

■ 論 文 ■

대형 교통사고 발생지점 유형화와 영향요인 분석에 따른 교통안전대책 방안에 관한 연구

Traffic Safety Countermeasures According to the Accident Area Patterns
and Impact Factor Analysis of the Large-scale Traffic Accident Locations

김 봉 기
(도로교통안전관리공단)

정 현 영
(부산대학교 도시공학과 교수)

고 상 선
(도로교통안전관리공단 교수)

목 차

- | | |
|---|---|
| <p>I. 서론</p> <p>1. 연구 목적</p> <p>2. 연구 방법 및 범위</p> <p>II. 교통사고의 일반적 현황</p> <p>1. 교통사고의 발생 추이</p> <p>2. 대형 교통사고의 발생 추이</p> <p>III. 발생지점 유형화와 영향요인 판별 및 분류분석</p> | <p>1. 판별 및 분류분석의 이론적 고찰</p> <p>2. 대형 교통사고 발생지점 유형화</p> <p>3. 유형별 충돌 형태별 영향요인 판별 및 분류분석</p> <p>4. 발생지점 유형별 특성과 교통사고 예방 대책</p> <p>IV. 결론
참고문헌</p> |
|---|---|

Key Words : 군집분석, 수량화 이론 II류, C&RT 방법, 판별분석, 분류분석, 외적 변수

요 약

본 연구는 대형 교통사고의 발생지점들을 특성별로 유형화하기 위하여 군집분석(Cluster Analysis)을 행하고, 아울러 충돌 형태에 미치는 영향요인에 대한 영향 정도를 판별할 수 있도록 하기 위하여 수량화 이론 II류(Quantification II)와 C&RT(Classification and Regression Trees) 방법에 의해 분석을 실시하여 이에 대한 적합성을 평가함으로써, 정량적 수준의 간략화를 도모하고자 하였다.

그 결과, 발생 지점별 유형화에 따른 4개 집단의 판별 및 분류분석의 충돌 형태별 제반 영향요인들 특성은 집단별로 명확한 차이를 보이는 것으로 나타나, 교통사고에 대해 우선 시행되어야 할 대책과 보완 대책들을 집단별로 체계적으로 제시할 수 있었다.

하지만 상당수 변수들에 결측치가 많아, 막대한 정보 손실이 초래되어 보다 심층적인 분석을 하기 어려웠으나, 이러한 문제점을 해결하기 위해서는 대형 교통사고 조사, 분석 시 표준화된 원 자료 시트의 작성을 의무화할 필요가 있는 것으로 나타났다.

This study divided the large-scale traffic accident locations into its own characteristics by using Cluster Analysis. Also, Quantification II and Classification and Regression Tree methods were used enabling evaluation for the amount of affecting influence by the crash type. After these analyses, we tested the fitness of the results and suggested the simplification of the quantification index.

With the results from the discussed procedure, obvious differences were observed by groups according to the characteristics of crash type from the Discrimination and Classification analysis of divided four groups. Thus, measures and supplementary measures for the traffic accidents could be suggested in groups systematically.

However, a lot of missing values in variables caused a huge loss of data and made this study difficult for more detailed analysis. With this difficulty, recording mandatory log files with a standardized format is also recommended to prevent this problem in advance.

1. 서론

1. 연구 목적

교통사고는 인적, 차량적, 도로·환경적 요인들에 의해 발생되고 있다.

이처럼 다양한 요인들에 의해 발생하는 교통사고 중에서 특히 대형 교통사고¹⁾는 전체 교통사고에서 매우 중요한 위치를 점하고 있음에도 불구하고 심도 있는 연구가 시행되어 오지 않았다. 단지 교통사고 발생 특성과 요인 간에 교차분석(Crosstabulation Analysis)에만 의존한 연구(이주형 외 2명, 1990, 김대웅 외 1명, 1990)가 몇몇 있을 뿐, 대형 교통사고 발생지점의 특성별 유형화와 충돌 형태에 미치는 영향요인에 대해서는 심층적인 분석 및 연구가 수행된 예는 전무한 실정에 있다.

또한 대형 교통사고에 있어 충돌 형태에 미치는 영향요인에 대한 분석은 차후 대형 교통사고 감소를 위한 대책 수립 시 가장 중요하게 고려되어야 할 주된 요인으로 그 무엇보다도 우선적으로 취급되어야 할 만큼 대단히 중요하다.

따라서 본 연구에서는 이러한 대형 교통사고의 예방을 위한 심층적 분석을 시행하였다. 우선 대형 교통사고에 있어 발생지점을 특성별로 유형화하고자 군집분석(Cluster Analysis)을 행하고, 아울러 충돌 형태에 미치는 영향요인에 대한 영향 정도를 판별할 수 있도록 하기 위하여 수량화 이론 II류(Quantification II Method)를 이용한 판별분석과 데이터 마이닝 기법의 일종인 의사결정나무분석법 중의 한가지인 C&RT(Classification and Regression Trees)를 이용한 분류분석을 동시에 구축하고, 이에 대한 적합성을 평가함으로써 대형 교통사고 감소를 위한 정량적 척도의 간략화를 도모하였다.

그 결과 두 분석에서 공통으로 나타난 영향요인과 관련한 대책²⁾은 중점 대책으로, 두 분석 중 한 곳에서만 나타난 영향요인과 관련한 대책은 보완 대책으로 구분하여 제시함으로써, 향후 이들 대책을 시행할 경우 예산의 효율적 집행이 가능한 토대를 제시하고자 하였다. 아울러 대형 교통사고 조사 자료의 구축 과정에서 나타난 문제점에 대한 개선 방안들도 언급 하였다.

2. 연구 방법 및 범위

일반적으로 교통사고는 단독 요인에 의해서 일어날 수도 있지만, 대부분의 경우 여러 가지 요인들이 복합적으로 작용하여 발생하는 경우가 많아 그 원인과 결과를 정확히 규명한다는 것이 매우 어려울 경우가 많다³⁾.

본 연구에서는 대형 교통사고에 관련된 여러 요인들의 복합적인 작용을 알아보기 위해 발생지점을 특성별로 유형화하고, 아울러 충돌 형태에 미치는 영향요인의 영향 정도를 판별 및 분류할 수 있는 분석을 구축하고자 하였다.

이를 위해, 먼저 총 9차례 발간된 "대형 교통사고 사례 분석" 보고서⁴⁾ 상에 나타난 대형 교통사고 사례 300건에 대하여 일차적으로 30개 항목의 자료를 구축하여 자료의 동질성을 분석해 본 결과, 자료가 비교적 충실하게 구축된 충돌 형태를 포함한 7개 항목에 대해 249개 지점을 분석대상으로 하였으며, 이들을 유형화하기 위하여 군집분석을 시행하였다.

또한 교통사고 충돌 형태별로 법규위반 유형, 사고 발생 시간대별, 총 차로수, 노면 포장상태, 도로 중앙부 형태, 타 차량 제공요인 등의 6가지의 영향요인을 중심으로 수량화 이론 II류 분석법을 이용하여 인자득점 및 편상관계수(partial correlation)⁵⁾ 등을 산출하였으며, 경계점(cutting score)⁶⁾과 적중률(hit ratio)⁷⁾ 등으

1) 경찰청, 교통사고 처리 지침 제2조, 제3호, '대형 교통사고라 함은 사망 3명 또는 부상 20명 이상의 사고와 기타 사회 물의를 야기한 사고를 말한다.'라고 규정하고 있음.

2) 현재까지 국내에서 제시되고 있는 각종 교통안전대책들의 대부분은 기 선진국에서 시행되고 있는 대책들과 거의 유사함.

3) 교통사고의 요인으로는 인적 요인과 차량적 그리고 도로·환경적 요인이 별개 내지는 복합적으로 작용하며, 선진국과 개발도상국 그리고 후진국별로 그 비율은 서로 다르게 나타나고 있으며, 이 중 우리나라의 경우가 유달리 인적 요인이 매우 높은 것으로 나타나고 있음. 이는 교통사고 조사 시 교통법규 위반 유형에 따라 가해자와 피해자를 구분하는 형태로 조사가 주로 이루어지고 있는 반면에, 차량의 결함이나 도로·환경적 요인에 대한 과학적인 분석이 미진함에 따른 것으로, 이를 교통사고 기록의 한계로 보는 견해가 있음.

4) '대형 교통사고 사례 분석' 보고서는 1989년, 1991년, 1992년, 1997년, 2000년, 2001년, 2002년, 2003년, 2004년까지 9차례 발간되었음.

5) 다중 상관분석에서 다른 변수들과의 관계는 고정되고 두 변수만의 관계에 대한 강도를 나타낸 것을 의미함.

6) Cutting Score(경계점) = $\frac{N_2 C_1 + N_1 C_2}{N_1 + N_2}$

N : 집단 1의 표본수, C : 집단 1의 중심값(Centroid)

7) Hit Ratio(적중율) = $\frac{\sum P_{ii}}{T} \times 100(\%)$

P_{ii} : 집단 1을 집단 1으로 올바르게 판별한 사례수, T : 전체 사례수

로 신뢰성을 검증한 후 최적의 판별분석을 제시하고자 하였다.

아울러 지니 지수(Gini Index)에 의한 이진분리(binary split)⁸⁾를 수행하는 C&RT 분석법을 이용하여 최적의 분류분석 구축을 도모하였다.

II. 교통사고의 일반적 현황

1. 교통사고의 발생 추이

교통사고에 대한 정의는 여러 학자들에 따라 약간의 차이가 있으나, 대부분 차량 등에 의하여 사람을 사상케 하거나 물건을 손상하는 것이라고 규정하고 있는 경우가 일반적이다.

이에 대한 정의를 도로교통법 제50조 제1항에 의하여 보면, 자동차의 교통으로 인하여 사람을 사상케 하거나 물건을 손상하는 것으로 정의하고 있으며, 이에 는 사망, 중상, 경상, 부상, 충돌, 추돌, 접촉, 전도, 전복, 추락 등의 10가지 유형으로 구분하고 있다.

이에 따라 1994년을 기점으로 과거 10년간 우리나라 교통사고 발생 추이를 <표 1>에서 살펴보면, 교통사고 발생건수는 1994년 266,107건에서 2003년 240,832

건으로 10년 연평균 0.8%씩 감소하였고, 사망자 역시도 10,087명에서 7,212명으로 3.6%씩 감소하였으나, 부상자의 경우는 350,892명에서 376,503명으로 1.1%씩 증가한 것으로 나타나고 있다.

특히 자동차 등록대수가 1,000만대를 돌파한 1997년도에는 교통사고가 총 246,452건이 발생하여, 사망자 11,605명, 부상자 343,159명의 피해가 야기 되었으며, 이를 1일 평균으로 환산하여 보면, 615건의 사고가 발생하여, 32명이 사망하고, 786명이 부상을 입었던 것으로 나타나고 있다.

또한 인구 10만 명 당 사망자수와 차량 1만대 당 사망자수에 있어서는 각각 4.2%와 12.0%씩 감소하는 것으로 나타났다.

2. 대형 교통사고의 발생 추이

다음에 제시된 <표 2>는 1994년부터 2003년까지의 10년간 발생한 전체 교통사고와 대형 교통사고의 발생 추이를 나타낸 것이다.

이를 보면, 대형 교통사고는 매년 전체 교통사고의

<표 2> 전체 교통사고와 대형 교통사고 발생 추이 (단위 : 건, %)

구분 \ 연도별	전체 교통사고 ^{자료1}		대형 교통사고 ^{자료2}	
	발생 건수	사망자	발생 건수	사망자
1994	266,107	627	221(0.08)	627
1995	248,865	772	259(0.10)	772
1996	265,052	732	256(0.10)	732
1997	256,452	595	215(0.09)	595
1998	239,721	477	182(0.08)	477
1999	275,938	399	159(0.06)	399
2000	290,481	624	215(0.07)	624
2001	260,579	301	142(0.06)	301
2002	231,026	256	108(0.05)	256
2003	240,832	247	119(0.04)	247
연평균 증가율(%)	-0.8	-8.4	-6.1	-8.4

주 : () 내는 전체 교통사고에 대한 대형 교통사고의 비율(%)을 나타낸 것임
 자료 : 1. 도로교통안전관리공단, 2004년판 교통사고 통계분석, 2004, p.4.
 2. 도로교통안전관리공단, 2004년판 대형 교통사고 사례 분석, 2004, p.16.

<표 1> 교통사고 발생 추이

구분 \ 연도별	교통사고			사망자	
	발생건수 (건)	사망 (명)	부상 (명)	인구 10만명당 (명/10만명)	차량 1만대당 (대/1만대)
1994	266,107	10,087	350,892	23	14
1995	248,865	10,323	331,747	23	12
1996	265,052	12,653	355,962	27	13
1997	246,452	11,603	343,159	25	11
1998	239,721	9,057	340,564	19	9
1999	275,938	9,353	402,967	20	8
2000	290,481	10,236	426,984	21	8
2001	260,579	8,097	386,539	17	6
2002	231,026	7,222	348,149	15	5
2003	240,832	7,212	376,503	15	4
연평균 증가율 ⁹⁾ (%)	-0.8	-3.6	1.1	-4.2	-12.0

자료 : 도로교통안전관리공단, 2004년판 교통사고 통계분석, 2004, p.17.

8) 부모마디로부터 자식마디가 2개만 형성되도록 함.

9) 10년 연평균 증가율 = $\left\{ \left(\frac{b}{a} \right)^{\frac{1}{n}} - 1 \right\} \times 100$

b: 당해 년도 값, a: 기준년도 값, n: 당해 년도-기준 년도

0.07%씩을 접하는 비율로 발생하였으며, 그 건수 면에서는 1995년과 1996년에 259건으로 가장 많이 발생하였다가, 2002년에는 최저치인 107건으로 10년 연평균 6.1%씩의 비율로 감소한 것으로 나타났다.

아울러 사망자의 경우도 1995년의 경우 772명으로 가장 높게 나타났다가, 2003년에는 247명으로 최저치를 나타내는 등 연평균 8.4%씩 감소한 것으로 나타나고 있으나 여전히 많은 사망자가 발생하고 있는 실정에 있다.

III. 발생지점의 유형화와 영향요인 판별 및 분류분석

1. 판별 및 분류분석의 이론적 고찰

1) 판별분석

수량화 모델은 자료(data)의 성격이 주로 정성적이어서, 정량적 형태의 자료를 적용하는 일반적인 통계 분석방법을 적용토록 함으로서, 수치로 나타낼 수 없는 질적인 것에 수량을 부여하여 목적변수에의 정성적 영향요인을 판별하기 위한 방법으로, 이러한 정성적 자료들에 대해서는 요인별 분석을 가능케 한다.

본 연구에서는 질적 요인에 의한 "질적 외적 기준"을 예측 혹은 판별하기 위한 방법인 수량화 이론 II류를 적용하여 분석하고자 하였다.

이러한 수량화 이론 II류는 R개의 정성적 속성에 관한 지식을 이용하여, 각각의 개체가 T군의 어디에 속하는 것인가를 판별하는 판별문제를 해석하는 모델이다.

즉, 정성적 설명변수인 항목(item)의 범주(category)로부터 외적 기준의 분류를 실시하는 수법으로, 일반적으로 어떤 무한 모집단에서 추출한 자료에 대하여 외적 기준을 가장 잘 설명할 수 있도록 범주에 득점을 부여한다. 이때의 설명변수는 상대비이고, 그 득점은 유형 점수(category score)이며, 편상관계수의 범위(range)가 클수록, 외적 기준에 미치는 영향이 크다고 판단할 수 있다.

그리고 표본수가 적은 경우에도 항목의 범주에 반응하면 분석이 가능해지고, 외적 기준과 설명변수의 인과관계를 정량화 할 수 있다는 점에서 대단히 편리하다.

각 개체가 R개 범주의 k개 선택치 중에서 어느 하나에 반응한다고 하는 상황을 가정할 때, i 번째의 개체가 j 번째의 속성에 관하여 k 번째의 범주에 반응(회답)

했을 때만을 1, 기타 kj-1개의 범주에 반응을 했을 때에는 0을 취하는 $\delta_i(jk)$ 가 되는 양을 도입하면,

$$\delta_i(jk) = \begin{cases} 1: \text{Category } k \text{에 반응(회답)했을 때} \\ 0: \text{이 외의 Category에 반응(회답)했을 때} \end{cases} \quad (1)$$

이 $\delta_i(jk)$ 에 관해서는 다음과 같은 관계식이 성립한다.

$$\sum_{k=1}^{k_i} \delta_i(jk) = 1 \quad (2)$$

$$\sum_{i=1}^n \delta_i(jk) = n_{jk} \quad (3)$$

$$\sum_{k=1}^{k_i} \sum_{i=1}^n \delta_i(jk) = n \quad (4)$$

또한 개체 i가 j속성의 k 범주에 반응했다고 하면,

$$\delta_i(jk)\delta_i(jk) = \begin{cases} 0 & (k \neq k) \\ 1 & (k = k) \end{cases} \quad (5)$$

지금 R개의 각각의 속성을 kj개의 각각의 범주에 대하여 $X_{jk}(j=1, \dots, k, k=1, \dots, kj)$ 가 되는 수치를 부여 했을 때, 개체 i에 대한 새로운 합성변수를 다음과 같이 정의한다.

$$\alpha_i = \sum_{j=1}^R \sum_{k=1}^{k_j} \delta_i(jk)X_{jk} \quad (6)$$

- 여기에서 α_i : 합성변량
- $\delta_i(jk)$: dummy변수
- X_{jk} : 고유 vector

즉, R개의 정성적 속성에 관한 지식을 사용하여 각각의 개체가 T개 군의 어딘가에 속하고 있는가를 판별하는 판별문제를 해석하는 모델이다.

만약 식(6)의 X_{jk} 치의 부여방법이 완벽에 가깝게 되면, T개 군의 구분을 가로축에, α 를 세로축에 두면, T와 α 의 상관비 η 혹은 η^2 은 1에 가까운 수치를 취한다. 따라서 이때에는 급간 분산을 전 분산으로 나눈 값인 $\eta^2 = \frac{\delta b^2}{\delta^2}$ 이 되는 η^2 이 최대가 되도록 X_{jk} 의 수치를 결정하게 된다.

따라서 식(6)의 경우,

$$\delta^2 = \frac{1}{n} \sum_{i=1}^n \alpha_i^2 \bar{\alpha}^2 \quad (7)$$

$$\delta b^2 = \sum_{i=1}^T \frac{n_i}{n} (\bar{\alpha}_i - \bar{\alpha})^2 \quad (8)$$

$$\bar{\alpha} = \frac{1}{n} \sum_{i=1}^n \alpha_i \quad (9)$$

$$\bar{\alpha}_i = \frac{1}{n} \sum_{k(i)=1}^{n_i} \alpha_i(k) \quad (10)$$

로 된다.

여기에서 η^2 을 최대로 하는 X_{jk} 를 구하기 위해서는 X_{uv} 로 편미분하여 각각을 0으로 두면 된다.

즉,

$$\frac{\partial \eta^2}{\partial X_{uv}} = 0 \quad (u=1, \dots, R, v=1, \dots, ku) \quad (11)$$

$$\frac{\delta^2 \frac{\partial \delta b^2}{\partial X_{uv}} - \delta b^2 \frac{\partial \delta^2}{\partial X_{uv}}}{(\delta^2)^2} = 0 \quad (12)$$

$$\frac{\partial \delta b^2}{\partial X_{uv}} = \eta^2 \frac{\partial \delta^2}{\partial X_{uv}} \quad (u=1, \dots, R, v=1, \dots, ku) \quad (13)$$

여기서

$$\frac{\partial \delta b^2}{\partial X_{uv}} = \frac{2}{n} \sum_j \sum_k \left[\sum \frac{g^t(jk)g^t(uv)}{n_i} - \frac{1}{n} n_{jk}n_{uv} \right] X_{jk} \quad (14)$$

$$\frac{\partial \delta^2}{\partial X_{uv}} = \frac{2}{n} \sum_j \sum_k \left[\sum f_{jk}(uv) - \frac{1}{n} n_{jk}n_{uv} \right] X_{jk} \quad (15)$$

따라서 식(13)은

$$\sum_{i=1}^R \sum_{k=1}^{k_i} h_{uv}(jk) X_{jk} = \eta^2 \sum_{i=1}^R \sum_{k=1}^{k_i} \left[f_{uv}(lm) - \frac{1}{n} n_{lm}n_{uv} \right] X_{lm} \quad (16)$$

단,

$$h_{uv}(jk) = \sum_{i=1}^T \frac{g^t(jk)g^t(uv)}{n_i} - \frac{1}{n} n_{jk}n_{uv}$$

식(16)을 행렬식으로 나타내면

$$HX = \eta^2 FX \quad (17)$$

여기서 $H = [h_{uv}(jk)]$: 정방행렬

$$F = \left[f_{uv}(lm) - \frac{1}{n} n_{jk}n_{uv} \right] : \text{정방행렬}$$

식(15)를 변형하면

$$F^{-1}HX = \eta^2 X \quad (18)$$

$$|F^{-1}H - \eta^2 I| = 0 \quad (19)$$

여기서 η^2 을 최대로 하는 고유벡터를 구하면 된다.

2) 분류분석(C&RT)

분류모델의 대표적인 분석방법인 의사결정나무분석법은 의사결정규칙(decision rule)을 도표화하여, 관심대상이 되는 집단을 몇 개의 소집단으로 분류(classification)하거나 예측(prediction)을 수행하는 분석방법이다.

이러한 의사결정나무분석법 중 본 연구에서는 C&RT 분석법을 사용하였다. 이 분석법은 임의의 한 개체가 목표변수의 i 번째 범주로부터 추출되었고, 그 개체를 목표변수의 i 번째 범주에 속한다고 오분류(misclassification)할 확률은 $P(i)P(j)$ 가 된다. 여기에서 $P(i)$ 는 각 마디에서 한 개체가 목표변수의 i 번째 범주에 속할 확률이다. 이러한 오분류 할 확률을 모두 더하면

$$G = \sum_{i=1}^c \sum_{j \neq i} P(i)P(j) \quad (20)$$

을 얻을 수 있고, 이는 이와 같은 분류규칙 하에서 오분류 확률의 추정치(estimate)가 되며, C 는 목표변수의 범주수를 나타낸다.

또한 이산형 목표변수에 대해 분리 기준으로 사용하는 지니 지수는 각 마디에서 불순도(impurity) 또는 다양도(diversity)를 재는 측도 중의 하나로써

$$G = \sum_{j=1}^c P(j)(1 - P(j)) = 1 - \sum_{j=1}^c P(j)^2 = 1 - \sum_{j=1}^c \left(\frac{n_j}{n} \right)^2 \quad (21)$$

와 같이 표현될 수 있다. 여기에서 n 은 그 마디에 포함되어 있는 관찰치 수를 말하고, n_i 는 목표변수의 i 번째 범주에 속하는 관찰치 수를 말한다. 지니 지수는 n 개의 원소 중에서 임의로 2개를 추출하였을 때, 추출된 2개가 서로 다른 그룹에 속해 있을 확률을 의미하며, Simpson의 다양도 지수(diversity index)로도 알려져 있다. 목표변수의 범주가 2개인 경우에는 지니 지수는 다음과 같이 표현될 수 있다.

$$G = 2P(1)P(2) = 2 \left(\frac{n_1}{n} \right) \left(\frac{n_2}{n} \right) \quad (22)$$

이는 카이제곱 통계량을 사용하는 것과 같은 결과를 갖는다.

C&RT는 이 지니 지수를 가장 많이 감소시켜주는 예측변수와 그 변수의 최적 분리를 자식마디로 선택하는데, 지니 계수의 감소량은 다음과 같이 계산된다.

$$\Delta G = G - \frac{n_L}{n} G_L - \frac{n_R}{n} G_R \quad (23)$$

여기서 n 은 부모 마디의 관측치 수를 말하고, n_R 과 n_L 은 각각 자식 마디의 관측치 수를 의미한다. 즉, 자식 마디로 분리되었을 때의 불순도가 가장 작도록, 다시 말하자면, 순순도가 가장 크도록 자식 마디를 형성하는 것이며, 이는 자식 마디에서의 불순도의 가중 합을 최소화하는 것과 동일하다.

그리고 의사결정나무를 얻은 후에 사용자가 직접 가지치기를 수행할 수도 있지만, C&RT 알고리즘에서는 가지치기(pruning) 기준을 사전에 지정을 할 수 있다.

그 기준 중의 한가지인 표준오차(standard error) 규칙은 T 번째 마디에서의 위험을 $R(T)$ 라 하고, $R(T)$ 의 표준오차를 $SE(R(T))$ 라 할 때, T_k 번째까지의 나무구조가 T_0 번째까지의 마디를 모두 포함하고 있다고 가정하면,

$$R(T_k) \leq R(T_0) + \alpha SE(R(T_0)) \quad (24)$$

인 경우에 있어 T_k 번째까지의 나무구조를 선택하는 것을 의미한다. 즉, $\alpha SE(R(T_0))$ 정도의 위험 감소량은 유의하지 않은 것으로 생각하여 더 이상 나무구조를 형성하지 않게 된다는 것을 의미한다.

또한 C&RT 분석법은 최소 위험(minimum risk)을 가지는 나무구조를 의사결정나무로 선택하며, 각 마디에서 선택된 예측변수에 결손 값이 존재하는 경우에는 다른 예측변수를 이용하여 결손 값을 대체할 수 있다. 결손 값에 대한 대체규칙(surrogate rule)을 지정해 놓으면 의사결정나무의 형성과정 또는 새로운 자료의 분류과정에서 이 규칙에 따라서 각 예측변수의 결손 값을 대체할 수 있다.

대체규칙을 선택할 때, 선택된 변수 대신에 다른 변

수를 선택하였을 경우 분류 또는 예측이 일치하는 정도를 의미하는 연관성(association)이 큰 순서대로 다른 예측변수로 대체하면 그 변수가 가지는 값에 따라 각 개체를 자식 마디에 할당하였을 때의 분류가 어느 정도가 일치하는지를 나타내어 줄 수 있다.

2. 대형 교통사고 발생지점 유형화

대형 교통사고의 발생지점 유형화를 위한 군집분석에서 먼저, 지점 특성별 적정 집단 수는 <표 3>에서와 같이, 그 증가폭이 급격히 변화하는 단계에서 최적의 유형화 집단수를 결정하거나, 집단수를 판단하는데 사용되는 부분결정계수(semi-partial R square)값과 초입방향 군집판정 기준(cubic clustering criterion)의 값을 동시에 고려하여 판단하였다.

그 결과 적정 집단수는 4개의 집단인 것으로 나타났으며, 이를 근거로 대형사고 발생지점 249개를 대상으로 군집분석을 행해 집단을 분류하였다.

<표 3> 최적 유형화 수준의 지표

군집수	부편결정계수	초입방향 군집판정 기준
10	0.0223	7.81
9	0.0234	6.90
8	0.0276	5.82
7	0.0284	5.13
6	0.0424	3.31
5	0.0567	0.84
4	0.0739	-1.50
3	0.0905	-3.00
2	0.1214	-2.20
1	0.1819	0.00

여기서 분류된 집단의 각 지점들에 대해 공통적으로 나타난 집단별 충돌 형태별 주요 영향 요인들을 유형화와 영향요인 분석을 위하여 도입된 항목의 변수들을 중심으로 정리하여 <표 4>에 제시하였다.

이를 보면, 먼저 1집단에 속하는 87개 지점들의 충돌 형태별 특성을 살펴보면, 인적 요인의 범규 위반 유형에 있어서는 차량 충돌과 차량 외 충돌의 경우 주취 운전이 많았으며, 도로·환경적 요인 중 사고 발생 시간대별에 있어서는 차량 충돌과 차량 외 충돌의 경우 야간 시간대인 20시부터 익일 06시 59분 사이가 많았고, 특히 차량 충돌의 경우 총 차로수에서는 도로 폭이 매우 넓은 4차로 이상인 경우, 노면 포장상태로는 아스

〈표 4〉 집단별 충돌 형태별 주요 영향 요인

집단	변수 유형		충돌 형태별		
			차량 충돌	차량 외 충돌	전도, 전복, 추락 등
1집단 (87개 지점)	인적	법규위반 유형	주취운전	주취운전	
	도로·환경적	사고 발생 시간대별	20시~익일 06시59분	20시~익일 06시59분	
		총차로수	4차로 이상		
		노면 포장상태	아스팔트		
2집단 (35개 지점)	인적	법규위반 유형		주취운전	
	도로·환경적	사고 발생 시간대별		20시~익일 06시59분	20시~익일 06시59분
3집단 (28개 지점)	인적	법규위반 유형	중앙선 침범 제한속도 위반	제한속도 위반	제한속도 위반
	도로·환경적	사고 발생 시간대별	20시~익일 06시59분	20시~익일 06시59분	
4집단 (99개 지점)	인적	법규위반 유형	중앙선 침범 제한속도 위반	제한속도 위반	앞지르기 방법 및 금지 위반
	도로·환경적	총차로수	4차로 이상		
		노면 포장상태	아스팔트		

팔트인 경우가 많은 것으로 나타났다.

다음으로 2집단에 속하는 35개 지점들의 충돌 형태별 특성을 살펴보면, 인적 요인인 법규 위반 유형에 있어서는 주취운전이 차량 외 충돌인 경우 많았으며, 도로·환경적 요인 중 사고 발생 시간대별에 있어서는 야간 시간대인 20시부터 익일 06시 59분 사이가 차량 외 충돌과 전도, 전복, 추락 등의 경우 많았던 것으로 나타났다.

그리고 3집단에 속하는 26개 지점들의 충돌 형태별 특성을 살펴보면, 인적 요인인 법규 위반 유형에 있어서는 제한속도 위반이 모든 충돌 형태에서 많이 나타났으며, 중앙선 침범의 경우는 차량 충돌의 경우 많은 것으로 나타났다.

또한 도로·환경적 요인 중 사고 발생 시간대별에

있어서는 야간 시간대인 20시부터 익일 06시 59분 사이가 차량 충돌과 차량 외 충돌의 경우에 많았던 것으로 나타났다.

끝으로 4집단에 속하는 99개 지점들의 충돌 형태별 특성을 살펴보면, 인적 요인인 법규 위반 유형에 있어서는 제한속도 위반의 경우 차량 충돌과 차량 외 충돌의 경우 많은 것으로 나타났으며, 중앙선 침범의 경우는 차량 충돌의 경우, 앞지르기 방법 및 금지 위반은 전도, 전복, 추락 등의 경우 많은 것으로 나타났다. 또한 도로·환경적 요인 중 총 차로수에서는 도로 폭이 매우 넓은 4차로 이상인 경우, 노면 포장상태로는 아스팔트인 경우가 차량 충돌의 경우 많은 것으로 나타나는 등 각 집단별 충돌 형태별 특성에 따른 인적 및 도로·환경적 영향 요인들이 비교적 명확하게 구분되어 나타났다.

3. 유형별 충돌 형태별 영향요인 판별 및 분류 분석

1) 판별분석

249개의 조사대상 지점을 이상과 같이 4개의 집단으로 구분하여 수량화 이론 II류에 의하여 각 집단별 대형 교통사고 충돌 형태별 판별분석을 구축하여 본바, 〈표 5〉와 같이 나타났다.

먼저 외적 변수(external variable)인 충돌 형태의 빈도 순위를 보면, 제1집단에는 차량 충돌, 차량 외 충돌, 전도, 전복, 추락 등의 순으로, 제2집단은 차량 외 충돌, 전도, 전복, 추락 등의 순으로, 제3집단은 차량 외 충돌, 전도, 전복, 추락 등, 차량 충돌의 순으로, 그리고 제4집단은 차량 충돌, 전도, 전복, 추락 등, 차량 외 충돌의 순으로 나타났다.

그리고 제1축에 대한 판별분석의 상관비는 1집단은 0.43033에서, 2집단은 0.49648에서, 3집단은 0.80411에서, 4집단은 0.35493에서 외적 변수의 구분 요인을 가장 잘 설명하고 있어 판별분석으로서 유효하다고 판단되었다.

다음으로 편상관계수가 높은 순서로 볼 때, 1집단은 도로·환경적 요인 > 인적 요인의 순으로, 2집단은 인적 요인 > 도로·환경적 요인의 순으로, 3집단은 인적 요인 > 도로·환경적 요인의 순으로, 그리고 4집단은 인적 요인과 도로·환경적 요인이 순차적으로 영향을 미치고 있음을 알 수 있었다.

그리고 외적 변수의 유형 점수에 있어서는 1집단은

<표 5> 수량화 이론 II류에 의한 충돌 형태별 판별분석(집단별)

항목	범주	1집단				2집단				3집단				4집단			
		빈도	수량화값	범위	편상관 계수	빈도	수량화값	범위	편상관 계수	빈도	수량화값	범위	편상관 계수	빈도	수량화값	범위	편상관 계수
법규 위반 유형	1. 신호 위반									(2)	-0.69682			(11)	-0.67711		
	2. 중앙선 침범									(4)	-1.50292			(53)	-0.35916		
	3. 제한속도 위반									(21)	0.32626			(20)	-0.00855		
	4. 앞지르기방법, 금지 위반									(1)	0.55395			(13)	1.47719		
	5. 차로 위반													(2)	3.72576		
	6. 추월운전	(26)	-0.55023	2.14945	0.46877	(13)	-0.95856	3.54585	0.56562								
	7. 안전운전 의무 불이행	(51)	0.55668		(2)	(21)	0.47019		(1)		2.05687	0.77573	(1)			4.40287	0.44578
	8. 철길간널목 통과방법 위반	(9)	-1.59277														
	9. 무면허운전, 보행자보호 의무 불이행 등	(1)	0.25003			(1)	2.58729										
사고발생 시간대별	1. 07:00~08:59	(7)	0.33601			(1)	1.33930			(3)	1.62432			(21)	0.01845		
	2. 09:00~16:59	(28)	-0.10218	0.55288	0.15043	(12)	0.29851	2.46086	0.46604	(9)	0.10639	2.11484	0.76746	(38)	-0.13394	0.26088	0.07757
	3. 17:00~19:59	(8)	0.45070		(6)	-1.12156		(2)	(4)	0.01396		(2)	(18)	0.10609		(5)	
	4. 20:00~06:59	(44)	-0.07037		(6)	0.11299			(12)	-0.49052			(22)	0.12694			
충차로수	1. 1차로	(9)	2.54396	2.99384	0.52788	(2)	-0.07343	0.35628	0.13878	(1)	-0.73396	0.88865	0.39643	(4)	0.31992	0.49735	0.12690
	2. 2차로	(23)	0.08061		(1)	-0.15749		(5)	(8)	-0.27565		(6)	(46)	0.16118		(4)	
	3. 4차로 이상	(55)	-0.44998			(15)	0.19878			(19)	0.15469			(49)	-0.17743		
노면 포장상태	1. 아스팔트	(73)	0.01633	2.01135	0.22134	(33)	-0.06774	1.18544	0.29264	(26)	-0.08558	1.19816	0.45849	(91)	-0.09927	3.65111	0.35560
	2. 콘크리트	(12)	-0.37250		(5)	1.11770		(4)	(2)	1.11258		(5)	(6)	0.32158		(2)	
	3. 비포장	(2)	1.63885										(2)	3.55184			
도로 중앙부 형태	1. 중앙분리대	(21)	-0.27815	1.10642	0.24188	(7)	-1.08307	2.24360	0.36476	(8)	-0.32892	1.31983	0.54428	(14)	-0.35695	1.39203	0.18673
	2. 황색실선	(51)	0.21879		(4)	0.19569		(3)	(16)	-0.04512		(4)	(74)	0.09625		(3)	
	3. 황색점선	(5)	-0.88764			(1)	1.16053			(1)	0.99092			(4)	-0.67542		
	4. 철길 없음	(10)	-0.08787			(4)	0.48000			(3)	0.78743			(7)	0.71661		
타차량 제공요인	1. 동일방향 직진, 회전 중	(22)	0.00055	1.34660	0.38127							1.45209	0.70771	(21)	0.07430	0.11630	0.02452
	2. 대향 직진, 회전 중	(28)	0.58457		(3)								(3)	(66)	-0.01601		(6)
	3. 측면에서 직진, 회전 중	(22)	-0.76202											(12)	-0.04200		
	4. 신호대기, 주·정차 일시정지 등	(10)	-0.01420							(6)	-1.14093						
	5. 없음	(5)	0.10527							(22)	0.31116						
Outside variable	1. 차량 충돌	(81)	-0.12844							(6)	-1.51925			(90)	-0.15605		
	2. 차량 외 충돌	(3)	-0.00130			(21)	-0.58084			(15)	0.09230			(4)	-0.32290		
	3. 전도, 전복, 추락 등	(3)	3.46905			(14)	0.87125			(7)	1.10444			(5)	2.55066		
Eta-square (Correlation ratio)		0.43033				0.49648				0.80411				0.35493			

전도, 전복, 추락 등 > 차량 외 충돌 > 차량 충돌의 순으로, 2집단은 전도, 전복, 추락 등 > 차량 외 충돌의 순으로, 3집단은 전도, 전복, 추락 등 > 차량 외 충돌 > 차량 충돌의 순으로, 그리고 4집단은 전도, 전복, 추락 등 > 차량 외 충돌 > 차량 충돌의 순으로 나타났다.

유형점수에서 판단해 보면, 차량 충돌과 차량 외 충돌에 크게 영향을 미치고 있는 변수들은 각 항목 변수 중에 음의 부호를 가진 범주로서 예를 들면, 제1집단에서 편상관계수가 가장 높은 총 차로수 항목에 있어서는

4차로 이상의 경우가 여기에 해당된다.

반면 전도, 전복, 추락 등에 크게 영향을 미치는 변수는 각 항목 변수 중에 양의 부호를 가진 범주로서, 1, 2차로의 경우가 해당된다.

2집단의 경우는 편상관계수가 가장 높은 범주위반 유형 항목 변수에 있어서는 추월운전인 경우가 차량 외 충돌에 크게 영향을 미치고 있는 것으로 나타났다.

반면 전도, 전복, 추락 등에 크게 영향을 미치는 범주로는 안전운전 의무 불이행과 무면허 운전, 보행자

보호 불이행 등의 범주가 해당되는 것으로 나타났다.

3집단 역시 편상관계수가 가장 높은 항목 변수는 범규위반 유형으로 나타났으며, 차량 충돌에 크게 영향을 미치는 범주는 신호 위반과 중앙선 침범이었다. 차량의 충돌과 전도, 전복, 추락 등에 크게 영향을 미치는 범주는 제한속도 위반과 앞지르기 방법 및 금지 위반인 것으로 나타났다.

끝으로 4집단 역시 편상관계수가 가장 높은 항목 변수는 범규위반 유형으로 나타났으며, 차량 충돌과 차량의 충돌 등에 크게 영향을 미치는 범주는 신호 위반, 중앙선 침범, 제한속도 위반 등이 해당되며, 전도, 전복, 추락 등에 크게 영향을 미치는 범주는 앞지르기 방법 및 금지 위반, 차로 위반 등이 해당되는 것으로 나타났다.

2) 분류분석

다음으로는 전체 249개소 중 1집단에 속하는 87개소 지점에 대해 1개의 목표변수 그리고 6개의 예측변수를 사용하여 분류분석을 구축하고자 하였다.

특히 이산형 목표변수로, n 개의 원소 중에서 임의로 2개를 추출하였을 때 추출된 2개가 서로 다른 그룹에 속해 있을 확률인 지니 지수를 분리 기준으로 이지분리를 수행하여, 이 지니 지수를 가장 많이 감소 시켜주는 예측변수와 그 변수의 최적 분리를 자식마디로 선택하여 분석을 구축하는 C&RT 알고리즘을 이용하여 분석한 결과, <그림 1>과 같은 회귀나무(regression tree)가 형성되었다.

이를 보면 1집단의 경우는 충돌 형태별에 대한 영향요인 분류분석에서는 예측변수로 노면 포장상태, 타 차량 제공 요인, 총 차로수, 범규위반 유형의 4 가지 변수가 채택되었다.

이 중 충돌 형태별에 대해 가장 크게 영향을 미치는 요인변수는 노면포장상태로, 이 예측변수로 분리했을 때 분산의 감소 정도는 0.0095¹⁰⁾로 전체 분산을

0.95% 정도 밖에 줄여주지 못하고 있어 다음 단계에서 타 차량 제공 요인을 예측변수로 하여 재차 분리하였다.

하지만 이 경우도 역시 0.33% 정도 밖에 줄여주지 못하여 다음 단계로 총 차로수를 예측변수로 하여 재차 분리하였다. 그러나 이 경우 역시 0.37% 정도 밖에 줄이지 못해 다음 단계에서 범규위반 유형을 예측변수로 하여 재차 분리하였으나, 이 경우도 역시 0.64% 정도 밖에 줄여주지 못하였으며, 이후 단계에서는 전혀 줄어들지 못하여 범규위반 유형까지의 단계에서 이지분리는 정지하게 되었다.

이에 반해 2집단에 속하는 35개소의 관찰대상과 1개의 목표변수 그리고 5개의 예측변수를 사용하여 분류분석을 구축 하였는바, 예측변수로 범규위반 유형 한 가지 변수만 채택, 분리하여도 전체 분산을 7.16% 정도나 줄여 주고 있는 것으로 나타난 반면에 이후 단계에서는 줄어들지 못하여 이 단계에서 정지하게 되었다.

다음으로 3집단에 속하는 28개소 지점과 1개의 목표변수 그리고 6개의 예측변수를 사용하여 분류분석을 구축 하였는바, 이 경우 역시 2집단과 유사하게 예측변수로 사고발생 시간대별 한 가지 변수만 채택, 분리하여도 전체 분산을 8.4% 정도를 줄여 주고 있는 것으로 나타난 반면에 이후 단계에서는 줄어들지 못하여 이 단계에서 정지하게 되었다.

끝으로 4집단에 속하는 99개소의 지점과 1개의 목표변수 그리고 6개의 예측변수를 사용하여 분류분석을 구축 하였는바, 예측변수로 범규위반 유형, 도로 중앙부 형태의 두 가지 변수가 채택되었다.

이 중 범규위반 유형을 예측변수로 분리했을 때 전체 분산을 1.9% 정도로 미미하게 줄여주고 있어, 다음 단계에서 노면포장 상태를 예측변수로 하여 재차 분리하였다.

하지만 이 단계 역시 0.92% 정도밖에 줄여주지 못하였으며, 이후 단계에서는 전혀 줄어들지 못하여 노면포장 상태까지의 단계에서 분리는 정지하게 되었다.

10) <그림 1> 충돌 형태별 영향요인 분류모델(1집단)의 Improvement = 0.0095의 값임.

계산 과정(소수점 넷째자리까지)

부모마디와 각 자식마디에서의 지니 지수 값

$$G_0 = 1 - (81/87)^2 - (3/87)^2 - (3/87)^2 = 1 - (0.931)^2 - (0.0344)^2 - (0.0344)^2 = 1 - 0.8667 - 0.0011 - 0.0011 = 0.1311$$

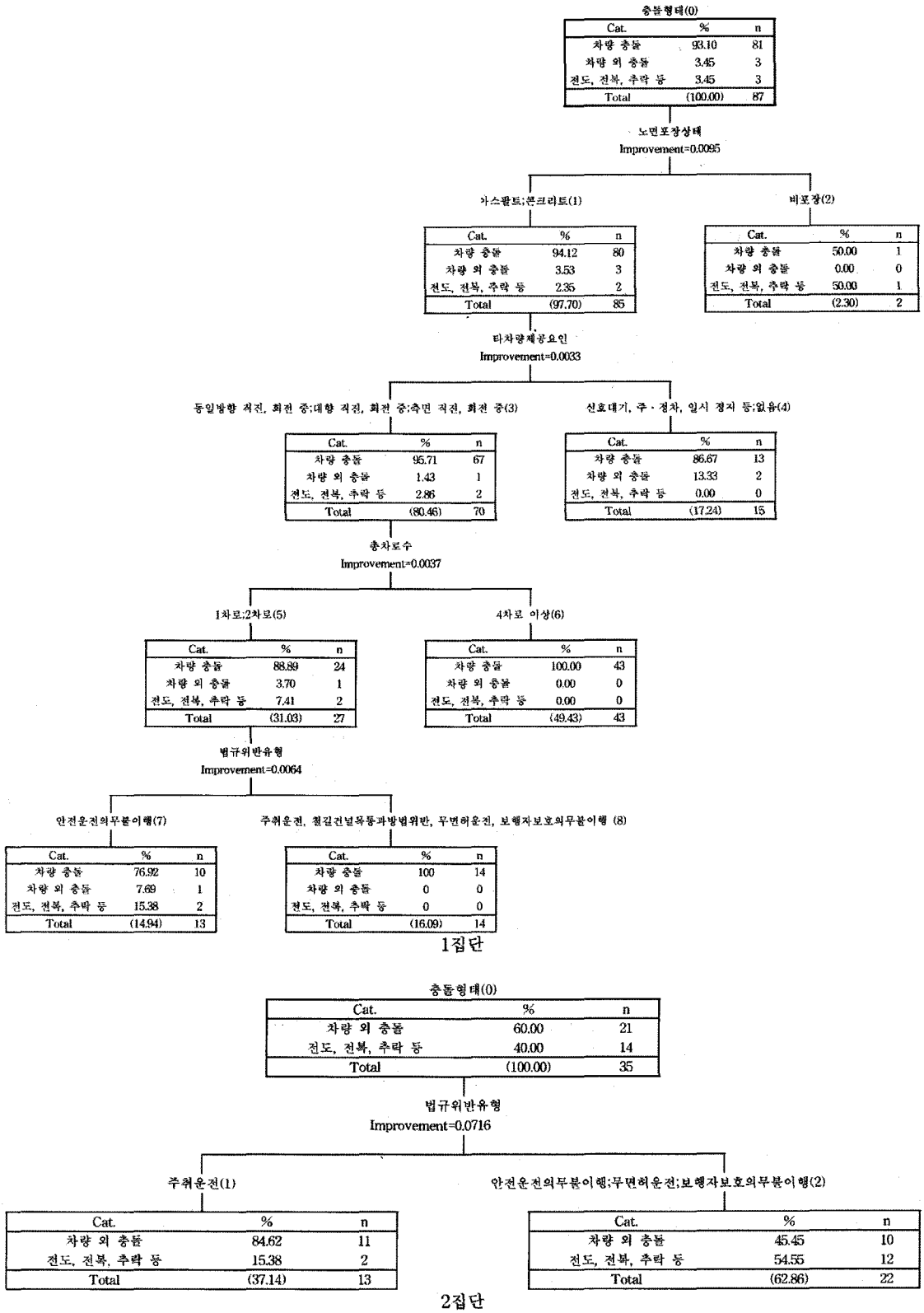
$$G_1 = 1 - (80/85)^2 - (3/85)^2 - (2/85)^2 = 1 - (0.9411)^2 - (0.0352)^2 - (0.0235)^2 = 1 - 0.8856 - 0.0012 - 0.0005 = 0.1127$$

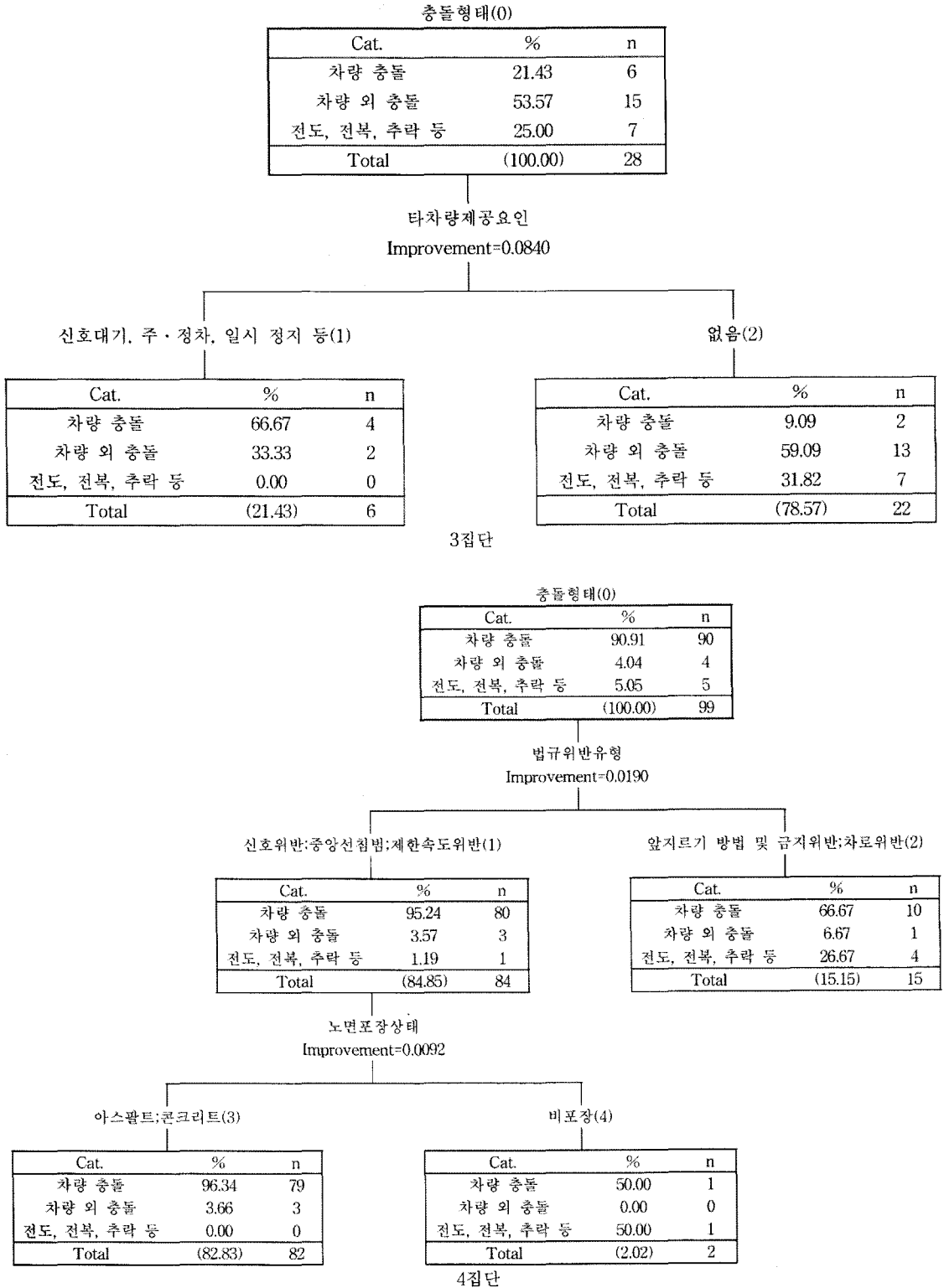
$$G_2 = 1 - (1/2)^2 - (0/2)^2 - (2/2)^2 = 1 - (0.5)^2 - (0)^2 - (0.5)^2 = 1 - 0.25 - 0 - 0.25 = 0.5$$

지니 지수 감소량

$$\Delta G = 0.1311 - (85/87) \cdot 0.1127 - (2/87) \cdot 0.5 = 0.1311 - (0.977 \times 0.1127) - (0.0229 \times 0.5) = 0.1311 - 0.1101 - 0.0114 = 0.0096$$

(1집단)의 Improvement = 0.0095와 0.0001의 차이가 나는 것은 계산 과정상의 소수점 자리 수 차이에 따른 반올림 결과 때문임.





〈그림 1〉 충돌 형태별 영향요인 분류분석(집단별)

주: ()숫자는 마디 번호임.

이러한 일련의 과정을 거쳐 형성된 나무지도 화면인 <그림 1>에서 집단별 분류 특성을 살펴보면, 1집단은 차량 충돌사고의 경우 주취운전, 철길건널목 통과방법 위반, 무면허 운전, 보행자 보호 의무 불이행 등을 한 상태에서, 진행 방향 전방에 타 차량이 동일 방향과 대향 혹은 측면에서 직진 또는 회전 중이어서 이를 피하는 과정에서 정면이나 측면 충돌의 형태로 나타나거나, 도로 폭이 넓어 운전 조건이 매우 좋은 4차로 이상, 아스팔트나 콘크리트 포장 도로 상에서도 많이 발생하는 것으로 나타났다.

또한 2집단은 차량 외 충돌사고의 경우 주취운전을 한 경우가 많이 발생하는 것으로 나타났다. 반면, 3집단은 차량 충돌사고의 경우 타 차량이 진행 방향 전방에서 신호 대기, 주·정차 또는 일시 정지 중인 경우에, 4집단 역시 차량 충돌사고의 경우 법규 위반 유형에 있어서는 신호위반, 중앙선 침범, 제한속도 위반 등을 한 경우 노면 포장상태에 있어서는 아스팔트와 콘크리트 포장인 경우에 많이 발생하는 것으로 나타났다.

4. 발생지점 유형별 특성과 교통사고 예방대책

이상으로 판별분석과 분류분석의 구축을 통해서 분석된 결과와 앞서 <표 4>에서 제시한 집단별 충돌 형

태별 주요 영향 요인을 동시에 고려하여 교통사고 예방 대책을 제시함에 있어, 두 분석에서 공통으로 나타난 부분에 대해서는 중점 대책으로, 두 분석 중 한 개 분석에서만 나타난 부분은 보완 대책으로 구분하여 제시하여 본 바, 다음의 <표 6>과 같이 나타났다.

먼저 1집단의 경우에는 인적 요인과 관련해서는 주취운전에 대한 단속 강화가 차량 충돌사고를 예방하기 위해서는 중점 대책으로, 차량 외 충돌의 경우는 보완 대책으로 시행되어야 할 것으로 나타났다.

도로·환경적 요인과 관련해서는 교통안전시설의 보강이 차량 충돌사고를 예방하기 위해서는 중점 대책으로, 시인성 향상을 위한 교통안전 시설의 개선은 차량 충돌 및 차량 외 충돌의 경우 보완 대책으로 시행되어야 할 것으로 나타났다.

2집단의 경우에는 인적 요인과 관련해서는 주취운전에 대한 단속 강화가 차량 외 충돌사고를 예방하기 위해서는 중점 대책으로, 도로·환경적 요인과 관련해서는 시인성 향상을 위한 교통안전 시설의 개선이 차량 외 충돌 및 전도, 전복, 추락 등의 경우 보완 대책으로 시행되어야 할 것으로 나타났다.

그리고 3집단의 경우에는 인적 요인과 관련해서는 과속 방지턱의 보강, 설치가 차량 충돌의 경우는 중점 대책으로, 차량 외 충돌 및 전도, 전복, 추락 등의 경

<표 6> 발생지점 유형별 교통사고 예방 대책

집단	요인	예방 대책 (충돌 형태별)					
		차량 충돌		차량 외 충돌		전도, 전복, 추락 등	
		중점 대책	보완 대책	중점 대책	보완 대책	중점 대책	보완 대책
1 (87)	인적	·주취운전에 대한 단속 강화			·주취운전에 대한 단속 강화		
	도로·환경적	·도로 조건이 좋는데 따른 교통안전시설의 보강	·시인성 향상을 위한 교통안전시설의 개선		·시인성 향상을 위한 교통안전시설의 개선		
2 (35)	인적			·주취운전에 대한 단속 강화			
	도로·환경적				·시인성 향상을 위한 교통안전시설의 개선		·시인성 향상을 위한 교통안전시설의 개선
3 (28)	인적	·과속 방지턱의 보강, 설치	·도로 중앙부의 정비		·과속 방지턱의 보강, 설치		·과속 방지턱의 보강, 설치
	도로·환경적		·시인성 향상을 위한 교통안전시설의 개선		·시인성 향상을 위한 교통안전시설의 개선		
4 (99)	인적	·도로 중앙부의 정비 ·과속 방지턱의 보강, 설치			·과속 방지턱의 보강, 설치		·도로 중앙부의 정비
	도로·환경적	·도로 조건이 좋는데 따른 교통안전시설의 보강					

주 : 집단 항목 내 ()안의 숫자는 그 집단에 속하는 지점수를 의미함

우는 보완 대책으로 시행되어져야 할 것으로 나타났다.

도로·환경적 요인과 관련해서는 시인성 향상을 위한 교통안전 시설의 개선이 차량 충돌 및 차량 외 충돌의 경우 보완 대책으로 시행되어져야 할 것으로 나타났다.

끝으로 4집단의 경우에는 인적 요인과 관련해서는 과속 방지턱의 보강, 설치가 차량 충돌의 경우는 중점 대책으로, 차량 외 충돌의 경우는 보완 대책으로 시행되어져야 할 것으로 나타났으며, 도로 중앙부의 정비는 차량 충돌의 경우는 중점 대책으로, 전도, 전복, 추락 등의 경우는 보완 대책으로 시행되어져야 할 것으로 나타났다.

또한 도로·환경적 요인과 관련해서는 교통안전시설의 보강이 차량 충돌사고를 예방하기 위해서는 중점 대책으로 시행되어져야 할 것으로 나타나는 등 각 집단별 충돌 형태별 특성에 따른 인적 및 도로·환경적 영향 요인 관련 중점 및 보완 대책들을 비교적 명확하게 구분하여 제시할 수 있었다.

IV. 결론

본 연구는 대형 교통사고에 있어 발생지점을 특성별로 유형화하고, 아울러 충돌 형태에 미치는 영향요인에 대한 영향 정도를 판별하는 판별분석과 분류분석을 구축하였는바, 그 결과는 다음과 같다.

- (1) 본 연구대상 지점들을 유형화 한 결과, 최적 집단수는 4개인 것으로 나타났으며, 분류된 각 집단별 충돌 형태별 특성에 따른 인적 및 도로·환경적 영향 요인들은 비교적 명확하게 구분되어 나타났다.
- (2) 발생지점별 유형화에 따른 4개 집단의 판별 및 분류분석에 있어 충돌 형태별 제반 영향 요인들의 특성은 집단별로 명확한 차이를 보이는 것으로 나타나, 각 집단별 충돌 형태별 특성에 따른 인적 및 도로·환경적 영향 요인 관련 중점 및 보완 대책들을 비교적 명확하게 구분하여 제시할 수 있었다. 아울러 향후 이들 대책을 시행할 경우 예산의 효율적 집행이 가능한 토대를 제시 하였던 점 또한 본 연구의 성과라 할 수 있을 것이다.

이상의 연구 결과 나타난 제반 문제점으로는 대형 교통사고를 조사, 분석하는 측면에서 볼 때 상당수의 변수들에 걸쳐측정이 많아, 분석 과정에서 막대한 정보 손실이 초래되게 되고 이로 인하여 보다 심층적인 분석

을 하기 어려웠다는 점이다.

따라서 이러한 문제점들을 해결하기 위해서는 대형 교통사고 조사, 분석 시 표준화된 원 자료 시트의 작성을 의무화할 필요가 있는 것으로 판단되었다.

참고문헌

1. Claude Seidman 저, 권오주, 이정무 역(2001), Microsoft SQL Server 2000 데이터 마이닝, 정보문화사, pp.115~201.
2. 경찰청, 교통사고 처리 지침
3. 김대용·임채문(1990), 도로교통에 있어서 운전자 시각특성 분석, 대한교통학회지, 제8권 제2호, 대한교통학회, pp.7~25.
4. 김종섭(2003), SAS v.8.2를 이용한 통계분석의 이해, 한울출판사, p.500.
5. 노형진(2003), SPSS/Amos에 의한 사회조사분석 - 범주형 데이터 분석 및 공분산구조분석 -, 형설출판사, pp.89~101.
6. 도로교통안전관리공단(2004), 2004년판 교통사고 통계 분석, p.4.
7. 도로교통안전관리공단(2004), 2004년판 대형 교통사고 사례 분석, p.16.
8. 도로교통안전관리공단(1989, 1991, 1992, 1997, 2000, 2001, 2002, 2003, 2004), 대형 교통사고 사례 분석.
9. 이주형·손동혁·윤문교(1990), 교통사고의 발생 특성과 그에 따른 사고 요인분석에 관한 연구, 대한국도계획학회지 국토계획, 제25권 제1호 통권 56호, pp.135~154.
10. 채서일·김종범(1988), SPSS/PC+를 이용한 통계분석, 법문사, p.112, p.115.
11. 최종성(2000), SPSS Ver 10을 이용한 현대통계분석, 북두출판사, pp.181~188, p.198, pp.341~354.
12. 최종후·한상태·강현철·김은석 공저(1998), Answer Tree를 이용한 데이터 마이닝 의사결정 나무분석, SPSS 아카데미, pp.42~48.
13. 허명희(2000), 사회과학을 위한 다변량 자료 분석, 자유아카데미, pp.117~118.
14. 허명희(1990), SAS 군집분석, 자유아카데미, pp.55~60.

15. 허명희(1992), 수량화 방법론의 이해, 자유아카데미, pp.2.1~2.18.
16. 田中 豊, 垂水共之 編(1996), Windows 版 统计解析ハンドブック - 多变量解析 -, 共立出版株式会社, pp.160~169.
17. SPSS(1998), Answer Tree 2.0 User's Guide, SPSS Inc., pp.56~57, pp.143~152, pp.189~190, pp.192~202.

♣ 주 작성자 : 김봉기

♣ 교신저자 : 고상선

♣ 논문투고일 : 2005. 9. 9

♣ 논문심사일 : 2005. 10. 12 (1차)

2005. 11. 11 (2차)

2005. 12. 13 (3차)

2005. 12. 27 (4차)

♣ 심사판정일 : 2005. 12. 27

♣ 반론접수기한 : 2006. 6. 30