CLAIM RESERVING USING TWEEDIE DISTRIBUTION

MBAUNI HARUN NDEGWA

I56/70320/2011

School of Mathematics

University of Nairobi

A research project submitted in partial fulfillment of the requirement for the degree of Master of Science in Actuarial Science of the University of Nairobi.

JULY 2013

DECLARATION

I declare that ***Claim Reserving using Tweedie Distribution*** is my own original work. This research project has never been presented for examination at any of the learning institution/University whether in Kenya or elsewhere as per my own knowledge and understanding. All the sources quoted have been indicated and acknowledged with complete reference.

STUDENT

MBAUNI HARUN NDEGWA

SIGN: ---------- DATE: ----------

This project has been submitted for examination with the approval of the following as University supervisor.

SUPERVISOR

Prof. JAM OTTIENO

SIGN: ---------- DATE: ----------

Prof. P.O. Weke

SIGN: ---------- DATE: ----------

ACKNOWLEDGEMENT

I wish to thank in particular my supervisors Prof. P.O WEKE and Prof. JAM OTTIENO, School of Mathematics for their priceless support without which this project would have not been successful. Special thanks Mr. Calvin Maina for regularly assisting whenever there was need.

DEDICATION

This project is dedicated to my family members for their infinity support and guidance.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Fitting mathematical regression models to insurance claims data has always been challenging. The problem is predominantly grave for data from individual policies where most of the losses are zero. In addition, those policies with anaffirmative loss, the losses are highly skewed. Most of the traditional regression models do not deal with a mixture of discrete losses of zero and continuous positive losses. One way of dealing with this problem is to fit separate models to the frequency and severity. We address this problem using a new stochastic model called the Tweedie distribution model to derive estimates of outstanding Claims liabilities which are close to the chain ladder estimates. We take a broad view the gamma cell distributions model which leads to Tweedie's compound Poisson model. Choosing a suitable parameterization, we estimate the parameters of our model within the framework of generalized linear model. We show that these methods lead to rational estimates of the outstanding claims liabilities.

CHAPTER 1

1.1: Introduction

Insurance companies are different from other industries since they are involved in offering promises rather than selling products. An insurance policy can be defined as a promise made by the insurer to the policy holder forthcoming claims for a received premium. Resulting from this, the insurers do not know the upfront cost of their service. Instead, they rely on past data analysis and judgment to obtain a justifiable price for their offering. When it comes to General Insurance most policies go for a period of approximately twelve months. Nonetheless, the claim payment process can run for years or decades. Therefore, not even the insurers are certain about the delivery date of the product. In general insurance, claims resulting from physical damage to (property including cars and buildings) and theft are reported and resolved relatively fast. Conversely, in some areas of general insurance there are significant delays between the determination of the exact amount the company will pay in settlement and the time of a claim prompting event. When an episode leading to a claim takes place, it may experience delays in reporting. In the case of an accident the incident may be quickly reported, but it may take some time before it is determined actually who is liable and to what extent. In some situation, one might have to wait for the outcome of legal action before damages can be properly ascertained.

It is important for an insurance company to know what to set aside in reserves at regular intervals so as to handle claims arising from episodes that have already taken place but for which it cannot actualize the extent of its liability. Claims resulting from episodes that have taken place but have not nonetheless been reported to the insurer are referred to as Incurred But Not Reported (IBNR) claims and a reserve set aside for these claims is called an IBNR Reserve. The most important item on the insurer's balance sheet liability side is provision or reserves for yet to come claims payments. These reserves can be broken down into Incurred But Not Reported (IBNR) claims and case reserves (outstanding claims) which are losses reported to the insurer. Overtime, numerous methods have been developed to approximate reserves for claims. Changes in regulatory conditions for instance, Solvency II in Europe have steered further research and development into topic with an emphasis on statistical

and stochastic techniques. Certainly, one of the common techniques for approximating reserves is the Chain Ladder Method. This method looks at how claims resulting from different origin years have developed over later development years and then use related ratios to forecast how future claims from these years will change.

The historical numerical integration and acceleration methods of the Chain Ladder are discussed in Chapter 2. Chapter 3 derives the Tweedie compound Poisson distribution models as a subclass of the exponential dispersion models. Implementation details are given in Chapter4.

CHAPTER 2

2.1: Literature Review

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. In addition, Taylor (1986), England and Verrall (2002) and Wüthrich and Merz (2008) give valuable general ideas of stochastic claims reserving. Many models have been proposed to deal with data which has been amassed in some way as this makes presentation more convenient. Yet, this aggregation of data may result to loss of information which in some instances can lead to relatively poor prediction and estimation of outstanding liabilities. This has been a prominent subject in some of the previous works on reserving. For example, Taylor and McGuire (2004) apply a generalized linear model outline to model the features of individual claims. Norberg (1999) put forward a structure for the claims occurrence, reporting and payment system at an individual level. Also related to this context is Norberg (1986). Models related to individual claims are likely to be very intricate and detailed and also use extensive data to approximate parameters. However, for the practicing actuary, there are certain limitations.

First, there is difficulty in implementation because the use of data at a personal level specifically computationally challenging. Second, elaborate and extensive data sets are usually hard to obtain from insurance companies and in most cases the model will only be applicable for practical situations if it can be used in a wide variety of data sets across a broad spectrum of business lines. From this it can be seen that there is a difficult choice to be made on whether to use individual data which is theoretically interesting but computationally complex or whether to use aggregate data which is better to deal with but from which some information has been missing.

The Chain ladder technique is well known, easy to apply and widely used in claim reserving. Many a times, the chain ladder method is the first technique to be applied, seconded by manual smoothing of the resulting development factors. This is followed by adjustment of the results in line with an actuary's opinion combined with additional information.

There has been a lot literature investigating the statistical origin of the Chain ladder method, notably, Mack & Venter (2000), England &Verrall (1999, 2001), Barnett & Zehnwirth (1998), Renshaw &Verrall (1998), Schmidt & Schnaus (1996), Murphy (1994), Mack (1993, 1994a, 1994b), Verrall (1989, 1990, 1991a, 1991b, 1994, 1996, 2000), Renshaw (1989), Taylor & Ashe (1983), and Kremer (1982). These literatures have made noteworthy advances in the understanding of the chain ladder technique, and this paper's objective is to convey this together in a suitable form, and show the possibilities of extension of the model. The primary focus is on stochastic methods founded on the framework of generalized linear models. Verrall & Renshaw (1998) were not the principal scholars to notice the relation between the chain ladder method and the Poisson distribution; however, they were the first to implement the model using typical procedures in statistical modelling, and to provide a relation with the scrutiny of contingency tables.

A similar model is described by Wright (1990). It includes a term to model claims inflation, but he did not consider the model in detail. Mack (1991) also arguments out that the chain ladder estimates can be found by maximizing Poisson likelihood by alluring to the so called `method of marginal totals'. A discussion of the stochastic origin of chain ladder models can be found in Verrall& England (2000), Mack & Venter (2000), Verrall (2000), and Mack (1994a). The relationship between the various models and whether they can admissibly be used to increase value to the deterministic chain ladder technique is the core of the discussion. Questions raised include the subject of whether incremental or cumulative claims should be modeled, the number of parameters estimated and what transpires when there are negative incremental claims. We are optimistic that, by setting out the models and so long as we give a set of examples in later sections, we are able to elucidate some of these concerns. Some of the questions raised are inapt.

This paper exhibits that the Chain Ladder method can be conveyed to a successful conclusion by linking advanced integration ways and means with some analytic analysis of the integrand. The method is used to assess the densities of an imperative class of distributions produced by exponential families. An exponential dispersion model is a two parameter family of distributions entailing of a linear exponential

family with an added dispersion parameter. Exponential dispersion models are significant in statistics because they are the retort distributions for generalized linear models (McCullagh & Nelder, 1989). Jørgensen (1987,1997) established exponential dispersion models as a field of study in its own right. He undertook a detailed study of their properties. An exponential dispersion model is characterized by its variance function $V(\mu)$. The variance function describes the mean variance association of the distribution when the dispersion parameter is held constant. If Y follows an exponential dispersion model distribution with mean μ and variance function $V(\mu)$ then $\mathrm{Var}(Y) = \phi V(\mu)$ where ϕ is the dispersion parameter. Exponential dispersion models with significant impact are those with power mean–variance relations, $V(\mu) = \mu^p$ for some p. Following Jørgensen (1987,1997), we refer these models as Tweedie distribution models. The class of Tweedie distribution models comprises most of the key distributions linked to generalized linear models, notably, normal distribution $(p = 0)$, Poisson distribution $(p = 1)$, gamma distribution $(p = 2)$ and the inverse Gaussian distribution $(p = 3)$. Though the other Tweedie distributions models are less well known, Tweedie models exist for values of p outside the interval (0, 1). Tweedie distributions with $p > 1$ have non-negative support values and positive means, $\mu > 0$.If $p > 2$, the Tweedie distributions are generated by stable distributions . The Tweedie distributions for $1 < p < 2$ can be epitomized as mixtures of Poisson and gamma distributions which have a mass at zero though they are continuous on the positive reals. Distributions with $p < 0$ are rare in that they support whole real line but they have positive means.

Tweedie distribution models are the only exponential dispersion models which are closed under rescaling and are natural for modeling continuous positive quantity data with a random measurement scale. The distributions with $1 < p < 2$ are particularly attractive for modeling quantity data where exact zeros are probable. Dunn and Smyth (2005) provided a survey of in print applications showing that Tweedie distribution models have been used in an assorted range of fields including meteorology, actuarial studies, consumption studies, assay analysis, time & expense studies and survival analysis. Latest applications include rainfall prediction (Dunn 2004) and fisheries (Candy 2004).

In many applications in which physical quantities are measured, the occurrence of continuous data with exact zeros is common. In such circumstances, Tweedie distribution models time and again are useful distributions and would be more frequently used in practice if high quality numerical computations were readily available. Dunn and Smyth (2005) also give two detailed data analyses, one for $p > 2$ and another for$1 < p < 2$, indicative of the practicality of the distributions and methodology on real data. Other than the well-known distributions with $p = 0,1,2,3$, none of the Tweedie distribution models have cumulative or probability density functions with obvious logical forms. This makes it difficult to use these distributions in statistical modeling. Actually, it inhibits their use with likelihood based estimation, testing or analytical procedures. Conversely, Tweedie distribution models have simple, analytic moment generating functions. The purpose of this paper is to provide fast, accurate computation of the Chain ladder using Tweedie densities.

CHAPTER 3

3.1: Methodology

3.1.1: Exponential dispersion models

Exponential dispersion families are sets of probability distributions. Many widely used densities can be parameterized as an exponential dispersion family. Exponential dispersion families are also the basis for Generalized Linear Models. This chapter presents a general construction of the Tweedie family of distribution based on the generalization of the exponential dispersion family. For this purpose the order of ideas in which our construction is made is based on Jorgensen (1986), but the names given to the different exponential dispersion families are taken from Jorgensen (1997).

3.1.2: One parameter exponential family

A random variable (Discrete or Continuous) has a distribution which is from the linear exponential family if its probability function may be parameterized in terms of a parameter $\theta$and expressed as

$$f(y;\theta) = \frac{p(y)e^{-\theta y}}{q(\theta)} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (3.1)$$

The function $p(y)$ depends only on $y\ not\ on\ \theta$, and the function $q(\theta)$ is a normalizing constant. Also, the support of the random variable must not depend on θ. The parameter $\theta$is called the natural parameter of the distribution.

Examples:

3.1.2.1: Exponential Distribution

The probability density function (pdf) is:

$$f(y,\beta) = \frac{1}{\beta} e^{-\frac{y}{\beta}}$$

If we let $\theta = \frac{1}{\beta}$, then the probability density function is

$$f(y,\theta) = \frac{1e^{-\theta y}}{\theta^{-1}}$$

This is of the form of equation (3.1) with $p(y) = 1\ and\ q(\theta) = \frac{1}{\theta}$

3.1.2.2: Poisson distribution

The probability distribution function is

$$f(y,\lambda) = \frac{\lambda^y e^{-\lambda}}{y!} = \frac{\left(\frac{1}{y!}\right) e^{(-(-\ln\lambda)y)}}{e^{\lambda}}$$

If we let $\theta = -\ln\lambda$, then the probability function is

$$f(y,\lambda) = \frac{\left(\frac{1}{y!}\right) e^{(-\theta y)}}{e^{e^{-\theta}}}$$

Which is of the form of equation (3.1) with $p(y) = \frac{1}{y!}$ and $q(\theta) = e^{e^{-\theta}}$

3.1.2.3: Normal Distribution with mean $\mu$and known variance v

The probability density function is

$$f(y;\mu,v) = (2\pi v)^{-\frac{1}{2}} \exp\left[\frac{1}{2v}(y-\mu)^2\right]$$

$$= (2\pi v)^{-\frac{1}{2}} \exp\left[-\frac{y^2}{2v} + \frac{\mu}{v}y + \frac{\mu^2}{2v}\right]$$

$$= \frac{(2\pi v)^{-\frac{1}{2}} \exp\left(-\frac{y^2}{2v}\right) \exp\left(\frac{\mu}{v}y\right)}{exp\left(\frac{\mu^2}{2v}\right)}$$

If we let $\theta = -\frac{\mu}{v}$, the probability density function is

$$f(y;\theta,v) = \frac{(2\pi v)^{-\frac{1}{2}} \exp\left(-\frac{y^2}{2v}\right) \exp(\theta y)}{exp\left(\frac{\theta^2 v}{2}\right)}$$

This is of the form of equation (3.1)

with $p(y) = (2\pi v)^{-\frac{1}{2}} \exp\left(-\frac{y^2}{2v}\right)$ and $q(\theta) = exp\left(\frac{\theta^2 v}{2}\right)$

Mean and Variance

We now find the mean and variance of the distribution defined by equation (3.1). First, note that

$$\ln f(y;\theta) = \ln p(y) - \theta y - \ln q(\theta)$$

Differentiate with respect to θ to obtain

$$\frac{\partial}{\partial\theta} f(y,\theta) = \left[-y - \frac{q'(\theta)}{q(\theta)}\right] f(y;\theta) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots .3.1.1$$

Integrate (or sum) over the range of y (known not to depend on θ) to obtain

$$\int \frac{\partial}{\partial\theta} f(y,\theta)dy = -\int y\, f(y,\theta)dy - \frac{q'(\theta)}{q(\theta)} \int f(y;\theta)dy$$

On the left hand side, interchange the order of differentiation and integration(or summation) to obtain

$$\frac{\partial}{\partial\theta}\left[\int f(y,\theta)dy\right] = -\int yf(y,\theta)dy - \frac{q'(\theta)}{q(\theta)} \int f(y;\theta)dy$$

We know that

$$\int f(y;\theta) = 1$$

And

$$\int yf(y;\theta) = E(y)$$

Thus

$$\frac{\partial}{\partial\theta}(1) = -E(y) - \frac{q'(\theta)}{q(\theta)}$$

In other words, the mean is

$$E(y) = \mu(\theta) = -\frac{q'(\theta)}{q(\theta)} = -\frac{d}{d\theta}.\ln q(\theta)$$

To obtain the variance, (3.1.1) may first be rewritten as

$$\frac{\partial}{\partial\theta} f(y;\theta) = -[y - \mu(\theta)]f(y;\theta)$$

Differentiating again with respect to θ to obtain

$$\frac{\partial^2}{\partial\theta^2} f(y;\theta) = \mu'(\theta)f(y;\theta) - [y - \mu(\theta)]\frac{\partial}{\partial\theta} f(y;\theta)$$

$$= \mu'(\theta)f(y;\theta) + [y - \mu(\theta)]^2 f(y;\theta)$$

Again, integrate over the range of x to obtain

$$\int \frac{\partial^2}{\partial\theta^2} f(y;\theta) = \mu'(\theta) \int f(y;\theta) + \int [y - \mu(\theta)] f(y;\theta)$$

In other words

$$\int [y - \mu(\theta)]^2 f(y;\theta) dy = -\mu'(\theta) + \frac{\partial^2}{\partial\theta^2} f(y;\theta) dy$$

Because $\mu(\theta)$ is the mean, the LHS is the variance (by definition) then because the 2nd term on the RHS is zero, we obtain

$$Var(y) = V(\theta) = -\mu'(\theta) = \frac{d^2}{d\theta^2} \ln q(\theta)$$

3.1.3: Two parameter exponential family

A distribution for a random variable y belongs to an exponential family if its density has the following form

$$f(y;\theta.\phi) = exp\left[\frac{y\theta - b(\theta)}{a(\phi)} + c(y,\phi)\right] \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3.2$$

Where $\theta \Rightarrow Natural\ Parameter,$

$$\phi \Rightarrow Scale\ parameter\ or\ Dispersion\ parameter$$

$$\theta\ is\ a\ function\ of\ \mu = E(y)\ only$$

Examples

3.1.3.1: Normal Distribution

$$f(y;\theta,\phi) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(y-\mu)^2}{2\sigma^2}\right]$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(y^2 - 2y\mu + \mu^2)}{2\sigma^2}\right]$$

$$= exp\left[-\frac{1}{2}\log 2\pi\sigma^2\right] exp\left[-\frac{y^2}{2\sigma^2} + \frac{y\mu}{\sigma^2} - \frac{\mu^2}{2\sigma^2}\right]$$

$$= exp\left[\frac{y\mu - \frac{\mu^2}{2}}{\sigma^2} - \frac{1}{2}\left(\frac{y^2}{\sigma^2} + \log 2\pi\sigma^2\right)\right]$$

This is of the form of (3.2). Where $\theta = \mu$

$$\phi = \sigma^2$$

$$a(\phi) = \sigma$$

$$b(\phi) = \frac{\mu^2}{2}$$

$$c(y, \phi) = -\frac{1}{2}\left(\frac{y^2}{\sigma^2} + \log 2\pi\sigma^2\right)$$

For member of an exponential family, we want to be able to find formulae for the mean and variance of the distribution from the general parameters. Consider the log-likelihood function,

$$l(y; \theta, \phi) = \log\big(f(y; \theta, \phi)\big)$$

$$\log\big(f(y; \theta, \phi)\big) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)$$

$$\frac{\partial \log f}{\partial \theta} = \frac{y - b'(\theta)}{a(\phi)}$$

$$E\left[\frac{\partial \log f}{\partial \theta}\right] = E\left[\frac{y - b'(\theta)}{a(\phi)}\right] = \frac{E(y) - b'(\theta)}{a(\phi)}$$

But

$$E\left[\frac{\partial \log f}{\partial \theta}\right] = \int f \frac{\partial \log f}{\partial \theta} dy$$

$$= \int f . \frac{1}{f}\frac{\partial f}{\partial \theta} dy$$

$$= \int \frac{\partial f}{\partial \theta} dy$$

$$= \frac{\partial}{\partial \theta}\int f dy$$

Since $\int f dy$ is a probability density function, its equal to 1

$$\frac{\partial}{\partial \theta}(1)$$

$$= 0$$

$$E\left[\frac{\partial \log f}{\partial \theta}\right] = \frac{E(y) - b'(\theta)}{a(\phi)} = 0$$

Therefore

$$E(y) = b'(\theta)$$

Computation of the variance

$$Var(y) = \frac{\partial^2 \log f}{\partial \theta^2} = -\frac{b''(\theta)}{a(\theta)}$$

Let

$$Q = E\left(\frac{\partial^2 l}{\partial \theta^2}\right) + E\left[\left(\frac{\partial l}{\partial \theta}\right)^2\right]$$

$$= E\left[\frac{\partial^2 l}{\partial \theta^2} + \left(\frac{\partial l}{\partial \theta}\right)^2\right]$$

$$\int f\left[\frac{\partial^2}{\partial \theta^2}\log f + \left(\frac{\partial}{\partial \theta}\log f\right)^2\right] dy$$

Using the function of a function rule,

$$Q = \int f\left[\frac{\partial^2}{\partial \theta^2}\log f + \left(\frac{\partial}{\partial \theta}\log f\right)^2\right] dy$$

$$= \int f.\frac{1}{f}\frac{\partial^2 f}{\partial \theta^2} dy$$

$$= \frac{\partial^2}{\partial \theta^2}\int f dy$$

$$= \frac{\partial^2}{\partial \theta^2}(1)$$

$$= 0$$

$$E\left[\frac{\partial^2 \log f}{\partial \theta^2}\right] + E\left[\left(\frac{\partial \log f}{\partial \theta}\right)^2\right] = 0$$

$$\frac{b''(\theta)}{a(\phi)} + E\left[\left(\frac{y - b'(\theta)}{a(\phi)}\right)^2\right] = 0$$

$$-\frac{b''(\theta)}{a(\phi)} + \frac{E\left(y - b'(\theta)\right)^2}{\left(a(\phi)\right)^2} = 0$$

$$-b''(\theta)a(\phi) + E\left(y - b'(\theta)\right)^2 = 0$$

$$E\left(y - b'(\theta)\right)^2 = b''(\theta)a(\phi)$$

$$Var(y) = a(\phi)b''(\theta)$$

Where $a(\phi) \Rightarrow Dispersion\ Parameter$

$$b''(\theta) \Rightarrow Variance\ function$$

3.1.3.2: Normal Distribution

Consider a normal distribution; we can derive the mean and variance.

$$b(\theta) = \frac{\mu^2}{2}$$

So

$$E(y) = b'(\theta) = \frac{2\mu}{2} = \mu$$

And

$$a(\phi) = \phi$$

So

$$Var(y) = a(\phi)b''(\theta)$$
$$= \phi(1) = \sigma^2$$

3.1.3.3: Poisson distribution

$$f(y;\theta,\phi) = \frac{\mu^y e^{-\mu}}{y!}$$

$$\exp[y\log\mu - \mu - \log y!]$$

This is in the form of (3.2), with: $\theta = \log\mu$

$$\phi = 1$$
$$a(\phi) = 1$$
$$b(\theta) = e^{\theta}$$
$$c(y,\theta) = -\log y!$$

Thus Natural parameter for the Poisson distribution is $\log\mu$

$$E(y) = b(\theta) = e^{\theta} = e^{\log\mu} = \mu$$
$$Var(y) = a(\phi)b''(\theta) = (1)e^{\theta} = (1)e^{\log\mu} = \mu$$

The variance function ($V(\mu) = b''(\theta)$) tells us that the variance function is proportional to the mean.

3.1.3.4: Binomial distribution

We have to first divide the binomial random variable by n.

Suppose $z \sim binomial\ (n, \eta)$. Let $y = \frac{z}{n}$ so that $z = ny$. The distribution of z is

$$f(z; \theta, \phi) = \binom{n}{z} \mu^z (1-\mu)^{n-z}$$

and by substitution for z, the distribution of y is

$$f(y; \theta, \phi) = \binom{n}{ny} \mu^{ny} (1-\mu)^{n-ny}$$

$$= \exp\left[ny \log \mu + (n - ny)\log(1-\mu) + \log\binom{n}{ny}\right]$$

$$= \exp\left[n[y \log \mu + (1-y)\log(1-\mu)] + \log\binom{n}{ny}\right]$$

$$= \exp\left[n\left\{y \log\left(\frac{\mu}{1-\mu}\right) + \log(1-\mu)\right\} + \log\binom{n}{ny}\right]$$

This is in the form of (3.2) with $\theta = \log\left(\frac{\mu}{1-\mu}\right)\left[Inverse;\ \mu = \frac{e^\theta}{1+e^\theta}\right]$

$$\theta = n$$

$$a(\phi) = \frac{1}{\phi}$$

$$b(\theta) = \log(1-\mu) = \log(1+e^\theta)$$

$$c(y, \phi) = \log\binom{n}{ny}$$

If $\theta = \log\frac{\mu}{1-\mu}$, then $e^\theta = \frac{\mu}{1-\mu}$.

This translates to $e^\theta - \mu e^\theta = \mu$

Hence $\mu = \frac{e^\theta}{1+e^\theta}$

$$b(\theta) = \log(1-\mu)$$

$$b(\theta) = \log\left(1 - \frac{e^\theta}{1+e^\theta}\right)$$

$$b(\theta) = \log\left(\frac{1}{1+e^\theta}\right) = -\log\left(1+e^\theta\right)$$

$$b(\theta) = \log\left(1+e^\theta\right)$$

$$b'(\theta) = \frac{d}{d\theta}\log\left(1+e^\theta\right)$$

$$= \frac{e^\theta}{1+e^\theta}$$

3.1.3.5: Gamma Distribution

The best way to consider the Gamma distribution is to change the parameters from α and λ to $\alpha$and $\mu = \frac{\alpha}{\lambda}$. Recall that: θ must always be expressed as a function of μ.

$$f(y, \theta, \phi) = \frac{\lambda^{\alpha} y^{\alpha-1} e^{-\lambda y}}{\sqrt{\alpha}}$$

$$= \frac{\alpha^{\alpha} y^{\alpha-1} e^{-y\frac{\alpha}{\mu}}}{\mu^{\alpha}\sqrt{\alpha}}$$

$$= \exp\left[\left(-\frac{y}{\mu} - \log\mu\right)\alpha + (\alpha - 1)\log y + \alpha\log\alpha - \log\sqrt{\alpha}\right]$$

This is in the form of equation (3.2).

Where $\theta = -\frac{1}{\mu}$

$$\phi = \alpha$$

$$a(\theta) = \frac{1}{\alpha}$$

$$b(\theta) = -\log(-\theta)$$

$$c(y, \phi) = (\phi - 1)\log y + \phi\log\phi - \log\sqrt{\phi}$$

The exponential distribution (y) is just a Gamma$(1, \lambda)$ distribution, so our results are consistent with those of

$$f(y) = \lambda\, e^{-\lambda y} = e^{\log\lambda - \lambda y}$$

Since$E(y) = \frac{1}{\lambda}$, this is in the appropriate form with

$$\theta = -\lambda = -\frac{1}{\mu}$$

$$b(\theta) = -\log(-\theta)$$

$$\phi = 1$$

$$a(\phi) = \phi$$

Thus, the natural parameter for the gamma distribution is$\frac{1}{\mu}$, ignoring the minus sign. The mean is$E(y) = b'(\theta) = -\frac{1}{\theta} = \mu$. The variance function is $V(\mu) = b''(\theta) = \frac{1}{\theta^2} = \mu^2$ and so the variance is$\frac{\mu^2}{\alpha}$.

3.1.3.6: Inverse Gaussian distribution

Consider the probability distribution function of inverse Gaussian which is given as

$$f(y;\theta,\phi)=\left(\frac{\lambda}{2\pi y^3}\right)^{\frac{1}{2}}\exp\left(\frac{-\lambda(y-\mu)^2}{2\mu^2 y}\right);y>0,\mu>0,\lambda>0$$

$$=\exp\left\{-\lambda\left(\frac{y^2-2y\mu+\mu^2}{2\mu^2 y}\right)+\frac{1}{2}(\ln\lambda-\ln 2\pi y^3)\right\}$$

$$=\exp\left\{\left(-\frac{y}{2\mu^2}+\frac{1}{\mu}\right)\lambda-\left(\frac{\lambda}{2y}+\frac{1}{2}\ln\frac{\lambda}{2\pi y^3}\right)\right\}$$

Letting $\phi=\frac{1}{\lambda}$, we have

$$=\exp\left\{\left(-\frac{y}{2\mu^2}+\frac{1}{\mu}\right)\frac{1}{\phi}-\left(\frac{1}{2\phi y}+\frac{1}{2}\ln\sqrt{\frac{1}{2\pi\phi y^3}}\right)\right\}$$

$$=\exp\left\{\left[\left(-\frac{y}{2\mu^2}\right)+\left(\frac{1}{\mu}\right)\right]\frac{1}{\phi}-\left(\frac{1}{2\phi y}+\ln\sqrt{2\pi\phi y^3}\right)\right\}$$

This is in the form of equation (3.2)

Where $\theta=\frac{1}{2\mu^2}$

$$\mu=(-2\theta)^{-\frac{1}{2}}$$

$$\phi=\sigma^2$$

$$a(\phi)=\phi$$

$$b(\theta)=-\frac{1}{\mu}=-(-2\theta)^{\frac{1}{2}}$$

$$c(y,\phi)=-\frac{1}{2}\left\{\ln 2\pi\phi y^3+\frac{1}{\phi y}\right\}$$

The mean is

$$E(y)=b'(\theta)=\frac{d}{d\theta}\left(-(-2\theta)^{\frac{1}{2}}\right)$$

$$=\frac{d}{d\theta}(-2\theta)^{\frac{1}{2}}$$

$$=(-2\theta)^{\frac{1}{2}}$$

$$=\mu$$

$$Var(\mu)=b''(\theta)=2^{-\frac{1}{2}}\frac{d}{d\theta}\left(\theta^{-\frac{1}{2}}\right)$$

$$=2^{-\frac{1}{2}}.-\frac{1}{2}\left(\theta^{-\frac{3}{2}}\right)$$

$$= -2^{-\frac{3}{2}}.\theta^{-\frac{3}{2}}$$

$$= -(2\theta)^{-\frac{3}{2}}$$

$$= \left[(-2\theta)^{-\frac{1}{2}}\right]^3$$

$$= \mu^3$$

Table 3.1: Summary of the exponential family of distribution with logical form

Distribution	$\boldsymbol{\theta}$	$\boldsymbol{\phi}$	$\boldsymbol{a(\phi)}$	$\boldsymbol{b(\theta)}$	$\boldsymbol{E(y) = b'(\theta)}$	$\boldsymbol{V(\mu) = b''(\theta)}$
Normal	μ	σ^2	ϕ	$\frac{\mu^2}{2}$	μ	$\mu^0 = 1$
Poisson	$\log \mu$	1	θ	e^θ	$\mu = e^\theta$	$e^\theta = \mu$
Binomial	$\log\left(\frac{\mu}{1-\mu}\right)$	η	$\frac{1}{\mu}$	$\log(1+e^\theta)$	$\mu = \frac{ne^\theta}{(1-e^\theta)}$	$\frac{ne^\theta}{(1-e^\theta)}$
Gamma	$-\frac{1}{\mu}$	α	$\frac{1}{\phi}$	$-\log(-\theta)$	$\mu = \frac{1}{\theta}$	$\frac{1}{\theta^2} = \mu^2$
Inverse Gaussian	$-\frac{1}{2\mu^2}$	σ^2	ϕ	$= -(2\theta)^{\frac{1}{2}}$	$\mu = -(2\theta)^{-\frac{1}{2}}$	$\left[(-2\theta)^{-\frac{1}{2}}\right]^3 = \mu^3$

From the above table, we can clearly see that the variance function is the power of the mean. $V(\mu) = \mu^P$

3.2: Exponential Dispersion models with Power Variance Functions

Distribution	$\boldsymbol{p}$	$\boldsymbol{\alpha}$
Generated by Extreme Stable Distributions	$p < 0$	$1 < \alpha < 2$
Normal Distribution	$p = 0$	$\alpha = 2$
Not exponential dispersion models	$0 < p < 1$	$\alpha > 2$
Poisson Distribution	$p = 1$	$\alpha = -\infty$
Tweedie Compound Poisson Distribution	$1 < p < 2$	$\alpha < 0$
Gamma Distribution	$p = 2$	$\alpha = 0$
Generated by Positive Stable Distribution	$2 < p < 3$	$0 < \alpha < \frac{1}{2}$
Inverse Gaussian Distribution	$p = 3$	$\alpha = \frac{1}{2}$
Generated by Positive Stable Distributions	$p > 3$	$\frac{1}{2} < \alpha < 1$

3.2: Tweedie Compound Poisson distribution, $1 < p < 2$

Tweedie distribution generalized linear model is the only model that has successfully been developed to require only one distribution. The Tweedie family of distributions isa class of distribution that is able to model both discrete and continuous probabilities together as one model. Tweedie distributions are based upon generalized linear models and are classified by their variances which takes the form

$$Var(y) = \phi\mu^p$$

Tweedie models exist for all values of p outside the interval (0,1), however only the 4 distributions (Normal, Poisson, Gamma and Inverse Gaussian) have density functions which have explicit analytic forms.

Tweedie distributions with $p > 1$ have strictly positive means, with $p > 2$ being continuous for positive Y and a shape similar to the gamma, but more right skewed. Distributions with $p < 0$ are continuous on the entire real axis. Finally, for $1 < p < 2$ the distributions are supported on non-negative real numbers, and the distributions are

mixtures of the Poisson and Gamma distribution, with a mass at zero. These distributions have been called 'Compound Poisson' distributions.

There are three main properties that make the Tweedie distributions exceptional for modeling claims:

i. The Tweedie distributions belong to the exponential family of distributions, and form part of a larger group of models called the generalized linear models; This is advantageous because fitting techniques and diagnostics are readily available for generalized linear model and thus Tweedie distribution models.
ii. The motivation behind the setup of these models is simple and logical: Total Claims is considered the sum of claims on some smaller time scale.
iii. The Tweedie distributions provide a mechanism in which finer-scale structures can be understood through courser-scale data.

Tweedie distributions models have a trait of being able to model both discrete and continuous combinations simultaneously. The mean, μ, and canonical parameter, θ can be found for a tweedie distribution by noting that

$$b''(\theta) = \frac{d\mu}{d\theta} = \mu^p$$

and the mean is given by $\mu = b'(\theta)$.

This allows the density function for a tweedie distribution to be specified. Hence,

$$\mu^p = \frac{\partial^2 b}{\partial \theta^2}$$

$$= \frac{\partial}{\partial \theta}\left(\frac{\partial b}{\partial \theta}\right)$$

$$= \frac{\partial \mu}{\partial \theta}$$

Taking reciprocals of both sides and integrating with respect to μ gives

$$\frac{\partial \theta}{\partial \mu} = \mu^p$$

$$\int \partial\theta = \int \mu^{-p}\, \partial\mu$$

$$\theta = \int \mu^{-p}\, d\mu$$

When $p = 1$, we have

$$\theta = \int \mu^{-1} d\mu$$

$$\theta = \int \frac{1}{\mu} d\mu = \ln \mu$$

When $p \neq 1$, we have

$$\theta = \int \mu^{-p}\, d\mu$$

$$\theta = \frac{\mu^{1-p}}{1-p}$$

Hence,

$$\theta = \begin{cases} \dfrac{\mu^{1-p}}{1-p} & p \neq 1 \\ \ln \mu & p = 1 \end{cases}$$

By Setting the arbitrary constant of integration to Zero and noting that $\mu = b'(\theta)$ gives,

$$\mu = b'(\theta)$$

$$\mu = \frac{d}{d\theta} b(\theta)$$

$$\mu\, d\theta = d\ b(\theta)$$

$$b(\theta) = \int \mu\ d\theta$$

But

$$d(\theta) = \mu^{-p}\ d\mu$$

$$= \int \mu . \mu^{-p}\ d\mu$$

$$= \int \mu^{1-p} \; d\mu$$

When $p = 2$

$$b(\theta) = \int \mu^{-1} \; d\mu$$

$$= \ln \mu$$

When $p \neq 2$

$$b(\theta) = \int \mu^{1-p} \; d\mu$$

$$= \frac{\mu^{2-p}}{2-p}$$

The Tweedie densities can thus be written as,

$$f(y;\mu,\phi) = a(y,\phi) \exp\left\{\frac{1}{a(\phi)}\left[y\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\right]\right\} \qquad for\ p \neq (1,2) \quad \ldots \quad (3.4)$$

Where:

$$a(y,\phi) = \begin{cases} \frac{1}{y}\sum_{k=1}^{\infty} \frac{\phi^k k_\alpha^k (-(\phi y)^{-1})}{\Gamma(-k\alpha)k!} & for\ y > 0 \\ 1 & for\ y = 0 \end{cases}$$

3.2.1: Quasi-likelihood Methods

In many situations, some details of the distribution governing the data is known, however the distribution may not be able to be specified exactly. In addition, there are some cases for which the distribution is known, however it's difficult to evaluate, such as the Tweedie distribution in order to construct the likelihood. The idea of quasi-likelihood addresses this concern. Quasi-likelihood methods are methodology for regression that requires few assumptions about the distribution of dependent variable. In likelihood analysis, the actual form of the distribution must be specified. However, in quasi-likelihood, only the relationship between the outcome mean and covariates, and the mean and variance, needs to be specified. The focus of quasi-

likelihood is on methods for inference about β, and hence ϕ can be treated as a nuisance parameter.

A quasi-likelihood can be used if the researcher does not know the density function of the distribution, but knows its mean and variance. It is defined for one observation, Q as

$$Q(y;\mu) = \int \frac{y-\mu}{v(\mu)} \quad d\mu \qquad \ldots \qquad 3.5$$

To define a quasi-likelihood function, only the relationship between the mean and variance needs to be specified through the variance function.

3.2.2: Tweedie Distribution and the Quasi-likelihood

Following equation (3.5) the Tweedie distribution has the following quasi-likelihood distribution (When setting the arbitrary constant of integration to Zero)

$$Q(\mu, y) = \int \frac{y-\mu}{v(\mu)} \quad d\mu$$

$$= \int \frac{y-\mu}{\mu^p} \quad d\mu$$

$$= \int \left(\frac{y}{\mu^p} - \mu^{1-p}\right) d\mu$$

$$= \int (y\mu^{-p} - \mu^{1-p}) d\mu$$

$$= \frac{y\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}$$

This equation has the same likelihood function as equation (3.4), expect now there is no need to estimate $a(y,\phi)$. This is extremely helpful as often the $a(y,\phi)$ term cannot be written in closed form, or is of a form which is extremely difficult to calculate.

CHAPTER 4

4.1: Methodology

4.1.1: Chain Ladder

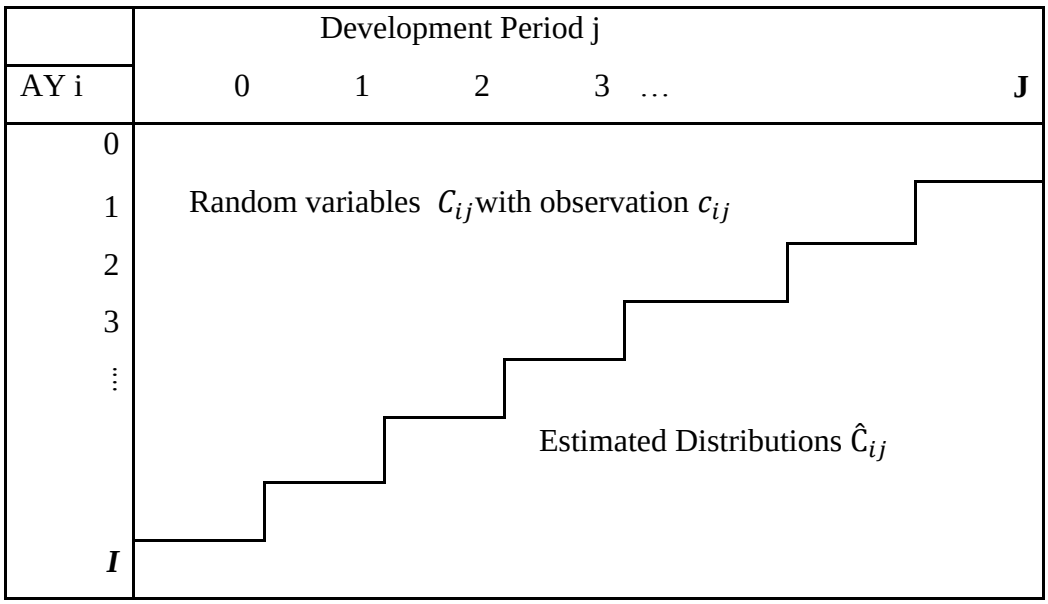
The typical Chain Ladder is a deterministic process to forecast claims based on historical data. It assumes that the proportional developments of claims from one development period to the next are the same for all origin years.

4.1.2: Definition of the Model

We use the following structure for the Chain Ladder run-off patterns: the origin years are denoted by $i \leq I$ and the development periods are denoted by$j \leq J$. We are interested in the random variables C_{ij}.C_{ij}denotesthe incremental payments for claims with origin in origin year i during development period j. Usually one has observation c_{ij} of C_{ij} for $i + j \leq I$ and one tries to complete(estimate) the triangle for $i + j \leq I$. The following illustration may be helpful

a) The number of paymentsC_{ij} is Poisson distributed with parameter$\lambda_{ij} w_i > 0$ and independent. The weight $w_i > 0$ is a suitable measure for the volume.

Table 4.1: Chain Ladder run-off triangle structure



b) The individual payments $X^{(k)}{}_{ij}$ are independent and gamma distributed with mean $\tau_{ij} > 0$ and shape parameter $\tau > 0$.

c) $C_{ij}$and$X^{(k)}{}_{ij}$ are independent for all indices. We define the incremental payments paid in cell(i,j) as follows

$C_{ij} = 1_{\{R_{ij}>0\}} \cdot \sum_{k=1}^{R_{ij}} X_{ij}^{(k)}andY_{ij} = {}^{C_{ij}}/_{w_i}$

Remarks

- There are several different possibilities to choose appropriate weights w_ie.g the number of policies or the total number of claims incurred. If one chooses the total number of claims incurred one needs first to estimate the number of Incurred But Not yet Reported Cases.
- Sometimes it is also convenient to define R_{ij} as the number of claims with origin in i which have at least one payment in period j
- Y_{ij}denotes the normalized incremental payments in cell (i, j).
- One easily sees that conditionally, given R_{ij}, the random variable C_{ij} is gamma distributed with mean $C_{ij}\tau_{ij}$ and shape parameter $C_{ij}\gamma$.

4.2.3: Claims Data Example

Consider the run off triangle below.

Table 4.2: Example of a Run-off triangle

	Development Period									
Origin Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2004	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
2005	106	4285	5396	10666	13782	15599	15702	16169	16704	
2006	3410	8992	13873	16141	18735	22214	22863	23466		
2007	5655	11555	15766	21266	23425	26083	27067			
2008	1092	9565	15836	22169	25955	26180				
2009	1513	6445	11702	12935	15852					
2010	557	4020	10946	12314						
2011	1351	6947	13112							
2012	3133	5395								
2013	2063									

This run off triangle shows the known values of loss from each origin year as of the end of the origin year and annual evaluations thereafter. For example, the known values of loss originating from the 2011 exposure period are 1,351, 6,947, and 13,112 as of year ends 2011, 2012, and 2013, respectively. The most recent diagonal, i.e., the vector 18,834, 16,704 . . . 2,063 from the upper right to the lower left, shows the most current evaluation available. The column headings show the years of the observations in the column relative to the beginning of the exposure period. For example, for the 2011 origin year, the age of the 1,351 value, evaluated as of December, 31st, 2011, is three years. The objective of a reserving exercise is to forecast the future claims development in the bottom right corner of the triangle and potential further developments beyond development year 10. Eventually all claims for a given origin period will be settled, but it is not always obvious to judge how many years or even decades it will take. We speak of long and short tail business contingent to the time it takes to pay all claims. We plot the data above to get an overview.

Figure 4.1: Claims development chart of the Claims Data triangle, with one line per origin period.

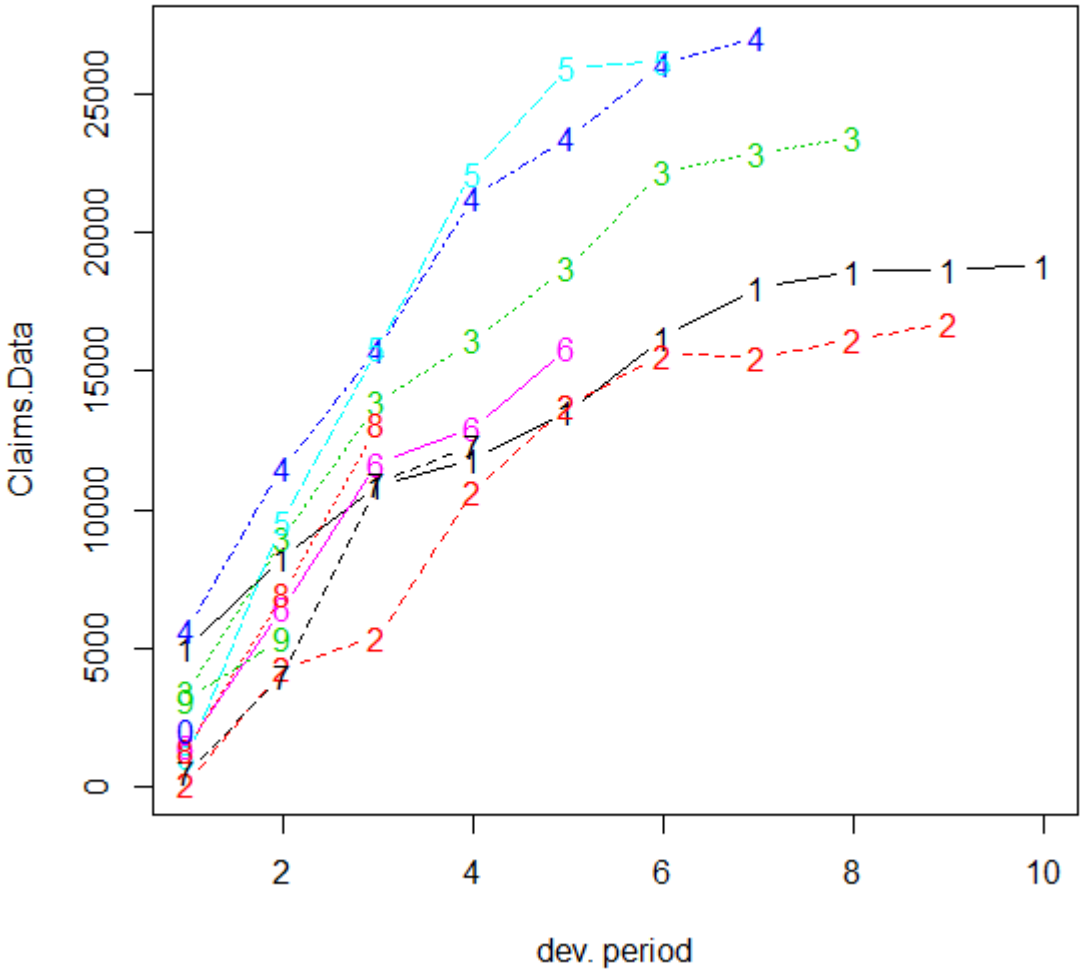
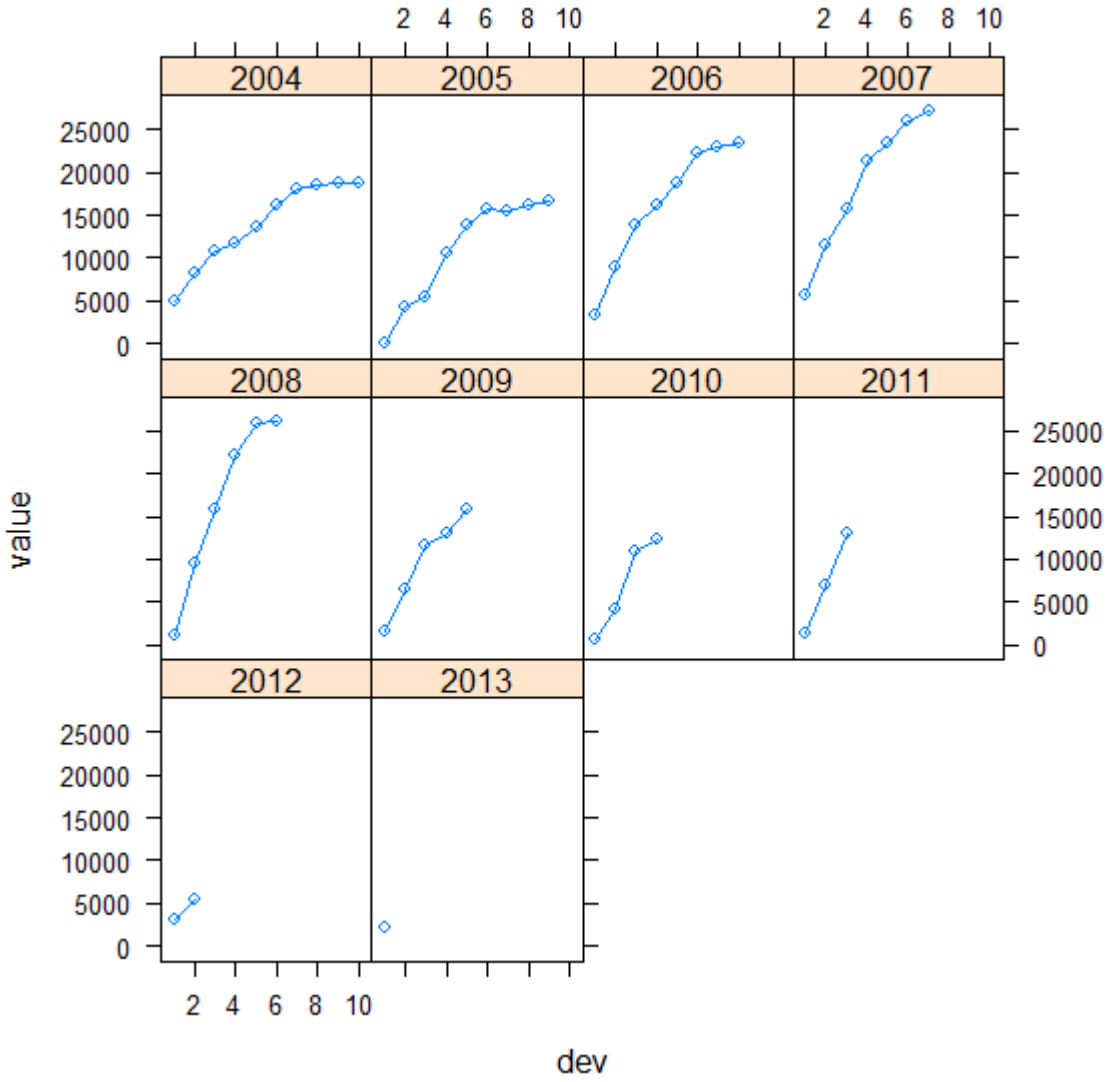


Figure 4.2: Claims development chart of the Claims Data triangle, with individual panels for each origin period.



4.2.4: Basic Chain Ladder

The most commonly used method as a first step is calculation of the age-to-age relation ratios as the volume weighted average development ratios of a cumulative loss development triangle from one development period to the next $C_{ik}, i, k = 1, \dots, n$.

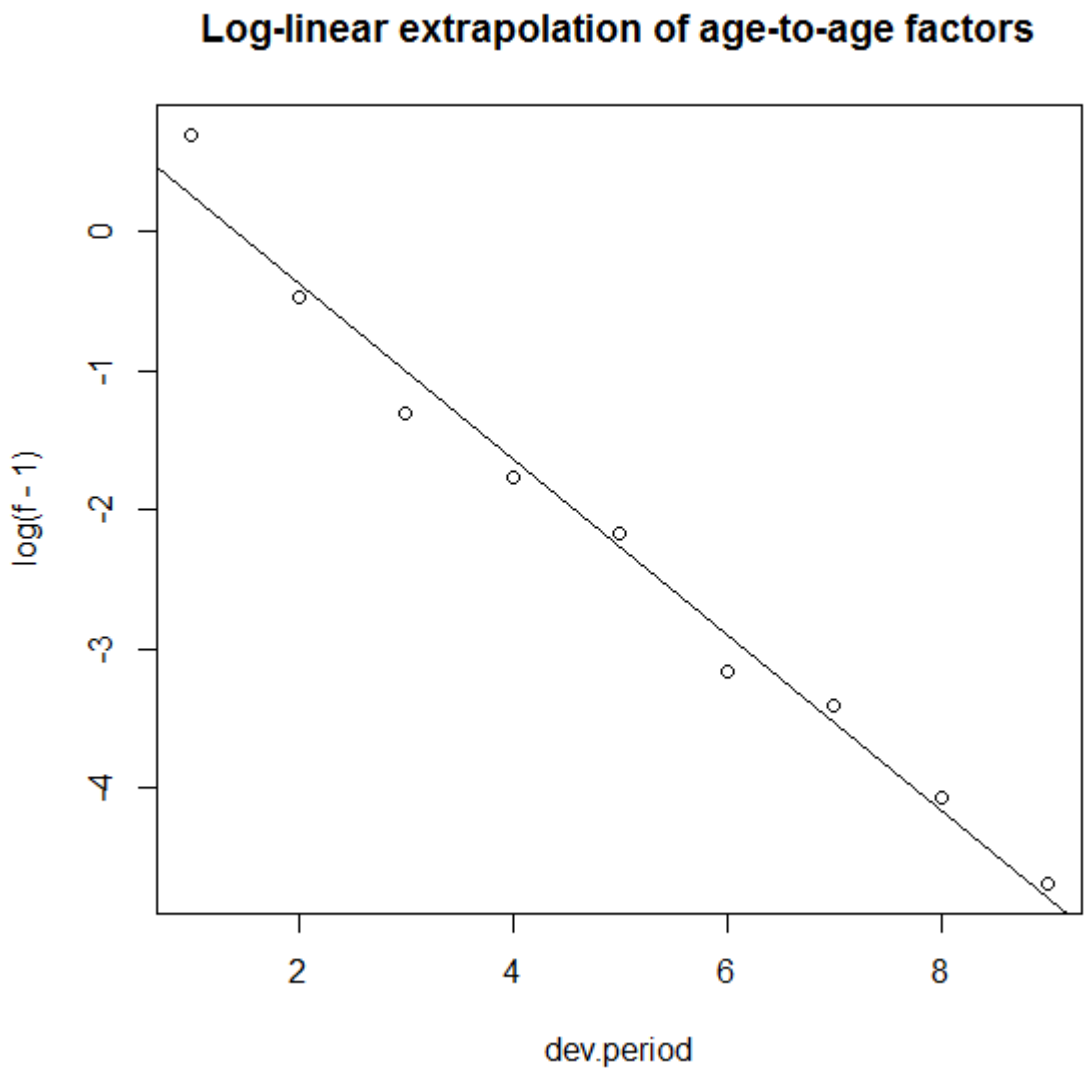
$$f_k = \frac{\sum_{i=1}^{n-k} C_{i,k+1}}{\sum_{i=1}^{n-k} C_{i,k}}$$

Table 4.3: Basic Chain Ladder Development factors

Development Period										
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Factors	2.999	1.624	1.271	1.172	1.113	1.042	1.033	1.017	1.009	1

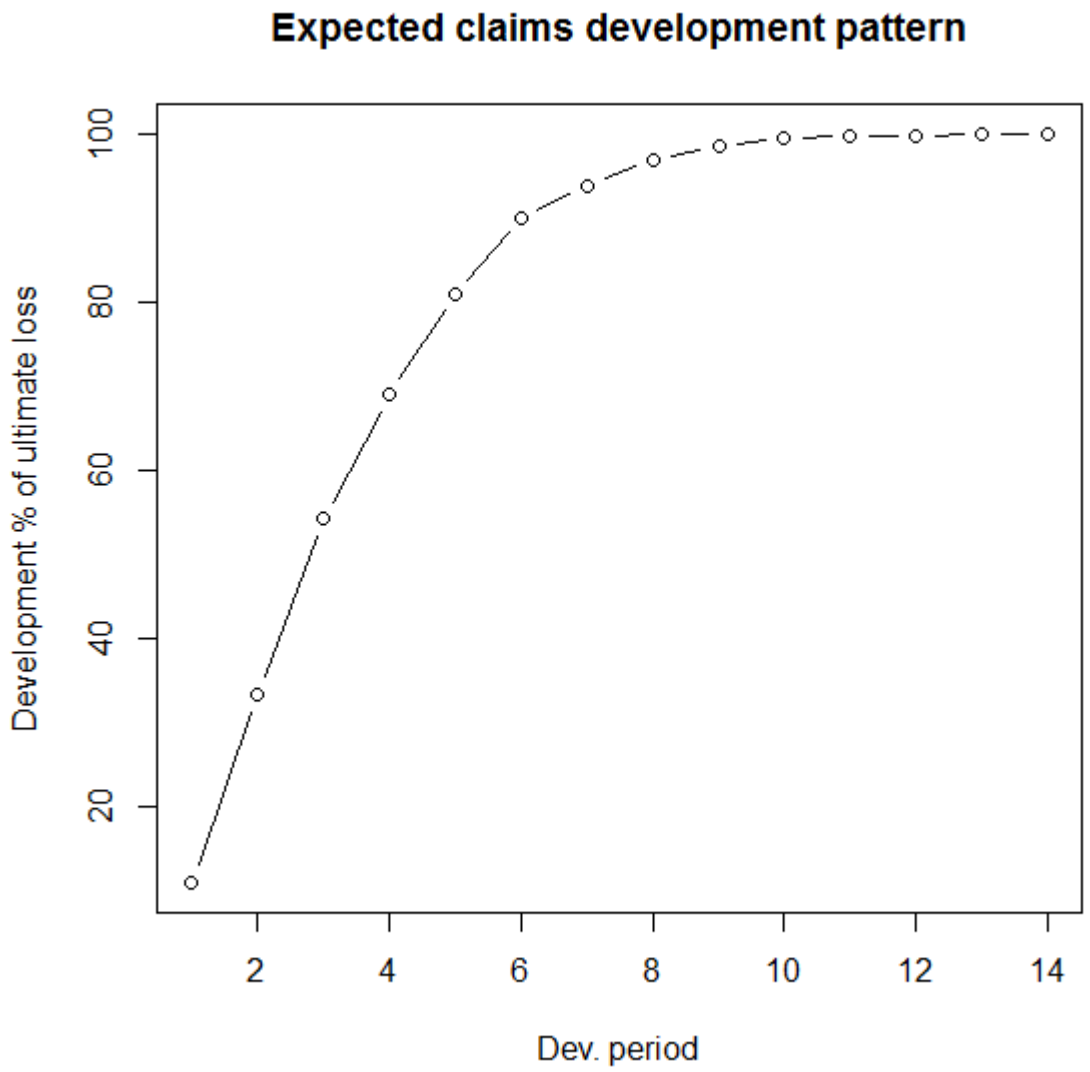
Time and again it is not suitable to assume that the oldest origin year is fully developed. A classical tactic is to extrapolate the development ratios, e.g. assuming a log-linear model.

Figure 4.3: Log-linear extrapolation of age-to-age factors



The age-to-age factors allow us to plot the expected claims development patterns.

Figure 4.4: Expected Claims development pattern



The link ratios are then applied to the latest known cumulative claims amount to forecast the next development period. The squaring of the Claims Data triangle is calculated below, where an ultimate column is appended to the right to accommodate the expected development beyond the oldest age (10) of the triangle due to the tail factor (1.005) being greater than unity.

Table 4.4: Basic Chain Ladder run-off triangle with forecasted outstanding claims

	Development Period										
Origin Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**Ultimate**
2004	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834	18928
2005	106	4285	5396	10666	13782	15599	15707	16169	16704	16858	16942
2006	3410	8992	13873	16141	18735	22214	22863	23466	23863	24083	24204
2007	5655	11555	15766	21266	23425	26083	27067	27967	28441	28703	28847
2008	1092	9565	15836	22169	25955	26180	27278	28185	28663	28927	29072
2009	1513	6445	11702	12935	15852	17649	18389	19001	19323	19501	19599
2010	557	4020	10946	12314	14428	16064	16738	17294	17587	17749	17838
2011	1351	6947	13112	16664	19525	21738	22650	23403	23800	24019	24139
2012	3133	5395	8759	11132	13043	14521	15130	15634	15898	16045	16125
2013	2063	6188	10046	12767	14959	16655	17353	17931	18234	18402	18495

The total estimated outstanding loss under this method is about 53,200. This approach is also called Loss Development Factor (LDF) method. More generally, the factors used to square the triangle need not always be drawn from the dollar weighted averages of the triangle. Other sources of factors from which the actuary may select link ratios include simple averages from the triangle, averages weighted toward more recent observations or adjusted for outliers, and benchmark patterns based on related, more credible loss experience. Also, since the ultimate value of claims is simply the product of the most current diagonal and the cumulative product of the link ratios, the completion of interior of the triangle is usually not displayed in favor of that multiplicative calculation. For example, suppose the actuary decides that the volume weighted factors from the Claims Data triangle are representative of expected future growth, but discards the 1.005 tail factor derived from the log-linear fit in favor of a five percent tail (1.05) based on loss data from a larger book of similar business.

4.2.5: Mack (1992) Chain Ladder

Thomas Mack published in 1993 a method which estimates the standard errors of the Chain Ladder forecast without assuming a distribution under three conditions.Mack's(1992) chain ladder method calculates the standard error for the reserves estimates. The method works for a cumulative triangle C_{ij} if the following assumptions hold:

- ✓ $E\left[\frac{C_{i,j+1}}{C_{ik}}|C_{i1}C_{i2},\dots,C_{ij}\right]=f_j$

- ✓ $Var\left[\frac{C_{i,j+1}}{C_{ij}}|C_{i1}C_{i2},\dots,C_{ij}\right]=\frac{{\sigma_j}^2}{w_{ij}{C_{ij}}^{\alpha}}$
- ✓ $\{C_{i1},\dots,C_{in}\},\{C_{j1},\dots,C_{jn}\}, are\ independent\ for\ origin\ period\ i\neq j$

with$w_{ij}\in[0;1],\alpha\in\{0,1,2\}$. If these assumptions hold, the Mack Chain Ladder model gives an unbiased estimator for IBNR (Incurred But Not Reported) claims. The Mack Chain Ladder model can be regarded as a weighted linear regression through the origin for each development period.

Table 4.5: Mack (1992) Chain Ladder forecast of outstanding claims

	Latest	Deviation To Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
2004	18,834	1	18,834	0	0	NaN
2005	16,704	0.991	16,858	154	206	1.339
2006	23,466	0.974	24,083	617	623	1.01
2007	27,067	0.943	28,703	1,636	747	0.457
2008	26,180	0.905	28,927	2,747	1,469	0.535
2009	15,852	0.813	19,501	3,649	2,002	0.549
2010	12,314	0.694	17,749	5,435	2,209	0.406
2011	13,112	0.546	24,019	10,907	5,358	0.491
2012	5,395	0.336	16,045	10,650	6,333	0.595
2013	2,063	0.112	18,402	16,339	24,566	1.503
Totals	**160,987**	**0.7553784**	**213,121**	**52,135**	**26,909**	**0.52**

The loss development factors are;

Table 4.6: Mack (1992) Chain Ladder development ratios

Development Years									
1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2.99935	**1.62352**	**1.27088**	**1.17167**	**1.11338**	**1.04193**	**1.03326**	**1.01693**	**1.00921**	**1**

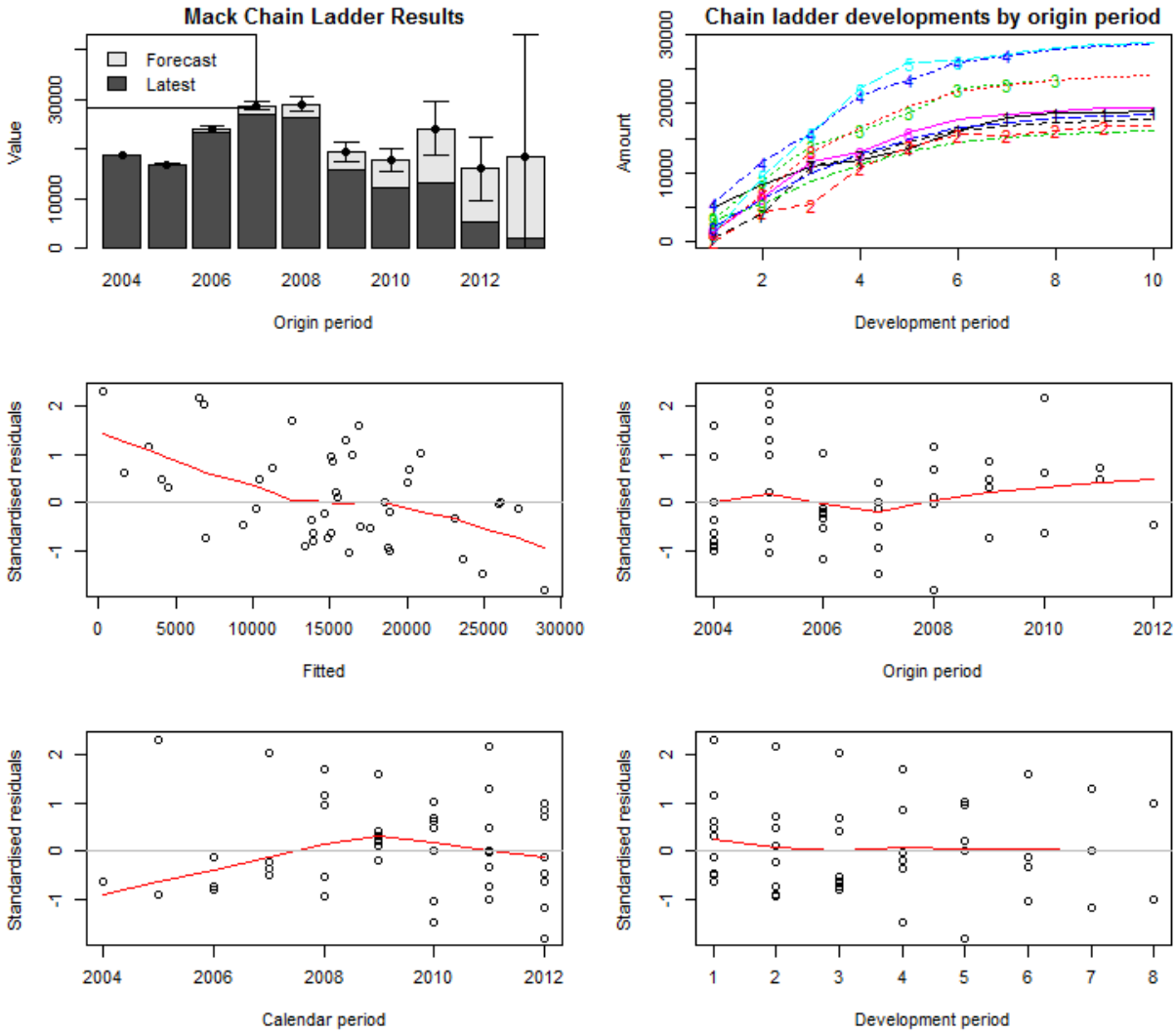
and the full triangle is given as

Table 4.7: Mack (1992) Chain Ladder full run-off triangle

	Development Years									
Origin Year	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2004	5012	8269	10907	11805	13539	16181	18009	18608	18662	18834
2005	106	4285	5396	10666	13782	15599	15707	16169	16704	16858
2006	3410	8992	13873	16141	18735	22214	22863	23466	23863	24083
2007	5655	11555	15766	21266	23425	26083	27067	27967	28441	28703
2008	1092	9565	15836	22169	25955	26180	27278	28185	28663	28927
2009	1513	6445	11702	12935	15852	17649	18390	19001	19323	19501
2010	557	4020	10946	12314	14428	16064	16738	17294	17587	17749
2011	1351	6947	13112	16664	19525	21738	22650	23403	23800	24019
2012	3133	5395	8759	11132	13043	14521	15130	15634	15898	16045
2013	2063	6188	10046	12767	14959	16655	17353	17931	18234	18402

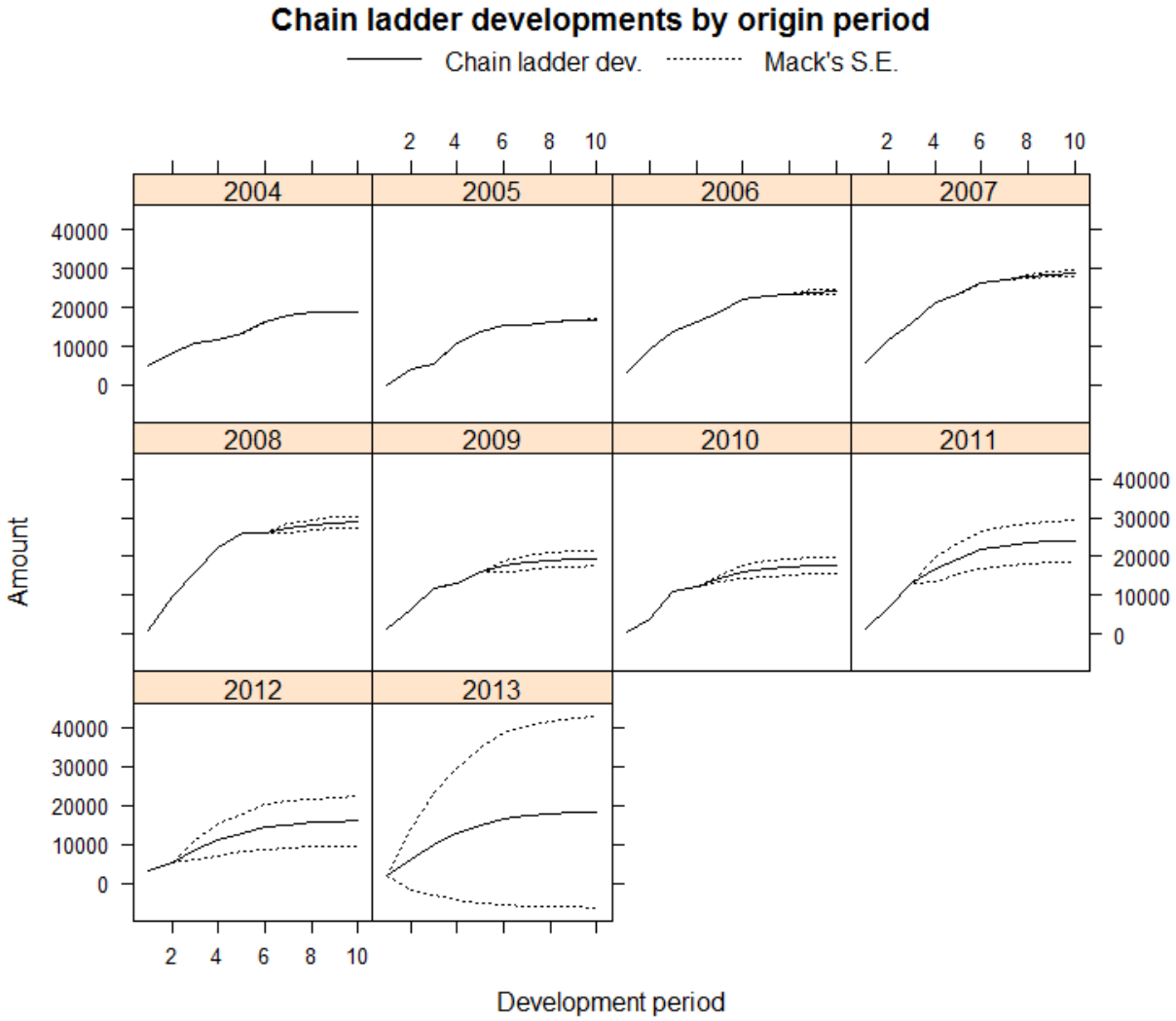
Figure 4.5: Mack (1992) Chain Ladder results

Mack's assumptions are valid as no trend show in the residual plots derived below.



We proceed to plot the development years, including the forecast and estimated standard errors by origin period

Figure 4.6: Mack (1992) Chain Ladder development patterns by origin period



4.2.6: Bootstrap Chain-Ladder

The Boot Chain Ladder function uses a two-stage bootstrapping/simulation approach following the paper by England and Verrall [PR02]. In the first stage an ordinary chain-ladder methods is applied to the cumulative claims triangle. From this we calculate the scaled Pearson residuals which we bootstrap R times to forecast future incremental claims payments via the standard Chain Ladder method. In the second stage we simulate the process error with the bootstrap value as the mean and using the

process distribution assumed. The set of reserves obtained in this way forms the predictive distribution, from which summary statistics such as mean, prediction error or quantiles can be derived.

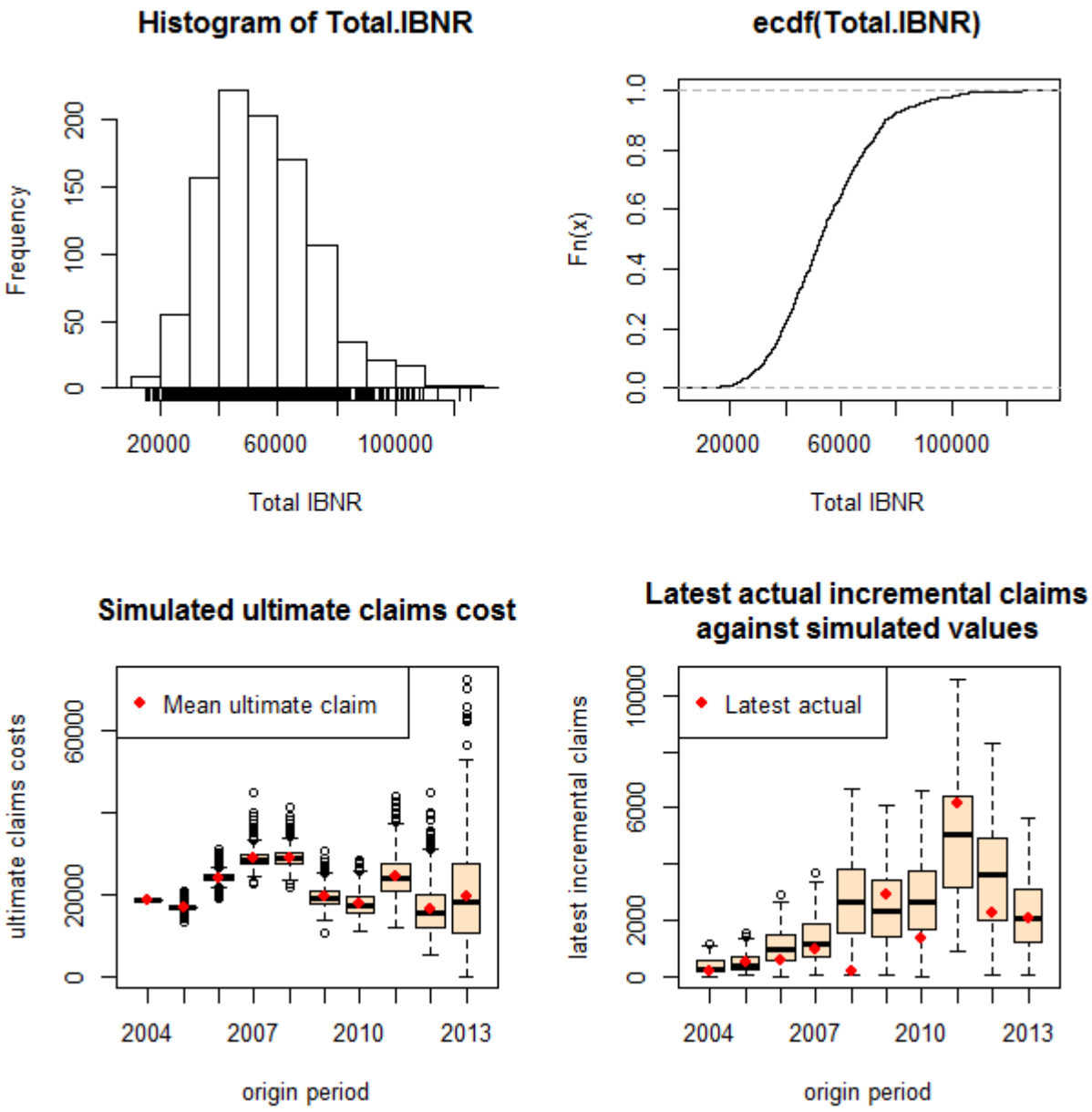
The bootstrap procedure is performed by completing the following steps, which can be performed without difficulty in a spreadsheet:

- ✓ Obtain the standard Chain-Ladder development factors from cumulative data.
- ✓ Obtain cumulative fitted values for the past triangle by backwards recursion, starting with the observed cumulative paid to date in the latest diagonal.
- ✓ Obtain incremental fitted values for the past triangle by differencing.
- ✓ Calculate the unscaled Pearson residuals for the past triangle.
- ✓ Calculate the Pearson scale parameter.

Table 4.8: Bootstrap chain ladder outstanding claims

	BootChainLadder(Triangle = Claims.Data, R = 999, process.distr = "gamma")					
Origin Year	**Latest**	**Mean Ultimate**	**Mean IBNR**	**SD IBNR**	**IBNR 75%**	**IBNR 95%**
2004	18,834	18,834	0	0	0	0
2005	16,704	16,868	164	590	201	1,269
2006	23,466	24,183	717	1,312	1,190	3,293
2007	27,067	28,862	1,795	1,968	2,834	5,235
2008	26,180	28,914	2,734	2,186	3,877	6,646
2009	15,852	19,535	3,683	2,371	4,976	8,082
2010	12,314	17,829	5,515	2,957	7,314	11,066
2011	13,112	24,425	11,313	4,938	14,383	20,221
2012	5,395	16,384	10,989	6,126	14,379	22,110
2013	2,063	19,470	17,407	13,118	25,050	41,827
Total	**160,987**	**215,304**	**54,317**	**17,807**	**65,456**	**86,952**

Figure 4.7: Bootstrap chart



Quantiles of the bootstrap IBNR are:

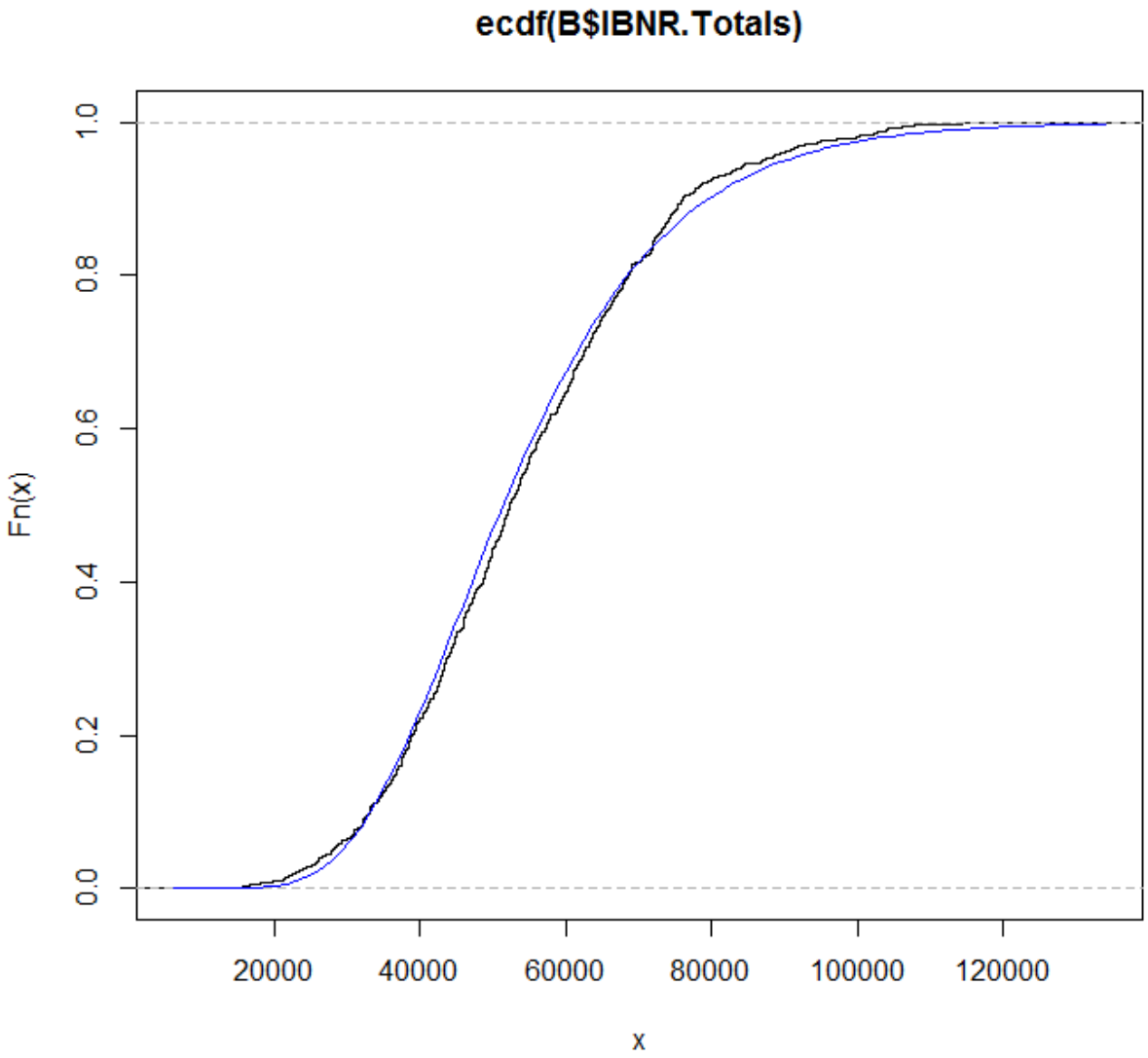
Table 4.9: Quantiles of the bootstrap IBNR

	IBNR 75%	IBNR 95%	IBNR 99%	IBNR 99.5%
2004	0	0	0	0
2005	201	1,269	2,509	2,852
2006	1,190	3,293	5,080	6,032
2007	2,834	5,235	7,841	10,297
2008	3,877	6,646	9,578	10,334
2009	4,976	8,082	10,853	11,328
2010	7,314	11,066	13,464	15,065
2011	14,383	20,221	24,888	26,684
2012	14,379	22,110	28,426	30,224
2013	25,050	41,827	54,589	61,388

	Totals
IBNR 75%:	65,456
IBNR 95%:	86,952
IBNR 99%:	104,042
IBNR 99.5%:	106,571

The distribution of the IBNR appears to follow a log-normal distribution, so let's fit it

Figure 4.8: Distribution of the IBNR



4.2.7: Munich Chain-Ladder

A reserving method that reduces the gap between IBNR projections based on paid losses and IBNR projections based on incurred losses. The IBNR reserve for a portfolio is usually calculated on the basis of both the run-off triangle of paid losses and the run-off triangle of incurred losses, i.e. the sum of paid losses and case reserves. Often, the problem arises that the projection based on paid losses is far different than the projection based on incurred losses. Even worse, paid losses may yield a higher ultimate loss projection than incurred losses in one origin year, but in the next origin year, the situation may be entirely reversed, with incurred losses yielding the higher projection of ultimate loss.

4.2.8: Clark's methods

Using a longitudinal analysis approach, Clark assumes that losses develop according to a theoretical growth curve. The LDF method is a special case of this approach where the growth curve can be considered to be either a step function or piecewise linear. Clark envisions a growth curve as measuring the percent of ultimate loss that can be expected to have emerged as of each age of an origin period. The LDF method assumes that the ultimate losses in each origin period are separate and unrelated. The goal of the method, therefore, is to estimate parameters for the ultimate losses and for the growth curve in order to maximize the likelihood of having observed the data in the triangle. The CapeCod method assumes that the a priori expected ultimate losses in each origin year are the product of earned premium that year and a theoretical loss ratio. The Cape Cod method, therefore, need estimate potentially far fewer parameters: for the growth function and for the theoretical loss ratio. One of the side benefits of using maximum likelihood to estimate parameters is that its associated asymptotic theory provides uncertainty estimates for the parameters. Observing that the reserve estimates by origin year are functions of the estimated parameters, uncertainty estimates of these functional values are calculated according to the Delta method, which is essentially a linearization of the problem based on a Taylor series expansion. The two functional forms for growth curves considered in Clark's paper are the log-logistic function (a.k.a., the inverse power curve) and the Weibull function, both being two-parameter functions. Clark uses the parameters ω and θ in his paper. Clark's methods work on incremental losses. His likelihood function is based on the assumption that incremental losses follow an over-dispersed Poisson (ODP) process.

4.2.8.1: Clark's LDF method

Consider again the Claims. Data triangle. Accepting all defaults, the Clark LDF Method would estimate total ultimate losses of 272,009 and a reserve (Future Value) of 111,022, or almost twice the value based on the volume weighted average link ratios and log linear fit in section 3.2.1(Basic Idea) above.

Table 4.10: Clark's chain ladder outstanding claims

Origin	Current Value	Ldf	Ultimate Value	Future Value	Std Error	CV%
2004	18,834	1.216	22,906	4,072	2,792	68.6
2005	16,704	1.251	20,899	4,195	2,833	67.5
2006	23,466	1.297	30,441	6,975	4,050	58.1
2007	27,067	1.36	36,823	9,756	5,147	52.8
2008	26,180	1.451	37,996	11,816	5,858	49.6
2009	15,852	1.591	25,226	9,374	4,877	52
2010	12,314	1.829	22,528	10,214	5,206	51
2011	13,112	2.305	30,221	17,109	7,568	44.2
2012	5,395	3.596	19,399	14,004	7,506	53.6
2013	2,063	12.394	25,569	23,506	17,227	73.3
Total	**160,987**		**272,009**	**111,022**	**36,102**	**32.5**

Most of the difference is due to the heavy tail, 21.6%, implied by the inverse power curve fit. Clark recognizes that the log-logistic curve can take an unreasonably long length of time to flatten out. If according to the actuary's experience most claims close as of, say, 20 years, the growth curve can be truncated accordingly.

Table 4.11: Clark's chain ladder outstanding claims truncated with a growth curve

Origin	Current Value	Ldf	Ultimate Value	Future Value	Std Error	CV%
2004	18,834	1.124	21,168	2,334	1,765	75.6
2005	16,704	1.156	19,314	2,610	1,893	72.6
2006	23,466	1.199	28,132	4,666	2,729	58.5
2007	27,067	1.257	34,029	6,962	3,559	51.1
2008	26,180	1.341	35,113	8,933	4,218	47.2
2009	15,852	1.471	23,312	7,460	3,775	50.6
2010	12,314	1.691	20,819	8,505	4,218	49.6
2011	13,112	2.13	27,928	14,816	6,300	42.5
2012	5,395	3.323	17,927	12,532	6,658	53.1
2013	2,063	11.454	23,629	21,566	15,899	73.7
Total	**160,987**		**251,369**	**90,382**	**26,375**	**29.2**

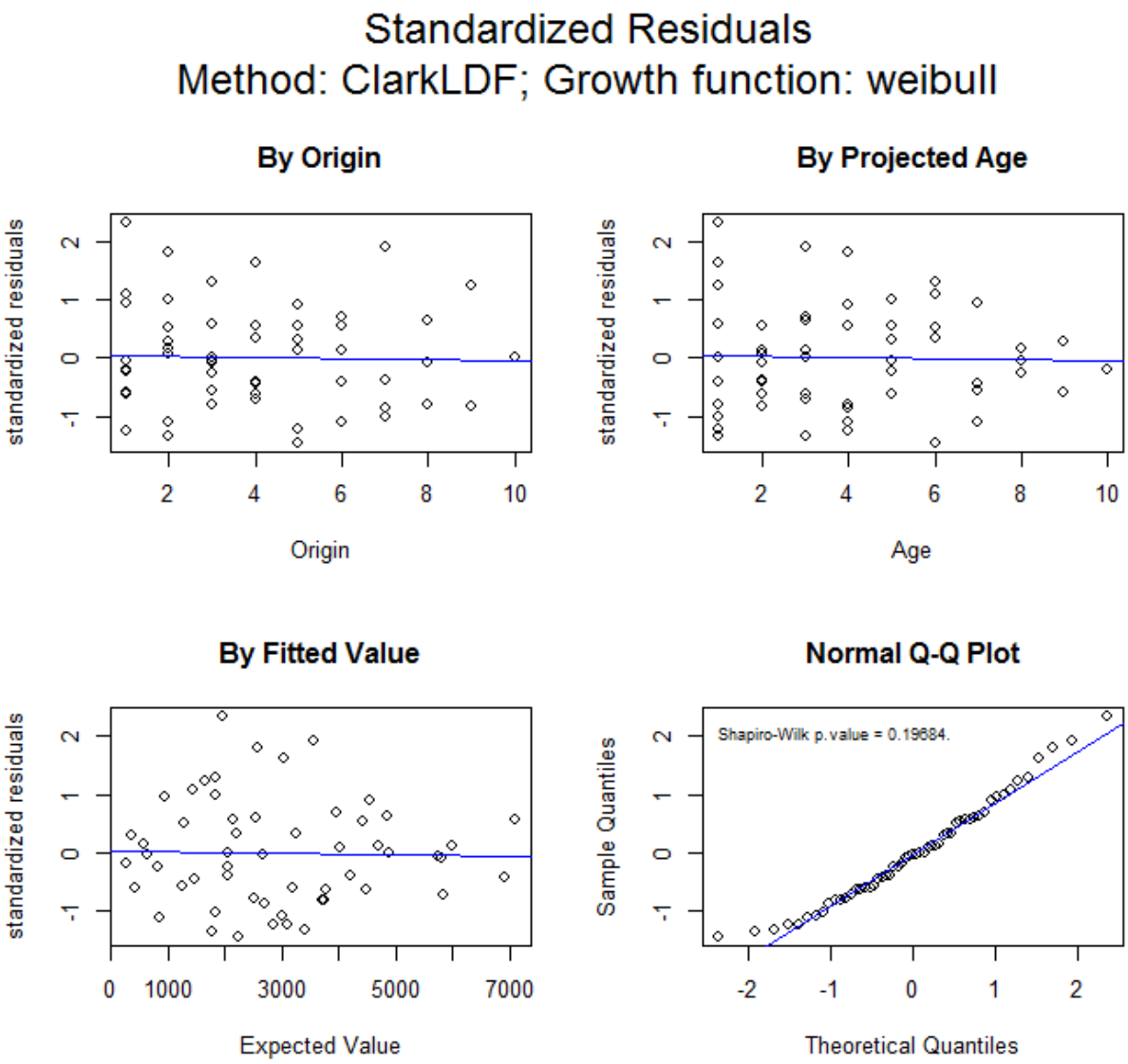
The Weibull growth curve tends to be faster developing than the log-logistic

Table 4.12: Clark's chain ladder outstanding claims using weibull growth curve

Origin	Current Value	Ldf	Ultimate Value	Future Value	Std Error	CV%
2004	18,834	1.022	19,254	420	700	166.5
2005	16,704	1.037	17,317	613	855	139.5
2006	23,466	1.06	24,875	1,409	1,401	99.4
2007	27,067	1.098	29,728	2,661	2,037	76.5
2008	26,180	1.162	30,419	4,239	2,639	62.2
2009	15,852	1.271	20,151	4,299	2,549	59.3
2010	12,314	1.471	18,114	5,800	3,060	52.8
2011	13,112	1.883	24,692	11,580	4,867	42
2012	5,395	2.988	16,122	10,727	5,544	51.7
2013	2,063	9.815	20,248	18,185	12,929	71.1
Total	**160,987**		**220,920**	**59,933**	**19,149**	**32**

It is recommend to inspect the residuals to help assess the reasonableness of the model relative to the actual data.

Figure 4.9: Clark's chart



Although there is some evidence of heteroscedasticity with increasing ages and fitted values, the residuals otherwise appear randomly scattered around a horizontal line through the origin. The q-q plot shows evidence of a lack of fit in the tails, but the p-value of almost 0.2 can be considered too high to reject outright the assumption of normally distributed standardized residuals.

4.2.8.2: Clark's Cap Cod method

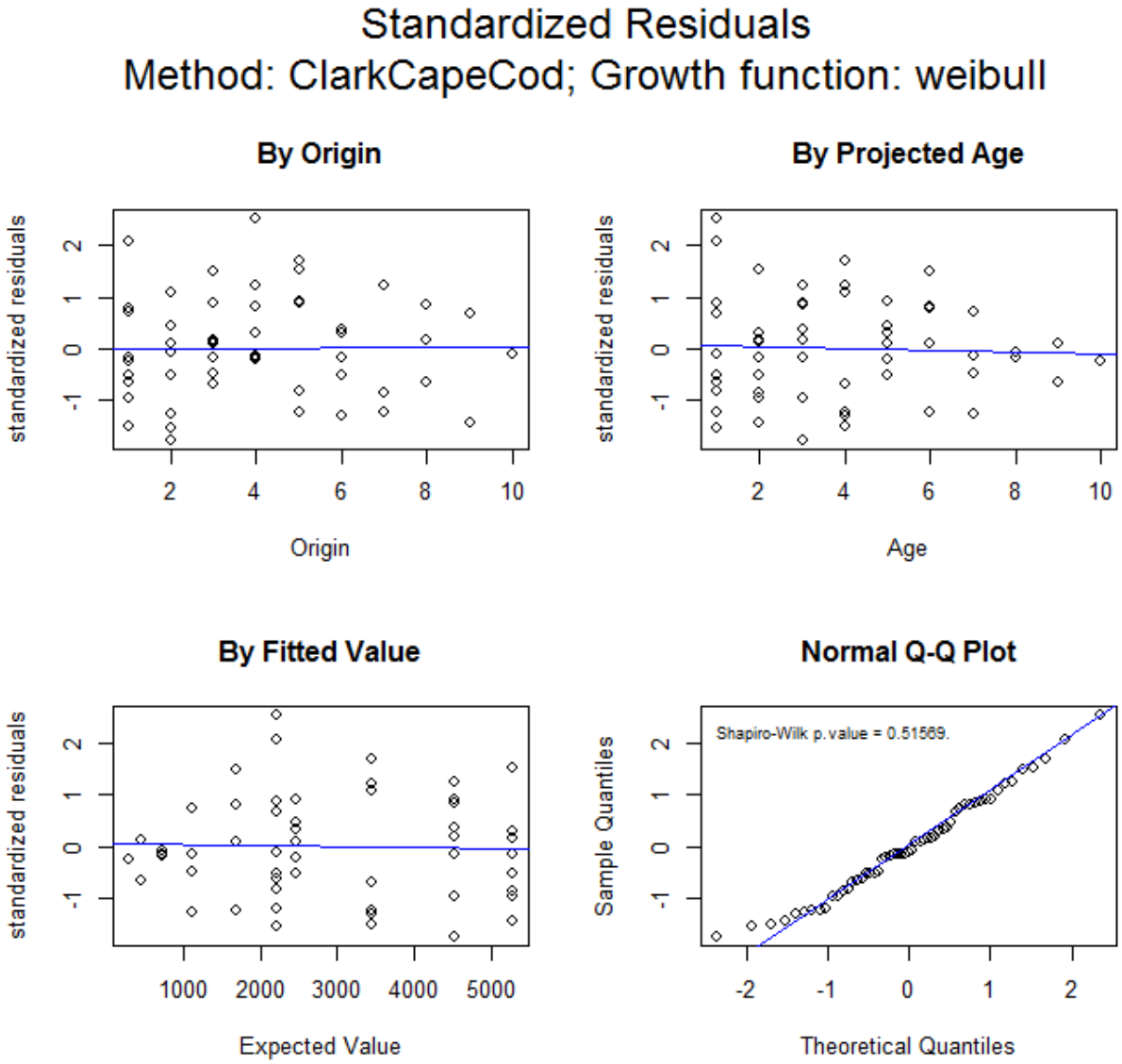
Consider the Claims.Data data set. Let's assume a constant earned premium of Kes.40,000 each year, and a Weibull growth function

Table 4.13: Clark's cap cod outstanding claims

Origin	Current Value	Premium	ELR	Future Growth Factor	Future Value	Ultimate Value	Std Error	CV%
2004	18,834	40,000	0.566	0.0192	436	19,270	692	158.6
2005	16,704	40,000	0.566	0.032	725	17,429	912	125.7
2006	23,466	40,000	0.566	0.0525	1,189	24,655	1,188	99.9
2007	27,067	40,000	0.566	0.0848	1,921	28,988	1,523	79.3
2008	26,180	40,000	0.566	0.1345	3,047	29,227	1,917	62.9
2009	15,852	40,000	0.566	0.2093	4,741	20,593	2,360	49.8
2010	12,314	40,000	0.566	0.3181	7,206	19,520	2,845	39.5
2011	13,112	40,000	0.566	0.4702	10,651	23,763	3,366	31.6
2012	5,395	40,000	0.566	0.6699	15,176	20,571	3,924	25.9
2013	2,063	40,000	0.566	0.9025	20,444	22,507	4,491	22
Total	**160,987**	**400,000**			**65,536**	**226,523**	**12,713**	**19.4**

The estimated expected loss ratio is 0.566. The total outstanding loss is about 10% higher than with the LDF method. The standard error, however, is lower, probably due to the fact that there are fewer parameters to estimate with the Cape Cod method, resulting in less parameter risk. A plot of this model shows similar residuals By Origin and Projected Age to those from the LDF method, a better spread By Fitted Value, and a slightly better q-q plot, particularly in the upper tail.

Figure 4.10: Clark's cap cod chart



4.3: Tweedie Compound Poisson Model

The distribution Y (for given weight w) is parameterized by the 3 parameters$\lambda, \tau \; and \; \gamma$. We now choose new parameters $\mu, \phi \; and \; p$ such that the density of Y can be written as, $y \geq 0$,

$$f\left(y; \mu, \frac{\phi}{w}, p\right) = c\left(y; \frac{\phi}{w}, p\right) \exp\left\{\frac{w}{\phi}\left(y\frac{\mu^{1-p}}{1-p} - \frac{\mu^{2-p}}{2-p}\right)\right\} \quad \ldots \quad 1$$

Where

$$c\left(y;\frac{\phi}{w},p\right)=\sum_{r}\frac{1}{r!\,\Gamma(r\gamma)y}\left\{\frac{(w/\phi)^{\gamma+1}y^{\gamma}}{(p-1)^{\gamma}(2-p)}\right\}^{r}$$

And

$$p=\left({}^{\gamma+2}/_{\gamma+1}\right)\in(1,2)$$

$$\mu=\lambda.\tau$$

$$\phi=\lambda^{1-p}\tau^{2-p}/(2-p)$$

If we set $\theta=\frac{\mu^{1-p}}{1-p}$ we see that the density of Y can be written as

$$f\left(y;\mu,\frac{\phi}{w},p\right)=c\left(y;\frac{\phi}{w},p\right)\exp\left\{\frac{w}{\phi}\big(y\theta-b(\theta)\big)\right\}\qquad\ldots\qquad 2$$

With

$$b(\theta)=\frac{1}{2-p}((1-p)\theta)^{\frac{2-p}{1-p}}$$

Hence, the distribution of Y belongs to the exponential family with parameters $\mu,\phi\ and\ p\epsilon(1,2)$. (McCullagh-Nelder)

We write for $p\epsilon(1,2)$

$$Y\sim ED^{(p)}(\mu,\phi/w)$$

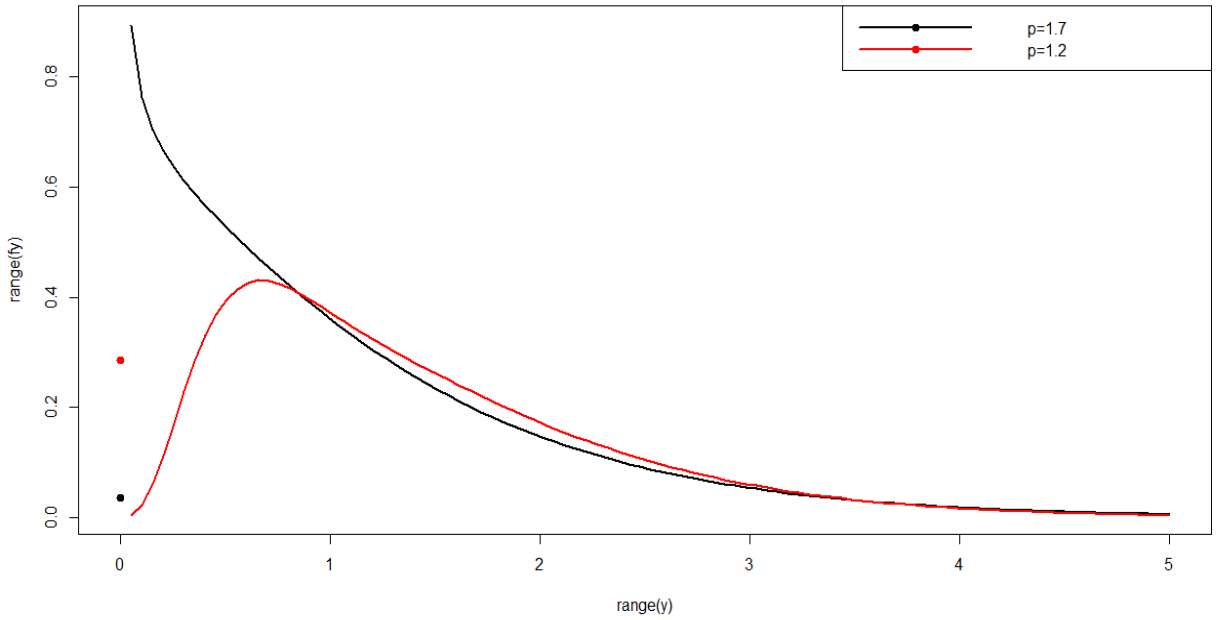
For $Y\sim ED^{(p)}(\mu,\phi/w)$ we have

$$E(Y)=b'(\theta)=\mu$$

$$Var(Y)=\frac{\phi}{w}.V(\mu)=\frac{\phi}{w}\mu^{p}$$

$\phi$is the so-called dispersion parameter and $V(\mu)$ the variance function with $p\in(1,2)$.

Figure 4.11: Tweedie density curve



4.3.1: Fitting Chain Ladder Method and Tweedie Distribution

Recent years have seen growing interest in using generalized linear models (GLM) for insurance loss reserving. The use of GLM in insurance loss reserving has many compelling aspects, e.g.

- When over-dispersed Poisson model is used, it reproduces the estimates from Chain Ladder
- It provides a more coherent modeling framework than the Mack method
- All the relevant established statistical theory can be directly applied to perform hypothesis testing and diagnostic checking

4.3.2: The Model

Let $y_1, \dots, y_n$ denote n independent observations on a response. We treat y_i as a realization of a random variable Y_i. In the general linear model we assume that Y_i has a normal distribution with mean $\mu_i$and variance σ^2 and we further assume that the expected value μ_i is a linear function of p predictors that take the values $x_i' =$

$(x_{i1}, \ldots, x_{ip})$ for the $i-th$ case, so that $Y_i = x_i'\beta$ where β is a vector of unknown parameters.

We will assume that the observations come from a distribution in the exponential family with probability density function

$$f\left(y;\mu,\frac{\phi}{w},p\right) = c\left(y;\frac{\phi}{w},p\right)\exp\left\{\frac{w}{\phi}\left(y\theta - b(\theta)\right)\right\}$$

4.3.3: The Link Function

The second element of the generalization is that instead of modeling the mean, as before, we will introduce a one-to-one continuous differentiable transformation $g(\mu_i)$ and focus on

$$\eta_i = g(\mu_i)$$

The function $g(\mu_i)$ will be called the link function. We further assume that the transformed mean follows a linear model, so that$\eta_i = X_i\beta$. The quantity η_i is called the linear predictor. Note that the model for η_i is pleasantly simple. Since the link function is one-to-one we can invert it to obtain

$$\mu_i = g^{-1}(X_i\beta)$$

The model for μ_i is usually more complicated than the model forη_i. Note that we do not transform the response y_i, but rather its expected value μ_i. A model where $\log y_i$ is linear on x_i, for example, is not the same as a generalized linear model where $\log \mu_i$ is linear onx_i.

Example: The standard linear model we have studied so far can be described as a generalized linear model with normal errors and identity link, so that

$$\mu_i = \eta_i$$

It also happens that μ_i, and therefore η_i, is the same as θ_i, the parameter in the exponential family density. When the link function makes the linear predictor η_i the same as the canonical parameter θ_i, we say that we have a canonical link. The identity is the canonical link for the normal distribution.

Error	**Link**	**Canonical Link**	**Variance Function $V(\mu)$**	**Scale ϕ**
Normal	Identity	μ	1	σ^2
Binomial	Logit	$\log \mu$	μ	1
Poisson	Log	$log\left[{}^{\mu}/_{1-\mu}\right]$	$\mu(1-\mu)$	1
Gamma	Inverse	${}^{-1}/_{\mu}$	μ^2	${}^{1}/_{\alpha}$

However, other combinations are also possible. An advantage of canonical links is that a minimal sufficient statistic for β exists, i.e. all the information about β is contained in a function of the data of the same dimensionality as β.

4.3.4: Maximum Likelihood Estimation

Given vector y, an observation of Y, maximum likelihood estimation of β is possible. Since the $Y_i$are mutually independent, the likelihood of β is

$$L(\beta) = \prod_{i=1}^{n} f\left(y_i; \mu_i, \frac{\phi}{w_i}, p\right),$$

and hence the log-likelihood of β is

$$l(\beta) = \sum_{i=1}^{n} \log(f\left(y_i; \mu_i, \frac{\phi}{w_i}, p\right)) = \sum_{i=1}^{n} \frac{w_i}{\phi}\left(y_i\theta_i - b(\theta_i)\right)$$

where the dependence of the right hand side on β is through the dependence of the θ_i on β. As discussed in the previous section, for practical work it suffices to consider only cases where we can write $\frac{\phi}{w_i}$, where w_i is a known constant.

Maximization proceeds by partially differentiating l with respect to each element of β, setting the resulting expressions to zero and solving forβ.

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\emptyset}\sum_{i=1}^{n} w_i\left(y_i\frac{\partial \theta_i}{\partial \beta_j} - b'(\theta_i)\frac{\partial \theta_i}{\partial \beta_j}\right)$$

and by the chain rule

$$\frac{\partial \theta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \beta_j},$$

so that differentiating (2), we get

$$\frac{\partial \mu_i}{\partial \theta_i} = b''(\theta_i) \Rightarrow \frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{b''(\theta_i)}$$

which then implies that

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\phi}\sum_{i=1}^{n} \frac{[y_i - b'(\theta_i)]}{b''(\theta_i)/w_i}\frac{\partial \mu_i}{\partial \beta_j}$$

Substituting from (2) and (3), into this last equation, implies that the equations to solve for β are

$$\sum_{i=1}^{n} \frac{(y_i - \mu_i)}{V(\mu_i)}\frac{\partial \mu_i}{\partial \beta_j} = 0 \quad \forall\ j. \qquad (4)$$

However, these equations are exactly the equations that would have to be solved in order to find β by non-linear weighted least squares, if the weights $V(\mu_i)$ were known in advance and were independent of β. In this case the least squares objective would be

$$S = \sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{V(\mu_i)}$$

Whereμ_i depends non-linearly on β, but the weights $V(\mu_i)$ are treated as fixed. To find the least squares estimates involves solving$\partial S/\partial \beta_j = 0 \quad \forall\, j$, but this system of equations is easily seen to be (4), when the $V(\mu_i)$ terms are treated as fixed.

This correspondence suggests a fitting method. Iterate the following two steps to convergence

1. Given the current$\hat{\mu}$ estimates, evaluate the$V(\hat{\mu})$ values

2. Find a value of β which reduces

$$\sum_{i=1}^{n} \frac{(y_i - \mu_i)^2}{V(\hat{\mu}_i)}$$

(the dependence on β is through μ, but not ^[1]). Let this improved parameter vector be denoted$\hat{\beta}$, and use it to update$\hat{\mu}$.

At convergence$\hat{\beta}$ must satisfy (4).

To implement this method we need to be able to find the required improved parameter vectors at step 2. To do this, just replace $\mu_i$by its first order Taylor expansion around μ_i, so that

$$y_i - \mu_i \simeq y_i - \hat{\mu} - \sum_j \frac{\partial \mu_i}{\partial \beta_j}(\beta_j - \hat{\beta}_j)$$

With exact equality at $\hat{\beta} = \beta$ (derivatives evaluated at current $\hat{\beta}$). Now, writing the linear predictor as $\eta_i = x_i\beta$

$$\frac{\partial \mu_i}{\partial \beta_j} = \frac{\partial \mu_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial \beta_i} = \frac{X_{ij}}{g'(\mu_i)}$$

Hence

$$\sum_i \frac{(y_i - \mu_i)^2}{V(\hat{\mu}_i)} \simeq \sum_i \frac{\{(g'(\hat{\mu}_i)y_i - g'(\hat{\mu}_i)\hat{\mu}_i) - X_i\beta + X_i\hat{\beta}\}^2}{g'(\hat{\mu}_i)^2 V(\hat{\mu}_i)} \qquad (6)$$

$$= \sum_i w_i(z_i = X_i\beta)^2 \qquad (7)$$

where$z_i = g'(\hat{\mu}_i)(y_i - \hat{\mu}_i) + X_i\beta$ and$w_i = g'(\hat{\mu}_i)^{-2}V(\hat{\mu}_i)^{-1}$. But (7) is just a weighted linear least squares problem, which is easily minimized w.r.t. β using standard least squares methods, making it easy to find an improved $\hat{\beta}$. Hence we arrive at the following GLM fitting algorithm. Iterate the following to convergence. . .

1. Given the current$\hat{\eta}\ and\ \hat{\mu}$estimates, calculate pseudo data z and weights w, as defined above.

2. Minimize

$$\sum_i w_i(z_i = X_i\beta)^2 \qquad (7)$$

with respect to β to obtain an improved estimate$\hat{\beta}$.

3. Evaluate a new linear predictor estimate $\hat{\eta} = X\hat{\beta}$ and new fitted values$\hat{\mu}_i = g^{-1}(\hat{\eta}_i)$.

If, as is usual, the method converges to fixed$\hat{\beta}$ then this must satisfy (4) and is hence the MLE of β. The iteration can be started by setting$\hat{\mu} = X\hat{\beta}$ (with modification to

avoid e.g. log (0)). The method is known as Iteratively Re-weighted Least Squares (IRLS).

The glm Reserve function in R takes an insurance loss triangle, converts it to incremental losses internally if necessary, transforms it to the long format and fits the resulting loss data with a generalized linear model where the mean structure includes both the origin year and the development lag effects. The function also provides both analytical and bootstrapping methods to compute the associated prediction errors. The bootstrapping approach also simulates the full predictive distribution, based on which the user can compute other uncertainty measures such as predictive intervals.

Only the Tweedie families of distributions are allowed. The variance power $p$is specified in the var.power argument, and controls the type of the distribution. When the Tweedie compound Poisson distribution $1 < p < 2$ is to be used, the user has the option to specify var.power = NULL, where the variance power p will be estimated from the data using the cplm package.

4.3.5: Over dispersed and Poisson Model

We first compare our result to the two boundary cases $p = 1\ and\ p = 2$. We obtain the following results:

Table 4.14: Estimated outstanding payments from the over-dispersed poisson model

	Latest	**Dev.To.Date**	**Ultimate**	**IBNR**	**S.E**	**CV**
2005	16704	0.9908649	16858	154	533.4338	3.463856
2006	23466	0.9743803	24083	617	1074.733	1.741869
2007	27067	0.9464317	28599	1532	1652.146	1.078425
2008	26180	0.9061018	28893	2713	2124.163	0.782957
2009	15852	0.8138413	19478	3626	2330.855	0.642817
2010	12314	0.6946074	17728	5414	2989.195	0.552123
2011	13112	0.5465383	23991	10879	4817.488	0.442825
2012	5395	0.3366405	16026	10631	5820.482	0.547501
2013	2063	0.1122355	18381	16318	12445.5	0.762685
Total	**142153**	**0.7326115**	**194036**	**51883**	**17400.03**	**0.335371**

Table 4.15: Estimated outstanding payments from the Gamma model

	Latest	**Dev.To.Date**	**Ultimate**	**IBNR**	**S.E**	**CV**
2005	16704	0.9918062	16842	138	134.2857	0.973085
2006	23466	0.9756361	24052	586	399.9768	0.682554
2007	27067	0.9472597	28574	1507	860.9569	0.571305
2008	26180	0.9215714	28408	2228	1153.503	0.51773
2009	15852	0.8268739	19171	3319	1760.844	0.530535
2010	12314	0.7261898	16957	4643	2417.333	0.52064
2011	13112	0.5717774	22932	9820	5216.219	0.531183
2012	5395	0.2856765	18885	13490	7955.764	0.589753
2013	2063	0.1045881	19725	17662	13293.14	0.752641
Total	**142153**	**0.726958**	**195545**	**53392**	**18010.38**	**0.337324**

It is not surprising that the over-dispersed Poisson model gives a better fit than the gamma model(especially for young accident years we have a huge estimation error term in Gamma model). Tweedie's compound poisson model converges to the Over-dispersed Poisson model for $p = 1$ and to the Gamma model for $p = 2$.

4.3.6: Tweedie Compound Poisson Model

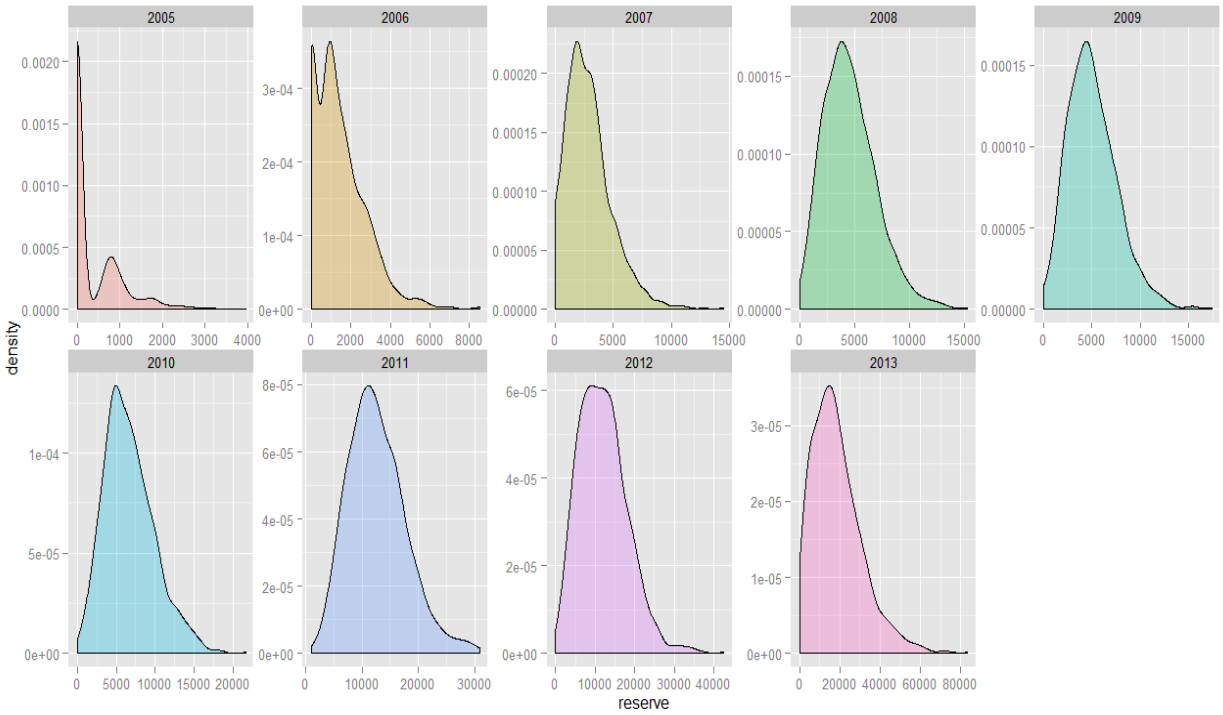
Table 4.16: Estimated outstanding payments from Tweedie compound model

	Latest	Dev.To.Date	Ultimate	IBNR	S.E	CV
2005	16704	0.991453	16848	144	238.1084	1.653531
2006	23466	0.9751496	24064	598	568.3887	0.950483
2007	27067	0.9470609	28580	1513	1022.54	0.675836
2008	26180	0.9130859	28672	2492	1378.224	0.55306
2009	15852	0.8183789	19370	3518	1781.237	0.506321
2010	12314	0.7064831	17430	5116	2452.993	0.479475
2011	13112	0.5562532	23572	10460	4656.158	0.445139
2012	5395	0.3135534	17206	11811	6168.482	0.522266
2013	2063	0.1103858	18689	16626	12091.37	0.727257
Total	**142153**	**0.7311269**	**194430**	**52277**	**16225.34**	**0.310373**

A look at the result shows that the Tweedie's compound Poisson model is close to the chain ladder estimates

The full predictive distribution of the simulated reserves for each year can be visualized as:

Figure 4.12: Full predictive distribution of the simulated reserves



CONCLUSION

Tweedie distribution method response provides an expedient method for the segmentation and pure premium estimation of an insurance portfolio.In Tweedie generalized linear models all the groups have the same dispersion parameter, while the standard method allocates different values to each group.This suggests that, in a certain situation, one technique will fit better than the other, depending if one common dispersion parameter is sufficient to fit the variance of the diverse homogeneous groups or not.

The fused approach has the advantage that its goodness of fit and distributional assumptions can be tested using the theory of generalized linear models. Conversely, the standard approach requires diverse techniques for its assessment. In conclusion, the Tweedie distribution model is a good challenger to the standard estimation method used in the industry. It should always be considered in the set of models to evaluate when the total outstanding claim size is thought to follow a Compound Poisson gamma distribution.

REFERENCES

Frees EW, Meyers G, Cummings DA (2011).*Summarizing Insurance Scores Using a Gini Index.*Journal of the American Statistical Association, 495, 1085-1098.

Jorgensen, B. (1986). *Some properties of exponential dispersion models.* Scandinavian Journal of Statistics 13 (3), 187-197.

Jorgensen, B. (1997). *The Theory of Dispersion Models (First ed.).*Chapman & Hall.

Klugman, S.A., P. H. and G. Willmot (2004).*Loss Models: From Data to Decisions (second ed.)*. Wiley.

Madsen, H. and P. Thyregod (2011).*Introduction to General and Generalized Linear Models*.CRC Press.

McCullagh, P., Nelder, J.A.*: Generalized Linear Models, 2nd edn.*Chapman and Hall, London (1989)

Nelder J, Pregibon D (1987). *An Extended Quasi-likelihood Function.*Biometrika, 74, 221-232.

Peters GW, Shevchenko PV,Wüthrich MV (2009).*Model Uncertainty in Claims Reserving within Tweedie's Compound Poisson Models.* ASTIN Bulletin, 39, 1-33.

R Development Core Team (2010).*R: A Language and Environment for Statistical Computing. Vienna, Austria*: R Foundation for Statistical Computing. ISBN 3-900051-07-0.http://www.R-project.org.

Shono H (2008*). Application of the Tweedie Distribution to Zero-catch Data in CPUE Analysis*.Fisheries Research, 93, 154-162.

Smyth G, Jorgensen B (2002). *Fitting Tweedie's Compound Poisson Model to Insurance Claims Data: Dispersion Modelling.* ASTIN Bulletin, 32, 143-157.

Yip KCH, Yau KKW (2005). *On Modeling Claim Frequency Data In General Insurance With Extra Zeros.* Insurance: Mathematics and Economics, 36, 153-163.

APPENDIX 1:R code

```
####LOADING THE CHAINLADDER PACKAGE


> library(stats)
> library(statmod)
> library(tweedie)
> getwd()
        [1] "C:/Users/hmbauni/Documents"
> setwd("C:/Users/hmbauni/Documents/project/data")



> ####CHAINLADDER
> library(proto)
> library(plyr)
> library(digest)
> library(gtable)
Loading required package: grid
> library(reshape2)
> library(scales)
> library(munsell)
> library(colorspace)
```

Attaching package: ‘colorspace’

The following object is masked from ‘package:munsell’:

    desaturate

> library(labeling)

> library(RColorBrewer)

> library(stringr)

> library(dichromat)

> library(testthat)

> library(evaluate)

> library(DBI)

> library(biglm)

> library(coda)

Loading required package: lattice

> library(ggplot2)

> library(lme4)

Loading required package: Matrix

Attaching package: ‘lme4’

The following object is masked from ‘package:coda’:

    HPDinterval

The following object is masked from ‘package:stats’:

    AIC, BIC

```
> library(Matrix)
> library(Rcpp)
> library(minqa)
> library(cplm)
Loading required package: splines
> library(lattice)
> library(MASS)
> library(zoo)

Attaching package: ‘zoo’

The following object is masked from ‘package:base’:

    as.Date, as.Date.numeric
```

```
> library(lmtest)
> library(car)
Loading required package: nnet
> library(sandwich)
> library(strucchange)
> library(dynlm)
> library(survival)
> library(Formula)
> library(AER)
> library(systemfit)
> library(Hmisc)
Hmisc library by Frank E Harrell Jr

Type library(help='Hmisc'), ?Overview, or ?Hmisc.Overview')
to see overall documentation.

NOTE:Hmisc no longer redefines [.factor to drop unused levels when
subsetting.  To get the old behavior of Hmisc type dropUnusedLevels().

Attaching package: 'Hmisc'
```

```
The following object is masked from 'package:survival':

    untangle.specials

The following object is masked from 'package:car':

    recode

The following object is masked from 'package:plyr':

    is.discrete, summarize

The following object is masked from 'package:base':

    format.pval, round.POSIXt, trunc.POSIXt, units

> library(actuar)

Attaching package: 'actuar'

The following object is masked from 'package:grDevices':
```

cm

```
> library(ChainLadder)

        ChainLadder version 0.1.5-6 by:
        Markus Gesmann <markus.gesmann@gmail.com>
        Wayne Zhang <actuary_zhang@hotmail.com>
        Daniel Murphy <danielmarkmurphy@gmail.com>

        Type ?ChainLadder to access overall documentation and
        vignette('ChainLadder') for the package vignette.

        Type demo(ChainLadder) to get an idea of the functionality of this
package.
        See demo(package='ChainLadder') for a list of more demos.

        More information is available on the ChainLadder project web-
site:
        http://code.google.com/p/chainladder/

        To suppress this message use the statement:
        suppressPackageStartupMessages(library(ChainLadder))
```

```
###LOAD DATA
> ppp <- read.table("ppp.txt",header=TRUE)
> summary(ppp)
        Origin          dev          value        lob
    Min.   :2004   Min.   : 1   Min.   :-103   RAA:55
    1st Qu.:2005   1st Qu.: 2   1st Qu.:1102
    Median :2007   Median : 4   Median :2638
    Mean   :2007   Mean   : 4   Mean   :2927
    3rd Qu.:2009   3rd Qu.: 6   3rd Qu.:4906
    Max.   :2013   Max.   :10   Max.   :8473

> ppp1 <- subset(ppp, lob %in% "RAA")
> head(ppp1)
     Origin dev value lob
1   2004   1  5012 IRA
2   2005   1   106 IRA
3   2006   1  3410 IRA
4   2007   1  5655 IRA
5   2008   1  1092 IRA
```

```
6   2009   1  1513 IRA
```

```
> ppp.tri <- as.triangle(ppp1, origin="Origin", dev="dev", value="value")
> ppp.tri
          dev
Origin    1    2    3    4    5    6    7   8   9  10
  2004 5012 3257 2638  898 1734 2642 1828 599  54 172
  2005  106 4179 1111 5270 3116 1817  103 673 535  NA
  2006 3410 5582 4881 2268 2594 3479  649 603  NA  NA
  2007 5655 5900 4211 5500 2159 2658  984  NA  NA  NA
  2008 1092 8473 6271 6333 3786  225   NA  NA  NA  NA
  2009 1513 4932 5257 1233 2917   NA   NA  NA  NA  NA
  2010  557 3463 6926 1368   NA   NA   NA  NA  NA  NA
  2011 1351 5596 6165   NA   NA   NA   NA  NA  NA  NA
  2012 3133 2262   NA   NA   NA   NA   NA  NA  NA  NA
  2013 2063   NA   NA   NA   NA   NA   NA  NA  NA  NA
```

```
##Conversion of data from Incremental to Cumulative
> Claims.Data <- incr2cum(ppp.tri)
> Claims.Data
```

```
    dev
Origin  1     2     3     4     5     6     7     8     9     10
  2004 5012  8269 10907 11805 13539 16181 18009 18608 18662 18834
  2005  106  4285  5396 10666 13782 15599 15496 16169 16704    NA
  2006 3410  8992 13873 16141 18735 22214 22863 23466    NA    NA
  2007 5655 11555 15766 21266 23425 26083 27067    NA    NA    NA
  2008 1092  9565 15836 22169 25955 26180    NA    NA    NA    NA
  2009 1513  6445 11702 12935 15852    NA    NA    NA    NA    NA
  2010  557  4020 10946 12314    NA    NA    NA    NA    NA    NA
  2011 1351  6947 13112    NA    NA    NA    NA    NA    NA    NA
  2012 3133  5395    NA    NA    NA    NA    NA    NA    NA    NA
  2013 2063    NA    NA    NA    NA    NA    NA    NA    NA    NA
> plot(Claims.Data)
> plot(Claims.Data, lattice=TRUE)
```

```
> ##Chain-ladder methods
> #Basic idea
> n <- 10
> f <- sapply(1:(n-1),function(i){
+ sum(Claims.Data[c(1:(n-i)),i+1])/sum(Claims.Data[c(1:(n-i)),i])
+ }
+ )
> f
[1] 2.999359 1.623523 1.270888 1.171675 1.113385
[6] 1.041935 1.033264 1.016936 1.009217
```

```
> #Often it is not suitable to assume that the oldest origin year is fully developed. A
> #typical approach is to extrapolate the development ratios, e.g. assuming a log-linear
> #model.
> dev.period <- 1:(n-1)
> plot(log(f-1) ~ dev.period, main="Log-linear extrapolation of age-to-age factors")
> tail.model <- lm(log(f-1) ~ dev.period)
> abline(tail.model)
> co <- coef(tail.model)
```

```
>
> # extrapolate another 100 dev. period
> tail <- exp(co[1] + c((n + 1):(n + 100)) * co[2]) + 1
> f.tail <- prod(tail)
> f.tail
[1] 1.005006
> plot(100*(rev(1/cumprod(rev(c(f, tail[tail>1.0001]))))), t="b",
+ main="Expected claims development pattern",
+ xlab="Dev. period", ylab="Development % of ultimate loss")
```

```
> #The link ratios are then applied to the latest known cumulative claims amount to
> #forecast the next development period
> f <- c(f, f.tail)
> fullClaims.Data <- cbind(Claims.Data, Ult = rep(0, 10))
> for(k in 1:n){
+ fullClaims.Data[(n-k+1):n, k+1] <- fullClaims.Data[(n-k+1):n,k]*f[k]
+ }
> round(fullClaims.Data)
```

	1	2	3	4	5	6	7	8	9	10	Ult

2004 5012 8269 10907 11805 13539 16181 18009 18608 18662 18834 18928

2005 106 4285 5396 10666 13782 15599 15496 16169 16704 16858 16942

2006 3410 8992 13873 16141 18735 22214 22863 23466 23863 24083 24204

2007 5655 11555 15766 21266 23425 26083 27067 27967 28441 28703 28847

2008 1092 9565 15836 22169 25955 26180 27278 28185 28663 28927 29072

2009 1513 6445 11702 12935 15852 17649 18389 19001 19323 19501 19599

2010 557 4020 10946 12314 14428 16064 16738 17294 17587 17749 17838

2011 1351 6947 13112 16664 19525 21738 22650 23403 23800 24019 24139

2012 3133 5395 8759 11132 13043 14521 15130 15634 15898 16045 16125

2013 2063 6188 10046 12767 14959 16655 17353 17931 18234 18402 18495

> #The total estimated outstanding loss

> sum(fullClaims.Data[,11] - getLatestCumulative(Claims.Data))

[1] 53202.12

```
> ####Mack ChainLadder
> mack <- MackChainLadder(Claims.Data, est.sigma="Mack")
> mack
        MackChainLadder(Triangle = Claims.Data, est.sigma = "Mack")
```

	Latest	Dev.To.Date	Ultimate	IBNR	Mack.S.E	CV(IBNR)
2004	18,834	1.000	18,834	0	0	NaN
2005	16,704	0.991	16,858	154	206	1.339
2006	23,466	0.974	24,083	617	623	1.010
2007	27,067	0.943	28,703	1,636	747	0.457
2008	26,180	0.905	28,927	2,747	1,469	0.535
2009	15,852	0.813	19,501	3,649	2,002	0.549
2010	12,314	0.694	17,749	5,435	2,209	0.406
2011	13,112	0.546	24,019	10,907	5,358	0.491
2012	5,395	0.336	16,045	10,650	6,333	0.595
2013	2,063	0.112	18,402	16,339	24,566	1.503

```
        Totals
Latest:     160,987.00
Dev:              0.76
Ultimate:   213,122.23
```

IBNR: 52,135.23

Mack S.E.: 26,909.01

CV(IBNR): 0.52

> mack$f

[1] 2.999359 1.623523 1.270888 1.171675 1.113385 1.041935 1.033264 1.016936 1.009217 1.000000

> mack$FullTriangle

dev

origin 1 2 3 4 5 6 7 8 9 10

2004 5012 8269.000 10907.000 11805.00 13539.00 16181.00 18009.00 18608.00 18662.00 18834.00

2005 106 4285.000 5396.000 10666.00 13782.00 15599.00 15496.00 16169.00 16704.00 16857.95

2006 3410 8992.000 13873.000 16141.00 18735.00 22214.00 22863.00 23466.00 23863.43 24083.37

2007 5655 11555.000 15766.000 21266.00 23425.00 26083.00 27067.00 27967.34 28441.01 28703.14

2008 1092 9565.000 15836.000 22169.00 25955.00 26180.00 27277.85 28185.21 28662.57 28926.74

2009 1513 6445.000 11702.000 12935.00 15852.00 17649.38 18389.50 19001.20 19323.01 19501.10

2010 557 4020.000 10946.000 12314.00 14428.00 16063.92 16737.55 17294.30 17587.21 17749.30

2011 1351 6947.000 13112.000 16663.88 19524.65 21738.45 22650.05 23403.47 23799.84 24019.19

2012 3133 5395.000 8758.905 11131.59 13042.60 14521.43 15130.38 15633.68 15898.45 16044.98

2013 2063 6187.677 10045.834 12767.13 14958.92 16655.04 17353.46 17930.70 18234.38 18402.44

> plot(mack)

> ClarkLDF(Claims.Data)

Origin	CurrentValue	Ldf	UltimateValue	FutureValue	StdError	CV%
2004	18,834	1.216	22,906	4,072	2,792	68.6
2005	16,704	1.251	20,899	4,195	2,833	67.5
2006	23,466	1.297	30,441	6,975	4,050	58.1
2007	27,067	1.360	36,823	9,756	5,147	52.8
2008	26,180	1.451	37,996	11,816	5,858	49.6
2009	15,852	1.591	25,226	9,374	4,877	52.0
2010	12,314	1.829	22,528	10,214	5,206	51.0
2011	13,112	2.305	30,221	17,109	7,568	44.2
2012	5,395	3.596	19,399	14,004	7,506	53.6
2013	2,063	12.394	25,569	23,506	17,227	73.3
Total	160,987		272,009	111,022	36,102	32.5

> ClarkLDF(Claims.Data,maxage=20)

Origin CurrentValue Ldf UltimateValue FutureValue StdError CV%

2004	18,834	1.124	21,168	2,334	1,765	75.6
2005	16,704	1.156	19,314	2,610	1,893	72.6
2006	23,466	1.199	28,132	4,666	2,729	58.5
2007	27,067	1.257	34,029	6,962	3,559	51.1
2008	26,180	1.341	35,113	8,933	4,218	47.2
2009	15,852	1.471	23,312	7,460	3,775	50.6
2010	12,314	1.691	20,819	8,505	4,218	49.6
2011	13,112	2.130	27,928	14,816	6,300	42.5
2012	5,395	3.323	17,927	12,532	6,658	53.1
2013	2,063	11.454	23,629	21,566	15,899	73.7
Total	160,987		251,369	90,382	26,375	29.2

> ClarkLDF(Claims.Data, G="weibull")

Origin	CurrentValue	Ldf	UltimateValue	FutureValue	StdError	CV%
2004	18,834	1.022	19,254	420	700	166.5
2005	16,704	1.037	17,317	613	855	139.5
2006	23,466	1.060	24,875	1,409	1,401	99.4
2007	27,067	1.098	29,728	2,661	2,037	76.5
2008	26,180	1.162	30,419	4,239	2,639	62.2
2009	15,852	1.271	20,151	4,299	2,549	59.3
2010	12,314	1.471	18,114	5,800	3,060	52.8
2011	13,112	1.883	24,692	11,580	4,867	42.0
2012	5,395	2.988	16,122	10,727	5,544	51.7

2013 2,063 9.815 20,248 18,185 12,929 71.1

Total 160,987 220,920 59,933 19,149 32.0

```
> plot(ClarkLDF(Claims.Data, G="weibull"))
> ClarkCapeCod(Claims.Data, Premium = 40000, G = "weibull")
```

Origin CurrentValue Premium ELR FutureGrowthFactor FutureValue UltimateValue StdError CV%

2004 18,834 40,000 0.566 0.0192 436 19,270 692 158.6

2005 16,704 40,000 0.566 0.0320 725 17,429 912 125.7

2006 23,466 40,000 0.566 0.0525 1,189 24,655 1,188 99.9

2007 27,067 40,000 0.566 0.0848 1,921 28,988 1,523 79.3

2008 26,180 40,000 0.566 0.1345 3,047 29,227 1,917 62.9

2009 15,852 40,000 0.566 0.2093 4,741 20,593 2,360 49.8

2010 12,314 40,000 0.566 0.3181 7,206 19,520 2,845 39.5

2011 13,112 40,000 0.566 0.4702 10,651 23,763 3,366 31.6

2012 5,395 40,000 0.566 0.6699 15,176 20,571 3,924 25.9

2013 2,063 40,000 0.566 0.9025 20,444 22,507 4,491 22.0

Total 160,987 400,000 65,536 226,523 12,713 19.4

> ClarkCapeCod(Claims.Data, Premium = 40000, G = "weibull")

Origin CurrentValue Premium ELR FutureGrowthFactor FutureValue UltimateValue StdError CV%

2004 18,834 40,000 0.566 0.0192 436 19,270 692 158.6

2005 16,704 40,000 0.566 0.0320 725 17,429 912 125.7

2006 23,466 40,000 0.566 0.0525 1,189 24,655 1,188 99.9

2007 27,067 40,000 0.566 0.0848 1,921 28,988 1,523 79.3

2008 26,180 40,000 0.566 0.1345 3,047 29,227 1,917 62.9

2009 15,852 40,000 0.566 0.2093 4,741 20,593 2,360 49.8

2010 12,314 40,000 0.566 0.3181 7,206 19,520 2,845 39.5

2011 13,112 40,000 0.566 0.4702 10,651 23,763 3,366 31.6

2012 5,395 40,000 0.566 0.6699 15,176 20,571 3,924 25.9

2013 2,063 40,000 0.566 0.9025 20,444 22,507 4,491 22.0

Total 160,987 400,000 65,536 226,523 12,713 19.4

```
> plot(ClarkCapeCod(Claims.Data, Premium = 40000, G = "weibull"))
```

#####Tweedie Distribution

```
> tra <- read.csv("tra.csv",header=TRUE)
> tra
   origin dev value lob
1    2004   1  5012 IRA
2    2005   1   106 IRA
3    2006   1  3410 IRA
4    2007   1  5655 IRA
5    2008   1  1092 IRA
6    2009   1  1513 IRA
7    2010   1   557 IRA
8    2011   1  1351 IRA
9    2012   1  3133 IRA
10   2013   1  2063 IRA
11   2004   2  8269 IRA
12   2005   2  4285 IRA
```

13 2006 2 8992 IRA

14 2007 2 11555 IRA

15 2008 2 9565 IRA

16 2009 2 6445 IRA

17 2010 2 4020 IRA

18 2011 2 6947 IRA

19 2012 2 5395 IRA

20 2004 3 10907 IRA

21 2005 3 5396 IRA

22 2006 3 13873 IRA

23 2007 3 15766 IRA

24 2008 3 15836 IRA

25 2009 3 11702 IRA

26 2010 3 10946 IRA

27 2011 3 13112 IRA

28 2004 4 11805 IRA

29 2005 4 10666 IRA

30 2006 4 16141 IRA

31 2007 4 21266 IRA

32 2008 4 22169 IRA

33 2009 4 12935 IRA

34 2010 4 12314 IRA

```
35   2004   5 13539 IRA
36   2005   5 13782 IRA
37   2006   5 18735 IRA
38   2007   5 23425 IRA
39   2008   5 25955 IRA
40   2009   5 15852 IRA
41   2004   6 16181 IRA
42   2005   6 15599 IRA
43   2006   6 22214 IRA
44   2007   6 26083 IRA
45   2008   6 26180 IRA
46   2004   7 18009 IRA
47   2005   7 15702 IRA
48   2006   7 22863 IRA
49   2007   7 27067 IRA
50   2004   8 18608 IRA
51   2005   8 16169 IRA
52   2006   8 23466 IRA
53   2004   9 18662 IRA
54   2005   9 16704 IRA
55   2004  10 18834 IRA
> tra.tri <- as.triangle(tra, origin="origin", dev="dev", value="value")
```

> tra.tri

dev

origin 1 2 3 4 5 6 7 8 9 10

2004 5012 8269 10907 11805 13539 16181 18009 18608 18662 18834

2005 106 4285 5396 10666 13782 15599 15702 16169 16704 NA

2006 3410 8992 13873 16141 18735 22214 22863 23466 NA NA

2007 5655 11555 15766 21266 23425 26083 27067 NA NA NA

2008 1092 9565 15836 22169 25955 26180 NA NA NA NA

2009 1513 6445 11702 12935 15852 NA NA NA NA NA

2010 557 4020 10946 12314 NA NA NA NA NA NA

2011 1351 6947 13112 NA NA NA NA NA NA NA

2012 3133 5395 NA NA NA NA NA NA NA NA

2013 2063 NA NA NA NA NA NA NA NA NA

####Fit the Poisson, Gamma and a Tweedie compound Poisson GLM reserving model by

changing the var.power argument:

```
> # Fit Poisson GLM
> (fit1 <- glmReserve(tra.tri, var.power = 1))
      Latest Dev.To.Date Ultimate  IBNR       S.E       CV
2005  16704  0.9908649    16858   154   533.4338 3.4638557
2006  23466  0.9743803    24083   617  1074.7330 1.7418688
2007  27067  0.9464317    28599  1532  1652.1463 1.0784245
2008  26180  0.9061018    28893  2713  2124.1634 0.7829574
2009  15852  0.8138413    19478  3626  2330.8553 0.6428172
2010  12314  0.6946074    17728  5414  2989.1954 0.5521233
2011  13112  0.5465383    23991 10879  4817.4876 0.4428245
2012   5395  0.3366405    16026 10631  5820.4819 0.5475009
2013   2063  0.1122355    18381 16318 12445.4966 0.7626852
total 142153  0.7326115   194036 51883 17400.0293 0.3353705
> (fit1 <- glmReserve(tra.tri))
      Latest Dev.To.Date Ultimate  IBNR       S.E       CV
2005  16704  0.9908649    16858   154   533.4338 3.4638557
2006  23466  0.9743803    24083   617  1074.7330 1.7418688
2007  27067  0.9464317    28599  1532  1652.1463 1.0784245
2008  26180  0.9061018    28893  2713  2124.1634 0.7829574
```

```
2009  15852  0.8138413   19478  3626  2330.8553 0.6428172
2010  12314  0.6946074   17728  5414  2989.1954 0.5521233
2011  13112  0.5465383   23991 10879  4817.4876 0.4428245
2012   5395  0.3366405   16026 10631  5820.4819 0.5475009
2013   2063  0.1122355   18381 16318 12445.4966 0.7626852
total 142153  0.7326115  194036 51883 17400.0293 0.3353705
```

```
> summary(fit1, type="model")

Call:
glm(formula = value ~ factor(origin) + factor(dev), family = fam,
    data = ldaFit, offset = offset)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-61.40  -22.63    0.00   14.51   53.46

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          7.65628    0.30643  24.985   <2e-16 ***
factor(origin)2005  -0.11084    0.33112  -0.335   0.7398
factor(origin)2006   0.24586    0.30566   0.804   0.4265
```

factor(origin)2007	0.41769	0.29739	1.405	0.1687
factor(origin)2008	0.42792	0.30027	1.425	0.1627
factor(origin)2009	0.03363	0.33948	0.099	0.9216
factor(origin)2010	-0.06050	0.36641	-0.165	0.8698
factor(origin)2011	0.24201	0.36320	0.666	0.5095
factor(origin)2012	-0.16145	0.49348	-0.327	0.7454
factor(origin)2013	-0.02436	0.74990	-0.032	0.9743
factor(dev)2	0.69283	0.25771	2.688	0.0108 *
factor(dev)3	0.62603	0.26720	2.343	0.0248 *
factor(dev)4	0.27695	0.29899	0.926	0.3605
factor(dev)5	0.06056	0.32798	0.185	0.8545
factor(dev)6	-0.19582	0.37290	-0.525	0.6027
factor(dev)7	-1.02355	0.56910	-1.799	0.0805 .
factor(dev)8	-1.39123	0.79830	-1.743	0.0899 .
factor(dev)9	-1.91711	1.30774	-1.466	0.1513
factor(dev)10	-2.50878	2.39009	-1.050	0.3009

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Tweedie family taken to be 966.4027)

Null deviance: 91801 on 54 degrees of freedom

Residual deviance: 36939 on 36 degrees of freedom

AIC: NA

Number of Fisher Scoring iterations: 5

```
> # Fit Gamma GLM
> (fit2 <- glmReserve(tra.tri, var.power = 2))
```

	Latest	Dev.To.Date	Ultimate	IBNR	S.E	CV
2005	16704	0.9918062	16842	138	134.2857	0.9730845
2006	23466	0.9756361	24052	586	399.9768	0.6825542
2007	27067	0.9472597	28574	1507	860.9569	0.5713052
2008	26180	0.9215714	28408	2228	1153.5032	0.5177304
2009	15852	0.8268739	19171	3319	1760.8440	0.5305345
2010	12314	0.7261898	16957	4643	2417.3333	0.5206404
2011	13112	0.5717774	22932	9820	5216.2189	0.5311832
2012	5395	0.2856765	18885	13490	7955.7640	0.5897527
2013	2063	0.1045881	19725	17662	13293.1434	0.7526409
total	142153	0.7269580	195545	53392	18010.3764	0.3373235

```
> summary(fit2, type="model")
```

Call:

```
glm(formula = value ~ factor(origin) + factor(dev), family = fam,
    data = ldaFit, offset = offset)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9579  -0.4419   0.0000   0.2560   0.9285

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         7.744375   0.315709  24.530  < 2e-16 ***
factor(origin)2005 -0.221657   0.308152  -0.719  0.47659
factor(origin)2006  0.095299   0.322274   0.296  0.76915
factor(origin)2007  0.306333   0.337690   0.907  0.37036
factor(origin)2008  0.145025   0.356022   0.407  0.68617
factor(origin)2009 -0.179546   0.379268  -0.473  0.63878
factor(origin)2010 -0.365953   0.410817  -0.891  0.37896
factor(origin)2011 -0.009263   0.457676  -0.020  0.98396
factor(origin)2012 -0.095863   0.538134  -0.178  0.85961
factor(origin)2013 -0.112459   0.725936  -0.155  0.87775
factor(dev)2        0.756081   0.308152   2.454  0.01912 *
factor(dev)3        0.759889   0.322274   2.358  0.02393 *
factor(dev)4        0.331548   0.337690   0.982  0.33274
```

```
factor(dev)5      0.165020   0.356022   0.464  0.64579
factor(dev)6     -0.121333   0.379268  -0.320  0.75088
factor(dev)7     -1.037863   0.410817  -2.526  0.01606 *
factor(dev)8     -1.387070   0.457676  -3.031  0.00450 **
factor(dev)9     -1.855797   0.538134  -3.449  0.00145 **
factor(dev)10    -2.596881   0.725936  -3.577  0.00101 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for Tweedie family taken to be 0.4273107)

    Null deviance: 48.67  on 54  degrees of freedom
Residual deviance: 22.67  on 36  degrees of freedom
AIC: NA

Number of Fisher Scoring iterations: 12
```

> #Fit Tweedie Compound Poisson GLM (variance function estimated from the data):

> (fit3 <- glmReserve(tra.tri, var.power = NULL))

	Latest	Dev.To.Date	Ultimate	IBNR	S.E	CV
2005	16704	0.9914530	16848	144	238.1084	1.6535308
2006	23466	0.9751496	24064	598	568.3887	0.9504828
2007	27067	0.9470609	28580	1513	1022.5400	0.6758361
2008	26180	0.9130859	28672	2492	1378.2244	0.5530595
2009	15852	0.8183789	19370	3518	1781.2373	0.5063210
2010	12314	0.7064831	17430	5116	2452.9925	0.4794747
2011	13112	0.5562532	23572	10460	4656.1582	0.4451394
2012	5395	0.3135534	17206	11811	6168.4820	0.5222659
2013	2063	0.1103858	18689	16626	12091.3681	0.7272566
total	142153	0.7311269	194430	52277	16225.3415	0.3103725

```
> summary(fit3, type="model")

Call:
cpglm(formula = value ~ factor(origin) + factor(dev), link =
link.power,
    data = ldaFit, offset = offset)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
```

-9.931 -3.210 0.000 2.039 6.686

	Estimate	Std. Error	t value	Pr(>|t|)	
(Intercept)	7.71827	0.30575	25.244	<2e-16	***
factor(origin)2005	-0.17590	0.31975	-0.550	0.5856	
factor(origin)2006	0.16832	0.31188	0.540	0.5927	
factor(origin)2007	0.35677	0.31151	1.145	0.2596	
factor(origin)2008	0.29287	0.32206	0.909	0.3692	
factor(origin)2009	-0.06975	0.35354	-0.197	0.8447	
factor(origin)2010	-0.19893	0.38466	-0.517	0.6082	
factor(origin)2011	0.12508	0.40178	0.311	0.7574	
factor(origin)2012	-0.14041	0.49986	-0.281	0.7804	
factor(origin)2013	-0.08635	0.72302	-0.119	0.9056	
factor(dev)2	0.70099	0.27584	2.541	0.0155	*
factor(dev)3	0.66228	0.28703	2.307	0.0269	*
factor(dev)4	0.28402	0.31309	0.907	0.3704	
factor(dev)5	0.09397	0.33519	0.280	0.7808	
factor(dev)6	-0.16214	0.36804	-0.441	0.6622	
factor(dev)7	-1.03331	0.46784	-2.209	0.0336	*
factor(dev)8	-1.40153	0.57873	-2.422	0.0206	*
factor(dev)9	-1.90951	0.78974	-2.418	0.0208	*
factor(dev)10	-2.57078	1.25155	-2.054	0.0473	*

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Estimated dispersion parameter: 13.548

Estimated index parameter: 1.5038

Residual deviance: 814.98 on 36 degrees of freedom

AIC: 970.68

Number of Fisher Scoring iterations: 6

####By default, the formulaic approach is used to compute the prediction errors.

####We can also carry out bootstrapping simulations by specifying mse.method = "bootstrap"

####(This argument supports partial match):

> set.seed(11)

> (fit5 <- glmReserve(tra.tri, mse.method = "boot"))

	Latest	Dev.To.Date	Ultimate	IBNR	S.E	CV
2005	16704	0.9908649	16858	154	625.1045	4.0591202
2006	23466	0.9743803	24083	617	1310.2226	2.1235375
2007	27067	0.9464317	28599	1532	2013.8105	1.3144977
2008	26180	0.9061018	28893	2713	2403.9029	0.8860681

2009	15852	0.8138413	19478	3626	2566.5170	0.7078094
2010	12314	0.6946074	17728	5414	3238.1459	0.5981060
2011	13112	0.5465383	23991	10879	5161.1307	0.4744122
2012	5395	0.3366405	16026	10631	6165.8296	0.5799858
2013	2063	0.1122355	18381	16318	12597.1503	0.7719788
total	142153	0.7326115	194036	51883	18002.2582	0.3469780

###components- “sims.par”, “sims.reserve.mean”, and “sims.reserve.pred” store the simulated

parameters, mean values and predicted values of the reserves for each year, respectively

```
> names(fit5)
[1] "call"              "summary"           "Triangle"
[4] "FullTriangle"      "model"             "sims.par"
[7] "sims.reserve.mean" "sims.reserve.pred"
```

####Hence compute the quantiles of the predictions based on the simulated samples

###in the “sims.reserve.pred” element as:

```
> pr <- as.data.frame(fit5$sims.reserve.pred)
> qv <- c(0.025, 0.25, 0.5, 0.75, 0.975)
```

```
> res.q <- t(apply(pr, 2, quantile, qv))
> print(format(round(res.q), big.mark = ","), quote = FALSE)
      2.5%   25%    50%    75%    97.5%
2005     0     0      0    735  2,105
2006     0     0  1,074  2,053  4,549
2007     0 1,473  2,548  3,862  7,665
2008   861 2,787  4,280  6,024  9,987
2009   969 3,202  4,741  6,613 11,041
2010 1,566 4,379  6,232  8,630 14,186
2011 4,143 8,996 12,043 15,860 24,645
2012 2,432 7,366 11,480 15,560 25,469
2013   886 9,388 16,526 25,332 49,172
```

####The full predictive distribution of the simulated reserves for each year can be visualized as:

```
> names(prm) <- c("year", "reserve")
> gg <- ggplot(prm, aes(reserve))
> gg <- gg + geom_density(aes(fill = year), alpha = 0.3) +
facet_wrap(~year, nrow = 2, scales = "free") +
+ theme(legend.position = "none")
> print(gg)
```