EXTREME VALUE THEORY

AND

ITS APPLICATION TO MOTOR INSURANCE

Fayek Hanna Tawdros

A Thesis Submitted for the Degree of
Doctor of Philosophy

The City University
Mathematics Department
(Actuarial Science Unit)
London

July 1983

TABLE OF CONTENTS

Page No.

Page No.

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor, Professor Bernard Benjamin, for his constant encouragement and guidance throughout all phases of the study.

I thank also Mrs Ann Sutcliff, the Librarian of the Institute of Actuaries' Library, for the library facilities afforded me.

My thanks go also to my colleagues in the Motor Department of the Egyptian Reinsurance Company, and to Mr Hammam Badr, the Company's representative in London, for the data supplied to me and all the information I required.

I would also like to thank Mr Ali Naweto, the Chairman of the National Insurance Company of Egypt, and the Board of Directors for their understanding of the importance of such a study and for their consent given for me to complete it.

My appreciation is also due to Mrs Maureen Brierley for the typing of the final manuscript.

Finally, I would like to thank my wife and my son for their understanding and their encouragement.

ABSTRACT

This research is concerned with the application of extreme value theory to motor insurance generally, and, in particular, to motor insurance in the Egyptian market. Since this theory was introduced there have been many attempts to use it in some insurance fields, but very few attempts have been made to apply it to motor insurance and this attempt is the first to be made in the Egyptian market.

In this particular attempt the aim is to study the likely benefit of the application of the theory within Egyptian experience; to attract the attention of underwriters to the shape of the claim amount distribution, especially at its extreme end so that they may be in a better position to protect their companies against unexpectedly severe claims.

The research includes formulae representing the 10 largest claims and an estimation of the 10 largest claims which are expected to occur in the Egyptian market during the period 1977 to 1981 inclusive. The detailed and complete calculations are presented.

CHAPTER ONE

1.1 Introduction

Anyone who uses a motor vehicle is liable to cause bodily injuries to third parties as well as damage to his own property. For this reason the Act of 1878 in the United Kingdom insisted that a locomotive might not pass along a public highway at a greater speed than four miles per hour, and then only if accompanied by an attendant who had to walk in front of the vehicle with a red flag (Alfred Eke, p.45). If someone caused a bodily injury to a third party, his financial position might not enable him to pay to the injured person or his estate the indemnity assessed by the court. Effectively, what the court does in assessing the lump sum payable is to estimate the amount of net earnings lost (assuming normally that the victim would retire at 65 if male, or 60 if female), subject to the usual hazards of mortality and redundancy (Hey and others Section 2.6.3, p.377). Even if the person can pay for one injured party, he may not be able to pay for two or three or more if the accident caused injuries to more than one person.

For this reason the necessity of motor insurance was created and made a legal requirement. In the United Kingdom the Acts of 1896 and 1901 were the start of this legal requirement (Abdel Rahman, G., p.56). Afterwards the law (specifically the Road Traffic Act 1972) required that anyone using a motor vehicle on the road must have, in force, an insurance in respect of his legal liability to third parties, including passengers, for bodily injury (Benjamin, B., p.18). In Egypt, the Law No. 652 for the 1955 made insurance a legal requirement (Abdel Rahman, G., p.117).

Motor insurance is divided according to the type of risk covered by the policy. The subdivisions of motor insurance in Egypt are similar to the subdivision of motor insurance in the United Kingdom and these are:

EGYPT	UNITED KINGDOM
1. Third party bodily injury only	Third party bodily injury only
2. Full third party 3. Full third party, Fire and theft	Full third party
4. Comprehensive	Comprehensive
(Abdela Rahman, G. p.116)	(Benjamin, B. p.18)

1. Third Party Bodily Injury

By law in both Egypt and in the United Kingdom, anyone using a motor vehicle must possess an insurance policy to protect him against any responsibility towards the bodily injury of a third party caused from using the car.

2. Full Third Party

In Egypt the full third party cover covers the responsibility for bodily injury as well as the property damage of a third party.

In the United Kingdom full third party cover almost invariably extends to legal liability for damage to the property of third parties as well as to injury; it may be extended to include protection against the loss of the insured's own vehicle by fire or theft (Benjamin, B. p.18).

3. Full Third Party, Fire and Theft

In Egypt, this type of cover is the same as the full third party in the United Kingdom.

4. Comprehensive Cover

A comprehensive policy provides indemnity for the loss of, or damage to, the insured's own private car, liability for injury to third parties or damage to their property, together with extra benefits such as personal accident cover for the insured, payment of medical expenses incurred by injured occupants of the insured's car and loss of, or damage (subject to low monetary limits) to rugs, clothing or personal effects carried in the insured car (Benjamin, B. p.18).

In Egypt, the comprehensive cover is similar to the corresponding cover in the United Kingdom, but excluding the loss of, or damage to, rugs, parcels, clothing or personal effects (Abdel Rahman, G. p.118). Moreover, the covers, 2, 3 and 4, except the bodily injury of the third party, are called "Balance of Risk Insurance" and are issued in a separate policy (System and Tariff of Balance of Risk Insurance in Egypt, p.1).

1.2 **Purpose of the Study**

In motor insurance, not all its constituents are indemnity contracts and the principle of indemnity does not apply in the case of third party bodily injury. The principle of indemnity is: "...that the policyholder who normally has full cover should be in the same financial position after the claim is settled as before the incident which gave rise to the claim" (Benjamin, B. p.202).

"The life of a human being, or any part of his body is priceless" (Abdel Rahman, G., p.155), but in spite of this, the court may assess the value of human life or injury to any part of his body. "Effectively what the court does in assessing lump sums is to estimate the number of years of net earnings lost (assuming normally that the victim would retire at 65 if male, and 60 if female) subject to the usual hazards of mortality and redundancy. The court then uses a multiplier which is roughly equal to an annuity for that period, assuming a non-inflationary rate of interest and applies it to the net loss of earnings as at the date of trial" (Hey and others, Section 2.6.3 p.377). Therefore, the size of the claim depends on the number of injured persons, the age of each person and their occupations. Subsequently, the size of the claim will range from a small amount to an indefinitely large amount of money in the case of multiple claims in one accident. Table 1. (1) shows the casualty figures for road accidents in Britain in 1975-79:

Table 1. (1)

Road Accidents in Britain (excluding Northern Ireland) in the period 1975-1979

Year	Vehicles on the road 1975=100	Number of casualties	Killed	Seriously injured	Slightly injured
1975	100	324,950	6,366	77,122	241,462
1976	104	339,673	6,570	79,531	253,572
1977	108	348,061	6,614	81,681	259,766
1978	112	349,795	6,831	82,518	260,446
1979	112	333,776	6,327	80,274	247,175

Source: British Insurance Association (BIA), Insurance Facts and Figures 1979.

The insurance company, in order to be safeguarded, must assess the risk undertaken and establish adequate premium rates leading to the development of reserves sufficient to preserve solvency. In fact, the insurance company does this in the case of normal claims that have a definite upper limit by applying methods and formulae that give a satisfactory approximation to the true risk. A few papers have been written to deal with the large claims or the extreme claims* (as it will be referred to in this study), and approximate methods have been suggested to deal with these claims (Beard and others, 1978, p.29).

*The definition of the extreme claims will be the subject of Chapter Two.

Therefore, the purpose of the study is to find a more accurate estimation by using a statistical theory called "The Theory of Extreme Values". This theory will be used in a trial to analyse the observed extreme claims in motor insurance in an attempt to find a solution to the many problems facing the insurance companies as a result of these extreme claims.

1.3 **Outline of the Study**

This research includes seven chapters. The first chapter is an introduction which includes an abridged history of motor insurance in Egpyt compared with its counterpart in the United Kingdom and includes also the purpose of the study and the outline of the research.

The definition of extreme claims will be discussed in Chapter Two.

The aim of Chapter Three is to discuss current experience and the effect of extreme claims on the companies' results and subsequently, the use of the loss ratio in comparing the results of two different markets or companies with different premium rates. In addition, the effect of large claims on the company's portfolio and reinsurance arrangements is discussed.

Extreme value theory is presented in Chapter Four in a way that makes it applicable to the insurance field. Some modifications have been suggested to suit the insurance statistics and to overcome the difficulties that may be encountered in collecting the data.

Chapter Five is concerned with the data collected from the Egyptian Reinsurance Company.

The application of the theory has been carried out in Chapter Six and detailed calculations have been presented in order to facilitate subsequent work. some estimations concerning the Egyptian market have been reached.

Chapter Seven provides the conclusions from the work which has been done and the results produced by applying extreme value theory.

CHAPTER TWO

THE DEFINITION OF EXTREME CLAIMS AND THE CAUSE OF ITS EXISTENCE

2.1 The definition of the extreme claims:

'Extreme' is defined in the Oxford Dictionary as 'something carried to excess' and in Webster's New Collegiate Dictionary as 'something situated at a marking one end or the other of a range'.

According to the Oxford Dictionary's definition, the extreme claims are defined to be the claims carried to excess, i.e. those claims which exceeding a normal figure, while according to Webster's Dictionary definition, they are the claims situated at one end or the other of the range of the claims, i.e. those claims which they are extremely small and extremely large.

Since the extreme large claims have more influence on the company's portfolio therefore only the extreme large claims will be the subject of this study and extreme claims will be used in this sense.

Our definition of the extreme claim will not depart too much from the abovementioned definitions, but will be more specific and appropriate to the study. Therefore it is defined to be that claim which exceeds a pre-estimated figure, as is the case in fire insurance where the extreme claim is defined as the claim which exceeds the Estimated Maximum Loss (EML). In motor insurance, the extreme claim could be defined as the claim which exceeds the retention of the company in an excess of loss reinsurance cover*. Also, the extreme claim could be defined as the claim which exceeds a given figure. In this study we will use this definition as a general case and the first definition as a special case.

2.2 The origin of extreme claims:

There are two causes of the existence of the extreme claims in motor insurance;

(a) The first cause comes from the nature of the motor insurance cover. "From the point of view of the insurers, the liability under the third party section is in some respects heavier than that under the damage section, for while the limit of indemnity under the latter is the value of the insured car as shown in the policy schedule, the liability under the third party section is unlimited in amount, and the cost of claims largely depends upon the awards of the court." (Batten A.G.M. & Dinsdale, W.A. p.161).

(b) "Furthermore, in some cases, a variety of insured events not independent of each other may occur and cause a chain of losses with cumulative effect on the insurer (e.g. catastrophes, conflagrations and earthquakes). Such adverse events can disrupt the balance of

---

*The excess of loss reinsurance cover will be defined in Chapter Three.

the insured portfolio of risks, so that considerable discrepancies occur between the initial forecasts based on probability counts and the actual gross results." (Reinsurance problems in developing countries, p.1).

CHAPTER THREE

THE EFFECTS OF THE EXTREME CLAIMS ON COMPANY'S PORTFOLIO

The occurrence of a large claim on a company's portfolio will make the average claim differ considerably from expectation and therefore to distortion in the company's results. Generally, the occurrence of an extreme claim will be bad for a company's portfolio unless:

1. It has been estimated to a certain degree of accuracy in amount and the period of its occurrence.

2. It has been taken into account when the risk premium was calculated.

3. Adequate reserves are built up to meet this contingency.

4. A sufficient reinsurance cover is arranged.

The first point will be dealt with in detail in Chapter Six, while the other points will be the subject of the following paragraphs.

3.1 **The adequacy of premium:**

Some companies do not include the large claims in the statistics from which they produce the risk premium on the grounds that the claim is an exceptional claim or derives from an unusual accident. As a result of this misjudgement, the calculated premium does not reflect the actual experience and will fall short of meeting the resultant claims, instead the company sets aside a lump sum out of its net profit (after tax) as a reserve to meet such a contingency (this will be discussed in detail later in the chapter). This procedure will reduce the profit attributable to the shareholders.

To measure the adequacy of premiums, an index called 'loss ratio' is produced.

3.1.1 **Loss ratio:**

In a Glossary of Terms Used in Insurance issued by The Victory Insurance Company Limited, on p.10 they defined the loss ratio as: 'The proportionate relationship of incurred losses to earned premiums expressed as a percentage.

Earned Premiums = Premiums written (current year) PLUS premium reserve (prior year) LESS premium reserve (current year)

Incurred Losses = Claims paid (current year) PLUS outstanding losses (current year) LESS outstanding losses (prior year).

3.1.2 **The fallacy of loss ratio in judging bad (good) underwriting:**

Most of those working in insurance and reinsurance fields and in a position to judge on the results of underwriting experience use the loss ratio as a measure of goodness of underwriting, i.e. if the loss ratio $\geq 100\%$ they say that either the underwriting policy of the company is bad or generally the results of the market are not good, but if it is $< 100\%$ they say that the company is doing reasonably well.

This judgement is not completely correct where the loss ratio (L/R) is constituted of two factors, namely: the claims (c) in the numerator and the premiums (P) in the denominator.

If the premium (P) is under-estimated then the L/R will be greater than 100% and if the claims are unusual and larger than expected then L/R will be greater than 100% as well, but if the premiums are correctly estimated to reflect the experience of the written business (including the large claims), then the L/R could be kept under 100%. Therefore, those who judge on the results of an underwriting experience must look behind the loss ratio to find whether the premium is inadequate, the experience as a whole is bad or an unusually large claim has occurred.

3.2.3 **The adequate premiums:**

'No company can continue to operate at a loss for long. Indeed a loss situation is self-deteriorating since bad results undermine confidence and this affects not only investors but also customers. The company must try to win on both counts and as Ratcliff (1976) has put it, must deal, inter alia, with the problem of establishing premium rates on which business can be obtained and written with a profitable result.' (Benjamin, B. p.156).

But how can the company establish this adequate and acceptable premium rate? In countries like the United Kingdom, 'Premiums, which are ultimately based upon claims experience, are computed accordingly' (Batten, A.G.M. & Dinsdale, W.A., p.161), even the No-claim Discount (NCD) is applied to the compulsory motor insurance as well as comprehensive motor insurance, 'It should be borne in mind that any N.C.D. scale will probably be applied to all private car policy holders and all types of cover' (Benjamin, B., p.187). This premium must be revised regularly to match the change in experience and inflation. 'In motor insurance which provides indemnity with no fixed monetary sums insured, a regular revision of premium rates is necessary to match inflation and experience'(Benjamin, B., p.139).

In Egypt, the situation is different since any increase in the rate of premium suggested by the insurance companies must be approved by the government which may refuse to increase it or may limit it because it considers the type of insurance is for the national benefit and must keep its price down. 'The nationalized industries could give away their products for nothing. They could sell at zero prices' (McCormick, B.J. & Others, p.385). In the case of nationalized insurance companies this policy works since when the company runs into difficulties, the Government can help it. In other words, we can say that this type of insurance is indirectly subsidized by the Government. But as private insurance companies have entered the insurance market, the situation has changed and the Government will find itself choosing between four alternatives: to take this insurance as its own to keep the nationalized companies transacting it or to subsidize the private companies if they are transacting it or to free it from its control and then the rate of premium will reflect the actual experience of the market and then competition between the companies will be created, otherwise the private companies will refrain from issuing this type of insurance which can turn customers away and subsequently lead to a loss of profitable business.

Going back to the nationalized companies, the following table represents the claim ratio for direct transactions (before reinsurance) in five years (1975-1979 inclusive) for the compulsory motor insurance compared with the comprehensive motor insurance in the Egyptian market:

Table 3 (1)

Claim ratio in motor insurance
in Egypt in the period 1975-79

Year	Motor Compulsory %	Motor Comprehensive %
1975	190.9	85.5
1976	175.4	61.6
1977	298.1	61.1
1978	232.2	63.5
1979	245.2	61.2

Source: The Egyptian General Insurance Organisation, Insurance Activities in the Arab Republic of Egypt for the years 1975, 1977 and 1979.

Table 3 (1) shows the imbalance in compulsory insurance between the earned premiums and the incurred claims. The reason for this is a mixed bag of inadequate premiums, inflationary effects and bad results. Conversely, comprehensive insurance resulted in profit. The incurred claims represent about two-thirds of the earned premiums.

How do the nationalized companies finance the loss in the compulsory insurance? These companies are run as if they are private enterprises and are required to maximize their profits, therefore, they should charge the loss in compulsory motor insurance to the other branches of general insurance which, in its rating policy, does not need the consent of the Government. This policy makes the other types of insurance - non-compulsory insurance - expensive compared with its actual cost, i.e. the premiums charged are much higher than the actual cost. In an economic language, if we apply the price-cost ratio (McCormick & Others, pp.344 & 368), we conclude that these companies are not competing with each other, but because they are the only three companies conducting the insurance business in Egypt, they constitute between themselves an oligoboly status.

In the absence of competition, this policy was working well in spite of the unfair premium charged. Recently, when the private companies came into existence they broke the united premium rate which is known as 'the tariff rate' and started to write business at lower rates. In this way they could attract a large number of customers from the nationalized companies. The nationalized companies found themselves facing rivals in the market. How can they face this challenge and what should their pricing policy be?

To preserve their share of the market they must reduce the premium rates in the other types of insurance - non-compulsory motor insurance - to the level that can compete with the rates of the private companies. 'It is clearly necessary for the ruling rates in the market for each class of insurance to be considered to ensure that the rates produced are competitive' (Benjamin, B. p.146). In respect of the compulsory motor insurance, there are two policies open to the nationalized companies to follow:

(a) The company may stick to the recent prices, financing the losses from the profit of the other branches, but this method could lead to an overall loss in the company business. If this loss lasts many years in succession it will ruin the company; or

(b) They may regard the office premium as constituted of two parts: the risk premium and the expenses represented in salaries, wages, staff benefits, property costs, general management expenses ...... etc., and these expenses can be considered fixed costs, 'expenses as normally considered in insurance are virtually all in the category of fixed expenses for pricing purposes' (Benjamin, B., p.146). From this point of view, the nationalized companies can continue writing compulsory motor insurance charging a minimum premium equal to the risk premium, assuming that no commission is paid. In general, the premium will equal the variable cost where the fixed cost will be met from the profit of the other branches of insurance.

3.2.4 From the above-mentioned, we conclude that if we freed the rating policy from the control of the Government and if we used the actual experience (including the large claims) to produce the premium rate then there would be a balance between the earned premiums and the incurred claims every year. This is not completely true because the large claims do not occur each year. Even so, the total amount of claims will differ from year to year and the company must protect itself from the fluctuation in the numbers and amounts of the large claims. This could be done by building up adequate reserves and arranging a sufficient reinsurance cover to meet this contingency.

3.3 The Reserves

In non-life insurance, an insurance company, in return for the payment of a premium, is liable to make a monetary payment to the insured or on his behalf, on the occurrence of a specified event which should happen within a specified period of time where both the event and period are covered by the insurance contract. Since the amount of claim and time of its occurrence are random, therefore, the insurance company must allocate a certain amount of money to meet these claims as they fall due. This amount is called 'a reserve' and the reserve can be assessed at any point of time and can be divided into the following categories:

(a) Unearned Premium Reserves (U.P.R.)

This reserve is the proportion of premiums received which is attributable to the period of cover still outstanding.

(b) Unexpired Risk Reserves:

If the unearned premium reserves are shown to be less than the expected liability resulting from the outstanding period then the difference is known as 'the unexpired risk reserve'.

(c) Outstanding Claims Reserve:

This is the outstanding liability for claims that have been notified to the company and not settled.

(d) The Incurred But Not Reported Claims Reserve (I.B.N.R)

Prior to the valuation date of reserves, a number of incidents will have occurred but will not then have been reported to the company. The I.B.N.R. reserve is the estimated liability for these claims. "Not infrequently, payments have to be made on claims thought to have been finally settled and a reserve for these re-opened claims needs to be established. Statistically, they can be dealt with on similar lines to the I.B.N.R. claims or included with them." (Benjamin, B., p.237).

3.3.1 **Unexpired Premium Reserves**:

Since the majority of the insurance policies fall into more than one accounting period, a proportion of the premiums written in a given accounting period are reserved to meet the liability arising after the closing date of the accounting period. There are different methods to calculate the U.P.R. which are outside the scope of this research and are given by Benjamin, B., p.235, and Abbott & Others, 1981, p.119. Since the U.P.R. is calculated by reference to the premium charges, therefore, "if the premiums are inadequate then any amount set aside from these premiums to cover future claims could also be inadequate" (Abbott & Others, 1981, p.119).

3.3.2 **Unexpired Risk Reserves (U.R.R.)**:

The unexpired risk reserve attempts to calculate the true cost of future claims and if this amount is greater than the unearned premium reserve, an additional reserve is required. Benjamin has presented a method of calculating these reserves which depends on the expected claims ratio and unearned premium and declared that this method suffers from the lack of knowledge of the total incurred claims liability for the most recent, and therefore, most relevant, claim years which will not yet have been fully developed, and he advised the addition of a prudent fluctuation margin which represents the degree of uncertainty in the elements of the estimation procedure (Benjamin, B., p.241). Abbott & Others have presented another method which uses claim frequency and average claim amount factors (Abbott & Others, 1981, p.119).

3.3.3 **Outstanding Claims Reserves**:

The outstanding claims reserve may be estimated by either the case estimation method or one of the statistical methods. The case estimation method consists of the examination of each outstanding claim by an individual experienced in claim administration while the statistical methods (Benjamin, B., pp.244-254) depend on statistical analysis of claims data collected from the actual experience of the company.

"Case estimation is less applicable to situations where claims may take a considerable period of time for the total extent of liability to become apparent, or where liability may be contested" (Reid, 1978, p.211). Moreover, Benjamin has mentioned that the method may be impracticable where there are large numbers of claims (Benjamin, B., p.243).

Statistical methods are impracticable where there is only a small number of claims.

In the case of a large number of claims, the very large claims which distort the data should be removed.

3.3.4 **I.B.N.R. Reserves:**

Longer periods may arise when the claimant, usually a third party, makes his claim against the insured, therefore, it may take several months before the claim is formally reported to the insurer.

The insurer, at his accounting date, should make an assessment of such incurred but not reported claims.

The I.B.N.R. reserves may be calculated by approximate methods or by a more accurate method suggested by Clarke, T.G. and his colleagues.

3.3.5 **Other Reserves:**

The insurance company may establish another reserve or reserves as Catastrophe Reserve or Claims Equalisation Reserve to meet the large claims at the tail of the frequency distribution or the claims of catastrophic nature or to meet any fluctuation in the actual claims experience.

There is no specific basis for the estimation of these reserves and they have to be set up out of taxed income so that the company has to consider the effect of setting up these reserves upon its results and investment policy and should find a way to assess such reserves to a certain degree of accuracy.

From the abovementioned considerations, any insurance company may find that the four categories of reserves are not enough to protect itself against unpredictable events so it may set up another category of reserves which has no basis for calculation. Extreme value theory may well be the basis for such a complementary reserve.

3.4 **Reinsurance** :

Since the insurance company is liable to incur unexpected very large claims or an unexpected large number of claims which can cause financial strain on the comapny's assets, or can jeopardize the solvency margin, each insurance company should make a reinsurance arrangement to spread the risk and protect itself. The reinsurance arrangements may be contracted between the original insurers and another insurance or reinsurance company in a formal reinsurance contract which is called a "treaty".

Reinsurance arrangements may take one of the following reinsurance forms:

3.4.1 **Facultative**:

The original form of reinsurance was facultative cession of business where each individual risk on which reinsurance was required was offered to the reinsurer without any obligation from the reinsurance company to accept it.

When the number of risks ceded by a facultative method had increased, the treaty reinsurance replaced the voluntary nature of cessions and acceptances as well as the individual risk-by-risk approach.

Under treaty reinsurance there is a compulsory cession and compulsory acceptance of all risks falling within the scope of the treaty. There are two types of treaty covers, namely: the proportional and non-proportional covers.

3.4.2 **Proportional Treaties** :

When this form of reinsurance is adopted, the reinsurer covers a proportion of the risk covered by the original policy and the premiums and claims are shared by the reinsurer and the ceding company in the same proportion. The proportional form may be subdivided into:

(a) Quota Share Treaty:

This is the simplest form of reinsurance under which the whole business written by the ceding company and covered by the treaty is apportioned pro rata between the ceding company and its reinsurers. The same percentage share is ceded to the reinsurers out of each and every risk within the scope of the treaty, therefore, the reinsurers' fortunes follow exactly those of the ceding company.

(b) Surplus Treaty:

Under this treaty each risk exceeding the retention line of the ceding company is apportioned as to premiums and claims between the company and the reinsurers. This proportion varies from one risk to another, therefore the reinsurer follows the fortunes of the ceding company in respect of each risk ceded and not the overall treaty.

3.4.3 **Non-Proportional Treaties**:

This type of treaty limits the loss of the ceding company. It is subdivided into:

(a) Excess of Loss Treaty:

This is an arrangement whereby a ceding company decides the maximum amount which it is prepared to bear in respect of any one claim or accumulation of claims arising out of one event. This maximum amount is called "retention". The reinsurance protection operates if, and only if, the ultimate net loss arising out of any one claim, or accumulation of claims arising out of one event exceeds the retention, the reinsurers' commitment being to answer for the excess over and above such retention up to an agreed limit. A first excess of loss cover may be followed by a second, operating after the first excess limit, and may further be followed by one or more other excess of loss layers.

(b) Stop Loss (or Excess of Loss Ratio) Treaty:

This is an arrangement "under which the insurers protect the total results out of a specific class of insurance from any excess in its loss ratio over and above an agreed loss ratio. Such a loss ratio is normally fixed at a level that would leave no margin of profit to the insurer on his net account of the class covered. The protection of the cover extends up to some prearranged limit fixed in terms of loss ratio or in an absolute amount, or in terms of both, whichever would give a lower figure. In most cases, the insurer has to retain for his net account, in addition to the excess loss ratio (over which the cover operates). A certain percentage of the cover itself" (Reinsurance Problems in Developing Countries, 1973, p.12).

3.4.4. **Motor Reinsurance:**

Since the results of third party motor insurance fluctuates in both the number of claims and the size of each claim and large claims may emerge, the most appropriate form of reinsurance is the excess of loss treaty. Even with such reinsurance the company has to choose between two risks, either to give away too much premium or to retain a large risk, therefore each insurance company has to have a reinsurance cover planned in such a manner that it retains for its net account the optimum business with the minimum cost. To achieve economy of cost implies primarily the choice of a plan which is not only adequate, but which is also expected to result in the minimum net outgo for reinsurance, i.e. choosing the suitable retention. "One of the greatest technical problems in reinsurance planning is the correct retention limits for each risk. In fact, if these limits are too low, part of the company's retention potential will be wasted and unnecessary reinsurance will be taken out. If limits are too high, fluctuations beyond the tolerable margins are bound to occur" (Reinsurance Problems in Developing Countries, p.8).

Many studies have been carried out in this regard (Benjamin, B. p.219).

Another approach by Beard (Beard, 1963b, p.6), Jung (Jung, 1965, p.178) and d'Hooge (Coe, p.17) has been made to solve the technical problem and to calculate the true excess of loss premium by using extreme value theory, as will be shown in the next chapter.

CHAPTER FOUR

The Extreme Value Theory

4.1.1 Introduction

The theory of extreme values arises as a special branch of the study of order statistics. L. von Bortkiewicz (1922) was the first to study extreme values [Gumbel, E.J., 1958]. Since then many authors have contributed to the theory. E.J. Gumbel has a major share in the development of this theory. He started his research in 1932 in Germany and developed it further in France until 1940. He has continued his research since then in the United States [Gumbel, E.J., 1954]. The theory has many applications in different fields of research such as floods, droughts, meteorological phenomena, gusts, breaking strength, quality control and the oldest ages in life tables. Galambas, J. has contributed to this theory by his recent book "The Asymptotic Theory of Extreme Order Statistics".

In the insurance field, in 1962 Beard, R.E., invited the actuaries to make their contributions. "There are a number of problems arising from insurance, particularly in non-life insurance to which the theory may be usefully applied." [Beard, R.E. 1963a p.313]. Since then many researches have been published investigating the usefulness of the theory in the calculation of an excess of loss premium. In recent times there are researches in connection with fire insurance problems. Among those who have contributed are Bearad [Astin Bulletin Vol. II, Part 3, Vol. III, Part 2]. Jung, J. [Astin Bulletin Vol. III, Part 2], Ramachandran, G. [Astin Bulletin Vol. III] and Leonard D. Coe who made his contribution by publishing his paper entitled "Extreme Values and Retentions".

4.1.2 Application of the theory in Motor Insurance

"The aim of a statistical theory of extreme values is to analyse observed extremes and to forecast further extremes. The essential condition in the analysis is the distribution from which the extremes have been drawn and its parameters must remain constant in time (or space), or the influence that time (or space) exercises upon them must be taken into account or eliminated. Another limitation of the theory is the condition that the observations from which the extremes are taken should be independent." [Gumbel, E.J. 1958, p.1.] There is another condition in applying the theory, i.e. that "the theory of extreme values will be centered about the exponential distribution" [Gumbel, E.J. 1958, p.113]; "there is therefore a prima facie case for expecting the theory to be very suitable for use with insurance claim distribution", [Leonard D. Coe, p.3].

By translating these conditions into insurance language, the conditions will be:

1. The largest claims arising from a portfolio in successive equal exposed periods are known.

2. The monetary values and other factors have been unchanged over the period.

3. Probability of claim has remained constant. [Beard, R.E., 1963a, p.313].

4. The distribution of the claim amount is of the exponential type.

Actuarial literature suggests that the log-normal distribution, for example, fits claim distribution in fire and motor insurance, whilst the Pareto distribution has been used extensively in demonstrating aspects of reinsurance, especially excess of loss problems [Leonard D. Doe, p.3]. Moreover, Lars-Gunnar Benckert has proved by practical examples that "The log-normal density function shows a good fit - especially for large values of the variable - in many branches of non-life insurance" [Astin Bulletin, Vol. II, Part 3, p.22]. One of his examples was Motor Third Party insurance (p.14). The claim distribution of Motor Third Party follows an exponential type so that the fourth condition is satisfied. To satisfy the first three conditions, we choose a suitable number of equal intervals (years) where the probability of claim can be assumed to be unchanged, then the observed claims are adjusted for changes in the value of money by using a suitable index number, e.g. Retail Price Index.

For the abovementioned reasons we find that the theory can be useful in dealing with the extreme claims arising from motor third party insurance.

The theory will be used to give the distribution of the number of excesses and the probability of the extreme values.

4.2.1 The distribution of the number of exceedances:

Suppose that ζ is a continuous variate and we have n observations of such variates ordered in a descending magnitude:

$$\zeta_1, \zeta_2, \zeta_3, \ldots, \zeta_m, \zeta_{m+1}, \ldots, \zeta_n$$

where ζ_1 is the largest and ζ_n is the smallest.

Let x be the number of observations which will be equal or exceed ζ_m out of N future claims, x is called the number of exceedances satisfies the condition: $0 \leq x \leq N$, N is not necessarily equal to n.

Suppose that the probabiliy for a value to be $< \zeta_m$ is $q = F_m$, and the probability of a value to be $\geq \zeta_m$ is $p = 1 - F_m$.

$\therefore$ If the probability F_m is known

$\therefore$ the probability that x among N future observations will be equal to or exceed ζ_m is the binomial formula:

$$w_1(1-F_m,N,x) = \binom{N}{x} (1-F_m)^x F_m^{N-x} \tag{1}$$

[Gumbel, E.J. 1958, p.58].

Since F_m is unknown and the only known data are the n past observations, then by a procedure analogous to the derivation of the Bernoulli distribution [Gumbel, E.J., 1954, p.9], the distribution of x will be given by:

$$w(n,m,N,x) = \frac{\binom{n}{m} m \binom{N}{x}}{(N+n) \binom{N+n-1}{m+x-1}} \tag{2}$$

$$1 \leq m \leq n, \; 0 \leq x \leq N$$

To simplify the numerical calculation, formula (2) may be written in the form:

$$\binom{N+n}{n} w(n,m,N,x) = \binom{N+n-m-x}{n-m} \binom{x+m-1}{m-1} \tag{3}$$

Formula (2) was derived by H.A. Thomas and he applied it to discharges of water out of dams and obtained very good results [Gumbell, E.J., 1954, p.9].

The probability that the largest value ζ_1 will be exceeded x times is given by:

$$w(n,1,N,x) = \frac{n\binom{N}{x}}{(N+n)\binom{N+n-1}{x}}$$

$$= \frac{n}{n+N}\ \frac{N!}{(N+n-1)!}\ \frac{(N+n-x-1)!}{(N-x)!} \qquad (4)$$

and this probability diminishes with x.

4.2.2 The Mean Number of Exceedances $\bar{x}_m$:

Since x_m is the mean number of exceedances over the mth largest value in N future trials, its value is given by:

$$\bar{x}_m = \sum_{x=0}^{\infty} x\, w(n,m,N,x) = \sum_{x=0}^{\infty} x \frac{\binom{n}{m} m \binom{N}{x}}{(N+n)\binom{N+n-1}{m+x-1}}$$

$$= m \frac{N}{n+1}$$

$$\bar{x}_m = m \frac{N}{n+1} \qquad (5)$$

The mean number of exceedances over the largest value x_1 is

$$\bar{x}_1 = \frac{N}{n+1} \qquad (6)$$

4.2.3 **The Variance of the Number of** Exceedances σ_m^2:

The variance of x is given by:

$$\sigma_m^2 = \sum x^2 w(n,m,N,x) - x_m^2$$

$$\sigma_m^2 = \frac{m(n-m+1)}{(n+1)^2} \cdot \frac{N(N+n+1)}{n+2} \qquad (7)$$

From (7) σ_m^2 is increasing with the increase of N and/or with the decrease of n, and it is a maximum when $m = \frac{n+1}{2}$ i.e. for the median of the original observations.

The variance of the number of exceedances over the largest value is given from (7) by putting $m = 1$

$$\sigma_1^2 = \frac{n}{(n+1)^2} \frac{N(N+n+1)}{n+2} \qquad (8)$$

The variance of the number of exceedances over the median is given from (7); by taking $m = \frac{n+1}{2}$, n is odd

$$\sigma^2_{\frac{1}{2}(n+1)} = \frac{(n+1)(n+1)}{2.2(n+1)^2} \cdot \frac{N(N+n+1)}{n+2}$$

$$= \frac{N(N+n+1)}{4(n+2)} \qquad (9)$$

from (8) and (9) $$\frac{\sigma^2_{\frac{1}{2}(n+1)}}{\sigma_1^2} = \frac{(n+1)^2}{4n}$$

$$\simeq \frac{n}{4} \qquad (10)$$

From (10) we deduce that the variance of the number of exceedances over the median is about 1/4 n times as large as the variance of the number of exceedances over the largest value.

4.2.4 The Distribution of Rare Exceedances:

Suppose that both n and N are large, therefore $n \simeq N$ and by substituting in (2)

$$w(n,m,n,x) = \frac{(m+x-1)!\,2^{-m-x}}{(m-1)!\,x!}$$

$$= \binom{m+x-1}{x}\left(\tfrac{1}{2}\right)^{m+x} = w(m,x)$$

$$w(n,m,n,x) = \binom{m+x-1}{x}\left(\tfrac{1}{2}\right)^{m+x} \qquad (11)$$

which is the probability that the mth largest value will be exceeded x times in n future trials and this probability does not depend on n.

Since m is small compared with n (when n is large), therefore equation (11) is called "the distribution of the rare exceedances".

For the largest value, when $m = 1$, the probability that the number of exceedances will be x is given by:

$$w(1,x) = (\tfrac{1}{2})^{x+1} \tag{12}$$

From (12), the probability that the largest value will not be exceeded, i.e. $x = 0$, is:

$$w(1,0) = (\tfrac{1}{2}) \tag{13}$$

From (11), the probability that the mth largest value will not be exceeded, i.e. $x = 0$, is:

$$w(m,0) = (\tfrac{1}{2})^{m} \tag{14}$$

4.2.5 The Mean and the Variances of the Number of Rare Exceedances:

From equation (5) take $n = N$, and since n is large, the mean of the number of rare exceedances $\bar{x}$ is given by

$$\bar{x} = m \tag{15}$$

In the same way, the variance of the number of rare exceedances σ is given from (7) by taking $n = N$

$$\sigma^2 = \frac{m(n-m+1)}{(n+1)^2} \quad \frac{n(2n+1)}{n+2}$$

$$= \frac{2m(n-m+1)}{(n+1)^2} \quad \frac{n(n+\frac{1}{2})}{n+2}$$

since n is large compared with m

$$n - m + 1 \simeq n + 2 \simeq n + 1 \simeq n + \tfrac{1}{2} \simeq n$$

$$\sigma^2 = 2m \qquad (16)$$

From (15) and (16) we can deduce that distribution (11) is similar to the Poisson Distribution with parameter m but the variance of (11) is double the variance of the Poisson Distribution.

4.2.6 **The Estimation of the Number of exceedances and Rare exceedances from the Portfolio of an Insurance Company:**

If an insurer has a portfolio of L policies at risk and if the expected mean value of claim frequency (crude rate of claims) for these policies during a specified period is h, then the expected number of claims $N^* = hL$ (17)

h is taken from past experience and modified to meet the changes in the risk factors.

From (17) and (2) we can find the distribution of the number of exceedances in the next period. Since "large claims exercise a critical effect on the performance of an insurer whose top risks are not cut away by reinsurance" [Ramachandran, G., 1974, pp.305-306], therefore, it is useful to apply the theory to control the insurer's portfolio by calculating N in order that there is a given probility α such that the largest claim among n claims will not be exceeded, so if we take $m = 1$, $x = 0$ and $w = \alpha$ in equation (2),

$$\therefore \frac{n}{N+n} = \alpha, \quad \text{therefore} \quad N = \frac{(1-\alpha)n}{\alpha} \qquad (18)$$

By substituting from (18) in (17)

$$\frac{(1-\alpha)n}{\alpha} = hL \quad \therefore L = \frac{(1-\alpha)n}{\alpha h} \qquad \underline{(19)}$$

Equation (19) now proposed by the author gives an estimate of the number of policies which should be written in order that no one claim in the next period will exceed the largest claim which has occurred out of n policies written in the last period (year).

Now we can answer the question: "What is the probability that in the next period (year) at least one claim $> \zeta_m$ will occur?".

If $n \simeq N$ and both of them are large, therefore from equation (11), the required probability $= \sum_{X=1}^{\infty} w(m,x)$ (20)

If n is not large, then from (19) we find an estimation of N and by substitution in (2), the required probability

$$= \sum_{X=1}^{\infty} w(n,m,N,x) \qquad (21)$$

According to this probability, the company can take the decision either to retain the whole risk or reinsure part of it on an Excess of Loss basis.

4.2.7 **The application of the Theory in the Reinsurance Business:**

"It may be that quite unexpected results may be shown and these may be due to an isolated large claim, ... The reinsurer will want to know the reasons and take them into account in his rating. He may well have countrywide statistics for the class concerned, and will use these in conjunction with the individual company's statistics to arrive at the expected statistics." [Excess of Loss Method of Reinsurance, p.11].

From the way in which the reinsurer collects his statistics and the type of statistics he collects, it is clear that the Statistical Method [Beard, R.E. and others, 1978, p.27] can reach an estimation of the distribution function of the claim amount more or less accurately. By using this method, we can find the average amount of claim in excess over certain retention, and by using the formula [Benjamin, B., 1977, p.219], we can calculate the net premium of the excess of loss reinsurance as:

$$P(x) = H(x) \quad M(x) \tag{22}$$

where x is the retention,

$P(x)$ is the net premium,

$H(x)$ = the expected number of claims that will exceed the retention x, and

$M(x)$ = the average excess claim over retention.

$H(x)$ in equation (22) is the correspondence to $\bar{x}_m$ in equation (5) and therefore it will be calculated from (5). $M(x)$ may be calculated by the Statistical Method mentioned above, but if there are not sufficient data to apply the Statistical Method, the reinsurer can use the mid range of the excess of loss layer as an estimation of $M(x)$ assuming a uniform distribution of claims over the range of the layer, or can use another estimation produced from his experience, but this method could not be used in the case of an unlimited layer.

Benjamin mentioned that "there is a difficulty to which Guaschi (1969) draws attention in that the reinsurer requires, for the reliable calculation of the premium, data covering either the whole of the market or at least a number of large ceding offices who would have to be prepared to arrange their records to assist the reinsurance market. Another difficulty is that in order to estimate $H(x)$ and $M(x)$ we need to consider a claim amount distribution which starts at x and is thus truncated." [Benjamin, B., 1977, p.219.]

However, the author has found no difficulties with $H(x)$ which can be estimated by using the theory, while $M(x)$ can be estimated by the Statistical Method or by another method of estimation.

The theory of the distribution of the number of exceedances not solve completely the uncertainty of the reinsurer but it may reduce it.

4.3.1 **The Distribution Function of the Largest Values. The Theory:**

If we define $F(x)$ as the probability that a variate is less than or equal to x, and if we have n independent observations, then the probability that all are less than or equal to x is $F(x)$ raised to the power n, and this is denoted by $\Phi_n(x)$

$$\Phi_n(x) = F^n(x) \tag{23}$$

where F is the initial probability function.

Now, if all the observations are less than x, then the largest of the n observations is less than x, therefore $\Phi_n(x)$ is the probability that the largest of the n observations is less than x.

To obtain the asymptotic distribution of the largest value, we order the n observations in a descending order of magnitude and by using Taylor's expansion of the initial probability function $F(x)$ about b_m where b_m is the characteristic mth largest value, where m is counted from the top.

$$F(x) = F(b_m) + \frac{x - b_m}{1!} F'(b_m) + \frac{(x-b_m)^2}{2!} F''(b_m) + \ldots\ldots + \frac{(x-b_m)^r}{r!} F^{(r)}(b_m) + \ldots \qquad (24)$$

where $F(b_m) = 1 - \frac{m}{n}$ (25)

$$F'(b_m) = \frac{d}{dx} F(b_m) = f(b_m) \; ; \; F''(b_m) = f'(b_m), \ldots$$

If we dropped the m from b_m and denoted $f(b_m)$ by f_m and substituted $F(b_m)$ in (24) by (25), therefore,

$$F(x) = 1 - \frac{m}{n} + \frac{x-b}{1!} f_m + \frac{(x-b)^2}{2!} f'_m + \ldots + \frac{(x-b)^r}{r!} f_m^{r-1} + \ldots \qquad (26)$$

By multiplying and dividing (26) by $\frac{m}{n}$

$$F(x) = 1 - \frac{m}{n} + \frac{m}{n} \frac{x-b}{1!} \frac{n}{m} f_m + \frac{m}{n} \frac{(x-b)^2}{2!} \frac{n}{m} f'_m + \ldots + \frac{m}{n} \frac{(x-b)^r}{r!} \frac{n}{m} f_m^{r-1} + \ldots \qquad (27)$$

If x becomes very large, the density of the probability $f(x)$ becomes very small and the same holds for $1-F(x)$, and if the variate is unlimited, the derivative $f'(x)$ converges to zero, whence L'Hopital's rule:

$$\frac{f(x)}{1-F(x)} = -\frac{f'(x)}{f(x)} \tag{28}$$

is applicable.

From (25) and (28) $\dfrac{f_m}{1-\left(1-\frac{m}{n}\right)} = -\dfrac{f'_m}{f_m}$

$$\therefore \frac{f'_m}{f_m} = \frac{n}{m} f_m = a_m$$

$$a_m = \frac{n}{m} f_m \tag{29}$$

a_m is a function of n, m and is known as the "Intensity Function" and it is analogous to the force of mortality in the actuarial statistics.

Since $-\dfrac{f'_m}{f_m} = a_m \qquad f'_m = -a_m f_m$

and from (29) $f'_m = -\dfrac{n}{m} f_m^2$

$$\frac{n}{m} f'_m = -\left(\frac{n}{m}\right)^2 f_m^2 = -a_m^2$$

substituting in (27)

$$F(x) = 1 - \frac{m}{n} + \frac{m}{n}(a-b)a_m - \frac{m}{n}\frac{(x-b)^2}{2!} a_m^2 + \ldots$$

$$+ \frac{m}{n}\frac{(x-b)^r}{r!} a_m^r + \ldots$$

$$= 1 - \frac{m}{n}\exp\left[-a_m(x-b)\right] \tag{30}$$

If we use the transformation

$$y_m = a_m(x-b_m) \tag{31}$$

$$\therefore F_m(x) = 1 - \frac{m}{n} \exp(-y_m) \tag{32}$$

y_m is called "the reduced mth largest value" and it is similar to the standardised normal deviate.

Equation (32) holds for any initial probability function of the exponential type and for large values of x, the distribution f(x) under the same conditions is given by differentiating (32) with respect to x

$$\therefore f_m(x) = \frac{m}{n} a_m e^{-y_m} \tag{33}$$

The probability that the mth largest value of the n observations (m is measured from the top) is less than x is given by:

$$\Phi_m(x) = \sum_m^n \binom{n}{m} (1-F)^m F^{n-m} \tag{34}$$

where F is given from (32).

By differentiating (34) the probability density of the mth largest value is given by:

$$\phi_m(x) = n\binom{n-1}{m-1} (1-F)^{m-1} F^{n-m} \, dF$$

$$= \frac{n!}{(m-1)!(n-m)!} (1-F)^{m-1} F^{n-m} f \qquad (35)$$

where f is given from (33).

If n is large and by substitution from (32) and (33)

$$\phi_m(x) = \frac{n^m}{(m-1)!} \left(1 - \frac{m}{n} e^{-y_m}\right)^{n-m} \left(\frac{m}{n}\right)^{m-1} e^{-(m-1)y_m} a_m \frac{m}{n} e^{-y_m}$$

by taking the limit as n tends to infinity, the asymptotic distribution of the mth largest value is given by:

$$\phi_m(x) = \frac{m^m}{(m-1)!} a_m \exp\left[-my_m - me^{-y_m}\right] \qquad (36)$$

where a_m, b_m and y_m are given by (29), (25) and (31) respectively.

a_m and b_m are called the extremal parameters where b_m is a certain average and a_m is a certain measure of dispersion [Gumbel, E.J., 1954, p.13] .

From equation (3) $F(b_m) = 1 - \frac{m}{n}$,

by differentiating w.r. to n

$$f(b_m)db_m = \frac{m}{n^2}dn$$

$$f(b_m)db_m = \frac{m}{n}\frac{dn}{n}$$

$$= \frac{m}{n} d \log_e n$$

$$\frac{db_m}{d \log n} = \frac{m}{nf(b_m)} = \frac{1}{a_m}; \text{ from } (29)$$

$$\therefore \frac{db_m}{d \log n} = \frac{1}{a_m} \qquad (37)$$

From (37) $\frac{1}{a_m}$ measures the increase of the expected largest value with the logarithm of the sample size. This relation is called "the trend of logarithmic increase of the extremes" [Gumbel, E.J., 1954, p.20].

4.3.2 The Asymptotic Probability of the mth Extreme:

The asymptotic probability $\Phi_m(y_m)$ of the mth reduced largest value y_m is given by integrating (36).

$$\Phi_m(y_m) = \int_{-\infty}^{y_m} (m\bar{e}^{y})^{m-1} \exp(-m\bar{e}^{y})\, m\bar{e}^{y}\, dy/_{(m-1)!} \tag{38}$$

$$\text{Put } \; m\bar{e}^{y} = z \tag{39}$$

$$\Phi_m(y_m) = \int_{m\bar{e}^{y_m}}^{\infty} z^{m-1} \bar{e}^{z}\, dz/\Gamma(m) \tag{40}$$

which is incomplete gamma function.

By integrating (40) by parts, we can express the successive probabilities for the mth extremes by the probability of the largest value,

$$\Phi_m(y_m) = \Phi_1^m(y) \sum_{r=0}^{m-1} \frac{m^r e^{-ry_m}}{r!} \tag{41}$$

where $\Phi_1^m(y)$ is the probability of the largest value raised to the power m and is given by taking $m = 1$ in (38) and by integrating, it gives the asymptotic probability of the first largest value $\Phi(y)$ as follows:

$$\Phi_1(y) = \exp(-\bar{e}^{y}) \tag{42}$$

by taking $m = 2, 3, \ldots$ and substituting from (42) in (41), we obtain the asymptotic probability of the second, third, fourth, ... largest values as follows:

$$\text{2nd largest} \quad \Phi_2(y_2) = (1+2e^{-y_2}) \exp(-2e^{-y})$$

$$\text{3rd largest} \quad \Phi_3(y_3) = (1+3e^{-y_3}+ \frac{9}{2} e^{-2y_3}) \exp(-3e^{-y})$$

$$\text{4th largest} \quad \Phi_4(y_4) = (1+4e^{-y_4}+ \frac{16}{2} e^{-2y_4}+ \frac{64}{6} e^{-3y_4}) \,.\, \exp(-4e^{-y})$$

The values of the cumulative probability function $\Phi(y)$ has been tabulated in the "Probability Tables for the Analysis of Extreme-Value Data", p. 17 for different values of y, while Gumbel, E.J., 1958, p.190 has tabulated the cumulative probability functions of the first three largest extremes for values of $y =$ -1.5 (0.5) 6.0. The cumulative probability functions for largest extremes of order four and over may be calculated from equation (41) and the values of the first extreme produced from the tables of the extreme-value data.

4.3.3 **Mean and Variance of the mth Largest Extreme**

The moment generating function of the mth largest reduced value y_m is obtained from equation (38) and defined as:

$$G_m(t) = \frac{m^{m-1}}{(m-1)!}\int_{-\infty}^{\infty} \exp\left[-(m-1-t)y_m - me^{-y_m}\right] me^{-y_m} dy_m \tag{43}$$

By using the transformation:

$$me^{-y_m} = z \quad , \quad me^{-y_m}\, dy_m = -dz$$

$$G_m(t) = \frac{m^t\,\Gamma(m-t)}{(m-1)!} \tag{44}$$

By taking the logarithms of $G_m(t)$

$$\log G_m(t) = t \log m + \sum_{r=1}^{m-1} \log\left(1 - \frac{t}{r}\right) + \log\Gamma(1-t) \qquad m \geq 2 \tag{45}$$

From equation (45), the means of the largest reduced values are given by:

$$\bar{y}_m = \log m - \sum_{r=1}^{m-1} \frac{1}{r} + \gamma \tag{46}$$

where γ = Euler's Constant

$$= \lim_{r \to \infty} [(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{r}) - \log (r)] = 0.577215$$

Hence, from equations (31) and (46), the mean of the variate itself is given by:

$$\bar{x}_m = b_m + \frac{\bar{y}_m}{a_m} \qquad (47)$$

The variance of the mth largest reduced values are given by:

$$\sigma^2_{y_m} = \frac{\pi^2}{6} - \sum_{r=1}^{m-1} \frac{1}{r^2} \qquad (48)$$

From equations (31) and (48), the variance of the variate itself is given by:

$$\sigma^2_{x_m} = \sigma^2_{y_m} / a^2_m \qquad (49)$$

The mean and variance of the first 10 largest reduced variates are tabulated in Gumbel, E.J., 1958, p.194, while Ramachandran, G. has made an extension to this table to m = 40 [Ramachandran, G., 1972, p.38]. The values given in table 4.(1)

.

The mean and variance of the first largest reduced value are

$$\begin{aligned} \bar{y} &= \gamma = 0.577216 \\ \sigma^2_y &= \frac{\pi^2}{6} = 1.64493 \end{aligned} \qquad (50)$$

Therefore, the mean and variance of the original variate are given by:

$$\left.\begin{aligned} \bar{x} &= b + \frac{\gamma}{a} \\ \sigma^2_x &= \frac{\pi^2}{6a^2} \end{aligned}\right\} \qquad (51)$$

Table 4. (1)

Order (m)	Extremes (y_m)	
	Expected Value $\bar{y}_m$	Variance (σ_m^2)
1	0.5772	1.6449
2	0.2704	0.6449
3	0.1758	0.3949
4	0.1302	0.2838
5	0.1033	0.2213
6	0.0857	0.1813
7	0.0731	0.1535
8	0.0637	0.1331
9	0.0565	0.1175
10	0.0508	0.1051
11	0.0461	0.0951
12	0.0422	0.0869
13	0.0390	0.0799
14	0.0362	0.0740
15	0.0337	0.0689
16	0.0316	0.0645
17	0.0297	0.0606
18	0.0280	0.0571
19	0.0265	0.0540
20	0.0252	0.0512
21	0.0240	0.0487
22	0.0228	0.0465
23	0.0219	0.0444
24	0.0210	0.0425
25	0.0201	0.0408
26	0.0193	0.0392
27	0.0186	0.0377
28	0.0179	0.0363
29	0.0173	0.0351
30	0.0167	0.0339
31	0.0162	0.0327
32	0.0156	0.0317
33	0.0151	0.0307
34	0.0147	0.0298
35	0.0143	0.0289
36	0.0138	0.0281
37	0.0134	0.0273
38	0.0131	0.0266
39	0.0128	0.0259
40	0.0125	0.0252

4.3.4.1 The Estimation of the Extremal Parameters a_m and b_m

If the initial distribution and its parameters and the size of the sample from which the largest observations are taken are known, then the parameters a_m and b_m may be obtained from equations (24) and (25) respectively. Gumbel has applied this method to the exponential, logistic, normal and double exponential distributions [Gumbel, E.J., 1958, pp. 113, 126, 129 and 201] [and Gumbel, E.J., 1954, p.35].

In practical applications, the initial distribution is usually unknown and even the number n may be unknown, but if we know that the initial distribution is of the exponential type, we may estimate the parameters b and a.

There are different methods suggested by Gumbel and by Kimball, B.F. [Gumbell, E.J., 1958, pp.34, 229-235] [and Kimball, B.F., 1946, p.299].

4.3.4.2 Estimation of the Parameters When the Initial Distribution is Unknown

Suppose that there are N samples each of size n of variate x_{mj}, $m = 1, \ldots, n$ and $j = 1, \ldots, N$, then estimation of the parameters could be obtained from:

$$b_m = \bar{x}_m - \frac{\bar{y}_m}{a_m} \tag{52}$$

$$a_m = \frac{\sigma_m}{\sigma_{mx}} \tag{53}$$

where $\bar{x}_m$ and σ_{mx} are the mean and the standard error of x_{mj}. The mean $\bar{y}_m$ and the standard error of y_m are given by equations (46) and (48) and are given in Table 4 (1). These limiting or asymptotic values are true only in the case when the number of samples (N) are large [Ramachandran, G., 1974, p.293].

4.3.5 If the Initial Distribution is Expected to Change Over a Number of Intervals and Only the First r Extremes are Known:

Suppose that the initial distribution is expected to change over a number of intervals, in this case it is wise to use only a small number N of samples each of size n and suppose that only the first r largest values are known, therefore, suppose x_{mj} is the mth largest in the jth sample, where

$$m = 1, 2, \ldots, r \ (r \leq n) \quad \text{and} \quad j = 1, 2, \ldots, N.$$

Since the initial distribution is known to be of the exponential type, therefore the parameters a_m and b_m may be estimated from the N observed mth largest values and the following method is suggested [Ramachandran, G., 1972, p.3].

We arrange each set of the r sets of x_{mj} in an increasing order of magnitude and let R_{mj} be the rank of x_{mj}. The empirical value of the cumulative relative frequency of x_{mj} is

$$\Phi_m (x_{mj}) = \frac{R_{mj}}{N+1} \tag{54}$$

$$R_{mj} = 1, 2, 3, \ldots, N$$

Since cumulative frequency are preserved under transformation

$$\Phi_m(x_{mj}) = \Phi_m(y_{mj}) = \Phi_m(U_{mj})$$

$$= \int_{z_{mj}}^{\infty} z^{m-1} e^{-z} dz/\Gamma_{(m)} \tag{55}$$

$$\text{where} \quad z_{mj} = m e^{-y_{mj}} \tag{56}$$

The values of $z_{mj}(m = 1,2,3,\ldots,r;\ j = 1,2,3,\ldots,N)$ corresponding to the cumulative frequencies (54) may be calculated from tables of incomplete gamma function, then using (56) the values of y_{mj} corresponding to the values of x_{mj} are obtained and from the equation:

$$x_{mj} = b_m + \frac{y_{mj}}{a_m} \tag{57}$$

and by the least square method, the parameters a_m and b_m , $m =$ 1, 2, 3, ..., r may be obtained.

Instead of using the incomplete gamma function to obtain the values of y_{mj} by using the equations (54), (55) and (56), it could be obtained directly from equation (54) by using Table 4 of "Probability Tables for the Analysis of Extreme-Value Data". To interpolate between the tabulated values of Φ, we use the Gregory-Newton interpolation formula:

$$f(x_0+ph) = f(x_0) + p\Delta f(x_0) + \frac{p(p-1)}{2}\Delta^2 f(x_0) +$$

$$\frac{p(p-1))p-2)}{6}\Delta^3 f(x_0) + \frac{p(p-1)(p-2)(p-3)}{24}\Delta^4 f(x_0) + \dots \tag{58}$$

where $f(x)$ is the tabulated value at intervals of $h = x_1 - x_0$ and $x = x_0 + ph$.

Since the intervals of Table 4 are not equal, therefore Δ^2, Δ^3, ... could not be derived from the table, therefore approximate values for different values of p are given on p.12 of the "Probability Tables for the Analysis of Extreme-Value Data". This method may be used when the tables of the incomplete gamma function are not available, otherwise the wide spacing of probability and its variation makes the interpolation a difficult task.

4.3.6 **Variation in Sample Size:**

In the previous study of the extreme value theory we assumed that the N samples are of equal size, but this assumption is not satisfied in many fields of applications, especially in insurance claim statistics where the number of claims in each year are not equal. "This variation would be expected to affect the values of extreme value parameters a_m and b_m as they are functions of n." [Ramachandran, G., 1972, p.3.]; Jung, J., has suggested that if the n_i values differ but represent "equally exposed intervals", then under certain conditions the method described for constant n may be used [Jung, J., p.178]. Ramachandran has suggested a modification to the y_{mj}(*) to be:

(*)In fact, Ramachandran (1972, p.36) has proved that if b_{mnj} be the characteristic mth large value in sample of size n_j, then

$$b_{mn_j} = b_{mn} + \frac{1}{a_{mn}} \log \frac{n_j}{n}$$

substituting in (57)

$$x_{mj} = b_{mn} + \frac{1}{a_{mn}} \log \left(\frac{n_j}{n}\right) + \frac{y_{mj}}{a_{mn}}$$

$$= b_{mn} + \frac{y'_{mj}}{a_{mn}} \text{ where } y'_{mj} = y_{mj} + \log \left(\frac{n_j}{n}\right)$$

$$y'_{mj} = y_{mj} + \log \frac{n_i}{n} \qquad (59)$$

$$j = 1,2,3,\ldots,N$$

where n_j is the size of the jth sample and n could be the sample size of the base period or the average sample size. As in the case of constant n, the parameters may be estimated by fitting the line:

$$x_{mj} = b_{mn} + \frac{y'_{mj}}{a_{mn}} \qquad (60)$$

or from (52) and (53) after modification to be:

$$b_{mn} = \bar{x}_{mj} - \frac{\bar{y}_m + \bar{p}_i}{a_{mn}} \qquad (61)$$

$$a^2_{mn} = \frac{\sigma^2_m + \sigma^2_p + 2\sigma_{yp}}{\sigma^2_{mx}} \qquad (62)$$

where $\bar{p}_j$ and σ^2_p are the mean and variance of

$p_j = \log \frac{n_j}{n}$, σ_{yp} is the covariance of y_{mj} and p_j

and could be included in the calculation though it is theoretically equal to zero [Ramachandran, E.J, 1974, p.293]. Moveover, the parameters may be estimated from (52) and (53) by taking $\bar{y}_m$ and σ_m as the mean and standard error of the corrected y. Leonard D. Coe has found that "satisfactory results for N not greater than 5 can be obtained by taking the mean of the x variates and associating with each the mean of the corresponding frequencies" [Leonard D. Coe, p.12].

4.3.7 **A Special Case in Insurance**

Sometimes the total number of claims in each year may not be available, especially to the reinsurers owing to the loss of records or for underwriting causes. In such cases the author suggests that the modification of y_{mj} could be by adding the logarithms of the ratio of the earned premium $\frac{P_j}{P}$ to be

$$y'_{mj} = y_{mj} + \log_e \frac{P_j}{P} \tag{63}$$

where P_j is the earned premium in year j and P is the earned premium in the base year and the earned premium is defined as:

Earned Premium = Premiums written (current year) PLUS premium reserve (Previous year) MINUS Premium reserve (current year).

This modification is preferable because the earned premium could be produced from the published statistics and the following conditions will be satisfied:

(1) The effect of inflation should be removed.

(2) There are no major changes in the underwriting policy which can affect th claim frequency distribution, especially the nil claims.

(3) There are no major changes in the premium rate which may not be removed by (1).

(4) The size of the portfolio is large enough to ensure that the number of claims resulted is large enough to apply the theory.

4.4.1 **The Return Period: The Theory:**

If we consider a discontinuous variate, the die, say, where the probability that a certain face to occur is $\frac{1}{6}$, then if we repeat casting the die many times we expect to get this face on the average once in each 6 trials, i.e. the return period of this face is 6, and if we denote the return period by T, then $T = 6$.

For a continuous variate there is no probability density function, therefore if we define $F(x)$ as the probability that a value is less than x, and let:

$$q = 1 - p = F(x)$$

so that

$$1 - F(x) = p = 1 - q \tag{64}$$

and this is the probability that the value is equal to or exceeds x. If the observations are made at regular intervals of time and the experiment stops when the value x has been exceeded once, and if the event happens at trial v and fails to happen in the first $v - 1$ trials, $v \geq 1$, then the probability that the exceedance happens for the first time at trial v is given by:

$$w(v) = pq^{v-1} \quad ; \quad v \geq 1 \tag{65}$$

which is a geometric distribution.

The moment generating function (m.g.f.) is:

$$G_v(t) = \sum_{v=1}^{\infty} e^{tv} pq^{v-1}$$

$$= pe^t \left[1 + (qe^t) + (qe^t)^2 + \quad \ldots\ldots \right]$$

$$= pe^t(1-qe^t)^{-1}$$

$$G_v(t) = \frac{p}{e^{-t}-q} \tag{66}$$

The mean value of $\bar{v}$ is the return period $T(x)$ and is defined as:

$$\bar{v} = T(x) = G_v'(0) \tag{67}$$

From (66) $G'_v(t) = -p(\bar{e}^t - q)^{-2} x - \bar{e}^t$

$$= p\bar{e}^t(\bar{e}^t - q)^{-2} = \frac{p\bar{e}^t}{(\bar{e}^t - q)^2}$$

$$G'_v(0) = \frac{p}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p} \quad \ldots \tag{68}$$

Therefore, $T(x) = \frac{1}{p}$; and from (67), (68) and (64) we get:

$$T(x) = \frac{1}{p} = \frac{1}{1-F(x)} > 1 \tag{69}$$

From equation (69) we deduce that if an event has a probability p, then its return period is the reciprocal of its probability, i.e. on the average we have to make $\frac{1}{p}$ trials in order that the event happens again.

The standard deviation of v may be defined as:

$$\sigma = \sqrt{G''_v(0) - \{G'_v(0)\}^2} \tag{70}$$

Since $G''_v(t) = - p\bar{e}^t(e^{-t} - q)^{-2} +$

$$p\bar{e}^t x - 2(\bar{e}^t - q)^{-3} \; x - \bar{e}^t$$

$$G''_v(0) = - \frac{p}{p^2} + \frac{2p}{p^3} = \frac{2}{p^2} - \frac{1}{p} \tag{71}$$

From (70), (68) and (71)

$$\sigma = \sqrt{\frac{1}{p^2} - \frac{1}{p}} = \sqrt{T^2 - T} \tag{72}$$

If p is very small then T becomes very large and (72) may be approximated to:

$$\sigma = T - \frac{1}{2} \tag{73}$$

4.4.2 The Return Period of the mth Extreme:

If the mth large observation follows each other in time and if the sample size is constant and if the distance between consecutive mth large values is approximately constant, then from equations (38) and (69) the return period of the mth large value is defined as:

$$T_m(y_m) = \frac{1}{1 - \Phi_m(y_m)} \tag{74}$$

From equation (74): if we are given the reduced values y_m then the corresponding cumulative probability Φ_m could be derived from the incomplete gamma function or from the Probability Tables for the Analysis of Extreme Value Data, and by substituting in (74) the average number of years necessary to obtain an mth value larger than y_m may be obtained where T is measured from the base year. Conversely, if we are given T the probaility that a value will occur larger than y_m could be obtained from:

$$\Phi_m(y_m) = 1 - \frac{1}{T} \tag{75}$$

In this concept we have assumed that the sample sizes are equal although in many applications of the theory this assumption is not fulfilled. Therefore, Rachamandran has suggested that y_m should be modified to match the expected change in the sample size, he assumes that there is a geometric increase in the sample size and the rate of increase could be defined from the given samples

$$n_N = n_1(1+h)^{N-1} \tag{76}$$

[Ramachandran, G., 1972, pp.8,9.]

From (76) we define the annual increase h in the size of the samples, therefore the expected sample size could be defined for any coming year, i.e. for year $N + k$ will be

$$n_{N+k} = n_1(1+h)^{n+k-1} \tag{77}$$

By substituting from (77) into (59) the modified $y_{m,(N+k)}$ may be obtained and the same procedure, as mentioned above, can be followed.

In insurance buseiness, if we use the earned premium instead of the number of claims, then h will be the annual rate of expansion of premium.

4.5.1 The Variance of the Estimation of the Extreme Values

In (4.3.4), (4.3.5) and (4.3.6) we presented the methods of estimating the parameters a_m and b_m from a sample and we substituted these values in equation (60) in order to find an estimation to x_m. These estimations have their variances, and these variances depend on m and on the method by which a_m and b_m are estimated. Gumbel has presented formulae for the standard error of x in the neighbourhood of the median [Gumbel, E.J., 1954, pp.17,18].

Using the maximum likelihood estimation functions for estimating the parameters a and b in order to find an estimation to x, corresponding to a given probability Φ, Kimball has shown that the variance of x for finite N may be approximated by:

$$\sigma_x^2 = \frac{1}{Na^2} \left[1+(1-\gamma+y)^2/(\pi^2/6)\right] \tag{78}$$

where N is the number of samples used in estimating a and b, γ is Euler's constant and y is corresponding to given probability Φ. The formula takes into account that a is unknown and what is called the "marginal distribution" of $\sqrt{N}(\hat{x}-x)$, the difference between the expected and observed, will be asymptotically normal [Kimball, B.F., 1949, p.110].

"σ_x^2 represents the sampling variance of the x coordinate of the best fitting distribution under the assumption that the universe from which the sample is drawn is truly a double exponential distribution of maximum values.

It is the type of variance whose square root is what has been known as "the standard error of forecast" [Hotelling]"
[Gumbel, E.J., 1958, p.235].

4.5.2 **The Variance of the Estimation of the Extreme**

Values from: $x_m = b_m + \dfrac{y_m + \log\frac{n_j}{n}}{a_m}$:

If we denote $\log \dfrac{n_j}{n}$ by p_j ;

then $$x_m = b_m + \frac{y_m + p_j}{a_m}$$

the variance of x_m is given by:

$$\sigma x_m^2 = \frac{1}{(a_m)^2}\left[\sigma y_m^2 + \sigma p^2 + 2\ \mathrm{Cov.}(y_m,\ p)\right] \tag{79}$$

where $\sigma^2_{y_m}$ is given from Table 4 (1) if N is large, but if N is small an estimation to the variance of y_m from the sample may be used,

$$\sigma^2_p = \frac{1}{N-1}\left[\sum_{j=1}^{N} p_j^{\,2} - \frac{\left(\sum_{j=1}^{N} p_j\right)^2}{N}\right] \qquad (80)$$

and

$$\mathrm{Cov}(y_m, p) = \frac{1}{N-1}\left[\sum_{j=1}^{N} y_{m_j} p_j - N\,\bar{y}_m \bar{p}\right] \qquad (81)$$

$\mathrm{Cov}(y_m, p)$ is theoreticaly equal to zero (see 4.3.6), therefore

$$\sigma^2_{x_m} = \frac{1}{(a_m)^2}\left[\sigma^2_{y_m} + \sigma^2_p\right] \qquad (82)$$

If the sample sizes are equal then

$$\sigma^2_{xm} = \frac{\sigma^2_{ym}}{\sigma^2_m} \qquad (83)$$

By using the variance given by equation (82) a control curve may be produced.

4.5.3 The Variance of the Estimated Value of x_m for A given y_m

Since the straight line (60) has been fitted by the least square method, therefore there is an error which equals the difference between the actual and expected values of x_m

$$\text{i.e.} \quad \varepsilon_{mj} = x_{mj} - \hat{x}_{mj} \tag{84}$$

where

$$E\varepsilon_m = 0 \quad \text{and} \quad E\varepsilon_m^2 = \frac{1}{N-2} \sum_j (x_{mj}-\hat{x}_{mj})^2 = S^2 \tag{85}$$

where the variance of the estimated value of x_m for a given y_m has a variance defined as:

$$S_{xm}^2 = S^2\left[1 + \frac{1}{N} + \frac{(y_m-\bar{y}_m)^2}{\sum(y_{mj}-\bar{y}_m)^2}\right] \tag{86}$$

The author suggests using (86) to place a confidence limit on the estimated value of x_m.

4.6 The Effect of Using a Different Distribution

In our study we said nothing about the distribution function of the parent distribution except that it should be of the exponential family, and since at the large extreme end of the distribution the curve appears to be straight, which makes the tails of the different curves of one family nearly coincident and the difference between the values of the ordinate produced from each distribution is practically negligible. Coe (p.19) has tried to investigate the difference between the estimated largest values by considering different distributions and he has found that there is no difference.

4.7 The Application of Extreme Value Theory to Estimate the Probability Distribution Function of the Parent Distribution

If the top n large values of a probability distribution are known, Rogers (1976) has suggested a model based on the theory of extreme values to estimate the parameters of the distribution. Moreover, Rogers (1977) has used the same procedure to estimate the parameters of the normal distribution.

Ramachandran and Taylor (1975) have developed a mathematical model for estimating the total losses in all fires by using only large losses.

4.8 The calculation of an excess of loss premium

The excess of loss premium has, for a long time, been the focus of much research and many papers have been published on this subject. The most recent papers are concerned with the use of extreme value theory to calculate the most reasonable premium. Different approaches have been made which will be discussed in the following paragraphs:

4.8.1 If the parent distribution and/or the total number of claims are known:

Suppose that the claims experienced by a portfolio could be regarded as independent random variables with a distribution function $F(x)$ and a probability density function

$$f(x) = \frac{d}{dx} F(x).$$

The net premium per claim for an excess loss cover above an amount L is given by:

$$P(L) = \int_{L}^{\infty} (x-L)f(x)dx \qquad (87)$$

This expression can also be written in the form:

$$P(L) = \int_{L}^{\infty} [1-F(x)]dx \qquad (88)$$

If $F(x)$ is not given but the total number of claims in a number of successive years and the first r largest claims in these years are given, then the parameters of the parent distribution may be estimated (Rogers, 1976 and 1977), and Ramachandran 1974 and 1982 - p.59, and the excess of loss premium could be calculated (Jung, p.178).

Ramachandran, 1982, p.59, has introduced a formula to calculate the approximate net premiums per claim for an excess of loss cover above L based on the extreme value parameters a_m abd b_m and the total number of claims n in the base year,

$$P(L) = \frac{m\ e^{L-y_L}}{n(a_m-1)} \qquad (89)$$

where m is the order of the extreme, a_m and b_m are the mth extreme value parameters, n is the total number of claims in the base year and

$$y_L = a_m(L-b_m) \qquad (90)$$

Equation (89) could be written in the form:

$$P(L) = \frac{m}{n} \cdot \frac{e^{-L(a_m-1)+a_m b_m}}{(a_m-1)} \qquad (91)$$

4.8.2 **If the parent distribution and the total number of claims are unknown:**

If the parent distribution is not known but can be assumed to be of the exponential type and the first r largest claims $x_i(i=1,2,\ldots,r)$ in each of N successive years are known, then the values of the parameters a_m and b_m for each value of $m = 1, \ldots, r$ may be calculated by one of the methods discussed earlier and Ramachandran (1974), p.293, has shown that the average claim $\bar{x}$ is given by:

$$x_m = \frac{e^{b_m} m^{1/a_m} \Gamma\left(m - \frac{1}{a_m}\right)}{\Gamma(m)} \qquad (92)$$

where x_i are considered to be above L and therefore the expected value of the net premium is given by:

$$\bar{P}(L) = \sum_m \bar{x}_m - rL \qquad (93)$$

Beard, 1963b, p.6, has introduced a formula to calculate the excess of loss premium in such cases as:

$$P(L) = \frac{1}{a_m} e^{-a_m(L-b_m)} \qquad (94)$$

He compared the premiums produced by his formula with the true premiums and found that the estimated premiums in all cases were greater than the true values and a reasonable close approximation is given if L is not too far removed from b, particularly if the number of claims is large.

The author has applied Beard's formula (equation 94) to data produced by Ramachandran (Ramachandran, 1982, p.59, Table 6) and compared the premiums produced by Beard's formula with those produced by Ramachandran's formula (equation 89) and found that the premiums produced by the first is greater than its counterpart produced by the second for the first four extremes, and less than their counterpart from the 5th to the 10th as shown in Table 4.2.

Table 4.2

The comparison between the estimated excess of loss premiums produced by equations 89 and 93 for different retentions

Extremes (m)	a_m	b_m	L = 3		L = 4		L = 5	
			Eq.89	Eq.94	Eq.89	Eq.94	Eq.89	Eq.94
1	2.247	5.214	5.013	64.410	1.441	6.809	0.416	0.720
2	1.785	4.829	2.880	14.663	1.316	2.460	0.594	0.413
3	1.626	4.534	2.507	7.450	1.343	1.465	0.712	0.288
4	1.460	4.327	2.607	4.753	1.643	1.104	1.039	0.256
5	1.387	4.113	2.613	3.376	1.774	0.843	1.202	0.211
6	1.424	3.988	2.497	2.867	1.632	0.690	1.069	0.166
7	1.239	3.749	3.200	2.042	2.522	0.591	1.989	0.171
8	1.163	3.564	4.085	1.657	3.478	0.518	2.953	0.162
9	1.212	3.448	3.156	1.420	2.555	0.423	2.063	0.126
10	1.034	3.259	16.607	1.264	16.051	0.449	15.509	0.160

where $n = 465$.

4.8.3 The direct method

As mentioned in (4.2.5) the net premium of excess of loss cover above retention X is given by:

$$P(X) = H(X) \cdot M(X) \qquad (95)$$

where $H(X)$ = the expected number of claims that exceed the retention X.

$M(X)$ = the average excess claim over the retention X.

If the r largest claims $x_i (i = 1,2,\ldots,r)$ could be estimated by the extreme value theory, then a method analogous to that used by Benjamin p.220, may be used to establish a formula to give the net premium of the excess of loss cover in excess of retention L.

4.9 Conclusion

Since there are different methods suggested to estimate the extremal parameters a_m and b_m, e.g. the least square method, the maximum likelihood method, ..., the values of a_m and b_m estimated by these methods are expected to be different; therefore "practical results based on extreme value theory would involve errors due to the particular method adopted for estimating the parameters" [Ramachandran, 1974, p.293].

CHAPTER FIVE

THE DATA

The data used in this thesis are collected from the Egyptian market (the source of the data is the Egyptian Reinsurance Company) and it relates to the ten largest claims in the third party motor insurance for all types of uses of vehicles collectively, which occurred in the period 1970-1976. Table 5(1) shows the amount of these claims after six years of their occurrence, e.g. the claims of 1970 as at 31.12.1976, the claims of 1971 as at 31.12.1977, ... etc.

Table 5(1)

The 10 largest claims occurred in the period 1970-1976

(in E.L. 000s)

Year / Order of claim	1970	1971	1972	1973	1974	1975	1976
1	28	44	472	66	86	49	151
2	22	30	55	44	50	41	56
3	20	23	31	40	43	39	47
4	19	22	23	39	27	39	46
5	17	21	20	32	22	33	40
6	14	20	17	28	21	31	37
7	14	20	17	26	21	31	35
8	13	17	16	22	20	30	20
9	12	10	14	21	17	27	18
10	12	7	13	21	16	20	15

N.B. E.L. = Egyptian Pounds.

Before 1970 it is difficult to obtain reliable data, while the data after 1976 will not include much of the incurred but not reported claims (IBNR). If the IBNR claims are not included in the data the results will vary in accuracy. By investigating a specimen of the data the author found that within six years of any year more than 90% of the claims which occurred in that year are reported.

The outstanding amount of any claim after six years of occurrence will be, on average, very small compared with the total amount of the claim and the estimation becomes more accurate.

Table 5 (2) shows the development of the earned premium in the period 1970-1976 which will be used as a corrective factor.

Table 5 (2)

The earned premium in the period 1970-1976

(in E.L. 000s)

Year	Earned Premium
1970	1145
1971	1368
1972	1593
1973	1682
1974	1772
1975	2079
1976	2457

CHAPTER SIX

THE APPLICATION OF THE THEORETICAL BASIS ON THE EGYPTIAN MARKET

6.1 The calculation of x_{mj}:

Table 5 (1) expresses the claims in a different values of pound. To remove the effect of inflation, the Retail Price Index of Egypt given in Table 6 (1) is applied to express the seven years claims in 1970 prices.

Since the claims are not settled at once but over a number of years, which in this study is about six years, it has been assumed that on average the claims are settled by one lump sum after three years of occurrence. Thus the amount of claims paid for an accident of a particular year could be considered as expressed in terms of the value of the pound three years after the year of its occurrence, e.g. the amount of 1970 claims are expressed in 1973 prices, the amount of 1971 claims are expressed in 1974 prices, and so on.

Table 6 (1)

The Retail Price Index of Egypt

1970 = 100

Year	R.P.I.
1970	100.0
1971	102.0
1972	102.1
1973	110.2
1974	126.0
1975	135.5
1976	146.1
1977	159.7
1978	183.2
1979	201.2

N.B. The figures of Table 6 (1) are taken from "Measurement of the trend of inflation in Egypt" published by Central Agency for Public Mobilization and Statistics, September 1981 (in Arabic) and recalculated with reference to 1970 as the basis year.

The corrected claims amount are given in Table 6 (2).

Table 6 (2)

The corrected claims expressed in 1970 prices (z_{mj})

(in E.L. 000s)

Extreme (m) \ Year (j)	1970	1971	1972	1973	1974	1975	1976
1	25	35	348	45	54	27	75
2	20	24	41	30	31	22	28
3	18	18	23	27	27	21	23
4	17	17	17	27	17	21	23
5	15	17	15	22	14	18	20
6	13	16	13	19	13	17	18
7	13	16	13	18	13	17	17
8	12	13	12	15	13	16	10
9	11	8	10	14	11	15	9
10	11	6	10	14	10	11	7

Since there is a belief that the natural logarithm of the claim amounts are of the exponential type, therefore the logarithms are taken and shown in Table 6. (3).

Table 6 (3)

The natural logarithm of the extremes: $x_{mj} = \log_e z_{mj}$

Extreme (m) \ Year (j)	1970	1971	1972	1973	1974	1975	1976
1	3.2189	3.5553	5.8522	3.8067	3.9890	3.2958	4.3174
2	2.9957	3.1781	3.7136	3.4012	3.4340	3.0910	3.3322
3	2.8904	2.8904	3.1355	3.2958	3.2958	3.0445	3.1355
4	2.8332	2.8332	2.8332	3.2958	2.8332	3.0445	3.1355
5	2.7081	2.8332	2.7081	3.0910	2.6391	2.8904	2.9957
6	2.5649	2.7726	2.5649	2.9444	2.5649	2.8332	2.8904
7	2.5649	2.7726	2.5649	2.8904	2.5649	2.8332	2.8332
8	2.4849	2.5649	2.4849	2.7081	2.5649	2.7726	2.3026
9	2.3979	2.0794	2.3026	2.6391	2.3979	2.7081	2.1972
10	2.3979	1.7918	2.3026	2.6391	2.3026	2.3979	1.9459

From Table 6. (3) for each $m = 1,2,\ldots,10$ we arrange x_{mj} in an increasing order and give it a rank R_{mj}, these ranks are shown in Table 6. (4).

Table 6. (4)

Ranks or Extremes (R_{mj})

Extreme (m) \ Year (j)	1970	1971	1972	1973	1974	1975	1976
1	1	3	7	4	5	2	6
2	1	3	7	5	6	2	4
3	1	2	4	6	7	3	5
4	1	2	3	7	4	5	6
5	2	4	3	7	1	5	6
6	1	4	2	7	3	5	6
7	1	4	2	7	3	5	6
8	2	4	3	6	5	7	1
9	4	1	3	6	5	7	2
10	5	1	3	7	4	6	2

6.2 The Calculation of the reduced extremes y_{mj}

From equation 4. (54):

$$\Phi_m(x_{mj}) = \int_{z_{mj}}^{\infty} z^{m-1} e^{-z} dz/\Gamma(m)$$

$$= 1 - \int_0^{z_{mj}} z^{m-1} e^{-z}\, dz/\Gamma(m)$$

$$\therefore \quad 1 - \Phi_m = \int_0^{z_{mj}} z^{m-1} e^{-z} dz/\Gamma(m) \tag{1}$$

The R.H.S. of this equation are known, from the tables of the incomplete gamma function, to be $I(u,p-1)$.

$$\therefore \quad 1 - \Phi_m = I(u,p) \tag{2}$$

where p is corresponding to $m-1$, i.e. $p+1 = m$ and

$$u = \frac{z_{mj}}{\sqrt{p+1}} = \frac{z_{mj}}{\sqrt{m}} \tag{3}$$

By interpolating in the tables of the incomplete gamma function for each integral value of $p = 0,1,2, \ldots, 9$ the values of u corresponding to $m = 1,2,3, \ldots, 10$ respectively could be produced.

From equations 6. (3) and 4. (56) we have:

$$u = \frac{z_{mj}}{\sqrt{m}} = \sqrt{m} \ e^{-y_{mj}}$$

$$\therefore e^{-y_{mj}} = \frac{u}{\sqrt{m}}$$

$$\therefore y_{mj} = -\log \frac{u}{\sqrt{m}}$$

The values of y_{mj}, $m = 1,2,\ldots, 10$ and $j = 1,2,\ldots,7$ are given in Table 6. (5)

Table 6.5
Reduced Extremes (y_{mj}) - uncorrected

Rank (j)	Cumulative frequency Φ=j/N+1	Extremes (m)				
		1	2	3	4	5
1	0.125	-0.7325	-0.5902	-0.5102	-0.4575	-0.4186
2	0.250	-0.3271	-0.2975	-0.2680	-0.2452	-0.2271
3	0.375	0.0186	-0.0574	-0.0723	-0.0754	-0.0751
4	0.500	0.3660	0.1751	0.1147	0.0853	0.0680
5	0.625	0.7529	0.4265	0.3139	0.2551	0.2183
6	0.750	1.2441	0.7328	0.5521	0.4563	0.3950
7	0.875	2.0047	1.1905	0.8996	0.7469	0.6485

Rank (j)	Cumulative frequency Φ=j/N+1	Extremes (m)				
		6	7	8	9	10
1	0.125	-0.3890	-0.3651	-0.3452	-0.3287	-0.3142
2	0.250	-0.2131	-0.2013	-0.1913	-0.1825	-0.1753
3	0.375	-0.0740	-0.0723	-0.0704	-0.0686	-0.0669
4	0.500	0.0565	0.0483	0.0422	0.0375	0.0336
5	0.625	0.1927	0.1738	0.1591	0.1473	0.1377
6	0.750	0.3523	0.3203	0.2951	0.2750	0.2583
7	0.875	0.5794	0.5278	0.4872	0.4539	0.4272

Assuming that the earned premiums given in Table 5 (2) are satisfying the conditions of 4.3.7, therefore the correction factors are given in Table 6.(6).

Table 6.(6)
The correction factor

Year	p_j in E.L. 000s	Correction factor $\log p_j/p_1$
1970	1145	0.0000
1971	1317	0.1397
1972	1530	0.2898
1973	1497	0.2677
1974	1380	0.1865
1975	1505	0.2734
1976	1650	0.3653

N.B. p_j are the earned premiums of Table 5(2) after removing the inflation effect

From Tables 6.(4) , 6.(5) and 6.(6) and by applying formula 4.(63) which states that

$$y_{mj} \text{ (corrected)} = y_{mj} + \log \frac{p_j}{p_1}$$

the corrected values of y_{mj} may be produced and are given in Table 6.(7) (for example, from Table 6.(4) the order of the first extreme claim in year 1973 is 4 and from Table 6.(5) the corresponding value of $y_{1,4}$ is 0.3660, from Table 6.(6) the correction factor is 0.2677, therefore, by applying formula 4 (63) the corrected value of $y_{1,4}$ is 0.6337, and so on).

Table 6.(7)

The corrected values of y_{mj}

Extreme (m) \ Year	1970	1971	1972	1973	1974	1975	1976
1	-0.7325	0.1583	2.2945	0.6337	0.9394	-0.0537	1.6094
2	-0.5902	0.0823	1.4803	0.6942	0.9193	-0.0241	0.5404
3	-0.5102	-0.1283	0.4045	0.8198	1.0861	0.2011	0.6792
4	-0.4575	-0.1055	0.2144	1.0146	0.2718	0.5285	0.8216
5	-0.2271	0.2077	0.2147	0.9162	-0.2321	0.4917	0.7603
6	-0.3890	0.1962	0.0767	0.8471	0.1125	0.4661	0.7176
7	-0.3651	0.1880	0.0885	0.7955	0.1142	0.4472	0.6856
8	-0.1913	0.1819	0.2194	0.5628	0.3456	0.7606	0.0201
9	0.0375	-0.1890	0.2212	0.5427	0.3338	0.7273	0.1828
10	0.1377	-0.1745	0.2229	0.6949	0.2201	0.5317	0.1900

By combining Table 6.(3)with Table 6.(7) Table 6.(8) is produced.

Table 6.(8)

Extreme (m)	Variables	1970	1971	1972	1973	1974	1975	1976	Mean Value	Variance
1	x_1	3.2189	3.5553	5.8522	3.8067	3.9890	3.2958	4.3174	4.0050	0.7045
	y_1	-0.7325	0.1583	2.2945	0.6337	0.9394	-0.0537	1.6094	0.6927	1.0574
2	x_2	2.9957	3.1781	3.7136	3.4012	3.4340	3.0910	3.3322	3.3065	0.0588
	y_2	-0.5902	0.0823	1.4803	0.6942	0.9193	-0.0241	0.5404	0.4432	0.4652
3	x_3	2.8904	2.8904	3.1355	3.2958	3.2958	3.0445	3.1355	3.0983	0.0272
	y_3	-05102	-0.1283	0.4045	0.8198	1.0861	0.2011	0.6792	0.3646	0.3105
4	x_4	2.8332	2.8332	2.8332	3.2958	2.8332	3.0445	3.1355	2.9726	0.0291
	y_4	-0.4575	-0.1055	0.2144	1.0146	0.2718	0.5285	0.8216	0.3268	0.2627
5	x_5	2.7081	2.8332	2.7081	3.0910	2.6391	2.8904	2.9957	2.8379	0.0267
	y_5	-0.2271	0.2077	0.2147	0.9162	-0.2321	0.4917	0.7603	0.3045	0.2008

Continued/...

Table 6.(8) (Continued)

Extreme (m)	Variables	1970	1971	1972	1973	1974	1975	1976	Mean Value	Variance
6	x_6	2.5649	2.7726	2.5649	2.9444	2.5649	2.8332	2.8904	2.7336	0.0242
	y_6	-0.3890	0.1962	0.0767	0.8471	0.1125	0.4661	0.7176	0.2896	0.1785
7	x_7	2.5649	2.7726	2.5649	2.8904	2.5649	2.8332	2.8332	2.7177	0.0177
	y_7	-0.3651	0.1880	0.0885	0.7955	0.1142	0.4472	0.6856	0.2791	0.1578
8	x_8	2.4849	2.5649	2.4849	2.7081	2.5649	2.7726	2.3026	2.5547	0.0193
	y_8	-0.1913	0.1819	0.2194	0.5628	0.3456	0.7606	0.0201	0.2713	0.1029
9	x_9	2.3979	2.0794	2.3026	2.6391	2.3979	2.7081	2.1972	2.3889	0.0483
	y_9	0.0375	-0.1890	0.2212	0.5427	0.3338	0.7273	0.1828	0.2652	0.0937
10	x_{10}	2.3979	1.7918	2.3026	2.6391	2.3026	2.3979	1.9459	2.2540	0.0691
	y_{10}	0.1377	-0.1745	0.2229	0.6949	0.2201	0.5317	0.1900	0.2604	0.0791

In order to fit equation 4.(57) from the data of Table 6.(8) the least square method is used for each value of m and the values of the parameters a_m and b_m and the correlation coefficient r_m are given in Table 6.(9).

Table 6.(9)

Extreme (m)	a_m	b_m	r_m
1	1.2335	3.4434	0.9250
2	2.8560	3.1513	0.9872
3	3.4582	2.9929	0.9571
4	3.0882	2.8668	0.8790
5	2.8391	2.7307	0.9520
6	2.8266	2.6311	0.8989
7	3.1223	2.6283	0.8660
8	2.4669	2.4447	0.8380
9	1.4960	2.2116	0.9080
10	1.1641	2.0303	0.8344

Since the correlation coefficient r_m is very high for all values of m there is therefore good grounds for accepting that the claim amount distribution can be represented by an exponential function and that extreme value theory is applicable to the data.

From Table 6.(9) we note that the values of b_m are decreasing with the increase of m and it is self-evident that b_m is the "characteristic mth largest value".

To find the control curves of the estimation, equation 4.(82) has been applied to calculate the variance of the estimates of the extreme values, these variance values for each $m = 1,2,...10$ are given in Table 6.(10).

Table 6.(10)

Extreme (m)	Variance
1	0.8690
2	0.1679
3	0.0940
4	0.0898
5	0.0758
6	0.0683
7	0.0552
8	0.0476
9	0.0723
10	0.0804

Assuming that the differences between the actual values of x and the estimated values by the least square method are normally distributed with mean zero and variance given from Table 6.(10) and with 95% confidence limits, the control curves for each value of m are:

$m = 1 \qquad x_1 = 3.4434 + 0.8107y_1 \pm 1.96(0.9322)$

$m = 2 \qquad x_2 = 3.1514 + 0.3501y_2 \pm 1.96(0.4098)$

$m = 3 \qquad x_3 = 2.9929 + 0.2892y_3 \pm 1.96(0.3066)$

$$m = 4 \qquad x_4 = 2.8668 + 0.3238y_4 \pm 1.96(0.2997)$$

$$m = 5 \qquad x_5 = 2.7307 + 0.3522y_5 \pm 1.96(0.2753)$$

$$m = 6 \qquad x_6 = 2.6311 + 0.3538y_6 \pm 1.96(0.2613)$$

$$m = 7 \qquad x_7 = 2.6283 + 0.3203y_7 \pm 1.96(0.2349)$$

$$m = 8 \qquad x_8 = 2.4447 + 0.4054y_8 \pm 1.96(0.2182)$$

$$m = 9 \qquad x_9 = 2.2116 + 0.6684y_9 \pm 1.96(0.2689)$$

$$m = 10 \qquad x_{10} = 2.0303 + 0.8590y_{10} \pm 1.96(0.2835) \qquad (4)$$

It is more appropriate to use another method to calculate the values of the parameters a_m and b_m, equations 4.(52) and 4.(53) have been used where $\bar{y}_m$ and σ_m are the mean and standard error of the corrected values of y. Results compared with those produced by the least square method are given in Table 6.(11).

Table 6.(11)

Extreme (m)	Least square method		Equations 4.(52) and 4.(53)	
	a_m	b_m	a_m	b_m
1	1.2335	3.4434	1.2252	3.4396
2	2.8560	3.1513	2.8128	3.1489
3	3.4582	2.9929	3.3796	2.9904
4	3.0882	2.8668	3.0047	2.8638
5	2.8391	2.7307	2.7424	2.7269
6	2.8266	2.6311	2.7153	2.6269
7	3.1223	2.6283	2.9872	2.6243
8	2.4669	2.4447	2.3103	2.4373
9	1.4960	2.2116	1.3926	2.1985
10	1.1641	2.0303	1.0696	2.0105

From Table 6.(11) we note that the a's and the b's have the same trend and the difference between the a's produced by the two methods are of order 10^{-2} and the difference between the b's are of the order 10^{-3}. Owing to these differences "practical results based on extreme value theory would involve errors due to the particular method adopted for estimating the parameters" (Ramachandran, G., 1974, p.295).

6.3 **Five years' planning:**

To make a forecasting about the 10 largest claims which are expected to occur within the five years' period from 1976 to 1981 we take $T = 12$ in 4.(75), where T starts from the base year 1970 and $T = 1981 - 1970 + 1$, therefore the probability that y_m will not be exceeded is $\Phi_m(y_m) = 1 - \frac{1}{12} = \frac{11}{12} = 0.91667$.

From the incomplete gamma function and following the procedure of 6.2, the value of y_m for $m = 1,2,\ldots,10$ are produced and given in Table 6.(12) in column 2.

To correct these values of y_m, equation 4.(76) has been applied to the data in Table 6.(6) and h has been found to equal 0.06 therefore, by substituting the value of h and taking $n_1 = 1145$ in equation 4.(77), the expected earned premium in 1981 expressed in 1970 prices are equal to 2237 (in E.L. 000s). Therefore the correction factor is $\log 1.954 = 0.67$.

By applying equation 4.(59), the corrected values of y_m are produced and given in column 3 of Table 6.(12). By substituting these values in equations 6.(4) the model values of x_m are produced and given in column 4 of Table 6.(12) (the confidence terms are not included in this calculation). Column 5 of this table shows the expected amount of the 10 largest claims expressed in 1970 prices. That the probability that the observed claims during the period 1970 to 1981 to exceed these figures is 0.0833, or one in 12 years, i.e. once before 1982.

Table 6.(12)

Extreme (m)	y_m	y_m corrected	x_m	$z_m = e^{x_m}$ (in E.L.000s)
1	2.4353	3.1053	5.9609	389
2	1.4384	2.1084	3.8894	49
3	1.0842	1.7542	3.5002	33
4	0.8976	1.5676	3.3744	29
5	0.7796	1.4496	3.2412	26
6	0.6962	1.3662	3.1145	23
7	0.6332	1.3032	3.0457	21
8	0.5852	1.2552	2.9536	19
9	0.5460	1.2160	3.0243	21
10	0.5125	1.1825	3.0461	21

By comparing the figures in Table 6.(2) and the expected amount of the 10 largest claims in Table 6.(12) we note that during the course of the period up to 1976 none of the actual claims exceeded the expected claims in column 5 of Table 6.(12). Therefore, we expect that the excesses are likely to happen during the five year period 1976 to 1981. This estimation is based on the current trend of the claims and if it does not actually happen then it would appear that the claim trend in the Egyptian market has been improved.

The estimation of the 10 largest claims during this five-year period may serve as a guide for the allocation of the financial reserves to meet these claims as it falls due or to make an arrangement for the required protection against these claims by means of reinsurance covers.

Although a five-year planning period was introduced, a longer period may be considered by giving T values greater than 12, but a long planned period is not advisable since it may embrace drastic changes in the experience which makes the study unrealistic. Hence a five-year planning period to be recurring is preferable.

The confidence limits of the expected value of x_m could be obtained from equations 6.(4) where the variances are calculated by formula 4.(82). Alternative values have been produced from equation 4.(86) which was suggested by the author. The two values are tabulated alongside in Table 6.(13).

Table 6.(13)

Extreme (m)	The variance from equation 4.(82)	The variance from equation 4.(86)
1	0.8690	0.2898
2	0.1679	0.0043
3	0.0940	0.0063
4	0.0898	0.0208
5	0.0758	0.0069
6	0.0683	0.0142
7	0.0552	0.0146
8	0.0476	0.0233
9	0.0723	0.0294
10	0.0804	0.0897

Equation 4.(82) gives the variance of the model value of x_m while equation 4.(86) gives the variance for each estimated value of x_m and the first method gives a wider range of limits than the second.

6.4 The estimation of the excess of loss premium

From Chapter Five, the data available do not include the number of claims incurred in each year of the investigation period. Therefore the method discussed in 4.8.1 is not applicable in this study.

By substituting the values of a_m and b_m from Table 6.(9) into equation 4.(93), an estimation of the excess premium for different values of excesses obtained is shown in Table 6.(14).

Table 6.(14)

Excess L \ Extremes	m	1	2	3	4	5	6	7	8	9	10
	a_m	1.2335	2.8560	3.4582	3.0882	2.8391	2.8266	3.1223	2.4669	1.4960	1.1641
	b_m	3.4434	1.1513	2.9929	2.8668	2.7307	2.6311	2.6283	2.4447	2.2116	2.0303
2.0		4.8094	9.3810	8.9612	4.7080	2.8040	2.1060	2.2777	1.2142	0.9174	0.8899
2.5		2.5956	2.2495	1.5901	1.0052	0.6781	0.5125	0.4781	0.3537	0.4342	0.4972
3.0		1.4008	0.5394	0.2822	0.2146	0.1640	0.1247	0.1003	0.1030	0.2055	0.2778
3.5		0.7560	0.1293	0.0501	0.0458	0.0396	0.0303	0.0211	0.0300	0.0973	0.1552
4.0		0.4080	0.0310	0.0089	0.0098	0.0096	0.0074	0.0044	0.0087	0.0460	0.0867
4.5		0.2202	0.0074	0.0016	0.0021	0.0023	0.0018	0.0009	0.0025	0.0218	0.0485
5.0		0.1188	0.0018	0.0003	0.0004	0.0006	0.0004	0.0002	0.0007	0.0103	0.0271

By transforming these figures into the original currency units, the net premiums for different excess limits are given in Table 6.(15).

Table 6.(15)

The excess of loss premium in E.L. 000s

Excess point	Net premium
7	19932
12	40
20	15
33	12
55	11
90	10
148	10

The net premium in Table 6.(15) is the summation of the net premium of each of the 10 largest claims for each excess point.

The net premium for a cover in excess of EL 20 millions is unrealistic, while if we increased the excess limit to EL 12000, the net premium will be reduced to EL 40000, i.e. reduced to about 2% of the preceding premium. The decrease in premium became smaller with high excess limits which can indicate the dangers of using this formula to estimate the premium in cases of very low limits. Generally, the application of the formula needs more research for it is still in its infancy stage.

The direct method suggested in 4.8.3 has been applied by the author to the expected 10 largest claims in Column 5 of Table 6.(12).

Berkstander (Benjamin, p.219), has suggested the distribution

$$M(Z) = z^{1-b}/a$$

as the average excess claim over a retention Z which converges more quickly. The case when $b = 0.5$ is considered so that:

$$M(Z) = \frac{z^{\frac{1}{2}}}{a} \qquad (5)$$

H(Z) = the number od claims in excess of z

$$= c.a\ z^{-\frac{1}{2}}\ \exp(-2\ a\sqrt{z}) \qquad (6)$$

P(Z) = the expected amount of claims in excess of z = net premium

$$= c\ \exp(-2a\sqrt{z}) \qquad (7)$$

From column 5 in Table 6.(12), we produce Table 6.(16).

Table 6.(16)

Excess point z E.L.000s	Number of claims H(z)	Average excess claim M(z) E.L.000s	Total of claims greater than P(z) E.L.000s
20	9	48	432
25	5	80	400
30	3	127	381
35	2	184	368
50	1	339	339
100	1	289	289
130	1	259	259

By applying equations (5), (6) and (7) the values of a and e could be obtained and they are given in Table 6.(17).

Table 6.(17)

z	a	c
20	.093	999
25	.063	745
30	.043	612
35	.032	540
50	.021	453
100	.035	575
130	.044	707

From Table 6.(17), the average values of a and c are approximately 0.045 and 660 respectively and by substituting these values in (5), (6) and (7) we get:

$$M(z) = \frac{\sqrt{z}}{.045} = 22.2^{\circ} \sqrt{z} \tag{8}$$

$$H(z) = 29.7 \frac{e^{-.09\sqrt{z}}}{\sqrt{z}} \tag{9}$$

$$P(z) = 660\, e^{-.09\sqrt{z}} \tag{10}$$

Substituting the different values of z in equations (8), (9) and (10), the estimated number of claims in excess of z, the estimated mean claim in excess of z and the estimated total of claims greater than z, are obtained and they are shown alongside the actual values in Table 6.(18).

Table 6.(18)

z	H(z)		M(z)		P(z)	
	Actual	Estimated	Actual	Estimated	Actual	Estimated
20	9	4	48	99	432	441
25	5	4	80	111	400	421
30	3	3	127	122	381	403
35	2	3	184	131	368	387
50	1	2	339	157	339	349
100	1	1	289	222	289	268
130	1	1	259	253	259	236

From Table 6.(18), we notice that the estimated H at z = 20 is less than half the actual value, while the estimated value of M is approximately double the actual value.

The difference between the actual and the estimated values of both M and H decreases with the increase of z except at $z = 50$. On the other hand, the difference between the actual and the estimated values of the net premium is not very much for all values of z, moreover, the premiums estimated by this method look reasonable.

From the way in which the 10 largest claims are produced and their interpretation, the premiums produced by this method may represent the average premium over a five-year period.

The application of 4.(92) to calculate the excess of loss premium is restricted by the condition that x_i should be greater than the excess point L, therefore it cannot be easily used to test the effect of changing the excess point on the premium payable, owing to the fact that the mth large claims over a period of years could include some claims less than, as well as greater than L. The author has applied these equations to the first 3 largest claims and the values of $\bar{x}_m$, $m = 1,2,3$ are: 152, 27 and 21. The net premium for different excess points are shown in Table 6.(19).

Table 6.(19)

Excess point L E.L. 000s	Net premium E.L. 000s
20	140
25	129
30	122
35	117
50	102

By comparing the three methods and their results, it could be that the direct method is the most appropriate one as its results are more realistic than the others. Nevertheless, the subject is open to more discussion.

CHAPTER SEVEN

CONCLUSIONS

"large claims exercise a critical effect on the performance of an insurer whose top risks are not cut away by reinsurance. A reinsurer, on the other hand, is worried about the fluctuations in the portfolio of the large risks he accepts. Hence large losses on claims merit special investigation" (Ramachandran, 1974, p.293). The necessity for special investigation increases with the size of claim as in liability insurance where the size of claim is unlimited.

The extreme value theory is found to be very useful to explore the uncertainty of this part of the claim amount distribution. Perhaps it will not give a full answer to all questions but at least it will throw light on the behaviour of the claim distribution and reduce the uncertainty about the expected amount of the extreme claims in future. Thus it will help to allocate the required reserves from the years when claims are lower.

By applying the theory to the experience of motor insurance in the Egyptian market it has been found that the correlation coefficient in Table 6.(9) is very large and this strengthens the belief that the claim amount distribution is of the exponential type. Moreover, in Table 6.(12) the expected 10 largest claims to occur in Egypt before 1982 expressed in 1970 prices have been calculated. To reach these results, the author has suggested the use of the earned premium, under certain conditions, instead of using the number of claims as a correction factor and has produced satisfactory results by this modification. Moreover, it is easy to produce such data from published statistics and this simplifies the application of the theory. To produce such results from a not very large amount of data as that used in this thesis gives strong support to the theory where previously it was proved to be useful only in situations where large losses alone are available for analysis.

Since there are different methods of estimating the parameters a_m snf b_m, the results would involve errors due to the method used and these results should be interpreted with caution.

It is appropriate to conclude this thesis with Kimball's words (1949), "As so often happens in research, the first analysis, although not giving the final answer, suggests the next step", as has been shown to be true.

REFERENCES

Abbott, M.A. & Others (1974),
Some Thoughts on Technical Reserves and Statutory Returns in General Insurance. Journal of the Institute of Actuaries vol. 101, Part 2, p.133.

Abbott, M.A. & Others (1981),
Some Financial Aspects of General Insurance Company, Journal of the Institute of Actuaries, vol. 108, Part 2, p.119.

Abdel Rahman, G.
The Principles of Insurance (In Arabic), Dar El Nahda - Cairo - Egypt.

Alfred Eke, A.
The Principles of Insurance, London; Sir Isaac Pitman & Sons Ltd.

Batten, A.G.M. & Dinsdale, W.A.
Motor Insurance, Stone & Cox Ltd., 3rd Edition.

Beard, R.E. (1963a)
4th Astin Colloquium. Subjects for discussion, Theory of Extreme Values and Application to Insurance Problems, The Astin Bulletin, vol. II, Part 3, p.313.

Beard, R.E. (1936b)
Some Notes on the Statistical Theory of Extreme Values, The Astin Bulletin, Vol. III, p.6.

Beard, R.E. & Others (1978)
Risk Theory (The Stochastic Basis of Insurance), Chapman and Hall, 2nd Edition.

Benckert, Lars-Gunnar,
The Lognormal Model for the Distribution of One Claim, The Astin Bulletin, Vol. II, p.9.

Benjamin, B. (1977),
General Insurance, Heinemann, London.

Bryant, E.C. (1966),
Statistical Analysis, 2nd Edition, McGraw-Hill Book Company, New York-London.

Clarke, T.G. & Others,
(Technical Reserves Working Party), IBNR Reserves, A paper presented at Giro Seminar - Kendal, November 1975, deposited in the Institute of Actuaries' Library.

Coe, Leonard D.
Extreme Values and Retentions, Unpublished paper.

Draper, N.R. and Smith, H.
Applied Regression Analysis, John Wiley & Sons Inc.

Excess of Loss - Methods of Reinsurance, Report by Advanced Study Group No. 201, The Insurance Institute of London, 1978.

Galambos, J. (1978),
The Asymptotic Theory of Extreme Order Statistics, John Wiley & Sons, New York.

Gerber, H.V.
An Introduction to Mathematical Risk Theory, Huebner Foundation Monograph 8.

Glossary of Terms Used in Insurance, Issued by The Victory Insurance Company Limited, London.

Gumbel, E.J. (1954),
Statistical Theory of Extreme Value and Some Practical Applications. U.S. Dept. of Commerce, National Bureau of Standards, Applied Mathematics, Series No. 33.

Gumbel, E.J. (1958),
Statistics of Extremes, Columbia University Press.

Hey, G.B. & Others (1980),
Compensation for Personal Injury, Journal of the Institute of Actuaries, Vol. 107, Part 4, p.373.

Insurance Facts and Figures, British Insurance Association (B.I.A.), 1979.

Jung, J. (1965),
On the Use of Extreme Values to Estimate the Premium from an Excess of Loss Reinsurance, The Astin Bulletin, Vol. III, Part 2, p.178.

Kimball, B.F. (1946),
Sufficient Statistical Estimation Functions for the Parameters of the Distribution of Maximum Values, Ann. of Mathematical Statistics, Vol. 17, p.299.

Kimball, B.F. (1949),
An Approximation to the Sampling Variance of an Estimated Maximum Value of Given Frequency Based on Fit of Doubly Exponential Distribution of Maximum Values, Ann. of Mathematical Statistics, Vol. 20, p.110.

McCormick, B.J. & Others (1977),
Introducing Economics, 2nd Edition, Penguin Modern Economics Texts.

Probability Tables for the Analysis of Extreme-Value Data, U.S. Dept. of Commerce, National Bureau of Standards, Series No. 22.

Ramachandran, G. (1970),
Some Possible Applications of the Theory of Extreme Values for the Analysis of Fire Loss Data, Fire Research Station, Fire Research Note No. 837, Sept. 1970.

Ramachandran, G. (1972),
Extreme Value Theory and Fire Losses - Further Results, Fire Research Station, Fire Research Note No. 910, Jan. 1972.

Ramachandran, G. (1974),
Extreme Value Theory and Large Fire Losses, The Astin Bulletin, Vol. III Part 3, p.293.

Ramachandran, G. (1981),
Extreme Order Theory and Third Party Motor Insurance, A paper presented to the 15th Astin Colloquium, Loen, Norway, June 1981.

Ramachandran, G. (1982),
Properties of Extreme Order Statistics and their Application to Fire Protection and Insurance Problems, Fire Safety Journal, May 1982, pp.59-76.

Ramachandran, G. & Taylor, H.D. (1975),
Management of Fire Risk and Extreme Value Theory, Papers in Risk Management, 3, The Munich Conference, October 1975.

Reid, D.H. (1978),
Claim Reserves in General Insurance, Jouranl of the Institute of Actuaries, Vol. 105, Part 3, p.211.

Reinsurance Problems in Developing Countries, Study by the UNCTAD Secretariat, United Nations, New York, 1973.

Rogers, F.E. (1976),
Investigation of a Technique for Estimating the Probability Distribution of Fire Loss Based on Extreme Value Theory, Fire Research Station, Fire Research Note No. 1062, December 1976.

Rogers, F.E. (1977),
The Estimation of Parameters of Normal Distribution from Data Restricted to Large Values, Fire Research Station, November 1977.

Sarhan, A.E. & Greenberg, B.G.
Contributions to Order Statistics, John Wiley & Sons.

Tables of the Incomplete Gamma Function,
Edited by Karl Pearson, F.R.S., Cambridge Univ. Press; Reissue 1934.

Wannacott, T.H. & Wannacott, R.J. (1977)
Introductory Statistics for Business and Economics, 2nd Edition, John Wiley & Sons.