Théorie de la crédibilité avec R

Vincent Goulet

UNIVERSITÉ LAVAL

Théorie de la crédibilité avec R

Vincent Goulet

Professeur titulaire
École d'actuariat, Université Laval

Édition 2020.05

Code source
Voir sur GitLab

Couverture
Bouquetin des Alpes (*Capra ibex*) du parc national autrichien du Hohe Tauern avec, en arrière-plan, le glacier Pasterze, le plus grand d'Autriche. À la tombée du jour, lorsque les randonneurs ont quitté la région, ces farouches chèvres de montagne s'approchent parfois plus qu'à l'habitude des sentiers.

Crédit photo : © Bernd Thaller, CC BY-SA 3.0 Autriche, via Wikimedia Commons.

Introduction

« Les mathématiques de l'hétérogénéité. » C'est parfois ainsi que l'on décrit la théorie de la crédibilité, pierre angulaire des mathématiques de l'assurance IARD. Cet ouvrage offre un traitement rigoureux et exhaustif des modèles de base de la crédibilité, soit la crédibilité de stabilité (*limited fluctuations*), la tarification basée sur l'expérience purement bayésienne et les modèles classiques de Bühlmann et de Bühlmann–Straub.

Le paquetage **actuar** (Dutang et collab., 2008) pour l'environnement statistique R (R Core Team, 2019) permet d'effectuer les calculs relatifs aux modèles de crédibilité abordés dans l'ouvrage. Nous expliquons comment utiliser la fonction `cm` du paquetage par le biais de code informatique distribué avec le document et reproduit à la fin des chapitres 3, 4 et 5.

L'ouvrage intègre également le recueil d'exercices et de solutions de Cossette et Goulet (2008). Tel que mentionné en introduction de ce document, la collection d'exercices est le fruit de la mise en commun d'exercices colligés au fil du temps pour des cours de théorie de la crédibilité à l'Université Laval et à l'Université Concordia. Certains exercices ont été rédigés par les Professeurs François Dufresne et Jacques Rioux, entre autres. Quelques exercices proviennent également d'anciens examens de la Society of Actuaries et de la Casualty Actuarial Society.

Le premier chapitre, tiré de Goulet (1994), trace l'historique et l'évolution de la théorie de la crédibilité, de ses origines jusqu'au début des années 1990. Ce chapitre ne comporte pas d'exercices.

L'annexe A offre un sommaire des principaux principes et résultats en estimation bayésienne. Un tableau synoptique des principaux résultats de crédibilité exacte se trouve à l'annexe B. L'annexe C présente la paramétrisation des lois de probabilité utilisée dans les exercices. En cas de doute, le lecteur est invité à la consulter. Il y trouvera également l'espérance, la variance et la fonction génératrice des moments (lorsqu'elle existe) des lois de probabilité rencontrées dans ce document.

Utilisation de l'ouvrage

L'étude de l'ouvrage implique des allers-retours entre le texte et le code R à la fin de chaque chapitre. Ce code informatique et les commentaires qui l'accompagnent vise à enrichir vos apprentissages. Assurez-vous donc de le lire attentivement, de l'exécuter pas-à-pas et de bien comprendre ses effets.

Le code informatique est distribué avec l'ouvrage sous forme de fichiers de script. De plus, à chaque fichier `.R` correspond un fichier `.Rout` contenant les résultats de son évaluation non interactive.

Fonctionnalités interactives

En consultation électronique, ce document se trouve enrichi de plusieurs fonctionnalités interactives.

- Intraliens du texte vers une ligne précise d'une section de code informatique et, en sens inverse, du numéro de la ligne vers le point de la référence dans le texte. Ces intraliens sont marqués par la couleur ▬.
- Intraliens entre le numéro d'un exercice et sa solution, et vice versa. Ces intraliens sont aussi marqués par la couleur ▬.
- Intraliens entre les citations dans le texte et leur entrée dans la bibliographie. Ces intraliens sont marqués par la couleur ▬.
- Hyperliens vers des ressources externes marqués par le symbole ↗ et la couleur ▬.
- Table des matières, liste des tableaux et liste des figures permettant d'accéder rapidement à des ressources du document.

Blocs signalétiques

Le document est parsemé de divers types de blocs signalétiques inspirés de AsciiDoc ↗ qui visent à attirer votre attention sur une notion. Vous pourrez rencontrer l'un ou l'autre des blocs suivants.

Astuce ! Ces blocs contiennent un truc, une astuce, ou tout autre type d'information utile.

Avertissement ! Ces blocs mettent l'accent sur une notion ou fournissent une information importante pour la suite.

Important ! Ces blocs contiennent les remarques les plus importantes. Veillez à en tenir compte.

Ces blocs contiennent des liens vers des vidéos dans la chaine YouTube liée à ce document. Les vidéos sont répertoriées dans la liste des vidéos.

Ces blocs vous invitent à interrompre la lecture du texte pour passer à l'étude du code R des sections d'exemples.

Document libre

Tout comme R et l'ensemble des outils présentés dans ce document, le projet « Théorie de la crédibilité avec R » s'inscrit dans le mouvement de l'informatique libre. Vous pouvez accéder à l'ensemble du code source en format LaTeX en suivant le lien dans la page de copyright. Vous trouverez dans le fichier `README.md` toutes les informations utiles pour composer le document.

Votre contribution à l'amélioration du document est également la bienvenue ; consultez le fichier `CONTRIBUTING.md` fourni avec ce document et voyez votre nom ajouté au fichier `COLLABORATEURS`.

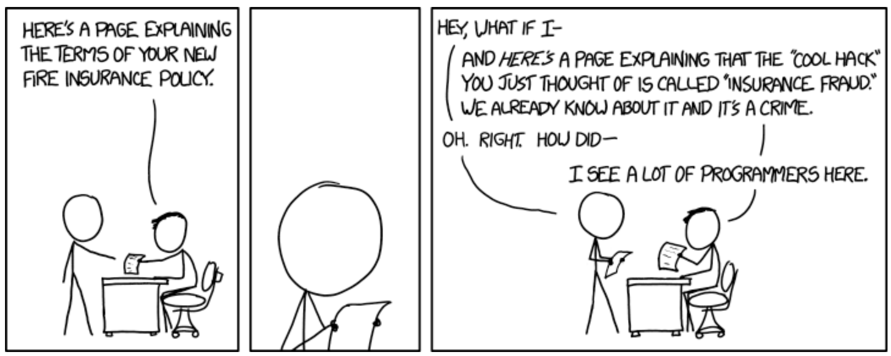
Table des matières

Liste des tableaux

Liste des figures

Liste des vidéos

Le numéro indiqué à gauche est celui de la section dans laquelle se trouve le bloc signalétique.



Tiré de XKCD.com

1 Introduction et perspective historique

Objectifs du chapitre

- Justifier l'utilisation d'un système de tarification basée sur l'expérience dans un contexte d'assurance IARD.
- Expliquer le rôle des modèles de crédibilité dans la tarification basée sur l'expérience.
- Établir le lien entre l'approche subjective des probabilités et la théorie de la crédibilité.
- Expliquer les différences entre la crédibilité de stabilité et la crédibilité de précision.

Le premier but d'un assureur — qu'il s'agisse d'un organisme gouvernemental ou d'une compagnie privée — au moment d'établir une structure de tarification consiste à charger un montant suffisant de primes pour payer les sinistres futurs. Pour différentes raisons, dont la compétitivité, l'assureur pourra par la suite chercher à répartir de manière optimale ces primes entre les assurés. L'une des premières techniques employée pour répartir les primes est la classification des assurés en fonction de différentes caractéristiques comme l'âge, le sexe, le type de véhicule, etc. Néanmoins, il subsistera dans tout regroupement d'assurés une certaine forme d'hétérogénéité, de telle sorte qu'il faudra encore affiner la redistribution des primes à l'intérieur du groupe. C'est là qu'entrent en jeu les systèmes de tarification basée sur l'expérience (*experience rating*) et, plus particulièrement, les méthodes de crédibilité.

TAB. 1.1 – Expérience du portefeuille de l'exemple 1.1 après une année

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

1.1 Tarification basée sur l'expérience

Comme le nom le laisse entendre, la tarification basée sur l'expérience a recours aux données de sinistres d'un assuré pour déterminer sa prime d'assurance. C'est une technique de tarification qui exige un volume d'expérience important. Elle est donc principalement utilisée en assurance automobile et en accidents du travail. En revanche, la tarification basée sur l'expérience n'est pas applicable en assurance-vie (on ne meurt qu'une fois !) ou encore en assurance habitation, où la fréquence des sinistres est trop faible.

De manière plus formelle, (Bühlmann, 1969) propose la définition suivante.

Définition 1.1. La tarification basée sur l'expérience vise à assigner à chaque risque sa prime juste et équitable. Cette prime pour une période dépend exclusivement de la distribution des sinistres (inconnue) de ce risque pour la période.

L'exemple suivant pose le contexte et le problème de base de la tarification basée sur l'expérience.

Exemple 1.1. Un portefeuille d'assurance est composé de dix contrats. Chaque contrat ne peut subir qu'un seul sinistre par année et le montant de ce sinistre est 1. Avant de recueillir des données d'expérience, l'assureur considère que, du point de vue de leur niveau de risque, tous les contrats sont équivalents. L'assureur prévoit devoir payer deux sinistres au cours de la première année. Il charge donc à tous les assurés un montant identique de 0,20. C'est la *prime collective*.

L'expérience au sein du portefeuille après une année se trouve au tableau 1.1. Nous constatons que le montant de sinistre moyen par contrat s'élève à $1/10 = 0{,}10$. La prime collective s'avère donc peut-être trop élevée. Cependant, la quantité de données demeure plutôt faible pour tirer une conclusion, tant sur le niveau de risque global du portefeuille que sur celui de chacun des contrats.

Tab. 1.2 – Expérience du portefeuille de l'exemple 1.1 après deux années

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

Le tableau 1.2 contient l'expérience du portefeuille après deux années. Le montant de sinistre moyen par contrat est maintenant de $4/20 = 0{,}20$. La prime collective semble donc globalement adéquate. L'assureur dispose toujours de peu de données pour juger du niveau de risque des différents contrats, mais nous remarquons que le contrat 9 a déjà subi deux sinistres.

Effectuons maintenant un bond dans le temps pour examiner la situation après dix années, qui est reproduite au tableau 1.3. Le montant de sinistre moyen par contrat de $23/100 = 0{,}23$. En revanche, ces sinistres ne sont pas uniformément répartis dans le portefeuille : le contrat 9 en a subi sept et les contrats 7, 8 et 10, aucun.

Si la prime collective de ce portefeuille s'avère globalement *adéquate*, elle n'est, en revanche, clairement pas *équitable*. Afin d'éviter l'anti-sélection, l'assureur devrait charger plus que la prime collective à certains assurés et moins à d'autres. L'expérience de sinistres a démontré que, contrairement à l'hypothèse de départ de l'assureur, le porfeuille est *hétérogène*. L'assureur devra donc avoir recours à une technique de tarification basée sur l'expérience pour adéquatement répartir les primes entre les assurés. □

Il existe de nombreux systèmes de tarification basée sur l'expérience, dont les systèmes bonus-malus, les polices avec participation ou les commissions en réassurance (Neuhaus, 2004 ; Bühlmann, 1967, 1969), mais les techniques les plus couramment employées demeurent les modèles de théorie de la crédibilité. La théorie de la crédibilité se divise en deux grandes branches.

1. Crédibilité de stabilité, ou américaine, *limited fluctuations*. Dans cette approche, l'assureur tient compte de l'expérience individuelle d'un contrat seulement si celle-ci est suffisamment stable dans le temps.
2. Crédibilité de précision, ou européenne, *greatest accuracy*. Ici, l'assureur tient compte de l'expérience d'un contrat de façon à obtenir la meilleure estimation de l'expérience future. Le poids de l'expérience individuelle augmente avec l'hétérogénéité du portefeuille.

TAB. 1.3 – Expérience du portefeuille de l'exemple 1.1 après dix années

	Contrat									
Année	1	2	3	4	5	6	7	8	9	10
1									1	
2	1	1	1						1	
3	1								1	
4			1						1	
5									1	
6		1								
7	1	1		1	1					
8	1			1		1			1	
9	1				1					
10	1								1	
$\bar{S}_i$	0,6	0,3	0,2	0,2	0,2	0,1	0	0	0,7	0
$\bar{S}$	0,23									

Nous pourrions tout à fait formuler l'exemple 1.1 dans un contexte tout autre que l'assurance. En effet, le tableau 1.3 établit simplement si un « évènement » est survenu ou non en dix « expériences » et pour dix « sujets ». Le but de l'exercice consiste à estimer la probabilité qu'un évènement survienne lors d'une onzième expérience en sachant que les sujets ne sont pas « équivalents ». Cela démontre que nous pouvons utiliser les modèles de crédibilité dans tout contexte de données hétérogènes.

1.2 Historique et évolution de la théorie de la crédibilité

Cette section reprend une partie du chapitre 1 de Goulet (1994).

La personne qui mentionne qu'elle étudie « la crédibilité » obtient habituellement de grands yeux incrédules et interrogateurs comme seule réponse de la part de son interlocuteur, fut-il même parfois actuaire. Pour l'actuaire, la crédibilité n'est souvent qu'un vague concept parmi tant d'autres rencontrés au fil des examens professionnels, concept qui eut tôt fait de fuir à toutes jambes sa mémoire une fois l'examen réussi. Quant au non actuaire, l'étude de la crédibilité ne lui apparaît d'aucun intérêt puisqu'il n'applique habituellement la notion qu'aux personnes physiques : le juge est crédible lorsqu'il interprète la loi, certaines critiques de cinéma sont crédibles alors que d'autres le sont moins, nul politicien n'est considéré crédible… Pourtant, la définition que donne le *Petit Robert* du mot crédibilité, « ce qui fait qu'une chose mérite d'être crue ; caractère de ce qui est croyable », est assez vaste pour laisser place à d'autres interprétations.

La population accorde donc de la crédibilité aux gens, et ce, selon leur notoriété, leurs réalisations, du fait d'expériences concluantes, de leur honnêteté et leur droiture reconnues ou tout simplement par affinité. En fin de compte, une personne est crédible à nos yeux s'il est *probable* que ce qu'elle dit ou fait se réalise ou soit un succès. Par exemple, il est fort probable que l'interprétation que le juge fait de la loi soit la bonne ; ou j'irai voir le nouveau film à l'affiche louangé par tel critique, car je sais qu'il est probable que je l'aime moi aussi.

On voit donc qu'il existe une étroite relation entre les notions de « probabilité » et de « crédibilité ». C'est pour cette raison qu'avant d'aborder l'historique de la théorie actuarielle de la crédibilité, il s'avérera intéressant d'étudier l'évolution de l'approche probabiliste privilégiée en actuariat, l'approche bayésienne. Avec en plus l'étude des réflexions sur le sujet du philosophe anglais Bertrand Russell, le lien entre les deux notions se clarifiera et la perspective historique n'en sera que meilleure.

1.2.1 L'émergence de l'approche bayésienne

Bien que Thomas Bayes eut présenté son fameux théorème dans un essai en 1763, ce n'est que dans la seconde moitié des années 1950 que la philosophie probabiliste basée sur ce théorème gagne suffisamment d'adeptes, et donc de popularité, pour aspirer au respect. L'approche classique monopolisait auparavant la foi des probabilistes et statisticiens, et c'est encore aujourd'hui la philosophie enseignée en premier lieu dans les écoles.

Le statisticien classique [1] a une vision très stricte, objective et mathématique de la probabilité et de l'usage qu'il est possible d'en faire pour inférer un résultat à partir de données. Pour lui, une probabilité est essentiellement la limite d'une série de fréquences relatives où la symétrie joue un rôle primordial : si un événement peut connaître n réalisations différentes mutuel-

1. Souvent appelé *objectivist* en anglais. L'expression est toutefois difficile à traduire adéquatement en français et c'est pourquoi on utilisera le terme « classique »

lement exclusives et ayant une chance égale de se réaliser, et si de ces réalisations m ont l'attribut A, alors la probabilité de A est m/n. Des énoncés du type « il est probable que le Québec devienne indépendant au cours des deux prochaines années » ou « il est probable qu'il neige demain » n'ont donc aucun sens à ses yeux. D'ailleurs, les statisticiens classiques croient que leur travail se limite exclusivement à l'analyse des données et qu'il ne leur appartient pas de prendre des décisions à la lumière des résultats obtenus. Il sera expliqué plus loin que, par opposition, les bayésiens voient en la prise de décision le but de tout travail statistique.

Un autre élément majeur qui caractérise l'approche classique — et c'est celui qui concerne particulièrement la théorie de la crédibilité — est l'absence de prise en compte de toute information à priori relative au problème étudié. On préfère « laisser les données parler d'elles-mêmes ». C'est ce qui fut à l'origine de la contestation devant mener à l'émergence de l'approche bayésienne.

Leonard J. Savage publie en 1954 (Savage, 1954) un des premiers livres consacrés à l'étude de la statistique dans une optique bayésienne. Il défend alors la vision individualiste de la probabilité (*personalistic view of probability*) qu'il qualifie de « seul concept de probabilité ». Comme son nom l'indique, cette vision tâche de se rapprocher le plus possible du raisonnement humain. La probabilité devient donc un indicateur de l'opinion d'une personne à propos d'un événement et, puisqu'une opinion est — habituellement ! — sujette à changement suite à l'ajout d'informations, l'inférence statistique à partir de nouvelles données constitue alors le mécanisme de révision de l'opinion. De plus, l'adhérent à cette vision refuse obstinément de rejeter toute information parallèle ou à priori au problème sous prétexte que, justement, l'humain tire profit de ces informations lorsqu'il prend une décision. Si la question « quel degré de conviction puis-je atteindre à la suite de l'ajout de cette information ? » est hors du champ de la statistique pour le statisticien classique, elle est au contraire au cœur même du problème pour l'individualiste. Ce dernier répond d'ailleurs à la question en utilisant le théorème de Bayes — d'où le qualificatif de bayésien. Cet algorithme donne la nouvelle probabilité d'un événement A à partir de sa probabilité originale et des nouveaux faits marquants venant s'y rattacher. Ainsi, si B constitue l'ensemble des nouveaux faits, alors

$$P[A|B] = \frac{P[B|A]P[A]}{P[B]},$$

où $P[A]$ est la probabilité à priori que l'événement A se réalise.

De plus, les bayésiens évitent l'apparente contradiction entre l'objectivité scientifique et l'irrationalité humaine en postulant un individu idéal, conséquent dans ses décisions. Par exemple, cet individu, lorsque confronté à un choix, choisira toujours l'option la plus probable.

Savage donne dans son ouvrage *The foundations of statistics* quelques expressions synonymes pour *personal probability* : *subjective probability*, *psychological probability* et, finalement, *degree of conviction*. Le lecteur est invité à conserver tout particulièrement en mémoire la dernière expression. En effet, la section suivante étudie sommairement les travaux du philosophe Bertrand Russell et il sera intéressant de constater à quel point les travaux de deux disciplines différentes ont su converger.

Les réflexions de Russell

Lord Bertrand Russell (1872-1970) est un éminent philosophe et logicien britannique, pacifiste invétéré et dont l'activité la plus marquante se rapporta aux mathématiques et à la logique. Il est entre autres le fondateur du logicisme, doctrine selon laquelle « les mathématiques seraient soumises à la formalisation de la logique et s'y réduiraient » (*Le Petit Larousse*). Il fut récipiendaire du Prix Nobel de littérature en 1950.

Joseph Butler (1692–1752) disait : « la probabilité est le guide de la vie d'une personne ». En effet, il est raisonnable de croire que lorsque deux événements ou plus ont des chances de se réaliser, la décision d'une personne sera basée sur celui qui se réalisera avec la plus forte probabilité. Mais ce type de probabilité est-il le même que dans l'énoncé : « la probabilité d'obtenir un double six au lancer de deux dés est de un sur trente-six » ? Russell soutient que non. Voici son raisonnement, que l'on peut retrouver de façon plus exhaustive dans son ouvrage de 1948 (Russell, 1948).

De toute évidence, cette dernière est la probabilité dite *mathématique*, obéissant aux axiomes de la théorie des probabilités, que l'on peut grossièrement réduire au quotient de deux nombres : le cardinal d'une classe spécifique (double six) et celui d'une classe fondamentale (les réalisations possibles du lancer de deux dés). Elle s'attarde aux énoncés d'ordre général, se référant à des classes « anonymes ». Un autre exemple serait, en assurance-vie : « la probabilité qu'un homme non-fumeur âgé de 30 ans décède dans l'année qui vient est, selon la table CSO 1958, de 0,00213 ». Jamais dans ces deux exemples il n'est fait référence à une personne ou un objet en particulier.

Cependant, en passant à un cas particulier du genre « la probabilité que je vive jusqu'à 90 ans est de 60 % », est-on toujours dans le domaine de la probabilité mathématique ? Au passage d'un cas général à un cas particulier (et en ne considérant pas le cas particulier comme une simple réalisation du phénomène général), une personne devrait être en mesure d'intégrer *tous* les éléments d'importance à sa prise de décision. Cette masse d'information lui permettrait alors de savoir avec certitude si oui ou non sa longévité atteindra les 90 ans. Et dès lors elle se retrouverait hors du champ de la probabilité.

Bien entendu, il est pour ainsi dire impossible d'atteindre un tel niveau de certitude. C'est pourquoi la probabilité demeure le guide de la vie de la personne, parce que *son savoir s'avère insuffisant*. Mais encore, si ce guide est la probabilité mathématique, il doit être possible de calculer et d'identifier *la* probabilité (c'est-à-dire *la* classe fondamentale), sinon le guide fait faux bon à son utilisateur ! La personne en quête de cette probabilité évaluera donc d'abord elle-même ses chances de vivre jusqu'à 90 ans, puis ira probablement consulter un médecin, ou même une voyante, qui chacun lui donnera son avis, pas nécessairement objectif. Cette personne ne pourra que croire en totalité ou en partie les divers jugements colligés, aussi lorsqu'elle affirmera « la probabilité que je vive jusqu'à 90 ans est de 60 % », elle attribuera en réalité une *crédibilité* de 60 % à l'expression « je vais vivre jusqu'à 90 ans ». C'est le mieux qu'elle pourra faire, tout le savoir de l'humain comportant une part de doute ou, à l'inverse, seulement un *degré de crédibilité*.

Crédibilité, voilà donc le nom qu'attribue Russell à ce second type de probabilité, la distinction entre la probabilité mathématique et la crédibilité se faisait au *passage du général au particulier*. Une relation subsiste néanmoins entre les deux notions, et c'est que la probabilité mathématique est un instrument de mesure de la crédibilité. De plus, toujours selon Russell, tout énoncé comporte un degré de crédibilité intrinsèque, si minime soit-il, car une personne a toujours un certain bagage de connaissances lui permettant de se faire une opinion à priori. Finalement, la crédibilité peut être augmentée par l'ajout d'information, le gain marginal allant en décroissant.

Cette théorie de Bertrand Russell date de 1948 (ou du moins sa publication). L'énoncé sommaire ci-dessus suffit pour constater à quel point cette théorie est près de l'approche bayésienne en statistique qui commençait alors à sortir de l'anonymat. Cela n'est pas surprenant si l'on se rappelle que les fondateurs de l'école bayésienne ont, comme des philosophes le font, basé leur théorie sur le comportement humain. Ainsi, ce que Russell nomme probabilité mathématique correspond en fait à l'approche classique en statistique et il ne la considère pas comme une fin en soi, mais bien comme un outil pour une prise de décision. C'est cette prise de décision, via la détermination d'un degré de crédibilité, qui constitue la fin de l'utilisation de la probabilité (ou de la statistique). D'ailleurs, l'expression *degree of conviction* proposée par Savage n'est-elle pas synonyme de celle de Russell, degré de crédibilité ?

Cependant, là où le philosophe fortifie le plus la thèse bayésienne, c'est lorsqu'il soutient que tout énoncé comporte un degré de crédibilité intrinsèque, car, ce faisant, il défend l'apport non négligeable de l'information à priori.

1.2.2 Et les actuaires dans tout ça ?

Longtemps avant l'apparition de l'approche bayésienne ou même les théories de Russell, l'information à priori était prise en compte par les actuaires. En effet, lorsqu'ils évaluaient une prime, jamais ils n'auraient considéré ne *rien* connaître du risque. À partir de différents critères, ils arrivaient au moins à le regrouper avec d'autres risques semblables. Toutefois, le mécanisme de prise en compte de cette information était souvent ad hoc et mal compris à la fois par les actuaires et les statisticiens. C'était la crédibilité.

Si la crédibilité a su se développer et faire le bonheur de ses utilisateurs pendant de nombreuses années sans les bases statistiques de l'approche bayésienne, il n'en demeure pas moins que l'arrivée de celle-ci marqua un virage important dans l'évolution de la crédibilité, contribuant d'ailleurs à en faire davantage une théorie. Les lignes suivantes tâcheront de retracer l'évolution historique de la théorie de la crédibilité, de ses balbutiements aux débuts du siècle à son adolescence actuelle, tout en faisant le lien avec les théories présentées plus haut.

La petite histoire suivante, inventée par le professeur François Dufresne pour un cours d'introduction à la théorie de la crédibilité à l'Université Laval, illustre sans doute bien comment cette théorie a pu voir le jour.

Vers les années 1910, aux États-Unis, la multinationale General Motors et le petit constructeur indépendant Tucker sont assurés chez Allstate contre les accidents de travail (*workers compensation*), avec quelques autres fabricants d'automobiles. Un taux moyen est calculé à partir de l'expérience de l'ensemble de ces fabricants et c'est ce taux qui est chargé à chacun. Or, la GM calcule elle-même son taux et s'aperçoit qu'il serait, année après année, inférieur à celui qu'on lui charge, et ce, grâce à une expérience meilleure que celle du groupe. Exaspérée par une telle situation, elle demande à Allstate qu'on lui charge son propre taux, cela sous prétexte que son nombre d'employés important est un gage de stabilité de l'expérience entre les années.

Les actuaires de Allstate sont intuitivement d'accord avec l'argumentation de GM, aussi s'apprêtent-ils à accéder à sa demande. Un petit problème les fait cependant hésiter : si le nombre d'employés de GM est clairement assez gros pour que l'on se fie à son expérience et celui de Tucker trop petit pour faire de même, où fixera-t-on la limite entre un nombre d'employés fiable et un non fiable ?

Il est généralement reconnu que le premier actuaire à avoir proposé une solution au problème dans la littérature est Arthur H. Mowbray, dans le premier numéro des Proceedings de la Casualty Actuarial Society, en 1914 (Mowbray, 1914). En supposant connue la probabilité q qu'un accident survienne, 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 de $100k$ % de la moyenne (ou du mode [2]) soit supérieure à $100p$ %. Si l'on note N la distribution du nombre d'accidents, alors l'énoncé précédent s'écrit en langage mathématique

$$P\left[(1-k)E[N] \leq N \leq (1+k)E[N]\right] \geq p,$$

où $N \sim \text{Binomiale}(n, q)$. Par la suite, l'utilisation de l'approximation normale pour la distribution de N permet d'éviter l'arbitrage entre le mode et la moyenne et d'obtenir aisément une réponse.

Les bases de la crédibilité de stabilité, ou *limited fluctuations*, étaient dès lors posées.

La solution de Mowbray était intéressante, car appuyée par des arguments probabilistes. Elle connut d'ailleurs de multiples adaptations qui lui permirent de rester d'actualité jusqu'à nos jours. Moyennant la détermination d'une distribution pour le nombre de sinistres, les praticiens pouvaient désormais résoudre par un calcul simple le problème auquel ils faisaient face : fixer un seuil d'admissibilité à la pleine crédibilité.

Les assurés eurent cependant tôt fait de soumettre un autre problème aux actuaires. Telle que présentée jusqu'à maintenant, la théorie de la crédibilité n'admettait que deux niveaux : 0 et 1. Or, cette situation pouvait se traduire, pour un employeur situé tout juste sous le seuil d'admissibilité, en une différence significative dans la prime à payer. De plus, la notion même de pleine crédibilité ne ralliait pas l'ensemble des actuaires, certains croyant que jamais les données ne sont fiables à 100 %.

C'est pour répondre à ces critiques que fut introduit le concept de crédibilité partielle, dont on attribue la première véritable théorie sur le sujet à Whitney (1918). Dès 1918, Whitney mentionne la « nécessité, par souci d'équité pour l'assuré, de pondérer d'un côté l'expérience collective et de

2. Mowbray privilégie la valeur la plus fréquente, le mode, mais s'en remet à la moyenne par simplicité et sous prétexte que ces deux valeurs sont presque égales.

l'autre l'expérience individuelle ». L'essence de la crédibilité consistait à calculer cette pondération.

À au moins deux reprises dans l'histoire, la crédibilité de précision (*greatest accuracy*) a eu la chance de s'imposer avant la contribution de Bühlmann en 1967. La première se trouve dans les travaux de Whitney. Dans son article, ce dernier retient quatre éléments qui influenceront la pondération à donner à l'expérience individuelle : l'exposition, le niveau de risque, la crédibilité de la prime collective et l'homogénéité du groupe. D'ailleurs, il mentionne :

> « Il n'y aurait pas de problème d'*experience rating* si chaque risque dans le groupe était typique du groupe, car dans ce cas les variations dans l'expérience ne seraient que purement aléatoires. » (Traduction libre)

Or, la notion d'homogénéité du collectif est au cœur même de la crédibilité de précision. Whitney modélise l'hétérogénéité du collectif en supposant que les moyennes des divers risques sont distribuées selon une loi normale. De là, par un développement mathématique lourd et laborieux, il obtient une expression pour la prime individuelle (P) de la forme

$$P = zX + (1 - z)C,$$

où X est l'expérience individuelle et C, l'expérience collective. Ces valeurs sont pondérées par le facteur de crédibilité, z, dont Whitney obtient une formule de la forme $n/(n+K)$. K n'est alors pas une constante arbitraire, mais bien une expression explicite dépendant des divers paramètres du modèle. Par souci de simplicité, Whitney suggère cependant de fixer K comme une constante à déterminer au jugement de manière à éviter les trop grandes fluctuations entre la prime individuelle et la prime collective. Ce faisant, il s'écarte d'une conception de la crédibilité visant la précision pour encourager plutôt celle visant la stabilité. La crédibilité de précision meurt dans l'œuf à cause de considérations d'ordre pratique.

Whitney fut de plus critiqué par Fisher (1919) pour avoir utilisé une méthode jugée hérétique à l'époque en statistique, la règle de Bayes. Fisher cite d'ailleurs quelques personnes faisant autorité s'étant prononcées contre l'usage de cette règle. En fait, Whitney utilise la version la plus simple de la règle, celle même proposée par Bayes[3], qui suppose qu'à priori tous les événements ont une chance égale de se réaliser. Cette modélisation était appelée par ses adeptes le « Principe de la raison insuffisante » et par ses détracteurs

3. Selon Bailey (1950).

« L'hypothèse de distribution uniforme de l'ignorance »[4]. Dans le développement de Whitney, cela revient à supposer que toutes les valeurs possibles pour la moyenne du groupe sont équiprobables. Les réserves de Fisher sur ce point étaient donc justifiées...

Il est difficile aujourd'hui de mesurer l'impact que la volte-face de Whitney en faveur des arguments de stabilité a pu avoir sur la pratique de la crédibilité. Toujours est-il que les actuaires ont utilisé intensivement pendant un demi-siècle la forme $z = n/(n+K)$ pour calculer des facteurs de crédibilité basés essentiellement sur la stabilité. Encore aujourd'hui, c'est l'approche privilégiée aux États-Unis. Pourtant, si les pratiques se sont quelque peu figées, la théorie, elle, n'a pas cessé d'évoluer.

L'approche de précision tente à nouveau sa chance par l'entremise de Arthur L. Bailey, et ce, à deux reprises (Bailey, 1945, 1950). En 1945 d'abord, Bailey obtient une expression pour la crédibilité dans ce qui semble être l'univers non paramétrique exploré plus tard par Bühlmann. Seulement, une notation confuse rend le texte difficilement déchiffrable et le condamne à la marginalité.

L'article de 1950 avait, lui, davantage le potentiel pour ébranler les acquis en théorie de la crédibilité. Bailey débute son article en relevant les confrontations historiques entre statisticiens et actuaires au sujet de la crédibilité. On apprend alors qu'il y a plus de trente ans, la polémique au sujet de la crédibilité était la même que de nos jours : les statisticiens « purs » crient au scandale, estimant les méthodes actuarielles contraires à toute théorie statistique, et les actuaires leur répondent que « peut-être, mais ça fonctionne ! » Le point majeur de discorde est l'utilisation par les actuaires de la règle de Bayes. Ainsi, Bailey (1950) écrit :

> « Présentement, presque toutes les méthodes d'estimation présentées dans les livres de méthodes statistiques ou enseignées dans les universités américaines sont basées sur un équivalent de l'hypothèse selon laquelle toute information parallèle ou connaissance à priori est inutile. (...) Des philosophes [Russell] ont récemment étudié la crédibilité à accorder à divers éléments de savoir, remettant par conséquent en doute la philosophie adoptée par les statisticiens. Par contre, il semble que ce ne soit que dans le domaine de l'actuariat qu'on ait assisté à une réelle révolte contre la mise de côté de tout savoir à priori au moment d'estimer une quantité à l'aide de nouvelles données. » (Traduction libre)

4. *Principe of insufficient reason* et *Assumption of the equal distribution of ignorance.*

Bailey se prononce donc clairement en faveur de la philosophie bayésienne. Il semble même qu'il fut le premier à en défendre si énergiquement la pertinence dans le processus de tarification. Mais ceci, fait-il remarquer, à condition d'utiliser une généralisation de la règle de Bayes, faite par Laplace en 1820, de manière à éviter le controversé « Principe de la raison insuffisante ». Cette généralisation rend la règle de Bayes applicable même si les événements n'ont pas à priori une chance égale de se réaliser.

À partir de ces prémisses, Bailey démontre qu'en minimisant l'erreur quadratique dans un contexte bayésien, l'estimateur obtenu est une fonction linéaire des observations [5] qui correspond exactement à la prime de crédibilité, et ce, pour les combinaisons de distributions binomiale/bêta, Poisson/-gamma et normale/normale. Il est possible d'en déduire un facteur de crédibilité encore une fois de la forme $z = n/(n+K)$ où K est une combinaison des paramètres du modèle. Contrairement à Whitney, Bailey ne propose pas d'évaluer K au jugement, mais bien de s'en tenir à l'expression développée algébriquement.

De plus, conscient que la procédure utilisée est une modélisation de l'hétérogénéité d'un groupe, Bailey mentionne pour la première fois que la crédibilité pourrait être utilisée hors du domaine de l'assurance générale. Malheureusement pour Bailey, le manque de popularité de l'approche bayésienne au sein de la communauté statistique au moment de la parution reléguera son article dans l'ombre. Sans doute l'approche très « théorique » de l'article, publié dans une revue essentiellement de praticiens, aura-t-elle aussi contribué à cet état de fait. Pourtant, dans un avenir pas si lointain, les travaux de Bailey seront reconnus pour leur importance et leur avant-gardisme. Cette reconnaissance proviendra cependant majoritairement d'outre-Atlantique.

L'apport de Bailey à la théorie de la crédibilité peut être résumé en deux points principaux : l'introduction explicite du principe de Bayes dans le processus de tarification et la découverte de la linéarité de l'estimateur bayésien sous certaines conditions [6]. Ces développements parvinrent aux oreilles de la communauté actuarielle européenne par la bouche de Bruno de Finetti au colloque ASTIN à Trieste, en 1963. Depuis quelques années déjà, les chercheurs européens tâchaient de trouver une justification théorique à ces formules de crédibilité américaines qui fonctionnaient si bien. L'approche bayésienne avait là aussi fait des progrès importants, notamment grâce à Bruno de Finetti dans les années 1930 et à Ove Lundberg dans les années 1940.

5. De la forme $c_0 + \sum_{i,j} c_{ij} X_{ij}$.

6. Sur ce point toutefois, il est difficile de savoir qui fut le pionnier. Norberg (1979) relève que Keffer (1929) a obtenu le même résultat pour le modèle Poisson/gamma. Toujours selon Norberg, il existerait aussi des références antérieures à 1929.

Parallèlement à tout cela, une nouvelle branche de la théorie bayésienne voyait le jour sous l'initiative de Herbert Robbins : l'approche bayésienne empirique. Celle-ci sera d'une importance capitale dans le développement futur de la crédibilité de précision puisqu'elle lui permettra de sauter le mur entre la théorie et la pratique. L'approche bayésienne empirique (Robbins, 1955, 1964) s'attaque au principal problème pratique de l'approche bayésienne pure, soit la fréquente ignorance de la distribution à priori. Jusqu'alors, lorsqu'une telle situation se présentait, il était coutume d'éviter le problème en réfutant l'aspect aléatoire d'une partie de la décision (Neyman, 1962). Grossièrement, la thèse de Robbins consiste à supposer que, bien qu'elle soit inconnue, la distribution à priori existe et qu'il est possible de l'estimer — ne serait-ce qu'indirectement — à partir de données issues de plusieurs expériences similaires. Selon lui (Robbins, 1964) :

> « L'approche bayésienne empirique de problèmes statistiques de prise de décision est applicable lorsque la même décision se présente à répétition et indépendamment avec une distribution fixe, mais inconnue du paramètre. » (Traduction libre)

Comme devait plus tard le mentionner Bühlmann, cela cadre admirablement bien avec le problème de l'*experience rating*.

Tous ces éléments — les travaux de Bailey, la popularité grandissante des approches bayésienne et bayésienne empirique — étaient réunis lors du congrès ASTIN de 1965, à Lucerne. Hans Bühlmann y redéfinit alors le problème fondamental en *experience rating* et présente la solution qui allait révolutionner la théorie de la crédibilité. En forçant la prime bayésienne à être linéaire, Bühlmann (1967, 1969) obtient, dans un cadre non paramétrique, un facteur de crédibilité de la forme $z = n/(n+K)$, avec une expression simple et générale pour K. Le virage sera alors définitivement pris en faveur de l'approche de précision et l'essentiel de la recherche se fera en Europe. C'est pourquoi l'approche traitant l'hétérogénéité est aujourd'hui souvent appelée « crédibilité européenne » malgré qu'elle est originaire des États-Unis (notamment dans les travaux de Bailey). L'approche de stabilité, *limited fluctuations*, a par conséquent reçu l'appellation « crédibilité américaine ».

On peut sans crainte poser que le célèbre modèle de Bühlmann marque le début de l'histoire contemporaine de la théorie de la crédibilité. Celle-ci étant bien davantage documentée et connue, on va maintenant se contenter de présenter chronologiquement les principaux modèles qui suivirent celui de Bühlmann[7]. La plupart de ces modèles se veulent des généralisations,

7. La liste ne comporte que les principaux modèles et n'est donc en rien exhaustive.

de plus en plus poussées, du modèle original de Bühlmann et bon nombre d'entres eux seront étudiés ultérieurement dans ce mémoire.

Le modèle de Bühlmann se décompose lui-même en deux parties communément appelées « modèle original » et « modèle classique ». Le premier pose les bases de la nouvelle théorie, tandis que l'apport de la théorie bayésienne empirique fait du second un modèle plus pratique. La première généralisation de ces modèles voit le jour en 1970, alors que Bühlmann s'adjoint son étudiant au doctorat Erwin Straub pour développer le très célèbre modèle qui portera leurs noms (Bühlmann et Straub, 1970). L'ajout de poids aux données et la définition d'estimateurs des paramètres de structure constituent les principales améliorations au modèle de Bühlmann. Elles sont toutefois de taille et permettront à la crédibilité de précision de véritablement faire une percée dans la pratique de l'assurance. Le modèle de Bühlmann–Straub constitue encore aujourd'hui un standard et est couramment utilisé dans les compagnies d'assurance, en Europe surtout.

En 1974, Jewell fait la première de ses deux plus grandes contributions au développement de la théorie de la crédibilité. Il démontre (Jewell, 1974) que l'estimateur bayésien est linéaire pour toute fonction de vraisemblance membre de la famille exponentielle utilisée avec sa conjuguée naturelle. Ce faisant, Jewell ne fait que confirmer certains des résultats obtenus avant lui par Bailey (1950) et Mayerson (1964), mais les unifie en une formulation générale.

Les deux années suivantes furent fastes pour la théorie de la crédibilité. D'abord, Hachemeister (1975) généralise le modèle de Bühlmann–Straub en incorporant la régression linéaire à la théorie de la crédibilité de précision. Plus tard, on poussera même cette idée un peu plus loin en étudiant la régression non-linéaire (De Vylder, 1985). Puis, non satisfait de la manière qu'a le modèle de Bühlmann–Straub d'intégrer à la prime les données collatérales, Jewell (1975) présente sa solution : la crédibilité hiérarchique. Ce qui est maintenant appelé le modèle hiérarchique de Jewell constitue une importante généralisation de plusieurs modèles de crédibilité. Toujours en 1975, Gerber et Jones (1975) définissent les propriétés et les conditions menant à des primes de crédibilité de type « mise à jour » (*updating type*). En 1976, (De Vylder, 1976b) présente ses modèles de crédibilité semi-linéaire et semi-linéaire optimal. La même année, il propose une formulation du problème de la crédibilité en termes d'espaces de Hilbert (De Vylder, 1976a). Norberg et Taylor (1977) se sont également intéressés à l'étude de la théorie de la crédibilité dans le cadre abstrait des espaces de Hilbert. Norberg (1979) a publié un imposant article révisant l'essentiel de la théorie de la crédibilité connu jusqu'alors. Il s'agit, encore aujourd'hui, d'une référence de choix pour qui désire approfondir des connaissances de base en théorie de la crédibilité.

Si les années 1970 furent celles de l'apparition de nombreux modèles de crédibilité de précision, les années 1980 furent plutôt celles de l'étude des estimateurs des paramètres de structure. De Vylder (1978, 1981, 1984) s'est alors avéré un acteur important, de même que Norberg (1980), Gisler (1980) et Dubey et Gisler (1981). On doit à ces derniers la première étude exhaustive des propriétés asymptotiques des estimateurs de variance dans le modèle de Bühlmann–Straub. La découverte de nouveaux estimateurs et leur maîtrise va permettre la plus large diffusion de la crédibilité de précision. Des volumes sur le sujet sont publiés (Goovaerts et collab., 1990) dont certains (Goovaerts et Hoogstad, 1987) sont expressément destinés aux actuaires œuvrant dans les compagnies d'assurance. De nombreux articles importants en théorie de la crédibilité sont publiés par des chercheurs au service de compagnies d'assurance : mentionnons Straub, Sundt, Dubey, Gisler, Reinhard ou même Goovaerts, à titre de consultant. Cela contribue donc à populariser l'utilisation de la théorie dans des applications pratiques.

La période s'étendant du milieu des années 1980 jusqu'au début des années 1990 a été marquée par un ralentissement dans la recherche en théorie de la crédibilité. Au cours des dernières années cependant, on note un regain d'intérêt pour la recherche d'estimateurs des paramètres de structure plus efficaces et dont les propriétés seraient davantage connues. Ainsi, Künsch (1992) et Gisler et Reinhard (1993) introduisent la statistique robuste en théorie de la crédibilité, ce qui ouvre par le fait même un nouveau et vaste champ de recherche. De leur côté, De Vylder et Goovaerts, toujours très actifs dans le domaine, proposent dans une série d'articles des estimateurs optimaux sous certaines conditions (De Vylder et Goovaerts, 1991, 1992a,b). Des recherches sont présentement en cours desquelles, nous promet-on, devraient ressortir de toutes nouvelles façons de trouver des estimateurs pour les principaux modèles de crédibilité.

La grande majorité des intéressés à la théorie de la crédibilité se rallient au modèle de base proposé par Bühlmann de même qu'à ses extensions. De plus, l'historique tracé ci-dessus correspondant à ce qui est généralement admis dans le domaine. À une exception près semble-t-il : Zehnwirth (1991) qui, sans renier les travaux de Bühlmann, a une manière bien à lui de concevoir et présenter la théorie de la crédibilité de précision. Selon lui, les formules de crédibilité sont si étroitement liées à la régression linéaire qu'elles n'en seraient que de simples dérivés. Karl Friedrick Gauss étant à l'origine de certains types de formules de régression, Zehnwirth en arrive à la conclusion suivante : « cela signifie que Gauss (1795) a dérivé la plupart des formules de crédibilité ».

1.2.3 En conclusion

Ce chapitre s'est ouvert sur une présentation de l'approche bayésienne en statistiques. Celle-ci allait en quelque sorte devenir le dénominateur commun des deux théories par la suite étudiées, celle de Bertrand Russell et la théorie actuarielle de la crédibilité. Ce pont permet l'intéressante comparaison de la définition donnée par Russell au mot « crédibilité » à sa signification actuarielle. On voit ainsi qu'en actuariat comme dans la philosophie de Russell, la crédibilité apparaît lors du passage du général au particulier, soit lors de la tarification d'un assuré pris isolément et non plus comme simple membre d'un groupe. Le traitement de l'information à priori ainsi que le gain marginal de crédibilité décroissant en théorie de la crédibilité de stabilité ou à la crédibilité de précision. Là n'en est pas le but de toute façon, puisque chacune de ces deux théories est valable en soi, seuls leurs buts étant différents.

2 Crédibilité de stabilité

Objectifs du chapitre

- Calculer le seuil de crédibilité complète à partir du modèle collectif du risque.
- Identifier la caractéristique d'un modèle la plus appropriée pour exprimer le seuil de crédibilité complète.
- Calculer un seuil de crédibilité complète basé sur le nombre d'années d'expérience.
- Calculer un facteur de crédibilité partielle selon l'approche de crédibilité de stabilité, et ce, en utilisant l'une ou l'autre des formules classiques (Whitney, racine carrée, puissance $2/3$).

La théorie de la crédibilité est apparue dans le domaine des accidents du travail au début des années 1900. Les actuaires cherchaient alors à déterminer la taille minimale (en nombre d'employés) qu'un employeur devait atteindre pour que son expérience soit considérée pleinement « fiable » (*dependable*) ou « crédible ».

Tel que relaté à la section 1.2, Mowbray (1914) fournit une première solution. Il fournit du même coup une définition de ce que doit être une prime pure crédible : « une prime pour laquelle la probabilité est forte qu'elle ne diffère pas de la vraie prime par plus d'une limite arbitraire ».

En termes mathématiques, nous souhaitons que

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

où S représente l'expérience d'un contrat, sous une forme ou une autre ; k est une petite valeur, habituellement autour de 5 % ; p une valeur près de 1, habituellement 0,90, 0,95 ou 0,99.

2.1 Crédibilité complète

En crédibilité complète, un contrat d'assurance est considéré *crédible* si son expérience est *stable*. Lorsque c'est le cas, seule l'expérience du contrat est utilisée dans la tarification de ce contrat.

> La notion de portefeuille de contrats ne joue aucun rôle pour le moment.

Intuitivement, la stabilité de l'expérience d'un contrat va de pair avec sa « taille », qu'elle soit exprimée en termes de : volume de prime, masse salariale, nombre d'employés, nombre de sinistres, nombre d'années d'expérience, etc.

De plus, la taille du contrat est généralement liée à la fréquence des sinistres, et non à la sévérité de ceux-ci.

Définition 2.1 (Crédibilité complète)**.** 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

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

est vérifiée.

La variable aléatoire S représente la somme de l'expérience de plusieurs risques. Par le Théorème central limite,

$$Z = \frac{S - E[S]}{\sqrt{\mathrm{Var}[S]}} \longrightarrow N(0,1)$$

lorsque le nombre de termes dans la somme S augmente. La relation de la définition 2.1 peut alors se réécrire

$$\begin{aligned}\Pr\left[-\frac{kE[S]}{\sqrt{\mathrm{Var}[S]}} \leq Z \leq \frac{kE[S]}{\sqrt{\mathrm{Var}[S]}}\right] &\approx \Phi\left(\frac{kE[S]}{\sqrt{\mathrm{Var}[S]}}\right) - \Phi\left(-\frac{kE[S]}{\sqrt{\mathrm{Var}[S]}}\right) \\ &= 2\Phi\left(\frac{kE[S]}{\sqrt{\mathrm{Var}[S]}}\right) - 1 \geq p.\end{aligned}$$

L'inégalité est satisfaite dès lors que

$$\frac{kE[S]}{\sqrt{\mathrm{Var}[S]}} \geq \zeta_{\varepsilon/2},$$

ou, de manière équivalente, lorsque

$$E[S] \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)\sqrt{\mathrm{Var}[S]}, \tag{2.1}$$

où $\varepsilon = 1-p$ et ζ_α est le $100(1-\alpha)^{\text{e}}$ centile d'une loi normale centrée réduite.

Exemple 2.1 (Binomiale pure)**.** Mowbray (1914) recherchait le nombre minimal d'employés pour considérer l'expérience d'un employeur pleinement crédible. Ses hypothèses étaient que la variable aléatoire S du nombre d'accidents par année pour un employeur est binomiale de paramètres n et θ, où n représente le nombre d'employés et θ la probabilité qu'un accident survienne au cours de l'année pour tout employé. Cette probabilité est supposée connue.

Avec $E[S] = n\theta$ et $\mathrm{Var}[S] = n\theta(1-\theta)$ et en isolant n dans l'inégalité (2.1), nous obtenons la relation suivante pour le critère de crédibilité complète :

$$n \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{1-\theta}{\theta}.$$

□

Exemple 2.2 (Poisson composée)**.** La taille d'un contrat est souvent exprimée en termes du nombre espéré de sinistres dans une période — typiquement une année. La distribution la plus populaire pour S dans un tel cas est alors la Poisson composée, c'est-à-dire

$$S = X_1 + \cdots + X_N$$

où $N \sim \text{Poisson}(\lambda)$ et la distribution de $X_1, X_2, \dots$ est $F_X(\cdot)$. Nous savons que

$$\begin{aligned}
E[S] &= E[N]E[X] \\
&= \lambda E[X] \\
\mathrm{Var}[S] &= \mathrm{Var}[N]E[X]^2 + E[N]\,\mathrm{Var}[X] \\
&= \lambda E[X]^2 + \lambda\,\mathrm{Var}[X] \\
&= \lambda E[X^2].
\end{aligned}$$

La relation (2.1) devient alors, en isolant le nombre espéré de sinistre λ :

$$\begin{aligned}
\lambda &\geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right) \\
&= \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 (1 + \mathrm{CV}(X)^2),
\end{aligned} \tag{2.2}$$

où

$$\mathrm{CV}(X) = \frac{\sqrt{\mathrm{Var}[X]}}{E[X]}$$

est le coefficient de variation de la variable aléatoire X.

Vous remarquerez que, dans l'inégalité (2.2), le seuil de crédibilité complète λ augmente avec la variabilité des montants de sinistres (exprimée en fonction de coefficient de variation $\mathrm{CV}(X)$).

Si $k = 0{,}05$ et $p = 0{,}90$, alors $\zeta_{0{,}05} = 1{,}645$ et

$$\lambda \geq 1\,082{,}41 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right).$$

Si, de plus, X est une variable aléatoire dégénérée (c'est-à-dire que $\Pr[X = M] = 1$ pour M quelconque), alors

$$\lambda \geq 1\,082{,}41.$$

Ce cas revient en définitive à poser $S \sim \mathrm{Poisson}(\lambda)$. □

Le seuil de pleine crédibilité 1 082 est un nombre célèbre en théorie de la crédibilité. Malheureusement, il est très souvent utilisé à tort et à travers et sans aucunement tenir compte des hypothèses qui nous ont permis d'obtenir cette valeur.

Dans les deux exemples précédents, nous avons déterminé le seuil de pleine crédibilité d'un contrat pour une seule période. Nous pourrions aussi choisir de fixer le seuil en fonction du nombre d'années d'expérience. Pour cela, il suffit de définir

$$W = \frac{S_1 + \dots + S_n}{n},$$

où $S_1, \dots, S_n$ sont des variables aléatoires indépendantes et identiquement distribuées et S_t est l'expérience de l'année $t = 1, \dots, n$. Nous avons alors

$$\begin{aligned} E[W] &= E[S_t] \\ \mathrm{Var}[W] &= \frac{\mathrm{Var}[S_t]}{n}. \end{aligned}$$

En appliquant le critère de crédibilité complète sur la variable aléatoire W, nous obtenons une expression en tous points équivalente à l'inégalité

(2.1) dans laquelle nous remplaçons l'espérance et la variance par les expressions ci-dessus. L'expérience d'un contrat est considérée pleinement crédible après

$$n \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{\mathrm{Var}[S_t]}{E[S_t]^2} \tag{2.3}$$

périodes.

La distribution de S n'étant en général pas symétrique, il serait légitime de nous demander s'il est correct d'utiliser le Théorème central limite (TCL) pour calculer le seuil de crédibilité complète. En fait, puisque le critère impose que la distribution de S soit très concentrée autour de sa moyenne, elle s'en trouve presque symétrique et le TCL s'avère donc bien assez précis (Goulet, 1998).

Il y a en fait très peu d'applications légitimes de la crédibilité de stabilité. Un bon exemple, toutefois : la fixation du seuil d'admissibilité à un régime de tarification rétrospectif, où la stabilité de l'expérience joue un rôle important.

2.2 Crédibilité partielle

La crédibilité complète ne connait que deux états : aucune crédibilité ou crédibilité complète. Le besoin (ou la pression) de tenir compte en partie de l'expérience individuelle d'un contrat se trouvant sous le seuil de crédibilité complète apparait rapidement chez les assureurs.

Whitney (1918) propose de pondérer l'expérience individuelle (S) et la prime collective (m) par un *facteur de crédibilité* ($0 \leq z \leq 1$) en une prime de la forme

$$\pi = zS + (1 - z)m.$$

Plusieurs formules différentes ont été proposées pour z. Soit n_0 le seuil de crédibilité complète. Les formules les plus populaires sont

$$z = \min\left\{\sqrt{\frac{n}{n_0}}, 1\right\},$$

$$z = \min\left\{\left(\frac{n}{n_0}\right)^{2/3}, 1\right\},$$

et la formule de Whitney

$$z = \frac{n}{n+K},$$

où K est une constante déterminée au jugement de façon à limiter les fluctuations dans la prime d'une année à l'autre (*swing*).

Le but de cette approche consiste à incorporer autant d'expérience individuelle que possible dans la prime sans trop en affecter la stabilité d'une année à l'autre.

> La distribution des primes dans cette approche de crédibilité partielle est basée uniquement sur la taille des assurés. Rien n'assure que la tarification est précise ou équitable.

2.3 Exercices

2.1 Le montant total des sinistres, S, a une distribution binomiale composée de paramètres $n = 1\,000$ et $\theta = 0{,}6$. La distribution du montant des sinistres est une log-normale de paramètres $\mu = 3$ et $\sigma^2 = 4$. Calculer $E[S]$ et $\mathrm{Var}[S]$.

2.2 Sachant que le nombre de sinistres a une distribution binomiale négative de paramètres $r = 4$ et $\theta = 0{,}5$ et que la sévérité d'un sinistre a une distribution gamma de paramètres $\alpha = 2$ et $\lambda = 0{,}5$ calculer l'espérance et la variance du montant total des sinistres, S.

2.3 Interpréter l'inégalité suivante :

$$\Pr\left[0{,}96E[S] \leq S \leq 1{,}04E[S]\right] \geq 0{,}95.$$

2.4 Dériver une formule générale pour déterminer le niveau de crédibilité complète d'ordre (k, p) pour $\bar{S} = (S_1 + \cdots + S_n)/n$ lorsque la distribution du montant total des sinistres est une Poisson composée.

2.5 Soit $N \sim \text{Poisson}(256)$, $X \sim \text{Pareto}(3, 0{,}05)$ et $S = X_1 + \cdots + X_N$.

a) Quelle est la plus petite marge d'erreur admissible autour de $E[S]$ faisant toujours en sorte que S a une crédibilité complète à 90 % ? Interpréter brièvement le résultat.

b) Quelle est la plus petite marge d'erreur admissible autour de $E[\bar{S}]$ faisant toujours en sorte que $\bar{S}$ a une crédibilité complète à 90 % après dix années ?

2.6 Sachant que la variable aléatoire S de l'expérience des contrats obéit à une loi $N(\mu, \sigma^2)$, trouver, pour une période d'expérience, la relation entre μ et σ pour avoir une crédibilité complète d'ordre (k, p) pour chacune des combinaisons de k et p suivantes.

a) $(0{,}04, 0{,}95)$

b) $(0{,}05, 0{,}90)$

c) $(0{,}01, 0{,}98)$

2.7 Sachant que $S \sim N(\mu, \sigma^2)$ et $\bar{S} = (S_1 + \cdots + S_9)/9$, où $S_1, \dots, S_9$ sont toutes des variables aléatoires mutuellement indépendantes distribuées comme S, déterminer la relation entre μ et σ faisant en sorte que $\bar{S}$ a une crédibilité complète d'ordre (k, p) pour chacune des combinaisons de k et p suivantes.

a) $(0{,}05, 0{,}90)$

b) $(0{,}05, 0{,}95)$

c) $(0{,}01, 0{,}90)$

d) $(0{,}01, 0{,}95)$

2.8 Juliette travaille au département de tarification d'une compagnie d'assurance très active en assurance automobile. À la suite d'une analyse de données exhaustive, Juliette peut affirmer que la fréquence des sinistres pour ce type de produit a une distribution binomiale négative de paramètres r et $\theta = 0{,}01$. La sévérité des sinistres, quant à elle, suit une loi gamma de paramètres $\alpha = 0{,}02$ et $\lambda = 1$. Trouver la plus petite valeur de r telle que le montant total des sinistres d'un contrat d'assurance automobile sera à plus ou moins 5 % égal à sa moyenne, 19 fois sur 20.

2.9 Soit S_j le montant total des sinistres pour un assuré à la période $j = 1, \dots, n$ tel que

$$S_j = X_1 + X_2 + \cdots + X_{N_j},$$

où $X_1, X_2, \dots$ sont les montants individuels des sinistres dont la distribution est dégénérée en M et N_j suit une loi de Poisson de paramètre λ. Calculer le nombre total de sinistres espéré minimal pour accorder une crédibilité complète d'ordre $(0{,}04, 0{,}90)$ à l'expérience individuelle $\bar{S}$.

2.10 Soit S_j le montant total des sinistres pour un assuré à la période $j = 1, \dots, n$ et $S_j \sim$ Poisson composée$(\lambda, F_X(\cdot))$. Le montant d'un sinistre individuel a une variance de 100 et une espérance de 5. Déterminer le nombre minimal espéré de sinistres pour accorder une crédibilité complète selon les critères ci-dessous.

a) Le montant total des sinistres demeure à 3 % ou moins du montant espéré avec une probabilité de 95 %.

b) Le nombre total de sinistres demeure à 3 % ou moins du nombre espéré avec une probabilité de 95 %.

2.11 Une compagnie assure deux groupes ayant la même loi de Poisson pour la fréquence de leurs sinistres individuels. Cependant, les individus du groupe A ne peuvent avoir que des sinistres de 50 alors que les individus du groupe B ont des sinistres obéissant à une loi gamma de moyenne 50. Sachant que l'observation de 1 000 sinistres est suffisante pour accorder une crédibilité complète au groupe A et que 3 000 sinistres sont nécessaires pour accorder une crédibilité complète de même ordre au groupe B, calculer les paramètres de la loi gamma en question.

2.12 Dans un modèle Poisson composée, on accorde une crédibilité complète d'ordre $(0{,}05, p)$ au nombre total de sinistres si le nombre de sinistres observé est supérieur à 1 000. Quel doit être le nombre minimal de sinistres pour accorder une crédibilité d'ordre $(0{,}25, p)$ au montant total des sinistres si le montant d'un sinistre individuel suit une loi $\text{Gamma}(2, 2)$?

2.13 Pour ses calculs relatifs à l'admission au régime rétrospectif, la CNESST peut n'employer que la fréquence des sinistres, ou encore la fréquence ainsi que la sévérité. Dans le premier cas, le nombre minimal d'employés pour être admissible au régime rétrospectif est de 6 494 en supposant que la probabilité d'avoir un accident est de 0,04. Sachant que le coefficient de variation de la sévérité des sinistres est 2, que devient le seuil d'admissibilité lorsque la sévérité est également prise en compte dans les calculs ?

2.14 Le montant total des sinistres a une distribution Poisson composée où les montants de sinistres individuels proviennent d'une loi de Pareto de paramètres $\alpha = 3$ et $\lambda = 100$. Lorsque la largeur de l'intervalle de confiance autour de la moyenne est 10 % de celle-ci, le seuil de crédibilité complète est 2 500. On décide de changer la distribution de la fréquence des sinistres pour une binomiale négative avec $\theta = 0{,}5$. Si le seuil de crédibilité complète et le niveau de confiance de l'intervalle autour de la moyenne demeurent tous deux inchangés, quelle est la nouvelle largeur de l'intervalle de confiance ?

2.15 On vous donne les renseignements suivants :

i) $S = \sum_{j=1}^{N} X_j$ et les variables aléatoires X_j sont mutuellement indépendantes et indépendantes de N.

ii) $X_j \sim \text{Pareto}(3, 3)$

iii) $N \sim \text{Binomiale négative}(r, 1/3)$.

Calculer la valeur minimale de r pour accorder une crédibilité complète d'ordre $(0{,}05, 0{,}90)$ à S. Utiliser l'approximation normale.

2.16 Soit S la variable aléatoire du montant total des sinistres pour un portefeuille d'assurance. Sachant que $S \sim \text{Poisson composée}(\lambda, F_X(\cdot))$, trouver le seuil de crédibilité complète en termes du nombre espéré de sinistres selon les hypothèses de distribution suivantes.

a) $\Pr[X = 1] = 1$ (dégénérée en 1), $k = 0{,}05$ et $p = 0{,}90$.

b) $X \sim \text{Exponentielle}(2)$, $k = 0{,}04$, $p = 0{,}95$.

c) Si le seuil de crédibilité complète selon les conditions en b) n'est pas atteint, quel facteur de crédibilité partielle accorderait-on à $\bar{S}$ pour une période d'expérience (en fonction de λ) ?

2.17 L'actuaire de la compagnie ABC croit qu'il faudrait 3 000 sinistres pour accorder une crédibilité complète à un assuré d'un groupe si la sévérité est constante. Après une étude, il se rend compte que la sévérité suit plutôt une loi de Pareto de moyenne 1 000 avec $\alpha = 3$. Si le nombre de sinistres obéit à une loi de Poisson, combien de sinistres doit-on avoir observé si l'on a accordé, après une période d'expérience, une crédibilité de 0,5 à l'expérience individuelle de l'assuré ?

2.18 On vous donne les informations suivantes :

i) Le nombre de sinistres suit une loi de Poisson.

ii) Le montant des sinistres a une distribution log-normale avec un coefficient de variation de 3.

iii) Le nombre et les montants de sinistres sont indépendants.

iv) Le nombre de sinistres la première année fut de 1 000.

v) Le montant total des sinistres la première année fut de 6,75 millions.

vi) La prime collective pour la seconde année est de 5,00 millions.

vii) Le volume du contrat est le même pour la première et la seconde année.

viii) Le niveau de crédibilité complète assure que le montant total des sinistres sera à 5 % de la moyenne 95 % du temps.

Déterminer la prime de crédibilité (partielle) selon l'approche de crédibilité de stabilité. Calculer le niveau de crédibilité complète sur le nombre d'années d'expérience et utiliser la formule de la racine carrée pour calculer le facteur de crédibilité.

2.19 On vous dit que $S \sim$ Poisson composée$(\lambda, F_X(\cdot))$. La fréquence annuelle espérée des sinistres est évaluée à 0,035. La grandeur minimale du portefeuille de l'assureur pour accorder une crédibilité complète à l'expérience est 103 500. Pour accorder une crédibilité de 0,67 à l'expérience d'une période, quelle doit être la grandeur minimale du portefeuille ?

2.20 Les contrats d'une compagnie d'assurance pour un certain type de produit ont les caractéristiques ci-dessous.

	Nombre de sinistres	Montant des sinistres
Espérance	10	5 000
Variance	10	6 250 000

Une crédibilité complète est accordée à l'expérience d'un contrat après n années si celle-ci se concentre dans un intervalle de 10 % autour de sa moyenne avec probabilité de 90 %. Déterminer après combien d'années d'expérience le facteur de crédibilité partielle sera de 0,54 selon chacune des formules de crédibilité partielle ci-dessous.

a) $z = \left(\dfrac{n}{n_0}\right)^{1/2}$

b) $z = \left(\dfrac{n}{n_0}\right)^{2/3}$

3 Tarification bayésienne

Objectifs du chapitre

- Expliquer le modèle d'hétérogénéité de Bühlmann, notamment le rôle de la variable aléatoire représentant le niveau de risque.
- Calculer la prime de risque, la prime collective, la prime bayésienne et la prime de crédibilité.
- Calculer la distribution à postériori pour des modèles discrets et continus.
- Calculer la distribution prédictive pour des modèles discrets et continus.
- Calculer la prime bayésienne pour des modèles discrets et continus, en particulier pour les combinaisons de distributions débouchant sur une prime linéaire.
- Énoncer les conditions du modèle de Jewell pour obtenir une prime bayésienne linéaire.
- Utiliser la fonction `cm` du paquetage R **actuar** pour calculer des primes bayésiennes linéaires et interpréter les résultats de la fonction.

Ce chapitre constitue notre porte d'entrée vers la crédibilité de précision. Il est très important pour la suite puisque nous y poserons les bases sur lesquelles nous construirons les modèles de crédibilité, notamment le modèle d'hétérogénéité.

Avant d'aller plus loin, rappelons que l'objectif que nous poursuivons consiste à établir une tarification optimale pour un regroupement (ou portefeuille) hétérogène de contrats d'assurance. Par « optimale », nous entendons que la tarification est la plus précise pour chaque contrat du regroupement.

L'approche purement bayésienne explorée dans ce chapitre est plutôt théorique. Vous constaterez que la notion de regroupement de contrats y joue un rôle très mineur, dans la mesure où nous pouvons effectuer toute notre modélisation sans vraiment y faire référence. Le portefeuille devient incontournable lorsque nous devons estimer certaines quantités à partir de données. Nous n'aborderons toutefois cette approche empirique qu'au chapitre 4.

Le chapitre débute par une mise en situation très simple, mais qui permet de bien illustrer le genre de problèmes auxquels nous sommes confrontés en tarification basée sur l'expérience, ainsi que les solutions que nous proposerons avec l'approche bayésienne. À la section 3.2, nous mettrons en place un modèle mathématique permettant de décrire l'état d'hétérogénéité du portefeuille. De ce modèle découleront, à la section 3.3, diverses prévisions, pour employer la terminologie statistique. En termes actuariels, ces prévisions constituent des primes d'assurance. À l'aide de plusieurs exemples, nous démontrerons que le calcul de la prime bayésienne peut parfois s'avérer ardu et nous établirons les conditions faisant en sorte que cette prime soit non seulement simple à calculer, mais qu'elle adopte une forme linéaire facile à interpréter.

Si le contexte et la terminologie de l'approche bayésienne en statistique ne vous sont pas familiers ou frais à la mémoire, révisez le contenu de l'annexe A avant d'aller plus loin.

3.1 Mise en situation

Un portefeuille d'assurance automobile est composé de 75 % de bons conducteurs et de 25 % de mauvais conducteurs. Les bons conducteurs ont une probabilité de $1/15$ d'avoir un accident, alors que la probabilité est de $1/10$ pour les mauvais conducteurs. Le coût d'un accident est de 1 000 et nous supposons qu'un conducteur ne peut subir qu'un accident dans une année.

1. Nous choisissons un assuré au hasard. Quelle est la probabilité que cet assuré ait un accident dans l'année qui suit ?

   Pour effectuer les calculs, nous définissons les deux événements suivants :

$$\begin{aligned} A &: \text{avoir un accident} \\ B &: \text{être un bon conducteur.} \end{aligned}$$

   Nous noterons $\bar{B}$ le fait d'être un mauvais conducteur.

Nous cherchons $\Pr[A]$ sachant que

$$\begin{aligned}\Pr[A|B] &= \frac{1}{15}\\ \Pr[A|\bar{B}] &= \frac{1}{10}\\ \Pr[B] &= \frac{3}{4}\\ \Pr[\bar{B}] &= \frac{1}{4}.\end{aligned}$$

Par la loi des probabilités totales,

$$\begin{aligned}\Pr[A] &= \Pr[A|B]\Pr[B] + \Pr[A|\bar{B}]\Pr[\bar{B}]\\ &= \left(\frac{1}{15}\right)\left(\frac{3}{4}\right) + \left(\frac{1}{10}\right)\left(\frac{1}{4}\right)\\ &= \frac{3}{40}.\end{aligned}$$

> Nous pouvons aussi exprimer cette probabilité sous la forme
>
> $$\Pr[A] = E[\Pr[A|\text{type de conducteur}]],$$
>
> ce qui permet d'accentuer le fait qu'il s'agit d'une moyenne (pondérée) des deux probabilités d'avoir un accident.

2. Quelle est la prime pure de cet assuré la première année ?

   Soit S le montant total des sinistres de cet assuré. La prime pure correspond à la quantité $E[S]$. Pour calculer cette espérance, nous pouvons procéder de deux façons.

   Tout d'abord, dans la mesure où l'assuré ne peut avoir qu'un seul sinistre dans l'année, que le montant de celui-ci est 1 000 et que la probabilité que ce sinistre survienne est 3/40, tel que calculé ci-dessus, la fonction de masse de probabilité de la variable aléatoire S est :

$$\Pr[S = x] = \begin{cases}\frac{3}{40}, & x = 1\,000\\ \frac{37}{40}, & x = 0.\end{cases}$$

   Par conséquent,

$$E[S] = 1\,000\left(\frac{3}{40}\right) = 75.$$

L'autre approche pour calculer la prime pure d'un conducteur choisi au hasard consiste à calculer la prime pure pour chaque type de conducteur et à calculer la moyenne de ces primes. De l'énoncé de la mise en situation, nous pouvons établir

$$\Pr[S = x|B] = \begin{cases} \frac{1}{15}, & x = 1\,000 \\ \frac{14}{15}, & x = 0 \end{cases}$$
$$\Pr[S = x|\bar{B}] = \begin{cases} \frac{1}{10}, & x = 1\,000 \\ \frac{9}{10}, & x = 0, \end{cases}$$

d'où

$$E[S|B] = \frac{200}{3}$$

et

$$E[S|\bar{B}] = 100.$$

Ainsi, nous obtenons la même prime pure que ci-dessus :

$$\begin{aligned} E[S] &= E[E[S|\text{type de conducteur}]] \\ &= E[S|B]\Pr[B] + E[S|\bar{B}]\Pr[\bar{B}] \\ &= \frac{200}{3}\left(\frac{3}{4}\right) + 100\left(\frac{1}{4}\right) \\ &= 75. \end{aligned}$$

3. L'assuré choisi précédemment a subi un accident dans la première année. Quelle est la probabilité qu'il s'agisse d'un bon conducteur ?

   Nous cherchons $\Pr[B|A]$, que nous pouvons interpréter comme la révision de notre opinion sur le niveau de risque du conducteur suite à l'ajout d'information. Or, par la règle de Bayes,

$$\begin{aligned} \Pr[B|A] &= \frac{\Pr[A|B]\Pr[B]}{\Pr[A]} \\ &= \frac{(1/15)(3/4)}{3/40} \\ &= \frac{2}{3} < \frac{3}{4}. \end{aligned}$$

   Nous estimons donc moins probable que l'assuré soit un bon conducteur puisqu'il a subi un accident dans la première année.

4. Quelle est la prime pure de cet assuré pour la seconde année ?

   Nous cherchons $E[S|A]$. Encore ici, nous pouvons procéder de deux manières qui s'avèrent équivalentes en bout de ligne.

   La première approche consiste à calculer la fonction de masse de probabilité de la variable aléatoire $S|A$. Nous avons

$$\begin{aligned}\Pr[A|A] &= \Pr[A|B]\Pr[B|A] + \Pr[A|\bar{B}]\Pr[\bar{B}|A] \\ &= \left(\frac{1}{15}\right)\left(\frac{2}{3}\right) + \left(\frac{1}{10}\right)\left(\frac{1}{3}\right) \\ &= \frac{7}{90},\end{aligned}$$

   d'où

$$\Pr[S = x|A] = \begin{cases}\frac{7}{90}, & x = 1\,000 \\ \frac{83}{90}, & x = 0\end{cases}$$

   et

$$E[S|A] = 1\,000\left(\frac{7}{90}\right) = \frac{700}{9}.$$

   Dans la seconde approche, nous calculons l'espérance en conditionnant sur le type de conducteur :

$$\begin{aligned}E[S|A] &= E[S|B]\Pr[B|A] + E[S|\bar{B}]\Pr[\bar{B}|A] \\ &= \frac{200}{3}\left(\frac{2}{3}\right) + 100\left(\frac{1}{3}\right) \\ &= \frac{700}{9}.\end{aligned}$$

3.2 Modèle d'hétérogénéité

Cette section introduit le modèle que nous utiliserons dorénavant en crédibilité de précision. Établi par Bühlmann (1967, 1969), ce modèle permet de décrire remarquablement bien la situation à laquelle nous sommes confrontés, à savoir que nous avons regroupé des contrats d'assurance aux caractéristiques semblables, mais que subsiste dans le groupe de l'hétérogénéité dans les niveaux de risque. C'est cette hétérogénéité qui exige que nous ayons recours à l'expérience des contrats pour établir une tarification juste et équitable.

Ainsi, nous avons un portefeuille (groupe) hétérogène de I contrats. Le niveau de risque du contrat $i = 1, \dots, I$ est inconnu, mais des données $S_{i1}, \dots, S_{in}$ sont disponibles pour fins de tarification.

L'élément clé dans notre modèle est le suivant : nous posons l'existence d'une variable aléatoire Θ_i qui représente le *niveau de risque* du contrat i. Cette variable aléatoire est bien évidemment non observable — il n'y aurait pas de problème de tarification autrement — et nous supposons qu'elle demeure constante dans le temps. En d'autres termes, la variable aléatoire Θ_i constitue la synthèse de tous les facteurs de risque associé au contrat i que nous ne savons prendre en compte par ailleurs dans la structure de classification.

La fonction de répartition de la variable aléatoire Θ — aussi appelée la *fonction de structure* du portefeuille — est $U(\theta)$ et $u(\theta)$ est la fonction de densité (ou de masse) de probabilité correspondante.

Nous posons les hypothèses suivantes pour compléter notre modèle d'hétérogénéité.

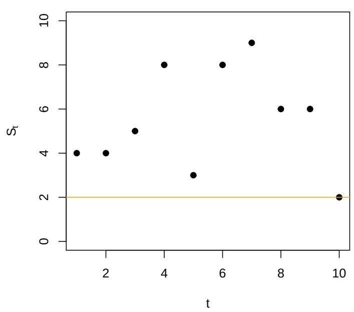
1. Les observations du contrat i sont *conditionnellement* indépendantes et identiquement distribuées avec fonction de répartition $F(x|\theta_i)$.
2. Les variables aléatoires $\Theta_1, \dots, \Theta_I$ sont identiquement distribuées avec fonction de répartition $U(\theta)$.
3. Les contrats sont indépendants.

La troisième hypothèse est assez standard : elle établit que le dossier d'un contrat n'a pas d'influence sur celui des autres contrats.

La portée des deux autres hypothèses est beaucoup plus grande. La première dicte que les sinistres d'un contrat sont indépendants et identiquement distribués seulement lorsque nous connaissons le niveau de risque de ce contrat. Autrement, les sinistres sont dépendants. C'est là un phénomène dit de *contagion apparente*. La figure 3.1 illustre ce phénomène. Dans la figure de gauche, les sinistres d'un contrat, représentés par des points, sont tous supérieurs à la moyenne du groupe, représentée par la ligne horizontale. C'est là une forme de dépendance. Pourtant, lorsque la ligne représente plutôt la moyenne (normalement inconnue) du contrat, nous constatons que les sinistres fluctuent simplement aléatoirement autour de leur moyenne.

Les contrats du portefeuille sont différents en ce qu'ils possèdent chacun leur propre niveau de risque. Cependant, en vertu de la seconde hypothèse, ces niveaux de risque proviennent tous de la même *distribution*. Autrement dit, ils sont suffisamment similaires pour justifier de les regrouper dans un même portefeuille.

Exemple 3.1. Dans la mise en situation de la section 3.1, nous avons deux niveaux de risque : « bon » et « mauvais ». La variable aléatoire Θ prend donc



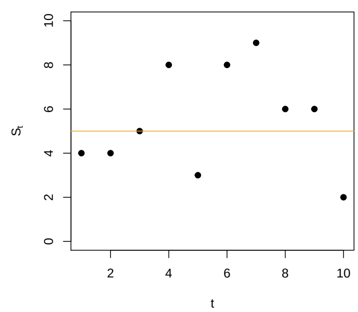


FIG. 3.1 – Illustration du phénomène de contagion apparente. À gauche, la ligne horizontale représente la moyenne du groupe. À droite, la ligne représente le niveau de risque du contrat.

deux valeurs : $\theta = \frac{1}{15}$ et $\theta = \frac{1}{10}$. Sa fonction de répartition et sa fonction de masse de probabilité sont, dans l'ordre :

$$U(\theta) = \begin{cases} 0, & \theta < \frac{1}{15} \\ \frac{3}{4}, & \frac{1}{15} \leq \theta < \frac{1}{10} \\ 1, & \theta \geq \frac{1}{10} \end{cases}$$

et

$$\Pr[\Theta = \theta] = \begin{cases} \frac{3}{4}, & \theta = \frac{1}{15} \\ \frac{1}{4}, & \theta = \frac{1}{10}. \end{cases}$$

□

En théorie des probabilités, notre modèle d'hétérogénéité est un modèle dit « urne d'urne » en ce sens que les montants de sinistres proviennent d'un processus en deux étapes ; voir la figure 3.2 pour une illustration.

Un modèle où la distribution d'une variable aléatoire S dépend d'un paramètre θ et que celui-ci est considéré comme une réalisation d'une variable aléatoire Θ est aussi connu sous le nom de *mélange continu*. La fonction $f(x|\theta)$ est la fonction de vraisemblance et θ est le paramètre de mélange.

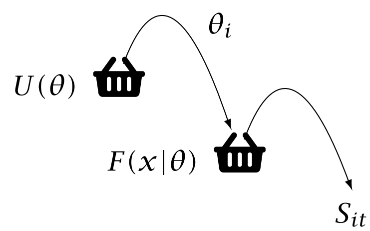


FIG. 3.2 – Modèle urne d'urne. Pour obtenir un montant de sinistre S_{it}, il faut d'abord choisir un niveau de risque θ_i de la distribution $U(\theta)$ (première urne), puis générer un montant de sinistre de la distribution $F(x|\theta_i)$ (deuxième urne).

3.3 Prévision

Tel que mentionné en introduction du chapitre, notre but en crédibilité de précision consiste à calculer la « meilleure » prime (pure) pour chacun des contrats de notre regroupement. D'un point de vue actuariel, cette prime est le montant espéré des sinistres d'un contrat quelconque, éventuellement en tenant compte des observations des années précédentes. D'un point de vue statistique, la prime peut aussi s'interpréter comme la *prévision* du montant total des sinistres de la prochaine année.

Dans la suite, nous considérerons que la meilleure prime (ou prévision) est celle qui minimise l'erreur quadratique moyenne.

Nous définissons trois primes : la prime de risque, la prime collective et la prime bayésienne. Elles diffèrent par la quantité d'information dont nous disposons sur les contrats.

3.3.1 Prime de risque

Si le niveau de risque du contrat i est connu, alors la meilleure prévision de ses sinistres futurs est l'espérance

$$\mu(\theta_i) = E[S_{it}|\Theta_i = \theta_i] = \int_0^\infty x f(x|\theta_i)\, dx. \qquad (3.1)$$

Cette fonction est appelée la *prime de risque*.

La prime de risque serait la prime idéale à charger. Cependant, le niveau de risque θ_i et, donc, la prime de risque sont inconnus. Dès lors, prévoir le niveau des sinistres de la prochaine période, $S_{i,n+1}$, et prévoir la prime de risque $\mu(\theta_i)$ deviennent des problèmes équivalents.

3.3.2 Prime collective

Nous cherchons à estimer — ou prévoir — la prime de risque. Comme première approximation, nous pouvons utiliser la moyenne pondérée de toutes les primes de risque :

$$m = E[\mu(\Theta_i)] = \int_{-\infty}^{\infty} \mu(\theta)u(\theta)\,d\theta. \tag{3.2}$$

Cette approximation sera la même pour tous les contrats ; c'est la *prime collective*.

Vous remarquerez que

$$m = E[\mu(\Theta_i)] = E[E[S_{it}|\Theta_i]] = E[S_{it}].$$

La prime collective est donc aussi égale au montant espéré des sinistres dans le portefeuille.

3.3.3 Prime bayésienne

Comme nous l'avons vu au chapitre 1, la prime collective ci-dessus est globalement adéquate, mais elle n'est pas nécessairement équitable ou *optimale*. Nous ne pouvons établir ce fait qu'à partir du moment où des données de sinistres deviennent disponibles. En termes statistiques, cela signifie qu'il existe une meilleure approximation des primes de risque dans ces circonstances.

La meilleure approximation (ou estimation, ou prévision) de la prime de risque $\mu(\theta_i)$ est la fonction $g^*(x_{i1}, \dots, x_{in})$ des observations $x_{i1}, \dots, x_{in}$ du contrat i qui minimise l'erreur quadratique moyenne

$$E[(\mu(\Theta_i) - g(x_{i1}, \dots, x_{in}))^2],$$

où $g()$ est une fonction quelconque.

Vous trouverez dans tout livre de statistique mathématique la démonstration à l'effet que

$$g^*(x_{i1}, \dots, x_{in}) = E[\mu(\Theta_i)|x_{i1}, \dots, x_{in}].$$

Nous nommerons cette espérance la *prime bayésienne* :

$$B_{i,n+1} = E[\mu(\Theta_i)|S_{i1} = x_{i1}, \dots, S_{in} = x_{in}] = \int_{-\infty}^{\infty} \mu(\theta)u(\theta|x_{i1}, \dots, x_{in})\,d\theta. \tag{3.3}$$

La prime bayésienne est la meilleure prévision de $S_{i,n+1}$ que l'on puisse calculer.

Ci-dessus, $u(\theta|x_{i1},\dots,x_{in})$ est la distribution à postériori des niveaux de risque. En d'autres termes, $U(\theta|x_{i1},\dots,x_{in})$ est la fonction de structure révisée après l'observation de l'expérience $S_{i1} = x_{i1},\dots,S_{in} = x_{in}$. Or, par la règle de Bayes et étant donné l'indépendance conditionnelle des observations,

$$\begin{aligned} u(\theta_i|x_{i1},\dots,x_{in}) &= \frac{f(x_{i1},\dots,x_{in}|\theta_i)u(\theta_i)}{\int_{-\infty}^{\infty} f(x_{i1},\dots,x_{in}|\theta)u(\theta)\,d\theta} \\ &= \frac{\prod_{t=1}^{n} f(x_{it}|\theta_i)u(\theta_i)}{\int_{-\infty}^{\infty}\prod_{t=1}^{n} f(x_{it}|\theta)u(\theta)\,d\theta} \\ &\propto u(\theta_i)\prod_{t=1}^{n} f(x_{it}|\theta_i). \end{aligned}$$

Comme la prime collective, la prime bayésienne est une moyenne pondérée des primes de risque, sauf que cette dernière utilise la distribution à postériori de Θ plutôt que la distribution à priori. En effet, comparez les deux équations ci-dessous :

$$\begin{aligned} m &= \int_{-\infty}^{\infty} \mu(\theta)u(\theta)\,d\theta \\ B_{i,n+1} &= \int_{-\infty}^{\infty} \mu(\theta)u(\theta|x_{i1},\dots,x_{in})\,d\theta. \end{aligned}$$

À l'inverse, nous pouvons interpréter la prime collective comme la prime bayésienne de première année, lorsque aucune expérience n'est disponible.

Puisque les montants de sinistres apparaissent seulement sous forme de produit dans la fonction de vraisemblance

$$f(x_{i1},\dots,x_{in}|\theta_i) = \prod_{t=1}^{n} f(x_{it}|\theta_i),$$

leur ordre ne joue aucun rôle dans les calculs.

Dans la suite, nous n'allons considérer qu'un seul contrat et, par conséquent, nous laisserons tomber l'indice i dans la notation. En effet, dans l'approche entièrement paramétrique adoptée dans ce chapitre, le portefeuille de contrats ne joue aucun rôle concret.

Exemple 3.2. Soit $S_t|\Theta = \theta \sim \text{Poisson}(\theta)$ et

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}3, & \theta = \frac{1}{2} \\ 0{,}5, & \theta = 1 \\ 0{,}2, & \theta = 2. \end{cases}$$

Nous allons calculer les trois primes définies précédemment.

a) Puisque $E[S_t|\Theta = \theta] = \theta$, les primes de risque pour chacun des trois niveaux de risque sont :

$$\begin{aligned} \mu(\tfrac{1}{2}) &= \tfrac{1}{2} \\ \mu(1) &= 1 \\ \mu(2) &= 2. \end{aligned}$$

b) La prime collective est la moyenne pondérée des primes de risque :

$$\begin{aligned} m &= E[\mu(\Theta)] \\ &= E[\Theta] \\ &= \frac{1}{2}(0{,}3) + 1(0{,}5) + 2(0{,}2) \\ &= 1{,}05. \end{aligned}$$

Nous pourrions également calculer la prime collective avec la formule $m = E[S_t]$, mais la distribution marginale de S_t est inconnue.

c) Supposons que $S_1 = 2$ et $S_2 = 1$. Pour calculer la prime bayésienne pour la troisième année, nous devons d'abord calculer la distribution à postériori de Θ :

$$\Pr[\Theta = \theta|S_1 = 2, S_2 = 1] = \frac{\Pr[S_1 = 2, S_2 = 1|\Theta = \theta]\Pr[\Theta = \theta]}{\sum_\theta \Pr[S_1 = 2, S_2 = 1|\Theta = \theta]\Pr[\Theta = \theta]}$$

Or,

$$\begin{aligned} \Pr[S_1 = 2, S_2 = 1|\Theta = \theta] &= \Pr[S_1 = 2|\Theta = \theta]\Pr[S_2 = 1|\Theta = \theta] \\ &= \frac{\theta^2 e^{-\theta}}{2!}\frac{\theta e^{-\theta}}{1!} \\ &= \frac{\theta^3 e^{-2\theta}}{2} \end{aligned}$$

et donc

$$\Pr[\Theta = \theta | S_1 = 2, S_2 = 1] = \begin{cases} 0{,}1245, & \theta = \frac{1}{2} \\ 0{,}6109, & \theta = 1 \\ 0{,}2646, & \theta = 2. \end{cases}$$

La prime bayésienne est donc :

$$\begin{aligned} B_3 &= E[\mu(\Theta) | S_1 = 2, S_2 = 1] \\ &= E[\Theta | S_1 = 2, S_2 = 1] \\ &= \frac{1}{2}(0{,}1245) + 1(0{,}6109) + 2(0{,}2646) \\ &= 1{,}20. \end{aligned}$$

Remarquez que, dans le cas présent, la prime bayésienne se trouve entre la prime collective et la moyenne des observations au cours des deux premières années : $1{,}05 = m < B_3 < \bar{S} = 1{,}5$. Ce n'est pas toujours le cas avec la prime bayésienne. □

Exemple 3.3 (Bernoulli/bêta)**.** Examinons de nouveau le portefeuille simplifié de l'exemple 1.1, cette fois dans le cadre théorique que nous venons de mettre en place. Dans cet exemple, un contrat ne peut avoir au maximum qu'un seul sinistre de montant 1 par année. (En d'autres termes, l'expérience consiste en une suite de 1 et de 0 selon qu'il y a eu un sinistre dans une année ou non.) La probabilité d'avoir un sinistre est toutefois inconnue et potentiellement différente pour chaque contrat.

Nous avons $S_t | \Theta = \theta \sim \text{Bernoulli}(\theta)$ et nous supposons que le paramètre θ est une réalisation d'une variable aléatoire avec distribution bêta de paramètres a et b. Ainsi,

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0, 1$$

et

$$u(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, \theta^{a-1}(1-\theta)^{b-1}, \quad 0 < \theta < 1.$$

a) La prime de risque pour un contrat est $E[S_t | \Theta = \theta] = \theta$.

b) La prime collective est, comme toujours, la moyenne des primes de risque :

$$\begin{aligned} m &= E[\mu(\Theta)] \\ &= E[\Theta] \\ &= \frac{a}{a+b}. \end{aligned}$$

c) Nous souhaitons maintenant calculer la prime bayésienne après n années d'observations $x_1, \dots, x_n$. Pour ce faire, nous déterminons d'abord la distribution à postériori de Θ. À une constante de proportionnalité, celle-ci est :

$$\begin{aligned}
u(\theta|x_1, \dots, x_n) &\propto u(\theta_i) \prod_{t=1}^{n} f(x_t|\theta_i) \\
&\propto \theta^{a-1}(1-\theta)^{b-1} \prod_{t=1}^{n} \theta^{x_t}(1-\theta)^{1-x_t} \\
&= \theta^{a+\sum_{t=1}^{n} x_t - 1}(1-\theta)^{b+n-\sum_{t=1}^{n} x_t - 1}.
\end{aligned}$$

Nous reconnaissons là une distribution bêta, mais avec de nouveaux paramètres $\tilde{a} = a + \sum_{t=1}^{n} x_t$ et $\tilde{b} = b + n - \sum_{t=1}^{n} x_t$. Par conséquent, la prime bayésienne pour l'année $n+1$ est

$$\begin{aligned}
B_{n+1} &= E[\mu(\Theta)|S_1, \dots, S_n] \\
&= E[\Theta|S_1, \dots, S_n] \\
&= \frac{\tilde{a}}{\tilde{a} + \tilde{b}} \\
&= \frac{a + \sum_{t=1}^{n} S_t}{a + b + n}.
\end{aligned}$$

□

Exemple 3.4. Reprenons le modèle de l'exemple 3.3, soit une Bernoulli de paramètre θ pour la distribution de la variable aléatoire $S_t|\Theta = \theta$, mais en changeant la distribution de la variable aléatoire Θ pour une uniforme sur l'intervalle (a, b).

Soit $n\bar{S} = \sum_{t=1}^{n} S_t$ le montant total des sinistres d'un contrat après n années. Norberg (1979) démontre que la prime bayésienne avec ce modèle est

$$B_{n+1} = \frac{\sum_{j=1}^{n-n\bar{S}} (-1)^j \frac{b^{n\bar{S}+j+2} - a^{n\bar{S}+j+2}}{(n-n\bar{S}-j)!\, j!\, (n\bar{S}+j+2)}}{\sum_{j=1}^{n-n\bar{S}} (-1)^j \frac{b^{n\bar{S}+j+1} - a^{n\bar{S}+j+1}}{(n-n\bar{S}-j)!\, j!\, (n\bar{S}+j+1)}}.$$

Cette prime bayésienne n'est pas linéaire et, de plus, elle ne se trouve pas nécessairement entre l'expérience individuelle $\bar{S}$ et la prime collective m. □

3.4 Approche par la distribution prédictive

Cette section introduit une manière alternative de calculer la prime bayésienne. Nous avons déjà effleuré l'idée à la section 3.3.2 lorsque nous avons présenté la prime collective à la fois comme la moyenne des primes de risque ou comme le montant espéré des sinistres :

$$m = E[\mu(\Theta)] = E[S_{it}].$$

De manière similaire, nous pouvons démontrer que

$$B_{i,n+1} = E[\mu(\Theta)|S_{i1} = x_1, \dots, S_{in} = x_n] = E[S_{i,n+1}|S_{i1} = x_1, \dots, S_{in} = x_n]. \tag{3.4}$$

La distribution de la variable aléatoire $S_{n+1}|S_1 = x_1, \dots, S_n = x_1$ avec fonction de densité de probabilité $f(x|x_1, \dots, x_n)$ est appelée la *distribution prédictive* de la variable aléatoire S_{n+1}.

Théorème 3.1. *La fonction de densité de probabilité de la distribution prédictive de S_{n+1} est*

$$f(x|x_1, \dots, x_n) = \int_{-\infty}^{\infty} f(x|\theta)u(\theta|x_1, \dots, x_n)\, d\theta.$$

Démonstration. Par la règle de Bayes et par l'indépendance conditionnelle des montants de sinistres, nous obtenons :

$$\begin{aligned}
f(x|x_1, \dots, x_n) &= \frac{f(x, x_1, \dots, x_n)}{f(x_1, \dots, x_n)} \\
&= \frac{\int_{-\infty}^{\infty} f(x, x_1, \dots, x_n|\theta)u(\theta)\, d\theta}{\int_{-\infty}^{\infty} f(x_1, \dots, x_n|\theta)u(\theta)\, d\theta} \\
&= \int_{-\infty}^{\infty} f(x|\theta)\, \frac{f(x_1, \dots, x_n|\theta)u(\theta)}{\int_{-\infty}^{\infty} f(x_1, \dots, x_n|\theta)u(\theta)\, d\theta}\, d\theta \\
&= \int_{-\infty}^{\infty} f(x|\theta)u(\theta|x_1, \dots, x_n)\, d\theta.
\end{aligned}$$

□

Puisque la fonction de densité de probabilité de la distribution marginale de la variable aléatoire S_t est

$$f(x) = \int_{-\infty}^{\infty} f(x|\theta)u(\theta)\, d\theta,$$

alors la seule différence entre l'expression de $f(x)$ et celle de $f(x|x_1, \dots, x_n)$ réside dans l'utilisation de la distribution à priori de Θ pour la première et de la distribution à postériori pour la seconde.

> La conséquence de la remarque précédente, c'est que si la distribution à posteriori de Θ est du même type que la distribution à priori, alors la distribution prédictive est du même type que la distribution marginale.

Nous avons posé la relation (3.4) de manière quelque peu intuitive. Démontrons maintenant qu'elle est vraie. Nous savons déjà que

$$E[\mu(\Theta)|S_{i1} = x_1, \dots, S_{in} = x_n] = \int_{-\infty}^{\infty} \mu(\theta) u(\theta|x_1, \dots, x_n)\, d\theta$$

et

$$\mu(\theta) = \int_0^{\infty} x f(x|\theta)\, dx.$$

Par conséquent,

$$\begin{aligned} E[\mu(\Theta)|S_{i1} = x_1, \dots, S_{in} = x_n] &= \int_{-\infty}^{\infty} \int_0^{\infty} x f(x|\theta) u(\theta|x_1, \dots, x_n)\, dx\, d\theta \\ &= \int_0^{\infty} x \left(\int_{-\infty}^{\infty} f(x|\theta) u(\theta|x_1, \dots, x_n)\, d\theta \right) dx \end{aligned}$$

d'où, par le théorème 3.1,

$$\begin{aligned} E[\mu(\Theta)|S_{i1} = x_1, \dots, S_{in} = x_n] &= \int_0^{\infty} x f(x|x_1, \dots, x_n)\, dx \\ &= E[S_{i,n+1}|S_{i1} = x_1, \dots, S_{in} = x_n]. \end{aligned}$$

Avec cette approche, la prime collective et la prime bayésienne s'interprètent toutes deux comme le montant moyen des sinistres dans le portefeuille, mais avec des pondérations différentes.

Exemple 3.5 (Poisson/gamma)**.** Le présent exemple joue un rôle important dans la suite. Soit

$$\begin{aligned} S_t|\Theta = \theta &\sim \text{Poisson}(\theta) \\ \Theta &\sim \text{Gamma}(\alpha, \lambda), \end{aligned}$$

c'est-à-dire

$$f(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \dots$$
$$u(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\, \theta^{\alpha-1} e^{-\lambda\theta}, \quad \theta > 0.$$

Comme dans les exemples 3.2 et 3.3, nous allons calculer les primes de risque, collective et bayésienne. Nous disposons toutefois maintenant de l'approche par la distribution prédictive pour effectuer nos calculs.

a) La prime de risque est

$$\mu(\theta) = E[S_t|\Theta = \theta] = \theta.$$

Dans la suite, nous aurons également besoin de la variance conditionnelle

$$\sigma^2(\theta) = \mathrm{Var}[S_t|\Theta = \theta] = \theta.$$

b) La prime collective est

$$m = E[\mu(\Theta)] = E[\Theta] = \frac{\alpha}{\lambda}.$$

c) Calculons maintenant la prime bayésienne à partir de la distribution à postériori de Θ. Tout d'abord, nous avons que

$$\begin{aligned} u(\theta|x_1, \dots, x_n) &\propto u(\theta) \prod_{t=1}^{n} f(x_t|\theta) \\ &\propto \theta^{\alpha-1} e^{-\lambda\theta} \prod_{t=1}^{n} \theta^{x_t} e^{-\theta} \\ &= \theta^{\alpha+\sum_{t=1}^{n} x_t - 1} e^{-(\lambda+n)\theta}, \end{aligned}$$

d'où la distribution à postériori de Θ est aussi une distribution gamma, mais avec de nouveaux paramètres :

$$\Theta|S_1, \dots, S_n \sim \text{Gamma}\left(\tilde{\alpha} = \alpha + \sum_{t=1}^{n} S_t, \tilde{\lambda} = \lambda + n\right).$$

Par conséquent, la prime bayésienne est

$$\begin{aligned} B_{n+1} &= E[\mu(\Theta)|S_1, \dots, S_n] \\ &= E[\Theta|S_1, \dots, S_n] \\ &= \frac{\tilde{\alpha}}{\tilde{\lambda}} \\ &= \frac{\alpha + \sum_{t=1}^{n} S_t}{\lambda + n} \end{aligned}$$

d) Avant d'examiner la distribution prédictive, calculons la distribution marginale de S_t. Nous avons

$$\begin{aligned}
f(x) &= \int_0^\infty f(x|\theta)u(\theta)\, d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)x!}\int_0^\infty \theta^x e^{-\theta}\theta^{\alpha-1}e^{-\lambda\theta}\, d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)\Gamma(x+1)}\int_0^\infty \theta^{\alpha+x-1}e^{-(\lambda+1)\theta}\, d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)\Gamma(x+1)}\frac{\Gamma(\alpha+x)}{(\lambda+1)^{\alpha+x}} \\
&= \frac{\Gamma(\alpha+x)}{\Gamma(\alpha)\Gamma(x+1)}\left(\frac{\lambda}{\lambda+1}\right)^\alpha \left(\frac{1}{\lambda+1}\right)^x \\
&= \binom{\alpha+x-1}{\alpha-1}\left(\frac{\lambda}{\lambda+1}\right)^\alpha \left(\frac{1}{\lambda+1}\right)^x \quad x = 0, 1, \dots .
\end{aligned}$$

Nous identifions cette fonction de masse de probabilité comme celle d'une binomiale négative de paramètres $r = \alpha$ et $p = \lambda/(\lambda+1)$ (voir la section C.1.2).

Nous pouvons obtenir le même résultat par les fonctions génératrices des moments. En effet,

$$\begin{aligned}
M_S(t) &= E[e^{tS}] \\
&= E[E[e^{tS}|\Theta]] \\
&= E[M_{S|\Theta}(t)] \\
&= E[e^{\Theta(e^t-1)}] \\
&= M_\Theta(e^t-1) \\
&= \left(\frac{\lambda}{\lambda+1-e^t}\right)^\alpha \\
&= \left(\frac{\lambda/(\lambda+1)}{1-e^t/(\lambda+1)}\right)^\alpha,
\end{aligned}$$

ce qui est la fonction génératrice des moments d'une binomiale négative de paramètres α et $\lambda/(\lambda+1)$.

Ce lien entre les lois de Poisson, gamma et binomiale négative constitue un résultat classique en théorie des probabilités. Répétons-le pour le mettre en exergue : si

$$\begin{aligned}
S|\Theta = \theta &\sim \text{Poisson}(\theta) \\
\Theta &\sim \text{Gamma}(\alpha, \lambda)
\end{aligned}$$

alors
$$S \sim \text{Binomiale négative}\left(\alpha, \frac{\lambda}{\lambda+1}\right).$$

Des résultats ci-dessus découle que la prime collective est
$$m = E[S_t] = \frac{\alpha(1-\lambda/(\lambda+1))}{\lambda/(\lambda+1)} = \frac{\alpha}{\lambda},$$
soit le même résultat que précédemment.

e) Les distributions à priori et à postériori de Θ sont toutes deux des distributions gamma, la seconde avec des paramètres mis à jour suite à l'ajout de données :
$$\Theta \sim \text{Gamma}(\alpha, \lambda)$$
et
$$\Theta | S_1 = x_1, \dots, S_n = x_n \sim \text{Gamma}(\tilde{\alpha}, \tilde{\lambda}).$$
C'est le contexte de la remarque de la page 43. Nous pouvons directement établir que la distribution prédictive de S_{n+1} est :
$$S_{n+1} | S_1 = x_1, \dots, S_n = x_n \sim \text{Binomiale négative}(\tilde{r}, \tilde{p})$$
avec
$$\begin{aligned}\tilde{r} &= \tilde{\alpha} = \alpha + \sum_{t=1}^{n} x_t \\ \tilde{p} &= \frac{\tilde{\lambda}}{\tilde{\lambda}+1} = \frac{\lambda+n}{\lambda+n+1}.\end{aligned}$$
Par conséquent, la prime bayésienne, calculée cette fois à partir de la distribution prédictive, est :
$$\begin{aligned}B_{n+1} &= E[S_{n+1} | S_1, \dots, S_n] \\ &= \frac{\tilde{\alpha}}{\tilde{\lambda}} \\ &= \frac{\alpha + \sum_{t=1}^{n} S_t}{\lambda + n}.\end{aligned}$$

f) Afin de fournir un autre exemple de calculs équivalents, mais qui empruntent un chemin différent, calculons $\text{Var}[S]$. En premier lieu, nous pouvons procéder en conditionnant sur Θ :

$$\begin{aligned}\mathrm{Var}[S] &= E[\mathrm{Var}[S|\Theta]] + \mathrm{Var}[E[S|\Theta]] \\ &= E[\sigma^2(\Theta)] + \mathrm{Var}[\mu(\Theta)] \\ &= E[\Theta] + \mathrm{Var}[\Theta] \\ &= \frac{\alpha}{\lambda} + \frac{\alpha}{\lambda^2} \\ &= \frac{\alpha(\lambda+1)}{\lambda^2}.\end{aligned}$$

En second lieu, nous pouvons calculer la variance directement depuis la marginale :

$$\begin{aligned}\mathrm{Var}[S] &= \frac{r(1-p)}{p^2} \\ &= \frac{\alpha(\lambda+1)}{\lambda^2}.\end{aligned}$$

g) Le tableau 3.1 contient les résultats de la prime bayésienne des dix premières années si la distribution à priori de la variable aléatoire Θ est une gamma de paramètres $\alpha = 3$ et $\lambda = 3$, et que les montants de sinistres au cours de ces années sont les suivants : $5, 3, 0, 1, 1, 2, 0, 2, 0, 2$.

La figure 3.3 illustre que la distribution à postériori de Θ est de plus en plus concentrée autour de sa moyenne au fur et à mesure que l'expérience s'accumule. En d'autres termes, l'incertitude sur le niveau de risque du contrat sous étude diminue. Du même coup, la précision de la prime bayésienne s'améliore. En effet, la vraie valeur de θ est utilisée pour générer les sinistres ci-dessus est 1,48. □

3.5 Crédibilité bayésienne linéaire

Examinons de plus près la prime bayésienne obtenue à l'exemple 3.3 avec une fonction de vraisemblance Bernoulli et une distribution du paramètre de mélange bêta. Nous avons obtenu

$$B_{n+1} = \frac{\alpha + \sum_{t=1}^{n} S_t}{\alpha + \beta + n}.$$

Or, cette expression se réécrit ainsi :

$$\begin{aligned}B_{n+1} &= \frac{n}{n+\alpha+\beta}\bar{S} + \frac{\alpha+\beta}{n+\alpha+\beta}\frac{\alpha}{\alpha+\beta} \\ &= z\bar{S} + (1-z)m,\end{aligned}$$

Tab. 3.1 – Primes bayésiennes de l'exemple 3.5

n	x_n	$\sum x_t$	$\alpha + \sum x_t$	$\lambda + n$	B_{n+1}
0	–	–	3	3	1,0
1	5	5	8	4	2,0
2	3	8	11	5	2,2
3	0	8	11	6	1,83
4	1	9	12	7	1,71
5	1	10	13	8	1,625
6	2	12	15	9	1,667
7	0	12	15	10	1,5
8	2	14	17	11	1,54
9	0	14	17	12	1,42
10	2	16	19	13	1,46

où $\bar{S} = n^{-1} \sum_{t=1}^{n} S_t$, m est la prime collective et

$$z = \frac{n}{n+K}, \quad K = \alpha + \beta.$$

Nous pouvons répéter le même exercice avec les résultats de l'exemple 3.5 pour la combinaison d'une loi de Poisson et d'une loi gamma. Dans ce cas,

$$B_{n+1} = \frac{\alpha + \sum_{t=1}^{n} S_t}{\lambda + n},$$

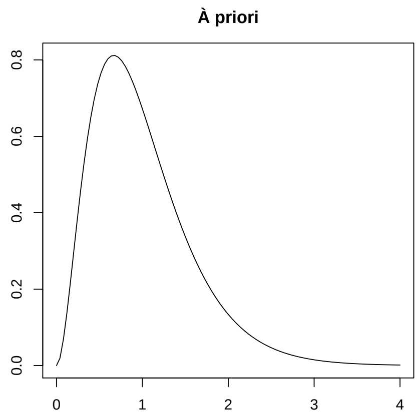
soit

$$\begin{aligned} B_{n+1} &= \frac{n}{n+\lambda}\bar{S} + \frac{\lambda}{n+\lambda}\frac{\alpha}{\lambda} \\ &= z\bar{S} + (1-z)m, \end{aligned}$$

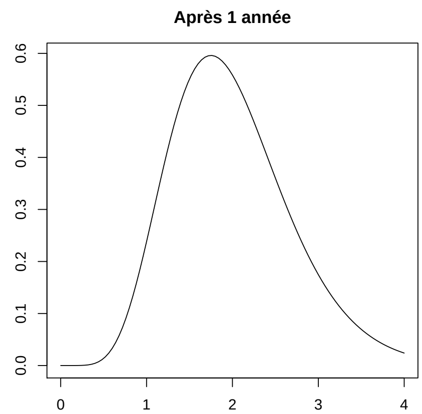
avec

$$z = \frac{n}{n+K}, \quad K = \lambda.$$

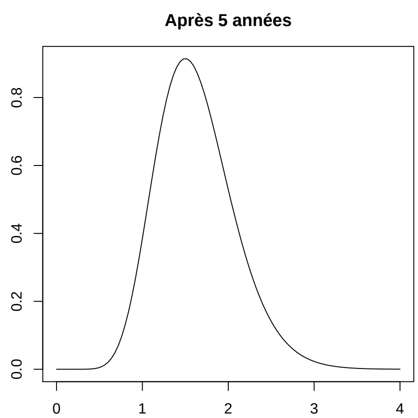
Ces résultats sont très séduisants sous plusieurs aspects. Premièrement, la prime bayésienne est linéaire, ce qui est à la fois simple à calculer et facile à interpréter — rappelez-vous à cet effet le cas tout à fait contraire de l'exemple 3.4. Ensuite, puisque $0 < z < 1$ pour $n > 0$, la prime bayésienne se trouve être une combinaison convexe de $\bar{S}$ et m. Elle se situe donc toujours entre l'expérience individuelle d'un contrat ($\bar{S}$) et la prime collective (m). D'un point de vue actuariel, il s'agit d'une propriété importante. Imaginez seulement, si ce n'était pas le cas, devoir expliquer à un client dont



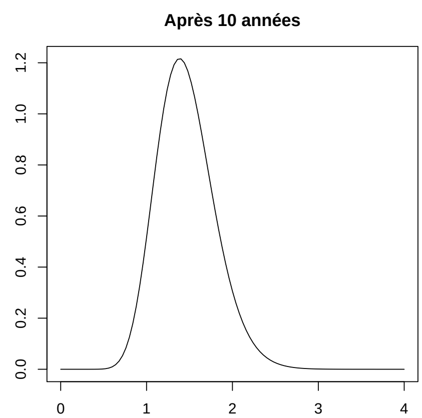
(a) Gamma avec $\alpha = 3$ et $\lambda = 3$. L'espérance est égale à 1.



(b) Gamma avec $\alpha = 8$ et $\lambda = 4$. L'espérance est égale à 2.



(c) Gamma avec $\alpha = 13$ et $\lambda = 8$. L'espérance est égale à 1,625.



(d) Gamma avec $\alpha = 19$ et $\lambda = 13$. L'espérance est égale à 1,46.

FIG. 3.3 – Distributions à priori et à postériori de la variable aléatoire Θ dans l'exemple 3.5

l'expérience est pire que la moyenne que vous devez non seulement lui charger plus que la prime collective, mais aussi plus que sa propre expérience !

Les primes de forme

$$\pi_{n+1} = z\bar{S} + (1 - z)m,$$

où $0 \leq z \leq 1$, jouent un rôle fondamental en théorie de la crédibilité. Nous les appelons *primes de crédibilité* et z est le *facteur de crédibilité*.

Whitney (1918) et Bailey (1950) furent les premiers à démontrer que la prime bayésienne est une prime de crédibilité pour certaines combinaisons de distributions. Nous en avons déjà identifié deux. Il y a en fait cinq combinaisons de distributions qui résultent en une prime bayésienne linéaire (en plus de leurs convolutions) :

1. $S|\Theta = \theta \sim \text{Poisson}(\theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$;
2. $S|\Theta = \theta \sim \text{Exponentielle}(\theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$;
3. $S|\Theta = \theta \sim \text{Normale}(\theta, \sigma_2^2)$ et $\Theta \sim \text{Normale}(\mu, \sigma_1^2)$;
4. $S|\Theta = \theta \sim \text{Bernoulli}(\theta)$ et $\Theta \sim \text{Bêta}(a, b)$;
5. $S|\Theta = \theta \sim \text{Géométrique}(\theta)$ et $\Theta \sim \text{Bêta}(a, b)$.

Nous effectuons les calculs de prime pour les cas exponentielle/gamma et normale/normale dans les deux exemples suivants. Vous devez faire de même pour le cas géométrique/bêta à l'exercice 3.18.

Exemple 3.6 (Exponentielle/gamma)**.** Soit

$$\begin{aligned} S_t|\Theta = \theta &\sim \text{Exponentielle}(\theta) \\ \Theta &\sim \text{Gamma}(\alpha, \lambda). \end{aligned}$$

La prime de risque est

$$\mu(\theta) = \frac{1}{\theta}$$

et la variance conditionnelle est

$$\sigma^2(\theta) = \frac{1}{\theta^2}.$$

La prime collective est

$$\begin{aligned} m &= E\left[\frac{1}{\Theta}\right] \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \theta^{\alpha-1-1} e^{-\lambda\theta}\, d\theta \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha-1)}{\lambda^{\alpha-1}} \\ &= \frac{\lambda}{\alpha-1}. \end{aligned}$$

La distribution à postériori de Θ est

$$\begin{aligned} u(\theta|x_1,\dots,x_n) &\propto \theta^{\alpha-1}e^{-\lambda\theta}\prod_{t=1}^{n}\theta e^{-\theta x_t} \\ &= \theta^{\alpha+n-1}e^{-(\lambda+\sum_{t=1}^{n}x_t)\theta}, \end{aligned}$$

d'où $\Theta|S_1,\dots,S_n \sim \text{Gamma}(\tilde{\alpha}=\alpha+n, \tilde{\lambda}=\lambda+\sum_{t=1}^{n}S_t)$.

Pour les distributions marginale et prédictive, nous avons

$$\begin{aligned} f(x) &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty \theta e^{-\theta x}\theta^{\alpha-1}e^{-\lambda\theta}\,d\theta \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\int_0^\infty \theta^{\alpha+1-1}e^{-(\lambda+x)\theta}\,d\theta \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\frac{\Gamma(\alpha+1)}{(\lambda+x)^{\alpha+1}} \\ &= \frac{\alpha\lambda^\alpha}{(\lambda+x)^{\alpha+1}}, \quad x>0, \end{aligned}$$

d'où $S_t \sim \text{Pareto}(\alpha,\lambda)$ et

$$S_{n+1}|S_1=x_1,\dots,S_n=x_n \sim \text{Pareto}(\tilde{\alpha}=\alpha+n, \tilde{\lambda}=\lambda+\textstyle\sum_{t=1}^{n}x_t).$$

La prime bayésienne est donc

$$\begin{aligned} B_{n+1} &= E[\Theta^{-1}|S_1,\dots,S_n] \\ &= E[S_{n+1}|S_1,\dots,S_n] \\ &= \frac{\tilde{\lambda}}{\tilde{\alpha}-1} \\ &= \frac{\lambda+\sum_{t=1}^{n}S_t}{\alpha+n-1} \\ &= \frac{n}{n+\alpha-1}\bar{S} + \frac{\alpha-1}{n+\alpha-1}\frac{\lambda}{\alpha-1} \\ &= z\bar{S}+(1-z)m \end{aligned}$$

avec

$$z = \frac{n}{n+\alpha-1}.$$

La prime bayésienne est donc linéaire. □

Exemple 3.7 (Normale/normale)**.** Soit

$$\begin{aligned} S_t|\Theta &\sim \text{Normale}(\Theta, \sigma_2^2) \\ \Theta &\sim \text{Normale}(\mu, \sigma_1^2). \end{aligned}$$

La prime de risque, la variance conditionnelle et la prime collective sont :

$$\begin{aligned} \mu(\Theta) &= \Theta \\ \sigma^2(\Theta) &= \sigma_2^2 \quad \text{(constante)} \end{aligned}$$

et

$$m = E[\Theta] = \mu.$$

Trouver la distribution à postériori n'est toutefois pas une sinécure. Tout d'abord,

$$u(\theta|x_1, \dots, x_n) \propto e^{-(\theta-\mu)^2/2\sigma_1^2} e^{-\sum_{t=1}^n (x_t-\theta)^2/2\sigma_2^2}.$$

En développant l'exposant tout en laissant de côté tous les termes non fonction de θ, nous obtenons

$$\begin{aligned} \text{Exposant} &= -\frac{1}{2}\left[\frac{\theta^2 - 2\theta\mu}{\sigma_1^2} + \sum_{t=1}^n \frac{-2\theta x_t + \theta^2}{\sigma_2^2} + \text{cte}\right] \\ &= -\frac{1}{2}\left[\theta^2\left(\frac{1}{\sigma_1^2} + \frac{n}{\sigma_2^2}\right) - \frac{2\theta\mu}{\sigma_1^2} - \frac{-2\theta\sum_{t=1}^n x_t}{\sigma_2^2} + \text{cte}\right] \\ &= -\frac{1}{2}\underbrace{\left(\frac{1}{\sigma_1^2} + \frac{n}{\sigma_2^2}\right)}_{\phi}\left[\theta^2 - \frac{2\theta\mu/\sigma_1^2}{(\frac{1}{\sigma_1^2} + \frac{n}{\sigma_2^2})} - \frac{-2\theta\sum_{t=1}^n x_t/\sigma_2^2}{(\frac{1}{\sigma_1^2} + \frac{n}{\sigma_2^2})} + \text{cte}\right] \\ &= -\frac{1}{2}\phi\left[\theta^2 - 2\theta\left(\frac{\mu/\sigma_1^2 + \sum_{t=1}^n x_t/\sigma_2^2}{\phi}\right) + \text{cte}\right] \\ &= -\frac{1}{2}\frac{\left(\theta - \left(\frac{\mu}{\phi\sigma_1^2} + \frac{\sum_{t=1}^n x_t}{\phi\sigma_2^2}\right)\right)^2}{1/\phi} + \text{cte}. \end{aligned}$$

Par conséquent,

$$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Normale}(\tilde{\mu}, \tilde{\sigma}_1^2)$$

avec

$$\begin{aligned}
\tilde{\sigma}_1^2 &= \frac{1}{\phi} \\
&= \frac{\sigma_1^2\sigma_2^2}{\sigma_2^2 + n\sigma_1^2} \\
&= \frac{\sigma_1^2}{1 + n\sigma_1^2/\sigma_2^2} < \sigma_1^2
\end{aligned}$$

et

$$\begin{aligned}
\tilde{\mu} &= \frac{\mu}{\phi\sigma_1^2} + \frac{\sum_{t=1}^n x_t}{\phi\sigma_2^2} \\
&= \frac{\mu\sigma_2^2 + \sigma_1^2\sum_{t=1}^n x_t}{\sigma_2^2 + n\sigma_1^2}.
\end{aligned}$$

La distribution marginale de S est plus aisée à trouver à l'aide des fonctions génératrice des moments. En effet,

$$\begin{aligned}
M_S(t) &= E[E[e^{tS}|\Theta]] \\
&= E[e^{\Theta t + \sigma_2^2 t^2/2}] \\
&= e^{\sigma_2^2 t^2/2} E\Big[e^{\Theta t}\Big] \\
&= e^{\sigma_2^2 t^2/2} e^{\mu t + \sigma_1^2 t^2/2} \\
&= e^{\mu t + (\sigma_1^2 + \sigma_2^2)t^2/2}
\end{aligned}$$

et donc

$$S \sim \text{Normale}(\mu, \sigma_1^2 + \sigma_2^2).$$

Puisque la distribution à postériori de Θ est du même type que la distribution à priori, nous pouvons immédiatement conclure que

$$S_{n+1}|S_1, \dots, S_n \sim \text{Normale}(\tilde{\mu}, \tilde{\sigma}_1^2 + \tilde{\sigma}_2^2).$$

La prime bayésienne est donc, en utilisant indifféremment l'approche de la distribution à postériori ou celle de la prédictive,

$$\begin{aligned}
B_{n+1} &= \tilde{\mu} \\
&= \frac{n\sigma_1^2}{n\sigma_1^2 + \sigma_2^2}\bar{S} + \frac{\sigma_2^2}{n\sigma_1^2 + \sigma_2^2}\mu \\
&= z\bar{S} + (1 - z)m
\end{aligned}$$

où

$$z = \frac{n}{n + \sigma_2^2/\sigma_1^2}.$$

□

Les formules de crédibilité linéaire — ou exacte — pour les combinaisons de distributions issues de la famille exponentielle sont rassemblées à l'annexe B. Les commentaires suivants se rapportent à ces résultats.

1. Dans la combinaison normale/normale, nous avons

$$\tilde{\sigma}_1^2 = \frac{\sigma_1^2}{n\dfrac{\sigma_1^2}{\sigma_2^2} + 1} \leq \sigma_1^2,$$

   avec égalité seulement lorsque $\sigma_1^2 = 0$ (le cas $\sigma_2^2 = \infty$ ne présentant aucun intérêt). Cette inégalité s'interprète comme une baisse de l'incertitude quant au niveau de risque d'un contrat au fur et à mesure que l'expérience s'accumule.
2. La combinaison normale/normale est celle considérée par Whitney (1918), mais aussi celle pour laquelle les formules sont les plus compliquées. Cela explique sans doute en partie pourquoi l'auteur a recommandé de fixer la constante K au jugement.
3. Dans tous les cas, $z \to 1$ lorsque $n \to \infty$. Le poids accordé à la prime individuelle d'un contrat va donc croissant avec le nombre d'années d'expérience disponible.
4. De plus, $z = n/(n+K) \to 1$ lorsque $K \to 0$ et $z \to 0$ lorsque $K \to \infty$. Dans la combinaison Poisson/gamma, où $K = \lambda$, une petite valeur de λ correspond à une grande incertitude quant au niveau de risque θ (la courbe gamma sera très évasée, voir la figure 3.4). L'assureur accorde donc peu de poids à la prime collective, d'où un grand facteur de crédibilité.
5. Dans la combinaison normale/normale, K est grand si σ_2^2 est également grand ou alors si σ_1^2 est petit. Respectivement, cela signifie que ou bien l'expérience est potentiellement si volatile que l'on ne peut s'y fier, ou bien que le niveau de risque θ est presque connu avec certitude. Dans un cas comme dans l'autre, il convient de charger la prime collective. Nous pouvons répéter une telle analyse pour chacune des autres combinaisons de distributions.
6. À un haut niveau de risque ne correspond pas nécessairement une grande valeur de θ, comme en fait foi la combinaison exponentielle/gamma.

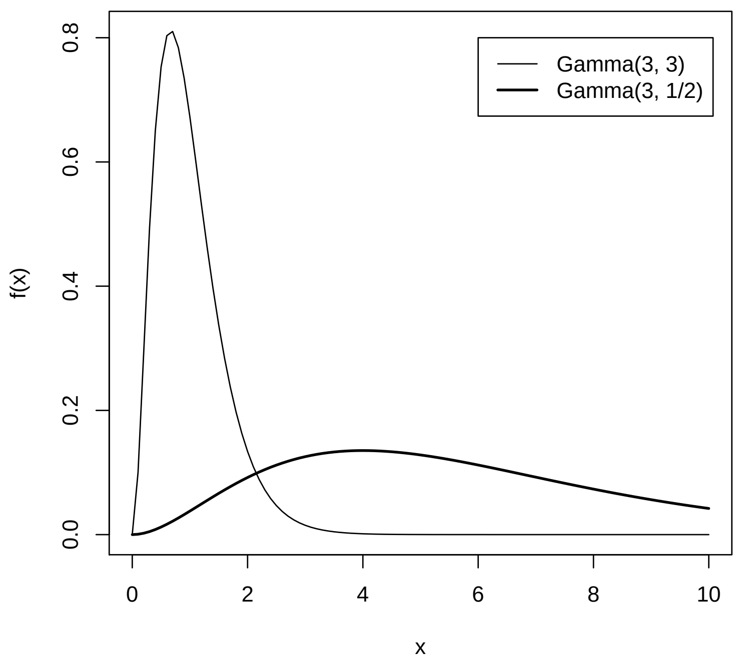


FIG. 3.4 – Distributions gamma avec différents paramètres d'échelle λ.

3.6 Évaluation numérique avec R

Nous abordons ici l'évaluation numérique de primes de crédibilité avec R, un sujet qui reviendra périodiquement dans les chapitres subséquents. Nous effectuerons les calculs de crédibilité avec la fonction `cm` du paquetage **actuar** (Dutang et collab., 2008).

Disponible dans le site Comprehensive R Archive Network (CRAN) depuis 2005 et toujours en développement, **actuar** ajoute au système R de base des fonctionnalités spécifiques pour les applications actuarielles : modélisation des distributions de sinistres ; théorie du risque et de la ruine ; simulation de mélanges et de modèles composés hiérarchiques ; théorie de la crédibilité. Le paquetage fournit également des fonctions pour le traitement de 19 lois de probabilité continues à valeurs extrêmes et les versions zéro tronquée et zéro modifiée des lois discrètes usuelles. Nous vous encourageons vivement à consulter l'exhaustive documentation du paquetage sous forme de *vignettes*. La commande ci-dessous permet d'afficher la liste des vignettes disponibles.

```
> vignette(package = "actuar")
```

La fonction `cm` de **actuar** sert d'interface unique pour l'ajustement de plusieurs modèles de crédibilité de précision : les modèles bayésiens linéaires de la section précédente, les modèles de Bühlmann (1969) et de Bühlmann et Straub (1970), le modèle hiérarchique de Jewell (1975) et le modèle de régression de Hachemeister (1975). L'interface de la fonction est fortement inspirée de `lm`, la fonction d'ajustement de modèles linéaires de R.

Pour la tarification bayésienne pure telle qu'étudiée dans le présent chapitre, la fonction `cm` prend en charge toutes les combinaisons de distributions mentionnées à la section 3.5, les convolutions binomiale/bêta (exercice 3.19), gamma/gamma (exercice 3.20) et binomiale négative/bêta, ainsi que la combinaison quelque peu périphérique Pareto translatée/gamma (exercice 3.32).

Nous expliquons l'utilisation de la fonction `cm` pour le calcul de primes bayésiennes linéaires à même le code informatique de la section 3.8. Étudiez attentivement l'intégralité du code.

3.7 Modèle de Jewell

Le modèle de crédibilité exacte de Jewell unifie les résultats des cinq cas spéciaux étudiés précédemment. Nous n'en traçons que les grandes lignes ici.

- En analyse bayésienne, si $u(\theta|x_1, \dots, x_n)$ appartient à la même famille que $u(\theta)$, on dit de $u(\theta)$ et $f(x|\theta)$ qu'elles sont des *conjuguées naturelles*.
- Les lois de Poisson, exponentielle, normale, Bernoulli et géométrique appartiennent toutes à la *famille exponentielle univariée*, c'est-à-dire que leur fonction de densité (ou de masse) de probabilité peut s'écrire sous la forme

$$f(x|\theta) = \frac{p(x)e^{-\theta x}}{q(\theta)}.$$

- Jewell (1974) démontre que lorsqu'une fonction de vraisemblance de la famille exponentielle est combinée avec sa conjuguée naturelle, alors la prime bayésienne est toujours une prime de crédibilité exacte.

- Goel (1982) conjecture que ceci n'arrive qu'avec les membres de la famille exponentielle, c'est-à-dire qu'il ne peut le prouver, mais qu'il ne peut non plus fournir de contre-exemple.

3.8 Code informatique

Fichier d'accompagnement bayesienne.R

```
## Chargement du paquetage actuar. La prise en charge des
## modèles bayésiens linéaires requiert une version du
## paquetage supérieure ou égale à 2.3.0.
if (packageVersion("actuar") >= "2.3.0")
    library(actuar)

## La fonction 'cm' de actuar constitue l'interface unifiée
## pour les calculs de crédibilité. La fonction ajuste un
## modèle de crédibilité de précision à des données et elle
## retourne un objet à partir duquel il est possible de
## réaliser différentes opérations, notamment le calcul de
## primes de crédibilité avec la fonction 'predict'.
##
## Dans son utilisation pour le calcul de primes bayésiennes
## linéaires, la fonction 'cm' accepte trois arguments
## principaux ainsi qu'un nombre variable d'autres arguments
## selon le modèle bayésien utilisé:
##
## formula: chaine de caractères "bayes";
## data: NULL, un vecteur simple de données, ou une matrice ou
##   un data frame de données (contrats sur les lignes,
##   périodes dans les colonnes);
## likelihood: nom (en minuscules) de la distribution de la
##   variable aléatoire S|Theta = theta;
## ...: valeurs des paramètres fixes des distributions des
##   variables aléatoires S|Theta = theta (s'il y en a) et
##   Theta.
##
## Pour spécifier les paramètres d'une distribution <loi>, il
## s'agit d'utiliser les noms des arguments de la fonction
## 'd<loi>' ('dgamma', 'dbeta', etc.)
##
## Lorsque l'argument 'data' est NULL (ou manquant), 'cm'
## ajuste le modèle à priori. L'ajustement du modèle à
## postériori requiert des données, soit pour un seul contrat
## (un vecteur), soit pour plusieurs contrats (matrice ou data
```

```r
## frame).
##
## Après ce long préambule, allons-y de quelques exemples.

###
### Modèle Bernoulli/bêta
###

## Ajustement du modèle à priori
##
##   S|Theta = theta ~ Bernoulli(theta)
##   Theta ~ Bêta(2, 1).
fit <- cm("bayes", likelihood = "bernoulli",
          shape1 = 2, shape2 = 1)

## Modèle ajusté.
fit

## Prédiction: il n'y a qu'une seule prime à calculer et c'est
## la prime collective.
predict(fit)

## Simulons 10 périodes d'observations de ce modèle pour 5
## contrats différents.
a <- 2                     # paramètre shape1 de la bêta
b <- 1                     # paramètre shape2 de la bêta
n <- 10                    # nombre de périodes
(theta <- rbeta(5, a, b))  # niveaux de risque
(x <- matrix(rbinom(5 * n, 1, theta), 5, n)) # données

## Ajustement des modèles après 10 périodes.
fit <- cm("bayes", data = x, likelihood = "bernoulli",
          shape1 = 2, shape2 = 1)

## Sommaire contenant le modèle et les primes de crédibilité.
summary(fit)

## Vérifions les calculs de primes à partir des résultats de
## l'exemple 3.3.
(a + rowSums(x))/(a + b + n)

## Écarts quadratiques entre les primes bayésiennes et les
## vraies valeurs des primes de risque (qui sont connues,
## ici).
(predict(fit) - theta)^2
```

```
###
### Modèle Poisson/gamma
###

## Modèle de l'exemple 3.5, partie g):
##
##   S|Theta = theta ~ Poisson(theta)
##   Theta ~ Gamma(3, 3).
##
## Calculons la prime de crédibilité après les cinq premières
## années d'observations.
(fit <- cm("bayes", data = c(5, 3, 0, 1, 1),
           likelihood = "poisson", shape = 3, rate = 3))
predict(fit)

## Pour réaliser quelque chose de plus amusant, reproduisons
## le tableau 3.1 des primes bayésiennes pour les 10 premières
## années d'observations.
##
## Tout d'abord, les données.
x <- c(5, 3, 0, 1, 1, 2, 0, 2, 0, 2)
n <- seq(0, length(x))

## Ensuite, le calcul des primes. Nous devons ajuster un
## nouveau modèle après chaque période d'observations.
primes <- sapply(n,
                 function(x, n) predict(cm("bayes", head(x, n),
                                          likelihood = "poisson",
                                          shape = 3, rate = 3)),
                 x = x)

## Enfin, le tableau.
cbind(n = n,                        # nombre de périodes
      x = c(NA, x),                 # observation de la période
      "sum(x)" = c(NA, cumsum(x)), # total des observations
      shape = 3 + cumsum(c(0, x)), # paramètre alpha révisé
      rate = 3 + n,                 # paramètre lambda révisé
      "B[n + 1]" = primes)          # primes bayésiennes

###
### Modèle normale/normale
###

## Certains modèles demandent de spécifier un paramètre de la
```

```
## distribution de S|Theta = theta. Dans le cas
## normale/normale, le paramètre à spécifier, 'sd' porte le
## même nom que le paramètre de la distribution de Theta.
## Comme le précise la rubrique d'aide de 'cm', il faut dans
## ce cas utiliser 'sd.lik' pour le paramètre de la
## vraisemblance.
cm("bayes", data = c(5, 3, 0, 1, 1),
   likelihood = "normal", sd.lik = 2,
   mean = 2, sd = 1)
```

3.9 Exercices

3.1 Un portefeuille d'assurance automobile est composé de 35 % de bons conducteurs, 40 % de conducteurs moyens et 25 % de mauvais conducteurs. L'actuaire a estimé que les bons conducteurs ont, en moyenne, un accident par 10 ans, les conducteurs moyens, deux accidents et les mauvais conducteurs, six accidents. L'actuaire suppose de plus que la fréquence des accidents a une distribution de Poisson. Par souci de simplicité, les sinistres sont tous d'un montant de 1.

a) Quelle est la probabilité qu'un assuré choisi au hasard ait un accident ?

b) Calculer la prime de risque pour chacun des trois types de conducteurs.

c) Calculer la prime collective.

d) Calculer la prime bayésienne de sixième année d'un contrat ayant le dossier suivant au cours des cinq premières années : 1, 0, 1, 1, 0.

3.2 Le *Zchoulp* se joue avec deux dés à six faces. Le premier est un dé usuel, dont les faces sont numérotées de 1 à 6. Le second dé a deux faces numérotées 5 et les autres numérotées de 1 à 4.

a) Un dé est choisi au hasard et lancé. Quelle est la probabilité d'obtenir un 5 ?

b) Si le résultat du lancer en a) est un 5 et que le même dé est relancé, quelle est la probabilité d'obtenir à nouveau un 5 ?

3.3 Deux urnes contiennent chacune une pièce de monnaie. La pièce dans l'urne A tombe sur face 40 % des fois et celle dans l'urne B, 80 % des fois. Nous choisissons une urne au hasard (la probabilité de choisir l'urne A est la même que celle de choisir l'urne B) et nous prenons la pièce qui s'y trouve. Nous lançons la pièce en l'air cinq fois et elle retombe sur pile quatre fois. Si nous lançons la pièce à cinq autres reprises, quel est le nombre espéré de fois que la pièce retombera sur pile ?

3.4 Les employeurs couverts par le régime d'assurance médicament d'un assureur sont classés par ce dernier dans trois groupes de taille égale. La probabilité de subir un sinistre dans une période pour chacun de ces groupes est donnée dans le tableau ci-dessous.

Groupe	Probabilité de sinistre
Fréquence faible	10 %
Fréquence moyenne	20 %
Fréquence élevée	40 %

L'assureur suppose de plus que les sinistres sont indépendants à l'intérieur de chacun des groupes.

a) Quelle est la probabilité qu'un employeur choisi au hasard dans ce portefeuille ait un sinistre ?

b) Après deux années d'expérience, l'employeur choisi en a) présente un dossier de sinistre vierge. À la lumière de ces résultats, quelle est la probabilité que cet employeur fasse partie du groupe à fréquence de sinistre faible ?

c) Quelle est maintenant la probabilité que l'employeur mentionné ci-dessus ait un sinistre lors de la troisième année ?

3.5 Nous souhaitons vérifier si une pièce de monnaie est équilibrée ou non (probabilité de $\frac{1}{2}$ de tomber sur pile ou face). L'incertitude quant à la probabilité θ d'obtenir, disons, face lors d'un lancer de la pièce est traduite en une variable aléatoire Θ. Celle-ci est distribuée selon une loi Bêta de paramètres $a > 0$ et $b > 0$.

a) Trouver $E[\Theta]$.

b) La pièce de monnaie est lancée n fois. La variable aléatoire S représente le nombre de fois que la pièce est tombée sur face. Trouver la distribution à postériori de Θ.

c) Une petite expérience pratique maintenant, aussi ludique qu'enrichissante. Jouez à pile ou face une bonne dizaine de fois avec une pièce quelconque. Enregistrer une valeur de 0 pour pile et de 1 pour face. Après chaque lancer, calculez la distribution à postériori de Θ et tracez-en un graphique approximatif. Vous pouvez aussi utiliser les fonctions `curve` et `dbeta` de R pour faire les graphiques. Par exemple, pour tracer la densité d'une bêta avec paramètres $a = 4$ et $b = 2$ l'expression à utiliser est

```
> curve(dbeta(x, 4, 2), from = 0, to = 1)
```

Observez les déplacements de la densité en fonction des résultats. Commencez l'expérience avec $a = b = 1$, soit $\Theta \sim U(0, 1)$.

d) Quelle est votre estimation de la probabilité d'obtenir pile au onzième lancer de la pièce si vous ignorez les résultats des dix premiers lancers ?

e) Quelle est maintenant votre estimation si vous connaissez les résultats des dix lancers effectués en c) ?

3.6 Votre opinion à priori quant à la distribution du montant des sinistres est une loi de Pareto de paramètres $\lambda = 10$ et $\theta = 1, 2$ ou 3, ces valeurs de θ étant toutes équiprobables. Pour un contrat choisi au hasard, vous observez par la suite un sinistre d'un montant de 20. Déterminer la probabilité que le montant du prochain sinistre de ce contrat soit supérieur à 30.

3.7 On demande à Camille d'élaborer un modèle pour la fréquence des sinistres au sein d'un portefeuille composé de dix contrats. Incertaine quant à la probabilité d'avoir un accident, Camille estime à 20 % la possibilité que la probabilité soit de 0,04, 60 % qu'elle soit de 0,10 et 20 % qu'elle soit de 0,16.

a) Quel est le modèle de Camille pour N, le nombre total d'accidents du portefeuille au cours d'une année ?

b) Quelle est la probabilité qu'il y ait 0, 1 et 2 accidents au cours d'une année ?

c) Comparer ces résultats avec la situation où Camille serait certaine que la probabilité d'accident est de 0,10.

3.8 Soit $S|\Theta \sim \text{Gamma}(2, \Theta)$ et $\Theta^{-1} \sim \text{Bêta}(2, 1)$. Trouver $\text{Var}[S]$.

3.9 Soit S la variable aléatoire représentant le nombre de sinistres d'un contrat d'assurance au cours d'une année. Le nombre de sinistres a une distribution de Poisson de paramètre inconnu Θ. La fonction de densité de probabilité de Θ est la suivante :

$$u(\theta) = \frac{5}{4}\frac{1}{\theta^2}, \quad 1 < \theta < 5.$$

a) Calculer $\Pr[S = 2]$.

b) Calculer $\Pr[S_3 = 0 | S_1 = 1, S_2 = 1]$.

c) Calculer la prime bayésienne de troisième année étant donné l'expérience en b).

3.10 La compagnie YARD assure un groupe de maisons contre les incendies. Son actuaire a divisé les maisons en trois classes de risque équiprobables : A, B et C. La probabilité qu'une maison prenne feu dans une année est de $\frac{1}{4}$, quelle que soit la classe. La distribution du montant à payer sachant qu'il y a eu incendie est donnée dans le tableau ci-dessous. Quelle est la prime bayésienne pour la deuxième année pour un assuré qui a eu un sinistre de 30 000 \$ l'année dernière ?

	Probabilité		
Montant du sinistre	Classe A	Classe B	Classe C
10 000	3/5	0	1/5
20 000	1/5	1/2	1/5
30 000	1/5	1/2	3/5

3.11 Vous savez que le nombre de sinistres pour un assuré est distribué selon une loi de Poisson de paramètre aléatoire. Ce paramètre est distribué selon une loi gamma avec moyenne 2 et variance 2. De plus, tout sinistre est d'un montant de 1 \$.

a) Trouver la prime qui devrait être exigée d'un nouvel assuré.

b) Trouver la prime bayésienne pour la quatrième année si cet assuré a eu huit sinistres au cours des trois premières années.

3.12 Le nombre annuel d'accidents d'un assuré suit une loi binomiale de paramètre $n = 2$. Il y a toutefois incertitude quant à la probabilité θ que cet assuré ait un accident. Trois valeurs sont jugées possibles : $\frac{1}{4}$, $\frac{1}{2}$ et $\frac{1}{8}$, et ce, avec probabilité 25 %, 25 % et 50 %, respectivement. L'assuré n'a eu aucun accident la première année et deux accidents la deuxième année.

a) Calculer la distribution révisée de Θ à la lumière des deux premières années d'expérience.

b) Calculer le nombre espéré d'accidents de cet assuré pour la troisième année.

3.13 On simule une expérience de sinistre comme suit : on lance un dé, et on pose le résultat égal à θ. Connaissant θ, on simule un nombre aléatoire d'une distribution uniforme sur $[0, 100\theta]$.

a) Trouver la prime bayésienne si des résultats de 80 et 340 ont été obtenus lors des deux premiers essais.

b) Est-il possible d'écrire la prime bayésienne sous forme d'une prime de crédibilité ? Si oui, trouver le facteur de crédibilité.

3.14 Une actuaire suppose que $S_t|\Theta = \theta$ suit une loi Gamma$(2, \theta)$ et elle fait l'hypothèse que la fonction de densité de probabilité de Θ est la suivante :

$$u(\theta) = \begin{cases} \theta/50, & 0 < \theta < 10 \\ 0, & \text{ailleurs.} \end{cases}$$

Elle simule ensuite deux valeurs de S et obtient $x_1 = 1$ et $x_2 = 1$. Quelle est maintenant la distribution à postériori de Θ pour ce simulateur ?

3.15 On vous donne les informations ci-dessous au sujet d'un régime d'assurance dentaire.

i) La fréquence des sinistres des assurés suit une loi de Poisson.

ii) La moitié des assurés a en moyenne deux sinistres par année.

iii) L'autre moitié a en moyenne quatre sinistres par année.

Un assuré choisi au hasard au sein du portefeuille a eu quatre sinistres dans chacune des deux premières années.

a) Énoncer le modèle complet utilisé ici par l'assureur.

b) Déterminer l'estimateur bayésien du nombre de sinistres de cet assuré pour la troisième année.

3.16 Un portefeuille d'assurance est composé de 25 % de bons risques, 60 % de risques moyens et 15 % de mauvais risques. Tous les risques ont une distribution de sinistres de type gamma, mais dont les paramètres diffèrent selon le tableau ci-dessous.

Type de risque	α	λ
Bon	4	2
Moyen	4	1
Mauvais	10	2

Le dossier de sinistre d'un risque choisi au hasard est de 1 et 2 au cours des deux premières années. Calculer la prime bayésienne de ce risque pour la troisième année.

3.17 On vous fournit l'information suivante au sujet de deux groupes de contrats.

i) La fréquence des sinistres des contrats du groupe A a une distribution de Poisson de moyenne 1 par année.

ii) La fréquence des sinistres des contrats du groupe B a une distribution de Poisson de moyenne 3 par année.

iii) Les montants de sinistres des contrats du groupe A a une distribution exponentielle de moyenne 1.

iv) Les montants de sinistres des contrats du groupe B a une distribution exponentielle de moyenne 3.

v) Les deux groupes sont composés d'un nombre égal de contrats.

vi) À l'intérieur de chaque groupe, la fréquence et la sévérité des sinistres sont indépendantes.

Un contrat choisi au hasard a deux accidents au cours de la première année. Le montant de ces sinistres est de 1 et 3.

a) Énoncer le modèle pour la fréquence et la sévérité des sinistres dans ce portefeuille.

b) Calculer l'espérance à postériori du montant total des sinistres du contrat choisi ci-dessus. (*Note* : calculer l'espérance du montant total des sinistres comme le produit de l'espérance de la fréquence et de l'espérance de la sévérité des sinistres.)

3.18 Soit un modèle géométrique/bêta, c'est-à-dire

$$\Pr[S = x | \Theta = \theta] = \theta(1-\theta)^x, \quad x = 0, 1, \dots$$

et

$$u(\theta) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\, \theta^{a-1}(1-\theta)^{b-1}, \quad 0 < \theta < 1.$$

a) Calculer la prime de risque.

b) Calculer la prime collective.

c) Calculer la distribution à postériori de Θ après n années d'expérience $S_1, \dots, S_n$.

d) Calculer la distribution prédictive de S_{n+1}.

e) Calculer la prime bayésienne à partir du résultat en c) ou celui en d). Pourquoi avoir choisi une approche plutôt qu'une autre ?

f) Exprimer la prime bayésienne en e) comme une prime de crédibilité.

3.19 Soit $S|\Theta \sim \text{Binomiale}(\nu, \Theta)$ et $\Theta \sim \text{Bêta}(a, b)$.

a) Déterminer la prime de risque.

b) Déterminer la prime collective.

c) Déterminer la distribution à postériori de Θ après n années.

d) Déterminer la prime bayésienne pour la $(n+1)^{\text{e}}$ année et vérifier si celle-ci peut s'exprimer comme une prime de crédibilité ou non.

3.20 Soit $S|\Theta \sim \text{Gamma}(\tau, \Theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$.

a) Déterminer la distribution marginale de S. Identifier cette distribution à trois paramètres.

b) Calculer la prime de risque.

c) Calculer la prime collective, d'abord à l'aide de la distribution marginale de S, puis comme la moyenne des primes de risque.

d) Déterminer la distribution à postériori de Θ après n années d'expérience $S_1, \dots, S_n$.

e) Déterminer la distribution prédictive de S_{n+1}.

f) Calculer la prime bayésienne, d'abord à partir de la distribution à postériori de Θ, puis à partir de la distribution prédictive.

g) La prime bayésienne est-elle une prime de crédibilité ?

3.21 Soit $S|\Theta = \theta \sim \text{Exponentielle}(\theta)$, $\Theta \sim \text{Gamma}(7, 42)$, la prime bayésienne de cinquième année est 9 et celle de sixième année, 8,5. Trouver x_5.

3.22 Vous utilisez un modèle Poisson/gamma pour la tarification bayésienne d'un contrat d'assurance. Les paramètres du modèle sont tels qu'après quatre années le facteur de crédibilité de ce contrat serait de 0,8. Si vous changez les hypothèses de telle sorte que la variance de la distribution de Θ est doublée, mais que l'espérance demeure inchangée, combien d'années faudra-t-il au contrat pour atteindre un niveau de crédibilité de 0,8 ?

3.23 Pour un modèle Poisson/gamma, on vous donne

$$\Pr[S_3 = x | S_1 = 1, S_2 = 2] = \binom{6+x}{x} (0{,}9)^7 (0{,}1)^x, \quad x = 0, 1, \dots.$$

Quelle est l'espérance de la distribution à priori de Θ ?

3.24 La distribution de $S_t|\Theta = \theta$ est une exponentielle de paramètre (θ) et la distribution à priori de Θ est une gamma de paramètre de forme 2 et de moyenne $1/4$. Calculer $\Pr[S_4 \leq 5|S_1 = 2, S_2 = 3, S_3 = 7]$.

3.25 Un contrat d'assurance a encouru les sinistres suivants sur une période de cinq années : 3, 1, 5, 4, 2. Vous utilisez un modèle de tarification bayésienne Poisson/gamma. Calculez la prime de crédibilité de ce contrat pour la sixième année pour chacune des combinaisons de paramètres de la distribution gamma ci-dessous. Interprétez les différences entre les primes de crédibilité.

a) $\alpha = 10$, $\lambda = 5$

b) $\alpha = 50$, $\lambda = 25$

c) $\alpha = \frac{1}{2}$, $\lambda = \frac{1}{4}$

3.26 La distribution marginale du montant total des sinistres d'un contrat d'assurance est

$$f(x) = \frac{1\,500}{(100 + x)^{2,5}}, \quad x > 0.$$

Sous les hypothèses usuelles en théorie de la crédibilité, quel est le montant total des sinistres espéré après cinq années sans accident ?

3.27 Les montants de sinistres d'un contrat furent de $S_1 = 7$, $S_2 = 13$, $S_3 = 1$, $S_4 = 4$ au cours des quatre premières années qu'il était couvert par votre compagnie d'assurance. Votre expérience antérieure avec ce type de contrat vous permet de postuler le modèle suivant pour les montants de sinistres de ce contrat :

$$\Pr[S = x|\Theta = \theta] = \binom{x+4}{4}\theta^5(1-\theta)^x, \quad x = 0, 1, \dots,$$
$$u(\theta) = 504\,\theta^5(1-\theta)^3, \quad 0 < \theta < 1.$$

Calculer la prime bayésienne de cinquième année.

3.28 Pour un certain modèle Poisson/gamma, nous avons

$$\Pr[S_3 = s_3|S_1 = 1, S_2 = 2] = \binom{6+s_3}{s_3}(0{,}9)^7(0{,}1)^{s_3}, \quad s_3 = 0, 1, \dots.$$

Trouver la covariance entre S_1 et S_2.

3.29 Cet exercice sert à démontrer le résultat obtenu par Jewell (1974), à savoir que la prime bayésienne issue d'un mélange d'une distribution de la famille exponentielle avec sa conjuguée naturelle est une prime de crédibilité. Soit la variable aléatoire $S|\Theta = \theta$ dont la distribution est membre de la famille exponentielle univariée, c'est-à-dire que

$$f(x|\theta) = \frac{p(x)e^{-\theta x}}{q(\theta)},$$

où $p(\cdot)$ et $q(\cdot)$ sont des fonctions quelconques.

a) Démontrer que la conjuguée naturelle de $f(x|\theta)$ est

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

où $t_0 > 0$, $x_0 > 0$ et $d(t_0, x_0) = \int_{-\infty}^{\infty} q(\theta)^{-t_0}e^{-\theta x_0}\,d\theta$. (*Astuce* : démontrer que la distribution à postériori $u(\theta|x_1, \dots, x_n)$ est de la même famille que la distribution à priori $u(\theta)$.)

b) Démontrer que

$$\mu(\theta) = -\frac{q'(\theta)}{q(\theta)} = -\frac{d}{d\theta}\ln q(\theta).$$

c) Démontrer que

$$\frac{d}{d\theta}u(\theta) = (t_0\mu(\theta) - x_0)u(\theta).$$

d) En intégrant l'équation en c) de part et d'autre par rapport à θ et en supposant que $u(\theta) = 0$ aux deux extrémités de son domaine de définition, démontrer que la prime collective est $E[\mu(\Theta)] = x_0/t_0$.

e) Avec ce qui précède, trouver la prime bayésienne et démontrer qu'il s'agit d'une prime de crédibilité.

3.30 La distribution de la variable aléatoire $S|\Theta = \theta$ est exponentielle de moyenne θ et celle de Θ est une gamma de paramètres α et λ. Démontrer que, dans un tel cas, la distribution gamma n'est pas la conjuguée naturelle de l'exponentielle.

3.31 Démontrer que $\Theta \sim \text{Gamma}(\tau, \lambda)$ est la conjuguée naturelle de $S|\Theta = \theta \sim \text{Gamma}(\alpha, \theta)$ et trouver les paramètres de la distribution à postériori de Θ.

3.32 Cet exercice est dérivé de Bühlmann et Gisler (2005, section 2.6). La distribution de Pareto translatée (ou « Pareto à un paramètre », dans la terminologie de Klugman et collab. (2012) aussi utilisée dans **actuar**) est très utile en réassurance pour la modélisation des montants extrêmes de sinistres au-delà d'un certain seuil. La fonction de densité de probabilité de la Pareto translatée de paramètre θ est

$$f(x) = \frac{\theta x_0^\theta}{x^{\theta+1}}, \quad x > x_0.$$

et son espérance est $x_0\theta/(\theta-1)$. Dans la suite, nous nous intéresserons à l'estimation bayésienne non pas de la prime de risque, mais bien seulement du paramètre θ.

a) Démontrer que l'estimateur du maximum de vraisemblance du paramètre θ de la loi de Pareto translatée à partir d'un échantillon aléatoire $X_1, \dots, X_n$ est

$$\hat{\theta}^{\text{EMV}} = \frac{n}{\sum_{i=1}^n \ln(X_i/x_0)}.$$

b) Démontrer que la loi gamma de paramètres α et λ est la conjuguée naturelle de la loi de Pareto translatée pour le paramètre θ.

c) Démontrer que l'estimateur bayésien du paramètre θ peut s'exprimer sous la forme

$$\hat{\Theta} = \eta\hat{\theta}^{\text{EMV}} + (1-\eta)\frac{\alpha}{\lambda}.$$

C'est un estimateur linéaire, mais s'agit-il d'une prime de crédibilité ?

4 Modèle de crédibilité de Bühlmann

Objectifs du chapitre

- Expliquer le lien entre la prime de crédibilité de Bühlmann et la prime bayésienne.
- Calculer la covariance entre deux variables aléatoires du modèle de Bühlmann.
- Minimiser l'espérance quadratique moyenne dans le cadre du modèle de Bühlmann.
- Calculer l'espérance des estimateurs des paramètres de structure dans le modèle de Bühlmann.
- Calculer une prime de crédibilité de Bühlmann dans un cadre paramétrique.
- Calculer une prime de crédibilité de Bühlmann à partir de données de sinistres.
- Utiliser la fonction `cm` du paquetage R **actuar** pour calculer des primes de crédibilité de Bühlmann et interpréter les résultats de la fonction.

L'approche purement bayésienne que nous avons étudiée au chapitre précédent s'avère élégante, certes, mais elle recèle également deux inconvénients qui la rendent peu applicable en pratique. Tout d'abord, la prime bayésienne n'est une prime de crédibilité — facile à calculer et à interpréter — que pour certaines combinaisons de distributions seulement. Ensuite, tout le calcul de la prime repose sur des hypothèses très subjectives pour les distributions de Θ et $S_t|\Theta$. L'effet du choix de distributions s'estompe avec le volume d'expérience, mais nécessite un nombre d'années d'expérience bien plus grand que ce dont nous disposons habituellement en pratique.

Nous avons déjà vu à la section 3.2 que l'une des contributions importantes du père de la théorie du risque moderne, Hans Bühlmann, à la théorie de la crédibilité de précision est le modèle d'hétérogénéité. Dans ce chapitre, nous étudions deux autres de ses contributions qui visent à pallier aux inconvénients de l'approche bayésienne mentionnés ci-dessus. Pour résoudre l'enjeu de linéarité de la prime bayésienne, Bühlmann (1967) a proposé de restreindre l'approximation de la prime de risque aux fonctions linéaires des observations, c'est-à-dire de la forme $c_0 + \sum_{t=1}^{n} c_t S_{it}$. Nous verrons que la meilleure approximation se trouve à être une prime de crédibilité.

Bühlmann (1969) a ensuite proposé d'avoir recours à la statistique bayésienne empirique pour contourner le problème du choix des distributions. Dans cette approche non paramétrique, le calcul de la prime de crédibilité repose sur l'estimation de paramètres à partir des données accumulées.

4.1 Notation et relations de covariance

Avant d'aller plus loin, nous allons définir quelques éléments de notation essentiels pour la suite, en plus d'établir des relations de covariance nécessaires pour calculer la meilleure approximation linéaire de la prime de risque.

Nous avons déjà défini au chapitre 3 la prime collective $m = E[\mu(\Theta)]$. Ajoutons maintenant les deux variances suivantes :

$$\begin{aligned} s^2 &= E[\mathrm{Var}[S_{it}|\Theta_i]] \\ &= E[\sigma^2(\Theta_i)] \\ a &= \mathrm{Var}[E[S_{it}|\Theta_i]] \\ &= \mathrm{Var}[\mu(\Theta_i)]. \end{aligned}$$

> Dans les textes de la Casualty Actuarial Society, les variances s^2 et a sont souvent appelées, respectivement, *expected process variance* (EPV) et *variance of the hypothetical means* (VHM).

Le théorème suivant permet de calculer la covariance entre deux variables aléatoires en conditionnant sur une troisième.

Théorème 4.1 (Théorème de la covariance totale)**.** *Soit X, Y et Θ des variables aléatoires dont la densité conjointe existe. Alors*

$$\mathrm{Cov}(X, Y) = \mathrm{Cov}(E[X|\Theta], E[Y|\Theta]) + E[\mathrm{Cov}(X, Y|\Theta)].$$

Démonstration. Pour la présente démonstration, vous devez toujours garder en tête qu'une espérance conditionnelle est une variable aléatoire. De plus, nous utiliserons à répétition le fait que, pour toute variable aléatoire

Y, $E[Y] = E[E[Y|\Theta]]$ et le premier moment centré de Y est nul, c'est-à-dire que

$$E[Y - E[Y]] = 0.$$

Par définition de la covariance, puis en ajoutant et en retranchant des espérances conditionnelles, nous avons :

$$\begin{aligned}\mathrm{Cov}(X,Y) &= E[(X - E[X])(Y - E[Y])] \\ &= E[(X - E[X|\Theta] + E[X|\Theta] - E[X])\cdot \\ &\qquad (Y - E[Y|\Theta] + E[Y|\Theta] - E[Y])].\end{aligned}$$

Nous pouvons ensuite décomposer l'espérance en quatre termes, puis conditionner sur Θ :

$$\begin{aligned}\mathrm{Cov}(X,Y) = {} & E[E[(X - E[X|\Theta])(Y - E[Y|\Theta])|\Theta]] \\ & + E[E[(X - E[X|\Theta])(E[Y|\Theta] - E[Y])|\Theta]] \\ & + E[E[(E[X|\Theta] - E[X])(Y - E[Y|\Theta])|\Theta]] \\ & + E[(E[X|\Theta] - E[X])(E[Y|\Theta] - E[Y])]\end{aligned}$$

À ce stade, il faut réaliser qu'à l'intérieur des espérances conditionnelles, les termes $(E[X|\Theta] - E[X])$ et $(E[Y|\Theta] - E[Y])$ sont des constantes. Nous pouvons donc les sortir des espérances conditionnelles pour obtenir :

$$\begin{aligned}\mathrm{Cov}(X,Y) = {} & E[E[(X - E[X|\Theta])(Y - E[Y|\Theta])|\Theta]] \\ & + E[(E[Y|\Theta] - E[Y])E[X - E[X|\Theta]|\Theta]] \\ & + E[(E[X|\Theta] - E[X])E[Y - E[Y|\Theta]|\Theta]] \\ & + E[(E[X|\Theta] - E[X])(E[Y|\Theta] - E[Y])]\end{aligned}$$

Dans le second terme, $E[X - E[X|\Theta]|\Theta] = 0$ puisqu'il s'agit du premier moment centré de la variable aléatoire $X|\Theta$. De même, $E[Y - E[Y|\Theta]|\Theta] = 0$. Par conséquent, les second et troisième termes ci-dessus sont nuls. Ne reste donc que

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

ce qui complète la démonstration. □

Corollaire 4.1. *En posant $X \equiv Y$ dans le théorème 4.1, on obtient*

$$\mathrm{Var}[X] = E[\mathrm{Var}[X|\Theta]] + \mathrm{Var}[E[X|\Theta]].$$

Un dernier résultat pour cette section : le théorème suivant établit les relations de variance entre des variables aléatoires dans un contexte analogue à notre modèle d'hétérogénéité.

Théorème 4.2. *Soit $S_1, \dots, S_n$ des variables aléatoires conditionnellement indépendantes sachant la variable aléatoire Θ, de même espérance et de même variance, c'est-à-dire que*

$$\begin{aligned} E[S_t|\Theta] &= \mu(\Theta), \quad t = 1, \dots, n \\ \operatorname{Var}[S_t|\Theta] &= \sigma^2(\Theta), \quad t = 1, \dots, n. \end{aligned}$$

Alors

$$\begin{aligned} \operatorname{Cov}(S_t, S_u) &= \begin{cases} a, & t \neq u \\ a + s^2, & t = u \end{cases} \\ &= a + \delta_{tu} s^2, \quad t, u = 1, \dots, n \\ \operatorname{Cov}(\mu(\Theta), S_t) &= a, \end{aligned}$$

où δ_{tu} est le delta de Kronecker :

$$\delta_{tu} = \begin{cases} 1, & t = u \\ 0, & t \neq u. \end{cases}$$

Démonstration. Pour le premier résultat, nous avons, en utilisant le théorème 4.1,

$$\begin{aligned} \operatorname{Cov}(S_t, S_u) &= \operatorname{Cov}(E[S_t|\Theta], E[S_u|\Theta]) + E[\operatorname{Cov}(S_t, S_u|\Theta)] \\ &= \operatorname{Cov}(\mu(\Theta), \mu(\Theta)) + E[\delta_{tu} \operatorname{Var}[S_t|\Theta]] \\ &= \operatorname{Var}[\mu(\Theta)] + \delta_{tu} E[\sigma^2(\Theta)] \\ &= a + \delta_{tu} s^2. \end{aligned}$$

Pour le second résultat, toujours en conditionnant sur la variable aléatoire Θ,

$$\begin{aligned} \operatorname{Cov}(\mu(\Theta), S_t) &= \operatorname{Cov}(\mu(\Theta), E[S_t|\Theta]) + E[\operatorname{Cov}(\mu(\Theta), S_t|\Theta)] \\ &= \operatorname{Var}[\mu(\Theta)] + E[0] \\ &= a, \end{aligned}$$

ce qui complète la démonstration. □

La démonstration pas à pas du théorème 4.1 et d'une partie du théorème 4.2 est également disponible en vidéo.

TAB. 4.1 – Représentation schématique des variables aléatoires dans le modèle de Bühlmann

Variables non observables	Observations 1	$\cdots$	t	$\cdots$	n
Θ_1	S_{11}	$\cdots$	S_{1t}	$\cdots$	S_{1n}
$\vdots$	$\vdots$				$\vdots$
Θ_i	S_{i1}	$\cdots$	S_{it}	$\cdots$	S_{in}
$\vdots$	$\vdots$				$\vdots$
Θ_I	S_{I1}	$\cdots$	S_{It}	$\cdots$	S_{In}

4.2 Modèle et prévision

Le modèle d'hétérogénéité utilisé dans le modèle de crédibilité de Bühlmann est très similaire à celui du chapitre 3, sauf que nous relâchons légèrement certaines hypothèses.

Le contexte est toujours celui d'un portefeuille de I contrats d'assurance, chacun caractérisé par un niveau de risque θ_i, $i = 1, \dots, I$ réalisation d'une variable aléatoire non observable Θ_i. Nous disposons pour fins de tarification de n périodes d'expérience $(S_{i1}, \dots, S_{in}) \equiv \mathbf{S}_i$ pour chaque contrat. Vous trouverez au tableau 4.1 une représentation schématique de l'ensemble de variables aléatoires dans le modèle de Bühlmann.

Dans leur version la moins restrictive, les hypothèses du modèle de Bühlmann sont les suivantes.

(B1) Les contrats $(\Theta_i, \mathbf{S}_i)$, $i = 1, \dots, I$ sont indépendants, les variables aléatoires $\Theta_1, \dots, \Theta_I$ sont identiquement distribuées et les variables aléatoires S_{it} ont une variance finie.

(B2) Les variables aléatoires S_{it}, sont telles que

$$\begin{aligned} E[S_{it}|\Theta_i] &= \mu(\Theta_i) \quad i = 1, \dots, I \\ \operatorname{Cov}(S_{it}, S_{iu}|\Theta_i) &= \delta_{tu}\sigma^2(\Theta_i), \quad t, u = 1, \dots, n. \end{aligned}$$

L'hypothèse d'indépendance entre les contrats peut ne pas être réaliste, mais elle simplifie les calculs et il s'agit d'une bonne approximation dans plusieurs cas.

L'hypothèse (B1) établit l'indépendance inter contrats (*between*). L'hypothèse (B2), quant à elle, stipule l'homogénéité temporelle des contrats — leur prime de risque est constante dans le temps — ainsi que l'« indépendance » intra contrats (*within*) — les observations d'un contrat sont conditionnellement non corrélées.

Avec le modèle et les hypothèses en main, nous pouvons maintenant procéder au calcul de la prévision de la prime de risque à partir des observations. Au chapitre 3, nous n'avons posé aucune restriction sur la forme de cette prévision, avec pour conséquence qu'elle peut s'avérer difficile à calculer. Le trait de génie de Bühlmann (1967), ç'aura été de contraindre la prévision à être une fonction linéaire des observations. Le résultat se trouve dans le théorème suivant.

Théorème 4.3. *Pour un portefeuille tel qu'illustré ci-dessus et sous les hypothèses (B1) et (B2), la meilleure approximation linéaire non homogène de la prime de risque $\mu(\Theta_i)$ est*

$$\pi^B_{i,n+1} = z\bar{S}_i + (1-z)m, \tag{4.1}$$

où

$$\bar{S}_i = \frac{1}{n}\sum_{t=1}^{n} S_{it}$$

$$z = \frac{n}{n+K}, \quad K = \frac{s^2}{a}.$$

Démonstration. Nous recherchons la meilleure approximation — au sens des moindres carrés — de la prime de risque du contrat i, $i = 1, \dots, I$, de la forme $c_0^i + \sum_{j=1}^{I}\sum_{t=1}^{n} c_{jt}^i S_{jt}$. Nous devons donc trouver les constantes $c_0^i, c_{11}^i, \dots, c_{kn}^i$ qui minimisent

$$E\left[\left(\mu(\Theta_i) - c_0^i - \sum_{j=1}^{I}\sum_{t=1}^{n} c_{jt}^i S_{jt}\right)^2\right].$$

Par indépendance entre les contrats, nous savons déjà que la prime de crédibilité du contrat i sera une fonction de ses observations seulement. Nous pouvons donc immédiatement réduire le problème à la recherche des constantes $c_0, c_1, \dots, c_n$ qui minimisent

$$E\left[\left(\mu(\Theta_i) - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)^2\right].$$

En calculant les dérivées partielles, d'abord par rapport à c_0, puis par rapport à c_u, $u = 1, \dots, n$, nous obtenons :

$$c_0 = E[\mu(\Theta_i)] - \sum_{t=1}^{n} c_t E[S_{it}] \tag{4.2}$$

$$\mathrm{Cov}(\mu(\Theta_i), S_{iu}) = \sum_{t=1}^{n} c_t \, \mathrm{Cov}(S_{it}, S_{iu}). \tag{4.3}$$

Or, à l'aide des résultats du théorème 4.2, nous pouvons réécrire l'équation (4.3) ainsi :

$$\begin{aligned} a &= \sum_{t=1}^{n} c_t (a + \delta_{tu} s^2) \\ &= a \sum_{t=1}^{n} c_t + c_u s^2. \end{aligned}$$

Par symétrie de cette égalité pour $u = 1, \dots, n$, nous pouvons conclure que

$$c_1 = c_2 = \cdots = c_n = c = \frac{a}{an + s^2}.$$

De l'équation (4.2),

$$c_0 = (1 - nc)m$$

et donc

$$\begin{aligned} \pi^{\mathrm{B}}_{i,n+1} &= c_0 + \sum_{t=1}^{n} c_t S_{it} \\ &= \frac{an}{an + s^2} \sum_{t=1}^{n} \frac{S_{it}}{n} + \left(1 - \frac{an}{an + s^2}\right) m \\ &= z\bar{S}_i + (1 - z)m \end{aligned}$$

avec $z = n/(n + s^2/a)$. □

Calculer la meilleure approximation de la prime de risque ou la meilleure prévision du montant des sinistres de la prochaine période sont des problèmes en tous points équivalents. En effet, remplacer $\mu(\Theta_i)$ dans le théorème 4.3 par $S_{i,n+1}$ ne change rien puisque $E[\mu(\Theta_i)] = E[S_{i,n+1}] = m$ et $\mathrm{Cov}(\mu(\Theta_i), S_{it}) = \mathrm{Cov}(S_{i,n+1}, S_{it}) = a$ pour $t = 1, \dots, n$.

La prime de crédibilité (4.1) possède deux belles propriétés. En premier lieu, elle est sans biais, c'est-à-dire que

$$E[\pi^{\mathrm{B}}_{i,n+1}] = zE[\bar{S}_i] + (1-z)m = m.$$

En moyenne, l'assureur perçoit donc suffisamment de primes pour payer les sinistres. L'absence de biais fait également en sorte qu'une mauvaise estimation du facteur de crédibilité n'a pas d'impact négatif sur le montant des primes perçu par l'assureur.

En second lieu, par la Loi des grands nombres, $\bar{S}_i \to \mu(\Theta_i)$ quand $n \to \infty$ et $z \to 1$ dans les mêmes circonstances. Par conséquent, $\pi^{\mathrm{B}}_{i,n+1} \to \mu(\Theta_i)$ quand $n \to \infty$.

L'énoncé du théorème 4.3 fait référence à une approximation *non homogène* de la prime de risque. Le qualificatif tient au terme constant c_0 dans l'approximation. Une approximation linéaire *homogène* de $\mu(\Theta_i)$ serait donc de la forme

$$\sum_{j=1}^{I} \sum_{t=1}^{n} c^i_{jt} S_{jt}.$$

Il est facile de démontrer que la meilleure approximation dans ce cas est

$$z\bar{X}_i + (1-z)\bar{S}, \quad \bar{S} = \frac{1}{I} \sum_{i=1}^{I} \bar{S}_i.$$

La prime de crédibilité peut aussi s'écrire

$$\pi^{\mathrm{B}}_{i,n+1} = m + z(\bar{S}_i - m).$$

Sous cette forme, la prime d'un contrat apparait comme la prime collective à laquelle on ajoute ou soustrait un ajustement, pondéré par un facteur de crédibilité, selon que l'expérience du contrat est pire ou meilleure que celle du groupe.

Le théorème suivant établit un résultat intéressant : la meilleure approximation linéaire de la prime de risque est également la meilleure approximation linéaire de la prime bayésienne. En d'autres termes, nous pouvons interpréter le calcul de la prime de crédibilité comme une minimisation en deux étapes : d'abord trouver la meilleure approximation de la prime de risque (la prime bayésienne), puis trouver la meilleure approximation linéaire de la prime bayésienne (la prime de crédibilité).

Théorème 4.4. *Si $\pi^B_{i,n+1}$ est la combinaison linéaire des observations minimisant*

$$E\left[\left(\mu(\Theta_i) - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)^2\right],$$

alors $\pi^B_{i,n+1}$ minimise également

$$E\left[\left(B_{i,n+1} - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)^2\right],$$

où $B_{i,n+1} = E[\mu(\Theta_i)|\mathbf{S}_i]$.

Démonstration. En ajoutant et soustrayant $B_{i,n+1}$ dans la première espérance du théorème, en développant le carré, puis en conditionnant sur Θ dans le terme mixte, nous obtenons :

$$\begin{aligned}
E\left[\left(\mu(\Theta_i) - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)^2\right] &= E\left[\left(\mu(\Theta_i) - B_{i,n+1}\right)^2\right] \\
&\quad + 2E\left[E\left[\left(\mu(\Theta_i) - B_{i,n+1}\right)\left(B_{i,n+1} - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)\Big|\mathbf{S}_i\right]\right] \\
&\quad + E\left[\left(B_{i,n+1} - c_0 - \sum_{t=1}^{n} c_t S_{it}\right)^2\right]
\end{aligned}$$

Or, le premier terme du côté droit de l'équation ne dépend pas du choix des constantes $c_0, c_1, \dots, c_n$ et le second terme est nul du fait que, par définition, $E[(\mu(\Theta_i) - B_{i,n+1})|\mathbf{S}_i] = 0$. Les constantes qui minimisent l'espérance du côté gauche minimisent donc aussi le troisième terme du côté droit. □

Corollaire 4.2. *Nous avons*

$$E[(\mu(\Theta_i) - \pi^B_{i,n+1})^2] \geq E[(\mu(\Theta_i) - B_{i,n+1})^2],$$

l'égalité survenant lorsque la prime bayésienne est une prime de crédibilité.

La vidéo sur le lien entre la prime de Bühlmann et la prime bayésienne propose une explication visuelle des résultats ci-dessus.

4.3 Approche paramétrique

Afin de comparer l'approche du modèle de crédibilité de Bühlmann à celle de la tarification purement bayésienne, nous allons pour un moment supposer que les distributions de $S_{it}|\Theta_i$ et de Θ_i sont connues. La notion de portefeuille n'est toujours pas nécessaire puisque nous déterminons les distributions pour chaque contrat. Nous omettons donc l'indice i dans les formules.

Le théorème 4.3 nous permet de calculer directement le facteur de crédibilité pour les combinaisons de distributions menant à une prime bayésienne linéaire ou, autrement, l'approximation linéaire de la prime bayésienne.

Exemple 4.1 (Bernoulli/uniforme)**.** Nous avons vu à l'exemple 3.4 que si $S_t|\Theta = \theta \sim \text{Bernoulli}(\theta)$ et $\Theta \sim U(a, b)$, alors la prime bayésienne est très compliquée. Or, $\mu(\theta) = \theta$ et $\sigma^2(\theta) = \theta(1 - \theta)$, d'où

$$\begin{aligned}
m &= E[\mu(\Theta)] \\
&= E[\Theta] \\
&= \frac{a+b}{2} \\
s^2 &= E[\sigma^2(\Theta)] \\
&= E[\Theta] - E[\Theta^2] \\
&= \frac{a+b}{2} + \frac{a^2+ab+b^2}{3} \\
a &= \text{Var}[\mu(\Theta)] \\
&= \text{Var}[\Theta] \\
&= \frac{(b-a)^2}{12},
\end{aligned}$$

donc

$$\begin{aligned}
K &= \frac{s^2}{a} \\
&= \frac{6(a+b) - 4(a^2+ab+s^2)}{(b-a)^2}
\end{aligned}$$

et la meilleure approximation linéaire de la prime bayésienne est

$$\pi_{n+1}^{\text{B}} = \frac{n}{n+K}\bar{S} + \left(1 - \frac{n}{n+K}\right)\frac{a+b}{2}.$$

Cette prime est facile à calculer. □

Exemple 4.2 (Poisson/gamma)**.** Examinons de nouveau la combinaison de distributions $S_t|\Theta = \theta \sim \text{Poisson}(\theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$. Nous savons déjà que $\mu(\theta) = \sigma^2(\theta) = \theta$. Par conséquent,

$$\begin{aligned} m &= E[\Theta] = \frac{\alpha}{\lambda} \\ s^2 &= E[\Theta] = \frac{\alpha}{\lambda} \\ a &= \text{Var}[\Theta] = \frac{\alpha}{\lambda^2}, \end{aligned}$$

d'où

$$K = \frac{s^2}{a} = \lambda$$

et

$$\pi^{\text{B}}_{n+1} = \frac{n}{n+\lambda}\bar{S} + \frac{\lambda}{n+\lambda}\frac{\alpha}{\lambda}.$$

C'est le même résultat qu'à l'exemple 3.5. □

Exemple 4.3 (Exponentielle/gamma)**.** Soit $S_t|\Theta = \theta \sim \text{Exponentielle}(\theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$. Le calcul de la prime de Bühlmann représente un défi plus grand pour cette combinaison de distributions puisque nous devons calculer des moments de l'inverse de la variable aléatoire Θ. En effet,

$$\begin{aligned} \mu(\Theta) &= \frac{1}{\Theta}, \\ \sigma^2(\Theta) &= \frac{1}{\Theta^2} \end{aligned}$$

et donc

$$\begin{aligned} s^2 &= E\left[\frac{1}{\Theta^2}\right] \\ a &= E\left[\frac{1}{\Theta^2}\right] - E\left[\frac{1}{\Theta}\right]^2. \end{aligned}$$

Or, pour $k \in \mathbb{Z}$,

$$\begin{aligned}
E[\Theta^k] &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty \theta^{\alpha+k-1} e^{-\lambda\theta}\, d\theta \\
&= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+k)}{\lambda^{\alpha+k}} \int_0^\infty \frac{\lambda^{\alpha+k}}{\Gamma(\alpha+k)}\, \theta^{\alpha+k-1} e^{-\lambda\theta}\, d\theta \\
&= \frac{\Gamma(\alpha+k)}{\Gamma(\alpha)\lambda^k} \\
&= \begin{cases}
\dfrac{\alpha(\alpha+1)\cdots(\alpha+k-1)}{\lambda^k}, & k \geq 1 \\
1, & k = 0 \\
\dfrac{\lambda^{|k|}}{(\alpha-1)(\alpha-2)\cdots(\alpha-|k|)}, & k \leq -1, |k| < \alpha.
\end{cases}
\end{aligned}$$

Nous avons donc

$$\begin{aligned}
E[\Theta^{-1}] &= \frac{\lambda}{\alpha-1} \\
E[\Theta^{-2}] &= \frac{\lambda^2}{(\alpha-1)(\alpha-2)}
\end{aligned}$$

et, ainsi,

$$\begin{aligned}
s^2 &= \frac{\lambda^2}{(\alpha-1)(\alpha-2)} \\
a &= \frac{\lambda^2}{(\alpha-1)(\alpha-2)} - \frac{\lambda^2}{(\alpha-1)^2} \\
&= \frac{\lambda^2}{(\alpha-1)^2(\alpha-2)}.
\end{aligned}$$

Finalement,

$$K = \frac{s^2}{a} = \alpha - 1$$

et le facteur de crédibilité dans la prime de Bühlmann est

$$z = \frac{n}{n+\alpha-1},$$

tel qu'obtenu à l'exemple 3.6. □

Petit secret de fabrication : lorsque la fonction `cm` du paquetage **actuar** calcule des primes bayésiennes linéaires, elle effectue les calculs à partir des formules des paramètres m, s^2 et a pour chacune des combinaisons de distributions. Ce sont d'ailleurs ces valeurs qui sont affichées à l'écran avec l'objet du modèle.

> Étudiez et exécutez les lignes 13-39 du fichier de script `buhlmann.R` reproduit à la section 4.6 et comparez les résultats obtenus avec `cm` aux formules des exemples 4.2 et 4.3.

4.4 Approche non paramétrique

En pratique, l'approche paramétrique s'avère d'un intérêt limité puisqu'elle nécessite toujours de déterminer les distributions des variables aléatoires $S_{it}|\Theta_i$ et Θ_i. Avec l'approche non paramétrique, nous délaissons l'approche bayésienne pure pour l'approche *bayésienne empirique*. Cela nécessite quelques ajustements à notre vocabulaire et à l'interprétation que nous faisons des différentes composantes de notre modèle d'hétérogénéité.

Le portefeuille de contrats joue — finalement — un rôle central. D'abord, chaque contrat, par son niveau de risque, représente une réalisation de la variable aléatoire Θ. La fonction de répartition $U(\theta)$ de Θ est la *fonction de structure* du portefeuille. Au chapitre 3, nous la voyions comme l'opinion à priori de l'assureur sur le niveau de risque d'un contrat. Dans l'approche empirique, la fonction de structure représente la proportion de contrats dans le portefeuille avec un niveau de risque inférieur ou égal à θ ou, de manière équivalente, la distribution des niveaux de risque entre les contrats.

Auparavant, les concepts d'homogénéité ou d'hétérogénéité étaient liés au niveau de certitude de l'assureur quant au niveau de risque d'un contrat. Ce n'était pas très intuitif. Dorénavant, nous parlerons d'un portefeuille homogène ou hétérogène selon que les moyennes des contrats s'avèrent semblables ou non.

Pour calculer les primes de crédibilité des contrats, nous devrons estimer les *paramètres de structure* du portefeuille, soit :

1. $m = E[\mu(\Theta)]$, la moyenne collective du portefeuille ;
2. $s^2 = E[\sigma^2(\Theta)]$, la variance intra (*within*) ou la variabilité moyenne du portefeuille, qui représente l'homogénéité temporelle ;
3. $a = \text{Var}[\mu(\Theta)]$, la variance inter (*between*) ou la variabilité entre les moyennes des contrats, qui représente l'homogénéité du portefeuille.

Pour effectuer l'estimation, nous allons développer des estimateurs sans biais des paramètres.

4.4.1 Estimation de la prime collective

Intuitivement, un estimateur de la prime collective est la moyenne empirique des moyennes individuelles :

$$\hat{m} = \bar{S} = \frac{1}{In} \sum_{i=1}^{I} \sum_{t=1}^{n} S_{it}. \tag{4.4}$$

Cet estimateur est bel et bien sans biais :

$$\begin{aligned} E[\hat{m}] &= \frac{1}{In} \sum_{i=1}^{I} \sum_{t=1}^{n} E[S_{it}] \\ &= \frac{1}{In} \sum_{i=1}^{I} \sum_{t=1}^{n} m \\ &= m. \end{aligned}$$

4.4.2 Estimation de la variance intra

Un estimateur sans biais de la variance du contrat i, $i = 1, \dots, n$, est

$$\frac{1}{n-1} \sum_{t=1}^{n} (S_{it} - \bar{S}_i)^2, \quad n \geq 2.$$

Pour obtenir un estimateur sans biais de la variance moyenne s^2, il suffit ensuite de calculer la moyenne de tous ces estimateurs :

$$\hat{s}^2 = \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} (S_{it} - \bar{S}_i)^2. \tag{4.5}$$

Pour démontrer l'absence de biais, remarquons d'abord que

$$\begin{aligned} E[(S_{it} - \bar{S}_i)^2 | \Theta_i] &= \mathrm{Var}[S_{it} - \bar{S}_i | \Theta_i] \\ &= \mathrm{Var}[S_{it} | \Theta_i] + \mathrm{Var}[\bar{S}_i | \Theta_i] - 2\,\mathrm{Cov}(S_{it}, \bar{S}_i | \Theta_i) \\ &= \sigma^2(\Theta_i) + \frac{\sigma^2(\Theta_i)}{n} - 2\frac{\sigma^2(\Theta_i)}{n} \\ &= \frac{n-1}{n}\, \sigma^2(\Theta_i) \end{aligned}$$

puisque $E[\bar{S}_i | \Theta_i] = E[S_{it} | \Theta_i] = \mu(\Theta_i)$. Par conséquent,

$$\begin{aligned} E[(S_{it} - \bar{S}_i)^2] &= E[E[(S_{it} - \bar{S}_i)^2 | \Theta_i]] \\ &= \frac{n-1}{n}\, E[\sigma^2(\Theta_i)] \\ &= \frac{n-1}{n}\, s^2 \end{aligned}$$

et

$$\begin{aligned} E[\hat{s}^2] &= \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} \frac{n-1}{n}\, s^2 \\ &= s^2. \end{aligned}$$

4.4.3 Estimation de la variance inter

Un estimateur intuitif de la variance entre les moyennes des contrats $a = \mathrm{Var}[\mu(\Theta)]$ serait

$$\frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2.$$

Or, cet estimateur est biaisé. En effet, $E[\bar{S}] = E[\bar{S}_i] = m$ et

$$\begin{aligned} E[(\bar{S}_i - \bar{S})^2] &= \mathrm{Var}[\bar{S}_i - \bar{S}] \\ &= \mathrm{Var}[\bar{S}_i] + \mathrm{Var}[\bar{S}] - 2\,\mathrm{Cov}(\bar{S}_i, \bar{S}). \end{aligned}$$

Par indépendance entre les contrats, nous avons

$$\begin{aligned} \mathrm{Cov}(\bar{S}_i, \bar{S}) &= \frac{1}{I} \sum_{j=1}^{I} \mathrm{Cov}(\bar{S}_i, \bar{S}_j) \\ &= \frac{1}{I}\, \mathrm{Var}[\bar{S}_i] \end{aligned}$$

et $\mathrm{Var}[\bar{S}] = I^{-1}\,\mathrm{Var}[\bar{S}_i]$, d'où

$$E[(\bar{S}_i - \bar{S})^2] = \frac{I-1}{I}\, \mathrm{Var}[\bar{S}_i]$$

et

$$E\left[\frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2\right] = \mathrm{Var}[\bar{S}_i].$$

Or,

$$\begin{aligned} \mathrm{Var}[\bar{S}_i] &= \mathrm{Var}[E[\bar{S}_i|\Theta_i]] + E[\mathrm{Var}[\bar{S}_i|\Theta_i]] \\ &= \mathrm{Var}[\mu(\Theta_i)] + E\left[\frac{\sigma^2(\Theta_i)}{n}\right] \\ &= a + \frac{s^2}{n}. \end{aligned}$$

TAB. 4.2 – Données du portefeuille de l'exemple 4.4

	Années					
Contrat	1	2	3	4	5	6
1	0	1	2	1	2	0
2	3	4	2	1	4	4
3	3	3	2	1	2	1

Le biais de l'estimateur est donc s^2/n. Pour obtenir un estimateur sans biais, nous allons simplement retrancher cette valeur de la somme des écarts au carré. Un estimateur sans biais de a est donc

$$\hat{a} = \frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2 - \frac{1}{n} \hat{s}^2. \tag{4.6}$$

L'estimateur $\hat{a}$ peut être négatif. En pratique, nous devons utiliser $\hat{a}' = \max(\hat{a}, 0)$, qui est un estimateur biaisé du paramètre a. C'est là un problème récurrent — et déroutant ! — en statistique que pour que l'estimateur d'une variance soit sans biais, il doit pouvoir être négatif.

4.4.4 Prime de crédibilité

Nous ne pouvons estimer directement sans biais la prime de crédibilité. Nous remplaçons simplement chaque paramètre de structure par son estimateur :

$$\begin{aligned} \hat{\pi}^{\mathrm{B}}_{i,n+1} &= \hat{z}\bar{S}_i + (1-\hat{z})\hat{m} \\ \hat{z} &= \frac{n}{n + \hat{s}^2/\hat{a}}. \end{aligned}$$

Bien que tous les estimateurs sont sans biais, nous ne pouvons conclure que $E[\hat{K}] = K$ et donc que $E[\hat{z}] = z$. Par conséquent, l'estimateur de la prime de crédibilité est fort probablement biaisé.

Exemple 4.4. Nous avons $n = 6$ années d'expérience d'un portefeuille de $I = 3$ contrats. Les données du portefeuille se trouvent au tableau 4.2. Calculons la prime de crédibilité pour la septième année pour chacun des contrats.

Tout d'abord, nous avons $\bar{S}_1 = 1$, $\bar{S}_2 = 3$, $\bar{S}_3 = 2$ et $\bar{S} = 2$. Ainsi, les estimateurs des paramètres de structure sont :

$$\begin{aligned}
\hat{m} &= \bar{S} = 2 \\
\hat{s}^2 &= \frac{1}{I} \sum_{i=1}^{I} \frac{1}{n-1} \sum_{t=1}^{n} (S_{it} - \bar{S}_i)^2 \\
&= \frac{1}{3}\left(\frac{4}{5} + \frac{8}{5} + \frac{4}{5}\right) = \frac{16}{15} \\
\hat{a} &= \frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2 - \frac{1}{n}\,\hat{s}^2 \\
&= \frac{2}{2} - \frac{1}{6} \cdot \frac{16}{15} = \frac{37}{45}.
\end{aligned}$$

Par conséquent, $\hat{K} = 48/37 \approx 1{,}30$ et

$$\hat{z} = \frac{6}{6 + 1{,}30} = 0{,}82,$$

d'où

$$\begin{aligned}
\hat{\pi}^{\mathrm{B}}_{1,7} &= 0{,}82(1) + 0{,}18(2) = 1{,}18 \\
\hat{\pi}^{\mathrm{B}}_{2,7} &= 2{,}82 \\
\hat{\pi}^{\mathrm{B}}_{3,7} &= 2.
\end{aligned}$$

□

Étudiez les lignes 42–87 du fichier de script `buhlmann.R` reproduit à la section 4.6 et comparez les résultats avec ceux de l'exemple 4.4.

4.5 Interprétation des résultats

Nous nous attardons principalement au facteur de crédibilité

$$z = \frac{n}{n+K}, \quad K = \frac{s^2}{a} = \frac{E[\sigma^2(\Theta)]}{\mathrm{Var}[\mu(\Theta)]}.$$

Le facteur de crédibilité augmente — un plus grand poids est donné à l'expérience individuelle — dans les situations suivantes :

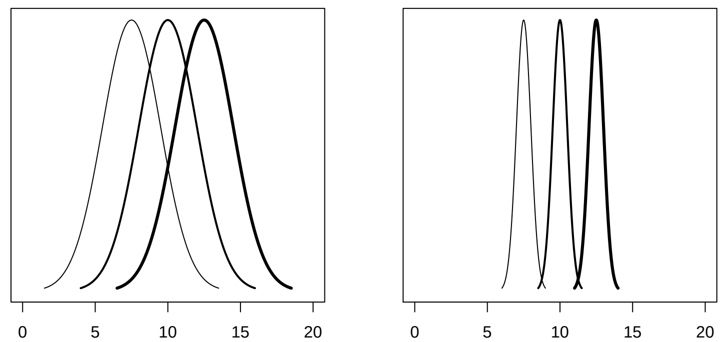


FIG. 4.1 – Effet de $s^2 = E[\sigma^2(\Theta)]$ sur le facteur de crédibilité. Gauche : grand s^2, l'expérience est trop volatile pour être fiable. Droite : petit s^2, les moyennes individuelles sont fiables. Les moyennes sont identiques dans les deux graphiques.

1. le nombre d'années d'expérience est grand, $n \to \infty$. À long terme, l'expérience d'un contrat représente exactement son niveau de risque. C'est la même situation qu'en crédibilité de stabilité, c'est-à-dire que le niveau de crédibilité augmente avec le volume d'expérience ;
2. le paramètre s^2 est petit, $s^2 \to 0$, l'expérience est globalement stable dans le temps. Les moyennes $\bar{S}_i$ représentent alors bien les niveaux de risque des contrats, ce qui réduit l'utilité de la prime collective.
3. le paramètre a est grand, $a \to \infty$, le portefeuille est hétérogène. Dans un tel cas, les moyennes individuelles sont de meilleures approximations des primes de risque que la prime collective. On notera au passage que a est en général le paramètre le plus intéressant et celui qui fluctue le plus d'un portefeuille à un autre.

Les figures 4.1 et 4.2 illustrent les points 2 et 3 ci-dessus. Chaque courbe représente l'expérience d'un contrat. Dans les deux cas, le facteur de crédibilité est plus grand dans le graphique de droite.

Si s^2 et a varient en des directions opposées, il devient difficile d'interpréter les résultats.

Lorsque toutes les moyennes individuelles sont identiques, $\bar{S}_1 = \dots = \bar{S}_I = \bar{S}$, alors $a = 0$ et la prime de crédibilité est $\pi^{\mathrm{B}}_{i,n+1} = \bar{S} = \bar{S}_i$ pour chaque contrat. Autrement dit, les primes individuelles et la prime collective sont identiques. Il serait donc numériquement équivalent d'octroyer une crédibilité de 1 à l'expérience individuelle. Comment, alors, justifier un facteur de

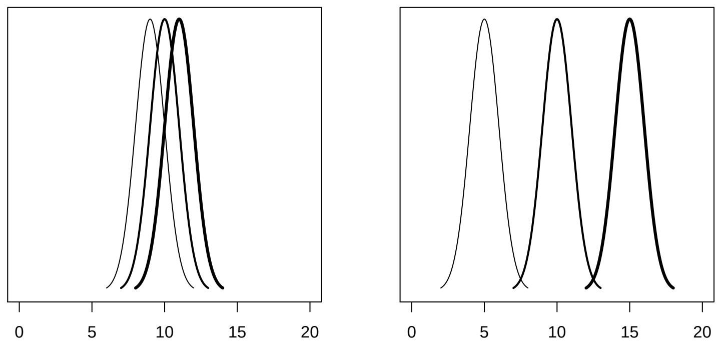


FIG. 4.2 – Effet de $a = \mathrm{Var}[\mu(\Theta)]$ sur le facteur de crédibilité. Gauche : petit a, le portefeuille est homogène. Droite : grand a, le portefeuille est hétérogène. Les variances sont identiques dans les deux graphiques.

crédibilité de 0 ? Tout simplement par le fait que, dans une telle situation, il n'est pas nécessaire de faire de la tarification basée sur l'expérience.

Parce qu'il s'agit d'une étape importante dans l'étude des modèles de crédibilité, l'interprétation des résultats du modèle de Bühlmann fait également l'objet d'une vidéo. Celle-ci utilise un contexte différent de celui de l'assurance qui peut aider à la bonne compréhension des concepts.

4.6 Code informatique

Fichier d'accompagnement `buhlmann.R`

```
###
### APPROCHE PARAMÉTRIQUE
###

## Charger le package actuar dans la session de travail.
library(actuar)

## Les primes bayésiennes linéaires et les primes de
## crédibilité de Bühlmann sont identiques dans l'approche
## paramétrique.
##
## Tel que mentionné dans le texte du chapitre, 'cm' effectue
```

```
## le calcul des premières à partir des formules des secondes.
## Comparez les résultats ci-dessous aux formules des
## combinaisons de distribution Poisson/gamma et
## exponentielle/gamma.
cm("bayes", likelihood = "poisson",
   shape = 3, rate = 2)
cm("bayes", likelihood = "exponential",
   shape = 3, rate = 2)

## Un coup d'oeil au code source de la fonction 'bayes' révèle
## le secret de fabrication. Comme cette fonction est interne
## au paquetage (autrement dit, vous ne pouvez l'appeler
## directement), vous devez pour y accéder ajouter le préfixe
## 'actuar:::' devant le nom de l'objet pour afficher son
## contenu.
actuar:::bayes

###
### APPROCHE NON PARAMÉTRIQUE
###

## Nous illustrons ici comment utiliser la fonction 'cm' pour
## calculer des primes de crédibilité pour des modèles non
## paramétriques.
##
## L'interface de 'cm' pour ce type de modèle est fortement
## inspirée de 'lm', la fonction d'ajustement de modèles
## linéaires de R.
##
## La fonction suppose d'abord que les données se trouvent
## dans une matrice ou dans un data frame à raison d'une ligne
## par contrat. La matrice ou le data frame doit également
## comporter une colonne (nommée) pour identifier les contrats
## à l'aide d'une valeur numérique ou texte.
(x <- data.frame(contract = 1:3,
                 matrix(c(0, 3, 3,
                          1, 4, 3,
                          2, 2, 2,
                          1, 1, 1,
                          2, 4, 2,
                          0, 4, 1), nrow = 3)))

## Les arguments de 'cm' obligatoires pour l'ajustement du
## modèle de Bühlmann sont les suivants:
##
```

```
## formula: formule de la forme '~ termes' où, dans le modèle
##   de Bühlmann, 'termes' est simplement le nom de la colonne
##   des données contenant les identifiants des contrats;
## data: les données
##
## Si la matrice ou le data frame contient des données autres
## que celles servant à la modélisation, l'argument 'ratios'
## permet d'indiquer dans quelles colonnes se trouvent les
## données. Par défaut, la fonction considérera que toutes les
## colonnes (autres que celle présente dans la formule)
## contiennent des données.
(fit <- cm(~ contract, x))                 # appel simple
(fit <- cm(~ contract, x, ratios = X1:X6)) # équivalent ici

## Calcul des primes de crédibilité.
predict(fit)

## Résultats détaillés.
summary(fit)
```

4.7 Exercices

4.1 Un assureur couvre une proportion égale d'hommes et de femmes contre les accidents d'automobile. Les femmes ont une probabilité d'accident de 20 % par année alors que les hommes ont une probabilité d'accident de 40 % par année. La distribution de la sévérité des sinistres est la suivante :

$$\Pr[X = x] = \begin{cases} 0{,}8, & x = 100 \\ 0{,}1, & x = 200 \\ 0{,}1, & x = 400. \end{cases}$$

On suppose qu'une personne ne peut avoir qu'un seul accident par année. Calculer le facteur de crédibilité pour un assuré selon le modèle de Bühlmann.

4.2 Vous roulez un dé régulier. Si vous obtenez 1, vous pigez dans l'urne A, si vous obtenez 2, 3 ou 4, vous pigez dans l'urne B et si vous obtenez 5 ou 6, vous pigez dans l'urne C. Chaque urne contient des balles rouges et blanches dans les proportions ci-dessous.

	Balles rouges	Balles blanches
Urne A	75 %	25 %
Urne B	50 %	50 %
Urne C	25 %	75 %

Vous roulez le dé et pigez dans l'urne correspondante cinq balles avec remise dont vous avez pris la couleur en note. Vous emmenez ensuite l'urne en question à un ami et il y pige à son tour cinq balles, toujours avec remise. Si vous ne vous souvenez plus ni du nombre indiqué par le dé, ni de quelle urne les balles ont été tirées, mais que vous avez toujours en note que vous avez pigé trois balles rouges, quelle est l'estimation du nombre de balles rouges que pigera votre ami selon le modèle de crédibilité de Bühlmann ?

4.3 On vous donne les informations suivantes sur un portefeuille de risques indépendants.

i) Les risques sont de deux types : type A et type B.

ii) Le nombre de risques de type A est le même que le nombre de risques de type B.

iii) Pour chaque risque, la probabilité d'avoir exactement un sinistre au cours d'une année est de 20 %, alors que la probabilité de n'avoir aucun sinistre est de 80 %.

iv) Le montant des sinistres des risques de type A est 2.

v) Le montant des sinistres des risques de type B est c, une constante inconnue.

Un risque est choisi au hasard au sein de ce portefeuille et le montant total des sinistres de ce risque est observé pour la première année. Vous souhaitez estimer le montant total des sinistres espéré pour la seconde année.

a) Quel est le modèle utilisé ici ?

b) Déterminer la limite du facteur de crédibilité dans le modèle de Bühlmann lorsque c tend vers l'infini.

4.4 Un portefeuille d'assurance est composé de 25 % de bons risques, 60 % de risques moyens et 15 % de mauvais risques. Tous les risques ont une distribution de sinistres de type gamma, mais dont les paramètres diffèrent selon le tableau ci-dessous.

Type de risque	α	λ
Bon	4	2
Moyen	4	1
Mauvais	10	2

Le dossier de sinistre d'un risque choisi au hasard est de 1 et 2 au cours des deux premières années. Calculer la prime de crédibilité de ce risque pour la troisième année selon le modèle de Bühlmann.

4.5 Un portefeuille d'assurance compte quatre classes d'assurés. Les caractéristiques de la distribution du montant total des sinistres annuels d'un assuré pour chacune des classes sont données dans le tableau suivant.

		Montant total des sinistres	
Classe	Nombre d'assurés	Moyenne	Variance
A	1 000	50	100 000
B	2 000	200	500 000
C	1 000	500	500 000
D	1 000	1 000	500 000

Calculer la prime de crédibilité chargée à un assuré ayant subi pour 800 $ de sinistres en quatre ans.

4.6 Un portefeuille d'assurance est composé de 30 % de bons risques, 50 % de risques moyens et 20 % de mauvais risques. La distribution des montants de sinistres est telle que présentée dans le tableau ci-dessous.

Type de risque	Distribution des sinistres
Bons	$\text{Pareto}(4, 1\,200)$
Moyens	$\text{Gamma}(1\,000, 2)$
Mauvais	$\text{Exponentielle}(0{,}00125)$

Selon le modèle de Bühlmann, quel facteur de crédibilité convient-il d'accorder après quatre années à un risque issu de ce portefeuille ?

4.7 La sévérité des sinistres pour un certain type d'assurance obéit à une loi gamma de paramètre de forme $\alpha = 5$ et de paramètre d'échelle θ variable par assuré. Les bons assurés ont une prime de risque de 2 500 $, les assurés moyens, une prime de risque de 4 000 $, et les mauvais assurés une prime de risque de 5 000 $. L'actuaire responsable de la tarification pour ce type d'assurance estime que le portefeuille est composé de 30 % de bons assurés, de 50 % de moyens et de 20 % de mauvais. Calculer à l'aide du modèle de Bühlmann la prime de crédibilité d'un assuré ayant encouru un total de 20 000 $ de sinistres au cours des quatre années précédentes.

4.8 Soit $N_t|\Theta = \theta \sim$ Géométrique(θ), où la fonction de densité de probabilité de Θ est

$$u(\theta) = 30(\theta^2 - 2\theta^3 + \theta^4), \quad 0 < \theta < 1.$$

a) Calculer $E[N_t]$.

b) Sachant que $N_1 = 0$, estimer N_2 par l'approche bayésienne pure.

c) Répéter la partie b) à l'aide du modèle de Bühlmann.

4.9 Déterminer la prime de crédibilité dans le modèle de Bühlmann selon les hypothèses suivantes :

a) $S|\Theta \sim$ Bernoulli(Θ), $\Theta \sim$ Bêta(a, b).

b) $S|\Theta \sim$ Géométrique(Θ), $\Theta \sim$ Bêta(a, b).

c) $S|\Theta \sim$ Gamma(τ, Θ), $\Theta \sim$ Gamma(α, λ).

d) $S|\Theta \sim$ Binomiale négative(r, Θ), $\Theta \sim$ Bêta(a, b).

e) $S|\Theta \sim$ Normale$(5\Theta, \sigma^2)$, $\Theta \sim U(a, b)$.

f) $S|\Theta \sim$ Exponentielle(Θ), $\Theta^{-1} \sim$ Bêta(a, b).

4.10 a) Soit $X_1, \dots, X_n$ un échantillon aléatoire et Y une variable aléatoire dont le second moment existe. Trouver les constantes α et β qui minimisent

$$E[(Y - \alpha - \beta\bar{X})^2],$$

où $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$. Ce problème est analogue à celui des moindres carrés ordinaires.

b) Considérer maintenant le modèle en théorie de la crédibilité de précision, soit où $X_1, \dots, X_n$ sont des variables aléatoires conditionnellement indépendantes sachant Θ. Trouver $\mathrm{Cov}(\mu(\Theta), \bar{X})$ et $\mathrm{Var}[\bar{X}]$, où $\mu(\Theta) = E[X_t|\Theta]$.

c) Toujours dans le même contexte qu'en b), utiliser le résultat en a) pour trouver les constantes α et β qui minimisent

$$E[(\mu(\Theta) - \alpha - \beta\bar{X})^2].$$

4.11 On vous donne les informations suivantes :

i) X_i est le nombre de sinistres du conducteur i au cours d'une année.

ii) La distribution de X_i est une binomiale négative de paramètres r_i et $p = 0{,}4$.

iii) μ_i est l'espérance du nombre de sinistres du conducteur i pour une année.

iv) La distribution de μ_i est une exponentielle de moyenne 0,2.

Déterminer le facteur de crédibilité d'un conducteur ayant une année d'expérience selon le modèle de Bühlmann.

4.12 Un certain groupe de conducteurs a une fréquence espérée de sinistres (par année) distribuée uniformément entre 0,1 et 0,3. Le nombre de sinistres observés (par année) pour chaque conducteur obéit à une loi de Poisson. Un conducteur du groupe a eu trois accidents au cours des cinq dernières années. Estimer, à l'aide du modèle de Bühlmann, la fréquence future des sinistres pour cet assuré.

4.13 La distribution conditionnelle du montant des sinistres est normale de moyenne Θ et de variance $4\Theta^2$. La distribution de la variable aléatoire $\sigma^2(\Theta)/\mu(\Theta)$ est uniforme sur l'intervalle $(0, 40)$. Calculer la prime de crédibilité pour la cinquième année dans le modèle de Bühlmann si la somme des sinistres des quatre premières années est égale à 12.

4.14 En utilisant le modèle de Bühlmann et les hypothèses de distributions ci-dessous, trouver la prime de crédibilité pour la sixième année sachant qu'un assuré a eu au cours des cinq premières années les montants de sinistre suivants : 3, 1, 5, 4, 2.

a) $S|\Theta \sim \text{Poisson}(\theta)$ et

   i) $\Theta \sim \text{Gamma}(10, 5)$.

   ii) $\Theta \sim \text{Gamma}(50, 25)$.

   iii) $\Theta \sim \text{Gamma}(\frac{1}{2}, \frac{1}{4})$.

b) $S|\Theta \sim U(0, 2\theta)$ et

   i) $\Theta \sim \text{Gamma}(10, 5)$.

   ii) $\Theta \sim \text{Gamma}(50, 25)$.

   iii) $\Theta \sim \text{Gamma}(\frac{1}{2}, \frac{1}{4})$.

4.15 On pose $S_j|\Theta = \theta \sim N(\theta, \theta^2)$ et $\Theta \sim N(150, 4\,500)$. Trouver, selon le modèle de Bühlmann, la prime de crédibilité pour la cinquième année si $S_1 = 100$, $S_2 = 125$, $S_3 = 75$ et $S_4 = 120$.

4.16 Une actuaire souhaite calculer la crédibilité de l'expérience de chaque police de deux portefeuilles d'assurance en utilisant le modèle de Bühlmann. Pour le portefeuille A, la variance de la prime de risque représente 25 % de la variance du montant des sinistres. Pour le portefeuille B, le montant des sinistres pour un assuré quelconque suit une loi normale de moyenne μ et de variance 25. La prime de risque est elle-même distribuée dans la population selon une loi normale de moyenne α et variance 16. Quel portefeuille mérite la plus grande crédibilité ?

4.17 a) Soit $S|\Theta \sim \text{Poisson}(\Theta)$ avec $\Theta \sim \text{Exponentielle}(\lambda)$. Calculer la constante K du modèle de Bühlmann sachant que la probabilité que la prime de risque soit supérieure à 5 est e^{-10}.

b) Soit maintenant $S_1|\Theta_1 \sim \text{Poisson}(\Theta_1)$ avec $\Theta_1 \sim \text{Exponentielle}(\alpha)$, et $S_2|\Theta_2 \sim \text{Poisson}(\Theta_2)$ avec $\Theta_2 \sim \text{Exponentielle}(\lambda)$. On suppose de plus que

i) S_1 est indépendante de S_2 et de Θ_2 ;

ii) S_2 est indépendante de S_1 et de Θ_1 ;

iii) Θ_1 et Θ_2 sont indépendantes.

Trouver la constante de crédibilité K pour $S = S_1 + S_2$.

4.18 Pour un groupe d'assurés détenant une police d'assurance contre le vol, un actuaire sait que 25 % de la variance totale du montant des sinistres est expliquée par la variance de la prime de risque. Après n années d'observation, un membre du groupe a eu des vols annuels de 125 $ en moyenne, et il a payé une prime de crédibilité de 110 $. Après $n + 1$ années, sa moyenne est passée à 150 $ et il a payé une prime de crédibilité de 125 $. À partir de ces données, quelle prime de crédibilité l'actuaire devra-t-il charger à un autre membre du groupe pour l'année $n + 1$ si celui-ci a eu des vols totalisant 500 $ après n années ?

4.19 La distribution conditionnelle de $S|\Theta$ n'est pas connue, mais on sait que $\mu(\Theta)$ suit une loi Gamma$(6, 4)$. De plus, $\sigma^2(\Theta) = \mu(\Theta)^2$. Le tableau ci-dessous contient les primes de crédibilité selon le modèle de Bühlmann.

	Montant total des sinistres						
Années d'expérience	0	1	2	3	4	5	6
2	$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$	$a_{2,4}$	$a_{2,5}$	$a_{2,6}$
3	$a_{3,0}$	$a_{3,1}$	$a_{3,3}$	$a_{3,3}$	$a_{3,4}$	$a_{3,5}$	$a_{3,6}$
4	$a_{4,0}$	$a_{4,1}$	$a_{4,4}$	$a_{4,3}$	$a_{4,4}$	$a_{4,5}$	$a_{4,6}$

Lesquels des énoncés suivants sont vrais ?

I. $a_{2,3} \leq a_{3,5}$

II. $a_{2,2} \leq a_{4,4}$

III. $a_{2,4} < a_{3,6}$

4.20 On vous donne les informations suivantes :

$$\Pr[S = x|\Theta = \theta] = e^{-\theta}\theta^x/x!, \quad x = 0, 1, \dots$$
$$u(\theta) = c\theta^2 e^{-4\theta}, \quad \theta > 0.$$

a) Calculer la prime collective.

b) Sachant que $S_1 = 4$, $S_2 = 1$ et $S_3 = 3$, calculer les primes de crédibilité pour les deuxième, troisième et quatrième années.

4.21 La distribution du montant total des sinistres est une exponentielle de paramètre de risque inconnu. La moyenne de cette distribution suit une loi gamma de moyenne 4 et de variance 8. Trouver la prime de crédibilité de troisième année dans le modèle de Bühlmann si les sinistres totaux au cours des deux premières années furent respectivement 1 et 3.

4.22 Démontrer que la prime de crédibilité dans le modèle de Bühlmann est de type « mise à jour » (Gerber et Jones, 1975), c'est-à-dire qu'elle peut s'écrire sous la forme

$$\pi_{n+1} = \zeta_n \pi_n + (1 - \zeta_n) S_n,$$

où $0 \leq \zeta_n \leq 1$ est une fonction de n, le nombre d'années d'expérience disponible, et π_n est la prime de crédibilité de l'année précédente.

4.23 Démontrer que l'estimateur $\hat{s}^2$ du modèle de Bühlmann est sans biais en calculant $E[(S_{it} - \bar{S}_i)^2]$ sans d'abord conditionner sur Θ_i.

4.24 Proposer un modèle pouvant servir pour la simulation de données dans une application du modèle de Bühlmann. Le modèle devrait :

i) respecter les deux principales hypothèses du modèle de Bühlmann, à savoir $E[S_{it}|\Theta_i] = \mu(\Theta_i)$ et $\mathrm{Var}[S_{it}|\Theta_i] = \sigma^2(\Theta_i)$;

ii) permettre de simuler des montants individuels de sinistres (et non pas seulement le montant total).

4.25 À partir des données de sinistres ci-dessous, calculer les primes de crédibilité de Bühlmann de l'année 7 pour chacun des trois contrats.

	Années					
Contrat	1	2	3	4	5	6
1	0	1	2	1	2	0
2	3	4	2	1	4	4
3	3	3	2	1	2	1

5 Modèle de Bühlmann–Straub

Objectifs du chapitre

- Comprendre l'utilité des poids dans le modèle de Bühlmann-Straub et leur implication sur les hypothèses du modèle.
- Calculer la covariance entre des variables aléatoires du modèle de Bühlmann-Straub, notamment entre diverses moyennes pondérées.
- Minimiser l'espérance quadratique moyenne dans le cadre du modèle de Bühlmann-Straub.
- Calculer l'espérance des estimateurs des paramètres de structure dans le modèle de Bühlmann-Straub.
- Justifier le choix des estimateurs des paramètres de structure.
- Calculer une prime de crédibilité de Bühlmann-Straub à partir de données de sinistres.
- Utiliser la fonction `cm` du paquetage R **actuar** pour calculer des primes de crédibilité de Bühlmann-Straub et interpréter les résultats de la fonction.

Le modèle de Bühlmann–Straub (Bühlmann et Straub, 1970) est une généralisation du modèle de Bühlmann tenant compte de l'exposition au risque des contrats. Ceci est particulièrement important dans les situations où la taille des contrats varie beaucoup.

Par exemple, en accidents de travail, l'exposition au risque d'une entreprise de 1 000 employés est beaucoup plus grande que celle d'une autre ne comptant que 10 employés. Toutes autres choses étant égales par ailleurs, l'expérience de la plus grande entreprise devrait se montrer plus stable dans le temps et, par conséquent, se mériter une plus grande crédibilité dans un système de tarification basée sur l'expérience. Or, avec le modèle de Bühlmann qui octroie la crédibilité uniquement sur la base du nombre d'années d'expérience, on attribue la même crédibilité aux deux entreprises, même si leur taille diffère grandement.

Tab. 5.1 – Représentation schématique des variables aléatoires dans le modèle de Bühlmann-Straub

Variables non observables	Observations					Poids				
	1	…	t	…	n	1	…	t	…	n
Θ_1	X_{11}	…	X_{1t}	…	X_{1n}	w_{11}	…	w_{1t}	…	w_{1n}
$\vdots$	$\vdots$				$\vdots$	$\vdots$				$\vdots$
Θ_i	X_{i1}	…	X_{it}	…	X_{in}	w_{i1}	…	w_{it}	…	w_{in}
$\vdots$	$\vdots$				$\vdots$	$\vdots$				$\vdots$
Θ_I	X_{I1}	…	X_{It}	…	X_{In}	w_{I1}	…	w_{It}	…	w_{In}

5.1 Modèle et prévision

Dans la forme la plus générale du modèle de Bühlmann–Straub, on associe un poids w_{it} à chaque donnée, qui sera maintenant notée X_{it}. La représentation schématique du portefeuille de données se trouve au tableau 5.1.

Afin de refléter le fait que l'expérience d'un contrat avec une grande exposition au risque sera plus stable dans le temps que celle dont l'exposition au risque est plus réduite, nous devons modifier l'hypothèse de variances conditionnelles identiques du modèle de Bühlmann. Les hypothèses du modèle de Bühlmann–Straub sont les suivantes.

(BS1) Les contrats $(\Theta_i, \mathbf{X}_i)$, $i = 1, \dots, I$ sont indépendants, les variables aléatoires $\Theta_1, \dots, \Theta_I$ sont identiquement distribuées et les variables aléatoires X_{it} ont une variance finie.

(BS2) Les variables aléatoires X_{it}, sont telles que

$$\begin{aligned} E[X_{it}|\Theta_i] &= \mu(\Theta_i) \quad i = 1, \dots, I \\ \operatorname{Cov}(X_{it}, X_{iu}|\Theta_i) &= \delta_{tu}\frac{\sigma^2(\Theta_i)}{w_{it}}, \quad t, u = 1, \dots, n. \end{aligned}$$

Nous postulons à l'hypothèse (BS2) que la variance de l'observation X_{it} est inversement proportionnelle au poids qui lui est rattaché :

$$\operatorname{Var}[X_{it}|\Theta_i] = \frac{\sigma^2(\Theta_i)}{w_{it}}.$$

Cette relation implique que les variables aléatoires X_{it} dans le modèle de Bühlmann–Straub sont des *ratios*.

La définition la plus usuelle des ratios X_{it} est

$$X_{it} = \frac{S_{it}}{w_{it}}$$

où, par exemple :

- S_{it} est le montant total des sinistres et w_{it} est la prime totale payée (*loss ratio*) ;
- S_{it} est le montant total des sinistres et w_{it} est la masse salariale ;
- S_{it} est le nombre d'accidents dans une flotte de véhicules et w_{it} est le nombre de véhicules ;
- etc.

> Le modèle de Bühlmann–Straub permet que l'exposition au risque varie par contrat et par période, mais ne l'oblige pas. L'exposition au risque pourrait être différente d'une période à l'autre, mais identique pour tous les contrats à l'intérieur d'une période. Inversement, l'exposition au risque peut être stable dans le temps pour un contrat, mais différente d'un contrat à l'autre. C'est seulement lorsque l'exposition au risque est la même à la fois dans le temps et entre les contrats que le modèle de Bühlmann–Straub se ramène au modèle de Bühlmann.

5.1.1 Notation et relations de covariance

Avec l'introduction des poids dans notre modèle, nous devrons dorénavant calculer des moyennes pondérées plutôt que des moyennes arithmétiques. Nous aurons recours à la notation suivante :

$$\begin{aligned}
X_{iw} &= \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} X_{it} & w_{i\Sigma} &= \sum_{t=1}^{n} w_{it} \\
X_{ww} &= \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} X_{iw} = \sum_{i=1}^{I} \sum_{t=1}^{n} \frac{w_{it}}{w_{\Sigma\Sigma}} X_{it} & w_{\Sigma\Sigma} &= \sum_{i=1}^{I} w_{i\Sigma} = \sum_{i=1}^{I} \sum_{t=1}^{n} w_{it} \\
X_{zw} &= \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} X_{iw} & z_\Sigma &= \sum_{i=1}^{I} z_i.
\end{aligned}$$

Le théorème suivant établit des relations de covariance analogues à celles du théorème 4.2 pour le modèle de Bühlmann–Straub. Ces relations sont nécessaires pour calculer la meilleure prévision de la prime de risque ainsi que l'espérance des estimateurs des paramètres de structure.

Théorème 5.1. *Soit X_{it}, $i = 1, \dots I$, $t = 1, \dots, n$ des variables aléatoires satisfaisant les hypothèses (BS1) et (BS2) ci-dessus. Alors,*

$$\begin{aligned}
\operatorname{Cov}(X_{it}, X_{ku}) &= \delta_{ik}\left(a + \delta_{tu}\frac{s^2}{w_{it}}\right) \\
\operatorname{Cov}(\mu(\Theta_i), X_{ku}) &= \delta_{ik} a \\
\operatorname{Cov}(X_{it}, X_{kw}) &= \delta_{ik}\left(a + \frac{s^2}{w_{i\Sigma}}\right).
\end{aligned}$$

Démonstration. Tout d'abord, toutes les covariances sont nulles lorsque $i \neq k$ par indépendance entre les contrats. Les deux premiers résultats sont équivalents à ceux du théorème 4.2 à la seule différence que

$$\begin{aligned}
E[\operatorname{Var}[X_{it}|\Theta_i]] &= E\left[\frac{\sigma^2(\Theta_i)}{w_{it}}\right] \\
&= \frac{s^2}{w_{it}}.
\end{aligned}$$

Pour le troisième résultat, nous avons

$$\begin{aligned}
\operatorname{Cov}(X_{it}, X_{iw}) &= \sum_{u=1}^{n} \frac{w_{iu}}{w_{i\Sigma}} \operatorname{Cov}(X_{it}, X_{iu}) \\
&= \sum_{u=1}^{n} \frac{w_{iu}}{w_{i\Sigma}}\left(a + \delta_{tu}\frac{s^2}{w_{it}}\right) \\
&= a + \frac{w_{it}}{w_{i\Sigma}}\frac{s^2}{w_{it}} \\
&= a + \frac{s^2}{w_{i\Sigma}}.
\end{aligned}$$

□

5.1.2 Meilleure prévision linéaire

Tout comme dans le modèle de Bühlmann, nous recherchons la meilleure approximation linéaire de la prime de risque d'un contrat.

Théorème 5.2. *Pour un portefeuille tel qu'illustré précédemment et sous les hypothèses (BS1) et (BS2), la meilleure approximation linéaire non homogène de la prime de risque $\mu(\Theta_i)$ — ou de $X_{i,n+1}$ — est*

$$\pi_{i,n+1}^{BS} = z_i X_{iw} + (1 - z_i) m \tag{5.1}$$

où

$$z_i = \frac{w_{i\Sigma}}{w_{i\Sigma} + K}, \quad K = \frac{s^2}{a}. \tag{5.2}$$

Démonstration. La démonstration est en tous points similaire à celle du théorème 4.3. Comme précédemment, nous pouvons nous restreindre à trouver les constantes $c_0, c_1, \dots, c_n$ qui minimisent l'erreur quadratique moyenne

$$E\left[\left(\mu(\Theta_i) - c_0 - \sum_{t=1}^{n} c_t X_{it}\right)^2\right].$$

En calculant les dérivées partielles, d'abord par rapport à c_0, puis par rapport à c_u, $u = 1, \dots, n$, nous obtenons les équations normales

$$c_0 = E[\mu(\Theta_i)] - \sum_{t=1}^{n} c_t E[X_{it}] \tag{5.3}$$

$$\operatorname{Cov}(\mu(\Theta_i), X_{iu}) = \sum_{t=1}^{n} c_t \operatorname{Cov}(X_{it}, X_{iu}). \tag{5.4}$$

Or, l'équation (5.4) peut se réécrire

$$\begin{aligned} a &= \sum_{t=1}^{n} c_t \left(a + \delta_{tu} \frac{s^2}{w_{iu}}\right) \\ &= a c_\Sigma + c_u \frac{s^2}{w_{iu}}. \end{aligned}$$

Par symétrie, les ratios c_u / w_{iu}, $u = 1, \dots, n$, sont tous égaux et, par conséquent, ils sont aussi égaux à $\sum_{u=1}^{n} c_u / \sum_{u=1}^{n} w_{iu}$:

$$\frac{c_1}{w_{i1}} = \frac{c_2}{w_{i2}} = \dots = \frac{c_n}{w_{in}} = \frac{c_\Sigma}{w_{i\Sigma}} = R_i.$$

Nous avons donc

$$a = a w_{i\Sigma} R_i + R_i s^2,$$

d'où

$$R_i = \frac{a}{a w_{i\Sigma} + s^2} = \frac{c_t}{w_{it}}$$

et

$$\begin{aligned}c_t &= \frac{w_{it}}{w_{i\Sigma}} \frac{a w_{i\Sigma}}{a w_{i\Sigma} + s^2} \\ &= \frac{w_{it}}{w_{i\Sigma}} z_i.\end{aligned}$$

Enfin, de l'équation (4.2),

$$\begin{aligned}c_0 &= m - \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} z_i m \\ &= (1 - z_i) m\end{aligned}$$

et, donc,

$$\begin{aligned}\pi_{i,n+1}^{\mathrm{BS}} &= c_0 + \sum_{t=1}^{n} c_t X_{it} \\ &= (1 - z_i) m + \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} z_i X_{it} \\ &= z_i X_{iw} + (1 - z_i) m.\end{aligned}$$

□

5.2 Estimation des paramètres de structure

Nous nous intéressons uniquement à l'approche non paramétrique avec le modèle de Bühlmann–Straub. Le calcul des primes de crédibilité requiert donc d'estimer les mêmes paramètres de structure qu'auparavant, soit la prime collective m, la variance intra s^2 et la variance inter a.

5.2.1 Estimation de la prime collective

Un estimateur intuitif de la prime collective m est

$$X_{ww} = \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} X_{iw}.$$

Or, on peut démontrer qu'en théorie de la crédibilité, l'estimateur linéaire à variance minimale est plutôt

$$\hat{m} = X_{zw} = \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} X_{iw}. \tag{5.5}$$

Formellement, X_{zw} n'est pas un estimateur puisqu'il dépend des paramètres inconnus s^2 et a. Nous appellerons de tels estimateurs des *pseudo-estimateurs* (De Vylder, 1981).

5.2.2 Estimation de la variance intra

Par une généralisation directe de l'estimateur obtenu dans le modèle de Bühlmann, nous obtenons un estimateur sans biais de la variance intra s^2 :

$$\hat{s}^2 = \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{t=1}^{n} w_{it}(X_{it} - X_{iw})^2. \tag{5.6}$$

5.2.3 Estimation de la variance inter

Après la discussion sur l'estimation de la variance inter au chapitre 4, nous pouvons soupçonner que l'estimateur intuitif

$$\sum_{i=1}^{I} w_{i\Sigma}(X_{iw} - X_{ww})^2$$

est biaisé. En effet, puisque $E[X_{iw}] = E[X_{ww}] = m$, il n'est pas très difficile de démontrer que

$$\begin{aligned} E[(X_{iw} - X_{ww})^2] &= \mathrm{Var}[X_{iw}] + \mathrm{Var}[X_{ww}] - 2\,\mathrm{Cov}(X_{iw}, X_{ww}) \\ &= a\left(1 - 2\,\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}}\right)^2\right) + s^2\left(\frac{1}{w_{i\Sigma}} - \frac{1}{w_{\Sigma\Sigma}}\right), \end{aligned}$$

d'où

$$E\left[\sum_{i=1}^{I} w_{i\Sigma}(X_{iw} - X_{ww})^2\right] = a\left(\frac{w_{\Sigma\Sigma}^2 - \sum_{i=1}^{I} w_{i\Sigma}^2}{w_{\Sigma\Sigma}}\right) + (I-1)s^2.$$

Un estimateur sans biais de la variance inter a est donc

$$\hat{a} = \frac{w_{\Sigma\Sigma}}{w_{\Sigma\Sigma}^2 - \sum_{i=1}^{I} w_{i\Sigma}^2}\left(\sum_{i=1}^{I} w_{i\Sigma}(X_{iw} - X_{ww})^2 - (I-1)\hat{s}^2\right). \tag{5.7}$$

Cet estimateur peut être négatif. La version tronquée à 0

$$\hat{a}' = \max(\hat{a}, 0),$$

est biaisée.

5.2.4 Pseudo-estimateur de la variance inter

L'estimateur de Bichsel–Straub du paramètre a est sans biais et toujours positif :

$$\tilde{a} = \frac{1}{I-1} \sum_{i=1}^{I} z_i (X_{iw} - X_{zw})^2. \tag{5.8}$$

Il n'y a toutefois rien de gratuit. D'abord, $\tilde{a}$ est un pseudo-estimateur puisqu'il dépend, via le facteur de crédibilité z_i, du paramètre a que nous cherchons justement à estimer. Ensuite, $\tilde{a}$ s'avère sans biais seulement lorsque nous considérons connus les facteurs de crédibilité. Autrement, l'espérance de l'estimateur est impossible à calculer.

Vous remarquerez qu'en pratique, l'estimateur est une fonction de lui-même : $\tilde{a} = f(\tilde{a})$. Nous devons donc effectuer le calcul de l'estimateur par la méthode du point fixe.

> On peut démontrer que si $\hat{a} < 0$, alors $\tilde{a}$ converge vers 0.

5.2.5 Sommaire des calculs

Vous trouverez ci-dessous sous forme d'algorithme les étapes de calcul des primes de crédibilité dans le modèle de Bühlmann–Straub en supposant que c'est l'estimateur de Bichsel–Straub (5.8) qui est utilisé pour estimer la variance inter.

1. Calculer $w_{i\Sigma}$, $i = 1, \dots, I$ et $w_{\Sigma\Sigma}$.
2. Calculer X_{iw}, $i = 1, \dots, I$ et X_{ww}.
3. Calculer $\hat{s}^2$.
4. Calculer $\hat{a}$.
5. Si $\hat{a} > 0$:
   i) calculer $\tilde{a}$ et poser $\hat{a} = \tilde{a}$;
   ii) calculer
   $$\hat{z} = \frac{w_{i\Sigma}}{w_{i\Sigma} + \hat{s}^2/\hat{a}};$$
   iii) calculer
   $$\hat{m} = \sum_{i=1}^{I} \frac{\hat{z}_i}{\hat{z}_\Sigma} X_{iw}.$$

   Sinon :

i) poser $\hat{a} = 0$;

ii) poser $\hat{z}_i = 0$, $i = 1, \dots, I$;

iii) poser $\hat{m} = X_{ww}$.

6. Calculer les primes de crédibilité

$$\hat{\pi}^{\text{BS}}_{i,n+1} = \hat{z}_i X_{iw} + (1 - \hat{z}_i)\hat{m}, \quad i = 1, \dots, I.$$

5.3 Données manquantes

Dans l'application du modèle de Bühlmann–Straub en pratique, il arrive fréquemment que le nombre d'observations ne soit pas le même pour tous les contrats. Les données et les poids sont alors disponibles pour $i = 1, \dots, I$ et $t = 1, \dots, n_i$ (en supposant les données contiguës). On aura donc, par exemple,

$$w_{i\Sigma} = \sum_{t=1}^{n_i} w_{it}$$

ou

$$X_{iw} = \sum_{t=1}^{n_i} \frac{w_{it}}{w_{i\Sigma}} X_{it}.$$

La seule formule affectée par ce changement est celle de l'estimateur de la variance intra, $\hat{s}^2$: (5.6) devient

$$\hat{s}^2 = \frac{1}{\sum_{i=1}^{I}(n_i - 1)} \sum_{i=1}^{I} \sum_{t=1}^{n} w_{it}(X_{it} - X_{iw})^2. \tag{5.9}$$

5.4 Exemple numérique

Les résultats de cette section sont tirés de Goovaerts et Hoogstad (1987), eux-mêmes basés sur les données de Hachemeister (1975). Ces données sont composées de montants de sinistres moyens au chapitre de la responsabilité civile en assurance automobile entre juillet 1970 et juin 1973 dans cinq États américains. Nous disposons donc des données de $I = 5$ contrats pendant $n = 12$ périodes d'expérience. Les montants de sinistres moyens X_{it} sont présentés au tableau 5.2 (à noter que le tableau est transposé par rapport à la notation usuelle).

TAB. 5.2 – Montants de sinistres moyens (ratios X_{it}) dans le portefeuille de Hachemeister

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
$t=1$	1738	1364	1759	1223	1456
$t=2$	1642	1408	1685	1146	1499
$t=3$	1794	1597	1479	1010	1609
$t=4$	2051	1444	1763	1257	1741
$t=5$	2079	1342	1674	1426	1482
$t=6$	2234	1675	2103	1532	1572
$t=7$	2032	1470	1502	1953	1606
$t=8$	2035	1448	1622	1123	1735
$t=9$	2115	1464	1828	1343	1607
$t=10$	2262	1831	2155	1243	1573
$t=11$	2267	1612	2233	1762	1613
$t=12$	2517	1471	2059	1306	1690

Le tableau 5.3 contient les poids w_{it} associés aux données. Il s'agit ici du nombre total de sinistres dans chaque période pour chaque état, soit le dénominateur des ratios X_{it}. Vous remarquerez que le nombre de sinistres est beaucoup plus élevé dans l'État 1, surtout, ainsi que dans l'État 5.

5.4.1 Résultats avec le modèle de Bühlmann

Pour fins de comparaison, calculons d'abord les primes de crédibilité avec le modèle de Bühlmann en ignorant simplement les poids rattachés aux données. Les estimateurs des paramètres de structure sont les suivants :

$$\begin{aligned}\hat{m} &= 1\,671 \\ \hat{s}^2 &= 46\,040 \\ \hat{a} &= 72\,310,\end{aligned}$$

ce qui mène aux résultats du tableau 5.4. Le facteur de crédibilité est plutôt élevé. Une analyse rapide des données suffit pour constater que l'expérience des états est relativement stable dans le temps. Il en résulte une valeur de $\hat{s}^2$ petite par rapport à celle de $\hat{a}$ et, donc, un grand facteur de crédibilité.

Pour l'État 1, dont l'expérience est la pire du portefeuille, un grand facteur de crédibilité a pour effet de ne réduire la prime de crédibilité (2 044) que de 1 % par rapport à la prime individuelle (2 064).

TAB. 5.3 – Nombres totaux de sinistres (poids w_{it}) dans le portefeuille de Hachemeister

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
$t=1$	7861	1622	1147	407	2902
$t=2$	9251	1742	1357	396	3172
$t=3$	8706	1523	1329	348	3046
$t=4$	8575	1515	1204	341	3068
$t=5$	7917	1622	998	315	2693
$t=6$	8263	1602	1077	328	2910
$t=7$	9456	1964	1277	352	3275
$t=8$	8003	1515	1218	331	2697
$t=9$	7365	1527	896	287	2663
$t=10$	7832	1748	1003	384	3017
$t=11$	7849	1654	1108	321	3242
$t=12$	9077	1861	1121	342	3425

TAB. 5.4 – Résultats avec le modèle de Bühlmann pour le portefeuille de Hachemeister

	$i=1$	$i=2$	$i=3$	$i=4$	$i=5$
Prime individuelle $\bar{X}_i$	2 064	1 511	1 822	1 360	1 599
Prime de crédibilité $\pi^{B}_{i,13}$	2 044	1 519	1 814	1 376	1 602
Facteur de crédibilité z	0,95	0,95	0,95	0,95	0,95

5.4.2 Résultats avec le modèle de Bühlmann–Straub

Le tableau 5.3 montre que le poids relatif de chacun des cinq états (poids mesuré en nombre de sinistres) est très différent : l'État 1 compte pour 57,5 % des sinistres du portefeuille, alors qu'à l'autre bout du spectre, l'État 4 ne compte que pour 2,4 %. Dans une telle situation, il convient d'utiliser le modèle de Bühlmann–Straub dans la tarification afin de tenir compte des volumes très différents d'un contrat à un autre.

Les estimateurs des paramètres de structure sont les suivants :

$$\begin{aligned}\hat{m} &= X_{zw} = 1\,689\\ \hat{s}^2 &= 139\,120\,026\\ \hat{a} &= 89\,639\\ \tilde{a} &= 64\,367.\end{aligned}$$

Tab. 5.5 – Résultats avec le modèle de Bühlmann-Straub pour le portefeuille de Hachemeister

	$i = 1$	$i = 2$	$i = 3$	$i = 4$	$i = 5$
Prime individuelle X_{iw}	2 061	1 511	1 806	1 353	1 600
Prime de crédibilité $\pi^{\mathrm{BS}}_{i,13}$	2 053	1 529	1 790	1 468	1 605
Facteur de crédibilité z	0,98	0,90	0,86	0,66	0,94

Ces résultats sont conformes aux étapes de calcul de la section 5.2.5, c'est-à-dire que l'estimateur X_{zw} de la moyenne collective ainsi que les primes de crédibilité du tableau 5.5 ont été calculés avec l'estimateur de Bichsel–Straub $\tilde{a}$ de la variance intra.

C'est pour l'État 4 que les différences entre les résultats des tableaux 5.4 et 5.5 s'avèrent les plus marquées. La prime de crédibilité de cet état augmente en effet de 1 376 à 1 468. Ceci est en partie dû à l'augmentation de l'estimateur de la prime collective, mais, surtout, à la forte baisse de son facteur de crédibilité. Le modèle de Bühlmann–Straub permet donc de reconnaître le rôle minime joué par cet état dans les résultats du portefeuille. C'est pourquoi le modèle y accorde peu de poids lors de la répartition des primes.

Le code du fichier de script `buhlmann-straub.R` reproduit à la section 5.5 permet d'obtenir les résultats des tableaux 5.4 et 5.5 avec la fonction `cm` du paquetage **actuar**. Étudiez attentivement ce code.

5.4.3 Limitations des modèles précédents

L'examen des données de l'État 1 montre que le montant moyen des sinistres va en augmentant d'une période à l'autre. Or, la prime de crédibilité du modèle de Bühlmann–Straub se trouve environ au niveau de la période $t = 8$. Il semble donc évident que la prime de crédibilité sera trop peu élevée.

C'est afin de pouvoir traiter de tels cas que Hachemeister (1975) a proposé son modèle de crédibilité avec régression. L'utilisation d'un tel modèle est particulièrement indiquée dans des situations de forte inflation ou d'augmentation ou de diminution structurelle des coûts.

La vignette `credibility` du paquetage **actuar** contient des informations succinctes sur le modèle de régression de Hachemeister. Pour un traitement plus exhaustif, consultez l'excellent ouvrage de Bühlmann et Gisler (2005).

5.5 Code informatique

Fichier d'accompagnement `buhlmann-straub.R`

```
## Charger le package actuar dans la session de travail.
library(actuar)

## Les données de Hachemeister (1975) sont fournies avec le
## paquetage actuar sous forme d'une matrice de 5 lignes et de
## 25 colonnes. La première colonne contient un identifiant
## d'État, les colonnes 2-13 les montants de sinistres moyens
## et les colonnes 14-25, les nombres de sinistres.
hachemeister

## Dès lors qu'un jeu de données comporte des colonnes autres
## que l'identifiant du contrat et les données de sinistres,
## il faut avoir recours aux arguments suivants de 'cm'.
##
## ratios: expression indiquant les colonnes dans lesquelles
##  se trouvent les données de sinistres;
## weights: expression indiquant les colonnes dans lesquelles
##  se trouvent les poids associés aux données de sinistres.
##
## Ces deux arguments sont utilisés comme l'argument 'select'
## de la fonction 'subset'. Cela permet d'indicer de manière
## très intuitive les colonnes du jeu de données.
##
## Par exemple, l'expression ci-dessous permet d'extraire les
## ratios des cinq premières années des données de
## Hachemeister.
subset(hachemeister, select = ratio.1:ratio.5)

## Ajustement du modèle de Bühlmann aux données de
## Hachemeister (sans tenir compte des volumes).
fit.B <- cm(~state, hachemeister,
            ratios = ratio.1:ratio.12)
summary(fit.B)

## L'argument 'method' de 'cm' permet de choisir l'estimateur
## de la variance inter. Trois choix sont disponibles.
##
## "Buhlmann-Gisler": notre estimateur sans biais;
## "Ohlsson": équivalent à "Buhlmann-Gisler" dans le modèle de
##   Buhlmann-Straub;
## "iterative": pseudo-estimateur de Bichsel-Straub.
```

```r
##
## La vignette "credibility" du paquetage contient tous les
## détails sur les méthodes d'estimation de la variance inter.
vignette("credibility")

## Ajustement du modèle de Bühlmann-Straub avec l'estimateur
## sans biais de la variance inter (la méthode par défaut,
## donc non spécifiée ici). Il faut spécifier des poids.
fit.BS1 <- cm(~state, hachemeister,
              ratios = ratio.1:ratio.12,
              weights = weight.1:weight.12)
summary(fit.BS1)

## Même chose, mais avec le pseudo-estimateur de la variance
## inter.
fit.BS2 <- cm(~state, hachemeister,
              ratios = ratio.1:ratio.12,
              weights = weight.1:weight.12,
              method = "iterative")
summary(fit.BS2)
```

5.6 Exercices

5.1 Démontrer les relations suivantes.

a) $\operatorname{Cov}(X_{iw}, X_{ww}) = a\,\dfrac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \dfrac{s^2}{w_{\Sigma\Sigma}}$

b) $\operatorname{Var}[X_{ww}] = a\sum_{i=1}^{I}\left(\dfrac{w_{i\Sigma}}{w_{\Sigma\Sigma}}\right)^2 + \dfrac{s^2}{w_{\Sigma\Sigma}}$.

c) $\operatorname{Var}[X_{zw}] = \dfrac{a}{z_\Sigma}$

5.2 Démontrer que l'estimateur $\hat{s}^2$ est sans biais.

5.3 Démontrer que l'estimateur $\hat{a}$ est sans biais.

5.4 Démontrer que l'estimateur $\tilde{a}$ est sans biais.

5.5 Pourquoi n'a-t-on pas un équivalent de l'estimateur $\tilde{a}$ dans le modèle de Bühlmann ? Démontrer que $\hat{a} = \tilde{a}$ dans le modèle de Bühlmann, c'est-à-dire lorsque tous les poids sont égaux.

5.6 À la section 5.2.5, il est précisé d'utiliser X_{ww} comme estimateur de la prime collective lorsque l'estimateur de la variance inter est nul. Justifier ce choix en démontrant que lorsque tous les facteurs de crédibilité sont nuls (situation qui survient principalement lorsque $a = 0$), alors $X_{zw} = X_{ww}$. Autrement dit, démontrer que

$$\lim_{a \to 0} X_{zw} = X_{ww}$$

5.7 Proposer un modèle pouvant servir pour la simulation de données dans une application du modèle de Bühlmann–Straub. Le modèle devrait

i) respecter les deux principales hypothèses du modèle de Bühlmann–Straub, à savoir $E[X_{it}|\Theta_i] = \mu(\Theta_i)$ et $\mathrm{Var}[X_{it}|\Theta_i] = \sigma^2(\Theta_i)/w_{it}$;

ii) permettre de simuler des montants individuels de sinistres (et non pas seulement le montant total).

Supposer les poids w_{it} connus.

5.8 Considérer le tableau suivant où S_{it} est le montant total des sinistres et w_{it} la masse salariale pour l'employeur i dans l'année t.

S_{it}

	Année			
Employeur	1	2	3	4
1	14	21	12	18
2	4	0	4	6
3	3	0	1	6

w_{it}

	Année			
Employeur	1	2	3	4
1	2	3	4	3
2	4	2	1	2
3	3	3	1	3

Calculer la prime de crédibilité (totale) de l'employeur 1 pour la cinquième année pour trois unités de masse salariale. Utiliser $\hat{a}$ comme estimateur de la variance inter.

5.9 Calculer la prime de crédibilité pour chacun des trois contrats du portefeuille ci-dessous.

Ratios X_{it}

	Année				
Contrat	1	2	3	4	5
1	3	5	–	–	4
2	6	8	8	14	4
3	–	2	0	3	6

Poids w_{it}

	Année				
Contrat	1	2	3	4	5
1	1	1	–	–	1
2	2	2	2	2	2
3	–	3	3	3	3

a) Faire les calculs « à la main ». Utiliser $\hat{a}$ comme estimateur du paramètre a.

b) Faire les calculs avec la fonction `cm` de **actuar**. Utiliser $\tilde{a}$ comme estimateur du paramètre a. Comparer les réponses avec celles obtenues en a).

5.10 On vous donne les informations suivantes :

$$
\begin{aligned}
K &= 15 & X_{zw} &= 0{,}8 \\
X_{1w} &= 0{,}89 & w_{1\Sigma} &= 8 \\
X_{2w} &= 0{,}85 & w_{2\Sigma} &= 6 \\
X_{3w} &= 0{,}65 & & \\
X_{4w} &= 0{,}70 & w_{4\Sigma} &= 9.
\end{aligned}
$$

Calculer le facteur de crédibilité du contrat 3.

A Estimation bayésienne

L'estimation ponctuelle par l'approche bayésienne diffère passablement des approches classiques comme la méthode des moments ou la méthode du maximum de vraisemblance. Cette annexe propose un résumé de la philosophie et l'estimation bayésienne ainsi que les principaux résultats.

A.1 Cas continu

Supposons que nous souhaitons estimer le paramètre inconnu θ d'une distribution continue avec fonction de densité de probabilité $f(x;\theta)$ (une loi normale avec moyenne inconnue θ et variance connue σ^2, par exemple) à partir d'un échantillon aléatoire $X_1, \dots, X_n$.

Les statisticiens classiques développeront des estimateurs à partir d'un critère objectif quelconque : absence de biais, maximum de vraisemblance, etc. Vous remarquerez que l'analyste ne pose aucune hypothèse à priori sur la valeur de θ. On « laisse parler les données ».

Dans l'approche bayésienne, l'opinion à priori d'une personne sur la valeur du paramètre θ est prise en compte dans l'estimation de ce dernier. Le truc consiste à considérer le paramètre comme une réalisation d'une variable aléatoire Θ avec densité $u(\theta)$ qui représente l'opinion de l'analyste quant à la valeur du paramètre. Au fur et à mesure que les données de l'échantillon aléatoire (l'information) s'accumulent, l'analyste améliore son opinion et, par conséquent, révise la distribution de la variable aléatoire Θ en calculant sa densité à postériori $u(\theta|x_1, \dots, x_n)$ à l'aide de la règle de Bayes :

$$\begin{aligned} u(\theta|x_1, \dots, x_n) &= \frac{f(x_1, \dots, x_n, \theta)}{f(x_1, \dots, x_n)} \\ &= \frac{f(x_1, \dots, x_n|\theta)u(\theta)}{f(x_1, \dots, x_n)}. \end{aligned}$$

Par la loi des probabilités totales,

$$f(x_1, \dots, x_n) = \int_{-\infty}^{\infty} f(x_1, \dots, x_n|\theta) u(\theta)\, d\theta,$$

et donc

$$u(\theta|x_1, \dots, x_n) = \frac{f(x_1, \dots, x_n|\theta) u(\theta)}{\int_{-\infty}^{\infty} f(x_1, \dots, x_n|\theta) u(\theta)\, d\theta}.$$

Le dénominateur du côté droit de l'expression ci-dessus est une constante de normalisation qui est souvent omise dans les calculs.

Pour obtenir un estimateur ponctuel $\hat{\theta} = g(X_1, \dots, X_n)$ du paramètre θ, nous minimisons l'espérance à postériori d'une *fonction de perte*. La fonction de perte la plus fréquemment employée est l'erreur quadratique

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2.$$

Or, l'estimateur bayésien qui minimise

$$E[L(\hat{\theta}, \theta)|X_1, \dots, X_n] = E[(\hat{\theta} - \theta)^2|X_1, \dots, X_n]$$

est

$$\begin{aligned}\hat{\theta} &= E[\Theta|X_1, \dots, X_n] \\ &= \int_{-\infty}^{\infty} \theta\, u(\theta|x_1, \dots, x_n)\, d\theta,\end{aligned}$$

soit l'espérance de Θ calculée par rapport à la distribution à postériori.

A.2 Cas discret

Les idées de la section précédente demeurent exactement les mêmes dans le cas discret ; seule la notation change légèrement. Pour simplifier la notation, posons $\mathbf{X} = (X_1, \dots, X_n)$ et $\mathbf{x} = (x_1, \dots, x_n)$.

Si la variable aléatoire Θ ne prend que des valeurs discrètes, la distribution à priori est exprimée sous forme d'une fonction de masse de probabilité $\Pr[\Theta = \theta]$. Nous pouvons toujours calculer la fonction de masse de probabilité conjointe de $X_1, \dots, X_n$ par la loi des probabilités totales :

$$\Pr[\mathbf{X} = \mathbf{x}] = \sum_{\theta=-\infty}^{\infty} \Pr[\mathbf{X} = \mathbf{x}|\Theta = \theta] \Pr[\Theta = \theta].$$

La règle de Bayes permet de calculer la distribution à postériori de Θ :

$$\begin{aligned}
\Pr[\Theta = \theta | \mathbf{X} = \mathbf{x}] &= \frac{\Pr[\mathbf{X} = \mathbf{x} | \Theta = \theta] \Pr[\Theta = \theta]}{\Pr[\mathbf{X} = \mathbf{x}]} \\
&= \frac{\Pr[\mathbf{X} = \mathbf{x} | \Theta = \theta] \Pr[\Theta = \theta]}{\sum_{\theta=-\infty}^{\infty} \Pr[\mathbf{X} = \mathbf{x} | \Theta = \theta] \Pr[\Theta = \theta]}.
\end{aligned}$$

Enfin, l'estimateur bayésien minimisant l'erreur quadratique moyenne demeure inchangé :

$$\begin{aligned}
\hat{\theta} &= E[\Theta | X_1, \dots, X_n] \\
&= \sum_{\theta=-\infty}^{\infty} \theta \Pr[\Theta = \theta | \mathbf{X} = \mathbf{x}].
\end{aligned}$$

A.3 Cas mixtes

Il est simple de dériver les formules des cas mixtes, les plus courants étant ceux où la distribution de $X|\Theta$ est discrète et celle de Θ, continue.

B Formules de crédibilité linéaire

Cette annexe contient les principaux résultats de crédibilité bayesienne pour les combinaisons de distributions de la famille exponentielle univariée. Outre les noms des distributions conditionnelles de $S|\Theta = \theta$ et à priori de Θ, on trouvera ci-dessous, pour chaque combinaison :

- la distribution à postériori de $\Theta|S_1 = x_1, \dots, S_n = x_n$, toujours du même type que la distribution à priori, mais avec des paramètres mis à jour;
- la prime de risque $\mu(\theta) = E[S|\Theta = \theta]$;
- la prime collective $m = E[\mu(\Theta)]$;
- la prime bayesienne $B_{n+1} = E[\mu(\Theta)|S_1, \dots, S_n]$, toujours égale à la prime collective évaluée avec les paramètres de la distribution à postériori;
- le facteur de crédibilité dans l'expression de la prime bayésienne sous forme de prime de crédibilité.

Les fonctions de masse ou de densité de probabilité des distributions se trouvent à l'annexe C.

B.1 Combinaison Bernoulli/bêta

$S|\Theta = \theta \sim \text{Bernoulli}(\theta)$
$\Theta \sim \text{Bêta}(a, b)$
$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Bêta}(\tilde{a}, \tilde{b})$

$$\tilde{a} = a + \sum_{t=1}^{n} x_t$$

$$\tilde{b} = b + n - \sum_{t=1}^{n} x_t$$

Prime de risque

$$\mu(\theta) = \theta$$

Prime collective

$$m = \frac{a}{a+b}$$

Prime bayésienne

$$B_{n+1} = \frac{a + \sum_{t=1}^{n} S_t}{a+b+n}$$

Facteur de crédibilité

$$z = \frac{n}{n+a+b}$$

B.2 Combinaison géométrique/bêta

$S|\Theta = \theta \sim \text{Géométrique}(\theta)$
$\Theta \sim \text{Bêta}(a, b)$
$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Bêta}(\tilde{a}, \tilde{b})$

$$\tilde{a} = a + n$$

$$\tilde{b} = b + \sum_{t=1}^{n} x_t$$

Prime de risque

$$\mu(\theta) = \frac{1-\theta}{\theta}$$

Prime collective

$$m = \frac{b}{a-1}$$

Prime bayésienne

$$B_{n+1} = \frac{b + \sum_{t=1}^{n} S_t}{a+n-1}$$

Facteur de crédibilité

$$z = \frac{n}{n+a-1}$$

B.3 Combinaison Poisson/gamma

$S|\Theta = \theta \sim \text{Poisson}(\theta)$
$\Theta \sim \text{Gamma}(\alpha, \lambda)$
$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Gamma}(\tilde{\alpha}, \tilde{\lambda})$

$$\tilde{\alpha} = \alpha + \sum_{t=1}^{n} x_t$$

$$\tilde{\lambda} = \lambda + n$$

Prime de risque

$$\mu(\theta) = \theta$$

Prime collective

$$m = \frac{\alpha}{\lambda}$$

Prime bayésienne

$$B_{n+1} = \frac{\alpha + \sum_{t=1}^{n} S_t}{\lambda + n}$$

Facteur de crédibilité

$$z = \frac{n}{n + \lambda}$$

B.4 Combinaison exponentielle/gamma

$S|\Theta = \theta \sim \text{Exponentielle}(\theta)$
$\Theta \sim \text{Gamma}(\alpha, \lambda)$
$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Gamma}(\tilde{\alpha}, \tilde{\lambda})$

$$\tilde{\alpha} = \alpha + n$$
$$\tilde{\lambda} = \lambda + \sum_{t=1}^{n} x_t$$

Prime de risque

$$\mu(\theta) = \frac{1}{\theta}$$

Prime collective

$$m = \frac{\lambda}{\alpha - 1}$$

Prime bayésienne

$$B_{n+1} = \frac{\lambda + \sum_{t=1}^{n} S_t}{\alpha + n - 1}$$

Facteur de crédibilité

$$z = \frac{n}{n + \alpha - 1}$$

B.5 Combinaison normale/normale

$S|\Theta = \theta \sim \text{Normale}(\theta, \sigma_2^2)$
$\Theta \sim \text{Normale}(\mu, \sigma_1^2)$

$\Theta|S_1 = x_1, \dots, S_n = x_n \sim \text{Normale}(\tilde{\mu}, \tilde{\sigma}_1^2)$

$$\tilde{\mu} = \frac{\sigma_1^2 \sum_{t=1}^{n} x_t + \sigma_2^2 \mu}{n\sigma_1^2 + \sigma_2^2}$$

$$\tilde{\sigma}_1^2 = \frac{\sigma_1^2 \sigma_2^2}{n\sigma_1^2 + \sigma_2^2}$$

Prime de risque

$$\mu(\theta) = \theta$$

Prime collective

$$m = \mu$$

Prime bayésienne

$$B_{n+1} = \frac{\sigma_1^2 \sum_{t=1}^{n} S_t + \sigma_2^2 \mu}{n\sigma_1^2 + \sigma_2^2}$$

Facteur de crédibilité

$$z = \frac{n}{n + \sigma_2^2/\sigma_1^2}$$

C Paramétrisation des lois de probabilité

Cette annexe précise la paramétrisation des lois de probabilité utilisée dans les énoncés des exercices des chapitres 2 à 5 ainsi que dans le tableau des formules de crédibilité exacte de l'annexe B; la racine des noms de fonctions pour calculer les caractéristiques des lois dans R; l'espérance, la variance et la fonction génératrice des moments (lorsqu'elle existe) des différentes lois.

La prise en charge de certaines lois dans R requiert le paquetage **actuar** (Dutang et collab., 2008).

C.1 Distributions discrètes

C.1.1 Binomiale

Racine : `binom`
Paramètres : `size` ($n = 0, 1, 2, \dots$), `prob` ($0 \leq \theta \leq 1$)
Cas spécial : Bernoulli(θ) lorsque $n = 1$.

$$\begin{aligned}
\Pr[X = x] &= \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, \dots \\
E[X] &= n\theta \\
\mathrm{Var}[X] &= n\theta(1-\theta) \\
M(t) &= (1 - \theta + \theta e^t)^n
\end{aligned}$$

C.1.2 Binomiale négative

Racine : `nbinom`
Paramètres : `size` ($r \geq 0$), `prob` ($0 < \theta \leq 1$), `mu` ($r(1-\theta)/\theta$)
Cas spécial : Géométrique(θ) lorsque $r = 1$.

$$\begin{aligned}
\Pr[X = x] &= \binom{x+r-1}{r-1} \theta^r (1-\theta)^x, \quad x = 0, 1, \dots \\
E[X] &= \frac{r(1-\theta)}{\theta} \\
\mathrm{Var}[X] &= \frac{r(1-\theta)}{\theta^2} \\
M(t) &= \left(\frac{\theta}{1-(1-\theta)e^t} \right)^r
\end{aligned}$$

C.1.3 Poisson

Racine : `pois`
Paramètre : `lambda` ($\lambda \geq 0$)

$$\begin{aligned}
\Pr[X = x] &= \frac{\lambda^x e^{-\lambda}}{x!}, \quad x = 0, 1, \dots \\
E[X] &= \lambda \\
\mathrm{Var}[X] &= \lambda \\
M(t) &= e^{\lambda(e^t - 1)}
\end{aligned}$$

C.2 Distributions continues

C.2.1 Bêta

Racine : `beta`
Paramètres : `shape1` ($a > 0$), `shape2` ($b > 0$)
Cas spécial : $\text{Uniforme}(0, 1)$ lorsque $a = b = 1$.

$$\begin{aligned}
f(x) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, \quad 0 < x < 1 \\
E[X] &= \frac{a}{a+b} \\
\mathrm{Var}[X] &= \frac{ab}{(a+b)^2 (a+b+1)}
\end{aligned}$$

C.2.2 Gamma

Racine : `gamma`
Paramètres : `shape` ($\alpha > 0$), `rate` ($\lambda = 1/\theta$), `scale` ($\theta > 0$)
Cas spéciaux : $\text{Exponentielle}(\lambda)$ lorsque $\alpha = 1$, $\chi^2(r)$ lorsque $\alpha = r/2$ et $\lambda = 1/2$.

$$\begin{aligned}
f(x) &= \frac{\lambda^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1} e^{-\lambda x}, \quad x > 0 \\
E[X] &= \frac{\alpha}{\lambda} \\
\mathrm{Var}[X] &= \frac{\alpha}{\lambda^2} \\
M(t) &= \left(\frac{\lambda}{\lambda - t}\right)^\alpha
\end{aligned}$$

C.2.3 Normale

Racine : `norm`
Paramètres : `mean` ($-\infty < \mu < \infty$), `sd` ($\sigma > 0$)

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}, \quad -\infty < x < \infty \\
E[X] &= \mu \\
\mathrm{Var}[X] &= \sigma^2 \\
M(t) &= e^{\mu t + \sigma^2 t^2/2}
\end{aligned}$$

C.2.4 Log-normale

Racine : `lnorm`
Paramètres : `meanlog` ($-\infty < \mu < \infty$), `sdlog` ($\sigma > 0$)

$$\begin{aligned}
f(x) &= \frac{1}{\sqrt{2\pi}\sigma} \frac{1}{x} \exp\left\{-\frac{(\ln x - \mu)^2}{2\sigma^2}\right\}, \quad x > 0 \\
E[X] &= e^{\mu + \sigma^2/2} \\
\mathrm{Var}[X] &= e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)
\end{aligned}$$

C.2.5 Pareto

Racine : `pareto`, `pareto2` (**actuar**)
Paramètres : `shape` ($\alpha > 0$), `scale` ($\theta > 0$)

$$\begin{aligned}
f(x) &= \frac{\alpha\theta^\alpha}{(x+\theta)^{\alpha+1}}, \quad x > 0 \\
E[X] &= \frac{\theta}{\alpha - 1}, \quad \alpha > 1
\end{aligned}$$

$$\mathrm{Var}[X] = \frac{\alpha\theta^2}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2$$

C.2.6 Pareto généralisée

Racine : `genpareto` (**actuar**)
Paramètres : `shape1` ($\alpha > 0$), `shape2` ($\tau > 0$), `rate` ($\lambda = 1/\theta$), `scale` ($\theta > 0$)
Cas spécial : Pareto(α, λ) lorsque $\tau = 1$.

$$\begin{aligned}
f(x) &= \frac{\Gamma(\alpha+\tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\theta^\alpha x^{\tau-1}}{(x+\theta)^{\alpha+\tau}}, \quad x > 0 \\
E[X] &= \frac{\theta\tau}{\alpha-1}, \quad \alpha > 1 \\
\mathrm{Var}[X] &= \frac{\theta^2\tau(\alpha+\tau-1)}{(\alpha-1)^2(\alpha-2)}, \quad \alpha > 2
\end{aligned}$$

C.3 Distributions composées

Soit $S = X_1 + \dots + X_N$, où $X_1, \dots, X_n$ sont des variables aléatoires mutuellement indépendantes, identiquement distribuées et toutes indépendantes de N. On a toujours

$$\begin{aligned}
E[S] &= E[N]E[X] \\
\mathrm{Var}[S] &= \mathrm{Var}[N]E[X]^2 + E[N]\,\mathrm{Var}[X].
\end{aligned}$$

C.3.1 Binomiale composée

Distribution de S lorsque $N \sim$ Binomiale(n, θ) et $\Pr[X \leq x] = F_X(x)$.

$$\begin{aligned}
E[S] &= n\theta E[X] \\
\mathrm{Var}[S] &= n\theta(1-\theta)E[X]^2 + n\theta\,\mathrm{Var}[X]
\end{aligned}$$

C.3.2 Binomiale négative composée

Distribution de S lorsque $N \sim$ Binomiale négative(r, θ) et $\Pr[X \leq x] = F_X(x)$.

$$\begin{aligned}
E[S] &= \frac{r(1-\theta)}{\theta}\, E[X] \\
\mathrm{Var}[S] &= \frac{r(1-\theta)}{\theta^2}\, E[X]^2 + \frac{r(1-\theta)}{\theta}\, \mathrm{Var}[X]
\end{aligned}$$

C.3.3 Poisson composée

Distribution de S lorsque $N \sim \text{Poisson}(\lambda)$ et $\Pr[X \leq x] = F_X(x)$.

$$\begin{aligned} E[S] &= \lambda E[X] \\ \text{Var}[S] &= \lambda E[X^2] \end{aligned}$$

D Solutions

Chapitre 2

2.1 Nous avons $S = X_1 + \cdots + X_N$ avec $N \sim$ Binomiale$(1\,000, 0{,}6)$ et $X \sim$ Log-normale$(3, 4)$, d'où

$$\begin{aligned}
E[N] &= (1\,000)(0{,}6) = 600 \\
\mathrm{Var}[N] &= (1\,000)(0{,}6)(1 - 0{,}6) = 240 \\
E[X] &= e^{3+4/2} = e^5 \\
\mathrm{Var}[X] &= e^{2\cdot 3+4}(e^4 - 1) = e^{10}(e^4 - 1)
\end{aligned}$$

et, par conséquent,

$$\begin{aligned}
E[S] &= E[N]E[X] \\
&= 600\, e^5 \\
&= 89\,047 \\
\mathrm{Var}[S] &= E[N]\,\mathrm{Var}[X] + \mathrm{Var}[N]E[X]^2 \\
&= 600\, e^{10}(e^4 - 1) + 240\, e^{10} \\
&= 713\,633\,042.
\end{aligned}$$

2.2 Nous avons $E[S] = E[N]E[X]$ et $\mathrm{Var}[S] = E[N]\,\mathrm{Var}[X] + E[X]^2\,\mathrm{Var}[N]$. Or, $E[N] = 4(1-0{,}5)/(0{,}5) = 4$, $\mathrm{Var}[N] = 4(1-0{,}5)/(0{,}5)^2 = 8$, $E[X] = 2/0{,}5 = 4$ et $\mathrm{Var}[X] = 2/0{,}5^2 = 8$, d'où

$$E[S] = (4)(4) = 16$$

et

$$\mathrm{Var}[S] = (4)(8) + (4^2)(8) = 160.$$

2.4 Puisque $E[\bar{S}] = E[S]$ et $\mathrm{Var}[\bar{S}] = \mathrm{Var}[S]/n$, l'inégalité

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

est satisfaite lorsque

$$n \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{\mathrm{Var}[S]}{E[S]^2}.$$

Nous savons que dans le cas d'une Poisson composée, $E[S] = \lambda E[X]$ et $\mathrm{Var}[S] = \lambda(\mathrm{Var}[X] + E[X]^2)$. Ainsi,

$$\begin{aligned} n &\geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{\mathrm{Var}[X] + E[X]^2}{\lambda E[X]^2} \\ &= \frac{1}{\lambda}\left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right) \\ &= \frac{1}{\lambda}\left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 (1 + \mathrm{CV}(X)^2), \end{aligned}$$

où $\mathrm{CV}(X)$ est le coefficient de variation de X.

2.5 a) Nous savons de l'exemple 2.2 que dans le cas d'une Poisson composée, la crédibilité complète est donnée par

$$\lambda \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right).$$

Or, nous avons $E[X] = \theta/(\alpha-1) = 1/40$ et $\mathrm{Var}[X] = \alpha\theta^2/((\alpha-1)^2(\alpha-2)) = 3/1\,600$. Comme nous exigeons un niveau de crédibilité de 90 %, $\zeta_{\varepsilon/2} = 1{,}645$. En isolant k dans la formule, nous obtenons

$$k \geq \frac{1{,}645}{\sqrt{256}}\sqrt{1 + \frac{3/1\,600}{1/1\,600}} = 0{,}2056 = 20{,}56\ \%.$$

La valeur de k obtenue est grande parce que la distribution du montant des sinistres est très asymétrique. Pour obtenir une valeur de k plus usuelle (de l'ordre de 5 %), il faudrait augmenter le paramètre de Poisson, de manière à sommer un plus grand nombre de lois de Pareto.

b) De manière similaire, mais en utilisant plutôt la formule

$$n \geq \frac{1}{\lambda}\left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right)$$

obtenue à l'exercice 2.4 avec $n = 10$, nous obtenons

$$k \geq \frac{1}{\sqrt{256}} \frac{1{,}645}{\sqrt{10}} \sqrt{1 + \frac{3/1\,600}{1/1\,600}} = 0{,}065 = 6{,}5\ \%.$$

La marge d'erreur est plus petite qu'en a) puisque la distribution de $\bar{S}$ après 10 années est le résultat de la somme d'un beaucoup plus grand nombre de sinistres.

2.6 Dans un tel cas normale pure, la formule générale (2.1) devient simplement

$$\mu \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)\sigma.$$

a) On a $k = 0{,}04$ et $p = 0{,}95$ d'où $\zeta_{0{,}025} = 1{,}96$ et donc $\mu \geq 49\sigma$.

b) On a $k = 0{,}05$ et $p = 0{,}90$ d'où $\zeta_{0{,}05} = 1{,}645$ et donc $\mu \geq 32{,}9\sigma$.

c) On a $k = 0{,}01$ et $p = 0{,}98$ d'où $\zeta_{0{,}01} = 2{,}326$ et donc $\mu \geq 232{,}6\sigma$.

2.7 On nous donne que $\bar{S} \sim N(\mu, \sigma^2/9)$. La crédibilité complète d'ordre (k, p) est donc atteinte lorsque

$$E[\bar{S}] \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)\sqrt{\mathrm{Var}[\bar{S}]},$$

soit

$$\mu \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)\frac{\sigma}{3}.$$

a) On a $k = 0{,}05$ et $p = 0{,}90$ d'où $\zeta_{0{,}05} = 1{,}645$ et $\mu \geq 10{,}97\sigma$.

b) On a $k = 0{,}05$ et $p = 0{,}95$ d'où $\zeta_{0{,}025} = 1{,}96$ et $\mu \geq 13{,}07\sigma$.

c) On a $k = 0{,}01$ et $p = 0{,}90$ d'où $\zeta_{0{,}05} = 1{,}645$ et $\mu \geq 54{,}83\sigma$.

d) On a $k = 0{,}01$ et $p = 0{,}95$ d'où $\zeta_{0{,}025} = 1{,}96$ et $\mu \geq 65{,}33\sigma$.

2.8 Déterminons tout d'abord une formule générale pour le cas de la binomiale négative composée. Sachant que $E[N] = r(1-\theta)/\theta$ et que $\mathrm{Var}[N] = r(1-\theta)/\theta^2$, nous obtenons

$$r \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(\frac{1}{1-\theta} + \frac{\theta}{1-\theta}\frac{\mathrm{Var}[X]}{E[X]^2}\right).$$

Or, dans cet exercice, $\theta = 0{,}01$, $\mathrm{Var}[X]/E[X]^2 = 1/\alpha = 50$ et $\zeta_{\varepsilon/2}/k = 1{,}96/0{,}05 = 39{,}2$. Par conséquent, $r \geq 2\,328{,}24$.

2.9 Remarquez tout d'abord que le contexte est celui de l'exercice 2.4, sauf que le critère de crédibilité complète est exprimé non pas en nombre d'années d'expérience, n, mais plutôt en nombre total de sinistres espéré, soit $n\lambda$. Nous avons donc

$$n\lambda \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 (1 + \mathrm{CV}(X)^2).$$

Or, dans le cas présent, $\mathrm{CV}(X) = 0$ puisque $\Pr[X = M] = 1$ (distribution dégénérée en M), $k = 0{,}04$ et $\zeta_{0{,}05} = 1{,}645$, d'où $n\lambda \geq 1\,692$.

2.10 a) Le montant total des sinistres a une distribution Poisson composée avec $E[X] = 5$, $\mathrm{Var}[X] = 100$, $k = 0{,}03$ et $\zeta_{\varepsilon/2} = \zeta_{0{,}025} = 1{,}96$. De la formule générale (2.2), nous obtenons directement $\lambda \geq 21\,343$.

b) La distribution du nombre de sinistres est une Poisson. Nous pouvons utiliser la formule (2.2) avec $\mathrm{Var}[X] = 0$, ce qui donne $\lambda \geq 4\,269$.

2.11 Soit X_A la variable aléatoire du montant des sinistres du groupe A et X_B la variable aléatoire du montant des sinistres du groupe B. On nous donne $E[X_A] = 50$, $\mathrm{Var}[X_A] = 0$ et $X_B \sim \mathrm{Gamma}(\alpha, \lambda)$ avec $\alpha/\lambda = 50$. De plus, la distribution du montant total des sinistres est une Poisson composée, donc le seuil de crédibilité complète est donné par (2.2).

Pour le groupe A, $\lambda_A = 1\,000$ et $\mathrm{CV}(X_A) = 0$, d'où $(\zeta_{\varepsilon/2}/k)^2 = 1\,000$. Ce facteur est identique pour le groupe B, mais $\lambda_B = 3\,000$ et $\mathrm{CV}(X_B) = 1/\sqrt{\alpha}$, d'où

$$3\,000 = 1\,000\left(1 + \frac{1}{\alpha}\right),$$

soit $\alpha = 1/2$. Enfin, de $\alpha/\lambda = 50$ découle que $\lambda = 1/100$.

2.12 Deux situations sont décrites :

1) la crédibilité complète est accordée à la fréquence des sinistres seulement ;

2) la crédibilité complète est accordée au montant total des sinistres.

On nous donne donc d'abord le seuil de crédibilité complète dans un modèle avec distribution de Poisson pure (ou Poisson composée avec une distribution des montants de sinistres dégénérée en 1), puis on nous demande de déterminer le seuil dans un modèle avec distribution Poisson composée où la distribution du montant des sinistres a un coefficient de variation de $1/\sqrt{2}$. Toujours à partir de la formule (2.2), nous trouvons que, dans un premier temps,

$$1\,000 = \left(\frac{\zeta_{\varepsilon/2}}{0{,}05}\right)^2,$$

soit $\zeta_{\varepsilon/2} = 0{,}05\sqrt{1\,000}$. Dans un deuxième temps, nous avons

$$\begin{aligned}
\lambda &\geq \left(\frac{\zeta_{\varepsilon/2}}{0{,}25}\right)^2 \left(1 + \frac{1}{2}\right) \\
&= 1{,}5\left(\frac{0{,}05\sqrt{1\,000}}{0{,}25}\right)^2 \\
&= 60.
\end{aligned}$$

2.13 Il faut d'abord identifier le modèle pour la fréquence comme étant une Binomiale(n, θ), où n est le nombre d'employés et θ, la probabilité qu'un employé ait un accident. Nous devons étudier deux cas : la binomiale pure et la binomiale composée. La formule de crédibilité de stabilité générale est

$$E[S] \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)\sqrt{\mathrm{Var}[S]}.$$

Si $S \sim$ Binomiale(n, θ), alors

$$n \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{1-\theta}{\theta}$$

et, si $S \sim$ Binomiale composée$(n, \theta, F_X(\cdot))$,

$$n \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(\frac{1}{\theta}\,\mathrm{CV}(X)^2 + \frac{1-\theta}{\theta}\right).$$

(À noter que cette dernière équation revient à la première si $X = 1$ avec probabilité 1, c'est-à-dire si $\mathrm{CV}(X) = 0$.) Nous avons $\theta = 0{,}04$ et, dans le cas d'une binomiale pure, un niveau de crédibilité complète de $n = 6\,494$. Par conséquent, $(\zeta_{\varepsilon/2}/k)^2 = 270{,}5833$. En insérant cette valeur dans la formule pour la binomiale composée avec $\mathrm{CV}(X) = 2$, nous obtenons un niveau de crédibilité complète de $n = 33\,552$ employés.

2.14 Avec la distribution Poisson composée, le seuil de crédibilité complète est

$$\lambda \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(1 + \frac{\mathrm{Var}[X]}{E[X]^2}\right).$$

Or, on nous donne $E[X] = 100/(3-1) = 50$ et $\text{Var}[X] = 3 \cdot 100^2/((3-1)^2(3-2)) = 7\,500$. Si le seuil de crédibilité complète dans le cas Poisson composée est 2 500 quand $k = 0{,}05$, alors

$$\zeta_{\varepsilon/2} = 0{,}05\sqrt{\frac{2500}{1 + 7500/50^2}} = 1{,}25.$$

Dans le cas binomiale négative composée, le seuil de crédibilité complète est plutôt donné par

$$r \geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(\frac{1}{1-\theta} + \frac{\theta}{1-\theta}\frac{\text{Var}[X]}{E[X]^2}\right).$$

Avec $\text{Var}[X]/E[X]^2 = 3$, $\theta = 0{,}5$, $r = 2\,500$ et $\zeta_{\varepsilon/2} = 1{,}25$, nous obtenons $k = 0{,}0559$, d'où la largeur de l'intervalle de confiance autour de la moyenne est $2kE[S] = 2(0{,}0559)(2\,500)(50) = 13\,975$.

2.15 Nous avons un cas binomiale négative composée. À l'aide de la formule développée à l'exercice 2.8, nous obtenons directement que

$$\begin{aligned} r &\geq \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \left(\frac{1}{1-\theta} + \frac{\theta}{1-\theta}\frac{\text{Var}[X]}{E[X]^2}\right) \\ &= \left(\frac{1{,}645}{0{,}05}\right)^2 \left(\frac{3}{2} + \frac{1}{2} \cdot 3\right) \\ &= 3\,247{,}23. \end{aligned}$$

2.16 a) Avec $X \equiv 1$, S est une Poisson pure et le seuil de crédibilité complète est $\lambda \geq (\zeta_{\varepsilon/2}/k)^2 = (1{,}645/0{,}05)^2 = 1082{,}41$.

b) Dans le cas Poisson composée avec $\text{CV}(X) = 1$, le seuil de crédibilité complète est $\lambda \geq (\zeta_{\varepsilon/2}/k)^2(1 + \text{CV}(X)^2) = 2(1{,}96/0{,}04)^2 = 4\,802$.

c) Sans plus de précisions sur la formule de crédibilité partielle à utiliser, nous allons utiliser sur la formule de la racine carrée. Pour une seule période d'expérience, nous avons $\bar{S} = S$, donc la crédibilité partielle z serait

$$z = \min\left\{\sqrt{\frac{\lambda}{4\,802}}, 1\right\}.$$

2.17 Nous avons un modèle Poisson composée pour le montant total des sinistres. Le seuil de crédibilité complète de 3 000 est atteint lorsque $\text{CV}(X) = 0$, donc $(\zeta_{\varepsilon/2}/k)^2 = 3\,000$. Si nous avons plutôt $\text{CV}(X)^2 = \alpha/(\alpha - 2) = 3$, alors le seuil de crédibilité complète est $\lambda_0 = 3\,000(1 + 3) = 12\,000$. Enfin, si l'on a accordé à l'assuré une crédibilité partielle de $z = \sqrt{\lambda/12\,000} = 0{,}5$, alors c'est que l'assuré a eu $\lambda = 3\,000$ sinistres.

2.18 Il faut déduire de l'énoncé que la crédibilité complète est donnée en années d'expérience et que $\zeta_{\varepsilon/2} = 1{,}96$, $k = 0{,}05$, $\mathrm{CV}(X) = 3$ et $\lambda = 1\,000$. Puisque nous avons un modèle Poisson composée, le niveau de crédibilité complète n_0 est

$$n_0 = \frac{1}{1\,000}\left(\frac{1{,}96}{0{,}05}\right)^2 (1 + (3)^2) = 15{,}3664.$$

Après $n = 1$ année d'expérience, le facteur de crédibilité est

$$z = \sqrt{1/15{,}3664} = 0{,}2551.$$

Puisque $\bar{S} = S_1 = 6{,}75$ et $m = 5{,}00$, la prime de crédibilité partielle est

$$P = 0{,}2551(6{,}75) + (1 - 0{,}2551)(5) = 5{,}45.$$

2.19 Soit λ la grandeur du portefeuille. Le seuil de crédibilité complète est $\lambda_0 = 103\,500$. En utilisant la formule de la racine carrée, on a donc

$$z = 0{,}67 = \sqrt{\frac{\lambda}{103\,500}},$$

d'où $\lambda = 46\,462$.

2.20 Nous avons $E[S] = E[N]E[X] = 50{,}000$ et $\mathrm{Var}[S] = \mathrm{Var}[N]E[X]^2 + \mathrm{Var}[X]E[N] = 312\,500\,000$. Avec la formule générale (2.2), le seuil de crédibilité complète est

$$\begin{aligned} n_0 &= \left(\frac{\zeta_{\varepsilon/2}}{k}\right)^2 \frac{\mathrm{Var}[S]}{E[S]^2} \\ &= 270{,}6025(0{,}125) \\ &= 33{,}8253. \end{aligned}$$

a) Avec la formule de la racine carrée,

$$0{,}54 = \sqrt{\frac{n}{33{,}8253}} \Rightarrow n = 9{,}86.$$

b) Avec la formule « puissance 2/3 »,

$$0{,}54 = \left(\frac{n}{33{,}8253}\right)^{2/3} \Rightarrow n = 13{,}42.$$

Le temps requis pour atteindre la pleine crédibilité est donc plus élevé avec cette formule.

Chapitre 3

3.1 Soit S le montant total des sinistres (ou le nombre de sinistres, car chaque sinistre est d'un montant de 1). Alors $S|\Theta \sim \text{Poisson}(\Theta)$ et Θ a une fonction de masse de probabilité

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}35, & \theta = 1/10 = 0{,}1 \\ 0{,}40, & \theta = 2/10 = 0{,}2 \\ 0{,}25, & \theta = 6/10 = 0{,}6. \end{cases}$$

a) Par la loi des probabilités totales,

$$\begin{aligned} \Pr[S = 1] &= \sum_\theta \Pr[S = 1|\Theta = \theta] \Pr[\Theta = \theta] \\ &= (0{,}1\, e^{-0{,}1})(0{,}35) + (0{,}2\, e^{-0{,}2})(0{,}40) + (0{,}6\, e^{-0{,}6})(0{,}25) \\ &= 0{,}1795. \end{aligned}$$

b) Ici, $\mu(\theta) = E[S|\Theta = \theta] = \theta$. Nous avons donc $\mu(0{,}1) = 0{,}1$, $\mu(0{,}2) = 0{,}2$ et $\mu(0{,}6) = 0{,}6$.

c) $E[S] = E[\mu(\Theta)] = E[\Theta] = 0{,}35(0{,}1) + 0{,}40(0{,}2) + 0{,}25(0{,}6) = 0{,}265$.

d) La première étape consiste à calculer la fonction de masse de probabilité à postériori de Θ. Soit $\mathbf{S} = (S_1, S_2, S_3, S_4, S_5)$ et $\mathbf{x} = (1, 0, 1, 1, 0)$. Alors, par l'indépendance conditionnelle de $S_1, \dots, S_5$,

$$\Pr[\mathbf{S} = \mathbf{x}|\Theta = \theta] = \prod_{j=1}^{5} \Pr[S_j = x_j|\Theta = \theta] = \theta^3 e^{-5\theta}.$$

Nous obtenons donc $\Pr[\mathbf{S} = \mathbf{x}|\Theta = 0{,}1] = 0{,}000607$, $\Pr[\mathbf{S} = \mathbf{x}|\Theta = 0{,}2] = 0{,}002943$, $\Pr[\mathbf{S} = \mathbf{x}|\Theta = 0{,}6] = 0{,}010754$ et, par la loi des probabilités totales, $\Pr[\mathbf{S} = \mathbf{x}] = 0{,}004078$. En utilisant la règle de Bayes,

$$\Pr[\Theta = \theta|\mathbf{S} = \mathbf{x}] = \begin{cases} 0{,}0521, & \theta = 0{,}1 \\ 0{,}2887, & \theta = 0{,}2 \\ 0{,}6592, & \theta = 0{,}6 \end{cases}$$

et, finalement, la prime bayésienne de sixième année est $E[\Theta|\mathbf{S} = \mathbf{x}] = 0{,}0521(0{,}1) + 0{,}2887(0{,}2) + 0{,}6592(0{,}6) = 0{,}4585$.

3.2 Soit S_i le résultat du i^{e} lancer d'un dé, D_1 l'événement « choisir le dé régulier » et D_2 l'événement « choisir le dé irrégulier ».

a) Par la loi des probabilités totales,

$$\begin{aligned}\Pr[S_1 = 5] &= \Pr[S_1 = 5|D_1]\Pr[D_1] + \Pr[S_1 = 5|D_2]\Pr[D_2] \\ &= \left(\frac{1}{6}\right)\left(\frac{1}{2}\right) + \left(\frac{2}{6}\right)\left(\frac{1}{2}\right) \\ &= \frac{1}{4}.\end{aligned}$$

Plus simplement, nous avons trois faces numérotées 5 sur un total de douze, d'où une probabilité de $\frac{1}{4}$.

b) Nous cherchons la valeur de la fonction de masse de probabilité de la distribution prédictive de S_2 au point $x = 5$:

$$\begin{aligned}\Pr[S_2 = 5|S_1 = 5] &= \frac{1}{\Pr[S_1 = 5]}\,(\Pr[S_1 = 5, S_2 = 5|D_1]\Pr[D_1] \\ &\quad + \Pr[S_1 = 5, S_2 = 5|D_2]\Pr[D_2]) \\ &= \frac{(\frac{1}{36})(\frac{1}{2}) + (\frac{4}{36})(\frac{1}{2})}{\frac{1}{4}} \\ &= \frac{5}{18}.\end{aligned}$$

3.3 Soit S le nombre de piles obtenus après n lancers d'une pièce de monnaie et Θ l'urne choisie. Nous identifions l'urne A par $\theta = 0{,}6$ et l'urne B par $\theta = 0{,}2$. Ainsi, $S|\Theta = \theta \sim \text{Binomiale}(n, \theta)$. Nous devons calculer $E[S_2|S_1 = 4]$. La distribution à postériori de $\Theta|S_1 = 4$ est

$$\Pr[\Theta = \theta|S_1 = 4] = \frac{\Pr[S_1 = 4|\Theta = \theta]\Pr[\Theta = \theta]}{\sum_\theta \Pr[S_1 = 4|\Theta = \theta]\Pr[\Theta = \theta]}.$$

Par conséquent,

$$\begin{aligned}\Pr[\Theta = 0{,}6|S_1 = 4] &= \frac{\binom{5}{4}(0{,}6)^4(0{,}4)(0{,}5)}{\binom{5}{4}(0{,}6)^4(0{,}4)(0{,}5) + \binom{5}{4}(0{,}2)^4(0{,}8)(0{,}5)} \\ &= \frac{0{,}1296}{0{,}1328} \\ &= 0{,}9759.\end{aligned}$$

De la même façon, $\Pr[\Theta = 0{,}2|S_1 = 4] = 0{,}0032/0{,}1328 = 0{,}02410$. Enfin,

$$\begin{aligned}E[S_2|S_1 = 4] &= E[S_2|\theta = 0{,}6]\Pr[\Theta = 0{,}6|S_1 = 4] \\ &\quad + E[S_2|\Theta = 0{,}2]\Pr[\Theta = 0{,}2|S_1 = 4] \\ &= 5(0{,}6)(0{,}9759) + 5(0{,}2)(0{,}02410) \\ &= 2{,}9518.\end{aligned}$$

3.4 Soit S une variable indiquant si un employeur a eu un sinistre dans une année et Θ une variable identifiant le groupe auquel cet employeur appartient. Identifions le groupe à fréquence faible par $\theta = 0{,}1$, le groupe à fréquence moyenne par $\theta = 0{,}2$ et le groupe à fréquence élevée par $\theta = 0{,}4$. Nous avons donc $S|\Theta = \theta \sim \text{Bernoulli}(\theta)$.

a) Par la loi des probabilités totales,

$$\begin{aligned}
\Pr[S = 1] &= \sum_\theta \Pr[S = 1|\Theta = \theta]\Pr[\Theta = \theta] \\
&= (0{,}1 + 0{,}2 + 0{,}4)\left(\frac{1}{3}\right) \\
&= \frac{7}{30}.
\end{aligned}$$

b) En utilisant la règle de Bayes et l'hypothèse d'indépendance des sinistres,

$$\begin{aligned}
\Pr[\Theta = 0{,}1|S_1 = 0, S_2 = 0] &= \frac{\Pr[S_1 = 0, S_2 = 0|\Theta = 0{,}1]\Pr[\Theta = 0{,}1]}{\Pr[S_1 = 0, S_2 = 0]} \\
&= \frac{(0{,}9)^2(\frac{1}{3})}{(0{,}9)^2(\frac{1}{3}) + (0{,}8)^2(\frac{1}{3}) + (0{,}6)^2(\frac{1}{3})} \\
&= \frac{81}{181} \\
&= 0{,}4475.
\end{aligned}$$

c) De la même façon qu'en b), nous trouvons $\Pr[\Theta = 0{,}2|S_1 = 0, S_2 = 0] = 64/181 = 0{,}3536$ et $\Pr[\Theta = 0{,}4|S_1 = 0, S_2 = 0] = 36/181 = 0{,}1989$. Ainsi,

$$\begin{aligned}
&\Pr[S_3 = 1|S_1 = 0, S_2 = 0] \\
&\quad= \sum_\theta \Pr[S_3 = 1|\Theta = \theta]\Pr[\Theta = \theta|S_1 = 0, S_2 = 0] \\
&\quad= \left(\frac{1}{181}\right)[(0{,}1)(81) + (0{,}2)(64) + (0{,}4)(36)] \\
&\quad= \frac{35{,}3}{181} \\
&\quad= 0{,}1950.
\end{aligned}$$

3.5 a) Nous avons

$$\begin{aligned} E[\Theta] &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \theta^{(a+1)-1}(1-\theta)^{b-1}\,d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)} \\ &= \frac{a}{a+b}. \end{aligned}$$

b) Nous avons $S|\Theta = \theta \sim \text{Binomiale}(n, \theta)$, c'est-à-dire $\Pr[S = x|\Theta = \theta] = \binom{n}{x}\theta^x(1-\theta)^{n-x}$. Par la règle de Bayes,

$$\begin{aligned} u(\theta|x) &\propto \theta^x(1-\theta)^{n-x}\theta^{a-1}(1-\theta)^{b-1} \\ &= \theta^{a+x-1}(1-\theta)^{b+n-x-1}, \end{aligned}$$

que nous savons être la densité d'une bêta avec de nouveaux paramètres. Donc, $\Theta|S = x \sim \text{Bêta}(\tilde{a} = a + x, \tilde{b} = b + n - x)$.

c) Une pièce a été lancée dix fois et les résultats suivants furent obtenus : P, F, F, P, F, P, F, F, P, P. En utilisant les résultats obtenus précédemment, la distribution à postériori de Θ est une bêta et les paramètres (a, b) après chaque lancer sont les suivants, dans l'ordre : $(1, 2)$, $(2, 2)$, $(3, 2)$, $(3, 3)$, $(4, 3)$, $(4, 4)$, $(5, 4)$, $(6, 4)$, $(6, 5)$, $(6, 6)$. Les graphiques de quelques distributions à postériori sont présentés aux figures D.1–D.6. Une première observation est le déplacement vers la gauche ou vers la droite de la distribution selon le nombre de résultats piles et de résultats faces obtenus. Une deuxième observation est la concentration graduelle de la distribution autour de $\theta = \frac{1}{2}$, ce qui indique une pièce équilibrée.

d) Soit Y le résultat du onzième lancer, où $Y = 1$ correspond à face. Nous avons

$$\begin{aligned} \Pr[Y = 1] &= \int_0^1 \Pr[Y = 1|\Theta = \theta]\Pr[\Theta = \theta]\,d\theta \\ &= \int_0^1 \theta \Pr[\Theta = \theta]\,d\theta \\ &= E[\Theta] \\ &= \frac{a}{a+b}. \end{aligned}$$

Le résultat n'est pas nécessairement $\frac{1}{2}$ puisque nous ignorons si la pièce est équilibrée ou non.

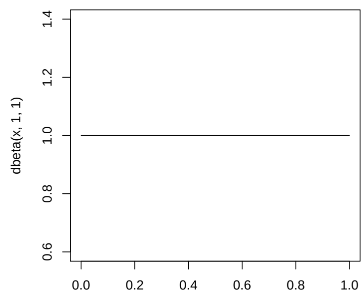


FIG. D.1 – Distribution à priori (uniforme)

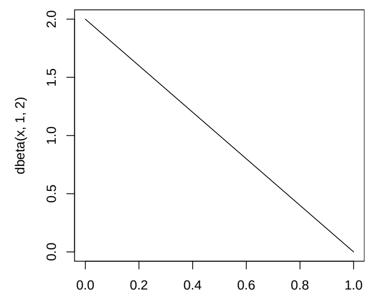


FIG. D.2 – Distribution après un lancer (P)

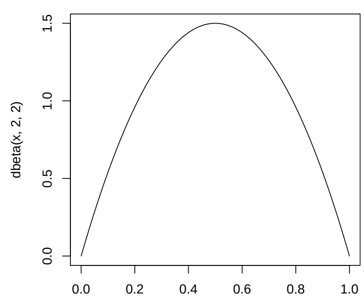


FIG. D.3 – Distribution après deux lancers (P, F)

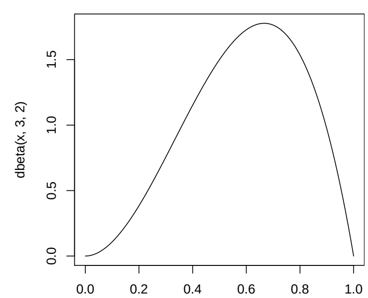


FIG. D.4 – Distribution après trois lancers (P, F, F)

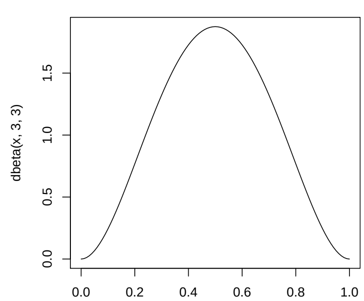


FIG. D.5 – Distribution après quatre lancers (P, F, F, P)

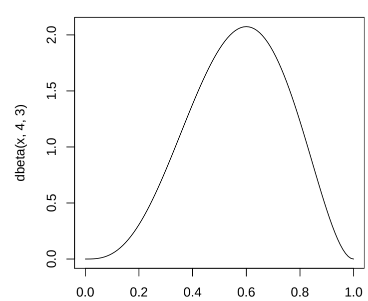


FIG. D.6 – Distribution après cinq lancers (P, F, F, P, F)

e) Nous avons

$$\begin{aligned}
\Pr[Y = 1|S = x] &= \int_0^1 \Pr[Y = 1|\Theta = \theta] \Pr[\Theta = \theta|S = x]\, d\theta \\
&= \int_0^1 \theta \Pr[\Theta = \theta|S = x]\, d\theta \\
&= E[\Theta|S = x] \\
&= \frac{a + x}{a + b + n}.
\end{aligned}$$

3.6 Nous avons le modèle suivant : $X|\Theta \sim \text{Pareto}(\Theta, 10)$ et $\Theta \sim U(1, 2, 3)$. La fonction de répartition de $X|\Theta = \theta$ est

$$F(x|\theta) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\theta}.$$

Ainsi,

$$\begin{aligned}
\Pr[X_2 > 30|X_1 = 20] &= \sum_{\theta=1}^{3} \Pr[X_2 > 30|\Theta = \theta] \Pr[\Theta = \theta|X_1 = 20] \\
&= \sum_{\theta=1}^{3} (1 - F(30|\theta)) \frac{f(20|\theta)(\frac{1}{3})}{\sum_{\theta=1}^{3} f(20|\theta)(\frac{1}{3})} \\
&= \frac{\sum_{\theta=1}^{3} \left(\frac{10}{10+30}\right)^{\theta} \frac{\theta 10^{\theta}}{(10+20)^{\theta+1}}}{\sum_{\theta=1}^{3} \frac{\theta 10^{\theta}}{(10+20)^{\theta+1}}} \\
&= 0{,}1484.
\end{aligned}$$

3.7 a) $N|\Theta = \theta \sim \text{Binomiale}(10, \theta)$ et la fonction de masse de probabilité de Θ est

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}2, & \theta = 0{,}04 \\ 0{,}6, & \theta = 0{,}10 \\ 0{,}2, & \theta = 0{,}16. \end{cases}$$

b) En général, nous avons

$$\begin{aligned}
\Pr[N = n] &= \sum_{\theta} \Pr[N = n|\Theta = \theta] \Pr[\Theta = \theta] \\
&= \sum_{\theta} \binom{10}{n} \theta^n (1 - \theta)^{10-n} \Pr[\Theta = \theta],
\end{aligned}$$

ce qui donne $\Pr[N = 0] = 0{,}377$, $\Pr[N = 1] = 0{,}354$ et $\Pr[N = 2] = 0{,}184$.

c) Nous réutilisons la formule en b) avec $\Pr[\Theta = 0{,}10] = 1$. Les résultats sont : $\Pr[N = 0] = 0{,}349$, $\Pr[N = 1] = 0{,}387$ et $\Pr[N = 2] = 0{,}194$.

3.8 Soit $S|\Theta \sim \text{Gamma}(\tau, \Theta)$ et $\Theta^{-1} \sim \text{Bêta}(\alpha, \beta)$. Alors $E[S|\Theta] = \tau\,\Theta^{-1}$ et $\text{Var}[S|\Theta] = \tau\,\Theta^{-2}$, d'où

$$\begin{aligned}
\text{Var}[S] &= \text{Var}[E[S|\Theta]] + E[\text{Var}[S|\Theta]] \\
&= \tau^2\,\text{Var}[\Theta^{-1}] + \tau E[\Theta^{-2}] \\
&= \frac{\tau^2\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} + \frac{\tau\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.
\end{aligned}$$

Avec $\tau = 2$, $\alpha = 2$ et $\beta = 1$, $\text{Var}[S] = 11/9$.

3.9 Nous avons

$$\begin{aligned}
\Pr[S = x|\Theta = \theta] &= \frac{\theta^x e^{-\theta}}{x!}, \quad x = 0, 1, \dots, \\
u(\theta) &= \left(\frac{5}{4}\right)\left(\frac{1}{\theta^2}\right), \quad 1 < \theta < 5.
\end{aligned}$$

a) Par la loi des probabilités totales,

$$\begin{aligned}
\Pr[S = 2] &= \int_1^5 \Pr[S = 2|\Theta = \theta]u(\theta)\,d\theta \\
&= \frac{5}{4}\int_1^5 \frac{\theta^2 e^{-\theta}}{2!\,\theta^2}\,d\theta \\
&= \left(\frac{5}{8}\right)(e^{-1} - e^{-5}) \\
&= 0{,}2257.
\end{aligned}$$

b) Déterminons d'abord la densité à postériori de Θ. Nous avons :

$$\begin{aligned}
u(\theta|x_1 = 1, x_2 = 1) &= \frac{\Pr[S_1 = 1|\Theta = \theta]\Pr[S_2 = 1|\Theta = \theta]u(\theta)}{\int_1^5 \Pr[S_1 = 1|\Theta = \theta]\Pr[S_2 = 1|\Theta = \theta]u(\theta)} \\
&= \frac{e^{-2\theta}}{\int_1^5 e^{-2\theta}\,d\theta} \\
&= \frac{2e^{-2\theta}}{e^{-2} - e^{-10}} \\
&= 14{,}7831e^{-2\theta}.
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
\Pr[S_3 = 0|S_1 = 1, S_2 = 1] &= \int_1^5 \Pr[S_3 = 0|\Theta = \theta]u(\theta|x_1 = 1, x_2 = 1)\,d\theta \\
&= \frac{2}{e^{-2} - e^{-10}} \int_1^5 e^{-3\theta}\,d\theta \\
&= \frac{2}{3}\,\frac{e^{-3} - e^{-15}}{e^{-2} - e^{-10}} \\
&= 0{,}2453.
\end{aligned}$$

c) Avant tout, il faut noter que $\mu(\Theta) = \Theta$. Puis, en intégrant par parties,

$$\begin{aligned}
E[\Theta|S_1 = 1, S_2 = 1] &= \int_1^5 \theta u(\theta|x_1 = 1, x_2 = 1)\,d\theta \\
&= \frac{2}{e^{-2} - e^{-10}} \int_1^5 \theta e^{-2\theta}\,d\theta \\
&= \frac{2}{e^{-2} - e^{-10}} \left[-\frac{\theta}{2}e^{-2\theta} - \frac{1}{4}e^{-2\theta}\right|_1^5 \\
&= \frac{2}{e^{-2} - e^{-10}} \left(\frac{3}{4}e^{-2} - \frac{11}{4}e^{-10}\right) \\
&= 1{,}4987.
\end{aligned}$$

3.10 Dans ce problème, il faut réaliser qu'il y a une probabilité de $3/4$ de ne pas avoir de sinistre ou, de manière équivalente, un sinistre de montant 0. Définissons les variables aléatoires S représentant le montant d'un sinistre dans une année et Θ identifiant la classe de risque. Par abus de notation, nous allons considérer que les valeurs possibles de Θ sont $\theta = A$, B et C. À partir des données du tableau, nous pouvons construire la fonction de masse de probabilité de la variable aléatoire conditionnelle $S|\Theta = \theta$:

$$\Pr[S = x|\Theta = A] = \begin{cases} \frac{3}{4}, & x = 0 \\ (\frac{1}{4})(\frac{3}{5}) = \frac{3}{20}, & x = 10\,000 \\ (\frac{1}{4})(\frac{1}{5}) = \frac{1}{20}, & x = 20\,000 \\ (\frac{1}{4})(\frac{1}{5}) = \frac{1}{20}, & x = 30\,000 \end{cases}$$

$$\Pr[S = x|\Theta = B] = \begin{cases} \frac{3}{4}, & x = 0 \\ 0, & x = 10\,000 \\ (\frac{1}{4})(\frac{1}{2}) = \frac{1}{8}, & x = 20\,000 \\ (\frac{1}{4})(\frac{1}{2}) = \frac{1}{8}, & x = 30\,000 \end{cases}$$

$$\Pr[S = x|\Theta = C] = \begin{cases} \frac{3}{4}, & x = 0 \\ (\frac{1}{4})(\frac{1}{5}) = \frac{1}{20}, & x = 10\,000 \\ (\frac{1}{4})(\frac{1}{5}) = \frac{1}{20}, & x = 20\,000 \\ (\frac{1}{4})(\frac{3}{5}) = \frac{3}{20}, & x = 30\,000 \end{cases}$$

et $\Pr[\Theta = \theta] = 1/3$, $\theta = A, B, C$. Pour la suite, il y a deux solutions possibles.

1. Avec la distribution à postériori de Θ. Tout d'abord, nous avons

$$\begin{aligned} \Pr[\Theta = \theta|S_1 = 30\,000] &= \frac{\Pr[S_1 = 30\,000|\Theta = \theta]\Pr[\Theta = \theta]}{\sum_\theta \Pr[S_1 = 30\,000|\Theta = \theta]\Pr[\Theta = \theta]} \\ &= \begin{cases} \frac{\frac{1}{20}}{\frac{1}{20}+\frac{1}{8}+\frac{3}{20}} = \frac{2}{13}, & \theta = A \\ \frac{\frac{1}{8}}{\frac{1}{20}+\frac{1}{8}+\frac{3}{20}} = \frac{5}{13}, & \theta = B \\ \frac{\frac{3}{20}}{\frac{1}{20}+\frac{1}{8}+\frac{3}{20}} = \frac{6}{13}, & \theta = C. \end{cases} \end{aligned}$$

Or,

$$\begin{aligned} \mu(A) &= E[S|\Theta = A] = 4\,000 \\ \mu(B) &= E[S|\Theta = B] = 6\,250 \\ \mu(C) &= E[S|\Theta = C] = 6\,000, \end{aligned}$$

d'où la prime bayésienne pour la deuxième année est

$$\begin{aligned} E[\mu(\Theta)|S_1 = 30\,000] &= 4\,000\left(\frac{2}{13}\right) + 6\,250\left(\frac{5}{13}\right) + 6\,000\left(\frac{6}{13}\right) \\ &= 5\,788{,}46. \end{aligned}$$

2. Avec la distribution prédictive. Nous avons

$$\begin{aligned}
\Pr[S_2 = x | S_1 = 30\,000] &= \frac{\Pr[S_2 = x, S_1 = 30\,000]}{\Pr[S_1 = 30\,000]} \\
&= \frac{\sum_{\theta=A,B,C} \Pr[S_2 = x | \Theta = \theta] \Pr[S_1 = 30\,000 | \Theta = \theta] \Pr[\Theta = \theta]}{\sum_{\theta=A,B,C} \Pr[S_1 = 30\,000 | \Theta = \theta] \Pr[\Theta = \theta]} \\
&= \begin{cases}
\dfrac{(\frac{3}{4})(\frac{1}{20}) + (\frac{3}{4})(\frac{1}{8}) + (\frac{3}{4})(\frac{3}{20})}{\frac{1}{20} + \frac{1}{8} + \frac{3}{20}}, & x = 0 \\
\dfrac{(\frac{3}{20})(\frac{1}{20}) + (0)(\frac{1}{8}) + (\frac{1}{20})(\frac{3}{20})}{\frac{1}{20} + \frac{1}{8} + \frac{3}{20}}, & x = 10\,000 \\
\dfrac{(\frac{1}{20})(\frac{1}{20}) + (\frac{1}{8})(\frac{1}{8}) + (\frac{1}{20})(\frac{3}{20})}{\frac{1}{20} + \frac{1}{8} + \frac{3}{20}}, & x = 20\,000 \\
\dfrac{(\frac{1}{20})^2 + (\frac{1}{8})^2 + (\frac{3}{20})^2}{\frac{1}{20} + \frac{1}{8} + \frac{3}{20}}, & x = 30\,000
\end{cases} \\
&= \begin{cases}
\frac{390}{520}, & x = 0 \\
\frac{6}{130}, & x = 10\,000 \\
\frac{41}{520}, & x = 20\,000 \\
\frac{65}{520}, & x = 30\,000,
\end{cases}
\end{aligned}$$

d'où

$$\begin{aligned}
E[S_2 | S_1 = 30\,000] &= 10\,000 \left(\frac{6}{130} \right) + 20\,000 \left(\frac{41}{520} \right) \\
&\quad + 30\,000 \left(\frac{65}{520} \right) \\
&= 5\,788{,}46.
\end{aligned}$$

3.11 La sévérité des sinistres ne joue aucun rôle dans ce problème. Posons $S|\Theta = \theta \sim \text{Poisson}(\theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$. Nous savons que $E[\Theta] = \alpha/\lambda = 2$ et $\text{Var}[\Theta] = \alpha/\lambda^2 = 2$, d'où $\alpha = 2$ et $\lambda = 1$.

a) Nous chargeons la prime collective à un nouvel assuré. Or, $E[S] = E[E[S|\Theta]] = E[\Theta] = 2$.

b) Nous cherchons $E[\mu(\Theta) | S_1 + S_2 + S_3 = 8]$. Or,

$$\begin{aligned}
u(\theta | x_1, x_2, x_3) &\propto u(\theta) \prod_{t=1}^{3} f(x_t | \theta) \\
&\propto \theta^{2-1} e^{-\theta} \prod_{t=1}^{3} \theta^{x_t} e^{-\theta} \\
&= \theta^{\sum_{t=1}^{3} x_t + 2 - 1} e^{-4\theta},
\end{aligned}$$

d'où $\Theta|S_1 = x_1, S_2 = x_2, S_3 = x_3 \sim \text{Gamma}(\sum_{t=1}^{3} x_t + 2, 4)$. (Vous remarquerez que seul le nombre total de sinistres dans les trois premières années est important, pas les fréquences annuelles.) Enfin, $E[\mu(\Theta)|S_1 + S_2 + S_3 = 8] = E[\Theta|S_1 + S_2 + S_3 = 8] = 10/4 = 2{,}5$.

3.12 Nous avons $N|\Theta = \theta \sim \text{Binomiale}(2, \theta)$ et

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}25, & \theta = 1/4 \\ 0{,}25, & \theta = 1/2 \\ 0{,}50, & \theta = 1/8. \end{cases}$$

a) La fonction de masse de probabilités est :

$$\begin{aligned} \Pr[\Theta = \theta|N_1 = 0, N_2 = 2] \\ &= \frac{\Pr[N_1 = 0, N_2 = 2|\Theta = \theta] \Pr[\Theta = \theta]}{\sum_\theta \Pr[N_1 = 0, N_2 = 2|\Theta = \theta] \Pr[\Theta = \theta]} \\ &= \frac{\theta^2(1-\theta)^2 \Pr[\Theta = \theta]}{\sum_\theta \theta^2(1-\theta)^2 \Pr[\Theta = \theta]} \\ &= \begin{cases} \frac{0{,}008789}{0{,}030396}, & \theta = 1/4 \\ \frac{0{,}015625}{0{,}030396}, & \theta = 1/2 \\ \frac{0{,}005981}{0{,}030396}, & \theta = 1/8 \end{cases} \\ &= \begin{cases} 0{,}2891, & \theta = 1/4 \\ 0{,}5141, & \theta = 1/2 \\ 0{,}1968, & \theta = 1/8. \end{cases} \end{aligned}$$

b) Nous devons calculer $E[N_3|N_1 = 0, N_2 = 2]$, ce qui peut se faire de deux façons. Tout d'abord, en trouvant d'abord la distribution prédictive :

$$\begin{aligned} \Pr[N_3 = n|N_1 = 0, N_2 = 2] \\ &= \sum_\theta \Pr[N_3 = n|\Theta = \theta] \Pr[\Theta = \theta|N_1 = 0, N_2 = 2] \\ &= \begin{cases} 0{,}4418, & n = 0 \\ 0{,}4085, & n = 1 \\ 0{,}1497, & n = 2. \end{cases} \end{aligned}$$

Ainsi,

$$\begin{aligned} E[N_3|N_1 = 0, N_2 = 2] &= 0{,}4085 + 2(0{,}1497) \\ &= 0{,}7078. \end{aligned}$$

L'autre méthode, plus rapide, consiste à calculer la prime bayésienne à partir de la prime de risque :

$$\begin{aligned}
E[N_3|N_1 = 0, N_2 = 2] &= E[\mu(\Theta)|N_1 = 0, N_2 = 2] \\
&= E[2\Theta|N_1 = 0, N_2 = 2] \\
&= 2E[\Theta|N_1 = 0, N_2 = 2] \\
&= 2\left[\frac{1}{4}(0{,}2891) + \frac{1}{2}(0{,}5141) + \frac{1}{8}(0{,}1968)\right] \\
&= 0{,}7078.
\end{aligned}$$

3.13 Soit S le résultat du tir aléatoire et Θ le résultat du dé. Nous avons $S|\Theta = \theta \sim U(0, 100\theta)$ et $\Pr[\Theta = \theta] = 1/6$.

a) Déterminons d'abord la densité à postériori de Θ :

$$\begin{aligned}
\Pr[\Theta = \theta|S_1 = 80, S_2 = 340) &= \frac{f(80|\theta)f(340|\theta)\Pr[\Theta = \theta]}{\sum_{\theta=1}^{6} f(80|\theta)f(340|\theta)\Pr[\Theta = \theta]} \\
&= \begin{cases}
\frac{(\frac{1}{100})(0)}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 1 \\
\frac{(\frac{1}{200})(0)}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 2 \\
\frac{(\frac{1}{300})(0)}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 3 \\
\frac{(\frac{1}{400})(\frac{1}{400})}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 4 \\
\frac{(\frac{1}{500})(\frac{1}{500})}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 5 \\
\frac{(\frac{1}{600})(\frac{1}{600})}{(\frac{1}{400})^2+(\frac{1}{500})^2+(\frac{1}{600})^2}, & \theta = 6
\end{cases} \\
&= \begin{cases}
0, & \theta = 1 \\
0, & \theta = 2 \\
0, & \theta = 3 \\
0{,}4798, & \theta = 4 \\
0{,}3070, & \theta = 5 \\
0{,}2132, & \theta = 6.
\end{cases}
\end{aligned}$$

Ainsi, puisque $\mu(\theta) = E[S|\Theta = \theta] = 50\theta$, la prime bayésienne est

$$\begin{aligned}
E[\mu(\Theta)|S_1 = 80, S_2 = 340] &= 50[(4)(0{,}4798) + (5)(0{,}3070) \\
&\quad + (6)(0{,}2132)] \\
&= 236{,}67.
\end{aligned}$$

b) Il faudrait que

$$\begin{aligned}236{,}67 &= z\bar{S} + (1-z)E[\mu(\Theta)] \\ &= z210 + (1-z)(50)(1+2+3+4+5+6)\left(\frac{1}{6}\right) \\ &= 175 + 35z,\end{aligned}$$

soit $z = 1{,}76$. Comme z ne peut être plus grand que 1 par définition, il est impossible d'écrire la prime bayésienne sous la forme d'une prime de crédibilité.

3.14 Par le théorème de Bayes,

$$\begin{aligned}u(\theta|S_1 = 1, S_2 = 1) &= \frac{\prod_{t=1}^{2} f(x_t|\Theta = \theta)u(\theta)}{\int_0^{10} \prod_{t=1}^{2} f(x_t|\Theta = \theta)u(\theta)\,d\theta} \\ &= \frac{\frac{\theta}{50}\prod_{t=1}^{2} \theta^2 x_t e^{-\theta x_t}}{\int_0^{10} \frac{\theta}{50}\prod_{t=1}^{2} \theta^2 x_t e^{-\theta x_t}\,d\theta} \\ &= \frac{\theta^5 e^{-2\theta}}{\int_0^{10} \theta^5 e^{-2\theta}\,d\theta},\end{aligned}$$

pour $0 < \theta < 10$. Or,

$$\begin{aligned}\int_0^{10} \theta^5 e^{-2\theta}\,d\theta &= \frac{\Gamma(6)}{2^6}\int_0^{10} \frac{2^6}{\Gamma(6)}\theta^{6-1}e^{-2\theta}\,d\theta \\ &= 1{,}875G(10; 6, 2),\end{aligned}$$

où $G(x; \alpha, \lambda)$ est la fonction de répartition d'une loi Gamma(α, λ). On obtient avec R la valeur de $G(10; 6, 2)$:

```
> pgamma(10, 6, 2)
[1] 0.9999281
```

Par conséquent, la distribution à postériori de Θ est

$$u(\theta|S_1 = 1, S_2 = 1) = \begin{cases}\dfrac{\theta^5 e^{-2\theta}}{1{,}8737}, & 0 < \theta < 10 \\ 0, \text{ailleurs.}\end{cases}$$

3.15 a) Le modèle est le suivant : $N|\Theta \sim \text{Poisson}(\Theta)$ et

$$\Pr[\Theta = \theta] = \begin{cases}0{,}5, & \theta = 2 \\ 0{,}5, & \theta = 4.\end{cases}$$

b) Vous devez d'abord trouver la distribution à postériori de Θ en utilisant

$$\Pr[N_1 = 4, N_2 = 4|\Theta = 2] = \left(\frac{2^4 e^{-2}}{4!}\right)^2 = 0{,}008140$$
$$\Pr[N_1 = 4, N_2 = 4|\Theta = 4] = \left(\frac{4^4 e^{-4}}{4!}\right)^2 = 0{,}038168,$$

d'où $\Pr[N_1 = 4, N_2 = 4] = 0{,}008140(0{,}5) + 0{,}038168(0{,}5) = 0{,}023154$. Alors

$$\Pr[\Theta = \theta|N_1 = 4, N_2 = 4] = \begin{cases} \frac{0{,}008140(0{,}5)}{0{,}023154} = 0{,}175779, & \theta = 2 \\ \frac{0{,}038168(0{,}5)}{0{,}023154} = 0{,}824221, & \theta = 4. \end{cases}$$

Ainsi, puisque $\mu(\Theta) = \Theta$, la prime bayésienne pour la troisième année est $E[\Theta|N_1 = 4, N_2 = 4] = 2(0{,}175779) + 4(0{,}824221) = 3{,}6484$.

3.16 Soit θ_1 représentant les bons risques, θ_2 les moyens risques et θ_3 les mauvais risques. Nous avons donc le modèle suivant :

$$S|\Theta = \theta_1 \sim \text{Gamma}(4, 2) \Leftrightarrow f(x|\theta_1) = \frac{8}{3} x^3 e^{-2x}$$
$$S|\Theta = \theta_2 \sim \text{Gamma}(4, 1) \Leftrightarrow f(x|\theta_2) = \frac{1}{6} x^3 e^{-x}$$
$$S|\Theta = \theta_3 \sim \text{Gamma}(10, 2) \Leftrightarrow f(x|\theta_3) = \frac{8}{2\,835} x^9 e^{-2x}$$

et

$$\Pr[\Theta = \theta_i] = \begin{cases} 0{,}25, & i = 1 \\ 0{,}60, & i = 2 \\ 0{,}15, & i = 3. \end{cases}$$

De plus $\mu(\theta_1) = 2$, $\mu(\theta_2) = 4$ et $\mu(\theta_3) = 5$. Le calcul de la prime bayésienne requiert la distribution à postériori de Θ :

$$\begin{aligned} \Pr[\Theta = \theta_i|S_1 = 1, S_2 = 2] &= \frac{f(1|\theta_i) f(2|\theta_i) \Pr[\Theta = \theta_i]}{\sum_{j=1}^{3} f(1|\theta_j) f(2|\theta_j) \Pr[\Theta = \theta_j]} \\ &= \begin{cases} 0{,}841507, & i = 1 \\ 0{,}158457, & i = 2 \\ 0{,}000036, & i = 3. \end{cases} \end{aligned}$$

La prime bayésienne pour la troisième année est donc $E[\mu(\Theta)|S_1 = 1, S_2 = 2] = 2(0{,}841507) + 4(0{,}158457) + 5(0{,}000036) = 2{,}3170$.

3.17 a) Soit Θ représentant le niveau de risque d'un contrat, avec $\theta = 1$ identifiant un risque de classe A et $\theta = 3$ identifiant un risque de classe B. Comme chaque classe contient le même nombre de risques, la probabilité qu'un risque choisi au hasard provienne d'une classe est de 50 % :

$$\Pr[\Theta = \theta] = \begin{cases} \frac{1}{2}, & \theta = 1 \\ \frac{1}{2}, & \theta = 3. \end{cases}$$

Si N est le nombre de sinistres pendant un an pour un contrat et X est le montant d'un sinistre pour ce même contrat, nous avons le modèle suivant :

$$\begin{aligned} N|\Theta &\sim \text{Poisson}(\Theta) \\ X|\Theta &\sim \text{Exponentielle}(1/\Theta). \end{aligned}$$

En effet, il est précisé dans l'énoncé que l'espérance de la loi exponentielle dans chaque groupe est égale à celle de la loi de Poisson.

b) Pour simplifier la notation, A représente l'événement $\{N_1 = 2, X_1 = 1, X_2 = 3\}$. Nous voulons calculer

$$E[S|A] = \sum_{\theta=1,3} E[S|\Theta = \theta] \Pr[\Theta = \theta|A].$$

Premièrement, comme la fréquence et la sévérité sont indépendantes sachant Θ à l'intérieur de chaque classe, nous avons

$$E[S|\Theta] = E[N|\Theta]E[X|\Theta] = \Theta^2.$$

Puis, en utilisant la règle de Bayes,

$$\begin{aligned} \Pr[\Theta = \theta|A] &= \frac{\Pr[A|\Theta = \theta]\Pr[\Theta = \theta]}{\sum_{\theta=1,3} \Pr[A|\Theta = \theta]\Pr[\Theta = \theta]} \\ &= \frac{(\theta^2 e^{-\theta})(\frac{1}{\theta} e^{-1/\theta})(\frac{1}{\theta} e^{-3/\theta})(\frac{1}{2})}{\sum_{\theta=1,3} (\theta^2 e^{-\theta})(\frac{1}{\theta} e^{-1/\theta})(\frac{1}{\theta} e^{-3/\theta})(\frac{1}{2})} \\ &= \frac{e^{-(\theta^2+4)/\theta}}{\sum_{\theta=1,3} e^{-(\theta^2+4)/\theta}} \\ &= \begin{cases} 0{,}3392, & \theta = 1 \\ 0{,}6608, & \theta = 3, \end{cases} \end{aligned}$$

d'où

$$E[S|A] = (1)^2(0{,}3392) + (3)^2(0{,}6608) = 6{,}286.$$

3.18 Nous avons $S|\Theta \sim$ Géométrique(Θ) et $\Theta \sim$ Bêta(a, b).

a) $\mu(\Theta) = E[S|\Theta] = \sum_{x=0}^{\infty} x\theta(1-\theta)^x = (1-\theta)/\theta$. (Ce résultat est aisément obtenu en dérivant $\sum_{x=0}^{\infty} \theta(1-\theta)^x = 1$.)

b) La prime collective est

$$\begin{aligned} m = E[\mu(\Theta)] &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \theta^{a-2}(1-\theta)^b \, d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a-1)\Gamma(b+1)}{\Gamma(a+b)} \\ &= \frac{b}{a-1}. \end{aligned}$$

c) Soit $\mathbf{S} = (S_1, \dots, S_n)$ et $\mathbf{x} = (x_1, \dots, x_n)$. Premièrement, $\Pr[\mathbf{S} = \mathbf{x}|\Theta = \theta] = \prod_{t=1}^{n} \Pr[S_t = x_t|\Theta = \theta] = \theta^n (1-\theta)^{\sum x_t}$. Ensuite,

$$\begin{aligned} u(\theta|\mathbf{x}) &\propto \theta^n (1-\theta)^{\sum x_t} \theta^{a-1}(1-\theta)^{b-1} \\ &= \theta^{a+n-1}(1-\theta)^{b+\sum x_t - 1}, \end{aligned}$$

d'où $\Theta|\mathbf{S} = \mathbf{x} \sim$ Bêta$(\tilde{a} = a + n, \tilde{b} = b + \sum x_t)$.

d) Comme la densité à postériori de Θ est du même type que sa densité à priori, nous pouvons calculer en premier la densité marginale de S. La distribution prédictive sera du même type, mais avec des paramètres modifiés. Or,

$$\begin{aligned} \Pr[S = x] &= \int_0^1 \Pr[S = x|\Theta = \theta] u(\theta) \, d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 \theta^a (1-\theta)^{b+x-1} \, d\theta \\ &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+1)\Gamma(b+x)}{\Gamma(a+b+x+1)} \end{aligned}$$

et donc

$$\Pr[S_{n+1} = x|\mathbf{S} = \mathbf{x}] = \frac{\Gamma(\tilde{a}+\tilde{b})}{\Gamma(\tilde{a})\Gamma(\tilde{b})} \frac{\Gamma(\tilde{a}+1)\Gamma(\tilde{b}+x)}{\Gamma(\tilde{a}+\tilde{b}+x+1)},$$

avec $\tilde{a} = a + n$ et $\tilde{b} = b + \sum x_t$.

e) Utilisons le résultat en c) car trouver l'espérance de la distribution prédictive trouvée en d) peut être quelque peu compliqué. Des résultats de b) et c),

$$E[\mu(\Theta)|\mathbf{S}] = \frac{\tilde{b}}{\tilde{a}-1} = \frac{b + \sum_{t=1}^{n} S_t}{a+n-1}.$$

f) $E[\mu(\Theta)|\mathbf{S}] = z\bar{S} + (1-z)m$ avec $z = n/(n+a-1)$.

3.19 Nous avons $S|\Theta \sim \text{Binomiale}(\nu, \Theta)$ et $\Theta \sim \text{Bêta}(a, b)$.

a) La prime de risque est $\mu(\Theta) = E[S|\Theta] = \nu\Theta$.

b) La prime collective est $E[\mu(\Theta)] = \nu E[\Theta] = \nu a/(a+b)$.

c) La distribution à postériori de Θ sachant $S_1 = x_1, \dots, S_n = x_n$ est

$$\begin{aligned} u(\theta|x_1,\dots,x_n) &\propto u(\theta)\prod_{t=1}^{n}\Pr[S = x_t|\Theta = \theta] \\ &\propto \theta^{a-1}(1-\theta)^{b-1}\prod_{t=1}^{n}\theta^{x_t}(1-\theta)^{\nu-x_t} \\ &= \theta^{a+\sum x_t-1}(1-\theta)^{b+n\nu+\sum x_t-1}. \end{aligned}$$

Il s'agit d'une distribution bêta avec paramètres $\tilde{a} = a + \sum_{t=1}^{n} x_t$ et $\tilde{b} = b + n\nu - \sum_{t=1}^{n} x_t$.

d) Comme la distribution à postériori est de même famille que la distribution à priori, la prime bayésienne est de la même forme que la prime collective, mais avec des paramètres modifiés :

$$\begin{aligned} E[\mu(\Theta)|S_1,\dots,S_n] &= \frac{\nu\tilde{a}}{\tilde{a}+\tilde{b}} \\ &= \frac{\nu(a+\sum_{t=1}^{n}S_t)}{n\nu+a+b} \\ &= \left(\frac{n\nu}{n\nu+a+b}\right)\left(\frac{\sum_{t=1}^{n}S_t}{n}\right) \\ &\quad+\left(\frac{a+b}{n\nu+a+b}\right)\left(\frac{\nu a}{a+b}\right) \\ &= z\bar{S} + (1-z)E[\mu(\Theta)] \end{aligned}$$

où

$$z = \frac{n}{n+(a+b)/\nu}.$$

Cette combinaison de distributions est la convolution du cas Bernoulli/Bêta, la binomiale étant une somme de Bernoulli.

3.20 Nous avons $S|\Theta \sim \text{Gamma}(\tau, \Theta)$ et $\Theta \sim \text{Gamma}(\alpha, \lambda)$.

a) La densité marginale de S est calculée comme suit :

$$\begin{aligned} f(x) &= \int_0^\infty f(x|\theta)\,u(\theta)\,d\theta \\ &= \frac{\lambda^\alpha x^{\tau-1}}{\Gamma(\tau)\Gamma(\alpha)} \int_0^\infty \theta^\tau e^{-\theta x} \theta^{\alpha-1} e^{-\lambda\theta}\,d\theta \\ &= \frac{\lambda^\alpha x^{\tau-1}}{\Gamma(\tau)\Gamma(\alpha)} \frac{\Gamma(\alpha+\tau)}{(x+\lambda)^{\alpha+\tau}}, \end{aligned}$$

laquelle est une densité de Pareto généralisée avec paramètres α, λ et τ.

b) La prime de risque est $\mu(\Theta) = E[S|\Theta] = \tau/\Theta$.

c) De la distribution marginale de S, nous avons $m = E[S] = \tau\lambda/(\alpha - 1)$ (section C.2.6). De la prime de risque, $m = \tau E[1/\Theta] = \tau\lambda/(\alpha-1)$. Les deux approches sont équivalentes.

d) La densité de la distribution à postériori de Θ est

$$\begin{aligned} u(\theta|x_1, \dots, x_n) &\propto u(\theta) \prod_{t=1}^n f(x_t|\theta) \\ &\propto \theta^{\alpha-1} e^{-\lambda\theta} \prod_{t=1}^n \theta^\tau e^{-\theta x_t} \\ &= \theta^{\alpha+n\tau-1} e^{-(\lambda + \sum x_t)\theta}, \end{aligned}$$

d'où la distribution à postériori de Θ est une gamma avec paramètres $\tilde{\alpha} = \alpha + n\tau$ et $\tilde{\lambda} = \lambda + \sum_{t=1}^n x_t$. La distribution à priori de Θ est la conjuguée naturelle de la fonction de vraisemblance $f(x|\theta)$.

e) La distribution prédictive est une Pareto généralisée avec paramètres $\tilde{\alpha} = \alpha + n\tau$, $\tilde{\lambda} = \lambda + \sum_{t=1}^n x_t$ et τ.

f) En utilisant la distribution à postériori ou la distribution prédictive, la prime bayésienne pour l'année $n + 1$ est simplement la prime collective avec les paramètres modifiés. Par conséquent,

$$\begin{aligned} E[\mu(\Theta)|S_1, \dots, S_n] &= \tau \left(\frac{\tilde{\lambda}}{\tilde{\alpha} - 1} \right) \\ &= \tau \left(\frac{\lambda + \sum_{t=1}^n S_t}{\alpha + n\tau - 1} \right). \end{aligned}$$

g) La prime bayésienne ci-dessus peut être réécrite comme suit :

$$\begin{aligned} E[\mu(\Theta)|S_1,\dots,S_n] &= \left(\frac{n\tau}{n\tau+\alpha-1}\right)\sum_{t=1}^{n}\frac{S_t}{n} + \left(\frac{\alpha-1}{n\tau+\alpha-1}\right)\left(\frac{\tau\lambda}{\alpha-1}\right) \\ &= z\bar{S} + (1-z)m \end{aligned}$$

avec

$$z = \frac{n}{n+(\alpha-1)/\tau}.$$

3.21 Dans le cas exponentielle/gamma, la prime bayésienne est

$$B_{n+1} = \frac{\lambda + \sum_{t=1}^{n} x_t}{\alpha+n-1}.$$

Ici $\alpha = 7$ et $\lambda = 42$. Si $B_5 = 9$, alors $42 + \sum_{t=1}^{4} x_t = 9(7+4-1) = 90$. Par conséquent,

$$\begin{aligned} B_6 &= \frac{42 + \sum_{t=1}^{4} x_t + x_5}{7+5-1} \\ &= \frac{90 + x_5}{11} \\ &= 8{,}5, \end{aligned}$$

d'où $x_5 = 3{,}5$.

3.22 Dans le modèle Poisson/gamma, $z = n/(n+\lambda)$. Si $z = 0{,}8$ quand $n = 4$, alors $\lambda = 1$. Pour doubler la variance d'une distribution gamma sans changer son espérance, il faut diminuer de moitié chacun des paramètres. Le nouveau paramètre λ est donc $1/2$. Nous cherchons la nouvelle valeur de n telle que $n/(n+1/2) = 0{,}8$, d'où $n = 2$.

3.23 Dans le modèle Poisson/gamma, nous savons que la distribution prédictive est une binomiale négative avec paramètres $r = \alpha + \sum_{t=1}^{n} x_t$ et $\theta = (\lambda+n)/(\lambda+n+1)$. Ici, nous pouvons identifier $r = 7$, $\theta = 0{,}9$ et nous avons $n = 2$ et $x_1 + x_2 = 3$. Par conséquent, $\alpha = 4$, $\lambda = 7$ et l'espérance de la distribution à priori — une gamma — est $4/7$.

3.24 Nous avons un modèle exponentielle/gamma. Comme le confirme le tableau de l'annexe B, la distribution marginale de S dans un tel cas est une Pareto(α, λ) et la distribution de $\Theta|S_1 = x_1, \dots, S_n = x_n$ une Gamma$(\alpha+n, \lambda + \sum_{t=1}^{n} x_t)$. Par conséquent, la distribution de $S|S_1 = x_1, \dots, S_n = x_n$ est une Pareto$(\alpha+n, \lambda + \sum_{t=1}^{n} x_t)$ et

$$\begin{aligned} \Pr(S_4 \leq 5|S_1 = 2, S_2 = 3, S_3 = 7) &= 1 - \left(\frac{8+12}{8+12+5}\right)^{2+3} \\ &= 0{,}6723. \end{aligned}$$

3.25 Nous savons que, dans le cas Poisson/gamma, $z = n/(n+\lambda)$. Ici, $n = 5$, et $\bar{S} = (3 + 1 + 5 + 4 + 2)/5 = 3$. De plus, la prime collective, m, est égale à 2 dans tous les cas ci-dessous. Nous effectuons les calculs de primes bayésiennes avec ces résultats ainsi qu'avec la fonction `cm` du paquetage **actuar**.

a) Ici, $z = 0{,}5$ et donc $\pi_6 = 2{,}5$.

```
> x <- c(3, 1, 5, 4, 2)
> predict(cm("bayes", x, likelihood = "poisson",
+            shape = 10, rate = 5))
[1] 2.5
```

b) Ici, $z = 1/6$ et donc $\pi_6 = 2{,}17$.

```
> predict(cm("bayes", x, likelihood = "poisson",
+            shape = 50, rate = 25))
[1] 2.166667
```

c) Ici, $z = 0{,}9524$ et donc $\pi_6 = 2{,}9524$.

```
> predict(cm("bayes", x, likelihood = "poisson",
+            shape = 0.5, rate = 0.25))
[1] 2.952381
```

Plus le paramètre λ de la distribution gamma est petit, moins certaine est la valeur du paramètre de risque θ. Par conséquent, nous accordons plus d'importance à l'expérience individuelle en augmentant le facteur de crédibilité.

3.26 La distribution marginale donnée est une Pareto$(1{,}5, 100)$. Or, nous savons que la Pareto est la distribution marginale dans le mélange exponentielle/gamma et que, dans ce cas, la distribution prédictive sera une Pareto$(1{,}5 + n, 100 + \sum_{t=1}^{n} S_t)$. Par conséquent,

$$E[S_6|S_1 = S_2 = \cdots = S_5 = 0] = \frac{100}{1{,}5 + 5 - 1} = \frac{200}{11}.$$

3.27 Il est utile de remarquer ici que $X|\Theta \sim$ Binomiale négative$(5, \Theta)$ et $\Theta \sim$ Bêta$(6, 4)$. Par conséquent, $\mu(\Theta) = r(1 - \Theta)/\Theta = 5(1 - \Theta)/\Theta$ et

$$\begin{aligned} u(\theta|x_1 = 7, x_2 = 13, x_3 = 1, x_4 = 4) &\propto \theta^5(1-\theta)^3 \prod_{t=1}^{4} \theta^5(1-\theta)^{x_t} \\ &= \theta^{25}(1-\theta)^{3+\sum x_t} \\ &= \theta^{25}(1-\theta)^{28}, \end{aligned}$$

d'où $\Theta|X_1 = 7, X_2 = 13, X_3 = 1, X_4 = 4 \sim \text{Bêta}(26, 29)$. Ainsi,

$$\begin{aligned}
E[\mu(\Theta)|X_1 = 7, X_2 = 13, X_3 = 1, X_4 = 4] \\
&= 5\,\frac{\Gamma(55)}{\Gamma(26)\Gamma(29)} \int_0^1 \theta^{24}(1-\theta)^{29}\,d\theta \\
&= 5\left(\frac{\Gamma(55)}{\Gamma(26)\Gamma(29)}\right)\left(\frac{\Gamma(25)\Gamma(30)}{\Gamma(55)}\right) \\
&= \frac{29}{5} \\
&= 5{,}8.
\end{aligned}$$

3.28 Nous avons le modèle suivant :

$$\begin{aligned}
S|\Theta = \theta &\sim \text{Poisson}(\theta) \\
\Theta &\sim \text{Gamma}(\alpha, \lambda)
\end{aligned}$$

La covariance entre S_1 et S_2 est

$$\begin{aligned}
\text{Cov}(S_1, S_2) &= E[\text{Cov}(S_1, S_2|\Theta)] + \text{Cov}(E[S_1|\Theta], E[S_2|\Theta]) \\
&= \text{Var}[\mu(\Theta)] \\
&= \text{Var}[\Theta] \\
&= \frac{\alpha}{\lambda^2}.
\end{aligned}$$

Or, nous savons que dans le modèle Poisson/gamma, la distribution prédictive est une binomiale négative de paramètres $r = \alpha + \sum x_j$ et $p = (\lambda + n)/(\lambda + n + 1)$. Nous avons donc $7 = \alpha + 3$ d'où $\alpha = 4$ et $0{,}9 = (\lambda + 2)/(\lambda + 3)$ d'où $\lambda = 7$. Par conséquent, $\text{Cov}\,(S_1, S_2) = \frac{4}{49}$.

3.29 a) Nous devons démontrer que la densité à postériori $u(\theta|x_1, \dots, x_n)$ est de la même famille que la densité à priori $u(\theta)$. De manière usuelle,

$$\begin{aligned}
u(\theta|x_1, \dots, x_n) &\propto u(\theta) \prod_{t=1}^{n} f(x_t|\theta) \\
&\propto q(\theta)^{-t_0} e^{-\theta x_0} \prod_{t=1}^{n} \frac{e^{-\theta x_t}}{q(\theta)} \\
&= q(\theta)^{-t_0 - n} e^{-\theta(x_0 + \sum x_t)}
\end{aligned}$$

qui est, en effet, de la même famille que $u(\theta)$ avec paramètres modifiés $\tilde{t}_0 = t_0 + n$ et $\tilde{x}_0 = x_0 + \sum_{t=1}^{n} x_t$.

b) Premièrement, remarquons que $q(\theta) = \int_{-\infty}^{\infty} p(x)e^{-\theta x}\,dx$ pour faire de $f(x|\theta)$ une densité. Ensuite,

$$\begin{aligned}
\mu(\theta) &= \int_{-\infty}^{\infty} x f(x|\theta)\,dx \\
&= \frac{1}{q(\theta)} \int_{-\infty}^{\infty} x p(x) e^{-\theta x}\,dx \\
&= \frac{1}{q(\theta)} \int_{-\infty}^{\infty} \frac{d}{d\theta}(-p(x)e^{-\theta x})\,dx \\
&= -\frac{1}{q(\theta)} \frac{d}{d\theta} \int_{-\infty}^{\infty} p(x)e^{-\theta x}\,dx \\
&= -\frac{q'(\theta)}{q(\theta)} = -\frac{d}{d\theta} \ln q(\theta).
\end{aligned}$$

c) Nous avons

$$\begin{aligned}
\frac{d}{d\theta} u(\theta) &= \frac{1}{d(x_0, t_0)} \frac{d}{d\theta} q(\theta)^{-t_0} e^{-\theta x_0} \\
&= \frac{1}{d(x_0, t_0)} \left[-t_0 q(\theta)^{-t_0 - 1} q'(\theta) e^{-\theta x_0} - x_0 q(\theta)^{-t_0} e^{-\theta x_0} \right] \\
&= \left[t_0 \left(-\frac{q'(\theta)}{q(\theta)} \right) - x_0 \right] \frac{q(\theta)^{-t_0} e^{-\theta x_0}}{d(x_0, t_0)} \\
&= (t_0 \mu(\theta) - x_0) u(\theta).
\end{aligned}$$

d) Supposons, sans perte de généralité, que le domaine de définition de la densité obtenue en c) est $(-\infty, \infty)$. En supposant que $u(-\infty) = u(\infty) = 0$, nous obtenons

$$\int_{-\infty}^{\infty} \frac{d}{d\theta} u(\theta)\,d\theta = 0.$$

Or,

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{d}{d\theta} u(\theta)\,d\theta &= t_0 \int_{-\infty}^{\infty} \mu(\theta) u(\theta)\,d\theta - x_0 \int_{-\infty}^{\infty} u(\theta)\,d\theta \\
&= t_0 E[\mu(\Theta)] - x_0,
\end{aligned}$$

d'où $E[\mu(\Theta)] = x_0 / t_0$.

e) Nous savons que la prime bayésienne est de la même forme que la prime collective, mais avec des paramètres modifiés. Nous avons donc

$$\begin{aligned} E[\mu(\Theta)|X_1,\dots,X_n] &= \frac{\tilde{x}_0}{\tilde{t}_0} \\ &= \frac{x_0 + \sum_{t=1}^n X_t}{t_0 + n} \\ &= \left(\frac{n}{n+t_0}\right)\bar{X} + \left(\frac{t_0}{n+t_0}\right)\left(\frac{x_0}{t_0}\right), \end{aligned}$$

qui est une prime de crédibilité.

3.30 Nous avons

$$f(x|\theta) = \frac{1}{\theta}e^{-x/\theta}, \quad x > 0$$

et

$$u(\theta) = \frac{\lambda^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\lambda\theta}, \quad \theta > 0.$$

La distribution à postériori de Θ est

$$\begin{aligned} u(\theta|x_1,\dots,x_n) &\propto \theta^{\alpha-1}e^{-\lambda\theta}\prod_{t=1}^n \theta^{-1}e^{\theta^{-1}x_t} \\ &= \theta^{\alpha-n-1}e^{-\lambda\theta+\theta^{-1}\sum x_t}, \end{aligned}$$

qui n'est clairement pas une gamma. Donc, la gamma n'est pas la conjuguée naturelle de l'exponentielle avec moyenne Θ. (La gamma inverse l'est, par contre.)

3.31 Voir l'exercice 3.20.

3.32 a) La fonction de log-vraisemblance est

$$\begin{aligned} \ell(\theta) &= \sum_{i=1}^n \ln f(x_i;\theta) \\ &= \sum_{i=1}^n (\ln\theta + \theta\ln x_0 - (\theta+1)\ln x_i) \\ &= n\ln\theta + n\theta\ln x_0 - (\theta+1)\sum_{i=1}^n \ln x_i, \end{aligned}$$

d'où

$$\begin{aligned}
\ell'(\theta) &= \frac{n}{\theta} - n\ln x_0 - \sum_{i=1}^{n} \ln x_i \\
&= \frac{n}{\theta} - \sum_{i=1}^{n} (\ln x_i - \ln x_0) \\
&= \frac{n}{\theta} - \sum_{i=1}^{n} \ln\left(\frac{x_i}{x_0}\right).
\end{aligned}$$

En résolvant pour θ l'équation $\ell'(\theta) = 0$, nous obtenons aisément l'estimateur du maximum de vraisemblance

$$\hat{\theta}^{\text{EMV}} = \frac{n}{\sum_{i=1}^{n} \ln(X_i/x_0)}.$$

b) Nous devons démontrer si $S|\Theta = \theta \sim$ Pareto translatée(θ) et que $\Theta \sim$ Gamma(α, λ), alors la distribution de $\Theta|S_1 = x_1, \dots, S_n = x_n$ est aussi une gamma avec de nouveaux paramètres. Il est plus simple, ici, de travailler avec la densité de la Pareto translatée exprimée sous la forme

$$f(x|\theta) = \frac{\theta}{x_0}\left(\frac{x}{x_0}\right)^{-(\theta+1)}.$$

Ensuite, nous avons, comme d'habitude,

$$\begin{aligned}
u(\theta|x_1, \dots, x_n) &\propto u(\theta) \prod_{i=1}^{n} f(x_i|\theta) \\
&\propto \theta^{\alpha-1} e^{-\lambda\theta} \prod_{i=1}^{n} \theta \left(\frac{x_i}{x_0}\right)^{-\theta} \\
&= \theta^{\alpha+n-1} e^{-\lambda\theta} \prod_{i=1}^{n} \left(\frac{x_i}{x_0}\right)^{-\theta} \\
&= \theta^{\alpha+n-1} e^{-\lambda\theta} e^{-\theta \sum_{i=1}^{n} \ln(x_i/x_0)} \\
&= \theta^{\alpha+n-1} \exp\left\{-\theta\left(\lambda + \sum_{i=1}^{n} \ln\left(\frac{x_i}{x_0}\right)\right)\right\},
\end{aligned}$$

d'où nous identifions que la distribution à postériori de Θ est gamma de paramètres

$$\begin{aligned}
\tilde{\alpha} &= \alpha + n \\
\tilde{\lambda} &= \lambda + \sum_{i=1}^{n} \ln\left(\frac{x_i}{x_0}\right).
\end{aligned}$$

c) L'estimateur à priori de θ (ou la « prime collective ») est

$$E[\Theta] = \frac{\alpha}{\lambda}.$$

Dans la mesure où la distribution à postériori est du même type que la distribution à priori, nous avons directement que l'estimateur bayésien de θ est

$$\begin{aligned}
\hat{\Theta} &= \frac{\tilde{\alpha}}{\tilde{\lambda}} \\
&= \frac{\alpha + n}{\lambda + \sum_{i=1}^{n} \ln(x_i/x_0)} \\
&= \eta \hat{\theta}^{\text{EMV}} + (1 - \eta)\frac{\alpha}{\lambda}
\end{aligned}$$

avec

$$\eta = \frac{\sum_{i=1}^{n} \ln(x_i/x_0)}{\lambda + \sum_{i=1}^{n} \ln(x_i/x_0)}.$$

Contrairement aux cas de crédibilité bayésienne linéaire, le facteur η dépend des observations et il ne se trouve pas confiné à l'intervalle $(0, 1)$. Il ne s'agit donc pas d'un facteur de crédibilité. Par conséquent, $\hat{\Theta}$ ne constitue pas une prime de crédibilité. Une seule des deux conditions de Jewell (1974) est vérifiée dans le cas présent.

Chapitre 4

4.1 Nous pouvons résoudre ce problème de deux façons.

a) De manière similaire à l'exercice 3.10, nous trouvons la distribution du montant total des sinistres S en considérant la probabilité de ne pas avoir d'accident. En identifiant les femmes par $\theta = 0{,}2$ et les hommes par $\theta = 0{,}4$, nous avons

$$\Pr[S = x | \Theta = 0{,}2] = \begin{cases} 0{,}8, & x = 0 \\ (0{,}8)(0{,}2) = 0{,}16, & x = 100 \\ (0{,}1)(0{,}2) = 0{,}02, & x = 200 \\ (0{,}1)(0{,}2) = 0{,}02, & x = 400 \end{cases}$$

$$\Pr[S = x | \Theta = 0{,}4] = \begin{cases} 0{,}6, & x = 0 \\ (0{,}8)(0{,}4) = 0{,}32, & x = 100 \\ (0{,}1)(0{,}4) = 0{,}04, & x = 200 \\ (0{,}1)(0{,}4) = 0{,}04, & x = 400 \end{cases}$$

et

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}5, & \theta = 0{,}2 \\ 0{,}5, & \theta = 0{,}4. \end{cases}$$

Ainsi,

$$\mu(0{,}2) = 28 \qquad \sigma^2(0{,}2) = 4\,816$$
$$\mu(0{,}4) = 56 \qquad \sigma^2(0{,}4) = 8\,064$$

et

$$\begin{aligned} s^2 &= E[\sigma^2(\Theta)] \\ &= \frac{4\,816 + 8\,064}{2} \\ &= 6\,440 \\ a &= \mathrm{Var}[\mu(\Theta)] \\ &= \frac{28^2 + 56^2}{2} - \left(\frac{28 + 56}{2}\right)^2 \\ &= 196, \end{aligned}$$

d'où le facteur de crédibilité est $z = n/(n + s^2/a) = n/(n + 32{,}86)$.

b) Nous posons $S = X_1 + \cdots + X_N$ avec $N|\Theta = \theta \sim \text{Bernoulli}(\theta)$,

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}5, & \theta = 0{,}2 \\ 0{,}5, & \theta = 0{,}4. \end{cases}$$

et

$$\Pr[X = x] = \begin{cases} 0{,}8, & x = 100 \\ 0{,}1, & x = 200 \\ 0{,}1, & x = 400. \end{cases}$$

Nous avons alors

$$\begin{aligned} \mu(\Theta) &= E[S|\Theta] \\ &= E[N|\Theta]E[X] \\ &= 140\Theta \\ \sigma^2(\Theta) &= \mathrm{Var}[S|\Theta] \\ &= \mathrm{Var}[N|\Theta]E[X]^2 + E[N|\Theta]\,\mathrm{Var}[X] \\ &= 19\,600\Theta(1 - \Theta) + 8\,400\Theta \\ &= 28\,000\Theta - 19\,600\Theta^2, \end{aligned}$$

d'où

$$\begin{aligned} s^2 &= E[\sigma^2(\Theta)] \\ &= 28\,000E[\Theta] - 19\,600E[\Theta^2] \\ &= 28\,000(0{,}3) - 19\,600(0{,}1) \\ &= 6\,440 \end{aligned}$$

et

$$\begin{aligned} a &= \mathrm{Var}[\mu(\Theta)] \\ &= 19\,600\,\mathrm{Var}[\Theta] \\ &= 19\,600(0{,}01) \\ &= 196 \end{aligned}$$

Le facteur de crédibilité est donc $z = n/(n + s^2/a) = n/(n + 32{,}86)$.

4.2 Soit S le nombre de balles rouges tirées et Θ la probabilité de tirer une balle rouge. On identifie par $\theta = 0{,}75$ l'urne A, $\theta = 0{,}50$ l'urne B et $\theta = 0{,}25$ l'urne C. Nous avons le modèle $S|\Theta = \theta \sim \text{Binomiale}(5, \theta)$ et

$$\Pr[\Theta = \theta] = \begin{cases} \frac{1}{6}, & \theta = 0{,}75 \\ \frac{3}{6}, & \theta = 0{,}50 \\ \frac{2}{6}, & \theta = 0{,}25. \end{cases}$$

Ainsi, $\mu(\Theta) = E[S|\Theta] = 5\Theta$ et $\sigma^2(\Theta) = \mathrm{Var}[S|\Theta] = 5\Theta(1-\Theta)$, d'où

$$\begin{aligned} m &= E[\mu(\Theta)] = 5E[\Theta] = 2{,}2917 \\ s^2 &= E[\sigma^2(\Theta)] = 5(E[\Theta] - E[\Theta^2]) = 1{,}0938 \end{aligned}$$

et

$$a = \mathrm{Var}[\mu(\Theta)] = 25\,\mathrm{Var}[\Theta] = 0{,}7378.$$

Après une expérience lors de laquelle nous avons obtenu $\bar{S} = S_1 = 3$, le facteur de crédibilité est $z = 1/(1 + 1{,}0938/0{,}7378) = 0{,}4028$. L'estimateur de Bühlmann du nombre de balles rouges lors du prochain tir dans la même urne est donc

$$\begin{aligned} \pi_2 &= 0{,}4028(3) + (1 - 0{,}4028)(2{,}2917) \\ &= 2{,}5770. \end{aligned}$$

4.3 a) Soit θ_A le niveau de risque de la classe A et θ_B celui de la classe B. Le modèle est le suivant :

$$\Pr[S = x|\Theta = \theta_A] = \begin{cases} 0{,}2, & x = 2 \\ 0{,}8, & x = 0 \end{cases}$$

$$\Pr[S = x|\Theta = \theta_B] = \begin{cases} 0{,}2, & x = c \\ 0{,}8, & x = 0 \end{cases}$$

$$\Pr[\Theta = \theta] = \begin{cases} \frac{1}{2}, & \theta = \theta_A \\ \frac{1}{2}, & \theta = \theta_B. \end{cases}$$

b) Avec l'information de a), nous avons

$$E[S|\Theta = \theta_A] = 0{,}4 \qquad \mathrm{Var}[S|\Theta = \theta_A] = 0{,}64$$
$$E[S|\Theta = \theta_B] = 0{,}2c \qquad \mathrm{Var}[S|\Theta = \theta_B] = 0{,}16c^2.$$

Ainsi,

$$s^2 = E[\mathrm{Var}[S|\Theta]] = 0{,}32 + 0{,}08c^2$$
$$a = \mathrm{Var}[E[S|\Theta]] = 0{,}04 - 0{,}04c + 0{,}01c^2$$

et, donc,

$$K = \frac{0{,}32 + 0{,}08c^2}{0{,}04 - 0{,}04c + 0{,}01c^2}.$$

Comme il n'y a qu'une année d'expérience, le facteur de crédibilité est $z = 1/(1 + K)$. Or, $\lim_{c\to\infty} K = 8$ d'où $\lim_{c\to\infty} z = 1/9$.

4.4 Nous utiliserons θ_1 pour représenter les bons risques, θ_2 les risques moyens et θ_3 les mauvais risques. Nous avons le modèle suivant :

$$X|\Theta = \theta_1 \sim \mathrm{Gamma}(4, 2)$$
$$X|\Theta = \theta_2 \sim \mathrm{Gamma}(4, 1)$$
$$X|\Theta = \theta_3 \sim \mathrm{Gamma}(10, 2)$$

et

$$\Pr[\Theta = \theta_i] = \begin{cases} 0{,}25, & i = 1 \\ 0{,}60, & i = 2 \\ 0{,}15, & i = 3. \end{cases}$$

Ainsi,

$$\begin{aligned} \mu(\theta_1) &= 2, & \sigma^2(\theta_1) &= 1 \\ \mu(\theta_2) &= 4, & \sigma^2(\theta_2) &= 4 \\ \mu(\theta_3) &= 5, & \sigma^2(\theta_3) &= 2{,}5 \end{aligned}$$

et

$$\begin{aligned} m &= E[\mu(\Theta)] \\ &= 0{,}25(2) + 0{,}60(4) + 0{,}15(5) = 3{,}65 \\ s^2 &= E[\sigma^2(\Theta)] \\ &= 0{,}25(1) + 0{,}60(4) + 0{,}15(2{,}5) = 3{,}025 \\ a &= \mathrm{Var}[\mu(\Theta)] \\ &= 0{,}25(2^2) + 0{,}60(4^2) + 0{,}15(5^2) - m^2 = 1{,}0275. \end{aligned}$$

Le facteur de crédibilité après deux années est

$$z = \frac{2}{2 + 3{,}025/1{,}0275} = 0{,}4045$$

et, enfin, la prime de Bühlmann pour la troisième année après une expérience individuelle moyenne de $(1+2)/2 = 1{,}5$ est

$$\pi_3 = z(1{,}5) + (1-z)(3{,}65) = 2{,}78.$$

4.5 Les valeurs des paramètres de structures sont :

$$\begin{aligned} m &= E[\mu(\Theta)] \\ &= \frac{1\,000}{5\,000}(50) + \frac{2\,000}{5\,000}(200) \\ &\quad + \frac{1\,000}{5\,000}(500) + \frac{1\,000}{5\,000}(1\,000) \\ &= 390, \\ s^2 &= E[\sigma^2(\Theta)] \\ &= \frac{1\,000}{5\,000}(100\,000) + \frac{2\,000}{5\,000}(500\,000) \\ &\quad + \frac{1\,000}{5\,000}(500\,000) + \frac{1\,000}{5\,000}(500\,000) \\ &= 420\,000 \end{aligned}$$

et

$$\begin{aligned}
a &= \mathrm{Var}[\mu(\Theta)] \\
&= E[\mu(\Theta)^2] - m^2 \\
&= 114\,400.
\end{aligned}$$

Le facteur de crédibilité après quatre ans est donc

$$z = \frac{4}{4 + 420\,000/114\,400} = 0{,}5214$$

et la prime de crédibilité est

$$\pi_5 = (0{,}5214)\left(\frac{800}{4}\right) + (0{,}4786)(390) = 290{,}93.$$

4.6 Le paramètre θ_1 représente les bons risques, θ_2 les risques moyens et θ_3 les mauvais risques. Nous avons donc le modèle suivant :

$$\begin{aligned}
X|\Theta = \theta_1 &\sim \text{Pareto}(4, 1200) \\
X|\Theta = \theta_2 &\sim \text{Gamma}(1000, 2) \\
X|\Theta = \theta_3 &\sim \text{Exponentielle}(0{,}00125)
\end{aligned}$$

et

$$\Pr[\Theta = \theta_i] = \begin{cases} 0{,}3, & i = 1 \\ 0{,}5, & i = 2 \\ 0{,}2, & i = 3. \end{cases}$$

Ensuite,

$$\begin{aligned}
\mu(\theta_1) &= 400, & \sigma^2(\theta_1) &= 320\,000 \\
\mu(\theta_2) &= 500, & \sigma^2(\theta_2) &= 250 \\
\mu(\theta_3) &= 800, & \sigma^2(\theta_3) &= 640\,000
\end{aligned}$$

et

$$\begin{aligned}
m &= E[\mu(\Theta)] = 530 \\
s^2 &= E[\sigma^2(\Theta)] = 224\,125 \\
a &= \mathrm{Var}[\mu(\Theta)] = 20\,100.
\end{aligned}$$

Le facteur de crédibilité après quatre années d'expérience est donc

$$z = \frac{4}{4 + 224\,125/20\,100} = 0{,}2640.$$

4.7 Nous avons $S|\Theta = \theta \sim \text{Gamma}(5, \theta)$, d'où $\mu(\theta) = 5/\theta$ et $\sigma^2(\theta) = 5/\theta^2 = \mu(\theta)^2/5$. On nous donne la fonction de masse de probabilité de la prime de risque :

$$\Pr[\mu(\Theta) = \mu] = \begin{cases} 0{,}3, & \mu = 2\,500 \\ 0{,}5, & \mu = 4\,000 \\ 0{,}2, & \mu = 5\,000 \end{cases}$$

ou, de manière équivalente,

$$\Pr[\Theta = \theta] = \begin{cases} 0{,}3, & \theta = 0{,}002 \\ 0{,}5, & \theta = 0{,}00125 \\ 0{,}2, & \theta = 0{,}001. \end{cases}$$

Par conséquent,

$$\begin{aligned} m &= E[\mu(\Theta)] \\ &= 2\,500(0{,}3) + 4\,000(0{,}5) + 5\,000(0{,}2) \\ &= 3\,750, \\ s^2 &= E[\sigma^2(\Theta)] \\ &= \frac{1}{5}\,E[\mu(\Theta)^2] \\ &= \frac{(2\,500)^2(0{,}3) + (4\,000)^2(0{,}5) + (5\,000)^2(0{,}2)}{5} \\ &= 2\,975\,000, \\ a &= \text{Var}[\mu(\Theta)] \\ &= (2\,500)^2(0{,}3) + (4\,000)^2(0{,}5) + (5\,000^2)(0{,}2) - m^2 \\ &= 812\,500, \\ K &= \frac{s^2}{a} \\ &= 3{,}6615 \end{aligned}$$

et

$$\begin{aligned} \pi_5 &= \frac{4}{4+K}\,\frac{20\,000}{4} + \frac{K}{4+K}\,3750 \\ &= 4\,402{,}61. \end{aligned}$$

4.8 Nous sommes en présence d'une combinaison géométrique/bêta avec $a = b = 3$. Voir les formules de l'annexe B. La prime bayésienne et la prime de Bühlmann sont égales dans un tel cas.

4.9 Dans tous les cas ci-dessous, le facteur de crédibilité est $z = n/(n+K)$, où $K = s^2/a$, $s^2 = E[\sigma^2(\Theta)]$, $a = \mathrm{Var}[\mu(\Theta)]$ et la prime de crédibilité est $z\bar{S} + (1-z)m$.

a) Nous avons $S|\Theta \sim \text{Bernoulli}(\Theta)$, $\Theta \sim \text{Bêta}(a, b)$, d'où $\mu(\Theta) = \Theta$ et $\sigma^2(\Theta) = \Theta(1-\Theta) = \Theta - \Theta^2$. Alors,

$$\begin{aligned} m &= \frac{a}{a+b} \\ s^2 &= \frac{a}{a+b} - \frac{a(a+1)}{(a+b)(a+b+1)} = \frac{ab}{(a+b)(a+b+1)} \\ a &= \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2} = \frac{ab}{(a+b)^2(a+b+1)}, \end{aligned}$$

et $K = a + b$.

b) Nous avons $S|\Theta \sim \text{Géométrique}(\Theta)$, $\Theta \sim \text{Bêta}(a, b)$, d'où $\mu(\Theta) = (1-\Theta)/\Theta$, $\sigma^2(\Theta) = (1-\Theta)/\Theta^2$. Or, il est facile de démontrer que si Θ a une distribution bêta, alors

$$E\left[\frac{(1-\Theta)^j}{\Theta^k}\right] = \frac{\Gamma(a+b)\Gamma(a-k)\Gamma(b+j)}{\Gamma(a)\Gamma(b)\Gamma(a+b+j-k)}.$$

Par conséquent,

$$\begin{aligned} m &= \frac{b}{a-1} \\ s^2 &= \frac{b(a+b-1)}{(a-1)(a-2)} \\ a &= \frac{b(a+b-1)}{(a-1)^2(a-2)}, \end{aligned}$$

et $K = a - 1$.

c) Nous avons $S|\Theta \sim \text{Gamma}(\tau, \Theta)$, $\Theta \sim \text{Gamma}(\alpha, \lambda)$, d'où $\mu(\Theta) = \tau/\Theta$ et $\sigma^2(\Theta) = \tau/\Theta^2$. Il n'est pas difficile de démontrer que

$$E[\Theta^k] = \begin{cases} \dfrac{\alpha(\alpha+1)\dots(\alpha+k-1)}{\lambda^k}, & k \geq 1 \\ 1, & k = 0 \\ \dfrac{\lambda^{|k|}}{(\alpha-1)(\alpha-2)\dots(\alpha-|k|)}, & k \leq -1, |k| < \alpha. \end{cases}$$

Par conséquent,

$$\begin{aligned} m &= \frac{\tau\lambda}{\alpha-1} \\ s^2 &= \frac{\tau\lambda^2}{(\alpha-1)(\alpha-2)} \\ a &= \frac{\tau^2\lambda^2}{(\alpha-1)^2(\alpha-2)}, \end{aligned}$$

et $K = (\alpha-1)/\tau$. Vous remarquerez qu'il s'agit d'une généralisation du cas exponentielle/gamma.

d) Nous avons $S|\Theta \sim$ Binomiale négative(r,Θ), $\Theta \sim$ Bêta(a,b), d'où $\mu(\Theta) = r(1-\Theta)/\Theta$ et $\sigma^2(\Theta) = r(1-\Theta)/\Theta^2$. En utilisant le résultat trouvé en b), nous obtenons

$$\begin{aligned} m &= \frac{rb}{a-1} \\ s^2 &= \frac{rb(a+b-1)}{(a-1)(a-2)} \\ a &= \frac{r^2b(a+b-1)}{(a-1)^2(a-2)}, \end{aligned}$$

et $K = (a-1)/r$. Il s'agit là de la généralisation du cas géométrique/-bêta.

e) Nous avons $S|\Theta \sim$ Normale$(5\Theta, \sigma^2)$, $\Theta \sim U(a,b)$, d'où $\mu(\Theta) = 5\Theta$ et $\sigma^2(\Theta) = \sigma^2$. Alors

$$\begin{aligned} m &= \frac{5}{2}(a+b) \\ s^2 &= \sigma^2 \\ a &= \frac{25(b-a)^2}{12}, \end{aligned}$$

et

$$K = \frac{12\sigma^2}{25(b-a)^2}.$$

f) Nous avons $S|\Theta \sim \text{Exponentielle}(\Theta)$, $\Theta^{-1} \sim \text{Bêta}(a, b)$, d'où $\mu(\Theta) = \Theta^{-1}$ et $\sigma^2(\Theta) = \Theta^{-2}$. Alors

$$\begin{aligned} m &= \frac{a}{a+b} \\ s^2 &= \frac{a(a+1)}{(a+b)(a+b+1)} \\ a &= \frac{ab}{(a+b)^2(a+b+1)} \end{aligned}$$

et

$$K = \frac{(a+b)(a+1)}{b}.$$

4.10 a) En posant égales à zéro les dérivées partielles de $E[(\mu(\Theta) - \alpha - \beta\bar{X})^2]$ par rapport à α et β, on obtient les équations normales

$$\begin{aligned} E[Y] - \alpha - \beta E[\bar{X}] &= 0 \\ E[Y\bar{X}] - \alpha E[\bar{X}] - \beta E[\bar{X}^2] &= 0. \end{aligned}$$

La solution de ce système d'équation est

$$\begin{aligned} \hat{\alpha} &= E[Y] - \hat{\beta} E[\bar{X}] \\ \hat{\beta} &= \frac{\text{Cov}(Y, \bar{X})}{\text{Var}[\bar{X}]}. \end{aligned}$$

b) Premièrement, nous avons que $\text{Cov}(\mu(\Theta), X_t) = a$. Il en découle directement que $\text{Cov}(\mu(\Theta), \bar{X}) = a$. Deuxièmement,

$$\begin{aligned} \text{Var}[\bar{X}] &= \text{Var}[E[\bar{X}|\Theta]] + E[\text{Var}[\bar{X}|\Theta]] \\ &= \text{Var}[\mu(\Theta)] + E\left[\frac{\sigma^2(\Theta)}{n}\right] \\ &= a + \frac{s^2}{n}. \end{aligned}$$

c) Dans un contexte de crédibilité, nous avons que $E[\mu(\Theta)] = E[\bar{X}] = m$, d'où

$$\beta = \frac{an}{an + s^2} = z$$

et $\alpha = (1 - z)m$. Cet exercice illustre deux choses : d'abord que la prime de Bühlmann minimise l'espérance en c), où l'on suppose déjà un poids égal donné à chaque contrat ; ensuite que la seule réelle différence entre la théorie de la crédibilité et les modèles statistiques usuels réside dans la structure de covariance.

4.11 Nous avons (en laissant tomber les indices i qui ne jouent aucun rôle ici) : $X|\Theta = \theta \sim$ Binomiale négative$(\theta, 0{,}4)$ et $\mu(\Theta) \sim$ Exponentielle(5). Par conséquent,

$$\begin{aligned}\mu(\Theta) &= \Theta\left(\frac{1-0{,}4}{0{,}4}\right)\\ &= 1{,}5\,\Theta\end{aligned}$$

et

$$\begin{aligned}\sigma^2(\Theta) &= \Theta\left(\frac{1-0{,}4}{0{,}4^2}\right)\\ &= 3{,}75\,\Theta\\ &= 2{,}5\mu(\Theta).\end{aligned}$$

Il est simple de trouver la distribution de Θ. En effet, $\mu(\Theta) = 1{,}5\Theta \sim$ Exponentielle(5), d'où $\Theta \sim$ Exponentielle$(7{,}5)$. Cela dit, sans même cette information, nous obtenons $s^2 = E[\sigma^2(\Theta)] = 2{,}5E[\mu(\Theta)] = 0{,}5$ et $a = \text{Var}[\mu(\Theta)] = 0{,}04$, d'où $K = 25/2$ et $z = 1/(1+25/2) = 0{,}0741$.

4.12 Soit N le nombre de sinistres. Nous avons $N|\Theta = \theta \sim$ Poisson(Θ), $E[N|\Theta] = \Theta \sim U(0{,}1, 0{,}3)$ et donc

$$\begin{aligned}m &= E[\mu(\Theta)] = E[\Theta] = 0{,}2\\ s^2 &= E[\sigma^2(\Theta)] = E[\Theta] = 0{,}2\\ a &= \text{Var}[\mu(\Theta)] = \text{Var}[\Theta] = \frac{1}{300}.\end{aligned}$$

Par conséquent, $K = s^2/a = 60$ et

$$\pi_6 = \left(\frac{5}{5+60}\right)\left(\frac{3}{5}\right) + \left(\frac{60}{5+60}\right)(0{,}2) = 0{,}2308.$$

4.13 Nous avons $X|\Theta \sim N(\Theta, 4\Theta^2)$ avec $\mu(\Theta) = \Theta$ et $\sigma^2(\Theta) = 4\Theta^2$, d'où $\sigma^2(\Theta)/\mu(\Theta) = 4\Theta \sim U(0, 40)$ et donc $\Theta \sim U(0, 10)$. Ainsi, $m = E[\Theta] = 5$, $s^2 = E[4\Theta^2] = 4(\text{Var}[\Theta] + E[\Theta]^2) = 1{,}600/12$ et $a = \text{Var}[\Theta] = 100/12$, d'où $K = s^2/a = 16$ et, finalement,

$$\pi_5 = \left(\frac{4}{4+16}\right)\left(\frac{12}{4}\right) + \left(1 - \frac{4}{4+16}\right)(5) = 4{,}6.$$

4.14 a) $S|\Theta = \theta \sim \text{Poisson}(\theta)$. Nous avons donc $\mu(\Theta) = \Theta$, $\sigma^2(\Theta) = \Theta$,

$$m = E[\mu(\Theta)] = E[\Theta] = \frac{\alpha}{\lambda}$$
$$s^2 = E[\sigma^2(\Theta)] = E[\Theta] = \frac{\alpha}{\lambda}$$
$$a = \text{Var}[\mu(\Theta)] = \text{Var}[\Theta] = \frac{\alpha}{\lambda^2}$$

et $z = n/(n+\lambda)$. Dans tous les cas ci-dessous, $m = 2$ et $\bar{S} = (3+1+5+4+2)/5 = 3$.

i) Si $\Theta \sim \text{Gamma}(10, 5)$, alors $z = 1/2$ et $\pi_6 = 2{,}5$.

ii) Si $\Theta \sim \text{Gamma}(50, 25)$, alors $z = 1/6$ et $\pi_6 = 2{,}17$.

iii) Si $\Theta \sim \text{Gamma}(\frac{1}{2}, \frac{1}{4})$, alors $z = 20/21$ et $\pi_6 = 2{,}95$.

b) $S|\Theta = \theta \sim U(0, 2\theta)$. Nous avons donc $\mu(\Theta) = \Theta$, $\sigma^2(\Theta) = \Theta^2/3$,

$$m = E[\mu(\Theta)] = E[\Theta] = \frac{\alpha}{\lambda}$$
$$s^2 = E[\sigma^2(\Theta)] = \frac{E[\Theta^2]}{3} = \frac{\alpha(\alpha+1)}{3\lambda^2}$$
$$a = \text{Var}[\mu(\Theta)] = \text{Var}[\Theta] = \frac{\alpha}{\lambda^2}$$

et $z = (3n)/(3n+\alpha+1)$. Dans tous les cas ci-dessous, $m = 2$ et $\bar{S} = (3+1+5+4+2)/5 = 3$.

i) Si $\Theta \sim \text{Gamma}(10, 5)$, alors $z = 15/26$ et $\pi_6 = 2{,}58$.

ii) Si $\Theta \sim \text{Gamma}(50, 25)$, alors $z = 15/66$ et $\pi_6 = 2{,}23$.

iii) Si $\Theta \sim \text{Gamma}(\frac{1}{2}, \frac{1}{4})$, alors $z = 30/33$ et $\pi_6 = 2{,}91$.

4.15 Nous avons

$$\begin{aligned} s^2 &= E[\sigma^2(\Theta)] \\ &= E[\Theta^2] \\ &= 4\,500 + 150^2 \\ &= 27\,000 \end{aligned}$$

et

$$\begin{aligned} a &= \text{Var}[\mu(\Theta)] \\ &= \text{Var}[\Theta] \\ &= 4\,500, \end{aligned}$$

d'où $z = 4/(4 + 27\,000/4\,500) = 0{,}4$ et $\pi_5 = (0{,}4)(100 + 125 + 75 + 120)/4 + (0{,}6)(150) = 132$.

4.16 Nous avons $\mathrm{Var}[S] = \mathrm{Var}[\mu(\Theta)] + E[\sigma^2(\Theta)] = a + s^2$. Pour le portefeuille A, $a = 0{,}25\,\mathrm{Var}[S]$, d'où $s^2 = 0{,}75\,\mathrm{Var}[S]$ et

$$K_A = \frac{0{,}75}{0{,}25} = 3.$$

Pour le portefeuille B, $S \sim N(\mu, 25)$ et $\mu(\Theta) \sim N(\alpha, 16)$. Par conséquent,

$$\begin{aligned} a &= \mathrm{Var}[\mu(\Theta)] = 16 \\ s^2 &= \mathrm{Var}[S] - \mathrm{Var}[\mu(\Theta)] = 25 - 16 = 9 \end{aligned}$$

et

$$K_B = \frac{9}{16}.$$

Puisque $K_B < K_A$, une plus grande crédibilité est accordée au portefeuille B.

4.17 a) Nous avons $\mu(\Theta) = \Theta$ et $\sigma^2(\Theta) = \Theta$. Par conséquent,

$$\Pr[\mu(\Theta) > 5] = \Pr[\Theta > 5] = e^{-5\lambda} = e^{-10},$$

d'où $\lambda = 2$. Nous savons (voir l'annexe B) que dans le modèle Poisson/gamma, la constante de crédibilité est $K = \lambda = 2$.

b) Nous avons $\mu(\Theta_1, \Theta_2) = E[S_1 + S_2|\Theta_1, \Theta_2] = E[S_1|\Theta_1] + E[S_2|\Theta_2] = \Theta_1 + \Theta_2$ et $\sigma^2(\Theta_1, \Theta_2) = \mathrm{Var}[S_1|\Theta_1] + \mathrm{Var}[S_2|\Theta_2] = \Theta_1 + \Theta_2$. Par conséquent,

$$\begin{aligned} s^2 &= E[\sigma^2(\Theta_1, \Theta_2)] \\ &= E[\Theta_1] + E[\Theta_2] \\ &= \frac{1}{\alpha} + \frac{1}{\lambda} \end{aligned}$$

et

$$\begin{aligned} a &= \mathrm{Var}[\mu(\Theta_1, \Theta_2)] \\ &= \mathrm{Var}[\Theta_1] + \mathrm{Var}[\Theta_2] \\ &= \frac{1}{\alpha^2} + \frac{1}{\lambda^2}, \end{aligned}$$

d'où

$$K = \frac{\alpha\lambda(\alpha + \lambda)}{\alpha^2 + \lambda^2}.$$

4.18 On nous donne les informations suivantes : $\pi_{n+1} = 110$, $\pi_{n+2} = 125$, $\bar{S}_n = 125$, $\bar{S}_{n+1} = 150$ et $\mathrm{Var}[\mu(\Theta)] = 0{,}25\ \mathrm{Var}[S]$. Tout d'abord, nous avons $E[\sigma^2(\Theta)] = 0{,}75\ \mathrm{Var}[S]$, d'où la constante de crédibilité est $K = 3$. Ainsi,

$$\begin{aligned} 110 &= z_n \bar{S}_n + (1 - z_n) m \\ &= \left(\frac{n}{n+3}\right)(125) + \left(\frac{3}{n+3}\right) m \end{aligned}$$

et

$$\begin{aligned} 125 &= z_{n+1} \bar{S}_{n+1} + (1 - z_{n+1}) m \\ &= \left(\frac{n+1}{n+1+3}\right)(150) + \left(\frac{3}{n+1+3}\right) m. \end{aligned}$$

En résolvant le système d'équations on trouve $n = 2$ et $m = 100$. Par conséquent, la prime de crédibilité d'un assuré ayant encouru 500 \$ de sinistres en $n = 2$ années est

$$\begin{aligned} \pi_3 &= \frac{2}{5}\left(\frac{500}{2}\right) + \frac{3}{5}(100) \\ &= 160. \end{aligned}$$

4.19 Nous avons $\mu(\Theta) \sim \mathrm{Gamma}(6, 4)$ et $\sigma^2(\Theta) = \mu(\Theta)^2$, d'où

$$\begin{aligned} m &= E[\mu(\Theta)] = 1{,}5 \\ s^2 &= E[\sigma^2(\Theta)] \\ &= E[\mu(\Theta)^2] = \frac{21}{8} \\ a &= \mathrm{Var}[\mu(\Theta)] = \frac{3}{8} \end{aligned}$$

et

$$K = \frac{s^2}{a} = 7.$$

Ainsi,

$$\begin{aligned} \pi_{n+1} &= \frac{n}{n+7}\bar{S} + \frac{7}{n+7}(1{,}5) \\ &= \frac{\sum_{t=1}^{n} S_t + 10{,}5}{n+7} \end{aligned}$$

et donc

$$\begin{aligned} a_{2,3} &= 1{,}50 & a_{3,5} &= 1{,}55 \\ a_{2,2} &= 1{,}39 & a_{4,4} &= 1{,}32 \\ a_{2,4} &= 1{,}61 & a_{3,6} &= 1{,}65. \end{aligned}$$

Les énoncés I et III sont donc vrais.

4.20 On reconnait dans $u(\theta)$ la fonction de densité de probabilité d'une loi gamma de paramètres $\alpha = 3$ et $\lambda = 4$ (d'où $c = 4^3/\Gamma(3) = 32$). Nous avons donc un modèle Poisson/gamma pour lequel toutes les formules pertinentes se trouvent à l'annexe B.

a) Nous avons $m = \alpha/\lambda = 0{,}75$.

b) La prime de crédibilité est égale à la prime bayésienne, soit

$$\begin{aligned} \pi_{n+1} &= \frac{\alpha + \sum_{t=1}^{n} S_t}{n + \lambda} \\ &= \frac{3 + \sum_{t=1}^{n} S_t}{n + 4}. \end{aligned}$$

Avec $S_1 = 4$, $S_2 = 1$ et $S_3 = 3$, nous avons donc $\pi_2 = 7/5 = 1{,}4$, $\pi_3 = 8/6 = 1{,}33$ et $\pi_4 = 11/7 = 1{,}57$.

4.21 Nous avons $S|\Theta \sim \text{Exponentielle}(\Theta)$ et $\mu(\Theta) = \Theta^{-1} \sim \text{Gamma}(\alpha, \lambda)$, de telle sorte que $\alpha/\lambda = 4$ et $\alpha/\lambda^2 = 8$. Or,

$$\begin{aligned} m &= E[\Theta^{-1}] = \frac{\alpha}{\lambda} = 4 \\ s^2 &= E[\Theta^{-2}] = \text{Var}[\Theta^{-1}] + E[\Theta^{-1}]^2 = \frac{\alpha}{\lambda^2} + \left(\frac{\alpha}{\lambda}\right)^2 = 24 \end{aligned}$$

et

$$a = \text{Var}[\Theta^{-1}] = 8$$

donnent $K = s^2/a = 3$, d'où le facteur de crédibilité après 2 ans est $z = 2/5 = 0{,}4$. La prime de Bühlmann est donc $0{,}4[(1+3)/2] + 0{,}6(4) = 3{,}2$.

4.22 Premièrement,

$$\pi_n = \frac{1}{n-1+K} \sum_{t=1}^{n-1} S_t + \frac{K}{n-1+K} m.$$

Alors,

$$\begin{aligned}
\pi_{n+1} &= \frac{1}{n+K}\sum_{t=1}^{n} S_t + \frac{K}{n+K}\,m \\
&= \frac{1}{n+K}\left(\sum_{t=1}^{n-1} S_t + S_n\right) + \frac{K}{n+K}\,m \\
&= \frac{n-1+K}{n+K}\left(\frac{1}{n-1+K}\sum_{t=1}^{n-1} S_t + \frac{K}{n-1+K}\,m\right) + \frac{1}{n+K} S_n \\
&= \zeta_n \pi_n + (1-\zeta_n) S_n,
\end{aligned}$$

avec $\zeta_n = (n-1+K)/(n+K)$.

4.23 Soit

$$\hat{s}^2 = \frac{1}{I(n-1)}\sum_{i=1}^{I}\sum_{t=1}^{n}(S_{it}-\bar{S}_i)^2.$$

Nous voulons démontrer que $E[\hat{s}^2] = s^2$ sans conditionner sur Θ_i. Le problème se réduit au calcul de $E[(S_{it}-\bar{S}_i)^2]$. Puisque $E[S_{it}] = E[\bar{S}_i]$, nous avons

$$E[(S_{it}-\bar{S}_i)^2] = \mathrm{Var}[S_{it}] + \mathrm{Var}[\bar{S}_i] - 2\,\mathrm{Cov}(S_{it},\bar{S}_i).$$

Or, nous savons que, par hypothèse dans le modèle de Bühlmann,

$$\mathrm{Cov}(S_{it}, S_{iu}) = a + \delta_{tu} s^2.$$

Par conséquent, nous avons

$$\begin{aligned}
\mathrm{Cov}(S_{it},\bar{S}_i) &= \frac{1}{n}\sum_{u=1}^{n}\mathrm{Cov}(S_{it},S_{iu}) \\
&= a + \frac{s^2}{n} \\
\mathrm{Var}[\bar{S}_i] &= \mathrm{Cov}(\bar{S}_i,\bar{S}_i) \\
&= \frac{1}{n}\sum_{t=1}^{n}\mathrm{Cov}(S_{it},\bar{S}_i) \\
&= a + \frac{s^2}{n}
\end{aligned}$$

et

$$\begin{aligned}
\mathrm{Var}[S_{it}] &= \mathrm{Cov}(S_{it},S_{it}) \\
&= a + s^2.
\end{aligned}$$

Ainsi

$$E[(S_{it} - \bar{S}_i)^2] = \frac{n-1}{n}\, s^2,$$

d'où $E[\hat{s}^2] = s^2$.

4.24 Voici des pistes de solutions. Afin de respecter la seconde exigence, nous pourrions poser

$$S_{it} = C_{it1} + \cdots + C_{itN_{it}},$$

où C_{itu} est le montant individuel du sinistre u de la période t du contrat i et N_{it} est le nombre total de sinistres du contrat i dans la période t. Il s'agit, par la suite, de déterminer des modèles de simulation pour C_{itu} et N_{it} de telle sorte que $E[S_{it}|\Theta_i] = \mu(\Theta_i)$ et $\mathrm{Var}[S_{it}|\Theta_i] = \sigma^2(\Theta_i)$. Comme il est raisonnable de supposer que les contrats se distinguent davantage par leur fréquence de sinistres que par la sévérité de ceux-ci, le modèle pour la simulation de N_{it} dépendra de la variable aléatoire Θ_i. Autrement dit, nous aurons un modèle de simulation pour $N_{it}|\Theta_i = \theta_i$ et un autre pour Θ_i.

4.25 Nous avons

$$\begin{aligned}
\hat{s}^2 &= \frac{1}{I(n-1)} \sum_{i=1}^{I} \sum_{j=1}^{n} (S_{ij} - \bar{S}_i)^2 \\
&= \frac{1}{(3)(5)}[(1+1+1+1) + (1+1+4+1+1) + (1+1+1+1)] \\
&= \frac{16}{15} = 1{,}0667 \\
\hat{a} &= \frac{1}{I-1} \sum_{i=1}^{I} (\bar{S}_i - \bar{S})^2 - \frac{\hat{s}^2}{n} \\
&= \frac{1}{2}[1+1] - \frac{16/15}{6} \\
&= \frac{74}{90} = 0{,}8222,
\end{aligned}$$

d'où $K = 48/37$ et $\hat{z} = 0{,}8222$. Par conséquent,

$$\begin{aligned}
\pi_{(1,7)} &= (0{,}8222)(1) + (0{,}1778)(2) = 1{,}1778 \\
\pi_{(2,7)} &= (0{,}8222)(3) + (0{,}1778)(2) = 2{,}8222 \\
\pi_{(3,7)} &= 2.
\end{aligned}$$

Nous pouvons obtenir les mêmes résultats avec la fonction `cm`.

```
> x <- data.frame(contract = 1:3,
+                 matrix(c(0, 1, 2, 1, 2, 0,
+                          3, 4, 2, 1, 4, 4,
+                          3, 3, 2, 1, 2, 1),
+                        nrow = 3, byrow = TRUE))
> fit <- cm(~ contract, x)
> summary(fit)
Call:
cm(formula = ~contract, data = x)

Structure Parameters Estimators

  Collective premium: 2

  Between contract variance: 0.8222222
  Within contract variance: 1.066667

Detailed premiums

    contract Indiv. mean Weight Cred. factor
    1        1           6      0.8222222
    2        3           6      0.8222222
    3        2           6      0.8222222
    Cred. premium
    1.177778
    2.822222
    2.000000
```

Chapitre 5

5.1 Du théorème 5.1, nous savons que $\mathrm{Cov}(X_{it}, X_{iw}) = a + s^2/w_{i\Sigma}$. Ensuite,

$$\begin{aligned}
\mathrm{Cov}(X_{iw}, X_{iw}) &= \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} \mathrm{Cov}(X_{it}, X_{iw}) \\
&= \sum_{t=1}^{n} \frac{w_{it}}{w_{i\Sigma}} \left(a + \frac{s^2}{w_{i\Sigma}} \right) = a + \frac{s^2}{w_{i\Sigma}} = \frac{a}{z_i}.
\end{aligned}$$

Par indépendance des contrats, $\mathrm{Cov}(X_{iw}, X_{kw}) = 0$, $k \neq i$.

a) Ici,

$$\begin{aligned}
\mathrm{Cov}(X_{iw}, X_{ww}) &= \sum_{k=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} \mathrm{Cov}(X_{iw}, X_{kw}) \\
&= \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} \mathrm{Cov}(X_{iw}, X_{iw}) = a \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \frac{s^2}{w_{\Sigma\Sigma}}.
\end{aligned}$$

b) En utilisant le résultat obtenu en a),

$$\begin{aligned}
\mathrm{Var}[X_{ww}] &= \mathrm{Cov}(X_{ww}, X_{ww}) \\
&= \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} \mathrm{Cov}(X_{iw}, X_{ww}) \\
&= \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} \left(a \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \frac{s^2}{w_{\Sigma\Sigma}} \right) \\
&= a \sum_{i=1}^{I} \left(\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} \right)^2 + \frac{s^2}{w_{\Sigma\Sigma}}.
\end{aligned}$$

c) Premièrement,

$$\begin{aligned}
\mathrm{Cov}(X_{iw}, X_{zw}) &= \sum_{k=1}^{I} \frac{z_k}{z_\Sigma} \mathrm{Cov}(X_{iw}, X_{kw}) \\
&= \frac{z_i}{z_\Sigma} \mathrm{Cov}(X_{iw}, X_{iw}) = \frac{a}{z_\Sigma}.
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
\mathrm{Var}[X_{zw}] &= \mathrm{Cov}(X_{zw}, X_{zw}) \\
&= \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} \mathrm{Cov}(X_{iw}, X_{zw}) \\
&= \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} \frac{a}{z_\Sigma} = \frac{a}{z_\Sigma}.
\end{aligned}$$

5.2 Calculons d'abord l'espérance de la variable aléatoire $(X_{it} - X_{iw})^2$. En utilisant les résultats de l'exercice 5.1, on a que

$$\begin{aligned}
E[(X_{it} - X_{iw})^2] &= \mathrm{Var}[X_{it}] + \mathrm{Var}[X_{iw}] - 2\,\mathrm{Cov}(X_{it}, X_{iw}) \\
&= a + \frac{s^2}{w_{it}} + a + \frac{s^2}{w_{i\Sigma}} - 2\left(a + \frac{s^2}{w_{i\Sigma}}\right) \\
&= \frac{s^2}{w_{it}} - \frac{s^2}{w_{i\Sigma}}.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
E\left[\sum_{i=1}^{I}\sum_{t=1}^{n} w_{it}(X_{it} - X_{iw})^2\right] &= \sum_{i=1}^{I}\sum_{t=1}^{n} w_{it} E[(X_{it} - X_{iw})^2] \\
&= (In - I)s^2,
\end{aligned}$$

d'où $E[\hat{s}^2] = s^2$.

5.3 En premier lieu,

$$\begin{aligned}
E[(X_{iw} - X_{ww})^2] &= \mathrm{Var}[X_{iw}] + \mathrm{Var}[X_{ww}] - 2\,\mathrm{Cov}(X_{iw}, X_{ww}) \\
&= a + \frac{s^2}{w_{i\Sigma}} + a\sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}}\right)^2 + \frac{s^2}{w_{\Sigma\Sigma}} - 2\left(a\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \frac{s^2}{w_{\Sigma\Sigma}}\right) \\
&= a\left(1 - 2\,\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} + \sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}}\right)^2\right) + s^2\left(\frac{1}{w_{i\Sigma}} - \frac{1}{w_{\Sigma\Sigma}}\right).
\end{aligned}$$

Par conséquent,

$$\begin{aligned}
E\left[\sum_{i=1}^{I} w_{i\Sigma}(X_{iw} - X_{ww})^2\right] &= a\sum_{i=1}^{I}\left(w_{i\Sigma} - 2\,\frac{w_{i\Sigma}^2}{w_{\Sigma\Sigma}} + w_{i\Sigma}\sum_{i=1}^{I}\left(\frac{w_{i\Sigma}}{w_{\Sigma\Sigma}}\right)^2\right) \\
&\quad + (I-1)s^2 \\
&= a\left(w_{\Sigma\Sigma} - \frac{\sum_{i=1}^{I} w_{i\Sigma}^2}{w_{\Sigma\Sigma}}\right) + (I-1)s^2 \\
&= a\left(\frac{w_{\Sigma\Sigma}^2 - \sum_{i=1}^{I} w_{i\Sigma}^2}{w_{\Sigma\Sigma}}\right) + (I-1)s^2.
\end{aligned}$$

Puisque $E[\hat{s}^2] = s^2$,

$$\hat{a} = \frac{w_{\Sigma\Sigma}}{w_{\Sigma\Sigma}^2 - \sum_{i=1}^{I} w_{i\Sigma}^2}\left(\sum_{i=1}^{I} w_{i\Sigma}(X_{iw} - X_{ww})^2 - (I-1)\hat{s}^2\right)$$

est un estimateur sans biais du paramètre a.

5.4 On a

$$\begin{aligned}
E[(X_{iw} - X_{zw})^2] &= \mathrm{Var}[X_{iw}] + \mathrm{Var}[X_{zw}] - 2\,\mathrm{Cov}(X_{iw}, X_{zw}) \\
&= \frac{a}{z_i} + \frac{a}{z_\Sigma} - 2\frac{a}{z_\Sigma} \\
&= \frac{a}{z_i} - \frac{a}{z_\Sigma}
\end{aligned}$$

et, donc,

$$\begin{aligned}
E\left[\sum_{i=1}^{I} z_i (X_{iw} - X_{zw})^2\right] &= \sum_{i=1}^{I} z_i E[(X_{iw} - X_{zw})^2] \\
&= (I-1)a,
\end{aligned}$$

d'où $E[\tilde{a}] = a$.

5.5 Si $w_{it} = w$ pour tout i et t, alors $X_{iw} = \bar{X}_i$, $X_{zw} = \bar{X}$ et $z_i = z = an/(an + s^2)$. Par conséquent, $\tilde{a}$ est la solution de

$$a = \frac{1}{I-1} \sum_{i=1}^{I} \frac{an}{an + s^2} (\bar{X}_i - \bar{X})^2$$

ou

$$an + s^2 = \frac{n}{I-1} \sum_{i=1}^{I} (\bar{X}_i - \bar{X})^2$$

ou encore

$$a = \frac{1}{I-1} \sum_{i=1}^{I} (\bar{X}_i - \bar{X})^2 - \frac{s^2}{n}.$$

Ceci démontre que $\tilde{a} = \hat{a}$ dans le modèle de Bühlmann.

5.6 Puisque $\lim_{a \to 0} z_i = 0$, nous avons

$$\begin{aligned}
\lim_{a \to 0} X_{zw} &= \lim_{a \to 0} \sum_{i=1}^{I} \frac{z_i}{z_\Sigma} X_{iw} \\
&= \left[\frac{0}{0}\right],
\end{aligned}$$

une indétermination. Or, puisque

$$\begin{aligned}\frac{\partial}{\partial a} z_i &= \frac{\partial}{\partial a} \frac{a w_{i\Sigma}}{a w_{i\Sigma} + s^2} \\ &= \frac{w_{i\Sigma}(a w_{i\Sigma} + s^2) - a w_{i\Sigma}^2}{(a w_{i\Sigma} + s^2)^2},\end{aligned}$$

alors, par une simple application de la règle de l'Hôpital,

$$\begin{aligned}\lim_{a \to 0} X_{zw} &= \sum_{i=1}^{I} \frac{s^2 w_{i\Sigma}}{\sum_{i=1}^{I} s^2 w_{i\Sigma}} X_{iw} \\ &= \sum_{i=1}^{I} \frac{w_{i\Sigma}}{w_{\Sigma\Sigma}} X_{iw} \\ &= X_{ww}.\end{aligned}$$

5.7 La façon la plus simple de simuler des ratios consiste à simuler des montants totaux de sinistres S_{it} et à poser

$$X_{it} = \frac{S_{it}}{w_{it}}.$$

Pour la simulation des montants de sinistres, se reporter à la solution de l'exercice 4.24. Il faut toutefois ajouter les poids au modèle de simulation. Puisque la taille des contrats a une influence sur leur nombre de sinistres et non sur le montant de ceux-ci, on intégrera les poids w_{it} au modèle de simulation de $N_{it}|\Theta_i = \theta_i$.

5.8 Il y a un piège dans cette question : on vous donne des montants totaux de sinistres, aussi faut-il convertir ces données en ratios pour pouvoir utiliser le modèle de Bühlmann–Straub. Les tableaux ci-dessous contiennent les données du portefeuille avec les ratios $X_{it} = S_{it}/w_{it}$.

X_{it}

	Année			
Employeur	1	2	3	4
1	7	7	3	6
2	1	0	4	3
3	1	0	1	2

w_{it}

	Année			
Employeur	1	2	3	4
1	2	3	4	3
2	4	2	1	2
3	3	3	1	3

Les moyennes pondérées sont :

$$\begin{aligned}
w_{1\Sigma} &= 12 & X_{1w} &= 5{,}4167 \\
w_{2\Sigma} &= 9 & X_{2w} &= 1{,}5556 \\
w_{3\Sigma} &= 10 & X_{3w} &= 1
\end{aligned}$$

et

$$w_{\Sigma\Sigma} = 31 \qquad X_{ww} = 2{,}8710.$$

Par conséquent,

$$\begin{aligned}
\hat{s}^2 &= \frac{1}{12-3}(36{,}9167 + 16{,}2222 + 6) \\
&= 6{,}5710 \\
\hat{a} &= \frac{31}{961-325}(128{,}3460 - (2)(6{,}5710)) \\
&= 5{,}6153
\end{aligned}$$

et $\hat{K} = \hat{s}^2/\hat{a} = 1{,}1700$. Ainsi,

$$\begin{aligned}
z_1 &= \frac{12}{13{,}1700} = 0{,}9112 \\
z_2 &= \frac{9}{10{,}1700} = 0{,}8850 \\
z_3 &= \frac{10}{11{,}1700} = 0{,}8953
\end{aligned}$$

et $X_{zw} = 2{,}6779$. Enfin,

$$\begin{aligned}
\hat{X}_{1,5} &= z_1 X_{1w} + (1 - z_1) X_{zw} \\
&= 5{,}1735
\end{aligned}$$

et pour trois unités de masse salariale, la prime totale payée par l'employeur est 15,52.

Nous pouvons obtenir les résultats ci-dessus (à l'exception de la prime totale) avec la fonction `cm` du paquetage **actuar**.

```
> X <- cbind(employeur = 1:3,
+   matrix(c(7, 7, 3, 6,
+            1, 0, 4, 3,
+            1, 0, 1, 2),
+          nrow = 3, byrow = TRUE, dimnames =
+            list(NULL,
+                 paste("ratio", 1:4, sep = "."))),
+   matrix(c(2, 3, 4, 3,
+            4, 2, 1, 2,
+            3, 3, 1, 3),
+          nrow = 3, byrow = TRUE, dimnames =
+            list(NULL,
+                 paste("poids", 1:4, sep = "."))))
> summary(cm(~ employeur, X,
+            ratios = ratio.1:ratio.4,
+            weights = poids.1:poids.4,
+            method = "Buhlmann-Gisler"))
Call:
cm(formula = ~employeur, data = X, ratios = ratio.1:ratio.4,
    weights = poids.1:poids.4, method = "Buhlmann-Gisler")

Structure Parameters Estimators

  Collective premium: 2.677935

  Between employeur variance: 5.615241
  Within employeur variance: 6.570988

Detailed premiums

    employeur Indiv. mean Weight Cred. factor
    1         5.416667    12     0.9111475
    2         1.555556     9     0.8849378
    3         1.000000    10     0.8952387
    Cred. premium
    5.173323
    1.684699
    1.175783
```

5.9 a) Il faut utiliser la formule de l'estimateur $\hat{s}^2$ pour données manquantes (5.9). Tout d'abord, nous avons

$$\begin{aligned} w_{1\Sigma} &= 3 & X_{1w} &= 4 \\ w_{2\Sigma} &= 10 & X_{2w} &= 8 \\ w_{3\Sigma} &= 12 & X_{3w} &= \frac{11}{4} \end{aligned}$$

et

$$w_{\Sigma\Sigma} = 25 \qquad X_{ww} = 5.$$

Par conséquent,

$$\begin{aligned} \hat{s}^2 &= \frac{1}{12-3}\bigg[(1)(1+1+0) + (2)(4+36+16) \\ &\quad + (3)\left(\frac{9}{16} + \frac{121}{16} + \frac{1}{16} + \frac{169}{16}\right)\bigg] \\ &= 18{,}9167, \\ \hat{a} &= \frac{25}{625-253}\left[(3)(1) + (10)(9) + (12)\left(\frac{81}{16}\right) - (2)(18{,}9167)\right] \\ &= 7{,}7901 \end{aligned}$$

et $\hat{K} = \hat{s}^2/\hat{a} = 2{,}4283$. Ainsi,

$$\begin{aligned} \hat{z}_1 &= \frac{3}{5{,}4283} = 0{,}5527 \\ \hat{z}_2 &= \frac{10}{12{,}4283} = 0{,}8046 \\ \hat{z}_3 &= \frac{12}{14{,}4283} = 0{,}8317 \end{aligned}$$

et $X_{zw} = 4{,}9954$. Enfin,

$$\begin{aligned} \hat{\pi}_{1,6} &= \hat{z}_1 X_{1w} + (1-\hat{z}_1)X_{zw} = 4{,}4453 \\ \hat{\pi}_{2,6} &= \hat{z}_2 X_{2w} + (1-\hat{z}_2)X_{zw} = 7{,}4129 \\ \hat{\pi}_{3,6} &= \hat{z}_3 X_{3w} + (1-\hat{z}_3)X_{zw} = 3{,}1279. \end{aligned}$$

b) La fonction `cm` permet de choisir entre l'estimateur sans biais $\hat{a}$ et l'estimateur itératif $\tilde{a}$. Il faut d'abord placer les données dans une matrice ou un *data frame* avec une colonne d'étiquettes pour les contrats :

```
> db <- cbind(contrat = 1:3,
+             matrix(c(3, 6, NA,  5, 8, 2, NA, 8, 0,
+                      NA, 14, 3, 4, 4, 6),
+                    nrow = 3, dimnames = list(NULL,
+                      paste("r", 1:5, sep = "."))),
+             matrix(c(1, 2, NA, 1:3, NA, 2, 3,
+                      NA, 2, 3, 1:3),
+                    nrow = 3, dimnames = list(NULL,
+                      paste("w", 1:5, sep = "."))))
> db
     contrat r.1 r.2 r.3 r.4 r.5 w.1 w.2 w.3 w.4
[1,]       1   3   5  NA  NA   4   1   1  NA  NA
[2,]       2   6   8   8  14   4   2   2   2   2
[3,]       3  NA   2   0   3   6  NA   3   3   3
     w.5
[1,]   1
[2,]   2
[3,]   3
```

Tout d'abord, vérifions les réponses de la partie a).

```
> summary(cm(~contrat, db,
+            ratios = r.1:r.5,
+            weights = w.1:w.5,
+            method = "Buhlmann-Gisler"))
Call:
cm(formula = ~contrat, data = db, ratios = r.1:r.5, weights = w.1:w.5,
    method = "Buhlmann-Gisler")

Structure Parameters Estimators

  Collective premium: 4.99537

  Between contrat variance: 7.790099
  Within contrat variance: 18.91667

Detailed premiums

    contrat Indiv. mean Weight Cred. factor
    1       4.00        3      0.5526596
```

```
    2       8.00        10      0.8046155
    3       2.75        12      0.8316990
    Cred. premium
    4.445269
    7.412942
    3.127898
```

Pour utiliser plutôt l'estimateur $\tilde{a}$ du paramètre a, l'argument `method` doit être `"iterative"`.

```
> summary(cm(~contrat, db,
+              ratios = r.1:r.5,
+              weights = w.1:w.5,
+              method = "iterative"))
Call:
cm(formula = ~contrat, data = db, ratios = r.1:r.5, weights = w.1:w.5,
    method = "iterative")

Structure Parameters Estimators

  Collective premium: 5.008418

  Between contrat variance: 5.565321
  Within contrat variance: 18.91667

Detailed premiums

    contrat Indiv. mean Weight Cred. factor
    1       4.00         3     0.4688214
    2       8.00        10     0.7463229
    3       2.75        12     0.7792701
    Cred. premium
    4.535650
    7.241104
    3.248500
```

Évidemment, les réponses sont alors différentes de celles obtenues en a).

5.10 On a $z_1 = 8/23$, $z_2 = 6/21$, $z_4 = 9/24$ et

$$\begin{aligned}
X_{zw} &= \sum_{i=1}^{4} \frac{z_i}{z_\Sigma} X_{iw} \\
&= \frac{\frac{8}{23}(0{,}89) + \frac{6}{21}(0{,}85) + \frac{9}{24}(0{,}7) + z_3(0{,}65)}{\frac{8}{23} + \frac{6}{21} + \frac{9}{24} + z_3} \\
&= 0{,}8.
\end{aligned}$$

Par conséquent, $z_3 = 0{,}0539$.

Bibliographie

Bailey, A. L. 1945, «A generalized theory of credibility», *Proceedings of the Casualty Actuarial Society*, vol. 32, p. 13–20.

Bailey, A. L. 1950, «Credibility procedures, Laplace's generalization of Bayes' rule and the combination of collateral knowledge with observed data», *Proceedings of the Casualty Actuarial Society*, vol. 37, p. 7–23.

Bühlmann, H. 1967, «Experience rating and credibility», *ASTIN Bulletin*, vol. 4, p. 199–207.

Bühlmann, H. 1969, «Experience rating and credibility», *ASTIN Bulletin*, vol. 5, p. 157–165.

Bühlmann, H. et A. Gisler. 2005, *A Course in Credibility Theory and its Applications*, Springer, ISBN 3-5402575-3-5.

Bühlmann, H. et E. Straub. 1970, « Glaubgwürdigkeit für Schadensätze », *Bulletin of the Swiss Association of Actuaries*, vol. 70, p. 111–133.

Cossette, H. et V. Goulet. 2008, *Exercices en théorie de la crédibilité : Avec solutions*, 2[e] éd., Document libre sous licence Creative Commons BY-SA, ISBN 978-2-9809136-9-3. URL http://libre.act.ulaval.ca/cours/act_2008.

De Vylder, F. 1976a, «Geometrical credibility», *Scandinavian Actuarial Journal*, vol. 1976, p. 121–149.

De Vylder, F. 1976b, «Optimal semilinear credibility», *Bulletin of the Swiss Association of Actuaries*, vol. 76, p. 27–40.

De Vylder, F. 1978, «Parameter estimation in credibility theory», *ASTIN Bulletin*, vol. 10, p. 99–112.

De Vylder, F. 1981, «Practical credibility theory with emphasis on parameter estimation», *ASTIN Bulletin*, vol. 12, p. 115–131.

De Vylder, F. 1984, «Practical models in credibility theory, including parameter estimation», dans *Premium calculation in insurance*, *NATO ASI series ; Series C, Mathematical and physical sciences*, vol. 121, édité par F. De Vylder, M. Goovaerts et J. Haezendonck, D. Reidel Pub. Co., Dordrecht, p. 133–150.

De Vylder, F. 1985, «Non-linear regression in credibility theory», *Insurance : Mathematics and Economics*, vol. 4, p. 163–172.

De Vylder, F. et M. J. Goovaerts. 1991, «Estimation of the heterogeneity parameter in the Bühlmann–Straub credibility theory model», *Insurance : Mathematics and Economics*, vol. 10, p. 233–238.

De Vylder, F. et M. J. Goovaerts. 1992a, «Optimal parameter estimation under zero-excess assumptions in a classical model», *Insurance : Mathematics and Economics*, vol. 11, p. 1–6.

De Vylder, F. et M. J. Goovaerts. 1992b, «Optimal parameter estimation under zero-excess assumptions in the Bühlmann–Straub model», *Insurance : Mathematics and Economics*, vol. 11, p. 167–171.

Dubey, A. et A. Gisler. 1981, «On parameter estimation in credibility», *Bulletin of the Swiss Association of Actuaries*, vol. 81, p. 187–211.

Dutang, C., V. Goulet et M. Pigeon. 2008, «actuar: An R package for actuarial science», *Journal of Statistical Software*, vol. 25, n° 7. URL http://www.jstatsoft.org/v25/i07.

Fisher, A. 1919, «The theory of experience rating: Discussion», *Proceedings of the Casualty Actuarial Society*, vol. 5, p. 139–145.

Gerber, H. U. et D. A. Jones. 1975, «Credibility formulas of the updating type», *Transaction of the Society of Actuaries*, vol. 27, p. 31–46.

Gisler, A. 1980, «Optimal trimming of data in the credibility model», *Bulletin of the Swiss Association of Actuaries*, vol. 80, p. 313–325.

Gisler, A. et P. Reinhard. 1993, «Robust credibility», *ASTIN Bulletin*, vol. 23, p. 117–143.

Goel, P. K. 1982, «On implications of credible means being exact bayesian», *Scandinavian Actuarial Journal*, p. 41–46.

Goovaerts, M. J. et W. J. Hoogstad. 1987, *Credibility Theory*, n° 4 dans Surveys of actuarial studies, Nationale-Nederlanden N.V., Netherlands.

Goovaerts, M. J., R. Kaas, A. E. van Heerwaarden et T. Bauwelinckx. 1990, *Effective Actuarial Methods*, North-Holland, Amsterdam, ISBN 0-4448839-9-1.

Goulet, V. 1994, *Théorie de la crédibilité : histoire, principes et applications*, mémoire de maîtrise, Université Laval.

Goulet, V. 1998, «On approximations in limited fluctuation credibility», *Proceedings of the Casualty Actuarial Society*, vol. 84, p. 533–552.

Hachemeister, C. A. 1975, «Credibility for regression models with application to trend», dans *Credibility, theory and applications*, Proceedings of the Berkeley actuarial research conference on credibility, Academic Press, New York, p. 129–163.

Jewell, W. S. 1974, «Credible means are exact bayesian for exponential families», *ASTIN Bulletin*, vol. 8, p. 77–90.

Jewell, W. S. 1975, «The use of collateral data in credibility theory: A hierarchical model», *Giornale dell'Istituto Italiano degli Attuari*, vol. 38, p. 1–16.

Keffer, R. 1929, «An experience rating formula», *Transactions of the Actuarial Society of America*, vol. 30, p. 130–139.

Klugman, S. A., H. H. Panjer et G. Willmot. 2012, *Loss Models: From Data to Decisions*, 4e éd., Wiley, New York, ISBN 978-1-118-31532-3.

Künsch, H. R. 1992, «Robust methods for credibility», *ASTIN Bulletin*, vol. 22, p. 33–49.

Mayerson, A. L. 1964, «A bayesian view of credibility», *Proceedings of the Casualty Actuarial Society*, vol. 51, p. 85–104.

Mowbray, A. H. 1914, «How extensive a payroll exposure is necessary to give a dependable pure premium?», *Proceedings of the Casualty Actuarial Society*, vol. 1, p. 25–30.

Neuhaus, W. 2004, «Experience rating», dans *Encyclopedia of Actuarial Science*, vol. 2, édité par J. L. Teugels et B. Sundt, Wiley, ISBN 0-4708467-6-3, p. 639–646.

Neyman, J. 1962, «Two breakthroughs in the theory of statistical decision making», *Revue de l'Institut International de Statistique*, vol. 30, p. 11–27.

Norberg, R. 1979, «The credibility approach to ratemaking», *Scandinavian Actuarial Journal*, vol. 1979, p. 181–221.

Norberg, R. 1980, «Empirical Bayes credibility», *Scandinavian Actuarial Journal*, vol. 1980, p. 177–194.

R Core Team. 2019, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. URL `https://www.r-project.org`.

Robbins, H. 1955, «An empirical Bayes approach to statistics», dans *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*, vol. 30, University of California Press, Berkeley and Los Angeles, p. 157–163.

Robbins, H. 1964, «An empirical Bayes approach to statistics», *Annals of Mathematical Statistics*, vol. 30, p. 1–20.

Russell, B. 1948, *Human Knowledge, Its Scope and Limits*, G. Allen and Unwin, London.

Savage, L. J. 1954, *The Foundations of Statistics*, Wiley, New York.

Taylor, G. C. 1977, «Abstract credibility», *Scandinavian Actuarial Journal*, vol. 1977, p. 149–168.

Whitney, A. W. 1918, «The theory of experience rating», *Proceedings of the Casualty Actuarial Society*, vol. 4, p. 275–293.

Zehnwirth, B. 1991, « Credibility (Rough study guide) », Notes d'une conférence prononcée lors du 1991 CAS Seminar on Ratemaking.

Ce document a été produit avec le système de mise en page XƎLaTeX. Le texte principal est composé en Lucida Bright OT 11 points, les mathématiques en Lucida Bright Math OT, le code informatique en Lucida Grande Mono DK et les titres en Fira Sans. Des icônes proviennent de la police Font Awesome. Les graphiques ont été réalisés avec R.