

L'approche bayésienne en crédibilité: (Modèle de Jewell)

Abdelouahab LATRECHE ¹

Mots clés: Actuariat ; Théorie de crédibilité ; Tarification de Bayes.

Résumé :

La théorie de crédibilité a pour objet le calcul de la prime d'assurance pour un contrat faisant partie d'un portefeuille hétérogène.

Les concepts utilisés en crédibilité bayésienne constituent la base des modèles développés plus récemment, comme par exemple les modèles de Bühlmann (1967,1969).

Dans ce travail, nous allons d'abord mettre en évidence la tarification a posteriori de la prime selon Bayes.

D'un point de vue théorique, cette prime ne pose aucun problème. Mais malheureusement, il y a des inconvénients du point de vue pratique (calcul explicite de la prime). Jewell propose alors le modèle qu'on va étudier ici dont le but est de remédier à ces problèmes pratiques.

La prime bayésienne sera donc notée $E[\mu(\Theta)/\underline{S}]$. Son expression est obtenue par un processus à deux étapes. On doit d'abord obtenir la fonction de densité a posteriori de paramètre de risque, soit $u(\theta/x_1, K, x_n)$, pour ensuite en évaluer l'espérance. Donc cette prime peut être exprimée comme une fonction linéaire des observations $\underline{S}$. On obtient $E[\mu(\Theta)/\underline{S} = s] = z\bar{S} + (1-z)E[\mu(\Theta)]$ (z : coefficient de crédibilité et $\mu = E[\mu(\Theta)]$: prime collective d'un portefeuille des contrats d'assurance).

1. Introduction

Un problème primordial auquel les compagnies d'assurances sont confrontées, est la détermination de la prime relative à un contrat d'assurance.

La théorie de crédibilité a pour objet le calcul de la prime d'assurance pour un contrat faisant partie d'un portefeuille hétérogène.

Cette théorie a donné naissance à différents modèles, dans lesquels interviennent des paramètres, appelés paramètres structurels. Ils sont au nombre de trois. Le premier s'interprète comme étant le montant moyen des sinistres de tout le portefeuille, le deuxième mesure la dispersion des primes de risque individuel et le troisième représente une mesure globale de la dispersion totale de données du portefeuille.

Depuis le début des années quatre-vingt, l'estimation des paramètres structurels a reçu une attention considérable dans la littérature actuarielle. Ce sujet est traité par des articles qui ont été abordés par De Vylder, Dubey et Gisler.

2. Inférence bayésienne

¹ Chargé de cours à l'INPS

2.1. La règle bayésienne

Soit $(X, \Theta)^T$ un vecteur aléatoire à valeurs dans l'espace $\mathcal{X} \times \Omega$ et soit $p(x, \theta)$ sa densité. Alors

$$\pi(\theta) = \int_{\mathcal{X}} p(x, \theta) dx \text{ et } q(x) = \int_{\Omega} p(x, \theta) d\theta \quad (2.1)$$

Sont les densités marginales de Θ et X , respectivement. L'approche bayésienne suppose que pendant l'expérience on n'observe que des réalisations de X , c'est-à-dire on suppose que X est une variable observable appelée un échantillon. Par contre la deuxième composante Θ est inconnue et non observée et est considérée comme un paramètre. Supposons que la densité conditionnelle de X sachant la valeur de Θ est connue. Notons

$$\pi(x/\theta) = \frac{p(x, \theta)}{\pi(\theta)} \quad (2.2)$$

La densité conditionnelle de X sachant que $\Theta = \theta$, et soit

$$q(\theta/x) = \frac{p(x, \theta)}{q(\theta)} \quad (2.3)$$

La densité conditionnelle de Θ sachant que $X = x$. Puisque

$$p(x, \theta) = \pi(x/\theta)\pi(\theta) \quad (2.4)$$

de (2.1) –(2.4) on tire les formules de Bayes:

$$q(\theta/x) = \frac{\pi(x/\theta)\pi(\theta)}{q(x)} = \frac{\pi(x/\theta)\pi(\theta)}{\int_{\Omega} \pi(x/\theta)\pi(\theta) d\theta} \quad (2.5)$$

et

$$\pi(x/\theta) = \frac{q(\theta/x)q(x)}{\pi(\theta)} = \frac{q(\theta/x)q(x)}{\int_{\mathcal{X}} q(\theta/x)q(x) dx} \quad (2.6)$$

La densité marginale $\pi(\theta)$ de Θ est appelée la densité à priori et la densité conditionnelle $q(\theta/x)$ de $X = x$ est appelée la densité à posteriori.

L'inférence statistique sur Θ dans l'optique de l'approche bayésienne est donnée en utilisant la densité à posteriori $q(\theta/x)$ basée sur l'échantillon X , puisque toute information probabiliste sur Θ est exprimée en termes de $q(\theta/x)$. S'il est nécessaire d'estimer la valeur $U(\theta)$, où θ est une réalisation non observée du paramètre aléatoire Θ , alors on utilise l'espérance conditionnelle $E(U(\Theta)/X)$ comme l'estimateur ponctuel pour $U(\theta)$.

2.2. Estimation ponctuelle

Supposons que pendant une expérience une réalisation de X est observée et la réalisation correspondante de Θ est inconnue. Il faut estimer la valeur θ de la réalisation non observée de Θ . Soit

$$\Theta^* = \Theta^*(X)$$

Un estimateur ponctuel de θ .

L'erreur systématique de Θ^* est

$$E(\Theta^* - \Theta / X) = E(\Theta^* / X) - E(\Theta / X),$$

Où

$$E(\Theta / X = x) = \int_{\Omega} \theta q(\theta / x) d\theta \quad \text{et} \quad E(\Theta^* / X = x) = \Theta^*(x)$$

Définition 1: L'estimateur $\hat{\Theta}(X)$ est sans biais si l'erreur systématique est à zéro, c'est-à-dire, si

$$\hat{\Theta}(x) = E(\Theta / X = x)$$

Il s'ensuit que l'estimateur sans biais est unique presque sûrement.

Théorème 1: L'estimateur sans biais $\hat{\Theta}(X) = E(\Theta / X) = \int_{\Omega} \theta q(\theta / X) d\theta$ est le meilleur au sens de moindres carrés.

Définition 2: $\hat{\Theta}(X)$ est appelé l'estimateur bayésien.

3. Prime de risque et Prime collective

Une prime est un montant fixe perçu par la compagnie, dans le but de compenser les dépenses résultant de sinistres. Désignons par S la variable aléatoire représentant le montant des sinistres au cours d'une année, par $G(x)$ sa fonction de répartition. Si la distribution de la variable S , pour chaque contrat, est parfaitement connue, il serait facile de déterminer une prime proportionnelle pour chaque risque. Cependant cette hypothèse, en pratique, n'est jamais vérifiée.

Chaque risque peut être repéré au moyen d'un paramètre λ (S_{λ} : montants de sinistre) de fonction de répartition $G_{\lambda}(x)$. Le collectif (ensemble de risque) peut ainsi être identifié avec l'ensemble Λ des valeurs que peut prendre le paramètre λ .

Un principe de calcul des primes H est une règle qui permet d'associer à toute fonction de répartition $G(x)$ un nombre P

$$P = H[G(x)]$$

où H est une fonctionnelle

Si $G_{\lambda}(x)$ est la fonction de répartition correspondant à un risque particulier λ , la prime correspondante est appelée prime de risque (prime individuelle)

$$P(\lambda) = H[G_{\lambda}(x)]$$

Si $G(x)$ est la fonction de répartition du collectif, nous parlons de prime collective

$$P = H[G(x)]$$

Pour étudier les rapports existants entre la prime de risque et la prime collective, il est nécessaire de décrire la structure du collectif. Cette information nous est fournie par une fonction de structure $U(\lambda)$ (indiquant la distribution de λ au sein de Λ). Si λ est un paramètre de risque réel (unidimensionnel), alors :

$$U(\theta) = P(\lambda \leq \theta)$$

Nous avons par la théorie des probabilités totales, une relation entre la distribution S des montants de sinistres dans le collectif et aux distributions S_λ par :

$$G(x) = \int_{\Lambda} G_\lambda(x) dU(\lambda)$$

Par le principe de l'espérance mathématique, la prime réclamée à l'assuré est égale à l'espérance mathématique du risque augmentée d'un chargement de sécurité proportionnel à cette espérance.

3.1. Prime de risque :

$$P(\lambda) = (1 + \alpha)\mu(\lambda)$$

où $\mu(\lambda) = \int_0^{+\infty} x dG_\lambda(x)$ est l'espérance de S_λ

3.2. Prime collective :

$$P = (1 + \alpha)\mu$$

où $\mu = E_\lambda[\mu(\lambda)] = \int_{\Lambda} \mu(\lambda) dU(\lambda)$ est l'espérance mathématique prise par rapport à la fonction de structure $U(\lambda)$.

4. Théorie de la crédibilité

Pour un portefeuille d'assurance hétérogène, les assurés ne sont pas tous égaux devant le risque, certains présentant un profil plus dangereux que d'autres. Appliquer la même prime pour tous pourrait donc paraître inéquitable ce qui induit une surtarification de certains assurés. On peut diminuer l'hétérogénéité du portefeuille en le divisant en sous portefeuille de risques homogènes (sur base de caractéristiques, telles que le sexe, l'âge, etc.). Les caractéristiques dans chaque sous portefeuille n'expliquent pas parfaitement la dangerosité des assurés. Il est donc assez naturel d'utiliser la sinistralité relative à un individu pour réévaluer le montant de sa prime. Cette pratique relève de la théorie de la crédibilité.

L'idée fondamentale de cette théorie peut se résumer comme suit: Supposons que, pour une police particulière, on ait observé, pour les n premières périodes, les montants des sinistres suivants : $x_1, \dots, x_n$; x_i est le montant de sinistre généré par cette police durant la i ème année d'observation. La pure observée est donc

$$p_n = \frac{x_1 + x_2 + \dots + x_n}{n}$$

Si l'assureur applique cette prime, quelle prime devrait être payer par un assuré dont le montant de sinistre est nul ($x_1 = x_2 = \dots = x_n = 0$). Donc l'assureur est confronté à un problème, soit

qu'il exige une prime collective p_{coll} identique pour tous les assurés (mais pourrait mécontenter les « bons » assurés qui, s'estimant lésés), en plus l'assureur n'a pas d'intérêt d'appliquer la prime p_n . Les actuaires ont alors songé à réclamer une prime entre p_n et p_{coll} :

$$p_{créd} = zp_n + (1-z)p_{coll}$$

Où z est un facteur de crédibilité

Remarque 1 : Ce facteur z mesure la crédibilité que l'on peut accorder à la prime p_n .

4.1. La crédibilité bayésienne

Considérons un portefeuille de contrats segmenté en k contrats (police)

a/ Cas d'un portefeuille homogène : ce cas n'est jamais rencontré en pratique, l'homogénéité se traduit par le fait que toutes les contrats sont indépendantes et identiques. C'est-à-dire, si on considère que S_i est le montant annuel des sinistres de la i ème contrat ($i = 1, \dots, k$), alors les variables aléatoires S_i sont indépendantes.

b/ Cas d'un portefeuille hétérogène : ce cas est souvent rencontré en pratique, l'hétérogénéité se traduit le fait que les risques des sinistres ne sont pas les mêmes pour toutes les contrats. (en assurance automobile par exemple, les assurés ne conduisent pas de la même façon, ce qui explique bien que le risque d'accident est plus ou moins grand selon les assurés).

4.2 La prime bayésienne

Supposons que, pour une police particulière, on ait observé, pour les n premières périodes, les montants des sinistres suivants : $S_1 = x_1, \dots, S_n = x_n$

Théorème 2:

La prime pure a posteriori (appelée aussi prime de Bayes) pour la période $n+1$ est donnée par : (au sens de moindres carrés)

$$\prod(x_1, x_2, \dots, x_n) = E[\mu(\Lambda) / x_1, x_2, \dots, x_n]$$

4.3. Famille exponentielle et crédibilité exacte

La prime bayésienne $E(\mu(\Theta)/S)$ constitue la prime (au sens de moindres carrés) à charger à un assuré compte tenu de l'expérience accumulée. Le calcul de la prime dans le cas où les distributions de S/Θ et Θ sont connues, devient « plus ou moins facile ».

Exemple 1: (Bailey 1950)

Soit $S/\Theta \sim \beta(\Theta)$, $\Theta \sim \text{beta}(\alpha, \beta)$. Alors :

$$\Theta / S_1 = x_1, S_2 = x_2, \dots, S_n = x_n \sim \beta\left(\alpha + \sum x_i, \beta + n - \sum x_i\right)$$

On obtient la prime bayésienne suivante :

$$E(\mu(\Theta) / S_1 = x_1, S_2 = x_2, \dots, S_n = x_n) = \frac{\alpha + \sum x_i}{\alpha + \beta + n} = \frac{n}{\alpha + \beta + n} \bar{S} + \frac{\alpha + \beta}{\alpha + \beta + n} \frac{\alpha}{\alpha + \beta}$$

$$\text{Avec } z = \frac{n}{\alpha + \beta + n} \text{ et } E(\mu(\Theta)) = \frac{\alpha}{\alpha + \beta}$$

Dans certains cas la forme de la prime peut après une seule année être complexe, voir impossible à calculer analytiquement.

Exemple 2:

Soit $S / \Theta \sim \beta(\Theta)$, $\Theta \sim U_{(a,b)}$. Alors

$$E(\mu(\Theta) / S_1 = x_1, S_2 = x_2, \dots, S_n = x_n) = \frac{\sum_{j=1}^{n-\bar{S}} (-1)^j \frac{b^{n\bar{S}+j+2} - a^{n\bar{S}+j+2}}{(n-n\bar{S}-j)! j! (n\bar{S}+j+2)}}{\sum_{j=1}^{n-\bar{S}} (-1)^j \frac{b^{n\bar{S}+j+1} - a^{n\bar{S}+j+1}}{(n-n\bar{S}-j)! j! (n\bar{S}+j+1)}}$$

La prime bayésienne est compliqué (non linéaire), de plus elle n'est pas nécessairement entre l'expérience individuelle $\bar{S}$ et la prime collective μ

4.4. Modèle de Jewell

Définition 3: (conjuguée naturelle)

Soient $u(\theta)$ et $u(\theta / x_1, \Lambda, x_n)$ les distributions a priori et a posteriori de Θ respectivement.

On dit que $u(\theta)$ est la conjuguée naturelle de la fonction $f(x/\theta)$ si $u(\theta / x_1, \Lambda, x_n)$ est du même type (avec de nouveau paramètre) que sa distribution a priori.

Théorème 3:

Soit $f(x/\theta)$ un membre de la famille exponentielle unidimensionnelle (Lehmann (1983)), dont la densité peut s'écrire

$$f(x/\theta) = \frac{p(x)e^{-\theta x}}{q(\theta)}, \quad x > 0 \quad (4.1)$$

Alors :

1/ La conjuguée naturelle de $f(x/\theta)$ est :

$$u(\theta) = \frac{q(\theta)^{-t_0} e^{-\theta x_0}}{c(t_0, x_0)}$$

où $t_0 > 0$ et $x_0 > 0$ seront définies dans la preuve qui suit et $c(t_0, x_0)$ est un facteur de normalisation dépendant des deux paramètres t_0, x_0 .

2/ L'estimateur bayésien $E[\mu(\Theta)/\underline{S}]$ est une fonction linéaire des observations et peut s'écrire comme une prime de crédibilité

$$\bar{\mu}(x_1, x_2, \dots, x_n) = E[\mu(\Lambda) / x_1, x_2, \dots, x_n] = \frac{t_0}{n + t_0} \mu + \frac{n}{n + t_0} \bar{S}$$

Avec $z = \frac{n}{n + t_0}$

Proposition 1:

Soit $\Sigma^2 = E(V(X / \Theta))$, $M^2 = V(\mu(\Theta))$. Alors :

$$t_0 = \frac{\Sigma^2}{M^2} = \frac{E(V(X / \Theta))}{V(\mu(\Theta))}$$

Le problème est résolu d'un point de vue théorique, mais malheureusement, pour faire intervenir ce résultat en pratique, il y a des inconvénients qui surgissent:

- Il faut pour chaque valeur θ du paramètre de risque Θ déterminer la forme analytique de la distribution $f(s/\theta)$
- Il faut aussi déterminer les fonctions de distribution du portefeuille $u(\theta)$
- Même dans l'hypothèse où ces fonctions de distribution seraient connues, les calculs à effectuer pour déterminer la prime bayésienne seraient difficiles à résoudre (sauf pour certaines distributions particulières).

Ces inconvénients majeurs à l'encontre de l'utilisation des primes pures a posteriori, ont fait naître le modèle construit par Bühlmann (1967, 1969) dont l'idée de base est de se restreindre à des primes linéaires "appelées estimateur de crédibilité".

5. La prime de crédibilité

5.1 Considération théoriques

5.1.1 Principe

La distinction entre prime de risque et prime collective est théorique. En effet, dans beaucoup de cas (notamment en assurance automobile) le paramètre de risque n'est pas parfaitement connu. D'après les statistiques réunies dans un ensemble de contrats, on calcule une prime collective. Pour que la prime soit « la plus individuelle » et donc la plus juste possible, le risque assuré doit être cerné en le suivant au cours du temps.

La prime de crédibilité répond à ce besoin. C'est une approche séquentielle de la prime de risque à partir de la prime collective.

Supposons que l'on ait pu suivre un assuré pendant n périodes et que l'on se propose de définir sa prime à la période $n+1$.

5.1.2 Estimateurs de crédibilité

Définition 4:

Soit X une variable aléatoire, l'estimateur de crédibilité de la variable aléatoire X en fonction des variables aléatoires $X_1, \dots, X_n$ est défini comme étant le meilleur estimateur linéaire de la forme :

$$c_0 + \sum_{i=1}^n c_i X_i$$

où c_i sont des constantes $i = 0, 1, \dots, n$

Soient $X_1, \dots, X_n$ des variables aléatoires dont les réalisations sont connues, et soit X_{n+1} une variable aléatoire inconnue à estimer en fonction des $X_1, \dots, X_n$

Proposition 2:

Un estimateur linéaire $\hat{X}_{n+1} = c_0 + \sum_{i=1}^n c_i X_i$ est estimateur de crédibilité de X_{n+1} , si et seulement si, il satisfait le système des équations normales suivants :

$$\begin{cases} E(\hat{X}_{n+1}) = E(X_{n+1}) \\ \text{Cov}(\hat{X}_{n+1}, X_j) = \text{Cov}(X_{n+1}, X_j) \quad \forall j = 1, \dots, n \end{cases}$$

Proposition 3:

Soient X_{n+1} et $(X_1, \dots, X_n)$ conditionnellement indépendantes par rapport à Λ . Soit $\mu(\Lambda) = E(X_{n+1} / \Lambda)$.

Si $\hat{X}_{n+1}$ estimateur de crédibilité de X_{n+1} dépendant linéairement de $X_1, \dots, X_n$, alors $\hat{X}_{n+1}$ l'est aussi de $\mu(\Lambda)$

5.2 Modèle de crédibilité linéaire

Considérons un portefeuille de contrats d'assurance hétérogène. Chaque contrat du portefeuille est caractérisé par un paramètre de risque Λ (inobservable). Soit X_t étant le montant des sinistres relatifs à un contrat durant la période t , $t = 1, \dots, n$

Hypothèse :

H_1 : Les variables X_t ($t = 1, \dots, n$) sont indépendantes et identiquement distribuées conditionnellement au paramètre de risque Λ .

H_2 : $\mu(\Lambda) = E(X_t / \Lambda)$ Indépendante de t

H_3 : $V(X_t / \Lambda) = \sigma^2(\Lambda)$ Indépendante de t

Notations :

a/ $\mu(\Lambda) = E(X_t / \Lambda)$ est la prime de risque individuelle. C'est la prime, qui devrait être utilisée si le paramètre de risque était connu.

b/ $\mu = E(\mu(\Lambda)) = E(E(X_t / \Lambda)) = E(X_t)$ est une valeur moyenne¹ calculée sur le portefeuille entier

c/ $M^2 = V(\mu(\Lambda))$ est une mesure de dispersion des primes de risque individuelle. C'est un indicateur de l'hétérogénéité du portefeuille.

d/ $\Sigma^2 = E(\sigma^2(\Lambda))$ est une mesure de la fluctuation due au hasard.

Théorème 4:

Etant donné qu'on a observé les montants des sinistres pour les n premières périodes, alors la prime de crédibilité est :

$$\prod(x_1, x_2, \dots, x_n) = c_0 + \sum_{i=1}^n c_i x_i$$

$$\text{où } c_0 = \frac{\Sigma^2}{\Sigma^2 + nM^2} \mu \text{ et } c_i = \frac{M^2}{\Sigma^2 + nM^2} \quad \forall i = 1, \dots, n$$

On voit bien que $\prod(x_1, x_2, \dots, x_n)$ est de la forme :

$$(1 - \alpha)\mu + \alpha \cdot \frac{1}{n} \sum_{i=1}^n x_i$$

$$\text{Avec } 1 - \alpha = \frac{\Sigma^2}{\Sigma^2 + nM^2} \text{ et } \alpha = \frac{nM^2}{\Sigma^2 + nM^2}, (\alpha : \text{facteur de crédibilité})$$

5.3 Estimation des paramètres structurels

5.3.1 Estimation de la moyenne μ

Théorème 5:

1/ Pour le contrat i :

$\hat{\mu}_i = \frac{1}{n} \sum_{t=1}^n x_{it} = x_{i\bullet}$ est l'estimateur conditionnellement sans biais de $\mu(\Lambda_i)$ où Λ_i un paramètre de risque caractérisant le contrat i et x_{it} le montant des sinistres pour le contrat i durant la période t ; ($i = 1, \dots, k$, $t = 1, \dots, n$)

2/ Pour le portefeuille :

$\hat{\mu} = \frac{1}{k} \sum_{i=1}^k \hat{\mu}_i = \frac{1}{nk} \sum_{i=1}^k \sum_{t=1}^n x_{it} = x_{\bullet\bullet}$ est l'estimateur sans biais de μ où $x_{\bullet\bullet}$ la moyenne générale des observations

¹ La prime de risque collective est en fait une moyenne pondérée des primes de risque individuelles. Elle est indépendante du paramètre de risque Λ . Un nouveau contrat n'ayant pas de statistique de sinistre sera chargé de cette prime moyenne

5.3.2 Estimateur de Σ^2

Théorème 6:

1/ $\hat{\sigma}_i^2 = \frac{1}{n-1} \sum_{t=1}^n (x_{it} - x_{i\bullet})^2$ est un estimateur conditionnellement sans biais de $\sigma_i^2(\Lambda_i)$

2/ $\hat{\Sigma}^2 = \frac{1}{k} \sum_{i=1}^k \hat{\sigma}_i^2 = \frac{1}{(n-1)k} \sum_{i=1}^k \sum_{t=1}^n (x_{it} - x_{i\bullet})^2$ est un estimateur sans biais de Σ^2

Contrat\Année	1	2	3	4	5	6	7	8	9	10	$\hat{\mu}_i$	$\hat{\sigma}_i^2$	Π_i
1	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
2	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
3	1	0	1	0	0	0	0	0	0	0	0.20	0.178	0.1822
4	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
5	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
6	0	0	0	0	0	1	0	0	1	0	0.20	0.178	0.1822
7	0	1	1	0	0	0	0	0	0	0	0.20	0.178	0.1822
8	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
9	0	1	1	0	1	2	1	0	0	1	0.70	0.456	0.5568
10	0	0	1	0	0	0	0	0	0	0	0.10	0.100	0.1146
11	1	1	0	0	1	0	0	0	1	0	0.40	0.267	0.3174
12	0	0	0	0	0	1	0	1	0	1	0.30	0.233	0.2498
13	0	0	0	0	0	0	0	0	1	0	0.10	0.100	0.1146
14	0	0	0	0	0	0	1	0	0	0	0.10	0.100	0.1146
15	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
16	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470
17	1	1	0	1	3	0	1	0	0	1	0.80	0.844	0.6300
18	1	0	0	0	0	0	0	0	0	0	0.10	0.100	0.1146
19	0	0	0	0	0	0	0	0	0	1	0.10	0.100	0.1146
20	0	0	0	0	0	0	0	0	0	0	0.00	0.000	0.0470

5.3.3 Estimateur de M^2

Théorème7:

$\hat{M}^2 = \frac{1}{k-1} \sum_{i=1}^k (x_{i\bullet} - x_{\bullet\bullet})^2 - \frac{1}{n} \hat{\Sigma}^2$ est un estimateur sans biais de M^2

5.4 Exemple : Application de la formule de crédibilité

Tableau 1 :

Estimateur moyenne collective : $\hat{\mu} = 0.145$

Estimateur de Σ^2 : $\hat{\Sigma}^2 = 0.10388$

Estimateur de M^2 : $\hat{M}^2 = 0.02169$

Facteur de crédibilité : $\alpha = 0.6761$

Conclusion :

Les principales conclusions qui se dégagent de notre analyse sont les suivantes:

1. La méthode de linéarisation de la prime de crédibilité a l'avantage de simplifier les calculs.
2. Le modèle de Jewell a permis de démontrer que lorsque utilisées avec leur conjuguée naturelle, les fonctions de vraisemblance membres de la famille exponentielle unidimensionnelle résultent en des primes bayésiennes linéaires que l'on peut écrire comme des primes de crédibilité.
3. Le modèle de Buhlmann a permis à la théorie de la crédibilité d'effectuer un grand pas. En effet, en approximant la prime bayésienne par une fonction linéaire des observations, Buhlmann a obtenu une expression qui est un compromis entre l'expérience individuelle d'un assuré et celle de l'ensemble du portefeuille
4. La prime de crédibilité est sans biais, c'est-à-dire $E(\Pi_i) = (1 - \alpha)\mu + \alpha \cdot E(\bar{X}_i) = \mu$.
-En moyenne, l'assureur perçoit donc suffisamment de primes pour payer les sinistres
-Que Π_i soit sans biais signifie aussi qu'une mauvaise estimation du facteur de crédibilité n'a pas d'impact négatif sur les montant des primes perçu par l'assureur.
5. Le facteur de crédibilité croît avec M^2 , (i.e avec l'hétérogénéité du portefeuille)

Bibliographie :

- Buhlmann H. (1967): Experience rating and credibility. Astin Bulletin, 4, 199-207.
- Buhlmann H. (1970): Mathematical methods in risk theory. Springer.
- Buhlmann, H., GISLER, A. (2005). A Course in Credibility Theory and its Applications. Springer-Verlag, Berlin.
- Denuit, M. CHARPENTIER, A. (2005): Mathématiques de l'assurance non-vie (TOME II). Economica
- De Vylder, F. (1978) : Parameter Estimation in credibility Theory. Astin Bulletin, 10, 1, pp 99-112
- Dionne.G, C.Vanasse (1989): A généralisation of Automobile Insurance Rating Models: The Negative Binomial Distribution With a Regression component. Astin Bulletin, 19, pp199-212
- Dubey, A., Gisler., A (1981) : On Parameter Estimation In Credibility. MVSVM, Heft 2, pp. 187-222
- Gourieroux C. (1999): Statistique de l'assurance. Economica.

Jewell, W.S. (1974): Credible means are exact Bayesian for exponential families; Astin bulletin, Vol. 8, part 1, pp. 77-90

Lemaire J. (1995): Bonus-Malus systems in automobile insurance. Huebner Int. series on risk, insurance and economic security. Kluwer Acad. Publ.

ANNEXE:

Preuve du théorème 3:

De l'équation (4.1) on a :

$$q(\theta) = \int_0^{+\infty} p(x) e^{-\theta x} dx \quad \text{et} \quad -\frac{q'(\theta)}{q(\theta)} = \int_0^{+\infty} x \frac{p(x) e^{-\theta x}}{q(\theta)} dx = \mu(\theta) \quad (4.2)$$

Dans ce cas, la loi a posteriori du paramètre de risque prend la forme

$$\begin{aligned} u(\theta / x_1, \Lambda, x_n) &= \frac{u(\theta) \prod_{i=1}^n f(x_i / \theta)}{f(x_1, \Lambda, x_n)} \\ &= \frac{u(\theta) \prod_{i=1}^n f(x_i / \theta)}{\int_{\lambda \in \Theta} u(\lambda) \prod_{i=1}^n f(x_i / \lambda) d\lambda} \\ &= \frac{u(\theta) \left(\frac{1}{q(\theta)} \right)^n e^{-\theta \sum_{i=1}^n x_i}}{\int_{\lambda \in \Theta} u(\lambda) \left(\frac{1}{q(\lambda)} \right)^n e^{-\lambda \sum_{i=1}^n x_i} d\lambda} \end{aligned}$$

On remarque que $u(\cdot / x_1, \Lambda, x_n)$ ne dépend des observations qu'au travers de n et $t = \sum_{i=1}^n x_i$, c'est-à-dire

$$u(\theta / x_1, \Lambda, x_n) = \frac{u(\theta) \left(\frac{1}{q(\theta)} \right)^n e^{-\theta t}}{\int_{\lambda \in \Theta} u(\lambda) \left(\frac{1}{q(\lambda)} \right)^n e^{-\lambda t} d\lambda}$$

Si on prend

$$u(\theta) = \frac{q(\theta)^{-t_0} e^{-\theta x_0}}{c(t_0, x_0)} = u(\theta; t_0, x_0) \quad (4.3)$$

A-t-on que $u(\cdot / x_1, \Lambda, x_n)$ et $u(\cdot)$ ont même forme analytique

De (4.3) on a : $c(t_0, x_0) = \int_{\theta \in \Theta} q(\theta)^{-t_0} e^{-\theta x_0} d\theta$

$$\begin{aligned}
u(\theta / x_1, \Lambda, x_n) &= \frac{u(\theta) \left(\frac{1}{q(\theta)} \right)^n e^{-\theta t}}{\int_{\lambda \in \Theta} u(\lambda) \left(\frac{1}{q(\lambda)} \right)^n e^{-\lambda t} d\lambda} \\
&= \frac{\frac{(q(\theta))^{-t_0} e^{-\theta x_0}}{c(t_0, x_0)} \left(\frac{1}{q(\theta)} \right)^n e^{-\theta t}}{\int_{\lambda \in \Theta} \frac{(q(\lambda))^{-t_0} e^{-\lambda x_0}}{c(t_0, x_0)} \left(\frac{1}{q(\lambda)} \right)^n e^{-\lambda t} d\lambda} \\
&= \frac{\left(\frac{1}{q(\theta)} \right)^{n+t_0} e^{-\theta(t+x_0)}}{\int_{\lambda \in \Theta} \left(\frac{1}{q(\lambda)} \right)^{n+t_0} e^{-\lambda(t+x_0)} d\lambda}
\end{aligned}$$

On déduit que :

$$u(\theta / x_1, \Lambda, x_n) = \frac{\left(\frac{1}{q(\theta)} \right)^{n+t_0} e^{-\theta(t+x_0)}}{\int_{\lambda \in \Theta} \left(\frac{1}{q(\lambda)} \right)^{n+t_0} e^{-\lambda(t+x_0)} d\lambda} = \frac{\left(\frac{1}{q(\theta)} \right)^{n+t_0} e^{-\theta(t+x_0)}}{c(n+t_0, t+x_0)} = u(\theta; n+t_0, t+x_0) \quad (4.4)$$

On retrouve donc une loi de la famille des fonctions de structure conjuguées mais les paramètres $n+t_0$ et $t+x_0$.

2/

$$\begin{aligned}
\mu_{P_{coll}} &= E_{P_{ind}} \left(E(S / \Theta) \right) = E(S) = \int_0^{+\infty} x f(x) dx = \int \mu(\theta) u(\theta; t_0, x_0) d\theta \\
&= - \int \frac{q'(\theta)}{q(\theta)} \frac{e^{-\theta x_0}}{(q(\theta))^{t_0} c(t_0, x_0)} d\theta \\
&= \frac{1}{t_0} [u(\theta; t_0, x_0)]_{\Theta} + \frac{x_0}{t_0}
\end{aligned}$$

En faisant l'hypothèse que $u(\cdot; t_0, x_0)$ est nulle aux extrémités de Θ alors on a :

$$\mu = \frac{x_0}{t_0} \quad (4.5)$$

La prime bayésienne :

$$\begin{aligned}
\Pi(x_1, x_2, \dots, x_n) &= E[\mu(\Lambda) / x_1, x_2, \dots, x_n] = \int_{\theta \in \Theta} \mu(\theta) u(\theta / x_1, x_2, \dots, x_n) d\theta \\
&= \int_{\theta \in \Theta} \mu(\theta) u(\theta; n+t_0, t+x_0) d\theta \\
&= \frac{t+x_0}{n+t_0} = \frac{t_0}{n+t_0} \frac{x_0}{t_0} + \frac{n}{n+t_0} \frac{t}{n}
\end{aligned}$$

Alors

$$\prod (x_1, x_2, \dots, x_n) = E[\mu(\Lambda) / x_1, x_2, \dots, x_n] = \frac{t_0}{n+t_0} \mu + \frac{n}{n+t_0} \bar{S}$$

Avec $z = \frac{n}{n+t_0}$

Preuve de la proposition 1:

De l'équation (4.4)

$$u(\theta; t_0, x_0) = \frac{\left(\frac{1}{q(\theta)}\right)^{t_0} e^{-\theta x_0}}{c(t_0, x_0)}$$

On a :

$$\frac{d}{d\theta} u(\theta; t_0, x_0) = u(\theta; t_0, x_0) \left[-t_0 \frac{q'(\theta)}{q(\theta)} - x_0 \right] \quad (4.6)$$

De (4.2)-(4.6) on déduit que :

$$\frac{d}{d\theta} u(\theta; t_0, x_0) = u(\theta; t_0, x_0) [t_0 \mu(\theta) - x_0]$$

Et

$$\frac{d^2}{d\theta^2} u(\theta; t_0, x_0) = [t_0 \mu(\theta) - x_0] u'(\theta; t_0, x_0) + t_0 \mu'(\theta) u(\theta; t_0, x_0)$$

De (4.2) on a : $E(X^2 / \theta) = q''(\theta) / q(\theta)$ et par la suite, on trouve que :

$$\frac{d}{d\theta} \mu(\theta) = -V(X / \theta)$$

Alors

$$\frac{d^2}{d\theta^2} u(\theta; t_0, x_0) = [t_0 \mu(\theta) - x_0]^2 u(\theta; t_0, x_0) - t_0 V(X / \theta) u(\theta; t_0, x_0) \quad (4.7)$$

Si $u(\theta; n+t_0, t+x_0)$ est nulle aux extrémités de Θ et en intégrant deux fois l'équation (4.7) par rapport

à θ , on obtient : $(\mu = \frac{x_0}{t_0})$

$$\begin{aligned} 0 &= E[t_0 \mu(\theta) - x_0]^2 - t_0 E[V(X / \theta)] = t_0^2 E\left[\mu(\theta) - \frac{x_0}{t_0}\right]^2 - t_0 E[V(X / \theta)] \\ &= t_0^2 V[\mu(\theta)] - t_0 E[V(X / \theta)] \end{aligned}$$

Donc $t_0 = \frac{E(V(X / \Theta))}{V(\mu(\Theta))}$

6- Bibliographie

Cohen A. (1992): ondelettes et traitement numérique du signal, édition Masson.

Daubechies I.: orthonormal bases of compactly supported wavelet, pure app Math, 41,909-996, 1988.

Donoho D.L. & Johnston I.M.: Minimax risk over l_p balls for l_q errors. Department of statistics, Stanford university, ca 94305. July 20, 1994.

Donoho D.L. & Jonstone I.M.: Minimax estimation via wavelet shrinkage. Technical report n° 402. Department of statistics, Stanford University, 1992.

Donoho D.L. & Johnston I.M.: wavelet shrinkage asymptotia, 1995.

Donoho D.L. & Johnston I.M.: Adapting to unknown smoothness via wavelet shrinkage. Department of statistics, Stanford University, July, 1994.

Donoho D.L. & Johnston I.M.: Ideal adaptation by wavelet shrinkage. Department of statistics, Stanford University, ca 94305-4065, USA.

Donoho D.L. & Silverman B.W.: Wavelet threshold estimations for data with correlated noise. Department of statistics, Stanford University, Decembre 1994.