ANALYSIS OF FACTORS AFFECTING MOTORCYCLE-MOTOR VEHICLE CRASH CHARACTERISTICS

Thesis

Submitted to

The School of Engineering of the UNIVERSITY OF DAYTON

In Partial Fulfillment of the Requirements for

The Degree

Master of Science in Civil Engineering

By

Di Zhu

Dayton, Ohio

August, 2014
ANALYSIS OF FACTORS AFFECTING MOTORCYCLE-MOTOR VEHICLE CRASH CHARACTERISTICS

Name: Zhu, Di

APPROVED BY:

Deogratias Eustace, Ph.D., P.E., PTOE Advisory Committee Chairperson
Associate Professor, Department of Civil and Environmental Engineering and Engineering Mechanics

Peter Hovey, Ph.D.
Committee Member
Associate Professor, Department of Mathematics

Gary Shoup, P.E.
Committee Member, Senior Engineer
Montgomery County Engineering Department

John G. Weber, Ph.D.
Associate Dean
School of Engineering

Eddy M. Rojas, Ph.D., M.A., P.E.
Dean, School of Engineering
ABSTRACT

ANALYSIS OF FACTORS AFFECTING MOTORCYCLE-MOTOR VEHICLE CRASH CHARACTERISTICS

Name: Zhu, Di
University of Dayton

Advisor: Dr. Deogratias Eustace

As everybody knows, there are many traffic crashes happening every day. Traffic crashes may result in injury, death, and property damage. A number of factors contribute to the risk of a crash, including vehicle design, speed of operation, road design, road environment, driver skill and/or impairment, and driver behavior. Worldwide, motor vehicle crashes lead to death and disability as well as financial costs to both society and the individuals involved. The objective of this study was to analyze crash data of motorcycle-motor vehicle collisions to identify possibly influential factors that cause these crashes and to study the magnitude of influence of each factor to these crashes. This study tested appropriate regression models to accurately model the factors that significantly influence motorcycle-motor vehicle crashes.

A nominal multinomial logistic regression model was built. From stepwise selection procedure, the influential factors included age, time of crash, number of units, vehicle in error, road contour, collision type, alcohol used, posted speed, and helmet used.
Number of units involved in a crash impacts the crash severity level, such as two units mostly result in injury and three or more units mostly result in fatal. If the driver of the motor vehicle causes the crash it will more likely result into injury than if the driver of the motorcycle causes the crash. Driver of motorcycle or vehicle that uses alcohol will certainly increase the chance of a fatality or injury. Crashes that occur on highways or freeways with higher speed limits are more likely to result in injuries and fatalities. The occupants of motorcycle use helmet will significantly be protected in the crash. These factors can be applied to reduce the severity of motorcycle-motor vehicle crashes.
ACKNOWLEDGEMENTS

My special thanks are in order to Dr. Deogratias Eustace, my advisor, for providing the time and advise necessary for the work contained herein, and for directing this thesis and bringing it to its conclusion with patience and expertise. I also deeply appreciate the help given by Dr. Peter Hovey and Mr. Gary Shoup, who helped me with the methodology in multinomial logistic regression model and reviewed the analysis results. For my friends and classmates, thank you for the help and contribution to my success in this program of study. Finally, my thanks go to my parents, for their continued contributions and bearing with me throughout this program of study.
# TABLE OF CONTENTS

ABSTRACT .......................................................................................................................... iii

ACKNOWLEDGEMENTS ................................................................................................. v

TABLE OF CONTENTS ..................................................................................................... vi

LIST OF FIGURES ........................................................................................................... ix

LIST OF TABLES ................................................................................................................ x

CHAPTER ONE INTRODUCTION ...................................................................................... 1

1.1 Introduction ............................................................................................................... 1

1.2 Problem Statement ................................................................................................. 2

1.3 Study Questions to be addressed ........................................................................... 2

CHAPTER TWO LITERATURE REVIEW ........................................................................... 3

2.1 Introduction ............................................................................................................... 3

2.2 Influential Factors for Traffic Crashes .................................................................... 3

2.3 Motorcycle-Motor Vehicle Crashes ....................................................................... 5

2.4 Motorcycle Crashes ............................................................................................... 6

2.5 Traffic Safety ......................................................................................................... 7

2.6 Summary ................................................................................................................ 7

CHAPTER THREE METHODOLOGY AND DATA COLLECTION .................................. 9

3.1 Introduction ............................................................................................................. 9
3.2 Methodology .............................................................................................................................9
  3.2.1 Logistic Regression Introduction ..........................................................................................10
  3.2.2 Logistic Regression Equation ..............................................................................................10
  3.2.3 Logistic Regression Model Fit ..............................................................................................11
    3.2.3.1 Deviance ........................................................................................................................11
    3.2.3.2 Maximum Likelihood ......................................................................................................12
    3.2.3.3 The Likelihood Ratio Tests ............................................................................................12
  3.3 Data Collection ..........................................................................................................................13
    3.3.1 Crash Data ..........................................................................................................................13
    3.3.2 Merging Files ......................................................................................................................14
    3.3.3 Creating of Motorcycle-Motor Vehicle Crash Database ..................................................15
  3.4 Description of Selected Variables for Study ............................................................................15
    3.4.1 Dependent Variable ..........................................................................................................16
    3.4.2 Explanatory Variables ........................................................................................................17
      3.4.2.1 Driver Behavioral Factors ............................................................................................17
      3.4.2.2 Geometric and Environmental Factors .........................................................................17
      3.4.2.3 Traffic Factors ............................................................................................................18
  3.5 Analysis of Data .......................................................................................................................18
  CHAPTER FOUR RESULTS ............................................................................................................19
  4.1 Introduction ..............................................................................................................................19
  4.2 Descriptive Results of Motorcycle-Motor Vehicle Crashes ..................................................20
    4.2.1 Road Condition Groups Frequency ....................................................................................20
4.2.2 Day of Week Groups Frequency ........................................................... 21
4.2.3 Weather Condition Groups Frequency .............................................. 22
4.2.4 Time of Crash Groups Frequency .................................................... 23
4.2.5 Light Condition Groups Frequency .................................................. 24
4.2.6 Age Groups Frequency ..................................................................... 25
4.3 Results of Logistic Regression Modeling for Predictor Variables .......... 25
4.3.1 General .............................................................................................. 26
4.3.2 Dependent Variable Crash Severity Checking in SPSS ..................... 26
  4.3.2.1 Test of Parallel Lines Procedure ............................................... 26
  4.3.2.2 Crash Severity Frequency .......................................................... 28
4.3.3 Nominal Logistic Regression Model Building in SAS ...................... 28
  4.3.3.1 Stepwise Selection Procedure .................................................. 28
  4.3.3.2 Maximum Likelihood Estimates ............................................... 30
  4.3.3.3 Odds Ratio Estimates ................................................................. 34
4.4 Discussion of Results .......................................................................... 36
  4.4.1 Influential Factors Impact Crash Severity Levels ............................. 36
  4.4.2 Influential Factors Comparison Analysis ........................................ 38
  4.4.3 Summary .......................................................................................... 42
CHAPTER FIVE CONCLUSIONS AND RECOMMENDATIONS ................... 43
REFERENCES ............................................................................................... 47
LIST OF FIGURES

Figure 3.1: Dependent Variable Crash Severity Frequency ........................................... 16

Figure 4.1: Explanatory Variable Road Condition Frequency ........................................ 20

Figure 4.2: Explanatory Variable Day of Week Groups Frequency ................................. 21

Figure 4.3: Explanatory Variable Weather Condition Groups

Frequency ....................................................................................................................... 22

Figure 4.4: Explanatory Variable Time of Crash Groups Frequency ............................... 23

Figure 4.5: Explanatory Variable Light Condition Groups

Frequency ....................................................................................................................... 24

Figure 4.6: Explanatory Variable Age Groups Frequency .............................................. 25

Figure 4.7: Ordinal Regression Dialog in SPSS .............................................................. 27

Figure 4.8: SAS Code of Stepwise Selection ................................................................. 28

Figure 4.9: Alcohol Non-used Crashes Distribution ...................................................... 38

Figure 4.10: Alcohol Used Crashes Distribution .......................................................... 39

Figure 4.11: Helmet Non-used Crashes Distribution .................................................... 40

Figure 4.12: Helmet Used Crashes Distribution ............................................................ 41
LIST OF TABLES

Table 4.1: Test of Parallel Lines in SPSS Output.......................................................... 27
Table 4.2: Crash Severity Frequency in SPSS Output................................................... 28
Table 4.3: Crash Severity Profile in SAS Output............................................................ 29
Table 4.4: Summary of Stepwise Selection in SAS Output............................................ 30
Table 4.5: Maximum Likelihood Estimates in SAS Output............................................ 32
Table 4.6: Odds Ratio Estimates in SAS Output............................................................ 35
CHAPTER ONE

INTRODUCTION

1.1 Introduction

As everybody knows, there are many traffic crashes happening every day. Traffic crashes may result in injury, death, and property damage. A number of factors contribute to the risk of crash, including vehicle design, speed of operation, road design, road environment, driver skill and/or impairment, and driver behavior. Worldwide, motor vehicle crashes lead to death and disability as well as financial costs to both society and the individuals involved. The objective of this study was to analyze crash data of motorcycle-motor vehicle collisions to identify possible factors that cause these crashes and to study the magnitude of influence of each factor to these crashes. This study tested appropriate regression models to accurately model the factors that significantly influence motorcycle-motor vehicle crashes.
1.2 Problem Statement

Motorcycle-motor vehicle crashes may cause by many factors such as traffic controls, speed and route type, road characteristic, weather impacts, road classification, and human behavior. The proposed research will develop an appropriate regression model that will help find the most influential factors to the crashes.

1.3 Study Questions to be addressed

1. Which statistical regression model will best present the results?
2. What are influential factors affecting to the motorcycle-motor vehicle crashes?
3. What possible solutions can be applied to minimize motorcycle-motor vehicle crashes?
CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction

A detailed literature review on related studies was preformed to collect information and build background knowledge for the thesis. This literature review researched for available published research notes and journal papers on factors affecting motorcycle-motor vehicle crashes. During the literature review, also a task of finding available methodologies for analyzing influential factors for traffic crashes was undertaken. In research studies conducted on this topic, so far logistic regression and logit regression models are the most common models used in analyzing influential factors for traffic crashes.

2.2 Influential Factors for Traffic Crashes

A study by Peng and Boyle [1] sought to gain insights on the impact of commercial driver factors on crash severity with respect to single-vehicle, run-of-road (ROR) crashes. This study mentioned seat belt use significantly reduced the likelihood of injury and fatal
ROR crashes. Driver distraction and inattention increased the likelihood of an ROR crash. Drowsiness, fatigue and speeding significantly increased the likelihood for injury and fatal ROR crashes. Commercial motor vehicle (CMV) drivers who drove a non-defective truck were associated with a lower likelihood of injury and fatal ROR crashes. An ROR crash was about 3.8 times more likely to be injurious and fatal if it happened on rural roads or dry roads. No other explanatory variables were observed as significant. The results of this study suggest that several driver factors: drowsiness and fatigue, speeding, distraction, inattention, and seat belt use did influence the likelihood of an ROR crash being injurious or fatal. Thus, the study of Peng and Boyle [1] provides insights on the magnitude of effect of these driver factors on ROR crashes that involve large trucks. The results have implications for behavioral safety countermeasures that can help mitigate the impacts of driver distractions and speeding [1].

Transient factors related to the failure to detect motorcycles might include alcohol, fatigue/lack of sleep, inattention, and information overload, whereas factors that are more permanent might include “cognitive” conspicuity and field dependence [2]. The model with the best fitting and highest predictive capability was used to identify the influence of roadway, environmental, vehicle, and driver related factors on severity. Travel speed, restraint device usage, point of impact, use of alcohol and drugs, personal condition, gender, whether the driver is at fault, urban/rural nature and grade/curve existence at the crash location were identified as the important factors for making an injury severity difference to older drivers involved in single vehicle crashes [3]. Logistic regression was applied to accident-related data collected from traffic police records in
In order to examine the contribution of several variables to accident severity \[^4\]. The ordered probit model was used to evaluate the effect of roadway and area type features on injury severity of pedestrian crashes in rural Connecticut \[^5\]. Failure to wear seat belts did not predict accidents but did significantly influence the severity of accidents that did occur; that is, those who had earlier reported using seat belts ‘always’ were less likely than others to be injured when accidents did occur. Financial stress increased the likelihood of involvement in more serious accidents \[^6\]. These results will influence the urban traffic police enforcement measures, which will change inappropriate behavior of drivers and protect the least experienced road users \[^7\].

### 2.3 Motorcycle-Motor Vehicle Crashes

Alcohol and excessive speed were common factors associated with motorcyclist crash involvement. Left turns and failure to yield were common factors associated with the involvement of other motorists. Suggested countermeasures include helmet use and enforcement of speed and impaired driving laws \[^8\]. Past motorcycle crash history, number of riding days, average riding distance, risk-taking level, alcohol consumption, and traffic violations were all significantly associated with an increased risk of being involved in a crash. Conversely, increasing age, riding experience, and automobile licensure were related to a decreased risk of crashing. Furthermore, helmet use was not independently related to the risk of crashing. In conclusion, a high-risk group predisposed to involvement in a motorcycle crash, including both non-injury and injury-related crashes, can be identified using selected risk factors for crash prevention among young riders \[^9\]. Nested logit (estimated with full information maximum
likelihood) and standard multinomial logit model results show a wide-range of factors significantly influence injury-severity probabilities. Key findings show that increasing motorcyclist age is associated with more severe injuries and that collision type, roadway characteristics, alcohol consumption, helmet use, unsafe speed and other variables play significant roles in crash-injury outcomes \[^{10}\]. Helmets and helmet use laws have been shown to be effective in reducing head injuries and deaths from motorcycle crashes. Alcohol is the major contributing factor to fatal crashes \[^{11}\].

### 2.4 Motorcycle Crashes

Motorcycles are overrepresented in road traffic crashes and particularly vulnerable at signalized intersections. The objective of this study is to identify causal factors affecting the motorcycle crashes at both four-legged and T signalized intersections \[^{12}\]. The fatalities usually involved a single vehicle crash and young men. The roadside barriers predominantly involved were steel W-beams, typically on a bend in the horizontal alignment of the road. A majority of fatalities occurred on a weekend, during daylight hours, on clear days with dry road surface conditions indicating predominantly recreational riding. Speeding and driving with a blood alcohol level higher than the legal limit contributed to a significant number of these fatalities \[^{13}\]. Strong effects of motorcycle type were observed on driver death rates and on the likelihood of risky driving behaviors such as speeding and alcohol impairment. Although the current study could not completely disentangle the effects of motorcycle type and rider characteristics such as age on driver death rates, the effects of both motorcycle type and rider age on the likelihood of risky driving behaviors were observed among fatally injured motorcycle
drivers [14].

2.5 Traffic Safety

A model is presented for assessing the effects of traffic safety measures, based on a breakdown of the process in underlying components of traffic safety (risk and consequence), and five (speed and conflict related) variables that influence these components, and are influenced by traffic safety measures [15]. A traffic encounter between individual road users is a process of continuous interplay over time and space and may be seen as an elementary event with the potential to develop into an accident [16]. It is concluded that only rules that are possible to enforce should be implemented and that police surveillance should be visible to the drivers. Also, the traffic system should be seen as a social system where drivers are interacting with other drivers and road users. Rules and regulations are important to help the actors of the system to function in a safe and effective way [17].

2.6 Summary

After doing the literature review by researching for available published research notes and journal papers on factors affecting motorcycle-motor vehicle crashes, these factors included drowsiness and fatigue, speeding, distraction, inattention, seat belt use, dry roadway surfaces, rural roads, dark and unlit road condition, and traveling between midnight and 6 a.m. were considered as influential factors for the crashes. The methodology for analyzing influential factors for motorcycle-motor vehicle crashes was the logistic regression model or logit regression model, it is the most common model
used in analyzing influential factors for traffic crashes.
CHAPTER THREE

METHODOLOGY AND DATA COLLECTION

3.1 Introduction

The current study analyzed motorcycle-motor vehicle crash data or motorcycle crash data from the Ohio Department of Public Safety (ODPS) database. The data covered five years through 2008 to 2012 traffic crashes in the State of Ohio. In the model building part of the methodology, factors such as driver behavioral factors, geometric and environmental factors, and traffic factors were used in both SPSS and SAS to come up with the factors that influence most to the crashes. After model building, the model was tested for regression analysis to evaluate the results. A multinomial logistic regression model was built by using both SPSS and SAS in this study.

3.2 Methodology

In this section, the logistic regression introduction, the logistic regression equation, and the logistic regression model fit are discussed.
3.2.1 Logistic Regression Introduction

In statistics, logistic regression or logit regression is a type of probabilistic statistical classification model. It is also used to predict a binary response from a binary predictor, used for predicting the outcome of a categorical dependent variable based on one or more predictor variables. That is, it is used in estimating empirical values of the parameters in a qualitative response model. The probabilities describing the possible outcomes of a single trial are modeled, as a function of the predictor variables, using a logistic function. Frequently logistic regression is used to refer specifically to the problem in which the dependent variable is binary—that is, the number of available categories is two—and problems with more than two categories are referred to as multinomial logistic regression or, if the multiple categories are ordered, as ordered logistic regression.

Logistic regression measures the relationship between a categorical dependent variable and one or more independent variables, which are usually continuous, by using probability scores as the predicted values of the dependent variable. A chi-square test is used to indicate how well the logistic regression model fits the data.

3.2.2 Logistic Regression Equation

The logistic formulas are stated in terms of the probability that Y=1, which is referred to as P. The probability that Y is 0 is 1-P.

\[
\ln\left(\frac{P}{1-P}\right) = \beta_0 + \beta_1 x
\]
The \( \ln \) symbol refers to a natural logarithm and \( \beta_0 + \beta_1 X \) is our familiar equation for the regression line.

\( P \) can be computed from the regression equation also. Therefore, if we know the regression equation, we could, theoretically, calculate the expected probability that \( Y = 1 \) for a given value of \( X \).

\[
P = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}
\]

\( \exp \) is the exponent function, sometimes written as \( e \). Therefore, the equation on the right is just the same thing but replacing \( \exp \) with \( e \).

### 3.2.3 Logistic Regression Model Fit

This sub-section discusses the deviance, the maximum likelihood, and the likelihood ratio tests.

#### 3.2.3.1 Deviance

With logistic regression, instead of \( R^2 \) as the statistic for overall fit of the model, the deviance is used instead. When the chi-square analyses are used, chi-square is said to be a measure of “goodness of fit” of the observed and the expected values. The chi-square is used as a measure of model fit here in a similar way. It is the fit of the observed values \( (Y) \) to the expected values \( (Y) \). The bigger the difference (or “deviance”) of the observed values from the expected values, the poorer the fit of the model. Therefore, a smaller deviance if possible is expected. As more variables are added to the equation/model the deviance should get smaller, indicating an improvement in fit.
3.2.3.2 Maximum Likelihood

Instead of finding the best fitting line by minimizing the squared residuals, as the ordinary least squares (OLS) regression is involved, a different approach with logistic—Maximum Likelihood (ML) is used. ML is a way of finding the smallest possible deviance between the observed and predicted values (kind of like finding the best fitting line) using calculus (derivatives specifically). With ML, the computer uses different "iterations" in which it tries different solutions until it gets the smallest possible deviance or best fit. Once it has found the best solution, it provides a final value for the deviance, which is usually referred to as "negative two log likelihood" (shown as "-2 Log Likelihood" in SPSS). The deviance statistic is called –2LL by Cohen et al., 2003 [18] and Pedazur and D by Hosmer and Lemeshow, 1989 [19], and it can be thought of as a chi-square value.

3.2.3.3 The Likelihood Ratio Tests

Instead of using the deviance (–2LL) to judge the overall fit of a model, however, another statistic is usually used that compares the fit of the model with and without the predictor(s). This is similar to the change in $R^2$ when another variable has been added to the equation. But here, the deviance is expected to decrease, because the degree of error in prediction decreases as another variable is added. To do this, the comparison of the deviance with just the intercept (–2LLnull referring to –2LL of the constant-only model) to the deviance is made when the new predictor or predictors have been added (–2LLk referring to –2LL of the model that has k number of predictors). The difference between these two deviance values is often referred to as G for goodness of fit.
(important note: G is referred to as “chi-square” in SPSS printouts).

\[ G = \chi^2 = D(\text{for the model without the variable}) - D(\text{for the model with the variable}) \]

or, using the Cohen et al., 2003 \cite{18} notation,

\[ G = \chi^2 = D_{null} - D_k = (-2LL_{null}) - (-2LL_k) \]

where \( D_{null} \) is the deviance for the constant only model and \( D_k \) is the deviance for the model containing \( k \) number of predictors. An equivalent formula sometimes presented in textbooks is:

\[ G = \chi^2 = -2 \ln \left( \frac{L_{null}}{L_k} \right) \]

where the ratio of the ML values is taken before taking the log and multiplying by -2. This gives rise to the term “likelihood ratio test” to describe \( G \).

3.3 Data Collection

This section discusses crash data, merging files, and the creating of motorcycle-motor vehicle crash database.

3.3.1 Crash Data

The aim of this thesis study was to analyze motorcycle-motor vehicle crash data or motorcycle crash data from the Ohio Department of Public Safety (ODPS) database. The data covered five years through 2008 to 2012 traffic crashes in the State of Ohio.
data contained comprehensive information on each crash, including driver demographics; driver behaviors; vehicle characteristics; crash types; crash severity; geometric characteristics, and environmental conditions at the time of crash.

### 3.3.2 Merging Files

The three related files in ODPS traffic crash database are crash records, unit records, and person records.

The “crash records” file contains information specific to each crash that occurred such as crash severity, vehicle in error, date of crash, time of crash, name of city, village or township where the crash occurred, FIPS place code, crash location, type of road, if alcohol or drug was used, if speeding was involved, etc.

The “unit records” file contains information on each unit/vehicle that was involved in a particular crash incident. Information recorded includes unit type (e.g., passenger vehicle, motorcycle, large vehicle, etc.), point of impact, and number of occupants in the unit, etc.

The “person records” file contains information on each person involved in each crash with the exception of hit-and-run cases where information is always not available. Information recorded in this file includes person type (e.g., driver, occupant, and pedestrian), age, gender, severity of injury sustained by an individual, safety equipment used, etc.

A single-to-many merging technique in SPSS (Version 20.0) software was used to merge the three files mentioned above. Corresponding records in the “crash” and “unit” files were joined by using a common variable, DOCNO (document number). The records
in the joint “crash-unit” file were then joined together with their corresponding records in the “person” file by using the two common variables DOCNO and UNITNO (unit number) to obtain a joint “crash-unit-person” file. The five created joint “crash-unit-person” files for each calendar year through 2008 to 2012 were again joined together to obtain a “final merge” file that contained a five-years traffic crash data. Records in the final file were systematically checked for consistency and to make sure that all records were correctly joined.

3.3.3 Creating of Motorcycle-Motor Vehicle Crash Database

By using the final joined file, a motorcycle-motor vehicle crash database was created by using several selected conditions. The first is “motorcycle related”, all the motorcycle related crash data were selected. The second is “collision type”, 1 and 9 were removed from the database. The third is “person type”, occupant and pedestrian were removed from the database. The fourth is “number of units”, 1 was removed from the database. The fifth is “unit type”, 24, 25 and 99 were removed from the database. Thus a file containing a total of 10418 records of motorcycle-motor vehicle crash data with complete crash-related information from the database of traffic crashes that occurred on Ohio’s public freeways and highways through five years period 2008 to 2012 was created.

3.4 Description of Selected Variables for Study

In this section, the dependent variable crash severity and the explanatory variables are discussed.
3.4.1 Dependent Variable

The dependent variable was crash severity, which was coded into three categories within the Ohio Department of Public Safety (ODPS) crash database. They are fatal, injury, and property-damage-only (PDO). This variable means number for the most severe injury, or in the absence of injury, property damage involved in the crash. The percentage of fatal crash is 3.95%, the percentage of injury crash is 70.45%, and the percentage of property-damage-only (PDO) crash is 25.61%. A multinomial variable was created for crash severity and used as the dependent variable.

![Crash Severity](image)

**Figure 3.1: Dependent Variable Crash Severity Frequency**

From the crash severity frequency shown above in Figure 3.1, the injury had the highest proportion of all the three crash severity categories.
3.4.2 Explanatory Variables

Several geometric, environmental, traffic, and driver behavioral factors, which are considered to have an effect on the occurrence of motorcycle-motor vehicle related traffic crashes were analyzed by using some powerful statistical modeling techniques in order to determine the most significant ones. A total of eighteen variables (include six grouped variables) were selected for exploratory analysis to investigate characteristics of predictor variables and screen out the most influential ones. These variables were considered as potentially influential factors on crash severity and used as the explanatory variables.

3.4.2.1 Driver Behavioral Factors

Driver behavioral factors considered in the model included the driver’s gender and age groups; drug or alcohol use; use of safety equipment; and vehicle in error. Driver’s gender and age means the gender and age of driver. Drug or alcohol use means the driver apparent condition at time of crash. Use of safety equipment means safety restraint equipment in use by the occupant at the time of crash. Vehicle in error means the driver which had the most causative bearing on the crash.

3.4.2.2 Geometric and Environmental Factors

Geometric and environmental factors considered in the model included road surface (type of road surface where crash occurred); light condition groups (lighting conditions at the time of crash); posted speed (posted speed limit of road); road contour (contour of road at crash scene); road condition groups (road conditions at crash scene);
3.4.2.3 Traffic Factors

Traffic factors considered in the model included vehicle groups, number of units, and collision type.

3.5 Analysis of Data

A multinomial logistic regression model was developed by using both SPSS and SAS. The test of parallel lines was made in SPSS to check which multinomial logistic regression model we should use, ordinal or nominal. The stepwise selection was made in SAS to select the best explanatory variables to build the model, these variables were considered as the influential factors. The significant level and chi-square were presented for each explanatory variable in the likelihood ratio tests and maximum likelihood estimates. They are important criterions for checking whether the explanatory variable was an influential factor on dependent variable crash severity or not.
CHAPTER FOUR

RESULTS

4.1 Introduction

Several geometric, environmental, traffic, and driver behavioral factors, which are considered to have an effect on the occurrence of motorcycle-motor vehicle related traffic crashes were analyzed by using logistic regression in order to determine the most significant ones. A total of eighteen variables (include six grouped variables) were selected for exploratory analysis to investigate characteristics of predictor variables and screen out the most influential ones. The multinomial logistic regression model which uses maximum likelihood estimation method was applied to estimate statistically the effects of these variables in contributing to the occurrence of motorcycle-motor vehicle crash severity levels. Predictor variables were tested at a 95% significance level. The first part of this chapter presents the descriptive results of motorcycle-motor vehicle crashes based on the data analyzed in this study; the second part discusses the results of the multinomial logistic regression modeling by using both SPSS and SAS; and lastly, in the third part, how the influential factors impact the results is discussed.
4.2 Descriptive Results of Motorcycle-Motor Vehicle Crashes

Motorcycle-motor vehicle crashes are usually complicated events normally contributed by a number of various types of interactions. Examining the characteristics of crash data provides a fairly decent inference at some underlying aspects of the motorcycle-motor vehicle crashes and thus can assist the analysts in devising possible safety countermeasures. For the objective of this study, a total of 10418 observations of complete data points were used whereby, 3.95% involved in fatal injury, 70.45% involved in injury, and 25.61% involved in property-damage-only (PDO). The main objective of the descriptive results is to provide a better view of characteristics of motorcycle-motor vehicle crashes. In particular, this section pays more attention to the injury crashes of fatal injury, injury and PDO.

4.2.1 Road Condition Groups Frequency

![Road Condition Groups Frequency](image)

Figure 4.1: Explanatory Variable Road Condition Groups Frequency
Figure 4.1 shows the road condition frequency distribution of the motorcycle-motor vehicle crashes. For the two categories of road condition groups, the percentage of dry is 95.87%, and the percentage of other is 4.13%. From Figure 4.1, the percentage of fatal in dry is higher than in other and the percentage of PDO in dry is lower than in other.

4.2.2 Day of Week Groups Frequency

![Day of Week Groups](image)

**Figure 4.2: Explanatory Variable Day of Week Groups Frequency**

The distribution of day of week frequency of motorcycle-motor vehicle crashes is displayed in Figure 4.2. For the two categories of day of week groups, the percentage of weekends is 32.22%, and the percentage of weekdays is 67.78%. From Figure 4.2, the percentage of fatal in weekends is higher than in weekdays and the percentage of PDO in weekends is lower than in weekdays.
4.2.3 Weather Condition Groups Frequency

The distribution of weather condition frequency of motorcycle-motor vehicle crashes is displayed in Figure 4.3. For the two categories of weather condition groups, the percentage of clear is 81.82%, and the percentage of other is 18.18%. From Figure 4.3, the percentage of fatal in clear is lower than in other and the percentage of PDO in clear is higher than in other.
4.2.4 Time of Crash Groups Frequency

![Time of Day Groups](image)

Figure 4.4: Explanatory Variable Time of Crash Groups Frequency

The distribution of time of crash frequency of motorcycle-motor vehicle crashes is displayed in Figure 4.4. For the five categories of time of crash groups, the percentage of morning is 7.42%, the percentage of midday is 38.29%, the percentage of evening is 30.37%, the percentage of early night is 18.44%, and the percentage of late night is 5.49%. From Figure 4.4, the percentage of fatal in late night is the highest in all categories and the percentage of PDO in evening is the highest in all categories.
4.2.5 Light Condition Groups Frequency

Figure 4.5 shows the light condition frequency distribution of the motorcycle-motor vehicle crashes. For the two categories of light condition groups, the percentage of daylight is 80.19%, and the percentage of other is 19.81%. From Figure 4.5, the percentage of fatal in daylight is lower than in other and the percentage of PDO in daylight is higher than in other.
4.2.6 Age Groups Frequency

Figure 4.6 shows the age frequency distribution of the motorcycle-motor vehicle crashes. For the five categories of age groups, the percentage of younger than 20 is 4.39%, the percentage of 20 to 25 is 12.22%, the percentage of 26 to 30 is 8.52%, the percentage of 31 to 64 is 69.82%, and the percentage of older than 65 is 5.04%. From Figure 4.6, the percentage of fatal in 65+ is the highest in all categories and the percentage of PDO in 20-25 is the highest in all categories.

4.3 Results of Logistic Regression Modeling for Predictor Variables

In this section, the dependent variable crash severity checking in SPSS and the nominal logistic regression model building in SAS are discussed.
4.3.1 General

The response variable in the dataset was the crash severity, which consisted of three levels of crash and it was modeled as a nominal variable. Eighteen predictor variables where selected from the motorcycle-motor vehicle crash dataset and six of these were grouped variables. The main objective of this task was to investigate the complex relationships between the crash severity and the nineteen selected predictor variables by using the logistic regression modeling. The final product from this logistic regression analysis was to identify the influential factors, which should be used in the regression modeling of predicting the crash severity of motorcycle-motor vehicle crashes.

The first part of this analysis was checking ordinal or nominal of dependent variable crash severity by using test of parallel lines in SPSS. In the second part of this analysis, the logistic regression model was built by using stepwise selection in SAS. The final results contained the variables that were selected from the stepwise selection procedure.

4.3.2 Dependent Variable Crash Severity Checking in SPSS

In this sub-section, the test of parallel lines and the crash severity frequency are discussed.

4.3.2.1 Test of Parallel Lines Procedure

Figure 4.7 shows a SPSS screen shot of ordinal regression output request menu used to test the parallel lines, which determines if the data can be treated as ordinal.
From the test of parallel lines results shown in Table 4.1 above, the significant level is less than 0.0001, so it is extremely significant. We can conclude that the dependent variable crash severity is nominal, so the nominal multinomial logistic regression model was built.
4.3.2.2 Crash Severity Frequency

<table>
<thead>
<tr>
<th>Crash Severity</th>
<th>Frequency</th>
<th>Percent</th>
<th>Valid Percent</th>
<th>Cumulative Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Valid</td>
<td>411</td>
<td>3.9</td>
<td>3.9</td>
<td>3.9</td>
</tr>
<tr>
<td>Fatal Injury</td>
<td>7339</td>
<td>70.4</td>
<td>70.4</td>
<td>74.4</td>
</tr>
<tr>
<td>Injury</td>
<td>2068</td>
<td>25.6</td>
<td>25.6</td>
<td>100.0</td>
</tr>
<tr>
<td>PDO</td>
<td>10418</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Crash Severity Frequency in SPSS Output

From the dependent variable crash severity frequency table shown above, a total of 10418 observations of the motorcycle-motor vehicle crash data points were used whereby, 3.95% observations involved in fatal injury, 70.45% observations involved in injury, and 25.61% observations involved in property-damage-only (PDO).

4.3.3 Nominal Logistic Regression Model Building in SAS

In this sub-section, the stepwise selection procedure, the maximum likelihood estimates, and the odds ratio estimates are discussed.

4.3.3.1 Stepwise Selection Procedure

```sas
proc logistic data=mth543.motorcycleonly;
   missing ’99;
   class age_grp weather_cond light_cond road_cond time_grp day_grp roadcountourid collisiontypepid isworkzone related roadsurface isalschoolrelated isdrugrelated dividedundividedroadway gender helmet;
   model crashseverityflag = age_grp weather_cond light_cond road_cond time_grp day_grp numbberofunits vehicleerror roadcountourid collisiontypepid isworkzone related roadsurface isalschoolrelated isdrugrelated dividedundividedroadway postedspeed gender helmet / selection=stepwise link= glogit;
   run;
```

Figure 4.8: SAS Code of Stepwise Selection
A nominal multinomial logistic regression model was built in SAS to predict the likelihood of a motorcycle-motor vehicle crash being injurious or fatal. The generalized logit model and Newton-Raphson optimization technique were used to develop the multinomial logistic regression model. From the response profile shown in Table 4.3, the response variable crash severity has three levels which are “1” refers to fatal injury, “2” refers to injury, and “3” refers to property-damage-only (PDO).

Table 4.3: Crash Severity Profile in SAS Output
From the summary of stepwise selection shown in Table 4.4, the stepwise selection procedure selected nine explanatory variables to build the multinomial logistic regression model and model building terminates because the last effect entered is removed by the Wald statistic criterion. Therefore the nine selected variables are significant to the model and have influential relationship with the dependent variable crash severity.

### 4.3.3.2 Maximum Likelihood Estimates

The nominal multinomial logistic regression model was developed to predict the likelihood of a motorcycle-motor vehicle crash being injurious or fatal. The chi-square is the important criterion in maximum likelihood estimates for identifying the influential factors for motorcycle-motor vehicle crashes.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>CRASHSEVERITYFLAG</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald Chl-Square</th>
<th>Pr &gt; Chisq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1</td>
<td>-12.6029</td>
<td>1.2936</td>
<td>0.0097</td>
<td>0.9214</td>
</tr>
<tr>
<td>Intercept</td>
<td>2</td>
<td>1</td>
<td>-1.5409</td>
<td>0.2834</td>
<td>29.5708</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>1</td>
<td>1</td>
<td>-0.3574</td>
<td>0.2845</td>
<td>1.5780</td>
<td>0.2091</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>1</td>
<td>2</td>
<td>-0.0465</td>
<td>0.1036</td>
<td>0.1827</td>
<td>0.6607</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>2</td>
<td>1</td>
<td>0.1244</td>
<td>0.1652</td>
<td>0.5669</td>
<td>0.4515</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>2</td>
<td>2</td>
<td>0.00619</td>
<td>0.0720</td>
<td>0.0074</td>
<td>0.9315</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>3</td>
<td>1</td>
<td>-0.2985</td>
<td>0.2005</td>
<td>2.2169</td>
<td>0.1366</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>3</td>
<td>2</td>
<td>0.0224</td>
<td>0.0826</td>
<td>0.0735</td>
<td>0.7663</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>4</td>
<td>1</td>
<td>-0.1475</td>
<td>0.1181</td>
<td>1.5996</td>
<td>0.2117</td>
</tr>
<tr>
<td>AGE_GRP</td>
<td>4</td>
<td>2</td>
<td>-0.0648</td>
<td>0.0497</td>
<td>1.5968</td>
<td>0.1927</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>1</td>
<td>1</td>
<td>-0.0299</td>
<td>0.2010</td>
<td>0.0221</td>
<td>0.8817</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>1</td>
<td>2</td>
<td>-0.0669</td>
<td>0.0840</td>
<td>0.4357</td>
<td>0.4253</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>2</td>
<td>1</td>
<td>0.0500</td>
<td>0.1132</td>
<td>0.1947</td>
<td>0.6591</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>2</td>
<td>2</td>
<td>0.0129</td>
<td>0.0494</td>
<td>0.0668</td>
<td>0.7960</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>3</td>
<td>1</td>
<td>-0.3412</td>
<td>0.1296</td>
<td>8.0043</td>
<td>0.0047</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>3</td>
<td>2</td>
<td>-0.1266</td>
<td>0.0610</td>
<td>6.0725</td>
<td>0.0137</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>4</td>
<td>1</td>
<td>0.1952</td>
<td>0.1258</td>
<td>2.3692</td>
<td>0.1237</td>
</tr>
<tr>
<td>TIME_GRP</td>
<td>4</td>
<td>2</td>
<td>0.1186</td>
<td>0.0627</td>
<td>3.8189</td>
<td>0.0584</td>
</tr>
<tr>
<td>NUMBEROFUNITS</td>
<td>1</td>
<td>1</td>
<td>1.1366</td>
<td>0.1514</td>
<td>56.3269</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>NUMBEROFUNITS</td>
<td>2</td>
<td>1</td>
<td>0.7358</td>
<td>0.1072</td>
<td>47.1394</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>VEHICLEERROR</td>
<td>1</td>
<td>1</td>
<td>0.00162</td>
<td>0.00286</td>
<td>0.3228</td>
<td>0.5693</td>
</tr>
<tr>
<td>VEHICLEERROR</td>
<td>2</td>
<td>1</td>
<td>-0.00292</td>
<td>0.00139</td>
<td>4.4304</td>
<td>0.0363</td>
</tr>
</tbody>
</table>
If the chi-square is less than 0.05, it means this factor is significant to the dependent variable crash severity. If the chi-square is less than 0.0001, it means this factor is extremely significant to the dependent variable crash severity.

The results of maximum likelihood estimates are shown in the Table 4.5. The
influential factors include age, time of crash, number of units, vehicle in error, road contour, collision type, alcohol used, posted speed, and helmet used. Helmet used significantly reduced the likelihood of an injurious and fatal motorcycle-motor vehicle crash. However, the likelihood of a motorcycle-motor vehicle crash being injurious and fatal was increased when motor vehicle caused the crash. Road contour significantly influences the likelihood of an injurious and fatal motorcycle-motor vehicle crash. Posted speed limit, collision type, and number of units also influence the likelihood of a motorcycle-motor vehicle crash being injurious and fatal. Alcohol used significantly increased the likelihood of an injurious and fatal motorcycle-motor vehicle crash. Also time of crash and age influence the likelihood of a motorcycle-motor vehicle crash being injurious and fatal. These factors contribute significantly to the model and they are considered as influential factors for the likelihood of an injurious and fatal motorcycle-motor vehicle crash. No other factors were considered as significant for the likelihood of a motorcycle-motor vehicle crash being injurious and fatal.

From the motorcycle-motor vehicle crashes or traffic crashes in this study, there are many influential factors that cause the crashes to take place and make the crash severity level to be fatal or injury. The influential factors from this study are typical factors from driver behavioral factors, geometric and environmental factors, and traffic factors. These factors are mostly existed when we are driving on highways or freeways and they all influence significantly to the crash severity levels when traffic crash occurred. Therefore from the transportation perspective, we should study these influential factors so that the severity of traffic crashes occurred can be reduced.
### 4.3.3.3 Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>CRASHSEVERITYFLAG</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE_GRP 1 vs 5</td>
<td>1</td>
<td>0.355</td>
<td>0.158</td>
</tr>
<tr>
<td>AGE_GRP 1 vs 5</td>
<td>2</td>
<td>0.881</td>
<td>0.631</td>
</tr>
<tr>
<td>AGE_GRP 2 vs 5</td>
<td>1</td>
<td>0.574</td>
<td>0.327</td>
</tr>
<tr>
<td>AGE_GRP 2 vs 5</td>
<td>2</td>
<td>0.927</td>
<td>0.704</td>
</tr>
<tr>
<td>AGE_GRP 3 vs 5</td>
<td>1</td>
<td>0.376</td>
<td>0.200</td>
</tr>
<tr>
<td>AGE_GRP 3 vs 5</td>
<td>2</td>
<td>0.942</td>
<td>0.703</td>
</tr>
<tr>
<td>AGE_GRP 4 vs 5</td>
<td>1</td>
<td>0.438</td>
<td>0.271</td>
</tr>
<tr>
<td>AGE_GRP 4 vs 5</td>
<td>2</td>
<td>0.864</td>
<td>0.679</td>
</tr>
<tr>
<td>TIME_GRP 1 vs 5</td>
<td>1</td>
<td>0.856</td>
<td>0.447</td>
</tr>
<tr>
<td>TIME_GRP 1 vs 5</td>
<td>2</td>
<td>0.880</td>
<td>0.648</td>
</tr>
<tr>
<td>TIME_GRP 2 vs 5</td>
<td>1</td>
<td>0.927</td>
<td>0.564</td>
</tr>
<tr>
<td>TIME_GRP 2 vs 5</td>
<td>2</td>
<td>0.953</td>
<td>0.740</td>
</tr>
<tr>
<td>TIME_GRP 3 vs 5</td>
<td>1</td>
<td>0.627</td>
<td>0.380</td>
</tr>
<tr>
<td>TIME_GRP 3 vs 5</td>
<td>2</td>
<td>0.830</td>
<td>0.544</td>
</tr>
<tr>
<td>TIME_GRP 4 vs 5</td>
<td>1</td>
<td>1.072</td>
<td>0.555</td>
</tr>
<tr>
<td>TIME_GRP 4 vs 5</td>
<td>2</td>
<td>1.059</td>
<td>0.811</td>
</tr>
<tr>
<td>NUMBEROFUNITS</td>
<td>1</td>
<td>3.116</td>
<td>2.316</td>
</tr>
<tr>
<td>NUMBEROFUNITS</td>
<td>2</td>
<td>2.087</td>
<td>1.692</td>
</tr>
<tr>
<td>VEHICLEERROR</td>
<td>1</td>
<td>0.002</td>
<td>0.002</td>
</tr>
<tr>
<td>VEHICLEERROR</td>
<td>2</td>
<td>0.997</td>
<td>0.994</td>
</tr>
<tr>
<td>ROADCONTOURID 1 vs 9</td>
<td>1</td>
<td>&gt;999.999</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>ROADCONTOURID 1 vs 9</td>
<td>2</td>
<td>1.528</td>
<td>0.565</td>
</tr>
</tbody>
</table>
The odds ratio presents levels comparison in each influential factor to indicate how each factor affects the analysis results. When the odds ratio is more than 1, it means the level before versus is more dangerous than the level behind versus. Only the conditions of 1 is not in the confidence interval will be discussed because there is no difference in

### Table 4.6: Odds Ratio Estimates in SAS Output

<table>
<thead>
<tr>
<th>ROADCONTOURID 2 vs 9</th>
<th>1</th>
<th>&gt;999.999</th>
<th>&lt;0.001</th>
<th>&gt;999.999</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROADCONTOURID 2 vs 9</td>
<td>2</td>
<td>1.647</td>
<td>0.604</td>
<td>4.488</td>
</tr>
<tr>
<td>ROADCONTOURID 3 vs 9</td>
<td>1</td>
<td>&gt;999.999</td>
<td>&lt;0.001</td>
<td>&gt;999.999</td>
</tr>
<tr>
<td>ROADCONTOURID 3 vs 9</td>
<td>2</td>
<td>2.101</td>
<td>0.740</td>
<td>5.962</td>
</tr>
<tr>
<td>ROADCONTOURID 4 vs 9</td>
<td>1</td>
<td>&gt;999.999</td>
<td>&lt;0.001</td>
<td>&gt;999.999</td>
</tr>
<tr>
<td>ROADCONTOURID 4 vs 9</td>
<td>2</td>
<td>1.838</td>
<td>0.665</td>
<td>5.157</td>
</tr>
<tr>
<td>COLLISIONTYPEID 2 vs 8</td>
<td>1</td>
<td>0.379</td>
<td>0.179</td>
<td>0.801</td>
</tr>
<tr>
<td>COLLISIONTYPEID 2 vs 8</td>
<td>2</td>
<td>0.606</td>
<td>0.445</td>
<td>0.824</td>
</tr>
<tr>
<td>COLLISIONTYPEID 3 vs 8</td>
<td>1</td>
<td>11.431</td>
<td>5.254</td>
<td>24.871</td>
</tr>
<tr>
<td>COLLISIONTYPEID 3 vs 8</td>
<td>2</td>
<td>2.347</td>
<td>1.557</td>
<td>3.540</td>
</tr>
<tr>
<td>COLLISIONTYPEID 4 vs 8</td>
<td>1</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&gt;999.999</td>
</tr>
<tr>
<td>COLLISIONTYPEID 4 vs 8</td>
<td>2</td>
<td>0.300</td>
<td>0.119</td>
<td>0.761</td>
</tr>
<tr>
<td>COLLISIONTYPEID 5 vs 8</td>
<td>1</td>
<td>&lt;0.001</td>
<td>&lt;0.001</td>
<td>&gt;999.999</td>
</tr>
<tr>
<td>COLLISIONTYPEID 5 vs 8</td>
<td>2</td>
<td>0.192</td>
<td>0.124</td>
<td>0.298</td>
</tr>
<tr>
<td>COLLISIONTYPEID 6 vs 8</td>
<td>1</td>
<td>3.571</td>
<td>1.745</td>
<td>7.309</td>
</tr>
<tr>
<td>COLLISIONTYPEID 6 vs 8</td>
<td>2</td>
<td>1.810</td>
<td>1.329</td>
<td>2.465</td>
</tr>
<tr>
<td>COLLISIONTYPEID 7 vs 8</td>
<td>1</td>
<td>1.111</td>
<td>0.491</td>
<td>2.513</td>
</tr>
<tr>
<td>COLLISIONTYPEID 7 vs 8</td>
<td>2</td>
<td>0.993</td>
<td>0.704</td>
<td>1.439</td>
</tr>
<tr>
<td>ISALCOHOLRELATED 0 vs 1</td>
<td>1</td>
<td>0.095</td>
<td>0.064</td>
<td>0.140</td>
</tr>
<tr>
<td>ISALCOHOLRELATED 0 vs 1</td>
<td>2</td>
<td>0.561</td>
<td>0.423</td>
<td>0.744</td>
</tr>
<tr>
<td>POSTEDSPEED</td>
<td>1</td>
<td>1.088</td>
<td>1.076</td>
<td>1.101</td>
</tr>
<tr>
<td>POSTEDSPEED</td>
<td>2</td>
<td>1.031</td>
<td>1.026</td>
<td>1.036</td>
</tr>
<tr>
<td>helmet 0 vs 1</td>
<td>1</td>
<td>1.657</td>
<td>1.281</td>
<td>2.144</td>
</tr>
<tr>
<td>helmet 0 vs 1</td>
<td>2</td>
<td>1.213</td>
<td>1.090</td>
<td>1.352</td>
</tr>
</tbody>
</table>
odds when 1 is included in the confidence interval.

For the collision type, the odds ratio of head-on versus sideswipe, opposite direction for crash severity fatal level is 11.431 and the odds ratio of head-on versus sideswipe, opposite direction for crash severity injury level is 2.347. This means head-on is more dangerous than sideswipe, opposite direction for both crash severity levels. The odds ratio of angle versus sideswipe, opposite direction for crash severity fatal level is 3.571 and the odds ratio of angle versus sideswipe, opposite direction for crash severity injury level is 1.810. Therefore, it means angle is more dangerous than sideswipe, opposite direction for both crash severity levels.

For the helmet used, the odds ratio of helmet non-used versus helmet used for crash severity fatal level is 1.657 and the odds ratio of helmet non-used versus helmet used for crash severity injury level is 1.213. Therefore, it means helmet non-used is more dangerous than helmet used for both crash severity levels.

4.4 Discussion of Results

In this section, the influential factors impact crash severity levels and the influential factors comparison analysis are discussed.

4.4.1 Influential Factors Impact Crash Severity Levels

For the time of crash, evening is significant for the crash severity fatal and injury level. So it means the crash which is taken place in evening will result in fatal more than result in injury.

Number of units is extremely significant for both fatal and injury crash severity
levels. Therefore, it means the number of units get involved in crash will impact the 
crash severity level, such as two units mostly result in injury and three or more units 
mostly result in fatal.

Vehicle in error is significant for the crash severity injury level. Therefore, it means if 
the driver of the motor vehicle causes the crash, the crash will more likely result into 
injury than if the driver of the motorcycle was the one who causes the crash.

For collision type, rear-end has higher probability in fatal crash than sideswipe, 
opposite direction and lower probability in injury crash. Head-on has higher probability 
in fatal crash than sideswipe, opposite direction and higher probability in injury crash. 
Rear-to-rear has lower probability in fatal crash than sideswipe, opposite direction and 
lower probability in injury crash. Backing has lower probability in fatal crash than 
sideswipe, opposite direction and lower probability in injury crash. Angle has higher 
probability in fatal crash than sideswipe, opposite direction and higher probability in 
injury crash. Sideswipe, same direction has higher probability in fatal crash than 
sideswipe, opposite direction and higher probability in injury crash.

Alcohol used is extremely significant for the crash severity fatal level and the crash 
severity injury level. Therefore, it means driver of motorcycle or vehicle that uses alcohol 
will certainly increase the chance of a fatality or injury.

Posted speed limit is extremely significant for the crash severity fatal level and the 
crash severity injury level. Therefore, it means crashes that occur on highways or 
freeways with higher speed limits are more likely to result in injuries and fatalities.

Helmet used of motorcycle occupants is significant for the crash severity fatal level
and the crash severity injury level. Therefore, it means the occupants of motorcycle use helmet will significantly be protected in the crash.

4.4.2 Influential Factors Comparison Analysis

The distribution of crashes for alcohol non-used is shown in the Figure 4.9. The percentage of fatal crashes is 2.94%, the percentage of injury crashes is 70.64%, and the percentage of PDO crashes is 26.42%. The alcohol non-used crashes indicate the occupants of motorcycle did not use alcohol before driving.

![Alcohol Non-used Crashes Distribution](image)
The distribution of crashes for alcohol used is shown in the Figure 4.10. The percentage of fatal crashes is 18.67%, the percentage of injury crashes is 67.62%, and the percentage of PDO crashes is 13.70%. The alcohol used crashes indicate the occupants of motorcycle did use alcohol before driving.

From comparing Figure 4.9 and Figure 4.10, with alcohol used, the percentage of fatal crashes increased and the percentage of injury and PDO crashes decreased. Especially the percentage of fatal crashes with alcohol used is six times of the percentage of fatal crashes with alcohol non-used. Therefore, the conclusion is alcohol influence significantly to the crash severity levels and the occupants should be advised not to drive after drinking, this is an effective way to reduce the crash severity levels or even the occurrences of crashes.
The distribution of crashes for helmet non-used is shown in the Figure 4.11. The percentage of fatal crashes is 4.90%, the percentage of injury crashes is 72.83%, and the percentage of PDO crashes is 22.27%. The helmet non-used crashes indicate the occupants of motorcycle did not use helmet in driving.
Figure 4.12: Helmet Used Crashes Distribution

The distribution of crashes for helmet used is shown in the Figure 4.12. The percentage of fatal crashes is 3.25%, the percentage of injury crashes is 70.58%, and the percentage of PDO crashes is 26.17%. The helmet used crashes indicate the occupants of motorcycle did use helmet in driving.

From comparing Figure 4.11 and Figure 4.12, with helmet used, the percentage of fatal and injury crashes decreased and the percentage of PDO crashes increased. Especially the percentage of PDO crashes with helmet used is increased by 3.9% of the percentage of PDO crashes with helmet non-used. So the conclusion is helmet can effectively protect the occupants of motorcycle and there can be more traffic signs on highways and freeways telling occupants to use helmet in driving, this is an effective way to protect occupants in crash.
4.4.3 Summary

Based on the above analysis, the whole procedure included ordinal or nominal checking of dependent variable in SPSS, modeling building by using stepwise selection in SAS, more specifically, stepwise selection procedure, maximum likelihood estimates, and odds ratio estimates. The transportation professionals can apply the analysis results by considering these influential factors such as road contour, posted speed limit, helmet used, vehicle in error, and alcohol used. On one hand, there can be more traffic safety signs set at the highways and freeways to warn the drivers. The sharp turn sign, number of lanes change sign, speed limit sign, and helmet used sign can be set. On the other hand, every driver should study safety knowledge. Alcohol used is a big problem and all the drivers should know not to drive after drinking. Careful driving is also very important and every driver should have a good driving behavior. The contribution of the analysis results to the transportation profession is based on the two applications mentioned above. By applying these two applications, the final goal of this study is going to reduce the occurrences of motorcycle-motor vehicle crashes and effectively protecting the occupants in driving.
CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

This study examined the influential factors for motorcycle-motor vehicle crashes in the state of Ohio. The main objective of the current study was to determine the influential factors that contribute significantly to the crash severity levels when motorcycle-motor vehicle crashes occur. Based on this main objective, a five years crash data through 2008 to 2012 obtained from the Ohio Department of Public Safety (ODPS) was used for this analysis. In this study, the nominal multinomial logistic regression model was built by using both SPSS and SAS to investigate characteristics of injury and fatality of motorcycle-motor vehicle crashes in the state of Ohio.

The multinomial logistic regression model was used because it has the ability to detect influential factors for crash severity. The dependent variable was crash severity, which was coded into three categories within the Ohio Department of Public Safety (ODPS) crash database. They were fatal, injury, and property-damage-only (PDO). This variable means number for the most severe injury, or in the absence of injury, property
damage involved in the crash. The explanatory variables include several geometric, environmental, traffic, and driver behavioral factors. They are considered to have an effect on the severity of motorcycle-motor vehicle related traffic crashes and they were analyzed by using some powerful statistical modeling techniques in order to determine the most significant ones. A total of eighteen variables (include six grouped variables) were selected for exploratory analysis to investigate characteristics of predictor variables and screen out the most influential ones. A multinomial logistic regression model was built by using both SPSS and SAS to investigate the influential factors for motorcycle-motor vehicle crashes.

The results of the current study suggested that several explanatory variables did influence the likelihood of a motorcycle-motor vehicle crash being injurious or fatal. The influential factors include age, time of crash, number of units, vehicle in error, road contour, collision type, alcohol used, posted speed, and helmet used. Helmet used significantly reduced the likelihood of an injurious and fatal motorcycle-motor vehicle crash. However, the likelihood of a motorcycle-motor vehicle crash being injurious and fatal was increased when especially motor vehicle caused the crash. Road contour influences significantly to the likelihood of an injurious and fatal motorcycle-motor vehicle crash. Posted speed limit, collision type, and number of units also influence the likelihood of a motorcycle-motor vehicle crash being injurious and fatal. Alcohol used significantly increased the likelihood of an injurious and fatal motorcycle-motor vehicle crash. Time of crash and age also influence the likelihood of a motorcycle-motor vehicle crash being injurious and fatal. These factors contribute significantly to the model and
they are considered as influential factors for the likelihood of an injurious and fatal motorcycle-motor vehicle crash. No other factors were considered as significant for the likelihood of a motorcycle-motor vehicle crash being injurious and fatal.

Significant influential factors that increase the likelihood of motorcycle-motor vehicle crashes injury severity, which have been identified in the current study and are in agreement with the previous study include helmet used, vehicle in error, posted speed limit, and time of crash (Peng and Boyle, 2012). Drowsiness and fatigue, distraction, and inattention were significant in the previous study (Peng and Boyle, 2012) and they can be considered as vehicle in error for the current study. Also number of units, road contour, collision type, alcohol used, and age were not significant in the previous study (Peng and Boyle, 2012), but they were significant in the current study. Because of the previous study analyzed the run-of-road crashes and the current study analyzed the motorcycle-motor vehicle crashes, the previous study used the state of Washington’s data and the current study used the state of Ohio’s data. Furthermore, the explanatory variables for the two studies were also different, so this is why there are some influential factors differences between the two studies.

The current study analyzed injuries sustained by all victims of motorcycle-motor vehicle crashes that occurred in the state of Ohio during the five years period through 2008 to 2012. The results from the current study are important in identifying influential factors that increase the likelihood of motorcycle-motor vehicle crashes sustaining various injury severity levels. The contribution is the likelihood of motorcycle-motor vehicle crashes occurred can be reduced effectively by studying the influential factors.
Therefore, it is recommended that characteristics of drivers who get involved, or more correctly, drivers who cause the occurrences of motorcycle-motor vehicle crashes should be educated on the influential factors affecting such crashes in order to take effective countermeasures. That is, taking notice of traffic signs including roadway signs and speed limit signs, using safety equipment in driving, having a good driving behavior, and driving without drinking for reducing the severity of motorcycle-motor vehicle crashes that occur on Ohio’s roadways and streets.
REFERENCES


[14] Eric R. Teoh, Marvin Campbell. Role of motorcycle type in fatal motorcycle crashes


