Statistical Modelling on the Severity of Road Accidents in Great Britain

Nurhidayah Hamdan

A thesis presented for degree of

Doctor of Philosophy

Department of Mathematical Sciences

University of Essex

February 2022

I would like to dedicate this thesis to my loving parents,
my supportive husband, my son and my daughter.

Declaration

The work in this thesis is based on research carried out with my supervisor Dr Hongsheng Dai and Dr. Haslifah Hashim at the Department of Mathematical Sciences, University of Essex, United Kingdom. No part of this thesis has been submitted elsewhere for any other degree or qualification, and it is all my own work, unless referenced, to the contrary, in the text.

Nurhidayah Hamdan
February 2022

Acknowledgements

All praise to God for the blessings and opportunity bestowed upon the undertake my PhD in Statistics. I would like to express my utmost gratitude to individuals who had contributed directly and indirectly during the producing this dissertation.

I would like to express my appreciation to the Department of Mathematical Sciences, especially to Dr. Hongsheng Dai and Dr. Haslifah Hashim for the continuous support, patience and willingness to guide me in overcoming obstacles in the completion of this thesis.

I wish to express my heartiest gratitude to my beloved husband, son and daughter Zarul Ikhwan, Ar-Razi and Alania for their support and constant encouragement throughout my study. I am also indebted to my family, especially my parents, who have really support me spiritually in all possible ways.

I am also grateful to all lecturers and loyal friends for their constant help and advice whenever I face difficulties on finishing this research project. Without them, the completion of this thesis will never come true.

Thank you.

Abstract

Great Britain has a modern road network and is well-known with the advanced technology in road engineering. Although with excellent road infrastructure, road accidents remain one of the main concerns in road safety literature among researchers and policymakers. One of the main strategies for improving road safety is to identify the contributing factors and then to develop countermeasures. There have been numerous studies that analyse road accident severity including binary outcome models, ordered discrete outcome models, unordered multinomial discrete outcome models, and other data mining approaches. The aim of this thesis is to identify the contributing factors affecting road accident severity in Great Britain and estimate the accident cost for all types of accident severity. For accident severity study, three statistical models are selected: multinomial logistic regression (MNL) model, log-linear graphical model and multinomial logistic with random effects (MNLRE) model. Markov Chain Monte Carlo (MCMC) simulation method by applying random walk Metropolis-Hastings (M-H) algorithm is used for parameter estimation in the MNLRE model. Accident cost study is investigated by applying three models: Gamma, Weibull and Log-normal distribution.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background of the Study

Motorization has enhanced the quality of life to many individuals and nations as a whole. One of the considerable burden lies in road transportation is road accident. Every day around the world, the number of injuries and deaths in road accident is raising together with the growing number of population and rapid change of motorization industry. The real fact is, wherever and whenever there are people and motorized transport, there will be road accidents in that area. Road accident is deemed as unavoidable risk that we have to endure in maintaining high levels of mobility.

Social and economic costs involved in road accident are undeniably high and it creates a development crisis within the nation. Despite of large magnitude of loss, road accident still receives low attention for higher authorities and other associated agencies. Hence, it requires concerted effort at the international, national and local levels for effective and sustainable prevention to reduce road casualties.
Road accident is a global major public health problem, and should be treated as a shared multidisciplinary responsibility. Worldwide, it is estimated almost

1.35 million deaths in 2016, and relative world's population showed constant rate of death[54]. Obviously it forms a large population, this figure suggests that road safety progress to control the situation from getting worse. Forecast estimation showed that, between 2000 and 2020 low and middle income countries will experience 80% increment of total number of road accidents and injuries [57, 58].

Currently, road traffic injuries rank as the eighth leading cause of death across all age groups, ahead of HIV/AIDS, tuberculosis OR diarrhoeal diseases (See Table 1.1). The detrimental impact of road accidents is more likely hazardous than real disease since there is neither medication nor right cure to prevent it [57, 58].

Table 1.1: World's Leading Causes of Death for Year 2004 and 2016

Rank	Leading Cause (2004)	Leading Cause (2016)
1	Ischaemic heart disease	Ischaemic heart disease
2	Cerebrovascular disease	Stroke
3	Lower respiratory infections	Chronic obstructive pulmonary disease
4	Chronic obstructive pulmonary disease	Lower respiratory infections
5	Diarrhoeal disease	Alzheimer's disease and other dementias
6	HIV/AIDS	Trachea, bronchus, lung cancers
7	Tuberculosis	Diabetes mellitus
8	Trachea, bronchus, lung cancers	Road traffic injurie
9	Road traffic injuries	Diarrhoeal disease
10	Premature and low birth weights	Tuberculosis

WHO Global Health Estimates [53] [54]

Road accidents affect all age groups, but analysis showed that young generation is highly exposed to the traffic accident risk. From out of 85% global road deaths, WHO recorded 96% of all children killed in road crashes worldwide, occur in low and middle income countries. Table 1.2 shows that road accidents injuries are the main contributor to world's fatality of people aged between 15 to 29 years old.

Table 1.2: Ranking of Road Traffic Injuries as World's Causes of Death

Age Group (Years)	Number of Rank
0-4	14
5-14	2
15-29	1
30-44	3
45-69	8
More than 70	20

Table is extracted from [53]

Road accident resulted in numerous negative effects to all parties involved, including the victims, society and the authority. Despite of pain, grief and suffering, the victim and victim's family have to swallow several consequences financially, physically and emotionally. It reveals that road accident tragedies resulting with catastrophic loss in terms of inconsolable losses of human life as well as destruction of properties incurred.

In economic and monetary perspective, loss is described as cost that involved when an accident happened directly and indirectly. Serious concern and double effort from both authority and society are needed to reduce the occurrence of road crashes.

1.1.1 Overview of Road Accidents in Great Britain (GB)

GB is a developed country and known as one of the leader of road safety worldwide. Overall, the total number of road casualties in GB is estimated between 630 to 800 thousand cases annually, including the unreported ones [16]. In 2013, 1,713 people killed and 21,657 people injured or disabled from road accident [16]. It found that

138,660 road accident cases reported throughout GB with 252,913 vehicles involved and 183,670 casualties occurred. It shows reduction compared to year 2012 with 145,571 accident cases, 265,877 vehicles and 194,723 casualties reported. During the 1-year period, 117,428 (84.7%) cases involved slight injury, 19,624 (14.2%) resulted in serious injury and 1,608 (1.2%) turned out to be fatalities. The distribution of the accident severity is shown in Figure 1.1.

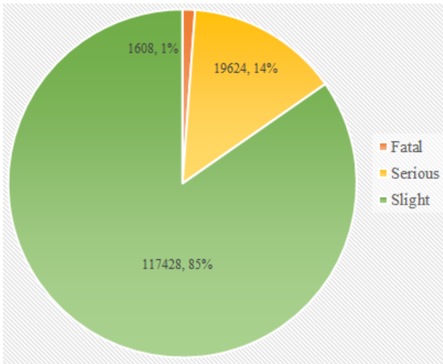


Figure 1.1. Percentage of Road Accident Severity in GB for Year 2012

Great Britain (GB) has a modern road network and known with the advanced technology for road engineering. Although with good road infrastructure, road traffic accidents still remain as one of the main concerns in transportation problem. GB categorise its road system into two types, motorways and non-motorways. Around 20% of GB motor traffic used motorways in 2013, but accounted for only 6% of road deaths (100 deaths) and 3% of serious injury (660 serious casualties) [16]. For non-motorway, it is divided into two road types which are rural and urban roads. Rural roads carried 53% of traffic, but accounted for two thirds of road fatalities in 2013. Meanwhile, urban roads carried 47% of traffic with 57% seriously injured casualties occurred [16]. Positive improvement in GB road accident is observed in year 2013 such as the lowest record of road deaths since 1926, serious injury accident have decreased consistently with the current record shows it is 43% lower than 2000, and total number of casualties has also decreased by 6%

compared to 2012 [16].

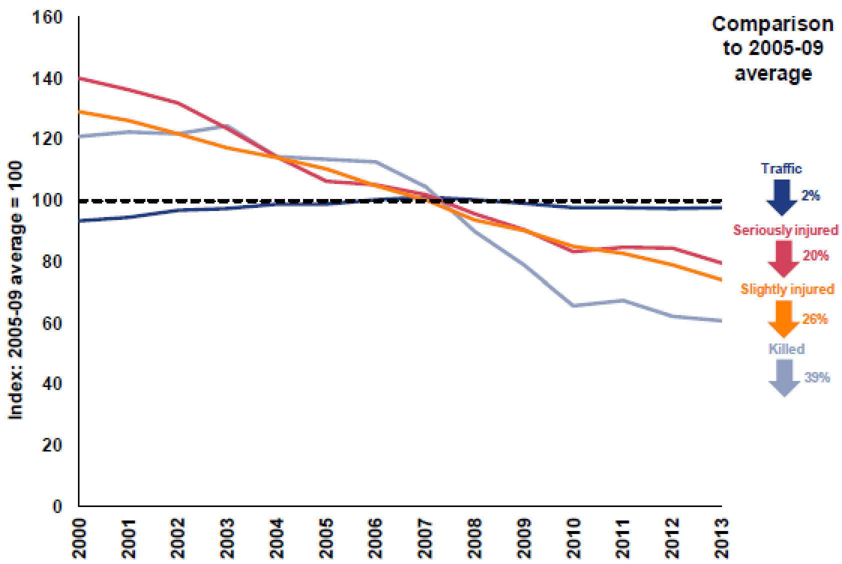


Figure 1.2. Traffic and reported casualties by severity (2000-2013)

Department for Transport (DfT) of GB has found out that the downward trend in road fatalities especially among vulnerable road users in 2012 and 2013 due to the weather patterns [16]. The most common recorded contributory factor in 2013 for all types of accident severity is from the driver/rider negligence. "Driver/rider failed to look properly" was mostly recorded factor in serious and slight injury with 36% of serious and 43% of slight cases. Meanwhile, "loss of control" is a common factor contributed 34% of fatal accident cases [16].

As described by Ogden (1996), road traffic may be considered as a system that consists of various components. These components, such as the human, the vehicle and the roads, interact with each other.

Road accident is defined as an occurrence of personal injury on the public highway (including footways) due to the negligence or omission by any party concerned resulting in a collision which involved at least one vehicle or vehicle collision with a pedestrian and is reported to the police within 30 days. In GB, damage-only accidents, involves no human casualties or accidents on private roads or car park are not considered in this category. Casualty involved when a person killed or injured in an accident. In this context, there are three categories of casualties in GB, which are fatal, serious injury and minor injury [16].

Types of Severity	**Definition**
Fatal	A road accident in which one or more person were killed less than 30 days after the date of event.
Serious	A road accident in which at least one person sustained serious injury but none killed.
Slight	A road accident in which at least one person is injured but none killed or seriously injured.

While positive improvement has been incline towards the government aims, further reductions in casualties are needed. Number of casualties are closely related to the numbers of traffic injuries. Therefore, any reduction in casualties is associated with a reduction in injury accidents. In addition, to improve road safety, causes of road accidents and the relationship between accident severity and relevant factors need to be investigated and studied.

DfT has declared to spend approximately £14.7 billion in 2013 for accident prevention excluding unreported cases. It showed 3% decrease compared to 2012 estimation. With large gap reduction of road casualties that occurred, it is important to keep the downward trend continues parallel with the cost involved. Therefore, it highlight the in-depth study requires for identifying and estimating both the severity and cost of road crashes by implementing cost-effective countermeasure.

1.2 Problem Statement

1.2.1 Public Health Perspective

Public health comprises of knowledge from variety disciplines such as medicine, health services, epidemiology, sociology, engineering, economics, behavioural science, education, statistics and to name a few. It is a unique approach since it is science based, and population focus. Public health model for prevention has been applied to a wide variety of infectious and chronic diseases with remarkable success [72].

Framing the road crashes problem as a predictable and preventable matter, makes road crashes injury a salient issue in public health and preventive medicine. WHO [57, 58] reported total number of road traffic deaths worldwide and injuries forecast to rise by 65% between 2000 and 2020, and in low and middle income countries deaths are expected to increase as much as 80%. Hence, there is an urgent need to observe the worsening situation in road deaths and injuries therefore requires integrated efforts for effective and sustainable prevention action.

Over the past years, a range of methods have been used for accident prevention. Nevertheless, there is lack of action or effort has been done by applying public health knowledge. Public health are the popular tools among the developed countries to analyse the cause and impact of road crashes. There is many tools of public health discussed among health practitioners but the foundation of public health study is derived on pre-crash phases, crash phases, and post-crash phases [21].

Numerous benefits offered by public health, but less application in solving road accident problem. Therefore, this study perhaps be a starting point to discover the problem in public health perspective, then comes out with the preventive measure in the context of public health, with the aim of maximizing benefits for the entire population.

1.2.2 Modelling the Severity and Cost of Road Accidents in GB

The trend of road accident injuries and deaths has become common transport issue in GB. Gaining a better understanding of the factors that affect the likelihood of a vehicle crash has been an area of research focus for many decades with the aims to improve road accident prediction model and to provide direction for policies and countermeasures aimed at alleviating number of crashes [64]. To address this issue, there has numerous research conducted to identify unobserved explanatory variables (random effects) that are significantly affect road crash frequency and severity and each research has its own recommendation to lessen the impact of accident frequency and severity. However, the number of studies focusing on accident severity still limited in GB. Most of the successful study conducted in United States and the data used is quite old. This study attempts to fill that gap and extends the related literature by applying appropriate statistical prediction models. Latest road accident data in GB is used and method exploration is executed to find which methods are the most fit to the objective. The purpose of this research in the first part is to present the modeling and analysis results of the relationship between roadway design, environment factor, and traffic attributes with the severity of road accidents in GB. Bayesian method by including random effects is applied as a new approach to model road accident. The findings of this study could also provide sufficient statistical evidence and good support to the road authority to implement impactful intervention and improve road safety in GB. Moreover, this study is a useful guide to researchers and practitioners on evaluating the methodological issues attributing to road crashes.

The second part of this study is to model accident cost after the contributory factors are known. Proper model valuation of accident cost can address the road crashes problem from actuarial monetary perspective. This model will diverge into three categories according to type of severity. Economic cost of road accidents can be evaluated from the cost involved such as property damage, administration costs, lost output, medical costs, and also pain, grief and suffering cost. As the accident cost is expressed in monetary term, continuous data will be used in modelling road crashes cost.

Accident costing is choose to meet research aim as it had been intensively practised by developed countries such as UK, Canada, United States, Australia, Japan, France, Germany, etc. Each country has different method to calculate the accident cost in order to meet certain country's aim. In Indonesia, the approximate accident cost ratio between fatal and non-fatal is 11:1 [17], while for UK is 29:1 [17], Jordan 15:1 [3], Bangladesh 7:1 [19], Australia 5:1 [11] and Vietnam 2:1 [6]. Overall, average accident cost for fatal resulting with higher cost than non-fatal. Since the early 1990s, UK has consistently used willingness to pay (WTP) method to evaluate road casualty.

Yet there is no study conducted in GB modelling both severity and cost of road accidents in one research to improve road safety and allocate accident cost accordingly. By establishing this model, sustainable and effective prevention action can be implemented to enhance the road safety level in GB. Conclusion can be derived in the perspective of public health as overall with the support from quantitative analysis.

1.3 Research Questions

The research questions of the study are as follows:

i. What are the factors that contribute to road accidents casualty in GB?
ii. How much is the cost involved for fatal, serious and slight road accident? How much is the ratio difference?
iii. What is the conclusion and recommendation in the perspective of road safety given the statistical and cost evidence?

1.4 Research Objectives

The objectives to be achieved from this study are:

i. To identify factors that contribute to road accidents in GB by applying appropriate statistical prediction model.
ii. To study and compare the severity of GB road accident by using accident cost model as reference.
iii. To advice and derive conclusion from the road safety perspective with the evidence from statistical and actuarial method.

1.5 Significance of the Study

Road crashes problem is a complicated problem that requires teamwork and intervention from various directions. The responsibility should not solely handle by police nor transport department. It requires integrated efforts and knowledge from many fields, hence this study is done by combining several major disciplines which are statistics, actuarial, economy and public health.

The findings of this study could also provide sufficient statistical evidence and good support in order to determine the dynamics of the crash, then propose the road authority to implement impactful intervention and improve road safety in GB. Since there is small amount of research in statistics and actuarial crossed with road accidents problem in GB, it is expected to bring the development of countermeasures in minimizing human and economic cost impact among society. Moreover, the findings are also useful to guide researchers and practitioners to further improvise the methods applied.

Chapter 2

Literature Review

2.1 Introduction

The attempt to reduce road crashes have emerged as a vital public health endeavour for the past decades. In 2014, number of road accidents which involved casualties are reduced, but the increment percentage of fatalities worldwide is increased.

2.2 Public Health Perspective

Generally, the responsibility of road crashes problems has been assumed to bear by the transport sector only. With the growing population and motorization, this assumption is not applicable anymore since the fast rate of road crashes has overcome the preventive measure done. It affects not only transportation system, but also economic systems, health systems, jobs, families and civil society [72]. As C. Everett Koop, former US Surgeon General said about childhood injuries, if a disease were killing our children in the proportion that accidents are, people would be outraged and demand this killer be stopped (National SAFE KIDS Campaign, 2006). From this statement, it is clearly stated that road accident problem has reached its alarming state worldwide, thus it needs attention from multi-disciplinary sectors to find the effective preventive measure to ameliorate the losses.

There are many public health tools that has been used by researchers and health practitioners to relate traffic safety and science-based population study. The early public health tool introduced to fight against road crashes problem is the classis epidemiologic triad, which involved host, agent and environment. Sleet et al. (2007) [72] explained traffic injury as the results from the interaction between injury-producing agents (the causative element and the vehicle or vector carrying it), host factors (the person affected) and the environment (conditions in which the host and agent find themselves) (Refer Figure 2.1).

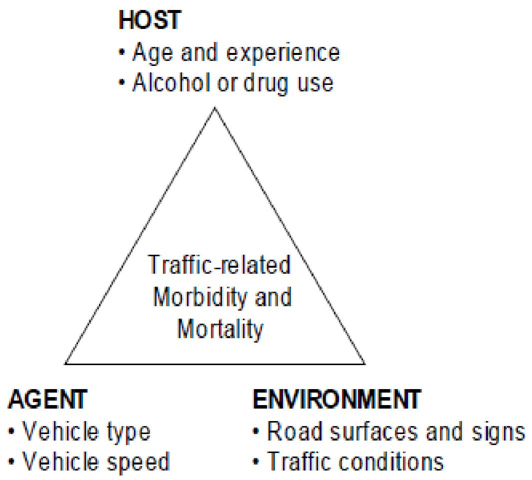


Figure 2.1. Classic Epidemiologic Triad on Traffic Safety

Haddon (1968) [21] came out with the scientific idea to prevent road accidents by adapting public health concept with nine-cell Haddon matrix models. Haddon described the model into three phases: pre-crash, crash and post-crash (Figure 2.2).

		FACTORS		
PHASE		HUMAN	VEHICLES AND EQUIPMENT	ENVIRONMENT
Pre-crash	Crash prevention	Information Attitudes Impairment Police enforcement	Roadworthiness Lighting Braking Handling Speed management	Road design and road layout Speed limits Pedestrian facilities
Crash	Injury prevention during the crash	Use of restraints Impairment	Occupant restraints Other safety devices Crash-protective design	Crash-protective roadside objects
Post-crash	Life sustaining	First-aid skill Access to medics	Ease of access Fire risk	Rescue facilities Congestion

Figure 2.2. The Nine-cell Haddon Matrix

Centers for Disease Control and Prevention (CDC) (1985) has brought four-step model to adapt in road safety promotion by applying it as public health framework with epidemiologic perspective. This model has been improved from the previous study and the action taken from defining problem to build solution must be in sequential manner (Figure 2.3).

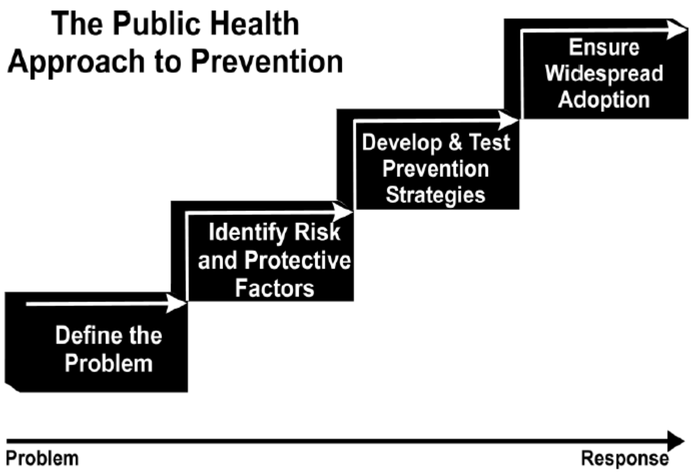


Figure 2.3. Four-step Model by CDC

2.3 Modelling Road Accident Severity

Hosseinpour et al. (2014) [25] has discussed the roadway geometric design and traffic characteristics on the frequency and severity of head-on crashes in Malaysia over the period of 2007-2010. Random-effect generalized ordered probit model were used and the results indicated that the severity of head-on crashes were affected by the horizontal curvature, paved shoulder width, terrain type and side friction, whereas access points, land use, and the presence of median reduced the probability of severe crashes. This study also showed that random-effect generalized ordered probit model outperform the standard ordered probit model according to goodness of fit measures.

Ma and Kockelman (2006) [39] assessed the effects of roadway characteristics with a joint model of crash frequency and severity at Washington State highways in 1996. Based on the Bayesian multivariate Poisson regression modelling results, three variables were found to be significantly associated with fatal crashes: posted speed limit, degree of curvature and right shoulder width; while five variables contributed to disabling injury crashes: posted speed limit, number of road lanes, presence of median, right shoulder width and presence of rolling terrain. Oddly, some of the significant variables in disabling injury crashes are not contributing any impact on fatal injury counts.

Anastasopoulos and Mannering (2009) [5] explored negative binomial random-parameters count models as another methodological alternative to analyse accident frequency that occurred on rural interstate highways in Indiana for 5-year period (1995-1999). The analysis results indicated that the roadway segment's international roughness index (IRI), rutting indicator, road segment length, median barrier indicator, interior shoulder width, horizontal curve's degree of curvature per mile, and the average annual daily traffic (AADT) of passenger cars were found to have significant influence towards number of accident occurrence.

Lee and Mannering (2002) [36] demonstrated that there was a significant difference in the factors that determined run-off-roadway accident frequencies in urban and rural areas. This researcher studied on the impact of roadside features on the frequency and severity of run-off-roadway accidents at 96.6 km section of highway in Washington State. This analysis applied two different prediction models which are zero-inflated count models and nested logit models are applied to identify the factors that significantly influence the frequency and severity of accidents. In rural model estimation, 5 factors that contributed to accident frequency which are speed limit, median width, distance from the outside shoulder edge to light poles, number of isolated trees in a section, and the presence of cut-slopes in the roadway; whereas factors that influenced accident severity are speed limit, weather factor, driver behaviour (driver's age and alcohol abused), the presence of roadway indicator such as asphalt shoulder, narrow shoulder indicator, instrumented guardrail, miscellaneous fixed object, sign support, tree group and utility pole.

Using data in Texas, Schneider et al. (2009) [65] studied about the impacts of factors associated on the degree of driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways by applying multinomial logit model. Horizontal curvature was found to affect driver injury severity and were more likely to occur at moderate radius and lower-speed curves. Run-off-the-road crashes that involve collisions with roadside objects, found to give the most significant impact to injury severity. The results also showed that severe injuries were more likely to affect female and older drivers. Other than that, accident severity is significantly determined by driver actions and behaviors such as seat belt application, drug or alcohol abuse and driver's lethargy. Many of the previous study has suggested that the presence of median traffic barrier has significantly affected crash severity.

Therefore, Miaou et al. (2005) [45] engaged a study to estimate the effect of installing median traffic barrier installation particularly related with crash frequencies and severities that occurred in Texas highways. Ordered multinomial logit model were employed to model for various type of median crashes with and without barrier. No factor was found to be statistically significant in the severity model for cross-median crashes (most likely due to small sample size); while median width, AADT, number of lanes and posted speed limit were found to be significantly associated with accident severity.

Khorashadi et al. (2005) [30] conducted a research to investigate the causes of injury severity that focus on large trucks in rural and urban roadways in California from year 1997 to 2002. Multinomial logit approach was used for analysis and it was found that severe/fatal accidents involving single-unit trucks is more likely to happen in urban areas compared to rural areas. Alcohol and drug factor also showed it was highly influenced the severity among drivers in urban areas, compared to rural areas. Most of the significant factors that associated with injury severity were similarly observed in both urban and rural location. However, there were some variables that significantly contributed in rural but not urban areas, and vice versa. As summary, below are the several methods used by previous researchers to analyse road crash frequency data:

Model Type	Previous Research
Artificial Neural Networks	Delen et al. (2006) [14]
Bayesian Hierarchical Binomial Logit	Helai et al. (2008) [26]
Bayesian Ordered Probit	Xie et al. (2009) [80]
Binary Logit and Binary Probit	Rifaat and Tay (2009) [62]; Haleem and Abdel-Aty (2010) [22]; Peek-Asa et al. (2010) [59]; Kononen et al. (2011) [33]; Moudon et al. (2011) [48]
Bivariate Binary Probit	Lee and Abdel-Aty (2008)
Poisson	Jones et al. (1991) [29]; Miao and Lum (1993) [46]; Miao (1994) [44]
Negative binomial/ Poisson-gamma	Maher and Summersgill (1996) [40]; Lee et al. (2002) [36]; Lord et al. (2009) [37]
Poisson-lognormal	Miao et al. (2005) [45]; Lord and Miranda-Moreno (2008) [38]
Zero-inflated Poisson and zero-inflated negative binomial	Lee and Mannering (2002) [36]; Shankar et al. (2003) [69]
Gamma	Oh et al. (2006) [51]; Noorizam and Kamaruzzaman (2005)
Random-effects model	Hausman et al. (1984) [24]
Random-parameters model	Anastasopoulos and Mannering (2009) [5]; Hosseinpour et al. (2014) [25]
Bivariate/multivariate model	Ma and Kockelman (2006) [39]; Ye et al. (2009); Park et al. (2010) [56]
Finite mixture/Markov switching	Malyshkina et al. (2009)[43]; Park and Lord (2009)[55]
Hurdle Poisson/ Hurdle negative binomial	Hosseinpour et al. (2014) [25]

2.4 Accident Costing

Countries such as United Kingdom, Australia and United States had implemented the accident costing in the past few decades. Official estimate of road accident costs have been prepared in most of the highly motorized countries all over the world for a number of years [19]. UK has superior infrastructure quality for roads and highways in order to support more than 500,000 vehicles daily. Assumption made, since motorization rate in UK is quite high, this will lead to the increasing accident rate also.

The cost of traffic accidents was estimated without sufficient detailed information or it was estimated based on the unit accident costs derived for other developed countries [3]. Hence, it is believed that every country has different accident costing evaluation approach and estimation result according to their economic conditions, population growth and developmental factor.

Costing road accidents has numerous advantages. According to Silcock (2003) [70], accident costing can be used for resource allocation at a national level to ensure road safety is ranked equitably, in terms of investment in the improvement of road safety. Accident costing also can be utilized to ensure the best use is made of any investment, through economic appraisal and cost benefit analysis.

Downing (1997) [17] also mentioned that costing approach allows for all negative consequences to be valued and comparisons made with other national problems. Other than that, cost benefit analysis of alternative road improvement schemes enables the expenditure on road safety to be optimized. Accident costing highlights the socioeconomic burden of road accidents. Furthermore, knowledge on accident costs allows safety impacts to be economically justified (Road Safety Guidelines for the Asian and Pacific Region, 2010).

In the study of 'How Much Do Road Accidents Cost the National Economy?' done by Elvik (2000) [20], it found that all economies would recommend to include an economic evaluation of lost quality of life in the road accident costs. Al-Masaeid et al. (1999) [3] in the research of accident costing in Jordan suggest that the 1996 traffic accident in Jordan cost the country about JD 103 million ($US 146.3) and this cost would be double if appropriate willingness to pay was used.

Downing (1997) [17] had made comparison between accident costs in Indonesia and United Kingdom, the study revealed that UK accident costs are much higher per accident than for the same severity in Indonesia which are 68:1 for fatal accidents, 27:1 for non-fatal accidents, 14:1 for slight injury and 12.5:1 for damage only cases. In the research of official economic valuations of accident fatalities in 20 motorized countries studied by Elvik (1995) [19], it conclude that, when the economic evaluation of lost quality of life is included, the economic valuation of a traffic accident fatality is more than twice as high as in countries where lost quality of life is excluded. Furthermore, Trawen et al. (2002) [73] had studied the international comparison of accident cost of a fatal casualty of road accidents in 1990 and 1999 said that, each countries used different method in calculating their accident cost that suits the government and society objective.

Silcock (2003) [70] declared that if decision makers are genuinely concerned about the quality of life and social well-being of their citizens, then the willingness to pay method should be used. The choice of accident costing method depends on the purpose, objective, capacity and data available of each country [6]. However, this method is extremely difficult to apply in developing countries, based as it is on the completion of complex questionnaires which relate to perceived risk and payment by individuals to avoid a given level of risk [70].

2.5 Accident Cost Methods

According to Silcock (2003) [70], there two common methods to use in evaluating accident cost which are:

i. Lost-output/human capital method
ii. Value of risk change/willingness to pay method

2.5.1 Lost-output/human capital method

Lost output/human capital method is classified into two cost components, which is direct and indirect cost. Direct cost is the cost of vehicle damage, medical treatment, police and administration [17]. Generally, this cost is the major financial burden to the accident victim in short and long term. Vehicle damage is the largest portion of this cost component, followed by medical treatment cost while police and administration cost is the lowest. Indirect part of this cost is costs that are due to a loss of future output. It refers to the loss to the economy of the productive capacity from those affected by road accident tragedies [70]. Loss of a key person in an institution will affect many activities in the long run. It will lead to a great loss impact to the family, organization and society.

Therefore, the evaluation of accident cost using this method only concentrate on the actual accident cost involved. In other aspect, this method does not include the evaluation of psychology cost to portray the pain, grief and suffering experienced by the accident victim.

2.5.2 Value of risk change/willingness to pay method

The human capital method is well suited to the objective of maximizing the wealth of the country but is not so appropriate for cost benefit analysis [17]. Around 1970 the human capital approach was criticised by several economists as being inconsistent with the theoretical principle of cost benefit analysis [19]. In a recent research, sum to reflect 'pain, grief and suffering' is added to the human capital method to capture some of the human considerations of the willingness to pay method [17]. However, the willingness to pay method is not applicable to be applied in all cases since every individual has their own perspective in deciding the cost that they willing to pay for a reduced risk in preventing road accident tragedy.

Chapter 3

Methodology

3.1 Introduction

This chapter describes and justifies the research methodology and framework that was adopted to develop statistical analysis in this study. There are three statistical analysis employed for accident severity models, and another three models for accident cost. To understand the basis of research methodology applied, firstly the scope of study is explained, data sources is described, then followed with the discussion of chosen models. Arguments from literature reviews are presented justifying the choice of specific research methods selected. The methodology employed is considered to be the most appropriate strategy in the context of this study objective and the data fittings.

3.2 Scope of Study

Great Britain (GB) has a modern road network and known with the advanced technology for road engineering. Although with modern road infrastructure, road traffic accidents still remain as one of the main concerns in transportation and public health concern. Therefore, in-depth research is worth to carry out in GB.

Furthermore, GB is chosen because of the ease to acquire an official data. When this study commenced, year 2013 is selected as it was the latest year that the official data has published and available that time. Due to time constraints, only one-year data is used for accident severity and accident cost model research. However, this study can be extended to other countries data with wider time horizon.

3.3 Accident Severity Data

Road accident statistics in the Great Britain are officially collected by the police in STATS19 form, then maintained by the Department for Transport (DfT). The source of the road accident data used in this research is the national STATS19 database for 2013 was extracted from the archive dataset using Microsoft Excel. The data were provided by the UK Data Archive at the University of Essex and can be obtained for replication after authorisation from the Data Archive. STATS19 provides detailed data about each accident including date and time, location, accident severity levels, local conditions at the time of the accident such as weather and visibility, personal details about the driver and type of vehicle involved. Accident severity is classified using the following definitions:

- slight, an accident in which at least one person is slightly injured but no-one is killed;
- serious, in which at least one person is seriously injured but no-one is killed;
- fatal, in which at least one person is killed in the accident.

During 2013, there were 138,660 road accidents reported to the police in Great Britain involving 252,913 vehicles and 183,670 casualties ([16]). STATS19 record is divided into three sets of database which called as 'Accident', 'Vehicle' and 'Casualty'. 'Accident' database compiled the information of each accident regarding the location, date, time, weather and light conditions, etc. 'Vehicle' database focused on the all vehicles that involved in the accident, such as vehicle type, vehicle manoeuvre, vehicle location, junction location from vehicle, etc. 'Casualty' database explained in details about the casualties reported in the accident such as casualty class, sex and age of casualty, pedestrian location, etc. All accidents have their own 'Accident Index Number' in the record. For example, accident with index number 201301BS70 has one record in 'Accident' database, two records in 'Vehicle' database (because of two vehicles involved in that accident), and four records in 'Casualty' database (because there were four accident victims in the scene). Some of the database has incomplete information. Hence, accident index number is useful to double check and match case between database. In this study, only 'Accident' and 'Vehicle' database will be used for further analysis.

3.3.1 Accident Data

The study of accident severity frequency model commenced by explaining 'Accident' data that used in this study. From the total accidents reported, 117,428 (84.7%) were slight injuries, 19,624 (14.2%) were serious injuries and 1,608 (1.2%) were fatalities ([16]). The missing value cases were discarded from the original data in order to achieve a better prediction model. Thus, the trimmed sample contains 82,570 road accidents, 71,574 slight injuries, 10,421 serious injuries and 575 fatal cases. Table 3.1 and 3.2 shows the summary statistics in the 'Accident' data.

Table 3.1: Summary Statistics for Accident Data

Variable		Fatal		Serious		Slight		Total
		n	%	*n*	%	*n*	%	
1st Road Class	Motorway	9	1.57	63	0.60	698	0.98	770
	A(M)	1	0.17	5	0.05	93	0.13	99
	A	331	57.57	5062	48.57	35628	49.78	41021
	B	88	15.30	1438	13.80	9109	12.73	10635
	C	37	6.43	845	8.11	6209	8.67	7091
	Unclassified	109	18.96	3008	28.86	19837	27.72	22954
Road Type	Roundabout	24	0.01	934	8.96	8362	11.68	9320
	One way street (from 2005)	8	0.01	218	2.09	1494	2.09	1720
	Dual carriageway	63	0.01	988	9.48	7388	10.32	8439
	Single carriageway	471	0.01	8177	78.47	53214	74.35	61862
	Slip road (from 2005)	8	0.01	79	0.76	916	1.28	1003
	Unknown	1	0.01	25	0.24	200	0.28	226

Table 3.2: Summary Statistics for Accident Data (Continued)

Variable		Fatal		Serious		Slight		Total
		n	%	n	%	n	%	
Speed Limit (permanent)	1	4	0.70	233	2.24	1365	1.91	1602
	2	308	53.57	7419	71.19	54256	75.80	61983
	3	53	9.22	888	8.52	6042	8.44	6983
	4	38	6.61	362	3.47	2119	2.96	2519
	5	139	24.17	1227	11.77	5718	7.99	7084
	6	33	5.74	292	2.80	2074	2.90	2399
Junction Detail	Not at junction or within 20 metres	0	0.00	0	0.00	1	0.00	1
	Roundabout	37	6.43	1181	11.33	11177	15.62	12395
	Mini-roundabout	8	1.39	189	1.81	1589	2.22	1786
	T or staggered junction	363	63.13	6012	57.69	37801	52.81	44176
	Slip road	19	3.30	217	2.08	1729	2.42	1965
	Crossroads	86	14.96	1657	15.90	11920	16.65	13663
	More than 4 arms (not roundabout)	5	0.87	136	1.31	1082	1.51	1223
	Private drive or entrance	32	5.57	663	6.36	4019	5.62	4714
	Other junction	25	4.35	366	3.51	2256	3.15	2647
Junction Control	Authorised person	0	0.00	21	0.20	161	0.22	182
	Auto traffic signal	77	13.39	1574	15.10	12745	17.81	14396
	Stop sign	5	0.87	71	0.68	594	0.83	670
	Give way or uncontrolled	493	85.74	8755	84.01	58074	81.14	67322
2nd Road Class	Motorway	6	1.04	79	0.76	818	1.14	903
	A(M)	3	0.52	8	0.08	76	0.11	87
	A	71	12.35	1405	13.48	12380	17.30	13856
	B	36	6.26	659	6.32	4864	6.80	5559
	C	62	10.78	786	7.54	5736	8.01	6584
	Unclassified	397	69.04	7484	71.82	47700	66.64	55581
Light Conditions	Daylight	361	62.78	7445	71.44	53290	74.45	61096
	Darkness - lights lit	163	28.35	2502	24.01	15691	21.92	18356
	Darkness - lights unlit	3	0.52	57	0.55	348	0.49	408
	Darkness - no lighting	41	7.13	273	2.62	1235	1.73	1549
	Darkness - lighting unknown	7	1.22	144	1.38	1010	1.41	1161
Weather Conditions	Fine no high winds	488	84.87	8745	83.92	58561	81.82	67794
	Raining no high winds	50	8.70	944	9.06	7597	10.61	8591
	Snowing no high winds	1	0.17	79	0.76	760	1.06	840
	Fine + high winds	14	2.43	137	1.31	929	1.30	1080
	Raining + high winds	7	1.22	135	1.30	1002	1.40	1144
	Snowing + high winds	1	0.17	14	0.13	152	0.21	167
	Fog or mist	0	0.00	41	0.39	229	0.32	270
	Other	6	1.04	162	1.55	1129	1.58	1297
	Unknown	8	1.39	164	1.57	1215	1.70	1387
Road Surface Conditions	Dry	416	72.35	7673	73.63	51910	72.53	59999
	Wet or damp	151	26.26	2587	24.82	18167	25.38	20905
	Snow	2	0.35	64	0.61	731	1.02	797
	Frost or ice	5	0.87	94	0.90	729	1.02	828
	Flood over 3cm. deep	1	0.17	3	0.03	37	0.05	41
Special Conditions at Site	None	567	98.61	10231	98.18	70343	98.28	81141
	Auto traffic signal - out	2	0.35	14	0.13	160	0.22	176
	Auto signal part defective	0	0.00	4	0.04	37	0.05	41
	Road sign or marking defective	0	0.00	14	0.13	127	0.18	141
	Roadworks	4	0.70	79	0.76	568	0.79	651
	Road surface defective	0	0.00	27	0.26	113	0.16	140
	Oil or diesel (from 2005)	1	0.17	32	0.31	146	0.20	179
	Mud (from 2005)	1	0.17	20	0.19	80	0.11	101
Carriageway Hazards	None	569	98.96	10317	99.00	70932	99.10	81818
	Vehicle load on road	1	0.17	9	0.09	35	0.05	45
	Other object on road	3	0.52	42	0.40	250	0.35	295
	Previous accident	1	0.17	7	0.07	41	0.06	49
	Pedestrian in carriageway - not injured	1	0.17	23	0.22	199	0.28	223
	Any animal in carriageway (except ridden horse)	0	0.00	23	0.22	117	0.16	140

3.3.2 Vehicle Data

Vehicle data is the second database used in accident severity analysis that provide details of all vehicles involved in GB road accident . The purpose of using another type of data and repeating the analysis is to acquire the overall measure from different angle of perspective given the same event. To clarify, the ’Vehicle’ data does not have the information of accident severity. Therefore, ’Accident’ and ’Vehicle’ database is matched using accident index number to get a complete information. The vehicle database contains 252,836 vehicles that involved in road accidents, resulting with 216,824 slight injuries, 33,167 serious injuries and 2,845 fatal accident. Table 3.3, and 3.4 shows the summary statistics in ’Vehicle’ data.

Table 3.3: Summary Statistics for Vehicle Data

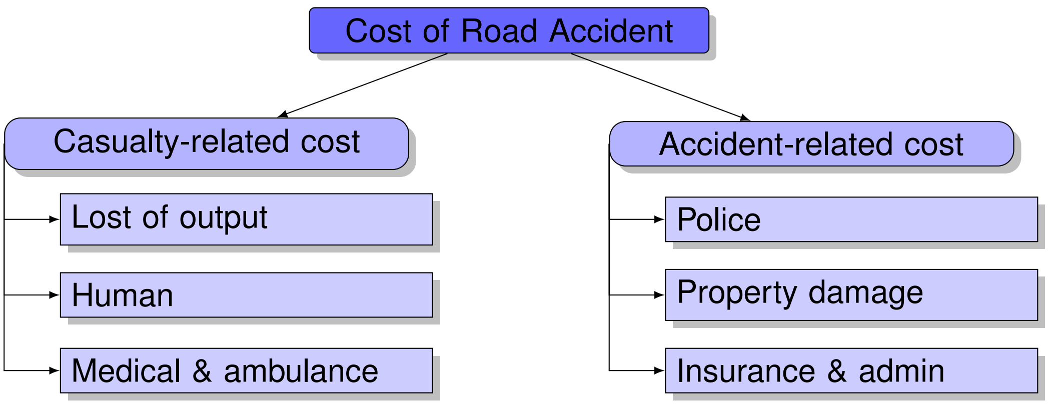
Variable		Fatal		Serious		Slight		Total
		n	%	*n*	%	*n*	%	
Type of Vehicle	Agricultural vehicle	24	0.84	123	0.37	411	0.19	558
	Bus/coach (17 or more seats)	69	2.43	698	2.10	5129	2.37	5896
	Car (from 2005)	1760	61.86	20249	61.04	158104	72.91	180113
	Electric m/cycle (from 2011)	0	0.00	2	0.01	5	0.00	7
	Goods 7.5 tonnes >	225	7.91	760	2.29	3761	1.73	4746
	Goods >3.5t. and <7.5t	45	1.58	247	0.74	1486	0.69	1778
	Goods vehicle - unknown weight (self rep only)	1	0.04	11	0.03	94	0.04	106
	M/cycle over 50 and up to 125cc (from 1999)	55	1.93	1594	4.81	5944	2.74	7593
	Minibus (8 - 16 seats)	12	0.42	98	0.30	485	0.22	595
	Mobility scooter (from 2011)	5	0.18	26	0.08	125	0.06	156
	M/cycle - unknown cc (from 2011)	0	0.00	17	0.05	38	0.02	55
	M/cycle 50cc and under	3	0.11	468	1.41	2061	0.95	2532
	M/cycle over 125cc and up to 500cc (from 2005)	33	1.16	605	1.82	1578	0.73	2216
	M/cycle over 500cc (from 2005)	265	9.31	2443	7.36	4427	2.04	7135
	Other vehicle	34	1.20	254	0.77	1168	0.54	1456
	Pedal cycle	121	4.25	3350	10.10	16578	7.64	20049
	Ridden horse (from 1999)	2	0.07	30	0.09	83	0.04	115
	Taxi / Private hire car (from 2005)	38	1.34	645	1.94	4378	2.02	5061
	Tram (from 1999)	0	0.00	2	0.01	18	0.01	20
	Van / Goods 3.5 tonnes mgw or under	153	5.38	1551	4.68	10982	5.06	12686
Towing and Articulation	Articulated vehicle	129	0.95	410	1.24	1954	0.90	2493
	Caravan	8	0.35	29	0.09	104	0.05	141
	Double or multiple trailer	0	0.56	6	0.02	36	0.02	42
	No tow/articulation	2679	9.59	32530	98.06	213988	98.67	249197
	Other tow	3	9.28	34	0.10	154	0.07	191
	Single trailer	27	1.19	165	0.50	642	0.30	834
Vehicle Manoeuvre	Changing lane to left	10	0.35	225	0.68	2035	0.94	2270
	Changing lane to right	16	0.56	264	0.80	2307	1.06	2587
	Going ahead left-hand bend	273	9.60	1650	4.97	6805	3.14	8728
	Going ahead other	1645	57.82	17178	51.78	99809	46.03	118632
	Going ahead right-hand bend	264	9.28	1942	5.85	7969	3.67	10175
	Moving off	34	1.20	1016	3.06	9317	4.30	10367
	Overtaking - nearside	13	0.46	228	0.69	1437	0.66	1678
	Overtaking moving vehicle - offside	103	3.62	998	3.01	3979	1.83	5080
	Overtaking static vehicle - offside	26	0.91	532	1.60	2685	1.24	3243
	Parked	162	5.69	1396	4.21	8124	3.75	9682
	Reversing	15	0.53	397	1.20	3206	1.48	3618
	Slowing or stopping	38	1.34	1291	3.89	18339	8.46	19668
	Turning left	21	0.74	938	2.83	7796	3.60	8755
	Turning right	150	5.27	3615	10.90	21857	10.08	25622
	U-turn	10	0.35	286	0.86	1609	0.74	1905
	Waiting to go - held up	47	1.65	809	2.44	15055	6.94	15911
	Waiting to turn left	1	0.04	58	0.17	1179	0.54	1238
	Waiting to turn right	17	0.60	349	1.05	3345	1.54	3711

Table 3.4: Summary Statistics for Vehicle Data (Continued)

Variable		Fatal		Serious		Slight		Total
		n	%	*n*	%	*n*	%	
Vehicle Location at Restricted Lane	Bus lane (from 1999)	5	0.18	121	0.36	795	0.37	921
	Busway (including guided busway) (from 1999)	0	0.00	7	0.02	75	0.03	82
	Cycle lane (on main carriageway) (from 1999)	3	0.11	88	0.27	616	0.28	707
	Cycleway or shared use footway	8	0.28	43	0.13	230	0.11	281
	Entering lay-by or hard shoulder	5	0.18	18	0.05	113	0.05	136
	Footway (pavement) (from 1999)	32	1.12	360	1.09	1908	0.88	2300
	Leaving lay-by or hard shoulder	2	0.07	43	0.13	267	0.12	312
	On lay-by or hard shoulder	23	0.81	82	0.25	472	0.22	577
	On main c'way - not in restricted lane (from 1999)	2767	97.22	32405	97.68	212384	97.92	247556
	Tram/Light rail track (from 1999)	1	0.04	7	0.02	31	0.01	39
Junction Location	Approaching junct or waiting/parked at junct apprch	390	13.70	6056	18.26	50927	23.48	57373
	Cleared junction or waiting/parked at junction exit	154	5.41	1891	5.70	11706	5.40	13751
	Entering from slip road	6	0.21	83	0.25	864	0.40	953
	Entering main road	77	2.71	1378	4.15	9309	4.29	10764
	Entering roundabout	11	0.39	613	1.85	5944	2.74	6568
	Leaving main road	44	1.55	983	2.96	5316	2.45	6343
	Leaving roundabout	12	0.42	395	1.19	3075	1.42	3482
	Mid Junction - on roundabout or on main road	335	11.77	7043	21.23	49966	23.04	57344
	Not at, or within 20 metres of, junction	1817	63.84	14727	44.40	79752	36.78	96296
Skidding_and_Overturning	Jackknifed	2	0.07	23	0.07	72	0.03	97
	Jackknifed and overturned	8	0.28	12	0.04	47	0.02	67
	None	2224	78.14	28061	84.59	194373	89.62	224658
	Overturned	102	3.58	763	2.30	2900	1.34	3765
	Skidded	409	14.37	3412	10.29	15902	7.33	19723
	Skidded and overturned	101	3.55	903	2.72	3592	1.66	4596
Hit Object in Carriageway	Any animal (except ridden horse) (from 2005)	7	0.25	76	0.23	250	0.12	333
	Bollard or refuge	14	0.49	159	0.48	913	0.42	1086
	Bridge (roof)	0	0.00	4	0.01	17	0.01	21
	Bridge (side)	4	0.14	25	0.08	115	0.05	144
	Central island of roundabout	4	0.14	50	0.15	209	0.10	263
	Kerb	94	3.30	671	2.02	2989	1.38	3754
	None	2650	93.11	31600	95.26	208637	96.19	242887
	Open door of vehicle	1	0.04	62	0.19	428	0.20	491
	Other object	13	0.46	77	0.23	456	0.21	546
	Parked vehicle	53	1.86	408	1.23	2746	1.27	3207
	Previous accident	4	0.14	24	0.07	70	0.03	98
	Road works	2	0.07	18	0.05	62	0.03	82
Vehicle Leaving Carriageway	Did not leave carriageway	1991	69.96	27836	83.91	195340	90.06	225167
	Nearside	452	15.88	2768	8.34	11343	5.23	14563
	Nearside and rebounded	39	1.37	364	1.10	1458	0.67	1861
	Offside	254	8.92	1505	4.54	5501	2.54	7260
	Offside - crossed central reservation	13	0.46	48	0.14	154	0.07	215
Vehicle Leaving Carriageway	Offside and rebounded	32	1.12	216	0.65	742	0.34	990
	Offside on to central reservation	23	0.81	146	0.44	973	0.45	1142
	Offside on to centrl res + rebounded	16	0.56	99	0.30	711	0.33	826
	Straight ahead at junction	26	0.91	192	0.58	668	0.31	886
Hit Object off C/way	Bus stop or bus shelter	1	0.04	19	0.06	117	0.05	137
	Central crash barrier	31	1.09	229	0.69	1581	0.73	1841
	Entered ditch	49	1.72	373	1.12	1596	0.74	2018
	Lamp post	33	1.16	266	0.80	1327	0.61	1626
	Near/Offside crash barrier	52	1.83	241	0.73	1414	0.65	1707
	None	2234	78.50	29364	88.52	200720	92.54	232318
	Other permanent object	148	5.20	1060	3.20	4179	1.93	5387
	Road sign or traffic signal	49	1.72	317	0.96	1549	0.71	1915
	Submerged in water	6	0.21	4	0.01	14	0.01	24
	Telegraph or electricity pole	20	0.70	117	0.35	515	0.24	652
	Tree	170	5.97	788	2.38	2224	1.03	3182
	Wall or fence (from 2011)	53	1.86	396	1.19	1655	0.76	2104
First Point of Impact	Back	229	8.05	3293	9.93	40752	18.79	44274
	Did not impact	235	8.26	2710	8.17	13403	6.18	16348
	Front	1663	58.43	17341	52.29	103485	47.72	122489
	Nearside	307	10.79	4637	13.98	28505	13.15	33449
	Offside	412	14.48	5180	15.62	30696	14.16	36288
Sex of Driver	Female	508	17.85	7618	22.96	64221	29.61	72347
	Male	2245	78.88	23689	71.41	139049	64.11	164983
	Not known	93	3.27	1867	5.63	13623	6.28	15583
Hit and Run	Hit and run	58	2.04	1488	4.49	14279	6.58	15825
	Non-stop vehicle not hit	24	0.84	428	1.29	2879	1.33	3331
	Other	2764	97.12	31258	94.22	199730	92.09	233752
Was Vehicle Left Hand Drive	No	2828	99.37	33026	99.56	215831	99.51	251685
	Yes	18	0.63	147	0.44	1054	0.49	1219
Journey Purpose	Commuting to/from work	246	8.64	3390	10.22	21263	9.80	24899
	Journey as part of work	585	20.56	5348	16.12	36406	16.79	42339
	Not known (from 2011)	1983	69.68	23920	72.11	155018	71.47	180921
	Other (from 2011)	19	0.67	175	0.53	1068	0.49	1262
	Pupil riding to/from school	2	0.07	87	0.26	803	0.37	892
	Taking pupil to/from school	11	0.39	251	0.76	2334	1.08	2596

3.4 Accident Cost Data

The second part of this thesis will discuss in depth about the accident cost data and the methodology used to model accident cost. Willingness-to-Pay (WTP) approach is used by GB since 1993 for the accident cost valuation [15]. The WTP method basically estimates the economic value of lost quality of life (VOSL), known as human cost, resulting from an accident event. The method simply valuate the amount of money an individual or the society willing to pay for accident risk reduction [66]. The components in WTP encompasses casualty-related cost, (lost output, medical & ambulance and human cost) and accident-related cost (police, property damage, insurance and administration cost). All types of cost is briefly discuss as below:



i. Lost of output cost

   Lost output refers to the production loss as a result from road accident. It means that expected loss of capacities due to fatality, disability or sick leave that inhibits an accident victim to contribute to the economy with its future earnings. Lost output cost calculation is studied in detail by [52] in 1993 and final valuation done by a group of researcher [10] in 1997. Since then, that value has been used as a benchmark then updating it over time based on inflation rate for that corresponding year. This cost is measured by the average earnings multiplied by the number of working years lost, activity rates and life expectancy, and then weighted according to current GDP per head ([52], [7]).

ii. Medical and ambulance cost

Medical costs is the costs incurred for medical treatment resulting from road casualties. Below are list down of cost items ([78]):

(a) First aid at the accident location and the transportation to hospital
(b) Medical treatment at the emergency department of hospital
(c) In-patient hospital treatment
(d) Out-patient hospital treatment
(e) Other treatment such as rehabilitation, physiotherapy, home care and carer service.
(f) Aids and appliances such as wheelchair, crutches, hearing aid, etc.

iii. Human cost

This cost constitutes the biggest portion in accident cost component. This intangible cost reflects the pain, grief, suffering and loss quality of life. Human cost cannot be estimated based on market price nor market transaction ([66]). Therefore, there are many ways to calculate human cost, but the most recommended method is Willingness To Pay approach ([78], [4]). Basically, WTP is a method to estimate the economic value that an individual (individual WTP) or the society (social WTP) is willing to pay to avoid certain accident risk. WTP method reflect the economic welfare theory, which means welfare is determined by individual preferences based on their consumption, and including factors that affect their loss quality of life. Value of a statistical life

(VOSL) also can be derived from WTP calculation. Human cost is describes as below ([78]):

(a) Human costs of fatalities (lost life years - this cost also known as VOSL)

(b) Human costs of injuries (loss of quality of life)

(c) Human costs for relatives and friend

iv. Police cost

Police costs are usually low when compared to other cost components. Police costs relate to the cost of time spent attending and reporting accidents by police officers. 51 police forces in England, Wales and Scotland are interviewed to estimate this cost in 2009. Since then, the national police cost is updated annually according to the market price. The cost estimation takes account the number and rank of police officer involved and time spent during the accident event. The estimated value will then convert into monetary unit to find the cost for each police force region for each type of accident severity in year 2009 ([16]).

v. Property damage cost

Damage to property includes both damages to vehicles and other third party property. The largest proportion of property damage comes from vehicle damage. Other property damage related cost is rather small, like infrastructure, fixed roadside objects and buildings, freight carried by lorries and personal property ([78]). In GB, property damage costs is estimated based on accident location and severity level from a survey of insurance claims in the early 1990s ([71]). Similar like other cost above, the national cost value is adjusted annually depends on the current market price.

vi. Insurance and admin cost

Insurance cost is the administrative costs of insurance company for the vehicle insurance. This cost is estimated by calculating the average handling cost per claim. It consists of the cost of personnel processing claims, overhead costs and allowance for expenses. The calculation of insurance cost was done in 1995 using insurance data, and it is adjusted annually till today. ([16]).

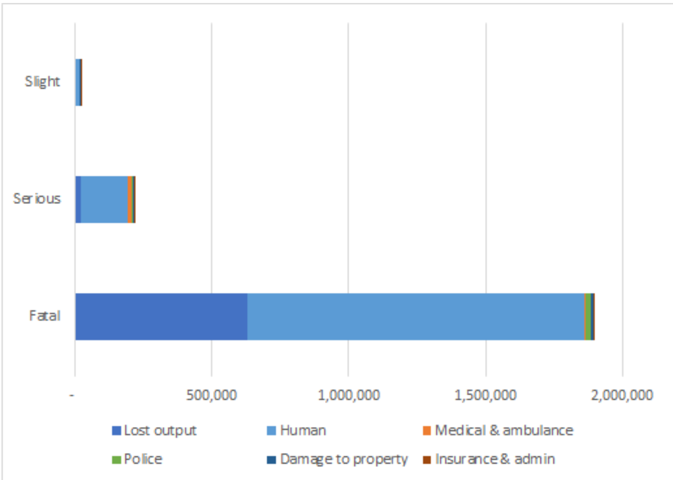


Figure 3.1. Comparison of Cost Components by Accident Severity [16]

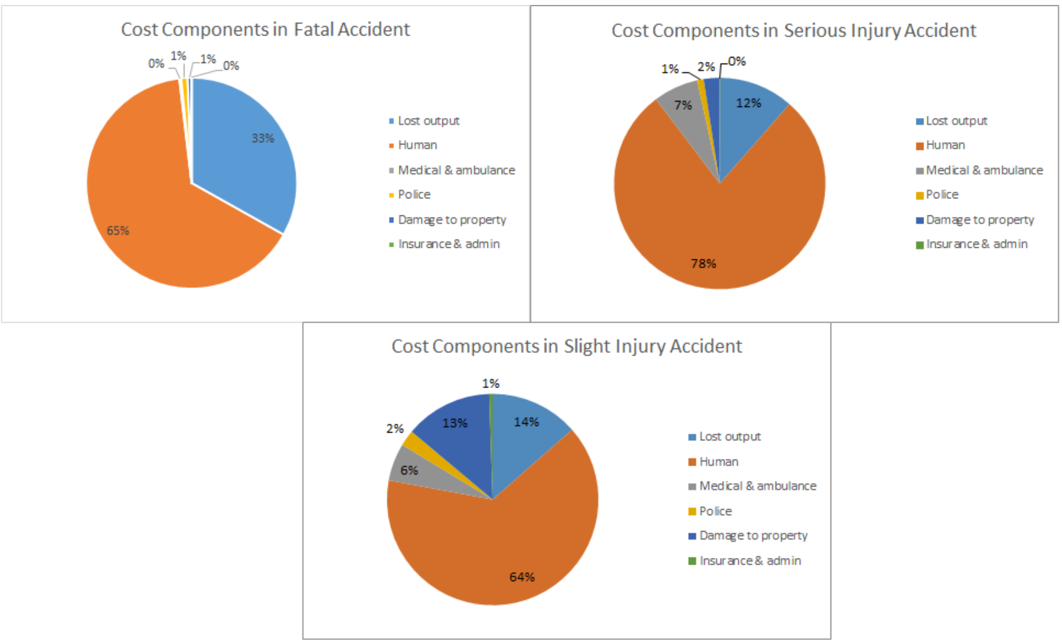


Figure 3.2. Percentage of Cost Components by Accident Severity [16]

There are many methods of accident cost valuation discussed by scholars around the world. Each method is chosen based on country's objective. In GB, there are three valuation methods used by DfT for accident costing ([77]):

i. The Restitution Costs (RC) approach

   This method estimates the amount of cost needed in order to restore the accident victim and their family back to the prior situation if they had not been involved in a road crash. It comprises of direct cost related with road accidents such as medical treatment, property damage, police cost and administrative cost. Direct cost is usually common in RC method because this cost is crucial in restoring the consequences of road accident. Market price is used to value these costs.

ii. Human capital (HC) approach

   This approach is suitable for estimation of loss of output cost. There are two types of lost of output cost, actual and potential loss. Actual loss indicates current productive capacity loss of a worker, while potential loss refers to the future productive capacity loss of unemployed to the person and their family due to road accident. Future loss is evaluated by using social discount rate.

iii. Willingness to pay (WTP) approach

   This method is the most recommended to calculate human cost ([4]) since there is no market price can represent human cost. WTP is an approach to estimate amount of cost of individual or society willing to pay for accident risk reduction. WTP has two methods, which are Stated Preference (SP) and Revealed Preference (RP) method. SP method uses questionnaires to ask people how much they are willing to pay for accident risk reduction. RP method calculate risk based on actual behaviour, such as buying airbags and car seat.

Below is the summary of the cost components and the method used to value it:

Table 3.5: Summary of Accident Cost Method

Cost Component	Method
Lost output	Human Capital
Medical & ambulance	Restitution Cost
Human	Willingness-To-Pay
Police	Restitution Cost
Property damage	Restitution Cost
Insurance & admin	Restitution Cost

Below is the summary of literature review for each cost component:

Cost Component	Literature Review
Lost of output	O'Reilly (1992) [52]; Carthy et al. (1998) [10]
Medical & ambulance	Department for Transport (2013) [16]
Human	Alfaro et al. (1994) [4]; Carthy et al. (1998) [10]
Police	Department for Transport (2013) [16]
Property damage	Simpson & O'Reilly (1994) [71]; Department for Transport (2013) [16]
Insurance & admin	Department for Transport (2013) [16]; Simpson & O'Reilly (1994) [71]

Accident cost data in GB is officially compiled in the DfT records in the WebTAG databook which can be assessed online. According to DfT personnel, the survey data that has been collected in 1990s by accident researcher is used until today and the value is adjusted from time to time based on current market price ([10], [15]). Raw survey data is not available for public use although formal request has been applied to Dft supported by university approval. The other alternative is by acquiring insurance claim data and use it as a proxy as accident cost. Unfortunately, all of the insurance companies did not give cooperation at all cost as all of their data were private and confidential. Questionnaire survey will involve variety of ethical issues as this topic is consider as sensitive issue among citizens.

Therefore, accident cost data from DfT record is considered to use in this study after much efforts for data request has failed. The accident cost records contain the values of accident cost occurred for the year 2013. However , the data obtained is in average form, hence the alternative way is to simulate data for accident cost by applying some assumptions to the data obtained. The database will be used as proxy to represent the actual cost involved. Below is the summary of accident cost data obtained from DfT as a main reference:

Table 3.6: Reported Accident Cases by Severity for Year 2012 [16]

Types of severity	Number of reported cases
Fatal	1,608
Serious	19,624
Slight	117,428
Total	**138,660**

Table 3.7: Accident Cost Data

Accident severity	Casualty related costs			Accident related cost			Total
	Lost output	Human	Medical & ambulance	Police cost	Damage to property	Insurance & admin	
Fatal	628,215	1,232,773	5,342	18,436	11,302	310	1,896,377
Serious	24,867	169,544	14,935	2,149	5,103	193	216,791
Slight	3,061	14,586	1,299	556	3,029	117	22,648
All injury	14,903	57,161	3,715	1,063	3,495	133	80,469
Damage only	-	-	-	36	1,930	56	2,022

For the above data, a simulation is done using R-software to find accident cost data. Each type of accident cost data is simulate based on number of accident occurrence. For example, fatal accident has 1,608 (refer Table 3.6) cases and the total output cost for each fatal accident is £628,215 (Table 3.7). For example, The simulation of lost output cost of fatal accident is set based on the boundary value. Assumption made that minimum value for lost of output cost for fatal accident will be equal to lost of output cost for serious injury accident which is £24,867 (refer Table 3.7). Therefore, a simulation between £24,867 to £628,215 is done for 1,608 data. The simulated data is checked by using scatter plot to ensure that it followed the three distributions.

Each cost component in fatal accident is simulate by applying the same method for three distribution - Gamma, Weibull and Log-normal. Then, a simulated database is prepared for three set of data for each distribution. It comprises 1,608 cases with six cost components each which generated a 1608×6 table. Accident cost for each case will be sum up to find the total cost will then be used further in modeling accident cost. The same process is repeated for other accident severity - serious and slight. Simulated accident cost data for serious injury has 19,624 cases with 6 cost components (19,624×6) while slight accident has 117,428 cases with 6 cost components (117,428×6).

3.5 Road Accidents Severity Model

Over the years, researchers have used a various range of methodological technique that have modelled the crash frequency and severity either independently or simultaneously. It is noteworthy to know that different countermeasures require to reduce crash frequency and crash severity [64]. The methods that selected in this study relied on the nature of the dependent variable, which is the severity of traffic accident. Injury severity data are classified by discrete categories such fatal, serious and slight injury. The responses data are in categorical form, ranging from less severe to more severe. Hence, discrete choice models (e.g., multinomial, ordinal or probit models) are selected. Models such as multinomial, nested and mixed logit models applied nominal response probability method. Furthermore, these models offer more flexible functional form and provide better parameter estimates [41].

As discussed previously in the literature review part, numerous research have been conducted to model the accident frequency and severity either independently or simultaneously. These are the statistical models that will be considered to apply in modelling road crashes severity:

i. Multinomial logistic regression
ii. Hierarchical log-linear graphical models
iii. Multinomial logistic random parameters model
iv. Bayesian inference - consensus Metropolis-Hastings algorithm

3.6 Multinomial Logistic Regression (MNL)

The methodological framework in this study aimed at identifying variables which may contribute to explaining accident severity and any variables that explain the unobserved random effects. Several researchers applied ordered discrete model to reflect the increasing level of injury severity like the study done by [32], [1], [2], [50], [18], [60], [61] and [22]. However, the degree of injury severity is not a major concern as less restriction model is required to obtain better prediction results. Therefore, MNL model is preferred over ordered models. This model is adopted to assess accident severity given that an accident has occurred. MNL is categorical outcome models with three or more nominal outcomes and will be developed to investigate the significance of chosen covariates towards each severity. This model has been utilized by many researchers previously to model accident severity ([79]; [31]; [81]; [30]; [28]; [65]; [41]; [63]; [42]; [64]; [68]; [2]; [9]). Essentially, in the MNL model formulation, a linear function of polytomous covariates that determine accident n's severity outcome i is given by,

$$S_{ij} = \alpha_i + \beta_i X_j, \tag{3.1}$$

where i is a set of all discrete outcomes (accident severity). S_{ij} stands for linear function determining the probability of severity outcome i for accident j, X_j denotes the exogenous measurable characteristics, e.g., road class, road type and speed limit for accident j with accident severity i, α_i is the intercept of accident severity i, β_i represents the vector of estimable parameters for accident severity i. The parameters in the model are estimated by using standard maximum likelihood approach. Data preparation and data analysis were performed using R software.

3.7 Log-linear Graphical Model

Graphical model is the second step of explorative study since it provides the conditional restrictions among variables and improves the high correlations issue arises in MNL model. It has become one of important tools in statistical modelling as an exploratory multivariate method. It is used for identifying direct and indirect association among random variables and able to break a large complex model into a simplified one. Furthermore, it provides illustrative presentation with some terminology for interpreting models. The result can be observed straightforward from the graph.

Graphical models are commonly used to explain conditional independence, where all variables are treated as response variables. UK road accident database contains large number of variables that may associate with road crashes. An important problem is then to identify the association, or in a complementary way, the conditional independence relationships between the variables under study [74]. The conditional independence is discussed as below:

Let X, Y be random variable. X and Y are independent if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y)$$

or equivalently, if

$$f_{Y|X}(y|x) = f_Y(y)$$

Let X, Y, Z be random variables, X and Y are conditionally independent given Z if for each value of z of Z, X and Y are independent in the conditional distribution given $Z = z$. That is if

$$f_{X,Y|Z}(x, y|z) = f_{X|Z}(x|z) f_{Y|Z}(y|z)$$

or equivalently

$$f_{Y|X,Z}(y|x, z) = f_{Y|Z}(y|z)$$

It also can be written in this form:

$$X \perp\!\!\!\perp Y|Z$$

Graphical models have been developed for categorical variable as a subclass of hierarchical log-linear models ([12],[35],[34]). In this study, all log-linear models are assumed to be hierarchical. For example, considering only three variables, x, y and z for simplicity, the saturated hierarchical log-linear model can be parameterised as:

$$\log p_{xyz} = u + u_{1(x)} + u_{2(y)} + u_{3(z)} + u_{12(xy)} + u_{13(xz)} + u_{23(yz)} + u_{123(xyz)} \tag{3.2}$$

The mutual independence model, that is the model which specifies that all variables are independent is,

$$\log p_{xyz} = u + u_{1(x)} + u_{2(y)} + u_{3(z)} \tag{3.3}$$

And the all two-way interaction model is

$$\log p_{xyz} = u + u_{1(x)} + u_{2(y)} + u_{3(z)} + u_{12(xy)} + u_{13(yz)} + u_{23(xz)} \tag{3.4}$$

The interaction graph of a graphical model for categorical variables is equivalent to the conditional independence graph, which is the main tool of graphical modelling [76]. The conditional independence graph is an undirected graph $G = (V, E)$ where it is a structure consisting of finite set V of vertices $V = \{1, 2, \ldots , d\}$ is corresponding to the set of variables in this study; and a finite set of E of edges between these vertices. It is called undirected graph since all the edges are undirected.

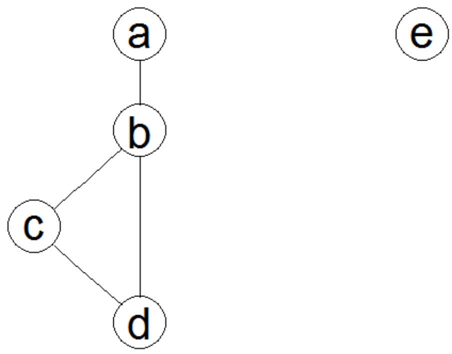


Figure 3.3. Example of Undirected Graph

Referring to Figure 3.1 above, it is an example of undirected graph, $G = (V, E)$, where $V = \{a, b, c, d, e\}$ and $E = \{ab, bcd, e\}$

3.8 Multinomial Logistic with Random Effects model (MNLRE)

In road accident, data commonly in clusters form and the suitable method to model this kind of data is by including random parameters in the MNL model [23]. One of the limitations of the traditional MNL models is its inability to accommodate any random effects and assumed that each observation is independent. Moreover, correlation is not well considered in the model.

Fixed effects is a parameter term used to explain factor's effect in a generalized linear model (GLM). It applies to all categories of interest, such as road type, light and weather conditions, type of first road class, etc. To overcome these limitations, in our model extension, we include random effects in the MNL model specification.

Random effects usually apply to a sample; however, in this case, the road accident model is observed to have spatial random effects on the accident-prone location for each severity outcome i. Generalized linear mixed model (GLMM) is an updated regression model that allows both random and fixed effects on the same scale in the linear estimation.

Location cluster j has T_j categorical observations. Let $\boldsymbol{X}_{jk}$ denote a column vector of independent variables for the kth observation in location j. For cluster j, u_j denote the location random effects. The linear part of the generalized linear mixed model (GLMM), for severity type i is:

$$S_{ijk} = \alpha_i + \beta_i X_{jk} + u_j \quad , \quad j = 1, 2, \ldots, J \tag{3.5}$$

where u_j are assumed to follow a multivariate normal distribution with $N(0, \Sigma)$, where Σ is an arbitrary covariance matrix, which allow more flexibility to the behaviour of the random effects.

3.9 Bayesian Inference

Bayesian inference is a method to find unknown parameters by using probability model conditionally on known data. The probability model is denote by $p(\theta|y)$, conditional probability density function where y is the known data and parameter θ is assume to be random and unknown. In Bayesian analysis, the likelihood function $L(\theta|y) = p(y|\theta)$, represent as the information carried by the observed data, combined with the prior distribution $p(\theta)$, are summarized in the posterior distribution:

$$p(\theta \mid y) = \frac{p(\theta)p(y \mid \theta)}{p(y)} \tag{3.6}$$

where $p(y)$ is the normalising constant $\int_\theta p(\theta)p(y \mid \theta)d\theta$. The posterior distribution can be represented as

$$p(\theta \mid y) \propto p(\theta)p(y \mid \theta) \tag{3.7}$$

which equal to $posterior \propto prior \times likelihood$. The posterior $p(\theta \mid y)$ is the probability distribution that holds the key solution to an inference problem of unknown parameters. The prior distribution is the early subjective assumption for unknown parameters. The prior distribution assumption can be made up by referring to previous analysis, theoritical reasoning or subjective judgement.

Many multivariate probability distributions are complex and difficult to sample. Markov Chain Monte Carlo (MCMC) is a method of simulating from such a distribution to generate a Markov chain sequence. MCMC algorithm constructs a sequence of random dependent variables θ_i from a normalised density $f(\theta) = g(\theta)/\int g(\theta)d\theta$, initiating from θ_0, where the next state θ_{i+1} depends only on the current state θ_i and not on the previous sequence θ_{i-1}.

The Markov chain sequence of θ_i converges to the stationary distribution if the specified conditions are met, and i reach a certain level to allow the chain converge.

3.9.1 Metropolis Hastings Algorithm

When direct sampling is complicated, Metropolis-Hastings (M-H) algorithm is used to estimate the MNLRE model's unknown parameters. M-H is a Markov

Chain Monte Carlo (MCMC) method where an arbitrary proposal probability distribution is used as a sample to generate a stationary distributed random samples after it has converged. Suppose that the likelihood function is

$$L(\theta \mid X, Y, u) = \prod_{j=1}^{J} \prod_{k=1}^{T_j} \prod_{i=1}^{I-1} \left\{ \frac{\exp\left(\alpha_i + \beta_i^T X_{jk} + u_j\right)}{1 + \sum_{i=1}^{I-1} \exp(\alpha_i + \beta_i^T X_{jk} + u_j)} \right\}^{I[Y_{jk}=i]} \tag{3.8}$$

where $u_i \sim N(0, \Sigma_u)$. The unknown parameters are $\theta_i = (\alpha_i, \beta_i)$. The chosen prior for α_i and β_i is assumed independent and normally distributed. These priors are designed to be fairly uninformative, described as below:

$$\alpha_i \sim N(0,1) \quad , \quad \beta_i \sim N(\mu_{\beta_i}, \Sigma_{\beta_i}), \tag{3.9}$$

for some known hyper-parameters μ_{β_i} and Σ_{β_i}. Define $\mu_\beta = (\mu_{\beta_1}, \mu_{\beta_2}, \ldots, \mu_{\beta_{I-1}})'$, $\Sigma_{\beta_i} = \begin{bmatrix} \Sigma_{\beta_1} & 0 & \cdots & 0 \\ 0 & \Sigma_{\beta_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_{\beta_{I-1}} \end{bmatrix}$ then the posterior density kernel can simply written as *posterior* $\propto$ *prior* $\times$ *likelihood* which equal to:

$$\begin{aligned} \pi(\theta \mid Y, X, u) &\propto \pi_0(\theta) L(\theta \mid X, Y, u) \\ &\propto \left[\prod_{i=1}^{I-1} N(\alpha_i; 0, 1)\right] N(\beta; \mu_\beta, \Sigma_\beta) L(\theta \mid X, Y, u) \end{aligned}$$

Let the proposal density for $q(\alpha_i | \alpha_i^{'})$ to be a normal distribution with mean, μ_{α_i} and standard deviation, σ_{α_i}, where μ_{α_i} is drawn from the perturbation chain generated in the previous chain step, while σ_{α_i} is obtained from the previous chain iteration. A proposal value α_i^* is then generated from $q(\alpha_i | \alpha_i^{'})$. The chain moves from α_i to α_i^* with probability:

$$\varphi\left(\alpha_i, \alpha_i^* \mid X_{jk}, Y_{ijk}, \beta_i, u_j\right) = \min\left\{ \frac{\pi(\alpha_i^*, \beta_i \mid Y_{ijk}, X_{jk}, u_j)}{\pi(\alpha_i, \beta_i \mid Y_{ijk}, X_{jk}, u_j)}, 1 \right\}, \tag{3.10}$$

where φ represents an M-H acceptance probability. If Eq. 3.10 is greater than U (where U is uniformly distributed on [0,1]), the proposal value α_i^* is accepted; otherwise, the current value α_i is kept as the new draws for the respective Markov chain.

Similar like the above case, the sample β_i is drawn from the posterior distribution $\pi(\theta_i \mid Y_{ijk}, X_{jk}, u_j)$ and the proposal value, β_i^* is obtained from the proposal distribution $q(\beta_i|\beta_i^{'})$, that follows a normal distribution with mean, μ_{β_i} generated from the chain value in the previous chain step, and standard deviation, σ_{β_i} is obtained from the previous iteration. The chain moves from β_i to β_i^* with probability:

$$\varphi\left(\beta_i, \beta_i^* \mid X_{jk}, Y_{ijk}, \alpha_i, u_j\right) = \min\left\{\frac{\pi(\beta_i^*, \alpha_i \mid Y_{ijk}, X_{jk}, u_j)}{\pi(\beta_i, \alpha_i \mid Y_{ijk}, X_{jk}, u_j)}, 1\right\}, \tag{3.11}$$

where φ represents an M-H acceptance probability. If Eq. 3.11 is greater than U (where U is uniformly distributed on [0,1]), the proposal value β_i^* is accepted; otherwise, the current value β_i will be applied for the next chain.

3.9.2 Consensus MCMC

The main constraint in performing the M-H simulation procedure for this dataset is the computational time and complexity in application. When the data is too large to process on a single machine, the solution is to divide the task among multiple machines by running a separate M-H algorithm on each machine, and then averaging individual draws across machines. This method is called the consensus MCMC introduced by [67].

In this method, the data are split among multiple databases into groups called shards. Each shard is assigned to a particular machine which run separate simulation on each machine from a posterior distribution given its own data without any inter-machine communication, i.e., each shard is independent of other database shards. The posterior simulations from each machine are then aggregated to produce a set of global average parameters representing the consensus of all the machines.

Let y denote the full data, y_s represent shard s for $s = 1, \ldots, S$ and θ be the global parameters. Assuming y to be exchangeable, the posterior distribution can be written as

$$p(\theta \mid y) = \prod_{s=1}^{S} p(y_s \mid \theta) p(\theta)^{\frac{1}{S}} \tag{3.12}$$

Note that the prior distribution $p(\theta) = \prod_s p(\theta)^{\frac{1}{S}}$ in Eq. 3.12 is divided into S components to preserve the total amount of prior information. This consensus system assumes each shard is conditionally independent between other shards.

Several methods have been proposed to aggregate the estimation results such as density estimation [49, 75], geometric median [47], and weighted averaging [67] that is used in this paper. The estimation results of each parameter $\theta_{s1}, \ldots, \theta_{sg}$ from all shards are combined and aggregated. Each shard is assigned a weight, which is a matrix W_s. Thus, the consensus posterior for g is

$$\theta_g = \left(\sum_s W_s \right)^{-1} \sum_s W_s \theta_{sg} \tag{3.13}$$

Note that since the posterior $p(\theta \mid y_s)$ is assumed to be Gaussian, then the joint posterior $p(\theta \mid y)$ is also Gaussian.

There are three different weighting schemes to combine consensus MCMC which are matrix, scalar, and equal schemes. According to [67], matrix and scalar weighting methods perform better when combining the estimation results from consensus MCMC algorithm in the application of logistic regression analysis. In this paper, the matrix weighting scheme is selected for the implementation of our model. Therefore, the appropriate weighted averages selected for W_s will be $W_s = \Sigma^{-1}$, where $\Sigma_s = CoVar(\theta \mid y_s)$. The Σ_s is sample covariance of $\theta_{s1}, \cdots, \theta_{sg}$ generated from consensus M-H simulation. Thus, referring to Eq. 3.13, the weighted average equation can be written as:

$$\theta_g = \left(\sum_{s=1}^{S} \Sigma_s \right)^{-1} \sum_{s=1}^{S} \Sigma_s^{-1} \theta_{sg} \tag{3.14}$$

3.10 Road Accidents Cost Model

Accident cost is the continuation study of road accidents severity model. After the contributing factors of each accident severity are known, it is worth to study in depth regarding how the accident cost influence each accident severity. Previous accident cost study showed all of the cost study used basic calculation with the application of cost economic model. Hence, a new approach of accident costing by using statistical model is proposed. As discussed previously, simulated data is used to model accident cost. Three statistical models are used to find the cost difference for each accident severity then compared with accident severity model result. There are three suitable distributions to estimate accident cost which are [27]:

i. Gamma distributions

ii. Weibull distributions

iii. Lognormal distributions

3.10.1 Gamma Distribution

Study of fitting accident cost ratio into Gamma distribution between fatal and non-fatal accident has been done to declare the marginal Gamma distribution and assumption made [27]. The model is described as by assuming that the accident cost for fatal accident C_1 and while accident cost for non-fatal accident C_2 are following Gamma distributions with parameters C_1~Gamma(γ_1, η_1) and C_2 ~Gamma(γ_2, η_2). Letting $Z = \frac{C_1}{C_2}$ to represent the cost ratio between fatal and non-fatal accident. Considering C_1 and C_2 are two independent random variables, thus, the marginal Gamma distribution obtained by applying transformation technique could be written as:

$$f_R(z) = \frac{\Gamma(\gamma_1 + \gamma_2)\eta^{\gamma_1}\eta^{\gamma_2}z^{\gamma_1 - 1}}{\Gamma(\gamma_1)\Gamma(\gamma_2)[z\eta_1 + \eta_2]^{\gamma_1+\gamma_2}} \quad , \quad z > 0 \tag{3.15}$$

Since there are three types of accident severity to estimate the cost, the above Equation 3.15 is not fit to apply in this study. Therefore, a normal Gamma model is used to find the estimated cost for each accident severity then comparison will be made.

The propose model is by assuming that the accident cost for fatal accident C_1, accident cost for serious injury accident C_2 and accident cost for slight injury accident C_3 are following Gamma distributions with parameters C_1~Gamma(γ_1, η_1), C_2 ~Gamma(γ_2, η_2) and C_3 ~Gamma(γ_3, η_3). Considering C_1, C_2 and C_3 are three independent random variables, thus, the probability distribution function of Gamma distribution for three accident cost could be written as:

$$f_{C_1}(x) = \frac{x^{\gamma_1 - 1} e^{-\frac{x}{\eta_1}}}{\Gamma(\gamma_1)\eta_1^{\gamma_1}} \quad , \quad x > 0; \qquad \gamma_1, \eta_1 > 0 \tag{3.16}$$

$$f_{C_2}(x) = \frac{x^{\gamma_2 - 1} e^{-\frac{x}{\eta_2}}}{\Gamma(\gamma_2)\eta_2^{\gamma_2}} \quad , \quad x > 0; \qquad \gamma_2, \eta_2 > 0 \tag{3.17}$$

$$f_{C_3}(x) = \frac{x^{\gamma_3 - 1} e^{-\frac{x}{\eta_3}}}{\Gamma(\gamma_3)\eta_3^{\gamma_3}} \quad , \quad x > 0; \qquad \gamma_3, \eta_3 > 0$$

Similar like in the previous part, the parameter estimation is done by using Bayesian inference, which is consensus MCMC with random walk M-H algorithm method. In this model, the unknown parameters are γ_i and η_i. The chosen prior for γ_i and η_i is assumed independent and normally distributed. The proposal density for $q(\gamma_i|\gamma_i^{'})$ and $q(\eta_i|\eta_i^{'})$ followed normal distribution with mean, μ_{γ_i} and μ_{η_i} is drawn from the perturbation chain generated in the previous chain step, meanwhile, standard deviation, σ_{γ_i} and σ_{η_i} is obtained from the previous chain iteration. The estimated parameter value from the simulation work will be used further in finding the accident cost for each level of severity.

3.10.2 Weibull Distribution

Weibull distribution is a very popular distribution and widely applied in modelling data in science field. Up to this day, this model has never been applied in road accident cost analysis before. Assuming that the cost for fatal accident (C_1) while cost for non-fatal accident (C_2) are following Weibull distributions with parameters C_1 ~Weibull (λ_1, k_1) and C_2 ~Weibull (λ_2, k_2). Considering new random variable with, $Z = \frac{C_1}{C_2}$, to represent the cost ratio between fatal and non-fatal accident. C_1 and

C_2 are two independent random variables, thus, probability distribution function of the ratio of two Weibull random variables is given by [8]:

$$f_R(z) = \frac{\theta\gamma(1-\rho)(1+\gamma z^{\frac{\theta}{a}})z^{\frac{\theta}{a}-1}}{c[1-2\gamma(2\rho-1)z^{\frac{\theta}{a}}+\gamma^2 z^{\frac{2\theta}{a}}]^{\frac{3}{2}}} \quad , \quad z > 0 \tag{3.18}$$

Where $Z = \frac{C_1}{C_2}, \theta = \frac{k_1}{k_2}, \gamma = \frac{\lambda_1}{\lambda_2} > 0$

The formula 3.19 above further simplify by letting $a = 1$ and $\rho = 0$ since C_1 and C_2 are two independent random variables.

$$f_R(z) = \frac{\theta\gamma(1+\gamma z^{\theta})z^{\theta-1}}{[1+2\gamma z^{\theta}+\gamma^2 z^{2\theta}]^{\frac{3}{2}}} \quad , \quad z > 0 \tag{3.19}$$

The above Equation 3.20 is only applicable for two levels of severity. Therefore, the original Weibull model is referred to find the accident cost for fatal, serious and slight accident. Assuming that the cost for fatal accident (C_1), cost for serious accident (C_2) and cost for slight accident (C_3) are following Weibull distributions with parameters C_1 ~Weibull (λ_1, k_1), C_2 ~Weibull (λ_2, k_2) and C_3 ~Weibull (λ_3, k_3). Considering C_1, C_2 and C_3 are all independent random variables, the probability distribution function of three Weibull models are given by:

$$f_{C_1}(x) = \frac{k_1}{\lambda_1}\left(\frac{x}{\lambda_1}\right)^{k_1-1} e^{-\left(\frac{x}{\lambda_1}\right)^{k_1}} \quad , \quad x \geq 0 \tag{3.20}$$

$$f_{C_2}(x) = \frac{k_2}{\lambda_2}\left(\frac{x}{\lambda_2}\right)^{k_2-1} e^{-\left(\frac{x}{\lambda_2}\right)^{k_2}} \quad , \quad x \geq 0 \tag{3.21}$$

$$f_{C_3}(x) = \frac{k_3}{\lambda_3}\left(\frac{x}{\lambda_3}\right)^{k_3-1} e^{-\left(\frac{x}{\lambda_3}\right)^{k_3}} \quad , \quad x \geq 0 \tag{3.22}$$

The unknown parameters for this models are λ_i and k_i. The parameter estimation is done by applying Bayesian technique using consensus MCMC random walk

M-H algorithm. The prior for λ_i and k_i are assumed independent and normally distributed. The proposal density for $q(\lambda_i|\lambda_i^{'})$ and $q(k_i|k_i^{'})$ followed normal distribution with mean, μ_{λ_i} and μ_{k_i} is drawn from the perturbation chain generated in the previous chain step, meanwhile, standard deviation, σ_{λ_i} and σ_{k_i} is obtained from the previous chain iteration. The estimated parameter value from the simulation work will be used further in finding the accident cost for each level of severity.

3.10.3 Log-Normal Distribution

Log-normal distribution is another suitable distribution to model accident cost [13]. There is no research nor literature found to discuss regarding the probability density function of the ratio of two Log-normal random variables. Therefore, the basic Log-normal model is used in this study.

Letting the fatal accident cost as C_1, serious accident cost as C_2) and slight accident cost as C_3). All these three cost variables are following Log-normal distributions with parameters C_1 ~Log-normal(μ_1, σ_1), C_2 ~Log-normal(μ_2, σ_2) and C_3 ~Log-normal(μ_3, σ_3). The probability distribution function of Log-normal distribution for C_1, C_2 and C_3 are given by:

$$f_{C_1}(x) = \frac{1}{x\sigma_1\sqrt{2\Pi}}\exp\left(-\frac{(\ln x - \mu_1)^2}{2\sigma_1^2}\right) \tag{3.23}$$

$$f_{C_2}(x) = \frac{1}{x\sigma_2\sqrt{2\Pi}}\exp\left(-\frac{(\ln x - \mu_2)^2}{2\sigma_2^2}\right) \tag{3.24}$$

$$f_{C_3}(x) = \frac{1}{x\sigma_3\sqrt{2\Pi}}\exp\left(-\frac{(\ln x - \mu_3)^2}{2\sigma_3^2}\right) \tag{3.25}$$

Where C_1, C_2 and C_3 are three independent random variables.

The unknown parameters for this models are μ_i and σ_i. Same like in the previous two models, parameter estimation is done by applying Bayesian technique using consensus MCMC random walk M-H algorithm. The prior for μ_i and σ_i are assumed to be independent and normally distributed. The proposal density for $q(\mu_i|\mu_i^{'})$ and $q(\sigma_i|\sigma_i^{'})$ followed normal distribution with mean, μ_{μ_i} and μ_{σ_i} is drawn from the perturbation chain generated in the previous chain step, meanwhile, standard deviation, σ_{μ_i} and σ_{σ_i} is obtained from the previous chain iteration. The estimated parameter value from the simulation work will be used further in finding the accident cost for each level of severity.

Chapter 4

Accident Severity Model

4.1 Introduction

There are two different approaches to model road accident severity which are by analysing through statistical frequency and also accident cost model. Separate models were used because this research aims to determine the contributing factors that associated with accident occurrence and the cost involved for each type of severity. This chapter focuses on finding the contributing accident factor to three types of accident severity - fatal, serious and slight.

In this study, the discussion is based on the result of three statistical models produced, which are multinomial logistic regression (MNL), log-linear graphical model and multinomial logistic regression with random effect (MNLRE). MNL and MNLRE are the same models with different approach and parameter estimation. MNL used all variables in the data source and the parameter estimation used is maximum likelihood estimation. Meanwhile, MNLRE is also used all variables, plus random effects location, and using Bayesian method as its parameter estimation. These three models will be explored and present in the next section. Accident cost model will be discussed in the next chapter.

4.2 Multinomial Logistic Regression

A multinomial logistic regression analysis was conducted to predict factors that contribute to the probability of accident severity given fatal, serious and slight. Two datasets are used to perform this analysis which is Accident and Vehicle, both for year 2013. Each dataset has different types of predictors and explanatory variables, hence all of the variables are treated as factors.

4.2.1 Accident Data

Results of the multinomial logit regression analysis was fitted using R software package and presented in Table 4.1. The R packages that used for this analysis are *mlogit*, *foreign*, *nnet* and *stargazer* from Comprehensive R Archive Network (CRAN) site. For this analysis, a total of 70 predictors are used and the nominal response outcomes are the severity of road crashes - fatal, serious and slight. For the response variable, fatal accident is selected as the reference group and set as zero, thus the interpretation results for other outcomes are depended on this base case (see Table 4.1). Significant variables with Wald statistics test greater than 1.96 corresponding to the level of significance 0.05 are discussed as below.

Using the previously described data, a MNL model that included all crashes with a wide variety of variables was developed. The calculated likelihood ratio (LR) value is less than level of significance 0.05, indicating that the model is statistically significant.

Table 4.1: Multinomial Logit Model Results for Accident Data

Variable	Factor	Estimated coefficient	Standard error	Wald-Statistic	*p*-value
Constant [1]		10.674	7.340	1.454	0.146
Constant [2]		12.736	7.318	1.740	0.082
First Road Class [1]	C [1]	0.426	0.180	2.359	0.018
	Unclassified [1]	0.417	0.121	3.436	0.001
First Road Class [2]	C [2]	0.515	0.177	2.910	0.004
	Unclassified [2]	0.357	0.119	2.994	0.003
Speed limit (mph) [1]	30 [1]	-1.195	0.515	-2.321	0.020
	40 [1]	-1.758	0.525	-3.347	0.001
	50 [1]	-1.804	0.504	-3.580	<0.001
	60 [1]	-2.124	0.563	-3.769	<0.001
Speed limit (mph) [2]	30 [2]	-1.209	0.51	-2.370	0.018
	40 [2]	-1.954	0.519	-3.762	<0.001
	50 [2]	-2.128	0.499	-4.261	<0.001
	60 [2]	-2.451	0.556	-4.407	<0.001
Road Type [2]	Single carriageway [2]	-0.407	0.173	-2.351	0.019
Junction detail [2]	Roundabout [2]	0.934	0.267	3.495	<0.001
Weather condition [1]	Fog or mist [1]	12.738	0.097	130.981	<0.001
	Raining + no high winds [1]	0.661	0.324	2.043	0.041
	Other	1.098	0.514	2.139	0.032
Weather condition [2]	Fog or mist [2]	12.662	0.097	130.204	<0.001
	Raining + no high winds [2]	0.842	0.313	2.692	0.007
	Snowing + no high winds [2]	2.661	1.34	1.986	0.047
	Raining + high winds [2]	0.929	0.468	.881	0.047
	Other [2]	1.124	0.503	2.236	0.025
Special conditions [1]	Road sign/marking defective [1]	7.079	0.319	22.158	<0.001
	Road surface defective [1]	10.185	0.285	35.683	<0.001
Special conditions [2]	Road sign/marking defective [2]	7.239	0.326	22.184	<0.001
	Road surface defective [2]	9.571	0.286	33.439	<0.001
Log likelihood with constants only		-34655			
Log likelihood at final		-34141.883			
Number of observations		82570			
p-value (from likelihood ratio test)		<0.0001			

* [1] serious injury, [2] slight injury. The fatal injury is the reference category and set to zero.

The result of significant variables for accident data is presented in Table 4.1. In terms of contributing factor identification, a wide variety of variables is found to significantly influence driver injury severity levels with a total of 20 coefficients estimated.

i. **First Road Class**

Among the six ′first road class′ groups, road crash is significant to occur at ′first road class type C′ and ′unclassified first road class′. The relative risk (fatal accident as base case) for serious and slight injury would be expected to increase by a factor of 1.53 and 1.67 for ′first road class type C′, while 1.52 and 1.43 for ′unclassified first road class′, given the other variables in the model held constant. Lower class of first road class increase the probability of serious and slight injury may be because of the poor road condition and less street light, road sign/markings. From, the analysis, fatal accident has lower likelihood in these road class compare to serious and slight injury.

ii. **Speed Limit**

Somewhat surprisingly, traffic crashes occurring on roads with lower speed limits led to increases in the road crash frequency and severity. Busy area commonly address with lower speed limit for safety precaution since it has high population of vulnerable road users such as kids, cyclists, older people, etc. This indicates that the posted speed limit is not adhere by the driver, hence more accidents occurred. MNL model results show similar risk relative value both for serious and slight injury, albeit the value is decreasing against the increasing speed limit. The risk relative for serious injury (fatal accident as reference case) is 0.303, 0.172, 0.165 and 0.120 for speed limit 30, 40, 50 and 60 mph respectively; while risk relative for slight injury is 0.298, 0.142, 0.119 and 0.086 for speed limit 30, 40, 50 and 60 mph.

iii. **Weather Conditions**

Weather conditions affected each severity in different way, notice that three variables are significant for serious injury while five variables for slight injury. 'Fog or mist' is substantially most likely to increase the relative risk of serious injury (fatal accident as reference group) by a factor of 340,441.982 and slight injury by a factor of 315,527.12. Under adverse weather, the limited visible distance and foggy windscreen make it more difficult to manoeuvre the vehicle. 'Raining with no high winds' affected serious and slight injury with the relative risk 1.937 and 2.321 if compare with fatal accident. 'Other' weather conditions also significant in both serious and slight injury with a relative risk of 2.998 and 3.077 respectively. Two other variables are not significant in serious injury but contributed for slight injury: 'snowing with no high winds' affected by relative risk of 14.311 and 'raining with high winds' with 2.532. This finding is likely because bad weather conditions reflect the reduced reaction times upon upcoming road hazards, hence causing more accidents occurrence.

iv. **Special Conditions at Site**

From MNL model results, 'road sign/marking' and 'road surface' conditions play an important role to explain accident severity. This can be observed by high relative risk result for both injury severity. For 'defective road sign/marking' , it would be expected to increase by a factor of 1,186.781 for serious injury and 1,392.70 for slight injury if compare with fatal group as reference, given the other variables in the model held constant. 'Defective road surface' also shows high value of relative risk (fatal accident as base group) with 26,502.65 for serious and 14,342.75 for slight accident. Road sign/marking worked as early indicator to the driver and other road user to be more alert on the surroundings. Defective road surface led to many accident since it affect vehicle movement.

v. **Junction Detail**

The relative risk for slight injury accident vs. fatal accident to occur, would be expected to increase by a factor of 2.545 for 'roundabout'; if moving from 'crossroads' to the mentioned category, given the other variables in the model held constant. It also means that slight injury accident more likely to happen at 'roundabout' compare to fatal accident.

4.2.2 Vehicle Data

Results of the MNL regression analysis is presented in Table 4.2, 4.3 and 4.4. For this analysis, a total of 122 predictors are used and the dependent variable is the severity of road crashes - fatal, serious and slight where fatal accident is selected as the reference group. Level of significance 0.05 is also used to find the significant variables that contribute to each accident severity. The likelihood ratio from this model is less than 0.05, indicating that this model is significant.

Table 4.2: Multinomial Logit Model Results for Vehicle Data

Variable	Factor	Estimated coefficient	Standard error	*p*-value
Constant [1]		7.318	0.330	<0.0001
Constant [2]		10.822	0.330	<0.0001
Age [1]		-0.005	0.001	0.001
Age [2]		-0.012	0.001	<0.0001
Vehicle Type [1]	Pedal cycle [1]	1.806	0.162	<0.0001
	M/cycle over 50 and up to 125 cc (from 1999) [1]	1.828	0.191	<0.0001
	Mobility scooter (from 2011) [1]	5.943	0.135	<0.0001
	M/cycle - unknown cc (from 2011) [1]	1.220	0.203	<0.0001
	M/cycle - 50 cc and under [1]	29.725	0.065	<0.0001
	M/cycle - over 125cc and up to 500 cc (from 2005) [1]	1.573	0.238	<0.0001
	M/cycle - over 500 cc (from 2005) [1]	0.691	0.138	<0.0001
	Taxi/private hire car (from 2005) [1]	1.026	0.226	<0.0001
	Car [1]	1.070	0.124	<0.0001
	Goods over 3.5 t and under 7.5 t (from1999) [1]	-0.500	0.187	0.007
	Goods vehicle - unknown weight [1]	-0.535	0.250	0.032
	Mini bus (8-16 passenger seats) (from 1999) [1]	1.918	0.178	<0.0001
	Bus or coach [1]	0.638	0.191	0.001
	Van/Goods 3.5 t mgw or under [1]	0.504	0.146	0.001
	Other vehicle [1]	0.624	0.261	
Vehicle Type [2]	Pedal cycle [2]	1.962	0.160	<0.0001
	M/cycle over 50 and up to 125 cc (from 1999) [2]	1.692	0.189	<0.0001
	Mobility scooter (from 2011) [2]	5.943	0.135	<0.0001
	M/cycle - 50 cc and under [2]	29.602	0.065	<0.0001
	M/cycle - over 125cc and up to 500 cc (from 2005) [2]	1.206	0.235	<0.0001
	Taxi/private hire car (from 2005) [2]	1.586	0.221	<0.0001
	Car [2]	1.705	0.121	<0.0001
	Mini bus (8-16 passenger seats) (from 1999) [2]	2.449	0.180	<0.0001
	Bus or coach [2]	1.350	0.186	<0.0001
	Van/Goods 3.5 t mgw or under [2]	1.103	0.142	<0.0001
	Other vehicle [2]	1.001	0.254	<0.0001
Towing and articulation [1]	Double or multiple trailer [1]	1.586	0.298	<0.0001
	Other tow [1]	1.636	0.111	<0.0001
Towing and articulation [2]	Double or multiple trailer [2]	2.409	0.304	<0.0001
	Caravan [2]	-1.188	0.467	0.011
	Other tow [2]	1.552	0.113	<0.0001
Vehicle Manoeuvre [1]	Reversing [1]	2.414	1.846	<0.0001
	Parked [1]	-1.939	0.159	<0.0001
	Waiting to go - held up [1]	10.793	0.045	<0.0001
	Slowing or stopping [1]	8.727	0.043	<0.0001
	Turning left [1]	1.514	0.047	<0.0001
	Waiting to turn left [1]	4.060	0.088	<0.0001
	Turning right [1]	-0.798	0.186	<0.0001
	Waiting to turn right [1]	-1.071	0.418	0.010
	Changing lane to right [1]	-0.941	0.329	0.004
	Overtaking moving vehicle - offside [1]	-2.458	0.133	<0.0001
	Overtaking static vehicle - offside [1]	-1.429	0.248	<0.0001
	Overtaking - nearside [1]	-2.018	0.275	<0.0001
	Going ahead left-hand bend [1]	-2.357	0.114	<0.0001
	Going ahead right-hand bend [1]	-2.137	0.114	<0.0001
	Going ahead other [1]	-2.108	0.093	<0.0001

* [1] serious injury, [2] slight injury. The fatal injury is the reference category and set to zero.

Table 4.3: Multinomial Logit Model Results for Vehicle Data (Continued)

Variable	Factor	Estimated coefficient	Standard error	p-value
Vehicle Manoeuvre [2]	Reversing [2]	1.846	0.051	<0.0001
	Parked [2]	-2.309	0.155	<0.0001
	Waiting to go - held up [2]	11.171	0.045	<0.0001
	Slowing or stopping [2]	8.946	0.043	<0.0001
	U-turn [2]	-1.136	0.451	0.012
	Waiting to turn left [2]	4.747	0.088	<0.0001
	Turning right [2]	-1.389	0.185	<0.0001
	Waiting to turn right [2]	-1.388	0.415	0.001
	Changing lane to right [2]	-1.061	0.324	0.001
	Overtaking moving vehicle - offside [2]	-3.024	0.130	<0.0001
	Overtaking static vehicle - offside [2]	-1.815	0.245	<0.0001
	Overtaking - nearside [2]	-2.401	0.271	<0.0001
	Going ahead left-hand bend [2]	-2.855	0.112	<0.0001
	Going ahead right-hand bend [2]	-2.674	0.112	<0.0001
	Going ahead other [2]	-2.523	0.092	<0.0001
Vehicle Location - Restricted Lane/Away from Main Carriageway [1]	Tram/Light rail track (from 1999) [1]	-0.956	0.313	0.002
	Cycle lane (on main carriageway) (from 1999) [1]	3.570	0.079	<0.0001
	Cycleway or shared use footway [1]	-1.884	0.374	<0.0001
	On lay-by or hard shoulder [1]	-1.079	0.302	<0.0001
	Leaving lay-by or hard shoulder [1]	-0.300	0.137	0.028
Vehicle Location - Restricted Lane/Away from Main Carriageway [2]	Tram/Light rail track (from 1999) [2]	-1.458	0.276	<0.0001
	Cycle lane (on main carriageway) (from 1999) [2]	3.399	0.079	<0.0001
	Cycleway or shared use footway [2]	-2.006	0.351	<0.0001
	On lay-by or hard shoulder [2]	-1.030	0.277	<0.0001
	Entering lay-by or hard shoulder [2]	-1.050	0.408	0.100
	Leaving lay-by or hard shoulder [2]	-0.550	0.137	<0.0001
Junction Location of Vehicle (from 2005) [1]	Not at, or within 20 metres of, junction [1]	-0.597	0.073	<0.0001
	Leaving roundabout [1]	1.804	0.039	<0.0001
	Entering roundabout [1]	5.925	0.025	<0.0001
	Mid Junction - on roundabout or on main road [1]	0.679	0.109	<0.0001
Junction Location of Vehicle (from 2005) [2]	Not at, or within 20 metres of, junction [2]	-0.904	0.071	<0.0001
	Leaving roundabout [2]	2.236	0.039	<0.0001
	Entering roundabout [2]	6.091	0.025	<0.0001
	Entering main road [2]	-0.393	0.195	0.044
	Mid Junction - on roundabout or on main road [2]	0.791	0.107	<0.0001
Skidding / Overturning [1]	None [1]	-1.502	0.168	<0.0001
	Skidded [1]	-1.374	0.172	<0.0001
	Skidded and overturned [1]	-1.405	0.186	<0.0001
	Jackknifed and overturned [1]	-2.704	0.493	<0.0001
	Overturned [1]	-1.749	0.184	<0.0001
Skidding / Overturning [2]	None [2]	-1.540	0.168	<0.0001
	Skidded [2]	-1.424	0.172	<0.0001
	Skidded and overturned [2]	-1.438	0.185	<0.0001
	Jackknifed and overturned [2]	-1.976	0.395	<0.0001
	Overturned [2]	-1.979	0.182	<0.0001
Hit Object In Carriageway [1]	Bridge (roof) [1]	1.809	0.286	<0.0001
	Bollard or refuge [1]	4.163	0.082	<0.0001
	Open door of vehicle [1]	2.072	0.097	<0.0001
	Central island of roundabout [1]	-1.682	0.492	0.001
Hit Object In Carriageway [2]	Previous accident [2]	-1.126	0.473	0.017
	Road works [2]	-0.702	0.178	<0.0001
	Bridge (roof) [2]	1.623	0.291	<0.0001
	Bollard or refuge [2]	4.010	0.082	<0.0001
	Open door of vehicle [2]	2.321	0.098	<0.0001
	Central island of roundabout [2]	-1.589	0.478	0.001
Vehicle Leaving Carriageway [1]	Nearside [1]	-0.908	0.080	<0.0001
	Straight ahead at junction [1]	-1.263	0.260	<0.0001
	Offside on to central reservation [1]	-0.811	0.267	<0.0001
	Offside on to centrl res + rebounded [1]	-0.835	0.372	0.025
	Offside - crossed central reservation [1]	-1.287	0.380	<0.0001
	Offside [1]	-0.639	0.101	<0.0001
Vehicle Leaving Carriageway [2]	Nearside [2]	-1.316	0.077	<0.0001
	Nearside and rebounded [2]	-0.446	0.207	0.031
	Straight ahead at junction [2]	-2.158	0.253	<0.0001
	Offside on to central reservation [2]	-1.004	0.255	<0.0001
	Offside on to centrl res + rebounded [2]	-0.728	0.353	0.039
	Offside - crossed central reservation [2]	-1.748	0.355	<0.0001
	Offside [2]	-1.098	0.098	<0.0001
	Offside and rebounded [2]	-0.680	0.242	0.005
Hit Object Off Carriageway [1]	Lamp post [1]	0.849	0.265	0.001
	Submerged in water [1]	-1.619	0.628	<0.0001
	Entered ditch [1]	0.802	0.214	<0.0001
	Other permanent object [1]	0.451	0.172	0.009
	Wall or fence (from 2011) [1]	0.475	0.208	0.022
Hit Object Off Carriageway [2]	Lamp post [2]	0.671	0.264	0.011
	Tree [2]	-0.640	0.168	<0.0001
	Submerged in water [2]	-2.109	0.538	<0.0001
	Entered ditch [2]	0.838	0.213	<0.0001

* [1] serious injury, [2] slight injury. The fatal injury is the reference category and set to zero.

Table 4.4: Multinomial Logit Model Results for Vehicle Data (Continued)

Variable	Factor	Estimated coefficient	Standard error	p-value
First Point of Impact [1]	Front [1]	-0.521	0.119	<0.0001
First Point of Impact [2]	Did not impact [2]	-0.891	0.145	<0.0001
	Front [2]	-0.947	0.117	<0.0001
	Offside [2]	-0.620	0.128	<0.0001
	Nearside	-0.463	0.134	<0.0001
Sex of Driver [1]	Male [1]	-0.618	0.070	<0.0001
Sex of Driver [2]	Male [2]	-0.790	0.069	<0.0001
Hit and Run [1]	Other [1]	-0.679	0.274	0.013
	Non-stop vehicle not hit [1]	4.579	0.057	<0.0001
Hit and Run [2]	Other [2]	-1.062	0.270	<0.0001
	Non-stop vehicle not hit [2]	4.335	0.057	<0.0001
Log likelihood with constants only		-103312		
Log likelihood at final		-98916		
Number of observations		82570		
p-value (from likelihood ratio test)		<0.0001		

* [1] serious injury, [2] slight injury. The fatal injury is the reference category and set to zero.

i. **Age**

Age of driver is significant in explaining serious and slight injury accident. Higher age reduce the relative risk by a factor of 0.995 for serious and 0.988 for slight injury. This indicates that driver's age is significant, but slightly contribute to the accident severity.

ii. **Vehicle Type**

Almost all types of vehicles are statistically significant and shows positive increase relative risk in explaining accident severity both for serious and slight injury. The highest relative risk for both accident severity is belong to motorcycle with 50 cc and below. Motorcycle with low cc commonly designed with lack of safety features such as no Air Braking System (ABS), poor handling and suspension. Despite from that, relative risk of 'goods vehicle' for serious injury is expected to decrease by a factor of 0.61 and 0.59, while 'Goods vehicle' variable is not significant factor in slight injury, given the other variables in the model held constant. This situation probably because 'goods vehicle' is usually drive by professional experienced driver who expert in driving heavy vehicle.

iii. **Vehicle Manoeuvre**

From the result presented above (Refer Table 4.2), both accident severity have reduce relative risk when it involves ’overtaking’, ’going ahead’ and ’right lane movement’. The highest relative risk is belongs to ’waiting to go-held up’ with a factor of 48,678 for serious and 71,040 for slight injury accident. Most of UK drivers drive manual transmission vehicle. When the engine is in idle mode, there is high probability that the driver is careless while shifting the gear that cause the accident.

iv. **Vehicle Location**

All accident factor in the location of junction shows reduce relative risk except for ’cycle lane’. The relative risk increase by 35.52 and 29.93 both for serious and slight accident. Cycle lane commonly use by vulnerable road user resulting with more severe accident.

v. **Junction Location and Skidding/Overturning**

The contributing factor that increase the probability of accident is ’roundabout’ both for serious and slight injury. Roundabout has busy road flow that cause higher accident risk. However, no evidence for ’Skidding/Overturning’ can increase the risk of accident.

vi. **Hit Object in Carriageway**

The relative risk of serious and slight accident is increase when a vehicle hit ’bridge’, ’bollard’ and ’open door of vehicle’ in carriageway. Among these three objects, hitting ’bollard’ in carriageway has increase accident risk the most, with a factor of 64.26 for serious and 55.15 for slight injury accident. Bollard is a short and sturdy post used as a divider or barrier in road traffic. It is made from steel and concrete. Therefore, the consequences of hitting a bollard is expected to be severe and catastrophic because of the high impact involved.

vii. **First Object Hit Off Carriageway**

From MNL model results at Table 4.3, 'lamp post' plays an important role to explain the increase relative risk of accident severity with a factor of 2.34 for serious and 1.96 for slight injury accident. Comparing these values to the above case, it is found that hitting an object off carriageway is not as severe as happened in carriageway. More variables are significant in serious accident ('entered ditch', 'other permanent object', 'wall or fence'), but not significant in slight accident.

xi. **Sex of Driver**

The relative risk for male driver is reduced by a factor of 0.54 and 0.45 for serious and slight injury accident, given the other variables in the model held constant.

4.3 Log-linear Graphical Models

4.3.1 Accident Data

From the result generated from multinomial logistic regression, it is found that all the variables are highly correlated and it is might be the one of the failing cause to obtain the parsimonious model. Therefore, it may be useful to find the graphical model fitting to the variables of the data and then try to refine the analysis.

Accident data for this analysis used is the same data discussed previously. 11 variables taken into account which are:

(a) Accident severity (fatal, serious, slight)

(b) First road class

(c) Road type

(d) Junction detail

(e) Junction control

(f) Second road class

(g) Light conditions

(h) Weather conditions

(i) Road surface conditions

(j) Special conditions at site

(k) Carriageway hazards

Details for each variable used can be found in Appendix. The choice of variables does not affect the principle of graphical modelling and all variables considered as response variables. Hierarchical log linear model is applied, with the aim of finding which graph consists of variables linked in cliques, that correspond to the severity of road crashes. This analysis is done using R software package *Rgraphviz*, *RBGL*, *grid*, *gRbase*, *lcd*, *ggm*, *rgl*, *sna*, and *gRim*. All of this package is retrieved from Bioconductors and CRAN site.

The analysis is proceed by selecting parsimonious model using backward selection, initiating from the saturated model. Since the table is very sparse, asymptotic chi-square distribution may be questionable. The graph generated is presented in Figure 4.1.

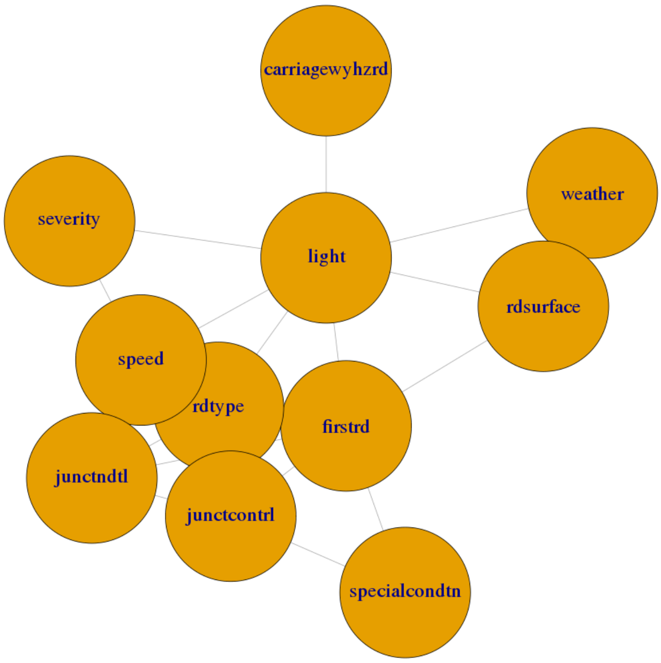


Figure 4.1. The Step-wise Log-linear Graphical Model for GB Accident Data with 11 Variables

Directly on the independence graph (Figure 4.1) it can be observed that, a clique of Speed limit - Light Conditions - Accident Severity was formed. Accident severity is independent of other variables in the model. It does not mean that the other variables are not associate with accident severity, but conditioning on the fact that an accident has occurred, the information provided by these three variables are important. Thus, this analysis highlighted that only two variables are crucial to explain accident severity, which are Speed Limit and Light Conditions.

The graphical model presented above suggests that there is a four-way interaction between accident severity, types of first road class, junction control and light condition. This result is also parallel with the previous regression output and validate this findings. Thus, studying those three variables can be further continued to understand their association in depth.

4.3.2 Vehicle Data

Vehicle data consisted of 252, 836 observations with 150 factor variables. After several attempts, log-linear graphical result for this data could not be obtained because of the size issue of the data. Numerous trials of data manipulation and sampling has been done to find the solution but failed. The R-software graphical model package cannot handle big data. Vehicle data is then proceed with MNLRE applying Bayesian analysis for further examination.

4.4 Multinomial Logistic Random Effects Model

4.4.1 Accident Data

Study is proceed with MNLRE by applying consensus M-H algorithm to accident data. All data is recoded into binary form, 1 and 0. Numerous M-H simulation technique is done repeatedly to find the most appropriate value for unknown parameters and proposal. The proposal density both for $q(\alpha_i^{'}|\alpha_i)$ and $q(\beta_i^{'}|\beta_i)$ to be a normal distribution with mean value equal to the chain value generated in previous step, while standard deviation value is equal to the value obtained from previous iteration. The following variables were incorporated in this analysis:

- y_{ijk} is a matrix of $k \times 3$ represent by Accident Severity (with 3 levels: fatal, serious, slight. Reference group is fatal)
- x_{jk} is a matrix of $k \times 75$ represent by all parameters in Accident dataset, details in Appendix

- u_j is a matrix of $k \times 1$ represent by Police Force location (with 51 levels: details in Appendix)

The analysis begin with assumption value of α_i and β_i. Since there are three categories of response variables, hence, two values of α_i should be selected to start this analysis. As for the beginning, α_1 and α_2 value are equal to 1 to denote the intercept value for serious and slight injury. α_1 and α_2 are in matrix form of $k \times 1$. The starting values used were adjusted accordingly after every simulations until the convergence is achieved.

Meanwhile, β_1 (serious injury accident) and β_2 (slight injury accident) are denoted as parameter value for x_{ij} variables. Therefore, there are 75 different assumption values to start with. β_1 is set in matrix form of 75×1 with all row values equal to 2, while β_2 is in matrix form of 75×1 with all row values equal to 3. Same like the above case, starting values were updated after every simulations until it converge well. The above explanation can be summarised as below:

$$\underset{k \times 1}{\begin{pmatrix} y_{ij;1,1} \\ \vdots \\ y_{ij;k,1} \end{pmatrix}} = \alpha_i \underset{k \times 1}{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}} + \underset{k \times 75}{\begin{pmatrix} x_{j;1,1} & \cdots & x_{j;1,75} \\ \vdots & \ddots & \vdots \\ x_{j;k,1} & \cdots & x_{j;k,75} \end{pmatrix}} \underset{75 \times 1}{\begin{pmatrix} \beta_{i;1,1} \\ \vdots \\ \beta_{i;75,1} \end{pmatrix}} + \underset{k \times 1}{\begin{pmatrix} u_{j;1,1} \\ \vdots \\ u_{j;k,1} \end{pmatrix}}$$

Prior simulations, data are divided to 20 shards and each shard is run with M-H algorithm simulation independently to generate samples from posterior distribution as discussed in Chapter 3. Several simulations are done to find the most optimum starting and proposal value for each shard. 10,000 iterations is ran and the first 5,000 as burn-in stage for trials. Generated simulation samples are then plotted into chain graph to check for the stationarity feature. After the stationary distribution chain is achieved for each parameter, final value for each chain is used to run longer iteration for each parameter of every shard. Then final concensus simulation is performed by applying matrix weighting scheme to combine all the 20 shards. The final combined results are illustrated as below:

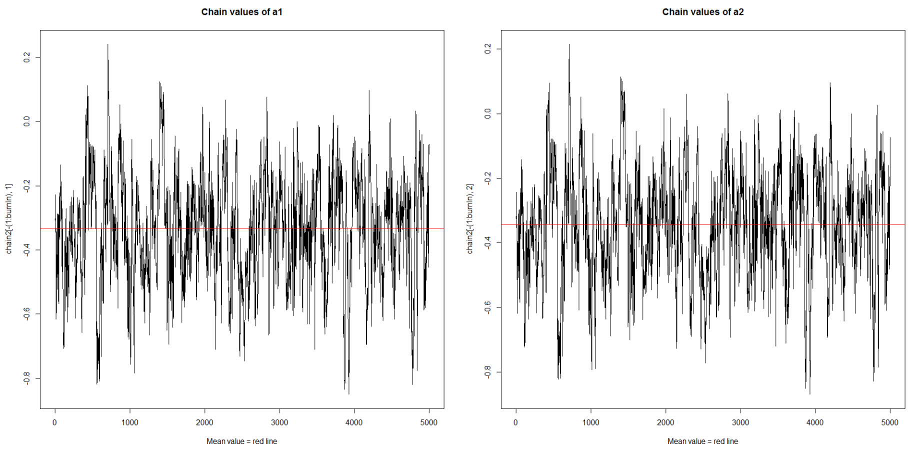


Figure 4.2. Plots for α_1 and α_2 chains

Both α_1 and α_2 are the intercept value for serious and slight injury MNLRE prediction. Chain plot in Figure 4.2 are stationary indicated that the chains are successfully converged. Summarized statistical value for α_1 and α_2 are as below:

Table 4.5: Summary Value of α_1 and α_2

Variable definition	Mean	Standard deviation	95% sample-based credible sets	
			2.5%	97.5%
α_1	-0.3336	0.1660	-0.6067	-0.0605
α_2	-0.3422	0.1664	-0.6160	-0.0684

Parameter $\beta_{1,j}$ is used to explain all variables in serious accident estimation, while $\beta_{2,j}$ represented slight injury parameters. The M-H results of $\beta_{i,j}$ is presented below:

Table 4.6: Summary Value of β_1 and β_2

Variable definition	β_1				β_2			
	Mean	Standard deviation	95% sample-based credible sets 2.5%	97.5%	Mean	Standard deviation	95% sample-based credible sets 2.5%	97.5%
First road class								
Motorway	-0.1102	0.7223	-1.2983	1.0780	-1.3605	0.3898	-2.0018	-0.7192
A (M)	-0.8461	0.8507	-2.2456	0.5533	-0.8007	0.6398	-1.8532	0.2518
A	-0.0387	0.3146	-0.5562	0.4788	-1.2366	0.1496	-1.4827	-0.9905
B	-0.0294	0.3891	-0.6695	0.6107	-1.1684	0.1563	-1.4256	-0.9112
C	-0.4712	0.4606	-1.2289	0.2865	-1.3577	0.1721	-1.6408	-1.0747
Unclassified	-0.4201	0.3783	-1.0424	0.2022	-1.1618	0.1504	-1.4091	-0.9145
Road type								
Roundabout	-0.4413	0.5562	-1.3563	0.4737	0.1391	0.1862	-0.1671	0.4454
One way street (from 2005)	-0.4521	0.7146	-1.6277	0.7234	-0.0619	0.2419	-0.4598	0.3359
Dual carriageway	-0.1622	0.4021	-0.8237	0.4993	0.1192	0.1692	-0.1591	0.3975
Single carriageway	-0.3666	0.3335	-0.9152	0.1820	0.1057	0.1472	-0.1364	0.3479
Slip road	-0.7671	0.6542	-1.8433	0.3090	-0.6942	0.3285	-1.2345	-0.1538
Unknown	-0.4070	0.8238	-1.7622	0.9482	-0.4672	0.5094	-1.3053	0.3708
One way street/slip road (1974-2004)	-0.1706	0.9303	-1.7010	1.3597	0.0263	0.9486	-1.5340	1.5867
Speed limit (mph)								
20	-0.8816	0.6862	-2.0105	0.2472	0.2864	0.2338	-0.0981	0.6709
30	-1.0775	0.3101	-1.5876	-0.5674	0.2026	0.1345	-0.0186	0.4238
40	-0.9675	0.4452	-1.6998	-0.2351	0.1883	0.1614	-0.0773	0.4538
50	-0.3041	0.4604	-1.0613	0.4532	0.4285	0.1990	0.1012	0.7559
60	-0.2452	0.3323	-0.7917	0.3014	0.6563	0.1584	0.3957	0.9169
70	-0.0780	0.5218	-0.9364	0.7803	0.4602	0.2222	0.0947	0.8256
Junction detail								
Not at junction	-0.6957	0.8931	-2.1648	0.7735	0.1193	0.8961	-1.3548	1.5933
Roundabout	-0.7546	0.4485	-1.4924	-0.0167	-0.3702	0.1575	-0.6293	-0.1111
Mini-roundabout	-0.6249	0.6986	-1.7741	0.5244	-0.2047	0.2370	-0.5945	0.1852
T or staggered junction	-0.0415	0.2803	-0.5025	0.4196	-0.0264	0.1117	-0.2102	0.1574
Slip road	-0.2233	0.5805	-1.1781	0.7316	-0.0866	0.2265	-0.4592	0.2860
Crossroads	-0.2048	0.3585	-0.7946	0.3850	-0.0829	0.1252	-0.2889	0.1231
More than 4 arms	-0.5113	0.6974	-1.6584	0.6359	0.0480	0.2416	-0.3494	0.4454
Private drive or entrance	-0.6142	0.4836	-1.4096	0.1813	-0.0215	0.1530	-0.2732	0.2301
Other junction	-0.3262	0.4909	-1.1336	0.4813	-0.1070	0.1678	-0.3830	0.1690
Junction control								
Unathorised person	0.0342	0.8710	-1.3985	1.4670	-1.1275	0.5066	-1.9607	-0.2942
Auto traffic signal	-1.2079	0.4345	-1.9226	-0.4931	-1.2380	0.1313	-1.4539	-1.0220
Stop sign	-0.9905	0.7323	-2.1952	0.2142	-1.1826	0.2936	-1.6656	-0.6996
Give way or uncontrolled	-0.7131	0.3704	-1.3225	-0.1038	-1.2839	0.1118	-1.4678	-1.1000
Second road class								
Motorway	-0.9902	0.7316	-2.1936	0.2133	0.2718	0.3217	-0.2574	0.8011
A(M)	0.3681	0.8240	-0.9873	1.7235	-0.0767	0.6569	-1.1573	1.0039
A	-0.6010	0.4476	-1.3373	0.1354	0.3505	0.1414	0.1178	0.5832
B	-0.5350	0.4521	-1.2788	0.2088	0.4022	0.1581	0.1422	0.6623
C	-0.4293	0.4411	-1.1549	0.2963	0.4254	0.1621	0.1588	0.6919
Unclassified	-0.5780	0.3278	-1.1171	-0.0388	0.4745	0.1256	0.2678	0.6811
Light conditions								
Daylight	-0.5570	0.3229	-1.0882	-0.0259	0.6797	0.0987	0.5173	0.8422
Darkness - lights lit	-0.1437	0.3537	-0.7254	0.4381	0.9421	0.1111	0.7594	1.1248
Darkness - lights unlit	-0.5827	0.7390	-1.7983	0.6329	0.4414	0.3851	-0.1921	1.0749
Darkness - no lighting	-0.0609	0.4869	-0.8619	0.7400	0.4324	0.1975	0.1076	0.7573
Darkness - lighting unknown	-0.8728	0.7336	-2.0795	0.3339	0.6938	0.2585	0.2686	1.1189
Weather conditions								
Fine no high winds	-0.1300	0.3514	-0.7081	0.4481	-0.0459	0.0988	-0.2084	0.1165
Raining no high winds	-0.1460	0.4597	-0.9022	0.6102	-0.2945	0.1344	-0.5157	-0.0733
Snowing no high winds	-0.4691	0.8081	-1.7985	0.8603	-0.6623	0.3785	-1.2849	-0.0397
Fine + high winds	-0.2521	0.6292	-1.2871	0.7829	-0.3214	0.2522	-0.7362	0.0935
Raining + high winds	-0.8997	0.6785	-2.0158	0.2164	-0.2491	0.2567	-0.6714	0.1731
Snowing + high winds	-0.6902	0.8264	-2.0496	0.6692	-0.4265	0.5562	-1.3415	0.4884
Fog or mist	-0.7096	0.7926	-2.0133	0.5942	0.2156	0.4310	-0.4933	0.9245
Other junction	-0.3421	0.6800	-1.4606	0.7765	-0.0474	0.2426	-0.4465	0.3518
Unknown	-0.7152	0.6355	-1.7605	0.3301	-0.1351	0.2300	-0.5135	0.2433
Road surface								
Dry	-0.7788	0.3401	-1.3383	-0.2193	-0.6445	0.0977	-0.8052	-0.4837
Wet /damp	-0.9205	0.3795	-1.5448	-0.2962	-0.5602	0.1117	-0.7438	-0.3765
Snow	-1.4623	0.7859	-2.7550	-0.1695	-0.6527	0.4073	-1.3227	0.0172
Frost/ice	-0.8139	0.7274	-2.0105	0.3826	-0.3061	0.2806	-0.7678	0.1555
Flood	0.1301	0.8302	-1.2356	1.4958	-0.4946	0.7718	-1.7642	0.7750
Oil/diesel	0.0328	0.9053	-1.4565	1.5220	-0.3581	0.8994	-1.8376	1.1214
Mud	-0.1999	0.9063	-1.6907	1.2910	0.3261	0.9557	-1.2460	1.8983
Special conditions at site								
None	-0.4266	0.3520	-1.0057	0.1524	-0.6534	0.0817	-0.7878	-0.5190
Auto traffic signal - out	-0.7885	0.8537	-2.1929	0.6159	-0.5648	0.5241	-1.4269	0.2973
Auto signal part defective	-0.1108	0.8962	-1.5851	1.3635	-0.2329	0.7998	-1.5485	1.0827
Road sign or marking defective	-0.2701	0.8815	-1.7201	1.1800	-0.7758	0.5842	-1.7368	0.1852
Roadworks	-0.2128	0.7113	-1.3829	0.9573	-0.5061	0.3006	-1.0005	-0.0116
Road surface defective	-0.7756	0.8071	-2.1032	0.5520	-0.1837	0.5830	-1.1428	0.7754
Oil or diesel	-0.4995	0.8668	-1.9253	0.9263	-0.2769	0.4972	-1.0948	0.5410
Mud	-0.4149	0.8668	-1.8408	1.0109	-0.3057	0.5651	-1.2352	0.6238

Table 4.7: Summary Value of β_1 and β_2 (Continued)

Variable definition	β_1				β_2			
	Mean	Standard deviation	95% sample-based credible sets		Mean	Standard deviation	95% sample-based credible sets	
			2.5%	97.5%			2.5%	97.5%
Carriageway hazards								
None	0.4880	0.3343	-0.0619	1.0379	0.1831	0.0732	0.0628	0.3034
Vehicle load on road	-0.1967	0.9005	-1.6780	1.2847	0.4792	0.7068	-0.6835	1.6419
Other object on road	-0.4538	0.8008	-1.7711	0.8634	0.3209	0.3660	-0.2810	0.9229
Previous accident	-0.0638	0.8326	-1.4334	1.3057	0.1379	0.6864	-0.9913	1.2671
Dog on road	-0.6045	0.9027	-2.0895	0.8806	-0.0241	0.8529	-1.4271	1.3789
Other animal on road	-0.6151	0.8738	-2.0525	0.8223	-0.1383	0.8894	-1.6014	1.3248
Pedestrian in carriageway	-0.4174	0.8671	-1.8438	1.0089	0.0046	0.4605	-0.7530	0.7622
Any animal in carriageway	-0.5524	0.8188	-1.8993	0.7944	-0.0806	0.5526	-0.9897	0.8285

The highlighted rows represent the statistically significant parameters, based on the 95% (2.5-97.5%) sample-based credible sets. The parameter estimates in Table 4.6 and 4.7 show that more variables are significant for slight injury than for serious injury accidents given fatal accident as the reference group.

From the results, clearly, road class is an important contributory factor in slight injury accidents for both first and second road classes. Another variable, road type of slip road is also a significant factor associated with slight injury accidents. Junction detail of roundabout increases the likelihood of road accidents involving both slight and serious injuries.

The MNLRE model results are consistent with the earlier MNL results that majority of serious injury accidents occurred on roads with low speed limits, i.e., 30 and 40 mph; while road accidents happen at higher speed limits areas, i.e., 50, 60 and 70 mph could lead to less severe accidents and injuries for drivers.

The variable, junction control is not significant in the MNL models, but the MNLRE results show that two variables, auto traffic signal and uncontrolled, i.e., no stop light or sign with only give way sign are significant in explaining serious injury accidents. However, all variables are significant in the occurrence of road accidents involving slight injuries.

Interestingly, the results indicate that light conditions are significant factors determining slight injuries accidents in both daylight and darkness (roads with and without road lighting), but it is not a significant cause of serious injuries accidents.

As discussed earlier, the effects of weather conditions could lead both to serious and slight injuries with three and five variables are significant, respectively. The results, as shown in Table 4.6 show that the likelihood of road accidents increases during raining and snowing with no high winds conditions. These challenging conditions could lead to slight injuries. In the adverse weather conditions, for example, high winds, a road weather warning not to travel unless necessary can be issued to drivers living in areas affected by poor weather conditions. Furthermore, drivers may theoretically acknowledge the need to reduce their speed in high winds weather conditions, hence, reduce road accidents.

The slipperiness of the road surface related to wetness, e.g., wet/damp and snow influence driving conditions to a much greater extent than other conditions like frost/ice, flood, oil/diesel or mud. Water on the road surface can lead to hydroplaning and skidding which may result in slight and serious injuries accidents. Slippery roads due to snow during the winter is only significant in serious injuries accident, but not for slight accidents.

In addition, surprisingly, the analysis shows that the dry road surface significantly contributed to road accidents for both slight and serious injuries accidents. A plausible explanation is that drivers are more likely to drive faster on a dry road surface when they think it is safe, thus increasing the likelihood of road accidents.

4.4.2 Vehicle Data

Similar like 'Accident' data, 'Vehicle' data also is in categorical form, all data is again recoded into binary form, 1 and 0. The same method applied as in the previous analysis, noting that 'Vehicle' data has more variables and bigger in size. 'Age of Driver' variable is excluded from the analysis as the measure does not conform to other variables which have been recoded into binary form. Therefore, it left 121 variables to be analysed in MNLRE for 'Vehicle' data.

Simulated Study

Unknown parameters and proposal value is estimated by applying M-H simulation repeatedly until the appropriate converged value is obtained. The same assumption in 'Vehicle' data also applied for the proposal density, $q(\alpha_i'|\alpha_i)$ and $q(\beta_i'|\beta_i)$ to be a normal distribution with mean value equal to the value generated in previous chain step, while standard deviation is equal to the value obtained from previous iteration. Below are the summary details of variables involved:

- y_{ijk} is a matrix of $k \times 3$ represent by Accident Severity (with 3 levels: fatal, serious, slight. Reference group is fatal)

- x_{jk} is a matrix of $k \times 121$ represent by all parameters in Accident dataset, details in Appendix

- u_j is a matrix of $k \times 1$ represent by Police Force location (with 51 levels: details in Appendix)

The simulation commenced with setting α_1 and α_2 equal to 1 as the intercept value for serious and slight injury. β_1 initial value is set equal to 2, while β_2 equal to 3. This can be summarised as below:

$$\underset{k\times 1}{\begin{pmatrix} y_{ij;1,1} \\ \vdots \\ y_{ij;k,1} \end{pmatrix}} = \alpha_i \underset{k\times 1}{\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}} + \underset{k\times 121}{\begin{pmatrix} x_{j;1,1} & \cdots & x_{j;1,121} \\ \vdots & \ddots & \vdots \\ x_{j;k,1} & \cdots & x_{j;k,121} \end{pmatrix}} \underset{121\times 1}{\begin{pmatrix} \beta_{i;1,1} \\ \vdots \\ \beta_{i;121,1} \end{pmatrix}} + \underset{k\times 1}{\begin{pmatrix} u_{j;1,1} \\ \vdots \\ u_{j;k,1} \end{pmatrix}}$$

Data are divided to 20 shards and each shard is run with M-H algorithm simulation independently to generate samples from posterior distribution. Several simulations are done to find the most optimum starting and proposal value for each shard. The final combined results are presented as below:

Table 4.8: Summary Value of α_1 and α_2

Variable definition	Mean	Standard deviation	95% sample-based credible sets	
			2.5%	97.5%
α_1	0.6865	0.0599	0.5881	0.7850
α_2	0.6871	0.0600	0.58845	0.7858

Table 4.9: Summary Value of β_1 and β_2

Variable definition	β_1				β_2			
	Mean	Standard deviation	95% sample-based credible sets		Mean	Standard deviation	95% sample-based credible sets	
			2.5%	97.5%			2.5%	97.5%
Vehicle type								
Pedal cycle	-1.2992	0.2366	-1.68843035	-0.90991765	-0.8309	0.0769	-0.95733805	-0.70436995
Motorcycle 50cc and under (from 1999)	-1.2287	0.5307	-2.10164025	-0.35580175	-0.6908	0.1370	-0.9162338	-0.4653722
M/cycle over 50 to 125cc (from 2005)	-1.1217	0.3156	-1.6408432	-0.6026508	-0.4277	0.0929	-0.5804917	-0.2749823
M/cycle over 125cc and up to 500cc (from 2005)	-0.3770	0.4225	-1.07202895	0.31802895	-0.2263	0.1339	-0.4465804	-0.0059836
M/cycle over 500cc (from 2005)	0.0906	0.2066	-0.24919875	0.43035075	-0.0333	0.0870	-0.1763012	0.1097972
Electric m/cycle (from 2011)	-0.6453	0.7484	-1.87641955	0.58578355	-0.4911	0.7624	-1.7452092	0.7629552
M/cycle - unknown cc (from 2011)	-0.3351	0.8052	-1.65958855	0.98948655	0.0896	0.6661	-1.0060827	1.1852547
M/cycle - scooter (1979-1998)	-0.8396	0.8287	-2.2028403	0.5237143	-0.2520	0.7276	-1.4489709	0.9448989
M/cycle (1979-1998)	-0.3960	0.7954	-1.70448635	0.91247835	-0.2210	0.7653	-1.4799453	1.0380233

Table 4.10: Summary Value of β_1 and β_2 (Continued)

Variable definition	β_1				β_2			
	Mean	Standard deviation	95% sample-based credible sets 2.5%	95% sample-based credible sets 97.5%	Mean	Standard deviation	95% sample-based credible sets 2.5%	95% sample-based credible sets 97.5%
M/cycle - Combination (1979-1998)	-0.4334	0.7888	-1.73094665	0.86410665	-0.9354	0.7833	-2.2239087	0.3530167
M/cycle over 125cc (1999-2004)	-0.1308	0.8081	-1.4601873	1.1985933	-0.6250	0.8087	-1.9553045	0.7053185
Taxi / Private hire car (from 2005)	-0.9340	0.3583	-1.52331925	-0.34467675	-1.0647	0.1186	-1.2598278	-0.8695022
Taxi (excl private hire cars) (1979-2004)	-0.4859	0.7747	-1.76024915	0.78841515	0.1445	0.7587	-1.1035993	1.3926553
Car (from 2005)	-1.3743	0.1447	-1.6122915	-1.1362285	-1.4133	0.0644	-1.519197	-1.307321
Car (incl private hire cars) (1979-2004)	-0.3317	0.7906	-1.632228	0.968846	-0.0600	0.7545	-1.30109805	1.18117405
Minibus (from 1999)	-0.6704	0.5438	-1.56493445	0.22420045	-1.0120	0.2663	-1.45012585	-0.57390015
Minibus/Motor caravan (1979-1998)	-0.3845	0.7821	-1.6710526	0.9019906	-0.2788	0.7936	-1.58428245	1.02669445
Bus or coach (17 or more pass seats)	-0.7478	0.2932	-1.23015945	-0.26549855	-1.4894	0.1152	-1.67887755	-1.29990245
Van / Goods 3.5 tonnes mgw or <	-1.0474	0.2254	-1.4182098	-0.6765122	-1.4198	0.0927	-1.5722665	-1.2672835
Goods >3.5t. and <7.5t. (from 1999)	-0.5840	0.3675	-1.1884647	0.0204787	-1.0704	0.1758	-1.3596468	-0.7811332
Goods 7.5 tonnes >(from 1999)	0.1531	0.2447	-0.24945295	0.55564295	-0.9321	0.1391	-1.16085315	-0.70331285
Goods vehicle - unknown weight	-0.6651	0.6865	-1.7943186	0.4642006	-0.8298	0.6216	-1.85234635	0.19281635
Goods >3.5 tonnes (1979-1998)	-0.6750	0.7627	-1.9295976	0.5796196	0.2354	0.7906	-1.065171	1.535903
Ridden horse (from 1999)	-0.9618	0.7120	-2.13306745	0.20944545	-0.6901	0.4606	-1.44777645	0.06763045
Agricultural vehicle (from 1999)	0.1979	0.4797	-0.5913133	0.9870313	-0.9324	0.2660	-1.36985375	-0.49487825
Tram (from 1999)	-0.4387	0.7772	-1.7171732	0.8396832	-0.4222	0.7305	-1.62386815	0.77937815
Mobility scooter (from 2011)	-0.4025	0.6329	-1.44362105	0.63858705	-0.6394	0.4181	-1.3271337	0.0482837
Other vehicle	-0.5029	0.4364	-1.2207911	0.2148991	-0.8726	0.1815	-1.17118205	-0.57407995
Towing and Articulation								
No tow/articulation	-0.4352	0.1509	-0.6833666	-0.1869714	0.5354	0.0566	0.44223265	0.62854535
Articulated vehicle	-0.3360	0.3117	-0.84874615	0.17664815	0.4843	0.1792	0.1895659	0.7790681
Double/multiple trailer	-0.9799	0.7253	-2.1729357	0.2131697	-0.2522	0.6697	-1.35387605	0.84940405
Caravan	0.1669	0.6308	-0.8707322	1.2044682	0.4839	0.4319	-0.2265884	1.1944284
Single trailer	-0.3906	0.4412	-1.1164648	0.3352148	0.5600	0.2068	0.2197911	0.9002289
Other tow	-0.4057	0.6569	-1.48623415	0.67486815	-0.0876	0.4202	-0.77874065	0.60361865
Vehicle Manoeuvre								
Reversing	-0.4388	0.4581	-1.19237585	0.31487185	-0.2283	0.1440	-0.46518435	0.00867435
Parked	0.0471	0.2622	-0.3841681	0.4784041	-0.4139	0.1051	-0.58676595	-0.24095405
Waiting to go - held up	-0.5517	0.3308	-1.09585545	-0.00749055	-1.0569	0.1060	-1.23121265	-0.88257135
Slowing or stopping	-0.9462	0.3347	-1.49675305	-0.39562295	-1.1076	0.0904	-1.2563119	-0.9588301
Moving off	-1.1544	0.4163	-1.8392055	-0.4695785	-0.5587	0.0977	-0.7193965	-0.3979635
U-turn	-0.5950	0.5765	-1.54337785	0.35340585	-0.2696	0.1671	-0.54442125	0.00517325
Turning left	-0.8618	0.4481	-1.5989654	-0.1246506	-0.5758	0.1036	-0.74611175	-0.40543225
Waiting to turn left	-0.6530	0.7196	-1.8368398	0.5307758	-1.2812	0.3138	-1.79742945	-0.76499455
Turning right	0.0077	0.2437	-0.3930907	0.4085507	-0.2752	0.0775	-0.40268305	-0.14774095
Waiting to turn right	-0.3003	0.5139	-1.14566705	0.54503105	-0.6899	0.1479	-0.93325995	-0.44663605
Changing lane to left	-0.6779	0.4986	-1.49812055	0.14224055	-0.9547	0.1933	-1.27269585	-0.63664015
Changing lane to right	-0.6787	0.4490	-1.4172471	0.0598971	-0.5498	0.1790	-0.8442651	-0.2554209
Overtaking moving veh - offside	0.4462	0.2584	0.021106	0.871242	-0.3619	0.1094	-0.5419118	-0.1818542
Overtaking static veh - offside	-0.8282	0.4834	-1.6233191	-0.0329989	-0.5510	0.1279	-0.7613677	-0.3407083
Overtaking - nearside	-0.3611	0.5272	-1.2284219	0.5061319	-0.6797	0.1858	-0.9854508	-0.3740372
Going ahead left-hand bend	0.7201	0.1928	0.4029262	1.0373698	-0.3295	0.0851	-0.46951705	-0.18957095
Going ahead right-hand bend	0.3642	0.1808	0.0668469	0.6616131	-0.3578	0.0803	-0.48994185	-0.22565615
Going ahead other	0.5647	0.1331	0.34578695	0.78365305	-0.4680	0.0600	-0.5667458	-0.3692142
Vehicle Location								
On main c'way - not in res lane (from 1999)	0.2932	0.1577	0.03387785	0.55261215	0.7997	0.0495	0.71827295	0.88109505
Tram/Light rail track (from 1999)	-0.6453	0.7859	-1.9380386	0.6475066	-0.0311	0.6846	-1.157315	1.095019
Bus lane (from 1999)	-0.8209	0.6660	-1.9164601	0.2746141	0.5819	0.2266	0.2091431	0.9547229
Busway (from 1999)	-0.3349	0.7698	-1.60128635	0.93145435	-0.2472	0.5739	-1.19120925	0.69675725
Cycle lane (on main cway) (from 1999)	-0.3803	0.6633	-1.4714325	0.7108245	0.2526	0.2738	-0.1978219	0.7030459
Cycleway or shared use footway	0.0841	0.6784	-1.0318392	1.1999652	0.1179	0.3861	-0.5172784	0.7530564
On lay-by or hard shoulder	0.2040	0.5004	-0.6190522	1.0271322	0.1491	0.2951	-0.33636405	0.63448205
Entering lay-by or hard shoulder	-0.0094	0.6652	-1.1037448	1.0848948	0.3037	0.4667	-0.4640983	1.0714763
Leaving lay-by or hard shoulder	-0.7352	0.6972	-1.88199965	0.41168965	0.2504	0.3755	-0.36738085	0.86811285
Footway (pavement) (from 1999)	-0.1809	0.4517	-0.92392195	0.56220395	0.5812	0.1591	0.3194157	0.8429863
Not on carriageway (1979-1998)	-0.5123	0.8197	-1.86064305	0.83613705	-0.7474	0.7632	-2.002911	0.508017
Junction Location of Vehicle								
Not at junction	0.4099	0.1747	0.1226003	0.6972317	-0.1913	0.0525	-0.27766115	-0.10503485
Approachg junct	-0.5858	0.2023	-0.91848925	-0.25308675	-0.4602	0.0594	-0.5579039	-0.3624121
Cleared junction	-0.0939	0.2727	-0.54246015	0.35462415	-0.0804	0.0789	-0.2102195	0.0493615
Leaving roundabout	-0.7118	0.4640	-1.4751079	0.0515179	-0.5745	0.1438	-0.8110042	-0.3380338
Entering roundabout	-1.0902	0.4620	-1.85015465	-0.33027335	-0.5250	0.1126	-0.71012375	-0.33983425
Leaving main road	-0.2380	0.3974	-0.89171045	0.41576845	-0.3498	0.1046	-0.5219818	-0.1777162
Entering main road	-0.0931	0.3117	-0.6058344	0.4197244	-0.4319	0.0889	-0.57806715	-0.28568485
Entering from slip road	-0.3612	0.5713	-1.30095325	0.57845925	-0.6763	0.2681	-1.1173604	-0.2352456
Mid Junction	-0.4710	0.2104	-0.81704775	-0.12499625	-0.4694	0.0580	-0.5648489	-0.3739631
Skidding / Overturning								
None	0.4289	0.1659	0.15600805	0.70185195	0.2466	0.0486	0.1665422	0.3265678
Skidded	0.0946	0.1968	-0.229096	0.418376	0.1695	0.0646	0.0631722	0.2758378
Skidded and overturned	-0.1376	0.2638	-0.57142975	0.29630775	0.0045	0.0995	-0.15916705	0.16815505
Jackknifed	-0.6675	0.6431	-1.72543095	0.39040095	-0.1096	0.4494	-0.84876465	0.62966265
Jackknifed and overturned	0.2690	0.6186	-0.7485951	1.2865331	-0.1502	0.4956	-0.96544045	0.66511645
Overturned	0.2590	0.2722	-0.1888139	0.7067899	0.2784	0.0993	0.1150107	0.4418393
Hit Object In Carriageway								
None	0.1098	0.1611	-0.15524485	0.37487285	-0.7343	0.0445	-0.80746025	-0.66121975
Previous accident	-0.1065	0.6622	-1.19585235	0.98288435	0.0533	0.4413	-0.67269285	0.77928285
Road works	-0.0627	0.6937	-1.20386785	1.07850385	-0.6835	0.5090	-1.52078745	0.15385545
Parked vehicle	-0.4360	0.8101	-1.7685507	0.8965467	-0.3062	0.7918	-1.6086282	0.9962622

Table 4.11: Summary Value of β_1 and β_2 (Continued)

Variable definition	β_1				β_2			
	Mean	Standard deviation	95% sample-based credible sets		Mean	Standard deviation	95% sample-based credible sets	
			2.5%	97.5%			2.5%	97.5%
Bridge (roof)	0.4412	0.3089	-0.0669286	0.9492866	-0.6872	0.1387	-0.91539395	-0.45903805
Bridge (side)	-0.0663	0.7513	-1.30220295	1.16960695	-0.7115	0.7139	-1.8857847	0.4628147
Bollard or refuge	-0.3960	0.6059	-1.3927026	0.6006426	0.5959	0.4109	-0.08004805	1.27178005
Open door of vehicle	-0.4044	0.5047	-1.2346844	0.4258444	-0.7105	0.2203	-1.0729304	-0.3480776
Central island of roundabout	-0.5209	0.6867	-1.6505743	0.6088003	0.7524	0.3238	0.2196821	1.2850499
Kerb	-0.5841	0.6579	-1.66628515	0.49810715	-0.3439	0.3652	-0.94466655	0.25680855
Other object	0.3183	0.2896	-0.15805565	0.79462965	-0.7296	0.1123	-0.91437495	-0.54487505
Any animal (x ridden horse) (from 2005)	-0.4215	-0.4215	0.271843635	-1.114769635	-1.0641	0.2853	-1.5334226	-0.5948514
Vehicle Leaving Carriageway								
Did not leave cway	-0.7413	0.5690	-1.67740135	0.19470735	-1.0358	0.2958	-1.52236565	-0.54928235
Nearside	-1.2571	0.1416	-1.4900619	-1.0241321	-0.9012	0.0453	-0.9757086	-0.8267374
Nearside and rebounded	-0.4288	0.1684	-0.70588935	-0.15175465	-0.5739	0.0684	-0.68647135	-0.46133665
Straight ahead at junction	-1.3040	0.4032	-1.9673198	-0.6406602	-0.7479	0.1461	-0.9881905	-0.5075215
Offside on to central reservation	-0.0125	0.4834	-0.8077061	0.7826141	-0.2932	0.2118	-0.6417128	0.0552408
Offside on to centrl res + rebounded	-0.6148	0.4930	-1.4257212	0.1961172	-0.8650	0.2314	-1.24563045	-0.48429155
Offside - crossd cntrl res	-0.6978	0.4996	-1.51954965	0.12403565	-0.7323	0.2788	-1.1909519	-0.2736341
Offside	-0.1233	0.5957	-1.10324185	0.85670985	-0.3590	0.3949	-1.00854505	0.29064305
Offside and rebounded	-0.5045	0.2007	-0.8346277	-0.1744563	-0.5636	0.0843	-0.70235885	-0.42491315
First Object Hit Off Cway								
None	-0.4855	0.3752	-1.10265475	0.13158875	-0.5274	0.1996	-0.8558548	-0.1990392
Road sign or traffic signal	-0.9269	0.1658	-1.1995641	-0.6541479	-0.9636	0.0467	-1.0404493	-0.8866747
Lamp post	0.7248	0.3612	0.13059555	1.31897645	-0.9917	0.1554	-1.2472462	-0.7361118
Telegraph or electricity pole	-0.8341	0.3757	-1.45209595	-0.21601005	-0.9721	0.1714	-1.25408045	-0.69014155
Tree	-0.7998	0.4714	-1.575301	-0.024395	-1.0522	0.2532	-1.4687528	-0.6355932
Bus stop or bus shelter	-0.1564	0.2323	-0.53851405	0.22572005	-0.7817	0.1117	-0.96550295	-0.59797705
Central crash barrier	-0.7984	0.6660	-1.8939211	0.2971531	-1.0004	0.4897	-1.8059414	-0.1947626
Near/Offside crash barrier	-0.9115	0.4301	-1.61894015	-0.20400985	-1.0310	0.2037	-1.36610585	-0.69583415
Submerged in water	-0.5544	0.3714	-1.1654108	0.0566268	-1.0942	0.1815	-1.39281985	-0.79558615
Entered ditch	0.3339	0.7083	-0.8312616	1.4989796	-0.7385	0.6718	-1.84363545	0.36661945
Other permanent object	-1.3641	0.3519	-1.9429964	-0.7851796	-1.1708	0.1459	-1.4107885	-0.9307775
Wall or fence (from 2011)	-0.7699	0.2342	-1.15517945	-0.38462855	-0.8838	0.1027	-1.0527713	-0.7147567
First Point of Impact								
Did not impact	-0.9603	0.3141	-1.47698395	-0.44356205	-0.9116	0.1309	-1.1268857	-0.6963563
Front	-0.7508	0.2199	-1.1124537	-0.3891143	-0.0029	0.0720	-0.12125875	0.11545675
Back	-0.8529	0.1566	-1.11047445	-0.59522755	-0.2050	0.0458	-0.2802642	-0.1297138
Offside	-1.2142	0.2182	-1.5731541	-0.8553419	-0.6902	0.0616	-0.7915859	-0.5888561
Nearside	-0.8033	0.1878	-1.11213565	-0.49437235	-0.2685	0.0551	-0.3590967	-0.1779493
Sex								
Male	-0.9626	0.1963	-1.28554395	-0.63968405	-0.3529	0.0557	-0.44456585	-0.26121415
Female	-0.9969	0.1498	-1.24334745	-0.75047255	-0.3654	0.0383	-0.4282997	-0.3024243
Not known	-1.3773	0.1752	-1.66541765	-1.08910835	-0.5729	0.0451	-0.64711485	-0.49863715
Hit and Run								
Other	-0.6564	0.4642	-1.4199871	0.1071651	-0.3698	0.2064	-0.7092862	-0.0303618
Hit and run	-0.0840	0.1369	-0.3092084	0.1412584	0.9064	0.0449	0.8325784	0.9802336
Non-stop vehicle not hit	-0.3602	0.3598	-0.9520102	0.2316002	0.4820	0.1219	0.2814056	0.6825224

The highlighted rows represent the statistically significant parameters, based on the 95% (2.5-97.5%) sample-based credible sets. Similar like previous result, parameter estimate in Table 4.9, 4.10 and 4.11 show that more variables are significant for slight injury than for serious injury accidents given fatal accident as the reference group.

The MNLRE model results are consistent with the earlier MNL results, it is clearly showed that pedal cycle and motorcycle are significant in both accident severity. Vulnerable road user is belong to this category as the accident risk is greater for two-wheels vehicle. Pedal cycle and motorcycle lack of safety features, the rider has high tendency to get injured because of the direct contact between vehicle and the surroundings. Apart from that, public transport such as taxi/private hire car, and bus/coach increase the likelihood of road accidents involving both slight and serious injuries. Reckless driver and poor vehicle maintenance may one of the reason contribute to this problem. Another unexpected variable that is significant in both serious and slight injury accident is involving car vehicle that was produced from year 2005 onwards. Although the car is still consider as new condition (not more 8 years old) and equipped with better safety features, driver factor may be one of the factor for this cause. Meanwhile, goods vehicle is significant in slight accident case, but not in serious accident.

Contradict from MNL results in Table 4.2, serious and slight injury accident has increase probability when there is no towing/articulation involved. Analysis showed that articulated vehicle and single trailer is significant in slight injury accident only.

For vehicle maneuver variable, almost all of the variables are significant in slight injury accident. Meanwhile, for serious injury accident, lesser significant variables spotted which are, 'waiting-to go held up', 'slowing/stopping', 'moving off', 'turning left', 'overtaking' and 'going ahead'. As mentioned before, most of GB drivers drive manual transmission vehicle. Experience driving skill is required to balance between clutch, brake and gas pedal to drive. Shifting gear transmission can be very tricky for inexperienced driver. There is high probability that the driver is careless while shifting the

gear during above mentioned condition that cause the accident. This result validate the MNL findings previously.
Result above showed that both accident severity is significant to occur at the main carriageway. This factor is a common accident risk since carriageway has busy road network and most of the drivers tend to drive fast resulting with more severe accident.

Slight injury accident showed many significant outcomes in junction location variable. The contributing factors that increase the probability of slight injury accident are 'roundabout', 'junction', 'main road' and 'slip road'. Drivers and pedestrians should alert more when passing these mentioned area for accident prevention. More road signs/warnings should be displayed for extra precaution. Improving traffic light system at the risky area can benefit the road users as a whole.

Analysis proved that hitting object in carriageway is not significant in serious accident, but it influence slight injury accident greatly. This findings is contra from the MNL analysis (Table 4.3).

Another notable variables that associate with both accident severity are 'First Object Hit Off Carriageway' and 'First Point of Impact'. Both accident severity found the same risk factor that cause the accident. Driver's gender reported has significant effect in explaining both accident severity.

Chapter 5

Accident Severity Cost Model

5.1 Introduction

This chapter discusses in depth about the road accident cost model applied in this study. Contrary from the previous statistical models, this chapter focuses on the actuarial method to analyse accident severity in terms of monetary unit. There are three actuarial models employed in analysing accident cost which are Gamma, Weibull and Log-normal distribution model. The elaboration is begin with the source of data and assumptions made for data simulation, parameter estimation method used for each model, explanation of the selected models and lastly the analysis of results. There are limited source of literature reviews discussing accident cost model using actuarial approach. Hence, this study is considered as pioneer in road accident analysis application.

5.2 Source of Data

Below is the scatterplot of simulated cost data obtained for all types accident severity.

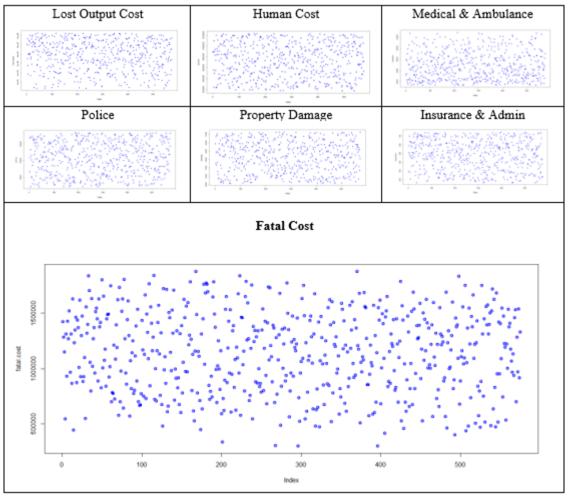


Figure 5.1. Fatal Accident Cost Data

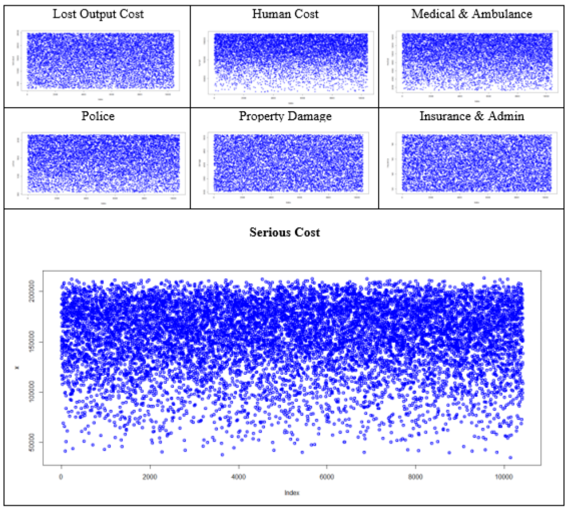


Figure 5.2. Serious Accident Cost Data

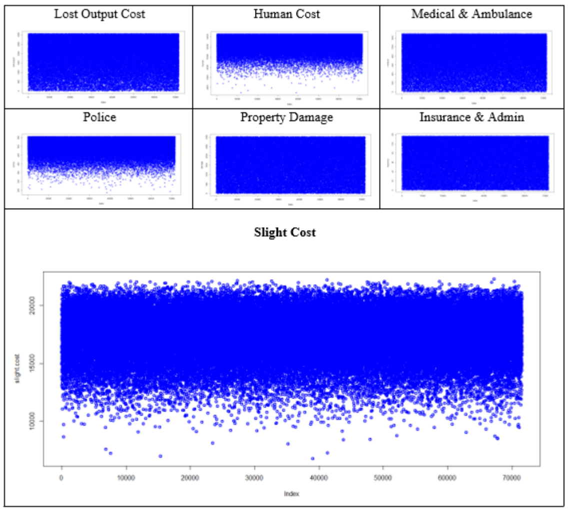


Figure 5.3. Slight Accident Cost Data

5.3 Gamma Distribution

The first accident cost model is Gamma distribution. This model is chosen as there are three suitable distributions to estimate insurance claim cost which are Gamma, Weibull and Lognormal distributions [27], [13]. Due to lack of insurance claim cost data, simulated accident cost data that followed Gamma distribution will be used instead as explained previously.

With limited data information, random walk M-H algorithm is chosen to estimate parameter of α and β in Gamma distribution. It is the same parameter estimation method that used previously in MNLRE but Gamma likelihood is applied. Normal distribution is used for prior to obtain the respected posterior value for α and β. Then, proposal value is compared with the posterior value using M-H algorithm. 50,000 iterations is ran and the first 10,000 is discarded as 'burn-in'. All these parameter values will be used to find the cost ratio between different type of accident severity. The MCMC results is presented as below:

Table 5.1: Summary Value of α and β

Accident severity	α	β
Fatal	0.6985	0.6885
Serious	0.7999	0.7780
Slight	0.7854	0.7786

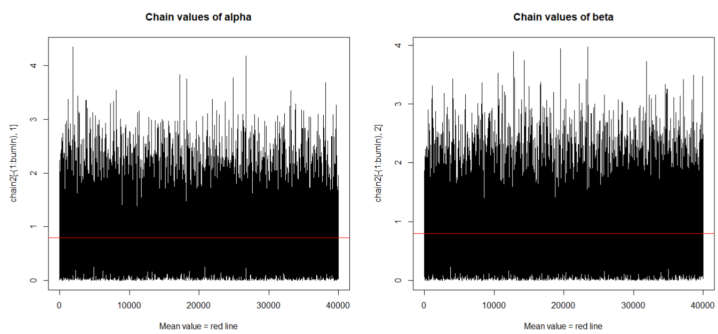


Figure 5.4. Plots of α and β for Fatal Accident

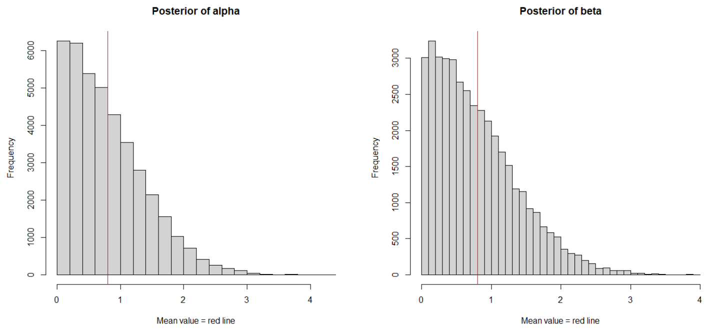


Figure 5.5. Histogram Plots of α and β for Fatal Accident

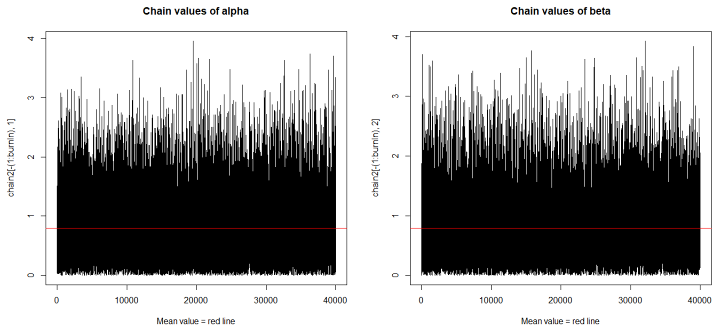


Figure 5.6. Plots of α and β for Serious Accident

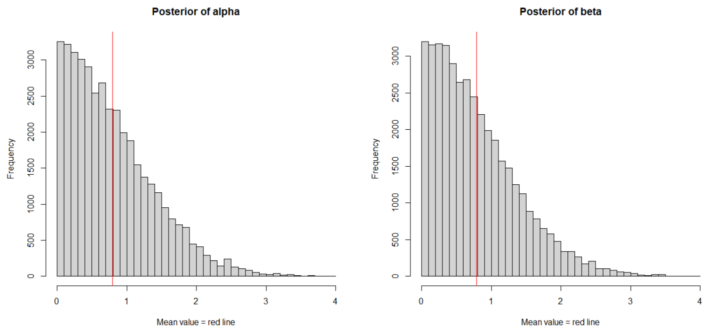


Figure 5.7. Histogram Plots of α and β for Serious Accident

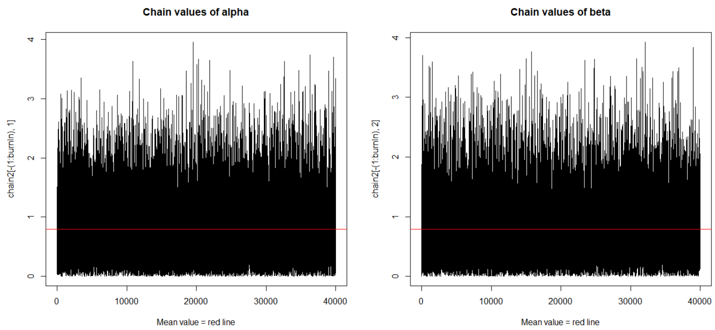


Figure 5.8. Plots of α and β for Slight Accident

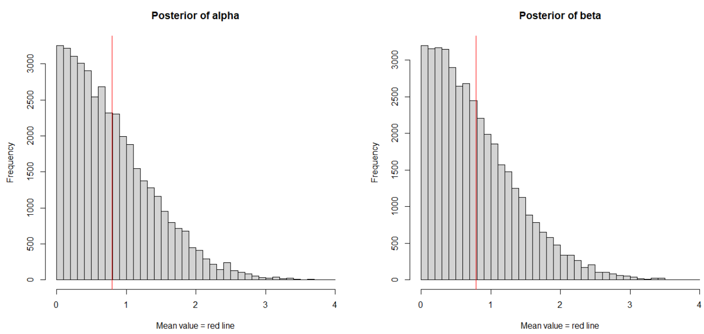


Figure 5.9. Histogram Plots of α and β for Slight Accident

From the parameter value estimated above, the value is not differed much for each type of accident severity. The M-H simulations is done numerous time by adjusting with different prior and proposal distribution. However, this is the best result achieved using the simulated data. The MCMC chain generated is later assess for its convergence fitness.

Table 5.2: MCMC Diagnostic Test of Gamma Distribution Chain

	Root Mean Square Error	Mean of the Sum of the Modulus of the Bias	Monte Carlo Standard Error
Fatal	0.002741	0.000011	0.005013
Serious	0.001934	0.000012	0.004398
Slight	0.001934	0.000011	0.004289

Table 5.3: MCMC Diagnostic Test of Gamma Distribution Chain (Continued)

		Acceptance Rate	Effective Sample Size	Integrated Autocorrelation Time
Fatal	α	0.4789	14008.32	2.7813
	β	0.5010	13668.79	3.0256
Serious	α	0.4876	11332.34	3.0014
	β	0.4888	13172.41	3.1186
Slight	α	0.4876	11332.34	3.0014
	β	0.4888	13172.41	3.1186

Referring to Table 5.1, the parameter values of α and β are almost equal between each type of accident severity. The diagnostic test as above (Table 5.2 and 5.3) also generated the same results. The chain showed a good fit with low value of Root Mean Square Error (RMSE) and Monte Carlo Standard Error (MCSE). Acceptance rate of α and β showed that it converged well with high value of Effective Sample Size (ESS) and low value of Integrated Autocorrelation Time (IAT).

By using the parameter value above, accident cost for each accident severity is estimated (refer Table 5.4. The cost value is obtained by finding the highest probability in Gamma distribution. Again, the results showed not much difference. The cost estimated for fatal accident is slightly higher than serious and slight accident. Therefore, cost ratio of all accident severity can be summarised equal to 1.01:1:1 for Gamma distribution model.

Table 5.4: Summary of Accident Cost: Gamma Distribution

Accident severity	Cost (GBP)
Fatal	9,300
Serious	9,120
Slight	9,120

5.4 Weibull Distribution

The second accident cost model chosen is Weibull distribution. The data set for Weibull distribution is used to find the parameter estimate by applying M-H algorithm in Weibull distribution. After numerous trial, the estimation is failed to obtain. This part will further discuss in Chapter 6.

5.5 Log-normal Distribution

The third accident cost model is Log-normal distribution. The accident cost data for Log-normal is used for estimating parameter of μ and σ in Log-normal distribution via M-H algorithm. For this model, normal distribution is used for prior and proposal. M-H algorithm is ran for 50,000 iterations and the first 10,000 iterations is eliminated because it is considered as unstable.

The simulation is ran various times until the chain generated is converged. The chain result is presented as below:

Table 5.5: Summary Value of μ and σ

Accident severity	μ	σ
Fatal	0.7026	0.7043
Serious	0.7026	0.7043
Slight	0.7026	0.7043

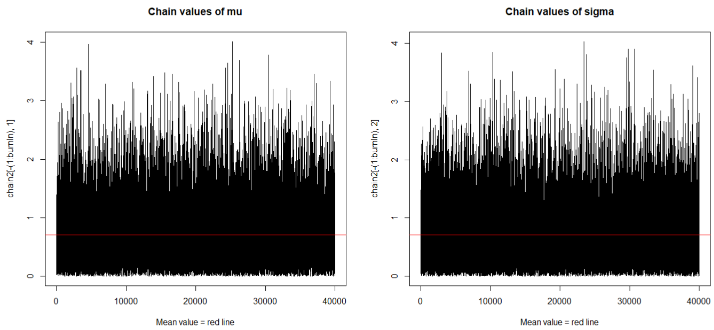


Figure 5.10. Plots of μ and σ for Fatal Accident

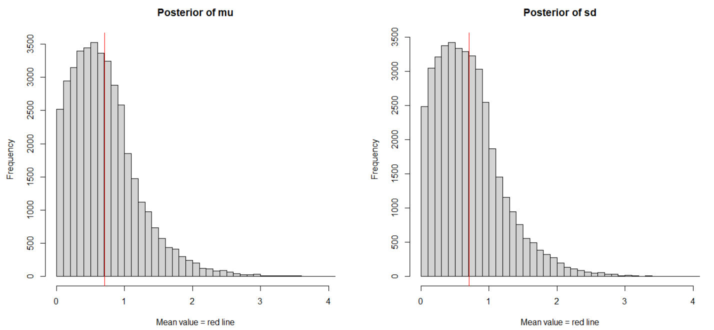


Figure 5.11. Histogram Plots of μ and σ for Fatal Accident

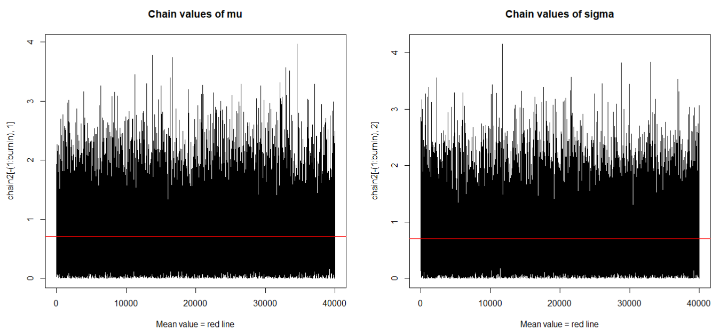


Figure 5.12. Plots of μ and σ for Serious Accident

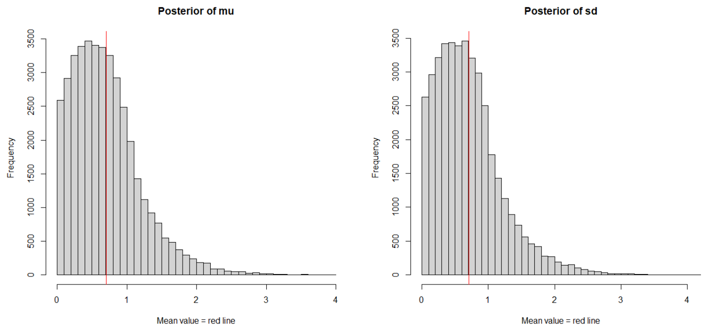


Figure 5.13. Histogram Plots of μ and σ for Serious Accident

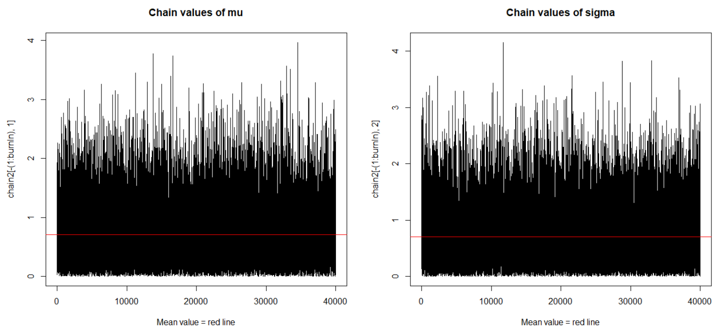


Figure 5.14. Plots of μ and σ for Slight Accident

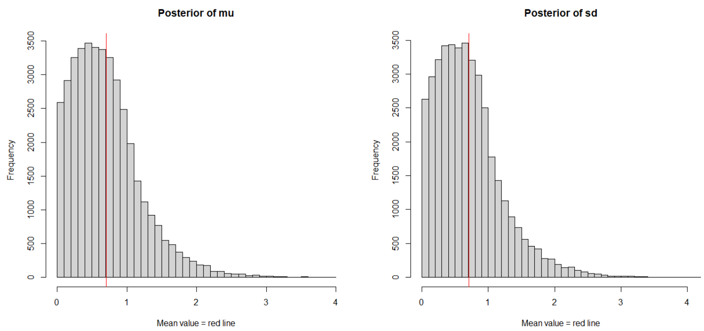


Figure 5.15. Histogram Plots of μ and σ for Slight Accident

From the parameter value estimated above, the value is equal across all accident severity. The simulations is done multiple time applying different prior and proposal distribution. Unfortunately, this is the best result obtained. The chain is later check using different type of diagnostic test to ensure its reliability. The result is shown below:

Table 5.6: MCMC Diagnostic Test of Log-Normal Distribution Chain

	Root Mean Square Error	Mean of the Sum of the Modulus of the bBias	Monte Carlo Standard Error
Fatal	0.000627	0.000004	0.002164
Serious	0.000627	0.000004	0.002164
Slight	0.000627	0.000004	0.002164

Table 5.7: MCMC Diagnostic Test of Log-Normal Distribution Chain (Continued)

		Acceptance Rate	Effective Sample Size	Integrated Autocorrelation Time
Fatal	μ	0.819421	26581.95	1.4741
	σ	0.820046	27848.37	1.4212
Serious	μ	0.819421	26581.95	1.4741
	σ	0.820046	27848.37	1.4212
Slight	μ	0.819421	26581.95	1.4741
	σ	0.820046	27848.37	1.4212

The diagnostic check in Table 5.6 and 5.7 suggested a good fit of result. The generated chain has low value of RMSE and MCSE, high value of acceptance rate and ESS, and also low IAT value.

Table 5.8: Summary of Accident Cost: Log-normal Distribution

Accident severity	Cost (GBP)
Fatal	44,212
Serious	44,212
Slight	44,212

Based on the parameter estimate, accident cost is calculated by finding the highest probability in log-normal model (refer Table 5.8). The cost for fatal, serious and slight injury accident is equal to £44,212 and the cost ratio of all accident severity is 1:1:1 for log-normal distribution model. Comparing these two models, log-normal gave higher cost value compared to Gamma.

Chapter 6

Conclusion & Future Research Recommendation

6.1 Introduction

The two primary objective of this current research has been to investigate the factors that were associated with the road accident severity and the accident cost in GB. This chapter presents a summary of the main results and conclusions obtained from the research. Furthermore, some recommendations based on the findings of this thesis for future research in the field of road safety are discussed.

6.2 Conclusion

6.2.1 Summary for Accident Data

The results in MNL, graphical model and MNLRE model suggested substantial differences in the statistically significant variables when random effects is included. In MNL model, there are 24 significant variables that are associated with serious accident while 27 significant variables accounted for slight injury accident. The correlation check found that the variables are highly correlated and this may affect the performance of the model. Hence, log-linear graphical model is applied to find the cluster information that is important to accident severity. The results showed that 'Speed Limit' and 'Light Conditions' are crucial in explaining severity conditioning that an accident has occurred. This results is not conclusive enough, therefore, MNLRE is introduced to investigate deeper by including random effects factor, which is accident location (Police Force).

Referring to MNL and MNLRE results, the contributing factors that are significant in both models. Slight accident is statistically significant and more likely to happen at First Road Class Type C and Unclassified First Road Class. It indicates that outskirts and abandoned road needs consistent improvement for long-term accident prevention.

Low Speed Limit (30 and 40 mph) is significant in serious accident while higher Speed Limit (50, 60 and 70 mph) is significant in slight accident. As mentioned before, low Speed Limit area is designated for the busy road, therefore more serious accident happen involving vulnerable road users. Slight accident is prone to happen in higher Speed Limit road because driver tend to be careless when they are driving fast. More warning signs should be placed to capture the vulnerable road users attention to be more alert on the surroundings.

MNL and MNLRE findings proved that roundabout is another factor that contribute to road accident occurrence. Drivers has potential to be less cautious when driving at the roundabout. The safety of roundabout with four or more junctions can be improved by installing traffic light at every junction to control the traffic effectively.

Two categories of Weather Conditions affected slight accident greatly, which is raining and snowing without high winds. Prior journey, drivers should be more cautious when driving in bad weather. Hence, there is an urgent need to implement a system to assist drivers regarding traffic flow and weather condition. It can be a mobile application that combine Global Positioning System (GPS) navigator and weather forecast feature. Rather than to open these two apps in two applications separately, this system will plan the driver's journey efficiently with the aim to avoid traffic congestion and road accident.

6.2.2 Summary for Vehicle Data

Vehicle data used two models for the analysis, which are MNL and MNLRE. Graphical model is excluded in vehicle data because R-package for this model has limited data size, thus cannot accommodate the analysis. To recall, the vehicle data has 252,836 observations with 121 variables. Since graphical model is not a major concern in this study, the analysis is proceeded with MNLRE.

When random effects (police force location) is included in the analysis, similar like accident data, both models showed big differences in the results. In MNL model, there are 67 significant variables that are associated with serious accident while 70 significant variables accounted for slight injury accident.

The contributing factors that are significant in both models are discussed by referring MNL (Table 4.2 and 4.3) and MNLRE (Table 4.9, 4.10 and 4.11) output. Serious and slight accident are statistically significant and more likely to happen when it involved pedal cycle, motorcycle with 125 cc and below, car and bus/coach. This can be summarised that types of vehicle is not a major concern as all vehicles has equally likely risk of road accident.

For the Vehicle Maneuver variable, ’waiting to go - held up’, ’slowing or stopping’, ’overtaking’ and ’going ahead’ are significant in serious and slight injury both in MNL and MNLRE. This results provided solid statistical proof that accident is likely to occur during gear transmission. Driver tend to be careless when shifting gear especially during overtaking other vehicle.

MNL and MNLRE gave different result for Vehicle Location, Junction Location, Skidding/Overturning and First Point of Impact’ variables. But, a notable point to highlight is, accident risk is higher and more severe when a vehicle ’entering roundabout’ or ’hitting a lamp post’. Other than that, both models also suggest that skidding/overturning is not a contributing factor that lead to accident event.

Three categories of Vehicle Leaving Carriageway affected both accident severity greatly, which are ’nearside’, ’nearside and rebounded’ and ’straight ahead at junction’. Results from MNL and MNLRE showed that there is no gender difference because both genders are associated in road accident.

The accident location showed great impact to analysis peformed and it would be of interest for future studies to further examine the spatial effect of each category of accident severity.

6.2.3 Summary for Accident Cost

Three accident cost models are used in this study which are Gamma, Weibull and Log-normal. However, only two models can be studied. As discussed previously, this study is lack of accident cost data. Therefore, simulated data is used to proceed for cost modelling.

Gamma model showed that accident cost for fatal is only £9,251 while serious and slight accident cost has same cost, which is £9,149. This results is not as expected as it is very low to represent the actual event.
Log-normal model gave higher cost value, which is £44,212 for all accident severity. This findings also deviated from expectation because each accident severity has different cost value involved.
The analysis of cost accident is not successful due to the data. The simulated data does not represent the actual accident cost. This resulting with inaccurate parameter estimation and cost analysis. The parameter estimation using M-H algorithm in Gamma and Log-normal distribution showed a good fit which means it is applicable to be used with different cost database.

6.3 Future Research Recommendation

6.3.1 Accident Severity Model

The scope of this current research was limited to the analyses of available data from DfT. Due to fund and time costraints, it is impossible to extend this current research by analysing data from other datasets. The data used also for year 2012 only. Therefore, the following issues are recommended for future research:

- Further research for specific type of road accident severity
- Refining this study by focusing on the order of accident severity (e.g: ordinal logit with random effects)
- Improving model specification by extending period of study and adding more variables.
- Exploring spatial effects of accident severity in GB
- Applying different parameter estimation by applying frequentist and other Bayesian method

This present research was limited to road accident data for year 2012, which were not true relative risks and may not be generalisable to the entire spectrum of road accident injuries in GB.

6.3.2 Accident Severity Cost Model

The only source of data for accident cost is the average cost data from DfT. Due to the restrictions on funding and time, it is impossible to obtain actual accident cost data by conducting a study survey. There is also ethical issue to deal with as accident cost is sensitive and confidential topic. Below are the several recommendations for future study:

- Obtaining an actual data for accident cost
- Applying different parameter estimation in Bayesian method
- Improving the predictability of the existing models

Accident cost study will be accurate if the actual data is used. However, this cost study is worth to investigate further in order to compare and validate the results with a real case study.

Bibliography

[1] Abdel-Aty, M. (2003). Analysis of driver injury severity levels at multiple locations using ordered probit models. *Journal of safety research*, 34(5):597–603.

[2] Abdel-Aty, M. and Abdelwahab, H. (2004). Modeling rear-end collisions including the role of driver's visibility and light truck vehicles using a nested logit structure. *Accident Analysis & Prevention*, 36(3):447–456.

[3] Al-Masaeid, H. R., Al-Mashakbeh, A. A., and Qudah, A. M. (1999). Economic costs of traffic accidents in jordan. *Accident Analysis & Prevention*, 31(4):347–357.

[4] Alfaro, J., Chapuis, M., and Fabre, F. (1994). Cost 313. socioeconomic cost of road accidents. *Report EUR*, 15464.

[5] Anastasopoulos, P. C. and Mannering, F. L. (2009). A note on modeling vehicle accident frequencies with random-parameters count models. *Accident Analysis & Prevention*, 41(1):153–159.

[6] Anh, T. T., Dao, N., and Anh, T. (2005). The cost of road traffic accident in vietnam. In *Proceedings of the Eastern Asia Society for Transportation Studies*, volume 5, pages 1923–1933.

[7] Bank, A. D. (1999). *Road Safety Guidelines for the Asian and Pacific Region*.

[8] Bithas, P. S., Sagias, N. C., Tsiftsis, T. A., and Karagiannidis, G. K. (2007). Products and ratios of two gaussian class correlated weibull random variables. In *Proceedings of the 12th International Conference on Applied Stochastic Models and Data Analysis (ASMDA 2007)*. Citeseer.

[9] Carson, J. and Mannering, F. (2001). The effect of ice warning signs on ice-accident frequencies and severities. *Accident Analysis & Prevention*, 33(1):99–109.

[10] Carthy, T., Chilton, S., Covey, J., Hopkins, L., Jones-Lee, M., Loomes, G., Pidgeon, N., and Spencer, A. (1998). On the contingent valuation of safety and the safety of contingent valuation: Part 2-the cv/sg" chained" approach. *Journal of risk and uncertainty*, 17(3):187–214.

[11] Connelly, L. B. and Supangan, R. (2006). The economic costs of road traffic crashes: Australia, states and territories. *Accident Analysis & Prevention*, 38(6):1087–1093.

[12] Darroch, J. N., Lauritzen, S. L., and Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *The Annals of Statistics*, pages 522–539.

[13] Daud, N. and Ibrahim, K. (2009). Hierarchical bayesian approach for ranking of accident blackspots with reference to cost of accidents. In *Advances in Numerical Methods*, pages 173–179. Springer.

[14] Delen, D., Sharda, R., and Bessonov, M. (2006). Identifying significant predictors of injury severity in traffic accidents using a series of artificial neural networks. *Accident Analysis & Prevention*, 38(3):434–444.

[15] Department for Transport (2011). A valuation of road accidents and casualties in greatbritain: Methodology note. Technical report, Department for Transport.

[16] Department for Transport (September 2014). Reported road casualties great britain: 2013 annual report. Technical report, Department for Transport.

[17] Downing, A. (1997). Accident costs in indonesia: A review. In *International Conference on Road Safety, Andhra University, Visakhapatnam, India*.

[18] Duncan, C., Khattak, A., and Council, F. (1998). Applying the ordered probit model to injury severity in truck-passenger car rear-end collisions. *Transportation Research Record: Journal of the Transportation Research Board*, (1635):63–71.

[19] Elvik, R. (1995). An analysis of official economic valuations of traffic accident fatalities in 20 motorized countries. *Accident Analysis & Prevention*, 27(2):237–247.

[20] Elvik, R. (2000). How much do road accidents cost the national economy? *Accident Analysis & Prevention*, 32(6):849–851.

[21] Haddon Jr, W. (1968). The changing approach to the epidemiology, prevention, and amelioration of trauma: the transition to approaches etiologically rather than descriptively based. *American Journal of Public Health and the Nations Health*, 58(8):1431–1438.

[22] Haleem, K. and Abdel-Aty, M. (2010). Examining traffic crash injury severity at unsignalized intersections. *Journal of safety research*, 41(4):347–357.

[23] Hartzel, J., Agresti, A., and Caffo, B. (2001). Multinomial logit random effects models. *Statistical Modelling*, 1(2):81–102.

[24] Hausman, J. A., Hall, B. H., and Griliches, Z. (1984). Econometric models for count data with an application to the patents-r&d relationship.

[25] Hosseinpour, M., Yahaya, A. S., and Sadullah, A. F. (2014). Exploring the effects of roadway characteristics on the frequency and severity of head-on crashes: Case studies from malaysian federal roads. *Accident Analysis & Prevention*, 62:209–222.

[26] Huang, H., Chin, H. C., and Haque, M. M. (2008). Severity of driver injury and vehicle damage in traffic crashes at intersections: a bayesian hierarchical analysis. *Accident Analysis & Prevention*, 40(1):45–54.

[27] Ibrahim, N. D. . K. (2007). Taburan nisbah kos kemalangan maut dan cedera parah. In *International Statistics Conference Proceeding. Kuala Lumpur*.

[28] Islam, S. and Mannering, F. (2006). Driver aging and its effect on male and female single-vehicle accident injuries: Some additional evidence. *Journal of Safety Research*, 37(3):267–276.

[29] Jones, B., Janssen, L., and Mannering, F. (1991). Analysis of the frequency and duration of freeway accidents in seattle. *Accident Analysis & Prevention*, 23(4):239–255.

[30] Khorashadi, A., Niemeier, D., Shankar, V., and Mannering, F. (2005). Differences in rural and urban driver-injury severities in accidents involving large-trucks: an exploratory analysis. *Accident Analysis & Prevention*, 37(5):910–921.

[31] Kim, J.-K., Kim, S., Ulfarsson, G. F., and Porrello, L. A. (2007). Bicyclist injury severities in bicycle–motor vehicle accidents. *Accident Analysis & Prevention*, 39(2):238–251.

[32] Kockelman, K. M. and Kweon, Y.-J. (2002). Driver injury severity: an application of ordered probit models. *Accident Analysis & Prevention*, 34(3):313–321.

[33] Kononen, D. W., Flannagan, C. A., and Wang, S. C. (2011). Identification and validation of a logistic regression model for predicting serious injuries associated with motor vehicle crashes. *Accident Analysis & Prevention*, 43(1):112–122.

[34] Lauritzen, S. L. (1996). *Graphical models*. Oxford University Press.

[35] Lauritzen, S. L. and Wermuth, N. (1989). Graphical models for associations between variables, some of which are qualitative and some quantitative. *The Annals of Statistics*, pages 31–57.

[36] Lee, J. and Mannering, F. (2002). Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. *Accident Analysis & Prevention*, 34(2):149–161.

[37] Lord, D. and Mahlawat, M. (2009). Examining application of aggregated and disaggregated poisson-gamma models subjected to low sample mean bias. *Transportation Research Record: Journal of the Transportation Research Board*, 2136(1):1–10.

[38] Lord, D. and Miranda-Moreno, L. F. (2008). Effects of low sample mean values and small sample size on the estimation of the fixed dispersion parameter of poisson-gamma models for modeling motor vehicle crashes: a bayesian perspective. *Safety Science*, 46(5):751–770.

[39] Ma, J. and Kockelman, K. M. (2006). Bayesian multivariate poisson regression for models of injury count, by severity. *Transportation Research Record: Journal of the Transportation Research Board*, 1950(1):24–34.

[40] Maher, M. J. and Summersgill, I. (1996). A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis & Prevention*, 28(3):281–296.

[41] Malyshkina, N. and Mannering, F. (2008). Effect of increases in speed limits on severities of injuries in accidents. *Transportation Research Record: Journal of the Transportation Research Board*, (2083):122–127.

[42] Malyshkina, N. V. and Mannering, F. L. (2010). Empirical assessment of the impact of highway design exceptions on the frequency and severity of vehicle accidents. *Accident Analysis & Prevention*, 42(1):131–139.

[43] Malyshkina, N. V., Mannering, F. L., and Tarko, A. P. (2009). Markov switching negative binomial models: an application to vehicle accident frequencies. *Accident Analysis & Prevention*, 41(2):217–226.

[44] Miaou, S.-P. (1994). The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis & Prevention*, 26(4):471–482.

[45] Miaou, S.-P., Bligh, R. P., and Lord, D. (2005). Part 1: Roadside safety design: Developing guidelines for median barrier installation: Benefit-cost analysis with texas data. *Transportation Research Record: Journal of the Transportation Research Board*, 1904(1):2–19.

[46] Miaou, S.-P. and Lum, H. (1993). Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis & Prevention*, 25(6):689–709.

[47] Minsker, S., Srivastava, S., Lin, L., and Dunson, D. (2014). Scalable and robust bayesian inference via the median posterior. In *International conference on machine learning*, pages 1656–1664.

[48] Moudon, A. V., Lin, L., Jiao, J., Hurvitz, P., and Reeves, P. (2011). The risk of pedestrian injury and fatality in collisions with motor vehicles, a social ecological study of state routes and city streets in king county, washington. *Accident Analysis & Prevention*, 43(1):11–24.

[49] Neiswanger, W., Wang, C., and Xing, E. P. (2014). Asymptotically exact, embarrassingly parallel mcmc. In *Proceedings of the Thirtieth Conference on Uncertainty in Artificial Intelligence*, pages 623–632. AUAI Press.

[50] O'donnell, C. and Connor, D. (1996). Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis & Prevention*, 28(6):739–753.

[51] Oh, J., Washington, S. P., and Nam, D. (2006). Accident prediction model for railway-highway interfaces. *Accident Analysis & Prevention*, 38(2):346–356.

[52] O'REILLY, D. (1992). Costing road traffic accidents: the value of lost output. Technical report, TRL Working Paper, Crowthorne.

[53] Organization, W. H. (2009). *Global status report on road safety: time for action*. World Health Organization.

[54] Organization, W. H. et al. (2018). Global status report on road safety 2018: Summary. Technical report, World Health Organization.

[55] Park, B.-J. and Lord, D. (2009). Application of finite mixture models for vehicle crash data analysis. *Accident Analysis & Prevention*, 41(4):683–691.

[56] Park, B.-J., Lord, D., and Hart, J. D. (2010). Bias properties of bayesian statistics in finite mixture of negative binomial regression models in crash data analysis. *Accident Analysis & Prevention*, 42(2):741–749.

[57] Peden, M. et al. (2004). World report on road traffic injury prevention.

[58] Peden, M. and Sminkey, L. (2004). World health organization dedicates world health day to road safety. *Injury prevention*, 10(2):67–67.

[59] Peek-Asa, C., Britton, C., Young, T., Pawlovich, M., and Falb, S. (2010). Teenage driver crash incidence and factors influencing crash injury by rurality. *Journal of safety research*, 41(6):487–492.

[60] Quddus, M. A., Noland, R. B., and Chin, H. C. (2002). An analysis of motorcycle injury and vehicle damage severity using ordered probit models. *Journal of safety research*, 33(4):445–462.

[61] Quddus, M. A., Wang, C., and Ison, S. G. (2009). Road traffic congestion and crash severity: econometric analysis using ordered response models. *Journal of Transportation Engineering*, 136(5):424–435.

[62] Rifaat, S. M. and Tay, R. (2009). Effects of street patterns on injury risks in two-vehicle crashes. *Transportation Research Record*, 2102(1):61–67.

[63] Rifaat, S. M., Tay, R., and de Barros, A. (2011). Effect of street pattern on the severity of crashes involving vulnerable road users. *Accident Analysis & Prevention*, 43(1):276–283.

[64] Savolainen, P. T., Mannering, F. L., Lord, D., and Quddus, M. A. (2011). The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accident Analysis & Prevention*, 43(5):1666–1676.

[65] Schneider IV, W., Savolainen, P., and Zimmerman, K. (2009). Driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways. *Transportation Research Record: Journal of the Transportation Research Board*, (2102):85–92.

[66] Schoeters, A., W. W. C. L. W. W. E. R. J. H. V. B. W. F. A. and Daniels, S. (2017). Costs related to serious injuries. Technical report, D7.3 of the H2020, project SafetyCube.

[67] Scott, S. L., Blocker, A. W., Bonassi, F. V., Chipman, H. A., George, E. I., and McCulloch, R. E. (2016). Bayes and big data: The consensus monte carlo algorithm. *International Journal of Management Science and Engineering Management*, 11(2):78–88.

[68] Shankar, V. and Mannering, F. (1996). An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *Journal of Safety Research*, 27(3):183–194.

[69] Shankar, V., Ulfarsson, G. F., and Pendyala, R. (2003). og mb nebergall, 2003: Modeling crashes involving pedestrians and motorized traffic. *Safety Science*, 41(7):627–640.

[70] Silcock, R. (2003). Guidelines for estimating the cost of road crashes in developing countries. *London, Department for International Development Project*, 7780:2003.

[71] Simpson, H. F. and O'REILLY, D. M. (1994). Revaluation of the accident related costs of road accidents. *TRL project Report*, (PR 56).

[72] Sleet, D. A., Dinh-Zarr, T. B., and Dellinger, A. M. (2007). Traffic safety in the context of public health and medicine. *Improving Traffic Safety Culture in the United States*, page 41.

[73] Trawen, A., Maraste, P., and Persson, U. (2002). International comparison of costs of a fatal casualty of road accidents in 1990 and 1999. *Accident Analysis & Prevention*, 34(3):323–332.

[74] Tunaru, R. (1999). *Statistical modelling of road accident data via graphical models and hierarchical Bayesian models.* PhD thesis, Middlesex University.

[75] White, S., Kypraios, T., and Preston, S. P. (2015). Piecewise approximate bayesian computation: fast inference for discretely observed markov models using a factorised posterior distribution. *Statistics and computing*, 25(2):289–301.

[76] Whittaker, J. (2009). *Graphical models in applied multivariate statistics*. Wiley Publishing.

[77] Wijnen, W., Weijermars, W., Schoeters, A., Van den Berghe, W., Bauer, R., Carnis, L., Elvik, R., and Martensen, H. (2019). An analysis of official road crash cost estimates in european countries. *Safety science*, 113:318–327.

[78] Wijnen, W., Weijermars, W., Van den Berghe, W., Schoeters, A., Bauer, R., Carnis, L., Elvik, R., Theofilatos, A., Filtness, A. J., Reed, S., et al. (2017). Crash cost estimates for european countries, deliverable 3.2 of the h2020 project safetycube. *Loughborough: Loughborough University, SafetyCube*.

[79] Wu, Q., Zhang, G., Zhu, X., Liu, X. C., and Tarefder, R. (2016). Analysis of driver injury severity in single-vehicle crashes on rural and urban roadways. *Accident Analysis & Prevention*, 94:35–45.

[80] Xie, Y., Zhang, Y., and Liang, F. (2009). Crash injury severity analysis using bayesian ordered probit models. *Journal of Transportation Engineering*, 135(1):18–25.

[81] Ye, F. and Lord, D. (2014). Comparing three commonly used crash severity models on sample size requirements: multinomial logit, ordered probit and mixed logit models. *Analytic methods in accident research*, 1:72–85.

Appendix

Variables of Accident Data

Accident Table

Accident Index Number (13 chars)
Accident Year (YYYY) + Accident Ref (9 chars) *1.3*

Police Force Code *1.2*
For codes see page 11

Accident Severity
1 Fatal
2 Serious
3 Slight

Number of Vehicles *1.5*

Number of Casualties *1.6*

Date *1.7*
Accident Day

Month
1 January
2 February
3 March
4 April
5 May
6 June
7 July
8 August
9 September
10 October
11 November
12 December

Year

Day of Week
1 Sunday
2 Monday
3 Tuesday
4 Wednesday
5 Thursday
6 Friday
7 Saturday

Time of Day (HHMM) *1.9*
Hour of Accident (24 hour)
Minute of Accident

Local Authority (Borough/District) *1.10*
For codes see pages 12-15

Location (OSGR) *1.11*
Easting
Northing

1st Road Class *1.12*
1 Motorway
2 A(M)
3 A
4 B
5 C
6 Unclassified

1st Road Number *1.13*

Road Type *1.14*
1 Roundabout
2 One way street (from 2005)
3 Dual carriageway
6 Single carriageway
7 Slip road (from 2005)
9 Unknown
12 One way street/Slip road (1979-2004)

Speed Limit (permanent) *1.15*
MPH

Junction Detail *1.16*
0 Not at junction or within 20 metres
1 Roundabout
2 Mini-roundabout
3 T or staggered junction
5 Slip road
6 Crossroads
7 More than 4 arms (not roundabout)
8 Private drive or entrance
9 Other junction

Junction Control *1.17*
1 Authorised person
2 Auto traffic signal
3 Stop sign
4 Give way or uncontrolled

2nd Road Class *1.18*
1 Motorway
2 A(M)
3 A
4 B
5 C
6 Unclassified

2nd Road Number *1.19*

Pedestrian Crossing - Human Control *1.20A*
0 None within 50 metres
1 Control by school crossing patrol
2 Control by other authorised person

Pedestrian Crossing - Physical Facilities *1.20B*
0 No physical crossing facilities within 50 metres
1 Zebra
4 Pelican, puffin, toucan or similar non-junction pedestrian light crossing
5 Pedestrian phase at traffic signal junction
7 Footbridge or subway
8 Central refuge

Light Conditions *1.21*
1 Daylight
4 Darkness - lights lit
5 Darkness - lights unlit
6 Darkness - no lighting
7 Darkness - lighting unknown

Weather Conditions *1.22*
1 Fine no high winds
2 Raining no high winds
3 Snowing no high winds
4 Fine + high winds
5 Raining + high winds
6 Snowing + high winds
7 Fog or mist
8 Other
9 Unknown

Road Surface Conditions *1.23*
1 Dry
2 Wet or damp
3 Snow
4 Frost or ice
5 Flood over 3cm. deep
6 Oil or diesel (1999-2004)
7 Mud (1999-2004)

Special Conditions at Site *1.24*
0 None
1 Auto traffic signal - out
2 Auto signal part defective
3 Road sign or marking defective or obscured
4 Roadworks
5 Road surface defective
6 Oil or diesel (from 2005 - see A23)
7 Mud (from 2005 - see A23)

Carriageway Hazards *1.25*
0 None
1 Vehicle load on road
2 Other object on road
3 Previous accident
4 Dog on road (1979-2004)
5 Other animal on road (1979-2004)
6 Pedestrian in carriageway - not injured (from 2005)
7 Any animal in carriageway (except ridden horse) (from 2005)

Did a Police Officer Attend the Scene and Obtain the Details for this Report
(from 1999) *1.26*
1 Yes
2 No
3 No - accident was reported using a self completion form

Variables of Vehicle Data

Vehicle Table

Accident Index Number (13 chars)

Vehicle Reference Number *2.4*

	Vehicle Type *2.5*
1	Pedal cycle
2	Motorcycle 50cc and under
3	M/cycle over 50 and up to 125cc (from 1999)
4	Motorcycle over 125cc and up to 500cc (from 2005)
5	Motorcycle over 500cc (from 2005)
23	Electric motorcycle (from 2011)
97	Motorcycle - unknown cc (from 2011)
103	Motorcycle - Scooter (1979-1998)
104	Motorcycle (1979-1998)
105	Motorcycle - Combination (1979-1998)
106	Motorcycle over 125cc (1999-2004)
8	Taxi / Private hire car (from 2005)
108	Taxi (excluding private hire cars) (1979-2004)
9	Car (from 2005)
109	Car (including private hire cars) (1979-2004)
10	Minibus (8 - 16 passenger seats) (from 1999)
110	Minibus/Motor caravan (1979-1998)
11	Bus or coach (17 or more pass seats)
19	Van / Goods 3.5 tonnes mgw or under
20	Goods over 3.5t. and under 7.5t. (from 1999)
21	Goods 7.5 tonnes mgw and over (from 1999)
98	Goods vehicle - unknown weight (self rep only)
113	Goods over 3.5 tonnes (1979-1998)
16	Ridden horse (from 1999)
17	Agricultural vehicle (from 1999)
18	Tram (from 1999)
22	Mobility scooter (from 2011)
90	Other vehicle

	Towing and Articulation *2.6*
0	No tow/articulation
1	Articulated vehicle
2	Double or multiple trailer
3	Caravan
4	Single trailer
5	Other tow

	Vehicle Manoeuvre *2.7*
1	Reversing
2	Parked
3	Waiting to go - held up
4	Slowing or stopping
5	Moving off
6	U-turn
7	Turning left
8	Waiting to turn left
9	Turning right
10	Waiting to turn right
11	Changing lane to left
12	Changing lane to right
13	Overtaking moving vehicle - offside
14	Overtaking static vehicle - offside
15	Overtaking - nearside
16	Going ahead left-hand bend
17	Going ahead right-hand bend
18	Going ahead other

Vehicle Movement Compass Point *2.8*

	Direction - From **Direction - To**
1	N
2	NE
3	E
4	SE
5	S
6	SW
7	W
8	NW
M	Undefined
00	Parked

	Vehicle Location - Restricted Lane/ Away from Main Carriageway *2.9*
0	On main c'way - not in restricted lane (from 1999)
1	Tram/Light rail track (from 1999)
2	Bus lane (from 1999)
3	Busway (including guided busway) (from 1999)
4	Cycle lane (on main carriageway) (from 1999)
5	Cycleway or shared use footway
6	On lay-by or hard shoulder
7	Entering lay-by or hard shoulder
8	Leaving lay-by or hard shoulder
9	Footway (pavement) (from 1999)
10	Not on carriageway (1979-1998)

	Junction Location of Vehicle (from 2005) *2.10*
0	Not at, or within 20 metres of, junction
1	Approaching junct or waiting/parked at junct apprch
2	Cleared junction or waiting/parked at junction exit
3	Leaving roundabout
4	Entering roundabout
5	Leaving main road
6	Entering main road
7	Entering from slip road
8	Mid Junction - on roundabout or on main road

	Skidding / Overturning *2.11*
0	None
1	Skidded
2	Skidded and overturned
3	Jackknifed
4	Jackknifed and overturned
5	Overturned

	Hit Object In Carriageway *2.12*
0	None
1	Previous accident
2	Road works
4	Parked vehicle
5	Bridge (roof)
6	Bridge (side)
7	Bollard or refuge
8	Open door of vehicle
9	Central island of roundabout
10	Kerb
11	Other object
12	Any animal (except ridden horse) (from 2005)

	Vehicle Leaving Carriageway *2.13*
0	Did not leave carriageway
1	Nearside
2	Nearside and rebounded
3	Straight ahead at junction
4	Offside on to central reservation
5	Offside on to centrl res + rebounded
6	Offside - crossed central reservation
7	Offside
8	Offside and rebounded

	First Object Hit Off Carriageway *2.14*
0	None
1	Road sign or traffic signal
2	Lamp post
3	Telegraph or electricity pole
4	Tree
5	Bus stop or bus shelter
6	Central crash barrier
7	Near/Offside crash barrier
8	Submerged in water
9	Entered ditch
10	Other permanent object
11	Wall or fence (from 2011)

	1st Point of Impact *2.16*
0	Did not impact
1	Front
2	Back
3	Offside
4	Nearside

	Sex of Driver *2.21*
1	Male
2	Female
3	Not known

Age of Driver *2.22*

	Hit and Run *2.24*
0	Other
1	Hit and run
2	Non-stop vehicle not hit

	Was Vehicle Left Hand Drive? (from 2005) *2.35*
1	No
2	Yes

	Journey Purpose of Driver/Rider (from 2005) *2.29*
1	Journey as part of work
2	Commuting to/from work
3	Taking pupil to/from school
4	Pupil riding to/from school
5	Other (from 2011)
6	Not known (from 2011)
15	Other/Not known (2005-2010)

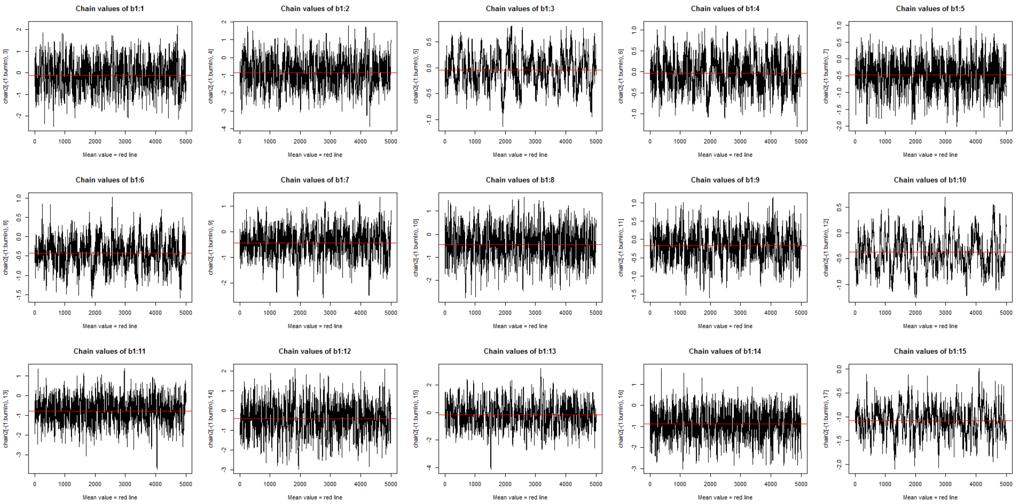
Police Force Code

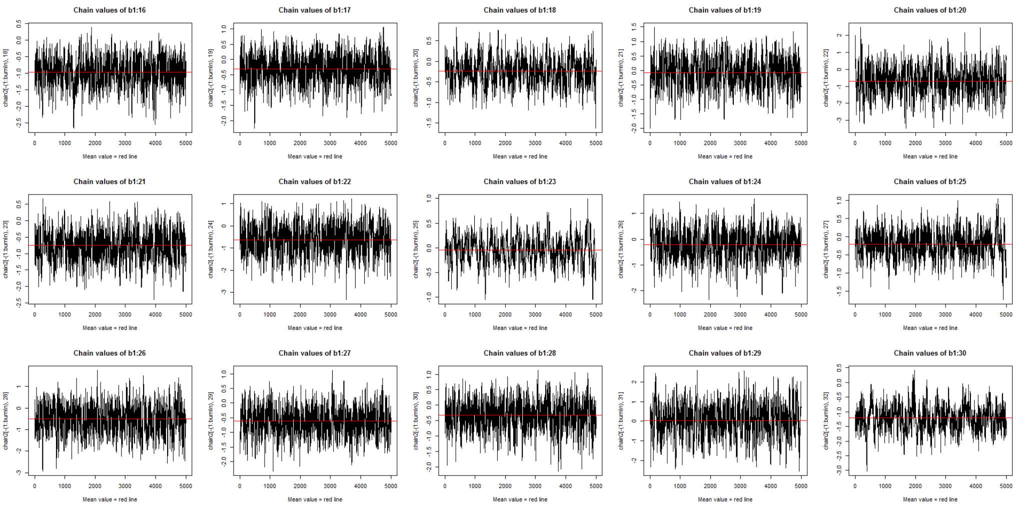
Value Codes

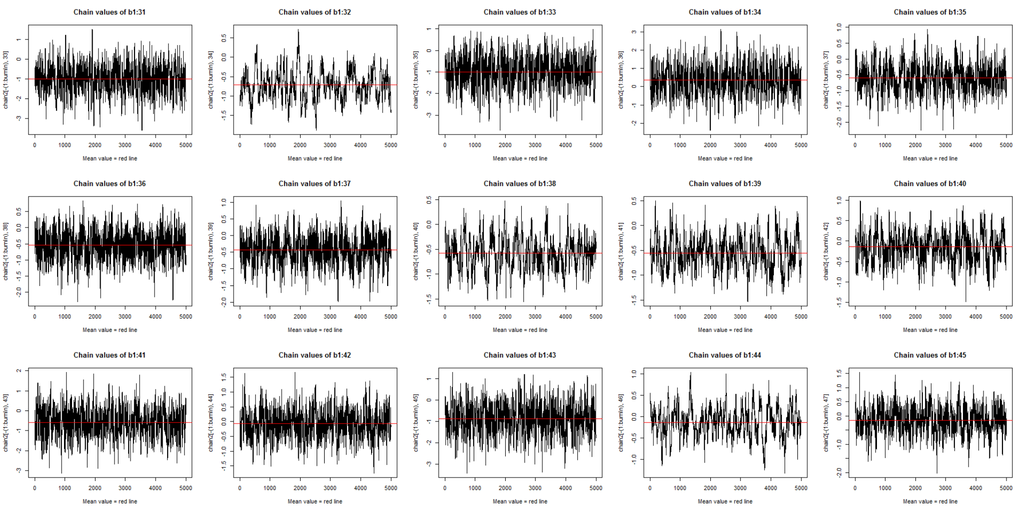
	Police Force Code
1	Metropolitan Police*
3	Cumbria
4	Lancashire
5	Merseyside
6	Greater Manchester
7	Cheshire
10	Northumbria
11	Durham
12	North Yorkshire
13	West Yorkshire
14	South Yorkshire
16	Humberside
17	Cleveland
20	West Midlands
21	Staffordshire
22	West Mercia
23	Warwickshire
30	Derbyshire
31	Nottinghamshire
32	Lincolnshire
33	Leicestershire
34	Northamptonshire
35	Cambridgeshire
36	Norfolk
37	Suffolk
40	Bedfordshire
41	Hertfordshire*
42	Essex*
43	Thames Valley
44	Hampshire
45	Surrey*
46	Kent
47	Sussex
48	City of London
50	Devon and Cornwall
52	Avon and Somerset
53	Gloucestershire
54	Wiltshire
55	Dorset
60	North Wales
61	Gwent
62	South Wales
63	Dyfed-Powys
91	Northern
92	Grampian
93	Tayside
94	Fife
95	Lothian and Borders
96	Central
97	Strathclyde
98	Dumfries and Galloway

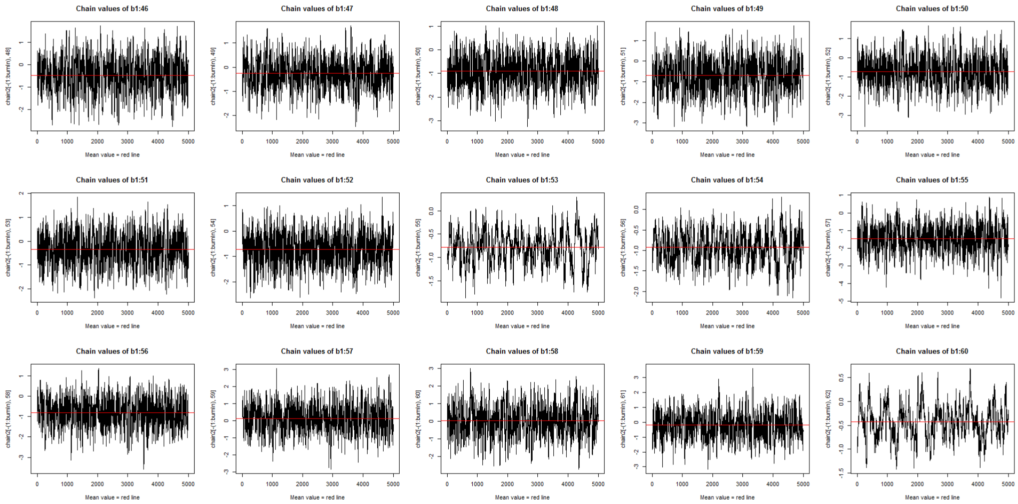
* - Note - Until 2000, Metropolitan Police patrolled areas in Herts, Essex and Surrey. In these cases the local authority code relates to the area patrolled but the Police Force Code is "1" - Metropolitan.

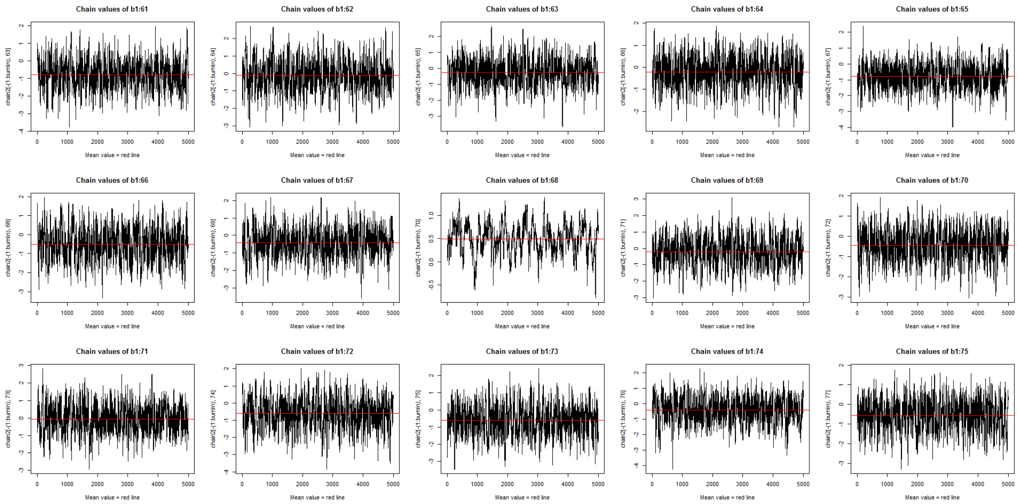
Plots for β_1 chains



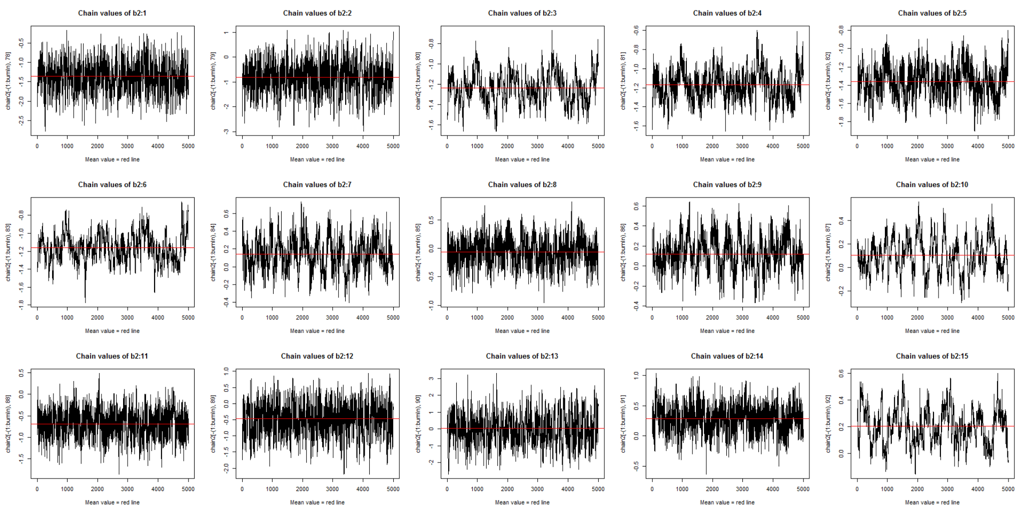


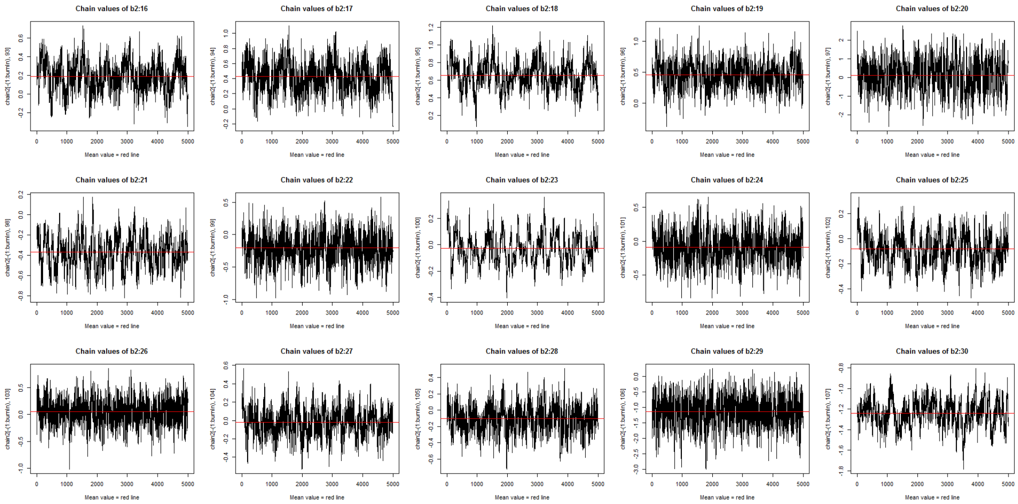


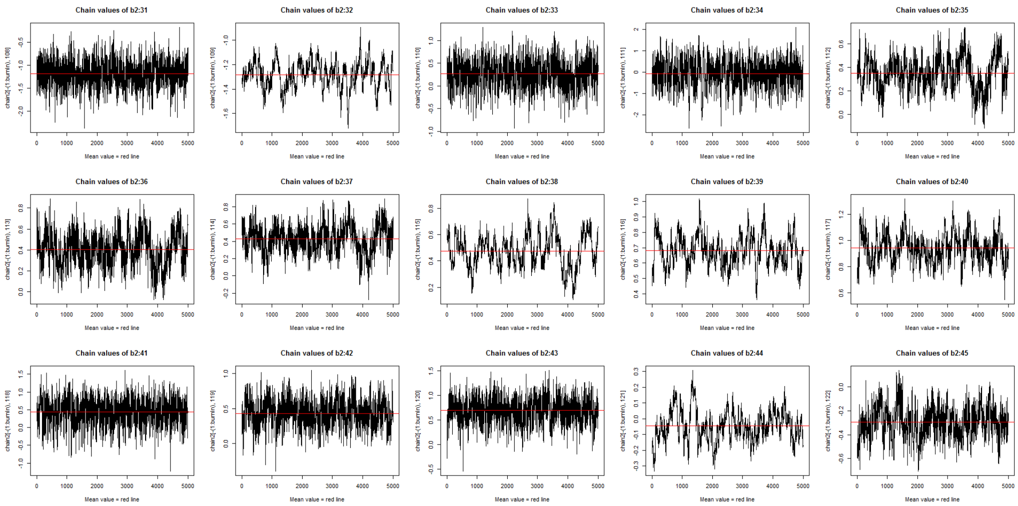


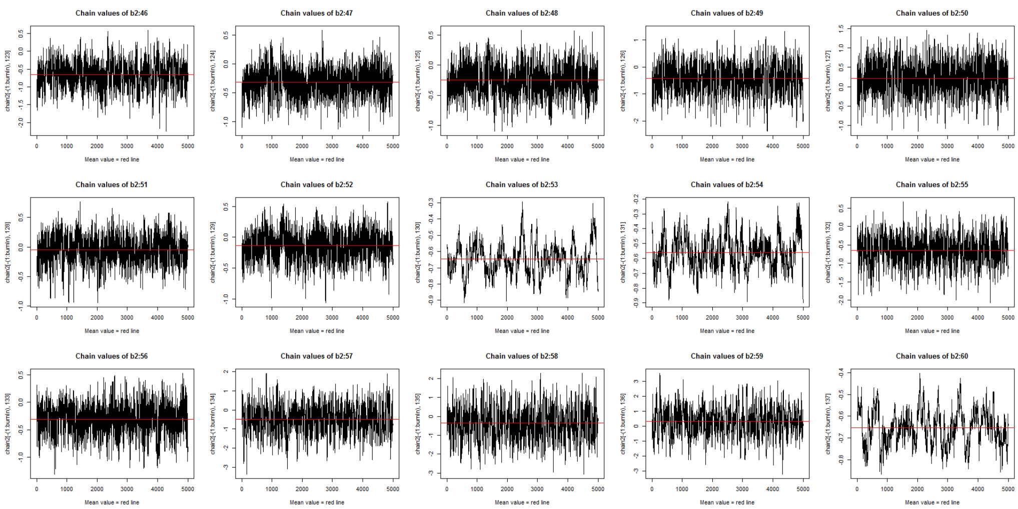


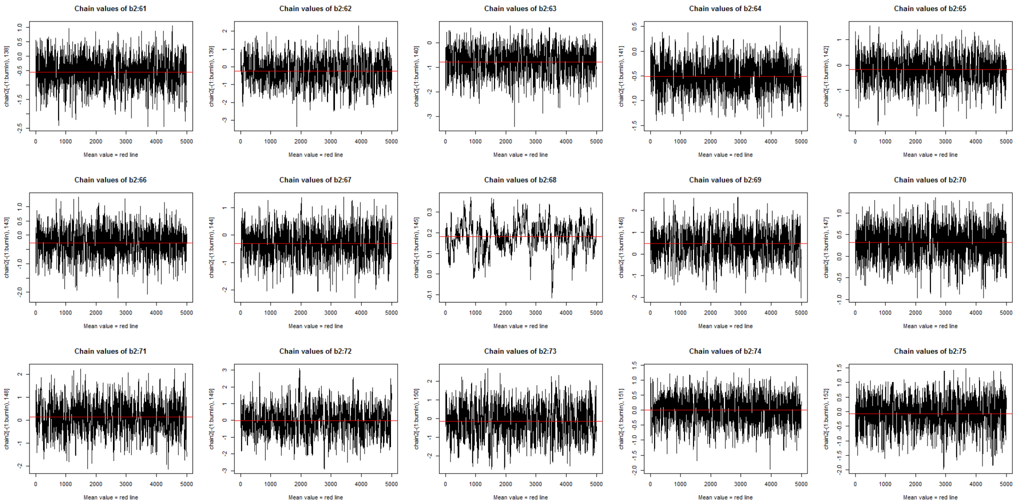
Plots for β_2 chains




Chain values of b2:16
Chain values of b2:17
Chain values of b2:18
Chain values of b2:19
Chain values of b2:20
Chain values of b2:21
Chain values of b2:22
Chain values of b2:23
Chain values of b2:24
Chain values of b2:25
Chain values of b2:26
Chain values of b2:27
Chain values of b2:28
Chain values of b2:29
Chain values of b2:30


Chain values of b2:31
Chain values of b2:32
Chain values of b2:33
Chain values of b2:34
Chain values of b2:35
Chain values of b2:36
Chain values of b2:37
Chain values of b2:38
Chain values of b2:39
Chain values of b2:40
Chain values of b2:41
Chain values of b2:42
Chain values of b2:43
Chain values of b2:44
Chain values of b2:45


Chain values of b2:46
Chain values of b2:47
Chain values of b2:48
Chain values of b2:49
Chain values of b2:50
Chain values of b2:51
Chain values of b2:52
Chain values of b2:53
Chain values of b2:54
Chain values of b2:55
Chain values of b2:56
Chain values of b2:57
Chain values of b2:58
Chain values of b2:59
Chain values of b2:60


Chain values of b2:61
Chain values of b2:62
Chain values of b2:63
Chain values of b2:64
Chain values of b2:65
Chain values of b2:66
Chain values of b2:67
Chain values of b2:68
Chain values of b2:69
Chain values of b2:70
Chain values of b2:71
Chain values of b2:72
Chain values of b2:73
Chain values of b2:74
Chain values of b2:75
Mean value = red line