



國立臺灣大學公共衛生學院

流行病學與預防醫學研究所 生物醫學統計組

碩士論文

Division of Biostatistics, Graduate Institute of Epidemiology and

Preventive Medicine

College of Public Health

National Taiwan University

Master Thesis

二元有序間隔時間之交叉分位數比分析

Analysis of cross quantile ratio for two serial gap times

林莞筑

Wan-Chu Lin

指導教授：張淑惠 博士

Advisor: Shu-Hui Chang, Ph.D.

中華民國 108 年 8 月

August, 2019

誌謝



這篇論文得以完成，首先要感謝指導教授張淑惠老師兩年來的指導。老師除了教導我如何設定主題與做研究、如何自我學習，並且在遇到問題時總是一語道破並適時給予協助。老師開明的作風以及對學生的期許，提供了很好的學習環境以及方向，對此我由衷感謝。

其次要感謝蘇登煌學長，對於我的論文細心指導，指出問題並提供建議；平日亦十分照顧研究室的學弟妹，不時叮囑我們，並對未來生涯方向給予建議。

在此感謝戴政老師、陳秀熙老師以及丘政民老師於口試時的指導與建議，讓我獲益良多，重新檢視論文出發點與動機。感謝研究室的同學憲周，一路上的相互提醒與幫忙，還有學弟們不時的幫助。感謝流預所以及所辦這兩年來幫助過我大大小小的事情。最後要感謝家人以及咪仔的支持，感謝你們對我的包容與遷就，讓我能無罣礙做自己想做的事。

有以上這些人的幫助，這篇論文才得以完成，誠摯地感謝這一切。這兩年我學習到很多，也得到很多善意與幫助。期許自己未來仍能不斷進步，盡我所能在社會的某個角落努力發光，並且回饋給社會與台灣。

摘要



序列事件常見於現代醫學和流行病學縱向研究中，而有序間隔時間資料常為研究者感興趣之主題。研究者欲了解病人的疾病進程而進行早期預防治療，而有序間隔時間之間相關性蘊含病人疾病歷程資訊，是自然的預測因子。本篇提出交叉分位數比 (cross quantile ratio, CQR) 測量隨時間改變之相依性，並以前兩段間隔時間為例。利用無母數方法，設限分位數迴歸估計交叉分位數比以及使用設限機率倒數權重調整相依設限，而不須額外假設聯合分布。本文另外討論交叉分位數比估計之一致性以及大樣本下共變異數估計。模擬結果顯示，交叉分位數比表現受限於設限率，在一定範圍的分位數相對偏誤都在 5% 以下；本文提供之標準差估計方法在大樣本下會有較好表現，最後以 rhDNase 資料為例，進行實際資料分析。

關鍵字：間隔時間、相關性、交叉分位數比、分位數迴歸

Abstract



Serial event data are often encountered and of interest in the follow-up studies of chronic diseases and gap times between successive events. The relationship between serial gap times may provide predictive information on disease progression. In this thesis, the cross quantile ratio (CQR) is introduced to measure the time-varying dependence between the first and second gap times without specifying their joint distribution. Nonparametric estimation of cross quantile ratio can be carried out through censored quantile regression approach. In addition, the inverse probability of censoring weights is used to tackle the induced dependent censoring. The asymptotic properties of the proposed estimators are investigated and the corresponding asymptotic variance estimators are provided as well. Simulation results suggest good performance of the proposed methods within a certain range of quantile due to censoring. The rhDNase dataset is analyzed for further illustration of CQR method

Keywords: gap times, association, cross quantile ratio, quantile regression



目錄

誌謝.....	I
摘要.....	II
Abstract.....	III
第一章 導論.....	1
第一節 前言.....	1
第二節 研究動機.....	1
第二章 文獻回顧.....	3
第一節 分位數迴歸模型於存活資料.....	3
第二節 交叉殘差分位數比(非單次事件).....	4
第三節 分位數勝算比.....	7
第四節 使用誘導平滑方法估計分位數迴歸模型的變異數.....	9
第三章 方法.....	11
第一節 交叉分位數比.....	12
第二節 有序間隔時間交叉分位數比估計.....	14
第三節 交叉分位數比估計量之共變異數估計.....	16
第四章 模擬.....	20
第一節 阿基米德耦合.....	20
第二節 對數常態模型.....	25
第三節 分段常數克萊頓 (Piecewise constant Clayton).....	28
第五章 實際資料分析.....	32
第一節 資料探索.....	34
第二節 納入過去歷史.....	36
第六章 討論與總結.....	41
第七章 參考文獻.....	43
附錄.....	45
A. 理論 A 之證明 (consistency)	45
B. 理論 B 之證明 (asymptotic normal)	47
C. 共變異數之估計步驟.....	51
D. Resampling 之理論證明	54
E. 模擬結果表格.....	55

圖 目錄

圖 1、分位數迴歸之示意圖	13
圖 2、資料結構及符號，C 為設限時間， T_m 為第 m 次感染事件與治療結束時間， $m=1,...,5$	33
圖 3、第一次感染間隔時間對治療分組之存活曲線(調整 fev)	34
圖 4、第二次感染間隔時間對治療分組之存活曲線(調整 fev)	35
圖 5、 $\beta_1(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間(虛線)。	36
圖 6、 $\tau_1 = 0.7$ ， $\beta(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間(虛線)。	38
圖 7、 $\tau_1 = 0.7$ ， $\hat{\beta}_1(\tau_1, \tau_2) + \hat{\beta}_3(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間(虛線)。	39
圖 8、Kaplan Meier 圖	40



表 目錄

表 1、Clayton 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	55
表 2、Clayton 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	56
表 3、Clayton 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果.....	57
表 4、Clayton 樣本數 600，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	58
表 5、Frank 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果	59
表 6、Frank 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果	60
表 7、Frank 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果	61
表 8、Gumbel 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	62
表 9、Gumbel 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	63
表 10、Gumbel 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果.....	64
表 11、Lognormal 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	65
表 12、Lognormal 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果.....	66
表 13、Lognormal 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果.....	67
表 14、Piecewise constantClayton 樣本數 200，相關性 $\theta_1 = 0.5$ ， $\theta_2 = 0.2$ ， β_1 模擬結果	68
表 15、Piecewise constantClayton 樣本數 200，相關性 $\theta_1 = 0.2$ ， $\theta_2 = 0.5$ ， β_1 模擬結果	69
表 16、Piecewise constantClayton 樣本數 200，相關性 $\theta_1 = 0.2$ ， $\theta_2 = -0.5$ ， β_1 模擬結果	70



第一章 導論


第一節 前言

序列事件(serial event)資料常見於現代醫學和流行病學縱向研究中(longitudinal study)，因醫療技術的進步使得人類壽命延長，病人在死亡前通常會經歷多次的感染，而有序間隔時間 (gap time) 之間相關性則蘊含病人疾病歷程。例如囊狀纖維化 (Fuchs et al. 1994)，病人因黏液積聚呼吸道導致感染，患者每次感染可能在肺部留下纖維化的疤痕組織，使得肺部功能逐漸下降，甚至死亡，了解有序間隔時間之間相關性有助於調整病人之後治療。研究者希望了解，較早發生感染患者，在治療過後是否很快會再次感染；抑或者，治療期間較長的病患可能代表感染較為嚴重，病人在治療過後是否很快會再次感染。若兩段間隔時間為高度正相關，則醫療人員可早期介入高風險族群。

已有許多文獻探討二元事件時間相關性。Clayton (1978) 提出常數交叉比例函數 (constant cross ratio)，即條件風險比例函數，探討二元事件時間相關性。但序列間隔時間有誘導相依設限問題，無法直接使用上述方法。Lakhal-Chaieb et al. (2010) 使用 Kendall's tau 測量兩連續間隔時間相關性，利用設限機率倒數權重 (Inverse probability censoring weight, IPCW) 調整誘導相依設限問題。Fu et al. (2016) 將其延伸至多個復發間隔時間並可彈性納入過去復發歷史。

第二節 研究動機

在一些臨床情境中，二元事件時間相關性不隨時間改變的假設不成立。例如女性月經週期改變發生之年齡與停經年齡研究中，研究者欲了解女性月經週期改



變與停經年齡此二事件發生年齡的相關性，發現在不同月經週期改變的年齡分組中，二事件發生時間的相關性明顯不同。Nan et al. (2006) 即提出分段常數交叉比例函數 (piecewise constant cross ratio) ，假設交叉比例函數隨著女性月經週期改變發生年齡改變。Li et al. (2017) 提出二元存活資料的分位數勝算比 (quantile odds ratio) ，研究心肌梗塞與中風事件時間在個別不同分位數下勝算比，發現勝算比隨分位數增加而下降，顯示較早發生心肌梗塞或中風事件之一的病人有較高風險會發生另一事件。

我們也可利用時間加速模型 (accelerated failure time model, AFT) 的概念來探討二元時間相關性。Yang and Peng (2016) 提出交叉殘差分位數比 (cross-quantile residual ratio) 用於研究半競爭風險資料，比較在特定時間點前發生非終止事件比上未發生終止事件兩組剩餘時間 (Residual time) 分位數。該方法因不須分配假設，所以較耦合(copula)架構更為彈性。本文延伸 Yang and Peng (2016) 交叉殘差分位數比至前兩段序列間隔時間存活資料，探討序列間隔時間之間的分位數相關性。期望獲得前段序列間隔時間對其後間隔時間長度之影響，並以交叉分位數比做表示。之所以採用交叉分位數比，是因為存活時間長度比能提供較直觀解釋，將使醫病溝通、以及醫療人員內部整合更加順暢。例如囊狀纖維化病人前 20%復發的病人，其下一次復發時間中位數，可能為後 80%復發病人的 0.5 倍，則病人即可了解到自己下一次可能的感染時間及早準備。

分位數迴歸提供二元存活時間彈性的相關性測量指標，無須額外假設，但相對的該方法需要大量的資料推論高分位數結果。而現代醫療健康登記檔提供大量且方便的長期追蹤資料，使本方法可作為初步分析之可行方法。

第二章 文獻回顧



本章首先回顧 Peng and Huang (2008) 發展應用於存活資料的分位數迴歸方法，接著回顧 Yang and Peng (2016) 所提出之交叉殘差分位數比 (cross-quantile residual ratio)，該方法利用分位數迴歸模型測量半競爭風險資料非終止事件時間與剩餘時間之相關性，本文參考其架構並延伸至有序間隔時間資料結構。再接著回顧 Li et al. (2014) 所提出二元存活時間分位數勝算比。最後回顧 Pang et al. (2012) 使用誘導平滑方法估計分位數迴歸模型變異數，本文使用誘導平滑方法在估計漸進共變異數。

第一節 分位數迴歸模型於存活資料

分位數迴歸模型為 $Q(\tau | \mathbf{Z}_i) = \exp(\mathbf{Z}_i^T \boldsymbol{\beta}(\tau))$, $\tau \in (0,1)$ ，

其中 $Q(\tau | \mathbf{Z}_i) = \inf \{t : \Pr(T \leq t | \mathbf{Z}) \geq \tau\}$ ， τ 為百分率。Peng and Huang (2008) 利用平賭過程(martingale process)估計迴歸係數。令 T 為存活時間， C 為設限時間， $\tilde{T} = \min(T, C)$ ， $\delta = I(T \leq C)$ 。 $\Lambda(t | \mathbf{Z}) = -\log(1 - \Pr(T \leq t | \mathbf{Z}))$ 為累積風險函數， $N(t) = I(\tilde{T} \leq t, \delta = 1)$ 為計數過程， $Y(t) = I(\tilde{T} > t)$ 為冒風險過程，

$M(t) = N(t) - \Lambda(t \wedge \tilde{T} | \mathbf{Z})$ 為平賭過程(martingale process)，

$H(x) = -\log(1 - x)$ ，作者利用平賭過程一階動差建構估計式。

因為 $E[M_i(t) | \mathbf{Z}_i] = 0$ ，對於所有 $t \geq 0$ ，

$$E \left[n^{-1/2} \sum_{i=1}^n N_i \left(\exp \{ \mathbf{Z}_i^T \boldsymbol{\beta}(\tau) \} \right) - \Lambda \left(\exp \{ \mathbf{Z}_i^T \boldsymbol{\beta}(\tau) \} \wedge \tilde{T} | \mathbf{Z}_i \right) \right] = 0$$

改寫



$$\Lambda\left(\exp\left\{\mathbf{Z}_i^T \boldsymbol{\beta}(\tau)\right\} \wedge \tilde{T} \mid \mathbf{Z}_i\right) = H(\tau) \wedge H\left(F_T\left(\tilde{T} \mid \mathbf{Z}_i\right)\right) = \int_0^{\tau} Y_i\left(\exp\left\{\mathbf{Z}_i^T \boldsymbol{\beta}(u)\right\}\right) dH(u)$$

所以可寫下估計式，

$$S_n(\boldsymbol{\beta}(\tau), \tau) = n^{-1/2} \sum_{i=1}^n \mathbf{Z}_i \left\{ N_i\left(\exp\left\{\mathbf{Z}_i^T \boldsymbol{\beta}(\tau)\right\}\right) - \int_0^{\tau} Y_i\left(\exp\left\{\mathbf{Z}_i^T \boldsymbol{\beta}(u)\right\}\right) dH(u) \right\} = 0$$

但隨機過程只有在事件發生時才會有機率跳動， $\boldsymbol{\beta}(\tau)$ 也才會隨之改變，為一右連續函數，所以將分位數切割成多段，而於高分位數時 $\boldsymbol{\beta}(\tau)$ 受設限影響而無法估計，因此將分位數改為 $\tau \in (0, \tau_U)$ ，並等量切割成 L 段將其離散化，使得 $\boldsymbol{\beta}(\tau)$ 只在每 $\tau_j, j=1, 2, \dots, L$ 分位數時有變化，且 $0 < \tau_L < \tau_1 < \dots < \tau_U < 1$ 。因此原估計方程式便修正為，

$$S_n(\boldsymbol{\beta}(\tau), \tau) = n^{-1/2} \sum_{i=1}^n \mathbf{Z}_i \left\{ N_i\left(\exp\left\{\mathbf{Z}_i^T \boldsymbol{\beta}(\tau_j)\right\}\right) - \sum_{k=0}^{j-1} Y_i\left(\exp\left\{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\tau_k)\right\}\right) \times \{H(\tau_{k+1}) - H(\tau_k)\} \right\}$$

但估計方程式為非連續函數，可能無法解根，因此利用 Fyngenson 及 Ritov (1994) 所提出的通解，極小化以下 L_1 形凸函數(L_1 -type convex function)，

$$l_j(\mathbf{b}) = \sum_{i=1}^n \left| \delta_i \log \tilde{T}_i - \delta_i \mathbf{b}^T \mathbf{Z}_i \right| + \left| M - \mathbf{b}^T \sum_{i=1}^n (-\delta_i \mathbf{Z}_i) \right| + \left| M - \mathbf{b}^T \sum_{i=1}^n \left(2 \mathbf{Z}_i \sum_{k=0}^{j-1} Y_i\left(\exp\left\{\mathbf{Z}_i^T \hat{\boldsymbol{\beta}}(\tau_k)\right\}\right) \times \{H(\tau_{k+1}) - H(\tau_k)\} \right) \right|$$

\mathbf{b} 是在分位數為 τ_j 時的回歸係數估計向量， M 為一極大數值(例如 10^7)使得後上式後兩項絕對值內為正。

第二節 交叉殘差分位數比(非單次事件)

Yang and Peng (2016) 提出交叉殘差分位數比 (cross quantile residual ratio) 用於研究半競爭風險資料中終端事件時間 (time to terminal event) 和非終端事件時間 (time to non-terminal event) 的關係，該方法比較在 t_0 時間點有無發生非終端事



件兩組的剩餘壽命分位數，

$$CQRR(\tau; t_0) = \frac{Q_\tau(T_2 - t_0 | T_2 > t_0, T_1 > t_0)}{Q_\tau(T_2 - t_0 | T_2 > t_0, T_1 \leq t_0)}, \tau \in (0, 1), t_0 > 0$$

$Q_\tau(T_2 - t_0 | T_2 > t_0, T_1 > t_0) = \inf \{t : Pr(T_2 - t_0 \leq t | T_2 > t_0, T_1 > t_0) \geq \tau\}$ ， T_1 為發生非終端事件時間， T_2 為死亡時間。 $CQRR(\tau | t_0)$ 越大，代表在 t_0 仍活著的人當中，尚未發生非終端事件的人比上發生非終端事件的人，其剩餘存活時間 τ th 分位數較長；反之亦然。在沒有左截切的情況下，半競爭風險觀察資料包含

(X, Y, δ, η) 。 C 為右設限時間， $X = \min(T_1, T_2, C)$ ， $Y = \min(T_2, C)$ 為觀察時間，

$\delta = I(T_1 < Y)$ 標記有無觀察到非終端事件發生， $\eta = I(T_2 < C)$ 標記有無觀察到終端

事件發生。當有左截切情況時，也就是只有在 $Y > L$ 時才可觀察到資料，觀察資

料符號表示為 $(X^*, Y^*, \delta^*, \eta^*, L^*)$ 。 $\tilde{A}(t_0) = (1, I(T_1 > t_0))^T$ 為模型中共變量，

$A^*(t_0) = (1, I(X^* > t_0))^T$ 為觀察共變量。

作者使用分位數回歸模型來估計 $CQRR(\tau | t_0)$ ，

$$Q_\tau(T_2 - t_0 | T_2 > t_0, I(T_1 > t_0)) = \exp\{\beta_0^{(1)}(\tau, t_0) + I(T_1 > t_0)\beta_0^{(2)}(\tau, t_0)\}$$

兩個參數意義分別為在 t_0 時已發生非終端事件組剩餘時間分位數

$\exp(\beta_0^{(1)}(\tau, t_0)) = Q_\tau(T_2 - t_0 | T_2 > t_0, T_1 \leq t_0)$ ，以及

$\exp(\beta_0^{(2)}(\tau, t_0)) = CQRR(\tau, t_0)$ 。

並使用 IPCW 調整具有完整訊息資料的貢獻，即 $(X, Y) = (\min(T_1, T_2), T_2)$ 且

$T_2 > t_0$ ，可觀察到

$$E\left(I(L^* \leq t_0)I(Y^* > t_0)\eta^* A^*(t_0)\left\{I[\log(Y^* - t_0) \leq A^{*T}(t_0)\mathbf{b}] - \tau\right\}\right) = E\left(\frac{I(T_2 > t_0)\tilde{A}(t_0)\left\{I[\log(T_2 - t_0) \leq \tilde{A}^T(t_0)\mathbf{b}] - \tau\right\}I(L \leq t_0, C > T_2)}{P(Y > L)}\right)$$



因此加入權重 $D(t_0, Y^*) = \frac{P(L \leq t_0, C > Y^*)}{P(Y > L)}$ ，假設 (L, C) 與 (T_1, T_2) 獨立，可推得

$$\begin{aligned} & E \left(\frac{I(L^* \leq t_0) I(Y^* > t_0) \eta^*}{D(t_0, Y^*)} A^*(t_0) \{ I[\log(Y^* - t_0) \leq A^{*T}(t_0) \mathbf{b}] - \tau \} \right) \\ &= E \left(I(T_2 > t_0) \tilde{A}(t_0) \{ I[\log(T_2 - t_0) \leq \tilde{A}^T(t_0) \mathbf{b}] - \tau \} \times \frac{E\{I(L \leq t_0, C > T_2) | T_1, T_2\}}{P(Y > L) D(t_0, T_2)} \right) \\ &= E \left(I(T_2 > t_0) \tilde{A}(t_0) \{ I[\log(T_2 - t_0) \leq \tilde{A}^T(t_0) \mathbf{b}] - \tau \} \right) = 0 \end{aligned}$$

依此建立估計式，

$$S_n(\mathbf{b}, \tau, t_0) = \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{I(L_i^* \leq t_0) I(Y_i^* > t_0) \eta_i^*}{\hat{D}(t_0, Y_i^*)} A_i^*(t_0) \{ I[\log(Y_i^* - t_0) \leq A_i^{*T}(t_0) \mathbf{b}] - \tau \} = 0$$

權重依條件機率拆解成以下式子，

$$\begin{aligned} D(s, t) &= \frac{P(L \leq s, C > t)}{P(Y > L)} = \frac{P(L \leq s < C) P(C > t | L \leq s < C)}{P(Y > L)} \\ &= \frac{P(L^* \leq s < Y^*) P(C > t | L \leq s < C)}{P(T_2 > s)} \end{aligned}$$

分別估計帶入。

$$P(L^* \leq s < Y^*) \text{ 可使用經驗估計 } n^{-1} \sum_{i=1}^n I(L_i^* \leq s < Y_i^*) ,$$

$$\hat{P}(T_2 > s) = \prod_{Y_j^* \leq s} \left\{ 1 - \frac{\sum_{i=1}^n I(Y_i^* = Y_j^*, \eta_i^* = 1)}{\sum_{i=1}^n I(L_i^* \leq Y_j^* \leq Y_i^*)} \right\} ,$$

$S(t | s) = P(C > t | L \leq s < C) = \exp\{-\Lambda(t | s)\}$ 使用 Nelson-Aalen 類型估計法，

$$\Lambda(t | s) = \int_s^t \frac{W_s(du)}{C_s(du)} , \quad W_s(u) = \frac{1}{n} \sum_{i=1}^n I(L_i^* \leq s < Y_i^* \leq u, \eta_i^* = 0) ,$$

$$C_s(u) = \frac{1}{n} \sum_{i=1}^n I(L_i^* \leq s < u \leq Y_i^*) .$$

因 $S_n(\mathbf{b}, \tau, t_0)$ 為非連續函數，可能無法解根，因此利用 Fygenon 及 Ritov (1994) 所提出的通解，極小化以下 L_1 型凸函數(L_1 -type convex function)，

$$U_n(\mathbf{b}, \tau, t_0) = \sum_{i=1}^n I(L_i^* \leq t_0) I(Y_i^* > t_0) \eta_i^* \left| \frac{\log(Y_i^* - t_0)}{\hat{D}(t_0, Y_i^*)} - \mathbf{b}^T \frac{A_i^*(t_0)}{\hat{D}(t_0, Y_i^*)} \right| \\ + |M - (2\tau - 1) \mathbf{b}^T \sum_{i=1}^n I(L_i^* \leq t_0) I(Y_i^* > t_0) \eta_i^* \frac{A_i^*(t_0)}{\hat{D}(t_0, Y_i^*)}|$$

M 為一極大數值(例如 10^7)使得後上式後兩項絕對值內為正。

第三節 分位數勝算比

Li et al. (2014) 提出分位數勝算比，利用勝算比表達相關性程度，計算隨分位數改變的區域相關性。 X_1 、 X_2 為二元連續變數， \mathbf{Z} 為共變數，

$F_j(X) = \Pr(X_j \leq x)$, $j=1, 2$ 為累積發生函數，

$Q_j(\tau | \mathbf{Z}) = \inf\{x: F_j(x | \mathbf{Z}) \geq \tau\}$, $\tau \in (0, 1)$, $j=1, 2$ 為給定共變數下的條件分位

數。我們定義分位數勝算比為以下

$$qor(\tau | \mathbf{Z}) = \frac{\text{odds}\{X_1 \leq Q_1(\tau | \mathbf{Z}) | X_2 \leq Q_2(\tau | \mathbf{Z}) | \mathbf{Z}\}}{\text{odds}\{X_1 \leq Q_1(\tau | \mathbf{Z}) | X_2 > Q_2(\tau | \mathbf{Z}) | \mathbf{Z}\}}$$

整理一下即知分位數勝算比相當於一致配對機率除以不一致配對的機率。

$$qor(\tau | \mathbf{Z}) = \frac{\Pr(X_1 \leq Q_1(\tau | \mathbf{Z}), X_2 \leq Q_2(\tau | \mathbf{Z}) | \mathbf{Z}) \times \Pr(X_1 > Q_1(\tau | \mathbf{Z}), X_2 > Q_2(\tau | \mathbf{Z}) | \mathbf{Z})}{\Pr(X_1 \leq Q_1(\tau | \mathbf{Z}), X_2 > Q_2(\tau | \mathbf{Z}) | \mathbf{Z}) \times \Pr(X_1 > Q_1(\tau | \mathbf{Z}), X_2 \leq Q_2(\tau | \mathbf{Z}) | \mathbf{Z})}$$

Li et al. (2017) 延續以上想法將 $qor(\tau | \mathbf{Z})$ 應用於右設限資料。令 T_1 、 T_2 為二元存活時間， $(\{\tilde{T}_{ij}, \delta_{ij}, j=1, 2\}, C_{ij}, \mathbf{Z}_i)$ 為觀察資料， $i=1, \dots, n$ 為個體指標，

$j=1, 2$ 為事件指標， C_{ij} 為個體 i 之右設限時間， $\tilde{T}_{ij} = \min(T_{ij}, C_i)$ 為觀察時間，

$\eta_{ij} = I(T_{ij} \leq C_{ij})$ 為設限指標，定義條件於共變數之存活耦合(survival copula)為



$$P(\tau) = \Pr(T_1 > Q_1(\tau | \mathbf{Z}), T_2 > Q_2(\tau | \mathbf{Z}) | \mathbf{Z})$$

存活耦合與分位數勝算比有一單調轉換關係 $qor(\tau) = \phi\{P(\tau); \tau\}$,

$$\phi\{x; \tau\} = \frac{x(2\tau + x - 1)}{(1 - \tau - x)^2} , \text{ 因此可透過帶入存活耦合估計值估計二元時間分位數勝算}$$

比 $\phi\{\hat{P}(\tau); \tau\}$ 。上述無母數估計適合資料量多且估計單一分位數相關性，若資料量少且欲估計一段區間的分位數相關性，則使用半參數化估計。假定

$$\log qor(\tau) = f(\tau, \beta_0), \tau \in [m_L, m_U] , f(\tau, \beta) = \beta_0^{(0)} + \sum_{k=1}^p \beta_0^{(k)} \tau^k , \text{ 利用}$$

$$\int_{m_L}^{m_U} W(\tau) (\hat{P}(\tau) - \phi^{-1}[\exp\{f(\tau, \beta)\}])^2 d\tau \text{ 解根, } W(\tau) \text{ 為權重函數。以上兩種方式}$$

皆需帶入存活耦合估計值，存活耦合估計可分為無母數及條件於共變數兩種方式。實務上經常研究同一個體不同事件相關性，此時兩事件通常有相同的設限時間，即 $C_1 = C_2$ 。Lin et al. (1993) 利用設限權重調整設限問題，

$$\hat{P}(\tau) = \frac{1}{n} \sum_{i=1}^n \frac{I(\tilde{T}_{i1} > Q_1(\tau), \tilde{T}_{i2} > Q_2(\tau))}{\hat{G}(\tilde{T}_{i1} \vee \tilde{T}_{i2})}$$

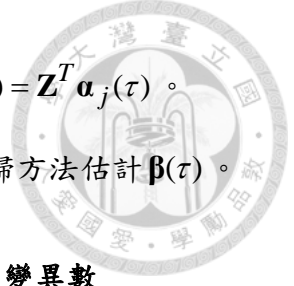
此方法適用單一設限時間， $G(\cdot)$ 是設限時間 C 的存活函數， $\hat{G}(\cdot)$ 是其 Kaplan-Meier 估計量，為二元觀察時間，上式中 $Q_1(\tau)$ 與 $Q_2(\tau)$ 可透過 1 減去 KM 法估計邊際存活函數估計累積發生函數估計帶入。Prentice and Cai (1992) 利用平賭過程 (martingale process) 估計二元存活時間共變異數與聯合存活函數，可適用更廣泛的設限情境，

$$\hat{P}(\tau) = (1 - \tau)^2 \hat{L}(\tau)$$

$$\text{其中 } L(\tau) = \frac{\Pr(U_1 > \tau, U_2 > \tau)}{\Pr(U_1 > \tau) \Pr(U_2 > \tau)} , U_1 = F_1(T_1 | \mathbf{Z}) , U_2 = F_2(T_2 | \mathbf{Z}) , \hat{L}(\tau) \text{ 透過}$$

Prentice and Cai (1992) 的方式帶入次分佈 $H_{nlk}(\tau, \hat{\alpha}_1, \hat{\alpha}_2)$ 估計。

$$H_{nlk}(\tau) = n^{-1} \sum_{i=1}^n I\{\tilde{T}_{i1} \leq Q_1(\tau | Z_i), \tilde{T}_{i2} \leq Q_2(\tau | Z_i), \eta_{i1} = l, \eta_{i2} = k\} , l, k \in \{0, 1\}$$



其中 $Q_j(\tau | Z_i), j=1,2$ 帶入分位數迴歸模型估計值， $\log Q_j(\tau | \mathbf{Z}) = \mathbf{Z}^T \boldsymbol{\alpha}_j(\tau)$ 。

可使用例如 Portnoy (2003) 或 Peng and Huang (2008) 等分位數迴歸方法估計 $\boldsymbol{\beta}(\tau)$ 。

第四節 使用誘導平滑方法估計分位數迴歸模型的變異數

Pang et al. (2012) 採用了 Brown and Wang (2005) 的誘導平滑方法 (induced smoothing)，將其延伸至設限資料之分位數迴歸係數的變異數估計。

令 T 為事件時間， C 為設限時間， $\tilde{T} = \min(T, C)$ 為觀察時間， $\delta = I(T \leq C)$

為發生事件指標函數， \mathbf{Z} 為共變量矩陣， $Q(\tau | \mathbf{Z}) = \inf \{t : \Pr(T \leq t | \mathbf{Z}) \geq \tau\}$

假設無設限時，利用動差估計法可建構估計式

$$n^{-1} \sum_{i=1}^n \mathbf{Z}_i \left\{ I(\log T_i \leq \mathbf{Z}_i^T b) - \tau \right\} = 0,$$

其同樣相當於求 $n^{-1} \sum_{i=1}^n r_i(b(\tau)) \left\{ \tau - I(r_i(b(\tau)) < 0) \right\}$ 的最小值，

$r_i(b) = \log T_i - \mathbf{Z}_i^T b$ 。而有設限時，可使用設限機率倒數權重 (IPCW) 調整設限

問題。假設設限時間 C 與事件時間 T 在給定共變量 \mathbf{Z} 之下條件獨立，

$$\begin{aligned} E \left\{ \frac{I(\tilde{T} \leq t, \delta = 1)}{G(\tilde{T} | \mathbf{Z})} \middle| \mathbf{Z} \right\} &= E \left[E \left\{ \frac{I(T \leq t, C \geq T)}{G(T | \mathbf{Z})} \middle| T, \mathbf{Z} \right\} \middle| \mathbf{Z} \right] \\ &= E \left\{ \frac{I(T \leq t) G(T | \mathbf{Z})}{G(T | \mathbf{Z})} \middle| \mathbf{Z} \right\} = \Pr(T \leq t) \end{aligned}$$

因此在有設限情況下，估計式可改寫成以下，

$$S_n(\mathbf{b}, \tau) = n^{-1} \sum_{i=1}^n \frac{\delta_i \mathbf{Z}_i}{\hat{G}(\tilde{T}_i)} \left\{ I(\log \tilde{T}_i \leq \mathbf{Z}_i^T b) - \tau \right\} = 0$$

其中 $\hat{G}(\cdot)$ 可用 KM 估計帶入。

Bang and Tsiatis (2002) 證明估計值 $\hat{\boldsymbol{\beta}}$ 是一致的且漸近常態的，

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} N(0, \Xi), \quad \Xi = H^{-1} \Sigma H^{-1}, \quad \Sigma = \lim_{n \rightarrow \infty} \text{cov}(n^{1/2} S_n(\beta, \tau)),$$

$$H = \lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n Z_i Z_i^T f_i(0 | Z_i), \quad \text{其中 } f_i(\cdot | Z_i) \text{ 為未知機率密度函數。}$$

$$\text{由於 } \hat{\beta} \text{ 為漸近常態，因此 } \hat{\beta} = \beta + \Gamma^{1/2} V, \quad \Gamma = n^{-1} \Xi, \quad V \sim N(0, I_p)$$

依此建立平滑估計式

$$\tilde{S}_n(\mathbf{b}, \tau, \Gamma) = E_V \left\{ S_n(\mathbf{b} + \Gamma^{1/2} V, \tau) \right\} = n^{-1} \sum_{i=1}^n \frac{\delta_i \mathbf{Z}_i}{\hat{G}(\tilde{T}_i)} \left\{ \Phi \left(\frac{\mathbf{Z}_i^T \mathbf{b} - \log \tilde{T}_i}{\sqrt{\mathbf{Z}_i^T \Gamma \mathbf{Z}_i}} \right) - \tau \right\}$$

ϕ 以及 Φ 為標準常態的機率密度函數與累積分布函數。

因 Γ 未知，文章使用正定矩陣 $\tilde{\Gamma}_0 = O(n^{-1})$ 作為初始值，然後通過迭代更新以估計 Γ ，從而得到 Ξ 。

$$\text{令 } \hat{\Sigma}(\mathbf{b}) = n^{-1} \sum_{i=1}^n \hat{\zeta}_i^{\otimes 2},$$

$$\hat{\zeta}_i = \frac{\delta_i \mathbf{Z}_i}{\hat{G}(\tilde{T}_i)} \left\{ I(\log \tilde{T}_i \leq \mathbf{Z}_i^T \mathbf{b}) - \tau \right\} \\ - I(\delta_i = 0) \left\{ \frac{\sum_{j=1}^n I(\tilde{T}_j \geq \tilde{T}_i) \frac{\delta_j}{\hat{G}(\tilde{T}_j)} \left\{ I(\log \tilde{T}_i \leq \mathbf{Z}_i^T \mathbf{b}) - \tau \right\}}{\sum_{j=1}^n I(\tilde{T}_j \geq \tilde{T}_i)} \right\}^{-1}$$

$$\text{I. } \tilde{A}_n(\mathbf{b}, \Gamma) = \frac{\partial \tilde{S}_n(\mathbf{b}, \Gamma)}{\partial \mathbf{b}} = n^{-1} \sum_{i=1}^n \frac{\delta_i \mathbf{Z}_i}{\hat{G}(\tilde{T}_i)} \left\{ \phi \left(\frac{\mathbf{Z}_i^T \mathbf{b} - \log \tilde{T}_i}{\sqrt{\mathbf{Z}_i^T \Gamma \mathbf{Z}_i}} \right) \right\} \frac{\mathbf{Z}_i^T}{\sqrt{\mathbf{Z}_i^T \Gamma \mathbf{Z}_i}}$$

給定初始值 $\tilde{\Gamma}_0 = 1/n$

$$\text{II. } \tilde{\beta}_k = \tilde{\beta}_{k-1} + \left\{ -\tilde{A}_n(\tilde{\beta}_{k-1}, \tilde{\Gamma}_{k-1}) \right\}^{-1} \tilde{S}_n(\tilde{\beta}_{k-1}, \tilde{\Gamma}_{k-1})$$

$$\tilde{\Xi}^k = \left\{ \tilde{A}_n(\tilde{\beta}_k, \tilde{\Gamma}_{k-1}) \right\}^{-1} \hat{\Sigma}(\tilde{\beta}_k) \left\{ \tilde{A}_n(\tilde{\beta}_k, \tilde{\Gamma}_{k-1}) \right\}^{-1}$$

重複步驟 II. 直至收斂。

第三章 方法



病人疾病進展歷程包含著病人經歷不同疾病階段之相關醫療資訊，如囊狀纖維化病人，由於肺部感染會形成結痂組織，進而降低肺功能，因此可以合理地假設每次肺部感染嚴重性都會影響下一次的感染。rhDNase 資料(Fuchs et al. 1994)亦顯示 rhDNase 治療對於延緩第一次肺部感染有顯著效果，但對於第二次感染則無明顯效果，且效果似乎隨時間而改變(於 80 天交叉)。用力呼氣容積(FEV)同樣對於第一次肺部感染有很好的解釋力，但對於第二次感染則沒有很好的解釋效果。推測可能因為，在追蹤期間內(約 169 天)經歷第一次感染之族群為篩選過之群體，且肺部惡化程度較嚴重的一群病患。由於肺部感染會形成結痂組織，進而降低肺功能，因此可以合理地假設每次肺部感染嚴重性都會影響下一次的感染。所以在條件於經歷第一次感染群體中，觀察治療對第二次感染效果，應加入前一次感染情形分析。在缺乏病人年齡、肺功能等資料下，此時上一段治療間隔時間包含病程資訊即可作為良好異質性指標。

本文欲延伸 Yang and Peng (2016) 應用於半競爭風險資料的交叉殘差分位數比至前兩段有序間隔時間存活資料，探討第一次間隔時間及第二次間隔時間兩兩相關性。符號定義如下：

下標 i 為個體指標， j 為序列的指標，定義完整資料如下：

T_j : 第 i 個人從起始事件到第 j 次事件時間，稱作個體 i 的第 j 次事件時間。

$Y_j = T_j - T_{j-1}$: 第 i 個人從第 $j-1$ 次事件到第 j 次事件的間隔時間，稱作個體 i 的第 j 次間隔時間。

右設限觀察資料如下：

C : 第 i 個人的右設限時間。

$\tilde{T}_j = \min(T_j, C_i)$: 第 i 個人第 j 次觀察到事件時間。



$\tilde{Y}_j = \min(Y_j, C - \tilde{T}_{j-1})$: 第 i 個人觀察到第 j 次間隔時間。

$\delta_j = I(Y_j \leq C - \tilde{T}_{j-1})$: 第 i 個人觀察到第 j 次事件的指標函數， $j=1,2$ 。

$Q_{\tau_1} = \inf \{t: \Pr(Y_1 \leq t) \geq \tau_1\}$: 第一次間隔時間累積到 τ_1 百分率的時間長度。

$Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1}) = \inf \{t: \Pr(Y_2 \leq t | Y_1 > Q_{\tau_1}) \geq \tau_2\}$: 給定第一次間隔時間長度大

於 τ_1 分位數，第二次間隔時間累積到 τ_2 百分率的時間長度。

第一節 交叉分位數比

我們利用時間長度比測量兩段間隔時間相關性，其直觀的想法為以第一段間隔時間做分組，若第二段間隔時間差異越大，代表以第一段間隔時間做分組之組別間異質性很高，也代表說兩段間隔時間相關性越高。定義，交叉分位數比(cross quantile ratio, CQR)，

$$CQR_{Y_1, \tau_1}(\tau_2) = \frac{Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1})}{Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})},$$

是在給定第一次間隔時間長度大於或小於 τ_1 分位數，第二次間隔時間累積到 τ_2 百分率的時間長度比。此相關性測量方法能夠發現隨時間改變相關性，且提供直觀地時間長度比解釋，不須過多模式假設。若兩段間隔時間呈現正相關，則

$CQR_{Y_1, \tau_1}(\tau_2) > 1$ ，代表 $Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1}) > Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})$ ，也就是說較晚發生

第一次事件病人($Y_1 > Q_{\tau_1}$)後續發生第二次事件間隔時間也較長。若兩段間隔時間

呈現負相關，則 $CQR_{Y_1, \tau_1}(\tau_2) < 1$ ，也就是說較晚發生第一次事件病人($Y_1 > Q_{\tau_1}$)後

續第二次事件間隔時間較短。若兩段間隔時間獨立，則 $CQR_{Y_1, \tau_1}(\tau_2) = 1$ 。

$CQR_{Y_1, \tau_1}(\tau_2)$ 可視為兩組條件機率反函數的相除。令

$\tau_2^{(1)} = F_{Y_2}(Y_2 \leq y_2 | Y_1 > Q_{\tau_1})$ 為條件於第一次間隔時間長度大於 τ_1 分位數而第二次間隔時間長度小於 y_2 的機率， $\tau_2^{(2)} = F_{Y_2}(Y_2 \leq y_2 | Y_1 \leq Q_{\tau_1})$ 為條件於第一次間隔時間長度小於 τ_1 分位數而第二次間隔時間長度小於 y_2 的機率，則我們可知 $\tau_2^{(1)} / \tau_2^{(2)}$ 為給定不同條件，累積至 y_2 時間的條件機率相除，相當於流行病學中的相對風險 (relative risk, RR)，即暴露與否的兩組在一定時間內發生事件之風險比。

$CQR_{Y_1, \tau_1}(\tau_2)$ 取條件機率反函數，即為條件時間長度。在控制兩組發生比例 (incidence proportion) 皆為 τ_2 之下，所需時間長度比。 $CQR_{Y_1, \tau_1}(\tau_2)$ 使用時間分位數比可較為彈性地測量兩段時間相關性，如圖 1，兩組的間隔時間差異隨著累積機率越來越大，甚至出現翻轉的情況。而 $CQR_{Y_1, \tau_1}(\tau_2)$ 會隨著分量 τ_2 而改變，意即 $CQR_{Y_1, \tau_1}(\tau_2)$ 可捕捉與時間有關相依模式。在缺乏第一次間隔時間分組依據之文獻資料下或希望彈性地探索兩段間隔時間相關性，使用分位數作為第一段間隔時間分組依據不失為一良好方式。但相對上缺點就是資料驅動，在臨床意義上解釋需更多資料佐證，所以若有具有臨床意義上的時間切點，則亦可作為第一段間隔時間分組依據。因有序間隔時間資料有先後順序特性，使 $CQR_{Y_1, \tau_1}(\tau_2)$ 可作為一量化預測指標。臨床上我們希望預測病人下一次事件發生時間，讓病人與家屬

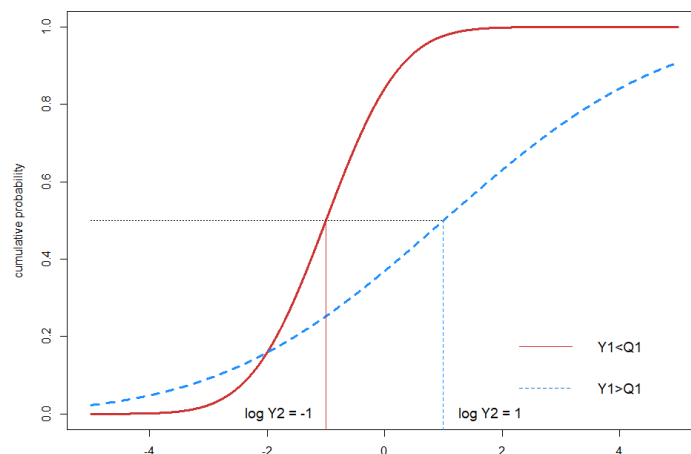


圖 1、分位數迴歸之示意圖

可事先做好準備。而間隔時間長度可視為病人健康狀況、生活環境以及照護等綜合資訊，因此前一事件發生間隔時間是下一間隔時間之可能預測因子。



第二節 有序間隔時間交叉分位數比估計

首先探討前兩段有序間隔時間 Y_1 與 Y_2 之交叉分位數比 (cross quantile ratio, CQR)。基本想法是在給定第一段時間不同的分位數下，比較第二段復發時間的分位數長度，即估計，

$$CQR_{Y_1, \tau_1}(\tau_2) = \frac{Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1})}{Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})},$$

另外定義 $\mathbf{Z}(\tau_1) = \{1, I(Y_1 > Q_{\tau_1})\}$ ， $\mathbf{Z}^*(\tau_1) = \{1, I(\tilde{Y}_1 > Q_{\tau_1})\}$ ，可重新改寫為

$$Q_{\tau_2}(Y_2 | I(Y_1 > Q_{\tau_1})) = \exp\{\mathbf{Z}^T(\tau_1)\boldsymbol{\beta}(\tau_1, \tau_2)\}$$

其中 $\boldsymbol{\beta}(\tau_1, \tau_2) = \{\beta_0(\tau_1, \tau_2), \beta_1(\tau_1, \tau_2)\}$ ，因 $I(Y_1 > Q_{\tau_1})$ 為唯一共變量，參考 Yang and Peng (2016) 方法建構一工作分位數迴歸模型 (working quantile regression

model)，得到估計 $\beta_1(\tau_1, \tau_2) = \log Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1}) - \log Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})$ ，即為估計 $\log CQR_{Y_1, \tau_1}(\tau_2)$ 。因此兩參數分別為， $\exp\{\beta_0(\tau_1, \tau_2)\} = Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})$ ， $\exp\{\beta_1(\tau_1, \tau_2)\} = CQR_{Y_1, \tau_1}(\tau_2)$ 。

根據以下式子可發現，若單純使用完整觀察 (δ_1, δ_2) 之資料來估計參數 $\boldsymbol{\beta}$ 可發現其並非不偏估計式，

$$\begin{aligned} & E\left\{\delta_2 \mathbf{Z}^{*T}(\tau_1) \left(I\left[\log \tilde{Y}_2 \leq \mathbf{Z}^{*T}(\tau_1)\boldsymbol{\beta}(\tau_1, \tau_2)\right] - \tau_2 \right)\right\} \\ &= E\left\{I(Y_2 \leq C - \tilde{T}_1) \mathbf{Z}^{*T}(\tau_1) \left(I\left[\log Y_2 \leq \mathbf{Z}^T(\tau_1)\boldsymbol{\beta}(\tau_1, \tau_2)\right] - \tau_2 \right)\right\} \\ &= E\left\{I(Y_2 \leq C - T_1) I(T_1 \leq C) \mathbf{Z}^T(\tau_1) \left(I\left[\log Y_2 \leq \mathbf{Z}^T(\tau_1)\boldsymbol{\beta}(\tau_1, \tau_2)\right] - \tau_2 \right)\right\} \\ &= E\left\{I(C \geq T_2) \mathbf{Z}^T(\tau_1) \left(I\left[\log Y_2 \leq \mathbf{Z}^T(\tau_1)\boldsymbol{\beta}(\tau_1, \tau_2)\right] - \tau_2 \right)\right\} \end{aligned}$$



因此加入設限機率倒數權重來建構估計式，如以下式子所示，

$$\begin{aligned}
& E \left\{ \frac{\delta_2}{P(C \geq \tilde{T}_2)} \mathbf{Z}^{*T}(\tau_1) \left(I[\log \tilde{Y}_2 \leq \mathbf{Z}^{*T}(\tau_1) \boldsymbol{\beta}(\tau_1, \tau_2)] - \tau_2 \right) \right\} \\
&= E \left\{ \frac{I(Y_2 \leq C - \tilde{T}_1)}{P(C \geq T_2)} \mathbf{Z}^T(\tau_1) \left(I[\log Y_2 \leq \mathbf{Z}^T(\tau_1) \boldsymbol{\beta}(\tau_1, \tau_2)] - \tau_2 \right) \right\} \\
&= E \left\{ \mathbf{Z}^T(\tau_1) \left(I[\log Y_2 \leq \mathbf{Z}^T(\tau_1) \boldsymbol{\beta}(\tau_1, \tau_2)] - \tau_2 \right) \times \frac{E[I(C \geq T_2) | T_1, T_2]}{P(C \geq T_2)} \right\} \\
&= E \left\{ \mathbf{Z}^T(\tau_1) \left(I[\log Y_2 \leq \mathbf{Z}^T(\tau_1) \boldsymbol{\beta}(\tau_1, \tau_2)] - \tau_2 \right) \right\} = 0
\end{aligned}$$

因此，我們可使用動差估計法獲得漸進不偏估計方程式 $S_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2)$ ，

$$S_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2) = n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \mathbf{b}(\tau_1, \tau_2)) - \tau_2 \right\} = 0,$$

G 為 C 之存活函數， \hat{G} 為根據資料所估計之 Kaplan-Meier 估計值，

$\hat{\mathbf{Z}}^*(\tau_1) = \{1, I(\tilde{Y}_1 > \hat{Q}_{\tau_1})\}$ ，其中 $\hat{Q}_{\tau_1} = \inf(t: \hat{P}(\tilde{Y}_1 \leq t) \geq \tau_1)$ ， $\hat{P}(Y_1 \leq t)$ 為根據資料所估計之 Kaplan-Meier 估計值。

由於 $S_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2)$ 為非連續函數，不一定能找到根。所以改求下列凸函數最小值，

$$\begin{aligned}
U_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2) &= \sum_{i=1}^n I(\delta_{2i} = 1) \left| \frac{\log \tilde{Y}_{2i}}{\hat{G}(\tilde{T}_{2i})} - \mathbf{b}^T(\tau_1, \tau_2) \frac{\hat{\mathbf{Z}}_i^{*T}(\tau_1)}{\hat{G}(\tilde{T}_{2i})} \right| \\
&+ \left| M - (2\tau - 1) \mathbf{b}^T(\tau_1, \tau_2) \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) I(\delta_{2i} = 1)}{\hat{G}(\tilde{T}_{2i})} \right|
\end{aligned}$$

其中 M 為極大的數使得 $M - (2\tau - 1) \mathbf{b}^T(\tau_1, \tau_2) \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) I(\delta_{2i} = 1)}{\hat{G}(\tilde{T}_{2i})}$ 恆大於 0。可用

Barrodale-Roberts 演算法找尋 $U_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2)$ 最小值。

第三節 交叉分位數比估計量之共變異數估計



本節首先說明交叉分位數比估計量 $\hat{\beta}(\tau_1, \tau_2)$ 的一致性(consistency)，接著提供二種 $\hat{\beta}(\tau_1, \tau_2)$ 的變異數估計方式，第一種為大樣本推導之變異數估計，第二種為重新取樣變異數估計。

3.3.1 大樣本性質

給定以下條件，

C1. 存在 $v > 0$ 使得 $P(C = v) > 0$ 和 $P(C > v) = 0$ 。 $0 < \tau_L < \tau_U < 1$

C2. $Z(\tau_1)$ 為均勻有界(uniformly bounded)。

C3. $\beta(\tau_1, \tau_2)$ 在 $\tau \in [\tau_L, \tau_U]$ 是利普希茨連續(Lipschitz continuity)，

$$0 < \tau_L < \tau_U < 1;$$

$f(t | Z(\tau_1))$ 是連續的並且在給定 τ_1 、 $Z(\tau_1)$ 、 t 為均勻上有界(bounded

above uniformly)， $f(t | Z(\tau_1)) = dF(t | Z(\tau_1)) / dt$ ，

$$F(t | Z(\tau_1)) = E[I(Y_2 < t) | Z(\tau_1)]$$

C4. 對於某些常數 $\rho_0 > 0$ 且 $c_0 > 0$ ，有 $\inf_{\{b \in B(\rho_0), \tau_1 \in [\tau_{1L}, \tau_{1U}], \tau_2 \in [\tau_{2L}, \tau_{2U}]\}} \text{eig min } H(b, \tau_1) \geq c_0$ 。

$$B(\rho_0) = \left\{ b \in R^2, \inf_{\tau_1 \in [\tau_{1L}, \tau_{1U}], \tau_2 \in [\tau_{2L}, \tau_{2U}]} \|b - \beta(\tau_1, \tau_2)\| \leq \rho_0 \right\},$$

$$H(b, \tau_1) = E \left[\mathbf{Z}(\tau_1)^{\otimes 2} f(\exp(\mathbf{Z}^T(\tau_1)b) | \mathbf{Z}^T(\tau_1)) \exp(\mathbf{Z}^T(\tau_1)b) \right]$$

另外定義以下符號， $h_1(y_1, y_2) = \partial P(\log Y_1 \leq y_1, \log Y_2 \leq y_2) / \partial y_1$ 為聯合分布對 y_1 微分函數，

$f_1(y_1) = dP(\log Y_1 \leq y_1) / dy_1$ 為 y_1 邊際機率密度函數，



$$\mu'_C(\log Q_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) = -h_1(\log Q_{\tau_1}, \beta_0(\tau_1, \tau_2) + \beta_1(\tau_1, \tau_2)) ,$$

$$\mu'_D(\log Q_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) = h_1(\log Q_{\tau_1}, \beta_0(\tau_1, \tau_2)) , \quad \mu'_E(\log Q_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) = -f_1(\log Q_{\tau_1}) .$$

$\iota_i(\tau_1, \tau_2) = \xi_{1,i} - \xi_{2,i} + \xi_{3,i}$ 為估計式 $S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2))$ 拆解成 i.i.d. 的漸進形式表示，其中，

$$\xi_{1,i} \equiv \left[\frac{\mu'_C + \mu'_D}{\mu'_C - \mu'_E} \right] \times f_1^{-1}(\log Q_{\tau_1}) \left[I(\log Y_{1,i} \leq \log Q_{\tau_1}) - \tau_1 \right] , \quad \xi_{2,i} \equiv \int_0^\infty \omega(\boldsymbol{\beta}, \tau_1, \tau_2) \frac{dM_i^G(s)}{y(s)}$$

$$\xi_{3,i}(\tau_1, \tau_2) \equiv \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \boldsymbol{\beta}(\tau_1, \tau_2)) - \tau_2 \right\} .$$

有以下理論：

A. 給定 C1~C4 的條件，

$$\lim_{n \rightarrow \infty} \sup_{\boldsymbol{\tau} \in [\boldsymbol{\tau}_L, \boldsymbol{\tau}_U]} \left\| \hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2) \right\| \xrightarrow{p} 0$$

B. 給定 C1~C4 的條件，

$\sqrt{n} \left\{ \hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2) \right\}$ 會弱收斂 (weak convergence) 至一期望值為零的

高斯過程 (Gaussian Process)，其共變異數矩陣 $\Omega(\tau_1, \tau_2)$

$$= H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1)^{-1} E \left[\iota(\tau_1, \tau_2) \iota(\tau_1, \tau_2)^T \right] \left[H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1)^{-1} \right]^T , \quad \boldsymbol{\tau} \in [\boldsymbol{\tau}_L, \boldsymbol{\tau}_U]$$

定理 A 說明交叉分位數比估計量 $\hat{\boldsymbol{\beta}}(\tau_1, \tau_2)$ 在 $\boldsymbol{\tau} \in [\boldsymbol{\tau}_L, \boldsymbol{\tau}_U]$ 中是一致的。定理 B 提

供 $\hat{\boldsymbol{\beta}}(\tau_1, \tau_2)$ 的漸近分佈封閉表達式 (closed-form)。證明詳見附錄 B。

3.3.2 共變異數估計

由於 $\Omega(\tau_1, \tau_2)$ 包含未知的機率密度函數 $f(t|Z(\tau_1))$ ，一般的作法為使用拔

靴法 (bootstrapping) 或重新取樣法 (resampling) 直接估計變異數，而免去估計

未知機率密度函數，但這兩種方法都需要重複抽樣而相對耗費時間。本文參考

Li et al. (2014) 使用誘導平滑方法 (induced smoothing) 將估計式

$S_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2)$ 拆成 i.i.d. 之結構並估計其共變異數，再參考 Peng and Fine

(2009) 的一致插入估計法 (plug-in estimate) 來估計估計值 $\hat{\beta}(\tau_1, \tau_2)$ 共變異數矩

陣。

由上節可得以下漸進等式，

$$n^{1/2} \left(\hat{\beta}(\tau_1, \tau_2) - \beta(\tau_1, \tau_2) \right) \approx -H(\beta(\tau_1, \tau_2), \tau_1)^{-1} S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) ,$$

簡易步驟如下，詳細過程請見附錄 C。

1. 估計 $n^{1/2} \left\{ \log \hat{Q}_{\tau_1} - \log Q_{\tau_1} \right\}$ 變異數矩陣。
2. 估計 $S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) \approx n^{-1/2} \sum_{i=1}^n \left\{ \xi_{1,i} - \xi_{2,i} + \xi_{3,i} \right\} \equiv n^{-1/2} \sum_{i=1}^n \iota_i$ 。
3. 估計 $H(\beta(\tau_1, \tau_2), \tau_1)^{-1} \approx \sqrt{n} \Lambda_n(\tau_1, \tau_2) E_n^{-1}(\tau_1, \tau_2)$ 。

3.3.3 重新取樣變異數估計

另外介紹 Jin et al. (2001) 提出之擾動最小化函數的重新取樣方法(resampling method by perturbing the minimand)，簡稱重新取樣法。重新取樣法的概念如下：

令 ψ_1, \dots, ψ_n 抽自標準指數分布之一隨機樣本，重複抽樣 M 次，在給定觀察資料下，隨機擾動估計式，

$$V_n(\mathbf{b}(\tau_1, \tau_2), \tau_1, \tau_2) = \sum_{i=1}^n I(\delta_{2i} = 1) \left| \frac{\psi_i \log \tilde{Y}_{2i}}{\hat{G}(\tilde{T}_{2i})} - \mathbf{b}^T(\tau_1, \tau_2) \frac{\psi_i \hat{\mathbf{Z}}_i^{*T}(\tau_1)}{\hat{G}(\tilde{T}_{2i})} \right| + \left| M - (2\tau - 1) \mathbf{b}^T(\tau_1, \tau_2) \sum_{i=1}^n \frac{\psi_i \hat{\mathbf{Z}}_i^*(\tau_1) I(\delta_{2i} = 1)}{\hat{G}(\tilde{T}_{2i})} \right|$$

，將得出之估計值記作 $\boldsymbol{\beta}^*(\tau_1, \tau_2)$ 。 $n^{1/2}(\boldsymbol{\beta}^*(\tau_1, \tau_2) - \hat{\boldsymbol{\beta}}(\tau_1, \tau_2))$ 在給定資料下與 $n^{1/2}(\hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2))$ 有相同的漸進分布，因此可透過 $\boldsymbol{\beta}^*(\tau_1, \tau_2)$ 的樣本變異數或者經驗百分位數 (empirical percentile) 建構 $\hat{\boldsymbol{\beta}}(\tau_1, \tau_2)$ 信賴區間。



第四章 模擬

本章根據三種阿基米德耦合(Archimedean Copula: Clayton, Frank, Gumbel)、對數常態模型(Lognormal)以及分段常數克萊頓五種不同的耦合架構生成資料,比較不同樣本數、設限率以及相關性下估計值相對偏誤以及三種標準差估計法比較,此三種方法分別為常見的拔靴法(bootstrapping)、3.3.3所簡介之重新取樣法(resampling)以及3.3.2所簡介之大樣本估計(Asymptotic)。事件時間以及設限時間皆來自韋伯分布, $T_i \sim weibull(\alpha_i, \lambda_i), i=1,2$, $C \sim weibull(\beta, \gamma)$

第一節 阿基米德耦合

以下分別介紹三種阿基米德耦合真值以及產生資料方法,接著討論模擬結果,包含估計值相對偏誤以及三種標準差估計法比較。

Clayton、Frank、Gumbel有以下模擬情境:

1. 第一間隔時間設限率($\delta_1 = 0$)約為0.24, 第二間隔時間設限率($\delta_2 = 0$)約為0.4。 $T_1 \sim Weibull(2, 0.5)$, $T_2 \sim Weibull(3.2, 0.35)$, $C \sim Weibull(1.18, 1.3)$
2. 第一間隔時間設限率($\delta_1 = 0$)約為0.54, 第二間隔時間設限率($\delta_2 = 0$)約為0.76。 $T_1 \sim Weibull(2, 0.5)$, $T_2 \sim Weibull(3.2, 0.35)$, $C \sim Weibull(1.2, 0.5)$

分別測試 (T_1, T_2) 分位數 τ 從0.1至0.7,每0.1一個間隔,模擬1000次,測試樣本數為200, 400, kendall等級相關係數(τ)=0.3, 0.5。

使用R"quantreg"套件中的 $rq.fit.br$ (Barrodale–Roberts algorithm)求 $U_n(\mathbf{b}, \tau_1, \tau_2)$ 最小值。



4.1.1 克萊頓耦合 (Clayton Copula)

$$P(U_1 \leq \mu_1, U_2 \leq \mu_2) = [\mu_1^{-\theta} + \mu_2^{-\theta} - 1]^{-1/\theta}$$

U_1 、 U_2 為均勻分布，參數 θ 與其 kendall 等級相關係數(τ)有一對一關係，

$$\theta = \frac{2\tau}{1-\tau}。$$

第一節首先計算真值，我們可以在給定(i) Y_1 邊際機率(τ_1)，(ii)條件於 Y_1 之下 Y_2

的條件機率(τ_2)，以及(iii)Clayton 的耦合架構下計算 Y_2 邊際機率。結果

如下：

$$\text{I. } \tau_2 = \Pr(U_2 \leq \mu_{20} | U_1 \leq \tau_1) = \frac{\Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20})}{\Pr(U_1 \leq \tau_1)}$$

$$P(U_2 \leq \mu_{20}) = \left\{ (\tau_1 \mu_{20})^{-\theta} - (\tau_1)^{-\theta} + 1 \right\}^{-1/\theta}$$

$$\beta_0 = \log F_2^{-1}(u_{20})$$

$$\text{II. } \tau_2 = \Pr(U_2 \leq \mu_{21} | U_1 > \tau_1) = \frac{\Pr(U_1 > \tau_1, U_2 \leq \mu_{21})}{\Pr(U_1 > \tau_1)}$$

$$P(U_1 > \tau_1, U_2 > \mu_{21}) = 1 - \tau_1 - \mu_{21} + P(U_1 \leq \tau_1, U_2 \leq \mu_{21}) = (1 - \tau_1)(1 - \tau_2)$$

$$\text{可解得 } \mu_{21}, \beta_1 = \log(F_2^{-1}(u_{21}) / F_2^{-1}(u_{20}))$$

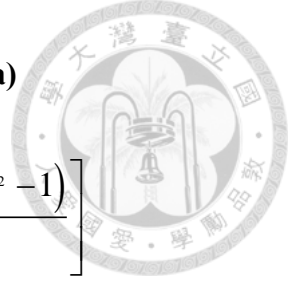
接著產生資料，步驟如下，

1. 產生 $u_1, v \sim \text{unif}(0,1)$

$$2. \quad u_2 = \left\{ u_1^{-\theta} \times \left(v^{\frac{-\theta}{1+\theta}} - 1 \right) + 1 \right\}^{-1/\theta}$$

$$3. \quad T_1 = F_1^{-1}(u_1), \quad T_2 = F_2^{-1}(u_2)$$

4.1.2 法蘭克耦合 (Frank Copula)



$$P(U_1 \leq \mu_1, U_2 \leq \mu_2) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta\mu_1} - 1)(e^{-\theta\mu_2} - 1)}{e^{-\theta} - 1} \right]$$

U_1 、 U_2 為均勻分布，參數 θ 與其 kendall 等級相關係數(τ)有一對一關係，

$$\tau = 1 + \frac{4}{\theta} (D(\theta) - 1), \quad D(\theta) = \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt \quad .$$

首先計算真值，利用相同的概念，可得，

$$I. \quad \tau_2 = \Pr(U_2 \leq \mu_{20} | U_1 \leq \tau_1) = \frac{\Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20})}{\Pr(U_1 \leq \tau_1)},$$

$$u_{20} = -\log \left\{ \frac{(e^{-\tau_1\tau_2\theta} - 1)(e^{-\theta} - 1)}{(e^{-\theta\tau_1} - 1)} + 1 \right\} / \theta, \quad \beta_0 = \log F_2^{-1}(u_{20}).$$

II. 由於 Frank Copula 為對稱的，即 Frank Copula 的存活耦合仍為相同結構，

$$P(U_1 > \mu_1, U_2 > \mu_2) = -\frac{1}{\theta} \log \left[1 + \frac{(e^{-\theta(1-\mu_1)} - 1)(e^{-\theta(1-\mu_2)} - 1)}{e^{-\theta} - 1} \right]$$

$$\tau_2 = \Pr(U_2 \leq \mu_{21} | U_1 > \tau_1) = \frac{\Pr(U_1 > \tau_1, U_2 \leq \mu_{21})}{\Pr(U_1 > \tau_1)},$$

$$\mu_{21} = 1 + \log \left\{ \frac{(e^{-(1-\tau_1)(1-\tau_2)\theta} - 1)(e^{-\theta} - 1)}{e^{-\theta(1-\tau_1)} - 1} + 1 \right\} / \theta,$$

可解得 μ_{21} ， $\beta_1 = \log(F_2^{-1}(u_{21}) / F_2^{-1}(u_{20}))$ 。

接著產生資料，步驟如下，

1. 產生 $u_1, v \sim \text{unif}(0,1)$

$$2. \quad u_2 = -\frac{1}{\theta} \log \left[1 + \frac{v_2(1 - e^{-\theta})}{v_2(e^{-\theta\mu_1} - 1) - e^{-\theta\mu_1}} \right]$$

3. $T_1 = F_1^{-1}(u_1)$ ， $T_2 = F_2^{-1}(u_2)$

4.1.3 岡貝爾耦合 (Gumbel Copula)



$$P(U_1 \leq \mu_1, U_2 \leq \mu_2) = \exp \left\{ - \left[(-\log \mu_1)^\theta + (-\log \mu_2)^\theta \right]^{1/\theta} \right\}$$

U_1 、 U_2 為均勻分布，參數 θ 與其 kendall 等級相關係數(τ)有一對一關係， $\theta = \frac{1}{1-\tau}$ 。

首先計算真值，利用相同的概念，可得，

$$\text{I. } \tau_2 = \Pr(U_2 \leq \mu_{20} | U_1 \leq \tau_1) = \frac{\Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20})}{\Pr(U_1 \leq \tau_1)}$$

$$\mu_{20} = \exp \left\{ - \left[(-\log \tau_2 \tau_1)^\theta - (-\log \tau_1)^\theta \right]^{1/\theta} \right\}$$

$$\beta_0 = \log F_2^{-1}(u_{20})。$$

$$\text{II. } \tau_2 = \Pr(U_2 \leq \mu_{21} | U_1 > \tau_1) = \frac{\Pr(U_1 > \tau_1, U_2 \leq \mu_{21})}{\Pr(U_1 > \tau_1)}$$

$$P(U_1 > \tau_1, U_2 > \mu_{21}) = 1 - \tau_1 - \mu_{21} + P(U_1 \leq \tau_1, U_2 \leq \mu_{21}) = (1 - \tau_1)(1 - \tau_2)$$

$$\text{可解得 } \mu_{21}, \beta_1 = \log(F_2^{-1}(u_{21}) / F_2^{-1}(u_{20}))$$

接著產生資料，步驟如下，

$$1. \text{ 產生 } v_1, v_2 \sim \text{unif}(0,1)$$

$$2. \quad K_c(w) = w \left(1 - \frac{\log w}{\theta} \right) = v_2, 0 < w < 1 \text{ 數值解根}$$

$$3. \quad u_1 = \exp \{ v_1^{1/\theta} \log w \}, \quad u_2 = \exp \{ (1 - v_1)^{1/\theta} \log w \}$$

$$4. \quad T_1 = F_1^{-1}(u_1), \quad T_2 = F_2^{-1}(u_2)$$



4.1.4 阿基米德耦合 模擬結果

模擬結果顯示，在有限的樣本數、中等相關性以及中等的設限率下，阿基米德耦合在 $\tau_1 \in [0.1, 0.7]$ ， $\tau_2 \in [0.1, 0.7]$ 的表現相對偏誤都在 3% 以下，且大部分為低估。三種耦合的偏誤隨著 τ_1 、 τ_2 而增加，而真值隨著 τ_1 、 τ_2 而減小。究其原因，因 CQR 為兩段時間長度比， τ_1 、 τ_2 越小則其相對基準小而有較大比值。當設限率提高時，偏誤皆明顯上升， $\tau_1 \in [0.4, 0.7]$ ， $\tau_2 \in [0.4, 0.7]$ 的表現相對偏誤都在 5% 以上，相對偏誤絕對值大小順序為 Gumbel > Frank > Clayton，Clayton 在高設限率下偏誤相對較低。而當樣本數提高至 400 時，除 $\tau = (0.6, 0.7)$ 及 $\tau = (0.7, 0.7)$ 外表現相對偏誤都在 5% 以下。當相關性提升時，真值隨著增大，其相對偏誤縮小，尤其在高分位數的結果尤其明顯。

標準差同樣隨著 τ_2 上升，隨著 τ_1 下降，但在 τ_1 大於觀察到資料比例時 ($\delta_1 = 1$) 則逐漸走高，可以想見，在高設限率下高分位數 (τ_1) 的估計值並不穩定。當樣本數或相關性提升時，標準差則較小。Bootstrapping 與 resampling 兩種方法估計標準差都較 1000 次模擬標準差略大，且 resampling 方法比 bootstrapping 方法更小，但 resampling 方法在高設限率高分位數時有低估情形，此情況在樣本數提升至 400 時有改善。而本文提供之標準差估計皆高於 Bootstrapping 與 resampling 兩種方法估計標準差，尤其在低分量時 ($\tau_1 = 0.1$)，亦高於 1000 次模擬標準差，且受到樣本數以及設限率很大的影響。當設限率為 (0.15, 0.3)，樣本數 800 時，本文提供之標準差估計與 1000 次模擬標準差越來越接近，但估計值仍為三種方法中最大。

總結來說，阿基米德耦合在有限的樣本數、中等相關性以及中等的設限率下，相對偏誤皆在 5% 以下，使用 resampling 方法估計之標準差為三種方式中最小的。當設限率提高或高分位數的估計上則需更大樣本數。

第二節 對數常態模型



多元常態是一常使用假設分布，利用二元常態(Bivariate Normal)分布來產生兩事件時間的相關性，再將二元常態分布生成的二元事件時間資料做指數轉換後，即得二元對數常態分布的時間資料。以下介紹二元常態產生資料方法，接著討論模擬結果，包含估計值相對偏誤以及三種標準差估計法比較。

模擬情境:

1. 第一間隔時間設限率($\delta_1 = 0$)約為 0.18，第二間隔時間設限率($\delta_2 = 0$)約為 0.4。

$$T_1 \sim \text{Lognormal}(-1, 0.25), T_2 \sim \text{Lognormal}(-0.8, 0.25),$$

$$C \sim \text{Weibull}(1.33, 1.48)$$

2. 第一間隔時間設限率($\delta_1 = 0$)約為 0.38，第二間隔時間設限率($\delta_2 = 0$)約為

$$0.65。T_1 \sim \text{Lognormal}(-1, 0.25), T_2 \sim \text{Lognormal}(-0.8, 0.25),$$

$$C \sim \text{Weibull}(1.05, 0.79)$$

分別測試(T_1, T_2) 分位數 τ 從 0.1 至 0.7，每 0.1 一個間隔，模擬 1000 次，測試樣本數為 200, 400，相關性 $kendall's \tau = 0.3, 0.5$ 。

首先計算真值，

$$\mathbf{V} \sim MVN\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_1\sigma_2 \times \tau \\ \sigma_1\sigma_2 \times \tau & \sigma_2^2 \end{pmatrix}\right), \tau \text{ 為 kendall's tau}$$

$$\text{I. } \Pr(V_1 \leq v_1, V_2 \leq v_2) = \Phi_{\Sigma}(v_1, v_2), V_1 \sim N(\mu_1, \sigma_1^2), V_2 \sim N(\mu_2, \sigma_2^2)$$

$$\tau_2 = \Pr(V_2 \leq v_{20} | V_1 \leq v_1) = \frac{\Pr(V_1 \leq v_1, V_2 \leq v_{20})}{\Pr(V_1 \leq v_1)}$$

$$\text{解 } \Phi_{\Sigma}(v_1, v_{20}) - \tau_1 \tau_2 = 0 \text{ 找出 } v_{20}, \beta_0 = \log Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1}) = v_{20}$$



$$\text{II. } \tau_2 = \Pr(V_2 \leq v_{21} | V_1 > v_1) = \frac{\Pr(V_1 > v_1, V_2 \leq v_{21})}{\Pr(V_1 > v_1)}$$

$$\Pr(V_1 \leq v_1, V_2 \leq v_{21}) + \Pr(V_1 > v_1, V_2 \leq v_{21}) = \Pr(V_2 \leq v_{21})$$

$$\Phi_{\Sigma}(v_1, v_{21}) + (1 - \tau_1) \tau_2 = \Pr(V_2 \leq v_{21})$$

可解得 v_{21} ， $\beta_1 = v_{21} - v_{20}$

接著產生資料，步驟如下，

1. 產生 $Z_1 \sim N(0,1)$ ， $Z_2 \sim N(0,1)$

2. 將共變異數矩陣做 cholesky 分解，

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \sin(\tau \times \frac{\pi}{2}) \\ \sigma_1 \sigma_2 \sin(\tau \times \frac{\pi}{2}) & \sigma_2^2 \end{pmatrix} = LL^T$$

3. $T_1 = \exp\{\mu_1 + LZ_1\}$ ， $T_2 = \exp\{\mu_2 + LZ_2\}$

4.2.1 對數常態模型 模擬結果

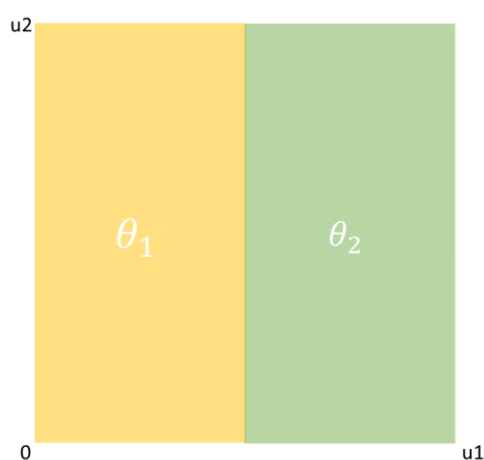
模擬結果顯示，在有限的樣本數、中等相關性以及中等的設限率下，對數常態模型在 $\tau_1 \in [0.1, 0.7]$ ， $\tau_2 \in [0.1, 0.7]$ 的表現都穩定低估 0.012~0.014 左右，並沒有隨著 τ_1 、 τ_2 而增加。當設限率高時，偏誤上升至 0.02~0.04 左右，隨著 τ_1 、 τ_2 而增加，相對偏誤都在 5% 以上。而當樣本數提高至 400 時，偏誤隨著 τ_1 、 τ_2 而增加，相對偏誤在 5% 以下，而高設限率以及高分量時相對偏誤達 4~6%。與阿基米德耦合不同的是，對數常態模型之真值並無隨著分位數提高而明顯下降，因此可觀察到篇舞以及相對偏誤隨著分量上升幅度不若阿基米德耦合。當相關性提升至 kendall's $\tau = 0.5$ 時，偏誤與 kendall's $\tau = 0.3$ 時接近，但略為上升，且隨著 τ_1 、 τ_2 而增加，因此觀察到相對偏誤較小。總結來說，對數常態模型偏誤主要受樣本數大小以及設限率影響，相關性以及分量提升同樣會使偏誤上升，但幅度不比前二項因素。

標準差隨著 τ_2 上升，隨著 τ_1 下降，但當 τ_1 大於第一段間隔時間完整資料比例 (δ_1) 時亦會上升。高設限會使標準差增加，而大樣本以及相關性提昇。會使標準差降低。Bootstrapping 與 resampling 兩種方法估計標準差都較 1000 次模擬標準差略大，與阿基米德耦合不同的是，resampling 方法不一定比 bootstrapping 方法更小，尤其在高分位數時。而本文提供之標準差估計皆高於 Bootstrapping 與 resampling 兩種方法估計標準差，亦高於 1000 次模擬標準差，且受到樣本數以及設限率很大的影響。當設限率為(0.15, 0.3)，樣本數 800 時，本文提供之標準差估計與 1000 次模擬標準差越來越接近，但估計值仍為三種方法中最大。

總結來說，對數常態模型在有限的樣本數、中等相關性以及中等的設限率下，相對偏誤皆在 5% 以下，使用 bootstrapping 方法估計之標準差為三種方式中最穩健。當設限率提高或高分位數的估計上則需更大樣本數。

第三節 分段常數克萊頓 (Piecewise constant Clayton)

本節模擬二元時間相關性隨分位數改變的情況，參考 Lin et al. (2006)，假設二元時間的相關性參數 θ 與第一段時間 Y_1 有關，也就是說其為 Y_1 的函數 $\theta(y_1)$ 。假設 Y_1 小於中位數的部分與 Y_2 的相關性結構仍為 Clayton Copula 參數為 θ_1 ， Y_1 大於中位數的部分與 Y_2 的相關性結構亦為 Clayton Copula 參數為 θ_2 。如下圖所示: (u_1 、 u_2 為邊際累積機率)



以下介紹二元常態產生資料方法，接著討論模擬結果，包含估計值相對偏誤以及三種標準差估計法比較。

模擬情境:

3. 第一間隔時間設限率($\delta_1 = 0$)約為 0.24，第二間隔時間設限率($\delta_2 = 0$)約為 0.4。

$$T_1 \sim \text{Weibull}(2, 0.5), T_2 \sim \text{Weibull}(3.2, 0.35), C \sim \text{Weibull}(1.18, 1.3)$$

4. 第一間隔時間設限率($\delta_1 = 0$)約為 0.54，第二間隔時間設限率($\delta_2 = 0$)約為

$$0.76。T_1 \sim \text{Weibull}(2, 0.5), T_2 \sim \text{Weibull}(3.2, 0.35), C \sim \text{Weibull}(1.2, 0.5)$$

分別測試(T_1, T_2) 分位數 τ 從 0.1 至 0.7，每 0.1 一個間隔，模擬 1000 次，測試樣本數為 200，相關性前高後低($\theta_1 = 0.5$ ， $\theta_2 = 0.2$)、前低後高($\theta_1 = 0.2$ ， $\theta_2 = 0.5$)以及相關性一正一負 ($\theta_1 = 0.2$ ， $\theta_2 = -0.5$)。

根據以上假設可寫下，

1. 在給定 $U_1 \leq 0.5$ 之下，

$$P(U_1 \leq u_1, U_2 \leq \mu_2 | U_1 \leq 0.5) = \frac{P(U_1 \leq u_1, U_2 \leq \mu_2)}{P(U_1 \leq 0.5)} = [\tilde{\mu}_1^{-\theta_1} + \tilde{\mu}_2^{-\theta_1} - 1]^{-1/\theta_1},$$

$$\tilde{\mu}_1 = P(U_1 \leq u_1) / P(U_1 \leq 0.5), \quad \tilde{\mu}_2 = P(U_1 \leq 0.5, U_2 \leq \mu_2) / P(U_1 \leq 0.5)$$

2. 在給定 $U_1 > 0.5$ 之下，

$$P(U_1 \leq u_1, U_2 \leq \mu_2 | U_1 > 0.5) = \frac{P(U_1 \leq u_1, U_2 \leq \mu_2)}{P(U_1 > 0.5)} = [\check{\mu}_1^{-\theta_2} + \check{\mu}_2^{-\theta_2} - 1]^{-1/\theta_2}$$

$$\check{\mu}_1 = (u_1 - 0.5) / P(U_1 > 0.5), \quad \check{\mu}_2 = (\mu_2 - P(U_1 \leq 0.5, U_2 \leq \mu_2)) / P(U_1 > 0.5)$$

首先計算真值，在 $u_1 \leq 0.5$ 的情況下，其邊際聯合分布仍為 Clayton Copula，

$$P(U_1 \leq \mu_1, U_2 \leq \mu_2 | U_1 \leq 0.5) P(U_1 \leq 0.5) = [\mu_1^{-\theta_1} + \mu_2^{-\theta_1} - 1]^{-1/\theta_1}$$

$$\text{I. } \tau_2 = \Pr(U_2 \leq \mu_2 | U_1 \leq \tau_1) = \frac{\Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20})}{\Pr(U_1 \leq \tau_1)}$$

$$\tau_1 \tau_2 = [\tau_1^{-\theta_1} + \mu_2^{-\theta_1} - 1]^{-1/\theta_1}$$

$$P(U_2 \leq \mu_{20}) = \left\{ (\tau_1 \tau_2)^{-\theta_1} - (\tau_1)^{-\theta_1} + 1 \right\}^{-1/\theta_1}$$

$$\beta_0 = \log F_2^{-1}(u_{20})$$

$$\text{II. } \tau_2 = \Pr(U_2 \leq \mu_{21} | U_1 > \tau_1) = \frac{\Pr(U_1 > \tau_1, U_2 \leq \mu_{21})}{\Pr(U_1 > \tau_1)}$$

$$P(U_1 > \tau_1, U_2 > \mu_2) = 1 - \tau_1 - \mu_2 + P(U_1 \leq \tau_1, U_2 \leq \mu_2) = (1 - \tau_1)(1 - \tau_2)$$

$$\text{可解得 } \mu_{21}, \quad \beta_1 = \log(F_2^{-1}(u_{21}) / F_2^{-1}(u_{20}))$$

在 $u_1 > 0.5$ 的情況下，其邊際聯合分布如下，

$$P(U_1 \leq \mu_1, U_2 \leq \mu_2) =$$

$$P(U_1 \leq \mu_1, U_2 \leq \mu_2 | U_1 \leq 0.5)P(U_1 \leq 0.5) + P(U_1 \leq \mu_1, U_2 \leq \mu_2 | U_1 > 0.5)P(U_1 > 0.5)$$

$$= \left[0.5^{-\theta_1} + \mu_2^{-\theta_1} - 1 \right]^{-1/\theta_1} + \left[\left(P(U_1 \leq \mu_1) - 0.5 \right)^{-\theta_2} + \left(\mu_2 - P(U_1 \leq 0.5, U_2 \leq \mu_2) \right)^{-\theta_2} - 0.5^{-\theta_2} \right]^{-1/\theta_2}$$

$$\text{I. } \tau_2 = \Pr(U_2 \leq \mu_{20} | U_1 \leq \tau_1) = \frac{\Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20})}{\Pr(U_1 \leq \tau_1)}$$

$$\tau_1 \tau_2 = \Pr(U_1 \leq \tau_1, U_2 \leq \mu_{20}) , \text{ 解得 } \mu_{20}$$

$$\beta_0 = \log F_2^{-1}(u_{20})$$

$$\text{II. } \tau_2 = \Pr(U_2 \leq \mu_{21} | U_1 > \tau_1) = \frac{\Pr(U_1 > \tau_1, U_2 \leq \mu_{21})}{\Pr(U_1 > \tau_1)}$$

$$P(U_1 > \tau_1, U_2 > \mu_{21}) = 1 - \tau_1 - \mu_2 + P(U_1 \leq \tau_1, U_2 \leq \mu_{21}) = (1 - \tau_1)(1 - \tau_2)$$

$$\text{可解得 } \mu_{21} , \beta_1 = \log(F_2^{-1}(u_{21}) / F_2^{-1}(u_{20}))$$

接著產生資料，步驟如下，

1. 產生 $u_1, v \sim \text{unif}(0,1)$
2. 根據 u_1 大於或小於 0.5 產生 u_2

若 $u_1 \leq 0.5$ ，利用條件機率 $v = P(U_2 \leq u_2 | U_1 = u_1, u_1 \leq 0.5)$

$$\tilde{u}_2 = \left\{ \tilde{u}_1^{-\theta_1} \times \left(v^{\frac{-\theta_1}{1+\theta_1}} - 1 \right) + 1 \right\}^{-1/\theta_1}, \quad \tilde{u}_1 = u_1 / 0.5$$

$$\mu_2 = \left[\{ \tilde{\mu}_2 \times 0.5 \}^{-\theta_1} + 1 - 0.5^{-\theta_1} \right]^{-1/\theta_1}$$

若 $u_1 > 0.5$ ，利用條件機率 $v = P(U_2 \leq u_2 | U_1 = u_1, u_1 > 0.5)$

$$\tilde{u}_2 = \left\{ \tilde{u}_1^{-\theta_2} \times \left(v^{\frac{-\theta_2}{1+\theta_2}} - 1 \right) + 1 \right\}^{-1/\theta_2}, \quad \tilde{u}_1 = \left(\frac{\mu_1 - 0.5}{0.5} \right)$$

$$\tilde{\mu}_2 = (\mu_2 - P(U_1 \leq 0.5, U_2 \leq \mu_2)) / P(U_1 > 0.5) , \text{ 可解得 } \mu_2$$

4.3.1 分段常數克萊頓模型 模擬結果



模擬結果顯示，在有限的樣本數、中等的設限率下，若前段相關性大於後段相關性($\theta_1 = 0.5$, $\theta_2 = 0.2$)，於切點分位數($\tau_1 = 0.5$) 相對偏誤約高估 1%；反之，前低後高($\theta_1 = 0.2$, $\theta_2 = 0.5$)則於切點分位數($\tau_1 = 0.5$)相對偏誤約大幅高估至接近 20%；而其餘分位數則與 Clayton 模擬結果相似，皆為低估。當設限率高時，相關性前高後低($\theta_1 = 0.5$, $\theta_2 = 0.2$)相對偏誤都在 5% 以下，除 $\tau = (0.6, 0.7)$ 的相對偏誤達 6% 之外，且於切點分位數($\tau_1 = 0.5$)相對偏誤從高估變為低估 1%；前低後高($\theta_1 = 0.2$, $\theta_2 = 0.5$)同樣低估幅度更大，除切點分位數($\tau_1 = 0.5$)因為原先高估而向下調整，因此相對偏誤反而下降。而當樣本數提高至 400 時，整體偏誤皆有改善。若相關性一正一負時，只有 $\tau_1 < 0.5$ 的相對偏誤在 5% 以下，大於等於切點($\tau_1 \geq 0.5$) 的相對偏誤都低估許多。

標準差估計與全段常數 Clayton 模型相似，Bootstrapping 與 resampling 兩種方法估計標準差都較 1000 次模擬標準差略大，且 resampling 方法比 bootstrapping 方法更小，但 resampling 方法在高設限率且高分位數時有低估情形。而本文提供之標準差估計皆高於 Bootstrapping 與 resampling 兩種方法估計標準差，尤其在低分量時($\tau_1 = 0.1$)，亦高於 1000 次模擬標準差，且受到樣本數以及設限率很大的影響。

總結來說，分段常數克萊頓模型與全段常數克萊頓模型偏誤趨勢相似，但在切點位置有較大偏誤；另外若兩段的相關性有較大的差異，甚或為相反相關性，則相對偏誤較大。

第五章 實際資料分析



簡介：囊狀纖維化(Cystic Fibrosis, CF)為體染色體隱性遺傳疾病，患者由於第七對染色體長臂上 CFRT (cystic fibrosis transmembrane conductance regulator) 基因的缺陷，造成呼吸道、胰臟、腸胃道、汗腺等外分泌腺體器官的功能異常。白血球釋放細胞外 DNA，這種過量的 DNA 會使粘液變稠，因此細菌容易生長而增加感染跟發炎的危險。患者經常受反覆呼吸道感染之苦，且肺部問題的惡化常是主要的死因。去氧核糖核酸酶 I (deoxyribonuclease I, DNase I) 是一種可以消化單鏈或雙鏈 DNA 的核酸內切酶。Genentech, Inc 製造高度純化的重組去氧核糖核酸酶 I (recombinant human deoxyribonuclease, rhDNase)，當吸入肺部時可切割細胞外 DNA，降低氣管分泌物的粘性並改善清除率。1992 年，該公司進行了一項比較 rhDNase 與安慰劑的隨機雙盲試驗 (Fuchs et al. 1994)，然後監測患者的肺部惡化情形，以及肺容量和血流量的測量。該試驗的主要事件 (primary end point) 是第一次肺部感染的時間，感染的定義為需要使用靜脈注射抗生素。以下介紹在 R “survival” 套件中的 rhDNase 資料欄位以及自行衍伸欄位，“survival” 套件中的 rhDNase 資料欄位：

Id: 編號	Inst: 研究機構	Trt: 0=對照組, 1= rhDNase
entry.dt: 進入研究日期	end.dt: 最後追蹤日期	fev: 用力呼氣容積(baseline)
ivstart: 從進入研究到開始使用抗生素的天數	ivstop: 從進入研究到開始停止使用抗生素的天數	

衍伸欄位：

$Y_{1,1}$: 從進入研究到第一次開始使用抗生素的天數，第一個下標 1 代表感染事件，第二個下標 1 代表為感染事件第一次間隔時間。

$Y_{2,1}$: 從第一次開始使用抗生素至第一次停止使用抗生素的天數，第一個下標 2

代表治療事件，第二個下標 1 代表為治療事件第一次間隔時間。

$Y_{1,2}$: 從第一次停止使用抗生素到第二次開始使用抗生素的天數

$Y_{2,1}$: 從第二次開始使用抗生素至第二次停止使用抗生素的天數

δ_1 : 觀察到第一次感染事件與治療事件指標

δ_2 : 觀察到第二次感染事件與治療事件指標

該筆資料中若有發生感染事件則皆有治療事件，因此只使用兩個事件指標代表發生第一次與第二次感染事件與治療事件。

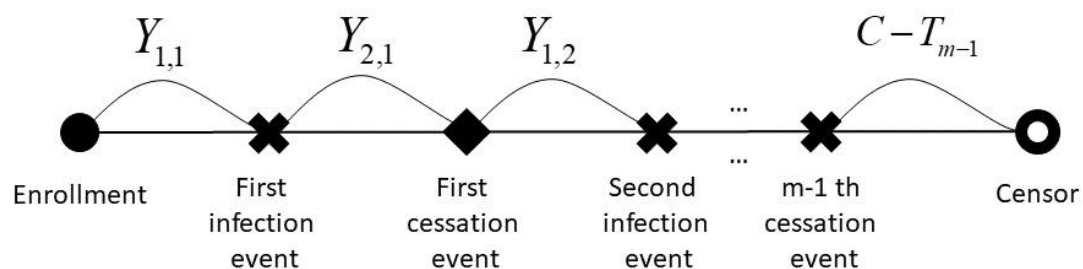


圖 2、資料結構及符號，C 為設限時間， T_m 為第 m 次感染事件與治療結束時間， $m=1, \dots, 5$

第一節 資料探索



大部分患者追蹤時間為 169 天，最長追蹤時間為 189 天，大部分受試者在追蹤期間沒有發生感染(62%)，最多發生感染 5 次(1 位)。總共 647 位患者，其中 318 位接受 rhDNase 治療，323 位為對照組，共有 6 位受試者在加入研究時處於感染狀態，我們將其從分析資料刪除，剩下 641 位。38% (241/641)的患者至少發生一次感染，13% (80/641)的患者至少有兩次肺部感染惡化的事件發生，28、8、1 位 的患者分別發生 3 到 5 次感染惡化事件。

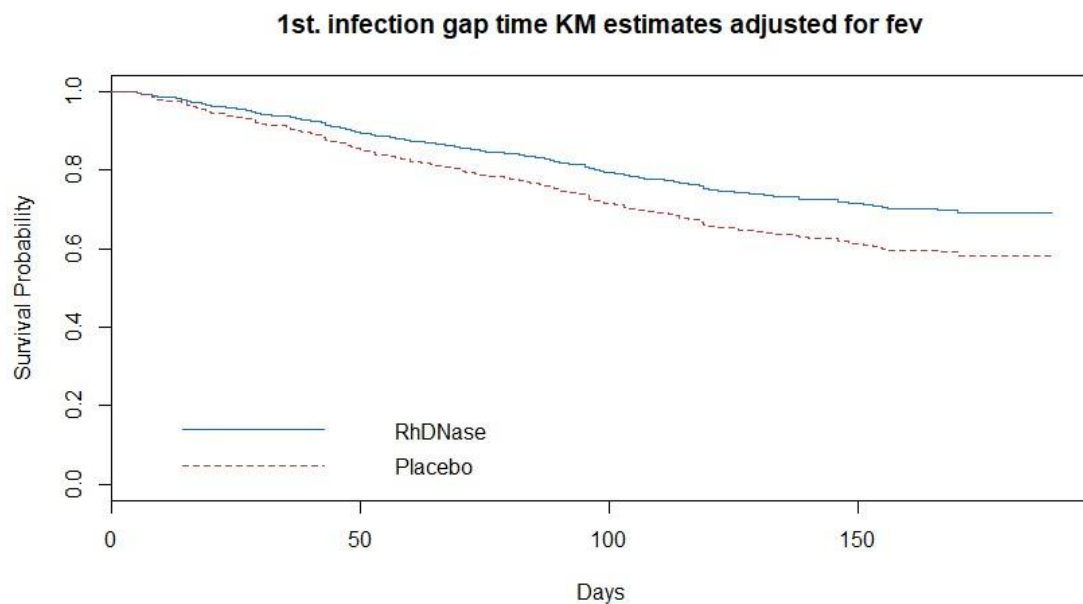


圖 3、第一次感染間隔時間對治療分組之存活曲線(調整 fev)

首先觀察 rhDNase 治療對於延緩第一次感染的效果，納入進入研究時的第一秒用力呼氣量(fev1) 建立 Cox 模型，發現治療降低 CF 病患 32% 肺部感染的風險，且統計上皆呈顯著效果。 $\hat{h}(t|trt, fev) = \hat{h}_0(t) \exp\{-0.38trt - 0.02fev\}$

由於 rhDNase 是長期使用，研究者想要了解 rhDNase 對於長期延緩 CF 病患感染是否有效。接著考慮已經有第一次肺部感染患者，rhDNase 治療對於延緩第二次感染效果，發現治療增加 CF 病患 37% 肺部感染的風險，且統計上皆不顯著。

$$\hat{h}(t | trt, fev) = \hat{h}_0(t) \exp\{0.32trt - 0.0003 fev\}$$

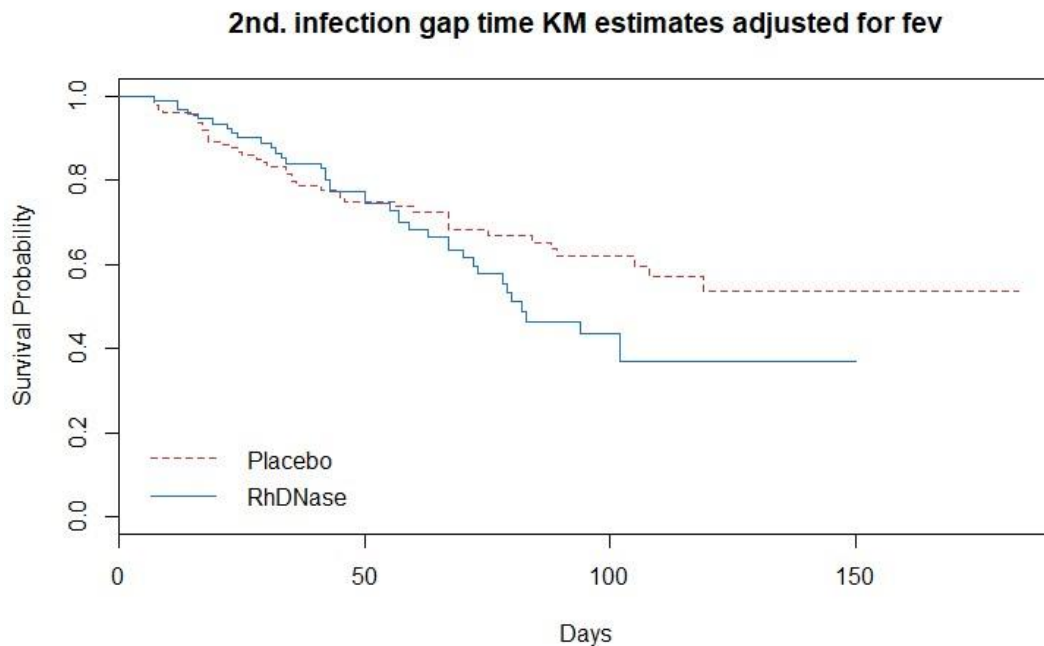


圖 4、第二次感染間隔時間對治療分組之存活曲線(調整 fev)

第一次感染事件後治療效果明顯減弱，且兩組的存活曲線交叉，顯示治療對於延緩 CF 病患第二次感染並無幫助，與原先預想結果不同。推測可能原因為，在追蹤期間內(約 169 天)經歷第一次感染之族群為篩選過之群體，且肺部惡化程度較嚴重的一群病患。由於肺部感染會形成結痂組織，進而降低肺功能，因此可以合理地假設每次肺部感染嚴重性都會影響下一次的感染。所以在條件於經歷第一次感染群體中，觀察治療對第二次感染效果，應加入前一次感染情形分析。

本文接著利用上一次治療間隔時間所提供之感染歷史資訊，觀察其與第二次感染時間的關係。

第二節 納入過去歷史



一、 首先觀察第一段治療間隔時間與第二段感染間隔時間相關性，因為第一段治療時間長短其實代表著上一次感染的嚴重程度或肺部惡化情況，

$$Q_{\tau_2}(Y_{1,2} | I(Y_{2,1} > Q_{\tau_1})) = \exp\{\beta_0(\tau_1, \tau_2) + \beta_1(\tau_1, \tau_2)I(Y_{2,1} > Q_{\tau_1})\}$$

因為有第一次肺部感染皆有第一段治療間隔時間，因此第一段治療時間為完整資料，第二段治療時間設限率為 0.33，我們探討 $\tau_1 = 0.3, 0.5, 0.7$ ，

$$\tau_2 = (0.1, 0.9) \text{ 時， } \hat{\beta}_1(\tau_1, \tau_2) = \log CQR_{Y_{2,1}, \tau_1}(\tau_2)$$

$$= \log \hat{Q}_{\tau_2}(Y_{1,2} | \tilde{Y}_{2,1} > \hat{Q}_{\tau_1}) - \log \hat{Q}_{\tau_2}(Y_{1,2} | \tilde{Y}_{2,1} \leq \hat{Q}_{\tau_1})，$$

我們發現在某些分位數顯著，變異數估計為透過重新取樣法(resampling)而取得，結果如圖 5 所示，

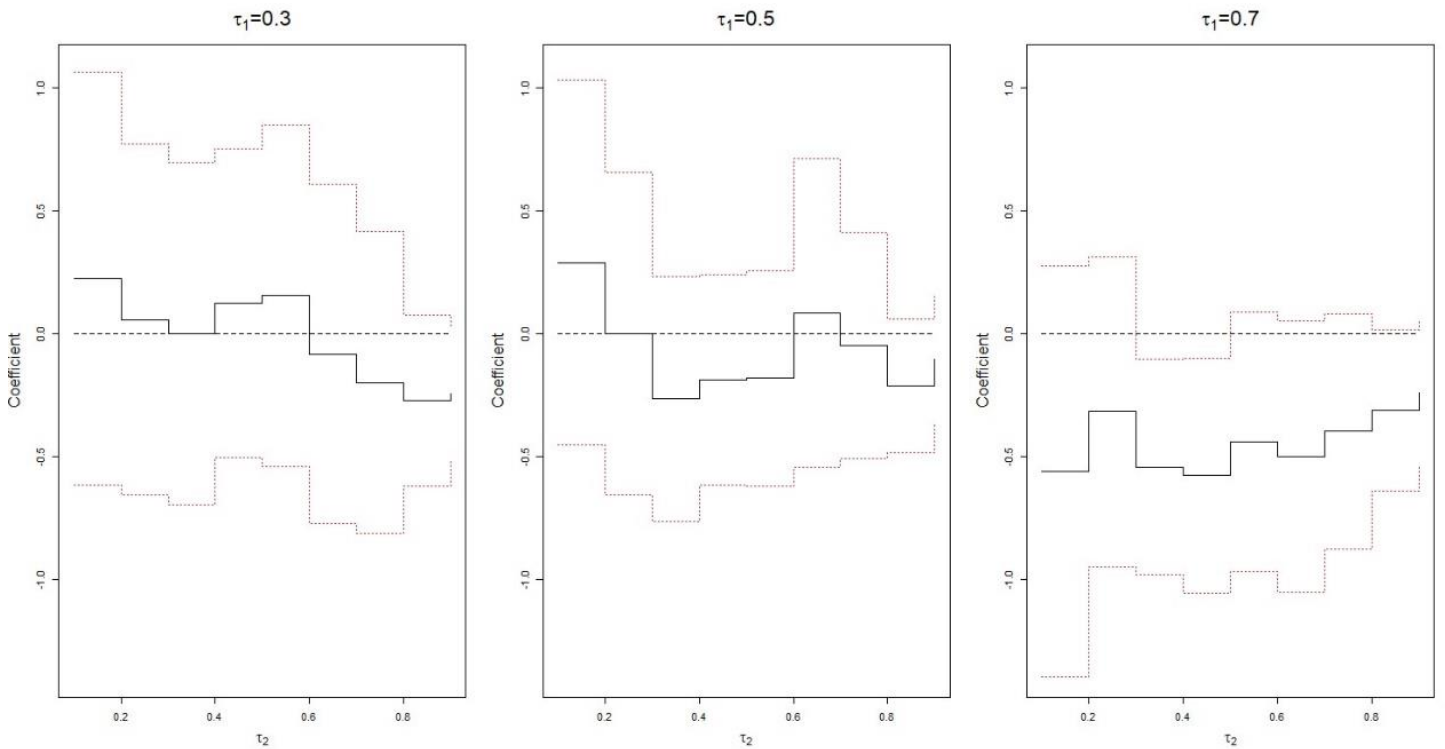


圖 5、 $\beta_1(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間

當 $\tau_1 = 0.3, 0.5$ 時， $\hat{\beta}_1(\tau_1, \tau_2)$ 都在 0 附近且不顯著。當 $\tau_1 = 0.7$ 時，

$\hat{\beta}_1(\tau_1, \tau_2)$ 皆小於 0 且在某些 τ_2 分位數顯著。此結果支持先前的假設，也就是先

前感染較嚴重者，很快會發生下一次感染。在這筆資料中，70 百分位數可以區分出感染嚴重與輕微者。 $\hat{\beta}_1(\tau_1=0.7, \tau_2=0.5)=(-0.923, 0.044)$ ，即 $CQR_{Y_{1,2},0.7}(0.5)=(0.40, 1.04)$ ，代表第一次治療間隔時間長度長於 70 百分位的病人相較於第一次治療間隔時間長度低於 70 百分位的病人，其第二段感染間隔時間長度中位數縮短 36%，雖然結果並不顯著，但觀察治療間隔時間與第二段感染間隔時間呈負相關，顯示感染較嚴重病患很快會發生第二次感染。

該研究欲探討使用 rhDNase 是否能延長 CF 病患每次肺部感染的間隔時間，由於治療對第二次感染的效果與預期不符，推測條件於有發生第一次感染的病患中，應納入前次感染嚴重情況進行分析，而 $CQR_{Y_{1,2},\tau_1}(\tau_2)$ 顯示第一次治療間隔時間長度 $Y_{2,1}$ 與第二次感染間隔時間長度 $Y_{1,2}$ 有負相關，因此在分析治療對第二次感染的效果應加入第一次治療間隔時間長度 $Y_{2,1}$ 進行調整。

二、 接著加入治療變項及交互作用項，觀察 trt 與第一次治療間隔時間是否有交互作用。根據前面的探討將切點設為 70 百分位數，結果如圖 6，

$$Q_{\tau_2}(Y_{1,2} | I(Y_{2,1} > Q_{\tau_1}), trt) = \exp\left\{\beta_0(\tau_1, \tau_2) + \beta_1(\tau_1, \tau_2)I(Y_{2,1} > Q_{\tau_1}) + \beta_2(\tau_1, \tau_2)trt + \beta_3(\tau_1, \tau_2)trt * I(Y_{2,1} > Q_{\tau_1})\right\}$$

因為加入交互作用項，所以需要分成 rhDNase 組與控制組兩組之 $CQR_{Y_{1,2},\tau_1}(\tau_2)$ 來看。 $\exp\{\hat{\beta}_1(0.7, \tau_2)\}$ 為安慰劑組之 $CQR_{Y_{1,2},0.7}(\tau_2)$ ，代表第一次治療間隔時間長於 70 百分位數者 $I(Y_{2,1} > Q_{\tau_1})$ 相較於第一次治療間隔時間短於 70 百分位數的組別，其第二次感染間隔時間長度比。可發現 $\hat{\beta}_1(0.7, \tau_2)$ 皆小於 0，且在 $\tau_2 \geq 0.6$ 時有顯著結果，顯示安慰劑組中，感染嚴重者普遍較輕微者更快感染第二次，且差距隨分位數累積越來越大。 $\exp\{\hat{\beta}_1(0.7, \tau_2) + \hat{\beta}_3(0.7, \tau_2)\}$ 為治療

組之 $CQR_{Y_1,0.7}(\tau_2)$ ，代表第一次治療間隔時間長於 70 百分位數的 $I(Y_{2,1} > Q_{\tau_1})$

相較於第一次治療間隔時間短於 70 百分位數的組別，其第二次感染間隔時間中位數長度比。但效果幾乎在 0 附近且不顯著(圖 7)，也就是說相較於安慰劑

組，治療組之 $CQR_{Y_1,0.7}(\tau_2)$ 並無呈現負相關，顯示 rhDNase 有延長感染較嚴重之病患第二次感染間隔時間。

$\exp\{\hat{\beta}_0(\tau_1, \tau_2)\}$ 為使用安慰劑且感染輕微組 $I(Y_{2,1} \leq Q_{\tau_1})$ 的第二次感染間隔時間長度取對數值，因此隨 τ_2 分位數的累積而越來越大。

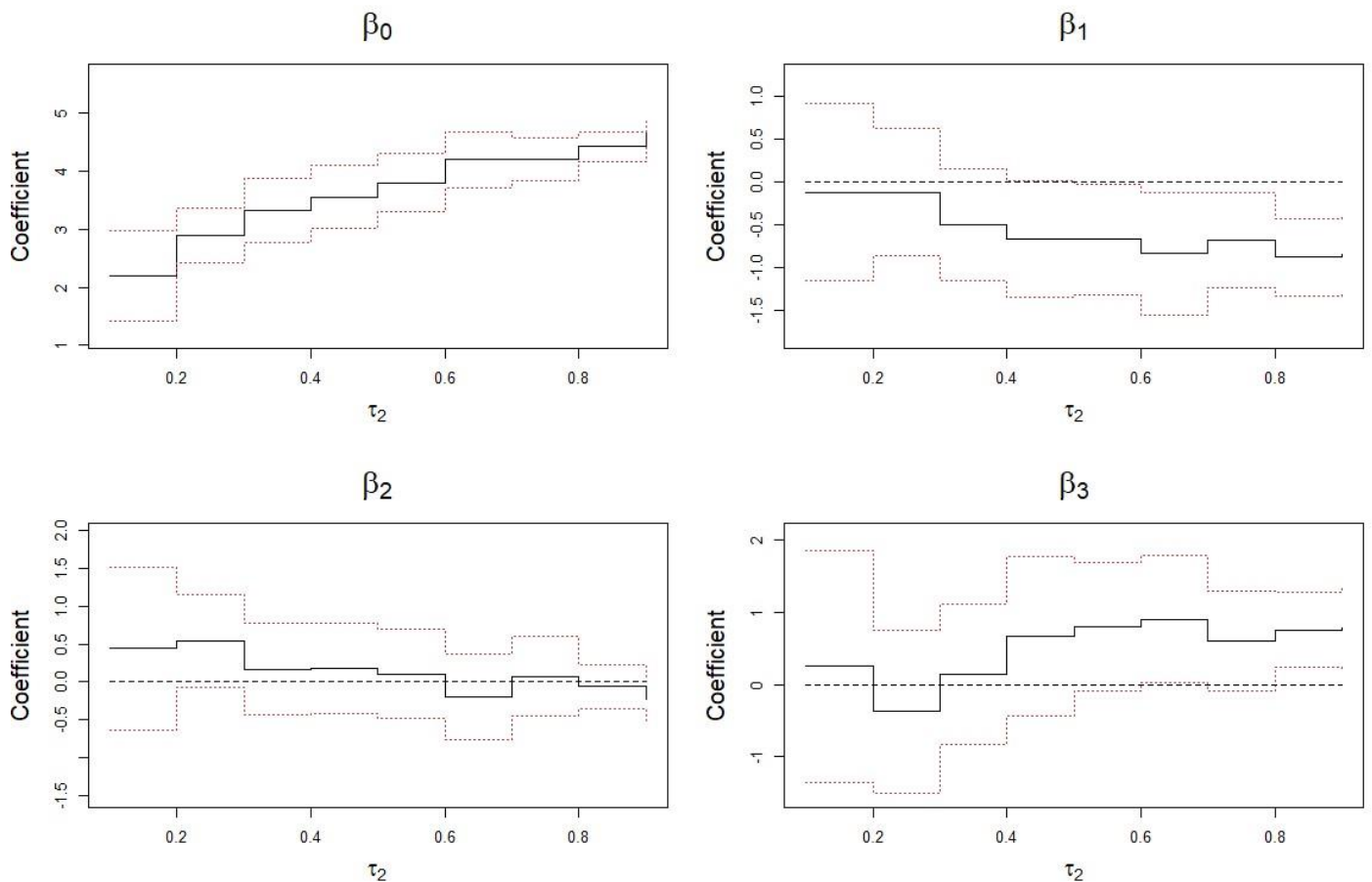


圖 6、 $\tau_1 = 0.7$ ， $\beta(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間(虛線)。

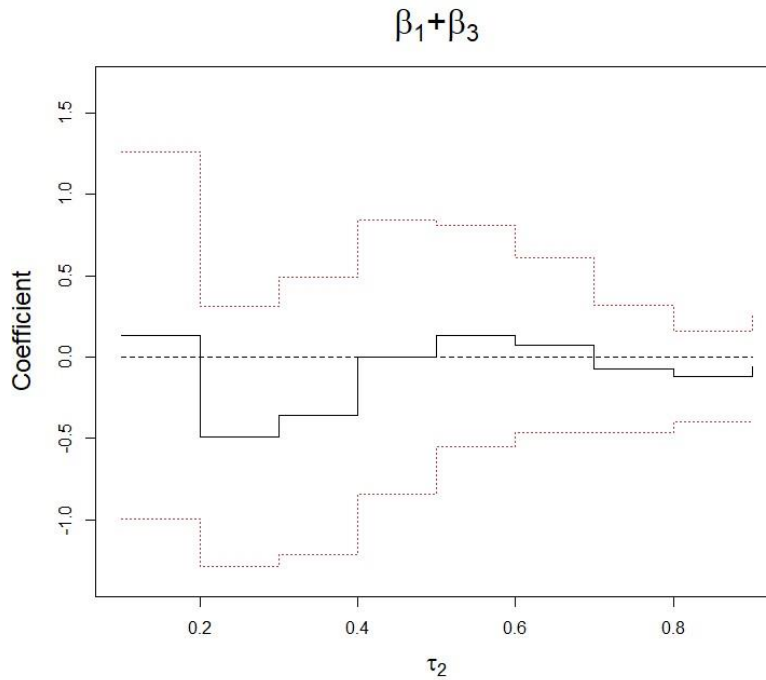


圖 7、 $\tau_1 = 0.7$ ， $\hat{\beta}_1(\tau_1, \tau_2) + \hat{\beta}_3(\tau_1, \tau_2)$ 之估計值(實線)與 resampling 建立之信賴區間(虛線)。

觀察以上結果可發現分位數迴歸在 τ_2 為低分位數時變異都比較大，因此不容易達到顯著，與模擬結果相符。且分位數迴歸本身就需要較大樣本數才達到顯著，且在高設限情況下(67%)更是如此。因此在一般樣本數下(241)且高設限情況下(67%)，我們仍觀察到 rhDNase 延緩感染嚴重組第二次感染間隔時間。根據分組重新畫圖後發現(錯誤! 找不到參照來源。)，兩組治療效果明顯不同且前次感染嚴重程度是有效分類指標。

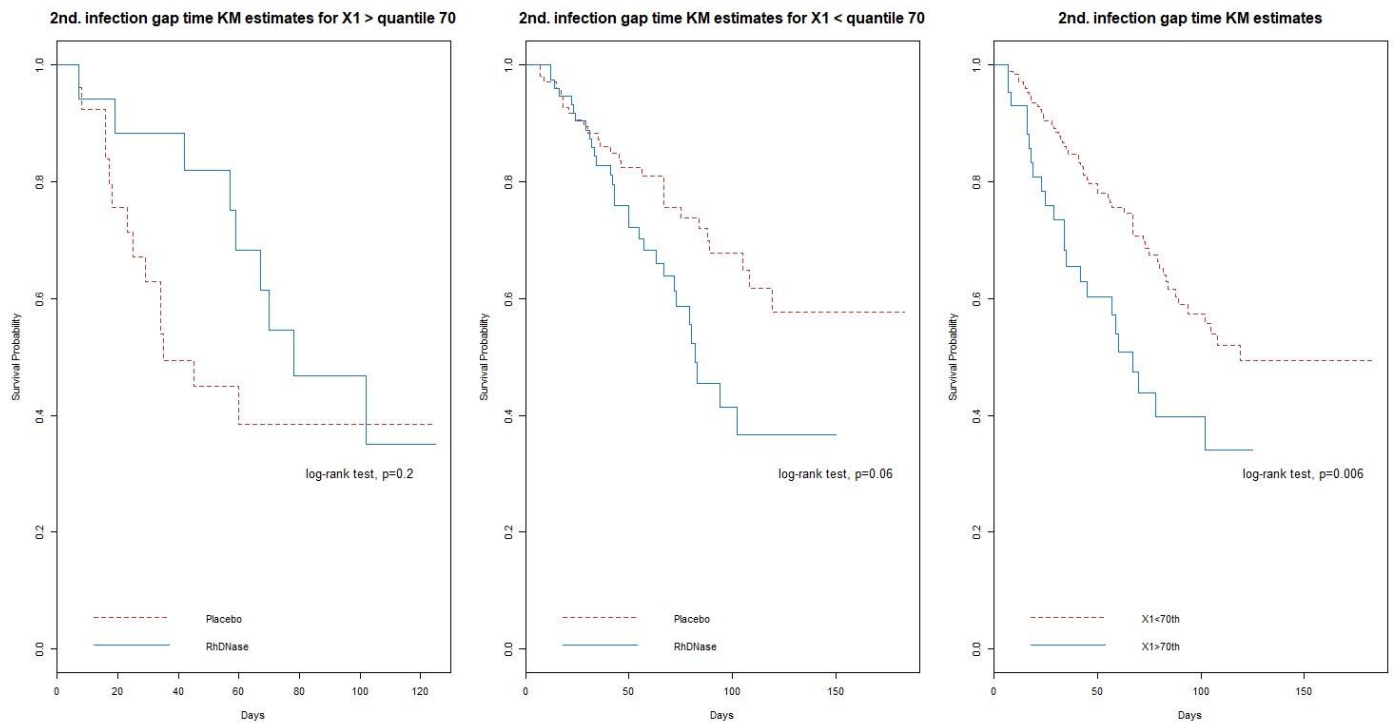


圖 8、Kaplan Meier 圖

第六章 討論與總結



本文探討有序間隔時間資料，並以條件分位數長度比作為二元間隔時間相關性測量指標，旨在提供一較彈性分析指標，可探索二元間隔時間在給定不同的 τ_1 、 τ_2 下區域相關性以及相關性隨分位數變動情形。過去在探討二元時間相關性多使用條件風險比(Clayton, 1978)或分位數勝算比(Li, 2014)，本文使用條件分位數長度比，在解釋上較為直觀易懂，估計上利用設限機率倒數權重加權建構動差估計式，接著尋找 L_1 凸函數最小值。工作迴歸模型(working quantile regression model) 使得交叉分位數比(CQR)可加入其他共變量，並且可使用現有套件進行估計。

模擬假設五種不同相關性結構，三種阿基米德 Copula，Lognormal 以及分段常數克萊頓。分別模擬不同設限率、相關性強度、樣本數下的相對偏誤以及不同標準差估計值比較。結果顯示，在有限的樣本數(200)、中等相關性(0.3)以及中等的設限率(0.24, 0.4)下， $\beta_1(\tau_1, \tau_2)$ 之相對偏誤均低於 5%，標準差估計在小樣本情況下，如高設限率以及高分位數等情境以拔靴法最好，中等樣本數情況下以重新取樣法估計最接近 1000 次模擬估計值之標準差，而本文提供之方法在大樣本下有較好表現。高設限或高分位數需較大樣本數才有較可靠之點估計。

實際資料分析顯示在缺乏資料欄位時，間隔時間之長短可提供病人過去感染的歷史資訊，而交叉分位數比(CQR)可提供研究者一彈性的框架探索資料。研究者可就全部分位數進行探討或就感興趣分位數進行解釋，可惜本文並無探討交叉分位數比(CQR)為常數之假設檢定，已有許多過去研究進行相似探討。若能更進一步進行假設檢定，則可更進一步精簡模型。另外，在分析時，第一分位數(τ_1)不應高於有觀察資料比例($\delta_1=1$)。

在過去已有許多二元事件時間相關性，如文獻回顧提及之分位數勝算比

(qor)以及常數交叉比。比較 $CQR_{Y_1, \tau_1}(\tau_2)$ 與 Li et al. (2014) 提出分位數勝算比

$$(\text{quantile odds ratio}), \quad qor(\tau | Z) = \frac{\text{odds}\{Y_1 \leq Q_1(\tau | Z) | Y_2 \leq Q_2(\tau | Z) | Z\}}{\text{odds}\{Y_1 \leq Q_1(\tau | Z) | Y_2 > Q_2(\tau | Z) | Z\}},$$

可發現 $CQR_{Y_1, \tau_1}(\tau_2) = \frac{Q_{\tau_2}(Y_2 | Y_1 > Q_{\tau_1})}{Q_{\tau_2}(Y_2 | Y_1 \leq Q_{\tau_1})}$ 對比於 $qor(\tau | Z)$ 在概念上是屬於相對風

險 (relative risk, RR) 的比較，只是其為累積分佈函數(cumulative distribution function)的反函數比例，具有時間長度解釋意義以及可作為預測模型。序列間隔時間資料是屬於縱向研究資料，因此 $CQR_{Y_1, \tau_1}(\tau_2)$ 隨著時間改變保留資料時序性結構並具預測性解釋。

總結來說， $CQR_{Y_1, \tau_1}(\tau_2)$ 具有彈性、不須模式假設、保留時序性架構以及時間長度比較值觀解釋等優點，但相對的，需大樣本以及 τ_1 、 τ_2 選擇為資料驅動(data driven)為其缺點。在有限的樣本數(200)、中等相關性(0.3)以及中等的設限率(0.24, 0.4)下估計偏誤在 5%以下，resampling 是估計的標準差最小，而本文提出之標準差估計法最為保守。

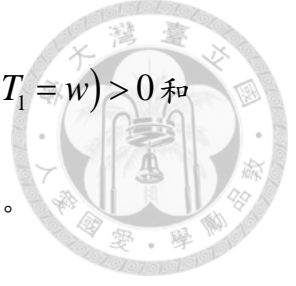
第七章 參考文獻



- Brown, B. M., Wang, Y.G. (2005). Standard errors and covariance matrices for smoothed rank estimators. *Biometrika* 92, 149–158.
- Clayton, D. G. (1978). A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* 65, 141–151.
- Fu, T.C., Su, D.H., Chang, S.H. (2016) Serial association analyses of recurrent gap time data via Kendall's tau. *Biostatistics* 17, 188–202.
- Fuchs, H.J., Borowitz, D., Christiansen, D., Morris, E., Nash, M., Ramsey, B., Rosenstein, B.J., Smith, A.L. and Wohl, M.E. (1994). The effect of aerosolized recombinant human DNase on respiratory exacerbations and pulmonary function in patients with cystic fibrosis. *New England Journal of Medicine* 331, 637–642.
- Jin, Z., Ying, Z., and Wei, L. J. (2001). A Simple Resampling Method by Perturbing the Minimand. *Biometrika* 88, 381–390.
- Lakhal-Chaieb, L., Cook, R. J. and Lin, X. (2010). Inverse probability of censoring weighted estimated of Kendall's tau for gap time analyses. *Biometrics* 66, 1145–1152.
- Li, R., Cheng, Y., Chen, Q. and Fine, J. (2017). Quantile association for bivariate survival data. *Biometrics*, 73, 506–516.
- Lin, D. Y. and Ying, Z. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika* 80, 573–581.
- Nan, B., Lin, X., Lisabeth, L. D., and Harlow, S. D. (2006). Piecewise constant cross-ratio estimation for association of age at a marker event and age at menopause. *Journal of the American Statistical Association* 101, 65–77.

- Pang, L., Lu, W., and Wang, H. (2012). Variance Estimation in Censored Quantile Regression via Induced Smoothing. *Computational Statistics & Data Analysis* 56, 785–796.
- Peng, L., Fine, J. (2009). Competing risks quantile regression. *Journal of the American Statistical Association* 104, 1440–1453.
- Peng, L., and Huang, J. (2008). Survival Analysis with Quantile Regression Models. *Journal of the American Statistical Association* 103, 637-649.
- Yang, J., Peng, L., (2016). A new flexible dependence measure for semi-competing risks. *Biometrics* 72, 770– 779

附錄



C1. (I) 存在 $v > 0$ 、 $w > 0$ 使得 $P(C = v) > 0$ 和 $P(C > v) = 0$ 以及 $P(T_1 = w) > 0$ 和

$P(T_1 > w) = 0$ 。(II) $f_1(y_1) = \frac{d \Pr(\log Y_1 \leq y_1)}{dy_1}$ 對於有界導數為連續可微。

C2. $Z(\tau_1)$ 為均勻有界(uniformly bounded)。

C3. (I) $\beta(\tau_1, \tau_2)$ 在 $\tau \in [\tau_L, \tau_U]$ 是利普希茨連續(Lipschitz continuity)， $0 < \tau_L < \tau_U < 1$;

$f_2(t | Z(\tau_1))$ 是連續的並且在給定 τ_1 、 $Z(\tau_1)$ 、 t 為均勻上有界(bounded above uniformly)，

$f_2(t | Z(\tau_1)) = dF_2(t | Z(\tau_1)) / dt$ ， $F_2(t | Z(\tau_1)) = E[I(Y_2 < t) | Z(\tau_1)]$ ；(II) 存在 $k_0 > 0$ 使得

$\tilde{A}_n(q, \Gamma)$ 的最小特徵值均勻 $> k_0$

C4. 對於某些常數 $\rho_0 > 0$ 、 $c_0 > 0$ ，有 $\inf_{\{b \in B(\rho_0), \tau_1 \in [\tau_{1L}, \tau_{1U}], \tau_2 \in [\tau_{2L}, \tau_{2U}]\}} \text{eig min } H(b, \tau_1) \geq c_0$ 。

$$B(\rho_0) = \left\{ b \in \mathbb{R}^2, \inf_{\tau_1 \in [\tau_{1L}, \tau_{1U}], \tau_2 \in [\tau_{2L}, \tau_{2U}]} \|b - \beta(\tau_1, \tau_2)\| \leq \rho_0 \right\},$$

$$H(b, \tau_1) = E \left[\mathbf{Z}(\tau_1)^{\otimes 2} f_2(\exp(\mathbf{Z}^T(\tau_1)b) | \mathbf{Z}^T(\tau_1)) \exp(\mathbf{Z}^T(\tau_1)b) \right]$$

A. 理論 A 之證明 (consistency)

$$\text{令 } S_n(\mathbf{b}, \tau_1, \tau_2) = n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \mathbf{b}) - \tau_2 \right\},$$

$$S_n^G(\mathbf{b}, \tau_1, \tau_2) = n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \mathbf{b}) - \tau_2 \right\},$$

$$\boldsymbol{\mu}(\mathbf{b}, \tau_1, \tau_2) = E \left[n^{-1/2} S_n^G(\mathbf{b}, \tau_1, \tau_2) \right]$$

根據條件 C1， $G(t)$ 和 $P(Y_1 > t)$ 對於所有 $t < v$ 及 $t < w$ 為有界，可寫下

$$\sup_{t < v} |\hat{G}(t) - G(t)| = o(n^{-1/2+r}) \text{ 對於所有 } r > 0, a.s.$$

$$\sup_{t < w} |\hat{P}(Y_1 > t) - P(Y_1 > t)| = o(n^{-1/2+r}) \text{ 對於所有 } r > 0, a.s.$$

因此， $\sup_{\tau_1, \tau_2, b} \left| n^{-1/2} S_n(\mathbf{b}, \tau_1, \tau_2) - n^{-1/2} S_n^G(\mathbf{b}, \tau_1, \tau_2) \right| = o(n^{-1/2+r}), a.s.$

另外定義， $F = \left\{ \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \mathbf{b}) - \tau_2 \right\}, \mathbf{b} \in R^2, \tau_1 \in [\tau_{1L}, \tau_{1U}], \tau_2 \in [\tau_{2L}, \tau_{2U}] \right\}$

因 Z_i 與 $G(T_{2i})$ 為均勻有界，且指標函數(indicator functions)為 Donsker 類函數，所以 F 為 Donsker 類函數也為 Glivenko–Cantelli 類函數。根據 Glivenko–Cantelli 理論，

$\sup_{\tau_1, \tau_2, b} \left| n^{-1/2} S_n^G(\mathbf{b}, \tau_1, \tau_2) - \mu(\mathbf{b}, \tau_1, \tau_2) \right| = o(1), a.s.$ ，因此

$$\sup_{\tau_1, \tau_2, b} \left| n^{-1/2} S_n(\mathbf{b}, \tau_1, \tau_2) - \mu(\mathbf{b}, \tau_1, \tau_2) \right| = o(1) \dots (A.1)$$

對於任何 $\mathbf{u} \in R^2$ 滿足 $\|\mathbf{u}\|^2 = 1$ ，且 $\delta \geq \rho_0$ ，

$$\mathbf{u}^T \left\{ \mu(\beta + \mathbf{u}\delta, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \geq \mathbf{u}^T \left\{ \mu(\beta + \mathbf{u}\rho_0, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \geq 0$$

根據柯西不等式，

$$\begin{aligned} \|\mathbf{u}\|^2 \left\| \left\{ \mu(\beta + \mathbf{u}\delta, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \right\|^2 &\geq \\ \left(\mathbf{u}^T \left\{ \mu(\beta + \mathbf{u}\delta, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \right)^2 &\geq \\ \left(\mathbf{u}^T \left\{ \mu(\beta + \mathbf{u}\rho_0, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \right)^2 &= \\ \left(\mathbf{u}^T \left\{ H(\beta + \mathbf{u}\rho^*, \tau_1) \mathbf{u}\rho_0 \right\} \right)^2 &\geq c_0^2 \rho_0^2, \quad \rho^* \in [0, \rho_0] \end{aligned}$$

根據條件 C4 且 $\mu(\mathbf{b}, \tau_1, \tau_2)$ 為 \mathbf{b} 之單調函數，

$$\inf_{b \in B(\rho_0)} \left| \mu(\mathbf{b}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right| \geq c_0 \rho_0$$

又 $\mu(\beta, \tau_1, \tau_2) = 0$ ， $n^{-1/2} S_n(\hat{\beta}, \tau_1, \tau_2) = o(1), a.s.$

$$\begin{aligned} &\mu(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \\ &= n^{-1/2} S_n(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) - \left[n^{-1/2} S_n(\hat{\beta}, \tau_1, \tau_2) - \mu(\hat{\beta}, \tau_1, \tau_2) \right] \end{aligned}$$

根據 A.1 可知，

$$\sup_{\tau_1, \tau_2} \left| \mu(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right| = o(1) \dots (A.2)$$

因此存在一 $N_0 > 0$ ，對於任何 $n \geq N_0$ ，使得 $\sup_{\tau_1, \tau_2} \left| \mu(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right| \leq c_0 \rho_0 / 2$ 的機率為 1。

在 $\beta(\tau_1, \tau_2)$ 附近使用泰勒展開式展開 $\mu(\hat{\beta}(\tau_1, \tau_2), \tau_1, \tau_2)$,

$$\begin{aligned} & \sup_{\tau_1, \tau_2} \|\hat{\beta}(\tau_1, \tau_2) - \beta(\tau_1, \tau_2)\| \\ &= \sup_{\tau_1, \tau_2} \left\| H \left\{ \tilde{\beta}(\tau_1, \tau_2), \tau_1 \right\}^{-1} \left\{ \mu(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\} \right\| \\ &\leq c_0^{-1} \sup_{\tau_1, \tau_2} \left\| \mu(\hat{\beta}, \tau_1, \tau_2) - \mu(\beta, \tau_1, \tau_2) \right\| \end{aligned}$$

$\tilde{\beta}(\tau_1, \tau_2)$ 為介於 $\hat{\beta}(\tau_1, \tau_2)$ 和 $\beta(\tau_1, \tau_2)$ 之間值，也因此當 n 很大時 $\tilde{\beta}(\tau_1, \tau_2) \in B(\rho_0)$ 。

根據 A.1 以及 C4 可得 $\hat{\beta}(\tau_1, \tau_2)$ 均勻一致性



B. 理論 B 之證明 (asymptotic normal)

將估計式拆解成以下三項，逐項討論

$$\begin{aligned} S_n(\hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) &= \\ & n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} - n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} \dots (i) \\ &+ n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} - n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} \dots (ii) \\ &+ n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} \dots (iii) \end{aligned}$$

$$\begin{aligned} \text{i. } & n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \beta) - \tau_2 \right\} \\ &= n^{-1/2} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ \hat{\mathbf{Z}}_i^*(\tau_1) I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \beta) - \hat{\mathbf{Z}}_i^*(\tau_1) \tau_2 \right\} \end{aligned}$$

令 $\hat{Q}_{\tau_1} = \inf(t : \hat{P}(\tilde{Y}_1 \leq t) \geq \tau_1)$ ，將截距項與共變量拆成以下兩個式子

$$n^{-1/2} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{li} > \log \hat{Q}_{\tau_1}, \log \tilde{Y}_{2i} \leq \beta_0 + \beta_1) + I(\log \tilde{Y}_{li} \leq \log \hat{Q}_{\tau_1}, \log \tilde{Y}_{2i} \leq \beta_0) - \tau_2 \right\} \dots (1)$$

$$n^{-1/2} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{li} > \log \hat{Q}_{\tau_1}, \log \tilde{Y}_{2i} \leq \beta_0 + \beta_1) - I(\log \tilde{Y}_{li} > \log \hat{Q}_{\tau_1}) \tau_2 \right\} \dots (2)$$

$$\text{令 } C_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) = n^{-1} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{1i} > \log \hat{Q}_{\tau_1}, \log \tilde{Y}_{2i} \leq \beta_0 + \beta_1) \right\},$$



$$D_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) = n^{-1} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{1i} \leq \log \hat{Q}_{\tau_1}, \log \tilde{Y}_{2i} \leq \beta_0) \right\},$$

$$E_n(\log \hat{Q}_{\tau_1}) = n^{-1} \sum_{i=1}^n \frac{\delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{1i} > \log \hat{Q}_{\tau_1}) \tau_2 \right\},$$

$$\mu_C(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) = E \left[C_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) \right],$$

$$\mu_D(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) = E \left[D_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) \right],$$

$$\mu_E(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) = E \left[E_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) \right].$$

Lemma 1. 對任何數列 $\{\beta_{1n}(\tau_1), \beta_{2n}(\tau_1)\}_{n=1}^{\infty}$ 滿足

$$\sup_{\tau \in [\tau_L, \tau_U]} \|\tilde{\beta}_{jn}(\tau) - \beta_{j0}(\tau)\| = o_p(1), j=1, 2,$$

則有

$$\sup_{\tau \in [\tau_L, \tau_U]} \|C_n(\tilde{\beta}_{1n}(\tau_1), \tilde{\beta}_{2n}(\tau_2)) - C_n(\beta_{10}(\tau_1), \beta_{20}(\tau_2)) - \mu_C(\tilde{\beta}_{1n}(\tau), \tilde{\beta}_{2n}(\tau)) + \mu_C(\beta_1(\tau), \beta_2(\tau))\| = o_p(n^{-1/2})$$

$$\text{令 } X(b_1, b_2; \tau) = I(\log Y_1 \leq b_1(\tau_1), \log Y_2 \leq Z^T b_2(\tau_1, \tau_2)) - I(\log Y_1 \leq \beta_1(\tau_1), \log Y_2 \leq Z^T \beta_2(\tau_1, \tau_2))$$

$$\sigma^2(b_1, b_2; \tau) = \text{var}(X(b_1, b_2; \tau)), \text{ 則}$$

$$\begin{aligned} \sigma^2(\tilde{\beta}_{1n}, \tilde{\beta}_{2n}; \tau) &\leq E \left\{ I(\log Y_1 \leq \tilde{\beta}_{1n}(\tau_1), \log Y_2 \leq Z^T \tilde{\beta}_{2n}(\tau_1, \tau_2)) - I(\log Y_1 \leq \beta_1(\tau_1), \log Y_2 \leq Z^T \beta_2(\tau_1, \tau_2)) \right\}^2 \\ &= E \left\{ I(\log Y_1 \leq \tilde{\beta}_{1n}(\tau_1), \log Y_2 \leq Z^T \tilde{\beta}_{2n}(\tau_1, \tau_2)) - I(\log Y_1 \leq \beta_1(\tau_1), \log Y_2 \leq Z^T \beta_2(\tau_1, \tau_2)) \right\} \\ &\leq E \left\{ I(\log Y_1 \leq \tilde{\beta}_{1n}(\tau_1)) - I(\log Y_1 \leq \beta_1(\tau_1)) \right\} + \\ &\quad E \left\{ I(\log Y_2 \leq Z^T \tilde{\beta}_{2n}(\tau_1, \tau_2)) - I(\log Y_2 \leq Z^T \beta_2(\tau_1, \tau_2)) \right\} \end{aligned}$$

$$\text{因為 } f_1(t) \text{ 及 } f_2(t|Z) \text{ 為有界函數，因此 } E \left\{ I(\log Y_1 \leq \tilde{\beta}_{1n}(\tau_1)) - I(\log Y_1 \leq \beta_1(\tau_1)) \right\} = o_p(1),$$

$$E \left\{ I(\log Y_2 \leq Z^T \tilde{\beta}_{2n}(\tau_1, \tau_2)) - I(\log Y_2 \leq Z^T \beta_2(\tau_1, \tau_2)) \right\} = o_p(1).$$

故 $\sup_{\tau} \sigma^2(\tilde{\beta}_{1n}, \tilde{\beta}_{2n}; \tau) = o_p(1)$, 得證 Lemma 1. ◦

根據 Lemma 1. 以及 $\log \hat{Q}_{\tau_1}$ 均勻一致性(uniform consistency) , 利用泰勒展示可發現 ,

$$\begin{aligned} & \sqrt{n} \left\{ C_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) - C_n(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \right\} \approx \sqrt{n} \left\{ \mu_C(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) - \mu_C(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \right\} \\ & \approx \mu'_C(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \times \sqrt{n} \left\{ \log \hat{Q}_{\tau_1} - \log Q_{\tau_1} \right\} \\ & \approx \mu'_C(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \times -n^{-1/2} \sum_{i=1}^n f_1^{-1}(\log Q_{\tau_1}) \left[I(\log Y_{1,i} \leq \log Q_{\tau_1}) - \tau_1 \right] \end{aligned}$$

同理

$$\begin{aligned} & \sqrt{n} \left\{ D_n(\log \hat{Q}_{\tau_1}, \beta(\tau_1, \tau_2)) - D_n(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \right\} \\ & \approx \mu'_D(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \times -n^{-1/2} \sum_{i=1}^n f_1^{-1}(\log Q_{\tau_1}) \left[I(\log Y_{1,i} \leq \log Q_{\tau_1}) - \tau_1 \right] \\ & \sqrt{n} \left\{ E_n(\log \hat{Q}_{\tau_1}) - E_n(\log Q_{\tau_1}) \right\} \\ & \approx \mu'_E(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) \times -n^{-1/2} \sum_{i=1}^n f_1^{-1}(\log Q_{\tau_1}) \left[I(\log Y_{1,i} \leq \log Q_{\tau_1}) - \tau_1 \right] \\ & \left[\begin{matrix} \mu'_C + \mu'_D \\ \mu'_C - \mu'_E \end{matrix} \right] \times -n^{-1/2} \sum_{i=1}^n f_1^{-1}(\log Q_{\tau_1}) \left[I(\log Y_{1,i} \leq \log \hat{Q}_{\tau_1}) - \tau_1 \right] \equiv n^{-1/2} \sum_{i=1}^n \xi_{1,i} \end{aligned}$$

ii. 令 $N_i^G(t) = I(\tilde{T}_{2i} \leq t, \delta_{2i} = 0)$, $Y_i(t) = I(\tilde{T}_2 \geq t)$, $y(t) = \Pr(\tilde{T}_2 \geq t)$,

$$\lambda^G(t) = \lim_{\Delta \rightarrow 0} \frac{\Pr(\tilde{T}_2 \in (t, t + \Delta), \delta_{2i} = 0 | \tilde{T}_2 \geq t)}{\Delta} , \quad \Lambda^G(t) = \int_0^t \lambda^G(s) ds ,$$

$$M_i^G(t) = N_i^G(t) - \int_0^t Y_i(s) d\Lambda^G(s) ,$$

$$\omega(\beta, \tau_1, \tau_2) = E \left\{ \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i} Y_i(s) \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \beta) - \tau_2 \right\}}{G(\tilde{T}_{2i})} \right\}$$

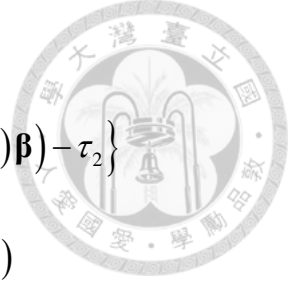
$$\text{根據 Pepe (1991), } \sup_{t \in [0, v)} \left\| n^{1/2} (\hat{G}(t) - G(t)) - n^{-1/2} \sum_{i=1}^n G(t) \int_0^t y(s)^{-1} dM_i^G(s) \right\| \rightarrow 0$$

$$\begin{aligned}
& n^{-1/2} \sum_{i=1}^n \frac{\hat{G}(\tilde{T}_{2i}) - G(\tilde{T}_{2i})}{\hat{G}(\tilde{T}_{2i}) G(\tilde{T}_{2i})} \mathbf{Z}_i^*(\tau_1) \delta_{2i} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \boldsymbol{\beta}) - \tau_2 \right\} \\
& \approx n^{-1/2} \sum_{i=1}^n \frac{n^{-1} \sum_{j=1}^n \int_0^\infty Y_j(s) y(s)^{-1} dM_j^G(s)}{G(\tilde{T}_{2i})} \mathbf{Z}_i^*(\tau_1) \delta_{2i} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \boldsymbol{\beta}) - \tau_2 \right\} \\
& = n^{-1/2} \sum_{j=1}^n \int_0^\infty \left\{ \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i} Y_i(s) \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \boldsymbol{\beta}) - \tau_2 \right\}}{n G(\tilde{T}_{2i})} \right\} \frac{dM_j^G(s)}{y(s)} \\
& \approx n^{-1/2} \sum_{j=1}^n \int_0^\infty \omega(\boldsymbol{\beta}, \tau_1, \tau_2) \frac{dM_j^G(s)}{y(s)} \equiv n^{-1/2} \sum_{i=1}^n \xi_{2,i}
\end{aligned}$$

$$\text{iii. } n^{-1/2} \sum_{i=1}^n \xi_{3,i} \equiv n^{-1/2} \sum_{i=1}^n \frac{\mathbf{Z}_i^*(\tau_1) \delta_{2i}}{G(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \mathbf{Z}_i^{*T}(\tau_1) \boldsymbol{\beta}) - \tau_2 \right\}$$

因此

$$\begin{aligned}
S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) & \approx n^{-1/2} \sum_{i=1}^n \left\{ \xi_{1,i} - \xi_{2,i} + \xi_{3,i} \right\} \equiv n^{-1/2} \sum_{i=1}^n t_i \\
\Sigma(\tau_1, \tau_2) & = E \left[\boldsymbol{\iota}(\tau_1, \tau_2) \boldsymbol{\iota}(\tau_1, \tau_2)^T \right] \\
S_n(\hat{Q}_{\tau_1}, \hat{\boldsymbol{\beta}}(\tau_1, \tau_2)) - S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) & \approx H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1) n^{1/2} (\hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2)) \\
n^{1/2} (\hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2)) & \approx -H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1)^{-1} S_n(\hat{Q}_{\tau_1}, \boldsymbol{\beta}(\tau_1, \tau_2)) \\
n^{1/2} (\hat{\boldsymbol{\beta}}(\tau_1, \tau_2) - \boldsymbol{\beta}(\tau_1, \tau_2)) & \text{隨機收斂至高斯過程，其共變異數為} \\
& H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1)^{-1} E \left[\boldsymbol{\iota}(\tau_1, \tau_2) \boldsymbol{\iota}(\tau_1, \tau_2)^T \right] H(\boldsymbol{\beta}(\tau_1, \tau_2), \tau_1)^{-T}
\end{aligned}$$



C. 共變異數之估計步驟

1. $n^{1/2} \{ \log \hat{Q}_{\tau_1} - \log Q_{\tau_1} \}$ 變異數估計

參考 Pang et al. (2012) 誘導平滑方法 (induced smoothing) ,

令 ϕ 以及 Φ 為標準常態的機率密度函數與累積分布函數, Ξ 為 $n^{1/2} \{ \log \hat{Q}_{\tau_1} - \log Q_{\tau_1} \}$ 變異數, $\Gamma = \Xi / n$. $\Sigma = \text{var}(n^{1/2} S_n(q, \tau_1))$,

$$\hat{\Sigma}(q) = n^{-1} \sum_{i=1}^n \hat{\varsigma}_i^{\otimes 2} \quad ,$$

$$\hat{\varsigma}_i = \frac{\delta_{1,i}}{\hat{G}(\tilde{T}_{1,i})} \left\{ I(\log \tilde{Y}_{1,i} \leq \log \hat{Q}_{\tau_1}) - \tau_1 \right\} - I(\delta_{1,i} = 0) \left\{ \frac{\sum_{j=1}^n I(\tilde{T}_{1,j} \geq \tilde{T}_{1,i}) \frac{\delta_{1,j}}{\hat{G}(\tilde{T}_{1,j})} \left\{ I(\log \tilde{Y}_{1,j} \leq \log \hat{Q}_{\tau_1}) - \tau_1 \right\}}{\sum_{j=1}^n I(\tilde{T}_{1,j} \geq \tilde{T}_{1,i})} \right\}$$

$$\text{建立平滑估計式 } \tilde{S}_n(q, \Gamma) = n^{-1} \sum_{i=1}^n \frac{\delta_{1,i}}{\hat{G}(\tilde{T}_{1,i})} \left\{ \Phi \left(\frac{q - \log \tilde{T}_{1,i}}{\sqrt{\Gamma}} \right) - \tau_1 \right\} ,$$

$$\tilde{A}_n(q, \Gamma) = \frac{\partial \tilde{S}_n(q, \Gamma)}{\partial q} = n^{-1} \sum_{i=1}^n \frac{\delta_{1,i}}{\hat{G}(\tilde{T}_{1,i}) \sqrt{\Gamma}} \left\{ \phi \left(\frac{q - \log \tilde{T}_{1,i}}{\sqrt{\Gamma}} \right) \right\} .$$

迭代步驟如下 ,

$$\text{I. 給定初始值 } \tilde{\Gamma}_0 = 1/n \quad , \quad \tilde{q}_0 = \log \hat{Q}_{\tau_1}$$

$$\text{II. } \tilde{q}_k = \tilde{q}_{k-1} + \left\{ -\tilde{A}_n(\tilde{q}_{k-1}, \tilde{\Gamma}_{k-1}) \right\}^{-1} \tilde{S}_n(\tilde{q}_{k-1}, \tilde{\Gamma}_{k-1})$$

$$\tilde{\Gamma}_k = n^{-1} \left\{ \tilde{A}_n(\tilde{q}_k, \tilde{\Gamma}_{k-1}) \right\}^{-1} \hat{\Sigma}(\tilde{q}_k) \left\{ \tilde{A}_n(\tilde{q}_k, \tilde{\Gamma}_{k-1}) \right\}^{-1}$$

重複步驟 II. 直至收斂, 估計值記作 $\hat{\Gamma}$.

$$2. \quad S_n(\hat{Q}_{\tau_1}, \mathbf{\beta}(\tau_1, \tau_2)) \approx n^{-1/2} \sum_{i=1}^n \{ \xi_{1,i} - \xi_{2,i} + \xi_{3,i} \} \equiv n^{-1/2} \sum_{i=1}^n \iota_i \text{ 之估計}$$

$$\text{I. } \xi_{1,i} \text{ 之估計}$$



參考 Li et al. (2014)，因 $\hat{\xi}_{1,i}$ 計算包含未知機率密度函數， $h_1(y_1, y_2)$

$= \partial P(\log Y_1 \leq y_1, \log Y_2 \leq y_2) / \partial y_1$ ， $f_1(y_1) = dP(\log Y_1 \leq y_1) / dy_1$ 等，再次利用 Pang et al. (2012)

誘導平滑概念作估計。令 $\mu'_C(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) = -h_1(\log Q_{\tau_1}, \beta_0 + \beta_1)$ ，

$\mu'_D(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) = h_1(\log Q_{\tau_1}, \beta_0)$ ， $\mu'_E(\log Q_{\tau_1}, \beta(\tau_1, \tau_2)) = -f_1(\log Q_{\tau_1})$ ，

$$\hat{\mu}'_C(\log \hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) = n^{-1} \sum_{i=1}^n -\phi\left(\frac{\log \hat{Q}_{\tau_1} - \log \tilde{Y}_{1,i}}{\sqrt{\hat{\Gamma}}}\right) \times \frac{I(\log \tilde{Y}_{2,i} \leq \hat{\beta}_0 + \hat{\beta}_1) \delta_{2,i}}{\hat{G}(\tilde{T}_{2,i}) \sqrt{\hat{\Gamma}}},$$

$$\hat{\mu}'_D(\log \hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) = n^{-1} \sum_{i=1}^n \phi\left(\frac{\log \hat{Q}_{\tau_1} - \log \tilde{Y}_{1,i}}{\sqrt{\hat{\Gamma}}}\right) \times \frac{I(\log \tilde{Y}_{2,i} \leq \hat{\beta}_0) \delta_{2,i}}{\hat{G}(\tilde{T}_{2,i}) \sqrt{\hat{\Gamma}}},$$

$$\hat{\mu}'_E(\log \hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) = n^{-1} \sum_{i=1}^n -\phi\left(\frac{\log \hat{Q}_{\tau_1} - \log \tilde{Y}_{1,i}}{\sqrt{\hat{\Gamma}}}\right) \frac{\delta_{2,i}}{\hat{G}(\tilde{T}_{2,i}) \sqrt{\hat{\Gamma}}},$$

$$\hat{f}_1(\log \hat{Q}_{\tau_1}) = n^{-1} \sum_{i=1}^n \phi\left(\frac{\log \hat{Q}_{\tau_1} - \log \tilde{Y}_{1,i}}{\sqrt{\hat{\Gamma}}}\right) \frac{\delta_{1,i}}{\hat{G}(\tilde{T}_{1,i}) \sqrt{\hat{\Gamma}}},$$

$$\hat{\xi}_{1,i} = \begin{bmatrix} \hat{\mu}'_C + \hat{\mu}'_D \\ \hat{\mu}'_C - \hat{\mu}'_E \end{bmatrix} \times \hat{f}_1^{-1}(\log \hat{Q}_{\tau_1}) \left[I(\log \tilde{Y}_{1,i} \leq \log \hat{Q}_{\tau_1}) - \tau_1 \right]$$

II. $\xi_{2,i}$ ， $\xi_{3,i}$ 之估計

參考 Yang and Peng(2016)，

$$\hat{\xi}_{3,i} - \hat{\xi}_{2,i} = \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2,i}}{\hat{G}(\tilde{T}_{2,i})} \left\{ I(\log \tilde{Y}_{2,i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\}$$

$$- I(\delta_{2,i} = 0) \left\{ \frac{\sum_{j=1}^n I(\tilde{T}_{2,j} \geq \tilde{T}_{2,i}) \frac{\hat{\mathbf{Z}}_j^*(\tau_1) \delta_{2,j}}{\hat{G}(\tilde{T}_{2,j})} \left\{ I(\log \tilde{Y}_{2,j} \leq \hat{\mathbf{Z}}_j^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\}}{\sum_{j=1}^n I(\tilde{T}_{2,j} \geq \tilde{T}_{2,i})} \right\}$$

3. $H(\beta(\tau_1, \tau_2), \tau_1)^{-1} \approx \sqrt{n} \Lambda_n(\tau_1, \tau_2) E_n^{-1}(\tau_1, \tau_2)$ 之估計

參考 Yang and Peng(2016)，

$$\hat{\Sigma}(\tau_1, \tau_2) = n^{-1} \sum_{i=1}^n \hat{t}_i^{\otimes 2}, \quad \hat{t}_i = \left\{ \hat{\xi}_{1,i} - \hat{\xi}_{2,i} + \hat{\xi}_{3,i} \right\}$$

I. 特徵分解找出一對稱矩陣使得 $\hat{\Sigma}(\tau_1, \tau_2) = E_n^2(\tau_1, \tau_2)$

II. 計算 $\Lambda_n(\tau_1, \tau_2) = [\hat{\beta}_{e_{n,1}} - \hat{\beta}, \hat{\beta}_{e_{n,2}} - \hat{\beta}]$

其中 $\hat{\beta}_{e_{n,j}}$ 為 $S_n(b, \tau_1, \tau_2) - e_{n,j} = 0$ 的解，而 $e_{n,j}$ 為 $E_n(\tau_1, \tau_2)$ 第 j 欄。

$$\begin{aligned} \text{變異數 } \hat{\Omega}(\tau_1, \tau_2) &= \sqrt{n} \Lambda_n(\tau_1, \tau_2) E_n^{-1}(\tau_1, \tau_2) \hat{\Sigma}(\tau_1, \tau_2) \left\{ \sqrt{n} \Lambda_n(\tau_1, \tau_2) E_n^{-1}(\tau_1, \tau_2) \right\}^T \\ &= n \Lambda_n^{\otimes 2}(\tau_1, \tau_2) \end{aligned}$$

補充: $S_n(b, \tau_1, \tau_2) - e_{n,j} = 0$

$$\left[n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\} \right] - e_{n,j} = 0$$

等同於找尋以下凸函數的最小值，

$$U_n(b, \tau) = \sum_{i=1}^n \delta_{2i} \left| \frac{\log \tilde{Y}_{2i}}{\hat{G}(\tilde{T}_{2,i})} - b^T \frac{\hat{\mathbf{Z}}_i^*(\tau_1)}{\hat{G}(\tilde{T}_{2,i})} \right| + \left| M - b^T \left\{ (2\tau_2 - 1) \sum_{l=1}^n \frac{\hat{\mathbf{Z}}_l^*(\tau_1) I(\delta_{2,l}=1)}{\hat{G}(\tilde{T}_{2,l})} + 2\sqrt{n} e_{n,j} \right\} \right|$$



$$\text{令 } S_n^*(\mathbf{b}, \tau_1, \tau_2) = n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \zeta_i \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \mathbf{b}) - \tau_2 \right\} = 0$$

$$\text{給定 } E[\zeta_i] = 1$$

$$S_n^*(\hat{Q}_{\tau_1}, \beta^*(\tau_1, \tau_2)) - S_n^*(\hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) \approx H(\beta(\tau_1, \tau_2), \tau_1) n^{1/2} (\beta^*(\tau_1, \tau_2) - \hat{\beta}(\tau_1, \tau_2))$$

$$n^{1/2} (\beta^*(\tau_1, \tau_2) - \hat{\beta}(\tau_1, \tau_2)) \approx -H(\beta(\tau_1, \tau_2), \tau_1)^{-1} S_n^*(\hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) \dots (1)$$

$n^{1/2} (\beta^*(\tau_1, \tau_2) - \hat{\beta}(\tau_1, \tau_2))$ 隨機收斂至高斯過程，其共變異數為

$$H(\beta(\tau_1, \tau_2), \tau_1)^{-1} \text{var} \left[S_n^*(\hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) \right] H(\beta(\tau_1, \tau_2), \tau_1)^{-T}$$

$$S_n^*(\beta^*, \tau_1, \tau_2) = o(1) \text{ a.s.}$$

$$S_n(\hat{\beta}, \tau_1, \tau_2) = o(1) \text{ a.s.}$$

$$-n^{1/2} S_n^*(\hat{\beta}, \tau_1, \tau_2) = \left\{ n^{1/2} S_n(\hat{\beta}, \tau_1, \tau_2) - n^{1/2} S_n^*(\hat{\beta}, \tau_1, \tau_2) \right\} + o(1)$$

$$= n^{-1/2} \sum_{i=1}^n \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\}$$

$$-n^{-1/2} \sum_{i=1}^n \zeta_i \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\} + o(1)$$

$$= n^{-1/2} \sum_{i=1}^n (1 - \zeta_i) \hat{l}_i + o(1)$$

$$\hat{l}_i = \frac{\hat{\mathbf{Z}}_i^*(\tau_1) \delta_{2i}}{\hat{G}(\tilde{T}_{2i})} \left\{ I(\log \tilde{Y}_{2i} \leq \hat{\mathbf{Z}}_i^{*T}(\tau_1) \hat{\beta}) - \tau_2 \right\}$$

$$E \left[S_n^*(\hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2)) S_n^*(\hat{Q}_{\tau_1}, \hat{\beta}(\tau_1, \tau_2))^T \mid \{\tilde{Y}_{1,i}, \tilde{Y}_{2,i}, \delta_{1,i}, \delta_{2,i}\}_{i=1}^n \right] = n^{-1} \sum_{i=1}^n \hat{l}_i^{\otimes 2}(\tau_1, \tau_2)$$

$$\xrightarrow{P} \Sigma(\tau_1, \tau_2) = E \left[\iota(\tau_1, \tau_2) \iota(\tau_1, \tau_2)^T \right] \dots (2)$$

根據(1)式與(2)式，可得 $n^{1/2} (\beta^*(\tau_1, \tau_2) - \hat{\beta}(\tau_1, \tau_2))$ 在給定資料下與 $n^{1/2} (\hat{\beta}(\tau_1, \tau_2) - \beta^*(\tau_1, \tau_2))$ 有相同的漸進分布。

E. 模擬結果表格

表 1、Clayton 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.4405	0.4376	-0.0029	-0.6684	0.1151	0.1289	0.1207	0.1449
	0.2	0.6	0.3988	0.3988	0.0000	0.0028	0.1124	0.1277	0.1198	0.1499
	0.2	0.7	0.3539	0.3534	-0.0005	-0.1385	0.1150	0.1303	0.1226	0.1612
	0.4	0.5	0.3011	0.2994	-0.0017	-0.5495	0.0819	0.0939	0.0905	0.1018
	0.4	0.6	0.2657	0.2637	-0.0020	-0.7388	0.0795	0.0920	0.0883	0.1028
	0.4	0.7	0.2305	0.2260	-0.0046	-1.9764	0.0808	0.0911	0.0871	0.1030
	0.6	0.5	0.2333	0.2308	-0.0024	-1.0441	0.0770	0.0879	0.0855	0.0936
	0.6	0.6	0.2024	0.1973	-0.0051	-2.5246	0.0751	0.0852	0.0831	0.0951
	0.6	0.7	0.1731	0.1676	-0.0055	-3.1899	0.0758	0.0833	0.0808	0.0921
(0.54,0.76)	0.2	0.5	0.4405	0.4324	-0.0082	-1.8511	0.1627	0.1899	0.1800	0.2466
	0.2	0.6	0.3988	0.3917	-0.0071	-1.7752	0.1619	0.1942	0.1811	0.2563
	0.2	0.7	0.3539	0.3495	-0.0044	-1.2489	0.1699	0.2022	0.1859	0.2709
	0.4	0.5	0.3011	0.2831	-0.0179	-5.9568	0.1451	0.1682	0.1599	0.1993
	0.4	0.6	0.2657	0.2491	-0.0166	-6.2526	0.1459	0.1692	0.1582	0.1995
	0.4	0.7	0.2305	0.2115	-0.0191	-8.2669	0.1491	0.1724	0.1582	0.2154
	0.6	0.5	0.2333	0.2175	-0.0158	-6.7631	0.1805	0.1896	0.1722	0.2188
	0.6	0.6	0.2024	0.1790	-0.0235	-11.5841	0.1795	0.1878	0.1664	0.2150
	0.6	0.7	0.1731	0.1461	-0.0271	-15.6312	0.1803	0.1863	0.1611	0.2144

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 2、Clayton 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.4405	0.4368	-0.0037	-0.8364	0.0831	0.0904	0.0862	0.1028
	0.2	0.6	0.3988	0.3947	-0.0041	-1.0229	0.0816	0.0887	0.0837	0.0982
	0.2	0.7	0.3539	0.3516	-0.0023	-0.6442	0.0816	0.0896	0.0848	0.1015
	0.4	0.5	0.3011	0.2991	-0.0020	-0.6624	0.0628	0.0645	0.0628	0.0700
	0.4	0.6	0.2657	0.2648	-0.0009	-0.3323	0.0606	0.0623	0.0604	0.0667
	0.4	0.7	0.2305	0.2302	-0.0003	-0.1239	0.0583	0.0623	0.0606	0.0692
	0.6	0.5	0.2333	0.2333	0.0001	0.0363	0.0574	0.0599	0.0587	0.0633
	0.6	0.6	0.2024	0.2033	0.0008	0.4096	0.0537	0.0582	0.0569	0.0620
	0.6	0.7	0.1731	0.1740	0.0009	0.5125	0.0523	0.0576	0.0562	0.0624
(0.54,0.76)	0.2	0.5	0.4405	0.4292	-0.0114	-2.5820	0.1200	0.1292	0.1252	0.1588
	0.2	0.6	0.3988	0.3905	-0.0082	-2.0661	0.1191	0.1317	0.1276	0.1687
	0.2	0.7	0.3539	0.3431	-0.0108	-3.0459	0.1238	0.1385	0.1327	0.1755
	0.4	0.5	0.3011	0.2917	-0.0094	-3.1269	0.1031	0.1121	0.1109	0.1277
	0.4	0.6	0.2657	0.2559	-0.0098	-3.6820	0.1022	0.1133	0.1110	0.1367
	0.4	0.7	0.2305	0.2237	-0.0068	-2.9568	0.1073	0.1177	0.1138	0.1419
	0.6	0.5	0.2333	0.2240	-0.0092	-3.9508	0.1147	0.1304	0.1245	0.1501
	0.6	0.6	0.2024	0.1949	-0.0076	-3.7497	0.1141	0.1300	0.1224	0.1507
	0.6	0.7	0.1731	0.1618	-0.0113	-6.5284	0.1187	0.1308	0.1219	0.1559

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 3、Clayton 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.6246	0.6175	-0.0071	-1.1298	0.0915	0.1027	0.0941	0.1189
	0.2	0.6	0.6084	0.6025	-0.0060	-0.9817	0.0914	0.1032	0.0930	0.1228
	0.2	0.7	0.5908	0.5851	-0.0057	-0.9650	0.0933	0.1090	0.0974	0.1381
	0.4	0.5	0.4550	0.4501	-0.0049	-1.0705	0.0704	0.0802	0.0767	0.0903
	0.4	0.6	0.4272	0.4228	-0.0044	-1.0237	0.0693	0.0803	0.0757	0.0912
	0.4	0.7	0.3988	0.3927	-0.0061	-1.5285	0.0717	0.0841	0.0785	0.0978
	0.6	0.5	0.3637	0.3591	-0.0046	-1.2711	0.0663	0.0764	0.0743	0.0838
	0.6	0.6	0.3299	0.3228	-0.0071	-2.1557	0.0679	0.0756	0.0738	0.0856
(0.54,0.76)	0.6	0.7	0.2968	0.2915	-0.0052	-1.7673	0.0718	0.0774	0.0750	0.0877
	0.2	0.5	0.6246	0.6112	-0.0133	-2.1321	0.1277	0.1492	0.1428	0.2037
	0.2	0.6	0.6084	0.5928	-0.0156	-2.5612	0.1325	0.1565	0.1457	0.2222
	0.2	0.7	0.5908	0.5733	-0.0175	-2.9573	0.1392	0.1718	0.1567	0.2582
	0.4	0.5	0.4550	0.4387	-0.0163	-3.5739	0.1202	0.1424	0.1361	0.1749
	0.4	0.6	0.4272	0.4090	-0.0182	-4.2541	0.1291	0.1484	0.1395	0.1868
	0.4	0.7	0.3988	0.3774	-0.0214	-5.3700	0.1358	0.1580	0.1448	0.2103
	0.6	0.5	0.3637	0.3442	-0.0195	-5.3572	0.1485	0.1618	0.1484	0.1992
	0.6	0.6	0.3299	0.3069	-0.0230	-6.9839	0.1533	0.1640	0.1470	0.2001
	0.6	0.7	0.2968	0.2702	-0.0266	-8.9618	0.1609	0.1675	0.1466	0.1928

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 4、Clayton 樣本數 600，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.4405	0.4331	-0.0075	-1.6922	0.0642	0.0679	0.0648	0.0742
	0.2	0.6	0.3988	0.3931	-0.0057	-1.4216	0.0630	0.0663	0.0630	0.0724
	0.2	0.7	0.3539	0.3494	-0.0045	-1.2800	0.0642	0.0666	0.0634	0.0735
	0.4	0.5	0.3011	0.2996	-0.0015	-0.4964	0.0456	0.0476	0.0462	0.0501
	0.4	0.6	0.2657	0.2651	-0.0006	-0.2269	0.0442	0.0462	0.0450	0.0491
	0.4	0.7	0.2305	0.2308	0.0003	0.1320	0.0436	0.0454	0.0443	0.0486
	0.6	0.5	0.2333	0.2324	-0.0008	-0.3539	0.0413	0.0437	0.0430	0.0455
	0.6	0.6	0.2024	0.2024	-0.0001	-0.0301	0.0405	0.0418	0.0415	0.0439
	0.6	0.7	0.1731	0.1742	0.0010	0.5934	0.0401	0.0419	0.0412	0.0453

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 5、Frank 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.3053	0.3041	-0.0012	-0.3846	0.0956	0.1072	0.1041	0.1184
	0.2	0.6	0.2901	0.2924	0.0023	0.7875	0.0929	0.1036	0.1007	0.1152
	0.2	0.7	0.2729	0.2742	0.0013	0.4864	0.0917	0.1044	0.1019	0.1195
	0.4	0.5	0.2894	0.2864	-0.0029	-1.0104	0.0745	0.0845	0.0828	0.0910
	0.4	0.6	0.2697	0.2663	-0.0034	-1.2498	0.0711	0.0821	0.0801	0.0915
	0.4	0.7	0.2490	0.2440	-0.0051	-2.0396	0.0723	0.0812	0.0793	0.0912
	0.6	0.5	0.2739	0.2711	-0.0028	-1.0324	0.0727	0.0831	0.0815	0.0884
	0.6	0.6	0.2501	0.2438	-0.0063	-2.5369	0.0708	0.0798	0.0785	0.0886
	0.6	0.7	0.2262	0.2198	-0.0064	-2.8129	0.0715	0.0786	0.0770	0.0874
(0.54,0.76)	0.2	0.5	0.3053	0.2972	-0.0081	-2.6610	0.1466	0.1683	0.1636	0.2019
	0.2	0.6	0.2901	0.2812	-0.0089	-3.0833	0.1441	0.1679	0.1607	0.2032
	0.2	0.7	0.2729	0.2649	-0.0079	-2.9071	0.1472	0.1745	0.1634	0.2207
	0.4	0.5	0.2894	0.2693	-0.0201	-6.9476	0.1349	0.1622	0.1549	0.1831
	0.4	0.6	0.2697	0.2523	-0.0175	-6.4704	0.1369	0.1600	0.1499	0.1856
	0.4	0.7	0.2490	0.2288	-0.0202	-8.1267	0.1386	0.1615	0.1496	0.1937
	0.6	0.5	0.2739	0.2485	-0.0254	-9.2758	0.1896	0.1983	0.1757	0.2074
	0.6	0.6	0.2501	0.2203	-0.0298	-11.9170	0.1843	0.1924	0.1655	0.2043
	0.6	0.7	0.2262	0.1934	-0.0328	-14.4849	0.1856	0.1890	0.1577	0.2000

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 6、Frank 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.3053	0.3036	-0.0018	-0.5779	0.0663	0.0746	0.0734	0.0808
	0.2	0.6	0.2901	0.2886	-0.0016	-0.5361	0.0630	0.0715	0.0701	0.0783
	0.2	0.7	0.2729	0.2719	-0.0010	-0.3638	0.0647	0.0709	0.0693	0.0779
	0.4	0.5	0.2894	0.2875	-0.0019	-0.6407	0.0563	0.0580	0.0575	0.0619
	0.4	0.6	0.2697	0.2699	0.0001	0.0554	0.0548	0.0554	0.0548	0.0597
	0.4	0.7	0.2490	0.2495	0.0004	0.1768	0.0525	0.0556	0.0552	0.0611
	0.6	0.5	0.2739	0.2746	0.0006	0.2307	0.0543	0.0566	0.0559	0.0601
	0.6	0.6	0.2501	0.2505	0.0004	0.1459	0.0505	0.0544	0.0537	0.0582
	0.6	0.7	0.2262	0.2266	0.0004	0.1891	0.0491	0.0539	0.0531	0.0594
(0.54,0.76)	0.2	0.5	0.3053	0.2970	-0.0083	-2.7134	0.1008	0.1127	0.1134	0.1347
	0.2	0.6	0.2901	0.2816	-0.0085	-2.9455	0.0997	0.1124	0.1127	0.1362
	0.2	0.7	0.2729	0.2664	-0.0064	-2.3574	0.1038	0.1169	0.1153	0.1378
	0.4	0.5	0.2894	0.2813	-0.0081	-2.7833	0.0933	0.1061	0.1065	0.1207
	0.4	0.6	0.2697	0.2589	-0.0108	-4.0158	0.0946	0.1055	0.1051	0.1270
	0.4	0.7	0.2490	0.2397	-0.0093	-3.7501	0.1001	0.1100	0.1074	0.1340
	0.6	0.5	0.2739	0.2637	-0.0102	-3.7267	0.1122	0.1291	0.1229	0.1449
	0.6	0.6	0.2501	0.2402	-0.0099	-3.9444	0.1127	0.1262	0.1188	0.1469
	0.6	0.7	0.2262	0.2115	-0.0147	-6.4781	0.1152	0.1266	0.1175	0.1501

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 7、Frank 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.4746	0.4728	-0.0017	-0.3660	0.0855	0.0976	0.0952	0.1084
	0.2	0.6	0.4607	0.4591	-0.0015	-0.3309	0.0837	0.0938	0.0915	0.1072
	0.2	0.7	0.4478	0.4471	-0.0007	-0.1625	0.0819	0.0939	0.0916	0.1085
	0.4	0.5	0.4309	0.4269	-0.0040	-0.9230	0.0649	0.0739	0.0729	0.0804
	0.4	0.6	0.4095	0.4059	-0.0036	-0.8766	0.0627	0.0720	0.0707	0.0806
	0.4	0.7	0.3910	0.3848	-0.0062	-1.5802	0.0645	0.0722	0.0706	0.0823
	0.6	0.5	0.3967	0.3917	-0.0050	-1.2508	0.0637	0.0701	0.0694	0.0763
	0.6	0.6	0.3675	0.3607	-0.0069	-1.8676	0.0611	0.0679	0.0672	0.0763
(0.54,0.76)	0.6	0.7	0.3421	0.3358	-0.0062	-1.8216	0.0641	0.0690	0.0678	0.0805
	0.2	0.5	0.4746	0.4634	-0.0112	-2.3509	0.1233	0.1477	0.1468	0.1772
	0.2	0.6	0.4607	0.4473	-0.0133	-2.8946	0.1261	0.1484	0.1452	0.1916
	0.2	0.7	0.4478	0.4360	-0.0118	-2.6388	0.1313	0.1547	0.1474	0.2039
	0.4	0.5	0.4309	0.4144	-0.0165	-3.8288	0.1161	0.1376	0.1337	0.1611
	0.4	0.6	0.4095	0.3918	-0.0177	-4.3200	0.1168	0.1379	0.1317	0.1707
	0.4	0.7	0.3910	0.3713	-0.0197	-5.0506	0.1256	0.1417	0.1322	0.1848
	0.6	0.5	0.3967	0.3748	-0.0218	-5.4989	0.1505	0.1603	0.1433	0.1834
	0.6	0.6	0.3675	0.3451	-0.0225	-6.1202	0.1497	0.1579	0.1375	0.1768
	0.6	0.7	0.3421	0.3159	-0.0262	-7.6539	0.1567	0.1580	0.1337	0.1667

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 8、Gumbel 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.2597	0.2618	0.0021	0.8138	0.0973	0.1097	0.1072	0.1210
	0.2	0.6	0.2451	0.2472	0.0021	0.8567	0.0947	0.1037	0.1005	0.1160
	0.2	0.7	0.2308	0.2336	0.0027	1.1892	0.0922	0.1019	0.0997	0.1185
	0.4	0.5	0.2514	0.2489	-0.0025	-1.0060	0.0778	0.0859	0.0851	0.0926
	0.4	0.6	0.2376	0.2346	-0.0030	-1.2485	0.0732	0.0814	0.0808	0.0902
	0.4	0.7	0.2242	0.2228	-0.0015	-0.6500	0.0713	0.0798	0.0789	0.0892
	0.6	0.5	0.2648	0.2589	-0.0059	-2.2305	0.0794	0.0865	0.0861	0.0924
	0.6	0.6	0.2495	0.2436	-0.0059	-2.3753	0.0762	0.0816	0.0810	0.0911
	0.6	0.7	0.2345	0.2289	-0.0056	-2.4040	0.0726	0.0791	0.0792	0.0910
(0.54,0.76)	0.2	0.5	0.2597	0.2492	-0.0105	-4.0455	0.1476	0.1700	0.1669	0.2049
	0.2	0.6	0.2451	0.2335	-0.0116	-4.7446	0.1444	0.1674	0.1620	0.2078
	0.2	0.7	0.2308	0.2191	-0.0117	-5.0713	0.1518	0.1672	0.1592	0.2084
	0.4	0.5	0.2514	0.2370	-0.0144	-5.7198	0.1418	0.1620	0.1557	0.1826
	0.4	0.6	0.2376	0.2187	-0.0189	-7.9444	0.1406	0.1571	0.1490	0.1885
	0.4	0.7	0.2242	0.2035	-0.0207	-9.2525	0.1391	0.1540	0.1434	0.1860
	0.6	0.5	0.2648	0.2292	-0.0356	-13.4525	0.1800	0.2007	0.1820	0.2187
	0.6	0.6	0.2495	0.2118	-0.0377	-15.1134	0.1775	0.1920	0.1693	0.1926
	0.6	0.7	0.2345	0.1918	-0.0427	-18.2293	0.1693	0.1857	0.1560	0.1833

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 9、Gumbel 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.2597	0.2594	-0.0003	-0.1295	0.0701	0.0763	0.0749	0.0827
	0.2	0.6	0.2451	0.2449	-0.0003	-0.1060	0.0652	0.0719	0.0706	0.0793
	0.2	0.7	0.2308	0.2309	0.0000	0.0151	0.0619	0.0695	0.0683	0.0763
	0.4	0.5	0.2514	0.2521	0.0007	0.2829	0.0540	0.0596	0.0594	0.0638
	0.4	0.6	0.2376	0.2357	-0.0019	-0.8192	0.0516	0.0558	0.0554	0.0602
	0.4	0.7	0.2242	0.2242	0.0000	-0.0167	0.0510	0.0544	0.0544	0.0590
	0.6	0.5	0.2648	0.2619	-0.0029	-1.1129	0.0538	0.0590	0.0590	0.0625
	0.6	0.6	0.2495	0.2459	-0.0036	-1.4453	0.0516	0.0550	0.0551	0.0588
	0.6	0.7	0.2345	0.2296	-0.0049	-2.0799	0.0486	0.0538	0.0543	0.0605
(0.54,0.76)	0.2	0.5	0.2597	0.2515	-0.0082	-3.1445	0.1061	0.1159	0.1166	0.1342
	0.2	0.6	0.2451	0.2381	-0.0070	-2.8720	0.1006	0.1131	0.1139	0.1334
	0.2	0.7	0.2308	0.2180	-0.0128	-5.5392	0.1034	0.1141	0.1143	0.1411
	0.4	0.5	0.2514	0.2345	-0.0169	-6.7262	0.0983	0.1080	0.1097	0.1252
	0.4	0.6	0.2376	0.2234	-0.0143	-6.0000	0.0977	0.1060	0.1068	0.1212
	0.4	0.7	0.2242	0.2059	-0.0183	-8.1643	0.0972	0.1082	0.1068	0.1351
	0.6	0.5	0.2648	0.2437	-0.0211	-7.9775	0.1129	0.1333	0.1301	0.1482
	0.6	0.6	0.2495	0.2258	-0.0237	-9.5059	0.1123	0.1279	0.1228	0.1479
	0.6	0.7	0.2345	0.2064	-0.0281	-11.9990	0.1158	0.1254	0.1171	0.1507

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 10、Gumbel 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.4400	0.4398	-0.0002	-0.0435	0.0934	0.1065	0.1046	0.1231
	0.2	0.6	0.4198	0.4173	-0.0025	-0.5966	0.0908	0.1007	0.0982	0.1128
	0.2	0.7	0.4007	0.3998	-0.0009	-0.2178	0.0928	0.0987	0.0963	0.1151
	0.4	0.5	0.3962	0.3938	-0.0024	-0.6100	0.0697	0.0772	0.0775	0.0848
	0.4	0.6	0.3747	0.3727	-0.0021	-0.5484	0.0668	0.0740	0.0740	0.0824
	0.4	0.7	0.3557	0.3523	-0.0034	-0.9505	0.0658	0.0732	0.0733	0.0847
	0.6	0.5	0.3900	0.3874	-0.0025	-0.6440	0.0650	0.0725	0.0729	0.0803
	0.6	0.6	0.3629	0.3590	-0.0039	-1.0842	0.0637	0.0686	0.0697	0.0775
(0.54,0.76)	0.6	0.7	0.3394	0.3354	-0.0040	-1.1889	0.0634	0.0676	0.0687	0.0798
	0.2	0.5	0.4400	0.4290	-0.0110	-2.5070	0.1403	0.1570	0.1566	0.1928
	0.2	0.6	0.4198	0.4065	-0.0134	-3.1812	0.1380	0.1561	0.1530	0.1993
	0.2	0.7	0.4007	0.3865	-0.0142	-3.5398	0.1448	0.1583	0.1507	0.2139
	0.4	0.5	0.3962	0.3795	-0.0168	-4.2290	0.1248	0.1399	0.1377	0.1667
	0.4	0.6	0.3747	0.3561	-0.0186	-4.9598	0.1227	0.1381	0.1333	0.1713
	0.4	0.7	0.3557	0.3342	-0.0214	-6.0201	0.1284	0.1370	0.1291	0.1697
	0.6	0.5	0.3900	0.3668	-0.0232	-5.9421	0.1410	0.1623	0.1469	0.1756
	0.6	0.6	0.3629	0.3327	-0.0302	-8.3253	0.1383	0.1557	0.1374	0.1706
	0.6	0.7	0.3394	0.3032	-0.0362	-10.6676	0.1383	0.1518	0.1280	0.1545

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 11、Lognormal 樣本數 200，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.18,0.4)	0.2	0.5	0.3934	0.3796	-0.0138	-3.5134	0.1215	0.1358	0.1338	0.1519
	0.2	0.6	0.3988	0.3848	-0.0140	-3.5191	0.1291	0.1423	0.1406	0.1652
	0.2	0.7	0.4051	0.3907	-0.0144	-3.5655	0.1408	0.1540	0.1537	0.1806
	0.4	0.5	0.3615	0.3476	-0.0139	-3.8507	0.1105	0.1174	0.1187	0.1331
	0.4	0.6	0.3633	0.3516	-0.0117	-3.2093	0.1177	0.1235	0.1256	0.1408
	0.4	0.7	0.3658	0.3515	-0.0143	-3.8979	0.1276	0.1337	0.1375	0.1548
	0.6	0.5	0.3615	0.3496	-0.0119	-3.2982	0.1191	0.1288	0.1313	0.1474
	0.6	0.6	0.3602	0.3495	-0.0107	-2.9715	0.1283	0.1351	0.1392	0.1564
(0.38,0.65)	0.6	0.7	0.3594	0.3457	-0.0137	-3.8072	0.1326	0.1488	0.1538	0.1848
	0.2	0.5	0.3934	0.3736	-0.0199	-5.0512	0.1521	0.1826	0.1819	0.2173
	0.2	0.6	0.3988	0.3749	-0.0240	-6.0090	0.1648	0.1997	0.1988	0.2503
	0.2	0.7	0.4051	0.3829	-0.0223	-5.4983	0.1857	0.2248	0.2206	0.2820
	0.4	0.5	0.3615	0.3379	-0.0236	-6.5257	0.1502	0.1677	0.1718	0.2035
	0.4	0.6	0.3633	0.3373	-0.0260	-7.1537	0.1619	0.1836	0.1891	0.2325
	0.4	0.7	0.3658	0.3374	-0.0284	-7.7690	0.1846	0.2065	0.2055	0.2689
	0.6	0.5	0.3615	0.3390	-0.0225	-6.2333	0.1790	0.1993	0.1965	0.2331
	0.6	0.6	0.3602	0.3318	-0.0284	-7.8829	0.1927	0.2122	0.2076	0.2610
	0.6	0.7	0.3594	0.3237	-0.0357	-9.9322	0.2221	0.2302	0.2181	0.2981

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為偏誤再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 12、Lognormal 樣本數 400，相關性 Kendall's $\tau = 0.3$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.18,0.4)	0.2	0.5	0.3934	0.3938	0.0004	0.0959	0.0859	0.0932	0.0928	0.1048
	0.2	0.6	0.3988	0.3971	-0.0017	-0.4199	0.0886	0.0972	0.0969	0.1091
	0.2	0.7	0.4051	0.4018	-0.0033	-0.8207	0.0959	0.1053	0.1062	0.1193
	0.4	0.5	0.3615	0.3552	-0.0063	-1.7344	0.0748	0.0800	0.0817	0.0885
	0.4	0.6	0.3633	0.3550	-0.0083	-2.2817	0.0788	0.0844	0.0867	0.0955
	0.4	0.7	0.3658	0.3582	-0.0076	-2.0799	0.0859	0.0925	0.0957	0.1070
	0.6	0.5	0.3615	0.3551	-0.0064	-1.7823	0.0793	0.0865	0.0893	0.0970
	0.6	0.6	0.3602	0.3511	-0.0091	-2.5328	0.0871	0.0923	0.0962	0.1052
	0.6	0.7	0.3594	0.3470	-0.0124	-3.4510	0.0922	0.1016	0.1085	0.1222
(0.38,0.65)	0.2	0.5	0.3934	0.3893	-0.0041	-1.0403	0.1108	0.1217	0.1233	0.1422
	0.2	0.6	0.3988	0.3900	-0.0088	-2.2052	0.1190	0.1313	0.1344	0.1570
	0.2	0.7	0.4051	0.3933	-0.0118	-2.9114	0.1326	0.1486	0.1522	0.1757
	0.4	0.5	0.3615	0.3514	-0.0101	-2.7912	0.1070	0.1116	0.1156	0.1275
	0.4	0.6	0.3633	0.3486	-0.0147	-4.0571	0.1128	0.1214	0.1286	0.1445
	0.4	0.7	0.3658	0.3474	-0.0184	-5.0406	0.1259	0.1385	0.1470	0.1720
	0.6	0.5	0.3615	0.3458	-0.0157	-4.3467	0.1180	0.1298	0.1354	0.1505
	0.6	0.6	0.3602	0.3407	-0.0195	-5.4203	0.1244	0.1434	0.1501	0.1816
	0.6	0.7	0.3594	0.3362	-0.0231	-6.4398	0.1397	0.1627	0.1669	0.2078

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 13、Lognormal 樣本數 200，相關性 Kendall's $\tau = 0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.18,0.4)	0.2	0.5	0.6008	0.5966	-0.0042	-0.6959	0.1052	0.1144	0.1141	0.1273
	0.2	0.6	0.6165	0.6073	-0.0092	-1.4846	0.1096	0.1186	0.1195	0.1361
	0.2	0.7	0.6358	0.6213	-0.0146	-2.2893	0.1156	0.1289	0.1311	0.1514
	0.4	0.5	0.5450	0.5324	-0.0126	-2.3149	0.0922	0.1024	0.1065	0.1143
	0.4	0.6	0.5462	0.5326	-0.0136	-2.4968	0.0971	0.1086	0.1158	0.1256
	0.4	0.7	0.5504	0.5318	-0.0185	-3.3684	0.1035	0.1199	0.1295	0.1444
	0.6	0.5	0.5500	0.5362	-0.0138	-2.5042	0.0995	0.1109	0.1161	0.1278
	0.6	0.6	0.5469	0.5321	-0.0147	-2.6939	0.1038	0.1183	0.1255	0.1346
	0.6	0.7	0.5463	0.5235	-0.0228	-4.1814	0.1136	0.1312	0.1404	0.1670
(0.38,0.65)	0.2	0.5	0.6008	0.5831	-0.0176	-2.9327	0.1270	0.1498	0.1533	0.1760
	0.2	0.6	0.6165	0.5931	-0.0234	-3.7935	0.1398	0.1608	0.1659	0.1948
	0.2	0.7	0.6358	0.6057	-0.0301	-4.7348	0.1522	0.1812	0.1860	0.2327
	0.4	0.5	0.5450	0.5173	-0.0277	-5.0830	0.1289	0.1518	0.1569	0.1771
	0.4	0.6	0.5462	0.5166	-0.0296	-5.4150	0.1451	0.1638	0.1684	0.1979
	0.4	0.7	0.5504	0.5126	-0.0377	-6.8551	0.1554	0.1809	0.1812	0.2348
	0.6	0.5	0.5500	0.5234	-0.0266	-4.8364	0.1463	0.1706	0.1719	0.1960
	0.6	0.6	0.5469	0.5155	-0.0314	-5.7377	0.1546	0.1830	0.1811	0.2248
	0.6	0.7	0.5463	0.5107	-0.0356	-6.5218	0.1717	0.1961	0.1895	0.2479

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 14、Piecewise constant Clayton 樣本數 200，相關性 $\theta_1 = 0.5$ ， $\theta_2 = 0.2$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.6246	0.6177	-0.0068	-1.0910	0.0936	0.1026	0.0937	0.1148
	0.2	0.6	0.6084	0.5995	-0.0090	-1.4717	0.0939	0.1031	0.0933	0.1210
	0.2	0.7	0.5908	0.5802	-0.0106	-1.7981	0.0940	0.1086	0.0982	0.1321
	0.5	0.5	0.4040	0.4071	0.0031	0.7604	0.0661	0.0744	0.0725	0.0813
	0.5	0.6	0.3728	0.3777	0.0050	1.3384	0.0633	0.0745	0.0726	0.0827
	0.5	0.7	0.3414	0.3458	0.0044	1.2859	0.0672	0.0771	0.0746	0.0886
	0.6	0.5	0.3821	0.3770	-0.0050	-1.3161	0.0631	0.0740	0.0718	0.0802
	0.6	0.6	0.3484	0.3448	-0.0036	-1.0471	0.0672	0.0745	0.0719	0.0828
	0.6	0.7	0.3148	0.3122	-0.0026	-0.8107	0.0708	0.0769	0.0740	0.0850
(0.54,0.76)	0.2	0.5	0.6246	0.6099	-0.0146	-2.3409	0.1276	0.1493	0.1436	0.1848
	0.2	0.6	0.6084	0.5892	-0.0192	-3.1584	0.1326	0.1574	0.1493	0.2094
	0.2	0.7	0.5908	0.5696	-0.0212	-3.5841	0.1406	0.1721	0.1600	0.2417
	0.5	0.5	0.4040	0.4011	-0.0030	-0.7339	0.1291	0.1470	0.1383	0.1772
	0.5	0.6	0.3728	0.3691	-0.0036	-0.9762	0.1326	0.1522	0.1406	0.1917
	0.5	0.7	0.3414	0.3359	-0.0055	-1.6078	0.1390	0.1577	0.1439	0.2055
	0.6	0.5	0.3821	0.3692	-0.0129	-3.3762	0.1391	0.1591	0.1450	0.1894
	0.6	0.6	0.3484	0.3340	-0.0144	-4.1394	0.1480	0.1620	0.1446	0.1913
	0.6	0.7	0.3148	0.2953	-0.0194	-6.1687	0.1538	0.1653	0.1446	0.2025

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 15、Piecewise constant Clayton 樣本數 200，相關性 $\theta_1 = 0.2$ ， $\theta_2 = 0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.2913	0.2920	0.0006	0.2168	0.1144	0.1320	0.1274	0.1495
	0.2	0.6	0.2530	0.2500	-0.0029	-1.1657	0.1078	0.1269	0.1223	0.1437
	0.2	0.7	0.2154	0.2115	-0.0040	-1.8445	0.1071	0.1241	0.1193	0.1421
	0.5	0.5	0.1758	0.2095	0.0337	19.1711	0.0878	0.0949	0.0888	0.1000
	0.5	0.6	0.1504	0.1758	0.0254	16.8863	0.0814	0.0902	0.0855	0.0965
	0.5	0.7	0.1268	0.1458	0.0189	14.9407	0.0783	0.0876	0.0843	0.0971
	0.6	0.5	0.3037	0.2971	-0.0067	-2.1959	0.0795	0.0865	0.0817	0.0917
	0.6	0.6	0.2602	0.2548	-0.0054	-2.0846	0.0792	0.0856	0.0816	0.0943
	0.6	0.7	0.2173	0.2144	-0.0029	-1.3400	0.0800	0.0855	0.0826	0.0979
(0.54,0.76)	0.2	0.5	0.2913	0.2781	-0.0133	-4.5571	0.1669	0.1963	0.1901	0.2384
	0.2	0.6	0.2530	0.2415	-0.0115	-4.5276	0.1677	0.1972	0.1892	0.2502
	0.2	0.7	0.2154	0.2069	-0.0085	-3.9427	0.1667	0.2005	0.1876	0.2733
	0.5	0.5	0.1758	0.1976	0.0219	12.4316	0.1603	0.1777	0.1683	0.2120
	0.5	0.6	0.1504	0.1661	0.0157	10.4069	0.1600	0.1774	0.1652	0.2198
	0.5	0.7	0.1268	0.1380	0.0112	8.8231	0.1685	0.1761	0.1627	0.2183
	0.6	0.5	0.3037	0.2795	-0.0243	-7.9942	0.1606	0.1741	0.1598	0.2131
	0.6	0.6	0.2602	0.2369	-0.0232	-8.9315	0.1688	0.1743	0.1581	0.2158
	0.6	0.7	0.2173	0.1965	-0.0207	-9.5440	0.1722	0.1736	0.1560	0.2064

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。

表 16、Piecewise constant Clayton 樣本數 200，相關性 $\theta_1 = 0.2$ ， $\theta_2 = -0.5$ ， β_1 模擬結果

設限率	τ_1	τ_2	真值	估計值	偏誤	相對偏誤 (%)	標準差	標準差 (Bootstrapping)	標準差 (Resampling)	標準差 (Asymptotic)
(0.24,0.4)	0.2	0.5	0.2913	0.2919	0.0006	0.1974	0.1176	0.1325	0.1272	0.1493
	0.2	0.6	0.2530	0.2509	-0.0021	-0.8229	0.1089	0.1272	0.1219	0.1427
	0.2	0.7	0.2154	0.2126	-0.0028	-1.3169	0.1070	0.1240	0.1186	0.1411
	0.5	0.5	0.1758	0.1478	-0.0280	-15.9189	0.0952	0.1037	0.0897	0.1073
	0.5	0.6	0.1504	0.1161	-0.0343	-22.7972	0.0944	0.1014	0.0851	0.1041
	0.5	0.7	0.1268	0.0860	-0.0408	-32.1516	0.0964	0.1000	0.0815	0.1035
	0.6	0.5	-0.0032	-0.0086	-0.0054	167.8717	0.1059	0.1072	0.0895	0.1114
	0.6	0.6	-0.0545	-0.0445	0.0101	-18.4940	0.0976	0.0998	0.0825	0.1029
	0.6	0.7	-0.0808	-0.0662	0.0145	-18.0089	0.0910	0.0963	0.0806	0.1038
(0.54,0.76)	0.2	0.5	0.2913	0.2895	-0.0018	-0.6168	0.1729	0.1997	0.1891	0.2378
	0.2	0.6	0.2530	0.2552	0.0022	0.8859	0.1676	0.1957	0.1844	0.2396
	0.2	0.7	0.2154	0.2162	0.0007	0.3307	0.1642	0.1961	0.1822	0.2612
	0.5	0.5	0.1758	0.1387	-0.0370	-21.0675	0.1622	0.1965	0.1757	0.2147
	0.5	0.6	0.1504	0.1071	-0.0433	-28.7721	0.1603	0.1900	0.1668	0.2060
	0.5	0.7	0.1268	0.0780	-0.0489	-38.5241	0.1589	0.1865	0.1587	0.2203
	0.6	0.5	-0.0032	-0.0096	-0.0064	200.8116	0.1810	0.2138	0.1905	0.2296
	0.6	0.6	-0.0545	-0.0428	0.0118	-21.5524	0.1750	0.2049	0.1781	0.2273
	0.6	0.7	-0.0808	-0.0630	0.0177	-21.9650	0.1740	0.1984	0.1670	0.2337

註 1：估計值為 1000 次模擬平均值。 註 2：相對偏誤為估計值減去真值再除以真值，以百分比呈現。註 3：標準差為 1000 次模擬估計值的標準差。

註 4：Bootstrapping、resampling、Asymptotic 為 1000 次模擬標準差估計平均值。