

# 고속도로 교통사고 분석모형 연구

## A Study of Accident Models for Highway Interchange Ramps

노 창 균

박 종 서

손 봉 수

(연세대학교 도시공학과 석사과정)

(세인시스템 대표이사/사장)

(연세대학교 도시공학과 교수)

### 목 차

#### I. 서론

1. 연구의 배경 및 목적
2. 문헌고찰

#### II. 분석자료

1. 분석자료의 선정
2. 자료분석
3. 사고자료의 분포

#### III. 모형도출

1. 변수선정
2. 트럼펫인터체인지
3. 클로버인터체인지

#### IV. 요약 및 결론

#### 참고문헌

## I. 서론

### 1. 연구의 배경 및 목적

고속도로의 안전을 증진시킬 수 있는 설계를 하기 위해서는 고속도로사고와 기하구조와의 관계를 규명하는 것이 우선이다. 하지만 안전과 기하구조와의 관계는 규정짓기 힘든 요인이 존재하며, 따라서 도로의 포장상태, 기상, 운전자 특성 및 차량의 특성과 같은 환경적 요인은 배제하고 정량적 지표로 규정 가능한 변수를 사용하였다. 본 연구의 목적은 고속도로 인터체인지에서 발생하는 교통사고를 수집이 용이한 자료만을 이용하여 신뢰성 있는 결과를 예측할 수 있는 모형으로 정립하는 것에 있다. 신뢰도 높은 모형을 통하여 현실과 가장 유사한 수학적 확률 모형을 개발함으로써 설계단계 및 평가단계에 적용가능한 모형을 제시하며, 다음과 같은 내용을 포함하고 있다. 고속도로 인터체인지 유형별 및 연결로 유형별 교통사고 발생빈도 분석을 바탕으로 교통사고 발생빈도 예측 및 평가 모형을 정립한다.

### 2. 문헌고찰

#### 1) 모형의 이론적 고찰

기존 연구는 설계속도, 곡률, 곡률차, 편구배 등을 변수로 이용하여 다중선형회귀모형을 구축(Multiple linear regression model)하였다. 그러나 불연속이며 의도하지 않아도 산발적인 형태로 발생하는 교통사고를 분석하는데에는 한계를 가질 수 밖에 없다. 따라서 사고를 이산적 확률변수로 해석하기 위하여 포아송회귀식을 도입하였으며, 이는 개념적으로 선형회귀식에 비해 사고분석에 합리적인 모형이다. 그러나 대부분의 사고분포는 과분산현상(Overdispersion)을 보이고 있으며 이는 평균과 분산이 같다는 성질을 가지고 있는 포아송분포를 위배하기 때문에 본 연구에서는 사용하지 않았다. 본 연구에서 사용한 모형은 음이항회귀모형으로 이론적 구조는 다음과 같다. 사고건수  $Y_i$ 가 음이항 분포를 따른다고 가정했을 때 사고가 일어날 확률  $P(Y_i)$ 은 parameter  $\alpha(0 \leq \alpha \leq 1)$ ,  $\kappa(\kappa \geq 0)$ 와 함께 식(1)과 같은 관계식을 가진다. ( $y_i = 0, 1, 2, \dots$ )

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

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

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

모형에서의 Parameter는 로그우도의 음수를 최소화 시켜서 얻을 수 있다.  $\alpha$ (overdispersion parameter)의 값이 '0'이라면 평균과 분산이 같아져 포아송분포에 적합하다는 결론이 나올 수 있는데 사고건수의 경우  $\alpha$ 가 0이 되는 경우는 거의 없다고 할 수 있다. log likelihood의 음이항분포는 식(4)와 같은 관계식을 가진다.

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

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

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

와 같다. 음이항 분포의 Parameter  $\alpha$ 와  $\kappa$ 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 GENMOD의 기능을 통해 식(1), 식(5)을 시행하고 회귀계수 값을 얻는다.

## 2) 기존연구 고찰

Bauer와 Harwood(1997)는 인터체인지 연결로와 가/감속 차로구간의 사고예측모형을 개발하였다. 이 모형에 적용된 인터체인지 연결로 교통사고에 영향을 주는 주요인자로 연결로를 지나는 교통량, 연결로의 형식, 가·감속차로 길이, 연결로 길이, 연결로의 유·출입여부라는 결론을 제시했다.

박병호와 류승욱(2002)은 트럼펫 인터체인지의 기하구조와 인터체인지에서 사고 발생빈도의 상관관계를 규명했다. 트럼펫 A형과 B형을 구분한 나눈 뒤 각 연결로의 최소곡선반경, 본선 최소곡선반경, 본선과 진입도로의 교차각 등을 독립 변수로하여 모형을 개발하였다. 하지만 통계적 유의성이 다소 낮고 다중선형회귀 분석시 고려해야할 분산분석, 다중 공선성등의 분석을 수행하지 못하였다는 한계점을 가진다.

윤병조 외 3인(2005)은 트럼펫 인터체인지 유

출연결로에서 연결로 형식별 사고예측모형을 개발하였다. 최소곡선반경이 아닌 곡률차를 이용하여 곡선간의 관계를 설명하였지만, 연구에서 비선형성을 가지는 교통사고의 특성을 감안하지 않고 다중선형회귀모형으로 개발한 모형의 한계점을 가진다.

김태영, 박병호(2007)는 트럼펫램프의 운전조건과 교통사고 분석에서 램프연결로의 회전방향 즉 운전방향에 따른 교통사고를 통계적인 가설검정을 통해 회전방향과 종단경사, 회전방향과 곡선반경의 조합에 따른 사고율과 사고건수의 평균을 비교분석하여 우측으로 회전하는 램프가 좌측으로 회전하는 램프보다 사고율과 사고건수가 높은 것으로 분석하였다.

박효신(2006)은 트럼펫인터체인지의 연결로형식별 사고예측모형을 개발하였다. 곡선장, 교통량, 중차량 비율, 곡선을 독립변수로 다루어 음이항회귀모형을 통한 모형을 개발하였으나, 사고예측측면에서 설계단계에서 예측이 힘든 중차량비율등의 변수로 인하여 개발된 모형의 적용성에 한계점을 가지고 있다.

## II. 분석자료

### 1. 분석자료의 선정

한국도로공사관내 인터체인지 총 314개소에서 4년간 674건의 교통사고가 발생했다. 그 중 연결로의 주행행태에 영향을 미치는 도로공사가 최소한으로 일어나고 사고발생 연결로의 위치가 명확하게 파악되었으며 모형 구축에 필요한 데이터의 수집이 가능한 총 164개의 인터체인지를 최종 분석대상으로 하여 인터체인지별, 인터체인지 연결로 형식별 총 508건의 교통사고에 대하여 분석하였다.

<표 1> 분석자료 선정 결과

구분		분석자료 선정	
트럼펫 I.C. (129개소)	직결연결로	258개 램프	196건
	준직결연결로	129개 램프	138건
	루프연결로	129개 램프	130건
	계	516개 램프	464건
클로버 I.C. (35개소)	직결연결로	35개 램프	25건
	루프연결로	27개 램프	29건
	계	62개 램프	44건
총 계		532개 램프	508건

이 중 준직결형 연결로 및 평면교차로 33개소와 12건의 사고가 발생한 6개소의 다이아몬드형 인터체인지의 경우 모델구축에 필요한 최소 30건의 Sample이 구축되지 못하여 분석대상에서 제외하였다.

<표 1>에서 보는 바와 같이 트럼펫 인터체인지 129개소에서 464건이 발생하였고 클로버 인터체인지 35개소에서 44건이 발생하였다.

트럼펫 인터체인지의 연결로 형식별로는 258개의 직결연결로에서 196건, 129개의 루프연결로 130건, 129개의 준직결연결로에서 138건이 발생하였다. 또한 클로버 인터체인지의 연결로 형식별로는 35개 직결연결로에서 25건, 27개의 루프연결로에서 총 19건의 사고가 발생하였다.

## 2. 자료 분석

Ivan(2004)은 교통량과 사고율은 선형적인 관계가 아닌 상당히 복잡한 비선형관계가 도출되는 것으로 분석하였다. 오철, 장명순 등(1999)은 신갈-안산 고속도로를 대상으로 기본구간, 터널구간, 영업소 구간 분류하여 교통혼잡도와 교통사고율의 관계를 분석한 결과에서 V/C와 사고율과의 관계가 U형의 곡선 형태를 나타냄을 보였다. 교통량에 비하여 사고건수는 매우 작기 때문에 사고율을 백만 차량 당 사고건수로 식(6)과 같이 정의 하였다.

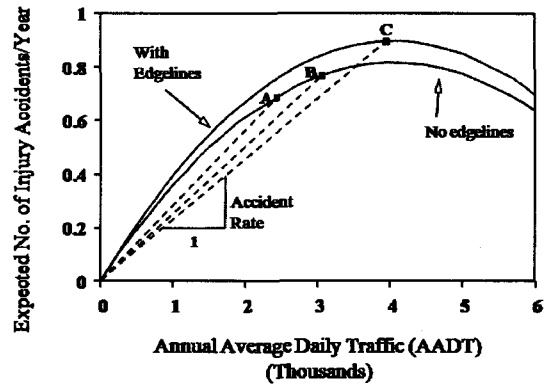
$$AR = \frac{\text{사고건수}(N) \times \text{백만}}{365 \times 3\text{년} \times AADT} \quad (6)$$

사고빈도와 사고율의 관계를 검토한 결과는 다음과 같다. 일반적으로 사고율과 사고빈도에 대한 기대값은 다음과 같이 정의되어 진다.

$$\text{Accident Rate} = \frac{\text{Accident Frequency}}{(\text{Exposure} / \text{Unit of time})} \quad (7)$$

$$E\{\text{Accident Frequency}\} = E\{\text{No. of Accidents on entity} / \text{Unit of time}\} \quad (8)$$

일반적으로 사고모형에 적용하는 사고율은 위에서 제시한 식(7)과 같다. 식에서 알 수 있듯이 사고빈도의 증가량보다 교통량이 증가하게 된다면 사고율은 감소하게 된다.



<그림 2> 교통류와 안전과의 관계(Hauer 1997)

또한 사고빈도의 변화 없이 교통량이 증가하게 되면 사고율은 감소하게 된다. 따라서 사고율로만 안전도를 측정하기에는 무리가 있으며, 이에 대한 실험적 근거는 Mahalel (1986), Brundell-Freij and Ekman (1991), Hauer (1993 & 1995)의 연구에 제시되어 있다. <그림 2>의 A점은 AADT가 2,500대 일 때 연간 0.69건의 사고의 기대값을 가지게 된다. 도로 및 모든 다른 조건은 동일할 때, AADT가 3,000대로 증가하면(점 B) 기울기에 해당하는 사고율은 감소하게 된다 하지만 <그림 2>에서 볼 수 있듯이 사고빈도에 대한 기대값은 증가하였고 이를 바탕으로 안전도가 증가하였다고 평가하기에는 논리적인 비약이 있다. 따라서 교통안전측면의 모형에서는 사고율이 아닌 사고빈도를 사용하는 것이 보다 신뢰도 높은 결과를 얻을 수 있다고 판단 할 수 있다.

## 3. 사고자료의 분포

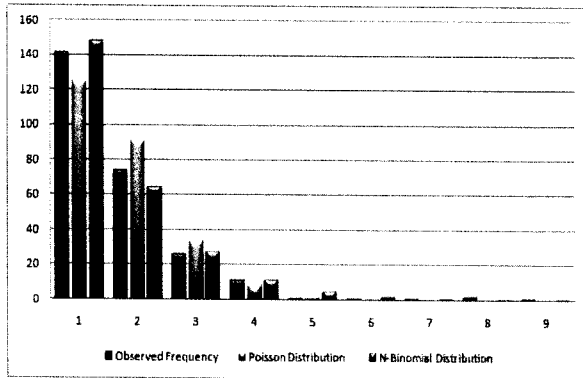
분석대상 트럼펫 인터체인지 129개소, 클로버 인터체인지 35개소의 연결로 전체에서 일어나는 교통사고와 인터체인지 연결로 형태별 직결연결로, 루프연결로, 준직결연결로에서 일어나는 교통사고를 분석했다. 사고건수를 이산적 확률변수로 해석하는 회귀식의 모형 추정을 위해서는 이산형분포의 특성을 분석하여 적절한 분포를 선택하여야 하며, 이때 분포의 선택은 평균(M)과 분산(S<sup>2</sup>)을 비교하여 결정할 수 있다. 일반적으로 교통사고모형추정에 이용될 수 있는 분포의 특성은 다음 <표 2>와 같다.

<표 2> 분석자료 선정 결과

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

(1) 트럼펫인터체인지

총 196건의 사고에 대하여 사고건수별 인터체인지 연결로의 분석결과 <그림 3>과 같다. 또한 <표 3>에서의 분석결과에서도 과분산분포(M=0.7597,  $S^2=1.4284$ )로 음이항분포가 가장 적합한 분포로 나타났다. 평균이 같은 경우의 포아송분포(M=0.7597  $S^2=0.7597$ )를 가상으로 그려본 결과 음이항 분포에 비해 꼬리부분이 짧은 특성을 나타내는데 이는 기준점으로부터 사고의 분포가 많이 흩어져 있다는 것을 의미한다. 또한 사고건수의 분포가 음이항 분포를 따름으로 음이항회귀식으로 모형을 개발하는 것이 타당하며 모형개발 이후의 모형의 적합도 대한 검증에서도 모형의 구조가 과산포를 따름을 보였다.



<그림 3> 트럼펫인터체인지 직결연결로의 사고분포

<표 3> 트럼펫 인터체인지 직결연결로의 음이항분포 적합도 검증

사고건수 ( $X_i$ )	관측값	음이항분포	$\chi^2$ 분포 $\nu$ (자유도) 34-1=33
	교통사고건수 ( $O_i$ )	예측된 사고건수 ( $E_i$ )	
0	141	148	$\chi^2=5.6306$ P-value=0.5530 (fail to reject $H_0$ )
1	74	64	
2	26	27	
3	11	11	
4	1	5	
5	1	2	
6	1	1	
7	2	0	
+8	1	0	
계	258	258	

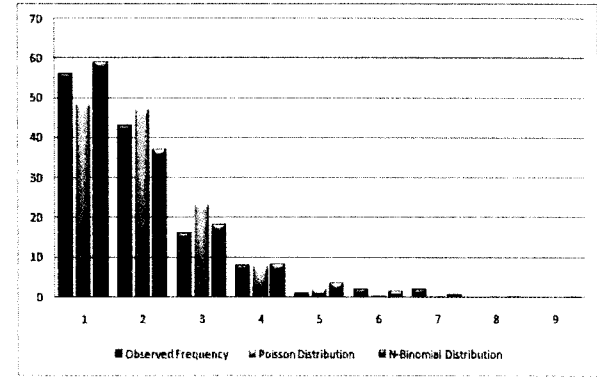
실제 사고건수와 음이항분포 및 포아송분포에 의해 추정된 사고건수는 <표 3>과 같다.

트럼펫 직결연결로에서의 음이항분포가 적합한 확률분포인지 검증하기 위하여 p-Value(유의확률)을 도출하였다. 다음은  $\chi^2$ 검증을 위해 설정한 가설(hypothesis)이다.

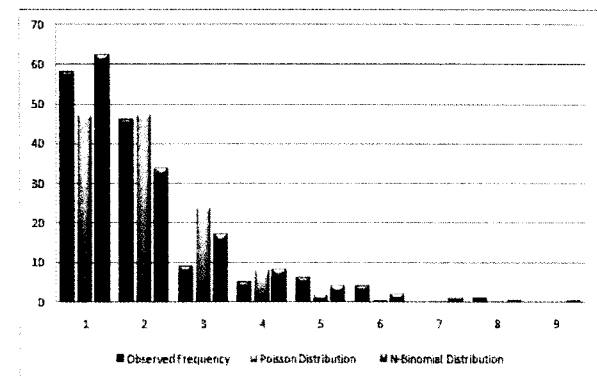
$H_0$  = 사고건수는 음이항분포를 따른다  
 $H_1$  = 사고건수는 음이항분포를 따르지 않는다.

검정통계량 값인 p-Value값이 0.553이므로 귀무가설  $H_0$ 을 따른다. 따라서 유의수준  $\alpha=0.05$ 에서 음이항분포를 따름을 알 수 있으며 음이항회귀식으로 모형을 개발하는 것이 타당하다.

준직결연결로에서 발생한 총 138건의 사고에 대하여 분석결과는 <그림 4>과 같다. <표 4>에서의 분석결과에서도 과분산분포(M=1.0698,  $S^2=2.6435$ )로 직결연결로와 동일한 결과를 보였다.



<그림 4> 트럼펫인터체인지 준직결연결로의 사고분포



<그림 5> 트럼펫인터체인지 루프연결로의 사고분포

<표 4> 트럼펫 인터체인지 준직결/루프연결로의 음이항분포 적합도 검증

사고건수 ( $X_i$ )	준직결연결로			루프연결로		
	관측값	음이항분포	$\chi^2$ 분포	관측값	음이항분포	$\chi^2$ 분포
	교통사고건수 ( $O_i$ )	예측된 사고건수 ( $E_i$ )	$\nu$ (자유도) 34-1=33	교통사고건수 ( $O_i$ )	예측된 사고건수 ( $E_i$ )	$\nu$ (자유도) 34-1=33
0	56	59	$\chi^2_i=4.6811$  P-value=0.3161 (fail to reject $H_0$ )	58	62	$\chi^2_i=13.38306$  P-value=0.2425 (fail to reject $H_0$ )
1	43	37		46	34	
2	16	18		9	17	
3	8	8		5	8	
4	1	3		6	4	
5	2	1		4	2	
+6	2	1		0	1	
계	128	127		128	128	

루프연결로에서 발생한 총 130건의 사고에 대하여 사고건수별 인터체인지 연결로의 분석 결과는 <그림 5>와 같다. 위의 두 연결로와 마찬가지로 과분산분포가 나타남을 <표 4>를 통하여 알 수 있다. 검정통계량 값인 p-Value 값이 0.2425이므로 귀무가설  $H_0$ 을 따름을 알 수 있다. 따라서 유의수준  $\alpha=0.05$ 에서 음이항분포를 따르므로 트럼펫 인터체인의 모든 연결로는 음이항회귀식으로 모형을 개발하는 것이 타당하다고 할 수 있다.

(2) 클로버인터체인지

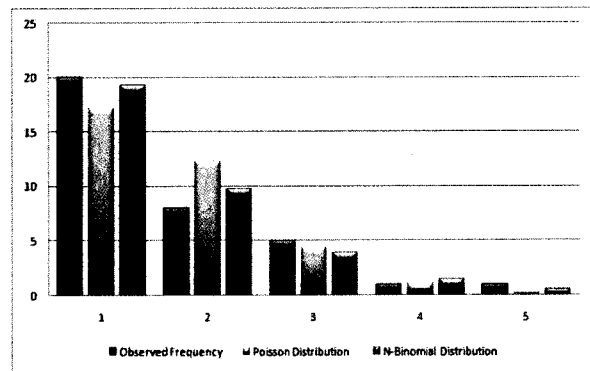
<그림 6>는 클로버인터체인지 직결연결로에서 발생한 총 25건의 사고에 대한 분석결과이다. <표 5>의 분석결과에 따라 음이항분포가 가장 적합한 것으로 나타났다. 따라서 사고건수의 분포가 음이항 분포를 따름으로 음이항회귀식으로 모형을 개발하는 것이 타당하며 모형개발 이후의 모형의 적합도 대한 검증에서도 모형의 구조가 과산포를 따름을 보였다.

루프연결로의 경우 총 19건의 사고에 대한 결과를 분석하였으며, 그 결과는 <표 5>와 같다.

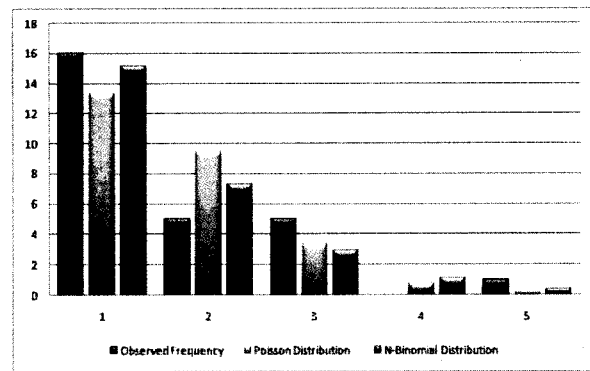
<표 5> 클로버 인터체인지 직결/루프연결로의 음이항분포 적합도 검증

사고건수 ( $X_i$ )	직결연결로			루프연결로		
	관측값	음이항분포	$\chi^2$ 분포	관측값	음이항분포	$\chi^2$ 분포
	교통사고건수 ( $O_i$ )	예측된 사고건수 ( $E_i$ )	$\nu$ (자유도) 34-1=33	교통사고건수 ( $O_i$ )	예측된 사고건수 ( $E_i$ )	$\nu$ (자유도) 34-1=33
0	16	15	$\chi^2_i=2.9714$  P-value=0.3669 (fail to reject $H_0$ )	20	19	$\chi^2_i=0.7026$  P-value=0.3273 (fail to reject $H_0$ )
1	5	7		8	10	
2	5	3		5	4	
3	0	1		1	1	
+4	1	0		1	0	
계	27	26		35	34	

클로버인터체인의 경우 트럼펫인터체인과 마찬가지로 모든 연결로에서 음이항 회귀식으로 모형을 개발하는 것이 타당하다.



<그림 6> 클로버인터체인지 직결연결로의 사고분포



<그림 7> 클로버인터체인지 루프연결로의 사고분포

### III. 모형도출

#### 1. 변수의 설정

고속도로안전과 관련된 기존 연구에서 사용한 변수를 검토한 결과 크게 3가지로 구분할 수 있으며, 도로기하구조요인, 교통 및 환경요인, 운전자 및 차량요인으로 나눌 수 있다. 이중 교통 및 환경요인과 운전자 및 차량요인의 경우 모형적용에 있어 어려움이 많으며, 현실 적용성이 떨어지는 단점이 있다. 본 연구에서는 기존 연구고찰과 기초상관분석을 통해 사고에 미치는 요인이 미약하거나 다중공선성이 있다고 판단되는 변수들은 모형 개발에 사용되지 않았고, 실제 개발된 모형에서도 본 연구에서 선정한 변수들이 교통사고를 설명하는데 의미 있는 변수들이라는 것이 입증되었다. 이를 바탕으로 모형개발에 사용한 종속변수와 독립변수는 아래 <표 6>과 같이 정리 할 수 있다.

기존문헌고찰과 본 연구의 특성에 맞게 제시한 변수의 설정방법에 따라 사고건수(Y), 곡률 또는 곡률(차)(X<sub>1</sub>), 속도차(X<sub>2</sub>), 유출입여부(X<sub>3</sub>), 교통량(X<sub>4</sub>)등의 변수를 사용하여 교통사고 예측모형을 추정하고 추정된 모형의 적합도를 통계적 방법을 이용하여 검증하였다. 트럼펫 인터체인지의 직결연결로의 경우 곡률차가 존재하지 않으므로 곡률을 변수로 사용하였고, 그 이외의 트럼펫 인터체인지와 클로버인터체인지의 루프, 준직결연결로는 연결로에서는 곡률차를 변수로 사용하였다. 유출입여부는 유입연결로를 '1'로, 유출연결로를 '0'으로 Dummy변수처리를 하였으며, 교통량의 경우 Log(AADT)값을 변수로 사용하여 분석하였다.

<표 6> 변수 선정 결과

변수	내용		
dummy	X <sub>3</sub>	램프형태(유출입)	유출램프=0 유입램프=1
continuous	X <sub>1</sub>	곡률, 곡률차	곡선반경의 역수
	X <sub>2</sub>	속도차	고속도로본선과 램프구간의 속도차
	X <sub>4</sub>	교통량(AADT)	4년간의 평균AADT
	Y	사고건수	4년간의 사고지점의 사고건수

#### 2. 트럼펫인터체인지

종속변수인 사고건수와 유입을 '1', 유출을 '0'으로 Dummy처리한 유출입여부, 고속도로본선과 연결로의 설계속도차, 교통량, 곡률 등을 포함한 총 4개의 독립변수의 기초통계값들을 <표 7>에서 분석하였다. 총 129개 트럼펫인터체인지 유입연결로 129개, 유출연결로 129개를 포함한 전체 258개 직결연결로에서 발생한 196건의 사고를 대상으로 분석하였으며, 141개의 미사고 연결로를 포함한 결과이다.

둘 이상의 독립변수들간에 상관관계가 높을 때 이들 변수들 간에 다중공선성이 있다고 하였다. 독립변수간의 높은 다중공선성이 존재할 경우 회귀모형의 결과가 왜곡되는 현상이 나타날 수 있다. 모형전체의 유의성이 통계적 검증을 통해 유의한 것으로 나타나더라도 추정된 계수의 표준오차가 커져 검정결과에서 일부 회귀계수들이 통계적으로 유의하지 않게 나타나는 것이다.

공차한계란  $(1-R_i^2)$ 을 의미하며 VIF는 공차한계의 역수로 표현된다. 여기서 공차한계 0.10 이하, VIF값 10이상이면 공선성이 높다고 본다. <표 9>에서 알 수 있듯이 독립변수들간의 다중공선성에 문제가 없다고 판단된다.

모형개발결과에 의한 예측치와 관측치의 차(residual)를 비교하는 통계적 방법을 이용, 모형의 적합도를 분석하였다. <그림 8>에서 예측치와 관측치의 차의 제곱한 값을 점으로 찍었을 때의 값들과 식 (9)의 variance 함수에 의해 추정되는 곡선과의 비교 결과 점들이 곡선을 따르므로 모형의 적합도는 높다고 할 수 있다.

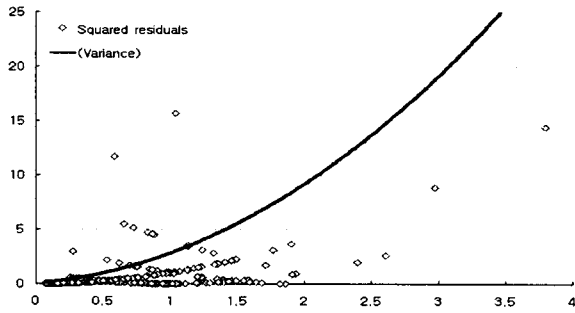
$$Var(Y)_{CDi} = E(Y)_{CDi} + \frac{E(Y)_{CDi}^2}{\kappa} \quad (9)$$

$E(Y)_{CDi}$  : 직결연결로에서의 어떤 지점에서 예측되는 교통사고건수

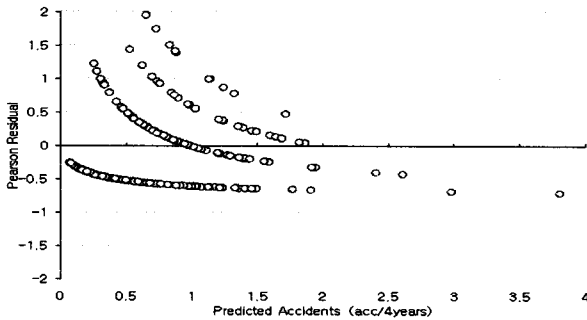
$\kappa$  : parameter

<표 7> 트럼펫 인터체인지 직결연결로의 기초통계분석

변수	종속변수		독립변수		
	사고건수	유출입여부	설계속도차	교통량	곡률
표본수	258	258	258	258	258
평균	0.7597	0.5000	54.6674	3.3822	0.0091
표준편차	1.1952	0.5010	7.4389	0.5091	1.1952
합계	196.0000	129.0000	14104.1812	872.6024	2.3377
최소값	0.0000	0.0000	30.0000	2.3012	0.0005
최대값	8.0000	1.0000	82.4819	4.9054	0.0290



<그림 8> 트럼펫 I.C. 직결연결로의 Estimated variance and squared residuals



<그림 9> 트럼펫 I.C. 직결연결로 Pearson residual

식 (10)에서는 예측된  $E(Y)$ 를 축으로 하여 <그림 9>를 통해 구해진  $PR_i$ 값들(Pearson Residual)을 점으로 찍었을 때 '0'에 대칭에 가까우므로 모형의 적합도는 높다고 할 수 있다.

$$Pearson\ Residual(PR_i) = \frac{E(Y)_{CD_i} - y_i}{\sqrt{var(y_i)}} \quad (10)$$

$y_i$  : 연결로에서의 어떤 지점  $i$ 에서 관측되는 교통 사고건수  
 $var(y_i)$  : 직결연결로에서의 어떤 지점  $i$ 에서 관측되는 교통사고건수의 분산

SAS의 PROC GENMOD에서 제공하는 임의의 parameter  $\kappa$  0.5532 으로 <표 10>에서 1에 가까운 deviance/(n-p)값을 얻을 수 있다. 또한 Pearson chi-square 통계값을 n-p로 나누었을 때의 값이 1에 아주 근접하고 0.8~1.2 사이의 값을 보이므로 모형은 매우 적합 할 수 있다.

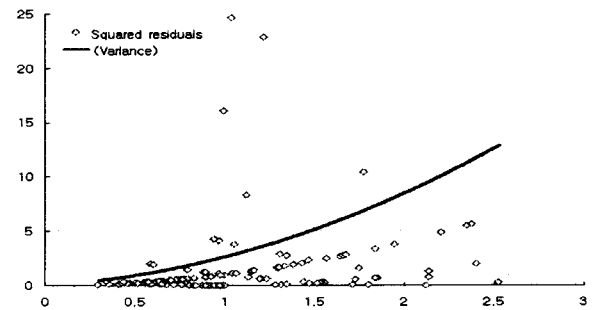
변수들의 유의성을 검증하기 위하여 Chi-square likelihood ratio 통계값 분석결과 곡률

<표 8> 트럼펫 인터체인지 준직결연결로의 기초통계분석

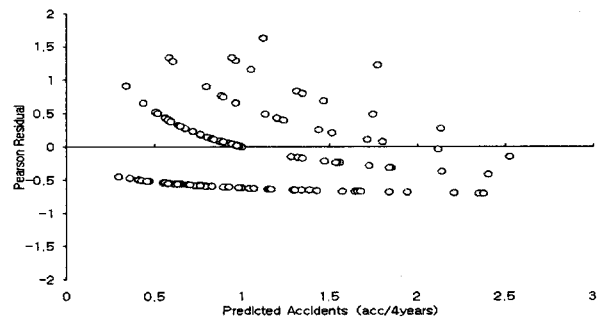
변수	종속변수	독립변수			
	사고건수	유출입여부	입계속도차	교통량	곡률
표본수	129	129	129	129	129
평균	1.0698	0.5116	61.6061	3.4602	0.7840
표준편차	1.6259	0.5018	6.6142	0.4689	0.3258
합계	138.0000	66.0000	7947.1871	446.3720	101.1350
최소값	0.0000	0.0000	37.3792	2.2076	0.1293
최대값	13.0000	1.0000	79.6861	4.5061	1.7969

( $X_1$ )과 유출입여부( $X_3$ )가 가장 유의성있는 변수로 분석되었다. 독립변수 모두 5%의 유의수준을 만족하는 결과를 보이고 있다.

129개 준직결연결로에서 발생한 138건의 사고의 경우, 56개의 미사고연결로를 포함하여 분석한 결과 다중공선성에 문제가 없으며, 유출입여부( $X_3$ )와 곡률차( $X_4$ )가 가장 유의성있는 변수로 분석되었다. 유의수준이 5%를 만족하지 못하는 변수가 존재 하기는 하나, 모두 고속도로 인터체인지 연결로를 설명하는데 필요한 변수라 판단하였기 때문에 모형 추정에 이용하였다.



<그림 10> 트럼펫 I.C. 준직결연결로의 Estimated variance and squared residuals



<그림 11> 트럼펫 I.C. 준직결연결로 Pearson residual

<표 9> 인터체인지 형식별, 연결로 유형별 다중공선성 분석 결과

변수	트럼펫 인터체인지						콜로버 인터체인지			
	직결연결로		준직결연결로		부프연결로		직결연결로		부프연결로	
	tolerance	VIF	tolerance	VIF	tolerance	VIF	tolerance	VIF	tolerance	VIF
램프유형	0.881	1.135	0.949	1.054	0.963	1.039	0.985	1.005	0.947	1.056
속도차	0.645	1.551	0.691	1.446	0.981	1.020	0.863	1.158	0.774	1.292
교통량(AADT)	0.978	1.023	0.962	1.040	0.967	1.034	0.860	1.163	0.840	1.191
곡률(차)	0.598	1.671	0.6896	1.451	0.985	1.015	0.754	1.326	0.915	1.093

<표 10> 트럼펫인터체인지 직결연결로 및 준직결연결로의 모형추정 결과

Parameter	직결연결로			준직결연결로		
	Estimate (β)	Chi-square	Pr>Chi-square	Estimate (β)	Chi-square	Pr>Chi-square
Intercept	-0.7918 (β <sub>0</sub> )	0.67	0.4123	-2.7475 (β <sub>0</sub> )	2.54	0.1107
곡률(차) (X <sub>1</sub> )	138.8968 (β <sub>1</sub> )	27.46	<.0001	0.1101 (β <sub>1</sub> )	0.06	0.8028
속도차 (X <sub>2</sub> )	-0.0549 (β <sub>2</sub> )	11.89	0.0006	-0.0021 (β <sub>2</sub> )	0.01	0.9273
유출입여부(ON/OFF) (X <sub>3</sub> )	1.3379 (β <sub>3</sub> )	31.37	<.0001	-0.5542 (β <sub>3</sub> )	5.40	0.0201
LOG(AADT) (X <sub>4</sub> )	0.3906 (β <sub>4</sub> )	0.9764	0.0263	0.8795 (β <sub>4</sub> )	10.23	0.0014
Dispersion	0.5532			0.6165		
Goodness of fit measures	DF	Value	Value/DF	DF	Value	Value/DF
Deviance	253	213.9745	0.8455	124	125.8923	1.0153
Scaled Deviance	253	217.0049	0.8579	124	118.9764	0.9595
Pearson chi-square	253	249.3510	0.9856	124	131.2079	1.0581
Scaled Pearson X <sup>2</sup>	253	253.0000	1.0000	124	124.0000	1.0000
Log Likelihood		-230.4715			-97.6229	
P-Value		0.5530			0.3161	

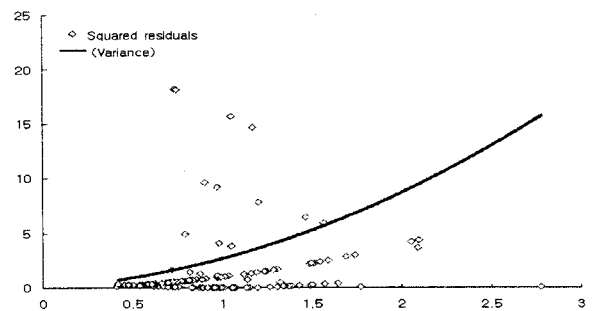
루프연결로의 경우 129개 연결로에서 발생한 130건의 사고와 58개의 미사고연결로를 포함하여 분석하였으며 가장 유의성 있는 변수로는 교통량(X<sub>4</sub>)으로 분석되었다. 교통량 이외의 변수의 경우 유의수준은 떨어지나 변수 선정 당시 고속도로 인터체인지 연결로를 설명하는데 꼭 필요한 변수라 판단하였기에 본 연구에서는 모형추정에 이용하였다.

<표 11> 트럼펫 인터체인지 루프연결로의 기초통계분석

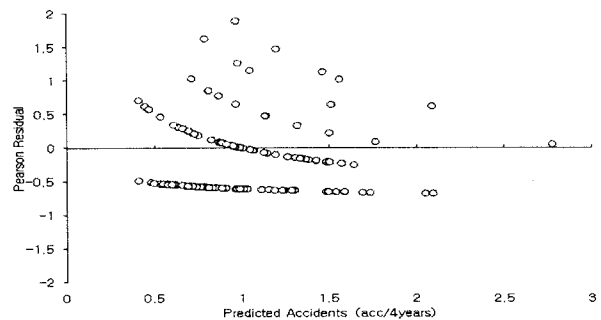
변수	종속변수		독립변수			
	사고건수	유출입여부	설계속도차	교통량	곡률	
표본수	129	129	129	129	129	
평균	1.0078	0.4884	67.7328	3.2806	0.4153	
표준편차	1.3721	0.5018	7.0731	0.5011	0.2984	
합계	130.0000	63.0000	8737.5348	423.2014	53.5714	
최소값	0.0000	0.0000	46.5659	2.3560	0.0000	
최대값	7.0000	1.0000	85.4498	4.6467	17.7501	

<표 12> 트럼펫인터체인지 루프연결로의 모형추정 결과

Parameter	루프연결로		
	Estimate(β)	Chi-square	Pr>Chi-square
Intercept	-4.2834 (β <sub>0</sub> )	7.31	0.0069
곡률(차) (X <sub>1</sub> )	0.3803 (β <sub>1</sub> )	1.05	0.3044
속도차 (X <sub>2</sub> )	0.0259 (β <sub>2</sub> )	2.28	0.1308
유출입여부(ON/OFF) (X <sub>3</sub> )	0.1516 (β <sub>3</sub> )	0.38	0.5353
LOG(AADT) (X <sub>4</sub> )	0.6809 (β <sub>4</sub> )	7.99	0.0047
Dispersion	0.5934		
Goodness of fit measures	DF	Value	Value/DF
Deviance	124	128.5274	1.0365
Scaled Deviance	124	118.3849	0.9547
Pearson chi-square	124	134.6235	1.0857
Scaled Pearson X <sup>2</sup>	124	124.0000	1.0000
Log Likelihood		-103.0281	
P-Value		0.2425	



<그림 12> 트럼펫I.C. 루프연결로의 Estimated variance and squared residuals



<그림 13> 트럼펫I.C. 루프연결로 Pearson residual

### 3. 클로버인터체인지

총 35개 트럼펫인터체인지중 유입연결로 18개, 유출연결로 17개의 직결연결로에서 발생한 25건의 사고와 20개의 미사고연결로를 포함하여 분석하였다. 가장 유의성있는 변수로는 유출입여부(X<sub>3</sub>)로 분석되었다.

또한 총 35개 클로버인터체인지 27개 루프연결로에서 발생한 19건의 사고와 16개 미사고연결로를 대상으로 분석한 결과 가장 유의성있는 변수로는 교통량(X<sub>4</sub>)으로 분석되었으며, 모든 변수를 모형 추정에 이용하였다.

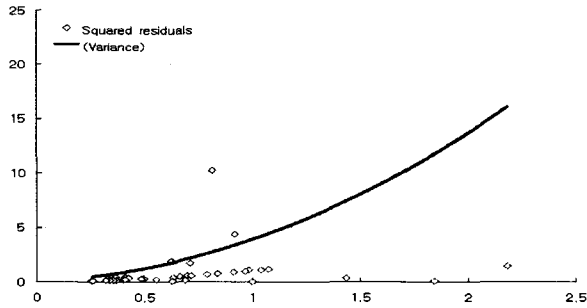


<표 13> 클로버 인터체인지 직결연결로의 기초통계분석

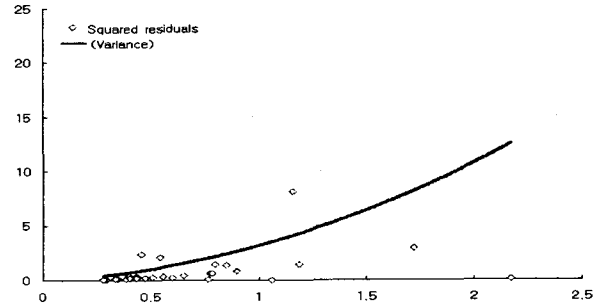
변수	종속변수		독립변수		
	사고건수	유출입여부	설계속도차	교통량	곡률
표본수	35	35	35	35	35
평균	0.7143	0.5143	52.3057	3.8870	0.0053
표준편차	1.0167	0.5071	7.6508	0.9960	0.0032
합계	25.0000	18.0000	1830.6996	30.0000	66.5002
최소값	0.0000	0.0000	30.0000	2.6132	0.0022
최대값	4.0000	1.0000	66.5002	6.2038	0.0168

<표 14> 클로버 인터체인지 루프연결로의 기초통계분석

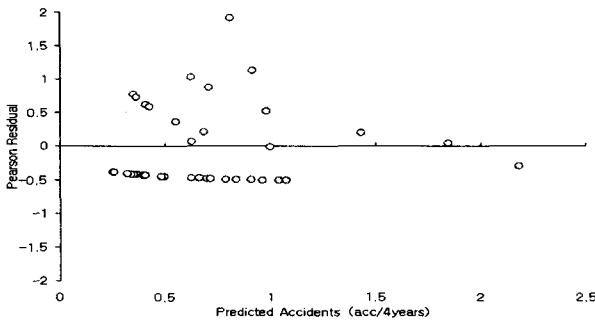
변수	종속변수		독립변수		
	사고건수	유출입여부	설계속도차	교통량	곡률
표본수	27	27	27	27	27
평균	0.7037	0.4444	71.0709	3.9567	0.2330
표준편차	1.0309	0.5064	5.3508	1.0231	0.2953
합계	19.0000	12.0000	1918.9155	106.8300	6.2908
최소값	0.0000	0.0000	60.9925	2.6368	0.0000
최대값	4.0000	1.0000	80.4391	6.1632	1.3078



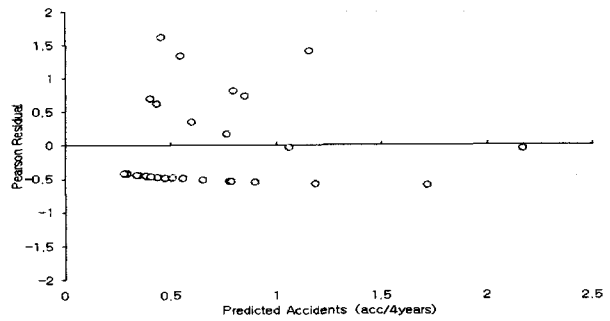
<그림 14> 클로버 I.C. 직결연결로의 Estimated variance and squared residuals



<그림 16> 클로버 I.C. 루프연결로의 Estimated variance and squared residuals



<그림 15> 클로버 I.C. 직결연결로 Pearson residual



<그림 17> 클로버 I.C. 루프연결로 Pearson residual

<표 15> 클로버 인터체인지 직결연결로 및 루프연결로에서의 모형추정 결과

Parameter	직결연결로			루프연결로		
	Estimate (β)	Chi-square	Pr>Chi-square	Estimate (β)	Chi-square	Pr>Chi-square
Intercept	-0.5332 (β <sub>0</sub> )	0.07	0.7889	-3.7069 (β <sub>0</sub> )	0.42	0.5148
곡률(차) (X <sub>1</sub> )	50.6039 (β <sub>1</sub> )	0.40	0.5277	-0.0975 (β <sub>1</sub> )	0.01	0.9323
속도차 (X <sub>2</sub> )	-0.0145 (β <sub>2</sub> )	0.15	0.6971	0.0183 (β <sub>2</sub> )	0.07	0.7897
유출입여부(ON/OFF) (X <sub>3</sub> )	-0.8229 (β <sub>3</sub> )	2.69	0.1010	0.5909 (β <sub>3</sub> )	0.97	0.3244
LOG(AADT) (X <sub>4</sub> )	0.2505 (β <sub>4</sub> )	1.10	0.2942	0.4252 (β <sub>4</sub> )	2.29	0.1304
Dispersion	0.3419			0.4558		
<b>Goodness of fit measures</b>	<b>DF</b>	<b>Value</b>	<b>Value/DF</b>	<b>DF</b>	<b>Value</b>	<b>Value/DF</b>
Deviance	30	34.1230	1.1374	22	25.8535	1.1752
Scaled Deviance	30	31.1225	1.0374	22	24.0674	1.0940
Pearson chi-square	30	32.8923	1.0964	22	23.6327	1.0742
Scaled Pearson X <sup>2</sup>	30	30.0000	1.0000	22	22.0000	1.0000
Log Likelihood		-27.0864			-21.1321	
P-Value		0.3273			0.3669	

최종모형식의 형태에 추정된 변수의 계수값 (X<sub>4</sub>),을 의미한다.

(β<sub>0</sub>, β<sub>1</sub>, β<sub>2</sub>, β<sub>3</sub>, β<sub>4</sub>)를 대입하면 다음과 같다. 여기서 E(A)는 연결로에서의 모형추정에 의한 예측 사고건수를 의미하고 AADTramp는 교통량

$$E(A) = \exp(\beta_0) \cdot (\text{AADT}_{\text{ramp}})^{\beta_4} \cdot \exp(\beta_1 X_1) \cdot \exp(\beta_2 X_2) \cdot \exp(\beta_3 X_3)$$

### 3) 모형의 적합도 분석

일반적으로 이와 같은 모형을 검증하는 방법은 2가지가 있다. 그 첫 번째는 데이터구축에 사용하지 않은 다른 데이터군을 사용하여 구축된 모형을 검증하는 방법이고, 다른 한가지 방법은 포아송 분포를 기반으로 한 모형과 비교하는 방법이다. 본 연구에서는 두 번째 방법을 통하여 모형의 적합도를 분석하였으며, 그 내용은 다음과 같다. 포아송분포와 음이항분포를 통하여 도출된 모형의 p-Value(유의확률) 값을 비교하여 모형의 적합도를 판단할 수 있다. 모형의 유의수준을 5%의 유의수준으로 검정을 할 때, p-Value는 0.05보다 클 경우 귀무가설  $H_0$ 을 지지한다고 한다. 이때의 가정은 다음과 같다.

$H_0$  = 관측된 자료들이 특정분포를 따른다.  
 $H_1$  = 관측된 자료들이 특정분포를 따르지 않는다.

음이항분포모형의 경우 귀무가설( $H_0$ )은 관측된 자료들이 음이항분포를 따르는 것이 되고, 포아송분포모형의 경우 귀무가설이 관측된 자료들이 포아송분포를 따르는 것을 의미한다. 각 분포모형의 유의확률은 SAS결과에는 도출되지 않으며, 본 연구에서는 도출된 결과값 중 Pearson Chi-Square의 Value와 Degree of Freedom(DF)을 이용하여 MATLAB 7.1에서 산출하였다. 그 결과값은 다음 <표16>과 같다.

음이항분포모형의 유의확률은 5개 연결로 모두 0.05이상으로 '음이항분포를 따른다'는 귀무가설을 만족한다. 하지만 포아송분포의 경우 클로버 인터체인지의 루프 연결로를 제외하고는 모두 유의확률값이 도출되지 않았으며, 따라서 '포아송분포를 따르지 않는다'는 대립가설을 만족하므로 본연구에서 음이항분포로 도출된 모형의 적합도가 높다고 판단할 수 있다.

<표 1> 음이항분포와 포아송분포의 P-value 비교

구 분		P-value of NBRM	P-value if PDM
트럼펫 I.C.	직결연결로	0.5530	0.0000
	준직결연결로	0.3161	0.0000
	루프연결로	0.2425	0.0000
클로버 I.C.	직결연결로	0.3401	0.0000
	루프연결로	0.3273	0.0870

## IV. 요약 및 결론

본 연구는 고속도로의 안전에 영향을 미치는 설계요소를 바탕으로 음이항분포를 이용한 사고예측모형을 도출하였다. 기하구조, 운전자 및 차량특성 사이에 상호작용은 교통안전측면에 있어 상당히 복잡한 상호관계를 맺고 있다. 또한 도로의 기하구조 및 차량의 성능이 지속적으로 향상되고 있으며, 이에 따라 모형 도출을 위한 통제가 힘들다. 이러한 한계를 극복하기 위하여 본 연구에서는 운전자의 특성, 기상 조건, 도로포장상태 등의 변수를 제외하고 설계 및 계획단계에서 수집 가능한 정보를 최소한으로 이용하여 모형을 도출하였다.

기존에 주로 사용되었던 포아송분포모형은 평균과 분산이 동일하다는 기본가정을 포함하고 있으나, 사고형태의 경우 과분산현상이 발생하므로 이와 같은 사항을 보완하기 위하여 본 연구에서는 감마분포를 따르는 분석자료의 특성에 따라 음이항분포를 이용한 모형을 도출하였다.

## 참고문헌

- 김태영, 박병호(2007), 트럼펫 IC램프의 운전조건과 교통 사고 분석, 대한교통학회지 제 25권, 제1호, pp. 73~79
- 박병호, 류승욱(2002), 트럼펫 인터체인지의 형식과 교통 사고율의 관계, 한국 지역개발 학회지 제 14권, 제1호, pp. 107~124
- 박효신(2006), 고속도로 인터체인지 연결로에서의 교통사고 예측모형 개발, 연세대학교
- 윤병조, 오영태, 이승환, 지동한(2006), 트럼펫 IC형식 연결로 교통사고 특성분석에 관한 연구, 대한교통학회지 제24권, 제7호, pp.41~51
- Bauer K. M. and D. W. Harwood (1997), Statistical Models of accidents in Interchange Ramp and Speed-Change Lanes, FHWA-RD-97-106
- Brundell-Freij, K. and Ekman, L., (1991), Flow and safety. Paper 910006, presented at the 70th Annual Meeting of the TRB
- Hauer, E., (1997), Observation before-after studies in road safety, Pergamon
- Ivan J. (2004), A new approach for including traffic volumes in crash rate analysis and forecasting, Annual Meeting of the TRB
- Mahalel, D. (1986), A note on accident risk. Transportation Research Board 1068, pp.85-89