

東海大學

統計研究所

碩士論文

伽瑪脆弱模型下多組存活函數差異之聯合信賴區間估計

Estimation of the simultaneous confidence interval for the differences between several survival functions under a gamma frailty model

指導教授：張玉媚 博士

學 生：陳永村

中華民國一〇二年六月

致謝辭

這本論文的完成那一刻起，也代表著研究生生涯告一段落，不禁想到當初對統計的陌生到現在可以寫出這篇論文，尤其是當初進研究所時面對統計軟體的徬徨，到現在完成這篇文章模擬的部分需要用到的 R 軟體，在這兩年我從教授們的身上學到了很多。

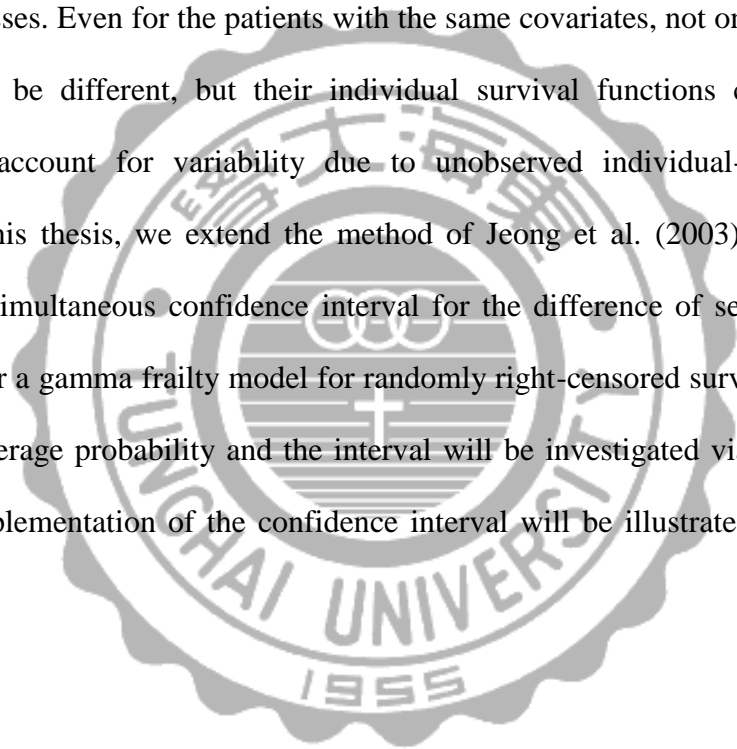
在心底誠摯的感謝指導教授張玉媚博士，承蒙張老師的悉心教導，讓我得以拓展統計學的視野，並建立正確的處事態度，令我受益匪淺。另外，本論文蒙陳春樹博士與俞一唐博士於百忙之中審核，斧正謬誤，並給予精闢的建議，使本文更完善，心中不勝感激。

在研究所修業期間，感謝統計系諸多老師傳授課業知識，東海統計所扎實的教學內容也讓我收穫頗豐。也幸虧有學長姐的指點，可以更快的解決不只課業上的問題還有行政上的幫助。還有我的同窗夥伴們，在面對困難的共同砥礪、閒暇時的閒話家常、熬夜趕作業的革命情感，這些都深深的記憶在我腦海裡。感謝家人的支持，使我無後顧之憂的在外地求學。



Abstract

Medical studies for drug development are frequently conducted to evaluate the treatment effects of a drug, where subjects or patients are randomly allocated to receive different treatment groups of the drug and a control. For right-censored survival data, it is usually interested in comparing the difference between two survival functions at a fixed time point rather than comparing the entire survival curves, when the survival functions are crossing. Moreover, in practice, a treatment effect may fade as time progresses. Even for the patients with the same covariates, not only might their survival times be different, but their individual survival functions could also be different. To account for variability due to unobserved individual-level factors, therefore, in this thesis, we extend the method of Jeong et al. (2003) and consider construct the simultaneous confidence interval for the difference of several survival functions under a gamma frailty model for randomly right-censored survival data. The associated coverage probability and the interval will be investigated via a simulation study. The implementation of the confidence interval will be illustrated using a real data set.



摘要

藥物反應研究中，為了探討某一種藥物的效果，經常進行數個不同的治療方式當作處理組和一個對照組來做比較。醫學上所蒐集到的資料經常出現右設限資料，導致分析的困難，針對右設限存活資料，當兩組存活函數呈現交叉時，通常不會檢定兩組的存活函數有無差異，而是針對特定的時間點下兩組存活函數是否有差異。此外，藥物的藥效可能隨時間而改變，具相同條件的病人，其療效也不見得相同。這是因為病人本身的異質性所造成，而引起變異的因素不見得觀察的到。因此，本文考慮推廣 Jeong et al. (2003)方法，在伽瑪脆弱(frailty)模型下，建立多組存活函數差異的聯合信賴區間。藉用模擬的方法來說明在特定時間點的多組存活函數差異聯合信賴區間的覆蓋機率以及面積的表現。最後將所提的方法應用於一筆實際資料做說明。



目錄

1.背景及研究目的	1
2.文獻回顧	2
3.多組存活函數差異之聯合信賴區間估計	4
4.模擬研究	6
5.實例分析	10
6.結論與建議	13
參考文獻	14
附錄	16



1. 研究背景及目的

在醫學研究中，經常感興趣的問題是接受不同處理治療的病人，其存活的時間是否有差異，針對右設限存活資料的多對一比較問題，已有相當多的文獻提出，例如：Chakraborti 和 Desu (1991)利用 Gehan (1965)的雙樣本統計量提出一個多對一的多重比較。Chen(1994, 2000)則根據數個 Gehan 雙樣本統計量提出一個 Steel 檢定(SMAX)，進一步根據加權對數秩(weighted logrank statistic)統計量，提出廣義的 steel 檢定即封閉降階檢定程序，進行多對一的多重比較。

在醫藥研究中，針對具共變數的右設限存活資料，Cox (1972)比例風險模型(proportional hazards model)經常被使用來描述病人的存活情形。但實際上，具不同共變數的病人其風險不見得成比例，且藥效會隨時間遞減。甚至具相同條件的病人，其療效也不見得相同，原因是病人本身的異質性所造成，而引起此變異的因素也不見得觀察的到。我們稱此因素為脆弱(frailty)效應。通常脆弱因子是個隱藏的效應，難以觀察得出來，因此通常會假設一穩定且合理的機率分配來描述他的影響。最常用的為伽瑪分配，因為伽瑪分配之脆弱模型可藉由拉普拉斯(Laplace transform)轉換，輕易的導出存活函數與累積風險函數，在先前的研究中，已被廣泛使用，例如：Greenwood 和 Yule (1920)、Vaupel et al. (1979)、Congdon (1995)、dos Santos et al. (1995)、Hougaard (2000)等。Jeong et al. (2003)則根據伽瑪脆弱模型，提出存活函數的區間估計。因此，本文考慮推廣 Jeong et al. (2003)的結果至多組存活函數差異的聯合信賴區間估計。

在本文第 2 章我們進行文獻回顧。第 3 章則描述如何建構多組存活函數差異的聯合信賴區間。第 4 章利用模擬研究所提方法在覆蓋機率及面積的表現。第 5 章引用一筆實際資料做說明。第 6 章則為結論與建議。

2. 文獻回顧

在醫學藥物試驗當中，在患者接受藥物治療結束後，會持續追蹤患者存活情形，但是治療效果會隨著時間而降低。在這樣的情況下，很容易違反比例風險模型之假設。於是，使用脆弱模型，藉由加入一個脆弱因子，使得預測存活機率更為合理。Jeong et al. (2003)根據伽瑪脆弱模型，配適違反比例風險模型假設的第三期乳癌之長期追蹤臨床試驗資料，針對不符合風險比例假設的資料作有效的估計與推論。以下回顧 Jeong et al. (2003)所提方法：在脆弱模型的假設下，第 i 個個體的風險函數定義為

$$h(t|w, z_i) = h_0(t)w e^{\beta z_i}$$

其中，在脆弱風險模型下 w 為脆弱因子，而 z_i 則代表個體 i 的共變數， $h_0(t)$ 為未知的基準風險函數。當假設脆弱因子 w 為伽瑪分配， $w \sim \Gamma\left(\frac{1}{\gamma}, \frac{1}{\gamma}\right)$ ， $E(w)=1$ ，由 Hougaard (1984) 及 Dabrowska and Doksum (1988)，可得

$$S(t|z_i) = \{1 + \gamma(\rho t)^\kappa \exp(\beta z_i)\}^{-1/\gamma} \quad (1)$$

在 Jeong et al. (2003) 的文章中已證明：

[1]. 當 $\gamma = 1$ 時將滿足比例勝算的特性：
$$\frac{1 - S(t|z_i=1)}{S(t|z_i=1)} = e^\beta \frac{1 - S(t|z_i=0)}{S(t|z_i=0)}$$

[2]. 當 γ 趨近於 0，對於任何假設下的基準存活函數 $S_0(t)$ ， $z=0$ 和 $z=1$ 的兩個邊際存活函數關係將近似於比例風險函數： $\log\{S(t|z_i = 1)\} = e^\beta \log\{S(t|z_i = 0)\}$ 。

假設 T 表示存活時間， C 表示設限時間，定義 $X_i = \min(T_i, C_i)$; $\delta_i = I(T_i \leq C_i)$; $\psi = (\gamma, \beta, \rho, \kappa)$ 。對數概似函數(log-likelihood) 可寫成

$$L(\psi) = \sum_{i=1}^n \delta_i \log\{h(x_i | \psi, z_i)\} + \sum_{i=1}^n \log\{S(x_i | \psi, z_i)\}$$

其中

$$\log\{h(x_i | \psi, z_i)\} = \beta z_i + \log(\rho) + (\kappa - 1) \log(\rho x_i) - \log\{1 + \gamma(\rho x_i)^\kappa \exp(\beta z_i)\} + \log(\kappa)$$

和

$$\log\{S(x_i | \psi, z_i)\} = (-1/\gamma) \log\{1 + \gamma(\rho x_i)^\kappa \exp(\beta z_i)\}$$

可藉由對 $L(\psi)$ 的一階和二階偏微分及參數估計值分別得到分數函數(score function)和觀察資訊矩陣(observed information matrix)，詳列於附錄 A。利用數值分析的方法求取 $\hat{\Psi}=(\hat{\gamma}, \hat{\beta}, \hat{\rho}, \hat{\kappa})$ ，把估計出來的 $\hat{\Psi}$ 代入(1)式可求得在時間點 t 的存活函數

$$\hat{S}(t|z_i, \hat{\Psi}) = \{1 + \hat{\gamma}(\hat{\rho}t)^{\hat{\kappa}} \exp(\hat{\beta}z_i)\}^{-1/\hat{\gamma}}$$

使用 delta method 估算存活函數的變異數為

$$\widehat{Var}\{\hat{S}(t|z_i, \hat{\Psi})\} = \left(\frac{\partial S(t|\psi, z_i)}{\partial \psi}\right) \Big|_{\psi=\hat{\Psi}} I^{-1}(\hat{\Psi}) \left(\frac{\partial S(t|\psi, z_i)}{\partial \psi}\right)' \Big|_{\psi=\hat{\Psi}}$$

其中 $\left(\frac{\partial S(t|\psi, z_i)}{\partial \psi}\right) = \left(\frac{\partial S(t|\psi, z_i)}{\partial \gamma}, \frac{\partial S(t|\psi, z_i)}{\partial \beta}, \frac{\partial S(t|\psi, z_i)}{\partial \kappa}, \frac{\partial S(t|\psi, z_i)}{\partial \rho}\right)$ 是存活函數一階偏微分，

為一個向量。所以在時間 t 的 $S(t|z_i, \psi)$ 的 95% 信賴區間為

$$\hat{S}(t|z_i, \hat{\Psi}) \pm 1.96 \times \sqrt{\widehat{Var}\{\hat{S}(t|z_i, \hat{\Psi})\}}$$



3. 多組存活函數差異之聯合信賴區間估計

Jeong et al. (2003) 的文章中主要是針對處理組 $K=1$ 的狀況在討論，即只考慮到一個對照組和一個治療組，而且重點放在參數的估計以及對照組和治療組各自的信賴區間，所以本文延續 Jeong et al. (2003) 的做法，研究當共變數項不再是 $K=1$ ，而是 $K=2$ 或更多的時候，該如何去建構固定時間點 t 下的存活函數差距之聯合信賴區間。

當有 K 組處理組及一個對照組的時候，令第 j 組的存活函數 $S_j(t) = S(t|z_j = 1, \psi)$ ， $j=1, 2, \dots, K$ ； $S_0(t)$ 則表示對照組。令 $\mathbf{Z} = (z_1, z_2, \dots, z_K)$ ，當 $z_j = 1$ ，為第 j 種治療方式；當 $z_j = 0$ ，為其它， \mathbf{Z}_i 表示第 i 個個體所接受的處理組別之向量。本文考慮的模式為

$$h(t|\mathbf{w}, \mathbf{Z}) = h_0(t) w e^{\beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k z_k}$$

其中參數 $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_K)$ ，存活函數為

$$S(t|\mathbf{Z}) = \{1 + \gamma(\rho t)^\kappa \exp(\beta_1 z_1 + \beta_2 z_2 + \dots + \beta_k z_k)\}^{-1/\gamma}$$

對數概似函數為

$$L(\psi) = \sum_{i=1}^n \delta_i \log\{h(x_i|\psi, \mathbf{Z}_i)\} + \sum_{i=1}^n \log\{S(x_i|\psi, \mathbf{Z}_i)\}$$

其中

$$\log\{h(x_i|\psi, \mathbf{Z}_i)\} = \boldsymbol{\beta}\mathbf{Z}_i + \log(\rho) + (\kappa - 1)\log(\rho x_i) - \log\{1 + \gamma(\rho x_i)^\kappa \exp(\boldsymbol{\beta}\mathbf{Z}_i)\} + \log(\kappa)$$

和

$$\log\{S(x_i|\psi, \mathbf{Z}_i)\} = (-1/\gamma)\log\{1 + \gamma(\rho x_i)^\kappa \exp(\boldsymbol{\beta}\mathbf{Z}_i)\}$$

同樣地方法藉由對 $L(\psi)$ 的一階和二階偏微分分別得到分數函數和觀察資訊矩陣，

利用數值分析的方法求取 ψ 的估計 $\hat{\psi} = (\hat{\gamma}, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_k, \hat{\rho}, \hat{\kappa})$ 。同理，

$$\widehat{Var}\{\hat{S}(t|\mathbf{Z}_i, \hat{\psi})\} = \left(\frac{\partial S(t|\psi, \mathbf{Z}_i)}{\partial \psi}\right) \Big|_{\psi=\hat{\psi}} I^{-1}(\hat{\psi}) \left(\frac{\partial S(t|\psi, \mathbf{Z}_i)}{\partial \psi}\right)' \Big|_{\psi=\hat{\psi}}$$

$$\widehat{cov}\{\hat{S}_j, \hat{S}_0\} = \left(\frac{\partial S_j(t)}{\partial \psi}\right) \Big|_{\psi=\hat{\psi}} I^{-1}(\hat{\psi}) \left(\frac{\partial S_0(t)}{\partial \psi}\right)' \Big|_{\psi=\hat{\psi}}$$

其中

$$\left(\frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \Psi}\right) = \left(\frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \gamma}, \frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \beta_1}, \dots, \frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \beta_k}, \frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \kappa}, \frac{\partial S(t|\Psi, \mathbf{Z}_i)}{\partial \rho}\right)$$

在建立治療組對對照組存活函數差異之聯合信賴區間之前，首先要先考慮兩個存活函數相減的變異數。就如上述 $\widehat{Var}\{\hat{S}_j(t)\}$ 可使用 delta method 和所估算出來的 $\hat{\Psi}$ 求得，所以 $\widehat{Var}\{\hat{S}_j(t) - \hat{S}_0(t)\} = \widehat{Var}\{\hat{S}_j(t)\} + \widehat{Var}\{\hat{S}_0(t)\} - 2\widehat{Cov}\{\hat{S}_j(t), \hat{S}_0(t)\}$ 。

在虛無假設 $H_0: S_j(t) = S_0(t)$ 之下， $j=1,2,\dots,K$ ， $\{S_j(t)-S_0(t); j=1,2,\dots,K\}$ 會漸近於 K 維常態分配(Andersen et al., 1993)，標準化後會服從 $N_K(\mathbf{0}, \Sigma)$ ，其中變異數共變異數矩陣

$$\Sigma = \begin{bmatrix} 1 & \dots & 0.5 \\ \vdots & \ddots & \vdots \\ 0.5 & \dots & 1 \end{bmatrix}$$

根據 K 維常態分配，便可計算其臨界值，記為 $C_{\alpha/2}$ ，滿足

$$\int_{-C_{\alpha/2}}^{C_{\alpha/2}} \int_{-C_{\alpha/2}}^{C_{\alpha/2}} \dots \int_{-C_{\alpha/2}}^{C_{\alpha/2}} f(\mathbf{X}) d\mathbf{X} = 1 - \alpha, \quad f(\mathbf{X}) \text{ 為 } K \text{ 維標準常態分配函數，因此}$$

$\{S_j(t) - S_0(t), j=1,2,\dots,K\}$ 的 $100(1-\alpha)\%$ 聯合信賴區間為

$$\{ \hat{S}_j(t) - \hat{S}_0(t) \pm C_{\alpha/2} \sqrt{\widehat{Var}\{\hat{S}_j(t)\} + \widehat{Var}\{\hat{S}_0(t)\} - 2\widehat{Cov}\{\hat{S}_j(t), \hat{S}_0(t)\}}, \quad j=1,2,\dots,K \}$$

4. 模擬研究

本章利用模擬方法，研究本文所提方法與 Bonferroni 的多重比較方法在覆蓋機率及面積的表現。考慮的存活函數為

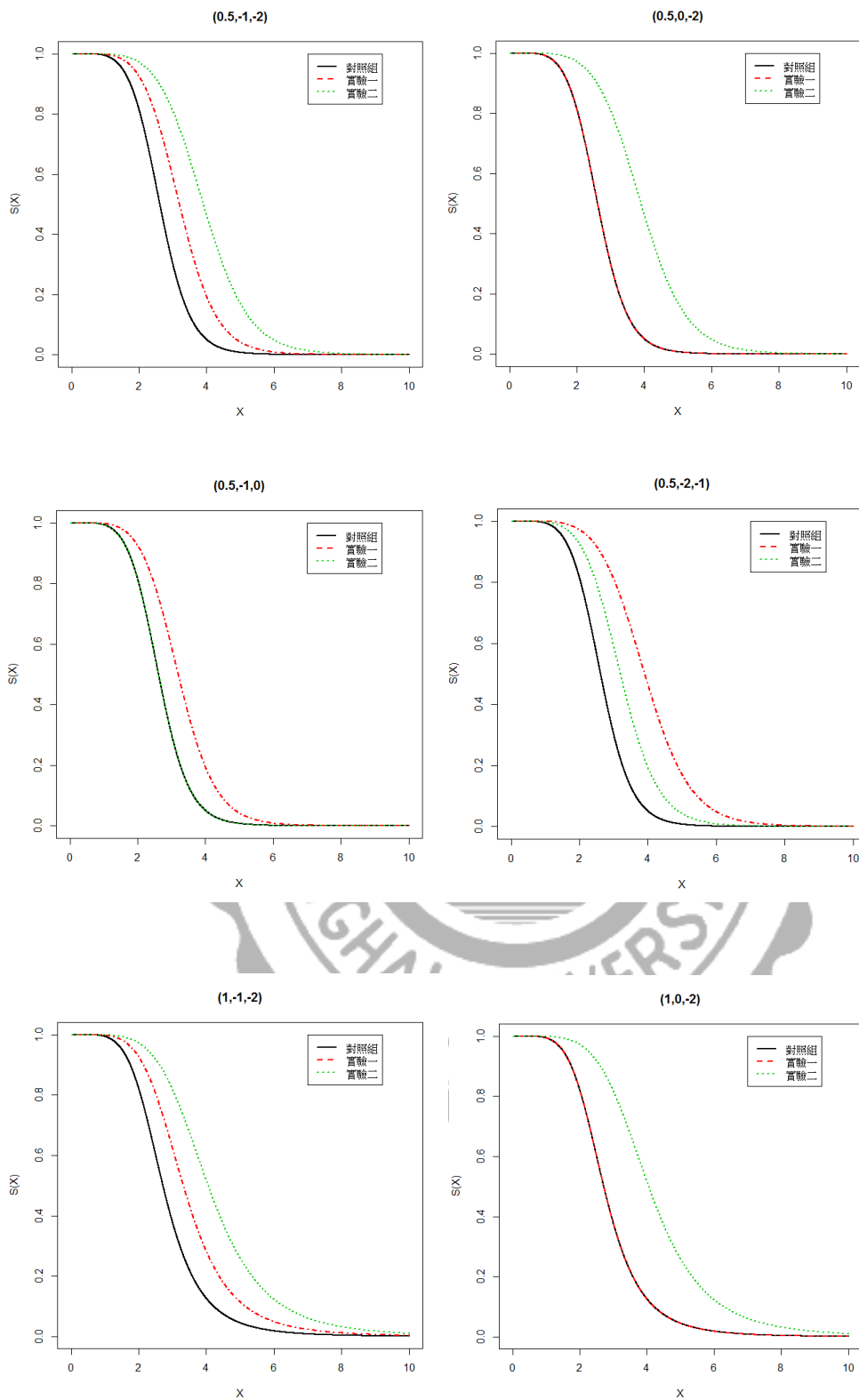
$$S(t|z_1, z_2) = \{1 + \gamma(\rho t)^\kappa \exp(\beta_1 z_1 + \beta_2 z_2)\}^{-1/\gamma}$$

其中 $z_i=1$ 表示第 i 組；0 表其他， $i=1,2$ ，即 $\mathbf{Z} = (0, 0), (1, 0), (0, 1)$ 分別為對照組，處理組一，處理組二。此外令伽瑪脆弱的參數 $\gamma = 0.5, 1, 2$ ；以及 $\beta = (-1, -2), (0, -2), (-1, 0), (-2, -1)$ 分別表示當藥劑濃度越高越好、只有高濃度有效、只有低濃度有效、低濃度比高濃度好，其存活函數圖列於圖一。假設設限分配為 $U(4,10)$ 。各組樣本數 $n_0=n_1=n_2=100$ 、 $\rho=0.37$ 、 $\kappa=5$ 。執行模擬 10000 次，在信心水準 95% 下之誤差率估計量的標準誤差約為 $0.002(\approx \sqrt{(0.05)(0.95)/10000})$ ；90% 下之誤差率估計量的標準誤差約為 $0.003(\approx \sqrt{(0.1)(0.9)/10000})$ 。結果呈現於表一至表二。

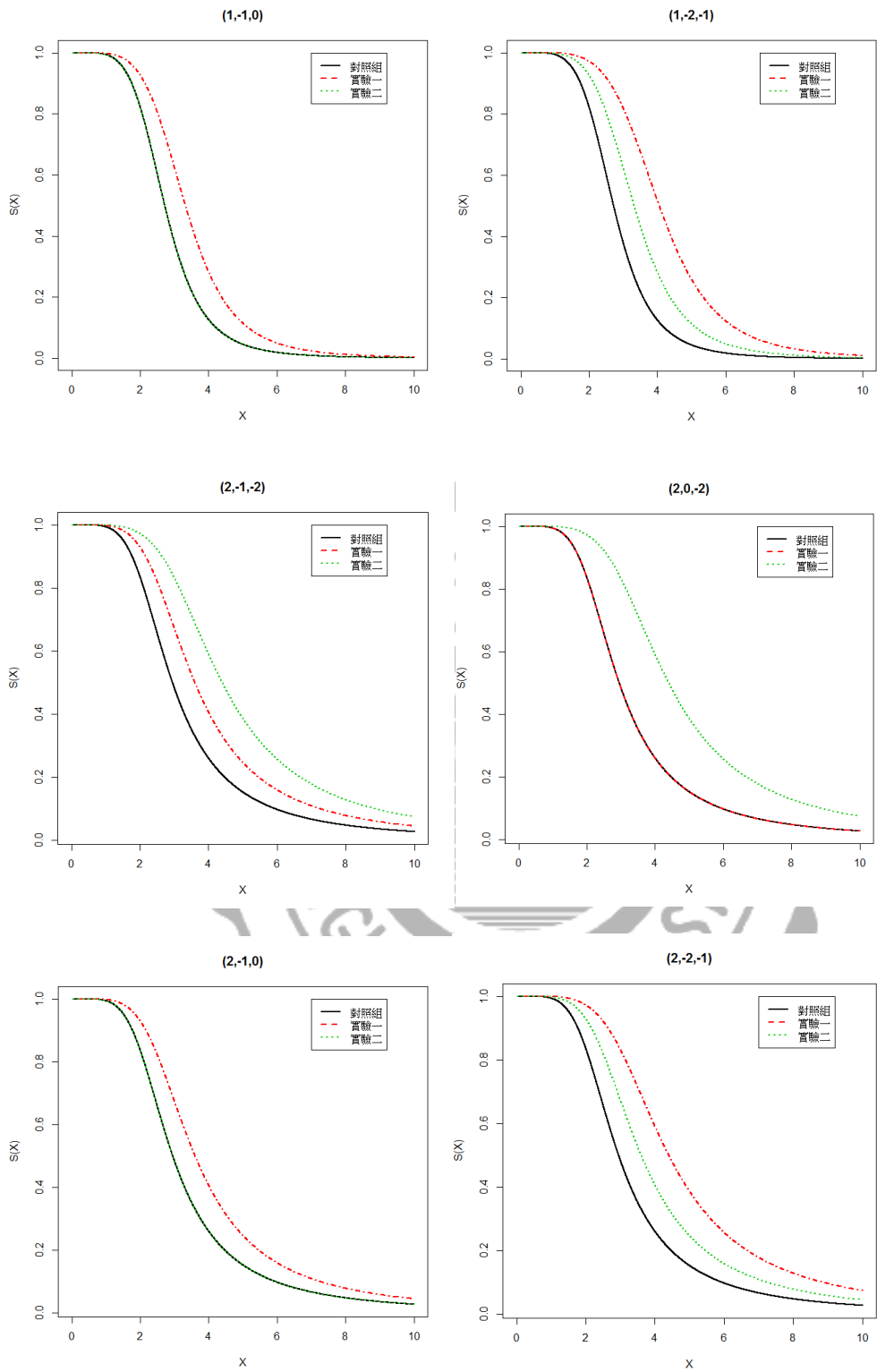
由表一可以看出估計出來的參數值差不多會落在真實的參數值附近。由表二可以看出在給定的時間點 $T=t$ 之下，本文所提 MNor 的方法比 Bonferroni 法來得好，覆蓋機率較接近信心水準。尤其，當 $T=5$ 時，Bonferroni 的覆蓋機率明顯地高於信心水準。

表一：抽取其中一種真實參數組合的估計及 10000 次的參數估計平均、標準差

	γ	β_1	β_2	ρ	κ
真實值	2	-2	-1	0.37	5
其中一組	2.50	-2.22	-0.53	0.36	6.34
平均值	2.06	-2.05	-1.02	0.36	5.13
標準差	0.461	0.406	0.351	0.021	0.598



圖一_1: 不同參數(γ, β_1, β_2)的存活函數曲線



圖一_2: 不同參數 $(\gamma, \beta_1, \beta_2)$ 的存活函數曲線

表二：在信心水準分別為 95%和 90%下估計的覆蓋機率

γ	β_1	β_2	$1-\alpha$	T=3		T=4		T=5	
				Bonfer	MNor	Bonfer	MNor	Bonfer	MNor
0.5	-1	-2	0.95	0.942	0.952	0.949	0.950	0.949	0.947
			0.90	0.900	0.906	0.903	0.896	0.909	0.899
	0	-2	0.95	0.943	0.954	0.961	0.944	0.970	0.935
			0.90	0.895	0.901	0.923	0.886	0.943	0.886
	-1	0	0.95	0.947	0.950	0.971	0.942	0.980	0.938
			0.90	0.898	0.898	0.938	0.892	0.963	0.892
-2	-1	0.95	0.940	0.953	0.943	0.949	0.943	0.947	
		0.90	0.893	0.905	0.892	0.896	0.908	0.898	
1	-1	-2	0.95	0.943	0.951	0.951	0.948	0.974	0.945
			0.90	0.899	0.904	0.905	0.897	0.940	0.893
	0	-2	0.95	0.949	0.954	0.958	0.947	0.976	0.944
			0.90	0.899	0.902	0.923	0.892	0.953	0.890
	-1	0	0.95	0.951	0.952	0.973	0.950	0.995	0.944
			0.90	0.903	0.899	0.940	0.90	0.984	0.893
-2	-1	0.95	0.943	0.952	0.949	0.948	0.973	0.947	
		0.90	0.898	0.906	0.899	0.894	0.942	0.896	
2	-1	-2	0.95	0.944	0.953	0.953	0.946	0.978	0.946
			0.90	0.897	0.903	0.910	0.897	0.949	0.895
	0	-2	0.95	0.947	0.951	0.960	0.946	0.980	0.945
			0.90	0.901	0.899	0.924	0.894	0.957	0.892
	-1	0	0.95	0.949	0.950	0.971	0.947	0.991	0.946
			0.90	0.902	0.898	0.936	0.898	0.977	0.897
-2	-1	0.95	0.941	0.949	0.954	0.947	0.978	0.946	
		0.90	0.897	0.903	0.908	0.897	0.948	0.897	

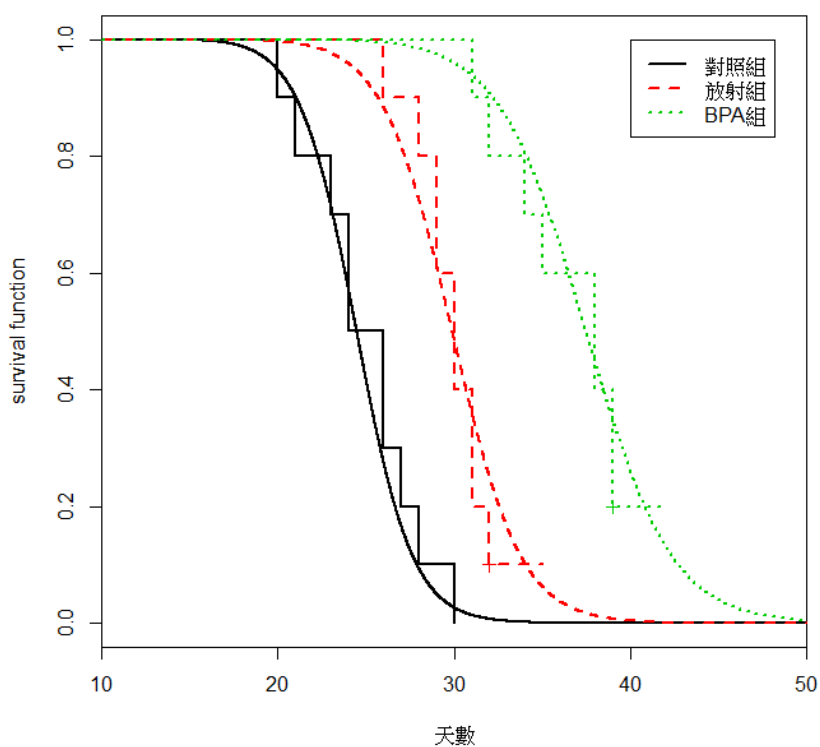
5. 實例分析

Klein 與 Moeschberger (1997)提出一個動物實驗，主要在探討對其他放射線治療無效之 F98-神經膠質瘤(F98-Glioma),若使用 boronophenylalanine(BPA)作為放射線擷取媒介物質時，是否能延長 F98-神經膠質瘤患者的生命。研究者將 F98-神經膠質瘤細胞植入老鼠腦部，研究對象為三組老鼠，第一組為對照組不予治療，第二組僅用放射線治療，第三組接受適當注射 BPA 後，再加上放射線治療，三組老鼠死亡時間(以天為單位紀錄)。資料列於表三。三組老鼠的 Kaplan-Meier(1958)存活函數估計繪於圖二。由圖二跟脆弱模型下之結果可以看出三種處理有的存活機率明顯的差別，而且放射線治療雖然有效果，但是有植入 BPA 的老鼠比只做放射線治療的老鼠有更高的存活機會。

參數 $\psi = (\gamma, \beta_1, \beta_2, \rho, \kappa)$ 的估計值為 $\hat{\gamma}=0.63$ 、 $\hat{\beta}_1=-2.73$ 、 $\hat{\beta}_2=-5.83$ 、 $\hat{\rho} = 0.04$ 、 $\hat{\kappa} = 13.67$ 。在第 25 天、第 30 天、第 35 天，令 D_1 表示放射線組的存活函數減對照組的存活函數， D_2 表示 BPA 組對對照組存活函數的差距，計算 D_1 、 D_2 的聯合信賴區間估計，結果列於表四。由表四看的出來，在前期 $T=25$ 及 30 天時兩種治療皆有效果，但隨著時間增加，BPA 明顯地比只做放射線的好。然而，在後期 $T=35$ 天的時候，由信賴區間可以知道只有放射線治療並不能提高存活機會。此外，本文所提方法 MNor 估出來的信賴區間會比較窄，可見 Bonfferoni 是比較保守的估計方法。

表三: F98-神經膠質瘤老鼠動物實驗:三組治療之存活時間(+表示設限觀測值)

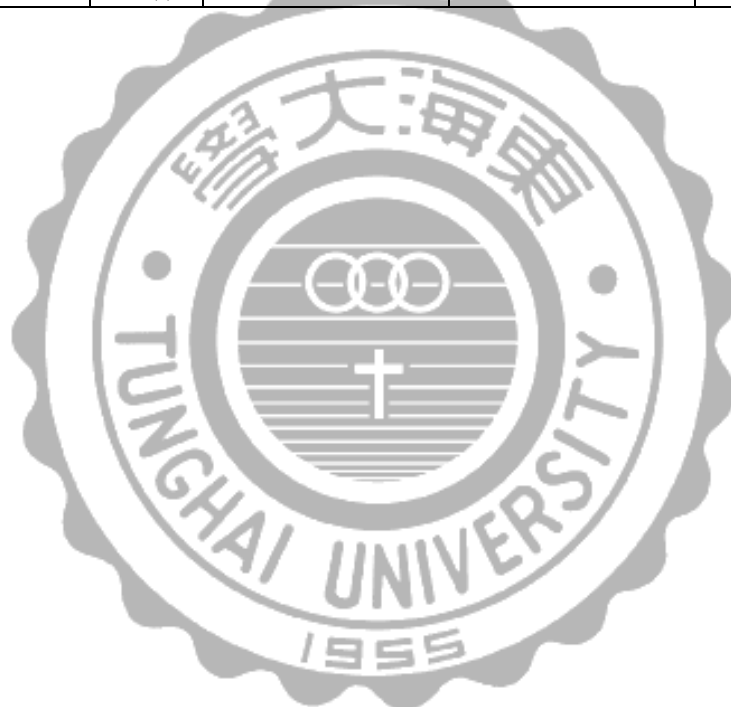
對照組	放射線組	放射線+BPA
20	26	31
21	28	32
23	29	34
24	29	35
24	30	36
26	30	38
26	31	38
27	31	39
28	32	42+
30	35+	42+



圖二: F98-老鼠的存活曲線

表四：治療組存活函數和對照組的差距之聯合信賴區間

			T=25	T=30	T=35
	點估 計	$\widehat{D}_1(t)$	0.51	0.460	0.061
		$\widehat{D}_2(t)$	0.57	0.933	0.733
Bonferroni	95%	$\widehat{D}_1(t)$	(0.154, 0.867)	(0.169, 0.751)	(-0.057, 0.181)
		$\widehat{D}_2(t)$	(0.238, 0.918)	(0.842, 1.024)	(0.490, 0.976)
	90%	$\widehat{D}_1(t)$	(0.199, 0.822)	(0.206, 0.715)	(-0.042, 0.166)
		$\widehat{D}_2(t)$	(0.281, 0.875)	(0.853, 1.012)	(0.520, 0.946)
MNor	95%	$\widehat{D}_1(t)$	(0.159, 0.862)	(0.173, 0.747)	(-0.055, 0.179)
		$\widehat{D}_2(t)$	(0.242, 0.913)	(0.843, 1.023)	(0.493, 0.973)
	90%	$\widehat{D}_1(t)$	(0.206, 0.815)	(0.211, 0.709)	(-0.040, 0.164)
		$\widehat{D}_2(t)$	(0.287, 0.868)	(0.855, 1.011)	(0.525, 0.941)



6. 結論與建議

本文主要是在針對多組右設限資料，推廣 Jeong et al. (2003)的結果，在 Cox 模型中引入一個伽瑪脆弱因子，並建立多組存活函數差異的聯合信賴區間，用以了解病人接受治療後的療效差異範圍。由模擬及資料分析結果皆顯示本文所提方法較既有的 Bonferroni 方法提供更多訊息，因為 Bonferroni 是取 $\alpha/(2K)$ 所對應的臨界值，尤其是在高維度的時候更為保守。

若收集的資料亦包含其他的共變數訊息，例如：腫瘤的大小、人種，癌症的期數...等，後續可以更進一步探討當加入其他變數的影響後，多個處理組存活函數差異的聯合信賴區間估計問題。



參考文獻

Andersen PK., Borgan Q., Gill, RD., and Keiding N. (1993). Statistical models based on counting processes. New York: Springer-Verlag.

Chakraborti S. Desu MM. (1991). Linear rank tests for comparing treatments with a control when data are subject to unequal patterns of censorship. *Statistical Neerlandica* 45,227-254.

Chen YI. (2000). Multiple comparisons in carcinogenesis study with right-censored survival data. *Statistics in Medicine* 19,353-367.

Congdon P. (1995). Modelling frailty in area mortality. *Statistics in Medicine* 14,1859-74.

Cox DR. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society Series B Methodological* Vol. 34,187-220.

Dabrowska DM, Doksum KA. (1988). Estimation and testing in a two-sample generalized odds-rate model. *Journal of the American Statistical Association* 83,744-749.

dos Santos DM., Davies RB., Francis B. (1995). Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. *Journal of Statistical Planning and Inference* 47,111-127.

Gehan EA. (1965). A generalized Wilcoxon test for comparing arbitrary singly-censored samples. *Biometrika* 52,2003-223.

Greenwood M., Yule GU. (1920) An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society* 83,255-279.

Hougaard P. (1986). Survival models for heterogeneous populations derived from stable distribution. *Biometrika* 73,387-396.

Hougaard P.(1984). Life table methods for heterogeneous populations: distributions

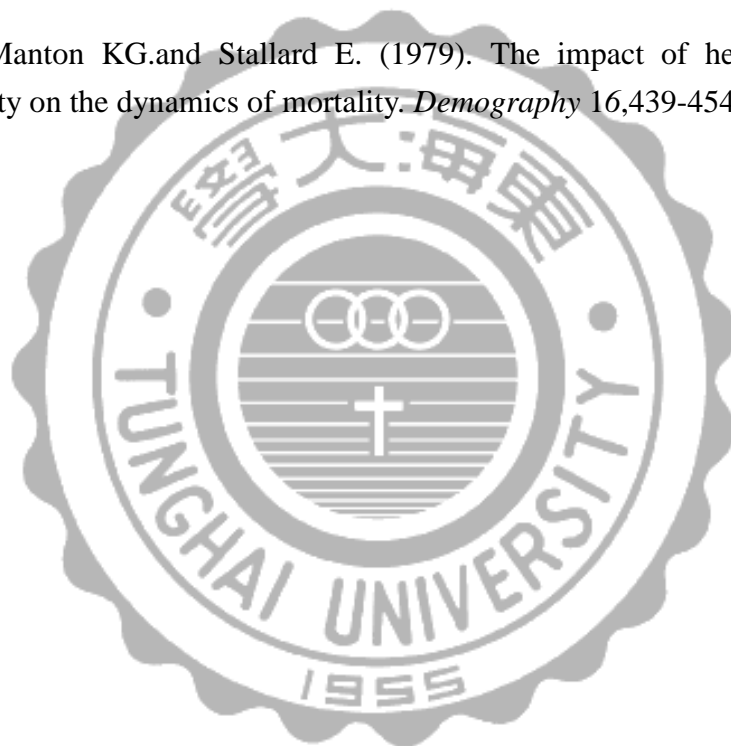
describing the heterogeneity. *Biometrika* 61,17-30.

Jeong JH., Jung SH. and Wieand S. (2003). A parametric model for long-term follow-up data from phase III breast cancer clinical trials. *Statist in Medicine* 22,339-352.

Kaplan EL. and Meier P. (1958). Nonparametric estimation from incomplete observations. *Journal of the American statistical association* 53,457-481.

Klein JP. And Moeschberger ML. (1997). Survival analysis: techniques for censor and truncated data. *Spinger-Verlag*.

Vaupel JW., Manton KG. and Stallard E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* 16,439-454.



附錄

附錄 A: 令 $\mu(t)=(\rho t)^\kappa$ ，當對對數概似函數 $L(\psi)$ 的各個參數一階微分可得以下的分數函數:

$$\frac{\partial L(\psi)}{\partial \gamma} = \frac{1}{\gamma} \sum_{i=1}^n \left[\left(\frac{1}{\gamma} \right) \log\{1 + \gamma \exp(\beta z_i) \mu(t)\} - \frac{\exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right] - \sum_{i=1}^n \delta_i \left\{ \frac{\exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\}$$

$$\frac{\partial L(\psi)}{\partial \beta} = - \sum_{i=1}^n \left\{ \frac{z_i \exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\} + \sum_{i=1}^n \delta_i \left\{ \frac{\gamma z_i \exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\}$$

$$\frac{\partial L(\psi)}{\partial \kappa} = - \frac{1}{\kappa} \sum_{i=1}^n \left\{ \frac{\exp(\beta z_i) \mu(t) \log\{\mu(t)\}}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\} + \frac{1}{\kappa} \sum_{i=1}^n \delta_i \left\{ \log\{\mu(t)\} - \frac{\gamma \exp(\beta z_i) \mu(t) \log\{\mu(t)\}}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\}$$

$$\frac{\partial L(\psi)}{\partial \rho} = - \frac{\kappa}{\rho} \sum_{i=1}^n \left\{ \frac{\exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\} + \frac{\kappa}{\rho} \sum_{i=1}^n \delta_i \left\{ 1 - \frac{\gamma \exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\}$$

同樣地，在 $L(\psi)$ 對各個參數二階偏微分，將 $\hat{\psi}$ 代入 ψ 可得以下的觀察資訊矩陣

$$I(\hat{\psi}) = \begin{bmatrix} i_{\gamma\gamma} & i_{\gamma\beta} & i_{\gamma\kappa} & i_{\gamma\rho} \\ i_{\gamma\beta} & i_{\beta\beta} & i_{\beta\kappa} & i_{\beta\rho} \\ i_{\gamma\kappa} & i_{\beta\kappa} & i_{\kappa\kappa} & i_{\kappa\rho} \\ i_{\gamma\rho} & i_{\beta\rho} & i_{\kappa\rho} & i_{\rho\rho} \end{bmatrix},$$

其中

$$i_{\gamma\gamma} = - \frac{1}{\gamma^2} \sum_{i=1}^n \left[\frac{2 \exp(\beta z_i) \mu(t) \{1 + \frac{\gamma}{2} \exp(\beta z_i) \mu(t)\}}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} - \frac{2}{\gamma} \log\{1 + \gamma \exp(\beta z_i) \mu(t)\} \right]$$

$$- \sum_{i=1}^n \delta_i \left\{ \frac{\exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} \right\}^2$$

$$i_{\beta\beta} = \sum_{i=1}^n \left[\frac{z_i^2 \exp(\beta z_i) \mu(t)}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \sum_{i=1}^n \delta_i \left[\frac{\gamma z_i^2 \exp(\beta z_i) \mu(t)}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} \right]$$

$$i_{\kappa\kappa} = \frac{1}{\kappa^2} \sum_{i=1}^n \left[\frac{\exp(\beta z_i) [\log\{\mu(t)\}]^2}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \frac{1}{\kappa^2} \sum_{i=1}^n \delta_i \left[\frac{\gamma \exp(\beta z_i) \mu(t) [\log\{\mu(t)\}]^2}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} \right]$$

$$i_{\rho\rho} = \frac{\kappa}{\rho^2} \sum_{i=1}^n \left[- \frac{\exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} + \frac{\kappa \exp(\beta z_i) \mu(t)}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \frac{\kappa}{\rho^2} \sum_{i=1}^n \delta_i \left[- \frac{\gamma \exp(\beta z_i) \mu(t)}{1 + \gamma \exp(\beta z_i) \mu(t)} + \right.$$

$$\left. \frac{\gamma \kappa \exp(\beta z_i) \mu(t)}{\{1 + \gamma \exp(\beta z_i) \mu(t)\}^2} + 1 \right]$$

$$\begin{aligned}
i_{\gamma\beta} &= -\sum_{i=1}^n \left[\frac{z_i^2 \{\exp(\beta z_i) \mu(t)\}^2}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \sum_{i=1}^n \delta_i \left[\frac{z_i \exp(\beta z_i) \mu(t)}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
i_{\gamma\kappa} &= -\frac{1}{\kappa} \sum_{i=1}^n \left[\frac{\{\exp(\beta z_i) \mu(t)\}^2 \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \frac{1}{\kappa} \sum_{i=1}^n \delta_i \left[\frac{\exp(\beta z_i) \mu(t) \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
i_{\gamma\rho} &= -\frac{\kappa}{\rho} \sum_{i=1}^n \left\{ \frac{\exp(\beta z_i) \mu(t)}{1+\gamma \exp(\beta z_i) \mu(t)} \right\}^2 + \frac{\kappa}{\rho} \sum_{i=1}^n \delta_i \left[\frac{\exp(\beta z_i) \mu(t)}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
i_{\beta\kappa} &= \frac{1}{\kappa} \sum_{i=1}^n \left[\frac{z_i \exp(\beta z_i) \mu(t) \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \frac{1}{\kappa} \sum_{i=1}^n \delta_i \left[\frac{\gamma z_i \exp(\beta z_i) \mu(t) \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
i_{\beta\rho} &= \frac{\kappa}{\rho} \sum_{i=1}^n \left[\frac{z_i \exp(\beta z_i) \mu(t)}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] + \frac{\kappa}{\rho} \sum_{i=1}^n \delta_i \left[\frac{\gamma z_i \exp(\beta z_i) \mu(t)}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
i_{\kappa\rho} &= \frac{1}{\rho} \sum_{i=1}^n \left[\frac{\exp(\beta z_i) \mu(t)}{1+\gamma \exp(\beta z_i) \mu(t)} + \frac{\exp(\beta z_i) \mu(x_i) \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} \right] \\
&\quad + \frac{1}{\rho} \sum_{i=1}^n \delta_i \left[\frac{\gamma \exp(\beta z_i) \mu(t)}{1+\gamma \exp(\beta z_i) \mu(t)} + \frac{\gamma \exp(\beta z_i) \mu(x_i) \log\{\mu(t)\}}{\{1+\gamma \exp(\beta z_i) \mu(t)\}^2} - 1 \right]
\end{aligned}$$

$S(t|\psi, z_i)$ 對各個參數的一階偏微分：

$$\begin{aligned}
\frac{\partial S(t|\psi, z_i)}{\partial \gamma} &= \frac{1}{\gamma} \left[\frac{\log\{1+\gamma \exp(\beta z_i) \mu(t)\}}{\gamma \{1+\gamma \mu(t) \exp(\beta z_i)\}^{-1/\gamma}} - \frac{\exp(\beta z_i) \mu(t)}{\{1+\gamma \mu(t) \exp(\beta z_i)\}^{1+1/\gamma}} \right] \\
\frac{\partial S(t|\psi, z_i)}{\partial \beta} &= -z_i \mu(t) \exp(\beta z_i) \{1+\gamma \mu(t) \exp(\beta z_i)\}^{-(1+1/\gamma)} \\
\frac{\partial S(t|\psi, z_i)}{\partial \kappa} &= -(1/\kappa) \mu(t) \exp(\beta z_i) \log\{\mu(t)\} \{1+\gamma \mu(t) \exp(\beta z_i)\}^{-(1+1/\gamma)} \\
\frac{\partial S(t|\psi, z_i)}{\partial \rho} &= -(\kappa/\rho) \mu(t) \exp(\beta z_i) \{1+\gamma \mu(t) \exp(\beta z_i)\}^{-(1+1/\gamma)}
\end{aligned}$$