

Random Effects Tobit 회귀모형을 이용한 교차로 교통사고 요인 분석

An Analysis on Vehicle Accident Factors of Intersections using Random Effects Tobit Regression Model

이 상 혁* · 이 정 범**

* 주저자 : 한국건설기술연구원 수석연구원

** 교신저자 : 대전세종연구원 책임연구원

Sang Hyuk Lee* · Jung-Beom Lee**

* Highway and Transportation Research Institute, Korea Institute of Civil Engineering and Building Tech.

** Urban Infrastructure Division, Daejeon Sejong Research Institute

† Corresponding author : Jung-Beom Lee, rutgerskg@gmail.com

Vol.16 No.1(2017)

February, 2017

pp.26~37

ISSN 1738-0774(Print)

ISSN 2384-1729(On-line)

<https://doi.org/10.12815/kits>

2017.16.1.26

Received 31 August 2016

Revised 27 September 2016

Accepted 18 January 2017

© 2017. The Korea Institute of
Intelligent Transport Systems. All
rights reserved.

요 약

본 연구는 random effects Tobit 회귀모형을 이용하여 도심지 교차로에 대한 교통사고모형을 개발하여 교통사고와 요인간의 상관관계를 파악하는 것이 목적이다. Random effects Tobit 회귀모형의 적용성을 비교·분석하기 위하여 fixed effect Tobit 회귀모형을 산정하였다. 산정 결과, 교통량, 제한속도, 차로수, 토지이용, 우회전차로, 전방신호등이 유의한 변수로 나타났으며, 총 교통사고율에 대한 random effects 모형의 모형 적합도(결정계수: 0.418, 로그-우도함수값: -3210.103, 우도비: 0.056)와 모형 설명력(MAD: 19.533, MAPE: 75.725, RMSE: 26.886)은 fixed effects 모형의 모형 적합도 (결정계수: 0.298, 로그-우도함수값: -3276.138, 우도비: 0.037)와 모형 설명력(MAD: 20.725, MAPE: 82.473, RMSE: 27.267)보다 우수한 것으로 나타났으며, 부상교통사고율에 대한 교통사고모형에서도 총 교통사고율의 산정결과와 동일하게 나타나 두 모형에서 random effects Tobit 회귀모형이 다소 우수한 것으로 분석되었다.

핵심어 : 교통사고모형, random effects, fixed effects, Tobit 회귀모형, 도심지 교차로

ABSTRACT

The study is to develop safety performance functions(SPFs) for urban intersections using random effects Tobit regression model and to analyze correlations between crashes and factors. Also fixed effects Tobit regression model was estimated to compare and analyze model validation with random effects model. As a result, AADT, speed limits, number of lanes, land usage, exclusive right turn lanes and front traffic signal were found to be significant. For comparing statistical significance between random and fixed effects model, random effects Tobit regression model of total crash rate could be better statistical significance with R_p^2 : 0.418, log-likelihood at convergence: -3210.103, ρ^2 : 0.056, MAD: 19.533, MAPE: 75.725, RMSE: 26.886 comparing with R_p^2 : 0.298, log-likelihood at convergence: -3276.138, ρ^2 : 0.037, MAD: 20.725, MAPE: 82.473, RMSE: 27.267 for the fixed model. Also random effects Tobit regression model of injury crash rate has similar results of model statistical significant with random effects Tobit regression model.

Key words : Safety Performance Functions, Random Effects, Fixed Effects, Tobit Regression, Urban Intersections

I. 서론

1. 개요

도로의 교통사고는 인적요인, 차량요인, 시설요인, 환경요인 등에 의해 발생하고 있으며, 이들 교통사고 발생에 영향을 미치는 요인은 통계학적 분석을 통하여 분석할 수 있다. 기존의 많은 연구에서는 교통사고 발생 원인이 되는 교통시설, 기하구조, 교통운영 요소를 파악하고 이를 개선하여 교통사고발생빈도를 줄이기 위한 방안을 마련하기 위한 연구가 진행되어져 왔다.

교통사고 발생 원인을 분석하기 위하여 기존 연구에서는 선형회귀모형(linear regression model), 포아송회귀모형(Poisson regression model)과 음이항회귀모형(negative binomial model)과 같은 가산모형(count model)이나 로짓모형(logit model)과 프로빗모형(probit model)과 같은 이산선택모형(discrete choice model) 등을 이용하여 교통사고모형을 구축하고 구축된 교통사고모형을 통해 교통사고 원인을 분석하였다.

하지만, 교통사고모형 개발 시 종속변수로 이용되는 교통사고데이터가 연도별/교차로 · 가로구간별 교통사고발생건수를 이용하는 경우에 가산모형을 이용할 수 있지만, 연도별/교차로 · 가로구간별 교통사고율이 종속변수일 경우에는 가산모형의 적용이 적합하지 않다. 또한, 교통사고율은 항상 양의 수(+)만 가지게 되므로 교통사고데이터는 영(zero)을 기준으로 왼편중도절단(left-censored at zero)형태가 된다. 이러한 일정한 구간의 값을 가진 교통사고데이터를 종속변수로 하는 교통사고모형은 Tobit 회귀모형으로 개발하는 것이 가능하다 (Anastasopoulou et al., 2008).

또한, 기존 모형은 교통사고발생빈도에 영향을 미치는 독립변수에 대한 추정된 모수(parameters)가 고정되어 있다고 가정하고 있다(Lee et al., 2015). 그러나 독립변수인 기하구조와 도로지형 등은 이질성(heterogeneity)이 존재함에도 불구하고 많은 연구에서는 동일하다고 가정(fixed effects)하여 교통사고모형을 개발하였다. 이렇게 개발된 모형의 계수는 표준오차 값이 과소 추정되거나 각 계수의 t -값이 과도하게 산정되어 모형의 설명력이 저하될 수 있다(Lee et al., 2015). 이러한 한계를 극복하기 위하여 모형구축과정에서 fixed effects 대신 확률적 모수인 random effects를 적용하는 것이 가능하다.

따라서 본 연구에서는 random effects를 고려한 Tobit regression model을 이용하여 대전광역시 주요 교차로에 대한 교통사고모형을 개발하고 fixed effects Tobit regression model과 비교 · 분석하였다.

II. 기존연구 고찰

Calzolari et al.(2001)은 random effects를 고려한 패널데이터(panel data) Tobit 모형을 위한 시뮬레이션 기반 추정의 작동 여부를 Monte Carlo 실험과 함께 평가하였다. Calzolari and Magazzini(2008)의 연구는 Calzolari et al.(2001)의 연구에 random effects뿐만 아니라 자기 상관 오류(autocorrelated errors)에 대한 고려를 포함시킨 것으로 본 모델을 특허와 R&D의 관계에 대한 분석에 적용시켰다. 분석결과, Monte Carlo 실험에서 본 모델은 추정효과가 뛰어난 것으로 나타났다. 자기 상관 오류의 경우에는 낮은 측정가능성의 문제가 발생했으며 이에 대한 해결책으로 제한된 간접추정법(constrained indirect estimation)이 제안되었다.

Bruno(2004)는 generalized method of moments (GMM), 혹은 Markov Chain Monte Carlo (MCMC) 방법론들의 비교를 통하여 Probit과 Tobit 모형과 같은 비선형 패널데이터 모형을 해결하고자 하였다. 본 연구 결과 데이터의 크기가 큰 경우에는 MCMC 방법론의 적용이 보다 용이한 것으로 나타났다.

Greene(2004)은 패널의 길이가 작거나 고정되어있을 때에 비선형 패널데이터 fixed effects 모형에서 최우추정량 (maximum likelihood estimator)이 편향되어있으며 일관성이 없다고 일반적으로 여겨지는 사실에 대한 학술적 근거가 거의 없다는 점에 주목하였다. 따라서 본 연구에서는 Monte Carlo method를 활용하여 fixed effects Tobit 모형의 최우추정량의 추정정도를 측정하였다.

Qi et al.(2007)은 고속도로에서의 사고가능성을 예측할 수 있는 모형의 정립을 위하여 패널데이터를 위한 discrete response 모형인 random effects ordered probit 모형을 개발하였다. 개발결과, 본 연구에서 개발한 모형은 교통사고와 관련된 요인을 발견하는 데에 적합한 것으로 나타났다. 또한, 개발된 모형은 예측모형으로써 적용되었을 때에 다양한 시간대 및 장소에 기반을 둔 예측결과를 도출하는 장점을 갖고 있다.

Henningsen(2010)은 표준 Tobit 모형을 포함한 절단회귀모형(Censored Regression Models)을 censReg을 활용하여 통계 프로그램 R을 통해 분석하였다. 또한, Islam과 Hernandez(2011)는 교통사고율에 있어서 대형트럭과 같은 화물차의 이동 효과에 대한 연구가 취약하다는 점을 지적하였다. 따라서 random parameters tobit regression 모형을 활용하여 미국의 대형트럭 운행거리 등을 고려하여 트럭 교통사고율에 대해 연구하였다. 실증연구 결과 Tobit 모형이 대형트럭의 교통사고율을 결정하는 요인들을 파악하는 데에 효과가 있다고 나타났다.

Anastasopoulos et al.(2012)은 random effects tobit model을 이용하여 고속도로 교통사고에 미치는 요인에 대하여 연구하였다. 본 연구의 경우 기존 연구에서 다뤄지지 않았던 관찰되지 않는 이질성(unobserved heterogeneity)에 대해 설명하고자 하였다. 실증연구결과 본 모형은 특정한 도로구간에서의 사고율을 결정하는 요인에 대한 보다 깊은 이해를 제공할 수 있는 가능성을 발견하였다.

Xu et al.(2014)의 연구는 신호교차로의 안전성 향상을 위하여 교통사고 심각도(severity)에 따른 two-stage bivariate logistic Tobit 모형을 제시하였다. 본 모형은 사고빈도와 심각도의 상관관계를 고려한 것으로 첫 번째 단계에서 교통사고 심각도를 결정하고 두 번째 단계에서 경상을 입은 경우와 사망을 포함한 중상을 입은 경우의 위험성을 평가하였다. 본 연구결과 전반적인 교통사고 심각도는 연평균 하루 교통사고 및 횡단보도 개수에 영향을 받는다고 나타났다. 또한, 조업차량비율이 경상 및 중상이상의 사고 수에 영향을 주는 것으로 나타났다.

Ⅲ. 이론적 배경

1. Tobit 회귀모형

대전광역시 주요교차로의 교통사고모형을 개발하기 위하여 Tobit 회귀모형을 이용하였다. James Tobin에 의해 제안된 Tobit 모형은 교통사고발생빈도와 같이 종속변수가 일정범위(≥ 0) 안에서 관측되는 제한된 값을 가지는 경우, 즉 중도절단 되었을 경우 사용되는 모형으로 Tobit 모형의 일반식은 다음과 같다(Henningsen, 2010)

$$y_i^* = x_i' \beta + \varepsilon_i$$

$$y_i = \begin{cases} 0 & \text{if } y_i^* \leq 0 \\ y_i^* & \text{if } y_i^* > 0 \end{cases} \quad (1)$$

여기서,

y_i^* : 가상의 종속변수 ($i = 1, 2, \dots, N$, N : 표본의 크기)

x_i : 독립변수의 벡터

β : 추정해야 할 모수벡터

ε_i : 오차항

또한, 중도절단회귀모형인 Tobit 회귀모형의 일반식은 다음과 같다.

$$y_{it}^* = x_{it}'\beta + \varepsilon_{it} = x_{it}'\beta + \mu_i + \nu_{it} \tag{2}$$

$$y_{it} = \begin{cases} a & \text{if } y_{it}^* \leq a \\ y_{it}^* & \text{if } a < y_{it}^* < b \\ b & \text{if } y_{it}^* \geq b \end{cases}$$

여기서,

y_{it}^* : 가상의 종속변수 ($i = 1, 2, \dots, N$, $t = 1, 2, \dots, T_i$, N : 표본의 크기, T_i : 데이터 수집 연도수)

μ_i : i 교차로에서 예측된 교통사고율

ν_{it} : 오차항

2. Fixed Effects Tobit 회귀모형

일반적으로 중도절단회귀모형인 Tobit 회귀모형은 우도함수(likelihood function)를 적용한 최우추정법(maximum likelihood method)에 의해 개발이 가능하다. 최우추정법에 의한 Tobit 회귀모형은 오차항 ν 는 정규분포(평균 0, 분산 σ^2)을 따른다는 전제를 만족해야 하며, 로그 우도함수(log-likelihood function)는 다음과 같다.

$$\log L = \sum_{i=1}^N \sum_{t=1}^{T_i} \left[I_{it}^a \log \Phi \left(\frac{a - x_{it}'\beta - \mu_i}{\sigma} \right) + I_{it}^b \log \Phi \left(\frac{x_{it}'\beta + \mu_i - b}{\sigma} \right) \right. \\ \left. + (1 - I_{it}^a - I_{it}^b) \left(\log \phi \left(\frac{y_{it} - x_{it}'\beta - \mu_i}{\sigma} \right) - \log \sigma \right) \right] \tag{3}$$

여기서,

L : 우도추정값

x_{it}' : 각 교차로와 년도별 독립변수의 벡터

β : 추정해야 할 모수벡터

μ_i : i 교차로에서 예측된 교통사고율

$\phi(\cdot)$ 와 $\Phi(\cdot)$: 표준정규분포의 확률밀도함수와 누적분포함수

I_{it}^a 와 I_{it}^b : 표시함수(indicator functions)

$$I_{it}^a = \begin{cases} 1 & \text{if } y_{it} = a \\ 0 & \text{if } y_{it} > a \end{cases} \quad I_{it}^b = \begin{cases} 1 & \text{if } y_{it} = b \\ 0 & \text{if } y_{it} < b \end{cases}$$

3. Random Effects Tobit 회귀모형

Random effects는 교통사고모형을 개발할 때 교차로나 가로구간 등이 도로지형 및 기하구조의 이질성 (heterogeneity)을 설명하기 위하여 적용하는 방법으로 Tobit 회귀모형 적용 시 시뮬레이션 최우추정법 (simulated maximum likelihood method)을 이용하여 모형을 개발할 수 있다. Random effects Tobit 회귀모형에서 일반식을 다음과 같다.

$$L_i = \int_{-\infty}^{\infty} \prod_{t=1}^{T_i} \left[\Phi \left(\frac{a - x'_{it}\beta - \mu_i}{\sigma_\nu} \right) \right]^{I_{it}^a} \left[\Phi \left(\frac{x'_{it}\beta + \mu_i - b}{\sigma_\nu} \right) \right]^{I_{it}^b} \left[\frac{1}{\sigma_\nu} \phi \left(\frac{y_{it} - x'_{it}\beta - \mu_i}{\sigma_\nu} \right) \right]^{(1 - I_{it}^a - I_{it}^b)} \phi \left(\frac{\mu_i}{\sigma_\mu} \right) d\mu_i \quad (4)$$

$$\log L = \sum_{i=1}^n \log L_i \quad (5)$$

IV. 모형개발 및 분석

1. 데이터 수집 및 분석

본 연구에서 도심지 교차로 교통사고모형을 개발하기 위하여 총 교통사고율(사망교통사고+중상교통사고+경상교통사고)와 부상교통사고율(중상교통사고+경상교통사고)을 종속변수로 하고 교통량정보와 기하구조정보를 독립변수로 하여 데이터를 수집·분석하였다.

교통사고데이터는 일 년을 시간단위로 하여 반복 관찰된 횡단면 시계열데이터(cross sectional time series data)인 패널데이터를 이용하였는데 패널데이터는 데이터수집과정에서 발견할 수 없어 통제할 수 없는 데이터로 인한 영향을 최소화 할 수 있기 때문에 연도별/교차로별 교통사고율을 종속변수로 정하였다. 교통사고 데이터는 경찰청과 도로교통공단에서 개발하여 교통사고발생정보를 제공하고 있는 교통사고분석시스템 (traffic accident analysis system, TAAS)에서 2007년부터 2014년까지 8년간 대전광역시 주요 교차로 88개에 대한 연도별 교통사고발생건수를 수집하였으며, 각 연도별/교차로별 교통량을 이용하여 교통사고율을 산정하였다.

또한, 교통사고모형개발에 사용된 독립변수는 교통사고율에 영향을 미친다고 판단되는 요인을 지정하여 수집하였으며, 교통량의 경우 현장조사와 ‘대전광역시 교통조사·분석보고서’를 통하여 수집하였으며, 교차로 기하구조정보의 경우 현장방문조사와 Google Map의 Street View를 통해 수집하였다. 본 연구에서 교차로 교통사고모형을 개발하기 위해 이용한 독립변수는 교통량(AADT), 교차로에 접근하는 주간선도로 교통량(MAADT), 교차로에 접근하는 보조간선도로 교통량(SAADT), 보행신호 카운트다운기 운영여부(CD), 교차로 주변 토지이용현황(LAND), 교차로에 접근하는 주간선도로 제한속도(SPLMT), 교통섬 설치 유무(ISLAND), 교차로에 접근하는 주간선도로 차선수(NLANE), 중앙분리대 설치유무(BAR), 교차로 내 횡단보도수(PED), 좌회전 전용차로 설치 유무(LP), 우회전 전용차로 설치 유무(RP), 전방신호등 설치 유무(FS) 등으로 각 변수의 정의와 기초통계 분석결과는 <Table 1>과 같다.

〈Table 1〉 Definition and Descriptive Statistics of Variables

Variables	Definition	Mean	Standard Deviation	Minimum	Maximum
TOTAL	Total Accidents per 100-million Vehicles (/Intersection/Year)	42.085	30.332	0	230.320
FATAL	Fatal Accidents per 100-million VMT (/Intersection/Year)	0.651	1.673	0	11.604
IA	Serious Injury Accidents per 100-million VMT (/Intersection/Year)	14.726	11.953	0	76.7734
IB	Slight Injury Accidents per 100-million VMT (/Intersection/Year)	26.737	21.682	0	40
IAB	Injury Accidents per 100-million VMT (/Intersection/Year)	41.434	30.137	0	230.320
AADT	Annual Average Daily Traffic (/Intersection/Year)	72975.700	30435.300	21561	239925
LAADT	Ln(AADT)	11.108	0.426	9.979	11.908
MAADT	Major Road AADT	49091.800	21486.100	15308	221810
SAADT	Minor Road AADT	23884.000	14920.900	2587	71989
CD	Intersections with Countdown Timer	0.592	0.492	0	1
LAND	Land Usage (Residential 1, Commercial 2, Industrial 3, Other 4)	1.432	0.823	1	4
SPLMT	Speed Limits at Major Road	64.318	8.770	50	80
ISLAND	Traffic Island in Intersection	0.443	0.497	0	1
NLANE	Number of Lanes at Major Road	4.352	1.235	2	7
BAR	Median Barrier	0.591	0.492	0	1
PED	Number of Pedestrian Signal	3.284	1.271	0	6
LP	Left Turn with an Exclusive Turn Lane	0.636	0.481	0	1
RP	Right Turn with an Exclusive Turn Lane	0.591	0.492	0	1
GL	Turning Guideline	0.943	0.232	0	1
FS	Front Signal	0.409	0.492	0	1

교차로 교통사고모형개발 시 모형의 설명력 및 적합도를 높이기 위하여 독립변수간의 상호 관련성이 없거나 통계학적으로 낮아야 한다. 만약 독립변수간의 상호 관련성이 통계학적으로 존재하거나 높다면 독립변수의 고유한 특성이 중복적으로 종속변수인 교통사고발생빈도에 영향을 미쳐 표준오차가 큰 계수(coefficient)가 산정될 가능성이 커진다. 따라서 본 연구에서는 모형개발에 이용할 독립변수의 상호 관련성을 분석하기 위하여 다중공선성 분석(multicollinearity test)을 시행하였다.

다중공선성은 회귀분석에서 독립변수 중 서로 상관성이 높은 것이 포함되어 있는 경우 분산과 공분산의 행렬의 행렬식이 0에 가까운 값이 되어 회귀계수의 산정 정확도가 낮아지는 경우가 발생하게 되는데 이런 현상을 다중공선성이라 정의할 수 있다.

다중공선성 분석에서 독립변수간의 상호 관련성을 공차(tolerance)와 분산팽창요인(variance inflation factor, VIF)을 통하여 진단할 수 있다. 다중공선성 분석으로 산정된 독립변수의 공차는 0.01과 같거나 커야하고, VIF는 10.00과 같거나 작아야 각 독립변수간의 상호 관련성이 없거나 낮다고 판단할 수 있다. 본 연구에서 시행한 다중공선성 분석결과는 <Table 2>와 같으며, 모든 독립변수의 공차는 0.01보다 크고, VIF는 10.00보다 작아 상호 관련성이 매우 약한 것으로 분석되었다.

〈Table 2〉 Results of Multicollinearity Test

Variables	Total & Injury Accidents	
	Collinearity Statistics	
	Tolerance	VIF
LAADT	0.265	3.766
MAADT	0.448	2.230
SAADT	0.331	3.017
CD	0.860	1.163
LAND	0.910	1.099
SPLMT	0.575	1.739
ISLAND	0.775	1.290
NLANE	0.389	2.574
BAR	0.734	1.362
PED	0.764	1.310
LP	0.584	1.713
RP	0.687	1.457
GL	0.820	1.220
FS	0.885	1.130

2. Tobit 회귀모형을 이용한 교통사고모형 개발

본 연구에서는 대전광역시 주요 교차로의 교통사고모형을 개발하기 위하여 교통사고율(TOTAL과 IAB)을 종속변수로 하고 교통량과 교차로 기하구조를 독립변수로 하여 모형을 개발하였다. 모형개발은 계량경제(econometrics)와 통계분석 소프트웨어인 LIMDEP의 LNOGIT 4.0을 이용하였다.

주요 교차로의 교통사고모형은 중도절단회귀모형(censored regression model)인 Tobit 회귀모형을 이용하였으며, 수집된 데이터에서 관찰되어지지 않은 요인인 이질성(heterogeneity)을 극복하기 위하여 random effects Tobit 회귀모형을 적용하여 모형을 개발하였으며, random effects Tobit 회귀모형의 모형 적합성과 설명력을 분석하기 위하여 fixed effect Tobit 회귀모형을 각각 개발하여 모형 적용성을 비교·분석하였다.

또한, 모형개발 시 모든 독립변수를 모형에 적용하는 것이 불가능하기 때문에 F-통계량의 유의확률이 가장 낮게 하는 독립변수를 하나씩 선택하여 모형에 적용하는 단계적 방식(stepwise manner)으로 모형을 개발하였다(Lee et al., 2015)

본 연구에서 Random effects Tobit 회귀모형을 개발하기 위하여 200 Halton Draws를 이용한 시뮬레이션 기반 최우추정법(simulation-based maximum likelihood method)을 이용하였다. 그리고 random effects 모형의 계수 추정 시 밀도함수의 함수형태는 여러 확률분포를 고려할 수 있는데, 본 연구에서는 정규분포(normal distribution)가 가장 유효한 것으로 나타났다.

본 연구에서 대전광역시 주요 교차로의 교통사고모형개발 결과, 총 교통사고율과 부상교통사고율 모두에서 교통량(Logarithm of AADT), 제한속도(SPLMT), 차로수(NLANE), 토지이용현황(LAND), 우회전 전용차로 설치 유무(RP), 전방신호등 설치 유무(FS)가 교통사고 발생빈도에 영향을 미치는 변수로 나타났으며, 각 교통사고모형개발 결과는 <Table 3>과 <Table 4>와 같다.

3. 모형개발 결과 및 분석

<Table 3>과 <Table 4>는 random effects와 fixed effects Tobit 회귀모형을 이용하여 총 교통사고율과 부상교통사고율에 대한 교통사고모형 개발결과이며, 개발된 2가지 모형의 모형 적합도를 비교·분석하기 위하여 모형결정계수(Madalla Pseudo R_p^2), 로그-우도함수값(log-likelihood at convergence), 우도비(likelihood ratio ρ^2)를 산정하였다. 또한, 개발된 모형에서의 예측치와 실제 관측된 실측치와의 오차분포 특성을 분석하여 모형의 설명력을 파악하기 위하여 절대평균편차(mean absolute deviation, MAD), 평균절대백분율오차(mean absolute percentage error, MAPE), 그리고 평균제곱근오차(root mean square error, RMSE)을 산정하여 분석하였다.

<Table 3>에서 보는 바와 같이, 총 교통사고율에 대한 교통사고모형에서 random effects Tobit 회귀모형의 모형 적합도는 결정계수 0.418로 fixed effects Tobit 회귀모형의 결정계수 0.298보다 우수한 것으로 나타났다. 또한 로그-우도함수값과 우도비는 random effects 모형에서 각각 -3210.103과 0.056으로 fixed effects 모형의 -3276.138과 0.037보다 약간 개선이 된 것으로 분석되었다. 또한, 모형의 설명력 분석에서 random effects 모형의 MAD, MAPE, RMSE는 각각 19.533, 75.725, 26.886으로 fixed effects 모형의 20.725, 82.473, 27.267보다 우수하게 나타나 random effects 모형의 예측치와 실측치간의 오차가 작은 것으로 나타났다.

<Table 3> Model Estimation Results of Total Crash Rate

Variables	Tobit with Fixed Effects		Tobit with Random Effects				
	Coefficients	t-Statistics (P-value)	Distribution	Mean	t-Statistics (P-value)	Standard Deviation	t-Statistics (P-value)
Constant	175.944	5.331 (0.000)	-	122.403	64.275 (0.000)	-	-
Ln(AADT)	-7.715	-2.566 (0.010)	Normal	-4.629	-26.751 (0.000)	6.529	143.490 (0.000)
SPLMT	-1.386	-9.542 (0.000)	Normal	-1.038	-120.486 (0.000)	0.588	125.234 (0.000)
NLANE	10.270	8.522 (0.000)	Normal	10.195	143.123 (0.000)	0.058	1.829 (0.067)
LAND	4.158	3.094 (0.002)	Normal	1.307	16.004 (0.005)	8.377	134.270 (0.000)
RP	-12.014	-5.126 (0.000)	Normal	-11.505	-83.107 (0.000)	16.147	182.606 (0.000)
FS	-7.337	-3.307 (0.000)	Normal	-4.255	-32.610 (0.000)	2.811	28.827 (0.000)
Number of Observations	704						
AIC	9.330		9.202				
BIC	9.382		9.390				
Madalla Pseudo R_p^2	0.298		0.418				
Log-likelihood with Constant Only	-3400.612		-3400.612				
Log-likelihood at Convergence	-3276.138		-3210.103				
ρ^2	0.037		0.056				
MAD	20.725		19.533				
MAPE	82.473		75.725				
RMSE	27.267		26.886				

그리고 <Table 4>에서 보는 바와 같이, 부상교통사율에 대한 교통사고모형은 총 교통사고율 모형의 결과와 마찬가지로 random effects Tobit 회귀모형의 모형 적합도(결정계수: 0.429, 로그-우도함수값: -3198.905, 우도비: 0.058)와 모형 설명력(MAD: 19.238, MAPE: 79.548, RMSE: 26.499)이 fixed effects Tobit 회귀모형의 모형 적합도(결정계수: 0.303, 로그-우도함수값: -3268.845, 우도비: 0.037)와 모형 설명력(MAD: 20.548, MAPE: 83.238, RMSE: 27.120)보다 우수한 것으로 나타나 random effects Tobit 회귀모형이 fixed effects Tobit 회귀모형에 비해 교차로 교통사고율 대한 모형 적용성이 높은 것으로 나타났다.

또한, 교통사고모형에서 산정된 계수를 살펴보면, Random effects Tobit 회귀모형에서 산정된 계수(coefficient)는 변수밀도(parameter density)의 표준편차가 통계적으로 유의하여야 모형에서 유의미한 결과를 도출할 수 있다. 따라서 random effects 모형에서 산정된 변수의 계수는 t-값이 1.96보다 크거나 같고 표준편차가 0이 아니어야 한다. 만약 t-값이 1.96보다 작고 표준편차가 0과 같을 경우 산정된 계수는 fixed effects 모형에서 산정한 계수와 동일한 값을 가지게 된다(Lee et al., 2015)

<Table 4> Model Estimation Results of Injury Crash Rate

Variables	Tobit with Fixed Effects		Tobit with Random Effects				
	Coefficients	t-Statistics (P-value)	Distribution	Mean	t-Statistics (P-value)	Standard Deviation	t-Statistics (P-value)
Constant	175.319	5.339 (0.000)	-	111.197	92.030 (0.000)	-	-
Ln(AADT)	-7.779	-2.601 (0.009)	Normal	-2.398	-21.789 (0.000)	6.632	221.636 (0.000)
SPLMT	-1.373	-9.503 (0.000)	Normal	-1.006	-183.789 (0.000)	0.312	106.070 (0.000)
NLANE	10.302	8.590 (0.000)	Normal	6.905	152.846 (0.000)	5.122	239.075 (0.000)
LAND	3.995	2.986 (0.003)	Normal	2.068	39.094 (0.000)	8.112	197.766 (0.000)
RP	-11.863	-5.087 (0.000)	Normal	-11.824	-136.413 (0.000)	5.629	95.391 (0.000)
FS	-7.632	-3.457 (0.001)	Normal	-9.959	-120.633 (0.000)	2.820	44.219 (0.000)
Number of Observations	704						
AIC	9.309		9.170				
BIC	9.361		9.358				
Madalla Pseudo R_p^2	0.303		0.429				
Log-likelihood with Constant Only	-3396.082		-3396.082				
Log-likelihood at Convergence	-3268.845		-3198.905				
ρ^2	0.037		0.058				
MAD	20.548		19.238				
MAPE	83.238		79.548				
RMSE	27.120		26.499				

본 연구에서 산정한 random effects Tobit 회귀모형을 이용한 교통사고모형의 결과를 살펴보면, 총 교통사고율과 부상교통사고율에 대한 교통사고모형에서 교통량(Logarithm of AADT)의 평균은 각각 -4.629와 -2.398로 나타나 높은 유의함을 보였다. 이는 교통량이 증가할수록 교통사고율이 감소하는 것을 의미하며, 표준편차가 각각 6.529와 6.632로 유의한 값을 나타내는 것은 본 연구대상 88개 교차로의 교통량이 교통사고율에 각기 다른 영향을 미치는 것으로 분석이 가능하다.

제한속도(SPLMT)의 경우, 전체교통사고율과 부상교통사고율에 대한 교통사고모형에서 평균은 각각 -1.038과 -1.006이며, 표준편차는 각각 0.588과 0.312로 제한속도가 높아질수록 교통사고 발생빈도가 낮아지는 것으로 나타났다. 이는 제한속도가 1 unit 증가 시 교통사고율이 각각 1.038과 1.006 만큼 감소한다는 의미로 보통은 제한속도가 증가하면 교통사고율이 높아지는 현상과 반대되는 결과를 도출하였다. 하지만 대도시 주요교차로의 경우 교통량 증가, 좁은 차선폭 등으로 인해 사망사고와 부상사고는 감소하고 대물사고(PDO)는 증가할 수 있으므로 본 연구에서는 유의미한 것으로 판단된다. 또한, 다른 연구에서도 비슷한 결과를 도출한 경우가 있어 결과는 유의미하다고 할 수 있다(Lee et al., 2015; Greibe, 2003).

또한, 산정된 교통사고모형에서 차로수(NLANE)와 토지이용현황(LAND) 계수의 평균은 총 교통사고율에서 각각 10.195와 1.307, 부상교통사고율에서 각각 6.905와 2.068로 나타나 차로수가 많을수록 토지이용현황이 주거지역보다 산업지역일수록 교통사고발생빈도가 높은 것으로 나타났다.

이러한 교통사고율에 부정적인 영향을 미치는 요소와는 반대로 우회전 전용차로 설치 유무(RP)와 전방신호등 설치 유무(FS)는 총 교통사고율에서 각각 -11.505와 -4.255, 부상교통사고율에서 각각 -11.824와 -9.959로 나타나 우회전 전용차로를 설치할수록, 그리고 전방신호등을 설치할수록 교통사고율이 낮아지는 것으로 나타났다. 또한, 표준편차는 총교통사고율에서 각각 16.147과 2.811, 부상교통사고율에서 각각 5.629와 2.820으로 나타나 높은 유의함을 보이고 있는 것으로 나타났다.

V. 결 론

본 연구에서는 기존에 교통사고모형개발에 많이 적용하였던 포아송회귀모형이나 음이항회귀모형과 같은 가산모형과 로짓모형과 프로빗 모형과 같은 이산선택모형 대신 random effects Tobit 회귀모형을 적용하여 대도시 교차로에 대한 교통사고모형을 개발하여 교통사고 원인을 분석하였다. 연구대상 교차로의 교통량, 교차로 기하구조 등의 종류 및 설치형태가 각기 달라 교차로마다 이질성(heterogeneity)이 존재하는 것을 고려하여 random effects 모형을 적용하였으며, random effects의 적용성을 분석하기 위하여 fixed effects Tobit 회귀모형을 개발하여 두 모형을 비교하였다.

총 교통사고율과 부상교통사고율을 종속변수로 하는 교차로 교통사고모형의 개발결과, 두 모형 모두에서 교통량(LAADT), 제한속도(SPLMT), 차로수(NLANE), 토지이용현황(LAND), 우회전전용차로 설치유무(RP), 전방신호등 설치 유무(FS)가 교통사고율에 영향을 미치는 것으로 나타나 도심지 교통사고 특성을 반영하는 것으로 나타났다.

모형 적합도는 총 교통사고율에 대한 교통사고모형에서 random effects Tobit 회귀모형의 모형결정계수(R_p^2), 로그-우도함수값, 우도비(ρ^2)가 각각 0.418, -3210.103, 0.056으로 fixed effects Tobit 회귀모형의 0.298, -3276.138, 0.037보다 개선된 것으로 나타났다. 또한 모형의 설명력을 분석하기 위하여 산정한 MAD, MAPE, RMSE는 random effects Tobit 회귀모형에서 각각 19.533, 75.725, 26.886으로 fixed effects Tobit 회귀모형의 20.725, 82.473, 27.267보다 우수한 것으로 나타났다. 이는 random effects Tobit 회귀모형의 예측치와 실측치간

차이가 fixed effects Tobit 회귀모형의 차이보다 작다는 것으로 random effects Tobit 회귀모형의 설명력이 다소 우수하다고 판단할 수 있다.

또한, 부상교통사고율에 대한 교통사고모형에서 random effects Tobit 회귀모형의 모형 적합도(결정계수: 0.429, 로그-우도함수값: -3198.905, 우도비: 0.058)와 모형 설명력(MAD: 19.238, MAPE: 79.548, RMSE: 26.499)이 fixed effects Tobit 회귀모형의 모형 적합도 (결정계수: 0.303, 로그-우도함수값: -3268.845, 우도비: 0.037)와 모형 설명력(MAD: 20.548, MAPE: 83.238, RMSE: 27.120)보다 다소 우수한 것으로 나타나 총 교통사고율에 대한 교통사고모형과 마찬가지로 random effects Tobit 회귀모형이 fixed effects Tobit 회귀모형에 비해 교차로 교통사고율에 대한 모형 적용성이 높은 것으로 나타났다.

기존 연구의 교통사고모형개발은 교통사고 발생요인이 되는 독립변수에 대한 추정된 모수가 고정되어있다는 가정 하에 교통량과 기하구조의 종류 및 제원에 관계없이 동일한 값을 적용한 fixed effects 모형을 적용하여 교통사고모형을 개발하였다. 하지만 본 연구결과에서 본 바와 같이 기하구조의 이질성을 고려한 random effects Tobit 회귀모형이 기존 fixed effects Tobit 회귀모형보다 모형의 설명력과 적합도가 다소 뛰어난 것으로 분석되었다. 이러한 결과는 향후 도심지 교차로에서 발생하는 교통사고에 대한 발생원인 분석 시 실제 도로현실을 현실적으로 반영하여 개선한다면 도심지 교차로 교통안전성 향상 및 교통사고발생 감소에 기여할 수 있을 것으로 판단된다.

REFERENCES

- Anastasopoulos C. P., Mannering L. F., Shankar N. V. and Haddock E. J.(2012), "A study of factors affecting highway accident rates using the random-parameters tobit model," *Accident Analysis & Prevention*, vol. 45, pp.628-633.
- Anastasopoulos Ch. P., Tarko P. A. and Mannering L. F.(2008), "Tobit analysis of vehicle accident rates on interstate highways," *Accident Analysis and Prevention*, vol. 40, no. 2, pp.768-775.
- Bruno G.(2004), "Limited dependent panel data models: a comparative analysis of classical and Bayesian inference among econometric packages," *In Computing in Economics and Finance*, 41.
- Calzolari G. and Magazzini L.(2008), "Estimating Tobit models for panel data with autocorrelated errors," *In European Meeting of the Econometric Society*, Milan, pp.27-318.
- Calzolari G. Magazzini L. and Mealli F.(2001), *Simulation-based estimation of Tobit model with random effects*, *Econometric Studies: A Festschrift in Honour of Joachim Frohn*, pp.349-369.
- Greene W.(2004), "Fixed effects and bias due to the incidental parameters problem in the Tobit model," *Econometric Reviews*, vol. 23, no. 2, pp.125-147.
- Greibe P.(2003), "Accident prediction models for urban roads," *Accident Analysis and Prevention*, vol. 35, pp.273-285.
- Henningsen A.(2010), "Estimating censored regression models in R using the censReg package," Available at <<http://cran.r-project.org/web/packages/censReg/vignettes/censReg.pdf>>.
- Islam B. M. and Hernandez S.(2011), "An empirical analysis of fatality rates for large truck involved crashes on interstate highways," *In 3rd International Conference on Road Safety and Simulation*, pp.14-16.
- Lee S. H., Park M. H. and Woo Y. H.(2015), "A study on developing crash prediction model for urban

intersections considering random effects,” *Journal of the Korea Institute of ITS*, vol. 14, no. 1, pp.85-93.

Qi Y., Smith L. B. and Guo J.(2007), “Freeway accident likelihood prediction using a panel data analysis approach,” *Journal of transportation engineering*, vol. 133, no. 3, pp.149-156.

Xu X., Wong C. S. and Choi K.(2014), “A two-stage bivariate logistic-Tobit model for the safety analysis of signalized intersections,” *Analytic Methods in Accident Research*, vol. 3, pp.1-10.