UNIVERSITY OF NAIROBI

COLLEGE OF BIOLOGICAL AND PHYSICAL SCIENCES

SCHOOL OF MATHEMATICS

CLAIMS RESERVING USING OVERDISPERSED POISSON MODEL

SUBMITTED BY:

OGUTU JULIE AMOLO

REG NO: 156/60494/2010

A PROJECT SUBMITTED IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF MASTER OF SCIENCE IN ACTUARIAL SCIENCE

OCTOBER 2011

DECLARATION

I the undersigned declare that this project report is my original work and that to the best of my knowledge has not been presented for the award of a degree in any other university.

NAME: **OGUTU JULIE AMOLO**

DATE:.........28/10/.2011..................

SIGNATURE:...........................

This project has been submitted for examination with my approval as supervisor.

NAME: **PROF. PATRICK G.O.WEKE**

DATE:.........28/10/2011..................

SIGNATURE:...........................

ACKNOWLEDGEMENT

First and foremost, I would like to thank the Almighty God for giving me the knowledge, strength and wisdom to undertake this project.

I cannot also fully express my gratitude to my supervisor Prof Patrick G. O Weke for his generosity, invaluable positive assistance, faith and superb guidance throughout the project.

I also wish to thank the University of Nairobi for an opportunity to undertake this Masters degree at the institution. My gratitude goes to my fellow classmates and my lecturers for their constant inspiration by words of criticism and encouragement.

My heartfelt gratitude and thanks are to Peter Clark Memorial Fund and AMMSI.Thank you for giving me a chance in order to complete my Masters program. Your input towards my education was invaluable and I am indebted to you. Thank you.

I would be remiss if I did not mention my family members. It is their moral support and encouragement that spurred me on whenever I felt fatigued. And for that I am eternally grateful.

DEDICATION

To my mum Keryne Atieno, my late dad Benjamin Wambogo Ogutu, and my brothers Henry and Byron your love, loyalty, support and enthusiasm are second to none. Thank you for believing in me before I believed in myself

ABSTRACT

Traditionally, outstanding claims reserves were settled using deterministic methods which resulted in point estimates of the reserves. The primary advantage of stochastic reserving models is the availability of measures of precision of reserve estimates, and in this respect, attention is focused on the root mean squared error of prediction (prediction error). Of greater interest is a full predictive distribution of possible reserve outcomes, and different methods of obtaining that distributions are described. This study considers the Over- dispersed Poisson model for claims reserving in general insurance. In the over -dispersed Poisson model for loss reserving, it is assumed that the incremental claims are independent and Poisson distributed with the expectations being the product of two factors, depending on the occurrence year and the development year, respectively. The model is cast in the form of a generalized linear model, and a quasi-likelihood approach is used. The model presented here allows the actuary to provide point estimates and measures of dispersion, as well as the complete distribution for outstanding claims from which the reserves can be derived.

Keywords: Claims reserving, chain ladder method, Over-dispersed Poisson model, Generalized Linear Models

TABLE OF CONTENT

CHAPTER ONE

INTRODUCTION

1.1 Background

Although the last twenty years have witnessed increasing interest in stochastic claims reserving methods, they are still only used by a limited number of practitioners. A number of reasons for this could be suggested, including: a general lack of understanding of the methods; lack of flexibility in the methods; lack of suitable software; and so on. However, the main reason is probably lack of need for the methods, when traditional methods suffice for the calculation of a best estimate of outstanding claims reserves.

Forecasting outstanding claims and setting up suitable reserves to meet these claims is an important part of the business of a general insurance company. Indeed, the published profits of these companies depend not only on the actual claims paid, but on the forecasts of the claims which will have to be paid. It is essential, therefore, that a reliable estimate is available of the reserve to be set aside to cover claims, in order to ensure the financial stability of the company and its profit and loss account. There are a number of methods which have proved useful in practice, one of which is extensively used and is known as the chain ladder technique. In recent years, a statistical framework for analyzing this data has been built up, which encompasses the actuarial method, extending and consolidating it.

Traditionally, outstanding claims reserves were settled using deterministic methods which resulted in point estimates of the reserves, i.e. the present values of the expected future costs of claims. Increasing demand for further insight about the variability of the reserves has lead to the development of stochastic models for the reserve calculations.

The setting and monitoring of claims reserves is a vital task required of the general insurance actuary. To aid in the setting of reserves, the actuary can make use of a variety of techniques, the most familiar of which is the chain ladder model or variation. The principal aim of a reserving exercise is to provide an *estimate* of the amount of money a company should set aside now to meet claims arising in the future on the policies already written. The actuary cannot predict with certainty and knows that there is a distribution of possible outcomes, but uses the techniques at his or her disposal to arrive at the best estimate of the reserve (even if the *best* estimate is not that which is carried in the accounts). Knowledge of the *precision* of that estimate is also desirable. Traditional reserving techniques can help provide a best estimate (a measure of *location* in the distribution of possible outcomes), but cannot help with measures of precision. Of course, the actuary knows that the reserve estimate associated with a well-behaved class of business will be more precise than that of a poorly-behaved class and that the reserve estimate associated with a short-tailed class is likely to be more precise than that of a long-tailed class, but measuring that precision is difficult.

Notwithstanding certain well-known shortcomings, the chain ladder technique continues to occupy a prominent position with many practitioners as a claims reserving tool. There would appear to be some doubt as to the precise origin of the chain-ladder technique, but claims reserving techniques, in general, have generated a large body of research literature in the intervening years. One substantial strand of this literature is concerned with the development of stochastic claims reserving techniques which have clear advantages over deterministic techniques, such as provision for conducting diagnostic checks and the production of confidence intervals.

The aims of this study are threefold: to review the over-dispersed Poisson model which has been suggested, highlighting the connections between it and the chain ladder model; to show how the method can be implemented in practice; and to discuss the characteristics of the model and interpretation of the results.

This study also brings together the results and illustrates how the chain ladder technique can be improved and extended, without altering the basic foundations upon which it has been built. These improvements are designed to overcome two problems with the chain ladder technique. Firstly, that not enough connection is made between the accident years, resulting in an over-parameterized model and unstable forecasts. Secondly, that the development pattern is assumed to be the same for all accident years. No allowance is made by the chain ladder technique for any change in the speed with which claims are settled, or for any other factors which may change the shape , and show how it can be used to give upper prediction bounds on total outstanding claims of the run-off pattern. Before describing the methods for overcoming these problems, we first define the chain ladder model.

Stochastic reserving has long been ignored by many practitioners due to lack of easily implementable models; the ones available often required expensive specialist software. This situation is about to change due to the regulatory constraints in several countries where the regulators expect knowledge about not only the point estimate of the reserve but also about its variability. Development of own models are encouraged and often incentivized by lower statutory capital requirement. The “embedded value” of having own models instead of using the so called standard model created according to the one size fits all principle is thus easily quantifiable. Furthermore, the theories have become more accessible thanks to some authors who apparently have laid a lot of emphasis on the pedagogical aspects when serving the often heavy theories in their papers which became thus fully digestible even for practitioners who have left school a long time ago. Also computers are significantly faster and even standard softwares are better qualified to cope with the complexity involved in the calculations. In fact, one of the constraints imposed on the stochastic model selection has been the possibility of performing the calculations without needing expensive software.

The estimation of adequate reserves for outstanding claims is one of the main activities of actuaries in property/casualty insurance. The need to estimate future claims has led to the development of many loss-reserving techniques. Probably the oldest and most widely used of these techniques is the well known chain ladder. It is frequently used as a benchmark because of its generalized use and ease of application (Hess and Schmidt 2002). In its original form the chain ladder is a non stochastic algorithm for producing estimates of outstanding claims. There are many variations of the method; a description of one of them that will be useful in what follows will now be provided.

Negative incremental values can arise in the run-off triangle as a result of salvage recoveries, payments from third parties, total or partial cancellation of outstanding claims due to initial overestimation of the loss or to a possible favorable jury decision in favor of the insurer, rejection by the insurer, or just plain errors. England and Verrall (2002) argue that it is probably better to use paid claims rather than incurred claims (paid losses and aggregate case reserve estimates combined) since negative values are less likely to appear in the former. That is because case reserve estimates, the amount set aside by the claims handlers (see Chamberlin 1989; Brown and Gottlieb 2001), are set individually and often tend to be conservative, resulting in overestimation in the aggregate. Adjusting for this overestimation in the later stages of development may lead to negative incremental amounts. Whatever their cause, the presence of these negative incremental values in the data may cause problems when applying some claims-reserving methods. Thus, ideally, before applying claims-reserving methods, the actuary will revise and correct the data to eliminate negative incremental values. In this respect de Alba and Bonilla (2002) provide a list of potential adjustments frequently used in practice. However, even after correcting the data it is not always possible to eliminate all the negative values. Hence it is convenient to have available claims-reserving methods that will allow the actuary to compute the necessary reserves even in the presence of the negative values that may remain in the data.

Stochastic claims reserving models aim to provide measures of location (best estimates) and measures of precision (measures of variability) by treating the reserving process as a data analysis exercise and building a reserving model within a statistical framework. Once within a statistical framework, diagnostic checks of the fitted models are possible, such as goodness-of-fit tests and analysis of residuals (which highlight systematic and isolated departures from the fitted model). Various stochastic reserving models have been proposed over the last two decades, and work progresses as new techniques in the field of statistical modeling become available.

Considerable attention has been given to the relationship between various stochastic models and the chain ladder technique. Stochastic models have been constructed with the aim of producing exactly the same reserve estimates as the traditional deterministic chain ladder model. This might seem like a futile exercise, but has the advantages that measures of precision are readily available, and the assumptions underlying the chain ladder model are clarified. More importantly, it provides a bridge between traditional methods and stochastic methods, which is useful for the practitioner who is familiar with traditional methods and needs a starting point for exploring stochastic methods.

It is sometimes rather naively hoped that stochastic methods will provide solutions to problems when deterministic methods fail. Indeed, sometimes stochastic models are judged on whether they can help when simple deterministic models fail. This rather misses the point. The usefulness of stochastic models is that they can, in many circumstances, provide more information which may be useful in the reserving process and in the overall management of the company.

This paper identifies a statistical procedure which is exactly equivalent to the chain-ladder technique, in almost all circumstances. It should be noted that the method cannot be applied if the column sum of incremental claims for any development year is negative. We have always held the view that it is vital to subject any claims reserving procedure to a full statistical review, and we believe that the model presented in this paper provides an important framework to do this for the chain-ladder technique.

Broadly speaking, a stochastic claims reserving process involves three stages:

- *stage one:* the specification of a flexible parameterized model structure;

- *stage two:* a means of fitting the structure to the run-off data coupled with the means to conduct diagnostic checks on the fitted model; and

- *stage three:* a means of projecting the fitted structure into the target triangle.

Note that we assign the specification of a modeling distribution to stage two, since this leads to the construction of a likelihood or quasi-likelihood function which is maximized to obtain parameter estimates.

For the specific purpose of this paper, we find it both helpful and informative to relate the stochastic claims reserving process, whenever possible, to the generalized linear modeling technique, especially in relation to the first two stages described above

In the over dispersed Poisson model for loss reserving it is assumed that the incremental claims are independent and Poisson distributed with the expectations being the product of two factors, depending on the occurrence year and the development year, respectively. It is well known that maximum –likelihood estimation in the over dispersed Poisson model yields the chain ladder estimators of the expected ultimate aggregate claims.

The very nature of this paper means that a high technical content is unavoidable. Because of its pre-eminent position in claims reserving, and because it is well-known, widely used, and easy to apply, we begin by concentrating on the basic chain-ladder technique, then follow it up with the Over- dispersed Poisson model.

1.2 Objectives of the study

The main objective of this study is to provide measures of location (best estimates) and measures of precision (measures of variability) by treating the reserving process as a data analysis exercise and building a reserving model within a statistical framework. The specific objectives are:

- To review the Over-dispersed Poisson model
- To show how it can be implemented in practice

- To discuss the characteristics of the model, interpret the results and its wider usefulness.

1.3 Significance of the study

The significance of the study is to show that the Over-dispersed Poisson model is analogous to the chain ladder model since the results of the predicted values (reserves estimates) are the same as those of the chain ladder technique. It also applies the Over-dispersed model using the stochastic method and deterministic method and looks at the differences of both methods.

1.4 Organization of Report

Chapter two examines related works that have been done with respect to the Over-dispersed Poisson model and other stochastic models.

Chapter three looks at the chain ladder model, Over-dispersed Poisson model and the prediction errors.

Chapter four discusses the data analysis, estimates the parameters for Over-dispersed Poisson model.

Conclusions and recommendations for further research are then made; a list of references is also given.

CHAPTER TWO

LITERATURE REVIEW

Stochastic models for claims reserving improve on the classical approach by allowing the actuary to obtain measures of uncertainty and sometimes the complete distribution of outstanding claims. For a comprehensive, although not exhaustive, review of existing stochastic methods for claims reserving see England and Verrall (2002) or Hess and Schmidt (2002).

Mack (1993) presents one of the earliest attempts at formalizing a stochastic model for claims reserving. He proposes a nonparametric model that reproduces the chain ladder and obtains distribution-free expressions for the standard of reserve estimates. The use of the model is not limited by the existence of negative incremental claims. What may be considered a limitation is that it is directed at reproducing the chain-ladder reserves. England and Verrall (2002) have proposed the use of bootstrapping to compute the prediction errors. A significant landmark in the development of stochastic versions of the chain ladder technique was made by Kremer (1982), in which, in our opinion, he establishes the nature of the parameterized model structure which is inherent in the chain-ladder technique.

England and Verrall (2002) emphasize the framework of generalized linear models (GLMs; Anderson et al. 2004). They provide predictions and prediction errors for the different methods discussed and show how the predictive distributions may be obtained by bootstrapping and Monte Carlo methods. Among the models considered by England and Verrall there are several that can handle negative values: an (over-dispersed) Poisson, a negative binomial, and a normal approximation to the negative binomial. They also mention the log-normal model that was introduced by Kremer (1982) and analyzed in detail in Verrall (1991) when there are some negative incremental claims. Referring to the Poisson model they argue that "this does not imply that it is only suitable for data consisting exclusively of positive integers. That constraint can be overcome using a 'quasi-likelihood' approach, which can be applied to non-integer data, positive and negative." A similar argument is used for the negative binomial. It is not the purpose of this paper, but one could question whether they are really using those distributions or some continuous approximation. Also, since long-tailed distributions are used for modeling claim data, the normal distribution is not appropriate.

The log-Normal model was introduced by Kremer (1982). Christofides (1990) showed how spreadsheets could be used to analyze the data using log-Normal models.

It is often the case that parametric curves are too rigid (in some ways the opposite problem to the chain-ladder technique, which assumes no prior shape on the run-off), and England and Verrall (2001) proposed using non-parametric smoothing methods as an alternative. England and Verrall showed that it is possible to use a wide range of models with a non-parametric approach, with the chain-ladder technique at one end of the range, and the Hoerl curve at the other. The non-parametric smoothing models move seamlessly between these two extremes, and allow the practitioner to choose a model somewhere between the two. It is straightforward to examine the effect on the run-off pattern. Another example of the use of non-parametric smoothing was given in Verrall (1996). In that paper, the stochastic chain-ladder model of Renshaw and Verrall (1998) was extended to incorporate smoothing of parameter estimates over origin years, while leaving the model describing the run-off pattern alone.

Reserving specialists are probably more familiar with the practice of first fitting a chain-ladder model (or variation thereof), then smoothing the resultant development factors using a model with a fixed parametric form. Using that approach, the development factors themselves become the focus, and a model is fitted to development factors with equal weight (usually) being given to each development factor

Renshaw and Verrall (1998) were not the first to notice the link between the chain-ladder technique and the Poisson distribution, but were the first to implement the model using standard methodology in statistical modeling, and to provide a link with the analysis of contingency tables. Wright (1990) also describes a similar model, including a term to model claims inflation, but did not consider the model in detail. Mack (1991) also points out that the chain-ladder estimates can be obtained by maximizing Poisson likelihood by appealing to the so called `method of marginal totals'.

A discussion of the stochastic basis of chain-ladder models can be found in Mack (1994a), Verrall (2000), Mack and Venter (2000) and Verrall and England (2000). At the heart of the discussion is the relationship between the various models, and whether they can justifiably be used to add value to the deterministic chain-ladder technique.

In the context of GLMs the first stochastic version of the chain-ladder method that can be applied in the presence of negative incremental claim values is defined as a generalized linear model with an over-dispersed Poisson distribution (Renshaw and Verrall 1998). In the over-dispersed Poisson model the mean and variance are not the same.

A significant landmark in the development of stochastic versions of the chain ladder technique was made by Kremer (1982), in which, in our opinion, he establishes the nature of the parameterized model structure which is inherent in the chain-ladder technique. This is equivalent to stage one of the stochastic claims reserving process, but it should be noted that Kremer only identified one of two possible ways of building the structure into the process as a whole. The structure is comprised of the linear predictor: based on parameters corresponding to accident year i and delay j, which is connected to the logged incremental claim amounts.. Implicit in Kremer's work is the suggestion that the log transformation should be applied to the incremental claim amounts. The other possibility is to apply the log transformation to the expected values. Thus, Kremer, in specifying stage two of the stochastic claims reserving process, elects to model the incremental data by imposing the log-normal distribution. Renshaw (1989), Verrall (1989, 1990, 1991), motivated by Kremer (1982), have investigated many different facets of a stochastic claims reserving process based on the log-normal assumption, taken in conjunction with the predictor

Mack (1994) has criticized certain aspects of this work. This criticism is justified in so far as the model referred to above, derived by Kremer (1982), is not exactly equivalent to the chain-ladder technique. In Chapter 2 of this paper we will derive a generalized linear model which is exactly equivalent to the model which underpins the chain-ladder technique (noting the exception mentioned above). We believe that this equivalence is well known to a number of actuaries, but has not before been expressed in terms of generalized linear models. The contributions that this paper makes are to relate the chain-ladder technique directly to a generalized linear model, to show that it is not the most appropriate model for claims data, and to show how the model may be adapted in a straightforward way in order to be appropriate for claims data.

CHAPTER THREE

TECHNIQUES OF CLAIM RESERVING

3.1 PREDICTIONS AND PREDICTION ERRORS

ERROR TYPES

Before we are ready to take care of the chain ladder model approximation by the Over-dispersed Poisson, we need to introduce the different error types which play an important role in our analysis.

Since mathematical models are an only idealizations of the real world, these are associated with uncertainties or errors. According to (Daykin, Pentikāinen and Pesosnen, 1994), these errors can be divided into the following three categories:

- **Model errors** arise due to the fact that models are not known with certainty and are only approximations to the real world phenomena which they intend to model.
- **Parameter errors** are due to that the observations are limited in quantity so parameters are not known with certainty, and finally
- **Process error** (stochastic error) which arises due to the random fluctuations of the target quantities even in an ideal situation where the model and the parameters are correct.

In spite of its importance, model error is omitted in this paper and focus is laid entirely on the two latter errors or variances; the parameter error, even called estimation variance and the process errors. The sum of the process variance and the estimation variance is called prediction variance and is a measure of the variability of the prediction calculated as the root means squared error of the prediction (RMSEP).

Claims reserving is a predictive process: given the data, we try to predict future claims. In this context, we use the expected value as the prediction. When considering variability, attention is focused on the root mean squared error of prediction (RMSEP), also known as the prediction error.

Consider a random variable y and a predicted value $\hat{y}$. The mean squared error of prediction (MSEP) is:

$$E[(y-\hat{y})^2] = E[((y-E[y])-(\hat{y}-E[y]))^2] \tag{3.1}$$

Plugging in $\hat{y}$ instead of y in the final expectation and expanding gives:

$$E\{(y-\hat{y})^2\} \approx E\{(y-E[y])^2\} - 2E\{(y-E[y])(\hat{y}-E[\hat{y}])\} + E\{(\hat{y}-E[\hat{y}])^2\} \tag{3.2}$$

Assuming future observations are independent of past observations gives:

$$E\{(y-\hat{y})^2\} \approx E\{(y-E[y])^2\} + E\{(\hat{y}-E[\hat{y}])^2\} \tag{3.3}$$

Which, in other words, is

Prediction Variance = Process Variance + Estimation Variance

When trying to estimate the prediction error of future payments and reserve estimates using classical statistical methods, the problem reduces to estimating the two components: the process variance and the estimation variance. Alternatively, if the full predictive distribution can be found, the RMSEP can be obtained directly by calculating its standard deviation

It is important to understand the difference between the prediction error and the standard error. Strictly, the standard error is the square root of the estimation variance. The prediction error is concerned with the variability of a forecast, taking account of uncertainty in parameter estimation and also of the inherent variability in the data being forecast. Unfortunately, there is confusion in the literature over terminology, with the RMSEP also being called the standard error of prediction, or simply the standard error.

3.2 CHAIN LADDER MODEL

The chain-ladder technique was conceived as a deterministic method for predicting claim amounts. It is applied to cumulative claim amounts, and is designed to predict future incremental claim amounts in the empty cells of the well-defined triangular region to the immediate south-east of the run-off triangle. This region is called the target triangle.

The straightforward chain-ladder technique uses cumulative data, and derives a set of 'development factors' or 'link ratios'. It will be shown that to a large extent, it is irrelevant whether incremental or cumulative data are used when considering claims reserving in a stochastic context, and it is easier for the explanations here to use incremental. In order to keep the exposition as straightforward as possible, and without loss of generality, we assume that the data consist of a triangle of incremental claims. This is the simplest shape of data that can be obtained, and it is often the case that data from early origin years are considered fully run-off or that other parts of the triangle are missing. Using a triangle simply avoids us having to introduce complicated notation to cope with all possible situations. Thus, we assume that we have the following set of incremental claims data:

$$\{C_{ij} : i = 1,...,n; j = 1,..,n-i+1\} \qquad (3.4)$$

The suffix i refers to the row, and could indicate accident year or underwriting year, for example. The suffix j refers to the column, and indicates the delay, here assumed also to be measured in years. It is straightforward to consider data collected more frequently for all models discussed in this paper.

The cumulative claims are defined by:

$$D_{ij} = \sum_{k=1}^{j} C_{ik} \qquad (3.5)$$

and the development factors of the chain-ladder technique are denoted by $\{\lambda_j : j = 2,...,n\}$. The chain ladder technique estimates the development factors as:

$$\hat{\lambda} = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}} \tag{3.6}$$

These are then applied to the latest cumulative claims in each row ($D_{i,n-i+1}$) to produce forecasts of future values of cumulative claims:

$$\begin{aligned} \hat{D}_{i,n-i+2} &= D_{i,n-i+1}\hat{\lambda}_{n-i+2} \\ \hat{D}_{i,k} &= \hat{D}_{i,k-1}\hat{\lambda}_k, k = n-i+3, n-i+4,...,n. \end{aligned} \tag{3.7}$$

Thus, the chain-ladder technique, in its simplest form, consists of a way of obtaining forecasts of ultimate claims only. Here 'ultimate' is interpreted as the latest delay year so far observed, and does not include any tail factors. From a statistical viewpoint, given a point estimate, the natural next step is to develop estimates of the likely variability in the outcome so that assessments can be made, for example, of whether extra reserves should be held for prudence, over and above the predicted values. In this respect, the measure of variability commonly used is the prediction error, defined as the standard deviation of the distribution of possible reserve outcomes. It is also desirable to take account of other factors, such as the possibility of unforeseen events occurring which might increase the uncertainty, but which are difficult to model.

The first step to obtaining the prediction error is to formulate an underlying statistical model making assumptions about the data. If the aim is to provide a stochastic model which is analogous to the chain-ladder technique, then an obvious first requirement is that the predicted values should be the same as those of the chain-ladder technique. There are two ways in which this has been attempted: specifying distributions for the data; or just specifying the first two moments. As we wish to estimate the prediction error of the reserve estimate given by the chain ladder model, we have to find a suitable model which approximates the data in the triangle - both past and future data and hopefully result in the same or very similar estimate as that given by the chain ladder model. This is where the Over-Dispersed Poisson model enters the scene

3.3 OVER-DISPERSED POISSON MODEL

3.3.1 STRUCTURE ONE:

Non –linear model

The over-dispersed Poisson distribution differs from the Poisson distribution in that the variance is not equal to the mean, but, instead, is proportional to the mean. In claims reserving, the over-dispersed Poisson model assumes that the incremental claims C_{ij} are distributed as independent over-dispersed Poisson random variables, with mean and variance:

i.e

$$E[C_{ij}] = m_{ij} = x_i y_j$$
$$Var[C_{ij}] = \phi x_i y_j$$

where

$$\sum_{k=1}^{n} y_k = 1$$

Here, x_i is the expected ultimate claims (where ultimate means up to the latest development year observed in the triangle), and y_j is the proportion of ultimate claims to emerge in each development year. Over-dispersion is introduced through the parameter ϕ , which is unknown and estimated from the data. Allowing for over-dispersion does not affect estimation of the parameters, but does have the effect of increasing their standard errors is the proportion of ultimate claims to emerge in each development year.

It should be noted that, since y_j appears in the variance, the restriction that y_j must be positive is automatically imposed. This implies that the sum of incremental claims in column j must also be positive, which is a limitation of the model. Note that some negative incrementals are allowed, as long as any column sum is not negative.

In this formulation, the mean has a multiplicative structure, that is, it is the product of the row effect and the column effect. Both the row effect and the column effect have specific interpretations (being the expected ultimate claims in each origin year and proportion of ultimate to emerge in each development year, respectively), and it is sometimes useful to preserve the model in this form.

With the first structure the model is non-linear in the parameters and non- linear modeling techniques are required to obtain estimates of the parameters.

However, for estimation purposes, it is often better to re-parameterize the model so that the mean has a linear form. In the terminology of generalized linear models, we use a log link function so that:

$$\log(m_{ij}) = c + \alpha_i + \beta_j$$

This predictor structure is still a chain-ladder type, in the sense that there is a parameter for each row i, and a parameter for each column j. This structure is discussed in details below as a generalized linear model

3.3.2 STRUCTURE TWO:

Generalized Linear Model

The Over Dispersed Poisson (ODP) model assumes that the incremental claims have an over dispersed Poisson distribution. An ODP looks like a Poisson distribution but it's variance is not equal to the mean but proportional to it where the proportionality factor is called the over dispersion parameter. We have

$$C_{ij} \sim iid\ ODPo(\mu_{ij}) \qquad (3.8)$$

where iid denotes independent, identically distributed with $E[C_{ij}] = \mu_{ij}$ and

$Var[C_{ij}] = \phi E[C_{ij}]$. Let

$$log\mu_{ij} = \mu + \alpha_i + \beta_j \qquad (3.9)$$

which is recognized as a generalized linear model in which the responses C_{ij} are modeled with a logarithmic link function and linear predictor, η_{ij}. This model is linear in the parameters and is thus suitable for fitting the chain ladder model since we have one parameter for each row *i* and each column *j*. Due to this overparametrisation of the model, we have as many parameters as values to fit, we apply the corner constraints as follows:

$$\begin{cases} \eta_{ij} = \mu + \alpha_i + \beta_j \\ \alpha_1 = 0 \\ \beta_1 = 0 \\ log\mu_{ij} = \eta_{ij} \end{cases} \qquad (3.10)$$

i.e. the first two parameters are zeroised.

β_j, the column parameter determines the run off structure of the data. Since we have one parameter for each column, we assume that there is no particular shape of the run off pattern, which is in line with the general assumptions imposed on the traditional chain ladder model. The ODP model is robust for a small number of negative incremental claims. Indeed, this is an important feature especially for the product lines where case reserves are set by the claims handlers who quite often tend to be over conservative in their judgments. Adjusting for this overestimation at a later stage may lead to negative incremental values.

ESTIMATION OF THE DISPERSION PARAMETER

The next step is to estimate the parameters of the ODP model. We begin with the estimation of the dispersion parameter Φ. Given that

$$\sum\left(\frac{\frac{C_{ij}}{\phi}-\frac{\widehat{\mu_{ij}}}{\phi}}{\sqrt{\frac{\widehat{\mu_{ij}}}{\phi}}}\right)^2 \sim \chi^2_{n-p} \qquad (3.11)$$

which is a result of that standard residuals, i.e. the quotient which is squared in the expression above is N(0,1) distributed and the sum of squared N(0,1) variables are $\chi 2$ distributed with the degree of freedom equal to the number of observations less the number of parameters.

The method of moments gives:

$$\sum\left(\frac{C_{ij}-\widehat{\mu_{ij}}}{\sqrt{\phi}\sqrt{\widehat{\mu_{ij}}}}\right)^2 = n-p \qquad (3.12)$$

since Φ is constant it is readily available from (3.12):

$$\hat{\phi} = \frac{1}{n-p}\sum\left(\frac{C_{ij}-\hat{\mu}_{ij}}{\sqrt{\hat{\mu}_{ij}}}\right)^2 \qquad (3.13)$$

One may interpret Φ as the average claim size while C_{ij} is the number of claims.

PARAMETER ESTIMATION

The estimation of the other parameters of the model is done by calculating the log-likelihood function, *l*, and maximizing it:

$$l = \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \frac{1}{\phi}(C_{ij} \log \mu_{ij} - \mu_{ij}) + \dots\dots\dots \qquad (3.14)$$

$$= \frac{1}{\phi}\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \left[\left(C_{ij}(\mu + \alpha_i + \beta_j\right) - e^{\mu+\alpha_i+\beta_j}\right] + \cdots \qquad (3.15)$$

ESTIMATION OF THE PREDICTION VARIANCE

The estimation variance for each estimated value is

$$Var(\hat{\mu}_{ij}) = Var(e^{\hat{\eta}_{ij}}) = \left(\frac{\partial e^{\hat{\eta}_{ij}}}{\partial \hat{\eta}_{ij}}\right)^2 Var(\hat{\eta}_{ij}) \qquad (3.16)$$

Taylor estimation of (3.16) gives:

$$Var(\hat{\mu}_{ij}) = (e^{\hat{\eta}_{ij}})^2 Var(\hat{\eta}_{ij}) = \hat{\mu}_{ij}^{\,2} Var(\hat{\eta}_{ij}) \qquad (3.17)$$

According to (3.3) the prediction variance for each value in a cell is:

$$PE = \hat{\emptyset}\hat{\mu}_{ij} + \hat{\mu}_{ij}^{\,2} Var(\hat{\eta}_{ij}) \qquad (3.18)$$

In the overall prediction variance even the covariance should be taken into account:

$$PE = \sum \hat{\phi}\hat{\mu}_{ij} + \sum \hat{\mu}_{ij}^{\,2}\, Var(\hat{\eta}_{ij}) + 2\sum Cov(\hat{\eta}_{ij}, \hat{\eta}_{ik})\, \hat{\mu}_{ij}\hat{\mu}_{ik} \qquad (3.19)$$

The indexes under the summations above have been omitted. It is not hard to realize, that calculation of the prediction error is quite cumbersome.

An alternative which is relatively easy to implement in a spreadsheet was presented by professor R.J.Verrall at a seminar held in Stockholm in September 2007. The method is best understood when setting up the model in matrix form. The design matrix is denoted by X, the parameter vector by $\hat{\beta}$ and the fitted values by $\hat{\mu}$:

$$\hat{\mu} = e^{(X\hat{\beta})} \tag{3.20}$$

the covariance matrix of $\hat{\beta}$ is

$$\hat{\Sigma} = \hat{\phi}(X^T W X)^{-1} \tag{3.21}$$

where

$$W = diag\left[\frac{1}{V(\mu_{ij})}\left(\frac{\partial \mu_{ij}}{\partial \eta_{ij}}\right)^2\right] \tag{3.22}$$

If the design matrix of future values is denoted with F then the covariance matrix of the linear predictors can be written in analogy with 14 as F $\hat{\Sigma}$ F^T and the covariance matrix of fitted values as diag ($\hat{\mu}$) F ΣF^T diag ($\hat{\mu}$).The data for the forecast vector is being placed in the same column (only one column). The goal is to calculate the prediction of the reserve.

The procedure can be summarized by the following scheme:

1. Set up the chain ladder triangle of cumulative data
2. Calculate the linear predictors
3. Calculate the mean for each cell
4. Calculate the log-likelihood for each cell
5. Sum up the log-likelihood

6. Estimate the parameters by minimizing the log-likelihood above

7. Calculate the Pearson residuals

8. Calculate the over dispersion parameter

9. Calculate the prediction error using the matrix set up.

However, instead of following the above procedure one may opt to write a code using standard software packages that would give the parameter estimates, calculate the total reserves and prediction error.

Thus the generalized linear model for the over-dispersed Poisson model can be solved using standard software packages. eg R, SPPS

R is one of the statistical packages that can be used to fit GLMs.

The over-dispersion manifests itself in the fact that the variance of the claims is *proportional* to the mean, rather than equal to it (as in the Poisson distribution). The log of the mean claim (in GLM-speak this means we are using a log link function) is equal to a linear function of both the origin period and the development period. These are usually referred to as *factors*.

For this project I will use the Generalized Linear Model, use R program and also Excel to calculate total reserves and the prediction error of the reserve

CHAPTER FOUR

DATA ANALYSIS

To illustrate the methodology, consider the claims amounts in Table 1, shown in incremental form. The data is taken from a paper by Renshaw (1989) and consists of claims from a portfolio of general insurance. The data is analyzed using R and Excel

Table 1: Run – off Claims Data

35784	766940	610542	482940	527326	574398	146342	139950	227229	67948
352118	884221	933894	1183289	445745	320996	527804	266172	280405	
290507	1001799	926219	1016654	750816	146923	495992	2480405		
310608	1108250	776189	1562400	272482	352053	206286			
443160	693190	991983	769488	504841	470639				
396132	937085	847498	805037	705960					
440832	847631	1131398	1063269						
359480	1061648	1443370							
376686	986608								
344014									

4.1 Estimation of parameters

From the equation (3.9) ie $log\mu_{ij} = \mu + \alpha_i + \beta_j$ we solve the parameter estimates

setting the corner constraints as $\begin{cases} \eta_{ij} = \mu + \alpha_i + \beta_j \\ \alpha_1 = 0 \\ \beta_1 = 0 \\ log\mu_{ij} = \eta_{ij} \end{cases}$

Where μ is a constant

Table 2 shows the parameter estimates and their standard errors obtained by fitting the over-dispersed Poisson model. For many of the parameters, the standard errors are small relative to the estimates themselves. This does not provide evidence that those estimates can be set to zero, since doing so may ignore a trend, which itself may be statistically significant. Ideally, the strength of that trend should be tested, and modeled directly, but, in this example, I ignore that feature, since the purpose is to fit a model which reproduces chain-ladder estimates.

Table 2: Over-dispersed Poisson model; parameter estimates

	Parameter Estimate	Standard Error
Constant(u)	12.17558	0.27788
Alpha 2	0.39160	0.24079
Alpha 3	0.76545	0.22923
Alpha 4	0.53650	0.25255
Alpha 5	0.45149	0.26324
Alpha 6	0.50397	0.26739
Alpha 7	0.60873	0.27290
Alpha 8	0.79669	0.29092
Alpha 9	0.62606	0.36968
Alpha 10	0.57285	0.65600
Beta 2	1.01435	0.23469
Beta 3	1.06443	0.24032
Beta 4	1.13509	0.24676
Beta 5	0.54718	0.28670
Beta 6	0.19507	0.33264
Beta 7	0.11264	0.36785
Beta 8	1.16757	0.30295
Beta 9	0.05394	0.54395
Beta 10	-1.0490	1.3656

4.2 Practical Application to a general insurance company

Estimates of future payments can be obtained from the parameters and inserting them into equation 2 and exponetiating.

ie

$$E[C_{ij}] = \mu_{ij}$$

But

$log\mu_{ij} = \mu + \alpha_i + \beta_j$

Thus

$$E(C_{ij}) = e^{c+\alpha_i+\beta_j}$$

Table 3 shows the observed claim amounts and estimated future claims (Excel used to solve estimated future payments)

Table 3: Estimated claims amounts and future claims.

35784	766940	610542	482940	527326	574398	146342	139950	227229	67948
352118	884221	933894	1183289	445745	320996	527804	266172	280405	100518
290507	1001799	926219	1016654	750816	146923	495992	2480405	440194	146084
310608	1108250	776189	1562400	272482	352053	206286	1066241	350116	116191
443160	693190	991983	769488	504841	470639	341025	979346	321583	106722
396132	937085	847498	805037	705960	390280	359400	1032114	338910	112472
440832	847631	1131398	1063269	616301	433384	399093	1146105	376341	124894
359480	1061648	1443370	1338901	743742	523002	481620	1383103	454162	150720
376686	986608	1051859	1128873	627074	440960	406070	1166140	382920	127077
344014	9486356	997353	1070375	594580	418110	385028	1105712	363077	120492

Total reserve Estimates is given by

$$\hat{C}_{++} = \sum_{i,j<\in\Delta} \hat{C}_{ij}$$

The total reserve can be calculated using either the deterministic method (using Excel) or the stochastic method(using R)

Table 4 shows the cumulative claims amount and projections of cumulative claims used to calculate total reserve.

1. Using Excel

To calculate total reserve we will use the cumulative claims data and estimated cumulative future claims.

Table 4: Cumulative claims and estimated future cumulative claims.

35784	802724	1413266	1896206	2423532	2997930	3144272	3284222	3511451	3579399
352118	1236339	2170233	3353522	3799267	4120263	4648067	4914239	5194644	5295162
290507	1292306	2218525	3235179	3985995	4132918	4628910	7109315	7549509	7695593
310608	1418858	2195047	3757447	4029929	4381982	4588268	5654509	6004625	6120816
443160	1136350	2128333	2897821	3402662	3873301	4214326	5193671	5515254	5621976
396132	1333217	2180715	2985752	3691712	4081992	4441392	5473506	5812416	5924888
440832	1288463	2419861	3483130	4099431	4532815	4931908	6078013	6454354	6579248
359480	1421128	2864498	4203399	4947142	5470144	5951763	7334866	7789028	7939749
376686	1363294	2415153	3544026	4171100	4612060	5018130	6184270	6567190	6694267
344014	1292650	2290003	3360378	3954958	4373068	4758095	5863807	6226884	6347376

Thus

Total reserve = (5295162-5194644)+(7695593-7109315)+(6120816-4588268)+(5621976-3873301)+(5924888-3691712)+(6579248-3483130)+(7939749-2864498)+(6694267-1363294)+(6347376-344014)

=25706898

Table 5 shows the reserves for each year .Note that no projection can be done for the first accident year because it is not possible to project beyond the highest development year.

Table 5 : Reserves for each year of origin

Year 1	0
Year 2	100518
Year 3	586278
Year 4	1532548
Year 5	1748675
Year 6	2233176
Year 7	3096118
Year 8	5075251
Year 9	5330973
Year 10	6003362
Total reserve	**25706898**

From Excel the total reserve is estimated to be 25706898

2. Using R

Total reserve can be calculated using R. ie

Total reserve= 25706974

Total RMSEP is given by

$$PE = \sum \hat{\phi}\hat{\mu}_{ij} + \sum \hat{\mu}_{ij}{}^{2}\, Var(\hat{\eta}_{ij}) + 2\sum Cov(\hat{\eta}_{ij}, \hat{\eta}_{ik})\, \hat{\mu}_{ij}\hat{\mu}_{ik}$$

Using R

Total RMSEP = 5854802

And

Total Prediction Error = 23 %

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

In conclusion, the assessment of the financial strength of a general insurance enterprise includes a thorough analysis of the outstanding claims reserves, including an assessment of the possible variability in the reserves.

Table 5 shows the total estimated future claims (reserves).Practitioners have keen interest in the values of table 5 since these values are estimates of the outstanding claims provision at the present time (i.e at the end of accident year 10) with respect to year of origin and the total overall outstanding claims provision for the entire year .The estimates are of significant use in forecasting the IBNR claims provision and in general organization of business.

It can be seen from our model above that the value for total reserve varies slightly when calculated using the deterministic method (25706898) and the stochastic method (25706974). However, the difference is small and can be considered negligible. The difference is due to round up differences in the calculation.

Therefore our model suggests that the reserve is 25706974 with a standard error of prediction of 23% (5854802)

This model provides a simple method whose application in claims reserving is nearly as simple to execute as the chain ladder method but has the advantage of providing goodness-of-fit test statistic and the estimation error. (This is so because the goodness-of-fit statistic of the model can be performed using the R package since both the deviance and degree of freedom are given in the output). It should also be noted that despite not mentioning the dispersion parameter in the data analysis above, the software R calculates it and uses it to calculate the prediction error.

In addition to the above conclusions, I would like to recommend further study to be done to check how the model behaves when we incorporate exposure measures and extend to include an inflationary trend (i.e. modeling of claims inflation)

BIBLIOGRAPHY

Anderson, Duncan, Sholom Feldblum, Claudine Modlin, Doris Schirmacher, Ernesto Schirmacher, and Neeza Thandi (2004). A Practitioner's Guide to Generalized Linear Models. 2004 Discussion Paper Program, Casualty Actuarial Society, Arlington, VA

Brown, Robert L., and Leon R. Gottlieb (2001). *Introduction to Ratemaking and Loss Reserving for Property and Casualty Insurance*. 2nd ed. Providence, RI: ACTEX Publications

Chamberlin, Gary F., ED. (1989) *Claims Reserving Manual, Vol. 1*. London: Institute of Actuaries

Christofides, S. (1990). *Regression models based on log-incremental payments. Claims Reserving Manual*, 2, Institute of Actuaries, London

Daykin, C. D., Pentikäinen, T., & Pesosnen, M. (1994). *Practical Risk Theory for Actuaries,*. London: Chapman & Hall.

De Alba, Enrique, and Roberto Bonilla (2002). Un Modelo para el Tratamiento de Valores Negativos en el Tria´ngulo de Desarrollo Utilizado en la Estimacio´n de Reservas para SONR. *In Transactions of the 27th International Congress of Actuaries*, Cancu´n, Me´xico

England, P. D. and Verrall, R. J. (2001). A flexible framework for stochastic claims reserving. *Proceedings of the Casualty Actuarial Society*.

England, Peter, and Richard J. Verrall. (2002). Stochastic Claims Reserving in General Insurance (with discussion). *British Actuarial Journal* 8: 443–544.

England, Peter (2002). Addendum to Analytic and Bootstrap Estimates of Prediction Errors in Claims Reserving. *Insurance: Mathematics and Economics* 31: 461–66

Hess, Klaus T., and Klaus D. Schmidt (2002). A Comparison of Models for the Chain-Ladder Method. *Insurance: Mathematics and Economics* 31: 351–64

Kremer, E. (1982). IBNR claims and the two way model of ANOVA. *Scandinavian Actuarial Journal*, pg 47-55

Mack, Thomas (1993). Distribution-Free Calculation of the Standard Error of Chain Ladder Reserve Structure. *ASTIN Bulletin* 23: 213–225

Mack, T. (1994). Which stochastic model is underlying the chain-ladder model? *Insurance: Mathematics and Economics*, 15, 133-138.

Mack, T. and Venter, G. (2000). A comparison of stochastic models that reproduce chain-ladder reserve estimates. *Insurance: Mathematics and Economics*, 26, 101-107

McCullagh, P. and Nelder, J. (1989). *Generalized linear models*, 2nd edition. Chapman and Hall

Renshaw, A. E. and Verrall, R. J. (1998). A stochastic model underlying the chain-ladder technique. *British Actuarial Journal* 4, 903-923.

Renshaw, A. E. (1994). On the second moment properties and the implementation of certain GLIM based stochastic claims reserving models. Actuarial Research Paper No.

65, Department of Actuarial Science and Statistics, City University, London.

Verrall, R. J. (1996). Claims reserving and .generalized additive models. *Insurance: Mathematics and Economics*, 19, 31-43

Verrall, Richard J. (1991). Negative Incremental Claims: Chain Ladders and Their Models. *Journal of the Institute of Actuaries* 120(1): 171–183

Verrall, R. J. and England, P. D. (2000). Comments on: `A comparison of stochastic models that reproduce chain-ladder reserve estimates', by Mack & Venter. *Insurance: Mathematics and Economics*, 26, 109-111.

APPENDIX

R SCRIPT

```
# Load up data
data <- read.csv("research.csv", header=F)
data <- as.matrix(data)
data
# Prepare data in correct form
claims <- as.vector(data)
n.origin <- nrow(data)
n.dev <- ncol(data)
origin <- factor(row <- rep(1:n.origin, n.dev))
dev <- factor(col <- rep(1:n.dev, each=n.origin))
# Put into a data frame (no need, but easier to visualise)
research <- data.frame(claims=claims, origin=origin, dev=dev)
research[1:5, ] # Print first five rows

# New quasi-poisson family
quasipoisson <- function (link = "log")
## Amended by David Firth, 2003.01.16, at points labelled ###
## to cope with negative y values
##
## Computes Pearson X^2 rather than Poisson deviance
##
## Starting values are all equal to the global mean

{
linktemp <- substitute(link)
if (!is.character(linktemp)) {
linktemp <- deparse(linktemp)
if (linktemp == "link")
linktemp <- eval(link)
}
if (any(linktemp == c("log", "identity", "sqrt")))
stats <- make.link(linktemp)
else stop(paste(linktemp, "link not available for
poisson",
"family; available links are", "\"identity\", \"log\"
and \"sqrt\""))
variance <- function(mu) mu
validmu <- function(mu) all(mu > 0)
dev.resids <- function(y, mu, wt) wt*(y-mu)^2/mu ###
aic <- function(y, n, mu, wt, dev) NA
initialize <- expression({
n <- rep(1, nobs)
mustart <- rep(mean(y), length(y)) ###
```

```
})
structure(list(family = "quasipoisson", link = linktemp,
linkfun = stats$linkfun, linkinv = stats$linkinv,
variance = variance,
dev.resids = dev.resids, aic = aic, mu.eta =
stats$mu.eta,
initialize = initialize, validmu = validmu, valideta =
stats$valideta),
class = "family")
}

# Fit model
model <- glm(claims ~ origin + dev, family = quasipoisson(),
subset=!is.na(claims), data=research)
summary(model)

# Extract useful info from the model
coef <- model$coefficients # Get coefficients
disp <- summary(model)$dispersion # Get dispersion parameter
cov.param <- disp * summary(model)$cov.unscaled
# Get covariance matrix of parameters
# To determine future uncertainty, need to create a
# design matrix for future payments. Build up in stages.
# Assume start from bottom left of future triangle.
n.fut.points <- length(claims[is.na(claims)])
fut.design <- matrix(0, nrow = n.fut.points, ncol=length(coef))
fut.points <- claims
fut.points[!is.na(claims)] <- 0
fut.points[is.na(claims)] <- 1:n.fut.points
for(p in 1:n.fut.points){
# All points and a constant in the predictor
fut.design[p, 1] <- 1
# Row factor
fut.design[p, 1 +
as.numeric(origin[match(p, fut.points)]) - 1] <- 1
# Col factor
fut.design[p, 1 + (n.origin-1) +
as.numeric(dev[match(p, fut.points)]) - 1] <- 1
}

# Determine fitted future values (as a diagonal matrix)
fitted.values <- diag(as.vector(exp(fut.design %*% coef)))
total.reserve <- sum(fitted.values)
total.reserve

[ 1] 25706974
```

```
# Determine covariance matrix of linear predictors
cov.pred <- fut.design %*% cov.param %*% t(fut.design)
# Determine covariance matrix of fitted values
cov.fitted <- fitted.values %*% cov.pred %*% fitted.values
# Determine uncertainty statistics
total.rmse <- sqrt(disp*total.reserve+sum(cov.fitted))
total.predictionerror <- round(100*total.rmse/total.reserve)
total.rmse
[1] 5854802

total.predictionerror

[1] 23
```