

二分因變數之會計學研究 應採用 Probit 或 OLS 模型之探討

洪欣慧
淡江大學

蔡彥卿
國立台灣大學

摘要

國內二分因變數的會計研究大部份屬於小樣本的情況，因而 Probit 模型之估計值不一定會較 OLS (ordinary least square) 更有效率。本研究嘗試由實證的角度，用模擬的方式，希望能知道那種模型較適合國內二分因變數的會計研究？

吾人首先由顯著水準及檢定力兩方面著手，比較此二模型在台灣證券市場之實際資料下，何者表現較佳。此外，本研究並分析此二方法在預測能力上之優劣。期望透過這三種比較方式，推論出何種模型較適合國內的會計資料，以期使研究者在面對模型選擇時，能有所依據。

本研究利用民國 76 年至 80 年的資料做模擬測試，結果顯示：當虛無假設為真時，OLS 檢定的統計量之真實顯著水準與理論顯著水準的差距比 Probit 小；而當虛無假設不正確時，OLS 的檢定力也比 Probit 為高。另外，在預測能力的比較分析上，也發現 OLS 表現比 Probit 為佳。所以整體而言，在本研究的樣本量和方法下，OLS 模型比 Probit 模型更適合國內二分因變數的會計研究。亦即當我們面臨這類會計研究之模型選擇時，簡單的線性迴歸似乎比複雜的 Probit 模型更為合適。

作者感謝蔡信夫教授及蔡揚宗教授之意見，本文受益良多。此外，林嬋娟教授與吳安妮教授也對作者幫助甚多。文中若有任何缺失，當為作者之責任。

An Empirical Comparison of Probit and OLS Regression -- Hypothesis Tests and Prediction Accuracy

Hsin-hui, Hung
Tamkang University
Yann-ching Tsai
National Taiwan University

Abstract

Noreen (1988) compared the relative performance of Probit and OLS regression significance tests. He concluded that for the kind of data and sample sizes commonly found in accounting classificatory studies, OLS regression performs at least as well as Probit.

In this paper, we compared the two models in two respects, significance tests and prediction accuracy. Using empirical data from firms listed on Taiwan Stock Exchange, our simulation indicates the following results: (I) the empirical distributions of the OLS regression conform closer to the theoretical distributions, when the null hypothesis is true, (II) OLS regression is more powerful, when the alternative hypothesis is true, and (III) the correct prediction ratio of OLS regression is higher than that of Probit model.

Based on the empirical results, we suggest the use of OLS regression in accounting classificatory studies for the sample sizes similar to our Monte Carlo experiments.

壹、前言

會計學研究中以 **Probit** 模型從事二分因變數屬性研究時，其樣本量一般來說並不大（註一）。在小樣本情況下，**Probit** 模型之估計值不一定會比 **OLS** (**ordinary least square**) 估計值更有效率。故本研究之目的即嘗試由實證的角度，用模擬的方式，希望能知那種模型較適合國內二分因變數的會計研究？

1988 年時 **Noreen** 曾利用美國之實際資料，從真實顯著水準與名目顯著水準 (**real and nominal significance levels**) 的差距，及檢定力之差異上，比較 **OLS** 與 **Probit** 模型在小樣本情況下，何法較佳？其結果為：(1) 顯著水準部份—**OLS** 的表現較 **Probit** 模型佳；即 **OLS** 模型之 **t** 檢定的真實顯著水準與名目顯著水準較接近。(2) 檢定力部份—兩模型不相上下。所以 **Noreen** 認為在小樣本的情況下，似乎可以 **OLS** 取代 **Probit** 模型，來處理二分因變數之研究。

本研究首先由顯著水準及檢定力兩方面著手，比較此二模型在台灣證券市場之實際資料下，何者表現較佳。此外，吾人認為欲判斷模型之優劣，預測力之大小也是很重要的一个指標。所以，本研究並分析此二方法在預測能力上之優劣。因此，本文比 **Noreen** 更為周延。作者期望透過這三種模型比較方式，推論出何種模型較適合國內的會計資料，以期使研究者在面對模型選擇時，能有所依據。

吾人利用民國 76 年至 80 年的資料做模擬測試，結果顯示：當虛無假設為真時，**OLS** 檢定的統計量之真實顯著水準與理論顯著水準的差距比 **Probit** 小；而當虛無假設不正確時，**OLS** 的檢定力也比 **Probit** 為高。另外，在預測能力的比較上，也發現 **OLS** 表現比 **Probit** 為佳。所以整體而言，在本研究的樣本量和方法下，**OLS** 模型比 **Probit** 模型更適合國內二分因變數的會計研究。亦即當我們面臨這類會計研究之模型選擇時，簡單的線性迴歸似乎比複雜的 **Probit** 模型更為合適。

註一：若解釋變數有 **k** 個，而樣本量小於 **30k** 則稱為小樣本 (**McFadden [1984, P. 1441-42]**)。

貳、研究期間、樣本與資料

本研究係以台灣上市公司為研究對象，故採用台灣市場之實際財務資料（金融股除外）為主要之解釋變數。然而財務資料項目繁多，何者應為本研究分析之對象呢？吾人將國內會計學研究中曾使用二分因變數模型作彙總分析，統計各項財務變數在文獻中被使用之次數，選用其中使用較頻繁的四個變數為本研究之解釋變數。下表即對國內七篇研究〔沈遠芳，民81；施春成，民81；彭秋萍，民81；程佩瑜，民81；鄧洪，民81；蕭博仁，民81；及蘇裕惠，民80〕所選用之變數做一彙總：

表一 國內相關研究使用變數之彙總

變數	1.	2.	3.	4.	5.	6.	7.	合計
負債比率	✓	✓	✓			✓	✓	5
銷貨成長率	✓	✓	✓	✓				4
營業淨利成長率				✓				1
總資產		✓	✓				✓	3
資本額				✓				1
銷貨額				✓		✓		2
稅後（前）淨利	✓							1
營業淨利				✓				1
稅後淨利/淨值	✓							1
流動比率	✓							1
應收帳款/總資產			✓					1
存貨週轉率	✓							1
存貨/總資產			✓					1
產業別	✓							1
公司成立年數	✓	✓				✓		3
非財務性資料	✓	✓		✓	✓	✓	✓	—
其他財務資料	✓			✓		✓	✓	—

根據上表之結果顯示，有四項財務資料使用次數較為頻繁，其分別為：

1. 總資產 (X1)
2. 負債比率 (X2) : $\frac{\text{負債總額}}{\text{股東權益總額}}$
3. 銷貨成長率 (X3) : $\frac{(\text{當期銷貨額} - \text{前期銷貨額})}{\text{前期銷貨額}}$
4. 公司成立年數 (X4)

本研究即依此結果，選用上述四項財務資料為解釋變數。因為國內二分因變數屬性之會計學研究中，多以七十六至八十年之上市公司為研究對象，故本研究選用相同期間。故總樣本數為五年內的 1006 個資料點。以下即對所選用之四項解釋變數，在這五年期間各年之情況作一簡單之描述。（資料來源：台灣經濟新報社財務資料庫）。但資料不全者不包含在研究資料內。

表二 描述性統計資料

變數名稱：總資產 (X1)

單位：百萬

時間	樣本量	平均數	標準差	最大值	中位數值	最小值
76	197	4,399	10,786	130,810	1,968	128
77	199	5,179	11,581	136,400	2,266	210
78	203	5,961	12,087	140,330	2,799	267
79	203	6,646	12,886	149,340	3,091	300
80	204	7,444	13,632	155,620	3,457	290
總和	1006	5,939	12,269	155,620	2,698	128

表三 描述性統計資料

變數名稱：負債比率 (X2)

時間	樣本量	平均數	標準差	最大值	中位數值	最小值
76	197	2.798	17.570	245.985	1.138	.09
77	199	1.231	1.089	8.408	0.968	.049
78	203	0.996	0.880	5.871	0.807	.085
79	203	0.795	0.649	4.217	0.610	.053
80	204	0.742	0.533	3.518	0.630	.051
總和	1006	1.303	7.830	245.985	0.784	.049

表四 描述性統計資料

變數名稱：銷貨成長率 (X3)

時間	樣本量	平均數	標準差	最大值	中位數值	最小值
76	196	0.537	1.586	19.808	0.080	-.81
77	199	0.788	8.058	113.667	0.092	-.31
78	203	0.141	0.333	2.268	0.074	-.41
79	203	0.219	1.279	16.502	0.058	-.54
80	204	0.120	0.318	2.790	0.088	-.47
總和	1006	0.323	3.705	113.667	0.079	-.81

表五 描述性統計資料

變數名稱：公司成立年數 (X4)

時間	樣本量	平均數	標準差	最大值	中位數值	最小值
76	197	19.638	9.135	41	19.0	1
77	199	20.492	9.109	42	20.0	2
78	203	21.606	9.955	77	21.0	2
79	203	22.606	9.955	78	22.0	3
80	204	23.598	9.932	79	23.0	4
總和	1006	21.586	9.720	79	21.0	1

由描述性統計資料來分析：

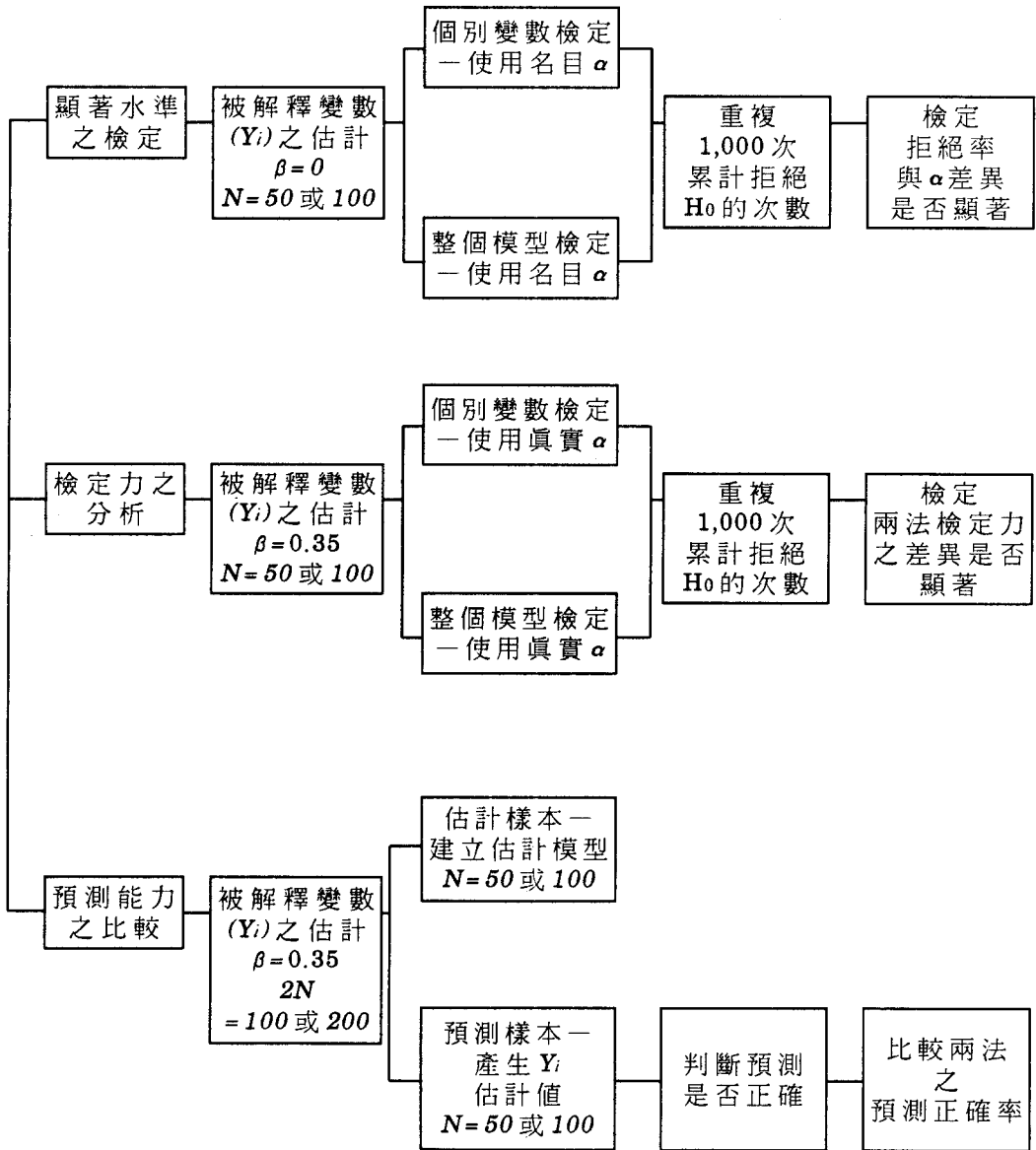
1. 總資產：呈現逐年上升的趨勢，且其標準差亦逐年擴大。
2. 負債比率：除七十九及八十年之平均數略低外，其餘各年大致平穩。但七十六年之標準差因受異常值 (outliers) 之影響即比其它各年大 (註二)。
3. 銷貨成長率：七十七年略高以外，其餘各年平均數大致穩定。七十七年之標準差與平均數受異常值的影響而偏高。
4. 公司成立年數：各年平均數逐年穩定增加，標準差則無太大差異。

註二：一般研究均將異常值刪除，再觀察結果是否與刪除前差異很大。在本研究中，因為每次取出之樣本數不大，所以異常值對結果之影響力可能不大；然而為求結果之完整，本研究將與平均數差異達三個標準差以上的樣本刪除，重複試驗，結果與未刪除前類似。

參、研究方法

本研究係先以台灣上市公司之實際財務資料作為解釋變數 ($X1 \sim X4$)，並加上一個虛擬變數 (**dummy variable**($X5$))，再以隨機抽樣產生誤差項，如此即可「製造」出被解釋變數值 (Y 值)，最後以「製造」出之 Y 值和實際之 X 值作模型之估計與檢定；前述由資料至統計檢定的一個完整程序為一次實驗。然後再重複實驗，以檢驗比較二個模型在以下三方面的差異：顯著水準、檢定力及預測能力。本研究之研究方法的主要架構如下圖：

圖一 研究方法的主要架構圖

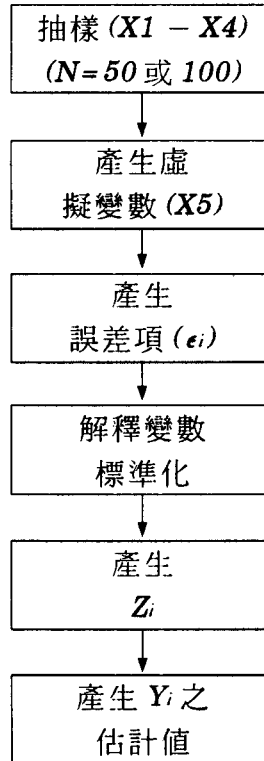


以下分三部份解釋研究架構圖：第一為被解釋變數 (Y_i) 之估計過程，第二為顯著水準 (α) 及檢定力 ($1 - \beta$) 之比較，第三為模擬預測的方法。

一、被解釋變數 (Y_i) 之估計過程

被解釋變數之估計過程，以流程圖方式表列於下：

圖二 Y_i 之估計過程



1. 抽樣—自資料集，以取出不放回的方式抽出 $N(N=50$ 及 100 等二種樣本量) 個樣本。包含總資產 ($X1$)、負債比率 ($X2$)、銷貨成長率 ($X3$) 及公司成立年數 ($X4$) 等四項變數。

2. 產生虛擬變數 ($X5$)—因為解釋變數常包含虛擬變數，故每一個樣本的第五個解釋變數以隨機的方式任意給予 0 或 1 之值，而其中 0 與 1 之比例為 0.6:0.4。此一比例乃比照 Noreen 之研究，代表兩組公司之間的比率相差不大的情況。但一般研究的組間樣本數的差異可能較大，因此吾人，亦對 0.9:0.1 之比例分配作敏感性分析。

3. 產生誤差項 (ϵ_i)—以標準常態分配隨機產生。

4. 解釋變數 ($X1 - X4$) 標準化—即將各解釋變數標準化 ($(X_i - \text{平均數}) / \text{標準差}$)，其目的是為減少不同單位之影響，如總資產之單位為元，數額可能相當大，而負債比例係為百分比，數值相對於總資產而言，顯得相當小；若未做標準化，總資產所占權數較重。

5. 產生 Z_i 值— Z_i 之產生乃依下列方程式得到：

$$Z_i = \beta_0 + \beta_1 X1 + \beta_2 X2 + \beta_3 X3 + \beta_4 X4 + \beta_5 X5 + \epsilon_i,$$

$X1$ ：總資產

$X2$ ：負債比例

$X3$ ：銷貨成長率

$X4$ ：公司成立年數

$X5$ ：虛擬變數。

6. 產生 Y_i 之估計值—設 Z^* 為各 Z_i 值扣除 ϵ_i 後之平均，若 $Z_i > Z^*$ 則 $Y_i = 1$ ；若 $Z_i < Z^*$ 則 $Y_i = 0$ 。在做顯著水準（比較）時， β 以 0 代入，在做檢定力分析時， β 以 0.35 代入（註三）。

二、顯著水準及檢定力之比較

1. 各模型中個別解釋變數之檢定：
 OLS 採用 t 檢定
 Probit 採用近似 t 檢定。
2. 整個模型檢定：
 OLS 採用 F 檢定
 Probit 採用 log-likelihood ratio 檢定。

3. 將 1 與 2 中之檢定重覆 1,000 次後，即可得 OLS 與 Probit 模型，拒絕 H_0 的真實比率，即真實 α 。再將之與名目之拒絕率（設定之顯著水準 α ，即名目 α ）相比，可以瞭解那一個模型較佳。

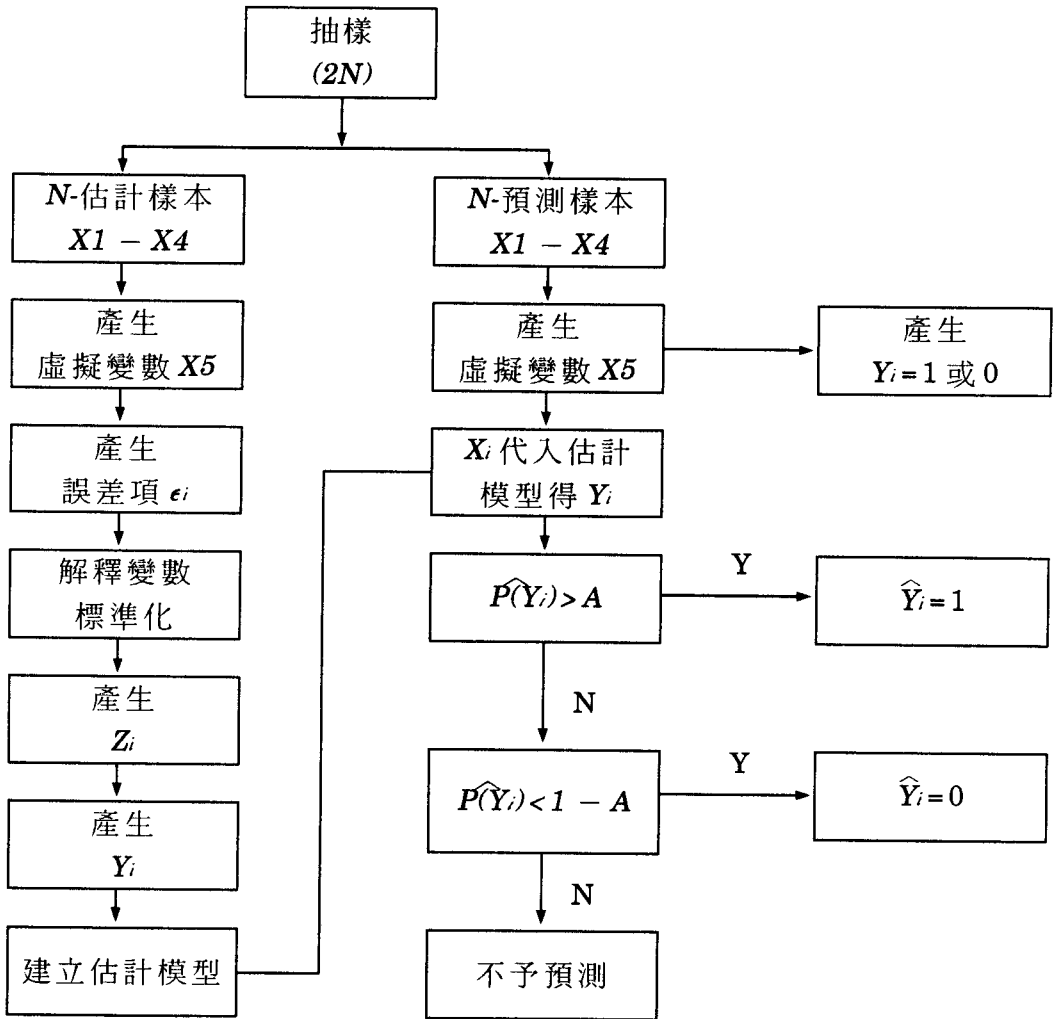
4. 檢定力之比較—為比較二個模型之檢定力何者較佳，必須控制真實 α ，使之相等。所以使用真實之 α 作為界定臨界值 (critical value) 之基礎，重覆 1,000 次實驗後，可以計算當對立假設正確時，統計量超越顯著水準之比率，此即實驗而得之檢定力。依此可比較二個模型檢定力之高低。

三、模擬預測方法：

本研究分別測試 OLS 與 Probit 模型兩法在預測能力上何者較佳？同樣地，係利用重覆模擬的方式進行，其預測之步驟如下圖：

註三： $\beta = 0.35$ 使得檢定力不至於過大或過小（參考 Noreen）。

圖三 模擬預測之流程圖



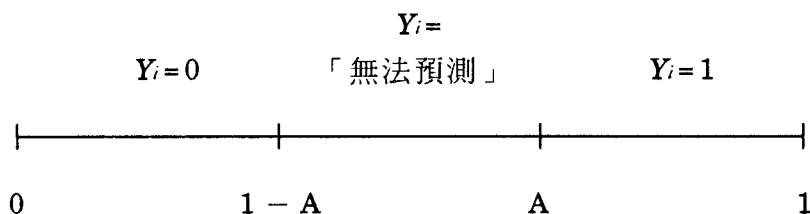
圖三之解釋如下：

1. 抽樣—自資料集中，以取出不放回的方式，抽取出 $2N$ 組 ($X1 - X4$)。再將該樣本一分為二，第一組 N 個樣本為估計樣本，第二組 N 個樣本為預測樣本。本研究將分別以 $N=50$ 及 $N=100$ ，二種不同之樣本量進行研究。

2. 建立估計模型—由估計樣本來建立估計模型，同樣地，須先經由前第一段所述之被解釋變數之產生過程產生 Y_i 。由 X_i 與 Y_i 可經由迴歸過程

建立 OLS 模型，另經由 Probit 之解釋變數估計程序建立 Probit 模型。

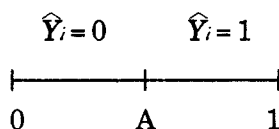
3. 預測過程—(1) 同樣地，須先依虛擬變數產生原則，產生預測樣本組中之虛擬變數。(2) 將 X_i 代入估計模型中，即可得 Y_i 等於 1 之機率 $P(\hat{Y}_i)$ 。(3) 若 $P(\hat{Y}_i)$ 大於某一特定值 A ，則令 $\hat{Y}_i = 1$ ；若 $P(\hat{Y}_i)$ 小於 $1 - A$ ，則令 $\hat{Y}_i = 0$ ；若 Y_i 係介於 A 與 $1 - A$ 之間，則不對 Y_i 做任何預測。其關係如下：



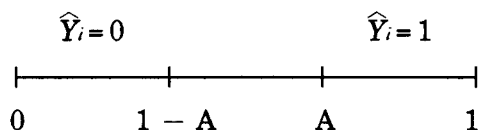
本研究為求完整起見，將 A 值分別設定為 0.5，0.6，0.7，0.8，0.9（故相對之 $1 - A$ 值為 0.5，0.4，0.3，0.2，0.1）。

現以 $A=0.5$ 及 $A=0.9$ 兩種情況來作說明：

〔例 $A=0.5$ 〕



〔例 $A=0.9$ 〕



在 $A=0.9$ 的情況下，係將以該法 (OLS 或 Probit) 計算出之 \hat{Y}_i ，較有把握 $Y_i=0$ 或 $Y_i=1$ 的部份篩選出來，再與 Y_i 作比較，以決定預測是否正確。 $A=0.5$ 的情形下，則對每一個預測樣本做預測，以比較模型之優劣。

肆、研究結果與分析

一、OLS 法與 Probit 模型顯著水準之測試

本節之主要目的是爲了檢驗 OLS 與 Probit 兩種方法，其真實之顯著水準是否與檢定時使用之顯著水準無差異。所以希望在重複實驗 1,000 次後，所得拒絕虛無假設次數的比率（即真實之顯著水準）會與檢驗時使用之顯著水準（即名目顯著水準）無差異，若結果如此即表示該方法較佳。

既然本階段是研究當 H_0 爲真，而拒絕 H_0 之比率與名目 α 是否差異很大，在作計算 Z_i 及 Z_i^* 時，當然應以 $\beta=0$ 代入計算，此時之假設爲：

$$H_0 : \beta_i = 0$$

$$H_a : \beta_i \neq 0$$

累計拒絕 H_0 之次數除以總實驗次數（1,000 次），即爲真實之顯著水準。檢定結果如表六。

表六 顯著水準(α)檢定結果(真實 $\beta=0$)
(虛擬變數配置比率 0.6:0.4)

樣本量：50

名顯著水準	模型	個別變數係數之檢定的拒絕率					整模之檢定	個型
		總資產	負債比率	銷貨成長率	成立年數	虛擬變數		
1%	PT	0.0%***	0.1%***	0.2%**	0.6%	0.6%	2.1%***	
	OLS	0.3%**	0.7%	0.6%	1.1%	0.8%		
5%	PT	1.6%***	4.0%	3.1%***	5.2%	5.3%	8.3%***	
	OLS	2.2%***	4.3%	2.9%***	5.3%	5.0%		
10%	PT	6.0%***	10.0%	6.8%***	11.3%	11.4%	16.1%***	
	OLS	8.0%**	9.2%	7.6%**	10.9%	10.5%		

樣本量：100

名顯著水準	模型	個別變數係數之檢定的拒絕率					整模之檢定	個型
		總資產	負債比率	銷貨成長率	成立年數	虛擬變數		
1%	PT	0.0%***	0.5%	0.3%**	0.5%	1.1%	1.8%**	
	OLS	0.3%**	1.1%	0.5%	0.7%	1.2%		
5%	PT	2.7%***	4.1%	2.4%***	5.3%	3.7%	9.7%***	
	OLS	3.3%**	4.3%	3.2%***	5.2%	5.6%		
10%	PT	7.2%***	9.5%	7.0%***	10.9%	11.3%	17.7%***	
	OLS	9.6%	10.9%	7.9%**	10.8%	11.2%		

- 附註：1. *** 係表示拒絕率與顯著水準之差距在 $\alpha=1\%$ 為顯著者
 2. ** 係表示拒絕率與顯著水準之差距在 $\alpha=5\%$ 為顯著者
 3. * 係表示拒絕率與顯著水準之差距在 $\alpha=10\%$ 為顯著者
 4. PT = Probit

由表六中得知：在樣本量為 50 的情況下，整體而言，OLS 法的表現較 Probit 模型好些；就個別變數而言，總資產的表現均不甚理想，公司成立年數與虛擬變數二項，其拒絕率與檢定之顯著水準，並無顯著差異。在樣本量為 100 的情況下，就整個模型檢定結果而言，OLS 法的表現則明顯地比 Probit 模型好；至於個別變數，其結果仍與樣本量為 50 時相同。由此可知，當樣本量由 50 增加至 100 時，OLS 法的表現則明顯改善，但 Probit 模型卻無明顯的改善，其拒絕率仍過高。本研究之顯著水準檢定結果，與 Noreen(1988) 之結果相似。另本研究對虛擬變數配置比例作一極端配置 (0.9:0.1)，看是否對原先結果造成影響。

其結果彙總如表七及表八。

表七 顯著水準檢定結果之彙總表 (樣本量 = 50)
(真實 $\beta=0$)

期間	α	模型	虛擬變數配置比率 (0.6:0.4)					虛擬變數配置比率 (0.9:0.1)						
			X1	X2	X3	X4	X5	M	X1	X2	X3	X4	X5	M
76 至 80	1 %	PT	***	***	**			***	***		***		***	***
		OLS	**					*			*			
	5 %	PT	***		***			***	***		***		***	***
		OLS	***		***			***		*				**
	10 %	PT	***		***			***	*				***	***
		OLS	**		**			***						

- 附註：1. *** 係表示拒絕率與顯著水準之差距在 $\alpha=1\%$ 為顯著者
 ** 係表示拒絕率與顯著水準之差距在 $\alpha=5\%$ 為顯著者
 * 係表示拒絕率與顯著水準之差距在 $\alpha=10\%$ 為顯著者
2. X1：總資產 X4：公司成立年數
 X2：負債比率 X5：虛擬變數
 X3：銷貨成長率 M：整個模型

表八 顯著水準檢定結果之彙總表 (樣本量 = 100)
(真實 $\beta=0$)

期間	α	模型	虛擬變數配置比率 (0.6:0.4)					虛擬變數配置比率 (0.9:0.1)						
			X1	X2	X3	X4	X5	M	X1	X2	X3	X4	X5	M
76 至 80	1 %	PT	***		**			**	***	**	***		**	***
		OLS	**						***		**			**
	5 %	PT	***		***			***	***		***		***	**
		OLS	**		***				***		***			***
	10 %	PT	***		***			***	***	**	***			***
		OLS	**		**				**	***	***			***

- 附註：1. *** 係表示拒絕率與顯著水準之差距在 $\alpha=1\%$ 為顯著者
 ** 係表示拒絕率與顯著水準之差距在 $\alpha=5\%$ 為顯著者
 * 係表示拒絕率與顯著水準之差距在 $\alpha=10\%$ 為顯著者
2. X1：總資產 X4：公司成立年數
 X2：負債比率 X5：虛擬變數
 X3：銷貨成長率 M：整個模型

由表七得知：在樣本量為 50 的情況下，整個模型的表現仍是 OLS 法比 Probit 模型來的佳，至於個別變數，則可發現，當虛擬變數配置比率由 0.6:0.4 改變為 0.9:0.1 時，Probit 模型下虛擬變數的表現，明顯地變差。公司成立年數，仍是在此表現最佳的變數。

在表八中，樣本量增為 100 時，整個模型的表現，仍是 OLS 法的表現較好，虛擬變數的表現仍同樣本量為 50 時的情況，而銷貨成長率為表現最差的變數，公司成立年數則仍然表現最佳。此外，樣本量由 50 增加至 100 時，OLS 法的表現仍比 Probit 模型佳，但銷貨成長率的表現則出乎意料的變差了。

二、OLS 法與 Probit 模型之檢定力分析

本節之目的即為比較 OLS 法與 Probit 模型兩法之檢定力，何者較佳？但在進行檢定力比較之前，須先將各模型在不同檢定使用之顯著水準調整至真實顯著水準。例如，在檢定時， $N=50$ ，OLS 法下所謂 $\alpha=0.05$ 係指真實水準為 0.09；而 Probit 模型所謂 $\alpha=0.05$ 係指真實顯著水準為 0.04

(註四及註五)。此乃為控制住型 I 誤差的水準，使能真正地比較出 OLS 法與 Probit 模型之檢定力。

在計算 Z_i 及 Z^* 時，以 $\beta=0.35$ 為真實 β 代入計算，而檢驗之假設為：

$$H_0: \beta_i = 0$$

$$H_a: \beta_i \neq 0$$

其中虛無假設不實，所以拒絕的次數當是愈高愈好。結果如下：

註四：各情況下之真實 α 係利用逐一測試法獲得，其結果如下：

			OLS	Probit
0.6 :	N=50	1%	1%	1%
		5%	9%	4%
		10%	12%	6%
0.4 :	N=100	1%	1%	1%
		5%	6%	3%
		10%	12%	5%
0.9 :	N=100	1%	1%	1%
		5%	8%	3%
		10%	12%	7%

註五：由於 SAS 中，最小為 1%，故若其真實 $\alpha < 1\%$ ，則無法獲得。

表九 檢定力 ($1 - \beta$) 之檢定結果 (真實 $\beta = 0.35$)
(虛擬變數配置比率 0.6:0.4)

樣本量：50

顯著水準	模型	個別變數係數之檢定					整模之檢定	個型之檢定
		總資產	負債率	銷貨成長率	成立年數	虛擬變數		
1%	PT	1.0%	4.3%	2.5%	7.7%	15.7%	61.1%***	44.1%
	OLS	6.5%***	9.4%***	6.3%***	13.3%***	19.3%**		
5%	PT	8.0%	20.5%	13.0%	26.1%	37.0%	81.2%	82.1%
	OLS	34.3%***	36.8%***	32.8%***	42.8%***	50.6%***		
10%	PT	14.6%	27.9%	19.2%	33.6%	43.6%	85.9%	84.7%
	OLS	42.4%***	42.5%***	40.1%***	49.0%***	57.2%***		

樣本量：100

顯著水準	模型	個別變數係數之檢定					整模之檢定	個型之檢定
		總資產	負債率	銷貨成長率	成立年數	虛擬變數		
1%	PT	5.6%	21.6%	10.9%	34.4%	45.6%	96.4%***	91.5%
	OLS	15.5%***	24.8%*	16.8%***	38.5%*	47.5%		
5%	PT	18.7%	37.6%	22.1%	53.1%	63.2%	98.8%	98.8%
	OLS	49.5%***	48.5%***	40.5%***	68.2%***	74.2%***		
10%	PT	29.0%	45.8%	28.7%	62.1%	72.6%	99.2%	99.6%
	OLS	67.9%***	64.4%***	55.1%***	78.2%***	85.1%***		

附註：1. *** 係表示拒絕率與顯著水準之差距在 $\alpha = 1\%$ 顯著且較高者
 2. ** 係表示拒絕率與顯著水準之差距在 $\alpha = 5\%$ 顯著且較高者
 3. * 係表示拒絕率與顯著水準之差距在 $\alpha = 10\%$ 顯著且較高者

表十 檢定力 ($1 - \beta$) 之檢定結果 (真實 $\beta = 0.35$)
(虛擬變數配置比率 0.9:0.1)

樣本量：100

顯著水準	模型	個別變數係數之檢定					整體模型之檢定
		總資產	負債率	銷貨成長率	成年數	虛擬變數	
1%	PT	5.1%	19.8%	8.9%	32.6%	8.0%	96.5%***
	OLS	12.6%***	24.0%**	12.8%***	38.2%***	31.4%***	90.6%
5%	PT	16.3%	38.3%	18.4%	52.2%	30.4%	98.2%
	OLS	56.3%***	53.6%***	44.2%***	72.8%***	73.0%***	98.8%
10%	PT	37.3%	52.2%	34.3%	68.0%	55.1%	99.4%
	OLS	67.6%***	62.6%***	54.5%***	80.0%***	81.2%***	99.4%

- 附註：1. *** 係表示拒絕率與顯著水準之差距在 $\alpha = 1\%$ 顯著且較高者
 2. ** 係表示拒絕率與顯著水準之差距在 $\alpha = 5\%$ 顯著且較高者
 3. * 係表示拒絕率與顯著水準之差距在 $\alpha = 10\%$ 顯著且較高者

由表九中得知：在樣本量為 50 時，除在顯著水準 = 1% 時，Probit 模型之檢定力略較 OLS 法來的高外，其餘顯著水準下，二法的檢定力並無顯著差異。在樣本量為 100 時，其結果仍同，除在顯著水準 = 1% 時，Probit 模型略佳外，其餘二法則無顯著差異。但當樣本量由 50 增加至 100 時，則發現二法的檢定力均有增加，且二法之檢定力都相當地高。本研究檢定力分析之結論，與 Noreen(1988) 一致。但本研究之檢定力不論 OLS 法或 Probit 模型，均較 Noreen(1988) 結果之檢定力高。表十中所列結果係將虛擬變數配置比率由 0.6:0.4 改至 0.9:0.1 時，其結論並無大改變。總而言之，就檢定力而言，OLS 法的表現與 Probit 模型，相差不遠。

三、OLS 與 Probit 模型預測能力之比較

一個好的模型除了能解釋資料結構外，更重要的是能預測未來情況，一個較能準確預測未來情況的模型，對使用者而言，才是較佳的模型。故本節即對 OLS 法與 Probit 模型兩法之預測能力作一模擬，以比較其結果。

在第三節中曾介紹，本研究共採五種不同 A 值，來判斷 \hat{Y}_i 的數值，此

五種不同標準分別為 0.5, 0.6, 0.7, 0.8, 0.9, 若大於此 A 值則 $\hat{Y}_i=1$, 若小於 $1-A$, 則 $\hat{Y}_i=0$ 。

表十一則為 $A=0.5 \sim 0.9$, 不同樣本量 (50,100), 不同虛擬變數配置比率 (0.6:0.4, 0.9:0.1) 的情形下, 預測能力之結果。由表中可發現, OLS 法的預測正確率略較 Probit 模型的正確率高。

表十一 預測能力檢定—正確率之彙總表

樣 本 量	期 間	模 型	虛 擬 變 數 配 置 比 率 (0.6:0.4)					虛 擬 變 數 配 置 比 率 (0.9:0.1)				
			0.5 VS 0.5	0.6 VS 0.4	0.7 VS 0.3	0.8 VS 0.2	0.9 VS 0.1	0.5 VS 0.5	0.6 VS 0.4	0.7 VS 0.3	0.8 VS 0.2	0.9 VS 0.1
			$N=50$	76 80	PT OLS	67.6% 67.8%	71.9% 72.9%	76.6% 79.0%***	81.7% 83.7%**	87.0% 87.5%	64.6% 65.5%	69.9% 71.1%
$N=100$	76 80	PT OLS	67.7% 67.8%	72.6% 74.5%***	79.1% 81.7%***	83.2% 86.9%***	88.2% 91.9%***	65.8% 65.9%	71.5% 73.4%***	77.5% 81.2%***	84.4% 88.2%***	89.9% 92.2%***

附註：

1. *** 係表示兩法正確率之差距在 $\alpha=1\%$ 時為顯著且較高者
2. ** 係表示兩法正確率之差距在 $\alpha=5\%$ 時為顯著且較高者
3. * 係表示兩法正確率之差距在 $\alpha=10\%$ 時為顯著且較高者

伍、結論與建議

本研究共分三部份來探討 OLS 法與 Probit 模型，在處理二分因變數的資料時，其結果是否有差異。這三部份之結果彙總如下：

1. OLS 法的真實顯著水準，與檢驗時所使用的顯著水準，較為接近，即 OLS 法導出的統計量實際分配 (empirical distributions) 與理論分配 (theoretical distributions)，較為接近。

2. 就檢定力而言，Probit 模型的表現略優於 OLS，但當顯著水準增大時，Probit 模型的表現則與 OLS 法無差異。整體而言，二法之檢定力並無太大差異，且二法之檢定力均相當高。

3. 在預測正確率的比較上，OLS 大致比 Probit 模型來得好。

由上述之結果可發現，OLS 法在處理二分因變數的資料時，其表現並不比 Probit 模型差，至於預測能力上則是 OLS 法優於 Probit 模型。結果顯示：若以國內市場實際資料作二分性質的研究，且樣本為小樣本時，OLS 法可能比 Probit 模型好。

本研究對於解釋變數之選用，係以國內曾經使用 Probit 模型或 Logit 模型之研究，所選用的變數較頻繁者。但就國內研究而言，至目前為止，使用 Probit 模型或 Logit 模型之財務或會計的研究仍不多，故本研究所選用變數的代表性，可能不是那麼高。若日後使用 Probit 模型或 Logit 模型之研究增多了，後續研究者或許可選取到更具代表性的研究變數，再進行研究。

本研究所使用之電腦軟體為 SAS 套裝軟體，而 SAS 在使用上有些限制。如對殘差項的分配，SAS 僅能假設為常態分配，但或許偏態分配的假設，能更符合真實狀況，故日後若對本研究主題有興趣者，不妨可使用其他軟體，使其殘差項之分配能符合偏態分配，再檢驗其結果是否與本研究之結果不同。

參考文獻

- 沈遠芳，民 81，會計師業務評核結果與審計品質替代性指標之分析，國立台灣大學會計學研究所未出版碩士論文。
- 施春成，民 81，台灣新上市公司特有風險與會計師選擇之研究，國立台灣大學會計學研究所未出版碩士論文。
- 彭秋萍，民 81，台灣上市公司損益平穩化行為之動機研究，國立台灣大學會計學研究所未出版碩士論文。
- 程佩瑜，民 81，我國會計師財務及融資簽證市場之研究，國立台灣大學會計學研究所未出版碩士論文。
- 鄧 洪，民 81，會計師保留意見之研究，國立台灣大學會計學研究所未出版碩士論文。
- 蕭博仁，民 81，我國股票發行市場上市審查之研究，國立台灣大學商學研究所未出版碩士論文。
- 蘇裕惠，民 80，台灣會計師界使用電腦輔助審計技術之研究，國立台灣大學商學研究所未出版碩士論文。
- Maddalla, G. S. 1983. *Limited dependent and qualitative variables in econometrics*. New York : Cambridge University Press.**
- Maddalla, G. S. 1991. A Perspective on the use of limited-dependent and qualitative variables models in accounting research. *The Accounting Review* Vol.66 (October) pp.788-807.**
- McFadden, , D. 1982. Conditional logit analysis of qualitative choice behavior *Frontiers in Econometrics*, edited by P. Zarembka, 105-142. New York: Academic Press.**
- Noreen, E. 1988. An empirical comparison of Probit and OLS regression hypothesis tests. *Journal of Accounting Research* 26 (Spring): 119-33.**