

# **ANALYSIS OF PEDESTRIAN-VEHICLE CRASHES IN KOREA:**

## **Focused on developing probabilistic pedestrian fatality model**

**Cheol Oh, Youn-soo Kang, Beomil Kim**

Center for Advanced Transportation Technology

The Korea Transport Institute

**Wonkyu Kim**

Hankuk Aviation University, Korea

Paper number: 05-0131

### **ABSTRACT**

Microscopic analysis of pedestrian-vehicle accident data is a backbone of devising various intelligent functionalities of vehicles to mitigate the fatality and injury severity of pedestrian in pedestrian-vehicle crashes. Worldwide significant effort has been directed at developing advanced vehicles for protecting pedestrian by the assistance of analyzing very detailed pedestrian accident data. As a part of the multi-year project titled with 'Development of Advanced Vehicle for Pedestrian Protection', this study analyzes pedestrian-vehicle crash data. Firstly, overview of the characteristics of pedestrian-involved crashes in Korea is presented. Another major focus of the study is to develop a probabilistic pedestrian fatality model. The logistic regression approach, one of the multivariate statistical modeling techniques, is applied in the model development. The developed model is expected to support various safety policies and evaluations of advanced systems of vehicles toward enhancing pedestrian safety. The findings of this study would be an invaluable linkage between pedestrian accident data and the development of various countermeasures for pedestrian protection.

### **INTRODUCTION**

A variety of research efforts including accident

data analyses have been performed to enhance safety on highways significantly. In particular, the most valuable finding from analyzing pedestrian accident data would be to discover effective countermeasures to alleviate both fatality and injury severity of pedestrian. Among various safety-related studies based on accident data analysis, to deal with pedestrian would be considered one of the keen interests. It is mainly because pedestrians are the most vulnerable element in transportation systems, which need to be primarily protected. From this perspective, thorough understanding of causes and effects of pedestrian accidents is a fundamental to prepare for feasible solutions to save human lives.

According to a study dealing with comparing accident statistics [1], it has been identified that Korea has been highly ranked in terms of frequency, fatality, and injury severity among OECD (Organization for Economic Cooperation and Development) countries. Hence, more rigorous and active efforts need to be performed to avoid such dishonor, which motivate this study. This study focuses on analyzing pedestrian-vehicle crashes, which ultimately attempts to find underlying clues to reduce both pedestrian fatality and injury severity. Two view points are taken into consideration in this study. Firstly aggregated pedestrian accident statistics is further analyzed to identify the accident

characteristics. The second view of this study is to develop a probabilistic pedestrian fatality model using data obtained from accident reconstruction. The model would contribute to evaluating advanced safety systems of vehicles and deriving safety policies.

After pedestrian accident data are analyzed statistically to discover the implications on safety policies, how to develop a model of collision speed in a statistical manner is presented. The binary logistic regression modeling approach is employed in this study. Then, the interpretation of the model is given with discussions about findings and future research. Conclusions are described along with the direction of future research.

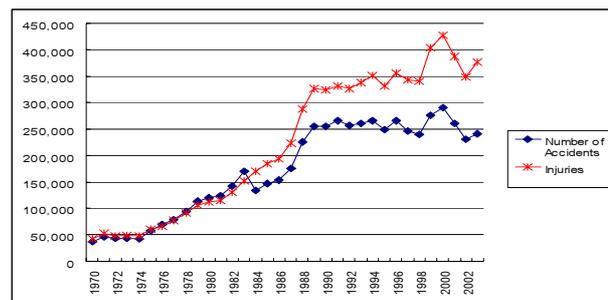
## OVERVIEW OF PEDESTRIAN ACCIDENT IN KOREA

The total number of accidents a year has steadily declined since 1970 in Korea as shown in Figure 1. Furthermore, 1990 can be seen as a turning point of trends in both fatalities and injuries, which shows the trend does not increase rapidly any more. It can be seen that huge national efforts to reduce the accident occurrence have been effective.

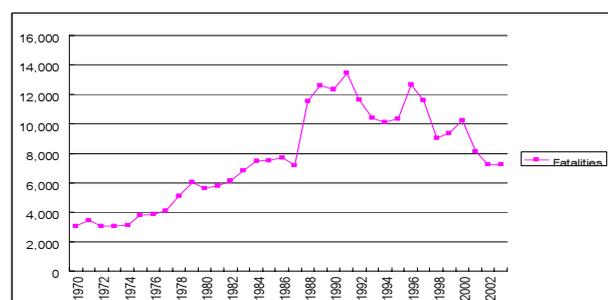
Although the transportation safety environment represented by a couple of accident statistics as the above seems to be improved, we still should pay attention to the international accident statistics. It is stated by the literature [2] that the number of accidents per 10,000 vehicles in Korea is the highest, in comparing with other OECD countries. Figure 2 depicts more details on the comparison.

Another important statistics that should be concerned is the pedestrian-involved accident.

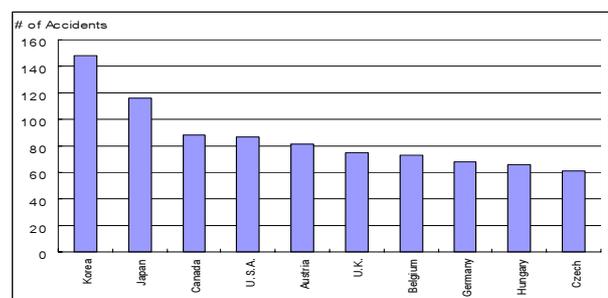
Among 30 OECD countries, Korea has been unfortunately positioned at the first rank, as shown in Figure 3, in comparing the death rate of pedestrian, which is approximately 43.0%. Therefore, it is truly apparent that protecting pedestrian in the transportation system needs to be taken care of with the highest priority by researchers and practitioners.



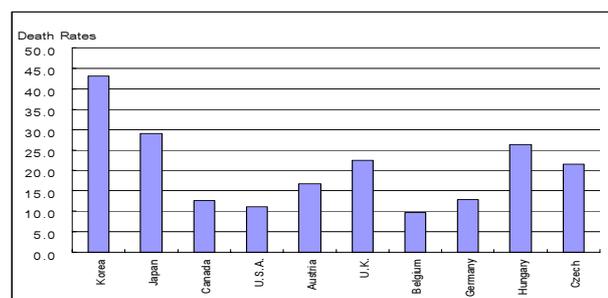
<Figure 1-(a) Trends of accidents and injuries>



<Figure 1-(b) Trend of fatalities>

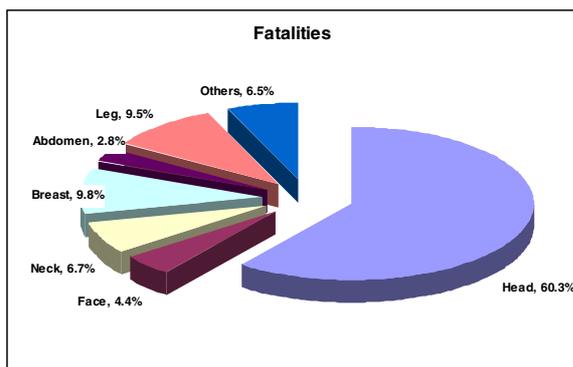


<Figure 2 International comparison of accidents per 10,000 vehicles>

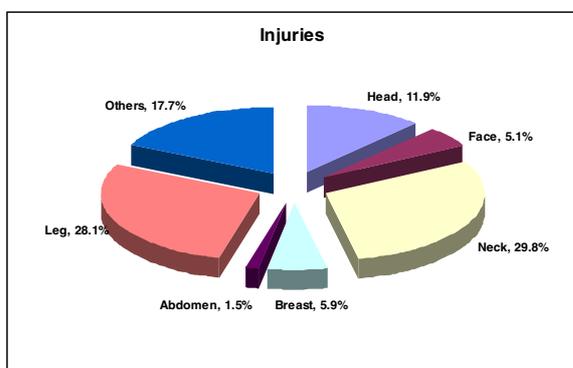


<Figure 3 Death rates of pedestrian for OECD countries>

Identifying injury body regions of pedestrian would be invaluable for further development of countermeasures to protect pedestrian that is the most vulnerable element than any other one in transportation systems. Head has been identified as the most critical injury region resulting in fatalities. In 2003, 1,745 pedestrians passed away mainly due to head injury in pedestrian-vehicle crashes, which corresponds to 60.3% of total pedestrian fatalities. Chest injury was another major body region resulting in pedestrian fatalities. On the other hand, neck and leg region have been recognized as major body region for the pedestrian injury. Neck and leg regions occupy 57.9% of pedestrian injury accidents. Details of injury body regions for fatalities and injuries are presented in Figure 4.



<Figure 4-(a) Injury body regions: Fatalities>



<Figure 4-(b) Injury body regions: Injuries>

Insight into the injury body regions provides us with a precious opportunity of how to cope with

pedestrian accidents. It can be concluded that special cares on both the head part including neck and the leg part needs to be performed. An example of countermeasures toward protecting pedestrian would be to develop an advanced vehicle capable of protecting head and leg of pedestrians.

## DATA PREPARATION FOR MODEL DEVELOPMENT

An accident report form, which was developed for this study, was distributed various agents that are in charge of reporting accidents on June of 2003. The form includes information on pedestrian, vehicle, highway, and environment. A total of 92 pedestrian accident cases have been collected as of now (December, 2004). Acquired accident cases were analyzed by National Institute of Scientific Investigation (NISI) and Center for Accident Analysis of Hanyang University, specialized in accident reconstruction. A major outcome of accident reconstruction was that collision speeds in pedestrian-vehicle crashes were discovered to be further utilized in establishing a model of pedestrian fatality.

As a part of accident reconstruction, estimating collision speed is of keen interest. Useful applications using the model of collision speed are plentiful. One of the nice examples is to be utilized in exploring the analysis of pedestrian trajectories in pedestrian-vehicle crashes, which can be further applied to develop advanced safety systems of vehicles for pedestrian protection. Also, researchers and engineers in the field of pedestrian safety have been recently working on establishing a global technology regulation to protect pedestrians in pedestrian-vehicle crash. Analyses on accident data associated with colliding pedestrians with vehicles are also the backbone of developing the regulation. Recently,

Korea has joined the pedestrian working group for international harmonization activities.

Various studies were performed for estimating more reliable and accurate collision speed. As examples, three methods used widely to estimate collision speed are introduced.

• Schmidt and Nagel (1971)

Schmidt and Nagel [3] found that collision speed is related with the distance from initial collision spot to final location of pedestrian on the ground.

$$V_x = \sqrt{\mu^2 \times h + e - \mu \times h}$$

where

$V_x$  : collision speed (m/s)

$\mu$  : coefficient of friction

$h$  : height of center of pedestrian weight

$d_i$  : distance from initial collision spot to final location of pedestrian

$e$  :  $2 \times \mu \times g \times d_i$

$g$  : gravitational constant of 9.8 m/s<sup>2</sup>

• Stcherbatchef et al. (1975)

It was revealed by Stcherbatchef et al. [4] that the distance from initial collision speed to final location of pedestrian is a function of the collision speed and the deceleration rate of vehicle.

$$d_i = \frac{V_e^2 - V_i^2}{2a} + kV_x$$

where

$k = \lambda \times a$ ,  $\lambda = 0.03$

(empirical value found by experiments)

$a$  : deceleration rate of vehicle (m/s<sup>2</sup>)

$V_i$  : initial vehicle speed (m/s)

$V_e$  : final vehicle speed (m/s)

$t$  : time (sec)

• Collins and Moris (1979)

Collins and Moris [5] reported that the distance from initial collision spot to final location of pedestrian can be estimated by the collision speed and the height of center of pedestrian weight.

$$d_t = \frac{V_x \times \sqrt{h}}{7.97} + \frac{V_x^2}{254 \times \mu} (m)$$

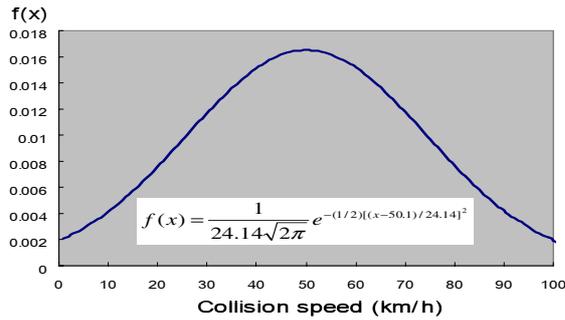
$$x = \frac{-B \pm \sqrt{B^2 - 4 \times A \times C}}{2 \times A}$$

$$A = \frac{1}{254 \times \mu}, B = \frac{\sqrt{-h}}{7.97}, C = -d_i$$

Basic statistics on collision speeds derived from the accident reconstruction are given in Table 1. Additionally the estimated normal distribution of collision speed obtained from accident reconstruction in this study is depicted in Figure 5.

Table 1 Statistics on collision speed

Mean	Std. Error of Mean	Median	Mode	Std. Deviation	Min	Max
<b>50.10 kmp</b>	<b>2.28</b>	<b>45.00</b>	<b>40.00</b>	<b>24.14</b>	<b>4.00</b>	<b>103.00</b>



<Figure 5 Normal distribution of collision speed>

### MODEL DEVELOPMENT: PROBABILISTIC PEDESTRIAN FATALITY MODEL

A statistical model, which is able to measure fatality of pedestrian struck by vehicle in a probabilistic manner, was developed in this study. The developed model relates the fatality of pedestrian in pedestrian-vehicle crashes to collision speed, vehicle characteristics and pedestrian characteristics. The model is expected to be very useful for researchers and practitioners in the field of both transportation safety and automobile. Firstly, the model could assist in estimating safety benefits quantitatively in case that a certain countermeasure associated with collision speed is employed for safety enhancement. An example includes that the effect of change in speed limit, which established from the consideration of collision speed and its effect on fatality, can be more comprehensively quantified. Taking vehicle characteristics into consideration in modeling pedestrian fatality could provide invaluable implications on developing new vehicles for pedestrian protection. Of course, it should be noted that extensive data analyses are required in order to obtain reliable outcomes prior to drawing conclusions.

A logistic regression modeling approach was applied to the pedestrian fatality problem. The

relationship between a dependant variable, which is a non-metric variable in particular, and one or more independent variables is modeled. Unlike discriminant analysis, when the dependant variable has only two groups, logistic regression may be preferred for several reasons [6]. One of the major reasons is that discriminant analysis relies on strictly meeting the assumptions of multivariate normality and equal variance, which logistic regression does not face such an assumption. In a case of binary logistic regression, 1 or 0 is taken as a dependent variable in modeling. The binary logistic regression model predicts the probability that the dependent variable would take 1. Therefore, what the dependant variable takes 1 means that the probability of pedestrian death is 1 in this application.

The form of a logistic regression model in this study is

$$P ( F_i = 1 | X_i ) = \frac{\exp [ f ( X_i , \beta ) ]}{1 + \exp [ f ( X_i , \beta ) ]}$$

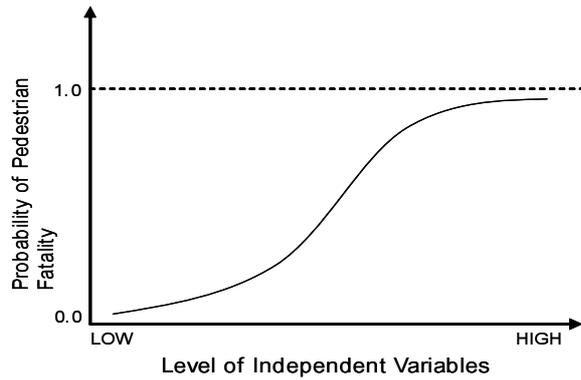
where

$F_i$  : dependant variable representing the pedestrian death ( $F_i = 1$ ) or survival ( $F_i = 0$ ) of the event for pedestrian-vehicle crash  $i$

$X_i$  : vector form of selected independent variables affecting pedestrian fatality

$f ( X_i , \beta )$  : a function of  $X_i$  and a parameter vector  $\beta$  to be estimated

Fatality probability values can be any value between 0 (survival) and 1 (death), but the predicted fatality value can not be out of the range of 0 and 1. To define a relationship bounded by 0 and 1, logistic regression uses an assumed relationship between the independent and dependent variables that resembles an S-shaped curve as shown in Figure 6 [6].



**<Figure 6 The logistic relationship between pedestrian fatality and independent variables>**

SPSS statistical analysis package was employed in modeling the pedestrian fatality. A variety of independent variables in the form of accident report

were considered as candidates for representing the pedestrian fatality. Those include collision speed and characteristics of pedestrian, vehicle, highway and environment. Details are presented in Table 2. Based on large efforts on modeling, three independent variables which are collision speed ( $V_x$ ), vehicle type ( $VT$ ), and pedestrian age ( $AGE$ ), were chosen to build up a probabilistic pedestrian fatality model. To investigate the potential of multicollinearity among the variables, the correlation analysis was conducted. As shown in Table 3, any pair of variables did not show high correlation, which correlation coefficients exceed 0.5. Some descriptive statistics of the independent variables are also given in Table 3.

Table 2 Candidates of independent variables

Item	Detailed information
General	Time and location of accident
Pedestrian	Gender, Age, Height, Weight, Drinking, Injury severity, Contact points (1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> )
Vehicle	Vehicle type(1: passenger car, 2: suv, 3: 1box, 4: bus/truck ), Production year
Highway and Environment	Super-elevation, Pavement condition, Accident location (intersection, mid-block, pedestrian crossing), Speed limit, Contact points (1 <sup>st</sup> , 2 <sup>nd</sup> , 3 <sup>rd</sup> ), Weather condition

Table 3 Correlation coefficients and descriptive statistics of the independent variables

Variable/Statistic	$V_x$	$VT$	$AGE$
Correlation Coefficient	$V_x$	1.000	-0.074
	$VT$	-	1.000
	$AGE$	-	-
Mean	50.100	40.123	1.384
Standard deviation	24.138	0.892	17.853
Minimum	4.000	1.000	6.000
Maximum	107.000	4.000	75.000
Skewness	0.181	2.295	-0.053
Kurtosis	-0.580	3.959	-0.622

Table 4 summarizes the results of the pedestrian fatality model with the specification of a binary logistic regression. The model pertains to the effect of collision speed, vehicle type, and age on the pedestrian fatality. One of the major findings from the analysis is that we statistically verified collision speed is the most significant variable affecting the pedestrian fatality with the largest Wald statistic and the lowest significance level. The model was able to correctly classify 84.4 percent of the pedestrian-vehicle crashes as pedestrian death or survival, with a classification cutoff threshold of 0.5 membership value.

Table 4 Results of logistic regression

Statistic	$V_x$	$VT$	$AGE$	Constant
Wald statistic	13.700	3.252	2.499	7.880
Standard error	0.025	0.381	0.025	2.094
$\beta$	0.092	0.687	-0.039	-5.878
Significance	0.000	0.071	0.114	0.005

- Correct classification rate: 84.4%
- Model chi-square: 37.562
- -2 log likelihood: 40.287
- Nagellerke R-square: 0.631

## DISCUSSION OF RESULTS

### Findings

The logistic regression model developed in this paper predicts the probability of pedestrian fatality in pedestrian-vehicle crashes. The variables employed as the independent variables of the model and their coefficients are listed in Table 4. Three independent variables include collision speed, pedestrian age, vehicle type. Collision speed denoted as  $V_x$ , which was recognized as the most significant factor to determine the pedestrian fatality, represents the situation of accident. Age and vehicle type can be viewed as the representatives of pedestrian characteristics and vehicle characteristics, respectively. As shown in Table 4, it can be interpreted by identifying the negative coefficient of  $AGE$  that as pedestrian age increases the probability of fatality decreases. That is, children are exposed to higher fatality in the transportation system. In addition, heavy vehicles have more influence on the fatality than light vehicles.

In a view point, the findings would be natural. However, the contribution is that the model allows us to quantify the pedestrian fatality in a probabilistic manner, which can assist in measuring the safety effects of countermeasures. Further analysis of the probability of pedestrian fatality can be also

converted into monetary value, which leads us to be able to conduct economic appraisals of countermeasures. As an example, the model would support how to estimate the benefits of brake assistant systems (BAS) resulting possibly in reducing collision speeds.

### Future Research

In addition to the considerations in developing the probability model of pedestrian fatality, we still have a variety of research opportunities to improve model development and to derive countermeasures for the pedestrian protection. First of all, a lot of crash data should be collected and further analyzed in order to obtain more statistically reliable and accurate model. Regarding the selection of independent variables, vehicle types can be divided into more substantial variables representing vehicle characteristics. Parameters representing the frontal shape of vehicles can be considered independent variables of the model. Those parameters include bumper lead (BL), bumper center height (BCH), leading edge height (LEH), bonnet length (BL), and bonnet angle (BA) etc. If those parameters can be incorporated into the fatality model, we will be able to see how vehicle frontal shapes affect the pedestrian injury. Establishing fine and huge crash dataset is a matter of course as a prerequisite for future research.

## CONCLUSION

Overview of the characteristics of pedestrian-involved accidents in Korea was presented. The major focus of the study was to develop a probabilistic pedestrian fatality model. The logistic regression approach, one of the multivariate statistical modeling approaches, was applied in the model development. The developed model is expected to support various safety policies and evaluations of advanced systems of vehicles to protect pedestrian. Collision speed, pedestrian age, and vehicle type were used as independent variables of the model. Findings and future research were also discussed. In-depth further study with a lot of crash data will be performed in the future.

## ACKNOWLEDGMENT

This study was conducted as a part of the multi-year project titled with “Development of Advanced Vehicle for Pedestrian Protection” funded by the Ministry of Construction and Transportation (MOCT) under the Korean government.

## REFERENCES

- [1] *Statistical Analysis on Traffic Accidents*, Korea Road Traffic Safety Authority, 2004.
- [2] *A Comparative Study of Accident Statistics for OECD Countries*, Center for Accident Analysis, Korea Road Traffic Safety Authority, 2004.
- [3] Schmidt, D. N, and Nagel, D. A. (1971), ‘Pedestrian Impact Case Study’, *Proceedings of 15<sup>th</sup> Conference Association for Automotive Medicine*.
- [4] Stcherbatcheff, G. et al. (1975), ‘Simulations of Collisions between Pedestrian and Vehicles using Adult and Child Dummies’, *SAE Paper*, NO. 751167.
- [5] Collins, J.C., and Moris. J. L. (1979), ‘*Highway*

*Collision Analysis*’, Thomas Publishing.

- [6] Hair, J.F. et al. (2003), ‘*Multivariate Data Analysis*’, Pearson Education.