International Journal of Engineering and Science Application
ISSN 2406-9833

IJEScA

A Relationship Model between Accident Factors and the Traffic Accident Severity Using Logistic Regression Model

[1]H. Halim[1], [2]M.I. Ramli, [2]S.A. Adisasmita, [3]S.H. Aly, [4]J. Prasetijo

[1]Doctoral Student, Department of Civil Engineering, Faculty of Engineering, Hasanuddin University, Jl. Poros Malino Km. 6 Gowa, 92172, Indonesia
[2]Lecturer, Department of Civil Engineering, Faculty of Engineering, Hasanuddin University
Jl. Poros Malino Km. 6 Gowa, 92172, Indonesia
[3]Lecturer, Department of Environmental Engineering, Faculty of Engineering, Hasanuddin University
Jl. Poros Malino Km. 6 Gowa, 92172, Indonesia
[4] Lecturer, Department of Rail Transportation, Faculty of Engineering Technology, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia

email: hasmar29@gmail.com (correspondence author).

ABSTRACT

The present paper purposes to develop the relationship model between the factors of accidents and severity level of traffic accidents by using multinomial logistic regression model approach, for a case study the traffic accident in Makassar City, Indonesia. In further, the study evaluates the traffic accident factors which significantly influence the traffic accident severity level. In this regard, the outcome variable is the severity level of the traffic accident which has three attributes, i.e., death, serious injury, and minor injury. The explanatory variables involve victim characteristics and traffic accident characteristics. The present study used the traffic accident database during 2012 – 2015 which recorded by the traffic police agency in the city. The model calibration results show that the relationship model has a good accuracy level. The victim position and the collision types significantly influence the severity accident level. The results provide basic information for efforts in reducing traffic accidents.

Keywords: Accident factors, traffic accidents severity, logistic regression model.

1. INTRODUCTION

A traffic accident is an occurrence that is hard at guessing where and when they occurred involving road users and impact damages or fatalities. According to Ngo et al. (2012) [1], the number of accidents in Vietnam recorded 15,000-18,000 deaths due to traffic accidents. In Thailand, the death toll reached 13,000 and 1,000,000 vehicles were an injury in traffic accidents [2]. Particularly in Indonesia, 61% mortality caused by traffic accidents is caused by drivers of two-wheeled and three-wheeled in three provinces [3].

Accidents happen influenced by the human factor, vehicle and road/environment as well as the interaction and combination of two or more of these factors [4]. Accidents can also be influenced by the age of a driver, young driver a range of traffic accidents [5]. In other studies, revealed that there is a significant relationship between demographic conditions of the driver and the distance traveled by accident [6]. Accidents are also strongly influenced by the behavior of drivers. Fitroh et al. (2015) examined the behavior among others, argues that there are differences between the behavior of drivers, men, and women against accidents [7].

The severities as the impact of the crash victims are composed of several levels, namely death, serious injuries, minor injuries, and properties damage [8]. Studies on the severity of traffic casualties, among other things stated by Zhang, Yau, & Chen (2013) [9]. The severity of traffic accidents can be influenced by sex, street lighting conditions, seat belts, weekday, and time of the accident vehicle age is a significant factor determining severity due to accidents [10]. Factors Overloading and obstruction is the most significant factor in the severity of the victim in Ghana [11]. Other studies have shown that not wearing a seatbelt, overtaking and speeding is the most important factor associated with the severity of the injury [12].

The model that is widely used in transportation research to explain the relationship between the variables is the logistic regression model. Logistic regression is used to estimate the effects of a statistically significant factor. Logistic regression distinguished between binomial and multinomial [13, 14, 15, 16].

The present study purposes to overview of the characteristics of victims of traffic accidents, the model the severity of traffic accident victims based on the factors that influence it, and a high risk of severity of each of the victims of traffic accidents to the accident victims in Makassar City in 2015 using logistics regression model approach.

2. THE STUDY METHODS

Often in a study, the researchers wanted to model the relationship between variables X (independent) and Y (dependent). The method most often used in such cases is a linear regression or Multicollinearity. However, sometimes linear regression with OLS (Ordinary Least Square) are less suitable for use. Said to be less suitable because if the linear regression used to be a violation of the Gauss-Markov assumptions. In cases where the response variable (Y) has the type of nominal data, while the predictor variable (X) has the type of interval data or ratio [17].

The use of logistic regression method is highly recommended to overcome the problem. As usual regression, logistic regression consists of two types, namely: Binary Logistic Regression and Multinomial Logistic Regression. Binary logistic regression is used when there are only two possibilities for the response variable (Y) for example buying and not buying. Multinomial Logistic Regression is used when the response variable (Y) is not more than two categorizations. Logistic regression is not limited only apply in cases where the interval ratio variable type X or alone. But logistic regression can also be applied to a case where the variable X nominal or ordinal data type. The variable is by the linear regression with dummy variables.

Multinomial Logistic Regression can also be used to analyze the relationship between one or more independent variables both nominal scale and categorical outcome data from more than two categories. So, the difference between binary logistic multinomial logistics lies only in the number of categories that the desired result.

2.1 Multinomial Logistic Regression

Logistic regression equation can be described by the following equations [18]:

$$P(Y = j|x) = \mu_j(x) = \frac{[g_j(x)]}{\sum_{h=0}^{2}[g_h(x)]} = \frac{(\beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p)}{\sum_{h=0}^{2} e^{(\beta_{h0} + \beta_{h1}x_1 + \beta_{h2}x_2 + \cdots + \beta_{hp}x_p)}}$$

……….1)

Where: $\beta_0 = 0$ than $g_0(x) = 0$

P (Y = $j|x$) = conditional probability of response variable j in the vector x

$\mu_j(x)$ = logistic regression equation for the response variable j

$g_j(x)$ = logit in the dependent variable j, $j = 0,1,2$

x_m= the value of the independent variable to - m, $m = 1,2,3, \ldots, p$

β_{jm} = coefficients/parameters of the model

Thus, if the outcome of the dependent variables in the form of three categories that are coded as 0, 1 and 2, then the equation is as follows:

$$P(Y = 0|x) = \mu_0(x) = \frac{1}{1 + e^{(\beta_{10} + \beta_{11}x_1 + \cdots + \beta_{1p}x_p)} + e^{(\beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p)}}$$

….2)

$$P(Y = 1|x) = \mu_1(x) = \frac{e^{(\beta_{10} + \beta_{11}x_1 + \cdots + \beta_{1p}x_p)}}{1 + e^{(\beta_{10} + \beta_{11}x_1 + \cdots + \beta_{1p}x_p)} + e^{(\beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p)}}$$

….3)

$$P(Y = 2|x) = \mu_2(x) = \frac{e^{(\beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p)}}{1 + e^{(\beta_{10} + \beta_{11}x_1 + \cdots + \beta_{1p}x_p)} + e^{(\beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p)}}$$

….4)

From the equation, it is known that the dependent variable with three categories will form two logit equations, where each equation is forming a binary logistic regression to compare a group of the reference category, is as follows:

$$g_1(x) = \ln\frac{P(Y=1|x)}{P(Y=0|x)} = \ln\frac{\mu_1(x)}{\mu_0(x)} = \beta_{10} + \beta_{11}x_1 + \cdots + \beta_{1p}x_p$$

…………………………………………5)

$$g_1(x) = \ln\frac{P(Y=2|x)}{P(Y=0|x)} = \ln\frac{\mu_2(x)}{\mu_0(x)} = \beta_{20} + \beta_{21}x_1 + \cdots + \beta_{2p}x_p$$

………………………………………...6)

So generally, the shape of the logit function with variable response consists of three categories is:

$g_j(x) = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \beta_{jp}x_p$; $j = 0, 1, 2$

If there is the independent variable in the form of categorical data with more than two categories, the necessary transformation by inserting dummy variables into the model. Suppose an explanatory variable to- "m", namely "*Xm*" which have many categories as "*hm*", there will be a dummy variable as "*hm* - 1". Thus, logistics functions with explanatory variables and a dummy variable would be [19]:

$$g_j(x) = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \cdots + \sum_{v=1}^{h_m-1} \beta_{jmv} D_{jmv} + \cdots + \beta_{jp} x_p$$

……………………………………..7)

where D_{jmv} are dummies of variables to *-m* logit function to*-j*

Analysis of the data in this study, using multinomial logistic regression. In general, the steps undertaken in the multinomial logistic regression analysis were three steps [20]. Firstly, testing parameters simultaneously to determine the suitability of the analysis model. Secondly, perform partial testing parameter to determine the most influential variables in the model. Thirdly, interpreting the value of the likelihood ratio is formed.

2.2 Parameter estimation

In a logistic regression model, the expected value of the response variable is not linear and has variants that are not the same, so that the β parameter estimates obtained by Maximum Likelihood. The solution for troubleshooting the system of nonlinear equations is the estimating β through prose Newton Raphson iteration. Because the response variable Yi is assumed to be independent, then we obtain the conditional likelihood function for a sample of n observations as follows:

$$l(\beta) = \prod_{i=1}^{n} [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \pi_2(x_i)^{y_{2i}}]$$ ……….8)

Mathematically, it would be easier to get the value that will maximize the likelihood function above through the logs of these functions is the log likelihood. Thus, the log-likelihood function, it is:

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^{n} \left(y_{1i} g_1(x_i) + y_{2i} g_2(x_i) - \ln\left[1 + \exp(g_1(x_i)) + \exp(g_2(x_i))\right] \right)$$

…………….9)

To get the maximum value L (β) then does the differentiation of L (β) with the proviso $\frac{\partial L}{\partial \beta} = 0$ dan $\frac{\partial^2 L}{\partial^2 \beta} < 0$.

2.3 Testing Parameters

Testing of the model parameters is done to examine the role of the independent variables in the model. Tests carried out is twofold [18]:

[1] The test parameters with a likelihood ratio test (simultaneous test or a test G)
G test statistics, the test used to test the role of the independent variables in the model together. The hypothesis tests consist of the following steps:
- H_0: $\beta_1 = \beta_1 = \ldots = \beta_p = 0$ which means there is no influence of a set of independent variables and the dependent variable.
- H1: at least one $\beta_j \neq 0$, which means at least one independent variables that affect the dependent variable.
- The statistical test $G = -2\ln\left[\frac{l_o}{l_k}\right]$ where l_o is likehood without the independent variable and l_k is the

likelihood with the independent variable.

G test statistic follows a chi-square distribution when n approaches infinity with degrees of freedom p where *$p = (r – 1) (c – 1)$*, *r* and *c* each is some categories on the independent variable and the dependent variable. H_0 will be rejected at a significance level α if the value $G > X^2_{p;a}$ or (*p*-value) < α, thus the conclusion that the independent variables jointly or overall affect the dependent variable, it can also be said that at least one coefficient $\beta j \neq 0$. To find out which ones have a significant effect βj do test parameters βj partially with the Wald test.

[2] The test parameters with the Wald test (Test Partial)

Tests conducted variables one by one using Wald test statistic. This test is performed by comparing the best models produced by the simultaneous test of the model without independent variables in the model the best. The hypothesis to be tested are as follows:

- H_0: $\beta_j = 0$, meaning there is no influence of the independent variables to the dependent-j
- H_1: $\beta_j \neq 0$, meaning that there is influence of independent variables on the dependent variable
- Test statistics are $W = \left[\frac{\hat{\beta}_j}{S\ (\hat{\beta}_j)}\right]^2$; $j = 1, 2, \ldots, p$, where $\hat{\beta}_j$ is the estimator of β*j* and $Se(\hat{\beta}_j)$ is the standard error of the estimate β*j*

W assumed to follow a chi-square distribution with degrees of freedom 1. H_0 will be rejected if the value $W > X^2_{(1;\alpha)}$ or (*p-value*) < α. If H_0 rejected, it concluded that significant βj. In other words, the independent variable X is the partially significant effect on the dependent variable.

2.4 Model Interpretation

The likelihood ratio is a measure that estimates how likely it is independent of the dependent variables. The odds ratio is a measure to determine the ratio of the tendency to experience a certain situation between categories with each other in a variable denoted by θ, which is as the ratio of the odds for x = 1 to x = 0. In other words, the risk propensity observation influence x = 1 is m-fold risk compared with observation x = 0, or a tendency to influence the risk of x = 0 is 1 / m-fold compared with the observation x = 1.

For θ = 0 means that x = 1 has a value The odds ratio for Y = j to Y = k calculated at two values (eg x = 1 and x = 0 is:

$$\theta = \frac{F(Y = j|x = 1)/F(Y = k|x = 1)}{F(Y = j|x = 0)/F(Y = k|x = 0)} = \exp[\beta_j]$$

….10)

the same trend with x = 0 to produce Y = j. If 1 <θ < means x = 1 have a greater tendency θ times dibandingakn x = 0 to produce Y = j and vice versa for 0 <θ <1.

3. DESCRIPTION OF MODEL

3.1 Specifications variable

The specification variables of the model in this study consist of two variable categories. Firstly, the severity of traffic accidents as the dependent variable. In this study, the dependent variable is the severity of traffic accidents. This variable is dichotomy into three categories: fatal accident (Y=0), serious injury (Y=1) and a slight injury (Y=3). Needs to be defined that the purpose of the fatal accident is an accident victim who certainly died as a result of traffic accidents within a maximum period of 30 days after the accident. Serious injury is the victim of an accident for injuries suffered permanent disability or hospitalization at the hospital in more than 30 days since the accident occurred. An event is classed as a permanent disability if something limbs are missing or can not be used at all, and cannot be cured or recovered forever. Minor injuries are accident victims who suffered injuries that do not require hospitalization or to be admitted to hospital than 30 days.

The next variable is the independent variable. This variable form of characteristic accidents associated with accident victims, such as gender, the day of the accident, age, time of the accident, the position of the victim, accident type, vehicle type, education, type of collision and location. This factor was the key factor that contributed to the severity of traffic accidents. This factor was taking account of the model as independent variables, namely X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8, X_9, and X_{10}.

Development of a model with descriptive statistical analysis of data to determine the characteristics of victims of accidents and a multinomial logistic regression model to get the factors that affect the severity of traffic accident victims by using IBM SPSS 21 software.

3.2 Description of Data

The data used in this paper is secondary data consisting of 810 data of victims of traffic accidents during the period 2015 in Makassar. This data does not portray in detail about the victims of accidents due to lack of facilities and training for the police to record the accident victims. Besides, it is difficult to obtain because of the medical report of accident data, and medical data are not stored together. Therefore, it is impossible to get details of the severity in this study.To that end, the data used in this study is data from the Traffic Accident Unit Polrestabes Makassar. The variables used in this study include response variable and predictor variables all variables described in the following categories. Determination of the response variable (Y), determined by the purpose of this study was to determine the pattern of the severity of traffic accident victims. In this regard, other predictor variables were category variables, while, the variable response was the severity of traffic accident victims. Variables used as described in Table 1.

Tabel 1. Description of the study variables

Variables	Variable attributes					
Severity of accident (Y)	a)	Death	b)	Serious injury	c)	Minor injury
Gender (X_1)	a)	Male	b)	Female		
Day of the accident (X_2)	a)	Weekday	b)	Off day		
Age (X_3)	a)	Children	b)	Juvenile	c)	Adult
	d)	Elderly				
Time of accident (X_4)	a)	Morning	b)	Daylight	c)	Afternoon
	d)	Night	e)		f)	
The position of the victim (X_5)	a)	Pedestrian	b)	Cyclist	c)	Motorcyclist
	d)	Driver	e)	A ride on a motorcycle	f)	Passenger
Accident types (X_6)	a)	Single Accident	b)	double collision	c)	collision streak
Vehicle types (X_7)	a)	un-motorcycle	b)	Motorcycle	c)	Wheel 3
	d)	Wheel 4	e)	Wheel 6	f)	>Wheel 10
Education (X_8)	a)	Not school	b)	Elementary School	c)	Junior high school
	d)	Senior High School	e)	College		
Type of collision (X_9)	a)	Rear-End	b)	Head-On	c)	Hit and Run
	d)	Hit side	e)	Sideswipe		
Location (X_{10})	a)	straight ahead	b)	3-way junction	c)	4-way junction
	d)	5-way junction	e)	Turn		

4. RESULT AND DISCUSSION

4.1 The traffic accident characteristics in Makassar City

Description of data related to the characteristics of traffic accidents and severity of traffic accidents as in Figure 1.Figure 1 shows that of the total of 810 occurrences of traffic accidents in 2015, accidents resulting in as many as 1,089 people suffered death (11%), serious injury (5%) and minor injuries (84%). Based on the severity of the victim, a dying man as much as 79.8%. The scene of the accident that caused serious injury on weekdays poses a risk of 78.6%. The type of double collision will cause a slight injury 73.3%.

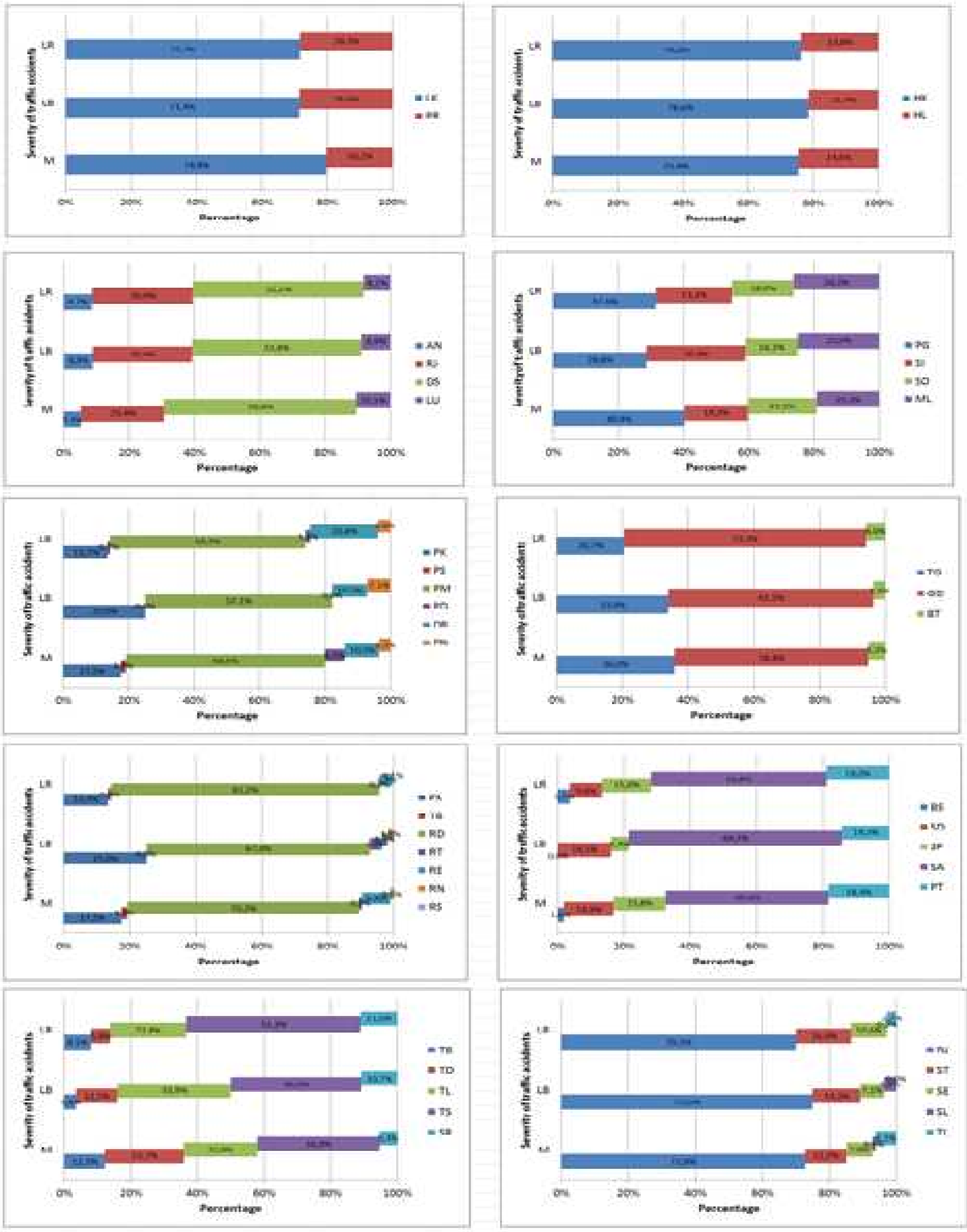

Figure 1. Characteristics of traffic accidents

For these types of vehicles, motorcycles have the highest percentage of the severity of traffic accidents. It can be seen that the motorcycle picture 1 cause of death by 70.2%, amounting to 67.9% serious injuries and slight injuries amounted to 81.2%. While based on the time of the accident, the accident scene on the morning of causing death by 40.4% and causing serious injury was 28.6% and causing slight injuries by 31.6%. This fact is due to the increase of traffic volume in the morning, especially on weekdays, causing a greater risk of accidents.

4.2 Reduction variables for the logistic regression model

The present study conducted three-stage analysis in statistical analysis side to select the variables that taking account in the logistic regression model. The first stages carried out an independence test, the second stage conducted a cross-tabulation analysis regarding the results, and the last stage repeated the first stage analysis regarding the result of the second stage. The analysis results of the three stages analysis are presented in Table 2, Table 3, and Table 4 for each stage, respectively.

Table 2 shows that there was four variables which significant in considering to taking account into the next step analysis. The variables involve the position of the victim (X_5), the type of accident (X_6), vehicle types (X_7), and the type of collision (X_9). In this regard, the four variables have indicator values of P-value which less than 0.05. It means that the variables are significant at the 95% level.

The results of the cross-tabulation analysis as the second stage analysis as seen in Table 3 show that there were some categories which have a number of the event less than 5%. The categories were part of the victim position and the vehicle type variables. Then, we have to reduce the number of categories of both variables to fulfill the 5% requirement. In this regard, the number of categories was a reduction from seven categories to three categories, and from six categories to three categories, for the victim position variable and the vehicle type variable, respectively. The three categories of the victim position are pedestrian, driver, and passenger. As well as, the three categories of the vehicle type are unmotorized, motorcycle, and the other motor vehicles.

Table 2. The first stage analysis result of the independence test

Variables		Pearson Chi-Square	Likelihood Ratio	P-Value
Gender	X_1	3,415	3,613	0,181
Day of the accident	X_2	0,208	0,212	0,901
Age	X_3	3,907	4,085	0,689
Time of accident	X_4	7,147	7,087	0,307
The position of the victim	X_5	29,817	26,931	0,019*
Accident types	X_6	17,617	16,335	0,001*
Vehicle types	X_7	33,235	24,380	0,004*
Education	X_8	13,411	15,945	0,329
Collision types	X_9	58,092	47,383	0,000*
Location	X_{10}	21,721	15,813	0,360

* The significance level of the variables is 95%

Table 4 shows the result of the second stage analysis for the independence test of the left four variables. Table 4 shows that there were only two variables, the victim position (X_5) and the collision type (X_9), which significant in considering to taking account into the next step analysis. In this regard, the four both variables have *P-value* indicators which less than 0.05.

Table 3. Cross-tabulation of the Variable Response with Variable Predictors

Variable Predictors	Severity levels of the traffic accident (%)			Number
	Death	Serious injury	Slight injury	(%)
The victim positions (X_5)				
Pedestrian (0)	1,87	1,31	11,50	14,67
Cyclist (1)	0,19	-	0,56	0,75
Motorcyclist (2)	6,45	2,99	50,19	59,63
Driver (3)	0,65	-	1,21	1,87
Get a ride motorcycle (4)	1,12	0,56	17,48	19,16
Passenger (5)	0,37	0,37	3,18	3,93
Accident types (X_6)				
Single Accident (0)	3,83	1,78	17,38	22,99
Double collision (1)	6,26	3,27	61,68	71,21
Collision streak (2)	0,56	0,19	5,05	5,79
Vehicle types (X_7)				
Unmotorized (0)	2,06	1,31	12,06	15,42
Motorcycle (1)	7,48	3,55	68,32	79,35
Motor tricycle (2)	0,09	0,19	0,28	0,56
Car (3)	0,93	0,09	3,27	4,30
Bus (4)	0,09	0,09	0,09	0,28
Truck (5)	0,00	0,00	0,09	0,09
Collision types (X_9)				
Rear-End (0)	1,31	0,19	6,82	8,32
Head-On (1)	2,52	0,65	4,86	8,04
Hit and Run (2)	2,34	1,78	19,16	23,27
Hit side (3)	3,93	2,06	44,02	50,00
Sideswape (4)	0,56	0,56	9,25	10,37

Table 4. The second stage analysis result of the independence test

Effect	Likelihood Ratio Tests		
	Chi-Square	df	Sig.
Intercept	0,000	0	
The victim positions (X_5)	18,664	4	,000
Accident types (X_6)	9,114	4	,361
Vehicle types (X_7)	4,333	6	,323
Collision types (X_9)	22,495	8	,000

4.2 The Estimation Result of the Severity Level Model of the Traffic Accident

The estimation result of the severity level model of the traffic accident which used multinomial logit model approach is showed in Table 5. Table 5 shows that the three types of Likelihood ratio (°^{2}) values are small. However, the hit ratio value is very good, where the hit ratio value is 84.1%. The parameters of the model have significant values which accepted.

Table 5. Parameter estimation of the model

The severity levels*	Variables	=	df	Sig.	Exp( )
Serious injury	Intercept	2.237	1	0.000	
	[X_5=0]	-1.649	1	0.003	0.192
	[X_5=1]	-1.858	1	0.000	0.156
	[X_5=2]	0^b	0		
	[X_9=0]	-3.625	1	0.002	0.026
	[X_9=1]	-2.542	1	0.000	0.079
	[X_9=2]	-1.470	1	0.014	0.230
	[X_9=3]	-1.584	1	0.005	0.205
	[X_9=4]	0^b	0		
Minor injury	Intercept	3.998	1	0.000	
	[X_5=0]	-1.157	1	0.002	0.315
	[X_5=1]	-0.870	1	0.003	0.419
	[X_5=2]	0^b	0		
	[X_9=0]	-1.664	1	0.001	0.189
	[X_9=1]	-2.717	1	0.000	0.066
	[X_9=2]	-1.106	1	0.019	0.331
	[X_9=3]	-0.912	1	0.042	0.402
	[X_9=4]	0^b	0		

a. The reference category : Death
b. Number of observation : 1,070
c. Pseudo R-Square:
- Cox and Snell : 0.070
- Nagelkerke : 0.107
- McFadden : 0.068
d. Overall Percentage (%) : 84.1
e. Pearson : 0.039

Regarding the results, the multinomial logit model for the severity levels of the traffic accident is obtained to predict the probability of each of the traffic accident. The probability prediction results of the severity levels using the model are presented in Table 6.

Table 6. The probability prediction of each severity level of the traffic accident

Collision types	The victim position	Probability of each severity level of the traffic accident (%)		
		Death	Serious injury	Minor Injury
Rear-End	Pedestrian	23.3	1.1	75.6
	Driver	18.7	0.7	80.6
	Passenger	8.7	2.1	89.2
Head-On	Pedestrian	44.0	6.2	49.8
	Driver	38.1	4.4	57.5
	Passenger	18.7	13.8	67.4
Hit and Run	Pedestrian	14.1	5.8	80.0
	Driver	11.2	3.8	85.0
	Passenger	4.7	10.2	85.1
Hit side	Pedestrian	12.1	4.5	83.4
	Driver	9.5	2.9	87.6
	Passenger	4.0	7.7	88.2
Sideswape	Pedestrian	5.0	9.0	86.0
	Driver	4.0	5.8	90.3
	Passenger	1.5	14.4	84.0

Table 6 shows that the victim position as a pedestrian has the highest probability value on the death severity level for all collision type. On the other side for all collision type, the victim position as the passenger has the highest probability value on the serious injury severity level, as well as on the minor injury level. The phenomena of the prediction results are in line with our common sense, where the pedestrian position is very sensitive to experiencing death severity level on a traffic accident event, due to the pedestrian does not has barrier directly to reduce the severity impact of the accident when a traffic accident occurred. Also, the passenger position has become more sensitive for the others severity levels of a traffic accident due to the passenger position do not chance to react in avoiding the impact of an accident event which almost occurs suddenly.

5. CONCLUSIONS

The relationship between characteristics of the traffic accidents, and their severity levels, as well as the assessment of the accident factors which significantly influenced the traffic accident severity have been explored in this study. By using the traffic accident data which recorded during 2012 - 2015 by the traffic police agency in Makassar City, Indonesia, the multinomial logistic regression model approach has been constructed and estimated in a good accuracy level.

The victim position and the collision types have played an important rule in the traffic accident severity level in the city. In further, the victim's position as a pedestrian has become the most sensitive attribute of the accident severity level. Also, the position as a passenger of the traffic accident victims has the largest sensitivity for the serious and minor injury level of the traffic accident severity levels.

Briefly, the results provide a basic view in developing a risk model for the severity level of the accident in Makassar City in further study.

ACKNOWLEDGMENT

The preliminary form of the paper has been presented at the EASTS Conference

2017. Therefore, we would like many thanks for the Conference's reviewers and participants, who were suggested some discussion to revise the essential materials of the paper.

REFERENCES

[1] Ngo, A. D., Rao, C., Hoa, N. P., Hoy, D. G., Trang, K. T., & Hill, P. S. (2012). Road traffic related mortality in Vietnam: Evidence for policy from a national sample mortality surveillance system. *BMC Public Health, 12*, 561.

[2] Sriwattanapongse, W., Khanabsakdi, S., Prasitwattanaseree, S., & Wongtra-ngan, S. (2013). Mortality Rate Model due to Transportation Accidents in Thailand. *Silpakorn U Science & Tech J, 7 (1)*, 9-18.

[3] Sutanto, S. (2009). Road Accidents In Indonesia. *IATSS Research, 33 (2)*, 122-124.

[4] Lum, H., & Reagan, J. (1995). Interactive Highway Safety Design Model: Accident Predictive Module. *Public Road, 59 (2)*.

[5] Masuri, M. G., Isa, K. A., & Tahir, M. P. (2010). Children, Youth and Road Environment: Road Traffic Accident. *Asia Pacific International Conference on Environment Behaviour Studies*, (pp. 213-218). Kuching.

[6] Lourens, P. F., Vissers, J. A., & Jessurun, M. (1999). Annual Mileage, Driving Violations, and Accident Involvement In Relation To Drivers' Sex, Age, and Level Of Education. *Accident Analysis and Prevention, 31*, 593-597.

[7] Fitroh, Lubis, H. A.-R., Frazila, R. B., & Kusumawati, A. (2015). Analysis of Motorcyclists Driving Behavior in Bandung City. *Journal of the Eastern Asia Society for Transportation Studies, 11*, 2070-2086.

[8] Adisasmita, S. A. (2010). *Road Planning and Road Network.* Makassar: Faculty Of Shipping UNHAS. (in Indonesian)

[9] Zhang, G., Yau, K. K., & Chen, G. (2013). Risk Factors Associated With Traffic Violations And Accident Severity In China. *Accident Analysis and Prevention*.

[10] Yau, K. K. (2004). Risk Factors Affecting The Severity Of Single Vehicle Traffic

[11] Haadi, A. R. (2014). Identification of Factors that Cause Severity of Road Accidents in Ghana: a Case Study of the Northern Region. *International Journal of Applied Science and Technology, 4 (3)*, 242-249.

[12] Kashani, A. T., Mohaymany, A. S., & Ranjbari, A. (2011). A Data Mining Approach to Identify Key Factors of Traffic Injury Severity. *Promet – Traffic &Transportation, 23 (1)*, 11-17.

[13] Al-Ghamdi, A. S. (2002). Using Logistic Regression To Estimate The Influence Of Accident Factors On Accident Severity. *Accident Analysis and Prevention, 34*, 729-741.

[14] Ramli, M. I., Sumi, T., & Oeda, Y. (2010). Study on Mode Choice Model of Trip for Daily Household Logistic based on Binomial Logit Model. *Proceedings of T-LOG.*

[15] Aly, S. H., Ramli, M. I., Selintung, M., & Sumi, T. (2011). Study on Emission Measurement of Vehicle on Road Based on Binomial Logit Model. *Proceedings of The 9th International Conference of Eastern Asia Society for Transportation Studies. 8.* Jeju: EASTS.

[16] Asri, A., Ramli, M. I., & Samang, L. (2011). Motorcyclist Acceptability On Road Safety Policy: Motorcycle Exclusive Lane In Makassar. *The 14th FSTPT International Symposium.* Pekanbaru: FSTPT.

[17] Kurniawan, D. (2007, Agustus 7). *Regresi Logistik*. Retrieved April 23,

2015, from https://ineddeni.wordpress.com/2007/08/07/regresi-logistik/.

[18] Wiguna, C. (2015, December 11). *Multinomial Logistic Regression Analysis*. Retrieved April 23, 2016, from http://duniaiptek.com/analisis-regresi-logistik-multinomial/. (in Indonesian)

[19] Saragih, S. R., & Yahya, K. (2014). *Severity Analysis of Traffic Accident Victims In Surabaya In 2012*, *the Statistical Analysis Log Linear and Logistic.* Surabaya: Institute of Technology Sepuluh Nopember. (in Indonesian)

[20] Afidah, L. N., & Susilaningrum, D. (2011). *Severity patterns Victims of Traffic Accidents Using Multinomial Logistic Regression (Case Study: Traffic Accidents In Surabaya).* Surabaya: Institute of Technology Sepuluh Nopember. (in Indonesian)