International Transport Forum

IRTAD
International Traffic Safety Data and Analysis Group

OECD

Analysis of Design Elements for Urban Highway Safety

SooBeom, Lee

4th IRTAD CONFERENCE
Road safety data: collection and analysis for target setting and monitoring performances and progress

Seoul, 16-17 September 2009

Contents

- Introduction
- Literature Review
- Data Description
- Model Development
- Conclusions

Introduction

- It is essential to identify and analyze the accident inducing factors and come up with accident reduction policies.
- This study attempts to develop an accident frequency prediction model.
- The major role of the developed model is to judge the safety level of highway segments by predicting the accident potential .

Introduction

- The developed model is only for urban highways.
- The developed model is focused only roadways excluding intersections.

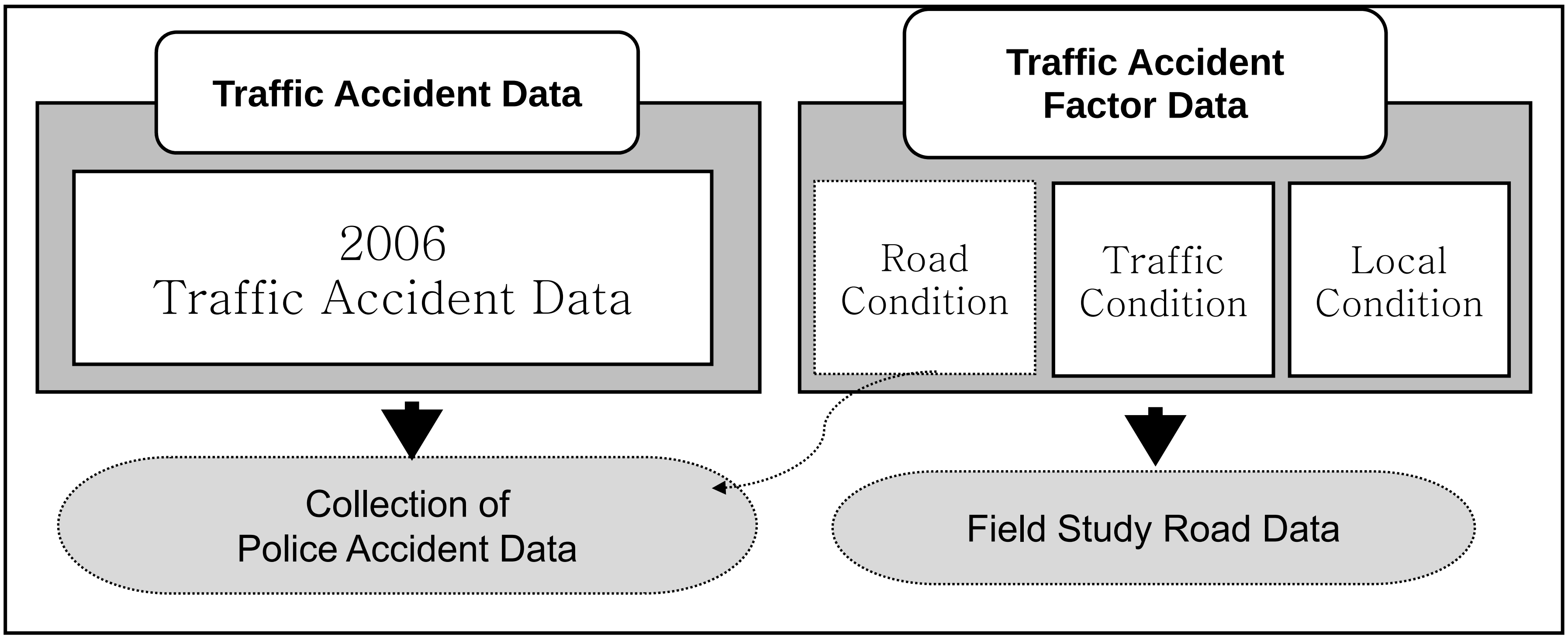
Literature Review

Model Variable	Typical Value	Safety Prediction Model Developer			
		(1)	(2)	(3)	(4)
Segment length, mi	0.5	√	√	√	√
Driveway density, driveways/mile	50 for undivided, 50 for TWLTL, 25 for raised curb	√	-	√	√
Number of intersections	5	-	√	-	-
Truck percentage	6.0	√	-	√	
Adjacent land use	Varied	√	-	√	√
Outside edge of pavement	Curb	-	√	√	-
Shoulder width, ft	1.5	√	√	√	-
Lane width, ft	12	-	√	-	-
Curve radius, ft	infinite (i.e., tangent)	-	-	√	-
Presence of curb parking	No	-	-	√	√
Presence of TWLTL	Varied	√	-	√	√
Presence of median	Varied	√	-	-	√
Speed limit, mph	35 residential and industrial; 40 business and office	-	√	√	-

Quotation: **(1)Harwood,** *Multilane Design Alternatives for Improving Suburban Highways* ,1986
(2)Hadi et al., Estimating Safety Effects of Cross- Section Design for Various Highway Types Using Negative Binomial Regression , 1995
(3)Hauer et al., Safety Models for Urban Four-Lane Undivided Road Segments, 2004
(4)Bonneson&McCoy, *Capacity and Operational Effects of Midblock Left-Turn Lanes.,* 1997

Data Description

- The vehicle-to-vehicle accident data are extracted from the police accident database



Data Description

- Field studies were conducted on 13 routes of urban high-way which showed typical urban highway characteristics ; Seoul Metropolitan Areas.

Research Section (Total Segment Length)	
Gyeonggi Sungnamsi	Sujeongro (3km), Beonyungro (4km)
Gyeonggi Goyangsi	Ilsanro (4km), Kangseokro (3km)
Gyeonggi Gwangmyeongsi	Oriro (5km)
Gyeonggi Gurisi	Kyungchunro (3km), (4km)
Dongdaemun	Jeonnongro (3km), Jegiro (3km), Hongneungro (2km)
Eunpyeong	Tongilro (5km)
Seocho	Unjuro (5km), Nambu Loop (33km)

Total Length : 77km   Number of Homogeneity Segment: 246

Data Description

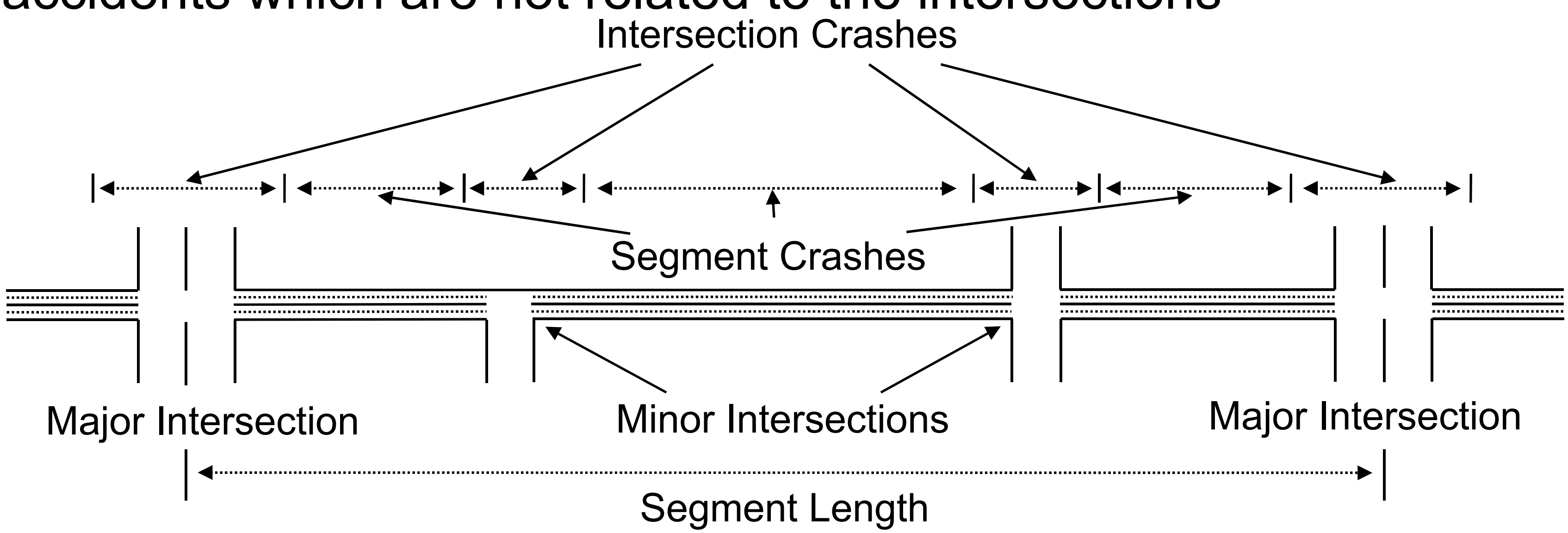
- Geometric & Traffic Condition DB for the Survey Result

Variable	Description of Variable	Variable	Description of Variable
Frequency	the number of traffic accidents	Speed Control	the number of speed control systems/ devices
AADT	annual average daily traffic volume [veh/day]	Terrain	terrain [level=1, rolling=2, mountainous=3]
Length	highway segment length [m]	Delineation	the presence of delineation systems [Yes=1, No=2]
H.C	the presence of horizontal curves [m]	Md_Type	the median type [None=0, Concrete=1, Guardrail=2, greenbelt=3, Others=4]
V.C	the presence of vertical curves [%]	Md_Width	the median width [m]
Radius	horizontal curvature [m]	Speed limit	posted speed [kph]
G	vertical grades [%]	Crosswalk	the number of crosswalks
Driveway	the number of driveways	Bus Stop	the number of bus stops
Lighting	the number of lights	On Street Parking	presence and type of on-street parking [Absence=0, Presence=1]
Terrain	terrain type [level, rolling, mountainous]	Ex_Buslane	exclusive bus lanes (Roadside)
T_width	traveled lane average Width [m]	Num_Bus	number of bus service routes
Sh_Width	the shoulder width [m]	Ill_Parking	presence/absence of illegal parking [Absence=0, Presence=1]
Sh_Type	the shoulder type [Non=0, Pavement=1, Non Pavement=2, Others=3]	Land-Use	area-use around highway segments [residential=1, commercial=2, farmland=3, industrial=4, residential-commercial=5, farmland-industrial=6, residential-farmland=7, industrial-farmland=8, others=9]
Lane	the number of lanes		

Data : Homogeneous Highway

Homogeneous Highway Segments in the Highway Safety Manual

- A highway is considered to be homogeneous if it is not divided into segments and if there is no change in ADT, lane width, shoulder width, presence of median, median width, side slope, and major intersection.
- Traffic accidents in the highway segment include only those accidents which are not related to the intersections



Data : Homogeneous Highway Segments

Homogeneous Highway Segment in the IHSDM
(Interactive Highway Safety Design Model)

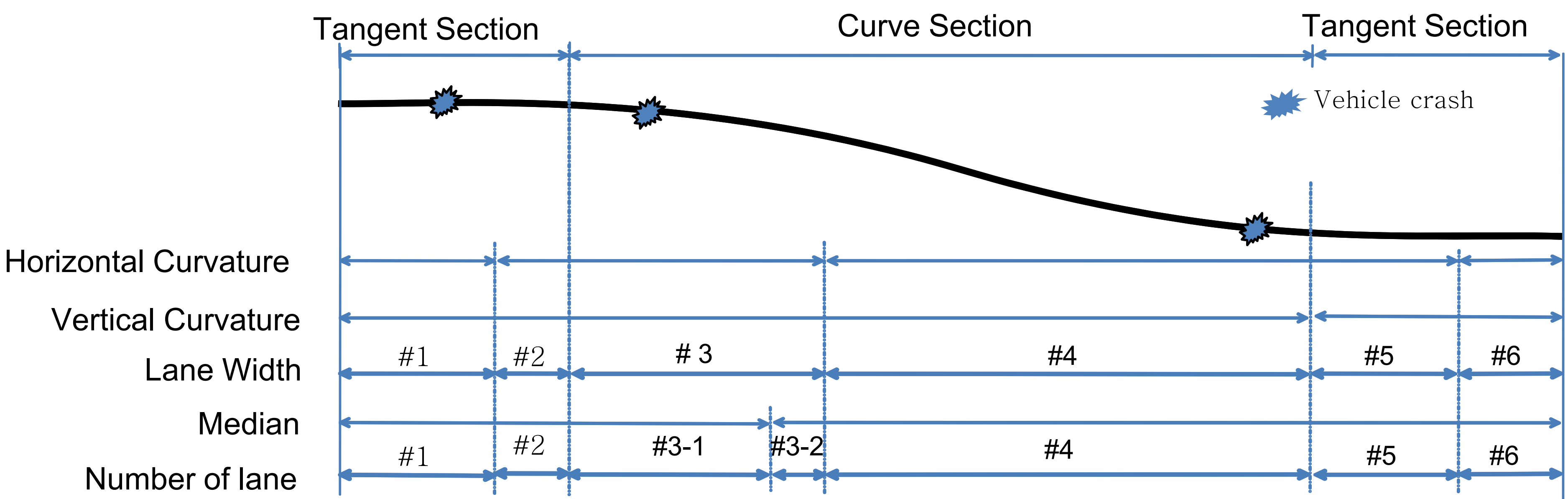
- The roadway must be divided into homogeneous segments.
- A new homogeneous segment begins at any of following location

  - Intersection.
  - Beginning or end of a horizontal curve.
  - Point of vertical intersection (PVI) for a crest vertical curve, a sag vertical curve, or an angle point at which two different roadway grades meet.
  - Beginning or end of a passing lane or short four-lane section provided for the purpose of increasing passing opportunities.
  - Beginning or end of a center two-way left-turn lane.

Data : Homogeneous Highway Segments

Homogeneous highway segment in this study

- The highway segment is considered to be homogeneous as far as none of the following design elements – horizontal curvature, vertical curvature, and lane width - is changing. Below figure shows the homogeneous highway segment in this study.



Model: Poisson Regression Model

- A Poisson distribution is given by:

$$\Pr[Y = y] = \frac{e^{\lambda} \lambda^{y}}{y\ !}, y \ = 0,1,2..$$

Where, λ is the average number of occurrences in a specified interval

- Poisson regression is modeled on the log scale
- Log (# of counts) = $B_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n$
- where B's are the regression coefficients, and the x's are the predictors.

Model: Negative binomial Regression Model

- A Negative binomial distribution is given by:

$$P(\mu_i = \lambda_i) = \frac{(\kappa / \mu_i)\lambda_i^{\kappa - 1} e^{-(\kappa / \mu_i)\lambda_i}}{\Gamma(\kappa)}$$

Where, μ_i is always a positive value

- Negative binomial distribution is always corrects for overdispersion
- Log (# of events) = $B_o + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n. + \sigma * e$, where e are the errors.
- The term $\sigma * e$ corrects for overdispersion

Model: Zero Inflated Poisson Regression Model

- Two sources of zeroes
  - a point mass of zeroes
  - zeroes from standard Poisson distribution
- The probability density function of a ZIP model can be expressed as follows :

$$\begin{cases} Y_i = 0, \text{ with probability: } P(Y=0) = p_i + (1-p_i)e^{-\lambda_i} \\ Y_i = y_i^*, \text{with probability: } P(Y=y_i) = (1-p_i)\dfrac{\lambda_i^{y_i} e^{-\lambda_i}}{y_i!} \end{cases}$$

Model Development : Validation of Model

ρ^2 (Rho-Squared)

- Maximum likelihood ratio ρ^2 is used to test the fit-ness of the model developed in this study. ρ^2 Is the number between 0 and 1, and the model shows the higher fitness as ρ^2 is close to 1.
- In the relationship between the traffic accidents an d the highway geometrics and facilities, the value of between 0.2 and 0.4, however, represents rath er excellent fitness (McFadden, 1978)

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

$L(\beta)$: Log likelihood function

$L(0)$: Restricted log likelihood

Model Development : Validation of Model

MPB (Mean Prediction Bias)

- MPB is used to judge the skewness of the results of the model due to the dependent variables.
- As long as the value is small, the value predicted by the model tends to be correct.

$$MPB = \frac{\sum_{j=1}^{n}(Y_j - \hat{Y}_j)}{n}$$

Where, Y_j : Actually Dependent Variable

$\hat{Y}_j$: Predicted Dependent Variable for Model

Model Development : Validation of Model

MAB (Mean Absolute Bias)

- MAB indicates how much error is contained on a verage in the predicted value of the model.
- MAB is different from MPB in which the values are not canceled out by the negative and positiv e numbers.

$$MAB = \frac{\sum_{j=1}^{n} \left| Y_j - \hat{Y}_j \right|}{n}$$

Where, Y_j : Actually Dependent Variable

$\hat{Y}_j$: Predicted Dependent Variable for Model

Model Development :Data Summary of Statistics

- Below table shows the summary statistics of the data used in this study

	Four or less(84)				Six(74)				Eight or more(88)			
	Min.	Max.	Mean	S.D	Min.	Max.	Mean	S.D	Min.	Max.	Mean	S.D
Frequency	0	10	1.6	2.0	0	8	2.41	2.1	0	13	4.1	3.0
The number of Driveway	0	14	2.1	2.5	0	10	2.9	1.8	0	12	3.8	2.1
Roadway segment length[m]	26	890	126.3	146.9	31	956	140.1	126.2	39.5	449	142.9	76.00
horizontal curve[m]	0	870	122.7	214.8	0	1800	160.0	310.7	0	857	88.4	233.6
vertical curve[%]	0	4.4	0.9	1.2	0	4	1.1	1.2	0	4.5	1.0	1.2
Traveled Lane Width[m]	2.4	4.46	3.3	0.4	2.53	5.25	3.7	0.6	2.58	4.1	3.3	0.2

Model Development : Selection of the Optimum Traffic Accident Frequency Prediction Model

- Decision rule for model selection under ZINB-NB comparis ons using the Vuong statistic and overdispersion parameter criteria. A value less than -1.96 rejects the ZIP model.

		NB Overdispersion parameter a (t-statistic)	
		<2	>2
Vuong Statistic for ZINB-NB comparison	< |1.96|	ZIP	NB
	> 1.96	ZIP	ZINB

Quotation: Vhon milton and Fred Mannering(1998), "Modeling Accident Frequencies as Zero-Altered Probability Processes: An Empirical Inquiry" , TRB

Model Development : Result

Type: 4 Lane and Under

$$Y = EXPO \times \exp(-7.345 + 0.017X_2 + 0.524X_3 + 0.108X_4)$$

Variable name			Four or Less
Constant		Parameter	-7.345
		t-statistic	-37.658
		P-value	0.000
Expo (LengthXADTX365X10^{-6})	X_1	Parameter	Fixed Parameter
		t-statistic	
		P-value	
Density of Driveways (km/Num)	X_2	Parameter	0.017
		t-statistic	2.581
		P-value	0.010
Presence of Vertical Curvature (%)	X_3	Parameter	0.524
		t-statistic	2.779
		P-value	0.006
Horizontal Curvature (1/R)	X_4	Parameter	0.108
		t-statistic	2.114
		P-value	0.345

Log likelihood function	-110.542
Restricted log likelihood	-133.913
α (t-statistic)	0.000
ρ^2	0.175
MPB	0.081
MAD	0.979
Vuong Statistic	-3.542

Variable Name	Tolerance Limit	VIF
EXPO (LengthXADTX365X10^{-6})	0.97	1.03
Density of Driveways (km/Num)	0.84	1.18
Presence of Vertical Curvature (%)	0.92	1.09
Horizontal Curvature (1/R)	0.92	1.09

Model Development : Result

Type: 6 Lane

$$Y = EXPO \times \exp(-5.769 + 0.027X_2 - 0.450X_3 - 0.009X_4 + 0.568\,X_5)$$

Variable name			Four or Less
Constant		Parameter	-5.769
		t-statistic	-10..425
		P-value	0.000
Expo (LengthXADTX365X10^{-6})	X_1	Parameter	Fixed Parameter
		t-statistic	
		P-value	
Density of Driveways (km/Num)	X_2	Parameter	0.027
		t-statistic	6.591
		P-value	0.000
Lane Width (m)	X_3	Parameter	-0.450
		t-statistic	-2.853
		P-value	0.004
Lighting Density (km/num)	X_4	Parameter	-0.009
		t-statistic	-2.456
		P-value	0.014
Ilegal Parking	X_5	Parameter	0.568
		t-statistic	2.892
		P-value	0.004

Log likelihood function	-125.203
Restricted log likelihood	-157.910
α (t-statistic)	0.001
ρ^2	0.207
MPB	-0.073
MAD	1.190
Vuong Statistic	-2.846

Variable Name	Tolerance Limit	VIF
EXPO (LengthXADTX365X10^{-6})	0.92	1.09
Density of Driveways (km/Num)	0.79	1.26
Lane Width (m)	0.83	1.20
Lighting Density (km/num)	0.81	1.24
Illegal Parking	0.81	1.23

Model Development : Result

Type: 8 Lane and Over

$Y = EXPO \times \exp(-4.514 + 0.019X_2 - 1.106X_3 + 0.358X_4)$

Variable name			Eight or More
Constant		Parameter	-4.514
		t-statistic	-4.908
		P-value	0.000
Expo (LengthXADTX365X10^{-6})	X_1	Parameter	Fixed parameter
		t-statistic	
		P-value	
Density of Driveways (km/Num)	X_2	Parameter	0.019
		t-statistic	5.988
		P-value	0.000
Lane Width (m)	X_3	Parameter	-1.106
		t-statistic	-4.050
		P-value	0.001
Land-use (Commercial)	X_4	Parameter	0.358
		t-statistic	2.958
		P-value	0.031

Log likelihood function	-187.007
Restricted log likelihood	-229.543
α (t-statistic)	0.003
ρ^2	0.185
MPB	0.185
MAD	1.833
Voung	-2.984

Variable Name	Tolerance Limit	VIF
EXPO (LengthXADTX365X10^{-6})	0.91	1.08
Density of Driveways (km/Num)	0.84	1.18
Presence of Vertical Curvature (%)	0.92	1.09
Land-use (Commercial)	0.92	1.09

Model Development : Result

- A statistical package called LIMDEP 8.0 was used to analyze the data. To decide which model fits the data better, the Poisson regression model, the Negative Binomial regression model, and the Zero Inflated Poisson regression model were tested.
- The Vuong statistic is computed as less than -1.96 due to the fact that there are few 0 values in the data set.
- Also, the overdispersion parameter statistic(=t-value of α) is less than 1.96, so the Poisson regression model is better than the Negative Binomial regression model in this study.

Model Development : Result

- When the value of ρ^2 is between 0.2 and 0.4, the fit is excellent. In this study, however, there is a case where ρ^2 is less than 0.2, Even though the value is not statistically significant, the model is still applicable because it can be logically explained.
- Multicollinearity testing was done to ensure the independence of the independent variables. Multicollinearity doesn't occur when the tolerance limit is close to 1 and the VIF is under 10.

Conclusions

- The zero count data portion is under 10% for all categorized type in this study. So, ZIP model doesn’t fit our dataset.
- The number of driveways are included in the three different highway groups – 4-lane and under, 6-lane, and 8-lane and over.
- It means that to reduce the vehicle-to-vehicle conflicts is important the design of urban highways.
- In the 4-lane and under highway group, as the horizontal curvature gets smaller and vertical curvature gets greater, more traffic accidents happen.

Conclusions(Cont.)

- For the 6-lane highway group, more accidents happen with more illegal street parking that is likely to increase the accident potential. Because, the drivers to avoid the illegally parked vehicles bring about sudden lane changes.
- For the 8-lane and above highway group, more accidents occurs when the land use pattern next to the highway is commercial.
- There exists a limitation in this study due to the fact that human factors are not considered in the development of the prediction model.

Thank you!

Questions?

E-mail :
mendota@uos.ac.kr