DELTA METHOD and RESERVING

By C. Partrat, N. Pey, J. Schilling

Abstract: This paper considers stochastic reserving models within the Generalized Linear Models (in particular the over-dispersed Poisson model). We give a simple procedure, based on the Delta method and products of matrix, to obtain the estimation risk of reserving parameters. We provide with alternative methods, based on m.g.f. and approximated distributions, to estimate the percentiles of reserves amounts.

The fact that such approaches can be used within a quasi-likelihood framework potentially increases the scope of modelling a claims amounts triangle.

Keywords: ***Stochastic reserving, Generalized Linear Models, Delta method, standard errors, predictive distribution, moment generative function, percentiles, approximations, quasi-likelihood.***

Contact addresses:

Partrat Christian: ISFA, Université Claude Bernard Lyon 1,
50, avenue Tony Garnier 69007 LYON
France
☎ : + 33 (0)6.60.50.49.35
Fax: + 33 (0)1.45.41.33.10
✉: partrat@univ-lyon1.fr

Pey Noémie: AXA Canada,
1100 boul René Lévesque, 19ième étage
MONTRÉAL, QUÉBEC, H3B4P4
☎ : (001) 514 392 6000
(001) 514 392 6033 poste 5383
Fax: (001) 514 392 6088
✉: noemie.pey@axa-assurances.ca

Schilling Jérôme : GIE AXA, Group Risk Management
21 Avenue Matignon, 75008 PARIS
☎ : (00 33) 1 40 75 58 66
Fax : (00 33) 1 40 75 58 27
✉ : jerome.schilling@axa.com

1. INTRODUCTION

Stochastic methods consist in applying parametric stochastic models to the claims payments rectangle. Model parameters are estimated with the upper triangle data.

Thus, it is assumed that components (cumulative or incremental) of the payments rectangles are real random variables.

In addition to a “best estimate” of reserve amounts the stochastic approach allows:

(1) To explain the underlying assumptions of the model,
(2) To validate the assumptions with a residuals analysis (at least partially),
(3) To assess the volatility of the predicted amounts of reserve provided by the model,
(4) To build confidence intervals,
(5) To simulate with Monte-Carlo methods the loss reserve for future accident years, as required for a DFA analysis.

A probability distribution of the total amount of reserve R can be also estimated. As a consequence some of its characteristics, which were up to now difficult to calculate analytically, can also be estimated: default risk, Value-at-Risk,…Such statistics are or will be fundamental to determine sufficient statutory claims reserves.

However, uncertainty in estimating parameters is significant with the stochastic approach. This risk consists in using an inadequate model which leads consequently to wrong results.

Two models, whose results are closed to chain ladder estimates, are well adapted to handle the analyses described below. The first one is Mack’s recursive model (1993) and some subsequent developments like the Munich Chain Ladder method.

The second model has been proposed by Kremer in 1982 and extensively used. It considers the factorial LogNormal models or Generalized Linear Models (GLM) and specifically the over dispersed Poisson model (Renshaw et al., 1998). Bootstrapping the Pearson residuals of an over dispersed Poisson model gives an estimate of the reserves distribution (England et al., 2002) and the corresponding percentiles. However it is commonly admitted that bootstrap is less efficient for extreme values and tail distribution.

For this paper the factorial GLM approach has been chosen. First we show the efficiency and the generality of the Delta method (detailed in appendix) for the calculation of the estimation risks. We extend with the building of confidence intervals for parameters linked to the reserves. The calculation of such confidence intervals is straightforward within an Excel spreadsheet thanks to products of matrix.

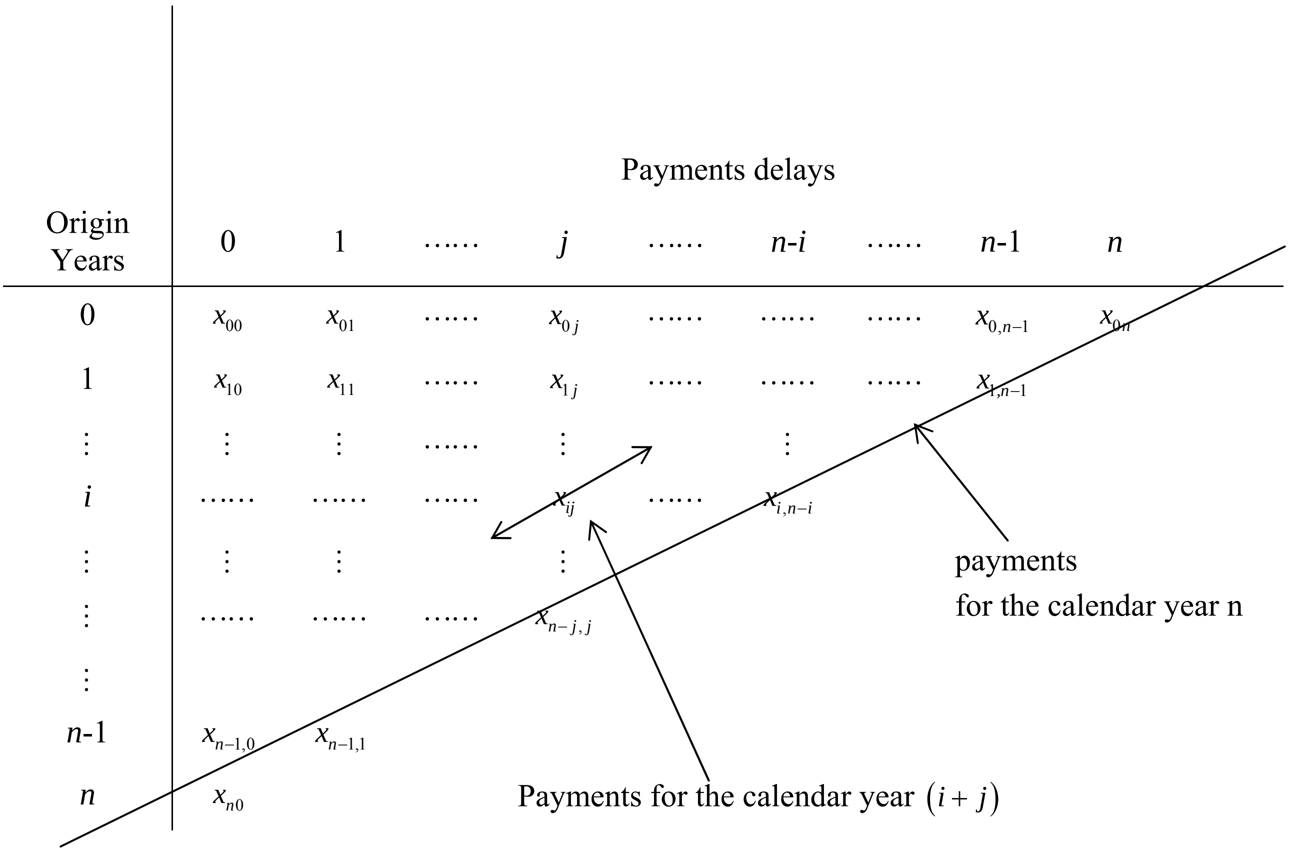
This paper deals also with approaches which can provide us with an estimation of reserve distribution and the corresponding percentiles. They are concerned with moments and moment generating functions (m.g.f.) which are stable under addition of independent random variables. Standard approximation formulas or inversion of the m.g.f. can provide us with such estimates but also with estimation risk and confidence intervals if needed.

2. NOTATIONS, MAIN ISSUES

For a given line of business, claims are assumed to be closed in (n + 1) years. Generalized linear models (GLM) are based on **incremental** payments which are supposed to be random variables $X_{ij},\ i, j = 0,...,n$ under a strong independence assumption:

The random variables $X_{ij},\ i, j = 0,...,n$ **are assumed to be independent.**

In particular, for a given origin year, compensation payments between one delay to another are excluded.
Among the $(n+1)^2$ variables of the payments rectangle, those located in the upper triangle[1] $\left(X_{ij}\right)_{i+j\le n}$ have been observed. Their observations are noted $\left(x_{ij}\right)_{i+j\le n}$.



[1] We count $m = \dfrac{(n+1)(n+2)}{2}$ observations

Most of those models are parametric models as the variables X_{ij} have the same kind of distributions. Their distributions are assumed to belong to the exponential family of GLM (Poisson, Normal, Gamma …) or to their respective transformations (Log Normal, Pareto…). Those distributions depend on real or vector parameters θ_{ij} .

The reserve for the i^{th} origin year is the sum of $X_{i,n-i+1},...,X_{in}$:

$$R_i = \sum_{h=n-i+1}^{n} X_{ih}$$

and the total reserve (random variable) is calculated as :

$$R = \sum_{i=1}^{n} R_i$$

Remark: To analyze future annual cash flows by integrating new business, the cash flows may be calculated as $\sum_{i+j=n+k}^{n} X_{ij}$ for the calendar year $(n+k)$.

2.1 *Interest Parameters*

Most of interest parameters are linked to probability distributions of R_i or R. If F_R is the cumulative distribution function (c.d.f.) of R and $\Pi(F_R)$ a real or vector interest parameter related to F_R, then $\Pi(F_R)$ could be:

- An indicator of location for R : mean $E(R)$, median, percentile…
- An indicator of dispersion: variance $V(R)$, standard deviation $\sigma(R)$…
- An indicator of margin: for instance $E(R)+\gamma\,\sigma(R)$,
- The probability of insufficiency for a given reserve amount R_o : $P(R > R_0) = 1 - F_R(R_0)$
- A tail indicator: the Value-at-Risk for a fixed $\varepsilon > 0$: VaR_ε is defined by: $P(R > VaR_\varepsilon) = \varepsilon$, in other words VaR_ε is the $(1-\varepsilon)^{\text{th}}$ percentile of R.

Alternatively we could use a hyper tail parameter: the Tail VaR defined by $E(R / R \succ VaR_\varepsilon)$.

More ambitiously the whole distribution of R could be obtained by its c.d.f. F_R, or by its m.g.f. $M_R(s) = E(s^R)$ if defined. Due to the assumed independence of the X_{ij} , we calculate the generating function as follow:

$$M_R(s) = \prod_{i=1}^{n} M_{R_i}(s) = \prod_{i=1}^{n} \prod_{i+j \geq n} M_{X_{ij}}(s)$$

2.2 *Estimation*

If $\hat{\Pi} = \hat{\Pi}\left[\left(X_{ij}\right)_{i+j\le n}\right]$ is an estimator of $\Pi\left(F_R\right)$[2], uncertainty related to this estimation is classically measured by the Mean Square Error:

$$\begin{aligned} MSE\left(\hat{\Pi}\right) &= E\left\{\left[\hat{\Pi} - \Pi\left(F_R\right)\right]^2\right\} \\ &= V\left(\hat{\Pi}\right) + \left[E\left(\hat{\Pi}\right) - \Pi\left(F_R\right)\right]^2 \\ &= V\left(\hat{\Pi}\right) \text{ if } \hat{\Pi} \text{ is an unbiased estimator} \end{aligned}$$

or asymptotically defined by $V_{as}(\widehat{\Pi})$.

We then calculate the standard error as the square root of estimation variance to obtain the estimation risk:

$$s.e.\left(\hat{\Pi}\right) = \sqrt{MSE\left(\hat{\Pi}\right)}.$$

or asymptotically defined by $s.e._{as}(\widehat{\Pi}) = \sqrt{V_{as}(\widehat{\Pi})}$.

Those functions are themselves respectively estimated by $\widehat{MSE\left(\hat{\Pi}\right)}$ and $\widehat{s.e.\left(\hat{\Pi}\right)}$.

In addition, a level (for instance) 95% (respectively asymptotic) confidence interval for $\Pi\left(F_R\right)$, is defined by its upper and lower bounds $A\left[\left(X_{ij}\right)_{i+j\le n}\right]$ and $B\left[\left(X_{ij}\right)_{i+j\le n}\right]$ such as

$$P\left\{A\left[\left(X_{ij}\right)_{i+j\le n}\right] \le \Pi\left(F_R\right) \le B\left[\left(X_{ij}\right)_{i+j\le n}\right]\right\} = 0,95\,(resp. \to 0,95).$$

3. GLM FACTORIAL MODELS

Generalized Linear Models have been introduced by J. Nelder and R.Wedderburn in 1972. Many papers set out in details the underlying statistical theory[3]. One of the best reference is Mc Cullag and Nelder, 1989.

---

[2] An indicator of location is frequently chosen: mean or median of *R*. Then the resulting reserves estimates should be closed or strictly the same as those obtained with the chain ladder technique

3 Models can be implemented with specific procedures from standard statistical packages (for instance Proc GENMOD or REG for SAS)

3.1 Random Component

We consider independent «response» $X_{ij}\left(i,j=0,\dots,n\right)$ with probability distributions belonging to the exponential family. The density[4] function of $X_{ij}\left(i,j=0,\dots,n\right)$ is defined by

$$f\left(x_{ij};\theta_{ij},\phi\right)=\exp\left\{\left[\theta_{ij}x_{ij}-b\left(\theta_{ij}\right)\right]/\phi+c\left(x_{ij},\phi\right)\right\}$$

where

- θ_{ij} is a real parameter, called natural parameter,
- $\phi>0$ is a dispersion parameter (eventually known) independent of *i* and *j*,
- *b* and *c* are functions specifically defined by the distributions, *b* being "regular".

It can be shown that :

$$\mu_{ij}=E\left(X_{ij}\right)=b'\left(\theta_{ij}\right) \text{ or, if } b' \text{ is invertible, } \theta_{ij}=b'^{-1}(\mu_{ij})=\theta(\mu_{ij})$$

$$V\left(X_{ij}\right)=\phi b'\left(\theta_{ij}\right)=\phi b''\left[b'^{-1}\left(\mu_{ij}\right)\right]=\phi V\left(\mu_{ij}\right)$$

The function *V* is called the variance function of the distribution. Moreover, third order central moment and skewness can be written as follow:

$$\mu_3(X_{ij})=\phi^2 b'''(\theta_{ij})=\phi^2 b'''\left[\theta(\mu_{ij})\right] \text{ and } \mu_3(X_{ij})=\phi^2 W(\mu_{ij}) \text{ with } W(\mu)=V(\mu)V'(\mu)$$

$$\gamma_1(X_{ij})=\frac{\mu_3(X_{ij})}{\left[V(X_{ij})\right]^{3/2}}=\sqrt{\phi}\,\frac{b'''(\theta_{ij})}{\left[b''(\theta_{ij})\right]^{3/2}}=\sqrt{\phi}\,\frac{W(\mu_{ij})}{\left[V(\mu_{ij})\right]^{3/2}}$$

Finally, the expressions of m.g.f. and cumulant generating function are:

$$M_{X_{ij}}(s)=\exp\left\{\frac{1}{\phi}\left[b(\theta_{ij}+s\phi)-b(\theta_{ij})\right]\right\}$$

$$C_{X_{ij}}(s)=\log M_{X_{ij}}(s)=\frac{1}{\phi}\left[b(\theta_{ij}+s\phi)-b(\theta_{ij})\right]$$

By successive derivations it would be possible to deduce the moments of X_{ij} firstly depending on (θ_{ij},ϕ) and secondly on (μ_{ij},ϕ).

---

[4] True probability density function in the continuous case, simple probability in the discrete case.

Examples: **(1) Discrete case**

- Bernoulli distribution $B(1,\pi)$:

$$P(X=x)=e^{x\ln\frac{\pi}{1-\pi}+\ln(1-\pi)}.\ x=0,1:\theta=\ln\frac{\pi}{1-\pi},\phi=1,b(\theta)=\ln(1+e^{\theta})$$

$$\mu=E(Y)=\pi\ ,\ V(\mu)=\mu(1-\mu)\ ,W(\mu)=\mu(1-\mu)(1-2\mu).$$

- Poisson distribution $P(\lambda)$:

$$P(X=x)=e^{(x\ln\lambda-\lambda)+c(x)}.\ x\in\mathbb{N}:\theta=\ln\lambda,\ \phi=1,\ b(\theta)=e^{\theta}$$

$$\mu=E(X)=\lambda,\ V(\mu)=\mu,\ W(\mu)=\mu$$

- Over-dispersed Poisson distribution $P_{surd}(\lambda,\phi)$:

Formally it would be written $X\approx P_{surd}(\lambda,\phi)\ if\ \dfrac{X}{\phi}\approx P(\dfrac{\lambda}{\phi})$

In section 4.4 the reader will find additional rigorous presentation of this specific case.

(2) **Continuous case**[5]

- Gamma distribution $\gamma\left(\upsilon,\dfrac{\upsilon}{\mu}\right)$:

$$f(x)=\exp\left[\left(-\frac{x}{\mu}-\ln\mu\right)\upsilon+c(x,\upsilon)\right].\ x>0\ :\theta=-\frac{1}{\mu},\ \phi=\frac{1}{\upsilon},\ E(X)=\mu,\ V(\mu)=\mu^2$$

- Inverse Gaussian distribution $IG(\mu,\sigma^2)$:

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma x^{\frac{3}{2}}}\exp\left[-\frac{(x-\mu)^2}{2\mu\sigma^2 x}\right]=\exp\left[\left(-\frac{x}{2\mu^2}+\frac{1}{\mu}\right)\frac{1}{\sigma^2}+c(x,\sigma^2)\right].\ x>0$$

$$\theta=-\frac{1}{2\mu^2},\ \phi=\sigma^2,\ E(X)=\mu,\ V(\mu)=\mu^3$$

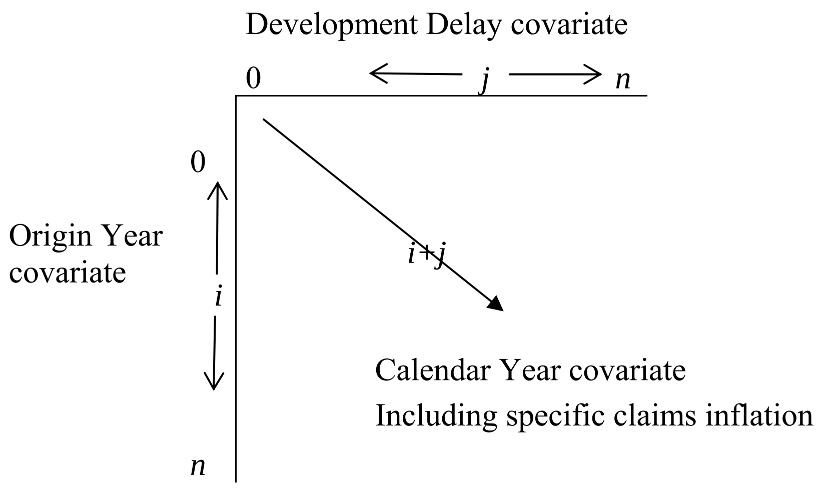
- Tweedie's compound Poisson distribution (see Wüthrich, 2003)

[5]For a Normal distribution $\mathrm{N}(\mu,\sigma^2):\theta=\mu,\ \phi=\sigma^2,\ E(X)=\mu,\ V(\mu)=1$

Remark: A similar approach could be envisaged for transforms of GLM random components. It would generate Log Normal, Pareto, Log Gamma …

3.2 Covariates

The covariates used by the different models represent the three natural directions of the payments triangle:



The "Year" covariates (origin or calendar) are (qualitative) ordinal covariates. The levels 0, 1 …n which are used within this paper are a pure arbitrary codification. However procedures which are used to implement covariates methods do not allow integrating such covariates. That's why in the following part, the Origin Year covariates will be considered as pure qualitative covariates, or in other words as factors.

If the origin year is assumed to be a factor, it will count (n+1) levels: 0, 1 …, n. And it will be replaced by n auxiliary covariates in 0, 1 associated to parameters $\alpha_1, \alpha_2, ..., \alpha_n$. The level 0 will be used as a reference ($\alpha_0 = 0$).
In the same way, the calendar year would be associated to parameters μ_{i+j} ($i, j = 0, ..., n$). Here the annual inflation is assumed to be constant for the 2n considered years: $\mu_{i+j} = \mu$. If this assumption is invalid, amounts have to be deflated.

The delay covariate would be differently modeled. By definition this covariate is discrete quantitative covariate with values equal to 0, 1… A way to model it is to use a linear combination of functions of j : j, j^r, $\log(1+j)$ … (see England et al.. 2002). However in this paper this covariate will be considered as a (n+1) levels factor associated to parameters $\beta_1, \beta_2, ..., \beta_n$ ($\beta_0 = 0$).

To sum up, a modeling using categorical covariates related to parameters $(\alpha_i)_{i=0,\ldots,n}, (\beta_j)_{j=0,\ldots,n}$ and to a constant parameter μ will be used in this paper[6]. To make it easier, no interactions between origin year and delay have been introduced.

As opposed to Mack's recursive model of which assumptions are only concerned with the two first (conditional) moments of X_{ij}, factor models require an a priori link between distribution of variables X_{ij} and parameters $(\alpha_i),(\beta_j),\mu$. A function which links $\mu_{ij} = E(X_{ij})$ to the systematic components of the GLM is extensively used.

3.3 Systematic Component, link function

Within the context of reserving, the systematic component is:

$$\eta_{ij} = \mu + \alpha_i + \beta_j \quad (i, j = 0,\ldots,n) \text{ with } \alpha_0 = \beta_0 = 0 \text{ (to identify the model).}$$

The link function allows to link together the random and systematic component. It is a strictly monotone and derivable real function g such as:

$$\eta_{ij} = g(\mu_{ij}) \text{ or } \mu_{ij} = g^{-1}(\eta_{ij}).$$

Standard links[7] are:

- Identity link: $\eta_{ij} = \mu_{ij}$ or $\mu_{ij} = \mu + \alpha_i + \beta_j$ (additive model)
- Log link: $\eta_{ij} = \log \mu_{ij}$ or $\mu_{ij} = e^{\eta_{ij}}$ or $\mu_{ij} = e^{\mu+\alpha_i+\beta_j}$ (multiplicative model).

The parameters θ_{ij} which appear within the density function have only a temporary interest. Finally a GLM can be characterized by:

- A probability distribution for the response variable,
- A variance function *V* and a dispersion parameter ϕ with:
  $$E(X_{ij}) = \mu_{ij}.\ V(X_{ij}) = \phi V(\mu_{ij}), \mu_3(X_{ij}) = \phi^2 V(\mu_{ij}) V'(\mu_{ij}).$$
- A link function $\eta_{ij} = g(\mu_{ij})$

[6] With *p*=2*n*+1 parameters

[7] A specific "canonical " link exists for each distribution (Mc Cullagh et al., 1989).

4. INFERENCE

4.1 *Estimation*

As the X_{ij} have been assumed independent the likelihood function related to variables of the upper triangle $L\left[\left(x_{ij}\right)_{i+j\le n};\mu,\left(\alpha_i\right),\left(\beta_j\right),\phi\right]$ is straightforward.

Putting the partial derivatives of *log L* with respect to parameters $\mu,(\alpha_i),(\beta_j)$ to 0 leads to the Wedderburn system of which equations are independent of the parameter ϕ :

$$(S)\left\|\sum_{\substack{i,j=1\\ i+j\le n}}^{n}\frac{\left(x_{ij}-\mu_{ij}\right)}{V\left(\mu_{ij}\right)}\frac{\delta\mu_{ij}}{\delta\eta_{ij}}b_{ij}^{(k)}=0\quad .\qquad k=1,\ldots,p\right.$$

where $b_{ij}^{(k)}$ is not depending on parameters $\mu,(\alpha_i),(\beta_j)$.

Remark: putting $\frac{\delta\log L}{\delta\phi}=0$ in addition within the system *(S)* would give an estimator of ϕ.

The system (S) can only be numerically solved with the standard Newton-Raphson or score algorithms. It allows to obtain the maximum likelihood estimate (m.l.e.) $\hat{\xi}=\left[\hat{\mu},\left(\hat{\alpha}_i\right),\left(\hat{\beta}_j\right)\right]$ of $\xi=\left[\mu,\left(\alpha_i\right),\left(\beta_j\right)\right]$.

Second-order partial derivatives of *log L* with respect to parameters allow obtaining the Fisher information matrix $I\left(\xi\right)$.

Its inverse $I^{-1}\left(\xi\right)$ is the asymptotic variance-covariance matrix[8], denoted $\Sigma_{as}(\hat{\xi})$, of the estimator $\hat{\xi}$. Its diagonal elements are the variances $\sigma_{as}^2(\widehat{\mu})$, $\sigma_{as}^2(\widehat{\alpha_i})$, $\sigma_{as}^2(\widehat{\beta_j})$ and its other elements are the covariances:

$$\mathrm{cov}_{as}(\widehat{\mu},\widehat{\alpha_i}),\ \mathrm{cov}_{as}(\widehat{\mu},\widehat{\beta_j}),\ \mathrm{cov}_{as}(\widehat{\alpha_i},\widehat{\beta_j})\ .$$

If the model contains a dispersion parameter ϕ, we have $\Sigma_{as}(\hat{\xi})=\phi\Sigma_{as}$ where Σ_{as} is a matrix which is independent of ϕ.

Under standard conditions concerning the maximum likelihood (Shao. 1999), $\hat{\xi}$ is asymptotically Normal AN [$\xi,\Sigma_{as}(\hat{\xi})$].

---

[8] Estimated asymptotic variances and covariances are extensively reported as GLM output of statistical packages.

The functional invariance property of the maximum likelihood gives: $\hat{\eta}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$ and $\hat{\mu}_{ij} = g^{-1}\left(\hat{\eta}_{ij}\right)$ which are respectively the m.l.e. of η_{ij} and $\mu_{ij} = E\left(X_{ij}\right)$.

$\hat{\mu}_{ij}$ is the predicted value by the model.

Remark : If g is a derivable function with $g'(\mu) \neq 0$, the Delta Method (see appendix) leads to the asymptotic variance of $\hat{\mu}_{ij}$:

$$\sigma_{as}^2\left(\hat{\mu}_{ij}\right) = \left[\left(g^{-1}\right)'\left(\eta_{ij}\right)\right]^2 \sigma_{as}^2\left(\hat{\eta}_{ij}\right).$$

Specifically for the reserves:

$$R_i = \sum_{j>n-i} X_{ij} \text{ gives } E\left(R_i\right) = \sum_{j>n-i} E\left(X_{ij}\right) = \sum_{j\succ n-i} \mu_{ij}$$

$$\widehat{E(R_i)} = \sum_{j>n-i} \hat{\mu}_{ij} \text{ is the m.l.e. of } E\left(R_i\right).$$

Consequently $R = \sum_{i=1}^{n} R_i = \sum_{i+j\succ n} X_{ij}$ gives $E\left(R\right) = \sum_{i+j\succ n} E\left(X_{ij}\right) = \sum_{i+j\succ n} \mu_{ij}$

and $\widehat{E\left(R\right)} = \sum_{i+j\succ n} \hat{\mu}_{ij}$ is the m.l.e.. of $E\left(R\right)$.

4.2 *Estimation risk*

Let us consider the following transformation:

$$\eta : \mathbb{R}^{2n+1} \to \mathbb{R}^{(n+1)^2} : \xi \sim\!\!\to \eta(\xi) = (\eta_{ij})$$

The Jacobian matrix J_η related to this transformation is defined for its line (i, j) by:

$$\frac{\partial \eta_{ij}}{\partial \mu} = 1,\ \frac{\partial \eta_{ij}}{\partial \alpha_k} = \begin{cases} 1 & \text{if } k = i \\ 0 & k \neq i \end{cases},\ \frac{\partial \eta_{ij}}{\partial \beta_l} = \begin{cases} 1 & \text{if } l = j \\ 0 & l \neq j \end{cases}$$

Thanks to the Delta method, $\hat{\eta} = (\widehat{\eta_{ij}})$ is asymptotically Normal AN [η, $\Sigma_{as}(\hat{\eta})$] with:

$$\Sigma_{as}(\hat{\eta}) = J_\eta \Sigma_{as}(\hat{\xi}) J_\eta'$$

for the transposed matrix J_η' of J_η.

Assuming that the link function g is strictly monotone and derivable, we will call D the Jacobian matrix related to the transformation $(\eta_{ij}) \sim\!\!\rightarrow (\mu_{ij})$:

$$\mathbb{R}^{(n+1)^2} \rightarrow \mathbb{R}^{(n+1)^2} : (\eta_{ij}) \sim\!\!\rightarrow \left[g^{-1}(\eta_{ij}) \right]$$

This matrix is diagonal with the following standard components: $(g^{-1})'(\eta_{ij})$.

Thanks to the Delta method, the estimation $\widehat{\mu} = (\widehat{\mu_{ij}})$ is asymptotically Normally distributed AN[$\mu, \Sigma_{as}(\widehat{\mu})$] with:

$$\Sigma_{as}(\widehat{\mu}) = D\Sigma_{as}(\widehat{\eta})D$$

As $E\left(R_i\right) = \sum_{j \succ n-i} \mu_{ij} \; (i = 1,...,n)$, the Jacobian matrix J_μ related to the transformation

$$\mathbb{R}^{(n+1)^2} \rightarrow \mathbb{R}^n : (\mu_{ij}) \sim\!\!\rightarrow \left[E(R_i) \right]$$

has an i[th] line defined as:

$$\frac{\partial E(R_i)}{\partial \mu_{kl}} = \begin{cases} 0 & if \quad\quad k \neq i \\ 0 & if \quad k = i, l \leq n-i \\ 1 & if \quad k = i, l \succ n-i \end{cases}$$

The random vector $\left[\widehat{E(R_i)} \right]_{i=1,...,n}$ is asymptotically Normally distributed

$$\text{AN} \left\{ \left[E(R_i) \right]_{i=1,...,n}, \Sigma_{as} \left\{ \left[\widehat{E(R_i)} \right] \right\} \right\} \text{ with } \Sigma_{as} \left\{ \left[\widehat{E(R_i)} \right] \right\} = J_\mu \Sigma_{as}(\widehat{\mu}) J_\mu'$$

Finally $J_R = (1,1,...,1)$ is the Jacobian matrix of the transformation:

$$\mathbb{R}^n \rightarrow \mathbb{R} : \left[E(R_i) \right] \sim\!\!\rightarrow E(R)$$

As a result $\widehat{E(R)}$ is asymptotically normally distributed $AN\left\{ E(R), \sigma_{as}^2 \left[\widehat{E(R)} \right] \right\}$ with:

$$\sigma_{as}^2 \left[\widehat{E(R)} \right] = J_R \Sigma_{as} \left\{ \left[\widehat{E(R_i)} \right] \right\} J_R'$$

or using elements of the matrix $\Sigma_{as} \left\{ \left[\widehat{E(R_i)} \right] \right\}$ (see above):

$$\sigma_{as}^2 \left[\widehat{E(R)} \right] = \sum_i \sigma_{as}^2 \left[\widehat{E(R_i)} \right] + 2 \sum_{i \neq j} \text{cov}_{as} \left[\widehat{E(R_i)}, \widehat{E(R_j)} \right]$$

Those results allow measuring estimation risk thanks to the estimated asymptotic relative standard error:

$$\frac{\widehat{\sigma_{as}}\left[\widehat{E(R)}\right]}{\widehat{E(R)}}$$

From the asymptotic Normality of $\widehat{E(R)}$ a asymptotic level (1-η) confidence interval of $E(R)$ (with η such as 0< η <1) can be deduced:

$$\left[\widehat{E(R)} - q_{1-\eta}\widehat{\sigma_{as}}\left[\widehat{E(R)}\right], \widehat{E(R)} + q_{1-\eta}\widehat{\sigma_{as}}\left[\widehat{E(R)}\right]\right]$$

where $q_{1-\eta}$ is the $(1\text{-}\eta)^{th}$ percentile of the standard Normal distribution.

If the variance function *V* is strictly monotone and derivable, the same approach could apply to the variance of *R* because:

$$V(R) = \sum_{i=1}^{n}\sum_{j \succ n-i} V(X_{ij}) = \phi \sum_{i=1}^{n}\sum_{j \succ n-i} V(\mu_{ij})$$

And more generally, it could be applied to the cumulants and then to the moments of this variable *R*.

Estimation risk and confidence intervals for the margin $E(R) + \gamma\sigma(R)$ could be obtained by using the Delta method as well.

4.3 *Deviance*

For $i + j \leq n$, we consider for the cell (i, j), the following residuals[9]:

(i) Raw residual: $r_{ij} = x_{ij} - \hat{\mu}_{ij}$

(ii) Deviance residual: $r_{ij}^{(D)} = \text{sgn}\left(x_{ij} - \hat{\mu}_{ij}\right)\sqrt{d_{ij}}$ where:

$$d_{ij} = 2\left\{x_{ij}\left(\tilde{\theta}_{ij} - \hat{\theta}_{ij}\right) - \left[b\left(\tilde{\theta}_{ij}\right) - b\left(\hat{\theta}_{ij}\right)\right]\right\}$$

is the $(i, j)^{\text{ith}}$ deviance term by using the notations introduced in section § 3.1:

$$\tilde{\theta}_{ij} = b'^{-1}\left(x_{ij}\right) \text{ and } \hat{\theta}_{ij} = b'^{-1}\left(\hat{\mu}_{ij}\right).$$

The aim of residuals analyses is to detect outlier cells and deviation from the hypotheses (Fahrmeir et al, 2001).

By considering these residuals globally, they lead to the deviance[10], which is a goodness-of-fit indicator of the model:

---

[9] Alternatively we could consider the Pearson residual : : $r_{ij}^{(p)} = \dfrac{x_{ij} - \hat{\mu}_{ij}}{\sqrt{V\left(\hat{\mu}_{ij}\right)}}$

$$D = 2\sum_{i+j\le n} d_{ij}$$

The statistic $D^* = \dfrac{D}{\phi}$ is the standardized deviance.

Let emphasize that minimizing D or D^* will give an optimal combination of covariates but only within the same random component (same V and ϕ).

Particular cases:

- Poisson $P(\lambda)$: $D = 2\sum_{i+j\le n}\left[x_{ij}\ln\dfrac{x_{ij}}{\hat{\mu}_{ij}} - \left(x_{ij}-\hat{\mu}_{ij}\right)\right]$
- Normal $N(\mu,\sigma^2)$: $D = X^2 = \sum_{i+j\le n}\left(x_{ij}-\hat{\mu}_{ij}\right)^2$ or the sum of squared residuals of the standard case.
- Gamma $\gamma\left(\upsilon,\dfrac{\upsilon}{\mu}\right)$: $D = 2\sum_{i+j\le n}\left(\dfrac{x_{ij}-\hat{\mu}_{ij}}{\hat{\mu}_{ij}} - \ln\dfrac{x_{ij}}{\hat{\mu}_{ij}}\right)$
- Inverse-Gaussian $IG(\mu,\sigma^2)$: $D = 2\sum_{i+j\le n}\dfrac{(x_{ij}-\widehat{\mu_{ij}})^2}{x_{ij}\,\widehat{\mu_{ij}}^2}$

4.4 *Quasi-likelihood, selection of models*

If the variance function $V(\mu) \succ 0$ and possibly the dispersion parameter ϕ are specified without any reference to an underlying distribution of the exponential family, Wedderburn's equations and asymptotic properties of the m.l.e. $\hat{\xi}$ remain valid (Mc Cullagh et al.. 1989). It leads to quasi-likelihood models that are used for ratemaking (Renshaw, 1994).

The quasi-likelihood is defined by $q(\underline{x},\underline{\mu}) = \sum_{i+j\le n}\int_{x_{ij}}^{\mu_{ij}}\dfrac{x_{ij}-u}{\phi V(u)}du$

The system of partial derivative functions gives the Wedderburn's equations.

Remark : Under certain conditions, such approach can be used with negative increments within the payments triangle.

By analogy the deviance defined before can be extended to the quasi-deviance *D*[11]

---

[10] Alternatively we could obtain the generalized Pearson statistic χ^2 : $X^2 = \sum_{i+j\le n} r_i^{(P)^2}$

[11] Alternatively we could use the quasi-likelihood coming from Pearson's chi-squared statistic

$$D = \sum_{i+j \le n} d_{ij} = -2\phi q(\underline{x}, \underline{\widehat{\mu}}) \qquad \text{with} \qquad d_{ij} = 2\int_{\widehat{\mu_{ij}}}^{x_{ij}} \frac{x_{ij} - u}{V(u)} du,$$

and can be extended also to the standardized deviance $D^* = \dfrac{D}{\phi}$.

If ϕ is a parameter, it is replaced everywhere by its deviance estimate $\hat{\phi} = \dfrac{1}{m-p} D$.

Particular case : Over-dispersed Poisson distribution
we obtain $\mu = E(X) = \lambda, V(\mu) = \mu, W(\mu) = \mu,\ \phi$
and $V(X) = \phi\mu, \mu_3(X) = \phi^2\mu$ (see section § 3.1)

To compare different models, Nelder et al. (1987), have introduced the extended quasi-likelihood q^+, defined by:

$$-2q^+(\underline{x}, \underline{\widehat{\mu}}) = \frac{1}{\phi} \sum_{i+j \le n} d_{ij} + \sum_{i+j \le n} \log\left[2\pi\phi V(x_{ij})\right]$$

Maximizing the statistic q^+ allows comparing models with different structures (V, ϕ, g). Minimizing $-2q^+$ is then equivalent.

Remark: In this approach, the family of power variance functions would be the main case :
$V(\mu) = \mu^\tau$, where τ is a real parameter, and $V(\mu) \xrightarrow[\tau \to 0]{} \log \mu$.
It would be the same with additive and multiplicative links.

5. ESTIMATION OF THE PREDICTIVE DISTRIBUTION

Knowing the moments of *R* will not be sufficient to obtain additional characteristics of *R* (see section § 2.1) : probability of insufficiency, percentiles (or VaR) …That's why it is necessary to get the distribution of *R*.

The distribution of *R* can be obtained using the additive properties of the m.g.f. of the X_{ij} in some very specific cases : Poisson, over dispersed Poisson and Normal distributions.

If the variables X_{ij} are over-dispersed Poisson distributed $P_{surd}(\mu_{ij}, \phi)$, then *R* is formally over-dispersed Poisson distributed and its distribution is :

$$P_{surd}(\mu_R, \phi) \text{ with } \mu_R = \sum_{i=1}^{n} \sum_{i+j \succ n} \mu_{ij}.$$

Commonly if $\mu_R \ge 50$, the c.d.f. of the Poisson distribution is approximated by a standard Normal distribution with correction for continuity :

$$P(R \le r) = \Phi\left(\frac{r + 1/2 - \mu_R}{\sqrt{\phi\mu_R}}\right) \quad \text{and percentile} \quad q_{1-\eta}(R) = \mu_R - \frac{1}{2} + \sqrt{\phi\mu_R}\, q_{1-\eta}$$

From the m.l.e. $(\widehat{\mu_{ij}})$ of (μ_{ij}), we deduce $\widehat{\mu_R}$ and $\widehat{q_{1-\eta}(R)}$ for μ_R and $q_{1-\eta}(R)$:

$$\widehat{q_{1-\eta}(R)} = \widehat{\mu_R} - \tfrac{1}{2} + \sqrt{\hat{\phi}\widehat{\mu_R}}\, q_{1-\eta}$$

We would obtain the estimation risk and confidence intervals.

Usually we cannot obtain directly such predictive distribution. Therefore we suggest some appropriate approaches described in the next section.

5.1 *Inversion of the moment generating function*

As the independence of the variables X_{ij} has been assumed, the m.g.f. of *R* can easily be obtained by:

$$M_R(s) = \prod_{i=1}^{n} M_{R_i}(s) = \prod_{i=1}^{n} \prod_{i+j \ge n} M_{X_{ij}}(s)$$

$$M_R(s) = \exp\left\{\frac{1}{\phi}\sum_{i=1}^{n}\sum_{j=n-i+1}^{n}\left\{b\left[\theta(g^{-1}(\eta_{ij})) + s\phi\right] - b\left[\theta(g^{-1}(\eta_{ij}))\right]\right\}\right\}$$

If this generating function is not standard, this one must be inverted to get the c.d.f. of *R*. The Fast Fourier Transform may allow resolving numerically the problem (Klugman et al.,1998)

5.2 *Approximated distributions*

Approximated distributions based on the first moments of *R* use classical methods in actuarial sciences: Normal Power, Esscher, translated Gamma, Bower's Gamma approximation..... (Gerber, 1979, Partrat et al.. 2005)
Such methods enable to approximate quite precisely the c.d.f. F_R and the insufficiency probability. Half of them provide also with a simple analytical approximation of percentiles. These quantities are then estimated by maximum likelihood. The Delta method could give the estimation risks and asymptotic confidence intervals.
With $R = \sum_{i=1}^{n}\sum_{j=n-i+1}^{n} X_{ij}$ and independence of X_{ij}, the moments of *R* can be expressed as a function of (μ_{ij}) (see section § 3.1) :

$$\mu = E(R) = \sum_{i=1}^{n}\sum_{j=n-i+1}^{n} \mu_{ij}$$

$$\sigma^2 = V(R) = \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} V(X_{ij}) = \phi \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} V(\mu_{ij})$$

$$\mu_3 = \mu_3(R) = \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} \mu_3(X_{ij}) = \phi^2 \sum_{i=1}^{n} \sum_{j=n-i+1}^{n} W(\mu_{ij})$$

$$\gamma_1 = \gamma_1(R) = \frac{\mu_3}{\sigma^3} = \sqrt{\phi}\, \frac{\sum_i \sum_j W(\mu_{ij})}{\left[\sum_i \sum_j V(\mu_{ij})\right]^{3/2}}$$

The invariance property of the maximum likelihood would give a m.l.e. $(\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma_1})$ for (μ, σ, γ_1) by replacing μ_{ij} by $\widehat{\mu_{ij}}$ in the above formulas.

Under the condition of second order derivability of the function *V*, we would deduce the asymptotic Normality of the estimator $(\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma_1})$ and its variance-covariance matrix using $\Sigma_{as}(\widehat{\mu})$.

5.2.1 Normal Power Approximation

We obtain directly $\quad F_R(x) \approx F^{(NP)}(x) = \Phi\left[\frac{-3}{\gamma_1} + \sqrt{\frac{9}{\gamma_1^2} + 1 + \frac{6}{\gamma_1}\left(\frac{x-\mu}{\sigma}\right)}\right]$

From the equation $F^{(NP)}(x) = 1 - \eta$ or $\frac{-3}{\gamma_1} + \sqrt{\frac{9}{\gamma_1^2} + 1 + \frac{6}{\gamma_1}\left(\frac{x-\mu}{\sigma}\right)} = q_{1-\eta}$

we deduce the approximated value of the $(1-\eta)^{\text{th}}$ percentile $q_{1-\eta}^{(NP)} = \mu + \sigma\left[\frac{\gamma_1}{6} q_{1-\eta}^2 + q_{1-\eta} - 1\right]$

estimated (maximum likelihood) by $\widehat{q_{1-\eta}^{(NP)}} = \widehat{\mu} + \widehat{\sigma}\left[\frac{\widehat{\gamma_1}}{6} q_{1-\eta}^2 + q_{1-\eta} - 1\right]$.

Finally we would obtain the estimated standard error of $\widehat{q_{1-\eta}^{(NP)}}$ and asymptotic confidence interval for $q_{1-\eta}^{(NP)}$.

Remark: Esscher approximation (Gerber,1979, Denuit et al.. 2004)

The Esscher transform of F_R with parameter h is the c.d.f. F_h defined by:

$$F_h(x) = \frac{1}{M(h)} \int_0^x e^{hy} dF_R(y)$$

with m.g.f..

$$M_h(s) = \frac{M(s+h)}{M(h)}$$

Expectation of F_h is $E(h) = \frac{M'(h)}{M(h)}$

For $x \succ \mu$ the Esscher approximation of $F_R(x)$ is obtained by applying the third order Edgeworth approximation to $F_h(x)$ for h such as $E(h) = x$. Hence for the tail function :

$$\overline{F_R}(x) \approx \overline{F}^{(ES)}(x) = M(h)e^{-hx}\left[E_0(h\sigma_h) - \frac{\gamma_1(h)}{6}E_3(h\sigma_h)\right]$$

$\sigma_h^2, \gamma_1(h)$ being respectively the variance and skewness of F_h and the Esscher functions:

$$E_0(u) = e^{\frac{u^2}{2}}\left[1-\Phi(u)\right]$$

$$E_3(u) = \frac{1-u^2}{\sqrt{2\pi}} + u^3 E_0(u)$$

Contrary to the NP approximation, the estimation of percentiles requires a numerical solution.

Particular case : over-dispersed Poisson distribution

As $\frac{R}{\phi}$ follows a Poisson distribution $P(\frac{\mu}{\phi})$, we have formally :

$$M(s) = \exp\left[\frac{\mu}{\phi}\left(e^{\phi s} - 1\right)\right]$$

$$M_h(s) = \frac{M(s+h)}{M(h)} = \exp\left[\frac{\mu}{\phi}e^{\phi h}(e^{\phi s} - 1)\right]$$

which is the m.g.f. of an over-dispersed Poisson model $P_{surd}(\mu e^{\phi h}, \phi)$

This means that $h = \frac{1}{\phi}\ln(\frac{x}{\mu})$, $\sigma_h^2 = \phi\mu e^{\phi h}$, $\gamma_1(h) = \sqrt{\frac{\phi}{\mu}} e^{\frac{1}{2}\phi h}$ and we deduce the Esscher approximation of $F_h(x)$.

5.2.2 Gamma approximation

A. Translated Gamma

The distribution of *R* is approximated by a translated Gamma distribution $\gamma(\upsilon, \beta, x_0)$. The corresponding parameters are obtained from identification of the three first moments:

$$E(R) = x_0 + \frac{\nu}{\beta} \ , \ V(R) = \frac{\nu}{\beta^2} \ , \ \mu_3(R) = \frac{2\nu}{\beta^2}$$

so that
$$\nu = \frac{4}{\gamma_1^2} \ , \ \beta = \frac{2}{\sigma\gamma_1} \ , \ x_0 = \mu - \frac{2\sigma}{\gamma_1}$$

and
$$F_R(x) \approx F_R^{(GT)}(x) = \Gamma\left[\nu, \beta(x - x_0)\right] = F_{\chi^2_{2\nu}}\left[2\beta(x - x_0)\right]$$

where $\Gamma(\nu, y)$, $F_{\chi^2_{2\nu}}(y)$ are respectively the incomplete Γ function and the c.d.f. of a Chi-squared distribution with 2ν degrees of freedom. This distribution is available in most statistical packages.

From the equation defining the percentiles $F_R^{(GT)}(x) = 1 - \eta$ we deduce :

$$q_{1-\eta}^{(GT)}(R) = x_0 + \frac{1}{2\beta}\chi^2_{2\nu}(1-\eta)$$

where $\chi^2_{2\nu}(1-\eta)$ is the $(1-\eta)^{th}$ percentile of the Chi-squared distribution with 2ν degrees of freedom. Such value is easily available.

Estimations (maximum likelihood), estimation risk and confidence intervals of parameters could be obtained as described in section 5.2.1.

B. Bowers' Gamma

This method is based on Laguerre orthogonal polynomials and the distribution $\gamma(\nu, \beta)$ where the two first moments coincide with those of *R*.

This means that
$$\nu = \frac{\mu^2}{\sigma^2} \ , \ \beta = \frac{\mu}{\sigma^2}$$

We deduce
$$F_R(x) \approx F^{(GB)}(x) = G(\beta x)$$

with
$$G(y) = \Gamma(\nu, y) - \frac{1}{6}(\beta^2\mu_3 - 2\nu)e^{-y}y^{\nu}\left[\frac{y^2}{\Gamma(\nu+3)} - \frac{2y}{\Gamma(\nu+2)} + \frac{1}{\Gamma(\nu+1)}\right]$$

As for the Esscher approximation, the approximation of percentiles requires a numerical solution.

.

6. NUMERICAL EXAMPLE ON MARINE BUSINESS

The table 1 gives the incremental claims amounts for some line of Marine business with *underwriting* years from 1984 to 1991.

Table 1: triangle of incremental payments

Years	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
0	1 381	4 399	4 229	435	465	205	110	67
1	859	6 940	2 619	1 531	517	572	287	
2	6 482	6 463	3 995	1 420	547	723		
3	2 899	16 428	5 521	2 424	477			
4	3 964	15 872	8 178	3 214				
5	6 809	24 484	27 928					
6	11 155	38 229						
7	10 641							

By applying the standard chain ladder method we obtain the following future incremental predicted values. In addition the last column provides with the reserve amounts for each underwriting year and the total reserves amount (table 2) :

Table 2 : Chain ladder reserves

1 381	4 399	4 229	435	465	205	110	67	0
859	6 940	2 619	1 531	517	572	287	***80***	80
6 482	6 463	3 995	1 420	547	723	***323***	***119***	442
2 899	16 428	5 521	2 424	477	***984***	***472***	***174***	1 631
3 964	15 872	8 178	3 214	***921***	***1 141***	***547***	***202***	2 811
6 809	24 484	27 928	***5 923***	***1 921***	***2 379***	***1 141***	***421***	11 786
11 155	38 229	***26 719***	***7 611***	***2 469***	***3 057***	***1 467***	***541***	41 864
10 641	***35 782***	***25 117***	***7 155***	***2 321***	***2 874***	***1 379***	***509***	75 137
							Total	133 750

The table 3 gives the useful part of the statistic $-2q^{+}$ for some standard GLM distributions according to the canonical, identity and log links. By minimizing this statistics we keep over-dispersed and Gamma models.

Table 3 : values of the statistic $-2q^{+}$

	Canonical	**Identity**	**Log**
Poisson	15 319	37 201	15 319
Over dispersed Poisson	536	569	536
Normal	620	620	2 369
Gamma	530	534	478
Normal – Inverse	797	798	788

6.1 ***Over-dispersed Poisson model***

The table 4 gives the parameters estimates and their associated standard errors. We present two ways of estimating the dispersion parameter ϕ : either with the deviance, or with the Pearson residuals. The standard error is impacted but not the parameters estimates.

Table 4: Parameters estimates

	Over-dispersed Poisson model			
	Estimation of ϕ : Deviance		**Estimation of ϕ : Pearson**	
Parameter	Estimates	Standard errors	Estimates	Standard errors
μ	7.2447	0.2914	7.2447	0.3083
α_0	0.0000	0.0000	0.0000	0.0000
α_1	0.1716	0.3429	0.1716	0.3627
α_2	0.5753	0.3174	0.5753	0.3358
α_3	0.9563	0.3011	0.9563	0.3186
α_4	1.1035	0.2968	1.1035	0.3140
α_5	1.8388	0.2793	1.8388	0.2954
α_6	2.0896	0.2881	2.0896	0.3048
α_7	2.0278	0.3902	2.0278	0.4128
β_0	0.0000	0.0000	0.0000	0.0000
β_1	1.2127	0.1664	1.2127	0.1761
β_2	0.8588	0.1936	0.8588	0.2048
β_3	-0.3969	0.3261	-0.3969	0.3450
β_4	-1.5229	0.6223	-1.5229	0.6584
β_5	-1.3090	0.7173	-1.3090	0.7588
β_6	-2.0434	1.3617	-2.0434	1.4406
β_7	-3.0400	3.2824	-3.0400	3.4725
ϕ	716.1832	-	801.5148	-

Hereafter is the estimated variance-covariance matrix $\Sigma_{as}(\hat{\xi})$ (section 4.1) obtained with ϕ estimated by deviance

0,08	-0,06	-0,06	-0,06	-0,07	-0,07	-0,07	-0,08	-0,02	-0,02	-0,03	-0,03	-0,04	-0,05	-0,08
- 0,06	0,12	0,06	0,06	0,06	0,06	0,06	0,06	0,00	0,00	0,00	0,00	0,00	0,00	0,06
- 0,06	0,06	0,10	0,06	0,06	0,06	0,06	0,06	0,00	0,00	0,00	0,00	0,00	0,03	0,06
- 0,06	0,06	0,06	0,09	0,06	0,06	0,06	0,06	0,00	0,00	0,00	0,00	0,02	0,03	0,06
- 0,07	0,06	0,06	0,06	0,09	0,07	0,07	0,07	0,00	0,00	0,00	0,01	0,02	0,03	0,07
- 0,07	0,06	0,06	0,06	0,07	0,08	0,07	0,07	0,00	0,00	0,01	0,01	0,02	0,03	0,07
- 0,07	0,06	0,06	0,06	0,07	0,07	0,08	0,07	0,00	0,01	0,01	0,01	0,02	0,03	0,07
- 0,08	0,06	0,06	0,06	0,07	0,07	0,07	0,15	0,02	0,02	0,03	0,03	0,04	0,05	0,08
- 0,02	0,00	0,00	0,00	0,00	0,00	0,00	0,02	0,03	0,02	0,02	0,02	0,02	0,02	0,02
- 0,02	0,00	0,00	0,00	0,00	0,00	0,01	0,02	0,02	0,04	0,02	0,02	0,02	0,02	0,02
- 0,03	0,00	0,00	0,00	0,00	0,01	0,01	0,03	0,02	0,02	0,11	0,03	0,03	0,03	0,03
- 0,03	0,00	0,00	0,00	0,01	0,01	0,01	0,03	0,02	0,02	0,03	0,39	0,03	0,03	0,03
- 0,04	0,00	0,00	0,02	0,02	0,02	0,02	0,04	0,02	0,02	0,03	0,03	0,51	0,04	0,04
- 0,05	0,00	0,03	0,03	0,03	0,03	0,03	0,05	0,02	0,02	0,03	0,03	0,04	1,85	0,05
- 0,08	0,06	0,06	0,06	0,07	0,07	0,07	0,08	0,02	0,02	0,03	0,03	0,04	0,05	10,77

Products of matrix (see section 4.2) give the relative standard errors in table 5.

Table 5: Reserve estimates and estimation risks

	Over-dispersed Poisson model				**Bootstrap (1000 samples)**	
	Estimation of ϕ : Deviance		**Estimation of ϕ : Pearson residuals**			
	Estimates of $E(R_i)$	**$se(R_i)/R_i$**	**Estimates of $E(R_i)$**	**$se(R_i)/R_i$**	**Estimates of $E(R_i)$**	**$se(R_i)/R_i$**
1	80	329%	80	348%	71	331%
2	442	134%	442	142%	427	123%
3	1 631	70%	1 631	74%	1 624	62%
4	2 811	53%	2 811	56%	2 777	46%
5	117 86	32%	11 786	34%	11 706	28%
6	41 864	20%	41 864	21%	41 799	18%
7	75 137	31%	75 137	33%	75 595	29%
Total	133 750	19%	133 750	20%	134 000	20%

In the tables 6 and 7 we give the estimates of the percentiles of *R* using additivity and Normal Power approximation.

A. Additivity

$\hat{\mu} = 133750$, $\sqrt{\hat{\phi}} = 26,762$ (28,311) where $\hat{\phi}$ is the deviance estimate (or Pearson Chi-squared) :

Table 6: Percentiles estimates

$1-\eta$	0.50	0.75	0.80	0.90	0.95	0.99
$q_{1-\eta}$	0	0.6745	0.8416	1.2816	1.6449	2.3263
$\widehat{q_{1-\eta}(R)}$	133 750 (133 750)	140 351 (140 733)	141 987 (142 463)	146 293 (147 019)	149 849 (150 781)	156 518 (157 836)

B. NP Approximation

$\widehat{\mu} = 133750$, $\sqrt{\widehat{\phi}} = 26,762(28,311)$, $\widehat{\sigma} = \sqrt{\widehat{\phi}\widehat{\mu}} = 9787,358\ (10353,856)$

With $\widehat{\gamma_1} = \sqrt{\dfrac{\widehat{\phi}}{\widehat{\mu}}} = 0,0732\ (0,0774)$

we obtain the table 7 giving the estimated percentiles of R :

Table 7: Percentiles estimates

$1-\eta$	0.50	0.75	0.80	0.90	0.95	0.99
$q_{1-\eta}$	0.0000	0.6745	0.8416	1.2816	1.6449	2.3263
$\widehat{q_{1-\eta}^{(NP)}(R)}$	123 963 (123 396)	130 619 (130 441)	132 284 (132 205)	136 702 (136 885)	140 385 (140 789)	147 377 (148 205)

C. Gamma Approximation

We observe a very small value of γ_1 (0.0753). As a consequence the approximation based on a Gamma distribution is not suitable.

6.2 Gamma model

The table 8 gives the parameters estimates of a Gamma model with Log link and their respective standard errors. ϕ has been estimated with maximum likelihood, deviance and Pearson residuals.

Table 8: Parameters Estimates and standard errors

	Gamma Model					
	Estimation of ϕ : likelihood		**Estimation of ϕ : Deviance**		**Estimation of ϕ : Pearson**	
parameters	**Estimates**	**Standard errors**	**Estimates**	**Standard errors**	**Estimates**	**Standard errors**
μ	7.2097	0.1783	7.2097	0.2355	7.2097	0.2409
α_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
α_1	0.4076	0.1791	0.4076	0.2366	0.4076	0.2421
α_2	0.8203	0.1880	0.8203	0.2483	0.8203	0.2540
α_3	0.9075	0.1983	0.9075	0.2619	0.9075	0.2679
α_4	1.2144	0.2151	1.2144	0.2841	1.2144	0.2906
α_5	1.9319	0.2365	1.9319	0.3124	1.9319	0.3196
α_6	2.1280	0.2781	2.1280	0.3673	2.1280	0.3757
α_7	2.0627	0.3727	2.0627	0.4923	2.0627	0.5036
β_0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
β_1	1.1958	0.1783	1.1958	0.2355	1.1958	0.2409
β_2	0.7055	0.1901	0.7055	0.2511	0.7055	0.2568
β_3	-0.5224	0.2005	-0.5224	0.2648	-0.5224	0.2709
β_4	-1.4714	0.2158	-1.4714	0.2850	-1.4714	0.2915
β_5	-1.5017	0.2379	-1.5017	0.3142	-1.5017	0.3214
β_6	-2.1960	0.2844	-2.1960	0.3756	-2.1960	0.3842
β_7	-3.0050	0.3727	-3.0050	0.4923	-3.0050	0.5036
ϕ	0.1071	2.1620	0.1869		0.1956	

With these estimates we get the predicted values by the models and the relative standard errors (table 9):

Table 9: Reserve estimates and estimation risks

	Gamma Model					
	Estimation of ϕ : likelihood		**Estimation of ϕ : Deviance**		**Estimation of ϕ : Pearson**	
	Estimates of $E(R_i)$	**$se(R_i)/R_i$**	**Estimates of $E(R_i)$**	**$se(R_i)/R_i$**	**Estimates of $E(R_i)$**	**$se(R_i)/R_i$**
1	101	37%	101	49%	101	50%
2	494	25%	494	33%	494	34%
3	1 286	22%	1 286	29%	1 286	30%
4	2 793	22%	2 793	28%	2 793	29%
5	11 262	23%	11 262	30%	11 262	31%
6	36 702	27%	36 702	35%	36 702	36%
7	69 563	36%	69 563	48%	69 563	49%
Total	122 200	22%	122 200	29%	122 200	30%

We deduce the estimates $\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma_1}$ for the reserve *R*:

$$\widehat{\mu} = 122192\,,\ \widehat{\sigma} = 21129\,,\ \widehat{\gamma_1} = 0.4927$$

.

Percentiles estimates are given in tables 10 (Normal Power) and 11 (Translated Gamma).

Results in brackets are the percentiles estimates using the chain ladder estimate $\widehat{\mu} = 133750$ of μ :

A. Normal Power Approximation

Table 10: Percentiles estimates (Normal Power approximation**)**

$1-\eta$	0.50	0.75	0.80	0.90	0.95	0.99
$q_{1-\eta}$	0	0.6745	0.8416	1.2816	1.6449	2.3263
$\widehat{q_{1-\eta}^{(NP)}(R)}$	101 063 (112 621)	116 104 (127 662)	120 074 (131 632)	130 992 (142 550)	140 513 (152 071)	159 606 (171 164)

B. Translated Gamma approximation

With the values above we deduce $\nu = 16.4750$, $\beta = 0.000192$, $x_0 = 36431$ and the estimated percentiles in table 11 :

Table 11: Percentiles estimates (Gamma approximation)

$1-\eta$	0.50	0.75	0.80	0.90	0.95	0.99
$\chi^2_{2\nu}(1-\eta)$	32.2878	37.9965	39.5081	43.6774	47.3338	54.7324
$\widehat{q_{1-\eta}^{(GT)}(R)}$	120 469 (132 027)	135 327 (146 885)	139 262 (150 820)	150 113 (161 671)	159 630 (171 188)	178 887 (190 445)

7. CONCLUSION

Within the general framework of GLM models, by definition asymptotic, this paper gives a simple procedure to determinate the estimation risk of any reserving parameter. It provides also with alternative methods to estimate the percentiles of reserves amounts.

The fact that such approaches can be used within a quasi-likelihood framework potentially increases the scope of modeling based on a claims amounts triangle.

Appendix: the Delta Method[12]

1. Real parameter case

Given a random variable *X* with distribution depending on a parameter $\theta \in \Theta$, open in $\mathbb{R}$ and $(X_1,...,X_n)$ a *n*-sample (i.i.d.) from *X*.

Given an estimator $T_n(X_1,...,X_n)$ of θ asymptotically Normal[13] with asymptotic variance and standard error:

$$V_{as}(T_n) = \frac{\sigma^2(\theta)}{n} \quad \text{and} \quad se_{as}(T_n) = \frac{\sigma(\theta)}{\sqrt{n}}$$

meaning that:

$$T_n \ AN(\theta, \frac{\sigma^2(\theta)}{n}) \text{ or } \frac{\sqrt{n}(T_n - \theta)}{\sigma(\theta)} \xrightarrow[n\to+\infty]{L} N(0,1)$$

Remark: if σ is a continuous function $\widehat{V_{as}(T_n)} = \frac{\sigma^2(T_n)}{n}$ estimates $V_{as}(T_n)$

Given a regular function π [14], then $\pi(T_n)$ is an asymptotically Normal estimator of $\pi(\theta)$ with asymptotic variance:

$$V_{as}(T_n) = \frac{\left[\pi'(\theta)\sigma(\theta)\right]^2}{n}$$

$$\pi(T_n) \ AN\left[\pi(\theta), \frac{[\pi'(\theta)\sigma(\theta)]^2}{n}\right]$$

Remark: if σ is a continuous function and if π is continuous and derivable then $V_{as}(T_n)$ and $se_{as}(T_n)$ are estimated by :

$$\widehat{V_{as}(T_n)} = \frac{\left[\pi'(T_n)\sigma(T_n)\right]^2}{n} \text{ and } \widehat{se_{as}(T_n)} = \frac{\left|\pi'(T_n)\sigma(T_n)\right|}{\sqrt{n}}$$

As the Slutsky theorem gives $$\frac{\left[\pi(T_n) - \pi(\theta)\right]}{\widehat{se_{as}(T_n)}} \xrightarrow[n\to+\infty]{L} N(0,1)$$

we deduce an level 95% asymptotic confidence interval for $\pi(\theta)$:

$$\left[\pi(T_n) - 1{,}96\,\widehat{se_{as}(T_n)}\ ; \pi(T_n) + 1{,}96\,\widehat{se_{as}(T_n)}\right]$$

---

12 Proofs of these results are given in Shao,1998.

13 As every maximum likelihood estimator under some regularity conditions

14 Derivable function with $\pi'(\theta) \neq 0$ for $\theta \in \Theta$.

2. General case

Given a random variable T_n whose distribution depends on a parameter $\theta \in \Theta$ (with Θ open in $\mathbb{R}^r$) and is asymptotically Normal :

$$T_n \ AN_r[\theta, \frac{1}{n}\Sigma(\theta)] \ \ ie \ \sqrt{n}(T_n - \theta) \xrightarrow[n\to+\infty]{L} N_r[0, \Sigma(\theta)]$$

where $\Sigma_{as}(T_n) = \frac{1}{n}\Sigma(\theta)$ is the asymptotic variance-covariance matrix of T_n.

Given $\pi = (\pi_1, ..., \pi_s)$ a regular function on Θ with values in $\mathbb{R}^s$ and the Jacobian matrix of π, $\Delta_\theta = \left(\frac{\partial \pi_i}{\partial \theta_j}\right) i = 1,...,s; j = 1,...,r$, then

$$\pi(T_n) \quad \text{is } AN\left[\pi(\theta), \frac{1}{n}\Delta_\theta \Sigma(\theta)\Delta_\theta{}'\right]$$

so that the variance-covariance matrix of $\pi(T_n)$ is $\Sigma_{as}\left[\pi(T_n)\right] = \frac{1}{n}\Delta_\theta \Sigma(\theta)\Delta_\theta{}'$.

Denoting $\Sigma(\theta) = \left[\sigma_{kl}(\theta)\right]_{k,l=1,...,r}$, the $(i, j)^{th}$ order component of the matrix is defined by:

$$\sigma_{ij}^{(\pi)}(\theta) = \frac{1}{n}\sum_{k=1}^{r}\sum_{l=1}^{r}\sigma_{kl}(\theta)\frac{\partial \pi_i}{\partial \theta_k}(\theta)\frac{\partial \pi_j}{\partial \theta_l}(\theta)$$

with

$$\sigma_{ii}^{(\pi)}(\theta) = V_{as}\left[\pi_i(T_n)\right] \text{ and } \sigma_{ij}^{(\pi)}(\theta) = Cov_{as}\left[\pi_i(T_n), \pi_j(T_n)\right]$$

Remark: if π is a C^1 function and if σ_{kl} is a continuous function for $k, l = 1,...,r$ then $\Sigma_{as}\left[\pi(T_n)\right]$ is estimated by $\widehat{\Sigma_{as}(T_n)} = \frac{1}{n}\Delta_{T_n}\Sigma(T_n)\Delta_{T_n}{}'$

REFERENCES

Denuit M. Charpentier A. (2004) : *Mathématiques de l'assurance non-vie,* tome1, Economica

England P.D.. Verrall R. J. (1999) : "Analytic and bootstrap estimates of prediction error in claims reserving" *Ins. : Math. & Econ*. Vol. 25. 281-293.

England P.D.. Verrall R. J. (2001) : "A flexible framework for stochastic claims reserving" *Cas. Act. Soc*.

England P.D.. Verrall R. J. (2002) : "Stochastic claims reserving in General Insurance" *Institute of Actuaries*.

Fahrmeir L., Tutz G. (2001) : *Multivariate statistical modeling based on Generalized linear models.* 2nd ed. Springer.

Gerber H.U. (1979) : *An introduction to mathematical risk theory.* Huebner Foundation for Insurance Education.

Kaas R.. Goovaerts M.. Dhaene J.. Denuit M. (2001) : *Modern Actuarial Risk Theory* Kluwer Acad. Press.

Klugman S.A., Panjer H.H., Willmot G.E. (1998) : *Loss models : from data to decisions.* Wiley.

Kremer E. (1982) : "IBNR claims and the two way model of Anova" *Scand. Act. J.*. 47-55.

Mack T. (1993) : "Distribution free calculation of the standard error of Chain ladder reserve estimates" *Astin Bull*. Vol. 23. 213-225.

Mack T. (1994) : "Which stochastic model is underlying the chain ladder model" *Ins.* : *Math. & Econ*. Vol. 15. 133-138

Mc Cullagh P.. Nelder J. (1989) : *Generalized Linear Models,* 2nd ed. Chapman &Hall

Partrat C.. Besson J-L. (2005) : *Assurance non-vie : modélisation, simulation.* Economica.

Nelder J.A., Pregibon D. (1987) : An extended quasi-likelihood function. *Biometrika,*vol.74.221-232.

Partrat C., Besson J-L. (2005) : *Assurance non-vie : modélisation, simulation.* Economica

Renshaw A.E. (1994) : Modelling the claim process in the presence of covariates. *Astin Bull.*vol. 24. 265-285

Renshaw A. E.. Verrall R. J. (1998) : "A stochastic model underlying the chain ladder technique" *British Act. J.* Vol. 4. 903-923

Shao J.. Tu D. (1995) : "The Jackknife and bootstrap" Springer

Taylor G. (2000) : *Loss reserving : an actuarial perspective.* Kluwer Acad. Press

Verrall R. J. (1991) : "On the estimation of reserves from Log linear models" *Ins. : Math. & Econ.* Vol. 10. 75-80

Verrall R. J. (2000) : " An investigation into stochastic claims reserving models and the chain ladder techniques" *Ins. : Math. & Econ.* Vol. 26. 91-99.

Wüthrich M.V. (2003) : Claims reserving using Tweedie's compound Poisson model. *Astin Bull.,* vol.33, 331-346.