論邏輯迴歸模式的估計

李 隆 安*　　葉 昭 瑛**

提 要

邏輯迴歸模式是一個廣泛應用的模式，現多採用最大概似估計法來估計參數，但這個估計方法會有不收斂的問題出現。文獻上有一些估計方法可避免這個現象，其中之一是判別函數估計法，其有效性Efron (1975)指出比最大概似估計法優越。本文提出改進的判別函數估計法，並用電腦模擬來實證改進的程度。

This paper provides an improved Discrimination Function Approach to logistic regression model. Some computer simulation demonstrates the new methods performance.

一、前　言

邏輯迴歸模式的架構如下：

$$\begin{aligned}\pi(x)&=\Pr(y=1\mid X=x)\\&=\frac{1}{1+e^{x'\beta}}\\\Pr(y=0\mid X=x)&=1-\pi(x)\end{aligned}\tag{1}$$

其中自變量x前面的係數向量β是未知的，需要從觀察到的樣本來估計。

*中央研究院統計科學研究所專任，政治大學應用數學所兼任。
**政治大學應用數學所研究生，現任教於聖功女中。

邏輯迴歸模式中的應變數y只可能有兩種數值$y=0$或$y=1$，所以和一般的迴歸模式的估計問題不盡相同。一般的迴歸模式估計是採用最小平方估計法(Least Squared Estimation)，而邏輯迴歸模式估計多採用最大概似估計法(Maximum Likelihood Estimation)。邏輯迴歸模式的最概似估計法，經過適當的整理後(見McCullagh, Nelder1983)，計算過程類似加權(Weighted) 最小平方估計法，需要多次疊代(interative)步驟來求解。

最大概似估計法的嚴重缺陷是「收斂」問題，極有可能各次疊代出來的數值最後發散，造成無解。在文獻上，另外有兩種估計法可以避免這項「收斂」問題。其中一種Grizzle, Starmer, Koch (1969)提出的非疊代加權最小平方估計法，可惜的是使用這個方法，需要先對$\pi(x)$有估計值，實用上這個方法受到了限制。

另外一個可避免最大概似估計法所遭遇的發散問題的估計法，是由Cornfield (1962)所提出的，稱爲判別函數估計法(Discrimination Function Estimation) 。判別函數估計法假設應變數 $y=0$ 及 $y=1$ 族群(population) 的自變數 x 各自爲一常態分配，則其驗後分配(posterior distribution) 恰爲邏輯迴歸模式(1)。這個方法可直接將樣本數據代入公式後即求出參數的估計值來。

Efron (1975)指出在$y=0$及$y=1$族群分別具有常態分配的架構下，判別函數估計法比最大概似估計法更具有效性。所以從有效性的觀點而言，邏輯迴歸模式應採用判別函數估計法。判別函數估計法既是由$y=0$及$y=1$族群各具常態分配推得而出，這兩個常態分配不一定具有相同的變異數（矩陣），求解的公式也有所不同。

一般在運用判別函數估計法時，先檢定兩個變異數是否相同。若檢定的結果是兩個變異數相同，則採用變異數相同的公式；若非，則採用變異數不同的公式。換言之，在運用判別函數估計法，要多一道程序：檢定變異數是否相同。

實際的日常生活中，$y=0$及$y=1$族群能具有相同變異數的可能性非常的小；即使檢定的結論是接受相同的變異數，仍極有可能是假陽性(false positive) 。Lee, Gurland(1975)， Lee, Fineberg(1991)，及Lee (1992) 等在比較兩個常態樣本時特別指出上述現象，並建議一律採用不同變異數的立場來處理，認爲這是較爲合理的方法。他們發現縱使在真正是相同變異數的狀況下，結果也不會太差。

本文主要的研究目的，就是探討一律採用不同變異數的立場來處理判別函數估計法的優劣。下一節中敘述邏輯迴歸模式的一般判別函數估計法，而於第三節提出改進的判別函數估計法。電腦模擬的比較及結果陳述在第四節內，最後一節是本文的結論。

二、判別函數估計法

這節敘述根據判別函數估計法的驗後分配，導出邏輯迴歸模式，比較彼此間的係數關係，得出所需的參數估計。先分別令$\bar{x}_j$，s_j^2，n_j，表示$y=j$的樣本平均值，樣本變異數（矩陣），及樣本大小，$\forall_j=0$，1，又令$n=n_0+n_1$。由兩母體變異數的相等與否，有以下情況產生：

㈠在單變量的情況下：

(A)考慮二母體變異數相同時，

已知$x_j \mid y=j \sim N(\mu j, \sigma^2)$，$\theta_j=P_r(y=j)$；$\forall_j=0$，1，則$y=1$時之條件分配函數爲

$$
\begin{aligned}
P_r(y=1 \mid X=x) &= \frac{P_r(y=1, X=x)}{P_r(X=x)} \\
&= \frac{P_r(x \mid y=1)P_r(y=1)}{P_r(x \mid y=0)P_r(y=0)+P_r(x \mid y=1)P_r(y=1)} \\
&= \frac{\theta_1 \cdot \frac{1}{\sqrt{2\pi\sigma}}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma^2}\right\}}{\theta_0 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma^2}\right\}+\theta_1 \cdot \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma^2}\right\}} \\
&= \frac{\exp\left\{(\ln\frac{\theta_1}{\theta_0}-\frac{\mu_1^2-\mu_0^2}{2\sigma^2})+\frac{(\mu_1-\mu_0)}{\sigma^2}x\right\}}{1+\exp\left\{(\ln\frac{\theta_1}{\theta_0}-\frac{\mu_1^2-\mu_0^2}{2\sigma^2})+\frac{(\mu_1-\mu_0)}{\sigma^2}x\right\}} \\
&= \frac{\exp\{\beta_0+\beta_1 x\}}{1+\exp\{\beta_0+\beta_1 x\}} \\
&= \pi(x)
\end{aligned}
$$

經邏輯轉換，$ln\frac{\pi(x)}{1-\pi(x)} = \beta_0+\beta_1 x$ 爲線性函數，其中

$$\beta_0 = ln\frac{\theta_1}{\theta_0} - \frac{(\mu_1^2-\mu_0^2)}{2\sigma^2} \tag{2}$$

$$\beta_1 = \frac{\mu_1-\mu_0)}{\sigma^2} \tag{3}$$

欲估計β_0，β_1則以$\hat{\mu}_j=\bar{x}_j$，$\hat{\theta}_j=\frac{n_j}{n}$，$\hat{\sigma}^2=\frac{(n_0-1)s_0^2+(n_1-1)s_1^2}{n_0+n_1-2}$
代入(2)，(3)之μ_j，θ_j，σ^2；$\forall_j = 0$，1，可得參數之估計值爲

$$\hat{\beta}_0 = ln\frac{n_1}{n_0} - \frac{1}{2}(\frac{x_1^2-x_0^2}{\hat{\sigma}^2}) \tag{4}$$

$$\hat{\beta}_1 = \frac{x_1-x_0}{\hat{\sigma}^2} \tag{5}$$

(B)考慮二母體變異數不相同時，

已知$x_j \mid y=j \sim N(\mu_j,\sigma_j^2)$，$\theta_j=P_r(y=j)$；$\forall_j=0$，1，則$y=1$ 時之條件分配函數爲

$$P_r(y=1 \mid X=x)=\frac{P_r(y=1,X=x)}{P_r(X=x)}$$

$$=\frac{P_r(x \mid y=1)P_r(y=1)}{P_r(x \mid y=0)P_r(y=0)+P_r(x \mid y=1)P_r(y=1)}$$

$$=\frac{\theta_1 \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}}{\theta_0 \cdot \frac{1}{\sqrt{2\pi\sigma_0^2}}\exp\left\{-\frac{(x-\mu_0)^2}{2\sigma_0^2}\right\}+\theta_1 \cdot \frac{1}{\sqrt{2\pi\sigma_1^2}}\exp\left\{-\frac{(x-\mu_1)^2}{2\sigma_1^2}\right\}}$$

$$=\frac{\exp\left\{(\ln\frac{\theta_1}{\theta_0} \cdot \frac{\sigma_0}{\sigma_1}+\frac{1}{2}(\frac{\mu_0^2}{\sigma_0^2}-\frac{\mu_1^2}{\sigma_1^2})+(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_0}{\sigma_0^2})x-\frac{1}{2}(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2})x^2\right\}}{1+\exp\left\{(\ln\frac{\theta_1}{\theta_0} \cdot \frac{\sigma_0}{\sigma_1}+\frac{1}{2}(\frac{\mu_0^2}{\sigma_0^2}-\frac{\mu_1^2}{\sigma_1^2})+(\frac{\mu_1}{\sigma_1^2}-\frac{\mu_0}{\sigma_0^2})x-\frac{1}{2}(\frac{1}{\sigma_1^2}-\frac{1}{\sigma_0^2})x^2\right\}}$$

$$=\frac{e^{\beta_0+\beta_1 X+\beta_2 X^2}}{1+e^{\beta_0+\beta_1 X}}$$

$$=\pi(x)$$

因此，參數之正確值爲

$$\beta_0 = ln\frac{\theta_1}{\theta_0} \cdot \frac{\sigma_0}{\sigma_1} - \frac{1}{2}(\frac{\mu_1^2}{\sigma_1^2} - \frac{\mu_0^2}{\sigma_0^2}) \tag{6}$$

$$\beta_1 = \frac{\mu_1}{\sigma_1^2} - \frac{\mu_0}{\sigma_0^2} \tag{7}$$

$$\beta_2 = -\frac{1}{2}(\frac{1}{\sigma_1^2} - \frac{1}{\sigma_0^2}) \tag{8}$$

欲估計β_0，β_1，β_2 則以$\hat{\mu}_j = \bar{x}_j$，$\hat{\theta}_j = \frac{n_j}{n}$，$\hat{\sigma}_j^2 = s_j^2$；$\forall_j = 0,1$，代入(6)～(8)之$\mu_j$，$\theta_j$，$\sigma_j^2$；$\forall_j = 0,1$，可得參數之估計值爲

$$\hat{\beta}_0^* = ln\frac{n_1}{n_0} \cdot \frac{s_0}{s_1} - \frac{1}{2}(\frac{\bar{x}_1^2}{s_1^2} - \frac{\bar{x}_0^2}{s_0^2}) \tag{9}$$

$$\hat{\beta}_1^* = \frac{\bar{x}_1}{s_1^2} - \frac{\bar{x}_0}{s_0^2} \tag{10}$$

$$\hat{\beta}_2^* = -\frac{1}{2}(\frac{1}{s_1^2} - \frac{1}{s_0^2}) \tag{11}$$

(二)在多變量的情況下：

$X' = (X_1, X_2, \cdots, X_p)$，邏輯迴歸模式爲

$$\pi(x) = \frac{e^{x'\beta}}{1+e^{x'\beta}}$$

邏輯轉換爲

$$ln\frac{\pi(x)}{1-\pi(x)} = \beta_0 + \beta_1 x_1 + \cdots\cdots + \beta_p x_p = x'\beta$$

(A)考慮二母體變異數相同時，

已知$X_j|y=j \sim N(\mu_j, \Sigma)$，$\theta_j = P_r(y=j)$；$\forall_j = 0, 1$，其中$X_j' = (X_{j1}, \cdots, X_{jp})$爲P個獨立變數，$\mu_j' = (\mu_{j1}, \cdots, \mu_{jp})$爲P個獨立變數之平均值，Σ爲此P個獨立變數之P×P維共變異矩

陣，則

$$P_r(y=1 \mid X=x)=\frac{P_r(y=1,X=x)}{P_r(X=x)}$$

$$=\frac{P_r(x \mid y=1)P_r(y=1)}{P_r(x \mid y=0)P_r(y=0)+P_r(x \mid y=1)P_r(y=1)}$$

$$=\frac{Pr(\mathrm{x}|y=1)Pr(y=1)}{Pr(\mathrm{x}|y=0)Pr(y=0)+Pr(\mathrm{x}|y=1)Pr(y=1)}$$

$$=\frac{\theta_1\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(\mathrm{x}-\mu_1)'\Sigma^{-1}(\mathrm{x}-\mu_1)}}{\theta_0\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(\mathrm{x}-\mu_0)'\Sigma^{-1}(\mathrm{x}-\mu_0)}+\theta_1\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma|^{-\frac{1}{2}}e^{-\frac{1}{2}(\mathrm{x}-\mu_1)'\Sigma^{-1}(\mathrm{x}-\mu_1)}}$$

$$=\frac{exp[(ln\frac{\theta_1}{\theta_0}-\frac{1}{2}(\mu_1-\mu_0)'\Sigma^{-1}(\mu_1+\mu_0))+(\mu_1-\mu_0)'\Sigma^{-1}\mathrm{x}]}{1+exp[(ln\frac{\theta_1}{\theta_0}-\frac{1}{2}(\mu_1-\mu_0)'\Sigma^{-1}(\mu_1+\mu_0))+(\mu_1-\mu_0)'\Sigma^{-1}\mathrm{x}]}$$

$$=\pi(x)$$

因此，參數之正確值爲

$$\beta_0=ln\frac{\theta_1}{\theta_0}-\frac{1}{2}(\mu_1-\mu_0)'\Sigma^{-1}(\mu_1+\mu_0) \tag{12}$$

$$\beta_1=(\mu_1-\mu_0)'\Sigma^{-1} \tag{13}$$

欲估計β_0，β_1則以$\hat{\mu}_j=\bar{x}_j=\frac{1}{n_j}\sum_{i=1}^{n_j}x_{ji}$，$\hat{\theta}_j=\frac{n_j}{n}$，$\hat{\Sigma}=\frac{(n_0-1)s_0^2+(n_1-1)s_1^2}{n_0+n_1-2}$，代入(12)，(13)之$\mu_j$，$\theta_j$，$\Sigma$；$_j=0,1$，其中$s_j=\frac{1}{n_j-1}\Sigma_{i=1}(x_{ji}-\bar{x}_j)(x_{ji}-\bar{x}_j)'$爲$x_j$之P×P維共變異矩陣的不偏推定。

(B)考慮二母體變異數不相同時，

已知$X_j \mid y=j \sim N(\mu_j,\Sigma)$，$\theta_j=P_r(y=j)$；$_j=0$，1，其中$X'_j=(X_{j1},\cdots,X_{jp})$，$\mu'_j=(\mu_{j1},\cdots,\mu_{jp})$，$\Sigma_j$爲二母體各別之P×P維共變異矩陣；$_j=0$，1

則

$$P_r(y=1 \mid X=x)=\frac{P_r(y=1,X=x)}{P_r(X=x)}$$

$$=\frac{Pr(x|y=1)Pr(y=1)}{Pr(x|y=0)Pr(y=0)+Pr(x|y=1)Pr(y=1)}$$

$$=\frac{\theta_1\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma_1|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu_1)'\Sigma^{-1}(x-\mu_1)}}{\theta_0\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma_0|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu_0)'\Sigma_0^{-1}(x-\mu_0)}+\theta_1\cdot(\frac{1}{\sqrt{2\pi}})^p|\Sigma_1|^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu_1)'\Sigma_1^{-1}(x-\mu_1)}}$$

$$=\frac{exp[ln(\frac{\theta_1}{\theta_0}|\Sigma_1|^{-\frac{1}{2}}|\Sigma_0|^{\frac{1}{2}})-\frac{1}{2}(\mu_1'\Sigma_1^{-1}\mu_1-\mu_0'\Sigma_0^{-1}\mu_0)}{1+exp[ln(\frac{\theta_1}{\theta_0}|\Sigma_1|^{-\frac{1}{2}}|\Sigma_0|^{\frac{1}{2}})-\frac{1}{2}(\mu_1'\Sigma_1^{-1}\mu_1-\mu_0'\Sigma_0^{-1}\mu_0)}$$

$$\frac{-(\mu_0'\Sigma_0^{-1}-\mu_1'\Sigma_1^{-1})x-\frac{1}{2}x'(\Sigma_1^{-1}-\Sigma_0^{-1})x]}{-(\mu_0'\Sigma_0^{-1}-\mu_1'\Sigma_1^{-1})x-\frac{1}{2}x'(\Sigma_1^{-1}-\Sigma_0^{-1})x]}$$

$$=\pi(x)$$

經邏輯換得

$$ln\frac{\pi(x)}{1-\pi(x)}=ln(\frac{\theta_0}{\theta_1}|\Sigma_1|^{-\frac{1}{2}}|\Sigma_0|^{\frac{1}{2}})-\frac{1}{2}(\mu_1'\Sigma_1^{-1}\mu_1'-\mu_0'\Sigma_0^{-1}\mu_0')$$

$$-(\mu_0'\Sigma_0^{-1}-\mu_1'\Sigma_1^{-1})x-\frac{1}{2}x'(\Sigma_1^{-1}-\Sigma_0^{-1})x \tag{14}$$

此時，邏輯轉換模式爲

$$\pi(x)=\frac{e^{\beta_0+\beta_1 x+x'\beta_2 x}}{1+e^{\beta_0+\beta_1 x+x'\beta_2 x}}$$

其中，參數之正確值爲

$$\beta_0=ln(\frac{\theta_0}{\theta_1}|\Sigma_1|^{-\frac{1}{2}}|\Sigma_0|^{\frac{1}{2}})-\frac{1}{2}(\mu_1'\Sigma_1^{-1}\mu_1-\mu_0'\Sigma_0^{-1}\mu_0) \tag{14}$$

$$\beta_1=(\mu_1'\Sigma_1^{-1}-\mu_0'\Sigma_0^{-1}) \tag{15}$$

$$\beta_2=(\frac{1}{2}\Sigma_1^{-1}-\Sigma_0^{-1}) \tag{16}$$

欲估計β_0 ，β_1 ，β_2 ，則以$\hat{\mu}_j=\bar{x}_j$，$\theta_j=\frac{n_j}{n}$，$\hat{\Sigma}_j=s_j$，代入(14)～(16)之μ_j，θ_j，Σ_j；$_j=0$，1。

三、改進判別函數估計法

在前言中，曾說明邏輯迴歸模式的判別函數估計法，通常需要進行兩道程序。第一道程序是檢定 y = 0 及 y = 1 兩個族群的自變數 x 是否具有共同的變異數，然後進行第二道程序。第二道程序是根據檢定的結果將樣本數據代入適當的公式求解。這裡的係數求解公式有兩組：一組是爲相同變異數者使用，其中的共同變異數是合併兩個樣本的數據而得；另一組是爲不同變異數者使用，各自樣本的數據計算各自的變異數。

從生物的角度來看，y = 0 及 y = 1 這兩個族群各自的 x 變量沒有理由必需要有相同的變異數（矩陣）。相同的變異數（矩陣）的假設，主要是根據建模簡單化的原則而來的，實際的數據多顯示 y = 0 及 y = 1 這兩個族群並不具有相同的變異數矩陣。即使在一般的判別函數估計法第一道檢定程序發現結果是正面的答案，仍極有可能爲假陽性(false positive) 的假象。Lee, Gurland (1975) ，Lee, Fineberg (1991) ，及Lee (1992)等觀察到這個問題，他們在處理兩個常態樣本時提出一個建議：一律採用不同變異數（矩陣）的立場，他們也從電腦模擬的結果發現這種立場表現得相當好。本文想了解邏輯迴歸模式的參數估計如果也採用這種永遠站在不同變異數（矩陣）的立場會有怎樣的表現。

本文建議的改進判別函數估計法，是一律採用不同變異數（矩陣）的那組公式，即去除一般判別函數估計法的第一道檢定程序。在 x 是單變量的情況下，改進的判別函數估計法是直接將 y = 0 及 y = 1 的各自樣本平均數$\bar{x}j$，樣本變異數s_j^2，及樣本大小n_j代入公式(6)至(8)，即採用公式(9)至(11) ；在 x 是多變量的情況下，則直接將 y = 0 及 y = 1 的各自樣本平均向量$\bar{x}j$，樣本變異數矩陣s_j^2，及樣本大小n_j代入公式(14)至(16)。在下一節中，建議的改進判別函數估計法與一般傳統的判別函數估計法將運用電腦模擬進行比較。

四、電腦模擬與比較

爲了比較改進判別函數估計法與傳統判別函數估計法兩者之間的差異，本文運用電腦模

擬來觀察彼此的表現。配合各種不同的組合，各個組合皆模擬500 次，最後將500 次模擬的結果以偏差(bias) ，均方誤(mean squared error) 表示出來，見表一至表七。由於均方誤的是變異數與偏差平方和，故均方誤、變異數、及偏差僅列出兩個即可。

改進判別函數估計法與傳統判別函數估計法彼此主要的相異之處，是當傳統判別函數估計法遇上檢定結果是正面的時候。這個時候，認定 $y=0$ 及 $y=1$ 兩個族群有共同的變異數（矩陣），傳統的選用變異數（矩陣）相同的公式，而改進的則仍採用變異數（矩陣）相異的公式。令N表示在模擬500 次中出現F檢定結果是正面的個數，表一至表七中的N當然隨著各種不同的組合而有改變。

表一至表七的方法Ⅰ是表示傳統的判別函數估計法，而方法Ⅱ是表示改進後的估計法。表一至表七也特別列出兩個方法之間的比例，Ⅱ／Ⅰ，以供比較，均方誤及偏差是各由下列公式算出來的：

$$\text{均方誤}(\hat{\beta}_j)=\sum_{i=1}^{N}\frac{(\hat{\beta}_{ji}-\beta_j)^2}{N}$$

$$\text{偏　差}(\hat{\beta}_j)=\sum_{i=1}^{N}\frac{(\hat{\beta}_{ji}-\beta_j)}{N}$$

其中j 表示方法Ⅰ或Ⅱ，i是N次正面F檢定結果的各次指標，估計值用帽子^ 加在欲估計的係數β_j上面。由於這是模擬的緣故，真正的係數βj是已知的，以上的計算是代入真正的係數β_j。

$y=0$及 $y=1$兩個族群的 x 變量各爲常態分配$x_0\sim N(\mu_0,\sigma_0^2)$及$x_1\sim N(\mu_1,\sigma_1^2)$，經轉換後可得$x_0'=\frac{x_0-\mu_1}{\sigma_1}\sim N(\frac{x_0-\mu_1}{\sigma_1},\frac{\sigma_0^2}{\sigma_1^2})$及$x_1'=\frac{x_1-\mu_1}{\sigma_1}\sim N(0,1)$。故在不失一般性的原則下，所有模擬可令 $y=1$族群的 x 變量之$\mu_1=0$ 且$\sigma_1=1$。本次模擬所考慮$y=0$族群的 x 變量的平均數μ_0有兩種，$\mu_0=1$及$\mu_0=4$，分別代與 $y=1$族群“靠近”或“疏遠”的情形。而$y=0$族群 x 變量之變異數σ_0則取爲$y=1$族群者的兩倍，或一半，以及兩兩的幾何平均中間數$\sqrt{0.5}$與$\sqrt{2}$，分別代表$y=0$及$y=1$兩個族群間 x 變異數不同的變化。

$y=0$及$y=1$兩個族群樣本大小，本次模擬考慮的有小樣本 $n=21$者，也考慮兩者差異很大如$n_1=21$與$n_0=210$ ，彼此有十倍之遠者，其他考慮的有$n=21$的兩倍$n=42$及五倍者

n＝105 ，分別表示y＝0及y＝1族群樣本大小彼此比例的不同變化。到底多大才算是所謂的“大樣本”，這完全決定在數據本身。從這次模擬中看出：若 y ＝0及 y ＝1兩個族群樣本大小分別在n_0＝105 及n_1＝42或以上時，已有“大樣本”的跡象了（見表六或表七）。

模擬的結果顯示，不論 y ＝0及 y ＝1兩個族群x 變量的平均數相距疏遠與否，在500 次模擬中出現F 檢定正面答案的次數N是與兩個族群x 變量的變異數比較有關。當兩者變異數相近時，F 檢定是正面的機會就會增加，故$\sigma_0=\sqrt{0.5}$或$a_0=\sqrt{2}$都比σ_0＝0.5或σ_0＝2時擁有較多的N。見表一，在小樣本$n_0=n_1$＝21時，在$\sigma_0=\sqrt{0.5}$或$a_0=\sqrt{2}$時的N都高達400 次以上；甚至在表四，N仍超過500 次的一半以上。這個現象隨著樣本大小的增加而有所改善，這點可由表六及表七中N會是0 看出來。在兩者變異數相同時，N都保持在490 左右，這是比較合理的現象，因爲這次模擬中的檢定是採用第一型錯誤(Type one error)在0.02。

以上的現象說明，方法Ⅰ，即傳統的判別函數估計法所需經過的檢定程序只有在兩個族群變異數相同時才能發揮功能。在其他變異數不相同時，傳統的判別函數估計法是犯下嚴重的錯誤，而且會有很高的機會（會有500 次中超過400 次的可能性）錯誤的採用變異數相同的公式去估計參數。特別是在σ_0與σ_1比較接近，但實際上並非相等的時候，這種現象是急速的惡化！因爲檢定出正面答案的假陽性會隨著σ_0與σ_1接近程度而增加。故使用傳統判別函數估計法時不得不戒懼小心：也許σ_0 只是和σ_1 非常的接近，但並不是相等！

在 y ＝0及 y ＝1族群 x 變量的平均數相當“疏遠”的情形，不論變異數相異程度是如何，方法Ⅱ，即是改進的判別函數估計法會隨著樣本大小的增加而表現得越來越好。不僅在變異數相同時改進的估計法隨著樣本大小的增加而變得更好；甚至在變異數不相同時，方法Ⅱ仍有如此良好的表現。見表一至表七的μ_0＝4及μ_0＝1的情況，不論是各係數估計值的均方誤或偏差，方法Ⅱ與方法Ⅰ之比例，Ⅱ／Ⅰ，可從 2 左右隨著樣本大小的增加而降至 1 左右，這是說方法Ⅱ即使在變異數相同時，幾乎可以和方法Ⅰ一樣的好！

在y＝0及y＝1族群 x 變量的平均數相當“接近”的情形，改進的判別函數估計法在y＝0及y＝1族群 x 變量的變異數不相同時，也是隨著樣本大小的增加而表現得越來越好。在兩個族群變異數相同時，大多數的模擬結果也顯示方法Ⅱ仍幾乎可以和方法Ⅰ同樣的好。見表一至表七的μ_0＝1及σ_0＝1的情況，絕大多數的機會，方法Ⅱ與方法Ⅰ的比例Ⅱ／Ⅰ都在 1 右

左，表示方法Ⅱ相當匹配方法Ⅰ。

五、結　論

邏輯迴歸模式的參數估計，傳統的判別函數估計法中的檢定程序，在y＝0及y＝1兩個族群 x 變量的變異數相同時的確發揮功效。但一旦這兩個族群變異數有些許的不相同時，檢定的結果仍傾向去接受與事實相反的答案，導致傳統的判別函數估計法是犯下嚴重的錯誤。這項錯誤的嚴重性從電腦模擬的結果看出有時會有500次中超過400次的機會！故傳統的判別函數估計法中所含的檢定程序並非一個非常合理的程序。

本文在電腦模擬傳統判別函數估計法中的變異數檢定方法是採用兩個樣本變異數相除形成的F檢定。這檢定是眾所周知的標準步驟，在基本的統計課程必定介紹此一檢定。F檢定在兩個母體是常態分配，而樣本是互相獨立時，具有最大檢力(power)，故傳統判別函數估計法中改用其他的變異數檢定方法，其結果不會比使用F檢定好。

在實際的生活中，y＝0及 y＝1兩個族群 x 變量的變異數是很相近，但不是相同的看法是比較合理的生物真象。因此，使用傳統的判別函數估計法是冒著觸犯檢定錯誤的風險，爲何不使用去除檢定程序的改進判別函數估計法呢？改進的判別函數估計法雖然在任何狀況都直接運用變異數相同的公式，從本文的模擬顯示，建議的改進判別函數估計法在大多數的情形下比傳統的表現優良，但少數的情形下仍可與傳統的相當。故本文提出的改進判別函數估計法是一種合理的建議。

本文進行的模擬中之檢定，第一型錯誤是在0.02上，所以在500次的模擬中期望有500×0.02＝10次左右的錯誤。實際模擬得到的結果，在真正是變異數相同的情況下，檢定結果是正面的次數N都在490　次左右。可見本次電腦模擬的結果與事實真象應是非常吻合的，本文的結論及建議是值得信賴的。

作者也曾模擬其他的組合，其結果均印證本文的結論，爲免贅述，故不在本文重覆，歡迎有興趣的讀者與作者連絡討論。

參 考 文 獻

Cornfield, J (1962). Joint dependence of the risk of coronary heart disease on serum cholesterol and systolic blood pressure: A discriminant function analysis. *Federation Proceedings,* 21, 58-61.

Cox, DR, & Snell, EJ (1989). *Analysis of binary data. Chapman* and Hall, London

Efron, B (1975). The efficiency of logistic regression compared to normal discriminant function analysis. *Journal of the American Statistical Association,* 70, 892-898.

Grizzle, J, Starmer, F, & Koch, G (1969). Analysis of categorical data by linear models. *Biometrics,* 25,489-504.

Hosmer, DW, & Lemeshow, S (1989). *Applied logistic regression.* Wiley, New York

Lee, AFS, (1992). Optimal Sample sizes determined by two-sample Welch's test. *Communications in Statistics,* No.3.

Lee, AFS, & Fineberg, NS (1991). A fitted test for the Behrens-Fisher problem. *Communications in Statistics,* No.2, 653-666.

Lee, AFS, & Gurland (1975). Size and power and tests for equality of means of two normal populations with unequal variances. *Journal of the American Statistical Association,* 70, 933-941.

McCullagh, P & Nelder, JA (1983). *Generalized Linear Models.* Chapman Hall, London.

		$\mu_1=0$, $\sigma_1=1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
21	21	4	1	I	5.1238	0.4255	1.1848	-0.2171	496
				II	10.5949	0.8978	2.6952	-0.4551	
				II/I	2.0678	2.1100	2.2747	2.0986	
			$\sqrt{0.5}$	,,	25.4463	-4.0618	6.9675	2.1821	413
					20.6130	0.1142	5.2584	-0.0389	
					0.8101	-0.0281	0.7547	-0.0179	
			0.5	,,	244.781	-15.128	67.6470	7.9904	130
					70.648	-5.2521	18.1773	2.8597	
					0.2886	0.3472	0.2687	0.3579	
			$\sqrt{2}$	,,	4.5887	1.3456	1.2867	-0.8688	417
					3.0421	0.5868	0.7924	-0.3634	
					0.6630	0.4361	0.6158	0.4183	
			2	,,	3.7660	1.3160	1.3857	-1.0077	119
					1.8324	0.6773	0.5947	-0.4972	
					0.4866	0.5146	0.4292	0.4934	
		1	1	,,	0.0773	0.0096	0.2073	-0.0639	491
					0.0992	0.0367	0.3004	-0.1207	
					1.2823	3.8471	1.4496	1.8885	
			$\sqrt{0.5}$	,,	0.0953	0.0557	0.6082	0.5297	414
					0.1965	0.0888	0.6227	-0.0456	
					2.0626	1.5925	1.0239	-0.0861	
			0.5	,,	0.2280	-0.2666	4.2456	1.8711	123
					0.3763	0.0249	1.3748	0.5436	
					1.6504	-0.0934	0.3238	0.2905	
			$\sqrt{2}$	,,	0.1128	-0.2367	0.1356	-0.1948	401
					0.0528	-0.0419	0.1231	-0.0637	
					0.4687	0.1769	0.9078	0.3269	
			2	,,	0.3125	-0.5010	0.1475	-0.2691	114
					0.0778	-0.1969	0.0815	-0.1263	
					0.2490	0.3928	0.5527	0.4692	

表 一

		$\mu_1 = 0$, $\sigma_1 = 1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
21	42	4	1	I	3.1587	0.3469	0.7682	-0.1874	489
				II	4.0528	0.4390	1.0731	-0.2269	
				II/I	1.2830	1.2656	1.3970	1.2112	
			$\sqrt{0.5}$	,,	14.7291	-2.8961	4.1435	1.5989	337
					9.8823	-0.4406	2.4400	0.2545	
					0.6709	0.1521	0.5889	0.1592	
			0.5	,,	104.760	-9.6381	28.5370	4.9968	44
					34.714	-3.4680	9.0830	1.8003	
					0.331	0.3598	0.3183	0.3603	
			$\sqrt{2}$	,,	1.6642	0.6686	0.5251	-0.5152	376
					1.1657	0.2871	0.3309	-0.1958	
					0.7004	0.4294	0.6301	0.3800	
			2	,,	1.0926	0.6315	0.5725	-0.6752	72
					0.6592	0.4308	0.2656	-0.3912	
					0.6033	0.6822	0.4640	0.5794	
		1	1	,,	0.0372	-0.0050	0.1120	-0.0152	485
					0.0512	0.0132	0.1392	-0.0369	
					1.3752	-2.6615	1.2428	2.4360	
			$\sqrt{0.5}$	,,	0.0787	0.1373	0.4231	0.3398	358
					0.1010	0.0867	0.3291	-0.0014	
					1.2835	0.6312	0.7778	-0.0040	
			0.5	,,	0.1158	-0.1188	3.0569	1.5547	50
					0.1530	-0.0190	1.0300	0.7059	
					1.3220	0.1600	0.3369	0.4540	
			$\sqrt{2}$	,,	0.1082	-0.2946	0.0694	-0.1150	356
					0.0320	-0.0627	0.0818	-0.0343	
					0.2957	0.2129	1.1789	0.297	
			2	,,	0.3966	-0.6169	0.0449	-0.1513	67
					0.0856	-0.2599	0.0413	-0.0911	
					0.2159	0.4213	0.9208	0.6022	

表 二

		$\mu_1=0$, $\sigma_1=1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
21	105	4	1	I	1.2632	0.1448	0.3493	-0.0870	489
				II	1.3658	0.1920	0.4097	-0.1036	
				II/I	1.0813	1.3257	1.1730	1.1904	
			$\sqrt{0.5}$	,,	5.2974	-1.2849	1.6709	0.8174	295
					4.1584	-0.2432	1.0824	0.1874	
					0.7850	0.1892	0.6478	0.2292	
			0.5	,,	51.453	-5.9431	14.1371	3.1016	21
					26.125	-3.2165	6.9597	1.6474	
					0.5077	0.5412	0.4923	0.5311	
			$\sqrt{2}$	,,	0.4222	0.1191	0.1506	-0.2319	363
					0.3658	0.0967	0.1295	-0.0766	
					0.8663	0.8115	0.8596	0.3305	
			2	,,	0.2274	-0.1409	0.1228	-0.2783	43
					0.1696	0.0288	0.0808	-0.1667	
					0.7455	0.2040	0.6577	0.5988	
		1	1	,,	0.0160	-0.0083	0.0747	-0.0242	491
					0.0396	0.0084	0.0884	-0.0299	
					1.9366	-1.0145	1.1839	1.3386	
			$\sqrt{0.5}$	,,	0.0652	0.1910	0.2671	0.2265	289
					0.0451	0.0728	0.1686	0.0693	
					0.6925	0.3812	0.6312	0.3060	
			0.5	,,	0.1857	0.2847	1.2917	0.8091	29
					0.1833	0.2338	0.7254	0.4306	
					0.9872	0.8213	0.5616	0.5322	
			$\sqrt{2}$	,,	0.1206	-0.3359	0.0258	-0.0252	341
					0.0262	-0.0765	0.0507	0.0211	
					0.2169	0.2279	1.9665	-0.8367	
			2	,,	0.4546	-0.6726	0.0111	-0.0751	30
					0.0900	-0.2880	0.0223	-0.0642	
					0.1979	0.4282	2.0072	0.8551	

表 三

		$\mu_1 = 0$, $\sigma_1 = 1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
21	210	4	1	I	0.6177	0.0483	0.1907	-0.0324	488
				II	0.6322	0.0735	0.2263	-0.0324	
				II/I	1.0235	1.5212	1.1870	1.0021	
			$\sqrt{0.5}$	,,	2.3716	-0.5290	0.7831	0.4053	275
					2.2383	-0.0698	0.6037	0.0981	
					0.9434	0.1320	0.7709	0.2421	
			0.5	,,	19.184	-2.5899	5.7723	1.4942	16
					12.111	-1.0573	3.3211	0.6327	
					0.6313	0.4082	0.5753	0.4234	
			$\sqrt{2}$	,,	0.1956	-0.1258	0.0639	-0.1254	342
					0.1613	-0.0037	0.0793	-0.0514	
					0.8244	0.0292	1.2418	0.4100	
			2	,,	0.2528	-0.4180	0.0320	-0.1080	30
					0.0895	-0.3029	0.0353	-0.0228	
					0.3540	0.3180	1.1014	0.2112	
		1	1	,,	0.0082	-0.0158	0.0613	-0.0174	491
					0.0277	0.0007	0.0786	-0.0184	
					3.3568	-0.0444	1.2818	1.0590	
			$\sqrt{0.5}$	,,	0.0912	0.2567	0.2240	0.1001	246
					0.0435	0.1086	0.1568	0.0086	
					0.4770	0.4231	0.6999	0.0855	
			0.5	,,	0.2125	0.4267	0.6576	0.2985	15
					0.1305	0.3142	0.2660	0.1350	
					0.6143	0.7362	0.4045	0.4522	
			$\sqrt{2}$	,,	0.1241	-0.3466	0.0205	-0.0210	357
					0.0232	-0.0917	0.0473	0.0038	
					0.1868	0.2647	2.3130	-0.1817	
			2	,,	0.4559	-0.6721	0.0110	-0.0197	17
					0.0956	-0.2973	0.0258	0.0116	
					0.2096	0.4423	2.3439	-0.5895	

表 四

		$\mu_1=0$, $\sigma_1=1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
42	42	4	1	I	1.9194	0.1967	0.4341	-0.0760	489
				II	3.4151	0.3861	0.8722	-0.1701	
				II/I	1.7793	1.9630	2.0091	2.2397	
			$\sqrt{0.5}$	,,	23.5159	-4.3507	6.6037	2.3307	277
					9.4249	-0.9434	2.4146	0.5051	
					0.4008	0.2168	0.3656	0.2167	
			0.5	,,	276.420	-16.470	75.6454	8.5681	6
					134.660	-11.239	36.8745	5.8058	
					0.487	0.682	0.4875	0.6776	
			$\sqrt{2}$	,,	3.1143	1.3408	0.9676	-0.8374	292
					1.4063	0.5338	0.3741	-0.3145	
					0.4516	0.3982	0.3866	0.3756	
			2	,,	4.2640	1.4868	1.6859	-1.1475	6
					1.9071	0.8422	0.7228	-0.6802	
					0.4473	0.5664	0.4287	0.5928	
		1	1	,,	0.0361	0.0240	0.0869	-0.0507	490
					0.0438	0.0330	0.1201	-0.0802	
					1.2152	1.3783	1.3828	1.5822	
			$\sqrt{0.5}$	,,	0.0446	0.0484	0.4836	0.5874	273
					0.0716	0.0374	0.2216	0.1486	
					1.6059	0.7725	0.4581	0.2531	
			0.5	,,	0.0769	-0.1684	3.0154	1.6938	12
					0.0923	0.0233	0.7419	0.7136	
					1.2001	-0.1386	0.2460	0.4213	
			$\sqrt{2}$	,,	0.0880	-0.2375	0.0945	-0.1923	262
					0.0293	-0.0656	0.0624	-0.0613	
					0.3330	0.2764	0.6600	0.3187	
			2	,,	0.3682	-0.5867	0.0703	-0.2152	6
					0.1338	-0.3431	0.0294	-0.1182	
					0.3634	0.5848	0.4182	0.5492	

表 五

		$\mu_1=0$, $\sigma_1=1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
42	105	4	1	I	1.0715	0.0721	0.2560	-0.0494	489
				II	1.3573	0.1642	0.3653	-0.0989	
				II/I	1.2667	2.2772	1.4267	2.0029	
			$\sqrt{0.5}$	,,	9.8347	-2.7141	2.9174	1.5147	170
					4.4399	-0.9962	1.1493	0.5592	
					0.4515	0.3671	0.3939	0.3692	
			0.5	,,	102.060	-9.9210	30.2177	5.4707	3
					44.553	-6.2712	12.4893	3.4400	
					0.437	0.6321	0.4133	0.6288	
			$\sqrt{2}$	,,	0.6289	0.4735	0.2686	-0.4151	210
					0.3960	0.2044	0.1320	-0.1649	
					0.6296	0.4316	0.4914	0.3972	
			2	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	
		1	1	,,	0.0141	0.0078	0.0516	-0.0158	494
					0.0229	0.0189	0.0624	-0.0188	
					1.6242	2.4137	1.2081	1.1916	
			$\sqrt{0.5}$	,,	0.0447	0.1481	0.2536	0.3549	204
					0.0356	0.0717	0.1354	0.1262	
					0.7961	0.4844	0.5341	0.3556	
			0.5	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	
			$\sqrt{2}$	,,	0.0936	-0.2887	0.0382	-0.1155	207
					0.0210	-0.0938	0.0367	-0.0531	
					0.2239	0.3251	0.9609	0.4593	
			2	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	

表 六

		$\mu_1=0$, $\sigma_1=1$			β_0		β_1		
n_1	n_0	μ_0	σ_0	方法	均方誤	偏差	均方誤	偏差	N
42	210	4	1	I	0.5867	0.0551	0.1571	-0.0387	487
				II	0.6149	0.0789	0.1746	-0.0478	
				II/I	1.0479	1.4320	1.1120	1.2357	
			$\sqrt{0.5}$	,,	3.8165	-1.3950	1.1860	0.8350	127
					2.2070	-0.4810	0.5559	0.2913	
					0.5783	0.3448	0.4687	0.3488	
			0.5	,,	*	-3.8852	*	1.9236	1
					*	-1.1318	*	0.5328	
					*	0.2913	*	0.2770	
			$\sqrt{2}$	,,	0.2335	0.1525	0.1114	-0.2519	179
					0.1963	0.1108	0.0665	-0.1180	
					0.8406	0.7267	0.5965	0.4683	
			2	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	
		1	1	,,	0.0063	-0.0062	0.0374	-0.0120	491
					0.0151	0.0028	0.0429	-0.0124	
					2.4065	-0.4481	1.1482	1.0303	
			$\sqrt{0.5}$	,,	0.0606	0.2155	0.1806	0.2740	127
					0.0296	0.1012	0.0940	0.1281	
					0.4888	0.4696	0.5208	0.4675	
			0.5	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	
			$\sqrt{2}$	,,	0.1267	-0.3205	0.0155	-0.0570	197
					0.0171	-0.1066	0.0217	-0.0235	
					0.1605	0.3326	1.4060	0.4115	
			2	,,	*	*	*	*	0
					*	*	*	*	
					*	*	*	*	

表 七