

■ 論 文 ■

EB기법을 이용한 사고찾은 곳 개선사업 우선순위 판정기법 개발

Development of Evaluation Model for Black Spot Improvement Priorities
by using Emperical Bayes Method

정성봉

(한국교통연구원 국가교통조사분석사업단 책임연구원)

황보희

(한국교통연구원 도로교통연구실 연구원)

성낙문

(한국교통연구원 도로교통연구실 연구실장)

이선하

(공주대학교 건설환경공학부 교수)

목 차

- I. 서론
 - 1. 연구의 배경 및 목적
 - 2. 연구의 범위 및 방법
- II. 선행연구 고찰
 - 1. 사고예측모형 관련 선행연구고찰
 - 2. 교통사고 잦은 지점 관련 선행연구고찰
- III. 이론적 고찰
 - 1. Empirical Bayes Method
- IV. 자료수집 및 모형개발
 - 1. 자료수집
 - 2. 모형개발
- V. 결론 및 향후 연구과제
참고문헌

Key Words : EB 방법, 포아송회귀모델, 음이항회귀모델, 교통사고 잦은 지점, 개선우선순위
Empirical Bayes Method, Poisson Regression Model, Negative Binomial Regression Model, Black Spot, Improvement Priority

요 약

도로 네트워크의 안전 관리는 기본적으로 교통안전 조사를 위한 사이트(교통사고 잦은 지점) 선정, 안전문제에 대한 진단, 잠재적 위험요소들에 대한 가능한 대안 선정, 주어진 예산 제약 하에 대안간의 우선순위 결정과 같은 순서로 이루어진다(Persaud, 2001). 효율적인 안전 조사를 위해 요구되는 지점(교통사고 잦은 지점) 선정에 대한 과정은 매우 중요하다. 이에, 본 연구는 인천시 4지 신호 교차로 중 교통사고 잦은 지점으로 선정된 지점의 3년간(2004~2006년) 사고 자료와 기하구조 자료를 이용하여 EB 기법 이용 시 필요한 사고예측모형을 개발하였다. 또한, 교통사고 잦은 지점을 선정하는데 있어 현재 우리나라에서 적용되고 있는 단순사고건수와 심각도를 고려하여 선정된 우선순위와 단순사고건수 대신 EB 기법을 이용하여 예측된 사고건수를 이용하여 선정된 우선순위 비교를 통해 기존 방법의 한계를 제시하고 합리적인 예측모형 개발 필요성을 보여주고자 하였다. 분석 결과, 총 사고건수 추정 시 기존의 사고 예측 방법인 비선형 회귀모형과 EB 기법의 결과 값 모두 예측력이 높은 것으로 나타났지만 지점별 사고건수 예측력을 함께 고려할 경우엔 EB 기법이 비선형 회귀 모형(포아송)의 결과보다 예측력이 좋은 것으로 나타났다. 또한, 도출된 우선순위 비교 결과 대부분의 지점의 우선순위는 크게 변동이 없었으나, 서해4거리 등 몇 개 지점의 개선우선순위에는 상당한 변동이 발생하는 것으로 나타났다. 이는 실제 사고건수를 이용할 경우 발생하는 RTM문제를 본 연구에서 제안한 기법을 사용할 경우 해결가능하다는 것을 시사한다.

The safety management of a road network comprises four basic inter-related components : identification of sites(black spot) requiring safety investigation, diagnosis of safety problems, selection of feasible treatments for potential treatment candidates, and prioritization of treatments given limited budgets(Persaud, 2001). Identification process of selecting black spot is very important for efficient investigation of sites. In this study, the accident prediction model for EB method was developed by using accident data and geometric conditions of black spots selected from four-leg signalized intersections in In-cheon City for three years (2004-2006). In addition, by comparing the rank nomination technique using EB method to that by using accident counts, we managed to show the problems which the existing method have and the necessity for developing rational prediction model. As a result, in terms of total number of accidents, both the counts predicted by existing non-linear regression model and that by EB method have high good of fitness, but EB method, considering both the accident counts by sites and total number of accident, has better good of fitness than non-linear poisson model. According to the result of the comparison of ranks nominated for treatment between two methods, the rank for treatment of almost sites does not change but SeoHae intersection and a few other intersections have significant changes in their rank. This shows that, with the technique proposed in the study, the RTM problem caused by using real accident counts can be overcome.

본 연구는 국토해양부 교통체계효율화사업(06교통핵심C01)의 연구비 지원에 의해 수행되었습니다.

I. 서론

1. 연구의 배경 및 목적

도로 네트워크의 안전 관리는 기본적으로 교통안전 조사를 위한 사이트(교통사고 잦은 지점) 선정, 안전문제에 대한 진단, 잠재적 위험요소들에 대한 가능한 대안 선정, 주어진 예산 제약 하에 대안간의 우선순위 결정과 같은 순서로 이루어진다. 이 중 교통안전 조사를 위한 사이트(교통사고 잦은 지점)의 경우는 안전 관리를 위한 첫 번째 단계인 동시에 가장 중요한 과정 중 하나이다.(Persaud, 2001)

국내 교통사고 잦은 지점 선정 시 과거에 발생한 교통사고 발생건수가 중요한 지표로 활용되고 있으나, 관측 기간 동안 특정지역에 임의적으로 사고가 대폭 증가하였을 경우 이를 반영하지 못하는 관계로 사이트를 선정하는 방법으로는 바람직하지 못하다는 평가를 받고 있다.

이에, 최근에는 다양한 사고예측모형 개발을 통해 사고 위험을 측정하고자 하는 노력이 이루어지고 있으나 이 또한 사고의 임의성과 모형의 예측력 부족 등으로 적용하기에는 많은 한계가 존재한다. 기존 사고예측모형은 모형의 적합도가 그다지 높지 않으며 또한, 다양한 요인에 의해서 교통사고가 일시적으로 대폭 증가 혹은 감소했을 때 이를 반영하지 못한다는 단점이 있다. 이는 최종적으로 교통안전개선 사업 평가 시 효과를 지나치게 과대 혹은 과소평가하는 결과를 초래할 수 있다(RTM Bias)¹⁾. 이러한 교통사고의 불확실성으로 인한 도로안전도 평가의 왜곡현상은 교통특성 및 도로의 구조가 비슷한 유사지점(Reference sites)의 교통사고 현황을 이용하여 보정할 수 있는데, 이때 사용되는 개념이 Bayesian 방법이다. 특히, Hauer의 베이지언 방법은 어떤 지점에서 발생한 교통사고는 교통사고의 여러 가지 요인들로 인해 변동적일 수 있으므로 Reference sites에서 발생한 교통사고를 토대로 보정되어야 한다는 이론이다. 이 방법은 교통사고 예측 시 노출도인 교통량을 고려할 수 있으며, 교통사고의 불확실성을 반영할 수 있다는 측면에서 기존의 사고예측모형보다 우월한 것으로 알려져 있다. 또한 교통사고의 Reference sites와 관련된 이론적 혹은 적용상의 문제점을 사고예측모형을 이용함으로써 문제점을 해결하고 있다.

본 연구에서는 인천시 4지 신호 교차로 지점 중 3년(2004년~2006년) 동안 교통사고 잦은 지점으로 선정된 교차로를 대상으로 EB(Empirical Bayes) Method 적용 시 필요한 사고예측모형을 개발하고 교통사고 잦은 지점을 선정하는데 있어 현재 우리나라에서 적용되고 있는 단순사고건수와 심각도를 고려하여 선정된 우선순위와 단순사고건수 대신 EB 기법에 의해 예측된 사고건수를 이용하여 선정된 우선순위 비교를 통해 기존 방법의 한계를 제시하고 합리적인 예측모형 개발 필요성을 보여주고자 한다.

2. 연구의 범위 및 방법

본 연구에서는 첫 번째 단계로 문헌고찰을 통하여 전통적인 사고예측방법론에 대한 문제점을 파악하고, EB 방법을 적용하기에 필요한 사고예측모형에 대해 심층적으로 분석하였다. 두 번째 단계로 공간적 범위로는 인천시 사고 잦은 지점 중 4지 신호 교차로를 대상으로 사고에 영향을 미치는 교통조건, 도로조건, 환경조건과 같은 변수를 선정하고 현장조사를 실시하여 각 교차로의 변수를 수집하였다. 세 번째 단계로 두 번째 단계에서 선정된 Reference sites의 사고이력자료 및 현장조사를 통해 구축된 변수를 바탕으로 사고 자료와의 상관관계를 분석하고 영향력 있는 변수를 추출하여, 이 변수를 적용하여 비선형 회귀분석을 이용한 사고예측모형(P)을 개발하였다. 마지막으로 EB방법을 이용하여 사고건수를 추정하여 실제 사고건수와 비교하고 기존의 우선순위 선정 시 사고건수에 대한 가중 값을 EB방법에 의해 추정된 사고건수 가중 값으로 계산하여 기존의 우선순위와 비교해 보았다.

II. 선행연구고찰

1. 사고예측모형 관련 선행연구고찰

본 절에서는 일반적으로 사고예측 시 적용되고 있는 모형인 다중회귀모형, 음이항모형, 그리고 포아송모형을 이용한 기존 연구사례를 검토한다.

지금까지 사고건수예측을 위해 많이 활용된 모형으로

1) 통계학적으로는 표본의 수가 커질수록 표본의 평균값은 모집단의 평균값과 유사해진다는 의미를 갖는다. 다시 이야기하면, 어떤 주어진 연도에 교통사고의 불확실성으로 교통사고가 일시적으로 증가 혹은 감소하는 현상이 발생하지만 장기적으로는 평균값에 수렴하는 현상을 말한다.

다중회귀모형을 들 수 있다. 이 모형은 종속변수의 변화를 설명하기 위하여 두 개 이상의 독립변수가 사용되는 선형회귀모형을 말하며, 실제 자료를 이용한 모형 적합의 경우에는 대부분이 중회귀모형을 이용한 모형적합이 이용된다. Scott E. Davis, H. Douglas Robertson은 다중회귀분석을 사용하여 개발한 시애틀 3그룹모델은 상충수, 보행교통량, 교통량, 차선수를 독립변수로 하여 개발하였다.

하태준(2001)은 모형개발에 앞서 사고유형별 발생비율을 단순통계분석을 통해 정리하였으며, 또한 4지 신호교차로 교통사고 예측모형을 개발하기 위해 차대차 교통사고를 18가지 유형으로 정의하였다. 그리고 모형식을 도출하기 위해 자료분석 기법을 이용하여 다중회귀분석을 하였고, 사고유형에 따른 교통류의 형태를 분류하여 사고와 관련이 없는 교통류는 제외시키고 교통류의 독립성과 상호작용 효과를 분석하였다.

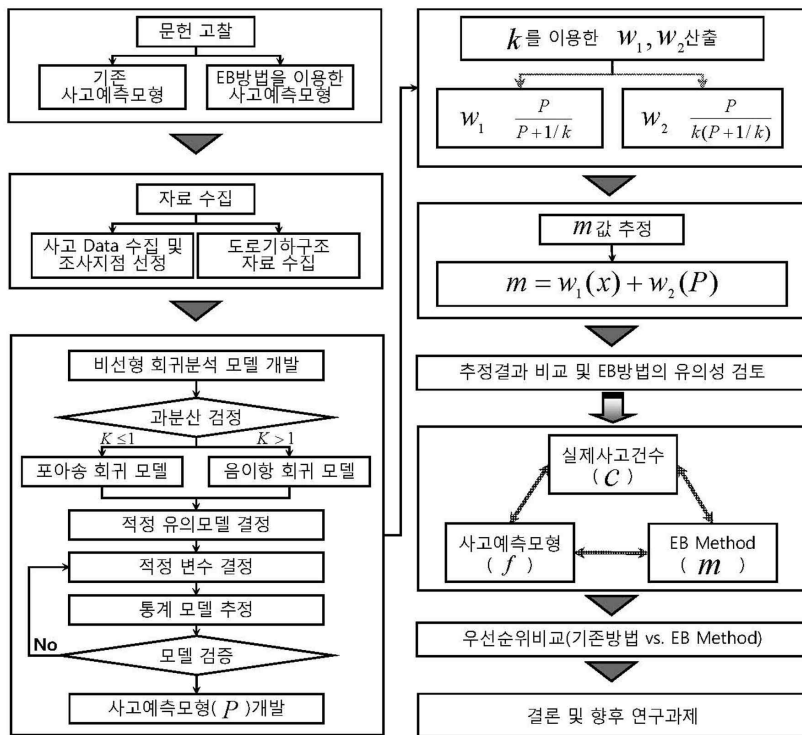
김원철(2000)은 수량화 1류를 이용하여 구축한 사고 예측모형은 독립변수를 교통량, 주행차량속도, 도로의 종단구배, 교통섬 유무로 정하였으며, 이 독립변수들의 각각의 경우를 조합하여 변수들간의 상대적 중요도, 아

이템과 카테고리의 수량, 아이템의 범위, 중상관계수 등으로 사고예측모형을 구축하였다.

강민욱(2002)는 호남고속도로의 전체 구간, 1996~2000년까지의 5년 간 교통사고 자료를 이용하여 고속도로 본선에서 발생하는 교통사고와 기하구조와의 관계를 분석하였으며 이를 통해 고속도로 구간에서 구간분할법과 최우추정법을 이용한 음이항 회귀 모델을 개발하였다.

이재명(2004)은 교통사고 요인들 간의 상호교호작용을 고려한 분석 방법인 CART(Classification Regression Tree)분석을 이용하여 교통사고발생도표(Traffic Accident Diagram)를 작성하여 사고 예측 모형을 제시 하였으며, 기존의 선형회귀분석모형 및 포아송 모형과 비교 검토하여 연구에서 제시된 모형의 적용 타당성을 평가하였다.

하오근(2005)의 연구에서는 신호 교차로의 도로시설 조건, 교통 환경조건, 주변 환경조건 등을 분석하여 안전성에 방해가 되는 요소들을 찾아내고, 그 요소들과 사고와의 상관관계를 이용하여 교차로 교통사고에 영향을 미치는 요소들을 추출하여 비선형 회귀분석 중 포아송 회귀 모형을 이용하여 우리나라 교차로 실정에 맞는 사고 예측모형을 개발하였다.



<그림 1> 연구의 흐름도

2. 교통사고 잦은 지점 관련 선행연구고찰

교통사고 잦은 지점을 찾기 위한 방법에 초점을 맞춘 상당히 포괄적인 논문이 존재한다. 사고 수나 빈도에 기초한 방법에 대한 연구논문(Deacon et al., 1975), 사고율(AR)과 rate-quality control을 적용한 논문도 있으며(Norden et al., 1956; Rudy, 1962; Morin, 1967; Stokes and Mutabazi, 1996), 사고빈도와 사고율을 함께 고려한 논문도 있다(Laughland et al., 1975).

그 중 Abbess and Wright(1981), Hauer 외 (1986, 1987, 1988, 1991), Hagle and Witkowski(1988)은 사고다발지역 우선순위 선정방법과 관련하여 연구하였으며, 이들 중에는 RTM오류를 극복하기 위해 Empirical Bayes(EB) 방법의 사용을 제안해 왔다. 이 방법은 특정지역에서의 사고발생 이력과 유사지역에서의 예상되는 사고 자료를 결합하는 방법으로 어쩌면 발견되지 않았을지도 모를 안전 문제를 설명할 수 있다는 장점이 있다. EB 기법의 적용은 철도와 고속도로 교차점, 신호교차로, 운전자 등의 안전에 대한 예측을 포함하고 있다.

Persaud and Hauer (1984)는 before-and-after 분석의 단점을 보완하기 위해 EB방법과 비모수 방식을 비교 및 평가하였다. 여러 데이터의 분석을 통해 대부분의 경우 Bayesian 방식이 우수하다는 결론을 냈다.

Cheng과 Washington (2005)에 의한 많은 연구들은 실험상에서 교통사고 잦은 지점을 선정하는 방법들을 비교하고 최적의 사고이력 지속시간을 찾기 위해 Arizona 주에서 얻은 여러 가지 실제 경험상의 사고 분포 및 여러 단계의 사고 이질성을 이용하여 분석하였다.

III. 이론적 고찰

1. Empirical Bayes Method

베이저언 이론은 사전분포(Prior distribution)의 불확실성을 사후분포(Posterior distribution)를 통하여 보정할 수 있다는 논리를 기반으로 하고 있다. 이는 다시 말하면 위험도를 가진 어떤 특정 지점은 과거의 사고기록 뿐만 아니라 그와 유사한 지점들의 집단인 Reference sites에도 속해 있으며 이러한 Reference sites는 개체의 위험도에 관한 정보를 제공한다.

도로의 안전도를 평가할 경우, 어떤 대상 지점의 교통사고 분포는 사전분포에 해당하며, Reference sites의 교통사고분포는 사후분포에 해당한다. 이 방법은 어떤 특정

지역의 교통사고 현황을 그 대상 지점과 Reference sites의 교통사고 기록과 결합하여 교통사고의 불확실성으로 인한 문제를 극복할 수 있고 확률적 해석이 가능하다는 측면 때문에 전통적인 사고예측방법과는 많이 다르다. 또한, EB Approach는 RTM bias 제거에 매우 효과적인 방법으로 알려져 있다.

Hauer(1992)는 위와 같은 개념을 바탕으로 아래와 같은 기본식을 제안하였다.

$$E\{m|x\} = \alpha E(m) + (1-\alpha)x \quad (1)$$

여기서,

$$\alpha = \frac{E(m)}{E(m) + VAR(m)} \quad (2)$$

식(1)과 같이 교통사고 건수는 그 지점에서 발생한 교통사고건수(x)와 Reference sites에서 발생한 교통사고 건수($E(m)$)의 가중치에 의해 결정된다. 이 때, 가중치(α)는 Reference sites의 교통사고건수와 교통사고건수의 분산값을 이용하여 결정할 수 있으나, 현실적으로 Reference sites의 사고건수와 분산값을 예측하기 어렵기 때문에 Hauer는 교통사고예측모델을 이용하여 이를 해결하는 방안을 제시하였다. 여러 연구자들은 Hauer(1992)에 의해 제안된 EB 기법을 이용하였고, 그 중 Persaud 외(2001, 2007)는 식(3)을 이용하여 대상 지점의 사고건수를 예측하였다.

$$m : w_1(x) + w_2(P) \quad (3)$$

여기서,

m : 장래 대상 지점(Specific Sites)의 사고건수

x : 대상 지점(Specific Sites)의 실제 사고건수

P : Reference sites의 사고이력자료를 통해 추정된 사고건수

여기서, w_2 는 Hauer(1992)에 의해 제안된 식(1)의 α 와 같고 w_1 는 $(1-\alpha)$ 와 같은 값으로 표현될 수 있다. w_1 과 w_2 는 식(4)와 같이 추정된다.(Persaud, 2007)

$$w_1 = \frac{P}{P+1/k}, w_2 = \frac{1}{k(P+1/k)} \quad (4)$$

여기서,

k : overdispersion parameter

2. Poisson Regression Model

포아송 분포는 각 사건이 서로 독립적이고 일정한 발생확률을 가지고 있다는 전제하에 일정기간에 어떤 사건이 몇 번 일어날지의 확률을 묘사한다. 일반적으로 포아송 분포의 모양은 그 평균값에 달려 있으며, 이는 분산값과 동일하다. 만약 평균이 0에 접근하면 그 분포는 수직축에 크게 치우친 모양의 그래프를 보여주고, 평균이 매우 크면 수직축으로부터 멀리 떨어져 있으며, 포아송 분포는 대략 정규 분포와 유사하게 될 수 있다. 교통사고의 발생은 산발적이고 이산적인 형태의 분포를 나타내므로 포아송 분포를 사용하는 것이 일반적이며, 교차로의 사고건수 y_i 가 포아송 분포를 따른다는 가정 하에 교차로 i 에서 m 개의 변수에 의해 발생할 사고의 확률에 대한 일반식은 식(5)와 식(6)으로 표현될 수 있다.

$$P(y_i) = \frac{\text{EXP}(-\lambda_i)\lambda_i^{y_i}}{y_i!}, y = 0, 1, 2.. \tag{5}$$

여기서,

$P(y_i)$: 매년 i 교차로에서 y_i 번의 사고가 날 확률
 λ_i : 매년 i 교차로에서 일어날 사고건수(기대값)

$$\lambda_i = \exp(\beta_0 X_0 + \beta_1 X_1 + \dots + \beta_m X_m) = \exp\left(\sum_{j=0}^m \beta_j X_j\right) \tag{6}$$

여기서,

β_j : 회귀추정계수

3. Negative Binomial Regression Model

음이항 회귀모형은 각 분포의 분산이 같아야 한다는 포아송 회귀모형의 제약조건을 완화시키는데 사용할 수 있다. 사고예측모델 개발에 있어 포아송 회귀분석은 추정된 자료의 평균을 실제자료의 분산이 초과하면 과분산으로 나타난다. 과분산은 조사되지 않은 다른 성질의 구간이 다양한 결과를 나타내기 때문이며 음이항 회귀분석은 포아송 회귀분석과 비슷하게 i 교차로에서의 사고발생건수는 m 개의 변수와 연관성을 가지며 식(7)과 같이 표현할 수 있다.

$$\lambda_i = \exp(\beta_0 X_0 + \beta_1 X_1 + \dots + \beta_m X_m + \epsilon_i) = \exp\left(\sum_{j=0}^m \beta_j X_j\right) \tag{7}$$

여기서,

β_j : 회귀추정계수

ϵ_i : 오차항으로 평균이 1이고 분산이 α 인 감마분포로 가정

식(7)을 토대로 음이항 확률 분포(Negative Binomial Distribution)에 적용된 사고건수의 평균과 분산과의 관계는 식(8)과 같다.

$$\begin{aligned} \text{VAR}[y_i] &= E[y_i][1 + KE[y_i]] \\ &= E[y_i] + KE[y_i]^2 \end{aligned} \tag{8}$$

4. 모형검증

위 방법을 이용하여 개발된 모형의 검증을 위해 다양한 적합도 검증 방법(Goodness-of-fit)을 사용한다.

1) 우도비(ρ^2 : likelihood ratio index)

개발된 모형이 얼마나 좋은 설명력을 가지는지를 판단하는 하나의 지표로서 일반적으로 0과 1사이의 값을 가지며, 1에 가까울수록 좋은 설명력을 나타낸다. McFadden (1976)에 의하면, ρ^2 의 값이 0.2~0.4 사이의 값을 가지면 모형은 아주 좋은 설명력을 가진다고 평가하고 있다.

$$\rho^2 = 1 - \frac{L(\beta)}{L(0)} \tag{9}$$

여기서,

$L(\beta)$: log likelihood function

$L(0)$: restricted log likelihood function

2) MPB(Mean Prediction Bias)

이 통계수치는 모형을 만들기 위하여 종속변수로 사용된 자료에 대해서 모형에 의한 결과 값이 어느 정도 그리고 어떻게 치우쳐져 있는지를 판단할 수 있는 기준을 제공해준다. 이 방법에 의한 결과 값이 작을수록 모형의 예측 값은 정확한 것을 의미한다.

$$\text{MPB} = \frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)}{n} \tag{10}$$

여기서,

Y_i : 모형을 만들기 위하여 종속변수로 사용된 자료

\hat{Y}_i : 모형에 의한 결과값

3) MAD(Mean Absolute Deviation)

MAD는 모형의 예측 값이 평균적으로 얼마나 잘못 예측되었는지를 판단할 수 있는 척도를 제공해준다. 이 방법이 MPB와 다른 점은 각 수치의 음과 양의 차이로 인해 상쇄되지 않는다는 점이다. 결과 값이 0에 가까울수록 모형이 실제의 관측된 자료에 부합하는 결과를 나타냄을 의미한다.

$$MAD = \frac{\sum_{i=1}^n |Y_i - \hat{Y}_i|}{n} \tag{11}$$

여기서,

Y_i : 모형을 만들기 위하여 종속변수로 사용된 자료

\hat{Y}_i : 모형에 의한 결과값

IV. 자료수집 및 모형개발

1. 자료 수집

3년간(2004~2006년) 인천시 교통사고 잦은 지점으로 선정된 비슷한 유형의 4지 신호 교차로 조사 후 2002년~2006년까지 교통사고 개선 사업이 이루어진 곳은 제외한 총 71개 교차로에 대해 사고건수를 조사하였다. 또한 2004년~2006년 인천시 교통량 자료를 통해 지점별 교통량 자료를 수집하였다. 사고 자료와 교통량 자료를 통해 구축된 총 71개 지점과 향후 순위선정을 하기 위한 지점에 대해 25개 항목에 대해 현장조사를 실시하였다. 25개 항목에 대한 기하구조 조사 후 데이터 값에 변화를 주어 50개 항목으로 추출 후 최종적으로 35개 항목만 모형개발에 이용되었다.

2. 모형개발

1) 변수 선정

본 연구에서는 Reference sites의 사고예측모형(P)을 개발하기 위해 통계프로그램은 Limdep 8.0을 이용하였

다. 모형개발을 위해 사용된 변수로는 교통량, 차로수, 중차량 비율, 유출입구 수, 제약시설 유무 등 35개의 변수를 이용하였고, 이러한 변수들을 이용하여 조사된 자료의 과분산 검정을 시행한 결과, 과분산 계수(k)의 값이 0에 가까움을 알 수 있었다. 이는 교차로 사고예측모형을 개발하는데 있어 음이항 회귀모형보다는 포아송 회귀모형이 적합함을 나타낸다.

변수들의 상관분석의 결과, 종속변수인 교통사고건수와 상관관계가 높은 변수로는 ln교통량(주도로+부도로) ln교통량(주도로+부도로), 부도로 교통량, 부도로 중차량 교통량, 주도로/부도로 차선수, 부도로 버스베이, 부도로 우회전 차로수, 부도로 길어깨 등으로 나타났다. 종속변수와의 상관계수는 <표 1>와 같으며, 이 중 독립변수들간의 상관분석을 통해 최종적으로 선택된 변수와 변수에 대한 정의는 <표 3>과 같다. 또한, 최종적으로 선택된 변수들의 다중공선성 분석 결과, 변수들간 다중공선성은 존재하지 않는 것으로 분석되었다<표 2>.

<표 1> 독립변수 및 상관계수

독립변수	상관계수	독립변수	상관계수
	유의수준		유의수준
주도로 교통량	0.409	부도로 중차량교통량	0.586
	0.000		0.000
부도로 교통량	0.621	주도로 중차량교통량(ln)	0.358
	0.000		0.002
주도로 교통량(ln)	0.453	부도로 중차량교통량(ln)	0.358
	0.000		0.002
부도로 교통량(ln)	0.518	중차량 교통량(합)	0.554
	0.000		0.000
교통량(합)	0.647	중차량 교통량(합ln)	0.552
	0.000		0.000
교통량 (합ln)	0.662	중차량 교통량(곱ln)	0.344
	0.000		0.003
교통량 (곱ln)	0.661	주도로 차선수	0.309
	0.000		0.009
주도로 중차량교통량	0.415	부도로 차선수	0.491
	0.000		0.000
부도로 우회전차로수	0.270	부도로 버스베이	0.208
	0.023		0.082
부도로 길어깨	0.220	-	-
	0.043		

※ 90% 유의확률

<표 2> 최종 변수의 다중공선성 분석

변수	공차한계	VIF
ADT	0.642	1.558
부도로 버스베이	0.691	1.447
부도로 길어깨	0.548	1.826

<표 3> 모형개발을 위한 주요 변수

변수	변수의 정의
사고 건수	교차로에서 발생한 사고건수[건]
ADT	교통량 [ln교통량(주도로+부도로),대]
부도로 버스베이	부도로 버스베이 유무[0/1]
부도로 길어깨	부도로 길어깨[m]

이에, 최종적으로 선정된 모델의 모형식은 식(12)와 같다.

- Model 1 : $y = \exp(\alpha + \beta_1 X_1)$
 - Model 2 : $y = \exp(\alpha + \beta_1 X_1 + \beta_2 X_2)$
 - Model 3 : $y = \exp(\alpha + \beta_1 X_1 + \beta_2 X_3)$
 - Model 4 : $y = \exp(\alpha + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3)$
- (12)

여기서,

- X_1 : LN교통량(주도로+부도로)
- X_2 : 부도로 버스베이 유무
- X_3 : 부도로 길어깨

2) 모형개발 및 선정

상관분석 결과를 이용하여 교통사고에 영향을 미치는 변수들에 대해 살펴보고, 변수들간의 상관구조에 대해서도 파악하였다. 이러한 데이터의 특성을 고려하여 교차로 교통사고 예측모델에 대해 적합도 검증(Goodness-of-fit)을 통하여 ρ^2 (우도비), MAD, MPB를 알아보고 가장 적합한 결과가 도출된 모델을 주 모델로 선정하였다.

관련식은 식(13)과 같으며 각 모델에 대한 결과 값과 검증 값은 <표 4>와 같다.

$$Y = \exp(-7.586 + 0.930 \times X_1 + 0.296 \times X_2) \quad (13)$$

여기서,

- X_1 : LN교통량(주도로+부도로)
- X_2 : 부도로 버스베이 유무

EB 기법의 적용성을 검토하기 위해 앞에서 개발된 Reference site의 모형에 의해 도출된 사고건수와 EB 방법을 이용하여 도출된 사고건수를 71개 지점별 실제 사고건수와 비교해 보았다. 분석 결과, 총 사고건수 예측

<표 4> 교통사고예측 모델(Reference site)

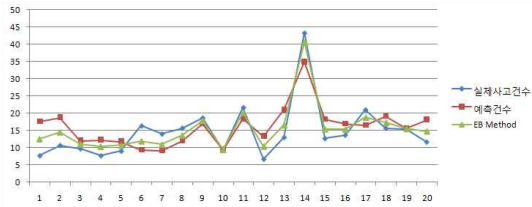
변수	주모델	예비 모델			
		모델 1	모델 2	모델 3	
상수	coeff.	-7.586	-7.633	-7.636	-7.581
	t-value	-8.207	-8.244	-8.245	-8.195
	P-value	0.000	0.000	0.000	0.000
LN(주도로+부도로 교통량)	coeff.	0.930	0.937	0.910	0.912
	t-value	11.324	11.395	10.934	10.963
	P-value	0.000	0.000	0.000	0.000
부도로 버스베이 유무	coeff.	0.296			0.258
	t-value	3.508			2.923
	P-value	0.005			0.004
부도로 길어깨	coeff.			0.374	0.249
	t-value			2.442	1.548
	P-value			0.015	0.122
ρ^2	0.23	0.22	0.23	0.24	
MPB/year	0.05	0.09	0.07	0.00	
MAD/year	4.78	5.01	4.90	4.69	
Alpha	0.06	0.07	0.06	0.05	

에서는 비슷한 경향을 보이나, 지점별 예측력에서는 EB 방법을 이용하여 도출된 결과 값이 사고모형을 이용하여 도출한 결과 값보다 더욱 우세한 결과 나타났다. <표 5>와 <그림 2>에서 자세한 결과값을 확인할 수 있다.(20개 지점만 표시)

또한 본 연구에서는 현 개선사업 우선순위 선정 방법으로 도출된 지점과 본 연구에서 EB 기법에 의해 도출

<표 5> 지점별 사고예측(71개 중 20개 지점만)

ID	지점명	실제 사고건수	사고 예측모형	EB 기법
1	사리골 4거리	8	18	12
2	화전 4거리	11	19	14
3	산곡 4거리	10	12	11
4	목재단지 4거리	8	12	10
5	구월시장 4거리	9	12	11
6	영아다방 4거리	16	9	12
7	청천 4거리	14	9	11
8	청수 4거리	16	12	14
9	석남 4거리	19	17	18
10	효성 4거리	9	9	9
11	계산 4거리	22	18	20
12	가구단지 4거리	7	13	10
13	장수 4거리	13	21	17
14	공촌 4거리	43	35	41
15	심곡 4거리	13	18	15
16	계산초교 4거리	14	17	15
17	암학 4거리	21	16	19
18	능해 IC	16	19	17
19	중합터미널 4거리	15	16	15
20	독곡 4거리	12	18	15



<그림 2> 지점별 사고 예측

된 지점을 2006년 인천시 4지 교차로 사고 잦은 지점을 대상으로 하여 비교해 보았다. 즉, 기존의 사고건수에 대한 가중 값을 EB 기법을 이용한 가중 값으로 대체하였을 경우 변화되는 순위를 살펴보았다. 기존 개선사업 우선순위 선정 방법은 식(14)와 같다.

$$ROPI_i = \left(\frac{NOA_i}{\sum_{j=1}^n NOA_j} + \frac{EPDO_i}{\sum_{j=1}^n EPDO_j} \right) \times 100 \quad (14)$$

여기서,

$ROPI_i$: i 지점의 사고우선순위

NOA_i : i 지점의 사고건수

$EPDO_i$: i 지점의 사고심각도

2006년 인천시 4지 교차로 중 교통사고 잦은 지점으로 선정된 인हा대 병원 4거리, 서해 4거리, 구 시민회관 4거리, 주안 4거리, 박문 4거리, 남동공단입구 4거리 부개수협 4거리, 부내초교 4거리, 연희 4거리, 서운 4거리를 대상으로 비교해 보았다.

<표 6>에서 볼 수 있듯이 서해4거리의 경우 실제 사고건수는 7건으로 우선순위는 타 지점에 비해 떨어지나 본 연구에서 적용한 방법으로 위험도를 예측한 결과 예

<표 6> 대상 지점 사고건수 예측

사고지점	실제 사고건수	순위	EB 기법	순위
인하대병원 4거리	20	3	21	3
서해 4거리	7	9	14	4
구시민회관 4거리	25	2	24	2
주안 4거리	12	4	12	6
박문 4거리	8	7	8	9
남동공단입구 4거리	28	1	29	1
부개수협 4거리	7	10	7	10
부내초교 4거리	8	8	11	7
연희 4거리	12	5	11	8
서운 4거리	10	6	12	5

측사고건수는 14건으로 개선우선순위 4위를 차지하는 것으로 나타났다. 이는 해당 교차로의 경우 실제 사고발생은 일어나지 않았지만 비슷한 여건의 교차로 사고특성을 고려할 경우 사고발생 가능성이 높다는 것을 의미한다. 물론 이와 반대의 경우도 발생할 수가 있는데 이처럼 단순한 사고건수를 이용하여 개선사업 우선순위를 판단할 경우 해당 지점의 교통여건을 정확하게 반영하지 못하는 한계가 있으며, 이러한 한계는 본 연구에서 제안하고 있는 EB 기법으로 극복이 가능할 것이다.

V. 결론 및 향후 연구과제

본 연구에서는 인천시 교통사고 잦은 지점으로 선정된 4지 교차로의 3년간(2004년~2006년) 사고 자료를 이용하여 Reference site의 사고예측모형을 개발하였다. 개발된 사고예측모형과 지점별 실제사고건수를 이용하여(EB 기법) 사고건수를 추정하고 결과의 유의성을 검토하였다. 검토 결과 총 사고건수 추정 시 기존의 사고예측 방법인 비선형 회귀모형과 EB 기법의 결과 값 모두 예측력이 높은 것으로 나타났지만 지점별 사고건수 예측력을 함께 고려할 경우엔 EB 기법이 비선형 회귀모형(포아송)의 결과보다 예측력이 좋은 것으로 나타났다. 또한 기존 방법으로 도출된 우선순위와 기존의 우선순위 선정 방법에서 사고건수에 대한 가중값을 EB 기법을 통해 추정된 사고건수에 대한 가중값으로 대체하여 도출된 값의 우선순위 비교 결과 대부분의 지점의 우선순위는 크게 변동이 없었으나, 서해4거리 등 몇 개 지점의 개선우선순위에는 상당한 변동이 발생하는 것으로 나타났다. 이는 실제 사고건수를 이용할 경우 발생하는 RTM문제를 본 연구에서 제안한 기법을 사용할 경우 해결가능하다는 것을 시사한다.

최근 사고예측모형은 활용도는 지속적으로 증가하고 있는 추세이다. 특히 도로 안전개선 사업 시 사업효과를 판단하기 위해 다양한 기법이 개발되어 적용되고 있다. 다시 말해 현재 사고다발지점 선정 시 과거 사고건수를 바탕으로 한 방법론이 가진 한계를 극복하고 잘못된 사업개선지점 선정에 따른 막대한 사회적 비용과 또 다른 사고를 야기 시키는 등의 문제점을 극복할 수 있으며, 또한 교통사고 잦은 지점 개선사업에 대한 효과평가를 위해 활용될 수 있다. 향후 후속연구가 지속적으로 진행된다면 EB 기법에 의한 사고예측모형은 사고 잦은 지점 선정 시 합리적인 기준을 제공할 수 있을 뿐만 아니라,

사고 잦은 지점 개선사업에 대한 효과분석 또한 가능할 것이며, 이를 통해 효율적인 교통안전대책을 개발·시행하여 교통사고로부터 국민들의 생명과 재산을 보호하는데 큰 역할을 할 수 있을 것이다.

알림 : 본 논문은 대한교통학회 제58회 학술발표회(2008. 2.23)에서 발표된 내용을 수정·보완하여 작성된 것입니다.

참고문헌

1. 강민욱, 2002, “고속도로 곡선 구간에서의 사고예측 모형 개발”, 제22권 제6호, 대한토목학회, pp.1077~1088.
2. 김원철(2001), “교차로 안전진단을 위한 교통사고건수예측모델화 수법에 관한 연구”, 21, 4(2001.7), 대한토목학회, pp.427~435.
3. 성낙문(2002), “교통사고 잦은 지점 및 구간선정방법 개선에 관한 연구”, 정책연구 2002-01, 한국교통연구원.
4. 성낙문(2003), “교통사고예측모델을 이용한 도로의 안전도 평가방법 연구”, 정책연구 2003-02, 한국교통연구원.
5. 오주택(2004), “도로설계시 안전성 평가모형 개발을 위한 기초연구”, 정책연구 2004-08, 한국교통연구원.
6. 하오근(2005), “국도변 신호교차로 안전성 향상을 위한 사고예측모형개발”, 제25권 제1호, 대한토목학회, pp.9~15.
7. 하태준·강정규·박제진(2001), “신호교차로 교통사고 예측모형의 개발 및 적용 (광주광역시 4-지 신호교차로를 중심으로)”, 대한교통학회지, 제19권 제6호, 대한교통학회, pp.207~218.
8. Abbess, C., Jarrett, D., Wright, C.C, (1981), “Accident at black-spot : estimating the effectiveness of remedial treatment, with special reference to the regression-to-mean effect”, Traffic Engineering and Control, Vol. 22, No. 10, pp.535~542.
9. Deacon, J.A., C.V.Zegeer, and R.C.Deen, (1975), “Identification of Hazardous Rural Highway Location”, In Transportation Research Record 543, TRB, pp.16~33.

10. Hauer, Ezra, Persaud, B., (1984), “Problem of Identifying Harzardous Locations Using Accident Data”, In Transportation Research Record 975, TRB, pp.36~43.
11. Hauer, Ezra, (1986), “On the Estimation of the Expected Number of Accident”, Accident Anal. & Prev., Vol. 18, No. 1, pp.1~12.
12. Hauer, Ezra, Persaud, B.,(1987), “How to Estimate the Safety of Rail-highway Grade Crossings and the Safety Effect of Warning Devices”, In Transportation Research Record 1114, TRB, pp.131~140.
13. Hauer, Ezra, J.C.N.Ng. and J. Lovell, (1988), “Estimation of Safety at Signalized Intersections”, In Transportation Research Record 1185, TRB, pp.48~61.
14. Hauer, Ezra, (1997), “Observational before-after studies in road safety”, Pergamon, Oxford, U.K.
15. Hauer, Ezra, Harwood, Douglas W., Council, Forrest M., Griffith, Michael S., (2002), “Estimating Safety by Empirical Bayes Method: A Tutorial”, In Transportation Research Record 1784, TRB, pp.126~131.
16. Hauer, Ezra, (1992), “Empirical bayes approach to the estimation of “UNSAFETY” : the multivariate regression method”, Accident Anal. & Prev., Vol. 24, No. 5, pp.457~477.
17. Hagle, J.L. and J.M.Hecht, (1988), “A Comparison of Techniques for the Identification of Hazardous Location”. In Transportation Research Record 1185, TRB, pp.24~36.
18. Laughland, J.C., L.E.Haefner, J.W.Hall, and D.R. Clough, (1975), “NCHRP Report 162 : Methods for Evaluating Highway Safety Improvements”, TRB.
19. Persaud, B., Lyon, C., and Nguyen, T., (1999), “Empirical Bayes Procedure for Ranking Sites for Safety Investigation by Potential for Safety Improvement”, In Transportation Research Record 1665, TRB, pp.7~12.
20. Persaud, B., Retting, R., Garder, P., and Nord,

- D., (2001), "Safety Effects of Roundabout Conversions in the United States", In Transportation Research Record 1751, TRB, pp.1~8.
21. Persaud, B., Lyon, C., (2007), "Empirical Bayes before-after safety studies : Lessons learned from two decades of experience and future directions", Accident Anal. & Prev., Vol. 39, No. 3, pp.546~555.
22. Stokes, R.W. and I.M.M utabazi, (1996), "Rate-Quality Control Method of Identifying Hazardous Road Locations". In Transportation Research Record 1542, TRB, pp.44~48.
23. Wen, Cheng, Simon, Washington, (2005), "Experimental evaluation of hotspot identification methods", Accident Anal. & Prev., Vol. 37, pp.870~881.
24. Wen, Cheng, Simon, Washington, (2009), "New Criteria for Evaluation Hotspot Identification Methods", TRB, 2008, pp.1~25.

✉ 주 작 성 자 : 정성봉

✉ 교 신 저 자 : 황보희

✉ 논문투고일 : 2008. 7. 2

✉ 논문심사일 : 2008. 8. 19 (1차)

2009. 3. 27 (2차)

2009. 5. 15 (3차)

✉ 심사판정일 : 2009. 5. 15

✉ 반론접수기한 : 2009. 10. 31

✉ 3인 익명 심사필

✉ 1인 abstract 교정필