KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLGY, KUMASI

DEPARTMENT OF MATHEMATICS

FACULTY OF PHYSICAL SCIENCES

COLLEGE OF SCIENCE

A LOGISTIC REGRESSION MODEL FOR MOTOR INSURANCE CLAIMS IN GHANA: A CASE STUDY OF SIC INSURANCE COMPANY LIMITED (SIC)

BY

BASHIRU IMORO IBN SAEED

MARCH, 2009

A LOGISTIC REGRESSION MODEL FOR MOTOR INSURANCE CLAIMS IN GHANA: A CASE STUDY OF SIC INSURANCE COMPANY LIMITTED(SIC)

BY

BASHIRU IMORO IBN SAEED BSc. MATHEMATICS

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS, KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF

MASTER OF SCIENCE

FACULTY OF PHYSICAL SCIENCES,

COLLEGE OF SCIENCE

MARCH, 2009

DECLARATION

I hereby declare that this submission is my own work towards the M.Sc. and that, to the best of my knowledge, it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University, except where due acknowledgement has been made in the text.

Bashiru Imoro Ibn Saeed (8864205)	---------------	04-09-09
(Student/ID)	**(Signature)**	**(Date)**
Certified by:		
Dr. F.T Oduro	-----------	11/11/09
(Supervisor)	**(Signature)**	**(Date)**
S.K. Appiah	---------------	11/09/09
For:		
(Supervisor)	**(Signature)**	**(Date)**
Dr. S.K. Amponsah	---------------	28/09/09
(Head of Department)	**(Signature)**	**(Date)**

ABSTRACT

Motor insurance is only the insurance which is compulsory by law. The National Insurance Commission (NIC) which has the responsibility for the insurance industry currently supervises about 21 insurance companies, 2 reinsurance companies, 37 insurance brokerage companies and about 4,500 insurance agents. This number is currently not enough for the ever growing population of the nation. With the collapse of some insurance companies, there has been the problem of how to handle most of the claims. Also the mere fact that industries in Ghana do not put emphasis on some of the risk factors in estimating motor insurance claims would give room for unfairness in making claim.

It is in line with the foregoing claim problems of motor insurance that this study is designed to consider and critically examine some pertinent characteristics (called here risk factors) of the an insured vehicle attributing to motor claims. This was achieved by establishing a logistic regression model using these risk factors: sex of the policyholder, make of vehicle, age of vehicle at point of entry, and cubic capacity of vehicle.

The data were obtained from the Motor Claims Department of SIC Insurance Company Limited (SIC) for the period, 2007 – 2008. The results from the various analyses revealed that only the two risk factors: the vehicle make (FIAT, NISSAN, TOYOTA, BMW), and vehicles aged at least 6 years at point of entry contribute significantly to the prediction of a motor insurance claim. The odds ratios analysis also showed that these makes of vehicle make more claims compared to OPEL while all the relatively old insured vehicles are less likely to make a claim compared to insured vehicles not more than five years old.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

ACKNOWLEDGEMENT

My appreciation first and foremost goes to the Almighty Allah for His grace, wisdom, revelation and understanding to produce this work.

I am most grateful to my Thesis Advisors, Dr. F.T. Oduro and Mr. S.K. Appiah, for their critical guidance, extremely valuable advice, and encouragement in time of difficulties. Indeed, they have introduced me to exciting area of research and also provided me with their brilliant ideas.

My sincere thanks to Mr. Samuel Duah Oduro for reviewing the first draft of this dissertation very carefully. I would also like to thank him for the indispensable guidance as far as the computations are concerned. It would have been impossible to complete this dissertation without the help of Mr. Paul Quarshie and his outfit of the Claims Department of the SIC Insurance Company Limited (SIC), who kindly assisted me for needed data and other information. They were really of immense help to me.

I would also like to thank Mr. Nakua Emmanuel for his valuable comments in the initial draft of the thesis. Finally, I would like to express my gratitude to my beloved wife Ayisha Mohammed who has been the greatest inspiration in my life and to my Mother Zainab Alhassan for showing me the way.

CHAPTER ONE

INTRODUCTION

1.1 BACKGROUND

The motor insurance business dates back to Ghana's colonial past. Prior to country's independence in 1957, the motor insurance sector was largely dominated by British insurance companies which introduced strict and professional underwriting standards. Following the withdrawal of the colonial insurers about 50 years ago and with little or no official supervision of the local market insurers at that time, the entire industry struggled to successfully market their products. In the process, public confidence in motor insurance, in particular, was severely affected by opportunistic companies that entered the market with a strategy of low premiums and poor, or non-existent claims service (Amoo, 2002). Several Government enquiries into the insurance industry in the 1980's led to the establishment of the National Insurance Commission (NIC) in the early 90's. The NIC was empowered to protect Policyholders' and Third Parties' rights under insurance contracts and to ensure that insurance companies were financially sound and conducted their business in a professional manner.

Since its creation more than a decade ago, the NIC has taken a number of bold initiatives that have significantly restored public confidence in the insurance industry, particularly in the motor vehicle insurance sector. Now the motor insurance business accounts for about 50% of the premium income of Ghana's insurance industry. Since certain level of motor insurance cover is compulsory, it is the most popular and, regrettably, the most controversial of all services provided by the country's insurance sector (Amoo, 2002).

The national economy is run on all the four modes of transportation, namely; Road, Rail, Aviation and Maritime & lake. However, the road transport sub-sector constitutes the major mode of transportation for both passengers and goods within the country. This sub-sector accounts for 97% and 94% of all national passenger and freight traffic respectively. A total of 215,537 vehicles operated in the economy in 1998. However, this figure far exceeds the total number (154,347) of insurance policies written (TRI-STAR, 2001). This discrepancy can be attributed to either some of the operating vehicles contravened the law by not taking any insurance policy within the year, or data capturing and/or processing by the authorities are improper thereby producing wrong statistics.

The road transport industry is, however, beset with a number of related and unrelated risks that must be borne by one or more of the stakeholders. These could be the owner of the vehicle, passengers on the vehicle, a third party who may be affected in one way or the other by the user of the vehicle or an insurance company. Fortunately, some of these risks can be shared among some of the stakeholders. The government enacted laws, Motor Vehicles (Third Party Insurance) Act 1957, making it compulsory for every mechanical propelled vehicle plying Ghanaian roads to have some form of an insurance cover. It is primarily aimed at protecting the third party who may be affected by the use of the said vehicle (TRI-STAR, 2001). This has been the basis upon which the risks associated with motor transportation have been categorized and insurance products developed thereto. Based on the categorization, the Ghanaian motor insurance market has some various insurance covers in its portfolio. These are the *Act Only Cover* (which is compulsory), *Third Party, Third Party Fire & Theft &Liability*, and *Comprehensive Cover.*

The Act Only Cover is the most basic motor insurance product in the Ghanaian market. It allows the policyholders to pass on the risk of providing unlimited compensation for injury or death to a third party by the insured vehicle to the insurance underwriter. The underwriting company assumes only the risk of injury or death to a third person, arising out of bodily injury. This third party can either be passenger in the said insured vehicle or a pedestrian. In Ghana this cover is compulsory for all vehicles except those of the police service, government departments or the military (TRI-STAR, 2001).

The Third Party Cover enables the underwriting company to assume the risk of providing unlimited compensation for bodily injury/ death or proper damage, or both, caused by the insured vehicle to a third party.

The Third Party Fire & Theft & Liability Cover is that segment of motor insurance coverage that encompasses the third cover for liabilities to third parties plus damage or loss of the policyholder's own vehicle caused by the fire or theft (TRI-STAR, 2001).

The Comprehensive Cover is the most extensive form of motor insurance cover available in the marketplace. It permits the policyholder to transfer all the risk associated with the use of insured vehicle onto the underwriting company. These risks include bodily injury to passengers and pedestrians, third party property damage, and accidental loss or damage of the whole or part thereof of the insured vehicle as a result of fire, theft, or accident (TRI-STAR, 2001)..

The premium paid for these covers are based on the type and/or use of the said motor vehicle. The insured vehicles are broadly classified either Private or Commercial Use. Further classification is made on the engine capacity of the vehicle (TRI-STAR, 2001).

The risk factors associated with claims have been the topic of many research papers appeared in the actuarial literature. The longitudinal (or panel) data model is often used in this context due to the randomness of the annual expected claim frequency, influencing the number of accidents reported to any company. However, this research work is interested in a policyholder ever making a claim or otherwise, which would be used as a basis for developing a generalised linear regression model, predicting frequency of motor insurance claims.

1.2 PROBLEM STATEMENT

Insurance is by nature an ultra competitive business. There is nonetheless, concern about the "over desire" by some insurers to win market share at a cost, thus leading to underwriting indiscipline. This is not uncommon in other insurance environment, but the use of tariff system in Ghana has, for now, militated against such tendencies. However, concerning the aggressive nature with which some companies are clamoring for increase market share, these problems could still come about if measures are not taken to stem such tendencies.

The motor Tariff Regime being employed in Ghana comprises the rating guidelines set down by the NIC. These guidelines set out minimum rates that all insurers must follow, with adjustments for discount. Thus under this regime, underwriting skills have mainly been relied upon in assessing those factors affecting level of discounting appropriate for a particular policy. The various discounts applicable are No Claim Bonus Discount, Fleet Discount on Fleet Policy and "Special" Discount, which a discretionary discount offered by insurers under special circumstances. For instance, where an insured has been a client successively with an insurer for a long time or where a client holds multiples policies with one insurer, a special discount could be offered

Arising out of the problem of minimal utilization/application of underwriting principle (inherent with the Tariff regime) as earlier mentioned, is the fact that the only data analysis done at present is not based on accurate figures or specific information. However, it is recognized that eventually the Tariff may have to be done away with (as it has in other development markets). This may be followed by a rating "war" between insurers as they strive to increase market share. The mere fact that industries in Ghana do not put emphasis on some of the risk factors in estimating motor insurance claims would give room for unfairness in making claim.

It is in line with the foregoing claim problems of motor insurance that this study tries to consider and critically examine some important characteristics of the insured vehicle and its driver. The study therefore seeks to address issues of the attitudes and views of the insurance companies regarding estimation of motor insurance claim as well as factors that underwriting companies consider in premium rating. The pertinent risk factors or characteristics being considered here to model the problem understudy include the sex of the policyholder, make of vehicle, age of vehicle at point of entry, and cubic capacity of vehicle.

1.3 RESEARCH OBJECTIVES

To ensure sanity and professionalism in the way motor insurance is underwritten and sustained for profitable growth of the industry while meeting clients' expectation, it has undoubtedly become imperative to put forth some quantitative arrangements or measures that will accelerate efficient and prudent underwriting. This, it is believed, would help assess clients fairly and equitably, based on their risk characteristics, which will then allow the underwriting companies to charge adequate and sustainable premiums. Undercutting of rates based on spurious unproven rationale would be checked, thus improving the financial solvency of such companies.

The consideration of future for the estimation of motor insurance claims would entail a variety of areas of investigation. This research is conducted as a way of undertaking this exploration. In particular, the goals of this research study are:

- To have a thorough knowledge and understanding of the motor insurance industry and how the industry operates in Ghanaian market.
- To establish a logistic regression model with the ultimate goal of determining the key risk factors associated with the motor insurance claim as well as predicting the frequency of making claims.
- To make some recommendations based upon the various analyses performed and conclusion of the study for policy implementation.
- To contribute to promotion of knowledge in motor insurance claims with a view to, among others, stimulate further research.

1.4 JUSTIFICATION OF STUDY

A lot of people have researched into motor claims insurance and contributed to knowledge immensely. The study is to bring to attention such contributions as well as that of mine with a view of, among other things, stimulating further research.

It would be suggested that SIC Insurance company Limited give relatively high premium to new cars as compared to older ones due to their higher demand of claim, thus helping SIC to increase their financial gains. The different various influential effects of risk factors should motivate the drivers or policyholders to be cautious in driving so as to avoid high premiums. This will help the professionals in the motor insurance industry in their informed decisions by appreciating the problems underground and also advise the underwriters from this position of knowledge.

1.5 RESEARCH METHODOLOGY

The data for the analysis of the study were obtained from the Motor Department of SIC Insurance Company Limited (SIC). The data collected were on insured private vehicles for the period, 2007–2008, indicating their make or type, cubic capacity, age of entry, and sex of driver (the policyholder). The National Insurance Commission (NIC) was also contacted for some literature materials concerning the estimation of motor insurance claims in Ghana.

To achieve the objectives of this research, the logistic regression analysis was employed as the main statistical methodology for analyzing the data. The response variable, in this case, indicates whether a claim is made or not by an insured vehicle. This binary variable is regressed on the risk factors (make, cubic capacity, age of entry, and sex of policyholder of vehicle) as the predictor variables for the modeling process. These predictor variables are individually categorized into different levels to suit some required analyses like the estimation of the odds ratios.

The results from the various analyses performed were obtained using the R software package (version 2.8.1) which implements the logistic regression analysis. The results show that sex of policyholder and the cubic engine capacity of vehicle have no significant impact on claims made, though female drivers are 1.26 times as likely to make a claim as their male counterparts. The odds ratios of the vehicle cubic capacity indicate lesser claims as size of engine increases. Also, after adjusting for all the levels of a car age, it was found that they were all less likely to make a claim as compared to cars that are not more than five years. They are nevertheless statistically significant. The established model used only two risk factors: the make and age of vehicle, indicating the various levels found to be significant to the prediction of a motor insurance claim.

1.6 SOME BASIC CONCEPTS IN INSURANCE RISKS

1.6.1 Insurance

Insurance is a contract (policy) according to which one party (a policy holder) pays an amount of money (premium) to another party (insurer) in return for an obligation to compensate some possible losses of the policy holder. The aim of such a contract is to provide a policy holder with some protection against certain *risks*. Death, sickness, disability, motor vehicle accident, loss of property, etc. are some typical examples of such risks. Each policy contract specifies the policy term and the method of compensation. Usually compensation is provided in the form of payment of an amount of money. Any event specified in the policy contract that takes place during its term can result in such an insurance claim. If none of the events specified in the policy contract happen during the policy term, then the policy holder has no monetary compensation for the paid premiums. Insurance helps individuals and organisations to reduce the financial risk of adverse event. An example is the cost of hospital care in the event of illness.

Insurance involves two parties, namely the insurance company, also known as the *insurer*, and policy holder, *insured.*

The policy holder purchases an insurance policy by paying a premium to the insurance company and reflects the level of risk. In return for the premium the insurance company promises to pay the policyholder an amount of money (called a benefit or claim payment) if the policyholder were to suffer a financial lost due to a specific event. The premium for each risk should commensurate with the risk. Those who are more likely to have losses should pay higher rates; those insuring for higher amount should pay higher premiums. The outgoings of the funds are mainly the payment of loses.

Insurance is therefore a method of sharing the losses of the few individuals in a group who suffer them among the many members of that group who do not (IIC, 2006). Insurance is based on risk. If there is no risk there can be no insurance.

Risk

The word risk implies both doubt about the future and the fact that the outcome could leave us in a worse position that we are in right now. In our daily lives we are constantly exposed to risk. Very little is certain. We risk our lives when we cross the street, being responsible for injury to others when we own or drive an automobile, losing our property if we are careless with fire or the security of our home, the lost of fortunes when we go into business for ourselves, and we risk failure when we do not study for an examination.

There are two classes of risks, *speculative risk* and *pure risk*. Speculative risk exists where there is a chance of loss and a chance of profit. A business operation is a speculative risk, it could fail or be successful. Such risks are not insurable. Pure risk entails a chance of loss but no chance of profit.

1.6.2 Insurable Risk

So far we have talked of risk in a very general way using a wide range of examples. As we said earlier the study of risk is to be a foundation for a more detailed examination of insurance and so we must now turn our attention to those risks that are insured. Insurable risks may be classed as *personal*, *property*, or *liability risks*.

- The *personal risks* are based on the loss of life or loss of income arising from premature death, or physical disability.

- The *property* risks are those arising from the destruction of, or damage to property. These losses may be *direct*- the cost to repair or replace the property, or *indirect* - the loss of use of the property caused by the damage or destruction of the said property.

- The *liability risk* arises from one's obligation to pay damages because of bodily injury to or death of another person (known as the third party) or damage to his property. This obligation would be based on one's negligent acts in relation to the operation of vehicles or equipment, the ownership or occupancy of property, the manufacture of products and the rendering of services. A person may take action against the insured for damages allegedly caused by the insured. The insured may have insurance protection (legal liability) for such an action and will be indemnified by the insurer for any damages found against him. The person who brings the action is called the third party.

1.6.3 Risk Management

Risk management is the minimization (at a minimum cost) of the detrimental effects of risk by identifying, measuring and controlling the risk. Risks are identified by inspection, through the use of inspectors, engineers and surveyors employed by the business. Once the risk have been identified and measured the risk manager must develop a method of controlling them. There are several alternatives available. These are:

- *Reduction of Risk by Preventive Effort*: If the risk is such that it can be reduced or eliminated by preventive effort it would be wise to do so. While the initial cost to reduce the risk may be high, when these costs are amortized over a period of time, there can be a substantial savings in lives and money. An example is the use of sprinkler system to reduce the possible effects of tire.
- *Assume or Retain the Risk*: A business operation may due to circumstances, choose to assume or retain the risk, self-insurance may be used where the business's operations are geographically scattered, when the type of risk is difficult or impossible to insure, when insurance is provided but the deductible is quite large, or where it is found to be the most economical method of controlling the risk, some businesses may set up a fund into which money is periodically paid to provide money should the event occur.

- *Transfer the Risk*: Risk managers frequently decide that the best method to deal with the specific risk is by transferring it to another person or corporation, which may have better financial resources and thus a better ability to withstand loss. This transfer is called insurance.

The individual or business seeking to formulate a method to control the risk will be strongly influenced by the *perils* to which it may be subjected to and the *hazard* that exists. A *peril* is an event, which may cause a loss, such as a fire, lightening or explosion. The peril may cause a direct loss - damage to property, or/and indirect loss of income resulting from the damaged or destroyed property.

A *hazard* is a condition, which may cause a peril to occur. There are two categories of hazards, physical and moral. Physical hazards are those which pertain to physical state of the property and which may lead to the occurrence of a peril. Moral hazards are characteristics of individual or firms, which may increase the probability or severity of loss while the physical hazard can usually be reduced, a moral hazard is not only difficult to determine but difficult to prevent because it relates to human element of the risk.

1.6.4 Reinsurance

An insurer tries to develop a book of business by spreading its liabilities in such a way that no single incident will affect all its risk at the same time. Insurers would rather write 1% of all homes, businesses or vehicles in all the towns and cities in Ghana than write 100% in say Accra or Kumasi. The amount of premium that would be generated either way may be quite similar but the possibility of being wiped out in one accident is disproportionately higher with the greater concentration of exposure

Reinsurance is an arrangement between the underwriting insurer and a professional reinsurance company. The insurer transfers all or part of a risk to another insurer known as a reinsurer in return for a portion of the premium generations by the underwriting insurer. The insurer would absorb an agreed proportion of the claims. The reinsurer transacts business only with insurance companies. The reasons for reinsurance are:

- To increase the insurer's capacity to write business: Reinsurance allows the insurer to write a higher level of risk than it would be to its own accounts.
- To reduce the effect of a catastrophic loss: While every insurer is prepared to accept ordinary loss in a course of its business, events like the 1995 floods in Accra may seriously strain its financial resources particularly insurers with large retention (they keep more of each risk than they reinsure). Reinsurance on a catastrophe basis is available for such an eventuality.
- To provide stability in a fluctuating market: Reinsurance will not prevent underwriting losses but it can smoothen out the effect on the insurers in any given year.
- To enable an insurer to cease operation: For reasons other than bankruptcy an insurer may wish to withdraw from the market. To do so, it must remain until all business has run off and the outstanding claims have been settled which could be a very long term and expensive proposition or to reinsure its outstanding portfolio of business with another insurer. This is not reinsurance in the strict sense but it is a transfer of risk from one carrier to another.

The insured is not a party to a reinsurance contract. Reinsurance arrangements are contracts between the primary and other insurers or professional reinsures. The original insured is not a party to this contract and usually knows nothing about it until the time that the insurer is financially unable to pay a claim and the reinsurer may be called upon in most cases, the arrangement made remain private. The spirit of utmost good faith with which insurers transact business with their agents, brokers and insured also extend to their transactions with their reinsurers and they therefore feel that these transactions should remain private because they would only enter into agreement that were in the best interest of their clients

1.7 BACKGROUND OF STUDY AREA

1.7.1 Introduction

SIC Insurance Company Limited (SIC) is one of the oldest non-life insurance companies in Ghana. It traces its roots to the year 1955, when the Gold Coast Insurance Company was established. It was renamed Ghana Insurance Company (GIC) in 1957, when Ghana attained her independence. In 1960, GIC which was primarily a life assurance company, set up a subsidiary company – Ghana General Insurance Company (Ghana General) to underwrite fire and motor businesses. The Government of Ghana in February 1962, per an executive instrument, took over the Ghana Cooperative Insurance Company and reconstituted it into the State Insurance Corporation to await the completion of takeover negotiations with Ghana Insurance Company and its subsidiary Ghana General. Subsequently after a successful takeover of the two private companies – Ghana Insurance and Ghana General – the new company, State Insurance Corporation, commenced business in November, 1962.

In 1995, SIC was converted into a public limited liability company as part of the Government of Ghana's divestiture programme. The company became known as State Insurance Company of Ghana Limited with the Government of Ghana as the sole shareholder. The State Insurance Company of Ghana Limited took over all the business assets and liabilities of the State Insurance Corporation of Ghana. By a special resolution passed on 22nd October, 2007 the name of the Company was changed to SIC Insurance Company Limited. In accordance with the provisions of the current insurance legislation, the Insurance Act, 2006, SIC has duly separated its general business from the life business with the incorporation of SIC Life Limited. By a special resolution, SIC transferred the life business and assets to SIC Life Limited and in consideration of those 80,000,000 ordinary shares were issued to the Government.

SIC is a leading provider of general or non-life insurance products in Ghana. Its business operations cover fire, motor, marine and aviation, and accident insurance. SIC also provides specialty insurance products such as hoteliers and leisure policy, a policy for the hospitality industry. With roots dating back to 1955, the Company has been operating for more than 52 years and has developed long-standing relationships with insurance brokerage firms and some independent agents, who constitute some of its primary distribution channels.

SIC's business is national in scope with a visible presence in all ten regions of Ghana. The Company has consistently, over the span of its business life, maintained steady market leadership. In 2006, SIC had approximately 40% of the insurance industry's total market share. SIC's vision is to maintain its dominance in the insurance industry and position itself as the most profitable, innovative and customer friendly insurance company. Its mission is to provide innovative and competitive insurance and other financial services to its clients through a highly skilled and motivated workforce with a commitment to be a good corporate citizen.

The Technical Operations committee consists of three (3) directors. The Head of Technical Operations of SIC is required to attend meetings of the committee. The committee holds bimonthly meetings to discuss various issues to do with the core business activities of the Company. The principal duties and responsibilities of the Technical Operations committee are as follows:

1. Formulate policy on technical operations for the attention of the board of directors;
2. Monitor the performance of the Marketing and Technical departments;
3. Ensure that all expenditure relating to technical matters is within the approved budgetary limits;
4. Ensure that the appropriate infrastructure and equipment is available for to the various departments to assist in the achievement of the goals and targets;
5. Monitor compliance with the relevant legal and regulatory framework;
6. Discuss any problems or reservations arising from any relevant reports and any matters the relevant supervising authority may wish to discuss.

The Technical Operations committee has the authority of the Board of Directors to investigate any activity within its principal duties outlined above. It also has authority to seek any information it requires from any employee, past and present, and such employee is required to cooperate with any request made by the committee. The Technical Operations committee has the authority of the board of directors to obtain outside legal or other independent professional advice and if it considers it necessary to secure the attendance of outsiders with relevant experience and expertise. SIC's investment portfolios are managed by SIC-FSL, a wholly-owned subsidiary of SIC, and an authorized investment advisor with a broker-dealer certification from the SEC. SIC-FSL provides investment advisory and asset management services to SIC, other corporate clients, and high net worth individuals. SIC's investment philosophy is to adhere to the guidelines under the Insurance Law and the Company's investment objectives.

1.7.2 The Structure of the Investment Portfolio

The Insurance Act, 2006 (Act 724) requires that general insurance companies invest 25% of their gross premiums written in Government bills and notes, and the remainder in securities approved by the National Insurance Commission. Investments are also structured to meet the company's profitability objectives and solvency requirements.

SIC places emphasis on optimizing investment returns over the long term. SIC's investment portfolio mix as of May 31, 2007 consisted largely of high quality, fixed deposits and foreign currency cash equivalent and short-term investments, as well as a significant amount of listed and unlisted equity securities. The company believes that prudent levels of investments within its investment portfolio are likely to enhance long term total returns without significantly increasing the risk profile of the portfolio. The table below presents the composition of SIC's investment portfolio as of May 31, 2007.

Table 1.1: Composition of SIC's Investment Portfolio at May 31, 2007

Investment Type	Percent of Total
Cash and Cash Equivalents (Domestic)	10%
Government Securities (Treasury bills and Notes)	25%
Fixed Deposits and Foreign Currency Cash Equivalent	40%
Long-Term Investments (excluding Real Estate)	25%
Total	100%

1.7.3 Insurance Products and Marketing Channels

Despite a highly competitive insurance market, it is significant to note that SIC continues to dominate the market with a market share of about 40%. The table below shows SIC's market share for the various insurance business lines for the years ended December 31, 2002 to 2005.

Table 1.2: SIC's Market of Insurance Products

Product	At Year Ending December 31			
	2005	2004	2003	2002
Fire	26	16	23	19
Motor	41	42	43	47
Accident	23	25	24	33
Marine & Aviation	70	68	69	60
Total	**40%**	**37%**	**38.1%**	**39.3%**

SIC offers its products through registered brokers, company-employed sales representatives, and marketing field staff. It also, to a large extent, depends on walk-in customers. SIC's marketing strategy is to appeal to customers who desire a feature-rich product at a competitive price. The Company uses broadcast, print media, community events, and the internet to offer its products to its target markets.

The company has a broad marketing reach because of its comparatively larger number of agents. It has approximately 600 in-house trained agents who market its products all over the country to cater for the special needs of its geographical markets.

SIC's marketing field officers do not only offer products to its target markets, but also collect market data for the company to use to enhance its existing product base and develop newer lines of policies. Additionally, the Company uses thirty-seven registered independent insurance brokers to offer its products, particularly for its fire, marine and general accident lines of business.

SIC has a consistent history of paying dividends to its shareholder. The Board of Directors, in accordance with the provisions of the company's code has the authority to recommend and pay dividends on ordinary shares. As a policy the minimum dividend payable is 40% of profit after tax. Any recommendation to pay dividends will be at the discretion of the board of directors and will be dependent on SIC's operational results, cash flows, financial position, capital requirements and any other factors the board of directors deem relevant. The average dividend payout ratio over the last five years has been approximately 43% of profit after tax.

1.7.4 Dividend History

Set forth in Table 3.3 is SIC's dividend history for the years ended December 31, 2002 to 2006. The dividend trend reveals a company that has continuously rewarded its shareholder for capital provided. During the five year period, spanning from 2002 to 2006, dividend payout ratio averaged 43%.

Payment of claims is perhaps one scope in which SIC continues to excel thereby assuring clients of her dependability and swiftness in responding to their plights. SIC has a consistent history of paying dividends to its shareholder.

Table 1.3: SIC Dividend (2002 -2006)

Year	Dividend Paid (GH¢ '000)	Dividend per share (GH¢)	Payout Ratio (%)
2006	1,300	0.03	44
2005	800	0.02	35
2004	400	0.01	23
2003	600	0.01	40
2002	400	0.01	73

1.7.5 Documents Required When Liability Is Accepted Injury Cases

The insurance industry is regulated by the National Insurance Commission (NIC), which administers the Insurance Act, 2006 (Act 724). The NIC has broad supervisory and administrative powers over such matters as licenses, standards of solvency, premium rates, policy forms, investments, minimum capital and surplus requirements, periodic examinations, annual and other report filings.

Claims management is the carrying out of the entire claims process with particular emphasis on the monitoring and lowering of claims review, investigation and negotiation. A claim is a request to be reimbursed (or compensated) filed by the insured and addressed to the insurer. Effective claims management is a critical factor in achieving satisfactory underwriting results. The procedure for handling claims varies according to the type of cover, amount of claim, and whether it is a personal or commercial claim. SIC maintains in-house investigators, engineers and surveyors to handle most of the claims, in order to detect any fraud or abuse in the processing of claims. Furthermore, in-house lawyers aid claim management procedures and law enforcement investigations in an attempt to reduce insurance fraud and litigation.

1.7.6 SIC Motor Insurance

SIC's motor insurance business is the biggest in Ghana. More than 50% of its gross premiums written come from this category with an average industry market share of 43% for the period 2002-2005 (Latest data available from the NIC). The Company's motor insurance business is national in scope and covers both private and commercial motor users. Motor premium rates vary according to usage and the less extensive the cover, the smaller the premiums. SIC's dependence on motor premiums is diversified across a broad customer base and no one single customer influences more than 2% of gross premiums.

Apart from the standard coverage of third-party liability and comprehensive insurance, the Company also handles additional perils for both private and commercial motor users under both its comprehensive and third-party liability coverage. Examples of perils include, flood, volcanic eruption, earthquake or other convulsions of nature, strikes, riots and civil commotion. SIC has motor gratuitous passengers/members of insured's household coverage, which pays full compensation for bodily injury or death to occupants of the motor vehicle other than the insured or the person driving.

CHAPTER TWO

2 MOTOR INSURANCE

2.1 INTRODUCTION

The first mechanically propelled motor vehicle appeared on the British roads in 1894 and by 1898 the Law Accident and Insurance Society Limited were offering policies to motor vehicle owners. The business was new and many of the early companies did not survive in the competition. The mushrooming use and development of the motor car followed the First World War during which the advantages of motor vehicles had been established. By the 1920s there were so many motor vehicles on the roads and so the legislation which passed in 1930, the Road Traffic Act, was inevitable.

In Ghana, insurance business started in the colonial era where it was done through foreign trading companies in the country acting as chief agents of insurance companies in the United Kingdom and elsewhere. This was concentrated in few towns such as Accra, Sekondi and Kumasi. However, the growth of Ghana's insurance industry took a turning point in 1962 when bold and revolutionary steps were taken to improve the insurance industry in the country. This resulted in the formation of the State Insurance Corporation leading to the spread of insurance centres to all the regional capitals and to some of the district capitals and beyond in the late 1980's. In recent times there has been an increase in the number of insurance companies in the country which totals 22.

The insurance market consists of insurers (insurance companies), intermediaries (insurance brokers and agents), and the insuring public (government institutions, corporate bodies, and individuals). Products that are offered on the insurance market are broadly categorised into two – Life and Non-Life (also referred to as General Business).

While the range for Life products varies, underwriting for Non-Life centres on Motor, Marine, Aviation, Fire, and General Accident.

A developed and functioning insurance sector is a fundamental condition for economic success. The objective of insurance is to provide financial stability to individuals, organizations and businesses. As a risk pooling and transfer mechanism, insurance allows the insured to mitigate pure risks (i.e. risks that involve only the possibilities of loss or no loss). Examples of such risks are fires, flooding, ill health and unintentional damage to a third party. Insurance helps business to stay open and individuals to continue their work or education by providing financial compensation if an insured risk occurs and causes damage. Even when no loss occurs, insurance provides peace of mind, a service of considerable, if un-quantifiable, value. (Yuengert,, 1993). As financial sector, insurance is a major investor. The insurance sector covers long and short-term risk activities. It comprises three basic activities: "*life insurance*" which includes common life insurance and life reinsurance with/without a saving component, "*Non-life insurance*", comprising insurance and reinsurance of non-life insurance business, e.g. accident, fire, health, property, motor, marine, aviation, transport, pecuniary loss and liability insurance, and "*Pension funding*" which includes the provision of retirement incomes, but non-contributory schemes where the funding is largely derived from public sources. Reinsurance activities are included in one of the three sections, according to the kind of the risk reinsured (Fecher,et al, 1993)

The insurance sector is one of the most important service sectors regarding its basic function for the whole economy and society. Modern, highly industrialized and technology driven economies are threatened by higher risk than ever; and individuals need to protect themselves against private risks as well as saving individually for their retirement. Insurance companies also play an important role as investors and shareholders.

2.2 *MOTOR INSURANCE CLAIMS AND COMPLIANCE*

This section presents some research works done in motor insurance claim and compliance of rules.

2.2.1 Motor Insurance Claims

The service of claims settlement provided by insurance firms are correlated with loss aggregates. Nonetheless, this empirical proxy also has its drawbacks (Cho, 1989). First, the stochastic nature of non-life losses is likely to create an 'errors-in-variables problem' that would distort the measurement of scale economies, particularly where the insurance includes catastrophe risk. Second, when claims incurred are used, the output quality of loss control and risk management is ignored, which is undesirable. As a result, an insurer that is very successful in loss-prevention would be measured as having less output. Given that both premiums and claims incurred are imperfect measures of insurance output, we follow the suggestion that says alternate between both measures of insurance output (Cummins et al, 2000).

Estimating loss severity distributions (i.e. distributions of individual claim sizes from historical data is an important actuarial activity in insurance. In this situation it is essential to find a good statistical model for the largest observed historical losses. It is less important that the model explains smaller losses. In fact, a model chosen for its overall fit to all historical losses may not provide a particularly good fit to the large losses. Such a model may not be suitable for pricing a high-excess layer. It is obvious that extreme-value theory is the most appropriate tool for this job, either by using extreme value distribution to model large claims or generalized Pareto distribution to model exceedances over a high threshold.

In his comprehensive paper entitled "A General Survey of Problems Involved in Motor Insurance", Dr. Carl Philipson discusses two methods of statistical estimation of Third Party Motor Insurance claim reserves. These methods can also be used for Car Damaged Insurance, since reserve determination for these rapid settlement property coverage is simpler than for Third Party lines. Dr. Philipson mentions the two main purposes of claim reserves –balance sheet loss reserves which shall be estimated on the safe side for financial reasons, and those reserves needed for risk statistics. However, this research work would confine its subject to risk factors associated with claims.

There is another paper that examines the validity of some stylized statements that can be found in the actuarial literature about random effects models. Specifically, the actual meaning of the estimated parameters and the nature of the residual heterogeneity are considered. A numerical illustration performed on a Belgian motor third party liability portfolio supports it. It is one of the important principles that applies to the contract of insurance thus, contract *uberrimae fidei*. It refers that both the parties involved in insurance contract should make the disclosure of all material facts and figures relating to the subject matter of the insurance contract. If either party does not disclose the utmost good faith the other party may avoid the contract. The insured's duty is to disclose all material facts known to him but unknown to the insurer. Similarly, the insured's duty of utmost good faith is disclosing the scope of insurance at the time of contract. Any concealment, misrepresentation, fraud or mistake concerning the material facts to the risk should be disclosed. No important material facts and figures must be concealed. Thus, responsibility of disclosure of both parties should be a reciprocal duty i.e. disclosure is absolute and positive. In an efficient, competitive insurance market, higher risk drivers face higher insurance premiums to reflect the higher expected accident costs that they impose on the insurance system (**Englund et al,** (1997). But a system of purely cost-based pricing may lead to degrees of price variation that are politically unacceptable or that undermine the goal of premium affordability for some drivers. Therefore regulators intervene in automobile insurance markets to reduce rate levels for high-risk drivers, to limit average rate increases for all drivers, and in some cases to reduce rate variation across drivers. The end result is that high-risk drivers pay less and low-risk drivers pay more than they would under a purely cost-based rating system, and the links between driver loss experience and rate changes are weakened.

With limited premium income, there will inevitably be restricted amounts available for compensation awards. In addition to the adequacy of compensation awards, there also appears to be a problem with few casualties claiming and collecting compensation. Victims will also be dependent on solicitors for their claims and it will not always be in the solicitor's interest to settle quickly.

2.2.2 Motor Insurance Compliance

Of the four LIC case studies, the insurance industry is involved in three of the national road safety management bodies. In terms of funding road safety, South Africa's RAF has agreed to contribute 2.5 per cent to prevention activities and Ghana's motor insurers are expected to donate a percentage of premium income. The state insurance provider in

Costa Rica has also invested in large publicity campaigns while the insurance industry in Karnataka has not been involved in either promoting road safety or investing in prevention efforts (Aeron-Thomas, 2002). Very little road safety research appears to be underway in the LIC case studies with the motor insurers still in the process of computerising claim data. Research centres do exist elsewhere, (Motor Insurance and Road Safety Scoping Study Final Report TRL 44 April 2002), especially in Latin America, and while focused on vehicle repair, they are also concerned with casualty and crash reduction.

Many countries, especially LICs, have problems with motor insurance compliance. However, it has been possible for several HICs to achieve good results. Earlier World Bank project reports included insurance coverage estimates of Peru (22%), Pakistan (10-20%), Zambia (15%) and in Vietnam, one quarter of motorcycles and half of four wheel motor vehicles were estimated to be insured (Eijbergen 2001, Aeron-Thomas 1999). In many countries, the owners of older commercial vehicles can have problems obtaining insurance. In France, if a motor vehicle owner has difficulty finding an insurer, the Central Rating Office will require a company to insure it (A Aeron-Thomas 1999).

As with vehicle inspection, many countries use a windscreen decal to show insurance coverage. This measure has been popular in Europe for several years and has recently been implemented in Ghana. The UK rejected the windscreen decal approach because it was considered too easy to replicate false decals. Instead, the insurance industry collaborated and funded the establishment of a Motor Insurance Database where insurance details are centralised and access allowed by police, insurers, etc.

With one agency, ICBC, responsible for licensing as well as mandatory insurance in British Columbia, insurance compliance can be checked any time a license is issued and renewed every five years. British Columbia is believed to be the only jurisdiction which has these two responsibilities under one agency.

The enforcement of insurance coverage is rarely a priority of traffic police. In Karnataka, the police announce motor insurance checks in advance. Driving while uninsured in Ghana can result in one-year imprisonment or a one year driving disqualification. These penalties are not believed to be implemented and the number of motorists fined for driving while uninsured was not available from the Traffic Police.

As in Ghana, driving in the UK while uninsured can incur more than just the fines, for example, between 6-8 penalty points on a driving license. However, a survey of the UK Magistrates courts found two third of the fines for uninsured driving were for £200 or less. The UK government arguably suggested that community sentences and permanent or temporary vehicle forfeiture be extended for driving while uninsured (Home Office, et al, 2000). A survey funded by the Insurer Direct Line found 44% drivers want more roadside checks to trap insurance cheats, 30% drivers say that the fines for uninsured drivers are too low and 28% favour imprisonment for persistent offenders (Williams, 2002). Sweden does not fine drivers for use of an uninsured motor vehicle but adds a surcharge of up to 10 per cent to the next policy which must be purchased from the state.

2.2.3 Premium Rating of Motor Insurance

In another development, following widespread positive feedback on the achievements of the General Insurance Reserving Issues Taskforce (GRIT), the General Insurance Board of the Faculty and Institute of Actuaries established GRIP in 2005 to review actuarial involvement in premium rating issues, pricing being one of the key areas in which actuaries work. The initial terms of reference as suggested by the General Insurance Board (GI Board) were refined following discussion within GRIP and wider consultation with the Faculty and Institute of Actuaries (the Profession) at GIRO in October 2005. The final terms of reference agreed with the GI Board can be summarised as follows:

- To review the areas in which UK actuaries are currently involved within the overall premium rating process, and to identify any areas where actuaries might be able to improve their contribution and/or add further value.
- To summarize, in broad terms, current methods used by actuaries in GI premium rating, to identify areas where types of methods and approaches could potentially be improved or used more appropriately, and to suggest potential areas for further research.
- To consider whether and how improvements could be made to the way GI pricing actuaries communicate with others.
- To consider whether existing professional guidance should be modified/clarified to make its application to premium rating clearer and to consider whether there is a need for more detailed best practice guidance from the Profession setting minimum standards for a direct business pricing assignment.

- To consider whether there are any implications for professional guidance or communication resulting from commercial pressures within organizations.
- To consider trends in the area of Treating Customers Fairly and to consider what the Profession might need to consider in preparation for issues arising in this area.
- To consider whether the content of the exam syllabus is adequate to prepare actuaries to work in the pricing area.

Although it was initially proposed that GRIP should not consider London Market reinsurance pricing, feedback from the profession clearly suggested that there was a desire that this area also be covered. Thus it was agreed that GRIP should consider "issues of relevance to Faculty and Institute members involved in pricing direct insurance (individual and account level products) and reinsurance".

In most countries, the government after consulting with the insurance industry determines the cost of motor insurance premiums. In many developing countries, the third party premium charges are influenced by transport fleet operators. The lowest third party premium for a private car was reported to be approximately £16 while in India, it was about £10 for a motor vehicle with greater than 1500cc (£7 for less than 1500 cc). Such low premiums obviously affect not only the potential compensation amounts available but also on the sensitivity of the premium to any pricing incentives. Adjusting insurance premiums to reflect perceived risk is the traditional, if not necessarily effective, road safety intervention adopted by insurers. It is standard practice to base the insurance premium on the vehicle type, and many countries also consider geographical location. Both Sweden and the UK allow premiums to be set by the insurers and many factors can influence the price. UK insurers offer premium reductions on the basis of age, sex,

- To consider whether there are any implications for professional guidance or communication resulting from commercial pressures within organizations.
- To consider trends in the area of Treating Customers Fairly and to consider what the Profession might need to consider in preparation for issues arising in this area.
- To consider whether the content of the exam syllabus is adequate to prepare actuaries to work in the pricing area.

Although it was initially proposed that GRIP should not consider London Market reinsurance pricing, feedback from the profession clearly suggested that there was a desire that this area also be covered. Thus it was agreed that GRIP should consider "issues of relevance to Faculty and Institute members involved in pricing direct insurance (individual and account level products) and reinsurance".

In most countries, the government after consulting with the insurance industry determines the cost of motor insurance premiums. In many developing countries, the third party premium charges are influenced by transport fleet operators. The lowest third party premium for a private car was reported to be approximately £16 while in India, it was about £10 for a motor vehicle with greater than 1500cc (£7 for less than 1500 cc). Such low premiums obviously affect not only the potential compensation amounts available but also on the sensitivity of the premium to any pricing incentives. Adjusting insurance premiums to reflect perceived risk is the traditional, if not necessarily effective, road safety intervention adopted by insurers. It is standard practice to base the insurance premium on the vehicle type, and many countries also consider geographical location. Both Sweden and the UK allow premiums to be set by the insurers and many factors can influence the price. UK insurers offer premium reductions on the basis of age, sex, additional driver training, and just recently, an insurer is offering to charge on the basis of mileage with a black box fitted to the vehicle (Thomas, 2002).

Sweden was the only case study identified which gives a discount to teetotaller drivers. Other countries use penalties to discourage drink driving. In the UK, drivers convicted of a drink driving conviction will experience difficulty in finding an insurer and their premiums will double in price. The impact of a drink driving conviction will also affect the insurance premium for several years.

Five southern African countries (South Africa, Botswana, Swaziland, Namibia, and Lesotho) currently collect third party bodily injury motor insurance through the fuel levy, which virtually eliminates the possibility of driving without insurance. In another development, the 'bonus malus' system refers to the use of premium discounts for claim-free driving and surcharges for crash involvement. No-claims discounts (NCD) are still popular in the UK, British Columbia and Sweden, with discounts up to 75 per cent available in the UK. However, NCDs are easier to justify as a marketing tool rather than as an effective safety intervention. The effectiveness of NCDs has been doubted, if not rejected, for many years (OECD, 1990). Even in countries where NCDs are popular, such as the UK, it is acknowledged that NCDs are not thought to be effective in reducing collisions. NCDs are believed by many to encourage non-reporting of claims, especially minor claims, rather than safer driving (Thomas, 2002).

The risk factors of motor insurance in Ghana, by reports of the National Insurance Commission are Marital Status, No Claims Discount percentage, Policyholder age, Policyholder sex, Type of Insurance (e.g. comprehensive), Vehicle age, Engine capacity, Value of Car, Model Year, Accident in the past year, Vehicle use by driver – pleasure or business, Previous accidents, and Vehicle Make.

2.3 *MOTOR INSURANCE COVERS*

2.3.1 Introduction

Many different classes of insurance have evolved and will continue to evolve because of the diverse needs of the insured. The prominent ones are motor, property, liability, fire, marine, inland transportation, crime, business interruption, boiler and machinery and life insurance. The intention of the 1930 Act, inter alia, was to ensure that funds would be available to compensate the innocent victims of motor accidents. This was to be provided by means of insurance against legal liability to pay damages to injured persons. The insurance requirement applied to all users of motor vehicles except where some special legal arrangement was in force. Further legislation followed in the Road Traffic Act 1960 and the Motor vehicles (Passenger Insurance) Act 1972, so that today insurance must be in force to cover legal liability to pay damages to any person, including others in the car, arising out of injury.

2.3.2 Types of Motor Insurance Cover

The motor insurance includes private cars, commercial vehicles, motorcycles, motor trade, and special types. The policies of these motor vehicles are also classified as into various covers which are discussed in the following sequel.

- *Private Cars Insurance*: The minimum requirement by law is to provide insurance in respect of legal liability to pay damage arising out injury caused to any person. A policy for this risk only is available and is termed an "Act only" policy. Such policies are not at all common and are usually reserved for a situation where the risk is exceptionally high. A "third party only" policy would satisfy the minimum legal requirement and in addition would include cover for legal liability where damage was caused to some other person's property. An addition to this form of cover is where damage to the car itself from fire or theft is included, the familiar "third party, fire and theft" policy. The most popular form of cover, accounting for about 66 per cent of all private car polices, is the "comprehensive policy". It covers all that has been said above with the addition of loss of or damage to the car itself. This policy also includes certain personal accident benefits for the insured and in some cases the insured's spouse. It also provides cover for loss of or damage to personal effects and medical expenses for passengers in the car.
- *Commercial Vehicles, Motorcycles and Motor Trade*: All vehicles used for commercial purposes, lorries, taxis, vans, hire cars, milk floats, police cars, etc, are not insured under private car policies but under special contracts known as commercial vehicle policies.

  This is a growing sector of motor insurance business and may well continue to be so if fuel prices become more and more expensive. The type of policy depends upon the machine, whether it is a moped or a high-powered motorbike, and on the age and experience of the cyclist. The cover is comparatively inexpensive relative to motor car insurance.

Special policies are offered to garages and other people within the motor trade to ensure that their liability is covered while using vehicles on the road. Damage to vehicles in garages and showrooms can also be included under such policies.

- *Special Types:* The present classification of insurance business refers to "land vehicles other than railway rolling stock" and many such vehicles fall under a category known to insurers as "Special types". These will include forklift trucks, mobile cranes, bulldozers and excavators. Such vehicles may travel on roads as well as building sites and other private ground. Where special type vehicles are not used on roads, and are transported from site to site, it is more appropriate to insure the liability under a public liability policy as the vehicle is really being used as a "tool of trade" rather than a motor vehicle.

The common types of insurance covers are:

- *Third Party Cover*: This policy insures you against claims for bodily injuries or deaths caused to other persons (known as the third party), as well as loss or damage to third party property caused by your vehicle.
- *Third Party, Fire and Theft Cover*: This policy provides insurance against claims for third party bodily injury and death, third party property loss or damage, and loss or damage to your own vehicle due to accidental fire or theft.

- *Comprehensive Cover*: This policy provides the widest coverage, i.e. third party bodily injury and death, third party property loss or damage and loss or damage to your own vehicle due to accidental fire, theft or an accident.

Automobile insurance provides protection against financial loss incurred because of personal injury and property damage claims involving insured's motor vehicle. The significant number of vehicles on the road combined with the increasing number of claims arising out of auto accidents makes it virtual necessity that every car owner protects him/herself with auto insurance. In most countries, there are mandatory minimum amounts of coverage that every car owner must purchase. Essentially there are two categories of automobile insurance.

- Liability or Compulsory (or sometimes third party) insurance, which includes coverage for bodily injury and property damage *and*
- Optional insurance, which includes collision and comprehensive coverage along with medical payments, optional bodily injury, optional property damage and uninsured motorist coverage.

The liability insurance is sometimes called third party cover. The insurance company is the first party, the vehicle is the second party and the passengers and other peoples' properties is the third party. The bodily injury coverage means the insurance company will pay damages to people injured or killed by your motor. The limits of the coverage are expressed as a fraction, for example 10/20. Coverage such as this would mean that the insurer will pay up to 10,000 ~~for injury~~ to one person per accident or a total of 20,000 for injuries suffered by two or more persons per accident.

The property damage coverage means the insurance company will pay for damages done to other people autos, buildings, trees, etc. Coverage stated as 10,000 per accident means that the insurance company will pay up to 10,000 for any damages to other property. Insurers base their auto insurance rates on the territory lived in, the driver classification, use of auto, how much the car is driven, and age and model of the car.

Motor insurance is probably the largest single class of insurance sold in Ghana. It provides coverage for legal liability to third parties and direct damage to the named vehicle. Motor insurance is actually a combination of three distinctly different coverage, - liability, accident and disability, and property damage.

A standard motor insurance will not cover certain losses, such as your own death or bodily injury due to a motor accident, your liability against claims from passengers in your vehicle (except for passengers of hired vehicles such as taxis and buses) and loss or damage arising from an act of nature, such as flood, storm and landslide. However, you may pay additional premiums to extend your policy to cover flood, landslide, land slip as well as cover your passengers. It is important to check your policy for the exclusions.

2.3.3 Purchase of Motor Insurance

The following gives some important points to consider when buying a motor insurance policy.

- *Insured Value/Sum Insured*: If buying a policy against loss/damage to your vehicle, ensure that your vehicle is adequately insured as it will affect the amount you can claim in the event of loss/damage. For a new vehicle, the insured value will be the purchase price while for other vehicles, the insured value is the market value of the vehicle at the point you apply for the insurance policy.

- *Under-insurance*: If you insure your vehicle at a lower sum than its market value, you will be deemed as self-insured for the difference, i.e. in the event of loss/damage, you will only be partially compensated (up to the proportion of insurance) by your insurance company.
- *Over-insurance*: Should you insure your vehicle at a higher sum than its market value, the maximum compensation you will receive is the market value of the vehicle as the policy owner cannot 'profit' from a motor insurance claim.
- *Duty of Disclosure*: You should disclose fully all material facts, including previous accidents (if any), modification to engines, etc. When in doubt as to whether a fact is relevant or not, it is best to ask your insurance company. If you fail to disclose any material fact, your insurance company may refuse to pay your claim or any claim made by a third party against you. In such cases, you are personally liable for such claims.
- *Price*: The price you pay for your motor insurance will depend on the type of policy selected. The insurance premium charged by your insurance company is the standard minimum rate in accordance with the Motor Tariff. However, in addition to the standard minimum rate, your insurance company may impose additional premiums known as loadings to the premium payable in view of higher risk factors involved such as age of vehicle and claims experience. Loadings are governed by Bank Negara Malaysia (BNM) and no insurance company may charge loadings higher than the levels permitted by BNM.

- *No-claim-discount*: The premium payable may be reduced if you have no-claim-discount (NCD) entitlement. NCD is a 'reward' scheme for you if no claim was made against your policy during the preceding 12 months of policy. Different NCD rates are applicable for different classes of vehicles. For a private car, the scale of NCD ranges from 25% to 55% as provided in the policy.
- *Excess*: Excess is also known as *deductible*. This is the amount of loss you have to bear before your insurance company will pay for the balance of your vehicle damage claim. The types of excess applicable are the *compulsory excess of RM400* and *other excess*. The Compulsory Excess of RM400 a vehicle is driven by a person not named in your policy or a person named in the policy but under the age of 21, the holder of a provisional (L) driving licence or the holder of a full driving licence of less than two years. The other excess is the applicable at the discretion of your insurance company and in some cases, no excess is imposed. You can negotiate with your insurance company on this excess.

With the implementation of e-cover note in 2005, insurance companies will transit motor insurance information electronically to the Road Transport Department (RTD) and you will receive confirmation slip containing details of your motor cover as confirmation of the purchase of your motor insurance. Thereafter, within one month, you should receive:

- the schedule which shows your name and address, details of the vehicle, the sum insured (for comprehensive and third party fire & theft policies), the period of insurance, the policy number, your NCD entitlement, premium breakdown, excess and named drivers;

- the certificate of insurance which shows your name, vehicle model, registration number and cubic capacity, period of insurance, authorised drivers and limitations of use. In some cases, this may be issued at the point of purchase in place of the cover note; and a motor policy which shows the terms and conditions of cover provided by your insurance company. If you do not receive your policy within one month, you should check with your insurance company.

2.4 *INSURANCE CLAIM PROCEDURE*

In the event of an accident/loss:

- Take notes of the accident, that is, the names and addresses of all drivers and passengers involved, vehicle registration numbers, make and model of each vehicle involved, the drivers' license numbers and insurance identification as well as the names and addresses of as many witnesses as possible.
- Lodge a police report within 24 hours of a road accident. This is required by law.
- Notify your insurance company in writing with full details as soon as possible. Depending on the type of claim you intend to make, you may have to notify the other insurance company. If you fail to report the accident, you will be liable for your own loss as well as any third party claim against you.
- Send the damaged vehicle to a workshop approved by your insurance company. If the accident occurs during office hours, you may call the hotline/emergency assistance numbers provided by your insurance company. Otherwise, you may call your insurance company for the nearest approved workshop. Should the accident occur outside office hours and you are making a claim against your own

policy or a third party claim, you should ensure that the vehicle is towed to an approved workshop.

When involved in an accident, you may either make an *own damage claim* or a *third party claim* against the insurance company.

- The own damage claim refers to making a claim on your own insurance policy. That is, you have a comprehensive policy. However, you will lose your NCD entitlement. In notifying your insurance company of the accident, enquire about the names of approved workshops to send your vehicle for repair. Submit the fully completed Motor Accident Report Form together with all supporting documents as soon as possible to your insurance company. The workshop will commence repairs on your vehicle upon the approval of your insurance company. Upon completion of repairs, you will be informed by the workshop to collect your vehicle.
- For the third party claim you may make a third party claim if you are not the party at fault in the accident and you can retain your NCD entitlement. There are two ways of making a third party claim, i.e. submit the claim directly to the insurance company of the party at fault or, if you have a comprehensive policy, submit the claim to your insurance company. You are encouraged to submit your claim to your own insurance company for speedier claims processing. As the third party claimant, you are required to mitigate your loss (i.e., you must act to minimize your loss). Appoint a licensed adjuster to assess the loss. The workshop or the third party insurance company may advise you on this. Submit the adjuster's report and the fully completed Motor Accident Report Form together

with all supporting documents as soon as possible. You are eligible to claim from the third party insurance company for `compensation for actual repair time' (CART) and compensation of excess. For the actual repair time of your vehicle, this is based on the adjuster's recommendation on the number of days required for your car to be repaired. Insurance companies, at their own discretion, may allow an additional seven working days for any unforeseen or unavoidable delay.

- The theft claims can also be made after submission of the claim form, you must cooperate fully with your insurance company or its representative during the course of investigation of the theft claim. In view that the police and your insurance company will require time to investigate your claim, you will receive the offer of settlement from your insurance company within six months from the theft notification or upon completion of police investigations, whichever is earlier.

Generally all insurance policies contain a condition, which lays down the procedure to be adopted in the event of a loss. An insured which fails to comply with this procedure places him/herself at the mercy of the company. Normally the Claims Department is established in an insurance company to see to the claims of the insured. Claims settlement as they apply broadly to all classes of insurance entail the following procedures

- *Notification:* You will notice from all insurance policies that the insured is required to notify the company of any occurrence that might give rise to a claim under the policy. This notice must be sent to the company immediately. Some policies stipulate that the notice must be sent within a specified number of days.

Failure to give the notice within the stipulated period constitutes a breach of the terms of the policy by virtue of which the company could repudiate liability.

The notice could be in writing or by telephone. Most often the insured or his representative calls to report an accident or a loss. On the receipt of notice of an incident, the Claims Department must ascertain whether the policy is in force at the material time and that it covers the occurrence, which is the subject of the claim. If satisfied, a claim file is opened and a number is allocated which would feature on all correspondence pertaining to the claim.

An estimate is allocated for the claim and a claim form is dispatched to the insured for completion. The notification enables the company to make provision for the loss, take the pre-requisite steps to ascertain the course and the extent of the loss. It makes it possible for immediate steps to be taken to mitigate the consequences of the loss. When the completed claim form is returned then the stage is then set for the investigation of the claim. In some cases it is even necessarily for the investigation to commence before the completion of the form.

- *Proof of Loss:* It is the responsibility of the insured to prove his loss and he must do this by satisfying the company that the loss has occurred and that is a loss caused by an insured peril. If the insured proves a loss within the provision of the policy any allegation by the company that the loss was caused by an expected peril must be substantiated by the company.
- *Settlement:* When a loss is proved and the amount of the loss is ascertained then the claim is settled. Claims are usually settled by cheque - a crossed cheque for

that matter, and normally marked account payee only. This is done after a discharge has been obtained from the insured. Legally the insured is entitled to be paid in cash but as what could be called the tradition of the trade, cheques have become the accepted mode of payment.

- *Ex-Gratia Payments:* These payments are made out of grace most often the company make payments to its policyholders or other claimants in respect of losses for which the company is not legally liable. These payments are made either in cases of hardship and out of sympathy for the insured or claimant as a means of promoting the reputation of the company for fair dealings.

The principle of indemnity is very crucial in insurance claim. Based on this principle, the insurance cover will compensate your loss by putting you back to the same position you were in immediately before the loss. As you will be compensated only for the loss suffered, you cannot 'profit' from a motor insurance claim. Therefore, if your vehicle is more than five years old, betterment will apply. Betterment occurs when in the course of repairing an accident-damaged vehicle, an old part is replaced with a new franchise part. In line with the principle of indemnity, you will have to bear the difference in costs as you are in a better position after the accident with the new franchise part. However, the application of betterment is at the discretion of your insurance company. Should your insurance company apply betterment, it will be in accordance with the standard scale of betterment adopted by the industry as follows.

CHAPTER THREE

REVIEW OF BASIC MODELS

3.1 REVIEW OF PROBABILITY MODELS

3.1.1 Probability Models for Claims Estimation

In situations where each policyholder is observed over a period of time, longitudinal data model could be used to estimate the frequency of claims. This consists of repeated observations of individual units that are followed over time. Each individual is assumed to be independent of the others but correlation between observations relating to the same individual is permitted. Here, it is assumed that the number of claims per year obeys to a Poisson distribution with a parameter specific to each policyholder. Specifically, $N_{i,t}$ is assumed to be Poisson distributed with mean $\theta_i \lambda_{i,t}$ $i = 1, \ldots, T$. The expected annual claim frequency is a product of a static factor θ_i and a dynamic factor $\lambda_{i,t}$ $(\theta_i \lambda_{i,t})$. The former accounts for the dependence between observations relating to the same insured. The latter introduces the observable characteristics (that are allowed to vary with time). In general, $\ln \lambda_{i,t}$ is expressed as a linear combination of the observable characteristics, that is $\lambda_{i,t} = \exp(\beta_0 + \beta' x_{i,t})$, where β_0 is the intercept, $\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$ is a vector of regression parameters for explanatory variables $x_{i,t} = (x_{i,t1}, x_{i,t2}, \ldots, x_{i,tp})^T$ (Boucher and Denuit, 2005).

Another technique, known as *Accident Year*, embraces the entire population of claims incurred with accidents dates in a particular calendar year, whether reported to the company in that year or subsequently. Under the Accident Year method of establishing claim reserves, there is an automatic segregation of all like accident year claims through successive calendar years as the total population of reported claims for each accident year moves to ultimate settlement status (Masterson, 1960).

Again, the Accident Year Method of establishing aggregate reserves involves procedures for making successive annual estimations of probable settlement costs of remaining unpaid claims in each accident year grouping of claims received. The general method is to determine, for each accident year, the estimated total losses incurred, say (*m* cedis). From this total, *m* cedis is subtracted from the actual total losses paid for each respective accident year giving a balance, which is the claim reserve for unpaid claims (Masterson,1960).

The Poisson-Gamma model assumes the random variable θ_i^{FE} which follows the Gamma distribution with mean 1 and variance $\frac{1}{v}$, the joint probability function of $N_{i,1},\ldots,N_{i,T}$ is given by

$$P\left[N_{i,1}=n_{i,1},\ldots,N_{i,T}=n_{i,T}\right]$$

$$=\left(\prod_{t=1}^{T}\frac{\left(\lambda_{i,t}^{RE}\right)^{n_{i,t}}}{n_{i,t}!}\right)\frac{\Gamma\left(n_{i,\bullet}+v\right)}{\Gamma(v)}\left(\frac{v}{\sum_{t=1}^{T}\lambda_{i,t}^{RE}+v}\right)^{v}\left(\sum_{t=1}^{T}\lambda_{i,t}^{RE}+v\right)^{-n_{i,\bullet}}$$

where $n_{i,1\bullet}=\sum_{t=1}^{T}n_{i,t}$, and $E\left[N_{i,t}\right]=\lambda_{i,t}^{RE}<V\left[N_{i,t}\right]=\lambda_{i,t}^{RE}+\left(\lambda_{i,t}^{RE}\right)^{2}/v.$

The maximum likelihood estimation of parameters and variances can be obtained as follows:

The first order conditions for parameters β_0^{RE} and β^{RE} and v lead to the system:

$$\sum_{i=1}^{n}\sum_{t=1}^{T} n_{i,t} - \lambda_{i,t}\frac{\sum_t n_{i,t}+v}{\sum_t \lambda_{i,t}+v} = 0$$

$$\sum_{i=1}^{n}\sum_{t=1}^{T} x_{i,t}\left(n_{i,t} - \lambda_{i,t}\frac{\sum_t n_{i,t}+v}{\sum_t \lambda_{i,t}+v}\right) = 0$$

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n_{i,\bullet}-1}\frac{1}{j+v}\right) - \log\left(1+\frac{\sum_t \hat{\lambda}_{i,t}}{v}\right) + \sum_t \frac{\lambda_{i,t}+n_{i,t}}{\sum_t \lambda_{i,t}+v} = 0$$

The solution to the above system is obtained by numerical procedures.

The Inverse Gaussian distribution is another good candidate to model the heterogeneity parameter (Willmot (1987), Dean, et al (1989) and Tremblay (1992) for an application with insurance data. Shoukri et al. (2004) showed that the Poisson with Inverse-Gaussian heterogeneity of mean and variance equal to 1 and τ respectively, has the joint probability function

$$P\left[N_{i,1} = n_{i,1} \ldots, N_{i,T} = n_{i,T}\right]$$

$$= \left(\prod_{t=1}^{T}\frac{\left(\lambda_{i,t}^{RE}\right)^{n_{i,t}}}{n_{i,t}!}\right)\left(\frac{2}{\pi\tau}\right)^{0.5} e^{1/\tau}\left(1+2\tau\sum_{t=1}^{T}\left(\lambda_{i,t}^{RE}\right)\right)^{-s_i/2} K_{s_i}(z_i)$$

where $K_j(.)$ is the modified Bessel function of the second kind, $s_i = n_{i,\bullet} - 0.5$, and $z_i = \frac{1}{\tau}\sqrt{1+2\tau\sum_{t=1}^{T}\lambda_{i,t}^{RE}}$

The modified Bessel function of the second kind has some useful properties that can be used to find maximum likelihood estimators (Shoukri et al., 2004).

Now, $V\left[N_{i,t}\right]=\lambda_{i,t}^{RE}+\tau\left(\lambda_{i,t}^{RE}\right)^{2}$ and the maximum likelihood estimators of β_0^{RE} and β^{RE} are solved by the equation:

$$\sum_{i=1}^{n}\sum_{t=1}^{T} n_{i,t}-\lambda_{i,t}M\left(n_{i,\cdot}\right)=\sum_{i=1}^{n}\sum_{t=1}^{T} x_{i,t}\left(n_{i,t}-\lambda_{i,t}M\left(n_{i,\bullet}\right)\right)=0$$

where the function M(.) can be expressed as the ratio of the modified Bessel function of the second kind.

$$\frac{K_{n_{i,\bullet}+\frac{1}{2}(a)}}{K_{n_{i,\bullet}-\frac{1}{2}(a)}}=M\left(n_{i,\bullet}\right)\sqrt{1+2\tau\sum_{t=1}^{T}\lambda_{i,t}^{RE}}.$$

The estimator of the parameter τ canbe found by solving

$$-\frac{1}{2\tau}-\frac{1}{\tau^{2}}-\frac{s_1\sum_t \lambda_{i,t}}{1+2\tau\sum\lambda_{i,t}}+\frac{\partial \log K_{s_1}\left(z_t\right)}{\partial T}=0$$

where the derivative of the function K is equal to:

$$\frac{\partial \log K_{s_1}\left(z_t\right)}{\partial T}=-M\left(n_{i,\bullet}\right)\sqrt{1+2\tau\sum_t \lambda_{i,t}}+\frac{\left(n_{i,\bullet}-\frac{1}{2}\right)\tau}{\sqrt{1+2\tau\sum_t \lambda_{i,t}}}$$

$$\frac{\partial z_1}{\partial T}=\frac{\sum_t \lambda_{i,t}}{\tau\sqrt{1+2\tau\sum_t \lambda_{i,t}}}-\frac{\sqrt{1+2\tau\sum_t \lambda_{i,t}}}{\tau^{2}}$$

Again, numerical procedures are needed to obtain the solutions.

In the fixed effects model, all characteristics that are not time-varying are captured by the individual heterogeneity term θ_i^{FE}. In our case, the intercept β_0 has to be removed (and 6 is included in θ_i^{FE}). As sex of the driver has been removed from the model, the remaining explanatory variables do vary with time and enter the fixed effects model.

The Poisson fixed effects model has been proposed by Palmgren (1981) and Hausman et al. (1984). The standard way of evaluating the parameters of this model is the conditional maximum likelihood of Andersen (1970). The idea of the conditional method is to obtain an estimator of β^{FE} without having to estimate each θ_i^{FE}.

As proved in Cameron and Trivedi (1998), the maximum likelihood and conditional maximum likelihood estimation methods always yield identical estimates for covariates parameter β^{FE} in case of Poisson distribution. Specifically, the estimated parameters are obtained by solving the equation:

$$\sum_{i=1}^{n}\sum_{t=1}^{T} x_{i,t}\left(n_{i,t} - \lambda_{i,t}^{FE}\frac{\sum_t n_{i,t}}{\sum_t \lambda_{i,t}^{FE}}\right) = 0$$

Note that only insureds with varying characteristics (and at least one claim) are used in the estimation of β^{FE}. Estimates of θ_i^{FE} can then be obtained from

$$\theta_i^{FE} = \frac{\sum_t n_{i,t}}{\sum_t \lambda_{i,t}^{FE}}$$

Both estimates of β^{FE} and θ_i^{FE} are consistent when $T \to \infty$ and $N \to \infty$, but only the estimate of β^{FE} is consistent for fixed T and $N \to \infty$, as for insurance data.

3.1.2 Probability Models for Premium Calculation

The problem of premium calculation is one of the key issues in the insurance business if the premium rate is too high, an insurance company will not have enough clients for successful operation. If the premium rate is too low, the company also may not have sufficient funds to pay all the claims. To study this problem we need some of the following basic and widely used probability distribution functions. The techniques reviewed in section 2.5.1 are all variations of these basic probability models.

- The Poisson distribution,

  $$P(\{\omega : X = x\}) = e^{-\lambda}\frac{\lambda^{x}}{x!}, x = 0, 1, 2, \ldots$$

  often used for modeling the number of claims.

- The Binomial distribution,

  $$P(\{\omega : X = x\}) = {}^{n}C_{x}\, p^{x}(1-p)^{m-x},\ x = 0, 1, 2, \ldots, m$$

  used for the number of claims for a portfolio of m independent policies, where p is probability of receiving a claim (if $m = 1$, then it is called Bernoulli distribution);

- The Exponential distribution,

  $$P(\{\omega : X \leq x\}) = 1 - e^{-\lambda x}, x \geq 0, \lambda > 0.$$

  has various applications, for example, models the distribution of jumps of a Poisson process with intensity λ;

- The Pareto distribution,

$$P(\{\omega : X \le x\}) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha}, x \ge 0, \alpha > 0, \lambda > 0,$$

has a 'heavy' tail, and hence is often used in modeling large claims;

- The Lognormal distribution,

$$P(\{\omega : X \le x\}) = \int_0^x \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(\log x - \mu)^2} dx,$$

Often used in the distribution of the risk process.

- The Gamma distribution,

$$P(\{\omega : X \le x\}) = \int_0^x \frac{\beta^{\alpha}}{\Gamma(a)} x^{\alpha-1} e^{-\beta x} dx, \beta > 0;$$

- The Normal distribution,

$$P(\{\omega : X \le x\}) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx;$$

3.1.3 Distribution of Risk Process

The distribution of the risk process is denoted by

$$F_X(X(t)) = P\left(\{\omega : X \leq x\}\right) = P\left(\{\omega : \sum_{t=1}^{N(t)} X_t \leq x\}\right),$$

To compute $F_X(X(t))$ one needs some additional assumptions. Usually processes, X_n and $N(\cdot)$ are assumed to be independent. Then we can write

$$F_x(t) = p(\{\omega : X \leq x\}) = \sum_{k=0}^{\infty} pk(t) F_x(x), \; where$$

$$F_x(x) = P(\{\omega : X_1 + \ldots + X_k \leq x\}).$$

Premium calculation of process $\pi(t)$ is one of the most essential and complex tasks of an insurer. Premium flow must guarantee payments of claims, but on the other hand, premiums must be competitive. One of the most widely used ways of computing π on interval $[0, t]$ is given by

$$\pi(t) = (1+\theta)E(N(t))E(X)$$

where X is a random variable with the same distribution as X_i and θ is the *security loading coefficient*. This formula says that the average premium income should be greater than the average aggregate claims payment. If they are equal, then such premium is called *net-premium* and the method of its computing is referred to as *equivalence principle*.

The *bonus-malus* system is an example of a different approach to premium calculations. In this case, all policy holders are assigned certain ratings according to their claims history, and they can be transferred from one group to another. This system is typically used by the motor vehicle insurance companies. Calculation of adequate premium consists in construction of process $\pi(t)$ given $F(X(t))$, the distribution function of the risk process. The process, π which will be written $\pi(x)$, has the following properties:

- $\pi(a) = a$, for any constant a if $\theta = 0$;
- $\pi(aX) = a\pi(X)$, for any constant a;
- $\pi(X+Y) \leq \pi(X) + \pi(Y)$
- $\pi(X+a) = \pi(X) + a$, for any constant a;
- *If* $X \leq Y$, then $\pi(X) \leq \pi(Y)$
- for any $p \in [0,1]$ and any random variable $\pi(X) = \pi(Y)$, implies that

  $\pi(pFX + (1-p)FZ) = \pi(pFY + (1-p)FZ)X)$

3.2 LOGISTIC REGRESSION MODEL

3.2.1 Introduction

The generalised linear regression (GLR) models together with the statistical package R have revolutionised graduate statistics and the carrying out of data analysis in practice. The (GLR) models are a natural generalisation of the familiar classical linear models. The class of GLR models includes, as special cases, linear regression, analysis-of-variance models, log-linear models for the analysis of contingency tables, logit models for binary data in the form of proportions and many others.

The use of classical linear models in actuarial work is not new. Thus, such models have become an established part of the description of claim frequency rates and average claim costs in motor insurance - as evidenced by a number of papers in the British actuarial literature, including (Mullahy, 2001). However, the use of generalised linear models in actuarial work is relatively new. Thus, McCullagh and Nelder(1989) in their excellent and comprehensive monograph give a number of examples of the fitting of generalised linear models to different types of data. One of these relates to data from Baxter et al on the average claim costs in a motor insurance portfolio (originally modeled by Baxter et al using a weighted least squares approach).

In this thesis we apply the use of generalised linear model, specifically, logistic regression model to the estimation of motor insurance premiums claims. This is being done using age of car, cubic capacity of vehicles, type of vehicle and sex of policyholder as the predictor or influential variables of a claim.

3.2.2 Definitions of Logistic Regression Model

The logistic regression model is a generalized linear model with two components namely, random and systematic. The random component is the response variable which is binary variable, Y_i=1 or 0 (claim made or claim not made). In this case, we are interested in the probability that Y_i=1, (that is, $\pi(x_i) = P(Y_i = 1)$), which has the Binomial distribution. The systematic component is a linear predictor, $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_q x_q$. The explanatory variables may be quantitative (continuous), qualitative (discrete), or both (mixed). The outcome variable in this analysis is claim. The fitted regression models takes the form:

$$\log it(\pi(x_i)) = \log\left(\frac{\pi(x_i)}{1-\pi(x_i)}\right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_q x_q. \tag{3.1}$$

Clearly the above is the log of the odds the response taking the value one. The above model could be rewritten as

$$\pi(x_1, x_2, \ldots, x_q) = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_q x_q)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_q x_q)} \tag{3.2}$$

The logit function can take any real value, but the associated probability always lies in the required [0, 1] interval. In a logistic regression model, the parameter β_j associated with explanatory variable x_j is such that $\exp(\beta_j)$ is the odds that the response variable takes the value one when x_j increases by one, conditional on the other explanatory variables remaining constant. The regression coefficients, $\beta = \left(\beta_0, \beta_1, \beta_2, \ldots, \beta_q\right)^T$, which are the parameters of the logistic regression model estimated by maximum likelihood method.

In this thesis the response and predictor variables are defined as follows, considering the problem under study of which the logistic regression model is being applied variables:

$$Y_i = \begin{cases} 1, \textit{if a claim has been made} \\ 0, \textit{if a claim has not been made} \end{cases}$$

$X_i\text{'s} = \textit{risk factors associated with the claims.}$ These are sex of a policyholder, vehicle age, make of a vehicle and vehicle cubic capacity

The standard multiple linear regression model is inappropriate to model this data for the following reasons:

- The model's predicted probabilities could fall outside the range 0 to 1.
- The dependent variable is not normally distributed. It follows the Binomial probability model.

- If we consider the normal distribution as an approximation for the binomial model, the variance of the dependent variable is not constant across all risk variables.

The logistic regression model is therefore developed to account for all these difficulties.

3.2.3 Odds Ratio

The odds ratio compares the likelihood of an event between two groups. Consider the following data on survival of passengers on the Titanic. There were 462 female passengers: 308 survived and 154 died. There were 851 male passengers: 142 survived and 709 died (see table below).

Table 3.1 Survival Data

Sex	Alive	Dead	Total
Female	308	154	462
Male	142	709	851
Total	450	863	1,313

Clearly, a male passenger on the Titanic was more likely to die than a female passenger, but how much more likely? The odds ratio compares the relative odds of death in each group. For females, the odds were exactly 2 to 1 against dying (154/308=0.5). For males, the odds were almost 5 to 1 in favor of death (709/142=4.993). The odds ratio is 9.986 (4.993/0.5) which means that there is about ten folds greater odds of death for males than for females.

The logistic regression model

$$\log it\left(\pi(x)\right) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x \tag{3.3}$$

can be written as a multiplicative model by taking the exponential of both sides, thus

$$\exp\left(\alpha + \beta x\right) = e^{\alpha} e^{\beta x}$$

This gives us a model for odds, where the odds increase multiplicatively with X. A unit increase in X leads to an increase in the odds of $\exp(\beta)$.

So the odds ratio for a unit increase in x equals

$$\{[\pi(x+1)/(1-\pi(x+1))]/[\pi(x)/(1-\pi(x))]\} = e^{\beta} \tag{3.4}$$

When $\beta = 0$, $e^{0} = 1$ and so the odds do not change with X. The logarithm of the odds changes linearly with X.

3.2.4 Statistical Model of Claim

A closer look at statistical modeling of binary response variables for which the response measurement for each subject is "success" or "failure". Binary data are perhaps the most common form of categorical data and the methods are of fundamental importance. The most popular model for binary data is logistic regression. For a binary response Y and a quantitative explanatory variable X, let π (x) denote the "success" probability when X takes value x. This probability is the parameter for the binomial distribution. The logistic regression model has linear form for the logit of this probability

$$\pi(x) = \log\left(\frac{\pi(x)}{1-\pi(x)}\right) = \alpha + \beta x \qquad (3.5)$$

The formula implies that $\pi(x)$ increases or decreases as an S-shaped function of x. An alternative formula for logistic regression refers directly to the success probability. This formula uses the exponential function exp (x) = e^x in the form

$$\pi(x) = \frac{\exp(a + \beta x)}{1 + \exp(a + \beta x)}$$

(3.6)

The parameter β determines the rate of increase or decrease of the S- shaped curve. The sign of β indicates whether the curve ascends or descends, and the rate of change increases as | β | increases. When the model holds with β = 0 the right- hand side of equation (3.5) simplifies to a constant. Then π (x) is identical at all x so the curve becomes a horizontal straight line. The binary response Y is then independent of X .

3.2.5 Confidence Intervals for Effects and Significance Testing

A large-sample confidence interval for the parameter β in the logistic regression model, $\log it(\pi(x)) = \alpha + \beta x$ is $\hat{\beta} \pm Z_{\alpha/2}.se(\hat{\beta})$

Exponentiating the endpoints of this interval yields one(1) for e^{β}, the multiplicative effect on the odds of a unit increase in x.

We next discuss significance test for the effect of x on the binary response. For the logistic regression model, the null hypothesis $H_0 : \beta = 0$ states that the probability of success is independent of x. For large samples, the test statistic $Z = \frac{\hat{\beta}}{ASE(\hat{\beta})}$, which has the standard normal distribution when $\beta = 0$. One can refer Z to the standard normal table to get a one-sided or two-sided *p-value* in the usual manner. Equivalently, for the two-sided alternative $H_1 : \beta \neq 0$, $Z^2 = \left(\frac{\hat{\beta}}{ASE(\hat{\beta})} \right)^2$ becomes the Wald statistic having a large-sample chi-squared distribution with degrees of freedom 1.

Though the Wald test works well for very large samples, the likelihood-ratio test is more powerful and reliable for sample sizes used in practice. The test statistic compares the maximum L_0 of the log-likelihood function when $\beta = 0$ (thus, $\text{logit}(\pi(x)) = \alpha$) is when (i.e. π (x)) is forced to be identical at all x values) to the maximum L_1 of the log-likelihood function for unrestricted β (thus, $\text{logit}(\pi(x)) = \alpha + \beta$). The test statistic -2(L_0 $-L_1$), also has a large-sample chi-squared distribution with degree of freedom $(df) = 1$. Most software for logistic regression reports the maximized log- likelihood L_0 and L_1. and the likelihood-ratio statistic derived from those maxima.

3.2.6 Probability Distribution of Estimates

The estimated probability that $Y = 1$ at fixed setting x of X equals

$$(\hat{\pi}(x)) = \frac{\exp(\hat{\alpha} + \hat{\beta}x)}{1 + \exp(\hat{\alpha} + \hat{\beta}x)} \tag{3.7}$$

Most software for logistic regression can report such estimates as well as confidence intervals for the true probabilities. One can construct confidence intervals for the probabilities using the covariance matrix of the model parameter estimates. The term $(\hat{\alpha} + \hat{\beta}x)$ in the exponents of the prediction equation (3.7) is the estimated linear predictor in the logit transform (3.6) of $\pi(x)$. This estimated logit has large-sample ASE given by the estimated square root of

$$Var((\hat{\alpha}+\hat{\beta}x)=Var((\hat{\alpha})+x^2Var(\hat{\beta})+2xCov(\hat{\alpha},\hat{\beta}) \quad (3.8)$$

A 95% confidence interval for the true logit is $(\hat{\alpha}+\hat{\beta}x)\pm 1.96ASE(\hat{\beta})$. Substituting the endpoints of this interval for $\alpha + \beta x$ in the exponents gives a corresponding interval for the probability. One could ignore the model fit and simply use sample proportions to estimate such probabilities. When the logistic regression model truly holds, the model-based estimator of a probability is considerably better than the sample proportion. The model has only two parameters to estimate, whereas the non model based approach has a separate parameter for every distinct value of X. The estimated standard error is

$$\sqrt{\frac{p(1-p)}{n}}.$$

In practice, any model will not exactly represent the true relationship between $\pi(x)$ and x. Thus, as the sample size increase, the model-based estimator may not converge exactly to the true value of the probability. This does not imply, however that the sample proportion is actually a better estimator in practice. If the model approximates the true probabilities decently, its estimator still ends to be much closer than the sample proportion to the true value. The model smoothens the sample data, somewhat dampening the observed variability. The resulting estimators tend to be better unless each sample proportion is based on an extremely large sample. If the logistic regression model approximates well the true dependence of $\pi(x)$ on x, point interval estimates of $\pi(x)$ based on it are quite useful. (Nelder,et al (1989))

3.2.7 Goodness of Fit and Likelihood-Ratio Model Comparison Test

The likelihood-ratio statistic $-2(L_0 - L_1)$ for testing whether certain parameters in a model equal zero. The test compares the maximized log likelihood (L_1) for the model to the maximized log likelihood (L_0) for the simpler model that deletes those parameters. Denote the fitted model by M_1 the simpler model for which those parameters equal zero by M_0.The goodness-of- fit statistic G^2 for testing the fit of a logistic regression model M is the special case of the likelihood-ratio statistic in which M_0 = M and M_1 is the most complex model possible. That complex model has a separate parameter for each logit and provides a perfect fit to the sample logits. It is called the saturated model. In testing whether M fits, we test whether all parameters that are in the saturate model are not in M equal zero. Denote this statistic for testing the fit of M by G^2(M). In GLM terminology, it is called the deviance of the model. Let L_s denote the maximized log likelihood for the saturated model. Then, the deviances for models M_0 and M_1 are G^2 (

$$M_0) = -2(L_0 - L_s) \text{ and } G^2(M_1) = -2(L_0 - L_s) \tag{3.9}$$

Denote the likelihood-ratio statistic for testing M_0, given that M_1 holds, by $G^2(M_0|M_1)$. This statistic for comparing this model equals

$$G^2(M_0|M_1) = -2(L_0 - L_s) - (-2(L_0 - L_s)) = G^2(M_0) - G^2(M_1) \tag{3.10}$$

a difference in G^2 goodness-of –fit statistics for the two models. That is, the likelihood-ratio statistic for comparing ~~two models~~ is simply the difference in the variances of those models. This statistic is large when M_0 fits poorly compared to M_1 . It is a large sample chi-squared statistic, with *df* equal to the difference between residual *df* values for two models.

The likelihood-ratio method can be used to test hypotheses about parameters in multiple logistic regression models. More generally, one can compare maximized log- likelihood for any pair of models such that one is a special case of the other.

3.2.8 Residuals for Logit Models

Goodness-of fit statistics as G^2 and X^2 are summary indicators of the overall quality of fit. Additional diagnostic analyses are necessary to describe the nature of any lack of fit. Residuals comparing observed and fitted counts are useful for this purpose. Let y_i denote the number of "successes" for n_i trials at the *ith* setting of the explanatory variables. Let $\hat{\pi}$ denote the predicted probability of success for the model fit. Then $n_i\hat{\pi}$ is the fitted number of success. For a GLM with binomial random component, the

Pearson residual for the fit at setting is

$$e_i = \frac{y_i - n_i\hat{\pi}_i}{\sqrt{n_i\hat{\pi}_i(1-\hat{\pi}_i)}} \tag{3.11}$$

Each residual divides the difference between an observed count and its fitted value by the estimated binomial standard deviation of the observed count. The Pearson statistic for testing the model fit satisfies $X^2 = \sum_{i=1} e_i^2$.

Each squared Pearson residual is a component of X^2. When the binomial index n_i is large, the Pearson residual e_i has an approximate normal distribution. When the model holds, it has an approximate expected value of zero but a smaller variance than a standard normal variate. If the number of model parameters is small compared to the number of sample logits, Pearson residuals are treated like standard normal deviates, with absolute values larger than 2 indicating possible lack of fit. Graphical displays are also useful for showing lack of fit one can compare observed and fitted proportions by plotting them against each other, or by plotting both of them against explanatory variables. We have noted that X^2 and G^2 are invalid when fitted values are very small. Similarly residuals have limited meaning in that case. When explanatory variables are continuous, often n_i = 1 at many settings. Then, y_i can equal only 0 or 1, and e_i can assume only two values. One must then be cautious about regarding either outcome as "extreme" and single residual is usually uninformative.

3.2.9 Logit Models for Qualitative Predictors

Logistic regression, like ordinary regression, extends to models incorporating multiple explanatory variables. Moreover, some or all of those explanatory variables can be qualitative, rather than quantitative. Suppose that a binary response Y has two binary predictors, X and Z. If the two levels for each variable are denoted by (0, 1), then $\log it(\pi) = \alpha + \beta_1 X + \beta_2 Z$ has separate main effects for the two predictors. It assumes an absence of interaction, the effect of one factor being the same at each level of the other factor. The variables x and z in this model are dummy variables that indicate categories for the predictors. At a fixed level z of Z, the effect on the logit of changing from $x = 0$ to $x = 1$ is $[\alpha + \beta_1(1) + \beta_2 z] - [\alpha + \beta_1(0) + \beta_2 z] = \beta_1$

This difference between two logits equals the difference of log odds, which equals the log of the odds ratio between X and Y at a fixed level of Z. Thus, $\exp(\beta_i)$ describes the conditional odds ratio between X and Y. Controlling for Z, the odds of "success" at x = 1 equal $\exp(\beta_1)$ times the odds of success at $x = 0$. This conditional odds ratio is the same at each level z of Z.

3.2.10 Model Selection and Test for Adequacy

Several model selection procedures exist, no one of which is "best". Cautions that apply to ordinary regression modeling of normal data hold for any generalized linear model. For instance, a model with several predictors has the potential for multicollinearity, strong correlations among predictors, making it seems that no one variable is important when all the others are in the model. A variable may seem to have little effect simply because it "overlaps" considerably with other predictors in the model.

The following criteria are used to test for the model adequacy:

- *Deviance Residuals*: *Deviance residual* is the measure of deviance contributed from each observation and is given by $r_{D_i} = sign(r_i)\sqrt{d_i}$, where d_i is the individual deviance contribution. The deviance residuals can be used to check the model fit at each observation for generalized linear models. The standardized and studentized deviance residuals are

$$r_{Dsi} = \frac{r_{Di}}{\sqrt{\hat{\phi}(1-h_i)}}, \text{ and } r_{Dti} = \frac{r_{Di}}{\sqrt{\hat{\phi}_{(i)}(1-h_i)}} \tag{3.12}$$

- *Akaike's Information Criterion (AIC)*: This is used for choosing between competing statistical models. For, categorical data like in our case, this amounts to choosing the model that minimizes $G^2 - 2v$, where G^2 is the likelihood-ratio goodness-of-fit statistic and v is the number of degrees of freedom associated with the model.

- *Confidence Interval*: Confidence intervals are used to indicate the reliability of an estimate. How likely the interval is to contain the parameter is determined by the confidence level

3.3 THEORY OF ESTIMATION OF PARAMETERS.

3.3.1 Introduction

Suppose we have a random variable X which has probability density function, $f(x,\theta)$ where θ is either a real number or a vector of real numbers. Assume that $\theta \in \Omega$ which is a subset of R^p, for $p \geq 1$. For example, θ could be the vector (μ, σ^2) when X has a $N(\mu, \sigma^2)$ distribution or θ could be the probability of success p when X has a binomial distribution. Our information about θ comes from the sample $X_1, X_2, \ldots, X_n$. We often assume that this is a random sample which means that the random variables $X_1, X_2, \ldots, X_n$ are independent and have the same distribution as X; that is, $X_1, X_2, \ldots, X_n$ are independently and identically distributed (*iid*). We may use the statistic $T = T(X_1, X_2, \ldots, X_n)$, a function of the sample to estimate θ and say that T is a point estimator for θ. For example, suppose $X_1, X_2, \ldots, X_n$ is a random sample from a distribution with mean μ and variance, σ^2. Then the statistics $\overline{X}$ and S^2 (the sample mean and variance of this random sample) are point estimators of μ and σ^2 respectively.

In order to decide which point estimator of a parameter is the best one to use, we need to examine their statistical properties and develop some criteria for comparing estimators and also gives the most economical information. Ideally, we want an estimator which generates estimates that can be expected to be close in value to the parameter. These properties are presented in Section 3.3.3.

The two of the most commonly used approaches to the statistical estimation of parameters are the least squares method and method of maximum likelihood. The later is usually used for generalized linear models, where the estimates are usually obtained numerically by an iterative procedure which turns out to be closely related to weighted least squares estimation. This method would be used to estimate the parameters of the logistic regression for the motor insurance claim.

The section will then develop statistical inference (estimation and testing) based on likelihood methods. We show that these procedures are asymptotically optimal under certain conditions (regularity condition).

3.3.2 Method of Maximum Likelihood

Suppose that $x_1, x_2, \ldots, x_n$ are independent and identically distributed (iid) random variables with common probability density function, $f(x;\theta), \theta \in \Omega$. The basis of our inferential procedures is the likelihood function given by,

$$L(\theta, x) = \prod_{i=1}^{n} f(x_i;\theta), \quad \theta \in \Omega, \tag{3.13}$$

where $X = (x_1, x_2, \ldots, x_n)^T$, and the likelihood function which is a function of θ is simply denoted by $L(\theta)$. The maximum likelihood estimator $(\hat{\theta})$ of the parameter θ is obtained by maximizing $L(\theta)$. Usually, for mathematical convenience, we rather work with $l(\theta) = \log L(\theta)$, which interestingly gives us no loss of information in using $l(\theta)$ because the log is a one-to-one function. Thus

$$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta), \quad \theta \in \Omega. \tag{3.14}$$

For example, given the *iid* random sample, $x_1, x_2, \ldots, x_n$ from the logistic density ,

$$f(x,\theta) = \frac{\exp\{-(x-\theta)\}}{(1+\exp\{-(x-\theta)\})^2}, \quad -\infty < x < \infty, -\infty < \theta < \infty.$$

the log of the likelihood simplifies to

$$l(\theta) = \sum_{i=1}^{n} \log f(x_i;\theta) = n\theta - n\bar{x} - 2\sum_{i=1}^{n} \log(1+\exp\{-(x_i-\theta)\}).$$

Using this, the first partial derivative is

$$l^1(\theta) = n - 2\sum_{i=1}^{n} \frac{\exp\{-(x_i-\theta)\}}{1+\exp\{-(x_1-\theta)\}}.$$

Setting this equation to 0 and rearranging terms results in the equation,

$$\sum_{i=1}^{n} \frac{\exp\{-(x_i-\theta)\}}{1+\exp\{-(x_i-\theta)\}} = \frac{n}{2}$$

The above has a unique solution. The derivative of the left side of the above equation simplifies to,

$$\frac{\partial}{\partial\theta}\left(\sum_{i=1}^{n}\frac{\exp\{-(x_i-\theta)\}}{1+\exp\{-(x_i-\theta)\}}\right)=\sum_{i=1}^{n}\frac{\exp\{-(x_i-\theta)\}}{(1+\exp\{-(x_i-\theta)\})^2}>0.$$

Thus the left side of equation $\sum_{i=1}^{n}\frac{\exp\{-(x_i-\theta)\}}{1+\exp\{-(x_i-\theta)\}}.=\frac{n}{2}$ is a strictly increasing function of θ. Finally, the left side of the same equation approaches 0 as $\theta\rightarrow-\infty$ and approaches n as $\theta\rightarrow\infty$. Thus the equation $\sum_{i=1}^{n}\frac{\exp\{-(x_i-\theta)\}}{1+\exp\{-(x_i-\theta)\}}=\frac{n}{2}$ has a unique solution. Also the second derivative of $l(\theta)$ is strictly negative for all θ; so the solution is a maximum.

Also, *iid* random sample, $x_1,x_2,...,x_n$ from the random variable x which has the Bernoulli distribution gives an unbiased maximum likelihood estimator, $\hat{p}=\frac{x}{n}$ given as follows:

$$f(x,p)=p^x(1-p)^{1-x}$$

$$L(x_1,...,x_n,p)=p^{\sum x_i}(1-p)^{n-\sum x_i}$$

$$InL(x_1,...,x_n,p)=\left(\sum x_i\right)Inp+\left(n-\sum x_i\right)In(1-p)$$

$$\frac{dL}{dp}(x_1,...,x_n,p)=\frac{\sum x_i}{p}-\frac{\left(n-\sum x_i\right)}{1-p}$$

$$\Rightarrow\frac{\sum x_i}{p}=\frac{n-\sum x_i}{1-p}$$

$$(1-p)\sum x_i = p\left(n-\sum x_i\right)$$

$$\sum x_i - p\sum x_i = np - p\sum x_i$$

$$np = \sum x_i$$

$$\hat{p} = \frac{1}{n}\sum x_i$$

3.3.3 Concept of Likelihood

If the probability of an event X dependent on model parameters p is written $P(X|\theta)$, then we would talk about the likelihood, $L(\theta|X)$, that is, the likelihood of the parameters given the data. For most sensible models, we will find that certain data are more probable than other data. The aim of maximum likelihood estimation is to find the parameter value(s) that makes the observed data most likely. This is because the likelihood of the parameters given the data is defined to be equal to the probability of the data given the parameters. We are much more interested in the likelihood of the model parameters that underlie the fixed data. For example, how would we go about this in a simple coin toss experiment? That is, rather than assume that p is a certain value (0.5) we might wish to find the maximum likelihood estimate *(MLE)* of p, given a specific dataset. Beyond parameter estimation, the likelihood framework allows us to make *tests* of parameter values. For example, we might want to ask whether or not the estimated p differs *significantly* from 0.5 or not. This test is essentially asking: is there evidence that the coin is biased? We will see how such tests can be performed when we introduce the concept of a *likelihood ratio test* in Section 3.3.5.

Say we toss a coin 100 times and observe 56 heads and 44 tails. Instead of *assuming* that p is 0.5, we want to find the MLE for p. Then we want to ask whether or not this value differs significantly from 0.50. How do we do this? We find the value for p that makes the observed data most likely. As mentioned, the observed data are now fixed. They will be constants that are plugged into our binomial probability model:

- $n = 100$ (total number of tosses), $x = 56$ (total number of heads)

If $p = 0.5$, then our binomial probability model gives:

$$L(p = 0.5 \mid data) = \frac{100!}{56!44!} 0.5^{56} 0.5^{44} = 0.0389$$

But if $p = 0.52$ instead, the model gives:

$$L(p = 0.52 \mid data) = \frac{100!}{56!44!} 0.52^{56} 0.48^{44} = 0.0581$$

So from this we can conclude that p is more likely to be 0.52 than 0.5. Tabulating the likelihood for different parameter values to find the maximum likelihood estimate of p:

Table 3.2 Likelihood Estimates for Various Values of p

Probability, p	0.48	0.50	0.52	0.54	0.56	0.58	0.60	0.62
Likelihood, L	0.0222	0.0389	0.0581	0.0739	0.0801	0.0738	0.0576	0.0378

If we graph these data across the full range of possible values for p we see the following *likelihood surface*.

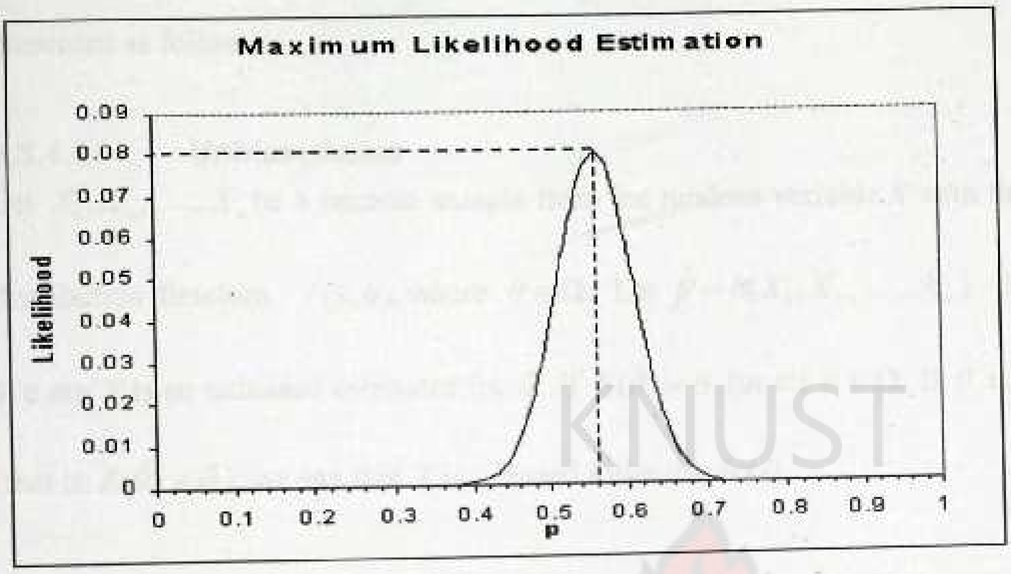

Figure 3.1: Maximum Likelihood Estimation of P

We see that the maximum likelihood estimate for p seems to be around 0.56. In fact, it is exactly 0.56, and it is easy to see why this makes sense in this trivial example. The best estimate for p from any one sample is clearly going to be the proportion of heads observed in that sample. (In a similar way, the best estimate for the population mean will always be the sample mean). So why did we waste our time with the maximum likelihood method? In such a simple case as this, nobody would use maximum likelihood estimation to evaluate p. But not all problems are this simple! As we shall see, the more complex the model and the greater the number of parameters, it often becomes very difficult to make even reasonable guesses at the MLEs. The likelihood framework conceptually takes all of this in its stride, however, and this is what makes it the work-horse of many modern statistical methods.(Craig, et al (2005)).

3.3.4 Properties of Point Estimators

The desirable properties of point estimators: unbiasedness, efficiency, consistency, and sufficiency, which can also be satisfied by a maximum likelihood estimators, are presented as follows:

3.3.4.1 Unbiasedness

Let $X_1, X_2, \ldots, X_n$ be a random sample from the random variable X with the probability distribution function, $f(x,\theta)$, where $\theta \in \Omega$. Let $\hat{\theta} = h(X_1, X_2, \ldots, X_n)$ be a statistic. We say T is an unbiased estimator for θ if $E(\hat{\theta}) = \theta$, for all $\theta \in \Omega$. If $\hat{\theta}$ is not unbiased (that is, $E(\hat{\theta}) \neq \theta$), we say that T is a biased estimator for θ.

3.3.4.2 Efficiency

Selecting among several unbiased estimators is to choose the one with minimum variance, minimum variance unbiased estimator (*MVUE*). The lower bound for the variance of an unbiased estimator can be established if the appropriate derivatives exist and can pass under the integration sign. A remarkable inequality, called the *Rao-Cramer Lower Bound* (*CRLB*), gives a lower bound on the variance of any unbiased estimator. We then show that, under regularity conditions, the variances of the maximum likelihood estimates achieve this lower bound asymptotically.

Assumptions (*Regularity Conditions*):

- The probability density functions ($pdfs$) are distinct.
- The $pdfs$ have common support for all θ.
- The point θ_0 (the true value of θ) is an interior point in Ω, the parameter space.
- ~~The~~ pdf, $f(x,\theta)$ is twice differentiable as a function of θ.

- The integral $\int_{R_x} f(x,\theta)dx$ can be differentiated twice under the integral sign as a function of θ.

Theorem (*Rao-Cramer Lower Bound*):

Let $x_1, x_2, \dots, x_n$ be independent and identically distribution (*iid*) with common *pdf* $f(x,\theta)$, for $\theta \in \Omega$ and T be an unbiased estimator for θ. then under the regularity conditions,

$$Var(T) \geq \left[nI(\theta)\right]^{-1},$$

where $I(\theta) = E\left[\left(\frac{\partial \log f(x,\theta)}{\partial\theta}\right)^2\right] = -E\left[\frac{\partial^2 \log f(x,\theta)}{\partial\theta^2}\right]$ (3.15)

The proof for the continuous *pdf* $f(x,\theta)$ is given as follows:

$1 = \int_{-\infty}^{\infty} f(x;\theta)dx$, since $f(x,\theta)$ is probability density function.

Taking the derivative with respect to θ results in,

$$0 = \int_{-\infty}^{\infty} \frac{\partial f(x;\theta)}{\partial\theta}dx$$

The latter expression can be rewritten as

$$0 = \int_{-\infty}^{\infty} \frac{\partial f(x;\theta)/\partial\theta}{f(x;\theta)} f(x;\theta)dx$$

Or, equivalently,

$$0 = \int_{-\infty}^{\infty} \frac{\partial \log f(x;\theta)}{\partial\theta} f(x;\theta)dx \qquad (3.16)$$

Writing this last equation as expectation, we have established,

$$E\left[\frac{\partial \log f(X;\theta)}{\partial \theta}\right] = 0, \tag{3.17}$$

that is, the mean of the random variable $\frac{\partial \log f(X;\theta)}{\partial \theta}$ is zero(0). If we differentiate equation (3.16) again, it follows that

$$0 = \int_{-\infty}^{\infty} \frac{\partial^2 \log f(x;\theta)}{\partial \theta^2} f(x;\theta)dx + \int_{-\infty}^{\infty} \frac{\partial \log f(x;\theta)}{\partial \theta} . \frac{\partial \log f(x;\theta)}{\partial \theta} f(x;\theta)dx \tag{3.18}$$

The second term of the right hand side of this equation (3.18) can be written as an expectation. This is *Fisher information*, denoted it by

$$I(\theta) = \int_{-\infty}^{\infty} \left(\frac{\partial \log f(x;\theta)}{\partial \theta}\right)^2 f(x;\theta)dx = E\left[\left(\frac{\partial \log f(x;\theta)}{\partial \theta^2}\right)^2\right] \tag{3.19}$$

Clearly, from equations (3.20) and (3.21), we have

$$I(\theta) = -\int \frac{\partial^2 \log f(x;\theta)}{\partial \theta^2} f(x;\theta)dx = Var\left(\frac{\partial \log f(X,\theta)}{\partial \theta}\right), \tag{3.20}$$

which means that the Fisher information is the variance of $\frac{\partial \log f(X;\theta)}{\partial \theta}$.

Note that the important function, $\frac{\partial \log f(x;\theta)}{\partial \theta}$ is called the *score function*. Recall that it determines the estimating equations for the maximum likelihood estimator, that is, the maximum likelihood estimator, $\hat{\theta}$ solves

$$\sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta} = 0 \text{ for } \theta.$$

Now let $T = h(x_1, x_2, \ldots, x_n)$ be the unbiased estimator for θ, $Z = \sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta}$, and define $T = g(T,\theta)$, where and define $E(T) = \int Tg(T,\theta)dT$. Then

$$E(Z) = E\left(\sum_{i=1}^{n} \frac{\partial \log f(x_i;\theta)}{\partial \theta}\right) = 0 \text{ and } Var(Z) = -n\sum_{i=1}^{n} \frac{\partial^2 \log f(x_i,\theta)}{\partial \theta^2}$$

Under the regularity conditions and differentiation under the integral sign:

$$\int_{-\infty}^{\infty} \frac{\partial^2}{\partial \theta^2} \log f(x_i,\theta)dx = \frac{\partial^2}{\partial \theta^2} \int_{-\infty}^{\infty} \log f(x_i,\theta)dx$$

$$Cov(T,Z) = E(TZ) - E(T)E(Z) \tag{3.21}$$

$$\rho = Corr = \frac{Cov(T,Z)}{\sigma_T \sigma_Z} \text{ and } \rho\sigma_T\sigma_Z + E(T)E(Z) = 1$$

$$\rho\sigma_T\sigma_Z = 1, \text{and } \rho^2 \sigma^2{}_T \sigma^2{}_Z \le 1, \text{since } E(Z) = 0 \text{ and } \rho^2 \le 1 \tag{3.22}$$

$$\sigma^2{}_T \ge \frac{1}{\sigma_Z^2} = \frac{1}{nE\left[\left(\frac{\partial \log f(x;\theta)}{\partial \theta}\right)^2\right]} = \frac{1}{-nE\left[\left(\frac{\partial^2 \log f(x;\theta)}{\partial \theta^2}\right)\right]} = \text{CRLB}$$

Definitions: The following are the consequences of the CRLB inequality:

- An unbiased estimator T is said to be efficient if it attains the CRLB.
- For any unbiased estimator T, its efficiency is defined by $eff(T) = \frac{CRLB}{Var(T)}$
- Often $T = t(x_1, x_2, \ldots, x_n)$ is efficient at $n \to \infty$. Specifically, T is said to be asymptotically efficient.

For example, if X has the Bernoulli distribution, $B(1,\theta)$, we obtain $I(\theta)$ as follows:

$$\log f(x;\theta) = x\log\theta + (1-x)\log(1-\theta),$$

$$\frac{\partial \log f(x;\theta)}{\partial \theta} = \frac{x}{\theta} - \frac{1-x}{1-\theta},$$

$$\frac{\partial^2 \log f(x;\theta)}{\partial \theta^2} = -\frac{x}{\theta} - \frac{1-x}{(1-\theta)^2}$$

$$I(\theta)=-E\left[\frac{-X}{\theta^2}-\frac{1-X}{(1-\theta)^2}\right]=\frac{\theta}{\theta^2}+\frac{1-\theta}{(1-\theta)^2}=\frac{1}{\theta}+\frac{1}{(1-\theta)}=\frac{1}{\theta(1-\theta)}$$

3.3.4.3 Consistency

Let $X_1, X_2, \ldots, X_n$ be a random sample from the random variable X with cumulative distribution function, $F(x,\theta)$, $\theta \in \Omega$.. The statistic $\hat{\theta}_n$ is said to be consistent if it converges to θ. In other words, the estimator θ_n is said to be a consistent estimator of θ if, for any positive ε,

$$\lim_{n\to\infty} \mathrm{P}\left(\left|\hat{\theta}_n-\theta\right|\le\varepsilon\right)=1, \text{ or equivalently, } \lim_{n\to\infty}\mathrm{P}\left(\left|\hat{\theta}_n-\theta\right|>\varepsilon\right)=0.$$

This means that an unbiased estimator $\hat{\theta}_n$ for θ is a consistent estimator of θ if $\lim_{x\to\infty} Var(\hat{\theta}_n)\to 0$

3.3.4.4 Sufficiency

The property of sufficiency provides methods for finding statistics that in a sense summarize all the information in a sample about a target parameter. Such statistics are said to be sufficient. Often, sufficient statistics are used to develop estimators that have minimum variance among all unbiased estimators.

- *Definition:*

  The statistic $T(x_1, x_2, \ldots, x_n)$ is said to be sufficient for θ if and only if, for each value of t, the conditional distribution of the random sample $(x_1, x_2, \ldots, x_n)$ given $T=t$ does not depend on θ. $f(x_1,x_2,\ldots,x_n \mid T=t)=\dfrac{f(x_1,x_2,\ldots,x_n,\theta)}{h(t,\theta)}=g(x_1,x_2,\ldots,x_n)$

  The following serves as an illustrative example: Let $x_1, x_2, \ldots, x_n$ be a sequence of independent random sample drawn from the Bernoulli distribution, $B(1,\theta)$, where

$P(x_i = 1) = \theta$ and $f(x_i, \theta) = \theta^{x_i}(1-\theta)^{1-x_i}$, $x_i = 0,1$. To show that $T = \sum x_i$ is sufficient for θ we have the following:

The statistic $T = \sum x_i \sim B(n,\theta) = \binom{n}{t}\theta^t(1-\theta)^{n-t}$

$$f(x_1, x_2, \dots, x_n, \theta) = \theta^{\sum x_i}(1-\theta)^{n-\sum x_i} = \theta^t(1-\theta)^{n-t}$$

$$P(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n | T = t)$$

$$= \frac{f(x_1, x_2, \dots, x_n \mid T = t)}{P(T = t)} = \frac{f(x_1, x_2, \dots, x_n, \theta)}{h(t,\theta)}$$

$$= \frac{\theta^t(1-\theta)^{n-t}}{\binom{n}{t}\theta^t(1-\theta)^{n-t}} = \frac{1}{\binom{n}{t}}, \text{ which is independent of } \theta$$

- *Neyman Factorisation Theorem:*

  This theorem which states as follows may also provide a convenient means of identifying sufficient statistics:

  A necessary and sufficient condition for the statistic $T(x_1, x_2, \dots, x_n)$ to be sufficient for θ is that the joint probability distribution function of the random sample $x_1, x_2, \dots, x_n$ can be factorised into two non-negative functions:

$$f(x_1, x_2, \dots, x_n, \theta) = h(T(x_1, x_2, \dots, x_n), \theta) \cdot g(x_1, x_2, \dots, x_n) \tag{3.25}$$

  where $g(x_1, x_2, \dots, x_n)$ is independent of θ.

  The Proof: ~~Let~~ $X = (X_1, \cancel{X_2, \dots, X_n})$ and $x = (x_1, x_2, \dots, x_n)$

$$P(T=t)=\sum_{T(x)=t} P(X=x)=h(t,\theta)\sum_{T(x)=t} g(x)$$

$$P(X=x \mid T=t)=\frac{P(X=x,T=t)}{P(T=t)}=\frac{h(x)}{\sum_{T(x)=t} h(x)}$$

which does not depend on θ. To show that the conclusion holds in other direction, suppose that the conditional distribution of X given T is independent of θ. Let

$h(t,\theta)=P(T=t,\theta)$ and $g(x)=P(X=x \mid T=t)$

$$P(X=x|\theta)=P(T=t,\theta).P(X=x \mid T=t$$

$$=h(t,\theta).g(x)$$

3.4 The Maximum likelihood Test

Let $X_1, X_2, \ldots, X_n$ be *iid* with *pdf* $f(X,\theta)$ for $\theta \in \Omega$, where θ is a scalar.

Consider the two-sided hypotheses $H_o : \theta = \theta_o$ versus $H_1 : \theta \neq \theta_o$,

where θ_o is a specified value. The maximum likelihood function and its log are given by:

$$L(\theta)=\prod_{i=1}^{n} f(x_i,\theta) \quad \text{and} \quad l(\theta)=\sum_{i=1}^{n} \log f(x_i,\theta) \; .$$

Let $\hat{\theta}$ denote the maximum likelihood estimate of θ. To motivate the test, if θ_o is the true value of θ then, asymptotically, $L(\theta_o)$ is the maximum value of $L(\theta)$. Consider the ratio of two likelihood functions, namely

$$\Lambda = \frac{L(\theta_0)}{L(\hat{\theta})}$$

Note that ≤ 1, but if H_0 is true Λ should be large (close to 1); while H_1 is true, Λ should be smaller. For a specified significance level α, this leads to the intuitive decision rule,

Reject H_0 in favour of H_1 if $\Lambda \leq c$ where c is such that $\alpha = P_{\theta_0}[\Lambda \leq c]$. This test is called the *likelihood ratio test.*

In the multi-parameter case, hypotheses of interest often specify θ to be in a sub-region of the space. Suppose X has a $N(\mu,\sigma^2)$ distribution. The full space is $\Omega = \{(\mu,\sigma^2):\sigma^2 > 0, -\infty < \mu < \infty\}$. This is a two-dimensional space. We may be interested through in testing that $\mu = \mu_0$, where μ_0 is a specified value. Under H_0, the parameter space is the one-dimensional space $\omega = \{(\mu,\sigma^2):\sigma^2 > 0\}$. We say that H_0, is defined in terms of one constraint on the space Ω.

In general, let $x_1,\ldots,x_n$ be *iid* with *pdf* $f(x;\theta)$ for $\theta \in \Omega \subset R^p$. The hypotheses of interest here are, $H_0:\theta \in \omega$ *versus* $H_1:\theta \in \Omega \cap \omega^c$, where $\omega \subset \Omega$ is defined in terms of $q, 0 < q \leq p$, independent constraints of the form, $g_1(\theta) = a_1,\ldots, g_p(\theta) = a_q$. The functions $g_1, g_2, \ldots, g_q$ must be continuously differentiable. This implies ω is a $p-q$ dimensional space. Based on Theorem;

$L(\theta, x) = \prod_{i=1}^{n} f(x_i;\theta), \quad \theta \in \Omega,$, the true parameter maximizes the likelihood function, so an intuitive test statistic is given by the likelihood ratio.

$$\Lambda = \frac{\max_{\theta \in \omega} L(\theta)}{\max_{\theta \in \Omega} L(\theta)}. \tag{3.26}$$

Large values (close to one) of Λ suggests that H_0 is true, while small values indicate H_1 is true. For a specified level α, $0 < \alpha < 1$, this suggests the decision rule reject H_0 in favor of H_1, if $\Lambda \leq C$, where C is such that $\alpha = \max_{\theta \in \omega} P[\Lambda \leq C]$. As in the scale case, this test often has optimal properties; To determine C, we need to determine the distribution of Λ or a function of Λ when H is true.

Let $\hat{\theta}$ denote the maximum likelihood estimator when the parameter space is the full space Ω and let $\hat{\theta}_0$ denote the maximum likelihood estimator when the parameter space is the reduced space ω. For convenience, define $L(\hat{\Omega}) = L(\hat{\theta})$ and $L(\hat{\omega}) = L(\hat{\theta}_0)$.

Then we can write the test statistic as $\Lambda = \dfrac{L(\hat{\omega})}{L(\hat{\Omega})}$. .(Mckean, et al 2005)

CHAPTER FOUR

LOGISTIC REGRESSION MODEL OF MOTOR CLAIMS

4.1 INTRODUCTION.

In this chapter we obtain the logistic regression model for the motor insurance claims of SIC. This will be done by first looking at the distribution of the various risk variables contributing to the claims. We then proceed to perform the odds ratio analysis, predicting claims for the various risk variables and then estimate the parameters in the model using the method of maximum likelihood. This will be done with the help of the R software package (version 2.8.1).

The data for the study were obtained from the Motor and Claims Department of the SIC Insurance Company Limited (SIC). The data included the following relevant information: claim by an insured vehicle, categorized into two - yes or no, and the characteristics of insured vehicle making the claim - sex of driver (policyholder), vehicle make, age and cubic capacity of engine.

The data collected are summarized using the variables Y and X_{ij}. The variable Y defines the claim by an insured vehicle and X_{ij}, defining *ith* level of risk factor j, where $i = 1, 2, \ldots, m_j$, and $j = 1, 2, \ldots, k$. Each risk factor or variable is also categorized into m levels.

A section of the summary of detailed data collected from the Motor and Claims Department of the SIC is displayed ~~in~~ Appendix A.1.

4.2 DESCRIPTIVE ANALYSIS OF CLAIM DATA

The descriptive analysis performed in this section gives the important features of the data obtained. The tables presented below give quick descriptive summary of all the variables under study.

Table 4.1: Distribution of Vehicle Make by Claim

Vehicle Make	No		Yes	
	Number	**Percentage**	**Number**	**Percentage**
OPEL	2915	23.4	38	15.6
AUDI	189	1.5	2	0.8
BMW	625	5	15	6.1
CHEVROLET	80	0.6	1	0.4
DAEWOO	162	1.3	1	0.4
FIAT	166	1.3	15	6.1
FORD	300	2.4	4	1.6
GEO	189	1.5	5	2
HONDA	227	1.8	8	3.3
HYUNDAI	364	2.9	4	1.6
ISUZU	90	0.7	0	0
KIA	308	2.5	4	1.6
LANCIA	59	0.5	1	0.4
LANDROVER	50	0.4	0	0
MAZDA	239	1.9	0	0
MERCEDES	858	6.9	21	8.6
MITSUBISHI	276	2.2	4	1.6
NISSAN	1539	12.3	43	17.6
OTHERS	164	1.3	2	0.8

PEUGEOT	180	1.4	1	0.4
RENAULT	186	1.5	5	2
SEAT	66	0.5	0	0
SUZUKI	76	0.6	2	0.8
TOYOTA	1870	15	49	20.1
VOLKSWAGEN	1195	9.6	19	7.8
VOLVO	100	0.8	0	0
Total	12473	100	244	100

Table 4.1 shows that NISSAN and OPEL cars made 43 and 38 claims respectively which appear to be the highest, accounting for about 33.2 percent of the total claims within the period, 2007 - 2008. It is also observed that ISUZU, LANDROVER, MAZDA, VOLVO and SEAT never made a claim. Generally, it appears that the vehicle make with very high number relatively made higher claims.

Table 4.2: Distribution of Car Age by Claim

Vehicle Age	No		Yes	
	Number	Percentage	Number	Percentage
1-5	3818	30.6	119	48.8
6-10	2427	19.5	40	16.4
11-15	3360	26.9	49	20.1
>15	2868	23	36	14.8
Total	12473	100	244	100

The trend in making claims, by Table 4.2, shows that relatively new vehicles made greatest claim.

3 **Table 4.3: Distribution of Engine Capacity by Claim**

Engine Capacity	No		Yes	
	Number	Percentage	Number	Percentage
1-1.6	5018	40.2	89	36.5
1.7-1.8	2877	23.1	76	31.1
1.9-2.0	2080	16.7	30	12.3
>2.0	2498	20	49	20.1
Total	12473	100	244	100

From Table 4.3 we noticed that 89 (36.5%) of cars that are not more than 1.6 cubic capacity made a claim which appears to be the highest. However, cars that have cubic capacity between 1.8 and 2.0 made the lowest claims due to the number that registered.

4 **Table 4.4: Distribution of Sex by Claim**

Sex of policyholder	No		Yes	
	Number	Percentage	Number	Percentage
Male	9735	78	180	**73.8**
Female	2738	22	64	**26.2**
Total	12473	100	244	**100**

Table 4.4 show that 180 (73.8%) of male policyholders made claims which appear to be higher than that of their female counterparts.

4.3 LOGISTIC REGRESSION MODEL OF MOTOR CLAIMS AT SIC

4.3.1 Estimation of Logistic Regression for Motor Claims

The generalized linear regression model, in this case logistic regression model, is being employed to predict the likelihood of making a motor insurance claim at SIC. The dependent variable (claim) is a binary variable and assumes the values, 0 and 1, therefore the entire claims conform to the Binomial probability distribution. The claim, Y is regressed on the risk factors, X_{ij}, which are some basic characteristics of an insured vehicle. These are sex of the driver or policyholder (X_{i1}), make or type of vehicle (X_{i2}), cubic capacity of vehicle (X_{i3}), and entry age of vehicle (X_{i4}).

The stepwise technique employed produced the first model which involves all the risk factors and the second model which gives the last iterative step. These results are shown in Table 4.5 and Table 4.6. The models obtained are adjusted for the effect of other variables. This is done using the odds ratio analysis.

Table 4.5: First Logistic Regression Model Results

glm(formula = claim ~ SEX + make1 + eng.ccap + car.age, family = binomial)

Predictor variable (X_{ij})	**Coefficients** $(\hat{\beta}_{ij})$	**Std. Error**	$Z-Value$	Pr(>|Z|)
Intercept	-0.87066	0.21219	-18.241	<2e-16 ***
Female (X_{21})	0.13242	0.14962	0.885	0.376157
AUDI (X_{22})	-0.27529	0.73482	-0.375	0.707935
BMW (X_{32})	0.65248	0.31485	2.072	0.038231 *
CHEVROLET (X_{42})	-0.47692	1.02409	-0.466	0.641430
DAEWOO (X_{52})	-0.90296	1.01772	-0.887	0.374949
FIAT (X_{62})	1.80998	0.32244	5.613	1.98e-08***
FORD (X_{72})	-0.05139	0.53356	-0.096	0.923265

GEO (X_{82})	0.46384	0.48606	0.954	0.339944
HONDA (X_{92})	0.76712	0.40067	1.915	0.055544
HYUNDAI ($X_{10\,2}$)	-0.18277	0.52976	-0.345	0.730088
ISUZU ($X_{11\,2}$)	-14.15397	681.65815	-0.021	0.983434
KIA ($X_{12\,2}$)	-0.13826	0.53213	-0.260	0.795000
LANCIA ($X_{13\,2}$)	0.28668	1.02422	0.280	0.779554
LANDROVER ($X_{14\,2}$)	-14.37077	912.32612	-0.016	0.987432
MAZDA ($X_{15\,2}$)	-14.16239	418.30122	-0.034	0.972991
MERCEDES ($X_{16\,2}$)	0.53235	0.28834	1.846	0.064858
MITSUBISHI ($X_{17\,2}$)	0.08311	0.53478	0.155	0.876502
NISSAN ($X_{18\,2}$)	0.66671	0.23035	2.894	0.003799 **
OTHERS ($X_{19\,2}$)	-0.16863	0.73408	-0.230	0.818313
PEUGEOT ($X_{20\,2}$)	-0.83719	1.01821	-0.822	0.410954
RENAULT ($X_{21\,2}$)	0.69050	0.48307	1.429	0.152887
SEAT ($X_{22\,2}$)	-14.29234	796.86662	-0.018	0.985690
SUZUKI ($X_{23\,2}$)	0.63331	0.73693	0.859	0.390127
TOYOTA ($X_{24\,2}$)	0.45263	0.22743	1.990	0.046569 *
VOLKSWAGEN ($X_{25\,2}$)	0.06158	0.28859	0.213	0.831015
VOLVO ($X_{26\,2}$)	-14.28967	647.57923	-0.022	0.982395
eng.ccap1.7-1.8 (X_{23})	0.31996	0.17281	1.851	0.064101
eng.ccap1.9-2.0(X_{33})	-0.25295	0.22570	-1.121	0.262403
eng.ccab>2.0 (X_{43})	-0.12946	0.19991	-0.648	0.517228
car.age6-10(X_{24})	-0.58051	0.18943	-3.064	0.002181 **
car.age11-15(X_{34})	-0.70092	0.18049	-3.883	0.000103***
car.age>15(X_{44})	-0.82087	0.20152	-4.073	4.63e-05***

AIC= 2367.2, Significance levels '***'= 0.001, '**'= 0.01, '*'= 0.05, '.'= 0.1

Table 4.6: Second Logistic Regression Model Results

glm(formula = claim ~ SEX +car. make + eng.ccap + car.age, family = binomial)

Predictor variable (X_{ij})	**Coefficients** $(\hat{\beta}_{ij})$	**Std. Error**	$Z-Value$	$\Pr(>\lvert Z\rvert)$
Intercept	-3.83518	0.20792	-18.446	< 2e-16 ***
AUDI (X_{22})	-0.27918	0.73468	-0.380	0.70395
BMW (X_{32})	0.64814	0.31478	2.059	0.03949 *
CHEVROLET (X_{42})	-0.46810	1.02401	-0.457	0.64758
DAEWOO (X_{52})	-0.89788	1.01770	-0.882	0.37763
FIAT (X_{62})	1.82408	0.32182	5.668	1.44e-08 ***
FORD (X_{72})	-0.04748	0.53349	-0.089	0.92909
GEO (X_{82})	0.47595	0.48588	0.980	0.32730
HONDA (X_{92})	0.77985	0.40035	1.948	0.05143
HYUNDAI $(X_{10\,2})$	-0.17853	0.52970	-0.337	0.73609
ISUZU $(X_{11\,2})$	-14.15633	681.81758	-0.021	0.98344
KIA $(X_{12\,2})$	-0.12442	0.53186	-0.234	0.81504
LANCIA $(X_{13\,2})$	0.28165	1.02411	0.275	0.78330
LANDROVER $(X_{14\,2})$	-14.37698	912.47690	-0.016	0.98743
MAZDA $(X_{15\,2})$	-14.16280	418.34567	-0.034	0.97299
MERCEDES $(X_{16\,2})$	0.52740	0.28823	1.830	0.06728
MITSUBISHI $(X_{17\,2})$	0.08395	0.53476	0.157	0.87525
NISSAN $(X_{18\,2})$	0.67094	023027	2.914	0.00357 **
OTHERS $(X_{19\,2})$	-0.17053	0.73408	-0.232	0.81630
PEUGEOT $(X_{20\,2})$	-0.83886	1.01822	-0.824	0.41002
RENAULT $(X_{21\,2})$	0.69906	0.48293	1.448	0.14775
SEAT $(X_{22\,2})$	-14.28235	797.11979	-0.018	0.98570
SUZUKI $(X_{23\,2})$	0.64590	0.73677	0.877	0.38066
TOYOTA $(X_{24\,2})$	0.45667	0.22731	2.009	0.04453 *
VOLKSWAGEN $(X_{25\,2})$	0.06272	0.28850	0.217	0.82789

VOLVO ($X_{36\,2}$)	-14.29467	647.57423	-0.022	0.98239
eng.ccap1.7-1.8 (X_{23})	0.32088	0.17276	1.857	0.06326
eng.ccap1.9-2.0(X_{33})	-0.25544	0.22562	-1.132	0.25758
eng.ccab>2.0 (X_{43})	-0.13854	0.19957	-0.694	0.48755
car.age6-10(X_{24})	-0.58381	0.18937	-3.083	0.00205 **
car.age11-15(X_{34})	-0.70687	0.18030	-3.921	8.83e-05 ***
car.age>15(X_{44})	-0.82087	0.20152	-4.073	4.63e-05 ***

AIC= 2366.0,. Significance levels '***'= 0.001, '**'= 0.01, '*'= 0.05, '.'= 0.1

We notice from the first model (in Table 4.5) that the inclusion of gender (female) and cubic capacity of engine did not produce any significant influence on the claim. The other predictor variables were significant, the highly being (at 0.001 significance level) the vehicle type: NISSAN and age of vehicle: 10-15 and more than 15 years. The model produced an AIC value of 2367.2.

After obtaining Model 1 (Table 4.5), we apply the stepwise regression technique where the iteration of the estimation process and other inferential analyses resume. The results obtained are similar to that of the first model but with a slight improved AIC value of 2366.0. These are shown in Table 4.6 (Model 2). The final model established is given by:

$$Y = \beta_0 + \lambda VehicleType + \gamma CarAge$$

Thus

$$Y_1 = -3.83518 + 0.64814BMW - 0.58381CarAge_{(6-10)}$$
$$Y_2 = -3.83518 + 0.64814BMW - 0.70687CarAge_{(11-15)}$$
$$Y_3 = -3.83518 + 0.64814BMW - 0.82087CarAge_{(>15)}$$
$$Y_4 = -3.83518 + 1.82408FIAT - 0.58381CarAge_{(6-10)}$$
$$Y_5 = -3.83518 + 1.82408FIAT - 0.70687CarAge_{(11-15)}$$
$$Y_6 = -3.83518 + 1.82408FIAT - 0.82087CarAge_{(>15)}$$
$$Y_7 = -3.83518 + 0.67094NISSAN - 0.58381CarAge_{(6-10)}$$
$$Y_8 = -3.83518 + 0.67094NISSAN - 0.70687CarAge_{(11-15)}$$
$$Y_9 = -3.83518 + 0.67094NISSAN - 0.82087CarAge_{(>15)}$$
$$Y_{10} = -3.83518 + 0.45667TOYOTA - 0.58381CarAge_{(6-10)}$$
$$Y_{11} = -3.83518 + 0.45667TOYOTA - 0.70687CarAge_{(11-15)}$$
$$Y_{12} = -3.83518 + 0.45667TOYOTA - 0.82087CarAge_{(>15)}$$

Where Y=response and β_0=constant parameter

4.3.2 Odds Ratios Analysis of Risk Factors

The computation of the crude and adjusted odds ratios for the risk factors, X_{ij}, is given by the estimate, $\exp(\hat{\beta}_{ij})$. The crude odds ratio of a risk factor determines the influence it has on the claim outcome while the adjusted odds ratio accounts for the inclusive effect of the other risk factors. The Wald's and log likelihood ratio tests are also performed to ascertain the significant effect of the risk factors. A probability value of less than or equal to 0.05 was considered to be statistically significant. Hence, the inclusion of that risk factor is important in determining the claim's outcome, $Y = 0$ or 1.

Table 4.7: Odds Ratio Analysis Predicting Claim for Sex Levels

Sex	Model 1			Model 2		
	Crude OR (95% CI)	Adjusted	P-value	Crude OR (95% CI)	Adjusted	P-value
Female	1.26 (0.95,1.69)	1.14 (0.85,1.53)	0.376	NIL	NIL	NIL
Ref. Male	P(LR-test) =0.381			NIL		

Table 4.7 shows an odds ratio of 1.26 and 95% interval estimate of (0.95, 1.69), indicating that the females are 1.26, as likely to make a claim as to their male counterparts. Finally, after odds ratio adjustment has been effected, the odds ratio slightly reduces to 1.14 with a 95% interval estimate of (0.85, 1.53), giving a similar statistically significance result. The confidence interval includes 1, which means that the sex predictor variable is not significant, so it is excluded in Model 2.

Table 4.8: Odds Ratio Analysis Predicting Claim for Vehicle Make Levels

Type of Vehicle	Model 1			Model 2		
	Crude OR (95% CI)	**Adjusted**	**P-Value**	**Crude OR (95% CI)**	**Adjusted**	**P-Value**
AUDI	0.81 (0.19, 3.39)	0.71 (0.18,3.21)	0.708	0.81 (0.19,3.39)	0.76 (0.18,3.19)	0.704
BMW	1.84 (1.01,3.37)	1.92 (1.04,3.56)	0.038	1.84 (1.01,3.37)	1.91 (1.03,3.54)	0.039
CHEVROLET	0.96 (0.13,7.07)	0.62 (0.08,4.62)	0.641	0.96 (0.13,7.07)	0.63 (0.08,4.66)	0.648
DAEWOO	0.47 (0.06,3.47)	0.41 (0.06,2.98)	0.375	0.47	0.41	0.378

				(0.06,3.47)	(0.06,2.99)	
FIAT	6.93 (3.74,12.86)	6.11 (3.25,11.5)	< 0.001	6.93 (3.74,12.86)	6.2 (3.3,11.64)	< 0.001
FORD	1.02 (0.36,2.89)	0.95 (0.33,2.7)	0.923	1.02 (0.36,2.89)	0.95 (0.34,2.71)	0.929
GEO	2.03 (0.79,5.22)	1.59 (0.61,4.12)	0.34	2.03 (0.79,5.22)	1.61 (0.62,4.17)	0.327
HONDA	2.7 (1.25,5.86)	2.15 (0.98,4.72)	0.056	2.7 (1.25,5.86)	2.18 (1,4.78)	0.051
HYUNDAI	0.84 (0.3,2.38)	0.83 (0.29,2.35)	0.73	0.84 (0.3,2.38)	0.84 (0.3,2.36)	0.736
ISUZU	0 (0,Inf)	0 (0,Inf)	0.983	0 (0,Inf)	0 (0,Inf)	0.983
KIA	1 (0.35,2.81)	0.88 (0.31,2.5)	0.815	1 (0.35,2.81)	0.88 (0.31,2.5)	0.815
LANDROVER	0 (0,Inf)	0 (0,Inf)	0.987	0 (0, Inf)	0 (0, Inf)	0.987
MAZDA	0 (0,Inf)	0 (0,Inf)	0.973	0 (0, Inf)	0 (0, Inf)	0.973
MERCEDES	1.88 (1.1,3.22)	1.7 (0.97,3)	0.065	1.88 (1.1, 3.22)	1.69 (0.96,2.98)	0.067
MITSUBISHI	1.11 (0.39,3.14)	1.09 (0.38,3.1)	0.877	1.11 (0.39,3.14)	1.09 (0.38,3.1)	0.875
NISSAN	2.14 (1.38,3.33)	1.95 (1.24,3.06)	0.004	2.14 (1.38,3.33)	1.96 (1.25,3.07)	0.004
OTHERS	0.94 (0.22,3.91)	0.84 (0.2,3.56)	0.818	0.94 (0.22,3.91)	0.84 (0.2,3.55)	0.816
PEUGEOT	0.43 (0.06,3.12)	0.43 (0.06,3.19)	0.411	0.43 (0.06,3.12)	0.43 (0.06,3.18)	0.41
RENAULT	2.06 (0.8,5.3)	1.99 (0.77,5.14)	0.153	2.06 (0.8,5.3)	2.01 (0.78,5.18)	0.148
SEAT	0 (0,Inf)	0 (0,Inf)	0.986	0 (0, Inf)	0 (0, Inf)	0.986
SUZUKI	2.02 (0.48,8.52)	1.88 (0.44,7.99)	0.39	2.02 (0.48,8.52)	1.91 (0.45,8.08)	0.381

TOYOTA	2.01 (1.31,3.08)	1.57 (1.01,2.46)	0.047	2.01 (1.31,3.08)	1.58 (1.01,2.46)	0.045
VOLKSWAGEN	1.22 (0.7,2.12)	1.06 (0.6,1.87)	0.831	1.22 (0.7,2.12)	1.06 (0.6,1.87)	0.828
VOLVO	0 (0,Inf)	0 (0,Inf)	0.982	0 (0, Inf)	0 (0, Inf)	0.982
Ref: OPEL	P(LR-test =0.001			P(LR-test) =0.001		

The odds ratios in Table 4.8, for BMW, FIAT, NISSAN, and TOYOTA clearly show that they are statistically significant for both models. The odds ratios recorded for Model 2 are respectively 1.91, 6.2, 1.96 and 1.58 as likely to make claim as OPEL. Their confidence intervals do not include 1 and so they are statistically significant.

Table 4.9: Odds Ratio Analysis Predicting Claim for Engine Capacity Levels

Engine Capacity	Model 1			Model 2		
	Crude OR (95% CI)	Adjusted	P-value	Crude OR (95% CI)	Adjusted	P-value
1.7-1.8	1.49 (1.09,2.03)	1.38 (0.98,1.93)	0.064	1.49 (1.09,2.03)	1.38 (0.98,1.93)	0.063
1.9-2.0	0.81 (0.54,1.23)	0.78 (0.5,1.21)	0.262	0.81 (0.54,1.23)	0.77 (0.5,1.21)	0.258
>2.0	1.11 (0.78,1.57)	0.88 (0.59,1.3)	0.517	1.11 (0.78,1.57)	0.87 (0.59,1.29)	0.488
Ref: > 1-1.6	P(LR-test) =0.026			P(LR-test) =0.024		

Table 4.9 shows that after adjusting for the effect of other levels of engine capacity, none was found to be significant at ~~significance~~ level of 0.05. The odds ratios together with their ~~95%~~ confidence intervals for engine capacity level 1.6-1.8, 1.8-2.0, and > 2.0 are

1.38 (0.98, 1.93), 0.77 (0.50, 1.21), and 0.87 (0.59, 1.29) respectively. The log likelihood test rather gave a significant result at significance level of 0.05.

Table 4.10: Odds Ratio Analysis predicting claim level for Car Age

CAR AGE	Model 1			Model 2		
	Crude OR (95% CI)	**Adjusted**	**P-value**	**Crude OR (95% CI)**	**Adjusted**	**P-value**
6-10	0.53 (0.37,0.76)	0.56 (0.39,0.81)	0.002	0.53 (0.37,0.76)	0.56 (0.38,0.81)	0.002
11-15	0.47 (0.33,0.65)	0.5 (0.35,0.71)	< 0.001	0.47 (0.33,0.65)	0.49 (0.35,0.7)	< 0.001
>15	0.4 (0.28,0.59)	0.45 (0.3,0.66)	< 0.001	0.4 (0.28,0.59)	0.44 (0.3,0.65)	< 0.001
Ref: 1-5	P(LR-test) =0.001			P(LR-test) =0.001		

In Table 4.10, the adjustment for all the levels of the vehicle age entry produces results which indicate they are all less likely to make a claim as compared to cars that are not more than five years. The odds ratios of increasing age vehicle are 0.56, 0.49 and 0.44 respectively with confidence intervals which do not include 1, showing a highly statistically significant inclusion of this risk factor.

CHAPTER FIVE

CONCLUSION AND RECOMMENDATIONS

5.1 INTRODUCTION

The thesis seeks to develop a logistic regression model with the ultimate goal of determining the main risk factors associated with the motor insurance claims at SIC Insurance Company Limited (SIC). This is tried to be achieved by reviewing pertinent materials for thorough knowledge and understanding of motor insurance industry and how it operates, especially in Ghana. The logistic regression modeling which was employed in the data analysis has its pertinent concepts well-discussed.

The Motor and Claims Department of the SIC Insurance Company Limited (SIC) provided the data for the study which were solely on private insured vehicles. The data included claim or otherwise by an insured vehicle, and its characteristics: sex of the driver (policyholder), make of vehicle, age and cubic capacity of engine. The claim is denoted by the response variable, $Y = 0, \text{or} 1$, the characteristics, called predictor variables or risk factors, and categorized levels, by X_{ij}, defining ith level of risk factor j. The R software package was used to perform the required logistic analysis of the data.

5.2 SUMMARY OF FINDINGS

The descriptive analysis showed 26.2% of female went for claims, compared to their male counterparts of more than 70% in the same period of study. Vehicles that have small cubic capacity relatively make ~~the larger~~ share of the claims (36.5%.). It is interesting to see the trend as the cubic capacity becomes great, the claim rate goes down. The frequency of claims made by OPEL (15.6%), NISSAN (17.6%), and TOYOTA (20.1%) are the relatively the high claims. However, cars like ISUZU, LANDROVER, MAZDA, SEAT and VOLVO never made claims within the registered year. The rest ranges from 0.4% to 8.6%. It was also noticed that 48% of cars not more than 5 years made claims. The claims made for cars aged 5-15 averaged 18.25% whereas claims by cars aged more than 15 years is 14.8%. It is also interesting to find that as vehicle gets old, the claim becomes less likely.

The predictive model established for the SIC motor claims was found to be attributed to make of vehicle and its entry age and were all statistically significant at most 0.05 level of significance. The prominent among vehicle make are FIAT, NISSAN, TOYOTA and BMW whose entry age is least 6 years. The sex of policyholder and engine capacity of vehicle did not seem to have significant impact on the claims. These results are as summarized in Table 5.1, shown below:

Table 5.1: Established Model Results

Risk Factor	*Coefficients*	*Std. Error*	*Z- value*	*Pr(>|Z|)*	*Odds Ratio*
Intercept	-3.83518	0.20792	-18.446	< 2e-16	0.02160
BMW	0.64814	0.31478	2.059	0.03949	1.91198
FIAT	1.82408	0.32182	5.668	1.44e-08	6.19709
NISSAN	0.67094	023027	2.914	0.00357	1.95608
TOYOTA	0.45667	0.22731	2.009	0.04453	1.57881
car.age6-10	-0.58381	0.18937	-3.083	0.00205	0.55777
car.age11-15	-0.70687	0.18030	-3.921	8.83e-05	0.49319
car.age>15	-0.82087	0.20152	-4.073	4.63e-05	0.44005

The motor insurance model for SIC (from Table 5.1) is given by:

$Y = \beta_0 + \lambda VehicleType + \gamma CarAge$

Thus

$$
\begin{aligned}
Y_1 &= -3.83518 + 0.64814BMW - 0.58381CarAge_{(6-10)} \\
Y_2 &= -3.83518 + 0.64814BMW - 0.70687CarAge_{(11-15)} \\
Y_3 &= -3.83518 + 0.64814BMW - 0.82087CarAge_{(>15)} \\
Y_4 &= -3.83518 + 1.82408FIAT - 0.58381CarAge_{(6-10)} \\
Y_5 &= -3.83518 + 1.82408FIAT - 0.70687CarAge_{(11-15)} \\
Y_6 &= -3.83518 + 1.82408FIAT - 0.82087CarAge_{(>15)} \\
Y_7 &= -3.83518 + 0.67094NISSAN - 0.58381CarAge_{(6-10)} \\
Y_8 &= -3.83518 + 0.67094NISSAN - 0.70687CarAge_{(11-15)} \\
Y_9 &= -3.83518 + 0.67094NISSAN - 0.82087CarAge_{(>15)} \\
Y_{10} &= -3.83518 + 0.45667TOYOTA - 0.58381CarAge_{(6-10)} \\
Y_{11} &= -3.83518 + 0.45667TOYOTA - 0.70687CarAge_{(11-15)} \\
Y_{12} &= -3.83518 + 0.45667TOYOTA - 0.82087CarAge_{(>15)}
\end{aligned}
$$

Where Y=response and β_0=constant parameter

5.3 CONCLUSION

The study has revealed that there is difference in the contribution level to making claims by the various categories of the risk variables or factors considered. This allowed for an in-depth analysis of claim variables which required statistical modeling, called logistic regression analysis to assist in understanding the motor insurance claims in Ghana.

The established model for the motor (given in sections 4.3.1 and 5.2) is mainly attributed to the relatively old vehicles of the following makes: FIAT, NISSAN, TOYOTA, and

BMW. It is our belief, the model if implemented, would go a long way of improving efficiency at the Motor Department of SIC.

5.4 RECOMMENDATIONS AND FURTHER WORK

In line with the detailed analysis of the data obtained from SIC and also other observations made in the course of this study, we make following recommendations:

- The SIC Insurance Company Limited should give relatively high premium to new cars as compared to older ones due to their higher demand of claims, thus helping SIC to increase their financial gains.

- The risk factors could be structured to take into account explicitly the underlying assumptions involved in claim. This could help actuaries concerned with motor claims to appreciate the problems and advise the underwriters accordingly.

- The users of OPEL, FIAT, NISSAN, TOYOTA, and BMW must be educated to be more careful since they consists majority (about 66%) of the claims made.

- The Information Bill which has been with our law-makers in Parliament (the legislature) for long be passed by the law makers with the view of removing the bottlenecks associated with data acquisition which was a a serious challenge of this study.

- It is also recommended that further study be done in other insurance companies with the view of appreciating the discrepancies that exist in valuing risk, thus help National Insurance Commission (NIC) to fairly price motor premiums.

REFERENCES

Actuarial & Management Consultants(Gh) Ltd. : Report on Creating an Industry wide market risk profile database(Vooacott T.H., Introduction to statistics, pp 445-450)

Aeron-Thomas (1997), Zambia Road Safety Final Report and Action Plan, Ministry of Transport,Lusaka

Aeron-Thomas et al. (2001), Review of Road Safety Management, TRL, Crowthorne.

Allen T.C., Joseph W.M., & Robert V. H., (2005) Introduction to Mathematical Statistics,6th Edition, Upper Saddle River, New Jersey

Amoo K. G. (2002). The challenge of providing insurance cover for loss of use of motor vehicle in a developing economy. Chartered Insurance Institute 4

Andersen, E.B. (1970). Asymptotic properties of conditional maximum likelihood estimators Journal of the Royal Statistical Society B 32, 283-301

Asyali, M.H., Kelton, D., Shoukri, M.M., & VanDorp, R., (2004). The Poisson Inverse Gaussian regression model in the analysis of clustered counts data. Journal of Data Science 2, 17-32.

Association of British Insurers (2001), Code of Best Practice on Rehabilitation, Early Intervention and Medical Treatment in Personal Injury Claims, London

Association for Chief Police Officers (2000), Effective Road Policing Volume 11, London

Badinter Loi Law. Extracts from article on http://www.spps.net/indexx2.2.htm

Botswana Motor Accident Fund (2001), http://www.raf-mu.com

Cameron, A. C. & Trivedi, P. K. (1998). Regression Analysis of Count Data.Cambridge University Press.

Cho, D., (1989), Some evidence of scale economies in workers' compensation insurance, *The Journal of Risk and Insurance* 55

Cummins, J.D. & M.A. Weiss, (2000), Analyzing firm performance in the insurance industry using frontier efficiency and productivity methods, in G. Dionne (ed), *Handbook of Insurance*, Kluwer, Boston.

Dean, C., Lawless, J.F., & Willmot, G.E. (1989). A mixed Poisson-Inverse Gaussian regression model. Canadian Journal of Statistics 17, 171-182.

Englund A and Pettersen HE (1997), The effect of traffic and vehicle insurance on traffic safety, VTI Report 415A,

Fecher, F., D. Kessler, S. Perelman, P. & Pestieau, (1993), Productive performance of the French insurance industry, *Journal of Productivity Analysis* 4, pp. 73–93.

Griliches, Z. Hall, B.H., & Hausman, J.A., (1984). Econometric models for count data with application to the Patents-R&D Relationship. Econometrica 52, 909-938.

Insurance Institute for Highway Safety (2001). Status Report. *New head restraint rule would prevent many whiplash injuries, but proposed dynamic tests could compromise safety.* Vol.36,No.4 April 28, 2001.

Joffe J (1997), Second Draft White Paper on the Road Accident Fund, Pretoria.

Louvain-la-Neuve (2005).Fixed versus random effects in poisson Regression models for claim counts: a Case study with motor insurance Jean-philippe boucher & michel denuit Institut des Sciences Actuarielles Universit´e Catholique de Louvain 6 rue des Wallons B-1348 Belgium.

Manning W.G., & J. Mullahy, (2001), Estimating log models: to transform or not to transform, *Journal of Health Economics* 20, pp. 461-494

Neter, K.N.W. (1996) Applied Linear Statistical Models,4th Edition, Tom Casson, New York.

Palmgren, J. (1981). The Fisher information matrix for log-linear models arguing conditionnally in observed explanatory variables. Biometrika 68, 563-566.

PROBLEM OF MOTOR INSURANCE-CLAIM RESERVES, NORTON E. MASTERSON, Stevens Point, Wisconsin,(U.S.A)

Swank, J., (1996), How stable is the multiproduct translog cost function? Evidence from the Dutch bnking industry, *Kredit und Kapital* 29, pp. 153-172.

Tremblay, L. (1992). Using the Poisson Inverse Gaussian in bonus-malus systems. ASTIN Bulletin 22, 97-106.

Tri-Star (2001), Report on Creating an Industry wide Market Risk Profile Database,Management Consult,Accra

Willmot, G. (1987). The Poisson-Inverse Gaussian distribution as an alternative to the Negative

Binomial. Scandinavian Actuarial Journal, 113-127.

Yuengert, A.W., (1993), The measurement of efficiency in life insurance: estimates of a mixed normal-gamma error model, *Journal of Banking and Finance* 17, pp. 483-496

APPENDIX

A.1 Section of Data Collected

id	Make	make1	Model	Usage	Claim	no claims	SEX	engine cap	Car Entry age	Carage at exit
2017	MERCEDES	47	SALOON	Private	0	0	1	6.0	16	17
4628	DODGE	20	MINI BUS	Private	0	0	1	5.9	11	12
7635	DODGE	20	PICKUP	Private	0	0	1	5.9	11	12
3066	OPEL	50	SALOON	Private	0	0	1	5.8	11	12
1758	CHEVROLET	11	SALOON	Private	0	0	1	5.7	11	12
7774	CHEVROLET	11	JEEP	Private	0	0	1	5.7	12	13
9942	CHEVROLET	11	PICKUP	Private	0	0	1	5.7	16	17
12976	CHEVROLET	11	SALOON	Private	0	0	1	5.7	8	9
5977	DODGE	20	OTHERS	Private	0	0	1	5.7	1	2
314	GMC	26	PICKUP	Private	0	0	2	5.7	12	13
15796	GMC	26	SALOON	Private	0	0	1	5.7	4	5
8390	INFINITI	32	SALOON	Private	0	0	1	5.6	1	2
8694	INFINITI	32	SALOON	Private	0	0	2	5.6	3	4
7810	NISSAN	49	SALOON	Private	0	0	1	5.6	1	2
11232	NISSAN	49	SALOON	Private	0	0	1	5.6	3	4
10208	MERCEDES	47	SALOON	Private	0	0	1	5.5	1	2
4001	FORD	24	SALOON	Private	0	0	1	5.4	7	8
15322	LINCOLN	44	SALOON	Private	0	0	1	5.4	9	10
8473	CHEVROLET	11	SALOON	Private	0	0	1	5.3	6	7
9302	CHEVROLET	11	PICKUP	Private	0	0	1	5.3	4	5
7041	BMW	7	SALOON	Private	0	0	1	5.3	14	15
11174	BMW	7	SALOON	Private	0	0	1	5.3	17	18
724	CHEROKEE	10	SALOON	Private	0	0	1	5.2	11	12
5869	CHEROKEE	10	JEEP	Private	0	0	1	5.2	13	14

10541	CHEROKEE	10 SALOON	Private	0	0	1	5.2	11	12
15273	CHEROKEE	10 SALOON	Private	0	0	2	5.2	14	15
8439	DODGE	20 PICKUP	Private	0	0	2	5.2	8	9
11755	GRAND	27 SALOON	Private	0	0	1	5.2	10	11
10715	JEEP	35 JEEP	Private	0	0	2	5.2	13	14
15924	OTHERS	51 SALOON	Private	0	0	2	5.2	11	12
15559	BMW	7 SALOON	Private	0	0	1	5.0	17	18
2771	FORD	24 PICKUP	Private	0	0	1	5.0	18	19
4353	FORD	24 PICKUP	Private	0	0	1	5.0	14	15
5899	FORD	24 PICKUP	Private	0	0	2	5.0	15	16
6312	FORD	24 VAN	Private	0	0	1	5.0	13	14
9983	FORD	24 BUS	Private	0	0	1	5.0	11	11
10986	FORD	24 SALOON	Private	0	0	1	5.0	10	11
11811	FORD	24 PICKUP	Private	0	0	1	5.0	13	14
15685	FORD	24 JEEP	Private	0	0	1	5.0	11	12
15863	FORD	24 PICKUP	Private	0	0	1	5.0	10	11
8856	MERCEDES	47 SALOON	Private	0	0	1	5.0	5	6
16063	MERCEDES	47 SALOON	Private	0	0	1	5.0	20	21
5349	NISSAN	49 SALOON	Private	0	0	3	5.0	3	4
10059	VOLKSWAGEN	74 SALOON	Private	0	0	1	5.0	3	4
12761	DAEWOO	15 SALOON	Private	0	0	1	5.0	8	9
13980	OTHERS	51 SALOON	Private	0	0	2	5.0	12	13
12689	BMW	7 SALOON	Private	0	0	2	4.8	2	3
2671	MERCEDES	47 SALOON	Private	0	0	1	4.8	6	7
9993	CHEROKEE	10 JEEP	Private	0	0	2	4.7	8	9
9715	DODGE	20 PICKUP	Private	0	0	2	4.7	5	6
9759	LEXUS	42 SALOON	Private	0	0	1	4.7	5	6
12776	LEXUS	42 SALOON	Private	0	0	2	4.7	9	10
666	TOYOTA	71 SALOON	Private	1	1	3	4.7	7	8

1330	TOYOTA	71 SALOON	Private	0	0	1	4.7	4	5
3069	TOYOTA	71	Private	0	0	1	4.7	5	6
4087	TOYOTA	71 4X4WD	Private	0	0	1	4.7	3	4
5851	TOYOTA	71 SALOON	Private	0	0	1	4.7	9	10
6143	TOYOTA	71 SALOON	Private	0	0	1	4.7	4	5
6625	TOYOTA	71 SALOON	Private	0	0	1	4.7	8	9
6824	TOYOTA	71 PICKUP	Private	0	0	1	4.7	4	5
8524	TOYOTA	71 SALOON	Private	0	0	1	4.7	8	9
9250	TOYOTA	71 SALOON	Private	0	0	1	4.7	5	6
9906	TOYOTA	71 PICKUP	Private	0	0	1	4.7	2	3
10699	TOYOTA	71 STATION	Private	0	0	1	4.7	3	4
10831	TOYOTA	71 SALOON	Private	0	0	1	4.7	5	6
11266	TOYOTA	71 SALOON	Private	0	0	1	4.7	2	3
13975	TOYOTA	71 SALOON	Private	0	0	1	4.7	7	7
14020	TOYOTA	71 PICKUP	Private	0	0	1	4.7	7	8
15357	TOYOTA	71 PICKUP	Private	0	0	1	4.7	2	3
15598	TOYOTA	71 SALOON	Private	0	0	1	4.7	8	9
15834	TOYOTA	71 SALOON	Private	0	0	1	4.7	2	2
12773	TOYOTA	71 SALOON	Private	0	0	3	4.7	17	18
5662	OTHERS	51 OTHERS	Private	0	0	1	4.7	12	13
6855	LEXUS	42 SALOON	Private	0	0	1	4.7	10	11
7224	NISSAN	49 SALOON	Private	0	0	1	4.6	18	19
915	FORD	24 SALOON	Private	0	0	1	4.6	1	2
5970	FORD	24 PICKUP	Private	0	0	1	4.6	9	10
6192	FORD	24 PICKUP	Private	0	0	1	4.6	9	10
9303	FORD	24 SALOON	Private	0	0	1	4.6	4	5
10802	FORD	24 PICKUP	Private	0	0	1	4.6	10	11
11316	FORD	24 SALOON	Private	0	0	1	4.6	8	9
9492	LANDROVER	40 SALOON	Private	0	0	2	4.6	7	8

1177	LINCOLN	44 SALOON	Private	0	0	1	4.6	12	13
9588	LINCOLN	44 VAN	Private	0	0	1	4.6	3	4
1412	RANGE ROVER	57 SALOON	Private	0	0	1	4.6	8	9
13616	CHEROKEE	10 JEEP	Private	0	0	1	4.5	10	11
6428	INFINITI	32 SALOON	Private	0	0	1	4.5	3	4
8790	LINCOLN	44 SALOON	Private	0	0	1	4.5	4	5
9454	NISSAN	49 SPORTS	Private	0	0	1	4.5	2	3
5676	OTHERS	51 SALOON	Private	0	0	1	4.5	11	12
743	TOYOTA	71 SALOON	Private	0	0	1	4.5	11	13
1232	TOYOTA	71 SALOON	Private	0	0	1	4.5	11	12
1296	TOYOTA	71 SALOON	Private	0	0	1	4.5	8	9
2286	TOYOTA	71 SALOON	Private	0	0	1	4.5	13	14
7421	TOYOTA	71 SALOON	Private	0	0	1	4.5	16	17
8464	TOYOTA	71 SALOON	Private	0	0	1	4.5	6	7
9849	TOYOTA	71 SALOON	Private	0	0	1	4.5	12	13
6959	FORD	24 PICKUP	Private	0	0	1	4.5	10	11
949	LANDROVER	40 SALOON	Private	0	0	1	4.4	2	3
5962	LANDROVER	40 SALOON	Private	0	0	1	4.4	1	2
10048	RANGE ROVER	57 SALOON	Private	0	0	2	4.4	3	4
12699	NISSAN	49 SALOON	Private	0	0	1	4.3	9	10
2482	CHEVROLET	11 SALOON	Private	0	0	1	4.3	10	11
4331	CHEVROLET	11 SALOON	Private	0	0	1	4.3	9	10
6903	CHEVROLET	11 BUS	Private	0	0	1	4.3	14	15
6965	CHEVROLET	11 SALOON	Private	0	0	1	4.3	15	16
8449	CHEVROLET	11 SALOON	Private	0	0	2	4.3	12	13
10094	CHEVROLET	11 SALOON	Private	0	0	1	4.3	9	10
10287	CHEVROLET	11 SALOON	Private	0	0	1	4.3	7	8
9266	FORD	24 SALOON	Private	0	0	1	4.3	12	13
5010	MERCEDES	47 SALOON	Private	0	0	1	4.3	8	8