

로지스틱 회귀분석 기반 노인 보행자 교통사고 요인 분석

Analysis-based Pedestrian Traffic Incident Analysis Based on Logistic Regression

김시원* · 길정원** · 권재경*** · 황재성**** · 이철기*****

* 주저자 : 아주대학교 일반대학원 교통공학과 석사

** 공저자 : 아주대학교 일반대학원 교통공학과 석사과정

*** 공저자 : 아주대학교 일반대학원 교통공학과 박사과정

**** 공저자 : 아주대학교 교통연구센터 연구원

***** 교신저자 : 아주대학교 교통시스템공학과 교수

Siwon Kim* · Jeongwon Gil* · Jaekyung Kwon* · Jae seong Hwang** · Choul ki Lee*

* Dept. of Dept. of Transportation Eng., Ajou University

** Dept. of Dept. of Transportation Research Center., Ajou University

† Corresponding author : Choul ki Lee, cklee@ajou.ac.kr

Vol. 23 No.2(2024)

April, 2024

pp.15~31

pISSN 1738-0774

eISSN 2384-1729

<https://doi.org/10.12815/kits.2024.23.2.15>

2024.23.2.15

Received 6 March 2024

Revised 11 March 2024

Accepted 18 April 2024

© 2024. The Korea Institute of Intelligent Transport Systems. All rights reserved.

요약

초고령화 사회로 진입하고 있는 대한민국 노인 인구의 상황을 반영하여 노인 교통사고의 특성을 파악하고, 이항형 변수를 활용하여 노인 보행자 교통사고에서 발생한 중상 이상의 교통사고와 경상 이하의 교통사고를 구분하여 독립변수와 종속변수의 관계를 분석하였다. 도로교통공단 교통사고분석시스템(TAAS)에서 지난 10년간('13년~'22년)의 노인 보행자 교통사고 자료를 구득하여 데이터 수집, 가공, 변수 선정을 수행하였으며, 기초 통계 및 사고 요인별 분석을 실시하였다. 로지스틱 회귀분석 모델을 적용하여 총 15개의 영향 변수를 도출하였고 노인 보행자 중상 이상의 교통사고 발생 확률이 가장 큰 영향을 미치는 영향 변수들을 도출하였다. 이후, 로지스틱 모형의 정확도 분석을 위해 통계적 검정을 수행하였으며, 노인 보행자 교통사고에 영향을 미치는 변수를 도출 및 예측모형 구축에 따른 교통사고 발생확률 예측 방법을 제시하였다.

핵심어 : 노인 보행자 교통사고, 로지스틱 회귀분석, 교통안전

ABSTRACT

The characteristics of elderly traffic accidents were identified by reflecting the situation of the elderly population in Korea, which is entering an ultra-aging society, and the relationship between independent and dependent variables was analyzed by classifying traffic accidents of serious or higher and traffic accidents of minor or lower in elderly pedestrian traffic accidents using binomial variables. Data collection, processing, and variable selection were performed by acquiring data from the elderly pedestrian traffic accident analysis system (TAAS) for the past 10 years (from 13 to 22 years), and basic statistics and analysis by accident factors were performed. A total of 15 influencing variables were derived by applying the logistic regression model, and the influencing variables that have the greatest influence on the probability of a traffic accident involving severe or higher elderly pedestrians were derived. After that, statistical tests were performed to analyze the suitability of the logistic model, and a method for predicting the probability of a traffic accident according to the construction of a prediction model was presented.

Key words : Elderly Pedestrian Traffic Accidents, Logistic Regression Analysis, Traffic safety

I. 서론

1. 배경 및 목적

지난 2013년부터 2022년까지 10년간 대한민국의 인구 성장률은 -0.233% 로 줄어든 반면, 65세 이상인 노인 인구, 즉 노인층의 인구는 5.22% 의 증가율을 보였다. UN(United Nations) 국제연합은 65세 이상을 ‘노인’으로 규정하고, 총인구 대비 노인의 인구 구성비가 14% 이상이면 ‘고령사회’, 20% 이상이면 ‘초고령사회’로 분류하고 있다. 현재 우리나라는 2025년 노인 인구 비중이 20% 이상으로 ‘초고령사회’ 진입을 전망하고 있다. 또한, 1955년에 태어난 ‘베이비 붐’ 세대가 2020년부터 노인 인구로 편입되면서, 노인 인구의 증가율이 가속되고 있는 통계 자료도 도출되었다.(KOSIS, 2021)

최근 10년(2013년 ~ 2022년) 동안 노인 보행자 교통사고는 62,305건이 발생하였으며 이 중 3주 이상의 치료를 필요로 하는 부상을 입은 중상자 수가 39,707명, 교통사고 발생으로 인하여 30일 이내에 사망한 사고가 5,897명으로 분석되었다. 중상사고 이상 발생한 교통사고는 43,196건으로 전체의 약 69.33% 를 차지하며 중상사고 이상의 교통사고로 인한 피해 노인 보행자는 45,604명으로 전체 피해자의 70.24% 를 나타내고 있다.

도로교통사고 중 인적피해에 따른 1인당 평균 사고비용을 추산한 결과, 사망자 비용은 중상자의 8배, 경상자의 107배, 부상신고자의 212배의 차이를 나타냈으며, 사망 5억 6,360만원, 중상 7,404만원, 경상 529만원, 부상신고 266만원으로 추산되었다.(KoRoad, 2023) 또한, 삼성교통문화안전연구소는 노인 취업자 수가 3년 뒤에는 323만명으로 2년 전보다 23% 증가할 것으로 판단하였고, 5년간 증가되는 노인 취업자의 추세를 보았을 때, 노인의 교통사고 건수는 증가할 것으로 판단하였다. 특히, 특별·광역시보다 광역자치도에서의 노인 보행자 사고가 비교적 많을 것으로 전망하였다.

이처럼 통계청 자료에 따라, 현재 대한민국의 고령화 속도는 빠르게 진행 중이며, 이에 따른 노인 인구의 사회적 활동도 같이 증가하기 때문에 노인 보행자의 교통사고 발생 위험도 같이 증가할 가능성이 높다. 현재 초고령화 사회 진입에 직면한 시점에서 노인 보행자의 교통사고 주요인 파악과 특성분석, 더 나아가 분석된 자료를 기반으로 예방 방안도 요구되고 있다.

본 연구에서 사용된 노인 보행자 교통사고 자료는 도로교통공단의 교통사고분석시스템(TAAS)를 통하여 수집되었다. 노인의 경우 신체적 기능 저하 등의 이유로 중상 이상의 사고 피해시, 사망으로 이를 확률이 매우 높기 때문에, 노인 보행자 교통 사고를 중상 이상(사망, 중상), 경상이하(경상, 부상)로 구분하여 분석을 진행하였으며, 이러한 사고 원인과 결과로 설명할 수 있는 이항형 변수로 표현하였다. 이는 독립변수와 종속변수 간의 관계를 나타내는 로지스틱 회귀분석을 통하여 노인 보행자 교통사고에 영향을 미치는 변수들을 도출하고, 이를 토대로 노인 보행자 교통사고 예측모형을 개발하여 분석을 진행하였다.

2. 연구범위 및 절차

본 연구는 로지스틱 회귀분석 방법으로 노인 보행자 교통사고에 영향을 주는 변수를 도출하여, 노인 보행자 교통사고 원인과 특성을 분석하고, 향후 노인사고를 감소시키는 예방 방법 수립 목적으로 수행되었다. 본 연구의 수행 절차는 <Fig. 1>에서 나타난 바와 같다.

선행연구 고찰하기에 앞서 노인에 대한 정의, 고령 보행자의 행동 특성 및 사고유형, 노인 사고 분석에 활용되는 로지스틱 회귀분석을 이론적 검토하였다. 또한 노인 보행자 교통사고와 관련된 선행연구 고찰과 로지스틱 회귀분석으로 도출된 예측값에 관한 연구도 같이 고찰하였다.

노인사고 데이터는 도로교통공단 교통사고 분석시스템(TAAS)에서 지난 10년간(‘13년 ~ ‘22년)의 노인 보행자 교통사고 자료를 사용하였다. 사용된 자료들은 분석을 위해 가공 및 변수 선정의 단계를 거쳤으며, 이를 통하여 노인 보행자 교통사고 현황 및 사고요인별 기초 분석의 수행이 가능하였다.

노인 보행자 사고가 발생할 확률을 추정하기 위해 유의미한 독립변수를 도출하는 과정을 거쳤고, 우도비검정, Hosmer & Lemeshow 검정, Cox and Snel R² & Nagelkerke R²검정을 통하여 분석 모형의 정확도를 분석하였다. 이를 기반으로 노인 보행자 중상 이상 교통사고 예측 모형을 제시하였고, 특정 독립변수의 변화 여부에 따른 교통사고 발생확률의 변화를 확인 및 실제 데이터를 통한 교통사고 발생 예측률 분석을 수행하였으며 결론과 향후 연구 방향을 논하였다.



<Fig. 1> Research promotion Process

II. 관련 이론 고찰

1. 노인의 정의 및 특징

국제노년학회와 노인복지론의 정의에 따르면 노화의 과정 또는 그 결과로서 생물, 심리, 사회적 기능이 약화되어 자립적 생활능력과 환경에 대한 적응능력이 약화되고 있는 사람이라고 정의 내리고 있다. 법·제도적 측면에서는 “노인복지법” 또는 “국민기초생활보장법”에 적용되는 65세 이상의 사람을 일컫는다. 또한, 장래추계 인구, 교통사고 통계분석과 같은 통계청에 해당하는 노인 기준 또한 대부분 65세 이상을 일컫고 있다. 영양학에서도 노년층의 경우 시각, 청각, 운동신경 등 물리적 반응과 기능이 불가역퇴행성 변화가 자·타각적으로 나타나며, 순간적 그리고 연속적 망각으로 올바른 판단을 내리지 못하는 경우가 많다. 때문에, 노인 보행자들은 안전에 대한 원칙을 배제하는 경향이 있으며, 차량의 움직임을 인지하지 못하고 자신의 행동만을 중시하는 등 소통 능력에 있어 비고령자와는 현저하게 낮은 수준이다. 실제 장애물 회피나 반응시간에 대한 시간을 측정하는 조사에서도 오류횟수가 고령자에게서 많이 나타났다.

신체적, 정신적 제약으로 외상에 노출될 가능성이 농후하고, 대응 능력이 부족으로 중증 외상, 중증 뇌손상 및 흉부외상, 척추 골절 등의 노인 외상 환자가 노인사고의 주를 이루고 있다. 또한, 노인은 생리학적인 기능의 저하 및 심장 질환자, 뇌질환 등 기저질환으로 노인 교통사고 발생 이후에 사망으로 이어지는 확률이 약 40% 발생하였다. 그러므로, 노인 교통사고의 경우 신체적인 기능의 저하 및 기저질환으로 인하여 수술 시행률이 높으며, 중증 증상으로 내원하였던 환자들이 사망하게 되는 경우가 높은 특성을 보였다.

2. 선행 연구 고찰

1) 노인 보행자 교통사고 연구 고찰

Kim et al.(2014)는 노인이 나이가 들어감에 따라 인지기능의 저하, 청력장애, 균형감각 및 신체조절 능력, 운동신경 저하 등 신체 기능 저하로 인해 위험요인에 대처하는 능력이 저하되어 외상에 쉽게 노출됨을 주장

하였다. 2008년 7월부터 2009년 3월까지 교통사고 후 원주세브란스기독병원 권역응급의료센터로 내원한 노인 외상환자 181명을 대상으로 분석한 결과, 노인군에서 사망률이 비 노인군 대비 사망률이 3배 이상 높으며 수술 시행여부 도 노인군에서 35% 더 많이 시행되었다.

Rosenbloom, et al.(2016)는 노인 보행자의 차량속도 판단 및 지각능력과 상황 대처능력이 비 노인군과 비교하였을 때 저하되며, 횡단 시 주변을 살피면서 진입하는 빈도가 비 노인군에 비해 약 30%정도 낮음을 분석하였다. 연구 결과에 따르면, 도로의 차선이 넓은 곳에서 횡단하는 노인 보행자, 특히 여성이거나 연령이 높으며 사회경제적 불안정한 경우에는 교통사고의 치명률이 높아짐을 도출하였다.

Gates, et al.(2006)는 시설·환경적인 측면에서 노인 보행자 교통사고에 영향을 미치는 영향요인을 연구하였다. 해당 연구에서 언급한 시설·환경적인 측면은 도로 종류 및 선형 등 도로구조와 교통 신호기, 중앙분리 시설, 안전시설물인 과속방지턱 등 뜻하고 있다. 노인 보행자는 단일로보다 교차로와 횡단보도 부근에서 교통사고가 더 빈번하게 발생함을 언급하였다.

Lee and Kim(2011)는 청·장년층과 노인을 비교그룹으로 설정한 다음, 노인의 교통사고 특성을 청·장년층과 비교하였다. 오전 6시부터 10시까지와, 초저녁 시간대인 18시~20시에 사망사고 건수가 3~5배 증가한 것으로 나타났다. 이는 해당 시간대에 노인복지센터를 방문하거나 집으로 귀가하는 시간대이며, 동시에 교통량이 많고 복잡한 교통 환경에 노출되기 때문이다.

Asher, et al.(2012)은 영국 노인을 기준으로 횡단보도 평균 보행속도는 남자 0.9m/s, 여자 0.8m/s로 65세 이상 남성의 84%와 여성 93%는 보행 속도 1.2m/s에 미치지 못하였으며, 노인에게 보행속도 1.2m/s는 부적절하며 횡단보도 보행 시간 연장에 대한 고려가 필요함을 주장하였다

Sin et al.(2022)은 일반 도로와는 특별한 경우로서 구분되는 교통약자 보호구역의 필요성을 중심으로 노인 교통사고 다발지역의 현황에 맞게 노인 보호구역이 지정되고 있는지와 어린이 보호구역이 적절히 일치되어 지정되고 있는지에 대해 파악과 비교 분석을 하였다. 분석결과 우리나라에서 어린이 인구 감소하고 노인의 인구가 증가하는 추세임에도 불구하고 어린이 보호구역이 노인 보호구역보다 약7.3배 많은 실정이다. 그리고 교통사고 사망자의 절반 이상이 노인이며, 노인 교통사고 발생은 심각한 사회의 문제라고 주장하였다.

Sung and Yoon(2023)은 대한민국은 2020년 15.7% 고령사회에 접어들고 있으며 이를 분석하기 위해

Random Forest와 Gradient Boosting regression을 활용하여 서울특별시를 분석하였다. 분석결과 횡단 중의 사고가 가장 큰 요인으로 작용하기에 횡단 중 노인보행자 교통사고를 낮추기 위해서는 과속차량의

단속 및 범규위반을 방지하는 대책 수립, 교통사고 예방과 불안감 저하를 위해 횡단시호 연장, 횡단거리 보행 시간 잔여기 신호등 설치 또한 노인보행 특성에 대한 운전자의 이해가 부족하여 노인 보행자 보호를 위한 운전자 교육이 필요할 것으로 언급하였다.

Nam(2020)은 고령자와 비고령자들을 대상으로 보행시물레이터를 활용하여 보행 패턴의 차이를 확인하고 시나리오 분석을 통해 분석한 결과, 고령자는 비고령자에 비해 횡단속도는 느렸고, 횡단실패 비율은 높았으며, 횡단 중 차량충돌 비율도 높은 결과를 나타낸다고 하였다. 이로 인해 횡단시 고령자가 비고령자에 비해 위험성이 증대될 수 있다고 시사하고 있다.

Aghdam et al.(2024)는 이란의 노인 보행자의 교통 행동과 그 결정 요인을 조사하는 것을 분석하였다. 다중 선형 회귀를 사용하여 고령 보행자의 90% 이상이 안전하지 않은 교통 행동을 보였으며, 거의 절반에 달하는 노인들은 건너나 길을 건널 때 공격적이고 주의가 산만했다고 주장했다. 이에 SES 수치가 낮은 보행자의 교통 행동을 개선하기 위한 개입이 필요하다고 주장했다.

2) 로지스틱 회귀분석 모델 기반의 교통사고 예측 연구 고찰

Yan et al.(2005)는 미국 플로리다의 교통사고 자료를 활용하여 차량의 유형, 도로 환경, 가해운전자의 특성 등에 따라 후미추돌사고에 대한 교통사고 발생확률을 추정하였다. 추정 방법으로 로지스틱 회귀모형을 활용하였으며, 교통량, 40mph 이상의 속도, 사고발생 시간대, 미끄러운 도로 표면, 가해운전자의 연령대, 음주 상태 등의 요인이 중요 변수라는 것을 도출하였다.

Park et al.(2009)는 충북지방경찰청에서 2004년 ~ 2005년까지 발생한 교통사고 자료를 바탕으로 청주시 4지 신호교차로의 사고위치별(유입/유출부, 교차로내, 횡단보도) 교통사고를 예측하기 위해 로지스틱 회귀분석 활용한 모형을 개발하였다. Hosmer & Lomeshow 검정과 정분류율 모두 정확한 모형으로 평가를 수행하였다. 교통량, 횡단거리 및 좌회전 전용차로가 교통사고 발생에 영향을 주는 공통 요인으로 분석되었다.

Kang et al.(2014)는 경상남도 창원시 도시부 신호교차로를 대상으로 유턴 사고모형 개발을 수행하였다. 연구에서는 3지, 4지 신호교차로의 유턴 교통사고 특성을 분석하고 로지스틱회귀분석 모델을 통해 사고예측모형을 구축하였다. 교통사고 건수에 대한 반응 척도를 이항으로 분류하였으며 3지 신호교차로의 경우, 상충교통량과 대향차로수가 (+)의 영향을, 4지교차로의 경우 교통사고 발생확률에 있어 상충교통량이 (+) 영향인 것으로 확인되었다.

Lee et al.(2015)는 부산광역시 해운대구 센텀남대로 및 해운대로에서 발생한 2013년도 교통사고 발생자료와 상세 기상 관측자료(시간당 강수량, 강수유무, 기온, 풍속), 시간대, 요일 자료 등을 로지스틱 회귀모형을 통하여 기상요인이 교통사고 발생에 미치는 영향 및 예측 모형을 산출하였다. 분석에 활용된 기상 요인 중, 기온과 강수유무 요인이 교통사고 발생에 유의미한 영향으로 작용하였다.

3. 연구의 차별성

기존의 선행연구는 노인 보행자 교통사고 현황, 발생 요인에 대한 단순 통계 비교와 이에 따른 해결책 위주로 제시가 되었으며, 로지스틱 회귀분석 기반의 사고 발생 위험성 분석 연구도 참조 집단에 대한 상대적인 효과 위주로 분석되었다. 본 연구는 로지스틱 회귀분석을 활용하여 노인 보행자 교통사고 단순 통계 비교가 아닌, 노인 보행자 교통사고 현황, 추이, 기초 통계분석을 통하여 사고시간, 도로환경(노면상태 및 도로구조 등), 운전자 특성과 같은 데이터로 노인 보행자 교통사고 확률을 추정하였다. 이를 기반으로 중상 이상 교통사고에 영향을 미치는 변수 도출 및 변수의 효과 분석을 통하여 예측모형을 제시하였다.

이항형 변수의 특성상 종속변수가 독립성, 정규분포, 등분산성을 따르지 않는 특징이 있어 선형회귀분석 방법은 부정확하다. 이후 개발된 모형식과 실제 데이터를 비교 및 분석하여 중상 이상 교통사고 발생확률을 예측하였다는 점에서 기존 선행연구와 차별성이 있다. 추가적으로, 사고 위치별, 운전자 상태별 등과 같은 변수들을 토대로 사고에 대한 위험도를 분석한 연구들은 영향 요인분석에서 그쳤다. 하지만 본 연구에서 제시하는 모형은 실제 발생하는 중상 이상 사고의 예방 및 안전 시스템 구축을 장려하는 연구로서 의의를 둘 수 있다고 할 수 있다. 모형에 의한 분석 결과로 노인 교통사고 발생에 가장 큰 영향을 미치는 변수를 도출해낸다. 또한, 도로 환경에 따른 위험 요소들과 법령 및 지리적 특성과 같은 도로 환경 외의 요소들까지 고려하여 노인 보행자 안전성 확보를 위한 정부 시책에 기여할 수 있는 점에서 차별성을 갖는다 판단된다.

4. 로지스틱 회귀분석

회귀분석은 두 가지 이상의 변수 간의 관계를 분석하여 종속변수의 값을 독립변수들로부터 설명하고 예

측하는 통계적 기법이다. 특히, 종속변수와 독립변수들의 관련성을 파악하는데 강력한 분석력을 가지며 적응성이 뛰어난 통계적 기법으로 알려져 있다. 회귀분석 식은 다음 <Equation 1>과 같다.

$$E(Y | X) = b_0 + b_1X_1 + b_2X_2 + \dots + b_{n-1}X_{n-1} + b_nX_n \dots\dots\dots (1)$$

절편은 b_0 이며 b_1 에서 b_n 까지의 계수는 각 독립변수의 회귀계수이다. 이 회귀계수는 각 독립변수가 종속변수에 미치는 영향의 크기를 의미한다. 독립변수 X 가 특정 값일 때 Y 의 예측값 또는 기댓값은 $E(Y | X)$ 이 된다. 그러나 종속변수가 중상 이상 교통사고, 경상 이하 교통사고와 같은 이항형 변수일 때 선형회귀분석을 적용하면 종속변수가 독립성, 정규분포, 등분산성 등과 같은 연속형 변수로 취급되지 않는다. 이로 인해 종속변수와 독립변수와의 관계의 명확성이 떨어지게 된다. 이러한 문제를 해결하기 위하여 로지스틱 함수가 활용된다. 로지스틱 함수를 통하여 확률의 개념을 지닌 종속변수 값이 0~1사이가 되도록 값을 정할 수 있다. 즉, 독립변수의 값이 아무리 클지라도 예측 확률값은 0보다 작거나 1을 초과할 수 없다. 로지스틱 함수를 수식으로 표현하면 <Equation 2>과 같다.

$$E(Y | X) = \frac{\exp(b_0 + b_1X)}{1 + \exp(b_0 + b_1X)} \dots\dots\dots (2)$$

독립변수 X 를 통해 도출된 Y 값은 $E(Y | X)$ 로 표현되며 종속변수 Y 가 이항형 변수일 경우 $E(Y | X)$ 는 확률 개념을 가지게 된다. 즉, 독립변수 X 의 값에 따라 도출된 확률값이 결정된다.

로지스틱 회귀모델의 추정에는 최대우도법(maximum likelihood) 추정 방법이 사용된다. 최대우도법에서는 자료의 발생 가능성을 최대화하여 미지수의 추정치를 구하는 방법이다. 이를 위해 우도함수(L)를 설정하고 이 함수를 최대화하는 값을 통해 미지수 추정치를 결정한다. 우도함수는 미지수에 따른 자료의 가능성을 함수로 표현한 것이다. 우도함수 L은 <Equation 3>과 같이 표현된다.

$$L = \prod [p_i^{y_i} [1 - p_i]^{1 - y_i}] \dots\dots\dots (3)$$

<Equation 3>에서 노인 보행자 중상 이상 교통사고 발생 또는 경상 이하 교통사고 발생여부로 나타내는 이항형 값은 y_i 이며, p_i 는 로지스틱 모형을 통해 도출된 발생확률을 뜻한다. 로지스틱 회귀모형의 정확도를 검증하기 위하여 분석모형(M)과 포화모형(S)을 비교하는 식은 <Equation 4>이다.

$$-2\ln L_M - (-2\ln L_S) \dots\dots\dots (4)$$

$\ln L_M$ 은 분석모형의 로그 우도 함수, $\ln L_S$ 는 포화모형의 로그 우도 함수를 의미한다. 모형의 정확도가 올라갈수록 $-2\ln L$ 의 값은 0에 가까워지며 모형의 정확도가 떨어질수록 양수로 값이 커진다. <Equation 4>를 로그 연산을 통해 <Equation 5>와 같이 변형할 수 있다.

$$-2\ln L \left[\frac{L_M}{L_S} \right] \dots\dots\dots (5)$$

$\frac{L_M}{L_S}$ 를 우도비라 칭하며, 우도비를 이용하여 얻은 통계값으로 우도비 검증의 절차를 거친다. 이를 수행하면 해당 변수를 포함한 모형과 포함하지 않은 모형의 $-2\ln L$ 차이를 산출하여 자유도 1인 x 분포에서 유의미한지 알아볼 수 있다. 타 검증 방법보다 우도비 검증의 검증력이 높기에 해당 방법이 사용되었다.

본 연구에서 로지스틱 회귀분석은 종속변수가 이항형 변수(중상 이상의 교통사고 발생, 경상 이하 교통사고 발생)일 때, 독립변수가 범주형 변수로 표현될 때 정확한 분석기법이다. 이와 같은 분석기법으로 각 변수가 분석 결과를 정량적으로 판단할 수 있으며, 변수 간 영향 관계를 분석할 수 있다. 분석을 위해, 종속변수가 이항형 변수일 때, 로지스틱 회귀분석을 활용하여 선형회귀분석의 한계를 극복하고 종속변수와 독립변수 간의 관계를 설명한다. 종속변수의 값을 0과 1사이로 제한하여 확률적 개념을 도입하고, 해당 모형으로 노인 보행자 교통사고 발생에 따른 유의미한 변수를 도출하여, 발생 요인을 도출한다.

III. 영향변수 도출 및 사고모형 구축

1. 데이터 검토 및 전처리

데이터는 도로교통공단에서 교통사고 분석시스템(TAAS)에서 제공하는 10년간(2013년 ~ 2022년) 발생한 노인 보행자 교통사고 데이터를 사용하였다. 노인 보행자 교통사고 분석을 진행하기에 앞서, 전처리 과정을 다음과 같이 진행하였다. 전체 교통사고 사고유형 중, 차대사람 사고유형과 연령이 65세 이상인 노인 보행자 사고를 분석 범위로 설정하였으며, 10년간 수집된 전체 사고 데이터 수는 108,299건으로 기타 범규위반, 기타불명의 운전자 차종, 미가입 등의 데이터를 제외한 69,232건을 이용하여 분석하였다. 전처리 후 데이터 테이블은 아래<Table 1>과 같다.

<Table 1> Data table after preprocessing

Category	Year of Accident	At-Fault Driver Gender	At-Fault Driver Age	At-Fault Vehicle Type	Road Type	Road Surface Condition	Season	Weather Condition	Day of the Week	Time of Accident	Accident Type	Traffic Violation
1	2013	male	20-30s	Passenger Car	Intersection	Dry	Winter	Clear	Weekend	Nighttime	Crossing the Road	Failure to Obey Traffic Rules
2	2013	male	20-30s	Motorcycle	Intersection	Dry	Winter	Clear	Weekday	Daytime	Pedestrian Crossing	Failure to Yield to Pedestrians
3	2013	Female	40-50s	Passenger Car	Intersection	Dry	Winter	Cloudy	Weekday	Daytime	Pedestrian Crossing	Failure to Obey Traffic Rules
4	2013	male	40-50s	Passenger Car	Intersection	Dry	Winter	Clear	Weekday	Daytime	Riding Along Roadside	Failure to Obey Traffic Rules
...
69229	2022	male	40-50s	Passenger Car Machinery	Intersection	Dry	Winter	Clear	Weekday	Daytime	Pedestrian Crossing	Failure to Obey Traffic Rules

Category	Year of Accident	At-Fault Driver Gender	At-Fault Driver Age	At-Fault Vehicle Type	Road Type	Road Surface Condition	Season	Weather Condition	Day of the Week	Time of Accident	Accident Type	Traffic Violation
69230	2022	Female	60s and Older	Passenger Car	Intersection	Dry	Winter	Clear	Weekday	Daytime	Crossing the Road	Failure to Obey Traffic Rules
69231	2022	male	60s and Older	Passenger Car	Intersection	Dry	Winter	Clear	Weekday	Daytime	Pedestrian Crossing	Failure to Yield to Pedestrians
69232	2022	male	40-50s	Construction Machinery	Intersection	Dry	Winter	Clear	Weekday	Daytime	Riding Along Roadside	Failure to Obey Traffic Rules

2. 변수 선정

노인 보행자 교통사고 자료를 바탕으로 노인 보행자 중 중상이상 사고(사망사고, 중상사고)와 경상이하 사고(경상사고, 부상신고)로 종속변수를 선정하였다. 도로교통공단 교통사고 분석시스템(TAAS)에서 수집한 자료와 선행연구 고찰을 통하여 인적요인, 차량요인, 도로 및 환경요인, 사고요인으로 총 네 가지 항목으로 독립변수를 선정하였다. 노인 보행자 교통사고의 가해자(운전자) 관련 자료를 인적요인 변수로 활용하였고 성별을 남자/여자, 연령은 20세 미만, 20~30대, 40~50대, 60대 이상으로 선정하였다. 차량요인은 선행연구 고찰을 통해 개인형 이동수단(PM), 건설기계, 승용, 승합, 원동기, 이륜, 자전거, 특수, 화물을 포함한 9가지 유형으로 구분하여 선정하였다. 도로 및 환경요인은 도로의 형태, 사고 시점의 노면 상태, 계절, 기상 상태, 발생 요일, 발생 시간대를 변수로 선정하였다. 도로형태는 교차로와 단일로 두 가지로 구분하고, 노면 상태는 건조, 결빙, 습기, 적설로 4가지의 유형으로 변수를 선정하였다. 계절은 4계절 형태로 구분하여 봄(3월~5월), 여름(6월~8월), 가을(9월~11월), 겨울(12월~2월)로 선정하였고 기상상태는 눈, 맑음, 비, 안개, 흐림으로 5가지의 날씨 유형으로 구분하였다. 발생 요일은 월요일부터 금요일은 주중, 토요일과 일요일은 주말로 구분하여 변수 선정하였고, 발생 시간대는 06시~18시를 주간, 18시~06시를 야간으로 구분하여 변수를 선정하였다. 사고요인은 사고유형과 법규위반을 변수로 선정하였으며 교통사고 분석시스템(TAAS)의 분류체계를 고려하여 변수를 설정하였다. 사고요인 중 사고유형은 길가장자리구역 통행 중, 보도 통행 중, 차도 통행 중, 횡단 중으로 4가지 유형으로 선정하였다. 두 번째 사고요인으로 법규위반은 과속, 교차로 운행 방법위반, 보행자 보호의무 위반, 불법유턴, 신호위반, 안전거리 미확보, 안전운전 불이행, 중앙선 침범, 직진 우회전 진행 방해, 총 9가지 유형으로 선정하였다. 위와 같은 자료들을 정리한 모습은 아래<Table 2>와 같다.

<Table 2> Selection Variables for the Prediction Model of Elderly Pedestrian Traffic Accidents

Division		Variable		
Dependent Variable		Severity of Injury (Severe or More, Minor)		
Independent Variables	Personal Factors	At-Fault Driver Gender	Male, Female	
		At-Fault Age Group	Under 20, 20-30s, 40-50s, 60 and Over	
	Vehicle Factors	At-Fault Vehicle Type	Personal Mobility, Construction Machinery, Passenger Car, Van, Motorcycle, Two-Wheeler, Bicycle, Special, Freight	
		Road and Environmental Factors	Road Type	Intersection, Single Road
			Road Surface Condition	Dry, Icy, Wet, Snowy
Season	Spring, Summer, Fall, Winter			

Division		Variable	
		Weather Condition	Snow, Clear, Rain, Fog, Cloudy
		Day of Occurrence	Weekday, Weekend
		Time of Occurrence	Daytime, Nighttime
	Accident Factors	Accident Type	Walking on Roadside, Walking on Sidewalk, Walking on Roadway, Crossing
		Traffic Violation	Speeding, Intersection Violation, Pedestrian Right-of-Way Violation, Illegal U-Turn, Signal Violation, Inadequate Following Distance, Failure to Drive in Proper Lane, Crossing Center Line, Obstructing Forward or Right Turn

(Hong, 2005)는 독립변수가 범주형일 때, 연속형 변수로 바꿔 회귀분석을 진행하기 위하여 더미변수 생성이 필요하다 하였다. 더미변수는 동일한 측정 단위 간격마다 동일한 차이를 부여하는 척도인 등간척도의 선형성 조건에 충족하지 못한 범주형 변수를 해결하기 위해 사용되는 방법이다. 각 범주의 집단을 구분하기 위해 더미변수를 생성하고 자신이 속한 집단에 1, 이외의 집단은 0으로 변수를 정한다. N개의 집단을 분석하는 경우, N-1개의 더미변수가 분석에 요구되며, 마지막 범주는 기준집단으로 지정한다. 이를 통해, 마지막 범주인 기준집단은 다른 더미변수들과의 관계를 설명하며, 각 범주가 독립변수와 종속변수 간에 어떠한 영향을 미치는지에 대한 정량적인 분석이 가능하게 된다.

3. 영향 변수 도출

위와 같이 선정된 변수들로 분류된 데이터는 교통사고 분석시스템(TAAS)에서 제공하는 자료와 선행연구 고찰을 통해 종속변수와 독립변수를 선정하였다. SPSS 21.0을 활용하여 선정된 47개의 독립 변수들을 로지스틱 회귀분석 기법으로 신뢰수준 95% 수준의 ($\alpha = 0.05$) 유의미한 변수 15개를 도출하였다. 해당 결과는 다음 <Table 3>이다.

<Table 3> Derivation of Influence Variables for Elderly Pedestrian Traffic Accidents

Independent Variables			Coefficient	Standard Error	Significance Probability	Exp(B)
Personal Factors	At-Fault Driver's Gender	Male (Reference: Female)	.098	.022	.016	1.035
	At-Fault Driver's Age Group	60 and Over	.101	.031	.031	1.036
Vehicle Factors	At-Fault Vehicle Type	Passenger Car	.233	.026	.004	1.087
		Van	.644	.040	.011	1.286
		Bicycle	-1.575	.072	.101	0.663
		Freight	.864	.085	.008	1.425
Road and Environmental Factors	Road Type	Intersection	.301	.019	.000	1.116
	Road Surface Condition	Dry	.143	.198	.034	1.052
	Season	Winter	.110	.022	.021	1.039
	Weather Condition	Clear	.288	.089	.000	1.110

Independent Variables			Coefficient	Standard Error	Significance Probability	Exp(B)
	Day of Occurrence	Weekday (Reference: Weekend)	.211	.031	.012	1.079
	Time of Occurrence	Nighttime (Reference: Daytime)	.424	.020	.008	1.172
Accident Factors	Accident Type	Crossing	.255	.011	.008	1.097
	Traffic Violation	Failure to Yield to Pedestrian	.421	.057	.021	1.170
		Failure to Drive Safely	.278	.071	.002	1.106

4. 사고모형 구축

본 연구에서는 노인 보행자 교통사고에 원인이 되는 독립변수가 2가지 이상이며, 로지스틱 회귀분석 모델은 식(6)과 식(7)과 같이 나타낼 수 있다.

$$Prob(\text{심각사고발생}) = \frac{1}{1 + e^{-z}} \dots\dots\dots (6)$$

독립변수의 선형결과 값으로 표현되는 z 는 $b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_p$ 으로 표현되며 p 는 노인 보행자 교통사고 중 심각도에 영향을 미치는 독립변수의 개수이다.

$$e^{-z} = \frac{1 - Prob(\text{중상이상사고발생})}{Prob(\text{중상이상사고발생})} = \frac{Prob(\text{경상이하사고발생})}{Prob(\text{중상이상사고발생})} \dots\dots\dots (7)$$

$$z = \ln \frac{Prob(\text{중상이상사고발생})}{Prob(\text{경상이하사고발생})} \dots\dots\dots (8)$$

노인 보행자 교통사고 중 중상 이상 교통사고가 발생할 확률과 경상 이하 교통사고가 발생할 확률을 식(8)과 같이 표현하면서, z 를 오즈(odds)라고 일컫으며 로그를 취한 값을 식(9)와 같이 도출할 수 있다.

$$z = \ln \frac{Prob(\text{중상이상사고발생})}{Prob(\text{경상이하사고발생})} = \ln Odds = b_0 + b_1x_1 + \dots b_p b_p \dots\dots\dots (9)$$

그러므로, 식(9)는 노인 보행자 교통사고 예측 모형식으로 적용할 수 있고 분석된 회귀계수가 0보다 크게 나타나면 오즈비는 1보다 큰 값이 된다. 로지스틱 회귀분석 결과, 독립변수의 오즈비가 1보다 크다는 뜻은 독립변수를 한 단위 증가시키면 노인 보행자의 중상 이상의 교통사고 발생 확률이 증가한다는 의미이다. 반대로 회귀 계수가 0보다 작으면 해당 독립변수의 오즈비는 1보다 작아지게 되며, 독립변수를 한 단위 증가시키면 중상 이상의 교통사고 발생 가능성이 감소하게 된다.

IV. 사고모형 정확도 분석 및 검증

1. 사고모형 정확도 분석

다음으로 모형 정확도 분석을 진행하기 위해 두 가지 관점에서 이를 수행하였다. 첫째, 전체 로지스틱 회귀모델의 유의성 검정하기 위해 우세한 정도를 나타내는 비율인 우도비(Likelihood ratio) 통계적 방법을 활용한다. 둘째, 개별 로지스틱 회귀계수에 대한 검정을 수행하며 최우법(Maximum-likelihood method)이라는 모집단의 모수를 추정하는 방법을 이용하여 로지스틱 회귀계수를 추정한다. 최우법은 로지스틱 회귀계수에서 관측값이 나타날 확률을 우도(Likelihood)라고 정의하고 해당 값을 최대화하여 회귀계수를 조정한다. 추정된 모델의 정확도는 3가지 검정 방법 중 첫 번째 방법인, 로그 값에 -2배를 한 지표(-2LL, -2log-likelihood)를 활용하여 정확성 검정 과정을 거친다. <Table 4>는 우도비 검정을 통해 로지스틱 회귀분석 모형에 새로운 변수의 추가 또는 제거에 따른 모형의 정확도 변화를 나타내며, -2LL값의 차이로 표현된다. 다시 말해, 독립변수들의 포함 여부로 로지스틱 회귀모형의 가변성을 의미한다. 총 3번의 검정 과정을 거치게 되는데, Step, Block, Model 단계로, 이는 모형에 대한 데이터 준비, 데이터 분석, 데이터 분석 결과의 과정에서 검정 절차를 거치는 것이다. 전체 독립변수들을 분석한 결과, 카이제곱 값이 1891.541이며, 유의확률(p)은 0.000으로 분석되어 신뢰수준 95%에서 해당 모형은 유의한 것으로 분석되었다.

<Table 4> Likelihood ratio test

Division	Chi-Square	Degrees of Freedom	Significance Probability
Step	1891.541	35	.000
Block	1891.541	35	.000
Model	1891.541	35	.000

<Table 5>의 Hosmer & Lemeshow 검정은 로지스틱 회귀분석 모형의 정확도를 평가하는 방법 중 하나이다. 해당 검정은 종속변수의 실제 값과 모형에 의한 예측값의 유사 정도를 나타내는 카이제곱 값으로 수행된다. 또한 귀무가설의 기각 여부와 모형의 유의수준을 확인하여 모형의 정확도를 분석할 필요가 있다. 여기서 귀무가설은 “노인 보행자 중상 이상 교통사고 발생에 대한 관측빈도와 예측빈도는 일치한다”로 설정되며, 유의수준은 0.531로 귀무가설을 기각하지 않음을 의미한다. 본 연구에서는 도출한 로지스틱 회귀분석 모형은 정확한 모형으로 구축되었음을 의미한다.

<Table 5> Hosmer & Lemeshow test

Chi-Square	Degrees of Freedom	Significance Probability
7.212	8	.531
7.212	8	.531
7.212	8	.531

추가적으로 Cox & Snel R²과, Nagelkerke R² 검정 방법은 로그우도 값을 활용하여 도출된 계수로 선형회귀분석의 R²과 대응된다. 이 값은 회귀식이 종속변수를 예측할 수 있는 설명력을 의미한다. 결정계수 값은

각각 0.277, 0.221으로 분석되었고, 이는 로지스틱 회귀분석 모형에서의 결정계수 값은 보편적으로 선형회귀 분석 모형 대비 상대적으로 매우 낮은 값을 지니기(Ko and Park) 때문에 참고하는 수준으로 활용된다.

<Table 6> Cox and Snel R² & Nagelkerke R² test

-2Log Likelihood	Cox & Snel R ²	Nagelkerke R ²
1188.824	.277	.221

2. 사고모형 정확도 검증

선정된 독립변수 중 인적요인, 차량요인, 도로 및 환경요인, 사고요인을 포함한 총 47개의 독립 변수에서 15개의 독립변수가 신뢰수준 95% 수준에서 통계적으로 유의미하게 분석되었다. 이렇게 분석된 15개의 독립 변수들을 활용하여 노인 보행자 중상 이상의 교통사고 로지스틱 모형을 개발하였고 해당 식은 <Equation 6>과 같다. 선정된 독립변수 중 인적요인, 차량요인, 도로 및 환경요인, 사고요인을 포함한 총 47개의 독립 변수에서 15개의 독립변수가 신뢰수준 95% 수준에서 통계적으로 유의미하게 분석되었다. 유의미한 독립변수 중 자전거를 제외한 14개의 변수에서 회귀계수 값은 양(+)의 부호를 보였고, 이는 독립변수의 단위 값이 증가될 때, 노인 보행자 중상 이상의 교통사고 발생 확률이 증가하는 정도를 의미한다.

$$Z = \ln \frac{\text{prob}(\text{중상이상사고})}{\text{prob}(\text{경상이하사고})} = \ln Odds = b_0 + b_1x_1 + \dots + b_px_p \dots\dots\dots (10)$$

$$\begin{aligned} Z = & 0.1821 + [0.098 * \text{Male}(x_1)] + [0.101 * \text{age 60 and older}(x_2)] + [0.233 * \text{Passenger car}(x_3)] \\ & + [0.644 * \text{Van}(x_4)] + [-1.575 * \text{Bicycle}(x_5)] + [0.864 * \text{Cargo}(x_6)] + [0.301 * \text{intersection}(x_7)] \\ & + [0.143 * \text{Dry}(x_8)] + [0.110 * \text{Winter}(x_9)] + [0.288 * \text{Clear}(x_{10})] + [0.211 * \text{Weekday}(x_{11})] \\ & + [0.424 * \text{nighttime}(x_{12})] + [0.255 * \text{Crossing}(x_{13})] + [0.421 * \text{Faliure to yield to pedestrian}(x_{14})] \\ & .278 * \text{Faliure to drive safely}(x_{15}) \end{aligned}$$

노인 보행자 교통사고 가중치 모형의 정확성을 검증하기 위해 오즈비 분석을 시행하였고, <Table 6>에서 제시된 Exp(B)(오즈비) 값을 활용하였다. <Table 7>은 비교적 큰 영향을 미치는 독립변수 ‘화물 차량’의 여부에 따른 노인 보행자 교통사고 발생확률의 변화를 의미한다. 즉, ‘화물 차량’을 동반한 노인 보행자 중상 이상의 교통사고 독립변수의 결정 값 Z(a)와 타 변수들은 유지된 상황에서 ‘화물 차량’이 없는 가정을 하였을 경우 독립변수 결정 값 Z(b)를 뜻한다.

<Table 7>의 결과를 <Table 6>의 오즈비 결과와 비교하면, ‘화물 차량’의 여부에 따라 Z값이 2.691에서 1.827로 감소 됨을 확인하였다. 이 감소된 차이는 ‘화물 차량’의 회귀계수인 0.864과 동일하다. 또한, 오즈비 변화에 있어, ‘화물 차량’ 여부에 대한 노인 보행자 사고 심각도는 $Z = \ln \frac{13.652}{5.7540} = 1.425$ 배 차이가 발생하며 이 역시, <Table 6>의 Exp(B)(오즈비) 1.425과 일치하는 결과를 보인다.

<Table 7> Verification of the accuracy of the elderly pedestrian traffic accident prediction model

Z(a) Prediction Model for Elderly Pedestrian Traffic Accidents(1)	Z(b) Prediction Model for Elderly Pedestrian Traffic Accidents(2)
$Z = -0.1821 + 0.098 + 0.101 - 1.575 + 0.233 + 0.644 + 0.864 + 0.301 + 0.43 + 0.11 + 0.288 + 0.211 + 0.424 + 0.255 + 0.421 + 0.278 = 2.614$	$Z = -0.1821 + 0.098 + 0.101 - 1.575 + 0.233 + 0.644 + (0.864 * 0) + 0.301 + 0.43 + 0.11 + 0.288 + 0.211 + 0.424 + 0.255 + 0.421 + 0.278 = 1.750$
$Prob = \frac{1}{1 + e^{-2.614}} = 93.18\%$	$Prob = \frac{1}{1 + e^{-1.750}} = 85.19\%$
$Z = \ln \frac{0.9318}{0.0682} = \ln 13.652 = 2.614$	$Z = \ln \frac{0.8519}{0.1481} = \ln 5.7540 = 1.750$

<Table 8>은 노인 보행자 교통사고 예측모형의 경상 이하 교통사고와 중상 이상의 교통사고에 대한 적합성을 검증한 결과이다. 적합도 검증표를 통하여 경상 이하 교통사고와 중상 이상의 교통사고에 대한 관측값과 예측값 차이를 분석하고, 경상이하 사고 16,509건, 중상이상 사고 40,596건을 분류하였으며 이에 대한 정확도는 91.65%로 분석되었다.

<Table 8> Verification of Accuracy of Traffic Accidents by Elderly Pedestrians

Division		Predicted Values			Accuracy (%)
		Minor Accidents	Severe and More Serious Accidents	Total Sum	
Observations	Minor Accidents	16,509	2,600	19,109	86.40%
	Severe and More Serious Accidents	2,600	40,596	43,196	94.00%
Total		19,109	43,196	62,305	91.65%

그러나, 분석에 활용된 변수 외에도 추가적으로 인적(보험가입, 면허여부, 음주 여부 등), 차종(각 차종별 소형, 중형, 대형 등의 구분), 도로 및 기상요인(도로선형, 시계 장애물 등), 사고요인 등을 포함한 다양한 변수를 분석한다면 보다 구체적인 교통사고 발생의 원인을 도출하고, 현실적인 교통사고 발생 확률을 예측할 수 있을 것으로 판단된다.

V. 실제 교통사고 데이터 기반 예측률 분석

교통사고 발생 예측률 분석은 2023년 기준 도로교통공단 교통사고 분석시스템 자료(2022년 교통사고 통계 및 GIS 정보)로 진행하였다. 현실의 데이터 기반으로 교통사고 발생확률을 예측할 수 있는지 분석하기 위하여 이와 같은 실데이터를 활용한 것이다. 예측률 분석 수행의 대상지로는 서울특별시 성동구를 선정하여 2022년 노인 보행자 교통사고 자료를 수집 및 데이터를 가공하였다. 성동구 선정 이유는 지하철 4가지 호선이 통과하는 곳이며, 서울 동북부의 강북, 강남을 연결하는 교통 중심지역으로, 과거 구축된 공업시설과 현재 밀집된 IT지식산업센터의 증가로 본 연구에서 다루고 있는 다양한 독립변수들을 포함하는 지점으로 판단되었기 때문

이다. 또한 다양한 독립변수가 존재하는 동시에 노인 인구가 약 3만의 높은 수치도 보여 분석하기 좋은 지점으로 판단하였다. 해당 지점의 노인 보행자 교통사고 발생 건수는 총 31건이며, 이 중 사망사고 1건, 중상사고 9건, 경상사고 20건, 부상신고 1건으로 나타났다. 이중 중상 이상의 교통사고는 10건으로 나타났으며, <Table 7>의 예측모형을 활용하여 2022년 서울특별시 성동구의 노인 보행자 교통사고 중 중상 이상 교통사고 발생확률을 분석하였다. 인적요인, 차량요인, 도로 및 기상요인, 사고요인에 대한 변수의 결정 값(Z)을 도출하여 <Table 9>와 같이 각 변수의 결정 값이 분석되었다. 분석된 결정 값의 범위는 최소 1.2369, 최대 2.589로 확인되었으며, 여기서 말하는 결정 값은 독립변수의 선형결과 값으로 표현되는 $\text{Exp}(B)$ (오즈비) 값을 의미한다. <Table 9>에는 중상 이상 교통사고 10건 중 3건의 사고 내용만을 대표적으로 표기하였으며 실제 10건 모두 분석 완료하였다. 합산된 변수의 결정 값을 활용하여 분석한 결과, 중상 이상의 노인 보행자 교통사고가 발생할 확률은 최소 77.50%, 최대 92.48%로 도출되었다. <Table 9>에서의 괄호 속 숫자는 각 변수들의 오즈비이며, 각각의 변수들로 종합적으로 계산하여 산출된 결정 값이 1 이상이라면, 노인 보행자의 중상 이상의 교통사고 발생확률이 높아지는 결정 값이라는 것을 의미한다.

<Table 9> Analysis of the elderly pedestrian traffic accidents

Division		1	2	3
The Factors Contributing to Elderly Pedestrian Traffic Accidents	Gender	Male(0.098)	Female (0.000)	Male(0.098)
	Age Group	40-50s (0.000)	40-50s (0.000)	40-50s (0.000)
	At-Fault Vehicle Type	Van (0.644)	Motorcycle (0.000)	Cargo (0.864)
	Road Configuration	Intersection (0.301)	Single Road (0.000)	Intersection (0.301)
	Road Surface Condition	Dry (0.143)	Dry (0.143)	Dry (0.143)
	Season	Fall (0.000)	Summer (0.000)	Winter (0.110)
	Weather Condition	Clear (0.288)	Clear (0.288)	Clear (0.288)
	Day of Occurrence	Weekend (0.000)	Weekday (0.211)	Weekday (0.211)
	Time of Occurrence	Nighttime (0.424)	Nighttime (0.424)	Daytime (0.000)
	Type of Accident	Crosswalk (0.255)	Crosswalk (0.255)	Crosswalk (0.255)
	Traffic Violation	Crossing the Center Line (0.000)	Crossing the Center Line (0.000)	Failure to Yield to a Pedestrian(0.421)
Accident Details		Severe Accident	Severe Accident	Fatal Accident
Constant		-.1821	-.1821	-.1821
Converted Sum of Z		2.0689	1.2369	2.5089
Occurrence Probability (%)		88.78	77.50	92.48

VI. 결론 및 향후연구

본 연구에서는 노인 보행자 교통사고 예측모형을 개발하고, 예측모형을 개발을 위해 노인 보행자의 행동 및 손상특성, 관련 이론과 선행연구를 고찰하였고, 예측 모형식 구축과정에서의 타당한 기반을 마련하였다.

데이터의 수집, 전처리 과정을 통하여 노인 교통사고 발생 요인 카테고리화 관련 변수를 도출하고, 활용된 변수를 정의 및 선정함으로써 노인 보행자 교통사고 현황, 요인별 분석을 수행하였다. 해당 모형은 로지스틱 회귀분석 모델을 통하여 노인 교통사고에 영향을 주는 인적요인, 차량요인, 도로 및 환경요인, 사고요인의 카테고리에서 총 15개의 영향 변수를 도출하였다. 또한 모형의 3가지의 우도비 검정, Hosmer & Lemeshow 검정, Cox and Snel R² & Nagelkerke R²검정의 검증방법을 수행하여 신뢰성과 정확성을 높이고 이러한 과정을 통해 개발된 모형으로 실제 교통사고 데이터를 기반으로 사고 예측을 시행하였다.

예측모형을 통해 노인 보행자 교통사고에 영향을 미치는 대표적인 변수는, ‘신호위반’, ‘보행자 보호의무위반’ ‘불법 유턴’등과 같은 인적 요인, 습기와 흐린날씨로 인한 가시성 저하의 기상과 도로 환경적 요인, 그리고 ‘화물 차량’과 같은 차량 요인 등이었습니다. 이 중 ‘화물 차량’이라는 독립변수가 노인 보행자 중상 이상의 교통사고 발생 확률에 가장 큰 영향을 미치는 것으로 분석되었다.

하지만 이는 절대적으로 ‘화물 차량’만이 사고 심각도를 상승시키는 요인으로 판단하기는 어렵다. <Table 7>의 결과를 <Table 6>의 오즈비 결과와 비교하였을 때, 화물차는 1.425배 위험한 요소로 도출되어 가장 영향을 미치는 화물 차량으로 검증되었지만, 수치상으로 가장 높을 뿐, ‘과속’이나 ‘보행자 보호 의무위반’과 같은 변수들 역시 1.393배로 ‘화물 차량’으로 인한 사고 심각도 수준과 비슷했기 때문이다. 분석을 진행한 성동구에서 역시 ‘화물 차량’의 사고 심각도가 높게 나왔지만, 해당 분석 결과의 ‘화물 차량’ 역시 ‘보행자 보호 의무위반’으로 높은 사고 확률을 보였기 때문에 향후 다양한 도로 및 도시 환경 형태에 맞는 분석을 진행할 필요성이 있다.

현재 행정안전부는 노인보행자 교통사고 다발 지역의 정비를 추진하고 있다. 증가하는 노인 보행자 사망자들의 위험요인을 확인하여, 고령화가 빠른 속도로 진행되고 있는 사회에서 고위험지점에 대한 지속적인 진단이 필요하다고 주장하였다. 현재 여러 도로 지점에 보행자를 위한 환경이 미흡하여 안전시설물 설치를 통하여 개선해 나갈 것임을 밝혔지만, 노인 보행자 교통사고에 영향을 미치는 요인에 따른 설치가 더 효율적일 것이라 판단된다.

따라서 본 연구를 통해, 노인 보행자 교통사고에 영향을 미치는 변수들을 확인하였으며, 교통안전 측면에서 노인 보행자 교통사고 예방 및 대응 전략에 유용할 것으로 기대된다. 초고령사회로 접어드는 시점에서 노인 보행자의 교통안전 문제는 심화되고 있으며, 노인 보행자 교통사고에 영향을 미치는 요인에 대한 대안을 통해 사고 심각도가 낮은 도로 환경 구축과 향상된 과학적인 사고 개선 사업을 추진할 수 있을 것으로 사료된다.

향후 활용한 독립변수 이외에, 설치된 안전시설물의 효과(중앙분리대, 신호체계 등), 보행환경의 변수(보행 도로 여부, 보행 환경 혼잡도 등), 교통운영 특성(교통량, 제한속도, 평균 속도 등), 교통 인프라 변수, 기상조건 변수, 사회 경제적 변수 등 추가적인 인과관계가 높은 독립변수를 찾아 예측모형 식에 적용하면, 보다 종합적이고 현실적인 고도화된 예측모형을 구축할 수 있을 것으로 판단된다. 이를 발전시켜, 도시부와 지방부로 공간적 범위 기준을 수립하여 각 지역적 특성에 따른 예측모형을 재정립하면 지역별로 효과적인 교통안전 정책과 개선방안을 도출할 수 있을 것으로 기대된다.

또한 본 연구의 경우 오직 노인 보행자 교통사고 요인분석을 진행하였다. 향후 연구에서는 비고령 보행자를 포함한 대상으로 로지스틱 회귀모형을 구축하여 결과를 비교하고, 차이점을 분석함으로써 노인 보행자만

의 특징이 아닌 광범위한 연령 그룹의 교통사고 요인을 도출할 수 있을 것으로 판단된다.

ACKNOWLEDGEMENTS

이 논문은 2023년도 정부(경찰청)의 재원으로 과학치안진흥센터의 지원을 받아 수행된 연구임(No.092021C 28S01000, 자율주행 혼재 시 도로교통 통합관제시스템 및 운영기술 개발)

REFERENCES

- Aghdam, F. B., Ghasemi, M., Zarghami, F., Jadidi, S. H., Sadeghi-Bazargani, H., Ponnet, K., Rostami, V. Z. and Kashgsaray, N. H.(2024), “The role of socioeconomic and cognitive status in determining traffic behaviour of elderly pedestrians in iran: A cross-sectional study”, *Journal of Road Safety*, vol. 35, no. 1, pp.40-51.
- Asher, L., Aresu, M., Falaschetti, E. and Mindell, J.(2012), “Most older pedestrians are unable to cross the road in time: a cross-sectional study”, *Age and Ageing*, vol. 41, no. 5, pp.690-694.
- Gates, T. J., Noyce, D. A., Bill, A. R. and Van Ee, N.(2006), “Recommended walking speeds for timing of pedestrian clearance intervals based on characteristics of the pedestrian population”, *Transportation Research Record*, vol. 1982, no. 1, pp.38-47.
- Hong, S.(2005), *Binomial and Multinomial Logistic Regression*, Kyoyookbook.
- Jin, S. B. and Lee, J. W.(2017), “Study on accident prediction models in urban railway casualty accidents using logistic regression analysis model”, *Journal of the Korean Society for Railway*, vol. 20, no. 4, pp.482-490.
- Kang, J. H., Kim, K. W. and Kim, S. M.(2014), “Development of the U-turn accident model at signalized intersections in urban areas by logistic regression analysis”, *KSCE(Korea Society of Civil Engineers) Journal of Civil and Environmental Engineering Research*, vol. 34, no. 4, pp.1279-1287.
- Kang, J. H., Kim, K. W. and Kim, S. M.(2014), “Development of the U-turn accident model at signalized intersections in urban areas by logistic regression analysis”, *KSCE(Korea Society of Civil Engineers)Journal of Civil and Environmental Engineering Research*, vol. 34, no. 4, pp.1279-1287.
- Kim, T. S., Lee, K. H., Kim, T. H., Kim, O. H., Cha, Y. S., Cha, K. C. and Hwang, S. O.(2014), “Clinical characteristics and prognostic factors of geriatric patients involved in traffic accidents”, *Journal of Trauma and Injury*, vol. 27, no. 4, pp.101-107.
- Ko, D. W. and Park, S. H.(2019), “The effects of individual accidents and neighborhood environmental characteristics on the severity of pedestrian traffic accidents in Seoul”, *Journal of the Architectural Institute of Korea Planning & Design*, vol. 35, no. 8, pp.101-109.
- Korea Road Traffic Authority, https://www.koroad.or.kr/main/board/6/87791/board_view.do?&cp=1&listType=list&bdOpenYn=Y&bdNoticeYn=N, 2023.03.15.

- KOSIS, https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1BPA002&conn_path=I2, 2021.12.09.
- Lee, H. J. and Kim, E. M.(2011), “Traffic Safety Countermeasures for the Elderly Considering Traffic Accidents Causes”, *Transportation Technology and Policy*, vol. 8, no. 5, pp.63-75.
- Lee, H. J. and Kim, E. M.(2011), “Traffic safety countermeasures for the elderly considering traffic accidents causes”, *Transportation Technology and Policy*, vol. 8, no. 5, pp.63-75.
- Lee, K., Jung, I., Noh, Y., Yoon, S. and Cho, Y.(2015), “The effect of road weather factors on traffic accident-Focused on Busan area”, *Journal of the Korean Data and Information Science Society*, vol. 26, no. 3, pp.661-668.
- Lee, T. H., Park, C. H., Park, H. H. and Kwak, D. H.(2019), “A study on accident prediction models for chemical accidents using the logistic regression analysis model”, *Fire Science and Engineering*, vol. 33, no. 6, pp.72-79.
- Nam, S. Y.(2020), “A study on the characterization of elderly pedestrians using a gait simulator”, *Journal of Traffic Safety*, vol. 39, pp.123-136.
- Oxley, J. A., Fildes, B. N., Ihsen, E., Charlton, J. L. and Day, R. H.(1999), “Age differences in road crossing decisions based on gap judgements”, *In Annual Proceedings/Association for the Advancement of Automotive Medicine, Association for the Advancement of Automotive Medicine*, vol. 43, p.279.
- Park, B. H., Yang, J. M. and Kim, J. Y.(2009), “Logistic regression accident models by location in the case of Cheong-ju 4-Legged signalized intersections”, *International Journal of Highway Engineering*, vol. 11, no. 2, pp.17-25.
- Rosenbloom, T., Sapir-Lavid, Y. and Perlman, A.(2016), “Risk factors in road crossing among elderly pedestrians and readiness to adopt safe behavior in socio-economic comparison”, *Accident Analysis & Prevention*, vol. 93, pp.23-31.
- Shin, J. H., Lee, D. Y., Kim, E. H., Hong, M. K. and Jang, H. S.(2022), “Verification of suitability for designation of traffic accident frequent areas for the elderly and children and protection zones for the elderly and children using GIS: focusing on the Seoul Metropolitan Government”, *Journal of the Korean Society of Sepled*, vol. 13, no. 1, pp.7-46.
- Sung, J. and Yoon, B.(2023), “Factor analysis of elderly pedestrian traffic accidents: Focusing on Seoul”, *Proceedings of the Korean Society for Disaster Information*, pp.261-262.
- Yan, X., Radwan, E. and Abdel-Aty, M.(2005), “Characteristics of rear-end accidents at signalized intersections using multiple logistic regression model”, *Accident Analysis & Prevention*, vol. 37, no. 6, pp.983-995.