A STOCHASTIC APPROACH IN DETERMINING CLAIMS RESERVE IN GENERAL INSURANCE

University of Nairobi

SCHOOL OF MATHEMATICS

Submitted by:

Samuel K. Muthee,
I46/60030/2007

Supervised by:

Dr. PGO Weke

A PROJECT SUBMITTED TO THE DEPARTMENT IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF POSTGRADUATE DIPLOMA IN ACTUARIAL SCIENCE

AUGUST 2009

DECLARATION

This project is my original work and has not been presented for a degree in any other university.

Date: 26/08/2009

SAMUEL KAMAU MUTHEE

This project has been submitted with my approval as the university supervisor.

Date: 26/08/2009

Dr. PGO WEKE

ACKNOWLEDGEMENT

I wish to acknowledge and thank all those who gave me the necessary support and made it possible for me to complete this work.

I am highly indebted to my supervisor Dr. Weke for his dedication to this work. His support and guidance has gone a long way in making this work a success.

Special thanks to mum and dad for their relentless support in my academics, it would be incomplete if I failed to mentioned the support and encouragement provided by my sister, Eunice.

TO ALL, THANK YOU.

TABLES OF CONTENTS

ABSTRACT.

This project illustrates the use of stochastic techniques in determining claims reserve in general insurance. The main objective being to obtain a 'best estimate' of the outstanding reserves and its variability with more emphasis being made on the application of stochastic techniques. An in-depth application of the Mack's model and the Negative binomial model has been considered. The application is in two dimensions; firstly, a distribution free method is considered and secondly, an assumption of the underlying distribution is made. The analysed data demonstrate that despite applying models that give reliable results, the integrity of the data should be guaranteed to give realistic results. In conclusion, stochastic techniques should be adopted because they allow the reserving actuary to determine the range within which he expects payments to fall with a certain level of confidence.

CHAPTER 1

INTRODUCTION.

1.1 Background of the study.

Reserving is a fundamental aspect of business management. The apparent profitability of a business as well as its solvency is highly dependent upon the value of the reserves and the reserving philosophy. Most of the key financial performance statistics used by insurance company analysts depend in some way upon the reserve value.

General insurance companies need to hold reserves because the timing of premiums receipt and claims payment does not coincide. Furthermore, there is a delay between the claim event and the claim settlement dates and this means that the insurer must set up reserves in respect of those claims still to be settled.

It is therefore important to recognize that 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 insurer must be able to quantify this liability if it is to assess its financial position correctly, both for statutory and for internal purposes.

Hence, the primary reasons to hold reserves are illustrated as follows:

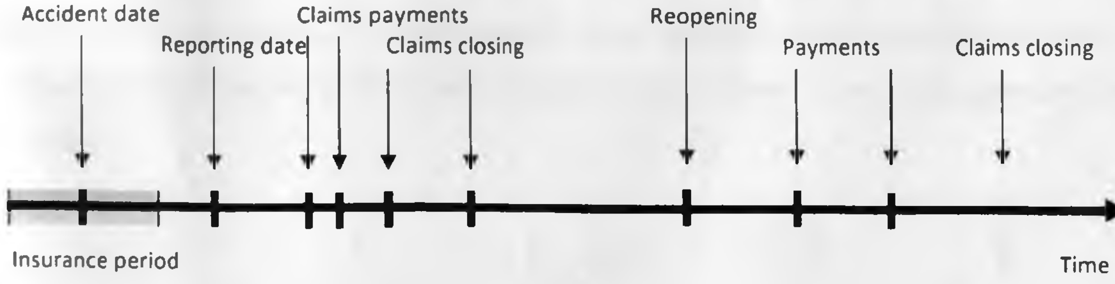
(i) premiums have been received but the cover relating to part of the premiums has not yet been completed ,

(ii) claims will have happened but will not yet have been paid, broken down into claims that have been reported and claims that have not yet been reported ,

(iii) other expenses incurred by the insurance company that have yet to be paid, such as tax payments,

(iv) It can also happen that a closed claim needs to be reopened due to (unexpected) new developments, or if a relapse occurs.

The history of a typical non-life insurance claim may take the form shown below:



The Insurance industry in Kenya has in the last decade experienced a series of collapse of insurance companies which has been attributed to mismanagement. The reserving exercise, if properly applied, can be used to address some of the insolvencies experienced. To guarantee solvency, the Government of Kenya doubled the minimum capital requirements for general insurance companies from KShs 150 million to KShs 300 million in the 2007/2008 budget. The main purpose for existence of insurance companies is to assume the risks inherent in the business environment for a premium and hence offer financial security in return. Therefore, stability of insurance companies should be guaranteed or else the whole purpose of insurance would be defeated.

The reserving philosophy of an insurance company should be reviewed and reserves should be maintained per class of business since short-tailed classes (e.g. motor vehicle and property) require less reserve than long-tailed classes. (e.g. asbestos or environmental pollution claims, liability claims)
Traditionally, deterministic methods have been applied to arrive at point estimates of reserves, however due to advancement of the computer age and more research on this area, stochastic techniques have gained popularity among practitioners not only for their degree of precision but also for describing the distribution underlying the claims pattern. With stochastic approach, the actuary can determine the range within which he expects payments to fall with a certain level of confidence. A good reserving model should be in a position to predict both the claim intensity (number of claims) and severity (amount paid on claims) of outstanding claims given the relevant data. It is hence a necessary requirement for insurance companies to hold "actuarially sound reserves."

1.2 **Motivation of the study.**

The greatest motivation to carry out this study was derived from the recent collapse of several insurance companies in Kenya with two companies being placed under receivership in the past twelve months. This has invariably raised questions on the stability, and hence solvency of other players in the industry which has dealt a heavy blow on the growth of the industry

1.3 Purpose of the study.

The purpose of this study is to address the inadequacy of loss reserving by the insurance industry in Kenya and to highlight actuarial techniques that can be applied to determine actuarially sound loss reserves to guarantee financial security.

1.4 Objectives of the study.

1. To determine the reserve to be maintained from a given data set.
2. To predict the estimate variance using stochastic techniques.
3. To highlight the major method for presenting data for the reserving exercise, i.e. the run-off triangle.

The various stochastic techniques demonstrated are:

- Negative Binomial Model
- The Mack's Model
- Normal Approximation to the Negative Binomial
- The Log-normal distribution
- The Hoerl curve
- The Wright's model

1.5 Significance of the study.

The study highlights stochastic methods available for adoption by the insurance companies to cushion them against cash-flow shocks arising from unanticipated claim liabilities.
It is hoped that an application of the techniques here-in would lead to:

- Adequate provision for outstanding liabilities arising from claims incurred but not yet reported or fully settled.
- Precise assessment of solvency and profitability.
- Continuous scientific evaluation of products.
- Correct pricing of insurance premiums and ratings.

CHAPTER 2

LITERATURE REVIEW

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 (*England and Verrall (2002)*). The financial condition of an insurance company cannot be adequately assessed without sound loss reserve estimates. Loss reserving is the term used to describe the actuarial process of estimating the amount of an insurance company's liabilities for loss and loss adjustment expenses. "Loss reserving is a major challenge to the casualty actuary because the estimation process not only involves complex technical tasks but considerable judgment as well. No formula will provide the correct answer."(*Wiser et al. - 2001*) Actuarial judgment is much needed since the mechanical application of stochastic and non-stochastic methods does not lead to a 'correct' result, and the result obtained will often need to be heavily qualified

The Casualty Actuarial Society defines an actuarially sound loss reserve as, "for a defined group of claims as of a given valuation date is a provision, based on estimates derived from reasonable assumptions and appropriate actuarial methods, for the unpaid amount required to settle all claims, whether reported or not, for which liability exists on a particular accounting date."

As our society is becoming more litigious there is an increase in the number and cost of new claims, making it essential for these claims to be correctly priced and reserved for. The reserving process provides insight into past claims performance and policy exposures and these can influence the terms and conditions offered on future business, including the basis of decisions to cease underwriting certain classes or to withdraw from insurance entirely in order to support alternative enterprises that may offer better rates of return on capital. Reserves that are conservative can lead to over-pricing, which may limit growth opportunities and establish a price umbrella for competitors. On the other hand, reserves that are deficient can lead to under-pricing, which may contribute to unprofitable growth. It is important to recognize both favorable and unfavorable development as quickly as possible, so that these inefficiencies are corrected. The reserves maintained by general insurance companies are usually not discounted providing a further safety margin.

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. The stochastic reserve estimate is expected to be *exactly* the same as that produced by the traditional deterministic chain ladder model.

The main preferences of stochastic models to traditional deterministic models for estimation of reserves are:

(i) The underlying distribution of the claims process can be fitted using the stochastic approach.

(ii) The likely variability of parameter estimates can be estimated.

(iii) With stochastic models one can estimate the goodness-of-fit of the model.

(iv) Allows smoothing of chain ladder development factors and estimation of tail factors

It is also useful to know where the data deviate from the fitted model, and to have a sound framework within which other models can be fitted and compared.
The Faculty and Institute of Actuaries reserving manual states that, "Deterministic reserving models are, broadly, those which only make assumptions about the expected value of future payments. Stochastic models also model the variation of those future payments. By making assumptions about the random component of a model, stochastic models allow the validity of the assumptions to be tested statistically, and produce estimates not only of the expected value of the future payments, but also of the variation about that expected value."

Since the prediction of outstanding or ultimate losses is a statistical problem, it is most helpful to formulate all methods in a statistical setting. This means that all losses are interpreted as random variables which are either observable or not. In particular, the data represented in a run-off triangle are interpreted as realizations of observable losses and the outstanding or ultimate losses are non-observable (except for the initial accident year). Stability in projections is to be sought by aiming to work with data groupings each containing a sufficient number of homogenous but independent risks on the assumption that they determine the characteristics of the resulting claims.

England and Verrall (2002) highlight the application of the negative binomial and the Normal approximation to the negative binomial models in determining reserve estimates and a prediction of errors. The negative binomial model is developed on the assumption that incremental claims D_{ij} are non-negative; this assumption is however relaxed in the normal approximation to the negative binomial model.

Mack (1993) also provides formulae for the prediction of errors of predicted payments and reserve estimates. Note that the mean and variance of C_{ij} under the Mack's model is similar to the mean and variance of C_{ij} in the Normal approximation to the negative binomial model, with the unknown scale parameters $\emptyset_j$ of the Normal approximation being replaced by σ_j^2 in Mack's model.

2.2 THE RUN-OFF TRIANGLE

Typically, data provided for a reserving exercise is in the form of a triangle of paid losses (see Table 1) in which the rows i denote accident years and the columns j delay or development years. Although we consider annual development here only, the methods can be extended easily to semiannual, quarterly or monthly development. The aim in reserving is to predict likely claim amounts in the missing southeast corner of the claims rectangle, the total reserve being the sum of these amounts. The run-off triangle forms the basis for most techniques used in the reserving exercise and it takes the form illustrated in table 1.

A year of account or origin year can be either:

(i) an underwriting year; that is, the calendar year in which cover commenced, or

(ii) an accident year; that is, the calendar year in which the event giving rise to a claim occurred, or

(iii) a reporting year, that is, the calendar year in which the event giving rise to a claim was reported.

Table 1.

Year of Origin	Development Year											
	1	2	...	k	...	$n-i$	...	$n-1$	...	n
1	$C_{1,1}$	$C_{1,1}$	...	$C_{1,k}$	...	$C_{1,n-i}$	...	$C_{1,n-1}$	...	$C_{1,n}$
2	$C_{2,1}$	$C_{2,1}$	...	$C_{2,k}$	...	$C_{2,n-i}$	...	$C_{2,n-1}$		
$\vdots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$				
i	$C_{i,1}$	$C_{i,1}$	...	$C_{i,k}$	...	$C_{i,n-i}$				
$\vdots$	$\vdots$	$\vdots$								
$n-k$	$C_{n-k,1}$	$C_{n-k,1}$	...	$C_{n-k,k}$						
$\vdots$	$\vdots$	$\vdots$								
$n-1$	$C_{n-1,1}$	$C_{n-1,1}$								
n	$C_{n,1}$									

The random variables $\{C_{i,k}\}_{i,k\epsilon\{0,1,\ldots n\}}$ are referred to as the incremental loss of accident year i and development year k.

The cumulative claims are defined by:

$$D_{ij} = \sum_{k=1}^{j} C_{ik} \tag{2.1}$$

The development factors are denoted by $\lambda_j : j = 2, \dots, n.$ and are estimated by:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} D_{ij}}{\sum_{i=1}^{n-j+1} D_{i,j-1}} \tag{2.2}$$

The development factor λ_j is weighted by the corresponding "volume" measure D_{ij-1}.

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 claim.

$$\widehat{D}_{i,n-i+2} = D_{i,n-i+1}\hat{\lambda}_{n-i+2}$$

$$\widehat{D}_{i,k} = \widehat{D}_{i,k-1}\hat{\lambda}_{k}, \qquad k = n - i + 3, n - i + 4, \dots, n.$$

By combining the development factors λ_j, an estimate of the ultimate total payments made is obtained for the year of origin j.

$$\beta_j = \prod_{z=j+1}^{n} \lambda_z \quad j = 0,1, \dots, (n-1) \tag{2.3}$$

Other methods have been proposed for estimating the development factors which include:

a) use of the log Normal distribution,
b) Geometric distribution,
c) Exponential Decay , by fitting an exponential curve of the form, $y = ae^{bt}$
d) Fitting the logarithmic curve of the form $y = a + b\ln t$
e) The inverse power curve (Richard Sherman, *Extrapolating, smoothing and Interpolating Development factors).*

2.3 THE NEGATIVE BINOMIAL MODEL

Assumption.

- The sum of incremental claims in column j should be positive, i.e. $\lambda_j > 1$

The model is similar to the Chain-ladder in that the parameters appear to be more 'like' the chain-ladder development factors. The Negative Binomial Model can be expressed as a model of either incremental or cumulative claims. C_{ij} has an over-dispersed negative binomial distribution, with mean and variance:

$$(\lambda_j - 1)D_{i,j-1} \text{ and } \emptyset\lambda_j(\lambda_j - 1)D_{i,j-1}, \text{ respectively.}$$

where λ_j is similar to the standard chain ladder development factor. *(P. England and R. Verrall-2002)*

We can write this model in terms of cumulative claims assuming that $D_{i,j-1}$ is known;

$$D_{i,j} = D_{i,j-1} + C_{ij}$$

where D_{ij} has an over-dispersed negative binomial distribution with mean and variance given as:

$$\lambda_j D_{i,j-1} \text{ and } \emptyset\lambda_j(\lambda_j - 1)D_{i,j-1} \text{ respectively.}$$

The variance holds only where $D_{i,j-1}$ is known. A recursive approach is required in approximating the variance, whereby, the calculation of the process variance involves estimating the variance of a k-steps-ahead forecast, using standard results from the analysis of conditional distributions. Since the negative binomial model is derived from the Poisson model, the predictive distributions are essentially the same, and give identical predicted values. The Negative Binomial model restricts itself to positive incremental claims. 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 due 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 since negative values are less likely to appear.

Negative Binomial Prediction Errors

The parameters in the negative binomial model relate to development years only.

Considering the model for incremental data:

Mean; $\mathrm{E}(C_{ij}) = (\lambda_j - 1)D_{i,j-1}$ and,

Variance; $\mathrm{Var}(C_{ij}) = \phi\lambda_j(\lambda_j - 1)D_{i,j-1}$

where the D_{ij} (where observed) are considered known.

Then, writing:

$$\mathrm{E}(C_{ij}) = m_{ij} = (\lambda_j - 1)D_{i,j-1} \tag{2.4}$$

and taking logs gives;

$$\log(m_{ij}) = \log(\lambda_j - 1) + \log(D_{i,j-1})$$

Writing;

$$\log(\lambda_j - 1) = c + \alpha_{j-1} \qquad \text{with: } \alpha_1 = 0,\ j \geq 2 \tag{2.5}$$

gives:

$$\log(m_{ij}) = c + \alpha_{j-1} + \log(D_{i,j-1}) \tag{2.6}$$

Equations 2.4, 2.5 and 2.6 specify a generalised linear model with logarithmic link function and negative binomial error structure. The $\log(D_{i,j-1})$ terms are derived from the known values $D_{i,j-1}$, and are specified as offsets in the model.

Estimates of the development factors can be obtained from the parameter estimates using equation 2.5, and their approximate standard errors can be obtained using:

$$\mathrm{Var}(\hat{\lambda}_j) = \mathrm{Var}(\hat{\lambda}_j - 1) \approx \exp(\hat{c} + \hat{\alpha}_{j-1})^2 \mathrm{Var}[\hat{c} + \hat{\alpha}_{j-1}] \quad j \geq 2$$

For the origin year reserve estimates, the ultimate claims, U_i, are the cumulative claims in the last development year. Hence,

$$U_i = D_{in} \tag{2.7}$$

The reserve estimate in origin year i, R_i, is $U_i - D_{i,n-i+1}$, where $D_{i,n-i+1}$ is the paid to date which is considered known. Therefore,

$$\mathrm{Var}[R_i] = \mathrm{Var}[U_i] = \mathrm{Var}[D_{in}] \tag{2.8}$$

and

$$\mathrm{Var}[\hat{R}_i] = \mathrm{Var}[\hat{U}_i] = \mathrm{Var}[\hat{D}_{in}] \tag{2.9}$$

The origin year process and estimation variances can be estimated by considering $\text{Var}[D_{in}]$ and $\text{Var}[\hat{D}_{in}]$ respectively.

Firstly, consider the process variance, $\text{Var}[D_{in}]$. Recursive procedures are required to obtain the estimation variance and the process variance since the negative binomial model is formulated as a recursive model. The calculation of the process variance involves estimating the variance of a k-steps ahead forecast, where $k = i - 1$.

For a negative binomial model,

$$\text{Var}[D_{in}] \approx D_{i,n-i+1} \prod_{k=n-i+2}^{n} \hat{\lambda}_k \left(\prod_{k=n-i+2}^{n} \hat{\lambda}_k - 1\right) \tag{2.10}$$

The estimation variance is calculated from

$$\text{Var}[D_{in}] \approx \text{Var}\left[D_{i,n-i+1} \prod_{k=n-i+2}^{n} \hat{\lambda}_k\right] = D_{i,n-i+1}^2 \text{Var}\left[\prod_{k=n-i+2}^{n} \hat{\lambda}_k\right] \tag{2.11}$$

Note that, the variance of the product of development factors in equation (2.11) can be calculated, since the parameters in the negative binomial model relate to the development factors λ_j, and the covariance matrix of the parameter estimates is readily available.

The overall reserve estimation and process variances can be estimated by considering $\text{Var}[R_+]$ and $\text{Var}[\hat{R}_+]$ respectively, where

$$R_+ = \sum_{i=2}^{n} R_i \tag{2.12}$$

The overall reserve process variance is the sum of the process variances of individual origin year reserves, assuming independence between years. The overall reserve estimation variance is given by,

$$\text{Var}[\hat{R}_+] \approx \sum_{i=2}^{n} \text{Var}[\hat{D}_{in}] + 2 \sum_{\substack{i=2 \\ j>i}}^{n} Cov[\hat{D}_{in}, \hat{D}_{jn}] \tag{2.13}$$

Hence, the estimation variance of overall reserves is the sum of the estimation variances of individual origin year reserves, with an additional component to take account of the covariance between years induced by dependence on the same parameters.

2.4 NORMAL APPROXIMATION TO THE NEGATIVE BINOMIAL MODEL.

When the assumption made above (positivity) on the Negative Binomial model doesn't hold, the model breaks down (England and Verrall 2002). To avoid such an eventuality, it is necessary to use a model that is not restrict to the positive real line and the Normal distribution is most suitable.

The Normal Distribution is used as an approximation of the Negative Binomial model with the mean remaining invariably unchanged and an adjustment being made on the variance to accommodate negative values. The Normal approximation for the distribution of incremental claims C_{ij} is approximately Normally distributed, with mean and variance:

$$(\lambda_j - 1)D_{i,j-1} \text{ and } \emptyset_j D_{i,j-1}, \text{ respectively.}$$

while D_{ij} is approximately Normally distributed, with mean and variance:

$$\lambda_j D_{i,j-1} \text{ and } \emptyset_j D_{i,j-1} \text{ respectively.}$$

The variance also holds if $D_{i,j-1}$ is known and is modelled as part of an extended fitting procedure known as joint modelling.

2.5 Mack's Model.

The model was proposed by Mack (1993) and made limited assumptions as to the distribution of underlying data by specifying the mean and variance of D_{ij} to be:

$$\lambda_j D_{i,j-1} \text{ and } \sigma_j^2 D_{i,j-1} \text{ respectively.}$$

Mack also produced estimators of the unknown parameters λ_j and σ_j^2 with further limited assumptions using:

$$\hat{\lambda}_j = \frac{\sum_{i=1}^{n-j+1} w_{ij} f_{ij}}{\sum_{i=1}^{n-j+1} w_{ij}} \quad \text{where:} \quad w_{ij} = D_{ij} \quad \text{and} \quad f_{ij} = \frac{D_{ij}}{D_{i,j-1}}$$

and

$$\hat{\sigma}_{ij}^2 = \frac{1}{n-j} \sum_{i=1}^{n-j+1} w_{ij}(f_{ij} - \hat{\lambda}_j)^2.$$

Mack also shows that the weighted average of individual development factors is preferable to the unweighted average since it has a lower variance, that is; it is the minimum variance unbiased estimator (UMVUE).

The variance component σ_j^2 is estimated as an average of weighted residuals, where the divisor is the number of residuals (used in calculating the estimator) minus one. The one is subtracted to provide an unbiased estimator of σ_j^2. These variance components are not used when estimating the development factors, but are required when considering the prediction errors of future payments. Like the Normal approximation to the negative binomial model, there is insufficient information to estimate the final variance component σ_n^2. Mack (1994b) proposed a method of estimating σ_n^2 by setting $\hat{\sigma}_n^2 = \hat{\sigma}_{n-1}^2$ and $\hat{\sigma}_n^2 = \hat{\sigma}_{n-2}^2$ and it was observed that the results in the earlier years are very sensitive to this single parameter.

Mack (1993) also provides formulae for the prediction of errors of predicted payments and reserve estimates. Note that the mean and variance of C_{ij} under the Mack's model is similar to the mean and variance of C_{ij} in the Normal approximation to the negative binomial model, with the unknown scale parameters $\emptyset_j$ of the Normal approximation being replaced by σ_j^2 in Mack's model.

According to *Mack (1993)*, the process variance of the origin year reserves is given by,

$$\mathrm{Var}[\hat{R}_i] \approx \hat{D}_{in}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \hat{D}_{ik}} \tag{2.14}$$

The estimation variance is given by,

$$\mathrm{Var}[\hat{R}_i] \approx \hat{D}_{in}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \sum_{q=1}^{n-k} \hat{D}_{qk}} \tag{2.15}$$

Therefore,

$$MSEP[\hat{R}_i] \approx \hat{D}_{in}^2 \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \hat{D}_{ik}} \left(\frac{1}{\hat{D}_{ik}} + \frac{1}{\sum_{q=1}^{n-k} \hat{D}_{qk}} \right) \tag{2.16}$$

For the overall reserve prediction error, a covariance adjustment is needed for the estimation variance, giving

$$MSEP[\hat{R}_+] = \sum_{i=2}^{n} \{ MSEP[\hat{R}_i] + \hat{D}_{in} \left(\sum_{q=i+1}^{n} \hat{D}_{qk} \right) \mathrm{x} \sum_{k=n-i+1}^{n-1} \frac{\hat{\sigma}_{k+1}^2}{\hat{\lambda}_{k+1}^2 \sum_{q=1}^{n-k} \hat{D}_{qk}} \} \tag{2.17}$$

Where,

$$\hat{R}_+ = \sum_{i=2}^{n} \hat{R}_i \tag{2.18}$$

2.6 OTHER MODELLING DISTRIBUTIONS

2.6.1 Log-Normal Distribution.

Early stochastic models for the chain-ladder technique focused on the logarithm of the incremental claims amounts $Y_{ij} = \log(C_{ij})$ and the log-normal class of models of the form;

$$Y_{ij} = m_{ij} + \varepsilon_{ij}$$

where; ε_{ij} is an independent random error. i.e.

$$\varepsilon_{ij} \sim \mathrm{IN}(0, \sigma^2) \text{ or } Y_{ij} \sim \mathrm{IN}(m_{ij}, \sigma^2)$$

A limitation is imposed in this class of models, in that incremental claim amounts have to be positive. Using the chain-ladder type structure, the mean is given as:

$$n_{ij} = c + \alpha_i + \beta_j$$

The log-Normal distribution has the advantage that it can be implemented without the need for specialist software; it also allows different assumptions to be incorporated concerning the run-off pattern and the connections between the origin year.

2.6.2 The Hoerl Curve

For models with a Log link function such as the Log-Normal model, over-dispersed poisson model and the Gamma model, the Hoerl Curve is provided by replacing the chain-ladder type linear prediction which has a parameter for each development year with:

$$n_{ij} = c + \alpha_i + \beta_j \log(j) + y_i j \qquad (j > 0). \tag{2.19}$$

The parametric form of the model on the untransformed scale can be seen by exponentiating equation (2.19), giving:

$$\exp(n_{ij}) = \mathrm{A}_i j^{\beta_j} e^{y_i j} \qquad \text{where: } \mathrm{A}_i = \exp(c + \alpha_i)$$

Here, development year j is considered as a continuous covariate, and the run-off pattern follows a fixed parametric form, being linear in development time and log development time on a log scale. The advantage of working on a log scale is that parameters can be readily estimated whereas the advantage of treating development time as a continuous covariate is that extrapolation is possible beyond the range of development time observed. This helps in estimating the tail factor.

2.6.2 Wright's Model

The stochastic model was proposed by Wright (1990) whereby the systematic and random components of the underlying model were based on a risk theoretic model of the claims generating process. Wright considered the incremental paid claims C_{ij} to be the sum of N_{ij} (independent) claims of amount X_{ij}. The claim numbers N_{ij} were assumed to be Poisson random variables, where:

$$\mathrm{E}[N_{ij}] = e_i a_j k_i j^{A_i} \exp(-b_i j)$$

and

$$\mathrm{Var}[N_{ij}] = \mathrm{E}[N_{ij}]$$

where k, A and b are unknown constants to be estimated, e is a measure of exposure, and a is a *known* adjustment term needed on technical grounds.
The values a are specified (in Appendix 1 of Wright) for each value of j. It is important to note that the variance of the number of claims equals the mean as a result of using the Poisson distribution.

CHAPTER 3:

DATA ANALYSIS

3.1 Introduction

The data used in this study (Table 3-1) is similar to the data used by P.England and R. Verall(2002). It refers to incurred claim amounts on general liability during each development year, before adjusting for inflation, the year of origin being the period in which the claims occurred. It is assumed that the claims are fully run-off in the first year of origin. Thus, when estimating the reserves, we do not consider claims that may have arisen after the last year of development.

The data was analysed using two stochastic models: that is, The Mack's model and the Negative Binomial model.

3.2 Mack's Model.

The data presented in Table 3-1 is to be analysed using the Mack's model.

Table 3-1: Incremental Paid Claims

Year of	Development years									
Origin	1	2	3	4	5	6	7	8	9	10
1	5,012	3,257	2,638	898	1,734	2,642	1,828	599	54	172
2	106	4,179	1,111	5,270	3,116	1,817	(103)	673	535	
3	3,410	5,582	4,881	2,268	2,594	3,479	649	603		
4	5,655	5,900	4,211	5,500	2,159	2,658	984			
5	1,092	8,473	6,271	6,333	3,786	225				
6	1,513	4,932	5,257	1,233	2,917					
7	557	3,463	6,926	1,368						
8	1,351	5,596	6,165							
9	3,133	2,262								
10	2,063									

The cumulative values and development factors are shown in Table 3-2, are calculated in the same way as the deterministic chain-ladder. The development factors are arrived at by applying equation 2.1

Table 3-2: Cumulative Paid Claims

Year of	Development years									
Origin	1	2	3	4	5	6	7	8	9	10
1	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
2	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
3	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
4	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
5	1,092	9,565	15,836	22,169	25,955	26,180				
6	1,513	6,445	11,702	12,935	15,852					
7	557	4,020	10,946	12,314						
8	1,351	6,947	13,112							
9	3,133	5,395								
10	2,063									

An application of equation 2.3 gives the ultimate claims and hence the reserves to be maintained for each origin year as outlined in Table 3-3.

Table 3-3: Ultimate claims and respective reserves

Year of	Development years											
Origin	1	2	3	4	5	6	7	8	9	10	Ultimate	Reserve
1	5.012	8,269	10.907	11.805	13,539	16,181	18,009	18.608	18.662	18,834	18.834	-
2	106	4,285	5.396	10,666	13.782	15,599	15,496	16,169	16,704		16.858	154
3	3,410	8,992	13.873	16,141	18,735	22,214	22.863	23,466			24,083	617
4	5,655	11,555	15,766	21,266	23,425	26,083	27,067				28,703	1,636
5	1.092	9,565	15,836	22,169	25,955	26,180					28,927	2,747
6	1,513	6,445	11.702	12,935	15,852						19,501	3,649
7	557	4,020	10.946	12,314							17,749	5,435
8	1,351	6,947	13,112								24,019	10,907
9	3,133	5,395									16,045	10,650
10	2,063										18,402	16,339
											Overall	**52,135**

The reserve to be maintained, assuming that the first year of origin is fully run-off, is 52,135.

Table 3-4 below shows the development factors and the variance component which were arrived at by applying Equation 2.15 in a Microsoft Excel spreadsheet.

Table 3-4: Development Factors and Variance Components

	Lambda	sigma squared
j=2	2.999	27,883.48
j=3	1.624	1,108.53
j=4	1.271	691.44
j=5	1.172	61.23
j=6	1.113	119.44
j=7	1.042	40.82
j=8	1.033	1.34
j=9	1.017	7.88
j=10	1.009	

Since 10 is the last development year, the variance of the estimate of development factors cannot be determined. The variance of the last development factor is hence approximated by setting either $\hat{\sigma}_{10}^2 = \hat{\sigma}_8^2$ or $\hat{\sigma}_{10}^2 = \hat{\sigma}_9^2$.

Tables 3-5 and 3-6 show the prediction error of the Mack's model determined by setting $\hat{\sigma}_{10}^2$ equal to $\hat{\sigma}_8^2$ and $\hat{\sigma}_9^2$, respectively.

Table 3-5: Reserve results setting $\hat{\sigma}_{10}^2 = \hat{\sigma}_8^2$

Year of Origin	Reserve	Prediction error	Prediction error %
2	154	206	134%
3	617	623	101%
4	1,636	747	46%
5	2.747	1,469	53%
6	3,649	2,002	55%
7	5,435	2,209	41%
8	10,907	5,358	49%
9	10,650	6,333	59%
10	16,339	24,566	150%
Overall	**52,135**	**26,909**	**52%**

Table 3-6: **Reserve results setting $\hat{\sigma}_{10}^2 = \hat{\sigma}_9^2$**

Year of Origin	Reserve	Prediction error	Prediction error %
2	154	500	325%
3	617	863	140%
4	1,636	1,014	62%
5	2,747	1,623	59%
6	3,649	2,065	57%
7	5,435	2,259	42%
8	10,907	5,391	49%
9	10,650	6,348	60%
10	16,339	24,571	150%
Overall	**52,135**	**27,172**	**52%**

3.3 The Negative Binomial Model

The negative binomial model can be fitted using incremental or cumulative data, and gives the same fitted values irrespective of which method is used. However, in this study, parameters have been estimated using cumulative paid data, in Table 3-2.

Table 3-7 gives the estimated parameters and the standard error.

Table 3-7: Parameter estimates

	Parameter Estimate	Standard Error
Constant	0.693	0.027
Alpha 2	-0.472	0.349
Alpha 3	-1.306	0.368
Alpha 4	-1.762	0.391
Alpha 5	-2.177	0.432
Alpha 6	-3.172	0.642
Alpha 7	-3.403	0.821
Alpha 8	-4.078	1.395
Alpha 9	-4.687	2.540

The development factors and their approximate standard errors are given in Table 3-8.

Table 3-8: Development factors and standard errors

	Estimate	Std error
Lambda 2	2.999	0.546
Lambda 3	1.624	0.135
Lambda 4	1.271	0.067
Lambda 5	1.172	0.048
Lambda 6	1.113	0.038
Lambda 7	1.042	0.024
Lambda 8	1.033	0.026
Lambda 9	1.017	0.023
Lambda 10	1.009	0.023

An application of the development factors to the cumulative data (Table 3-2) to obtain future cumulative payments shown in Table 3-9

Table 3-9: Future cumulative payments

Year of	Development years									
Origin	1	2	3	4	5	6	7	8	9	10
1	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
2	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	**16,858**
3	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466	**23,863**	**24,083**
4	5,655	11,555	15,766	21,266	23,425	26,083	27,067	**27,967**	**28,441**	**28,703**
5	1,092	9,565	15,836	22,169	25,955	26,180	**27,278**	**28,185**	**28,663**	**28,927**
6	1,513	6,445	11,702	12,935	15,852	**17,649**	**18,389**	**19,001**	**19,323**	**19,501**
7	557	4,020	10,946	12,314	**14,428**	**16,064**	**16,738**	**17,294**	**17,587**	**17,749**
8	1,351	6,947	13,112	**16,664**	**19,525**	**21,738**	**22,650**	**23,403**	**23,800**	**24,019**
9	3,133	5,395	**8,759**	**11,132**	**13,043**	**14,521**	**15,130**	**15,634**	**15,898**	**16,045**
10	2,063	**6,188**	**10,046**	**12,767**	**14,959**	**16,655**	**17,353**	**17,931**	**18,234**	**18,402**

An estimation of the future claim payments enables us to estimate the ultimate claims and hence the reserve to be maintained for each year of origin. The reserves to be maintained are shown in Table 3-10.

Table 3-10: Estimated Reserve

Year of Origin	Reserve
1	-
2	154
3	617
4	1,636
5	2,747
6	3,649
7	5,435
8	10,907
9	10,650
10	16,339
Total	**52,135**

Upon determination of the reserve estimate, it is imperative to determine its degree of precision by estimating the variance. An application of recursive models (*Appendix 1, P. England and R. Verrall – 2002*) on the data gives the prediction error and the prediction variance. The prediction error is outlined in Table 3-11.

Table 3-11: Prediction error

Year of Origin	Reserve	Prediction error	Prediction error %
2	154	566	367%
3	617	1,139	185%
4	1,636	1,807	110%
5	2,747	2,271	83%
6	3,649	2,483	68%
7	5,435	3,180	59%
8	10,907	5,122	47%
9	10,650	6,185	58%
10	16,339	13,227	81%
Overall	**52,135**	**18,528**	**36%**

The prediction variance as outlined in Table 3-12 is a sum of estimation variance and process variance which are also determined recursively. However, the overall estimation variance is determined differently by applying the covariance matrix of the development factors to equation 2.13. (*Appendix 2, P. England and R. Verrall – 2002*)

Table 3-12: Prediction Variance

Year of Origin	Estimation Variance	Process Variance	Prediction Variance
2	139	155	294
3	561	634	1,195
4	1,268	1,735	3,003
5	1,709	3,035	4,744
6	1,184	4,489	5,673
7	1,470	7,834	9,304
8	4,160	19,980	24,140
9	3,524	31,674	35,197
10	15,234	145,754	160,987
Overall	**100,598**	**215,288**	**315,886**

CHAPTER 4:

SUMMARY, CONCLUSION AND RECOMMENDATIONS.

4.1 Summary of the Study

A Comparison of Mack's Model and the Negative Binomial Model.

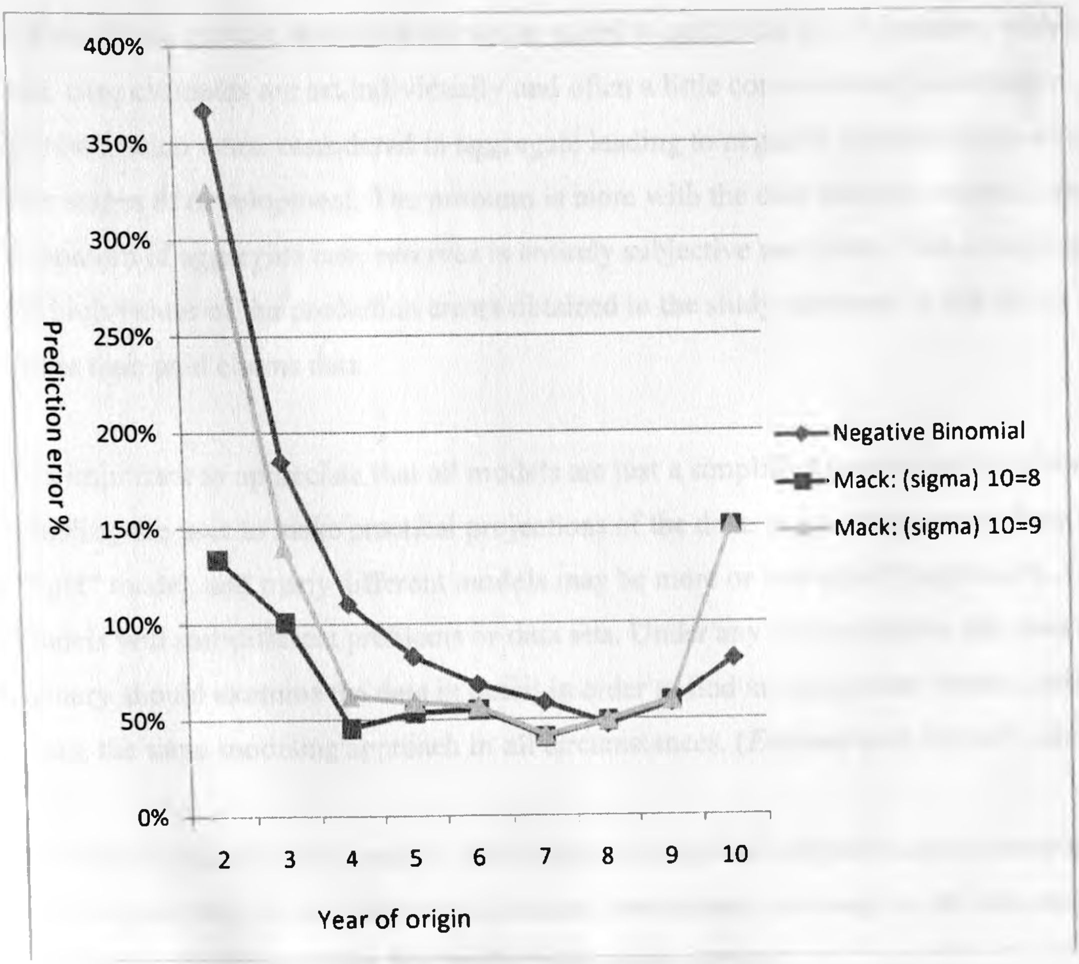
The estimates of the development factors are identical for both models. This implies that the reserve to be maintained thereof is also similar for both models.
A comparison of Tables 3-5, 3-6 and 3-11 reveals that the models have great similarity, given that these are raw results without any bias correction. The results of the above named tables are summarized in Table 3-13.

Table 3-13: A comparison of Prediction errors

Year of Origin	Negative Binomial	Mack: $\hat{\sigma}_{10}^2 = \hat{\sigma}_8^2$	Mack: $\hat{\sigma}_{10}^2 = \hat{\sigma}_9^2$
2	367%	134%	325%
3	185%	101%	140%
4	110%	46%	62%
5	83%	53%	59%
6	68%	55%	57%
7	59%	41%	42%
8	47%	49%	49%
9	58%	59%	60%
10	81%	150%	150%
Overall	**36%**	**52%**	**52%**

A graphical representation of the above data in Fig. 1 illustrates the similarity of the models further. The negative binomial model is quite similar to Mack's model with $\hat{\sigma}_{10}^2 = \hat{\sigma}_9^2$ up to the 8[th] development year. The results of work done by P. England and R. Verrall (2002) confirm that, upon correction for biasness, the two models are essentially the same.

Fig. 1: A graphical representation of prediction error for the Negative Binomial and Mack Models.



4.2 CONCLUSION.

The assessment of the financial strength of a general insurance company includes a thorough analysis of the outstanding claims reserves, including an assessment of the possible variability in the reserves. Any failure to do so will result in the insolvency of some insurers, as it has been witnessed in the recent past. It is important that the best method for reserve estimation is chosen in order to avoid any negative effects in the financial position of the company.

The stochastic models described are better suited to paid data; this is because, with incurred data, case estimates are set individually and often a little conservatively resulting in overestimation when considered in aggregate leading to negative incremental amounts in the later stages of development. The problem is more with the data than the methods, since, the estimation of aggregate case reserves is entirely subjective and faulty. This is highlighted by the high values of the prediction errors obtained in the study attributed to the use of incurred rather than paid claims data.

It is important to appreciate that all models are just a simplified representation of reality, enabling the user to make practical projections of the data. As a consequence, there is no one "right" model, and many different models may be more or less equally applicable. Different models will suit different problems or data sets. Under any circumstances, the reserving actuary should examine the data in detail in order to find an appropriate model, rather than using the same modeling approach in all circumstances. (*England and Verrall – 2002*)

The claim process across years of development should be analysed to spot discrepancies which could alter the development of claims. For-instance, a change in the time required for processing claims may alter the development across periods.

4.3 RECOMMENDATIONS.

This study considered undiscounted claims reserve estimate, discounting of the reserves may be considered due to their significance in the long tailed classes of business. The written premiums are expected to be invested for a given period of time before claims are made and subsequently paid; however, short-tailed classes of business are not greatly affected by discounting.

An application of smoothing models for the development factors should be considered since their application allows extrapolation for the estimation of tail factors.

The industry regulators could do more to secure the solvency of the industry by proposing legislation that would make it mandatory for general insurance industry to maintained "best estimate" reserves at defined confidence levels. Where a "best estimate" is intended to represent: "the expected value of the distribution of possible outcomes of the unpaid liabilities".

4.4 LIMITATIONS OF THE STUDY.

The data considered in this study relates to incurred claims composed of case estimates that are set individually and often a little conservatively resulting in overestimation when considered in aggregate leading to negative incremental amounts in the later stages of development.

Time taken to conduct the study was limited and as such, the models applied were fewer than had originally been intended.

References:

England, P.D., Verrall, R.J. (2002). Stochastic Claims Reserving in General Insurance. *British Actuarial Journal*, vol.8, pages 443-518.

Faculty of Actuaries; Institute of Actuaries. (1989). *Claims Reserving Manual.* Vol. 1, The Faculty and Institute of Actuaries.

Faculty of Actuaries; Institute of Actuaries. (1997). *Claims Reserving Manual.* Vol. 2, The Faculty and Institute of Actuaries.

Mack, T. (1993). Distribution-Free Calculation of the Standard Error of Chain-Ladder Reserve Estimates, *ASTIN Bulleting.* vol.23, pages 213-225.

Mack, T. (1994b). Measuring the Variability of Chain-Ladder Reserve Estimates, *Casualty Actuarial Society,* Spring Forum.

Mack, Thomas, "The Standard Error of Chain Ladder Reserve Estimates: Recursive Calculation and Inclusion of a Tail Factor," *ASTIN Bulletin*, 1999, Vol. 29, No. 2, pp. 361-366.

Schmidt, K. D., and Schnaus, A. [1996] An extension of Mack's model for the chain ladder method. *ASTIN Bulletin.* Vol. 26, pp. 247-262.

Wiser, R.F.; Cockley, J.E; and Gardner A., "Loss Reserving," *Foundations of Casualty Actuarial Science* (Fourth Edition), Casualty Actuarial Society. 2001

The Casualty Actuarial Society (CAS) website: http://www.casact.org