기술혁신학회지 제9권 2호
2006년 6월 pp.304~324

기술력평가 자료를 이용한 중소벤처기업 파산예측 판별모형에 관한 연구*

A Study on Predicting Bankruptcy Discriminant Model for Small-Sized Venture Firms using Technology Evaluation Data

성 웅 현(Sung, Oong-Hyun)**

국문요약

재무분석가들은 기업의 파산에 양향을 미치는 예측변수를 탐색하기 위해서 상당한 연구가 수행되어 왔다. 그러나 기술지향적 중소벤처기업은 일반적으로 역사적 재무자료가 부족하고, 기술경쟁력 수준에 따라 잠재적인 고성장과 고위험이 존재한다. 본 논문에서는 재무자료 대신에 기술력평가 자료를 이용하여 파산을 예측하기 위해서 파산예측 판별모형을 제안하였고, 모형의 정분류율을 통해서 예측력을 검증하기 위해서 교차타당성방법, 최대사후확률방법 등을 사용하였다. 분석결과 중소벤처기업의 파산예측모형으로 선형판별모형이 로지스틱판별모형보다 적합한 모형이고, 표본자료에 대한 정분류율 추정은 약 69% 이고 정분류율 예측은 약 67% 가 될 것으로 기대된다.

핵심어 : 파산예측, 판별모형, 정분류율, 예측분류율 검증

Abstract

There were considerable researches by finance people trying to find out business ratios as predictors of corporate bankruptcy. However, such financial ratios usually lack theoretical justification to predict bankruptcy for technology-oriented small sized venture firms. This study proposes a bankruptcy predictive discriminant model using technology evaluation data instead of financial data, evaluates the model fit by the correct classification rate, cross-validation method and M-P-P method. The results indicate that linear discriminant model was found to be more appropriate model than the logistic discriminant model and 69% of original grouped data were correctly classified while 67% of future data were expected to be classified correctly.

Key words : bankruptcy, discriminant model, correct classification rate, validation

* 이 논문은 2006년도 한신대학교 학술연구비에 의하여 연구되었음

** 한신대학교 정보과학대학 정보통계학과 정교수 (031) 370-6767 019-289-0636 soh@hs.ac.kr

I. 서 론

기업부실에 대한 정의는 경제적 부실(economic failure), 기술적 지급불능(technical insolvency), 파산(bankruptcy) 등으로 요약되고 있다. 경제적 부실이란 수익이 비용보다 적든지, 실제수익이 기대수익보다 적은 경우를 의미한다. 기술적 지급불능이란 총자산이 총부채를 상회하고 있지만, 만기가 도래한 채무를 상환하지 못하거나 어음이 부도 처리된 경우를 의미한다. 기업파산이란 기업의 총자산이 총부채보다 적게 되어 채권자들이 기업의 파산을 법원에 신청하여 기업의 존속을 종식시키는 경우를 의미한다. 부실기업을 예측하기 위한 통계적 판별모형에 관한 연구는 Atman(1968)에 의하여 처음으로 시도되었다. 이후 여러 학자들에 의하여 기업파산 예측모형(bankruptcy prediction model)은 기업의 재무상태(자본구성, 유동성, 수익성, 안정성), 사업성(경영성과, 성장성, 시장성, 상환력), 거래신뢰도(금융기관 여신거래 실적, 사업안정성), 인적사항(업력, 경영자능력), 신용등급평가 등을 종합적으로 고려한 판별모형(discriminant model)에 의하여 많은 연구가 되어왔다.

통계적 판별모형에 대한 이론적인 효율성 비교는 Efron(1975), Byth(1980), Brooks(1988)과 성웅현(1999)에 의해서 수행되었다. 전자는 두모집단에 대하여 다변량정규분포(multivariate normal distribution)에 근거한 선형판별모형(linear discriminant model)과 로지스틱판별모형(logistic discriminant model)의 효율성 비교하였고, 후자는 두 가지 이상의 모집단에 대하여 선형판별모형, 이차판별모형, 회귀모형, 로짓판별모형(logit discrimination model) 등을 비교 분석하였다. 연구결과 모집단에 대하여 다변량정규분포를 가정하였을 때 선형판별모형과 로짓판별모형의 효율성은 유사한 것으로 나타났고, 다변량정규분포에서 크게 위반된 경우에는 로짓판별모형이 다른 판별모형보다 상대적인 효율성이 높은 것으로 나타났다.

기업의 파산예측모형의 설명력을 높이고 예측오차를 줄이기 위해서는 기업의 재무적 요소와 비재무적 요소 등을 종합적으로 고려하는 것이 일반적이다. 그러나 기술지향적인 중소벤처기업의 특성은 충분한 수익이 발생되기까지에는 상당한 기간이 소요될 것으로 예상되고, 시장진입의 성공과 성장동력을 가속화하기 위해서는 지속적인 연구개발 투자가 병행되어져야만 한다. 또한 중소벤처기업의 업력은 일반적으로 짧기 때문에 연관된 재무적, 비재무적 시계열 정보가 극히 제한적이다. 최근 정부는 기술개발결과의 산업화를 촉진하고, 기술지향적인 중소기업을 지원하기 위하여 기술력평가에 의한 기술담보대출을 활성화하는 기술금융정책을 구축하고 있다. 기술력평가란 기업의 기술경영과 인적사항, 기술성, 시장성, 사업성 등과 관련된 세부항목에 의하여 측정된 결과에 근거해서 내재된 기술력 수준을 평

점 혹은 등급으로 평가하는 것을 의미한다. 국내에서는 개별 법령(기술촉진법, 신기술사업금융지원법,벤처육성특별조치법)에 의하여 신기술 개발 결과의 사업화 촉진 자금지원, 융자기업에 대한 기술보증서 발급, 벤처기업 확인 등의 목적으로 기술력 평가를 기술신용보증기금, 중소기업진흥공단, 한국기술거래소, 한국과학기술정보연구원 등 12개 전문평가기관에서 수행하고 있다.

기술지향적인 중소벤처기업에 대한 기술금융을 활성화하고 동시에 금융 부도위험을 줄이기 위해서 기술금융 주체들은 파산기업을 사전적으로 예측할 수 있는 판별모형을 요구하게 될 것이다. 그러나 기술지향적인 중소벤처기업에서 수집될 수 있는 유용한 정보는 재무적 정보보다는 기술력 전문기관에서 평가된 기술력평가 결과이다. 따라서 중소벤처기업에 대한 기술금융과의 연계성을 분석하기 위해서는 우선 기술력평가 수준과 미래 성과범주에 대한 실증적 연관관계를 분석할 필요가 있다. 그러나 국내 중소벤처기업의 역사가 짧고 연관된 자료 부족으로 두 가지 요인사이이 연관성에 대한 실증적 연구가 매우 미미한 실정이다.

본고에서는 기술력평가를 받은 중소벤처기업 중에서 일정한 기간이 지난 후 사후적인 성과범주로 정상과 파산 등 두 가지 범주로 구분하고, 파산예측 판별모형을 설정하기 위해서 재무정보 대신에 기술력평가 항목을 이용하였다. 판별분석의 핵심 관심사는 중소벤처기업의 파산예측에 영향을 미칠 수 있는 변수로 기술력평가 항목만을 고려했을 때 어느 정도 수준의 설명력과 예측력이 있는지 추정하는 것이고, 추정된 결과를 이용한 기술금융 의사결정에 대한 활용이다. 만약 사전 기술력평가자료를 이용하여 사후 기업의 경영성과(정상, 파산) 범주를 구분할 수 있는 설명력이 높은 판별함수를 설정할 수 있다면, 실무적으로 기술력 담보를 근거로 대출여부를 의사결정하는데 효율적으로 사용되어질 수 있을 것이다. 본고에 포함될 주요 분석내용은 (1) 다변량 분산분석(multivariate analysis of variance)을 이용한 성과범주에 따른 기술력평가 항목 평균벡터에 대한 유의성 차이 검정, (2) 성과범주를 판별하기 위한 선형판별모형과 로지스틱판별모형의 비교분석, (3) 선정된 판별모형의 설명력과 예측력 분석과 활용 등으로 구성되었다.

II. 선행연구 검토와 문제점

기업의 재무정보와 비재무정보를 이용한 기업의 파산예측모형은 수년간 상당한 연구가 이루어졌다. 기업의 파산예측모형에 대한 연구 영역은 기업의 파산에 유의한 영향을 미치는 주요 독립변수의 탐색과 더불어 다양한 판별모형의 예측력에 대한 비교이다. Atman(1968)은 선형판별함수를 이용하여 기업파산에 영향을 미치는 주요 재무비율로 (유동자산-유동부

채)/총자산, 이익잉여금, 경상이익/총자산, 주가/장부가 총부채, 매출액/총자산 등을 선택하였다. Atmam 의 연구의 연장선상에서 Platt 와 Platt(1991), Poston, Harmon 과 Gramlich(1994), Clark, Webster, Hogan 과 Foster(1997), Zhang, Hu, Patuwo 와 Indro(1999)는 미국 여러 산업별로 시계열 재무자료에 선형판별모형, 로짓모형, 신경망모형 등을 적합하여 기업의 파산예측모형을 개발하였다. 성웅현(1997)은 중소기업 파산예측모형으로 로지스틱 판별분석을 적용하여 정상기업과 부실기업을 판별하는데 통계적으로 유의한 변수로 총차입금 대 자기자본비율, 금융기관 여신 거래실적, 사업안정성 등을 선정하였다.

기업의 재무비율에 근거한 선행연구 결과에 파산기업의 재무지표는 정상기업에 비해서 파산시점에 가까이 갈수록 재무지표가 악화되는 경향을 보이고 있고, 이러한 현상은 평균적으로 2년 전부터 시작되는 것으로 나타났다. 그리고 기업의 1년 내 채권 신용등급과 부채비율의 급격한 변화는 기업의 파산가능성을 예측하는데 매우 유용한 정보로 나타났다. 특히 Tsukuda 와 Baba(1994)의 연구결과에 의하면 기업파산에 대한 예측 확률의 타당성은 최근 3년간 재무정보를 이용했을 경우에 높아지는 것으로 나타났다. 선행 연구에 의하면 판별모형의 파산 여부의 예측력은 사용된 주요 변수의 선택과 표본기업의 선택, 파산시점으로 부터 기간 설정, 적용된 판별모형에 따라 서로 다르게 나타났고, 그 예측 설명력의 대부분 범위는 약 65% - 75% 로 나타났다.

최근 지식기반 경제체제에서 기업의 성장과 경쟁력의 원천은 노동·자본에서 지식·정보 우위로 전환되고 있고, 선진국의 경우 기술, 정보 등 무형자산에 대한 투자가 생산설비투자를 상회하고 있는 것이 현실이다. 특히, 기술력이 기업경쟁력의 핵심요소로 대두됨에 따라 기술지향적 중소벤처기업들은 독자적인 기술력을 통하여 경쟁력 우위를 확보하고 시장적 기회를 창출하기 위한 기술투자 등에 우선순위를 두게 되었다. 무형자산(연구개발, 특허, 혁신 등)에 관한 연구개발 투자가 기업가치에 공헌하는 정도에 대한 국외 연구는 Brown 과 Svenson(1998), Aboody 와 Lev(1998), Griliches(1990) 등이 있고, 기술기업에서 연구개발의 비용-효익 관계를 분석한 결과 연구개발 투자와 기업가치 및 성과사이에 유의한 상관관계가 존재하고 그 공헌도가 다른 자산에 비해서 평균적으로 높다는 결론을 도출하였다. 임기철외 10인(2003)은 국내기업의 연구개발 투자와 생산성과의 관계를 분석하기 위해서 R&D Scoreboard 를 사용하였고, 분석결과 대부분의 기업에서 연구개발 투자와 생산성은 양의 상관관계가 있는 것으로 나타났다. 따라서 기술자산의 경쟁력은 미래 기대 현금흐름의 성장과 생산성에 유의한 영향을 미치는 주요 요인으로 판단할 수 있다.

양동우(2003, 2005)는 벤처 기술평가와 경영성과사이의 관계를 검정하기 위해서 표본기업 22개에 대하여 회귀분석한 결과 전반적인 기술평가 수준이 미래 매출액 성장에 유의한

영향을 미치고 있는 것으로 나타났다. 분석결과 2001년도에 기술력평가를 받은 기업 중에서 2005년도 사업성과를 알려진 중소벤처기업에 대하여 그 연관성을 분석하고자 요인분석, 로지스틱회귀분석, 선형회귀분석 등을 수행하여 설명력을 비교한 결과, 그 설명력은 범주별로 61%-84% 인 것으로 나타났다. 성웅현(2004)은 기술력평가 항목 내에서 기술성 변수 정보를 이용하여 사업성 수준 범주(상대적으로 높은 경우와 낮은 경우의 두 가지 범주와 상중하의 세 가지 범주)를 설명하기 위해서 로지스틱 판별모형과 로짓 판별모형을 이용하였다. 기술성 변수의 정보를 이용하여 사업성 범주를 판별하는데 설명력은 세범주인 경우 약 67.7%, 두 범주인 경우에는 95.6% 로 나타났다.

국내외 기업파산 예측을 위한 선행연구 분석결과의 주요 문제점은 세 가지로 요약된다. 첫째, 판별모형에 적합 시킬 다변량 자료의 설정된 가정의 적합성 여부이다. 예를 들면, 선형판별모형을 사용하고자 할 경우에는 개별 범주에 속한 다변량 자료가 다변량 정규분포에서 크게 위반되었는지 여부를 분석 전에 탐색할 필요가 있다. 왜냐하면 변량자료가 다변량 정규분포에서 크게 위반된 다변량 이상값(multivariate outliers)이 다량으로 존재하는 경우에는 판별모형의 신뢰성과 적합성에 문제가 제기될 수 있기 때문이다. 물론 이러한 현상은 다변량 정규분포 가정에서 크게 제한받지 않는 로지스틱 판별모형을 적합 시킬 경우에도 동일하게 적용된다. 둘째, 추정된 판별모형으로부터 개별 기업을 특정 범주로 분류하기 위한 기준이 되는 임계확률값 설정의 적절성 여부이다. 왜냐하면 임계확률값을 임의로 설정할 경우 해당 판별모형의 설명력인 정분류율이 크게 달라질 수 있기 때문이다. 셋째, 추정된 판별모형을 미래 자료에 적용했을 때 기대되는 분류 예측력에 대한 평가이다. 이론적으로 해당 표본자료를 분류에 적용하여 구한 정분류율은 해당 판별모형의 정분류율을 과대평가하는 경향이 있기 때문에, 해당 판별모형의 예측력의 편의(bias)를 추정하여 정분류율을 수정할 수 있는 방법론이 포함되어야 한다. 본고에서는 사전 기술력평가 13개 중항목을 포함한 판별모형의 추정과 설명력 평가 과정에서 이러한 문제들을 고려하여 분석하였다.

Ⅲ. 기술력평가 다변량 분산분석

사전 기술력평가와 사후 두 가지 경영성과(정상과 파산)와의 연관성을 분석하기 위한 표본기업은 중소기업청에서 2001년 당시 기술신용보증기금으로부터 기술력평가를 받은 중소벤처기업 482개를 표본기업으로 선정하고, 2005년도 시점에서 표본기업을 다시 경영성과 범주에 따라 조사한 결과 수익이 발생되고 있는 정상기업 361개와 파산상태에 있는 121개 기업으로 구분되었다. 본 절에서는 두 범주(정상과 파산)에 대한 기술력평가 13개 중항목의

평균벡터(mean vector)가 유의한 차이가 있는지 여부를 다변량 분산분석(multivariate analysis of variance)을 통하여 검정하고, 또한 두 범주간 평균벡터의 유의한 차이에 영향을 미친 중항목을 탐색하였다.

1. 범주별 다변량 이상값 탐색

두 가지 범주에 대한 다변량 분산분석을 수행하기 전에 다변량 자료에 대한 다변량 이상값을 탐색할 필요가 있다. 왜냐하면 다변량 분산분석의 기본적인 가정은 두 가지 범주에 대한 다변량자료는 평균벡터가 서로 다르고 동일한 공분산행렬을 가진 다변량 정규분포로부터 추출된 표본이라고 가정하기 때문이다. 기술력평가 13개 중항목은 (1) 기술경영능력(T1), (2) 인적자원 및 연구개발능력(T2), (3) 기술개발환경(T3), (4) 기술우수성(T4), (5) 기술경쟁성(T5), (6) 기술권리성(T6), (7) 시장특성(T7), (8) 시장환경(T8), (9) 산업환경성(T9), (10) 생산기반(T10), (11) 생산성(T11), (12) 수익성(T12), (13) 마케팅 전략성(T13)으로 구성되었다. 다변량 이상값을 탐색하기 위하여 Yang 과 Lee(1987)가 제안한 $F-$검정 방법을 사용하였다. $p=13$ 개 중항목에 대한 $j-$번째 자료벡터 $\boldsymbol{T}_j = [T_{1j}, T_{2j}, \cdots, T_{13j}]'$ 가 평균벡터가 $\boldsymbol{\mu}$ 이고 공분산행렬(covariance matrix)이 $\boldsymbol{\Sigma}$ 인 다변량 정규분포 $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$[1]에서 추출된 표본이라고 하자. 이때 $\boldsymbol{T}_j$ 와 $\boldsymbol{\mu}$ 사이의 Mahalanobis 거리제곱은 $\delta_j^2 = (\boldsymbol{T}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{T}_j - \boldsymbol{\mu})$ 이고, 표본자료 n 인 자료벡터 $\boldsymbol{T}_j's$ 에서 구한 표본평균벡터 $\overline{\boldsymbol{T}}$ 와 표본공분산행렬 $\boldsymbol{S}$ 로 표시한 표본 Mahalanobis 거리제곱은 $d_j^2 = (\boldsymbol{T}_j - \overline{\boldsymbol{T}})' \boldsymbol{S}^{-1} (\boldsymbol{T}_j - \overline{\boldsymbol{T}})$ 이 된다. 이때 다변량 이상값을 탐색하기위한 $F-$통계량은 식(1)과 같다.

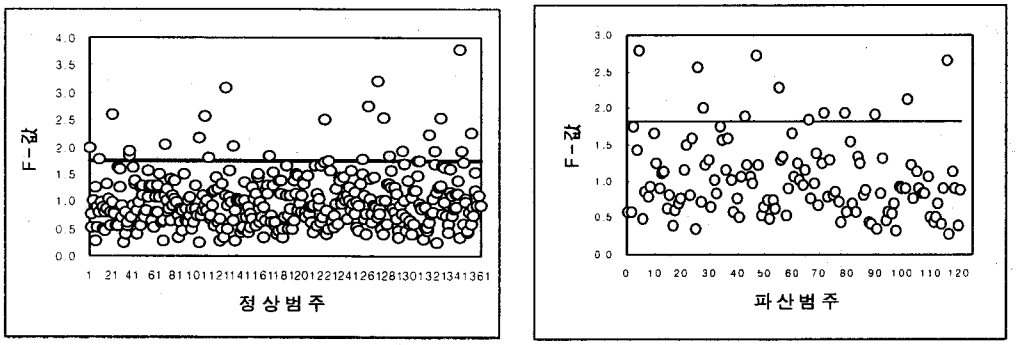
$$F_j = \frac{n-p-1}{p} \times \left[\frac{1}{1 - n\, d_j^2 / (n-1)^2} - 1 \right] \tag{1}$$

식(1)에서 F_j 통계량은 분자의 자유도가 p 이고 분모의 자유도가 $n-p-1$ 인

1) $p \times 1$ 확률벡터가 $\boldsymbol{T} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ 인 $p-$변량 정규분포에 따를 때 $p-$변량 정규확률밀도함수는 $f(\boldsymbol{t}) = \frac{1}{(2\pi)^{p/2} |\boldsymbol{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2} (\boldsymbol{T} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{T} - \boldsymbol{\mu}) \right\}$ 이 된다.

$F(p, n-p-1)$ 분포에 근사하게 된다. 유의수준 0.05 에서 다변량 이상값 검정은 검정통계량 F_j 가 기각치 $F(0.05, p, n-p-1)$ 보다 클 때 자료벡터 T_j 가 다변량 이상값이라고 판정한다. 다변량 이상값을 탐색하기 위해서 수평축을 자료 순서로 하고 수직축을 $F-$통계량으로 플롯한 산점도를 SAS 에서 IML 프로그램을 작성하여 〈그림 1〉과 같이 구하였다. 〈그림 1〉의 왼쪽 $F-$통계량 산점도에서 정상범주인 경우 유의수준 0.05 에서 기각치 $F(0.05, 13, 347)$=1.748 보다 큰 자료는 23개로 나타났고, 파산범주인 경우 기각치 $F(0.05, 13, 107)$=1.813 보다 큰 자료는 12개로 나타났다. 표본자료 482개 자료 중에서 $F-$통계량 검정결과 다변량 이상값으로 탐색된 자료는 35개로 전체자료의 7.3%로 나타났다. 기술력평가 자료에서 다변량 이상값이 발생된 원인을 여러 가지가 있겠지만 통계적으로 주요 원인을 요약하면 (1) 특정 자료가 대부분의 다른 자료와 비교해서 중항목 분포의 속성이 크게 다른 경우, (2) 특정 자료의 중항목사이의 점수 패턴이 전체 자료 중항목간 상관관계 패턴과 큰 차이가 있는 경우, (3) 전문가의 항목별 평가상의 일관성에 대한 오류도 고려될 수 있을 것이다. 따라서 다변량 분산분석에 사용된 자료는 35개 이상값 자료를 제외한 정상범주 338개 파산범주 109개 등 총 447개 이다.

〈그림 1〉 정상범주와 파산범주에서 $F-$검정을 이용한 다변량 이상값 탐색



2. 다변량 분산분석 결과

다변량 분산분석에서는 파산범주와 정성범주에 속한 13개 중항목은 평균벡터를 μ_o, μ_1 라고 하고 동일한 공분산행렬 Σ 을 가진 다변량 정규분포에 따른다고 가정하고, 두 가지 범주사이의 평균벡터에 유의한 차이가 있는지 여부를 검정한다. 우선 두 가지 범주의

공분산행렬이 동일한지 여부를 Box(1949)가 제안한 다변량 Box's M-검정 결과 검정통계량[2]은 $M^* = 82.303$ 이고 연관된 $p-$값이 0.731 이므로 두 범주에서 공분산행렬이 동일하다는 귀무가설을 채택할 충분한 근거가 있다고 판단된다. 다변량 분산분석에서 평균벡터의 유의한 차이를 검정하는 검정통계량 Wilk's Lambda Λ, Pillai's Trace V, Hotelling-Lawley U, Roy's Greatest Root θ 에 대한 $F-$근사 검정결과는 아래 〈표 1〉과 같다.

〈표 1〉 두범주에 대한 평균벡터 동일성 다변량 분산분석 검정 결과

검정통계량	검정통계량값	F-변환 값	분자 자유도	분모 자유도	$p-$값
Wilk's Lambda Λ	0.87078455	4.94	13	433	〈.0001
Pillai's Trace V	0.12921545	4.94	13	433	〈.0001
Hotelling-Lawley Trace U	0.14838969	4.94	13	433	〈.0001
Roy's Greatest Root θ	0.14838969	4.94	13	433	〈.0001

네 가지 검정통계량의 $F-$검정 결과 $p-$값이 0.0001 미만으로 0에 가까운 매우 작은 값을 갖는 것으로 나타났다. 따라서 파산과 정상 두 범주 사이의 평균벡터가 동일하다는 귀무가설 $H_o: \mu_o = \mu_1$ 을 강하게 기각할 수 있는 충분한 근거가 있기 때문에, 기술력평가 13개 중항목의 평균벡터가 정상범주와 파산범주에 따라 매우 유의한 차이가 있다고 판단된다. 다변량분산분석 결과 기술력평가 13개 중항목에 대한 종합적인 평균정보가 정상과 파산범주에서 유의한 차이가 있는 것으로 나타났기 때문에, 개별 중항목이 두 범주사이에 유의한 평균차이가 있는지 여부를 검정할 필요가 있다. 그러나 두 범주에서 기술력평가 개별 중항목에 대한 일변량 분산분석이나 두표본 $t-$검정을 할 경우 검정에 문제가 제기된다. 성웅현(2000)에 의하면 13개 중항목에 대한 두 범주간 다변량 검정을 사용하지 않고 일변량 개별 검정을 수행할 경우 세 가지 문제를 제기 하였다, 첫째, 13개 중항목에 대하여 다변량 검정을 수행할 경우 설정된 유의수준 α (제1종 오류)가 유지되지만, 일변량 개별 검정을 수행할 경우 유의수준이 팽창하게 된다. 둘째, 일변량 개별 검정인 경우 기술력

2) Box's M-검정에서 $v_i = n_i - 1$, $g=2$ 로 했을 때 검정통계량은 $M^* = -2(1-c)\ln M$ 이 되고, 검정통계량의 분포는 $M^* \approx \chi^2[0.5(g-1)p(p+1)]$ 에 근사한다. 여기서
$M = (|S_o|^{v_o/2}\ |S_1|^{v_1/2}) / |S_p|^{\Sigma v_i/2}$,
$c = [\Sigma_i 1/v_i - 1/\Sigma_i v_i]\ [(2p^2+3p-1)/(6(p+1)(g-1))]$ 이다.

평가 13개 중항목사이의 상관구조(correlation structure)가 무시되지만, 다변량 검정인 경우는 중항목사이의 공분산행렬을 검정에 이용한다. 셋째, 다변량 검정의 검정력(power of test)이 13개 일변량 개별검정의 검정력에 비해서 높다는 것이다.

이러한 일변량 문제점을 개선하기 위해서 다변량 검정결과 귀무가설 H_o: $\mu_o = \mu_1$ 이 강하게 기각되면, 두 범주에 대한 기술력평가 중항목 평균벡터 차이를 최대한 유의하게 구별할 수 있는 최대유의선형결합(maximum significant linear combination)을 구하게 된다. 두 가지 경영성과 범주를 구분하기 위한 최대유의선형결합을 추정하기 위해서 기술력 자료벡터 T_{ij} $i=0, 1,$ $j=1,2,\cdots,n_i$ 의 선형결합 $z_{ij}=a' T_{ij}$ 을 설정하여, 범주간 평균벡터의 차이를 최대로 하는 계수벡터 $a = [a_1, a_2, \cdots, a_{13}]'$ 를 구한다. 표본 범주간 행렬이 $H=\sum_{i=0}^{1} n_i (\overline{T_i} - \overline{T})$ $(\overline{T_i} - \overline{T})'$ 이고, 표본 범주내 행렬이 $E = \sum_{i=0}^{1} \sum_{j=1}^{n_i} (T_{ij} - \overline{T_i})(T_{ij} - \overline{T_i})'$ 일 때, $E^{-1}H$ 를 최대로 하는 추정된 계수벡터 $\hat{a}$ 는 첫 번째 고유값(eigen value) 0.148 에 대응되는 고유벡터(eigen vector)로 〈표 2〉와 같이 구해졌다.

〈표 2〉 추정된 최대유의선형결합 계수

기술력평가 대항목	기술력평가 중항목	표준화 계수
기술경영과 인적자원	기술경영능력(T1)	.469
	인적자원 및 연구개발 능력(T2)	-.063
	기술개발환경(T3)	-.005
기술성	기술의 우수성(T4)	.285
	기술의 경쟁성(T5)	.045
	기술의 권리성(T6)	-.067
시장성	시장특성(T7)	.130
	시장환경(T8)	-.063
	산업환경성(T9)	.059
사업성	생산기반(T10)	.312
	생산성(T11)	.531
	수익성(T12)	-.164
	마케팅전략(T13)	.155

여기서 구한 계수는 표준화된 값이기 때문에 절대값이 클수록 두 가지 범주를 구분하는데 상대적으로 기여도가 높다고 판단할 수 있다. 따라서 상대적으로 기여도가 가장 높은 중항목을 정리하면 생산성(T11), 기술경영능력(T1), 생산기반(T10), 기술의 우수성(T4)으로 구해졌다.

Ⅳ. 기술력평가 판별분석

기술력평가 판별분석의 두 가지 주요 목적은 다음과 같다. 첫 번째 목적은 두 가지 경영성과 범주(정상과 파산)의 차이를 설명하기 위해서 선형 혹은 비선형 판별함수로 하는 것이다. 두 번째 목적은 추정된 판별함수를 이용하여 임의의 기술력평가 자료에 대해서 특정 범주에 속할 사후확률(posterior probability)을 추정하여 사후확률이 큰 범주로 분류하는 것이다. 본장에서는 두 가지 범주의 변동을 설명하기 위한 적절한 판별모형을 결정하기 위해서 선형판별모형과 로지스틱판별모형에 대하여 비교 분석하였다.

1. 기술력평가 판별모형의 추정과 검정

정상과 파산 범주에 속한 기술력평가 13개 중항목의 표본평균벡터가 $\overline{T_1}$, $\overline{T_o}$ 이고 표본합동공분산행렬[3](sample pooled covariance matrix)이 S_p 일 때 표본선형분류함수(sample linear classification function) L_i, $i=0,1$ 은 식(2)와 같이 구해지고, 선형판별함수는 $L=L_1-L_o$ 식(3)과 같다. 그리고 두 가지 범주에 대한 자료에서 추정된 선형분류함수와 선형판별함수의 계수는 〈표 4〉와 같다.

$$L_i = \overline{T_i}' S_p^{-1} T - 0.5\ \overline{T_i}' S_p^{-1} \overline{T_i}, \quad i=0,1. \tag{2}$$

$$L = (\overline{T_1} - \overline{T_o})' S_p^{-1} T - 0.5 (\overline{T_1} - \overline{T_o})' S_p^{-1} (\overline{T_1} + \overline{T_o}) \tag{3}$$

3) 파산과 정상 관측범주에 속한 13개 중항목자료의 표본공분산행렬을 각각 S_o, S_1 로 표시하면, 두개 표본공분산행렬의 표본합동공분산행렬은 $S_p=[(n_o-1)S_o+(n_1-1)S_1]/(n_o+n_1-2)$이 된다.

〈표 3〉 추정된 선형분류함수와 판별함수

대항목	중항목	선형분류와 판별함수		
		파산(L_o)	정상(L_1)	L
		-78.448	-88.076	-9.628
기술경영과 인적자원	기술경영능력(T1)	1.305	1.455	0.150
	인적자원 및 연구개발 능력(T2)	1.118	1.089	-0.029
	기술개발환경(T3)	0.118	0.119	0.001
기술성	기술의 우수성(T4)	1.799	1.901	0.102
	기술의 경쟁성(T5)	-0.268	-0.251	0.017
	기술의 권리성(T6)	0.480	0.453	-0.027
시장성	시장특성(T7)	-0.147	-0.073	0.074
	시장환경(T8)	2.112	2.083	-0.029
	산업환경성(T9)	0.674	0.715	0.041
사업성	생산기반(T10)	1.347	1.459	0.112
	생산성(T11)	0.339	0.562	0.223
	수익성(T12)	0.465	0.407	-0.058
	마케팅전략(T13)	3.215	3.347	0.132

선형판별모형에서 두 가지 범주의 변동을 설명하는데 유의한 기술력평가 중항목을 탐색하기 위해서 스텝와이즈 선택법(stepwise selection)에서 $F-$검정과 Wilk's Λ 우도비 검정 결과는 〈표 4〉와 같다. 스텝와이즈 선택법에 의해서 상대적으로 영향이 큰 중항목을 유의수준 0.05 에서 검정한 결과 기술경영능력(T1), 생산성(T11), 기술의 우수성(T4), 생산기반(T10) 등 네 가지 중항목으로 나타났다.

〈표 4〉 스텝와이즈에서 유의하게 선택된 중항목 결과

단계	변수선택	$F-$값	$p-$값	Wilk's Λ	$p-$값
1	T1	26.12	〈 0.0001	0.9446	〈 0.0001
2	T11	18.57	〈 0.0001	0.9066	〈 0.0001
3	T4	10.11	0.0016	0.8864	〈 0.0001
4	T10	4.03	0.0454	0.8784	〈 0.0001

앞에서 설명한 선형분류함수는 13개 중항목들이 다변량 정규분포에 따른다고 가정하고 추정된 것이다. Byth 와 McLachlan(1980)의 이론적 근거와 성웅현(2004)에 의하면 기술력 평가 중항목에 대한 자료 분포를 다변량 정규분포에 따른다고 가정해야할 특별한 이유가 없거나, 가정에서 크게 위반되는 경우에는 로지스틱판별모형이 사용될 수 있다. 기술력평가 중항목에 근거하여 두 가지 경영성과 범주를 예측할 수 있는 로지스틱 판별모형에서 종속변수 Y 는 두개 범주를 표시하는 0(파산) 과 1(정상) 로 표시되고, 13개 중항목을 사용한 로지스틱판별모형은 식(4)와 같이 표시된다. 로지스틱 판별모형을 자료에 적합시켜 Newton-Raphson 방법으로 최대우도 추정된 결과는 〈표 5〉와 같다. 로지스틱 판별분석결과 경영성과 범주에 상대적으로 영향이 큰 중항목을 유의수준 0.05 에서 검정한 결과 기술경영능력(T1), 기술의 우수성(T4), 생산기반(T10), 생산성(T11) 등 네 가지 중항목으로 나타났다. 이러한 결과는 스텝와이즈 선택법에서 구한 〈표 4〉의 결과와 동일하게 나타났다.

$$Y = \frac{\exp(\beta_o + \beta_1 T_1 + \beta_2 T_2 + \cdots + \beta_{13} T_{13})}{1 + \exp(\beta_o + \beta_1 T_1 + \beta_2 T_2 + \cdots + \beta_{13} T_{13})} + \varepsilon \qquad (4)$$

〈표 5〉 추정된 로지스틱 판별함수 계수와 검정결과

대항목	중항목	계수	$p-$값
		-8.505	〈.0001
기술경영과 인적자원	기술경영능력(T1)	0.136	0.0020
	인적자원 및 연구개발 능력(T2)	-0.035	0.5719
	기술개발환경(T3)	0.006	0.9114
기술성	기술의 우수성(T4)	0.095	0.1058
	기술의 경쟁성(T5)	0.017	0.8554
	기술의 권리성(T6)	-0.033	0.6615
시장성	시장특성(T7)	0.094	0.3146
	시장환경(T8)	-0.014	0.9029
	산업환경성(T9)	0.039	0.6307
사업성	생산기반(T10)	0.109	0.1120
	생산성(T11)	0.237	0.0002
	수익성(T12)	-0.065	0.2069
	마케팅전략(T13)	0.160	0.1794

2. 사후확률에 근거한 임계확률값 비교

기술력평가 13개 중항목을 이용한 선형판별함수를 이용하여 두 가지 범주로 분류하고자 할 경우에는 개별 범주에 속할 예측가능성을 사후확률로 평가한 후 상대적으로 큰 사후확률을 가진 범주로 분류하는 절차를 따르게 된다. 13개 중항목에 대한 기술력 평점이 T^*로 주어졌을 때 개별 범주에 속할 사후확률은 식(5)와 같이 구해진다. 또한 로지스틱판별함수를 이용하여 분류될 범주를 예측할 때에도 개별 범주에 속할 가능성을 확률로 평가한 후 상대적으로 큰 확률을 가진 범주로 분류하는 절차를 식(6)을 이용한다.

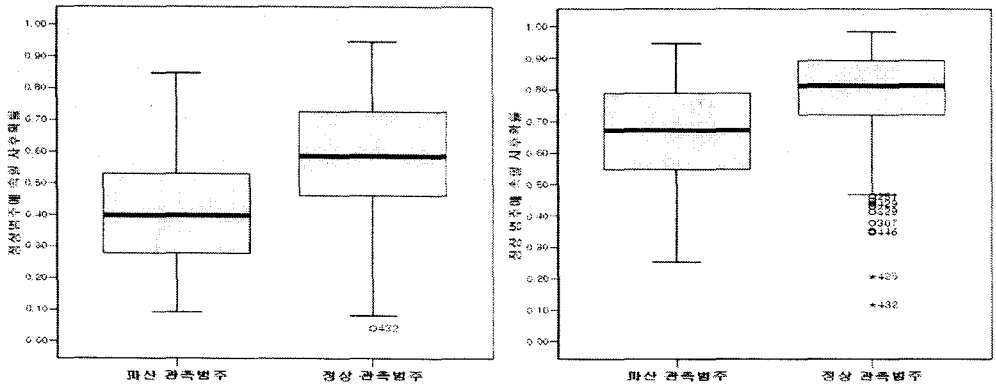
$$P(\text{정상범주}|\,T^*) = \frac{\exp(L_1)}{\exp(L_o)+\exp(L_1)} = \frac{1}{1+\exp(L)}$$
$$= \frac{1}{1+\exp(-9.628+0.150T1-0.029T2+\cdots+0.132T13)} \qquad (5)$$
$$P(\text{파산범주}|\,T^*) = 1 - P(\text{정상범주}|\,T^*)$$

$$P(\text{정상범주}|\,T_o) = \frac{\exp(-8.505+0.136T1-0.035T2+\cdots+0.160T13)}{1+\exp(-8.505+0.136T1-0.035T2+\cdots+0.160T13)} \qquad (6)$$

이때 두 가지 범주를 구분할 때 사용될 임계확률값(critical value of probability)은 특정 범주에 속할 사전확률이 알려져 있지 않은 경우 일반적으로 0.5 가 사용된다. 두 가지 관측범주에 속한 자료가 정상 범주에 속할 사후확률에 대한 요약통계는 〈표 6〉과 같다. 선형판별모형인 경우 사후확률분포는 대칭에서 약간 벗어나 있기 때문에 임계확률값으로 평균보다는 중앙값을 적용하는 것이 적절하다. 요약통계 분석결과 관측된 파산기업이 정상범주에 속할 사후확률의 중앙값은 0.398 로 관측된 정상기업이 정상범주에 속할 사후확률의 중앙값인 0.585 와 분명한 차이가 있음을 알 수 있다. 두 가지 범주의 사후확률 중앙값을 평균하면 0.49 [(0.398 +0.585)/2] 로 약 0.5 근처에 있기 때문에, 사전 기술력평가에 의해서 정상 혹은 파산범주로 분류할 적절한 임계확률값은 0.5 라고 판단된다. 그러나 로지스틱 판별모형인 경우 정상범주에 속할 사후확률의 자료분포가 왼쪽으로 크게 치우쳐 있기 때문에, 적절한 임계확률값을 설정하는 것이 쉽지 않다. 요약통계에 의하면 두개 관측범주(파산, 정상)에서 구한 정상범주에 속할 사후확률이 비대칭이기 때문에, 두 가지 범주를 판별할 수 있는 적절한 임계확률값은 약 0.74 (0.673+ 0.813 = 0.743)가 된다.

〈그림 2〉 두 가지 판별모형에 대한 사후확률 자료 상자그림

(a) 선형판별모형 (b) 로지스틱 판별모형



〈표 6〉 두 가지 판별모형에 대한 개별 범주별 사후확률 요약통계

구분	선형판별모형 사후확률		로지스틱 판별모형 사후확률	
	파산범주	정상범주	파산범주	정상범주
평균	0.415	0.586	0.652	0.790
중앙값	0.398	0.585	0.673	0.813
표준편차	0.182	0.188	0.173	0.140
일사분위수	0.277	0.460	0.548	0.722
삼사분위수	0.530	0.726	0.790	0.893
최대	0.847	0.946	0.947	0.984
최소	0.092	0.040	0.253	0.118
왜도	0.266	-0.345	-0.489	-1.295

3. 두 가지 판별모형의 분류 설명력 비교

선형판별함수와 로지스틱판별함수의 설명력을 비교하기 위해서는 전체 범주와 개별 범주에 대한 임계확률값의 설정의 적절성과 분류 설명력을 비교 검토하여야 한다. 선형판별함수에 의한 분류에서 임계확률값을 0.5 으로 설정했을 때 관측범주와 예측범주로 구분한 이원분류표는 〈표 7〉(a) 와 같다. 이원분류표에 의하면 파산범주로 관측된 109개 중 75개가 정

분류되었고 34개가 정상범주로 오분류되어, 파산 관측범주에 대한 정분류율은 약 68.8% 이다. 그리고 정상범주로 관측된 338개 중 235개가 정분류되었고 103개가 파산범주로 오분류되어, 정상 관측범주에 대한 정분류율은 약 69.5% 로 나타났다. 개별 범주에 속할 사전확률이 동일하다고 가정하면 전체 관측범주에 대한 정분류율[4]은 약 69.2% 가 된다. 로지스틱판별함수에 의한 분류에서 두 가지 임계확률값 0.5 와 0.74(이원분류표에서 ( ) 로 표시)를 동시에 적용하여 구한 이원분류표는 〈표 8〉(b) 와 같다. 임계확률값이 0.5 인 경우 파산 관측범주에 대한 정분류율은 약 19.3% 로 매우 낮게 나타난 반면에 정상범주에 대한 정분류율은 약 95.3% 로 매우 높게 나타났고, 전체 관측범주에 대한 정분류율은 약 57.3% 이 된다. 임계확률값을 0.74 로 다시 적용하면 파산 관측범주에 대한 정분류율은 약 65.1% 이고, 정상 관측범주에 대한 정분류율은 약 72.2% 이다. 따라서 전체자료에 대한 정분류율은 약 68.7% 이다.

〈표 7〉 두 가지 판별모형에 분류결과 비교

(a) 선형판별모형

관측 범주	예측 범주		전체
	파산	정상	
파산	75	34	109
정상	103	235	338

(b) 로지스틱 판별모형

관측 범주	예측 범주		전체
	파산	정상	
파산	21(71)	88(38)	109
정상	16(94)	322(244)	338

상기 분류결과를 종합하면 선형판별모형인 경우 전체 관측범주에 대한 정분류율은 약 69.2 % 로 로지스틱판별함수의 약 68.7 % 보다 약 0.5% 높게 나타났으나, 두 가지 판별모형간 분류력에는 유의한 차이가 있다고 판단할 수 없다. 여기서 유의할 점은 선형판별함수인 경우 파산 관측범주와 정상 관측범주의 정분류율이 거의 유사한 68.8%, 69.5% 로 나타났으나, 로지스틱판별함수(임계확률값이 0.74인 경우)인 경우 정상 관측범주의 정분류율이 72.2% 로 파산 관측범주의 정분류율 65.1% 보다 약 7.1% 의 큰 차이가 있는 것으로 나타났다. 판별모형의 적절성과 안정성은 전체 관측범주에 대한 정분류율이 일정 이상인 모형 중에서 개별 범주의 정분류율이 서로 비슷하게 구해진 모형이 적절할 것으로 판단된다. 이

4) 전체 관측범주에 대한 정분류율을 구할 때 두 가지 방법이 적용할 수 있다. 첫 번째 방법은 전체표본자료에 대한 정분류된 도수의 비율인 (75+235)/447 = 69.4% 이다. 이 경우는 개별 범주에 속한 자료의 수에 따라 정분류율이 왜곡될 가능성이 높다. 두 번째 방법은 개별 범주의 정분류율을 사전확률로 가중평균하여 사용한다. 이 경우 특정 범주에 속할 사전확률이 미지이므로 일반적으로 0.5를 사용하여 구하면 (68.8+69.5)/2 = 69.2% 가 된다.

때 로지스틱판별모형에서 제기될 수 있는 문제는 분류 기준값으로 사용될 임계확률값의 적절성 문제이다. 왜냐하면 정분률율을 개선하기 위해서 임계확률값을 임의로 사용할 가능성을 배제할 수 없기 때문이다. 따라서 해당 기술력평가 중항목자료를 이용하여 미래 중소벤처기업의 파산여부 가능성을 설명할 수 있는 판별모형은 로지스틱판별모형보다 선형판별모형이 더 적합하다고 판단된다.

V. 분류 예측력 평가와 활용

1. 선형판별모형의 분류 예측력 평가

선형판별함수를 이용한 분류결과 파산 관측범주의 정분류율은 약 68.8%, 정상 관측범주의 정분류율은 약 69.5% , 전체 관측범주에 대한 정분류율은 약 69.2% 로 나타났다. 여기서 핵심 관심사는 추정된 선형판별함수를 미래 기술력평가 자료에 적용할 때 기대되는 분류 예측력 수준이다. 앞에서 구한 정분률율을 단순정분류율(apparent correct classification rate)이라고 하며, 이론적으로 해당 표본자료에서 구한 단순정분류율은 실제정분류율(actual correct classification rate)을 과대 추정하는 경향이 있다. 즉, 단순정분류율은 실제정분류율과 비교해서 양의 편의(positive bias)를 갖게 된다. 왜냐하면 판별함수를 추정하는데 사용된 표본자료가 정분류율을 평가할 때에도 다시 사용되었기 때문이다. 실제정분류율을 추정하기 위한 방법은 표본분할방법(split sample method), 교차타당성방법(cross-validation method), 최대사후확률방법(maximum-posterior-probability method) 등이 사용된다.

표본분할방법은 각 집단에 속하는 표본자료가 충분히 클 때 사용될 수 있는 방법으로 전체표본을 판별함수를 추정하는데 사용될 분석표본(analysis sample)과 정분률율을 평가할 때 사용될 검증표본(holdout sample)으로 구분하여 예측력을 비교 평가한다. 그러나 표본분할방법을 사용할 경우 두 가지 표본의 비율에 대한 적절성 문제가 제기되기 때문에, 본 분석에서 파산범주에 속한 자료가 크지 않기 때문에 표본분할방법으로 편의를 추정할 경우 신뢰성문제가 제기될 수 있다. 표본분할방법을 사용할 경우 파산범주로 관측된 자료가 충분히 크지 않기 때문에 본 분석에서는 적절한 방법이 아니라고 판단된다. 교차타당성방법은 자료를 하나씩 제외하고 나머지 자료로 판별함수를 추정한 후 제외된 하나의 자료의 분류를 결정하는 절차로 모든 자료에 반복적으로 적용하여 예측력을 평가한다. 최대사후확률방법(M-P-P)은 특정 범주로 분류된 모든 자료의 사후확률의 평균을 이용한다. 성웅현(1999)은 정분류율의 편의를 추정할 때 표본분할방법의 편의 변동이 다른 방법에 비하여 상대적

으로 크기 때문에 교차타당성방법과 과 붓스트렙(bootstrap) 방법이 상대적으로 안정된 방법으로 평가하였다. Hora 와 Wilcox(1982)는 선형판별모형의 분류 예측력을 여러 방법에 대하여 비교한 결과 교차타당성방법과 최대사후확률방법을 결합하여 추정할 때 그 정확도가 가장 높다고 하였다.

따라서 본 절에서는 선형판별모형의 분류 예측력을 평가하기 위해서 교차타당성방법, M-P-P 방법, 교차타당성과 M-P-P 를 결합한 방법 등을 사용하여 실제정분류율을 추정한 결과는 아래 〈표 8〉과 같다. 분석결과 실제정분류율의 추정값은 파산범주인 경우 62.4% - 66.7%, 정상범주인 경우 64.8% - 67.1%, 전체범주에 대해서는 64.8% - 67.1% 로 나타났다. 가장 신뢰성이 높은 교차타당성/M-P-P 방법을 이용하여 분석한 결과 파산범주인 경우 분류 예측력은 약 63,6%, 정상범주인 경우 분류 예측력은 약 70.1%, 전체범주에 대하여는 분류 예측력은 약 66.9% 인 것으로 나타났다. 따라서 전체 범주에 대한 정분류율의 편의는 세 가지 방법에 따라 2.3% - 4.4% 범위로 구해졌다. 분석결과를 종합하면 추정된 선형판별모형의 미래 표본자료에 대한 예측력 수준은 약67% 정도인 것으로 기대할 수 있다.

〈표 8〉 여러 가지 방법에 의한 실제 정분류율 추정 결과

구분	기존 방법	교차타당성	M-P-P	교차타당성/M-P-P
파산범주	68.8%	62.4%	66.7%	63.6%
정상범주	69.5%	67.2%	67.4%	70.1%
전체	69.2%	64.8%	67.1%	66.9%

2. 선형판별모형의 파산 예측모형 활용

추정된 선형판별함수를 이용하여 두 가지 범주(정상, 파산)로 분류할 때 유의할 사항은 개별 범주에 속할 가능성에 관한 확실성이다. 만약 특정 기술력평가 자료에 서 두 가지 범주로 분류될 사후확률이 유사한 경우에는 단순히 사후확률의 대소 기준(사후확률 0.5)에 의하여 분류할 경우에는 의사결정의 확실성에 문제가 제기 될 수 있다. 이러한 문제를 개선하기 위해서는 개별 범주에 분류될 가능한 사후확률의 최소값인 경계값(threshold value)을 지정할 필요가 있다. 만약 개별 범주에 속할 사후확률이 경계값 t 보다 작을 경우에는 두 가지 범주와 별도로 유보범주(in-doubt group)로 구분하여 분류할 필요가 있다. 경계값의 결정은 두 가지 범주에 대한 중복성의 정도에 따라 달라질 수 있다. 만약 두 가지 범주간 중복성이 작을 경우에는 경계값을 0.5 정도로 설정하는 것이 적합하지만, 중복성

이 증가함에 따라 경계값이 0.5 보다 크게 설정할 필요가 있다. 〈그림 2〉(a) 에 의하면 선형판별함수인 경우 경계값을 0.55 로 설정하는 것이 적절할 것으로 판단되고, 사후확률이 0.55 보다 작을 경우에 유보범주를 포함하여 세 가지 범주로 구분한 이원분류표는 〈표 9〉와 같다.

〈표 9〉 경계값을 0.55 인 경우 선형판별모형 분류결과

관측 범주	예측 범주			전체
	파산	정상	유보	
파산	65	26	18	109
정상	79	200	59	338

경계값이 0.55 인 경우 파산 관측범주 109개 가운데 유보범주로 분류된 것은 18개로 약 16.5%(18/109)로 나타났다. 정상 관측범주 338개 가운데 유보범주로 분류된 것은 59개로 약 17.5%(59/338)로 나타났다. 그리고 세 가지 범주로 분류를 설정했을 때 유보범주이외에 다른 범주로 잘못 분류된 경우를 심각한 오분류율(serious misclassification rate)라고 할 때 파산 관측범주의 심각한 오분류율은 약 23.9% (26/109)가 되고, 정상 관측범주에 대한 심각한 오분류율은 약 23.4%(79/338) 인 것으로 나타났다. 따라서 추정된 선형판별함수에서 구한 사후확률을 이용하여 특정 범주에 속할 가능성을 평가할 경우 사후확률의 경계값 설정은 사후확률 자료의 분포와 전문가적 입장에서 설정하여 평가하는 것이 적절할 것으로 판단된다.

VI. 결론 및 한계

금융기관은 담보대출에 대한 위험을 최소화하려고 한다. 다시 말해서 금융기관이 감당할 만한 수준으로 리스크가 헷지(hedge)되지 않은 금융상품은 금융기관에 의해 수용될 수 없다는 의미이다. 금융기관 대출은 일반적으로 담보를 요구하는 것은 바로 이런 리스크 헷지의 일환이라고 할 수 있다. 국내외 금융시장에서 대출시 외부 신용평가회사의 신용평가등급(credit rating)의 정보가 필수적으로 활용되고 있다. 왜냐하면 금융기관들은 외부 신용평가회사가 제시하는 신용등급이 오랜 기간에 걸쳐 안정적인 부도율(default rate)과 손실률(loss rate)에 대한 충분한 통계가 확보되어 있기 때문에 신용등급을 신뢰하기 때문이다. 그러나 금융기관들은 투자 위험성이 높고 단기간에 투자자금을 회수하기 어려운 새로운 기술

의 속성에 대한 이해가 충분하지 못하여 중소벤처기업에 대하여도 일반금융과 같은 관행을 적용하고 있는 것이 현실이다.

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 중소벤처기업에 대한 기술금융 기반구축이 조성되고 있는 현실이다. 본고에서는 이러한 현실적인 문제에 매우 유용한 참고정보를 제공하기 위해서 사전 기술력평가와 사후 경영성과 범주(정상, 파산)사이의 관계성을 분석하였다. 이러한 관계성을 객관적으로 분석하기 위해서 (1) 다변량 이상값 탐색, (2) 다변량 분산분석, (3) 판별모형의 비교, (4) 실제정분류율 추정, (5) 판별모형의 활용 등을 포함하였다. 임계확률값의 설정과 개별 범주와 전체 범주에 대한 정분류율 비교 분석 결과 선형판별모형이 로지스틱 판별모형보다 적절한 모형으로 판단된다. 기술력평가 중항목을 이용한 선형판별모형에 의하면 전체 범주에 대한 단순정분류율은 약 69% 이었고, 예측 정분류율로 사용될 수 있는 실제 정분류율 추정치는 약67% 정도인 것으로 구해졌다. 이러한 결과는 중소벤처기업의 사후 경영성과의 변동을 설명하는데 기술력평가 자료가 핵심정보로 사용될 수 있음을 보여주고 있으며, 설명력을 추가적으로 개선하기 하기 위해서는 기술력평가 정보와 더불어 기업의 추가적인 위험요소를 고려할 필요가 있다고 판단된다.

중소벤처기업에 대한 기술금융을 위한 기반구축을 위해서는 관련 기업에 대한 중장기 시계열에 걸친 기술력평가 정보와 재무정보에 대한 데이터베이스의 구축이 절대적으로 필요하다. 이러한 데이터베이스에서 최근 2년-3년 시계열 정보를 활용하면 높은 신뢰성이 내재된 파산예측 가능성 정보를 제공할 수 있을 것이다. 또한 미래 중소벤처기업의 파산 예측 모형의 신뢰성과 설명력을 개선하기 위해서는 기술력평가 자료와 더불어 파산예측에 주요한 영향을 미치는 핵심 재무정보를 결합한 통계적 판별모형의 구축이 요구된다. 향후 이러한 구축 사업이 체계적으로 이루어진다면 추정된 판별모형에서 구한 특정 범주에 속할 사후확률에 근거하여 정상범주 혹은 파산범주에 속할 가능성을 사전에 예측할 수 있을 것이다.

참고문헌

성웅현(2000), 「응용다변량분석」, 제2판, 탐진, 1999, pp. 177-202.

_____(1997), "중소기업부실 예측을 위한 로지스틱 판별분석," 한신논문집, 제14권, pp621-636.

_____(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.

_____(2005), "기술종합지수를 이용한 기술등급평가에 관한 연구," 기술혁신학회지, 제8권 2호, pp. 583-604.

양동우(2003), "벤처의 기술평가와 경영성과의 관계에 관한 연구", 지식경영연구, 제4권 제1호, pp. 21-34.

_____(2005), "기술성지표와 기업성과의 관계비교 분석-초기중소벤처와 성장 중소벤처-," 기술혁신학회지, 제8권 제3호, pp. 1175-1198.

Aboody, David and Baruch Lev, "The Value Relevance of Intangibles: The Case Study Software Capitalization", *Journal of Accounting Research*, 36, 1998, pp 161-191.

Altmam, 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.

Brown, Mark G. and Raynold A. Svenson, "Measuring R&D Productivity", *Research Technology Management*, 1998, pp. 30-35.

Byth, K. and McLachlan, G. J(1980), "Logistic regression compared to normal discrimination for non-normal populations," Austral. *Journal of Statistics*, 22, pp. 188-196.

Clark, C. E., Webster, Hogan, and Foster, P. L., (1997), "Judgemental approach to

forecasting bankruptcy," *Business Sources Premier*, 16(2), pp.14-19.

Efron, B.(1975), "The efficiency of logistic regression compared to normal discriminant Analysis," *Journal of the American Statistical Association*, 70, pp. 892-898.

Griliches, Zvi, "Patent Statistics as Economic Literature", *Journal of Economic Literature*, 92, 1990, pp. 630-653.

Hora, S. C., and Wilcox, J. B.(1982), "Estimation of error rates in several discriminant analysis," *Journal of Marketing Research*, 19, pp. 57-61.

Ohlson, J. S.(1980), "Financial ratios and probabilistic prediction of bankruptcy," *Journal of Accounting Research*, Spring, pp.109-131.

Platt, H. D., and Platt, M. B.(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., Harmon, W. K., and Gramlich, J. D(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 Lee, Y.(1987), "Identification of a multivariate outliers," presented at the Annual Meeting of the *American Statistical Association*.

Zhang, G. Q., Hu, M. Y., Patuwo, B. E., and Indro, D.C.(1999), "Artificial neural networks in bankruptcy prediction-General framework and cross-validation analysis," *European Journal of Operation Research*, 116(1), pp. 16-32.

성웅현

성균관대학교 통계학과를 졸업하고, 미국 Ohio University 에서 경영학석사 및 Texas Tech University 에서 경영통계학 박사학위를 취득하였다. 현재 한신대학교 정보통계학과 정교수로 재직 중이다. 주요 관심분야는 다변량분석, 기술가치평가, 실물옵션, 기술경영 등이다. 주요저서로는 경영통계정보분석, SAS 경영통계, 회귀분석, 응용다변량분석, 응용로지스틱회귀분석 등이 있다. 또한 통계분석, 기술가치평가, 실물옵션, 기술력경쟁력지수 등에 관한 다수의 논문이 있다.