

國立中央大學

統計研究所

碩士論文

時間相依 ROC 曲線分析：修正與比較

Time-dependent ROC Curve

Analysis : Revise and Comparison

研究生：侯昱宇

指導教授：曾議寬 博士

中華民國一百一十一年六月

摘要

醫學研究中，常以受試者特徵曲線 (Receiver Operating Characteristic curve, 簡稱 ROC 曲線) 評估生物指標對疾病預測能力。其中以 Heagerty.& Zheng.(2005) 提出的時間相依 ROC 曲線最為常見，透過分析並提高預測效率，辨別各個生物指標的預測能力，進而找出適合的模型。過去文獻使用比例風險模型 (Proportional hazard model, 簡稱 PH 或 Cox 模型)，或使用 AFT 模型 (Accelerated failure time model)、PO 模型 (Proportional odds model)，建構時間相依 ROC 曲線，並獲得曲線下面積 (the area under the ROC curves, 簡稱 AUC)，透過對各個時間下的 AUC 加權平均，取得一致性指標 (Concordance)，並可將其推導為風險迴歸的函數，當使用不同模型時，只需代入對應的風險函數，此一致性指標在先前已被證明對預測精準度有一致性。本研究針對 AUC 的計算方式提出修正，透過原先的定義，將面積計算方式作調整，進而獲得更貼近真實值的結果。後續再使用過去文獻提出的各個模型，將其應用在同一筆模擬資料中，比較其中的差異與優劣，觀察一致性指標的準確性，最後以實際愛滋病資料分析，展示不同模型下的結果。

關鍵字: 比例風險模型、加速失敗模型、比例勝算模型、一致性指標、ROC 曲線計算修正、模型比較

目錄

摘要	i
目錄	ii
圖目錄	iv
表目錄	v
第一章 緒論	1
1.1 ROC 曲線	2
1.2 時間相依 ROC 曲線	5
1.3 累積敏感度與動態特異度	9
1.4 事件敏感度與靜態特異度	12
1.5 事件敏感度與動態特異度	13
第二章 統計模型	15
2.1 Cox 比例風險模型	15
2.2 AFT 模型	17
2.3 PO 模型	19
第三章 模擬研究	22
3.1 修正計算方式	23
3.2 Cox 模型模擬	26
3.3 AFT 模型模擬	31

3.4 PO 模型模擬	36
3.5 修正計算方式結論	40
3.6 Bivariate normal 模擬	42
第四章 資料分析	45
4.1 愛滋病資料	45
4.2 時間獨立共變數存活模型	46
第五章 結論	48
參考文獻	49
附錄 A.1 時間相依 AUC 與一致性指標推導	51

圖目錄

圖 3.1：原始計算方式重複點問題	24
圖 3.2：試驗後期 TP 與 FP 分佈	24
圖 3.3：試驗後期極端現象	25
圖 4.1：模型 III 在 AFT 模型下不同帶寬 AUC	47
圖 4.2：模型 III 下不同風險模型各時間點 AUC	47

表目錄

表 2.1: 各種模型、方法下 $\widehat{TP}_t(c)$ 形式	21
表 3.1: Lognormal-Cox 模擬結果	28
表 3.2: Weibull-Cox 模擬結果	30
表 3.3: Lognormal-AFT 模擬結果	33
表 3.4: Weibull-AFT 模擬結果	35
表 3.5: Lognormal-PO 模擬結果	37
表 3.6: Loglogistic-PO 模擬結果	39
表 3.7: Bivariate normal 模擬結果 (n=200)	42
表 3.8: Bivariate normal 模擬結果 (n=500)	43
表 4.1: 不同模型下一致性指標比較	46

第一章 緒論

醫學研究中，我們常會使用受試者特徵曲線 (Receiver Operating Characteristic curve，簡稱 ROC 曲線) 來衡量某一生物指標對疾病的區分能力，並使用 ROC 曲線下面積 (Area under the Receiver Operating Characteristic curve，簡稱 AUC)，比較不同生物指標的區分能力，其意義代表生物指標整體的預測能力。當資料與時間有關，可以使用 Heagerty & Zheng (2005) 提出的時間相依敏感度與特異度，並建構出時間相依的 AUC 與整體一致性指標 (Concordance)。實務上常使用比例風險模型 (Proportional hazard model, 簡稱 PH, 也稱為 Cox 模型) 描述共變量與存活時間的關係，而當資料不滿足比例風險時，可以考慮使用加速失敗模型 (Accelerated failure time, 簡稱 AFT)，或使用比例勝算模型 (Proportional Odds model，簡稱 PO) 取代。

本章將介紹基本的 ROC 曲線，接著推廣至存活時間的資料，透過三種不同的時間相依敏感度與特異度定義，建構時間相依 ROC 曲線與一致性指標，經由不同的定義方式、不同的模型，或是不同的參數計算，所得到的結果都會不同。第二章將會介紹本篇使用的模型，包含模型假設、參數估計及帶寬選擇。第三章首先介紹本篇針對 AUC 計算提出的修正方式，透過細微的修正，改善並優化在實驗後期的估計結果，接著透過模擬的方式，展示修正前後估計結果的差異，並比較各個模型與方法之間的差異。第四章以實際的愛滋病資料，分析並展示模型應用的結果。第五章為最後的結論。

1.1 ROC 曲線

受試者特徵曲線 (Receiver Operating Characteristic curve, 簡稱 ROC 曲線), 醫學上, 常用於評估生物指標對疾病的區分能力, 並利用 ROC 曲線下面積 (Area under the Receiver Operating Characteristic curve, 簡稱 AUC), 比較不同生物指標的區分能力, 其意義代表生物指標整體的預測能力。

若生物指標為連續型 (例如 : 血壓), 令 X 為連續型的生物指標, 定義門檻值 (threshold) c , 此處假設生物指標越高越有可能得病, 且 D 表示受試者真實情況, $D = 0$ 表示未得病, $D = 1$ 表示得病, 則我們可以定義敏感度與特異度:

$$Sensitivity = P(X \geq c \mid D = 1)$$

$$Specificity = P(X < c \mid D = 0)$$

在 ROC 曲線的定義中, 我們會使用

$$TPR(c) = P(X \geq c \mid D = 1) = Sensitivity$$

$$FPR(c) = P(X \geq c \mid D = 0) = 1 - Specificity$$

敏感度所代表的意義為, 已知其有病, 而生物指標也判定為陽性的機率, 稱其為真陽性率 (True Positive Rate, 簡稱 TPR), 而 (1 - 特異度) 代表的意義為, 已知其沒病, 而生物指標判定為陽性的機率, 稱為假

陽性率 (False Positive Rate，簡稱 FPR)，則我們可以定義 ROC 曲線為

$$ROC = \{(FPR(c), TPR(c)), c \in (-\infty, \infty)\}$$

其中 c 值為事先給定，每給定一個 c 值，即可獲得一組 $(FPR(c), TPR(c))$ ，將每組結果畫在直角坐標平面上，即成為 ROC 曲線，若令 $p = FPR(c)$ ， $ROC(p) = TPR(c)$ ，則 ROC 曲線亦可表示為

$$ROC = \{(p, ROC(p)), p \in (0, 1)\}$$

若 ROC 曲線為 45 度斜直線，則表示此生物指標的預測能力為隨機猜測，即不具有預測能力，故我們會希望 ROC 曲線越高越好，即表示此生物指標有較好的預測能力。

ROC 曲線的估計主要分兩種方法，首先介紹非參數化方法，使用經驗分配函數 (the empirical distribution function) 估計 TPR 與 FPR，再獲得 ROC 曲線，估計如下：

$$TPR(c) = \frac{1}{n_D} \sum_{i=1}^{n_D} I\{X_{D_i} \geq c\}$$

$$FPR(c) = \frac{1}{n_{\bar{D}}} \sum_{i=1}^{n_{\bar{D}}} I\{X_{\bar{D}_i} \geq c\}$$

其中 D 及 \bar{D} 分別表示有病與沒病，則 ROC 曲線可表示為：

$$ROC(c) = \hat{S}_D(\hat{S}_{\bar{D}}^{-1}(p))$$

接著介紹參數化方法，此處假設兩群體的生物指標為常態分配，即 $X_D \sim N(\mu_D, \sigma_D^2)$ ， $X_{\bar{D}} \sim N(\mu_{\bar{D}}, \sigma_{\bar{D}}^2)$ ，經推導可得：

$$ROC(p) = 1 - \Phi\left(\frac{\mu_{\bar{D}} - \mu_D - \Phi^{-1}(p)\sigma_{\bar{D}}}{\sigma_D}\right) = \Phi(b\Phi^{-1}(p) + a)$$

$$ROC = \{(p, ROC(p)), p \in (0, 1)\} = \{(p, \Phi(b\Phi^{-1}(p) + a)), p \in (0, 1)\}$$

其中 $a = \frac{\mu_{\bar{D}} - \mu_D}{\sigma_D}$ ， $b = \frac{\sigma_{\bar{D}}}{\sigma_D}$ ， $\Phi(\cdot)$ 為標準常態分佈的累積機率密度函數，未知參數利用最大概似估計法估計，帶入上式，即可得到 ROC 曲線。

為了能夠有效判別 ROC 曲線的預測精準度，通常會使用曲線下面積 (Area under curve，簡稱 AUC) 來衡量一 ROC 曲線的優劣，定義即推導如下：

$$\begin{aligned} AUC &= \int_0^1 ROC(p) dp = \int_0^1 \hat{S}_D(\hat{S}_{\bar{D}}^{-1}(p)) dp \\ &= \int_{-\infty}^{\infty} S_D(y) dS_{\bar{D}}(y) \\ &= \int_{-\infty}^{\infty} P(Y_D > y) f_{\bar{D}}(y) dy \\ &= \int_{-\infty}^{\infty} P(Y_D > y, Y_{\bar{D}} = y) dy \\ &= P(Y_D > Y_{\bar{D}}) \end{aligned}$$

最後所得的結果為真正有病的人的生物指標，大於真正沒病的人的生物指標的機率，其機率應介於 0.5 至 1 之間，當 ROC 曲線為一條通過原點的 45 度線時，AUC 為 0.5，代表真正沒病的人的生物指標大於真正有病的人的生物指標的機率為 0.5，表示此生物指標沒有預測價

值；當 AUC 為 1 時，代表此生物指標可以完全預測出是否患有疾病，為最佳情況，故 AUC 越接近 1，表示該生物指標的預測能力越好。

1.2 時間相依 ROC 曲線

由 Heagerty & Zheng (2005) 提出，作者將定義推廣至存活資料上，給定的條件由有病沒病，轉變成存活與否，首先定義時間相依敏感度與特異度，得到時間相依 ROC 曲線，進一步推導出時間相依 AUC，符號定義如下：

T_i ：第 i 位受試者存活時間

X_i ：第 i 位受試者共變量資料

$M_i = X_i^T \beta$ ：第 i 位受試者模型分數

C_i ：第 i 位受試者設限時間，與 T_i 獨立

$V_i = \min(T_i, C_i)$ ：第 i 位受試者觀測時間

$\Delta_i = I\{T_i \leq C_i\}$ ：設限指標

$R_i(t) = I\{V_i \geq t\}$ ：風險指標

$N_i^*(t) = I\{T_i \leq t\}$ ：計數函數

$dN_i^*(t) = N_i^*(t) - N_i^*(t-)$ ：增量函數

作者於該篇論文中使用 Cox 回歸模型，並以半參數方法估計回歸係數 β ，以下討論資料為時間獨立共變數的情況。

1.2.1 時間獨立共變數

當資料為時間獨立共變數時，令 $M_i = X_i^T \beta$ 為原始共變數的線性組合，將其視為一新的生物指標，作者提出的時間相依敏感度 (incident sensitivity，又稱事件敏感度) 與時間相依特異度 (dynamic specificity，又稱動態特異度)，定義如下：

$$\text{sensitivity}^{\mathbb{I}}(c, t) : P(M_i > c \mid T_i = t) = P(M_i > c \mid dN_i^*(t) = 1) = TP_t^{\mathbb{I}}(c)$$

$$\text{specificity}^{\mathbb{D}}(c, t) : P(M_i \leq c \mid T_i > t) = 1 - P(M_i > c \mid N_i^*(t) = 0) = 1 - FP_t^{\mathbb{D}}(c)$$

在此定義下，ROC 曲線可表示為：

$$ROC_t^{\mathbb{I}/\mathbb{D}}(p) = TP_t^{\mathbb{I}}\{[FP_t^{\mathbb{D}}]^{-1}(p)\}, p \in (0, 1)$$

以及任意時間 t 下的 AUC 可表示為：

$$AUC(t) = P(M_j > M_k \mid T_j = t, T_k > t) \quad (1.1)$$

其中各個受試者 j, k 彼此之間的生物指標與存活時間獨立，即 (M_j, T_j) 與 (M_k, T_k) 獨立。

其中 M 在事件發生時間 t 下的條件機率，可透過貝氏定理推導得到：

$$\begin{aligned}
 f_M(m | T = t) &= \frac{f(t | m)g(m)}{\int f(t | m)g(m)dm} \\
 &= \frac{\lambda(t | m)S(t | m)g(m)}{\int \lambda(t | m)S(t | m)g(m)dm} \\
 &= \frac{\lambda(t | m)f(m | T \geq t)S(t)}{\int \lambda(t | m)f(m | T \geq t)S(t)dm} \\
 &= \frac{\lambda(t | m)f(m | T \geq t)}{\int \lambda(t | m)f(m | T \geq t)dm} \tag{1.2}
 \end{aligned}$$

$g(m)$ 為 M 的機率密度函數， $S(t | M)$ 及 $f(t | m)$ 為 T 給定 M 的條件存活函數與條件機率密度函數。透過 (1.2) 式，可以根據不同模型的定義，替換其對應的風險函數估計式 $\hat{\lambda}(t | M_i)$ ，故 (1.2) 式可視為給定一事件發生時間 t 之下 M 的條件機率密度函數通式。(1.2) 式中給定事件發生時間 $T \geq t$ 下 M 的條件分配估計可透過一致估計得到：

$$\hat{P}(M \leq m | T \geq t) = \frac{\sum_{M_i \leq m} R_i(t)}{\sum_{j=1}^n R_j(t)}$$

時間相依敏感度可被估計為：

$$\begin{aligned}
 \widehat{TP}_t(c) &= \hat{P}(M > c | T = t) \\
 &= \frac{\sum_{M_i > c} \hat{\lambda}(t | M_i) R_i(t) / \sum_{i=1}^n R_i(t)}{\sum_{j=1}^n \hat{\lambda}(t | M_j) R_j(t) / \sum_{j=1}^n R_j(t)} \\
 &= \sum_{M_i > c} \frac{\hat{\lambda}(t | M_i) R_i(t)}{\sum_{j=1}^n \hat{\lambda}(t | M_j) R_j(t)}
 \end{aligned}$$

$\widehat{FP}_t(c)$ 以經驗估計法估計：

$$\widehat{FP}_t(c) = \widehat{P}(M > c \mid T > t) = \sum_{M_i > c}^n \frac{R_i(t+)}{\sum_{j=1}^n R_j(t+)}$$

其中 $R_i(t+) = \lim_{|\delta| \rightarrow 0} R_i(t + |\delta|)$ ，則 ROC 曲線的估計可表示成：

$$\widehat{ROC}_t(p) = \widehat{TP}_t\{\widehat{FP}_t^{-1}(p)\}$$

其中 $\widehat{FP}_t^{-1}(p) = \inf_c \{\widehat{FP}_t(c) \leq p\}$ 。根據 (1.1) 式，在事件發生時間 $T = t$ 下的 AUC 可以透過數值積分求出：

$$\widehat{AUC}(t) = \int \widehat{ROC}_t(p) dp$$

1.2.2 一致性指標

因實驗後期受試者數量會逐漸減少，甚至遠少於對照組的人數，若直接將所有時間下的 AUC 取平均，作為評斷區分能力好壞的標準，會造成一定的偏差，所以採用加權平均的方法較為合適。

此處採用 Heagerty & Zheng (2005) 提出的一種一致性指標 (the concordance summary)，令其為 C ，透過加權的方式，衡量整體時間下的區分能力，算法如下：

$$\begin{aligned} C &= P(M_j > M_k \mid T_j < T_k) \\ &= E_T[AUC(T) \times 2 \times S(T)] \end{aligned}$$

若估計範圍為一時間段，例如 $(0, \tau)$ ，則一致性指標將改寫為：

$$\begin{aligned} C^\tau &= P(M_j > M_k \mid T_j < T_k, T_j < \tau) \\ &= \int_0^\tau AUC(t) \times w^\tau(t) dt = E_{\{T < \tau\}} \left[\frac{AUC(T) \times 2 \times S(T)}{W^\tau} \right] \quad (1.3) \end{aligned}$$

其中 $w^\tau(t) = \frac{2f(t)S(t)}{W^\tau}$ ， $W^\tau = \int_0^\tau 2f(t)S(t)dt = 1 - S^2(\tau)$ 。根據 (1.3) 式，一致性指標的估計可以寫成：

$$\hat{C}^\tau = \int_0^\tau \widehat{AUC}(t) \hat{w}^\tau(t) dt \quad \hat{w}^\tau(t) = \frac{2f(t)S(t)}{1 - S^2(\tau)}$$

其中 $S(t)$ 以 Kaplan-Meier 方法估計得到。以上各個估計式中，除了 $\widehat{TP}_t(c)$ 需根據不同的模型定義做替換，其餘 $\widehat{FP}_t(c)$ 、 $\widehat{ROC}_t(p)$ 、 $\widehat{AUC}(t)$ 與一致性指標估計方式皆相同。

事實上，除了上述方法，時間相依敏感度與特異度還有其他定義方式，差異主要在於病例組與對照組的定義集合不同，以下分別介紹。

1.3 累積敏感度與動態特異度

累積敏感度與動態特異度 (Cumulative sensitivity and dynamic specificity，簡稱 C/D)，在每個時間點 t ，所有個體都被分類為病例組或對照組，病例組定義為在 $(0, t]$ 間發生事件的個體，對照組為直到時間 t 都沒發生事件的個體，此種定義的特點為病例組與對照組的人數，

會隨著時間不斷在變動，且個體在時間 t 較小時為對照組 (當 $T_i > t$)，而當時間 t 較大時，也可能轉變為病例組 (當 $T_i \leq t$)，故稱為累積敏感度與動態特異度，定義如下：

$$\text{sensitivity}^{\mathbb{C}}(c, t) : P(M_i > c \mid T_i \leq t) = P(M_i > c \mid N_i^*(t) = 1) = TP_t^{\mathbb{C}}(c)$$

$$\text{specificity}^{\mathbb{D}}(c, t) : P(M_i \leq c \mid T_i > t) = 1 - P(M_i > c \mid N_i^*(t) = 0) = 1 - FP_t^{\mathbb{D}}(c)$$

在此定義下，AUC 可表示為

$$AUC^{\mathbb{C}/\mathbb{D}}(t) = P(M_j > M_k \mid T_j \leq t, T_k > t)$$

此定義方式較適合使用於，當對一個特定的時間點感興趣，用於區分已發生事件的個體與在特定時間點之前沒有發生事件的個體，此種定義常用於臨床相關的應用，缺點為需要花更多時間在區分病例組及對照組，以及無法推導出一致性指標。在 C/D 定義下，利用不同的估計量，可以得到不同的時間相依敏感度與特異度的估計，以下分別介紹。

1.3.1 Kaplan-Meier 估計量

Heagerty 利用生存函數的 Kaplan-Meier 估計量，來估計時間相依敏感度與特異度，因其定義在 C/D 之下，故又被稱為 CD1，利用貝氏

定理，可得敏感度與特異度的估計：

$$\widehat{Se}(c, t) = \frac{\{1 - \hat{S}(t \mid M_i > c)\}(1 - \hat{F}_M(c))}{1 - \hat{S}(t)}$$

$$\widehat{Sp}(c, t) = \frac{\hat{S}(t \mid M_i \leq c)\hat{F}_M(c)}{\hat{S}(t)}$$

其中 $\hat{S}(t)$ 為存活函數的估計， $\hat{S}(t \mid M_i > c)$ 為條件存活函數的估計， $\hat{F}_M(c)$ 為生物指標 M 的累積經驗分配函數。此種方法的優點為計算簡單；缺點為，此估計的方式，會使得敏感度與特異度的估計不單調 (non-monotone)，且對於生物指標與設限時間具有相關性時 (M and C are dependent)，其估計結果較不 robust，原因為條件 Kaplan-Meier 估計量 $\hat{S}(t \mid M_i > c)$ ，是假設設限時間與生物指標是不相關的 (C does not depend on M)。

1.3.2 最鄰近點估計法

為了解決上述方法存在的問題，Heagerty 又提出了替代的估計方法，利用加權 Kaplan-Meier 估計式，權重使用最鄰近點核平滑 (nearest neighbour kernel) 估計式，此方法原先用於二元分佈 (M, Y) 的估計，且已被證明可以在設限時間 C 與生物指標 M 有相關的情況下使用，且估計的結果具有單調性 (monotone)，因其定義在 C/D 下，故又被稱為 CD2，加權 Kaplan-Meier 估計式如下：

$$\hat{S}_{\lambda_n}(t \mid M_i) = \prod_{a \in T_n, a \leq t} \left(1 - \frac{\sum_j K_{\lambda_n}(M_j, M_i) I(Z_j = a) \delta_j}{\sum_j K_{\lambda_n}(M_j, M_i) I(Z_j \geq a)} \right)$$

其中 $K_{\lambda_n}(M_j, M_i) = I(-\lambda_n < \hat{F}_M(M_i) - \hat{F}_M(M_j) < \lambda_n)$ 為 0/1 最近點核函數 (0/1 nearest neighbourhood kernel)，平滑參數為 λ_n ， $Z_j = \min(Y_j, C_j)$ ， $\delta_j = I(Y_j = Z_j)$ ，則敏感度與特異度的估計可表示為：

$$\widehat{Se}(c, t) = \frac{(1 - \hat{F}_M(c)) - \hat{S}_{\lambda_n}(c, t)}{1 - \hat{S}_{\lambda_n}(t)}$$

$$\widehat{Sp}(c, t) = 1 - \frac{\hat{S}_{\lambda_n}(c, t)}{\hat{S}_{\lambda_n}(t)}$$

1.4 事件敏感度與靜態特異度

事件敏感度與靜態特異度 (Incident sensitivity and static specificity，簡稱 I/S)，對於病例組的定義為，在時間 t 發生事件的個體，對照組為在一固定的時間段 $(0, t^*)$ 沒有發生事件的個體，此定義方式通常用於，區分在時間 t 發生事件的個體與在一適當長的時間後都沒有發生事件的個體 ($T_i \geq t^*$)，使用此定義方式的原因為， t^* 為事先給定，且這給定的時間被認為是觀察事件發生的時間已足夠長，即在 t^* 之內沒有發生事件的個體，就認定該個體不會發生事件，舉例來說， $t^* = 2$ 年通常用於乳腺癌檢查，因為通常在檢查後的兩年內未出現疾病，則假設該個體沒有該疾病。定義方式如下：

$$\text{sensitivity}^{\mathbb{I}}(c, t) : P(M_i > c \mid T_i = t) = P(M_i > c \mid dN_i^*(t) = 1) = TP_t^{\mathbb{I}}(c)$$

$$\text{specificity}^{\mathbb{S}}(c, t^*) : P(M_i \leq c \mid T_i > t^*) = 1 - P(M_i > c \mid N_i^*(t^*) = 0) = 1 - FP_{t^*}^{\mathbb{S}}(c)$$

則 AUC 可表示為

$$AUC^{\mathbb{I}/\mathbb{S}}(t) = P(M_j > M_k \mid T_j = t, T_k > t^*)$$

在此定義下，個體不會因為時間不同，而有從對照組轉變為病例組的現象發生。

1.5 事件敏感度與動態特異度

事件敏感度與動態特異度 (Incident sensitivity and dynamic specificity，簡稱 I/D)，為本篇最主要使用的定義方式，同時也是其對於病例組的定義為，在時間 t 發生事件的個體，對照組為在時間 t 沒有發生事件的個體，在此定義下，有些個體既不是病例組也不是對照組 ($T_i < t$)，且每個發生過事件的個體可能在較早的時間為對照組 ($T_i > t$)，而在發生事件的時間點 t 轉變為病例組 ($T_i = t$)，定義如先前所述：

$$\text{sensitivity}^{\mathbb{I}}(c, t) : P(M_i > c \mid T_i = t) = P(M_i > c \mid dN_i^*(t) = 1) = TP_t^{\mathbb{I}}(c)$$

$$\text{specificity}^{\mathbb{D}}(c, t) : P(M_i \leq c \mid T_i > t) = 1 - P(M_i > c \mid N_i^*(t) = 0) = 1 - FP_t^{\mathbb{D}}(c)$$

在此定義下，AUC 可表示為：

$$AUC^{\mathbb{I}/\mathbb{D}}(t) = P(M_j > M_k \mid T_j = t, T_k > t)$$

I/D 的定義通常用於，事件發生的確切時間點 t 是知道的，並且想要區分發生事件的個體與直到時間 t 都沒有發生事件的個體，定義方式為將時間 t 下的風險集 (risk set) 以二分法分為病例組與對照組，且在此定義下，我們可以推導出一致性指標 C (Heagerty & Zheng, 2005):

$$C^\tau = \int_0^\tau AUC^{\mathbb{I}, \mathbb{D}}(t) \times w^\tau(t) dt$$

其中 $w^\tau(t) = 2f(t)S(t)/W^\tau$ ， $W^\tau = \int_0^\tau 2f(t)S(t)dt = 1 - S^2(\tau)$. 其代表的意義為，對隨機一對個體的預測與其結果一致的機率，本篇論文主要使用此種定義方式，將其應用在各個模型上。

第二章 統計方法

前述介紹過各種時間相依敏感度與特異度的定義，只有 I/D 的定義下能夠推導出一致性指標，以下將介紹透過此種定義方式，應用在不同模型中的結果。

2.1 Cox 比例風險模型

由 Cox 於 1972 年提出的比例風險模型 (Proportional hazard model, 簡稱 PH)，用以描述共變數與存活時間的關係，在醫學上廣泛被使用，因模型不需要對基線風險進行估計，但是對預測變量對於風險的效果 (effect of the explanatory variable) 用模型中的 β 進行參數估計，故屬於半參數化模型 (semi-parametric model)，因放棄了對基線風險函數的估計，也就是放寬了對它的限制和假設，使模型估計的結果更高效可靠 (efficient and reliable)，因此至今仍然廣泛被使用，當資料為時間獨立共變數時，模型如下：

$$\lambda(t|M_i) = \lambda_0(t)e^{M_i}$$

其中 $M = X^T\beta$ ， β 為回歸係數， $\lambda(t|M_i)$ 為受時間獨立共變數影響的風險函數， $\lambda_0(t)$ 為基線風險函數，由於此模型假設風險比不隨時間變化，故又稱為比例風險模型，當符合比例風險假設時，此模型的部分概似方程式如下：

$$\sum_i \Delta_i \left[M_i - \left(\sum_k \pi_k(\beta, V_i) | M_k \right) \right]$$

其中 $\pi_k(\beta, t) = R_k(t) \exp(\beta M_i) / W(t)$ ， $W(t) = \sum_j R_j(t) \exp(\beta M_j)$ ， V_i 為觀測時間，透過以上方程式，可以推導出同時具有一致性 (consistent) 與常態漸進性 (asymptotically normal) 的最大部分概似估計量 (maximum partial likelihood estimator，簡稱 MPLE) $\hat{\beta}$ ，Xu & O'Quigley(2000) 提出上述的權重 $\pi_k(\beta, V_i)$ 可被用來估計在給定事件發生時間 t 之下共變數的條件分配。透過 (1.2) 式，代入 Cox 模型的風險函數，則時間相依敏感度估計可寫為：

$$\begin{aligned} \widehat{TP}_t(c) &= \sum_{M_i > c} \frac{\lambda_0(t) e^{M_i} R_i(t)}{\sum_{j=1}^n \lambda_0(t) e^{M_j} R_j(t)} \\ &= \sum_{M_i > c} \frac{e^{M_i} R_i(t)}{\sum_{j=1}^n e^{M_j} R_j(t)} \end{aligned} \quad (1.4)$$

$\widehat{FP}_t(c)$ 及後續 ROC、AUC、一致性指標的估計皆與前面所述相同。

2.1.1 Varying-coefficient model

Heagerty & Zheng (2005) 於論文中提到，當比例風險假設不滿足時，可以使用放寬假設的 varying-coefficient 模型，模型假設如下：

$$\lambda(t|M_i) = \lambda_0(t) e^{M_i \gamma(t)}$$

其中 $\gamma(t)$ 為時變係數函數 (time-varying coefficient function)，此函數可由 Grambsch and Therneau (1994) 提出的平滑殘差估計法 (smoothly residual-based methods) 估計，使得風險比隨著時間的推移平穩變化，在此方法下，時間相依敏感度的估計為：

$$\widehat{TP}_t(c) = \hat{P}(M_i > c \mid T_i = t) = \sum_k I(M_k > c) \pi_k[\hat{\gamma}(t), t]$$

2.2 AFT 模型

AFT 模型 (Accelerated failure time model，又稱加速失效模型) 的特點是不管所謂的風險概念，而是對每個患者真正的生存時間進行模型化處理，與 Cox 模型不同的是，此模型不需要滿足比例風險的假設，但是需要對基線風險進行參數估計，故當比例風險不滿足時，可以考慮使用 AFT 模型來取代 Cox 模型 (Wei, 1992)。AFT 模型直接利用存活時間的長度來描述共變量的影響，當資料屬於時間獨立共變數時，模型如下：

$$\log T = \mu - X^T \beta + e$$

其中 T 為存活時間， X 為時間獨立共變數， μ 為截距項， β 為迴歸係數， e 為隨機誤差項，令 $S_0(\cdot)$ 表 $X = 0$ 下的基線存活函數， $\lambda_0(\cdot)$

表 $X = 0$ 下的基線存活函數，可得此模型的風險函數形式：

$$S(t | X) = S_0(te^{X^T \beta})$$

$$\lambda(t | X) = \lambda_0(te^{X^T \beta})e^{X^T \beta}$$

令 $u_i = te^{X_i^T \beta}$ ， $M_i = X_i^T \beta$ ， $\psi_1^{-1}(u_i; \hat{\beta}) = t$ ，利用 (1.2) 式，時間相依敏感度可被估計為：

$$\begin{aligned} \widehat{TP}_t(c) &= \sum_{M_i > c} \frac{\lambda_0(u_i)e^{M_i}R_i(t)/\sum_{i=1}^n R_i(t)}{\sum_{j=1}^n \lambda_0(u_j)e^{M_j}R_j(t)/\sum_{j=1}^n R_j(t)} \\ &= \sum_{M_i > c} \frac{\lambda_0(u_i)e^{M_i}R_i(\psi_1^{-1}(u_i; \hat{\beta}))}{\sum_{j=1}^n \lambda_0(u_j)e^{M_j}R_j(\psi_1^{-1}(u_j; \hat{\beta}))} \end{aligned} \quad (1.5)$$

此處 AFT 模型基線風險函數估計使用 Zeng & Lin (2007) 提出的核平滑估計式，如下：

$$\hat{\lambda}_0(u) = \frac{(nu)^{-1} \sum_{i=1}^n \Delta_i K_h(L_i - \log u)}{n^{-1} \sum_{i=1}^n \int_{-\infty}^{L_i - \log u} K_h(s) ds} \quad (2.5)$$

其中 $K_h(s) = \frac{1}{h}K(\frac{s}{h})$ ， $K(\cdot)$ 為選定的核函數 (kernel function)，此處核函數選擇標準常態分配的 pdf， h 為帶寬，分子分母各別選擇兩種帶寬：Jones (1990) 及 Jones & Sheather (1991)，分別為 $(8\sqrt{2}/3)^{1/5}\sigma_1 n^{-1/5}$ 及 $4^{1/3}\sigma_2 n^{-1/3}$ ，其中 σ_1 為未設限個體對應 L_i 的標準差， σ_2 為所有個體對應 L_i 的標準差，估計式中 $L_i = \log(V_i) + M_i$ 。其餘 $\widehat{FP}_t(c)$ 、 $\widehat{ROC}_t(p)$ 、 $AUC(t)$ 及一致性指標 (concordance) 的估計方式與前面所述相同。

2.3 PO 模型

PO 模型 (Proportional Odds model, 簡稱 PO) 由 Bennett (1983) 提出，是二分類 logit 模型的擴展，當資料不符合比例風險的假設時，除了 AFT 模型，也可以考慮使用 PO 模型，模型假設如下：

$$\log\left(\frac{P(T \leq t | X)}{1 - P(T \leq t | X)}\right) = \log\left(\frac{P(T \leq t)}{1 - P(T \leq t)}\right) + X^T \beta \quad (2.1)$$

此模型用於描述順序性反映變數與資料的關係，故 PO 模型又稱為累積比例羅吉斯模型 (Cumulative odds logit models)。在本研究中，順序性資料為存活函數，模型 (2.1) 可以改寫為：

$$\log\left(\frac{1 - S_X(t)}{S_X(t)}\right) = \log(G(t)) + X^T \beta$$

其中基線存活時間勝算 $G(t) = \frac{1 - S(t|X=0)}{S(t|X=0)}$ ，此處利用由 Doksum 於 1987 年提出的半參數轉換模型：

$$h(T) = -X^T \beta + \epsilon$$

其中 $h(\cdot)$ 為非指定的嚴格遞增函數，誤差項 ϵ 服從已知的分佈 F_ϵ ，則透過上式，可推導出 PO 模型風險函數的形式：

$$\begin{aligned} \lambda(t | X) &= g(t) \exp(X^T \beta) \lambda_0(\exp(X^T \beta) G(t)) \\ &= \frac{g(t)}{e^{-X^T \beta} + G(t)} \end{aligned}$$

其中 $G(t)$ 為基線勝算， β 為共變數的迴歸係數。首先介紹 $G(t)$ 的估計，令 $N^*(t) = (N_1^*(t), \dots, N_n^*(t))$ ， $N^*(t) = \sum_{i=1}^n N_i^*(t)$ 為一 n 維度彼此間獨立的計數過程。令 $M_i(t)$ 為平賭計數過程，將其分解後， $dN_i(t)$ 可表示如下：

$$dN_i^*(t) = S_0(t, \beta, G) dG(t) + dM_i(t)$$

其中 $S_0(t, \beta, G) = \sum_{i=1}^n R_i(t) e^{X_i^T \beta} \lambda_0(e^{X_i^T \beta} G(t-))$ ， $\lambda_0(t) = \frac{1}{1+t}$ ，透過 Breslow 估計式 $\tilde{G}(t, \beta)$ ，藉由估計式在 β 固定下估計 $G(t)$

$$\tilde{G}(t, \beta) = \int_0^t \frac{1}{S_0(s, \beta, \tilde{G})} dN_i^*(s) \quad (2.2)$$

令 $\tilde{G}(0, \beta) = 0$ ，則 $\tilde{G}(t, \beta)$ 可利用遞迴關係式估計。

接著介紹 β 的估計，此模型的部分概似函數為：

$$l(\beta) = \prod_{i=1}^n \prod_{t \geq 0} [R_i(t) \exp(X_i^T \beta) dG(t) \lambda_0(\exp(X_i^T \beta) G(t-))]^{\Delta_i} \\ \times \exp \left\{ - \int_0^\infty R_i(t) \exp(X_i^T \beta) \lambda_0(\exp(X_i^T \beta) G(t-)) dG(t) \right\}$$

$dG(t)$ 、 $G(t)$ 可由估計式代入，對 β 微分後可得

$$\tilde{U}(\beta) = \sum_{i=1}^n \int_0^\infty \left\{ \frac{w'_i(t, \beta, \tilde{G})}{w_i(t, \beta, \tilde{G})} - \frac{S'_0(t, \beta, \tilde{G})}{S_0(t, \beta, \tilde{G})} \right\} dN_i(t) \quad (2.3)$$

其中 $w_i(t, \beta, \tilde{G}) = \exp(X_i^T \beta) \lambda_0(\exp(X_i^T \beta) \tilde{G}(t-, \beta))$ ， $w'_i(t, \beta, \tilde{G}) = \frac{\partial}{\partial \beta} w_i(t, \beta, \tilde{G})$ ， $S'_i(t, \beta, \tilde{G}) = \frac{\partial}{\partial \beta} S_i(t, \beta, \tilde{G})$ ，求解後即可獲得估計值 $\hat{\beta}$ 。則 PO 模型的時

間相依敏感度估計式為：

$$\begin{aligned}
\widehat{TP}_t(c) &= \hat{P}(M > c \mid T = t) \\
&= \frac{\sum_{M_i > c} (e^{-M_i} + G(t))^{-1} R_i(t) / \sum_{i=1}^n R_i(t)}{\sum_{i=1}^n (e^{-M_i} + G(t))^{-1} R_i(t) / \sum_{i=1}^n R_i(t)} \\
&= \frac{\sum_{M_i > c} (e^{-M_i} + G(t))^{-1} R_i(t)}{\sum_{i=1}^n (e^{-M_i} + G(t))^{-1} R_i(t)} \quad (2.4)
\end{aligned}$$

由上方推導的 $\hat{\beta}$ 及 $\hat{G}(t)$ 代入後，即可獲得時間相依敏感度的估計，後續 ROC、AUC 及一致性指標的估計與之前所述相同。

以下統整本研究所使用的模型與過去文獻中所提到的方法對應的 $\widehat{TP}_t(c)$ 。

表 2.1：各種模型、方法下 $\widehat{TP}_t(c)$ 形式

Model/Method	$\widehat{TP}_t(c)$
Cox	$\sum_{M_i > c} \frac{e^{M_i} R_i(t)}{\sum_{i=1}^n e^{M_i} R_i(t)}$
Residual smooth	$\sum_{M_i > c} \frac{e^{M_i \hat{\gamma}(t)} R_i(t)}{\sum_{i=1}^n e^{M_i \hat{\gamma}(t)} R_i(t)}$
AFT	$\sum_{M_i > c} \frac{\lambda_0(u_i) e^{M_i} R_i(\psi_1^{-1}(u_i; \beta))}{\sum_{i=1}^n \lambda_0(u_i) e^{M_i} R_i(\psi_1^{-1}(u_i; \beta))}$
PO	$\sum_{M_i > c} \frac{(e^{-M_i} + G(t))^{-1} R_i(t)}{\sum_{i=1}^n (e^{-M_i} + G(t))^{-1} R_i(t)}$
Kaplan-Meier (CD1)	$\frac{\{1 - \hat{S}(t M_i > c)\} (1 - \hat{F}_M(c))}{1 - \hat{S}(t)}$
Nearest neighbour (CD2)	$\frac{(1 - \hat{F}_M(c)) - \hat{S}_{\lambda_n}(c, t)}{1 - \hat{S}_{\lambda_n}(t)}$

第三章 模擬研究

此章節分兩個部分，第一部分針對舊有的方法提出修正，並以相同模擬方式，比較其不同的結果，使 AUC 的估計結果介於較合理的 0.5 至 1 的範圍內，首先生成存活時間 T 來自三種分配，分別是 Weibull、Lognormal 與 Loglogistic，以及透過三種不同的模型 (PH、AFT 與 PO)，得到不同的模型分數 (model score)，此處為時間獨立共變數的模擬，以下為模擬步驟：

步驟一：生成樣本數為 n 的隨機變數 $X \sim N(0, 1)$ ， $C \sim Exp(\lambda_c)$ ， $U \sim U(0, 1)$ 。

步驟二：存活時間 T 根據存活模型假設以逆轉換法生成。

步驟三：觀測時間 $V_i = \min(T_i, C_i)$ ，若 $V_i = T_i$ ，則 $\Delta_i = 1$ ；否則， $\Delta_i = 0$ 。

步驟四：根據不同風險模型時間相依敏感度定義，利用觀測時間 V 、設限指標 Δ 與共變數 X 得到計算預測精準度所需的 β 、 $\lambda_0(u)$ 與基線勝算估計 $G(t)$ 。

步驟五：將觀測時間 V 、設限指標 Δ 、共變數 X 、 β 、 $\lambda_0(u)$ 與 $G(t)$ ，帶入模型中估計 AUC 與一致性指標。

步驟二存活時間的生成方法根據不同分配與風險模型假設替換。

3.1 修正計算方式

關於修正部分，因 TP 與 FP 的估計，皆定義在事件發生的時間點上 (event time) 計算，所以實際畫出的 ROC 曲線是由多個離散點所構成的曲線，故在估算其面積 (AUC) 時，是採用梯形法累加面積獲得。

本研究針對原有的 AUC 計算方式做修正，根據 I/D 的定義，時間相依敏感度的估計是定義在風險集 (risk set) 下做計算，時間相依特異度為定義在風險集扣除時間 t 下做計算，故特異度的估計不應有時間 t 下的資料，而實際上，原始方法在計算時，並未將此筆數據去除 (如圖 3.1 紅點)，圖中紅點即為事件發生 t 下的資料點所估計出的 TP 與 FP，即使在 AUC 面積的估算上，此點的面積貢獻為 0，但仍然會影響鄰近點 (面積高度) 的面積計算，以圖 3.1 為例，對於 FP 為 0.5 至 0.6 間的面積計算，原先的計算方式為 $(0.6 - 0.5)(0.754 - 0.675)$ ，面積的高受到事件發生時間 t 下的資料點影響，修正後的計算方式為 $(0.6 - 0.5)(0.816 - 0.675)$ ，排除紅點造成的影響，此現象在試驗初期只造成一小塊面積的差異，所以影響不大，但是到了後期，因試驗人數下降，面積的影響便會逐漸放大 (如圖 3.2)，甚至會造成非常極端的現象 (如圖 3.3)，AUC 僅有 0.272，本研究針對此部分做修正。

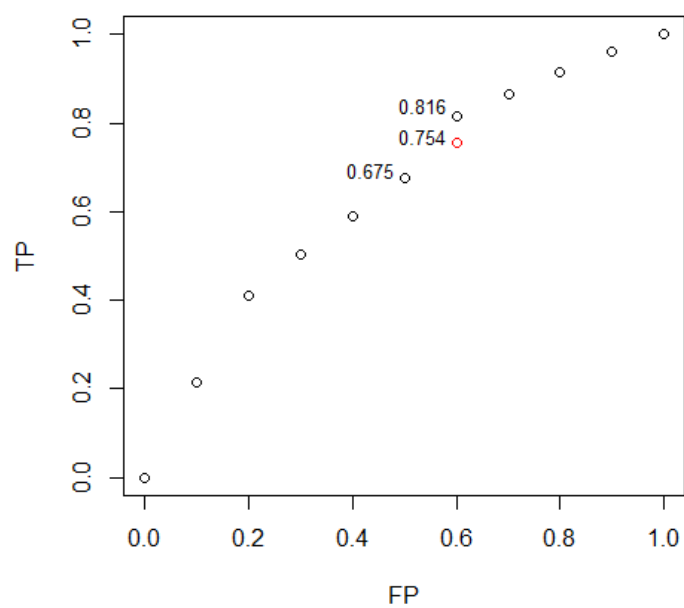


圖 3.1：原始計算方式重複點問題

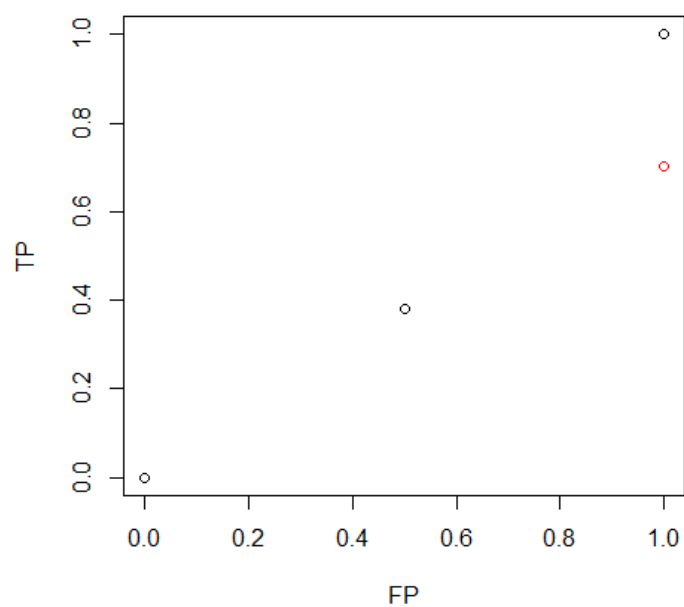


圖 3.2：試驗後期 TP 與 FP 分佈

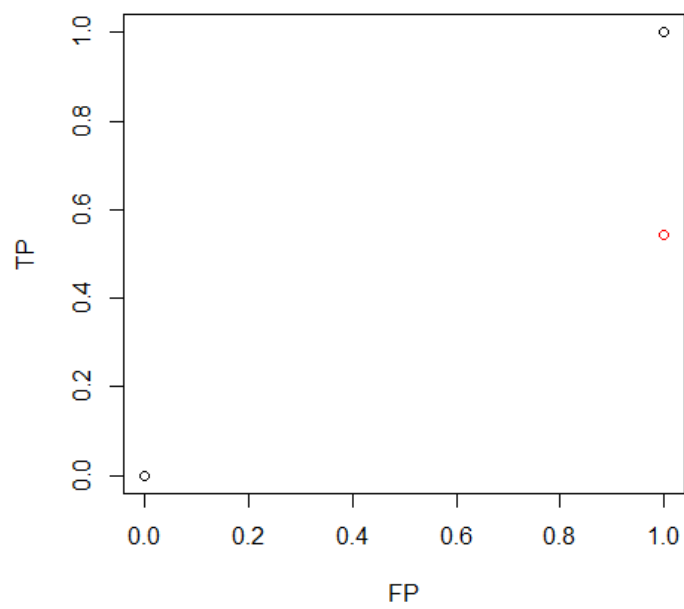


圖 3.3：試驗後期極端現象

3.2 Cox 模型模擬

模擬存活時間分別來自 Lognormal 與 Weibull，並使用半參數 Cox 模型，在不同設限率下，比較修正前後的 AUC 與一致性指標的差異。

3.2.1 Lognormal - Cox

模擬存間來自 Lognormal 分配，共變數 X 為標準常態分佈，參數為 $\mu = 0.5$ ， $\sigma = 2.5$ ， $\beta = 1$ ，設限時間為指數分配，設限率為 20% 及 40%，參數分別為 $\lambda_c = 0.01, 0.18$ ，故模型可表示為：

$$\lambda(t | X) = \frac{1}{\sigma t} \frac{\phi\left(\frac{\ln t - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)} e^{\beta x}$$

透過逆轉換法可求得存活時間

$$S(t | X) = \left(1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right)\right)^{\exp\{\beta x\}}$$
$$t = \exp\{\sigma \Phi^{-1}(1 - U^{\exp\{-\beta x\}}) + \mu\}, U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

根據表 3.1，可以看出當資料來自 Lognormal - Cox，在修正前，20% 設限率下，原先的估計方法，在後期（ex: $t = \exp(3)$ ）有低估的現象，而修正過後則略有高估，但更接近真實值；在 40% 設限率下，原

先的估計方法，在後期(ex : $t = \exp(2.5), \exp(3)$)有明顯低估的現象，修正過後，AUC 的估計更接近真實值，標準差也較小。

表 3.1 : Lognormal-Cox 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-2.50	0.759	0.759	0.026	0.759	0.026
-2.00	0.754	0.755	0.026	0.755	0.026
-1.50	0.746	0.747	0.024	0.747	0.024
-1.00	0.734	0.735	0.022	0.735	0.022
-0.50	0.721	0.721	0.021	0.721	0.021
0.00	0.709	0.709	0.022	0.709	0.022
0.50	0.698	0.697	0.023	0.697	0.023
1.00	0.689	0.688	0.026	0.688	0.027
1.50	0.681	0.682	0.036	0.684	0.036
2.00	0.675	0.671	0.069	0.679	0.065
2.50	0.670	0.640	0.191	0.686	0.113
3.00	0.665	0.531	0.323	0.669	0.141
C	0.726	0.726	0.019	0.726	0.019
40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-2.50	0.759	0.759	0.028	0.759	0.028
-2.00	0.754	0.755	0.027	0.755	0.027
-1.50	0.746	0.747	0.025	0.747	0.025
-1.00	0.734	0.735	0.022	0.735	0.022
-0.50	0.721	0.721	0.023	0.721	0.023
0.00	0.709	0.708	0.023	0.709	0.023
0.50	0.698	0.697	0.025	0.697	0.025
1.00	0.689	0.687	0.034	0.688	0.034
1.50	0.681	0.681	0.066	0.688	0.063
2.00	0.675	0.627	0.219	0.686	0.119
2.50	0.670	0.549	0.302	0.667	0.135
3.00	0.665	0.531	0.311	0.661	0.135
C	0.726	0.726	0.020	0.727	0.020

3.2.2 Weibull - Cox

模擬存間來自 Weibull 分配，共變數 X 為標準常態分佈，參數為 $\alpha = 1.5$ ， $\lambda = 0.25$ ， $\beta = 1$ ，設限時間為指數分配，設限率分別為 20% 及 40%，參數分別為 $\lambda_c = 0.007$ ， 0.027 ，故模型可表示為：

$$\lambda(t | X) = \alpha \lambda t^{\alpha-1} e^{\beta x}$$

透過逆轉換法可求得存活時間

$$S(t | X) = \exp\{-\lambda t^\alpha e^{\beta x}\}$$
$$t = \left(-\frac{1}{\lambda e^{\beta x}} \ln U\right)^{\frac{1}{\alpha}}, \quad U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

透過表 3.2，可以看出當資料來自 Weibull - Cox，不論設限率是 20% 或 40%，在後期，修正前都有明顯低估的現象，修正後更為接近真實值，且標準差也比修正前小，一致性指標同樣差異不大。

表 3.2 : Weibull-Cox 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-2.50	0.759	0.758	0.027	0.758	0.027
-2.00	0.757	0.757	0.026	0.757	0.026
-1.50	0.754	0.753	0.024	0.753	0.024
-1.00	0.748	0.748	0.023	0.748	0.023
-0.50	0.740	0.740	0.022	0.740	0.022
0.00	0.728	0.728	0.021	0.728	0.021
0.50	0.714	0.714	0.020	0.714	0.020
1.00	0.699	0.699	0.022	0.700	0.022
1.50	0.684	0.685	0.032	0.686	0.032
2.00	0.670	0.665	0.083	0.676	0.070
2.50	0.658	0.440	0.358	0.642	0.152
C	0.726	0.726	0.018	0.726	0.018
40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-2.50	0.759	0.759	0.029	0.759	0.029
-2.00	0.757	0.757	0.028	0.757	0.028
-1.50	0.754	0.754	0.027	0.754	0.027
-1.00	0.748	0.749	0.026	0.749	0.026
-0.50	0.740	0.740	0.024	0.740	0.024
0.00	0.728	0.728	0.023	0.728	0.023
0.50	0.714	0.715	0.024	0.715	0.024
1.00	0.699	0.701	0.028	0.702	0.028
1.50	0.684	0.687	0.054	0.691	0.053
2.00	0.670	0.592	0.284	0.687	0.142
2.50	0.658	0.373	0.380	0.627	0.159
C	0.726	0.726	0.020	0.727	0.020

3.3 AFT 模型模擬

模擬存活時間分別來自 Lognormal 與 Weibull，並使用半參數 AFT 模型，在不同設限率下，比較修正前後的 AUC 與一致性指標的差異。

3.3.1 Lognormal - AFT

模擬存活時間來自 Lognormal 分配，共變數 X 來自標準常態分佈，參數為 $\mu = 1.0$ ， $\sigma = 2.5$ ， $\beta = 1.0$ ，設限時間為指數分配，設限率為 20% 及 40%，參數分別為 $\lambda_c = 0.08, 0.15$ ，故模型可表示為：

$$\lambda(t | X) = \frac{1}{\sigma t} \frac{\phi\left(\frac{\ln t - \beta x - \mu}{\sigma}\right)}{1 - \Phi\left(\frac{\ln t - \beta x - \mu}{\sigma}\right)}$$

透過逆轉換法可求得存活時間

$$S(t | X) = 1 - \Phi\left(\frac{\ln t - \beta x - \mu}{\sigma}\right)$$
$$t = \exp\{\sigma \Phi^{-1}(1 - U) + \beta x + \mu\}, \quad U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

根據表 3.3，可以看出當資料來自 Lognormal - AFT，設限率為 20% 時，在後期（ex: $t = \exp(5)$ ），修正前有明顯的低估現象，修正後雖然有些微高估，但仍然較接近真實值，標準差也較小，一致性指

標差異不大；當設限率為 40% 時，同樣於後期，修正前明顯低估，修正後更接近真實值，值得注意的是，當 $t = \exp(4), \exp(5)$ 時，AUC 的估計值都相同，原因為後期的存活人數過少，導致 $t = \exp(4), \exp(5)$ 時的人數都相同 ($t = \exp(4)$ 及 $t = \exp(5)$ 間的人數、存活、死亡沒有任何改變)，所以估計出的 TP 與 FP 都相同。

表 3.3 : Lognormal-AFT 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-5.00	0.731	0.738	0.093	0.738	0.093
-4.00	0.701	0.704	0.063	0.704	0.063
-3.00	0.672	0.674	0.050	0.674	0.050
-2.00	0.646	0.646	0.040	0.646	0.040
-1.00	0.623	0.624	0.035	0.624	0.035
0.00	0.604	0.604	0.032	0.604	0.033
1.00	0.588	0.588	0.030	0.588	0.030
2.00	0.575	0.576	0.032	0.576	0.032
3.00	0.564	0.564	0.048	0.565	0.048
4.00	0.556	0.562	0.117	0.571	0.114
5.00	0.549	0.467	0.297	0.592	0.174
C	0.623	0.623	0.024	0.623	0.024

40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-5.00	0.731	0.738	0.096	0.738	0.096
-4.00	0.701	0.704	0.066	0.704	0.066
-3.00	0.672	0.673	0.052	0.673	0.052
-2.00	0.646	0.647	0.040	0.647	0.040
-1.00	0.623	0.623	0.038	0.623	0.038
0.00	0.604	0.604	0.035	0.604	0.035
1.00	0.588	0.585	0.040	0.585	0.040
2.00	0.575	0.572	0.063	0.573	0.063
3.00	0.564	0.554	0.203	0.592	0.162
4.00	0.556	0.490	0.264	0.581	0.167
5.00	0.549	0.490	0.264	0.581	0.167
C	0.623	0.626	0.027	0.628	0.027

3.3.2 Weibull-AFT

模擬存活時間來自 Weibull 分配，共變數 X 來自標準常態分佈，參數為 $\alpha = 1.5$ ， $\lambda = 0.05$ ， $\beta = 1$ ，設限時間為指數分配，設限率為 20% 及 40%，參數分別為 $\lambda_c = 0.02$ ， 0.07 ，故模型可表示為：

$$\lambda(t | X) = \alpha \lambda t^{\alpha-1} e^{\alpha \beta x}$$

透過逆轉換法可求得存活時間

$$S(t | X) = \exp\{-\lambda t^\alpha e^{\alpha \beta x}\}$$
$$t = \left(-\frac{1}{\lambda e^{\alpha \beta x}} \ln U\right)^{\frac{1}{\alpha}}, \quad U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

根據表 3.4，當資料來自 Weibull - AFT，在後期修正前的低估情況非常嚴重（ex：t = exp(4.25), exp(4.5)），修正後更合理且較接近真實值，標準差也明顯下降，一致性指標同樣差異不大。

表 3.4 : Weibull-AFT 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-1.00	0.840	0.845	0.040	0.845	0.040
0.00	0.817	0.820	0.026	0.820	0.026
1.00	0.785	0.786	0.024	0.786	0.024
2.00	0.751	0.750	0.026	0.750	0.027
3.00	0.720	0.720	0.047	0.722	0.047
3.50	0.707	0.710	0.104	0.721	0.089
4.00	0.695	0.572	0.344	0.704	0.166
4.25	0.690	0.409	0.397	0.649	0.180
4.50	0.684	0.324	0.399	0.624	0.176
C	0.790	0.788	0.016	0.788	0.016
40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-1.00	0.840	0.845	0.041	0.845	0.041
0.00	0.817	0.819	0.028	0.819	0.028
1.00	0.785	0.784	0.027	0.784	0.027
2.00	0.751	0.749	0.033	0.749	0.033
3.00	0.720	0.719	0.102	0.731	0.090
3.50	0.707	0.586	0.353	0.718	0.171
4.00	0.695	0.459	0.401	0.674	0.179
4.25	0.690	0.457	0.401	0.673	0.179
4.50	0.684	0.457	0.401	0.673	0.179
C	0.790	0.788	0.018	0.788	0.018

3.4 PO 模型模擬

模擬存活時間分別來自 Lognormal 與 Loglogistic，並使用半參數 PO 模型，在不同設限率下，比較修正前後的 AUC 與一致性指標的差異。

3.4.1 Lognormal - PO

模擬存活時間來自 Lognormal 分配，共變數 X 來自標準常態分佈，參數為 $\beta = 1$ ， $\mu = 1$ ， $\sigma = 0.3$ ，設限時間為指數分配，設限率為 20% 及 40%，參數分別為 $\lambda_c = 0.06$ ， 0.16 ，故模型可表示為：

$$\lambda(t | X) = \frac{\phi(\frac{\ln t - \mu}{\sigma})e^{\beta x}}{(1 - \Phi(\frac{\ln t - \mu}{\sigma}))^2((1 - \Phi(\frac{\ln t - \mu}{\sigma}))^{-1} - 1)e^{\beta x} + 1}$$

其中 $\phi(\cdot)$ 與 $\Phi(\cdot)$ 為標準常態分配的機率密度函數與累積機率密度函數，透過逆轉換法可求得存活時間

$$S(t | X) = (1 + (1 - \Phi(\frac{\ln t - \mu}{\sigma}))^{-1}e^{\beta x})^{-1}$$
$$t = \exp\{\sigma\Phi^{-1}(1 - ((U^{-1} - 1)e^{-\beta x} + 1)^{-1}) + \mu\}, U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

根據表 3.5，當資料來自 Lognormal - PO，修正前於後期有非常嚴重的低估，甚至在 40% 設限率下， $t = \exp(2)$ 的估計值僅 0.047 (真實值

為 0.501)，修正後為 0.512，標準差也較修正前低，一致性指標差異不大。

表 3.5：Lognormal-PO 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
0.00	0.760	0.757	0.034	0.757	0.034
0.25	0.755	0.752	0.033	0.752	0.033
0.50	0.730	0.728	0.028	0.728	0.028
0.75	0.683	0.682	0.024	0.682	0.024
1.00	0.631	0.631	0.020	0.631	0.020
1.25	0.579	0.580	0.020	0.580	0.020
1.50	0.533	0.535	0.040	0.540	0.038
1.75	0.506	0.345	0.298	0.570	0.110
2.00	0.501	0.345	0.298	0.507	0.040
C	0.669	0.668	0.022	0.668	0.022
40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
0.00	0.760	0.753	0.039	0.753	0.039
0.25	0.755	0.748	0.037	0.748	0.037
0.50	0.730	0.725	0.032	0.725	0.032
0.75	0.683	0.681	0.027	0.681	0.027
1.00	0.631	0.630	0.024	0.630	0.024
1.25	0.579	0.578	0.024	0.578	0.024
1.50	0.533	0.537	0.080	0.553	0.056
1.75	0.506	0.212	0.294	0.547	0.100
2.00	0.501	0.047	0.169	0.512	0.055
C	0.669	0.666	0.025	0.666	0.025

3.4.2 Loglogistic - PO

模擬存活時間來自 Loglogistic 分配，共變數 X 來自標準常態分佈，參數為 $\beta = 1$ ， $\mu = 0.5$ ， $\sigma = 0.2$ ，設限時間為指數分配，設限率為 20% 及 40%，參數分別為 $\lambda_c = 0.12, 0.30$ ，故模型可表示為：

$$\lambda(t | X) = \frac{1}{\sigma t} \frac{\exp\{\frac{\ln t - \mu}{\sigma}\}}{(1 + \exp\{\frac{\ln t - \mu}{\sigma}\})} e^{\beta x}$$

透過逆轉換法可求得存活時間：

$$S(t | X) = (1 + \exp(\frac{\ln t - \mu}{\sigma}) e^{\beta X})^{-1}$$
$$t = \exp \left\{ \sigma \left[\ln\left(\frac{1}{U} - 1\right) - \beta X \right] + \mu \right\}, \quad U \sim U(0, 1)$$

真實值則經由已知分配，透過定義計算出各時間點的 AUC 及一致性指標。

根據表 3.6，當資料來自 Loglogistic - PO，修正前的後期，低估的現象非常嚴重，修正後的估計更接近真實值，同時也降低標準差，一致性指標差異不大。

表 3.6 : Loglogistic-PO 模擬結果

20% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-1.00	0.760	0.755	0.035	0.755	0.035
-0.50	0.754	0.751	0.033	0.751	0.033
0.00	0.718	0.716	0.027	0.716	0.027
0.25	0.679	0.677	0.023	0.677	0.023
0.50	0.631	0.630	0.020	0.630	0.020
0.75	0.583	0.583	0.019	0.583	0.019
1.00	0.544	0.545	0.027	0.547	0.027
1.25	0.519	0.520	0.102	0.550	0.075
1.50	0.507	0.321	0.291	0.562	0.106
1.75	0.502	0.121	0.249	0.530	0.081
2.00	0.501	0.067	0.193	0.516	0.061
C	0.670	0.667	0.022	0.667	0.022

40% censoring					
Log time	Original			Modify	
	AUC(t)	EST	SD	EST	SD
-1.00	0.760	0.753	0.039	0.753	0.039
-0.50	0.754	0.749	0.037	0.749	0.037
0.00	0.718	0.715	0.029	0.715	0.029
0.25	0.679	0.677	0.026	0.677	0.026
0.50	0.631	0.630	0.024	0.630	0.024
0.75	0.583	0.583	0.024	0.584	0.024
1.00	0.544	0.543	0.044	0.550	0.041
1.25	0.519	0.460	0.233	0.577	0.103
1.50	0.507	0.224	0.304	0.555	0.104
1.75	0.502	0.124	0.253	0.531	0.084
2.00	0.501	0.111	0.241	0.528	0.079
C	0.670	0.667	0.024	0.667	0.024

3.5 修正計算方式結論

根據表 3.1~3.6，可以看出當實驗越後期，存活人數降低，估算 AUC 的梯形法，會明顯受到事件發生點 (event time) 的資料所影響，若無移除該點，則 AUC 的估計會有明顯的低估現象，透過修正，能夠使估計值更接近真實值，同時也降低標準差。一致性指標的部分，由於後期人數較少，估算時的權重也較低，所以沒有明顯的改變。

第二部分生成 $(M_i, \log T_i)$ 來自二元常態分佈，針對不同的統計方法、模型，估計其結果，比較各個模型之間的差異，以下為模擬步驟：

步驟一：生成樣本數為 n 的隨機變數 $(M_i, \log T_i) \sim BN(0, 1)$ ，其中 $\rho = -0.7$ 。

步驟二：觀測時間 $V_i = \min(T_i, C_i)$ ，若 $V_i = T_i$ ，則設限指標 $\Delta_i = 1$ ；否則， $\Delta_i = 0$ 。

步驟三：根據不同風險模型時間相依敏感度定義，利用觀測時間 V 、設限指標 Δ 與共變數 X 得到計算預測精準度所需的 β 、 $\lambda_0(u)$ 與基線勝算估計 $G(t)$ 。

步驟四：將觀測時間 V 、設限指標 Δ 、共變數 X 、 β 、 $\lambda_0(u)$ 與 $G(t)$ ，帶入模型中估計 AUC 與一致性指標。

步驟二存活時間的生成方法根據不同分配與風險模型假設替換。

比較的方法包括 Cox 模型、不符合比例風險時的殘差和平滑估計方法、PO 模型、AFT 模型、Kaplan-Meier 估計方法、最鄰近點估計方法，設限率為 20% 與 40% ，且分別考慮 $n = 200$ 與 $n = 500$ 兩種模擬。

3.6 Bivariate normal 模擬

模擬資料 $(M_i, \log T_i)$ 來自二元常態分佈，參數設定 $\mu_1 = 0$, $\mu_2 = 0$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\rho = -0.7$, 設限時間 C 來自 lognormal 分配，參數設定根據設限率 20% 及 40% , 分別為 $\mu = 1.19$, 0.36 , 標準差 $\sigma = 1$, 模擬次數 $m = 1000$ 。

表 3.1 : Bivariate normal 模擬結果 (n = 200)

Log time	AUC(t)	Cox model		Residual smooth		PO model		AFT model		CD1		CD2	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20% censoring													
-2.0	0.884	0.754	0.028	0.880	0.040	0.848	0.024	0.885	0.020	0.917	0.060	0.854	0.062
-1.5	0.833	0.746	0.026	0.840	0.032	0.821	0.021	0.836	0.016	0.884	0.044	0.818	0.048
-1.0	0.782	0.735	0.024	0.777	0.031	0.786	0.019	0.783	0.014	0.857	0.034	0.774	0.039
-0.5	0.734	0.726	0.021	0.727	0.034	0.743	0.019	0.736	0.015	0.836	0.030	0.731	0.033
0.0	0.693	0.717	0.021	0.688	0.037	0.698	0.020	0.694	0.016	0.831	0.031	0.704	0.029
0.5	0.660	0.711	0.024	0.663	0.041	0.655	0.023	0.662	0.020	0.837	0.032	0.692	0.029
1.0	0.634	0.703	0.039	0.645	0.060	0.615	0.032	0.636	0.032	0.858	0.041	0.697	0.035
1.5	0.614	0.693	0.092	0.653	0.100	0.607	0.076	0.639	0.071	0.880	0.062	0.714	0.070
2.0	0.598	0.647	0.163	0.642	0.159	0.598	0.126	0.640	0.166	0.876	0.105	0.628	0.253
C	0.741	0.728	0.020	0.744	0.018	0.745	0.018	0.749	0.011				
40% censoring													
-2.0	0.884	0.768	0.031	0.881	0.042	0.850	0.024	0.885	0.031	0.914	0.061	0.885	0.062
-1.5	0.833	0.759	0.028	0.836	0.034	0.823	0.022	0.836	0.027	0.888	0.043	0.819	0.048
-1.0	0.782	0.749	0.026	0.773	0.035	0.787	0.020	0.784	0.026	0.858	0.036	0.773	0.040
-0.5	0.734	0.738	0.023	0.726	0.039	0.744	0.020	0.735	0.027	0.837	0.033	0.731	0.035
0.0	0.693	0.729	0.024	0.689	0.044	0.700	0.022	0.693	0.030	0.830	0.036	0.703	0.031
0.5	0.660	0.721	0.032	0.666	0.053	0.657	0.029	0.661	0.049	0.837	0.040	0.690	0.031
1.0	0.634	0.710	0.068	0.658	0.084	0.625	0.056	0.646	0.099	0.857	0.063	0.695	0.048
1.5	0.614	0.683	0.149	0.666	0.150	0.634	0.125	0.657	0.169	0.857	0.103	0.650	0.196
2.0	0.598	0.626	0.161	0.618	0.158	0.590	0.127	0.615	0.167	0.854	0.114	0.536	0.285
C	0.741	0.740	0.021	0.744	0.020	0.748	0.019	0.751	0.020				

表 3.2 : Bivariate normal 模擬結果 (n = 500)

Log time	AUC(t)	Cox model		Residual smooth		PO model		AFT model		CD1		CD2	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
20% censoring													
-2.0	0.884	0.751	0.018	0.886	0.025	0.851	0.015	0.885	0.020	0.918	0.034	0.867	0.041
-1.5	0.833	0.743	0.017	0.840	0.020	0.824	0.013	0.835	0.016	0.886	0.026	0.820	0.035
-1.0	0.782	0.733	0.015	0.774	0.021	0.787	0.012	0.783	0.014	0.857	0.022	0.771	0.029
-0.5	0.734	0.724	0.014	0.725	0.023	0.745	0.012	0.734	0.015	0.836	0.019	0.729	0.025
0.0	0.693	0.716	0.014	0.688	0.025	0.699	0.012	0.693	0.016	0.829	0.019	0.703	0.022
0.5	0.660	0.710	0.015	0.662	0.028	0.654	0.014	0.661	0.020	0.837	0.021	0.696	0.023
1.0	0.634	0.704	0.023	0.642	0.037	0.613	0.018	0.634	0.032	0.856	0.027	0.705	0.027
1.5	0.614	0.695	0.047	0.635	0.059	0.584	0.035	0.619	0.071	0.884	0.040	0.732	0.044
2.0	0.598	0.676	0.137	0.653	0.137	0.609	0.109	0.638	0.166	0.883	0.093	0.723	0.176
C	0.741	0.726	0.013	0.743	0.012	0.746	0.011	0.747	0.011				
40% censoring													
-2.0	0.884	0.765	0.020	0.886	0.025	0.851	0.015	0.885	0.019	0.918	0.035	0.867	0.041
-1.5	0.833	0.756	0.018	0.833	0.021	0.824	0.014	0.835	0.016	0.885	0.027	0.820	0.035
-1.0	0.782	0.746	0.016	0.771	0.024	0.787	0.013	0.783	0.015	0.856	0.023	0.771	0.029
-0.5	0.734	0.736	0.015	0.725	0.025	0.745	0.013	0.734	0.016	0.836	0.021	0.729	0.025
0.0	0.693	0.728	0.015	0.689	0.029	0.700	0.014	0.694	0.019	0.829	0.021	0.703	0.023
0.5	0.660	0.721	0.019	0.666	0.035	0.655	0.017	0.660	0.028	0.836	0.026	0.695	0.024
1.0	0.634	0.711	0.037	0.648	0.053	0.615	0.028	0.635	0.057	0.855	0.038	0.704	0.033
1.5	0.614	0.708	0.103	0.665	0.108	0.610	0.088	0.656	0.140	0.873	0.071	0.727	0.093
2.0	0.598	0.641	0.164	0.634	0.160	0.594	0.123	0.630	0.177	0.852	0.122	0.600	0.263
C	0.741	0.738	0.014	0.743	0.013	0.747	0.012	0.748	0.012				

首先，當我們使用 Cox 模型，實際上資料並不符合比例風險，所以使用錯誤的模型假設下，會造成估計結果出現早期低估，而後期高估的現象，整體的一致性指標則略有低估。接著使用 Cox 所提出的半母數風險模型，假設風險比隨著時間平穩變化，時變系數函數 $\gamma(t)$ 以平滑殘差估計法估計 (residual smooth)，與原先的 Cox 模型相比，此方法的估計結果更接近真實值，但是到了後期，因為人數明顯減少，使的估計結果較不穩定且有高估的情況，一致性指標的部分則與真實值較為接近。接著使用 PO 模型，對於 AUC 的估計較前兩種方法穩定 (SD 較小)，但在早期略有低估，而後期也稍有不穩定的現象，一致性指標的部分則略有高估。AFT 模型的部分，除了後期實驗人數較少，估計結果較不穩定，其餘部分都與真實值非常接近。CD1 的部分，先前有提到此模型假設設限時間與生物指標是不相關的，而實際

上資料並不會完全無關，使用錯誤模型的結果，使得估計結果出現嚴重高估的現象，由於是在 C/D 定義之下，故不會有一致性指標。接著使用 Heagerty 提出的最鄰近點估計法，相較於 CD1，結果更接近真實值，但仍然於後期有高估的現象。另外，由以上兩表格可以看出，當模擬資料個數 n 越多，整體估計會更加穩定 (標準差變小)，以及當設限比例越高，標準差也會隨著增加。

第四章 資料分析

本章節中，以時間獨立共變數存活模型分析資料，並以 Heagerty & Zheng (2005) 所提出的一致性指標來探討何種風險模型透過使用生物指標預測效果較好，透過不同模型的選擇，找出較適合此筆資料的估計方法。

4.1 愛滋病資料

本研究使用資料取自西元 1990 年至 2003 年的台灣愛滋病患者資訊，病人須每隔一段時間回診追蹤一次 (3 至 4 個月)，每位患者至少有一次測量紀錄，至多 13 次，其中共有 1164 位病人感染 HIV 病毒，變數包含患者感染 HIV 病毒至死亡期間的 CD4 數量與使用的治療藥劑類別，包含雞尾酒療法 (HAART) 與標準療法 (Standard treatment)。根據過去文獻，我們使用其中 137 位愛滋病患者作為本研究的實例資料，其中 37 位患者使用雞尾酒療法，其餘 100 位患者使用標準療法，最後共有 68 位患者發病而死亡，本研究探討患者從被發現感染 HIV 病毒到因愛滋病病發而死亡為止，此段時間其共變數與存活時間的關係。

4.2 時間獨立共變數存活模型

以下分別考慮三種模型，其中以每位患者第一次測量的 CD4 指數與治療藥劑類別作為時間獨立共變數資料，比較三個模型之間的差異，三種模型如下：

模型 I：變數為 CD4

模型 II：變數為治療藥劑類別

模型 III：變數為 CD4 + 治療藥劑類別

首先考慮不同變數之下，各個模型的一致性指標，由下方表 4.1 中可以看出，不論在哪個模型假設之下，模型 III 的一致性指標都最高，表示將兩變數 CD4 數量與治療藥劑類別都放入模型中較為合適，所以後續都選擇以兩變數作為分析變數。

表 4.1: 不同模型下一致性指標比較

	Cox	AFT	PO
模型 I	0.611	0.567	0.737
模型 II	0.594	0.511	0.598
模型 III	0.675	0.646	0.797

接著我們考慮在 AFT 模型下，改變不同的帶寬 (bandwidth)，對其估計結果的影響，由圖 4.1 可看出，三種帶寬的一致性指標差異不大，其中以帶寬為 1/2 時的一致性指標最高，但是其 AUC 表現並不穩定，浮動較大，所以選擇先前提到的最佳帶寬較為合適。

根據圖 4.2，PO 模型於三種模型之中表現最好，一致性指標達到 0.797，AFT 模型於初期有不穩定的情形，原因為基線風險函數

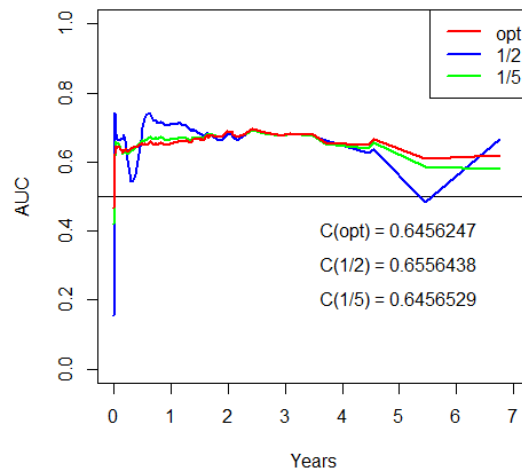


圖 4.1：模型 III 在 AFT 模型下不同帶寬 AUC

使用核平滑估計法，使得在估計第一點時會有偏差，一致性指標為 0.645。Cox 模型的曲線起伏與 PO 模型類似，但每一點的預測效果相比之下較差，整體而言，當資料屬於時間獨立共變數時，PO 模型有較好的預測能力 ($0.797 > 0.675 > 0.645$)。

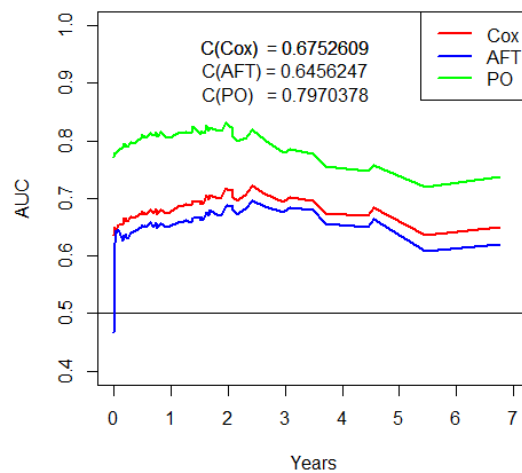


圖 4.2：模型 III 下不同風險模型各時間點 AUC

第五章 結論

過去文獻已將基於比例風險模型、加速失敗模型及比例勝算模型下的時間相依 ROC 曲線，推導為風險迴歸的函數。本研究針對時間相依 AUC 的計算方式提出修正，透過微調計算方式，改善實驗後期 AUC 估計結果低估的現象，一致性指標也隨之修正，不過由於實驗後期的權重較低，故修正前後無明顯差距。

從模擬結果中可知，當資料來自不同模型時，大多數實驗後期 AUC 的估計結果都會明顯低估，修正後雖然也有部分偏差，原因為後期實驗人數越來越少，導致估計的準確度下降，但整體的估計結果仍然較修正前更貼近真實值，標準差也明顯降低，表示此修正方法的可行性。接著將過去文獻提出的模型與估計方法，與本研究提出的修正方法，應用在模擬資料上，展示各個模型與估計方法下的差異，整體而言，後期的估計以 PO 模型的結果最為接近真實值，一致性指標以核平滑估計法最為接近真實值，可能較適合此筆模擬資料。最後以實際愛滋病資料，分析並找出最適合的模型，以 PO 模型包含所有變數 (CD4 + 治療藥劑類別) 的結果最為適合。

本篇並未使用大樣本性質估計，未來可考慮使用大樣本性質作為估計基礎，延伸套討是否能夠更有效更準確的估計，並且可以延伸到時間相依共變數的應用。

参考文献

- Bennett, S. (1983). Analysis of Survival Data by the Proportional Odds Model. *Statistics in Medicine* **2**, 273–277.
- Box, G., and Cox, D. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society Series B*, **26**, 211-252.
- Chiou, S. H., Kang, S. and Yan, J. (2014). Fitting Accelerated Failure Time Models in Routine Survival Analysis with R Package aftgee. *Journal of Statistical Software* **61**, 1–23.
- Cox, D. R. (1975). Partial likelihood. *Biometrics* **62**, 269-276.
- Heagerty, P. J., Saha-Chaudhuri, P. and Saha-Chaudhuri, M. P. (2012). Package ‘risksetROC’ : riskset ROC curve estimation from censored survival data.
- Heagerty, P.J. and Zheng, Y. (2005). Survival Model Predictive Accuracy and ROC Curves. *Biometrika* **61**, 92-105.
- Jones, M. C. (1990). The performance of kernel density functions in kernel distribution function estimation. *Statistics and Probability Letter* **9**, 129–132.
- Jones, M. C. and Sheather, S. J. (1991). Using non-statistic terms to advan-

tage in kernel-based estimation of integrated squared density derivatives.

Statistics and Probability Letter **11**, 511–514.

Kamarudin, A. N., Cox, T. and Kolamunnage-Dona, R. (2017). Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Medical Research Methodology*, 17-53.

Lin, Y. H. (2017). Model-base Time dependent AUC and Predictive Accuracy. *Graduate Institute of Statistics, National Central University*.

Wei, L. J. (1992). The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine* **11**, 1871–1879.

林威廷 (2021)。風險回歸模型下時間相依 ROC 曲線。國立中央大學統計研究所碩士論文。

附錄 A.1 時間相依 AUC 與一致性指標推導

Heagerty & Zheng (2005) 於文章中提及三種時間相依敏感度與特異度的定義，以下以 incident sensitivity 與 dynamic specificity 為例，可將任意時間 t 下的 AUC 推導為：

$$\begin{aligned}
 AUC(t) &= \int_0^1 ROC_t^{\mathbb{I}/\mathbb{D}}(p) dp \\
 &= \int_0^1 TP_t^{\mathbb{I}}\{[FP_t^{\mathbb{D}}]^{-1}(p)\} dp \\
 &= \int_{-\infty}^{\infty} TP_t^{\mathbb{I}}(c) f(M_k = c \mid T_k > t) dc \\
 &= \int_{-\infty}^{\infty} \left[\int_c^{\infty} f(s \mid T_j = t) ds \right] f(c \mid T_k > t) dc \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(s > c) f(s \mid T_j = t) f(c \mid T_k > t) ds dc \\
 &= P(M_j > M_k \mid T_j = t, T_k > t)
 \end{aligned}$$

其餘 C/D 及 I/D 的定義方式，只需將推導過程，以對應的定義代入即可。 $AUC(t)$ 的估計採用的是梯形法，計算方式為：

$$AUC(t) = \sum_{k=1}^D \frac{1}{2} (ROC_t(p_k) + ROC_t(p_{k-1})) (p_k - p_{k-1})$$

其中 $0 = p_0 \leq p_1 \leq \dots \leq p_D \leq 1$ 為 $\widehat{FP}_t(c)$ 取反函數後對應得到的切點。在 I/D 的定義之下，可以推導出一致性指標 (the concordance summary)，記為 C^τ ，用來衡量生物指標在時間 $(0, \tau)$ 內的預測能力，其中假設不同個體的生物指標與存活時間 (M_j, T_j) 和 (M_k, T_k) 為獨立。

詳細推導過程如下：

$$\begin{aligned}
C^\tau &= P(M_j > M_k \mid T_j < T_k, T_j < \tau) \\
&= P(M_j > M_k, T_j < T_k, T_j < \tau) \frac{2}{2P(T_j < T_k, T_j < \tau)} \\
&= \int_0^\tau P(\{M_j > M_k\} \mid \{T_j = t\} \cap \{t < T_k\}) \times \frac{2P(\{T_j = t\} \cap \{t < T_k\})}{2P(T_j < T_k, T_j < \tau)} dt \\
&= \int_0^\tau AUC(t) \times \frac{2f(t)S(t)}{W^\tau} dt \\
&= E_{T < \tau}(AUC(T) \times \frac{2S(T)}{1 - S^2(\tau)})
\end{aligned}$$

其中 $W^\tau = 2P(T_j < T_k, T_j < \tau) = \int_0^\tau 2f(t)S(t)dt = 1 - S^2(\tau)$