



교통사고 위험그룹 및 사고유형별 심각도 결정 연구

- 서울시 중심 -

The Determination of Risk Group and Severity by Traffic Accidents Types - Focusing on Seoul City -

심 관 보*

Shim, Kywan Bho

Abstract

This research wished to risk type and examine closely driver special quality and relation of traffic accidents by occurrence type of traffic accidents and traffic accidents seriousness examine closely relation with Severity. Fractionate traffic accidents type by eight, and driver's special quality for risk group's classification did to distinction of sex, vehicle type, age etc. analyzed relation with injury degree adding belt used putting on availability for security the objectivity with wave. Used log-Linear model and Logit model for analysis of category data. A head-on collision and overtaking accident, right-turn accident are high injury or death accident and possibility to associate in relation with accident type and seriousness degree. In risk group analysis The age less than 20 years in motor-cycle driver, taxi driver in 41 years to 50 years old are very dangerous. The woman also was construed to the more risk group than man from when related to car, mini-bus, goods vehicle etc. Therefore, traffic safety education and Enforcement for risk group that way that can reduce accident that produce to reduce a loss of lives at traffic accidents appearance a head-on collision and overtaking accidents, right-turn accidents should be studied and as traffic accidents weakness class may have to be solidified.

Keywords : risk group, risk type, traffic accidents severity, log-linear model, logit-model

요 지

본 연구는 교통사고의 발생 유형과 교통사고 심각도(Severity)와의 관계를 규명함으로써 위험유형을 제시하고, 운전자 특성과 교통사고의 관계를 규명하고자 하였다. 교통사고 유형을 여덟 가지로 세분하고, 결과의 객관성 확보를 위해 안전벨트 착용여부를 추가하여 상해정도와의 관계를 분석하였으며, 위험그룹의 분류를 위한 운전자의 특성은 성별, 차종, 연령 등을 대상으로 하였다. 카테고리 자료의 분석을 위하여 로그-선형 모형 및 로짓 모형을 사용하였다. 분석결과 사고유형과 심각도와의 관계에서는 정면충돌 사고와 앞지르기시, 우회전시 사고가 부상 또는 사망사고에 연루될 가능성이 높았다. 위험그룹 분석에서는 20세 미만의 이륜차 운전자, 41세에서 50세까지의 택시 운전자가 가장 위험한 집단으로 분석되었으며 또한 남자 보다는 여자가 승용차와 중형화물 등에 관계되었을 때 더 위험한 것으로 분석되었다. 따라서 교통사고 발생시 인명피해를 줄이기 위해서는 정면충돌 사고와 앞지르기시, 우회전시 발생하는 사고를 줄일 수 있는 방안이 연구되어야 하고, 교통사고 취약계층으로 분석된 위험그룹에 대한 교통안전 교육 및 단속이 강화되어야 할 것이다.

핵심용어 : 위험 그룹, 위험 유형, 교통사고 심각도, 로그-선형 모형, 로짓 모형

* 정희원 · 도로교통공단 교통과학연구원 책임연구원(E-mail: Shimkb4@paran.com)



1. 서론

1.1 연구의 배경 및 목적

교통사고와 관련된 사망자와 부상자를 줄이기 위한 많은 노력에도 불구하고 우리 나라의 도로교통 사고는 여전히 수많은 인명을 앓아가고 있으며, 교통사고로 인한 손실비용은 연간 수조원에 달하고 있다.

선진국의 경우 교통사고의 부상 심각도(Severity)를 낮추고, 사망자수를 감소시키기 위해 구체적인 방법론에 대한 연구가 활발히 진행되고 있다. 그 주요 내용은 운전자 특성 및 위험그룹 분석 연구, 차량의 안전도 개선 및 탑승자 보호 연구, 사고유형과 부상 정도 관계 정립, 도로환경개선 및 안전시설과 도로구조물의 안전도 연구 등으로 대별된다.

Evans(1991)는 “운전자 특성과 위험스런 운전 행동은 연령, 성별과 관련이 있으며, 젊은 운전자들은 특별한 안전문제의 대상으로 인식되고 있다고 하였다”. 즉, 어떤 계층이 더 위험한가에 대한 관계를 이해함으로써, 효과적인 교통안전 프로그램을 설계할 수 있다고 하였다. Karl Kim(1995)은 하와이주의 사고유형과 부상정도간의 관계를 정립하는 로그-선형모형 개발을 통해 사고유형과 부상 카테고리간의 유용한 비교 수단을 제공하고 있으나, 우리나라의 경우 관련 분야 연구가 거의 없는 실정이다.

따라서 본 연구는 성별, 연령 등의 운전자 특성이 교통사고와 어떠한 관계를 가지며 어떠한 특성을 포함한 운전자 집단이 교통사고에 가장 많이 연관되는가를 분석하고자 한다. 또한 ‘정면충돌’, ‘추돌’ 등과 같은 사고발생 유형과 각 유형에 따른 교통사고 부상정도와의 관계 규명을 통하여 위험유형 및 유형별 위험등급을 제시하고자 한다. 이를 위하여 본 연구에서는 위험그룹의 분류를 위한 운전자의 특성을 연구자의 주관적인 판단과 자료 수집상의 문제 등으로 성별, 차량종별, 연령 등에 국한하였으며, 사고유형과 상해정도와의 관계를 분석하기 위해 차대차 사고의 유형을 여덟 가지로 세분하고 결과의 신뢰성 확

보를 위해 안전벨트 착용 여부를 추가하였다.

또한 본 연구의 교통사고 자료는 다차원의 범주형 변수들로 이루어져 있어 SAS를 통한 로그-선형 모형 기법을 이용하였으며, 반응변수에서 변수들간의 상호 영향을 효과적으로 비교하기 위해 로짓모형 기법을 추가하였다. 연구의 진행은 분석방법론 고찰, 위험그룹을 분류·결정하며, 사고유형별 사고심각도 모형결정 등의 순으로 이루어졌다.

2. 방법론 고찰

로그-선형모형 기법은 교통사고 데이터와 같은 카테고리변수들간에 관계를 조명하는데 유용한 방법이다. 지수화된 로짓 파라메타(Odds Multiplier)의 계산은 반응변수에서 변수조합의 영향을 비교하는데 매우 효과적이다. 각각의 모형을 살펴보면 다음과 같다.

2.1 로그-선형 모형(Log-Linear Models)

교통사고에 관한 많은 자료는 빈도수로 나타나고, 이러한 빈도수가 상호 교차·분류되어 분할표의 형태로 나타난다. 두 변수가 교차 분류되는 2차원 분할표에 대해 피어슨 X^2 검정 또는 X^2 검정과 우도비 G^2 검정을 통하여 두변수간에 독립성 및 동질성을 검정한다. 여러 개의 범주형 변수들로 이루어진 다차원 분할표의 경우 여러 이차원 주변합표 (Contingency marginal tables)를 만들어 X^2 또는 G^2 분석을 통하여 변수간의 관계를 분석할 때 다음과 같은 단점이 있다

- 한 쌍의 범주형 변수들간의 주변 관련성과 다른 변수가 새로 추가되었을 때의 관련성을 복합적으로 설명할 수 없다.
- 여러 쌍에 대한 변수들의 관련성을 동시에 측정할 수 없다.
- 변수들의 3차나 그 이상 차수의 교호작용 가능

성이 무시된다.

이상의 단점을 보완하고 다차원 상황에서 교차 분류된 자료의 통계적 분석을 위해 로그-선형 모형을 이용하는 것이 바람직하다.

3차원 이상의 다차원 분할표는 복잡한 구조를 가질 수 있기 때문에 구조적 특성을 탐색하기 위해서 X^2 검정 이상의 체계적인 통계적 방법이 필요하다. 로그-선형 모형이 바로 그런 통계적 분석을 위한 모형이라고 할 수 있다.

세범주형 변수(분류변수) V_1, V_2, V_3 에 의하여 분할된 돛수표에서 Y_{ijk} 를 칸(i, j, k)에서의 관측돛수 Y_{ijk} 가 평균 m_{ijk} 를 갖는 포아송 확률변수라고 하고, 적합돛수에 로그(logarithm)를 씌우면 세변수 수준의 선형결합으로 나타내어진다. 일반적으로 $\log m_{ijk}$ 는 식 (1)과 같다.

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} + u_{123(ijk)} \quad (1)$$

여기서 $u_{1(i)} + u_{2(j)} + u_{3(k)}$ 는 각 분류변수 1, 2, 3의 주효과(main effect)를 나타내며, 제약식 $\sum_i u_{1(i)} = 0, \sum_j u_{2(j)} = 0, \sum_k u_{3(k)} = 0$ 을 만족하도록 정해진다.

$u_{12(ij)} + u_{13(ik)} + u_{23(jk)}$ 는 두 분류변수의 교호작용이며, 분류변수 V_1, V_2, V_3 범주의 수를 각각 I, J, K 라고 할 때 m_{ijk} 의 수는 IJK 이고, 선형 독립적인 모수의 수도 IJK 이기 때문에 이러한 모형을 무제한 모형(unrestricted model)이라고 할 수 있다. 실제로는 이보다 간단한 모형 즉 제한 모형(restricted model)이라야 모형으로서의 가치가 있다.

예를 들면 식 (2)는 3-인자 교호작용이 없는 모형(no three-factor interaction model)이다.

$$\log m_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} \quad (2)$$

여러 형태의 모형에서 우리는 로그-선형 모형의 기

대모수값을 추정할 수 있다. 그리고 추정된 기대값을 바탕으로 어떤 로그-선형 모형의 적합성을 검정하게 된다.

2.2 로짓 모형(Logit-Models)

로그-선형 모형에서 특정한 변수의 비율, 특히 다른 범주형 변수들은 '설계' 변수이고 두 개의 값을 가진 '반응' 변수(설명변수)에 대한 범주형 변수의 효과를 평가하는 모형을 개발하고자 할 때 반응변수의 비율에 대해 연구하기 위하여 자료에 적합한 모형을 찾는다. 로그-선형 모형식(3)에서

$$\log(m_{ijk}) = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)} \quad (3)$$

비율 m_{ijk} / m_{ojk} 에 대한 로그-선형 모형식은 다음과 같이 정의된다.

$$\begin{aligned} \log(m_{ijk} / m_{ojk}) &= [u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)}] - [u + u_{1(o)} + u_{2(j)} + u_{3(k)} + u_{12(oj)} + u_{13(ok)} + u_{23(jk)}] \\ &= [u_{1(i)} - u_{1(o)}] + [u_{12(ij)} - u_{12(ok)}] + [u_{13(ik)} - u_{13(ok)}] \\ &= W_1 + W_{2(j)} + W_{3(k)} \end{aligned}$$

위 모형식의 W 항들도 통상적 분산분석과 유사한 제약조건을 갖는다. 이러한 모형을 로짓 모형이라 한다. 여기서 W 는 추정 파라메타 값이고, 로그-선형 모형에서 추정된 u 로부터 계산된다. 이 로짓 모형의 양변을 지수화 하면 식 (4)와 같은 가법모형이 된다.

$$(m_{ijk} / m_{ojk}) = \exp[W_1] \exp[W_{2(j)}] \exp[W_{3(k)}] \quad (4)$$

이상에서 살펴본 바와같이 로그-선형 분석은 교통 사고 데이터를 분석하는데 좋은 방법이며, 특히 카테



고리 변수들간에 관계를 분석하는데 매우 유용한 방법이다.

로지스틱 회귀분석(Logistic Regression)과 같은 다른 기법이 추가로 적용될 수 있으며 Odd Multipliers (Exponentiated Logit Parameters)의 계산은 '+' 형태의 Model을 'x' 형태로 변형시켜 반응변수에서 변수의 영향을 비교하는데 있어 더욱 효과적이다.

3. 교통사고 위험그룹 결정

2007년 서울시에서 발생한 교통사고 1,599건을 대상으로 교통사고를 발생시키기 쉬운 특정 운전자 집단을 분석하였다. 성(性), 차종 그리고 연령층에 따른 교통사고 발생 현황은 다음의 표 1과 같다.

교통사고를 야기한 운전자의 86.6%는 남성이고 13.4%가 여성인 것으로 나타났으며, 차종별로는 승용차가 66.7%, 택시 10.5%, 소형화물 5.4%, 소형버스 5.1%, 버스 5.0%, 이륜차 4.3%, 대형화물

표 1. 교통사고 발생현황(2007, 서울시)

	차종	빈도	백분율	누적빈도	누적백분율
성별	남자	1,385	86.6	1,385	86.6
	여자	214	13.4	1,599	100.0
차종	승용차	1,066	66.7	1,066	66.7
	택시	168	10.5	1,234	77.2
	소형버스	82	5.1	1,316	82.3
	버스	80	5.0	1,396	87.3
	소형화물	86	5.4	1,482	92.7
	중형화물	15	0.9	1,497	93.6
	대형화물	34	2.1	1,531	95.7
	이륜차	68	4.3	1,599	100.0
연령	20세미만	60	3.8	60	3.8
	21~30세	545	34.1	605	37.8
	31~40세	508	31.8	1,113	69.9
	41~50세	342	21.4	1,455	91.0
	51~60세	144	9.0	1,599	100.0

주) 운전자 연령으로 61세 이상은 부적합으로 판단하여 분석집단에서 고려하지 않았다.

2.1% 등이었다. 한편, 연령층에 따른 교통사고는 21세에서 30세까지 34.1%(545건)가 가장 많은 교통사고를 일으키는 것으로 조사되었고, 다음이 31세에서 40세까지 31.8%(508건)로 나타났다.

성별은 남녀로, 차종유형은 승용차, 택시, 소형버스, 버스등 8개 카테고리, 연령은 5개 카테고리 그룹화 하였으며, 여러 모델 형태를 테스트하여 3변수간의 최적모델(best fitting model)을 선정하였다. 로그-선형 모형의 최적화 모형을 선정하기 위한 각 모형의 자유도(DF) 및 우도비 검정통계량 값(G^2), 그리고 대응하는 유의수준(P)을 정리하면 표 2와 같다.

표 2. 위험그룹 분석을 위한 모형의 비교

모형	DF	G^2	Prob.
① $u + u_1 + u_2 + u_3$	67	367.86	0.0000***)
② $u + u_1 + u_2 + u_3 + u_{23}$	39	184.99	0.0000*
③ $u + u_1 + u_2 + u_3 + u_{13}$	63	344.95	0.0000*
④ $u + u_1 + u_2 + u_3 + u_{21}$	60	224.80	0.0000*
⑤ $u + u_1 + u_2 + u_3 + u_{12} + u_{23}$	32	41.93	0.1125
⑥ $u + u_1 + u_2 + u_3 + u_{13} + u_{23}$	35	162.08	0.0000*
⑦ $u + u_1 + u_2 + u_3 + u_{12} + u_{13}$	56	201.88	0.0000*
⑧ $u + u_1 + u_2 + u_3 + u_{12} + u_{13} + u_{23}$	28	19.60	0.8786

주) *로 나타난 모형은 대응되는 자유도의 χ^2 분포와 비교하여 유의수준 5(%)에서 귀무가설 모형식을 기각한다는 것을 의미한다.

참고) u_1 : 성별, u_2 : 차종, u_3 : 연령, u_{12} : 성별과 차종의 조합,

u_{13} : 성별과 연령 조합, u_{23} : 차종과 연령 조합

표 2로부터 적합한 모형은 ⑤번과 ⑧번 모형이다. 특히, ⑧번 모형은 ⑤번 모형에 변수 1과 변수 3의 교호작용항이 첨가된 모형으로 자료를 더욱 잘 적합하고 있다. 즉 ⑧번 모형의 G^2 값이 ⑤번 모형의 통계량 값 보다 20 이상 감소되었다. 따라서 세변수의 일차 교호작용이 있고, 이차 교호작용항은 없는 ⑧번 모형이 자료를 가장 잘 설명하고 있다고 할 수 있다. 따라서 ⑧번 모형이 최적의 로그-선형 모형식으로 결정되었다(식 5).

χ^2 혹은 G^2 , "Log-likelihood ratio", "Goodness-of-fit statistic"이 모형의 적합성 여부



를 결정하기 위해 사용되었으며, 모형은 변수 S (성별), V (차종), a (연령)를 갖는 세변수 모형이고, 최적의 모형은 3개의 주효과와 3개의 양변수간 교호작용을 포함한다.

$$\log(m_{ijk}) = u + u_{s(i)} + u_{v(j)} + u_{a(k)} + u_{sv(ij)} + u_{va(ik)} + u_{sa(jk)} \quad (5)$$

여기서 m_{ijk} 는 기대되는 셀의 빈도이고 셀 빈도의 총괄평균은 “ u ”이다. $u_{a(k)}$ 는 카테고리화된 연령그룹(20세 미만, 21세에서 30세 등)에 대한 연령효과(age effects)이다. 또한 $u_{sv(ij)}$ 는 성별과 차종간의

표 3. 로그-선형 모형(성별, 차종별, 연령별)의 파라메타 예측

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob.
성 별	1	2.6518	0.7897	11.28	0.0008
차 종	2	4.4517	0.8296	28.79	0.0000
	3	0.2194	1.9277	0.01	0.9094
	4	1.1042	0.8747	1.90	0.2068
	5	-1.2169	2.8484	0.18	0.6692
	6	-1.7362	2.8581	0.37	0.5435
	7	0.1607	0.9090	0.03	0.8597
	8	-1.7389	2.8576	0.37	0.5428
	연 령	9	-2.1012	1.0415	4.07
10		0.9164	0.2808	10.65	0.0011
11		0.9901	0.2820	12.33	0.0004
12		0.6962	0.2894	5.79	0.0161
성별 × 차종	13	-1.8884	0.7889	5.73	0.0167
	14	-0.5328	0.8276	0.41	0.5197
	15	-1.0013	0.8283	1.46	0.2267
	16	1.5046	2.8354	0.28	0.5957
	17	1.4433	2.8371	0.26	0.6109
	18	-1.6956	0.8547	3.94	0.0473
	19	0.9749	2.8407	0.12	0.7315
성별 × 연령	20	0.2933	0.2453	1.43	0.2319
	21	0.1789	0.0908	3.88	0.0488
	22	-0.0886	0.0864	1.05	0.3056
	23	-0.2864	0.0894	10.27	0.0013
	24	0.1812	1.0252	0.03	0.8597

(표 계속)

Effect	Parameter	Estimate	Standard Error	Chi-Square	Prob.
차종 × 연령	25	-0.1236	0.2737	0.20	0.6514
	26	-0.1236	0.2769	0.20	0.6554
	27	-0.1148	0.2858	0.16	0.6877
	28	-4.7369	6.9638	0.46	0.4964
	29	0.3643	1.7484	0.04	0.8350
	30	1.4671	1.7462	0.71	0.4008
	31	1.4833	1.484	0.72	0.3962
	32	0.1352	1.1310	0.01	0.9048
	33	0.1689	0.3275	0.27	0.6059
	34	-0.2017	0.3425	0.35	0.5560
	35	-0.2558	0.3683	0.48	0.4873
	36	0.3752	1.0945	0.12	0.7318
	37	-1.0630	0.3512	9.16	0.0025
	38	-0.1388	0.3271	0.18	0.6714
	39	0.2759	0.3353	0.68	0.4105
	40	-0.1396	1.2419	0.01	0.9105
	41	0.7365	0.3812	3.73	0.0534
	42	0.2814	0.3953	0.51	0.4765
	43	0.4708	0.4089	1.33	0.2496
	44	1.0236	1.2503	0.67	0.4129
	45	-0.1113	0.4792	0.05	0.8163
	46	-0.3880	0.5178	0.56	0.4536
	47	-0.2525	0.5583	0.20	0.6511
	48	0.3313	1.2382	0.07	0.7890
	49	-0.0117	0.4007	0.00	0.9767
	50	0.2557	0.3995	0.41	0.5221
	51	-0.7920	0.5304	2.23	0.1354

상호작용 효과를 나타낸다. 모형식에 따른 모수 추정 결과는 표 3과 같다.

성별과 차종, 성별과 연령층, 차종과 연령층에 대한 모수 추정치 u 값은 표 4~표 6과 같다. 성별과 차종의 조합에서는 여성이 승용차, 소형버스, 중형화물 등에 관련 되었을 때 남성보다 교통사고 발생 가

표 4. 성별과 차종에 대한 모수추정치 (u)

구분	승용	택시	소형 버스	버스	소형 화물	중형 화물	대형 화물	이륜차
남	-1.8884	-0.5328	-1.0013	1.5046	1.4433	-1.6956	0.9749	1.1953
여	1.8884	0.5328	1.0013	-1.5046	-1.4433	1.6956	-0.9749	-1.1953



능성이 높은 것으로 나타났다. 또한 가장 위험한 경우는 '여성과 승용차'의 관계였다.

한편, 성별과 연령층의 조합에서는 '20세 미만의 남자'가 u 값 0.2933으로 가장 높아 위험한 집단으로 분류되었고, 다음은 '41세에서 50까지의 여자' (0.2864), '21세에서 30까지의 남자' (0.1789)의 순이었다.

표 5. 성별과 연령층에 대한 모수추정치 (u)

구분	20세 미만	21~30세	31~40세	41~50세	51~60세
남	0.2933	0.1789	-0.0886	-0.2864	-0.0972
여	-0.2933	-0.1789	0.0886	0.2864	0.0972

차종에 따른 위험연령층은 '20세 미만의 이륜차 운전자'가 2.830으로 가장 높고, '41세~50세, 31~40세, 51세~60세'까지의 택시운전자'가 1.483, 1.4671, 1.422로 가장 위험한 집단인 것으로 분석되었으며, 그 다음이 '20세 미만의 중형화물' 운전자 순으로 나타났다. 특히 택시의 경우 '20세 미만'인 경우를 제외하고 각 연령층에서 위험이 높은 것으로 나타났다.

표 6. 차종과 연령층에 따른 상호작용효과 모수추정치 (u)

구분	20세 미만	21~30세	31~40세	41~50세	51~60세
승용	0.1812	-0.1236	-0.1236	-0.1148	0.1808
택시	-4.7369	0.3643	1.4671	1.4833	1.4222
소형버스	0.1352	0.1689	-0.2017	-0.2558	0.1534
버스	0.3752	-1.0630	-0.1388	0.2759	0.5507
소형화물	-0.1396	0.7365	0.2814	0.4708	-1.3491
중형화물	1.0236	-0.1113	-0.3880	-0.2525	-0.2718
대형화물	0.3313	-0.0117	0.2557	-0.7920	0.2167
이륜차	2.8300	0.0399	-1.1521	-0.8149	0.9029

4. 교통사고 유형별 심각도 결정 모형적합

교통사고의 유형에 따른 사고심각도를 결정하기 위하여 차량단독 사고와 차대사람 사고를 제외하고 2007년 우리 나라에서 발생한 차대차 사고 중 안전

벨트의 착용 여부가 조사된 68,366건의 교통사고를 대상으로 하였다. 표 7은 안전벨트의 착용 여부, 사고유형, 심각도 등에 따라 분류한 빈도표이다. 사고유형은 '정면충돌', '추돌', '직각충돌' 등의 여덟 가지로 구분하였다. 전체 사고의 69.6%인 47,611건은 안전벨트를 착용한 상태에서 교통사고가 발생하였고, 30.4%는 미착용 상태에서 발생한 것으로 나타났다.

사고유형별로는 추돌사고가 48.3%로 가장 많은 사고유형이었고, 교행시 발생한 사고와 직각충돌사고 등이 비교적 많이 발생하였다. 또한 93.9%의 사고는 단순물피(Property Damage Only)이고 약 6%가 경상 이상의 사고인 것으로 조사되었다.

표 7의 카테고리 데이터 분석을 위해서 로그-선형 모형과 로짓 모형을 이용하였다. 사고심각도와 다른 변수간의 관계를 조사한 결과, 사고유형은 사고심각도를 결정하는 중요한 인자였다. 또한 사고유형의 카테고리를 유지하면서 차별성을 갖는 사고심각도를 분석하기 위해 변수간의 주효과와 교호작용을 통합하는 형태로 모형을 개발하였다.

표 7. 안전벨트·사고유형·심각도별 빈도표

구분		빈도	백분율	누적빈도	누적백분율
안전벨트	착용	47,611	69.6	47,611	69.6
	미착용	20,755	30.4	68,366	100.0
사고유형	정면충돌	4,828	7.1	4,828	7.1
	추돌	33,043	48.3	37,871	55.4
	직각충돌*	7,322	10.7	45,193	66.1
	앞지르기시	1,518	2.2	46,711	68.3
	진로변경시	5,361	7.8	52,072	76.2
	교행시	9,301	13.6	61,373	89.8
	좌회전시	5,001	7.3	66,374	97.1
	우회전시	1,992	2.9	68,366	100.0
사고심각도	상해없음	64,173	93.9	64,173	93.9
	경상	2,016	2.9	66,189	96.8
	중상	1,666	2.4	67,855	99.3
	사망	511	0.7	68,366	100.0

* 직각충돌: 비신호교차로에서 통행우선순위 미숙에 따른 사고



세변수의 범주형자료에 적합한 로그-선형 모형을 찾아 그들 상호간의 관련성 및 교호작용을 설명하기 위해 SAS통계 자료를 적합시킨 결과는 표 8과 같다.

모형의 적합성 여부를 결정하기 위해 X^2 혹은 G^2 , 자유도(DF) 등이 사용되었다. 위 표로부터 로그-선형 모형을 설명하기 위한 최적화 모형은 ⑦번 모형으로 나타났다. ⑦번 모형의 G^2 은 18.10 (Log-likelihood ratio estimate), DF=14, Prob= 0.2025를 보이고 있다.

모형은 변수 b (안전벨트), a (사고유형), s (사고피해 정도)를 갖으며, 대표적으로 최적의 모형은 3개의 주효과와 3개의 양변수간 교호작용을 포함하는 식 (6)과 같다.

$$\log(m_{ijk}) = u + u_{b(i)} + u_{a(j)} + u_{s(k)} + u_{ba(ij)} + u_{bs(ik)} + u_{as} \quad (6)$$

여기서, m_{ijk} : 기대된 셀 빈도
 u : 평가 파라메타

표 8. 로그-선형 모형 분석결과

Source	자유도	Chi-Square	Prob.
① 안전벨트(b)	1	351.61	0.0000
② 사고유형(a)	7	1335.56	0.0000
③ 사고피해정도(s)	2	300.70	0.0000
④ 안전벨트×사고유형(ba)	7	65.17	0.0000
⑤ 안전벨트×사고피해정도(bs)	7	23.10	0.0000
⑥ 사고유형×사고피해정도(as)	2	250.09	0.0000
⑦ Likelihood Ratio	14	18.10	0.2025

표 9에서 셀 빈도 전체의 총괄 평균은 “ u ”이다. 기술된 u 파라메타 각각은 그 효과에 기인한 총괄 평균으로부터의 편차이다. 예를 들어 $u_{a(j)}$ 는 각 사고유형 카테고리를 위한 독립된 파라메타 평가치를 갖는 사고유형 효과(effect)이다.

최적의 로그-선형 모형으로부터 파라메타 평가치 u 가 얻어졌으며, 파라메타의 예측치와 표준에러값은 표 9와 같다.

표 9. 사고유형, 안전벨트사용, 사고심각도의 로그-선형 모형 Parameter estimates

(Model : $[sa][sb][ba]$, $G^2=18.10$, 자유도=14, $p=0.2025$)

구분		상해 없음	경 상	중 상	사 망	
사 고 심각도	u	3.1262	-0.4229	-0.7321	1.9712	
	$s. e$	(0.0257)	(0.0360)	(0.3405)	(0.0771)	
안전벨트 착용	u	-0.1400	0.1715	0.0873	-0.1188	
	$s. e$	(0.0165)	(0.0257)	(0.0265)	(0.0495)	
미착용	u	0.1400	-0.1715	-0.0873	0.1188	
	$s. e$	(0.0165)	(0.0257)	(0.0265)	(0.0495)	
사 고 유 형	정면 충돌	u	-0.9663	-0.4701	0.3182	1.1182
		$s. e$	(0.0356)	(0.0573)	(0.0554)	(0.1068)
	추돌	u	0.0599	-0.1418	0.0577	0.0242
		$s. e$	(0.0312)	(0.0432)	(0.0479)	(0.0939)
	전측면	u	0.1344	0.2665	0.1187	-0.5196
		$s. e$	(0.0575)	(0.0726)	(0.0812)	(0.1725)
	앞지르기	u	-0.2743	-0.1415	0.1341	0.2797
		$s. e$	(0.0760)	(0.1126)	(0.1170)	(0.2295)
	진로 변경	u	0.3862	0.0901	-0.2922	-0.1841
		$s. e$	(0.0660)	(0.0867)	(0.1048)	(0.1980)
	교행	u	0.0336	-0.0191	0.0683	-0.0828
		$s. e$	(0.0450)	(0.0623)	(0.0682)	(0.1350)
좌회전	u	0.2614	0.2436	-0.0279	-0.4771	
	$s. e$	(0.0697)	(0.0870)	(0.0999)	(0.2091)	
우회전	u	-0.3631	-0.1723	0.3769	-0.1585	

주효과(main effect)의 행(raw)은 부상정도의 상 대분포를 나타낸다. 사고유형과 안전벨트의 사용효 과도 분석되었으며, 주효과의 대부분은 0.05수준에 서 통계적으로 유의하였다.

예측치 u 는 효과의 방향과 크기의 기준을 제공한 다. 즉 양의 값을 나타내는 관계는 가능성이 높음을 나타내고, 음의 값은 낮은 가능성을 나타낸다.

안전벨트를 착용하지 않았을 때 사망사고로 이어 질 가능성이 높은 것으로 분석되었다. 또한 사망자는 정면충돌과 앞지르기에 u 값이 각각 1.1182, 0.2797로 사망 가능성이 가장 높은 것으로 분석되었 다. 중상의 경우 우회전시(0.3769)와 정면충돌시 (0.3182)에 발생할 가능성이 높았다. 한편 상해가 없는 물피사고는 진로변경시와 좌회전시 교통사고와



관계 있는 것으로 나타났다. 이상의 관계를 보다 정확히 정립하기 위해 로짓 모형을 이용하였다. 로그-선형 모형으로부터 상해 카테고리 “O”를 이용하여 사고심각도에 대해 설정한 로짓 모형은 다음과 같다.

$$\log(m_{ijk}) = u + u_{b(i)} + u_{a(j)} + u_{s(k)} + u_{ba(ij)} + u_{bs(ik)} + u_{as(jk)}$$

$$\begin{aligned} \log(m_{ijk} / m_{ojk}) &= [u + u_{b(i)} + u_{a(j)} + u_{s(k)} + u_{ba(ij)} \\ &\quad + u_{bs(ik)} + u_{as(jk)}] - [u + u_{b(o)} + u_{a(j)} \\ &\quad + u_{s(k)} + u_{ba(oj)} + u_{bs(ok)} + u_{as(jk)}] \\ &= [u_{b(i)} - u_{b(o)}] + [u_{ba(ij)} - u_{ba(oj)}] \\ &\quad + [u_{bs(ik)} - u_{bs(ok)}] \\ &= W_i + W_{a(j)} + W_{s(k)} \end{aligned}$$

여기서 W 는 로그-선형 모형에서 추정된 모수 u 로부터 계산된다. 안전벨트의 사용, 교통사고 유형, 사고피해정도 등 세변수 모형의 변수간 상호작용과 주효과를 통합하기 위해 ‘Additive model’을 사용하였다. 표 10은 사고심각도에 대한 안전벨트 착용 여부, 사고유형의 “Odds Multipliers”를 나타낸 것이다.

표 10. 사고 심각도에 대한 안전벨트 착용여부, 사고유형의 “Odds Multipliers” 값

		상해없음	경상	중상	사망
Baseline Odds		1	0.0287	0.0211	0.3150
Odds Multipliers					
안전 벨트	미착용	1	1.3654	1.2552	0.9790
	착용	1	0.7094	0.7966	0.7719
사 고 유형	정면충돌	1	1.6424	3.6128	8.0400
	추돌	1	0.9418	0.9978	0.9649
	직각충돌	1	1.1410	0.9844	0.5199
	앞지르기	1	0.8756	1.5044	1.7401
	진로변경	1	0.7437	0.5074	0.5653
	교행	1	0.9486	1.0353	0.8901
	좌회전	1	0.9823	0.7487	0.4778
	우회전	1	1.2101	2.0959	1.2270

아래 식의 첫 번째 인자는 사고심각도의 “Baseline odds”를 나타내고, 나머지 인자인 $\exp[W_{a(j)}]$ 와 $\exp[W_{s(k)}]$ 는 사고유형과 사고심각도에 대한 “Odds multiplier”를 나타낸다.

$$\ln(m_{ijk} / m_{ojk}) = \exp[W] \exp[W_{a(j)}] \exp[W_{s(k)}] \quad (7)$$

사고유형과 상해정도의 관계를 보면, 사망 가능성의 경우 정면충돌시에 8.04로 가장 높았고, 다음이 앞지르기 1.74 순이었으며, 추돌과 교행사고는 사망 가능성이 낮았다. 중상 가능성은 정면충돌이 3.6, 우회전이 2.1 등으로 높았으나 진로변경 사고시에는 낮았다. 경상가능성은 직각충돌이 1.14로 높게 나타났다. 또한 우회전시 사고는 중상 가능성이 높았으며, 좌회전은 경상 가능성이 가장 높은 것으로 나타났다. 안전벨트의 착용 여부와 사고심각도의 관계는 미착용시에 부상 또는 사망할 가능성이 훨씬 높았다. 즉 안전벨트의 착용은 교통사고로 인한 부상을 경감할 수 있다는 사실을 설명하고 있는 것이다.

5. 결론

본 연구는 운전자 특성이 교통사고와 어떠한 관계를 가지며, 어떠한 특성을 포함한 운전자 집단이 교통사고와 가장 많이 연루되는 가를 분석하고자 하였으며, 또한 교통사고의 발생 유형과 각 유형에 따른 교통사고 심각도와의 관계 규명을 통하여 위험유형을 제시하고자 하였다. 우선 위험그룹 분류를 위한 운전자 특성으로 성별, 차종, 연령 등을 고려하였고, 교통사고 유형과 사고심각도와의 관계 규명을 위해 사고유형을 여덟 가지로 세분하고 결과의 신뢰성 확보를 위해 안전벨트 착용 여부를 추가하여 상해정도와의 관계를 분석하였다. 위험그룹 분석결과는 성별과 차종의 조합에서는 여성이 승용차, 택시, 이륜차 등에 관계되었을 때 남성보다 교통사고를 야기할 가



능성이 높은 것으로 나타났으며, 성별과 연령층의 조합에서는 '20세 미만의 남성'이 가장 높아 위험한 집단으로 분류되었다. 차종에 따른 위험연령층은 '20세 미만의 이륜차 운전자'와 '41세에서 50세까지의 택시운전자'가 가장 위험한 집단인 것으로 조사되었다. 특히 택시의 경우 '20세 미만'인 경우를 제외하고 각 연령층에서 위험이 높은 것으로 나타났다.

한편, 사고유형과 심각도와와의 관계는 '정면충돌' 사고가 사망으로 이어질 가능성이 가장 높은 것으로 분석되었고, '우회전시' 사고는 중상사고와 경상사고로 연계될 가능성이 높은 것으로 나타났다. 안전벨트의 착용 여부와 사고심각도의 관계는 미착용시에 부상 또는 사망할 가능성이 훨씬 높았다. 즉 안전벨트의 착용은 교통사고로 인한 부상을 경감할 수 있다는 사실을 설명하고 있는 것이다.

따라서 교통사고 발생시 인명피해를 줄이기 위해서는 '정면충돌' 사고와 '우회전시' 발생하는 사고를 줄일 수 있는 방안이 연구되어야 하고, 안전벨트 착용에 대한 단속이 강화되어야 한다. 또한 교통사고 취약계층으로 분석된 '여자', '20세 미만의 이륜차 운전자', '택시 운전자' 등에 대한 교통안전 교육 및 단속도 강화되어야 할 것으로 판단된다.

교통사고 위험그룹 분석에 보다 신뢰도 높은 연구 결과를 얻기 위해서는 성별, 차종, 연령 등과 같은 기초적인 자료 외에도 운전경력, 사고이력 등 운전자의 특성을 차별화시킬 수 있는 데이터가 고려되어야 할 것이다. 아울러 사고유형별 심각도 분석의 경우에도 전국단위의 데이터를 갖고 분석을 한다면 더욱 좋은 결과를 얻을 것으로 생각되어 향후 연구과제로 제안한다.

참고 문헌

경찰청, "2008년판 교통사고 통계", 2008
 도로교통공단, "2008년판 교통사고 통계분석", 2008
 유근영, 범주형 자료분석론, 1996
 허명희, 이태림, "통계적 자료분석 I", 1994
 허명희, "SAS 범주형 데이터 분석", 1989
 홍종선, "SAS 와 통계자료분석", 1996
 Evans, L. "Traffic Safety and Driver", Van Nostrand Reinhold, New York. 1991.
 Evans, L. P. Wasielewski. Risky Driving Behavior "Related to Driver and Vehicle Characteristics". Accident Analysis and Prevention. 1983.
 Feinberg, S. "The Analysis of Cross-Classified Categorical Data", The MIT Press, Cambridge, Mass, 1980.
 Karl Kim, "Driver's at Fault: Influence of Age, Sex, and Vehicle Type", 1995
 Karl Kim, Lawrence Nitz, James Richardson, and Lei Li, "Analyzing the Relationship Between Crash Types and Injuries in Motor Vehicle Collisions in Hawaii".
 Williams, A. F. and R. S. Karf, "Teenaged Drivers and Fatal Crash Responsibility", Law Quarterly, 1984

접 수 일: 2009. 3. 11

심 사 일: 2009. 3. 13

심사완료일: 2009. 5. 6