

개인신용평점에서 항목그룹화와 모형평가를 위한 교육용 소프트웨어의 개발[†]

정기문¹

¹경성대학교 정보통계학과

접수 2010년 9월 28일, 수정 2010년 11월 17일, 게재확정 2010년 11월 22일

요 약

개인신용평점에서 항목그룹화의 과정은 연속형 특성변수를 밴드로 분할하고 이산형 특성변수는 그룹으로 분할하는 것이다. 또한 평점표는 시간이 지나감에 따라 성능이 떨어지게 되고, 따라서 사용되고 있는 승인점을 조정하여야 한다. 그러나 개인신용평점에서 항목그룹화와 승인점의 조정은 매우 복잡하고 번거로운 과정이라고 할 수 있다. 따라서 본 논문에서는 비주얼베이직을 사용하여 개인신용평점에서 항목그룹화와 모형평가를 위한 소프트웨어를 개발하였다. 개발된 소프트웨어를 활용하면 항목그룹화에서 최적의 분할과 모형평가에서 최적의 승인점을 쉽게 찾을 수 있다.

주요용어: 개인신용평점, 비주얼베이직, 승인점, 항목그룹화.

1. 서 론

개인의 신용위험을 예측하는 문제는 통계학 및 확률모형화 분야에서 매우 중요한 문제로 연구되고 있다 (Thomas, 2000). 이러한 개인신용평점 (credit scoring)과 관련된 연구는 판별분석 (discriminant analysis), 로지스틱 회귀분석 (logistic regression analysis), 생존분석 (survival analysis) 등을 활용하여 개인의 신용을 평가하기 위한 신용평점모형을 구축하는 분야와 이렇게 구축된 신용평점모형의 평점 자료를 활용하여 모형을 평가하는 분야로 나뉘어져 활발히 진행되고 있다.

우선, 개인신용평가를 위한 모형화 분야에서는 각 특성들을 상대적으로 작은 몇 개의 그룹으로 묶어줌으로써 로버스트한 신용평가시스템이 구축되도록 하기 위해서 항목그룹화 (coarse classifying)를 활용하고 있다. 이러한 항목그룹화는 특성변수들의 항목을 유사한 성질을 가진 항목들로 그룹화를 하는 것을 의미하는데, 항목그룹화를 수행하는 과정이 매우 번거롭기 때문에 이를 위한 소프트웨어의 개발이 필요하다 (Jung 등, 2006). 또한, 구축된 기존의 신용평점모형으로부터 얻게 되는 평점자료를 활용하여 신용평점모형의 성능을 지속적으로 평가하여야 한다. 그러나 이러한 구축된 신용평점모형을 평가하기 위해서는 번거로운 계산과 많은 비교를 수행하여야 한다. 따라서 신용평점모형을 평가하기 위한 소프트웨어의 개발도 필요하다고 할 수 있다.

또한 대학의 통계학 관련 강좌에서 강의를 효과적으로 진행하기 위한 교육용 소프트웨어의 개발과 관련된 연구가 꾸준히 진행되고 있다 (Choi와 Kim, 2005; Choi, 2006a; Choi, 2006b). 따라서 위에서 언급한 항목그룹화와 모형평가를 위한 교육용 소프트웨어를 개발함으로써 최근에 통계학과에서 관심을 많

[†] 이 연구는 2009학년도 경성대학교의 연구년지원에 의하여 연구되었음.

¹ (608-736) 부산광역시 남구 대연 3동 110-1, 경성대학교 정보통계학과, 부교수.
E-mail: kmjung@ks.ac.kr

이 갖고 있는 금융통계와 관련된 학부 및 대학원 강좌에서 수업의 보조 자료로써 활용될 수 있을 뿐만 아니라 프로그래밍 언어 특히, 비주얼베이직을 활용한 통계소프트웨어 개발과 관련된 강좌에서는 훌륭한 사례로 활용할 수 있을 것으로 생각한다. Jung 등 (2006)은 항목그룹화를 위한 소프트웨어를 개발하였으나, χ^2 -통계량 하나만을 제공해 주고, 시각적인 분석이 제공되지 않는 등 아주 단순한 형태를 벗어나지 못하였다. 또한, 구축된 신용평점모형을 평가하기 위한 메뉴 또한 제공되지 않았다. 따라서 이러한 부분을 좀 더 보완한 향상된 소프트웨어의 개발이 필요하다고 할 수 있다.

따라서 이번에 수정 보완된 개인신용평점에서 항목그룹화와 모형평가를 위한 소프트웨어 SCCSM (Software for Coarse Classifying and Scorecard Maintenance) 2010은 통계학과 학부과정에서 개설되고 있는 통계소프트웨어 개발과 관련된 강좌에서 흔히 사용되고 있는 비주얼베이직을 이용하여 개발하였으며, 개발된 소프트웨어의 모든 절차는 메뉴방식으로 실행이 가능하여 프로그램에 관한 전문적인 지식이 없는 일반 사람들도 쉽게 사용할 수 있도록 하였다. 본 논문은 다음과 같이 구성된다. 2장에서는 항목그룹화와 개인신용평점의 모형평가에 대한 이론적 배경에 대해서 간단히 살펴본다. 그리고 3장에서는 개인신용평점에서 항목그룹화와 모형평가를 위해서 비주얼베이직으로 개발된 소프트웨어 SCCSM 2010을 소개하고 개발된 소프트웨어를 사용하는 방법에 대해서 예제를 통하여 자세히 설명하고자 한다. 그리고 기존의 항목그룹화를 위한 소프트웨어와 개발된 소프트웨어 SCCSM 2010을 간단히 비교하고자 한다.

2. 항목그룹화와 모형평가

2.1. 신용평점모형에서 항목그룹화

항목그룹화는 특성변수들의 항목을 유사한 성질을 가진 항목들로 그룹화 하는 것을 의미하는데, 이는 각 특성변수들을 상대적으로 작은 몇 개의 그룹으로 묶어줌으로써 로버스트한 신용평점시스템을 구축하도록 해준다. 즉, 항목그룹화를 통해서 명목형 특성의 경우 항목의 수가 너무 많으면 특정 속성에 대한 표본이 충분하지 않아 로버스트한 결과를 얻을 수 없는 문제점을 해결할 수 있으며, 연속형 특성의 경우에는 신용평점시스템을 이용하여 예측을 하기 위해서 비선형이 아닌 선형을 고려할 경우의 문제점을 해결할 수 있다. 예를 들어, 그림 2.1에서와 같이 우량고객의 비율이 나이 (age)에 대해 단조증가하지 않는 경우에 일반적인 회귀선은 이를 반영하지 못하는 문제점이 있고, 이러한 문제점을 항목그룹화를 통해서 해결할 수 있다 (Thomas 등, 2002). 따라서, 이러한 항목그룹화에서의 주된 관심은 어떠한 항목들을 하나의 동일한 그룹으로 만들겠느냐 하는 것인데, 이때, 가장 잘 분류된 항목그룹화를 최적분할이라고 한다. 이러한 최적분할을 찾기 위해서는 승산비를 이용하거나 χ^2 -통계량, 정보통계량, 일치통계량 등을 사용한다.

항목그룹화를 위해서 전통적으로 승산비가 사용되는데, 항목별로 승산비를 구하여 유사한 항목을 하나의 그룹으로 만들어 최적분할을 찾게 된다. 그러나 이 방법은 항목들 사이에 승산비의 차이가 뚜렷하지 않거나, 항목의 수가 상대적으로 많은 경우 최적분할을 찾는데 연구자의 주관적 의사가 많이 개입되어야 한다. 따라서 다음과 같이 정의되는 χ^2 -통계량을 이용하여 일반적으로 항목그룹화를 수행하게 된다.

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \quad (2.1)$$

여기서, O_{ij} 는 i, j 셀의 관측도수이고, E_{ij} 는 i, j 셀의 기대도수이다.

그러나 이러한 통계량을 사용하여 최적분할을 판단하기 위해서는 여러 번 반복해서 통계량을 구하고 이를 비교해야 하는 작업이 필요하다. 예를 들어, 어떤 특성변수의 항목이 C_1, C_2, C_3 의 3개인 경우에

두 그룹으로 항목그룹화를 수행하려면 $(C_1, C_2 + C_3)$, $(C_2, C_1 + C_3)$, $(C_3, C_1 + C_2)$ 에 대한 각 통계량 값을 구해서 이를 비교하여야 한다. 그리고 특성변수의 항목이 C_1, C_2, C_3, C_4 의 4개이면 $(C_1, C_2 + C_3 + C_4)$, $(C_2, C_1 + C_3 + C_4)$, $(C_3, C_1 + C_2 + C_4)$, $(C_4, C_1 + C_2 + C_3)$, $(C_1 + C_2, C_3 + C_4)$, $(C_1 + C_3, C_2 + C_4)$, $(C_1 + C_4, C_2 + C_3)$ 에 대한 각 통계량 값을 구해서 이를 비교하여야 한다. 따라서 항목의 수가 많아지면 상당히 많은 항목그룹화의 경우에 대하여 반복적으로 통계량 값을 구하고 이를 비교해야 한다. 그러므로 단순하지만 반복적인 작업이 필요한 항목그룹화를 수행하는 과정을 위한 소프트웨어의 개발이 필요하다 할 수 있다.

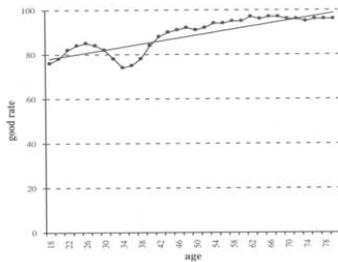


그림 2.1 나이 (age)에 따른 우량고객의 비율 (Thomas 등, 2002)

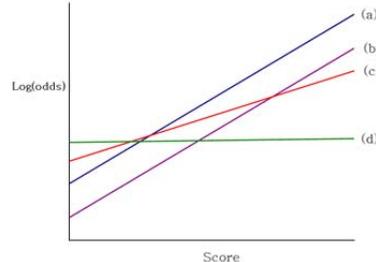


그림 2.2 오즈 대 평점 (Thomas와 Jung, 2010)

2.2. 신용평점모형의 평가

개인신용평점모형은 시간이 지나감에 따라 모형의 예측능력이 저하되기 때문에 평점모형은 구축되어 사용하기 시작하는 순간부터 지속적으로 관찰되어야 한다. 일반적으로 개인특성을 x , 신용평점을 $s(x)$ 그리고 신용이 좋을 확률을 $p = P(G|s(x))$ 라고 하면 평점과 로그오즈 사이에는 다음과 같은 관계가 있기 때문에 평점모형의 예측능력의 변화는 로그오즈 대비 평점의 그래프로 확인을 할 수 있다 (Thomas, 2009).

$$\ln \left(\frac{P(G|s)}{P(B|s)} \right) = \ln \left(\frac{p}{1-p} \right) = a + bs(x). \tag{2.2}$$

그림 2.2는 로그오즈 대비 평점의 전형적인 그래프를 보여준다. 여기서 선 (a)는 개발표본의 로그오즈 대비 평점을 나타내고, (b), (c) 그리고 (d)는 최근표본에 대하여 나타낼 수 있는 오즈 대비 평점이다. 선 (b)의 경우는 선 (a)와 기울기가 동일하기 때문에 위험순위를 매기는 상대적인 능력은 개발표본에 해당하는 선 (a)와 동일하다고 할 수 있다. 그리고 선 (c)는 선 (a)에 비하여 기울기가 완만하여 평점표의 능력이 저하되었음을 알 수 있다. 끝으로 선 (d)는 평점표로서의 의미가 전혀 없음을 나타낸다.

Thomas와 Jung (2010)은 시간에 따른 신용평점모형의 비용분석을 통하여 모형을 평가하고 관리하는 방법에 대하여 소개하였는데, 식 (2.2)를 이용하여 주어진 시점 t 에서 구축된 신용평점모형을 사용함으로써 발생하는 기대손실비용을 다음과 같이 구하였다.

$$E(a_t, b_t, c_t) = N \left(\int_{c_t}^{\infty} \frac{Df(s)}{1 + e^{a_t + b_t s}} ds + \int_{-\infty}^{c_t} \frac{Le^{a_t + b_t s} f(s)}{1 + e^{a_t + b_t s}} ds \right), \tag{2.3}$$

여기서 a_t 와 b_t 는 시점 t 에서 식 (2.2)를 만족하는 a 의 값과 b 의 값이고, c_t 는 시점 t 에서 사용할 승인점 (cut-off score)이며, N 은 전체 고객의 수이다. 그리고 D 는 신용이 나쁜 고객을 채택하여 발생하게

되는 단위비용이며, L 은 신용이 좋은 고객을 기각하여 발생하는 단위비용이고, $f(s)$ 는 평점분포 (score distribution)를 나타낸다. 따라서 주어진 a_t, b_t, c_t 에 대한 기대비용을 구할 수 있으며, 이러한 비용을 최소화하는 c_t 의 값을 찾으면 이 값이 주어진 a_t 와 b_t 에 대한 최적의 승인점인 된다.

3. SCCSM 2010

이 절에서는 항목그룹화와 모형평가를 위해서 비주얼베이직으로 개발된 소프트웨어 SCCSM 2010의 사용방법을 자세히 소개하고자 한다. 본 연구에서 개발된 소프트웨어 SCCSM 2010의 모든 절차는 메뉴방식으로 실행이 가능하여 프로그램에 관한 전문적인 지식이 없는 일반 사람들도 쉽게 사용할 수 있도록 하였다. 따라서 학부 및 대학원의 금융통계 관련 강좌에서 보조 자료로 충분히 활용될 수 있을 것이다.

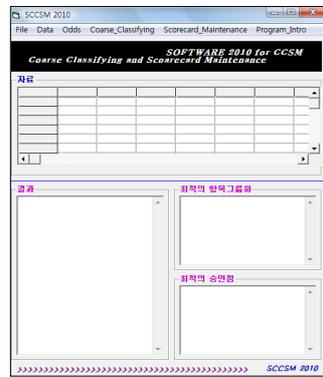


그림 3.1 소프트웨어의 초기화면

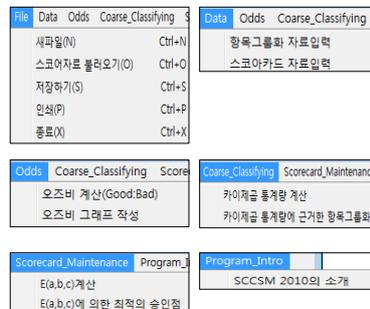


그림 3.2 소프트웨어의 주요 메뉴

먼저, 그림 3.1은 개발된 항목그룹화와 모형평가를 위한 소프트웨어 SCCSM 2010의 첫 화면인데, 화면구성이 메뉴 표시줄, 자료 창, 결과 창, 최적의 항목그룹화 창, 최적의 승인점 창 등으로 되어 있음을 알 수 있다. SCCSM 2010은 항목그룹화와 모형평가를 위하여 자료를 각각 불러오거나 직접 입력할 수 있도록 하였으며, 최적의 항목그룹화를 설정하고 시간에 따른 신용평점모형의 성능을 분석할 수 있는 메뉴를 포함하고 있다. 본 연구에서 개발된 소프트웨어에서 사용 가능한 주요 메뉴들이 그림 3.2에 소개되어 있으며, 다음 3.1절과 3.2절에서 항목그룹화와 모형평가를 위하여 개발된 소프트웨어 SCCSM 2010을 어떻게 사용할 수 있는지를 자세히 소개하고자 한다.

3.1. SCCSM 2010을 활용한 항목그룹화

이 절에서는 개발된 소프트웨어 SCCSM 2010에서 항목그룹화와 관련된 메뉴를 소개하고자 한다. 먼저, 그림 3.2에 있는 소프트웨어의 첫 화면에서 볼 수 있는 것처럼 항목그룹화를 위해서 자료를 불러오거나 소프트웨어에서 직접 입력할 수 있다. 그리고 각 항목에 대한 승산비를 구할 수 있으며, 구해진 승산비를 사용하여 시각적으로 표현이 가능하도록 그래프 작성 메뉴를 추가하였다. 또한 카이제곱 통계량을 기준으로 최적의 항목그룹화를 찾기 위한 절차가 메뉴 방식으로 실행되도록 만들었다. 이제, 표 3.1에 나타나 있는 특성변수에 대한 자료를 사용하여 개발된 소프트웨어를 어떻게 활용하여 항목그룹화를 수행할 수 있는지를 단계별로 자세히 살펴보고자 한다.

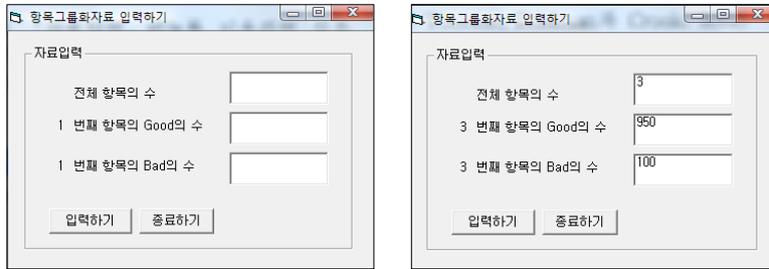
표 3.1 특성변수 주거형태별 신용상태

구분	자가	월세	부모와 거주	계
양호 (Good)	6,000	1,950	950	8,900
불량 (Bad)	300	540	100	940
계	6,300	2,490	1,050	9,840

(자료: Thomas 등, 2002)

(1) 자료입력

[Data]메뉴의 [항목그룹화 자료입력]메뉴를 이용하면 항목그룹화를 하고자하는 특성변수에 대한 자료를 쉽게 입력할 수 있다. 다음은 [항목그룹화 자료입력]메뉴를 이용하여 표 3.1의 자료를 입력하는 방법을 보여준다. 전체 항목 3개에 대하여 Good의 수와 Bad의 수를 입력한 후 종료하기 버튼을 클릭하면 그림 3.3과 같이 입력된 자료가 자료 창에 나타나게 된다.



(2) 로그승산비의 계산

로그승산비를 계산하기 위해서는 [Odds]메뉴의 [오즈비 계산 (Good:Bad)]메뉴를 선택하면 된다. 그러면 그림 3.4와 같이 소프트웨어의 자료 창에 로그승산비가 계산되어 나타나는 것을 볼 수 있다.

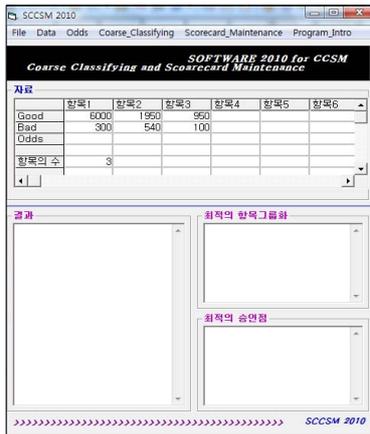


그림 3.3 자료입력 후의 화면그림

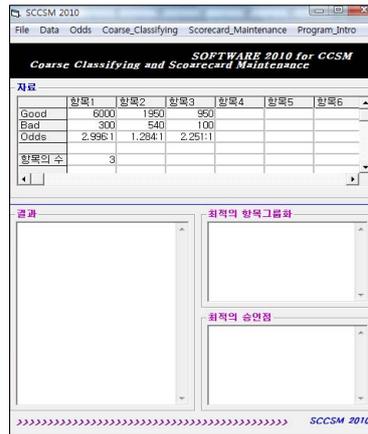
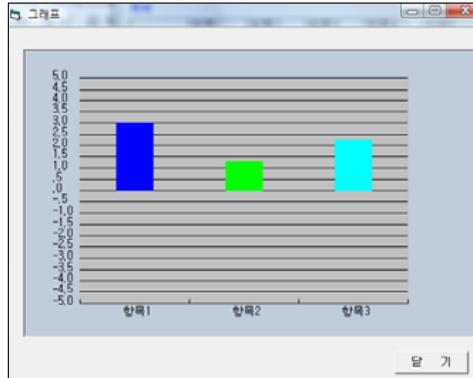


그림 3.4 입력된 자료의 승산비 계산

(3) 로그승산비 그래프 그리기

[Odds]메뉴의 [오즈비 계산 (Good:Bad)]메뉴를 이용하여 계산된 오즈비를 시각적으로 표현하기 위해서는 그래프를 작성하여야 하는데, 이를 위해서는 [Odds]메뉴의 [오즈비 그래프 작성]메뉴를 선택하면 된다. 그러면 오른쪽과 같은 그래프가 작성되는 것을 알 수 있다. 이 그래프로부터 대략적으로 항목1 (자가)과 항목3 (부모와 거주)을 하나의 그룹으로 만드는 것이 최적분할이 된다는 것을 알 수 있다.



(4) χ^2 통계량 값의 계산

최적의 분할을 찾기 위해서 χ^2 통계량 값을 계산해야 하는데, 이를 위해서는 [Coarse_Classifying]메뉴의 [카이제곱 통계량 계산]메뉴를 선택하면 된다. 그림 3.5의 결과 창에는 [카이제곱 통계량 계산]메뉴를 이용해서 얻게 되는 각 그룹에 대한 카이제곱 통계량 값을 보여주고 있다. 결과 창의 Y(1,1)=1은 항목1 (자가), 항목2 (월세)+항목3 (부모와 거주)를 의미하는 것으로 이때의 χ^2 통계량 값이 465.2071라는 의미이다. 마찬가지로 Y(1,1)=2는 항목2 (월세), 항목1 (자가)+항목3 (부모와 거주)를 의미하고 Y(1,1)=3은 항목3 (부모와 거주), 항목1 (자가)+항목2 (월세)를 의미하는 것이다.

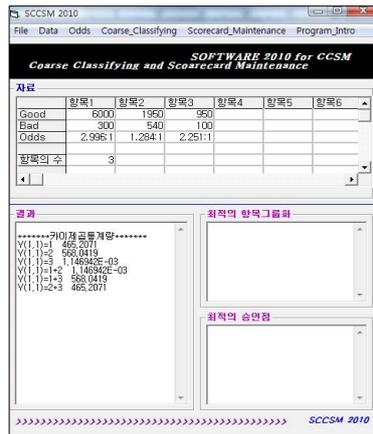


그림 3.5 χ^2 통계량 값의 계산그림

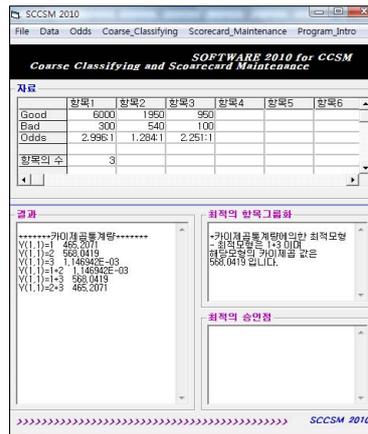


그림 3.6 최적분할을 통한 항목그룹화

(5) χ^2 통계량 값에 근거한 최적분할

최적분할을 통한 항목그룹화를 수행하기 위해서는 [카이제곱 통계량 계산]메뉴를 통해서 얻게 되는 각 항목그룹별 카이제곱 통계량 값을 비교해야 한다. 이를 위해서는 [Coarse_Classifying]메뉴의 [카이제곱 통계량에 근거한 항목 그룹화]메뉴를 선택하면 된다. 그러면, 그림 3.6과 같이 최적의 항목그룹화 창에 선택된 최적분할이 나타나게 된다. 즉, 항목1 (자가)과 항목3 (부와 거주)을 하나의 그룹으로 만드는 것이 최적분할이 된다는 것이다.

3.2. SCCSM 2010을 활용한 모형평가

이 절에서는 개발된 소프트웨어 SCCSM 2010에서 모형평가와 관련된 메뉴를 소개하고자 한다. 먼저, 그림 3.1에 있는 소프트웨어의 첫 화면과 그림 3.2에 있는 소프트웨어의 주요 메뉴에서 볼 수 있는 것처럼 모형평가를 위해서 자료를 불러오거나 소프트웨어에서 직접 입력할 수 있다. 그리고 입력된 자료를 사용하여 구축된 신용평점모형을 사용함으로써 발생하는 기대손실비용을 계산할 수 있도록 하였다. 또한 기대손실비용을 기준으로 최적의 승인점을 찾기 위한 절차가 메뉴 방식으로 실행되도록 만들었다. 이제, 표 3.2와 같이 주어지게 되는 평점자료를 사용하여 개발된 소프트웨어를 어떻게 활용하여 모형평가를 수행할 수 있는지를 자세히 살펴보고자 한다.

표 3.2 신용평점모형의 스코어 자료

구간	534-644	645-658	...	795-807	전체
도수	257,596	278,784	...	760,155	15,106,387

(자료: Thomas와 Jung, 2010)

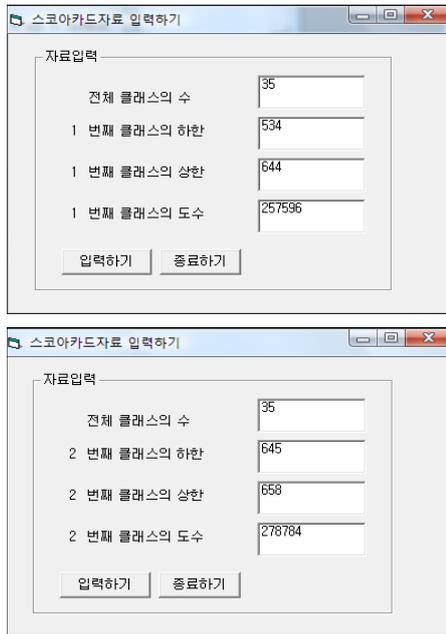


그림 3.7 자료입력 화면

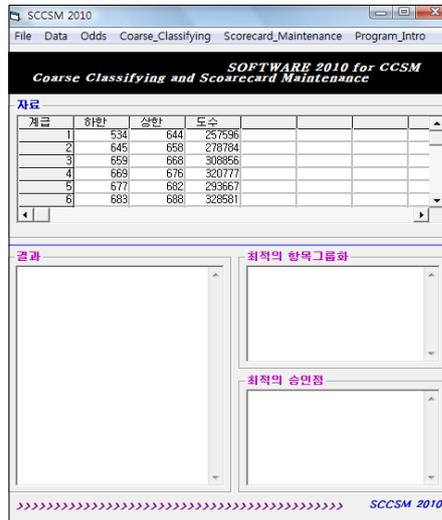


그림 3.8 자료입력 후의 화면

(1) 자료입력

[File]메뉴의 [스코어자료 불러오기]메뉴를 이용하거나 [Data]메뉴의[스코어카드 자료입력]메뉴를 이용하면 항목그룹화를 하고자하는 특성변수에 대한 자료를 쉽게 입력할 수 있다. 다음은[Data]메뉴의[스코어카드 자료입력]메뉴를 이용하여 표 3.2의 자료를 입력하는 방법을 보여준다. 전체 클래스 35개에 대하여 각 클래스의 하한과 상한 그리고 도수를 그림 3.7과 같이 입력한 후 입력하기 버튼을 클릭하면 그림 3.8과 같이 입력된 자료가 자료 창에 나타나게 된다.

(2) 기대손실비용의 계산

시점 t 에서 a , b 그리고 c 의 값이 주어지는 경우에, 구축된 신용평점모형을 사용함으로써 발생하는 기대손실비용의 계산을 위해서는 [Scorecard_Maintenance]메뉴의 [E(a,b,c)]메뉴를 선택하면 된다. [E(a,b,c)]메뉴를 선택하면 그림 3.9와 같이 a , b 그리고 c 의 값을 입력할 수 있는 대화상자가 나타나는데, 이를 사용하여 세 값을 입력하고 나면 그림 3.10과 같이 결과 창에 주어진 a , b 그리고 c 에 대한 기대손실비용이 구해진다.

cs2010
Input the value of intercept(a)
확인 취소
-1.2

cs2010
Input the value of slope(b)
확인 취소
0.024

cs2010
cut-off score(band)를 입력하세요.
확인 취소
3

그림 3.9 자료입력 화면

SCCSM 2010
File Data Odds Course_Classifying Scorecard_Maintenance Program_Intro
SOFTWARE 2010 for CCSM
Course Classifying and Scorecard Maintenance

자료

계급	하한	상한	도수
1	534	644	257596
2	645	668	278764
3	659	668	308856
4	669	676	320777
5	677	682	293667
6	683	696	329581

결과

E(a,b,c) ###
a(intercept) = -1.2
b(slope) = .024
c(cut-off band) = 3
E(-1.2, .024, 3) = 1.50902727175694

최적의 항목그룹화
최적의 승인점

SCCSM 2010

그림 3.10 기대손실비용의 계산

(3) 기대손실비용에 근거한 최적의 승인점

기대손실비용을 최소화하는 최적의 승인점을 찾기 위해서는 [Scorecard_Maintenance]메뉴의 [E(a,b,c)]메뉴를 통해서 얻게 되는 각 승인점별 기대손실비용을 비교해야 한다.

이를 위해서는 [Scorecard_Maintenance]메뉴의 [E(a,b,c)]에 의한 최적의 승인점]메뉴를 선택하면 된다. 이 때, a 와 b 의 값을 입력하게 되는 대화상자가 나타나며, 이를 이용하여 이 두 값을 입력하면 모든 가능한 c 의 값에 대하여 기대손실비용을 구한 후, 그림 3.11과 같이 기대손실비용이 가장 작게되는 최적의 승인점이 나타나게 된다.

즉, 승인점이 $c = 2$ 인 경우에 기대손실비용이 최소가 된다는 사실을 알 수 있다.

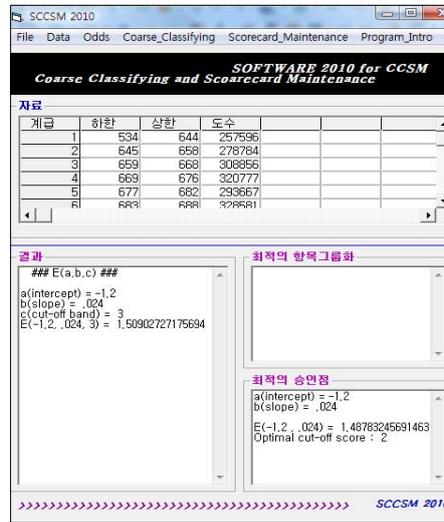


그림 3.11 최적의 승인점

3.3. 기존 소프트웨어와의 비교

본 연구에서는 비주얼 베이직을 활용하여 개인신용평점에서 평점모형을 구축하는데 사용되는 항목그룹화와 구축된 평점표를 평가하기 위한 소프트웨어 SCCSM 2010을 개발하였다. 이와 관련되어 Jung 등 (2006)은 항목그룹화를 위한 소프트웨어를 개발하였다. 따라서 본 절에서는 기존의 항목그룹화를 위한 소프트웨어와 이번에 개발된 SCCSM 2010을 추가된 메뉴를 중심으로 간단히 비교하고자 한다.

Jung 등 (2006)이 개발한 항목그룹화를 위한 소프트웨어는 χ^2 -통계량을 제공해 주고, 시각적인 분석이 제공되지 않는 등 아주 단순한 형태를 벗어나지 못하였다. 또한, 구축된 신용평점모형을 평가하기 위한 메뉴는 전혀 제공되지 않았다. 구체적으로 SCCSM 2010은 기존의 소프트웨어에 비해서 다음과 같은 부분들이 추가되고 변경되었다. 첫째, 전통적인 항목그룹화의 방법인 로그승산비를 구하고, 이를 사용하여 시각적인 방법에 의해서 대략적으로 항목그룹화를 수행할 수 있도록 그래프를 작성할 수 있는 메뉴를 추가하였다. 둘째, 항목그룹화 자료와 신용평점 자료를 개발된 소프트웨어의 자료 창에 생성할 수 있는 주 메뉴를 추가하였다. 셋째, 기존의 신용평점모형으로부터 얻은 평점자료를 활용하여 신용평점모형을 평가하기 위한 기대손실비용을 계산 할 수 있는 메뉴를 추가하였다. 넷째, 계산된 기대손실비용에 근거하여 최적의 승인점을 찾을 수 있는 메뉴를 추가하였다. 끝으로, 기타 수정 및 보완이 필요한 부분들을 수정하였다.

4. 결론

본 연구에서는 비주얼 베이직을 활용하여 개인신용평점에서 평점모형을 구축하는데 사용되는 항목그룹화와 구축된 평점표를 평가하기 위한 소프트웨어 SCCSM 2010을 개발하였다. SCCSM 2010에서의 모든 절차는 메뉴방식으로 실행할 수 있도록 하여 컴퓨터에 익숙하지 않은 초보자들도 쉽게 사용할 수 있도록 하였다.

이번 연구에서 개발된 SCCSM 2010을 활용하면 매우 많은 경우에 대하여 χ^2 통계량을 고려해야만

하는 항목그룹화의 비효율적인 절차들을 간단히 해결할 수 있다. 그리고 구축된 기존의 신용평점모형으로부터 얻게 되는 평점자료를 활용하여 기대손실비용을 계산하여 신용평점모형의 성능을 평가할 수 있다. 더불어 구해진 기대손실비용을 최소화하는 최적의 승인점을 찾을 수도 있다. 개발된 소프트웨어인 SCCSM 2010은 개인신용평점에서 항목그룹화를 수행하고 신용평점모형을 평가할 수 있다는 소프트웨어로서의 본질적인 기능 이외에, 통계학과의 금융통계와 관련된 학부 및 대학원 강좌에서 수업의 보조 자료로 활용한다면 강의의 매우 효과적으로 진행할 수 있을 것으로 기대한다. 뿐만 아니라 프로그래밍 언어 특히, 비주얼베이직을 활용한 통계소프트웨어 개발과 관련된 강좌에서는 교재에서 흔히 제공되는 전형적인 형태가 아닌 실제 개발 사례로 활용될 수 있을 것으로 생각한다.

참고문헌

- Choi, H. S. (2006a). A program for teaching Type I and Type II errors. *Journal of the Korean Data & Information Science Society*, **17**, 19-30.
- Choi, H. S. (2006b). The teaching of statistics using Excel VBA. *Journal of the Korean Data & Information Science Society*, **17**, 811-820.
- Choi, H. S. and Kim, T. Y. (2005). Computer simulation program for central limit theorem. *Journal of the Korean Data & Information Science Society*, **16**, 359-369.
- Jung, K. M., Kim, M. C. and Yum, J. K. (2006). Development of software for coarse classifying. *Journal of the Korean Data & Information Science Society*, **17**, 1085-1090.
- Thomas, L. C. (2000). A Survey of Credit and behavioral scoring: Forecasting financial risk of lending to consumers. *International Journal of Forecasting*, **16**, 149-172.
- Thomas, L. C. (2009). *Consumer credit models: pricing, profit and portfolios*, Oxford University Press, Oxford.
- Thomas, L. C., Edelman D. B. and Crook, J. N. (2002). *Credit scoring and its application*, SIAM.
- Thomas, L. C. and Jung, K. M. (2010). When to rebuild and when to recalibrate credit scorecards. *24th European Conference on Operational research*.

Development of educational software for coarse classifying and model evaluation in credit scoring[†]

Ki Mun Jung¹

¹Department of Informational Statistics, Kyungsoong University

Received 28 September 2010, revised 17 November 2010, accepted 22 November 2010

Abstract

The coarse classifying procedure in credit scoring splits the values of a continuous characteristic into bands and the values of a discrete characteristic into groups of values. Also, the scorecard degrades over time and thus we should adjust the cut-off score being used. However, the coarse classifying and the adjustment of cut-off score in credit scoring are very complicate and troublesome procedure. Thus, in this paper, we develop a software for the coarse classifying and the model evaluation by using Visual Basic Language. By using the developed software, we can find the best split in the coarse classifying and the optimal cut-off score in the model evaluation.

Keywords: Coarse classifying, credit scoring, cut-off score, Visual Basic.

[†] This work was supported by Kyungsoong University in 2009.

¹ Associate professor, Department of Informational Statistics, Kyungsoong University, Busan 608-736, Korea. E-mail: kmjung@ks.ac.kr