

PROVISIONNEMENT POUR SINISTRES A PAYER :
ANALYSES ET MODELISATIONS SUR DONNEES
DETAILLEES

Magali KELLE
Actuaire CEA
Chargée d'études à la Matmut

Ce travail est extrait d'un mémoire d'actuariat co-dirigé par Gilbert Macquart et Christian Partrat. La disparition de ce dernier laisse un grand vide dans nos travaux et nos cœurs...

ABSTRACT

The weight and the volatility of the outstanding claims liabilities in cars insurance make actuaries develop various statistical methods. Convergence of such methods should enlighten about loss reserves amount. But hypothesis aren't always checked and results differ one another. A few claims data processing can reduce difference and increase account accuracy.

This study deals with the integration of administrative and management data in modelisation process in order to detect instable claim valuation. In this way, we observe individual corporeal claims and forecast of outstanding liability in respect of a single claim. We realize these estimations at many times by general linear model or deterministic rules. And then we compare estimates made by assessor with modelisation forecasts. When an important gap is observed assessor valuation is refused and claim is reconsidered.

This method provides also an estimate of aggregate loss reserves, which makes it comparable to some aggregate statistical methods. But this is not our aim. Indeed our purpose is to obtain an individual reliability guide of assessor valuations that include some contextual claim data.

RESUME

Lors de la clôture des comptes, la charge de la sinistralité doit être évaluée. Pourtant les gestionnaires manquent parfois d'information pour anticiper le coût ultime de certains

dossiers sinistres à long déroulement. L'estimation des provisions pour sinistre à payer nécessite alors le déploiement d'un faisceau de méthodes, la convergence de ces méthodes devant éclairer le praticien sur le niveau de charge probable à attendre. Mais les données ne respectent pas toujours les hypothèses sur lesquelles s'appuient les méthodes et peuvent engendrer des résultats divergents.

Les coûts de dossiers sinistres corporels, en garantie de responsabilité civile automobile, peuvent s'avérer élevés après plusieurs années d'attente d'information. Dans ces circonstances les évaluations dossier-dossier sont soumises à de fortes variabilités et perdent en fiabilité. Ces dossiers, de valorisation instable, perturbent alors les méthodes d'estimation. C'est pourquoi nous cherchons à les détecter en intégrant l'historique des actes de gestion dans des méthodes de tri de dossiers aussi industrialisables que possible.

Pour cela, nous observons des coûts ultimes individuels et nous construisons un processus de modélisation de ce coût. En fonction de la maturité et du type de dossier sinistre considéré, des règles déterministes ou des modélisations sont appliquées. L'estimation obtenue est ensuite comparée à la valorisation dossier-dossier afin de conclure à l'acceptabilité de l'évaluation fournie par la gestion. Les dossiers refusés seront prioritairement réétudiés pendant la phase d'inventaire.

Cette méthode fournit également une charge ultime agrégée qui peut être comparée aux méthodes classiques sur triangles de données. Toutefois, notre principale préoccupation est l'obtention, au niveau individuel, d'un indicateur de cohérence de la provision dossier-dossier intégrant le contexte du sinistre.

Mots clés : provisionnement, données détaillées, préparation de données.

1. INTRODUCTION

Cette étude traite de la prise en compte des données détaillées lors des estimations de Provisions pour Sinistres A Payer (PSAP). En assurance automobile, ces estimations s'appuient, dans la plupart des cas, sur les évaluations dossier-dossier et s'accompagnent d'un faisceau de méthodes statistiques permettant de confirmer ou d'infirmer les résultats dossier-dossier.

Les méthodes statistiques présupposent des hypothèses de stabilité temporelle relativement lourdes. Ces méthodes sont d'autant plus pertinentes que :

- L'historique de données est important et fiable,
- Le passé, le présent et le futur sont structurellement proches,

- Le passé est régulier et peu volatile.

Pour vérifier ces hypothèses, il est courant d'actualiser les données et d'exclure les sinistres catastrophiques ou exceptionnels. Ce traitement ne suffit pas toujours à régulariser les données. On peut alors avoir recours à des facteurs externes, comme dans la méthode de Bornhuetter-Ferguson, pour introduire un peu de régularité. Le développement des méthodes stochastiques montre, en outre, l'intérêt porté aux estimations de variance des résultats fournis.

Toutefois, pour appliquer ces méthodes de façon optimale la connaissance des données, de l'historique des principales modifications de gestion (souscription, indemnisation) et des tendances externes (évolutions de sinistralité ou réglementation) est indispensable [IA]. Reste à savoir comment exploiter cette connaissance pour maîtriser les données et leur contexte. Cette question a inspiré nos travaux.

Cette étude se situe donc en amont des problématiques d'estimation de provisionnement puisque nous cherchons à préparer les données qui alimentent les modèles d'estimation des provisions. Plus concrètement, nous avons mis en place un processus de détection de sinistres de valorisation potentiellement incongrue. La détection de ces dossiers instables, permet de les réviser prioritairement lors des périodes d'inventaire et de fiabiliser les données.

Cette problématique peut être abordée de multiples manières. Après plusieurs essais, nous avons choisi de travailler en 2 étapes :

- Dans une première étape, nous construisons une évaluation de référence à chaque date d'étude et pour chaque sinistre en fonction du contenu du dossier et son historique de gestion.
- Dans une deuxième étape, nous comparons l'évaluation gestionnaire avec cette évaluation de référence. La comparaison nous permet de détecter les dossiers à réviser.

L'évaluation dossier-dossier étant une anticipation du coût ultime du dossier, notre évaluation de référence doit elle aussi être une approche de ce coût. Nous cherchons donc, dans la première étape, à construire une sorte d'estimateur du coût ultime d'un dossier.

2. PERIMETRE DE L'ETUDE

Nous nous intéressons plus particulièrement à la garantie responsabilité civile (RC) des sinistres automobiles ayant eu des conséquences corporelles. Notre choix s'est porté sur

cette garantie puisque, comme l'ensemble des garanties à longue durée de vie, elle pose des difficultés d'anticipation du coût ultime.

Afin de disposer de suffisamment de recul pour observer puis construire un coût ultime, nous nous appuyons sur les sinistres survenus en 1998 que nous observons jusqu'en décembre 2003.

3. MISE EN FORME DES DONNEES

Dans ce travail, les bases de données de l'entreprise, bases de gestion sinistre ou de gestion contrat, sont exploitées pour construire un modèle de données exploitable statistiquement. Ainsi, nous obtenons des bases de données donnant à **chaque instant** une vision globale et synthétique du sinistre en introduisant des variables qualitatives ou quantitatives représentant les actes de gestion (paiements, évaluations, courriers, expertises...).

La notion d'instant est définie par la maturité d'un dossier. La maturité s'apparente au nombre de mois écoulés depuis l'ouverture informatique du dossier (si on ne considère que des mois de 30 jours). On conserve une vision du sinistre de la maturité 1 à la maturité 61. Le choix de la maturité 61 vient du fait que les sinistres sont observés jusqu'au 31 décembre 2003, ce qui fait un maximum de 61 mois d'observation pour les dossiers ouverts en décembre 1998. Pour avoir le même recul sur tous les dossiers, l'observation est interrompue après 61 mois.

Sur l'échantillon d'étude, on remarque qu'après 61 mois de vie les dossiers ouverts représentent 17% des nombres et 64% des évaluations. Sur ces dossiers, toujours en cours, nous sommes obligés de faire une hypothèse : l'évaluation ultime est approximée par l'évaluation à 61 mois de vie. Cette hypothèse peut paraître forte mais elle semble vérifiée sur des dossiers d'exercices antérieurs à 1998 : pour ces exercices, l'évaluation dossier-dossier à 61 mois est une vision réaliste de l'évaluation ultime.

**Par la suite, nous confondons le coût ultime d'un dossier et sa valeur à 61 mois.
Cette valeur est notée V_{ult} .**

4. ETUDE PRELIMINAIRE

La notion de maturité permet de comparer les dossiers avec le même recul de gestion. Toutefois, elle écrase les caractéristiques calendaires (tel que le caractère tardif des dossiers ou les délais de déclaration). Dans une étude préliminaire, les caractéristiques

calendaires ont été étudiées de façon à mesurer leur pouvoir prédictif sur le coût ultime des dossiers. Lorsqu'une de ces variables possède un caractère prédictif, elle est réintégrée spécifiquement dans le modèle de données.

A titre d'exemple, nous présentons un extrait des travaux effectués sur le délai de déclaration. Dans cette partie, nous appliquons des tests de rang pour voir si le délai influe sur la valorisation ultime des sinistres. Pour cela, nous construisons une variable booléenne D telle que :

$$\begin{cases} D = 1 & \text{s'il existe un retard de déclaration} \Leftrightarrow \text{si délai} > 5 \text{ jours } ^1 \\ D = 0 & \text{sinon} \end{cases}$$

Les résultats fournis par le test de Wilcoxon et de la Médiane sont donnés ci-dessous :

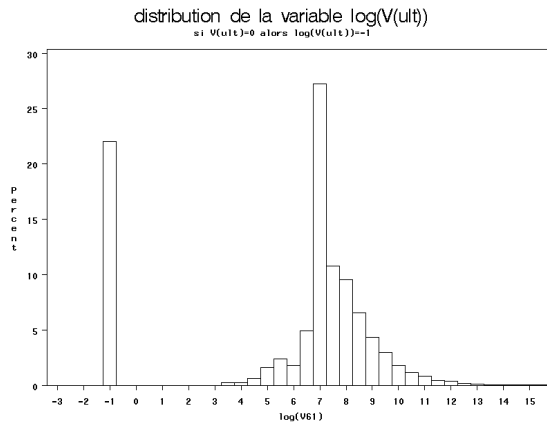
	Statistique observée	Statistique attendue	P(attendue > observée)
Wilcoxon	1 116 835	1 070 560	2.48%
Médiane	398	374	2.24%

Ces résultats montrent que, pour un seuil de test $\alpha \leq 2.2\%$, l'hypothèse nulle est acceptée, l'homogénéité des échantillons est donc acceptée, soit la non influence de D sur la distribution de $V_{ult.}$. Ce seuil de 2.2% ne permet pas de conclure avec certitude sur un éventuel lien entre D et $V_{ult.}$. Ceci étant, la variable D devra être testée comme variable explicative lors de la modélisation de $V_{ult.}$

5. ETUDE DESCRIPTIVE DE L'EVALUATION ULTIME

Pour simplifier les visualisations, nous nous intéressons au logarithme de l'évaluation ultime. On représente ci-après la distribution du logarithme de la valorisation ultime des dossiers.

¹ Ce délai correspond, en général, au délai légal de déclaration tel qu'il apparaît dans les conditions générales des contrats : Le délai légal est de 2 jours ouvrés pour les vols et de 5 jours ouvrés pour les autres types de déclaration.



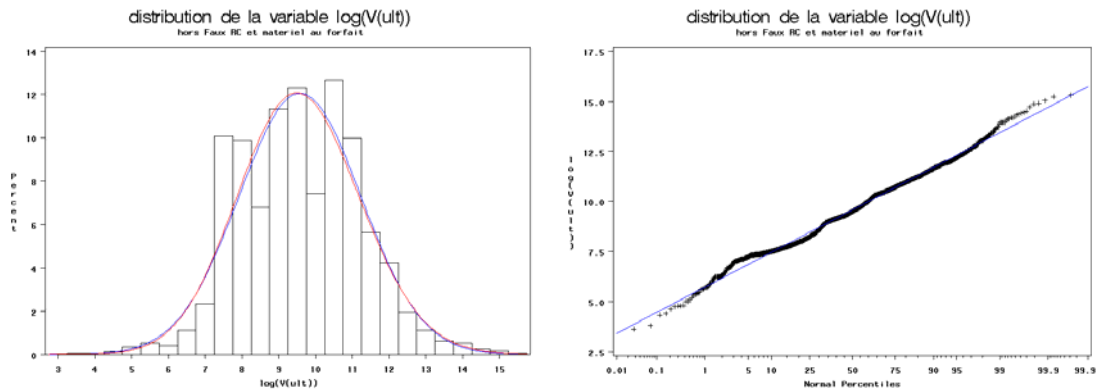
On remarque tout de suite deux pics de dossier :

- Un pic, placé en -1, correspondant aux dossiers de valorisation ultime nulle pour lesquels on a volontairement initialisé le logarithme à -1. Ce sont des dossiers ouverts à tort dans la garantie RC et dans lesquels on n'indemniserait pas au titre de cette garantie.
- Un pic, aux alentours de 7, correspondant aux sinistres purement matériels pour la garantie RC pour lesquels l'évaluation ultime correspond au versement du forfait CGIRSA¹ de 1998 soit 1 036 €.

Ce 2^{ème} pic résulte d'une des particularités du système d'information, à savoir que la notion de corporel est au niveau du dossier et non de la garantie. Ainsi, parmi les dossiers RC ayant eu des conséquences corporelles, on retrouve des dossiers entièrement matériels au titre de la garantie RC.

Si nous excluons ces deux familles de dossiers, nous obtenons une distribution plus homogène. Cette distribution regroupe l'ensemble des dossiers corporels et/ou matériels hors forfait. Cette distribution est présentée ci-dessous à gauche.

¹ La CGIRSA est une convention de règlement matériel entre assureurs telle que la plupart des dommages matériels se soldent par un paiement de forfait.



Il semble possible d'ajuster convenablement une distribution gaussienne ou gamma (les deux courbes se superposent) sur cet ensemble de dossier. Et si les tests refusent l'hypothèse de normalité du logarithme de la valorisation ultime, la droite d'Henry (figure de droite) montre que cette hypothèse est tout à fait plausible.

Ainsi, nous avons mis en évidence l'existence de 3 familles de dossiers dans l'échantillon :

- Les « Faux RC » dont la valorisation ultime est nulle, ils ont été valorisés à tort dans cette garantie.
- Les « matériels au forfait », qui sont purement matériels dans la garantie RC et dont la valorisation ultime est égale à 1 036 €.
- Les « Vrais RC » qui regroupent des dossiers matériels et/ou corporels dont le logarithme de la valorisation ultime peut être assimilé à une loi normale ou gamma.

Ainsi, à chaque maturité, lorsque l'appartenance d'un dossier à l'une des deux premières familles est vérifiée, l'évaluation ultime de référence s'en déduit (0 ou 1 036€). L'appartenance à la 3^{ème} famille ne donne pas directement l'évaluation de référence d'un dossier. Toutefois, au vu de la distribution de coût dans cette famille, il semble concevable d'y chercher une modélisation de la valorisation de référence.

Ce constat a déterminé la suite de l'étude puisque, dans un premier temps, nous cherchons à déterminer des règles d'appartenance à chacune des familles puis, dans une deuxième étape, nous nous intéressons aux dossiers de la 3^{ème} famille.

6. REGLES D'APPARTENANCE AUX TROIS FAMILLES

Après plusieurs tests, nous avons choisi d'utiliser les règles suivantes :

A une maturité donnée,

- Les « Faux RC » correspondent aux dossiers :
 - o sans paiement dans la garantie RC,
 - o valorisés par le gestionnaire à moins de 2 000€.
- Les « RC matériels au forfait » correspondent aux dossiers :
 - o n'appartenant pas au faux RC,
 - o sans paiement corporel dans la garantie RC,
 - o dans lesquels la variable « corporel grave » est à 0¹
 - o dans lesquels la variable « responsabilité sociétaria » est à 4,
 - o valorisés par le gestionnaire entre 1 000 et 2 000€.
- Les « Vrais RC » (corporels et/ou matériels) regroupent tous les autres dossiers.

En appliquant ces règles, on obtient les taux de reconnaissance suivants :

Type du dossier a posteriori	Taux de reconnaissance		
	3 mois	6 mois	12 mois
Faux RC	85%	86%	88%
RC Matériel Forfaitaire	35%	69%	78%
RC Corporel ou Matériel	87%	95%	97%

A partir de 6 mois, les taux de reconnaissance de ces 3 catégories sont acceptables.

7. RECHERCHE D'UNE VALORISATION DE REFERENCE SUR LA 3EME FAMILLE

La 3^{ème} famille, pour laquelle nous ne disposons pas directement de valorisation de référence, paraît assez hétérogène. En effet, elle regroupe des dossiers matériels et corporels. Malgré nos efforts, nous n'avons pu obtenir de règles a priori permettant de distinguer les dossiers matériels des dossiers corporels. Sur cette troisième famille, nous cherchons une valorisation de référence en modélisant le logarithme de la valorisation ultime en fonction du contenu des dossiers.

Ce contenu évolue au cours du temps et la valorisation de référence doit tenir

¹ Cette variable est un booléen répondant à l'obligation légale de déclarer les sinistres pouvant entraîner un décès, une hospitalisation ≥ 1 mois, une ITT ≥ 6 mois, une IPP $> 25\%$ ou un cumul IPP $\geq 40\%$. La modalité 0 signifie que le dossier ne rentre pas dans le cadre de cette obligation légale.

compte de cette évolution. C'est pourquoi nous avons tout d'abord tenté d'intégrer une dimension temporelle aux modélisations. Ceci n'a pas donné de résultats satisfaisants et nous avons finalement opté pour des modélisations conditionnées par la maturité des dossiers.

Nous avons créé 12 modèles indépendants à chaque maturité ayant pour variable d'analyse le logarithme de la valorisation ultime et pour variable explicative le contenu du dossier à la maturité choisie. Les modèles sont espacés de 3 mois, puisque c'est, semble-t-il, le délai de mise à jour des dossiers corporels.

Le premier modèle correspond aux dossiers de 6 mois de maturité (avant 6 mois le contenu ne semble pas significatif). Le dernier modèle se base sur les dossiers de maturité 42 mois car, au-delà, nous aboutissons à un modèle qui a pour seule variable explicative, la valorisation gestionnaire, qui correspond souvent à la variable à expliquer.

Plusieurs essais de modèles linéaires généralisés ont été réalisés avec des distributions Gamma ou Normale et différentes fonctions de lien. Pour comparer les modèles et arrêter notre choix, nous utilisons le critère AIC, une analyse des résidus et des influences ainsi qu'un contrôle de robustesse à l'aide d'un échantillon test. Ce croisement de critères révèle qu'un modèle linéaire standard comportant 3 variables explicatives est le plus approprié.

Le modèle retenu à une maturité t est : $LV(ult) = X_t \beta_t + \varepsilon_t$ avec

- $LV(ult)$, le logarithme de la valorisation ultime, vecteur des n variables à expliquer,
- X_t la matrice $n \times p$ des variables explicatives $X_t = (X_{ik})_{i=1 \dots n, k=1 \dots p}$.
- et une collection de n variables aléatoires de loi Normale $N(0, \sigma_t^2)$
- β_t un vecteur de p paramètres.

Après plusieurs essais de modèles, les variables explicatives X_t retenues sont les mêmes pour toutes les maturités étudiées. Ces variables sont :

- Le logarithme des valorisations gestionnaires à la maturité t , intégré en variable continue par morceau : $[-\infty ; \log(0.5M)]$, $[\log(0.5M) ; \log(1M)]$, $[\log(1M) ; +\infty]$,
- Le code barème CGIRSA¹,

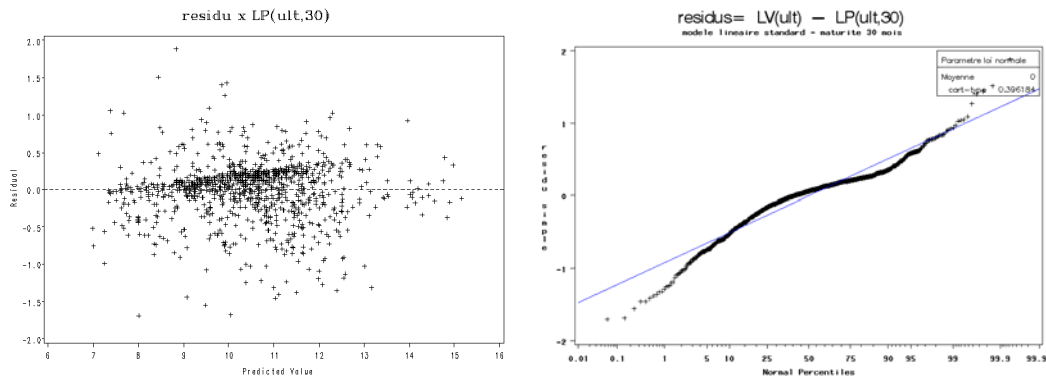
¹ Cette variable comporte à l'origine 14 modalités (13 codes CGIRSA et un code manquant) mais pour accroître son pouvoir prédictif, elle a été regroupée. Ainsi, on ne conserve que 2 modalités (circulation dans le même sens, autres cas). La circulation dans le même sens est, en règle générale, un indicateur de faibles dommages.

- Le nombre de réévaluations (positives) intégré dans le modèle comme variable continue.¹

Les paramètres associés à ces variables explicatives et l'estimateur de l'écart type sont obtenus par maximum de vraisemblance. On obtient ainsi 12 vecteurs de paramètres permettant d'attribuer à chaque dossier sinistre, i , et à chaque maturité, t , une estimation du logarithme de la valorisation ultime :

$$\hat{E}(LV(ult))_{it} = LP(ult, t)_i \text{ avec } t \in \{6; 9; 12; 15; 18; 21; 24; 27; 30; 33; 36; 39; 42\}.$$

On présente ci-après le graphe des résidus et leur droite d'Henry pour le modèle à 30 mois de maturité.



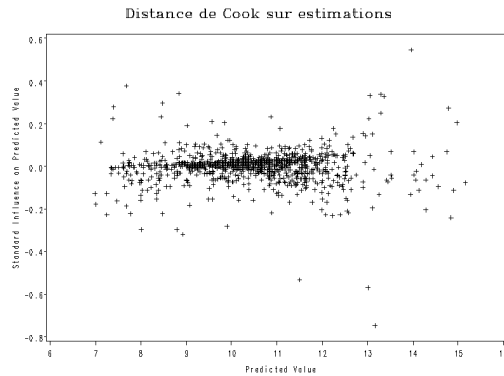
Le graphe (résidus ; estimation) ne révèle pas de structure cachée flagrante. Il est possible de repérer une légère tendance qui pourrait provenir de variances inégales ou d'un terme manquant mais celle-ci demeure faiblement marquée.

La droite d'Henry sur la distribution des résidus montre des queues de distribution un peu lourdes notamment à gauche. Ceci remet en cause l'hypothèse de normalité des résidus. La lourdeur constatée à gauche traduit une tendance à sur représenter les résidus de faibles valeurs, ce qui nous convient.

On s'intéresse à l'influence des observations sur la valeur estimée. Pour cela, on

¹ Cette variable comporte à l'origine 14 modalités (13 codes CGIRSA et un code manquant) mais pour accroître son pouvoir prédictif, elle a été regroupée. Ainsi, on ne conserve que 2 modalités (circulation dans le même sens, autres cas). La circulation dans le même sens est, en règle générale, un indicateur de faibles dommages.

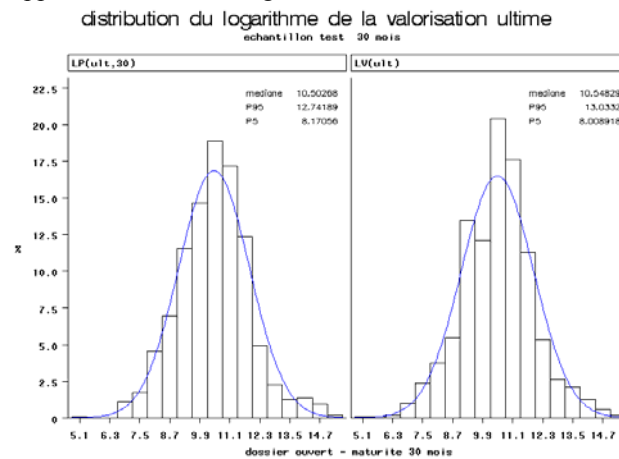
présente la distance de Cook¹ en fonction des estimations précédentes.



Le graphe ci-contre représente la distance de Cook en fonction des estimations. Ces mesures paraissent tout à fait convenables puisque faibles et proches de 0. On remarque toutefois une observation qui conserve une influence relativement forte bien que sa distance de Cook demeure faible. Nous conservons cette observation dans l'échantillon puisqu'elle ne nous paraît pas aberrante.

Une étude plus précise de l'influence sur les paramètres confirme ces résultats. La plupart des valeurs sont largement inférieures à 0.15. Le même dossier sinistre ressort avec une influence de 0.47 pour le paramètre des évaluations gestionnaires supérieur à 1 M€.

Pour valider la robustesse du modèle, les paramètres obtenus sont appliqués à un ensemble de dossiers n'ayant pas participé à la modélisation. Cet ensemble est appelé échantillon test. On présente ci-après les distributions à 30 mois issus de l'échantillon test. La distribution modélisée ($LP(ult, 30)$) est proche de la distribution de coût ultime observée ($LV(ult)$), le modèle apparaît donc cohérent globalement.



¹ La distance de Cook est : $D_i = \frac{(\hat{\beta} - \hat{\beta}_{(-i)})^T X^T X (\hat{\beta} - \hat{\beta}_{(-i)})}{(p+1)\hat{\sigma}^2}$ où $\hat{\beta}_{(-i)}$ est le vecteur de paramètre estimé à partir de l'échantillon privé de la $i^{ème}$ observation.

Cette étude de distribution permet de conclure que, globalement, les modélisations sont satisfaisantes. Toutefois, nous souhaitons utiliser le modèle de façon détaillée pour détecter les sinistres de valorisation incongrue et nous ne pouvons pas nous contenter d'une validation globale. Pour valider le modèle au niveau individuel nous utilisons un critère classique : l'erreur relative.

Avant de construire des indicateurs d'erreur relative nous abandonnons le logarithme et repassons à la valorisation ultime qui semble plus parlante. Puisque l'on sait que :

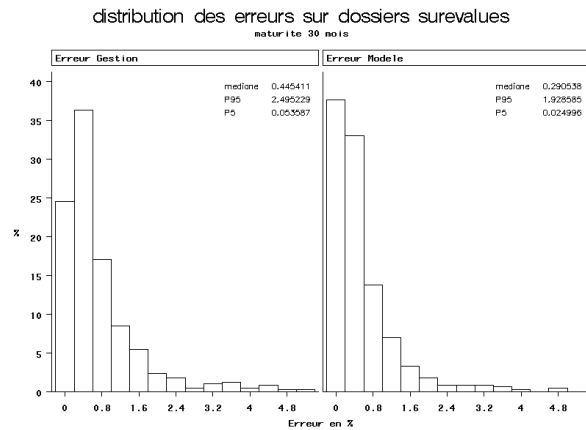
$$\begin{cases} \text{si } L V_{ult} \text{ suit une } N(\mu, \sigma^2) \\ \text{alors } E(V_{ult}) = \exp\left(\mu + \frac{\sigma^2}{2}\right) \end{cases}$$

Les propriétés de convergence des estimateurs du maximum de vraisemblance permettent de prendre comme estimateur de $E(V_{ult})_{it}$: $\hat{E}(V_{ult})_{it} = \exp(\hat{\mu}_{it} + \frac{\hat{\sigma}_{it}^2}{2})$ qui est asymptotiquement sans biais. Les travaux de Simar T. et Paris J. ([SIMAR 1996]) permettent de montrer que, dans notre étude, le biais est inférieur à 5% de l'estimation obtenue dans 95% des cas. Ce résultat nous paraissant satisfaisant nous ne cherchons pas à débiaiser cet estimateur.

En utilisant cet estimateur de l'évaluation ultime, nous calculons les erreurs relatives faites par les gestionnaires et par les modèles. En notant respectivement V_t et P_t , l'évaluation dossier-dossier et celle fournie par le modèle à la maturité t , on a :

$$\text{Erreur Gestion} = \frac{V_{ult} - V_t}{V_{ult}} \quad \text{Erreur de Modèle} = \frac{V_{ult} - P_t}{V_{ult}}$$

Ces erreurs peuvent être positives ou négatives. Ainsi, pour que les comparaisons aient un sens nous travaillons par catégorie de dossiers : surévalués par la gestion et surévalués par le modèle, sous évalués par la gestion et surévalués par le modèle, etc.



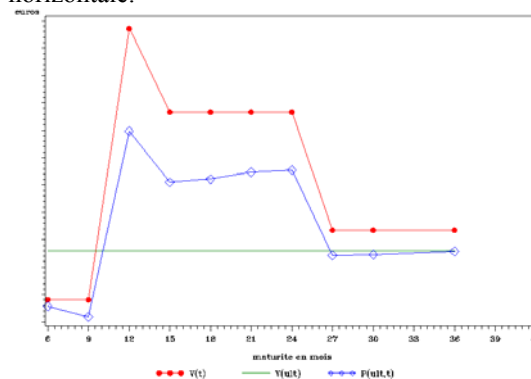
On retrouve ci-contre la distribution des erreurs sur les dossiers de maturité 30 mois surevalués par le modèle et par la gestion. Ces dossiers représentent 50% de l'échantillon.

On remarque que, sur cette catégorie de dossiers, les erreurs du modèle (histogramme de droite sur le graphique) sont plus faibles que les erreurs de gestion, la distribution de l'erreur du modèle est décalée à gauche).

Les constats effectués sur cette catégorie de dossiers se retrouvent sur les 3 autres catégories. Les modélisations apparaissent donc exploitables au niveau détaillé.

8. IMPACT SUR UNE TRAJECTOIRE D'ÉVALUATION

Avant de voir comment ces modélisations permettent de détecter les dossiers de valorisation potentiellement incongrue, nous proposons de suivre l'évolution du coût ultime d'un sinistre. Ce sinistre a été choisi pour la représentativité des trajectoires d'évaluation. Sur le graphique ci-dessous, l'évaluation fournie par les gestionnaires est représentée par des points, celle issue des modélisations par des losanges et le coût ultime par la trajectoire horizontale.



On remarque, sur cet exemple, que les modélisations corrigent une partie de la surestimation gestionnaire tout en conservant un niveau de prudence. Les deux trajectoires ont la même allure générale. Ceci provient du fait que l'évaluation gestionnaire est l'une des trois variables explicatives des modélisations.

9. DETECTION DES DOSSIERS A REVISER EN PRIORITE

Ainsi, on a montré que pour les vrais dossiers RC, les modèles permettent de construire une évaluation de référence en fonction de la maturité du dossier. Pour les autres

familles, nous attribuons une valeur de référence déterministe (0 et 1036€) et les taux de reconnaissance obtenus garantissent un faible niveau d'erreur. Nous disposons donc d'une évaluation de référence pour chaque dossier et à chaque maturité. Cette évaluation de référence permettrait de détecter les valorisations dossier-dossier potentiellement incongrues par comparaison. Dans cette partie, nous exposons les outils de comparaison retenus.

Pour les « Vrais dossiers RC », les modélisations nous permettent d'obtenir des intervalles de confiance en appliquant la méthode Delta rappelée ci-dessous.

Si $\hat{\eta} = (\hat{\eta}_1, \dots, \hat{\eta}_m)^T$ est l'estimateur du maximum de vraisemblance de $\eta = (\eta_1, \dots, \eta_m)^T$ obtenu sur un échantillon de taille n , et h , une fonction de $\mathcal{H}^m$ dans $\mathcal{H}$ dérivable à dérivée non nulle, alors $h(\hat{\eta})$ est l'estimateur du maximum de vraisemblance de $h(\eta)$.

Et l'intervalle de confiance symétrique de seuil α de $h(\eta)$ est :

$$IC_{\alpha}(h(\eta)) = \left[h(\hat{\eta}) - z_{1-\frac{\alpha}{2}} \left(\frac{\nabla_{h(\hat{\eta})} \Sigma_{\hat{\eta}} \nabla_{h(\hat{\eta})}^T}{n} \right)^{-\frac{1}{2}} ; h(\hat{\eta}) + z_{1-\frac{\alpha}{2}} \left(\frac{\nabla_{h(\hat{\eta})} \Sigma_{\hat{\eta}} \nabla_{h(\hat{\eta})}^T}{n} \right)^{-\frac{1}{2}} \right]$$

Où :

- $z_{1-\frac{\alpha}{2}}$ est le quantile de la loi normale d'ordre $\frac{\alpha}{2}$,
- $\Sigma_{\hat{\eta}}$ est la matrice de variance covariance de $\hat{\eta}$,
- $\nabla_{h(\hat{\eta})} = \left(\frac{\partial h(\hat{\eta})}{\partial \eta_1}, \frac{\partial h(\hat{\eta})}{\partial \eta_2}, \dots, \frac{\partial h(\hat{\eta})}{\partial \eta_m} \right)^T$ le vecteur jacobien.

Dans cette problématique, pour chaque dossier i , η est un vecteur de paramètre de taille $p+1$, tel que :

$$\eta = (\beta_1, \beta_2, \dots, \beta_p, \sigma)^T, \quad h_i(\eta) = \exp(X_i \beta + \frac{\sigma^2}{2}) \quad \text{et}$$

$$\nabla_{h_i(\hat{\eta})} = (X_{i1} h_i(\hat{\eta}), X_{i2} h_i(\hat{\eta}), \dots, X_{ip} h_i(\hat{\eta}), \hat{\sigma} h_i(\hat{\eta}))^T.$$

L'application des formules ci-dessus fournit des intervalles de confiance exploitables pour détecter les valorisations potentiellement incongrues (dossiers pour lesquels l'évaluation gestionnaire sortirait de cet intervalle). Toutefois, l'application directe de ces intervalles conduit à refuser un trop grand nombre de dossiers pour que cela soit utilisé lors des procédures d'inventaire : à 42 mois les $\frac{3}{4}$ des dossiers seraient à réviser. Nous décidons donc d'élargir les intervalles de confiance pour en faire des intervalles

d'acceptation :

$$IA = [0.90 * borne\ inférieure\ de\ IC_{1\%}(V_{ult}); 1.1 * borne\ supérieure\ de\ IC_{1\%}(V_{ult})].$$

Cet élargissement grossier permet d'ajuster facilement le nombre de dossiers détectés aux capacités de révision : avec cette définition, $\frac{1}{4}$ des dossiers de maturité 42 mois et $\frac{1}{2}$ des dossiers de 18 mois sont à réviser.

Ainsi, un dossier est refusé (à réviser) si l'évaluation gestionnaire n'appartient pas à l'intervalle d'acceptation.

L'utilisation d'une fonction des intervalles de confiance est possible pour les dossiers de la 3^{ème} famille pour laquelle il existe un modèle statistique. Pour les deux autres familles, la valeur de référence est fixée de façon déterministe à 0 ou à 1 036€. Pour détecter les dossiers à réviser au sein de ces familles, nous construisons des intervalles d'acceptation basés sur des règles déterministes.

Ainsi, lorsque l'on considère qu'un dossier est membre de la 1^{ère} famille, on le classe « à réviser » s'il est évalué par le gestionnaire à plus de 1 000 €. Si c'est un dossier de la deuxième famille, il est acceptable si l'évaluation dossier-dossier est inférieure à 2 000 €¹.

Après avoir défini ces intervalles d'acceptation, nous vérifions qu'ils permettent de détecter les dossiers problématiques à l'aide de l'*Erreur Gestion* (erreur relative à partir de la valorisation dossier-dossier). Comme précédemment, on effectue une étude de distribution qui révèle que, sur les dossiers acceptés, les erreurs de gestion sont inférieures à celles observées sur les dossiers refusés. Les dossiers refusés sont donc bien, en règle générale, des dossiers pour lesquels l'anticipation de l'évaluation ultime pose problème.

Cette validation étant faite, nous pouvons supposer que les dossiers acceptés peuvent être plus sereinement intégrés à des méthodes statistiques de provisionnement. Puisque leur acceptabilité nous permet de valider que :

- l'évaluation dossier-dossier est cohérente avec son contenu informatique,
- sa variabilité est limitée.

10. DÉTOURNEMENT DE LA MÉTHODE

Avant de résumer la démarche générale conduisant à la détection des dossiers à réviser, nous nous attardons sur un détournement possible de cette méthode qui peut s'avérer intéressant lors des estimations de provisionnement. En construisant des

¹ En croisant les règles d'appartenance et les règles d'acceptation on remarque qu'en pratique les dossiers de la deuxième famille ne peuvent être refusés.

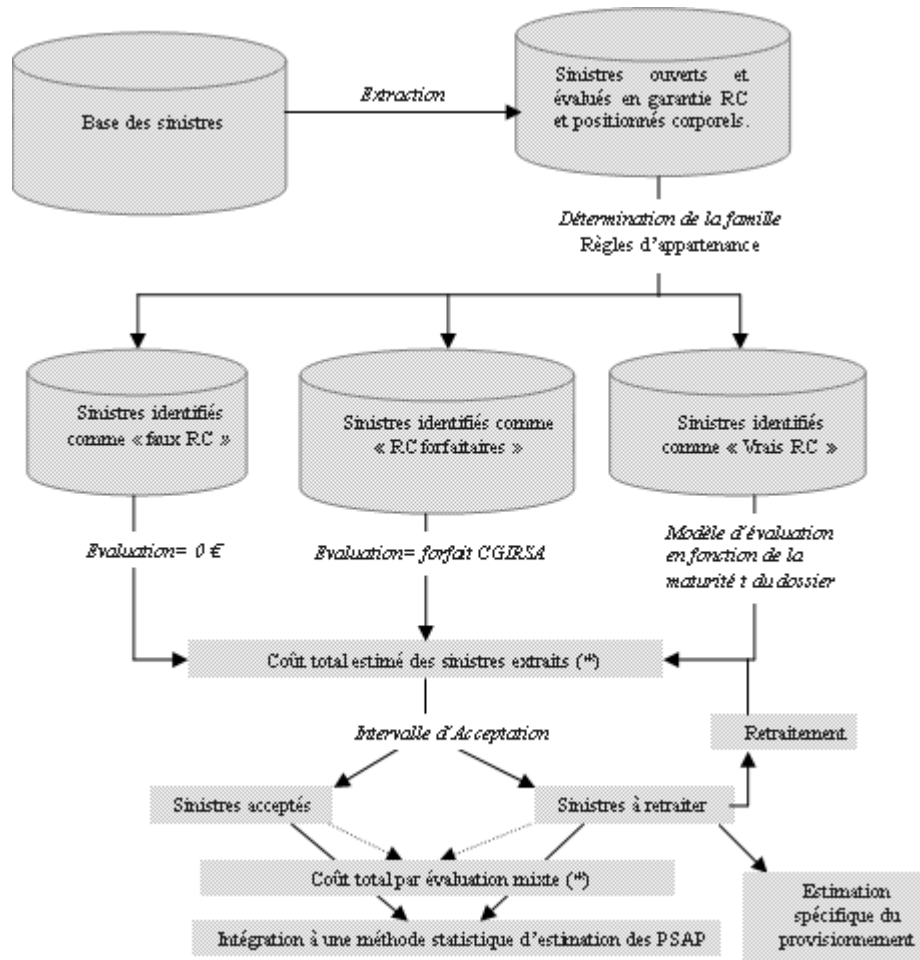
évaluations de référence, nous avons obtenu des estimations du coût de chaque dossier. Ces estimations peuvent être agrégées pour obtenir une estimation globale de la charge d'un exercice. Cette méthodologie étant totalement indépendante des méthodes classiques, l'obtention d'une estimation de la charge ultime globale permet d'enrichir le faisceau de méthodes qu'il est conseillé d'utiliser.

Il existe 2 façons d'obtenir une estimation du coût ultime à l'aide de ce dispositif :

- On peut, dans un premier temps, sommer toutes les estimations obtenues,
- On peut également sommer les valorisations dossier-dossier pour les dossiers acceptés et les estimations modélisées pour les dossiers refusés. On obtient alors une estimation de charge par la méthode mixte.

Remarque : *Les deux étapes du détournement de ce dispositif sont signalées par un astérisque dans le schéma global suivant.*

11. SCHÉMA GLOBAL



12. MISE EN PRATIQUE DE LA METHODE

Le schéma global étant posé, nous allons voir comment le dispositif peut s'intégrer à une problématique d'estimation du provisionnement et quelles en sont les conséquences.

Pour cela, nous considérons un triangle de charge très simple avec seulement 4 années de développement. Les données sont introduites dans le triangle sans traitement préalable (ni actualisation, ni révision des évaluations dossier-dossier) :

Charges en K€				
Année de survenance	Année de développement			
	1	2	3	4
1999	38 833	42 936	39 526	38 420
2000	43 228	52 324	45 515	
2001	47 654	53 207		
2002	33 093			

Sur ce triangle, nous appliquons tout d'abord des méthodes classiques telles que Chain Ladder standard et une méthode stochastique. Seule la méthode stochastique et ses résultats sont présentés en détail ci-après. En outre, nous proposons les résultats fournis par une méthode de Chain Ladder sur règlement cumulé afin d'alimenter le faisceau des méthodes actuarielles.

On note i les années de survenance et j les années de développement, variant de 1 à 4. Dans la méthode stochastique, la charge est supposée suivre une loi normale d'espérance μ_{ij} et de variance σ^2 . La fonction de lien choisie est logarithmique de telle sorte qu'en notant a et b les paramètres du modèle, on ait : $\ln \mu_{ij} = cste + ai + bj$.

L'application de ce modèle sur le triangle précédent donne les paramètres suivants :

	Paramètres			
	1	2	3	4
Constante	17.321	17.321	17.321	17.321
Année surv.	0.1431	0.2964	0.3501	0
Année dev.	0	0.1299	0	0

Les années de survenance semblent avoir plus de poids que les années de développement dans l'estimation de la charge. Le paramètre de l'exercice de survenance permet de prendre en compte l'accroissement du portefeuille assuré et les tendances de sinistralité de l'exercice.

Les paramètres des années de développement 1 et 3 apparaissent comme non significativement différents de 0. Ainsi, la charge de l'année de développement 3 est la même que celle de l'année de développement 4. Ceci se retrouve également sur la méthode Chain Ladder sur les charges avec un facteur de développement de 0.97.

Ces paramètres nous permettent de calculer pour l'ensemble des 4 années de développement et pour les 4 exercices considérés les charges sinistres suivantes :

Méthode stochastique sur Charge en K€				
Année de survenance	Année de développement			
	1	2	3	4
1999	38 420	43 749	38 420	38 420
2000	44 785	50 998	44 785	44 785
2001	47 256	53 811	47 256	47 256
2002	33 297	37 916	33 297	33 297

Sur les années écoulées, on remarque des différences avec les charges observées allant jusqu'à 1.3 M€ (3%) pour l'exercice 2000. La charge globale estimée est de 163.8M€. Son intervalle de confiance symétrique à 5%, obtenu par la méthode Delta, est [146.9M€ ; 180.6M€]. Cet intervalle a une largeur de 33.7M€, ce qui représente 21% de l'estimation de charge donnée par cette méthode.

Dans le tableau ci-dessous, on regroupe les estimations de la charge ultime globale fournies par les 4 méthodes (dossier-dossier, Chain Ladder sur charge, sur règlement, méthode stochastique) ainsi que les intervalles de confiance issus de la méthode stochastique.

Estimation sur échantillon non retraité en K€						
Estimation	Dossier-dossier	CL sur charge	Méthode Sto	Règlement cum	Interv. Conf. Inf.	Interv. Conf. Sup.
Charge	170 235	161 698	163 760	153 259	146 906	180 614
PSAP	130 515	121 978	124 040	113 539	107 186	140 894

On s'aperçoit que l'estimation dossier-dossier est largement supérieure à celle fournie par les méthodes statistiques. Les estimations de PSAP divergent de 17M€ (13% des PSAP dossier-dossier). Pourtant, toutes les estimations sont incluses dans l'intervalle de confiance donné par les méthodes stochastiques. L'écart le plus grand est mesuré entre la méthode dossier-dossier et les règlements cumulés qui apparaissent d'ailleurs un peu isolés.

La largeur du faisceau d'estimation des PSAP par méthodes statistiques est ici de 10.5M€ (8% des PSAP dossier-dossier). Cette divergence ne permet pas de conclure en toute sérénité puisque clôturer les comptes en intégrant une charge de 170.2M€, expose à un surprovisionnement de 17M€.

Au vu de ce constat, nous décidons de retravailler sur les données de façon à réduire la largeur du faisceau et à minimiser le risque de surprovisionnement en fiabilisant les pivots. Toutefois, pour optimiser la révision des dossiers, nous appliquons la méthodologie

de détection de dossiers instables. Ceci fait ressortir que 30% des dossiers non clos sont à retraiter.

Les révisions étant faites, nous disposons d'un nouveau triangle issu l'échantillon en partie révisé. Ces retraitements modifient les pivots de 5.8M€ (3.4%). Le pivot le plus touché est celui de l'exercice 2000 avec une baisse de 5% de sa valeur. L'application des 4 méthodes précédentes sur ce nouveau jeu de données donne :

Estimation sur échantillon retraité en K€						
Estimation	Dossier-dossier	CL sur charge	Méthode Sto	Règlement cum	Interv. Conf. Inf.	Interv. Conf. Sup.
Charge	164 372	151 821	157 757	153 259	132 923	182 590
PSAP	124 652	112 101	118 037	113 539	93 203	142 870

Sur ces nouvelles estimations, on remarque que l'écart s'est réduit entre les 4 méthodes tout comme entre les 3 méthodes statistiques.

Ainsi, on remarque qu'une diminution de 3.4% de la charge dossier-dossier perturbe fortement les estimations. La méthode Chain Ladder sur charge est la plus contrariée avec une différence de provisionnement de 9.9M€ (8.1%). Le retraitement a touché les pivots de cette méthode, ce qui explique ce résultat. La méthode stochastique semble mieux résister à une révision des pivots, puisque ses résultats sont impactés de 4.8%. Toutefois, l'élargissement de l'intervalle de confiance traduit la perturbation des données.

La largeur du faisceau formé des méthodes statistiques est ici de 5.9M€. Ainsi, une révision de 3.4% de la charge dossier-dossier diminue de 43.5% le faisceau des méthodes statistiques.

Dans la première configuration, la méthode des règlements cumulés paraissait sous estimer les PSAP, alors qu'après révision, cette méthode est au centre du faisceau. La révision a recentré les estimations et les a rendues cohérentes avec les règlements. Le résultat fourni par les règlements cumulés n'est plus du tout isolé.

CONCLUSION

Cette étude montre que les données détaillées permettent de repérer les dossiers posant des problèmes d'anticipation et ainsi d'améliorer la qualité des données sans avoir à développer des processus complexes. Toutefois, ces travaux conservent un caractère exploratoire.

La problématique de détection de valorisations dossier-dossier incongrues s'est vite

révélée vaste et confuse : de multiples manières d'aborder le sujet sont apparues. Plusieurs choix ont dû être faits, certains ont été abandonnés et de nombreuses bifurcations ont été prises. Nous ne prétendons pas avoir pris les meilleures décisions à chaque carrefour. Ainsi, il n'est pas certain que la modélisation GLM soit la plus adaptée. Cette modélisation s'appuie sur les principales caractéristiques des dossiers et les variables faiblement renseignées sont alors exclues des modèles. Cette approche nous a peut-être conduit à éliminer des variables, peu usitées, mais extrêmement utiles pour détecter les sinistres à valorisation instable. L'historique des actes de gestion réellement intégré aux modèles apparaît très restreint. Cette restriction est toutefois un avantage dans une optique de traitement massif puisqu'elle peut être garante de robustesse et de pérennité. Ce type de modélisation a, en outre, l'avantage de posséder un cadre statistique bien connu et d'être simple à appliquer industriellement.

La définition des intervalles d'acceptation peut aussi être discutée. Nous aurions pu les définir à partir de mesures de contributions d'erreur de valorisations globales plutôt qu'individuelles.

Malgré nos hésitations et au vu des résultats et de l'impact des retraitements sur les estimations de PSAP, revenir aux données détaillées paraît essentiel.

BIBLIOGRAPHIE

ARJAS E. (1989): The claims reserving problem in non life insurance : some structural ideas - *Astin Bulletin* vol 19 N°2.

ENGLAND P.D., VERRAL R.J. (2002): Stochastic claims reserving in general insurance - *British Actuarial Journal*. N°8, 443-518.

FAHRMEIR L., TUTZ G. (2001), *Multivariate Statistical Modelling Based on Generalized Linear Models*. Second Edition - Springer - Springer series in statistics.

IA (2004) : Institut des Actuaire, Cahier des règles professionnelles, Version 1. Novembre 2004.

HAASTRUP S. (1997): Some fully bayesian micro models for claims reserving. *Laboratory of Actuarial Mathematics – University of Copenhagen* - Ph.D. thesis.

HAASTRUP S., Arjas E. (1996): Claims reserving in continuous time ; a non parametric Bayesian approach - *Astin Bulletin* vol 26 N°2.

HACHEMEISTER C.A. (1976): Breaking down the loss reserving process. *Casualty Actuarial Society -Loss Reserve Symposium Chicago, Illinois - Friday, September 17, 1976*.

HACHEMEISTER C.A. (1978): A structural model for the analysis of loss reserves - *Bulletin de l'Association Royale des Actuaire Belges*, 73, 17-27.

HACHEMEISTER C.A. (1980): A stochastic model for loss reserving - *Transactions of the 21th International congress of actuaries 1*, 185-194.

HESSELLAGER O. (1994): A Markov model for loss reserving - *Astin Bulletin vol 24* N°2.

LEBARBIER E., MARY-HUARD T. (2004) : Le critère BIC : fondements théoriques et interprétation - *Rapport de recherche N°5315 (septembre 2004) - INRIA*.

NORBERG R. (1993): Prediction of outstanding liabilities in non life insurance - *Astin Bulletin vol 23* N°1.

NORBERG R. (1999): Prediction of outstanding liabilities (Model variations and extensions) – *Astin Bulletin vol 29* N°1.

PARTRAT C. (2003) : Evaluation de la provision pour sinistre à payer-support de cours-Formation continue IA.

PARTRAT C., BESSON J.L. (2004) : *Assurance non vie : modélisation, simulation* - Economica.

SHEN W.H. (1998): Estimation of parameters of a lognormal distribution - *Taiwanese Journal of Mathematics vol 2* N°2, p243-250, juin 1998.

SIMAR T., PARIS J. (1996) : Le nombre de sinistres nécessaires pour en estimer valablement le coût moyen, dans le cas lognormal - *Discussion Paper - Université Catholique de Louvain (UCL)*.

TAYLOR G. (2000): *Loss reserving, an actuarial perspective* - Kluwer Academic Publisher.

TAYLOR G., CAMPBELL M. (2002): Statistical case estimation - *research paper number 104, University of Melbourne*.