REPUBLIQUE DU CAMEROUN
**********
Paix-Travail-Patrie
**********
UNIVERSITE DE YAOUNDE I
***********
ECOLE NORMALE SUPERIEURE DE YAOUNDE
***********
DEPARTEMENT DE MATHEMATIQUES
***********

REPUBLIC OF CAMEROON
**********
Peace-Work-Fatherland
**********
UNIVERSITY OF YAOUNDE I
***********
HIGHER TEACHERS' TRAINING COLLEGE OF YAOUNDE
***********
DEPARTMENT OF MATHEMATICS
**********

# THÉORIE DE LA CRÉDIBILITÉ

**Mémoire de DI.P.E.S II de Mathématiques**

**De**

**YOBUNG TATSI Corielle Sylvie**

***Matricule* : 11Y698**

***Licenciée en Mathématiques***

**Sous la direction de :**

**Dr FOTSO Siméon**

***Chargé de Cours***

**École Normale Supérieure, Université de Yaoundé I**

***Année académique : 2015-2016***

# THÉORIE DE LA CRÉDIBILITÉ

Mémoire de DI.P.E.S II de mathématiques

De


**YOBUNG TATSI Corielle Sylvie**

Matricule: **11Y698**

*Licenciée en Mathématiques*

Sous la direction de:

**Dr FOTSO Siméon**

*Chargé de Cours*

École Normale Supérieure, Université de Yaoundé I



Année Académique 2015-2016

# ♦ Dédicace ♦

*À*

*mes parents*

*papa TATSI Jean et maman TCHAKOUNTEU Regine pour tout l'amour dont ils m'ont comblé et le sens du travail qu'ils m'ont appris.*

# ♦ Remerciements ♦


La rédaction d'un mémoire nécessite en plus de la motivation de l'auteur le support et la collaboration de plusieurs personnes. Je tiens à dire merci :

- ☞ Au Dieu tout puissant pour la grâce qu'il m'a accordée d'arriver jusqu'ici ;
- ☞ Au docteur FOTSO Siméon pour m'avoir accordé le privilège de diriger les travaux de ce mémoire. Il s'est toujours montré disponible à me transmettre ses connaissances en théorie de la crédibilité et a suscité en moi un enthousiasme pour la recherche ;
- ☞ À tous les enseignants du département de mathématiques de l'Ecole Normale Supérieure de Yaoundé qui m'ont tout appris depuis mon entrée dans cette prestigieuse école ;
- ☞ À tous ceux qui ont donné de leur temps pour lire ce mémoire et en faire des critiques et suggestions, je pense à KAMDEM Vidal, NKECK Leonard, NANGO Ulrich, MANTHO Tatiana ;
- ☞ À toute ma grande famille pour le soutien moral, matériel et financier qu'elle a eu à faire pour ma réussite académique ;
- ☞ À tous mes camarades de la $53^{\text{ème}}$ promotion en particulier TCHOUSSA Adèle, TACHAGO Wilfried, TCHOUPA Georges, TABO Raoul, NKOMBOU Franck qui n'ont jamais cessé de me soutenir malgré la distance qui nous a séparé ;
- ☞ À tous mes camarades de la $55^{\text{ème}}$ promotion ;
- ☞ À mes amis NGOUNGOURE Sorelle, MOUNGNY Audrey, NGONGANG Orelien pour leur présence effective à mes cotés ;
- ☞ À Tous ceux qui de près ou de loin de quelques façons que ce soit ont contribué à la rédaction de ce mémoire.

# ♦ Déclaration sur l'honneur ♦

**Le présent travail est une oeuvre originale du candidat et n'a été soumis nulle part ailleurs, en partie ou en totalité, pour une autre évaluation académique. Les contributions externes ont été dûment mentionnées et recensées en bibliographie.**

**Signature du candidat**

**YOBUNG TATSI Corielle Sylvie**

# ♦ Table des matières ♦

# ♦ RESUME ♦

En assurance, la tarification qu'établit une compagnie pour assumer les risques de son portefeuille est d'une importance capitale. Elle nécessite par conséquent d'être évaluée avec précision. L'objet de ce mémoire est de présenter quelques modèles en théorie de la crédibilité qui permettent de mieux évaluer la prime juste à payer par l'assuré. Dans le modèle de crédibilité de stabilité que nous présentons en premier, il est question de déterminer le seuil de crédibilité à partir duquel un assuré serait tarifé sur la base de sa propre expérience. Par la suite, nous présentons la crédibilité bayésienne dont l'objectif est d' évaluer la prime sur la base des informations a priori et a posteriori disponibles sur le paramètre de risque. Nous terminons par le modèle de crédibilité de Bühlmann où l'on pose la prime comme fonction linéaire des observations du portefeuille.
**Mots clés** : Crédibilité, Seuil de crédibilité, Facteurs de crédibilité, Risque, Prime de credibilité.

# ♦ Abstract ♦

In an insurance company, the price setting made in order to meet the risk of its portfolio is of higher importance. Thus, it needs to be estimated accurately. The aim of this report is to present some credibility models which ensures a better estimation of the premium. The first method is the limited-fluctuation credibility which focusis on determining the standard of credibility under which an insured will be taxed with regards to his personal experience. This is followed by the bayesian method which deals on estimating the premium according to available prior and posterior information of the risk parameter. We conclude with the Bühlmann approach where the premium is considered as the linear function of portfolio observations.
**Keys words** : Credibility, Standard of credibility, Risk, Premium of credibility, Credibility factor. .

# ♦ Abréviations ♦

v.a : variable aléatoire.

# ♦ Table des figures ♦

# ♦ Liste des tableaux ♦

# ♦ INTRODUCTION ♦

Une compagnie d'assurance est une entreprise assez particulière dont la spécificité réside dans l'inversion du cycle de production. En effet, en assurance, l'assureur perçoit d'abord le prix de son produit avant de connaitre son coût de production. Par conséquent, la compagnie doit établir une tarification qui reflète au maximum le potentiel de coûts qu'elle devra assumer pour les sinistres de son portefeuille. Après avoir fixé les prix, il faut pouvoir repartir les risques entre les différents assurés dudit portefeuille. Ainsi, l'assureur peut répartir de façon égalitaire les charges entre ces derniers. Mais, en réagissant de la sorte, il ne tient pas compte des caractéristiques particulières propres à chaque assuré. En assurance automobile par exemple, un homme âgé de $70$ ans ne conduirait pas de la même façon que celui âgé de $35$ ans. On doit donc dans ce cas prendre en compte dans le tarif, le paramètre âge du client. Un autre paramètre important dont il faut tenir compte dans le tarif d'un assuré est son passé sinistre. Il parait donc naturel qu'un assuré qui représente un grand risque paye plus cher. Cependant, les caractéristiques propres à chaque assuré sont en général inconnues de l'assureur sinon ce dernier pourrait déterminer une prime reflétant exactement le risque de cet assuré. Un tel constat explique clairement la naissance de la théorie de la crédibilité dont le but est d'attribuer à chaque contrat d'un portefeuille d'assurance une prime juste et équitable.

La théorie de la crédibilité est apparue dans le domaine des accidents du travail au début des années 1900 à la suite de l'histoire d'un gros employeur avec une meilleure expérience que celle du groupe. Mowbray en 1914 propose les premières solutions qui consistent à déterminer avec la plus grande précision les conditions sur l'expérience d'un contrat pour que celle-ci ne s'éloigne pas des valeurs observées. Ces solutions se heurtent très vite à un gros problème : les données ne sont jamais fiables à $100\%$. Pour résoudre ce problème, le concept de crédibilité partielle fut introduit en 1918 par Whitney qui, quant à lui, s'est écarté d'une vision de la crédibilité visant la précision vers une crédibilité visant la stabilité. Cette dernière est l'approche privilégiée aux États-Unis et porte aussi le nom de crédibilité américaine ou *limited fluctuation*. Cependant, Whitney a été critiqué par Bailey en 1919 pour avoir utilisé la règle de Bayes dans les conditions inappropriées. Les reproches faites par Bailey ont permis à la crédibilité de précision de prendre un nouveau départ. L'essentiel de la recherche sera faite en Europe et on lui donnera alors le nom de crédibilité européenne. Häns Bülhmann (1965) redéfinit le problème fondamental de *l'expérience rating* (tarification basée sur l'expérience) et présente la solution qui va révolutionner la théorie de la crédibilité.

Ce mémoire cherche à mettre à plat les réflexions sur la théorie de la crédibilité en exposant quelques modèles de crédibilité qui existent. Pour ce faire, notre travail s'articule autour de trois chapitres. Dans le chapitre un, nous présentons le cadre général de l'assurance, un exemple de situation ayant conduit à la théorie de la crédibilité, les origines de cette théorie et enfin la crédibilité de stabilité. Ensuite, nous présentons au chapitre deux le théorème de Bayes, l'inférence bayésienne et la crédibilité bayésienne sans oublier les limites de celle-ci. Le chapitre trois présente pour terminer le modèle original de Bühlmann qui est une réponse aux limites de la crédibilité bayésienne et le modèle classique de Bühlmann qui est une amélioration de son modèle original.

Chapitre Un

# Introduction à la théorie de la crédibilité

Les modèles de crédibilité ont été proposés pour la première fois au début du $20^{\text{ème}}$ siècle pour mettre à jour les prévisions sur les mesures de sinistres à la lumière des données disponibles dans le portefeuille de l'assureur. Dans ce chapitre, nous présentons le cadre général de l'assurance, un exemple de situation ayant entrainé la naissance de la théorie de la crédibilité, les origines de cette théorie et enfin la crédibilité de stabilité qui est la plus ancienne des approches de crédibilité.

## 1.1 Cadre général de l'assurance

Généralement connue sous le nom de **contrat d'assurance**, la **police d'assurance** est un accord signé entre deux parties à savoir : l'assureur et l'assuré. Ce contrat fixe en particulier la garantie de l'assureur, c'est-à-dire l'assistance apportée à l'assuré en cas de sinistre. Cette garantie se fait moyennant le paiement d'une cotisation de l'assuré que l'on appelle **prime**.
La prime se compose de trois parties : la partie risque, la partie frais et la partie bénéfice. Elle est en principe valable pour une période d'un an.

→ La prime de risque ou prime pure est le montant du sinistre moyen auquel devra faire face l'assureur pour couvrir le risque à assurer.

→ Les frais de gestion constituent la part qui s'ajoute à la prime pure. Ils permettent à l'assureur (en les répartissant sur tous les clients) de couvrir les charges opérationnelles (salaires, loyers,...).

→ La partie bénéfice correspond à la marge (positive ou négative) que l'assureur attribue à une population en fonction de ses objectifs commerciaux.

**Les types d'assurance**

On distingue deux types d'assurances : **l'assurance vie et l'assurance non-vie**.

→ **L'assurance vie** : elle a trait aux évènements de vie d'un individu (accidents de travail, décès, épargne et à la retraite de l'assuré). En effet, c'est un contrat de placement engageant un assureur à verser un capital à un assuré ou aux bénéficiaires d'un assuré en cas de survenance d'un risque (par exemple décès).

→ **L'assurance non-vie** : elle n'a pas directement partie liée avec la vie de l'assuré. Elle concerne les garanties des biens matériels, des dettes consenties par l'assuré ainsi que celle des personnes. En termes techniques, on parle d'assurances Incendie, Accident et Risques Divers (IARD).

Cette dernière est celle autour de laquelle nous nous appesantirons particulièrement.

**Les catégories d'assurance**

On distingue trois catégories d'assurance : **les assurances de dommages aux biens, les assurances de responsabilité et les assurances de personnes**.

**Les assurances de dommages aux biens** couvrent une grande variété de risques tels que l'incendie, le vol, le dégât des eaux, le bris de glaces, ou encore les dégradations consécutives dues aux intempéries.

**Les assurances de responsabilité** couvrent l'assuré, les personnes officiellement à sa charge, ainsi que les conséquences financières des dommages dont il est l'auteur.

**Les assurances de personnes** ont pour rôle de préserver l'être humain (maladie, accident, retraite, décès).

## 1.2 Exemple introductif

Un portefeuille d'assurance est composé de 10 contrats. Les contrats sont a priori considérés équivalents. On suppose que :

→ tout assuré ne peut avoir qu'au plus un sinistre par année ;

→ le montant du sinistre est de 200000 ;

→ la prime collective est de 40000, c'est-à-dire que l'assureur s'attend qu'en moyenne deux assurés sur 10 aient un accident au cours d'une année.

### Situation après une année

Après une année, voici l'expérience au sein du portefeuille.

| | Contrat | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Année | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | | | | | 1 | | | | | |

→ Le montant moyen de sinistre par contrat est $\dfrac{200000}{10} = 20000$

→ La prime collective est très élevée.

→ On n'a pas suffisamment de données pour tirer une conclusion.

## Situation après cinq années

Après cinq années, l'expérience du portefeuille est la suivante :

| | **Contrat** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Année | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | | | | | 1 | | | | | |
| 2 | 1 | | 1 | | 1 | | | | | |
| 3 | | | 1 | | 1 | | | | | |
| 4 | | | | | | 1 | | | | |
| 5 | 1 | | | | 1 | 1 | | | | |

→ Le montant moyen de sinistre par contrat est $\frac{2000000}{50} = 40000$

→ La prime collective est adéquate.

→ Le contrat 5 a déjà eu 4 sinistres tandis que les contrats $1$, $3$ et $6$ ont deux sinistres chacun et les autres n'en n'ont aucun.

## Situation après dix années

Après dix années, l'expérience du portefeuille est la suivante :

| | **Contrat** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Année | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | | | | | 1 | | | | | |
| 2 | 1 | | 1 | | 1 | | | | | |
| 3 | | | 1 | | 1 | | | | | |
| 4 | | | | | | 1 | | | | |
| 5 | 1 | | | | 1 | 1 | | | | |
| 6 | | 1 | | | 1 | 1 | | | | |
| 7 | | | | | 1 | 1 | | 1 | | |
| 8 | 1 | 1 | | | | | | | | |
| 9 | | | | | | | | | | |
| 10 | 1 | 1 | | | 1 | | | | | |
| Total | 4 | 3 | 2 | 0 | 7 | 4 | 0 | 1 | 0 | 0 |

→ Le montant moyen de sinistre par contrat est $\frac{4200000}{100} = 42000$.

→ La prime collective est raisonnablement adéquate.

→ Le contrat 5 a déjà eu 7 sinistres tandis que les contrats 4,7,9 et 10 n'ont jamais eu de sinistres.

## Conclusion

Sur les 10 années d'observation, on dénombre au total 21 sinistres.
Les contrats 1 et 5 représentent à eux seuls la moitié de la sinistralité globale tandis que les contrats 4,7,9 et 10 n'ont jamais eu de sinistres. On vient alors à se demander si les contrats sont réellement équivalents.
La tarification actuelle (même prime collective) est raisonnablement adéquate mais n'est pas équitable. De prime à bord on a l'impression que certains assurés payent trop pendant que les assurés 1 et 5 sont sous tarifés. Si cette situation n'est pas revue, cela amènerait les bons risques à partir. Le besoin d'une technique de tarification basée sur l'expérience est donc nécessaire : **c'est la théorie de la crédibilité**.

# 1.3 Les origines de la théorie de la crédibilité

Vers 1910 aux États-Unis, General Motors (GM) et le petit constructeur indépendant Tucker sont assurés chez Allstate contre les accidents de travail, avec quelques autres fabricants d'automobiles. Un taux moyen est calculé à partir de l'expérience de l'ensemble des fabricants et c'est ce taux qui est attribué à chacun. En calculant son propre taux, GM constate que ce taux serait année après année inférieure au taux appliqué à l'ensemble des entreprises assurées. En argumentant que son nombre d'employés est suffisamment important, il réclame un tarif fondé exclusivement sur sa propre sinistralité et non celui de l'ensemble des entreprises assurées.
Au même moment, l'entreprise Tucker plus petite que GM fait la même demande. De ces réclamations naissent un certain nombre de questionnement. On veut savoir la taille à partir de laquelle une entreprise peut être tarifée exclusivement sur sa propre expérience. Autrement dit, la limite entre un nombre d'employés fiable et un nombre d'employés non fiable.
Arthur H. Mowbray (1914) propose une première solution dans le premier numéro des *proceedings* de la *Casuality Actuarial Society*. Supposant la probabilité $q$ qu'un accident survienne connue, il désire calculer le nombre minimal d'employés assurés $n$ de telle sorte que la probabilité que le nombre d'accidents ne s'éloigne pas de plus $100k\%$ de la moyenne soit supérieure à $100p\%$. Si l'on note $N$ la distribution du nombre d'accidents, l'énoncé précédent s'écrit mathématiquement

$$Pr[(1-k)E[N] \leq N \leq (1+k)E[N]] \geq p$$

où $N$ suit la loi binomiale de paramètre $n$ et $q$ ($N \hookrightarrow \mathcal{B}(n,q)$).
Par cette équation, Mowbray fixait les bases de la crédibilité de stabilité encore connue sous le nom de *Limited fluctuations*. La solution de Mowbray a ainsi connu de multiples adaptations qui lui ont

permis de rester d'actualité jusqu'à nos jours. Cependant, dans un contexte où la notion de pleine crédibilité ne faisait plus l'unanimité au milieu des actuaires, certains pensant que les données ne sont jamais fiables à $100\%$, les assurés les ont très vite soumis à un nouveau problème, celui de savoir la solution à prendre lorsque les valeurs sont comprises entre 0 et 1.

Pour répondre à ces préoccupations, le concept de **crédibilité partielle** fut introduit par Whitney en 1918. Par là, il mentionne la nécessité d'établir la prime comme la pondération de l'expérience collective et de l'expérience individuelle de l'assuré. La pondération à donner à l'expérience individuelle est influencée par le niveau de risque, la crédibilité de la prime collective et l'homogénéité du groupe. La notion d'homogénéité du collectif est au cœur de la crédibilité de précision. Whitney modélise l'hétérogénéité du groupe en supposant que les moyennes des divers risques sont distribuées suivant une loi normale. On obtient alors une expression pour la prime individuelle $P$ de la forme

$$P = ZX + (1 - Z)C$$

où $X$ est l'expérience individuelle, $C$ l'expérience collective et $Z$ le facteur de crédibilité.
Dans ses travaux, Whitney s'est écarté d'une conception de la crédibilité visant la précision pour encourager plutôt celle visant la stabilité qui est celle privilégiée jusqu'aujourd'hui aux États-Unis. Il fut critiqué par Bailey en 1919 pour avoir utilisé la règle de Bayes dans des conditions qui n'intégraient pas la règle généralisée faite par Laplace en 1820. À partir de ces reproches, Bailey intègre de façon explicite à la théorie de la crédibilité le principe de Bayes dans le processus de tarification et la linéarité de l'estimateur bayésien sous certaines conditions.

Pendant que l'approche bayésienne faisait des progrès dans les années 1930 grâce à Bruno Finelli et Ove Lundberg dans les années 40, une nouvelle branche de la théorie bayésienne voyait le jour : c'est l'approche bayésienne empirique. Celle-ci est d'une importance capitale dans le développement de la crédibilité de précision puisqu'elle lui permet de sauter le mur entre la théorie et la pratique. Häns Bülhmann (1965) redéfinit alors le problème fondamental de *l'expérience rating* (tarification basée sur l'expérience) et présente la solution qui va révolutionner la théorie de la crédibilité. Le virage est alors définitivement pris en faveur de l'approche de précision et l'essentiel de la recherche sera fait en Europe et va ainsi prendre le nom de crédibilité européenne dont le célèbre Bülhmann marque le début de l'histoire contemporaine.

## 1.4 Crédibilité de stabilité

**Définition 1.1. La théorie de la crédibilité** peut être définie comme l'ensemble des techniques utilisées par les actuaires pour déterminer la prime d'un assuré dans un portefeuille hétérogène.
Cette théorie est principalement utilisée en assurance automobile et en accidents du travail et peut s'appliquer à différentes mesures de sinistres.

### Notations

Nous résumons ci-dessous les facteurs clés et définissons les notations qui seront utilisées par la suite :
**Nombre de sinistres** : le nombre de sinistres dans une période est désignée par $N$.
**Montant d'un sinistre** : le montant moyen d'un sinistre est représenté par $\overline{X}$.
**Charge totale sinistres** : on désigne par $X_i$ le montant du $i^{\text{ème}}$ sinistre et par $S$ la charge totale sinistres au cours d'une période. Ainsi $S = X_1 + X_2 + \cdots + X_N$ et $\overline{X} = S/N$ est le montant moyen d'un sinistre.
**Prime pure** : elle est notée $P$ et définie par $P = S/E$ où $E$ est le nombre de personnes exposées au risque dans un groupe.

Les mesures de sinistres $N$, $X_i$, $S$, $\overline{X}$ et $P$ sont des variables aléatoires déterminées par des évènements incertains, pendant que $E$ est une constante mesurant la taille du portefeuille.
De façon générale, la prime est une moyenne pondérée de l'expérience du contrat et d'une quantité spécifiée par le manuel d'assurance. Plus formellement, si $D$ est la perte qui vient de l'expérience d'un contrat et $M$ la quantité spécifiée par le manuel d'assurance, la prime s'écrit

$$P = ZD + (1 - Z)M, \quad 0 \leq Z \leq 1$$

$Z$ est appelé facteur de crédibilité.
-Si $Z = 1$, on parle de crédibilité complète (crédibilité totale).
-Si $Z \in [0; 1[$, on parle alors de crédibilité partielle.

### 1.4.1 Crédibilité complète

Un contrat d'assurance est considéré crédible si son expérience est stable. La stabilité de l'expérience va de pair avec la taille d'un contrat qu'elle soit exprimée en terme de volume de primes, masse salariale, nombre d'employés, nombre de sinistres, nombre d'années d'expérience, etc...
Dans cette approche, il est question de déterminer le minimum de données nécessaires pour attri-

buer à une expérience une crédibilité complète. Cette taille minimum de données est appelée **seuil de crédibilité complète**.

**Définition 1.2.** Une **crédibilité complète d'ordre** $(k, p)$ est attribuée à l'expérience $S$ d'un contrat si les paramètres de la distribution de $S$ sont tels que la relation

$$Pr\left[(1-k)E[S] \leq S \leq (1+k)E[S]\right] \geq p$$

est vérifiée.

### Crédibilité complète pour le nombre de sinistres

Supposons que la variable aléatoire $N$ désignant le nombre de sinistres observés a une espérance $\mu_N$ et une variance $\sigma_N^2$. Pour évaluer comment les valeurs observées de $N$ sont représentatives de la vraie espérance, on se pose la question suivante : quelle est la probabilité que le nombre de sinistres observé se concentre dans un intervalle de $100k\%$ autour de l'espérance ?
Cette probabilité est donnée par :

$$Pr\left[\mu_N - k\mu_N \leq N \leq \mu_N + k\mu_N\right] = Pr\left[\frac{-k\mu_N}{\sigma_N} \leq \frac{N-\mu_N}{\sigma_N} \leq \frac{k\mu_N}{\sigma_N}\right]. \tag{1.1}$$

Si nous supposons de plus que la variable aléatoire $N$ est normalement distribuée alors $\dfrac{N-\mu_N}{\sigma_N}$ suit la loi normale centrée réduite. L'expression ci-dessus devient :

$$\begin{aligned} Pr\left[\frac{-k\mu_N}{\sigma_N} \leq \frac{N-\mu_N}{\sigma_N} \leq \frac{k\mu_N}{\sigma_N}\right] &\simeq \phi\left(\frac{k\mu_N}{\sigma_N}\right) - \phi\left(\frac{-k\mu_N}{\sigma_N}\right) \\ &= \phi\left(\frac{k\mu_N}{\sigma_N}\right) - \left[1 - \phi\left(\frac{k\mu_N}{\sigma_N}\right)\right] \\ &= 2\phi\left(\frac{k\mu_N}{\sigma_N}\right) - 1. \end{aligned} \tag{1.2}$$

Si

$$\frac{k\mu_N}{\sigma_N} = t_{1-\frac{\alpha}{2}} \tag{1.3}$$

où $t_\beta$ est le $100\beta^e$ centile de la loi normale, c'est-à-dire $\phi\left(t_\beta\right) = \beta$ alors la probabilité en $(1.2)$ est donnée par $2(1-\alpha/2)-1 = 1-\alpha$. Ainsi, il y a une probabilité $p = 1-\alpha$ que le nombre de sinistres observé se concentre dans un intervalle de $100k\%$ autour de la moyenne.

L'application de l'équation $(1.3)$ requiert la connaissance de l'espérance $\mu_N$ et de l'écart type $\sigma_N$. Pour simplifier les calculs, on suppose souvent que le nombre de sinistres observé est distribué suivant une loi de Poisson de paramètre $\lambda_N$ très grand ainsi on pourra l'approximer à la loi normale. Suite aux hypothèses de la loi de Poisson, nous avons $\mu_N = \sigma_N^2 = \lambda_N$ et l'équation $(1.3)$ peut s'écrire

$$\frac{k\lambda_N}{\sqrt{\lambda_N}} = k\sqrt{\lambda_N} = t_{1-\frac{\alpha}{2}}. \tag{1.4}$$

La crédibilité complète est atteinte pour le nombre de sinistres si la relation
$Pr\left[\mu_N - k\mu_N \leq N \leq \mu_N + k\mu_N\right] \geq p$ est vérifiée où $p = 1 - \alpha$, $k$ et $p$ sont des valeurs données.
De l'équation $(1.4)$ on peut voir que sous l'hypothèse de la loi de Poisson la relation ci-dessus devient :

$$k\sqrt{\lambda_N} \geq t_{1-\frac{\alpha}{2}}. \tag{1.5}$$

La crédibilité complète est donc attribuée aux données si $\lambda_N \geq \left(t_{1-\frac{\alpha}{2}}/k\right)^2$.
Nous définissons maintenant

$$\lambda_F = \left(\frac{t_{1-\frac{\alpha}{2}}}{k}\right)^2, \tag{1.6}$$

le seuil de crédibilité complète pour le nombre de sinistres. La crédibilité complète est donc atteinte si $\lambda_N \geq \lambda_F$. En pratique, le nombre de sinistres pour une année en cours étant inconnu, l'exécution de ce modèle de crédibilité est de comparer les valeurs observés de $N$ dans une période récente (l'année dernière par exemple) à $\lambda_F$ calculé en utilisant l'équation $(1.6)$.

Nous donnons dans la table ci-dessous quelques valeurs du seuil de crédibilité complète $\lambda_F$ pour le nombre de sinistres.

| $\alpha$ | probabilité $p$ | $k$ | | |
|---|---|---|---|---|
| | | $10\%$ | $5\%$ | $1\%$ |
| $0.20$ | $80\%$ | $165$ | $666$ | $16641$ |
| $0.10$ | $90\%$ | $271$ | $1083$ | $27061$ |
| $0.05$ | $95\%$ | $385$ | $1537$ | $38416$ |
| $0.01$ | $99\%$ | $664$ | $2653$ | $66349$ |

TABLE 1.1 – Quelques valeurs du seuil de crédibilité complète pour le nombre de sinistres.

Cette table nous montre que pour une valeur de $k$ donnée, le seuil de crédibilité $\lambda_F$ augmente lorsque la probabilité $p$ augmente (ou que la valeur de $\alpha$ diminue).

**Exemple d'application 1.4.1.** Une compagnie d'assurance requiert une probabilité de couverture de $97\%$ pour être certaine que le nombre de sinistres espérés se concentre dans un intervalle de $k = 5\%$ autour de sa moyenne.
Déterminons le nombre de sinistres requis dans une période récente pour une crédibilité complète.

En utilisant l'équation $(1.6)$ nous obtenons

$$\lambda_F = \left(\frac{t_{0,985}}{0,05}\right)^2 = \left(\frac{2,17}{0,05}\right)^2 = 1883,56.$$

$1884$ sinistres sont nécessaires pour qu'une crédibilité complète soit accordée aux données.
La compagnie d'assurance a reçu $2590$ sinistres dans son portefeuille. On suppose que la distribution du nombre de sinistres suit une loi de Poisson qui peut être approximée à la loi normale. Déterminons le nombre espéré de sinistre mis à jour pour la période suivante.

Comme le nombre de sinistres observés $2590$ est plus grand que $1884$, la crédibilité complète est attribuée aux données. Le nombre de sinistres espérés pour la période suivante est donc estimé à $2590$.

**Remarque 1.4.1.** Le seuil de crédibilité que nous avons déterminé repose fortement sur l'hypothèse selon laquelle le nombre de sinistres observés suit une loi de Poisson. Toutefois cette hypothèse peut être relâchée et on peut supposer par exemple que $N$ suit la loi binomiale de paramètre $n$ et $\theta$ où $n$ est le nombre de personnes exposées au risque et $\theta$ la probabilité d'un sinistre.

### Crédibilité complète pour le montant d'un sinistre

Dans le cas précédent, nous avons déterminé le seuil de crédibilité complète en fonction du nombre de sinistres. On peut aussi fixer le seuil de crédibilité complète en fonction du montant d'un sinistre. Supposons que l'on dispose d'un échantillon de $N$ sinistres de montants $X_1,\ X_2, \ldots, X_N$. Dans la suite, nous supposons que $X_1,\ X_2, \ldots, X_N$ sont indépendantes et identiquement distribuées d'espérance $\mu_X$ et de variance $\sigma_X^2$ et nous utilisons la moyenne $\overline{X}$ de l'échantillon pour estimer $\mu_X$. Puisque $\overline{X} = \dfrac{X_1 +\ X_2 + \cdots + X_N}{N}$, on a alors :

$$\mu_{\overline{X}} = \frac{N\mu_X}{N} = \mu_X$$

$$\sigma^2_{\overline{X}} = \frac{N\sigma^2_{X_i}}{N^2} = \frac{\sigma^2_X}{N}.$$

La crédibilité complète est attribuée à $\overline{X}$ si on a $Pr\left[(1-k)\mu_X \leq \overline{X} \leq (1+k)\mu_X\right] \geq p$ pour des valeurs de $p$ et $k$ données. Nous supposons également que la taille de l'échantillon $N$ est suffisamment grande pour approximer $\overline{X}$ à une loi normale de moyenne $\mu_X$ et de variance $\dfrac{\sigma_X^2}{N}$. De là, on a :

$$\begin{aligned} Pr\left[\mu_X - k\mu_X \leq \overline{X} \leq \mu_X + k\mu_X\right] &= Pr\left[\frac{-k\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{\overline{X} - \mu_X}{\frac{\sigma_X}{\sqrt{N}}} \leq \frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right] \\ &\simeq 2\phi\left(\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}}\right) - 1. \end{aligned} \quad (1.7)$$

Pour que $Pr\left[(1-k)\mu_X \leq X \leq (1+k)\mu_X\right] \geq p$, on doit avoir

$$\frac{k\mu_X}{\frac{\sigma_X}{\sqrt{N}}} \geq t_{1-\frac{\alpha}{2}}. \tag{1.8}$$

Ainsi,

$$N \geq \left(\frac{t_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\sigma_X}{\mu_X}\right)^2. \tag{1.9}$$

On rappelle que le coefficient de variation d'une variable aléatoire $X$ est $C_X = \sigma_X/\mu_X$. En utilisant l'équation $(1.6)$, l'expression $(1.9)$ peut s'écrire

$$N \geq \lambda_F C_X^2. \tag{1.10}$$

Ainsi, $\lambda_F C_X^2$ est le seuil de crédibilité complète pour le montant d'un sinistre. Si le nombre de sinistres est supérieur à $\lambda_F C_X^2$, $\overline{X}$ deviendra la valeur de prédiction pour le montant moyen des sinistres pour la période suivante. Dans le calcul ci-dessus, $N$ est traité comme une constante et non pas comme une variable aléatoire dépendant de l'expérience du contrat. Pour mettre à exécution la méthodologie, $\mu_X$ et $\sigma_X^2$ doivent être estimées à partir de l'échantillon.

**Exemple d'application 1.4.2.** Déterminons le seuil de crédibilité complète pour le montant d'un sinistre avec $\alpha = 0.10$ et $k = 0.05$, sachant que l'espérance et la variance du montant d'un sinistre sont estimées à 1000 et 3000000 respectivement.
De la table $1.1$ nous avons $\lambda_F = 1083$. Ainsi, utilisant l'expression $(1.10)$, le seuil de crédibilité complète pour le montant d'un sinistre est

$$1083\left[\frac{3000000}{(1000)^2}\right] = 3249$$

Un assuré doit donc avoir au minimum 3249 sinistres pour se voir accorder une crédibilité complète sur le montant d'un sinistre.

### Crédibilité complète pour la charge totale sinistres

Pour obtenir le seuil de crédibilité complète pour la charge totale, nous déterminons le minimum du nombre de sinistres probable pour que la relation $Pr\left[\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S\right] \geq p$ soit vérifiée pour des valeurs de $p$ et $k$ données. Désignons respectivement par $\mu_S$ et $\sigma_S^2$, l'espérance et la variance de $S$. Nous avons besoin d'évaluer

$$Pr\left[\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S\right] = Pr\left[\frac{-k\mu_S}{\sigma_S} \leq \frac{S-\mu_S}{\sigma_S} \leq \frac{k\mu_S}{\sigma_S}\right]. \tag{1.11}$$

Pour calculer $\mu_S$ et $\sigma_S^2$, nous utilisons les formules de la distribution composée. Spécialement, si $N$ et $X_1, X_2, \ldots, X_n$ sont mutuellement indépendantes, nous avons $\mu_S = \mu_N\mu_X$ et $\sigma_X^2 = \mu_N\sigma_X^2 + \mu_X^2\sigma_N^2$.

Si nous supposons de plus que $N$ est distribuée suivant une loi de Poisson de moyenne $\lambda_N$, alors $\mu_N = \sigma_N^2 = \lambda_N$ et $\sigma_X^2 = \lambda_N\left(\mu_X^2 + \sigma_X^2\right)$. Ainsi

$$\frac{\mu_S}{\sigma_S} = \frac{\lambda_N \mu_X}{\sqrt{\lambda_N(\mu_X^2 + \sigma_X^2)}} = \frac{\mu_X \sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}}. \tag{1.12}$$

En approximant la charge totale sinistres par une loi normale, l'équation $(1.11)$ peut encore s'écrire

$$\begin{aligned} Pr\left[\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S\right] &\approx 2\phi\left(\frac{k\mu_S}{\sigma_S}\right) - 1 \\ &= 2\phi\left(\frac{k\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}}\right) - 1. \end{aligned} \tag{1.13}$$

Pour que $Pr\left[\mu_S - k\mu_S \leq S \leq \mu_S + k\mu_S\right] \geq p$, il faut que

$$\frac{k\mu_X\sqrt{\lambda_N}}{\sqrt{\mu_X^2 + \sigma_X^2}} \geq t_{1-\frac{\alpha}{2}}. \tag{1.14}$$

Ainsi,

$$\lambda_N \geq \left(\frac{t_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right).$$

Le seuil de crédibilité complète pour la charge totale est donc

$$\left(\frac{t_{1-\frac{\alpha}{2}}}{k}\right)^2 \left(\frac{\mu_X^2 + \sigma_X^2}{\mu_X^2}\right) = \lambda_F\left(1 + C_X^2\right). \tag{1.15}$$

Nous constatons de l'expression $(1.15)$ que le seuil de crédibilité complète pour la charge totale sinistres est toujours plus grand que le seuil de crédibilité complète pour le nombre de sinistres enregistré. En effet,

$$\lambda_F\left(1 + C_X^2\right) = \lambda_F + \lambda_F C_X^2 \geq \lambda_F.$$

On remarque que :

| Seuil de crédibilité complète pour la charge totale sinistres | = | Seuil de crédibilité complète pour le nombre de sinistres | + | Seuil de crédibilité complète pour le montant d'un sinistre |
|---|---|---|---|---|

**Exemple d'application 1.4.3.** (COSSETTE et GOULET 2010)

Un portefeuille de police d'assurance pour accidents de travail a pour espérance du montant des sinistres 5000 et pour variance 6250000. On suppose que le nombre de sinistres suit une loi de poisson qui peut être approximée à une loi normale. Pour $k = 10\%$ et $p = 98\%$, déterminons le seuil de la crédibilité complète pour cette compagnie d'assurance.

De l' équation $(1.6)$, le seuil de crédibilité complète pour le nombre de sinistres est

$$\lambda_F = \left(\frac{t_{1-\frac{0{,}02}{2}}}{0{,}1}\right)^2 = 542{,}49.$$

Le coefficient de variation du montant des sinistres est

$$\frac{\sqrt{6250000}}{5000} = 0,5.$$

De l'équation $(1.15)$, le seuil de la crédibilité complète est

$$\lambda_F\left(1 + C_X^2\right) = 542,49(1 + (0,5)^2) = 678,1125.$$

Cette compagnie attribue une crédibilité complète à une entreprise si celle ci a déjà enregistré au moins 679 accidents.

**Crédibilité complète pour la prime pure**

La prime pure notée $P$ est la prime chargée pour couvrir les pertes avant de prendre en compte les charges et les bénéfices de l'assureur. En gardant en mémoire tout ce qui précède, nous voulons évaluer

$$Pr\left[\mu_P - k\mu_P \leq P \leq \mu_P + k\mu_P\right] \simeq 2\phi\left(\frac{k\mu_P}{\sigma_P}\right) - 1$$

où $\mu_P$ et $\sigma_P^2$ sont respectivement l'espérance et la variance de $P$. Comme $P = S/E$ où E est une constante désignant le nombre de personnes exposées au risque, nous avons
$\mu_P = \left(\frac{1}{E}\right)\mu_S$ et $\sigma_P = \left(\frac{1}{E}\right)\sigma_S$ donc $\mu_P/\sigma_P = \mu_S/\sigma_S$.

Ainsi, l'expression finale de $\mu_S/\sigma_S$ dans l'équation $(1.12)$ peut être utilisée pour calculer la probabilité précédente. Cela nous permet de conclure que le seuil de crédibilité complète pour la prime pure est le même que celui de la charge totale sinistres, soit $\lambda_F\left(1 + C_X^2\right)$.

**Exemple d'application 1.4.4.** YIU-KUEN (2009)
Dans un portefeuille de police d'assurance, chaque police a une probabilité $0,05$ d'avoir un sinistre. On suppose que le nombre de sinistres est distribué selon une loi de Poisson. Si la distribution du montant d'un sinistre suit la loi log-normale avec $\mu = 5$ et $\sigma = 1$, calculons le nombre polices requises pour atteindre une crédibilité complète pour la prime pure avec $p = 0,98$ et $k = 0,05$.

L'espérance et la variance du montant d'un sinistre sont :

$$\mu_X = exp(\mu + \frac{\sigma^2}{2}) = exp(5,5) = 244,619$$

et

$$\sigma_X^2 = \left[exp\left(2\mu + \sigma^2\right)\right]\left[exp(\sigma^2) - 1\right] = 102880,6497.$$

Ainsi le coefficient de variation du montant d'un sinistre est

$$C_X = \frac{\sqrt{102880,6497}}{244,69} = 1,3108.$$

Or $t_{0,99} = 2,3263$ ; ainsi le seuil de crédibilité complète exigé pour la prime pure est

$$\lambda_F\left(1+C_X^2\right) = \left(\frac{2,3263}{0,05}\right)^2\left[1+(1,3108)^2\right] = 5884,2379.$$

D'où le minimum du nombre de polices pour une crédibilité complète est

$$\frac{5884,2379}{0,03} = 196142.$$

### 1.4.2 Crédibilité partielle

Lorsqu'un assuré ne dispose pas d'un volume de données au moins égal au seuil de crédibilité, la crédibilité de stabilité permet de déterminer pour ce dernier une prime qui sera un barycentre de la prime collective et de la prime d'expérience. On est ainsi amené à déterminer le facteur de crédibilité $Z$ telle que

$$Pr[Z\mu_W - k\mu_W \leq ZW \leq Z\mu_W + k\mu_W] = p \tag{1.16}$$

où $W$ est une mesure de perte et $k$ une valeur donnée. YIU-KUEN (2009).
Ainsi, dans le cas où la mesure de perte est le nombre de sinistres $N$, $(1.16)$ devient

$$Pr\left[Z\mu_N - k\mu_N \leq ZN \leq Z\mu_N + k\mu_N\right] = p.$$

Ce qui est équivalent à

$$Pr\left[\frac{-k\mu_N}{Z\sigma_N} \leq \frac{N-\mu_N}{\sigma_N} \leq \frac{k\mu_N}{Z\sigma_N}\right] = p. \tag{1.17}$$

Supposons que le nombre de sinistres est distribué suivant une loi de poisson de moyenne $\lambda_N$. En approximant cette loi à une loi normale, la relation $(1.17)$ se réduit à :

$$2\phi\left(\frac{k\sqrt{\lambda_N}}{Z}\right) - 1 = p.$$

Ainsi,

$$\left(\frac{k\sqrt{\lambda_N}}{Z}\right) = t_{1-\frac{\alpha}{2}},$$

ceci entraine

$$Z = \left(\frac{k}{t_{1-\frac{\alpha}{2}}}\right)\sqrt{\lambda_N} = \sqrt{\frac{\lambda_N}{\lambda_F}}. \tag{1.18}$$

L'équation $(1.18)$ établit que le facteur de crédibilité partielle $Z$ est la racine carrée du rapport du nombre de sinistres espérés et du seuil de crédibilité complète pour le nombre de sinistres.
Le principe de détermination du facteur de crédibilité pour le nombre de sinistres peut être appliqué aux autres mesures de perte et nous obtenons des résultats ci-après :

**facteur de crédibilité pour le montant d'un sinistre** : $Z = \sqrt{\dfrac{N}{\lambda_F C_X^2}}$.

**facteur de crédibilité pour les charges globales ou la prime pure** : $Z = \sqrt{\dfrac{\lambda_N}{\lambda_F\left(1+C_X^2\right)}}$.

<u>**Exemple d'application**</u> **1.4.5.** Dans un portefeuille de polices d'assurance, on a enregistré $1230$ sinistres l'année dernière avec une espérance de $45$ et une variance de $5035$. La crédibilité complète est basée sur une probabilité de couverture de $95\%$ pour $k = 5\%$. Chaque police a une probabilité $0,08$ d'avoir un sinistre et le portefeuille a $19400$ polices. Calculons le facteur de crédibilité $Z$ pour le nombre de sinistres, le montant d'un sinistre, et la charge totale pour cette année.

Le nombre de sinistres probable est $\lambda_N = 19400(0,08) = 1552$.
D'après la table $(1.1)$, pour $p = 95\%$ et $k = 5\%$ le seuil de crédibilité complète pour le nombre de sinistres est $\lambda_F = 1537 < 1552$. La crédibilité complète est donc atteinte pour le nombre de sinistres et par conséquent $Z = 1$.
La valeur estimée du coefficient de variation pour le montant d'un sinistre est

$$C_X = \frac{\sqrt{5035}}{45} = 1,5768,$$

donc le seuil de crédibilité complète pour le montant d'un sinistre est

$$\lambda_F C_X^2 = (1537)(1,5768)^2 = 3821,44$$

qui est plus grand que la taille de l'échantillon qui est de $1230$. À cet effet, la crédibilité complète n'est pas atteinte pour le montant d'un sinistre. Le facteur de crédibilité partiel est

$$Z = \sqrt{\frac{1230}{3821,44}} = 0,5673.$$

Pour la charge totale sinistres, le seuil de crédibilité complète est

$$\lambda_F(1 + C_X^2) = 5358,44 > 1552 = \lambda_N.$$

La crédibilité complète n'est pas atteinte pour la charge totale sinistres et le facteur de crédibilité est

$$Z = \sqrt{\frac{1552}{5358,44}} = 0,5382.$$

### 1.4.3 Résumé et limites de crédibilité de stabilité

#### Résumé sur l'approche de crédibilité de stabilité

La table $(1.2)$ suivante donne le seuil de crédibilité complète et le facteur de crédibilité $Z$ pour les différentes mesures de sinistres.

| Mesures de sinistres | seuil de crédibilité complète | facteur de crédibilité $Z$ partielle |
|---|---|---|
| Nombre de sinistres | $\left(\frac{t_{1-\frac{\alpha}{2}}}{k}\right)^2$ | $\sqrt{\frac{\lambda_N}{\lambda_F}}$ |
| Montant d'un sinistre | $\lambda_F C_X^2$ | $\sqrt{\frac{N}{\lambda_F C_X^2}}$ |
| Charge totale/Prime pure | $\lambda_F(1+C_X^2)$ | $\sqrt{\frac{N}{\lambda_F(1+C_X^2)}}$ |

TABLE 1.2 – Seuil de crédibilité complète et facteur de crédibilité

Les résultats ci-dessus supposent qu'une approximation à la loi normale pour toutes ces mesures de sinistres est possible. On suppose aussi souvent que le nombre de sinistres est distribué selon une loi de Poisson. Le coefficient de variation est supposé être donné ou estimé à partir des données disponibles. Le calcul des seuils de la crédibilité complète dépend de la donnée de la probabilité de couverture $(1-\alpha)$ et du paramètre de précision $k$.

**Limites de la crédibilité de stabilité**

L'approche de crédibilité de stabilité est certes simple à appliquer, mais elle présente certaines insuffisances que Pierre THEROND (2009) résume en les points suivants :

1. Cette approche met l'accent sur la valeur $D$ de l'expérience individuelle d'un contrat et n'attache aucune importance à l'information préalable que nous a donnée le paramètre $M$ du manuel d'assurance.
2. Le seuil de crédibilité complète dépend des valeurs de certains paramètres inconnus et l'approche ne rend pas compte de comment le calibrage des valeurs de ces paramètres affecte la crédibilité.
3. Certaines hypothèses sont compressées dans le but d'obtenir des résultats analytiques souples.

# Crédibilité Bayésienne

Lorsque l'on réalise une étude statistique, on a souvent des informations a priori provenant soit d'études antérieures soit d'avis d'experts. La statistique bayésienne permet de combiner ces connaissances a priori à l'information apportée par les données pour obtenir une information a posteriori. L'objectif de ce chapitre est de prévoir la prime (prime bayésienne) pour une période donnée en se servant des méthodes bayésienne. Dans ce chapitre, nous présentons l'approche de crédibilité bayésienne en rappelant tout d'abord la formule de Bayes (cas continu et cas discret) ensuite les distributions conjuguées et enfin les limites de cette approche. Les exemples de ce chapitre s'inspirent de ceux proposés par FOTSO (2014).

## 2.1 Théorème de Bayes

Soient $A$ et $B$ deux évènements de l'espace probabilisé $(\Omega, \mathcal{A}, Pr)$ ; par la formule des probabilités composées, si $Pr[A] \neq 0$ et $Pr[B] \neq 0$, on a

$$Pr[A \cap B] = Pr[A]Pr[B|A] = Pr[B]Pr[A|B]$$

d'où l'on a

$$Pr[B|A] = \frac{Pr[B]Pr[A|B]}{Pr[A]}. \tag{2.1}$$

Cette formule est connue sous le nom du **théorème de Bayes**.

$Pr[B]$ est la probabilité a priori de $B$ car elle précède toute information sur $A$. C'est aussi la loi marginale de $B$.

$Pr[B|A]$ est la probabilité a posteriori de $B$ car elle est postérieure à la réalisation de $A$. On améliore le théorème de Bayes de la manière suivante :

**Théorème 2.1.** Soit $(B_n)_{n\geq 1}$une suite d'évènements de $(\Omega, \mathcal{A}, Pr)$ qui forment une partition de $\Omega$. On suppose que $Pr[B_n] \neq 0 \ \forall n \geq 1$. Pour tout évènement $A$ de probabilité non nulle, on a

$$Pr[B_k|A] = \frac{Pr[B_k]Pr[A|B_k]}{\sum_{n\geq 1} Pr[B_n]Pr[A|B_n]}, \ \forall k \geq 1 \tag{2.2}$$

### 2.1.1 Distribution de probabilité a priori

**Définition 2.1.** Une distribution de probabilité a priori est l'ensemble des informations a priori disponibles sur le paramètre, ainsi que les imprécisions qui s'y rattachent.

**Remarque 2.1.1.** Le choix de la loi a priori est crucial pour l'analyse. Deux lois a priori différentes conduisent à deux systèmes d'inférence différents, même si l'accumulation de données finit par effacer cette différence.

### 2.1.2 Distribution de probabilité a posteriori

L'analyse a posteriori est effectuée après le recueil des informations sur l'échantillon.
L'expérience ayant été effectuée et l'échantillon $x_1, x_2, \ldots, x_n$ observé, la formule de Bayes est utilisée pour actualiser les probabilités a priori ( basées sur les croyances et les connaissances ).

**Exemple d'application 2.1.1.** Une machine fabrique des pièces par série de $1000$ dont une proportion $\theta$ est défectueuse. Cette proportion n'est pas connue mais on sait qu'elle peut prendre 5 valeurs : $0,03$ ; $0,05$ ; $0,12$ ; $0,18$ et $0,25$ avec les probabilités respectives de $0,2$ ; $0,4$ ; $0,2$ ; $0,1$ et $0,1$. Supposons que cette machine a produit $50$ pièces parmi lesquelles $4$ sont mauvaises. Déterminons la distribution de probabilité a posteriori sur $\Omega_\theta$.
Soit $A$ l'évènement "$4$ pièces sur $50$ sont défectueuses". Par le théorème de Bayes

$$Pr[\Theta = \theta_i|A] = \frac{Pr[A|\Theta = \theta_i]Pr[\Theta = \theta_i]}{\displaystyle\sum_{j=1}^{5} Pr[A|\Theta = \theta_j]Pr[\Theta = \theta_j]} \quad i = 1, 2, 3, 4, 5$$

avec

$$Pr[A|\Theta = \theta_j] = C_{50}^4 (\theta_j)^4 (1 - \theta_j)^{46}.$$

$$\begin{aligned}
Pr[A|\Theta = 0,03] &= 0,0460 \\
Pr[A|\Theta = 0,05] &= 0,1360 \\
Pr[A|\Theta = 0,12] &= 0,1334 \\
Pr[A|\Theta = 0,18] &= 0,0262 \\
Pr[A|\Theta = 0,25] &= 0,0016
\end{aligned}$$

La distribution a posteriori est alors

| $\Theta$ | $\theta = 0,03$ | $\theta = 0,05$ | $\theta = 0,12$ | $\theta = 0,18$ | $\theta = 0,25$ |
|---|---|---|---|---|---|
| $Pr[\Theta = \theta\|A]$ | $0,0988$ | $0,5846$ | $0,2867$ | $0,0281$ | $0,0017$ |

L'information que $A$ est réalisé a fortement modifiée les probabilités a priori.

## 2.2 Inférence Bayésienne

### Notations

Nous définissons ici les notations que nous allons utiliser dans toute la suite de notre chapitre.

- $X$ est une v.a. à valeurs dans $\Omega$, de loi $P_\theta$, $\theta \in \Omega_\theta$ et $\theta$ inconnu.
- $\theta$ est une réalisation d'une v.a $\Theta$; donc $P_\theta$ est considérée comme une loi conditionnelle $X|\Theta = \theta$ notée simplement $X|\theta$.
- $(X, \Theta)$ a pour densité conjointe $p(x, \theta)$.
- $X|\theta$ ou encore $P_\theta$ a pour densité $f(x|\theta)$.
- $f$ est la densité marginale de $X$.
- $g$ est la densité marginale de $\Theta$ (densité a priori).
- $\Theta|X = x$ a pour densité $g(\theta|x)$ (densité a posteriori).

### 2.2.1 Formule de Bayes

Avec les notations précédentes, on a :

- $f(x) = \int_{\Omega_\theta} p(x, \theta)\, d\theta$ ;
- $g(\theta) = \int_{\Omega} p(x, \theta)\, dx$ ;
- $p(x, \theta) = f(x|\theta)\, g(\theta) = g(\theta|x)\, f(x)$ donc

$$g(\theta|x) = \frac{f(x|\theta)\, g(\theta)}{f(x)} = \frac{f(x|\theta)\, g(\theta)}{\int_{\Omega_\theta} p(x, \theta)\, d\theta}.$$

C'est la formule de Bayes dans le cas continu.

### 2.2.2 Estimateur de Bayes

L'approche bayésienne suppose que pendant l'expérience on n'observe que des réalisations de $X$. En effet, $X$ dépend d'un paramètre $\theta$ inconnu et non observé .
Étant donnée une réalisation $x$ de $X$, il faut estimer la valeur $\theta$ de la réalisation non observée $\Theta$.

**Définition 2.2.** Étant donné $X$, **l'estimateur de Bayes** est l'espérance de la distribution a posteriori $\Theta|X$. C'est donc

$$\widehat{\Theta}(X) = E(\Theta|X).$$

**Définition 2.3.** Un estimateur $\widehat{\Theta}$ de $\Theta$ est au moins aussi bon qu'un estimateur $\widehat{\Theta}^*$ si

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

$\left(\widehat{\Theta}-\Theta\right)^2$ est l'erreur quadratique et $E\left[\left(\widehat{\Theta}-\Theta\right)^2\right]$ le risque quadratique commis en estimant $\Theta$ par $\widehat{\Theta}$.

**Théorème 2.2.** L'estimateur de bayes minimise le risque quadratique a posteriori.

**Preuve**. Soit $\widehat{\Theta}_1(X)$ un autre estimateur de $\theta$.

$$\begin{aligned}
E\left[\left(\widehat{\Theta}_1-\Theta\right)^2|X=x\right] &= E\left[\left(\widehat{\Theta}_1-\widehat{\Theta}+\widehat{\Theta}-\Theta\right)^2|X=x\right]\\
&= E\left[\left(\widehat{\Theta}_1-\widehat{\Theta}\right)^2|X=x\right]+2E\left[\left(\widehat{\Theta}_1-\widehat{\Theta}\right)\left(\widehat{\Theta}-\Theta\right)|X=x\right]+\\
&\quad E\left[\left(\widehat{\Theta}-\Theta\right)^2|X=x\right]\\
&= \left(\widehat{\Theta}_1(x)-\widehat{\Theta}(x)\right)^2+2\left(\widehat{\Theta}_1(x)-\widehat{\Theta}(x)\right)E\left[\left(\widehat{\Theta}-\Theta\right)|X=x\right]+\\
&\quad E\left[\left(\widehat{\Theta}-\Theta\right)^2|X=x\right]\\
&= \left(\widehat{\Theta}_1(x)-\widehat{\Theta}(x)\right)^2+E\left[\left(\widehat{\Theta}-\Theta\right)^2|X=x\right]\\
&\geq E\left[\left(\widehat{\Theta}-\Theta\right)^2|X=x\right]
\end{aligned}$$

■

**Exemple d'application 2.2.1.** Soit $X$ la v.a de Bernoulli de paramètre $\theta$. Si $\Theta$ suit une loi uniforme sur [0,1] c'est à dire $\Theta \hookrightarrow U_{[0,1]}$, déterminons l'estimateur bayésien $\widehat{\Theta}(X)$.

$$g(\theta)=\begin{cases}1 \text{ si } \theta\in[0,1],\\ 0 \text{ sinon}\end{cases} \qquad \text{et} \qquad f(x|\theta)=\begin{cases}\theta^x(1-\theta)^{1-x} \text{ si } x\in\Omega=\{0,1\}\,,\\ 0 \qquad\qquad \text{sinon.}\end{cases}$$

La loi a posteriori est

$$\begin{aligned}
g(\theta|x) &= \frac{f(x|\theta)g(\theta)}{\int_\Theta f(x|\theta)g(\theta)d\theta}\\
&= \begin{cases}\dfrac{\theta^x(1-\theta)^{1-x}}{\int_0^1\theta^x(1-\theta)^{1-x}d\theta} & \text{si } x\in\Omega\\ 0 & \text{sinon}\end{cases}\\
&= \begin{cases}2(1-\theta) & \text{si } x=0,\ \theta\in[0,1]\\ 2\theta & \text{si } x=1,\theta\in[0,1]\\ 0 & \text{sinon.}\end{cases}
\end{aligned}$$

- Si $X=0$, alors $\widehat{\Theta}(0)=\int_0^1\theta g(\theta|0)d\theta=2\int_0^1\theta(1-\theta)d\theta=\dfrac{1}{3}$.
- Si $X=1$, On a $\widehat{\Theta}(1)=\int_0^1\theta g(\theta|1)d\theta=2\int_0^1\theta^2d\theta=\dfrac{2}{3}$.

**Estimateur de Bayes dans le cas d'un échantillon de taille $n$**

Soit $X_1, X_2, \ldots, X_n$ un échantillon de taille $n$ de $X$, $x = (x_1, x_2, \ldots, x_n)$ une réalisation de cet échantillon et $L$ la vraisemblance. On a

$$L(x|\theta) = \prod_{i=1}^{n} f(x_i|\theta).$$

De la relation

$$L(x|\theta)g(\theta) = g(\theta|x)K(x)$$

où $K$ est la densité marginale de l'échantillon, on a la formule de Bayes

$$g\left(\theta|x\right) = \frac{L\left(x|\theta\right)g(\theta)}{K\left(x\right)} = \frac{\left(\prod_{i=1}^{n} f(x_i|\theta)\right)g(\theta)}{\int_{\Omega_\theta}\left(\prod_{i=1}^{n} f(x_i|\theta)\right)g(\theta)d\theta} \tag{2.3}$$

l'estimateur bayésien est alors

$$\widehat{\Theta}(X_1, X_2, \ldots, X_n) = E\left(\Theta|X_1, X_2, \ldots, X_n\right)$$

la moyenne de la distribution a posteriori.

**Estimateur de bayes dans le cas discret**

Pour simplifier la notation, posons plutôt $\mathbb{X} = (X_1, X_2, \ldots, X_n)$ et $x = (x_1, x_2, \ldots, x_n)$. Si $\Theta$ est discrète, la distribution a priori est exprimée sous forme d'une fonction de probabilité $Pr\left(\Theta = \theta\right), \theta \in \Omega_\theta$. La probabilité conjointe de $(X_1, X_2, \ldots, X_n)$ peut toujours être calculée par la loi des probabilités totales :

$$Pr(\mathbb{X} = x) = \sum_{\theta \in \Omega_\theta} Pr\left(\mathbb{X} = x|\Theta = \theta\right)Pr(\Theta = \theta).$$

La règle de Bayes permet de calculer la distribution a priori de $\Theta$

$$Pr(\Theta = \theta|\mathbb{X} = x) = \frac{Pr(\mathbb{X} = x|\Theta = \theta)Pr(\Theta = \theta)}{Pr(\mathbb{X} = x)}.$$

L'estimateur bayésien est alors

$$\widehat{\Theta}\left(X_1, X_2, \ldots, X_n\right) = E\left(\Theta|X_1, X_2, \ldots, X_n\right) = \sum_{\theta \in \Omega_\theta} \theta Pr\left(\Theta = \theta|\mathbb{X} = x\right).$$

**Exemple d'application 2.2.2.** Soit $X \looparrowright \mathcal{B}(1, \theta)$ et $(X_1, X_2, \ldots, X_n)$ un échantillon de taille $n$ de $X$. Si $\Theta \looparrowright Beta(\alpha, \beta)$, déterminons l'estimateur de Bayes de $\Theta$.

On a $f(x|\theta) = \theta^x(1-\theta)^{1-x}, x = 0,1$ et $g(\theta) = \dfrac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}, \theta \in [0,1]$.
Posons $x = (x_1, x_2, \ldots, x_n)$. La vraisemblance de l'échantillon est

$$L(x|\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} = \theta^{\sum_{i=1}^n x_i}(1-\theta)^{\sum_{i=1}^n (1-x_i)}$$

et la densité conjointe est

$$\begin{aligned} L(x,\theta) &= L(x|\theta)g(\theta) \\ &= \left[\theta^{\sum_{i=1}^n x_i}(1-\theta)^{\sum_{i=1}^n(1-x_i)}\right]\left[\frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)}\right] \\ &= \frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha,\beta)}. \end{aligned}$$

Comme

$$\begin{aligned} K(x) = \int_0^1 L(x,\theta)d\theta &= \int_0^1 \left[\frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha,\beta)}\right] d\theta \\ &= \frac{B(\alpha+n\bar{x}, \beta+n-n\bar{x})}{B(\alpha,\beta)}, \end{aligned}$$

la densité a posteriori est

$$g(\theta|x) = \frac{L(x,\theta)}{K(x)} = \frac{\dfrac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha,\beta)}}{\dfrac{B(\alpha+n\bar{x},\beta+n-n\bar{x})}{B(\alpha,\beta)}} = \frac{\theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+n-n\bar{x})-1}}{B(\alpha+n\bar{x},\beta+n-n\bar{x})}.$$

Donc la loi a posteriori de $\Theta$ est une loi $Beta(\alpha+n\bar{x}, \beta+n-n\bar{x})$. L'estimateur de Bayes est alors $\widehat{\Theta}(X_1, X_2, \ldots, X_n) = \dfrac{\alpha+n\bar{X}}{\alpha+\beta+n}$. Dans la formule (2.3) le dénominateur dépend seulement de $x$ et pas de $\theta$. En posant

$$N(x) = \frac{1}{K(x)} = \frac{1}{\displaystyle\int_{\Omega_\theta}\left(\prod_{i=1}^n f(x_i|\theta)\right) g(\theta)d\theta},$$

la densité a posteriori s'écrit $g(\theta|x) = N(x)L(x|\theta)g(\theta) \propto L(x|\theta)g(\theta)$. $N(x)$ est indépendant de $\theta$ et est considéré comme une constante de proportionnalité : c'est **une constante de normalisation**. L'expression $L(x|\theta)g(\theta)$ permet d'identifier la forme de la densité à posteriori sans calculer la densité marginale de $X$.

**Exemple d'application 2.2.3.** Soit $X \looparrowright \mathcal{B}(m,\theta)$ et $(X_1, X_2, \ldots, X_n)$ un échantillon de $X$. On suppose que $\Theta \looparrowright Beta(\alpha,\beta)$. Déterminons l'estimateur de Bayes de $\theta$.
on a

$$\begin{aligned} g(\theta|x) &\propto L(x,\theta)g(\theta) \\ &\propto \left[\theta^{n\bar{x}}(1-\theta)^{\sum_{i=1}^n(m-x_i)}\right]\left[\theta^{\alpha-1}(1-\theta)^{\beta-1}\right] \\ &\propto \theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+mn-n\bar{x})-1} \end{aligned}$$

donc

$$\Theta|x_1, x_2, \ldots, x_n \looparrowright Beta\left(\alpha + n\bar{x}, \beta + mn - n\bar{x}\right)$$

et l'estimateur bayésien est

$$\widehat{\Theta}\left(X_1, X_2, \ldots, X_n\right) = \frac{\alpha + n\bar{X}_n}{\alpha + \beta + mn} \text{ où } \overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

**Remarque 2.2.1.** Dans les exemples précédents, les distributions a priori et a posteriori appartiennent à la même famille de lois de probabilité, la loi bêta mais seuls les paramètres changent.

## 2.3 Distributions conjuguées

La détermination de la distribution a posteriori est une étape difficile dans l'application de l'inférence statistique bayésienne. En effet, l'égalité $(2.3)$ montre que la densité a posteriori est en général difficile à évaluer car c'est le rapport de deux termes dont le dénominateur est une intégrale. Il s'en suit la difficulté de déterminer l'estimateur bayésien.
Cependant, il existe des classes de distributions a priori qui sont conjuguées à certaines fonctions de vraisemblance en ce sens que la distribution a posteriori résultante appartient à la même famille de distributions que la loi a priori. Ainsi dans ces cas, les données observées ne changent pas le type de distribution, mais changent plutôt ses paramètres. La définition formelle est la suivante :

**Définition 2.4.** Soit $g(\theta|\gamma)$ une distribution a priori dépendant du paramètre inconnu $\gamma$ ( appelé hyper paramètre). $g(\theta|\gamma)$ est conjuguée à la fonction de vraisemblance $f(x|\theta)$ si la densité a posteriori est égale à $g(\theta|\gamma*)$. En d'autres termes les densités a priori et a posteriori appartiennent à la même famille de distributions.

Nous avons rencontré dans l'exemple précédent la distribution conjuguée beta-binomial et présentons dans les paragraphes qui suivent deux autres exemples.

### 2.3.1 Distribution conjuguée bêta-binomial (cas général)

Soient $X_1, X_2, \cdots, X_n$ $n$ v.a indépendantes telles que $\forall i = 1, \ldots, n\ \ X_i \looparrowright \mathcal{B}(m_i, \theta)$. On suppose que $\Theta \looparrowright Beta(\alpha, \beta)$. Déterminons la distribution a posteriori.
Posons $x = (x_1, x_2, \ldots, x_n)$. La vraisemblance s'écrit

$$\begin{aligned} L(x|\theta) &= \textstyle\prod_{i=1}^{n} C_{m_i}^{x_i} \theta^{x_i}(1-\theta)^{m_i - x_i} \\ &= \left[\textstyle\prod_{i=1}^{n} C_{m_i}^{x_i}\right]\left[\theta^{n\bar{x}}(1-\theta)^{\sum_{i=1}^{n}(m_i - x_i)}\right]. \end{aligned}$$

Posons $m = \sum_{i=1}^{n} m_i$. La densité a posteriori $g(\theta|x)$ est telle que

$$\begin{aligned} g(\theta|x) &\propto L(x|\theta)g(\theta|\alpha,\beta) \\ &\propto \left[\theta^{n\bar{x}}(1-\theta)^{(m-n\bar{x})}\right]\left[\theta^{\alpha-1}(1-\theta)^{\beta-1}\right] \\ &\propto \theta^{(\alpha+n\bar{x})-1}(1-\theta)^{(\beta+m-n\bar{x})-1} \end{aligned}$$

donc $\Theta|X_1, X_2, \ldots, X_n \looparrowright Beta(\alpha + n\bar{X}_n, \beta + m - n\bar{X}_n)$.

### 2.3.2 Distribution conjuguée Poisson-gamma

$X \looparrowright P(\lambda)$, $X_1, X_2, \ldots, X_n$ un échantillon de taille $n$ de $X$. $\lambda$ est une réalisation d'une v.a $\Lambda$ qui suit la loi gamma de paramètres $\alpha$ et $\beta$, $G(\alpha, \beta)$. Déterminons la distribution a posteriori.

Posons $x = (x_1, x_2, \ldots, x_n)$. La vraisemblance s'écrit

$$L(x|\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda}\lambda^{n\bar{x}}}{\prod_{i=1}^{n} x_i!}$$

et la densité a priori

$$g(\lambda) = \frac{\beta^{\alpha}\lambda^{\alpha-1}e^{-\lambda\beta}}{\Gamma(\alpha)}.$$

On a

$$\begin{aligned} g(\lambda|x) &\propto L(x|\lambda)g(\lambda) \\ &\propto \lambda^{\alpha+n\bar{x}-1}e^{-\lambda(n+\beta)}. \end{aligned}$$

En posant

$$\alpha^* = \alpha + n\bar{x} \text{ et } \beta^* = \beta + n,$$

on a

$$\Lambda|x_1, x_2, \ldots, x_n \looparrowright Gamma(\alpha^*, \beta^*).$$

La table suivante présente quelques distributions conjuguées.

| Distribution a priori | Distribution de $X$ | Distributions a posteriori |
|---|---|---|
| $Beta(\alpha, \beta)$ | $\mathcal{B}(1, \theta)$ | $Beta(\alpha + n\bar{x}, \beta + n - n\bar{x})$ |
| $Beta(\alpha, \beta)$ | $\mathcal{B}(m_i, \theta)$ | $Beta\left(\alpha + n\bar{x}, \beta + \sum_{i=1}^{n}(m_i - x_i)\right)$ |
| Gamma $(\alpha, \beta)$ | $\mathcal{P}(\lambda)$ | Gamma$(\alpha + n\bar{x}, \beta + n)$ |
| $Beta(\alpha, \beta)$ | $\mathcal{G}(\theta)$ | $Beta(\alpha + n, \beta + n\bar{x})$ |
| Gamma $(\alpha, \beta)$ | $\mathcal{E}(\lambda)$ | Gamma$\left(\alpha + n, \dfrac{1}{\beta + n\bar{x}}\right)$ |

TABLE 2.1 – Quelques distributions conjuguées

## 2.4 Crédibilité bayésienne

### 2.4.1 Exemple : bons risques/ mauvais risques

Considérons un portefeuille de polices d'assurance automobile dont les assurés sont divisés en 2 groupes : les bons risques et les mauvais risques. Notons $B$ l'évènement «être un bon risque » et $B^c$ son complémentaire « être un mauvais risque ». Notons par ailleurs $N_k$ le nombre de sinistres déclarés au cours de l'année $k$ et supposons que :

$$\begin{cases} Pr\left[N_k = 1|B\right] = 0,2 = 1 - Pr\left[N_k = 0|B\right] \\ Pr\left[N_k = 1|B^c\right] = 0,8 = 1 - Pr\left[N_k = 0|B^c\right]. \end{cases}$$

Les assurés ont donc au plus un sinistre par année et les bons risques en ont en moyenne quatre fois moins que les mauvais risques.

Supposons par ailleurs que le montant d'un sinistre est $1000F$. Soit $X$ le montant d'un sinistre sur un contrat. On a

$$X = \begin{cases} 1000 & \text{si } N_k = 1 \\ 0 & \text{si } N_k = 0. \end{cases}$$

La probabilité d'avoir un sinistre est

$$Pr\left[N_k = 1\right] = Pr\left[N_k = 1|B\right] Pr\left[B\right] + Pr\left[N_k = 1|B^c\right] Pr[B^c].$$

Quand un nouvel assuré arrive dans le portefeuille, comme l'assureur ne segmente pas son portefeuille a priori et qu'il ne dispose pas d'historique sinistre de cet individu, il va lui demander une prime collective. En supposant que les bons risques représentent $50\%$ des assurés, il vient que

$$\begin{aligned}\text{Prime collective} &= E(X)\\ &= 1000P(N_k=1)\\ &= 1000\left(Pr\left[N_k=1|B\right]Pr\left[B\right]+Pr\left[N_k=1|B^c\right]Pr\left[B^c\right]\right)\\ &= 1000(0,2\times 0,5+0,8\times 0,5)\\ &= 500.\end{aligned}$$

Ce calcul revient à faire l'hypothèse que la probabilité que le nouvel assuré soit un bon risque est identique à celle observée dans le portefeuille. Il nous a permis de déterminer la prime pure a priori. Intéressons-nous à présent à la mise à jour du tarif d'un assuré dont on dispose de trois années d'historiques $(N_1=1,N_2=1,N_3=0)$. Comme, conditionnellement à la qualité d'un risque, les variables aléatoires nombres annuels de sinistres $N_k$ sont indépendantes, on peut écrire

$$\begin{aligned}Pr\left[N_1=1,N_2=1,N_3=0\right] &= Pr\left[N_1=1,N_2=1,N_3=0|B\right]Pr\left[B\right]\\ &+ Pr\left[N_1=1,N_2=1,N_3=0|B^c\right]Pr[B^c]\\ &= Pr\left[N_1=1|B\right]Pr\left[N_2=1|B\right]Pr\left[N_3=0|B\right]Pr\left[B\right]\\ &+ Pr\left[N_1=1|B^c\right]Pr\left[N_2=1|B^c\right]Pr\left[N_3=0|B^c\right]Pr\left[B^c\right]\\ &= 0,2\times 0,2\times 0,8\times 0,5+0,8\times 0,8\times 0,2\times 0,5=0,08\end{aligned}$$

Ainsi en moyenne $8\%$ des assurés ont eu ce parcours sur les trois années.

La tarification bayésienne repose sur l'historique des sinistres pour réviser le montant de la prime pure. En effet, si l'assureur ne peut observer directement si tel assuré est un bon risque ou un mauvais risque, le passé sinistre va néanmoins lui fournir une information sur la qualité du risque. En effet,

$$\begin{aligned}Pr[B|N_1=1,N_2=1,N_3=0] &= \frac{Pr[B,N_1=1,N_2=1,N_3=0]}{Pr[N_1,N_2=1,N_3=0]}\\ &= \frac{Pr[N_1=1,N_2=1,N_3=0|B]Pr[B]}{Pr[N_1=1,N_2=1,N_3=0]}\\ &= 0,2.\end{aligned}$$

il y a donc $20\%$ de chances qu'un assuré ayant ce passé sinistre soit un bon risque. Rappelons que cette probabilité est de $50\%$ a priori.

La prime pure pour un bon risque est

$$p(B)=E\left[X|B\right]=1000Pr\left[N_k=1|B\right]$$

et la prime pure pour un mauvais risque est

$$p(B^c)=E[X|B^c].$$

La tarification bayésienne consiste à exiger de notre assuré, pour la période $[3,4]$ ; le montant

$$Pr\left[B|N_1=1,N_2,N_3=0\right]p(B)+Pr\left[B^c|N_1=1,N_2,N_3=0\right]p(B^c).$$

Soit

$$0,2\times 0,2\times 1000+0,8\times 0,8\times 1000=680.$$

La prime de cet assuré va donc passer de $500$frs la première année à $680$frs la quatrième année.

### 2.4.2 Prime bayésienne

On considère un portefeuille hétérogène de plusieurs polices d'assurance divisé en plusieurs groupes de risque .

$X$ une variable aléatoire représentant pour un groupe de risque

- le nombre de sinistre sur un contrat.
- le montant sinistre du contrat ou bien.
- la charge sinistre totale des contrats du groupe.

$X$ dépend d'un paramètre de risque $\theta$ considéré comme une réalisation d'une v.a.r $\Theta$ de densité de probabilité $g(\theta)$

Soit $\mathbb{X}=(X_1,X_2,\ldots,X_n)$ un échantillon de taille $n$ de $X$ et $x_1,x_2\ldots,x_n$ une réalisation de $\mathbb{X}$.

**Problème** : on veut estimer

$$\mu_X(\Theta)=E\left(X|\Theta\right).$$

Soit $T(x)$ un estimateur de $\mu_x(\Theta)$. En utilisant la fonction de perte quadratique

$$\mathcal{L}\left(\mu_X(\Theta),T(x)\right)=\left(\mu_X(\Theta)-T(x)\right)^2$$

le risque est

$$E\left(\mathcal{L}\left(\mu_X(\Theta),T(x)\right)\right)=\int_{\Omega_\theta}(\mu_X(\theta)-T(x))^2g(\theta|x)d\theta$$

donc l'estimateur de $\theta$ est

$$T^*(x)=E\left(\mu_X(\Theta)|x\right).$$

L'estimateur de bayes de $\mu_x(\Theta)$ que nous notons $\widehat{\mu}_X(x)$ est :

$$\widehat{\mu}_X(x)=E\left(\mu_X(\Theta)|x\right)=\int_{\Omega_\theta}\mu_X(\theta)g(\theta|x)d\theta. \tag{2.4}$$

Si $X$ est le montant d'un sinistre , $\widehat{\mu}_X(x)$ est appelée **prime bayésienne. La prime collective** ou **prime pure** est

$$E(X)=E(\mu_X(\Theta))=E(E(X|\Theta))=\int_{\Omega_\theta}\mu_X(\theta)g(\theta)d\theta.$$

**Remarque 2.4.1.**

- La prime collective est une moyenne pondérée de $\mu_X(\theta)$ en utilisant la distribution a priori, alors que la prime bayésienne est une moyenne pondérée de $\mu_X(\theta)$ en utilisant la distribution a posteriori.
- La prime collective ou prime pure peut être interprétée comme la prime bayésienne de la première année, lorsqu'aucune expérience n'est disponible.

**Exemple d'application 2.4.1.** Soit $X$ une v.a désignant le montant d'un sinistre. On suppose que $X|\Theta = \theta \hookrightarrow P(\theta)$ et la distribution a priori de $\Theta$ est

$$Pr\left[\Theta = \theta\right] = \begin{cases} 0,3 \text{ si } \theta = \dfrac{1}{2} \\ 0,4 \text{ si } \theta = 1 \\ 0,3 \text{ si } \theta = 2. \end{cases}$$

1. Calculons la prime pure dans chaque groupe de risque.

   $\mu_X(\Theta) = E(X|\Theta) = \Theta$ donc,

$$\begin{array}{ll} \text{si } \theta = \dfrac{1}{2} & \mu_X(\dfrac{1}{2}) = \dfrac{1}{2} \\ \text{si } \theta = 1 & \mu_X(1) = 1 \\ \text{si } \theta = 2 & \mu_X(2) = 2. \end{array}$$

2. Calculons la prime pure collective.

   On a

$$E(X) = E(\mu_X(\Theta)) = E(\Theta) = 0,5 \times 0,3 + 1 \times 0,4 + 2 \times 0,3 = 1,15.$$

3. Calculons la prime bayésienne pour la quatrième année si $X_1 = 2$, $X_2 = 1$, $X_3 = 1$.

   Déterminons tout d'abord la distribution a posteriori de $\Theta$.

   Posons $x = (2, 1, 1)$, en appliquant la formule de Bayes on a

$$Pr\left[\Theta = \theta|x\right] = \frac{Pr\left[\Theta = \theta\right] \prod_{i=1}^{i=3} Pr\left[X_i = x_i|\Theta = \theta\right]}{\sum_{\theta \in \Omega_\theta} Pr\left[\Theta = \theta\right] \prod_{i=1}^{i=3} Pr\left[X_i = x_i|\Theta = \theta\right]}$$

avec

$$\begin{aligned} Pr\left[X_1 = 2|\Theta = \theta\right] &= \frac{e^{-\theta}\theta^2}{2!} \\ Pr\left[X_2 = 1|\Theta = \theta\right] &= \frac{e^{-\theta}\theta^1}{1!} \\ Pr\left[X_3 = 1|\Theta = \theta\right] &= \frac{e^{-\theta}\theta^1}{1!}. \end{aligned}$$

On obtient donc

$$Pr\left[\Theta = 0,5|x\right] = 0,0113, \quad Pr\left[\Theta = 1|x\right] = 0,673 \text{ et } Pr\left[\Theta = 2|x\right] = 0,214.$$

D'où la prime bayésienne est

$$\begin{aligned} E(\mu_X(\Theta)|x) &= E(\Theta|x) \\ &= (0,5 \times 0,113) + (1 \times 0,673) + (2 \times 0,214) \\ &= 1,1575. \end{aligned}$$

On constate que

$$E(\mu_X(\Theta)) = 1,15 < E(\mu_X(\Theta)|x) < \bar{x} = 1,333.$$

## 2.5 Distribution prédictive

L'objectif dans cette partie est de prévoir la valeur future de $X_{n+1}$ étant donné $x = (x_1, x_2, \ldots, x_n)$

**Définition 2.5.** On appelle distribution prédictive de la variable aléatoire $X_{n+1}$, la distribution de $X_{n+1}|x = (x_1, x_2, \ldots, x_n)$ avec comme fonction densité de probabilité $f_{x_{n+1}}(x_{n+1}|x_1, x_2, \ldots, x_n)$.

**Théorème 2.3.** FOTSO(2014). La fonction densité de probabilité de la distribution prédictive de $X_{n+1}$ est :

$$f_{x_{n+1}}(x_{n+1}|x_1, x_2, \ldots, x_n) = \int_{\Omega_\theta} f(x_{n+1}|\theta) g(\theta|x_1, x_2, \ldots, x_n) d\theta. \tag{2.5}$$

**Preuve**.

$$\begin{aligned} f_{x_{n+1}}(x_{n+1}|x_1, x_2, \ldots, x_n) &= \frac{L_1(x_{n+1}, x)}{K(x)} \\ &= \frac{\int_{\Omega_\theta} L_1(x_{n+1}, x|\theta) g(\theta) d\theta}{\int_{\Omega_\theta} L(x|\theta) g(\theta) d\theta} \\ &= \int_{\Omega_\theta} f(x_{n+1}|\theta) \left[ \frac{L(x|\theta) g(\theta)}{\int_{\Omega_\theta} L(x|\theta) g(\theta) d\theta} \right] d\theta \\ &= \int_{\Omega_\theta} f(x_{n+1}|\theta) g(\theta|x_1, x_2, \ldots, x_n) d\theta. \end{aligned}$$

■

**Remarque 2.5.1.**

Puisque la densité de $X$ s'écrit

$$f(u) = \int_{\Omega_\theta} f(u|\theta) g(\theta) d\theta,$$

la seule différence entre l'expression de $f(u)$ et celle de $f_{x_{n+1}}(x_{n+1}|x_1, x_2, \ldots, x_n)$ vient du fait que l'on utilise la distribution a priori de $\Theta$ pour la première et la distribution a posteriori pour la seconde. Ainsi,

$$\begin{array}{ccc} g(\theta|x_1, x_2, \ldots, x_n) & \Longrightarrow & f_{x_{n+1}}(x_{n+1}|x_1, x_2, \ldots, x_n) \\ \text{du même type que } g(\theta) & & \text{du même type que } f(x) \end{array}$$

À partir du théorème précédent, montrons que la prime bayésienne $E(\mu_X(\Theta)|x)$ est la meilleure prévision de $X_{n+1}$.

$$\begin{aligned} E(\mu_X(\Theta)|x_1, x_2, \ldots, x_n) &= \int_{\Omega_\theta} \mu_X(\theta) g(\theta|x_1, x_2, \ldots, x_n) d\theta \\ &= \int_{\Omega_\theta} \left( \int_\Omega u f(u|\theta) du \right) g(\theta|x_1, x_2, \ldots, x_n) d\theta \\ &= \int_{\Omega_\theta} \int_\Omega u f(u|\theta) g(\theta|x_1, x_2, \ldots, x_n) du d\theta \\ &= \int_\Omega u \left( \int_{\Omega_\theta} f(u|\theta) g(\theta|x_1, x_2, \ldots, x_n) d\theta \right) du \\ &= \int_\Omega u f_{x_{n+1}}(u|x_1, x_2, \ldots, x_n) du \end{aligned}$$

donc

$$E(\mu_X(\Theta)|x_1, x_2, \cdots, x_n) = E(X_{n+1}|x_1, x_2, \ldots, x_n) \tag{2.6}$$

La prime bayésienne peut donc être interprétée comme la moyenne conditionnelle de $X_{n+1}$ sachant $\mathbb{X}$.

En resumé, l'estimateur de Bayes de la moyenne $E(X)$, appelé prime bayésienne, est la moyenne de $X$ sachant les données $x = (x_1, x_2, \cdots, x_n)$ en prenant la distribution a priori de $\Theta$ voir la formule (2.4). C'est également la moyenne conditionnelle de la valeur future $X_{n+1}$ (voir (2.6)).

**Exemple d'application 2.5.1.** Soit $X$ une v.a désignant le nombre de sinistres. On suppose que $X|\Theta = \theta \looparrowright P(\theta)$ et $\Theta \looparrowright Gamma(\alpha, \lambda)$.

1) Calculons la prime pure pour chaque niveau de risque.

   On a $\mu_X(\Theta) = E(X|\Theta) = \Theta$ donc $E(X|\Theta = \theta) = \theta$.

2) Calculons la prime pure collective.

   On a

$$E(\mu_X(\Theta)) = E(\Theta) = \frac{\alpha}{\lambda}.$$

3) Déterminons la prime bayésienne après $n$ années.

   Posons $x = (x_1, x_2, \ldots, x_n)$

$$\Theta|x \looparrowright \; Gamma(\alpha^*, \lambda^*)$$

   avec $\alpha^* = \alpha + n\bar{x}$ et $\lambda^* = \lambda + n$. La prime bayésienne est donc

$$E(\Theta|x) = \frac{\alpha^*}{\lambda^*} = \frac{\alpha + n\bar{x}}{\lambda + n}.$$

4) Déterminons la distribution marginale de $X$ puis calculons $E(X)$ et $Var(X)$.

La distribution marginale de X est :

$$\begin{aligned}
f(u) &= \int_0^{+\infty} f(u/\theta)g(\theta)d\theta \\
&= \int_0^{+\infty} \frac{\theta^u e^{-\theta}}{u!} \frac{\lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}}{\Gamma(\alpha)} d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)u!} \int_0^{+\infty} \theta^u e^{-\theta} \theta^{\alpha-1} e^{-\lambda\theta} d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)\Gamma(u+1)} \int_0^{+\infty} \theta^{\alpha+u-1} e^{-(\lambda+1)\theta} d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)\Gamma(u+1)} \times \frac{\Gamma(\alpha+u)}{(\lambda+1)^{\alpha+u}} \\
&= \frac{\Gamma(\alpha+u)}{\Gamma(\alpha)\Gamma(u+1)} \left(\frac{\lambda}{\lambda+1}\right)^\alpha \left(\frac{1}{\lambda+1}\right)^u \\
&= C_{\alpha+u-1}^{\alpha-1} p^\alpha (1-P)^u, u = 0, 1, 2, \ldots,
\end{aligned}$$

d'où

$$X \looparrowright BinNeg(r,p) \text{ avec } r = \alpha \text{ et } p = \frac{\lambda}{\lambda+1}.$$

Donc

$$E(X) = \frac{r(1-p)}{p} = \frac{\alpha}{\lambda} \quad \text{et} \quad Var(X) = \frac{r(1-p)}{p^2} = \frac{\alpha(\lambda+1)}{\lambda^2}.$$

5) Déterminons la distribution conditionnelle de $X_{n+1}$ après $n$ années et en déduisons la prime bayésienne.

Les distributions a priori et a posteriori sont du même type :

$$\Theta \looparrowright Gamma(\alpha, \lambda) \quad \text{et} \quad \Theta/x \looparrowright Gamma(\alpha + n\bar{x}, \lambda + n).$$

Donc

$$X_{n+1}|x \looparrowright BinNeg(r^*, p^*) \quad \text{avec} \quad r^* = \alpha + n\bar{x} \quad \text{et} \quad p^* = \frac{\lambda+n}{\lambda+n+1}.$$

La prime bayésienne est alors

$$E(X_{n+1}|x) = \frac{r^*(1-p^*)}{p^*} = \frac{\alpha+n\bar{x}}{\lambda+n}.$$

6) Calculons la prime Bayésienne pour les douze prochaines années si a priori $\Theta \looparrowright Gamma(2,3)$ et que les montants des sinistres au cours de ces années sont : $3, 4, 1, 0, 2, 0, 2, 3, 1, 1, 3, 5$.

| $n$ | $x_n$ | $n\bar{x}$ | $\alpha + n\bar{x}$ | $\lambda + n$ | $p_{bayes}$ |
|---|---|---|---|---|---|
| 0 | - | - | 2 | 3 | 0,666 |
| 1 | 3 | 3 | 5 | 4 | 1,25 |
| 2 | 4 | 7 | 9 | 5 | 1,8 |
| 3 | 1 | 8 | 10 | 6 | 1,666 |
| 4 | 0 | 8 | 10 | 7 | 1,428 |
| 5 | 2 | 10 | 12 | 8 | 1,5 |
| 6 | 0 | 10 | 12 | 9 | 1,333 |
| 7 | 2 | 12 | 14 | 10 | 1,4 |
| 8 | 3 | 15 | 17 | 11 | 1,545 |
| 9 | 1 | 16 | 18 | 12 | 1,5 |
| 10 | 1 | 17 | 19 | 13 | 1,461 |
| 11 | 3 | 20 | 22 | 14 | 1,571 |
| 12 | 5 | 25 | 27 | 15 | 1,8 |

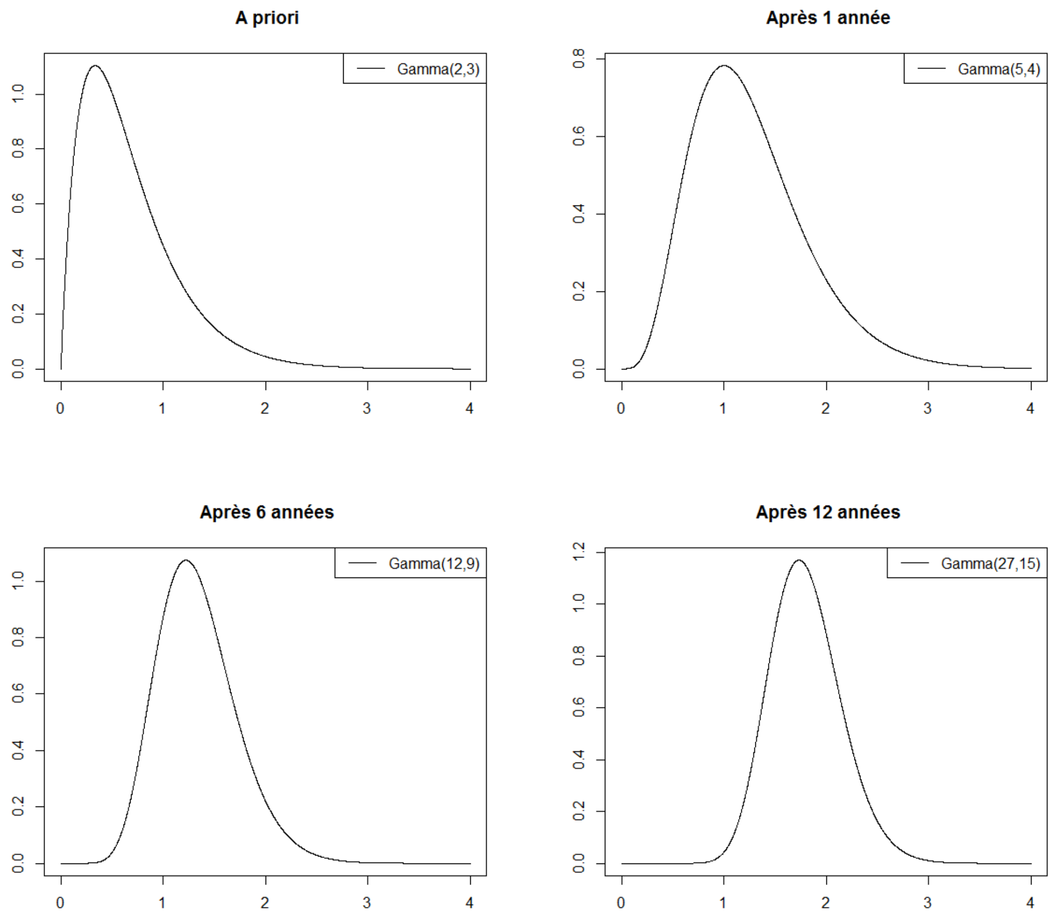


FIGURE 2.1 – Distribution a priori de $\Theta$ et après quelques années d'observation

La figure 2.1 présente la distribution a priori de $\Theta$ et la distribution de $\Theta$ après 1, 6 et 12 années d'observation. L'on constate que la distribution a posteriori de $\Theta$ est de plus en plus concentrée autour de sa moyenne au fur et à mesure que l'expérience s'accumule. Ainsi la précision de la prime bayésienne s'améliore.

### 2.5.1 Crédibilité bayésienne linéaire ou exacte

Une prime de la forme

$$P = Z\bar{x} + (1 - Z)E(X) \tag{2.7}$$

est appelée **prime de crédibilité** et $0 \leq Z \leq 1$ est le **facteur de crédibilité**.
Dans l'exemple $(2.5.1)$ la prime bayésienne est :

$$\begin{aligned} E(\Theta|x) &= \frac{\alpha + n\bar{x}}{\lambda + n} \\ &= \frac{n}{n+\lambda}\bar{x} + \frac{\lambda}{n+\lambda}.\frac{\alpha}{\lambda} \\ &= Z\bar{x} + (1 - Z)E(X) \end{aligned}$$

avec $Z = \dfrac{n}{n+\lambda}$.
La prime bayésienne est donc linéaire dans le cas Poisson/gamma. On peut également le vérifier dans les cas Bernoulli/bêta, exponentielle/gamma, Normale/normale et Géométrique/bêta.
Whitney $(1918)$ et Bailey $(1950)$ furent les premiers à démontrer que la prime bayésienne est une prime de crédibilité linéaire pour certaines combinaisons de distributions.
Jewell (1974) démontre que lorsque les distributions sont conjuguées alors la prime bayésienne est toujours une prime de crédibilité exacte.

## 2.6 Limites de la crédibilité bayésienne

La crédibilité bayésienne pose deux problèmes pratiques

- La prime bayésienne est une prime de crédibilité dans certains cas seulement.
- Le calcul de la prime bayésienne repose sur des hypothèses subjectives pour des distributions de $\Theta$ et de $X|\Theta$.

Pour le premier problème, Bühlmann $(1967)$ restreint l'approximation de la prime de risque aux fonctions linéaires des observations. En utilisant le critère des moindres carrés, il trouve que la meilleur approximation est une prime de crédibilité. Pour le second problème, Bühlmann $(1969)$ le contourne en utilisant une approche non paramétrique pour calculer la prime de crédibilité.

⋆ ⋆ **Chapitre Trois** ⋆ ⋆

# Crédibilité de Bühlmann

Le calcul de la prime selon une approche bayésienne nécessite la connaissance des lois a priori et a posteriori. Comme il a été mentionné dans le chapitre précédent, l'approximation de la prime de risque sous une approche bayésienne peut donc s'avérer complexe. Pour contourner le problème de complexité de calcul de la prime bayésienne, Bühlmann $(1967, 1969)$ propose des modèles qui reposent essentiellement sur une approximation linéaire de la prime de risque. Il obtient ainsi une approximation de la prime de risque qui peut s'exprimer sous la forme d'une prime de crédibilité. Son premier modèle, le modèle original (1967), s'applique à un portefeuille composé d'un seul contrat. Par là, il vient corriger une faiblesse de la crédibilité bayésienne notamment celle de la complexité du calcul. Cependant, en pratique, les lois a posteriori sont bien souvent inconnues de l'assureur. Bühlmann a évité ce problème par son modèle classique $(1969)$, où l'on considère plutôt un portefeuille composé de plusieurs contrats. Ainsi, les observations de l'ensemble du portefeuille pourront être utilisées pour estimer les paramètres nécessaires au calcul de la prime de crédibilité.
Ce chapitre présente le modèles original et classique de Buhlmann $(1967, 1969)$. Pour chacun de ces modèles, on expose les hypothèses permettant de déterminer la prime de crédibilité et nous en donnons une expression de cette prime.

## 3.1 Notations et relations de covariance

### 3.1.1 Notations

On considère un contrat d'un portefeuille de police d'assurance. La notation suivante sera utilisée tout au long de ce chapitre.

- $X$ est une v.a pouvant désigner pour un groupe à risque le nombre de sinistres, le montant d'un sinistre ou la charge totale sinistres.
- $\theta$ désigne le niveau de risque d'un contrat et est considéré comme une réalisation d'une variable aléatoire $\Theta$.

- L'espérance conditionnelle et la variance de $X$ sachant $\theta$ sont données respectivement par $E(X/\theta) = \mu(\theta)$ et $Var(X/\theta) = \sigma^2(\theta)$.
- $E(X|\Theta)$ et $Var(X|\Theta)$ sont des variables aléatoires en $\Theta$ et sont désignées respectivement par $\mu(\Theta) = E(X|\Theta)$ et $\sigma^2(\Theta) = Var(X|\Theta)$.
- $s^2 = E[\sigma^2(\Theta)]$.
- $a = Var[\mu(\Theta)]$.
- $m = E\left[\mu(\Theta)\right]$.

### 3.1.2 Relations de covariance

**Lemme 3.1.** Pour toute variable aléatoire $X$, on a :

i) $E\left[X\right] = E\left[E\left[X|\Theta\right]\right]$,

ii) $E\left[X - E\left[X\right]\right] = 0$.

GOULET (2010) énonce les théorèmes 3.1 et 3.2 suivants :

**Théorème 3.1.** Soient $X$, $Y$ et $\Theta$ des variables aléatoires donc la densité conjointe existe. Alors

$$Cov(X,Y) = Cov(E\left[X|\Theta\right], E\left[Y|\Theta\right]) + E\left[Cov(X,Y|\Theta)\right].$$

**Preuve**.

$$\begin{aligned}
Cov(X,Y) &= E[(X - E\left[X\right])(Y - E\left[Y\right])] \\
&= E[(X - E\left[X|\Theta\right] + E\left[X|\Theta\right] - E\left[X\right])(Y - E\left[Y|\Theta\right] + E\left[Y|\Theta\right] - E\left[Y\right])] \\
&= E[E[(X - E[X|\Theta])(Y - E[Y|\Theta])|\Theta]] \\
&\quad + E[E[(X - E[X|\Theta])(E[Y|\Theta] - E[Y])|\Theta]] \\
&\quad + E[E[(E[X|\Theta] - E[X])(Y - E[Y|\Theta])|\Theta]] \\
&\quad + E[(E[X|\Theta] - E[X])(E[Y|\Theta] - E[Y])] \\
&= E[E[(X - E[X|\Theta])(Y - E[Y|\Theta])|\Theta]] \\
&\quad + E[(E[Y|\Theta] - E[Y])E[X - E[X|\Theta]|\Theta]] \\
&\quad + E[(E[X|\Theta] - E[X])E[Y - E[Y|\Theta]|\Theta]] \\
&\quad + E[(E[X|\Theta] - E[X])(E[Y|\Theta] - E[Y])] \\
&= E[Cov(X,Y|\Theta] + 0 + 0 + Cov(E[X|\Theta, E[Y|\Theta]).
\end{aligned}$$

■

**Corollaire 3.1.** En posant $X = Y$ dans le théorème 3.1, on obtient

$$Var\left[X\right] = E\left[Var(X|\Theta)\right] + Var(E\left[X|\Theta\right]) = s^2 + a.$$

**Théorème 3.2.** [8] Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes et identiquement distribuée conditionnellement à la variable aléatoire $\Theta$ avec espérance $\mu(\Theta)$ et variance $\sigma^2(\Theta)$. Alors

$$\begin{aligned} Cov(X_j, X_r) &= \begin{cases} a, & j \neq r \\ a + s^2, & j = r \end{cases} \\ &= a + \delta_{jr} s^2, \quad j, r = 1, \ldots, t \\ Cov(\mu_X(\Theta), X_t) &= a, \end{aligned}$$

où $\delta_{jr}$ est le delta de Kronecker :

$$\delta_{jr} = \begin{cases} 1, & j = r \\ 0, & j \neq r. \end{cases}$$

**Preuve**. Pour le premier résultat, on a

$$\begin{aligned} Cov(X_j, X_r) &= Cov(E\left[X_j|\Theta\right], E[X_r|\Theta]) + E\left[Cov(X_j, X_r|\Theta)\right] \\ &= Cov(\mu(\Theta), \mu(\Theta)) + E\left[\delta_{jr} Var\left[X_j|\Theta\right]\right] \\ &= Var\left[\mu(\Theta)\right] + \delta_{jr} E\left[\sigma^2(\Theta)\right] \\ &= a + \delta_{jr} s^2. \end{aligned}$$

De plus,

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

■

## 3.2 Modèle original de Bühlmann

Le modèle original de Bühlmann $(1967)$ suppose un portefeuille composé d'un seul contrat, dont l'expérience est observée pendant $t$ périodes. Par conséquent, le calcul de la prime est développé essentiellement sous une approche bayésienne pure. Puisque la connaissance des lois a priori et a posteriori est nécessaire, ce modèle est dit paramétrique.

### 3.2.1 Hypothèse du modèle original de Bühlmann

Le modèle original de Bühlmann $(1967)$ est basé sur l' hypothèse suivante :

$(H)$ Pour un niveau de risque $\Theta = \theta$ donné, les variables $X_1, X_2, \ldots, X_t$ sont conditionnellement indépendantes et identiquement distribuées et ont une variance finie.

On aura donc :

1) $E[X_j|\Theta] = \mu(\Theta), \qquad j, r = 1, ..., t.$

2) $Cov(X_j, X_r|\Theta) = \delta_{jr}\sigma^2(\Theta), \quad j, r = 1, \ldots, t.$

**Remarque 3.2.1.** L'hypothèse $(H)$ établit l'indépendance pour le contrat entre chacune des périodes d'observation. Cette hypothèse stipule l'homogénéité des montants des sinistres dans le temps puisque pour une valeur de $\Theta$ donnée, les variables $X_1, X_2, \ldots, X_n$ sont identiquement distribuées.

### 3.2.2 Calcul de la prime

Le modèle original de Bühlmann (1967) propose une approximation de la prime de risque par des fonctions linéaires des observations. Comme l'a montré Jewell (1974), lorsqu'une fonction de vraisemblance est combinée avec sa conjuguée naturelle, la prime bayésienne est linéaire. Dans ces cas particuliers, la fonction linéaire obtenue par le modèle original de Bühlmann $(1967)$ est donc équivalente à la prime bayésienne.
Dans le modèle original de Bühlmann $(1967)$, l'approximation de la prime de risque est obtenue par une fonction linéaire non homogène des observations. Ainsi, la meilleure prime, au sens des moindres carré aura la forme suivante :

$$c_0 + \sum_{j=1}^{t} c_j X_j. \tag{3.1}$$

Afin d'obtenir la fonction linéaire non homogène qui constitue une approximation de la prime risque, on doit trouver les coefficients $c_0^*, c_1^*, \ldots, c_t^*$ qui minimisent

$$E\left[\left(\mu(\Theta) - c_0 - \sum_{j=1}^{t} c_j X_j\right)^2\right]. \tag{3.2}$$

Le théorème suivant détermine la prime de crédibilité non homogène pour la période $t+1$ selon le modèle original de Bühlmann $(1969)$.

**Théorème 3.3.** Soient $X_1, \ldots, X_t$ des variables aléatoires avec fonction de répartition $F(x|\theta)$. Sous les hypothèses du modèle original de Bühlmann $(1969)$, la prime linéaire $P_{i,t+1}$ de la prime de risque du contrat $i$ pour la période $t+1$ telle que $(3.2)$ soit minimale, est de la forme

$$P_{i,t+1} = z\overline{X} + (1-z)m, \tag{3.3}$$

où

$$z = \frac{at}{at + s^2}, \tag{3.4}$$

$$\overline{X} = \frac{1}{t}\sum_{j=1}^{t} c_j X_j, \tag{3.5}$$

$$m = E[\mu(\Theta)]. \tag{3.6}$$

**Preuve**. Les coefficients $c_0^*, c_1^*, \ldots, c_t^*$ qui minimisent $(3.2)$ s'obtiennent en posant nulles les dérivées partielles par rapport à chacun de ces coefficients. On a

$$\frac{\partial}{\partial c_0} E\left[\left(\mu(\Theta) - c_0 - \sum_{j=1}^{t} c_j X_j\right)^2\right] = -2E\left[\left(\mu(\Theta) - c_0 - \sum_{j=1}^{t} c_j X_j\right)\right] = 0,$$

d'où

$$c_0^* = E[\mu(\Theta)] - \sum_{j=1}^{t} c_j E[X_j] = E[\mu(\Theta)] - \sum_{j=1}^{t} c_j m \tag{3.7}$$

En insérant cette expression de $c_0^*$ dans l'équation $(3.2)$, on détermine les coefficients $c_1^*, \ldots, c_t^*$ qui la minimisent. Ainsi, pour $q = 1, \ldots, t$,

$$\frac{\partial}{\partial c_q} E\left[\left(\mu(\Theta) - m - \sum_{j=1}^{t} c_j (S_j - m)\right)^2\right] = -2E\left[\left(\mu(\Theta) - m - \sum_{j=1}^{t} c_j (X_j - m)\right)(X_q - m)\right] = 0.$$

Ce qui entraine

$$\mu(\Theta) = m + \sum_{j=1}^{t} c_j (X_j - m),$$

et donc

$$Cov(\mu(\Theta), X_q) = \sum_{j=1}^{t} c_j Cov(X_j, X_q). \tag{3.8}$$

Étant donné le théorème $(3.2)$, l'équation $(3.8)$ peut se réécrire

$$a = \sum_{j=1}^{t} c_j (a + \delta_{jq} s^2), \quad q = 1, \cdots, t. \tag{3.9}$$

Écrit sous forme matricielle, l'équation $(3.9)$ devient

$$\begin{pmatrix} a+s^2 & a & \cdots & a \\ a & a+s^2 & \cdots & a \\ \vdots & \vdots & \ddots & \vdots \\ a & a & \cdots & a+s^2 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_t \end{pmatrix} = \begin{pmatrix} a \\ a \\ \vdots \\ a \end{pmatrix}$$

Puisque ce système est symétrique, les coefficients $c_j$ doivent être égaux dans la solution. L'expression de $c_j$ sera donc

$$c_j^* = \frac{a}{at + s^2}, \quad j = 1, \ldots, t. \tag{3.10}$$

De l'équation $(3.7)$, on obtient

$$c_0^* = (1 - \frac{at}{at + s^2})m.$$

On obtient la prime de crédibilité en remplaçant les coefficients $c_0^*, \ldots, c_t^*$

Et donc

$$c_0^* + \sum_{j=1}^{t} c_j X_j = \frac{at}{at + s^2} \sum_{j=1}^{t} \frac{X_j}{t} + \left(1 - \frac{at}{at + s^2}\right) m$$

En posant $z = \dfrac{at}{at+s^2}$, on trouve

$$P_{i,t+1} = z\overline{X} + (1-z)m.$$

La prime de crédibilité consiste donc en un compromis entre l'expérience individuelle $\overline{X}$ d'un contrat et la prime collective $m$. Tel que défini dans les notations, le paramètre $a$ représente la variance des primes de risque du portefeuille. Puisque l'on ne considère qu'un seul contrat dans le modèle original, $a$ représente la variabilité de la prime de risque $\mu(\Theta)$. Une valeur élevée de $a$ indique que la prime de risque $\mu(\Theta)$ est portée à varier de façon importante autour de sa moyenne au cours des différentes périodes d'observation. Par conséquent, l'espérance de la prime de risque $E[\mu(\Theta)]$ n'est pas une approximation juste et il est préférable d'accorder une plus grande importance à l'expérience du contrat $\overline{X}$. La prime de crédibilité devra donc accorder un poids plus grand à $\overline{X}$, ce qui se traduit par un coefficient de crédibilité plus élevé. ∎

**Remarque 3.2.2.** La prime de crédibilité vu au chapitre précédent peut encore s'écrire sous la forme $P = E[X] + Z(\overline{x} - E[X])$. La meilleure approximation linéaire de la prime de risque est donc également la meilleure approximation linéaire de la prime bayésienne.

**Exemple d'application 3.2.1.** Un assuré a eu au cours des $5$ premières années les montants des sinistres suivants : $2, 1, 0, 3, 2$. Sachant que $X|\Theta \hookrightarrow$ Poisson($\theta$) et $\Theta \hookrightarrow$Gamma$(10, 5)$, déterminons la prime de crédibilité pour la sixième année.

On a $\mu(\theta) = E[X|\theta] = \theta$ et $\sigma^2(\theta) = Var[X|\theta] = \theta$. Par conséquent,

$$\begin{aligned} m &= E[\mu(\Theta)] = E[\Theta] = \frac{10}{5} = 2, \\ s^2 &= E[\Theta] = 2, \\ a &= Var[\mu(\Theta)] = \frac{10}{5^2} = 0,4. \end{aligned}$$

Ainsi

$$\frac{s^2}{a} = 5 \text{ et } z = \frac{5}{5+5} = 0,5.$$

La prime de crédibilité pour la sixième année est donc

$$P = 0,5 \times \left(\frac{2+1+0+3+2}{5}\right) + (1-0,5) \times 2 = 1,8.$$

## 3.3 Modèle classique de Bühlmann

Dans le modèle classique de Bühlmann (1969), on considère un portefeuille composé de $k$ contrats observés pendant $t$ périodes. Contrairement au modèle original, l'approximation linéaire $P_{i,t+1}$ de la prime de risque ne nécessite plus la connaissance des paramètres $m$, $s^2$ et $a$ car il est possible de les estimer à partir des observations des différents contrats. Le calcul de la prime est donc développé sous une approche bayésienne empirique. Comme il a été mentionné au chapitre précédent, on suppose que chaque contrat est caractérisé par un paramètre de risque $\Theta_i = \theta_i$ qui lui est propre. On considère que tous les paramètres de risque sont issus d'une même fonction de structure $U(.)$.

### 3.3.1 Hypothèses du modèle classique de Bühlmann

Soit $X_{ij}$ le montant total des sinistres attribué au $i^{\text{ème}}$ contrat pour la $j^{\text{ème}}$ période. Le modèle classique de Bühlmann$(1969)$ repose sur les hypothèses suivantes :

$(H_1)$ Pour $i = 1, \cdots, k$ les contrats $(\Theta_i, X_i)$ sont indépendants.

$(H_2)$ Pour un paramètre de risque $\Theta_i = \theta_i$ donné, les variables $X_{i1}, X_{i2}, \ldots, X_{it}$ sont conditionnellement indépendantes et identiquement distribuées.

$(H_3)$ $E[X_{ij}|\Theta_i] = \mu(\Theta_i), \qquad i = 1, \ldots, k,\ j = 1, \ldots, t.$

$(H_4)$ $Cov(X_{ij}, X_{ir}|\Theta_i) = \delta_{jr}\sigma(\Theta_i), \ \ i = 1, ..., k,\ j, r = 1, .., t.$

### 3.3.2 Calcul de la prime

Le modèle classique de Bühlmann $(1969)$ permet d'obtenir une approximation de la prime de risque par des fonctions linéaires homogènes et non homogènes des observations. Comme dans le modèle original, cette combinaison linéaire est de la forme

$$c_0^i + \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj}. \tag{3.11}$$

Il faut donc trouver les constantes $c_0^{i*}, c_{11}^{i*}, \ldots, c_{j't}^{i}$ qui minimisent

$$E\left[\left(\mu(\Theta_i) - c_0^i - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj}\right)^2\right]. \tag{3.12}$$

Notons que l'exposant $i$ désigne le contrat pour lequel on cherche à obtenir la prime de crédibilité. Le théorème[4] suivant présente la prime de crédibilité non homogène du modèle classique de Bühlmann $(1969)$.

**Théorème 3.4.** Soit un portefeuille satisfaisant aux hypothèses du modèle classique de Bühlmann. Alors la meilleure approximation linéaire $P_{i,t+1}$, au sens des moindres carrés, de la prime de risque du contrat $i$ pour la période $t+1$ telle que $(3.12)$ soit minimale est de la forme

$$P_{i,t+1} = z\overline{X_i} + (1-z)m. \tag{3.13}$$

où

$$z = \frac{at}{at+s^2}, \tag{3.14}$$

$$\overline{X_i} = \frac{1}{t}\sum_{j=1}^{t} X_{ij}, \tag{3.15}$$

$$m = E[\mu(\Theta_i)]. \tag{3.16}$$

**Preuve**. Pour chacun des $k$ contrats, on doit trouver les coefficients $c_{rj}^{i*}, (i, r = 1, \ldots, k; j = 1, \ldots, t)$ minimisant (3.12). Ces coefficients sont trouvés par la résolution du système obtenu en posant nulles les dérivées partielles par rapport à $c_0^i$ et $c_{rj}^i$, pour $i, r = 1, \ldots, k$ $j = 1, \ldots, t$. En posant nulle la dérivée partielle par rapport à $c_0^i$, on obtient

$$-2E\left[(\mu(\Theta_i) - c_0^i - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj})\right] = 0,$$

et par conséquent,

$$c_0^i = m - m\sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i \tag{3.17}$$

car $E[\mu(\Theta_i)] = E[X_{rj}] = m$. En remplaçant $c_0^i$ par son expression trouvée en (3.17) dans (3.12), les dérivées partielles par rapport à $c_{r'j'}^i, (i, r' = 1, \ldots, k; j' = 1, \ldots, t)$ sont

$$\frac{\partial}{\partial c_{r'j'}^i} E\left[\left(\mu(\Theta_i) - m - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i (X_{rj} - m)\right)^2\right]$$
$$= -2E\left[\left(\mu(\Theta_i) - m - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i (X_{rj} - m)\right)(X_{r'j'} - m)\right].$$

Ce qui conduit au système d'équations suivant :

$$Cov\left[\mu(\Theta_i), X_{r'j'}\right] = \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i Cov\left[X_{r'j'}, X_{rj}\right], \quad r' = 1, \ldots, k; j = 1, \ldots, t. \tag{3.18}$$

Ainsi, pour $r' = i$,

$$a = \sum_{j=1}^{t}(a + \delta_{j'j}s^2)c_{ij}^i. \tag{3.19}$$

Puisque les contrats sont supposés indépendants, pour chaque contrat, on a un système matriciel semblable à celui que l'on retrouve dans le modèle original. Par conséquent, les coefficients $c_{ij}^i$, $j = 1, \ldots, t$ sont égaux. En effectuant la sommation dans (3.19), on obtient

$$c_{ij}^i = \frac{a}{at+s^2}, \quad i = 1, \ldots, k. \tag{3.20}$$

Puisque cette expression de $c_{ij}^i$ ne dépend de $i$, ces coefficients sont tous égaux pour $r' = i$. Lorsque $r' \neq i$, le membre de gauche de l'égalité $(3.18)$ est nul puisque les contrats sont supposés indépendants. Ce système étant symétrique, les coefficients $c_{rj}^i$ sont donc nuls pour $r' \neq i$. Ainsi en remplaçant les expressions obtenues pour $c_0^i$ et $c_{rj}^i$ dans $(3.11)$, on trouve

$$\begin{aligned} c_0^i + \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj} &= m(1 - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i) + \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj} \\ &= m\left(1 - \frac{a}{at+s^2}\right) + \frac{a}{at+s^2}\overline{X_i}. \end{aligned}$$

En posant $z = \dfrac{at}{at+s^2}$, on obtient donc la prime linéaire suivante :

$$P_{i,t+1} = z\overline{X_i} + (1-z)m.$$

■

Le théorème $(3.4)$ présente la prime de crédibilité non homogène. De la même façon, on peut obtenir la prime de crédibilité homogène c'est-à-dire sans terme constant. L'approximation de la prime de risque du contrat $i$ est de la forme suivante :

$$\sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj}.$$

Le théorème suivant présente la prime de crédibilité homogène.

**Théorème 3.5.** Soit un portefeuille satisfaisant aux hypothèses du modèle classique de Bühlmann. Alors la prime linéaire homogène

$$P_{i,t+1}^h = \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj}$$

telle que

$$E\left[(\mu(\Theta) - \sum_{r=1}^{k}\sum_{j=1}^{t} c_{rj}^i X_{rj})^2\right] \tag{3.21}$$

est minimale, est de la forme

$$P_{i,t+1} = z\overline{X_i} + (1-z)\overline{X}, \tag{3.22}$$

où

$$z = \frac{at}{at+s^2}, \tag{3.23}$$

$$\overline{X_i} = \frac{1}{t}\sum_{j=1}^{t} X_{ij}, \tag{3.24}$$

$$\overline{X} = \frac{1}{k}\sum_{i=1}^{k} \overline{X_i}. \tag{3.25}$$

Ce théorème se démontre de la même façon que pour la prime de crédibilité non homogène. c'est-à-dire en posant nulles les dérivées partielles par rapport à chacun des coefficients $c_{rj}^i (i, r = 1, \ldots, k; j = 1, \ldots, t)$. La prime de crédibilité homogène est semblable à la prime de crédibilité non homogène, à l'exception près que le paramètre $m$ de cette dernière est estimé à partir des observations du portefeuille.

### 3.3.3 Estimation des paramètres de structures

Le calcul de la prime selon le modèle original de Bühlmann (1967) nécessite la connaissance des fonctions de vraisemblance et la densité a priori. Habituellement, celles-ci sont inconnues de l'assureur. Comme on l'a mentionné dans la section précédente, le modèle classique de Bühlmann $(1969)$, puisqu'il fait référence à un portefeuille composé de plusieurs contrats, vient contourner le problème de complexité du calcul de la prime. Ainsi, il est possible d'estimer les paramètres de structure $m$, $a$ et $s^2$ à partir des observations du portefeuille. Nous présentons dans cette section certains estimateurs des paramètres impliqués dans le calcul du coefficient de crédibilité $z$.

Le théorème suivant propose des estimateurs optimaux sans biais pour $m$, $s^2$ et $a$ du modèle classique de Bühlmann.

**Théorème 3.6.** Soit un portefeuille d'assurance composé de $k$ contrats et dont l'expérience $X_{ij}$ du contrat $i$ au cours de la période $j$ a été observée pendant $t$ périodes. Les variables aléatoires suivantes constituent des estimateurs sans biais des paramètres $m$, $s^2$ et $a$ du modèle classique de Bühlmann.

$$\hat{m} = \frac{1}{k}\sum_{i=1}^{k}\overline{X_i}, \tag{3.26}$$

$$\hat{s}^2 = \frac{1}{k(t-1)}\sum_{i=1}^{k}\sum_{j=1}^{t}(X_{ij} - \overline{X}_i)^2, \tag{3.27}$$

$$\hat{a} = \frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_i - \overline{X})^2 - \frac{\hat{s}^2}{t}. \tag{3.28}$$

**Preuve**. L'estimateur du paramètre $m$ présenté en $(3.26)$ est sans biais. En effet,

$$\begin{aligned} E\left[\hat{m}\right] &= \frac{1}{k}\sum_{i=1}^{k} E[\overline{X}_i], \\ &= \frac{1}{k}\sum_{i=1}^{k}\frac{1}{t}\sum_{j=1}^{t} E[X_{ij}], \end{aligned}$$

et puisque $E[X_{ij}] = E[\mu(\Theta_i)] = m$, on obtient

$$E\left[\hat{m}\right] = m.$$

Pour un contrat $i$ fixé, un estimateur sans biais de $s^2$ est :

$$\hat{s_i}^2 = \frac{1}{t-1}\sum_{j=1}^{t}(X_{ij} - \overline{X}_i)^2.$$

Pour démontrer l'absence de biais, on note d'abord que

$$\begin{aligned} E\left[(X_{ij} - \overline{X}_i)^2|\Theta_i\right] &= E\left[X_{ij}^2 - 2X_{ij}\overline{X}_i + \overline{X}_i^2|\Theta_i\right] \\ &= [Var[X_{ij}|\Theta_i] + E[X_{ij}|\Theta_i]^2 - 2Cov(X_{ij}, \overline{X}_i|\Theta_i) \\ &\quad -2E[X_{ij}|\Theta_i]E[\overline{X}_i|\Theta_i] + Var[\overline{X}_i|\Theta_i] + E[\overline{X}_i|\Theta_i]^2]. \end{aligned}$$

On rappelle que pour deux variables aléatoires $X$ et $Y$, on a

$$Var[X] = E[X^2] - (E[X])^2 \text{ et } Cov(X,Y) = E[XY] - E[X]E[Y].$$

Par l'hypothèse de covariance conditionnelle du modèle classique de Bühlmann $(1969)$ et puisque $E[X_{ij}|\Theta_i] = E[\overline{X}_i|\Theta_i] = \mu(\Theta_i)$, on obtient

$$\begin{aligned} E[s_i^2] &= \frac{1}{t-1}\sum_{j=1}^{t} E[(X_{ij} - \overline{X}_i)^2|\Theta_i] \\ &= \frac{1}{t-1}\sum_{j=1}^{t}[Var[X_{ij}|\Theta_i] + E[X_{ij}|\Theta_i]^2 - 2Cov(X_{ij}, \overline{X}_i|\Theta_i) \\ &\quad -2E[X_{ij}|\Theta_i]E[\overline{X}_i|\Theta_i] + Var[\overline{X}_i|\Theta_i] + E[\overline{X}_i|\Theta_i]^2] \\ &= \frac{1}{t-1}\sum_{j=1}^{t}[\sigma^2(\Theta_i) + \mu^2(\Theta_i) - 2\frac{1}{t}\sigma^2(\Theta_i) - 2\mu^2(\Theta_i) + \frac{1}{t}\sigma^2(\Theta_i) + \mu^2(\Theta_i)] \\ &= \frac{1}{t-1}\sum_{j=1}^{t}\left(\sigma^2(\Theta_i) - 2\frac{\sigma^2(\Theta_i)}{t} + \frac{\sigma^2(\Theta_i)}{t}\right) \\ &= \frac{1}{t-1}\sum_{j=1}^{t}\frac{t-1}{t}\sigma^2(\Theta_i) \\ &= \sigma^2(\Theta_i). \end{aligned}$$

Ainsi, on peut estimer $s^2$ en tenant compte des observations d'un contrat $i$ $(i = 1, \ldots, k)$ en particulier. On pourrait donc choisir un estimateur $\hat{s}_i^2$ parmi les $k$ estimateurs possibles. Cependant, on opte plutôt pour la moyenne arithmétique des $\hat{s}_i^2$ c'est-à-dire

$$\frac{1}{k(t-1)}\sum_{i=1}^{k}\sum_{j=1}^{t}(X_{ij} - \overline{X}_i)^2.$$

Disposant ainsi d'un volume d'information plus grand relié à tout le portefeuille, le degré de précision de $s^2$ en est augmenté. Cet estimateur est sans biais.

Le paramètre $a$ représentant la variance des moyennes individuelles peut être approximé par l'estimateur suivant :

$$\frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_i-\overline{X})^2.$$

Cet estimateur est biaisé. En effet en ajoutant et en retranchant la moyenne collective $m$, on a

$$\begin{aligned} E\left[\frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_i-\overline{X})^2\right] &= \frac{1}{k-1}\sum_{i=1}^{k}E[((\overline{X}_i-m)-(\overline{X}-m))^2] \\ &= \frac{1}{k-1}\sum_{i=1}^{k}\left[Var[\overline{X}_i]+Var[\overline{X}]-2Cov(\overline{X}_i,\overline{X})\right] \qquad (3.29) \end{aligned}$$

Par le corollaire $(3.1)$, on obtient les expressions suivantes pour $Var[\overline{X}_i]$, $Var[\overline{X}]$ et $Cov(\overline{X}_i,\overline{X})$ :

$$\begin{aligned} Var[\overline{X}_i] &= Var[E[\overline{X}_i|\Theta]]+E[Var[\overline{X}_i|\Theta_i]] \\ &= Var[\mu(\Theta_i)]+\frac{1}{t}E[\sigma^2(\Theta_i)] \\ &= a+\frac{s^2}{t}, \qquad (3.30) \end{aligned}$$

$$\begin{aligned} Var[\overline{X}] &= Var[E[\overline{X}|\Theta]]+E[Var[\overline{X}|\Theta_i]] \\ &= \frac{1}{k}Var[\mu(\Theta_i)]+\frac{1}{kt}E[\sigma^2(\Theta_i)] \\ &= \frac{1}{k}\left(a+\frac{s^2}{t}\right), \qquad (3.31) \end{aligned}$$

$$\begin{aligned} Cov(\overline{X}_i,\overline{X}) &= Cov(E[\overline{X}_i|\Theta_i],E[\overline{X}|\Theta_i])+E[Cov((\overline{X}_i,\overline{X})|\Theta_i] \\ &= \frac{1}{k}Cov(\mu(\Theta_i),\mu(\Theta_i))+\frac{1}{kt}E[\sigma^2(\Theta_i)] \\ &= \frac{1}{k}\left(a+\frac{s^2}{t}\right). \qquad (3.32) \end{aligned}$$

Ainsi en insérant ces expressions dans $(3.29)$, on obtient

$$E\left[\frac{1}{k-1}\sum_{i=1}^{k=1}(\overline{X}_i-\overline{X})^2\right]=a+\frac{s^2}{t}.$$

Par conséquent, cet estimateur est biaisé. Puisque l'on recherche un estimateur sans biais, on lui soustrait le biais $s^2/t$. Le paramètre $s^2$ étant inconnu, on prend son estimateur sans biais $\hat{s}^2$, et l'on obtient

$$\frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_i-\overline{X})^2-\frac{\hat{s}^2}{t}$$

comme estimateur sans biais de $a$. L'estimateur $\hat{a}$ de la variance des primes de risque proposé en $(3.27)$ peut parfois prendre des valeurs négatives. Une alternative proposée par Bühlmann et Straub $(1970)$, consiste à prendre $\hat{a}'=max(0,\hat{a})$ comme estimateur du paramètre $a$. Cependant, ce nouvel estimateur comporte le désavantage de ne plus être sans biais. ■

### 3.3.4 Estimation de la prime de crédibilité

Dans la section précédente, on a présenté des estimateurs optimaux sans biais pour les paramètres de structure $m$, $s^2$ et $a$ du modèle classique de Bühlmann $(1969)$. Généralement, la prime de crédibilité est estimée en remplaçant simplement chaque paramètre inconnu par son estimateur. Ainsi,

$$P_{i,t+1} = \hat{z}\overline{X}_i + (1-\hat{z})\hat{m}$$

avec

$$\hat{z} = \frac{t}{t+\hat{s}^2/\hat{a}}.$$

Bien que tous les estimateurs soient sans biais, on ne peut conclure que $E[\hat{z}] = z$. Par conséquent, l'estimateur de la prime de crédibilité est fort probablement biaisé.

**Exemple d'application 3.3.1.** Soit le portefeuille de police d'assurance constitué de $4$ contrats après $9$ années d'expérience.

| | **Années** | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Contrats | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 1 | 0 | 3 | 2 | 4 | 3 | 3 | 0 |
| 2 | 0 | 1 | 0 | 0 | 2 | 1 | 1 | 2 | 2 |
| 3 | 3 | 3 | 4 | 2 | 2 | 3 | 4 | 2 | 4 |

Calculons la prime de crédibilité pour la $10^{\text{ème}}$ année pour chacun des contrats.

Tout d'abord on a

$$\overline{X}_1 = 2, \qquad \overline{X}_2 = 1, \qquad \overline{X}_3 = 3.$$

Calculons les estimateurs des paramètres de structures.

$$\hat{m} = \overline{X} = \frac{2+3+1+}{3} = 2$$

$$\begin{aligned}
\hat{s}^2 &= \frac{1}{k(t-1)}\sum_{i=1}^{k}\sum_{j=1}^{t}(X_{ij}-\overline{X}_i)^2 \\
&= \frac{1}{3\times 8}(16+6+7) = \frac{29}{24}
\end{aligned}$$

$$\begin{aligned}
\hat{a} &= \frac{1}{k-1}\sum_{i=1}^{k}(\overline{X}_i-\overline{X})^2 - \frac{\hat{s}^2}{t} \\
&= \frac{1}{2}(0+1+1) - \frac{29}{24}.\frac{1}{9} \\
&= \frac{187}{216}
\end{aligned}$$

Par conséquent, $\hat{s}^2/\hat{a} = 1,40$ et

$$\hat{z} = \frac{9}{9+1,40} = 0,87,$$

d'où

$$\begin{aligned} P_{1,10} &= 2 \times 0,87 + (1-0,87) \times 2 = 2 \\ P_{2,10} &= 0,87 \times 1 + 0,26 = 1,13 \\ P_{3,10} &= 0,87 \times 3 + 0,26 = 2,87. \end{aligned}$$

**Exemple d'application 3.3.2.** L'information sur le montant sinistre dans une année d'un portefeuille de polices, divisée en 7 classes et pendant les dernières cinq années est la suivante :

| | 1 | | 2 | | 3 | | 4 | | 5 | | 6 | | 7 | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $P$ | $X$ | $P$ | $X$ | $P$ | $X$ | $P$ | $X$ | $P$ | $X$ | $P$ | $X$ | $P$ | $X$ |
| 1 | 5 | 0 | 14 | 11,3 | 18 | 8 | 20 | 5,4 | 21 | 9,7 | 43 | 9,7 | 70 | 9 |
| 2 | 6 | 0 | 14 | 25 | 20 | 1,9 | 22 | 5,9 | 24 | 8,9 | 47 | 14,5 | 77 | 9,6 |
| 3 | 8 | 4,2 | 13 | 18,5 | 23 | 7 | 25 | 7,1 | 28 | 6,7 | 53 | 10,8 | 85 | 8,7 |
| 4 | 10 | 0 | 11 | 14,3 | 25 | 3,1 | 29 | 7,2 | 34 | 10,3 | 61 | 12 | 92 | 11,7 |
| 5 | 12 | 7,7 | 10 | 30 | 27 | 5,2 | 35 | 8,3 | 42 | 11,1 | 70 | 13,1 | 100 | 7 |

où $P$ est volume de primes et $X$ la charge sinistres, à travers un taux pour cent des capitaux en risque.
Déterminons la valeur de $X$ pour la sixième année, pour chaque groupe.
Les calculs dont le script est donné en annexe ont été effectués avec le logiciel R.
Estimateurs des paramètres de structure.
$\hat{m} = 9,225714$
$\hat{s^2} = 12,587$
$\hat{a} = 29,22030$
$\hat{z} = 0,920681$
les valeurs estimées de la prime pour la sixième année pour chaque classe sont regroupées dans le tableau ci- après.

| Classe(i) | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| Prime($P_{i,6}$) | 2,922995 | 18,979673 | 5,372006 | 6,973991 | 9,330935 | 11,798360 | 9,202040 |

On constate que pour la sixième année, la classe 2 est celle qui a la charge totale sinistres la plus élevée et la classe 1 la charge totale sinistres la plus basse. Ce constat est tout à fait normal dans la mesure où, durant les cinq dernières années, la classe 2 est celle qui paye en moyenne plus que toutes les autres et la classe 1 moins que toutes les autres.

# ♦ Implication Pédagogigue ♦

Ce mémoire est rédigé dans le but d'obtenir un diplôme de professeur de l'enseignement secondaire. Il parait donc normal de terminer ce travail en mentionnant ce dont il nous apporte en matière d'enseignement.

Les outils mathématiques utilisés en théorie de la crédibilité pour exprimer les primes des assurés sont essentiellement les statistiques et les probabilités (beaucoup plus les probabilités conditionnelles). Dans les programmes de mathématiques de l'enseignement secondaire au Cameroun, les statistiques sont enseignées de la classe de $5^{\text{ième}}$ à la classe de Terminale et les probabilités quant à elles ne sont enseignées qu'en classe de Terminale. Travailler sur la théorie de la crédibilité nous a permis d'élargir les champs d'applications de ces deux notions mathématiques dans le domaine des assurances que nous avons eu l'occasion de découvrir. Cette découverte nous incite à :

- chercher d'autres domaines de la vie courante où l'on utilise les statistiques ou les probabilités afin de diversifier les exercices que l'on pourra donnés aux élèves.
- Proposer des sujets concrets afin que ces derniers comprennent que les statistiques ne se limitent pas au calcul de la moyenne et de la variance et les probabilités aux jeux de cartes mais bien plus encore.

Dans le cadre de ce travail, nous avons fait la connaissance du logiciel R qui facilite les calculs statistiques et probabilistes.

Tout en sachant que l'analphabète d'aujourd'hui n'est plus seulement celui qui ne sait pas lire, mais aussi celui qui ne sait pas se servir de l'outil informatique, nous trouvons intéressant de faire apprendre de tels logiciels aux élèves. Cependant le logiciel R étant d'un niveau un peu élevé, nous nous proposons de montrer aux élèves comment utiliser les logiciels Sine qua non, geogebra qui vont bien au delà de la statistique et de la probabilité,

# ♦ Conclusion générale et perspectives ♦

Ce travail de recherche portait sur les modèles de crédibilité utilisés en assurance pour établir la prime à payer par un assuré. Il s'est agit d'établir un instrument juste et objectif pouvant rendre compte de manière équitable du calcul de la prime des assurés. Pour ce faire, nous avons évoqué tour à tour la crédibilité de stabilité, le modèle de la crédibilité bayésienne et le modèle de Bühlmann.

En effet, la crédibilité de stabilité qui a fait l'objet de notre premier chapitre a permis de poser les bases de la théorie de la crédibilité en montrant comment les seuils de la crédibilité complète et le facteur de crédibilité interviennent dans le calcul de la prime pour les différentes mesures de sinistres. Toutefois, cette approche ne présente aucune hypothèse en ce qui concerne les paramètres intervenant dans le calcul de la prime. Ce défaut majeur a généré la crédibilité bayésienne qui considère que chaque contrat dépend d'un paramètre de risque dont on a une idée de la loi a priori et le calcul de la prime dépendant de la loi a posteriori de ce paramètre. L'on constate par là que lorsque les distributions sont conjuguées la loi a posteriori se trouve aisément car elle est la même que la loi a priori mais avec des paramètres différents. Cependant le problème de complexité de calcul de la prime bayésienne se pose lorsque les distributions ne sont pas conjuguées. Bühlmann par ses deux modèles (original et classique) propose comme solution à ce problème de formuler la prime comme fonction linéaire des observations du portefeuille. Il obtient par là une expression de la prime représentant un compromis entre l'expérience individuelle et l'expérience de l'ensemble du portefeuille.

Partant de ces modèles de crédibilité ci-dessus présentés, nous avons tendance de prime à bord de conseiller à un actuaire d'établir les primes d'assurance selon le modèle de Bühlmann. Toutefois, ce modèle n'est pas toujours applicable à toutes les situations de part son manque de représentativité de la réalité. En accidents de travail par exemple, l'exposition au risque d'un employeur ayant $5000$ employés est beaucoup plus grande que celle d'un employeur avec seulement $50$ employés. Mais, Bühlmann attribue un même poids à ces deux employeurs dans le calcul de leur prime, chose pas du tout acceptable. Le modèle de Bühlmann-Straub (1970) vient palier à ce problème en introduisant une pondération à chaque contrat. On obtient ainsi un facteur de crédibilité différent pour chaque police. D'autres modèles ont vu le jour à la suite du modèle de Bühlmann- Straub il s'agit des modèles de

Jewell (1975), Taylor (1979), Danneburg (1996). Une suite à nos travaux serait de voir comment la prime est formulée dans ces autres modèles et la véritable différence existant avec le modèle de Bühlmann.

# ♦ Bibliographie ♦


[1] **Arthur H MOWBRAY (1914)**, How extensive a payroll is necessary to give a dependable pure premium ? *Proceedings of the Casuality Actuarial Society*, vol 1 : 24-30.

[2] **BÜHLMANN (1967)**, Expérience rating and credibility. *ASTIN Bulletin*, vol 4 : 99-207.

[3] **BÜHLMANN (1969)**, Experience rating and credibility. *ASTIN Bulletin*, vol 5 : 157-165.

[4] **Geneviève GARON (1999)**, Théorie de la crédibilité : MODÈLES de BÜHLMANN ET BÜHLMANN-STRAUB ET APPLICATIONS. *Mémoire d'actuariat, Université LAVAL, QUÉBEC*.

[5] **Hélène COSSETTE, Vincent GOULET (2010)**, Exercices en théorie de la crédibilité. *San Francisco, Creative Commons*, 103 pages.

[6] **Pierre E.THEROND (2005)**, Théorie de la crédibilité. *Notes de cours de l'ISFA, Université Lyon* 1.

[7] **Pierre E.THEROND (2009)**, Théorie de la crédibilité. *Notes de cours de l'ISFA, Université Lyon* 1.

[8] **Pierre E.THEROND (2014)**, Théorie de la crédibilité. *Notes de cours de l'ISFA, Université Lyon* 1.

[9] **Vincent GOULET (2010)**, Théorie de la crédibilité. *San Francisco, Creative Commons*, 73 pages.

[10] **Siméon FOTSO (2014)**, Statistique bayésienne et Applications en assurance. *Quatrième Ecole Mathématique Africaine, Université de Douala, Faculté des Sciences, Cameroun*.

[11] **Whitney A. (1918)**, The theory of Experience rating. *Proceedings of the Casualty Actuarial Society*, vol 4 : 275-293.

[12] **YIU-KUEN TSE (2009)**, Nonlife Actuarial Models Theory, Methods and evaluation. *Cambridge, Cambrige University Press*, 542 pages.

# ♦ Annexe ♦

```
########## code credibilité Bühlmann ###############
library(actuar)
state <-  c(1,2,3,4,5,6,7)
 ratio.1 <- c(0,11.3,8,5.4,9.7,9.7,9)
 ratio.2 <- c(0,25,1.9,5.9,8.9,14.5,9.6)
 ratio.3 <- c(4.2,18.5,7,7.1,6.7,10.8,8.7)
 ratio.4 <- c(0,14.3,3.1,7.2,10.3,12,11.7)
 ratio.5 <- c(7.7,30,5.2,8.3,11.1,13.1,7)
 weight.1 <- c(5,14,18,20,21,43,70)
 weight.2 <- c(6,14,20,22,24,47,77)
 weight.3 <- c(8,13,23,25,28,53,85)
 weight.4 <- c(10,11,25,29,34,61,92)
 weight.5 <- c(12,10,27,35,42,70,100)
 dataBS <- cbind(state,ratio.1,ratio.2,ratio.3,ratio.4,ratio.5,
                  + weight.1,weight.2,weight.3,weight.4,weight.5); dataBS
state  ratio.1  ratio.2  ratio.3  ratio.4  ratio.5  weight.1  weight.2  weight.3  weight.4
weight.5
B <- cm(~state, dataBS, ratios = ratio.1:ratio.5)
summary(B)
```

FIGURE 3.1 – Code R pour le calcul de la prime selon le modèle de Bühlmann.

```
################# courbe de la densité de la loi gamma(2,3)##################
x<-seq(0,4,by=0.001)
y<-dgamma(x,2,3)
leg.txt<-c("Gamma(2,3)")
plot(x,y,main="Apres 1 année",type = "l",xlab=" ", yylab = " ", col="black")
legend("topright",leg=leg.txt,col="black",lty=1)
```

FIGURE 3.2 – Code R pour le tracé de la courbe de densité de la fonction gamma (2, 3).