

Validation Comparison of Credit Rating Models Using Box-Cox Transformation

Chong Sun Hong¹⁾ · Jeong Min Choi²⁾

Abstract

Current credit evaluation models based on financial data make use of smoothing estimated default ratios which are transformed from each financial variable. In this work, some problems of the credit evaluation models developed by financial experts are discussed and we propose improved credit evaluation models based on the stepwise variable selection method and Box-Cox transformed data whose distribution is much skewed to the right. After comparing goodness-of-fit tests of these models, the validation of the credit evaluation models using statistical methods such as the stepwise variable selection method and Box-Cox transformation function is explained.

Keywords: Box-Cox transformation; Credit rating model; Performance; Power transformation; Smoothing; Stepwise variable selection; Validation.

1. 서론

금융회사의 주요 업무이면서 핵심은 리스크(risk)를 적절히 수용하고 관리하면서 적절한 수익을 확보해야한다. 금융회사는 신용리스크(credit risk)를 관리함에 있어 엄격한 신용평가(credit assessment)를 통해 신용리스크를 관리하고 있다. 기업 등의 재무 및 비재무 상황을 종합적으로 평가하여 그 결과에 따라 일정 수준의 신용등급(credit rating)을 부여하는 모형을 신용평가모형(credit rating model, credit evaluation model)이라 한다. 일반적으로 금융기관에서 기업체를 대상으로 사용하는 신용평가모형은 재무정보에 기반을 둔 신용평가모형과 전문가 판단모형의 결합한 형태이다(검사지원국 2004).

재무정보를 이용한 신용평가모형은 거래상대방이 일정기간 내에 도산 또는 부실화

1) Correspondence Author, Professor, Department of Statistics, Sungkyunkwan University, 110-745, KOREA. E-mail : cshong@skku.ac.kr

2) Researcher, Research Institute of Applied Statistics, Sungkyunkwan University, Seoul 110-745 Korea.

될 가능성을 수치화된 자료를 이용하여 점수화함으로써 평가자의 주관적인 판단을 배제한 객관적 결론을 유도하여 산출하는 것이 목표이다. 따라서 과거 일정기간 동안 실제 도산한 기업들의 재무자료를 토대로 통계적인 분석을 수행하여 유의한 것으로 밝혀진 설명변수들을 추출하고 이 변수들을 이용한 부도확률함수를 제시함으로써 기업의 잠재적 부도가능성을 계량화하는 방법이 재무정보에 근거한 신용평가모형이다.

Hand와 Henley(1997)는 신용등급을 산출하기 위한 통계적 방법과 신용평가에 관하여 연구하였다. 전형적인 신용평가모형은 대부분 연속형 재무자료(financial data)를 수집하여 재무비율들을 생성한 후, 각각의 변수를 연도별 변동을 줄이면서 평활(smoothing)하게 변환한다. 그리고 변수를 선택하여 최적의 모형을 개발하는데, 특정 분야의 전문가들만이 다룰 수 있는 전문적 지식습득을 기초로 하는 분야이다. 본 논문에서도 비재무자료를 제외한 재무자료만을 사용한 신용평가모형에 대하여 연구한다. 재무변수를 평활하게 변환시키는 Loess 방법을 많이 사용하는데 본 연구에서는 Box-Cox가 제안한 파워변환(power transformation)을 확장한 Yeo와 Johnson(2000)의 새로운 파워변환(본 논문에서는 Box-Cox 변환이라고 한다)을 이용하는 방법을 제안하고 그 결과를 비교 설명하고자 한다. 이미 개발된 신용평가모형의 분류성과(performance)를 ‘평균차이’(D), ‘불량 누적비율 50%에 대한 우량의 누적비율’(1-PH), ‘일치성’(Concordant), ‘콜모고로브-스미노브’(K-S), ‘accuracy ratio’(AR), ‘area under receiver operating characteristic’(AUROC) 등의 통계량을 사용하여 구하고, 본 연구에서 제안한 모형의 결과를 살펴보면서 비교 평가한다.(박병수 2004, 임종건 2005, Engelmann와 그 외 2003, Joseph 2005, Stein 2005, Thomas와 그 외 2004 등 참조)

본 논문의 구성은 다음과 같다. 2절에서는 국내 모 은행에서 개발한 신용평가모형 설정과정을 설명한다. 3절에서는 현재 사용하고 있는 신용평가모형의 문제점을 지적한 후에, 각각의 재무변수를 Box-Cox가 개발한 변수변환방법을 대안적으로 제시한다. 그리고 변환된 재무자료를 신용평가모형에 적용해 기존의 모형과 비교한다. 일반적인 신용평가모형 개발과정에서 통계적 변수선택 방법 이외에 금융전문가들이 모형의 설명력, 직관성, 안정성 등을 고려하여 변수들을 선택한다. 4절에서는 기존에 사용한 변수 변환하는 방법과 Box-Cox 변환방법을 적용한 재무자료 각각에 대하여 통계적인 방법 중의 하나인 단계적 변수선택 방법(stepwise variable selection method)을 이용하여 신용평가모형을 제안하고 각 모형의 분류성과를 비교 검토한다. 마지막으로 5절에서는 본 연구에서 제안한 신용평가모형에 대하여 결론을 내린다.

2. 신용평가모형 개발과정

방대한 재무자료를 이용하여 신용평가모형을 개발하는 일반적인 과정은 자료수집, 변수생성, 변수변환, 변수선택 그리고 최종모형선정의 과정으로 이루어진다. 여기서는 예를 들어 국내 모 은행에서 개발한 신용평가모형의 과정에 대하여 설명한다.

첫번째 자료수집 및 변수생성 단계에서는 재무모형 개발에 적용 가능한 대기업 자료를 이용하는데, 1994년부터 2005년까지 외감 기업 중 매출액 1,000억 이상의 대기업에 관한 연도별로 4,268건(정상: 4101건, 부도: 167건)의 재무자료를 바탕으로 각 재무비율을 변수로 생성하고, 두번째 변수변환 단계에서는 ‘By year method’로 규모(Size), 안정성(Leverage), 안정성(Retained Earning), 부채상환능력(Debt Coverage), 수익성

(Profitability), 유동성(Liquidity), 성장성(Growth), 활동성(Activity)의 8개 구분범주로 나뉜 182개의 재무변수를 각각 연도별로 재무비율의 구성이 2%씩 되도록 50개의 구간(bucket)으로 나눈다. 그리고 각 연도에 해당하는 2%, 4%,..., 50%의 구간별 부도빈도수를 계산한 뒤 50개 동일 구간 내 연도별 자료를 묶어 각 구간을 하나의 부도율을 계산한다. 그리고 비모수적 평활(non-parametric smoothing) 방법으로 재무변수의 각 백분위수(percentile)에 해당하는 구간별 부도율은 변동이 심하므로 평활한 추세를 나타내는 Loess 방법(해당위치의 값을 주위 값으로 추정된 선형식 값으로 대체하여 주변의 값은 비슷한 값을 가지도록 하여 위치에 따른 변동을 줄이는 방법으로 자세한 설명은 Cleveland(1979)와 Cohen(1999)의 논문을 참조)을 이용하여 구간별 예측부도율로 변환한다. 그리고 세번째 단계에서는 예측부도율로 변환한 자료에 대한 신용평가 모형의 Concordant, AR 통계량 등을 통해 변수의 유의성을 살펴보면서 1차 변수선택을 하고, 변수들의 상관관계를 통해 2차 변수선택을 실시한다. 그리고 마지막으로 네번째 최종모형 선정 단계에서는 방법1과 방법2에서 선택된 재무변수들을 이용해 각 조합을 생성하고, 선택한 변수 수가 6~7개 사이일 때의 조합을 대상으로 각 조합별 모형들의 구성 위험 요인(risk factor)들이 모형 안에서 통계적으로 유의적인지 확인한다. 그리고 금융전문가들이 모형의 설명력, 선택된 변수의 이해가능성, 직관성, 모형의 안정성 등을 포괄적으로 고려하여 일부 새로운 변수들을 추가하거나 대체한 후 6개의 재무변수와 3개의 가변수를 이용하여 로지스틱 회귀분석을 실시하고, 그 모형의 분류성과를 비교하기 위하여 훈련자료(train dataset)와 타당성 검사자료(validation dataset)를 일반적인 분할 비율인 6:4로 분리해 여러 종류의 적합성 검정 방법(D, 1-PH, K-S, AR, AUROC, Concordant 통계량)을 통해 비교함으로써 최종적인 신용평가모형을 결정한다. 선정된 최종 신용평가모형(신평모형 I)의 결과와 분류성과는 [표 2.1]에 수록하였다.

[표 2.1] 최종 선택된 변수의 신평모형 I 결과와 분류성과(Loess)

범주항목	변수	Estimate	Wald Chi-Square	p-값
	상수항	-7.2094	420.5131	<.0001
	건설업	-0.7051	6.1315	0.0133
	서비스업	-1.1910	5.7442	0.0165
	IT	0.5629	4.1977	0.0405
안정성 (Leverage)	(부채총계-이익잉여금)/(유형자산)	0.1967	22.5573	<.0001
	차입금/자기자본	0.1155	20.1821	<.0001
부채 상환능력	총C/F대차입금비율	0.0768	8.5625	0.0034
	경상이익이자보상비율	0.0906	6.4049	0.0114
유동성	단기유동성비율	0.0581	4.3810	0.0363
활동성	유동부채/매출액	0.0960	13.5367	0.0002

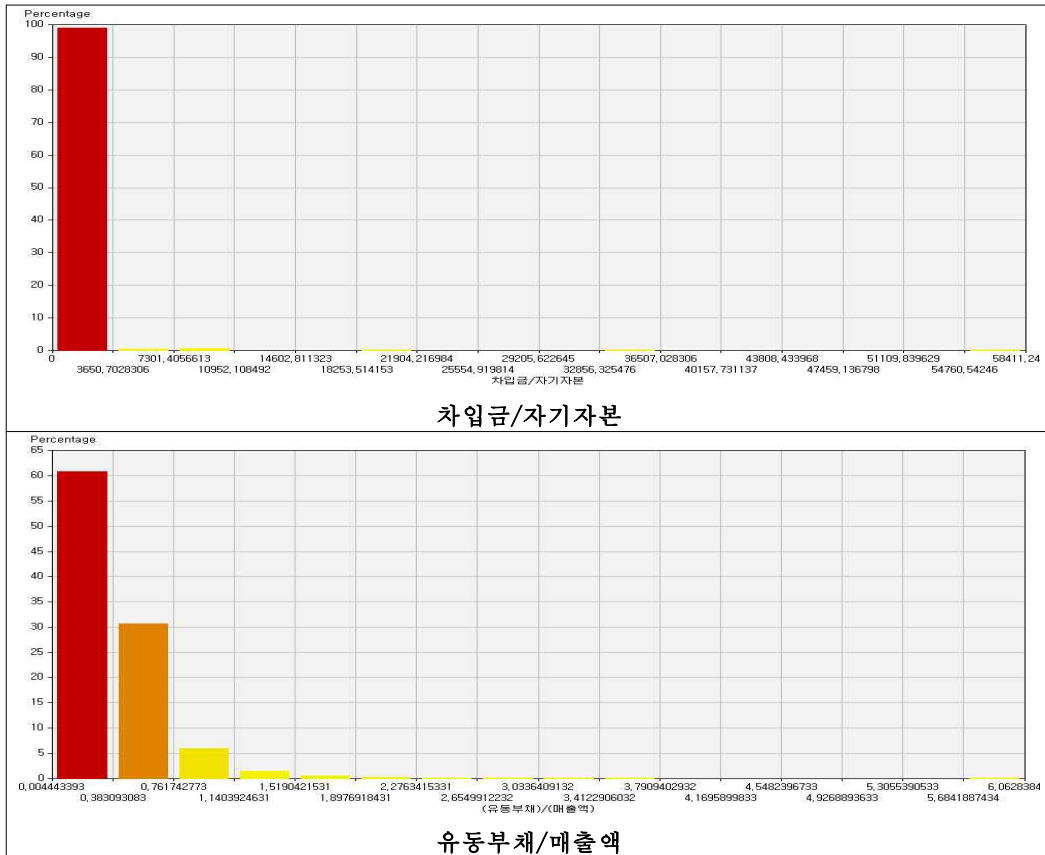
검정통계량	분류성과
<i>D</i>	1.4035
<i>1-PH</i>	0.9722
<i>K-S</i>	0.6676
<i>AR</i>	0.7819
<i>AUROC</i>	0.8909
<i>Concordant</i>	0.8830

최종 로지스틱 회귀모형의 회귀계수는 유의수준 5%에서 모두 유의하다는 것을 [표 2.1] 왼쪽의 모형 결과에서 파악할 수 있다. 최종 신용평가모형 분류성과를 살펴보면 1-PH, AR, AUROC, 그리고 Concordant 통계량들은 매우 큰 값을 갖고 있고 자유도에 의존하는 D와 K-S 통계량들의 값도 유의한 큰 값을 나타내고 있으므로 모형의 예측률이 높다고 판단할 수 있다.

3. Box-Cox 변환을 이용한 신용평가모형

기존의 신용평가모형은 재무평가항목들의 자료를 수집해 각 변수마다 Loess 방법을 이용하여 예측부도율로 변환하여 모형을 개발하였다. 2절에서 설명한 'By year method'에 의한 Loess 방법을 사용함으로써 연도별 재무변수의 흐름이 변동함에도 불구하고 그러한 특수상황이 무시됨에 따라 정보의 손실이 예상된다. 이것은 원자료의 고유한 특성을 고려하지 않고 182개 변수 각각에 대하여 50개 구간 부도율을 평활모수에 의해 추정된 예측 값 즉, 예측부도율로 변환함으로써 자료가 갖고 있는 고유한 정보의 손실을 초래하는 문제점이 있다. 또한 최근 수집되어 갱신하는 개별(기업) 관측 값에 대해서도 위와 같은 변환 작업을 다시 수행해야하는 불편함으로 모형의 안정성 측면에서 부족하다.

【그림 3.1】 최종모형에 선택된 6개 변수중 2개 변수의 분포



또한 최종모형에서 선택된 6개 독립변수를 비롯한 대부분의 독립변수인 재무변수들은 큰 양수 값의 왜도를 나타낸다. [표 2.1]에 선택된 6개의 변수 중 2개만을 [그림 3.1]에 표현하였다. 분포가 대칭적이 되도록 왜도를 줄이면서 그리고 원자료의 정보손

실을 방지하기 위해 Box와 Cox(1964)가 제안한 파워변환 (power transformation) 방법을 적용해 연구한다. Box-Cox 변환은 변수값이 양수에 대하여만 적용되었으나, Yeo와 Johnson(2000)에 소개된 새로운 파워변환(new power transformation)은 음수를 포함한 실수에 대하여 다른 변환모수를 고려하고 특히, 큰 음수를 갖는 자료나 심한 왜도를 갖는 자료에 대해 일반적이다. 새로운 파워변환의 함수 $\psi(\cdot, \cdot)$ 는 다음과 같다.

$$\psi(\lambda, x) = \begin{cases} \{(x+1)^\lambda - 1\}/\lambda & (x \geq 0, \lambda \neq 0), \\ \log(x+1) & (x \geq 0, \lambda = 0), \\ -\{(-x+1)^{2-\lambda} - 1\}/(2-\lambda) & (x < 0, \lambda \neq 2), \\ -\log(-x+1) & (x < 0, \lambda = 2). \end{cases}$$

X_1, X_2, \dots, X_n 은 확률표본(random sample)이고, $\psi(\lambda, X_1), \dots, \psi(\lambda, X_n)$ 에 의해 변환된 변수들을 정의한다. 변환된 관측값들은 평균이 u 분산이 σ^2 인 정규분포로 취급될 수 있음을 추정한다. 이런 추정 하에서 로그가능도함수는 다음과 같다:

$$l_n(\theta|x) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n [\psi(\lambda, x_i) - u]^2 \quad (3.1)$$

$$+ (\lambda - 1)\sum_{i=1}^n \text{sgn}(x_i)\log(|x_i| + 1),$$

여기서 $\theta = (\lambda, u, \sigma^2)'$, $x = (x_1, x_2, \dots, x_n)'$. λ 를 고정시킨 후, $l_n(\lambda, \cdot, \cdot |x)$ 를 최대화하면서 $\hat{u}(\lambda) = \sum \psi(\lambda, x_i)/n$ 와 $\hat{\sigma}^2(\lambda) = \sum \{\psi(\lambda, x_i) - \hat{u}(\lambda)\}^2/n$ 를 얻는다. λ 의 최대가능도추정량(MLE) $\hat{\lambda}$ 는 $l_n(\lambda, \cdot, \cdot |x)$ 를 최대화하면서 구한다. 추정량 $\hat{\theta} = (\hat{\lambda}, \hat{u}(\hat{\lambda}), \hat{\sigma}^2(\hat{\lambda}))$ 는 (3.1)식의 로그가능도함수를 최대화시킨다.

[표 3.1] 최종 선택된 모형 결과와 분류성과(Box-Cox)

범주항목	변수	Estimate	Wald Chi-Square	p-값
	상수항	-8.2645	336.4945	<.0001
	건설업	0.0182	0.0060	0.9381
	서비스업	-1.5017	7.8512	0.0051
	IT	0.6676	6.3396	0.0118
안정성 (Leverage)	(부채총계-이익잉여금)/(유형자산)	-0.0013	0.9401	0.3322
	차입금/자기자본	0.4360	107.8408	<.0001
부채상환능력	총C/F대차입금비율	0.0000	0.1050	0.7459
	경상이익이자보상비율	0.0000	0.0000	0.9993
유동성	단기유동성비율	-0.0053	2.3002	0.1294
활동성	유동부채/매출액	8.2355	38.7006	<.0001

검정통계량	분류성과
<i>D</i>	1.3304
<i>1-PH</i>	0.9604
<i>K-S</i>	0.6327
<i>AR</i>	0.7438
<i>AUROC</i>	0.8719
<i>Concordant</i>	0.8670

신용평가모형의 개발 과정에서는 각 재무자료를 예측부도율로 변환하여 설명변수로 사용하는데, 본 연구에서는 왜도가 심한 원자료를 Yeo와 Johnson(2000)의 새로운 파

위변환(본 논문에서는 Box-Cox 변환)라는 직접적인 변환방법을 사용하여 신용평가모형을 개발하고자 한다. Loess 변환방법은 복잡하면서도 연도별 특수상황이 무시되고, 재무변수를 예측부도율로 대체하면서 발생하는 정보손실의 문제점까지도 보완할 수 있다. 2절에서 살펴본 기존의 모형과 비교하기 위해 최종모형에 포함된 6개의 재무변수를 Loess 변환이 아닌 Box-Cox 변환하고 3개의 가변수와 같이 로지스틱 회귀분석을 실시한 결과를 [표 3.1]에 나타내었다. [표 3.1]을 [표 2.1]과 비교하면서 살펴보면, 모형에 포함된 변수의 계수들 중 일부가 유의하지 않게 나타났으며 분류성과도 낮은 값을 보여준다. 이것은 Box-Cox 변환방법과 Loess 변환방법을 모형과 분류성과를 통해 비교하기 전에 우선 변수를 선택하는 과정에서부터 다시 고려해야 하겠다. 따라서 4절에서는 통계적 변수선택 방법을 고려하여 선택된 2절과 3절에서의 모형에 포함된 변수와 다른 변수에 대해 Box-Cox와 Loess 변환방법을 적용하여보자.

4. 통계적 변수선택 방법을 이용한 신용평가모형

2절에서 설명한 기존의 신용평가모형 개발 과정에서 1차 변수선택은 일변량 검정을 그리고 2차 변수선택은 상관분석을 통해서 한 다음, 금융전문가들이 모형의 설명력, 선택된 변수의 이해가능성, 직관성, 모형의 안정성 등을 고려하여 변수를 선택하여 최종 신용평가모형을 구축하였다. 여기에서는 금융전문가의 의견을 수렴하지 않고 오직 통계적인 방법으로 선택한 변수를 Box-Cox가 제안한 방법을 이용해 변환해서 구한 모형과 Loess 방법을 이용한 기존 모형을 비교하고자 한다. 2절에서 설명한 1차 변수선택에서 선정된 13개의 변수들에 대하여 Loess 변환과 Box-Cox 변환방법을 적용한 다음에, 각각의 변환된 자료에 대하여 통계적인 변수선택 방법 중에서 가장 많이 사용하는 단계적 변수선택 방법을 이용하여 구한 두 모형을 비교한다.

4.1 Loess 변환한 자료에 대한 신용평가모형

우선 1차 변수선택 방법에서 선정된 13개의 변수를 Loess 변환하고 단계적 변수선택법($\alpha=0.05$, $\beta=0.05$)으로 구한 최종모형에는 부채상환능력에 관련된 재무변수가 2개((자본총계/cpi)/(금융비용/cpiavg), 총C/F대차입금비율), 유동성에 관련된 변수가 1개(단기유동성비율), 안전성에 관련된 변수가 1개((부채총계/cpi-이익잉여금/cpiavg)/(유형자산/cpi)), 수익성에 관련된 변수가 1개(매출 총이익/이자비용)로 총 5개의 재무변수가 선정되었다. 금융전문가들이 설정한 신용평가모형에 포함된 6개 변수와 단계적 변수선택법에 의하여 선정한 5개 변수들을 비교해보면, 대부분의 변수들은 중복되거나 연관성이 높은 변수들로 구성되어 있음을 알 수 있다. 3개의 가변수를 포함하면서 단계적 변수선택법에 의한 변수들의 신용평가모형(신평모형Ⅱ)의 결과와 분류성과는 [표 4.1]과 같이 나타난다.

[표 4.1]의 분류성과 결과를 [표 2.1]과 비교해보면, 단계적 선택 모형(신평모형Ⅱ)의 대부분의 검정 통계량값은 통계적으로 유의한 차이가 있다고 판단할 수는 없으나 신평모형Ⅰ의 값보다 약간 크다. 두 신용평가모형 모두 검정통계량 값이 크기 때문에 자료를 잘 설명하는 모형이라고 결론내릴 수 있으며, 통계적 방법만을 이용하여 구한 모형이 금융전문가가 선정한 모형에 비교해서 전혀 뒤지지 않는 좋은 성능을 갖는다고 판단할 수 있다.

4.2 Box-Cox 변환한 자료에 대한 신용평가모형

다음은 Box-Cox 방법을 이용해 변환시킨 13개의 재무변수를 단계적 변수선택 방법을 이용하여 모형을 설정하고 그 신용평가모형(신평모형Ⅲ)의 결과와 분류성과를 [표 4.2]에 정리하였다. 여기에서 선택된 변수들의 개수가 4.1절에서의 변수 개수와 동일하도록 유의수준을 $\alpha=0.2$ 로 조정하였다. Box-Cox 변환을 적용한 최종모형에는 안정성에 관련된 '고정자산/차입금'변수와 부채상환능력에 관련된 '((자본총계/cpi)/(금융비용/cpiavg)'변수, 수익성에 관련된 '총자본경상이익율'변수, 유동성에 관련된 '유동부채구성비율'변수, 그리고 활동성에 관련된 '(유동부채/cpi)/(매출액/cpiavg)'변수로 총 5개의 재무변수가 선정되었다. 8개 범주구분 중 5개의 범주구분에서 선택되었다는 것이 특징이다.

[표 4.1] 신평모형Ⅱ 결과와 분류성과(Loess)

범주구분	변수	Estimate	Wald Chi-Square	p-값
	상수항	-6.8184	462.0566	<.0001
	건설업	-0.5911	4.8038	0.0284
	서비스업	-1.0296	4.3865	0.0362
	IT	0.5732	4.2489	0.0393
부채상환능력	(자본총계/cpi) / (금융비용/cpiavg)	0.1378	31.9778	<.0001
	총C/F대차입금비율	0.0801	9.6142	0.0019
유동성	단기유동성비율	0.0901	14.3306	0.0002
안정성 (Leverage)	(부채총계/cpi-이익잉여금/cpiavg)/(유형자산/cpi)	0.1660	15.7424	<.0001
수익성	매출총이익/이자비용	0.0812	10.2922	0.0013

검정 통계량	분류성과
<i>D</i>	1.4081
<i>1-PH</i>	0.9727
<i>K-S</i>	0.6739
<i>AR</i>	0.7843
<i>AUROC</i>	0.8921
<i>Concordant</i>	0.8860

[표 4.2] 신평모형Ⅲ 결과와 분류성과(Box-Cox)

범주구분	변수	Estimate	Wald Chi-Square	p-값
	상수항	-2.7935	3.7858	0.0517
	건설업	-0.0223	0.0076	0.9303
	서비스업	-0.9502	3.5626	0.0591
	IT	0.7485	7.3190	0.0068
안정성 (Leverage)	고정자산/차입금	-0.5730	1.6944	0.1930
부채상환능력	(자본총계/cpi)/(금융비용/cpiavg)	-0.8539	41.0015	<.0001
수익성	총자본경상이익율	-0.0481	14.4302	0.0001
유동성	유동부채구성비율	0.0679	2.3607	0.1244
활동성	(유동부채/cpi)/(매출액/cpiavg)	7.4488	29.4653	<.0001

검정 통계량	분류성과
<i>D</i>	1.3899
<i>1-PH</i>	0.9658
<i>K-S</i>	0.6456
<i>AR</i>	0.7748
<i>AUROC</i>	0.8874
<i>Concordant</i>	0.8790

[표 4.2]를 [표 4.1]과 비교해보면, 단계별 선택모형(신평모형Ⅱ)의 모든 검정 통계량 값은 Loess 변환을 이용한 단계별 선택 모형의 통계량 값보다 약간 적은 값을 갖지만 그 차이는 매우 작기 때문에 신평모형Ⅱ와 신평모형Ⅲ의 성과는 유사하다고 할 수 있다. 통계적 방법을 이용한 신평모형Ⅱ는 금융전문가의 모형(신평모형Ⅰ)만큼 좋은 성

과를 갖고 있으며, Box-Cox 변환한 재무자료를 바탕으로 단계적 선택법을 이용하여 구한 신평모형Ⅲ의 성과가 유사하다면 변환 과정에서 복잡하고 번거로운 Loess 변환보다는 Box-Cox 변환을 이용한 방법이 효과적이라고 할 수 있다.

4.3 5년간 자료에 대한 두 신용평가모형 비교

Box-Cox 방법을 이용해 변환시킨 재무변수를 통계적인 방법인 단계적 변수선택법을 이용하여 구한 모형이 효과적임을 보이기 위하여 1995년부터 2004년까지의 10년간 재무자료를 5년 단위로 나누어 1995년부터 1999년까지의 자료(정상: 1574 부도: 120)와 2000년부터 2004년까지의 자료(정상: 2527 부도: 47)를 바탕으로 4.1절과 4.2절에서 구한 최종 신용평가모형에 적용하여 결과를 비교한다.

1995년부터 1999년까지의 모형

1995년부터 1999년까지의 재무자료를 4.1절과 4.2절에서 얻은 신평모형Ⅱ와 신평모형Ⅲ에 적용하여 구한 결과를 [표 4.3]과 [표 4.4]에 나타내었다.

[표 4.3] 신평모형Ⅱ 결과와 분류성과(Loess, 1995~1999년)

범주구분	변수	Estimate	Wald Chi-Square	p-값	검정 통계량	분류성과
	상수항	-0.2083	0.2440	0.6213	<i>D</i>	2.3258
	건설업	-0.8615	7.0263	0.0080		
	서비스업	-1.0553	1.9328	0.1644		
	IT	0.4387	1.1441	0.2848		
부채 상환능력	(자본총계/cpi) / (금융비용/cpiavg)	-1.0106	10.0023	0.0016	<i>I-PH</i>	0.9421
	총C/F대차입금비율	-0.8594	9.7552	0.0018	<i>K-S</i>	0.6529
유동성	단기유동성비율	-0.4971	10.9427	0.0009	<i>AR</i>	0.7918
안정성 (Leverage)	(부채총계/cpi-이익 잉여금/cpiavg)/(유형 자산/cpi)	0.1663	16.2975	<.0001	<i>AUROC</i>	0.8959
수익성	매출총이익/이자비용	-0.4635	5.3776	0.0204	<i>Concordant</i>	0.8910

[표 4.4] 신평모형Ⅲ 결과와 분류성과(Box-Cox, 1995~1999년)

범주구분	변수	Estimate	Wald Chi-Square	p-값	검정 통계량	분류성과
	상수항	2.2184	0.9864	0.3206	<i>D</i>	2.3032
	건설업	-0.3796	1.5302	0.2161		
	서비스업	-1.1815	2.4469	0.1178		
	IT	0.6441	2.5197	0.1124		
안정성 (Leverage)	고정자산/차입금	-2.4621	11.8963	0.0006	<i>I-PH</i>	0.9555
부채 상환능력	(자본총계/cpi)/(금융비용/cpiavg)	-1.1397	23.8587	<.0001	<i>K-S</i>	0.6208
수익성	총자본 경상이익율	-0.0435	3.6368	0.0565	<i>AR</i>	0.7851
유동성	유동부채 구성비율	0.0243	0.1469	0.7016	<i>AUROC</i>	0.8925
활동성	(유동부채/cpi)/(매출액/cpiavg)	9.5774	23.6534	<.0001	<i>Concordant</i>	0.8910

우선 2000년 이전의 Loess 변환 모형(신평모형Ⅱ)의 결과인 [표 4.3]을 전체년도의 Loess변환 모형(신평모형Ⅰ)의 결과인 [표 4.1]과 비교하면 여섯 개의 검정 통계량 중에서 두 종류(1-PH, K-S) 통계량이 더 큰 값을 보이고 나머지 네 종류(D, AR, AUROC, Concordant)의 통계량은 2000년 이전의 모형에서 더 좋은 결과를 보인다. 그리고 Box-Cox 변환 모형(신평모형Ⅲ)의 결과인 [표 4.4]를 [표 4.2]와 비교하면, 신평모형Ⅰ의 두 종류(1-PH, K-S,) 통계량이 더 큰 값을 보이고 나머지 네 종류(D, AR, AUROC, Concordant) 통계량은 2000년 이전의 모형이 더 좋은 결과를 보인다. 또한 1995년부터 1999년까지 5년간의 재무자료를 바탕으로 구한 [표 4.3]과 [표 4.4]를 비교해보면, 모든 검정 통계량 값에서 비슷한 결과를 보인다.

그러므로 10년간 방대한 자료가 아니라도 신용평가모형을 개발하는데 큰 무리가 없으며, 또한 복잡하고 번거로운 Loess 변환방법을 이용하기보다 비교적 변환이 간단하고 원자료의 특성을 손실하지 않는 Box-Cox 변환을 통해서 신용평가모형(신평모형Ⅲ)을 개발하는 것도 좋은 방법이라고 판단한다.

2000년부터 2004년까지의 모형

다음으로 2000년부터 2004년까지의 재무자료를 Loess와 Box-Cox 변환하여 구한 신용평가모형을 구하여 결과를 [표 4.5]와 [표 4.6]에 정리하였다.

2000년 이후 5년간의 재무자료를 바탕으로 구한 [표 4.5]와 [표 4.6]을 비교해보면, 정상에 비해 부도의 비율이 현저히 떨어지므로 [표 4.1]과 [표 4.2]의 결과와 종합적으로 비교하여 분류성과가 다소 낮게 나타났으며 모든 검정 통계량값에서 비슷한 결과를 보인다. 그리고 Loess 변환 모형(신평모형Ⅱ)의 결과인 [표 4.5]을 [표 4.1]과 살펴보면 여섯 개의 검정 통계량 중에서 하나의 통계량(D)만이 2000년 이후의 모형에서 더 좋은 결과를 보이고, 이런 현상은 Box-Cox 변환 모형(신평모형Ⅲ)의 결과인 [표 4.6]과 [표 4.2]에서도 파악된다.

그러므로 5년간 자료에서도 신용평가모형을 개발하는데 무리가 없으며, 나아가 Loess 변환보다 변환방법이 쉽고 자료의 정보를 손실하지 않는 Box-Cox 변환을 통해서 신용평가모형(신평모형Ⅲ)을 개발하는 것이 효율적인 방법이라고 제안한다.

[표 4.5] 신평모형Ⅱ 결과와 분류성과(Loess, 2000~2004년)

범주구분	변수	Estimate	Wald Chi-Square	p-값
	상수항	-3.3900	76.4251	<.0001
	건설업	-1.3964	5.3913	0.0202
	서비스업	-0.2017	0.1018	0.7497
	IT	0.8668	5.9066	0.0151
부채 상환능력	(자본총계/cpi) / (금융비용/cpiavg)	-0.0710	0.3989	0.5277
	총C/F대차입금비율	-0.0128	0.0202	0.8869
유동성	단기유동성비율	-0.1576	3.3734	0.0663
안정성 (Leverage)	(부채총계/cpi-이익 잉여금/cpiavg)/(유형 자산/cpi)	0.1947	20.3008	<.0001
수익성	매출총이익/이자비용	-0.2476	4.4256	0.0354

검정 통계량	분류성과
<i>D</i>	1.4273
<i>1-PH</i>	0.9418
<i>K-S</i>	0.5382
<i>AR</i>	0.6330
<i>AUROC</i>	0.8165
<i>Concordant</i>	0.8010

[표 4.6] 신평모형Ⅲ 결과와 분류성과(Box-Cox, 2000~2004년)

범주구분	변수	Estimate	Wald Chi-Square	p-값
	상수항	-6.4963	7.5707	0.0059
	건설업	-0.0200	0.0012	0.9723
	서비스업	-0.2048	0.0941	0.7591
	IT	0.9757	6.0854	0.0136
안정성 (Leverage)	고정자산/차입금	0.9336	1.3006	0.2541
부채 상환능력	(자본총계/cpi)/ (금융비용/cpiavg)	-0.5928	11.0062	0.0009
수익성	총자본경상이익율	-0.0493	9.3795	0.0022
유동성	유동부채구성비율	0.0743	1.2426	0.2650
활동성	(유동부채/cpi)/ (매출액/cpiavg)	3.6708	2.9794	0.0843

검정 통계량	분류성과
<i>D</i>	1.4135
<i>I-PH</i>	0.9663
<i>K-S</i>	0.5357
<i>AR</i>	0.6270
<i>AUROC</i>	0.8145
<i>Concordant</i>	0.8000

5. 결론

기업 신용평가모형은 경기 변동, 신용평가 절차의 변화, 자료의 질에 관한 문제 등 다양한 이유로 신용평가모형의 예측력 또는 안정성이 저해될 수 있으며, 특히 신용평가모형에 사용되는 연속형 재무자료는 최적 가중값 및 평균 부도율도 시간이 지남에 따라 변화할 수 있기 때문에 자료수집 및 최신 자료로 갱신하는 어려움 등의 한계성 그리고 재무자료의 특성상 심한 왜도의 발생이 생기는 문제점 등이 있다. 따라서 신용평가모형에 의해 예측된 부도율을 정기적으로 점검하고 그 적합성을 검정할 필요가 있다.

Loess 방법을 이용한 변환은 원래 재무자료를 예측부도율로 변환하여 모형에 적용한다. 이는 원자료의 고유한 특성을 배제하고 182개 자료 내의 50개 구간 부도율을 평활모수에 의해 추정된 예측부도율로 변환함으로써 자료가 갖고 있는 고유의 정보 손실을 초래하는 문제점이 발생한다. 그리고 'By year method'에 의한 Loess 방법을 사용함으로써 연도별 재무변수의 흐름이 변동함에도 불구하고 그러한 특수 상황이 무시됨에 따라 정보의 손실이 예상된다. 최근에 수집되어 갱신한 각 기업의 관측값에 대해 부도인지 정상인지를 판단하기 위해서 182개 재무변수의 50구간의 예측부도율을 산출하고 적지 않은 재무변수의 평활곡선을 보면서 적절한 평활모수를 찾아서 변환시키는 작업을 번거롭게 다시 수행해야하는 문제가 있어 모형의 안정성을 저해한다. 또한 수집된 대부분의 재무변수는 매우 심한 왜도를 갖고 있어 모형의 적합성을 저해하는 요인으로 판단된다.

본 연구에서는 기존에 개발된 신용평가모형의 문제점을 살펴보고, 재무자료를 Loess 변환한 후에 통계적 방법 중의 하나인 단계적 변수선택법을 이용하여 구한 신용평가모형(신평모형Ⅱ)은 금융전문가의 신용평가모형(신평모형Ⅰ)만큼 좋은 성과를 갖는다는 것을 발견하였다. 그리고 왜도가 심한 재무자료의 정규성 및 대칭성을 향상

시켜 모형의 적합성 및 안정성을 높이기 위해 Yeo와 Johnson(2000)의 새로운 파워변환(Box-Cox 변환)한 재무변수들을 통계적인 방법인 단계적 변수선택 방법을 사용하는 신용평가모형(신평모형Ⅲ)을 제안하였다. Box-Cox 변환한 재무자료를 바탕으로 단계적 변수선택법을 이용하여 구한 신용평가모형이 유사한 성과를 나타내므로 변환 과정에서 복잡하고 번거로운 Loess 변환보다는 Box-Cox 변환을 이용한 자료를 통계적 변수선택 방법을 사용한 신용평가모형(신평모형Ⅲ)을 개발하여 사용하는 것이 효과적임을 보였다. 10년간의 방대한 자료를 5년 단위로 나누어 5년간의 자료에 대하여 다양한 신용평가모형을 적용하여 성과를 비교해 본 결과에서도 Loess 변환보다 쉽고 자료의 정보를 손실하지 않는 Box-Cox 변환을 사용하고, 단계적 변수선택 방법으로 신용평가모형을 개발하는 효율적인 방법을 제안한다.

그러므로 'By year method' 단계를 통해 Loess 방법으로 변환하는 과정에서 년도에 따른 재무비율의 연동흐름의 무시와 재무비율을 예측부도율로 대체하여 사용하는 등 복잡하고 정보손실을 가져오는 과정을 실행하지 않더라도, 단순하게 재무비율을 Box-Cox 변환한 재무자료를 사용함으로써 효율성과 안정성을 높일 수 있다는 결론을 내릴 수 있다.

참고문헌

1. 감사지원국 (2004). 신용리스크 따라잡기, 금융감독원.
2. 박병수 (2004). 신용평가모형의 적합성 검증 방안 및 선결과제, 대구은행 경제리뷰, 2004, 7-8월호.
3. 임중건 (2005). 신용평가시스템에 대한 적합성 검증 방법론 소개, 금융감독원, Risk Review, 2005. 08.
4. Box, G. E. P., and Cox, D. R. (1964). An Analysis of Transformation, *Journal of the Royal Statistical Society*, B26, 211-252.
5. Cleveland, W. S. (1979). Robust locally-weighted regression and smoothing scatterplots, *Journal of the American Statistical Association*, 74, 829-836.
6. Cohen, R. A. (1999). *An introduction to PROC LOESS for Local Regression*, Paper 273, SAS Institute Inc., Cary, North Carolina, USA.
7. Engelmann B., Hayden E., and Tasche D. (2003). Measuring the Discriminative Power of Rating Systems, Discussion paper, Series 2: *Banking and Financial Supervision*, No. 01/2003.
8. Hand, D. J., and Henley, W. E. (1997). Statistical Classification Methods in Consumer Credit Scoring: A Review, *Journal of the Royal Statistical Society*, 160, Part 3, 523-541.
9. Joseph, M. P. (2005). A PD Validation Framework for Basel II Internal Ratings-Based Systems, *Credit Scoring and Credit Control IX*.
10. Stein, R. M. (2005). The relationship between default prediction and lending profits, *Journal of Banking & Finance*, 29, 1213-1236.

11. Thomas, L. C., Edelman, D. B., and Crook, J. N. (2004). *Readings in Credit Scoring*, Chapter 4, Oxford University Press.
12. Yeo, I. K., and Johnson, R. A. (2000). A new family of power transformations to improve normality or symmetry, *Biometrika*, 87, 4, 954-959.

[접수일(2008년 6월 25일), 수정일(2008년 7월 15일), 게재확정일(2008년 7월 20일)]