Wolfe, 2000)，在此先回顧 Chen 和 Wolfe (2000)所提之加權對數秩傘檢定。

在劑量反應研究中，若反應值隨著劑量漸增而遞增，但是受制於在高劑量下可能產生之反效果，之後反應值隨劑量漸增而遞減，此時，劑量反應呈先增後減之傘形型態。針對多樣本的右設限存活資料，Chen 和 Wolfe (2000)根據雙樣本加權對數秩統計量建構傘檢定。

令 k 組不同處理水準之存活資料為

$$T_{i1}, \dots, T_{in_i} \stackrel{iid}{\sim} S_i(x) \quad i = 1, \dots, k,$$

其設限資料為

$$C_{i1}, \dots, C_{in_i} \stackrel{iid}{\sim} C_i(x) \quad i = 1, \dots, k,$$

其中 $S_i(t)$ 和 $C_i(t)$ 分別為 T_{ij} 和 C_{ij} 之存活函數。又假設 T_{ij} 和 C_{ij} 為獨立之隨機變數。在右設限之下，我們真正觀察到的資料為

$$X_{ij} = \min\{T_{ij}, C_{ij}\}$$

和設限指標

$$\delta_{ij} = I\{T_{ij} \leq C_{ij}\} \quad j = 1, \dots, n_i, \quad i = 1, \dots, k。$$

令 $S_i = 1 - F_i$ 為第 i 組之存活函數。

在檢定 $H_0 : (S_1(t) = S_2(t) \text{ for all } t)$ 相對於 $H_1 : (S_1(t) \leq S_2(t) \text{ for all } t \text{ and}$

$S_1(t) < S_2(t)$ for some t 的問題中，雙樣本 WLR 統計量為

$$U_{12} = \int_0^{t_c} W(t) d\{\hat{\Lambda}_1(t) - \hat{\Lambda}_2(t)\} ,$$

其中 $t_c = \min\{t_1, t_2\}$ ， t_i 為各組最後的觀測值， $W(t)$ 為加權函數， $\hat{\Lambda}_i(t)$ 為第 i 組累積危險函數的估計量。為檢定虛無假說

$$H_0 : (S_1(t) = S_2(t) = \dots = S_k(t) \text{ for all } t)$$

相對於傘形對立假說

$$H_{1U} : (S_1(t) \leq \dots \leq S_p(t) \geq \dots \geq S_k(t) \text{ for some } t \text{ with at least one strict inequality}) ,$$

其中 p 為傘檢定之傘頂，Chen 和 Wolfe 在傘頂已知時推廣 Mack-Wolfe (1981) 的無母數傘檢定得到統計量

$$A_p = \sum_{1 \leq i < j \leq p} U_{ij} + \sum_{p \leq i < j \leq k} U_{ji} ,$$

其中 U_{ij} 為比較第 i 個和第 j 個樣本的雙樣本加權對數秩統計量。其利用鞅論證明出 A_p 的極限分布為常態分布，並求出變異數 $Var(A_p)$ 之一致估計 $\hat{Var}(A_p)$ ，然後計算

$$A_p^* = A_p / \sqrt{\hat{Var}(A_p)} .$$

最後獲得傘頂已知的檢定為

$$\text{棄卻 } H_0 \text{ 若 } A_p^* \geq z(\mathbf{a}) ,$$

其中 $z(\mathbf{a})$ 為標準常態分布之 $100(1-\mathbf{a})$ 百分位。

當傘頂未知時，Chen 和 Wolfe 推廣 Chen-Wolfe (1990) 的無母數傘檢定方法，考慮檢定統計量

$$A_{\max}^* = \max\{A_1^*, \dots, A_k^*\}$$

其中 $A_i^* = A_i / \sqrt{\hat{Var}(A_i)}$ ， $i = 1, \dots, k$ 。除利用鞅論證明 (A_1^*, \dots, A_k^*) 的極限分布為 k 維常態分布，並推導其相關矩陣，計算 A_{\max}^* 極限分布之 $100(1-\mathbf{a})$ 百分位，記為 $z_{\max}(\mathbf{a})$ 。最後得到在傘頂未知的檢定為

$$\text{棄卻 } H_0 \text{ 若 } A_{\max}^* \geq z_{\max}(\mathbf{a}) .$$

Chen 和 Wolfe 利用模擬研究得知，在等比例危險函數模式下，加權對數秩族中的對數秩傘檢定的檢定力是最高的。若是危險函數在早期出現差異時，

Peto-Prentice-Wilcoxon (Peto 和 Peto, 1972, Prentice, 1978)傘檢定的檢定力是最高的。同時，在檢定中錯估傘頂位置時，其錯估傘頂的檢定力會比使用傘頂未知的傘檢定的檢定力更低。因此，建議在傘頂不確定時，採用傘頂未知的傘檢定。

4.2 傘頂已知之傘檢定

針對傘頂已知之傘檢定問題，本文推廣 Mack-Wolfe (1981)的無母數傘檢定方法，考慮之檢定統計量為

$$WKM_p = \sum_{1 \leq i < j \leq p} WKM_{ij} + \sum_{p \leq i < j \leq k} WKM_{ji} , \quad (4.1)$$

其中 WKM_{ij} 為比較第 i 個和第 j 個樣本的加權 Kaplan-Meier 統計量(1.1)。當資料完整觀測時， WKM_p 可寫成 $\sum_{1 \leq i < j \leq p} c_{ij}(\bar{X}_j - \bar{X}_i) + \sum_{p \leq i < j \leq k} c_{ji}(\bar{X}_i - \bar{X}_j)$ ，其中 \bar{X}_i 為第 i 個樣本的均數， $i = 1, \dots, k$ 。所以， WKM_p 檢定統計量可視為利用均數建立之傘檢定在右設限存活資料的推廣。

為建構 WKM_p 檢定，我們將利用下面定理得到 WKM_p 統計量在虛無假說下的極限分布。

定理 4.1 令 $\hat{p}_i = n_i/N \rightarrow p_i$ ， $0 < p_i < 1$ ， $i = 1, \dots, k$ 。假設 $Y_i(t)/n_i$ 在 $t \in [0, t]$ 均勻地機率收斂至 $p_i(t) = C_i(t)S_i(t)$ ，其中 $Y_i(t) = \sum_{j=1}^{n_i} I_{[X_{ij} \geq t]}$ 為第 i 組中涉險(at-risk)人數， $i = 1, \dots, k$ ，且 $I_{[A]}$ 為事件 A 之指標函數， $t = \sup\{t : \min\{S(t), C_1(t), \dots, C_k(t)\} > 0\}$ ， $S(t)$ 為在 H_0 之存活函數。則 WKM_p 統計量在 H_0 下之極限分布為常態分布其均數為 0 變異數為 s_p^2 ，其中

$$s_p^2 = \sum_{j=1}^k \int_0^t h_j(t) \{1 - \Delta\Lambda(t)\} d\Lambda(t) , \quad (4.2)$$

$$h_j(t) = \begin{cases} \left[\sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \sqrt{p_i} \operatorname{sgn}(j-i) \left\{ \int_t^t w_{ij}(u) S(u) du \right\} \right]^2 / \mathbf{p}_j(t) & \text{if } 1 \leq j \leq p-1 ; \\ \left[\sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \sqrt{p_i} \left\{ \int_t^t w_{ip}(u) S(u) du \right\} \right]^2 / \mathbf{p}_p(t) & \text{if } j = p ; \\ \left[\sum_{\substack{i \neq j \\ p \leq i \leq k}} \sqrt{p_i} \operatorname{sgn}(i-j) \left\{ \int_t^t w_{ji}(u) S(u) du \right\} \right]^2 / \mathbf{p}_j(t) & \text{if } p+1 \leq j \leq k . \end{cases} \quad (4.3)$$

同時

$$w_{ij}(t) = \frac{C_i(t-)C_j(t-)}{p_i C_i(t-) + p_j C_j(t-)} ,$$

$\text{sgn}(x) = 1$ if $x > 0$ 和 -1 if $x < 0$, $\Lambda(\cdot)$ 為 H_0 下之累積危險函數 (cumulative hazard function) , $\Delta\Lambda(t) = \Lambda(t) - \Lambda(t-)$ 。

證明：

在 Fleming 和 Harrington (1991) 之定理 3.2.3 將 Kaplan-Meier 估計量 $\hat{S}_i(t)$ 表示為

$$S(t) - S(t) \int_0^t \frac{\hat{S}_i(v)}{S(v)} I_{[Y_i(v) > 0]} dM_i(v) + B_i(t) ,$$

其中

$$B_i(t) = I_{[T_c < t]} \frac{\hat{S}_i(T_c) \{S(T_c) - S(t)\}}{S(T_c)} ,$$

同時 $M_i(\cdot)$ 為獨立且均數為 0 之鞅 (martingale)。因此 , (4.1) 式可寫為

$$WKM_p = \sum_{j=1}^k \int_0^{T_c} H_j(t) dM_j(t) - \sum_{j=1}^k \int_0^{T_c} V_j(t) B_j(t) dt \quad (4.4)$$

其中

$$H_j(t) = \begin{cases} \sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \sqrt{\frac{n_i n_j}{N}} \text{sgn}(j-i) \left\{ \int_t^{T_c} \hat{w}_{ij}(u) S(u) du \right\} \frac{\hat{S}_j(t)}{S(t)} \frac{I_{[Y_j(t) > 0]}}{Y_j(t)} & \text{if } 1 \leq j \leq p-1 ; \\ \sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \sqrt{\frac{n_i n_p}{N}} \left\{ \int_t^t \hat{w}_{ip}(u) S(u) du \right\} \frac{\hat{S}_p(t)}{S(t)} \frac{I_{[Y_p(t) > 0]}}{Y_p(t)} & \text{if } j = p ; \\ \sum_{\substack{i \neq j \\ p \leq i \leq k}} \sqrt{\frac{n_i n_j}{N}} \text{sgn}(i-j) \left\{ \int_t^t \hat{w}_{ji}(u) S(u) du \right\} \frac{\hat{S}_j(t)}{S(t)} \frac{I_{[Y_j(t) > 0]}}{Y_j(t)} & \text{if } p+1 \leq j \leq k. \end{cases} \quad (4.5)$$

且

$$V_j(t) = \begin{cases} \sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \sqrt{\frac{n_i n_j}{N}} \text{sgn}(j-i) \hat{w}_{ij}(t) & \text{if } 1 \leq j \leq p-1 ; \\ \sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \sqrt{\frac{n_i n_p}{N}} \hat{w}_{ip}(u) & \text{if } j = p ; \\ \sum_{\substack{i \neq j \\ p \leq i \leq k}} \sqrt{\frac{n_i n_j}{N}} \text{sgn}(i-j) \hat{w}_{ji}(t) & \text{if } p+1 \leq j \leq k. \end{cases}$$

因為(4.4)式中第二項收斂至 0，且 $H_j^2(t)Y_j(t) \xrightarrow{p} h_j(t)$ 在 t 均勻收斂對所有 $j=1, \dots, k$ 。所以(4.3)式之 $h_j(t)$ 滿足 Fleming 和 Harrington (1991)之定理 6.2.1 之條件，根據鞅中央極限定理得證此定理。

令 $\hat{S}(t)$ 為合併 k 組資料之 Kaplan-Meier 存活函數估計。因為 Kaplan-Meier 存活函數估計量具有一致性，並且 $p_j(t)$ 之一致估計量為 $\hat{p}_j(t) = \hat{S}(t)\hat{C}_j(t-)$ ， $d\Lambda(t)$ 和 $1 - \Delta\Lambda(t)$ 之一致估計量為 $d\hat{\Lambda}(t) = -d\hat{S}(t) / \hat{S}(t-)$ 和 $1 - \Delta\hat{\Lambda}(t) = \hat{S}(t) / \hat{S}(t-)$ 。所以，估計量

$$\hat{Var}(WKM_p) = \sum_{j=1}^k \int_0^{T_c} \hat{h}_j(t) \frac{-\hat{S}(t)d\hat{S}(t)}{\hat{S}^2(t-)} \quad (4.6)$$

為 s_p^2 之一致估計量，其中

$$\hat{h}_j(t) = \begin{cases} \left[\sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \sqrt{\hat{p}_i} \operatorname{sgn}(j-i) \left\{ \int_t^t \hat{w}_{ij}(u) \hat{S}(u) du \right\} \right]^2 / [\hat{S}(t)\hat{C}_j(t-)] & \text{if } 1 \leq j \leq p-1; \\ \left[\sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \sqrt{\hat{p}_i} \left\{ \int_t^t \hat{w}_{ip}(u) \hat{S}(u) du \right\} \right]^2 / [\hat{S}(t)\hat{C}_p(t-)] & \text{if } j = p; \\ \left[\sum_{\substack{i \neq j \\ p \leq i \leq k}} \sqrt{\hat{p}_i} \operatorname{sgn}(i-j) \left\{ \int_t^t \hat{w}_{ji}(u) \hat{S}(u) du \right\} \right]^2 / [\hat{S}(t)\hat{C}_j(t-)] & \text{if } p+1 \leq j \leq k. \end{cases} \quad (4.7)$$

根據 Slutsky's 定理得到在虛無假說之下統計量

$$WKM_p^* = \frac{WKM_p}{\sqrt{\hat{Var}(WKM_p)}} \quad (4.8)$$

之極限分布為標準常態分配。最後獲得傘頂已知的檢定為

$$\text{棄卻 } H_0 \text{ 若 } WKM_p^* \geq z(\alpha),$$

其中 $z(\alpha)$ 為標準常態分布之 $100(1-\alpha)$ 百分位。

4.3 傘頂未知之傘檢定

當傘頂未知時，參考 Chen-Wolfe (1990)做法，將傘形對立假說視為 k 個傘頂分別為 $1, \dots, k$ 之傘型對立假說之聯集，得到傘頂未知之傘檢定統計量為

$$WKM_{\max}^* = \max\{WKM_1^*, \dots, WKM_k^*\}$$

其中 WKM_i^* 列於(4.8)式, $i = 1, \dots, k$, 。

若給定任意常數 a_1, \dots, a_k , 我們得到

$$\sum_{p=1}^k a_p WKM_p = \sum_{i=1}^k \int_0^{T_c} \left\{ \sum_{p=1}^k a_p H_i^{(p)} \right\} dM_i$$

其中 $H_i^{(p)}$ 同(4.5)式。利用鞅中央極限定理及 Cramer-Wold device 得知, 在虛無假說下

$$(WKM_1^*, \dots, WKM_k^*) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \mathbf{R}), \quad \text{as } N \rightarrow \infty ,$$

其中

$$\mathbf{R} = \left(\frac{\text{Cov}(WKM_p, WKM_q)}{\sqrt{\mathbf{s}_p^2 \mathbf{s}_q^2}} \right)_{p \times q} ,$$

且 \mathbf{s}_p^2 , $p = 1, \dots, k$, 同(4.2)式 ,

$$\text{Cov}(WKM_p, WKM_q) = \sum_{i=1}^k \int_0^t h_i^{(pq)} \{1 - \Delta\Lambda\} d\Lambda ,$$

其中 $h_i^{(pq)}$ 為 $H_i^{(p)} H_i^{(q)} Y_i$, $p \neq q = 1, \dots, k$, 之極限。因為, 對任意 $p \neq q = 1, \dots, k$ $\text{Cov}(WKM_p, WKM_q)$ 之一致估計量為

$$\hat{\text{Cov}}(WKM_p, WKM_q) = \sum_{i=1}^k \int_0^{T_c} \hat{h}_i^{(pq)}(t) \frac{-\hat{S}(t) d\hat{S}(t)}{\hat{S}^2(t-)} ,$$

所以, \mathbf{R} 之一致估計量為

$$\hat{\mathbf{R}} = \left(\frac{\hat{\text{Cov}}(WKM_p, WKM_q)}{\sqrt{\hat{\text{Var}}(WKM_p) \hat{\text{Var}}(WKM_q)}} \right)_{p \times q} ,$$

其中 $\hat{\text{Var}}(WKM_p)$ 同(4.6)式。

令 (Z_1, \dots, Z_k) 為一隨機向量且服從均數為 0 向量, 相關矩陣為 $\hat{\mathbf{R}}$ 之 k 維常態分布。並令 $z_{\max}(k, \mathbf{a})$ 為 $\max(Z_1, \dots, Z_k)$ 分布之第 \mathbf{a} 上百分位。所以, 得到近似水準為 \mathbf{a} 之傘頂未知之傘檢定為

$$\text{棄卻 } H_0 \text{ 若 } WKM_{\max}^* \geq z_{\max}(k, \mathbf{a})。$$

同時, 臨界值 $z_{\max}(k, \mathbf{a})$ 值當 $k \leq 7$ 時, 可利用 Schervish (1984) 的程式求得。

此外，在相同設限分布假設下， $Cov(WKM_p, WKM_q)$ 對 $p < q$ 時可簡化為

$$\begin{aligned}
Cov(WKM_p, WKM_q) = & \left\{ \sum_{j=1}^{p-1} \left(\sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(j-i) \right) \left(\sum_{\substack{i \neq j \\ 1 \leq i \leq q}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(j-i) \right) \right. \\
& + \left(\sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_p} \right) \left(\sum_{\substack{i \neq p \\ 1 \leq i \leq q}} \frac{\sqrt{p_i}}{p_i + p_p} \operatorname{sgn}(p-i) \right) \\
& + \sum_{j=p+1}^{q-1} \left(\sum_{\substack{i \neq j \\ p \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(i-j) \right) \left(\sum_{\substack{i \neq j \\ 1 \leq i \leq q}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(j-i) \right) \\
& + \left(\sum_{\substack{i \neq q \\ 1 \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_q} \right) \left(\sum_{\substack{i \neq q \\ 1 \leq i \leq q}} \frac{\sqrt{p_i}}{p_i + p_q} \operatorname{sgn}(i-q) \right) \\
& + \sum_{j=q+1}^k \left(\sum_{\substack{i \neq j \\ p \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(i-j) \right) \left(\sum_{\substack{i \neq j \\ q \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(i-j) \right) \left. \right\} \\
& \int_0^t \left[\int_t^t C(u)S(u)du \right]^2 \frac{-dS(t)}{C(t-)S(t-)S(t)} \circ
\end{aligned}$$

同樣的， s_p^2 ， $p=1,..k$ 同(4.2)式亦可簡化成

$$\begin{aligned}
s_p^2 = & \left\{ \sum_{j=1}^{p-1} \left(\sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(j-i) \right)^2 + \sum_{j=p+1}^k \left(\sum_{\substack{i \neq j \\ p \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_j} \operatorname{sgn}(i-j) \right)^2 + \left(\sum_{\substack{i \neq p \\ 1 \leq i \leq k}} \frac{\sqrt{p_i}}{p_i + p_p} \right)^2 \right\} \\
& \int_0^t \left[\int_t^t C(u)S(u)du \right]^2 \frac{-dS(t)}{C(t-)S(t-)S(t)} \circ
\end{aligned}$$

因此，便可導出在相同樣本數及相同設限分布下之相關矩陣，並可計算(當 $k \leq 7$) 或模擬(當 $k > 7$) k 維常態分布在已知相關矩陣為 \mathbf{R} 下之臨界值 $z_{\max}(k, \mathbf{a})$ 。同時，可得知 WKM_1 及 WKM_k 之相關係數為 -1，所以，臨界值 $z_{\max}(k, \mathbf{a})$ 值亦可視為來自 $\max(|Z_1|, Z_2, \dots, Z_{k-1})$ 分布之第 \mathbf{a} 上百分位。臨界值 $z_{\max}(k, \mathbf{a})$ 在相同設限分布相同樣本數， $k=2, \dots, 10$ 和 $\mathbf{a}=0.01, 0.05$ 和 0.1 分別列於表 4.1，並利用此臨界值在模擬研究上。此外，在資料分析時，可計算出相關矩陣之估計值 $\hat{\mathbf{R}}$ ，並計算或模擬 k 維常態分布在已知相關矩陣為 $\hat{\mathbf{R}}$ 下之臨界值 $z_{\max}(k, \mathbf{a})$ 。

表 4-1 在相同設限分布和相同樣本數下 $z_{\max}(k, \mathbf{a})$ 值

k									
\mathbf{a}	2	3	4	5	6	7	8	9	10
0.01	2.58	2.72	2.80	2.86	2.90	2.93	2.99	3.02	3.05
0.05	1.96	2.12	2.22	2.28	2.33	2.36	2.37	2.38	2.39
0.10	1.65	1.82	1.92	1.99	2.03	2.07	2.08	2.09	2.10

4.4 傘頂已知之漸進相對效率

在連續序列對立假說下，令

$$\sqrt{N}\{S_j^N(t) - S_i^N(t)\} \rightarrow D_{ij}(t)$$

在 $[0, t)$ 均勻收斂，對某些有界函數 $D_{ij}(\cdot)$ ， $i \neq j = 1, \dots, k$ 。則 WKM_p 之期望值為

$$\Delta = \sum_{1 \leq i < j \leq p} \sqrt{p_i p_j} \int_0^t w_{ij}(t) D_{ij}(t) dt + \sum_{p \leq i < j \leq k} \sqrt{p_i p_j} \int_0^t w_{ji}(t) D_{ji}(t) dt。$$

並且根據鞅中央極限定理(Martingale Central Limit Theorem)得到 WKM_p 之極限分布為 $N(\Delta, \mathbf{s}_p^2)$ ，其中 \mathbf{s}_p^2 為(4.2)式。也就是說，利用 WKM_p 統計量所建構之檢定當 Δ 值變大時檢定亦隨之增加。因此， WKM_p 之 Pitman 效力(efficacy)為

$$\begin{aligned} \text{eff}(WKM_p) = & \left[\sum_{1 \leq i < j \leq p} \sqrt{p_i p_j} \int_0^t w_{ij}(t) D_{ij}(t) dt \right. \\ & \left. + \sum_{p \leq i < j \leq k} \sqrt{p_i p_j} \int_0^t w_{ji}(t) D_{ji}(t) dt \right]^2 / \mathbf{s}_p^2。 \end{aligned} \quad (4.9)$$

為了計算(4.9)式之 Pitman 效力，我們考慮 Lehmann 對立假說：

$$H_0 : (S_i = S \text{ for } i = 1, \dots, k)$$

相對於

$H_1 : (S_i^N = S^{1-q_i/\sqrt{N}}$ and $\mathbf{q}_1 \leq \dots \leq \mathbf{q}_p \geq \dots \geq \mathbf{q}_k$ for some $p, 1 \leq p \leq k$, with at least one strict inequality)。

因為在 Lehmann 對立假說下，我們得到 $D_{ij}(t) \rightarrow (\mathbf{q}_j - \mathbf{q}_i)S(t)\ln(S(t))$ 對 $i, j = 1, \dots, k$ 。

因此，

$$\begin{aligned}
\text{eff}(WKM_p) = & \left[\sum_{1 \leq i < j \leq p} \sum_{p \leq i < j \leq k} \sqrt{p_i p_j} \int_0^t w_{ij}(t) (\mathbf{q}_i - \mathbf{q}_j) S(t) \ln(S(t)) dt \right. \\
& \left. + \sum_{p \leq i < j \leq k} \sqrt{p_i p_j} \int_0^t w_{ji}(t) (\mathbf{q}_j - \mathbf{q}_i) S(t) \ln(S(t)) dt \right]^2 / \mathbf{s}_p^2.
\end{aligned} \tag{4.10}$$

所以，在相同設限分布及樣本數假設下，加權 Kaplan-Meier 傘檢定之效力為

$$\begin{aligned}
\text{eff}(WKM_p) = & \left[\sum_{1 \leq i < j \leq p} (\mathbf{q}_i - \mathbf{q}_j) + \sum_{p \leq i < j \leq k} (\mathbf{q}_j - \mathbf{q}_i) \right]^2 \left[\int_0^t C(t) S(t) \ln(S(t)) dt \right]^2 \\
& / \left\{ k \left[\sum_{j=1}^{p-1} \left(\sum_{\substack{i \neq j \\ 1 \leq i \leq p}} \text{sgn}(j-i) \right)^2 + (k-1)^2 + \left(\sum_{j=p+1}^k \sum_{\substack{i \neq j \\ p \leq i \leq k}} \text{sgn}(i-j) \right)^2 \right] \right. \\
& \left. \int_0^t \left(\int_i^t C(u) S(u) du \right)^2 \frac{-dS(t)}{C(t) S^2(t)} \right\}.
\end{aligned}$$

而加權對數秩傘檢定之效力，則可參考 Chen 和 Wolfe (2000) 的第三節。

若 \mathbf{q}_i 為常數，則對應之 Lehmann 對立假說為比例危險模式。所以，當 Lehmann 對立假說之 $S(t) = \exp(-t)$ 和隨時間改變之 \mathbf{q}_i ，可得到不同時段危險函數成段指數 (piecewise exponential) 存活函數。例如，危險函數為

$$I_{i1} I_1(t) + I_{i2} \{1 - I_1(t)\},$$

其中 $I_1(t) = 1$ 若 $t < t_1$ ；或 0，其他。則對應之 \mathbf{q}_i 為

$$I_{i1} I_1(t) + \{I_{i2} + (I_{i1} - I_{i2}) t_1 / t\} \{1 - I_1(t)\}.$$

在這個例子中，若危險函數中之 I_{i2} (I_{i1}) 相同而 I_{i1} (I_{i2}) 不同，則得到早 (晚) 期差異之危險函數。相同的，若危險函數為

$$I_{i1} I_1(t) + I_{i2} I_2(t) + I_{i3} I_3(t) + I_{i4} \{1 - I_1(t) - I_2(t) - I_3(t)\},$$

其中 $I_j(t) = 1$ 若 $t_{j-1} \leq t < t_j$ ；或 0，其他， $t_0 = -\infty$ ， $j = 1, 2, 3$ 。則對應之 \mathbf{q}_i 為

$$\begin{aligned}
& I_{i1} I_1(t) + \{I_{i2} + (I_{i1} - I_{i2}) t_1 / t\} I_2(t) + \{I_{i3} + (I_{i1} - I_{i2}) t_1 / t + (I_{i2} - I_{i3}) t_2 / t\} I_3(t) \\
& + \{I_{i4} + (I_{i1} - I_{i2}) t_1 / t + (I_{i2} - I_{i3}) t_2 / t + (I_{i3} - I_{i4}) t_4 / t\} \{1 - I_1(t) - I_2(t) - I_3(t)\}.
\end{aligned}$$

因此，在成段指數分布下，不同的 I_{ij} 值可產生危險函數交叉之情形。

為了計算加權 Kaplan-Meier 傘檢定 (傘頂已知) 相對於加權對數秩傘檢定 (傘頂已知) 之漸進相對效率 (asymptotic relative efficiency, ARE)，我們考慮在 $k = 5$ 且樣本數相同，傘頂 $p = k$ 的情形。並考慮存活函數為指數分布，尺度參數為 (1) $\mathbf{q}_i = i$ ，(2) $\mathbf{q}_i = 2^{(i-1)}$ 和 (3) $\mathbf{q}_i = 1$ ，若 $i \leq k-1$ 且 $\mathbf{q}_i = 2$ ，若 $i = k$ ，對所有

$i = 1, \dots, k$ 。另外，利用成段指數分布中改變 I_{ij} 值而得到不同危險函數差異情形。圖 4-1 為所考慮的成段指數分布。設限分布假設為均勻分布從 0 到 R，記為 $U(0,R)$ 。在加權對數傘檢定中，權數採用 Fleming 和 Harrington (1982) 所建議之權數形式 $\{S(t)\}^r \{1-S(t)\}^g$ ，其中 $r \geq 0, g \geq 0$ 。在此，我們考慮的加權對數秩檢定統計量有：對數秩(LR, $r = g = 0$)統計量，Peto-Prentice-Wilcoxon (PPW, $r = 1, g = 0$)統計量或針對危險函數差異發生在晚期之統計量 WLRL ($r = 0, g = 1$)。ARE(WKM,WLR)值列於表 4-2。

從表 4-2 我們得知：在比例危險模式下，WKM 檢定與 LR 檢定之效率相當。在危險函數早期發生差異時，WKM 檢定與 PPW 檢定之效率相當。若危險函數差異發生在晚期時，WKM 檢定效率低於 LR 或 WLRL 檢定之效率，這是因為 WKM 檢定之權數隨時間遞減。但是，在危險函數交叉時，WKM 檢定效率卻是優於所討論之 WLR 檢定之效率。

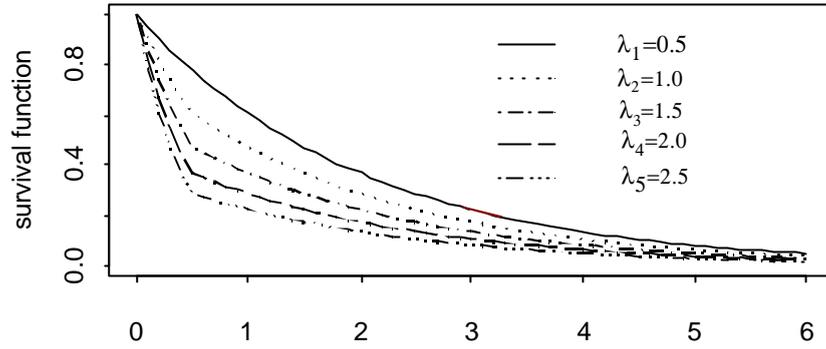
表 4-2 在 $k = 5, p = 5$ ，相同樣本數和共變數為 (q_1, \dots, q_5) 且設限分布為 $U(0,R)$ 之 ARE(WKM,WLR)值

q_i	R	WKM/LR	WKM/PPW	WKM/WLRL
(a) Standard exponential distribution				
(1,2,3,4,5)	3.185	1.06	1.50	1.25
	9.901	1.08	1.48	1.39
(1,1,1,1,2)	3.185	1.63	2.34	1.95
	9.901	1.69	2.31	2.17
(1,2,4,8,16)	3.185	1.31	1.85	1.54
	9.901	1.33	1.82	1.72
(b) Piecewise exponential distribution ¹				
(I) Early hazard difference		1.31	0.97	27.4
(II) Late hazard difference		0.70	1.86	0.21
(III) Crossing hazard difference		1.22	1.03	4.97

1: R=2 for piecewise exponential distribution

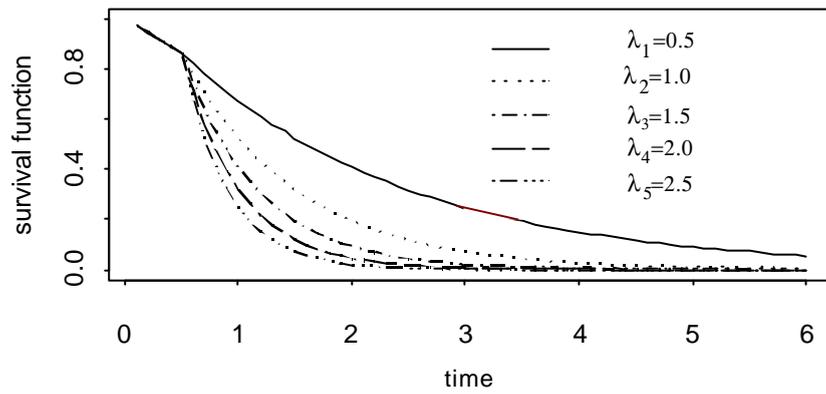
(I) Early hazard difference

$$\lambda(t) = \lambda_i | [t < 0.5] + 0.5 | [t > 0.5]$$



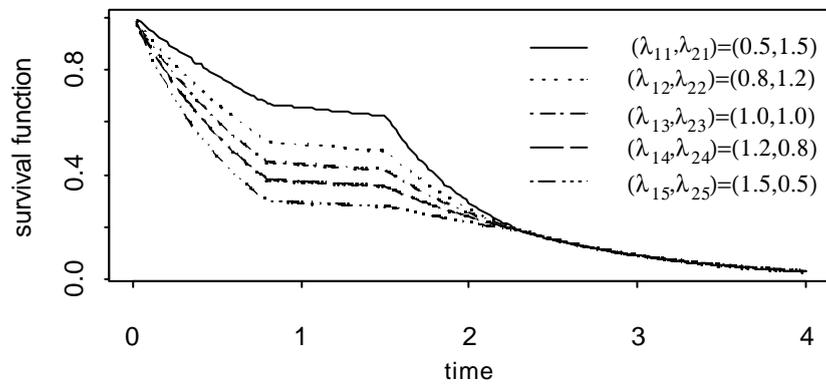
(II) Late hazard difference

$$\lambda(t) = 0.3 | [t < 0.5] + \lambda_i | [t > 0.5]$$



(III) Crossing hazard difference

$$\lambda(t) = \lambda_{1i} | [t < 0.8] + | [0.8 < t < 1.5] + \lambda_{2i} | [1.5 < t < 2.3] + | [t > 2.3]$$



■ 4-1. Survival configuration for alternatives.

4.5 模擬研究

本節採用模擬研究探討不同傘檢定之水準及檢定力表現。考慮組別數 $k = 4$ 和 5，且在水準研究時，假設各組樣本數 $n_1 = \dots = n_k = 10, 20$ 或 30，在檢定力研究時，假設各組樣本數均為 20 或 30。並假設存活函數來自指數分布或成段指數分布，設限時間服從均勻分布。在水準研究中，考慮標準指數分布及一些特定之成段指數分布。在檢定力研究中，改變指數分布之尺度 (scale) 參數 ($I_i, i = 1, \dots, k$) 得到不同型態之傘形對立假說，同樣的，利用改變成段指數分布之參數，建構危險函數發生早期差異或交叉之對立假說。此外，在水準研究中，設限分布服從均勻分布 $U(0, R)$ ，其設限比率為 0.1 和 0.3 其相對應之 R 值別為 9.901 和 3.185，在檢定力研究時亦考慮相同之設限分布。在成段指數分布考慮設限分布均為 $U(0, 2)$ ，因此，在水準研究中案例(1)-(3)之設限比率分別為 0.51, 0.56 及 0.51。但是，在檢定力研究時，因其對立假說之存活函數不同，其所相對應之設限比率亦不相同。

在這樣的配置下，使用名目水準 $\alpha = 0.05$ ，重覆 5000 次得到水準及檢定力之估計。從模擬結果得知 $k=4$ 和 $k=5$ 的結果相似，因此列出 $k=4$ 的結果。水準估計之結果列於表 4-3，檢定力估計之結果列於表 4-4，表 4-5 則提供在傘頂已知之傘檢定中，若選錯或選對傘頂時，檢定力之表現。

從表 4-3 中可知：在傘頂已知之傘檢定， WKM 檢定及 WLR 檢定水準的表現均在合理誤差內，而在傘頂未知之傘檢定 WKM 檢定水準的表現較 WLR 檢定佳。也就是說， WLR 檢定在傘頂未知傘檢定之水準維持較保守。表 4-4 中檢定力的表現與表 4-1 漸進相對效率的情形相同。此外，若選擇正確的傘頂位置，不論 WKM 檢定或 WLR 檢定均有較高的檢定力，此結果在表 4-5 中，若選擇錯誤的傘頂，其檢定之檢定力比使用傘頂未知傘檢定之檢定力更低。因此選擇正確的傘頂亦是重要的問題，若無明顯證據證明傘頂位置時，應使用傘頂未知傘檢定會較適當。

表 4-3 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之水準估計

		Peak-known				Peak-unknown			
n	R	WKM	LR	WLRL	PPW	WKM	LR	WLRL	PPW
(a) Standard exponential distribution									
10	9.901	0.055	0.045	0.044	0.044	0.058	0.044	0.050	0.039
	3.185	0.052	0.043	0.046	0.045	0.051	0.045	0.050	0.037
20	9.901	0.056	0.054	0.051	0.052	0.058	0.046	0.052	0.041
	3.185	0.050	0.049	0.045	0.047	0.049	0.044	0.049	0.037
30	9.901	0.051	0.048	0.049	0.044	0.053	0.043	0.044	0.039
	3.185	0.048	0.048	0.052	0.050	0.049	0.038	0.045	0.037
(b) Piecewise exponential distribution									
(1) $I(t) = I_{[t \leq 0.5]} + 0.5I_{[t > 0.5]}$									
10	2	0.058	0.040	0.038	0.040	0.051	0.038	0.040	0.034
20	2	0.052	0.045	0.043	0.047	0.050	0.040	0.040	0.036
30	2	0.053	0.046	0.043	0.050	0.042	0.045	0.040	0.044
(2) $I(t) = 0.3I_{[t \leq 0.5]} + I_{[t > 0.5]}$									
10	2	0.050	0.050	0.044	0.050	0.038	0.040	0.046	0.035
20	2	0.043	0.042	0.040	0.042	0.042	0.036	0.042	0.037
30	2	0.052	0.049	0.046	0.046	0.048	0.037	0.046	0.037
(3) $I(t) = I_{[t \leq 0.8]} + 0.1I_{[0.8 < t \leq 1.5]} + I_{[1.5 < t \leq 2.3]} + I_{[2.3 > t]}$									
10	2	0.051	0.048	0.047	0.045	0.050	0.041	0.038	0.040
20	2	0.045	0.045	0.044	0.044	0.042	0.039	0.038	0.037
30	2	0.042	0.044	0.043	0.048	0.046	0.039	0.040	0.037

表 4-4 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之檢定力估計

Parameters						Peak-known				Peak-unknown				
I_1	I_1	I_1	I_1	R	N	WKM	LR	WLRL	PPW	WKM	LR	WLRL	PPW	
(a) Exponential distribution														
1	1	1.5	2	9.901	20	0.631	0.619	0.509	0.544	0.420	0.386	0.292	0.300	
					30	0.795	0.774	0.673	0.673	0.627	0.561	0.438	0.447	
					3.185	20	0.580	0.507	0.392	0.458	0.355	0.284	0.205	0.242
					30	0.727	0.653	0.521	0.599	0.517	0.409	0.288	0.367	
1	1.5	2	1	9.901	20	0.607	0.741	0.653	0.647	0.472	0.575	0.491	0.476	
					30	0.787	0.882	0.792	0.806	0.671	0.765	0.649	0.654	
					3.185	20	0.560	0.627	0.527	0.569	0.405	0.453	0.730	0.400
					30	0.720	0.782	0.692	0.728	0.572	0.639	0.528	0.569	
(b) Early hazard difference														
$I(t) = I_t I_{[t \leq 0.5]} + 0.5 I_{[t > 0.5]}$														
2	1.5	1	0.5	2	20	0.798	0.686	0.370	0.752	0.599	0.451	0.197	0.506	
					2	30	0.922	0.838	0.501	0.887	0.788	0.645	0.290	0.724
2	1.5	0.5	1	2	20	0.698	0.598	0.351	0.640	0.591	0.496	0.236	0.557	
					2	30	0.842	0.759	0.448	0.796	0.783	0.682	0.332	0.742
(c) Late hazard difference														
$I(t) = 0.3 I_{[t \leq 0.5]} + I_t I_{[t > 0.5]}$														
2	1.5	1	0.3	2	20	0.600	0.713	0.749	0.564	0.379	0.481	0.512	0.317	
					2	30	0.769	0.867	0.904	0.721	0.562	0.687	0.736	0.489
2	1.5	0.3	1	2	20	0.541	0.683	0.777	0.530	0.396	0.532	0.610	0.381	
					2	30	0.693	0.828	0.914	0.671	0.579	0.722	0.808	0.547
(d) Crossing hazard difference														
$I(t) = I_{i1} I_{[t \leq 0.8]} + 0.1 I_{[0.8 < t \leq 1.5]} + I_{i2} I_{[1.5 < t \leq 2.3]} + I_{[2.3 > t]}$														
1.5	1	0.8	0.5 ¹	2	20	0.712	0.634	0.403	0.659	0.496	0.390	0.213	0.414	
0.5	1	1.2	1.5	2	30	0.865	0.798	0.534	0.828	0.702	0.585	0.322	0.617	
1.5	0.8	0.5	1 ¹	2	20	0.702	0.653	0.456	0.663	0.525	0.493	0.320	0.503	
0.5	1.2	1.5	1	2	30	0.832	0.796	0.570	0.809	0.702	0.656	0.417	0.671	

1: the first row is I_{i1} and the second row is I_{i2} .

表 4-5 當 $k = 4$, 名目水準 $\alpha = 0.05$, 設限分布為 $U(0, R)$ 且 $n_1 = \dots = n_4 = n$ 之檢定力估計

Parameters						$p=2$		$p=3$		Peak-unknown		
I_1	I_1	I_1	I_1	R	n	WKM	LR	WKM	LR	WKM	LR	
(a) Exponential distribution												
1	1.5	2	1	9.901	20	0.315	0.426	0.607	0.741	0.472	0.575	
					30	0.418	0.573	0.787	0.882	0.671	0.765	
					3.185	20	0.293	0.377	0.560	0.627	0.405	0.453
					30	0.393	0.512	0.720	0.782	0.572	0.639	
1	1.5	2	1.5	9.901	20	0.104	0.160	0.461	0.529	0.355	0.421	
					30	0.115	0.195	0.627	0.677	0.529	0.593	
					3.185	20	0.102	0.148	0.419	0.424	0.367	0.328
					30	0.111	0.179	0.562	0.579	0.449	0.469	
(b) Early hazard difference												
$I(t) = I_i I_{[t \leq 0.5]} + 0.5 I_{[t > 0.5]}$												
2	1	0.5	1.5	2	20	0.290	0.382	0.798	0.752	0.632	0.601	
					2	30	0.367	0.498	0.920	0.805	0.765	
1.5	1	0.5	2	2	20	0.382	0.436	0.734	0.690	0.567	0.531	
					2	30	0.485	0.574	0.856	0.853	0.735	0.730
(c) Late hazard difference												
$I(t) = 0.3 I_{[t \leq 0.5]} + I_i I_{[t > 0.5]}$												
2	1	0.3	1.5	2	20	0.176	0.323	0.604	0.803	0.428	0.628	
					2	30	0.226	0.418	0.784	0.927	0.613	0.822
1.5	1	0.3	2	2	20	0.225	0.347	0.558	0.757	0.383	0.579	
					2	30	0.258	0.456	0.717	0.892	0.552	0.761
(d) Crossing hazard difference												
$I(t) = I_{i1} I_{[t \leq 0.8]} + 0.1 I_{[0.8 < t \leq 1.5]} + I_{i2} I_{[1.5 < t \leq 2.3]} + I_{[2.3 > t]}$												
1.5	1	0.8	0.5 ¹	2	20	0.053	0.090	0.712	0.634	0.496	0.390	
0.5	1	1.2	1.5	2	30	0.048	0.103	0.865	0.798	0.702	0.585	
1.5	0.8	0.5	1 ¹	2	20	0.218	0.287	0.702	0.653	0.525	0.493	
0.5	1.2	1.5	1	2	30	0.288	0.368	0.832	0.796	0.702	0.656	

1: the first row is I_{i1} and the second row is I_{i2}

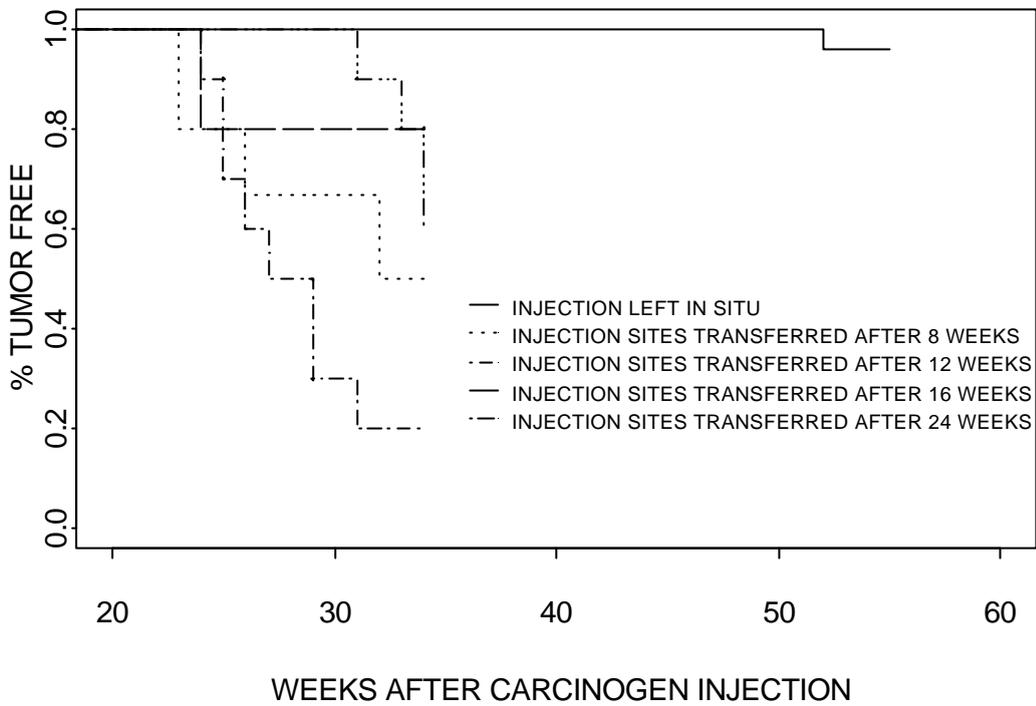
4.6 實例分析

在此考慮二個例子。首先，考慮 Chen 和 Wolfe (2000)所引用 Homburger 和 Treger (1970)之例子。本例主要研究在動物中注射致癌物質 (benz(**a**)-anthracene, BA, 500 **mg**) 後，將注射致癌物質之動物細胞在不同時間轉移至其他健康活體上，看是否會形成腫瘤。40 隻 C57BL/6J 的母鼠分別被植入注射致癌物質之動物細胞在注射後之 8 , 12 , 16 和 24 週後。對照組為 50 隻注射致癌物質 (benz(**a**)-anthracene, BA, 500 **mg**) 但不移植之母鼠。記錄從注射或植入後到研究人員觸摸到腫瘤之時間，此外，會有一些直到研究結束並未發生腫瘤或因其他因素死亡，這些與腫瘤無關的因素記為設限資料。這五組資料之 Kaplan-Meier 存活函數估計列於圖 4-2，傘形對立假說之傘檢定結果列於表 4-6。表 4-6 中列出

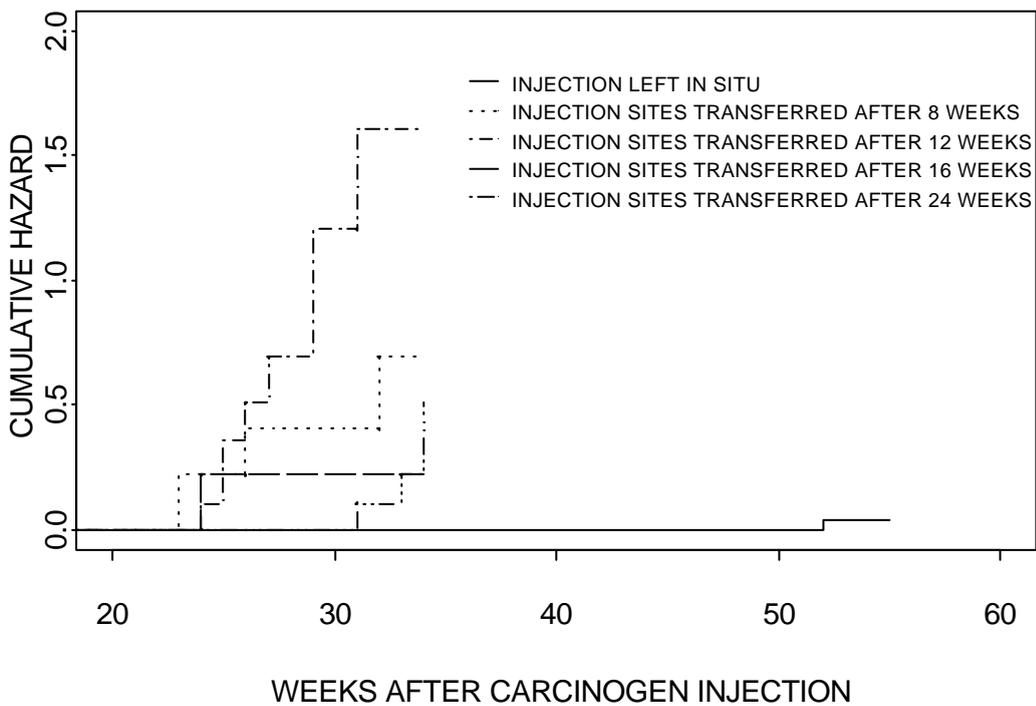
表 4-6 the injection sites-transfer data 之傘檢定統計量

Peak (p)	1	2	3	4	5
(a) Weighted Kaplan-Meier (WKM)					
WKM_p^*	1.016	2.995	3.824	1.324	-1.016
\hat{R}	1.000	0.636	0.114	-0.445	-1.000
		1.000	0.481	-0.060	-0.630
			1.00	0.419	-0.112
				1.000	0.472
$p - value = 3 \times 10^{-4}$					
(b) Logrank (LR) ¹					
A_p^*	-2.040	2.922	5.377	3.884	5.536
\hat{R}	1.000	0.300	0.152	-0.133	-0.754
		1.000	0.584	0.482	0.081
			1.000	0.681	0.327
				1.000	0.678
$p - value = 4 \times 10^{-7}$					
(c) Peto-Prentice-Wilcoxon (PPW) ¹					
A_p^*	-2.234	3.229	5.659	4.196	5.437
\hat{R}	1.000	0.285	0.115	-0.172	-0.803
		1.000	0.663	0.549	0.126
			1.000	0.753	0.384
				1.000	0.712
$p - value = 1 \times 10^{-7}$					

1: A_p^* 為 Chen and Wolfe (2000)所建議之加權對數秩傘檢定統計量



4-2 The Kaplan-Meier estimates for the injection sites-transfer data.



4-3 The cumulative hazard functions for the injection sites-transfer data.

各傘頂已知檢定統計量，其中粗黑字體者為傘頂未知檢定統計量，對應之 p 值則列於該檢定資料的最後一行。

從表 4-6 可得知，加權 Kaplan-Meier 傘頂未知之傘檢定與 Peto-Pentrice-Wilcoxon 傘未知傘檢定皆指出此五組呈傘形結構且傘頂位置於第三組，也就是注射 12 週移植。但是對數秩傘頂未知傘檢定所得到的結論卻是傘頂位置在第五組，也就是順序型態。同時，利用對數累積危險函數所繪製之圖 4-3 亦指出比例危險模式似乎並不恰當。因此，使用 WKM 和 PPW 傘檢定分析此筆資料似乎比較恰當。雖然 PPW 傘檢定之 p 值較 WKM 傘檢定之 p 值小，但是 WKM 傘檢定卻比 PPW 傘檢定更明確指出傘頂位置應該在第三組。所以由以上分析可知，注射致癌物質 (benz(a)-anthracene, BA, 500 μg) 12 週後移植比不移植或 8、16 週後移植較容易發生腫瘤。

其次，考慮的是第三章前列腺癌的例子。在此將分別對癌症三期和四期做傘檢定。圖 4-4(a)(b)分別為癌症三期和癌症四期之四組資料之存活函數估計及對數累積危險函數圖。表 4-7 和 4-8 則分別為癌症三期和癌症四期之傘形對立假說之傘檢定統計量之結果。由圖 4-4(a)可看出在癌症三期，第三組 (劑量為 1.0mg) 的反應值最高，且四組危險函數約成比例。由表 4-7 得知三種檢定統計量均指出傘頂位置應為第三組，而三種檢定統計量所得之 p 值亦相當，其結果與圖 4-4(a)相當吻合。也就是說，在前列腺癌症等級為三時使用雌性激素治療時，劑量值為 1.0mg (第三組) 時可得到最長的存活時間。圖 4-4(b)可看出在癌症四期，第三組 (劑量為 1.0mg) 的反應值和第四組 (劑量為 5.0mg) 的反應值相當，且四組之危險函數約略相交。表 4-8 中知三種檢定統計量均指出傘頂位置應為第四組，但 WKM 較 LR 及 PPW 之 p 值小，這是因為存活函數有交叉的情形，且在顯著水準若為 0.1 時，僅 WKM 有足夠證據拒絕四組存活函數相同之虛無假說。因此，可知在前列腺癌症等級為四時使用雌性激素治療時，劑量值為 5.0mg (第四組) 時可得到最長的存活時間。

表 4-7 the prostatic cancer data 之傘檢定統計量
the cancer stage = 3

Peak (p)	1	2	3	4
(a) Weighted Kaplan-Meier (WKM)				
WKM_p^*	-1.615	-0.629	1.984	1.614
\hat{R}	1.000	0.347	-0.363	-0.992
		1.000	0.280	-0.355
			1.00	0.372
<i>p-value</i> =0.0863				
(b) Logrank (LR) ¹				
A_p^*	-1.433	0.014	1.993	1.767
\hat{R}	1.000	0.507	-0.096	-0.914
		1.000	0.529	-0.027
			1.000	0.469
<i>p-value</i> =0.0844				
(c) Peto-Prentice-Wilcoxon (PPW) ¹				
A_p^*	-1.413	-0.079	1.952	1.724
\hat{R}	1.000	0.539	-0.067	-0.917
		1.000	0.537	-0.053
			1.000	0.455
<i>p-value</i> =0.0959				

1: A_p^* 為 Chen and Wolfe (2000)所建議之加權對數秩傘檢定統計量

圖 4-4 (a) 前列腺癌資料，癌症等級三之存活函數估計及對數累積危險函數圖

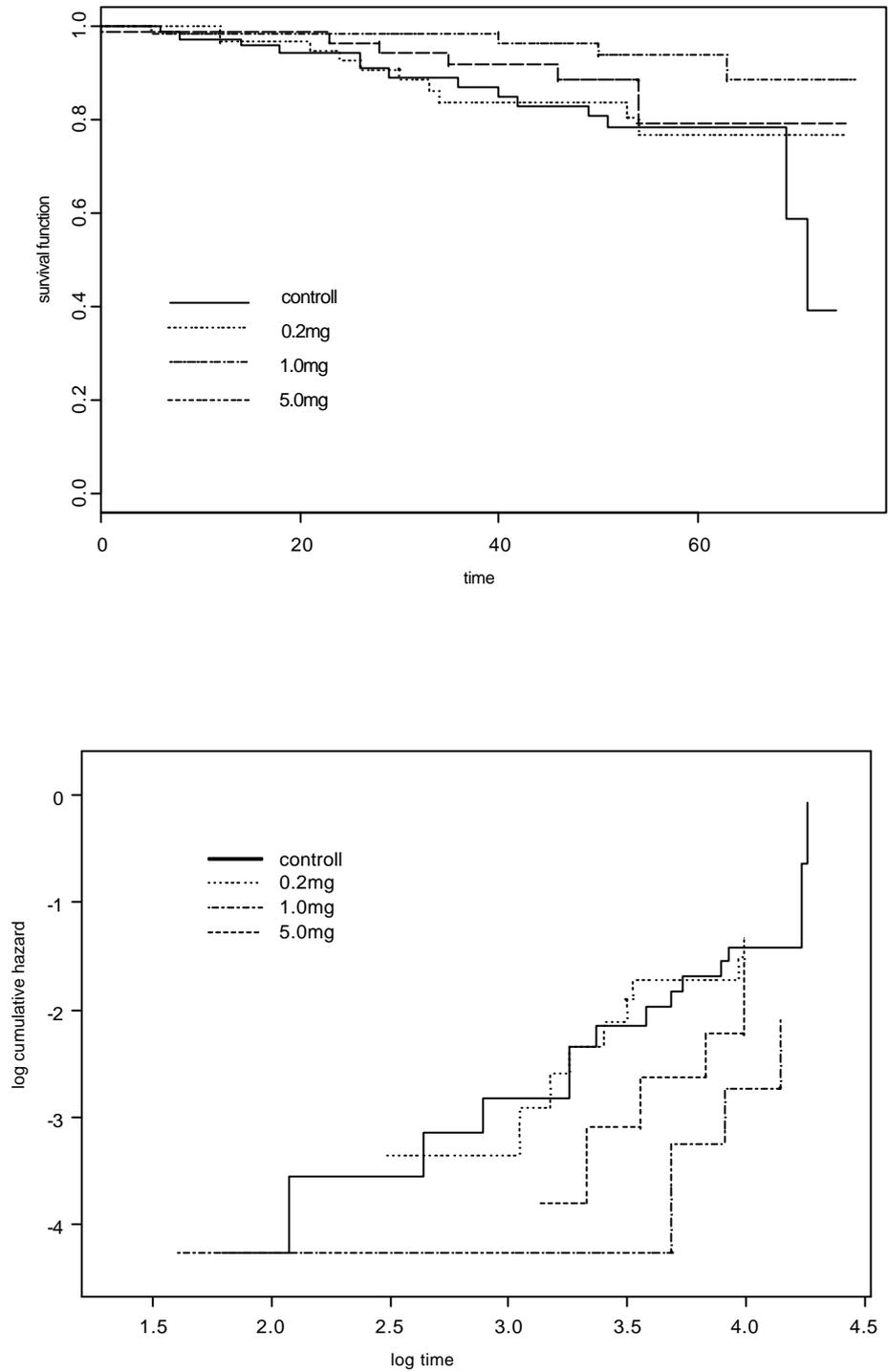
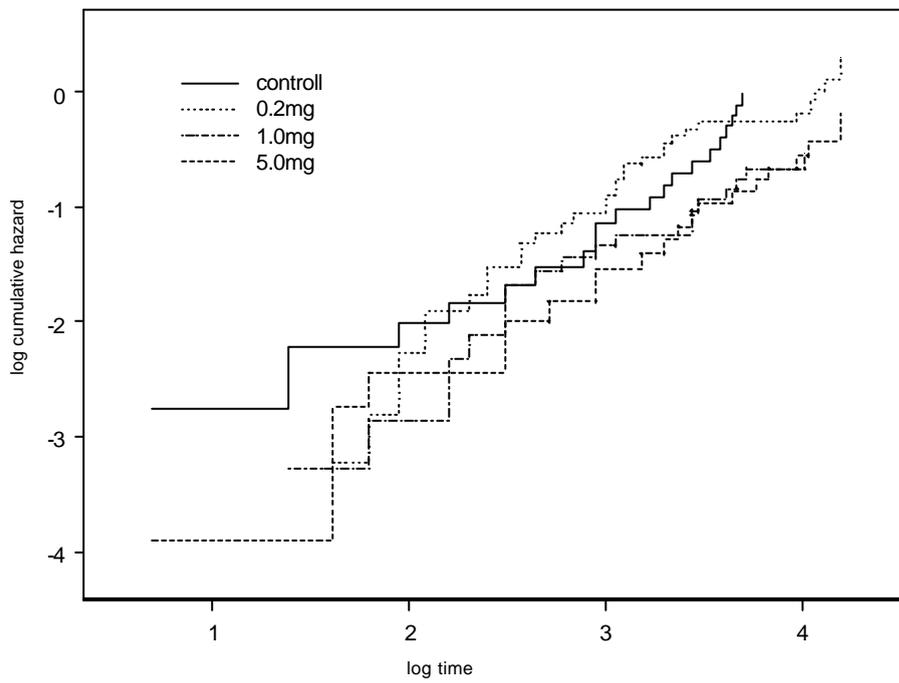
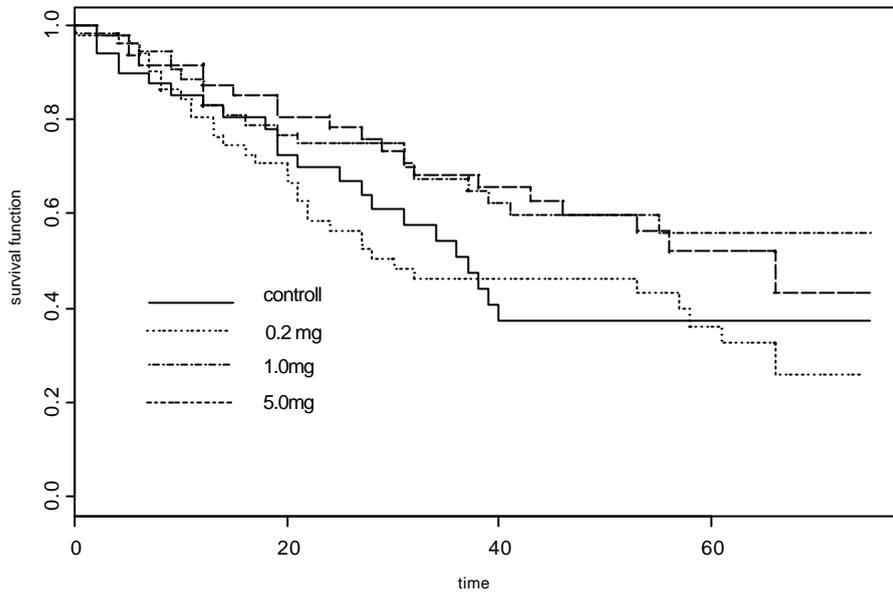


表 4-8 the prostatic cancer data 之傘檢定統計量
the cancer stage = 4

Peak (p)	1	2	3	4
(a) Weighted Kaplan-Meier (WKM)				
WKM_p^*	-2.139	-1.648	1.108	2.139
$\hat{\mathbf{R}}$	1.000	0.399	-0.372	-1.000
		1.000	0.259	-0.380
			1.000	0.351
<i>p-value</i> =0.0623				
(b) Logrank (LR) ¹				
A_p^*	-1.953	-1.428	0.996	1.729
$\hat{\mathbf{R}}$	1.000	0.554	-0.024	-0.889
		1.000	0.493	-0.126
			1.000	0.413
<i>p-value</i> =0.137				
(c) Peto-Prentice-Wilcoxon (PPW) ¹				
A_p^*	-1.977	-1.491	0.829	1.747
$\hat{\mathbf{R}}$	1.000	0.585	0.016	-0.893
		1.000	0.512	-0.153
			1.000	0.388
<i>p-value</i> =0.133				

1: A_p^* 為 Chen and Wolfe (2000)所建議之加權對數秩傘檢定統計量

圖 4-4 (b) 前列腺癌資料，癌症等級四之存活函數估計及對數累積危險函數圖



第五章 結論及未來研究方向

5.1 結論

本文主要是將雙樣本加權 Kaplan-Meier 統計量推廣至不同之檢定問題上。雖然 Pepe 和 Fleming(1989)指出當危險函數成比例時，WKM 單尾檢定與 LR 單尾檢定相當。此外，當危險函數早期差異時，WKM 單尾檢定之檢定力高於 LR 單尾檢定之檢定力，在本文第一章中利用模擬得知，當危險函數交叉時，WKM 單尾檢定仍優於 WLR 單尾檢定，但是危險函數晚期差異時，WKM 檢定之檢定力便劣於 LR 傘檢定。在第二章雙尾檢定問題中，所建議之 Renyi-type WKM 檢定並不比 Renyi-type LR 檢定更具檢定力，這是因為危險函數相交若在早期相交，之後不再相交時，Renyi 的方法並無效用，所得的統計量與不使用 Renyi 方法的統計量值一樣（不看正負符號）。這是因為計算方法是從前面向後累加，所得之最大值仍是在最後一個觀測值（仍有正負相抵的情形）。在第三章中可知共變數效應存在時，使用不調整共變數之加權 WKM 群序列檢定時，其檢定力比使用調整共變數後之 WMS 群序列檢定之檢定力小，因此，共變數的調整是必要的。在第四章中，由模擬及實例分析結果可知，WKM 傘檢定在比例危險模式下其檢定力與 LR 傘檢定相當，若是在危險函數早期差異，或交叉時，其檢定力均優於 WLR 傘檢定，但是在危險函數晚期差異時，WKM 傘檢定其檢定力便劣於 LR 傘檢定，其結論與第二章中模擬結果相符。同時，在傘頂不確定時，應使用傘頂未知之傘檢定，如此才不會因錯估傘頂而損失檢定力。

5.2 未來研究方向

在多樣本傘檢定問題中，若資料有共變數時，應可利用 Shen 和 Fleming(1997)的方法，建立調整共變數之 WMS 傘檢定。此外，在調整共變數之 WMS 群序列檢定中，若觀測資料之處理組不只二組時，可將雙樣本的檢定方法推廣至多樣本檢定問題上。上述問題均值得未來繼續研究。

參 考 文 獻

1. Andersen, P.K. and Gill, R.D. (1982). Cox's regression model for counting processes: a large sample study. *The Annals of Statistics* **10**, 1100-1120.
2. Andersen, P.K., Borgan, Ø., Gill, R. and Keiding, N. (1993). *Statistical Model Based on Counting Processes*, Springer-Verlag, New York.
3. Byar, D.P. (1985). Prognostic variables for survival in a randomized comparison of treatments for prostatic cancer. In *Data* (Edited by D.F. Andrews and A.M. Herzberg), 261-274, Springer-Verlag, New York.
4. Chen, Y.I. and Wolfe, D.A. (1990). A study of distribution-free tests for umbrella alternatives. *Biometrical Journal* **32**, 47-57.
5. Chen Y.I. (1991). Notes on the Mack-Wolfe and Chen-Wolfe tests for umbrella alternatives. *Biometrical Journal* **33**, 281-290.
6. Chen, Y.I. and Wolfe, D.A. (2000). Umbrella tests for right-censored survival data. *Statistica Sinica* **10**, 595-612.
7. Cox, D.R. (1972). Regression model and life tables (with Discussion). *Journal of the Royal Statistical Society, Series B* **34**, 187-220.
8. Cox, D.R. (1975). Partial likelihood. *Biometrika* **62**, 269-276.
9. Drezner, Z. (1992). Computation of the Multivariate normal integral. *AMC Transactions on Mathematical Software* **18**, 470-480.
10. Fleming, T.R. and Harrington, D. T. (1991). *Counting Processes and Survival Analysis*. New York; Wiley.
11. Gill, R.D. (1980). *Censoring and Stochastic Integrals, mathematical Centre Tracts* **124**, Amsterdam: Mathematisch Centrum.
12. Gu, M. and Ying, Z. (1995). Group sequential methods for survival data using partial likelihood score processes with covariate adjustment. *Statistica Sinica* **5**, 793-804.

13. Harrington, D.P. and Fleming, T.R. (1982). A class of rank test procedures for censored survival data. *Biometrika* **69**, 133-143.
14. Jonckheere, A.R. (1954). A distribution-free k-sample test against ordered alternatives. *Biometrika* **41**, 133-145.
15. Jennison, C. and Turnbull, B. (1997). Group-sequential analysis incorporating covariate information. *Journal of American Statistical Association* **92**, 1330-1341.
16. Kaplan, E.L. and Meier, P. (1958). Nonparametric estimation from incomplete observations. *Journal of American Statistical Association* **53**, 457-481.
17. Lan, K. K. G. and DeMets, D. L. (1983). Discrete sequential boundaries for clinical trials. *Biometrika* **70**, 659-663.
18. Li, Z. (1999). A group sequential test for survival trials: An alternative to rank-based procedures. *Biometrics* **55**, 277-283.
19. Lin, D.Y. and Wei, L.J. (1989). Robust inference for the Cox proportional hazards model. *Journal of American Statistical Association* **84**, 1074-1078.
20. Lin, D.Y. (1992). Sequential log rank tests adjusting for covariates with the accelerated life model. *Biometrika* **79**, 523-529.
21. Liu, P.Y., Green, S., Wolf, M., and Crowley, J. (1993). Testing against ordered alternatives for censored survival data. *Journal of American Statistical Association* **88**, 153-166.
22. Mack, G.A. and Wolfe, D.A. (1981). K-sample rank tests for umbrella alternatives. *Journal of American Statistical Association* **76**, 175-181.
23. O' Brien, P.C. and Fleming, T.R. (1979). A multiple testing procedure for clinical trials. *Biometrics* **35**, 549-556.
24. Pepe, M.S. and Fleming, T.R. (1989). Weighted Kaplan-Meier statistics: a class of distance tests for censored survival data. *Biometrics* **45**, 497-507.
25. Pepe, M.S. and Fleming, T.R. (1991). Weighted Kaplan-Meier statistics: Large sample

- and optimality considerations. *Journal of the Royal Statistical Society, Series B* **53**, 341-352.
26. Peto, R. and Peto, J. (1972). Asymptotic efficient rank invariant test procedures. *Journal of the Royal Statistical Society, Series A* **135**, 185-206.
 27. Pocock, S.J. (1977). Group sequential methods in the design and analysis of clinical trials. *Biometrika* **64**, 199-199.
 28. Prentice, R.L. (1978). Linear rank tests with right censored data. *Biometrika* **65**, 165-179.
 29. Schervish, M.J. (1984). Multivariate normal probabilities with error bound (with correction in 1985). *Applied Statistics* **33**, 81-94.
 30. Schumacher, M. (1984). Two-sample tests of the Cramer-von Mises and Kolmogorov-Smirnov type for randomly censored data. *International Statistical Review* **52**, 263-281.
 31. Shen, Y. and Fleming, T.R. (1997). Weighted mean survival test statistics – a class of distance tests for censored survival data. *Journal of the Royal Statistical Society, Series B* **59**, 269-280.
 32. Simpson, D.C. and Margolin, B.H. (1986). Recursive nonparametric testing for dose-response relationships subject to downturns at high dose. *Biometrika* **73**, 589-596.
 33. Slud, E.V. and Wei, L.J. (1982). Two-sample repeated significance tests based on the modified Wilcoxon statistic. *Journal of American Statistical Association* **77**, 862-868.
 34. Slud, E.V. (1984). Sequential linear rank tests for two-sample censored survival data. *Ann. Statist* **12**, 551-571.
 35. Terpstra, T.J. (1952). The asymptotic normality and consistency of Kendall's test against trend, when ties are present in one ranking. *Indagationes Math* **14**, 327-333.
 36. Tsiatis, A.A. (1982). Repeated significance testing for a general class of statistics used in censored survival analysis. *Journal of American Statistical Association* **77**, 855-863.
 37. Tsiatis, A.A., Rosner, G.L. and Tritchler, D.L. (1985). Group sequential tests with censored survival data adjusting for covariates. *Biometrika* **72**, 365-373.