An Injury Severity Analysis using Korean In-Depth Crash Data

Eunbi Jeong, Seolyoung Lee, Cheol Oh
Hanyang University

* This research was supported by a grant from the Transportation & Logistics Research Program, funded by Ministry of Land, Infrastructure and Transport Affairs of the Korean Government (Project No.: 15PTSI-C054118-07).

I. INTRODUCTION

A variety of contributing factors, such as drivers, vehicles, geometric conditions, environments and traffic conditions, are associated with crash occurrence. To date, significant research efforts have been made to identify the factors influencing traffic crashes in order to prevent crashes and mitigate injury severity. However, the many limitations associated with crash records do not allow us to analyze crash data in close detail. These limitations include missing information, inaccurate crash location, and inaccurate time, traffic, vehicular and weather information. To overcome these limitations, significant efforts have been made to collect in-depth crash data, which is more detailed crash data that integrates local-crash, vehicle-crash data and emergency medical service data.

This study introduces the characteristics of the recently established in-depth crash database, set up as part of a national research project funded by the Korean Ministry of Land, Infrastructure and Transport. In addition, crash severity analyses are conducted to identify contributing factors, which will be used to develop useful countermeasures. In order to do this, the frequency analysis and cross analysis are conducted to characteristics analysis of Korean in-depth crash data. Following this, an ordered logistic regression (OLR) is used to develop the injury severity statistical model.

II. METHODS

Data

This study focuses on freeway traffic crashes because crashes occurring on freeways are associated with high fatality rates and severe injuries. A total of 140 freeway front-collision data in a recent five-year period (2011–2015) were used in this study. The Korean in-depth crash data contains useful information on crashes. The crash record consists of three main categories. (1) Location-specific base information includes time, location and geometric conditions. (2) Vehicle information includes vehicle type, damage of vehicles and use of in-vehicle safety measures, such as seatbelt, airbags and advanced safety measures like adaptive cruise control. (3) Medical service data, including characteristics of victim and the level of injury severity, are also included in crash record files. Table I shows a set of variables to be used for statistical analyses, which is derived from the in-depth crash records. These independent variables include vehicle characteristics, roadway geometric conditions and environmental conditions. This study classifies injury severity into four levels: minor (ISS 0–9), moderate (ISS 10–15), severe (ISS 16–24), and critical (ISS$\geq$25) injury.

TABLE I
VARIABLE DESCRIPTIONS

Variable	Variable name	Descriptions
Injury	Injury severity	Minor (0), Moderate (1), Severe (2), Critical (3)
	Counter vehicle type	PC, SUV, Truck, Van, Bus
	Crush extent	Continuous
Victim-related	Gender	Female (0), Male (1)
	Height, Weight	Continuous (cm, kg)
	Location of boarding	Driver, Passenger, Rear
Vehicle-related	Vehicle type	PC, SUV, Truck, Van, Bus
	Safety belt	No (0), Yes (1)
	Front airbag	No (0), Yes (1)
Environments	Weather	Normal, Rain, Snow, Fog

Ordered Logistics Regression

This study develops a statistical model that predicts injury severity of crash victims. As crash injury severity is intrinsically ordered, the model must account for ordinal characteristics. For this reason, an ordered logistic regression technique is appropriate. The ordered model can be specified as $y_i^* = \beta' \ x_i + \varepsilon_i$, where y_i^* represents the injury severity of the crash victim and x_i represents the independent variables that affect injury severity. The probability that the injury severity y with the independent variable x_k is less than j is estimated using Equation (1):

$$\Pr(y \leq j|x) = \frac{e^{\mu_j - \sum_{k=1}^{K} \beta_k x_k}}{1 + e^{\mu_j - \sum_{k=1}^{K} \beta_k x_k}} \qquad (1)$$

More detailed information on ordered logistic models and the relevant derivations can be found in the literature (Maddala, 1983; Liao, 1994).

III. INITIAL FINDINGS

SPSS, a well-known statistical software package program, was used to predict injury severity. The results of the model development are summarized in Table II. The value of Chi-square for the logistic regression model is 44.96. A total of seven variables were identified as statistically significant based on evaluating Wald-statistic tests (sig. probability ≤0.1). The signs of the coefficients of the independent variables for front seats, including driver and passenger seats and crush extent, were positive. On the contrary, use of seatbelt, activation of front airbag and larger vehicle type showed positive coefficients. These results imply that the injury severity decreases when a driver is driving a larger vehicle equipped with in-vehicle safety measures. The odds ratio for safety belt was 0.232, which implies that when a driver fastened the safety belt, the probability of injury severity would be 0.232 times less than without safety belt. The probability model with the estimated coefficients is shown in Equations (2)–(5).

TABLE II
RESULTS OF ORDERED LOGISTIC REGRESSION

Parameter	β	Standard error	Wald statistic	Sig. Prob.	Odds ratio
μ_1	3.616	0.799	15.657	0.000	-
μ_2	3.986	0.830	23.039	0.000	-
μ_3	5.024	0880	32.614	0.000	-
Driver seat (X_1)	2.114	0.696	9.230	0.002	8.281
Passenger seat (X_2)	1.635	0.756	4.680	0.031	5.129
Safety belt (X_3)	-1.459	0.491	8.844	0.003	0.232
Front airbag (X_4)	-0.915	0.543	2.846	0.092	0.401
SUV (X_5)	-1.213	0.662	3.359	0.067	0.297
Truck (X_6)	-1.336	0.613	4.746	0.029	0.263
Crush extent (X_7)	0.605	0.125	23.364	0.000	1.831

$$\Pr(Minor) = \frac{\exp(3.616+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)}{1+\exp(3.616+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)} \qquad (2)$$

$$\Pr\ (Moderate) = \frac{\exp(3.986+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)}{1+\exp(3.986+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)} - \Pr(Minor) \qquad (3)$$

$$\Pr\ (Severe) = \frac{\exp(5.024+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)}{1+\exp(5.024+2.114X_1+1.635X_2-1.459X_3-0.915X_4-1.213X_5-1.336X_6+0.605X_7)} - \Pr\ (Minor) - \Pr(Moderate) \qquad (4)$$

$$\Pr\ (Critical) = 1 - \Pr\ (Minor) - \Pr\ \ (Moderate) - \Pr(Severe) \qquad (5)$$

IV. DISCUSSION

In general, in-depth crash data provide valuable insight into the development of countermeasures for reducing injury severity and establishing relevant traffic safety policies. Although we were not able to develop a very insightful statistical model to analyze injury severity, mainly due to limited crash samples, continuous efforts should be made to refine the developed injury severity statistical model in order to enhance the applicability of the Korean in-depth crash data. More systematic analyses need to be conducted with larger dataset for more generalized and insightful results.

V. REFERENCES

[1] Jeong, E. *et al*., IJAT, 2013.
[2] Choi, S. *et al*., *Traffic Inj. Prev*., 2014.
[3] Liao, T. F., *Appli. In the Social Sci*., 1994.
[4] Maddala, G.S. Cambridge Univ., 1983.