

기술금융을 위한 부실 가능성 예측 최적 판별모형에 대한 연구*

A Study on the Optimal Discriminant Model Predicting
the likelihood of Insolvency for Technology Financing

성용현(Oonghyun Sung)**

목 차

- | | |
|----------------|---------------|
| I. 서론 | IV. 연구 분석 결과 |
| II. 기존연구와 개선방향 | V. 분류 예측력과 활용 |
| III. 연구 설계 | VI. 결론 |

국 문 요 약

본 연구는 기술력평가에 근거해서 중소기업 부실예측 가능성을 사전에 예측할 수 있는 최적 판별 모형을 개발 제안하였다. 판별모형에 포함될 설명변수는 요인분석과 판별모형의 단계별 선택방법에 의하여 선정되었다. 분석결과 선형판별모형이 로지스틱판별모형보다 임계확률 관점에서 적절한 것으로 나타났다. 최적 선형판별모형의 분류 정분류율은 70.4%, 분류 예측력은 67.5% 로 나타났다. 최적 선형판별모형의 활용도를 높이기 위해서 확실 범주와 유보범주를 구분할 수 있는 경계값을 설정하였다. 분석결과를 활용하면 기술금융 취급기관은 부실위험 평가와 더불어 기술금융 신청 기업의 순위를 부여할 때 유용하게 사용할 수 있을 것으로 기대된다.

핵심어 : 기술금융, 최적 판별모형, 정분류율, 유보범주, 경계값

* 이 논문은 2007년도 한신대학교 학술연구비에 의하여 연구되었음

** 한신대학교 정보과학대학 정보통계학과 정교수, soh@hs.ac.kr 031-370-6767

ABSTRACT

An investigation was undertaken of the optimal discriminant model for predicting the likelihood of insolvency in advance for medium-sized firms based on the technology evaluation. The explanatory variables included in the discriminant model were selected by both factor analysis and discriminant analysis using stepwise selection method. Five explanatory variables were selected in factor analysis in terms of explanatory ratio and communality. Six explanatory variables were selected in stepwise discriminant analysis.

The effectiveness of linear discriminant model and logistic discriminant model were assessed by the criteria of the critical probability and correct classification rate. Result showed that both model had similar correct classification rate and the linear discriminant model was preferred to the logistic discriminant model in terms of criteria of the critical probability. In case of the linear discriminant model with critical probability of 0.5, the total-group correct classification rate was 70.4% and correct classification rates of insolvent and solvent groups were 73.4% and 69.5% respectively.

Correct classification rate is an estimate of the probability that the estimated discriminant function will correctly classify the present sample. However, the actual correct classification rate is an estimate of the probability that the estimated discriminant function will correctly classify a future observation. Unfortunately, the correct classification rate underestimates the actual correct classification rate because the data set used to estimate the discriminant function is also used to evaluate them. The cross-validation method were used to estimate the bias of the correct classification rate. According to the results the estimated bias were 2.9% and the predicted actual correct classification rate was 67.5%. And a threshold value is set to establish an in-doubt category. Results of linear discriminant model can be applied for the technology financing banks to evaluate the possibility of insolvency and give the ranking of the firms applied.

Key words : technology financing, optimal discriminant model, correct classification rate, in-doubt category, threshold value

I. 서 론

정부와 4개 시중은행(국민은행, 기업은행, 신한은행, 우리은행)이 공동으로 시행중에 있는 “혁신형 중소기업 기술금융 사업”의 목적은 기술력 있는 혁신형 중소기업을 대상으로 전문평가기관¹⁾의 기술력평가를 통해 우수 기술기업을 선별하여 금융기관이 특허담보 또는 신용대출을 지원함으로써, 혁신형 중소기업의 성장을 지원하고 개발 기술의 사업화를 촉진하는 것으로 지난해부터 시행되고 있다. 그러나 지난 일년동안 시행된 기술금융 사업내용은 기대한 만큼 활성화가 되지 못한 실정이다.

2006년 말 기준으로 산업자원부 통계에 의하면 232개 기업이 이 사업에 신청하여 139건에 대한 대출심사가 종료되었고, 심사종료 139건 중 83건에 대해 365억원이 대출되었다. 대출 근거로 분류하면 순수 신용대출이 168억원(46%), 보증·담보 결합 신용대출이 152억원(42%), 보증·담보 대출이 45억원(12%)으로 나타났다. 기술력평가 실시 여부로 분류하면 기술력평가를 거쳐 대출된 경우가 185억원(51%), 기술력평가를 생략하고 대출된 것은 180억원(49%)이다²⁾. 전체 대출규모 365억 중에서 기술력평가를 거쳐 순수 신용대출을 받은 기업은 단지 15건 77억원(21%)으로, 기술력평가와 연계된 기술금융사업의 취지에서 크게 벗어나 있음을 알 수 있다.

이러한 통계는 금융기관들이 대출여부를 결정할 때 기술력평가에 근거한 신용대출 보다는 보증·담보를 결합하여 시행하고 있음을 보여주고 있고, 위험이 헷지(hedge)되지 않은 금융상품은 수용하기 어렵다는 반증이 된다. 특히 대출을 전담하는 금융기관이 어떤 이유에서든지 보증·담보를 요구하는 것은 바로 이러한 위험 헷지의 일환이라고 할 수 있다. 금융기관에서 일반기업의 대출 결정시에는 외부의 신용등급 혹은 자체적인 신용등급체계를 이용하지만, 혁신형 중소기업과 연관된 기술금융에서는 기술력 평가에 초점이 맞춰지는 것이 취지에 부합된다.

기술력평가와 기술금융을 연계하기 위해서는 대출 위험을 탐색할 수 있는 핵심정보가 기술력평가로부터 도출될 수 있어야 한다. 그러나 금융기관은 기술력평가의 적절성을 충분히 신뢰할 수 없기 때문에, 혁신형 중소기업에 대해서도 신용평가를 우선적으로 고려하고 있는 실정이다. 왜냐하면 신용평가 기관에서 제공하는 신용등급은 기업부실 여부를 판정하는데 검증된 통계적 적절성을 가지고 있지만, 기술력평가와 부실위험 사이의 연관성에 대

1) 기술력평가 기관은 중전 6개 기관에서 기술보증기금, 한국기술거래소, 한국발명진흥회, 한국과학기술정보연구원 등 4개 기관으로 축소 운영되고 있음.

2) 금융기관 상담과정에서 기업의 신용도 등에 따라 기술평가를 생략하거나 담보대출상품으로 안내하여 대출을 실시.

한 신뢰성은 아직 미흡하기 때문이다.

기술력평가는 기업이 보유하고 있는 기술능력과 더불어 기술의 사업화에 따른 전반적인 경쟁력을 평가하는 것으로 목적으로 한다. 만약 부실위험을 사전적으로 탐색할 수 있는 정보를 기술력평가로부터 얻을 수 있다면, 중소기업에 대한 기술금융을 활성화할 수 있는 하나의 방안이 될 수 있을 것이다. 일반기업의 부실위험에 영향을 미치는 변수 중에는 재무정보와 비재무정보가 종합적으로 포함될 수 있지만, 중소기업의 업력은 상대적으로 짧기 때문에 재무정보보다는 기술력평가 결과가 유용한 정보로 사용될 수 있을 것이다. 만약 중소기업의 기술력(기술인적자원, 기술성, 시장성, 사업성의 종합)이 평균보다 높게 평가될수록, 다른 조건이 유사하다면 부실위험 가능성은 상대적으로 감소될 것으로 기대할 수 있을 것이다.

본 연구의 목적은 중소기업의 부실위험 가능성을 예측할 수 있는 최적 판별모형을 제시하는 것이다. 본 연구의 구성은 다음과 같다. II장은 부실예측모형에 대한 기존연구에 대한 문제점과 개선방향을 제안하였고, III장은 기술력평가 항목의 구성과 최적 판별모형을 도출하기 위한 단계별 연구 설계로 구성되었다. IV장은 요인분석과 단계별 변수선택 분석에 의하여 부실위험에 유의한 영향을 미치는 주요 설명변수를 선택하였고, 선형판별모형과 로지스틱판별모형의 분류 설명력을 비교하여 최적 판별모형을 도출하였다. V장에서는 최적 판별모형의 분류 예측력을 검토하기 위해서 교차타당성 방법을 적용하였고, 실무적 적용과정에서 분류 예측의 신뢰성을 확보하기 위해서 경계값을 이용한 확실범주와 유보범주의 개념을 도입하였다.

II. 기존연구와 개선방향

기업부실 예측모형은 기업의 재무정보와 비재무정보를 이용하여 수년간 상당한 연구가 이루어졌다. 기업부실 예측모형에 대한 연구 영역은 기업의 부실(혹은 파산)에 유의한 영향을 미치는 주요 변수의 탐색과 더불어 다양한 판별모형의 예측력에 대한 비교이다. Atman(1968)은 선형판별함수를 이용하여 기업파산에 영향을 미치는 주요 재무비율로 (유동자산-유동부채)/총자산, 이익잉여금, 경상이익/총자산, 주가/장부가 총부채, 매출액/총자산 등을 선택하였다. Atmam(1968) 연구의 연장선상에서 Platt(1991), Poston, Harmon, Gramlich(1994), Clark, Webster, Hogan, Foster (1997), Zhang, Hu, Patuwo, Indro(1999)는 미국 여러 산업별로 시계열 재무자료에 선형판별모형, 로짓모형, 신경망모형 등을 적합

하여 기업파산 예측모형을 개발하였다. 성용현(1997)은 중소기업파산 예측모형으로 로지스틱판별모형을 적용하여 정상기업과 부실기업을 판별하는데 통계적으로 유의한 변수로 총차입금 대 자기자본비율, 금융기관 여신 거래실적, 사업안정성 등을 선정하였다. 선행연구결과 판별모형의 파산예측력은 사용된 주요 변수의 선택과 표본기업의 선택, 파산시점으로부터 기간 설정, 적용된 판별모형에 따라 서로 다르게 나타났고, 분류 예측력 대부분은 65-75% 범위에서 구해졌다.

기술력평가와 경영성과사이의 연관관계를 분석한 국내논문은 다음과 같다. 양동우(2003, 2005)는 벤처기업 22개에 대한 회귀분석 결과 전반적인 기술평가 수준이 미래 매출 성장에 유의한 영향을 미치고 있는 것으로 나타났다. 양동우(2005, 2006)는 기술력평가 항목을 재구성한 4개 대항목(기술시장성, 기술성, 경제성, 생산성)과 기업성과사이의 관계를 로지스틱판별모형을 이용하여 분석하였다. 분석결과 통계적으로 유의한 변수는 기술성, 경제성, 생산성 등이고, 기술시장성은 유의하지 않은 것으로 나타났고, 설정된 로지스틱판별모형의 정분류율 범위는 59.4-65% 로 나타났다. 성용현(2004)은 로지스틱판별모형과 로지스틱판별모형을 이용하여 기술력평가에서 기술성변수와 사업성범주사이의 연관성을 분석하였다. 세계 사업성범주인 경우 설정된 판별모형의 분류 설명력은 67.7%, 두개 사업성 범주인 경우 95.6% 로 나타났다. 성용현(2006)은 기술력표준모형 13개 중항목을 모두 포함한 선형판별모형과 로지스틱판별모형을 이용하여 중소기업의 파산 가능성을 예측하였다. 분석결과 선형판별모형과 로지스틱판별모형의 정분류율은 각각 69.2%, 68.7% 로 나타났다.

상기 기술력평가와 경영성과사이의 연관성 분석에서 주요 문제점과 개선방향을 정리하면 다음과 같다. 첫째, 중소기업 부실위험에 유의한 영향을 미치는 주요 변수의 선택하는 방법을 개선할 필요가 있다. 기술력평가는 대항목, 중항목, 소항목으로 구성된다. 중항목 혹은 대항목은 연관된 항목들의 합으로 구해지기 때문에, 특정 세부항목의 심각한 평가가 서로 상쇄될 수 있는 가능성이 존재한다. 그러나 세부 소항목을 판별모형에 모두 포함시킨다면, 판별모형의 복잡성과 더불어 변수들 사이의 중복성 문제가 제기 될 수 있다. 그리고 통계적 선택방법에 의하여 변수를 선택할 경우 핵심변수가 제외될 가능성이 있다. 따라서 부실예측에 영향을 미칠 수 있는 주요 변수의 선택은 핵심 변수의 선택과 더불어 통계적 검정을 통한 변수 선택을 함께 고려할 필요가 있다.

둘째, 추정된 판별모형에서 특정 범주로 분류하기 위한 기준인 임계확률(critical probability)설정의 적절성과 분류 범주의 설정이다. 관측범주가 두개이고 사전확률이 미지일 때 특정범주로 분류하기 위한 임계확률은 0.5 를 적용하는 것이 일반적이다. 임계확률의 설정은 설정된 판별모형과 자료의 속성에 따라 달라질 수 있지만, 분류 예측력을 높이기

위해서 임의로 임계확률을 변경 설정하는 것은 적절하지 않다. 성웅현(2006)은 판별모형에서 특정 범주에 속한 임계확률을 사후확률분포 요약통계로부터 설정하는 것을 제안하였다. 또한 분류기준은 사후확률이 임계확률이상이면 특정 범주로 자동 분류되기 때문에, 사후확률이 임계확률 주위에서 산출되면 예측된 범주의 확실성에 문제가 제기될 수 있다. 따라서 특정범주에 속할 사후확률이 일정 기준보다 크지 않을 경우 유보범주를 설정할 필요가 있고, 확실범주와 유보범주를 구분할 수 있는 경계값(threshold value)의 설정이 필요하다.

셋째, 설정된 판별모형에 대한 분류 예측력 평가이다. 이론적으로 해당표본에서 구한 정분류율은 판별모형의 예측력을 과대평가하는 경향이 있기 때문에, 판별모형의 예측 편의(bias)를 추정하여 정분류율을 수정할 필요가 있다. 양동우(2005, 2006)는 전체자료를 분석용과 검증용으로 양분하여 분석하는 표본분할법을 사용하였고, 성웅현(2006)은 개별범주에 속한 자료가 충분하지 않기 때문에 교차타당성방법을 적용하였다. 본 연구에서는 상기 문제점을 고려하여 중소기업 부실예측 가능성을 탐색하기 위한 최적 판별(분류)모형의 설정과 활용방안을 제시하였다.

Ⅲ. 연구 설계

1. 기술력평가 항목구성과 표본자료

기술력평가 표준모형 항목은 <표 1>과 같이 4개 대항목, 13개 중항목, 46개 소항목으로 구성되어 있고, 개별 소항목은 5점 척도로 평가된다. 기술력평가 표준모형 대항목은 기술자원과 인적자원, 기술성, 시장성, 사업성 등 4개 대항목으로 구성되었다. 기술력평가 항목 개발은 5개 전문평가기관³⁾ 전문가 의견을 수렴하여 연역적으로 구성하였고, 13개 중항목에 대한 단일차원성 확인요인분석(confirmatory factor analysis)을 통하여 소항목 구성의 적절성이 검토되었다. 따라서 기술력평가 표준모형은 일련의 대항목, 중항목, 소항목 평가항목 구성의 연관성은 체계적인 틀을 갖추고 있다고 판단된다.

기술력평가와 경영성과사이의 연관성을 탐색하기 위해서 중소기업청에서 2001년 당시

3) 기술력평가 표준모형 개발에 참여한 기관은 한국기술거래소, 한국발명진흥회, 산업은행, 기술신용보증기금, 한국과학기술정보원 등 5개 기관이다.

기술신용보증기금으로부터 기술력평가를 받은 중소벤처기업 482개를 선정한 다음, 2005년 시점에서 경영성과를 조사하여 수익이 발생되고 있는 361개 정상기업과 파산상태로 관측된 121개 기업의 자료를 수집하였다. 본 연구에서 사용된 자료는 482개 중소기업의 2001년에 실시된 기술력평가 결과를 다시 기술력평가 표준모형으로 재평가하여 수집된 것이다.

〈표 1〉 중소기업 기술력평가 표준모형 평가항목 구성

대항목	중항목	소항목
기술자원과 인적자원	기술경영능력	기술경험 수준(X1), 경영자의 기술이해도(X2), 경영자의 기술전문성(X3), 위기관리능력(X4), 기술개발전담조직(X5)
	인적자원 및 연구개발능력	기술개발추진능력(X6), 기술개발실적(X7), 기술장비보유 현황(X8), 연구개발투자비율(X9)
	기술개발환경	연계현황(X10), 추진형태(X11), 정보수집능력(X12)
기술성	기술의 우수성	기술의 차별성(X13), 기술의 수명주기상 위치(X14), 모방용이성(X15), 기술의 완성도(X16), 기술의 자립도(X17), 사업전략과의 부합성(X18)
	기술의 경쟁성	대체기술 출현가능성(X19), 산업적 파급효과(X20), 기술의 응용 및 확장가능성(X21)
	권리성	권리안정성 또는 확보가능성(X22), 권리의 이전가능성(X23), 권리의 범위(X24)
시장성	시장특성	시장구조(X25), 시장점유율(X26), 시장의 성장성(X27)
	시장환경	시장에서의 선호도(X28), 관련산업 동향과의 부합성(X29), 법·규제 등 제약/장려요인(X30)
	산업 환경성	산업시장 진입성(X31), 대체품과의 우위성(X32), 기업간 경쟁강도(X33), 구매업자와의 교섭력(X34)
사업성	생산기반	생산시설 확보용이성(X35), 재료 및 부품조달용이성(X36), 자본조달용이성(X37), 표준화 적합성(X38)
	생산성	부가가치창출능력(X39), 노동생산성(X40), 원가우위성(X41)
	수익성	매출액영업이익률(X42), 투자이익률(X43), 매출성장성(X44)
	마케팅전략성	마케팅계획의 적절성(X45), 판로의 다양성 및 구축성(X46)

판별분석 결과 신뢰성은 자료에서 구한 표본공분산행렬(sample covariance matrix)의 적정성 여부에 의하여 좌우된다. 만약 자료에 다변량 이상값이 다수 포함되어있는 경우 표본공분산행렬 구성요소에 민감한 영향을 미칠 수 있기 때문에, 판별분석 전에 기술력평가 자료에 대한 다변량 이상값 존재 여부를 탐색하여야 한다. 성웅현(2006)은 482개 자료에서 다변량 이상값을 탐색하기 위하여 Yang, Lee(1987)가 제안한 식(1)과 같은 F -검정을 사용하였다. 식(1)에서 X_j 는 j -번째 다변량 자료벡터이고 \bar{X} 는 표본평균벡터, S 는 표본공분산행렬, n 은 자료수, p 는 변수 개수이다.

$$F_j = \frac{n-p-1}{p} \times \left[\frac{1}{1 - n d_j^2 / (n-1)^2} - 1 \right], \quad d_j = (X_j - \bar{X})' S^{-1} (X_j - \bar{X}) \quad (1)$$

유의수준 0.05 에서 $F_j > F_{0.05, p, n-p-1}$ 로 탐색된 다변량 이상값 자료는 35개로 전체 자료의 7.3% 로 나타났고, 범주별로는 정상범주에서 23개 파산범주에서 12개가 탐색되었다. 따라서 판별분석에 사용된 표본자료는 <표 2>와 같이 전체자료에서 35개 이상값 자료를 제외한 정상범주 338개(75.6%), 파산범주 109개(24.4%) 등 총 447개 이다.

<표 2> 경영성과 분류에 대한 자료

구 분		전체자료	이상값 자료	표본자료
경영성과	파산	121	12	109
	정상	361	23	338
합 계		482	35	447

2. 판별분석 단계별 설계

기술력평가 표준모형 46개 소항목 중에서 경영성과(파산, 정상)범주에 영향을 미치는 주요 변수를 선택하기 위해서 요인모형과 판별모형을 병행하여 사용하였고, 중속기업 부실 예측모형으로서 활용도를 개선하기 위해서 <표 3>과 같은 단계별 분석절차를 적용하였다.

〈표 3〉 연구 분석 단계별 설계 내용

분석 1단계	분석 2단계	분석 3단계
주요 변수선택	판별모형 설명력	판별(분류)모형 활용
(1) 요인분석을 이용한 변수 선택 (2) 단계별 선택방법을 이용한 변수 선택	(1) 특정범주 분류 기준인 임계확률 결정 (2) 두가지 판별모형의 정분류를 비교 분석	(1) 최적 판별모형의 선택과 예측력 평가 (2) 확실/유보범주 구분을 위한 경계값 설정

분석 1단계 내용은 경영성과 범주(파산, 정상)에 영향을 미칠 수 있는 주요 변수를 선택하는 것이다. 기술력평가 개별 중항목을 단일 요인으로 설정하고 연관된 소항목 중에서 연관성이 가장 높은 소항목을 요인분석을 통하여 선택하는 것이다. 판별모형에 포함될 변수는 요인분석에서 선택된 변수와 판별모형에서 단계별 선택방법(stepwise selection method)에서 선택된 변수로 구성된다. 이러한 절차를 적용한 이유는 후자와 같은 변수선택 방법만을 적용하였을 때, 주요 핵심 변수임에도 불구하고 단계별 선택과정에서 통계적 유의성이 상대적으로 작아 포함되지 못하는 문제가 발생될 수 있기 때문이다.

분석 2단계 내용은 추정된 판별함수를 이용하여 경영성과 두개 범주를 분류하기 위한 적절한 임계확률의 설정과 분류 설명력 평가이다. 특정범주에 속할 사전확률⁴⁾이 알려져 있지 않은 경우 임계확률로 0.5 가 일반적으로 적용된다. 미래 경영성과 범주에 속할 사전확률이 미지인 경우 표본 분석 결과에서 구한 특정 범주에 속할 사후확률분포⁵⁾ 중심의 대표값을 이용하여 임계확률을 설정할 수 있다. 만약 사후확률분포가 대칭인 경우에는 대표값으로 평균을 사용할 수 있지만, 비대칭인 경우에는 평균보다는 중앙값을 적용하는 것이 적절하다. 그러나 표본분석에서 구한 임계확률은 범주별 표본자료의 구성과 충분성에 따라 달라질 수 있다. 따라서 특정 범주에 속할 사전확률이 미지인 경우 0.5 보다 상당히 큰 임계확률을 적용하는 것은 적절하지 않다고 판단된다. 적절한 임계확률이 결정되면 추정된 판별함수의 분류 설명력을 평가하기 위해서 정분류를 적용하였다.

분석 3단계 내용은 최적 판별모형의 부실위험 예측력을 평가하고, 판별모형을 이용하여 개별 중소기업의 향후 부실 가능성을 예측하는 것이다. 만약 특정범주로 분류될 사후확률이

4) 특정 범주에 속할 사전확률은 장기간에 걸친 자료 수집을 통하여 추정될 수 있고, 모집단의 특성을 대표할 수 있다고 판단되면, 추정된 사전확률은 향후 특정범주에 속할 임계확률로 적용될 수 있다. 그러나 본 연구에서 사용한 표본자료가 충분하지 못하기 때문에 특정 범주에 속할 표본비율을 사전확률로 일반화하기에는 무리가 있다.

5) 사후확률(posterior probability)이란 자료가 주어졌을 때 판별(혹은 분류)함수에 의하여 사후적으로 추정된 특정범주에 속할 확률(가능성)을 의미한다.

임계확률과 유사한 경우 예측된 범주의 확실성에 문제가 제기된다. 이러한 문제를 해결하기 위해서 특정범주에 속할 사후확률이 경계값(임계확률보다 큰 값임)보다 크지 않으면 그 범주로 분류하지 않고 유보범주로 구분하여 설정하는 것이 적절하다. 경계값 수준의 설정은 활용 목적과 금융기관의 위험부담 수준에 따라 달라질 수 있지만, 본 연구에서는 사후확률의 요약통계를 활용하여 가능한 경계값을 탐색하여 결정하였다.

IV. 연구 분석 결과

1. 주요 설명변수의 선택

부실예측 판별모형에 포함될 주요 설명변수를 선택하기 위해서 두가지 절차를 적용하였다. 첫번째 절차는 <표 1> 기술력평가 개별 중항목을 단일요인으로 설정한 요인모형⁶⁾을 이용하여 가장 연관성이 높은 소항목을 설명변수로 선택하였다. 설명변수 선택 기준은 단일요인의 설명비율이 50% 이상이고 개별 소항목의 공통성(communality)이 0.6 이상인 것 중에서 가장 큰 것을 선택하였다. 요인분석 결과 선택된 설명변수는 기술개발 환경 중항목에서 '추진형태(X11)', 권리성 중항목에서 '권리안정성 또는 확보가능성(X22)', 생산성 중항목에서 '노동생산성(X40)', 수익성 중항목에서 '투자이익률(X43)', 마케팅 전략성에서 '마케팅계획의 적절성(X45)' 등으로 5개 중항목에서 하나씩 선택된 5개 소항목이다. 나머지 8개 중항목과 연관된 소항목은 선정된 기준에 미흡하기 때문에 제외하였다.

두번째 절차는 요인분석에서 선택된 5개 설명변수를 판별모형에 일단 포함시킨 다음, 나머지 41개 소항목 중에서 두개 경영성과 범주(파산 혹은 정상)에 통계적으로 유의한 설명변수를 단계별 선택방법에 의해서 선정하였다. 판별모형의 종속변수를 $Y=0$ (파산), $Y=1$ (정상), 자료벡터를 $X'=(X_1, X_2, \dots, X_p)$, 선형/로지스틱판별모형 계수벡터를

6) 예를 들면, 첫번째 중항목인 기술경영능력을 단일요인 F 으로 설정했을 때 연관된 5개 소항목 $X_i, i=1, 2, \dots, 5$ 의 요인모형은 $X_i = l_i F + \epsilon_i, i=1, 2, 3, 4, 5$ 이다. 단일요인의 설명비율은 5개 고유값의 합에 대한 첫 번째 고유값의 비율을 의미한다. $\lambda_1 = 1.947$ 이므로 단일요인의 설명비율은 $\lambda_1 / \sum_{i=1}^5 \lambda_i = 0.389$ 이 된다. 또한 공통성은 개별 소항목의 변동 중에서 단일요인에 의한 비율을 의미하고, 소항목의 변동은 1 이기 때문에 공통성이 $h_i^2 = l_i^2 > 0.6$ (사회과학에서는 0.5 기준도 적용함)인 경우에 적절한 변수로 선택한다. 첫번째 중항목 요인모형에서는 설명비율이 0.5 미만이기 때문에 설명변수를 선택하지 않았다.

$\mathbf{a}' = (a_1, a_2, \dots, a_p)$, $\boldsymbol{\beta}' = (\beta_0, \beta_1, \dots, \beta_p)$ 로 각각 표시하면 선형(정준)판별함수와 로지스틱판별함수는 식(2)과 식(3)과 같다.

$$Y = \mathbf{a}' \mathbf{X} = a_1 X_1 + a_2 X_2 + \dots + a_p X_p \quad (2)$$

$$P(Y = 1 | \mathbf{X}) = \frac{\exp(\boldsymbol{\beta}' \mathbf{X})}{1 + \exp(\boldsymbol{\beta}' \mathbf{X})} = \frac{\exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)}{1 + \exp(\beta_0 + \beta_1 X_1 + \dots + \beta_p X_p)} \quad (3)$$

식(2)의 선형판별함수는 두개 범주간 평균벡터 거리를 최대화할 수 있는 p 개 변수의 선형결합으로 표시되고, 계수벡터 추정은 $\hat{\mathbf{a}} = S_p^{-1}(\bar{X}_1 - \bar{X}_0)$ 로 구해진다. 여기서 \bar{X}_1 과 \bar{X}_0 는 정상범주와 파산범주에 속한 자료의 표본평균벡터, S_p^{-1} 는 합동표본공분산행렬⁷⁾의 역행렬을 의미한다. 식(3)의 로지스틱판별함수에서 $P(Y = 1 | \mathbf{X})$ 는 다변량자료 \mathbf{X} 가 주어졌을 때 정상범주로 분류될 사후확률을 의미하고, 계수벡터 $\boldsymbol{\beta}$ 의 추정은 Newton-Raphson 방법으로 근사해를 구한다. 판별모형에 대한 이론적 효율성 비교한 Brooks(1988)와 성웅현(1999)에 의하면 모집단에 대한 다변량 정규분포(multivariate normal distribution)를 가정하였을 때 선형판별모형과 로지스틱판별모형의 효율성은 유사한 것으로 나타났고, 다변량 정규분포에서 크게 위반된 경우에는 로지스틱판별모형이 선형판별모형보다 상대적으로 효율성이 높은 것으로 나타났다.

본 연구에서 다변량 자료가 다변량 정규분포로부터 크게 벗어난 이상값을 식(1)의 검정통계량을 통하여 제외하였기 때문에 두가지 판별모형을 함께 적용하였다. 추가적인 설명 변수를 선택하기 위해서 단계별 선택법⁸⁾을 적용한 결과 선형판별모형에서 선택된 10개 소항목은 <표 4>와 같고, 로지스틱판별모형에서 선택된 7개 소항목은 <표 5>와 같다.

7) 정상범주와 파산범주의 표본공분산행렬을 각각 S_1, S_0 라고 했을 때, 합동표본공분산행렬은 $S_p = ((n_1 - 1)S_1 + (n_0 - 1)S_0) / (n_1 + n_0 - 2)$ 로 구해진다.

8) 선형판별모형과 로지스틱판별모형의 단계별 변수선택과정에서 유의수준은 동일하게 0.15 를 적용하였다. 유의수준 0.15 를 선정한 이유는 변수 선택과정에서 영향을 미칠 수 있는 주요 변수를 포함시킨 다음 최종 검정과정에서 개별 계수의 p -값에 따라 그 유의성 여부를 판단하기 위해서다. 그리고 SAS 에서 선형판별의 단계별선택과정에서 적용하는 유의수준 default 는 0.15 이다.

〈표 4〉 선형판별모형에서 추가 설명변수 선택

중항목	선택 소항목	부분 F -값	p -값
기술경영능력	경영자의 기술전문성(X3)	3.28	0.0708
	위기관리 능력(X4)	16.64	<0.0001
인적자원 및 연구개발능력	기술개발 실적(X7)	2.18	0.1403
	연구개발 투자비율(X9)	4.37	0.0371
기술의 우수성	기술의 자립도(X17)	14.96	0.0001
산업환경성	산업시장 진입성(X31)	2.43	0.1195
	대체품과의 우위성(X32)	2.80	0.0953
	기업간 경쟁강도(X33)	3.13	0.0776
생산기반	표준화 적합성(X38)	8.34	0.0041
생산성	부가가치 창출능력(X39)	9.27	0.0025

〈표 5〉 로지스틱판별모형에서 추가 설명변수 선택

중항목	선택 소항목	Score Chi-Square	p -값
기술경영능력	위기관리 능력(X4)	15.7417	<0.0001
인적자원 및 연구개발능력	기술개발 실적(X7)	4.2436	0.0394
	연구개발 투자비율(X9)	4.3326	0.0374
기술의 우수성	기술의 자립도(X17)	14.1738	0.0002
산업환경성	기업간 경쟁강도(X33)	3.3320	0.0666
생산기반	자본조달 용이성(X37)	4.5171	0.0336
생산성	부가가치 창출능력(X39)	8.7643	0.0031

두가지 판별모형에서 선택된 설명변수가 서로 다르기 때문에 다음과 같은 절차를 통하여 추가적인 설명변수를 선택하였다. 두가지 판별모형의 변수선택에서 일치된 6개 소항목(X4, X7, X9, X17, X33, X39)을 우선 선택하였고, 일치되지 않은 나머지 소항목 중에서 변수선택은 전문가 입장에서 결정하였다. 선형판별모형에서 선택된 기술경영능력 중항목에

속한 '경영자의 기술전문성(X3)' 은 기술사업화에 따른 기술/사업 위협을 감소시킬 수 있는 변수이지만, 이미 동일한 중항목에서 선택된 위기관리능력(X4)과 연관성이 높다고 판단하여 제외하였다. 선형판별모형에서 산업환경성 중항목과 연관된 3개 소항목(X31, X32, X33)이 선택되었으나, 로지스틱 판별모형에서 1개 소항목(X33)만 선택하였다. 산업환경성 중항목과 연관된 3개 소항목간 상관관계는 0.308-0.356 범위에서 서로 유사하게 나타났고, 특히 '기업간 경쟁강도(X33)' 소항목은 기업의 향후 안정적인 시장점유와 연관이 높고 연관된 p -값이 가장 작기 때문에 선택하였다.

생산기반 중항목인 경우 선형판별모형에서 표준화 적합성(X38)이 로지스틱판별모형에서 자본조달용이성(X37)이 선택되었으나, 사업의 경영성과 범주와 연관된 사업위험을 평가할 때에는 '자본조달 용이성(X37)'이 '표준화 적합성(X38)'보다 상대적으로 적합할 것으로 판단되어 선택하였다. 따라서 요인분석과 판별분석의 단계별 선택방법에 의하여 선택된 12개 소항목을 정리하면 <표 6>과 같다.

<표 6> 최종 설명변수 선택 결과

중항목	소항목	중항목	소항목
기술경영능력	위기관리 능력(X4)	산업환경성	기업간 경쟁강도(X33)
인적자원 및 연구개발능력	기술개발 실적(X7)	생산기반	자본조달 용이성(X37)
	연구개발 투자비율(X9)		부가가치 창출능력(X39)
기술개발환경	추진형태(X11)	생산성	노동생산성(X40)
기술 우수성	기술의 자립도(X17)	수익성	투자이익률(X43)
권리성	권리 안정성 또는 확보 가능성(X22)	마케팅 전략성	마케팅계획 적절성(X45)

요인분석과 판별모형의 단계별 변수선택법에 의하여 설명변수를 <표 6>과 같이 선택하였다. 이러한 두가지 절차를 함께 적용한 이유는 후자와 같은 단계별 변수선택법만을 적용하였을 때, 주요 핵심 변수임에도 불구하고 단계별 선택과정에서 통계적 유의성 관점에서 포함되지 못하는 문제가 발생될 가능성이 있기 때문이다. 만약 두개 경영성과 범주에 유의한 영향을 미치는 설명변수를 요인분석 결과를 적용하지 않고 단계별 변수선택법에 의하여 선택한 결과는 <표 7>과 같다. 선택된 설명변수를 <표 6>과 비교하면 권리성, 수익성과 마케팅 전략성 등 중항목과 연관된 소항목은 선택되지 않았다.

통계적 분석 결과 생략된 변수 중에는 이론적으로 기업의 경영성과에 영향을 미칠 수 있는 주요 변수가 포함되어 있다고 판단된다. 권리성 중항목과 연관된 '권리안정성 또는 확보가능성(X22)' 소항목은 등록된 권리가 무효화되지 않고 안정적으로 유지될 가능성을 평가하는 것으로, 특히 종속적인 제품일 경우 권리의 안정성이 강할수록 상대적인 경쟁력이 높아져 시장에서 제품 경쟁력과 경제적 수명에 긍정적인 영향을 미칠 것으로 판단된다. 수익성 중항목과 연관된 '투자이익률(X42)' 소항목은 업종평균 자본수익률 대비 해당 기업 자본수익률을 평가하는 것으로, 기업의 수익성과 성장성에 긍정적인 영향을 미칠 수 있는 변수라고 판단된다. 또한 마케팅 전략성 중항목과 연관된 '마케팅 계획의 적절성(X45)' 소항목은 시장규모 분석, 경쟁자 분석, 시장세분화 전략, 광고 또는 홍보전략 등 전반적인 마케팅 전략이 객관적 자료를 근거로 추진하고 있는가를 평가하는 것으로, 마케팅 전략에 의한 시장침투와 대처능력은 경영성과에 긍정적인 영향을 미칠 것으로 기대된다. 따라서 상기와 같은 변수는 단계별 변수선택에서는 생략되었으나 최종 판별모형에 포함되는 것이 이론적으로 적절하다고 판단하였다.

〈표 7〉 판별모형에 의한 단계별 변수선택 결과

중항목	소항목	
	선형판별모형	로지스틱판별모형
기술경영능력	경영자의 기술전문성(X3) 위기관리 능력(X4)	위기관리 능력(X4)
인적자원 및 연구개발능력	기술장비보유 현황(X8)	기술개발 실적(X7) 연구개발 투자비율(X9)
기술개발환경	정보수집 능력(X12)	정보수집 능력(X12)
기술의 우수성	기술의 자립도(X17)	기술의 자립도(X17)
산업 환경성	산업시장 진입성(X31) 대체품과의 우위성(X32)	
생산기반	자본조달 용이성(X37) 표준화 적합성(X38)	자본조달 용이성(X37) 표준화 적합성(X38)
생산성	부가가치 창출능력(X39)	부가가치 창출능력(X39)

2. 임계확률의 설정과 판별모형의 설명력 비교

경영성과 범주를 판별하기 위하여 설정된 선형판별모형과 로지스틱판별모형의 분류 설명력을 비교하기 위해서 특정 범주로 분류될 기준인 임계확률을 설정하여야 한다. 특정 범주에 속할 사전확률이 알려져 있지 않은 경우 일반적으로 임계확률 0.5 를 적용하지만, 사후확률의 요약통계를 통해서 적절한 임계확률을 설정할 수 있다. 두가지 판별모형에 대한 분석결과 두개 관측범주(파산과 정상)에 속한 자료가 정상범주에 속할 사후확률에 대한 요약통계는 <표 8>과 같다.

<표 8> 정상범주에 속할 사후확률 요약통계

구분	선형판별모형 사후확률		로지스틱판별모형 사후확률	
	파산	정상	파산	정상
관측범주				
평균	0.39	0.61	0.62	0.80
중앙값	0.37	0.63	0.65	0.85
표준편차	0.18	0.21	0.19	0.16
일사분위수 ⁹⁾	0.25	0.47	0.51	0.73
삼사분위수	0.50	0.77	0.77	0.92
왜도	0.26	-0.37	-0.52	-1.32

<표 8>에서 범주별 사후확률분포의 왜도¹⁰⁾를 살펴보면 선형판별모형의 경우 대칭에서 약간 벗어나 있지만, 로지스틱판별모형의 경우 대칭에서 상대적으로 크게 벗어난 것으로 나타났다. 사후확률분포가 비대칭인 경우 일정 구간에서 임계확률을 변화하였을 때 정분류율의 영향을 평가하여 적절한 임계확률을 설정할 수 있다. 선형판별모형인 경우 임계확률의 범위를 0.35-0.65 로 하고, 로지스틱판별모형인 경우 0.50 -0.80 로 했을 때 개별 범주와 전체에 대한 정분류율은 <표 9>와 같다.

9) 일사분위수는 자료 분포에서 하위 25% 에 해당되는 값을 의미하고, 삼사분위수는 상위 25% 에 해당되는 값을 의미한다. 그리고 중앙값은 50% 에 해당되는 값이다.

10) 자료 분포의 좌우대칭 정도를 측정하기 위해서 왜도(skewness)는 -3 과 +3 사이의 값으로 표시된다. 왜도가 0인 경우 좌우대칭, 양수인 경우 오른쪽으로 긴 꼬리, 음수인 경우 왼쪽으로 긴 꼬리를 갖는 것을 의미한다.

〈표 9〉 다양한 임계확률에 대한 정분류율 결과

선형판별모형				로지스틱판별모형			
임계확률	파산	정상	전체	임계확률	파산	정상	전체
0.35	43.1%	84.9%	74.7%	0.50	23.9%	95.0%	77.6%
0.40	54.1%	81.1%	74.5%	0.55	34.0%	89.6%	76.1%
0.45	61.5%	76.3%	72.7%	0.60	40.4%	87.3%	75.8%
0.50	73.4%	69.5%	70.4%	0.65	51.4%	82.8%	75.2%
0.55	83.5%	61.2%	66.7%	0.70	60.6%	78.1%	73.8%
0.60	86.2%	53.6%	61.5%	0.75	70.6%	72.2%	71.8%
0.65	91.7%	46.5%	57.5%	0.80	83.5%	61.2%	66.7%

다양한 임계확률에서 사후적으로 적절한 임계확률을 설정하기 위해서는 개별 범주에 대한 정분류율이 가능한 균형을 이루면서 전체 정분류율이 일정 수준이상인 경우를 고려하였다. 〈표 9〉에 의하면 전체 정분류율이 70% 이상이면서 범주별 유사한 정분류율을 보이는 임계확률은 선형판별모형인 경우 0.5 이고 로지스틱판별모형인 경우 0.75 인 것으로 나타났다.

판별모형에서 사후적으로 적절한 임계확률을 설정할 수 있는 또 다른 방법은 〈표 8〉의 사후확률분포의 요약통계를 이용할 수 있다. 비대칭이 존재하는 경우 특정(정상)범주로 분류할 임계확률은 사후확률의 중앙값을 이용하여 구할 수 있다. 두개 관측범주에서 정상범주에 속할 사후확률분포의 중앙값을 다시 평균하면 선형판별모형인 경우 0.5 $[(0.37 + 0.63)/2]$ 이고, 로지스틱판별모형인 경우 0.75 $[(0.65 + 0.85)/2]$ 가 된다. 이와 같은 결과는 〈표 9〉 결과로 구한 임계확률과 동일하게 산출되었음을 알 수 있다. 임계확률 0.5 와 0.75 를 각각 적용한 선형판별모형과 로지스틱판별모형의 분류결과를 이원분류표로 표시하면 〈표 10〉과 같다. 〈표 10〉(b)에서 () 임계확률 0.5 를 적용한 경우이다.

〈표 10〉 판별모형 분류결과에 대한 이원분류표

(a) 선형판별모형				(b) 로지스틱판별모형			
관측 범주	예측 범주		전체	관측 범주	예측 범주		전체
	파산	정상			파산	정상	
파산	80	29	109	파산	77(26)	32(83)	109
정상	103	235	338	정상	94(17)	244(321)	338
전체	183	264	447	전체	171(43)	276(404)	447

선형판별모형(임계확률 0.5 적용)과 로지스틱판별모형(임계확률 0.75 적용)에서 전체 범주에 속한 표본자료에 대한 정분류율¹¹⁾은 각각 70.4 % 와 71.8% 로 나타났다. 두가지 판별모형사이의 정분류율 차이는 1.4% 로 매우 작게 나타났기 때문에, 표본자료에 대한 두가지 판별모형의 분류 설명력은 통계적으로 유의한 차이가 있다고 판단할 수 없다. 만약 로지스틱판별모형에 임계확률 0.5 를 적용한 분류결과를 살펴보면 파산범주에서 매우 심각한 오분류¹²⁾가 발생되었고, 반면에 정상범주에서 매우 높은 정분류율을 보이는 심각한 불균형을 보이고 있다. 로지스틱판별모형에 적용한 임계확률 0.75 는 사후적으로 표본분석 결과로부터 설정된 값이기 때문에 적절성 문제가 제기 될 수 있다. 왜냐하면 정분류율을 개선하기 위해서 임의의 임계확률을 사용할 가능성을 배제할 수 없고, 또한 해당 표본에서 설정된 임계확률이 다른 표본에 적용 가능한지 여부도 확신할 수 없기 때문이다. 반면에 선형판별모형에서 사후적으로 설정된 임계확률 0.5 는 사전확률이 알려져 있지 않을 때 일반적으로 적용하는 값과 일치하고 있다. 따라서 중소기업의 향후 부실여부 가능성을 판단할 수 있는 판별모형은 임계확률 설정과 정분류율을 고려할 때 선형판별모형이 로지스틱판별모형보다 상대적으로 적합하다고 판단된다.

성용현(2006)은 동일한 자료를 13개 중항목이 포함된 판별모형에 적합 시킨 결과 선형판별모형 정분류율은 69.2%, 로지스틱판별모형 경우 68.7% 로 나타났다. 이러한 정분류율 결과와 12개 소항목을 이용한 <표 10>의 결과와 비교하면, 전체 기술력 평가 46개 소항목 자료 대신에 12개 주요 소항목 자료를 활용하더라도 유사한 결과를 기대할 수 있는 근거가 될 수 있다.

V. 분류 예측력과 활용

1. 선형판별함수의 분류 예측력 평가

판별분석의 목적은 경영성과 범주의 차이를 최대한 분리할 수 있는 판별함수의 추정과 추정된 판별함수를 이용한 분류 예측력을 평가하는 것이다. IV장에서는 해당 표본에서 경

11) 이원분류표에서 정분류율은 전체자료에서 관측범주와 예측범주가 일치된 비율을 의미한다. 선형판별모형 정분류율은 $(80+235)/447=0.704$, 로지스틱판별모형 정분류율은 $(77+244)/447 = 0.718$ 이다.

12) 로지스틱판별모형에서 임계확률을 0.5 로 적용했을 때 파산범주의 정분류율은 $26/109=0.239$ 이고, 정상범주의 정분류율은 $321/338= 0.950$ 으로 범주간 심각한 불균형을 보이고 있다. 따라서 파산범주의 오분류율은 76.1% 로 매우 심각하고, 정상범주의 오분류율은 5% 로 매우 낮게 나타났다.

영성과 두가지 범주를 설명할 수 있는 적절한 판별함수로 선형판별함수를 선택하였다. 선형판별함수의 적절성을 확인하기 위해서 두개 범주의 공분산행렬이 동일한지 여부를 M -검정¹³⁾한 결과 선형판별함수가 적절한 것으로 나타났다.

〈표 10〉에서 선형판별함수의 파산범주와 정상범주에 대한 정분류율은 각각 73.4%, 69.5% 로 나타났고 전체범주에 대한 정분류율은 70.4% 이다. 이러한 결과는 해당표본에서 추정된 판별함수를 동일표본에 적용하여 구한 정분류율이기 때문에, 미래 다른 표본에서 대한 분류 예측력이라고 단정할 수 없다. 이론적으로 해당 표본에서 구한 정분류율은 실제 정분류율(예측력)을 과대 추정하는 경향이 있다. 실제 정분류율과 정분류율사이의 편의를 추정할 수 있는 방법은 해당표본을 분석표본과 검증표본으로 분할하여 추정하는 표본분할법(split sample method)과 교차타당성방법(cross validation method)¹⁴⁾ 이 있다. 본 연구에서는 개별범주에 속한 자료가 충분하지 않기 때문에 교차타당성방법을 적용하였고 분류결과는 〈표 11〉과 같다. 교차타당성 분석결과 실제 정분류율 추정값은 파산범주인 경우 66%, 정상범주인 경우 68%, 전체범주인 경우 67.5% 로 나타났다. 전체범주에 대한 정분류율 편의는 2.9% (70.4% -67.5%)로 구해졌기 때문에, 추정된 선형판별함수의 미래자료에 대한 분류 예측력 수준은 67.5% 정도로 기대할 수 있다.

〈표 11〉 교차타당성 방법에 의한 분류결과

관측 범주	예측 범주		전체
	파산	정상	
파산	72	37	109
정상	108	230	338
전체	180	267	447

13) Box (1949)가 제안한 M -검정통계량은 $M = \left[|S_0|^{n_0/2} |S_1|^{n_1/2} \right] / |S_p|^{\sum \nu_i / 2}$, $\nu_i = n_i - 1$ 이다. 범주가 $k=2$ 이고 변수가 $p=12$ 일 때 M 을 근사 χ^2 -분포로 변환하기 위해서 아래와 같은 c 을 이용하고, c 의 함수인 검정통계량 $u = -2(1-c) \ln M$ 은 자유도가 $0.5(k-1)p(p+1)$ 인 χ^2 -분포에 근사하게 된다. 검정결과 $u=92.470$ 이고 연관된 p -값이 $P(\chi_{78}^2 > 92.470) = 0.1258$ 이므로, 경영성과 두개 범주가 동일한 공분산행렬을 갖는다는 귀무가설을 기각할 수 없는 근거가 된다.

$$c = \left[\sum_{i=1}^k \frac{1}{\nu_i} - 1 / \sum_{i=1}^k \nu_i \right] \times \left[\frac{2p^2 + 3p - 1}{6(p+1)(k-1)} \right]$$

14) 교차타당성방법은 자료를 하나씩 제외하고 나머지 자료로 판별함수를 추정한 후 제외된 하나의 자료의 분류를 결정하는 절차로 모든 자료에 반복적으로 적용하여 예측력을 평가하는 방법이다.

2. 부실예측을 위한 선형판별함수의 활용

파산과 정상범주에 속한 기술력평가 12개 소항목 자료 벡터를 X_0 과 X_1 , 표본평균벡터를 \bar{X}_0 과 \bar{X}_1 , 표본합동공분산행렬을 S_p 라고 표시했을 때 추정된 선형판별(분류)함수 $\hat{Y}_i, i = 0, 1$ 은 식(4)와 같이 설정된다. 특정 기업의 기술력평가 자료 X 가 주어졌을 때 정상범주에 속할 사후확률 $P(Y=1|X)$ 은 <표 12>에서 구한 계수를 이용하면 식(5)와 같이 구해진다.

$$\hat{Y}_i = \hat{a}_i' X = \bar{X}_i' S_p^{-1} X - 0.5 \bar{X}_i' S_p^{-1} \bar{X}_i, \quad i = 0, 1. \quad (4)$$

$$P(Y=1|X) = \frac{\exp(\hat{Y}_1)}{\exp(\hat{Y}_0) + \exp(\hat{Y}_1)} = \frac{1}{1 + \exp(\hat{Y}_1 - \hat{Y}_0)} = \frac{1}{1 + \exp(\hat{a}' X)} \quad (5)$$

$$= \frac{1}{1 + \exp(-7.527 + 0.618 X_4 + 0.276 X_7 + \dots - 0.045 X_{45})}$$

<표 12> 추정된 선형분류함수의 계수

소항목	선형분류함수 계수		
	a_0	a_1	$a = a_1 - a_0$
	-55.110	-62.637	-7.527
위기관리 능력(X4)	5.433	6.051	0.618
기술개발 실적(X7)	0.937	1.213	0.276
연구개발 투자비율(X9)	2.625	2.360	-0.265
추진형태(X11)	0.773	0.707	-0.066
기술의 자립도(X17)	6.657	7.329	0.633
권리안정성 또는 확보가능성(X22)	3.083	2.960	-0.123
기업간 경쟁강도(X33)	2.712	2.996	0.284
자본조달 용이성(X37)	3.126	3.493	0.367
부가가치 창출능력(X39)	0.706	1.092	0.386
노동생산성(X40)	0.542	0.690	0.148
투자이익률(X43)	2.159	2.221	0.062
마케팅계획 적절성(X45)	4.621	4.576	-0.045

식(5)에서 구한 사후확률과 임계확률을 비교하여 분류할 경우 특정범주에 속할 확실성이 분명하지 못한 경우가 흔히 발생된다. 만약 특정 범주에 속할 사후확률이 임계확률보다 유의하게 크지 않는 경우 특정 범주로 분류하기 보다는 유보범주(in-doubt category)로 분류하는 것이 적절하다. 따라서 확실범주와 유보범주로 구분할 수 있는 경계값을 임계확률과 별도로 설정할 필요가 있다. 경계값 수준은 두개 경영성과 범주간 중복 정도에 의하여 결정되고, 중복성이 커질수록 경계값은 임계확률보다 크게 설정하는 것이 타당하다.

경계값 결정은 활용목적과 금융기관의 위험 부담 수준에 따라 정해지는 것이 적절하지만, <표 8>의 사후확률분포로부터 적용 가능한 값을 설정하였다. 선형판별모형에서 파산범주와 정상범주의 사후확률 중앙값이 각각 0.37, 0.63 로 분명하게 구분되기 때문에, 두개 범주사이의 중복성은 크지 않다고 판단된다. 따라서 경계값으로 적용 가능한 값은 임계확률 0.5 에서 크게 벗어나지 않을 것으로 예상되기 때문에, 경계값은 임계확률과 중앙값사이 범위 내에서 설정하는 것이 적절하다고 판단된다.

세가지 경계값 0.53, 0.55, 0.58 를 적용하여 확실범주와 유보범주¹⁵⁾로 구분하여 평가된 결과는 <표 13>-<표 15>와 같다. 분석결과 경계값이 0.53, 0.55, 0.58 로 증가함에 따라 전체자료에서 유보범주에 속한 비율은 8.6%, 15.4%, 23.2% 로 증가하고 있다. 만약 유보범주 비율을 10-15% 정도로 설정한다면 경계값은 0.55 로 결정할 수 있다. 경계값 0.55 로 적용했을 때 유보범주 비율은 전체의 15.4% 이고, 유보범주를 제외한 두개 확실범주의 정분류율은 73%(277/378)로 나타났다.

결론적으로 기술금융을 취급하는 금융기관은 식(5)와 같은 추정된 선형판별(분류)함수를 이용하면, 해당 중소기업이 향후 정상범주에 속할 사후확률을 사전에 예측할 수 있을 것이다. 분류 경계값으로 0.55 혹은 금융기관에서 설정한 임의값을 적용한다면, 예측범주를 확실범주와 유보범주로 구분하여 평가할 수 있을 것이다. 또한 금융기관은 기술금융 대상을 선정할 때 경계값을 정상범주에 속할 사후확률의 최소값으로 설정하여 활용할 수 있을 것이다.

<표 13> 경계값 0.53 을 사용한 삼원분류 결과

관측 범주	예측 범주		예측 합계	유보 범주	전체
	파산	정상			
파산	73	24	97	12	109
정상	85	223	308	30	338
전체	158	247	405	42	447

15) 유보범주는 사후확률이 경계값 미만으로 두개 범주로 확실히 분류할 수 없는 자료의 범주를 의미한다.

〈표 14〉 경계값 0.55 을 사용한 삼원분류 결과

관측 범주	예측 범주		예측 합계	유보 범주	전체
	파산	정상			
파산	67	21	88	21	109
정상	80	210	290	48	338
전체	147	231	378	69	447

〈표 15〉 경계값 0.58 을 사용한 삼원분류 결과

관측 범주	예측 범주		예측 합계	유보 범주	전체
	파산	정상			
파산	62	17	79	30	109
정상	71	193	264	74	338
전체	133	210	343	104	447

VI. 결 론

최근 정부는 '기술혁신 확산 및 전주기적 신기술 산업화'를 촉진하기 위해 기술력평가에 근거한 투자와 융자확대, 기술이전과 기술거래 활성화를 적극적으로 추진하고 있고, 금융기관과 연계하여 담보력이 열악한 중소기업에 대한 기술금융을 작년도부터 시행하고 있다. 그러나 지난 일년동안 기술력평가와 연계한 기술금융은 아직 활성화가 되지 못하고 있는 실정이다. 이러한 이유는 정책금융과의 구조적 문제, 시중은행의 기술금융 취급에 따른 리스크 완화 등 여러 문제가 복합적으로 내재 되어있다.

기술금융인 경우 핵심 기준은 보증·담보 보다는 기술력에 초점이 맞춰지는 것이 취지에 부합되기 때문에, 시중 금융기관의 중소기업용 기술금융의 참여를 유도하기 위해서는 무엇보다도 기술력에 근거한 중소기업 부실 가능성을 사전에 예측할 수 있는 판별(분류)모형이 필요하다. 따라서 본 연구에서는 향후 중소기업의 부실 가능성을 예측하기 위해서 기술력평가 12개 핵심 소항목을 설명변수로 포함한 최적 선형판별(분류)함수를 제시하였다. 분석결과 선형판별모형과 로지스틱판별모형의 분류 설명력은 유사하게 나타났고, 임계확률 설정의 적절성을 비교한 결과 전자가 후자보다 적합한 것으로 나타났다. 기술력평가 12개

핵심 소항목을 포함한 선형판별모형은 기술력평가 전체 항목을 적용한 것과 비교해서 설명력이 거의 유사한 것으로 나타났기 때문에, 부분 소항목 정보를 포함한 판별(분류)함수를 최적 선형판별(분류)함수로 결정하였다. 최적 선형판별함수의 분류 설명력인 정분류율은 70.4% 이고, 분류 예측력은 67.5% 로 나타났다.

본 연구에서 제안된 선형판별(분류)함수는 기술력평가에 근거하여 중소기업 부실 가능성을 사후확률로 사전에 예측할 수 있는 방법이 될 수 있을 것이다. 또한 분류범주를 확실범주와 유보범주로 구분할 수 있는 경계값 설정 근거를 제시하였다. 기술금융을 취급하는 금융기관은 경계값을 이용하여 예측범주를 확실범주와 유보범주로 구분하여 평가할 수 있을 것이고, 경계값은 기술금융 대상을 선정할 때 정상범주에 속할 사후확률의 최소값으로 활용할 수 있을 것이다. 본 연구의 제한점은 개별범주 표본자료의 불충분성과 경영성과 범주를 관측한 기간의 제한성 때문에, 분석 결과의 신뢰성이 충분하다고 판단할 수 없다. 따라서 기술력평가를 받은 중소기업들에 대한 향후 경영성과 정보가 장기에 걸쳐 지속적으로 수집될 수 있다면, 본 연구에서 제안한 분석 논리를 적용하여 보다 높은 신뢰성 있는 결과를 기대할 수 있을 것이다.

참고문헌

- 성용현(2001), 「응용다변량분석」, 제2판, 탐진, pp. 177-202.
- _____ (1997), "중소기업부실 예측을 위한 로지스틱 판별분석", 한신논문집, 제14권, pp. 621-636.
- _____ (1999), "Evaluation predictive ability of classification models with ordered multiple categories", The Korean Communications in Statistics, Vol. 6, No.2, pp. 383-395.
- _____ (2004), "기술력평가에서 사업성수준과 기술성변수간 연관성에 관한 실증연구", 품질경영학회지, 제32권 제3호, pp. 198-215.
- _____ (2006), "기술력평가 자료를 이용한 중소벤처기업 파산예측 판별모형에 관한 연구", 기술혁신학회지, 제9권 2호, pp. 304-324.
- 양동우(2003), "벤처의 기술평가와 경영성과의 관계에 관한 연구", 지식경영연구, 제4권 제1호, pp. 21-34.

- _____ (2005), "기술성지표와 기업성과의 관계비교 분석-초기중소벤처와 성장 중소기업-", 기술혁신학회지, 제8권 제3호, pp. 1175-1198.
- _____ (2006), "초기 중소기업의 기술혁신역량과 기업성과의 관계에 관한 연구", 지식경영연구, 제7권 제1호, pp. 49-63.
- Altman, E. I.(1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", Journal of Finance, pp. 589-609.
- Box, G. E .P(1949), "A general distribution Theory for a class of likelihood criteria", Biometrika, 36, pp. 317-346.
- Brooks, C.(1988), "The robustness of the logistic risk function", Communications in Statistics, Simulation, 17(1), pp. 1-24.
- Clark, C. E., Webster, Hogan and Foster(1997), "Judgemental approach to forecasting bankruptcy", Business Sources Premier, 16(2), pp.14-19.
- Griliches, Zvi(1990), "Patent Statistics as Economic Literature", Journal of Economic Literature", 92, pp. 630-653.
- Platt, H. D. and M. B. Platt(1991), "A Note on the use of industry-relative ratios in bankruptcy prediction", Journal of Banking and Finance, 15(5), pp. 1183-1194.
- Poston, K. M., W. K. Harmon and J. D. Gramlich(1994), "A test of financial ratios as predictors of turnaround versus failure among financially distressed firms", Journal of Applied Business Research, 10(1), pp. 41-51.
- Yang, S. S. and Y. Lee(1987), "Identification of a multivariate outliers", presented at the Annual Meeting of the American Statistical Association.
- Zhang, G. Q., M. Y. Hu, B. E. Patuwo and D. C. Indro(1999), "Artificial neural networks in bankruptcy prediction-General framework and cross-validation analysis", European Journal of Operation Research, 116(1), pp. 16-32.

성용현

성균관대학교 통계학과를 졸업하고 미국 Ohio University에서 경영학석사 및 Texas Tech University에서 경영통계학 박사학위를 취득하였다. 현재 한신대학교 정보통계학과 정교수로 재직 중이다. 주요 관심분야는 다변량분석, 기술가치평가, 실물옵션, 기술경영 등이다. 주요 저서로는 경영통계 정보분석, SAS 경영통계, 회귀분석, 응용다변량분석, 응용로지스틱회귀분석, 실물옵션과 R&D 과제의 성과분석 및 예측 등이 있다. 또한 기술가치평가, 실물옵션, 기술력평가 등에 관한 다수의 논문이 있다.