

東海大學

統計研究所

碩士論文

餘命分位數比值的聯合信賴域估計與應用

Estimation of the simultaneous confidence regions
for the ratios of quantile residual lifetimes

指導教授：張玉媚 博士

學生：江玉茹

中華民國一〇五年六月

致謝詞

經歷兩年的努力，終於完成本篇論文，從整個的學習及撰寫的過程中，收穫良多，不僅是知識上的吸收，更讓自我成長許多。此篇論文之完成，要感謝許多在碩士生涯中幫助過我的人，其中最要感謝的是我的指導教授張玉媚博士，老師認真、專注、嚴厲且耐心的指導，給予我許多重要的啟發與成長，讓我能夠在我存活分析領域獲得許多寶貴的經驗與學習。此外，也十分感謝沈葆聖老師在這些年多來給我指導與協助，在繁忙的日子裡，不忘細心指導我的論文，使我的整個思路及思想更為清晰明確。

論文審查及口試期間，承蒙陳春樹老師，沈葆聖老師提供許多寶貴的意見，讓我的論文能更加完整及明確，在此僅呈上最誠摯之謝意。此外我也感謝在東海及統計系所中所有教導過我的老師與助教們，有你們的協助及指教，才能造就出我今日之成就與成長。

在求學的期間，很幸運能夠擁有許多一起研究、努力、歡笑及扶持的同學們，讓煩悶的生活有了互相的陪伴，共同度過種種的阻礙及歡度的學生的快樂時光，我會永遠記得求學生涯的點點滴滴，雖然在未來大家要各奔東西，但我很高興與你們的歡樂時光。最後我要謝謝我的家人，一直默默的支持著我，容忍我因論文的研究及撰寫而減少陪伴他們的時間。謹將本論文獻給所有關心我的人。

目錄

Abstract

摘要

第一章 背景及研究目的.....	P.1~2
第二章 文獻回顧	
2.1 無母數推論.....	P.3~6
2.2 經驗概似比推論.....	P.6~8
第三章 多組餘命分位數比值的聯合信賴域	
3.1 無母數聯合信賴域.....	P.9~12
3.2 經驗概似比聯合信賴域	P.12~14
第四章 模擬研究	
4.1 模擬方法.....	P.15~18
4.2 模擬結果.....	P.19~23
第五章 資料分析	
5.1 資料介紹.....	P.24~25
5.2 資料分析	
5.2.1 整體的治療效果分析.....	P.26~27
5.2.2 不同性別之下治療方式的差異	P.27~31
5.2.3 不同年齡之下治療方式的差異	P.31~35
5.2.4 大腸癌侵犯淋巴結個數來看治療方式的差異 ...	P.36~40
第六章 結論與建議	P.41
參考文獻.....	P.42~43
附錄	
附錄 A-1: $W_{t_0}(\gamma_{t_0,0})$ 的極限分布	P.44~45
附錄 A-2: $Q(\gamma_{t_0,0})$ 的極限分布	P.45~46

Abstract

The estimation of mean residual lifetime or quantile residual lifetime is becoming more critical in cancer research, since it provides patients and physicians with information about the efficacy of the treatment therapies in a more straightforward way using the hazard function or probability of survival. In medical study, the multiple comparison procedures are the commonly used multiple decisions that assess the differences of the effects for several treatment groups, simultaneously. Because the simultaneous confidence regions maybe used for hypotheses testing, as well as providing the information about magnitudes of the differences between the survival indices of each treatment group. In this project for censored survival data, we consider the construction of the simultaneous confidence regions for the ratios of quantile residual lifetimes between the treatment group and the control group based on the nonparametric method and the empirical likelihood ratio tests. The performance of the associated coverage probability and the average coverage area will be investigated via a simulation study. The use of the proposed procedure will be illustrated using real data sets.

Keywords : censored survival data; confidence region; minimum dispersion statistic; multiple comparison; quantile residual lifetime

摘要

平均餘命(residual lifetime)或餘命的分位數(quantile)的估計在癌症醫學研究越來越重要，因為它提供了比風險函數或存活機率在治療效果上更直覺的訊息。在醫學上，經常利用多重比較法(multiple comparison procedure)評估多個處理效應的異同。因聯合信賴域不但能夠進行檢定，能提供更多個處理之間差異的範圍。因此，本文針對多組右設限存活資料，分別根據無母數推論方法(nonparametric method)及經驗概似比檢定(empirical likelihood ratio test)統計量建立各處理組與對照組的餘命分位數比值的聯合信賴域。接著透過模擬方法研究兩方法所建立的聯合信賴域在覆蓋機率(coverage probability)及平均覆蓋面積(coverage area)的表現。最後，引用實際資料說明此方法的應用。

關鍵詞：設限存活資料；信賴域；最小距離統計量；多重比較；餘命分位數

第一章 背景及研究目的

在癌症醫學研究中，經常感興趣的問題是接受不同處理治療相較於對照組的病人，其存活時間或復發時間是否有差異，例如：Moertel 等學者(1991)和 Lin (1994)之大腸癌存活分析的化學藥物治療輔助治療臨床試驗，共 924 位受試者患有大腸癌，隨機分配 314 位受試者接受安慰劑治療，310 位受試者到藥物治療 Levamisole (Lev)組，304 位受試者到合併治療 Levamisole + 5-fluorouracil (Lev+5FU)組，想了解藥物治療和藥物加化療的合併治療相較於安慰劑的病人是否可以延長病人壽命。相較於得知存活機率或復發機率，醫生或病人更想了解接受完治療 t_0 時間後，病人還能活多久？還有多久時間會再復發？因此，針對右設限存活資料，根據平均餘命(residual lifetime)或餘命的分位數(quantile)的多對一比較就很重要。

已經有很多學者研究平均餘命函數(Chen 等(1983)，Berger 等(1988)，Oakes 和 Dasu (1990)，Maguluri 和 Zhang (1994)，Chen 等(2005))，但是在右設限資料之下對餘命中位數函數的討論較少。對於右設限存活資料，針對兩組存活中位數差異之信賴區間估計，已經有許多學者提出，例如：Wang 和 Hettmansperger (1991)整併兩個單樣本存活中位數之信賴區間，建立一個雙樣本存活中位數差異之信賴區間。但是此一方法必須估計存活函數的密度函數(density function)，為了避免此一麻煩，Su 和 Wei (1993)利用 Basawa 和 Koul (1988)所提出的最小距離統計量(minimum dispersion statistics)，求取兩組存活中位數差異或比值的信賴區間。推廣上述方法至餘命分配，Berger 等學者(1988)根

據 Fligner 和 Rust's (1982)的檢定方法，提出一個修正的分數型態檢定 (a score-type test)來檢定兩組餘命中位數的差異。但是此方法依然需要密度函數的估計，於是 Jeong 等學者(2008)推廣 Su 和 Wei (1993)的方法，提出無母數餘命中位數的信賴區間估計，但檢定過程還是需要分數函數(score function)的變異數估計。因此，Zhou 和 Jeong (2011)根據經驗概似比檢定(Empirical likelihood (EL) ratio test)，建立兩組餘命中位數或平均餘命比值的信賴區間，此方法結合概似比的優點，在檢定過程中不需要任何變異數的估計。

Zhao 和 Qin (2006)及 Qin 和 Zhao (2007)根據 EL 的方法估計平均餘命，然而，Qin 和 Zhao (2007)的方法需要估計干擾參數(nuisance parameters)。此外，在醫藥或醫學研究中，餘命中位數不一定存在。因此，本文考慮推廣 Jeong 等學者(2008)及 Zhou 和 Jeong (2011)的雙樣本比值的推論結果，至多組餘命分位數比值的聯合信賴域估計。

在本文第二章我們進行文獻回顧。第三章描述如何建構多組餘命分位數比值的聯合信賴域。第四章利用模擬研究所提方法在覆蓋機率及平均覆蓋面積的表現。第五章引用一筆實際資料做說明。最後，第六章為結論與建議。

第二章 文獻回顧

本章節將分別回顧 Jeong 等學者(2008)所提出的無母數方法及 Zhoua 和 Jong 等學者(2011)所提出的經驗概似比理論，如何建立雙樣本餘命中位數比值之信賴區間。

2.1 無母數推論 (Nonparametric Inference)

令 n_k , $k = 0, 1$, 為第 k 組的樣本數 , T_{ki} 與 C_{ki} , $i = 1, 2, \dots, n_k$, 分別第 k 組之下第 i 位病人的死亡時間與設限時間 , 且假設兩個隨機變數 T 與 C 是獨立的。右設限資料的觀察值為 (X_{ki}, δ_{ki}) , $i = 1, 2, \dots, n_k$, $k = 0, 1$, 其中 $X_{ki} = \min(T_{ki}, C_{ki})$, 設限指標 $\delta_{ki} = I(T_{ki} \leq C_{ki})$ 。假設 $S_k(t)$ 及 $H_k(t)$ 分別為第 k 組病人的存活函數和累積風險函數。令 $Y_{ki}(t) = I(X_{ki} \geq t)$ 和 $N_{ki}(t) = \delta_{ki} I(X_{ki} \leq t)$ 是涉險 (at-risk) 和死亡 (death) 過程 , 也定義 $Y_k(t) = \sum_{i=1}^{n_k} Y_{ki}(t)$ 及 $N_k(t) = \sum_{i=1}^{n_k} N_{ki}(t)$ 。

在存活資料中 , 假設存活超過 $t = t_0$ 時間點的餘命為 $(T - t_0 | T > t_0)$, 因此餘命的存活函數可寫為 $S(t|t_0) = S(t + t_0) / S(t_0)$, $t_0 \geq 0$ 。令 $\theta_{t_0,0}^{(k)}$, $k = 0, 1$ 是第 k 組在 t_0 時間真實的餘命中位數且令 $S(t|t_0) = 0.5$, 得到下式 :

$$\hat{u}_k(\theta_{t_0}^{(k)}) = \hat{S}_k(t_0 + \theta_{t_0}^{(k)}) - \frac{1}{2} \hat{S}_k(t_0)$$

其中 $\hat{S}_k(t)$ 是利用 (X_{ki}, δ_{ki}) 計算的 Kaplan-Meier (1958) 估計式。令

$\hat{u}_k(\theta_{t_0}^{(k)}) = 0$ 便可求得第 k 組餘命中位數估計值為

$$\hat{\theta}_{t_0}^{(k)} = \sup\{t : \hat{S}_k(t + t_0) - \frac{1}{2}\hat{S}_k(t_0) \geq 0\}, k = 0, 1.$$

因為 $\theta_{t_0}^{(k)}$ 為真實的餘命中位數，所以

$$u_k(\theta_{t_0}^{(k)}) = S_k(t_0 + \theta_{t_0}^{(k)}) - \frac{1}{2}S_k(t_0) = 0, k = 0, 1.$$

Jeong 等學者(2008)已證明 $\hat{\theta}_{t_0}^{(k)}$ 為 $\theta_{t_0}^{(k)}$ 的一致性估計。

Fleming 和 Harrington (1991) 的引理 3.2.1 已證明，在真實的 $\theta_{t_0,0}^{(k)}$ 時，

$$\begin{aligned} \hat{u}_k(\theta_{t_0,0}^{(k)}) &= -S_k(t_0 + \theta_{t_0,0}^{(k)}) \sum_{i=1}^{n_k} \int_0^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{Y_{ki}(v)} \\ &\quad + \frac{1}{2} \{S_k(t_0) \sum_{i=1}^{n_k} \int_0^{t_0} \frac{dM_{ki}(v)}{Y_k(v)}\} + o_P(n^{-\frac{1}{2}}), \end{aligned}$$

其中 $M_{ki}(t) = N_{ki}(t) - \int_0^t Y_{ki}(v) dH_k(v)$ 為一平賭過程 (martingale process) 且 $o_P(n^{-\frac{1}{2}})$ 機率收斂到 0。因為 $u_k(\theta_{t_0,0}^{(k)}) = 0$ ，利用極限 $y_k(t)$ 來取代 $Y_k(t)/n$ 得

$$\hat{u}_k(\theta_{t_0,0}^{(k)}) = \sum_{i=1}^{n_k} e_{ki}(\theta_{t_0,0}^{(k)}) + o_P(n^{-\frac{1}{2}})$$

其中

$$\begin{aligned} e_{ki}(\theta_{t_0,0}^{(k)}) &= -S_k(t_0 + \theta_{t_0,0}^{(k)}) \int_0^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{ny_k(v)} \\ &\quad + \frac{1}{2} S_k(t_0) \int_0^{t_0} \frac{dM_{ki}(v)}{ny_k(v)} \end{aligned} \quad (2.1)$$

因為 $E\{dM_{ki}(t)\} = E[E\{dM_{ki}(t)|\mathcal{F}_t\}] = 0$ ，其中 $\{\mathcal{F}_t : t \geq 0\}$ 為過濾函數(filtration)，所以 $e_{ki}(\theta_{t_0,0}^{(k)})$ 為獨立隨機變數且期望值為 0。由於 $S_k(t_0 + \theta_{t_0,0}^{(k)}) = \frac{1}{2}S_k(t_0)$ ，所以(2.1)便可簡化為

$$e_{ki}(\theta_{t_0,0}^{(k)}) = -\frac{1}{2}S_k(t_0) \int_{t_0}^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{ny_k(v)}.$$

利用中央極限定理(CLT)證明 $\hat{u}_k(\theta_{t_0,0}^{(k)})$ 近似常態分配，期望值為 0，變異數為 $\sigma_{t_0}^2(\theta_{t_0,0}^{(k)}) = \sum_{i=1}^{n_k} e_{ki}^2(\theta_{t_0,0}^{(k)})$ ，假設 $e_{ki}(\theta_{t_0,0}^{(k)})$ 的分布相同。將 $e_{ki}(\theta_{t_0,0}^{(k)})$ 用一致性估計 $\hat{e}_{ki}(\hat{\theta}_{t_0}^{(k)})$ 來取代，即

$$\hat{e}_{ki}(\hat{\theta}_{t_0}^{(k)}) = -\frac{1}{2}\hat{S}_k(t_0) \int_{t_0}^{t_0 + \hat{\theta}_{t_0}^{(k)}} \frac{d\hat{M}_{ki}(v)}{Y_k(v)}$$

便可得變異數估計 $\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(k)}) = \sum_{i=1}^{n_k} \hat{e}_{ki}^2(\hat{\theta}_{t_0}^{(k)})$ ，其中 $\hat{M}_{ki}(t) = N_{ki}(t) - \int_0^t Y_{ki}(v)d\hat{H}_k(v)$ 和 $\hat{H}_k(t) = \int_0^t Y_k^{-1}(v)dN_k(v)$ 是 Nelson-Aalen (Nelson,1972 和 Aalen,1978)的風險函數(hazard function)估計。

令 $\gamma_{t_0} = \theta_{t_0,0}^{(1)}/\theta_{t_0,0}^{(0)}$ 為在 t_0 時間點第 1 組相對於第 0 組餘命中位數的比值。針對虛無假設 $H_0 : \gamma_{t_0} = \gamma_{t_0,0}$ v.s 對立假設 $H_1 : \gamma_{t_0} \neq \gamma_{t_0,0}$ ， $r_{t_0,0}$ 是一個特定的值，當 $\gamma_{t_0,0} = 1$ 時，便可檢定兩組餘命中位數是否相等。在 H_0 之下， $\theta_{t_0,0}^{(1)} = \gamma_{t_0,0}\theta_{t_0,0}^{(0)}$ 。Jeong 等學者(2008)針對 γ_{t_0} 的推論，建議如下估計式：

$$V_{t_0}(r_{t_0,0}, \theta_{t_0}^{(0)}) = \frac{\hat{u}_0^2(\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(0)})} + \frac{\hat{u}_1^2(\theta_{t_0}^{(1)})}{\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(1)})} = \frac{\hat{u}_0^2(\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(0)})} + \frac{\hat{u}_1^2(r_{t_0,0}\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(1)})}.$$

因為 $\theta_{t_0}^{(0)}$ 為一個干擾參數，所以令

$$W_{t_0}(r_{t_0,0}) = \min_{\theta_{t_0}^{(0)}} V_{t_0}(r_{t_0,0}, \theta_{t_0}^{(0)})$$

為最小距離統計量(Basawa 和 Koul (1988))。根據 Su 和 Wei (1993)的推論已證明 $W_{t_0}(r_{t_0,0})$ 會服從自由度為 1 的卡方分配，記為 χ_1^2 。所以 γ_{t_0} 的 $100(1 - \alpha)\%$ 的信賴區間為

$$CI_N = \{\gamma_{t_0} : W_{t_0}(r_{t_0,0}) < \chi_{1-\alpha,1}^2\},$$

其中 $\chi_{1-\alpha,1}^2$ 為卡方分配中的 $100(1 - \alpha)\%$ 分位數。

2.2 經驗概似比推論 (Empirical Likelihood Approach)

Zhou 和 Jeong (2011)提出，在存活資料中，假設存活超過 $t = t_0$ 時間點的餘命中位數函數，為 $P(T_k - t_0 > \theta_{t_0}^{(k)} | T_k > t_0) = 0.5$, $k = 0,1$ ，將餘命中位數定義為

$$\frac{1 - F_k(t_0 + \theta_{t_0}^{(k)})}{1 - F_k(t_0)} = 0.5 \quad (2.2)$$

其中， $F_k(\cdot)$ 為第 k 組死亡時間的累積分配函數， $\theta_{t_0}^{(k)}$ 是第 k 組在 t_0 時間真實的餘命中位數， $k = 0,1$ 。由(2.2)可得

$$0.5 = F_k(t_0 + \theta_{t_0}^{(k)}) - 0.5F_k(t_0).$$

定義一個函數為 $g_b(t)$

$$g_b(t) = I_{[t \leq (t_0+b)]} - 0.5I_{[t \leq t_0]} - 0.5,$$

因此，檢定假設 $H_0 : \theta_{t_0}^{(k)} = b$ 等同檢定

$$H_0 : \int_0^\infty g_b(t) dF_k(t) = 0.$$

令 p_{ki} 是一個觀測值 X_{ki} 的機率，經驗概似比函數表示為

$$EL_k = \prod_{i=1}^{n_k} \{p_{ki}\}^{\delta_{ki}} \left\{ \sum_{X_{kj} > X_{ki}} p_{kj} \right\}^{1-\delta_{ki}}, k = 0, 1.$$

將 EL_k 最大化滿足條件 $p_{ki} \geq 0$ 和 $\sum p_{ki} = 1$ ，其中 p_{ki} 是利用 (X_{ki}, δ_{ki}) 來計算 Kaplan-Meier 估計值 (Owen (2001), p. 142)，得到的最大值記為 $EL_k(KM)$ 。

為建構概似比檢定，最大概似解需在 $H_0: \theta_{t_0}^{(k)} = b$ 假設下求解，也就是須滿足以下限制式

$$\sum_{i=1}^{n_k} g(X_{ki}) p_{ki} = 0,$$

其中 $g(t)$ 是一個給定的函數， $0 < \text{Var } g(T) < \infty$ 。因為 $\sum g(X_{ki}) w_{ki}$ 的變異數不一定存在，其中 w_{ki} 為 Kaplan-Meier 估計值，所以需要額外加限制式

$$\int \frac{g^2(t)}{1 - G_k(t-)} dF_k(t) < \infty,$$

其中 $G_k(\cdot)$ 是 C_{ki} 設限時間的分配函數。

根據 EL 定理，在虛無假設 $H_0: \theta_{t_0}^{(k)} = \mathbb{E}_k g(T)$ 。Zhou 和 Jeong (2011) 針對 $\theta_{t_0}^{(k)}$ 的推論，建議如下式：

$$W_k \left(\theta_{t_0}^{(k)} \right) = -2 \log \frac{\max EL_k}{EL_k(KM)} \rightarrow \chi_1^2 \quad \text{as } n \rightarrow \infty$$

根據 Bathke 等學者 (2009) 的推論已證明 $W_k \left(\theta_{t_0}^{(k)} \right)$ 會服從自由度為 1

的卡方分配，記為 χ_1^2 。

若想建立雙樣本餘命中位數比值檢定，針對虛無假設 $H_0 : \gamma_{t_0} = \gamma_{t_0,0}$

v.s 對立假設 $H_1 : \gamma_{t_0} \neq \gamma_{t_0,0}$ ， $r_{t_0,0}$ 是一個特定的值，Zhou 和 Jeong

(2011)建議如下估計式，估計 γ_{t_0} ：

$$Q(r_{t_0,0}) = \min_{\theta_{t_0}^{(0)}} \left\{ W_0 \left(\theta_{t_0}^{(0)} \right) + W_1 \left(r_{t_0,0} \theta_{t_0}^{(0)} \right) \right\},$$

且已證明在 $\gamma_{t_0,0} = 1$ 時， $Q(r_{t_0,0})$ 會服從自由度為 1 的卡方分配。所以

γ_{t_0} 的 $100(1 - \alpha)\%$ 的信賴區間為

$$CI_E = \left\{ \gamma_{t_0} : Q(r_{t_0,0}) < \chi_{1-\alpha,1}^2 \right\},$$

其中 $\chi_{1-\alpha,1}^2$ 為卡方分配中的 $100(1 - \alpha)$ 分位數。

第三章 多組餘命分位數比值的聯合信賴域

本章節考慮推廣 Jeong 等學者(2008)提出的無母數方法和 Zhou 和 Jeong (2011)提出的經驗概似比值檢定法，建構多組餘命分位數比值的聯合信賴域。

3.1 無母數聯合信賴域

令 n_k 為第 k 組病人的樣本數， $k = 0, 1, \dots, m$ ，其中 $k = 0$ 時為對照組，其餘的是處理組。令 (T_{ki}, C_{ki}) ， $i = 1, 2, \dots, n_k$ ，分別第 k 組第 i 位病人的死亡時間與設限時間，且假設兩個隨機變數 T 與 C 是獨立的。右設限資料的觀察值為 (X_{ki}, δ_{ki}) ， $i = 1, 2, \dots, n_k$ ， $k = 0, 1, \dots, m$ ，其中 $X_{ki} = \min(T_{ki}, C_{ki})$ ，設限指標 $\delta_{ki} = I(T_{ki} \leq C_{ki})$ 。假設 $S_k(t)$ 及 $H_k(t)$ 分別為第 k 組病人的存活函數和累積風險函數。令 $Y_{ki}(t) = I(X_{ki} \geq t)$ 和 $N_{ki}(t) = \delta_{ki} I(X_{ki} \leq t)$ 是涉險(at-risk)和死亡(death)過程，也定義 $Y_k(t) = \sum_{i=1}^{n_k} Y_{ki}(t)$ 及 $N_k(t) = \sum_{i=1}^{n_k} N_{ki}(t)$ 。

在存活資料中，假設存活超過 $t = t_0$ 時間點的餘命為 $(T - t_0 | T > t_0)$ ，因此餘命的存活函數可寫為 $S(t|t_0) = S(t + t_0) / S(t_0)$ ， $t_0 \geq 0$ 。令 $\theta_{t_0,0}^{(k)}$ ， $k = 0, 1, \dots, m$ ，是第 k 組在 t_0 時間真實第 $100q$ 個餘命分位數且令 $S(t|t_0) = (1 - q)$ ，得到下式：

$$\hat{u}_k \left(\theta_{t_0}^{(k)} \right) = \hat{S}_k \left(t_0 + \theta_{t_0}^{(k)} \right) - (1 - q) \hat{S}_k(t_0)$$

其中 $\hat{S}_k(t)$ 是利用 (X_{ki}, δ_{ki}) 計算的 Kaplan-Meier (1958) 估計式。令

$\hat{u}_k(\theta_{t_0}^{(k)}) = 0$ 便可求得第 k 組餘命分位數估計值為

$$\hat{\theta}_{t_0}^{(k)} = \sup\{t : \hat{S}_k(t + t_0) - (1 - q)\hat{S}_k(t_0) \geq 0\}, k = 0, 1, \dots, m.$$

因為 $\theta_{t_0}^{(k)}$ 為真實的餘命分位數，所以

$$u_k(\theta_{t_0}^{(k)}) = S_k(t_0 + \theta_{t_0}^{(k)}) - (1 - q)S_k(t_0) = 0, k = 0, 1, \dots, m.$$

Jeong 等學者(2008)已證明 $\hat{\theta}_{t_0}^{(k)}$ 為 $\theta_{t_0}^{(k)}$ 的一致性估計。

Fleming 和 Harrington (1991)的引理 3.2.1 已證明，在真實的 $\theta_{t_0,0}^{(k)}$ 時，

$$\begin{aligned} \hat{u}_k(\theta_{t_0,0}^{(k)}) &= -S_k(t_0 + \theta_{t_0,0}^{(k)}) \sum_{i=1}^{n_k} \int_0^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{Y_{ki}(v)} \\ &+ (1 - q) \{S_k(t_0) \sum_{i=1}^{n_k} \int_0^{t_0} \frac{dM_{ki}(v)}{Y_k(v)}\} + o_P(n^{-\frac{1}{2}}), \end{aligned}$$

其中 $M_{ki}(t) = N_{ki}(t) - \int_0^t Y_{ki}(v) dH_k(v)$, $k = 0, 1, \dots, m$ 為一平賭過程 (martingale process) 且 $o_P(n^{-\frac{1}{2}})$ 機率收斂到 0。因為 $u_k(\theta_{t_0,0}^{(k)}) = 0$ ，利用極限 $y_k(t)$ 來取代 $Y_k(t)/n$ 得

$$\hat{u}_k(\theta_{t_0,0}^{(k)}) = \sum_{i=1}^{n_k} e_{ki}(\theta_{t_0,0}^{(k)}) + o_P(n^{-\frac{1}{2}}),$$

其中

$$\begin{aligned} e_{ki}(\theta_{t_0,0}^{(k)}) &= -S_k(t_0 + \theta_{t_0,0}^{(k)}) \int_0^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{ny_k(v)} \\ &+ (1 - q)S_k(t_0) \int_0^{t_0} \frac{dM_{ki}(v)}{ny_k(v)}, \end{aligned} \quad (3.1)$$

因為 $E\{dM_{ki}(t)\} = E[E\{dM_{ki}(t)|\mathcal{F}_t\}] = 0$ ，其中 $\{\mathcal{F}_t : t \geq 0\}$ 為過濾函

數(filtration)，所以 $e_{ki}(\theta_{t_0,0}^{(k)})$, $k = 0, 1, \dots, m$ ，為獨立隨機變數且期望值為 0。由於 $S_k(t_0 + \theta_{t_0,0}^{(k)}) = (1 - q)S_k(t_0)$ ，所以(3.1)便可簡化為

$$e_{ki}(\theta_{t_0,0}^{(k)}) = -(1 - q)S_k(t_0) \int_{t_0}^{t_0 + \theta_{t_0,0}^{(k)}} \frac{dM_{ki}(v)}{ny_k(v)}, k = 0, 1, \dots, m.$$

利用中央極限定理(CLT)證明 $\hat{u}_k(\theta_{t_0,0}^{(k)})$ 近似常態分配，期望值為 0，變異數為 $\sigma_{t_0}^2(\theta_{t_0,0}^{(k)}) = \sum_{i=1}^{n_k} e_{ki}^2(\theta_{t_0,0}^{(k)})$ ，假設 $e_{ki}(\theta_{t_0,0}^{(k)})$, $k = 0, 1, \dots, m$ ， k 個分布相同。將 $e_{ki}(\theta_{t_0,0}^{(k)})$ 用一致性估計 $\hat{e}_{ki}(\hat{\theta}_{t_0}^{(k)})$ 來取代，即

$$\hat{e}_{ki}(\hat{\theta}_{t_0}^{(k)}) = -(1 - q)\hat{S}_k(t_0) \int_{t_0}^{t_0 + \hat{\theta}_{t_0}^{(k)}} \frac{d\hat{M}_{ki}(v)}{Y_k(v)}, k = 0, 1, \dots, m.$$

便可得變異數估計 $\hat{\sigma}_{t_0}^2(\hat{\theta}_{t_0}^{(k)}) = \sum_{i=1}^{n_k} \hat{e}_{ki}^2(\hat{\theta}_{t_0}^{(k)})$ ，其中 $\hat{M}_{ki}(t) = N_{ki}(t) - \int_0^t Y_{ki}(v)d\hat{H}_k(v)$ 和 $\hat{H}_k(t) = \int_0^t Y_k^{-1}(v)dN_k(v)$ 是 Nelson-Aalen (Nelson, 1972 和 Aalen, 1978)的風險函數(hazard function)估計。

令 $\gamma_{t_0}^{(k)} = \theta_{t_0,0}^{(k)}/\theta_{t_0,0}^{(0)}$, $k = 1, \dots, m$ ，為在 t_0 時間點第 k 組相對於第 0 組餘命分位數的比值。我們感興趣的虛無假設為

$$H_0 : \gamma_{t_0}^{(1)} = \gamma_{t_0,0}^{(1)}, \gamma_{t_0}^{(2)} = \gamma_{t_0,0}^{(2)}, \dots, \gamma_{t_0}^{(m)} = \gamma_{t_0,0}^{(m)}, \quad (3.2)$$

$\gamma_{t_0,0}^{(k)}$, $k = 1, \dots, m$ ，是給定的值，當 $\gamma_{t_0,0}^{(k)} = 1$, $k = 1, \dots, m$ 時，便可檢定

多組餘命分位數是否相等。在 H_0 之下， $\theta_{t_0,0}^{(k)} = \gamma_{t_0,0}^{(k)}\theta_{t_0,0}^{(0)}$, $k = 1, \dots, m$ 。

根據 Jeong 等學者(2008)針對 $\gamma_{t_0}^{(k)}$ 的推論，因此考慮檢定統計量為下式：

$$V_{t_0}(\theta_{t_0}^{(0)}, \gamma_{t_0,0}) = \frac{\hat{u}_0^2(\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(0)})} + \sum_{k=1}^m \frac{\hat{u}_k^2(\theta_{t_0}^{(k)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(k)})}$$

$$= \frac{\hat{u}_0^2(\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(0)})} + \sum_{k=1}^m \frac{\hat{u}_k^2(\gamma_{t_0,0}^{(k)}\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(k)})},$$

其中 $\boldsymbol{\gamma}_{t_0,0} = (\gamma_{t_0,0}^{(1)}, \gamma_{t_0,0}^{(2)}, \dots, \gamma_{t_0,0}^{(m)})$ 。因為 $\theta_{t_0}^{(0)}$ 為一個干擾參數，所以令最小距離統計量為

$$W_{t_0}(\boldsymbol{\gamma}_{t_0,0}) = \min_{\theta_{t_0}^{(0)}} V_{t_0}(\theta_{t_0}^{(0)}, \boldsymbol{\gamma}_{t_0,0}).$$

於附錄 A-1 我們推得 $W_{t_0}(\boldsymbol{\gamma}_{t_0,0})$ 會服從自由度為 m 的卡方分配，記為 χ_m^2 。所以 $\boldsymbol{\gamma}_{t_0} = (\gamma_{t_0}^{(1)}, \gamma_{t_0}^{(2)}, \dots, \gamma_{t_0}^{(m)})$ 的 $100(1 - \alpha)\%$ 的聯合信賴域為

$$SCI_N = \{\boldsymbol{\gamma}_{t_0} : W_{t_0}(\boldsymbol{\gamma}_{t_0}) < \chi_{1-\alpha, m}^2\}, \quad (3.3)$$

其中 $\chi_{1-\alpha, m}^2$ 為卡方分配 χ_m^2 的 $100(1 - \alpha)\%$ 分位數。

3.2 經驗概似比聯合信賴域

Zhou 和 Jeong (2011) 提出，在存活資料中，假設存活超過 $t = t_0$ 時間點第 $100q$ 個的餘命分位數函數，為 $P(T_k - t_0 > \theta_{t_0}^{(k)} | T_k > t_0) = (1 - q)$, $k = 0, 1, \dots, m$ ，將第 $100q$ 個餘命分位數定義為

$$\frac{1 - F_k(t_0 + \theta_{t_0}^{(k)})}{1 - F_k(t_0)} = (1 - q) \quad (3.4)$$

其中 $F_k(\cdot)$ 為第 k 組死亡時間的累積分配函數， $\theta_{t_0}^{(k)}$ 是第 k 組在 t_0 時間真實的餘命分位數， $k = 0, 1, \dots, m$ 。由 (3.4) 可得

$$q = F_k(t_0 + \theta_{t_0}^{(k)}) - (1 - q)F_k(t_0).$$

定義一個函數為 $g_b(t)$

$$g_b(t) = I_{[t \leq (t_0 + b)]} - (1 - q)I_{[t \leq t_0]} - q$$

因此，檢定假設 $H_0 : \theta_{t_0}^{(k)} = b, k = 0, 1, \dots, m$ ，等同檢定

$$H_0 : \int_0^\infty g_b(t) dF_k(t) = 0.$$

令 p_{ki} 是一個觀測值 X_{ki} 的機率，經驗概似比函數表示為

$$EL_k = \prod_{i=1}^{n_k} \{p_{ki}\}^{\delta_{ki}} \left\{ \sum_{X_{kj} > X_{ki}} p_{kj} \right\}^{1 - \delta_{ki}}, \quad k = 0, 1, \dots, m.$$

將 EL_k 最大化滿足條件 $p_{ki} \geq 0$ 和 $\sum p_{ki} = 1$ ，其中 p_{ki} 是利用 (X_{ki}, δ_{ki}) 來計算 Kaplan-Meier 估計值 (Owen (2001), p. 142)，得到的最大值記為 $EL_k(KM)$ 。

為建構概似比檢定，最大概似解需在 $H_0 : \theta_{t_0}^{(k)} = b$ 假設下求解，也就是須滿足以下限制式

$$\sum_{i=1}^{n_k} g_b(X_{ki}) p_{ki} = 0,$$

其中 $g(t)$ 是一個給定的函數， $0 < \text{Var } g(T) < \infty$ 。因為 $\sum g(X_{ki}) w_{ki}$ 的變異數不一定存在，其中 w_{ki} 為 Kaplan-Meier 估計值，所以需要額外加限制式

$$\int \frac{g^2(t)}{1 - G_k(t-)} dF_k(t) < \infty,$$

其中 $G_k(\cdot)$ 是 C_{ki} 設限時間的分配函數。

根據 EL 定理，在虛無假設 $H_0 : \theta_{t_0}^{(k)} = \mathbb{E}_k g(T), k = 0, 1, \dots, m$ ，

之下，Zhou 和 Jeong (2011)針對 $\theta_{t_0}^{(k)}$ 的推論，建議如下式：

$$W_k \left(\theta_{t_0}^{(k)} \right) = -2 \log \frac{\max EL_k}{EL_k(KM)} \rightarrow \chi_1^2 \quad \text{as } n \rightarrow \infty$$

根據 Bathke 等學者(2009)的推論已證明 $W_k \left(\theta_{t_0}^{(k)} \right)$ 會服從自由度為 1 的卡方分配。

若想建立多組樣本餘命分位數比值檢定，我們感興趣的虛無假設為(3.2)式，根據 Zhou 和 Jeong (2011)針對 $\gamma_{t_0}^{(k)}$ 的推論，因此考慮檢定統計量為：

$$Q(\boldsymbol{\gamma}_{t_0,0}) = \min_{\theta_{t_0}^{(0)}} \left\{ W_0 \left(\theta_{t_0}^{(0)} \right) + W_1 \left(\gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right) + W_2 \left(\gamma_{t_0,0}^{(2)} \theta_{t_0}^{(0)} \right) + \dots + W_m \left(\gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right) \right\},$$

其中 $\boldsymbol{\gamma}_{t_0,0} = (\gamma_{t_0,0}^{(1)}, \gamma_{t_0,0}^{(2)}, \dots, \gamma_{t_0,0}^{(m)})$ 。於附錄 A-2 我們推得 $Q(\boldsymbol{\gamma}_{t_0,0})$ 在 $\boldsymbol{\gamma}_{t_0,0} = 1$ 時， $Q(\boldsymbol{\gamma}_{t_0,0})$ 會服從自由度為 m 的卡方分配。因此 $\boldsymbol{\gamma}_{t_0} = (\gamma_{t_0}^{(1)}, \gamma_{t_0}^{(2)}, \dots, \gamma_{t_0}^{(m)})$ 的 $100(1 - \alpha)\%$ 的聯合信賴域為

$$SCI_E = \left\{ \boldsymbol{\gamma}_{t_0} : Q(\boldsymbol{\gamma}_{t_0}) < \chi_{1-\alpha, m}^2 \right\}, \quad (3.5)$$

其中 $\chi_{1-\alpha, m}^2$ 為卡方分配 χ_m^2 的 $100(1 - \alpha)\%$ 分位數。

第四章 模擬研究

4.1 模擬方法

本章利用模擬方法比較 SCI_N (3.3)和 SCI_E (3.5)，其多組樣本餘命中位數比值的 90% 和 95% 聯合信賴域之覆蓋機率及平均覆蓋面積的表現。

本文考慮 $m = 2$ 個處理組與一個對照組在不同的樣本數 $n_0 = n_1 = n_2 = n = 50, 100, 200$ 及 500 ，以及存活超過時間點分別為 $t_0 = 1.5$ 和 2 之下。模擬存活資料為服從不同尺度參數(scale parameter)和不同形狀參數(shape parameter)的 Weibull 分布，其存活函數為

$$S_k(t) = \exp(-\lambda_k t^{\eta_k}), k = 0, 1, 2,$$

其中 $\lambda_k > 0, \eta_k > 0$ 分別為 Weibull 分布的尺度參數跟形狀參數。考慮 $\eta_0 = \eta_1 = \eta_2 = \eta = 1.2$ 或 1.5 ，及 $(\lambda_0, \lambda_1, \lambda_2) = (0.6, 0.6, 0.6), (0.6, 0.4, 0.4), (0.6, 0.4, 0.25), (0.6, 0.6, 0.25)$ ，分別對應於四種型態的假設：(1) $S_0 = S_1 = S_2$ (2) $S_0 < S_1 = S_2$ (3) $S_0 < S_1 < S_2$ (4) $S_0 = S_1 < S_2$ 。各種參數組合的存活函數圖形繪於圖一和圖二。

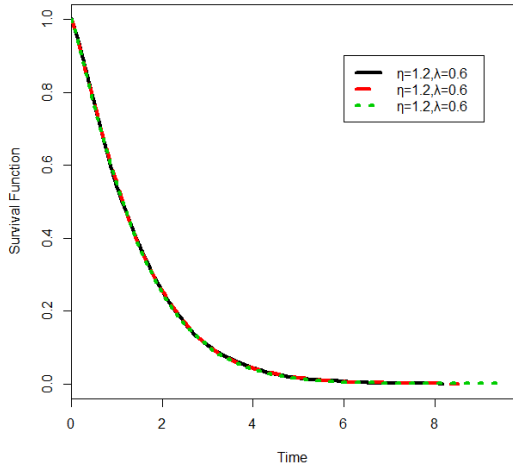
模擬次數為 5000 次，設限分布介於 1.5 到 8 的均勻分布，記為 $U(1.5, 8)$ ，此時的設限機率大約介於 0.07~0.25 之間。利用所提方法分別求出各處理組相對於對照組餘命中位數比值的 90% 和 95% 聯合信賴域，然後計數 5000 個聯合信賴域包含真正餘命中位數比值的比率，藉以計算聯合信賴域之覆蓋機率及平均覆蓋面積，其覆蓋面積公式如下：

$$\frac{(CI_{1U} - CI_{1L})(CI_{2U} - CI_{2L})}{4} * \pi$$

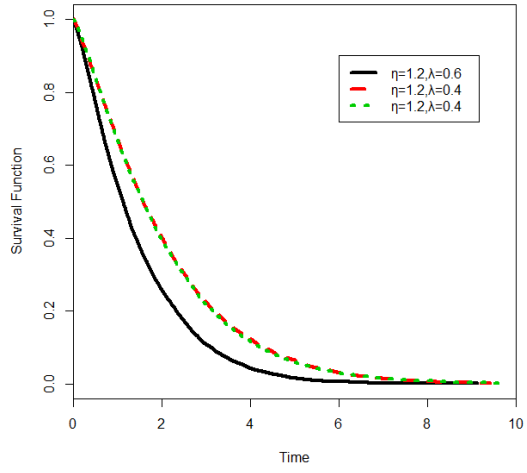
其中， (CI_{iU}, CI_{iL}) 為 $\gamma_{t_0}^{(i)}, i = 1, 2$ 的信賴區間估計， π 為圓周率

(3.14159...)，但此公式只能適用於 $m = 2$ ，若 $m > 2$ 則我們建議可用每組的處理組相對於對照組餘命分位數比值的聯合信賴域總和距離 $\sum_{i=1}^m (CI_{iU} - CI_{iL}), i = 1, 2, \dots, m$ 來比較。模擬結果列於表 1~4。

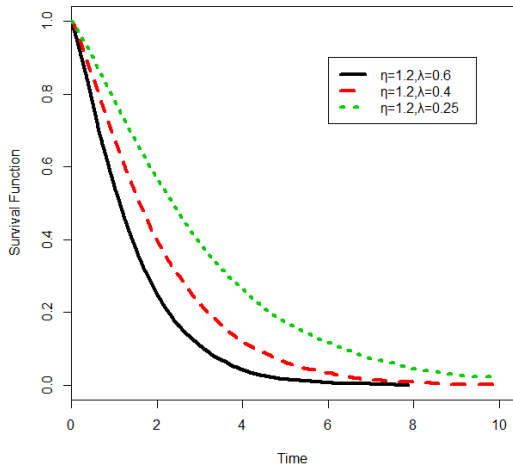
(1)



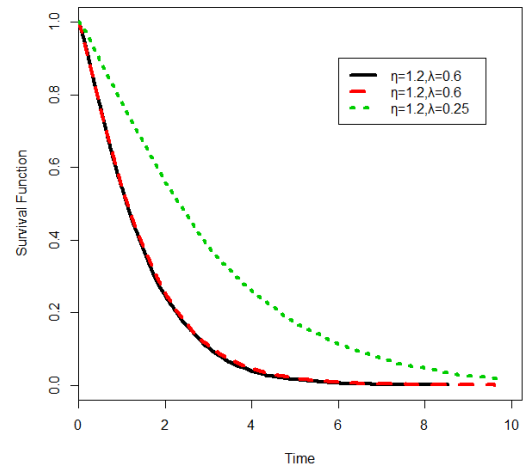
(2)



(3)

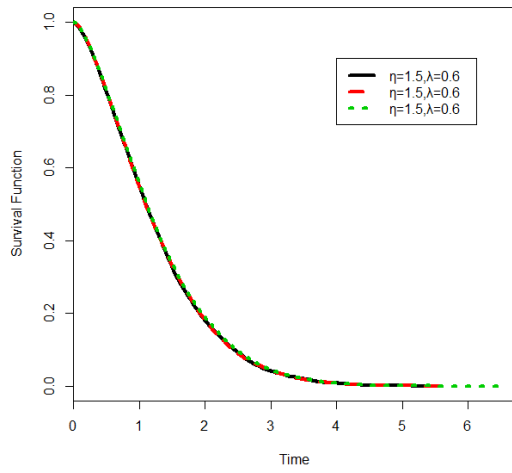


(4)

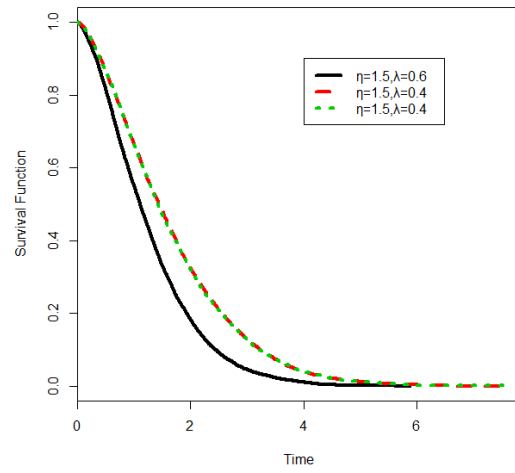


圖一、不同之 λ 及 $\eta = 1.2$ ，在四種型態下的存活函數曲線：(1) $S_0 = S_1 = S_2$ (2) $S_0 < S_1 = S_2$ (3) $S_0 < S_1 < S_2$ (4) $S_0 = S_1 < S_2$ 。

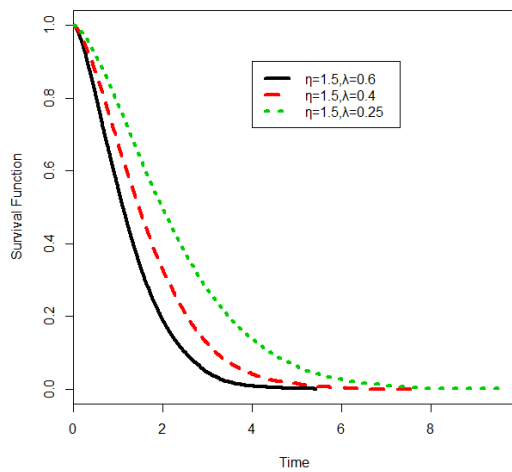
(1)



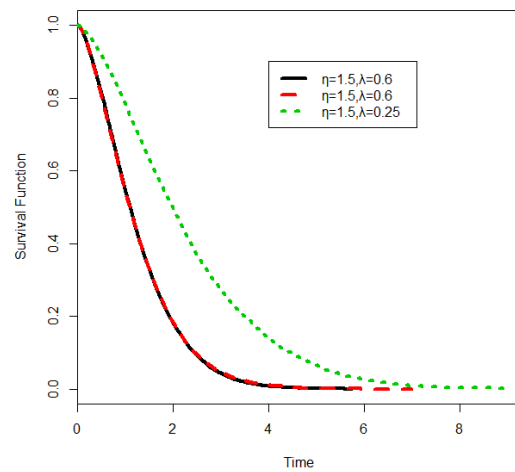
(2)



(3)



(4)



圖二、不同之 λ 及 $\eta = 1.5$ ，在四種型態下的存活函數曲線：(1) $S_0 = S_1 = S_2$ (2) $S_0 < S_1 = S_2$ (3) $S_0 < S_1 < S_2$ (4) $S_0 = S_1 < S_2$ 。

4.2 模擬結果

從表 1 及表 2 的結果得知，當樣本數小時，不管 $\eta = 1.2$ 或 1.5 ， SCI_N 和 SCI_E 的 90% 和 95% 覆蓋機率皆會略高於所設定的信心水準；但當樣本數提高時，不管 $\eta = 1.2$ 或 1.5 ， SCI_N 和 SCI_E 的 90% 和 95% 覆蓋機率皆能維持其信心水準，但是 SCI_E 趨近於設定的信心水準的速度比 SCI_N 快。

從表 3 及表 4 的結果得知，當 $t_0 = 1.5$ 及樣本數小時， SCI_N 的 90% 和 95% 平均覆蓋面積比 SCI_E 大，除了在此假設型態 $S_0 < S_1 < S_2$ 下， SCI_N 的 90% 和 95% 平均覆蓋面積比 SCI_E 小；但當樣本數提高時，不管 $\eta = 1.2$ 或 1.5 ， SCI_N 和 SCI_E 的 90% 和 95% 平均覆蓋面積皆有明顯縮減。然而，當 $t_0 = 2$ 時在不同型態假設下， SCI_N 的 90% 和 95% 平均覆蓋面積皆比 SCI_E 大。在此假設型態 $S_0 = S_1 = S_2$ 的信賴寬度最寬； $S_0 < S_1 < S_2$ 信賴寬度最窄。

表 1、當 $t_0 = 1.5$ 時，設限分布介為 $U(1.5,8)$ ，不同之 η 及 γ 參數組合及樣本數，之 90% 和 95% 信賴區間覆蓋機率

n	η	λ			90% 覆蓋機率		95% 覆蓋機率	
					SCI_N	SCI_E	SCI_N	SCI_E
50	1.2	0.6	0.6	0.6	0.935	0.932	0.980	0.979
		0.6	0.4	0.4	0.943	0.939	0.978	0.976
		0.6	0.4	0.25	0.939	0.935	0.979	0.977
		0.6	0.6	0.25	0.934	0.932	0.976	0.974
	1.5	0.6	0.6	0.6	0.934	0.931	0.975	0.975
		0.6	0.4	0.4	0.936	0.933	0.978	0.976
		0.6	0.4	0.25	0.935	0.934	0.976	0.976
		0.6	0.6	0.25	0.933	0.932	0.974	0.973
100	1.2	0.6	0.6	0.6	0.933	0.927	0.972	0.967
		0.6	0.4	0.4	0.935	0.932	0.973	0.969
		0.6	0.4	0.25	0.932	0.928	0.971	0.970
		0.6	0.6	0.25	0.931	0.924	0.970	0.965
	1.5	0.6	0.6	0.6	0.929	0.923	0.973	0.967
		0.6	0.4	0.4	0.928	0.921	0.974	0.969
		0.6	0.4	0.25	0.932	0.924	0.971	0.962
		0.6	0.6	0.25	0.930	0.925	0.967	0.960
200	1.2	0.6	0.6	0.6	0.927	0.922	0.968	0.961
		0.6	0.4	0.4	0.926	0.921	0.964	0.959
		0.6	0.4	0.25	0.921	0.915	0.968	0.961
		0.6	0.6	0.25	0.914	0.913	0.963	0.954
	1.5	0.6	0.6	0.6	0.923	0.917	0.965	0.956
		0.6	0.4	0.4	0.915	0.916	0.967	0.957
		0.6	0.4	0.25	0.918	0.911	0.962	0.952
		0.6	0.6	0.25	0.923	0.917	0.959	0.951
500	1.2	0.6	0.6	0.6	0.901	0.892	0.957	0.954
		0.6	0.4	0.4	0.910	0.903	0.954	0.952
		0.6	0.4	0.25	0.909	0.905	0.952	0.949
		0.6	0.6	0.25	0.901	0.899	0.951	0.947
	1.5	0.6	0.6	0.6	0.891	0.890	0.953	0.950
		0.6	0.4	0.4	0.904	0.901	0.948	0.947
		0.6	0.4	0.25	0.904	0.902	0.949	0.945
		0.6	0.6	0.25	0.903	0.900	0.947	0.944

表 2、當 $t_0 = 2$ 時，設限分布介為 $U(1.5,8)$ ，不同之 η 及 γ 參數組合及樣本數，之 90% 和 95% 信賴區間覆蓋機率

n	η	λ			90% 覆蓋機率		95% 覆蓋機率	
					SCI_N	SCI_E	SCI_N	SCI_E
50	1.2	0.6	0.6	0.6	0.933	0.930	0.979	0.974
		0.6	0.4	0.4	0.945	0.934	0.976	0.973
		0.6	0.4	0.25	0.938	0.930	0.974	0.971
		0.6	0.6	0.25	0.934	0.929	0.972	0.970
	1.5	0.6	0.6	0.6	0.932	0.930	0.973	0.973
		0.6	0.4	0.4	0.935	0.934	0.974	0.971
		0.6	0.4	0.25	0.937	0.933	0.972	0.970
		0.6	0.6	0.25	0.934	0.930	0.973	0.971
100	1.2	0.6	0.6	0.6	0.933	0.929	0.971	0.969
		0.6	0.4	0.4	0.933	0.932	0.972	0.968
		0.6	0.4	0.25	0.931	0.928	0.973	0.971
		0.6	0.6	0.25	0.934	0.932	0.976	0.964
	1.5	0.6	0.6	0.6	0.930	0.924	0.972	0.962
		0.6	0.4	0.4	0.929	0.922	0.976	0.962
		0.6	0.4	0.25	0.931	0.922	0.969	0.960
		0.6	0.6	0.25	0.929	0.923	0.963	0.959
200	1.2	0.6	0.6	0.6	0.924	0.919	0.964	0.960
		0.6	0.4	0.4	0.924	0.915	0.962	0.957
		0.6	0.4	0.25	0.920	0.911	0.967	0.960
		0.6	0.6	0.25	0.912	0.909	0.962	0.953
	1.5	0.6	0.6	0.6	0.913	0.907	0.964	0.952
		0.6	0.4	0.4	0.914	0.907	0.966	0.958
		0.6	0.4	0.25	0.912	0.909	0.959	0.951
		0.6	0.6	0.25	0.927	0.919	0.958	0.952
500	1.2	0.6	0.6	0.6	0.890	0.890	0.952	0.950
		0.6	0.4	0.4	0.904	0.903	0.953	0.949
		0.6	0.4	0.25	0.903	0.896	0.950	0.943
		0.6	0.6	0.25	0.892	0.890	0.949	0.945
	1.5	0.6	0.6	0.6	0.890	0.890	0.953	0.952
		0.6	0.4	0.4	0.901	0.893	0.947	0.943
		0.6	0.4	0.25	0.903	0.902	0.946	0.941
		0.6	0.6	0.25	0.905	0.900	0.947	0.945

表 3、當 $t_0 = 1.5$ 時，設限分布介為 $U(1.5,8)$ ，不同之 η 及 γ 參數組合及樣本數，之 90% 和 95%信賴區間平均覆蓋面積之估計

n	η	λ			90%平均覆蓋面積		95%平均覆蓋面積	
					SCI_N	SCI_E	SCI_N	SCI_E
50	1.2	0.6	0.6	0.6	4.513	1.293	3.612	0.946
		0.6	0.4	0.4	1.035	0.961	0.951	0.732
		0.6	0.4	0.25	0.784	0.863	0.654	0.694
		0.6	0.6	0.25	4.131	1.563	3.166	0.891
	1.5	0.6	0.6	0.6	2.036	0.946	1.942	0.874
		0.6	0.4	0.4	0.943	0.841	0.843	0.645
		0.6	0.4	0.25	0.721	0.751	0.622	0.637
		0.6	0.6	0.25	2.009	0.913	1.459	0.761
100	1.2	0.6	0.6	0.6	1.956	0.963	1.064	0.894
		0.6	0.4	0.4	0.974	0.746	0.841	0.578
		0.6	0.4	0.25	0.674	0.691	0.532	0.537
		0.6	0.6	0.25	1.045	0.923	0.941	0.717
	1.5	0.6	0.6	0.6	0.946	0.892	0.713	0.674
		0.6	0.4	0.4	0.842	0.763	0.611	0.498
		0.6	0.4	0.25	0.643	0.712	0.465	0.479
		0.6	0.6	0.25	0.892	0.792	0.622	0.562
200	1.2	0.6	0.6	0.6	0.943	0.713	0.844	0.623
		0.6	0.4	0.4	0.732	0.546	0.513	0.431
		0.6	0.4	0.25	0.543	0.463	0.362	0.361
		0.6	0.6	0.25	0.863	0.632	0.794	0.613
	1.5	0.6	0.6	0.6	0.845	0.761	0.746	0.595
		0.6	0.4	0.4	0.635	0.456	0.531	0.323
		0.6	0.4	0.25	0.547	0.312	0.316	0.232
		0.6	0.6	0.25	0.674	0.625	0.665	0.531
500	1.2	0.6	0.6	0.6	0.226	0.188	0.186	0.092
		0.6	0.4	0.4	0.094	0.062	0.084	0.045
		0.6	0.4	0.25	0.075	0.041	0.056	0.033
		0.6	0.6	0.25	0.189	0.095	0.087	0.065
	1.5	0.6	0.6	0.6	0.197	0.146	0.156	0.045
		0.6	0.4	0.4	0.089	0.066	0.056	0.021
		0.6	0.4	0.25	0.068	0.045	0.034	0.016
		0.6	0.6	0.25	0.145	0.084	0.091	0.035

表 4、當 $t_0 = 2$ 時，設限分布介為 $U(1.5,8)$ ，不同之 η 及 γ 參數組合及樣本數，之 90% 和 95% 信賴區間平均覆蓋面積之估計

n	η	λ			90% 平均覆蓋面積		95% 平均覆蓋面積	
					SCI_N	SCI_E	SCI_N	SCI_E
50	1.2	0.6	0.6	0.6	4.746	1.064	2.912	0.960
		0.6	0.4	0.4	1.123	0.923	0.946	0.746
		0.6	0.4	0.25	0.795	0.723	0.685	0.672
		0.6	0.6	0.25	4.132	1.156	2.163	0.889
	1.5	0.6	0.6	0.6	2.133	0.923	1.065	0.871
		0.6	0.4	0.4	0.943	0.842	0.891	0.684
		0.6	0.4	0.25	0.723	0.531	0.684	0.561
		0.6	0.6	0.25	1.952	0.861	1.092	0.812
100	1.2	0.6	0.6	0.6	1.832	0.952	1.048	0.897
		0.6	0.4	0.4	0.896	0.641	0.766	0.513
		0.6	0.4	0.25	0.662	0.513	0.513	0.431
		0.6	0.6	0.25	0.984	0.841	0.896	0.691
	1.5	0.6	0.6	0.6	0.923	0.813	0.742	0.615
		0.6	0.4	0.4	0.813	0.623	0.561	0.456
		0.6	0.4	0.25	0.623	0.423	0.532	0.415
		0.6	0.6	0.25	0.852	0.752	0.701	0.591
200	1.2	0.6	0.6	0.6	0.932	0.625	0.844	0.592
		0.6	0.4	0.4	0.685	0.423	0.512	0.359
		0.6	0.4	0.25	0.562	0.304	0.313	0.292
		0.6	0.6	0.25	0.723	0.513	0.623	0.489
	1.5	0.6	0.6	0.6	0.764	0.553	0.692	0.443
		0.6	0.4	0.4	0.620	0.312	0.531	0.256
		0.6	0.4	0.25	0.532	0.235	0.312	0.232
		0.6	0.6	0.25	0.631	0.456	0.562	0.359
500	1.2	0.6	0.6	0.6	0.123	0.092	0.097	0.084
		0.6	0.4	0.4	0.093	0.063	0.078	0.045
		0.6	0.4	0.25	0.063	0.045	0.054	0.039
		0.6	0.6	0.25	0.113	0.093	0.094	0.079
	1.5	0.6	0.6	0.6	0.137	0.092	0.095	0.051
		0.6	0.4	0.4	0.075	0.043	0.062	0.023
		0.6	0.4	0.25	0.056	0.035	0.030	0.017
		0.6	0.6	0.25	0.112	0.082	0.092	0.024

第五章 資料分析

5.1 資料介紹

大腸癌好發在西方人，而在亞洲較西方的國家少，可能和傳統飲食有關。大腸直腸癌之形成，與肉食，高脂肪食物，低纖維食物之飲食習慣有密切關係。台灣近年來老年人口增加，同時因為油脂的攝取量增加，食物中纖維的攝取量減少，飲食習慣逐漸西化，大腸癌的患者數目也持續的增加。手術治療是大腸癌唯一根治的療法，診斷後要盡早切除，加上輔助療法有放射治療，化學藥物治療及免疫療法等。大腸癌手術治療後，常有局部復發或遠處轉移，少數有局部再發於遠端轉移的病人，仍可給予手術治療，大多病人則需要加上化學療法或放射療法治療。

引用 Moertel 等 (1991) 和 Lin (1994) 之大腸癌存活分析的化學藥物治療輔助治療臨床試驗資料，共 924 位受試者患有大腸癌，隨機分配 314 位受試者到對照組，310 位受試者到 Levamisole (Lev) 組，304 位受試者到 Levamisole + 5-fluorouracil (Lev + 5FU) 組，資料紀錄隨機分配到復發時間，以及隨機分配到死亡時間，共有 414 位受試者兩個事件都有觀察到，423 位受試者兩個事件都沒有觀察到，54 位受試者只觀察到隨機分配到復發時間，復發後仍然存活，38 位受試者無復發事件，只觀察到隨機分配到死亡時間。本文考慮探討在不同年齡、性別等變數下，建立接受單一治療(Lev)及合併治療(Lev+5FU)相對於接受對照組治療病人其餘命分位數比值的聯合信賴域，部分資料如表 5 顯現，變數說明如表 6 顯現。

表 5、大腸癌的化學藥物治療輔助治療臨床試驗部分資料

ID	treat	sex	age	Progdate time	progcens	Lastdate time	lastcens	strat
1	3	1	0	2.654	1	4.170	1	1
2	3	1	1	0.023	0	8.458	0	1
3	1	0	1	1.486	1	2.640	1	2
4	3	0	1	0.674	1	0.806	1	1
5	1	1	1	1.431	1	1.806	1	1
					⋮			
					⋮			
925	3	1	1	0.014	0	5.131	0	1
926	2	0	1	0.016	0	5.896	0	1
927	2	1	1	2.324	1	2.788	1	1
928	3	0	0	0.016	0	5.672	0	1
929	2	0	1	0.014	0	4.981	0	1

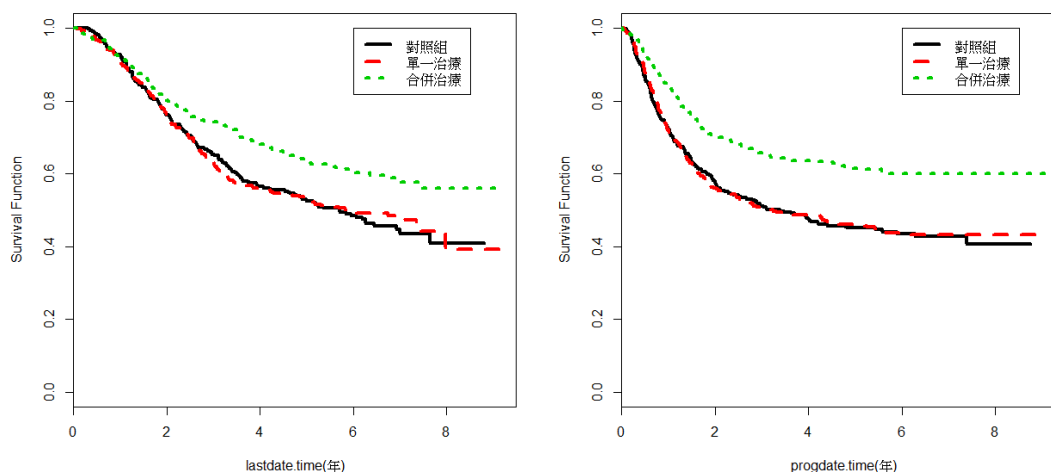
表 6、大腸癌的化學藥物治療輔助治療臨床試驗資料變數說明

變數	說明
ID	病患編碼
treat	治療組別： 1 = 觀察組；2 = Levamisole；3 = Levamisole + 5-fluorouracil
sex	性別：1=男；2=女
age	病患診斷大腸癌的年齡： 0=小於 60 歲；1=大於 60 歲
progdate.time	大腸癌復發時間(年)
progcens	大腸癌復發的狀態：0=無；1=有
lastdate.time	大腸癌死亡時間(年)
lastcens	大腸癌死亡的狀態：0=無；1=有
strat	0=大腸癌侵犯 1-4 個淋巴結；1=大腸癌侵犯大於 4 個淋巴結

5.2 分析結果

5.2.1 整體的治療效果分析

在實驗中 924 位受試者患有大腸癌，314 位受試者到對照組作為第 0 組，310 位受試者到 Levamisole (Lev) 作為第 1 組，304 位受試者到 Levamisole + 5-fluorouracil (Lev + 5FU) 作為第 2 組。將這三組死亡時間及復發時間的 Kaplan-Meier (1958) 存活函數估計繪於圖三。由圖三，可知，不管是死亡時間或復發時間，接受合併治療的病人相較對照組病人存活機率較高，但接受單一治療的病人與對照組病人的存活機率沒有差異。



圖三、死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計

於表 7 我們列出 $t_0 = 1.5$ 時，各組死亡時間的第 25 個百分位數餘命；及 $t_0 = 1$ 時，各組復發時間的第 20 個百分位數餘命。經由計算後列於表 7。

表 7、死亡時間和復發時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
死亡	1.73	1.41	3.09
復發	1	0.84	1.56

從表 7 得知，病人在接受治療後存活 1.5 年之下，其對照組、單一治療和合併療法的病人還有 75 % 的人分別可以存活 1.73、1.41 和 3.09 年。以及病人在接受治療 1 年後，其對照組、單一治療和合併療法的病人還有 80 % 的人，復發時間分別為 1、0.84 和 1.56 年。

表 8、死亡時間和復發時間餘命分位數比值的聯合信賴域估計

	死亡			復發		
	比 值	95%聯合信賴下界		比 值	95%聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.82	0.53	0.56	0.84	0.45	0.46
合併治療 V.S 對照組	1.79	1.07	1.09	1.56	0.84	0.84

從表 8 得知，病人在接受治療後存活 1.5 年之下，兩組治療處理相對於對照組的第 25 個百分位數比值，分別為 0.82、1.79；病人在接受治療 1 年後，兩組治療處理相對於對照組的第 20 個百分位數比值，分別為 0.84、1.56。從 95 % 聯合信賴下界得知，在死亡時間上無母數方法及經驗概似比方法皆建議合併治療的第 25 個百分位餘命比對照組大；但單一治療的第 25 個百分位餘命與對照組沒有差異；在復發時間上，不管是合併治療或單一治療相對於對照組的第 20 個百分位餘命都沒有差異。

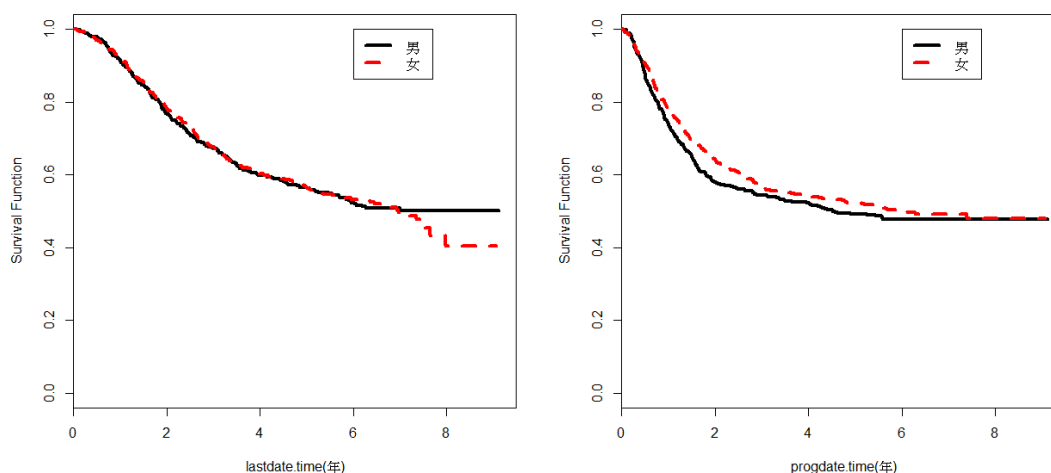
5.2.2 不同性別之下治療方式的差異

為探討治療是否因性別而有不同，因此我們分別在不同性別下建立兩治療組處理相對於對照組的餘命分位數之 95% 聯合信賴域。

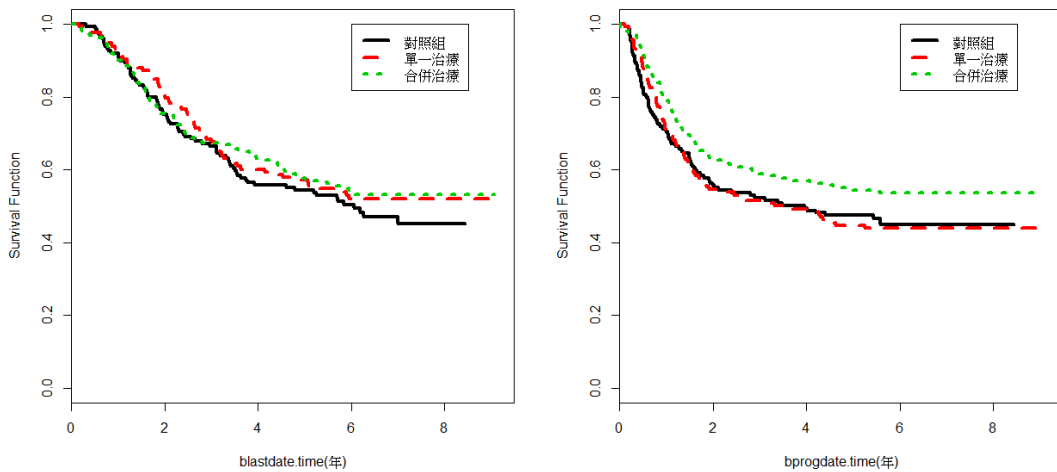
在實驗中 924 位受試者患有大腸癌，包含 445 位男性受試者患有大腸癌，其中 149 位受試者到對照組作為第 0 組，133 位受試者到

Levamisole (Lev)作為第 1 組，163 位受試者到 Levamisole + 5-fluorouracil (Lev+5FU) 作為第 2 組；女性則有 484 位，其中 166 位受試者到對照組作為第 0 組，177 位受試者到 Levamisole (Lev)作為第 1 組，141 位受試者到 Levamisole + 5-fluorouracil (Lev+5FU) 作為第 2 組，將性別別對這三組存活時間及復發時間的 Kaplan-Meier (1958)存活函數估計繪於圖四至圖六。

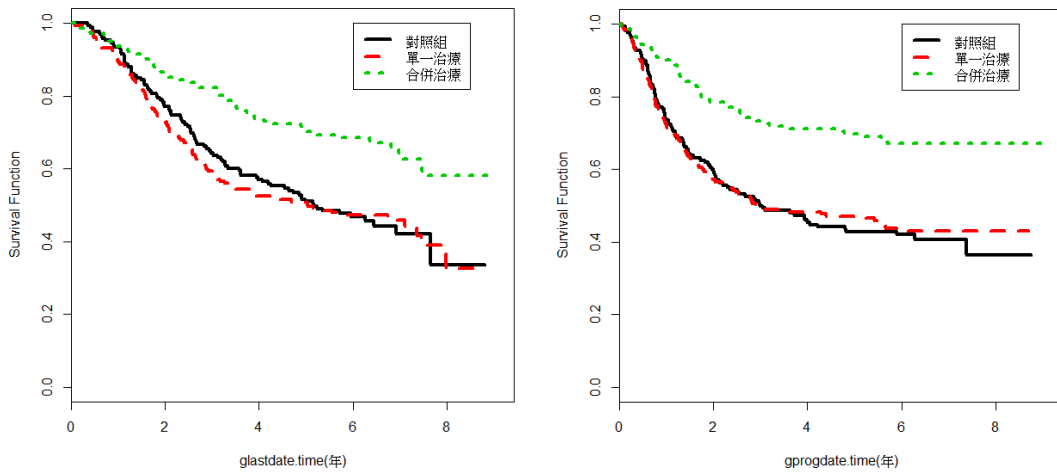
由圖四可知，不管是死亡時間或復發時間，在性別之間存活機率沒有差異下。由圖五發現，男生接受合併治療或單一治療的病人與對照組病人的存活機率沒有差異；由圖六則可看出，女生接受合併治療的病人相較對照組病人存活機率較高。但接受單一治療的病人與對照組病人的存活機率沒有差異。



圖四、性別之間死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖五、男性死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖六、女性死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計

接著分別在不同性別，計算出 $t_0 = 1.5$ 時，各組死亡時間的第 25 個百分位數餘命；及 $t_0 = 1$ 時，各組復發時間的第 20 個百分位數餘命。經由計算後列於表 9 和表 10。

表 9、不同性別之死亡時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
男性	1.89	1.64	2.64
女性	1.61	1.29	4.13

從表 9 得知，病人在接受治療後存活 1.5 年之下，男性病人之對照組、單一治療和合併療法的病人還有 75 % 的機會，分別可以存活 1.89、1.64 和 2.64 年；女性病人之對照組、單一治療和合併療法的病人還有 75 % 的機會，分別可以存活 1.61、1.29 和 4.13 年。

表 10、不同性別之復發時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
男性	0.93	0.80	0.92
女性	1	0.85	2.18

從表 10 得知，男性病人在接受治療 1 年後，其對照組、單一治療和合併療法的病人還有 80 % 的機會，復發時間分別為 0.92、0.8 和 0.92 年；女性病人之對照組、單一治療和合併療法病人還有 80 % 的機會，復發時間分別為 1、0.85 和 2.18 年。

表 11、不同性別之死亡時間餘命分位數比值的聯合信賴域估計

	男性			女性		
	比 值	95%聯合信賴下界		比 值	95%聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.87	0.49	0.52	0.80	0.40	0.40
合併治療 V.S 對照組	1.40	0.50	0.66	2.56	1.07	1.26

從表 11 得知，接受治療後存活 1.5 年之下，男性病人之兩組處理

組相對於對照組的第 25 個百分位數比值，分別為 0.87、1.4，從 95 % 聯合信賴下界得知，男性在死亡時間無母數方法及經驗概似比檢定法都顯示，上不管是合併治療或單一治療相對於對照組的第 25 個百分位餘命沒有差異；則女性分別為 0.8、2.56。然而，對女性而言，合併療法的第 25 個百分位餘命比對照組大，但單一治療於對照組的第 25 個百分位餘命則無差異。

別之復發時間餘命分位數比值的聯合信賴域估計

	男性			女性		
	比 值	95%聯合信賴下界		比 值	95%聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.86	0.24	0.36	0.84	0.42	0.47
合併治療 V.S 對照組	0.99	0.27	0.67	2.17	1.07	1.22

從表 12 得知，就復發時間而言，從表 10 得知結論與表 9 類似，皆是女性接受合併治療的第 20 個百分位餘命會比對照組好。

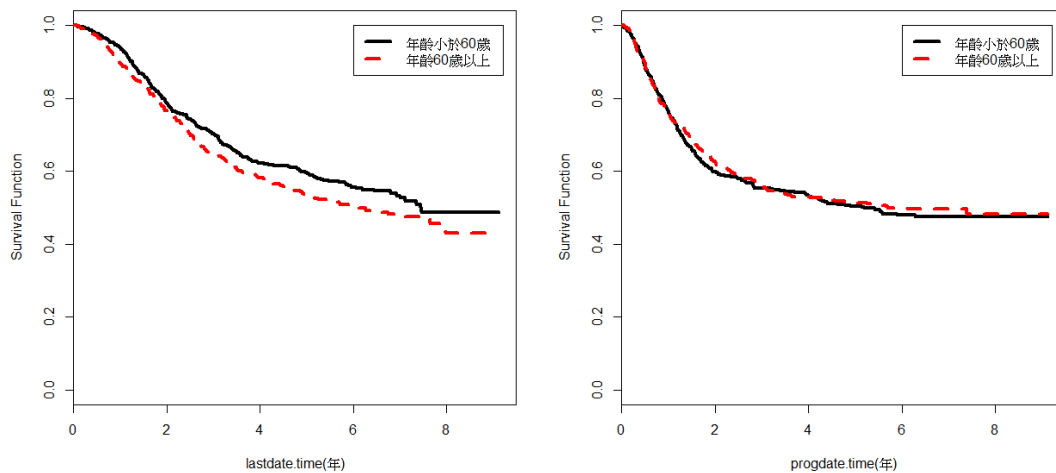
5.2.3 不同年齡之下治療方式的差異

為探討治療是否因年齡而有不同，因此我們分別在年齡小於 60 歲和年齡 60 以上之下，建立兩治療組處理相對於對照組的餘命分位數之 95% 聯合信賴域。

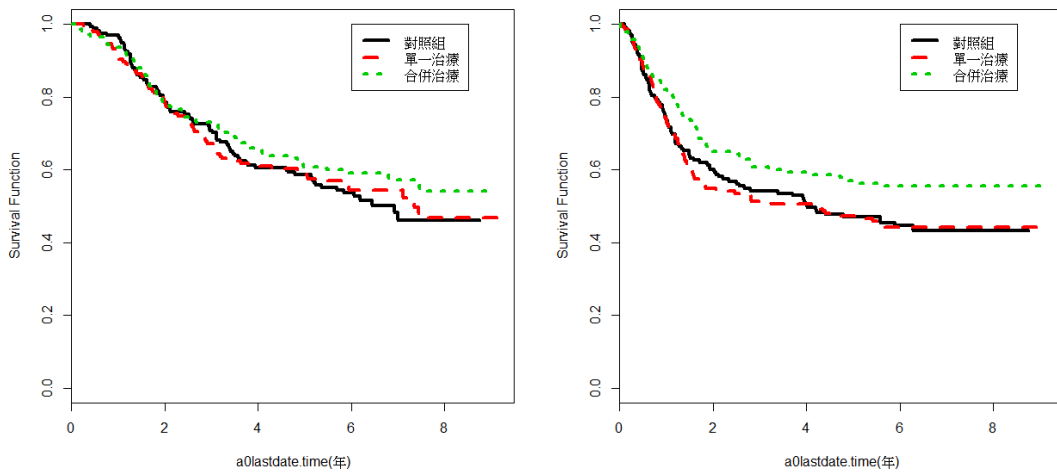
在實驗中 924 位受試者患有大腸癌，包含年齡小於 60 歲中有 445 位受試者患有大腸癌，其中 158 位受試者到對照組作為第 0 組，146 位受試者到 Levamisole (Lev) 作為第 1 組，141 位受試者到 Levamisole + 5-fluorouracil (Lev+5FU) 作為第 2 組；則年齡 60 歲以上中有 484 位，其中 157 位受試者到對照組作為第 0 組，164 位受試者到

Levamisole (Lev)作為第 1 組，163 位受試者 Levamisole + 5-fluorouracil (Lev +5FU) 作為第 2 組，將年齡和年齡分別對三組死亡時間及復發時間的 Kaplan-Meier (1958)存活函數估計繪於圖七至圖九。

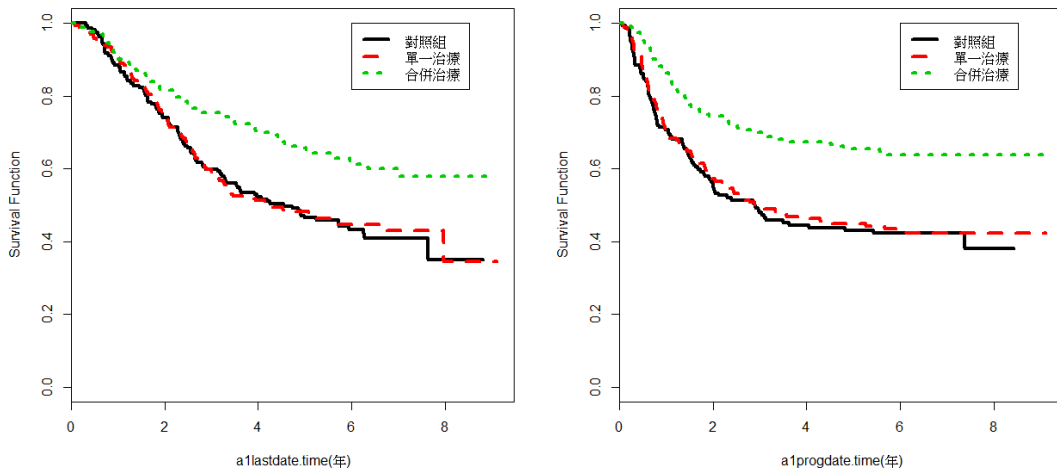
由圖七可知，不管是死亡時間或復發時間，年齡小於 60 歲及 60 歲以上的病人的存活機率沒有差異；圖八顯示，年齡小於 60 歲接受合併治療或單一治療的病人與對照組病人的存活機率沒有差異；圖九則可看出，年齡 60 歲以上接受合併治療的病人相較對照組病人存活機率較高，但接受單一治療的病人與對照組病人的存活機率沒有差異。



圖七、不同年齡組病人的死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖八、年齡小於 60 歲病人的死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖九、年齡 60 歲以上病人的死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計

為了解年齡的影響，我們分別就不同年齡組，計算出 $t_0 = 1.5$ 時，各組死亡時間的第 25 個百分位數餘命；及 $t_0 = 1$ 時，各組復發時間的第 20 個百分位數餘命。經由計算後列於表 13 和表 14。

表 13、不同年齡組之死亡時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
年齡小於 60 歲	1.99	1.64	2.30
年齡 60 歲以上	1.30	1.23	3.59

從表 13 得知，病人在接受治療後存活 1.5 年之下，年齡小於 60 歲且接受對照組、單一治療或合併療法的病人還有 75 % 機會，分別可以存活 1.99、1.64 和 2.30 年；但對年齡 60 歲以上的病人，則分別可以存活 1.30、1.23 和 3.59 年。

表 14、不同年齡之復發時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
年齡小於 60 歲	1.06	0.57	0.95
年齡 60 歲以上	0.95	1.19	1.25

從表 14 得知，病人在接受治療 1 年後，年齡小於 60 歲且接受對照組、單一治療和合併療法的病人還有 80 % 機會，復發時間分別為 1.06、0.57 和 0.95 年；年齡 60 歲以上，則分別復發時間為 0.95、1.19 和 1.25 年。

表 15、不同年齡組之死亡時間餘命分位數比值的聯合信賴域估計

	年齡小於 60 歲			年齡 60 歲以上		
	比 值	95% 聯合信賴下界		比 值	95% 聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.82	0.37	0.46	0.95	0.52	0.80
合併治療 V.S 對照組	1.16	0.50	0.66	2.76	1.39	1.65

從表 15 得知，病人在接受治療後存活 1.5 年之下，年齡小於 60

歲的病人之兩組處理組相對於對照組的第 25 個百分位數比值，分別為 0.82、1.16，從無母數方法及經驗概似比方法 95 % 聯合信賴下界得知，在死亡時間上不管是合併治療或單一治療對於對照組的第 25 個百分位餘命沒有差異；年齡 60 歲以上的病人之兩組處理組相對於對照組的比值，分別為 0.95、2.76。從無母數方法及經驗概似比方法 95 % 聯合信賴下界皆建議，在死亡時間上合併治療對於對照組的第 25 個百分位餘命有差異；單一治療對於對照組的第 25 個百分位餘命沒有差異。

表 16、不同年齡組之復發時間餘命分位數比值的聯合信賴域估計

	年齡小於 60 歲			年齡 60 歲以上		
	比 值	95% 聯合信賴下界		比 值	95% 聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.54	0.17	0.23	1.25	0.42	0.47
合併治療 V.S 對照組	0.90	0.42	0.46	2.25	1.23	1.31

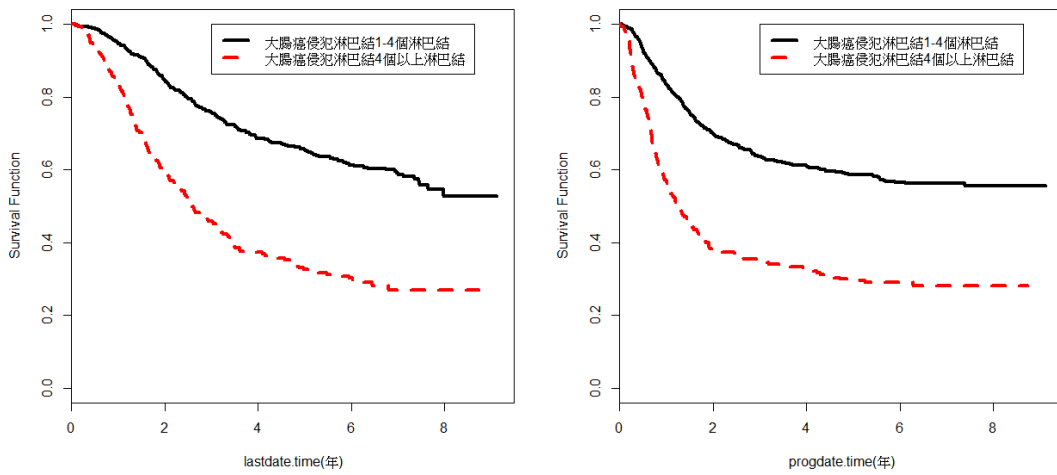
從表 16 得知，病人在接受治療後存活 1 年之下，年齡小於 60 歲的病人之兩組處理組相對於對照組的第 20 個百分位數比值，分別為 0.54、0.9，從無母數方法及經驗概似比方法 95 % 聯合信賴下界得知，在復發時間上不管是合併治療或單一治療對於對照組的第 20 個百分位餘命沒有差異；年齡 60 歲以上的病人之兩組處理組相對於對照組的比值，分別為 1.25、2.25。從無母數方法及經驗概似比方法 95 % 聯合信賴下界皆建議，在復發時間上合併治療對於對照組的第 20 個百分位餘命有差異；單一治療對於對照組的第 20 個百分位餘命沒有差異。

5.2.4 大腸癌侵犯淋巴結個數之下來看治療方式的差異

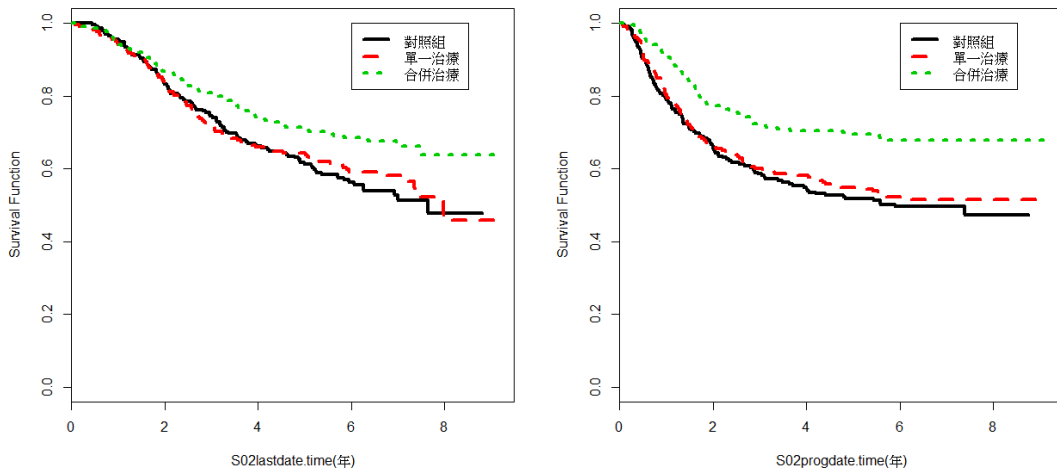
為探討治療效果是否因大腸癌侵犯淋巴結個數而有不同，因此我們分別在大腸癌侵犯 1-4 個淋巴結和大腸癌侵犯 4 個以上淋巴結之下，建立兩治療組處理相對於對照組的餘命分位數之 95% 聯合信賴域。

在實驗中 924 位受試者患有大腸癌，包含大腸癌侵犯 1-4 個淋巴結有 674 位受試者患有大腸癌，其中 228 位受試者到對照組作為第 0 組，221 位受試者到 Levamisole (Lev) 作為第 1 組，225 位受試者到 Levamisole + 5-fluorouracil (Lev+5FU) 作為第 2 組；則侵犯 4 個以上淋巴結，其中 87 位受試者到對照組作為第 0 組，89 位受試者到 Levamisole (Lev) 作為第 1 組，79 位受試者到 Levamisole + 5-fluorouracil (Lev+5FU) 作為第 2 組，將侵犯淋巴結個數和侵犯淋巴結個數分別對三組存活時間及復發時間的 Kaplan-Meier (1958) 存活函數估計繪於圖十至圖十二。

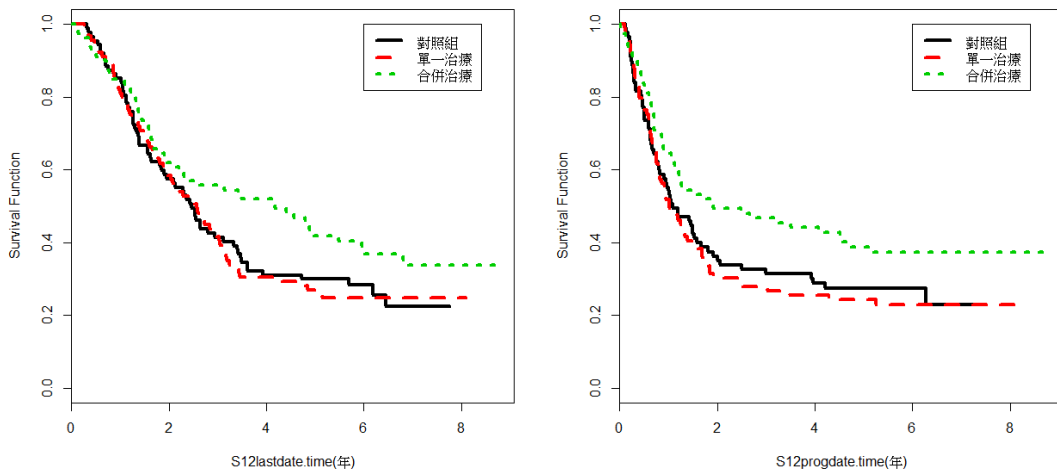
圖十可知，不管是死亡時間或復發時間，在侵犯淋巴結個數 1-4 個淋巴結相比 4 個以上的病人存活機率較高；圖十一可知，侵犯淋巴結個數 1-4 個且接受合併治療或單一治療的病人與對照組病人的存活機率沒有差異；圖十二可知，侵犯 4 個以上淋巴結且接受合併治療的病人相較對照組病人存活機率較高，但接受單一治療的病人與對照組病人的存活機率沒有差異。



圖十、大腸癌侵犯淋巴結個數之間死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖十一、大腸癌侵犯 1-4 個淋巴之結死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計



圖十二、大腸癌侵犯 4 個以上淋巴結之死亡時間(左)及復發時間(右)之 Kaplan-Meier 存活函數估計

為了解大腸癌侵犯淋巴結個數的影響，我們分別就不同淋巴結個數組，計算出 $t_0 = 1.5$ 時，各組死亡時間的第 25 個百分位數餘命；及 $t_0 = 1$ 時，各組復發時間的第 20 個百分位數餘命。經由計算後列於表 17 和表 18。

表 17、大腸癌侵犯淋巴結個數之死亡時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
1-4 個淋巴結	2.24	2.13	4.33
4 個以上淋巴結	0.97	0.79	1.64

從表 17 得知，病人在接受治療後存活 1.5 年之下，侵犯 1-4 個淋巴結的病人且接受對照組、單一治療和合併療法的病人還有 75 % 機會，分別可以存活 2.24、2.13 和 4.33 年；但對侵犯 4 個以上淋巴結，則分別可以存活 0.97、0.79 和 1.64 年。

表 18、大腸癌侵犯淋巴結個數之復發時間餘命分位數估計

	餘命分位數		
	對照組	單一治療	合併治療
1-4 個淋巴結	1.20	1.33	1.83
4 個以上淋巴結	0.50	0.38	0.69

從表 18 得知，病人在接受治療 1 年後，侵犯 1-4 個淋巴結的病人且接受對照組、單一治療和合併療法的病人還有 80 % 機會，復發時間分別為 1.2、1.33 和 1.83 年；侵犯 4 個以上淋巴結，則復發時間分別為 0.5、0.38 和 0.69 年。

表 19、大腸癌侵犯淋巴結個數之死亡時間餘命分位數比值的聯合信賴域估計

	侵犯 1-4 個淋巴結			侵犯 4 個以上淋巴結		
	比 值	95% 聯合信賴下界		比 值	95% 聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	0.95	0.45	0.81	0.81	0.37	0.42
合併治療 V.S 對照組	1.94	0.80	0.97	1.69	0.36	0.65

從表 19 得知，病人在接受治療後存活 1.5 年之下，侵犯 1-4 個淋巴結的病人之兩組處理組相對於對照組的第 25 個百分位數比值，分別為 0.95、1.94，從無母數方法及經驗概似比方法 95 % 聯合信賴下界得知，在死亡時間上不管是合併治療或單一治療對於對照組的第 25 個百分位餘命沒有差異；侵犯 4 個以上淋巴結的病人之兩組處理組相對於對照組的比值，分別為 0.76、1.69。從無母數方法及經驗概似比方法 95 % 聯合信賴下界皆建議，在死亡時間上合併治療或單一治療對於照組的第 25 個百分位餘命沒有差異。

表 20、大腸癌侵犯淋巴結個數之復發時間餘命分位數的聯合信賴域估計

	侵犯 1-4 個淋巴結			侵犯 4 個以上淋巴結		
	比 值	95%聯合信賴下界		比 值	95%聯合信賴下界	
		SCI _N	SCI _E		SCI _N	SCI _E
單一治療 V.S 對照組	1.10	0.36	0.43	0.76	0.35	0.37
合併治療 V.S 對照組	1.52	0.73	0.84	1.38	0.51	0.56

從表 20 得知，病人在接受治療後存活 1 年之下，侵犯 1-4 個淋巴結的病人之兩組處理組相對於對照組的第 20 個百分位數比值，分別為 1.10、1.52，從無母數方法及經驗概似比方法 95 %聯合信賴下界得知，在復發時間上不管是合併治療或單一治療對於對照組的第 25 個百分位餘命沒有差異；侵犯 4 個以上淋巴結的病人之兩組處理組相對於對照組的比值，分別為 0.76、1.38。從無母數方法及經驗概似比方法 95 %聯合信賴下界皆建議，在復發時間上合併治療或單一治療對於對照組的第 20 個百分位餘命沒有差異。

綜合而言，不管是死亡時間或復發時間，女性、年齡 60 歲以上和淋巴結 4 個以上且接受治療的病人，在無母數及經驗概似比檢定方法下，所建立的 95%信賴下界皆建議，合併療法比對照組的餘命分位數大；但是，根據經驗概似比檢定法所建立的 95%信賴下界比無母數方法大，所以提供較多的訊息。

第六章 結論與建議

本文主要針對右設限存活資料，根據無母數及經驗概似比方法建立多組餘命分位數比值之聯合信賴域估計再據此進行多對一的比較。由模擬研究結果得知，若樣本數小時，無母數及經驗概似比方法所建立多組餘命分位數比值在覆蓋機率皆會略高於所設定的信心水準；但當樣本數提高時，無母數及經驗概似比方法所建立多組餘命分位數比值在覆蓋機率皆能維持其信心水準，但是經驗概似比方法趨近於設定的信心水準的速度比無母數方法快。樣本數小時，無母數所建立多組餘命分位數比值在 90% 和 95% 的平均覆蓋面積比經驗概似比方法大，除了在此假設型態 $S_0 < S_1 < S_2$ 之外；但當樣本數提高時，無母數及經驗概似比方法所建立多組餘命分位數比值在平均覆蓋面積皆有明顯縮減。

從資料分析來看，不管是死亡時間或復發時間，女性、年齡 60 歲以上和淋巴結 4 個以上且接受治療的病人，在無母數及經驗概似比檢定方法下，所建立的 95% 信賴下界皆建議，合併療法比對照組的餘命分位數大；但是，根據經驗概似比檢定法所建立的 95% 信賴下界比無母數方法大，所以提供較多的訊息。本文資料可對應到模擬假設 $S_0 = S_1 < S_2$ ，且模擬與真實資料結果一致。此外經驗概似比方法所建立的多組餘命分位數之 95% 信賴域寬度比無母數方法窄，且避免變異數的估計，所以我們建議採用經驗概似比檢定法。

由於，在真實生活中疾病可能會因為某些因素造成死亡的加速，例如：本文資料分析結果來看，性別、年齡、淋巴結的個數，在死亡時間及復發時間且接受治療皆會影響大腸癌病人的餘命，所以在未來的研究中，可以加入共變數因子，利用 Cox 模型來做預測，建立多組餘命分位數比值的條件聯合信賴域。

參考文獻

1. Aalen, O. O., Nonparametric inference for a family of counting processes. *The Annals of Statistics* 6, 701-726 (1978).
2. Bathke A, Kim M, Zhou M. Combined multiple testing by censored empirical likelihood. *Journal of Statistical Planning and Inference* 139, 814-827 (2009).
3. Berger RL, Boos DD, Guess FM. Tests and confidence sets for comparing two mean residual life functions. *Biometrics* 44, 103-115 (1988.)
4. Basawa, I. V. and Koul, H. L. Large-sample statistics based on quadratic dispersion. *International Statistical Review* 56, 199-219 (1988).
5. Chen, Y.-Y., Hollander, M., and Langberg, N. A. Tests for monotone mean residual life, using randomly censored data. *Biometrics* 39, 119-127 (1983).
6. Chen, Y. Q., Jewell, N. P., Lei, X., and Cheng, S. C. Semiparametric estimation of proportional mean residual life model in presence of censoring. *Biometrics* 61, 170-178 (2005).
7. Fleming, T. R. and Harrington, D. P. *Counting Processes and Survival Analysis*. New York: Wiley (1991).
8. Fligner, M. A. and Rust, S. W. A modification of Mood's median test for the generalized Behrens-Fisher problem. *Biometrika* 69, 221-226 (1982).
9. Jeong JH, Jung SH, Costantino JP. Nonparametric inference on median residual life function. *Biometrics* 64, 157-163 (2008).
10. Maguluri, G. and Zhang, C.-H. Estimation in the mean residual life regression model. *Journal of the Royal Statistical Society Series B* 56,

- 477-489 (1994).
11. Nelson, W. Theory and applications of hazard plotting for censored failure data. *Technometrics* 14, 945-965 (1972).
 12. Oakes, D. and Dasu, T. A note on residual life. *Biometrika* 77, 409-410 (1990).
 13. Owen A. Empirical likelihood ratio confidence intervals for a single functional. *Biometrika* 75, 237-249 (1988).
 14. Qin G, Zhao Y. Empirical likelihood inference for the mean residual life under random censorship. *Statistics and Probability Letters* 77, 549-557 (2007).
 15. Su, J. Q. and Wei, L. J. Nonparametric estimation for the difference or ratio of median failure times. *Biometrics* 49, 603-607 (1993).
 16. Wang, J. L. and Hettmansperger, T. P. Two-sample inference for median survival times based on one-sample procedures for censored survival data. *Journal of the American Statistical Association* 85, 529-536 (1990).
 17. Zhao Y, Qin G. Inference for the mean residual life function via empirical likelihood. *Communications in Statistics: Theory and Methods* 35, 1025-1036 (2006).
 18. Zhou M, Jeong JH. Empirical likelihood ratio test for median and mean residual lifetime. *Statistics in Medicine* 30, 152-159 (2011).

附錄

附錄 A-1: $W_{t_0}(\boldsymbol{\gamma}_{t_0,0})$ 的極限分布

因為 $u_k(\theta_{t_0}^{(k)}) = S_k(t_0 + \theta_{t_0}^{(k)}) - (1 - q)S_k(t_0) = 0, k = 1, \dots, m$ ，所

$$\begin{aligned} \text{以得下式 } \hat{u}_0(\theta_{t_0}^{(0)}) &= \hat{S}_0(t_0 + \theta_{t_0}^{(0)}) - (1 - q)\hat{S}_0(t_0) \\ &= \hat{S}_0(t_0 + \theta_{t_0}^{(0)}) - S_0(t_0 + \theta_{t_0}^{(0)}) - (1 - q)\{\hat{S}_0(t_0) - S_0(t_0)\} \end{aligned}$$

利用泰勒展開式(Taylor series expansion) $u_0(\theta_{t_0}^{(0)})$ 在 $\hat{\theta}_{t_0}^{(0)}$ 進行展開，

其中 $\gamma_{t_0}^{(k)} = \hat{\theta}_{t_0}^{(k)} / \hat{\theta}_{t_0}^{(0)}$ 是 $\gamma_{t_0}^{(k)}$ 的估計值

$$u_0(\theta_{t_0}^{(0)}) = -f_0(t_0 + \theta_{t_0,0}^{(0)}) (\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)}) + o_P(n^{-\frac{1}{2}}) \quad (\text{A.1})$$

其中 $f_k(\cdot)$ ($k = 1, \dots, m$) 是存活機率 $S_k(\cdot)$ 的機率密度函數。

令 $\theta_{t_0,0}^{(k)} = \gamma_{t_0,0}^{(k)} \theta_{t_0,0}^{(0)}$ 相同地，

$$\begin{aligned} \hat{u}_k(\theta_{t_0}^{(k)}) &= \hat{S}_k(t_0 + \theta_{t_0}^{(k)}) - (1 - q)\hat{S}_1(t_0) \\ &= -f_k(t_0 + \theta_{t_0,0}^{(k)}) (\hat{\theta}_{t_0}^{(k)} - \theta_{t_0}^{(k)}) + o_P(n^{-\frac{1}{2}}) \\ &= -f_k(t_0 + \theta_{t_0,0}^{(k)}) \left\{ \theta_{t_0,0}^{(0)} (\hat{\gamma}_{t_0} - \gamma_{t_0,0}) + \gamma_{t_0,0} (\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)}) \right\} \\ &\quad + o_P(n^{-\frac{1}{2}}), k = 1, \dots, m \end{aligned} \quad (\text{A.2})$$

從(A.1)和(A.2)得到 $V_{t_0}(\theta_{t_0,0}^{(0)}, \boldsymbol{\gamma}_{t_0,0}) = K_{t_0}(\theta_{t_0,0}^{(0)}, \boldsymbol{\gamma}_{t_0,0}) + o_P(n^{-\frac{1}{2}})$

其中

$$\begin{aligned} K_{t_0}(\theta_{t_0,0}^{(0)}, \boldsymbol{\gamma}_{t_0,0}) &= \frac{\hat{u}_0^2(\theta_{t_0}^{(0)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(0)})} + \sum_{k=1}^m \frac{\hat{u}_k^2(\theta_{t_0}^{(k)})}{\hat{\sigma}_{t_0}^2(\theta_{t_0}^{(k)})} \\ &= \frac{f_0^2(t_0 + \theta_{t_0,0}^{(0)}) (\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)})^2}{\sigma_0^2(\theta_{t_0}^{(0)})} \end{aligned}$$

$$+ \sum_{k=1}^m \frac{f_k^2 \left(t_0 + \theta_{t_0,0}^{(k)} \right) \left\{ \theta_{t_0,0}^{(k)} (\hat{\gamma}_{t_0} - \gamma_{t_0,0}) + \gamma_{t_0,0} \left(\hat{\theta}_{t_0}^{(k)} - \theta_{t_0}^{(k)} \right) \right\}^2}{\sigma_k^2 \left(\theta_{t_0}^{(k)} \right)}$$

當 $\theta_{t_0}^{(0)} = \hat{\theta}_{t_0}^{(0)}$ ， $K_{t_0} \left(\theta_{t_0,0}^{(0)}, \gamma_{t_0,0} \right)$ 會有小值，所以 $V_{t_0} \left(\theta_{t_0,0}^{(0)}, \gamma_{t_0,0} \right)$ 也會有最小值，因為 $\hat{\theta}_{t_0}^{(0)}$ 是 $\theta_{t_0}^{(0)}$ 的一致性估計，當 $n \rightarrow \infty$ ， $\theta_{t_0}^{(0)}$ 會收斂到 $\hat{\theta}_{t_0}^{(0)}$ 。

$$W_{t_0} \left(r_{t_0,0} \right) = \min_{\theta_{t_0}^{(0)}} V_{t_0} \left(r_{t_0,0}, \theta_{t_0}^{(0)} \right)$$

$W_{t_0} \left(r_{t_0,0} \right)$ 會服從卡方分配，自由度 m 。

附錄 A-2: $Q(\gamma_{t_0,0})$ 的極限分布

$$\hat{\varphi}_0 \left(\theta_{t_0}^{(0)} \right) = \hat{F}_{KM} \left(x + \theta_{t_0}^{(0)} \right) - (1 - q) \hat{F}_{KM}(x) - q$$

$$\hat{\varphi}_1 \left(\gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right) = \hat{F}_{KM} \left(x + \gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right) - (1 - q) \hat{F}_{KM}(x) - q$$

⋮

$$\hat{\varphi}_m \left(\gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right) = \hat{F}_{KM} \left(x + \gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right) - (1 - q) \hat{F}_{KM}(x) - q$$

其中 $\hat{F}_{KM}(\cdot)$ 是一個存活函數，

$W_0 \left(\theta_{t_0}^{(0)} \right) + W_1 \left(\gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right) + W_2 \left(\gamma_{t_0,0}^{(2)} \theta_{t_0}^{(0)} \right) + \dots + W_m \left(\gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right)$ ，在

EL 比值(3.3)之下

$$\frac{n[\hat{\varphi}_0 \left(\theta_{t_0}^{(0)} \right) - \varphi_0 \left(\theta_{t_0}^{(0)} \right)]}{\xi_0^2} + \frac{n[\hat{\varphi}_1 \left(\gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right) - \varphi_1 \left(\gamma_{t_0,0}^{(1)} \theta_{t_0}^{(0)} \right)]}{\xi_1^2} + \dots$$

$$+ \frac{n[\hat{\varphi}_m \left(\gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right) - \varphi_m \left(\gamma_{t_0,0}^{(m)} \theta_{t_0}^{(0)} \right)]}{\xi_m^2}$$

其中 ξ_k^2 , $k = 0, 1, \dots, m$ 是 $\hat{\theta}_0(\theta_{t_0}^{(0)})$, $\hat{\theta}_1(\gamma_{t_0,0}^{(1)}\theta_{t_0}^{(0)})$, \dots , $\hat{\theta}_m(\gamma_{t_0,0}^{(m)}\theta_{t_0}^{(0)})$ 的變異數。

在假設 $\gamma_{t_0,0} = 1, \gamma_{t_0,0} = (\gamma_{t_0,0}^{(1)}, \gamma_{t_0,0}^{(2)}, \dots, \gamma_{t_0,0}^{(m)})$, 均勻一致性(uniform consistency) 和 近似常態 (asymptotic normality) 的條件下, 利用 Kaplan-Meier 估計和 delta method, 得到近似值如下:

$$U(\theta) = \frac{n(\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)})}{\sigma_0^2} + \frac{n(\gamma_{t_0,0}^{(1)}\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)})}{\sigma_1^2} + \dots + \frac{n(\gamma_{t_0,0}^{(m)}\hat{\theta}_{t_0}^{(0)} - \theta_{t_0}^{(0)})}{\sigma_m^2}$$

其中 σ_k^2 , $k = 0, 1, \dots, m$ 是 $\hat{\theta}_{t_0}^{(0)}$, $\gamma_{t_0,0}^{(1)}\hat{\theta}_{t_0}^{(0)}$, \dots , $\gamma_{t_0,0}^{(m)}\hat{\theta}_{t_0}^{(0)}$ 的變異數。

在

$$\alpha_{01} = \frac{a_0\hat{\theta}_{t_0}^{(0)} + a_1\gamma_{t_0,0}^{(1)}\hat{\theta}_{t_0}^{(0)}}{a_0 + \gamma_{t_0,0}^{(1)2} a_1}, \alpha_{02} = \frac{a_0\hat{\theta}_{t_0}^{(0)} + a_2\gamma_{t_0,0}^{(2)}\hat{\theta}_{t_0}^{(0)}}{a_0 + \gamma_{t_0,0}^{(2)2} a_1}, \dots, \alpha_{0m} = \frac{a_0\hat{\theta}_{t_0}^{(0)} + a_m\gamma_{t_0,0}^{(m)}\hat{\theta}_{t_0}^{(0)}}{a_0 + \gamma_{t_0,0}^{(m)2} a_1}$$

統計量 $U(\theta)$ 會有最小值, 其中 $a_k = 1/\sigma_k$, $k = 0, 1, \dots, m$, 整理

過後如下:

$$Q(\gamma_{t_0,0}) = \frac{(\hat{\theta}_{t_0}^{(0)} - \gamma_{t_0,0}^{(1)}\hat{\theta}_{t_0}^{(0)})^2}{1/a_0 + \gamma_{t_0,0}^{(1)2}/a_1} + \frac{(\hat{\theta}_{t_0}^{(0)} - \gamma_{t_0,0}^{(2)}\hat{\theta}_{t_0}^{(0)})^2}{1/a_0 + \gamma_{t_0,0}^{(2)2}/a_2} + \dots + \frac{(\hat{\theta}_{t_0}^{(0)} - \gamma_{t_0,0}^{(m)}\hat{\theta}_{t_0}^{(0)})^2}{1/a_0 + \gamma_{t_0,0}^{(m)2}/a_m}$$

$Q(\gamma_{t_0,0})$ 會服從卡方分配, 自由度 m 。