

조정된 ROC와 CAP 곡선

홍중선¹ · 김지훈² · 최진수³

¹성균관대학교 경제학부, ²성균관대학교 응용통계연구소, ³성균관대학교 응용통계연구소

(2008년 8월 접수, 2008년 11월 채택)

요약

신용평가 연구에서 부도율분포를 기반으로 부도기업과 정상기업의 판별력을 탐색하는 방법 중의 하나로 ROC와 CAP 곡선을 사용한다. 부도와 정상기업을 분류하는 절단점의 변동에 따라 구한 여러 부도비율을 통해 ROC와 CAP 곡선을 작성하는데, 곡선의 각 좌표에 대응하는 절단점을 탐색하기 어렵다. 본 연구에서는 ROC와 CAP 곡선을 나타내는 부도비율들의 함수를 이용하여 조정된 ROC와 CAP 곡선을 제안한다. 조정된 ROC와 CAP 곡선을 통해 절단점과의 관계를 파악할 수 있으며, 최적의 절단점을 식별할 수 있다. 또한 부도와 정상기업에 관한 분포함수의 동일성을 검정하는 Kolmogorov - Smirnov 통계량과 조정된 ROC와 CAP 곡선을 통해 얻은 최적 절단점의 관계를 토론한다.

주요용어: 부도율, 신용평가, 스코어, 오즈비, 절단점, K-S 통계량, 판별력.

1. 서론

확률표본 $\{X_1, X_2, \dots, X_n\}$ 의 확률밀도함수 $f(x)$ 가 다음과 같이 확률밀도함수 $f^1(x)$ 와 $f^2(x)$ 로 분할되었다고 가정하자:

$$f(x) = \alpha f^1(x) + (1 - \alpha) f^2(x),$$

여기서 $\alpha \in (0, 1)$ 이며 $x \in (-\infty, \infty)$ 이다. 이와 같은 상황은 신용평가에 관한 부도율(probability of default)의 연구에서 많이 찾아볼 수 있다. 확률변수 X 를 스코어라 하고, $f^1(x)$ 과 $f^2(x)$ 를 각각 $f_D(x)$ 와 $f_N(x)$ 으로 정의하여 부도(default)기업과 정상(non-default)기업의 부도율분포를 나타낸다고 가정하자. 또한 스코어의 누적분포함수를 $F(x)$ 로 정의하고 부도기업과 정상기업의 누적분포함수를 $F_D(x)$ 와 $F_N(x)$ 로 정의하면, 스코어의 누적분포함수는 다음과 같이 분할된다.

$$F(x) = \alpha F_D(x) + (1 - \alpha) F_N(x), \quad (1.1)$$

여기서 α 는 부도율총합(total portfolio-wide probability of default)이다.

ROC(Receiver Operating Characteristic) 곡선과 CAP(Cumulative Accuracy Profile) 곡선은 신용평가에서 예측력(prediction power)과 판별력(discriminatory power)을 탐색하는데 유용한 그래픽적인 방법이다. ROC 함수식과 CAP 함수식은 식 (1.1)에서의 분포함수를 이용하여 다음과 같이 정의한다 (Tasche, 2006).

$$\text{ROC}(u) = F_D(F_N^{-1}(u)), \quad u \in (0, 1),$$

$$\text{CAP}(u) = F_D(F^{-1}(u)), \quad u \in (0, 1).$$

¹교신저자: (110-745) 서울시 종로구 명륜동 3-53, 성균관대학교 경제학부 통계학전공, 교수.

E-mail: cshong@skku.ac.kr

표 1.1. 혼동행렬

		실제상황(actual class)	
		P (실제부도)	N (실제정상)
예측변수 (predicted class)	Y (부도로 예측)	True Positive	False Positive
	N (정상으로 예측)	False Negative	True Negative
합		p	n

ROC 곡선은 $(u, ROC(u))$, $u \in (0, 1)$ 또는 $(F_N(x), F_D(x))$, $x \in R$ 에 대응하는 점으로 작성하며, CAP 곡선은 $(u, CAP(u))$, $u \in (0, 1)$ 또는 $(F(x), F_D(x))$, $x \in R$ 에 대응하는 점으로 작성한다. 이를 이차원 분류 문제(classification problem)로 고려하여보자. 분류모형(classification model)은 실제상황(instance, actual class)을 예측변수(predicted class)로 표현하는 함수이다. 실제상황과 예측변수는 각각 $\{P, N\}$ 과 $\{Y, N\}$ 의 범주 수준을 갖고 있다고 하자. 그러면 표 1.1과 같은 네 개의 결과를 갖는 2×2 분할표 또는 혼동행렬(confusion matrix)형태로 나타난다 (Fawcett, 2004, 2005; Sonogo 등, 2008).

각각의 절단점(cutting point)의 변화에 따라 구한 표 1.1에서 tp 비율(true positive rate; hit rate 또는 recall이라고도 함), fp 비율(false positive rate; false alarm rate라고도 함) 그리고 부도율은 다음과 같이 정의한다.

$$tp\text{비율} = \frac{TP}{p},$$

$$fp\text{비율} = \frac{FP}{n},$$

$$\text{부도율} = \frac{TP + FP}{p + n},$$

여기서 ROC 곡선의 X축과 Y축 좌표는 각 절단점에서 구한 $\hat{F}_N(\cdot) = fp$ 비율과 $\hat{F}_D(\cdot) = tp$ 비율이고, CAP 곡선의 X축과 Y축 좌표는 $\hat{F}(\cdot) = \text{부도율}$ 과 $\hat{F}_D(\cdot) = tp$ 비율이다. 따라서 ROC 곡선과 CAP 곡선은 모두 X축이 동일한 점에서는 실제 부도를 부도로 예측하는 확률이 높을수록 좋은 모형이라 할 수 있으며, 곡선 밑의 궤적이 대각선에서 멀리 위치할수록 ROC 곡선에서는 예측력이 그리고 CAP 곡선에서는 판별력이 높다고 한다. ROC 곡선 아래의 면적을 AUROC(area under ROC)라고 하는데, ROC 곡선의 사용 목적은 최적 절단점을 발견하는데 있고, AUROC의 사용 목적은 여러 절단점을 모두 고려했을 때 척도들의 전반적인 유용성을 비교하는데 있다 (Hosmer와 Lemeshow, 2000; 임종건, 2005; 이근희, 2006). 이차원 그래프로 표현되는 ROC 곡선과 CAP 곡선은 각 절단점에서 얻는 비율들로 이루어져있는데 각 곡선의 점을 통해서 절단점을 탐색하기 어렵다는 문제점이 있다. 본 연구에서는 이러한 문제점을 개선하는 새로운 곡선을 제안한다.

본 연구의 구성은 다음과 같다. 2절에서는 ROC와 CAP 곡선을 구성하는 fp 비율, tp 비율 그리고 부도율의 함수로 오즈비(odds ratio)를 생성하고, 스코어와 오즈비로 표현하는 조정된 ROC와 CAP 곡선을 제안한다. 제안한 곡선에 관하여 이론적 배경과 특성을 설명하고 ROC와 CAP 곡선과 비교하면서 Kolmogorov - Smirnov(K-S) 통계량과의 관계를 유도한다. 3절에서는 모의실험을 통하여 정상과 부도 기업의 자료를 생성하고, 정상기업과 부도기업을 잘못 분류하는 오분류율에 따라 자료를 혼합한다. 모의실험으로 얻은 혼합자료를 바탕으로 조정된 ROC와 CAP 곡선의 최대값에 대응하는 스코어를 구한

후 이 스코어와 K-S 통계량에 대응하는 스코어가 일치함을 살펴보고, 모든 스코어가 일치하는 경우의 자료 특성을 정리한다. 4절에서 실제 신용평가모형의 자료를 바탕으로 조정된 ROC와 CAP 곡선을 작성하고, ROC와 CAP 곡선과 비교하며 K-S 통계량과의 관계를 설명한다. 그리고 5절에서 결론을 유도하고 향후연구를 제안한다.

2. 조정된 ROC와 CAP 곡선

ROC 곡선은 각 절단점에서 식 (1.2)의 tp 비율과 식 (1.3)의 fp 비율에 대한 곡선이므로 tp 비율과 fp 비율의 오즈비(odds ratio)를 OR_{AROC} 라고 하고 다음과 같이 정의한다.

$$OR_{AROC} = \frac{TN/FP}{FN/TP} = \frac{TP \cdot TN}{FP \cdot FN}. \quad (2.1)$$

OR_{AROC} 는 표 1.1의 혼동행렬에서 대각간의 곱의 비율로 나타나기 때문에 교차적비(cross-product ratio)라고도 할 수 있다. 또한 CAP 곡선은 각 절단점에서 구한 식 (1.4)의 부도율과 식 (1.2)의 tp 비율에 대한 곡선이므로 부도율과 tp 비율의 오즈비를 OR_{ACAP} 이라 하고 다음과 같이 정의한다.

$$OR_{ACAP} = \frac{(TN + FN)/(TP + FP)}{FN/TP} = \frac{TP \cdot (TN + FN)}{(TP + FP) \cdot FN}. \quad (2.2)$$

각 절단점 또는 스코어를 X 축에 그리고 식 (2.1)의 OR_{AROC} 와 식 (2.2)의 OR_{ACAP} 을 Y 축 좌표로 각각 표현한 그래프를 조정된 ROC와 CAP 곡선(adjusted ROC and CAP curves)이라하고 각 곡선의 이름을 ‘AROC’ 곡선과 ‘ACAP’ 곡선으로 제안한다.

AROC와 ACAP 곡선을 구현하기 위하여 임종건 (2005)의 자료를 참고한 표 2.1의 가상자료를 고려하자. 표 2.1은 부도율이 20%인 50개의 기업(정상기업: 40개, 부도기업: 10개)에 대한 스코어를 21에서 70까지로 50개로 가정한다. 첫 번째 열은 평가모형에 의해 얻어진 스코어으로써 각각의 스코어를 하나의 기업으로 간주한다. 스코어 31을 정상과 부도기업의 판별기준으로 정하고, 31미만의 부도기업의 오분류를 22, 24, 26으로 그리고 31이상의 정상기업의 오분류를 32, 34, 36으로 각 3개씩의 오분류를 임의로 설정하여 두 번째 열에 표현하였다.

AROC와 ACAP 곡선을 작성하기 위하여 각각의 스코어를 절단점으로 간주하여 TP, FN, FP, TN 의 값과 이들의 함수로 정의된 식 (2.1)과 (2.2)의 오즈비를 구하였다. 여기서 FP 와 FN 의 값이 0일 경우 OR_{AROC} 와 OR_{ACAP} 의 계산이 가능하도록 0.1로 대체하였다. 또한 K-S 통계량을 구하기 위하여 부도와 정상기업의 누적분포함수인 tp 비율과 fp 비율을 각각 구하고 이 값의 차이를 k_{s} 로 나타내었으며 이 값이 최대일 때가 K-S 통계량이 된다 (자세한 K-S 통계량의 설명은 송문섭 등 (2003) 참조). 따라서 표 2.1을 이용하면 AROC와 ACAP 곡선뿐만 아니라 ROC와 CAP 곡선도 작성할 수 있고, 부도와 정상기업의 누적분포함수를 작성할 수 있으며 K-S 통계량에 관한 정보도 얻을 수 있다.

표 2.1을 바탕으로 AROC와 ACAP 곡선을 그림 2.1에 구현하였다. 그림 2.1의 두 곡선 중에서 윗 곡선이 AROC 곡선이고 아래에 표현한 것이 ACAP 곡선이다. AROC와 ACAP 곡선의 최대값(peak)에 대응하는 스코어가 일치하는 것을 발견하였다. 이 스코어는 36으로 마지막 부도가 발생한 스코어이다.

식 (2.1)에서 정의한 오즈비 OR_{AROC} 의 값이 최대가 되기 위해서는 TN/FP (정상을 정상으로 예측하는 비율)가 큰 경우인 $FP = 0$ (0.1로 대체)인 경우이거나 또는 FN/TP (부도를 정상으로 예측하는 비율)가 작은 경우인 $FN = 0$ (0.1로 대체)인 경우이므로 OR_{AROC} 의 분모에 있는 FN 과 FP 의 값에 민감하다. 또한 AROC 곡선에서의 OR_{AROC} 가 최대화되는 경우에는 FN 과 FP 가 작고 TP 와 TN 이 큰 경우이므로 ROC 곡선이 예측력을 측정하는 것과 동일하게 모형의 예측력(prediction power)이 좋다고 평가할 수 있다.

표 2.1. AROC와 ACAP 곡선 계산과정

스코어	부도	TP	FN	FP	TN	OR _{AROC}	OR _{ACAP}	tp비율	fp비율	k _s
21	1	1	9	0.1	40	44.4	4.9	10	0.25	9.75
22	0	1	9	1	39	4.3	2.7	10	2.5	7.5
23	1	2	8	1	39	9.8	3.9	20	2.5	17.5
24	0	2	8	2	38	4.8	2.9	20	5	15
25	1	3	7	2	38	8.1	3.9	30	5	25
26	0	3	7	3	37	5.3	3.1	30	7.5	22.5
27	1	4	6	3	37	8.2	4.1	40	7.5	32.5
28	1	5	5	3	37	12.3	5.3	50	7.5	42.5
29	1	6	4	3	37	18.5	6.8	60	7.5	52.5
30	1	7	3	3	37	28.8	9.3	70	7.5	62.5
31	0	7	3	4	36	21.0	8.3	70	10	60
32	1	8	2	4	36	36.0	12.7	80	10	70
33	0	8	2	5	35	28.0	11.4	80	12.5	67.5
34	1	9	1	5	35	63.0	23.1	90	12.5	77.5
35	0	9	1	6	34	51.0	21.0	90	15	75
36	1	10	0.1	6	34	566.7*	213.1*	100	15	85*
37	0	10	0.1	7	33	471.4	194.7	100	17.5	82.5
38	0	10	0.1	8	32	400.0	178.3	100	20	80
39	0	10	0.1	9	31	344.4	163.7	100	22.5	77.5
40	0	10	0.1	10	30	300.0	150.5	100	25	75
41	0	10	0.1	11	29	263.6	138.6	100	27.5	72.5
42	0	10	0.1	12	28	233.3	127.8	100	30	70
43	0	10	0.1	13	27	207.7	117.8	100	32.5	67.5

이하중략

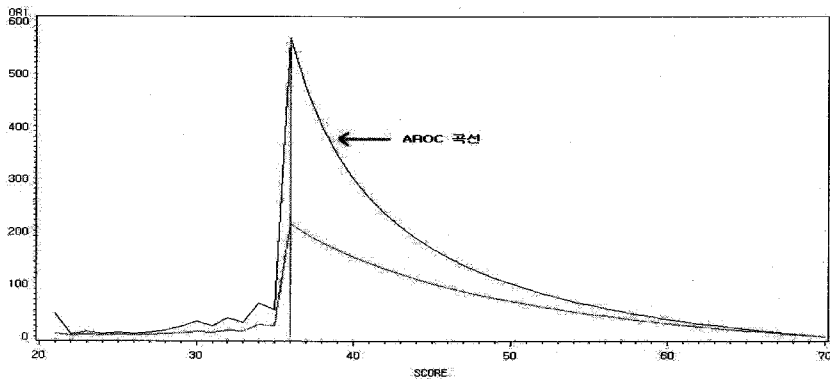


그림 2.1. AROC와 ACAP 곡선

OR_{AROC}에서 FN과 FP의 비용함수(cost function)를 각각 C(FN)과 C(FP)라 하면, C(FN)과 C(FP)가 모두 작은 값을 가지는 경우에 기대 비용함수(expected cost function)는 최소가 된다. FN과 FP이 0값에 수렴하는 경우에 OR_{AROC}가 최대화되고 기대 비용함수는 최소가 된다. 기대 비용함수가 최소가 되는 점을 최적의 절단점(optimal cut-off value)이라고 할 수 있으므로 AROC 곡선에서 OR_{AROC}가 최대일 때의 스코어를 최적의 절단점이라고 판단할 수 있다.

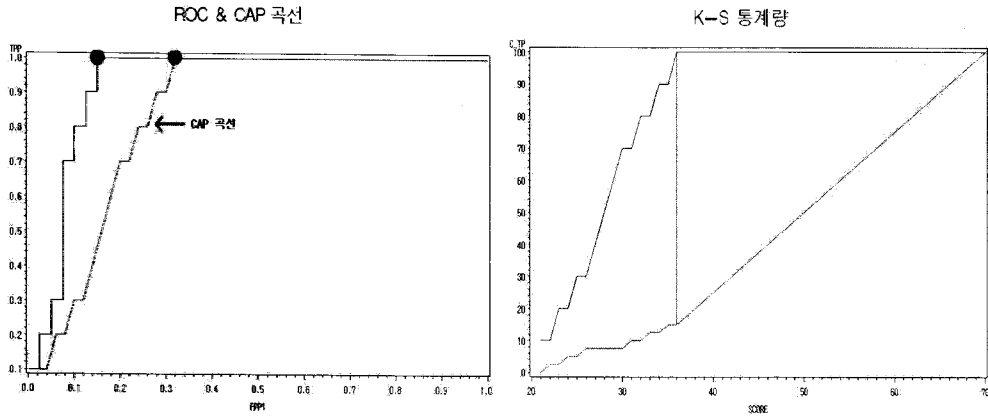


그림 2.2. ROC와 CAP 곡선 그리고 K-S 통계량과의 관계

오즈비 OR_{ACAP} 의 분모는 $(TP + FP) \cdot FN$ 이므로 OR_{ACAP} 이 큰 값을 갖는 경우는 FN 이 0에 수렴하는 경우이다. 따라서 OR_{ACAP} 은 오직 FN 의 값에 민감하게 반응한다. FN 이 0일 때의 의미는 마지막 부도가 발생하는 경우이고 이때에 OR_{ACAP} 은 최대값을 갖는다. 그러므로 ACAP 곡선에서 OR_{ACAP} 이 최대화되는 스코어에서 마지막 부도의 정보가 파악되므로 ACAP 곡선으로부터 신용평가 모형의 판별력(discriminatory power)을 탐색할 수 있다.

ROC와 CAP 곡선에서는 각 좌표에 대응하는 스코어를 파악할 수 없으나, 본 연구에서 제안한 AROC와 ACAP 곡선에서는 스코어를 수직축에 나타내기 때문에 곡선의 최대값에 대응하는 스코어를 통해 최적의 절단점, 예측력 그리고 판별력을 탐색할 수 있다는 장점이 있다.

그림 2.2의 왼쪽에는 ROC와 CAP 곡선(ROC 곡선이 위에 위치함)을 표현하였고 AROC와 ACAP 곡선이 최대값일 때 스코어에 대응하는 ROC와 CAP 곡선의 좌표에 점(·)으로 나타내었다. 그리고 그림 2.2의 오른쪽에는 부도기업과 정상기업의 누적분포함수(부도기업의 누적분포함수가 위에 위치함)를 작성하고 이 두 분포함수의 차이가 최대일 때 두 함수의 폭을 연결하는 직선으로 나타내었는데 이 직선의 길이가 K-S 통계량 값을 의미한다. 우선 그림 2.2의 ROC와 CAP 곡선을 그림 2.1의 AROC와 ACAP 곡선과 비교하여 살펴보면, AROC와 ACAP 곡선에서 최대값의 스코어인 36에 대응하는 ROC와 CAP 곡선은 모두 수직축의 좌표가 1일 때인 것을 발견할 수 있다. ROC와 CAP 곡선의 수직축은 $TP/p = 1 - FN/p$ 이므로 좌표값이 1이 되기 위해서는 $FN = 0$ (또는 $TP=p$)인 경우이며 AROC와 ACAP 곡선에서 $FN = 0$ 인 경우에 최대값을 갖기 때문이다. ROC와 CAP 곡선에서 수직축 좌표가 처음으로 1에 도착하는 스코어는 판별력을 판단하는데 중요한 역할을 한다. 예를 들어 이 스코어가 전체 중에서 순서가 빠른 스코어라면 ROC와 CAP 곡선이 빠르게 증가하여 곡선 AUROC 또는 AR이 큰 값을 갖는다. 따라서 예측력 또는 판별력을 식별하는데 중요한 정보를 제공한다.

그림 2.2의 오른쪽에 나타난 부도와 정상기업의 누적분포함수와 비교하면, AROC와 ACAP 곡선의 최대값에 대응하는 스코어 36에서 부도기업의 누적분포함수인 tp 비율과 정상기업의 누적분포함수인 fp 비율의 차이가 최대임을 발견할 수 있다. 즉 부도와 정상기업의 누적분포함수의 최대차이에 대응하는 스코어에서 AROC와 ACAP 곡선이 최대값을 갖는다. AROC 곡선의 최대값은 FP 가 0에 수렴하거나(fp 비율 감소) 또는 FN 이 0에 수렴하는 경우(tp 비율 증가)인데 tp 비율이 증가하거나 fp 비율이 감소하면 tp 비율과 fp 비율의 차이가 커지기 때문에 AROC 곡선의 최대값과 K-S 통계량에 대응하는 스코어는 일치한다. ACAP 곡선의 최대값은 부도가 마지막으로 발생하는 경우인 FN 이 0에 수렴하는 경

우(tp 비율 증가)이므로 tp 비율이 1인 경우에 ACAP 곡선의 최대값과 K-S 통계량에 대응하는 스코어는 일치하는 경향을 보인다.

AROC와 ACAP 곡선의 최대값 스코어에서 ROC와 CAP 곡선은 수직축의 값이 처음으로 1이 될 때이고, K-S 통계량이 결정되는 것을 발견할 수 있다. 이런 결론이 유도되는 이유는 AROC와 ACAP 곡선의 최대값이 발생하는 스코어에서 평가모형의 판별력의 차이가 가장 크게 나타나기 때문이라고 판단된다. 그러나 AROC와 ACAP 곡선을 통하여 얻은 결론이 모든 신용평가 자료에서 항상 동일하게 발생하지 않는다. 어떤 상황에서 동일한 결과가 발생하는지를 파악하기 위하여 모의실험을 통하여 관찰하고자 한다.

3. 모의실험

2절에서 제안한 AROC와 ACAP 곡선의 최대값이 발생하는 스코어에서 신용평가모형의 판별력 차이가 가장 크게 나타나기 때문에 최대값에 대응하는 스코어에서 ROC와 CAP 곡선은 수직축의 값이 가장 먼저 1이 되는 상황이고, 부도와 정상기업의 누적분포함수의 차이가 제일 큰 K-S 통계량이 결정되는 것을 발견할 수 있다. 이런 상황이 어떤 신용평가모형의 결과에서 적용되는지를 알아보기 위하여 현업에서 사용하는 신용평가 자료와 유사한 자료를 모의실험을 통해 생성하여 살펴보고자 한다. 현업에서 사용하는 자료와 유사하게 설계하기 위해서 다음의 절차에 따라 정상과 부도기업의 자료를 생성한다.

- 1 표준정규분포 $N(0, 1)$ 로부터 1,000개의 난수를 생성하여 1,000개 기업의 스코어로 간주한다.
- 2 정상과 부도기업의 비율을 결정하는 부도율을 0.05로 설정하고 정상과 부도기업의 스코어를 생성한다. 즉 950:50의 비율로 정상과 부도기업을 구분한다. 신용평가에 사용되는 실제 자료는 금융 서비스를 받은 대상자가 분석대상이 되며, 현실적으로 부도와 정상 대상자의 비율차이가 크게 나타나기 때문에 비율차이를 크게 설정하였다.
- 3 실제 부도기업을 정상기업으로 잘못 예측하여 분류하는 오분류율 α_1 과 실제 정상기업을 부도기업으로 잘못 예측하여 분류하는 오분류율 α_2 로 설정하여 자료를 혼합한다.
- 4 오분류된 자료를 효과적으로 혼합하기 위하여 부도율 0.05에 대한 $\Phi^{-1}(0.05) = -1.645$ 를 정상과 부도기업의 분류기준점으로 간주하고, 0.25의 구간폭으로 양쪽 방향의 두 구간씩을 설정하여 오분류율 α_{11} , α_{12} 와 α_{21} , α_{22} 를 할당한다. 예를 들어 오분류율 α_1 과 α_2 를 각각 5%와 10%로 설정하고 $(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}) = (0.02, 0.03, 0.06, 0.04)$ 라고 표기하면, 구간 $(-2.145, -1.895)$ 에서는 오분류된 자료 $\alpha_1 = 5\%$ 중에서 2%의 자료를, 구간 $(-1.895, -1.645)$ 에서는 오분류된 자료 $\alpha_1 = 5\%$ 중에서 3%의 자료를, 구간 $(-1.645, -1.395)$ 에서는 오분류된 자료 $\alpha_2 = 10\%$ 중에서 6%의 자료를 그리고 구간 $(-1.395, -1.145)$ 에서는 오분류된 자료 $\alpha_2 = 10\%$ 중에서 4%의 자료를 할당하여 오분류된 자료가 기준점 가까이에 보다 많이 발생하는 경우이다.
- 5 부도와 정상기업의 누적분포함수의 차이가 최대일 때의 스코어를 S_{KS} 로, AROC와 ACAP 곡선에서의 최대값 스코어를 각각 S_{AROC} 와 S_{ACAP} 로 표기하고 오분류율과 할당의 변화에 따라 $S_{AROC} = S_{ACAP} = S_{KS}$ 가 일치하는 경우를 살펴본다.
- 6 위 1-4의 과정을 10,000번 반복한다.

오분류율 α_1 과 α_2 를 5%와 10%로 설정하고 각각의 오분류율을 적절한 구간에 다양하게 할당하여 모의실험을 실시한 결과를 표 3.1에 나열하였다. AROC와 ACAP 곡선의 최대값에 대응하는 스코어 S_{AROC} 와 S_{ACAP} 그리고 K-S 통계량에 대응하는 스코어 S_{KS} 가 모두 일치하는 경우를 세 종류의 스코어 중에서 두 가지만 일치하거나 모두 일치하지 않는 경우와 비교하면서 살펴본다.

표 3.1. 오분류율 할당에 따른 모의실험 결과

A: (0.02, 0.03, 0.05, 0.05)			B: (0, 0.05, 0.10, 0)		
	빈도	백분율		빈도	백분율
$S_{AROC} = S_{ACAP} = S_{KS}$	3,311	33.11	$S_{AROC} = S_{ACAP} = S_{KS}$	4,255	42.55
$S_{AROC} = S_{KS} \neq S_{ACAP}$	27	0.27	$S_{AROC} = S_{KS} \neq S_{ACAP}$	6	0.06
$S_{ACAP} = S_{KS} \neq S_{AROC}$	2,430	24.30	$S_{ACAP} = S_{KS} \neq S_{AROC}$	5,335	53.35
$S_{ACAP} = S_{AROC} \neq S_{KS}$	1,982	19.82	$S_{ACAP} = S_{AROC} \neq S_{KS}$	141	1.41
$S_{AROC} \neq S_{ACAP} \neq S_{KS}$	2,250	22.50	$S_{AROC} \neq S_{ACAP} \neq S_{KS}$	263	2.63
계	10,000	100.00	계	10,000	100.00

C: (0.02, 0.03, 0.05, 0)			D: (0.02, 0.03, 0.10, 0)		
	빈도	백분율		빈도	백분율
$S_{AROC} = S_{ACAP} = S_{KS}$	7,952	79.52	$S_{AROC} = S_{ACAP} = S_{KS}$	9,456	94.56
$S_{AROC} = S_{KS} \neq S_{ACAP}$	238	2.38	$S_{AROC} = S_{KS} \neq S_{ACAP}$	7	0.07
$S_{ACAP} = S_{KS} \neq S_{AROC}$	479	4.79	$S_{ACAP} = S_{KS} \neq S_{AROC}$	180	1.80
$S_{ACAP} = S_{AROC} \neq S_{KS}$	1,179	11.79	$S_{ACAP} = S_{AROC} \neq S_{KS}$	353	3.53
$S_{AROC} \neq S_{ACAP} \neq S_{KS}$	152	1.52	$S_{AROC} \neq S_{ACAP} \neq S_{KS}$	4	0.04
계	10,000	100.00	계	10,000	100.00

표 3.1에서 A경우는 생성된 신용평가모형의 절단점을 기준으로 오분류율 $\alpha_1 = 5\%$ 를 α_{11} 과 α_{12} 에 각각 0.02와 0.03으로 할당하고, 오분류율 $\alpha_2 = 10\%$ 를 $\alpha_{21} = 0.05$, $\alpha_{22} = 0.05$ 로 할당하였다(0.02, 0.03, 0.05, 0.05). 부도를 정상으로 예측하는 오분류율보다 정상을 부도로 예측하는 오분류율을 두 배로 설정한 경우이다. 이 경우에 AROC와 ACAP 곡선의 최대값 그리고 K-S 통계량값에 대응하는 스코어가 모두 일치하는 경우($S_{AROC} = S_{ACAP} = S_{KS}$)가 전체의 약 33.11%로 일치하는 현상이 두드러지게 나타나지는 않지만 모든 스코어가 일치하지 않는 종류의 경우보다 상대적으로 비율이 조금 높다.

표 3.1에서 B경우는 오분류율 α_1 과 α_2 를 A경우와 동일하게 5%와 10%로 설정하고, 오분류율 할당을 (0, 0.05, 0.10, 0)으로 하였다. 이 경우는 절단점 가까이에 위치한 좌우의 한 구간에서만 오류가 집약적으로 발생하도록 설정하였다. B경우에서는 $S_{ACAP} = S_{KS} \neq S_{AROC}$ 인 경우가 $S_{AROC} = S_{ACAP} = S_{KS}$ 인 경우보다 조금 더 많이 나타났지만 $S_{AROC} = S_{ACAP} = S_{KS}$ 인 경우가 42.55%로 A경우의 33.11%보다 증가하였다.

표 3.1의 C경우는 오분류율 α_1 과 α_2 를 5%로 동일하게 설정하였는데 이는 A경우에서의 α_2 를 10%에서 5%로 축소하면서 오분류율 절단점 가까이에 밀집되게 할당한 경우이다(0.02, 0.03, 0.05, 0). ‘부도를 정상으로 잘못 분류하는 오분류율은 넓게 퍼져있고 정상을 부도로 잘못 분류하는 오분류율은 절단점 가까이에 많이 발생하는 경우이며, 이 경우에는 AROC와 ACAP 곡선의 최대값 그리고 K-S 통계량값에 대응하는 스코어가 모두 일치하는 경우($S_{AROC} = S_{ACAP} = S_{KS}$)의 비율이 전체의 약 79.52%로 크게 증가하였다. A경우와 비교하면, 부도를 정상으로 예측하여 오분류된 자료가 발생하는 위치에는 민감하지 않은 반면에 정상을 부도로 예측하여 오분류된 자료의 위치가 절단점 가까이에 있는지는 매우 민감하게 반응하며 이런 경우에 $S_{AROC} = S_{ACAP} = S_{KS}$ 인 경우가 많이 발생함을 판단할 수 있다.

표 3.1의 D경우 오분류율 α_1 을 C경우와 동일하게 5%로 설정하고 α_2 를 5%에서 10%로 증가시켜 오분류율 할당을 (0.02, 0.03, 0.10, 0)으로 설정한다. B경우의 오분류율 할당보다 부도를 정상으로 잘못 분류하는 오분류율을 넓게 퍼지도록 설정하는 경우이고, A경우의 오분류율 할당에서 정상을 부도로 잘못 분류하는 오분류율을 절단점 가까이로 밀집시킨 경우이다. AROC와 ACAP 곡선의 최대값 그리고 K-S 통계량값에 대응하는 스코어가 모두 일치하는 경우($S_{AROC} = S_{ACAP} = S_{KS}$)가 전체의 약 94.56%로 대부분을 차지한다. 그러므로 부도를 정상으로 잘못 분류하는 오분류된 자료 위치는 넓게 퍼져있어도

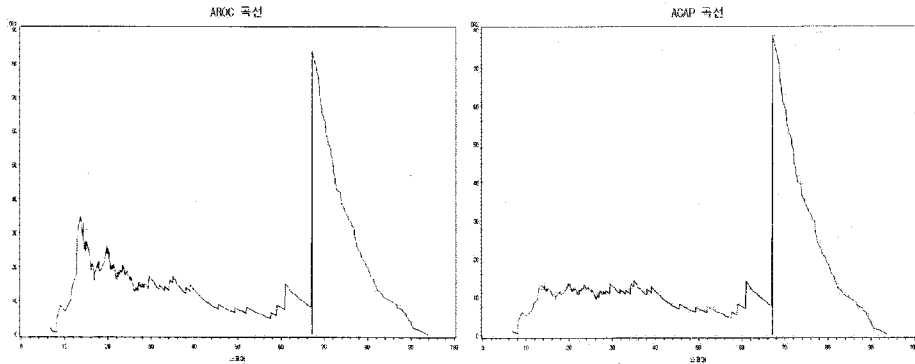


그림 4.1. AROC와 ACAP 곡선

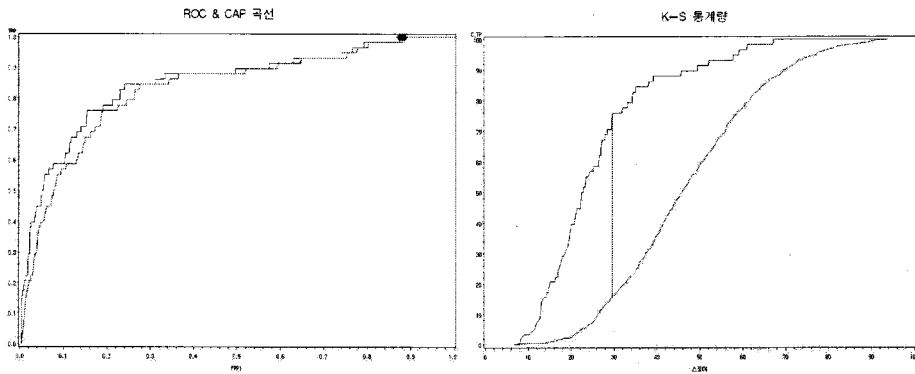


그림 4.2. ROC와 CAP 곡선과 K-S 통계량

민감하게 반응하지 않지만, 정상을 부도로 잘못 분류하는 오분류된 자료는 절단점 가까이에서 많이 발생할수록 모든 스코어가 동일한 $S_{AROC} = S_{ACAP} = S_{KS}$ 인 경우가 자주 발생하는 특성을 파악할 수 있다.

결과적으로 부도를 정상으로 잘못 분류하는 오분류율 α_1 은 α_{11} 과 α_{12} 도 균등하게 할당되고, 정상을 부도로 잘못 분류하는 오분류율 α_2 의 크기가 클수록 그리고 오분류된 자료가 절단점 가까이에 밀집되어 있는 경우에 모든 스코어가 동일한 $S_{AROC} = S_{ACAP} = S_{KS}$ 인 경우가 자주 발생한다.

4. 실증 자료

2절과 3절에서 살펴본 자료의 유형 이외에 AROC와 ACAP 곡선의 최대값에 대응하는 각각의 스코어 S_{AROC} 와 S_{ACAP} 그리고 K-S 통계량값에 대응하는 S_{KS} 가 동일한 스코어를 갖지 않는 경우가 자주 발생한다. 다시 말하면 오분류된 자료가 절단점을 중심으로 가까이 집중되어 발생하지 않고 산포되어 있는 자료의 경우에는 스코어 S_{AROC} , S_{ACAP} 그리고 S_{KS} 가 다른 값을 갖고 따라서 그들과의 관계를 발견하는데 한계가 있다.

예를 들어 1994년부터 2005년까지 외감 기업 중 매출액 1,000억 이상의 대기업에 관한 연도별로 4,268건(정상: 4101건, 부도: 167건)의 재무자료를 바탕으로 개발한 신용평가모형의 예측 부도율 자

료를 살펴보자. 최종적으로 채택된 신용평가모형은 좋은 판별력을 갖고 있으며 그 결과는 홍종선 등 (2008)에 수록되어있다. 그 중에서 K-S 통계량은 0.6676이고 AUROC는 0.8909의 높은 값을 가지는 좋은 모형이라고 판단할 수 있다. 이 모형에 대한 자료 중에서 자산규모가 4,500억 이상의 우수한 업체만을 선정하여 1,009건(정상: 951건, 부도: 58건)의 재무자료에 대하여 AROC와 ACAP 곡선을 그림 4.1에 구현하고, ROC와 CAP 곡선 그리고 K-S 통계량을 식별할 수 있는 누적분포함수를 그림 4.2에 작성하였다. 그림 4.1에서 AROC와 ACAP 곡선의 최대값에 대응하는 스코어는 일치한다 $S_{AROC}=67.18$ (889번째 스코어), $S_{ACAP}=67.18$ (889번째 스코어). 그리고 그림 4.2를 통하여 K-S 통계량에 대응하는 스코어는 $S_{KS}=29.52$ (191번째 스코어)로 S_{AROC} 와 S_{ACAP} 는 다른 스코어이며, 작은 스코어에서 K-S 통계량이 결정된다는 것을 확인할 수 있다. 따라서 이 신용평가모형의 자료를 표현한 AROC와 ACAP 곡선으로부터 최적의 절단점을 S_{AROC} 로 추정할 수 있으며, $S_{AROC}=S_{ACAP}$ 에서 판별력이 우수하다고 판단할 수 있다.

5. 결론

ROC 곡선은 각 절단점에서의 fp 비율을 X 축 좌표에 그리고 tp 비율을 Y 축 좌표로 나타내어 작성하고, CAP 곡선은 각 절단점에서의 부도율을 X 축 좌표에 그리고 Y 축 좌표는 ROC 곡선과 동일하게 tp 비율로 작성한다. 따라서 각 절단점에서 구한 비율들로 ROC와 CAP 곡선이 이루어졌지만 반면에 곡선을 통해서 절단점을 파악할 수 없다. 그러므로 ROC와 CAP 곡선을 통하여 최적의 절단점을 탐색하는 것은 쉽지 않다.

본 연구에서는 fp 비율과 tp 비율 그리고 부도율의 오즈비들을 식 (2.1)과 (2.2)와 같이 설정하고 이 오즈비를 이용하여 각 절단점과의 관계를 그림으로 작성한 AROC와 ACAP 곡선을 제안하였다. 즉 X 축 좌표에는 절단점에 대응하는 스코어를, Y 축 좌표에는 fp 비율과 tp 비율의 함수인 오즈비(OR_{AROC})를 나타내는 AROC 곡선을 제안하였고, ACAP 곡선에서는 X 축은 AROC 곡선과 동일하게 설정하고 Y 축 좌표에는 tp 비율과 부도율의 함수인 오즈비(OR_{ACAP})를 나타내었다.

절단점에 가까이에 오분류가 많이 발생하는 자료에 대하여는 AROC와 ACAP 곡선의 최대값에 대응하는 각각의 스코어들이 일치하며, 그 스코어에서 ROC와 CAP 곡선의 Y 축 좌표가 처음으로 1이 된다는 것을 발견하였다. 또한 이 스코어에서 부도와 정상기업의 누적분포함수의 차이가 최대가 되고 그 차이가 K-S 통계량값이 된다는 것을 발견하였다. 이런 현상은 대부분의 오분류된 자료가 절단점 가까이에 위치하는 경우로 판별력이 좋은 신용평가 자료인 경우에 발생한다. 그러나 신용평가모형의 판별력이 좋게 검증(validation)된 자료라 할지라도 AROC와 ACAP 곡선에서 최대값에 대응하는 최적의 절단점(스코어)이 ROC와 CAP 곡선의 Y 축 좌표가 처음으로 1이 되는 스코어와 K-S 통계량에 대응하는 스코어 모두가 항상 일치하지 않는다. 이 스코어들의 불일치는 전체 스코어 분포에 의존하기 때문에 낮은 스코어와 높은 스코어에서 빈도수가 작게 나타나는 경우에 OR_{AROC} 와 OR_{ACAP} 의 값이 매우 불안정하게 나타날 수 있는 단점이 발생한다. 그러나 본 연구를 통해 AROC 곡선에서의 OR_{AROC} 의 값이 클수록 예측력이 좋다고 판단할 수 있기 때문에 최대값에 대응하는 스코어를 최적의 절단점으로 간주할 수 있으며, ACAP 곡선에서 최대값을 갖는 스코어에서는 판별력이 좋다고 판단할 수 있다.

본 논문에서는 AROC와 ACAP 곡선의 최대값 그리고 K-S 통계량값에 대응하는 스코어가 모두 일치하는 경우($S_{AROC} = S_{ACAP} = S_{KS}$)만을 중점적으로 연구하였으나 표 3.1으로부터 세 종류의 스코어가 서로 일치하지 않는 다양한 경우가 존재함을 알 수 있다. 따라서 절단점을 중심으로 어떤 오분류의 상황 하에서 스코어들이 일치하고 일치하지 않는지를 탐색하여 $S_{AROC} = S_{ACAP} = S_{KS}$ 이외의 다양한 경우에 나타나는 자료의 형태를 파악하는 것은 필요하다. 또한 Joseph (2005)은 AUROC와 AR를

포함한 신용평가모형의 타당성 검증(validation) 통계량들의 평가기준을 모분산이 일정한 두 정규분포로부터 모평균 차이의 변화에 따라 제안하였다. AROC와 ACAP 곡선의 오즈비 OR_{AROC} 와 오즈비 OR_{ACAP} 의 값에 대한 일반적인 평가기준을 Joseph (2005)의 연구를 바탕으로 제시하고, 나아가 보다 현실적인 신용평가모형에 적합한 가정 하에서 평가기준을 제시하는 향후 연구를 제안한다.

참고문헌

- 승문섭, 박창순, 이정진 (2003). <S-LINK를 이용한 비모수통계학>, 자유아카데미, 경기.
- 이근희 (2006). 바젤II 협약 기반 신용평점체계의 계량적 점검에 대한 고찰, 예금보험공사, <금융리스크리뷰>, **3**, 70-95.
- 임종건 (2005). 신용평가시스템에 대한 적합성 검증, 금융감독원, <리스크리뷰>, **2005**.
- 홍중선, 이창혁, 김지훈 (2008). Validation comparison of credit rating models for categorized financial data, <한국통계학회논문집>, **15**, 615-631.
- Fawcett, T. (2004). *ROC Graphs: Notes and Practical Considerations for Researchers*, Kluwer Academic Publishers, Netherlands.
- Fawcett, T. (2005). An introduction to ROC analysis, *Pattern Recognition Letters*, **27**, 861-874.
- Hosmer, D. W. and Lemeshow, S. (2000). *Applied Logistic Regression*, 2nd ed, John Wiley & Sons, New York.
- Joseph, M. P. (2005). A PD validation framework for basel II internal ratings-based systems, *Credit Scoring and Credit Control IV*.
- Sonego, P., Kocsor, A. and Pongor, S. (2008). ROC analysis: Applications to the classification of biological sequences and 3D structures, *Briefings in Bioinformatics*, **9**, 198-209.
- Tasche, D. (2006). Validation of internal rating systems and PD estimates, *arXiv:physics/0606071*, **1**.

Adjusted ROC and CAP Curves

Chong Sun Hong¹ · Ji Hun Kim² · Jin Soo Choi³

¹Dept. of Statistics, Sungkyunkwan University;

²Research Institute of Applied Statistics, Sungkyunkwan University;

³Research Institute of Applied Statistics, Sungkyunkwan University

(Received August 2008; accepted November 2008)

Abstract

Among others, ROC and CAP curves are used to explore the discriminatory power between the defaults and non-defaults, based on the distribution of the probability of default in credit rating works. ROC and CAP curves are plotted in terms of various ratios of the probability of default. Each point on ROC and CAP curves is calculated according to cutting points (scores) for classifying between defaults and non-defaults. In this paper, adjusted ROC and CAP curves are proposed by using functions of ratios of the probability of default. It is possible to recognize the score corresponding to a point on these adjusted curves, and we can identify the best score to show the optimal discriminatory power. Moreover, we discuss the relationships between the best score obtained from the adjusted ROC and CAP curves and the score corresponding to Kolmogorov - Smirnov statistic to test the homogeneous distribution functions of the defaults and non-defaults.

Keywords: Credit rating, cross-product ratio, cutting point, discriminatory power, odds ratio, probability of default, score.

¹Corresponding author: Professor, Dept. of Statistics, Sungkyunkwan University, 3-53, Myungryun-dong 3, Jongro-gu, Seoul 110-745, Korea. E-mail: cshong@skku.ac.kr