

- **PARTRAT, CHRISTIAN**
- **ISFA, Université Claude Bernard Lyon 1**
50, avenue Tony Garnier
69366 Lyon Cedex 07
France
- **Home : 97-99, avenue du Général Leclerc**
75014 Paris
France
- **Tél. : 33 6 60 50 49 35**
- **Fax : 33 1 45 41 33 10**
- **E-mail : partrat@univ-lyon1.fr**

Résumé

Dans le cadre standard des Modèles Linéaires Généralisés, ayant comme cas particulier la méthode chain ladder, cet article présente un modèle stochastique de projection de triangles de liquidation directe et flexible. On peut en déduire des projections de résultats techniques ainsi que des évaluations de besoins en capital. Aux différents résultats obtenus peuvent être associés des risques d'estimation et des intervalles de confiance.

Mots clés

Assurance non-vie, provisionnement, Modèles Linéaires Généralisés, modèles Poisson surdispersé (chain ladder) et Gamma, méthode Delta, projection de triangles de liquidation et de résultats techniques, modèles internes.

Abstract

Inside the standard Generalized Linear Models frameworks, with the chain ladder method as a particular case, this paper presents a direct and simple stochastic model for projecting run-off triangles and technical results. Then we could deduce valuation of needs of capital. If necessary the model allows obtaining estimation risks and confidence intervals.

Keywords

General insurance, Reserving, Generalized Linear Models, Overdispersed Poisson (chain ladder) and Gamma models, Delta method, Projected run-off triangles and technical results, internal models.

PROJECTION DYNAMIQUE STOCHASTIQUE d'un TRIANGLE de LIQUIDATION

Christian PARTRAT

ISFA, Université Claude Bernard Lyon 1
France

1. PRELIMINAIRES

On analyse le déroulement temporel dans une (sous-) branche non-vie donnée d'une quantité liée à la sinistralité de celle-ci. Les quantités analysées sont de natures très diverses :

- paiements de sinistres, recours, ,...
- nombres de sinistres : déclarés, réglés, tardifs,...
- loss ratios (montants / primes acquises),...

Les sinistres sont rapportés à des périodes : année, semestre, trimestre, ... Pour fixer les idées, on utilisera dans la suite l'année.

L'année récurrente se déroule du 1/1 au 31/12, date d'inventaire (ou de fin d'exercice).

Les sinistres sont rattachés à des années d'origine (*origin year*) qui peuvent être :

- l'année de survenance (*accident year*)
- l'année de souscription (*underwriting year*)
- l'année de déclaration (*report year*)

Par commodité d'expression, nous supposons le plus souvent ci-dessous qu'il s'agit d'années de survenance, rattachement le plus fréquent dans la pratique.

On suppose que les sinistres de la (sous-)branche se déroulent sur $(n+1)$ années. On note

i l'année d'origine (*origin year*), $i = 0, \dots, n$

j le délai de développement (*development year*), $j = 0, \dots, n$

x_{ij} ou X_{ij} la mesure de sinistralité correspondant à l'année d'origine i et au délai j , qu'elle soit connue ou encore aléatoire.

Dans la suite on supposera, pour fixer les idées, que cette mesure correspond à des paiements de sinistres sous la forme non cumulée (incrément, *increments*).

Au 31/12/ n , il est nécessaire de prendre en compte les années de survenance 0 à n . Les paiements de sinistres antérieurs à cette date x_{ij} ($i + j \leq n$) sont classiquement mis sous la forme d'un triangle de liquidation de montants non cumulés dont la forme générale est donnée au tableau 1. L'année 0 est, par hypothèse, complètement déroulée.

Remarque : On peut noter que ce triangle contient $t = 1 + 2 + \dots + (n+1) = \frac{(n+1)(n+2)}{2}$ données soit 21 données pour $n=5$ (short tail) et 136 pour $n=15$ (long tail). Les résultats asymptotiques de cet article supposent t (et non la longueur de liquidation n) « grand ».

On se place dans une problématique de projection de sinistralité sur un horizon de m années au delà du 31/12/ n , soit, dans le tableau 1, les années $n+1$ à $n+m$ de survenances futures.

Les modèles stochastiques développés ci-dessous considèrent l'ensemble des montants non cumulés du tableau 1 comme des variables aléatoires $X_{ij}, i, j = 0, \dots, n+m$. Parmi les $(n+1)^2$ variables du rectangle des années 0 à n au cours desquelles les sinistres sont survenus, les v.a.r. $(X_{ij})_{i+j \leq n}$ du triangle supérieur ont été observées, $(x_{ij})_{i+j \leq n}$ étant leur réalisation.

Les règlements restant à effectuer au titre des années 1 à n sont les variables aléatoires X_{ij} ($i, j = 1, \dots, n, i+j > n$) du triangle inférieur.

Les variables $X_{ij}, i = n+1, \dots, n+m$ sont les montants de survenances futures.

Tableau 1 : Projection de triangle de liquidation

Année d'origine	0	...	j	...	l	...	n	Charges sinistres
0	x_{00}						x_{0n}	S_0
...								...
i			x_{ij}					S_i
...								...
k					X_{kl}			S_k
...								...
n	x_{n0}						$X_{n,n}$	S_n
$n+1$	$X_{n+1,0}$						$X_{n+1,n}$	S_{n+1}
...								...
...								...
$n+m$	$X_{n+m,0}$						$X_{n+m,n}$	S_{n+m}

D'autre part les diagonales du tableau 1 correspondent à des règlements calendaires : l'année calendaire n (indiquée sur le rectangle de liquidation), par exemple, a donné lieu au

paiement total $\sum_{i=0}^n x_{i,n-i}$, toutes années d'origine confondues.

Pour $i=1, \dots, n$, le montant cumulé des sinistres survenus lors du i ème exercice s'écrit

$$S_i = \sum_{j=0}^{n-i} x_{ij} + \sum_{j=n-i+1}^n X_{ij} = \sum_{j=0}^{n-i} x_{ij} + R_i$$

en notant $R_i = \sum_{h=n-i+1}^n X_{ih}$ la provision (aléatoire) correspondante au 31/12/ n .

La provision globale (aléatoire) à cette même date s'écrit

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

dont on notera F_R la fonction de répartition (f.r.).

Pour $k=1, \dots, m$, les montants de sinistres futurs sont les v.a.r. $S_{n+k} = \sum_{j=0}^n X_{n+k,j}$

Les modèles d'incrémentés utilisés dans cet article s'appuient sur l'hypothèse fondamentale de leur indépendance :

HI : Les v.a.r. X_{ij} , $i, j = 0, \dots, n$ sont indépendantes.

Cette hypothèse exclut en particulier tout phénomène de compensation des règlements de sinistres d'un délai au suivant au cours d'une même année d'origine. L'hypothèse d'indépendance ci-dessus est souvent invalidée par des variations annuelles d'inflation des montants de sinistre dans la (sous-) branche étudiée, susceptible d'introduire des « corrélations » entre les X_{ij} , $i, j = 0, \dots, n$. Une telle situation nécessiterait une mise as-if préalable des données.

Il faut noter l'existence de tests statistiques non paramétriques permettant de juger le caractère réaliste de cette hypothèse d'indépendance au vu des observations du triangle supérieur, avant ou après mise as-if.

La modélisation utilisée dans cet article est paramétrique au sens où les v.a.r. X_{ij} sont supposées munies de lois de probabilité de même « type » (Poisson, Normale, LogNormale, Gamma, loi prise dans la famille exponentielle des GLM, ...). Ces distributions dépendent de paramètres « naturels » θ_{ij} réels ou vectoriels et, pour certains d'entre eux, d'un paramètre additionnel ϕ dit de « dispersion », indépendant de (i, j) . Pour ne pas alourdir l'exposé, on se limitera dans la suite aux distributions de la famille des GLM.

2. MODELES GLM

Les modèles linéaires généralisés (GLM en abrégé) ont été introduits par J. Nelder et R. Wedderburn en 1972. De très nombreux ouvrages exposent, de manière approfondie, la théorie statistique sous-jacente. L'un des plus classiques est celui de Mc Cullag et Nelder, 1989.

On dispose des v.a.r « réponses » X_{ij} ($i, j = 0, \dots, n$), indépendantes, dont la loi de probabilité est de type exponentiel. La « densité » de X_{ij} ($i, j = 0, \dots, n$) a pour expression

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

où θ_{ij} est un paramètre réel, appelé paramètre naturel.

$\phi > 0$ (ϕ connu éventuellement) est un paramètre de dispersion indépendant de i et j . b et c sont des fonctions spécifiques de la distribution, b étant « régulière ».

On montre que $\mu_{ij} = E(X_{ij}) = b'(\theta_{ij})$ soit, si b' est inversible, $\theta_{ij} = b'^{-1}(\mu_{ij})$ et

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

La fonction V est appelée fonction variance de la distribution.

De plus le moment centré d'ordre 3 et le coefficient d'asymétrie s'écrivent :

$$\mu_3(X_{ij}) = \phi^2 b'''(\theta_{ij}) = \phi^2 V(\mu_{ij}) V'(\mu_{ij}) \quad \gamma_1(X_{ij}) = \frac{\mu_3(X_{ij})}{[V(X_{ij})]^{3/2}} = \sqrt{\phi} \frac{V'(\mu_{ij})}{[V(\mu_{ij})]^{1/2}}$$

Les moments d'ordre supérieur se déduiraient, par dérivations successives, en fonction de (θ_{ij}, ϕ) puis de (μ_{ij}, ϕ) , de la fonction génératrice des moments des X_{ij} :

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

Cas particuliers : Parmi les distributions de type « exponentiel » des GLM, nous utiliserons les lois Poisson surdispersée et Gamma.

– Loi de Poisson $P(\lambda)$

$$P(X = x) = e^{(x \ln \lambda - \lambda) + c(x)}, x \in \mathbb{N}$$

$$\theta = \ln \lambda, \phi = 1, b(\theta) = e^\theta \quad \mu = E(X) = \lambda, V(\mu) = \mu$$

– Loi de Poisson surdispersée $P_{surd}(\lambda, \phi)$

$$\text{Formellement on écrirait } X \approx P_{surd}(\lambda, \phi) \text{ ssi } \frac{X}{\phi} \approx P\left(\frac{\lambda}{\phi}\right)$$

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

Une présentation plus rigoureuse de cette distribution repose sur la quasi-vraisemblance (cf section 4.3).

– Loi Gamma $\left(\nu, \frac{\nu}{\mu}\right)$

$$f(x) = \exp\left[\left(-\frac{x}{\mu} - \ln \mu\right) \nu + c(x, \nu)\right], x > 0$$

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

– Loi Poisson composée de Tweedie

Sans expliciter le mode de génération de cette distribution ni sa densité (cf Wüthrich, 2004), on obtient :

$$\theta = \frac{\mu^{1-r}}{1-r}, \phi > 0, b(\theta) = \frac{1}{2-r} [(1-r)\theta]^{2-r}, E(X) = \mu, V(\mu) = \mu^r$$

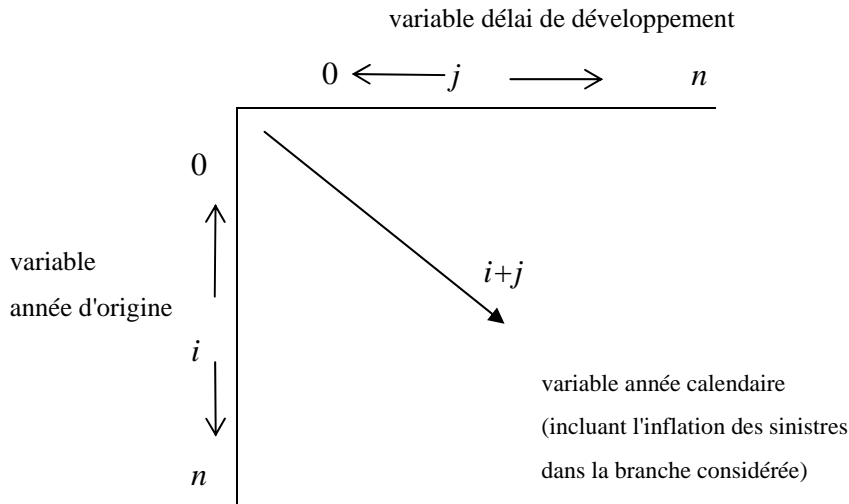
où r est supposé donné.

L'intérêt de cette distribution est de combler l'espace entre les deux modèles ci-dessus, très utilisés pratiquement : Poisson surdispersé ($r=1$) et Gamma ($r=2$).

Remarque : Une approche semblable s'envisagerait pour des transformées de composantes aléatoires GLM. La transformation Log donnerait les distributions LogNormale, Pareto, LogGamma,...

3. VARIABLES EXPLICATIVES

Les variables intervenant dans la modélisation d'un triangle de liquidation correspondent aux trois directions "naturelles" de celui-ci :



Les variables année (origine ou calendaire) sont, par nature, qualitatives *ordinales*, les modalités 0, 1, ... utilisées dans ce texte n'étant que des codages arbitraires.

Cependant les procédures de mise en œuvre des méthodes explicatives ne permettent pas d'intégrer simplement ce type de covariates. C'est pourquoi, dans la suite, les variables année seront, le plus souvent, considérées comme qualitatives pures i.e. des facteurs.

Si l'année d'origine est considérée comme un facteur, celui-ci est à $(n+1)$ modalités 0, 1, ..., n . On sait que l'on lui substitue n variables auxiliaires disjonctives en 0 ou 1 à paramétrage $\alpha_1, \alpha_2, \dots, \alpha_n$, la modalité 0 servant de référence ($\alpha_0 = 0$).

De même l'année calendaire serait paramétrée en μ_{i+j} ($i, j = 0, \dots, n$). Pour un triangle à inflation constante, en particulier déflaté, (ce que nous supposons), l'inflation annuelle est supposée constante : $\mu_{i+j} = \mu$.

Il en va autrement de la variable délai, naturellement quantitative discrète à valeurs 0, 1, ... Il en résulte deux approches selon qu'on la considère comme

- (i) facteur à $(n+1)$ modalités avec paramétrage $\beta_1, \beta_2, \dots, \beta_n$ ($\beta_0 = 0$).
- (ii) variable à valeurs entières.

Dans ce second cas, une modélisation de l'influence du délai sur la loi des X_{ij} serait modélisée, par exemple, à l'aide de fonctions paramétrées de j telles β_j , $\gamma \ln(j+1)$, ...

Dans la suite, on privilégiera la modélisation à variables qualitatives à paramétrages $(\alpha_i)_{i=0, \dots, n}$ pour les années d'origine, $(\beta_j)_{j=0, \dots, n}$ pour les délais de règlement et un paramètre constant μ .

Les modèles explicatifs nécessitent la donnée a priori d'un « lien » entre les distributions des X_{ij} et les paramètres de régression $(\alpha_i), (\beta_j), \mu$. Ce lien passe par la moyenne, sous la forme d'une fonction

$$\mu_{ij} = E(X_{ij}) = K(\mu, \alpha_i, \beta_j)$$

pouvant inclure le paramètre ϕ . Par référence à la méthode chain ladder et pour l'interprétation du paramétrage qu'il permet, nous privilégierons ci-dessous le lien multiplicatif

$$\mu_{ij} = e^{\mu + \alpha_i + \beta_j}.$$

4. ESTIMATION, RISQUE D'ESTIMATION

4.1 Estimation

Dans la modélisation des $(X_{ij})_{0 \leq i, j \leq n}$ à paramètre vectoriel $\xi = \left[\mu, (\alpha_i)_{i=1, \dots, n}, (\beta_j)_{j=1, \dots, n} \right]$, la méthode du maximum de vraisemblance, appliquée aux données du triangle supérieur et à la vraisemblance $L[(x_{ij}); \xi]$, fournit l'estimateur du maximum de vraisemblance (e.m.v. en abrégé) $\hat{\xi} = \left[\hat{\mu}, (\hat{\alpha}_i), (\hat{\beta}_j) \right]$ de ξ et, éventuellement, de ϕ , pour l'estimation duquel il serait seulement nécessaire de disposer d'un estimateur convergent.

Par invariance fonctionnelle du maximum de vraisemblance, on en déduit les e.m.v.

$$\hat{\mu}_{ij} = e^{\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j} \quad \text{des} \quad \mu_{ij} = E(X_{ij}) \quad (i, j = 0, \dots, n).$$

et, plus généralement, de toute fonction des μ_{ij} , en particulier des moments de X_{ij} . Il en est ainsi de $V(X_{ij}) = \phi V(\mu_{ij})$

L'estimation $\hat{E}(X_{ij}) = \hat{\mu}_{ij}$ est la valeur de μ_{ij} "prévue" par le modèle.

Il en résulte un e.m.v. $\hat{E}(R_i) = \sum_{h=n-i+1}^n \hat{\mu}_{ih}$ de $E(R_i)$, pour $i=1, \dots, n$, puis le best estimate de la provision R :

$$\hat{E}(R) = \sum_{i=1}^n \hat{E}(R_i) = \sum_i \sum_n \hat{\mu}_{ij}$$

Remarque : On rappelle le résultat fondamental de Renshaw et Verrall, 1994 et 1998 :

Dans un triangle de liquidation à increments positifs, les valeurs prévues par un modèle de Poisson (simple ou surdispersé) dans le triangle inférieur, les provisions par année d'origine et totale coïncident avec les éléments correspondants calculés à l'aide de la méthode chain ladder standard.

Enfin le triangle supérieur des valeurs rétrospectives calculées à l'aide des facteurs de développement chain ladder appliqués aux éléments de la diagonale principale du triangle coïncide avec le triangle supérieur des valeurs prévues par le modèle.

Remarque : L'hypothèse d'indépendance des increments permettrait d'exprimer aisément les moments de la provision R en fonction du paramètre $\xi = \left[\mu, (\alpha_i)_{i=1, \dots, n}, (\beta_j)_{j=1, \dots, n} \right]$ et ainsi d'en obtenir des estimateurs. Pour $V(R)$, par exemple, on aurait

$$V(R) = \sum_{i=1}^n V(R_i) = \sum_{i=1}^n \sum_{i+j > n} V(X_{ij}) \quad \text{avec} \quad V(X_{ij}) = \phi V(\mu_{ij})$$

4.2 Risque d'estimation

L'incertitude dans l'estimation de $E(R_i)$ et $E(R)$ est mesurée par $V[\widehat{E}(R_i)]$ ($i=1, \dots, n$) et $V[\widehat{E}(R)]$, dont l'évaluation nécessite la connaissance des variances et covariances des $\widehat{\mu}_{ij}$:

$$V(\widehat{E}(R_i)) = \sum_{h=n-i+1}^n V(\widehat{\mu}_{ih}) + \sum_{h_1=n-i+1}^n \sum_{\substack{h_2=n-i+1 \\ h_1 \neq h_2}}^n \text{cov}(\widehat{\mu}_{h_1 i}, \widehat{\mu}_{h_2 i})$$

Le calcul **exact** de ces variances et covariances des $\widehat{\mu}_{ij}$ est le plus souvent hors d'atteinte. On leur substitue des valeurs approchées obtenues soit par des simulations de Monte Carlo dans un cadre bootstrap (cf England et al., 2002) soit par une approche asymptotique justifiée par la méthode Delta. Cette dernière conduit à des expressions analytiques rappelées ci-dessous. Ces deux méthodes donnent, si besoin est, des intervalles de confiance pour $E(R)$ ou d'autres paramètres liés à R .

La matrice de variances-covariances asymptotique, notée $\Sigma_{as}(\widehat{\xi})$, de l'estimateur $\widehat{\xi}$ fournit, sur sa diagonale, les variances asymptotiques $\sigma_{as}^2(\widehat{\mu})$, $\sigma_{as}^2(\widehat{\alpha}_i)$, $\sigma_{as}^2(\widehat{\beta}_j)$ et, de part et d'autre, les covariances

$$\text{cov}_{as}(\widehat{\mu}, \widehat{\alpha}_i), \text{cov}_{as}(\widehat{\mu}, \widehat{\beta}_j), \text{cov}_{as}(\widehat{\alpha}_i, \widehat{\beta}_j).$$

Sous des conditions standard sur le maximum de vraisemblance (cf Shao, 1999). $\widehat{\xi}$ est asymptotiquement Normal $AN[\xi, \Sigma_{as}(\widehat{\xi})]$. Il en résulte, par la méthode Delta, que $\widehat{E}(R)$ est asymptotiquement Normal $AN\{E(R), \sigma_{as}^2[\widehat{E}(R)]\}$ avec

$$\sigma_{as}^2[\widehat{E}(R)] = M \Sigma_{as}(\widehat{\xi}) M'$$

où M est une matrice aisément calculable par tableur et explicitée dans Partrat et al., 2005.

La substitution de son estimation $\widehat{\xi}$ au paramètre ξ (et, si nécessaire, d'un estimateur convergent à ϕ) dans ces résultats conduisent à une mesure du risque d'estimation de la provision totale par l'erreur standard asymptotique relative estimée $\frac{\widehat{\sigma}_{as}[\widehat{E}(R)]}{\widehat{E}(R)}$.

De la Normalité asymptotique de $\widehat{E}(R)$, on déduit un intervalle de confiance pour $E(R)$ de niveau asymptotique $(1-\eta)$ ($0 < \eta < 1$) :

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

où $q_{1-\eta}$ est le quantile d'ordre $1-\eta$ de la loi Normale standard.

Remarque : Sous des conditions de régularité, la même approche pourrait s'appliquer à la variance, à tous les moments de R et, par exemple, à la marge $E(R) + \gamma \sigma(R)$.

4.3 Résidus, déviance

La qualité d'ajustement d'un modèle aux données du triangle supérieur, la détection de cellules « anormales » pour celui-ci et le contrôle des hypothèses du modèle reposent sur les résidus, écarts normalisés entre observations x_{ij} et valeurs prévues par le modèle $\hat{\mu}_{ij}$, pour $i + j \leq n$. Les résidus bruts $r_{ij} = x_{ij} - \hat{\mu}_{ij}$ n'étant pas pertinents dans les GLM, on utilise ceux de la déviance (ou alternativement ceux de Pearson) définis par :

$$r_{ij}^{(D)} = \text{sgn}(x_{ij} - \hat{\mu}_{ij}) \sqrt{d_{ij}}$$

où $d_{ij} = 2 \left\{ x_{ij} (\tilde{\theta}_{ij} - \hat{\theta}_{ij}) - [b(\tilde{\theta}_{ij}) - b(\hat{\theta}_{ij})] \right\}$, en posant $\tilde{\theta}_{ij} = b'^{-1}(x_{ij})$ et $\hat{\theta}_{ij} = b'^{-1}(\hat{\mu}_{ij})$.

La déviance du modèle est la somme des carrés de ces résidus :

$$D = \sum_{i+j \leq n} [r_{ij}^{(D)}]^2 = 2 \sum_{i+j \leq n} \left\{ x_{ij} (\tilde{\theta}_{ij} - \hat{\theta}_{ij}) - [b(\tilde{\theta}_{ij}) - b(\hat{\theta}_{ij})] \right\}.$$

Cas particuliers :

$$\text{Poisson } P(\lambda) : D = 2 \sum_{i+j \leq n} \left[x_{ij} \ln \frac{x_{ij}}{\hat{\mu}_{ij}} - (x_{ij} - \hat{\mu}_{ij}) \right]$$

$$\text{Gamma } \gamma \left(\nu, \frac{\nu}{\mu} \right) : D = 2 \sum_{i+j \leq n} \left(\frac{x_{ij} - \hat{\mu}_{ij}}{\hat{\mu}_{ij}} - \ln \frac{x_{ij}}{\hat{\mu}_{ij}} \right)$$

Remarque : D'un accès plus simple que celui du maximum de vraisemblance, l'estimateur suivant de ϕ , basés sur les résidus de déviance, est convergent

$$\hat{\phi} = \frac{1}{t-p} \sum_{i+j \leq n} (r_{ij}^{(D)})^2 = \frac{D}{t-p}.$$

où t et p sont les nombres de données du triangle supérieur et de paramètres de régression.

Le modèle est d'autant mieux ajusté aux données que l'indicateur D est faible. Cependant cette statistique ne permet la comparaison de modèles que s'ils sont issus d'une **même** composante aléatoire (mêmes V et ϕ).

Par extension de la log-vraisemblance, on introduit la quasi-(log) vraisemblance (McCullagh et al., 1989) par

$$q(\underline{x}, \underline{\mu}) = \sum_{i+j \leq n} \int_{x_{ij}}^{\mu_{ij}} \frac{x_{ij} - u}{\phi V(u)} du$$

déterminée par la seule donnée d'une fonction variance $V(\mu) > 0$ et, éventuellement, un paramètre de dispersion ϕ , sans référence à une distribution de type exponentiel sous-jacente.

Par analogie, la déviance ci-dessus s'étend à la quasi-déviance D :

$$D = \sum_{i+j \leq n} d_{ij} = -2\phi q(\underline{x}, \hat{\underline{\mu}}) \quad \text{en posant} \quad d_{ij} = 2 \int_{\hat{\mu}_{ij}}^{x_{ij}} \frac{x_{ij} - u}{V(u)} du.$$

Cas particuliers

- Loi de Poisson surdispersée $P_{surd}(\lambda, \phi)$
correspond au couple (λ, ϕ) tel que $\mu = E(X) = \lambda$, $V(\mu) = \mu$.
Il en résulte $V(X) = \phi\mu = \phi E(X)$ (cf section 2).
 - Fonction variance puissance
 $V(\mu) = \mu^\tau$ ($\tau \in \mathbb{R}$) ($V(\mu) = \ln \mu$ si $\mu = 0$)
dont la quasi-vraisemblance s'écrit $\int_x^\mu \frac{x-u}{V(u)} du = \frac{1}{\mu^\tau} \left(\frac{x\mu}{1-\tau} - \frac{\mu^2}{2-\tau} \right)$, $\tau \neq 0, 1, 2$
- Ce cas particulier est à mettre en relation avec le modèle composé de Tweedie.

Dans une finalité de comparaison de modèles, Nelder et al. (1987) ont introduit la quasi-vraisemblance étendue q^+ définie par l'équation

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

soit une quasi-déviance pénalisée par (ϕ, V) , analogue aux critères d'Akaike et de Schwarz.

La statistique q^+ , à maximiser, permet de comparer des modèles à structures (V, ϕ) différentes. Il est équivalent de minimiser $-2q^+$.

Exemple : Le tableau 2 donne les paiements non cumulés de sinistres effectués par une compagnie d'assurances pour son portefeuille Corps maritimes au titre des exercices de souscription 1984 à 1991. Les paramètres estimés quand sont ajustés les modèles Poisson surdispersé puis Gamma forment le tableau 3. Celui-ci est complété par les erreurs standard correspondantes.

De plus les estimations de ϕ par déviance sont respectivement :

$$\text{Poisson surdispersé : } \hat{\phi} = 716,18 \quad ; \quad \text{Gamma } \hat{\phi} = 0,1869$$

Les valeurs de la statistique $-2q^+$ de la quasi-vraisemblance étendue sont respectivement :

$$\text{Poisson surdispersé : } 536 \quad \text{Gamma : } 478$$

montrant une qualité d'ajustement acceptable, légèrement meilleure pour la modélisation Gamma. Cet indicateur prenait des valeurs nettement plus élevées pour des distributions GLM alternatives (Poisson simple, Normale, Inverse-Gaussienne, ...).

Tableau 2 - Triangle de liquidation des montants non cumulés

Exercices	0	1	2	3	4	5	6	7
0	1381	4399	4229	435	465	205	110	67
1	859	6940	2619	1531	517	572	287	
2	6482	6463	3995	1420	547	723		
3	2899	16428	5521	2424	477			
4	3964	15872	8178	3214				
5	6809	24484	27928					
6	11155	38229						
7	10641							

Tableau 3 - Estimation et risques d'estimation des paramètres de régression, modèles Poisson surdispersé et Gamma.

Paramètres de régression	E.m.v. Poisson surdispersé	Erreurs standard Poisson surdispersé	E.m.v. Gamma	Erreurs standard Gamma
μ	7,2447	0,2914	7.2097	0.2355
α_1	0,1716	0,3429	0.4076	0.2366
α_2	0,5753	0,3174	0.8203	0.2483
α_3	0,9563	0,3011	0.9075	0.2619
α_4	1,1035	0,2968	1.2144	0.2841
α_5	1,8388	0,2793	1.9319	0.3124
α_6	2,0896	0,2881	2.1280	0.3673
α_7	2,0278	0,3902	2.0627	0.4923
β_1	1,2127	0,1664	1.1958	0.2355
β_2	0,8588	0,1936	0.7055	0.2511
β_3	-0,3969	0,3261	-0.5224	0.2648
β_4	-1,5229	0,6223	-1.4714	0.2850
β_5	-1,3090	0,7173	-1.5017	0.3142
β_6	-2,0434	1,3617	-2.1960	0.3756
β_7	-3,0400	3,2824	-3.0050	0.4923

Nous explicitons, seulement dans le cas Poisson surdispersé, la matrice de variances-covariances asymptotique $\Sigma_{as}(\hat{\xi})$ de $\hat{\xi}$ en tableau 4. Nous déduisons, dans le tableau 5, les best estimates des provisions par exercice puis de la provision globale au 31/12/1991, pour chacun des deux modèles. Ce tableau est complété par les risques d'estimation correspondants sous la forme des erreurs standard relatives estimées.

Par Normalité asymptotique de $\hat{E}(R)$, on obtient, dans le modèle Poisson surdispersé, une estimation par intervalle de $E(R)$ de niveau asymptotique 0,95 :

$$\left[\hat{E}(R) - 1,96 \widehat{se}_{as} \left[\hat{E}(R) \right], \hat{E}(R) + 1,96 \widehat{se}_{as} \left[\hat{E}(R) \right] \right]$$

soit pour ce triangle [83941, 183559].

Dans la modélisation Poisson surdispersé, le tableau 6 fournit le rectangle des valeurs prévues $\hat{E}(X_{ij}) = \hat{\mu}_{ij} = e^{\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j}$. Celles-ci coïncident exactement avec les valeurs correspondantes issues des facteurs de développement du tableau 7 par la méthode chain ladder.

Tableau 4 - Matrice de variances-covariances des paramètres de régression. Poisson surdispersé.

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

Tableau 5 – Estimation et risque d'estimation des provisions, modèles Poisson surdispersé et Gamma.

Exercices	Poisson sur dispersé		Gamma	
	$\hat{E}(R_i)$	$\frac{\widehat{se}(\hat{R}_i)}{\hat{R}_i}$	$\hat{E}(R_i)$	$\frac{\widehat{se}(\hat{R}_i)}{\hat{R}_i}$
1	80	329%	101	49%
2	442	134%	494	33%
3	1631	70%	1286	29%
4	2811	53%	2793	28%
5	11786	32%	11262	30%
6	41864	20%	36702	35%
7	75137	31%	69563	48%
Total	133750	19%	122200	29%

Tableau 6 - Valeurs prévues, modèle Poisson surdispersé.

	0	1	2	3	4	5	6	7	Provisions
0	1401	4710	3306	942	305	378	182	67	
1	1663	5592	3925	1118	362	449	215	80	80
2	2490	8373	5877	1674	543	672	323	119	442
3	3645	12256	8602	2451	795	984	472	174	1631
4	4223	14199	9967	2839	921	1141	547	202	2811
5	8809	29621	20792	5923	1921	2379	1141	421	11786
6	11320	38064	26719	7611	2469	3057	1467	541	41864
7	10641	35782	25117	7155	2321	2874	1379	509	75137
							Total		133750

Tableau 7 – Facteurs de développement chain ladder

j	0	1	2	3	4	5	6
f_j	4,3627	1,5410	1,1000	1,0295	1,0355	1,0164	1,0060

6. DISTRIBUTIONS PREDICTIVES

Des caractéristiques de la loi de R , fondamentales pour la suite, ne peuvent pas être atteintes directement par la seule connaissance de ses moments : probabilité d'insuffisance ou quantiles (ou VaR). Il est donc nécessaire d'obtenir une estimation de la loi de R .

De manière exceptionnelle, celle-ci s'obtient directement par additivité de la loi générique des X_{ij} . Il en est ainsi pour la distribution Poisson surdispersée. Formellement si les X_{ij} suivent des lois Poisson (surdispersée) $P_{surd}(\mu_{ij}, \phi)$. il en est de même de R dont la distribution est $P_{surd}(\mu_R, \phi)$ avec $\mu_R = \sum_{i=1}^n \sum_{i+j>n} \mu_{ij}$ par $\frac{R}{\phi} \approx P\left(\frac{\mu_R}{\phi}\right)$. Dans le cas courant où

$\frac{\mu_R}{\phi} \geq 50$., on approxime la f.r. de la loi de Poisson par celle d'une loi Normale standard :

$$P(R \leq r) = \Phi\left(\frac{r - \mu_R}{\sqrt{\phi \mu_R}}\right)$$

et le quantile d'ordre $(1-\eta)$ de R : $q_{1-\eta}(R) = \mu_R + \sqrt{\phi \mu_R} q_{1-\eta}$ à l'aide du quantile de même ordre $q_{1-\eta}$ de cette loi standard.

Des e.m.v. ($\hat{\mu}_{ij}$) de (μ_{ij}) puis $\hat{\mu}_R$ de μ_R , on déduit les e.m.v. de la loi prédictive et des quantiles $q_{1-\eta}(R)$ de R :

$$\hat{P}(R \leq r) = \Phi\left(\frac{r - \hat{\mu}_R}{\sqrt{\hat{\phi} \hat{\mu}_R}}\right), \hat{q}_{1-\eta}(R) = \hat{\mu}_R + \sqrt{\hat{\phi} \hat{\mu}_R} q_{1-\eta}$$

Risque d'estimation et intervalle de confiance pour $q_{1-\eta}(R)$ asymptotiques s'obtiendraient, comme ci-dessus, par Normalité asymptotique.

Tableau 3 - Estimation Exemple : (suite) Dans le cas Poisson surdispersé, les estimations $\hat{\mu} = 133750$, $\sqrt{\hat{\phi}} = 26,762$, où $\hat{\phi}$ est l'estimation de la déviance, conduisent au tableau 8

Tableau 8 - Estimation des quantiles de la provision. Additivité.

$1-\eta$	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
$\widehat{q_{1-\eta}}(R)$	133750	140351	141987	146293	149849	156518

Dans le cas fréquent où la loi de R n'est pas directement accessible par additivité, on dispose de deux approches : l'approche bootstrap, basée sur le rééchantillonnage des résidus du modèle sélectionné, largement utilisée dans la pratique (cf England et al., 2002). La seconde, plus récente, fait appel à des distributions approchées basées sur les premiers moments estimés de R .

Celles-ci font appel aux méthodes d'approximation de distribution standard en sciences actuarielles : Normal Power, Esscher, Gamma translatée, Gamma – Bowers... (cf Partrat et al., 2005).

Elles conduisent à des approximations fermées de la f.r. F_R et de la probabilité d'insuffisance. La moitié d'entre elles fournit de plus une approximation analytique simple des quantiles de cette distribution. Pour l'autre moitié, ceux-ci sont solutions numériques d'équations réelles.

Ces quantités sont ensuite estimées par maximum de vraisemblance. La méthode Delta en donnerait risques d'estimation et intervalles de confiance asymptotiques.

Par $R = \sum_{i=1}^n \sum_{j=n-i+1}^n X_{ij}$ et l'indépendance des X_{ij} , les moments de R s'expriment en fonction des (μ_{ij}) (cf § 3.1) :

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

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

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

puis le coefficient d'asymétrie (skewness) de R :

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

Si nécessaire, on expliciterait les moments centrés d'ordre supérieur et le coefficient d'aplatissement (Kurtosis) de R : $\gamma_2 = \gamma_2(R) = \frac{\mu_4}{\sigma^4} - 3$.

La propriété d'invariance du maximum de vraisemblance donnerait l'emv $(\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma}_1)$ de (μ, σ, γ_1) , en remplaçant μ_{ij} par $\widehat{\mu}_{ij}$ dans les formules ci-dessus, puis l'e.m.v. de toute fonction réelle ou vectorielle de (μ, σ, γ_1) .

Par souci de simplicité, on se limitera, dans cet article, à l'approximation Normal Power (NP en abrégé). De semblables développements prendraient place pour les autres formules d'approximation mentionnées ci-dessus.

De manière directe, cette approximation donne, pour $\gamma_1 > 0$,

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

Par inversion de cette f.r., on déduit une valeur approchée du quantile de R d'ordre $1 - \eta$:

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

estimé (par maximum de vraisemblance) par $\widehat{q}_{1-\eta}^{(NP)} = \widehat{\mu} + \widehat{\sigma} \left[\frac{\widehat{\gamma}_1}{6} q_{1-\eta}^2 + q_{1-\eta} - 1 \right]$.

On en déduirait l'erreur standard estimée de $\widehat{q}_{1-\eta}^{(NP)}$ et un intervalle de confiance asymptotique pour $q_{1-\eta}^{(NP)}$.

Exemple : (suite) : Les tableaux 9 et 10 fournissent les quantiles estimés de R sous les modèles Poisson surdispersé et Gamma. Pour chacun d'eux, le paramètre ϕ est estimé par déviance.

1/ Modèle Poisson surdispersé

$$\widehat{\mu} = 133750; \sqrt{\widehat{\phi}} = 26,762; \widehat{\sigma} = \sqrt{\widehat{\phi}\widehat{\mu}} = 9787,36; \widehat{\gamma}_1 = \sqrt{\frac{\widehat{\phi}}{\widehat{\mu}}} = 0,073$$

Tableau 9 : Quantiles estimés, modèle Poisson surdispersé, approximation NP.

$1 - \eta$	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
$\widehat{q}_{1-\eta}(R)$	123963	130619	132284	136702	140385	147377

2/ Modèle Gamma

Nous obtenons les estimations $\widehat{\mu}, \widehat{\sigma}, \widehat{\gamma}_1$ suivantes pour les paramètres de la provision R :

$$\widehat{\mu} = 122192, \widehat{\sigma} = 21129, \widehat{\gamma}_1 = 0.4927 .$$

Tableau 10 : *Quantiles estimé, modèle Gamma, approximation NP.*

$1-\eta$	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
$\widehat{q_{1-\eta}^{(NP)}(R)}$	101063	116104	120074	130992	140513	159606

7. PROJECTIONS

7.1 Préliminaires

Le lien multiplicatif retenu, soit $E(X_{ij}) = e^{\mu+\alpha_i+\beta_j}$ ($i=1,\dots,n+m; j=0,\dots,n$), permet une interprétation simple et directe des paramètre α_i et β_j . En effet

$$S_i = \sum_{j=0}^n X_{ij} \text{ donne } E(S_i) = \sum_{j=0}^n E(X_{ij}) = \sum_{j=0}^n e^{\mu+\alpha_i+\beta_j}$$

puis

$$E(S_i) = e^{\mu+\alpha_i} \sum_{j=0}^n e^{\beta_j} = e^{\mu} \left(\sum_{j=0}^n e^{\beta_j} \right) e^{\alpha_i}$$

On en déduit

$$\ln E(S_i) = c + \alpha_i, \text{ en posant } c = \mu + \ln \left(\sum_{j=0}^n e^{\beta_j} \right)$$

La première égalité introduit une interprétation du paramètre α_i en terme de charge sinistres attendue $E(S_i)$.

A partir de l'e.m.v. $\hat{\xi}$ de ξ , déterminé sur le triangle supérieur, on obtient, par invariance fonctionnelle du maximum de vraisemblance, l'e.m.v. \hat{c} de c :

$$\hat{c} = \hat{\mu} + \ln \left(\sum_{j=0}^n e^{\hat{\beta}_j} \right)$$

D'autre part

$$e^{\beta_j} = \frac{e^{\mu+\alpha_i+\beta_j}}{e^{\mu+\alpha_i}} = \frac{E(X_{ij})}{E(S_i)} \left(\sum_{k=0}^n e^{\beta_k} \right)$$

conduit à

$$\frac{e^{\beta_j}}{\left(\sum_{k=0}^n e^{\beta_k} \right)} = \frac{E(X_{ij})}{E(S_i)}$$

le second terme de l'égalité étant la part des paiements attendus au cours du délai j dans la charge sinistres attendue de l'exercice i . Les β_j sont donc reliés à une cadence attendue de paiements.

On supposera, par la suite, les cadences de règlement stables au cours des exercices futurs, ce qui nous conduira à conserver l'estimation $(\hat{\beta}_j)_{0 \leq j \leq n}$ du vecteur $(\beta_j)_{0 \leq j \leq n}$. Il en sera de même de $\hat{\mu}$ et du paramètre de dispersion estimé $\hat{\phi}$.

7.2 Projection de triangle de liquidation

La projection du triangle au 31/12/ n pour former celui au 31/12/ $(n+1)$ s'opère par projection des charges sinistres attendues $[E(S_i)]_{0 \leq i \leq n}$ soit, par l'égalité $\ln E(S_i) = c + \alpha_i$, à l'aide des paramètres $(\alpha_i)_{0 \leq i \leq n}$.

La représentation graphique de la suite des estimations $(\hat{\alpha}_i)_{0 \leq i \leq n}$ conduit à sélectionner une fonction, paramétrée, $a(i, \gamma)$ susceptible d'ajuster cette suite. Ce choix est fait dans un souci de parsimonie, conduisant à concilier paramétrage minimum et bonne qualité d'ajustement. La méthode des moindres carrés, éventuellement pondérés pour privilégier les exercices récents (cf. annexe), conduit au meilleur ajustement $a(i, \hat{\gamma})$ de la famille choisie par minimisation en γ de

$$\sum_{i=0}^n \omega_i [\hat{\alpha}_i - a(i, \gamma)]^2$$

L'équation $\ln \hat{E}(S_i) = \hat{c} + \hat{\alpha}_i$ ($i = 0, \dots, n$) fournit l'ajustement équivalent pour la suite $[\hat{E}(S_i)]_{0 \leq i \leq n}$ des charges sinistres estimées. La projection de la fonction $a(i, \hat{\gamma})$ sur l'exercice $(n+1)$ en $\hat{\alpha}_{n+1} = a(n+1, \hat{\gamma})$ détermine la valeur prévue de la charge au cours de cet exercice :

$$\hat{E}(S_{n+1}) = e^{\hat{c} + \hat{\alpha}_{n+1}}$$

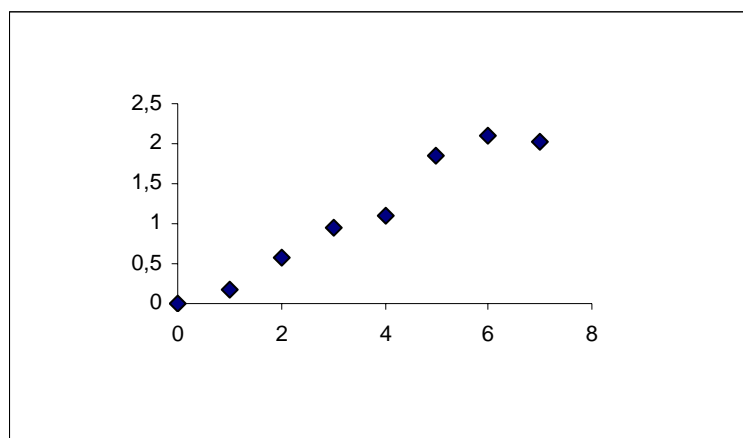
Remarque : Comme $\hat{\alpha}_{n+1}$ est fonction de $(\hat{\alpha}_i)_{0 \leq i \leq n}$ et \hat{c} de $[\hat{\mu}, (\hat{\beta}_j)_{0 \leq j \leq n}]$, $\hat{E}(S_{n+1})$ est une fonction régulière de $\hat{\xi}$. Il en résulte sa Normalité asymptotique et la possibilité d'obtenir, comme ci-dessus, risque d'estimation de $\hat{E}(S_{n+1})$ et intervalle de confiance pour $E(S_{n+1})$.

Exemple : Le graphique ci-dessous représente la suite des $(\hat{\alpha}_i)_{0 \leq i \leq n}$ dans le modèle Poisson surdispersé, On observe une haute qualité globale d'ajustement ($R^2=0,96$) par une droite de régression, plus volatile pour les valeurs élevées de i . La modélisation Gamma conduirait à des résultats identiques.

Un test de sensibilité au choix de la fonction d'ajustement mobiliserait les fonctions quadratiques.

Compte tenu de la position relative du point ultime $\hat{\alpha}_7$, nous avons choisi de lui appliquer une forte pondération dans la suite des poids (0,5 ; 1 ; 1,5 ; 2 ; 2,5 ; 3 ; 10 ; 1000) que nous utilisons. Il en résulte $\hat{\alpha}_8 = 2,2905$ puis $\hat{E}(S_8) = 111542$ par $\hat{c} = 9,3317$. On peut noter qu'une moindre attention portée au point ultime par une pondération uniforme donnerait $\hat{\alpha}_8 = 2,5805$ et $\hat{E}(S_8) = 149072$.

Graphique 1 : Représentation de la suite $(\hat{\alpha}_i)_{0 \leq i \leq n}$ et droite ajustée, modèle Poisson surdispersé.



En revenant au cas général, le triangle de liquidation au $31/12/(n+1)$ se déduit du triangle précédent de la manière suivante.

Les délais de règlement variant toujours de 0 à n et les seuls exercices d'origine concernés étant ceux de 1 à $(n+1)$, au sous-triangle des règlements $(x_{ij})_{\substack{i=1,\dots,n \\ i+j \leq n}}$ avant le $31/12/n$, on adjoint la diagonale des paiements $(x_{ij})_{\substack{i=1,\dots,n+1 \\ i+j=n+1}}$ de l'année calendaire $(n+1)$. Ces $x_{ij}(i+j=n+1)$ sont des réalisations simulées des v.a.r. X_{ij} distribuées selon la loi choisie pour la modélisation (Poisson surdispersée ou Gamma dans notre exemple) avec paramétrage $(\hat{\mu}, \hat{\alpha}_i, \hat{\beta}_{n+1-i}, \hat{\phi})$.

Ce processus pourrait être renouvelé, à partir de ce triangle au $31/12/(n+1)$, jusqu'à atteindre l'horizon $(n+m)$ choisi.

7.3 Projection de résultats techniques, besoin en capital

L'assureur dispose d'une mesure de « suffisance » de provision $PS(R)$, lui fournissant la provision à constituer pour tout engagement, aléatoire, de règlements futurs de sinistralité R . Cette mesure est un paramètre de la f.r. de R , choisi selon la finalité, économique ou réglementaire, de provisionnement.

Ce pourrait être $E(R), E(R) + \gamma\sigma(R), VaR_\varepsilon(R) = q_{1-\varepsilon}(R)$ pour ε donné (cf. Australian Prudential Regulation Authority, 2001).

L'assureur s'est, d'autre part, fixé un principe de calcul de primes H associant à toute sinistralité future S , de f.r. F_S , la *prime de risque* qu'il demanderait pour couvrir ce risque :

$$\Pi(S) = H(F_S)$$

H peut ne dépendre que des moments de S ou nécessiter la connaissance de la distribution complète de S (principe PH-transform par exemple).

Le résultat technique simplifié (hors produits financiers, frais d'acquisition et de gestion,...) de l'exercice $(n+1)$ en brut de réassurance est déterminé par les éléments suivants :

- *La prime de risque estimée au $1/1/(n+1)$ pour l'exercice $(n+1)$*

Soit $F_{S_{n+1}}$ la distribution prédictive de la sinistralité de cet exercice

$$S_{n+1} = \sum_{j=0}^n X_{n+1,j}$$

Cette distribution se déduit de celles des $X_{n+1,j}$, sous l'hypothèse de leurs indépendances. On note que la loi prédictive estimée de cette v.a.r. $X_{n+1,j}$ correspond à la modélisation générale choisie, à laquelle on applique le paramétrage $(\hat{\mu}, \hat{\alpha}_{n+1}, \hat{\beta}_j, \hat{\phi})$.

La prime de risque estimée $\hat{\Pi}_{n+1} = H(\hat{F}_{S_{n+1}})$ est obtenue directement si le principe $H(S)$ ne dépend que des moments de S , à l'aide de l'approximation NP de F_S , comme ci-dessus, sinon.

- *La sinistralité de l'exercice $(n+1)$*

Celle-ci s'écrit $x_{n+1,0} + PS\left(\sum_{j=1}^n X_{n+1,j}\right)$, où $x_{n+1,0}$ est la réalisation simulée de la v.a.r. $X_{n+1,0}$ à paramétrage $(\hat{\mu}, \hat{\alpha}_{n+1}, \hat{\beta}_0, \hat{\phi})$.

- *Le delta de provisions au titre des exercices 1 à n*

Cette différence s'écrit

$$\Delta = PS_{1,\dots,n}^{(n)} - PS_{1,\dots,n}^{(n+1)} - \sum_{i=1}^n x_{i,n-i+1}$$

où

$$PS_{1,\dots,n}^{(n)} = PS\left(\sum_{i=1}^n \sum_{j=n-i+1}^n X_{ij}\right) \text{ et } PS_{1,\dots,n}^{(n+1)} = PS\left(\sum_{i=2}^n \sum_{j=n-i+2}^n X_{ij}\right)$$

sont respectivement les provisions « suffisantes », au titre des exercices 1 à n , au $31/12/n$ et au $31/12/(n+1)$.

$\sum_{i=1}^n x_{i,n-i+1}$ représentent les paiements simulés de l'exercice $(n+1)$.

Le besoin en capital au 31/12/ n de l' assureur, qui a constitué une provision « suffisante » au titre des sinistres des exercices 1 à n , est déterminé par des simulations répétées de résultats de l'exercice ($n+1$) et application d'un critère de capital « suffisant », selon que la finalité est réglementaire ou économique.

8. CONCLUSION

Dans le cadre standard des Modèles Linéaires Généralisés, qui ont comme cas particulier la méthode chain ladder, cet article a présenté un modèle stochastique de projection de triangles direct et flexible, ne nécessitant pas la mise en œuvre du bootstrap. Il permet d'obtenir simplement des projections de résultats techniques ainsi que des évaluations de besoins en capital.

Il serait aisément généralisable à des cadences de règlements variant, de manière dynamique, en fonction des exercices. Il pourrait intégrer une approche fréquence-coûts de sinistralité par le modèle composé de Tweedie.

REFERENCES

- Daykin C.D., Pentikainen T., Pesonen M. (1994) : *Practical risk theory for actuaries*. Chapman & Hall.
- Denuit M., Charpentier A. (2004) : *Mathématiques de l'assurance non-vie*, Tome I, 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. (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.
- Kaufmann R., Gadmer A., Klett R. (2001) : Introduction to Dynamic Financial Analysis. *Astin Bull.* Vol.31, 213-251.
- Klugman S.A., Panjer H.H., Willmot G.E. (1998) : *Loss models : from data to decisions*. Wiley.
- Mack T. (1993) : "Distribution free calculation of the standard error of Chain ladder reserve estimates" *Astin Bull.* Vol. 23, 213-225.
- McCullagh P., Nelder J. (1989) : *Generalized Linear Models*, 2nd ed. Chapman & Hall
- 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*. Collection AAA, Economica.
- Partrat C., Braillard P., Lecoeur E., Nisipasu E., Nessi J.-M., Reiz O.: (2006) : *Evaluation des provisions techniques en assurance non-vie*. Collection AAA, Economica.

- Renshaw A. E., Verrall R. J. (1998) : "A stochastic model underlying the chain ladder technique". *British Act. J.* Vol. 4. 903-923
- Shao J. (1999) : *Mathematical statistics*. Springer.
- Taylor G. (2000) : *Loss reserving : an actuarial perspective*. Kluwer Acad. Press.
- 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.

ANNEXE

L'ajustement de la droite $y=a+bx$ aux points $(x_k, y_k)_{k=0, \dots, n}$, avec pondération (ω_k) , par minimisation de

$$\sum_{k=0}^n \omega_k [y_k - (a + bx_k)]^2$$

donne, en posant $\bar{x}_p = \frac{\sum_k \omega_k x_k}{\sum_k \omega_k}$, $\bar{y}_p = \frac{\sum_k \omega_k y_k}{\sum_k \omega_k}$, $\sigma_p^2 = \frac{\sum_k \omega_k x_k^2}{\sum_k \omega_k} - \bar{x}_p^2$,

$$\hat{b} = \frac{\frac{\sum_k \omega_k x_k y_k}{\sum_k \omega_k} - \bar{x}_p \bar{y}_p}{\sigma_p^2}, \quad \hat{a} = \bar{y}_p - \hat{b} \bar{x}_p$$