

第五章 預測效力比較

我們基於第二章第五節所介紹的 ROC 以及 CAP 曲線作為預測效力來作 Logit 模型和 Binary Regression Quantiles 模型的預測解釋做一比較。

我們運用 Grigorios Kordas(2004)的方法，來計算 Binary Regression Quantiles 所預測的違約機率，在以 ROC 與 CAP 的信用風險模型驗證。

則我們的應變數 y 為二元變數，例如違約與不違約的二元結果，而 y 是假設由可觀察到的一個隨機向量 x 所決定，而假設 $x \in R^K$ ，除了 x 的影響外， y 也受到無法觀測的分配 u 所影響，假設 $u \in R^1$ ，而一組 (x, u) 透過線性方程式影響 y ，該方程式設定如下：

$$y^* = x' \beta + u \quad (1)$$

(1) 式中， β 為一 $k \times 1$ 的參數向量。而我們再設定以下方程式：

$$y = 1\{x' \beta + u \geq 0\} \quad (2)$$

(2) 式中， $1\{A\}$ 為一 A 事件發生與否的指示函數，如果指示函數等於 1 代表 A 發生，如果指示函數等於 0，則代表 A 事件沒有發生。

在經濟的應用上， y 是個人在兩個選擇下 $\{1, 0\}$ 可觀察到的效用極大函數，(1) 式中的 y^* 是個人在選擇 1 或選擇 0 的效用函數相減的結果， $y^* = V^1 - V^0$ 。

即 $y^* \geq 0$ 則代表 $y=1$ ，相反 $y^* \leq 0$ 則代表 $y=0$ 。

對於每一個 $x \in X$ ， $\{y=1\}$ 在 $y=1$ 為信用卡違約時的情況，其機率為

$$P_{1|x} \equiv \Pr(y = 1 | x), \text{ 即 } \hat{P}_{1|x} = M^{-1} \sum_{m=1}^M 1\{x' b_N(\tau_m) / \|b_N(\tau_m)\| \geq 0\} \quad (3)$$

(3) 式中， $F_{u|x}$ 為在已知 x 的條件下， u 的分配函數。而這也是我們有興趣的機率

值，假設 τ 代表不同分量，我們將(3)式改寫為：

$$P_{1|x} = \int_0^1 1\{x'\beta(\tau) \geq 0\} d\tau \quad (4)$$

在不影響正負號下，我們將各分量的係數標準化，故將 $\beta(\tau)$ 以 $\beta(\tau)/\|\beta(\tau)\|$ 取

代，(4)式可寫為：

$$\hat{P}_{1|x} = M^{-1} \sum_{m=1}^M 1\{x'b_N(\tau_m)/\|b_N(\tau_m)\| \geq 0\} \quad (5)$$

(4) 式中的 $b_N(\tau_m)/\|b_N(\tau_m)\|$ 為標準化的 Binary Regression Quantiles 係數估計

值，而總共有 m 個分量。

由本文可利用 $b_N(\tau_m)/\|b_N(\tau_m)\|$ 求算出下表在各分量的係數的權數標準化

【表 5-1】 各分量的係數的權數標準化

各分量的係數	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
年收入	0.000	0.001	0.008	0.002	0.003	0.001	0.005	0.004	0.006
職級	0.302	0.353	-0.244	0.409	0.153	0.032	-0.602	0.509	-0.244
婚姻	0.522	0.280	0.392	-0.435	-0.406	0.352	0.242	0.315	0.553
每月限額	-0.274	-0.300	-0.323	-0.219	-0.253	-0.002	-0.011	-0.007	-0.012
6 個月交易金額	-0.174	-0.211	-0.388	-0.299	-0.481	-0.019	-0.276	-0.082	-0.046
過去 6 個月平均繳款金額	-0.519	-0.589	-0.506	-0.483	-0.479	-0.688	-0.638	-0.520	-0.554
6 個月平均期初應繳金額	0.343	0.375	0.384	0.316	0.386	0.041	0.083	0.047	0.055
過去 6 個月是否有循環息	-0.329	0.414	-0.248	-0.224	0.039	0.631	-0.061	0.453	0.161
過去 6 個月繳款金額/應繳金額	0.188	0.064	0.257	0.351	-0.371	-0.036	-0.292	0.397	-0.544

我們將使用(5)式做為違約估計值的計算，將【表 5-1】各分量係數的估計結果代

入(5)式，且計算出各分量每一個的觀察值違約機率，故本文使用十個分量

(0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9)且每個分量有 9 個解釋變數與 13123 的觀察值。

【表 5-2】 計算出平均分量的違約機率 \hat{P}_{lx} 與實際違約機率相比較

	0.1 分量	在 0.1 分量下是否違約(P_1)	0.9 分量	在 0.9 分量下是否違約(P_2)	平均分量的違約機率 \hat{P}_{lx}	實際違約機率
13123 個觀察值	利用(5)公式算出來的機率	可算出 1 或 0(1 為違約, 0 為不違約)	利用(5)公式算出來的機率	可算出 1 或 0(1 為違約, 0 為不違約)	把 $(P_1+P_2+P_3+P_4+P_5+P_6+P_7+P_8+P_9)/9$	0:無違約者 1:違約者

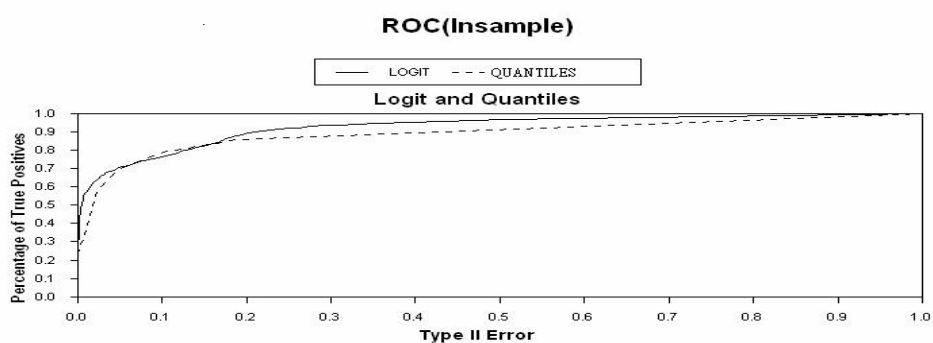
我在利用上面作出「平均分量的違約機率」與「實際違約機率」來劃出 ROC 與

CAP 的預測違約機率曲線圖並與 Logit 迴歸所求出的預測機率作比較。

第一節 利用 ROC 曲線與 CAP 曲線作樣本內的 Binary Quantiles Regression 與 Logit Regression 違約機率預測

壹、利用 ROC 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機率預測：ROC 曲線中 Logit 包含的面積比例為 0.92373 且 Binary Quantiles Regression 包含的面積比例為 0.89070。我們可以看出在樣本內兩模型其預測效力很好且 ROC 愈往左上，則代表模型的預測效力越好。

【圖 5-1】樣本內 ROC 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機率預測



貳、利用 CAP 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機率預測：CAP 曲線中 Logit regression 包含的面積比例為 0.86574 且 Binary Quantiles Regression 包含的面積比例為 0.85571。我們可以看出在樣本內 Binary Quantiles Regression 和傳統下的 Logit regression, 兩模型其預測效力都很好且 CAP 愈往左上，則代表模型的預測效力越好。

Cut off	LOGIT	QUANTILES
0.0000	0.0000	0.0000
0.0100	0.0697	0.0708
0.0200	0.1405	0.1416
0.0300	0.2108	0.2097
0.0400	0.2773	0.2686
0.0500	0.3454	0.3076
0.0600	0.4119	0.3546
0.0700	0.4724	0.4092
0.0800	0.5281	0.4605
0.0900	0.5746	0.5189
0.1000	0.6092	0.5741
0.1100	0.6384	0.6011
0.1200	0.6654	0.6346
0.1300	0.6859	0.6719
0.1400	0.7005	0.6968
0.1500	0.7130	0.7054
0.1600	0.7259	0.7114
0.1700	0.7395	0.7254
0.1800	0.7514	0.7389
0.1900	0.7611	0.7627
0.2000	0.7719	0.7914
0.2100	0.7827	0.7951
0.2200	0.7978	0.8022
0.2300	0.8065	0.8097
0.2400	0.8168	0.8189
0.2500	0.8281	0.8270
0.2600	0.8416	0.8362
0.2700	0.8562	0.8422
0.2800	0.8719	0.8492
0.2900	0.8822	0.8600
0.3000	0.8930	0.8686
0.3100	0.9016	0.8746
0.3200	0.9054	0.8832
0.3300	0.9086	0.8870
0.3400	0.9162	0.8930
0.3500	0.9189	0.8962

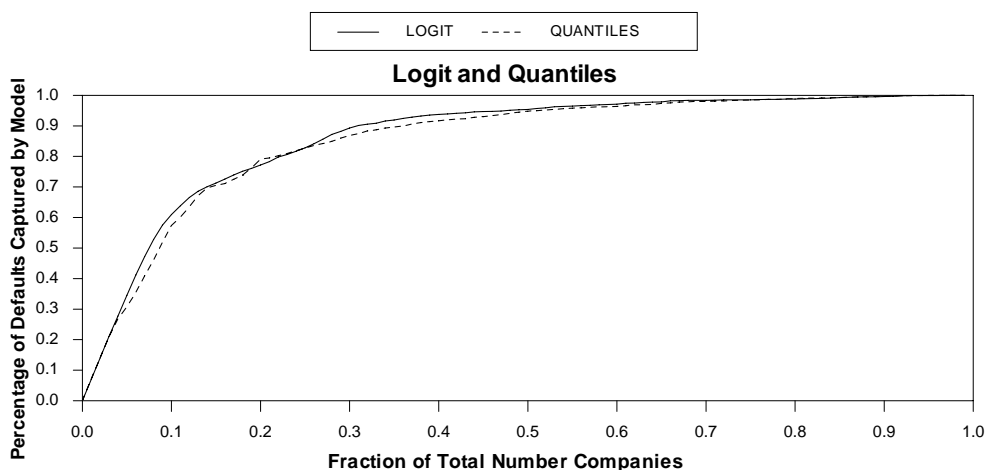
0.3600	0.9243	0.8995
0.3700	0.9286	0.9059
0.3800	0.9324	0.9108
0.3900	0.9351	0.9141
0.4000	0.9373	0.9168
0.4100	0.9389	0.9195
0.4200	0.9411	0.9222
0.4300	0.9427	0.9232
0.4400	0.9459	0.9281
0.4500	0.9470	0.9303
0.4600	0.9476	0.9335
0.4700	0.9486	0.9368
0.4800	0.9514	0.9411
0.4900	0.9524	0.9449
0.5000	0.9535	0.9481
0.5100	0.9568	0.9497
0.5200	0.9589	0.9519
0.5300	0.9627	0.9541
0.5400	0.9638	0.9562
0.5500	0.9649	0.9578
0.5600	0.9665	0.9589
0.5700	0.9676	0.9611
0.5800	0.9692	0.9622
0.5900	0.9697	0.9627
0.6000	0.9708	0.9638
0.6100	0.9741	0.9659
0.6200	0.9746	0.9692
0.6300	0.9768	0.9692
0.6400	0.9778	0.9703
0.6500	0.9784	0.9735
0.6600	0.9822	0.9751
0.6700	0.9827	0.9778
0.6800	0.9838	0.9789
0.6900	0.9838	0.9800
0.7000	0.9838	0.9800
0.7100	0.9843	0.9811
0.7200	0.9849	0.9816

0.7300	0.9849	0.9832
0.7400	0.9854	0.9838
0.7500	0.9854	0.9849
0.7600	0.9859	0.9854
0.7700	0.9865	0.9865
0.7800	0.9870	0.9865
0.7900	0.9876	0.9886
0.8000	0.9876	0.9892
0.8100	0.9886	0.9908
0.8200	0.9892	0.9908
0.8300	0.9897	0.9914
0.8400	0.9908	0.9924
0.8500	0.9924	0.9930
0.8600	0.9935	0.9935

0.8700	0.9946	0.9946
0.8800	0.9951	0.9951
0.8900	0.9951	0.9962
0.9000	0.9962	0.9968
0.9100	0.9968	0.9973
0.9200	0.9984	0.9989
0.9300	0.9995	0.9989
0.9400	0.9995	0.9995
0.9500	0.9995	1.0000
0.9600	0.9995	1.0000
0.9700	1.0000	1.0000
0.9800	1.0000	1.0000
0.9900	1.0000	1.0000
1.0000	1.0000	1.0000

【圖 5-2】樣本內 CAP 曲線作 Binary Quantiles Regression 與 Logit Regression 違約

機率預測



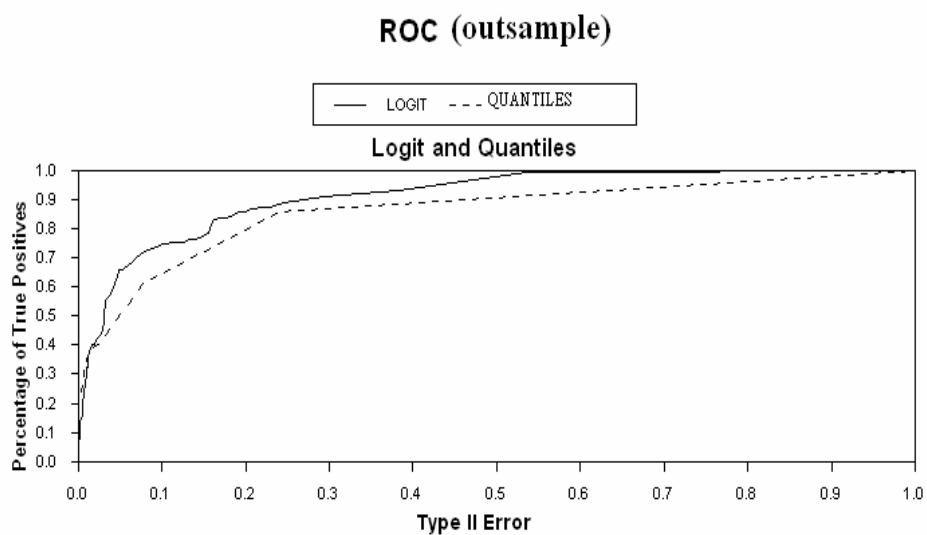
由於本文的資料高達 13123 筆，在每一分量分量中作一次信用卡的預測，然後在把每個分量所得的結果加權平均，再與實際違約作比較。然而資料高達 1 萬多

筆，則使得預測能與實際違約機率差異不大。

第二節 利用 ROC 曲線與 CAP 曲線作樣本外的 Binary Quantiles Regression 與 Logit Regression 違約機率預測

壹、利用 ROC 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機率預測：ROC 曲線中 Logit 包含的面積比例為 0.91041 且 Binary Quantiles Regression 包含的面積比例為 0.85940。我們可以看出在樣本內兩模型其預測效力很好且 ROC 愈往左上，則代表模型的預測效力越好。

【圖 5-3】樣本外 ROC 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機率預測



貳、利用 CAP 曲線作 Binary Quantiles Regression 與 Logit Regression 違約機

率預測：CAP 曲線中 Logit regression 包含的面積比例為 0.56163 且 Binary

Quantiles Regression 包含的面積比例為 0.54491。我們可以看出在樣本內 Binary

Quantiles Regression 和傳統下的 Logit regression, 兩模型其預測效力都很好且

CAP 愈往左上，則代表模型的預測效力越好。

【圖 5-4】樣本外 CAP 曲線作 Binary Quantiles Regression 與 Logit Regression 違約

機率預測

