http://www.therond.fr

Théorie de la crédibilité

Pierre Thérond
pierre@therond.fr

Institut de Science Financière et d'Assurances - Université Lyon 1

Année universitaire 2017-2018

Plan du cours

Bibliographie principale

Bühlmann H., Gisler A. (2005) *A course in Credibility Theory and its Applications*, Springer.

Denuit M., Charpentier (2005) *Mathématiques de l'assurance non-vie*, volume II, Economica.

Herzog T.N. (1999) *Introduction to Credibility Theory*, ACTEX Publications.

Partrat Ch., Besson J.L. (2004) *Assurance non-vie*, Economica.

Philbrick S.W. (1981) An examination of credibility concepts, *Proceedings of the CAS*, vol. 68, 195-219.

Bibliographie complémentaire

Bühlmann H. (1967) Experience Rating and Credibility, *ASTIN Bulletin*, vol. 4, 199-207.

Bühlmann H., Straub E. (1970) Glaubwürdigkeit für Schadensätze, *Mitteilungen der Vereinigung Schweizerischer Versicherungsmathematiker*, vol. 70.

Cohen A., Dupin G., Levy C. (1986) Tarification de l'incendie des risques industriels français par la méthode de la crédibilité, *ASTIN Bulletin*, vol. 4.

Denuit M., Kaas R., Goovaerts M., J. Dhaene (2005) *Actuarial Theory for Dependent Risks : Measures, Orders and Models*, Wiley : New York.

Droesbeke J.J., Fine J., Saporta G. (éditeurs) (2002) *Méthodes bayesiennes en statistique*, Technip.

Gerber H. (1995) A teacher's remark on exact Credibility, *ASTIN Bulletin*, vol. 16.

Goulet V. (1998) Principles and application of credibility therory, *Journal of Actuarial Practice*, vol. 6.

Longley-Cook L.H. (1962) An Introduction to Credibility Theory, *Proceedings of the CAS*, vol. 49, 194-221.

Exercices

Des exercices figurent en fin de chaque chapitre.

Les étudiants sont néanmoins invités à consulter le recueil d'exercices de Vincent Goulet et Hélène Cossette dont le lien est disponible sur la page dédiée à ce cours :

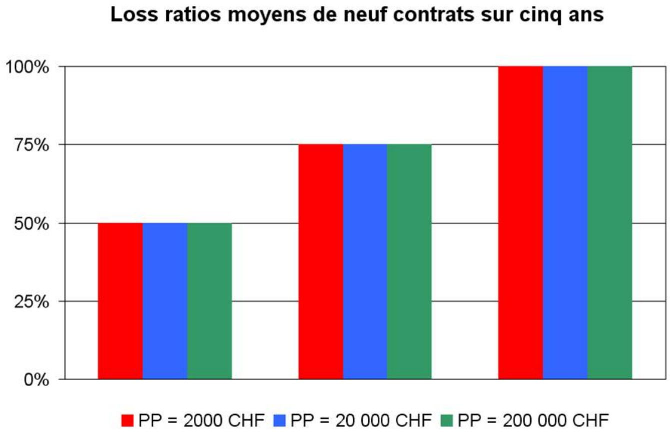
http://www.therond.fr

Cossette, H. et Goulet, V., *Exercices en théorie de la crédibilité : Avec solutions*, 2e éd. révisée, Document libre publié sous contrat Creative Commons, 2008. ISBN 978-2-9809136-9-3

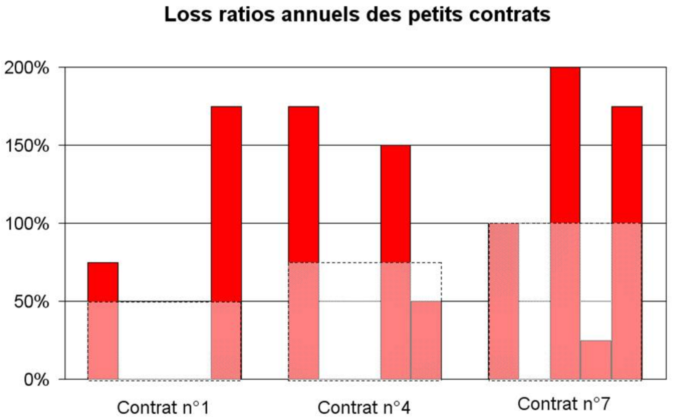
Sommaire

1. Illustrations introductives

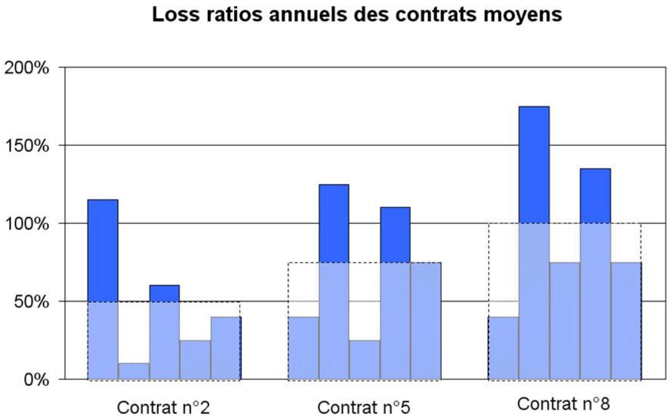
1.1. Une première illustration



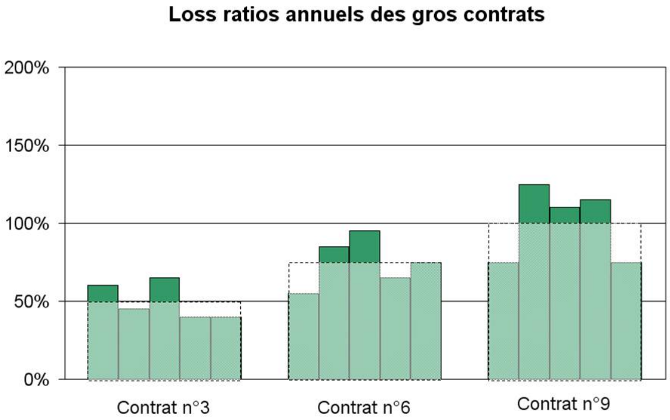
1.1. Une première illustration



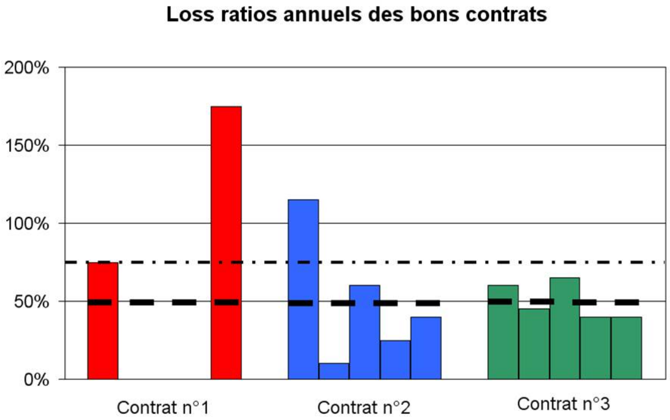
1.1. Une première illustration



1.1. Une première illustration



1.1. Une première illustration



1.2. Exemple de Ragnar Norberg

Considérons 20 contrats indépendants et *a priori* homogènes (la segmentation ne permet pas de les différencier). Ces contrats peuvent être touchés par des sinistres de montant 1 dont la survenance annuelle est distribuée selon une loi de Bernoulli.

Ces 20 contrats ont été observés sur une période de 10 ans.

1.2. Exemple de Ragnar Norberg

Sinistralité observée

Années	1	2	3	4	5	6	7	8	9	10	Total
contrat 1											0
contrat 2											0
contrat 3	1		1								2
contrat 4											0
contrat 5						1			1		2
contrat 6											0
contrat 7		1	1								2
contrat 8											0
contrat 9		1	1		1	1	1			1	6
contrat 10			1								1
contrat 11	1	1			1				1		4
contrat 12						1		1		1	3
contrat 13									1		1
contrat 14								1			1
contrat 15											0
contrat 16											0
contrat 17	1	1		1			1			1	5
contrat 18	1										1
contrat 19										1	1
contrat 20											0

1.2. Exemple de Ragnar Norberg

Sur les 10 années d'observation, on dénombre au total 29 sinistres. La sinistralité annuelle moyenne par contrat est donc de $29/200 = 0,145$.

Les contrats n°9, 11 et 17 représentent à eux-seuls 15 sinistres, soit plus de la moitié de la sinistralité globale. Sans eux, la prime pure annuelle moyenne serait de $14/170 = 0,083$

Il convient de se demander si le portefeuille est réellement homogène.

1.2. Exemple de Ragnar Norberg

Notons p_j le paramètre de la loi de Bernoulli du j-ème contrat et $\hat{p}_j$ son estimation naturelle à partir des observations. Notons de plus $\hat{p} := \frac{1}{20} \sum \hat{p}_j$.

On cherche à tester $H_0 : p_1 = p_2 = \ldots = p_{20}$.

Sous H_0, la statistique

$$X^2 = 10 \sum_{j=1}^{20} \frac{(\hat{p}_j - \hat{p})^2}{\hat{p}}$$

est distribuée selon une loi du χ^2 à 19 degrés de liberté. Comme $\chi^2_{19}(0,05) = 30,14 < X^2 = 57,9$, on rejette H_0.

A posteriori, le portefeuille n'est donc pas homogène.

1.2. Exemple de Ragnar Norberg

La tarification actuelle (même prime pure $\hat{p} = 0,145$ pour tous) n'est pas équitable : la plupart des assurés paient trop alors que les assurés 9, 11 et 17 sont sous-tarifés. Si cette situation perdurait, cela inciterait les *bons risques* à partir.

La société est confrontée à une alternative :

- segmenter davantage pour avoir des portefeuilles plus homogènes ; ou
- trouver un compromis entre p et p_j qui ne soit pas trop volatile au cours du temps.

2. Théorie de la fluctuation limitée

2.1. Les origines

En 1910, General Motors (GM) assurée chez Allstate pour les accidents de travail remarque que sa prime d'expérience est sensiblement inférieure à la prime réclamée à l'ensemble des entreprises assurées.

En argumentant que sa propre exposition au risque (nombre de salariés) est suffisament importante, GM réclame un tarif fondé exclusivement sur sa propre sinistralité et non celle de l'ensemble des sociétés assurées.

Dans le même temps, une plus petite entreprise (Tucker) fait une demande identique.

2.1. Les origines

Quelle est la taille à partir de laquelle une une entreprise peut être tarifée exclusivement sur la base de sa propre expérience ?

Mowbray [1914] puis Whitney [1918] apportent les premières réponses à cette question en fondant la *théorie de la fluctuation limitée* ou *crédibilité américaine*.

Mowbray A.H. [1914] "How Extensive a Payroll is Necessary to Give a Dependable Pure Premium ?", *Proceedings of the CAS*, vol. 1, 24-30.

Whitney A. [1918] "The Theory of Experience Rating ?", *Proceedings of the CAS*, vol. 4, 274-92.

2.1. Les origines

La théorie de la fluctuation limitée constitue la première approche de crédibilité qui a été utilisée pour mettre à jour les primes individuelles de la sinistralité individuelle observée.

L'idée est de tarifer le contrat n°j pour la $(t+1)$-ème période par

$$(1-Z_t)\,\hat{p} + Z_t\,{}_j\bar{p}_t,$$

où ${}_j\bar{p}_t = \frac{1}{t}\sum_{s=1}^{t} {}_jp_s$. Le coefficient Z est appelé *facteur de crédibilité*.

La théorie de la fluctuation limitée consiste à fixer Z de manière à ce que les variations aléatoires de la prime restent limitées.

2.2. Fluctuation limitée

Deux situations sont possibles :

- $Z = 1$, on parle de *crédibilité totale*.
- $Z \in]0;1[$, on parle alors de *crédibilité partielle*.

2.3. Crédibilité totale

On parle de crédibilité totale lorsque la prime pour la $(t+1)$-ème période dépend exclusivement de la sinistralité propre de l'assuré observée sur les t premières périodes.

La prime vaut donc ${}_j\bar{p}_t$.

Le problème réside dans la question de l'applicabilité de cette prime : à partir de quelle exposition au risque, peut-on appliquer à un assuré une prime basée exclusivement sur sa propre expérience ?

2.3. Crédibilité totale

Notons S la charge de sinistres d'un assuré au cours d'une période. Supposons que

$$S = \sum_{k=1}^{N} X_k,$$

où $N \sim \mathcal{P}(\lambda)$ et (X_k) est une suite de v.a.i.i.d. de moyenne μ et de variance σ^2 indépendante de N.

On accordera à l'assuré une crédibilité totale si la probabilité que la charge de sinistres s'éloigne de la moyenne est suffisament faible, *i.e.* si

$$\Pr\left[\frac{|S - \mathbf{E}[S]|}{\mathbf{E}[S]} \geq c\right] \leq \epsilon,$$

où c et ϵ sont des constantes fixées *a priori* (par exemple : $c = 5\%$ et $\epsilon = 5\%$).

2.3. Crédibilité totale

Le théorème central-limite (applicable à une somme aléatoire) permet, pour un nombre important d'observations de faire l'approximation

$$\frac{S - \mathbf{E}[S]}{\sqrt{\mathbf{Var}[S]}} \approx \mathcal{N}(0,1).$$

Ce qui conduit à

$$\frac{c\mathbf{E}[S]}{\sqrt{\mathbf{Var}[S]}} = \frac{c\lambda\mu}{\sqrt{\lambda\left(\sigma^2+\mu^2\right)}} \approx z_{\epsilon/2}.$$

D'où l'on déduit

$$\lambda_F := \frac{z_{\epsilon/2}^2}{c^2}\left(1+\left(\frac{\sigma}{\mu}\right)^2\right),$$

le nombre minimum de sinistres espérés durant la prochaine période pour que l'on puisse appliquer la crédibilité totale.

2.3. Crédibilité totale

Dans le cas particulier où les montants de sinistres sont constants (réparation forfaitaire), pour $c = 5\%$ et $\epsilon = 5\%$, le nombre minimum de sinistres espérés durant la prochaine période doit être supérieur à

$$\lambda_F = \left(\frac{1,96}{5\%}\right)^2 \left(1 + (0)^2\right) \approx 1537,$$

pour que l'assuré se voit appliquer une prime exclusivement fondée sur sa propre expérience.

En pratique, peu d'assurés peuvent répondre à ce genre de critère.

2.4. Crédibilité partielle

Lorsqu'un assuré ne dispose pas du volume de risque λ_F, la théorie de la fluctuation limitée permet de déterminer une prime qui sera un barycentre de la prime collective et de la prime d'expérience.

Ainsi on est amené à déterminer le facteur de crédibilité Z tel que

$$\Pr\left[\frac{|S-\mathbf{E}[S]|}{\mathbf{E}[S]} Z \geq c\right] \leq \epsilon.$$

En procédant de la même manière que précédemment, on obtient la relation

$$Z = \frac{\sqrt{\lambda}}{\frac{z_{\epsilon/2}}{c}\sqrt{\left(1+\left(\frac{\sigma}{\mu}\right)^2\right)}} = \sqrt{\frac{\lambda}{\lambda_F}}.$$

2.5. Quelques remarques

Si le principal intérêt de la fluctuation limitée réside dans sa simplicité de mise en oeuvre, cette théorie souffre d'inconvénients qui en font un outil peu utilisé en pratique :

- choix *a priori* des deux paramètres c et ϵ délicat (quelles justifications ?),
- approximation central-limite,
- approche paradoxale : pour obtenir le facteur de crédibilité, on fait l'hypothèse que l'on connaît μ, σ^2 et λ, *i.e.* que l'on est capable de déterminer la prime pure propre de l'assuré considéré. Dans une telle situation, pourquoi choisir une autre prime ?

3. Formalisation mathématique du problème de tarification

3.1. Risques individuel et collectif

Considérons un risque individuel qui produit des montants agrégés de sinistres $(X_j)_{j=1,\dots,n}$ où X_j est le montant correspondant à la période j.

On souhaite déterminer la prime pure $\mathbf{E}[X_{n+1}]$ pour la période future.
Pour cela, on fait les hypothèses classiques :

- de stationnarité : les X_j sont identiquement distribués de fonction de répartition F ;
- d'indépendance : les X_j sont indépendants.

En pratique, F est inconnue et varie d'un risque à l'autre.

Notation : $\theta \in \Theta$ désignera le profil de risque et F_θ la fonction de réparition correspondante.

3.1. Risques individuel et collectif

La *prime individuelle correcte* pour un assuré de profil de risque θ est

$$P^{ind}(\theta) := \mathbf{E}\left[X_{n+1}|\theta\right] =: \mu(\theta).$$

En pratique, θ et $\mu(\theta)$ sont inconnus. Il faut donc trouver un estimateur $\hat{\mu(\theta)}$ pour $\mu(\theta)$.

Les assurés ont des profils de risque θ_i $(i \in \{1, ..., I\})$ inconnus de l'assureur qui prennent leur valeur dans Θ. N.B. Dans le cas d'un portefeuille réellement homogène, Θ est réduit à un singleton.

En revanche, l'assureur dispose d'informations sur la structure collective du risque (ex : la plupart des conducteurs sont des bons risques qui ont rarement un accident et un petit nombre a fréquemment des sinistres). Cette information peut être résumée par la distribution de probabilité $U(\theta)$ sur Θ.

3.1. Risques individuel et collectif

Definition (Fonction de structure)

La distribution $U(\theta)$ est appelée fonction de structure du portefeuille.

La *prime collective* est donnée par

$$P^{coll} := \int_{\Theta} \mu(\theta) dU(\theta) =: \mu_0.$$

Cette prime correspond au montant moyen de sinistre espéré par risque sur l'ensemble du portefeuille.

Thèse La prime la plus compétitive est la prime individuelle correcte. Sinon il y a des opportunités d'arbitrage.

3.2. Formulation bayesienne

Le problème évoqué *supra* se décrit aisément en termes bayésiens. La problématique peut être résumée par un modèle à deux urnes.

La première urne représente le collectif de risques dont on sélectionne un élément θ qui conditionne le contenu de la deuxième urne dans laquelle on tire des variables aléatoires $X_1, X_2, \ldots$ de fonction de répartition F_θ.

Formellement :

- Θ est une v.a. de fonction de répartition U ;
- conditionnellement à $\Theta = \theta$, les v.a. $X_1, X_2, \ldots$ sont i.i.d. de fonction de répartition F_θ.

3.2. Formulation bayesienne

Remarques :

- Dans cette interprétation, la prime individuelle devient une v.a. puisque l'on ne sait pas distinguer *a priori* deux risques

$$P^{ind} := \mathbf{E}\left[X_{n+1}|\Theta\right] =: \mu(\Theta).$$

- La prime collective est, en revanche, un montant certain

$$P^{coll} := \mu_0 := \int_\Theta \mu(\theta) dU(\theta) = \mathbf{E}\left[X_{n+1}\right].$$

- Les $X_1, X_2, \ldots$ ne sont indépendants que conditionnellement à $\Theta = \theta$. Sinon ils sont positivement corrélés :

$$\begin{aligned}\mathbf{Cov}(X_1, X_2) &= \mathbf{E}\left[\mathbf{Cov}(X_1, X_2|\Theta)\right] + \mathbf{Cov}\left(\mathbf{E}[X_1|\Theta], \mathbf{E}[X_2|\Theta]\right)\\ &= 0 + \mathbf{Cov}[\mu(\Theta), \mu(\Theta)] = \mathbf{Var}[\mu(\Theta)].\end{aligned}$$

3.2. Formulation bayesienne

Notre objectif est d'estimer pour chaque risque la prime correcte $\mu(\Theta)$ aussi précisément que possible compte-tenu de l'historique $\mathbf{X} = (X_1, ..., X_n)'$ dont on dispose.

La *meilleure prime d'expérience* ou *prime de Bayes* est définie par

$$P^{Bayes} = \widetilde{\mu(\Theta)} := \mathbf{E}\left[\mu(\Theta)|\mathbf{X}\right].$$

Il s'agit d'une prime certaine à la différence de la prime individuelle.

Exercices

Exercice 1.1 Considérons la situation où, si les montants de sinistres sont constants, le nombre de sinistres attendu pour pouvoir utiliser la crédibilité totale vaut 2 670. Quel est le nombre minimum de sinistres attendus pour pouvoir utiliser la crédibilité totale lorsque les montants de sinistres sont distribués selon une loi log-normale d'espérance 1 000 et de variance 1 500 000 ?

Exercice 1.2 Soit x le nombre de sinistres espérés pour pouvoir utiliser la crédibilité totale en acceptant une erreur de 5 % avec une probabilité de 90 %. Notons y la même grandeur pour une erreur acceptée de 10 %. Que vaut le rapport $\frac{x}{y}$?

Exercice 1.3 Montrer que l'on attend au moins 2 401 accidents par an pour pouvoir utiliser la crédibilité totale en acceptant une erreur de 4 % avec une probabilité de 95 %. En supposant que l'on dispose d'une exposition au risque de 40 000 véhicules par an et que la fréquence de sinistres moyenne est 0,0441, déterminer le niveau maximal du facteur de crédibilité.

Sommaire

1. Langage bayesien

1. Langage bayesien

Considérons une hypothèse H et un événement B et notons $\Pr[H]$ notre connaissance initiale (*a priori*) de la vraisemblance de l'hypothèse H. Alors

- $\Pr[B|H]$ représente la probabilité conditionnelle de l'événement B sous l'hypothèse H ;
- considérée comme une fonction de H, $\Pr(B|H)$ est la vraisemblance de B sachant H ; et
- $\Pr[H|B]$ est la probabilité *a posteriori* de H.

1. Langage bayesien

Le théorème de Bayes nous permet d'écrire

$$\Pr[H|B] = \frac{\Pr[B|H]\Pr[H]}{\Pr[B]},$$

lorsque $\Pr[B] > 0$.

Ainsi lorsque $\Pr[B] > 0$, la probabilité *a posteriori* de H ($\Pr[H|B]$) est proportionnelle au produit de la probabilité *a priori* ($\Pr[H]$) et de la vraisemblance de B sachant H ($\Pr[B|H]$).

1. Langage bayesien

Rappels :

- Prime individuelle (variable aléatoire) :

$$P^{ind} = \mu(\Theta) = \mathbf{E}\left[X_{n+1}|\Theta\right].$$

- Prime collective (réel) :

$$P^{coll} = \mu_0 = \int_\Theta \mu(\theta)dU(\theta) = \mathbf{E}\left[X_{n+1}\right].$$

- Prime de Bayes (réel lorsque l'on a observé $\mathbf{X}$) :

$$P^{Bayes} = \widetilde{\mu(\Theta)} = \mathbf{E}\left[\mu(\Theta)|\mathbf{X}\right].$$

2. Meilleure prime d'expérience

2. Meilleure prime d'expérience

Definition (Critère de l'erreur quadratique moyenne)

Un estimateur $\widehat{\mu(\Theta)}$ de $\mu(\Theta)$ est au moins aussi bon qu'un estimateur $\widehat{\mu(\Theta)}^*$ si

$$\mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \mu(\Theta)\right)^2\right] \leq \mathbf{E}\left[\left(\widehat{\mu(\Theta)}^* - \mu(\Theta)\right)^2\right],$$

où $\mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \mu(\Theta)\right)^2\right]$ est l'erreur quadratique moyenne de $\widehat{\mu(\Theta)}$.

Théorème

Au regard du critère de l'erreur quadratique moyenne,

$$\widetilde{\mu(\Theta)} = \mathbf{E}\left[\mu(\Theta)|\mathbf{X}\right]$$

est le meilleur estimateur de $\mu(\Theta)$.

2. Meilleure prime d'expérience

Démonstration :

Soit $\widehat{\mu(\Theta)}$ un estimateur de $\mu(\Theta)$ et $\widetilde{\mu(\Theta)} = \mathbf{E}\left[\mu(\Theta)|\mathbf{X}\right]$ l'espérance *a posteriori* de $\mu(\Theta)$. On a

$$\mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \mu(\Theta)\right)^2\right] = \mathbf{E}\left[\mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \widetilde{\mu(\Theta)} + \widetilde{\mu(\Theta)} - \mu(\Theta)\right)^2 |\mathbf{X}\right]\right].$$

Or

$$\mathbf{E}\left[\left(\widehat{\mu} - \widetilde{\mu}\right)\left(\widetilde{\mu} - \mu\right)|\mathbf{X}\right] = \widetilde{\mu}\mathbf{E}\left[\widehat{\mu}|\mathbf{X}\right] - \mathbf{E}\left[\mu\widehat{\mu}|\mathbf{X}\right] - \widetilde{\mu}^2 + \widetilde{\mu}^2,$$

d'où

$$\mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \mu(\Theta)\right)^2\right] = \mathbf{E}\left[\left(\widehat{\mu(\Theta)} - \widetilde{\mu(\Theta)}\right)^2\right] + \mathbf{E}\left[\left(\widetilde{\mu(\Theta)} - \mu(\Theta)\right)^2\right].$$

$\square$

2. Meilleure prime d'expérience

Théorème (Erreur quadratique moyenne)

(1) L'erreur quadratique moyenne de la prime de Bayes vaut :

$$\mathbf{E}\left[\left(\widetilde{\mu(\Theta)} - \mu(\Theta)\right)^2\right] = \mathbf{E}\left[\mathbf{Var}(\mu(\Theta)|\mathbf{X})\right].$$

(2) L'erreur quadratique moyenne de la prime collective est donnée par :

$$\begin{aligned}\mathbf{E}\left[(\mu_0 - \mu(\Theta))^2\right] &= \mathbf{Var}\left[\mu(\Theta)\right] \\ &= \mathbf{E}\left[\mathbf{Var}(\mu(\Theta)|\mathbf{X})\right] + \mathbf{Var}\left[\mathbf{E}(\mu(\Theta)|\mathbf{X})\right].\end{aligned}$$

Remarque : L'erreur quadratique moyenne de la prime de Bayes correspond au premier terme de décomposition de la variance dans celle de la prime collective.

3. Construction d'une échelle bonus-malus

3.1. Contexte

- Fin des années 60, en Suisse, le niveau des primes d'assurance auto dépend uniquement de la puissance du véhicule.
- Les assureurs souhaitent augmenter le niveau des primes qu'ils jugent insuffisant pour couvrir les sinistres.
- Accord des autorités de contrôle sous réserve que le nouveau tarif tienne compte du passé-sinistre des assurés.
- F. Bichsel est chargé de mettre au point un système de mise à jour des primes mieux adapté aux profils de risques individuels.
- Bichsel estime que les différences de profil de risque entre les conducteurs sont mieux résumées par les nombres de sinistres que par leur taille.

3.2. Tarification a posteriori sans personnalisation a priori

- Notons N_j le nombre de sinistres engendrés par un conducteur particulier dans l'année j,
- X_j la charge agrégée de sinistres correspondante.

Bichsel fait implicitement l'hypothèse suivante :

Hypothèse (Hypothèse implicite de Bichsel)

Pour tout profil de risque θ, $\mathbf{E}\left[X_j|\Theta=\theta\right] = C\mathbf{E}\left[N_j|\Theta=\theta\right]$ où C est une constante qui dépend uniquement de la puissance du véhicule et où $\mathbf{E}\left[N_j|\Theta=\theta\right]$ dépend uniquement du conducteur.

Sous cette hypothèse, il suffit donc de modéliser le nombre de sinistres.

3.2. Tarification a posteriori sans personnalisation a priori

Hypothèse (Poisson-Gamma)

(PG1) Conditionnellement à $\Theta = \theta$, les v.a. N_j ($j = 1, ..., n$) sont indépendantes et identiquement distribuées selon une loi de Poisson de paramètre θ, i.e.

$$\Pr\left(N_j = k | \Theta = \theta\right) = e^{-\theta}\frac{\theta^k}{k!}, \;\; \textit{pour } k \in \mathbf{N}.$$

(PG2) Θ est distribué selon une loi Gamma de paramètres γ et β, i.e. Θ a pour densité

$$u(\theta) = \frac{\beta^\gamma}{\Gamma(\gamma)}\theta^{\gamma-1}e^{-\beta\theta}, \;\; \textit{pour } \theta \geq 0,$$

où $\Gamma(\nu) = \int_0^\infty e^{-x}x^{\nu-1}dx$, pour $\nu > 0$.

Remarque : $\mathbf{E}[\Theta] = \frac{\gamma}{\beta}$ et $\mathbf{Var}[\Theta] = \frac{\gamma}{\beta^2} \Rightarrow \beta = \frac{\mathbf{E}[\Theta]}{\mathbf{Var}[\Theta]}$ est une mesure de l'homogénéité du collectif.

3.2. Tarification a posteriori sans personnalisation a priori

Proposition

Sous l'hypothèse 2, on a pour la fréquence de sinistres (notée F)

$$\begin{aligned} F^{ind} &= \mathbf{E}\left[N_{n+1}|\Theta\right] = \Theta, \\ F^{coll} &= \mathbf{E}\left[\Theta\right] = \frac{\gamma}{\beta}, \\ F^{Bayes} &= \frac{\gamma + N_{\bullet}}{\beta + n} = \alpha\bar{N} + (1-\alpha)\frac{\gamma}{\beta}, \end{aligned}$$

où $\alpha = \frac{n}{n+\beta}$, $N_{\bullet} = \sum_{j=1}^{n} N_j$ *et* $\bar{N} = \frac{1}{n}\sum_{j=1}^{n} N_j$.

L'erreur quadratique moyenne de F^{Bayes} *est donnée par*

$$\mathbf{E}\left[\left(F^{Bayes} - \Theta\right)^2\right] = (1-\alpha)\mathbf{E}\left[\left(F^{coll} - \Theta\right)^2\right] = \alpha\mathbf{E}\left[\left(\bar{N} - \Theta\right)^2\right].$$

3.2. Tarification a posteriori sans personnalisation a priori

Remarques (1/2) :

- La prime de Bayes $P^{Bayes} = CF^{Bayes}$ est une fonction linéaire des observations (c'est donc une *prime de crédibilité*).
- F^{Bayes} est une moyenne pondérée de la fréquence individuelle observée $\bar{N}$ et de la fréquence espérée *a priori* $\mathbf{E}[\Theta] = \frac{\gamma}{\beta}$.
- $\alpha = \frac{n}{n+\beta}$ est appelé *facteur de crédibilité*. Il est croissant avec le nombre d'années d'observation n et décroissant avec β, *i.e.* plus on dispose d'informations individuelles, plus on accorde de crédit au passé-sinistres de l'assuré et plus le collectif est homogène, plus on accorde de crédit à la prime collective.
- α est également le facteur par lequel est réduite l'erreur quadratique moyenne lorsque l'on utilise F^{Bayes} au lieu de F^{coll}.

3.2. Tarification a posteriori sans personnalisation a priori

Remarques (2/2) :

- L'erreur quadratique moyenne de F^{coll} vaut :

$$\mathbf{E}\left[\left(F^{coll}-\Theta\right)^2\right]=\mathbf{E}\left[\left(\frac{\gamma}{\beta}-\Theta\right)^2\right]=\mathbf{Var}\left[\Theta\right]=\frac{\gamma}{\beta^2}.$$

- L'erreur quadratique moyenne de $\bar{N}$ vaut :

$$\mathbf{E}\left[\left(\bar{N}-\Theta\right)^2\right]=\mathbf{E}\left[\mathbf{E}\left[\left(\bar{N}-\Theta\right)^2|\Theta\right]\right]=\mathbf{E}\left[\mathbf{Var}\left(\bar{N}|\Theta\right)\right]$$
$$=\frac{\mathbf{E}[\Theta]}{n}=\frac{\gamma}{n\beta}.$$

- $\mathbf{E}\left[\left(F^{Bayes}-\Theta\right)^2\right]\leq\mathbf{E}\left[\left(\bar{N}-\Theta\right)^2\right]$ et $\mathbf{E}\left[\left(\bar{N}-\Theta\right)^2\right]\xrightarrow{n\to\infty}0$ donc

$$\mathbf{E}\left[\left(F^{Bayes}-\Theta\right)^2\right]\xrightarrow{n\to\infty}0.$$

3.2. Tarification a posteriori sans personnalisation a priori

Démonstration (1/2) :

Les deux premiers points sont immédiats.

Pour obtenir la prime de Bayes, écrivons la densité *a posteriori* de Θ sachant $\mathbf{N} = (N_1, ..., N_j)'$,

$$u(\theta|\mathbf{N}) = \frac{\frac{\beta^\gamma}{\Gamma(\gamma)}\theta^{\gamma-1}e^{-\beta\theta}\prod_{j=1}^{n} e^{-\theta}\frac{\theta^{N_j}}{N_j!}}{\int \frac{\beta^\gamma}{\Gamma(\gamma)}\theta^{\gamma-1}e^{-\beta\theta}\prod_{j=1}^{n} e^{-\theta}\frac{\theta^{N_j}}{N_j!}\,d\theta} \propto \theta^{\gamma+\sum_{j=1}^{n} N_j - 1} e^{-(\beta+n)\theta}.$$

Le membre de droite est le noyau de la densité d'une loi Gamma de paramètres $\gamma' = \gamma + N_\bullet$ et $\beta' = \beta + n$, d'où l'on obtient F^{Bayes}.

3.2. Tarification a posteriori sans personnalisation a priori

Démonstration (2/2) :

L'erreur quadratique moyenne de F^{Bayes} est donnée par

$$\begin{aligned}
\mathbf{E}\left[\left(F^{Bayes}-\Theta\right)^2\right] &= \mathbf{E}\left[\mathbf{Var}\left(\Theta|\mathbf{N}\right)\right] = \mathbf{E}\left[\frac{\gamma+N_\bullet}{(\beta+n)^2}\right] \\
&= \frac{\gamma}{(\beta+n)^2}\left(1+\frac{n}{\beta}\right) = \frac{\gamma}{\beta}\frac{1}{n+\beta} \\
&= \frac{\gamma}{\beta^2}(1-\alpha) = \frac{1}{n}\frac{\gamma}{\beta}\alpha.
\end{aligned}$$

$\square$

3.2. Tarification a posteriori sans personnalisation a priori

Tableau : *Distribution observée du nombre de sinistres par police en 1961*

k	Nombre de polices avec k sinistres
0	103 704
1	14 075
2	1 766
3	255
4	45
5	6
6	2
Total	119 853

3.2. Tarification a posteriori sans personnalisation a priori

Estimation des paramètres de structure par la méthode des moments :

$$\hat{\mu}_N = \bar{N}_{1961} := \frac{1}{I}\sum_{i=1}^{I} N^{(i)} = 0,1551$$

$$\hat{\sigma}_N^2 = \frac{1}{I-1}\sum_{i=1}^{I}\left(N^{(i)} - \bar{N}_{1961}\right)^2 = 0,1793$$

$$\mu_N = \mathbf{E}[N] = \mathbf{E}\left[\mathbf{E}[N|\Theta]\right] = \frac{\gamma}{\beta}$$

$$\sigma_N^2 = \mathbf{Var}[N] = \mathbf{E}\left[\mathbf{Var}[N|\Theta]\right] + \mathbf{Var}\left[\mathbf{E}[N|\Theta]\right]$$

$$= \mathbf{E}[\Theta] + \mathbf{Var}[\Theta] = \frac{\gamma}{\beta} + \frac{\gamma}{\beta^2} = \frac{\gamma}{\beta}\left(1 + \frac{1}{\beta}\right)$$

$$\Rightarrow \hat{\gamma} = 0,9956 \text{ et } \hat{\beta} = 6,4172.$$

3.2. Tarification a posteriori sans personnalisation a priori

En remplaçant les paramètres de structure γ et β par leurs estimateurs $\hat{\gamma}$ et $\hat{\beta}$, la prime de Bayes du conducteur ayant l'historique $(N_1, ..., N_n)$ est estimée par :

$$\widehat{F^{Bayes}}^{emp} = \frac{n}{n+\hat{\beta}}\bar{N} + \frac{\hat{\beta}}{n+\hat{\beta}}\frac{\hat{\gamma}}{\hat{\beta}}, \text{ où } \bar{N} = \frac{1}{n}\sum_{j=1}^{n} N_j.$$

Remarque : Le calcul de la prime de Bayes s'effectue fréquemment en deux étapes :

- détermination de l'expression de P^{Bayes} sous l'hypothèse que les paramètres de structure sont connus ;
- estimation des paramètres de structure à l'aide des données du collectif.

3.2. Tarification a posteriori sans personnalisation a priori

Description du risque dans le collectif

$$
\begin{aligned}
\mathbf{P}(N=k) &= \int \mathbf{P}\left(N=k|\Theta=\theta\right) u(\theta) d\theta \\
&= \int e^{-\theta}\frac{\theta^k}{k!}\frac{\beta^\gamma}{\Gamma(\gamma)}\theta^{\gamma-1}e^{-\beta\theta}d\theta \\
&= \frac{\beta^\gamma}{\Gamma(\gamma)}\frac{1}{k!}\underbrace{\int e^{-(\beta+1)\theta}\theta^{\gamma+k-1}d\theta}_{=\frac{\Gamma(\gamma+k)}{(\beta+1)^{\gamma+k}}} \\
&= \frac{\Gamma(\gamma+k)}{\Gamma(\gamma)k!}\left(\frac{\beta}{\beta+1}\right)^\gamma\left(\frac{1}{\beta+1}\right)^k \\
&= \begin{pmatrix} \gamma+k-1 \\ k \end{pmatrix} p^\gamma(1-p)^k, \text{ avec } p=\frac{\beta}{\beta+1}.
\end{aligned}
$$

On retrouve la distribution binomiale-négative.

3.2. Tarification a posteriori sans personnalisation a priori

Description du risque dans le collectif

Le modèle Poisson-Gamma représente mieux les observations que le modèle homogène dans lequel tous les assurés ont le même paramètre de Poisson.

Tableau : *Nombre de polices avec k sinistres*

k	Observé	$\mathcal{P}(\hat{\mu}_N)$	$\mathcal{BN}(\hat{\gamma}, \hat{\beta})$
0	103 704	102 630	103 761
1	14 075	15 922	13 927
2	1 766	1 235	1 874
3	255	64	252
4	45	2	34
5	6	0	5
6	2	0	1
Total	119 853	119 853	119 853

3.2. Tarification a posteriori sans personnalisation a priori

Construction de l'échelle bonus-malus (1/2)

L'échelle bonus-malus doit permettre d'ajuster la prime d'un conducteur en fonction de sa fréquence de sinistre espérée.

En pratique, cela passe par l'utilisation d'un coefficient à appliquer à la prime collective :

$$coef.\ bonus - malus = \frac{\widehat{F^{Bayes}}}{\lambda_0}$$

où $\lambda_0 = 0,1551$ est la fréquence de sinistres moyenne observée.

3.2. Tarification a posteriori sans personnalisation a priori

Construction de l'échelle bonus-malus (2/2)

Tableau : *Coefficients bonus-malus*

$n \backslash k$	0	1	2	3
1	87%	173%	260%	347%
2	76%	153%	229%	306%
3	68%	137%	205%	273%
4	62%	123%	185%	247%
5	56%	113%	169%	226%
6	52%	104%	156%	207%

Dans ce tableau, n représente le nombre d'années d'observation et k le nombre de sinistres observés.

3.3. Tarification a posteriori avec fréquences a priori différenciées

L'hypothèse 2 peut se réécrire sous la forme suivante :

- *(PG1) Conditionnellement à* $\Theta = \theta$*, les v.a.* N_j $(j = 1, ..., n)$ *sont indépendantes et identiquement distribuées selon une loi de Poisson de paramètre* $\theta\lambda_0$*, où* $\lambda_0 = \frac{\gamma}{\beta}$*.*
- *(PG2)* Θ *est distribué selon une loi Gamma avec* $\mathbf{E}[\Theta] = 1$ *et de paramètre de forme* γ*.*

Remarques (1/2) :

- $\mathbf{E}[\Theta] = 1 \Rightarrow$ le paramètre d'échelle de la distribution de Θ est égal au paramètre de forme γ.
- $F^{ind} = \mathbf{E}[N_j|\Theta] = \lambda_0\Theta$ donc Θ désigne directement le coefficient bonus-malus recherché.

3.3. Tarification a posteriori avec fréquences a priori différenciées

Remarques (2/2) :

- L'hétérogénéité du portefeuille peut être mesurée par le coefficient de variation de F^{ind}. Sous cette réécriture du modèle, $\mathbf{CoVa}(F^{ind}) = \sqrt{\mathbf{Var}[\Theta]}$ et $\gamma = \left(\mathbf{CoVa}(F^{ind})\right)^{-2}$.

- Pour estimer Θ, il est naturel de considérer la fréquence relative de sinistres $\widetilde{F}_j = \frac{N_j}{\lambda_0}$. En particulier, $\widetilde{F}^{ind} = \mathbf{E}[\widetilde{F}_j|\Theta] = \Theta$.

- En divisant F^{Bayes} par λ_0, on obtient :

$$\widetilde{F}^{Bayes} := \frac{F^{Bayes}}{\lambda_0} = \widetilde{\Theta} = \mathbf{E}[\Theta|\mathbf{N}] = 1 + \alpha\left(\frac{N_\bullet}{\nu_\bullet} - 1\right),$$

où $\nu_\bullet = n\lambda_0$ représente le nombre de sinistres espéré *a priori* et $\alpha = \frac{\nu_\bullet}{\nu_\bullet + \lambda_0\beta} = \frac{\nu_\bullet}{\nu_\bullet + \gamma}$. De plus,

$$\frac{N_\bullet}{\nu_\bullet} = \overline{\tilde{F}} := \frac{1}{n}\sum_{j=1}^{n}\tilde{F}_j \Rightarrow \tilde{\Theta} = 1 + \alpha\left(\overline{\tilde{F}} - 1\right).$$

3.3. Tarification a posteriori avec fréquences a priori différenciées

Considérons la situation dans laquelle la fréquence *a priori* varie selon les années et selon les risques et intéressons-nous à un risque de nombre de sinistres espéré λ_j et de nombre de sinistres observés N_j pour l'année j.

Hypothèse (Poisson-Gamma II)

(PG1') Conditionnellement à $\Theta = \theta$, les v.a. N_j $(j = 1, ..., n)$ sont indépendantes et identiquement distribuées selon une loi de Poisson de paramètre $\Theta\lambda_j$.
(PG2') Θ est distribué selon une loi Gamma avec $\mathbf{E}[\Theta] = 1$ et de paramètre de forme γ.

3.3. Tarification a posteriori avec fréquences a priori différenciées

Proposition

Sous l'hypothèse 3, on a :

$$\widetilde{F}^{Bayes} = \widetilde{\Theta} = \mathbf{E}[\Theta|\mathbf{N}] = 1 + \alpha\left(\frac{N_\bullet}{\nu_\bullet} - 1\right),$$

où

- $N_\bullet$ *est le nombre observé de sinistres,*
- $\nu_\bullet = \sum_{j=1}^{n} \lambda_j$ *le nombre de sinistres espéré a priori,*
- $\alpha = \frac{\nu_\bullet}{\nu_\bullet + \gamma}$.

Remarque : Ce résultat est plus fort que le précédent, car $\nu_\bullet$ peut varier entre les risques (par exemple dépendre de variables explicatives telles que la composition de la flotte de véhicules).

4. Classes de lois conjuguées

4. Classes de lois conjuguées

Dans l'exemple précédent, pour déterminer la prime de Bayes, il a été nécessaire de spécifier :

- la loi conditionnelle F_θ du nombre de sinistres pour $\Theta = \theta$,
- la fonction de structure $U(\theta)$.

Notons $\mathcal{F} = \{F_\theta : \theta \in \Theta\}$ la famille des distributions possibles et $\mathcal{U} = \{U_\gamma(\theta) | \gamma \in \Gamma\}$ la famille des fonctions de structure.

Si, en pratique, la spécification de $\mathcal{F}$ n'est pas évidente, celle de $\mathcal{U}$ l'est encore moins.

4.1. Choix de la loi de structure et de la loi conditionnelle

En pratique, les conditions suivantes doivent être vérifiées :

- La famille $\mathcal{U}$ doit être assez grande pour contenir les distributions qui peuvent décrire le collectif.
- La famille $\mathcal{U}$ doit être aussi petite qu'il est possible de représenter notre connaissance du collectif.
- Idéalement, les familles $\mathcal{U}$ et $\mathcal{F}$ devraient être choisies de manière à ce que la prime de Bayes puisse s'exprimer analytiquement.

4.1. Choix de la loi de structure et de la loi conditionnelle

Definition (Familles de distributions conjuguées)

La famille $\mathcal{U}$ est conjuguée à la famille $\mathcal{F}$ si, pour tout $\gamma \in \Gamma$ et pour toute réalisation $\mathbf{x}$ du vecteur des observations $\mathbf{X}$, il existe $\gamma' \in \Gamma$ tel que

$$U_\gamma(\theta|\mathbf{X} = \mathbf{x}) = U_{\gamma'}(\theta), \quad \text{pour tout } \theta \in \Theta.$$

Un des principaux intérêts d'utilisation des familles de distributions conjuguées est que la distribution *a posteriori* obtenue sur une période peut être utilisée comme *a priori* de la période suivante.

Ceci a notamment pour effet de simplifier les processus de mise à jour des primes : les méthodes mises en place peuvent être conservées pour les périodes suivantes.

4.1. Choix de la loi de structure et de la loi conditionnelle

Exemple (1/2)

Considérons une suite de n v.a. Bernoulli $N_1, N_2, ..., N_n$ qui ont toutes la même probabilité constante de succès Θ. Le nombre total de succès observé est donc distribué selon une loi binomiale $B(n, \Theta)$. Connaissant Θ, la probabilité d'observer r succès lors des n tirages vaut

$$\Pr[N_1 + ... + N_n = r|\Theta] = \binom{n}{r} \Theta^r (1 - \Theta)^{n-r}.$$

Supposons que la distribution *a priori* de Θ soit une loi $\beta(a, b)$ qui admet donc pour densité

$$f(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \theta^{a-1} (1-\theta)^{b-1}, \qquad \theta \in [0; 1].$$

4.1. Choix de la loi de structure et de la loi conditionnelle

Exemple (2/2)

D'après la règle de Bayes, la densité de la loi *a posteriori* de Θ est donnée par

$$\begin{aligned} f(\theta|n_1, ..., n_n) &\propto f(\theta)f(n_1, ..., n_n|\theta) = f(\theta)\prod_{i=1}^{n} f(n_i|\theta) \\ &\propto \theta^{a-1}(1-\theta)^{b-1}\theta^r(1-\theta)^{n-r} \\ &\propto \theta^{a+r-1}(1-\theta)^{b+n-r-1}. \end{aligned}$$

La distribution *a posteriori* de Θ est une loi $\beta(a+r, b+n-r)$.
Ainsi la famille des lois bêta est conjuguée à la famille des lois Bernoulli.

4.2. Famille exponentielle

Definition (Famille exponentielle)

Une distribution de probabilités appartient à la famille exponentielle $\mathcal{F}_{exp}$ si sa loi peut s'exprimer sous la forme :

$$dF(x) = \exp\left[\frac{x\theta - b(\theta)}{\sigma^2/\omega} + c(x, \sigma^2/\omega)\right] d\nu(x), \qquad x \in \mathcal{A} \subset \mathbf{R},$$

où ν désigne la mesure de Lebesgue ou la mesure de comptage, $b \in \mathcal{C}^2(\Theta, \mathbf{R})$, b' injective, $c : \mathbf{R}^2 \to \mathbf{R}$ et ω et σ^2 constantes.

Interprétation des paramètres :

- σ^2 est le paramètre de dispersion,
- ω le poids *a priori* associé à l'observation,
- θ le profil de risque.

4.2. Famille exponentielle

De nombreuses distributions usuelles font partie de la famille exponentielle, notamment :

- Poisson,
- Bernoulli,
- binomiale,
- gamma,
- gaussienne,
- inverse-gaussienne.

4.2. Famille exponentielle

Notons $\mathcal{F}_{exp}^{b,c}$ la famille associée aux fonctions b et c et considérons la famille $\mathcal{U}_{exp}^{b}$ des distributions ayant une densité de la forme

$$u_\gamma(\theta) = \exp\left[\frac{x_0\theta - b(\theta)}{\tau^2} + d\left(x_0, \tau^2\right)\right], \theta \in \Theta.$$

Théorème

La famille $\mathcal{U}_{exp}^{b}$ est conjuguée à $\mathcal{F}_{exp}^{b,c}$.

Remarques :

- x_0 et τ^2 sont des hyperparamètres,
- la famille conjuguée à $\mathcal{F}_{exp}^{b,c}$ ne dépend pas de c.

4.2. Famille exponentielle

Exemples de classes conjuguées dans la famille exponentielle :

- Poisson - gamma,
- binomial - bêta,
- gamma - gamma,
- geométrique - bêta,
- gausienne - gaussienne.

4.3. Prime de Bayes dans le cadre exponentiel

Supposons que, pour θ fixé, les composantes de $\mathbf{X} = (X_1, ..., X_n)'$ sont indépendantes, de distribution $F_\theta \in \mathcal{F}_{exp}^{b,c}$, toutes avec le même paramètre de dispersion σ^2 et avec les poids $\omega_j, j = 1, ..., n$.

Théorème

Pour la famille $\mathcal{F}_{exp}^{b,c}$ et sa famille conjuguée $\mathcal{U}_{exp}^{b}$,

$$P^{ind}(\theta) = b'(\theta), \qquad \mathbf{Var}\left[X_j|\Theta = \theta, \omega_j\right] = b''(\theta)\frac{\sigma^2}{\omega_j},$$

et sous certaines conditions de régularité,

$$P^{coll} = x_0,$$

$$P^{Bayes} = \alpha\bar{X} + (1-\alpha)P^{coll},$$

où $\bar{X} = \sum_{j=1}^{n} \frac{\omega_j}{\omega_\bullet} X_j$ *et* $\alpha = \frac{\omega_\bullet}{\omega_\bullet + \sigma^2/\tau^2}$.

La prime de Bayes est linéaire en les observations : on parlera de *crédibilité exacte*.

4.4. Construction de classes conjuguées

Théorème

Considérons les hypothèses suivantes :

(H1) Pour $\mathbf{x}$ *fixé, les fonctions de vraisemblance* $l(\theta) = f_\theta(\mathbf{x})$ *sont proportionnelles à un élément de* $\mathcal{U}$*, i.e.*

$$\forall \mathbf{x}, \exists u_\mathbf{x} \in \mathcal{U}, u_\mathbf{x}(\theta) = f_\theta(\mathbf{x}) \left(\int f_\theta(\mathbf{x}) d\theta \right)^{-1}.$$

(H2) $\mathcal{U}$ *est fermée sous l'opérateur produit, i.e.*

$$\forall u, v \in \mathcal{U}, u(.)v(.) \left(\int u(\theta)v(\theta) d\theta \right)^{-1} \in \mathcal{U}.$$

Si $\mathcal{F}$ *et* $\mathcal{U}$ *satisfont (H1) et (H2), alors* $\mathcal{U}$ *est conjuguée à* $\mathcal{F}$*.*

Démonstration : La distribution *a posteriori* vaut :

$$u(\theta|\mathbf{x}) = \frac{f_\theta(\mathbf{x})u(\theta)}{\int f_\theta(\mathbf{x})u(\theta)d\theta} = \frac{\int f_\theta(\mathbf{x})d\theta}{\int f_\theta(\mathbf{x})u(\theta)d\theta} u_\mathbf{x}(\theta)u(\theta) \in \mathcal{U}. \qquad \square$$

4.4. Construction de classes conjuguées

En considérant le théorème 5, pour construire une classe conjuguée à $\mathcal{F}$, il est naturel de construire

$$\mathcal{U}' = \left\{ u_{\mathbf{x}} : u_{\mathbf{x}}(\theta) = f_\theta(\mathbf{x}) \left(\int f_\theta(\mathbf{x}) d\theta \right)^{-1}, \mathbf{x} \in A \right\}.$$

Si $\mathcal{U}'$ est fermée sous l'opération produit alors $\mathcal{U}'$ est conjuguée à $\mathcal{F}$; sinon on lui recherchera une extension naturelle qui l'est.

4.4. Construction de classes conjuguées

Exemple : Considérons $\mathcal{F}$ la famille des distributions binomiales, $\mathcal{F} = \{f_\theta(\mathbf{x}) : \theta \in [0;1]\}$ avec

$$f_\theta(\mathbf{x}) = C_n^{x_\bullet} \theta^{x_\bullet} (1-\theta)^{n-x_\bullet}, \quad x_\bullet = \sum_{j=1}^{n} x_j, \quad x_j \in \{0;1\}.$$

$\mathcal{U}'$ est l'ensemble des distributions bêta de paramètre $(i, n+2-i)$, $i \in \{1, ..., n+1\}$.

$\mathcal{U}'$ n'est pas fermée sous l'opération produit mais $\mathcal{U}$ l'ensemble des distributions bêta en est une extension naturelle qui est fermée sous l'opération produit (*exercice*). Donc grâce au théorème 5, $\mathcal{U}$ est conjuguée à $\mathcal{F}$.

Exercices

Exercice 2.1 Considérons un portefeuille regroupant exclusivement des "bons" et des "mauvais" risques. Notons B l'événement "être un bon risque" et B^c son complémentaire "être un mauvais risque". Supposons que l'assureur ne sache pas, *a priori* distinguer les bons des mauvais risques mais qu'il sait qu'ils sont en proportions égales dans son portefeuille ($\Pr[B] = 1/2$). Notons N_k le nombre de sinistres causés par un assuré au cours de l'année k et supposons que

$$\Pr[N_k = 1|B] = 0,2 = 1 - Pr[N_k = 0|B]$$
$$\Pr[N_k = 1|B^c] = 0,8 = 1 - Pr[N_k = 0|B^c]$$

Supposons enfin que les sinistres sont tous de même montant 1 et que conditionnellement à la qualité du risque, les nombres de sinistres sont indépendants.
1. Déterminez la prime pure *a priori* d'un assuré de ce portefeuille.
2. Considérons un assuré avec l'historique de sinistres : $(N_1 = 0, N_2 = 1, N_3 = 0)$. Quelle est la probabilité qu'il s'agisse d'un bon risque ?
3. Quelle prime exigeriez-vous de cet assuré pour la 4^e année ?

Exercices

Exercice 2.2 Considérons un assuré dont le nombre annuel de sinistres est distribué selon une loi de Poisson de paramètre Θ. Les montants de sinistres sont constants. La distribution *a priori* de Θ est une loi uniforme sur l'intervalle $[0, 1]$.
1. Donnez la prime de Bayes de cet assuré.
2. Lors de la première année d'observation, l'assuré a causé un sinistre. Quelle prime lui réclameriez-vous pour la seconde année ?

Exercice 2.3 Notons N_i le nombre de sinistres causés par un assuré durant l'année i. Conditionnellement à Θ, les N_i sont indépendants et identiquement distribués selon une loi de Poisson de paramètre Θ. Θ admet pour densité $g(\theta) = e^{-\theta}, \theta > 0$.
1. Donnez la probabilité que l'assuré cause 2 sinistres durant la pemière année $\Pr[N_1 = 2]$.
2. Donnez la distribution prédictive de N_2 sachant que $N_1 = 2$ (il s'agit de déterminer $\Pr[N_2 = n | N_1 = 2]$ pour $n = 0, 1, ...$)

Exercices

Exercice 2.4 Considérons un assuré qui

- a exactement un sinistre par an,
- dont le coût est distribuée selon une loi exponentielle de densité $f(x|T=t) = te^{-tx}, x > 0$,
- où le paramètre T admet pour densité $h(t) = te^{-t}, t > 0$.

1. Donnez la prime de Bayes de cet assuré pour la première période.
2. Sachant qu'il a eu un sinistre de montant 5 la première année, déterminez la densité *a posteriori* de T.

Exercice 2.5 La distribution *a priori* d'une v.a. H est donnée par $\Pr[H = 0,25] = 0,8$ et $\Pr[H = 0,5] = 0,2$. Conditionnellement à H les v.a. D_i sont i.i.d. Le résultat d'une expérience D_i est distribué selon

$$\Pr[D_i = d|H = h] = h^d(1-h)^{1-d},$$

pour $d \in \{0;1\}$.
1. Sachant $D_1 = 1$, exprimez la loi *a posteriori* de H.
2. Sachant $D_1 = 1$, quelle est la distribution prédictive du résultat de l'expérience D_2 ?

Sommaire

Contexte

- La prime de Bayes $\widetilde{\mu(\Theta)} = \mathbf{E}\left[\mu(\Theta)|\mathbf{X}\right]$ est le meilleur estimateur possible de la prime individuelle $\mu(\Theta) = \mathbf{E}\left[X_{n+1}|\Theta\right]$.
- En général, elle ne peut s'exprimer de manière analytique ; son calcul nécessite alors l'utilisation de méthodes numériques.
- Sa détermination nécessite en outre de spécifier une distribution conditionnelle et une distribution *a priori*.

La théorie de la crédibilité repose sur l'idée de se restreindre à la classe des estimateurs qui sont linéaires en les observations.

Comme dans l'approche bayesienne, la qualité de ces estimateurs sera appréciée par le critère d'erreur quadratique moyenne.

1. Prime de crédibilité

1.1. Introduction dans un contexte simplifié

Hypothèse (Modèle de crédibilité simplifié)

(H1) Les variables aléatoires X_j ($j = 1, ..., n$) sont, conditionnellement à $\Theta = \theta$, indépendantes et identiquement distribuées selon une loi F_θ avec les moments conditionnels

$$\mu(\theta) = \mathbf{E}\left[X_j | \Theta = \theta\right],$$
$$\sigma^2(\theta) = \mathbf{Var}\left[X_j | \Theta = \theta\right].$$

(H2) Θ est une variable aléatoire de distribution $U(\theta)$.

Dans ce modèle, on a

$$P^{ind} = \mu(\Theta) = \mathbf{E}\left[X_{n+1} | \Theta\right],$$
$$P^{coll} = \mu_0 = \int_\Theta \mu(\theta) dU(\theta).$$

1.1. Introduction dans un contexte simplifié

On souhaite trouver un estimateur de la prime individuelle $\mu(\Theta)$ et l'on va se restreindre à ceux qui sont linéaires en les observations $\mathbf{X} = (X_1, ..., X_n)'$.

Notons P^{cred} ou $\widehat{\widehat{\mu(\Theta)}}$ le meilleur estimateur linéaire en $\mathbf{X}$ de $\mu(\Theta)$.

L'estimateur de crédibilité $\widehat{\widehat{\mu(\Theta)}}$ est de la forme

$$\widehat{\widehat{\mu(\Theta)}} = \hat{a}_0 + \sum_{j=1}^{n} \hat{a}_j X_j,$$

où les coefficients $(\hat{a}_0, \hat{a}_1, ..., \hat{a}_n)$ sont solution de

$$(\hat{a}_0, \hat{a}_1, ..., \hat{a}_n) = \arg\min \mathbf{E}\left[\left(\mu(\Theta) - a_0 - \sum_{j=1}^{n} a_j X_j\right)^2\right].$$

1.1. Introduction dans un contexte simplifié

La distribution de $\mathbf{X}$ est invariante par permutations des X_j, donc

$$\hat{a}_1 = ... = \hat{a}_n.$$

Aussi, l'estimateur de crédibilité est de la forme

$$\widehat{\widehat{\mu(\Theta)}} = \hat{a} + \hat{b}\bar{X},$$

où $\hat{a}$ et $\hat{b}$ sont solution de

$$\left(\hat{a}, \hat{b}\right) = \arg\min \mathbf{E}\left[\left(\mu(\Theta) - a - b\bar{X}\right)^2\right].$$

1.1. Introduction dans un contexte simplifié

Les e.d.p. de ce problème sont

$$\begin{aligned} &\mathbf{E}\left[\mu(\Theta) - a - b\bar{X}\right] = 0, \\ &\mathbf{Cov}\left(\bar{X}, \mu(\Theta)\right) - b\mathbf{Var}\left(\bar{X}\right) = 0. \end{aligned}$$

De par la structure de dépendance du modèle, on a

$$\begin{aligned} \mathbf{Cov}\left(\bar{X}, \mu(\Theta)\right) &= \mathbf{Var}\left[\mu(\Theta)\right] =: \tau^2, \\ \mathbf{Var}\left(\bar{X}\right) &= \frac{\mathbf{E}\left[\sigma^2(\Theta)\right]}{n} + \mathbf{Var}\left[\mu(\Theta)\right] =: \frac{\sigma^2}{n} + \tau^2, \end{aligned}$$

d'où l'on obtient

$$\begin{aligned} b &= \frac{\tau^2}{\tau^2 + \sigma^2/n} = \frac{n}{n + \sigma^2/\tau^2}, \\ a &= (1-b)\mu_0. \end{aligned}$$

1.1. Introduction dans un contexte simplifié

Théorème

Sous les hypothèses (H1) et (H2), l'estimateur de crédibilité est donné par

$$\widehat{\widehat{\mu(\Theta)}} = \alpha\bar{X} + (1-\alpha)\mu_0,$$

où $\mu_0 = \mathbf{E}\left[\mu(\Theta)\right]$ *et* $\alpha = \frac{n}{n+\sigma^2/\tau^2}$.

Remarques :

- On retrouve la même structure que pour la prime de Bayes dans le cadre exponentiel.
- P^{cred} est une moyenne pondérée de la prime collective et de la prime individuelle observée.

1.1. Introduction dans un contexte simplifié

Remarques :

- Le coefficient $\kappa := \sigma^2/\tau^2$ est appelé *coefficient de crédibilité*. En l'écrivant sous la forme
$$\kappa = \left(\frac{\sigma/\mu_0}{\tau/\mu_0}\right)^2,$$
et en remarquant que
$$\frac{\sigma}{\mu_0} = \frac{\sqrt{\mathbf{E}\left[\mathbf{Var}(X_j|\Theta)\right]}}{\mathbf{E}[X_j]},$$
$$\frac{\tau}{\mu_0} = \mathbf{CoVa}[\mu(\Theta)].$$
$\sqrt{\kappa}$ s'interprète comme le rapport entre la variabilité interne du risque et l'hétérogénéité du portefeuille.

1.1. Introduction dans un contexte simplifié

Remarques :

- Le *facteur de crédibilité* α croît avec le nombre d'années d'obervations *n*, croît avec l'hétérogénéité du portefeuille (mesurée par $\frac{\tau}{\mu_0}$) et décroît avec la variabilité interne au risque (mesurée par $\frac{\sigma}{\mu_0}$).

- Lorsque la prime de Bayes est linéaire, elle coïncide avec la prime de crédibilité, on parle alors de *crédibilité exacte*.

- La prime de crédibilité fait intervenir les paramètres de structure τ^2, σ^2 et μ_0. Ces paramètres pourront être estimés empiriquement sur le collectif ou fixés *a priori* (avis d'expert).

1.2. Interprétation de la prime de crédibilité

P^{cred} est une moyenne pondérée de P^{coll} et de $\bar{X}$.

- $P^{coll} = \mu_0$ est le meilleur estimateur *a priori*, son erreur quadratique moyenne vaut
$$\mathbf{E}\left[(\mu_0 - \mu(\Theta))^2\right] = \mathbf{Var}[\mu(\Theta)] = \tau^2.$$

- $\bar{X}$ est le meilleur estimateur linéaire et individuellement non-biaisé basé uniquement sur les observations, son erreur quadratique moyenne vaut
$$\mathbf{E}\left[(\bar{X} - \mu(\Theta))^2\right] = \mathbf{E}\left[\frac{\sigma^2(\Theta)}{n}\right] = \frac{\sigma^2}{n}.$$

P^{cred} est une moyenne pondérée dont les poids sont inversement proportionnels à l'erreur quadratique moyenne de ses deux composantes.

1.3. Erreur quadratique moyenne de la prime de crédibilité

Théorème

L'erreur quadratique moyenne de l'estimateur de crédibilité $\widehat{\widehat{\mu(\Theta)}}$ vaut

$$\mathbf{E}\left[\left(\widehat{\widehat{\mu(\Theta)}} - \mu(\Theta)\right)^2\right] = (1-\alpha)\tau^2 = \alpha\frac{\sigma^2}{n}.$$

Démonstration : *exercice.*

Remarques :

- α est le coefficient par lequel est réduite l'erreur quadratique moyenne lorsque l'on utilise P^{cred} au lieu de P^{coll}.
- $(1-\alpha)$ est le coefficient par lequel est réduite l'erreur quadratique moyenne lorsque l'on utilise P^{cred} au lieu de $\bar{X}$.

1.3. Erreur quadratique moyenne de la prime de crédibilité

Sous les hypothèses (H1) et (H2), on a

Prime	Erreur quadratique moyenne
$P^{coll} = \mu_0$	$\mathbf{Var}\left[\mu(\Theta)\right]$
$P^{cred} = \widehat{\widehat{\mu(\Theta)}}$	$(1-\alpha)\mathbf{Var}\left[\mu(\Theta)\right]$
$P^{Bayes} = \widetilde{\mu(\Theta)}$	$\mathbf{E}\left[\mathbf{Var}\left(\mu(\Theta)\right)\|\mathbf{X}\right]$

Le principal intérêt de la prime de crédibilité réside dans sa simplicité de mise en oeuvre : il n'est pas nécessaire de spécifier une structure (choix conjoint de l'*a priori* et de la distribution conditionnelle) qui ajoute un risque additionnel de modèle.

En contrepartie, la précision de la prime de crédibilité est moindre que celle de la prime de Bayes.

2. Modèle de Bühlmann

2. Modèle de Bühlmann

Jusqu'à présent, on s'est intéressé la tarification d'un risque particulier. On va dorénavant considérer un portefeuille de I risques similaires.

On notera $\mathbf{X}_i = (X_{i1}, ..., X_{in})'$ le vecteur des observations associé au risque i et Θ_i son profil de risque.

Ces I risques sont similaires dans la mesure où chaque risque i respecte les hypothèses (H1) et (H2).

On se retrouve ainsi dans le cadre décrit par Bühlmann [1967].

2.1. Hypothèses du modèle

Hypothèse (Modèle de Bühlmann)

(B1) Les variables aléatoires X_{ij} $(j = 1, ..., n)$ *sont, conditionnellement à* $\Theta_i = \theta$, *indépendantes et identiquement distribuées selon une loi* F_θ *avec les moments conditionnels*

$$\mu(\theta) = \mathbf{E}\left[X_{ij} | \Theta_i = \theta\right],$$
$$\sigma^2(\theta) = \mathbf{Var}\left[X_{ij} | \Theta_i = \theta\right].$$

(B2) Les couples $(\Theta_1, \mathbf{X}_1), ..., (\Theta_I, \mathbf{X}_I)$ *sont indépendants et identiquement distribués.*

Du fait de (B2), le portefeuille est hétérogène : les profils de risque $\Theta_1, ..., \Theta_I$ sont indépendants et ont la même distribution de structure $U(\theta)$.
Ces risques ont néanmoins en commun qu'on ne peut les différencier *a priori*.

2.1. Hypothèses du modèle

L'objectif est de trouver l'estimateur de crédibilité dans le modèle de Bühlmann. Cela amène deux questions :

- **Que cherche-t-on à estimer ?**
  Pour chaque risque individuel i, on souhaite estimer la prime individuelle $\mu(\Theta_i)$ par l'estimateur $\widehat{\widehat{\mu(\Theta_i)}}$ qui est linéaire en les observations. On estime donc I primes.
- **En quelles observations l'estimateur doit-il être linéaire ?**
  On utilisera toute l'information disponible (pas uniquement le passé sinistres du risque i) et l'on supposera que $\widehat{\widehat{\mu(\Theta_i)}}$ est le meilleur estimateur de la classe

$$\left\{ \widehat{\mu(\Theta_i)} : \widehat{\mu(\Theta_i)} = a + \sum_{k=1}^{I} \sum_{j=1}^{n} b_{kj} X_{kj} \ ; \ a, b_{kj} \in \mathbf{R} \right\}.$$

2.1. Hypothèses du modèle

La distribution de $\mathbf{X}_i$ est invariante par permutations des X_{ij} $(j = 1, \ldots, n)$, donc

$$\widehat{\widehat{\mu(\Theta_i)}} = \hat{a}_0^{(i)} + \sum_{k=1}^{I} \hat{b}_k^{(i)} \bar{X}_k,$$

où $\bar{X}_k = \frac{1}{n} \sum_{j=1}^{n} X_{kj}$ et

$$\left(\hat{a}_0^{(i)}, \hat{b}_k^{(i)}\right) = \arg\min \mathbf{E}\left[\left(\mu(\Theta_i) - a_0^{(i)} - \sum_{k=1}^{I} b_k^{(i)} \bar{X}_k\right)^2\right].$$

Les e.d.p. par rapport à $a_0^{(i)}$ et aux $b_k^{(i)}$ donnent (*exercice*)

$$\mathbf{Cov}\left(\mu(\Theta_i), \bar{X}_k\right) = b_k^{(i)} \mathbf{Var}\left[\bar{X}_k\right].$$

2.1. Hypothèses du modèle

D'après (B2), **Cov** $\left(\mu(\Theta_i), \bar{X}_k\right) = 0$ si $i \neq k$.

L'estimateur de crédibilité de $\mu(\Theta_i)$ ne dépend donc que des observations associées au risque i. Il est de la forme

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha \bar{X}_i + (1-\alpha)\mu_0,$$

où $\alpha = \frac{n}{n+\sigma^2/\tau^2}$.

Bien qu'ayant la même forme, ce résultat est plus fort que le précédent car l'estimateur est ici basé sur l'ensemble des observations du portefeuille $\mathbf{X} = (\mathbf{X}_1', ..., \mathbf{X}_I')'$.

2.2. Estimateur de crédibilité homogène

Definition (Estimateur homogène de crédibilité)

L'estimateur homogène de crédibilité $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$ de $\mu(\Theta_i)$ est le meilleur estimateur de la classe des estimateurs collectivement sans biais

$$\left\{\widehat{\mu(\Theta_i)} : \widehat{\mu(\Theta_i)} = \sum_{k=1}^{I}\sum_{j=1}^{n} b_{kj}X_{kj} \;;\; \mathbf{E}\left[\widehat{\mu(\Theta_i)}\right] = \mathbf{E}\left[\mu(\Theta_i)\right], b_{kj} \in \mathbf{R}\right\}.$$

Remarquons que

$$\mathbf{E}\left[\widehat{\widehat{\mu(\Theta_i)}}^{hom}\right] = \mathbf{E}\left[\mu(\Theta_i)\right] = \mathbf{E}\left[X_{kj}\right] = \mu_0,$$

pour tous i, j et k. On en déduit donc la contrainte sur les paramètres

$$\sum_{k=1}^{I}\sum_{j=1}^{n} b_{kj} = 1.$$

2.2. Estimateur de crédibilité homogène

Théorème

Dans le modèle de Bühlmann, l'estimateur homogène de crédibilité est donné par

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha \bar{X}_i + (1-\alpha)\bar{X},$$

où $\alpha = \frac{n}{n+\sigma^2/\tau^2}$.

Démonstration : *exercice.*

Remarque : Le terme $\bar{X}$ est l'estimateur naturel de μ_0.

2.3. Estimation des paramètres de structure

Théorème

Les estimateurs

$$\widehat{\sigma^2} = \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{j=1}^{n} \left(X_{ij} - \bar{X}_i\right)^2,$$

$$\widehat{\tau^2} = \frac{1}{I-1} \sum_{i=1}^{I} (\bar{X}_i - \bar{X})^2 - \frac{\widehat{\sigma^2}}{n},$$

sont sans biais et convergents.

Exercices

Exercice 3.1 Considérons un assuré dont le nombre annuel de sinistres est distribué selon une loi de Poisson de paramètre Θ. La distribution *a priori* de Θ est une loi uniforme sur l'intervalle $[0, 1]$.
Lors de la première année d'observation, l'assuré a causé un sinistre. Utilisez le modèle de Bühlmann pour estimer le nombre de sinistres espéré que va engendrer cet assuré pour la deuxième année puis comparez le résultat avec celui obtenu dans l'exercice 2.2.

Exercice 3.2 Une société d'assurance couvre deux contrats depuis 3 ans. Elle dispose des montants annuels de sinistres suivants :

Contrat	Année 1	Année 2	Année 3
1	5	8	11
2	11	13	12

Selon le modèle de Bühlmann, quelles primes réclameriez-vous à ces deux assurés pour la 4^e année ?

Exercices

Exercice 3.3 Un portefeuille de 340 assurés a produit 210 déclarations de vol au cours d'une année. La répartition des sinistres est la suivante :

Nombre de sinistres	Nombre d'assurés
0	200
1	80
2	50
3	10

En supposant que le nombre de sinistres déclarés par un assuré suit une loi de Poisson dont le paramètre peut varier d'un assuré à l'autre, déterminez le facteur de crédibilité (selon le modèle de Bühlmann) pour un assuré de ce portefeuille. Déduisez-en l'augmentation de la prime d'un assuré qui aurait déclaré deux sinistres.

Exercice 3.4 Soit N_i le nombre de sinistres (de montant 1) causés par un assuré au cours de l'année i. Conditionnellement à $\Theta = \theta$, les N_i sont supposés indépendantes et distribuées selon une loi de Poisson de paramètre θ. Θ est supposé de loi gamma de moyenne a/τ et de variance a/τ^2.

1. Un assuré a produit n_1, n_2 et n_3 sinistres les trois premières années. Calculez sa prime de Bayes pour la quatrième année.
2. Comparez avec la prime qui lui serait réclamée selon le modèle de Bühlmann et commentez.

Sommaire

Motivation

Considérons la situation d'un portefeuille de contrats qui doivent être tarifés. Ce portefeuille a fait l'objet d'une segmentation qui a conduit à la création de classes de risque à l'intérieur desquelles les risques ne peuvent être différenciés *a priori*.

La théorie de la crédibilité va permettre, pour chaque risque, d'enrichir l'information *a priori* grâce au passé sinistre.

Dans cette optique, le modèle de Bühlmann utilise une hypothèse peu réaliste :

$$\sigma^2(\Theta_i) = \mathbf{Var}\left[X_{ij}|\Theta_i\right].$$

Il apparaît en effet assez naturel que cette variance conditionnelle devrait être décroissante avec l'exposition du risque i.

Motivation

Le modèle de Bühlmann-Straub vient généraliser le modèle de Bühlmann en relachant cette hypothèse par l'introduction de poids.

Ce modèle est largement utilisé en pratique que ce soit en assurance non-vie, en assurance vie ou en réassurance.

Dans la suite, nous traiterons un portefeuille de I risques (ou classes de risques) et nous noterons X_{ij} le ratio de sinistres (ou la fréquence de sinistres ou la taille moyenne des sinistres ou tout autre indicateur) du risque i pendant l'année j et ω_{ij} le poids associé.

Démonstration des résultats

La plupart des résultats de ce chapitre se démontrent de la même manière que dans le modèle de Bühlmann (cf. Chapitre 3 : Estimateurs de crédibilité).

Pour les autres, le lecteur pourra se référer à Bühlmann & Gisler [2005] pour le détail des preuves.

1. Hypothèses du modèle

1. Hypothèses du modèle

Hypothèse (Modèle de Bühlmann-Straub)

Le risque i est caractérisé par le profil de risque θ_i qui est une réalisation de Θ_i et l'on a

(BS1) Les variables aléatoires X_{ij} $(j = 1, ..., n)$ sont, conditionnellement à Θ_i, indépendantes avec les moments conditionnels

$$\mu(\Theta_i) = \mathbf{E}\left[X_{ij}|\Theta_i\right],$$
$$\sigma^2(\Theta_i) = \omega_{ij}\mathbf{Var}\left[X_{ij}|\Theta_i\right].$$

(BS2) Les couples $(\Theta_1, \mathbf{X}_1), ..., (\Theta_I, \mathbf{X}_I)$ sont indépendants et les $\Theta_1, ..., \Theta_I$ sont indépendants et identiquement distribués.

Remarques générales :

- Le nombre d'années d'observations peut varier entre les risques. Formellement cela passe par l'introduction de poids nuls.
- L'indépendance conditionnelle dans (BS1) peut être relachée en exigeant seulement la non-corrélation conditionnelle, *i.e.* $\mathbf{E}\left[\mathbf{Cov}\left(X_{ik}, X_{il}|\Theta_i\right)\right] = 0$ pour $k \neq l$.

1. Hypothèses du modèle

Remarques sur (BS2) :

- Les risques sont indépendants : les sinistres engendrés par deux contrats différents sont indépendants.
- Les risques sont *a priori* égaux : cette hypothèse doit être testée. Bühlmann & Gisler [2005] proposent un ajustement du modèle dans le cas contraire.

Remarques sur (BS1) :

- Pour chaque risque, le "vrai" ratio de sinistres $\mu(\Theta_i)$ n'évolue pas au cours du temps : hypothèse classique en assurance non-vie, elle nécessitera le plus souvent l'élaboration préalable de données *as-if*.
- La variance conditionnelle est inversement proportionnelle aux poids que l'on se donne. En pratique les poids sont des mesures du volume de risque (durée d'exposition au risque, montants assurés en incendie, masse salariale en santé, chiffre d'affaire réalisé par la cédante en réassurance, etc.)

1. Hypothèses du modèle

Rappel des notations :

- Prime individuelle, $\mu(\Theta_i) := \mathbf{E}[X_{ij}|\Theta_i]$,
- Risque individuel normalisé ($\omega_{ij} = 1$), $\sigma^2(\Theta_i) := \omega_{ij}\mathbf{Var}[X_{ij}|\Theta_i]$,
- Prime collective, $\mu_0 := \mathbf{E}[\mu(\Theta_i)]$,
- Risque collectif normalisé ($\omega_{ij} = 1$),

$$\begin{aligned}\mathbf{Var}[X_{ij}] &= \mathbf{Var}[\mathbf{E}[X_{ij}|\Theta_i]] + \mathbf{E}[\mathbf{Var}[X_{ij}|\Theta_i]] \\ &= \underbrace{\mathbf{Var}[\mu(\Theta_i)]}_{=:\tau^2} + \underbrace{\mathbf{E}[\sigma^2(\Theta_i)]}_{=:\sigma^2}.\end{aligned}$$

σ^2 est une mesure du risque interne au risque individuel alors que τ^2 mesure l'hétérogénéité du portefeuille.

2. Estimateurs de crédibilité

2.1. Estimateur non-homogène

Pour chaque risque i, on cherche l'estimateur de crédibilité $\widehat{\widehat{\mu(\Theta_i)}}$ de son ratio de sinistres individuel $\mu(\Theta_i)$.

On dispose de données $\mathbf{X} = (\mathbf{X}_1', ..., \mathbf{X}_I')'$ sur l'ensemble du portefeuille.

Les risques étant indépendants (BS2), $\widehat{\widehat{\mu(\Theta_i)}}$ dépend uniquement des observations $\mathbf{X}_i$ rattachées au risque i, soit

$$\widehat{\widehat{\mu(\Theta_i)}} = a_{i0} + \sum_j a_{ij} X_{ij}.$$

2.1. Estimateur non-homogène

Par ailleurs, en notant $\omega_{i\bullet} := \sum_{j=1}^{n} \omega_{ij}$, le meilleur estimateur individuellement sans biais de $\mu(\Theta_i)$ est

$$X_i := \sum_j \frac{\omega_{ij}}{\omega_{i\bullet}} X_{ij},$$

pour lequel on a (*exercice*)

$$\begin{aligned}
\mathbf{E}[X_i|\Theta_i] &= \mu(\Theta_i), \\
\mathbf{Var}[X_i|\Theta_i] &= \sum_j \left(\frac{\omega_{ij}}{\omega_{i\bullet}}\right)^2 \mathbf{Var}[X_{ij}|\Theta_i] = \frac{\sigma^2(\Theta_i)}{\omega_{i\bullet}}, \\
\mathbf{Var}[X_i] &= \frac{\sigma^2}{\omega_{i\bullet}} + \tau^2.
\end{aligned}$$

2.1. Estimateur non-homogène

On prouve (cf. Bühlmann & Gisler [2005]) que l'estimateur de crédibilité ne dépend des observations que par X_i. Comme de plus, $\mathbf{E}[X_i] = \mu_0$, l'estimateur de crédibilité est de la forme

$$\widehat{\widehat{\mu(\Theta_i)}} = \alpha_i X_i + (1 - \alpha_i)\mu_0,$$

et doit satisfaire

$$\mathbf{Cov}\left(\widehat{\widehat{\mu(\Theta_i)}}, X_i\right) = \alpha_i \mathbf{Cov}(X_i, X_i) = \mathbf{Cov}\left(\mu(\Theta_i), X_i\right).$$

Donc

$$\begin{aligned} \alpha_i &= \frac{\mathbf{Cov}\left(\mu(\Theta_i), X_i\right)}{\mathbf{Var}[X_i]} \\ &= \frac{\mathbf{E}\left[\mathbf{Cov}\left(\mu(\Theta_i), X_i | \Theta_i\right)\right] + \mathbf{Cov}\left(\mu(\Theta_i), \mathbf{E}\left[X_i | \Theta_i\right]\right)}{\mathbf{Var}[X_i]} \\ &= \frac{0 + \tau^2}{\sigma^2 / \omega_{i\bullet} + \tau^2} = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \frac{\sigma^2}{\tau^2}}. \end{aligned}$$

2.1. Estimateur non-homogène

Théorème

Dans le modèle de Bühlmann-Straub, l'estimateur (non-homogène) de crédibilité est donné par

$$\widehat{\mu(\Theta_i)} = \alpha_i X_i + (1-\alpha_i)\mu_0,$$

où

$$X_i = \sum_j \frac{\omega_{ij}}{\omega_{i\bullet}} X_{ij},$$

$$\omega_{i\bullet} = \sum_j \omega_{ij},$$

$$\alpha_i = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \frac{\sigma^2}{\tau^2}} = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \kappa}.$$

κ est appelé *coefficient de crédibilité*.

2.1. Estimateur non-homogène

Théorème

L'erreur quadratique moyenne de l'estimateur de crédibilité est donnée par

$$\mathbf{E}\left[\left(\widehat{\widehat{\mu(\Theta_i)}} - \mu(\Theta_i)\right)^2\right] = (1-\alpha_i)\tau^2 = \alpha_i \frac{\sigma^2}{\omega_{i\bullet}}.$$

Remarques :

- τ^2 est l'erreur quadratique moyenne de μ_0,
- $\frac{\sigma^2}{\omega_{i\bullet}}$ est l'erreur quadratique moyenne de X_i (*exercice*).

2.2. Estimateur homogène

Théorème

Dans le modèle de Bühlmann-Straub, l'estimateur homogène de crédibilité est donné par

$$\widehat{\widehat{\mu(\Theta_i)}}^{hom} = \alpha_i X_i + (1-\alpha_i)\widehat{\widehat{\mu_0}},$$

où

$$\widehat{\widehat{\mu_0}} = \sum_{i=1}^{I} \frac{\alpha_i}{\alpha_\bullet} X_i,$$

$$\alpha_i = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \frac{\sigma^2}{\tau^2}},$$

$$\alpha_\bullet = \sum_{i=1}^{I} \alpha_i.$$

2.2. Estimateur homogène

Remarques :

- Pour estimer μ_0, il eut été intuitif d'utiliser la moyenne observée $\bar{X} = \sum \frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}} X_i$. On lui préfèrera la moyenne pondérée par les facteurs de crédibilité $\widehat{\widehat{\mu_0}} = \sum \frac{\alpha_i}{\alpha_\bullet} X_i$.
- Il est possible d'avoir simultanément $\widehat{\widehat{\mu(\Theta_i)}}^{hom} > X_i \left(\Leftrightarrow \widehat{\widehat{\mu_0}} > X_i\right)$ et $\bar{X} < X_i$.
- À la différence de l'estimateur de crédibilité (non-homogène), l'estimateur homogène utilise l'ensemble des observations disponibles.

2.2. Estimateur homogène

Théorème

Sous les hypothèses du modèle de Bühlmann-Straub, on a

$$\sum_{ij} \omega_{ij} \widehat{\widehat{\mu(\Theta_i)}}^{hom} = \sum_{ij} \omega_{ij} X_{ij}.$$

Démonstration : *exercice.*

Ce résultat nous indique que si la prime homogène de crédibilité avait été utilisée dans la passé, il y aurait eu équilibre entre primes et prestations.

2.2. Estimateur homogène

Théorème

Sous les hypothèses du modèle de Bühlmann-Straub, l'erreur quadratique moyenne de l'estimateur homogène de crédibilité est donnée par

$$\mathbf{E}\left[\left(\widehat{\widehat{\mu(\Theta_i)}}^{hom} - \mu(\Theta_i)\right)^2\right] = \tau^2(1-\alpha_i)\left(1+\frac{1-\alpha_i}{\alpha_\bullet}\right).$$

Remarque : L'erreur quadratique moyenne de l'estimateur homogène de crédibilité est égale à celle de l'estimateur (non-homogène) majorée par le coefficient $\frac{1-\alpha_i}{\alpha_\bullet}$. Cette perte de précision résulte de l'estimation de μ_0.

2.3. Estimation des paramètres de structure

L'estimation de la prime de crédibilité passe par l'estimation des paramètres de structure τ^2 et σ^2.

Rappelons que ces grandeurs interviennent :

- dans le calcul des facteurs de crédibilité pour l'estimation de $\widehat{\widehat{\mu(\Theta_i)}}$ et de $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$,
- dans l'estimation de μ_0 préalable à celle de $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$.

Leur estimation nécessite donc une attention particulière.

2.3. Estimation des paramètres de structure

Théorème

Les estimateurs

$$\widehat{\sigma^2} := \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{j=1}^{n} \omega_{ij} \left(X_{ij} - X_i\right)^2,$$

$$\widehat{\widehat{\tau^2}} := \frac{\omega_{\bullet\bullet}}{\omega_{\bullet\bullet}^2 - \sum_{i=1}^{I} \omega_{i\bullet}^2} \left\{ \sum_{i=1}^{I} \omega_{i\bullet} (X_i - \bar{X})^2 - (I-1)\widehat{\sigma^2} \right\},$$

sont sans biais et convergents (pour autant que $\sum_i \left(\frac{\omega_{i\bullet}}{\omega_{\bullet\bullet}}\right)^2 \to 0$ *lorsque* $I \to \infty$ *pour* $\widehat{\widehat{\tau^2}}$*).*

Remarque : L'estimateur $\widehat{\widehat{\tau^2}}$ peut donner une valeur négative ; en pratique, on utilisera donc $\widehat{\tau^2} = \max\left\{\widehat{\widehat{\tau^2}}, 0\right\}$.

2.4. Estimateur empirique de crédibilité

En pratique, on utilisera *l'estimateur de crédibilité empirique* $\widehat{\widehat{\mu(\Theta_i)}}^{emp}$ qui est obtenu en remplaçant les paramètres de structure par leurs estimateurs dans la formule de $\widehat{\widehat{\mu(\Theta_i)}}^{hom}$.

Théorème

L'estimateur empirique de crédibilité est donné par

$$\widehat{\widehat{\mu(\Theta_i)}}^{emp} := \widehat{\alpha}_i X_i + (1 - \widehat{\alpha}_i)\widehat{\widehat{\mu_0}},$$

où

$$\widehat{\alpha}_i := \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \widehat{\sigma}^2/\widehat{\tau}^2},$$

$$\widehat{\widehat{\mu_0}} := \frac{\sum \widehat{\alpha}_i X_i}{\sum \widehat{\alpha}_i}.$$

3. Récapitulatif de la démarche

3. Récapitulatif de la démarche

Récapitulatif de la démarche à suivre en pratique

1. Détermination des poids ω_{ij}
2. Calcul des $X_i = \sum_j \frac{\omega_{ij}}{\omega_{i\bullet}} X_{ij}$
3. Estimation de σ^2 par $\widehat{\sigma^2}$
4. Estimation de τ^2 par $\widehat{\tau^2}$
5. Estimation des α_i par $\widehat{\alpha}_i = \frac{\omega_{i\bullet}}{\omega_{i\bullet} + \widehat{\sigma}^2 / \widehat{\tau}^2}$
6. Estimation de μ_0 par $\widehat{\widehat{\mu_0}}$
7. Calcul des $\widehat{\widehat{\mu(\Theta_i)}}^{emp} = \widehat{\alpha}_i X_i + (1 - \widehat{\alpha}_i)\widehat{\widehat{\mu_0}}$

Exercices

Exercice 4.1 Utilisez le modèle de Bühlmann-Straub pour déterminer pour chaque groupe d'assurés la prime pour la 4^e année.
La sinistralité et les effectifs des trois premières années sont résumées ci-après.

	Montants de sinistres		
Groupe d'assurés	Année 1	Année 2	Année 3
1	8 000	11 000	15 000
2	20 000	24 000	18 000
3	10 000	15 000	13 500

	Taille des groupes			
Groupe d'assurés	Année 1	Année 2	Année 3	Année 4
1	40	50	75	75
2	100	120	120	95
3	50	60	60	60

Exercices

Exercice 4.2 Le tableau suivant reprend le nombre de sinistres causés par 300 assurés pendant la première année d'observation.

Nombre de sinistres	0	1	2	3	4	5
Nombre d'assurés	123	97	49	21	8	2

En supposant que le nombre de sinistres causés par un assuré est disribué selon une loi de Poisson dont la moyenne peut être différente selon les assurés, donnez, pour chaque assuré, une estimation de crédibilité du nombre de sinistres qu'il causera durant la prochaine année d'observation.

Exercices

Exercice 4.3 Considérer le tableau suivant où X_{ij} est le montant total de sinistres et ω_{ij} la masse salariale pour l'employeur i dans l'année j.

	X_{ij}			
Employeur i	Année 1	Année 2	Année 3	Année 4
1	14	21	12	18
2	4	0	4	6
3	3	0	1	6

	ω_{ij}			
Employeur i	Année 1	Année 2	Année 3	Année 4
1	2	3	4	3
2	4	2	1	2
3	3	3	1	3

A l'aide du modèle de Bühlmann-Straub, calculez la prime de crédibilité de l'employeur n$^\circ$ 1 pour la cinquième année sachant que $\omega_{1;5} = 3$.