

## 보험사 언더라이팅 기준 설정을 위한 스코어링 기법에 관한 연구

이창수<sup>1</sup> · 권혁성<sup>2</sup> · 김동광<sup>3</sup>

<sup>1</sup>승실대학교 정보통계보험수리학과, <sup>2</sup>승실대학교 정보통계보험수리학과,

<sup>3</sup>승실대학교 정보통계보험수리학과

(2009년 3월 접수, 2009년 3월 채택)

### 요약

언더라이팅은 보험자가 청약자에 대한 인수여부를 결정하는 보험계약 운영의 초기단계로서, 그 기준의 정교성 여부에 따라 회사에 안정적인 수익을 가져올 수도 있고 예기치 못한 큰 손실을 야기할 수도 있다. 일반적으로, 각 청약자의 위험요소를 파악, 평가하여 위험정도를 평가할 수 있는 점수를 이용하는 스코어링 시스템을 언더라이팅 기준으로 이용하는데, 점수를 산출하기 위한 방법은 각 위험요소별 점수 기준을 설정하고 위험요소별 점수를 합산하는 형태가 보편적이다. 최근 위험율차 이익의 중요성이 강조됨에 따라 기존의 방법에 비해 보다 효과적인 언더라이팅 기법이 요구되고 있다. 본 연구에서는 일반화 선형모형을 이용하여 보다 합리적으로 보험 계약자에 대한 위험도를 추정하는 방법을 제시하였다. 실증분석을 통해 위험집단의 위험도를 평가하고 점수화하는 구체적인 언더라이팅 모형을 설계하였고, 제시된 언더라이팅 모형의 적용효과를 반영하여 언더라이팅 기준을 설정하는 방법에 대해서도 설명하였다.

주요용어: 언더라이팅, 스코어링기법, 일반화선형모형, 생명보험.

### 1. 서론

보험회사에서 새로운 계약을 인수하거나 기존 계약을 갱신하는 기준을 마련하고 이를 실행하는 일련의 과정을 언더라이팅(underwriting)이라 한다. 언더라이팅의 기본 목적은 동일한 보험요율을 적용받는 피보험자 집단 내 실제 위험 정도의 동질성을 최대한 확보하여 보험사의 보험리스크를 안정적으로 운영하기 위한 것이다. 효과적인 언더라이팅 과정을 통해 리스크 수준이 상대적으로 낮은 피보험자들을 선별하여 회사의 재정적 이익을 증대할 수 있고, 부수적으로 회사의 평판을 높임과 동시에 고객의 높은 계약유지율을 달성할 수 있다. 반대로 언더라이팅의 부실화는 역선택 또는 보험사기에 대한 리스크를 증가시켜 회사의 손실을 증가시킬 수 있다. 일반적으로, 동일한 보장내용을 가진 보험상품일지라도 각 보험회사의 장기적인 목표나 전략에 따라 인수여부를 결정하는 기준이 상이하게 수립되는 것이 보통이다. 생명보험사의 언더라이팅과 관련한 내용은 Atkinson과 Dallas (2000)와 Bickley (2007)에 자세히 소개되어 있다. 또한, 미국, 캐나다의 경우 보험계리사 협회(Society of Actuaries)에서 생명보험 업계의 언더라이팅 현황을 조사하고 보고하는 작업을 수행하고 있는데, 예를 들어 우량제 기준을 회사마다 어떻게 설정하는지에 대한 조사도 중요한 정보를 제공하고 있다. 이와 관련한 최근 보고서로 Society of Actuaries (2005)를 참조하길 바란다.

본 연구는 승실대학교 교내연구비 지원으로 이루어졌음.

<sup>2</sup>교신저자: (156-743) 서울시 동작구 상도동 승실대학교 정보통계보험수리학과, 전임강사.

E-mail: hskwon@ssu.ac.kr

보험 언더라이팅에서는 특정 상품에서 보장해주는 담보에 영향을 미칠 수 있는 요소를 선택하고 각 요소가 미치는 중요도에 따라 보험 청약자의 위험 수준에 대한 점수를 매겨 일정 기준을 초과하는 위험 수준을 가지고 있는 청약자는 인수를 제한하거나 거절하는 방법을 보편적으로 이용하는데, 이를 언더라이팅 스코어링 시스템(scoring system)이라 한다. 유사한 방법을 사용하는 다른 분야의 적용사례로 신용평가와 관련한 스코어링 시스템에 대한 일반적 논의는 Lewis (1992)에 소개되어 있다. 국내의 경우 자동차보험에서 보험금 지급건수나 보험금 규모에 영향을 미치는 다양한 요소들을 이용하는 통계적 모형, 구체적으로 일반화 선형모형을 바탕으로 종합적 위험수준을 평가하여 언더라이팅에 반영하는 시스템을 구축하여 활용한 대형 보험사의 사례가 있다. 그러나, 생명보험사의 경우 수리적 언더라이팅 기법의 활용은 현재까지도 매우 초보적인 수준이며, 통계적 방법론에 의한 접근은 사례를 찾아볼 수 없다. 따라서, 손해보험에서 효과적으로 활용된 사례가 있는 일반화 선형모형을 이용하여 생명보험에의 적용 가능성을 생각해 보는 것은 의미있는 시도가 될 것이라 생각한다. 더욱이 최근 보험사간 경쟁과 수익구조의 악화를 극복하기 위해서는 보다 효율적인 언더라이팅 기법의 적용이 요구된다. 앞으로 다양한 언더라이팅 요소의 효과를 합리적으로 반영할 수 있는 통계적 모형을 이용하여 보다 정교하게 리스크 수준을 측정, 평가할 수 있는 언더라이팅 방법에 대한 연구가 확대되어야 한다고 생각한다.

본 연구에서는 생명보험사의 실제 자료를 이용하여 피보험자 집단의 위험 요소별 특성에 따른 위험수준을 보다 합리적으로 측정하고 비교할 수 있는 방법을 제시하였다. 구체적으로, 일반화선형모형을 이용하여 각 위험집단별 사고율을 추정한 결과를 기초로 스코어링 시스템을 설계하고, 이러한 스코어링 시스템을 이용하는 언더라이팅 방법의 적용에 따른 효과를 분석하였다.

본 논문의 구성은 다음과 같다. 제 1장 서론에 이어 제 2장에서는 본 연구에 활용된 실제 자료에 대해 설명하고 연구에 이용된 통계모형에 대하여 논의하였다. 제 3장에서는 통계모형의 적합 결과와 이를 이용한 수리적 분석 결과에 대하여 설명하였으며, 제 4장에서는 연구 내용과 결과를 요약, 정리하고 향후 개선방향과 과제를 제시하였다.

## 2. 분석 대상과 분석 방법

### 2.1. 분석 대상

통계적 모형을 이용한 언더라이팅 기준 설정에 대한 실증적 분석을 위하여 본 연구에서는 국내 생명보험사의 과거 계약자료를 활용하였다. 실증분석을 위한 자료는 담보의 성격에 따라 분류된 9개의 담보군에 대한 2004년부터 2007년까지의 계약별 실적자료를 포함하고 있다.

분석용 자료를 집계하기 위한 기초작업으로 먼저 언더라이팅 효과를 분석하기 적절한 담보군을 선정해야 하는데, 담보군의 실적에 따른 언더라이팅 기준의 수립만으로 접근하기 어려운 두 가지 경우가 있다. 첫 번째 경우로 다양한 리스크 수준을 가진 피보험자 군의 손해율이 모두 양호한 경우, 즉 100% 미만의 손해율을 가진 경우이다. 이러한 경우는 해당 담보군에 대하여 현재의 인수과정에 시급히 개선해야 할 문제가 없다고 볼 수 있고, 위험률차 이익을 앞으로 계속 실현해 나갈 수 있을 것이라 기대되기 때문에 추가적인 언더라이팅 기준의 도입을 당장 고려할 필요성이 적다. 두 번째는 다양한 리스크 수준을 가진 피보험자 군의 손해율이 모두 불량한 경우, 즉 100% 이상의 손해율을 가진 경우이다. 이러한 경우 새로운 언더라이팅 기준을 적용하더라도 향후 위험률차 손실의 폭을 줄이기만 할 뿐 위험률차 이익을 기대하기는 어렵다. 따라서, 언더라이팅에 의한 계약 선별 노력 보다 위험률 조정을 우선적으로 시행하는 것이 바람직할 것이다. 결론적으로, 새로운 언더라이팅 기준 수립을 위한 분석대상은 리스크 수준에 따라, 양호한 손해율을 가진 피보험자 집단과 불량한 손해율을 가진 피보험자 집단이 혼재되어 있는 담보군이 적절하다고 판단하였다.

표 2.1. 분석 대상 변수와 범주 단순화 결과

변수	기존 범주	범주 단순화
신출년도	2004~2007	좌동(연도별 구분)
성별	1. 남성 2. 여성	좌동
가입연령대	1. 0~219세, 2. 20~229세, 3. 30~239세, 4. 40~249세, 5. 50~259세, 6. 60~269세	좌동
판매채널	11. 개인채널, 13. 대리점, 15. 전화, 21. 법인, 31. 은행	1. 구채널(11, 13) 2. 신채널(15, 31)
직업군	1. 학자 및 전문직, 2. 행정, 국방, 사무직, 3. 농축산업, 4. 광업, 5. 음식료품, 담배제조업, 6. 섬유, 의류제조업, 7. 목재가공, 제지업, 8. 화학제품 제조업, 9. 비금속제품 제조업, 10. 금속산업, 11. 금속가공업, 12. 기계조립 및 설비업, 13. 전기, 통신 사업, 14. 건설업, 15. 판매, 숙박, 요식업, 16. 운수, 창고업, 17. 문화, 사회, 오락업, 18. 기타	1. 고 위험직업군 (2, 3, 14, 15, 16, 18) 2. 중 위험직업군 (6, 7, 11, 12, 13, 17) 3. 저 위험직업군 (1, 4, 5, 8, 9, 10)
지역	1. 서울, 2. 부산, 3. 대구, 4. 인천, 5. 광주, 6. 대전, 7. 울산, 8. 경기, 9. 강원, 10. 충북, 11. 충남, 12. 전북, 13. 전남, 14. 경북, 15. 경남, 16. 제주	1. 고 위험지역 (1, 2, 8, 13, 15) 2. 중 위험지역 (3, 4, 5, 12, 14) 3. 저 위험지역 (6, 7, 9, 10, 11, 16)
위험보험료	위험보험료 금액	좌동
지급보험금	지급보험금	좌동

본 연구에서는 그러한 특성을 가진 담보군 중 고도장해 담보군(불의의 사고로 말미암아 신체에 장애가 발생하고 장애율이 일정 기준 이상인 경우 계약 시 정한 금액을 지급하는 담보)을 선택하여 2004년부터 2007년까지의 약 222만 건의 계약 자료를 분석하였다. 해당 담보군은 분석 대상 기간 동안의 평균 손해율이 105.4%이고, 위험집단별로 손해율의 범위가 0%~150.9%이므로, 적절한 언더라이팅 기준의 적용에 의해 손해를 개선효과를 기대할 수 있어 본 연구의 목적에 적합하다고 볼 수 있다.

## 2.2. 모형의 설계 및 적합

모형의 설계를 위하여 먼저 언더라이팅 기준을 위한 통계분석에 이용될 수 있는 변수들을 정리하였다. 표 2.1은 경험통계 분석을 바탕으로 고도장해 발생에 영향을 미친다고 판단되는 변수들과 이들 변수에 대한 범주화 내용을 보여 주고 있다. 초기에 고려하였던 변수들 중, 상품유형과 판매지역 본부는 경험통계 분석 과정에서 고도장해 발생빈도에 뚜렷한 차이를 나타내지 않았기 때문에 분석 대상 변수에서 제외시켰다. 또한, 판매채널의 경우 법인채널은 인수기준과 요율 체계가 다른 판매채널과 다르기 때문에 분석에서 제외하였다. 통계적 모형에 포함되는 설명변수의 개수를 최소화하기 위하여, 해당 담보의 보험금이 정액형으로 설계되었다는 점을 감안하여 보험금 지급 사유인 고도장해 발생 빈도의 차이에 따라 적절한 범주를 단순화 하였다. 채널별 고도장해 발생 빈도를 기준으로 전통적인 보험판매 수단인 개인채널과 대리점은 구채널로, 비교적 최근에 도입된 텔레마케팅과 방카슈랑스 채널은 신채널로 구분하였다. 직업군과 지역의 경우는 기존 범주에 해당되는 집단별 사고 발생 빈도를 산출한 뒤 뚜렷한 차이를 나타낼 수 있는 3개의 위험군으로 범주를 단순화 하였다.

표 2.1에서 고도장해 위험수준에 영향을 미치는 변수 중 년도와 성별 및 가입연령은 위험률 산출, 즉 보험가격을 결정하는 요소이다. 따라서, 이들 변수를 제외한 나머지 변수들 중 위험수준에 영향을 미치는 변수들에 의해 결정되는 빈도의 상대도를 이용하여 위험수준의 순위를 평가하여 전략적 언더라이팅 기준을 수립할 수 있을 것이다.

스코어링 시스템 구축에 활용될 수 있는 통계모형 가운데 일반화선형모형(*generalized linear model*)은 반응변수에 영향을 미치는 여러 가지 요소를 다양한 구조로 표현하여 모형화 할 수 있어 그 활용도가 높은 통계모형이다. 이러한 일반화선형모형은 이미 금융 분야와 관련한 여러 연구에 광범위하게 이용되는 방법 중 하나이다. 특히, 보험 분야에서도 다양한 작용 방안들이 제시되어 있다. Haberman과 Renshaw (1996)은 일반화선형모형이 보험수리에 적용될 수 있는 사례에 대하여 소개하고 있다. Murphy 등 (2000)은 일반화선형모형을 동적 보험료 산출에 적용하고 있다. 또한 Kwon과 Jones (2006)는 다양한 사망률 위험요소를 이용하는 일반화선형모형과 마르코프(Markov)모형을 결합하여 위험률을 도출하고 생명보험과 연금에 적용하는 방안을 제시하였다.

본 연구에서는 통계적으로 유의한 설명변수를 이용하여 빈도의 상대도를 결정하기 위하여 일반화선형모형을 적용하였다. 빈도는 연간 고도장해 발생확률로 표현되므로 로지스틱 회귀모형을 적용할 수 있으나, 연령변수가 산출년도에서의 정확한 연령 대신 가입연령대를 사용하고 있으므로 정확한 실제 연령에서의 위험률을 도출하기는 어렵다. 그러나, 최종적으로 언더라이팅 기준에 의한 효과를 이해하기 위해서는 서로 다른 위험집단 간 오즈비(odds ratio) 보다는 빈도의 상대도를 활용해야 하므로, 발생빈도에 대해 이항분포 모형을 가정하되, 연결함수(link function)는 log 연결함수를 사용하여 모형을 다음과 같이 설계하였다.

$$\log \pi_j = \underline{\beta}^T \underline{Z}_j \quad (2.1)$$

$\pi_j$ 는 위험집단  $j$ 의 연간 고도장해 발생확률을 의미하고,  $\underline{Z}_j$ 는 유의한 설명변수와 교호변수의 값으로 구성된 열벡터로 각 위험집단별 위험수준을 결정짓는 요소이다. 여기서  $\pi_j$ 는 가입연령을 기준으로 한 발생확률임에 유의해야 한다. 또한  $j$ 의 값은 위험집단의 분류 개수만큼의 값을 가질 수 있다.  $\underline{\beta}$ 는  $\underline{Z}_j$ 의 각 성분에 대응되는 계수로 구성된 열벡터로 각 설명변수의 영향의 방향과 정도를 나타낸다. 설명변수를 구성하는 열벡터의 요소에 대한 기호와 변수값에 대한 설명은 표 2.2에 설명되어 있다. 위험집단별 위험보험료의 규모를 반영하기 위해 위험보험료를 가중치 변수로 이용하였다. 유의수준 1% 하에서 단계별 변수선택(stepwise variable selection) 과정을 적용하여 유의한 설명변수 및 교호변수를 찾고 각 변수에 대한 계수를 추정하여 다음의 표 2.3과 같은 결과를 얻을 수 있었다. 다양한 변수들 간의 교호효과를 고려함에 있어서는 특정한 변수들 간의 교호효과가 통계적으로 유의하다고 하더라도 모형의 적합도를 크게 향상시키지 않는 경우에는 모형의 단순화를 위하여 선택변수에서 제외하였다. 일반화선형모형에 대한 단계별 변수 선택 과정에 대해서는 McCullagh와 Nelder (1989)에 자세한 설명이 제시되어 있다. 변수선택 과정 마지막 단계의 모형에 포함된 설명변수와 그 교호작용에 대한 계수추정치에 대한 결과는 표 2.3에 제시하였다. 최종모형에서의  $\underline{Z}_j$ 와  $\underline{\beta}$ 의 추정치는 다음과 같다.

$$\begin{aligned} \underline{Z}_j^T &= (1, Z_{j1}, Z_{j2}, Z_{j2}^2, Z_{j2}^3, Z_{j3}, Z_{j4}, Z_{j5}, Z_{j6}, Z_{j7}, Z_{j8}, Z_{j1}Z_{j2}) \\ \underline{\beta}^T &= (-675.4190, 0.3366, -3.4121, 0.9489, -0.0807, -0.5525, -1.9651, 1.5200, -0.8204, \\ &\quad 0.3073, -0.4565, 0.1591) \end{aligned}$$

추정된 계수값을 통해 언더라이팅 기준에 반영될 수 있는 판매채널, 위험직업군, 위험지역의 효과를 비교해 보면, 판매채널, 직업군, 지역 순으로 고도장해 발생 빈도에 미치는 영향이 크게 나타남을 파악할 수 있다.

표 2.2. 유의한 설명변수와 기호 및 변수값

설명변수	변수값
산출년도( $Z_{j1}$ )	산출년도
가입 연령대( $Z_{j2}$ )	0~19세 = 1, 20~29세 = 2, 30~39세 = 3, 40~49세 = 4, 50~59세 = 5, 60~69세 = 6
성별( $Z_{j3}$ )	남성 = 0, 여성 = 1
판매채널( $Z_{j4}$ )	구채널 = 0, 신채널 = 1
고 위험직업군( $Z_{j5}$ )	고 위험직업군 = 1, 고 위험직업군이외 = 0
저 위험직업군( $Z_{j6}$ )	저 위험직업군 = 1, 저 위험직업군이외 = 0
고 위험지역( $Z_{j7}$ )	고 위험지역 = 1, 고 위험지역이외 = 0
저 위험지역( $Z_{j8}$ )	저 위험지역 = 1, 저 위험지역이외 = 0

표 2.3. 최종 모형 적합 결과

변수	계수추정치	표준오차	p-value
절편	-675.4190	0.0124	< 0.0001
산출년도	0.3366	< 0.0001	< 0.0001
가입연령대	-3.4121	0.0001	< 0.0001
가입연령대 <sup>2</sup>	0.9489	< 0.0001	< 0.0001
가입연령대 <sup>3</sup>	-0.0807	< 0.0001	< 0.0001
성별	-0.5525	< 0.0001	< 0.0001
가입연령대×성별	0.1591	< 0.0001	< 0.0001
판매채널	-1.9651	0.0002	< 0.0001
고 위험직업군	1.5200	0.0001	< 0.0001
저 위험직업군	-0.8204	0.0001	< 0.0001
고 위험지역	0.3073	< 0.0001	< 0.0001
저 위험지역	-0.4565	< 0.0001	< 0.0001

### 3. 보험수리적 분석

2.2절의 분석결과에 의하면, 앞서 언급한 요율에 반영되는 요소인 산출년도, 연령, 성별 이외의 유의한 설명변수로 판매채널과 직업군 및 지역이 있다. 판매채널과 직업군 및 지역 변수가 취할 수 있는 값들에 따라 구분되는 위험등급 조합의 개수는 18개이다. 이는 동일한 요율을 적용받는 같은 성별, 연령집단 내의 계약을 18개 등급으로 세분화 할 수 있다는 것을 의미한다. 18개의 등급조합 중 위험수준이 높은 집단의 비중을 언더라이팅 과정을 통해 계약 포트폴리오로부터 줄이거나 제거할 수 있다면 위험율차 이익을 확대하여 수익성을 개선할 수 있을 것이다.

언더라이팅 기준의 설계를 위하여, 위험집단별 위험보험료 규모를 고려하여 각 위험집단 별 점수를 부여할 수 있다. 우선 각 위험집단을 위험수준이 높은 순서대로 순위를 부여하고 위험수준이  $j$ 번째로 높은 집단의 위험보험료 비중을  $p_j(j = 1, 2, \dots, 18)$ 라 하자. 그러면, 위험수준이  $j$ 번째로 높은 집단의 점수를  $s_j = 100 \times \sum_{i=1}^j p_i$ 로 부여할 수 있다. 즉, 부여된 점수가 낮을수록 위험도가 상대적으로 높음을 의미하게 된다. 또한 언더라이팅 효과를 분석하기 위해서는 각 위험집단 간 연간 고도장해 발생률에 대한 상대도가 필요하다. 앞서 정의한 변수의 정의에 따라  $Z_{j4} = Z_{j5} = Z_{j6} = Z_{j7} = Z_{j8} = 0$ 인 집단 즉, 구채널, 중 위험직업군, 중 위험지역에 해당되는 피보험자 집단을 기준집단으로 하면 위험수준이  $j$ 번째로 높은 위험수준의 상대도 값  $r_j$ 는

$$\exp(-1.9651Z_{j4} + 1.5200Z_{j5} - 0.8204Z_{j6} + 0.3073Z_{j7} - 0.4565Z_{j8})$$

표 3.1. 위험집단별 위험보험료 비중, 위험 점수 및 위험률 상대도

순위	$(Z_{j4}, Z_{j5}, Z_{j6}, Z_{j7}, Z_{j8})$	$p_j$	$s_j$	$r_j$
1	(0, 1, 0, 1, 0)	0.0088	0.9	6.2171
2	(0, 1, 0, 0, 0)	0.0126	2.1	4.5722
3	(0, 1, 0, 0, 1)	0.1257	14.7	2.8965
4	(0, 0, 0, 1, 0)	0.0138	16.1	1.3597
5	(0, 0, 0, 0, 0)	0.0220	18.3	1.0000
6	(1, 1, 0, 1, 0)	0.0001	18.3	0.8713
7	(1, 1, 0, 0, 0)	0.0001	18.3	0.6408
8	(0, 0, 0, 0, 1)	0.1968	38.0	0.6335
9	(0, 0, 1, 1, 0)	0.0348	41.5	0.5986
10	(0, 0, 1, 0, 0)	0.0579	47.2	0.4403
11	(1, 1, 0, 0, 1)	0.0009	47.3	0.4059
12	(0, 0, 1, 0, 1)	0.5199	99.3	0.2789
13	(1, 0, 0, 1, 0)	0.0001	99.3	0.1906
14	(1, 0, 0, 0, 0)	0.0001	99.3	0.1401
15	(1, 0, 0, 0, 1)	0.0014	99.5	0.0888
16	(1, 0, 1, 1, 0)	0.0003	99.5	0.0839
17	(1, 0, 1, 0, 0)	0.0003	99.5	0.0617
18	(1, 0, 1, 0, 1)	0.0046	100.0	0.0391

로 표현할 수 있다. 제 2장에서 적합한 모형을 근거로 18개 집단에 대한 위험수준의 변수값과  $p_j$ ,  $s_j$ ,  $r_j$  값을 위험수준이 높은 순서대로 정리한 결과는 표 3.1과 같다. 앞서 언급한 바와 같이 고도장해 담보군의 경우 지난 4년간 해당 담보의 손해율은 105.4%로 위험률차 손실을 기록하고 있다. 전체적으로는 위험수준이 3, 8, 12번째로 높은 위험집단의 위험보험료 비중이 집중되어(약 84.2%) 있는 것을 볼 수 있다. 특히 위험수준이 12번째로 높은 위험집단(구채널로 판매되었고, 저 위험직업군, 저 위험지역에 해당되는 집단)의 위험보험료 비중이 전체 대비 절반정도를(약 52.0%) 차지하고 있는데, 이는 현재 이 회사의 계약 포트폴리오 품질이 전반적으로 양호함을 보여주는 증거이다. 하지만 위험수준이 3번째로 높은 집단(구채널로 판매되었고, 고 위험직업군, 저 위험지역에 속한 계약집단)의 비중이 다른 집단에 비하여 상대적으로 크기 때문에(약 12.6%) 해당 집단에 대한 언더라이팅 강화를 통해 수익성을 개선할 수 있을 것이라 기대할 수 있다.

전략적 언더라이팅은 위험수준이 높은 집단에 대하여 인수를 제한하거나 거절하여, 평균적인 위험수준을 낮게 만드는 것이다. 언더라이팅 강도의 선택을 위해서는 언더라이팅 기준의 적용에 따른 보험회사의 손해율(지급보험금/위험보험료)의 개선효과를 추정하는 과정이 선행되어야 한다. 현재 고도장해 담보군의 손해율을  $\alpha\%$ 라 하자. 또한, 향후 신계약에 대하여 위험집단별 위험보험료 비중과 언더라이팅 기준이 성별, 연령별로 동일하여, 향후 신계약 포트폴리오의 예상 손해율이 과거와 같은 수준으로 일정 기간 지속될 것이라 가정하자. 위험수준이 일정 기준 이상이 되는 위험집단을 인수 제한하거나 거절할 경우 손해율의 구성요소인 위험보험료 및 지급보험금이 모두 영향을 받게 될 것이다. 즉, 인수 제한과 거절로 인해 향후 신계약 수입 위험 보험료 수준이 기존보다 줄어들 것이고, 위험 수준이 높은 집단을 제한 또는 배제하기 때문에 향후 신계약에 대해서는 지급보험금 수준도 기존보다 줄어들게 될 것이다. 결국, 위험보험료의 조정이 없을 경우 언더라이팅 효과에 따른 향후 신계약에 대한 예상 손해율 수준은

$$\alpha \times \frac{(1 - \text{지급보험금 감소율})}{(1 - \text{수입보험료감소율})} \% \quad (3.1)$$

로 표현된다.

언더라이팅 기준을 위험수준이 높은 상위  $x\%$ 의 피보험자 집단, 즉 위험점수가  $100 - x$ 점 이상인 집단의 인수율을 기존의  $100\%$ 에서  $y\%$ 로 조정하였다고 가정하면, 수입보험료의 감소율은  $0.01x(1 - 0.01y)$ 가 될 것이다. 지급보험금의 감소비율은 계약자 집단의 평균위험률 감소 수준으로 추정할 수 있는데, 언더라이팅 기준의 변경 전과 변경 후의 평균위험률 수준의 추정을 통해 다음과 같이 계산할 수 있다. 특정연령, 성별에 대한 18개의 각 위험집단의 위험률을  $q_j$ 라 하자.  $q_j$ 는 모형의 종속변수인  $\pi_j$ 와 달리 해당년도의 정확한 실제연령을 반영한 위험률이다. 모형에서 적용한 계약연령의 빈도효과는 실제 연령의 빈도효과와 비슷할 것이라 가정하면, 연령 이외의 다른 설명변수의 효과는 모형에서 적합한 결과를 그대로 이용할 수 있다. 또한,  $j$ 는 표 3.1의 순위를 나타낸다. 그러면, 각 위험집단의 상대도의 기준이 되는 5번 집단의 위험률을  $q_5$ 라 할 때, 각 위험집단의 위험률은  $q_j = r_j q_5$  ( $j = 1, 2, \dots, 18$ )로 표현될 것이다. 언더라이팅 기준 변경 전의 평균 위험률 수준은 위험보험료 비중을 가중치로 한 위험률의 평균값으로 계산될 수 있으며,

$$\sum_{i=1}^{18} q_i r_i = q_5 \sum_{i=1}^{18} p_i r_i \tag{3.2}$$

로 표현된다. 식 (3.2)와 마찬가지로, 언더라이팅 기준 변경 후의 평균 위험률 수준은

$$q_5 \sum_{i=1}^{18} p'_i r_i \tag{3.3}$$

로 표현되고  $p'_j$ 은 위험보험료 비중의 변화에 따라서

$$p'_j = \frac{p_j(1 - 0.01y)I(s_j \leq x) + p_j I(s_j > x)}{\sum_{i=1}^{18} \{p_i(1 - 0.01y)I(s_i \leq x) + p_i I(s_i > x)\}}, \quad (j = 1, 2, \dots, 18) \tag{3.4}$$

로 표현된다. 따라서, 지급보험금의 감소율은

$$1 - \frac{q_5 \sum_{i=1}^{18} p_i r_i}{q_5 \sum_{i=1}^{18} p'_i r_i} = 1 - \frac{\sum_{i=1}^{18} p_i r_i}{\sum_{i=1}^{18} p'_i r_i} \tag{3.5}$$

로 나타낼 수 있고, 최종적으로 새로운 언더라이팅 기준 적용에 의한 손해율의 감소율은

$$1 - \frac{\sum_{i=1}^{18} p_i r_i / \sum_{i=1}^{18} p'_i r_i}{1 - 0.01x(1 - 0.01y)} \tag{3.6}$$

로 표현될 수 있다. 최종적으로 식 (3.5)와 (3.6)의 결과를 식 (3.1)에 대입하면 언더라이팅 기준 변경 후 예상 손해율을 도출할 수 있다. 이러한 결과는 각 연령과 성별에 해당하는 18개의 위험집단의 위험보험료 비중이 비슷하다고 가정된 경우이다. 보다 정밀한 계약자의 분포를 사용할 수 있다면, 보다 정확하게 손해율 개선효과를 추정할 수 있을 것이다.

언더라이팅 기준은 보험상품에 대한 규정이나 사회적 인식 및 회사의 장기 전략에 따라서 좌우될 수 있다. 위험수준이 높은 집단에 대한 인수 배제를 통해 언더라이팅 효과를 극대화할 수 있으나, 일반적으로 적용되는 규정에 따르면 위험률 결정요소 이외의 요소를 이용하여 가입을 거절할 수 없는 경우가 많다.

표 3.2. 언더라이팅 기준에 따른 효과 분석 결과 (단위: %)

언더라이팅 기준		위험보험료 감소율	지급보험금 감소율	손해율 감소율	예상손해율
$x$	$y$	-	-	-	-
0	0	0	0	0	105.4
15	0	19.3	53.5	43.1	60.0
	50	9.1	24.1	16.4	88.1
	70	5.5	13.9	8.9	96.0
20	0	14.7	49.7	41.1	62.1
	50	7.3	22.9	16.8	87.7
	70	4.4	13.3	9.3	95.6

따라서, 현실적인 제약을 반영한 다양한 언더라이팅 가능성과 전략에 대하여 언더라이팅 효과를 분석해 보는 것이 필요하다. 이러한 필요성을 감안하여 인수기준을 나타내는  $x$ 값과 인수율  $y$ 값의 몇 가지 조합에 대한 언더라이팅 효과를 분석하여 보았는데 그 결과는 표 3.2와 같다. 예상한 바와 같이 동일한  $x$ 값에 대하여는  $y$ 의 값이 증가할수록 손해를 개선효과가 증가함을 볼 수 있으나, 동일한  $y$ 값에 대하여는 손해를 개선효과가 동일하게 나타나지 않음을 파악할 수 있다. 따라서, 가능한  $x, y$ 값의 조합에 대하여 손해율의 개선효과를 파악하여 회사의 상황과 전략에 가장 적합한 언더라이팅 기준을 정하는 것이 중요하다. 또한, 앞서 제시한 언더라이팅 방법의 적용과 함께 위험수준이 낮은 집단에 대한 마케팅을 적극적으로 추진하여 해당 집단의 계약건수를 증가시킬 수 있다면 손해를 개선효과는 더욱 크게 나타날 것이라 기대할 수 있다.

#### 4. 결론 및 향후 연구과제

본 연구에서는 기존의 단순한 방법을 이용한 언더라이팅 스코어링 시스템을 개선할 수 있는 통계모형의 활용 방안에 대하여 살펴보았다. 일반화선형모형을 실제 자료에 적용하여 현재 이용되지 않는 추가적인 위험요소를 반영하고 각 위험요소가 보험금 지급발생에 영향을 미치는 효과를 종합적으로 분석할 수 있는 모형을 구체적으로 설계하였다. 또한, 새로운 언더라이팅 기준에 따른 손해를 개선효과를 분석하는 방안을 제시하였다. 이러한 스코어링 방법의 적용을 통해 일관성 있고 보다 객관적인 심사기준의 적용이 가능할 것이다. 또한, 이러한 기준을 시스템적으로 적용함으로써 기존의 심사과정을 단순화하고 자동화하여 비용절감의 효과를 추가적으로 기대할 수 있을 것이다.

향후 본 연구와 관련한 추가 또는 보완 연구를 다음과 같이 생각해 볼 수 있다. 본 연구에서는 정확한 연령에 대한 자료를 사용할 수 없었기 때문에, 연령별 정확한 보험료를 산출방법에 의한 위험집단 간 위험정도 수준의 비교대신, 위험요소를 반영하여 산출한 상대도를 이용하여 위험정도를 비교하고 손해를 효과를 분석하였다. 세분화된 연령 자료를 사용할 수 있다면 로지스틱 회귀모형 등을 이용한 보다 정밀한 분석도 적용할 수 있을 것이고, 손해를 효과에 대한 분석에서도 정확도를 향상시킬 수 있을 것이다. 또한, 본 연구에서 이용했던 설명변수 외에도 해외의 연구 등에서 밝혀진 중요한 사회인구학적 변수와 행태적 변수들에 대한 데이터도 추가로 확보하여 분석에 사용할 수 있다면, 보다 정밀한 언더라이팅 시스템을 구축하는데 도움이 될 것이라 생각한다. 따라서, 생명보험 계약과 관련한 신뢰할 수 있으며 보다 세밀하고 다양한 데이터의 확보는 언더라이팅 방법론의 개선을 위해 지속적인 노력이 필요한 기본적 과제인 것이다.

본 연구에서는 생명보험 담보에 대하여 언더라이팅 개선 방안을 제시하였는데, 실제 손실금액을 보상해주는 손해보험 담보에 대하여 보다 적합성이 높은 모형을 탐색하는 것도 중요한 향후 연구과제이다. 추



가적으로, 일반화선형모형과 더불어 데이터마이닝 등의 보완적 기법의 활용방안에 대한 검토도 모형을 더욱 발전시키는 데 도움이 될 것이다. 본 논문의 저자들은 수리적 언더라이팅의 개선을 위해 일반화선형모형과 데이터마이닝 기법을 적절히 결합하는 방법에 대해 앞으로 연구할 예정이다. 또한 수리적 모형화의 결과를 현업에 적용하기 위한 시스템 개발 및 도입 방법에 대한 검토도 중요한 현실적 과제라 볼 수 있다.

## 참고문헌

- Atkinson, D. B. and Dallas, J. W. (2000). *Life Insurance Products and Finance*, Society of Actuaries.
- Bickley, M. C. (2007). *Life and Health Insurance Underwriting*, LOMA
- Haberman, S. and Renshaw, A. E. (1996). Generalized linear models and actuarial science, *The Statistician*, **45**, 407-436.
- Kwon, H. S. and Jones, B. L. (2006). The impact of the determinants of mortality on life insurance and annuities, *Insurance: Mathematics and Economics*, **38**, 271-288.
- Lewis, E. M. (1992). *An Introduction to Credit Scoring*, Fair, Isaan and Company.
- McCullagh, P. and Nelder, J. A. (1989). *Generalized Linear Models*, Chapman & Hall/CRC, London.
- Murphy, K. P., Brockman, M. J. and Lee, P. K. W. (2000). Using Generalized Linear Models to Build Dynamic Pricing Systems, *Casualty Actuarial Forum*, Winter 2000.
- Society of Actuaries (2005). Report of the Society of Actuaries Preferred Underwriting Survey Subcommittee, March 2005.

# A Study on the Scoring Method for the Insurance Underwriting Using Generalized Linear Model

Chang Soo Lee<sup>1</sup> · Hyuk-Sung Kwon<sup>2</sup> · Dong-Kwang Kim<sup>3</sup>

<sup>1</sup>Department of Statistics and Actuarial Science, Soongsil University;

<sup>2</sup>Department of Statistics and Actuarial Science, Soongsil University;

<sup>3</sup>Department of Statistics and Actuarial Science, Soongsil University

(Received March 2009; accepted March 2009)

---

## Abstract

Underwriting is the first step for the administration of an insurance contract, which may result in stable profitability or unexpected loss for insurance company. Adequacy of underwriting criteria determines underwriting result. Generally, quantitative scoring system is used for underwriting. Method of evaluating risk for the scoring system is summing up scores for risk factors of a potential policyholder in consideration. Scores for each risk factor is predetermined. Current business environment for insurance companies makes underwriting profit more important, which means that insurance companies need more efficient underwriting method. This study suggests a reasonable approach to estimate risk relativities based on generalized linear model. Real data were used to quantify risk levels of groups of insureds for the design of underwriting model. Finally, effects in business volume and profitability of reflecting estimated underwriting scoring system are explained.

**Keywords:** Underwriting, scoring method, generalized linear model, life insurance.

---

---

This work was supported by the Soongsil University Research Fund.

<sup>2</sup>Corresponding author: Assistant Professor, Department of Statistics and Actuarial Science, Soongsil University, Seoul 156-743, Korea. E-mail: hskwon@ssu.ac.kr