

Tobit 모형을 이용한 간선도로 사고 요인 분석

Analysis of Accident Factors at Arterial Roads Using Tobit Model

김경환 Kim, Kyung Hwan | 정회원 · 충북대학교 공과대학 도시공학과 박사과정 · 주저자 (E-mail : as1836@nate.com)
박병호 Park, Byung Ho | 정회원 · 충북대학교 공과대학 도시공학과 교수 · 교신저자 (E-mail : bhpark@chungbuk.ac.kr)

ABSTRACT

PURPOSES : The intents of the study are to identify the accident factors and to demonstrate the potentials of tobit model as a tool to study the number of accidents on arterial roads segments.

METHODS : This paper uses a tobit regression as a methodology to analyze the factors affecting the number of accidents. In pursuing the above goal, this study gives particular attentions to analyzing the data of 2,446 accidents (1,610 in major arterial roads and 836 in minor arterial roads) occurred on arterial roads in 2007 to 2010.

RESULTS : First, 3 accident models which were classified by total arterial roads, major arterial roads and minor arterial roads, and were all statistically significant were developed. Second, the exclusive right-turn lane as common variable, and the number of accident, traffic volume, number of lanes, link length, rate of median, number of entrances, number of pedestrian crossings, number of curves, number of bus stops and exclusive left-turn as specific variables of the models were selected. Finally, the paired sample t-test could not be rejected the null hypotheses of three types of models.

CONCLUSIONS : Using data from vehicle accidents on arterial roads, the estimation results show that many factors related to roadway geometrics and traffic characteristics significantly affect to the number of accidents.

Keywords

accident model, arterial road, paired sample t-test, RMSE(Root Mean Square Error), tobit model

Corresponding Author : Park, Byung Ho
ChungBuk National University, 52 Naesudong-ro, Cheongju,
361-763, Korea
Tel : +82.43.261.2496 E-mail : bhpark@chungbuk.ac.kr

International Journal of Highway Engineering
<http://www.ijhe.or.kr/>
ISSN 1738-7159 (Print)
ISSN 2287-3678 (Online)

1. 서론

1.1. 연구 배경 및 목적

교통사고 예방활동 및 법규준수의식 향상으로 우리나라의 교통사고 사망자 수는 감소추세를 보이고 있다. 하지만 우리나라 자동차 1만 대당 교통사고 발생건수는 2009년 기준 111.4건으로 OECD 평균인 55.1건에 비해 2.0배 높은 것으로 나타났다. 또한 자동차 1만 대당

사망자 수는 2.8명으로 자료가 파악되지 않은 멕시코와 칠레를 제외한 32개국 중 30위를 차지하였으며, OECD 평균인 1.2명에 비해 약 2.3배 많이 발생한 것으로 나타났다(도로교통공단, 2011).

이에 교통사고를 감소시키고 사망자를 줄이기 위한 연구가 지속적으로 이루어지고 있다. 교통사고는 다양한 요인에 의해 발생하며, 교차로와 가로구간을 대상으

로 교통사고와 사고요인과의 관계를 분석하고 사고를 예측하는 연구가 활발히 진행되었다. 기존의 교통사고 관련 연구에 따르면 사고에 영향을 미치는 요인은 교통운영요인, 기하구조 요인, 도로 환경 요인 및 인적요인 등으로 구분할 수 있다.

대부분의 연구에서는 가산자료 모형(음이항 및 포아송 회귀모형), 다중선형 회귀모형, 로짓 및 프로빗 모형 등을 이용하여 예측 모형을 구축하였다. 하지만 사고건수는 항상 0 이상의 값을 가지므로 일정한 방향으로 한계 값을 가지는 분포를 띠고 있다. 즉, 사고건수는 제한된 위의 값을 갖기 때문에 일반적인 회귀모형에서 가정하고 있는 정규분포와 달라 회귀계수가 일관성 없는 추정치(inconsistent estimates)를 갖게 되므로 잘못된 예측결과를 가져올 수 있다.

이에 이 연구에서는 Tobin(1958)이 개발한 중도절단 회귀모형(censored regression model)인 Tobit 모형을 이용하여 간선도로의 사고 요인을 분석하고, 사고모형을 구축하는데 그 목적이 있다.

1.2. 연구 내용 및 방법

이 연구는 청주시 간선도로에서 발생한 교통사고를 대상으로 하였으며, 도로교통공단의 교통사고분석시스템(TAAS : traffic accidents analysis system)을 통해 2007~2010년에 발생한 사고 자료를 수집하였다. 항공사진과 AutoCAD를 이용하여 기하구조자료를 구축하고, 현장조사를 통해 보완하였다. 교통량은 교통량 조사자료, 교통영향평가 등 교통 관련 보고서의 자료를

활용하였다.

연구의 진행은 다음과 같다. 제1장에서 연구의 배경 및 목적을 제시하고, 제2장에서 기존 문헌을 검토하여 차별성을 제시하였다. 제3장에서 분석의 틀을 설정하였으며, 그 다음으로 제4장에서 Tobit 모형을 이용하여 사고 요인을 분석하고 사고모형을 구축하였다. 마지막으로 제5장 결론에서 분석 결과를 요약하였다.

2. 기존문헌 고찰

2.1. 교통사고 모형

김경환 등(2010)은 간선도로에서 발생한 사고를 조사하고, 운전자의 운전행태를 3개의 유형(직진, 차선변경 및 기타)로 구분하여 사고모형을 개발하였다. 다중선형 회귀모형을 이용하여 유형별 사고모형을 구축하였으며, 직진, 차선변경 및 기타 유형 모두 좌회전 차로 수가 사고건수와 EPDO에 영향을 미치는 것으로 분석되었다. 또한 RMSE 분석 결과, 유형을 더미변수로 처리한 통합 모형에 비해 운전 유형별 모형이 더 적합한 것으로 평가되었다.

김응철 등(2010)은 지방부 2차로 도로의 사고예측계수(AMF : accident modification factor)를 개발하였다. 사고예측계수 항목으로는 횡단보도와 진출입구의 밀도, 지형(산악지형), 토지이용도(주거), 중앙분리대의 설치여부를 선정하였으며, 모형을 이용하여 계수를 산출하였다. 산출된 사고예측계수는 국외의 사고예측계수를 적용했을 경우와 비교분석하였으며, 실제 사고발생 위험도가 높은 도로구간에 적용하였을 경우에 국외 계수보다 반영도가 더 좋은 것으로 평가되었다.

김상엽 등(2011)은 고속도로를 구간분할법을 사용하여 직선부, 곡선부, 연속곡선부로 나누고, 이를 군집분석과 판별분석을 이용하여 유형을 구분하여 총 9개의 사고모형을 구축하였다. 가산자료 모형(포아송 및 음이항)을 이용하여 사고빈도모형을 개발하였으며, 과대산포 검정 결과 포아송모형보다 음이항모형이 적합한 것으로 평가되었다.

한수산 등(2011)은 원형교차로의 교통사고를 대상으로 가산자료 모형(포아송 및 음이항)을 이용하여 사고유형별(차대차, 차대사람 등) 사고심각도 요인을 분석하였다. 사고건수 및 EPDO 모형의 공통변수로는 교통량이 채택되었으며, 차대차의 경우엔 우회전 별도차로 수가 사고건수 및 사고심각도에 가장 많은 영향을 주는 것으로 평가되었다.

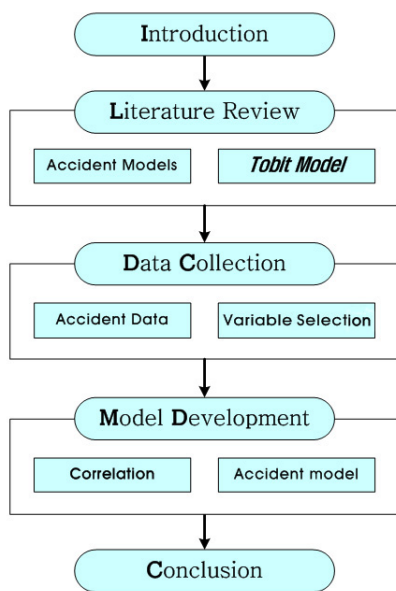


Fig. 1 Study Flowchart

박민수 등(2012)은 평면곡선 구간과 종단선형과의 조합을 통해 3개의 복합곡선 형태로 유형을 구분하여 고속도로의 교통사고 특성을 분석하였다. 분석결과 종단곡선이 없는 경우가 사고율이 가장 높았으며, 그 다음으로는 종단곡선이 1개인 경우, 그리고 종단곡선이 2개 이상인 경우로 나타났다.

오주택 등(2012)은 지방부 도로를 2차로와 다차로 구간으로 구분하여 사고특성을 분석하고 모형을 개발하였다. 도로구간의 교통사고는 “0”의 빈도가 높게 나타난다는 점을 고려하여, 사고의 비선형적 요소를 설명하는 허들모형(hurdle model)을 이용하여 모형을 구축하였다.

Anastasopoulos 등(2008)은 인디애나 도시부 주간선 도로의 사고자료를 이용하여 Tobit 모형의 가능성을 검토하였다. 도로구간을 337구간으로 구분하여 교통사고율 모형을 개발하였으며, Maddala Pseudo- R^2 가 0.948로 나타나, 모형이 적합한 것으로 평가하였다. 또한 사고가 0건인 구간과 1건 이상인 구간에 대한 민감도 분석을 수행하였다.

Anastasopoulos 등(2012a)은 도시부 주간선 도로를 대상으로 random-parameters Tobit 모형과 fixed-parameters Tobit 모형을 비교·분석하였다. 포장 특성, 도로 기하구조, 교통류 특성 등을 변수로 사용하여 Tobit 모형을 개발하였다. 개발된 두 모형을 비교한 결과, fixed-parameters Tobit 모형에 비해 random-parameters Tobit 모형이 더 적합한 것으로 평가되었다.

Anastasopoulos 등(2012b)은 워싱턴 다차로 도로의 사고자료를 이용하여 사고 심각도 모형을 구축하였다. 모형 구축 결과 univariate tobit 모형에 비해 multivariate tobit 모형이 더 적합한 것으로 나타났으며, MSE(mean square error)를 통해 multivariate 음이항 모형과 비교한 결과 multivariate tobit이 더 적합한 것으로 분석되었다.

2.2. Tobit 모형

사고건수와 같이 종속변수가 일정 범위 안에서는 관측될 수 없는 제한된 값, 즉 중도 절단되는 특성을 갖는 회귀모형에는 중도절단회귀모형(censored regression model)과 절단회귀모형(truncated regression model)이 있다. 이 중 전자를 Tobit 모형이라 부르며, 중도절단 회귀모형을 최초로 분석한 토빈(Tobin, 1958)에 의해 제안되었다. Tobit 모형은 상단수의 종속변수가 0인 경우에 유용하게 이용되는 것으로 알려져 있으며, 일반적으로 Tobit 모형은 Eq. (1)과 같이 표현될 수 있다.

$$y_i^* = \beta x_i + \varepsilon_i, \quad i = 1, 2, \dots, n$$

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

여기서, n 은 표본의 크기, y_i^* 는 가상적인 종속변수, y_i 는 관찰된 종속변수, x_i 는 설명변수, 그리고 ε_i 는 오차 항을 의미한다. 일반적인 회귀모형과 유사하지만, Tobit 모형에서는 관찰된 모든 종속변수가 y_i^* 값이 아니라라는 점에서 차이가 있다. 즉 $y_i^* > 0$ 이면 $y_i = y_i^*$ 가 관찰되고, $y_i^* \leq 0$ 이면 $y_i = 0$ 으로 관찰된다.

Eq. (1)에서 y_i^* 의 값을 관찰할 수 있다면 일반적인 최소자승법을 적용하여 β 의 값을 추정할 수 있으나, 관측되는 종속변수 $y_i \geq 0$ 에서만 관측가능하다. 즉 y_i^* 대신 y_i 를 이용하여 최소자승법으로 추정하면 $E(\varepsilon_i) \neq 0$ 이므로 β 의 일치추정량을 얻을 수 없게 된다.

이와 같이 절단된 자료에 대해서는 최대우도법(maximum likelihood)을 이용하여 일치추정량을 얻을 수 있다. 이 경우 전체표본에 대한 우도함수식은 Eq. (2)와 같다. 여기서, ϕ 는 표준정규분포의 밀도함수(density function)를 나타내며, Φ 는 표준정규분포의 누적분포함수를 나타낸다.

$$L = \prod_{y_i > 0} \phi(y_i) \prod_{y_i = 0} \Phi(y_i) \quad (2)$$

일반적으로 이 우도함수에 로그 값을 취해 얻어지는 $\log L$ 을 극대화하여 계수 추정치 β 를 구할 수 있으며, 이는 Eq. (3)으로 표현될 수 있다.

$$\ln L = \sum_0 \ln(1 - \Phi(\frac{x_i \beta}{\sigma_u})) + \sum_0 [\ln \sigma_u + \ln \phi(\frac{y_i - x_i}{\sigma_u})] \quad (3)$$

2.3. 연구의 차별성

이 연구는 중도절단회귀모형인 Tobit 모형을 이용하여 모형을 구축하였다는 점에서 차별성이 있다. 기존 연구에서는 다중선형회귀모형, 가산자료 모형(포아송 및 음이항), 로짓 및 프로빗 모형 등을 이용하여 사고건수 및 사고심각도 모형을 개발하였다. 그러나 사고건수의 경우에는 음의 값이 나올 수 없으므로 일정한 영역에서만 관찰되는 변수를 종속변수로 하는 중도절단회귀모형이 더 적합한 것으로 판단되었다.

또한 이 연구에서는 교차로를 기준으로 간선도로 구간을 세분화하여 교차로 사고를 제외하고자 하였다. 간

선도로를 세분화하는 과정에서 사고건수가 0인 구간이 다수 발생하였으나, Tobit 모형은 다량의 종속변수가 0인 경우에 유용하게 이용되는 것으로 알려져 있다.

238건에 비해 1.9배(194건) 높은 것으로 나타났으며, 간선도로의 총 연평균 사고건수는 612건으로 분석되었다.

3. 분석 틀의 설정

3.1. 자료 수집

교통사고 data는 도로교통공단의 교통사고분석시스템을 통해 청주시 간선도로에서 발생한 4년간(2007~2010년)의 사고 자료를 수집하였다. 구간 길이, 중앙분리대 비율, 진·출입구수, 횡단보도수, 굴곡점수, 정류장수, 전용차로 유·무 및 평지여부 등 기하구조 요소들은 항공사진, AutoCAD를 이용하여 구축한 후 현장조사를 통해 보완하였으며, 교통량은 교통 관련 보고서의 조사 자료를 활용하였다.

연도별 교통사고 건수는 Table 1과 같다. 2010년의 사고건수가 714건으로 가장 많았으며, 2007년부터 연평균 14.4% 증가하는 추세로 조사되었다. 연평균 사고건수는 주간선도로가 403건으로 보조간선도로

3.2. 변수 정의

종속변수와 독립변수는 기존 문헌 고찰을 통해 사고에 영향을 미칠 것으로 기대되고 자료 수집이 용이한 요인을 선정하였으며, 이는 Table 2와 같이 정의하였다. 종속변수는 가장 일반적으로 사용되어지는 연평균 사고건수를 사용하였다.

독립변수는 자료 수집이 어려운 인적요인을 제외한 교통운영 요인(교통량), 기하구조 요인 및 도로 환경 요인으로 선정하였다. 또한 기하구조 및 도로 환경 요인은 해당연도의 자료 수집 및 연도별 변화를 반영하는데 어려움이 많고 분석을 단순화하기 위하여, 변화가 없다는 가정 하에 2010년을 기준으로 자료를 구축하였다.

4. 모형개발

교통사고 모형은 도로유형에 따라 전체 간선도로(주간선+보조간선), 주간선도로 및 보조간선도로로 구분하여 3개의 모형을 구축하였다. 또한 주요교차로를 기준으로 세분화하여 교차로사고가 포함되지 않도록 고려하여, 총 474구간으로 세분화하였다. 도로유형별 연평균 사고건수를 종속변수로 선정하였으며, 사고건수에 영향을 미치는 요인을 분석하기 위해 NLOGIT 3.0을 이용하여 Tobit 모형을 구축하였다. 하한 값은 0으로 설정하였다.

Table 1. Number of Accidents by Year

(Unit : Number)

Classification	2007	2008	2009	2010	Mean
Major arterial roads	302	388	444	476	403
Minor arterial roads	172	205	221	238	209
SUM	474	593	665	714	612

Table 2. Definition of Variables

Variables	Symbol	Definition	Mean	Range
Number of accident	Y_1	Average annual number of accidents	0.73	0.00~11.00
Traffic volume	X_1	$\ln(\text{Traffic volume of link})$	9.31	8.01~10.69
Number of lanes	X_2	Number of lanes at link	2.15	1.00~4.00
Link length	X_3	Link length	435.33	25.28~1,866.23
Rate of median	X_4	Median length/Total length	7.84	0.00~91.22
Number of entrances	X_5	Number of entrances at link	2.55	0.00~16.00
Number of pedestrian crossings	X_6	Number of pedestrian crossings at link	2.43	0.00~13.00
Number of curves	X_7	Number of curves at link	0.25	0.00~8.00
Number of bus stops	X_8	Number of bus stops at link	0.75	0.00~5.00
Exclusive left-turn	X_9	0 if there is no exclusive left-turn, 1 otherwise	0.55	0.00~1.00
Exclusive right-turn	X_{10}	0 if there is no exclusive right-turn, 1 otherwise	0.10	0.00~1.00
Grade	X_{11}	0 if maximum grade is less than 2%, 1 otherwise	0.46	0.00~1.00

The value of the average and range were based on the total arterial roads(major and minor arterial roads)

Table 3. Result of Correlation Analysis for Total Arterial Roads

Classification	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
Y_1	1	-	-	-	-	-	-	-	-	-	-	-
X_1	0.138**	1	-	-	-	-	-	-	-	-	-	-
X_2	0.015	-0.721**	1	-	-	-	-	-	-	-	-	-
X_3	0.173**	0.232**	-0.218**	1	-	-	-	-	-	-	-	-
X_4	0.097*	-0.115*	0.156**	0.177**	1	-	-	-	-	-	-	-
X_5	0.193**	0.420**	-0.399**	0.552**	-0.076	1	-	-	-	-	-	-
X_6	0.188**	0.398**	-0.397**	0.598**	-0.152**	0.737**	1	-	-	-	-	-
X_7	-0.034	0.021	-0.060	0.404**	-0.020	0.157**	0.156**	1	-	-	-	-
X_8	0.196**	0.110*	-0.082	0.264**	-0.014	0.279**	0.195**	0.252**	1	-	-	-
X_9	0.161**	-0.194**	0.339**	-0.120**	0.142**	-0.076	-0.116*	0.025	0.092*	1	-	-
X_{10}	0.099*	-0.070	0.089	-0.006	0.128**	-0.018	-0.066	-0.116*	0.011	0.121**	1	-
X_{11}	-0.036	-0.043	0.061	-0.099*	0.042	-0.132**	-0.156**	-0.064	-0.102*	0.012	0.019	1

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

4.1. 전체 간선도로(주간선+보조간선)

Pearson 상관계수를 통해 종속변수와 독립변수간의 상관성을 분석하였으며, 그 결과는 Table 3과 같다. 사고건수와 유의한 상관관계를 갖는 변수로는 교통량(X_1)과 구간 길이(X_3), 중앙분리대 비율(X_4), 진출입구수(X_5), 횡단보도수(X_6), 정류장수(X_8), 좌회전 전용차로 유·무(X_9) 및 우회전 전용차로 유·무(X_{10})가 채택되었다. 유의한 모든 변수는 사고건수와 양의 관계를 갖는 것으로 평가되었다. 또한 모든 변수들의 VIF 값이 10 이하로 다중공선성의 문제는 없는 것으로 나타났다.

Table 4에서 보는 바와 같이 모형 구축 결과, 교통량(X_1), 차로수(X_2), 중앙분리대 비율(X_4), 횡단보도수(X_6), 굴곡점수(X_7), 정류장수(X_8), 좌회전 전용차로 유·무(X_9) 및 우회전 전용차로 유·무(X_{10})가 독립변수로 채택되었고, 이 변수들의 p값은 신뢰수준 90%(α

Table 4. Result of Tobit Model for Total Arterial Roads

Variables	Coefficient	z value	p value
Intercept	-14.675	-3.025	0.002
X_1	1.183	2.470	0.014
X_2	0.651	1.987	0.047
X_4	0.020	2.290	0.022
X_6	0.279	4.318	0.000
X_7	-0.403	-1.864	0.062
X_8	0.408	2.802	0.005
X_9	0.587	1.836	0.066
X_{10}	0.833	1.951	0.051
σ^2	2.464		
Log likelihood	-550.397		

=0.10) 기준에 유의한 것으로 평가되었다.

굴곡점 수를 제외한 모든 변수는 사고건수와 양의 관계를 갖는 것으로 분석되었다. 교통량은 기존 문헌의 결과와 동일한 의미 갖는 것으로 판단되었다. 차로수, 중앙분리대 비율, 횡단보도수가 증가하고, 좌회전 전용차로 및 우회전 전용차로가 설치되어 있을수록 차량과 차량, 차량과 보행자, 차량과 중앙분리대의 상충기회가 증가하여 사고가 많아지는 것으로 평가되었다. 반면 음의 관계에 있는 굴곡점 수가 증가할수록 차량의 속도가 감소하여 사고가 감소하는 것으로 판단되었다.

4.2. 주간선도로

전체 간선도로와 동일한 방법으로 Pearson 상관계수를 통해 종속변수와 독립변수간의 상관성을 분석하였다. 사고건수와 유의한 상관관계를 갖는 변수로는 교통량(X_1), 진출입구수(X_5), 횡단보도수(X_6) 및 정류장수(X_8)가 채택되었으며, 결과는 Table 5와 같다. 유의한 상관관계를 갖는 변수는 사고건수와 양의 관계를 갖는 것으로 평가되었으며, 다중공선성의 문제는 없는 것으로 나타났다.

Tobit 모형의 분석 결과는 Table 6과 같다. 교통량(X_1), 구간 길이(X_3), 중앙분리대 비율(X_4), 진출입구수(X_5), 정류장수(X_8) 및 우회전 전용차로 유·무(X_{10})가 독립변수로 채택되었다. 또한 이 변수들의 p값은 신뢰수준 90%($\alpha=0.10$) 기준에 유의한 것으로 평가되었다.

구간 길이를 제외한 모든 변수는 사고건수와 양의 관계를 갖는 것으로 분석되었다. 교통량은 기존 문헌의 결

Table 5. Result of Correlation Analysis for Major Arterial Roads

Classification	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
Y_1	1	-	-	-	-	-	-	-	-	-	-	-
X_1	0.240**	1	-	-	-	-	-	-	-	-	-	-
X_2	-0.061	-0.522**	1	-	-	-	-	-	-	-	-	-
X_3	0.057	0.033	0.061	1	-	-	-	-	-	-	-	-
X_4	0.056	0.063	0.082	0.518**	1	-	-	-	-	-	-	-
X_5	0.303**	-0.013	-0.048	0.204**	-0.061	1	-	-	-	-	-	-
X_6	0.183**	0.055	-0.084	0.090	-0.213**	0.490**	1	-	-	-	-	-
X_7	-0.021	-0.004	0.013	0.376**	0.136*	0.109	0.134*	1	-	-	-	-
X_8	0.171**	-0.117*	0.031	0.171**	-0.093	0.350**	0.336**	0.197**	1	-	-	-
X_9	0.044	-0.024	0.110	-0.113*	-0.197**	0.021	0.000	-0.037	0.039	1	-	-
X_{10}	0.019	-0.020	0.234**	0.168**	0.154**	-0.074	-0.171**	-0.117*	-0.09	-0.051	1	-
X_{11}	0.014	0.028	-0.048	-0.004	-0.016	0.041	0.001	0.022	0.046	-0.035	0.042	1

** . Correlation is significant at the 0.01 level (2-tailed)

* . Correlation is significant at the 0.05 level (2-tailed).

과와 동일한 것으로 나타났으며, 중앙분리대의 비율이 높고 진·출입구 및 정류장수가 증가할수록 상충기회가 상승하여 사고건수가 증가하는 것으로 평가되었다.

Table 6. Result of Tobit Model for Major Arterial Roads

Variables	Coefficient	z value	p value
Intercept	-17.970	-5.087	0.000
X_1	1.745	4.651	0.000
X_3	-0.001	-1.785	0.074
X_4	0.015	1.889	0.059
X_5	0.536	5.409	0.000
X_8	0.499	1.792	0.073
X_{10}	0.988	1.751	0.080
σ^2	2.524		
Log likelihood	-413.185		

4.3. 보조간선도로

종속변수와 독립변수간의 상관관계를 분석하기 위해 신뢰수준을 95%($\alpha=0.05$)로 하여, Pearson 상관계수를 통해 상관성을 분석하였으며, 그 결과는 Table 7과 같다. 사고건수와 유의한 상관관계를 갖는 변수로는 구간 길이(X_3), 진출입구수(X_5), 횡단보도수(X_6) 및 정류장수(X_8)가 채택되었다. 유의한 모든 변수는 사고건수와 양의 관계를 갖는 것으로 평가되었다. 또한 모든 변수들의 VIF 값이 10 이하로 다중공선성의 문제는 없는 것으로 나타났다.

Tobit 모형을 이용하여 분석한 결과는 Table 8과 같다. 횡단보도수(X_6)와 우회전 전용차로 유·무(X_{10})가 독립변수로 채택되었고, 이 변수들의 p값은 신뢰수준 90%($\alpha=0.10$) 기준에 유의한 것으로 평가되었다.

Table 7. Result of Correlation Analysis for Minor Arterial Roads

Classification	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}
Y_1	1	-	-	-	-	-	-	-	-	-	-	-
X_1	0.073	1	-	-	-	-	-	-	-	-	-	-
X_2	-0.007	-0.827**	1	-	-	-	-	-	-	-	-	-
X_3	0.224**	0.345**	-0.227**	1	-	-	-	-	-	-	-	-
X_4	0.057	0.110	-0.152*	0.015	1	-	-	-	-	-	-	-
X_5	0.260**	0.519**	-0.417**	0.623**	0.215**	1	-	-	-	-	-	-
X_6	0.343**	0.436**	-0.372**	0.710**	0.113	0.761**	1	-	-	-	-	-
X_7	-0.044	0.012	-0.017	0.455**	0.095	0.167*	0.155*	1	-	-	-	-
X_8	0.242**	0.138	-0.078	0.292**	-0.074	0.279**	0.200**	0.263**	1	-	-	-
X_9	0.141	-0.145	0.121	-0.055	0.093	0.006	0.072	0.095	0.126	1	-	-
X_{10}	0.146	-0.070	-0.001	-0.086	0.152*	0.006	-0.036	-0.097	0.059	0.117	1	-
X_{11}	-0.093	-0.053	0.030	-0.120	0.136	-0.151*	-0.154*	-0.079	-0.138	-0.084	-0.010	1

** . Correlation is significant at the 0.01 level (2-tailed)

* . Correlation is significant at the 0.05 level (2-tailed).

채택된 변수는 모두 양의 관계를 갖는 것으로 분석되었다. 횡단보도수가 증가하고 우회전 전용차로가 설치되어 있을수록 사고건수는 증가하는 것으로 나타났다. 이는 우회전 전용차로 설치에 따라 상충기회가 상승하며, 특히 횡단보도가 증가할수록 차량간 상충뿐 아니라 보행자와의 상충기회가 증가하고, 운전행태의 변화가 발생하여 사고에 영향을 미치는 것으로 평가되었다.

Table 8. Result of Tobit Model for Minor Arterial Roads

Variabe	Coefficient	z value	p value
Intercept	-2.374	-4.996	0.000
X_6	0.341	4.339	0.000
X_{10}	1.201	1.844	0.065
σ^2		2.137	
Log likelihood		-179.419	

4.4. 모형의 검증

이 연구에서는 모형의 개발과정에서 적용된 독립변수를 구축된 모형식에 적용하여 실제치와 예측치와의 적합정도를 판단하는 방법을 통해 사고모형을 검증하였다. 이를 위해 대응표본 t 검정(paired sample t-test)을 시행하여 실제치와 예측치 간의 차이 여부를 통계적으로 입증하였다.

대응표본 t 검정 결과는 Table 9와 같다. 3개 모형의 상관계수는 0.379~0.431사이로 있으며, 유의확률은 0.667~0.996사이로 분석되었다. 즉 모든 귀무가설을 기각하지 못하여 실제치와 예측치 간에 차이가 있다고 할 수 없는 것으로 평가되었다.

또한 평균제곱근오차(RMSE : root mean square error)를 통해 예측치와 실제치와의 오차분포 특성을 분석하였다. RMSE의 산정식은 다음과 같다.

$$RMSE = \sqrt{\frac{\sum_i (t_i - T_i)^2}{N}} \quad (4)$$

여기서, t_i 는 예측된 사고건수, T_i 는 실제 관측된 사고

건수, 그리고 N 은 구간수를 나타낸다.

RMSE 분석은 가산자료(음이항 및 포아송) 모형과의 비교분석을 병행하여 실시하였으며, 결과는 Table 10과 같다. Tobit 모형을 이용한 전체 간선도로 모형은 1.091, 주간선도로 모형은 1.322, 그리고 보조간선도로 모형은 0.868로 분석되었다. 즉 보조간선도로의 경우엔 보조간선도로 모형, 주간선도로의 경우엔 전체 간선도로 모형을 적용하는 것이 사고건수를 예측하는데 있어 오차를 더 줄일 수 있는 것으로 평가되었다. 하지만 변수가 서로 다르게 채택된 점을 감안하여, 자료의 확보 정도에 따라 모형을 선택하여 사고건수를 예측해야 할 것으로 판단된다.

또한 가산자료 모형 구축 결과 포아송보다 음이항 회귀모형이 더 적합한 것으로 분석되었으며, RMSE 분석 결과 음이항 모형보다 Tobit 모형의 RMSE가 더 낮은 것으로 평가되었다.

Table 10. Result of RMSE by Type

Classification	Tobit	Negative binomial
Arterial roads	1.091	1.267
Major arterial roads	1.322	2.144
Minor arterial roads	0.868	1.124

5. 결론

이 연구는 Tobit 모형을 이용하여 간선도로의 교통사고 요인을 분석하였다. 도로유형별 사고건수를 종속변수로 선정하여 사고모형을 구축하였으며, 주요 결과는 다음과 같다.

첫째, 전체 간선도로, 주간선도로 및 보조간선도로로 구분된 3개의 사고모형이 개발되었다. 변수들의 p값은 신뢰수준 90%($\alpha=0.10$) 기준에 유의한 것으로 분석되었다. 또한 Tobit 모형의 σ^2 가 2.464, 2.524 및 2.137로 나타나 통계적으로 유의한 것으로 평가되었다.

둘째, 3개 모형의 공통변수로는 우회전 전용차로가 채택되었다. 그리고 특정변수로는 전체 간선도로의 경우엔 교통량, 차로 수, 중앙분리대 비율, 횡단보도 수, 굴곡점

Table 9. Paired sample t-test by type

Classification	Paired sample t-test					t-value	Coefficient of correlation	p-value (two-tailed)
	Mean	Standard deviation	Error of mean	95% Confidence interval				
				Minimum	Maximum			
Total arterial roads	0.011	1.092	0.050	-0.088	0.109	0.213	0.397	0.831
Major arterial roads	0.000	1.325	0.076	-0.150	0.150	0.005	0.431	0.996
Minor arterial roads	0.029	0.870	0.066	-0.102	0.160	0.431	0.379	0.667

수, 정류장수 및 좌회전 전용차로 유·무. 주간선도로의 경우엔 교통량, 구간 길이, 중앙분리대 비율, 진·출입구 수 및 정류장수, 보조간선도로의 경우엔 횡단보도수가 선정되었다. 구간 길이와 굴곡점수를 제외한 모든 변수들은 사고건수와 양의 관계를 갖고 있으며, 이 변수들은 기존 연구의 결과와 동일한 것으로 판단되었다.

마지막으로 대응표본 t 검정 결과 구축된 3개 모형 모두 귀무가설을 기각하지 못하였다. 즉, 실측치와 예측치의 차이가 있다고 할 수 없는 것으로 평가되었다. 또한 RMSE 분석 결과 보조간선도로 모형이 가장 적합한 것으로 분석되었으며, 그 다음으로는 전체 간선도로 모형, 그리고 주간선도로 모형의 순으로 평가되었다.

이 연구에서는 Tobit 모형을 이용하여 사고 요인을 분석하였다. 그러나 단순 사고건수만을 종속변수로 사용하였으며, 인적요인을 고려하지 못하였다는 한계를 지니고 있다. 따라서 향후 더욱 설명력이 있는 연구를 위해서는 인적요인의 자료를 확보하여야 하며, 단순 사고건수뿐 아니라 사고심각도에 영향을 미치는 요인에 대한 분석 및 모형 개발도 함께 연구되어야 할 것이다. 또한 변수들의 표준화 계수를 산출하여 변수들간의 상대적 중요도에 대한 분석도 함께 병행되어야 할 것으로 판단된다.

References

Anastasopoulos, P.C., Mannering, F.L. and Tarko, A.P.(2008) "Tobit analysis of vehicle accident rates on interstate highways", *Accident Analysis & Prevention*, Volume 40, pp.628-633.

Anastasopoulos, P.C., Mannering, F.L., Shankar, V.N. and

Haddock, J.E.(2012a) "A study of factors affecting highway accident rates using the random-parameters tobit model", *Accident Analysis & Prevention*, Volume 45, pp.628-633.

Anastasopoulos, P.C., Shankar, V.N., Haddock, J.E. and Mannering, F.L.(2012b) "A multivariate tobit analysis of highway accident-injury-severity rates", *Accident Analysis & Prevention*, Volume 45, pp.110-119.

Han, Su San, Kim, Kyung Hwan and Park, Byung Ho(2011). "Accident Models of Circular Intersections by Type in Korea", *International Journal of Highway Engineering*, 13(3), pp.103-110.

Kim, Kyung Hwan and Park, Byung Ho (2010), "Developing the Traffic Accident Severity Models by Vehicle Type", *Journal of the Korean Society of Safety*, Vol.25 No.3, pp.131-136.

Kim, Eung Cheol, Choe, Eun Jin, Lee, Dong Min and Kim, Do Hun(2010), "Development and validation of Accident Modification Factors of Two-Lane Rural Roadways", *Journal of Korean Society of Transportation*, Vol.28 No.3, pp.131-143.

Kim, Sang Youp, Choi, Jai Sung, Lee, Soo Beom, Kim, Seong Min, Cho, Won Bum and Kim, Yong Seok(2011). "Freeway Crash Frequency Model Development Based on the Classification of Geometric Alignment Type", *Journal of the Korean Society of Highway Engineers*, 13(1), pp.97-105.

Oh, Ju Taek , Kim, Do Hoon and Lee, Dong Min(2012). "Development of the Expected Safety Performance Models for Rural Highway Segments", *International Journal of Highway Engineering*, 14(2), pp.131-143.

Park, Min Soo and Chang, Myung Soon(2012). "Analysis of Traffic Accident Characteristics for the Overlap Section of Horizontal and Vertical Alignment", *Journal of the Korean Society of Highway Engineers*, 14(1), pp.95-102.

(접수일 : 2012. 7. 12 / 심사일 : 2012. 7. 15 / 심사완료일 : 2012. 11. 28)