

고속도로 인터체인지 연결로에서의 교통사고 예측모형 개발

Development of Accident Prediction Models on Freeway Interchange Ramps
Analyses of Capacity and Left-Turn Adjustment Factors Permitted

박 효 신

손 봉 수

김 형 진

(연세대학교 도시공학과, 석사과정) (연세대학교 도시공학과, 교수) (연세대학교 도시공학과, 교수)

목 차

I. 서론

II. 문헌 및 선행연구고찰

III. 모형의 설정

IV. 교통사고 분석

V. 모형의 추정

VI. 결론 및 향후연구과제

참고문헌

I. 서론

우리나라에서는 설정된 기준 이상의 교통사고가 발생한 지점을 사고다발지점으로 설정하는데 이는 도로의 각 지점·구간에 내포된 고유의 사고위험성과 교통 및 도로특성의 변화로 생기는 교통사고의 불확실성을 간과하였다는 점에서 적용상 한계를 가질 수 있다. 변화하는 교통사고를 설명하고 교통사고 요인들에 대한 분석을 위해 사고모형개발이 선행되어야 한다.

고속도로 교통사고에 영향을 주는 주요 인자는 도로 기하구조요인, 교통 환경요인, 운전자 및 차량요인으로 구분할 수 있고 이들 상호간 복합적인 관계에 의해서 교통사고가 발생할 수 있다. 이 중 도로의 기하구조는 보편적인 운전자가 보통 수준의 주의력과 긴장감을 가지고 운전하더라도 사고의 위험성이 없도록 설계되어야 한다. 우리나라는 건설교통부에서 규정하는 지침에 의거하여 도로를 설계하는데 안전한 도로의 건설을 위해서는 사고와 연관된 기하구조요인의 정확한 이해와 분석을 통해 도로의 설계자가 도로 설계과정에 반영하는 것이 필요하다. 교통 환경요인 또한 기존연구자들의 분석에서 교통사고를 설명하는데 중요한 인자라는 결론을 얻었으며 실제적인 제어 또한 가능하다.

기존 사고예측모형개발은 주로 고속도로 본선에 대하여 수행되어져왔다. 한국도로공사관내 고속도로에서 발생한 총 2880건의 교통사고건수(2005년)중 분기점/인터체인지 연결로와 관련

하여 발생한 교통사고건수는 총 245건(8.5%)을 차지했다. 총 연장에 비해 사고건수의 구성비가 높은 연결로에서의 사고분석이 필요하다.

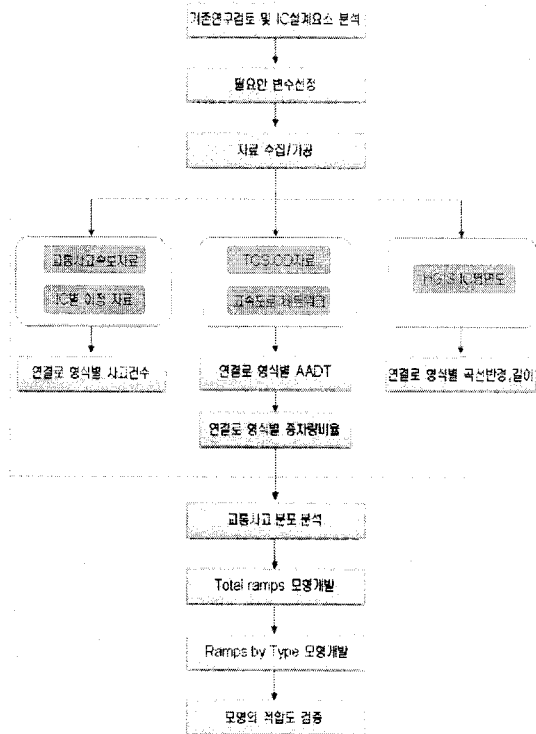
교통사고건수의 경우 과분산(overdispersion)의 특징을 가지고 있으므로 본 연구에서는 음이항 회귀모형(Negative binomial regression model)을 개발한다. 추정된 모형의 통계적 유의한 변수들을 파악을 통해 연결로의 도로 설계 및 운영에 적용하거나 해당도로의 교통특성을 반영한 교통사고 절감 대책등에 이용할 수 있다.

II. 문헌 및 선행연구고찰

1. 연구의 수행과정

본 연구에서는 교통사고에 가장 밀접하게 영향을 미칠 것이라고 판단되는 교통량이라는 교통요인을 고려하기위해 고속도로 인터체인지를 분석대상으로 하였다. 또한 한국도로공사 관내 폐쇄식구간의 인터체인지 중 85.1%를 차지하고 있는 트럼펫 인터체인지의 연결로를 대상으로 하였다. 2003~2005년 교통사고 속보자료 분석결과 한국도로공사관내 폐쇄식구간 트럼펫 인터체인지 총 188개소 중 122개소에서 총 326건의 교통사고가 발생했다. 이 중 모형 구축에 필요한 데이터의 수집이 가능한 총 31개소 인터체인지를 최종분석대상으로 하여 총 146건의 교통사고에 대해 분석하였다. 연결로의 도로 기하구조자료, 교통량자료, 사고이력자료는 한국도로공사 HGIS의 인터체인지 평면도, 전국 고속

도로 네트워크, TSC OD 교통량, 교통사고속보 자료를 이용하였다. 고속도로 인터체인지 설계와 관련한 요소는 입체교차로 설계지침(건설교통부 2005), 도로의 구조·시설 기준에 관한 규칙 해설 및 지침(건설교통부 2000)등을 참고하였다. 수행과정을 도식화하면 <그림 1>과 같다.



<그림 1> 연구 수행과정도

2. 트럼펫 인터체인지의 설계 기준

트럼펫 인터체인지의 설계기준은 본선과 연결로의 도로 기하구조를 중심으로 규정되고 있다.

1) 곡선반경

평면 곡선부를 주행하는 운전자 안전 확보를 위해 직선부와 같은 주행의 연속성을 갖도록 할 필요가 있다. 평면 곡선부주행시 발생하는 원심력으로 인하여 곡선부의 바깥쪽으로 미끄러지거나 전도할 위험을 방지할 수 있도록 타이어와 포장면사이의 횡 방향 마찰력이 원심력보다 크도록 하여야 하며 동시에 주행의 쾌적성을 확보할 수 있어야 한다. 규정 값에 얽매이지 않고 그 구간 앞뒤의 조건과 균형을 고려하여 지형조건에 순응할 수 있는 값을 적용하여야 한다.

<표 1> 최소평면곡선반경의 값

설계속도(km/h)	최소평면곡선반경(m)		
	최대편경사 6%	최대편경사 7%	최대편경사 8%
60	140	135	130
50	90	85	80
40	60	55	50

2) 곡선길이

평면 곡선부주행시 곡선의 길이가 짧으면 운전자는 곡선방향으로 핸들을 조작하였다가 직선부로 진입하기 위해 즉시 핸들을 반대방향으로 조작하여야 하기 때문에 고속 주행시 안전에 좋지 않은 영향을 줄 수 있다.

<표 2> 최소평면곡선길이의 값

설계속도(km/h)	최소평면곡선길이(m)	
	도로의 교각이 5도미만인경우	도로의 교각이 5도이상인경우
60	$350 / \theta$	70
50	$300 / \theta$	60
40	$250 / \theta$	50

주: * θ 는 도로 교각의 값(도)이며 2도 미만인 경우에는 2도로 한다.

3. 기존연구고찰

1) 고속도로 교통사고 기존연구

고속도로 교통사고에 영향을 주는 주요 인자는 크게 도로 기하구조 요인, 교통 환경요인, 운전자 및 차량요인으로 구분할 수 있으며 이와 관련한 기존 선행연구를 검토하였다.

강정규(1985)은 경부고속도로, 호남고속도로를 대상으로 종단구배가 6%이상이면 사고율이 급격히 증가하는 것으로 분석하였다. 임강원(1985)은 곡선반경이 짧아질수록 단독사고의 비율이 높아지지만 사고율 증가율 폭이 커지므로 MVK(백만차량-km)당 사고율은 증가하며 경부고속도로(4차선)보다 호남고속도로(2차선)곡선구간에서 곡선반경이 짧은 경우 사고가 더 많이 일어나는 것으로 분석했다. 건설기술연구원에서는 도로의 기하구조와 사고율의 관계에 관한 연구(1990)에서 곡선반경이 300m이상인 지점에서는 곡선반경이 커질수록 사고율이 낮아지는 결과를 얻었다. 강정규, 이성관(2002)은 호남고속도로를 대상으로 도로 선형요소 중 교통사고와 관계가 큰 것으로 분석된 곡선반경, 곡선 길이, 직선도로구간의 길이 등을 이용하여 고속도로 선형조건별 안전도평가모형을 구축하였다.

2) 인터체인지 교통사고 기존연구

(1) 국외 연구

Cirillo, Deitz등은 13개 인터체인지 구성요소를 통해 6개의 다중회귀모형을 이용하여 인터체인지 사고율 예측을 위한 모형을 개발하고 교통량과 인터체인지 형식별 기하구조에 의해 교통사고 발생이 영향을 받게 된다는 결과를 얻었다.

Bauer와 Harwood(1997)는 인터체인지 연결로

와 가/감속 차로구간의 사고예측모형을 개발하였다. 인터체인지 연결로 교통사고에 영향을 주는 주요인자로 연결로를 지나는 교통량, 연결로의 형식, 가·감속차로 길이, 연결로 길이, 연결로의 유·출입여부라는 결론을 제시했다.

(2) 국내 연구

박병호와 류승옥(2002)은 트럼펫 인터체인지의 트럼펫 A형과 B형을 각각 Under-pass와 Over-pass로 나눈 뒤 연결로형태별 최소곡선반경, 본선최소곡선반경, 본선과 진입도로의 교차각등을 독립 변수로 하여 모형을 개발하였다. 하지만 제한된 분석 자료로 인하여 통계적 유의성이 다소 낮고 다중선형회귀 분석시 고려해야 할 분산분석, 다중 공선성등의 분석을 수행하지 못하였다는 한계점을 가진다.

윤병조(2005)는 트럼펫 인터체인지 유출연결로에서 연결로 형식별 사고예측모형을 개발하였다. 기존에서 최소곡선반경만을 다루어 왔던 것과는 달리 대원과 소원의 곡률차를 이용하였지만 이후연구에서 비선형성을 가지는 교통사고의 특성을 감안하지 않고 다중선형회귀모형으로 개발하여 모형의 한계점을 가진다.

III. 모형의 설정

1. 모형의 구조설정

1) 비선형 회귀분석

교통사고를 분석하고 예측모형을 개발하기 위한 교통사고연구에 주로 다중선형회귀모형(Multiple linear regression model)이 쓰여 왔다. 하지만 사고는 random하고 불연속적이며 의도하지 않게 일어난다는 점에서 교통사고를 분석하는데 있어 설명력이 낮을 수 있다.

첫째, 선형회귀모형은 교통사고발생의 오차구조를 정규분포(normal distribution)로 간주하고 교통사고 건수와 관계없이 편차가 일정하다는 가정하에 접근하는데 교통사고는 임의적이고 산발적으로 일어나는 경향을 보임으로 정규분포를 따르지 않을 것이다.

둘째, 기울기가 표본의 구성에 따라 과소 또는 과대 추정될 수 있는 구조적 취약성을 가진다.

셋째, 일정기간 사고가 발생하지 않았거나 낮은 사고건수에 대해서 사고발생건수가 음수값(negative value)을 갖는 경우 이 문제를 임의로 해결하려면 사고 표본중 이에 해당하는 일부 샘플을 제거하고 분석하는 기법(Left-Truncating

the Accident Frequency at zero)을 사용하고 이에 따라 불합리한 결론이 도출될 오류가 있다.

또한 포아송회귀모형(Poisson regression model)을 이용할 경우 평균과 오차의 분산분포가 동일하다는 가정을 가지고 있어 과분산 현상(overdispersion)을 반영하지 못한다.

본 연구에서는 교통사고의 특성을 반영할 수 있는 음이항분포(negative binomial distribution) 비선형 교통사고예측모형을 제시한다.

2) 이론적 구조

Venkataraman Shankar(1995)등은 도로 기하구조 및 환경요소가 도시고속도로에서의 사고에 미치는 영향을 음이항회귀모형을 통해 도출하였다. 대부분의 사고분포가 과분산현상을 보이므로 포아송 분포를 보완하기 위한 것이다. 이 연구에서는 곡선구간의 수, 최소곡선반경, 도로구간의 최대 경사 등의 도로 기하구조요소를 모형식에 적용하였다. Rice(1997)또한 포아송 분포로 설명될 수 있는 사건이 과분산현상을 보일 때 음이항분포를 포아송분포의 대체수단으로 제시하였다. 이기영(2004)은 우리나라 전역 고속도로대상으로 교통사고 자료(1998)에 포아송분포와 음이항분포를 적용하여 버스과 화물차의 사고유형별 적합한 사고모형 개발하였다.

음이항회귀모형은 다음과 같은 이론적 구조를 가진다. 사고건수 Y_i 가 음이항분포를 따른다고 가정했을 때 사고가 일어날 확률 $P(Y_i)$ 은 parameter α 와 $\kappa(0 \leq \alpha \leq 1, \kappa \geq 0)$ 와 함께 <식 1>과 같은 관계식을 가진다. ($y_i = 0, 1, 2, \dots$)

$$P(Y_i = y_i; \alpha, \kappa) = \frac{(y_i + \kappa - 1)!}{y_i! (\kappa - 1)!} \frac{\alpha^{y_i}}{(1 + \alpha)^{y_i + \kappa}} \quad \text{---<식 1>}$$

$$\text{평균} = E(Y) = U_i = \kappa \alpha \quad \text{---<식 2>}$$

$$\text{분산} = \text{Var}(Y) = \kappa \alpha + \kappa \alpha^2 = \mu_i + \frac{\mu_i^2}{\kappa} \quad \text{---<식 3>}$$

모형에서의 각 Parameter는 로그우도의 음수를 최소화 시켜서 얻을 수 있다. α (overdispersion parameter)의 값이 '0'이라면 평균과 분산이 같아져 포아송분포에 적합하다는 결론이 나올 수 있는데 사고건수의 경우 α 가 0이 되는 경우는 거의 없다고 할 수 있다.

log likelihood의 음이항 분포는 <식 4>와 같은 관계식을 가진다.

$$\log(L) = \sum_{i=1}^n y_i \log[\alpha / (1 + \alpha)] - n \kappa \log(\alpha + 1) + \text{function } y_i, \kappa \quad \text{---<식 4>}$$

$\log[\alpha / (1 + \alpha)]$ 에 $\alpha = \mu_i / \kappa$ 를 대입하면

$$\log(\mu_i / (\mu_i + \kappa)) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \dots + \beta_q X_{iq} \quad \text{---<식 5>}$$

와 같다. 음이항 분포의 Parameter α 와 κ 는 일반화된 선형모형과 공식에 의하여 간접적으로 추정되고 회귀계수 $\beta_0, \beta_1, \beta_2, \dots, \beta_q$ 또한 같은 과정을 통하여 얻어지게 된다. SAS에서는 PROC GENMOD(a generalized linear model procedure) 기능을 통해 회귀계수값을 얻는다.

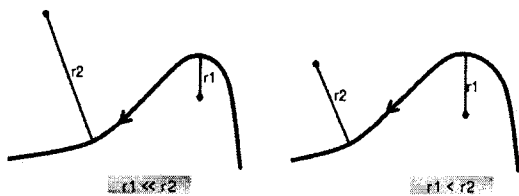
2. 변수의 설정

강민욱, 손봉수(2002)의 연구에서는 곡선구간의 경우 각 곡선구간마다 서로 다른 기하구조상의 특성이 있으므로 배향곡선구간과 단일곡선구간으로 나누어 분석하였다. 서로 다른 방향으로 굽은 곡선구간의 곡선반경이 상이할 때와 비슷할 때 안전도에 미치는 영향이 다를 것이라 판단하였기 때문이다. 평면선형상 곡선구간만을 별도로 하여 배향곡선을 분석할 경우 가능한 기하구조요소로는 다음 <표 3>과 같다.

<표 3> 배향곡선으로 분할시 적용가능한 기하구조변수

변수	내용
작은 곡선반경(km) X_{r1}	두 개의 곡선 중 곡선반경이 작은 값
큰 곡선반경(km) X_{r2}	두 개의 곡선 중 곡선반경이 큰 값
$X_{r1}X_{r2}(km^2)$	두 곡선반경의 곱
편경사 변화값(%/km)	각 배향곡선상에서 구간길이에 대한 편경사의 변화값

본 연구에서의 고속도로 인터체인지 연결로 형태별 준직결연결로와 루프연결로에서도 <그림 2>, <그림 3>과 같이 연속된 두 곡선이 연결된다. 연결로를 진입하여 지나는 첫 번째 곡선구간이 굽어보구간이라면 이 곡선구간의 기하구조가 사고에 미치는 영향은 연속된 다음곡선구간에도 영향을 미칠 것이라 판단된다. 따라서 차량이 첫 번째 주행하는 곡선구간의 곡선반경(r_1)과 두 번째 주행하는 (r_2)간의 관계를 설명할 수 있는 설계요소가 필요하다.



<그림 2> 준직결연결로



<그림 3> 루프연결로

기하구조 특성상 나타나는 두 곡선 반경의 역

수(곡률)의 차를 $(1/r_1 - 1/r_2) \times 100$ 으로 정의하여 연속된 곡선과 곡선사이의 관계를 설명하였다. 직결연결로는 단일 곡선구간으로 이루어져 있으므로 곡률 = $1/r_1$ 을 변수로 설정하였다.

강정규, 이성관(2002)의 연구에는 곡선길이가 늘어날수록 사고율은 점차 감소하다가 일정 길이 이상이 되면 사고율이 일정수준을 유지하거나 오히려 증가하는 것으로 나타났다. 입체교차로 설계 지침 (건설교통부, 2005)에서는 연결로 곡선의 길이가 짧고 도로 교각도 짧은 경우 평면곡선반경이 실제 크기보다 작게 느껴져 크게 곡선부를 회전하려는 경향으로 사고의 위험이 있다고 한다. 이는 사고에 영향을 미치는 요인에 최소곡선반경 뿐만 아니라 곡선길이도 함께 복합적인 영향을 미치는 것으로 판단되며 연결로의 곡선길이가 고속도로 연결로에도 영향을 미친다는 가정하에 본 연구에는 연결로의 곡선길이(m)를 변수로 설정하여 모형개발에 하였다.

고속도로 인터체인지 연결로는 인터체인지형태 및 연결로 형태에 따라 본선에서 틀게이트를 통해 진출하려는 차량(유출)과 틀게이트를 통해 본선으로 진입하려는 차량(유입)으로 구성되며 차량의 유입과 유출과정에 따라 사고에 미치는 사고발생 행태는 다를 것이라 판단된다. 윤병조(2005)의 연구에서도 유출 및 유입형태에 따라 사고가 미치는 영향이 다를 수 있고 유출연결로에서 사고가 더 많이 발생하므로 유출연결로에 대한 회귀모형을 개발하였다. 본 연구에서는 유출은 '1', 유입은 '0'으로 연결로를 표현하여 변수를 설정하였다..

기존 교통사고를 분석한 연구들 중 상당수가 교통량을 이용하여 교통사고 예측모형을 개발하고 교통량이 교통사고를 설명하는데 중요한 변수임을 보였다. 고속도로 인터체인지 연결로에서는 사고순간의 구체적인 교통량을 알 수 없고 교통량의 직접적 영향을 세부적으로 통제된 조사를 하는 것이 현실적으로 불가능하였다. 따라서 이를 대신하여 교통량을 대표할 수 있는 연평균일교통량(AADT)이 교통사고에 미치는 영향을 분석하고 기존 연구와 같이 교통사고와의 관계가 비선형의 관계를 나타냄을 보였다. 또한 기존 연구에서 포아송회귀모형과 음이항회귀모형에서 교통량 변수를 $\log(AADT)$ 로 처리할 때 모형의 적합도가 높아지게 됨을 보였다. 이는 교통량에 비해 사고건수가 매우 작아 발생할 수 있는 오차구조에 기인한 것으로 보이며 $\log(AADT)$ 를 변수로 설정하였다.

고속도로 연결로에서의 교통사고를 발생시킨 원인 차량으로 <표 4>의 2종이상 차량에 해당하는 중차량이 차지하는 비율은 59%로 연결로를 차지하는 차량비율에 비해 상당한 부분을 차지하였으며 치사율이 매우 높았다. 따라서 중차량비율이 높을수록 사고 및 사고의 심각도가 증가할 수 있다는 가정하에 중차량비율이 사고에 미치는 영향을 분석하였다. 해당 연결로를 이용하는 중차량AADT를 구하고 램프AADT와의 공선성을 피하기 위해 중차량의 비율을 구하여 변수로 적용하였다.

<표 4> Toll Collection System 차종분류

TCS분류	차종	차종분류(11종)
1종	승용차(6인 이하)	1
	승합차(16인승 이하)	2
	화물차(2.5톤 미만)	5
2종	승합차(17인승 이상~25인승 이하)	3
	화물차(2.5톤 이상~5톤 이하)	6
3종	승합차(25인승 초과)	4
	화물차(5톤 초과~10톤 미만)	7
4종	화물차(10톤 이상~20톤 미만)	8
5종	화물차(20톤 이상)	9
	세미트레일러	10
	풀트레일러	11

교통사고는 분석자가 용이하게 모형을 적용할 수 있도록 현실적 적용가능하며 유의한 인자만을 선별하여 변수로 이용해야 하므로 기존 연구고찰과 기초상관분석을 통해 사고에 미치는 요인이 미약하거나 다중공선성이 있다고 판단되는 변수들은 모형 개발에 이용하지 않았다. 개발된 모형에서도 본 연구에서 선정한 변수들이 교통사고를 설명하는데 의미 있는 변수들이라는 것이 입증된다. 종속변수는 2003~2005년 사고지점 연결로 형태별 사고건수로 정의한다.

<표 5> 변수 자료

변수		내용
dummy	X ₁	유출입지점여부 유출지점 '1', 유입지점 '0'
	X ₂	곡선장 곡선의 길이(m)
continuous	X ₃	교통량 평균 AADT
	X ₄	중차량비율 중차량 평균AADT/평균AADT
	X ₅	곡률or곡률차 곡선반경의 역수
	Y	사고건수 3년간 사고지점 사고건수

IV. 교통사고 분석

1. 분석개요

한국도로공사관내 폐쇄식구간 트럼펫 인터체인지중 122개소에서 2003~2005년간 326건의 교통사고가 발생했다. 연결로의 주행행태에 영향을 미치는 도로공사가 최소한으로 일어나고 사고발생 연결로

위치가 명확하게 파악되며 모형 구축에 필요한 자료 수집이 가능한곳을 분석대상으로 하였다.

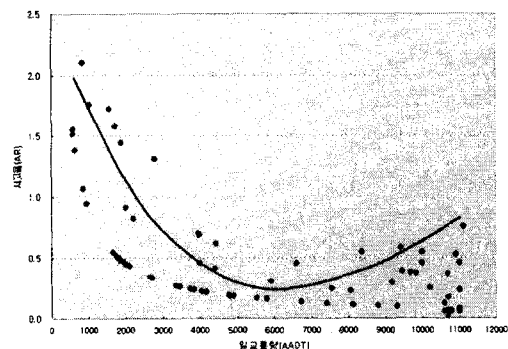
<표 6> 트럼펫 인터체인지 형식별 사고기록

인터체인지명	사고건수				Total	
	유입직결 연결로	유출직결 연결로	루프 연결로	준직결 연결로		
A형	경사	2	7	0	0	9
	경주	0	5	4	6	15
	광양	0	0	4	3	7
	기흥	0	1	0	1	2
	동대구	2	1	0	0	3
	목천	0	0	0	1	1
	서대전	0	0	0	1	1
	수원	0	2	4	0	6
	신탄진	0	2	0	2	4
	안성	0	2	0	0	2
	오산	2	0	1	1	4
	옥과	0	1	0	1	2
	천안	0	1	4	0	5
	현풍	0	1	4	1	6
	김곡	0	1	0	2	3
	김천	1	2	1	1	5
	단성	0	0	1	1	2
다평	0	0	3	1	4	
B형	동군산	3	0	1	1	5
	동김해	0	1	1	1	3
	무안	1	2	2	0	5
	북대전	0	2	0	1	3
	비봉	0	0	2	0	2
	서순천	1	4	0	1	6
	서진주	0	0	6	0	6
	서평택	0	1	3	8	12
	순천	0	3	1	0	4
	양지	0	0	0	2	2
	이천	1	1	1	3	6
진주	0	3	0	0	3	
청주	1	5	1	1	8	
Total 31개 IC	14	48	44	40	146	

2. 교통사고분석

1) 교통량과 사고율의 관계

Ivan(2004)은 교통량과 사고율은 상당히 복잡한 비선형관계가 도출되는 것으로 분석하였다. 오철, 장명순(1999)은 신갈-안산 고속도로를 기본,터널,영업소 구간으로 분류하여 교통혼잡도와 교통사고율의 관계를 분석한 결과 V/C와 사고율관계가 U형의 곡선 형태를 나타냄을 보였다. 본 연구에서 인터체인지 연결로를 지나는 연평균일교통량(3년간 AADT의 평균)과 교통사고율과의 관계를 분석한 결과 <그림 4>과 같다.



<그림 4> 교통량과 사고율의 관계

교통량이 적은 경우 사고율이 높지만 증가할수록 사고율은 감소하며 9000~1000대 이상부터 다시 사고율이 증가하는데 이는 일정한계이상의 교통량의 증가는 차량 간의 상충과 혼잡을 유발할 수 있기 때문으로 판단된다. 교통량에 비해 사고건수는 매우 작기 때문에 사고율을 백만 차량 당 사고건수로 정의 하였다.

$$AR = \frac{\text{사고건수}(N) \times \text{백만}}{365 \times 3\text{년} \times AADT} \quad \text{<식 6>}$$

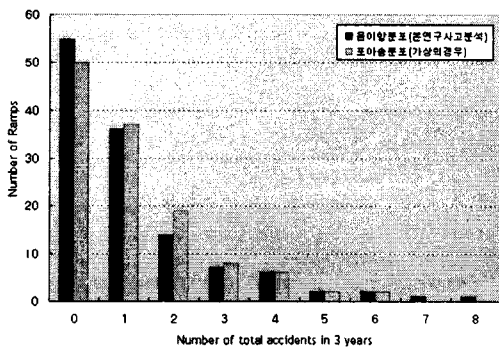
2) 교통사고 빈도분석 및 분포추정

사고건수를 이산적확률변수(Discrete Random Variable)로 해석하는 회귀식의 모형 추정을 위해서는 분포의 특성을 분석하여 적절한 분포를 선택하여야 한다. 교통사고모형추정에 이용될 수 있는 분포의 특성은 다음 <표 7>와 같이 평균(M)과 분산(S²)을 비교하여 결정할 수 있다.

<표 7> 음이항분포의 적합도 분석(전체연결로)

분포	S ² /M	산출식
포아송분포	=1	$P(x) = \frac{m^x e^{-m}}{x!}$
음이항분포	>1	$P(x) = C_{\kappa-1}^x p^\kappa (1-p)^x$

총 146건의 사고에 대하여 사고건수별 인터체인지 연결로의 분석결과 <그림 5>과 같다. 검정통계량 pearson χ^2 값이 3.38으로 임계값 $\chi_{0.05}^2(5) = 9.49$ 를 넘지 못하고 유의수준 $\alpha = 0.05$ 에서 음이항분포를 따름을 보였다. 평균이 같은 경우의 포아송분포(평균=1.7742 분산=1.7742)를 가상으로 그려본 결과 음이항 분포에 비해 꼬리부분이 짧은 특성을 나타내는데 이는 기준점으로부터 사고의 분포가 많이 흩어져 있다는 것을 의미한다. 사고건수의 분포가 음이항 분포를 따름으로 음이항회귀식으로 모형을 개발하는 것이 타당하며 모형개발 이후의 모형의 적합도 검증에서도 모형의 구조가 과산포를 따름을 보였다.



<그림 5> 전체연결로 사고빈도분석
이와 같은 기법을 이용하여 직결연결로, 루프

연결로, 준직결연결로 각각에서의 사고빈도분석과 검정통계량 값의 분석결과 모두 음이항 분포를 따름을 보였다.

V. 모형의 추정

본 연구의 특성에 맞게 제시한 변수설정방법에 따라 교통사고 예측모형을 추정하고 추정된 모형의 적합도를 통계적 방법을 이용하여 검증하였다. 전체연결로의 연결로 형식중 직결연결로는 곡률차가 존재하지 않으므로 전체연결로 모형추정에서는 곡률(X₅)변수를 이용하였다.

1. 전체연결로(Total ramps) 모형추정

1) 기초분석

<표 8> 기초통계분석

변수		표본수	평균	표준편차	합계	최소값	최대값
dummy	유출입여부(X ₁)	124	0.5081	0.5020	63	0	1
	곡선장(X ₂)	124	359.9594	93.9210	44635	185	768
	교통량(X ₃)	124	8.3060	0.9039	1029.94	6.36	10.52
continuous	중차량비율(X ₄)	124	18.0034	7.1068	2232.42	4.65	39.42
	곡률(X ₅)	124	0.3534	0.4130	43.82	0.01	1.89
	사고건수(Y)	124	1.1774	1.5980	146	0	8

독립변수간 높은 다중공선성이 존재할 경우 회귀모형의 결과가 왜곡되는 현상이 나타날 수 있다. 모형전체의 유의성이 통계적 검증을 통해 유의한 것으로 나타나도 추정된 계수의 표준오차가 커져 검정결과에서 일부 회귀계수들이 통계적으로 유의하지 않게 나타나는 것이다. 이러한 결과는 독립변수들간의 높은 상관관계로 인해 한 독립변수의 설명력이 다른 독립변수에 의해 흡수될 수 있기 때문이며 본 연구에서도 다중공선성 분석을 통해 독립변수들의 타당성을 검증한결과 문제가 없는 것으로 판단되었다.

2) 모형의 추정결과

SAS의 PROC GENMOD에서 제공하는 임의의 parameter κ 0.7316으로 <표 9>에서 1에 가까운 deviance/(n-p)값을 얻을 수 있다. 또한 Pearson chi-square 통계값을 n-p로 나누었을 때의 값이 1.0022로 1에 아주 근접하고 0.8~1.2사이의 값을 보이므로 모형은 매우 적합하다고 간주할 수 있다. 이 값이 1보다 크므로 모형의 구조가 과산포(overdispersion)를 따름을 의미한다.

<표 9> 모형의 적합도 검증

Criterion	Model diagnostics
κ factor	0.7316
deviance/(n-p)	1.0000
Pearson chi-square/(n-p)	1.0022

Chi-square likelihood ratio 통계값 분석결과 <표 10>과 같다. 교통량(X_3)이 교통사고를 판단하는데 가장 유의성있는 변수로 분석되었다. 또한 중차량 비율(X_4)과 곡률(X_5)도 유의수준 5%를 만족하였으며 유출입여부(X_1)는 유의수준 10%를 만족했다. 곡선장(X_2)은 유의수준 20%를 만족하였지만 변수로 이용했을때 모형의 적합도가 적절하게 나왔고 필요한 변수라고 판단되었기 때문에 모형 추정에 이용하였다.

<표 10> 변수의 추정치 분석

변수	DF	Estimate(β)	Chi-square	Pr > Chi-square
Intercept	1	-5.7037	15.26	0.0000
유출입여부(X_1)	1	0.3912	2.74	0.0976
곡선장(X_2)	1	0.0017	1.94	0.1634
교통량(X_3)	1	0.4854	12.38	0.0004
중차량비율(X_4)	1	0.0386	5.16	0.0231
곡률(X_5)	1	0.5333	3.86	0.0496

계수값을 대입하면 다음과 같다. $E(A)_T$ 는 전체 연결로 모형추정에 의한 예측 사고건수를 의미하고 $AADT_{ramp}$ 는 교통량(X_3)을 의미한다.

$$E(A)_T = \exp(\beta_0) \cdot (AADT_{ramp})^{\beta_3} \cdot \exp(\beta_1 X_1) \cdot \exp(\beta_2 X_2) \cdot \exp(\beta_4 X_4) \cdot \exp(\beta_5 X_5)$$

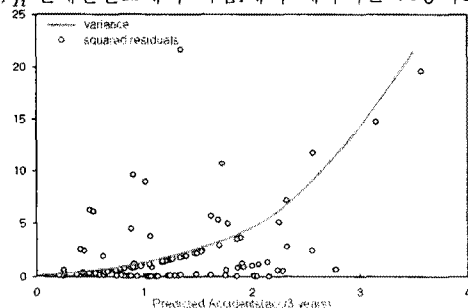
3) 모형의 적합도 분석

(1) 잔차분석

모형개발의 결과 예측치와 관측치의 차(residual)를 비교하는 통계적 방법을 이용하여 모형의 적합도를 분석하였다. <그림 6>의 예측치와 관측치의 차의 제곱한 값을 점으로 찍었을 때의 값들과 <식 7>의 variance 함수에 의해 추정되는 곡선과의 비교결과 전체적으로 점들이 곡선을 따르므로 모형의 적합도는 높다고 할 수 있다. (κ : parameter)

$$Var(Y)_{T_i} = E(Y)_{T_i} + \frac{[E(Y)_{T_i}^2]}{\kappa} \text{-----<식 7>}$$

$E(Y)_{T_i}$: 전체연결로에서 지점 i에서 예측되는 교통사고건수



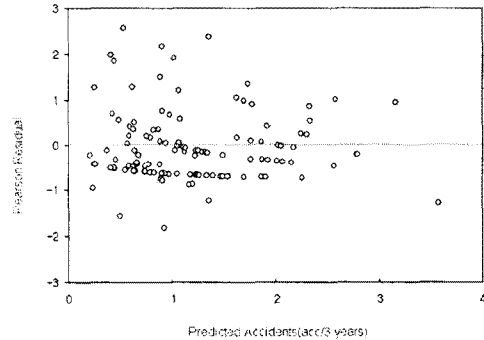
<그림 6> Estimated Variance and squared residuals

(2) Pearson Residual 분석

<그림 7>에서는 예측된 $E(Y)_T$ 를 축으로 하여 <식 8>를 통해 구해진 PR_i (Pearson Residual) 값들을 점으로 찍었을때 '0'을 중심으로 대칭에 가까우므로 모형의 적합도는 높다고 할 수 있다.

$$Pearson\ Residual\ (PR_i) = \frac{E(Y)_{T_i} - y_i}{\sqrt{var(y_i)}} \text{----<식 8>}$$

y_i : 전체연결로에서의 지점 i에서 관측되는 교통사고건수
 $var(y_i)$: 전체연결로에서의 지점 i에서 관측되는 교통사고건수의 분산



<그림 7> Pearson Residual 분석

2. 직결연결로(Direct ramps) 모형추정

1) 모형의 추정결과

직결 연결로에서는 도로 기하구조 특성상 단일곡선으로 구성되기 때문에 최소곡선반경도 하나이다. 따라서 곡률차는 존재하지 않게 되고 설명변수로 곡률(X_5)을 사용하였다. 독립변수들 간의 다중공선성에는 문제가 없다고 판단된다.

모형의 추정결과 parameter κ 0.6919 으로 1에 가까운 deviance/(n-p)값을 얻을 수 있다. Pearson chi-square 통계값을 n-p로 나누었을 때의 값이 1.1133로 1에 아주 근접하고 0.8~1.2 사이의 값을 보이므로 모형은 매우 적합하다고 간주할 수 있다.

<표 11> 변수의 추정치 분석

변수	DF	Estimate(β)	Chi-square	Pr > Chi-square
Intercept	1	-5.9949	9.64	0.0019
유출입여부(X_1)	1	1.0610	9.46	0.0021
교통량(X_3)	1	0.3983	3.93	0.0474
중차량비율(X_4)	1	0.0480	4.21	0.0401
곡률(X_5)	1	1.2790	7.25	0.0071

Chi-square likelihood ratio 통계값 분석결과 유출입여부(X_1)이 직결 연결로에서의 교통사고를 판단하는데 가장 유의성있는 변수로 분석되었다. 또한 교통량(X_3), 중차량 비율(X_4)과 곡률(X_5)도 유의수준 5%를 만족하였다. 곡선장(X_2)을 변수로 사용하여 분석한 결과 유의수준 20%를 만족시키지 않았고 모형의 적합도도 적절하

지 않게 나왔으며 고속도로 인터체인지 연결로를 설명하는데 필요하지 않은 변수라고 판단되었기 때문에 모형 추정에 이용하지 않았다.

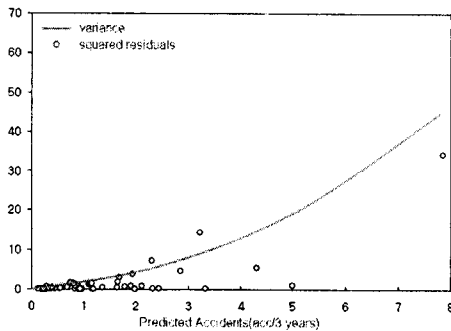
추정된 변수의 계수값($\beta_0, \beta_1, \beta_2, \beta_4, \beta_5$)를 대입하면 최종모형식의 형태는 다음과 같다. $E(A)_D$ 는 직결연결로 모형추정에 의한 예측 사고건수를 의미하고 $AADT_{ramp}$ 는 교통량(X_3)을 의미한다.

$$E(A)_D = \exp(\beta_0) \cdot (AADT_{ramp})^{\beta_3} \cdot \exp(\beta_1 X_1) \cdot \exp(\beta_4 X_4) \cdot \exp(\beta_5 X_5)$$

2) 모형의 적합도 분석

(1) 잔차분석

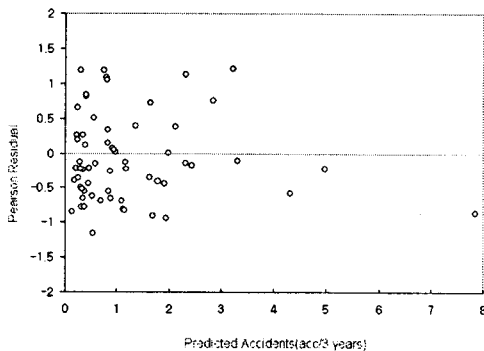
모형개발의 결과에 의한 예측치와 관측치의 차를 비교하는 통계적 방법을 이용하여 모형의 적합도를 분석결과 전체적으로 점들이 곡선을 따르므로 모형의 적합도는 높다고 할 수 있다.



<그림 8> Estimated Variance and squared residuals

(2) Pearson Residual 분석

예측된 $E(Y)_D$ 를 축으로 하여 PR_i (Pearson Residual) 값들을 점으로 찍었을 때 '0'을 중심으로 대칭에 가까우므로 모형의 적합도는 높다고 할 수 있다.



<그림 9> Pearson Residual 분석

3. 루프연결로(Loop ramps) 모형추정

1) 모형의 추정결과

독립변수들간의 다중공선성에 문제가 없다고 분석되었으며 parameter κ 0.5405 으로 1에 가까운 deviance/(n-p)값을 얻을 수 있고 Pearson

chi-square 통계값을 n-p로 나누었을때의 값이 1.1131로 1에 근접하고 0.8~1.2 사이의 값을 보 이므로 모형은 매우 적합하다. 이 값이 1보다 크므로 모형의 구조가 과산포를 따름을 의미한다. Chi-square likelihood ratio값 분석결과 교통량 (X_3)이 교통사고를 판단하는데 가장 유의성있는 변수로 분석되며 유출입여부(X_1), 곡선장(X_2)과 함께 유의수준 5%를 만족한다. 중차량비율(X_4), 곡률차(X_5)는 유의수준 10%를 만족한다.

<표 12> 변수의 추정치 분석

변수	DF	Estimate(β)	Chi-square	Pr>Chi-square
Intercept	1	-4.5630	1.80	0.1797
유출입여부(X_1)	1	2.6841	4.59	0.0321
곡선장(X_2)	1	-0.0117	5.51	0.0189
교통량(X_3)	1	0.7171	6.67	0.0098
중차량비율(X_4)	1	0.0693	2.38	0.0926
곡률차(X_5)	1	-1.7996	3.37	0.0663

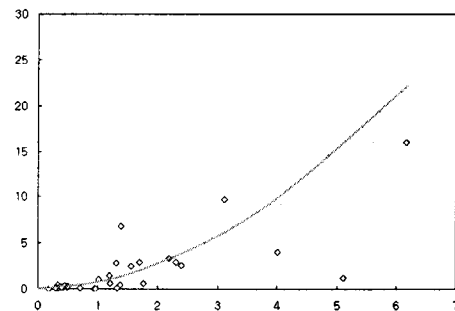
추정된 최종모형식의 형태는 다음과 같다.

$$E(A)_L = \exp(\beta_0) \cdot (AADT_{ramp})^{\beta_3} \cdot \exp(\beta_1 X_1) \cdot \exp(\beta_2 X_2) \cdot \exp(\beta_4 X_4) \cdot \exp(\beta_5 X_5)$$

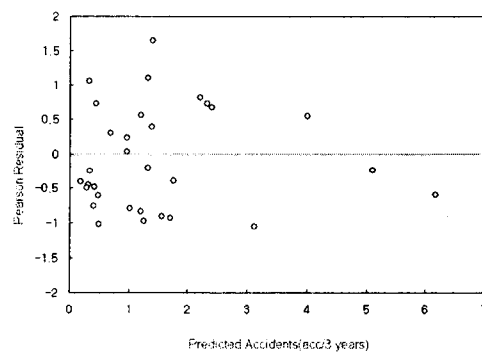
2) 모형의 적합도 분석

1) 잔차분석

예측치와 관측치차의 제곱한 값을 점으로 찍었을 때의 값들과 variance함수에 의해 추정되는 곡선과의 비교결과 전체적으로 점들이 곡선을 따르므로 모형의 적합도는 높다고 할 수 있다.



<그림 10> Estimated Variance and squared residuals



<그림 11> Pearson Residual 분석

2) Pearson Residual 분석

예측된 $E(Y)_L$ 를 축으로 하여 PR_i (Pearson Residual) 값들을 점으로 찍었을때 '0'을 중심으로 대칭에 가까우므로 모형의 적합도는 높다고 할 수 있다.

4. 준직결연결로(Semi-direct ramps) 모형 추정

1) 모형의 추정결과

독립변수들간의 다중공선성에 문제가 없다고 분석되었으며 parameter κ 0.2973 으로 1에 가까운 deviance/(n-p)값을 얻을 수 있다. 또한 Pearson chi-square 통계값을 n-p로 나누었을 때의 값이 1.0481로 1에 아주 근접하고 0.8~1.2 사이의 값을 보이므로 모형은 매우 적합하다고 간주할 수 있다. 또한 이 값이 1보다 크므로 모형의 구조가 과산포를 따름을 의미한다.

<표 13> 변수의 추정치 분석

변수	DF	Estimate(β)	Chi-square	Pr>Chi-square
Intercept	1	-8.3788	10.61	0.0011
유출입여부(X_1)	1	-0.0464	2.39	0.1255
곡선장(X_2)	1	0.0043	6.81	0.0091
교통량(X_3)	1	0.5852	5.52	0.0188
중차량비율(X_4)	1	0.0932	11.29	0.0008
곡률차(X_5)	1	0.6246	3.71	0.0799

Chi-square likelihood ratio 통계값 분석결과 중차량 비율(X_4)이 교통사고를 판단하는데 가장 유의한 변수로 판단되며 곡선장(X_2), 교통량(X_3), 곡률(X_5)과 함께 유의수준 5%를 만족하였다. 유출입여부(X_1)는 유의수준 10%를 만족하지 못하는데 이는 유출입여부(X_1)의 합계 자체가 사고건수에 비해 적기 때문에 1(유출)일때의 상황을 충분히 반영하지 못한 것으로 보인다. 하지만 유의수준 20%를 만족하고 모형 개발자의 의도에 따라 연구에서 유출입여부(X_1)를 변수로 사용했을때 모형의 적합도가 적절하게 나왔고 고속도로 인터체인지 연결로를 설명하는데 필요한 변수라고 판단되어 모형 추정에 이용하였다. 추정된 최종모형식의 형태는 다음과 같으며 $E(A)_S$ 는 준직결연결로에서의 모형추정에 의한 예측 사고건수를 의미한다.

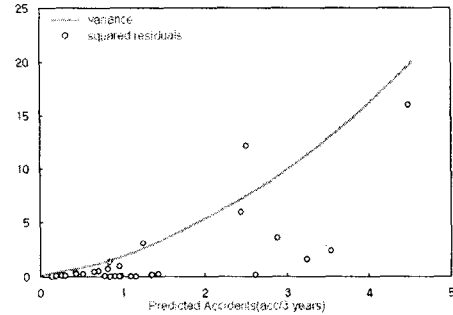
$$E(A)_S = \exp(\beta_0) \cdot (AADT_{ramp})^{\beta_3} \cdot \exp(\beta_1 X_1) \cdot \exp(\beta_2 X_2) \cdot \exp(\beta_4 X_4) \cdot \exp(\beta_5 X_5)$$

2) 모형의 적합도 분석

(1) 잔차분석

모형개발의 결과에 의한 예측치와 관측치의 차의 제공한 값을 점으로 찍었을 때의 값들과

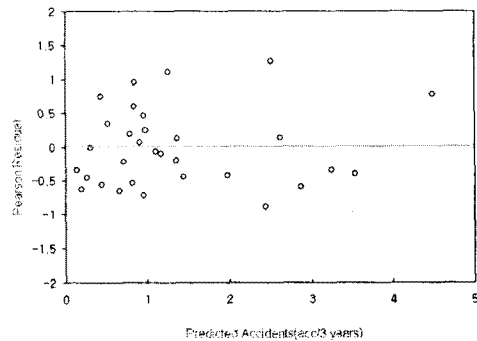
variance 함수에 의해 추정되는 곡선과의 비교 결과 전체적으로 점들이 곡선을 따르므로 모형의 적합도는 높다고 할 수 있다.



<그림 12> Estimated Variance and squared residuals

2) Pearson Residual 분석

예측된 $E(Y)_S$ 를 축으로 하여 PR_i (Pearson Residual) 값들을 점으로 찍었을때 '0'을 중심으로 대칭에 가까우므로 모형의 적합도는 높다고 할 수 있다.



<그림 13> Pearson Residual 분석

VI. 결론 및 향후연구과제

1. 결론

본 연구에서는 총 31개 고속도로 트럼펫 인터체인지상에서 연결로형식별로 일어나는 교통사고와 도로 기하구조 및 교통량등의 등의 교통사고 요인들과의 관계를 분석하기 위해 통계적 기법을 이용하였다.

교통사고의 분포의 특성을 분석하여 적합도 검증을 통해 모형추정 시 가장 적절한 분포를 찾은 결과 음이항분포가 선택되었다. 선택된 분포에 기반하여 트럼펫 인터체인지 연결로 전체, 연결로 형식별(직결, 준직결, 루프연결로) 각각의 음이항회귀모형을 개발하였다.

총 4개의 모형을 개발하고 모형의 적합도를 판단하는 여러가지 통계값과 모형에서 예측한 값과 실제 관측값과의 차이를 분석한 결과 예측모형이 적합하게 구축되었음을 보였다.

- ① Total ramps(124개)에서 사고 $E(A)_T$ 를 예측,
- ② Direct ramps(62개)에서 사고 $E(A)_D$ 를 예측,
- ③ Loop ramps(31개)에서 사고 $E(A)_L$ 를 예측,
- ④ Semi-direct ramps(31개)에서 사고 $E(A)_S$ 를 예측하는 모형을 개발하였다. 모형의 통계적으로 유의한 변수들을 분석한 결과 교통량이 교통사고를 설명하는데 가장 유의한 변수로 판단되었다. 또한 중차량 비율, 곡률 or 곡률차, 곡선장, 유출입지점 여부등도 통계적으로 유의한 변수로 판단되었다. 이러한 변수들을 도로의 설계자가 도로 설계 및 운영에 적용하거나 해당 도로의 교통특성을 반영한 교통사고 절감 대책 등에 이용할 수 있을 것이다.

2. 향후연구과제

국의 연구에서 사용하는 교통사고이력자료에 서는 한 연결로를 3구간으로 나누어 coding하여 이력자료를 산출한다. 따라서 사고가 연결로에 진입한 직후 일어났는지, 연결로 진입후 연결로 중간지점에서 일어났는지, 연결로 이용을 마치고 진출하려는 때인지 정확한 지점에 대한 신뢰성 있는 분석이 가능하다. 이러한 지점의 파악이 가능하다면 연결로 진입 전·후의 감·가속차로구간에서 또는 감·가속차로구간의 도로 기하구조로 인하여 발생하는 사고에 대한 설명이 가능할 것으로 판단된다.

참고문헌

1. 강정규(1985), "고속도로의 기하구조가 교통사고 발생에 미치는 영향에 관한 연구", 서울대학교 환경대학원.
2. 한국건설기술연구원(1985), "도로의 기하구조와 교통 안전성에 관한 미국 연방도로청의 보고서", 도로교통 안전 기술보고서
3. 오철, 장재남, 장명순(1999), "고속도로 시설물 구간의 교통혼잡도와 사고율의 관계분석(신갈-안산 고속도로를 중심으로)", 대한교통학회지, 제17권 제2호, pp.21~27
4. 이기영(1999), "고속도로 사고모형의 개발과 적용-확률회귀모형을 이용한 버스사고 요인 분석", 대한교통학회 제 36회 학술발표회
5. 건설교통부(2000), "도로의 구조 · 시설기준에 관한 규칙해설 및 지침"
6. 강정규, 이성관(2002), "고속도로 선형조건별 교통사고 위험도 평가모형 개발(호남고속도

- 로를 중심으로)", 대한교통학회지, 제 20권 제2호, pp.107~124
7. 강민욱, 손봉수, 도철웅(2002), "도로의 기하구조에 따른 고속도로 사고예측모형개발", 대한토목학회지, 제22권 6-D호, pp.1077~1088
8. 박병호, 류승옥(2002), "트럼펫 인터체인지의 형식과 교통사고율의 관계", 한국 지역개발학회지 제 14권, 제1호, pp. 107~124
9. 건설교통부(2005), "입체교차로 설계 지침"
10. May J. F.(1964), "A Determination of an accident prone location", Traffic Engineering, Vol. 34, no.5.
11. SAS Institute Inc.(1993), "SAS/STAT Software : The PROC GENMOD Procedure, Release 9.1", SAS Technical Report P-243, Cary, NC, p.88
12. Belanger C.(1994), "Estimation of safety of four-leg unsignalized intersection", Transportation Research Record, 1467
13. Venkataraman Shankar, Fred Mannering, and Woodrow Barfield(1995), "Effect of Road-way Geometrics and Environmental Factors on Rural Freeway Accident Frequencies", Accident Analysis and Prevention, Vol.27. No.3.
14. Bauer, K. M. and D. W. Harwood(1997) "Statistical Models of accidents in Interchange Ramp and Speed-Change Lanes", FHWA-RD-97-106, FHWA
15. Rice J. A.(1997) Mathematical statistics and data analysis: 2nd edition Wadworth, INC
16. Vogt and Bared J.(1998), Accident models for two lane rural segments and intersections. Transportation Research Record, 1635, Transportation Research Board, National Research Council, Washington D.C., pp.18~29
17. SAS Institute Inc.(1999), "SAS/STAT User's Guide Version 8 : Volume 2", Cary, NC
18. Alex Pedan, Vasca Inc., Tewksbury, MA., "Analysis of Count Data Using the SAS System"
19. Ivan J.(2004), "A new approach for including traffic volumes in crash rate analysis and forecasting", Annual Meeting of the Transportation Research Board.