

# K-평균 군집모형 및 순서형 로짓모형을 이용한 버스 사고 심각도 유형 분석 측면부 사고를 중심으로

이인식\* · 이현미\*\* · 장정아\*\*\* · 이용주\*\*\*\*

## Analysis of Bus Accident Severity Using K-Means Clustering Model and Ordered Logit Model

Insik Lee\*, Hyunmi Lee\*\*, Jeong Ah Jang\*\*\*, Yongju Yi\*\*\*\*

**Key Words:** *Bus accident repair cost*(버스 사고 수리 비용), *Accident severity*(사고 심각도), *K-means clustering*(K-평균 군집분석), *Ordered logit model*(순서형 로짓 모형)

### ABSTRACT

Although accident data from the National Police Agency and insurance companies do not know the vehicle safety, the damage level information can be obtained from the data managed by the bus credit association or the bus company itself. So the accident severity was analyzed based on the side impact accidents using accident repair cost. K-means clustering analysis separated the cost of accident repair into 'minor', 'moderate', 'severe', and 'very severe'. In addition, the side impact accident severity was analyzed by using an ordered logit model. As a result, it is appeared that the longer the repair period, the greater the impact on the severity of the side impact accident. Also, it is appeared that the higher the number of collision points, the greater the impact on the severity of the side impact accident. In addition, oblique collisions of the angle of impact were derived to affect the severity of the accident less than right angle collisions. Finally, the absence of opponent vehicle and large commercial vehicles involved accidents were shown to have less impact on the side impact accident severity than passenger cars.

### 1. 서론

우리나라는 교통안전법 제 52조, 제 59조에 의거 경찰청, 보험사, 공제 조합에서 교통사고 자료를 수집, 통합 분석하고 있다. 교통사고자료는 일반적으로 인적피해사

고와 물적피해사고로 구분되고 있으나, 이러한 기존 사고 자료에서는 버스 파손 정도 및 위치 등의 정보를 확인할 수 없다. 버스 파손과 관련된 실사고 데이터는 차량안전성을 판단하는데 매우 중요한 데이터임에도 불구하고 관련 자료가 부족한 실정이다.

본 연구에서는 버스정비회사 자체적으로 보유 및 관리하고 있는 버스 파손수리 자료에 대하여 수집 및 분석하여 차량안전성 측면에서 활용하기 위한 기초 연구를 수행하였다. 즉, 기존의 국내 경찰청 및 보험회사의 교통사고 데이터로는 차량 안전성을 알 수 없으며, 버스 공제조합이나 버스 회사 자체적으로 관리하는 데이터에서는 파손 정도 정보가 구득 가능하므로 이 중 사고 수리 비용을 이

\* 아주대학교 교통시스템공학과, 석사과정

\*\* 아주대학교 교통시스템공학과, 박사과정

\*\*\* 아주대학교 TOD 기반 지속가능 도시교통, 연구센터 연구교수

\*\*\*\* 아주대학교 TOD 기반 지속가능 도시교통, 연구센터 연구부교수

E-mail: azang@ajou.ac.kr

용하여 측면부 충돌 사고를 중심으로 사고 심각도를 분석하였다. 경찰청 교통사고 분석시스템(Traffic Accident Analysis System)의 2018년부터 2019년까지 사고 데이터에서는 전면부 충돌, 측면부 충돌 및 후면부 충돌 등 차대차 사고 유형 중 측면부 충돌 사고는 가장 많은 사망자 수 및 부상자 수를 발생시켰다. 이는 측면부 충돌 사고가 전면부 및 후면부 충돌 사고보다 상대적으로 취약하다고 볼 수 있으므로 측면부 충돌 사고를 중심으로 연구를 진행하였다.

버스 실사고 데이터는 전국을 대상으로 2018년부터 2019년까지 2개년 동안 수집된 A 운수사의 사고 수리 내역 정보를 기초로 한다. 수집 분석된 데이터에는 사고 심각도 관련 지표로 버스수리비용, 버스수리기간, 수리지점 수 등이 있는데 이 중 사고수리비용을 기반으로 사고 심각도 유형분류 및 분석을 실시하였다. 사고수리비용이 수리기간 및 수리지점 수보다 연속적인 변수인 것을 고려하였으며, 수리기간이 길다는 것은 그만큼 사고가 심각하다는 것을 의미할 수 있으나 운영사의 부품 수급 기간에 따라 수리기간이 길어질 수 있다는 문제점이 존재한다. 그러나 수리비용의 경우에는 부품 수급과 관계없이 버스 손상 정도에 따라 수리비용이 소모되기 때문에 버스 사고 심각도에 수리기간보다 직접적으로 영향을 미칠 것으로 판단하였다. 따라서 본 연구에서는 사고 심각도 지표로 수리비용을 선정하였다. 이후 사고 심각도 지표인 사고 수리 비용에 영향을 미치는 요인이 무엇인지 통계적 분석을 실시하였다.

## 2. 관련 연구 고찰

### 2.1. 버스 실사고 데이터

우리나라 버스 사고 데이터는 경찰청 및 보험사 등에서 집계하는데, TAAS에서는 차량 관련 데이터를 구분하여 Table 1과 같이 버스 사고 데이터를 제공하고 있다. 사고 데이터는 크게 가해운전자 차종별, 가해운전자 차량 용도별로 나뉘고 이 중 버스 관련 데이터는 승합차, 사업용의 노선버스 및 전세버스 등이 있다. 이러한 데이터에는 버스 파손위치, 버스파손사진, 상대차량의 정보 등은 포함하고 있지 않다.

### 2.2. 관련 이론 분석 연구

Barua and Tay(2010)의 연구에서는 순서형 프로빗

Table 1 Data Provided by TAAS

Vehicle-related	
By Vehicle Type of the Perpetrator	Monthly Traffic Accident
	Traffic Accident by Day of the Week
	Traffic Accident by Time Zone
	Traffic Accident by the Degree of Alcohol of the Driver
	Traffic Accident by Road Class
	Traffic Accident by Road Type
	Traffic Accident by Vehicle Type of the Victim
By Vehicle Use of the Perpetrator	Accident
	Monthly Traffic Accident
	Traffic Accident by Day of the Week
	Traffic Accident by Time Zone
	Traffic Accident by Driver's Law Violation
	Traffic Accident by the Degree of Alcohol of the Driver
	Traffic Accident by Road Class
	Traffic Accident by Road Type
Traffic Accident by Intersection Type	

모형을 활용하여 1998년부터 2005년까지 방글라데시의 다카에서 발생한 버스 교통사고에 대한 사고데이터를 분석하였다. 이를 통해 해당 기간 내 다카에서 발생한 버스 교통사고 심각도 영향요인을 도출하였다. 영향요인으로는 주말에 발생한 사고인 경우, 비침두시간에 발생한 사고인 경우, 왕복 2차로 구간에서 발생한 사고인 경우, 단독사고인 경우, 보행자 사고인 경우 등으로 나타났다. 이와 달리, 중앙분리대 설치 구간, 교통경찰의 신호 제어 구간의 사고 심각도는 낮은 것으로 나타났다.<sup>(1)</sup>

한수산과 박병호(2011)는 도로교통공단에서 제공하는 2007년 청주시 주간선도로 13개 구간 교통사고 자료를 수집하였다. 이 자료를 바탕으로 사고특성 분석 및 순서형 로짓모형을 이용해 사고 심각도 분석을 실시하였다. 계절 중 여름철, 건조 상태 및 교통 체증 시 사고 심각도가 높은 것으로 분석되었고, 사고 유형 및 기상 조건은 통계적으로 유의미하지 않은 것으로 나타났다.<sup>(2)</sup>

Feng et al.(2016)은 미국 내 다양한 유형의 운전자에 대한 버스 사고 심각도 위험요인을 분석하였다. 분석을 위하여 2006년부터 2010년까지 수집된 BIFA(Buses Involved in Fatal Accidents) 데이터를 활용하였고, 분석 방법론으로는 순서형 로짓 모형을 적용하였다. 또한,

운전자 유형을 분류하기 위해 k-평균 군집분석 방법론을 접목하였다. 분석 결과, 버스 사고 심각도 위험요인은 운전자 면허 유형, 사고 발생 요일, 사고 관여 차량 수, 충돌유형, 제한속도, 결빙 노면 등으로 나타났다. 또한, 도로교통법 위반 이력이 존재하는 운전자 영향요인으로는 구간유형, 차로 수, 도로경사, 젖은 노면, 자전거 관여 사고 등으로 나타났다.<sup>(3)</sup>

Briand et al.(2016)은 대중교통 이용자들의 통행패턴을 분석하기 위해 2014년 4월 프랑스 렌의 대중교통 스마트카드 데이터에 대하여 가우시안 모형 개념을 기반으로 한 군집분석을 수행하였다. 군집분석을 통해 도출한 통행패턴을 확인하여 대중교통 이용자들의 통행패턴 특성을 분석하였다. 분석 결과, 대중교통 이용자들은 오전, 오후 첨두시간대, 오전 비첨두시간대인 10시, 오후 비첨두시간대인 3시 등에 대중교통 이용이 높은 것으로 나타났다.<sup>(4)</sup>

Mondal and Rehana(2019)의 연구에서는 도로 구간상에 설치한 통행패턴 정보 수집 시스템을 통해 수집한 구간별 통행패턴 자료를 기반으로 교통정체 패턴을 분석하였다. 교통정체 패턴 분석을 위해 활용한 요인은 구간별 교통량과 속도이다. 분석에는 각 구간 데이터에 대하여 k-평균 군집분석을 접목하였고, 군집분석을 통해 나타난 군집별 특성 확인을 통해 통행패턴을 도출하였다. 분석 결과, 4개 군집이 적정 군집 수로 나타났고, 각 군집은 고밀도의 교통량-낮은 속도, 중밀도의 교통량-낮은 속도, 중밀도의 교통량-적정 속도, 저밀도의 교통량-높은 속도의 특성을 보이는 것으로 나타났다.<sup>(5)</sup>

빈미영과 손슬기(2019)는 버스유형 및 반복사고 여부별 교통사고 심각도를 분석하고, 그에 대한 개선방안으로 버스 첨단 운전자지원시스템 지원 방향을 제시하였다. 사고 심각도 분석을 위해 2015년부터 2017년에 발생한 경기도 내 버스 운전자 원인 교통사고 7,146건을 수집하였고, 이 중 2,293건을 추출하여 분석에 활용하였다. 해당 데이터에 순서형 로짓 모형을 접목하여 분석을 진행하였다. 분석 결과, 차대사람 사고, 속도위반, 신호위반, 승객 안전조치위반 등의 요인이 영향요인으로 선정되었다. 분석 결과를 토대로 이 연구에서는 향후 버스 대 사람 사고를 감소시킬 수 있는 보행자감지시스템, 긴급제동장치 등의 첨단 운전자지원시스템이 버스에 장착되어야 한다고 제안하였다.<sup>(6)</sup>

이인목 등(2020)은 2018년 4월~5월 대전광역시의 교통카드데이터에 k-평균 군집분석을 접목하여 대중교통 이용자들의 시간적 통행특성을 분석하였다. 분석 결

과, 8개의 시간적 통행패턴이 형성되었음을 확인할 수 있었다. 이를 통하여 시간적 통행패턴을 기준으로 대중교통 이용자를 구분할 수 있었다. 도출한 시간적 통행패턴을 군집별로 확인한 결과, 5개의 군집은 일반적인 오전, 오후 첨두시간대의 통행특성을 보이는 것으로 나타났고, 2개 군집은 낮 시간대의 통행특성을, 1개 군집은 첨두시간대의 통행특성과 낮 시간대의 통행특성을 모두 보이는 것으로 나타났다.<sup>(7)</sup>

김형준 등(2020)의 연구에서는 사고데이터로 2018년부터 2019년까지 수집된 1,416개의 버스 수리 내역을 사고데이터로 선정하여 대표 충돌 유형 상황을 분석하였다. 분석에는 k-평균 군집분석을 활용하였다. 군집분석 결과를 기반으로 파손부위별 사고 심각도가 높은 사고 유형을 대표 버스충돌유형으로 선정하였다. 일례로, 전면부 사고 유형의 대표 버스충돌유형은 버스 운행계통이 고속/시외버스인 경우, 피해사고인 경우, 기상상태가 맑음 또는 흐림인 경우, 노면상태가 마르고 포장된 경우, 차대차 사고인 경우, 상대차량이 버스 또는 트럭 등의 대형 상용차량인 경우, 사고발생 도로유형이 자동차 전용도로인 경우, 사고발생 지점유형이 합류부 또는 분류부인 경우로 나타났다.<sup>(8)</sup>

이상의 선행연구 검토를 통해 범주형 자료에 해당하는 교통사고 데이터의 사고 심각도 분석에는 주로 순서형 로짓 모형, 순서형 프로빗 모형 등의 통계적 모형을 활용한다는 것을 알 수 있었다. 일부 k-평균 군집분석 연구 사례 중 버스 교통사고 심각도가 높은 사고유형을 도출한 사례를 확인할 수 있었으나, 대중교통 이용자의 통행 시간대 등을 분석한 통행패턴 분석 사례가 주를 이루는 것을 알 수 있었다. 본 연구와 같이 k-평균 군집분석 기법을 활용하여 사고 심각도를 구분하고, 이를 기반으로 순서형 로짓 모형을 구축한 연구사례는 부족한 것으로 확인되었다.

따라서, 본 연구는 종속변수가 연속형인 버스 수리비용에 대해 k-평균 군집분석을 수행하여 군집별 평균 비용으로 사고 심각도를 구분한 뒤, 구분한 사고 심각도를 기반으로 순서형 로짓 모형을 활용하여 분석을 수행하였다.

### 3. 버스 측면부 실사고 데이터 구축

본 연구는 제주도를 제외한 전국을 대상으로 2018년부터 2019년까지 수집된 A 운수사 사고 수리 내역 데이터를 활용한다. 데이터는 약 1,900건으로 사고 발생 범위

는 전국이며 경기도가 전체의 40% 정도를 차지한다. 이 데이터에는 ‘차량번호’, ‘입고일자’, ‘출고일자’, ‘차종’, ‘공임액’, ‘부품액’, ‘사고주소’, ‘사고내용’, ‘일기’, ‘도로상태’ 등의 정보 및 사진이 포함되어 있다. 이 중 버스 측면부 충돌 사고 총 444건만을 활용하여 분석을 진행하였다. 사고 심각도 종속변수로 ‘사고 수리 비용(공임액+부품액)(Repair Cost)’을 선정하여 k-평균 군집분석을 통해 데이터 군집을 구축하였다.

또한, 사고 심각도의 연속형 요인으로는 ‘수리 기간(일)(Repair Term)’, ‘충돌지점 수(개)(Number of Collision Points)’로 설정하였으며, 범주형 요인으로는 전면~40% Offset, 40%~60% Offset, 60%~80% Offset, 80%~100% Offset, 뒷바퀴~후면이 포함된 ‘충돌위치(Collision Point)’, 사선충돌, 직각충돌이 포함된 ‘충돌각도(Collision Angle)’, 부재(단독사고), 대형차(승합차, 버스, 화물차), 소형차(승용차, 이륜차)가 포함된 ‘상대차량(Opponent Vehicle)’, 자동차전용도로, 도시부도로, 지방부도로가 포함된 ‘도로유형(Road Type)’, 본선, 교차로, 합류부/분류부, 기타(TG/교량/지하차도/터널 등)가 포함된 ‘지점유형(Location Type)’으로 설정한다. 충돌위치(Collision Point)의 Offset은 버스 측면부의 전륜과 후륜을 각각 0%와 100%로 정의한 후 각 충돌위치를 %로 나타낸 것이며, 위에 나열한 순서대로 Fig. 1의 왼쪽부터 부여된 번호와 매칭된다. CNG, 수소 등의 연료장치를 탑재한 버스의 경우 전륜과 후륜의 중앙에서부터 후륜까지 연료장치가 존재하여 후륜 중 위치 영역의 충돌유형에 대한 세부 분석 필요성이 있기 때문에 Fig. 1의 1, 2, 3을 세부적으로 나누어 분석하였다.



Fig. 1 Collision Point

#### 4. 버스 수리 비용 기반 심각도 유형 분류

##### 4.1. K-평균 군집분석 적용

군집분석이란 분석 데이터의 특성에 따라 데이터 집단인 군집을 정의하고, 각 군집의 대표점을 도출하는 방법을 말한다. 다수의 전체 데이터 특성을 대표점 확인으로

도 파악할 수 있다는 장점이 있다.

K-평균 군집분석은 MacQueen(1967)이 최초로 제안한 분석 방법론으로, 분석하고자 하는 전체 데이터를 각 군집 간 거리 차이의 분산을 최소화하는 k개의 군집으로 분석하는 방법론이다.<sup>(9)</sup>

K-평균 군집분석의 첫 번째 단계는 초기 중심값을 임의적으로 선택하는 단계이다. 두 번째 단계는 각 그룹 중심값과 해당 그룹에 구성되어 있는 데이터 간 거리를 유클리디안 거리 산출 방법을 활용하여 산출하는 단계이다. 유클리디안 거리 산출은 식 (1)을 활용한다. 이후에는 산출한 값이 최소값이 되는 군집에 해당 데이터를 할당한다. 세 번째 단계는 앞서 형성한 군집마다 새로운 중심값을 산출하여 새로운 군집을 형성하는 단계이다. 위의 과정을 반복함으로써 결과적으로 최종 군집을 형성하게 된다.

$$\begin{aligned} & \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \\ & = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \end{aligned} \quad (1)$$

여기서,  $p_n$  = P점의 좌표  
 $q_n$  = Q점의 좌표

K-평균 군집분석 수행을 위해서는 군집분석 수행 전에 군집 수(k)를 산출해야 한다. 군집 수 산출 방법론으로는 rule of thumb, elbow method, 정보 기준 접근법 등이 있다. 이들 중 본 연구에서는 elbow method를 선정하였다. Elbow method는 점진적으로 군집 수를 증가시키며 각 군집 수를 적용했을 경우의 군집분석 결과를 모니터링하여, n+1개 군집의 군집분석 결과가 n개 군집의 군집분석 결과보다 나은 결과를 나타내지 않는 경우, 적정 군집 수를 n으로 채택하는 방법이다. 적정 군집 수는 그 그래프를 활용하여 시각적으로도 확인 가능하다.

#### 4.2. K-평균 군집분석을 이용한 사고 심각도 수준 구분

사고 심각도 수준을 구분하기 위해 사고 심각도 지표인 사고 수리 비용(Repair cost)에 대하여 k-평균 군집분석을 실시하였다. 군집을 나누기에 앞서, 적절한 군집 수를 산출하기 위하여 elbow method를 활용하였다. 적정 군집 수를 Fig. 2로 확인한 결과, 군집이 4개 이상부터 간격이 일정하고 5개 군집의 군집분석 결과가 4개 군

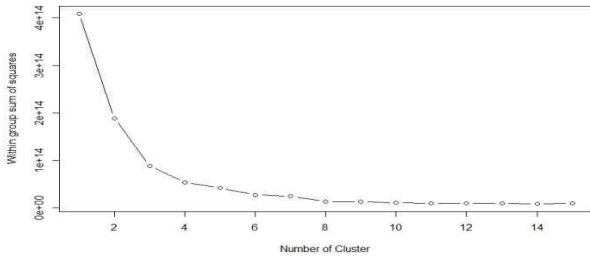


Fig. 2 Elbow Method Outcome

집의 군집분석 결과보다 나은 결과를 나타내지 못하여 적정 군집의 수를 4개로 채택한다. 또한, 기존 사고 심각도 연구들도 대부분 사고 심각도 지표를 사망, 중상, 경상, 부상의 총 4가지로 나누는데, 사고 수리 비용으로 사고 심각도 수준을 나눈 결과 기존 연구들과 같은 양상을 보임을 확인할 수 있었다. 따라서, 본 연구의 사고 심각도 수준을 군집별로 경미(Minor), 보통(Moderate), 심각(Severe), 매우 심각(Very Severe)으로 구분하였다.

4.3. 사고 심각도 정의

K-평균 군집분석 결과, 1번 군집은 평균 사고 수리 비용이 364,940원으로 사고 심각도 수준을 경미로 구분하고, 2번 군집은 평균 사고 수리 비용이 1,164,940원으로 사고 심각도 수준을 보통으로 구분하며, 3번 군집은 평균 사고 수리 비용이 2,452,868원으로 사고 심각도 수준을 심각으로 구분한다. 마지막으로, 4번 군집은 평균 사고 수리 비용이 6,090,330원으로 사고 심각도 수준을 매우 심각으로 구분한다. 결론적으로, 사고 수리 비용이 364,940원 이하는 경미, 364,940원 초과 1,164,251원 이하는 보통, 1,164,251원 초과 2,452,868원 이하의 심각, 2,452,868원 초과 6,090,330원 이하는 매우 심각으로 구분할 수 있다. 사고 심각도 수준이 높아짐에 따라 연속형 변수인 수리 기간과 충돌지점 수의 평균은 높아지는 경향을 보였으며, 그 외 범주형 변수들의 코딩정보 및 결과는 Table 2와 같다.

Table 2 Meaning of Variables

Category	Variable		Cluster			
			Cluster 1	Cluster 2	Cluster 3	Cluster 4
Average	Repair Cost		364,940	1,164,251	2,452,868	6,090,330
	Repair Term		0.30	0.92	1.79	6.00
	Number of Collision Points		1.23	1.54	2.06	3.86
Number of vehicle (N)	Collision Point	Front~40% Offset	97	61	14	4
		40% Offset~60% Offset	24	13	3	0
		60% Offset~80% Offset	29	12	5	2
		80% Offset~100% Offset	39	30	4	0
		100% Offset~Rear	72	26	8	1
	Collision Angle	Oblique Collision	243	119	25	6
		Right angle collision	18	23	9	1
	Opponent Vehicle	Unclassified	40	23	8	3
		Absence	77	38	4	3
		Large Commercial Vehicle	123	46	13	1
		Passenger Car	21	35	9	0
	Road Type	Motorway	25	20	8	3
		Urban Road	180	95	18	2
		Local Road	56	27	8	2
	Location Type	Others	141	50	12	2
		Intersection	79	58	16	3
Merging/Diverging Area		10	12	2	1	
Main Line		31	22	4	1	
Total (N=444)			261 (58.78%)	142 (31.98%)	34 (7.66%)	7 (1.58%)

5. 버스 측면부 사고 심각도 분석

5.1. 순서형 로짓 모형 적용

범주형 자료에 대한 영향요인 및 상관관계 분석을 수행하는 경우에는 주로 일반화 선형모형을 활용한다. 일반화 선형모형이란 기존 선형회귀모형에서 종속변수의 분포를 정규분포를 포함한 다양한 분포로 확장하고, 기존 종속변수의 평균과 독립변수의 선형 관계를 종속변수 평균 함수와 독립변수의 선형 관계로 확장한 모형이다. 본 연구에서는 일반화 선형모형 중 순서형 로짓 모형을 활용하였다. 순서형 로짓 모형은 McKelvey and Zavoina(1975)가 개발한 모형으로, 종속변수가 순서를 가지고 있는 범주형인 경우에 활용되는 분석 모형이다.  $y_i$ 를 교통사고 심각도로 정의하고,  $x_i$ 를 교통사고 심각도 영향요인으로 정의하면, 잠재변수(latent variable)  $y_i^*$ 는 식 (2)와 같고, 그에 따른 교통사고 심각도  $y_i$ 는 식 (3)과 같이 정의될 수 있다. 오차항  $\epsilon_i$ 의 확률분포가 표준 로지스틱 분포(standard logistic distribution)을 따른다고 가정하면 로짓 모형이 된다.<sup>(10)</sup>

$$y_i^* = \beta X_i + \epsilon_i \tag{2}$$

$$y_i = j \text{ if } \alpha_{j-1} < y_i^* \leq \alpha_j \tag{3}$$

또한,  $x_i$  값이 주어진 경우,  $i$ 번째 분류된 교통사고 심각도가  $j$ 에 속할 확률은 식 (4)와 같다. 여기서,  $F(z)$ 는 로지스틱 분포의 누적 확률 함수를 의미한다.

$$P(y_i = j) = \frac{P(\alpha_{j-1} < y_i^* \leq \alpha_j)}{F(\alpha_j - \chi_i' \beta) - F(\alpha_{j-1} - \chi_i' \beta)} \tag{4}$$

추정치는 최대우도추정치(maximum likelihood estimators)를 활용하여 산출한다. 추정치 산출에 활용되는 우도함수(likelihood function)는 식 (5)와 같다. 구축 모형의 적합도는  $\rho^2$  우도비(likelihood ratio)와 카이제곱( $\chi^2$ ) 값을 활용하여 확인한다.  $\rho^2$  우도비 산출식은 식 (6)과 같다.

$$L = \prod_{j=1}^J \prod_{i=1}^N [F(\alpha_j - \chi_i' \beta) - F(\alpha_{j-1} - \chi_i' \beta)]^{I(y_i = j)} \tag{5}$$

$$\rho^2 = 1 - (\ln L_0 / \ln L_1) \tag{6}$$

5.2. 순서형 로짓 모형 적용

버스 측면부 사고 심각도를 분석하기 위해 순서형 로짓 모형을 적용하기에 앞서, 독립변수간 다중공선성 문제가 발생하는지 여부를 확인하기로 한다. 다중공선성이란 독립변수간 강한 상관관계를 나타내는 것이다. 다중공선성 문제가 발생할 경우, 회귀분석의 전체 조건이 위배되는 것이므로 해결해야 한다. 다중공선성 진단 결과, 모든 변수에서 분석팽창계수(VIF, Variance Inflation Factor)가 2 미만으로 다중공선성 문제가 없다는 결과가 도출되었다. 독립변수의 VIF 결과는 Table 3과 같다.

Table 3 Multicollinearity Diagnostics Results

Independent Variable	VIF
Repair Term	1.085
Number of Collision Points	1.186
Collision Point	1.033
Collision Angle	1.092
Opponent Vehicle	1.032
Road Type	1.035
Location Type	1.043

다중공선성 문제가 발생하지 않는 것을 확인한 후, 순서형 로짓 모형의 적합도를 알아보기 위해 라인 평행성 검정(TPL test)을 실시하였다. 검정결과, 0.22로 0.05보다 크므로 모형이 적합함을 확인할 수 있었으며 -2 로그 우도값은 533.730으로 도출되었다. 그 결과는 Table 4와 같다.

6. 유의성 분석 결과

순서형 로짓 모형을 적용하여 버스 측면부 사고 심각도 분석 결과, 95%의 신뢰도를 기준으로 유의미한 변수는 수리기간(Repair Term), 충돌지점 수(Number of Collision Points), 충돌각도(Collision Angle), 상대차량(Opponent Vehicle) 등이 있다.

6.1. 수리 기간의 사고 심각도 분석

수리 기간은 통계적으로 유의한 결과가 나타났으며 수리 기간이 하루씩 증가할 때마다 사고 심각도가 높아질 위험이 1.843배 증가하였다. 이는 시간이 많이 소요될수

Table 4 Ordered Logit Model Results

Explanatory Variable	B	Standard Error	Wald	Degree of Freedom	P-Value	Exp(B)	95% Confidence Interval		
							Lower Limit	Upper Limit	
Repair Term	0.612	0.080	58.570	1.000	0.000***	1.843	0.455	0.768	
Number of Collision Points	1.058	0.160	44.026	1.000	0.000***	2.882	0.746	1.371	
Collision Point	Front~40% Offset	-0.074	0.285	0.068	1.000	0.794	0.928	-0.632	0.483
	40% Offset~60% Offset	0.217	0.415	0.274	1.000	0.601	1.242	-0.595	1.030
	60% Offset~80% Offset	0.238	0.383	0.385	1.000	0.535	1.269	-0.513	0.990
	80% Offset~100% Offset	0.366	0.337	1.183	1.000	0.277	1.442	-0.293	1.026
	100% Offset~Rear	0			0.000				
Collision Angle	Oblique Collision	-0.655	0.333	3.859	1.000	0.049**	0.519	-1.309	-0.001
	Right angle collision	0			0.000				
Opponent Vehicle	Unclassified	-0.503	0.352	2.048	1.000	0.152	0.604	-1.193	0.186
	Absence	-0.768	0.358	4.592	1.000	0.032**	0.464	-1.471	-0.066
	Large Commercial Vehicle	-0.958	0.325	8.702	1.000	0.003***	0.384	-1.594	-0.321
	Passenger Car	0			0.000				
Road Type	Motorway	0.508	0.427	1.414	1.000	0.234	1.662	-0.329	1.345
	Urban Road	0.048	0.267	0.032	1.000	0.858	1.049	-0.475	0.570
	Local Road	0			0.000				
Location Type	Others	-0.503	0.331	2.306	1.000	0.129	0.605	-1.152	0.146
	Intersection	0.111	0.346	0.103	1.000	0.749	1.117	-0.568	0.790
	Merging/Diverging Area	-0.624	0.572	1.190	1.000	0.275	0.536	-1.745	0.497
	Main Line	0			0.000				
-2 Log Likelihood				533.730					
TPL Test				0.220					

\*\* p<0.05, \*\*\* p<0.01

록 수리해야 할 부분이 많다는 뜻으로 사고 심각도인 사고 수리 비용이 높아지는 것으로 생각할 수 있다.

### 6.2. 충돌지점 수의 사고 심각도 분석

수리 기간은 통계적으로 유의한 결과가 나타났으며 충돌지점 수가 하나씩 증가할수록 사고 심각도가 높

아질 위험이 2.882배 증가하였다. 이는 충돌지점 수가 많을수록 파손된 부위가 많고 수리해야 할 부위가 많아지므로 사고 심각도인 사고 수리 비용이 높아지는 것으로 볼 수 있다.

### 6.3. 충돌위치의 사고 심각도 분석

충돌위치는 통계적으로 유의하지 않았으며, 본 연구는

측면부 사고 심각도 분석으로 측면부의 어느 위치든 파손의 정도에 따라 사고 수리 비용 가변 폭이 크기 때문에 유의미하지 않은 결과가 도출되었다고 판단된다.

### 6.4. 충돌각도의 사고 심각도 분석

충돌각도 중 사선충돌은 통계적으로 유의한 결과가 나타났으며 사고 심각도는 직각충돌의 0.519배 낮게 나타났다. 이는 버스 측면부에 사선으로 충돌했을 때보다 직각충돌 시 충격이 더 세게 가해지기 때문으로 판단할 수 있다.

### 6.5. 상대차량의 사고 심각도 분석

상대차량 교통사고 중 단독사고인 경우와 대형차량 관

여 사고는 통계적으로 유의한 결과가 나타났으며, 미분류 차량은 통계적으로 유의하지 않은 결과가 나타났다. 이는 소형차에 비하여 단독사고는 0.464배, 대형차는 0.384배 낮은 사고 심각도가 나타난 것으로 해석할 수 있다. 단독 사고가 소형차에 비해 사고 심각도가 낮은 이유는 단순히 공작물 충돌 시보다 차대차 사고에 의한 충격이 사고 수리 비용이 더 많이 나오는 것으로 볼 수 있기 때문이다. 또한, 대형차가 소형차에 비해 사고 심각도가 낮은 이유는 대형차는 속도제한 등 주행속도가 소형차에 비해 낮기 때문으로 볼 수 있다. 주행속도가 낮을수록 차량에 가해지는 충격이 적으므로 파손 정도도 적기 때문에 사고 수리 비용이 적게 드는 것으로 파악할 수 있다.

#### 6.6. 도로유형 및 지점유형의 사고 심각도 분석

도로유형과 지점유형은 통계적으로 유의하지 않은 결과가 나타났으며, 사고 수리 비용과의 관계성이 없는 것으로 해석된다.

### 7. 결 론

제주도를 제외한 전국을 대상으로 2018년부터 2019년까지 수집된 A 운수사의 사고 수리 내역 데이터 중 측면부 충돌 사고 총 444건을 활용하여 본 연구를 진행하였다. 연속형 요인으로는 수리 기간, 충돌지점 수로 설정하고, 범주형 요인으로는 전면~40% Offset, 40% Offset~60% Offset, 60% Offset~80% Offset, 80% Offset~100% Offset, 뒷바퀴~후면이 포함된 '충돌위치', 사선충돌, 직각충돌이 포함된 '충돌각도', 부재(단독사고), 대형차(승합차, 버스, 화물차), 소형차(승용차, 이륜차)가 포함된 '상대차량', 자동차전용도로, 도시부도로, 지방부도로가 포함된 '도로유형', 본선, 교차로, 합류부/분류부, 기타(TG/교량/지하차도/터널 등)가 포함된 '지점유형'으로 설정한 후, 사고 심각도 종속변수로 사고 수리 비용을 설정하여 k-평균 군집분석을 실시하였다. 그 결과, 사고 수리 비용이 364,940원 이하는 경미한 사고, 364,940원 초과 1,164,251원 이하는 보통인 사고, 1,164,251원 초과 2,452,868원 이하는 심각한 사고, 2,452,868원 초과 6,090,330원 이하는 매우 심각한 사고로 구분되었다.

순서형 로짓 모형을 적용하여 사고 심각도를 분석한 결과, 수리 기간, 충돌지점 수, 충돌각도의 사선충돌, 상대차량의 부재(단독사고), 대형차 등이 95% 수준에서 유의한 결과를 보였다. 수리 기간 및 충돌지점 수는 길고 많

을수록 사고 심각도에 큰 영향을 미치는 것으로 도출되었고, 충돌각도의 사선충돌은 직각충돌보다 사고 심각도에 영향을 덜 미치는 것으로 도출되었다. 마지막으로, 상대차량의 부재(단독사고) 및 대형차는 소형차에 비해 사고 심각도에 영향을 덜 미치는 것으로 나타났다.

본 연구에서는 k-평균 군집분석과 순서형 로짓 모형을 활용하여 버스 측면부 충돌 사고 심각도 분석을 진행하였는데, 선행연구의 대부분은 사고 심각도의 수준을 사망/중상/경상/부상의 4가지로 나누는 것과 다르게 사고 수리 비용을 심각도 수준으로 설정하여 분석하는 데에 차별성이 있다. 또한, 운전자 입장에서 비용은 중요한 변수이므로, 운전자 입장과 탑승자(안전) 입장을 동시에 고려 가능하다는 장점이 있다.

그러나, 한계점으로는 충돌 부위에 따라 사고 수리 비용이 상이할 수 있기 때문에, 이와 같은 문제는 이상치로 설정하는 방법을 고려해야 할 것으로 보인다.

### 후 기

본 연구는 국토교통부 수소버스 안전성 평가기술 및 장비 개발 사업의 연구비 지원(과제번호 21HBST-B158067-02)에 의해 수행되었습니다.

### 참고문헌

- (1) Barua, U. and Tay, R., 2010, "Severity of urban transit bus crashes in Bangladesh", *Journal of Advanced Transportation*, Vol. 44, No. 1, pp. 34~41.
- (2) 한수산, 박병호, 2011, "순서형 로짓 모형을 이용한 사고 심각도 비교 분석: 청주시를 사례로", *국토계획*, Vol. 46, No. 2, pp. 183~192.
- (3) Feng, S., Li, Z., Ci, Y. and Zhang, G., 2016, "Risk factors affecting fatal bus accident severity: Their impact on different types of bus drivers", *Accident Analysis & Prevention*, Vol. 86, pp. 29~39.
- (4) Briand, A. S., Côme, E., Mohamed, K. and Oukhellou, L., 2016, "A mixture model clustering approach for temporal passenger pattern characterization in public transport", *International Journal of Data Science and Analytics*, Vol. 1, No. 1, pp. 37~50.
- (5) Mondal, M. A. and Rehena, Z., 2019, "Identifying



traffic congestion pattern using K-means clustering technique”, In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU), pp. 1~5.

- (6) 빈미영, 손슬기, 2019, “버스운전자 안전운행지원을 위한 교통사고 분석 연구”, 한국 ITS 학회 논문지, Vol. 18, No. 1, pp. 14~26.
- (7) 이인목, 민재홍, 김경태, 고승영, 2020, “k-means 클러스터링을 활용한 교통카드데이터 기반의 대중교통 이용자 통행패턴 생성”, 한국철도학회 논문집, Vol. 23, No. 3, pp. 204~215.
- (8) 김형준, 장정아, 이인식, 이용주, 오세창, 2020, “버스 실사고 데이터 구축을 통한 대표 버스충돌유형 분석 연구”, 자동차안전학회지, Vol. 12, No. 4, pp. 39~47.
- (9) MacQueen, J., 1967, “Some methods for classification and analysis of multivariate observations”, In Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, Vol. 1, No. 14, pp. 281~297.
- (10) McKelvey, R. D. and Zavoina, W., 1975, “A statistical model for the analysis of ordinal level dependent variables”, Journal of mathematical sociology, Vol. 4, No. 1, pp. 103~120.