

東海大學統計碩士班

碩士論文

指導教授：黃連成博士

設限資料參數模型的適合度檢定
之模擬研究

Simulation Study on Goodness of Fit Test
for Parametric Model of Censored data

研究生：潘晨綱

中華民國 100 年 6 月

摘要

設限資料建構的參數模型適合度檢定 (goodness of fit test), 通常是在實驗結束之後, 並且以估計的參數建構檢定統計量, 而且已經出現在許多的文章當中。本文中的設限資料, 考慮不同的進入時間 (entry time), 而此設限資料會隨著觀測時間點的不同而增加減少, 因為當觀測的時間越久則所觀察得的資料越多, 反之則越少, 則此模型中的參數估計會與日程時間 t 的資料有關, 所以在日程時間 t 時, 將以最大概似估計 (maximum likelihood estimate) 來估計參數模型中的參數, 而因為資料的蒐集會隨著時間改變, 因此 Hwang (2002) 直觀的考慮一個隨機過程 (stochastic process), 並且利用它會弱收斂 (weak convergence) 到一個漸進高斯平賭 (Gaussian martingale) 過程的性質, 然後以一個成組序列檢定 (group sequential test) 的方法來做適合度檢定, 而成組序列檢定方法的臨界值是利用 Pocock (1977)、O'Brien 和 Fleming (1979) 以及 Lan 和 DeMets (1983) 三種方法來建構。本文主要的目的是使用模擬的方式來驗證 Hwang (2002) 的結果。

目錄

第 1 章	緒論	3
第 1.1 節	背景及研究動機	3
第 1.2 節	論文架構	5
第 2 章	設限資料參數模型與適合度檢定	6
第 2.1 節	設限資料參數模型	6
第 2.2 節	適合度檢定	9
第 2.3 節	成組序列方法	12
第 3 章	模擬過程與結果	14
第 3.1 節	模擬過程	14
第 3.2 節	模擬結果	15
第 4 章	結論	20
	參考文獻	22

第1章 緒論

第 1.1 節 背景及研究動機

存活的設限資料經常出現在各種不同的領域，像是在生物統計學、可靠度、生活史的資料、人口統計學和精算統計等領域，而以設限的存活資料來建構模型更是十分常見，對於建構這樣的模型有很多的統計學家提出方法來檢測，像是 Arjas (1988)對於以存活資料來建構 Cox 迴歸模型的適合度檢定 (goodness of fit)，提出一個易於使用的圖解方法。Akritas (1988)認為卡方檢定是被提議作為當資料可能或不可能以隨機設限為條件來檢定適合度，並提出在設限的情況下作檢定，需要假設設限分配沒有限制並且統計量具有相同簡單的形式。Hjort (1990)對檢查參數模型的正確性，比較累積風險率 (cumulative hazard rate) 和隨著時間的改變估計參數累積風險率 (parametric cumulative hazard rate)，使用計數過程 (counting processes) 來建構，而得到的理論被應用到適合度檢定的卡方類型的統計量。Lin 和 Wei (1991)藉由 White (1982)延伸信息矩陣 (information matrix) 的檢定，提出設定誤差到部分概似來檢測參數模型，並說明 Cox 模型的適合度檢定是基於一些漸進的結果來建構。而 McKeague 和 Utikal (1991) 對 Aalen 和 Cox 模型比較兩個模型在一個相等的立足點上最佳配適完全非參數模型 (fully nonparametric model) 來制定適合度檢定的形式。Myles 和 Edsel (1992)提出一個與 Karl Pearson 在隨機設限模型下的卡方適合度檢定相似的程序，並且與 Akritas (1988) 的方法做比較。Edsel (1998)對於 Cox 比例風險模型，考慮存在不完整的資料，並且以適合度檢定基線風險率 (baseline hazard rate) 函數是否等於一些具體指定的風險率函數。Li 和 Sun (2000)提出一個基於模擬的最小上界型檢定 (supremum-type tests)，檢定的提

出是在 Aalen 的乘性強度記數過程模型的架構下, 因此適用於一些存活模型, 包括那些通用的設限和截斷形式。David (2004)研究的適合度檢定是基於在分層的非殘餘過程, 並推導漸進分配以及建構 Kolmogorov-Smirnov 型檢定。Ma 和 Edsel (2005)對於檢定的漸近屬性, 是在一個臨近的序列對於局部選擇下從理論上研究, 運用到一個一般性的復發事件模型。Guosheng (2009)以右設限的存活資料對 Pearson 卡方適合度檢定提出一個貝氏 (Bayesian) 計算和推論方法, 並以馬可夫鏈蒙地卡羅 (Markov chain Monte Carlo) 的過程來對模型中的參數產生事後樣本, 最後以模擬分析研究來檢查提出的模型診斷方法。所有這些檢測模型的適合度檢定都是發生在實驗結束之後, 並以參數估計來建構。在我們的設限資料中, 考慮進入的時間並不是固定的, 而以設限資料來建構參數模型, 因為資料的蒐集會隨著觀測時間點的不同而有所增減, 因此若以在日程時間 t 時所觀測到的資料以最大概似估計來估計參數, 則可以知道模型中的參數會與日程時間 t 的資料有關, 而這樣隨著時間的改變來觀測資料可以算是一個隨機過程, 而隨機過程會漸進至一個高斯平賭過程。

因為參數估計的過程與時間 t 相依, 所以我們可以建構一組序列檢定, 並以成組序列方法來做為檢定的依據, 當中包含了 Pocock (1977)、O'Brien 和 Fleming (1979) 以及 Lan 和 DeMets (1983)的顯著水準花費函數 (α -spending function) 三種方法。

第 1.2 節 論文架構

本文主要介紹在不同進入時間的情況下，以存活的設限資料來建構參數模型，並且建構參數模型的適合度檢定，然後以成組序列方法來做檢定，最後以電腦模擬的方式來驗證理論假設，本文分成四章來討論。

第1章，簡單介紹設限資料模型的適合度檢定的背景與現況，並提及相關文獻。第1.2節敘述本文的架構。

第2章的第2.1節介紹設限資料參數模型，以及在不同設限時間或者存活時間可以觀設到的資料資訊，並介紹參數估計的過程。第2.2節介紹適合度檢定的定理與統計量。第2.3節介紹 Pocock (1977) 和 O'Brien 和 Fleming (1979) 的方法以及 Lan 和 DeMets (1983) 提出的顯著水準花費函數 (α -spending function)。

第3章的第3.1節是在說明模擬的過程以及資料的生成方式，並且驗證第2章所提出的適合度檢定的方法。第3.2節是在說明模擬的結果。

第4章，總結以上程式模擬的結果與結論。

第2章 設限資料參數模型與適合度檢定

第 2.1 節 設限資料參數模型

對於分析設限的存活資料，我們可以假設存活時間 (survival time) 的分配服從某個參數模型，那就可以確立參數模型的形式和參數的個數，因此只要估計模型中的參數，就能得知模型的性質，然而隨著實驗的進行，時間也在改變，因此資料資訊隨著我們觀測的時間點不同，得到的資訊也不同，因此估計的參數也會不同，也就是說估計的參數與日程時間 t 相依。

在這裡我們病思考慮進入實驗的時間是不固定的，因此我們令 $\{(Y_j, Z_j, X_j, C_j), j = 1, \dots, n\}$ 是一組包含 n 個獨立的四維隨機向量的隨機樣本，其中令 Y_j 表示第 j 個病患的進入時間， Z_j 表示第 j 個病患的影響因子， X_j 表示第 j 個病患的存活時間，而 C_j 表示第 j 個病患的設限時間 (censoring time)；假設第 j 個病患的進入時間和存活時間獨立，且給定進入時間和解釋變數時，存活時間和設限時間也是獨立，我們將假設存活時間的風險函數 (hazard function) 以下式表示：

$$\lambda(\cdot, Z_j, \theta) \quad (A)$$

其中 θ 為 p 維向量的模型參數。

依照我們在日程時間 t 所可能觀察到的情況，分成兩部分來做討論，第一部分為設限時間發生於存活時間之後 ($Y_j + X_j \leq Y_j + C_j$)，第二部分為設限時間發生於存活時間之前 ($Y_j + X_j > Y_j + C_j$)。第一部份，從圖 2.1-A 可知，在 $s < Y_j$ 的情況下，我們觀察到時間 s ，也就是沒有觀察到任何資訊。從圖 2.1-B，觀察的時間為 ($Y_j < s < Y_j + X_j$)，則我們可以觀測到進入的時間 Y_j ，若 ($Y_j + X_j < s < Y_j + C_j$)，則可以觀測得到進入的時間 Y_j 以及資料進入後的存活的時間 $Y_j + X_j$ ，

如圖 2.1-C。而若觀測到的時間已在設限之後，則我們可以觀測得所有資訊，如圖 2.1-D。

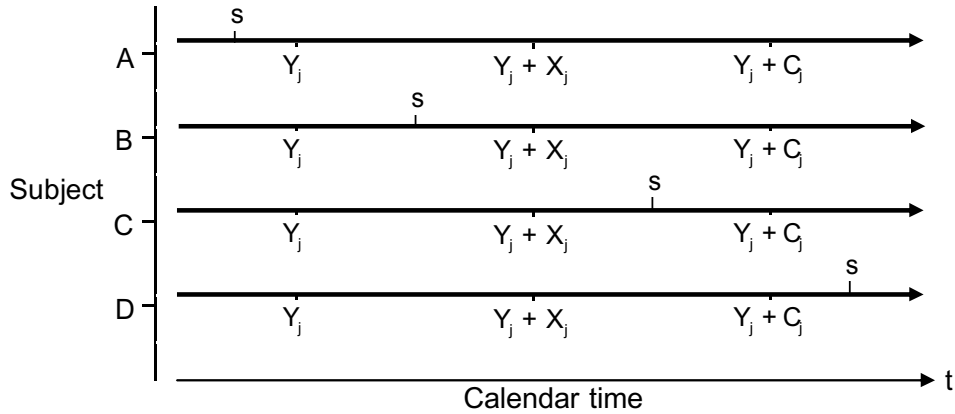


圖 2.1: $(Y_j + X_j) \leq (Y_j + C_j)$

若在 $(Y_j + X_j > Y_j + C_j)$ 的情況下，若觀測到的時間在進入時間之前，則只觀測到時間 s ，如圖 2.2-E，也就是沒有觀測到任何資訊，倘若觀測到時間 s 為 $Y_j < s < Y_j + C_j$ ，那就只能觀測到進入的時間 Y_j ，如圖 2.2-F。而如果觀測時間是在設限之後之後，那所就可以觀測得資料進入的時間 Y_j 以及設限的時間 $Y_j + C_j$ ，如圖 2.2-G，但若是在設限之後才發生的事件則無法觀測得，如圖 2.2-H。

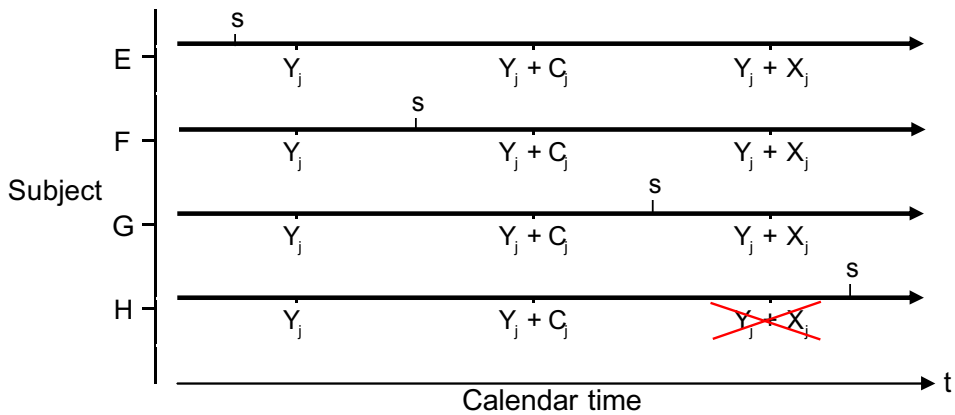


圖 2.2: $(Y_j + X_j) > (Y_j + C_j)$

我們將在日程時間 t 可能觀測到的情況以下式表示:

$$\{Y_j \wedge s, Z_j, (Y_j + X_j \wedge C_j) \wedge s, 1_{[(Y_j + X_j) \wedge s \leq (Y_j + C_j) \wedge s]} | s \leq t, j = 1, \dots, n\},$$

當中 $1_{[(Y_j + X_j) \wedge s \leq (Y_j + C_j) \wedge s]}$ 為一指標函數表示只要在資料進入之後的存活時間 $(Y_j + X_j)$ 可以觀測到且在 $(Y_j + X_j) \leq (Y_j + C_j)$ 的情況下, 則計為 1, 反之只要在 $(Y_j + X_j) > (Y_j + C_j)$ 的情況下, 則恆為零。令 \mathcal{F}_t^n 表示在日程時間 t 時, 所觀測到所有資訊的 σ -field, 且

$$M_j(t, \theta) = 1_{[Y_j, \infty)}(t \wedge (Y_j + C_j)) - \int_0^t \lambda(s - Y_j, Z_j, \theta) 1_{(Y_j, Y_j + X_j \wedge C_j]}(s) ds,$$

而 $M_j(t, \theta)$ 是 \mathcal{F}_t^n 的平賭過程 (martingale)。根據 Chang 和 Hsiung (1988), 可知在日程時間 t 時的對數概似 (log-likelihood) 為

$$L_n(t, \theta) = \sum_{j=1}^n \int_0^t \log \lambda(u, Z_j, \theta) dN_j^t(u) - \sum_{j=1}^n \int_0^t \lambda(u, Z_j, \theta) H_j^t(u) du,$$

當中 $N_j^t(u) = 1_{[X_j, \infty)}(u \wedge C_j \wedge (t - Y_j)^+)$ 且 $H_j^t(u) = 1_{(0, X_j \wedge C_j \wedge (t - Y_j)^+]}(u)$, 由上式可得概似分數 (likelihood score) 過程為

$$\begin{aligned} U_{n,l}(t, \theta) &= \frac{\partial}{\partial \theta_l} L_n(t, \theta) \\ &= \sum_{j=1}^n \int_0^t \frac{\frac{\partial \lambda}{\partial \theta_l}(u - Y_j, Z_j, \theta)}{\lambda(u - Y_j, Z_j, \theta)} 1_{(Y_j, \infty)}(u) dM_j(u, \theta), \end{aligned}$$

當中 $l = 1, \dots, p$, 可以將其完整的表示為 $U_n(t, \theta) = (U_{n,1}(t, \theta), \dots, U_{n,p}(t, \theta))$,

並以最大概似估計求得在日程時間 t 時的參數解 $\hat{\theta}_n(t)$, 滿足 $U_n(t, \hat{\theta}_n(t)) = 0$ 。

由於已經知道 $M_j(t, \theta)$ 是 \mathcal{F}_t^n 的平賭過程 (martingale), 因此本文中直觀的考慮隨機過程

$$H_n(t) = \frac{1}{\sqrt{n}} \sum_{j=1}^n M_j(t, \hat{\theta}_n(t)),$$

並利用它來建構一個成組序列檢定。

第 2.2 節 適合度檢定

令 \mathcal{Z} 是 Z_1 的值域 (range)、 Θ 是實際參數值 θ_0 的鄰域 (neighborhood), 且假設 $t_0 > 0$, 而在前文當中提到的風險函數 $\lambda(\cdot, \cdot, \cdot)$ 具有大於 0 的下限, 對於 θ 的第一次、第二次及第三次偏微在 $(0, t_0] \times \mathcal{Z} \times \Theta$ 是有界。令

$$\phi_{p+1,p+1}(t, \theta_0) = E\{\lambda(t - Y_1, Z_1, \theta_0) 1_{(Y_1, Y_1+X_1 \wedge C_1]}(t)\}$$

$$\phi_{l,p+1}(t, \theta_0) = \phi_{p+1,l}(t, \theta_0) = E\left\{\frac{\partial \lambda}{\partial \theta_l}(t - Y_1, Z_1, \theta_0) 1_{(Y_1, Y_1+X_1 \wedge C_1]}(t)\right\}$$

$$\phi_{l,k}(t, \theta_0) = E\left\{\frac{\frac{\partial \lambda}{\partial \theta_l}(t - Y_1, Z_1, \theta_0) \frac{\partial \lambda}{\partial \theta_k}(t - Y_1, Z_1, \theta_0)}{\lambda(t - Y_1, Z_1, \theta_0)} 1_{(Y_1, Y_1+X_1 \wedge C_1]}(t)\right\},$$

當中 $l, k = 1, \dots, p$, 而 $\Delta_{1l}(t, \theta_0) = \int_0^t \phi_{l,p+1}(u, \theta_0) du$ 是為 $1 \times p$ 的矩陣 $\Delta_1(t, \theta_0)$ 的第 l 個元素, $\Delta_{2lk}(t, \theta_0) = \int_0^t \phi_{l,k}(u, \theta_0) du$ 則表示一個 $p \times p$ 的矩陣 $\Delta_2(t, \theta_0)$ 的第 (l, k) 個元素。

下面的定理 1 及定理 2 的結果可參考 Hwang (2002) 的技術報告:

定理 1. 在模型(A)的假設下, 且 $a > 0$, 若 $\Delta_2(a, \theta_0)$ 為一可逆矩陣, 則隨機過程 $H_n(t)$ 在 $D[a, \infty)$ 上會弱收斂 (converge weakly), 至一個連續型的高斯平賭 (Gaussian martingale) 過程, 而其變異數為

$$g(t, \theta_0) = -\Delta_1(t, \theta_0) \Delta_2^{-1}(t, \theta_0) \Delta_1'(t, \theta_0) + \int_0^t \phi_{p+1,p+1}(u, \theta_0) du.$$

這個定理說明, 如果 H_n 與 0 有顯著的差異則拒絕以模型 (A) 來建構此設限資料, 而因為變異數 $g(t, \theta_0)$ 與未知參數 θ_0 相關, 因此我們需要估計 $g(t, \theta_0)$, 來建構適合度檢定。

爲了估計 $g(t, \theta_0)$, 所以我們令

$$\begin{aligned}\hat{\Delta}_{1l}(t, \hat{\theta}_n(t)) &= \frac{1}{n} \sum_{j=1}^n \int_0^t \frac{\frac{\partial \lambda}{\partial \theta_l}(u - Y_j, Z_j, \hat{\theta}_n(t))}{\lambda(u - Y_j, Z_j, \hat{\theta}_n(t))} d1_{[Y_j+X_j, \infty)}(u \wedge (Y_j + C_j)) \\ \hat{\Delta}_{2lk}(t, \hat{\theta}_n(t)) &= \frac{1}{n} \sum_{j=1}^n \int_0^t \frac{\frac{\partial \lambda}{\partial \theta_l}(u - Y_j, Z_j, \hat{\theta}_n(t)) \frac{\partial \lambda}{\partial \theta_k}(u - Y_j, Z_j, \hat{\theta}_n(t))}{\lambda^2(u - Y_j, Z_j, \hat{\theta}_n(t))} \\ &\quad d1_{[Y_j+X_j, \infty)}(u \wedge (Y_j + C_j)),\end{aligned}$$

當中 $l, k = 1, \dots, p$, 且 $\hat{\Delta}_{1l}(t, \hat{\theta}_n(t))$ 表示為在一個 $1 \times p$ 的矩陣 $\hat{\Delta}_1(t, \hat{\theta}_n(t))$ 的第 l 個元素, 而 $\hat{\Delta}_{2lk}(t, \hat{\theta}_n(t))$ 則是表示為在一個 $p \times p$ 的矩陣 $\hat{\Delta}_2(t, \hat{\theta}_n(t))$ 的第 (l, k) 個元素。

定理 2. 假設與定理1相同的條件下, 且 $t \geq a$, 則 $\hat{\Delta}_{1l}(t, \hat{\theta}_n(t))$ 、 $\hat{\Delta}_{2lk}(t, \hat{\theta}_n(t))$ 和 $\frac{1}{n} \sum_{j=1}^n \int_0^t d1_{[Y_j+X_j, \infty)}(u \wedge (Y_j + C_j))$ 分別為 $\Delta_{1l}(t, \theta_0)$ 、 $\Delta_{2lk}(t, \theta_0)$ 和 $\int_0^t \phi_{p+1, p+1}(u, \theta_0) du$ 的一致性估計量。

因此我們可以令

$$\begin{aligned}\hat{g}(t, \hat{\theta}_n(t)) &= -\hat{\Delta}_1(t, \hat{\theta}_n(t)) \hat{\Delta}_2^{-1}(t, \hat{\theta}_n(t)) \hat{\Delta}_1'(t, \hat{\theta}_n(t)) \\ &\quad + \frac{1}{n} \sum_{j=1}^n \int_0^t d1_{[Y_j+X_j, \infty)}(u \wedge (Y_j + C_j)),\end{aligned}$$

為變異數 $g(t, \theta_0)$ 的一致性估計量, 當中的 $t \geq a$, 那我們就可以定義檢定統計量

$$T_m = \sum_{i=1}^m \frac{(H_n(t_i) - H_n(t_{i-1}))^2}{\hat{g}(t_i, \hat{\theta}_n(t_i)) - \hat{g}(t_{i-1}, \hat{\theta}_n(t_{i-1}))},$$

當中 $0 = t_0 < a \leq t_1 < t_2 < \dots < t_m < \infty$, 且 $\hat{g}(0, \hat{\theta}_n(0)) = 0$, 而且在模型 (A) 的假設下, T_m 的極限分配服從自由度 m 的卡方分配。

對於上式 T_m , 若我們觀測的時間在 t_m , 則為了計算檢定統計量仍必須要回到前面的 t_1, \dots, t_{m-1} 計算其 H_n 與估計 $\hat{g}(t_i, \hat{\theta}_n(t_i))$, 因此我們利用隨機過程 $H_n(t)$ 的特性來建構一個成組序列檢定, 這樣就不需要再回溯前面的

t_1, \dots, t_{m-1} , 倘若

$$\left| \frac{H_n(\hat{g}^{-1}(t_i, \hat{\theta}_n(t_i)))}{\sqrt{t_i}} \right| > c_i$$

則拒絕模型 (A) 的假設, 當中

$$\hat{g}^{-1}(t_i, \hat{\theta}_n(t_i)) = \inf\{s : \hat{g}(s, \hat{\theta}_n(s)) > t_i\},$$

$i = 1, \dots, K$, 臨界值 c_1, \dots, c_K 是由服從 $N(0, 1)$ 的過程 $\{\frac{B(t_i)}{\sqrt{t_i}} : i = 1, \dots, K\}$ 挑選出來, 而且 $t_i = \frac{i}{K}$, 且 $B(t)$ 是一個標準布朗運動, 且 $0 \leq t \leq 1$, 因此 $(\frac{B(t_1)}{\sqrt{t_1}}, \dots, \frac{B(t_K)}{\sqrt{t_K}})$ 服從多變量常態分配 $N_K((0, \dots, 0), \sigma_{ij})$, 假如 $i < j$ 時

$$\begin{aligned} \sigma_{ij} &= Cov\left(\frac{B(t_i)}{\sqrt{t_i}}, \frac{B(t_j)}{\sqrt{t_j}}\right) = E\left(\frac{B(t_i)B(t_j)}{\sqrt{t_i t_j}}\right) - E\left(\frac{B(t_i)}{\sqrt{t_i}}\right)E\left(\frac{B(t_j)}{\sqrt{t_j}}\right) \\ &= \frac{1}{\sqrt{t_i t_j}} E(B(t_i)(B(t_j) - B(t_i)) + (B(t_i))^2) = \sqrt{\frac{t_i}{t_j}}, \end{aligned}$$

反之亦然, 所以

$$\sigma_{ij} = \begin{cases} 1 & \text{若 } i = j, \\ \sqrt{\frac{t_i}{t_j}} & i < j, \\ \sqrt{\frac{t_j}{t_i}} & i > j. \end{cases}$$

Pocock (1977) 與 O'Brien 和 Fleming (1979) 以及 Lan 和 DeMets (1983) 則以 $\{\frac{B(t_i)}{\sqrt{t_i}} : i = 1, \dots, K\}$ 來建構成組序列方法, 詳細的內容在第 2.3 節介紹。

第 2.3 節 成組序列方法

常見的成組序列方法為 Pocock (1977)、O'Brien 和 Fleming (1979) 以及 Lan 和 DeMets (1983) 三種方法。Pocock (1977) 用名義上的顯著水準 (nominal significance level) α' 當概念, 當第 k 個檢定階段的 p 值小於 α' , 亦即 $p_k < \alpha'$, 其中 $k = 1, \dots, K$, 則拒絕 H_0 , 表示模型的建構不適合, 反之若第 k 個檢定階段的 p 值不小於 α' , 則繼續檢定至第 $k + 1$ 個階段為止所累積的資料所建構之模型, 其中名目上的顯著水準 (nominal significance level) α' 是根據事先設定的檢定階段個數 K 和總顯著水準 α 所決定, 而且 Pocock (1977) 方法必須滿足每一階段有相同的樣本數或者是時間的區間必須相同的假設。而 O'Brien 和 Fleming (1979) 的方法與 Pocock (1977) 的方法類似, 不同的地方在於 O'Brien 和 Fleming (1979) 的方法會依據卡方統計量求出各個階段的臨界值, 且臨界值會隨著不同的階段而逐漸遞減, 而且 O'Brien 和 Fleming (1979) 的方法不須滿足每一階段有相同的樣本數的假設。Lan 和 DeMets (1983) 則提出顯著水準花費函數 $\alpha^*(t)$, 並且討論三種情況 $\alpha_1^*(t)$ 、 $\alpha_2^*(t)$ 和 $\alpha_3^*(t)$, 以求出各階段的臨界值。

對於顯著水準花費函數, 當中

$$\alpha_1^*(t) = \begin{cases} 0 & \text{若 } t = 0, \\ 2 - 2\Phi(z_{\frac{\alpha}{2}}/\sqrt{t}) & 0 < t \leq 1, \end{cases}$$

$z_{\frac{\alpha}{2}}$ 為顯著水準 $\frac{\alpha}{2}$ 之標準長態分配值, $P(Z > z_{\frac{\alpha}{2}}) = \frac{\alpha}{2}$, $\Phi(z_{\frac{\alpha}{2}}/\sqrt{t})$ 為標準常態分配值是 $z_{\frac{\alpha}{2}}/\sqrt{t}$ 之累積機率。t 的範圍是 $0 \leq t \leq 1$, 為一標準化過的時間, 且 $\alpha_1^*(t)$ 為一嚴格遞增函數, 當觀測尚未開始時 $t = 0$, $\alpha_1^*(0) = 0$, 隨著觀測的時間增加, 累積顯著水準即增加, 至觀測結束時 $t = 1$, $\alpha_1^*(1) = \alpha$, 亦即將顯著水準控制在

α , 如果給定 $t = 0.5$ 和 $\alpha = 0.5$, 則 $\alpha_1^*(0.5) < 0.006$, 表示不可能在早期觀測時即停止, 若在 $t_i = \frac{i}{K}$, 且 $i = 1, \dots, K$ 時, 結果將會近似於 O'Brien 和 Fleming (1979) 的方法。而 $\alpha_2^*(t) = \alpha \log\{1 + (e - 1)t\}$, 如果給定 $t = 0.5$, 則 $\alpha_2^*(0.5) \cong 0.62\alpha$, 亦即在 $t = 0.5$ 的時候我們花費 62% 的型一錯誤機率 α , 表示比較可能在早期觀測時即停止, 若在 $t_i = \frac{i}{K}$, 且 $i = 1, \dots, K$ 時, 結果會與 Pocock (1977) 的臨界值相近。而 $\alpha_3^*(t) = \alpha t$, 表示顯著水準 α 均勻的分佈在所有的時間 t , 且 $\alpha_3^*(t)$ 的結果介於 $\alpha_1^*(t)$ 與 $\alpha_2^*(t)$ 之間。

表 2.1: 臨界值

	$\alpha = 0.05$					$\alpha = 0.10$				
	c_1	c_2	c_3	c_4	c_5	c_1	c_2	c_3	c_4	c_5
O'Brien和 Fleming	4.56	3.23	2.63	2.28	2.04	3.92	2.77	2.26	1.96	1.75
$\alpha_1^*(t)$	4.90	3.35	2.68	2.29	2.03	4.23	2.89	2.30	1.96	1.74
Pocock	2.41	2.41	2.41	2.41	2.41	2.12	2.12	2.12	2.12	2.12
$\alpha_2^*(t)$	2.44	2.43	2.41	2.40	2.39	2.18	2.14	2.11	2.09	2.07
$\alpha_3^*(t)$	2.58	2.49	2.41	2.34	2.28	2.83	2.22	2.12	2.03	1.96

表1參考自 Lan 和 DeMets (1983), 是表示分別在顯著水準 0.05 和 0.1 的情況下, 對於三條顯著水準花費函數及 O'Brien 和 Fleming (1979) 與 Pocock (1977) 所設定五次觀測的臨界值, 可以看出 $\alpha_1^*(t)$ 與 O'Brien 和 Fleming (1979) 較為接近, 而 $\alpha_2^*(t)$ 與 Pocock (1977) 相似, $\alpha_3^*(t)$ 結果介於 $\alpha_1^*(t)$ 與 $\alpha_2^*(t)$ 之間, 且 $\alpha_1^*(t)$ 在前期時, 較 Pocock (1977) 和 O'Brien 和 Fleming (1979) 大, $\alpha_2^*(t)$ 與 $\alpha_3^*(t)$ 則在前期時較 Pocock (1977) 大, 表示水準花費函數需要比 O'Brien 和 Fleming (1979) 與 Pocock (1977) 累積更多的資訊。

第3章 模擬過程與結果

第 3.1 節 模擬過程

本節主要說明如何利用模擬來驗證第二章所提及的方法。首先根據模型 (A) 的假設，令存活時間為一服從指數分配之隨機變數，則風險函數為其參數 λ ，由於存活時間為指數分配，因此我們考慮 $\lambda = \beta_0 + \beta_1 Z$ ，而 Z 為服從某一機率分配之隨機變數， (β_0, β_1) 為一組實際的參數。由於我們考慮不同的進入時間與設限時間，因此我們個別令進入時間和設限時間為服從某一機率分配之隨機變數，或者考慮進入時間為 0，或者不設限。

在進行模擬時，變異數 $g(t, \theta_0)$ 的估計值 $\hat{g}(t, \hat{\theta}_n(t))$ 會出現負值的情況，因此我們對統計檢定做了些微修正，而修正後為

$$\left| \frac{H_n(\hat{g}^{-1}(t_i, \hat{\theta}_n(t_i)))}{\sqrt{t_i}} \right| > c_i,$$

當中 $\hat{g}^{-1}(t_i, \hat{\theta}_n(t_i)) = \inf\{s : \hat{g}(s, \hat{\theta}_n(s)) > t_i\}$ ， s 為連續時間，因此當模擬開始時必須先生成日程時間 t ，然而日程時間 t 並非為 $0 \leq t \leq 1$ ，所以我們考慮 $\{B(t) : 0 \leq t \leq t_0\}$ ，因此 $t_i = \frac{i}{K} t_0$ ， $i = 1, \dots, K$ ，而 $\{\frac{B(t_1)}{\sqrt{t_1}}, \dots, \frac{B(t_K)}{\sqrt{t_K}}\}$ 一樣是服從多變量常態分配 $N_K((0, \dots, 0), \sigma_{ij})$ ，當中 σ_{ij} 與第 2.2 節敘述相同。

根據上述方法，為了找出 $\hat{g}^{-1}(t_i, \hat{\theta}_n(t_i))$ ，先隨機生成一組 $\{Y_j, Z_j, X_j, C_j\}$ ， $j = 1, \dots, 300$ ，然後可以得到 \mathcal{F}_t^n 的平賭過程 $M_j(t, \theta)$ ，再來則可以估得對數概似 $L_n(t, \theta)$ ，以及概似分數 $U_{n,i}(t, \theta)$ ，並且以最大概似估計求得參數解，並且找到 5 個 $g(s, \hat{\theta}_n(s))$ 剛好大於 t_i 的 5 個 s 時間，當中 $i = 1, \dots, 5$ ，並且依序帶入修正後的統計檢定，若 5 個當中只要有一次得到拒絕的結果，則拒絕模型 (A) 的假設，最後重覆以上的步驟 10000 次，並且記錄其拒絕次數。

第 3.2 節 模擬結果

依照上節所敘述的過程做模擬, 得到表 3.1、表 3.2、表 3.3 和表 3.4。表 3.1 是在存活時間服從平均數為 $\frac{1}{\lambda}$ 之指數分配, $\lambda = \beta_0 + \beta_1 Z$, 令 $(\beta_0, \beta_1) = (2, 16)$ 的情況下, 考慮進入的時間服從 $(0, 14)$ 之均勻分配或進入時間為 0, 而設限時間服從平均數為 1 的指數分配或設限時間為 ∞ , 並且重複前文敘述的模擬過程一萬次來觀察拒絕次數。而表 3.2、表 3.3 和表 3.4 則分別假設參數 (β_0, β_1) 為 $(2, 2)$ 、 $(12, 2)$ 和 $(3, 5)$ 的情況下, 進行與上述表 3.1 相同的模擬過程。而在四個表當中也分別假設了不同的影響因子, 表 3.1 假設影響因子服從 $(0, 1)$ 之均勻分配, 或者服從 $(3, 5)$ 之對數常態分配。表 3.2 假設影響因子服從 1,2,3 之離散均勻分配, 或者服從 $(13, 4)$ 之對數常態分配。表 3.3 假設影響因子服從 1,2 之離散均勻分配, 或者服從 $(13, 6)$ 之對數常態分配。最後表 3.4 假設影響因子服從 1,2 之離散均勻分配, 或者服從 $(13, 5)$ 之對數常態分配。

最後我們可以發現當 $t_i = \frac{i}{K}t_0$, $i = 1, \dots, K$ 時, 以 $\alpha_1^*(t)$ 為臨界值的結果會近似於以 O'Brien 和 Fleming 為臨界值, 若以 $\alpha_2^*(t)$ 為臨界值的結果會近似於以 Pocock 為臨界值, 而 $\alpha_3^*(t)$ 的結果介於 $\alpha_1^*(t)$ 與 $\alpha_2^*(t)$ 之間, 而當 t 的設定相同時, 以 Pocock (1977)、 $\alpha_2^*(t)$ 和 $\alpha_3^*(t)$ 為臨界值時的拒絕次數會較以 O'Brien 和 Fleming (1979) 和 $\alpha_1^*(t)$ 為臨界值時高。而且 O'Brien 和 Fleming (1979)、 $\alpha_1^*(t)$ 、Pocock (1977)、 $\alpha_2^*(t)$ 和 $\alpha_3^*(t)$ 為臨界值檢定, 在不同的觀測時間下皆可以讓型一錯誤發生的機率維持在 5% 左右。

表 3.1: 參數 $\beta_0 = 2, \beta_1 = 16$

$\alpha = 0.05$					
	O'Brien 和 Fleming	$\alpha_1^*(t)$	Pocock	$\alpha_2^*(t)$	$\alpha_3^*(t)$
	$t = [0.012 \quad 0.024 \quad 0.036 \quad 0.048 \quad 0.060]$				
$Y_j \sim Uniform(0, 14)$					
$Z_j \sim Uniform(0, 1)$	503	459	1441	1413	1266
$C_j \sim Exp(1)$	$t = [0.02 \quad 0.04 \quad 0.06 \quad 0.08 \quad 0.10]$				
	200	186	596	574	502
	$t = [0.02 \quad 0.04 \quad 0.06 \quad 0.08 \quad 0.10]$				
$Y_j \sim Uniform(0, 14)$					
$Z_j \sim Uniform(0, 1)$	426	380	1470	1433	1284
$C_j \sim \infty$	$t = [0.032 \quad 0.064 \quad 0.096 \quad 0.128 \quad 0.160]$				
	175	165	613	586	499
	$t = [0.22 \quad 0.44 \quad 0.66 \quad 0.88 \quad 1.10]$				
$Y_j \sim 0$					
$Z_j \sim Lognormal(3, 5)$	497	433	6373	6221	5521
$C_j \sim Exp(1)$	$t = [0.54 \quad 1.08 \quad 1.62 \quad 2.16 \quad 2.70]$				
	66	58	644	556	329

表 3.2: 參數 $\beta_0 = 2, \beta_1 = 2$

		$\alpha = 0.05$				
		O'Brien 和 Fleming				
		$\alpha_1^*(t)$	Pocock	$\alpha_2^*(t)$	$\alpha_3^*(t)$	
		$t = [0.01 \quad 0.02 \quad 0.03 \quad 0.04 \quad 0.05]$				
$Y_j \sim Uniform(0, 14)$						
$Z_j \sim Discrete \ uniform(3)$		623	588	1579	1528	
$C_j \sim Exp(1)$		$t = [0.022 \quad 0.044 \quad 0.066 \quad 0.088 \quad 0.110]$				
		55	53	457	426	
		$t = [0.022 \quad 0.044 \quad 0.066 \quad 0.088 \quad 0.110]$				
$Y_j \sim Uniform(0, 14)$						
$Z_j \sim Discrete \ uniform(3)$		322	291	1548	1497	
$C_j \sim \infty$		$t = [0.04 \quad 0.08 \quad 0.12 \quad 0.16 \quad 0.20]$				
		32	32	486	453	
		$t = [0.19 \quad 0.38 \quad 0.57 \quad 0.76 \quad 0.95]$				
$Y_j \sim 0$						
$Z_j \sim Lognormal(13, 4)$		322	317	400	398	
$C_j \sim Exp(1)$					394	

表 3.3: 參數 $\beta_0 = 12, \beta_1 = 2$

		$\alpha = 0.05$				
		O'Brien 和 Fleming				
		$\alpha_1^*(t)$	Pocock	$\alpha_2^*(t)$	$\alpha_3^*(t)$	
		$t = [0.006 \quad 0.012 \quad 0.018 \quad 0.024 \quad 0.030]$				
$Y_j \sim \text{Uniform}(0, 14)$						
$Z_j \sim \text{Discrete uniform}(2)$		328	309	792	776	702
		$t = [0.008 \quad 0.016 \quad 0.024 \quad 0.032 \quad 0.040]$				
$C_j \sim \text{Exp}(1)$		166	155	538	518	458
		$t = [0.007 \quad 0.014 \quad 0.021 \quad 0.028 \quad 0.035]$				
$Y_j \sim \text{Uniform}(0, 14)$						
$Z_j \sim \text{Discrete uniform}(2)$		379	361	899	868	797
		$t = [0.0095 \quad 0.0190 \quad 0.0285 \quad 0.0380 \quad 0.0475]$				
$C_j \sim \infty$		147	136	630	596	511
		$t = [0.05 \quad 0.10 \quad 0.15 \quad 0.20 \quad 0.25]$				
$Y_j \sim 0$						
$Z_j \sim \text{Lognormal}(13, 6)$		671	561	4492	4377	3921
		$t = [0.19 \quad 0.38 \quad 0.57 \quad 0.76 \quad 0.95]$				
$C_j \sim \text{Exp}(2)$		201	181	513	503	439

表 3.4: 參數 $\beta_0 = 3, \beta_1 = 5$

		$\alpha = 0.05$				
		O'Brien				
		和 Fleming	$\alpha_1^*(t)$	Pocock	$\alpha_2^*(t)$	$\alpha_3^*(t)$
		$t = [0.008 \quad 0.016 \quad 0.024 \quad 0.032 \quad 0.040]$				
$Y_j \sim \text{Uniform}(0, 14)$						
$Z_j \sim \text{Discrete uniform}(2)$		532	494	1383	1350	1255
$C_j \sim \text{Exp}(2)$		$t = [0.012 \quad 0.024 \quad 0.036 \quad 0.048 \quad 0.060]$				
		112	100	635	609	510
		$t = [0.009 \quad 0.018 \quad 0.027 \quad 0.036 \quad 0.045]$				
$Y_j \sim \text{Uniform}(0, 14)$						
$Z_j \sim \text{Discrete uniform}(2)$		623	586	1620	1580	1428
$C_j \sim \infty$		$t = [0.019 \quad 0.038 \quad 0.057 \quad 0.076 \quad 0.095]$				
		85	75	577	544	442
		$t = [0.19 \quad 0.38 \quad 0.57 \quad 0.76 \quad 0.95]$				
$Y_j \sim 0$						
$Z_j \sim \text{Lognormal}(13, 5)$		444	433	667	664	644
$C_j \sim \text{Exp}(2)$						

第4章 結論

由第2章的理論與方法和第3章模擬的結果，可以看出 Hwang (2002) 直觀的以隨機過程來建構的適合度檢定，無論考慮進入的時間相同與不同，或者考慮不設限，皆可以獲得良好的結果。而且從原本服從卡方分配的檢定統計量 T_m 需要考慮每個觀測時間點往前推到時間 0 的所有隨機過程，到以成組序列檢定方法的檢定統計量，只需要記錄我們觀測時間點的隨機過程，而不必往回推到所有時間點，然後再去記錄所有時間點的隨機過程，這樣可以省去很多的時間成本，而且更有效率。本文中所討論的模擬都是在存活時間為指數分配的情況下，因此只有考慮一個解釋變數，未來可以再對進入時間、設限時間和解釋變數，給予更多不同的分配，或者是更改存活時間的分配，並增加解釋變數的參數個數，來觀察在拒絕次數上是否會有更好的結果，抑或是在增加解釋變數之後，固定所有給定的分配，然後增加每一次試驗的人數，觀察是否也會有更好的結果。

參考文獻

- Akritis, M. G. (1988), “Peaeson-type goodness-of-fit tests : the univariate case”, *J. Amer. Statist. Assoc.*, 83, 222–230.
- Arjas, E. (1988), “A graphical method for assessing goodness of fit in cox’s proportional hazards model”, *J. Amer. Statist. Assoc.*, 83, 204–212.
- Chang, I. S. and Hsiung, C. A. (1988), “Likelihood process in parametric model of censored data with staggered entry-asymptotic properties and applications”, *J. Multivariate Anal.*, 24, 31–45.
- David, K. (2004), “Goodness-of-fit inference for the coxaalen additivemultiplicative regression model”, *Statistics and Probability Letters*, 70, 285–298.
- Edsel, A. Pena (1998), “Smooth goodness-of-fit tests for the baseline hazard in cox’s proportional hazards model”, *Journal of the American Statistical Association*, 93, 673–692.
- Guosheng, Y. (2009), “Bayesian goodness-of-fittestforcensoreddata”, *Journal of Statistical Planning and Inference*, 139, 1474–1483.
- Hjort, N. L. (1990), “Goodness of fit tests in models for life history data based on cumulative hazard rates”, *Ann. Statist.*, 18, 1221–1258.
- Hwang, L. C. (2002), “A note on goodness-of-fit tests for censored data”, *Technical Report*.
- Lan, K. K. G. and DeMets, D. L. (1983), “Discrete sequential boundaries for clinical trials”, *Biometrika*, 70, 659–663.
- Li, G. and Sun, Y. (2000), “A simulation-based goodness-of-fit test for survival data”, *Statistics and Probability Letters*, 47, 403–410.
- Lin, D. Y. and Wei, L. J. (1991), “Goodness-of-fit tests for the general cox regression model”, *Statist. Sinica.*, 1, 1–17.
- Ma. Zenia N. Agustin, Edsel A. Pena. (2005), “A basis approach to goodness-of-fit testing in recurrent event models”, *Journal of Statistical Planning and Inference*, 133, 285–303.

- McKeague, I. W. and Utikal, K. J. (1991), “Goodness-of-fit tests for additive hazards and proportional hazards models”, *Scand. J. Statist.*, 18, 177–195.
- Myles, H. and Edsel, A. Pena (1992), “A chi-squared goodness-of-fit test for randomly censored data”, *Journal of the American Statistical Association*, 87, 458–463.
- O’Brien, P. C. and Fleming, T. R. (1979), “A multiple testing procedure for clinical trials”, *Biometrics*, 35, 549–556.
- Pocock, S. J. (1977), “Group sequential methods in the design and analysis of clinical trials”, *Biometrika*, 64, 191–199.
- White, H. (1982), “Maximum likelihood estimation of misspecified models”, *Econometrica*, 50, 1–25.