

UNIVERSITY OF GHANA



**PERFORMANCE MEASUREMENT OF PROBABILITY
DISTRIBUTIONS IN MODELLING NON-LIFE INSURANCE
CLAIMS**

BY

SAMUEL KWEKU-BIO DZIDZORNU

(10279227)

A THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS AND
ACTUARIAL SCIENCE, UNIVERSITY OF GHANA
IN PARTIAL FUFILLMENT OF THE REQUIREMENT FOR THE AWARD OF
MASTER OF PHILOSOPHY IN ACTUARIAL SCIENCE

May, 2019

DECLARATION

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

SAMUEL K. DZIDZORNU

.....

.....

Candidate

Signature

Date

(10279227)

Certified by:

DR. RICHARD MINKAH

.....

.....

Principal Supervisor

Signature

Date

DR. SAMUEL IDDI

.....

.....

Co-Supervisor

Signature

Date

DEDICATION

This work is dedicated to my parents, Peter Y. Dzidzornu and Ruth A. Ankrah, for the lifelong inspiration, nurturing, and blessings.

To my siblings, Godwin K. Dzidzornu and Angela E. Dzidzornu, for the bond of affection.

To my beloved, Marjorie O. Odei, for the depth of dedication in friendship.

ACKNOWLEDGEMENTS

To Jehovah, the Eternal Father, be the Glory through Jesus Christ, the Author and Finisher of my faith. His Love, Grace, and Providence has sustained me in this pursuit.

I am immensely grateful to my supervisors, Dr. Richard Minkah and Dr. Samuel Iddi, for the diligent academic support throughout the course of study. Their scholarly contributions, timely guidance and continued motivation has made this work possible.

I express gratitude to Dr. Doku-Amponsah, Head of Department, and all lecturers of the Department of Statistics and Actuarial Science, University of Ghana, for the consistent reviews and constructive remarks provided during the research. Also, I appreciate the entire departmental staff for the cordial work relationship.

Profound gratitude is extended to my colleague, Abdul-Rashid Abdul-Rahaman of National Insurance Commission, for the assistive role in obtaining relevant data to facilitate the research.

ABSTRACT

In this thesis, we model non-life insurance claims by using the two-parameter Negative Binomial (NB) and three-parameter Discrete Generalised Pareto (DGP) distributions. Data from National Insurance Commission (NIC) on Reported and Settled Claims counts for the period 2012 - 2016 were considered. The maximum likelihood estimation (MLE) was adopted to fit Negative Binomial and Discrete Generalised Pareto to the count data. In the latter case, the estimation involved two steps. First, the μ and $(\mu + 1)$ frequency method (Prieto et al., 2014) of generating initial estimators, was modified to suit the characteristics of the count data under study. Second, the parameter estimates were obtained by MLE, using the initial values from the modified μ and $(\mu + 1)$ frequency method. In addition, a bootstrap process was used to obtain the standard errors of the estimators of the DGP parameters. The models were compared using the information criteria, AIC and BIC. Under Reported and Settled Claims categories, each criteria was found to favour the DGP model. Therefore, the DGP model is recommended, as it provides a better fit to the non-life claims data.

CONTENTS

DECLARATION	i
DEDICATION	ii
ACKNOWLEDGEMENT	iii
ABSTRACT	iv
ABBREVIATION	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
1 INTRODUCTION	1
1.1 Background and Problem Statement	2
1.2 Objectives of Study	4
1.3 Scope and Contribution	5
1.4 Significance of Study	6
1.5 Organisation of Study	6
2 LITERATURE REVIEW	8
2.1 Distribution of Probability	8
2.2 Negative Binomial Distribution	9
2.2.1 Background of Development	9
2.2.2 Empirical Applications	11
2.3 Discrete Generalised Pareto Distribution	12

2.3.1	Background of Development	13
2.3.2	Empirical Applications	14
2.4	Overview of Insurance Industry	15
2.4.1	Historical Development in Ghana	16
2.4.2	Insurance Service Classifications	17
2.4.3	Challenges and Interventions	18
2.4.4	Local and Global Outlook	18
3	METHODOLOGY	20
3.1	Overview of Data	20
3.2	Axioms of Probability Distributions	21
3.2.1	Discrete Probability Distributions	21
3.2.2	Continuous Probability Distributions	22
3.3	Parameter Estimation Method	22
3.3.1	Maximum Likelihood Estimation	23
3.3.2	Theoretical Definitions	23
3.3.3	Application to Data	24
3.4	Negative Binomial Distribution	24
3.4.1	Theoretical Definition	25
3.4.2	Distributional Properties	26
3.4.2.1	Proof: 1st Moment	26
3.4.2.2	Proof: 2nd Moment	27
3.4.2.3	Results: 3rd & 4th Moments	28
3.4.3	Estimation of Parameters	29
3.5	Discrete Generalised Pareto Distribution	30
3.5.1	Theoretical Definition	30
3.5.2	Distributional Properties	32
3.5.3	Estimation of Parameters	33
3.5.4	Estimation of Standard Errors	35
3.6	Model Selection Criteria	36

3.6.1	Akaike Information Criterion	37
3.6.2	Bayesian Information Criterion	38
3.7	Statistical Computing Environment	38
4	DATA ANALYSIS AND DISCUSSIONS	39
4.1	Description of Data	39
4.1.1	Composition of Data	40
4.1.2	Descriptive statistics	41
4.1.2.1	Policy Subscriptions Count	41
4.1.2.2	Reported Claims Count	43
4.1.2.3	Settled Claims Count	45
4.1.3	Limitations of Data	47
4.2	Model Fitting	48
4.2.1	Estimation of NB Parameters	49
4.2.2	Estimation of DGP Parameters	49
4.2.3	Comparison of Models	52
4.3	Aggregation of Data	53
4.3.1	Estimation of Parameters	54
4.3.2	Comparison of Models	54
5	CONCLUSION AND RECOMMENDATIONS	57
5.1	Conclusion	57
5.2	Recommendations	59
	REFERENCES	61
	Appendices	68
A.1	Reported Claim Counts	69
B.2	Settled Claim Counts	74
C.3	R Source Codes	78
C.3.1	DGP Log-likelihood Function	78

C.3.2	DGP Parameter Estimator	78
C.3.3	Bootstrap Resampling	79
C.3.4	Information Criteria Statistics	79

LIST OF ABBREVIATION

AIC	Akaike Information Criterion
BIC	Beysian Information Criterion
BoG	Bank of Ghana
CII	Chartered Insurance Institute
CILA	Chartered Institute of Loss Adjusters
DGP	Discrete Generalised Pareto
DWCTP	Deep Water Cape Three Points
EOM	Encyclopaedia of Mathematics
EY	Ernst & Young
FPD	Fellar-Pareto Distribution
GNPC	Ghana National Petroleum Corporation
GNU	GNU's Not Unix
GoG	Government of Ghana
GSS	Ghana Statistical Service
IBNR	Incurred But Not Reported
ICT	Information Communication and Technology
MCMC	Markov Chain Monte Carlo
MLE	Maximum Likelihood Estimation

MNO	Mobile Network Operator
NB	Negative Binomial
NIC	National Insurance Commission
NIST	National Institute of Standards and Technology
OECD	Organisation for Economic Cooperation and Development
PD	Pareto Distribution
PNDC	Provisional National Defense Council
PPD	Posterior Predictive Distribution
PwC	PriceWaterHouseCoopers
RBNS	Reported But Not settled
SAS	Statistical Analysis System
SBC	Schwarz Beysian Criterion
SEBO	Settled But Outstanding
SIC	State Insurance Company
SPSS	Statistical Package for the Social Sciences
TRF	The R Foundation
WHO	World Health Organisation

LIST OF TABLES

4.1	Descriptive summary of Policy Subscriptions Count dataset	42
4.2	Descriptive summary of Reported Claims Count dataset	44
4.3	Descriptive summary of Settled Claims Count dataset	46
4.4	k_i labels for Reported and Settled Claims Count	49
4.5	Parameter estimates from Negative Binomial model	50
4.6	Relative frequencies of Claims Count categories	50
4.7	Parameter estimates from Discrete Generalised Pareto model	52
4.8	AIC statistics for NB and DGP model	53
4.9	BIC statistics for NB and DGP model	53
4.10	Parameter estimates from Negative Binomial model	54
4.11	Parameter estimates from Discrete Generalised Pareto model	55
4.12	AIC statistics for NB and DGP model	55
4.13	BIC statistics for NB and DGP model	56
1	Restructured Reported Claims Dataset	69
2	Restructured Settled Claims Dataset	74

LIST OF FIGURES

3.1	PDF of NBD for selected successes - r (Mathworks, 2015)	25
3.2	Pmf of DGP for selected α, λ, μ - Prieto et al. (2014)	32
4.1	Number of non-life policies recorded in the period 2012-2016 . . .	42
4.2	Number of non-life claims reported in the period 2012-2016	44
4.3	Number of non-life claims settled in the period 2012-2016	46

CHAPTER 1

INTRODUCTION

This thesis is a comparative study of Negative Binomial (NB) and Discrete Generalised Pareto (DGP) distributions in insurance. Utilising a practical dataset, the study seeks to assess the performances of both distributions in modelling non-life claim counts. To facilitate the discussions, an explanation of key terminologies in the title would be relevant. As such, Performance Measurement, Probability Distributions, Non-Life Insurance, and Claims, are briefly introduced in the following paragraphs.

First, Performance Measurement involves the collection, analysis, and reporting of information regarding the capabilities of individuals, groups, organisations, systems, or components (Upadhaya et al., 2014). Thus, it considers the study of processes, parameters or phenomena, to determine the match between output and expected results, for informed decision making. Furthermore, the exercise constitutes a series of actions, carried out to derive comparative insights, for establishing effective quality controls. In this study, performance measurement will require the estimation of distributional parameters by Maximum Likelihood Estimation (MLE). Afterwards, the Akaike (1974) and Bayesian (Schwarz, 1978) information criteria will provide a statistical basis for selecting a distribution, which provides the most suitable fit to the research data.

Second, Probability Distribution is described by Everitt B.S. and Skrondal A. (2010) in the Cambridge Dictionary of Statistics (4th Edition), in discrete and continuous terms. In the former case, it is regarded as a mathematical function that indicates the probability of each value of a discrete random variable. The Poisson and Binomial distributions are classical examples. In the latter case, it is considered as a curve which

describes the probability that a continuous random variable, falls within a specified interval under the curve. Examples include the Normal and Exponential distributions.

Third, non-life insurance addresses the socio-economic need for financial loss recovery, through the offering of protection against random events. Conditioned on periodic payment of a predetermined amount, called premium, non-life policies are designed to provide coverage against insured probabilistic events; for individuals, private, and public institutions. The random payments effected, in response to occurrences of such events are termed as insurance claims (Wutrich, 2014).

In chapters two and three of the study, concepts underlying the foregoing terms will be explored in detail. The remainder of this chapter is organised into three parts. Respectively, sections 1.1, 1.2, and 1.3, presents the Background and Problem Statement, Scope and Contribution, and Organisation of Study.

1.1 Background and Problem Statement

In Ghana, the insurance industry is governed under Insurance Act, 2006 (Act 724). In compliance with Core Principles of the International Association of Insurance Supervisors (IAIS), the Act confers regulatory powers on the National Insurance Commission (NIC). As at June 2016, the insurance industry was comprised of the following actors: 26 non-life insurance companies; 23 life companies; 3 reinsurance companies; 70 Broking companies; 1 reinsurance company; and 7000 insurance agents (NIC, 2016). Evidently, non-life institutions form a majority of the forty-nine direct insurance companies in the country. This justifies the focus of study on claims associated with non-life insurance. An inspection of non-life data on claims and premiums, reveals key sectoral developments, as well as challenges facing the Ghanaian insurance landscape. As such, the following problem statements are deductions of the authour, drawn from data produced by NIC.

First, from 2012 to 2016, number of non-life policy subscriptions increased from

626,522 to 931,603, and declined to 766,048. This corresponds to a fall in growth rate from 48.69% to -17.77% within the period. Yielding an average industrial gain of 7% year-on-year, the regressive growth in demand for non-life products marks the first major challenge facing the industry. Largely, the problem is attributable to low corporate awareness creation, on relevance of non-life insurance arrangements to the public.

Second, inadequate liquidity results in a difficulty of some establishments to duly finance claims obligation. In 2015 for instance, at least 10.71% of non-life companies faced varying forms of exposure to liquidity risks. Thus, three out of twenty-eight companies recorded a deficit in gross premiums against claims. Given a total receipt of GHC 152,177,806 reported by the distressed enterprises, gross claims exceeded premiums by 32.49%, equivalent to GHC 49,438,753. This condition often poses liquidity depletion risks, as affected companies resort to their reserves for claims settlement.

Third, access to reliable actuarial data on Ghana's demographic dynamics presents a major challenge. According to Abaitey and Oduro (2017), life table estimations and mortality models in Africa are mostly unavailable, inconsistent, or inaccurate. The authors noted that this misinforms the computational assumptions, underlying actuarial values such as premiums, claims, sums assured, and annuities. The study also indicated that Kpedekpo (1969) proposed functional life tables, based on the World Health Organisation (WHO) tables of 1960. WHO further published life tables in 2000, 2010, and 2012. Later, empirical life tables were developed in 2010 by the Ghana Statistical Service (GSS). However, the eventual life tables were mostly adopted from South Africa, among other countries, and modified accordingly (Abaitey & Oduro, 2017). This presents numerous challenges associated with contextualisation of data.

By extension, the problems discussed above can be stated as the following summary:

1. Access to reliable distributional research on Ghana's claims dynamics presents a challenge.

2. Limited actuarial information on claims count characteristics results in misinformed premium pricing decisions.
3. Difficulty in optimal funds allocation, leading to inability of some institutions to finance claims obligations when due.

The foregoing challenges can be addressed through a detailed comparative study of probability distributions. As such, the study seeks to examine the performances of NB and DGP distributions, in describing data on non-life insurance claim counts in Ghana. The distribution which exhibits a better fit to the claims data being investigated will be proposed.

1.2 Objectives of Study

The main objective of this study is to model the claim counts of licensed non-life insurance companies in Ghana. Specifically, the research seeks to:

1. Investigate claims data characteristics under Negative Binomial and Discrete Generalised Pareto distributional cases.
2. Select a suitable statistical model for describing insurance claim counts of non-life service providers.

In this study, the choice of the distributions was informed by exploratory descriptive analysis of the claims counts data. This required the computation of key statistical measures including mean, standard deviation, skewness, and kurtosis. Eventual decision was based on the data features identified; through interpretation of the measures, and outcomes of preliminary graphical exercises.

1.3 Scope and Contribution

A fair amount of research has been devoted to the study of insurance claims from varying perspectives. Taking randomised spatial effects into consideration, GschloBl and Czado (2007) studied the modelling of claim frequencies and sizes in a non-life insurance context. In contrast with the assumption of independence, posited by the classical Compound Poisson Model, the authors allowed for dependency between the two data classes. Estimating the parameters via Markov Chain Monte Carlo (MCMC), Bayesian analysis was followed, while the Posterior Predictive Distribution (PPD) was utilised for comparison of models. The study established that the dependence postulate for claim counts and sizes, considerably improves the model predictability of non-life insurance claims. Additionally, Pacakova (2011) investigated collective risk modelling and simulation, applicable to scenarios with limited data on non-life policies. Denoting the aggregate annual claims amount, individual claim sizes, and number of claims by the random variables S , $X_i (i \leq N)$, and N respectively, the paper used the compound model $S = X_1 + X_2 + \dots + X_N$, to determine the insurance premium and calculate value of risk. In the end, values, s_1, s_2, \dots, s_n , were generated from a simulation, with parameter estimation processes carried out, and Goodness of Fit tests duly conducted. Among others comprehensively discussed in Chapter two, the cited works provide an overview of notable scholarly developments in insurance claims analysis. However, there is little in research to illustrate the application of Discrete Generalised Pareto and Negative Binomial distributions to claims modelling. Therefore, this study contributes to the subject in three areas:

1. The study proposes an alternative data structure for organising industrial claims counts. Potentially, the proposal presents a more informative grouping of count observation, for enriching the results derived from modelling insurance count datasets.
2. A modification is made to the μ and $\mu + 1$ frequency method, proposed by Prieto

et al. (2014), for obtaining initial estimators. This extends its application to real count data which exhibits varying intervals between observations.

3. Algorithms written for modelling the count datasets, provide a guide for performing statistical computations, in further claims-related research.

To the best of our knowledge, limited studies on the Discrete Generalised Pareto distribution has been undertaken. As such, the findings presented in this work adds to the scarce amount of literature on the distribution. On a broader scale, application of the considered distributions to non-life claims modelling, provides a practical outlook on the relevance of probability models in actuarial contexts.

1.4 Significance of Study

The study is significant in broadening the scope of non-life claims research in three aspects. First, it will complement the research efforts of the limited number of Ghanaian actuaries, seeking to perform distributional count modelling on national claims data. Second, the findings will provide useful insights on liquidity management for effective claims management by non-insurance service providers. Third, the outcomes will assist the industry regulator, in identifying key performance indicators for monitoring and evaluation purposes. These areas of significance will enhance the operations of market players, resulting in sustainable growth of Ghana's insurance industry.

1.5 Organisation of Study

The rest of the study is organised into four chapters. Chapter two reviews academic and industrial literature on the subject, whereas Chapter three discusses the methodology for investigating the distributions under study. In addition, Chapter four employs data

from the insurance industry regulator, to perform claims analysis based on the approach outlined in the previous chapter. Chapter five concludes the study by presenting key findings from the analysis, and recommending measures for consideration by stakeholders.

CHAPTER 2

LITERATURE REVIEW

Conducting an in-depth research requires the study of earlier works on the subject. This chapter reviews relevant literature on selected probability distributions and Ghana's insurance industry. The chapter consists of four sections, which discuss key scholarships in the research scope. Section one provides an overview of probability distributions, as acknowledged in statistics. The backgrounds and applications of the NB and DGP distributions are explored in sections two and three respectively. Last, section four delves into Ghana's insurance ecosystem, and situates its developments in a broader global context.

2.1 Distribution of Probability

A probability approximates the possibility of occurrence of an event, conditioned on random factors. It can be represented by a graphical function which assigns a likelihood to each event. The functional representation of such assignments, for a series of random outcomes, is identified as Probability Distribution (Viti et al., 2015). In statistics, a distinction between the types of random variables, enhances the understanding of probability distributions. First, a variable is considered random, if it is expressible as a function which assigns unique values to each outcome of a random phenomenon, such as an experiment (Easton & McColl, 1997). Furthermore, Hitchcock (1997) draws the following distinction between discrete and continuous random variables. On one hand, discrete random variables take on values belonging to the set of natural numbers. As such, they have countable gaps within the range

of possible numerical values. Examples include, frequency of flights travelled by an aircraft within a week, and number of job applicants qualified for a company aptitude test. On the other hand, continuous random variables take on a range of possible values lying within an unbroken interval. The length of artificial hair produced daily by a manufacturing plant, and height of factory operators are examples.

2.2 Negative Binomial Distribution

The negative binomial distribution (NBD) is alternatively referred to as Pascal distribution and denoted by $X \sim \mathcal{NB}(r, p)$. Given X , a discrete random variable, r and p are parameters; with r as a fixed positive integer, and probability, p . In some statistical texts, slight variations in the definitions of NBD may be observed. That noted, readers are required to understand the particular parameterisation used in any NBD literature of interest. The development and applications of the NBD are discussed in the following subsections.

2.2.1 Background of Development

In statistics, the NBD is studied as a variant of the Binomial Distribution (BD). Among other important distributions, the BD and its variations such as the NBD, offer a wide range of applicable continuous models in probability theory. Pascal (1679) and Bernoulli (1713) (as cited in Philippou and Antzoulakos, 2011) were the first to derive special instances of the BD. The NBD indicates the number of failures before the n th success in a sequence of independent Bernoulli trials with probability p of success on each trial.

Historically, Fisher (1941) explored two algebraic cases of the binomial expansion, namely, positive binomial and negative binomial. According to the paper, the positive binomial, expressed as $(q + p)^n$, frequently occurs with a known integer, n , while the fractions p and $q = 1 - p$ remain unknown. However, certain mathematical

reasons may account for cases of unknown n . First, the expansion produces negative coefficients after a point in the series if n is not integral. However, the negation of coefficients cannot be regarded as negative frequencies. As a result, the expansion bears no correspondence with any distribution. Second, cases of an integral and unknown n may arise. In this regard, the paper argues that given a sufficiently large sample, the number of frequency classes is necessarily one more than n . This makes n determinable without reference to the actual frequencies (Fisher, 1941). On the other hand, the negative binomial, written as $(q - p)^{-k}$ after Haldane (1941) (as cited in Fisher, 1941), with $q = 1 + p$ and positive k , yields the term $q^{-k} \frac{(k+x-1)!}{x!(k-1)!} \left(\frac{p}{q}\right)^x$ during expansion. Given that the expression generates positive terms for all positive values of x , with k being integral or otherwise, Fisher claims that the simultaneous evaluation of p and k presents a challenge to which no recourse is provided by the positive binomial. Fisher's work further notes that an extension of the underlying series conditions for Poisson, results in the negative binomial with unknown exponent (Fisher, 1941).

In a more recent literature, Hilbe (2011) offered a brief background of the NBD. In his book, *Negative Binomial Regression* (2nd Edition), reference is made to Todhunter's (1865) report (as cited in Hilbe, 2011), stating that NBD was identified in 1713 by Pierre de Montmort, as the number of failures, y , preceding the k th success in a binary trial series. Early in the twentieth century, William Gosset also derived the NBD, building upon an earlier work of Gauss (1823) (as cited in Hilbe, 2011). Enhancing Montmort's derivation, Gosset's (1907) deduction of the NBD (as cited in Hilbe, 2011), arose from a study of the sampling error associated with counting yeast cells using a haemocytometer. Furthermore, Hilbe writes that the first work on fitting the NBD with a maximum likelihood algorithm was published in 1960 by Leroy Simon. Simon (1961) (as cited in Hilbe, 2011) and others explored the actuarial perspective by fitting the Poisson and NBD to insurance data.

The NBD derives its name from the fact that, the distribution is generated by a binomial with a negative exponent (Encyclopaedia of Mathematics [EOM], 2012). Thus, the coefficient of the expansion of $p^r(1 - qz)^{-r}$ in powers of z , produces the

required probabilities. The NBD is defined by generating and characteristic functions, respectively expressed as $P(z) = p^r(1-qz)^{-r}$ and $f(t) = p^r(1-qe^{it})^{-r}$ (EOM, 2012). The NBD has undergone several phases of development, taking unique functional forms over the course of centuries. In recent times, fitting the NBD is carried out via algorithms implemented in statistical computing environments. Particularly, the NBD is fitted in the R software using packages *foreign*, *ggplot2*, and *MASS* (UCLA Institute for Digital Research and Education [IDRE], n.d.). Among other resources, the packages enhance the modelling of practical data on counts, such as non-life insurance claims as considered under the study.

2.2.2 Empirical Applications

The NBD presents a wide array of cross-disciplinary applications. Johnson, Kotz, and Kemp (1992) (as cited in Hoffman, 2003), noted that the distribution has been explored in accident statistics, birth and death processes, market research, econometrics, biometrics, and ecology, among other fields of knowledge.

In pharmaceutical statistics, Hoffman employed the NBD in deriving control limits for count data with extra-Poisson variation. The authour applied the negative binomial utility to bacteria count data sampled from a water purification system. Using the SAS GENMOD procedure, the work considered the negative binomial model in comparison with poisson model, arguing that the former offers higher modelling flexibility in fitting count data.

In medical informatics, scholars have utilised the NBD for varying forms of predictive studies. Specifically, Carter and Potts (2014) investigated the predictability of hospital length of stay, using data extracted from an electronic patient record system. The NBD was applied to the study of significant factors influencing length of stay - age, gender, consultant, discharge destination, deprivation and ethnicity. Overall, the negative binomial model recorded a significant predictive accuracy, for the patients' length of stay over a five-year timeline.

Also, the distribution provides practical uses in the domain of hydrocarbon exploration. In April 2015, Government of Ghana (GoG), represented by Ghana National Petroleum Corporation (GNPC), entered into a memorandum of understanding (MoU) with Exxon Mobil to negotiate a Petroleum Agreement, for acquisition of exploration and production rights by the latter (Adombila, 2017). Lying in water depth between 2000 and 4000 meters, and situated at 150 kilometres offshore Ghana, the Deep Water Cape Three Points (DWCTP) block was envisioned to hold nineteen drilled exploratory wells, each of which would have varying probabilities of oil strike. To aid resource allocation, the negative binomial would be useful in answering critical geological questions, such as: what is the probability that the fourth well will produce the first strike, or what is the probability that the third strike would flow from the ninth well? The distribution can further assist GoG and Exxon, to determine the mean and variance of the number of wells that must be drilled, if the latter needs to set up a predetermined number of production points.

In general, the usefulness of NBD is evident in the modelling of real-life processes, involving independent successes and failures prior to reaching a final desired state.

2.3 Discrete Generalised Pareto Distribution

The Discrete Generalised Pareto distribution, denoted by $X \sim \mathcal{DGP}(\alpha, \lambda, \mu)$, is a parametric model with shape, scale, and location parameters. Popularly used in tail modelling, the DGP assumes different forms based on omission of one or two of its parameters. For instance, the DGP without a location parameter, thus $\mu = 0$, becomes a Discrete Lomax (DLO) distribution - written as $\mathcal{DLo}(\alpha, \lambda)$. Such discrete and/or continuous derivational forms, invariably belong to the family of generalised Pareto distributions. An overview of background and applications of the DGP distribution are provided in the subsections below.

2.3.1 Background of Development

The Discrete Generalised Pareto Distribution (DGP) is derived by discretising the generalised form of the Pareto Distribution (PD). As such, a review of literature on DGP requires a developmental overview of the PD. The PD was discovered by Vilfredo Pareto, following a study of data on income tax from numerous countries (Geerolf, 2017). Revealing a linear relationship, the study, according to Geerolf, plotted a cohort of income earners against a specified monetary threshold, on a double logarithmic paper. In comparison with bell-shaped curves, the author described the resulting income distribution as exhibiting more skewness and tail heaviness. This outcome led Pareto in the 1800s, to discover the fundamental theories of inequality as widely applied in economics.

Later in the 1900s and 2000s, the PD, based on the principles of inequality, underwent relevant modifications to suit various disciplines. One of the earliest adaptations, known as the Pareto principle or 80/20 rule was popularised by John M. Juran, a management consultant. The rule states that, about 80% of effects are derived from 20% of causes, for most events (Bunkley, 2008). In his earliest paper, *Cours d'économie politique*, Pareto situated the rule in a context of land ownership among the Italian population. He argued that, 80% of Italian land was owned by 20% of the population (Pareto, 1971). In the twenty-first century, Pareto's principle has witnessed several applications in technology-related disciplines. In particular, Lowell Arthur noted that, 80% of errors are contained in 20% of code written for software (Pressman, 2010). A classical example is recorded by Rooney (2002), stating that 80% of Microsoft's system issues were eliminated by fixing the top 20% of most reported bug cases.

Furthermore, the Pareto principle has undergone distinct phases of evolution, and gained relevance in the study of probability distributions. In statistics, PD is often used in fitting trends that describe relative proportions in human activities. Practically, the distribution provides a framework for modelling income distributions, as it was

originally borne out of Pareto's attempt to describe wealth allocation within societies. At least, five main types of PD exist, in accordance with the inherent levels of complexity, as determined by the range of support and parametric bounds. The types of PD are given hierarchically as Pareto Type I, II, III, IV, and Feller-Pareto distributions (FPD) (Arnold, 1983). The distributional parameters of the Pareto Types are μ , α , σ , and γ , respectively denoting the location, tail index or shape, scale, and inequality. While types I, II, and III are special cases of IV, the FPD forms a generalisation of the Pareto Type IV (Feller, 1971).

The Generalised Pareto Distribution (GPD) was explored by Davison, Smith (1984), and Castilo (1997, 2008), among other scholars, following an earlier introduction by Pikands (1975) (as cited in Meirlus-Mazilu, 2010). A family of continuous probability distributions, the GPD is traditionally employed in modelling the tails of other parametric distributions. However, further studies of the GPD, a continuous distribution, unveiled its usefulness and adaptability to discrete contexts. As such, the Discrete Generalised Pareto distribution (DGP) is derived by discretising the GPD through appropriate statistical techniques. To the best of the researcher's knowledge in R, there are no readily available software packages for fitting DGP to real data. Therefore, statistical programming will be performed to derive a set of algorithms, suitable for modelling non-life claim counts under the DGP distribution.

2.3.2 Empirical Applications

The Pareto distribution is traditionally applied in allocation of wealth among individuals in a given population (Pareto, Busino, & Bousquet, 1964). Thus, originally, usefulness of the Pareto model was evident in addressing issues related to equitable distribution of scarce resources. Therefore, all varying forms of the distribution, such as DGP , are widely applicable in areas such as stocks returns analysis, human settlement studies, and oil field reservations, among others. Also, given that road injury has been ranked among the top ten causes of global deaths, for the period 2000-2016

(World Health Organisation [WHO], 2018), emerging scholars have directed some research efforts towards road accident analysis. Particularly, Prieto et al. (2014) harnessed the Discrete Generalised Pareto distribution in modelling data on road accident blackspots. Defining blackspots as a 100m section of a road with at least three road crashes per annum, the study aimed at demonstrating the role of probabilistic models, focusing on DGP , in describing road traffic networks. In furtherance of the analysis, a five-year retrospective data on number of crashes and fatalities on blackspots in Spain, were obtained from Spanish General Directorate of Traffic. Spanning the period 2003 to 2007, the dataset was modelled over 16,552 accidents with DGP and DLo , and compared with Negative Binomial outcomes. Prieto et al. (2014) studied the respective distributional properties, and drew statistical conclusions in favour of DGP (and DLo), following the requisite parametric estimations, goodness-of-fit tests, and model selection based on appropriate information criteria.

2.4 Overview of Insurance Industry

Globally, the insurance sector of emerging economies are invariably affected by trends in developed markets. For instance, as the advanced economies of Europe and North America approach recovery, the developing markets of Asia and Latin America experience steadiness in growth (PriceWaterHouseCoopers [PwC], 2013). Later in 2016, a strong underwriting performance was achieved among countries belonging to the Organisation for Economic Cooperation and Development (OECD), as well as selected non-member countries in Africa, Latin America, Asia, and Southern Europe. This performance was the result of an increase in gross premiums in excess of liabilities associated with claims settlements. Among forty reporting countries studied by OECD in 2016, gross premiums averaged an increase of 3.7% in the life sector, and 2.0% in the non-life service. Specifically, Costa Rica, Turkey, and Russia, recorded over 30% real growth, representing the highest increase in life gross premiums between 2015 and 2016. In the same period, a fall in price of unit-linked products resulted in a dip in

life premiums by 27.6%, 24.0%, and 14.1% in Finland, Portugal, and Poland (OECD, 2018).

Furthermore, the global macroeconomic environment has immensely influenced the insurance industry in the past decade. This is reflected in the relationship between economic tides and insurance patronage; in that, high demand and low subscription are experienced in periods of economic growth and downturns respectively. For example, OECD further reports that the persistent low interest rates of 2016, in countries such as Switzerland presents bond reinvestment risks to insurers, arising from lower returns offered on new bonds, as the existing high yielding bonds mature. This reduced the fund reserves available to most insurers for designing and promoting products that suit consumers' tastes and preferences. Also, the occurrence of natural disasters such as the 2011 earthquakes in New Zealand adversely impacted claims levels.

Despite the high and low developments in the life and non-life sectors, PwC, OECD, and other international finance and insurance bodies, report positively on the industry in net real terms. Overall, the global industry has recorded positive returns on equity resulting in a significant increase in shareholder value. The following subsections summarise the industrial evolution, development challenges, and key reforms of the insurance sector in Ghana. They conclude with a discussion on future growth potential of the local industry in relation to trends in the global market.

2.4.1 Historical Development in Ghana

In 1924, insurance commenced in Gold Coast, now Ghana, with the establishment of Enterprise Insurance Company Limited, formerly known as Royal Guardian Enterprise. Subsequently, the State Insurance Company (SIC) was formed in 1962; seven years after the establishment of Gold Coast Insurance Company in 1955. The latter was the nation's first indigenous private insurance entity. A decade after Ghana had gained republic status, eleven additional companies were established in 1971. As at 1976 year-end, seven more insurance companies, one reinsurance institution,

and insurance brokerage firms were in operation. In 1989, the National Insurance Commission (NIC) was instituted through the Provisional National Defense Council (PNDC) Law 229, to exercise regulation and enhance governance of the insurance industry (Bank of Ghana [BoG], n.d.). However, the 1989 Insurance Law was replaced by Insurance Act, 2006 (Act 724), which among other provisions, prohibits composite insurance operations (NIC, 2017).

2.4.2 Insurance Service Classifications

In Ghana, Life Insurance and Non-Life or General Insurance are the two broad categories of insurance service operations. The differences between the two lie in the nature of risks covered. While life insurance pays a sum upon survival of a given term, or demise of an insured, non-life policies offer protection against risks of loss resulting from occurrence of specified events. Although both have recorded steady growth in the past years, the largest number of firms belong to the latter. The NIC places non-life insurance businesses into five core classes: Fire, Burglary and Property Damage; Marine and Aviation; Accident; Motor; and General Liability. The Chartered Insurance Institute [CII] (n.d.) describes the aforementioned classes as follows. First, property insurance provides protection against fire or theft, for physical property such as building, equipment, and inventory. Second, marine insurance offers coverage for interests insured against perils of sea such as fire, stranding, and bad weather. In addition, aviation policies address air-borne damages, as well as liabilities associated with passengers and cargo. Next, motor policies provide coverage for liabilities associated with the use of vehicle, whereas accident insurance offers compensation for injury-related costs in the event of an accident. The protection may take the form of third party or comprehensive insurance coverage. Last, liability insurance protects an insured from risks of incurring liabilities, and settling damages, arising from claims of legal suits. Therefore, non-life insurance provides diversity of coverage options, resulting in a wide range of products, intended to match the varying risk needs of

policy subscribers. Fewer remarks on life insurance are provided in this section, since the study focuses on claims analysis pertaining to non-life coverage.

2.4.3 Challenges and Interventions

The insurance industry has witnessed several problems throughout its development in Ghana. In particular, numerous non-life enterprises have experienced two major interrelated issues: low liquidity for claim settlements, and varying risks of business insolvency. Deductions from recent NIC data suggests that these challenges could have arisen from erroneous financial forecasts. However, this could be further attributed to the under-resourced supervision of non-life insurance companies. To address the issue, the NIC embarked on special initiatives which centred on integration of Information Communication and Technology (ICT) resources into its regulatory operations. Specifically, the commission invested in database digitisation measures, which led to a streamlined access to relevant national insurance data. The digitisation required an institutional shift from paper-based archiving mediums to electronic file transmission systems. As such, from 2012 to date, readily-accessible data insights have been harnessed by NIC, to strengthen its capacity to monitor, evaluate, and intervene in the affairs of service providers.

2.4.4 Local and Global Outlook

Leading insurance experts, such as Ernst & Young (EY), OECD, and PwC, have reported positive forecasts in various publications. Citing an eMarketer research, EY (2015) reported that the pace of technological evolution, and user adoption rate will grow mobile payments to \$27.5 billion in 2016. As such, modern insurers could partake in the potential transactional share, by integrating digital experiences into the insurance delivery process. As the trend approaches 2020, failure to adjust existing business models to digitised value chains, may result in risks of loss of

competitiveness to agile insurance enterprises (EY, 2015). In Ghana, expert studies on the future of insurance suggest that micro-insurance holds high prospects for growth. According to the Micro-Insurance Landscape survey, micro-insurance coverage grew from 7% in 2012, to 12% in 2013. This represents an increase from 1.8 million to 3.1 million people, out of the previous year population of 24.1 million. In 2014, 28% of the population had subscribed to varying forms of micro-insurance, constituting a coverage for 7.1 million lives and properties (CDC Consult, Microinsurance Centre, & PromIGH, 2015). Considering the growth highlights above, NIC continues to urge industry players to move towards the micro-insurance sub-sector, to enhance equitable distribution of insurance services in Ghana. In this regard, NIC aims at two primary objectives. First, it seeks to reorient society against the view, that insurance is solely applicable to middle income or elite classes. To achieve the desired change in perception, the commission encourages insurance operators to design quality low-premium products, which can be afforded by the low income bracket of the population. Second, NIC seeks to increase capitalisation of the local market, by sensitising insurance enterprises to take advantage of emerging opportunities in the informal sector. Given that partnerships with Mobile Network Operators (MNOs) account for the rise of micro-insurance (Eduku, 2017), similar service collaborations and resource integrations, may expand the scale of insurance delivery in the long term.

CHAPTER 3

METHODOLOGY

As stated in chapter one, the study seeks to assess the relative performances of Negative Binomial and Discrete Generalised Pareto distributions. The performance will be measured in terms of the extent to which the distributions fit data on non-life insurance claims in Ghana. This chapter focuses on the approach to investigating the distributions, and describes the statistical methods and tools, utilised for conducting the research. It further informs the reader of parametric estimation techniques, and model selection criteria, considered in carrying out the claims study. Considering Schneider's (2014) view on methodology, the chapter discusses key theoretical concepts that guide the author's application of selected methods in the broader academic work.

3.1 Overview of Data

The study employs secondary quantitative data on claims counts from the National Insurance Commission of Ghana. Spanning over a five-year period, the historical data covers insurance claims of twenty-nine non-life service providers from 2012 to 2016. Primarily, the NIC obtains such data through mandatory actuarial reports, submitted by non-life enterprises each fiscal year. Acting in a regulatory capacity, NIC further ensures credibility of the industry data collated, through enforcement of quality control measures, which require periodic audits of licensed service providers. Further description of the research data will be provided in chapter four of the study.

3.2 Axioms of Probability Distributions

An axiom is a statement of claim, proposed to construe a fundamental truth (Downing, 2009). Alternatively referred to as postulate, it forms the basis for further logical deduction and reasoning. Axioms can also describe the properties or characteristics that must be satisfied for a mathematical or statistical statement to be considered valid. Probability distributions can be discrete or continuous as determined by specific conditions. In that regard, (National Institute of Standards and Technology [NIST], 2012) highlights the axioms for discrete and continuous cases in the following subsections.

3.2.1 Discrete Probability Distributions

Suppose $x_1 \leq i \leq \dots$ are possible values of the discrete random variable X . The function $p(x)$ is a discrete probability distribution for X if:

1. $P[X = x] = p(x) = p_x$ is the probability that x takes on specific values.
2. $p(x) \geq 0, \forall i \in \mathbb{R}$. Thus, $p(x)$ is non-negative for every real i .
3. $\sum_i p_i = p(x_1) + p(x_2) + \dots + p(x_n) + \dots = 1$. Thus, for all possible values of x , the sum of $p(x)$ is one. As stated earlier, p_i is the probability at x_i , where i constitutes all possible numerical values of x .

The second and third properties place the probability at x between zero and one. Thus, $0 \leq p(x) \leq 1$ is a consequence yielded by both conditions. In furtherance of this, a discrete function which generates values of p - greater than one or less than zero fails to be a probability distribution.

3.2.2 Continuous Probability Distributions

Suppose Z , the continuous random variable takes on values in the closed continuous interval $[a, b]$. The function $f(x)$ is a continuous probability distribution or density for Z if:

1. $p[a \leq x \leq b] = \int_a^b f(x)dx$ is the probability that x lies between a and b .
2. $f(x) \geq 0, \forall i \in [a, b]$. Thus, $f(x)$ is non-negative for every real i in the interval.
3. $\int_{-\infty}^{\infty} f(x)dx = 1$. Thus, for all possible values of x , the integral of the probability function is one.

In contrast with the discrete case, probabilities in a continuous context are defined over measurable intervals. This has two implications. First, the area under a curve between any two points marks the probability of that interval. Second, the probability at any specified point is zero. The condition of integral unity, as stated in the third property, corresponds to the discrete distributional axiom that all probabilities must sum up to one.

While numerous distributions exist for discrete and continuous cases, the intended comparative appraisal focuses on two, based on findings from initial explorations. Also, an informative comparison can be drawn when distributions of same cases are considered. As such, the negative binomial - a discrete distribution, and the discretised form of the continuous Generalised Pareto distribution, are collectively considered for the modelling.

3.3 Parameter Estimation Method

In probability theory, parameter estimation involves the use of sample data to approximate the values of parameters for a statistical distribution. Conventionally

represented by Greek alphabets to ease identification, parameters describe the effect of a unit change in a predictor, on a response, holding all other predictors constant (Analyse-it Software Ltd., n.d.; Minitab Inc., 2017). In the performance of distributional analysis, several methods of parameter estimation are available to statisticians. Among others, the approaches include Rank Regression (or Least Squares), Probability Plotting, Bayesian Inferencing, and Maximum Likelihood Estimation (MLE). Regarded as the most widely-used estimation technique (Levy, 2012), MLE is employed to evaluate the parameters of the distributions considered in the study.

3.3.1 Maximum Likelihood Estimation

As the name suggests, MLE helps to estimate the model parameters which are most likely to characterise a given dataset. Thus, it seeks to maximise the functional agreement between a model and its corresponding data. Formal MLE definitions from a Pennsylvania State University (2008) online statistical resource are presented as follows.

3.3.2 Theoretical Definitions

Suppose X_1, X_2, \dots, X_n is a random sample from a distribution which depends on one or more unknown parameters $\theta_1, \theta_2, \dots, \theta_m$. Let $f(x_i; \theta_1, \theta_2, \dots, \theta_m)$ be the probability density (or mass) function, with $(\theta_1, \theta_2, \dots, \theta_m)$ restricted to a given parameter space Ω . Then:

1. When regarded as a function of $\theta_1, \theta_2, \dots, \theta_m$, the joint probability density (or mass) function of X_1, X_2, \dots, X_n :

$$L(\theta_1, \theta_2, \dots, \theta_m) = \prod_{i=1}^n f(x_i; \theta_1, \theta_2, \dots, \theta_m)$$

$((\theta_1, \theta_2, \dots, \theta_m)$ in Ω) is described as the likelihood function.

2. If: $[u_1(x_1, x_2, \dots, x_n), u_2(x_1, x_2, \dots, x_n), \dots, u_m(x_1, x_2, \dots, x_n)]$ is the m -tuple that maximises the likelihood function, then: $\hat{\theta}_i = [u_i(X_1, X_2, \dots, X_n)]$ is the maximum likelihood estimator of θ_i , for $i = 1, 2, \dots, m$.
3. The corresponding observed values of the statistics in (2), namely:

$$[u_1(x_1, x_2, \dots, x_n), u_2(x_1, x_2, \dots, x_n), \dots, u_m(x_1, x_2, \dots, x_n)]$$

are called the maximum likelihood estimates of θ_i , for $i = 1, 2, \dots, m$

3.3.3 Application to Data

To ensure computational precision, the maximum likelihood estimation of the NB and DGP parameters will be performed in R (See section 3.7). The NB parameters will be estimated using the *mle* function and its standard arguments. To the best of our knowledge, no statistical package is available for estimating DGP parameters in R. As a result, we provide R functions for computing the estimators of the parameters of the DGP distribution. The functions are presented in Appendix C.3.

3.4 Negative Binomial Distribution

Similar to the Poisson and Binomial distributions, the NB distribution is defined over a sample with non-negative integral values. However, their differences lie in the relationship between the first two moments of the distributions. In the binomial case, the variance is less than the mean, while the poisson exhibits equality between the two moments. On the other hand, the variance of NB is greater than its mean. NB is mostly fitted by two estimation approaches - maximum likelihood, and method of moments. Given that DGP parameters are estimated by MLE, the same technique is adopted for fitting the NB, to ensure procedural consistency in the study.

3.4.1 Theoretical Definition

Suppose a sequence of Bernoulli trials is observed. By definition, a turn of each trial yields two possible outcomes - success and failure, with respective probabilities of occurrence denoted by p , and $(1 - p)$. Also, the trials are independent and p remains constant for each observation. Assuming X represents the number of trials (or failures) prior to the r -th success, then X follows a negative binomial distribution with probability mass function:

$$f(x) = P[X = x] = \binom{x-1}{r-1} (1-p)^{x-r} p^r \quad (3.1)$$

for $x = r, r + 1, r + 2, \dots$

In effect, the distribution estimates the probability $r - 1$ successes and x failures in $x + r - 1$ trials, and success on the $(x + r)$ th trial (Weisstein, n.d.).

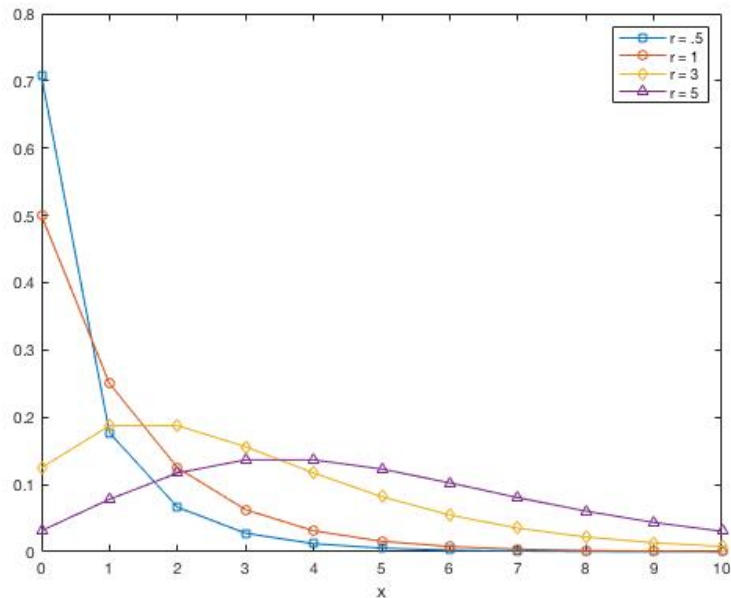


Figure 3.1: PDF of NBD for selected successes - r (Mathworks, 2015)

The plot above demonstrates the varying forms or shapes of NB distribution. Depending on the value assigned to r , the probability distribution function may appear

to be very skewed or nearly symmetric.

Although alternative NB parameterisations exist, a distinction can be determined by three factors; starting point of the support - whether at $x = 0$ or $x = r$, definition of p - whether representing probability of failure or success, and interpretation of r - whether denoting number of success or failure (DeGroot, 1986).

3.4.2 Distributional Properties

In statistics, the distributional properties, or moments, describe the primary characteristics of a distribution. The first two moments of the negative binomial distribution, mean and variance, are stated and proved as follows:

$$E(X) = \frac{r}{p} \quad (3.2)$$

and

$$var(X) = \frac{rq}{p^2} \quad (3.3)$$

where $q = 1 - p$

3.4.2.1 Proof: 1st Moment

Drawing inference from equation (3.1), the mean, which is the expectation of X can be defined as:

$$E(X) = \sum_{x=r}^{\infty} x \binom{x-1}{r-1} (1-p)^{x-r} p^r \quad (3.4)$$

Replacing $(1 - p)$ with q , and re-writing the product of x and $\binom{x-1}{r-1}$ from equation (3.4)

as:

$$x \binom{x-1}{r-1} = r \binom{x}{r} = r \binom{x+1-1}{r+1-1}$$

for computational convenience, equation (3.4) becomes further reducible as follows:

$$\begin{aligned} E(X) &= rp^r \sum_{x=r}^{\infty} \binom{x+1-1}{r+1-1} q^{x+1-(r+1)} \\ &= rp^r \sum_{\varphi=r+1}^{\infty} \binom{\varphi-1}{r+1-1} q^{\varphi-(r+1)} \\ &= rp^r \frac{1}{(1-q)^{r+1}} = \frac{rp^r}{p^{r+1}} \\ &= \frac{r}{p} \end{aligned}$$

3.4.2.2 Proof: 2nd Moment

On the other hand, variance of X is conventionally expressed by the definition:

$$\begin{aligned} \text{var}(X) &= E(X^2) - [E(X)]^2 \\ &= E[X(X+1)] - E(X) - E(X)^2 \end{aligned} \tag{3.5}$$

An examination of the terms in equation (3.5), indicates that evaluating $E[X(X+1)]$ would suffice for calculating the result for variance. As such, the expression becomes:

$$E[X(X+1)] = \sum_{x=r}^{\infty} x(x+1) \binom{x-1}{r-1} q^{x-r} p^r$$

Following a similar reduction process as used earlier, the product of $x(x+1)$ and $\binom{x-1}{r-1}$ can be re-written as:

$$x(x+1) \binom{x-1}{r-1} = r(r+1) \binom{x+1}{r+1} = r(r+1) \binom{x+2-1}{r+2-1}$$

Therefore, appropriate substitutions yield the following simplifications:

$$\begin{aligned} E[X(X+1)] &= r(r+1)p^r \sum_{x=r}^{\infty} \binom{x+2-1}{r+2-1} q^{x+2-(r+2)} \\ &= r(r+1)p^r \sum_{\varphi=r+2}^{\infty} \binom{\varphi-1}{r+2-1} q^{\varphi-(r+2)} \\ &= r(r+1)p^r \frac{1}{(1-q)^{r+2}} = \frac{r(r+1)p^r}{p^{r+2}} = \frac{r(r+1)}{p^2} \end{aligned} \quad (3.6)$$

Substituting results (3.2) and (3.6) in equation(3.5), the following conclusion is drawn:

$$\begin{aligned} var(X) &= \frac{r(r+1)}{p^2} - \frac{r}{p} - \left(\frac{r}{p}\right)^2 \\ &= \frac{r(1-p)}{p^2} = \frac{rq}{p^2} \end{aligned}$$

3.4.2.3 Results: 3rd & 4th Moments

Implementing the distribution in the Wolfram Language, on the other hand, Weisstein (n.d.) states results for the third and fourth moments, skewness (γ_1) and kurtosis excess (γ_2) as:

$$\gamma_1 = \frac{2-p}{\sqrt{rq}} \quad (3.7)$$

and

$$\gamma_2 = \frac{p^2 - 6p + 6}{rq} \quad (3.8)$$

where $q = 1 - p$

3.4.3 Estimation of Parameters

Given N independent and identically-distributed (iid) claims count observations, (k_1, \dots, k_N) , the likelihood function can be expressed as:

$$L(r, p) = \prod_{i=1}^N f(k_i; r, p) \quad (3.9)$$

The log-likelihood function can be obtained by transforming equation (3.9) as:

$$\ell(r, p) = \sum_{i=1}^N \log(\Gamma(k_i + r)) - \sum_{i=1}^N \log(k_i!) - N \log(\Gamma(r)) + \sum_{i=1}^N k_i \log(p) + Nr \log(1 - p) \quad (3.10)$$

To maximise equation (3.10), the partial derivative with respect to r and p are set to zero to derive:

$$\frac{\partial \ell(r, p)}{\partial p} = \left[\sum_{i=1}^N k_i \frac{1}{p} \right] - Nr \frac{1}{1 - p} = 0 \quad (3.11)$$

and

$$\frac{\partial \ell(r, p)}{\partial r} = \left[\sum_{i=1}^N \psi(k_i + r) \right] - N\psi(r) + N \log(1 - p) = 0 \quad (3.12)$$

where the digamma function $\psi(k) = \frac{\Gamma'(k)}{\Gamma(k)}$

Furthermore, solving for p in equation (3.11) produces:

$$p = \frac{\sum_{i=1}^N k_i}{Nr + \sum_{i=1}^N k_i} \quad (3.13)$$

Finally, substituting p in equation (3.12) yields:

$$\frac{\partial l(r, p)}{\partial r} = \left[\sum_{i=1}^N \psi(k_i + r) \right] - N\psi(r) + N \log \left(\frac{r}{r + \sum_{i=1}^N k_i / N} \right) = 0 \quad (3.14)$$

3.5 Discrete Generalised Pareto Distribution

The DGP distribution arises from a discretisation of the continuous Generalised Pareto distribution. However, relevant elementary forms of the Pareto distribution are briefly revisited, to provide a basis for further discussions on the DGP.

On one hand, Pareto Type-I has a distribution function of the form:

$$D(x) = 1 - \left(\frac{b}{x} \right)^a \quad (3.15)$$

defined over the interval $x \geq b$

On the other, Generalised Pareto is defined by the function:

$$G(x) = 1 - \left[1 + \frac{\alpha(x - \mu)}{\lambda} \right]^{-\frac{1}{\alpha}} \quad (3.16)$$

for $1 + \alpha(x - \mu)/\lambda > 0$ and $x > \mu$, where $\lambda > 0$

The Generalised Pareto distribution provides important tail modelling capabilities, inherited by DGP into the discrete probability space. As stated in previous sections, the study follows MLE in fitting the DGP distribution through statistical programming.

3.5.1 Theoretical Definition

In view of the elementary Pareto forms, this section presents a formal deduction of the pmf-based definition of DGP. First, the discrete generalised pareto distribution is expressed by the cumulative distribution function (cdf) as:

$$F(x) = Pr(X \leq x) = 1 - [1 + \lambda(x - \mu + 1)]^{-\alpha} \quad (3.17)$$

where $x = \mu, \mu + 1, \dots$, and $F(x) = 0$ if $x < \mu$

The parameters α, λ , and μ defined above, are strictly positive. However, the \mathcal{DGP} emanates from a discretisation of the continuous \mathcal{GPD} through a number of scholarly approaches. In particular, Krishna and Pundir (2008) addressed the discretisation of a continuous model by observing unit groupings on the failure time axis. The authors reasoned that for a continuous failure time X , with survival function $S(x) = P[X \geq x]$, and time groupings of intervals $dX = \lfloor X \rfloor$, the floor of X , the discrete observed variable, dX , would have the probability mass function (pmf):

$$p(x) = P[dX = x] = P[x \leq X < x + 1] = S(x) - S(x + 1) \quad (3.18)$$

for $x = 0, 1, 2, \dots$

Given the conventional survival expression historically proposed by (Xekalaki, 1983):

$$S(x) = Pr(X \geq x) = 1 - F(x - 1) \quad (3.19)$$

the survival function of the \mathcal{DGP} distribution can be further derived as:

$$S(x) = [1 + \lambda(x - \mu + 1)]^{-\alpha} \quad (3.20)$$

for $x = \mu, \mu + 1, \dots$

Therefore, with further determination of $S(x + 1)$, the \mathcal{DGP} yields a pmf expressed as:

$$Pr(X = x) = [1 + \lambda(x - \mu)]^{-\alpha} - [1 + \lambda(x - \mu + 1)]^{-\alpha} \quad (3.21)$$

for $x = \mu, \mu + 1, \dots$

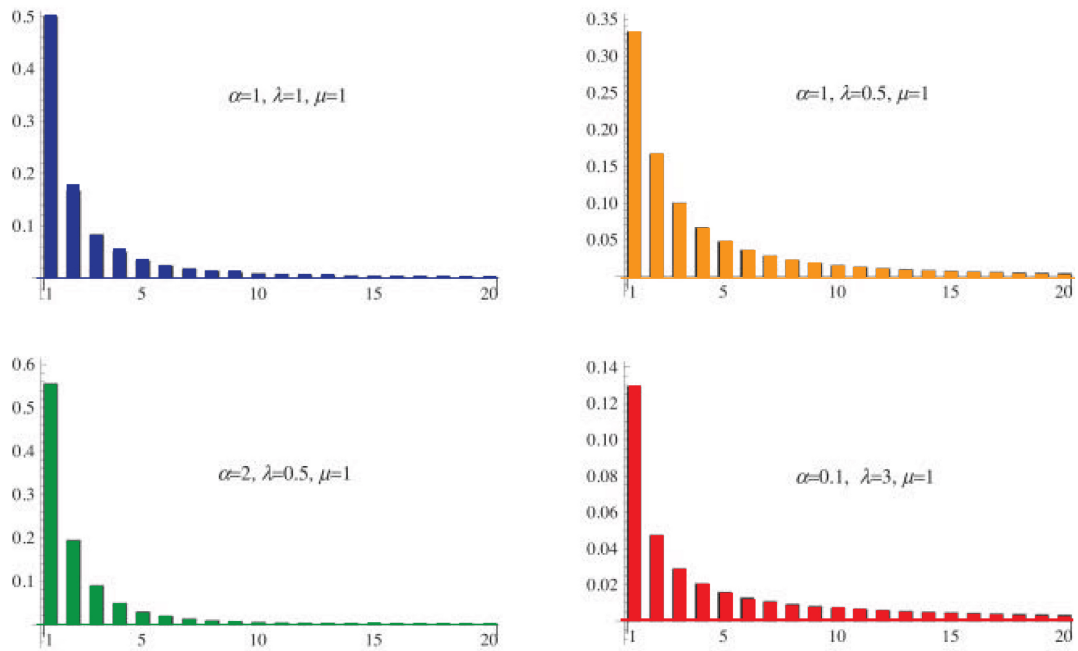


Figure 3.2: **Pmf of DGP for selected α, λ, μ** - Prieto et al. (2014)

It is worthy of note, that certain critical values of the GPD parameters, results in special distributions of statistical interest. In particular, the discrete Lomax distribution \mathcal{DLo} , a special two-parameter case of the DGP , results when the location parameter, $\mu = 0$. Obtained by a discretisation of the continuous Lomax distribution, the \mathcal{DLo} shares an equivalence relationship with the DGP . Thus $DGP(\alpha, \lambda, \mu) \equiv \mathcal{DLo}(\alpha, \lambda)$.

3.5.2 Distributional Properties

A deduction for r th-order moment, $E(X^r)$, of the DGP distribution is presented by Prieto et al., (2014) in the following expressions:

$$E(X^r) = \sum_{x=\mu}^{\infty} X^r P_r(X = x) = \sum_{x=\mu+1}^{\infty} [x^r - (x-1)^r] \bar{F}(x)$$

where $\bar{F}(x)$ is the survival function previously stated at equation (3.19)

$$\Rightarrow E(X^r) = \sum_{x=\mu+1}^{\infty} \frac{x^r - (x-1)^r}{[1 + \lambda(x - \mu)]^\alpha} \quad (3.22)$$

Thus, implicitly, the mean and second moments of the \mathcal{DGP} distribution are respectively deducible from equation (3.22) as:

$$E(X) = \sum_{x=\mu+1}^{\infty} \frac{1}{[1 + \lambda(x - \mu)]^\alpha} \quad (3.23)$$

and

$$E(X^2) = \sum_{x=\mu+1}^{\infty} \frac{2x - 1}{[1 + \lambda(x - \mu)]^\alpha} \quad (3.24)$$

Two observations are apparent from equations (3.23) and (3.24). First, if $\alpha > 1$ and $\alpha > 2$, then $E(X)$ and $E(X^2)$ are respectively finite. Second, a closer inspection shows that the mean decreases with both distributional parameters, α and λ , since $\frac{\partial E(x)}{\partial \alpha} < 0$ and $\frac{\partial E(x)}{\partial \lambda} < 0$.

3.5.3 Estimation of Parameters

Suppose x_i, \dots, x_n is a DGP distributional sample of size n . The parameters α and λ are estimated on the assumption that μ is known, and that $\hat{\mu} = x_{min} \cdot x_{min}$ is the sample minimum, denoting the smallest observation, such that $x_{min} \leq x_i, \forall i$. In adopting the μ -frequency and $(\mu+1)$ -frequency method proposed by Prieto et al. (2014), initial values, $(\alpha_0, \lambda_0, \mu_0)$ can be obtained and used as estimators in the subsequent maximum likelihood operation. As a result, the relative frequencies of $X = \mu$, and $X = (\mu + 1)$, respectively denoted by \hat{f}_μ and $\hat{f}_{\mu+1}$, are calculated from the sample data. Analogously, $X = \mu$ and $X = (\mu + 1)$ are substituted into the DGP pmf in equation (3.21), and the corresponding pmf outputs are equated to the respective \hat{f}_μ and $\hat{f}_{\mu+1}$ values, after which α and λ can be determined simultaneously.

However, the μ -frequency and $(\mu+1)$ -frequency method operates on the assumption that data used is observed in increasing steps of one. On the contrary, the count data under study appears in varying intervals. As such, applying the method strictly on the count data available will result in generating $\mu + 1 = 0$, leading to $\hat{f}_{\mu+1} = 0$ for most years of the reported and settled claims count data. In that regard, proceeding the computations with zero relative frequencies, will result in a loss of essential frequency information from the dataset. Therefore, in this study, the μ - and $(\mu+1)$ -frequency method is modified as μ and $(\mu + \epsilon)$; where $\epsilon > 0$ is the increment from the minimum, μ , to the next observation. As such, $(\mu + \epsilon)$ is the smallest observation greater than μ . In effect, α and λ can be obtained by solving the resulting expressions, (3.25) and (3.26) simultaneously.

$$\hat{f}_{\mu} = 1 - [1 + \lambda]^{-\alpha} \quad (3.25)$$

$$\hat{f}_{\mu+\epsilon} = [1 + \lambda]^{-\alpha} - [1 + 2\lambda]^{-\alpha} \quad (3.26)$$

The expression in (3.27) results, after α is eliminated in equations (3.25) and (3.26):

$$\frac{\log(1 + 2\lambda)}{\log(1 + \lambda)} = \frac{\log(1 - \hat{f}_{\mu} - \hat{f}_{\mu+\epsilon})}{\log(1 - \hat{f}_{\mu})} \quad (3.27)$$

The equation is monotonic in λ , and can be solved in \mathbb{R} , with a defined *uniroot* function.

Subsequently, equation (3.28) is obtained as follows, by appropriate substitutions into (3.27).

$$\hat{\alpha} = -\frac{\log(1 - \hat{f}_{\mu})}{\log(1 + \hat{\lambda})} \quad (3.28)$$

Next, the maximum likelihood method is employed to conclude the DGP parameter estimations. As such, the log-likelihood function can be constructed as:

$$\log \ell(\lambda, \alpha) = \sum_{i=1}^n \log Pr(X = x_i) = \sum_{i=1}^n \log [(1 + \lambda(x_i - \mu))^{-\alpha} - (1 + \lambda(x_i - \mu + 1))^{-\alpha}] \quad (3.29)$$

The partial derivatives of equation (3.29) are taken with respect to α and λ , and set to zero; noting that $Pr(X = x_i)$ refers to the pmf specified in (3.21). The resulting stationary equations are:

$$\begin{aligned} \frac{\partial \log \ell}{\partial \alpha} &= \sum_{i=1}^n \frac{\log [1 + \lambda(x_i - \mu + 1)]}{[1 + \lambda(x_i - \mu + 1)]^\alpha [1 + \lambda(x_i - \mu)]^{-\alpha} - 1} \\ &\quad - \sum_{i=1}^n \frac{\log [1 + \lambda(x_i - \mu)]}{1 - [1 + \lambda(x_i - \mu)]^\alpha [1 + \lambda(x_i - \mu + 1)]^{-\alpha}} = 0 \end{aligned} \quad (3.30)$$

$$\frac{\partial \log \ell}{\partial \lambda} = \sum_{i=1}^n \frac{\alpha(x_i - \mu + 1)}{[1 + \lambda(x_i - \mu + 1)]^{\alpha+1} [1 + \lambda(x_i - \mu)]^{-\alpha} - [1 + \lambda(x_i - \mu + 1)]}$$

$$- \sum_{i=1}^n \frac{\alpha(x_i - \mu)}{[1 + \lambda(x_i - \mu)] - [1 + \lambda(x_i - \mu)]^{\alpha+1} [1 + \lambda(x_i - \mu + 1)]^{-\alpha}} = 0 \quad (3.31)$$

3.5.4 Estimation of Standard Errors

From the above equations, it will be difficult to obtain explicit expressions for the estimators, and hence numerical methods are used. As a result, the standard errors of the estimators are obtained using the bootstrap resampling technique (Effron and Tibshirani, 1993). First, suppose there is a need to draw inferences about some statistic, $\hat{\theta} = s(\mathbf{x})$, given $\mathbf{x} = (x_1, \dots, x_n)'$, for which $x_i \stackrel{i.i.d}{\sim} F(x), \forall i \in \{1, \dots, n\}$. Then, adopting the Monte Carlo bootstrap procedure, a resampling algorithm will be written

to perform the following operations:

1. Sample x_i^* from $\{x_1, \dots, x_n\}, \forall i \in \{1, \dots, n\}$, with replacement.
2. Compute $\hat{\theta}^* = s(\mathbf{x}^*)$ for k -th sample, with $\mathbf{x}^* = (x_1^*, \dots, x_n^*)'$.
3. Repeat steps 1 and 2, K times, to obtain a bootstrap distribution of $\hat{\theta}$.
4. Compare $\hat{\theta}^* = s(\mathbf{x})$ to the derived bootstrap distribution.

Finally, the standard error of $\hat{\theta}$ is evaluated as standard deviation of $\{\hat{\theta}_k^*\}_{k=1}^K$ given by:

$$\hat{\sigma}_K = \sqrt{\frac{1}{K-1} \sum_{k=1}^K (\hat{\theta}_k^* - \bar{\theta}^*)^2} \quad (3.32)$$

where $\bar{\theta}^* = \frac{1}{K} \sum_{k=1}^K (\hat{\theta}_k^*)$ represents mean of the bootstrap distribution of $\hat{\theta}$.

3.6 Model Selection Criteria

In statistics, a model refers to a mathematical summary of the relationship among properties of measurements (Vrieze, 2013). Thus, models describe the behaviour of variables which influence a given phenomenon. In data studies, model selection involves actions which guide the choice of the best fitted option from a set of candidate models. Besides its usefulness in identifying a single best hypothesis, the selection provides a basis for drawing inferences among competing models using weighted support (Johnson & Omland, 2004). According to Occam's razor, which has been critiqued and justified by various authorities, the simplest model, bearing the fewest assumptions, presents the likeliest best choice. However, several statistical benchmarks, collectively referred to as selection criteria, have been proposed to provide scientific basis for model adoption. Possessing unique strengths and weaknesses, each selection criterion serves as a measure of model suitability necessary for informed decision making. Considering the scope of work, the researcher employs

Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) for model comparison as discussed in the following section.

3.6.1 Akaike Information Criterion

The Akaike Information Criterion (AIC) was first introduced by Hirotogu Akaike and formally published in 1974 (Akaike, 1974). Originally, the statistician referred to the AIC as an Entropy Maximisation Principle, since the approach is based on concepts underlying entropy in information theory. Given a set of candidate models, AIC measures the quality of individual models in relation to each of the alternatives available for consideration. In striving for balance between goodness-of-fit and simplicity of models, the criterion provides an estimate for the relative information lost, due to the usage of a particular model to fit a given data. The AIC is defined by Akaike (1974) as the following expression:

$$AIC = -2\log L + 2d \quad (3.33)$$

In equation (3.33) above, $2d$ measures the goodness of model fit, where d refers to the number of parameters estimated. On the other hand, $\log L = \log l(\hat{\theta}|x)$ is the log-likelihood of the model, evaluated at the maximum likelihood estimates. Thus, L is the maximum value assumed by the likelihood function associated with the model. The term, $-2\log L$, serves as a penalty function to discourage overfitting, since AIC generally enhances goodness of fit arising from increasing number of parameters, d . As such, the model which outputs the minimum AIC value is statistically preferred and selected.

3.6.2 Bayesian Information Criterion

The Bayesian Information Criterion (BIC) was formulated by Gideon E. Schwarz and published in 1978 (Schwarz, 1978). Otherwise referred to as Schwarz Bayesian Criterion (SBC), it presents an alternative means of model selection among a finite set of candidate options. BIC offers an approximation which is asymptotic to a transformed Bayesian posterior probability of a candidate model (Neath & Cavanaugh, 2011). Schwarz's (1978) BIC is stated as:

$$BIC = -2\log L + d\log(n) \quad (3.34)$$

In equation (3.34) above, n denotes the sample size, length of time series, or size of observations, while d , L and $\log L$ maintain the definitions assigned to them under the AIC statistic. In comparison with AIC, BIC addresses the issue of overfitting with a factor, $\log(n)$, thereby posing a higher penalty to model complexity (Dziak, Coffman, Lanza, & Li, 2012). In statistical decision making, a candidate model with minimum BIC value is selected.

3.7 Statistical Computing Environment

The study uses R in performing relevant analysis of the datasets under study. In particular, libraries such as *fitdistrplus* and *MASS* in the R environment, offer an integrated set of tools which facilitate model fitting to research data. These tools will enable us to perform effective distributional comparisons, deemed relevant to the focus of study.

CHAPTER 4

DATA ANALYSIS AND DISCUSSIONS

This chapter provides a detailed analysis of data considered for the study. Following the methods discussed in chapter three, chapter four demonstrates the fitting of NB and DGP distributions to non-life insurance claim counts data. The chapter is divided into four sections as follows. Section 4.1 introduces the dataset utilised, by describing the aspects of relevance to the study. In section 4.2, the researcher performs the model fitting exercise, driven by the data described in section 4.1. Furthermore, section 4.3 compares the relative distributional results obtained from section 4.2. The chapter concludes with a summary of key statistical deductions presented in section 4.4.

4.1 Description of Data

As briefly noted in chapter three, the data utilised in this study is obtained from the National Insurance Commission (NIC). The data indicates the observed claims counts of twenty-nine non-life insurance service providers from 2012 to 2016. Within the five-year period, the data covers claims categories across the five business classes of non-life insurance in Ghana. The classes are Fire, Burglary, and Property Damage; Accident; Marine and Aviation; Motor; and General Liability. For each fiscal year, the dataset discloses the total number of claims administered per licensed operator. The data is further described in terms of its composition, basic statistics, and limitations, in the subsections below.

4.1.1 Composition of Data

The non-life claims data is organised into three categories on the basis of clearance status. Clearance status refers to the state of claims payments, recorded at a particular time in the compensation process. In order of claims administration, the categories are Incurred But Not Reported (IBNR), Reported But Not settled (RBNS), and Settled But Outstanding (SEBO). In accordance with NIC's interpretations, the indicators are explained in the following paragraphs.

First, IBNR refers to claims arising from the occurrence of insured but unreported events. This is an actuarial provision made by insurers. Thus, upon the occurrence of an insured event, reports by victims may not be immediate. As such, not all claims will be requested for at the time of loss. Therefore, insurers assume the existence uncommunicated losses, which are likely to be reported in the future. This enables insurers to reserve funds for offsetting such provisions when they arise.

Second, RBNS refers to claims emerging from formally communicated events with pending valuation of losses. Cases of RBNS imply that insurers have received applications for claims by insured individuals, but yet to validate the proximate causes as grounds for processing of compensations. According to the Chartered Institute of Loss Adjusters [CILA] (2016), the case of *Pawsey v Scottish Union & National Insurance Company* (1908) led to the definition of proximate cause as the active and efficient cause that sets in motion a train of events which brings about a result, without the intervention of any force started and working actively from a new and independent source. Thus, proximate cause is an important insurance principle, verified by insurers upon receipt of event reports, for the purposes of justifying claims issuances.

Third, SEBO refers to claims associated with events and losses which have been verified and valued respectively. In practice, a claim is regarded as settled when an insurance company accepts liability and concludes on an amount payable to the policy holder. The settlements are qualified with the word *outstanding* to indicate

that, although valuation of liability has occurred, issuance of cheques to actually pay the claims may be pending. However, given that the settlement phase is technically considered as final of the compensation process, SEBO amounts are implicitly treated as payout values.

The difference between the IBNR and RBNS claims lies in the basis of counting. While IBNR values are projections based on varying actuarial assumptions, RBNS refers to the actual number of claims based on real events recorded. Overall, the dataset considered in this work consists of 3,878,355 non-life insurance policies; generating 39,563 reported claims, of which 5,210 claims were settled from 2012 to 2016.

4.1.2 Descriptive statistics

Given the above composition, a statistical summary of the dataset will provide useful insights to aid in analysis. In this regard, the mean, standard deviation, skewness, and kurtosis of the dataset are measured. Additionally, the yearly growth levels of the total number of policy covers, reported claims, and settled claims are worth discussing. Aided by graphs and tables, the stated features of the dataset are described in this subsection.

4.1.2.1 Policy Subscriptions Count

The graph in Figure 4.1 shows the annual totals of non-life policy subscriptions recorded by insurers. The number policies grew from 666,752 in 2012 to 869,544 in 2013, representing an increase of 30.02%. In the 2014 fiscal period, total policy covers moved to 951,620, resulting in a growth rate reduction to 9.44% relative to the previous year. Decreasing by 15.85% in 2015, the policy counts recovered slightly by 2.42% in 2016, following a movement from 800,764 in the former year to 820,180 in the latter year. In effect, the non-life insurance sector recorded an average annual growth (AAG) of 6.51% in policy subscriptions from 2012 to 2016.

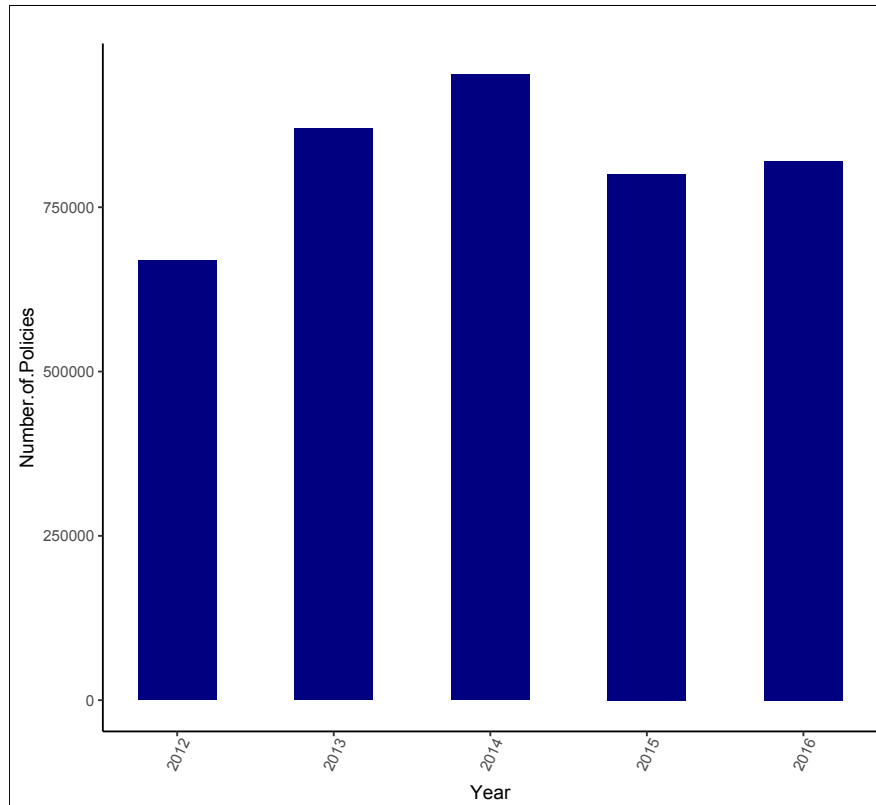


Figure 4.1: **Number of non-life policies recorded in the period 2012-2016**

On the other hand, a statistical summary of dataset, on number of policy subscriptions, is displayed in Table 4.1.

Table 4.1: **Descriptive summary of Policy Subscriptions Count dataset**

Year	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max
2012	23060.62	32224.08	2.5	7.24	0	155376
2013	29984.28	40565.08	1.43	0.86	0	141742
2014	32814.47	52232.65	3	10.08	0	262057
2015	27612.55	32281.07	1.45	1.43	0	120966
2016	28282.08	28366.85	1.18	0.69	2663	109506

From 2012 to 2016, the annual mean of the number of non-life policy subscriptions moved from 23,060.62 to 28,282.08, with the highest average of 32,814.47 recorded in 2014. During the period, the standard deviation or dispersion from mean, moved from 32,224.08 to 28,366.85. In 2014, the policy subscriptions count data exhibited the highest variation from expectation with a standard deviation of 52,232.65.

Given the positive skewness values across the years, the data can be regarded as positively skewed. In terms of symmetry, this means that the right tail of the distribution is longer than the left. In accordance with Bulmer's (1979) rule of thumb, the distribution is highly skewed since the associated values are the greater than 1.

With reference to tailed-ness of the data relative to a normal distribution, the distributions for the 2013, 2015, and 2016 policy subscriptions data are platykurtic, since their computed values for kurtosis are less than 3. Thus, in comparison with a normal distribution, the tails are shorter and thinner, often with lower and broader central peaks. On the other hand, the distributions for the 2012 and 2014 data are leptokurtic, since their computed values for kurtosis are greater than 3. Thus, in comparison with a normal distribution, the tails are longer and thicker, often with higher and sharper central peak. Prior to 2016, the minimum count observed was 0, due to the non-existence of some insurers in the 2012-2015 period. However, the maximum count observed grew from 155,376 in 2012 to a peak of 262,057 in 2014. The years 2015 and 2016 reported maximum count values of 120,966 and 109,506 respectively; both of which represent a decline from the 2014 figure.

4.1.2.2 Reported Claims Count

The graph in Figure 4.2 indicates the total number of claims reported by insureds. From 2012 to 2013, the count of reported claims grew from 7,420 to 12,032 in 2013, representing an increase of 62.16%. In 2014, the number fell to 8,175, resulting in a rate of -32.06% growth with reference to the previous year. Increasing by 21.22% in 2015, the figure rose significantly by 97.99% in 2016, corresponding to a movement from 9,910 to 19,621 of the respective years. Overall, the non-life insurance sector averaged a 37.33% annual growth of reported insurance claims, during the five-year time horizon under consideration.

Furthermore, a statistical summary of dataset on number of reported claims, is provided in Table 4.2.

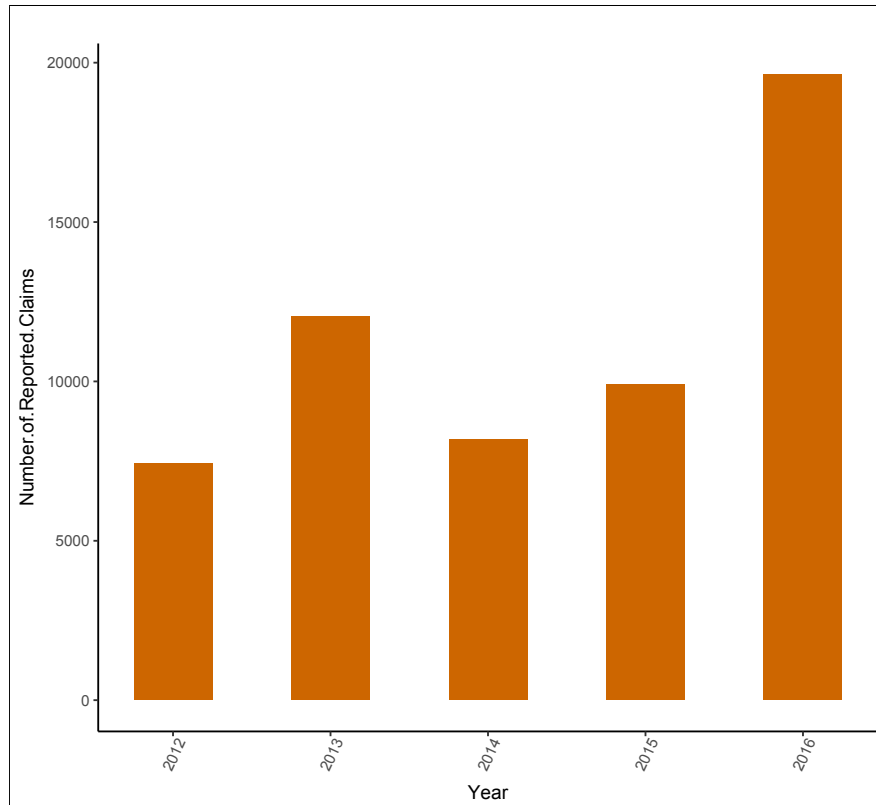


Figure 4.2: **Number of non-life claims reported in the period 2012-2016**

Table 4.2: **Descriptive summary of Reported Claims Count dataset**

Year	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max
2012	255.86	329.95	1.28	0.71	0	1187
2013	414.9	620.11	1.58	1.12	0	2053
2014	281.9	350.03	1.57	1.47	0	1335
2015	341.72	503.77	2.12	4.07	0	2151
2016	676.59	1839.98	4.34	18.81	2	9902

From 2012 to 2016, the annual mean of the number of non-life reported claims moved from 255.86 to 676.59, with the latter representing the highest average across the years. In the first year, the standard deviation or dispersion from mean, was 329.95. In the fifth year, the reported claims count data exhibited the highest variation from expectation with a standard deviation of 1839.98.

Given the positive skewness values across the years, the data can be regarded as positively skewed. In terms of symmetry, this means that the right tail of the

distribution is longer than the left. In accordance with Bulmer's (1979) rule of thumb, the distribution is highly skewed since the associated values are the greater than 1.

With reference to tailed-ness of the data relative to a normal distribution, the distributions for the 2012, 2013, and 2014 reported claims data are platykurtic, since their computed values for kurtosis are less than 3. Thus, in comparison with a normal distribution, the tails are shorter and thinner, often with lower and broader central peaks. On the other hand, the distributions for the 2015 and 2016 data are leptokurtic, since their computed values for kurtosis are greater than 3. Thus, in comparison with a normal distribution, the tails are longer and thicker, often with higher and sharper central peaks.

Prior to 2016, the minimum count observed was 0, due to the non-existence of some insurers in the 2012-2015 period. However, the maximum count observed grew from 1,187 in 2012 to a peak of 9,902 in 2016.

4.1.2.3 Settled Claims Count

The annual totals of claims settled are displayed in Figure 4.3. Moving from 2,297 to 1,341, the number of claim settlements declined by 41.62% between 2012 and 2013. However, the count rose steadily to 1,567 in 2014, and 1,865 in 2015, corresponding to a growth of 16.85% and 19.02% in the respective years. As at the 2016 fiscal year-end, the number had risen to 2,128, representing a 14.10% rise in the count of settled claims. Thus, the non-life sector experienced a 2.09% AAG in insurance claims settlement during the period 2012-2016.

Additionally, Table 4.3 presents a statistical summary of dataset on number of settled claims.

From 2012 to 2016, the annual mean of the number of non-life settled claims moved from 79.21 to 73.38, with the former representing the highest average across the years. During the period, the standard deviation or dispersion from mean, moved from 123.42

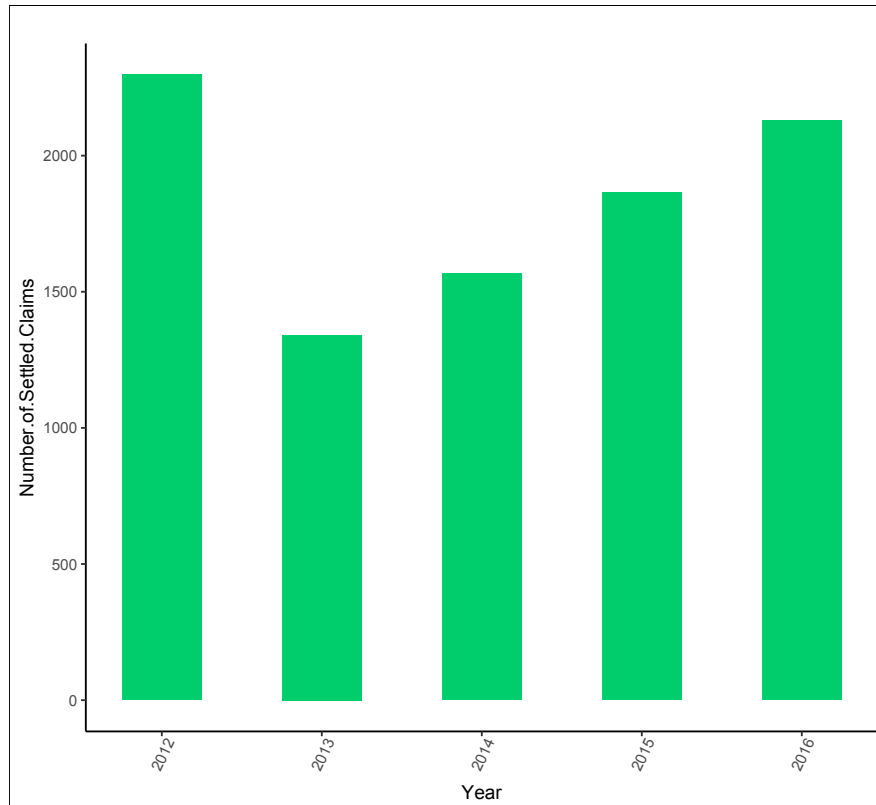


Figure 4.3: **Number of non-life claims settled in the period 2012-2016**

Table 4.3: **Descriptive summary of Settled Claims Count dataset**

Year	Mean	Std. Dev.	Skewness	Kurtosis	Min	Max
2012	79.21	123.42	1.85	2.34	0	452
2013	46.24	63.22	2.52	7.35	0	307
2014	54.03	65.38	2.24	5.42	0	307
2015	64.31	90.59	1.93	3.03	0	362
2016	73.38	106.75	2.19	4.44	2	452

to 106.75. In 2013, the settled claims count data exhibited the lowest variation from expectation with a standard deviation of 63.22.

Given the positive skewness values across the years, the data can be regarded as positively skewed. In terms of symmetry, this means that the right tail of the distribution is longer than the left. In accordance with Bulmer's (1979) rule of thumb, the distribution is highly skewed since the associated values are the greater than 1.

With reference to tailed-ness of the data relative to a normal distribution, the

distribution for the 2012 settled claims data is platykurtic, since its computed value for kurtosis is less than 3. Thus, in comparison with a normal distribution, the tails are shorter and thinner, often with lower and broader central peaks. On the other hand, the distributions for the 2013, 2014, 2015, and 2016 data are leptokurtic, since their computed values for kurtosis are greater than 3. Thus, in comparison with a normal distribution, the tails are longer and thicker, often with higher and sharper central peaks.

Prior to 2016, the minimum settled claims count observed was 0, due to the non-existence of some insurers in the 2012-2015 period. However, the highest maximum count of 452 occurred in 2012 and 2016, with 2013 and 2014 recording the lowest value of 307.

In summary, the annual counts for policy subscriptions, reported claims, and settled claims grew by 6.51%, 37.33%, and 2.09% respectively. This implies that the number of unsettled claims rose by approximately 35.24% during the period. Thus, the number unsettled claims rose at a rate which is 16.86 times faster than the growth of settled claims. This observation further justifies the need for appropriate statistical modelling, to improve the future counts of claim settlements by non-life service providers.

4.1.3 Limitations of Data

The data obtained for the study, provides a record of disaggregated claim counts reported by non-life insurers. However, particular states of the industry may have limited the coverage of data collated by the regulator. These are summarised as follows:

First, the time horizon covered by the data is five years. Although the researcher sought to obtain data on ten years of claims, to enhance the comprehensiveness of analysis, records for additional five years were challenging to access at the NIC. This is because, NIC commenced data digitisation in 2012, with the most current update recorded in

2016. Thus, data on claims for the years preceding 2012, are currently not available in readily usable forms.

Second, some companies were either non-existent or not fully functional during the years under consideration. Thus, between 2012 and 2016, certain presently operational companies, did not possess licenses to undertake commercial insurance activities. While such institutions are important market players in present times, their policy claims records are not captured in the study data.

Third, some insurers were acquired by larger insurance groups within the five-year period. Such actions resulted in three interrelated developments. These are, cessation of business by the smaller entity, transfer of ownership to the larger entity, and corporate restructuring of the acquiring party to accommodate the change. Such processes may influence the composition of a given dataset, due to the simultaneous dissolution and consolidation of entities. However, the study claims data, when considered in isolation, does not reflect such noteworthy developments.

In summary, the limitations of the research data border on inaccessibility to a longer timeframe, non-operational state of particular firms, and institutional restructuring of certain entities, during the period of coverage. While these issues are vital, they may not exert a significant influence on expected outcomes, given the specific objectives of the study.

4.2 Model Fitting

Considering the description above, two sets of discrete data are modelled in this section of the study. They are, number of claims reported (RBNS), and number of claims settled (SEBO), arising from the total volume of policies recorded in the period 2012 - 2016. As discussed earlier in the study, determination of NB and DGP distributional parameters will follow the MLE approach. The computational estimations in R will be based on the underlying closed form operations presented in the chapter three.

In the subsections ahead, the modelling will involve estimation of the NB and DGP parameters, and comparison of models.

4.2.1 Estimation of NB Parameters

In section 3.4.3 the theoretical estimation of NB parameters was discussed. To proceed with real claims count data, the *fitdist* function in the *fitdistrplus* package in R are used. Initially, the theoretical N iid claims count observations, (k_1, \dots, k_N) , are replaced by the actual counts for each of the five years. Thus, each year would have distinct $k_{1 \leq i \leq N}$'s under both count categories - Reported Claims and Settled Claims. This results in ten sets of k_i 's, organised and labelled in Table 4.4.

Table 4.4: k_i labels for Reported and Settled Claims Count

Reported Counts k_i 's	Settled Counts k_i 's
ReportedClaimsCount2012	SettledClaimsCount2012
ReportedClaimsCount2013	SettledClaimsCount2013
ReportedClaimsCount2014	SettledClaimsCount2014
ReportedClaimsCount2015	SettledClaimsCount2015
ReportedClaimsCount2016	SettledClaimsCount2016

Applying the appropriate functional R arguments to the respective count divisions in Table 4.4, NB parametric estimates and standard errors in Table 4.5 are obtained.

In Table 4.5, the NB parameters are estimated using an alternative parameterisation given by $X \sim \mathcal{NB}(r, m/(m+r))$; where \hat{m} and \hat{r} represent the mean and dispersion parameter respectively. The standard errors are placed in parenthesis.

4.2.2 Estimation of DGP Parameters

The relative frequencies calculated from the Reported and Settled Claims counts data are given in Table 4.6.

Table 4.5: **Parameter estimates from Negative Binomial model**

Variable	Parameters	2012	2013	2014	2015	2016
Reported Claims	\hat{r}	0.4915 (0.1234)	0.6874 (0.1661)	1.0370 (0.2605)	0.8603 (0.2121)	0.4518 (0.1052)
	\hat{m}	296.8624 (84.7788)	542.4572 (130.9043)	327.0975 (64.3623)	396.9273 (85.7927)	783.7698 (233.2420)
Settled Claims	\hat{r}	0.4225 (0.1063)	0.6735 (0.1764)	1.1312 (0.3002)	0.7453 (0.1878)	0.7815 (0.1912)
	\hat{m}	88.3519 (26.7214)	51.5716 (12.4032)	60.2751 (11.2194)	71.7227 (16.3761)	81.8596 (18.2495)

Table 4.6: **Relative frequencies of Claims Count categories**

	Reported Claims Count	Settled Claims Count
\hat{f}_{μ}	0.9756	0.9756
$\hat{f}_{\mu+\epsilon}$	0.9837	0.9837

Following the operations in section 3.5.3, and values provided in Table 4.6, the initial estimators for μ , α , and λ , is $(\mu_0, \alpha_0, \lambda_0) = (2, 0.0172, 1.6)$.

To proceed, the R functions in Appendix C.3 were used to obtain the DGP parameters. Sequentially, the algorithms sought to:

1. Specify the log-likelihood function (equation 3.29) in R, based on the DGP probability mass function (equation 3.21). The log-likelihood function is set to return a negation of the log-likelihood value, since the R *optim* function is a minimiser. As such, minimising the negated log-likelihood function at the initial estimates, produces the equivalent of maximising the log-likelihood value.
2. Optimise the log-likelihood function in (1) at the seed values, by Simulated Annealing (SANN), a variant of Belisle (1992). The optimisation function follows the procedure outlined in section, and solves equations (3.30) and (3.31). In addition to SANN, four convex optimisation methods were employed

- Conjugate Grading (CG) based on Fletcher and Reeves (1964); Brent; Nelder Mead (Nelder and Mead, 1965); and Broyden, Fletcher, Goldfarb and Shanno [BFGS] algorithms. In addition, the BFGS algorithm was considered with its extensions - Limited Memory BFGS (L-BFGS) and Limited Memory Boxed BFGS [L-BFGS-B] (Byrd et. al, 1995). While the former requires lower memory resources, the latter extends L-BFGS and imposes box constraints on the method. SANN was selected among the five because during the algorithmic run-time, the optimisation cycle does not terminate in cases of non-existent estimates at the initial parameter values. In such instances, the next candidate point is generated by simulation, and executed concurrently.

3. Extract the estimated parameters, α and λ from the output generated by (2). Among other results, the returned output contains the convergence and Hessian matrix. The latter is a symmetric matrix which estimates the Hessian at the solution found. After convergence of the algorithm, standard errors of α , λ , and μ can be derived by square-rooting the diagonal entries of the inverse hessian matrix. However, the limitation in this approach of obtaining standard errors, is evident in cases of negative diagonal entries. In such instances, the standard errors will be undefined. To address the difficulty, an alternative approach is followed in (4).
4. Iteratively compute the standard errors of estimates based on bootstrap resampling of 1000 replicates. Following the process outlined in section 3.5.4, functions written to perform the bootstrap resampling technique is shown in Appendix C.3.3.
5. Compute the AIC and BIC statistics for model selection based on functions specified in Appendix C.3.4.

Applying the five-stage process in section 4.2.2 on each count dataset across the five periods, the DGP maximum likelihood estimates and standard errors in Table 4.7 are obtained.

Table 4.7: Parameter estimates from Discrete Generalised Pareto model

Variable	Parameters	2012	2013	2014	2015	2016
Reported Claims	$\hat{\mu}$	7.3983 (3.4175)	4.8507 (3.9254)	5.8075 (3.8876)	6.2320 (5.6555)	2.2437 (2.2000)
	$\hat{\alpha}$	0.0128 (0.0104)	0.0044 (0.0011)	0.0020 (0.0019)	0.0013 (0.0010)	0.0102 (0.0017)
	$\hat{\lambda}$	3.3700 (0.1318)	2.7590 (0.8976)	2.8036 (0.0106)	2.4566 (0.0829)	4.9025 (2.3947)
Settled Claims	$\hat{\mu}$	3.5866 (1.0333)	3.3776 (0.1015)	4.6485 (3.4941)	3.6443 (0.6387)	3.0205 (2.5885)
	$\hat{\alpha}$	3.4157 (0.0532)	2.9258 (0.0200)	2.0675 (0.0235)	2.3356 (0.0064)	0.1121 (0.0311)
	$\hat{\lambda}$	3.4015 (0.9538)	9.4855 (3.0089)	3.1532 (0.2831)	3.1127 (1.5265)	2.8593 (0.9741)

In Table 4.7, $\hat{\mu}$, $\hat{\alpha}$, and $\hat{\lambda}$ represent the estimated location, shape, and scale parameters respectively. The bootstrap standard errors are placed in parenthesis.

4.2.3 Comparison of Models

As discussed in Chapter three of the study, comparison of the models will be performed, to aid selection based on the Akaike and Bayesian information criteria. Accordingly, smaller values of the AIC and/or BIC statistic are indicators of better fitted models. Table 4.8 and 4.9 display values of the AIC and BIC from the NB and DGP distributions fitted to data on reported and settled claims counts.

In terms of reported claims count, DGP model presents smaller AIC and BIC values, in comparison with NB model. This observation is consistent across the entire fiscal years under consideration. Therefore, on the basis of AIC and BIC, DGP model is preferred to NB model in the case of counts of reported claims. Regarding settled claims count, DGP exhibits smaller values of AIC and BIC, in the five-year period

Table 4.8: **AIC statistics for NB and DGP model**

Model	Criterion	2012	2013	2014	2015	2016
Reported Claims						
NB model	<i>AIC</i>	329.0927	366.1941	343.5535	352.7973	372.5812
DGP model	<i>AIC</i>	55.9362	55.9999	55.9998	55.9999	55.9972
Settled Claims						
NB model	<i>AIC</i>	274.866	259.0694	269.3573	277.0937	284.2884
DGP model	<i>AIC</i>	57.9813	57.9958	57.9476	57.9653	57.8863

Table 4.9: **BIC statistics for NB and DGP model**

Model	Criterion	2012	2013	2014	2015	2016
Reported Claims						
NB model	<i>BIC</i>	331.5304	368.6319	345.9913	355.2351	375.0189
DGP model	<i>BIC</i>	57.1299	57.1936	57.1935	57.1936	57.1909
Settled Claims						
NB model	<i>BIC</i>	277.3822	261.5856	271.8735	279.6099	286.8046
DGP model	<i>BIC</i>	59.1750	59.1895	59.1413	59.1590	59.0800

under study. Thus, NB model is associated with larger AIC and BIC values across the time horizon of the count data. Therefore, the AIC and BIC consider the DGP as the better fitted model, in the case of counts of settled claims. In effect, the DGP model is recommendable, as it provides a more reliable fit for count data on both classes of non-life insurance claims.

4.3 Aggregation of Data

In the previous sections, annual observed counts data for reported and settled claims were modelled independently. As such, five sets of NB and DGP parameter estimates were computed for each year - from 2012 to 2016. Final comparison of the fitted models considered the AIC and BIC statistics per year. In this section, the distributions are fitted to the five-year aggregated count data on reported and settled claims. In effect, one combined sample vector will be obtained for each category of claim counts.

The two resulting samples will be modelled accordingly, and compared to the earlier case of independent samples. This two-sided investigation ensures that, the general and specific characterisations of the datasets are fairly considered.

4.3.1 Estimation of Parameters

Applying the section 4.2.1 to the reported and settled five-Year Aggregate counts data, the NB parametric estimates and standard errors in Table 4.10 are obtained.

Table 4.10: **Parameter estimates from Negative Binomial model**

Variable	Parameters	5-Year Aggregate
Reported Claims	\hat{r}	0.6538 (0.0710)
	\hat{m}	477.0055 (53.2312)
Settled Claims	\hat{r}	0.8596 (0.0978)
	\hat{m}	74.7975 (7.3175)

Applying sections 3.4.3 and 4.2.1 to the reported and settled 5-Year Aggregate counts data, the DGP maximum likelihood estimates in Table 4.11 are produced.

4.3.2 Comparison of Models

Similar to 4.2.3 of the study, comparison of the aggregately fitted models will be performed, to aid selection based on the Akaike and Bayesian information criteria. Smaller values of the AIC and/or BIC statistics are indicators of better fitted models. Tables 4.12 and 4.13 display values of the AIC and BIC from the NB and DGP distributions, fitted to the 5-Year aggregated data on reported and settled claims counts.

Table 4.11: **Parameter estimates from Discrete Generalised Pareto model**

Variable	Parameters	5-Year Aggregate
Reported Claims	$\hat{\mu}$	2.3354 (1.9731)
	$\hat{\alpha}$	0.0017 (0.0012)
	$\hat{\lambda}$	2.8171 (2.0240)
Settled Claims	$\hat{\mu}$	1.9089 (1.6214)
	$\hat{\alpha}$	0.0077 (0.0015)
	$\hat{\lambda}$	1.9333 (1.4524)

Table 4.12: **AIC statistics for NB and DGP model**

Model	Criterion	5-Year Aggregate
Reported Claims		
NB model	<i>AIC</i>	1750.1770
DGP model	<i>AIC</i>	249.9887
Settled Claims		
NB model	<i>AIC</i>	1311.177
DGP model	<i>AIC</i>	251.9658

In terms of the aggregated reported claims count, DGP model presents smaller AIC and BIC values, in comparison with NB model. Therefore, on the basis of AIC and BIC, DGP model is preferred to NB model in the case of counts of reported claims. Regarding the aggregated settled claims count, DGP model exhibits smaller values of AIC and BIC. As a result, DGP is considered to provide a better fit to the counts of settled claims. In effect, the DGP model is recommendable, as it provides a more reliable fit for the count data on both classes of non-life insurance claims.

Table 4.13: BIC statistics for NB and DGP model

Model	Criterion	5-Year Aggregate
Reported Claims		
NB model	<i>BIC</i>	1755.8010
DGP model	<i>BIC</i>	251.1824
Settled Claims		
NB model	<i>BIC</i>	1316.801
DGP model	<i>BIC</i>	253.1595

CHAPTER 5

CONCLUSION AND RECOMMENDATIONS

This thesis examined the performance of Negative Binomial (NB) and Discrete Generalised Pareto (DGP) distributions in describing non-life insurance claims. Chapter one provided an introduction to the problems which motivated the study, contributions and objectives, and significance of the expected outcomes of the research. In chapter two, a review of relevant studies on claims modelling was conducted. Additionally, the review considered the background of NB and DGP distributions, as well as notable developments in the insurance industry of Ghana. Chapter three explored the theoretical concepts underlying the probability distributions, their applications, and methods employed in obtaining the parametric estimates. In chapter, the source and composition of the claims counts datasets were described. Fitting of NB and DGP distributions to the data followed, after which the selection of a suitable model was performed based on AIC and BIC. This chapter concludes the study, by summarising the key findings, and providing recommendations of academic and industrial interest.

5.1 Conclusion

Overall, the study demonstrated that non-life insurance claim counts can be described by two-parameter Negative Binomial and three-parameter Discrete Generalised Pareto probabilistic models. In the former distributional case, the mean and shape parameters were studied. In the latter case, the location, scale, and shape parameters were evaluated.

Data on reported and settled claim counts was obtained from National Insurance Company (NIC) to undertake the study. The data covers the 2012-2016 fiscal period, presenting annual records of claims counts for twenty-nine non-life insurance companies in Ghana. The data consists of 39,563 reported claims, and 5,210 settled claims observed across the five-year period.

The NB and DGP distributions were fitted to the discrete datasets on reported and settled claims counts. First, parameters of NB were estimated by the maximum likelihood approach. Secondly, DGP parameters were estimated in two steps. The μ and $(\mu + \epsilon)$ frequency technique produced initial estimators for obtaining the maximum likelihood estimates of the DGP parameters. To the best of our knowledge, there was no package available to compute the estimates. As such, R-based functions were written to facilitate the parameter estimation procedure.

Comparison of models was performed using the information criteria, AIC and BIC. In each case, the information criterion statistics associated with DGP recorded the minimum values. In view of their respective rules of selection, both criteria were found to favour the DGP model, in all years and categories of observed claim counts.

Specifically, the study concluded that:

1. Re-organising non-life claims counts data in terms of Number of Claims and the corresponding Frequency of Counts (see Appendices A.1 and B.2), provides an alternative data structure for efficient statistical modelling.
2. In evaluating initial estimators $(\alpha_0, \lambda_0, \mu_0)$ for determining the DGP parameters, the μ and $(\mu + 1)$ frequency method proposed by Prieto et al. (2014) is not valid in all cases. As a result, we propose the μ and $(\mu + \epsilon)$ frequencies; where $\epsilon > 0$ is the increment from the minimum, μ , to the next observation. Thus, $(\mu + \epsilon)$ is the smallest observation greater than μ . As demonstrated in the study, the modification provides a better estimation of initial values, by preserving the essential frequency information contained in the datasets.

3. Simulated Annealing (SANN) presents the most consistent optimisation routine for the DGP log-likelihood function with respect to a given set of initial parameters. The argument is based on outcomes of comparisons with CG, Brent, Nelder-Mead, and BFGS techniques as explored in the study.
4. Discrete Generalised Pareto demonstrates a better fit to the count data, relative to the Negative Binomial distribution. This is evident in both cases of disaggregated and aggregated yearly data on reported and settled claims counts.

5.2 Recommendations

Four recommendations are presented in this section, bordering on three areas of future studies, and one regulatory intervention.

First, adapting this study to a life insurance context would be beneficial to stakeholders. Although the thesis focused primarily on modelling data from the non-life insurance sector, a distributional enquiry in the life insurance sector will present unique insights. Presently, the latter sector consists of twenty-four licensed companies (NIC, 2018), making it the second largest division of insurance in Ghana. Therefore, exploring probability distributions in relation to life insurance claims, will complement the findings of this study, to provide NIC with a balanced view of both sectors.

Second, in future studies, introducing covariates in the modelling of claims counts and/or size data, may provide a better perspective. In terms of counts, identifying and investigating relevant explanatory variables will provide a framework for making informed claims forecasts. Furthermore, examining the sizes or amounts will enable insurers to measure the severity of claims and their impact. Outcomes of such research will contribute towards optimal funds allocation among companies.

Third, in further research, the model can be applied to data from countries with life table, to facilitate cross-comparisons between results generated by the selected model

and those derived from life tables. Benchmarking the model results against external life table readings can produce relevant findings for the insurance industry in Ghana, and provide statistical insights for model adjustments over the long-term.

Finally, management of NIC should regulate the offering of composite insurance covers by non-life companies. A composite cover is an insurance service, sold as a combination of multiple independent products. On one hand, such services maximise revenues of companies, by widening the bracket of product options available for public patronage. On the other, they minimise the specificity of data on claims, arising from occurrence of insured events. Enforcing the disaggregation of insurance products will ensure that claims data collected are informative, for enhanced industrial regulation and research.

REFERENCES

- Adombila, M. A. (2017, November 10). Gov't invites Exxon Mobil for direct negotiations on oil block. Retrieved from Graphic Online: <https://www.graphic.com.gh/news/general-news/gov-t-invites-exxon-mobil-for-direct-negotiations-on-oil-block.html>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6), 716–723. doi: 10.1109/TAC.1974.1100705
- Analyse-it Software Ltd. (n.d.). Parameter estimates. Retrieved May 13, 2018, from <https://analyse-it.com/docs/user-guide/fitmodel/linear/parameterestimates>
- Arnold, B. C. (2008). Pareto and Generalized Pareto Distributions. In S. E.-B. Economic Studies in Equality, & D. Chotikapanich (Ed.), *Modeling Income Distributions and Lorenz Curves* (Vol. 5). New York, NY: Springer.
- Arnold, B. C. (1983). Pareto Distributions. Maryland, MD, USA: International Co-operative Publishing House.
- Bank of Ghana. (n.d.). Insurance in Ghana - A Survey. Retrieved September 3, 2017, from <https://www.bog.gov.gh/privatecontent/Research/Articles/rpaper9.htm>
- Belisle, C. J. P. (1992). Convergence theorems for a class of simulated annealing algorithms on R^d . *Journal of Applied Probability*, 29 (4), 885–895. doi:10.2307/3214721
- Bulmer, M. G. (1979). Principles of Statistics. New York: Dover.
- Bunkley, N. (2008, March 3). Joseph Juran, 103, Pioneer in Quality Control, Dies. Retrieved December 21, 2017, from New York Times: <https://web.archive.org/web/20170906182706/http://www.nytimes.com/2008/03/03/business/03juran.html>

- Byrd, R. H., Lu, P., Nocedal, J. and Zhu, C. (1995). A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16, 1190–1208. doi:10.1137/0916069
- Carter, E. M., & Potts, H. W. (2014). Predicting length of stay from an electronic patient record system: a primary total knee replacement example. *BMC Medical Informatics and Decision Making*, 14 (26). doi: 10.1186/1472-6947-14-26
- CDC Consult, Microinsurance Centre, & PromIGH. (2015). Landscape Of Microinsurance In Ghana 2015. Retrieved from https://www.microfinancegateway.org/sites/default/files/publication_files/the_landscape_of_microinsurance_in_ghana_2015_0.pdf
- Chartered Institute of Loss Adjusters. (n.d.). The Principles of Insurance - Proximate. Retrieved from <http://www.cila.co.uk/cila/downloads/getting-qualified/certificate/files-2/205-chapter-8-proximate-cause/file>
- Chartered Insurance Institute. (n.d.). Types of insurance. Retrieved February 2018, 2, from <http://www.cii.co.uk/membership/new-starters/types-of-insurance/>
- DeGroot, M. H. (1986). Probability and Statistics (2 ed., 258–259). Addison-Wesley.
- Downing, D. (2009). Dictionary of Mathematics Terms (3 ed., pp.19). New York: Barron's Educational Series, Inc.
- Dziak, J. J., Coffman, D. L., Lanza, S. T., & Li, R. (2012). Sensitivity and specificity of information criteria. Retrieved from Pennsylvania State University: <https://methodology.psu.edu/media/techreports/12-119.pdf>
- Easton, V. J., & McColl, J.H. (1997). Random variables and probability distributions. *Statistics Glossary v1.1*. Retrieved from <http://www.stats.gla.ac.uk/steps/glossary/>
- Eduku, P. A. (2017, January 29). Insurers push for education et al to scale up microinsurance. Retrieved August 31, 2017, from Citifmonline Website:

<http://citifmonline.com/2017/01/29/insurers-push-for-education-et-al-to-scale-up-microinsurance/>

Efron, B. and Tibshirani, R. (1993) *An Introduction to the Bootstrap*. Chapman Hall.

Encyclopaedia of Mathematics. (2012, May 4). Negative Binomial Distribution. Retrieved February 17, 2018, from Encyclopaedia of Mathematics: https://www.encyclopediaofmath.org/index.php/Negative_binomial_distribution

Ernst & Young Global Limited. (2015). The future of insurance in a digital world. Retrieved from https://www.ey.com/Publication/vwLUAssets/The_future_of_insurance_in_a_digital_world/24FILE/EY-future-of-in-a-digital-world-design-comments.pdf

Everitt, B.S., & Skrondal, A. (2010). *The Cambridge dictionary of statistics* (4th ed., pp. 338). New York, NY: Cambridge University Press.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications* (2nd ed., Vol. 2). John Wiley & Sons, Inc.

Fisher, R.A. (1941). The negative binomial distribution. *Annals of Human Genetics*. 11(1), 182-187. doi: 10.1111/j.1469-1809.1941.tb02284.x

Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradients. *Computer Journal* 7, 148–154. doi: 10.1093/comjnl/7.2.149

Geerolf, F. (2017). A Theory of Pareto Distributions. Retrieved March 19, 2018 from <https://www.econ.ucla.edu/fgeerolf/geerolf-pareto.pdf>

Gschlöbl, S., & Czado, C. (2007). Spatial modelling of claim frequency and claim size in non-life insurance. *Scandinavian Actuarial Journal*. 2007(3), 202-225. doi: 10.1080/03461230701414764

Hilbe, J.M. (2011). *Negative Binomial Regression* (2nd ed.). New York, NY, USA: Cambridge University Press.

- Hitchcock, D., (1997). Discrete random variables, Retrieved from <http://people.stat.sc.edu/Hitchcock/notes515fall10chap4.pdf>
- Hoffman, D. (2003). Negative binomial control limits for count data with extra-Poisson variation. *The Journal of Applied Statistics in the Pharmaceutical Industry*, 2 (2), 127-132. doi: 10.1002/pst.51
- Insurance Act 724. (2006). Retrieved November 23, 2017, from <http://nicgh.org/wp-content/uploads/2016/07/InsuranceAct7242006VoC.pdf>
- Johnson, J. B., & Omland, K. S. (2004). Model selection in ecology and evolution. *Trends in Ecology & Evolution*, 19(2), 101-108. doi: 10.1016/j.tree.2003.10.013
- Kpedekpo, G. M. K. (1969). Working life tables for males in Ghana 1960. *Journal of the American Statistical Association*, 64(325), 102-110. doi: 10.1080/01621459.1969.10500957
- Krishna, H., & Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, 6(2), 177-188. doi: 10.1016/j.stamet.2008.07.001
- Levy, R. (2012, November 6). Probabilistic Models in the study of Language (draft).
- Mathworks Inc. (2015). Compute and Plot Negative Binomial Distribution PDF. Retrieved May 11, 2018, from <https://www.mathworks.com/examples/statistics/mw/stats-ex86650156-compute-and-plot-negative-binomial-distribution-pdf>
- Mierlus-Mazilu, I. (2010). On Generalized Pareto Distributions. *Romanian Journal of Economic Forecasting*, 107-117.
- Minitab Inc. (2017). What are parameters, parameter estimates, and sampling distributions? Retrieved May 19, 2018, from <https://support.minitab.com/en-us/minitab/18/help-and-how-to/statistics/basic-statistics/supporting-topics/data-concepts/what-are-parameters-parameter-estimates-and-sampling-distributions/>

- National Institute of Standards and Technology. (2012). What is a Probability Distribution. Retrieved August 30, 2017, from NIST/SEMATECH e-Handbook of Statistical Methods: <https://www.itl.nist.gov/div898/handbook/eda/section3/eda361.htm>
- National Insurance Commission. (2018). The industry: Types of companies. Retrieved June 14, 2017, from <http://nicgh.org/insurers-reinsurers/life-companies/>
- National Insurance Commission. (2016). The industry: Types of companies. Retrieved June 14, 2017, from <http://nicgh.org/the-industry/type-of-companies/>
- National Insurance Commission. (2017). Retrieved September 19, 2017, from <http://nicgh.org/the-industry/legal-environment/>
- Neath, A. and Cavanaugh, J. (2011). The Bayesian information criterion: Background, derivation, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(2), 199–203. doi: 10.1002/wics.199
- Nelder, J. A. and Mead, R. (1965). A simplex algorithm for function minimization. *Computer Journal*, 7, 308–313. doi: 10.1093/comjnl/7.4.308
- Oduro, F. (2017). Estimated life tables and mortality model for Ghana. *International Journal of Statistics and Applications*. 2017, 7(2), 121-130. doi: 10.5923/j.statistics
- Organisation for Economic Co-operation and Development. (2018). Global Insurance Market Trends 2017. Retrieved from <http://www.oecd.org/daf/fin/insurance/Global-Insurance-Market-Trends-2017.pdf>
- Pacáková, V. (2011, July 14-16). Modelling and simulation in non-life insurance. In N. Mastorakis, V. Mladenov, Z. Bojkovic, F. Topalis, and K. Psarris (Eds.), *ASM'11 Proceedings of the 5th international conference on Applied mathematics, simulation, modelling*, Corfu Island, Greece (pp. 129-133). Stevens Point, Wisconsin, USA: World Scientific and Engineering Academy and Society (WSEAS).
- Pareto, V. (1971). *Translation of Manuale di economia politica*. (A. S. Schwier, A. N. Page, Eds., & A. S. Schwier, Trans.) New York, A.M. Kelley.

- Pareto, V., Busino, G., & Bousquet, G.-H. (1964). *Cours d'Économie Politique*. Genève, Switzerland: Librairie Droz.
- Pennsylvania State University. (2018). Maximum Likelihood Estimation. Retrieved February 27, 2018, from <https://onlinecourses.science.psu.edu/stat414/node/191/>
- Philippou, A. and Antzoulakos, D. (2011). Binomial Distribution. *International Encyclopaedia of Statistical Science*, 152-154. doi: 10.1007/978-3-642-04898-2_46
- Pressman, R. S. (2010). *Software Engineering: A Practitioner's Approach*. Retrieved from http://dinus.ac.id/repository/docs/ajar/RPL-7th_ed_software_engineering_a_practitioners_approach_by_roger_s._pressman.pdf
- PriceWaterHouseCoopers. (2013). Insurance 2020: A quiet revolution – The future of insurance M&A. Retrieved from <https://www.pwc.com/gx/en/insurance/assets/pwc-insurance-2020-a-quiet-revolution-the-future-of-insurance-m-and-a.pdf>
- Prieto, F., Gómez-Déniz, E., & Sarabia, J. M. (2014). Modelling road accident blackspots data with the discrete generalized Pareto distribution. *Accident Analysis and Prevention*, 71, 38-49.
- Rooney, P. (2002, October 3). Microsoft's CEO: 80-20 rule applies to bugs not just features. Retrieved November 7, 2017, from The Channel Company: <https://www.crn.com/news/security/18821726/microsofts-ceo-80-20-rule-applies-to-bugs-not-just-features.htm>
- Schneider, F. (2014, February 18). What's in a methodology? Retrieved January 13, 2018, from Politics EastAsia Web site: <http://www.politicseastasia.com/studying/whats-methodology/>
- Schwarz, Gideon E. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6 (2), 461–464. doi:10.1214/aos/1176344136
- The R Foundation. (n.d.). What is R? Retrieved March 1, 2018, from <https://www.r-project.org/about.html>

- UCLA Institute for Digital Research and Education (IDRE). (n.d.). Negative Binomial Regression. Retrieved May 7, 2018, from UCLA IDRE web site: <https://stats.idre.ucla.edu/r/dae/negative-binomial-regression/>
- Upadhaya, B., Munir, R., & Blount, Y. (2014). Association between performance measurement systems and organisational Effectiveness. *International Journal of Operations & Production Management*, 34(7), 2-2.
- Viti, A., Terzi, A., & Bertolaccini, L. (2015). A practical overview on probability distributions. *Journal of Thoracic Disease*. 7(3), E7-E10. doi: 10.3978/j.issn.2072-1439.2015.01.37
- Vrieze, S. I. (2012). Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*, 17(2), 228-243. doi: 10.1037/a0027127
- Weisstein, E. W. (n.d.). Negative Binomial Distribution. Retrieved April 2018, 2, from MathWorld-A Wolfram Web Resource: <http://mathworld.wolfram.com/NegativeBinomialDistribution.html>
- World Health Organisation. (2018, May 24). The top 10 causes of death. Retrieved June 5, 2018, from <http://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- Wuthrich M.V. and Switzerland, R. (2014). Non-life insurance: mathematics & statistics.
- Xekalaki, E. (1983). Hazard functions and life distributions in discrete time. *Communications in Statistics - Theory and Methods*, 12 (21), 2503-2509. doi: 10.1080/03610928308828617

Appendices

A.1 Reported Claim Counts

Table 1: Restructured Reported Claims Dataset

Number of Reported Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
0	6	5	4	4	3
2	0	0	0	0	2
5	0	1	0	0	0
10	1	0	0	0	0
15	1	0	0	0	0
19	2	0	0	0	0
22	0	0	1	0	0
23	0	0	1	0	1
26	0	0	0	1	0
28	1	0	0	0	0
29	1	1	0	0	1
34	1	0	0	0	0
37	0	0	0	1	0
38	0	0	1	0	0
45	1	0	0	0	0
49	0	3	0	0	0

Continued on next page

Table 1 – *Continued from previous page*

Number of Reported Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
53	0	0	0	1	0
55	0	0	1	1	0
58	0	0	0	1	0
66	0	0	1	0	0
68	0	0	0	0	1
71	0	0	0	1	1
74	0	0	0	1	0
79	0	0	0	1	1
89	0	0	0	0	2
91	1	0	0	0	0
99	0	0	0	0	1
106	0	1	1	0	0
107	0	0	2	0	0
108	0	2	0	0	0
111	0	1	0	0	0
113	0	0	1	0	0
118	0	0	1	0	0
123	0	0	0	1	0

Continued on next page

Table 1 – *Continued from previous page*

Number of Reported Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
128	0	0	0	1	1
140	1	0	0	0	0
143	1	0	0	0	1
144	0	1	0	0	0
149	0	0	1	1	1
161	0	2	1	0	0
164	0	0	0	1	0
179	0	1	0	0	0
182	0	0	0	0	1
186	0	0	0	0	1
190	0	0	1	0	0
192	0	0	0	1	0
197	0	1	0	0	1
207	1	0	1	0	0
227	1	1	0	1	0
228	0	1	0	1	0
249	0	0	1	0	0
258	0	0	0	1	0

Continued on next page

Table 1 – *Continued from previous page*

Number of Reported Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
260	0	0	1	0	0
297	0	0	0	0	1
300	0	0	1	0	0
302	0	0	0	1	0
320	1	0	0	1	0
324	0	0	1	0	0
331	0	0	1	0	0
386	0	0	1	1	0
395	1	0	0	0	0
466	1	0	0	0	0
506	0	1	0	0	0
518	1	0	0	0	1
529	0	1	0	1	1
542	1	0	0	0	0
606	1	0	0	0	0
616	1	0	0	0	0
722	1	0	0	0	1
796	0	0	2	0	0

Continued on next page

Table 1 – *Continued from previous page*

Number of Reported Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
802	0	1	0	0	1
844	0	0	1	2	0
885	0	0	0	0	1
1041	1	0	0	0	0
1092	0	0	1	1	1
1187	1	0	0	0	1
1310	0	1	0	0	0
1335	0	1	1	0	0
1520	0	0	0	1	0
1533	0	1	0	0	0
2053	0	2	0	0	0
2151	0	0	0	1	1
9902	0	0	0	0	1
Total	29	29	29	29	29

B.2 Settled Claim Counts

Table 2: **Restructured Settled Claims Dataset**

Number of Settled Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
0	6	5	4	4	3
1	1	1	0	0	0
2	0	1	0	1	2
3	1	0	0	0	0
6	2	1	0	0	0
7	1	0	0	0	1
9	0	1	0	1	0
10	0	1	1	1	0
13	1	1	1	1	1
14	0	0	0	0	2
15	0	0	2	2	1
16	0	0	0	0	1
17	1	0	0	1	0
18	1	0	1	0	0
19	0	1	0	1	0
20	0	0	1	0	0

Continued on next page

Table 2 – *Continued from previous page*

Number of Settled Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
23	1	0	0	0	1
24	0	0	1	1	0
26	0	1	0	0	1
27	0	0	1	0	0
29	0	1	0	1	2
30	0	1	2	0	0
32	0	0	0	1	0
34	1	0	0	1	0
35	0	3	1	2	0
36	0	1	0	0	1
37	0	1	0	0	0
39	1	0	0	0	0
41	0	0	1	1	0
42	0	0	2	0	1
44	1	0	0	2	0
48	0	1	0	1	1
51	1	0	2	0	0
52	0	0	0	0	1

Continued on next page

Table 2 – *Continued from previous page*

Number of Settled Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
53	0	0	1	0	0
54	1	0	0	0	1
60	0	0	1	0	0
64	0	1	0	0	1
65	0	0	1	0	0
73	0	1	1	0	0
77	0	1	0	0	0
85	1	1	0	0	1
88	1	0	1	0	0
94	0	1	0	0	0
119	1	0	0	2	2
133	0	1	1	0	0
137	1	1	0	0	0
158	0	0	1	1	1
166	1	0	1	0	0
167	0	0	0	2	1
180	1	0	0	0	1
307	0	1	1	1	0

Continued on next page

Table 2 – *Continued from previous page*

Number of Settled Claims	Frequency of Counts by Year				
	2012	2013	2014	2015	2016
362	1	0	0	1	1
392	1	0	0	0	0
452	1	0	0	0	1
Total	29	29	29	29	29

C.3 R Source Codes

C.3.1 DGP Log-likelihood Function

Computes the log-likelihood for estimating DGP parameters:

```
dgp.fit<-function(params,x){
  alpha<-params[1]
  lambda<-params[2]
  mu<-params[3]
  logl<-sum(log((1+lambda*(x-mu))^-alpha)
            -(1+lambda*(x-mu+1))^-alpha)

  return(-logl)

#R optimisation routines, by default, perform minimisation
#on functions. However, since MLE entails maximisation of
#likelihood, the maximised DGP likelihood is obtained
#by negating the returned functional value.
}
```

C.3.2 DGP Parameter Estimator

Estimates simulated DGP parameters at initial functional values:

```
dgp.all<-function(data){
  data.params<-optim(par=c(Vector_Of_Initial_Estimators)
                    , dgp.fit, method="SANN", x=data, hessian = T)

  return(data.params)
```

```
}
```

C.3.3 Bootstrap Resampling

Estimates DGP SEs based on process in section 3.5.4:

```
ID<-function(data,R) sapply(1:R, function(i)
  sample(1:length(data), size=length(data),
  replace=T))

boot.samp<-function(idmat, data) sapply
  (1:ncol(idmat), function(i) data[idmat[,i]])

id<-ID(Vector_Of_Count_Data, 1000)
B.sample<-boot.samp(data=Vector_Of_Count_Data, idmat=id)

b.Par<-apply(B.sample, 2, function(x) tryCatch
  (dgp.est(x), error=function(err) rep(NA, 3)))

apply(b.Par, 1, function(x) sd(x, na.rm = T))
```

C.3.4 Information Criteria Statistics

Produces information criteria values for assessing DGP model suitability:

```
info.crit<-function(params) {
#Recall that the returned functional value in Appendix C.3.1
#was negated for the purpose of likelihood maximisation.
#Therefore, in computing AIC and BIC statistics for the DGP
#model, the functional value, represented by params$value,
```

```
#is negated to obtain the actual value of the function  
#at the maximum point.
```

```
logl<--params$value  
npar<-length(params$par)  
k<-2  
n_obs<-length(x)  
  
aic<--2*logl + k*npar  
  
bic<--2*logl + log(n_obs)*npar  
  
info.vals<-c(aic,bic)  
  
return(info.vals)  
}
```