

# ZAM 모형을 이용한 청주시 간선가로 구간의 사고모형 개발

## Developing the Accident Models of Cheongju Arterial Link Sections Using ZAM Model

박 병 호 Park, Byung Ho | 정회원 · 충북대학교 도시공학과 교수 (E-mail : bhpark@chungbuk.ac.kr)  
 김 준 용 Kim, Jun Yong | 비회원 · 충북대학교 도시공학과 석사과정 (E-mail : redblack21@nate.com)

### ABSTRACT

This study deals with the traffic accident of the Cheongju arterial link sections. The purpose of the study is to develop the traffic accident model. In pursuing the above, this study gives particular attentions to developing the ZAM(zero-altered model) model using the accident data of arterial roads divided by 322 small link sections. The main results analyzed by ZIP(zero inflated Poisson model) and ZINB(zero inflated negative binomial model) which are the methods of ZAM, are as follows. First, the evaluation of various developed models by the Vuong statistic and  $t$  statistic for overdispersion parameter  $\alpha$  shows that ZINB is analyzed to be optimal among Poisson, NB, ZIP(zero-inflated Poisson) and ZINB regression models. Second, ZINB is evaluated to be statistically significant in view of  $t$ ,  $p$  and  $\rho^2$  (0.63) values compared to other models. Finally, the accident factors of ZINB models are developed to be the traffic volume(ADT), number of entry/exit and length of median. The traffic volume(ADT) and the number of entry/exit are evaluated to be the '+' factors and the length of median to be '-' factor of the accident.

### KEYWORDS

*accident model, ZAM model, arterial link sections, negative binomial regression model, correlation analysis*

### 요지

본 연구는 청주시의 가로구간 교통사고를 다루고 있다. 연구의 목적은 가로구간의 사고모형을 개발하는 데 있다. 이를 위해서 이 연구에서는 전체 322개 세부구간으로 분리된 간선도로의 사고 자료를 이용하여 ZAM 모형을 개발하는데 중점을 두고 있다. ZAM 모형의 일종인 ZIP(zero inflated Poisson model)과 ZINB(zero inflated negative binomial model)를 중심으로 분석한 연구의 주요결과는 다음과 같다. 첫째, 모형의 적합성을 결정하는 Vuong 통계 값과 과분산계수  $\alpha$ 의  $t$  통계 값을 바탕으로 개발된 다양한 모형을 평가한 결과, 포아송, 음이항, ZIP 및 ZINB 회귀모형 중 ZINB 모형이 최적인 것으로 나타난다. 둘째, ZINB 모형은  $t$ ,  $p$  및  $\rho^2$  값 (0.63)의 관점에서 보면, 다른 모형에 비해서 통계적으로 매우 의미 있는 모형으로 평가된다. 마지막으로, 개발된 ZINB 모형의 사고 요인은 교통량, 진출입구 수 그리고 중앙분리대 길이로 분석된다. 교통량과 진출입구 수는 사고발생에 '+' 요인, 그리고 중앙분리대 길이는 '-' 요인으로 평가된다.

### 핵심용어

*사고모형, ZAM 모형, 간선가로 구간, 음이항 회귀모형, 상관관계 분석*

## 1. 서론

### 1.1. 연구의 배경 및 목적

오늘날 우리나라는 자동차의 대량 보급으로 가구당 1대 이상을 보유하고 있는 자동차 대국으로 자리잡고 있다. 자동차

가 대중화됨에 따라 우리나라도 마이카 시대에 접어들게 되었다. 하지만 차량의 급격한 증가는 과도한 혼잡과 함께 사고의 증가를 가져와 심각한 사회문제로 대두되고 있다.

자동차 사고는 그 발생건수가 급증하여 전국에서는 하루도

빠짐없이 사고발생 상황이 보도되고 있는 실정이다. 이에 본 연구에서는 청주시 간선가로 구간을 대상으로 2007년에 발생한 교통사고자료를 조사하고<sup>1)</sup>, 사고에 영향을 미칠 것이라 판단되는 도로조건, 교통조건 등의 자료를 조사하였다. 이 연구는 전문적인 조사 및 수집된 자료를 근거로 통계적 방법을 이용하여 가로구간에서 발생하는 사고모형을 개발하는데 그 목적이 있다.

## 1.2. 연구의 내용 및 방법

본 연구의 내용 및 방법은 다음과 같다.

첫째, 선행연구를 고찰함으로써 최근 연구동향을 파악하고 이를 검토한다.

둘째, 청주시의 간선가로 구간 중 주간선도로를 공간적 범위로 설정하고, 사고 자료의 특성(종속변수에 '0' 값이 다수 존재)에 적합한 분석방법으로 알려진 ZAM(zero-altered model) 모형의 이론을 검토한다.

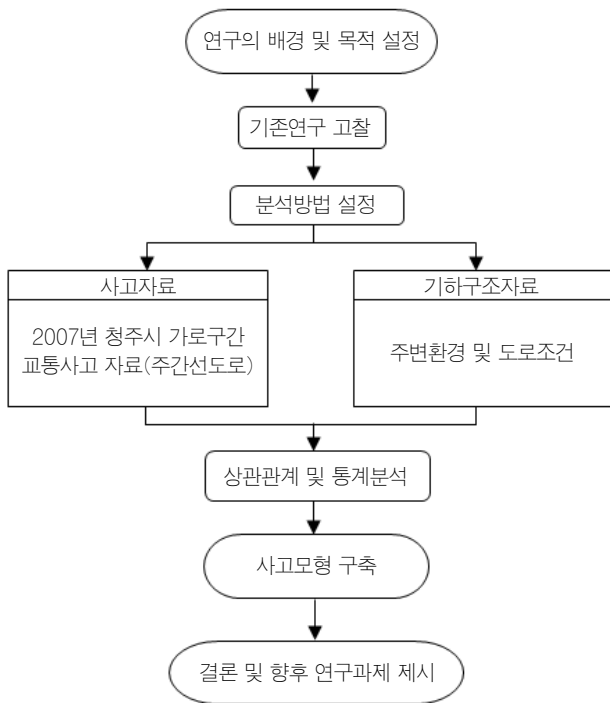


그림 1. 연구의 내용 및 순서

마지막으로 교통량, 차로 폭, 횡단거리 등과 같은 교통사고 관련 변수들을 현장조사를 통하여 수집 정리하고, 포아송 및 음이항 회귀분석과의 비교분석을 통해 통계적으로 가장 적합한 ZAM 모형을 이용한 사고모형을 개발한다.

자료의 정리 및 분석에는 SPSS 12.0을 이용하여 변수들 간의 상관관계를 분석하며, 교통사고 분석에는 LIMDEP

8.0을 이용한다.

## 2. 기존 연구의 고찰

이인성(2004)은 교통사고의 사망사고건수 중 0에 해당하는 빈도가 지나치게 많고 분산이 평균보다 큰 과분산 문제를 가지고 있다는 자료상의 문제를 극복하기 위하여 ZIP(zero-inflated Poisson)모형, 음이항 모형, ZINB(zero-inflated negative binomial) 회귀모형을 이용하였다. 그는 '0'의 빈도가 과다하게 많고 과분산 문제를 가진 교통사고 자료에는 ZIP 모형이나 음이항 회귀모형이 적합함을 밝히고 있다.

박정순 등(2007)은 충북 청주시 4지 신호교차로를 대상으로 다중선형, 다중비선형 및 다중비선형 회귀모형을 통하여 전체 교통사고와 후미추돌 사고에 관한 모형을 개발하였다. 여기에 사용된 변수는 4지 신호교차로 사고에 대하여 15개의 변수, 그리고 후미추돌사고에 대해서는 9개의 변수를 사용하여 회귀모형을 개발하고, 이에 대한 적합성 검정을 통하여 사고모형을 선정하였다.

박병호 등(2008)은 청주시의 3지 및 4지 비신호교차로를 대상으로 사고특성을 분석하고, ZAM 모형을 이용하여 사고모형을 개발하였다. 또한 비신호교차로의 사고요인은 교통량, 횡단보도수, 교차각, 시거 및 차로 폭으로 분석되었다.

Jinsun and Mannering(2002)은 갓길 폭, 도로 폭, 종단 경사, 교량 길이, 중앙분리대 폭 등 총 41개의 변수와 2,736개의 사고자료를 이용하여 사고의 심각도와 발생빈도에 대한 요소들의 파급효과에 대하여 ZINB 모형과 탄력성(elasticity)을 이용하여 분석하였다.

Dominique 등(2005)은 사고자료에 적합한 모형을 어떤 방법으로 적용할 것인지에 대한 걱정 지침을 마련하였고, 이를 위해 이론적인 원리와 기초 이해력을 시험하였다. 또한 사고에 적용하는 통계적인 모형을 평가하였으며, 그들 모형이 통계적으로 적합한지를 판단하였다.

Simon P. Washington 등(2003)은 사고모형의 적합성을 결정함에 있어 모형의  $r^2$  값, 과분산계수  $\alpha$ 를 이용하여 가이드 라인을 마련하고, 95% 신뢰수준에서 모형 결정을 하도록 모형 결정표를 제시하였다.

기존 연구는 교통사고 모형의 개발에 있어서 다양한 모형 중 사고자료에 따라 모형의 적합성 차이를 밝히고 있다. 또한 도로구간에서 사고가 없는 구간이 많이 포함되는 경우에는 ZAM 모형이 포아송 및 음이항 모형보다 통계적으로 의미있다는 것을 알 수 있다.

1) 「도로교통공단의 2007년 사고자료」

## 2.1. 기존 연구와의 차별성

이 연구의 차별성은 다음과 같다.

첫째, 교통사고를 다룬 기존의 선행연구는 대부분 교차로를 연구대상으로 선정하여 사고모형을 개발하였다. 그러나 본 연구는 사고분석의 대상지를 교차로가 아닌 간선가로 구간으로 설정하여 여기에서 발생하는 사고에 주목하여 관련변수를 분석하는데 중점을 두고 있다.

둘째, 청주시를 대상으로 사고건수가 '0'인 빈도가 과다하게 많은 간선가로 구간 사고자료의 특성을 감안하여 포아송과 음이항 회귀분석<sup>2)</sup>, ZAM 모형<sup>3)</sup>을 개발하고 가장 적합한 사고모형을 비교분석하였다는 점이 선행연구들과의 차별성이라 할 수 있다.

## 3. 분석의 틀 설정

### 3.1. 자료의 수집 및 분석

사고모형을 개발하기 위해서는 관련자료를 수집하는 작업이 선행되어야 한다. 본 연구에서의 조사 지점 및 수집방법은 그림 2와 같다.

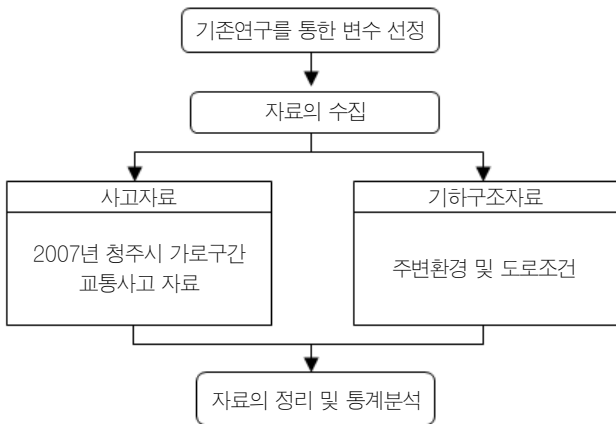


그림 2. 자료조사방법

사고자료는 최신자료인 도로교통공단의 2007년도 사고자료를 활용하여 전체 교통사고 중 간선가로 구간(단일로)의 사고건수를 코딩하였고, 관련변수에 대한 자료 수집은 청주시 CAD자료와 현장조사를 통하여 수집되었다. 청주시의 가로구간 중 주간선도로를 대상으로 교차로와 교차로 사이의 구간을 1개 구간으로 설정하여 전체 322개 세부구간으로 나누어 연구에 필요한 다양한 변수들을 수집하고 분석하였다.

2) 포아송 및 음이항 모형 : 교통사고 모형 중 가장 널리 사용되는 모형으로 사고가 드문 교통사고를 반영하기 위한 방법이며, 포아송 모형 자료의 평균과 분산이 동일해야한다는 전제가 요구됨.

3) ZAM 모형 : 교통사고 자료의 특성에 따라 사고가 많이 발생하는 구간과 사고가 없는 구간이 나타나고, 사고가 없는 구간이 많은 경우엔 ZAM(zero-altered model) 모형을 이용하는 것이 일반적임.

## 3.2. 변수의 설정

### 3.2.1. 종속변수 및 독립변수 선정

종속변수 선정에 있어서 이 연구는 구간별 사고의 특성을 가장 잘 반영할 수 있는 km당 사고건수<sup>4)</sup>를 종속변수로 선정하였다. 수집된 사고관련요소들 중에 연구의 대상지인 가로구간에서 발생한 교통사고와 밀접한 관련성이 있을 것으로 판단되는 변수를 독립변수로 정리하였다.

정리된 독립변수는 총 17개의 변수로서 사고와 가장 밀접한 관련이 있을 것이라 판단되는 일평균 교통량을 비롯해 도로 기하구조 중 교통운영에 주요한 역할을 담당하는 진출입구 수, 중앙분리대의 길이, 좌회전 및 우회전 전용차로 수와 길이, 도로의 경사도를 나타내는 종단경사, 차로 폭, 버스정류장의 수 등을 포함하였다. 정리된 종속변수와 독립변수는 다음 표 1과 같다.

표 1. 종속변수 및 독립변수

종속변수	단위	기호
km당 사고건수	건/km	$Y_1$
독립변수	단위	기호
1. 일평균 교통량	대/일	$X_1$
2. 횡단보도 수	개	$X_2$
3. 차로 수	개	$X_3$
4. 차로 폭	m	$X_4$
5. 진출입구 수	개	$X_5$
6. 중앙분리대 길이	m	$X_6$
7. 중앙분리대 비율	%	$X_7$
8. 버스정류장 수	개	$X_8$
9. 버스베이 길이	m	$X_9$
10. 굴곡점 수	개	$X_{10}$
11. 좌회전 전용차선 수	개	$X_{11}$
12. 좌회전 전용차선 길이	m	$X_{12}$
13. 우회전 전용차선 수	개	$X_{13}$
14. 우회전 전용차선 길이	m	$X_{14}$
15. 가로등 수	개	$X_{15}$
16. 평균 종단경사	%	$X_{16}$
17. 최급 종단경사	%	$X_{17}$

### 3.3. 사고특성 분석

전체 세부구간에 대한 사고특성을 분석한 결과, 322개 세부구간 중 사고가 발생한 구간은 124개 구간으로 38.5%이고, 무사고구간은 198개 구간으로 61.5%이었다. 또한 평균 사고건수는 2.31건으로 가장 사고가 많이 발생한 구간의 사고건수는 21건으로 나타났다.

4) 322개 세부구간의 길이가 달라 종속변수인 사고건수를 동일한 단위로 변환하기 위해서 km당 사고건수를 종속변수로 사용함.

구간별 횡단보도 수에 따른 사고분석에는 횡단보도가 2개 있는 구간에 사고건수가 53건으로 가장 많이 발생하는 것으로 분석되었다. 또한 1개와 3개인 구간에도 각각 25건과 29건으로 나타났다. 구간별 차로 수에 따른 사고건수에는 2차로인 경우에 69건으로 가장 많이 발생하는 것으로 분석되었고, 버스정류장 수에 따른 사고건수에는 정류장이 1개인 구간이 63건으로 발생확률이 가장 높은 것으로 분석되었다. 분석된 주요 결과는 표 2와 같다.

표 2. 사고특성분석

구 분	사고건수(건)
사고구간	124
무사고구간	198
평균사고건수	2.31
구간별 최대사고건수	21
구간별 최소사고건수	0

구 분	유형 및 사고건수(건)					
구간별 횡단보도수별 사고건수	0	1	2	3	4	5
	6	25	53	29	8	3
구간별 차선수별 사고건수	1	2	3	4		
	11	69	43	1		
구간별 버스정류장수별 사고건수	0	1	2	3		
	40	63	21	0		

#### 4. 사고모형 개발

교통사고는 임의적(random)이고 불규칙적으로 발생하게 된다. 이러한 불규칙적인 경우를 고려하면 여러 가지 회귀분석 중 비선형 회귀분석인 포아송과 음이항 회귀분석 그리고 이 두 모형의 보다 확장된 형태인 ZAM 모형이 우선 고려대상이라 판단된다. 포아송 및 음이항 모형은 교통사고 모형 중 가장 널리 사용되는 모형으로 사고가 드문 교통사고특성을 반영하기 위한 방법이며, ZAM 모형은 사고의 특성과 사고가 발생하는 구간이나 지점의 특성에 따라 사고가 없는 구간(지점)이 많은 경우에 이용하는 것이 일반적이다.

##### 4.1. 상관관계분석

변수간의 상관관계를 분석하기 위하여 신뢰수준을 90%로 설정하여 km당 사고건수( $Y_1$ )를 종속변수로 하고, 각각의 독립변수( $X_i$ )를 이용하여 상관분석을 실시하였다. 상관관계 분석결과, 횡단보도수, 차로수, 진출입구 수 그리고 교통량이 사고와 밀접한 관련이 있는 것으로 나타났다. 상관관계 분석 결과는 표 3과 같다.

##### 4.2. 다중공선성 분석

설명력이 높은 선형 종속관계(상관관계)가 존재하는 경우, 최소제곱추정량의 계산이 불가능할 수 있고, 추정량의 분산이 커지는 문제가 발생한다. 이 경우 설명변수들 간에 다중공

표 3. 상관관계분석 결과

	$Y_1$	$X_1$	$X_2$	$X_3$	$X_5$	$X_6$	$X_7$	$X_8$	$X_9$	$X_{10}$	$X_{11}$	$X_{12}$	$X_{13}$	$X_{14}$	$X_{15}$	$X_{16}$	$X_{17}$
$Y_1$	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$X_1$	0.382	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$X_2$	0.125	0.031	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
$X_3$	-0.064	0.060	-0.092	1	-	-	-	-	-	-	-	-	-	-	-	-	-
$X_5$	0.390	0.113	0.480	-0.084	1	-	-	-	-	-	-	-	-	-	-	-	-
$X_6$	-0.104	0.086	-0.197	0.034	-0.025	1	-	-	-	-	-	-	-	-	-	-	-
$X_7$	-0.020	0.101	-0.254	0.035	-0.075	0.798	1	-	-	-	-	-	-	-	-	-	-
$X_8$	0.086	-0.005	0.330	0.024	0.332	-0.138	-0.146	1	-	-	-	-	-	-	-	-	-
$X_9$	-0.043	0.031	0.053	-0.002	-0.008	-0.099	-0.087	0.369	1	-	-	-	-	-	-	-	-
$X_{10}$	-0.094	-0.021	0.130	-0.021	0.102	0.203	0.068	0.197	-0.001	1	-	-	-	-	-	-	-
$X_{11}$	-0.014	0.029	-0.019	0.093	0.068	-0.130	-0.254	0.032	0.027	-0.050	1	-	-	-	-	-	-
$X_{12}$	0.083	0.041	0.114	0.057	0.197	-0.143	-0.207	0.112	-0.023	-0.010	0.633	1	-	-	-	-	-
$X_{13}$	0.002	0.086	-0.145	0.239	-0.074	0.243	0.280	-0.083	0.064	-0.092	-0.092	-0.089	1	-	-	-	-
$X_{14}$	0.000	0.023	-0.124	0.050	0.019	0.386	0.452	-0.092	-0.017	-0.053	-0.149	-0.097	0.566	1	-	-	-
$X_{15}$	0.057	0.196	0.241	0.125	0.165	0.032	0.042	0.134	0.019	0.262	-0.099	0.025	-0.018	0.087	1	-	-
$X_{16}$	-0.046	0.077	-0.183	0.005	-0.003	0.268	0.122	-0.038	0.017	0.242	-0.052	-0.078	-0.014	0.136	0.121	1	-
$X_{17}$	-0.014	0.018	-0.001	0.050	0.021	0.016	0.018	-0.052	-0.049	-0.017	0.033	0.006	-0.030	-0.019	0.091	-0.006	1

주) pearson 상관계수

선성(multicollinearity)이 존재한다.

따라서 본 연구에서 관찰된 설명변수들 간의 상관관계를 분석하였으며, 이 때 설명변수들 간의 상관관계가 비정상적으로 높다면, 다중공선성을 가지고 있는지를 확인해 볼 필요가 있다. 다중공선성 분석결과는 표 4와 같다.

표 4의 결과를 보면, 일평균교통량은 진출입구수와 세부구간별 가로등 수, 차로 수는 우회전 전용차선 수, 버스정류장 수는 버스베이 길이와 굴곡점 수, 횡단보도 수는 진출입구 수, 버스정류장 수, 굴곡점수, 좌회전 전용차선 길이와 상관관계가 높다. 즉 자료의 설명변수들 간에는 부분적으로 강한 상관관계수가 존재하며, 이는 다중공선성의 존재를 의심해 볼 필요가 있음을 알 수 있다.

표 4. 독립변수의 다중공선성

독립변수	다중공선성 변수
일평균교통량( $X_1$ )	진출입구 수( $X_5$ ), 세부구간별 가로등 수( $X_{15}$ )
횡단보도 수( $X_2$ )	진출입구 수( $X_5$ ) 버스정류장 수( $X_8$ ) 굴곡점 수( $X_{10}$ ) 좌회전 전용차선 길이( $X_{12}$ ) 세부구간별 가로등 수( $X_{15}$ )
차로 수( $X_3$ )	우회전 전용차선 수( $X_{13}$ )
진출입구 수( $X_5$ )	버스정류장 수( $X_8$ ) 좌회전 전용차선 길이( $X_{12}$ )
중앙분리대 길이( $X_6$ )	중앙분리대 비율( $X_7$ ) 굴곡점 수( $X_{10}$ ) 우회전 전용차선 수( $X_{13}$ ) 우회전 전용차선 길이( $X_{14}$ ) 세부구간별 평균 종단경사( $X_{16}$ )
중앙분리대 비율( $X_7$ )	우회전 전용차선 수( $X_{13}$ ) 우회전 전용차선 길이( $X_{14}$ )
버스정류장 수( $X_8$ )	버스베이 길이( $X_9$ ) 굴곡점 수( $X_{10}$ )
굴곡점 수( $X_{10}$ )	세부구간별 가로등 수( $X_{15}$ ) 세부구간별 평균 종단경사( $X_{16}$ )
좌회전 전용차선 수( $X_{11}$ )	좌회전 전용차선 길이( $X_{12}$ )
우회전 전용차선 수( $X_{13}$ )	우회전 전용차선 길이( $X_{14}$ )

따라서 이 연구에서는 다중공선성이 존재한다고 판단되는 변수를 모형식에 적용할 때 이를 감안하여 제거하였다.

### 4.3. 사고모형 개발 및 결과 분석

#### 4.3.1. 포아송 및 음이항 모형분석

본 연구에서는 가로구간 전체 322개 세부구간을 중심으로 수집된 자료를 근거로 변수의 상관관계를 검토한 결과를 토대로 사회과학 통계프로그램인 LIMDEP 8.0을 이용하여 모형을 개발한다.

사고모형 개발에 있어서 가장 일반적인 포아송(Poisson) 및 음이항(negative binomial) 모형을 이용하여 사고모형을 개발한 결과, 포아송 사고모형에서 사고요인으로 교통량, 진출입구 수, 중앙분리대 길이 및 좌회전 전용차선 수가 분석되었다.

또한 음이항 회귀모형을 이용한 사고모형에서는 교통량, 진출입구 수, 중앙분리대 길이가 포아송 모형과 동일하게 포함된 것으로 나타나, 이들 변수가 사고에 영향력이 큰 것으로 분석되었다.

표 5. 포아송 모형

기호	변수	계수	t 값	p 값
-	상수항	-0.6815	-6.092	0.000
$X_1$	교통량	0.0007	15.613	0.000
$X_5$	진출입구 수	0.2115	14.327	0.000
$X_6$	중앙분리대 길이	-0.0011	-5.326	0.000
$X_{11}$	좌회전 전용차선 수	-0.3436	-4.053	0.000
통계값	관찰 수	322		
	로그 우도 함수	-826.1		
	제약 로그 우도	-1090.6		
	$\rho^2$	0.75		

표 6. 음이항 모형

기호	변수	계수	t 값	p 값
-	상수항	-0.9663	-3.860	0.000
$X_1$	교통량	0.0006	4.569	0.000
$X_5$	진출입구 수	0.2877	6.005	0.000
$X_6$	중앙분리대 길이	-0.0011	-3.039	0.002
	$\alpha$	1.5896	4.066	0.000
통계값	관찰 수	322		
	로그 우도 함수	-555.3		
	제약 로그 우도	-833.9		
	$\rho^2$	0.66		

#### 4.3.2. ZAM 모형분석

앞서 설명된 바와 같이, 대상구간의 무사고 구간이 198개 구간으로 전체의 61.5%이므로, 사고건수가 '0' 값이 매우 많은 문제를 극복하기 위하여 이 연구에서는 ZAM 모형의 일종인 ZIP(zero-inflated Poisson)와 ZINB(zero-inflated negative binomial) 회귀모형을 사용하였다. ZIP 모형은 포아송 모형의 확장된 형태이며, ZINB 모형은 음이항 모형에서 보다 확장된 모형을 의미한다. 또한 모형을 결정함에 있어 기존 연구 고찰에서 언급한 Simon P. Washigton 등이 제시한 모형 결정계수 표를 참조하여 최종적으로 모형의 적합성을 검증하였다.

표 7. 사고모형 결정계수 범위

구 분	NB 과분산계수 $\alpha$ 의 $t$ 통계 값		
		<[1.96]	>[1.96]
ZINB와 NB 비교를 위한 Vuong <sup>5)</sup> 통계 값	<-1.96	ZIP 혹은 Poisson	NB
	>1.96	ZIP	ZINB

자료 : Simon P. Washington(2003), statistical and econometric methods for transportation data analysis, pp. 250-254.

위의 분석방법을 통하여 통계적으로 의미있는 사고모형이 개발되었다. 모형의 Vuong 통계 값이 ZIP 7.1062이고, ZINB 8.3737로 1.96보다 큰 값을 나타내고 있다. 따라서 포아송이나 음이항 모형보다는 ZAM 모형의 일종인 ZIP모형과 ZINB모형이 통계적으로 더 적합한 것으로 분석되었다.

표 8. ZIP 모형

기호	변수	계수	$t$ 값	$p$ 값
-	상수항	1.1249	13.538	0.000
$X_1$	교통량	0.0003	10.887	0.000
$X_5$	진출입구 수	0.0481	4.313	0.000
$X_6$	중앙분리대 길이	-0.0013	-6.5465	0.0043
$X_{11}$	좌회전 전용차선 수	-0.1812	-2.8566	0.0354
통계 값	관찰 수	322		
	로그 우도 함수	-554.9		
	제약 로그 우도	-1090.6		
	$\rho^2$	0.51		
	Vuong 통계 값	7.1062		

표 9. ZINB 모형

기호	변수	계수	$t$ 값	$p$ 값
-	상수항	-0.6069	-2.8610	0.004
$X_1$	교통량	0.0007	4.9679	0.000
$X_5$	진출입구 수	0.2102	4.3024	0.000
$X_6$	중앙분리대 길이	-0.0007	-2.4537	0.014
-	$\alpha$	1.5896	7.3091	0.000
	Tau	-1.0803	-2.9906	0.003
통계 값	관찰 수	322		
	로그 우도 함수	-526.9		
	제약 로그 우도	-833.9		
	$\rho^2$	0.63		
	Vuong 통계 값	8.3737		

### 4.3.3. 결과분석

표 8과 표 9에 나타나듯이, ZIP 모형과 ZINB 모형의  $\rho^2$

5) Vuong Test : 1989년 전통적인 모형과 비교하여 ZAM 모형의 사용에 대한 적절성을 검증하기 위하여 Vuong이 제안한 통계 검정

값은 각각 0.51과 0.63으로 분석되었다. 또한 표 7에서 나타나듯이 과분산계수  $\alpha$  값의  $t$  통계 값이 1.96보다 큰 경우에는 ZIP 모형보다 ZINB 모형이 통계적으로 보다 유의한 것으로 분석된다.

ZINB 모형의  $t$ 값과  $p$ 값은 모두 95% 신뢰수준에서 유의하며, 모형의  $\rho^2$  값도 0.63으로, ZINB 모형은 통계적으로 매우 의미 있는 모형으로 판단된다.

ZINB 모형의 사고요인에는 ZIP 모형과 동일하게 좌회전 전용차선 수를 제외한 교통량, 진출입구수 그리고 중앙분리대 길이가 포함된 것으로 분석되었다. 즉, 교통량과 진출입구수가 많을수록 사고발생에 '+' 요인, 중앙분리대 길이가 길수록 사고발생에 '-' 요인인 것으로 분석되었다.

## 5. 결론

이 연구는 청주시 가로구간의 기하구조와 교통사고자료를 바탕으로 ZAM 모형을 이용하여 사고모형을 개발하였다. 주요 분석 결과는 다음과 같다.

첫째, ZINB와 NB의 비교를 위한 Vuong 통계 값과 NB의 과분산계수  $\alpha$ 의  $t$  통계 값을 바탕으로 모형을 평가한 결과, 포아송, 음이항, ZIP 및 ZINB 중 ZINB 모형이 최적모형으로 선정되었다.

둘째, ZINB 모형의  $t$ 값과  $p$ 값은 모두 95% 신뢰수준에서 유의하며, 모형의  $\rho^2$  값도 0.63으로, ZINB 모형은 통계적으로 매우 의미 있는 것으로 평가되었다.

마지막으로, ZINB 모형의 사고요인은 교통량, 진출입구수 그리고 중앙분리대 길이로 분석된다. 교통량과 진출입구수는 사고발생에 '+' 요인, 그리고 중앙분리대 길이는 '-' 요인인 것으로 분석된다.

향후 연구과제로는 보조간선도로를 포함한 청주시 전체 가로구간의 사고자료와 기하구조 자료를 수집하여 더욱 정확한 사고모형을 개발할 필요가 있는 것으로 판단된다.

## 감사의 글

이 논문은 2009년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었음.

## 참고 문헌

박병호 · 박상혁 · 이영민 · 인병철(2008), "ZAM 모형을 이용한 비신호차로 사고특성 분석 - 청주시 3지 · 4지 비신호차로를 사례로-", 국토계획, 대한국토 · 도시계획학회, 제43권 제6호, pp. 66~78.

박정순 · 김태영 · 유두선(2007), "도로환경요인과 교통사고의 상관

분석 및 사고추정모형 개발”, *대한교통학회지*, *대한교통학회*, 제 25권 제2호 pp. 63~72.

Dominique Lord, Simon P. Washington, John N IVAN(2005), Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory, *Accident Analysis and Prevention* 37, p. 35-46.

Dominique Lord, Simon Washington, John N IVAN(2007), Further notes on the application of zero-inflated models in highway safety, *Accident Analysis and Prevention* 39, p. 53-57.

Insung Lee(2004), *Application of ZIP regression to traffic accident data*, The Graduate School Yonsei University Department of Applied Statistics.

Jinsun Lee and Fred Mannering(2002), Impact of roadside features on the frequency and severity of run-off-roadway accidents : an empirical analysis, *Accident analysis & prevention* 34: 149-161.

Simon P. Washington, Matthew G. Karlaftis, Fred L. Mannering(2003), *Statistical and econometric methods for transportation data analysis*.

접 수 일 : 2009. 12. 20

심 사 일 : 2009. 12. 27

심사완료일 : 2010. 5. 18