

# An Analysis on the Gender Differences in the Level of Accident Risk using Generalized Linear and Heckman Methods

DaeHwan Kim<sup>a</sup> · HwaGyu Park<sup>b,1</sup>

<sup>a</sup>Department of Economics, Dong-A University

<sup>b</sup>Department of Health Administration and Management, Soonchunhyang University

(Received December 16, 2013; Revised February 6, 2014; Accepted February 6, 2014)

---

## Abstract

Women's roles have changed substantially in economically developed countries; subsequently, the ratio of female drivers has also increased. In such countries, there has been considerable interest in assessing gender differences in vehicle accident risks and reasons to explain the gender differences. This study investigates the gender differences in vehicle accident risk based on 500,000 drivers randomly selected from a population sample. A Heckman model is used for accident damage and a negative binomial model is used for the accident frequency. Empirical results show that male drivers are 8.3% riskier than female drivers in terms of accident damage; however, female drivers are 113% riskier than male drivers in term of accident frequency. We can implement more practical policies to reduce vehicle accidents if we can understand the reasons for the gender differences.

Keywords: Female drivers, vehicle accidents, Heckman model, negative binomial model.

---

## 1. 서론

경제성장에 기반한 소득증가의 원인으로 인해 우리나라 자동차등록대수도 빠르게 증가해 왔다. 20년 전인 1992년도에는 자동차등록대수가 520만대에 불과하였으나 그 후 10년 뒤인 2002년에는 1,400만대로 증가하였고 현재(2012년 12월 말 기준) 1,840만대에 달하고 있는 상황이다 (Ministry of Land, Infrastructures and Transport, 2012). 자동차대수 증가보다 더욱 눈에 띄는 것은 여성운전자의 증가이다. 주요 선진국에서처럼 여성의 사회적 참여 및 지위가 상승함에 따라 여성운전자의 비중은 자동차등록대수보다 더욱 빠르게 증가하고 있다. 2012년 말 기준으로 여성운전자는 374만 명에 달하며 (Ministry of Land, Infrastructures and Transport, 2012), 수도권외의 경우 운행되는 승용차의 여성운전자 비중이 1997년 13.4%에서 2010년 27.6%로 두 배 이상 증가하였다 (Ryu 등, 2013).

자동차 수가 증가함에 따라 자동차사고로 인한 사망자 수도 적지 않다. 2012년에 교통사고로 인한 사망자수는 6,502명으로 매일 17.8명이 사망하고 있다 (National Statistical Office, 2013). 우리나라가 경

---

This research was supported by a Dong-A University Research Fund.

<sup>1</sup>Corresponding author: Professor, Department of Health Administration and Management, Soonchunhyang University, Chungnam 336-745, Korea. E-mail: [hkpark1@sch.ac.kr](mailto:hkpark1@sch.ac.kr)

험하고 있는 것처럼 주요국에서도 여성의 사회적 참여와 지위가 상승함에 따라 여성운전자가 빠르게 증가하였으며, 자동차사고로 인한 사망이 자국의 주요 사망원인이기 때문에 성별에 따른 자동차사고 위험도를 평가하고 교통사고 예방책을 마련하는 연구가 활발히 진행되어 왔다 (Waldron 등, 2005).

OECD 국가 중 가장 높은 자동차사고 사망률을 고려하여 우리나라 정부는 그동안 사후관리 중심이던 자동차손해배상보장 기능의 한계를 극복하기 위해 자동차사고 예방사업을 마련해 내년부터 시행할 예정이며, 국가교통안전시행계획을 마련한 상황이다 (Ministry of Land, Infrastructures and Transport, 2013).

우리나라 교통사고 위험도가 성별에 따라 유의하게 다르다면 좀 더 실효성 높은 교통사고 정책이 마련될 수 있을 것이다. 그럼에도 불구하고 해외 주요국과 달리 우리나라는 성별 자동차사고 위험도를 평가한 연구가 매우 미진한 상황이다.

이에 본 연구는 성별 자동차사고 위험도를 분석하였으며, 특히 자동차사고 위험도를 심도와 빈도로 구분하여 면밀한 분석을 시도하였다. 무엇보다 여성 운전자 수가 빠르게 증가하고 있는 현황을 반영하여 남성 대비 여성운전자의 교통사고 빈도와 교통사고 심도를 중심으로 실증분석하였다.

본 연구의 구성은 다음과 같다. 2장에서는 관련 선행연구를 살펴보고, 3장에서는 실증분석 모형과 데이터를 소개한다. 4장에서는 실증분석 결과에 대해 논의하고, 마지막으로 5장은 결론과 함께 본 연구의 한계점에 대해 논의한다.

## 2. 선행연구

반면 해외 주요국의 경우 성별 자동차사고 위험도 차이가 존재하는지 그리고 왜 존재하는지에 대한 원인 규명을 위해 다양한 연구가 진행되어 왔다. Al-Balbissi (2003)은 자동차사고로 인한 피해액을 기준으로 성별 자동차사고 위험도를 분석한 결과 남성이 여성보다 자동차사고 위험도가 높음을 보였다. Massie와 Campbell (1993)도 치명상에 이를 정도의 교통사고 심도를 기준으로 성별 자동차사고 위험도를 분석한 결과 75세 이상 고령층을 제외하고 대부분의 연령대에서 여성보다 남성의 위험도가 높은 것으로 나타났다. 이렇듯 교통사고 피해액 및 사망률 등과 같은 심도를 기준으로 성별 자동차사고 위험도를 평가했을 때 여성보다는 남성의 위험도가 높다는 것이 일반적인 결론이며 다른 대부분의 연구에서도 동일한 결과가 도출되고 있다.

반면 교통사고 횟수를 기준으로 한 성별 교통사고 위험도는 다소 다른 양상을 보인다. Storie (1977)은 교통사고 심도를 기준으로 남성이 여성보다 교통사고 위험도가 높지만, 빈도를 기준으로 여성이 남성보다 교통사고 위험도가 높다고 주장하였다. 특정한 운전상황에서는 오히려 남성이 여성보다 교통사고를 유발할 가능성이 높다는 연구결과도 존재한다. Waylen과 McKenna (2002)는 굵은 도로에서, 어두운 상황에서, 그리고 추월상황에서는 남성이 여성보다 교통사고 빈도가 높음을 보였다.

한편 성별 간 자동차사고 위험도 차이의 원인을 규명하려는 연구도 활발한 상황이다. 자동차사고 심도를 기준으로 남성의 높은 자동차사고 위험도는 남성의 공격적인 운전행태 때문이라는 주장이 일반적이다. Storie (1977)는 특히 남성운전자가 여성운전자에 비해 과속을 하기 때문에 사고 시 피해액 및 사망률 등이 높다고 주장하였다. Lajunen과 Parker (2001)가 운전자 스스로 주관적인 공격성을 평가하도록 한 결과, 남성운전자가 여성운전자보다 운전 시 공격성이 훨씬 높은 것으로 나타났다. Zuckerman (1996)은 테스토스테론(testosterone) 호르몬이 공격성에 영향을 줄 수 있기 때문에 여성보다는 남성이 공격적인 운전을 하고 따라서 사고피해액과 사망률이 높다고 진단하였다. 또한 대표적인 남성호르몬인 테스토스테론은 젊은 남성에게 많이 분비되기 때문에 동일한 남성이더라도 젊은 운전자의 자동차사고

위험도가 높다고 주장하였다.

Doob와 Gross (1968)는 초록색 신호등이 켜졌음에도 불구하고 앞차가 출발하지 않을 때 남성과 여성의 경적을 울리는 시간을 계산함으로써 성별 운전공격성을 측정하였는데, 남성이 여성보다 3배나 빨리 경적을 울릴 정도로 공격적인 성향이 높다고 주장하였다. Baker 등 (1992)는 여성보다 높은 남성의 자동차운전 자살률이 남성의 높은 자동차사고 위험도를 어느 정도 설명할 수 있을 것이라고 주장하였다.

성별 교통사고 빈도를 설명하는 연구는 상대적으로 많지 않지만 Storie (1977)는 여성이 남성보다 지각 판단 능력이 상대적으로 취약하기 때문에 교통사고를 많이 발생시킨다고 주장하였다.

물론 여성보다는 남성의 자동차사고 위험도가 여전히 높기는 하지만 성별 자동차사고 위험도가 점차 감소하고 있다는 연구도 이루어지고 있다 (Waldron 등, 2005). 과거에는 남성이 여성보다 자동차사고 위험도가 높았지만 여성의 노동시장 참여를 비롯하여 사회적 지위가 증가함에 따라 성별 역할 간격이 감소하여 성별 자동차사고 위험도도 점차 감소하고 있다고 주장하였다.

이렇게 해외의 경우 성별에 따라 자동차사고 위험도에 차이가 있는지 그리고 그러한 차이가 왜 발생하는지에 대한 연구가 활발한 반면 국내의 경우 관련 연구가 부재한 상황이다. 또한 성별에 따른 자동차사고위험도를 분석한 해외 선행연구의 경우 교통사고 빈도의 경우 일반화선형모형(Generalized Linear Method) 중 포아송(Poisson)모형이 활용되었고, 교통사고 심도의 경우 최소자승법(Ordinary least square)이 활용되었는데 실증분석 모형으로 미흡한 측면이 존재한다. 이에 본 연구에서는 우리나라의 경우에도 성별에 따라 자동차사고 위험도가 존재하는지 여부를 분석하고, 해외 선행연구에서 활용된 실증분석모형보다 개선된 방법론을 활용하였다.

### 3. 실증분석 자료 및 모델

연구에 사용된 데이터는 보험개발원의 자동차보험 자료이다. 자동차보험 담보 중 자기신체사고 및 자차보험과 같은 담보의 경우 운전자가 선택적으로 가입할 수 있는 있다는 특성을 고려하여 분석대상에서 제외하고, 대신 운전을 하려면 반드시 가입해야 하는 대인배상I과 대물담보에서 발생한 교통사고 심도와 빈도를 종속변수로 설정하였다.

보험개발원은 자동차보험에 가입되어 있는 모든 운전자의 사고기록을 보유하고 있는데 2006년 국내 모든 운전자 중 영업용 자동차를 제외한 개인용 자동차 중 약 50만개의 샘플을 무작위로 추출하여 실증분석에 활용하였다. 그 중 결측치가 발생한 샘플을 누락시켜 488,139개의 샘플이 실증분석에 활용되었다. 현재까지 해외연구에서도 본 연구에서 사용된 데이터처럼 모집단으로부터 샘플을 추출한 연구는 부재한 상황이다.

Table 3.1은 실증분석에 활용된 변수명과 정의를 보여준다. 독립변수로 교통사고 위험도에 영향을 줄 수 있는 운전자의 성, 연령, 차종, 운전경력, 그리고 지역을 활용하였다. 지역의 경우에는 서울, 인천, 경기도와 각 광역시, 그리고 각 도지역 단위로 구분하였다. Table 3.1의 지역1은 서울, 지역2는 부산, 지역3은 경기, 지역4는 강원, 지역5는 충북, 지역6은 충남, 지역7은 전북, 지역8은 전남, 지역9는 경북, 지역10은 경남, 지역11은 제주, 지역12는 대구, 지역13은 인천, 지역14는 광주, 지역15는 대전, 그리고 지역16은 울산을 의미한다.

전술하였듯이 본 연구에서 활용된 종속변수는 자동차사고 심도와 빈도이다. 각 종속변수의 속성에 따라 적용한 분석모형도 선행연구와 차별화하였다. 먼저 자동차사고 빈도의 경우 자동차사고가 발생하지 않은 비중이 89.9%, 1번 사고를 낸 경우가 6.3%, 2번 사고를 낸 경우가 3.3% 등으로 정규분포(normal distribution)가 아닌 지수분포(exponential distribution)이다. 종속변수가 지수분포인 경우 실증분석

Table 3.1. Name and definition of variables

변수		정의	
종속변수	사고심도	교통사고 피해액(보험금)의 로그값	
	사고빈도	교통사고 횟수	
	성별	남성이면 1, 여성이면 0	
독립변수	연령	운전자 연령	
	운전 경력	경력1	운전경력이 1년 미만이면 1, 아니면 0
		경력2	운전경력이 1년 이상 2년 미만이면 1, 아니면 0
		경력3	운전경력이 2년 이상 3년 미만이면 1, 아니면 0
		경력4	운전경력이 3년 이상 4년 미만이면 1, 아니면 0
		경력5	운전경력이 4년 이상 5년 미만이면 1, 아니면 0
		경력6	운전경력이 5년 이상이면 1, 아니면 0
	차종	경차	배기량이 1000cc 미만 경차이면 1, 아니면 0
		소형	배기량이 1000~1599cc이면 1, 아니면 0
		중형	배기량이 1600~1999cc이면 1, 아니면 0
		대형	배기량이 2000cc 이상이면 1, 아니면 0
		벤 SUV	자동차가 벤(Van)이면 1, 아니면 0 자동차가 SUV이면 1, 아니면 0
	지역	지역1	자동차 등록 지역이 1이면 1, 아니면 0
지역2		자동차 등록 지역이 2이면 1, 아니면 0	
지역3		자동차 등록 지역이 3이면 1, 아니면 0	
지역4		자동차 등록 지역이 4이면 1, 아니면 0	
지역5		자동차 등록 지역이 5이면 1, 아니면 0	
지역6		자동차 등록 지역이 6이면 1, 아니면 0	
지역7		자동차 등록 지역이 7이면 1, 아니면 0	
지역8		자동차 등록 지역이 8이면 1, 아니면 0	
지역9		자동차 등록 지역이 9이면 1, 아니면 0	
지역10		자동차 등록 지역이 10이면 1, 아니면 0	
지역11		자동차 등록 지역이 11이면 1, 아니면 0	
지역12		자동차 등록 지역이 12이면 1, 아니면 0	
지역13		자동차 등록 지역이 13이면 1, 아니면 0	
지역14		자동차 등록 지역이 14이면 1, 아니면 0	
지역15		자동차 등록 지역이 15이면 1, 아니면 0	
지역16		자동차 등록 지역이 16이면 1, 아니면 0	

모형으로 일반화선형(Generalized Linear Method)모형이 적합하다. 일반화선형모형은 종속변수의 연속성의 성격에 따라 크게 두 분류로 구분되는데, 그 중에서도 일정기간 동안 사고(Events) 발생 건수와 같은 비연속변수인 카운트변수(Count Variables)의 경우 포아송(Poisson) 또는 음이항분포(Negative Binomial) 모형이 활용될 수 있다 (Cameron과 Trivedi, 2005). Figure 3.1의 경우 교통사고 빈도의 분포를 보여주고 있는데 카운트 변수이면서 지수분포를 보이고 있음을 알 수 있다.

포아송의 경우 평균과 분산이  $\lambda$ 로 같다는 가정에 기반한 모형이기 때문에 분산이 평균보다 큰지 여부, 즉 과대산포가 존재하는지 여부를 검증하여 음이항분포와 포아송 중 적합한 모형을 선택해야 한다. 음이항 분포에 따라 교통사고 빈도를 나타내는 확률변수  $Y$ 는 식 (3.1)로 표현된다.

$$P(Y = y) = \frac{\Gamma(\alpha^{-1} + y)}{\Gamma(y + 1)\Gamma(\alpha^{-1})} \left( \frac{\alpha^{-1}}{\alpha^{-1} + \alpha^{-1}} \right)^{\alpha^{-1}} \left( \frac{\lambda}{\alpha^{-1} + \lambda} \right)^y, \quad (3.1)$$

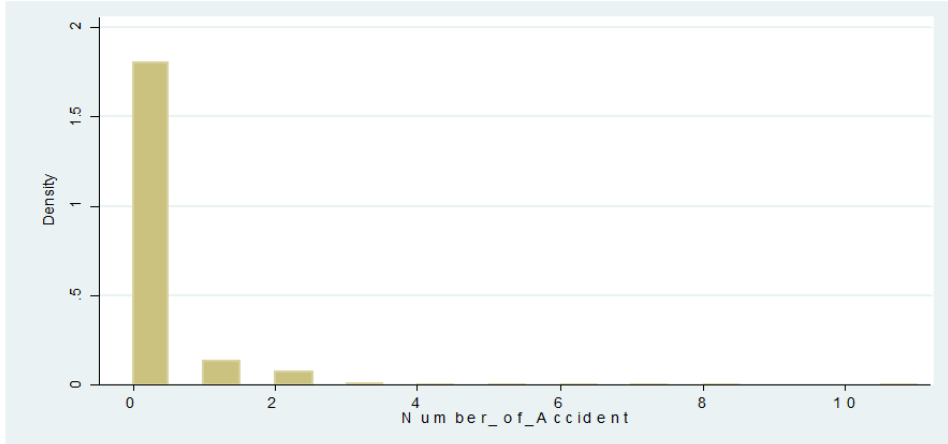


Figure 3.1. Distribution of the number of accidents

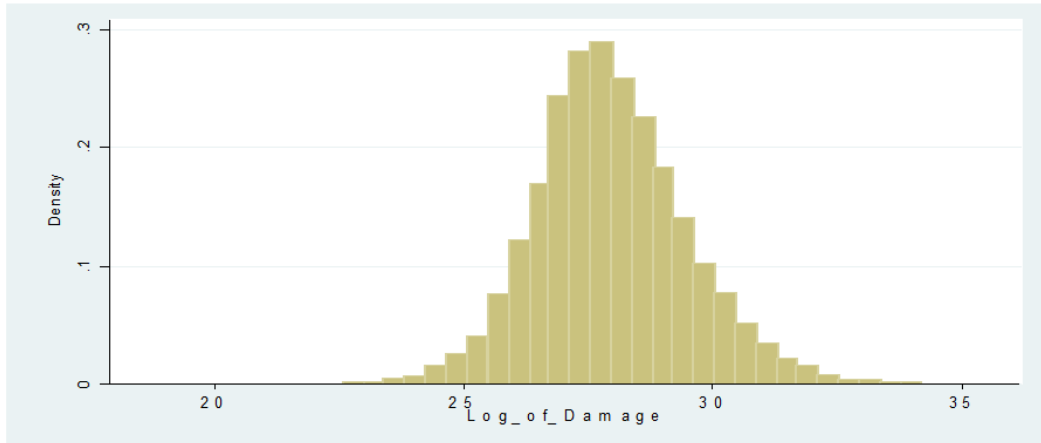


Figure 3.2. Distribution of the logarithmic values of accident damages

$\Gamma$ 는 감마(gamma)함수이며,  $\alpha$ 는 과대산포 여부를 나타내는 변수인데 음이항 분포의 경우 평균이 포아송처럼  $\lambda$ 로 동일하지만, 분산은  $\lambda + \alpha\lambda$ 으로 포아송의 분산  $\lambda$ 과 다르다. 만약 과대산포가 없어  $\alpha$ 가 0이 되면, 식 (3.1)은 포아송분포 식 (3.2)로 변경된다는 것을 알 수 있다.

$$\Pr(Y = y) = \frac{e^{-y} + \lambda^y}{y!}, \quad y! = y \times (y - 1) \times \dots \times 2 \times 1. \tag{3.2}$$

교통사고 심도의 경우에는 자동차사고 발생이 전제되어야만 발생하는 변수이기 때문에 최소자승법(Ordinary least square)이 아닌 토빗(Tobit)모형이나 헤크먼(Heckman)모형이 적합하다 (Cameron 과 Trivedi, 2005).

$$y1 = \begin{cases} 1, & \text{if } y_1^* > 0, \\ 0, & \text{if } y_1^* \leq 0. \end{cases} \tag{3.3}$$

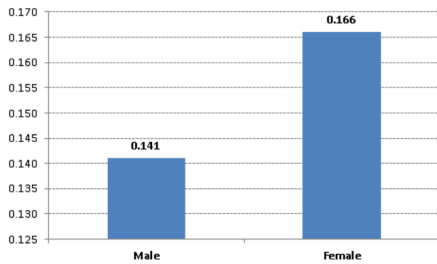


Figure 4.1. Accident frequency by gender

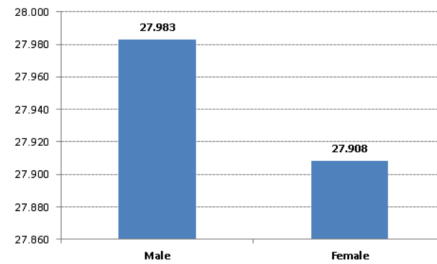


Figure 4.2. Accident severity by gender

잠재변수(latent variable)  $y^*$ 를 활용해 자동차사고 발생 가능성을 프로빗(Probit)모형으로 분석하게 되는데, 자동차사고가 발생한 사람들만 자동차사고 심도  $y_2^*$ 가 관측가능하게 된다.

$$y_2 = \begin{cases} y_2^*, & \text{if } y_1^* > 0, \\ 0, & \text{if } y_1^* \leq 0. \end{cases} \quad (3.4)$$

식 (3.4)를 분석하기 위해 토빗모형과 헤크먼모형이 고려될 수 있겠으나, 토빗모형의 경우 자가선택(sample selection)의 영향을 고려하지 않는다. 즉, 종속변수인 교통사고 심도는 교통사고 사고가 경험한 운전자에게만 발생하는 비용인데 Tobit모형은 교통사고를 경험한 운전자와 무사고 운전자를 동일한 성격의 운전자로 취급하기 때문에 내생성(endogenous)의 문제가 발생한다. 반면 헤크먼모형은 교통사고를 경험한 운전자와 무사고 운전자 간 차이를 고려하여 내생성의 문제를 해결한다 (Cameron과 Trivedi, 2005). 이러한 헤크먼 모형은 주로 노동경제학에서 활용되고 있는데, 종속변수가 임금일 경우 노동시장에 진입한 사람만 임금이 관찰되기 때문에 자가선택문제를 고려한 헤크먼 모형이 주로 활용된다 (Heckman, 1979; Lee, 2005).

본 연구에서 사고심도 분포를 확인한 결과 Figure 3.2에서 보여지는 것처럼 로그값으로 변환한 뒤 심도값의 분포가 정규분포에 가까워 사고피해액을 그대로 분석에 활용하지 않고 로그값으로 변환한 뒤에 헤크먼모형을 적용하였다.

## 4. 실증분석 결과

### 4.1. 기술적 통계

Table 4.1은 본 연구에 활용된 샘플을 남성운전자와 여성운전자로 구분하여 각 성별에 따른 변수들의 평균과 표준편차를 보여준다. 예를 들어, 남성의 평균 사고빈도는  $\sum_{i=1}^{N_m} Acc_i / N_m$ 으로 계산되었으며,  $N_m$ 은 총 남성운전자 수,  $Acc_i$ 는 각 남성운전자의 사고건수이다. 이러한 통계값으로는 남성과 여성의 평균 사고빈도만 비교할 수 있을 뿐 각 운전자의 연령, 운전경력 등 다른 요인들이 교통사고 빈도에 미치는 영향들이 반영되지 않았다.

총 488,139명 중 여성은 104,740으로 약 21%의 비중이다 (정부에서 발표하는 여성운전자 비중보다 낮은 이유는 정부의 경우 운전면허증 소지자를 기준으로 성별 비중을 계산하지만, 본 연구에서는 운전면허증을 보유한 사람 중 보험에 가입한 사람을 기준으로 성별 비중을 계산하기 때문이다.). 총 샘플의 경우 평균 0.146번의 교통사고를 유발하였으며 여성운전자는 0.166번으로 남성의 0.141번보다 사고빈도가 높았다. 교통사고 피해액인 사고심도의 경우 남성이 여성보다 다소 높기는 하였으나 큰 차이가 나지는 않고 있다.

**Table 4.1.** Descriptive statistics of the sample by gender

구분	총 샘플		남성 운전자		여성 운전자	
	평균	표준편차	평균	표준편차	평균	표준편차
사고빈도	0.146	0.482	0.141	0.474	0.166	0.513
사고심도	27.965	1.530	27.983	1.532	27.908	1.525
남성	0.785	0.411	1.000	0.000	0.000	0.000
연령	43.060	10.620	43.324	10.721	42.092	10.185
경력1	0.051	0.220	0.037	0.188	0.104	0.306
경력2	0.046	0.209	0.035	0.184	0.085	0.279
경력3	0.051	0.220	0.041	0.199	0.087	0.283
경력4	0.082	0.275	0.075	0.264	0.108	0.310
경력5	0.100	0.300	0.098	0.297	0.109	0.312
경력6	0.669	0.470	0.714	0.452	0.506	0.500
경차	0.075	0.263	0.063	0.243	0.117	0.322
소형	0.310	0.463	0.300	0.458	0.347	0.476
중형	0.309	0.462	0.316	0.465	0.280	0.449
대형	0.097	0.297	0.099	0.299	0.091	0.288
벤	0.004	0.065	0.005	0.068	0.003	0.053
SUV	0.205	0.403	0.216	0.412	0.161	0.368
지역1	0.188	0.390	0.185	0.389	0.196	0.397
지역2	0.064	0.244	0.064	0.245	0.062	0.242
지역3	0.234	0.423	0.232	0.422	0.239	0.426
지역4	0.033	0.177	0.032	0.176	0.034	0.182
지역5	0.032	0.176	0.032	0.177	0.031	0.173
지역6	0.040	0.197	0.041	0.197	0.040	0.196
지역7	0.035	0.183	0.035	0.183	0.035	0.184
지역8	0.031	0.172	0.031	0.174	0.028	0.166
지역9	0.056	0.231	0.058	0.234	0.051	0.220
지역10	0.068	0.251	0.069	0.254	0.062	0.241
지역11	0.012	0.108	0.011	0.105	0.015	0.120
지역12	0.059	0.235	0.058	0.235	0.060	0.238
지역13	0.054	0.225	0.054	0.226	0.053	0.224
지역14	0.032	0.175	0.032	0.175	0.032	0.176
지역15	0.038	0.192	0.038	0.191	0.039	0.195
지역16	0.026	0.160	0.027	0.163	0.022	0.147
샘플 수	488,139		383,399		104,740	

평균연령은 남성이 43.3세로 여성의 평균연령인 42.1세보다 다소 많고 운전경력도 현격하게 남성이 여성보다 많음을 알 수 있다. 총 샘플의 경우 차종은 소형차와 중형차가 60% 이상을 차지하고 있으며, 남성은 중형차를 가장 많이 운전하고 여성은 소형차를 가장 많이 운전하고 있는 것으로 나타났다. 특히 경차의 경우에는 여성운전자의 비중이 현격하게 높음을 알 수 있으며, 반대로 SUV의 경우에는 남성운전자의 비중이 현격하게 높다.

#### 4.2. 성별 교통사고 위험도

Table 4.2의 왼쪽 결과는 사고심도를 기준으로 성별 자동차사고 위험도를 분석한 결과를 보여준다. 교통사고 심도 및 빈도를 분석한 모형 모두, 성별의 기준변수로 여성을, 운전경력의 기준변수로 경력6을,

Table 4.2. Level of riskiness by gender

구분	사고심도		사고빈도		
	추정계수	표준오차	추정계수	IRR	표준오차
남성	0.083	0.032***	-0.119	0.888	0.012***
연령	0.003	0.001**	0.003	1.003	0.001***
경력1	0.291	0.134**	0.533	1.703	0.022***
경력2	0.131	0.100	0.313	1.367	0.024***
경력3	0.105	0.082	0.242	1.274	0.023***
경력4	0.125	0.053**	0.097	1.102	0.019***
경력5	0.125	0.044***	0.065	1.067	0.017***
경차	-0.241	0.066***	-0.212	0.809	0.021***
소형	-0.118	0.036***	-0.117	0.889	0.013***
대형	0.187	0.045***	-0.063	0.939	0.019***
벤	0.009	0.234	-0.268	0.765	0.083***
SUV	0.191	0.042***	0.141	1.151	0.014***
지역2	-0.110	0.070	-0.086	0.917	0.023***
지역3	0.169	0.036***	-0.043	0.958	0.015***
지역4	0.136	0.084	-0.203	0.816	0.031***
지역5	0.047	0.088	-0.191	0.826	0.031***
지역6	0.104	0.083	-0.132	0.877	0.028***
지역7	0.107	0.064*	-0.036	0.964	0.029
지역8	0.158	0.098	-0.182	0.834	0.032***
지역9	-0.016	0.077	-0.106	0.900	0.024***
지역10	0.026	0.086	-0.126	0.881	0.023***
지역11	-0.377	0.210*	-0.279	0.757	0.051***
지역12	-0.218	0.057***	-0.004	0.996	0.023
지역13	0.152	0.060***	0.084	1.088	0.024***
지역14	0.030	0.074	-0.042	0.959	0.030
지역15	-0.009	0.062	-0.050	0.951	0.028*
지역16	-0.118	0.094	-0.100	0.905	0.033***
_cons	27.711	1.127***	-1.977	0.139	0.030***
분석모형	헤크먼		음이항분포		
Log Likelihood	-104976.40(Prob>Chi2<0.000)		-204027.48(Prob>Chi2<0.000)		
유의성 표시	*, **, ***는 각각 유의수준 10%, 5%, 1%에서 통계적으로 유의함을 의미				
샘플수	488,139				

차종의 기준변수로 중형차를, 지역의 기준변수로 서울을 의미하는 지역1을 활용하였다.

사고심도를 분석한 헤크먼모형에서는 종속변수가 로그값으로 전환한 사고피해액이기 때문에 추정계수 값에 100을 곱하여 종속변수의 변화율로 해석하는 것이 용이하다 (즉, 특정 독립변수  $x$ 가 한 단위 변할 때 심도변화율은 다음과 같이 정의된다 (Wooldridge, 1999). 사고심도변화율  $\approx (100 \cdot \text{추정계수})\Delta x$ ).

사고심도의 경우 남성 운전자가 여성운전자에 비해 8.3% 높은 것으로 나타났으며 유의수준 1%에서 통계적으로 유의한 것으로 추정되었다. 이는 해외 선행연구와 일치하는 결과로 우리나라 남성운전자도 여성보다 교통사고가 발생할 경우 피해액이 크다는 것을 알 수 있다.

연령은 증가할수록 사고심도가 높은 것으로 나타났으며, 운전경력이 1년 미만인 초보운전자(경력1)는 운전경력이 5년 이상인 운전자(경력6)에 비해 교통사고가 발생할 경우 피해액이 29.1%나 높은 것으로



분석되어졌다.

차종에 따른 교통사고 피해액은 일반적으로 차의 크기와 비례하여 증가하였는데, 기준변수인 중형차에 비해서는 경차와 소형차의 피해액이 작은편이었고 대형차와 SUV 운전자는 교통사고 심도가 각각 18.7%, 19.1% 높은 것으로 나타났다. 또한 운전하는 지역도 교통사고 피해액에 유의한 영향을 주는 요인임을 알 수 있다. 기준변수인 서울과 함께 경기도와 인천의 사고심도가 높은 반면, 전북의 사고심도는 상대적으로 낮은 것으로 분석되어졌다.

Table 4.2의 오른쪽 결과는 사고빈도를 기준으로 성별 자동차사고 위험도를 분석한 결과를 보여준다. 사고빈도의 경우 포아송과 음이항분포 중 적절한 모형을 선택하기 위해  $\alpha > 0$ 인지 여부를 검증해야 한다.  $\alpha$ 의 값이 5.24로 추정되었으며 표준오차가 0.053으로 Likelihood-ratio 검증결과  $\alpha$ 가 유의하게 0보다 큰 것으로 검증되었다. 즉, 본 연구에 활용된 사고빈도자료의 경우 분산이 평균보다 유의적으로 크기 때문에 포아송이 아닌 음이항분포가 적절한 실증분석모형임을 알 수 있다.

사고빈도를 종속변수로 지정하여 성별 교통사고 위험도를 음이항분포로 추정한 결과 남성보다 오히려 여성의 교통사고 발생위험이 높은 것으로 나타났다. 하지만 Table 4.2에 기입된 추정계수는 식 (3.1)을 음이항분포로 추정된 것으로  $\beta = \log(\mu_{x+1}) - \log(\mu_x)$ 를 의미한다. 즉, 음이항분포의 추정계수는 특정 독립변수  $x$ 가 한 단위 변화할 때 교통사고 빈도의 기대 로그값이 얼마나 변하는지를 보여준다. 이러한 추정계수는 해석상 어려움이 있기 때문에 추정계수를 이해하기 쉬운  $\mu_{x+1}/\mu_x$ 로 변환할 필요가 있다. 로그값의 차이가 비율로 표현될 수 있는 특성을 이용해  $\beta = \log(\mu_{x+1}) - \log(\mu_x)$ 는  $\log(\mu_{x+1}/\mu_x)$ 로 변환될 수 있으며, 양변의 값을 자연지수화하는 방법으로  $\mu_{x+1}/\mu_x$ 를 구할 수 있다. 이렇게 음이항분포모형으로 추정된 계수값을 비율로 변화시킨 값을 IRR(Incidence rate ratio)이라고 한다. 예를 들어, Table 4.2의 남성이라는 독립변수의 IRR이 0.88이라는 의미는, 여성이 교통사고를 발생시킬 확률이 1이라면 남성의 경우에는 0.88정도로 낮다는 의미이다. 즉, 여성이 남성보다 교통사고 발생확률이 1.13배 높은 것으로 추정되었다.

연령이 증가함에 따라 사고빈도도 증가하여 연령은 사고심도뿐만 아니라 사고빈도도 높이는 요인임을 알 수 있다. 사고빈도는 사고심도와는 달리 운전경력이 증가하면 교통사고 발생확률이 지속적으로 감소하는 것으로 나타났다. 반면 운전경력 1년 미만인 운전자의 경우 5년 이상인 운전자에 비해 교통사고 발생확률이 1.7배나 높은 것으로 추정되었다.

차종에 따른 사고빈도 위험도는 사고심도와 다소 다른 양상을 보였다. 중형차와 SUV 운전자가 교통사고 발생확률이 가장 높았으며, 경차와 소형차는 사고심도처럼 사고빈도 측면에서도 위험도가 높지 않았다. 하지만 대형차는 교통사고 발생확률이 중형차보다 낮은 것으로 분석되어졌다. 즉, 대형차의 경우 교통사고 발생가능성은 낮으나 사고가 발생할 경우에는 교통사고 피해액이 많은 것으로 나타났다.

지역은 사고심도와 같이 교통사고 발생가능성에 영향을 주는 매우 유의한 요인임이 확인되었다. 특히 사고심도가 높은 서울과 인천은 사고빈도도 다른 지역보다 높은 것으로 분석되어졌다.

## 5. 결론

본 연구는 우리나라 모든 운전자 자료 중 50만명을 무작위 추출하여 성별 자동차사고 위험도를 분석하였다. 자동차사고 위험도는 사고심도와 사고빈도로 구분한 이후 각각에 대해 헤크먼모형과 일반화선형모형을 적용하여 분석하였는데, 분석결과 사고심도의 경우 남성이 여성보다 8.3% 높은 반면 사고빈도의 경우에는 반대로 여성에 비해 남성의 교통사고 발생 가능성이 현저하게 낮은 것으로 분석되어졌다.

이러한 결론은 해외의 주요 선행연구와 일치하는 결과이다. 우리나라 교통사고 위험도는 OECD 국가

중 가장 높은 상황으로 정부는 주기적으로 교통사고 경감을 위한 정책들을 제시해 왔다. 남성의 경우 다소 공격적인 운전습관으로 인해 교통사고가 발생할 경우 사고심도가 매우 높게 나타나고 있기 때문에 정속주행이나 교통신호 준수 등과 같은 교통법규에 대한 교육을 강화하는 정책들이 유용할 것으로 판단된다. 반면 여성의 경우에는 상대적으로 낮은 지각판단능력으로 인해 교통사고 발생빈도가 높기 때문에 커브길이나 주차장에 반사경을 설치할 경우 여성의 교통사고 발생을 크게 감소시킬 수 있을 것으로 기대된다. 또는 차선변경을 용이하게 하는 차량기술개발 등도 남정보다는 여성의 사고감소에 더욱 효과적일 것으로 판단된다.

본 연구가 우리나라 운전자들의 성별 교통사고 위험도를 사고심도와 사고빈도로 구분하여 실증적으로 분석은 하였지만, 데이터의 한계로 인해 성별로 교통사고 위험도가 차이나는 구체적인 원인에 대해서는 충분한 설명을 제공하지 못하는 한계가 존재한다. 현재 국토교통부, 경찰청, 그리고 보험산업이 서로 다른 교통사고 정보를 보유하고 있는데 향후 다양한 기관에서 독립적으로 축적하고 있는 사고정보들을 통합한다면 성별로 교통사고 빈도 및 심도가 차이나는 원인에 대해 규명하고 실효성 높은 교통사고 경감정책이 시행될 수 있을 것으로 기대된다.

## References

- Al-Balbissi, A. H. (2003). Role of gender in road accidents, *Traffic Injury Prevention*, **4**, 64–73
- Baker, S., O'Neill, B., Ginsburg, M. J. and Li, G. (1992). *The Injury Fact Book*, Oxford University Press, New York.
- Cameron, A. C. and Trevedi, P. K. (2005). *Microeconomics : Method and Applications*, Cambridge Press.
- Doob, A. N. and Gross, A. (1968). A status of frustrator as an inhibitor of horn-honking responses, *The Journal of Social Psychology*, **76**, 213–218.
- Heckman, J. J. (1979). Sample selection bias as a specification error, *Econometrica*, **47**, 153–161
- Lajunen, T. and Parker, D. (2001). Are aggressive people aggressive drivers?: A study of the relationship between self-reported general aggressiveness, driver anger and aggressive driving, *Accident Analysis & Prevention*, **33**, 243–255
- Lee, D. S. (2005). Training, wages, and sample selection: estimating sharp bounds on treatment effects, NBER, No. 11721
- Massie, D. L. and Campbell, K. L. (1993). Analysis of accident rates by age, gender, and time of day based on the 1990 nationwide personal transportation survey, Transportation Research Institute.
- Ministry of Land, Infrastructures and Transport (2012). Press release
- Ministry of Land, Infrastructures and Transport (2013). National plans of traffic safety for 2013.
- National Statistical Office (2013). Statistics of the cause of death in 2013.
- Ryu, S. G., Bin, M. Y., Kim, C. M. and Kim, J. S. (2013). *Gyeonggi Research Institute, Issue&Diagnosis*
- Storie, V. J. (1977). Male and Female Car drivers: Differences observed in accidents, UK, England: Transport and Road Research Laboratory.
- Waldron, I., McCloskey, C. and Earle, I. (2005). Trends in gender differences in accidents mortality: Relationships to changing gender roles and other societal trends, *Demographic Research*, **13**, 415–454.
- Waylen, A. and McKenna, F. (2002). The development of gender differences in risky attitudes and behavior in road use, UK, Basingstoke: AA Foundation for Road Safety.
- Wooldridge, J. M. (1999). *Introductory Econometrics : Modern Approach*, Cengage Learning EMEA, 43.
- Zuckerman, M. (1996). The psychobiological model for impulsive unsocialized sensation seeking: A comparative approach, *Neuropsychobiology*, **34**, 125–129.

# 일반화선형모형과 헤크먼모형을 활용한 성별 자동차사고 위험도 분석

김대환<sup>a</sup> · 박화규<sup>b,1</sup>

<sup>a</sup>동아대학교 경제학과, <sup>b</sup>순천향대학교 보건행정경영학과

(2013년 12월 16일 접수, 2014년 2월 6일 수정, 2014년 2월 6일 채택)

## 요약

최근 여성의 사회적 참여와 지위가 상승함에 따라 여성운전자도 급격히 증가해왔다. 이에 주요국에서는 성별 자동차사고 위험도를 분석하고, 성별 위험도 차이의 원인을 규명하려는 연구가 활발히 진행되고 있다. 본 연구는 국내 모든 운전자 중 50만 명을 무작위로 추출하여 성별 자동차사고 위험도를 분석하였다. 종속변수로 사고심도와 사고빈도를 활용하였는데, 사고심도 변수에는 헤크먼모형(Heckman Method)을, 사고빈도 변수에는 일반화선형모형(Generalized Linear Method)을 적용하였다. 분석결과 남성은 교통사고가 발생할 경우 사고피해액이 여성보다 8.3% 높은 것으로 추정되었으며, 반대로 사고빈도의 경우에는 여성의 위험도를 1로 설정할 경우 남성의 위험도는 0.88로 남성의 위험도가 오히려 낮은 것으로 분석되었다. 향후 성별로 상이한 자동차사고 위험도의 원인이 규명된다면 좀 더 실효성 높은 정책을 통해 자동차사고를 경감시킬 수 있을 것이다.

주요용어: 여성운전자, 자동차사고, 헤크먼모형, 음이항분포.

이 논문은 동아대학교 교내연구비 지원에 의하여 연구되었음.

<sup>1</sup>교신저자: (336-745) 충남 아산시 신창면 순천향로 22, 순천향대학교 보건행정경영학과, 교수.

E-mail: hkpark1@sch.ac.kr