

Institut de Mathématiques de Marseille
Institut des Sciences Financières et d'Assurance
ED 184: Ecole doctorale en Mathématiques et Informatique
de Marseille

Approximations polynomiales de densités de probabilité et applications en assurance

THESE DE DOCTORAT

présentée par

Pierre-Olivier Goffard

pour obtenir le grade de

Docteur de l'université d'Aix-Marseille

Discipline : Mathématiques appliquées

Soutenue le 29/06/2015 devant le jury :

Claude Lefèvre	Professeur à l'université libre de Bruxelles	Rapporteur
Partice Bertail	Professeur à l'université Paris X	Rapporteur
Dominique Henriot	Professeur à l'école Centrale Marseille	Examineur
Xavier Guerrault	Actuaire chez Axa France	Responsable entreprise
Denys Pommeret	Professeur à l'université d'Aix-Marseille	Directeur de thèse
Stéphane Loisel	Professeur à l'université de Lyon 1	Co-directeur de thèse

Table des matières

Remerciements	1
Introduction	1
1 Les distributions composées et la théorie de la ruine	7
1.1 Les distributions composées en dimension 1	8
1.2 Les distributions composées en dimension n	21
1.3 Théorie de la ruine : L'approximation de la probabilité de ruine ultime dans le modèle de ruine de Poisson composée	26
1.4 Références	34
2 Approximation de la densité de probabilité via une représentation polynomiale	39
2.1 Les familles exponentielles naturelles quadratiques et leurs polynômes orthogonaux	40
2.2 Représentation polynomiale de la densité en dimension 1	41
2.3 Représentation polynomiale de la densité en dimension n	50
2.4 Illustrations numériques : Application à la distribution Poisson composée	59
2.5 Perspectives en statistique	80
2.6 Références	88
3 A polynomial expansion to approximate ruin probability in the compound Poisson ruin model	93
3.1 Introduction	94
3.2 Polynomial expansions of a probability density function	97
3.3 Application to the ruin problem	99
3.4 Numerical illustrations	105
3.5 Conclusion	112
3.6 Références	112
4 Polynomial approximations for bivariate aggregate claims amount probability distributions	115
4.1 Introduction	116
4.2 Expression of the joint density	118
4.3 Application to a bivariate aggregate claims amount distribution	121
4.4 Downton Bivariate Exponential distribution and Lancaster probabilities .	125
4.5 Numerical illustrations	127

4.6	Conclusion	136
4.7	Références	137
5	Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points ?	139
5.1	Introduction	140
5.2	Cash flows projection models and best estimate liabilities assessment . . .	141
5.3	Presentation of the aggregation procedure	143
5.4	Illustration on a real life insurance portfolio	147
5.5	Conclusion	154
5.6	Références	154
	Conclusion	157
	Résumé	159
	Abstract	161

Liste des figures

1.1	Évolution des processus de richesse et d'excédent de sinistres au cours du temps.	27
2.1	Erreur relative de l'approximation polynomiale, avec différentes paramétrisations, de la densité de probabilité d'une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$. . .	63
2.2	Erreur relative de l'approximation polynomiale, avec différentes paramétrisations, de la Fonction De Survie (f.d.s.) d'une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$. . .	63
2.3	Erreur relative de l'approximation polynomiale de la f.d.s. d'une distribution $[\mathcal{P}(4), \Gamma(1, 1/2)]$ avec différentes méthodes.	66
2.4	Approximation polynomiale de la densité défaillante, de la f.d.s., de la Fonction De Répartition (f.d.r.) et de la prime <i>stop-loss</i> usuelle pour une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$	67
2.5	Erreur relative de l'approximation polynomiale, avec différentes paramétrisations, de la densité de probabilité d'une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$. . .	70
2.6	Erreur relative de l'approximation polynomiale, avec différentes paramétrisations, de la f.d.s. d'une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$	70
2.7	Erreur relative de l'approximation polynomiale de la f.d.s. d'une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$ suivant différentes méthodes numériques.	72
2.8	Approximation polynomiale de la densité défaillante, de la f.d.s., de la f.d.r. et de la prime <i>stop-loss</i> usuelle pour une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$. . .	74
2.9	Erreur relative de l'approximation polynomiale, avec différentes paramétrisations, de la f.d.s. d'une distribution $[\mathcal{P}(3), \mathcal{U}(0, 10)]$ par rapport à méthode d'inversion numérique de la transformée de Fourier.	76
2.10	Erreur relative, en %, sur la f.d.s. d'une distribution $[\mathcal{P}(2), \mathcal{U}(0, 10)]$ suivant la méthode numérique utilisée.	78
2.11	Approximation polynomiale de la densité défaillante, de la f.d.s., de la f.d.r. et de la prime <i>stop-loss</i> usuelle pour une distribution $[\mathcal{P}(3), \mathcal{U}(0, 10)]$. . .	79
3.1	Difference between exact and polynomial approximations of ruin probabilities for $\Gamma(2, 1)$ -distributed claim sizes using different parametrizations and an order of truncation $K = 40$	107
3.2	Difference between exact and approximated ruin probabilities for $\Gamma(2, 1)$ -distributed claim sizes.	108
3.3	Difference between exact and approximated ruin probabilities for $\Gamma(3, 1)$ -distributed claim sizes.	110
3.4	Relative error and ruin probabilities for $U(0, 100)$ -distributed claim sizes	111

4.1	Error map for the joint PDF of a DBVE(1/2,2,1/4) distribution approximated by polynomial expansion	128
4.2	Error map for the joint PDF of a DBVE(1/2,2,1/4) distribution approximated by the moment-recovered method	129
4.3	Error map for the joint survival function of a MOBVE(1/2,2,1) distribution approximated by polynomial expansion.	131
4.4	Error map for the joint survival function of a MOBVE(1/2,2,1) distribution approximated by the moment recovered method.	131
4.5	Discrete error map for the joint survival function of compound NEG – BIN(1,3/4) DBVE(1,1,1/4) distribution	133
4.6	Joint survival function of compound NEG – BIN(1,3/4) DBVE(1,1,1/4) distribution with an order of truncation equal to 10	133
4.7	Error map for the joint survival function of (S_1, S_2) obtained by polynomial expansion	135
4.8	Joint density and survival function of the aggregate claims in the two insurance portfolios	135
4.9	Survival function of the reinsurance cost obtained by polynomial expansion (blue) and Monte-Carlo simulations (red); Difference between the survival function obtained with Monte Carlo and the one obtained by polynomial expansion	136
5.1	Lapse rate curve depending on the seniority of contracts in a saving contract portfolio	145
5.2	Error on the BEL evaluation depending on the number of model points and the aggregation method	150
5.3	WWCI evolution depending on the number of clusters	151
5.4	Partitionning quality indicators variations depending on the number of clusters	152
5.5	Portfolio visualization through its trajectories of surrender	152
5.6	Expected present value of surrender during the projection	153

Liste des tableaux

2.1	Approximations de la f.d.s. d'une loi $(\mathcal{P}(4), \Gamma(1, 4))$ par la méthode d'approximation polynomiale	64
2.2	Erreur relative (%) sur la f.d.s. d'une loi $(\mathcal{P}(4), \Gamma(1, 2))$ associée à la méthode d'approximation polynomiale	64
2.3	Approximations de la f.d.s. d'une loi $(\mathcal{P}(4), \Gamma(1, 2))$	66
2.4	Erreur relative (%) sur la f.d.s. d'une loi $(\mathcal{P}(4), \Gamma(1, 2))$	67
2.5	Calcul de la prime <i>stop-loss</i> pour une loi $(\mathcal{P}(4), \Gamma(1, 1/2))$	68
2.6	Approximations polynomiales de la f.d.s. d'une loi $(\mathcal{P}(2), \Gamma(3, 1))$	71
2.7	Erreur relative (%) sur la f.d.s. d'une loi $(\mathcal{P}(2), \Gamma(3, 1))$ associée à la méthode d'approximation polynomiale	71
2.8	Approximations de la f.d.s. d'une loi $(\mathcal{P}(2), \Gamma(3, 1))$ via différentes méthodes	73
2.9	Erreur relative (%) liée à l'approximation de la f.d.s. d'une loi $(\mathcal{P}(2), \Gamma(3, 1))$	73
2.10	Calcul de la prime <i>stop-loss</i> pour une loi $(\mathcal{P}(2), \Gamma(3, 1))$	73
2.11	Approximations polynomiales de la f.d.s. d'une loi $(\mathcal{P}(4), \mathcal{U}(0, 10))$	76
2.12	Erreur relative (%) sur la f.d.s. d'une loi $(\mathcal{P}(4), \mathcal{U}(0, 10))$ associée à la méthode d'approximation polynomiale	77
2.13	Approximations de la f.d.s. d'une loi $(\mathcal{P}(4), \mathcal{U}(0, 10))$	77
2.14	Erreur relative (%) sur la f.d.s. d'une loi $(\mathcal{P}(4), \mathcal{U}(0, 10))$	78
2.15	Calcul de la prime <i>stop-loss</i> pour une loi $(\mathcal{P}(4), \mathcal{U}(0, 10))$	80
3.1	Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $\Gamma(2, 1)$ -distributed.	109
3.2	Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $\Gamma(3, 1)$ -distributed.	109
3.3	Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $U[0, 100]$ -distributed.	111
4.1	Survival function of the reinsurance cost obtained by polynomial expansion and Monte-Carlo simulations	136
5.1	Number of policies and amount of the initial surrender value of the portfolio	148
5.2	Statistical description of the variable AGE in the portfolio	148
5.3	Statistical description of the variable SENIORITY in the portfolio	148

5.4	Number of MP and best estimate liability of the aggregated portfolio . . .	149
5.5	Best estimate liabilities error with 28 model points depending on the aggregation method	149
5.6	Best estimate liabilities error with 3 model points depending on the aggregation method	151

Remerciements

« A Tata Annette. »

Je tiens à remercier Claude Lefevre et Patrice Bertail pour avoir accepté de rapporter cette thèse. Je remercie Dominique Henriet pour avoir accepté de faire partie du jury. J'en suis très honoré.

Je remercie bien entendu mes deux directeurs de thèse, Denys Pommeret et Stéphane Loisel. Sans vous ce travail n'aurait jamais existé. Merci pour le temps que vous avez pu me consacrer, les conseils toujours pertinents que vous m'avez prodigués et le goût de la recherche que vous avez su m'insuffler. Je garderai en mémoire les bons moments passés en marge des conférences ou encore les après-midi de travail à l'ISFA post restaurant, toujours très fructueuses...

J'adresse un grand merci à Xavier Guerrault, mon "responsable-entreprise" chez AXA, qui a su se montrer très patient avec moi. Je sais que je peux parfois être difficile à manager. J'ai beaucoup apprécié travailler avec toi, et je suis très triste à l'idée que ça s'arrête. Tu m'as fait confiance jusqu'au bout et je sais que sans toi cette thèse CIFRE n'aurait jamais vu le jour. Tu es un exemple à suivre professionnellement, peut-être as-tu raison un peu trop souvent ce qui est un peu énervant ! J'en profite pour remercier Mohammed Baccouche, pour le soutien qu'il a apporté au projet *Model Point*, et Isabelle Jubelin en charge de la gestion du fonds AXA pour la recherche qui m'a permis de terminer cette thèse dans les meilleures conditions.

J'ai une pensée pour mes parents et ma famille, qui m'ont toujours soutenu, et supporté malgré mes sauts d'humeur et mes phases de "débranchage" lors de mes retours à Lorient. Je remercie mes parents d'avoir fait le maximum pour que ma soutenance se déroule dans les meilleures conditions, je ne sais pas comment j'aurais fait sans vous. Je remercie Louis-Marie, Nathalie, Corentin, et Amélie d'avoir fait le déplacement jusque dans cette sulfureuse cité phocéenne, mes deux grands-parents dont la présence est toujours réconfortante, enfin mes deux frangins Simon et Vincent qui ont pris leur part de responsabilité dans l'organisation de cette merveilleuse soirée de soutenance !

Je remercie tous les gens que j'ai eu la chance de rencontrer au sein d'AXA. J'ai passé de très bons moments avec vous pendant les pauses café, les repas de midi, mais aussi les foots en salle, les barbeucs, les plages, les restos avec piscine et les soirées en tout

genre! Merci à Jennifer, Julie, Céline, Modou, Amin et David de l'équipe épargne individuelle. Merci à Patricio, Tom et JP de la retraite collective. Merci à Vince, Badr, Mathieu, Marilynne, Marion, Pierre-Etienne et Emmanuele dans l'équipe prévoyance individuelle. Merci à Anais, Loic, Sophie, Maud V, Maud M, Olive, Christine, et Thérèse de l'équipe des comptes IARD. Merci à Erwan qui aurait vraiment toute sa place dans cette équipe IARD. Merci à Alban, Fida, John, Laure, Anwar, Hugo, Leslie, Delphine, Corinne, et Carelle dans l'équipe RMMV. Merci à Raph et à Bénédicte de la formation, ainsi qu'au Poppy Chap des Ressources Humaines. Des remerciements spéciaux vont à Viken pour tous les BBQ que tu as organisés, ta bonne humeur, les discussions toujours passionnantes et conflictuelles (les fameux "débats-minute" des repas de midi), et l'animation générale que tu parviens à générer à AXA Marseille. Je remercie spécialement Romain Silvestri car tu m'as hébergé lors des semaines d'examen à Lyon, et incrusté aux soirées de fin d'examen, Good times!

Je veux remercier toutes les personnes rencontrées au laboratoire de mathématiques à Luminy, même si mon intégration a été difficile car je ne fais pas des "vraies" maths. Je garde énormément de très bons souvenirs des foots à Luminy, des séminaires, de la rando dans les calanques, des soirées au barberousse, des soirées chez Joël, des soirées Tonneaux-Bagel, des nombreuses parties de belotte contrée, et bien sûr l'organisation de cet immense événement qu'est le Pi Day! Merci à mes amis thésards Emilie, J-B ou Jean Ba, Kimo, Marc M, Marc B, Eugenia, Fabio, Jordi, Fra, Matteo, Lionel, Paolo I, Paolo II, Marcello, Florent, Guillaume, et Sarah l'américaine. Merci aux membres de la team stat de l'I2M Mohamed, Badhi et Laurence. Merci à Jean-Bruno, l'apple genius du labo, toujours là en cas de soucis informatique! J'adresse des remerciements spéciaux à Joël Cohen qui a pris le temps de relire ma thèse pour traquer les typos! Les discussions que l'on a pu avoir ont toujours été riches en enseignement, et montrent qu'une collaboration entre un matheux appliqué et un matheux fondamental est possible (même si dans le cas présent cela n'a marché que dans un sens, ce n'est pas comme si je pouvais aider à résoudre des problèmes dans les groupes p-adiques tordus...)! J'adresse aussi un gigantesque merci à Anna Iezzi pour toute la bonne humeur, et le dynamisme qui la caractérisent. Tu es la Chief Happiness Officer du labo. La vie sans toi y serait bien terne. Je te remercie car tu t'es occupée de moi lors de mes week-ends passer à Luminy (surtout en fin de rédaction). Tu m'as soutenu moralement et nutritivement à base de pâtes au moment où j'en avais le plus besoin. Tu bénéficies d'une infinité de Jokers.

Je remercie mes amis de Lorient, Romain, Jerem, Jo, Raph, et Gwenn bien entendu. Vous avez toujours su répondre présent lors de mes retours en Bretagne pour aller se pinter un brun dans les bars, et les boîtes de Lorient, chiller à la plage, au sonisphere et à Nantes (epic nights out). J'en viens naturellement à remercier mes amis du lycée, Julien, Awen, Lucas, Romain, Carole, Tonton, Greg, Julie, Mathieu, Solène, Laura, Manon, Sarah, Lucile, Camille, Maena, et Marjo. C'est vrai qu'il devient de plus en plus difficile de se réunir et qu'il faut attendre des événements comme le nouvel an, les EVG et les mariages. J'espère que "les vacs entre potes qui depotent et Ah que ça va être la grosse teuf" deviendront une tradition qui perdurera!

REMERCIEMENTS

Je remercie tous mes potes de l'ENSAI, Kik, Kub, Adam, Chlo, Tilde, Chris, Tiny, Virgule, Isa, Guigui, Picot, Lulu, Julie, Marie, Justine, Charlou, Gregou, Cyb, Agathe, Elise, Clem, et j'en oublie... Pour toutes les soirées sur Paris avec souvent hébergement à la clé, les vacs au ski, Solidays, les Week End à Marseille - WEM, et bien sur les Week end des Amis de l'Apéro - WAA! Un merci spécial à Boris mon frère de thèse, même si tu n'as pas pu venir à la soutenance je t'aime bien quand même. On aura réussi à vivre deux ou trois aventures durant ces trois ans et demi. Je citerais pèle-mêle le dry humming, Ange Marie, la ferme aux crocodiles et Billy.

Je remercie mes amis de Marseille, à commencer par Laurent mon ancien colocataire et partenaire du Marseille-Cassis, qui m'a introduit dans son groupe d'amis ("la gangrène"). Merci à Diane, Melanie, Magalie, Julien, et Marie. Que de bons moments passés ensemble lors des week end aux Lecques, les soirées chez Melanie Laurent, les indénombrables Brady's + Pizza, et tous les anniversaires surprise! Un merci spécial à Cédric, le quintal avec qui j'ai pu passer des soirées assez mémorables dans les bars de Marseille #TournéeDesGrandsDucs, et merci de m'avoir introduit à la team des blacks shorts de la blue coast. Je remercie donc mes amis surfeurs de Méditerranée, Cédric (dit "le gros"), Nico, Pimp, Antho, Flo, Quentin (même s'il fait du body et ne surf pas en Med), et bien sûr Ludo la maquina! Merci pour toutes ces sessions sur la côte bleue, les trips camion dans les landes et au pays Basque (voire "en el pais basco"), sans oublier les surf trips au Maroc chez notre ami Abdessamad! Je remercie pour finir les colocataires que j'ai eu la chance de me frapper lors de mes débuts à Marseille, Gilles, Lisa, Valérie, et Anne-Victoria! On était comme *un petit famille*.

En vous souhaitant une bonne lecture,

Pierre-O.

Introduction

Ce travail a été réalisé dans le cadre d'une thèse convention CIFRE, issue d'un partenariat entre la compagnie d'assurance AXA et Aix-Marseille Université. Il comprend une partie recherche opérationnelle concrétisée par un projet de recherche et développement au sein d'AXA France et une partie recherche académique visant à l'élaboration de méthodes numériques pour la gestion des risques des compagnies d'assurance.

La partie recherche opérationnelle de ce travail s'inscrit dans les préparatifs de l'entrée en vigueur de la directive européenne Solvabilité II. Cette directive définit un nouveau cadre prudentiel et oblige les compagnies d'assurance européennes à adopter une vue prospective de leurs engagements afin de bien définir leur niveau de provisionnement. Ayant été accueilli au sein d'une direction technique en charge de la gestion des contrats d'assurance vie du périmètre épargne individuelle, le projet de recherche et développement conduit au sein d'AXA France a eu pour but l'optimisation des temps de calcul des provisions *Best Estimate* associées aux contrats d'assurance vie de type épargne individuelle. L'assuré, lors de la souscription, effectue un premier versement pour constituer un capital initial. Ce capital est investi sur différents supports d'investissement dont la performance influence la valeur de l'épargne au cours du temps. L'assureur s'engage à verser à l'assuré ou à ses bénéficiaires désignés la valeur de l'épargne en cas de décès ou de rachat du contrat. Le rachat du contrat est une action volontaire de l'assuré qui décide de récupérer tout ou partie de son épargne afin de l'utiliser autrement. La provision *Best Estimate* d'un contrat épargne est égale à la moyenne des *cash-flows* sortants, actualisés et pondérés par leur probabilité d'occurrence au cours du temps. Le décès et le rachat sont les deux événements qui entraînent un *cash-flow* sortant. Ces deux événements sont respectivement associés à des probabilités de décès et de rachat fonction de l'âge de l'assuré, de son genre et aussi de l'ancienneté du contrat. Des lois de rachat et de décès sont calibrées statistiquement à l'aide d'un historique. La valeur du *cash-flow* sortant est fonction de la valeur de l'épargne au cours du temps et du rendement des actifs sur les marchés financiers. Le rendement des actifs est modélisé par l'intermédiaire de processus stochastiques. Des scénarios d'évolution de taux d'intérêts sont générés sur la base de ces processus stochastiques. La provision est évaluée pour chacun de ces scénarios, la moyenne des provisions sur les différents scénarios financiers renvoie la provision *Best Estimate*. La directive Solvabilité II préconise une approche contrat par contrat. Le temps de calcul pour un contrat est relativement faible, de l'ordre de quelques secondes, il devient problématique dans le cadre d'un portefeuille de contrats volumineux comme celui d'AXA France qui comprend 3 millions de contrats. L'autre problème réside dans la consommation de mémoire induite par le stockage des données de simulation. Les autorités

de contrôle et de réglementation sont conscientes de ces problèmes et cautionnent la mise en place de techniques pour réduire les temps de calculs sous réserve de ne pas fausser significativement la valorisation des provisions. L'objectif du projet de recherche et développement mené au cours de cette thèse est la mise au point d'une procédure visant à agréger le portefeuille de contrats d'assurance vie. Le portefeuille de contrats agrégé est alors utilisé en *input* du modèle interne de projection des *cash-flows* permettant l'évaluation des provisions *Best Estimate*.

La partie recherche académique de ce travail consiste en la mise au point d'une méthode numérique d'approximation des probabilités dans le cadre des modèles de risque en assurance non-vie. Pour un portefeuille de contrats d'assurance non vie, sur une période d'exercice donnée, le nombre de sinistres est une variable aléatoire de comptage et le montant des sinistres est une variable aléatoire positive. La charge totale du portefeuille, qui correspond au cumul des prestations à verser par l'assureur sur la période d'exercice considérée, est une quantité centrale dans la gestion des risques de la compagnie. Elle quantifie les engagements contractés par la compagnie d'assurance auprès de ses clients. Pour calibrer le montant de la prime et déterminer les marges de solvabilité, l'actuaire doit étudier la distribution du cumul des prestations et connaître son espérance, sa variance, sa fonction de survie ou encore ses quantiles. La charge totale du portefeuille admet une distribution de probabilité composée. Le problème de ces distributions de probabilité est que leur densité de probabilité n'est accessible que dans un nombre très restreint de cas. Les calculs de probabilité peuvent être effectués par le biais de méthodes de simulation de Monte-Carlo mais les temps de calcul importants associés à une bonne précision sont souvent prohibitifs et motivent la mise au point de méthodes numériques d'approximation de la densité de probabilité. Une compagnie d'assurance possède souvent plusieurs branches d'activité, chaque branche d'activité est associée à une charge totale et la gestion globale des risques implique la définition d'un vecteur aléatoire. En cas d'indépendance entre les branches d'activité, le problème se résume à considérer plusieurs problèmes en dimension 1. En revanche, s'il existe une corrélation entre les branches d'activité alors il est opportun de modéliser les charges totales conjointement via un vecteur aléatoire associé à une distribution de probabilité multivariée. Les méthodes numériques doivent alors pouvoir être étendues en dimension supérieure à 1. Cette corrélation est particulièrement pertinente dans la gestion des risques d'un réassureur qui réassure la même branche d'activité de plusieurs compagnies d'assurance. Un prolongement du modèle collectif est étudié en théorie de la ruine où un aspect dynamique est ajouté. La théorie de la ruine concerne la modélisation de la richesse globale d'une compagnie d'assurance ou la réserve financière allouée à une certaine branche d'activité ou portefeuille de contrats. La réserve financière est égale à une réserve initiale à laquelle s'ajoute les primes reçues par unité de temps moins les prestations. L'arrivée des sinistres est modélisée par un processus de comptage, le montant des prestations par des variables aléatoires positives. L'objectif est de déterminer la probabilité de ruine qui correspond à la probabilité de passage en dessous de 0 du processus en fonction de la réserve financière initiale. La complexité des variables aléatoires en jeu rend la probabilité de ruine difficile à évaluer via des formules fermées d'où la mise au point de méthodes numériques.

L'objet de ce travail est d'optimiser la précision et la rapidité des temps de calcul. La partie recherche opérationnelle traite d'une procédure à utiliser en amont du calcul alors que la partie recherche académique s'intéresse à la façon même de réaliser les calculs.

Le Chapitre 1 présente les distributions composées et leur utilité en actuariat, et en théorie de la ruine. La distribution composée en dimension 1 modélise le cumul du montant des prestations pour un portefeuille de contrats sur une période d'exercice donnée. Le concept de distribution composée s'étend en dimension supérieure à 1 pour permettre la modélisation jointe des risques supportés par plusieurs portefeuilles dont les charges totales sont corrélées. Le Chapitre 1 s'achève sur une introduction à la théorie de la ruine avec la présentation du modèle de ruine de Poisson composé et la définition de la probabilité de ruine ultime.

Le Chapitre 2 présente la méthode numérique d'approximation au centre de ce travail et comprend la plupart des résultats mathématiques de cette thèse. La densité de probabilité d'une variable aléatoire est projetée sur une base de polynômes orthogonaux. L'approximation est obtenue en tronquant à un certain ordre la représentation polynomiale de la densité de probabilité. La [f.d.r.](#) et la [f.d.s.](#) sont approchées par intégration de l'approximation polynomiale de la densité de probabilité. L'application de la méthode d'approximation aux distributions composées en dimension 1 conduit à opter pour une mesure de probabilité gamma associée aux polynômes de Laguerre généralisés. Pour que l'approximation polynomiale soit valide, il est nécessaire de vérifier une condition d'intégrabilité. La Proposition 6 définit une majoration de la densité de probabilité qui permet de choisir les paramètres de la mesure de référence, voir Corollaire 2, pour assurer la validité de l'approximation polynomiale. La qualité de l'approximation dépend de la décroissance des coefficients de la représentation polynomiale. La Proposition 7 indique une convergence en $1/k$, sous certaines conditions de régularité sur la densité de probabilité. La fonction génératrice des coefficients s'exprime en fonction de la transformée de Laplace de la densité de probabilité. L'étude de la forme de cette fonction génératrice permet dans certains cas de choisir les paramètres de la mesure de référence pour accélérer la convergence vers 0 des coefficients. L'approximation de fonctions basées sur des polynômes ou des fonctions orthogonales est un sujet déjà beaucoup étudié dans la littérature, la plus-value de ce travail réside en grande partie dans les façons de choisir les paramètres de la mesure de référence pour obtenir une meilleure approximation. La densité de probabilité multivariée d'un vecteur aléatoire peut aussi être projetée sur une base de polynômes orthogonaux. Ces polynômes sont orthogonaux par rapport à une mesure de référence construite via le produit de mesures de probabilité univariées appartenant aux [Familles Exponentielles Naturelles Quadratiques \(FENQ\)](#). La représentation polynomiale de la densité de probabilité multivariée est justifiée par le Théorème 7. L'approximation polynomiale découle d'une troncature des séries infinies qui définissent la représentation polynomiale. La [Fonction De Répartition Multivariée \(f.d.r.m.\)](#) et la [Fonction De Survie Multivariée \(f.d.s.m.\)](#) sont obtenues par intégration de l'approximation polynomiale. L'application de la méthode d'approximation aux distributions composées multivariées

conduisent à opter pour des mesures de références gamma et des polynômes multivariés issus du produit de polynômes de Laguerre généralisés respectivement orthogonaux par rapport aux mesures de probabilité marginales choisies. La Proposition 9 propose une majoration de la densité de probabilité multivariée. Le choix des paramètres des mesures de référence gamma est effectué conformément au Corollaire 3 de façon à satisfaire la condition d'intégrabilité qui sous-tend la validité de la représentation polynomiale. Le choix des paramètres est affiné via l'étude de la fonction génératrice des coefficients qui s'exprime en fonction de la transformée Laplace du vecteur aléatoire considéré. La représentation polynomiale en dimension 2 est reliée aux lois de probabilité de Lancaster. Le vecteur aléatoire introduit dans la Définition 12 admet une distribution particulière sur \mathbb{R}_+^2 . La Proposition 10 établit l'appartenance de cette distribution aux lois de Lancaster de type gamma-gamma. Une illustration numérique des performances de la méthode d'approximation polynomiale sur l'approximation d'une distribution Poisson composée est proposée et une étude comparative avec les autres méthodes numériques évoquées dans ce travail est conduite. En plus de l'approximation de la densité de probabilité et de la f.d.s., une approximation polynomiale de la prime *stop-loss* usuelle est étudiée. Le Chapitre 2 s'achève sur la description de l'estimateur de la densité de probabilité basé sur la formule d'approximation polynomiale. Une discussion autour de son application concrète à l'estimation des distributions composées en fonction des données disponibles conclut le Chapitre 2.

Le Chapitre 3 présente l'application de la méthode d'approximation via la représentation polynomiale à l'évaluation de la probabilité de ruine ultime dans le modèle de ruine de Poisson composé. Une comparaison avec d'autres méthodes numériques d'approximation est effectuée. Le Chapitre 3 est un article scientifique accepté, [GOFFARD et collab. \[2015b\]](#).

Le Chapitre 4 présente l'application de la méthode d'approximation via la représentation polynomiale à l'évaluation des probabilités associées à une distribution composée bivariée. L'usage de telles distributions en actuariat et plus particulièrement en réassurance est justifié. Une comparaison avec d'autres méthodes numériques d'approximation est également effectuée. Le Chapitre 4 est un article scientifique soumis, [GOFFARD et collab. \[2015a\]](#).

Le Chapitre 5 décrit la procédure d'agrégation des portefeuilles de contrat d'assurance vie de type épargne. Il s'agit d'une procédure en deux étapes, comprenant une classification des contrats en portefeuille via des algorithmes de classification classiques en analyse de données, suivie de la détermination d'un contrat représentatif, appelé *model point*, pour chacune des classes. Le portefeuille de contrats agrégé est utilisé comme input du modèle interne de projection des *cash-flows*. Le Chapitre 5 est un article scientifique accepté, [GOFFARD et GUERRAULT \[2015\]](#).

Mes Publications

- GOFFARD, P.-O. et X. GUERRAULT. 2015, «Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points?», *European Actuarial Journal*, p. 1–16. [4](#)
- GOFFARD, P.-O., S. LOISEL et D. POMMERET. 2015a, «Polynomial approximations for bivariate aggregate claim amount probability distributions», *Working Paper*. [4](#)
- GOFFARD, P.-O., S. LOISEL et D. POMMERET. 2015b, «A polynomial expansion to approximate the ultimate ruin probability in the compound Poisson ruin model», *Journal of Computational and Applied Mathematics*. [4](#)

Chapitre 1

Les distributions composées et la théorie de la ruine

Sommaire

1.1 Les distributions composées en dimension 1	8
1.1.1 Définition, propriétés et interprétations	8
1.1.2 Méthodes d'approximation des distributions composées via la discrétisation de la loi du montant des sinistres	11
1.1.3 Les méthodes numériques d'inversion de la transformée de La- place	16
1.2 Les distributions composées en dimension n	21
1.3 Théorie de la ruine : L'approximation de la probabilité de ruine ul- time dans le modèle de ruine de Poisson composée	26
1.3.1 Présentation du modèle de ruine et définition de la probabilité de ruine ultime	26
1.3.2 Approximation de la probabilité de ruine ultime : Revue de la littérature	31
1.4 Références	34

1.1 Les distributions composées en dimension 1

1.1.1 Définition, propriétés et interprétations

Le sujet au coeur de la gestion des risques d'une compagnie d'assurance est l'évaluation de ses engagements futurs vis-à-vis des assurés. Dans le cas d'une compagnie d'assurance non-vie, l'engagement de la compagnie d'assurance est égal au montant cumulé des prestations qui seront versées aux assurés sur une période d'exercice. Le nombre de sinistres subis par les assurés sur cette période est modélisé par une variable aléatoire de comptage N . Le coût unitaire des sinistres est modélisé par une variable aléatoire U positive et continue. Le cumul des prestations, aussi appelé charge totale du portefeuille, est une variable aléatoire X , qui admet une distribution de probabilité composée.

Définition 1. *La variable aléatoire X suit une loi de probabilité composée $(\mathbb{P}_N; \mathbb{P}_U)$ si elle admet la forme*

$$X = \sum_{i=1}^N U_i, \quad (1.1)$$

où N est une variable aléatoire de comptage de loi de probabilité \mathbb{P}_N et de densité $f_N(k) = \mathbb{P}(N = k)$, avec $k \in \mathbb{N}$. La suite $\{U_i\}_{i \in \mathbb{N}}$ est formée de variables aléatoires réelles, positives, *Indépendantes et Identiquement Distribuées (i.i.d.)* suivant une loi de probabilité, \mathbb{P}_U , et de densité f_U par rapport à la mesure de Lebesgue. La variable aléatoire N est supposée indépendante de la suite de variables aléatoires $\{U_i\}_{i \in \mathbb{N}}$. Enfin, il est supposé par convention que $X = 0$ si $N = 0$.

Cette approche globale dans laquelle le portefeuille est un ensemble de contrats anonymes porte le nom de modèle collectif en théorie du risque. Ce modèle est étudié dans le chapitre 12 de l'ouvrage de **BOWERS et collab. [1997]**, le chapitre 4 de l'ouvrage de **ROLSKI et collab. [1999]**, le chapitre 3 de l'ouvrage de **KAAS et collab. [2008]** et le chapitre 9 de l'ouvrage de **KLUGMAN et collab. [2012]**. Le portefeuille est vu comme un ensemble qui subit une série de chocs causés par l'occurrence des sinistres. La variable aléatoire X donne l'ampleur des engagements de la compagnie d'assurance auprès de ses assurés.

La variable aléatoire (1.1) admet une distribution de probabilité mixte avec une masse de probabilité non nulle en 0, avec

$$d\mathbb{P}_X(x) = f_N(0)\delta_0(x) + d\mathbb{G}_X(x), \quad (1.2)$$

où δ_0 désigne la mesure de Dirac en 0, et \mathbb{G}_X est une mesure de probabilité défailante qui forme la partie absolument continue de la distribution de probabilité de X par rapport à la mesure de Lebesgue, sa densité de probabilité est notée g_X . La mesure de probabilité \mathbb{G}_X est défailante au sens où

$$\int d\mathbb{G}_X(x) = 1 - f_N(0) \leq 1.$$

L'expression de la densité de probabilité défailante g_X est donnée par

$$g_X(x) = \sum_{k=1}^{+\infty} f_N(k) f_U^{(*k)}(x), \quad (1.3)$$

où $f_U^{(*k)}$ est la densité de la somme de k variables aléatoires *i.i.d.* de loi \mathbb{P}_U , son expression est obtenue par convolutions successives de f_U avec elle-même. L'expression de la densité de probabilité défaillante (1.3) est difficile à exploiter pour effectuer des calculs, en raison de la présence d'une série infinie dans son expression. Il est dès lors difficile d'effectuer des calculs de probabilité et d'avoir accès à la *f.d.r.* et à la *f.d.s.* de X qui sous-tendent des applications intéressantes.

La *f.d.s.* est par exemple très utile pour évaluer les primes *stop-loss* généralisées :

Définition 2. La prime *stop-loss* généralisée est définie par

$$\Pi_{c,d}(X) = \mathbb{E} \left[(X - c)_+^d \right] \quad (1.4)$$

où c est un réel positif et d un entier. L'opérateur $(\cdot)_+$ désigne la partie positive.

Certaines paramétrisations de la prime *stop-loss* généralisée correspondent à des quantités classiques en actuariat.

- La prime pure, définie par l'espérance de X , est obtenue en prenant $c = 0$ et $d = 1$, avec

$$\Pi_{0,1}(X) = \mathbb{E}(X). \quad (1.5)$$

- la *f.d.s.* de la variable aléatoire de X est obtenue lorsque $d = 0$, avec

$$\Pi_{0,c}(X) = \overline{F}_X(c).$$

- La prime *stop-loss* usuelle de seuil de rétention c est obtenue pour $d = 1$, avec

$$\Pi_{1,c}(X) = \mathbb{E}[(X - c)_+]. \quad (1.6)$$

La prime *stop-loss* usuelle (1.6) est la prime associée aux traités globaux de réassurance non-proportionnelle, réputés optimaux théoriquement. La prime *stop-loss* généralisée se calcule par intégration de la *f.d.s.* de X .

Proposition 1. La prime *stop-loss* généralisée s'écrit

$$\Pi_{c,d} = \int_0^{+\infty} y^{d-1} \Pi_{c+y,0}(X) dy. \quad (1.7)$$

Ainsi l'accès à la *f.d.s.* de X permet de calculer toutes les primes *stop-loss*.

Sur une période d'exercice donnée, la compagnie d'assurance reçoit des primes de la part des preneurs d'assurance, dont le montant cumulé est noté P , et doit déboursier X pour régler les prestations suite à l'occurrence des sinistres. Le principe de tarification par la moyenne consiste à déterminer le niveau de prime à partir de la prime pure (1.5). Le cumul des primes commerciales est égal à

$$P = \Pi_{0,1}(X)(1 + \eta), \quad (1.8)$$

où η désigne un chargement de sécurité qui indique de combien les primes reçues excèdent le coût moyen pour la compagnie d'assurance. La valeur de η est généralement

donnée en pourcentage. Une compagnie d'assurance souhaitant conquérir un marché peut utiliser le levier des prix en optant pour un η inférieur à celui des autres acteurs économiques du marché. Une fois que le niveau des primes (1.8) a été fixé, la f.d.s. de la charge totale permet l'évaluation des probabilités de ruine annuelle, définies par

$$\psi(u) = \mathbb{P}(X > P + u). \quad (1.9)$$

Le résultat technique annuel est donné par $P - X$ et u est un capital dont la fonction est d'amortir un résultat annuel déficitaire. Le montant du capital u est solution de l'équation

$$\psi(u) = \alpha,$$

où α désigne le niveau de risque pris par la compagnie d'assurance, ce niveau oscille typiquement entre 0.5% et 0.01%. Le choix du niveau de risque est le fruit d'un arbitrage car la réserve de capital u est en un certain sens gelé, il ne fait pas partie de la stratégie d'investissement de la compagnie d'assurance. Le choix d'une gestion des risques prudente, avec une faible probabilité de ruine, associée à un capital u élevé, affecte négativement le résultat financier de la compagnie d'assurance. Une connaissance précise de la f.d.s. de X permet la détermination de u associé à un niveau de risque α , avec

$$u_\alpha = \text{VaR}_X(1 - \alpha) - P,$$

où $\text{VaR}_X(1 - \alpha)$ désigne la **Value-at-Risk (VaR)** de niveau $1 - \alpha$ de la variable aléatoire X . Il s'agit du quantile d'ordre $1 - \alpha$ de la distribution de X . La connaissance de la distribution de X est importante pour que la compagnie d'assurance puisse déterminer le niveau des primes et ses marges de solvabilité. La densité de probabilité défailante (1.3), nécessaire pour effectuer les calculs, admet une forme explicite dans un nombre très restreint de cas. Ce constat motive la mise au point de méthodes numériques pour approcher cette fonction. Ces méthodes numériques se doivent notamment d'approcher finement la f.d.r. et la f.d.s. qui sont très utiles dans les applications. Contrairement à la densité, il est à noter que la **Fonction Génératrice des Moments (f.g.m.)** de la variable aléatoire X , définie par

$$\mathcal{L}_X = \mathbb{E}(e^{sX}), \quad (1.10)$$

admet une forme assez simple dans la plupart des cas.

Proposition 2. La f.g.m. de la variable aléatoire X est donnée par

$$\mathcal{L}_X(s) = \mathcal{G}_N[\mathcal{L}_U(s)], \quad (1.11)$$

où $\mathcal{G}_N(s) = \mathbb{E}(s^N)$ désigne la **Fonction Génératrice des Probabilités (f.g.p.)** de la variable aléatoire N . L'espérance de X est donnée par

$$\mathbb{E}(X) = \mathbb{E}(N)\mathbb{E}(U), \quad (1.12)$$

et sa variance par

$$\mathbb{V}(X) = \mathbb{E}(N)\mathbb{V}(U) + \mathbb{E}(U^2)\mathbb{V}(N).$$

Les notions de f.g.m. et de transformée de Laplace des variables aléatoires sont confondues dans ce document. La Proposition 2 conduit naturellement à penser aux méthodes numériques basées sur les moments et l'inversion de la transformée de Laplace pour accéder à la fonction g_X . Les sous-sections suivantes offrent un tour d'horizon des méthodes utilisées en actuariat.

1.1.2 Méthodes d'approximation des distributions composées via la discrétisation de la loi du montant des sinistres

Algorithme de Panjer : définition et mise en oeuvre

L'algorithme de Panjer est une méthode récursive permettant l'évaluation exacte des probabilités d'une distribution composée $(\mathbb{P}_N, \mathbb{P}_U)$ lorsque le montant des sinistres admet une loi de probabilité discrète caractérisée par $f_U(k) = \mathbb{P}(U = k)$, où $k \in \mathbb{N}$. Dans ce cas, la loi de X est aussi discrète, et l'algorithme de Panjer retourne les valeurs des probabilités

$$f_X(k) = \mathbb{P}(X = k), \quad k \in \mathbb{N}. \quad (1.13)$$

L'algorithme est applicable à condition que la loi de probabilité \mathbb{P}_N du nombre de sinistres N appartienne à la famille de Panjer. Cette famille de loi est caractérisée par une relation de récurrence d'ordre 1 sur les probabilités de N .

Définition 3 (Panjer 1981). *La distribution de N appartient à la famille de Panjer si ses probabilités vérifient*

$$f_N(k) = \left(a + \frac{b}{k}\right) f_N(k-1), \quad k = 1, 2, \dots, \quad (1.14)$$

où $a < 1$ et $b \in \mathbb{R}$.

Cette famille de lois de probabilité a été initialement introduite par KATZ [1965], et entièrement caractérisée par SUNDT et JEWELL [1981].

Theorème 1 (Sundt & Jewell 1981). *Soit N une variable aléatoire discrète à valeurs entières. Si la loi de probabilité de N vérifie la relation (1.14) alors N admet une distribution binomiale, binomiale négative ou Poisson. De plus,*

- (i) *Si $a = 0$, alors $b > 0$ et N suit une loi de Poisson $\mathcal{P}(b)$.*
- (ii) *Si $0 < a < 1$, alors $a + b > 0$, et N suit une loi binomiale négative $\mathcal{NB}(\alpha, p)$ avec $p = a$ et $\alpha = 1 + bp^{-1}$*
- (iii) *Si $a < 0$, alors $b = -a(n+1)$ pour $n \in \mathbb{N}$, et N suit une loi binomiale $\mathcal{B}(n, p)$ avec $p = a(a-1)^{-1}$ et $n = -1 - ba^{-1}$*

L'expression de la densité et de la f.g.p. des lois binomiale, binomiale négative et Poisson sont données ci-après.

Définition 4. *La densité d'une variable aléatoire N de loi binomiale $\mathcal{B}(n, p)$ est donnée par*

$$f_N(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k \in \{1, \dots, n\}, \quad (1.15)$$

et sa f.g.p. par

$$\mathcal{G}_N(x) = (1 - p + px)^n. \quad (1.16)$$

Le cas particulier où $n = 1$ correspond à la loi de Bernoulli.

Définition 5. La densité d'une variable aléatoire N de loi binomiale négative $\mathcal{BN}(\alpha, p)$ est donnée par

$$f_N(k) = \frac{\Gamma(\alpha + k)}{\Gamma(k + 1)\Gamma(\alpha)} (1 - p)^\alpha p^k, \quad k \in \mathbb{N}, \quad (1.17)$$

et sa f.g.p. par

$$\mathcal{G}_N(x) = \left(\frac{1 - p}{1 - ps} \right)^\alpha. \quad (1.18)$$

Le cas particulier où $\alpha = 1$ correspond au cas particulier de la loi géométrique.

Définition 6. La densité d'une variable aléatoire N de loi Poisson $\mathcal{P}(\lambda)$ est donnée par

$$f_N(k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k \in \mathbb{N}, \quad (1.19)$$

et sa f.g.p. par

$$\mathcal{G}_N(x) = e^{\lambda(s-1)}. \quad (1.20)$$

L'algorithme de Panjer est issu du travail de **PANJER [1981]**.

Theorème 2 (Panjer (1981)). Soit X une variable aléatoire de distribution composée $(\mathbb{P}_N, \mathbb{P}_U)$. Si U est une variable aléatoire discrète à valeurs entières et que N est une variable aléatoire de comptage dont les probabilités vérifient l'équation de récurrence (1.14), alors

$$f_X(k) = \begin{cases} \mathcal{G}_N[f_U(0)], & \text{pour } k = 0, \\ \frac{1}{1 - af_U(0)} \sum_{j=1}^k \left(a + \frac{bj}{k} \right) f_U(j) f_X(k - j), & \text{pour } k \geq 1. \end{cases} \quad (1.21)$$

Le premier inconvénient de l'algorithme de Panjer est la restriction à une certaine classe de distributions pour la fréquence des sinistres. **SUNDT [1992]** décrit une famille de distributions plus étendue, caractérisée par la relation de récurrence

$$f_N(k) = \sum_{i=1}^k \left[a(i) + \frac{b(i)}{k} \right] f_N(k - i), \quad (1.22)$$

permettant aussi la définition d'un algorithme à l'image de l'algorithme (1.21) pour déterminer les probabilités de X . L'autre inconvénient de l'algorithme de Panjer est de devoir se limiter à des lois de probabilités discrètes pour la loi de la sévérité des sinistres. Il est plus réaliste de modéliser le montant des sinistres par une loi de probabilité continue à support sur \mathbb{R}^+ . Si la loi de N appartient à la famille de Panjer et que la loi de U est continue alors la partie absolument continue de la loi de X , de densité g_X , est solution de l'équation intégrale

$$g_X(x) = f_N(1) f_U(x) + \int_0^x \left(a + b \frac{y}{x} \right) f_U(y) g_X(x - y) dx. \quad (1.23)$$

STRÖTER [1985] a proposé une méthode d'approximation en résolvant numériquement l'équation intégrale (1.23) via l'emploi d'un noyau défini à partir de la densité de la loi de U . L'autre solution consiste à discrétiser la loi de U , en choisissant un pas de discrétisation h . Le support de U devient un ensemble discret de la forme $\{0, h, 2h, \dots\}$.

L'application de l'algorithme de Panjer dans ce cas produit une approximation de la loi de probabilité de X , avec

$$f_X(x) \approx \tilde{f}_X(kh), \quad x \in [kh, (k+1)h[, \quad k = 0, 1, \dots \quad (1.24)$$

La précision est fonction décroissante de h , mais un pas de discrétisation faible augmente les temps de calcul. Le fonctionnement de l'algorithme (1.21) implique l'évaluation de $f_X(h), \dots, f_X[(k-1)h]$ préalablement au calcul de $f_X(kh)$. Ainsi, l'obtention d'une bonne approximation de $f_X(x)$, qui va de paire avec un faible pas de discrétisation, implique un grand nombre d'itérations et des temps de calculs parfois inacceptables lorsque x est grand. Cela explique pourquoi l'optimisation de la méthode de discrétisation de la loi de U visant à limiter la perte d'information est un sujet important. GERBER et JONES [1976] ont proposé de discrétiser la loi de U via

$$\tilde{f}_U(kh) = \begin{cases} F_U\left(\frac{h}{2}\right), & \text{pour } k = 0, \\ F_U\left(kh + \frac{h}{2}\right) - F_U\left(kh - \frac{h}{2}\right), & \text{pour } k \geq 1, \end{cases} \quad (1.25)$$

ce qui revient à discrétiser la distribution de U en arrondissant au plus proche multiple de h . Cette méthode simple et efficace présente l'inconvénient de ne pas conserver les moments de la distribution initiale. GERBER [1982] a mis au point la méthode **Local Moment Matching (LMM)**, qui permet de discrétiser la distribution en conservant les moments de la distribution jusqu'à un certain ordre n . L'idée est d'approcher la distribution de U sur des intervalles de la forme $[knh, (k+1)nh]$ par des masses de probabilités $f_U(knh + jh)$, où $j \in \{1, \dots, n\}$, vérifiant le système d'équations

$$\sum_{j=1}^n (knh + jh)^i f_U(knh + jh) = \int_{knh}^{(k+1)nh} x^i d\mathbb{P}_U(y), \quad i = 0, \dots, n. \quad (1.26)$$

Ce système d'équations a pour solution

$$\tilde{f}_U(knh + jh) = \int_{knh}^{(k+1)nh} \prod_{i \neq j} \frac{x - knh - ih}{(j - i)h} d\mathbb{P}_U(x), \quad j = 0, \dots, n. \quad (1.27)$$

Les probabilités (1.27) peuvent prendre des valeurs négatives, ce problème est étudié dans le travail de COURTOIS et DENUIT [2007]. VILAR [2000] propose une variante de la méthode LMM, appliquée à l'ordre 2, dans laquelle le moment d'ordre 2 de la distribution discrétisée est choisie le plus proche possible du moment d'ordre 2 de la véritable distribution de U sous une contrainte de positivité des probabilités (1.27). WALHIN et PARIS [1998] propose de trouver la distribution discrète qui minimise la distance de Kolmogorov.

Définition 7. Soit X et Y deux variables aléatoires réelles de loi de probabilité respectives \mathbb{P}_X et \mathbb{P}_Y . La distance de Kolmogorov entre \mathbb{P}_X et \mathbb{P}_Y est

$$\|\mathbb{P}_X - \mathbb{P}_Y\|_K = \sup_{x \in \mathbb{R}} |F_X(x) - F_Y(x)|. \quad (1.28)$$

La minimisation de (1.28) peut se faire sous la contrainte de conserver les moments. Le prix à payer est l'augmentation de la distance minimale entre la distribution discrétisée et la véritable distribution. L'algorithme de Panjer, couplé à une méthode de discrétisation de la distribution de U , donne accès à des quantités intéressantes se rapportant à la distribution de X , comme par exemple la *f.d.r.*, avec

$$F_X(x) \approx \sum_{j=0}^k \tilde{f}_X(jh), \quad x \in [kh, (k+1)h[, \quad (1.29)$$

ou la *f.d.s.* par

$$\bar{F}_X(x) \approx 1 - \sum_{j=0}^k \tilde{f}_X(jh), \quad x \in [kh, (k+1)h[. \quad (1.30)$$

Des formules récursives permettent aussi de déterminer directement certaines quantités comme les moments de la distribution composée avec ce résultat de DE PRIL [1986].

Proposition 3 (De Pril (1986)). *Soit X une variable aléatoire admettant une distribution composée $(\mathbb{P}_N, \mathbb{P}_U)$. Si la loi de probabilités de N vérifie la relation de récurrence (1.14), alors*

$$\mathbb{E}(X^{k+1}) = \frac{1}{1-a} \sum_{j=0}^k \binom{k}{j} \left(\frac{k+1}{j+1} a + b \right) \mathbb{E}(U^{j+1}) \mathbb{E}(X^{k-j}). \quad (1.31)$$

Il est aussi possible de mettre au point une récursion pour évaluer la prime *stop-loss* usuelle de seuil de rétention c (1.6). D'après la Proposition 1, la prime peut se réécrire en fonction de la *f.d.s.*. En prenant c , tel qu'il existe $k \in \mathbb{N}$ vérifiant $c = kh$, il vient

$$\begin{aligned} \Pi_{kh,1}(X) &= \int_0^{+\infty} \Pi_{kh+y,1}(X) dy \\ &= \int_{kh}^{+\infty} \bar{F}_X(y) dy \\ &\approx \sum_{j=k}^{+\infty} \bar{F}_X(jh) \\ &\approx \Pi_{(k-1)h,1}(X) - \bar{F}_X[(k-1)h]. \end{aligned} \quad (1.32)$$

L'équation (1.32) est une formule récursive dans laquelle l'itération est effectuée sur le seuil de rétention. La *f.d.s.* est approchée via l'équation (1.30) et l'initialisation est effectuée en observant que $\Pi_{0,1}(X) = \mathbb{E}(X)$. Une formule récursive plus générale et applicable sur un ensemble de fonctions définies à l'aide d'un opérateur est établie dans le papier de DHAENE et collab. [1999].

L'inversion de la transformée de Fourier discrète : l'algorithme Fast Fourier Transform

L'algorithme *Fast Fourier Transform* (FFT) est le concurrent historique de l'algorithme de Panjer. Il est né de la nécessité d'optimiser les temps de calcul et se fonde sur l'accessibilité de la transformée de Laplace associée aux distributions composées. A l'instar de l'algorithme de Panjer, l'algorithme FFT permet l'évaluation exacte des probabilités d'une distribution composée $(\mathbb{P}_N, \mathbb{P}_U)$ à condition que la variable aléatoire U

soit discrète. La contrainte supplémentaire pour une évaluation exacte est que le support de U soit fini.

Soit X une variable aléatoire discrète, de support $\{1, \dots, n\}$. La loi de probabilité de X est entièrement caractérisée par $f_X(k) = \mathbb{P}(X = k)$ pour $k = 1, \dots, n$. La transformée de Fourier de X est donnée par

$$\mathcal{L}_X(is) = \sum_{k=1}^n e^{isk} f_X(k), \quad s \in \mathbb{R}. \quad (1.33)$$

L'algorithme FFT permet d'inverser cette formule en exprimant la suite $\{f_X(j)\}_{j \in \{1, \dots, n\}}$ en fonction d'une suite de valeurs de la transformée de Fourier $\{\mathcal{L}_X(is_k)\}_{k \in \{1, \dots, n\}}$ sur une grille de points $\{s_k\}_{k \in \{1, \dots, n\}}$, définie par

$$s_k = \frac{2\pi(k-1)}{n} \quad k = 1, \dots, n. \quad (1.34)$$

Soit $\mathbf{p}_X = [f_X(1), \dots, f_X(n)]$ le vecteur des probabilités de X et $\mathbf{f}_X = [\mathcal{L}_X(is_1), \dots, \mathcal{L}_X(is_n)]$ le vecteur des valeurs de la transformée de Fourier de X . Alors

$$\mathbf{f}_X^T = \mathbf{F} \mathbf{p}_X^T, \quad (1.35)$$

où \mathbf{x}^T désigne la transposé du vecteur $\mathbf{x} = (x_1, \dots, x_n)$ et \mathbf{F} est une matrice de taille $n \times n$ définies par

$$\mathbf{F} = \left(e^{(j-1) \times is_k} \right)_{k, j \in \{1, \dots, n\}} = \begin{pmatrix} 1 & e^{1 \times is_1} & \dots & 1 \\ 1 & e^{1 \times is_2} & \dots & e^{(n-1) \times is_2} \\ \vdots & \vdots & & \vdots \\ e^{0 \times is_k} & e^{1 \times is_k} & \dots & e^{(n-1) \times is_k} \\ \vdots & \vdots & & \vdots \\ 1 & e^{1 \times is_n} & \dots & e^{(n-1) \times is_n} \end{pmatrix}.$$

La matrice \mathbf{F} est une matrice unitaire à une constante de normalisation près ce qui permet de l'inverser aisément via sa transconjuguée, avec

$$\mathbf{F}^{-1} = \left(\frac{1}{n} e^{-is_j(k-1)} \right)_{k, j \in \{1, \dots, n\}}. \quad (1.36)$$

L'équation (1.35) implique que $\mathbf{p}_X^T = \mathbf{F}^{-1} \mathbf{f}_X^T$, ce qui donne explicitement les probabilités en fonction de la transformée de Fourier, avec

$$f_X(j) = \frac{1}{n} \sum_{k=1}^n \mathcal{L}_X(is_k) e^{-is_k(j-1)} \quad j = 1, \dots, n, \quad (1.37)$$

et $f_X(n) = 1 - \sum_{j=1}^{n-1} f_X(j)$. Lorsque le support de la loi de probabilité de X est infini, l'algorithme FFT retourne une approximation des probabilités $f_X(1), \dots, f_X(n)$. La 2π -périodicité de l'application $k \mapsto e^{isk}$ pour $s \geq 0$ permet d'opérer à un ré-arrangement des termes dans la série infinie qui définit la transformée de Laplace de X , précisément

$$\mathcal{L}_X(is_k) = \sum_{j=1}^{+\infty} f_X(j) e^{is_k(j-1)} = \sum_{j=1}^{+\infty} \sum_{l=1}^n f_X[j + n(l-1)] e^{is_k(j-1+n(l-1))} = \sum_{l=1}^n \tilde{f}_X(l) e^{is_k(l-1)}, \quad (1.38)$$

où $\tilde{f}_X(l) = \sum_{j=1}^{+\infty} f_X(j + nl)$. La suite de termes $\tilde{f}_X(1), \dots, \tilde{f}_X(n)$ est obtenue via la relation (1.37), après évaluation de $\mathcal{L}_X(is_1), \dots, \mathcal{L}_X(is_n)$. Le terme $\tilde{f}_X(k)$ approche la probabilité $f_X(k)$ pour $k = 1, \dots, n$, avec une erreur égale à

$$|\tilde{f}_X(k) - f_X(k)| = \sum_{j=2}^{+\infty} f_X(j + nl). \quad (1.39)$$

L'erreur (1.39) est l'erreur d'*aliasing*, elle tend vers 0 lorsque n tend vers l'infini. L'algorithme FFT a été utilisé pour approcher la loi de probabilité d'une distribution composée dans le papier de HECKMAN et MEYERS [1983]. Comme pour l'algorithme de Panjer, si le montant des sinistres est gouverné par une loi de probabilité continue alors il faut préalablement procéder à la discrétisation de cette loi en usant des techniques évoquées précédemment. L'approximation de la f.d.r., de la f.d.s. ou encore de la prime *stop-loss* résulte de l'approximation de la loi de probabilité de X . L'avantage de l'algorithme FFT est qu'il peut s'appliquer avec n'importe quel type de loi pour la fréquence des sinistres. Le désavantage est l'erreur d'*aliasing* qui s'ajoute à l'erreur de discrétisation. L'erreur d'*aliasing* peut être contrôlée et diminuée en effectuant un changement de mesure comme le montre les travaux de GRÜBEL et HERMESMEIER [1999]. Cette considération est aussi à nuancer car l'algorithme FFT est moins gourmand en temps de calcul ce qui permet de diminuer le pas de discrétisation et donc l'erreur résultant de la phase de discrétisation. Les articles de BÜHLMAN [1984] et EMBRECHTS et FREI [2009] comparent l'efficacité des deux approches.

1.1.3 Les méthodes numériques d'inversion de la transformée de Laplace

Ce document présente de nouvelles méthodes pour concurrencer l'algorithme de Panjer et l'algorithme FFT. Ces méthodes doivent permettre une augmentation de la précision sans augmenter les temps de calculs. Le cahier des charges supprime notamment l'étape de discrétisation de la loi de probabilité pour le montant des sinistres. Ces méthodes numériques permettent une approximation de la densité de probabilité ou de la f.d.r. à partir de la transformée de Laplace, ce qui les rend particulièrement attractives dans le cadre des distributions composées puisqu'une expression complexe de la densité de probabilité s'oppose à une expression simple de la transformée de Laplace.

La méthode au centre de ce travail est basée sur une représentation polynomiale de la densité par rapport à une mesure de probabilité de référence. Cette mesure de référence appartient aux FENQ. Elle est choisie absolument continue par rapport à la loi de probabilité associée à la distribution composée. La densité de la distribution composée par rapport à la mesure de référence est projetée sur une base de polynômes orthonormaux par rapport à la mesure de référence. La représentation polynomiale de la densité par rapport à la mesure de référence est valide sous réserve que celle-ci soit de carré intégrable. Le Chapitre 2 décrit la méthode et fait le lien avec d'autres méthodes d'approximation basées sur les polynômes orthogonaux. L'approximation résulte de la troncature simple de la série infinie issue de la projection orthogonale. Dans le cadre de

l'application aux distributions composées, la mesure de référence est une loi gamma et les polynômes sont les polynômes de Laguerre généralisés. L'application de la méthode améliore le travail de BOWERS [1966] dans lequel la loi de probabilité de la variable aléatoire (1.1) est représentée par une somme finie de densités gamma. Les techniques de calcul pour les coefficients du développement sont inspirées de la méthode Laguerre, présentée dans ABATE et collab. [1995], qui consiste à projeter la densité sur une base de fonctions orthogonales construites par produit des polynômes de Laguerre simples et de la fonction exponentielle décroissante. La nouveauté réside dans le choix des paramètres de la loi gamma. Ce choix conditionne la validité du développement polynomial, certaines paramétrisations ne permettent pas la vérification de la condition d'intégrabilité ce qui rend la formule d'approximation caduque. Un choix judicieux peut permettre l'optimisation de la vitesse de convergence de la somme partielle. Cette méthode peut être qualifiée de méthode d'approximation semi-paramétrique. Ce terme est employé dans le papier de PROVOST [2005] qui lie une méthode générale pour approcher les densités de probabilités à partir des moments théoriques de la distribution et les méthodes d'approximation basées sur les polynômes orthogonaux. L'application de cette méthode pour les distributions composées conduit aussi à utiliser les polynômes de Laguerre et la loi gamma, voir JIN et collab. [2014]. Les coefficients de la représentation, définie par produit scalaire, sont donnés par l'espérance des polynômes avec en argument la variable aléatoire associée à la distribution composée. La formule d'approximation est ainsi basée sur l'évaluation de combinaisons linéaires de moments de la variable aléatoire (1.1). Le remplacement de ces quantités par leur contrepartie empirique doublé d'une procédure d'estimation pour les paramètres de la mesure de référence conduit à la mise au point d'un estimateur semi-paramétrique de la densité à partir de la formule d'approximation. La fin du Chapitre 2 est consacrée à l'étude de l'estimateur de la densité basé sur la représentation polynomiale.

La détermination d'une distribution de probabilité à partir de ces moments se rapporte au *Stieljes moment problem* si le support est dans \mathbb{R}^+ , au *Hamburger moment problem* si le support est dans \mathbb{R} et au *Hausdorff moment problem* si le support est dans $[0, 1]$. L'étude de ces problèmes conduit à la mise au point de formules d'approximation de la densité basée sur les moments de la distribution qui se déclinent naturellement en estimateurs non-paramétriques de la densité. Ces estimateurs sont adaptés à l'estimation statistique de quantités qui ne sont pas observables directement. Par exemple, si une variable aléatoire X est gouvernée par une loi de probabilité paramétrique \mathbb{P}_X où la valeur du paramètre Θ est une variable aléatoire de loi \mathbb{P}_Θ . Il existe souvent une relation simple entre les moments de X et de Θ . Il résulte de l'injection de la relation dans l'estimateur de la densité de \mathbb{P}_Θ , un estimateur basé sur des observations de X et non de Θ . Les distributions composées sont des cas particuliers de mélange infini de distributions. C'est pourquoi les estimateurs basés sur les méthodes d'approximations présentées ici, sont potentiellement adaptés à l'estimation des distributions composées. La fin du Chapitre 2 comprend une discussion portant sur l'exploitation de données sur les montants et la fréquence des sinistres pour obtenir une estimation de la densité de X à l'aide d'estimateurs basés sur les méthodes d'approximation étudiées dans ce document.

Approximations de la f.d.r. et de la densité de probabilité via les moments exponentiels

Soit Y une variable aléatoire réelle, continue, de loi de probabilité \mathbb{P}_Y à support dans $[0, 1]$. L'idée de [MNATSAKANOV et RUYMGAART \[2003\]](#) est d'approcher la f.d.r. de la variable aléatoire Y par la f.d.r. d'une variable aléatoire N de loi binomiale mélange $\mathcal{B}(n, Y)$, avec

$$F_Y(y) \approx \sum_{k=0}^{\lfloor ny \rfloor} \mathbb{P}_N(N = k), \quad (1.40)$$

où

$$\mathbb{P}_N(N = k) = \int_0^1 \binom{n}{k} z^k (1-z)^{n-k} d\mathbb{P}_Y(z). \quad (1.41)$$

La réinjection de l'expression des probabilités de N dans la f.d.r. de la loi binomiale mélange (1.40) conduit à

$$\begin{aligned} \sum_{k=0}^{\lfloor ny \rfloor} \mathbb{P}_N(N = k) &= \int_0^1 \sum_{k=0}^{\lfloor ny \rfloor} \binom{n}{k} z^k (1-z)^{n-k} d\mathbb{P}_Y(z) \\ &= \sum_{k=0}^{\lfloor ny \rfloor} \binom{n}{k} \mathbb{E} \left[Y^k (1-Y)^{n-k} \right] \\ &= \sum_{k=0}^{\lfloor ny \rfloor} \sum_{j=k}^n \binom{n}{j} \binom{j}{k} (-1)^{j-k} \mathbb{E} \left(Y^j \right). \end{aligned} \quad (1.42)$$

L'approximation (1.40) est justifiée par le résultat de convergence

$$\sum_{k=0}^{\lfloor ny \rfloor} \binom{n}{k} z^k (1-z)^{n-k} \rightarrow \begin{cases} 1 & \text{pour } z < y \\ 0 & \text{pour } z > y \end{cases}, \quad n \rightarrow +\infty, \quad (1.43)$$

qui implique que

$$\lim_{n \rightarrow +\infty} F_Y^n(y) = F_Y(y), \quad \forall y \in [0, 1],$$

où

$$F_Y^n(y) = \sum_{k=0}^{\lfloor ny \rfloor} \sum_{j=k}^n \binom{n}{j} \binom{j}{k} (-1)^{j-k} \mathbb{E} \left(Y^j \right). \quad (1.44)$$

Ainsi la précision de l'approximation (1.44) augmente avec n et donc avec le nombre de moments de la variable aléatoire Y utilisés. Cette méthode de résolution du *Hausdorff moment problem* est décrite exhaustivement dans les articles de [MNATSAKANOV \[2008a,b\]](#). La variable aléatoire X est à support dans \mathbb{R}^+ , l'approximation de sa f.d.r. est obtenue à partir de l'approximation (1.44) via le changement de variable $Y = e^{-\ln(b)X}$, avec

$$F_X^n(x) = 1 - \sum_{k=0}^{\lfloor ne^{-\ln(b)x} \rfloor} \sum_{j=k}^n \binom{n}{j} \binom{j}{k} (-1)^{j-k} \mathcal{L}_X(-j \ln(b)). \quad (1.45)$$

où $1 < b \leq \exp(1)$ est un paramètre à déterminer. Une approximation de la densité de X est obtenue par différentiation de l'approximation de la f.d.r. (1.45), avec

$$f_X^n(x) = \frac{n+1}{n} \frac{F_X^n(x_i) - F_X^n(x_{i-1})}{x_i - x_{i-1}}, \quad x \in [x_{i-1}, x_i[, \quad (1.46)$$

où $x_i = \frac{\ln(n) - \ln(n-i+1)}{\ln(b)}$ pour $i = 1, \dots, n$. L'approximation finale

$$f_X^n(x) = \frac{\lfloor ne^{-\ln(b)x} \rfloor \ln(b) \Gamma(n+2)}{n \Gamma(\lfloor ne^{-\ln(b)x} \rfloor + 1)} \sum_{k=0}^{n - \lfloor ne^{-\ln(b)x} \rfloor} \frac{(-1)^k \mathcal{L}_X[-\ln(b)(k + \lfloor ne^{-\ln(b)x} \rfloor)]}{k! \Gamma(n - \lfloor ne^{-\ln(b)x} \rfloor - k + 1)}, \quad x \in \mathbb{R}^+, \quad (1.47)$$

est obtenue après quelques calculs. L'application de cette méthode d'approximation aux distributions composées a été effectuée par **MNATSAKANOV et SARKISIAN [2013]**.

Approximations de la f.d.r. et de la densité de probabilité via les moments fractionnels

Dans la même veine, la méthode présentée dans le papier de **GZYL et TAGLIANI [2010]** pour résoudre l'*Hausdorff moment problem*, a été adaptée pour être appliquée aux distributions composées dans **GZYL et TAGLIANI [2012]**. Soit Y une variable aléatoire sur $[0, 1]$ de loi de probabilité \mathbb{P}_Y , l'idée est de chercher une densité de probabilité vérifiant des contraintes sur les moments, avec

$$\int_0^1 y^{\alpha_j} f_Y^K(y) dy = \mathbb{E}(Y^{\alpha_j}) \quad k = 0, \dots, K, \quad (1.48)$$

où $\alpha_0 = 0$ pour bien définir une densité. Il est supposé que le problème admet une solution de la forme

$$f_Y^K(y) = \exp\left(-\sum_{k=0}^K \lambda_k y^{\alpha_k}\right). \quad (1.49)$$

Ce choix est justifié dans l'ouvrage de **KAPUR et KESAVAN [1992]**. Les moments dans l'équation (1.48) sont appelés moments fractionnels. Le paramètre λ_0 joue le rôle de constante de normalisation, avec

$$\lambda_0 = \ln\left[\int_0^1 \exp\left(-\sum_{k=1}^K \lambda_k y^{\alpha_k}\right) dy\right]. \quad (1.50)$$

Les séquences $\{\lambda_k\}_{k \in \{1, \dots, K\}}$ et $\{\alpha_k\}_{k \in \{1, \dots, K\}}$ sont déterminées pour optimiser la distance de Kullback-Leibler

$$\mathcal{K} \mathcal{L}(f_Y, f_Y^K) = \int_0^1 \ln\left(\frac{f_Y(y)}{f_Y^K(y)}\right) f_Y(y) dy, \quad (1.51)$$

entre la densité approchée f_Y^K et la vraie densité f_Y . La distance (1.51) se réécrit en termes de différence d'entropie avec

$$\mathcal{K} \mathcal{L}(f_Y, f_Y^K) = \mathcal{H}[f_Y^K] - \mathcal{H}[f_Y], \quad (1.52)$$

où $\mathcal{H}[g] = -\int_0^1 \ln[g(x)] g(x) dx$ désigne l'entropie de la fonction $x \mapsto g(x)$. Ainsi, minimiser la distance de Kullback-Leibler revient à minimiser l'entropie de f_Y^K . Les séquences $\{\lambda_k\}_{k \in \{1, \dots, K\}}$ et $\{\alpha_k\}_{k \in \{1, \dots, K\}}$ sont donc solution des problèmes de minimisation imbriqués

$$\min_{\alpha_1, \dots, \alpha_n} \min_{\lambda_1, \dots, \lambda_n} \left\{ \ln\left[\int_0^1 \exp\left(-\sum_{j=1}^K \lambda_j y^{\alpha_j}\right) dy\right] + \sum_{k=1}^K \lambda_k \mathbb{E}(Y^{\alpha_k}) \right\}. \quad (1.53)$$

Une fois l'optimisation achevée, le changement de variable $Y = e^{-X}$ est utilisé et la densité de la variable aléatoire X est approchée par

$$f_X^K(x) = e^{-x} f_Y^K(e^{-x}). \quad (1.54)$$

Les moments exponentiels et les moments fractionnels peuvent être estimés statistiquement lorsque des données sont disponibles. Des estimateurs de la densité sont alors obtenus en remplaçant les moments théoriques par les moments empiriques.

Inversion de la transformée de Fourier

La méthode d'inversion numérique de la transformée de Fourier est une méthode éprouvée pour approcher la **f.d.s.** d'une distribution de probabilité. Cette méthode jouera le rôle de *benchmark* lorsqu'aucune formule fermée n'est disponible pour la **f.d.s.** de X . Cette méthode est décrite exhaustivement dans **ABATE et WHITT [1992]**. L'idée sous jacente est d'approcher l'intégrale qui fonde la formule d'inversion reliant la **f.d.s.** à sa transformée de Fourier via la méthode des trapèzes.

Soit une fonction $g : \mathbb{R} \rightarrow \mathbb{R}$, continue. Si g est bornée et que

$$\int_{-\infty}^{+\infty} |\mathcal{L}_g(is)| ds < \infty,$$

alors

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-isx} \mathcal{L}_g(is) ds, \quad x \in \mathbb{R}. \quad (1.55)$$

L'objectif est de retrouver la **f.d.s.** de la variable aléatoire X qui est une fonction continue de \mathbb{R}^+ dans $[0, 1]$, donc bornée. La fonction définie par

$$g(x) = \begin{cases} e^{-s_0 x} \overline{F}_X(x), & \text{pour } x \geq 0, \\ g(-x), & \text{pour } x < 0, \end{cases} \quad (1.56)$$

où $s_0 > 0$, constitue une candidate à l'inversion via la formule (1.55). L'application de la formule (1.55) à la fonction (1.56) renvoie

$$\overline{F}_X(x) = \frac{2e^{s_0 x}}{\pi} \int_0^{+\infty} \cos(xs) \Re \left[\mathcal{L}_{\overline{F}_X}(-s_0 - is) \right] ds. \quad (1.57)$$

La méthode d'approximation se fonde sur l'approximation de l'intégrale (1.57) par l'intermédiaire de la méthode des trapèzes, avec

$$\overline{F}_X(x) \approx \frac{2e^{x s_0}}{\pi} h \left\{ \frac{\mathcal{L}_{\overline{F}_X}(s_0)}{2} + \sum_{k=1}^{+\infty} \cos(xkh) \Re \left[\mathcal{L}_{\overline{F}_X}(-s_0 - ikh) \right] \right\}, \quad (1.58)$$

où h est le pas de discrétisation de l'approximation trapézoïdale. L'expression (1.58) ne peut pas être évaluée directement du fait de la série infinie. Une troncature doit être effectuée. L'approximation sera sujette à une erreur de discrétisation et une erreur liée à la troncature. L'idée développée dans **ABATE et WHITT [1992]** est de choisir h et s_0 de manière à, contrôler l'erreur de discrétisation d'une part, et faire apparaître une

série alternée pour pouvoir utiliser une technique d'accélération de la convergence des sommes partielles d'autre part. Le paramétrage en question est $h = \frac{\pi}{2x}$ et $s_0 = \frac{a}{2x}$, où $a > 0$. L'approximation (1.58) devient alors

$$\overline{F_X}^a(x) = \frac{e^{a/2}}{2x} \left\{ \mathcal{L}_{\overline{F_X}}\left(-\frac{a}{2x}\right) - 2 \sum_{k=1}^{+\infty} (-1)^{k+1} \Re \left[\mathcal{L}_{\overline{F_X}}\left(\frac{a+i\pi k}{2x}\right) \right] \right\}, \quad (1.59)$$

car

$$\cos\left(k\frac{\pi}{2}\right) = \begin{cases} 0 & \text{si } k \text{ est impaire,} \\ (-1)^{k/2} & \text{sinon.} \end{cases}$$

De plus, il est possible de montrer que

$$|\overline{F_X}(x) - \overline{F_X}^a(x)| < \frac{e^{-a}}{1 - e^{-a}}. \quad (1.60)$$

La majoration (1.60) suggère que la précision augmente avec celle du paramètre a . Toutefois un arbitrage est nécessaire car des difficultés numériques lors de la mise en pratique peuvent apparaître expliquées notamment par des erreurs d'arrondi lorsque le terme $e^{a/2}$ dans l'approximation (1.59) devient grand. Comme la série infinie dans l'approximation (1.59) est alternée, il est possible d'utiliser la formule de sommation d'Euler au lieu de tronquer simplement la série infinie. L'utilisation de cette procédure conduit à l'approximation finale

$$\overline{F_X}^{a,K,M}(x) = \sum_{m=0}^M \binom{M}{m} 2^{-M} \overline{F_X}^{a,K+m}(x), \quad (1.61)$$

où

$$\overline{F_X}^{a,K}(x) = \frac{e^{a/2}}{2x} \left\{ \mathcal{L}_{\overline{F_X}}\left(-\frac{a}{2x}\right) - 2 \sum_{k=1}^K (-1)^{k+1} \Re \left[\mathcal{L}_{\overline{F_X}}\left(-\frac{a+i\pi k}{2x}\right) \right] \right\}.$$

Dans le Chapitre 2, une illustration des méthodes sur l'approximation de la f.d.s. d'une distribution Poisson composée est proposée. Une étude comparative des performances de l'algorithme de Panjer, la méthode basée sur les moments exponentiels, la méthode d'inversion numérique de la transformée de Fourier et la méthode d'approximation polynomiale est menée.

Le modèle collectif 1 représente le risque supporté par un portefeuille de contrats appartenant à une branche d'activité de la compagnie d'assurance. Les compagnies d'assurance diversifient généralement leur activité pour pouvoir mutualiser les risques. La question de la corrélation entre les branches d'activité est une question d'actualité, sa prise en compte dans la gestion des risques passe par la considération de vecteurs aléatoires admettant des distributions composées multivariées.

1.2 Les distributions composées en dimension n

Le vecteur aléatoire

$$\mathbf{X} = (X_1, \dots, X_n) \quad (1.62)$$

est composé de plusieurs charges totales de sinistres sur une période d'exercice donnée. Il modélise conjointement les risques de plusieurs portefeuilles de contrats pour prendre en compte de façon rigoureuse leurs éventuelles corrélations. Une compagnie d'assurance diversifiée généralement son activité et dispose ainsi de plusieurs branches d'activité. le nombre n désigne alors le nombre de branches d'activité. Les composantes du vecteur aléatoire (1.62) peuvent aussi correspondre au cumul des prestations relatives à chaque assuré, et dans ce cas, n désigne le nombre d'assurés en portefeuille. Une compagnie de réassurance qui réassure la même ligne d'affaire pour n compagnies d'assurance doit également étudier la distribution de probabilité d'un vecteur aléatoire (1.62). Une modélisation classique des risques consiste souvent à considérer que les composantes du vecteur aléatoire (1.62) sont mutuellement indépendantes. Cette hypothèse d'indépendance est remise en question par l'observation de corrélations des sinistralités sur différentes branches d'activité, entre les assurés ou sur la même ligne d'affaire de différentes compagnies d'assurance.

Un évènement climatique peut engendrer des dégâts matériels sur les habitations des particuliers pris en charge par la branche habitation mais aussi sur des entrepôts ou des usines appartenant à des professionnels dont la gestion est effectuée par la branche professionnelle. Un accident de voiture engendre des dégâts matériels, pris en charge par la branche automobile, et des dommages corporels qui relèvent de la branche responsabilité civile. Dans le cadre d'un contrat couvrant les accidents du travail, la compagnie d'assurance est amenée à verser des prestations couvrant les jours d'arrêts de travail du salarié et les éventuelles dépenses médicales. Ces deux derniers exemples sont reportés dans le papier de CUMMINS et WILT BANK [1983]. Ce papier est le premier à évoquer la possibilité d'opter pour des modèles de montants agrégés de sinistres multivariés. L'étude statistique, produite dans CUMMINS et WILT BANK [1983], met en évidence l'existence de corrélations et quantifie leur importance. Le montant des prestations impactant plusieurs portefeuilles suite à un évènement dépend intuitivement de l'amplitude de cet évènement. Dans le domaine de la réassurance, un réassureur qui réassure la même branche d'activité pour différentes compagnies d'assurance est aussi confronté au problème de la corrélation car un évènement climatique localisé géographiquement engendre des sinistres importants pour plusieurs assureurs qui déclenchent le versement de plusieurs prestations pour le réassureur, dont les montants sont corrélés positivement.

Dans le souci d'affiner la gestion des risques, il est pertinent de modéliser la dépendance entre les risques en proposant des modèles multivariés. Les composantes du vecteur (1.62) représentent des sommes agrégées de prestations et sont par conséquent des variables aléatoires d'une forme similaire à (1.1). Le travail d'AMBAGASPITIYA [1999] met en exergue les difficultés liées à la modélisation du vecteur (1.62). Trois modèles se distinguent actuellement.

Le modèle 1 suppose un même nombre de sinistres pour les différentes composantes et l'existence d'une corrélation entre les montants associés aux sinistres.

Modèle 1. *Le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_n)$ admet une distribution composée multi-*

variée $(\mathbb{P}_N, \mathbb{P}_U)$ suivant le modèle 1 s'il est de la forme

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ \vdots \\ U_{nj} \end{pmatrix}, \quad (1.63)$$

où N est une variable aléatoire de comptage de loi \mathbb{P}_N et $\{U_j\}_{j \in \mathbb{N}}$ une suite de vecteurs aléatoires *i.i.d.* suivant une loi de probabilité \mathbb{P}_U . La loi de U est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n , sa densité jointe est notée f_U . La variable aléatoire N est indépendante de la suite de vecteurs aléatoires $\{U_j\}_{j \in \mathbb{N}}$.

Pour ce modèle, il est montré dans le papier de [AMBAGASPITIYA \[1999\]](#), que les méthodes récursives en dimension 1 se transposent pour approcher la loi du vecteur aléatoire (1.63). Le papier de [SUNDT \[1999\]](#) développe d'ailleurs une extension multivariée de l'algorithme de Panjer adaptée à ce modèle.

Le modèle 2 suppose une corrélation sur la fréquence des sinistres affectant les différentes composantes. Les montants unitaires des sinistres sur les différentes composantes sont indépendants.

Modèle 2. Le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_n)$ admet une distribution composée multivariée $(\mathbb{P}_N, [\mathbb{P}_{U_1}, \dots, \mathbb{P}_{U_n}])$ suivant le modèle 2 s'il est de la forme

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N_1} U_{1j} \\ \vdots \\ \sum_{j=1}^{N_n} U_{nj} \end{pmatrix}, \quad (1.64)$$

où $\mathbf{N} = (N_1, \dots, N_n)$ est un vecteur aléatoire de comptage de loi discrète multivariée \mathbb{P}_N , la suite $\{U_{ij}\}_{j \in \mathbb{N}}$ est une suite de variables aléatoires *i.i.d.* de loi \mathbb{P}_{U_i} qui modélise le montant des sinistres pour la composante i avec $i = 1, \dots, n$. Le montant des sinistres sur les différentes composantes sont mutuellement indépendants et indépendants du nombre de sinistres.

Le modèle 2 est étudié en dimension 2 dans les papiers de [HESSELAGER \[1996\]](#) et [VERNIC \[1999\]](#). Les vecteurs aléatoires de comptage $\mathbf{N} = (N_1, N_2)$ considérés dans ces deux papiers admettent des lois de probabilité caractérisées par des relations de récurrence qui permettent la mise en place d'algorithmes récursifs pour déterminer les probabilités du vecteur $\mathbf{X} = (X_1, X_2)$ proches de ceux utilisés en dimension 1. Le papier de [SUNDT \[2000\]](#) étend le travail de [VERNIC \[1999\]](#) en dimension supérieure à 2.

Récemment, un modèle mêlant les modèles 1 et 2 a été proposé.

Modèle 3. Le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_n)$ admet une distribution composée multivariée $(\mathbb{P}_{N,M}, [\mathbb{P}_{U_1}, \dots, \mathbb{P}_{U_n}, \mathbb{P}_V])$ suivant le modèle 3 s'il est de la forme

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^{N_1} U_{1j} \\ \vdots \\ \sum_{j=1}^{N_n} U_{nj} \end{pmatrix} + \sum_{j=1}^M \begin{pmatrix} V_{1j} \\ \vdots \\ V_{nj} \end{pmatrix}, \quad (1.65)$$

où $(\mathbf{N}, \mathbf{M}) = (N_1, \dots, N_n, M)$ est un vecteur aléatoire de comptage de loi discrète multivariée $\mathbb{P}_{\mathbf{N}, \mathbf{M}}$, la suite $\{U_{ij}\}_{j \in \mathbb{N}}$ est une suite de variables aléatoires *i.i.d.* suivant une loi \mathbb{P}_{U_i} qui représente le montant des sinistres affectant spécifiquement la $i^{\text{ème}}$ composante, et $\{\mathbf{V}_j\}_{j \in \mathbb{N}}$ est une suite de vecteurs aléatoires *i.i.d.* suivant une loi de probabilité $\mathbb{P}_{\mathbf{V}}$ qui modélise les sinistres qui affectent simultanément toutes les composantes. La loi de \mathbf{V} est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n , sa densité jointe est notée $f_{\mathbf{V}}$. Le vecteur aléatoire (\mathbf{N}, \mathbf{M}) , la suite de vecteurs aléatoires $\{\mathbf{V}_j\}_{j \in \mathbb{N}}$ et les suites de variables aléatoires $\{U_{ij}\}_{j \in \mathbb{N}}$ sont supposés mutuellement indépendants.

Le modèle 3 est une généralisation en dimension supérieure à 2 du modèle proposé dans le papier de JIN et REN [2014]. Ce papier propose une méthode récursive pour approcher la loi de probabilité de \mathbf{X} .

Comme en dimension 1, les méthodes récursives permettent une évaluation exacte des probabilités associées aux distributions composées multivariées dans le cas où les sinistres admettent des lois de probabilité discrètes. Lorsque les montants associés aux sinistres sont modélisés par des lois continues, alors les algorithmes récursifs fournissent des approximations suite à la discrétisation des lois de probabilités via les méthodes évoquées dans la Section 1.1.2. L'erreur de discrétisation se propage de manière plus importante en dimension supérieure à 1. Pour un même niveau de précision, il est nécessaire d'opter pour un pas de discrétisation plus petit. Cela va de paire avec une augmentation considérable des temps de calculs. Il faut ajouter à cela la lourdeur des expressions engagées dans les algorithmes récursifs en dimension supérieure à 1. Il est raisonnable de considérer des méthodes numériques d'inversion de la transformée de Laplace pour concurrencer les méthodes récursives. L'application de méthodes numériques d'inversion de la transformée de Laplace est rendue possible par la définition des transformées de Laplace multivariées ou des [Fonction Génératrice des Moments Multivariée \(f.g.m.m.\)](#) pour les vecteurs aléatoires (1.63), (1.64), et (1.65).

Proposition 4. 1. La *f.g.m.m.* du vecteur \mathbf{X} dans le modèle 1 est donnée par

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) = \mathcal{G}_{\mathbf{N}}[\mathcal{L}_{\mathbf{U}}(\mathbf{s})], \quad (1.66)$$

où $s \mapsto \mathcal{G}_{\mathbf{N}}(s)$ est la *f.g.p.* de la variable aléatoire de comptage \mathbf{N} et $\mathbf{s} \mapsto \mathcal{L}_{\mathbf{U}}(\mathbf{s})$ est la *f.g.m.m.* du vecteur aléatoire \mathbf{U} .

2. La *f.g.m.m.* du vecteur \mathbf{X} dans le modèle 2 est donnée par

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) = \mathcal{G}_{\mathbf{N}}[\mathcal{L}_{U_1}(s_1), \dots, \mathcal{L}_{U_n}(s_n)], \quad (1.67)$$

où $\mathbf{s} \mapsto \mathcal{G}_{\mathbf{N}}(\mathbf{s})$ est la *Fonction Génératrice des Probabilités Multivariée (f.g.p.m.)* du vecteur aléatoire de comptage \mathbf{N} et $s \mapsto \mathcal{L}_{U_i}(s)$ est la *f.g.m.* de la variable aléatoire U_i pour $i = 1, \dots, n$.

3. La *f.g.m.m.* du vecteur \mathbf{X} dans le modèle 3 est donnée par

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) = \mathcal{G}_{\mathbf{N}, \mathbf{M}}[\mathcal{L}_{U_1}(s_1), \dots, \mathcal{L}_{U_n}(s_n), \mathcal{L}_{\mathbf{V}}(\mathbf{s})], \quad (1.68)$$

où $\mathbf{s} \mapsto \mathcal{G}_{\mathbf{N}, \mathbf{M}}(\mathbf{s})$ est la *f.g.p.m.* du vecteur aléatoire de comptage $(\mathbf{N}, \mathbf{M}) = (N_1, \dots, N_n, M)$, $s \mapsto \mathcal{L}_{U_i}(s)$ est la *f.g.m.* de la variable aléatoire U_i pour $i = 1, \dots, n$ et $\mathbf{s} \mapsto \mathcal{L}_{\mathbf{V}}(\mathbf{s})$ est la *f.g.m.m.* du vecteur aléatoire \mathbf{V} .

L'algorithme FFT se généralise naturellement en dimension supérieure à 1. Le papier de JIN et REN [2010] applique cette généralisation sur le modèle 2 en dimension 2. Les mêmes hypothèses que dans HESSELAGER [1996] sont adoptées concernant le vecteur \mathbf{N} . Le papier JIN et REN [2014] applique cette généralisation au modèle 3 en dimension 2. Dans chacun des deux papiers, une analyse comparative des performances relatives aux algorithmes récursifs et à l'algorithme FFT est menée. L'algorithme FFT multivarié est toujours sujet à l'erreur d'aliasing, ainsi qu'à l'erreur de discrétisation. Le débat sur le choix de la méthode est sensiblement le même qu'en dimension 1. Le problème des algorithmes récursifs et de l'algorithme FFT demeure la nécessité de discrétiser préalablement la loi de probabilité régissant le montant des prestations.

L'extension des méthodes numériques d'inversion de la transformée de Laplace aux dimensions supérieures à 1 représente donc une alternative valable. La méthode basée sur l'inversion de la transformée de Laplace via les moments exponentiels est étendue en dimension 2 dans MNATSAKANOV [2011]. La méthode n'est cependant pas illustrée sur les distributions composées multivariées. L'extension multivariée de la méthode d'inversion de la transformée de Laplace via les moments fractionnels est simplement évoquée dans GZYL et TAGLIANI [2012], aucun papier ne traite explicitement du sujet. La méthode d'inversion de la transformée de Fourier est étendue aux dimensions supérieures à 1 dans CHOUDHURY et collab. [1994] avec une application en théorie des files d'attente. La méthode d'approximation via une représentation polynomiale est étendue en dimension supérieure $n > 1$ dans le Chapitre 2, et appliquée concrètement sur un modèle proche du modèle 3 en dimension 2 dans GOFFARD et collab. [2015a] qui constitue le Chapitre 4 de ce manuscrit.

Les algorithmes récursifs, l'algorithme FFT et les méthodes numériques d'inversion de la transformée de Laplace donnent accès à la densité de probabilité multivariée, à la f.d.r.m. et à la f.d.s.m. du vecteur aléatoire (1.62). Dans le travail de AMBAGASPITIYA [1998], l'accent est mis sur la distribution de la somme des charges totales qui compose le vecteur 1.62, avec

$$X = \sum_{i=1}^n X_i. \quad (1.69)$$

L'importance de l'inclusion d'une hypothèse de corrélation dans l'étude de la variable aléatoire (1.69) est relativisée lorsque l'une des composantes domine les autres. Ce qui peut être le cas, par exemple en assurance auto, où les dommages corporels sont souvent plus coûteux que les dégâts matériels. De plus, le problème se résume à un problème univarié dans lequel il n'est pas toujours nécessaire de connaître la loi jointe pour étudier la loi de probabilité de la variable aléatoire (1.69).

Dans le cas d'un réassureur qui propose un traité global de réassurance non proportionnelle pour la même branche d'activité de plusieurs compagnies d'assurance concurrentes, la quantification de l'exposition au risque du réassureur passe par le calcul de la f.d.s. du coût de l'opération de réassurance. Les portefeuilles de contrats réassurés admettent des montants agrégés de prestations $\mathbf{X} = (X_1, \dots, X_n)$. Le réassureur propose aux compagnies un traité de réassurance non proportionnelle de priorités $\mathbf{b} = (b_1, \dots, b_n)$ et de portées $\mathbf{c} = (c_1, \dots, c_n)$. Le coût total de l'opération de réassurance

est donné par

$$Z = \sum_{i=1}^n \max[(X_i - b_i)_+, c_i]. \quad (1.70)$$

Le réassureur souhaite avoir accès aux quantiles de la distribution de Z pour définir ses marges de solvabilité sur la période d'exercice. Le calcul de la f.d.s. de la variable aléatoire Z nécessite la connaissance de la loi de probabilité jointe du vecteur aléatoire \mathbf{X} .

Les modèles collectifs univariés et multivariés, qui ont fait l'objet des sections 1.1 et 1.2, adoptent un point de vue statique, la section suivante traite d'un modèle qui incorpore un caractère dynamique et modélise l'équilibre financier d'un assureur à plus long terme. Le portefeuille de contrats est toujours considéré comme un collectif de risques que les sinistres viennent choquer. La différence est que l'instant d'arrivée de ces chocs revêt une importance particulière.

1.3 Théorie de la ruine : L'approximation de la probabilité de ruine ultime dans le modèle de ruine de Poisson composée

1.3.1 Présentation du modèle de ruine et définition de la probabilité de ruine ultime

La théorie de la ruine consiste en la modélisation de la richesse d'une compagnie d'assurance ou de la réserve financière allouée à l'une de ses branches d'activité. Cette modélisation est adaptée à la gestion des risques des compagnies d'assurance non-vie. La compagnie d'assurance est supposée capable de suivre l'évolution de sa richesse continuellement dans le temps. L'arrivée des sinistres au cours du temps est modélisée via un processus de comptage $\{N_t\}_{t \geq 0}$ égale au nombre de sinistres décomptés jusqu'à l'instant t . Les montants de sinistre sont modélisés par une suite de variables aléatoires $\{U_i\}_{i \in \mathbb{N}}$ strictement positives, i.i.d. suivant une loi de probabilité \mathbb{P}_U et indépendantes de $\{N_t\}_{t \geq 0}$. Un niveau de richesse initial $u \geq 0$ et un niveau de prime périodique $c \geq 0$ sont définis pour aboutir au modèle suivant :

Modèle 4. *Le processus de richesse ou de réserve financière est donné par*

$$R_t = u + ct - \sum_{i=1}^{N_t} U_i. \quad (1.71)$$

Le processus d'excédent de sinistres est défini par

$$S_t = u - R_t. \quad (1.72)$$

La Figure 1.1 offre une visualisation graphique de l'évolution des processus de richesse et d'excédent de sinistres. Cette modélisation est dynamique par opposition à la modélisation statique présentée dans les Sections 1.1 et 1.2. En théorie de la ruine,

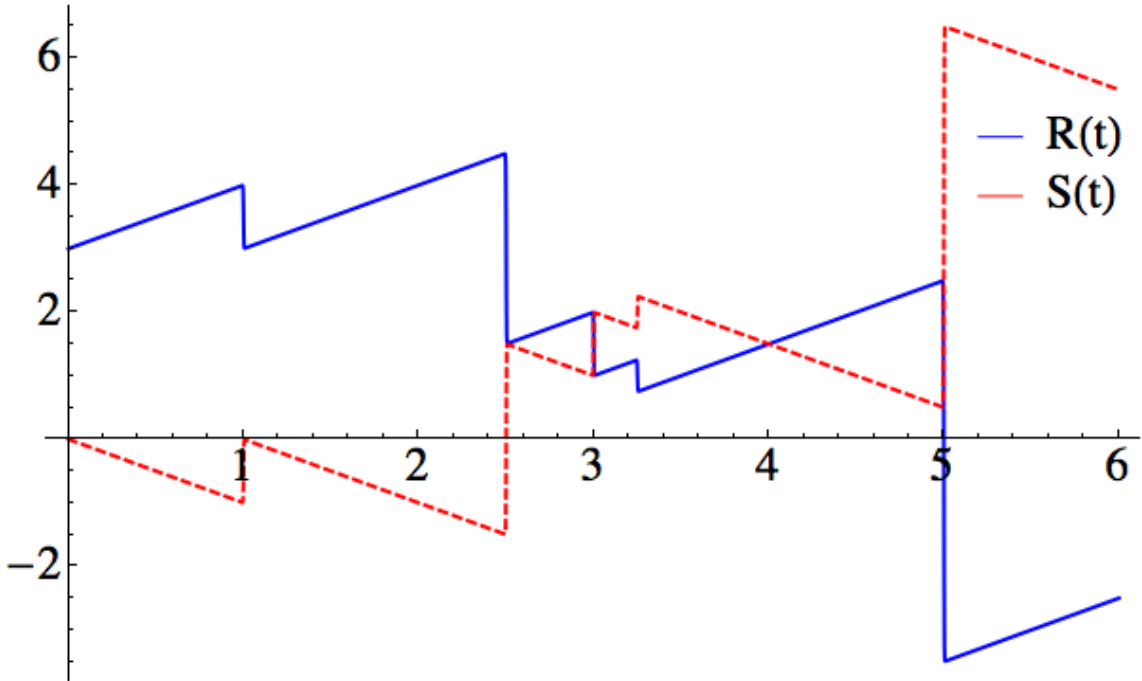


FIGURE 1.1 – Évolution des processus de richesse et d'excédent de sinistres au cours du temps.

l'objectif est d'évaluer la probabilité que le processus de richesse passe en dessous de 0, en fonction de son point de départ u . Cette situation correspond à une situation critique dans laquelle la compagnie n'est plus en mesure de remplir ses engagements vis-à-vis de ses clients, de ses actionnaires ou d'une autorité de contrôle.

Définition 8. La probabilité de ruine à horizon de temps fini T , de réserve initiale u est définie par

$$\psi(u, T) = \mathbb{P} \left(\inf_{t \in [0, T]} R_t < 0 \mid R_0 = u \right). \quad (1.73)$$

La probabilité (1.73) correspond à la probabilité que les réserves deviennent strictement négatives à un instant précédant l'horizon T . La probabilité de ruine ultime ou à horizon de temps infini est définie par

$$\psi(u) = \mathbb{P} \left(\inf_{t \geq 0} R_t < 0 \mid R_0 = u \right). \quad (1.74)$$

L'instant où la réserve passe en dessous de 0 coïncide avec l'instant où le processus d'excédent de sinistres passe au-dessus de u , cette remarque permet une définition alternative des probabilités de ruine avec

$$\psi(u, T) = \mathbb{P} \left(\sup_{t \in [0, T]} S_t > u \mid S_0 = 0 \right),$$

et

$$\psi(u) = \mathbb{P} \left(\sup_{t \geq 0} S_t > u \mid S_0 = 0 \right).$$

La probabilité de non-ruine est définie à horizon de temps fini et infini par

$$\phi(u, T) = 1 - \psi(u, T),$$

et

$$\phi(u) = 1 - \psi(u).$$

Pour une revue exhaustive des résultats de la théorie de la ruine, le lecteur peut se référer par exemple aux ouvrages de [GRANDELL \[1991\]](#), [ROLSKI et collab. \[1999\]](#), et [ASMUSSEN et ALBRECHER \[2010\]](#).

Il convient d'insister sur le caractère extrêmement simplifié du modèle 4, la probabilité de ruine calculée dans le cadre de ce modèle s'envisage comme une mesure de risque qui va permettre de juger de la pertinence de la tarification de l'assureur ou encore du niveau de fonds propres à allouer au vue de la sinistralité. Le modèle 4 fait office de base à l'élaboration de modèles plus complexes qui incorporent des taux d'inflation, une prise en compte de l'évolution de la sinistralité ou du niveau des primes périodiques. L'étude des modèles de ruine est aussi rendue intéressantes par leur lien avec d'autres champs d'application des probabilités. En effet, les modèles de ruine tels que le Modèle 4 apparaissent dans d'autres domaines d'application des probabilités que l'actuariat. La théorie des files d'attente étudie les solutions optimales de gestion des files d'attente créées par l'accumulation de clients à des guichets. L'équivalent du Modèle 4 en théorie des files d'attente est noté M/G/1. Le "M" indique que l'arrivée des clients est gouvernée par un processus ponctuel (Markovien), le "G" indique que les temps d'occupation du guichet par les clients forment une suite de variables aléatoires réelles, positives, et *i.i.d.* suivant une loi de probabilité qui n'est pas spécifiée (Générale), et le "1" indique que la file d'attente modélise l'arrivée des clients sur un seul guichet. La théorie des files d'attente s'intéresse à des quantités telles que le temps nécessaire au traitement de toutes les demandes en attente ou encore le temps d'attente du client n . Historiquement, ces modèles ont été introduits pour étudier l'accumulation de l'eau retenue par des barrages. L'utilisation actuelle est plutôt l'optimisation de la gestion du temps de traitement des données par des serveurs informatiques, pour plus de détails voir le livre d'[ASMUSSEN \[2010\]](#). Récemment, les modèles de ruine et de files d'attente ont été employés dans l'étude de l'évolution de la quantité de contaminants alimentaires présents dans l'organisme. Le modèle en question a été baptisé [Kinetic Dietary Exposure Model \(KDEM\)](#). L'ingestion de contaminants par un individu est un évènement supposé ponctuel et aléatoire, l'arrivée des contaminants dans l'organisme est par conséquent modélisée via un processus stochastique de comptage. L'augmentation de la quantité de contaminants dans le corps à chaque ingestion est aléatoire et modélisée par une variable aléatoire réelle, et continue. L'accumulation de contaminants est un phénomène identique à l'arrivée de clients à un guichet comme l'explique le papier de [BERTAIL et collab. \[2008\]](#). Une fonction d'élimination des contaminants au cours du temps est définie, et il en résulte un processus de contamination alimentaire dont l'évolution est inverse à celle de la richesse d'une compagnie d'assurance suivant le modèle 4, voir [BERTAIL et LOISEL \[2014\]](#).

Le travail effectué dans cette thèse se focalise sur l'évaluation de la probabilité de ruine ultime (1.74) dans un cas particulier du modèle de ruine 1.71, appelé modèle de

ruine de Poisson composé. Dans le cadre de ce modèle, le processus d'arrivée des sinistres est un processus de Poisson simple d'intensité λ . Le calcul des probabilités de ruine s'effectue à niveau de prime périodique fixé. L'importance de ce paramètre est à nuancer. La probabilité de ruine est une quantité réaliste lorsqu'elle est faible. Les probabilités de ruine calculées sont associées à des niveaux de réserves initiales très importants relativement au niveau des primes. Le niveau des primes c est contraint par le marché de l'assurance fortement concurrentiel. Une prime périodique élevée entraînerait mécaniquement la fuite des assurés vers les compagnies d'assurance moins chères. Il semble tout de même logique de prélever aux assurés un peu plus que ce qu'ils seront indemnisé en moyenne afin d'avoir une activité rentable. Dans le modèle de ruine de Poisson composé, les assurés coûtent en moyenne

$$\lim_{t \rightarrow +\infty} \frac{1}{t} \sum_{i=1}^{N_t} U_i = \lambda E(U), \quad (1.75)$$

par unité de temps. La limite (1.75) est une conséquence de la loi des grands nombres. Le niveau de prime est donné par

$$c = (1 + \eta)\lambda E(U), \quad (1.76)$$

où η est généralement exprimé en pourcentage, et indique de combien la prime doit excéder le coût moyen des sinistres par unité de temps. Une activité rentable est caractérisée par la condition $\eta > 0$, appelée *Net Benefit Condition*, et justifiée par le résultat suivant :

Proposition 5. *Si $\eta > 0$, alors*

$$\lim_{t \rightarrow +\infty} R_t = +\infty, \quad (1.77)$$

$\psi(u) < 1$ et l'activité est dite rentable. Si $\eta < 0$, alors

$$\lim_{t \rightarrow +\infty} R_t = -\infty, \quad (1.78)$$

$\psi(u) = 1$ et l'activité n'est pas rentable.

Il existe un lien entre les distributions composées et la probabilité de ruine ultime dans le modèle de Poisson composé qui se matérialise à travers la *formule de Pollaczek-Khinchine*.

Theorème 3 (Pollaczek & Khinchine). *Dans le modèle de ruine de Poisson composé, la probabilité de ruine ultime est égale à la f.d.s. d'une distribution géométrique composée ($\mathbb{P}_N, \mathbb{P}_V$). Soit $X = \sup_{t \geq 0} S_t$ alors*

$$\psi(u) = \mathbb{P}(X > u) = (1 - p) \sum_{i=1}^{+\infty} p^i \bar{F}_V^{(*n)}(u), \quad (1.79)$$

avec

$$X = \sum_{i=1}^N V_i, \quad (1.80)$$

où N est une variable aléatoire de comptage de loi géométrique $\mathcal{NB}(1, p)$ de paramètre $p = \frac{\lambda \mathbb{E}(U)}{c}$, et $\{V_i\}_{i \in \mathbb{N}}$ est une suite de variables aléatoires, positives, *i.i.d.* suivant une loi \mathbb{P}_V et indépendantes de N . La distribution de probabilité \mathbb{P}_V est appelée *integrated tail distribution* de la distribution \mathbb{P}_U . La densité de probabilité de \mathbb{P}_V est donnée par

$$f_V(x) = \frac{\bar{F}_U(x)}{\mathbb{E}(U)}. \quad (1.81)$$

La transformée de Laplace de la probabilité de ruine est la transformée de Laplace de la *f.d.s.* de la variable aléatoire (1.80), soit

$$\mathcal{L}_\Psi(s) = \frac{1}{s} (\mathcal{L}_X(s) - 1), \quad (1.82)$$

avec

$$\mathcal{L}_X(s) = \frac{1-p}{1-p\mathcal{L}_V(s)}, \quad (1.83)$$

et

$$\mathcal{L}_V(s) = \frac{1}{s\mathbb{E}(U)} (\mathcal{L}_U(s) - 1). \quad (1.84)$$

La présence de la série infinie dans la *formule de Pollaczek-Khinchine* (1.79) la rend difficilement exploitable pour procéder au calcul de la probabilité de ruine ultime. Celle-ci admet une formule fermée dans un nombre très restreint de cas particuliers, la plupart étant englobés par le résultat suivant.

Theorème 4. *Si le montant des sinistres est distribué suivant une loi de probabilité phase-type $\mathcal{PT}(E, \alpha, \mathbf{T})$, alors*

$$\psi(u) = \alpha_+ \exp[(\mathbf{T} + \mathbf{t}\alpha_+)u], \quad (1.85)$$

où $\alpha_+ = -\beta\alpha\mathbf{T}^{-1}$.

Ce résultat est démontré dans le chapitre 8 du livre d'ASMUSSEN et ALBRECHER [2010]. Ainsi, une formule fermée est disponible pour calculer la probabilité de ruine ultime dans le modèle de ruine de Poisson composé lorsque les sinistres sont distribués suivant une loi *phase-type*.

Définition 9. *Soit X une variable aléatoire de loi phase-type $\mathcal{PT}(E, \alpha, \mathbf{T})$. X représente le temps d'absorption d'un processus de Markov absorbant d'espace d'état E , de générateur \mathbf{T} , et de loi initiale α . La densité de X est donnée par*

$$f_X(x) = \alpha e^{\mathbf{T}x} \mathbf{t}, \quad x \in \mathbb{R}_+, \quad (1.86)$$

et sa *f.g.m.* par

$$\mathcal{L}_X(s) = \alpha (-s\mathbf{I} - \mathbf{T})^{-1} \mathbf{t}, \quad (1.87)$$

où \mathbf{I} est la matrice identité et \mathbf{t} vérifie $\mathbf{t} = \mathbf{T}\mathbf{e}$, avec $\mathbf{e} = (1, \dots, 1)$.

Un résultat fondamental en théorie de la ruine est l'encadrement de la probabilité de ruine par des fonctions exponentielles décroissantes. Ce résultat fondamental découle du Théorème 5, relatif aux distributions géométriques composées.

Théorème 5. Si X admet une distribution géométrique composée $[\mathcal{N} \mathcal{B}(1, p), \mathbb{P}_U]$ telle que l'équation

$$\mathcal{L}_U(s) = \frac{1}{p} \quad (1.88)$$

admette une unique solution positive, notée γ , alors

$$a_- e^{-\gamma x} < \bar{F}_X(x) < a_+ e^{-\gamma x}, \quad (1.89)$$

avec

$$a_- = \inf_{x \in \text{Supp } U} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^{+\infty} e^{\gamma x} \bar{F}_U(x) dx}, \quad a_+ = \sup_{x \in \text{Supp } U} \frac{e^{\gamma x} \bar{F}_U(x)}{\int_x^{+\infty} e^{\gamma x} \bar{F}_U(x) dx}. \quad (1.90)$$

Une démonstration du Théorème 5 est donnée au chapitre 4 de l'ouvrage de **ROLSKI et collab.** [1999]. L'application du Théorème 5 permet d'établir le résultat suivant :

Corollaire 1. Dans le modèle de ruine de Poisson composé, la probabilité de ruine vérifie

$$a_- e^{-\gamma u} < \psi(u) < a_+ e^{-\gamma u}, \quad (1.91)$$

où

$$a_- = \inf_{x \in \text{Supp } V} \frac{e^{\gamma x} \bar{F}_V(x)}{\int_x^{+\infty} e^{\gamma x} \bar{F}_V(x) dx}, \quad a_+ = \sup_{x \in \text{Supp } V} \frac{e^{\gamma x} \bar{F}_V(x)}{\int_x^{+\infty} e^{\gamma x} \bar{F}_V(x) dx}. \quad (1.92)$$

et γ est l'unique solution positive de l'équation

$$\lambda (\mathcal{L}_U(s) - 1) = cs, \quad (1.93)$$

appelée équation fondamentale de Cramer-Lundberg.

L'existence d'une solution $\gamma > 0$ pour l'équation (1.93) implique que la **f.g.m.** de la variable aléatoire U est définie sur $[0, \gamma]$. Ce résultat n'est pas valable si la **f.g.m.** de U n'est pas définie lorsqu'elle prend en argument des valeurs positives. Cette situation est typique d'une modélisation du montant des sinistres via une loi de probabilité à queue lourde. Les distributions *phase-type* sont d'ailleurs des distributions de probabilité à queue légère. Le travail effectué ici, se focalise sur l'approximation de la probabilité de ruine lorsque les montants de sinistres sont modélisés par des lois de probabilité à queue légère associées à des transformées de Laplace bien définies. Il s'agit d'une condition nécessaire à l'application des méthodes numériques d'inversion de la transformée de Laplace.

1.3.2 Approximation de la probabilité de ruine ultime : Revue de la littérature

Le constat selon lequel la probabilité de ruine ultime n'est accessible que dans un nombre de cas restreint motive la mise au point de méthodes numériques pour l'évaluer. La première approximation proposée est celle de *Cramer-Lundberg*, elle est directement inspirée de l'encadrement (1.91). La probabilité de ruine ultime est approchée par

$$\psi(u) = \frac{c - \mathbb{E}(U)}{\lambda \mathcal{L}_U(\gamma) - c} e^{-\gamma u}. \quad (1.94)$$

Elle se déduit de l'étude de la limite

$$\lim_{u \rightarrow +\infty} e^{Y u} \psi(u) = \frac{c - \mathbb{E}(U)}{\lambda \mathcal{L}_U(\gamma) - c}. \quad (1.95)$$

Elle est performante pour les grandes valeurs de réserves initiales.

L'idée de BOWERS [1966] d'approcher la densité de probabilité par des sommes de densités gamma a inspiré le travail de BEEKMAN [1969], duquel a résulté l'*approximation de Beekman-Bowers* pour la probabilité de ruine ultime. Cette approximation consiste à approcher la loi de probabilité de la variable aléatoire Z dont la loi est caractérisée par sa f.d.r.

$$F_Z(u) = 1 - \frac{\psi(u)}{p},$$

par une loi gamma $\Gamma(\alpha, \beta)$.

Définition 10. Soit X une variable aléatoire de loi gamma $\Gamma(\alpha, \beta)$, de paramètre de forme α et de paramètre d'échelle β . La densité de la variable aléatoire X est donnée par

$$f_X(x) = \frac{e^{-x/\beta} x^{\alpha-1}}{\Gamma(\alpha) \beta^\alpha}, \quad x \in \mathbb{R}^+, \quad (1.96)$$

et sa f.g.m. par

$$\mathcal{L}_X(s) = \left(\frac{1}{1 - s\beta} \right)^\alpha. \quad (1.97)$$

A noter que le cas particulier $\alpha \in \mathbb{N}$ correspond à la loi de Erlang et que le cas très particulier $\alpha = 1$ correspond à la loi exponentielle.

L'identification des paramètres fait coïncider les moments d'ordre 1 et 2 de la loi gamma avec ceux de la variable aléatoire Z .

Le Théorème 4 indique que la probabilité de ruine ultime est donnée explicitement par (1.85) lorsque le montant des sinistres est distribué suivant une loi *phase-type*. La loi exponentielle est un cas particulier des lois *phase type*, la probabilité de ruine ultime lorsque le montant des sinistres suit une loi exponentielle admet une formule fermée. L'approximation de DE VYLDER [1978] consiste à définir un nouveau processus de réserve financière avec une loi exponentielle $\Gamma(1, \beta)$ pour modéliser la sévérité des sinistres. L'identification des paramètres du nouveau processus de richesse, à savoir l'intensité du processus de Poisson, le niveau de prime périodique, et le paramètre β de la loi du montant des sinistres assure la correspondance des trois premiers moments du nouveau processus de ruine avec ceux de l'ancien. La probabilité de ruine est approchée par la probabilité de ruine ultime associée au nouveau processus de réserve par la formule (1.85). Le papier de GRANDELL [2000] propose une étude comparative dans laquelle l'approximation de DE VYLDER [1978] se distingue de ses concurrentes.

En effet, l'idée sous-tendue par cette approche est le remplacement de la loi pour le montant de sinistres par une loi permettant le calcul de la probabilité de ruine via une formule fermée. Cette idée a inspiré bon nombre de méthodes d'approximation.

Les distributions *phase-type* englobent beaucoup de lois de probabilité à support dans \mathbb{R}^+ comme les mélanges de lois exponentielles et de Erlang. La procédure de **DE VYLDER** [1978] est utilisée, en prenant comme substitut à la loi de probabilité du montant des sinistres, un mélange de lois exponentielles à deux composantes ou une Erlang de paramètre de forme $\alpha = 2$ dans le papier de **RAMSAY** [1992] et une loi gamma plus générale dans le papier de **BURNECKI et collab.** [2005]. Cette idée a été approfondie dans le travail de **AVRAM et collab.** [2011]. La loi de probabilité du montant des sinistres est approchée par un mélange fini de lois exponentielles. Les paramètres du mélange s'obtiennent en résolvant un système d'équation, dont la complexité est fonction du nombre de composantes dans le mélange. La résolution du système est facilitée par l'utilisation d'une approximation de Padé pour la *f.g.m.* de la variable aléatoire U sous la forme d'une fonction fractionnelle. Une fois les paramètres identifiés, la loi du montant des sinistres est remplacée par la loi issue du mélange fini de lois exponentielles et la probabilité de ruine ultime est approchée par la formule (1.85). Cette méthode est intéressante car il s'agit d'une procédure d'inversion de la transformée de Laplace basée sur l'étude des moments de la distribution.

Le Théorème 3 rend possible l'application de l'algorithme de Panjer, car la loi géométrique est un cas particulier de la loi binomiale négative qui appartient à la famille de Panjer. La mise en application de l'algorithme de Panjer pour approcher la probabilité de ruine est effectuée dans le papier de **DICKSON** [1995]. D'autres récursions proches de la version continue de l'algorithme de Panjer sont étudiées dans les papiers de **DICKSON** [1989] et **GOOVAERTS et DE VYLDER** [1984]. L'expression de la transformée de Laplace de la probabilité de ruine (1.82) permet l'application de l'algorithme FFT, voir **EMBRECHTS et collab.** [1993]. L'algorithme de Panjer et l'algorithme FFT admettent les mêmes limites dans l'approximation de la probabilité de ruine ultime que dans le cas des distributions composées. Les méthodes d'inversion numérique de la transformée de Laplace évoquées dans la Section 1.1.3 ont aussi été utilisées dans la littérature pour approcher la probabilité de ruine ultime. La méthode des moments exponentiels est appliquée dans l'article de **MNATSAKOV et collab.** [2014] et celle des moments fractionnels dans le papier de **GZYL et collab.** [2013]. La méthode d'inversion numérique de la transformée de Fourier est appliquée dans le chapitre 5 du livre de **ROLSKI et collab.** [1999]. Récemment, **ALBRECHER et collab.** [2010] ont proposé une méthode d'inversion qui repose sur une quadrature fondée sur l'approximation rationnelle de la fonction exponentielle dans le plan complexe.

L'application de la méthode basée sur la représentation polynomiale à l'approximation de la probabilité de ruine fait l'objet du papier **GOFFARD et collab.** [2015b] et constitue le Chapitre 3 de ce manuscrit. Ce travail comprend une comparaison avec la méthode d'inversion de la transformée de Fourier, la méthode des moments exponentiels, et l'algorithme de Panjer. La méthode repose sur la combinaison de la loi gamma et des polynômes de Laguerre généralisés. Elle peut être vue comme le prolongement des travaux de **BOWERS** [1966] et **BEEKMAN** [1969]. La méthode polynomiale peut aussi être reliée aux travaux de **TAYLOR** [1978] et de **ALBRECHER et collab.** [2001], dans lesquels il est supposé que la densité de probabilité gouvernant le montant des sinistres admet un développement sous la forme d'une série infinie de densité gamma. Ce dé-

veloppement est valide si la densité de probabilité est analytique. L'objectif affiché par ces travaux était cependant d'obtenir des formules fermées pour la probabilité de ruine à horizon de temps fini et infini. Ces probabilités admettent aussi des développements en série de densité gamma dont les coefficients sont liés aux développements de la densité de probabilité des montants de sinistre. Des relations de récurrence sont obtenues après réinjection du développement de la densité du montant de sinistres dans les équations vérifiées par la probabilité de ruine. L'exploitation de ces relations de récurrence conduisent dans certains cas à des formules fermées. A noter TAYLOR [1978] traite du modèle de ruine de Poisson composé alors que ALBRECHER et collab. [2001] considèrent un modèle de ruine plus général qui comprend un taux d'intérêt, un taux d'inflation et un taux d'actualisation constants. La *formule de Picard-Lefèvre*, voir PICARD et LEFÈVRE [1997] et RUILLERE et LOISEL [2004], donne une formule exacte pour la probabilité de ruine à horizon de temps fini lorsque le montant des sinistres admet une loi discrète. Cette formule est basée sur l'exploitation des propriétés des polynômes d'Appel, qui sont orthogonaux, voir POMMERET [2001] par exemple.

Ce chapitre a présenté les défis que doivent relever les méthodes numériques d'approximation et en particulier la méthode polynomiale, décrite exhaustivement dans le Chapitre 2. Ce chapitre comprend aussi une illustration des performances de la méthode polynomiale sur l'exemple de la distribution de Poisson composée. Le Chapitre 3 traite de l'approximation polynomiale de la probabilité de ruine ultime dans le modèle de ruine de Poisson composé, et le Chapitre 4 étudie l'approximation polynomiale de fonctions associées aux distributions composées bivariées.

1.4 Références

- ABATE, J., G. CHOUDHURY et W. WHITT. 1995, «On the Laguerre method for numerically inverting Laplace transforms», *INFORMS Journal on Computing*, vol. 8, n° 4, p. 413–427. [17](#)
- ABATE, J. et W. WHITT. 1992, «The Fourier-series method for inverting transforms of probability distributions.», *Queueing Systems*, vol. 10, p. 5–88. [20](#)
- ALBRECHER, H., F. AVRAM et D. KORTSCHAK. 2010, «On the efficient evaluation of ruin probabilities for completely monotone claim distributions», *Journal of Computational and Applied Mathematics*, vol. 233, n° 10, p. 2724–2736. [33](#)
- ALBRECHER, H., J. TEUGELS et R. TICHY. 2001, «On a gamma series expansion for the time dependent probability of collective ruin», *Insurance : Mathematics and Economics*, , n° 29, p. 345–355. [33](#), [34](#)
- AMBAGASPITIYA, R. 1998, «On the distribution of a sum of correlated aggregate claims», *Insurance : Mathematics and Economics*, vol. 23, n° 1, p. 15–19. [25](#)
- AMBAGASPITIYA, R. 1999, «On the distributions of two classes of correlated aggregate claims», *Insurance : Mathematics and Economics*, vol. 24, n° 3, p. 301–308. [22](#), [23](#)

- ASMUSSEN, S. 2010, *Applied probability and queues*, vol. 51, Springer Science and Business Media. [28](#)
- ASMUSSEN, S. et H. ALBRECHER. 2010, *Ruin Probabilities, Advanced Series on Statistical Science and Applied Probability*, vol. 14, World Scientific. [28](#), [30](#)
- AVRAM, F., D. CHEDOM et A. HORVARTH. 2011, «On moments based Padé approximations of ruin probabilities», *Journal of Computational and Applied Mathematics*, vol. 235, n° 10, p. 321–3228. [33](#)
- BEEKMAN, J. 1969, «Ruin function approximation», *Transaction of Society of Actuaries*, vol. 21, n° 59 AB, p. 41–48. [32](#), [33](#)
- BERTAIL, P., S. CLÉMENÇON et J. TRESSOU. 2008, «A storage model with random release rate for modelling exposure to food contaminants», *Mathematical Biosciences and Engineering*, vol. 5, n° 1, p. 35. [28](#)
- BERTAIL, P. et S. LOISEL. 2014, *Théorie de la ruine et applications*, Approches Statistiques du risque, Journée d'étude statistique. [28](#)
- BOWERS, N. 1966, «Expansion of probability density functions as a sum of gamma densities with applications in risk theory», *Transaction of Society of Actuaries*, vol. 18, n° 52, p. 125–137. [17](#), [32](#), [33](#)
- BOWERS, N. L., H. U. GERBER, J. C. HICKMAN, D. A. JONES et C. J. NESBITT. 1997, *Actuarial Mathematics*, 2^e éd., The Society of Actuaries, Schaumburg, Illinois. [8](#)
- BÜHLMAN, H. 1984, «Numerical evaluation of the compound poisson distribution : recursion or fast fourier transform?», *Scandinavian Actuarial Journal*, , n° 2, p. 116–126. [16](#)
- BURNECKI, K., P. MISTA et A. WERON. 2005, «A new gamma type approximation of the ruin probability», *Acta Physica Polonica B*, vol. 36, n° 5, p. 1473. [33](#)
- CHOUDHURY, G., D. LUCANTONI et W. WHITT. 1994, «Multidimensional transform inversion with application to the transient M / G / 1 queue», *The Annals of Applied Probability*, vol. 4, n° 3, p. 719–740. [25](#)
- COURTOIS, C. et M. DENUIT. 2007, «Local moment matching and s-convex extrema», *Astin Bulletin*, vol. 37, n° 2, p. 387. [13](#)
- CUMMINS, J. et L. J. WILTBANK. 1983, «Estimating the total claims distribution using multivariate frequency and severity distributions», *Journal of Risk and Insurance*, p. 377–403. [22](#)
- DE PRIL, N. 1986, «Moments of a class of compound distributions», *Scandinavian Actuarial Journal*, , n° 2, p. 117–120. [14](#)
- DE VYLDER, F. 1978, «A practical solution to the problem of ultimate ruin probability», *Scandinavian Actuarial Journal*, , n° 2, p. 114–119. [32](#), [33](#)

- DHAENE, J., G. WILLMOT et B. SUNDT. 1999, «Recursions for distribution functions and stop-loss transforms», *Scandinavian Actuarial Journal*, , n° 1, p. 52–65. [14](#)
- DICKSON, D. C. M. 1989, «Recursive calculation of the probability and severity of ruin», *Insurance : Mathematics and Economics*, vol. 8, n° 2, p. 145–148. [33](#)
- DICKSON, D. C. M. 1995, «A review of Panjer’s recursion formula and it’s applications», *British Actuarial Journal*, vol. 1, n° 1, p. 107–124. [33](#)
- EMBRECHTS, P. et M. FREI. 2009, «Panjer recursion versus FFT for compound distribution», *Mathematical Methods of Operations Research*, vol. 69, p. 497–508. [16](#)
- EMBRECHTS, P., R. GRÜBEL et S. M. PITTS. 1993, «Some applications of the fast fourier transform algorithm in insurance mathematics», *Statistica Neerlandica*, vol. 47, n° 1, p. 59–75. [33](#)
- GERBER, H. U. 1982, «On the numerical evaluation of the distribution of aggregate claims and its stop-loss premiums», *Insurance : Mathematics and Economics*, vol. 1, n° 1, p. 13–18. [13](#)
- GERBER, H. U. et D. JONES. 1976, «Some practical considerations in connection with the calculation of stop-loss premiums», *Transaction of the Society of Actuary*, vol. 28, p. 215–231. [13](#)
- GOFFARD, P.-O., S. LOISEL et D. POMMERET. 2015a, «Polynomial approximations for bivariate aggregate claim amount probability distributions», *Working Paper*. [25](#)
- GOFFARD, P.-O., S. LOISEL et D. POMMERET. 2015b, «A polynomial expansion to approximate the ultimate ruin probability in the compound Poisson ruin model», *Journal of Computational and Applied Mathematics*. [33](#)
- GOOVAERTS, M. et F. DE VYLDER. 1984, «A stable recursive algorithm for evaluation of ultimate ruin probabilities», *Astin Bulletin*, vol. 14, n° 1, p. 53–59. [33](#)
- GRANDELL, J. 1991, *Aspects of risk theory*, Springer. [28](#)
- GRANDELL, J. 2000, «Simple approximations of ruin probabilities», *Insurance : Mathematics and Economics*, vol. 26, n° 2, p. 157–173. [32](#)
- GRÜBEL, R. et R. HERMESMEIER. 1999, «Computation of compound distributions i : Aliasing errors and exponential tilting», *Astin Bulletin*, vol. 29, n° 2, p. 197–214. [16](#)
- GZYL, H., P. NOVI-INVERARDI et A. TAGLIANI. 2013, «Determination of the probability of ultimate ruin probability by maximum entropy applied to fractional moments», *Insurance : Mathematics and Economics*, vol. 53, n° 2, p. 457–463. [33](#)
- GZYL, H. et A. TAGLIANI. 2010, «Hausdorff moment problem and fractionnal moments», *Applied Mathematics and Computation*, vol. 216, n° 11, p. 3319–3328. [19](#)
- GZYL, H. et A. TAGLIANI. 2012, «Determination of the distribution of total loss from the fractional moments of its exponential», *Applied Mathematics and Computation*, vol. 219, n° 4, p. 2124–2133. [19](#), [25](#)

- HECKMAN, P. E. et G. G. MEYERS. 1983, «The calculation of aggregate loss distributions from claims severity and claims count distributions», *Proceedings of the Casualty Actuarial Society*, vol. 70, p. 133–134. [16](#)
- HESSELAGER, O. 1996, «Recursions for certain bivariate counting distributions and their compound distributions», *Astin Bulletin*, vol. 26, n° 1, p. 35–52. [23](#), [25](#)
- JIN, T., S. B. PROVOST et J. REN. 2014, «Moment-based density approximations for aggregate losses», *Scandinavian Actuarial Journal*, vol. (ahead-of-print), p. 1–30. [17](#)
- JIN, T. et J. REN. 2010, «Recursions and fast fourier transforms for certain bivariate compound distributions», *Journal of Operational Risk*, , n° 4, p. 19. [25](#)
- JIN, T. et J. REN. 2014, «Recursions and fast fourier transforms for a new bivariate aggregate claims model», *Scandinavian Actuarial Journal*, , n° 8, p. 729–752. [24](#), [25](#)
- KAAS, R., M. GOOVAERTS, J. DHAENE et M. DENUIT. 2008, *Modern actuarial risk theory : using R*, vol. 128, Springer Science and Business Media. [8](#)
- KAPUR, J. N. et H. K. KESAVAN. 1992, «Entropy optimization principles with applications», *Academic Press*. [19](#)
- KATZ, L. 1965, «Unified treatment of a broad class of discrete probability distribution», *Classical and contagious discrete distributions*, vol. 1, p. 175–182. [11](#)
- KLUGMAN, S. A., H. H. PANJER et G. E. WILLMOT. 2012, *Loss models : From data to decision*, vol. 715, John Wiley and sons. [8](#)
- MNATSAKANOV, R. et F. H. RUYMGAART. 2003, «Some properties of moment empirical cdf's with application to some inverse estimation problem», *Mathematical Methods of Statistics*, vol. 12, n° 4, p. 478–495. [18](#)
- MNATSAKANOV, R. M. 2008a, «Hausdorff moment problem : Reconstruction of probability density functions», *Statistics and Probability Letters*, vol. 78, n° 13, p. 1869–1877. [18](#)
- MNATSAKANOV, R. M. 2008b, «Hausdorff moment problem : Reconstruction of probability distribution», *Statistics and Probability Letters*, vol. 78, n° 12, p. 1612–1618. [18](#)
- MNATSAKANOV, R. M. 2011, «Moment-recovered approximations of multivariate distributions : The Laplace transform inversion», *Statistics and Probability Letters*, vol. 81, n° 1, p. 1–7. [25](#)
- MNATSAKANOV, R. M. et K. SARKISIAN. 2013, «A note on recovering the distribution from exponential moments», *Applied Mathematics and Computation*, vol. 219, p. 8730–8737. [19](#)
- MNATSAKANOV, R. M., K. SARKISIAN et A. HAKOBYAN. 2014, «Approximation of the ruin probability using the scaled Laplace transform inversion», *Working Paper*. [33](#)

- PANJER, H. H. 1981, «Recursive evaluation of a family of compound distributions», *Astin Bulletin*, vol. 12, n° 1, p. 22–26. [12](#)
- PICARD, P. et C. LEFÈVRE. 1997, «The probability of ruin in finite time with discrete claim size distribution», *Scandinavian Actuarial Journal*, p. 58–69. [34](#)
- POMMERET, D. 2001, «Orthogonal and pseudo-orthogonal multi-dimensional Appell polynomials», *Applied Mathematics and Computation*, vol. 117, n° 2, p. 285–299. [34](#)
- PROVOST, S. B. 2005, «Moment-based density approximants», *Mathematica Journal*, vol. 9, n° 4, p. 727–756. [17](#)
- RAMSAY, C. 1992, «A practical algorithm for approximating the probability of ruin», *Transaction of Society of Actuaries*, vol. 44, p. 443–461. [33](#)
- ROLSKI, T., H. SCHMIDLI, V. SCHMIDT et J. TEUGELS. 1999, *Stochastic Processes for Insurance and Finance*, Wiley series in probability and statistics. [8](#), [28](#), [31](#), [33](#)
- RUILLERE, D. et S. LOISEL. 2004, «Another look at the Picard-Lefèvre formula for finite-time ruin probability», *Insurance : Mathematics and Economics*, vol. 35, n° 2, p. 187–203. [34](#)
- STRÖTER, B. 1985, «The numerical evaluation of the aggregate claim density function via integral equations», *Blätter der DGVM*, vol. 17, n° 1, p. 1–14. [12](#)
- SUNDT, B. 1992, «On some extensions of panjer's class of counting distributions», *Astin Bulletin*, vol. 22, n° 1, p. 61–80. [12](#)
- SUNDT, B. 1999, «On multivariate Panjer recursions», *Astin Bulletin*, vol. 29, n° 1, p. 29–45. [23](#)
- SUNDT, B. 2000, «On multivariate Vernic recursions», *Astin Bulletin*, vol. 30, n° 1, p. 111–122. [23](#)
- SUNDT, B. et W. S. JEWELL. 1981, «Further results on recursive evaluation of compound distributions», *Astin Bulletin*, vol. 12, n° 1, p. 27–39. [11](#)
- TAYLOR, G. 1978, «Representation and explicit calculation of finite-time ruin probabilities», *Scandinavian Actuarial Journal*, p. 1–18. [33](#), [34](#)
- VERNIC, R. 1999, «Recursive evaluation of some bivariate compound distributions», *Astin Bulletin*. [23](#)
- VILAR, J. L. 2000, «Arithmetization of distributions and linear goal programming», *Insurance : Mathematics and Economics*, vol. 1, n° 27, p. 113–122. [13](#)
- WALHIN, J. F. et J. PARIS. 1998, «On the use of equispaced discrete distributions», *Astin Bulletin*, vol. 28, n° 2, p. 241–255. [13](#)

Chapitre 2

Approximation de la densité de probabilité via une représentation polynomiale

Sommaire

2.1 Les familles exponentielles naturelles quadratiques et leurs polynômes orthogonaux	40
2.2 Représentation polynomiale de la densité en dimension 1	41
2.2.1 Cas général	41
2.2.2 Application aux distributions composées en dimension 1	43
2.3 Représentation polynomiale de la densité en dimension n	50
2.3.1 Cas général	50
2.3.2 Application aux distributions composées en dimension n	52
2.3.3 Cas $n = 2$: La connexion avec les probabilités de Lancaster	56
2.4 Illustrations numériques : Application à la distribution Poisson composée	59
2.4.1 Le cas des sinistres de loi exponentielle $\Gamma(1, \beta)$	61
2.4.2 Le cas des sinistres de loi gamma $\Gamma(\alpha, \beta)$	68
2.4.3 Le cas des sinistres de loi uniforme $\mathcal{U}[\alpha, \beta]$	74
2.5 Perspectives en statistique	80
2.5.1 Estimation de la densité de probabilité via la représentation polynomiale	80
2.5.2 Application aux distributions composées	84
2.6 Références	88

2.1 Les familles exponentielles naturelles quadratiques et leurs polynômes orthogonaux

Soit $\mathcal{P} = \{\mathbb{P}_\theta, \theta \in \Theta\}$ avec $\Theta \subset \mathbb{R}$ une **Famille Exponentielle Naturelle (FEN)** générée par la mesure de probabilité ν sur \mathbb{R} . Soit X une variable aléatoire de loi de probabilité $\mathbb{P}_\theta \in \mathcal{P}$, alors

$$\begin{aligned} \mathbb{P}_\theta(X \in A) &= \int_A \exp [x\theta - \kappa(\theta)] d\nu(x) \\ &= \int_A f(x, \theta) d\nu(x), \end{aligned}$$

où $A \subset \mathbb{R}$, l'application $s \mapsto \kappa(s) = \log[\mathcal{L}_X(s)]$ est la **Fonction Génératrice des Cumulants (f.g.c.)** et $f(x, \theta)$ est la densité $\frac{d\mathbb{P}_\theta}{d\nu}$, de la loi de X par rapport à ν . L'espérance de X est donnée par

$$E(X) = \kappa'(\theta) = m,$$

et sa variance par

$$V(X) = \kappa''(\theta) = V(m),$$

où $m \mapsto V(m)$ est appelée fonction de variance. L'application $\theta \mapsto \kappa'(\theta)$ est bijective, sa fonction réciproque est notée $m \mapsto h(m)$, et est définie sur $\mathcal{M} = \kappa'(\Theta)$. La bijectivité de la dérivée de la **f.g.c.** permet de reparamétriser la famille de loi \mathcal{P} en caractérisant chaque élément par sa moyenne m . Chaque membre \mathbb{P}_m de la famille $\mathcal{P} = \{\mathbb{P}_m, m \in \mathcal{M}\}$ admet une moyenne m et une fonction de densité $f(x, m) = \exp \{h(m)x - \kappa[h(m)]\}$ par rapport à ν . Cette définition des **FEN** est donnée dans **BARNDORFF-NIELSEN [1978]**.

Une **FEN** est dite quadratique lorsque la fonction de variance $V(m)$ est un polynôme de degré 2 en m , avec

$$V(m) = v_2 m^2 + v_1 m + v_0, \quad (2.1)$$

où v_2, v_1 et v_0 sont des réels. Les **FENQ** comprennent les distributions normales, gammas, hyperboliques, Poisson, binomiales et négative binomiales. Les polynômes définis par

$$Q_k(x, m) = V^k(m) \left[\frac{\partial^k}{\partial m^k} f(x, m) \right] / f(x, m), \quad \forall k \in \mathbb{N}, \quad (2.2)$$

sont de degré k en m et en x . De plus, si ν génère une **FENQ**, alors $\{Q_k\}_{k \in \mathbb{N}}$ est une suite de polynômes orthogonaux par rapport à ν au sens où

$$\langle Q_k, Q_l \rangle = \int Q_k(x, m) Q_l(x, m) f(x, m) d\nu(x) = \delta_{kl} \|Q_k(\cdot, m)\|^2, \quad k, l \in \mathbb{N}, \quad (2.3)$$

où

$$\delta_{kl} = \begin{cases} 1 & \text{Si } k = l \\ 0 & \text{Sinon} \end{cases} \quad (2.4)$$

est le symbole de Kronecker. Si ν admet une moyenne m_0 alors $f(x, m_0) = 1$ et les polynômes orthogonaux associés à ν sont définis par

$$Q_k(x, m_0) = V^k(m) \left[\frac{\partial^k}{\partial m^k} f(x, m) \right]_{m=m_0}, \quad \forall k \in \mathbb{N}. \quad (2.5)$$

Dans le reste de ce document, il sera question de développement via les polynômes orthonormaux par rapport à la mesure de référence ν . C'est à dire une version normalisée des polynômes (2.5), notée

$$Q_k(x) \doteq \frac{Q_k(x, m_0)}{\|Q_n(\cdot, m_0)\|}. \quad (2.6)$$

Le papier de MORRIS [1982] propose une caractérisation des FENQ, de leurs propriétés, et de leurs polynômes orthogonaux.

2.2 Représentation polynomiale de la densité en dimension 1

2.2.1 Cas général

Soit $L^2(\nu)$ l'ensemble des fonctions de carré intégrable par rapport à ν , mesure de probabilité génératrice d'une FENQ, au sens où

$$\|f\|^2 = \int f^2(x) d\nu(x) < +\infty, \quad (2.7)$$

implique que $f \in L^2(\nu)$. L'ensemble des polynômes est dense dans l'ensemble des fonctions de carré intégrable par rapport à ν . La suite de polynômes $\{Q_k\}_{k \in \mathbb{N}}$, définie dans l'équation (2.6), forme un système orthonormal complet de $L^2(\nu)$. La formule d'approximation de la densité de probabilité est fondée sur le résultat suivant :

Theorème 6. *Soit ν une mesure de probabilité sur \mathbb{R} , génératrice d'une FENQ. Soit $\{Q_k\}_{k \in \mathbb{N}}$ la suite de polynômes orthonormaux par rapport à ν . Soit X une variable aléatoire réelle, de loi de probabilité \mathbb{P}_X absolument continue par rapport à ν . Si $\frac{d\mathbb{P}_X}{d\nu} \in L^2(\nu)$ alors*

$$\frac{d\mathbb{P}_X}{d\nu}(x) = \sum_{k=0}^{+\infty} a_k Q_k(x), \quad x \in \mathbb{R}, \quad (2.8)$$

avec

$$a_k = E[Q_k(X)], \quad \forall k \in \mathbb{N}. \quad (2.9)$$

De plus, l'identité de Parseval

$$\left\| \frac{d\mathbb{P}_X}{d\nu} \right\|^2 = \sum_{k=0}^{+\infty} a_k^2 \quad (2.10)$$

est vérifiée.

Démonstration. Comme $\{Q_k\}_{k \in \mathbb{N}}$ forment une base orthogonale de $L^2(\nu)$ et que $\frac{d\mathbb{P}_X}{d\nu} \in L^2(\nu)$ alors par projection orthogonale

$$\frac{d\mathbb{P}_X}{d\nu}(x) = \sum_{k=0}^{+\infty} \left\langle \frac{d\mathbb{P}_X}{d\nu}, Q_k \right\rangle Q_k(x), \quad (2.11)$$

et

$$\begin{aligned}
 a_k &\doteq \left\langle \frac{d\mathbb{P}_X}{d\nu}, Q_k \right\rangle \\
 &= \int \frac{d\mathbb{P}_X}{d\nu}(x) Q_k(x) d\nu(x) \\
 &= \int Q_k(x) d\mathbb{P}_X \\
 &= \mathbb{E}[Q_k(X)].
 \end{aligned}$$

L'identité (2.10) est une conséquence immédiate de l'orthogonalité des polynômes $\{Q_k\}_{k \in \mathbb{N}}$. □

Remarque 1. L'expression des coefficients (2.9) montre qu'ils sont égaux à une combinaison linéaire de moments de la distribution de X . Une approximation très semblable est proposée dans le papier de PROVOST [2005]. La façon de définir la représentation polynomiale diffère cependant.

Les mesures de probabilité ν et \mathbb{P}_X sont absolument continues par rapport à une mesure positive commune λ . Leurs densités respectives sont notées f_ν et f_X . L'application du Théorème 6 permet d'écrire la densité de X sous la forme

$$\begin{aligned}
 f_X(x) &= \sum_{k=0}^{+\infty} a_k Q_k(x) f_\nu(x) \\
 &= f_{X,\nu}(x) f_\nu(x), \quad x \in \mathbb{R},
 \end{aligned} \tag{2.12}$$

où $f_{X,\nu}(x) = \frac{d\mathbb{P}_X}{d\nu}(x)$. L'égalité 2.8 du Théorème 6 implique que

$$\lim_{K \rightarrow +\infty} \int [f_{X,\nu}(x) - f_{X,\nu}^K(x)]^2 d\nu(x) = 0, \tag{2.13}$$

avec

$$f_{X,\nu}^K(x) = \sum_{k=0}^K a_k Q_k(x). \tag{2.14}$$

La convergence en moyenne quadratique (2.13) des sommes partielles justifie l'emploi de $f_{X,\nu}^K$ pour approcher $f_{X,\nu}$. La séquence $\{a_k\}_{k \in \mathbb{N}}$ représente les coefficients du développement polynomial. L'identité de Parseval (2.10) implique que

$$\lim_{k \rightarrow +\infty} a_k = 0. \tag{2.15}$$

La qualité de l'approximation (2.14) est fonction de la vitesse de convergence vers 0 de la suite $\{a_k\}_{k \in \mathbb{N}}$. L'Erreur Quadratique Intégrée (EQI) de l'approximation de $f_{X,\nu}$ par $f_{X,\nu}^K$ est définie par

$$L(f_{X,\nu}, f_{X,\nu}^K) = \int (f_{X,\nu}(x) - f_{X,\nu}^K(x))^2 d\nu(x), \tag{2.16}$$

et

$$L(f_{X,\nu}, f_{X,\nu}^K) = \sum_{k=K+1}^{+\infty} a_k^2. \tag{2.17}$$

L'approximation d'ordre K de la densité de X est donnée par

$$f_X^K(x) = f_{X,v}^K(x) f_v(x), \quad x \in \mathbb{R}. \quad (2.18)$$

Le polynôme de degré 0 vérifie $Q_0(x) = 1$ pour tout $x \in \mathbb{R}$ par construction, et comme $a_0 = 1$ alors

$$\begin{aligned} \int f_X^K(x) dx &= \sum_{k=0}^{+\infty} a_k \int Q_0(x) \times Q_k(x) dv(x) \\ &= \sum_{k=0}^{+\infty} a_k \delta_{0k} \\ &= a_0 \\ &= 1. \end{aligned}$$

Ainsi l'intégrale de l'approximation (2.18) sur le support de X vaut 1. Garantir la positivité de l'approximation (2.18) pour un ordre de troncature K est un problème difficile. L'approximation peut être négative si l'ordre de troncature n'est pas suffisamment élevé. En ce sens, il ne s'agit pas d'une approximation *bona-fide* de la densité de probabilité. Une solution consiste à procéder à la modification de l'approximation suivante

$$\tilde{f}_{X,v}^K(x) = \frac{\max[f_{X,v}^K(x), 0]}{\lambda_K}, \quad x \in \mathbb{R}, \quad (2.19)$$

où λ_K est une constante de normalisation définie par

$$\lambda_K = \int \max[f_{X,v}^K(x), 0] dv(x). \quad (2.20)$$

L'approximation modifiée (2.19) est automatiquement associée à une EQI plus faible que l'approximation initiale. L'approximation pour un $x \in \mathbb{R}$ est cependant altérée voire dégradée.

La vérification de la condition d'intégrabilité et la vitesse de convergence des coefficients de la représentation polynomiale sont des sujets centraux de ce travail. Ces deux problèmes sont liés aux choix de la mesure de référence et de ses paramètres.

2.2.2 Application aux distributions composées en dimension 1

Soit X une variable aléatoire de la forme

$$X = \sum_{i=1}^N U_i, \quad (2.21)$$

où N est une variable aléatoire de comptage de loi de probabilité \mathbb{N} , $\{U_i\}_{i \in \mathbb{N}}$ est une suite de variables aléatoires réelles, continues, positives et *i.i.d.* suivant une loi de probabilité \mathbb{P}_U . La variable aléatoire (2.21) admet une distribution composée $(\mathbb{P}_N, \mathbb{P}_U)$, et sa loi de probabilité s'écrit

$$d\mathbb{P}_X(x) = f_N(0) \delta_0(x) + dG_X(x). \quad (2.22)$$

Choix de la mesure de référence

L'application du Théorème 6 nécessite de choisir une mesure ν appartenant aux FENQ, telle que la mesure de probabilité \mathbb{P}_X soit absolument continue par rapport à ν . La masse de probabilité en 0 de la loi de X rend cette condition impossible à remplir. L'indétermination de la loi de probabilité de X repose sur la mesure de probabilité défailante \mathbb{G}_X , absolument continue par rapport à la mesure de Lebesgue. L'idée est d'utiliser la méthode d'approximation polynomiale pour retrouver la densité défailante g_X donnée par

$$g_X(x) = \sum_{k=1}^{+\infty} f_N(k) f_U^{(*k)}(x). \quad (2.23)$$

Pour appliquer le Théorème 6, il est nécessaire que $\text{Supp}(\mathbb{G}_X) \subset \text{Supp}(\nu)$. Comme $\text{Supp}(\mathbb{G}_X) = \mathbb{R}_+$, alors il est logique d'opter pour une loi gamma $\Gamma(r, m)$ en guise de mesure de référence. La densité de la mesure de référence ν est

$$f_\nu(x) = \frac{e^{-x/m} x^{r-1}}{m^r \Gamma(r)}, \quad x \in \mathbb{R}_+. \quad (2.24)$$

Les polynômes orthogonaux par rapport à ν sont les polynômes de Laguerre généralisés. L'expression des polynômes (2.5) est donnée par

$$Q_k(x) = (-1)^n \binom{n+r-1}{n}^{-1/2} L_n^{r-1}(x/m), \quad (2.25)$$

où

$$L_n^{r-1}(x) = \sum_{k=0}^n \binom{n+r-1}{n-k} \frac{(-x)^k}{k!}, \quad (2.26)$$

d'après la définition donnée dans le livre de SZEGÖ [1939]. Ce choix est cohérent avec les résultats asymptotiques obtenus par SUNDT [1982] et EMBRECHTS et collab. [1985], dans lesquels la queue de distribution de la charge totale d'un portefeuille, lorsque la fréquence des sinistres est distribuée suivant une loi binomiale négative, s'apparente à la queue de la distribution d'une loi gamma. Les approximations présentées dans le papier de PAPUSH et collab. [2001] font intervenir la loi gamma, normale, et log-normale. Sur les 7 situations considérées, l'approximation basée sur la loi gamma renvoie systématiquement les meilleurs résultats. L'application de la méthode de PROVOST [2005] conduit aussi à utiliser la loi gamma et les polynômes de Laguerre lorsqu'elle est appliquée aux distributions composées, voir JIN et collab. [2014].

Si $\frac{d\mathbb{G}_X}{d\nu} \in L^2(\nu)$ alors, par application du Théorème 6, la densité g_X de la mesure de probabilité défailante \mathbb{G}_X admet la représentation

$$\begin{aligned} g_X(x) &= \sum_{k=0}^{+\infty} a_k Q_k(x) f_\nu(x) \\ &= g_{X,\nu}(x) f_\nu(x), \end{aligned} \quad (2.27)$$

où $g_{X,\nu}(x) = \frac{d\mathbb{G}_X}{d\nu}(x)$ est la densité défailante de X par rapport à la mesure de référence ν . L'approximation polynomiale est obtenue en tronquant la série infinie (2.27)

à l'ordre $K \geq 0$, avec

$$\begin{aligned} g_X^K(x) &= \sum_{k=0}^K a_k Q_k(x) f_v(x) \\ &= g_{X,v}^K(x) f_v(x), \end{aligned} \quad (2.28)$$

où $g_{X,v}^K(x)$ est l'approximation d'ordre de troncature K de la densité défaillante de X par rapport à la mesure de référence v . L'intégrale de l'approximation de la densité (2.28) donne accès à la **f.d.r.**, la **f.d.s.** et la prime *stop loss* généralisée.

Vérification de la condition d'intégrabilité

L'approximation est valide sous réserve de vérifier la condition d'intégrabilité

$$\int_0^{+\infty} \left(\frac{dG_X}{dv}(x) \right)^2 dv(x) < +\infty \Leftrightarrow \int_0^{+\infty} g_X(x)^2 e^{x/m} x^{1-r} dx < +\infty. \quad (2.29)$$

Le résultat suivant majore la densité défaillante g_X pour choisir sereinement les paramètres de la mesure de référence et assurer la validité des approximations.

Proposition 6. *Soit X une variable aléatoire continue, réelle, positive, de loi de probabilité \mathbb{P}_X et de densité f_X . Soit $\gamma_X = \inf\{s > 0 : \mathcal{L}_X(s) = +\infty\}$, si $x \mapsto f_X(x)$ est continuellement dérivable et qu'il existe $a \in \mathbb{R}^+$ tel que $x \mapsto f_X(x)$ soit strictement décroissante pour $x > a$, alors*

$$f_X(x) < A(s_0) e^{-s_0 x}, \quad \forall x > a, \quad (2.30)$$

où $s_0 \in [0, \gamma_X[$, et la quantité $A(s_0)$ ne dépend pas de la valeur de x .

Démonstration. Soit $x > a$ et $s_0 \in [0, \gamma_X[$, la **f.g.m.** de X est minorée par

$$\begin{aligned} \mathcal{L}_X(s_0) &> \int_a^x e^{s_0 y} f_X(y) dy \\ &> \left[\frac{e^{s_0 y}}{s_0} f_X(y) \right]_a^x - \frac{1}{s_0} \int_a^x e^{s_0 y} f_X'(y) dy \\ &> \frac{1}{s_0} [f_X(x) e^{s_0 x} - f_X(a) e^{s_0 a}]. \end{aligned}$$

La densité de X est alors majorée par

$$f_X(x) < A(s_0) e^{-s_0 x},$$

où $A(s_0) = [s_0 \mathcal{L}_X(s_0) + f_X(a) e^{s_0 a}]$. □

La Proposition 6 est applicable directement à la majoration d'une densité de probabilité défaillante telle que g_X . Le résultat suivant explicite le choix des paramètres m et r de la mesure de référence (2.24) pour assurer la validité des représentations polynomiales.

Corollaire 2. *Si la densité défaillante $x \mapsto g_X(x)$ vérifie les hypothèses de la Proposition 6 alors la paramétrisation*

$$r \leq 1 ; \quad m \in \left] \frac{1}{2\gamma_X}, +\infty \right[, \quad (2.31)$$

assure la vérification de la condition d'intégrabilité (2.29).

Démonstration. La fonction de densité de probabilité défaillante g_X vérifie les hypothèses de la Proposition 6, soit $a \in \mathbb{R}^+$ tel que $x \mapsto g_X(x)$ soit strictement décroissante pour $x > a$. D'après la Proposition 6, la densité de probabilité défaillante g_X est majorée par

$$g_X(x) < A(s_0)e^{-s_0x}, \quad (2.32)$$

pour $s_0 \in [0, \gamma_0[$ et $x > a$. L'intégrale de droite dans l'équivalence (2.29) s'écrit

$$\int_0^{+\infty} g_X(x)^2 e^{x/m} x^{1-r} dx = \int_0^a g_X(x)^2 e^{x/m} x^{1-r} dx \quad (2.33)$$

$$+ \int_a^{+\infty} g_X(x)^2 e^{x/m} x^{1-r} dx. \quad (2.34)$$

Pour $r \leq 1$, l'intégrale (2.33) est finie en tant qu'intégrale d'une fonction continue sur un intervalle borné. L'application de la Proposition 6 permet de majorer l'intégrale (2.34) par

$$\int_a^{+\infty} g_X(x)^2 e^{x/m} x^{1-r} dx < A(s_0) \int_a^{+\infty} e^{-x(2s_0-1/m)} x^{1-r} dx. \quad (2.35)$$

Ce qui implique que l'intégrale (2.34) est finie si $2s_0 - 1/m > 0$, ce qui équivaut à $m \in \left] \frac{1}{2s_0}, +\infty \right[$. Comme $s_0 \in [0, \gamma_X[$, alors la condition d'intégrabilité (2.29) est vérifiée pour

$$r \leq 1 ; m \in \left] \frac{1}{2\gamma_X}, +\infty \right[.$$

□

L'application de l'approximation de la méthode des polynômes orthogonaux à la probabilité de ruine est un cas particulier de l'application dans le cadre d'une distribution composée générale. Un résultat proche de la Proposition 6 est établi dans **GOFFARD et collab. [2015]**, voir Chapitre 3 de ce manuscrit. L'idée sous-jacente est que la quantité $\gamma_X = \inf\{s > 0 : \mathcal{L}_X(s) = +\infty\}$ coïncide avec l'unique solution positive de l'équation de Cramer-Lundberg (1.93). La définition particulière de la variable aléatoire X , donnée par le théorème 3, dans le cadre de la probabilité de ruine ultime permet de relaxer l'hypothèse de décroissante sur fonction $x \mapsto g_X(x)$. Un résultat identique au Corollaire 2 est aussi établi dans **GOFFARD et collab. [2015]**.

Vitesse de convergence de l'approximation et calcul des coefficients de la représentation polynomiale

Une bonne approximation est caractérisée par une grande précision pour un ordre de troncature faible. Ces deux critères reposent intégralement sur la vitesse à laquelle les coefficients du développement polynomial, la suite des $\{a_k\}_{k \in \mathbb{N}}$, tendent vers 0 lorsque k augmente. Le résultat suivant s'applique dans le cadre de la combinaison mesure de référence gamma et polynômes de Laguerre généralisés.

Proposition 7. *Soit ν la mesure de probabilité d'une loi gamma $\Gamma(m, r)$ associée à la suite de polynômes orthonormaux $\{Q_k\}_{k \in \mathbb{N}}$. Soit X une variable aléatoire réelle, continue*

et positive de loi \mathbb{P}_X et de densité $f_{X,v}$ par rapport à v . Si $x \mapsto f_{X,v}(x)$ est continue, deux fois dérivable et que $f_{X,v}$, $f_{X,v}^{(1)}$, et $f_{X,v}^{(2)}$ appartiennent à $L^2(v)$ alors

$$a_k = o\left(\frac{1}{k}\right), \quad k \rightarrow +\infty. \quad (2.36)$$

Démonstration. Comme $f_{X,v} \in L^2(v)$ alors par application du Théorème 6, la densité $f_{X,v}$ admet la représentation polynomiale

$$f_{X,v}(x) = \sum_{k=0}^{+\infty} a_k Q_k(x). \quad (2.37)$$

L'identité de Parseval

$$\|f_{X,v}\|^2 = \sum_{k=0}^{+\infty} a_k^2 < +\infty,$$

est vérifiée, et par conséquent $\lim_{k \rightarrow +\infty} a_k = 0$. Comme les fonctions $x \mapsto f_{X,v}^{(1)}(x)$, et $x \mapsto f_{X,v}^{(2)}(x)$ appartiennent à $L^2(v)$, alors l'application $x \mapsto (x - rm)f_{X,v}^{(1)}(x) - xf_{X,v}^{(2)}(x)$ appartient à $L^2(v)$ et admet la représentation

$$(x - rm)f_{X,v}^{(1)}(x) - xf_{X,v}^{(2)}(x) = \sum_{k=0}^{+\infty} b_k Q_k(x), \quad (2.38)$$

en vertu du Théorème 6. L'identité de Parseval

$$\|(x - rm)f_{X,v}^{(1)}(\cdot) - xf_{X,v}^{(2)}(\cdot)\|^2 = \sum_{k=0}^{+\infty} b_k^2 < +\infty,$$

est vérifiée, ce qui implique que $\lim_{k \rightarrow +\infty} b_k = 0$. Par dérivation de la représentation (2.37), il vient

$$(x - rm)f_{X,v}^{(1)}(x) - xf_{X,v}^{(2)}(x) = \sum_{n=0}^{+\infty} a_n \left[(x - rm)Q_n^{(1)}(x) - xQ_n^{(2)}(x) \right]. \quad (2.39)$$

Les polynômes de Laguerre généralisés $\{L_k^{r-1}\}_{k \in \mathbb{N}}$ sont solutions de l'équation différentielle du second ordre

$$x(L_k^{r-1})^{(2)}(x) + (r-x)(L_k^{r-1})^{(1)}(x) + kL_k^{r-1}(x) = 0, \quad k \geq 2, \quad (2.40)$$

voir SZEGÖ [1939]. La suite de polynômes $\{Q_k\}_{k \in \mathbb{N}}$ est définie par

$$Q_k(x) = (-1)^k c_k L_k^{r-1}\left(\frac{x}{m}\right), \quad (2.41)$$

avec $c_k = \binom{k+r-1}{k}^{-1/2}$. La réinjection de l'expression des polynômes (2.41) dans l'équation différentielle (2.40) implique que

$$kQ_k(x) = (x - rm)Q_k^{(1)}(x) + mxQ_k^{(2)}(x), \quad k \geq 2. \quad (2.42)$$

La réinjection de l'équation (2.42) dans la représentation polynomiale (2.38) conduit à

$$(x - rm)f_{X,v}^{(1)}(x) - xf_{X,v}^{(2)}(x) = \sum_{k=0}^{+\infty} ka_k Q_k(x). \quad (2.43)$$

Par unicité de la projection orthogonale de la fonction $x \mapsto (x - rm)f_{X,v}^{(1)}(x) - xf_{X,v}^{(2)}(x)$ sur la base de polynômes $\{Q_k\}_{k \in \mathbb{N}}$ et identification avec les coefficients de la représentation (2.38), il vient $b_k = ka_k$ puis $\lim_{k \rightarrow +\infty} ka_k = 0$, ce qui est équivalent à l'assertion (2.36). \square

La vitesse de décroissance (2.36) des coefficients de la représentation polynomiale est acceptable dans un contexte d'estimation de la densité de probabilité. Une vitesse de décroissance plus importante est souhaitable pour l'approximation de la densité de probabilité. Les paramètres de la mesure de référence peuvent être choisis de façon à accélérer considérablement la convergence de la suite des coefficients via l'étude la fonction génératrice des coefficients de la représentation polynomiale.

Soit X une variable aléatoire réelle, continue, positive, de loi \mathbb{P}_X , et de densité f_X . Si $\frac{d\mathbb{P}_X}{dv} \in L^2(v)$, alors la densité f_X admet la représentation

$$f_X(x) = \sum_{k=0}^{+\infty} a_k Q_k(x) f_v(x), \quad (2.44)$$

par application du Théorème 6. La transformée de Laplace de la représentation (2.44) s'écrit

$$\begin{aligned} \mathcal{L}_X(s) &= \int_0^{+\infty} e^{sx} \sum_{k=0}^{+\infty} a_k Q_k(x) f_v(x) dx \\ &= \sum_{k=0}^{+\infty} a_k \int_0^{+\infty} e^{sx} Q_k(x) f_v(x) dx \\ &= \sum_{k=0}^{+\infty} a_k \int_0^{+\infty} e^{sx} \frac{(-1)^k}{\sqrt{\binom{k+r-1}{k}}} L_k^{r-1}\left(\frac{x}{m}\right) f_v(x) dx \\ &= \sum_{k=0}^{+\infty} a_k \frac{(-1)^k}{\sqrt{\binom{k+r-1}{k}}} \int_0^{+\infty} e^{sx} \sum_{i=0}^k \binom{k+r-1}{k-i} \frac{(-x)^i}{i! m^i} \frac{x^{r-1} e^{-x/m}}{\Gamma(r) m^r} dx \\ &= \sum_{k=0}^{+\infty} a_k \frac{(-1)^k}{\sqrt{\binom{k+r-1}{k}}} \sum_{i=0}^k \frac{(-1)^i \Gamma(k+r)}{\Gamma(k-i+1) i! \Gamma(r)} \int_0^{+\infty} e^{sx} \frac{x^{i+r-1} e^{-x/m}}{\Gamma(r+i) m^{i+r}} dx \\ &= \sum_{k=0}^{+\infty} a_k (-1)^k \sqrt{\binom{k+r-1}{k}} \sum_{i=0}^k \binom{k}{i} (-1)^i \left(\frac{1}{1-sm}\right)^{i+r} \\ &= \sum_{k=0}^{+\infty} a_k c_k \left(\frac{1}{1-sm}\right)^r \left(\frac{sm}{1-sm}\right)^k \\ &= \left(\frac{1}{1-sm}\right)^r \mathcal{C}\left(\frac{sm}{1-sm}\right), \end{aligned} \quad (2.45)$$

où $z \mapsto \mathcal{C}(z) = \sum_{k=0}^{+\infty} a_k c_k z^k$ est la fonction génératrice de la séquence $\{a_k c_k\}_{k \in \mathbb{N}}$. La fonction génératrice s'exprime en fonction de la transformée de Laplace de X via

$$\mathcal{C}(z) = (1+z)^{-r} \mathcal{L}_X\left(\frac{z}{m(1+z)}\right). \quad (2.46)$$

Les coefficients de la représentation polynomiale $\{a_k\}_{k \in \mathbb{N}}$ s'obtiennent par dérivation de la fonction génératrice (2.46), avec

$$a_k = \frac{1}{k! c_k} \left[\frac{d^k}{dz^k} \mathcal{C}(z) \right]_{z=0}. \quad (2.47)$$

Le but du jeu est alors de choisir m et r pour simplifier au maximum l'expression de la fonction génératrice (2.46), et prévoir le comportement de la suite $\{a_k\}_{k \in \mathbb{N}}$. Dans certains cas, ce procédé permet d'obtenir une formule fermée comme le montre le résultat suivant :

Proposition 8. *Soit X une variable aléatoire admettant une distribution de probabilité composée $(\mathbb{P}_N, \mathbb{P}_U)$, où \mathbb{P}_N est la mesure de probabilité associée à une loi binomiale négative $\mathcal{NB}(\alpha, p)$, et \mathbb{P}_U est une mesure de probabilité associée à une loi gamma $\Gamma(1, \beta)$, la loi de probabilité de X est donnée par*

$$\mathbb{P}_X(x) = (1-p)^\alpha \delta_0(x) + \sum_{k=0}^{\alpha-1} a_k Q_k(x) \nu(x),$$

où ν est la mesure de probabilité d'une loi exponentielle $\Gamma\left(1, \frac{\beta}{1-p}\right)$. La suite $\{Q_k\}_{k \in \mathbb{N}}$ est la suite des polynômes orthonormaux par rapport à ν et les coefficients $\{a_k\}_{k \in \mathbb{N}}$ de la représentation polynomiale sont donnés par

$$a_k = \begin{cases} c_k^{-1} p (1-p)^{\alpha-1} \sum_{i=k}^{\alpha-1} \binom{i}{k} k! \frac{(-p)^k}{(1-p)^i}, & \text{si } k < \alpha, \\ 0, & \text{si } k \geq \alpha. \end{cases}$$

Démonstration. La loi de probabilité de X admet une singularité en 0, avec une masse de probabilité égale à $f_N(0) = (1-p)^\alpha$. Soit \mathbb{G}_X la partie continue de la loi de probabilité de X . Soit ν la mesure de probabilité associée à la loi $\Gamma(m, r)$, et $\{Q_k\}_{k \in \mathbb{N}}$ sa suite de polynômes orthonormaux. Si $\frac{d\mathbb{G}_X}{d\nu} \in L^2(\nu)$, alors, par application du Théorème 6, la densité de probabilité défailante associée à \mathbb{G}_X admet la représentation

$$g_X(x) = \sum_{k=0}^{+\infty} a_k Q_k(x) f_\nu(x). \quad (2.48)$$

La transformée de Laplace de la partie continue g_X de la loi de X est donnée par

$$\mathcal{L}_{g_X}(s) = (1-p)^\alpha \left\{ \left[\frac{(1-s\beta)}{(1-p)-s\beta} \right]^\alpha - 1 \right\}. \quad (2.49)$$

La réinjection dans l'expression de la fonction génératrice des coefficients (2.46) conduit à

$$\mathcal{C}(z) = (1+z)^{-r} (1-p)^\alpha \left\{ \left[\frac{z(m-\beta) + m}{z[m(1-p)-\beta] + m(1-p)} \right]^\alpha - 1 \right\}. \quad (2.50)$$

L'expression de la fonction génératrice (2.50) se simplifie en choisissant $m = \beta/(1-p)$ avec

$$\mathcal{C}(z) = (1+z)^{-r} [(1+zp)^\alpha - (1-p)^\alpha]. \quad (2.51)$$

L'emploi de l'identité remarquable

$$a^\alpha - b^\alpha = (a-b) \sum_{i=0}^{\alpha-1} a^i b^{\alpha-1-i}, \quad \forall a, b \in \mathbb{R},$$

avec $a = 1+zp$ et $b = 1-p$ dans l'expression de la fonction génératrice (2.51) donne

$$\mathcal{C}(z) = (1+z)^{1-r} p (1-p)^{\alpha-1} \left[\sum_{i=0}^{\alpha-1} \left(\frac{1+zp}{1-p} \right)^i \right]. \quad (2.52)$$

L'expression de la fonction génératrice (2.52) en choisissant $r = 1$, avec

$$\mathcal{C}(z) = p(1-p)^{\alpha-1} \left[\sum_{i=0}^{\alpha-1} \left(\frac{1+zp}{1-p} \right)^i \right]. \quad (2.53)$$

Les coefficients du développement sont alors donnés par

$$a_k = \begin{cases} c_k^{-1} p(1-p)^{\alpha-1} \sum_{i=k}^{\alpha-1} \binom{i}{k} k! \frac{(-p)^k}{(1-p)^i}, & \text{si } k < \alpha, \\ 0, & \text{si } k \geq \alpha. \end{cases}$$

□

Malheureusement, le choix des paramètres de la mesure de référence n'est pas toujours aussi aisé. Dans la pire des situations, il faut se cantonner aux valeurs de paramètres prévues par le Corollaire 2 pour une application sur les distributions composées. La pire des situations correspond à une fonction génératrice des coefficients très complexe avec la nécessité de prendre un ordre de troncature très grand. Cette situation peut engendrer des difficultés d'ordre numérique lors de l'évaluation de la dérivée

$$a_k = \frac{1}{k!c_k} \left[\frac{d^k}{dz^k} \mathcal{C}(z) \right]_{z=0}, \quad (2.54)$$

lorsque k est grand. Cette difficulté numérique peut être contournée par l'emploi de la formule d'intégrale de contour de Cauchy permettant d'exprimer la dérivée (2.54) comme une intégrale, avec

$$a_k = \frac{1}{c_k 2\pi i} \int_{\mathcal{C}_r} \frac{\mathcal{C}(z)}{z^{k+1}} dz,$$

où \mathcal{C}_r est un cercle de centre l'origine et de rayon $0 < r < 1$. Le changement de variable $z = re^{iw}$ conduit à

$$a_k = \frac{1}{c_k 2\pi r^k} \int_0^{2\pi} \mathcal{C}(re^{iw}) e^{-iw} dw. \quad (2.55)$$

L'intégrale (2.55) est approchée par une méthode des trapèzes, ce qui donne

$$\begin{aligned} a_k &\approx \frac{1}{c_k 2\pi r^k} \sum_{j=1}^{2k} (-1)^j \Re \left[\mathcal{C} \left(r e^{\pi j i / k} \right) \right] \\ &\approx \frac{1}{c_k 2\pi r^k} \left\{ \mathcal{C}(r) + (-1)^k \mathcal{C}(-r) + 2 \sum_{j=1}^k (-1)^j \Re \left[\mathcal{C} \left(r e^{\pi j i / k} \right) \right] \right\}. \end{aligned}$$

La pertinence d'une telle approximation est étudiée dans le papier d'ABATE et collab. [1995].

2.3 Représentation polynomiale de la densité en dimension n

2.3.1 Cas général

Soit $\mathbf{X} = (X_1, \dots, X_n)$ un vecteur aléatoire réel, dont la loi de probabilité $P_{\mathbf{X}}$ est absolument continue par rapport à une mesure multivariée λ , et de densité jointe $f_{\mathbf{X}}$. Une

mesure de probabilité multivariée est définie via le produit de mesure de probabilité appartenant aux FENQ avec

$$\nu(x_1, \dots, x_n) = \prod_{i=1}^n \nu_i(x_i), \quad (2.56)$$

où ν_i est une mesure de probabilité univariée de moyenne m_i , appartenant aux FENQ pour $i = 1, \dots, n$. Les mesures de probabilité $\{\nu_i\}_{i \in \{1, \dots, n\}}$ sont choisies de telle sorte que ν soit absolument continue par rapport à λ . La densité jointe de ν est définie par

$$f_\nu(x_1, \dots, x_n) = \prod_{i=1}^n f_{\nu_i}(x_i). \quad (2.57)$$

Soit $\{Q_k^{\nu_i}\}_{k \in \mathbb{N}}$ le système de polynômes orthonormaux par rapport à la mesure ν_i pour $i = 1, \dots, n$. Un système de polynômes orthonormaux par rapport à ν est construit via

$$Q_{\mathbf{k}}(x_1, \dots, x_n) = \prod_{i=1}^n Q_{k_i}^{\nu_i}(x_i), \quad (2.58)$$

où k_i désigne le degré du polynôme associé à la mesure de probabilité ν_i pour $i = 1, \dots, n$. Soit $\mathbf{k}, \mathbf{l} \in \mathbb{N}^n$, les polynômes (2.58) sont orthonormaux par rapport à ν au sens où

$$\begin{aligned} \langle Q_{\mathbf{k}}, Q_{\mathbf{l}} \rangle &= \int Q_{\mathbf{k}}(x_1, \dots, x_n) Q_{\mathbf{l}}(x_1, \dots, x_n) d\nu(x_1, \dots, x_n) \\ &= \prod_{i=1}^n \delta_{k_i l_i}. \end{aligned}$$

Soit $L^2(\nu)$ l'espace des fonctions de carré intégrable par rapport à ν .

Théorème 7. Soit ν une mesure de probabilité multivariée de la forme (2.56). Soit $\{Q_{\mathbf{k}}\}_{\mathbf{k} \in \mathbb{N}^n}$ la séquence de polynômes orthonormaux par rapport à ν . Soit $\mathbf{X} = (X_1, \dots, X_n)$ un vecteur aléatoire réel, de loi de probabilité $P_{\mathbf{X}}$, absolument continue par rapport à ν . Si $\frac{dP_{\mathbf{X}}}{d\nu} \in L^2(\nu)$ alors

$$\frac{dP_{\mathbf{X}}}{d\nu}(\mathbf{x}) = \sum_{k_1, \dots, k_n=0}^{+\infty} a_{\mathbf{k}} Q_{\mathbf{k}}(\mathbf{x}),$$

avec

$$a_{\mathbf{k}} = \mathbb{E}[Q_{\mathbf{k}}(\mathbf{x})], \quad \mathbf{k} \in \mathbb{N}^n.$$

De plus, l'identité de Parseval

$$\left\| \frac{dP_{\mathbf{X}}}{d\nu} \right\|^2 = \sum_{k_1, \dots, k_n=0}^{+\infty} a_{\mathbf{k}}^2,$$

est vérifiée.

Démonstration. La démonstration est identique à celle du Théorème 6. □

2.3.2 Application aux distributions composées en dimension n

Quel modèle pour la distribution composée ?

L'application de la méthode aux distributions composées en dimension n nécessite de choisir un modèle à étudier parmi les Modèles 1, 2, et 3. Les distributions de probabilité qui gouvernent dans les Modèles 2, et 3 sont associées à de nombreuses singularités. En effet, les probabilités qu'une ou plusieurs des composantes soient égales à 0 sont non nulles. L'hypothèse d'absolue continuité, pré-requis de l'application du Théorème 7, implique l'isolation des parties singulières des distributions de probabilité. Ce travail a été effectué en dimension 1, il est bien plus fastidieux en dimension supérieure. Pour une première application de l'approximation polynomiale, le Modèle 1 est considéré. Dans le Modèle 1, le vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_n)$ est de la forme

$$\begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ \vdots \\ U_{nj} \end{pmatrix}, \quad (2.59)$$

où N est une variable aléatoire de comptage de loi \mathbb{P}_N et $\{\mathbf{U}_j\}_{j \in \mathbb{N}}$ est une suite de vecteurs aléatoires *i.i.d.* suivant une loi de probabilité $\mathbb{P}_{\mathbf{U}}$. La loi de \mathbf{U} est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n , sa densité jointe est notée $f_{\mathbf{U}}$. La variable aléatoire N et la suite de vecteurs aléatoires $\{\mathbf{U}_j\}_{j \in \mathbb{N}}$ sont indépendantes. La loi de probabilité du vecteur (2.59) est définie par

$$d\mathbb{P}_{\mathbf{X}}(\mathbf{x}) = f_N(0)\delta_{\mathbf{0}}(\mathbf{x}) + d\mathbb{G}_{\mathbf{X}}(\mathbf{x}), \quad (2.60)$$

où $\delta_{\mathbf{0}}(x_1, \dots, x_n) = \prod_{i=1}^n \delta_0(x_i)$ et $\mathbf{0} = (0, \dots, 0) \in \mathbb{R}^n$. Pour choisir la loi du vecteur aléatoire \mathbf{U} , il est possible de s'inspirer des lois de probabilité sur $\mathbb{R}_+ \times \mathbb{R}_+$ décrites dans l'ouvrage de [BALAKRISHNAN et LAI \[2009\]](#).

Choix de la mesure de référence

La loi de probabilité défailante $\mathbb{G}_{\mathbf{X}}$ est absolument continue par rapport à la mesure de Lebesgue sur \mathbb{R}^n , le support est $\text{Supp}(\mathbb{G}_{\mathbf{X}}) = \mathbb{R}_+^n$. La mesure de référence ν est construite via le produit de n mesures de probabilité ν_i associée à la loi gamma $\Gamma(m_i, r_i)$ pour $i = 1, \dots, n$. La mesure de probabilité s'écrit

$$\nu(x_1, \dots, x_n) = \nu_1(x_1) \times \dots \times \nu_n(x_n). \quad (2.61)$$

Les polynômes orthonormaux par rapport à cette mesure sont définis par

$$Q_{\mathbf{k}}(\mathbf{x}) = \prod_{i=1}^n Q_{k_i}^{\nu_i}(x_i), \quad (2.62)$$

où $Q_{k_i}^{\nu_i}(x_i)$ est le polynôme orthonormal de degré k_i , par rapport à la mesure ν_i , et défini dans l'équation (2.25) pour $i = 1, \dots, n$.

Si $\frac{dG_{\mathbf{X}}}{d\nu} \in L^2(\nu)$, alors la densité $g_{\mathbf{X}}$ de la mesure de probabilité défailante $G_{\mathbf{X}}$ admet la représentation polynomiale

$$\begin{aligned} g_{\mathbf{X}}(\mathbf{x}) &= \sum_{\mathbf{k} \in \mathbb{N}^n} a_{\mathbf{k}} Q_{\mathbf{k}}^{\nu}(\mathbf{x}) f_{\nu}(\mathbf{x}) \\ &= g_{\mathbf{X},\nu}(\mathbf{x}) f_{\nu}(\mathbf{x}), \end{aligned} \quad (2.63)$$

par application du Théorème 7. L'approximation est caractérisée par un vecteur d'ordres de troncature $\mathbf{K} = (K_1, \dots, K_n)$. L'ordre de troncature peut différer selon les directions de l'espace. L'approximation de la densité défailante $g_{\mathbf{X}}$ est obtenue par troncature multiple des séries infinies de la représentation polynomiale (2.63), avec

$$g_{\mathbf{X}}^{\mathbf{K}}(\mathbf{x}) = \sum_{k_1=0}^{K_1} \dots \sum_{k_n=0}^{K_n} a_{\mathbf{k}} Q_{\mathbf{k}}^{\nu}(\mathbf{x}) f_{\nu}(\mathbf{x}). \quad (2.64)$$

$$= g_{\mathbf{X},\nu}^{\mathbf{K}}(\mathbf{x}) f_{\nu}(\mathbf{x}). \quad (2.65)$$

L'intégration de l'approximation de la densité (2.64) permet d'approcher la f.d.r. jointe, la f.d.s. jointe ainsi que d'autres probabilités intéressantes suivant le problème étudié.

Vérification de la condition d'intégrabilité

La représentation polynomiale (2.63) est valide sous réserve que la condition d'intégrabilité

$$\int_{\mathbb{R}^n} \left[\frac{dG_{\mathbf{X}}}{d\nu}(\mathbf{x}) \right]^2 d\nu(\mathbf{x}) < +\infty,$$

qui est équivalent à

$$\int_0^{+\infty} \dots \int_0^{+\infty} [g_{\mathbf{X}}(x_1, \dots, x_n)]^2 \exp\left(\sum_{i=1}^n \frac{x_i}{m_i}\right) \prod_{i=1}^n x_i^{1-r_i} dx_1 \dots dx_n < +\infty, \quad (2.66)$$

soit satisfaite. Le Proposition 6 doit être adaptée en dimension supérieure à 1. Il est nécessaire d'introduire une relation d'ordre sur les vecteurs de façon à donner un sens à l'infimum

$$\inf\{\mathbf{s} \in \mathbb{R}_{+*}^n : \mathcal{L}_{\mathbf{X}}(\mathbf{s}) = +\infty\}.$$

Définition 11. Soient $\mathbf{s} = (s_1, \dots, s_n)$ et $\mathbf{t} = (t_1, \dots, t_n)$ deux vecteurs de \mathbb{R}_{+*}^n . Le vecteur \mathbf{s} est supérieur au vecteur \mathbf{t} , avec

$$\mathbf{s} \geq \mathbf{t}$$

si

$$\mathbf{s} \cdot \mathbf{w} \geq \mathbf{t} \cdot \mathbf{w}, \quad \forall \mathbf{w} \in \mathbb{R}_{+*}^n. \quad (2.67)$$

Ce qui équivaut à $s_i \geq t_i$ pour $i = 1, \dots, n$.

Une conséquence immédiate est que si $\mathbf{s} \geq \mathbf{t}$ alors

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) \geq \mathcal{L}_{\mathbf{X}}(\mathbf{t}). \quad (2.68)$$

L'ensemble $\Gamma_{\mathbf{X}} = \inf\{\mathbf{s} \in \mathbb{R}_{+*}^n : \mathcal{L}_{\mathbf{X}}(\mathbf{s}) = +\infty\}$ est l'ensemble des minorants maximaux de $\{\mathbf{s} \in \mathbb{R}_{+*}^n : \mathcal{L}_{\mathbf{X}}(\mathbf{s}) = +\infty\}$. Soit $\boldsymbol{\gamma}_{\mathbf{X}} \in \Gamma_{\mathbf{X}}$. Soit $\mathbf{s} \in \mathbb{R}_{+*}^n$, si $\mathbf{s} \leq \boldsymbol{\gamma}_{\mathbf{X}}$ alors $\mathcal{L}_{\mathbf{X}}(\mathbf{s}) < +\infty$. La proposition suivante généralise la Proposition 6 en dimension supérieure à 1.

Proposition 9. Soit \mathbf{X} un vecteur aléatoire réel, continu, dont les composantes sont strictement positives, et de loi de probabilité $\mathbb{P}_{\mathbf{X}}$. La densité de probabilité du vecteur \mathbf{X} est notée $f_{\mathbf{X}}$. Soit $\Gamma_{\mathbf{X}} = \inf\{\mathbf{s} \in \mathbb{R}_{+,*}^n : \mathcal{L}_{\mathbf{X}}(\mathbf{s}) = +\infty\}$, l'ensemble limite de vecteurs pour lesquels la f.g.m.m. de \mathbf{X} n'est pas définie. Si $\Gamma_{\mathbf{X}}$ est non vide, et qu'il existe un vecteur $\mathbf{a} = (a_1, \dots, a_n)$ tel que pour tout $\mathbf{x} \geq \mathbf{a}$, les applications $t \mapsto f_{\mathbf{X}}(x_1, \dots, t, \dots, x_n)$ sont strictement décroissantes pour $i = 1, \dots, n$, alors

$$f_{\mathbf{X}}(\mathbf{x}) \leq A_{\mathbf{X}}(\mathbf{s}) e^{-\mathbf{x} \cdot \mathbf{s}}, \quad \mathbf{s} \leq \gamma_{\mathbf{X}} \quad (2.69)$$

où $\gamma_{\mathbf{X}} \in \Gamma_{\mathbf{X}}$.

Démonstration. Soit $\mathbf{x} \geq \mathbf{a}$, la f.g.m.m. du vecteur aléatoire \mathbf{X} est minorée par

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) \geq \int_{a_1}^{x_1} \dots \int_{a_{n-1}}^{x_{n-1}} e^{s_1 y_1 + \dots + s_{n-1} y_{n-1}} \left(\int_{a_n}^{x_n} e^{s_n y_n} f_{\mathbf{X}}(y_1, \dots, y_n) dy_n \right) dy_{n-1} \dots dy_1. \quad (2.70)$$

La partie entre parenthèses du second membre de l'équation (2.70) est minorée via une intégration par partie, avec

$$\begin{aligned} \int_{a_n}^{x_n} e^{s_n y_n} f_{\mathbf{X}}(y_1, \dots, y_n) dy_n &= \left[\frac{e^{s_n y_n}}{s_n} f_{\mathbf{X}}(y_1, \dots, y_n) \right]_{a_n}^{x_n} - \frac{1}{s_n} \int_{a_n}^{x_n} e^{s_n y_n} \frac{\partial}{\partial y_n} f_{\mathbf{X}}(y_1, \dots, y_n) dy_n \\ &\geq \frac{1}{s_n} \left[e^{s_n x_n} f_{\mathbf{X}}(y_1, \dots, y_{n-1}, x_n) - e^{a_n s_n} f_{\mathbf{X}}(y_1, \dots, y_{n-1}, a_n) \right]. \end{aligned} \quad (2.71)$$

La procédure de minoration (2.71) est appliquée dans chacune des directions de l'espace.

Soit $\Omega_n = \{0, 1\}^n$ l'ensemble des vecteurs de taille n , composés exclusivement de 0 et de 1. Soit $\mathbf{e}_n = (1, \dots, 1)$ le vecteur de taille n dont toutes les composantes sont égales à 1. Soit

$$\Omega_{i,n} = \{\mathbf{w} \in \Omega_n : \mathbf{w} \cdot \mathbf{e}_n = i\},$$

l'ensemble des vecteurs de $\Omega_n = \{0, 1\}^n$ dont i composantes sont égales à 1. La matrice identité de taille n est notée \mathbf{I}_n , et la matrice $\mathbf{W} = \text{Diag}(\mathbf{w})$ est la matrice diagonale dont les coefficients diagonaux sont les composantes du vecteur \mathbf{w} . Le vecteur défini par $\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}$ est un vecteur de taille n comprenant i composantes du vecteur \mathbf{x} et $n - i$ composantes du vecteur \mathbf{a} . La répétition de l'étape de minoration (2.71) conduit à

$$\begin{aligned} \mathcal{L}_{\mathbf{X}}(\mathbf{s}) &\geq \frac{1}{\prod_{i=1}^n s_i} \sum_{i=0}^n (-1)^{n-i} \sum_{\mathbf{w} \in \Omega_{i,n}} \exp\{[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \cdot \mathbf{s}\} \\ &\times f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \end{aligned} \quad (2.72)$$

$$\begin{aligned} &= \frac{1}{\prod_{i=1}^n s_i} \sum_{i=1}^{n-1} (-1)^{n-i} \sum_{\mathbf{w} \in \Omega_{i,n}} \exp\{[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \cdot \mathbf{s}\} \\ &\times f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \\ &+ \frac{(-1)^n}{\prod_{i=1}^n s_i} e^{\mathbf{s} \cdot \mathbf{a}} f_{\mathbf{X}}(\mathbf{a}) + \frac{1}{\prod_{i=1}^n s_i} e^{\mathbf{s} \cdot \mathbf{x}} f_{\mathbf{X}}(\mathbf{x}). \end{aligned} \quad (2.73)$$

Il en résulte une minoration de la densité $f_{\mathbf{X}}$ avec

$$\begin{aligned} f_{\mathbf{X}}(\mathbf{x}) &\leq \left[\mathcal{L}_{\mathbf{X}}(\mathbf{s}) \prod_{i=1}^n s_i - (-1)^n e^{\mathbf{s} \cdot \mathbf{a}} f_{\mathbf{X}}(\mathbf{a}) \right] e^{-\mathbf{x} \cdot \mathbf{s}} \\ &\quad - \sum_{i=1}^{n-1} (-1)^{n-i} \sum_{\mathbf{w} \in \Omega_{i,n}} \exp[(\mathbf{I}_n - \mathbf{W})\mathbf{a} \cdot \mathbf{s}] \exp[-(\mathbf{I}_n - \mathbf{W})\mathbf{x} \cdot \mathbf{s}] \\ &\quad \times f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}]. \end{aligned} \quad (2.74)$$

Soit $\mathbf{w} \in \Omega_{i,n}$, avec $i = 1, \dots, n-1$. L'application

$$\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a} \mapsto f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \quad (2.75)$$

est une application définie de \mathbb{R}_+^i dans \mathbb{R}_+ . Comme $\mathcal{L}_{\mathbf{X}}(\mathbf{s}) < +\infty$ alors par application du théorème de Fubini-Tonelli, l'application

$$\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a} \mapsto \exp[\mathbf{W}\mathbf{x} \cdot \mathbf{s}] f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \quad (2.76)$$

est intégrable sur \mathbb{R}_+^i . Le problème de départ était la majoration d'une fonction de n variables. Les étapes de minoration successives conduisent à réduire la dimension du problème, car des fonctions de $n-i$ variables, avec $i=1, \dots, n-1$, vérifiant les mêmes hypothèses que $f_{\mathbf{X}}$ sont à majorer. Un raisonnement par récurrence sur i , dont l'initialisation au rang $i = 1$ est consécutive à l'application de la Proposition 6, s'applique pour montrer que

$$f_{\mathbf{X}}[\mathbf{W}\mathbf{x} + (\mathbf{I}_n - \mathbf{W})\mathbf{a}] \leq A_{\mathbf{w}}(\mathbf{s}) e^{-\mathbf{W}\mathbf{x} \cdot \mathbf{s}}, \quad \forall \mathbf{w} \in \Omega_{i,n}. \quad (2.77)$$

L'utilisation de l'inégalité (2.77) dans l'inégalité (2.74) conduit à la minoration de la densité $f_{\mathbf{X}}$ par

$$f_{\mathbf{X}}(\mathbf{x}) \leq \left[\mathcal{L}_{\mathbf{X}}(\mathbf{s}) \prod_{i=1}^n s_i - \sum_{i=0}^{n-1} (-1)^{n-i} \sum_{\mathbf{w} \in \Omega_{i,n}} \exp\{[\mathbf{I}_n - \mathbf{W}]\mathbf{a} \cdot \mathbf{s}\} A_{\mathbf{w}}(\mathbf{s}) \right] e^{-\mathbf{s} \cdot \mathbf{x}}. \quad (2.78)$$

□

La Proposition 9 est une généralisation à la dimension n du résultat énoncé au Chapitre 4 pour $n = 2$, les raisonnements employés sont identiques. La Proposition 9 est applicable à la majoration d'une densité de probabilité défaillante. Le résultat suivant éclaire le choix des paramètres $\mathbf{m} = (m_1, \dots, m_n)$ et $\mathbf{r} = (r_1, \dots, r_n)$ pour la mesure de référence (2.61).

Corollaire 3. *Si la densité défaillante $\mathbf{x} \mapsto g_{\mathbf{X}}(\mathbf{x})$ vérifie les hypothèses de la Proposition 9 alors la paramétrisation*

$$\mathbf{r} \leq \mathbf{e}_n ; \mathbf{m}^{-1} < 2\boldsymbol{\gamma}_{\mathbf{X}}, \quad (2.79)$$

où $\mathbf{m}^{-1} = \left(\frac{1}{m_1}, \dots, \frac{1}{m_n}\right)$ assure la vérification de la condition d'intégrabilité (2.66).

Démonstration. Soit $\boldsymbol{\Gamma}_{\mathbf{X}} = \inf\{\mathbf{s} \in \mathbb{R}_{+*}^n : \mathcal{L}_{\mathbf{X}}(\mathbf{s}) = +\infty\}$ et $\boldsymbol{\gamma}_{\mathbf{X}} \in \boldsymbol{\Gamma}_{\mathbf{X}}$. Soit $\mathbf{s} \in \mathbb{R}_{+*}^n$ tel que $\mathbf{s} < \boldsymbol{\gamma}_{\mathbf{X}}$. L'application de la Proposition 9 permet de majorer le membre à gauche du signe équivalent de la condition d'intégrabilité (2.66), avec

$$\int_{\mathbb{R}_+^n} g_{\mathbf{X}}^2(\mathbf{x}) \exp[\mathbf{m}^{-1} \cdot \mathbf{x}] \prod_{i=1}^n x_i^{1-r_i} d\mathbf{x} < \int_{\mathbb{R}_+^n} \exp[-(2\mathbf{s} - \mathbf{m}^{-1}) \cdot \mathbf{x}] \prod_{i=1}^n x_i^{1-r_i} d\mathbf{x}.$$

La condition d'intégrabilité (2.66) est vérifiée si $\mathbf{m}^{-1} < 2\boldsymbol{\gamma}_{\mathbf{X}}$ et $\mathbf{r} < \mathbf{e}_n$. □

La fonction génératrice des coefficients de la représentation polynomiale

Soit \mathbf{X} un vecteur aléatoire réel, continu et positif de loi $\mathbb{P}_{\mathbf{X}}$, et de densité de probabilité $f_{\mathbf{X}}$. Si $\frac{d\mathbb{P}_{\mathbf{X}}}{d\nu} \in L^2(\nu)$. La densité $f_{\mathbf{X}}$ admet la représentation polynomiale

$$f_{\mathbf{X}}(\mathbf{x}) = \sum_{\mathbf{k} \in \mathbb{N}^n} a_{\mathbf{k}} Q_{\mathbf{k}}(\mathbf{x}) f_{\nu}(\mathbf{x}). \quad (2.80)$$

Par analogie avec le calcul (2.45) effectué en dimension 1, la transformée de Laplace de la représentation polynomiale (7) est donnée par

$$\mathcal{L}_{\mathbf{X}}(\mathbf{s}) = \mathcal{C} \left(\frac{s_1 m_1}{1 - s_1 m_1}, \dots, \frac{s_n m_n}{1 - s_n m_n} \right) \prod_{i=1}^n \left(\frac{1}{1 - s_i m_i} \right), \quad (2.81)$$

où $\mathbf{z} = (z_1, \dots, z_n) \mapsto \mathcal{C}(\mathbf{z}) = \sum_{k_1=0}^{+\infty} \dots \sum_{k_n=0}^{+\infty} a_{k_1 \dots k_n} c_{k_1}^{\nu_1} \dots c_{k_n}^{\nu_n} z_1^{k_1} \dots z_n^{k_n}$ est la fonction génératrice multivariée de la séquence $\{a_{\mathbf{k}} \prod_{i=1}^n c_{k_i}^{\nu_i}\}_{\mathbf{k} \in \mathbb{N}^n}$. La fonction génératrice des coefficients de la représentation polynomiale s'exprime en fonction de la **f.g.m.m.** du vecteur aléatoire \mathbf{X} , avec

$$\mathcal{C}(\mathbf{z}) = (1 + z_1)^{-r_1} \times \dots \times (1 + z_n)^{-r_n} \times \mathcal{L}_{\mathbf{X}} \left(\frac{z_1}{m_1(1 + z_1)}, \dots, \frac{z_n}{m_n(1 + z_n)} \right). \quad (2.82)$$

Les coefficients $\{a_{\mathbf{k}}\}_{\mathbf{k} \in \mathbb{N}^n}$ de la représentation polynomiale (2.80) s'obtiennent par dérivation de la fonction génératrice (2.82), avec

$$a_{\mathbf{k}} = \frac{1}{\prod_{i=1}^n k_i! c_{k_i}^{\nu_i}} \left[\frac{d^{k_1} \dots d^{k_n}}{dz_1 \dots dz_n} \mathcal{C}(\mathbf{z}) \right]_{\mathbf{z}=(0, \dots, 0)}. \quad (2.83)$$

L'usage de la fonction génératrice des coefficients (2.82) est le même qu'en dimension 1. La forme de celle-ci en dimension n est cependant plus complexe ce qui rend le choix des paramètres parfois difficile.

2.3.3 Cas $n = 2$: La connexion avec les probabilités de Lancaster

Un problème classique en statistique, connu sous le nom de problème de Lancaster, voir **LANCASTER [1958]** et **LANCASTER et SENETA [2005]**, est la construction de lois de probabilité sur \mathbb{R}^2 de lois marginales données. Soit ν une mesure de probabilité bivariée définie par

$$\nu(x, y) = \nu_1(x) \times \nu_2(y), \quad (2.84)$$

où ν_1 et ν_2 sont des mesures de probabilité appartenant aux **FENQ**. Les mesures ν_1 et ν_2 sont associées à des séquences de polynômes orthonormaux notées respectivement $\{Q_k^{\nu_1}\}_{k \in \mathbb{N}}$ et $\{Q_k^{\nu_2}\}_{k \in \mathbb{N}}$. Les mesures de probabilités de Lancaster bivariées échangeables admettent la forme

$$\sigma(x, y) = \left(\sum_{k=1}^{+\infty} a_k Q_k^{\nu_1}(x) Q_k^{\nu_2}(y) \right) \nu(x, y), \quad (2.85)$$

où $\{a_k\}_{k \in \mathbb{N}}$ est une suite de réels telle que

$$a_k = \int \int Q_k^{\nu_1}(x) Q_k^{\nu_2}(y) d\nu(x, y),$$

avec $a_0 = 1$. La mesure (2.85) est une mesure de probabilité sous réserve de la positivité de (2.85), plus précisément

$$\left(\sum_{k=1}^{+\infty} a_k Q_k^{v_1}(x) Q_k^{v_2}(y) \right) \geq 0, \quad \forall x, y \in \mathbb{R}^2. \quad (2.86)$$

Ainsi le problème de Lancaster se résume à trouver des séquences $\{a_k\}_{k \in \mathbb{N}}$ permettant de vérifier la condition (2.86). La loi de probabilité (2.85) est celle d'un vecteur aléatoire dont les composantes sont distribuées selon v_1 et v_2 . La structure de corrélation est caractérisée par la partie entre parenthèses de (2.85) qui s'apparente à une densité de copule. EAGLESON [1964] a montré que le vecteur $(W_1 + W_2, W_2 + W_3)$ admet une mesure de probabilité de Lancaster si W_1, W_2 et W_3 sont des variables aléatoires indépendantes distribuées suivant une loi appartenant aux FENQ. Il s'agit d'une méthode de construction classique des probabilités de Lancaster. La caractérisation rigoureuse des probabilités de Lancaster, basées sur des lois marginales appartenant aux FENQ, et de leur séquences admissibles est effectuée dans les travaux de KOUDOU [1995, 1996]. En particulier, des combinaisons particulières de loi marginales, telles que Gamma-Gamma, Poisson-Poisson, Binomiale Négative-Binomiale Négative et même Gamma-Binomiale Négative sont traités dans l'article KOUDOU [1998]. Le cas binomial-binomial est étudié dans le travail de DIACONIS et GRIFFITHS [2012], où une interprétation probabiliste de ces lois est donnée. La modélisation de vecteur aléatoire à l'aide des probabilités de Lancaster facilite la simulation par échantillonneur de Gibbs, voir DIACONIS et collab. [2008]. Il est aussi simple de calculer l'indicateur de corrélation introduit dans SZÉKELY et collab. [2007], le lecteur peut se référer au pré-print de DUECK et collab. [2015].

Afin de relier l'approximation polynomiale aux probabilités de Lancaster, le cas d'un vecteur aléatoire (X, Y) est considéré. Sa loi de probabilité $\mathbb{P}_{X,Y}$ est supposée absolument continue par rapport à la mesure ν , définie dans l'équation (2.84). Si $\frac{d\mathbb{P}_{X,Y}}{d\nu} \in L^2(\nu)$, alors l'application du Théorème 7 permet d'écrire

$$\mathbb{P}_{X,Y}(x, y) = \sum_{k=1}^{+\infty} \sum_{l=1}^{+\infty} a_{k,l} Q_k^{v_1}(x) Q_l^{v_2}(y) \nu(x, y), \quad (2.87)$$

où

$$a_{k,l} = \mathbb{E} [Q_k^{v_1}(X) Q_l^{v_2}(Y)], \quad k, l \in \mathbb{N}.$$

Les mesures de probabilités de Lancaster (2.85) sont des cas particuliers des mesures de probabilité (2.87). Les mesures de probabilité (2.87) sont des lois de Lancaster lorsque la condition de bi-orthogonalité

$$\mathbb{E} [Q_k^{v_1}(X) Q_l^{v_2}(Y)] = a_k \delta_{kl} \quad (2.88)$$

est vérifiée. L'application de la méthode d'approximation polynomiale pour retrouver la loi de probabilité d'un vecteur aléatoire gouverné par une loi de Lancaster est facilitée puisque K fois moins de coefficients sont à évaluer, où K désigne l'ordre de troncature de l'approximation polynomiale. L'utilisation de la fonction génératrice des coefficients (2.81) permet de constater qu'une famille de lois de probabilité exponentielles bivariées absolument continues sur $\mathbb{R}_{+,*}^2$ sont en fait des lois de Lancaster.

Définition 12. Un vecteur aléatoire (X, Y) est de loi Lancaster Gamma-Gamma $\mathcal{L}_{\Gamma-\Gamma}(\alpha, p, \beta_1, \beta_2)$ si le vecteur aléatoire (X, Y) est défini par

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \sum_{j=1}^{N+\alpha} \begin{pmatrix} U_j \\ V_j \end{pmatrix}, \quad (2.89)$$

où N est une variable aléatoire de comptage de loi binomiale négative $\mathcal{NB}(\alpha, p)$, et la suite $\{(U_i, V_i)\}_{i \in \mathbb{N}}$ est une suite de vecteurs aléatoires *i.i.d.*, composés de deux variables aléatoires indépendantes de loi exponentielle $\Gamma(1, \beta_1)$ et $\Gamma(1, \beta_2)$. La variable aléatoire N est indépendante de la suite de vecteurs aléatoires $\{(U_i, V_i)\}_{i \in \mathbb{N}}$. La loi de probabilité $\mathbb{P}_{X,Y}$ du vecteur aléatoire (2.89) est une loi de probabilité absolument continue de \mathbb{R}_{+*}^2 .

Proposition 10. Soit (X, Y) un vecteur aléatoire de loi $\mathcal{L}_{\Gamma-\Gamma}(\alpha, p, \beta_1, \beta_2)$. La *f.g.m.m.* du vecteur (X, Y) est donnée par

$$\mathcal{L}_{X,Y}(s, t) = \left[1 - \frac{s\beta_1}{1-p} - \frac{t\beta_2}{1-p} - \frac{st\beta_1\beta_2}{1-p} \right]^{-\alpha} \quad (2.90)$$

Démonstration. Soit (X, Y) un vecteur aléatoire de loi $\mathcal{L}_{\Gamma-\Gamma}(\alpha, p, \beta_1, \beta_2)$, sa *f.g.m.m.* est définie par

$$\begin{aligned} \mathcal{L}_{X,Y}(s, t) &= [\mathcal{L}_{U,V}(s, t)]^\alpha \mathcal{G}_N[\mathcal{L}_{U,V}(s, t)] \\ &= \left[\frac{1}{(1-s\beta_1)(1-t\beta_2)} \right]^\alpha \mathcal{G}_N \left[\frac{1}{(1-s\beta_1)(1-t\beta_2)} \right] \\ &= (1-p)^\alpha \left[\frac{1}{(1-s\beta_1)(1-t\beta_2)} \right]^\alpha \left[\frac{(1-s\beta_1)(1-t\beta_2)}{(1-s\beta_1)(1-t\beta_2)-p} \right]^\alpha \\ &= \left[1 - \frac{s\beta_1}{1-p} - \frac{t\beta_2}{1-p} - \frac{st\beta_1\beta_2}{1-p} \right]^{-\alpha}. \end{aligned}$$

□

Pour montrer que la loi du vecteur aléatoire (2.89) est une loi de Lancaster, il faut montrer que sa loi de probabilité $\mathbb{P}_{X,Y}$ admet la forme (2.85). La démonstration consiste à définir la représentation polynomiale (2.87) de la loi $\mathbb{P}_{X,Y}$ via le Théorème 7 en utilisant comme mesure de référence le produit de deux mesures gamma. Le choix des paramètres des mesures gamma est effectué en étudiant la fonction génératrice des coefficients (2.82) de la représentation polynomiale. Il est ainsi démontré que la loi du vecteur aléatoire (X, Y) est une loi de Lancaster du type gamma-gamma.

Proposition 11. Soit (X, Y) un vecteur aléatoire de loi $\mathcal{L}_{\Gamma-\Gamma}(\alpha, p, \beta_1, \beta_2)$. La loi de probabilité de (X, Y) est une loi de Lancaster, avec

$$\mathbb{P}_{X,Y}(x, y) = \sum_{k=0}^{\infty} a_k Q_k^{v_1}(x) Q_k^{v_2}(y) v_1(x) v_2(y), \quad (2.91)$$

où

$$a_k = \frac{p^k \Gamma(\alpha + k - 1)}{k! \Gamma(\alpha)} (c_k^{v_1} c_k^{v_2})^{-1}, \quad k \in \mathbb{N}. \quad (2.92)$$

Les mesures v_1 et v_2 sont des mesures de probabilité gamma $\Gamma\left(\alpha, \frac{\beta_1}{1-p}\right)$ et $\Gamma\left(\alpha, \frac{\beta_2}{1-p}\right)$, et les suites $\{Q_k^{v_1}\}_{k \in \mathbb{N}}$ et $\{Q_k^{v_2}\}_{k \in \mathbb{N}}$ sont les suites de polynômes orthonormaux par rapport à v_1 et à v_2 respectivement.

Démonstration. Soit (X, Y) un vecteur aléatoire de loi $\mathcal{L}_{\Gamma-\Gamma}(\alpha, p, \beta_1, \beta_2)$. Si $\frac{d\mathbb{P}_{X,Y}}{d\nu} \in L^2(\nu)$ alors la mesure de probabilité $\mathbb{P}_{X,Y}$ admet la représentation polynomiale

$$\mathbb{P}_{X,Y}(x, y) = \sum_{k=1}^{+\infty} \sum_{l=1}^{+\infty} a_{k,l} Q_k^{\nu_1}(x) Q_l^{\nu_2}(y) \nu_1(x) \nu_2(y), \quad (2.93)$$

par application du Théorème 7. La mesure ν_1 est une mesure de probabilité gamma $\Gamma(m_1, r_1)$ et ν_2 est une mesure de probabilité gamma $\Gamma(m_2, r_2)$. Le choix des paramètres m_1, r_1, m_2 et r_2 est effectué via la fonction génératrice des coefficients de la représentation polynomiale (2.93). La f.g.m.m. du vecteur aléatoire (X, Y) est donnée par

$$\mathcal{L}_{X,Y}(s, t) = \left[1 - \frac{s\beta_1}{1-p} - \frac{t\beta_2}{1-p} - \frac{st\beta_1\beta_2}{1-p} \right]^{-\alpha}. \quad (2.94)$$

En injectant la f.g.m.m. (2.94) dans l'expression de la fonction génératrice des coefficients (2.81), il vient

$$\mathcal{C}(z_1, z_2) = \left[1 - \frac{z_1\beta_1}{m_1(1+z_1)(1-p)} - \frac{z_2\beta_2}{m_2(1+z_2)(1-p)} - \frac{z_1z_2\beta_1\beta_2}{m_1m_2(1+z_1)(1+z_2)(1-p)} \right]^{-\alpha}. \quad (2.95)$$

En prenant $m_1 = \frac{\beta_1}{(1-p)}$ et $m_2 = \frac{\beta_2}{1-p}$, la fonction génératrice des coefficients (2.95) se simplifie, avec

$$\mathcal{C}(z_1, z_2) = \frac{(1+z_1)^{\alpha-r_1} (1+z_2)^{\alpha-r_2}}{[(1+z_1)(1+z_2) - z_1(1+z_2) - z_2(1+z_1) + (1-p)z_1z_2]^\alpha}. \quad (2.96)$$

En prenant $r_1 = r_2 = \alpha$, la fonction génératrice des coefficients (2.96) se simplifie, avec

$$\mathcal{C}(z_1, z_2) = \left(\frac{1}{1 - z_1z_2p} \right)^\alpha. \quad (2.97)$$

Les coefficients du développement, définis par (2.83), sont donnés par

$$a_{k,l} \doteq a_k = \frac{p^k \Gamma(\alpha + k - 1)}{k! \Gamma(\alpha)} (c_k^{\nu_1} c_k^{\nu_2})^{-1}, \quad \forall k \in \mathbb{N}. \quad (2.98)$$

□

Le résultat 11 donne un moyen de construire des lois de Lancaster de type gamma-gamma autre qu'à la mode *random element in common* qui est la méthode classique. Cette construction est associée à une interprétation probabiliste et basée sur la composition d'un vecteur aléatoire (U, V) , dont les composantes sont indépendantes et de loi exponentielle, par une variable aléatoire de comptage N de loi binomiale négative. Les paramètres α et p de la loi $\mathcal{NB}(\alpha, p)$ vont permettre d'ajuster la corrélation souhaitée pour le vecteur aléatoire (X, Y) .

2.4 Illustrations numériques : Application à la distribution Poisson composée

L'objectif principal de cette section est de montrer l'efficacité de la méthode d'approximation par les polynômes orthogonaux. Soit X une variable aléatoire distribuée

suivant une loi de probabilité composée $(\mathbb{P}_N, \mathbb{P}_U)$. La méthode d'approximation est utilisée pour approcher la densité défailante g_X , la f.d.s. \overline{F}_X , et la prime *stop-loss* généralisée $\Pi_{c,d}(X)$. La précision de l'approximation est mesurée par un critère d'erreur relative par rapport à la valeur exacte ou à défaut une approximation *benchmark*. Soit \tilde{f} , une approximation de la fonction f . L'erreur relative de \tilde{f} , en $x \in \mathbb{R}$, est définie par

$$\Delta \tilde{f}(x) = \frac{\tilde{f}(x) - f(x)}{f(x)}. \quad (2.99)$$

La densité défailante associée à la loi de probabilité composée $(\mathbb{P}_N, \mathbb{P}_U)$ est approchée par

$$g_X^K(x) = \sum_{k=0}^K a_k Q_k(x) f_V(x), \quad x \geq 0, \quad (2.100)$$

où K désigne l'ordre de troncature. Les coefficients du développement polynomial $\{a_k\}_{k \in \mathbb{N}}$ sont définis par $a_k = \mathbb{E}[Q_k(X)]$, pour $k = 1, \dots, K$, et calculés via la formule (2.54). La suite de polynômes $\{Q_k\}_{k \in \mathbb{N}}$, définie dans l'équation (2.25), forme un système orthonormal de polynômes par rapport à une mesure de probabilité gamma $\Gamma(r, m)$ dont f_V désigne la densité de probabilité. L'approximation de la f.d.s. de X s'obtient en intégrant l'approximation de la densité défailante (2.100), avec

$$\overline{F}_X^K(x) = \int_x^{+\infty} g_X^K(y) dy, \quad x \geq 0. \quad (2.101)$$

L'approximation de la prime *stop-loss* généralisée s'obtient en réinjectant l'approximation de la f.d.s. (2.101) dans l'expression de la prime *stop-loss* (1.7), avec

$$\Pi_{c,d}(X) = \int_0^{+\infty} y^{d-1} \overline{F}_X^K(y+c) dy \quad c \geq 0, \quad d \geq 0. \quad (2.102)$$

Seule la prime *stop-loss* usuelle, caractérisée par $d = 1$, est considérée dans les illustrations. L'application de la méthode d'approximation polynomiale passe par le choix des paramètres m et r . Ces paramètres sont choisis en fonction de la forme de la fonction génératrice des coefficients donnée par l'équation (2.46). Cette fonction, suivant les cas, peut prendre une forme très compliquée. Dans ce cas, les paramètres m et r sont choisis conformément au Corollaire 2. Dans les exemples considérés, cette façon de paramétrer revient à rapprocher la mesure de référence de la loi \mathbb{P}_U gouvernant le montant des sinistres en faisant coïncider leurs moments respectifs. L'approximation décrite dans JIN et collab. [2014] est très proche de celle présentée dans ce chapitre. Les paramètres sont choisis de façon à faire coïncider les moments d'ordre 1 et 2 de la variable aléatoire X avec ceux de la mesure de référence. Il en résulte la paramétrisation suivante

$$m = \frac{\mathbb{V}(X)}{\mathbb{E}(X)}; \quad r = \frac{\mathbb{E}(X)^2}{\mathbb{V}(X)}. \quad (2.103)$$

Une paramétrisation plus naïve que la précédente consiste à prendre une mesure de référence exponentielle, avec $r = 1$, et à faire coïncider simplement le moment d'ordre 1. Il en résulte la paramétrisation suivante

$$m = \mathbb{E}(X); \quad r = 1. \quad (2.104)$$

Les paramétrisations (2.103) et (2.104) sont testées dans chacun des exemples étudiés même si elles rentrent parfois en contradiction avec le Corollaire 2.

Lorsque la valeur exacte de la fonction approchée n'est pas disponible pour la f.d.s., la méthode d'inversion de la transformée de Fourier est utilisée en guise de référence. L'erreur associée à cette méthode est connue et peut être contrôlée, voir ABATE et WHITT [1992]. La formule d'approximation (1.61) est utilisée avec $a = 18.5$, $M = 11$ et $K = 15$, conformément au Chapitre 5 de l'ouvrage de ROLSKI et collab. [1999]. L'approximation polynomiale est aussi comparée aux approximations renvoyées par la méthode des moments exponentiels et l'algorithme de Panjer.

L'application de la méthode des moments exponentiels conduit à utiliser la formule d'approximation de la f.d.s. basée sur la formule d'approximation de la f.d.r. (2.105), avec

$$\bar{F}_X^{n,b}(x) = \sum_{k=0}^{\lfloor ne^{-\ln(b)x} \rfloor} \sum_{j=k}^n \binom{n}{j} \binom{j}{k} (-1)^{j-k} \mathcal{L}[-j \ln(b)], \quad (2.105)$$

où $n = 32$, et $b = 1.115$, ce qui est conforme aux paramètres utilisés dans le papier de MNATSAKANOV et SARKISIAN [2013].

L'approximation de la f.d.s. par l'algorithme de Panjer est obtenue via la formule (1.30). La loi des sinistres est discrétisée suivant le procédé (1.25) avec un pas de discrétisation $h = \frac{1}{100}$.

L'approximation polynomiale de la prime *stop-loss* usuelle est comparée à une valeur de référence obtenue par une méthode de Monte-Carlo classique pour approcher l'espérance mathématique (1.4), avec $c = 1$. Un échantillon (X_1, \dots, X_n) de taille $n = 10^6$ issu de la loi de X est produit pour évaluer la prime *stop-loss* usuelle par

$$\Pi_{1,c}(X) \approx \frac{1}{n} \sum_{i=1}^n (X_i - c)_+.$$

Les résultats et graphiques présentés dans cette section ont été obtenus en utilisant le logiciel Mathematica.

Soit X une variable aléatoire distribuée suivant une loi de probabilité composée $(\mathbb{P}_N, \mathbb{P}_U)$, où N est une variable aléatoire de comptage distribuée suivant une loi de Poisson $\mathcal{P}(\lambda)$. La f.g.m. d'une variable aléatoire de loi de probabilité composée $[\mathcal{P}(\lambda), \mathbb{P}_U]$ est donnée par

$$\mathcal{L}_X(s) = e^{\lambda[\mathcal{L}_U(s)-1]}, \quad (2.106)$$

où \mathcal{L}_U désigne la f.g.m. de la variable aléatoire U .

2.4.1 Le cas des sinistres de loi exponentielle $\Gamma(1, \beta)$

Les montants de sinistre sont modélisés par une suite $\{U_i\}_{i \in \mathbb{N}}$ de variables aléatoires i.i.d. suivant une loi exponentielle $\Gamma(1, \beta)$. Ce cas particulier est intéressant car la den-

sité défaillante g_X , de la variable aléatoire X , est donnée par

$$g_X(x) = \exp \left[- \left(\lambda + \frac{x}{\beta} \right) \right] 2 \sqrt{\frac{\lambda}{\beta x}} \mathbf{I}_1 \left(2 \sqrt{\frac{\lambda x}{\beta}} \right), \quad (2.107)$$

où $x \mapsto \mathbf{I}_1(x)$ est la fonction de Bessel modifiée de première espèce.

Définition 13. La fonction de Bessel modifiée de première espèce est définie par

$$\mathbf{I}_\alpha(x) = \sum_{k=0}^{+\infty} \frac{1}{k! \Gamma(k + \alpha + 1)} \left(\frac{x}{2} \right)^{2k + \alpha}, \quad x \in \mathbb{R}, \alpha \in \mathbb{R}.$$

La densité et la f.d.s. de la variable aléatoire X sont disponibles explicitement, ce qui va permettre de mesurer la précision des méthodes d'approximation.

La f.g.m. de X de loi composée $[\mathcal{P}(\lambda), \Gamma(1, \beta)]$ est donnée par

$$\mathcal{L}_X(s) = \exp \left[\lambda \left(\frac{1}{1 - s\beta} - 1 \right) \right], \quad (2.108)$$

ce qui implique que $\gamma_X = \frac{1}{\beta}$. La fonction génératrice des coefficients est obtenue par réinjection de la f.g.m. (2.108) dans (2.46), avec

$$\mathcal{C}(z) = (1 + z)^{-r} \exp \left\{ \lambda \left[\frac{m(1 + z) - [m + z(m - \beta)]}{m + z(m - \beta)} \right] \right\}. \quad (2.109)$$

La paramétrisation

$$m_1 = \beta; r_1 = 1, \quad (2.110)$$

simplifie la fonction génératrice des coefficients, avec

$$\mathcal{C}(z) = (1 + z)^{-1} \exp [\lambda(1 + z)].$$

La paramétrisation (2.110) est conforme au Corollaire 2. Pour l'illustration numérique, le paramètre de la loi de Poisson est égal à 4 et le paramètre de la loi exponentielle est égal à 2. La paramétrisation 2 est définie par l'équation (2.104), ce qui donne

$$m_2 = 8; r_2 = 1,$$

après applications numériques. La paramétrisation 3 est définie par l'équation (2.103), ce qui donne

$$m_3 = 5; r_3 = \frac{8}{5},$$

après applications numériques. La paramétrisation 2 est conforme au Corollaire 2, ce n'est pas le cas pour la paramétrisation 3. L'ordre de troncature est $K = 30$ pour chacune des trois paramétrisations. La Figure 2.1 montre l'erreur relative associée à l'approximation polynomiale de la densité de probabilité $g_X(x)$, en fonction des valeurs de x , suivant les trois paramétrisations. La Figure 2.2 montre l'erreur relative associée à l'approximation polynomiale de la f.d.s. $\bar{F}_X(x)$, en fonction des valeurs de x , suivant les

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE REPRÉSENTATION POLYNOMIALE

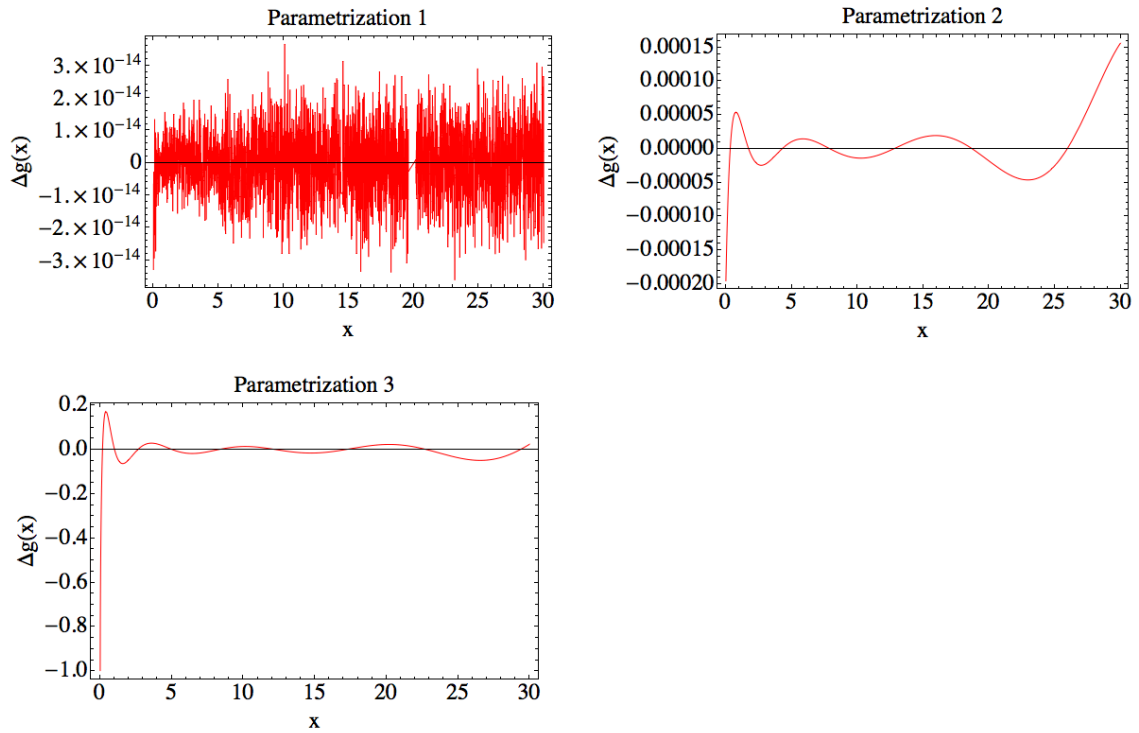


FIGURE 2.1 – Erreur relative de l’approximation polynomiale, avec différentes paramétrisations, de la densité de probabilité d’une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$.

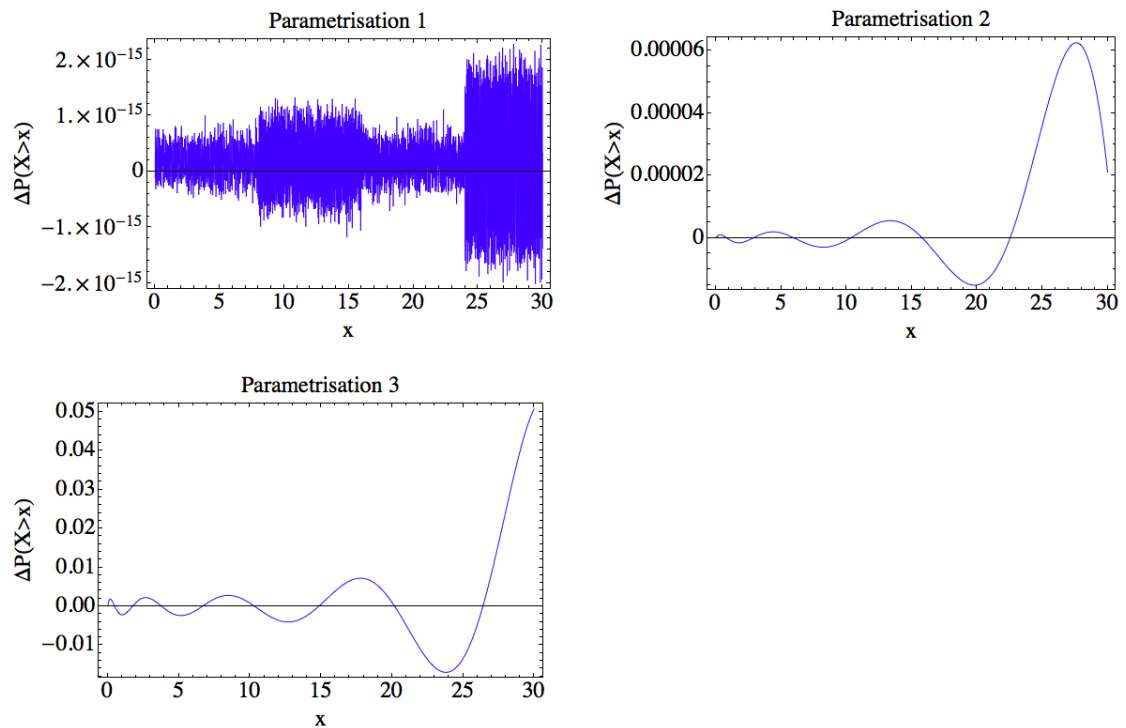


FIGURE 2.2 – Erreur relative de l’approximation polynomiale, avec différentes paramétrisations, de la f.d.s. d’une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$.

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE
REPRÉSENTATION POLYNOMIALE

TABLEAU 2.1 – f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \Gamma(1, 1/2)]$ approchée par la méthode d'approximation polynomiale suivant différentes paramétrisations.

x	Exact	Param. 1	Param. 2	Param. 3
3.	0.806382	0.806382	0.806382	0.807966
6.	0.573092	0.573092	0.573092	0.572258
9.	0.364357	0.364357	0.364356	0.365284
12.	0.21241	0.21241	0.212411	0.21165
15.	0.115555	0.115555	0.115555	0.115619
18.	0.0594094	0.0594094	0.0594088	0.0598344
21.	0.0291366	0.0291366	0.0291363	0.028992
24.	0.0137285	0.0137285	0.0137288	0.0134971
27.	0.00624886	0.00624886	0.00624924	0.00630265
30.	0.00275979	0.00275979	0.00275985	0.00289996

TABLEAU 2.2 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \Gamma(1, 2)]$ approchée par la méthode polynomiale suivant différentes paramétrisations.

x	Param. 1	Param. 2	Param. 3
3.	0.	0.00003	0.19645
6.	0.	0.	-0.14558
9.	0.	-0.00025	0.25437
12.	0.	0.00041	-0.3579
15.	0.	0.0003	0.05592
18.	0.	-0.00105	0.71535
21.	0.	-0.00124	-0.49631
24.	0.	0.00203	-1.6853
27.	0.	0.00607	0.8607
30.	0.	0.0021	5.07896

trois paramétrisations. Le Tableau 2.1 donne des valeurs numériques de la f.d.s. obtenues via l'approximation polynomiale en fonction de la paramétrisation. Le Tableau 2.2 donne des valeurs numériques de l'erreur relative sur la f.d.s. consécutives à l'approximation polynomiale en fonction de la paramétrisation.

La paramétrisation 1 permet une approximation nettement meilleure que les deux autres paramétrisations. La paramétrisation 2 est associée à des niveaux d'erreur acceptables qui semblent constants en fonction de x . Les résultats pour la paramétrisation 3 ne sont pas désastreux mais l'erreur semble augmenter avec x . En effet, le niveau d'erreur atteint 5% lorsque $x = 30$. Dans ce cas de figure, le respect du Corollaire 2 est indispensable pour obtenir de bons résultats comme le prouvent les erreurs obtenues avec les paramétrisations 1 et 2. La Figure 2.1 révèle aussi un problème en $x = 0$ pour l'approximation de la densité avec la paramétrisation 3. L'autre enseignement est qu'une bonne approximation passe par une mesure de référence proche de la mesure de probabilité gouvernant le montant des sinistres plutôt que proche de la mesure de probabilité de la distribution composée. Cette conclusion est intéressante car contre-intuitive. La méthode d'approximation polynomiale est parfois assimilée à une approximation paramétrique réajustée par un développement polynomial. Dans ce cas, il semblerait logique que la mesure de référence soit aussi proche que possible de la densité inconnue. Dans le cas de la distribution composée $[\mathcal{P}(\lambda), \Gamma(1, \beta)]$, l'expression de la densité (2.107) est très proche de la formule d'approximation polynomiale (2.28) avec la paramétrisation 1.

Ces remarques conduisent à opter pour la paramétrisation 1 pour comparer l'approximation polynomiale aux autres méthodes. La Figure 2.3 représente l'erreur relative sur la f.d.s. de la distribution composée $[\mathcal{P}(4), \Gamma(1, 2)]$ avec l'approximation polynomiale, l'approximation basée sur l'inversion de la transformée de Fourier, la méthode des moments exponentiels et l'algorithme de Panjer. Le Tableau 2.4 donne les valeurs numériques de ces erreurs relatives pour certaines valeurs de x . Le Tableau 2.3 donne les valeurs numériques de la f.d.s. obtenues via les différentes méthodes numériques.

L'approximation polynomiale est bien meilleure que ses concurrentes dans ce cas. L'inversion de la transformée de Fourier arrive en deuxième position avec de très bonnes performances également. L'algorithme de Panjer donne des résultats acceptables, l'erreur relative semble augmenter linéairement avec x . La méthode des moments exponentiels ferme la marche, l'erreur explose littéralement pour $x \geq 21$.

La Figure 2.4 représente la densité, la f.d.s., la f.d.r. et la prime *stop-loss* associées à la variable aléatoire X résultant de l'approximation polynomiale avec la paramétrisation 1 et un ordre de troncature $K = 30$. Le Tableau 2.5 donne les valeurs numériques de la prime *stop-loss* usuelle calculées via l'approximation polynomiale et la méthode de Monte Carlo ainsi que l'écart relatif entre les deux.

Cette illustration est importante car elle montre une utilisation très pratique de l'approximation polynomiale de la f.d.s.. L'approximation polynomiale est manipu-

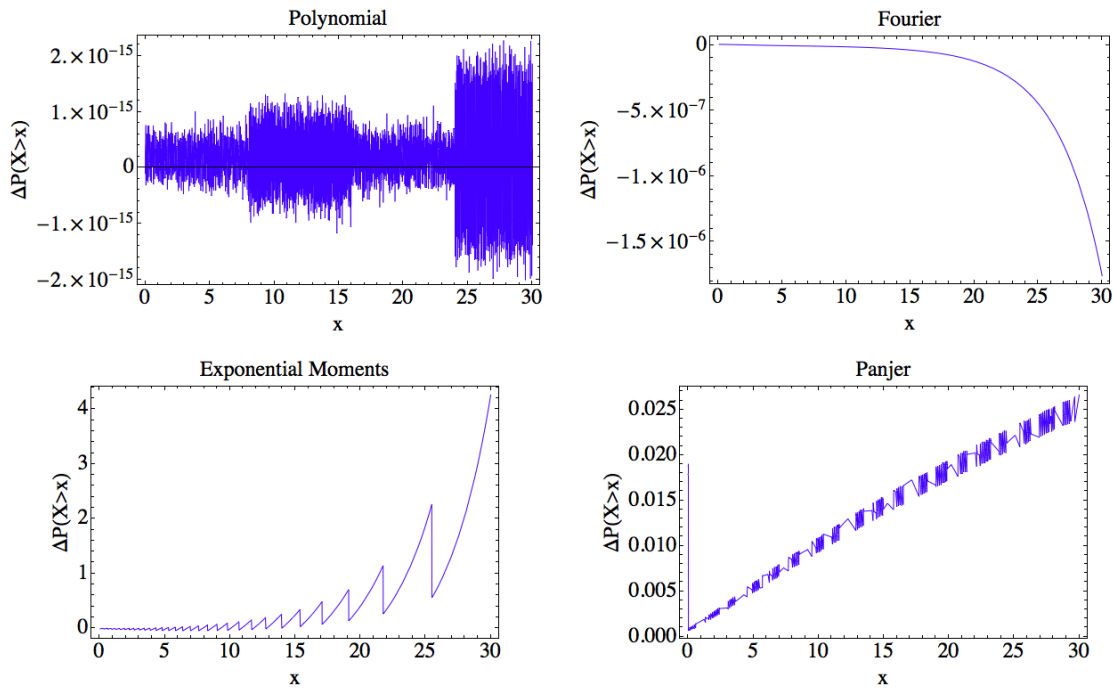


FIGURE 2.3 – Erreur relative de l’approximation polynomiale de la f.d.s. d’une distribution $[\mathcal{P}(4), \Gamma(1, 1/2)]$ avec différentes méthodes.

TABEAU 2.3 – f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \Gamma(1, 2)]$ approchée par différentes méthodes.

x	Exact	Polynomial	Fourier	Exponential Moment	Panjer
3.	0.806382	0.806382	0.806382	0.805233	0.808772
6.	0.573092	0.573092	0.573092	0.560555	0.576346
9.	0.364357	0.364357	0.364357	0.390845	0.367403
12.	0.21241	0.21241	0.21241	0.220259	0.214731
15.	0.115555	0.115555	0.115555	0.14352	0.117098
18.	0.0594094	0.0594094	0.0594094	0.0785728	0.0603396
21.	0.0291366	0.0291366	0.0291366	0.0521104	0.0296565
24.	0.0137285	0.0137285	0.0137285	0.0305607	0.014002
27.	0.00624886	0.00624886	0.00624886	0.0145444	0.00638581
30.	0.00275979	0.00275979	0.00275979	0.0145444	0.00282554

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE REPRÉSENTATION POLYNOMIALE

TABLEAU 2.4 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \Gamma(1, 2)]$, approchée par différentes méthodes.

x	Polynomial	Fourier	Exponential Moment	Panjer
3.	0.	0.	-0.14243	0.29637
6.	0.	0.	-2.18774	0.56776
9.	0.	0.	7.26981	0.83609
12.	0.	0.	3.69503	1.09255
15.	0.	0.	24.2007	1.33558
18.	0.	-0.00001	32.2566	1.56579
21.	0.	-0.00002	78.8485	1.78436
24.	0.	-0.00003	122.608	1.99258
27.	0.	-0.00008	132.753	2.19157
30.	0.	-0.00018	427.012	2.38237

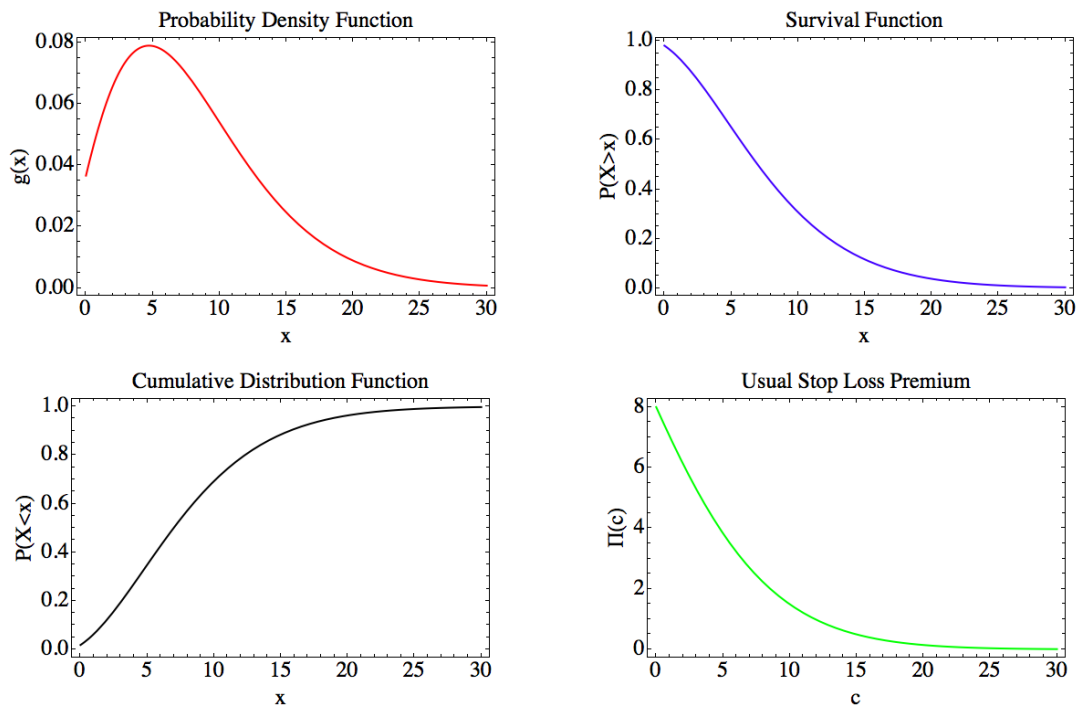


FIGURE 2.4 – Approximation polynomiale de la densité défailante, de la f.d.s., de la f.d.r. et de la prime *stop-loss* usuelle pour une distribution $[\mathcal{P}(4), \Gamma(1, 2)]$.

TABLEAU 2.5 – Evaluation de la prime *stop-loss* usuelle pour une loi composée $[\mathcal{P}(4), \Gamma(1, 1/2)]$ via l'approximation polynomiale et des simulations de Monte-Carlo.

c	Monte-Carlo	Polynomial	Relative Error (%)
3.	5.27937	5.28977	0.19704
6.	3.21013	3.2187	0.26682
9.	1.81787	1.82481	0.38157
12.	0.969839	0.974639	0.49493
15.	0.49123	0.494874	0.74191
18.	0.237786	0.240618	1.1907
21.	0.110721	0.112688	1.77607
24.	0.0498302	0.0510734	2.49487
27.	0.0217394	0.0224882	3.44426
30.	0.00939161	0.00965032	2.75466

lable, il est possible de l'intégrer pour approcher d'autres quantités avec une bonne précision. Sur cet exemple, les primes *stop loss* issues de l'approximation polynomiale prennent des valeurs numériques proches des valeurs obtenues par Monte-Carlo. Le calcul de la prime *stop-loss* $\Pi_{1,c}(X)$, par Monte-Carlo, pour différentes valeurs de c nécessite la manipulation et la sauvegarde de jeux de données contenant 10^6 observations. Ce processus est coûteux en temps de calcul et en mémoire. Une visualisation graphique à l'image de celle de la Figure 2.4 serait obtenue en interpolant les valeurs calculées pour différentes valeurs de c . L'utilisation de l'approximation polynomiale consiste à ré-injecter l'approximation de la fonction de survie dans l'expression intégrale (1.7) de la prime *stop-loss*. Le résultat du calcul intégral est une approximation de $\Pi_{1,c}(X)$ valable pour toute valeur de c .

2.4.2 Le cas des sinistres de loi gamma $\Gamma(\alpha, \beta)$

Les montants de sinistre sont modélisés par une suite $\{U_i\}_{i \in \mathbb{N}}$ de variable aléatoires **i.i.d.** suivant une loi gamma $\Gamma(\alpha, \beta)$. Dans ce cas, la densité et la **f.d.s.** de la variable aléatoire X ne sont pas disponibles explicitement. Une approximation de la densité défaillante est obtenue en tronquant simplement la série infinie

$$g_X(x) = \sum_{l=0}^{+\infty} \frac{e^{-\lambda}}{\lambda^l l!} \times \frac{e^{-x/\beta} x^{l \times \alpha - 1}}{\beta^{l \times \alpha} \Gamma(l \times \alpha)},$$

à l'ordre L . Cette approximation est particulièrement efficace pour λ faible eu égard à la décroissance rapide des probabilités de la loi de Poisson. Ce proxy va servir de référence pour apprécier les performances des méthodes d'approximations.

La **f.g.m.** de X de loi de probabilité composée $[\mathcal{P}(\lambda), \Gamma(\alpha, \beta)]$ est donnée par

$$\mathcal{L}_X(s) = \exp \left[\lambda \left(\frac{1}{1 - s\beta} \right)^\alpha - 1 \right], \quad (2.111)$$

ce qui implique que $\gamma_X = \frac{1}{\beta}$. La fonction génératrice des coefficients est obtenue par

réinjection de la f.g.m. (2.111) dans (2.46), avec

$$\mathcal{C}(z) = (1+z)^{-r} \exp \left\{ \lambda \left[\frac{m^\alpha(1+z)^\alpha - [m+z(m-\beta)]^\alpha}{[m+z(m-\beta)]^\alpha} \right] \right\}. \quad (2.112)$$

La paramétrisation

$$m_1 = \beta ; r_1 = 1, \quad (2.113)$$

simplifie la fonction génératrice des coefficients, avec

$$\mathcal{C}(z) = (1+z)^{-1} \exp [\lambda(1+z)^\alpha].$$

La paramétrisation (2.113) est conforme au Corollaire 2. La paramétrisation 2, définie par

$$m_2 = \beta ; r_2 = \alpha, \quad (2.114)$$

simplifie la fonction génératrice des coefficients, avec

$$\mathcal{C}(z) = (1+z)^{-\alpha} \exp [\lambda(1+z)^\alpha].$$

La paramétrisation (2.114) n'est pas conforme au Corollaire 2. L'idée est de rapprocher autant que possible la mesure de référence de la mesure de probabilité du montant des sinistres. La paramétrisation 1 fait correspondre le moment d'ordre 1 et la paramétrisation 2 fait correspondre les moments d'ordre 1 et 2. Pour l'illustration numérique, le paramètre de loi de Poisson est $\lambda = 2$ et les paramètres de la loi gamma sont $\alpha = 3$ et $\beta = 1$. La paramétrisation 3 est définie par l'équation (2.104), ce qui donne

$$m_3 = 6 ; r_3 = 1, \quad (2.115)$$

après applications numériques. La paramétrisation 4 est définie par l'équation (2.103), ce qui donne

$$m_4 = 5 ; r_4 = \frac{6}{5}, \quad (2.116)$$

après applications numériques. La paramétrisation (2.115) est conforme au Corollaire 2, ce n'est pas le cas de la paramétrisation (2.124). L'ordre de troncature est $K = 85$ pour chacune des quatre paramétrisations. La Figure 2.5 représente l'erreur relative de l'approximation polynomiale sur la densité défaillante $g_X(x)$, d'une loi de probabilité composée $[\mathcal{P}(2), \Gamma(3, 1)]$, associée aux différentes paramétrisations, en fonction des valeurs de x . La Figure 2.6 représente l'erreur relative de l'approximation polynomiale sur la f.d.s. $\overline{F}_X(x)$, d'une loi de probabilité composée $[\mathcal{P}(2), \Gamma(3, 1)]$, associée aux différentes paramétrisations, en fonction des valeurs de x . Le Tableau 2.6 donne les valeurs approchées grâce à la méthode d'approximation polynomiale suivant les différentes paramétrisations. Le Tableau 2.7 donne les valeurs numériques des erreurs relatives liées à l'emploi de l'approximation polynomiale suivant les différentes paramétrisations.

L'erreur relative explose pour la paramétrisation 4. La paramétrisation (2.114) bien que non conforme au Corollaire 2 est celle qui fonctionne le mieux, tallonnée par la paramétrisation (2.113). Cela montre que le Corollaire 2 est une condition suffisante

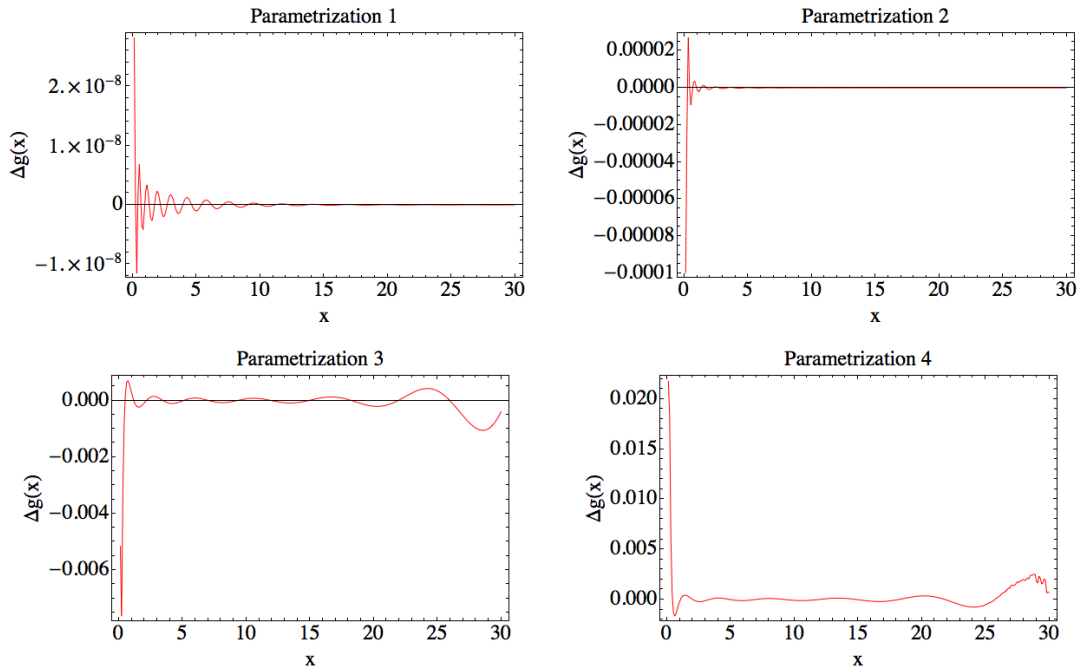


FIGURE 2.5 – Erreur relative de l’approximation polynomiale, avec différentes paramétrisations, de la densité de probabilité d’une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$.

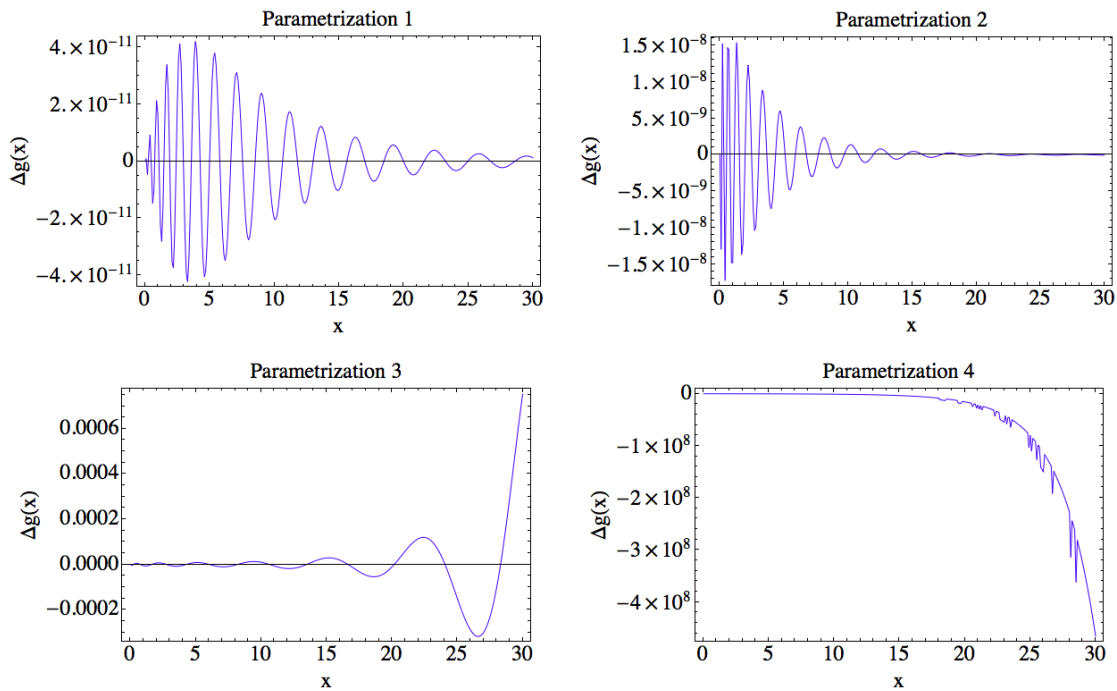


FIGURE 2.6 – Erreur relative de l’approximation polynomiale, avec différentes paramétrisations, de la f.d.s. d’une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$.

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE REPRÉSENTATION POLYNOMIALE

TABLEAU 2.6 – Evaluation de la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ via l'approximation polynomiale avec différentes paramétrisations.

x	Exact	Param. 1	Param. 2	Param. 3	Param. 4
3.	0.685132	0.685132	0.685132	0.685129684	-196608.
6.	0.4313	0.4313	0.4313	0.4312998	-196608.
9.	0.238763	0.238763	0.238763	0.2387659551	-196608.
12.	0.118895	0.118895	0.118895	0.11889306	-196608.
15.	0.0542376	0.0542376	0.0542376	0.0542391	-196608.
18.	0.0229767	0.0229767	0.0229767	0.0229756268	-196608.
21.	0.00913388	0.00913388	0.00913388	0.00913441	-196608.
24.	0.00343513	0.00343513	0.00343513	0.0034351619	-196608.
27.	0.00123021	0.00123021	0.00123021	0.00122984	-196608.
30.	0.000421752	0.000421752	0.000421752	0.000422070029	-196608.

TABLEAU 2.7 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ approchée par méthode polynomiale suivant différentes paramétrisations.

x	Param. 1	Param. 2	Param. 3	Param. 4
3.	0.	0.	-0.0004	∞
6.	0.	0.	0.00006	∞
9.	0.	0.	0.00111	∞
12.	0.	0.	-0.00181	∞
15.	0.	0.	0.00287	∞
18.	0.	0.	-0.00469	∞
21.	0.	0.	0.00586	∞
24.	0.	0.	0.0008	∞
27.	0.	0.	-0.03009	∞
30.	0.	0.	0.07546	∞

pour que l'approximation polynomiale (2.100) soit valide, cette condition n'est cependant pas nécessaire. Ce cas particulier montre aussi que plus la mesure de référence est proche de la mesure de probabilité des montants de sinistres, plus la méthode d'approximation est performante. La comparaison avec les autres méthodes d'approximation est effectuée avec la paramétrisation 1. La Figure 2.7 représente l'erreur relative sur la f.d.s. de la loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ suivant la méthode numérique utilisée. Le Tableau 2.8 donne les valeurs numériques de la f.d.s. résultant de l'emploi des différentes méthodes. Le Tableau 2.9 donne les valeurs numériques de l'erreur relative sur la f.d.s. suite à l'approximation via les différentes méthodes.

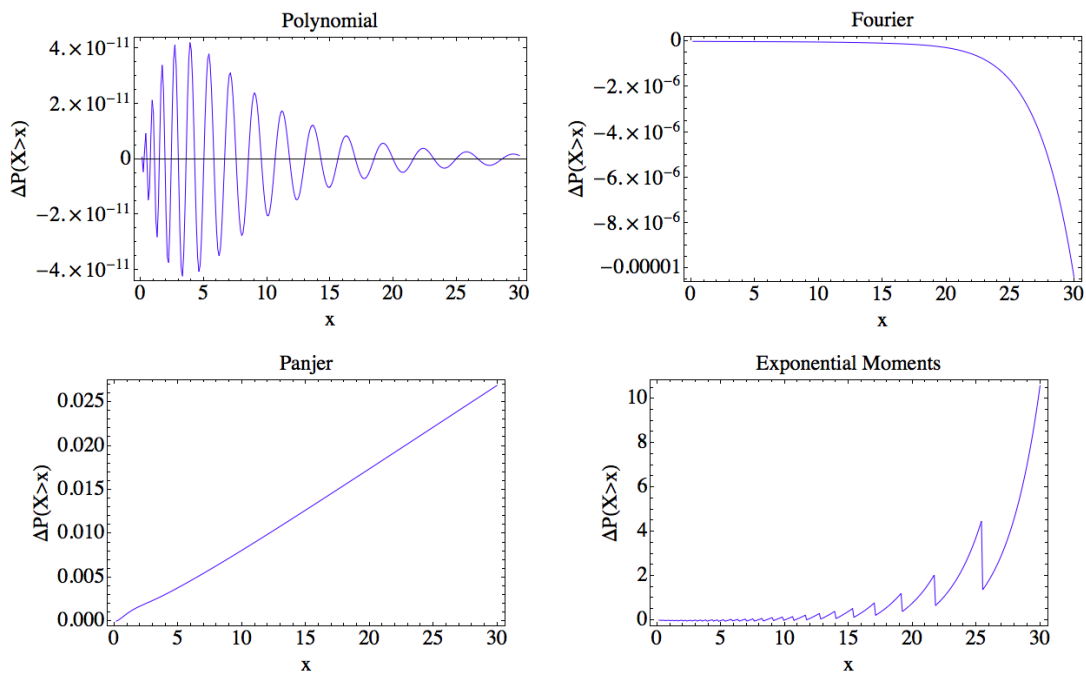


FIGURE 2.7 – Erreur relative de l'approximation polynomiale de la f.d.s. d'une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$ suivant différentes méthodes numériques.

Le classement en termes de performance est le même que dans le cas où les montants de sinistres sont distribués exponentiellement. La méthode polynomiale arrive en tête, suivie de la méthode d'inversion de la transformée de Fourier, l'algorithme de Panjer puis pour finir la méthode des moments exponentiels. La Figure 2.8 propose une visualisation graphique de l'approximation polynomiale, obtenue en utilisant la paramétrisation 1 et un ordre de troncature $K = 85$, de la densité de probabilité, de la f.d.s., de la f.d.r., et de la prime *stop-loss* usuelle de la loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$. Le Tableau 2.10 donne les valeurs numériques de la prime *stop-loss* usuelle calculées par la méthode polynomiale et la méthode de Monte-Carlo.

La prime *stop-loss* explose pour un seuil de rétention $c \geq 9$, ce qui est assez surprenant au vu des bons résultats obtenus dans l'approximation de la densité de probabilité, de la f.d.s., et de la f.d.r.. L'emploi d'un ordre de troncature trop grand entraîne des problèmes numériques lors de l'évaluation de l'intégrale (2.102). Il n'y a aucune raison pour que l'intégration en elle-même engendre une telle erreur. Cette erreur est

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE REPRÉSENTATION POLYNOMIALE

TABLEAU 2.8 – Evaluation de la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ par différentes méthodes.

x	Exact	Polynomial	Fourier	Exponential Moment	Panjer
3.	0.685132	0.685132	0.685132	0.685203	0.686771
6.	0.4313	0.4313	0.4313	0.422254	0.433313
9.	0.238763	0.238763	0.238763	0.266124	0.24049
12.	0.118895	0.118895	0.118895	0.129356	0.120076
15.	0.0542376	0.0542376	0.0542376	0.076121	0.0549268
18.	0.0229767	0.0229767	0.0229767	0.0364145	0.0233334
21.	0.00913388	0.00913388	0.00913387	0.0222035	0.00930159
24.	0.00343513	0.00343513	0.00343513	0.011718	0.003508
27.	0.00123021	0.00123021	0.00123021	0.00489717	0.00125982
30.	0.000421752	0.000421752	0.000421747	0.00489717	0.000433105

TABLEAU 2.9 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ via différentes méthodes.

x	Polynomial	Fourier	Exponential Moment	Panjer
3.	0.	0.	0.01036	0.2392
6.	0.	0.	-2.0972	0.46687
9.	0.	0.	11.4594	0.72337
12.	0.	0.	8.79806	0.99333
15.	0.	-0.00001	40.3474	1.27079
18.	0.	-0.00001	58.4844	1.55239
21.	0.	-0.00004	143.089	1.83622
24.	0.	-0.00012	241.122	2.12116
27.	0.	-0.00036	298.076	2.40651
30.	0.	-0.00104	1061.15	2.69184

TABLEAU 2.10 – Evaluation de la prime *stop-loss* usuelle pour une loi composée $[\mathcal{P}(2), \Gamma(3, 1)]$ via l'approximation polynomiale et des simulations de Monte-Carlo.

c	Monte-Carlo	Polynomial	Relative Error (%)
0	5.99316	6.	0.114172
3	3.59554	3.60192	0.177357
6	1.93243	1.9526	1.04382
9	0.94746	29.3862	3001.58
12	0.429122	834.273	194314.
15	0.181003	-14846.1	∞
18	0.0716428	-415589.	∞
21	0.0269423	∞	∞
24	0.00976161	∞	∞
27	0.00342669	∞	∞
30	0.00113488	∞	∞

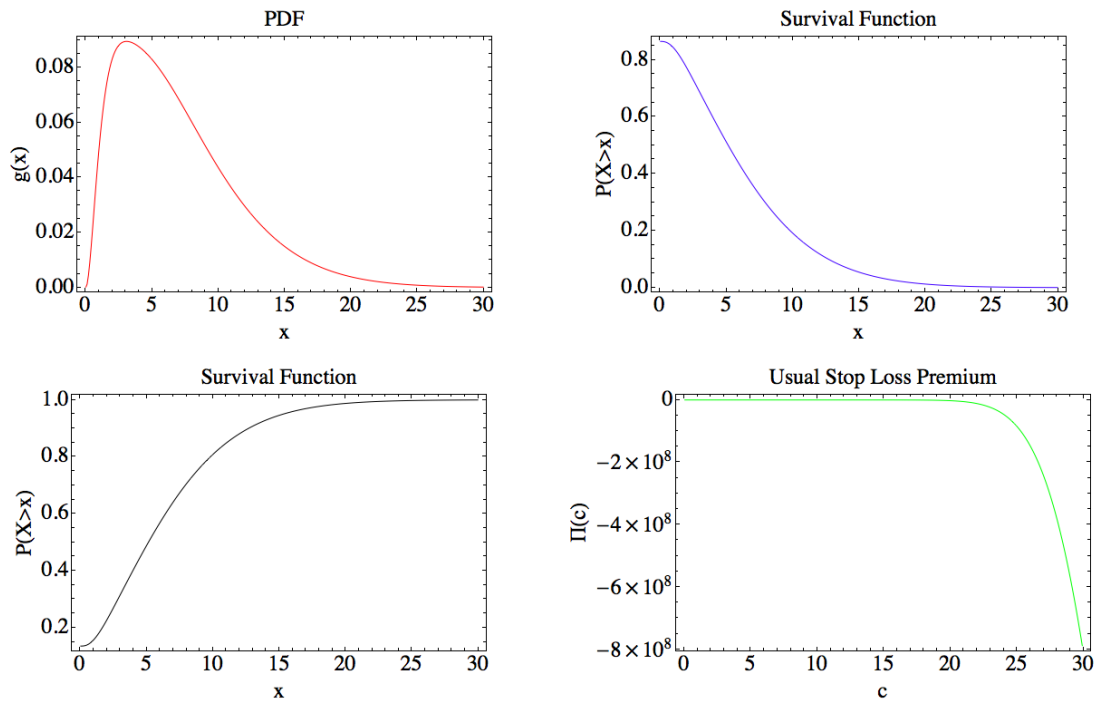


FIGURE 2.8 – Approximation polynomiale de la densité défaillante, de la f.d.s., de la f.d.r. et de la prime *stop-loss* usuelle pour une distribution $[\mathcal{P}(2), \Gamma(3, 1)]$.

sûrement due à des erreurs d'arrondi et à une maladresse dans la façon d'évaluer numériquement les intégrales via Mathematica. Ces problèmes seront à résoudre dans les suites de ce travail.

2.4.3 Le cas des sinistres de loi uniforme $\mathcal{U} [\alpha, \beta]$

Le montant des sinistres est modélisé par une suite $\{U_i\}_{i \in \mathbb{N}}$ de variables aléatoires i.i.d. suivant une loi uniforme $\mathcal{U}(\alpha, \beta)$.

Définition 14. Soit X une variable aléatoire de loi uniforme $\mathcal{U}(\alpha, \beta)$, avec $\alpha, \beta \in \mathbb{R}$ et $\alpha < \beta$. La densité de X est donnée par

$$f_X(x) = \frac{1}{\beta - \alpha}, \quad x \in [\alpha, \beta], \quad (2.117)$$

et sa f.g.m. par

$$\mathcal{L}_X(s) = \frac{e^{s\beta} - e^{s\alpha}}{\beta - \alpha}. \quad (2.118)$$

Dans ce cas, la densité et la f.d.s. de la variable aléatoire X ne sont pas disponibles explicitement, la méthode d'inversion numérique de la transformée de Fourier est utilisée pour mesurer la précision des méthodes numériques sur la f.d.s. de la variable aléatoire X . La f.g.m. de X de loi composée $[\mathcal{P}(\lambda), \mathcal{U}(\alpha, \beta)]$ est donnée par

$$\mathcal{L}_X(s) = \exp \left\{ \lambda \left[\frac{e^{s\beta} - e^{s\alpha}}{(\beta - \alpha)s} - 1 \right] \right\}, \quad (2.119)$$

ce qui implique que $\gamma_X = +\infty$. La fonction génératrice des coefficients admet une forme qui ne permet pas la prise de décision en ce qui concerne la paramétrisation. Au vu des résultats obtenus sur les deux exemples précédents, il est logique d'opter pour des paramétrisations basées sur les moments de la distribution de probabilité du montant des sinistres. La paramétrisation 1, définie par

$$m_1 = \frac{\mathbb{V}(U)}{\mathbb{E}(U)} = \frac{(\beta - \alpha)^2}{6(\alpha + \beta)}; r_1 = \frac{\mathbb{E}(U)^2}{\mathbb{V}(U)} = \frac{3(\alpha + \beta)^2}{(\beta - \alpha)^2}, \quad (2.120)$$

fait coïncider les moments d'ordre 1 et 2 de la mesure de référence avec ceux de la distribution du montant des sinistres. La paramétrisation 2, définie par

$$m_2 = \mathbb{E}(U) = \frac{(\beta - \alpha)}{2}; r_2 = 1$$

fait coïncider le moment d'ordre 1 de la mesure de référence avec celui de la distribution du montant des sinistres. Pour l'illustration, le paramètre de la loi de Poisson est $\lambda = 3$ et les paramètres de la loi uniforme $\mathcal{U}(\alpha, \beta)$ sont $\alpha = 0$ et $\beta = 10$. Ce qui donne

$$m_1 = \frac{5}{3}; r_1 = 3, \quad (2.121)$$

et

$$m_2 = 5; r_2 = 1, \quad (2.122)$$

après applications numériques. La paramétrisation 3 est définie par l'équation (2.104), ce qui donne

$$m_3 = 15; r_3 = 1, \quad (2.123)$$

après applications numériques. La paramétrisation 4 est définie par l'équation (2.103), ce qui donne

$$m_4 = \frac{23}{3}; r_4 = \frac{45}{23}, \quad (2.124)$$

après applications numériques. Les paramétrisations 2 et 3 sont conformes au Corollaire 2, ce n'est pas le cas des paramétrisations 1 et 4. L'ordre de troncature est $K = 75$ pour chacune des paramétrisations étudiées. La Figure 2.9 montre l'écart relatif de l'approximation polynomiale de la f.d.s. par rapport à l'approximation issue de l'inversion de la transformée de Fourier, suivant la paramétrisation. Le Tableau 2.11 donne les valeurs numériques de la f.d.s. approchée par la méthode d'inversion numérique de la transformée de Fourier et la méthode polynomiale. Le Tableau 2.12 donne les valeurs numériques des écarts entre les approximations polynomiales et celles renvoyées par l'inversion numérique de la transformée de Fourier.

Les écarts pour les paramétrisations 3 et 4 sont très importants. Les deux autres paramétrisations renvoient des valeurs proches de celles obtenues via l'inversion numérique de la transformée de Fourier, notamment la paramétrisation 2. Si il est admis que la méthode d'inversion de la transformée de Fourier est aussi précise que dans les cas d'étude précédent, alors le niveau d'erreur de l'approximation polynomiale est bien plus important que dans les cas précédent. Ce constat s'explique par le fait que la

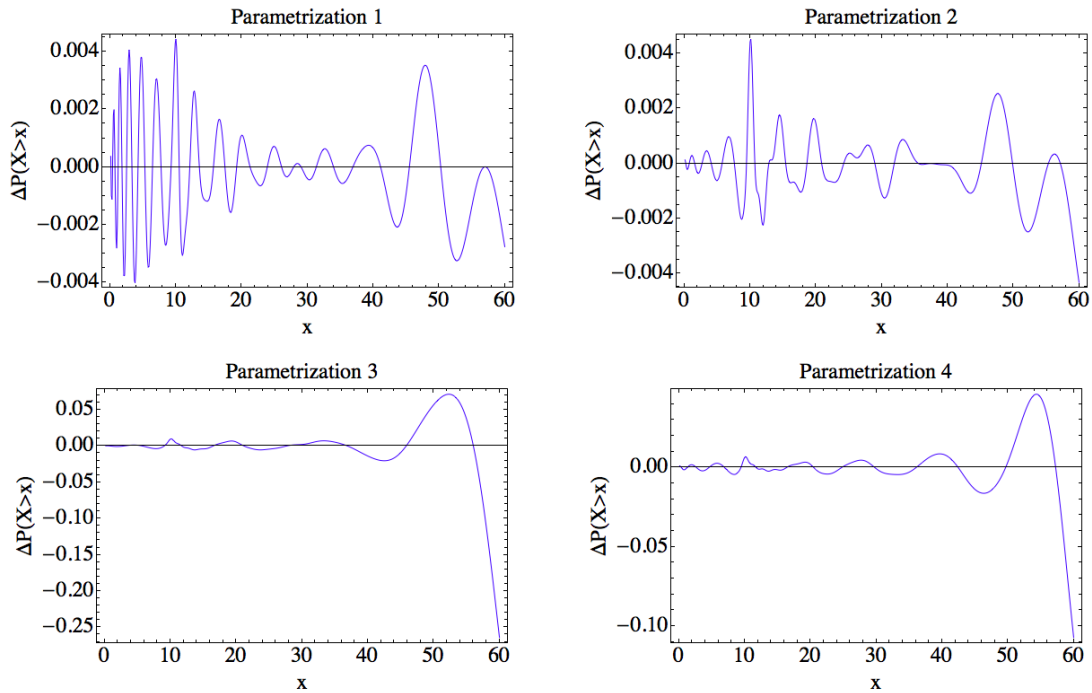


FIGURE 2.9 – Erreur relative de l’approximation polynomiale, avec différentes paramétrisations, de la f.d.s. d’une distribution $[\mathcal{P}(3), \mathcal{U}(0, 10)]$ par rapport à méthode d’inversion numérique de la transformée de Fourier.

TABLEAU 2.11 – Evaluation de la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \mathcal{U}(0, 10)]$ via la méthode d’approximation polynomiale avec différentes paramétrisations.

x	Fourier	Param. 1	Param. 2	Param. 3	Param. 4
6.	0.811227	0.80865	0.811507	0.810796205	0.81332372
12.	0.56834	0.568012	0.567072	0.56746466	0.5678477
18.	0.34191	0.341455	0.341587	0.3434973	0.3426809
24.	0.17857	0.178591	0.178539	0.177635	0.1782032
30.	0.081776	0.0817428	0.0816851	0.081952935	0.0817116
36.	0.033478	0.0334635	0.0334768	0.033576449	0.03347229796
42.	0.0124404	0.0124292	0.0124332	0.012192117	0.01247017
48.	0.00420535	0.00422016	0.00421571	0.0043344158	0.0041541479
54.	0.00132527	0.00132197	0.00132356	0.0014034753	0.00138585
60.	0.000386308	0.000385234	0.000384629	0.00028410264	0.0003448

CHAPITRE 2. APPROXIMATION DE LA DENSITÉ DE PROBABILITÉ VIA UNE REPRÉSENTATION POLYNOMIALE

TABLEAU 2.12 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \mathcal{U}(0, 10)]$ via la méthode d'approximation polynomiale avec différentes paramétrisations.

x	Param. 1	Param. 2	Param. 3	Param. 4
6.	-0.31771	0.03449	-0.05311	0.25845
12.	-0.05772	-0.2231	-0.15394	-0.08653
18.	-0.13288	-0.09454	0.4643	0.22553
24.	0.01139	-0.01759	-0.52331	-0.2055
30.	-0.04068	-0.11119	0.21631	-0.07873
36.	-0.04315	-0.00353	0.29409	-0.01701
42.	-0.0902	-0.05834	-1.99604	0.23904
48.	0.35227	0.24625	3.06909	-1.21755
54.	-0.24908	-0.1292	5.90079	4.57148
60.	-0.27808	-0.43455	-26.457	-10.7411

loi uniforme est bien plus éloignée de la mesure de référence que la loi gamma. L'augmentation de l'ordre de troncature devrait pouvoir résoudre ce problème, il pourrait cependant s'accompagner de difficultés d'ordre numérique. La paramétrisation 2 est retenue pour effectuer la comparaison avec les autres méthodes numériques. La Figure 2.10 permet la visualisation des écarts entre l'approximation de la f.d.s. de X via l'inversion de la transformée de Fourier et les autres méthodes numériques. Le Tableau 2.13 donne les valeurs numériques de la f.d.s. approchée par les différentes méthodes. Le Tableau 2.14 donne les valeurs numériques des écarts entre l'approximation de la f.d.s. via l'inversion de la transformée de Fourier et les autres méthodes numériques.

TABLEAU 2.13 – Evaluation de la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \mathcal{U}(0, 10)]$ par différentes méthodes.

x	Fourier	Polynomial	Exponential Moment	Panjer
6	0.811227	0.811507	0.798633	0.811957
12	0.56834	0.567072	0.557872	0.569321
18	0.34191	0.341587	0.35222	0.343157
24	0.17857	0.178539	0.216235	0.179621
30	0.081776	0.0816851	0.140781	0.0823872
36	0.033478	0.0334768	0.0644221	0.0338055
42	0.0124404	0.0124332	0.0644221	0.0125867
48	0.00420535	0.00421571	0.0644221	0.00428459
54	0.00132527	0.00132356	0.0644221	0.00134572
60	0.000386308	0.000384629	0.0644221	0.000393232

Dans ce cas particulier, l'algorithme propose une approximation concurrentielle avec la méthode polynomiale. La méthode des moments exponentiels performe nettement moins bien que les deux autres méthodes. La Figure 2.11 montre l'approximation polynomiale de la densité, de la f.d.s., de la f.d.r., et de la prime *stop-loss* usuelle, obtenues avec la paramétrisation 2 et un ordre de troncature $K = 75$. Le Tableau 2.15 donne les valeurs numériques des approximations par la méthode polynomiale et la méthode

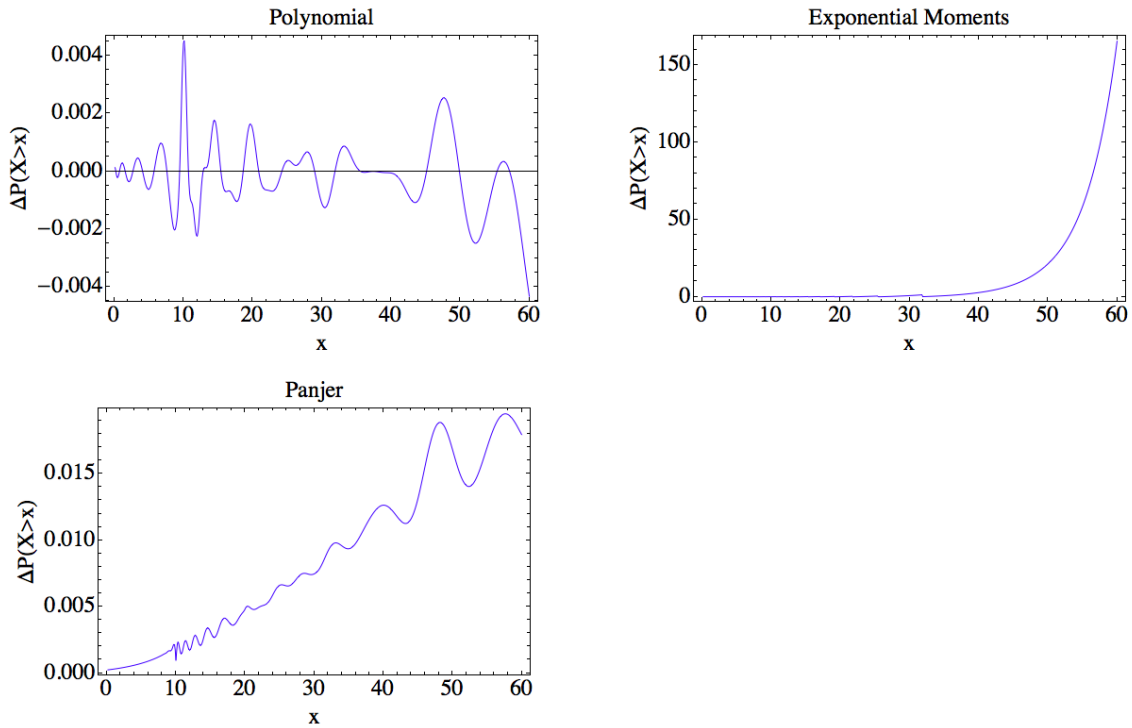


FIGURE 2.10 – Erreur relative, en %, sur la f.d.s. d'une distribution $[\mathcal{P}(2), \mathcal{U}(0, 10)]$ suivant la méthode numérique utilisée.

TABLEAU 2.14 – Erreur relative en %, à 10^{-5} près, sur la f.d.s. de la variable aléatoire X de loi composée $[\mathcal{P}(4), \mathcal{U}(0, 10)]$ par rapport à la méthode d'inversion de la transformée de Fourier.

x	Polynomial	Exponential Moment	Panjer
6	0.03449	-1.55242	0.08995
12	-0.2231	-1.84175	0.17261
18	-0.09454	3.01549	0.36475
24	-0.01759	21.0926	0.58869
30	-0.11119	72.1542	0.74738
36	-0.00353	92.4311	0.97826
42	-0.05834	417.844	1.17585
48	0.24625	1431.91	1.88423
54	-0.1292	4761.04	1.543
60	-0.43455	16576.4	1.79229

de Monte-Carlo de la prime *stop-loss* usuelle ainsi que l'écart relatif entre les deux.

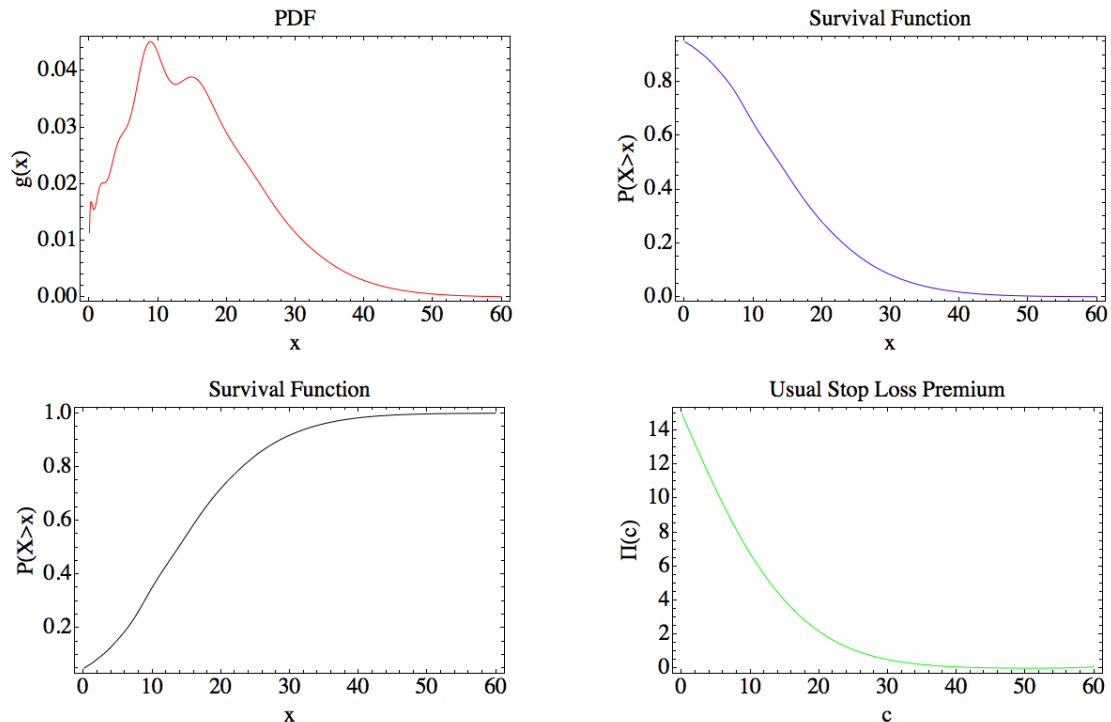


FIGURE 2.11 – Approximation polynomiale de la densité défailante, de la f.d.s., de la f.d.r. et de la prime *stop-loss* usuelle pour une distribution $[\mathcal{P}(3), \mathcal{U}(0, 10)]$.

La méthode d'approximation polynomiale se heurte comme dans le cas d'étude précédent à des difficultés d'ordre numérique dans l'approximation de la prime *stop-loss* usuelle pour des valeurs de seuils de rétention élevé $c \geq 9$.

La méthode d'approximation polynomiale est une méthode efficace, qui concurrence l'inversion numérique de la transformée de Fourier dans certains cas. La restriction sur le choix des paramètres conformément au Corollaire 2 assure la validité de l'approximation et l'obtention de résultats au moins comparables à l'algorithme de Panjer. Son avantage réside dans l'exploitation directe de l'approximation de la densité pour l'évaluation de la f.d.s. et de la prime *stop-loss*. Les difficultés numériques doivent pouvoir être contournées en prêtant soin à l'implémentation de la méthode polynomiale. Les temps de calcul sont aussi un frein pour obtenir l'approximation de la f.d.s. pour un ordre de troncature supérieur à 100. Encore une fois, une implémentation plus adroite doit permettre de diminuer les temps de calcul. Le principal avantage de la méthode polynomiale réside dans la perspective d'une application statistique lorsque des données sont disponibles.

TABLEAU 2.15 – Evaluation de la prime *stop-loss* usuelle pour une loi composée [$\mathcal{P}(4), \mathcal{U}(0, 10)$] via l’approximation polynomiale et des simulations de Monte-Carlo.

c	Monte-Carlo	Polynomial	Relative Error (%)
0	5.99316	6.	0.114172
3	3.59554	3.60192	0.177357
6	1.93243	1.9526	1.04382
9	0.94746	29.3862	3001.58
12	0.429122	834.273	194314.
15	0.181003	∞	∞
18	0.0716428	∞	∞
21	0.0269423	∞	∞
24	0.00976161	∞	∞
27	0.00342669	∞	∞
30	0.00113488	∞	∞

2.5 Perspectives en statistique

2.5.1 Estimation de la densité de probabilité via la représentation polynomiale

Soit X_1, \dots, X_n un échantillon *i.i.d.*, de taille n , issu de la loi de probabilité \mathbb{P}_X . Un estimateur de la densité est obtenu à partir de l’approximation (2.18), avec

$$\widehat{f}_X^K(x) = \widehat{f}_{X,v}^K(x) \widehat{f}_v(x). \quad (2.125)$$

L’écriture \widehat{f}_v sous-entend une estimation statistique des paramètres de la mesure de probabilité de référence. La distribution de référence représente une distribution a priori. Un ordre de troncature $K = 0$ équivaut à faire l’hypothèse d’un modèle paramétrique basé sur les [FENQ](#). Le choix d’un ordre de troncature supérieur à 0 permet d’ajuster et de corriger l’erreur de modèle via un développement polynomial. L’estimateur (2.125) est semi-paramétrique. La technique d’estimation du paramètre dépend du choix de la mesure de référence et du problème étudié. Afin de conserver un caractère général pour l’instant, il est supposé que les paramètres de la mesure de référence ont été estimés. Le système de polynômes $\{Q_k\}_{k \in \mathbb{N}}$ forme un système orthonormal par rapport à la mesure de probabilité caractérisée par la densité \widehat{f}_v . Cette idée de démarrer avec une première estimation paramétrique et d’ajuster à l’aide d’une projection sur une base de fonctions orthogonales remonte aux travaux de [WHITTLE \[1958\]](#) et [BRUNK \[1978\]](#). Le terme *prior distribution* est d’ailleurs employé pour parler de la mesure associée à f_v . Les travaux plus récents de [BUCKLAND \[1992\]](#) et de [PROVOST et JIANG \[2012\]](#) décrivent une méthode basée sur des polynômes orthogonaux par rapport à une mesure de référence, appelée fonction de poids. La méthode présentée dans [PROVOST et JIANG \[2012\]](#) est d’ailleurs basée sur l’approximation de la densité par les moments de la distribution qui fait l’objet de l’article [PROVOST \[2005\]](#) déjà cité. Dans le papier de [HJORT et GLAD \[1995\]](#), la densité par rapport à une mesure de probabilité paramétrique de référence est estimée via un estimateur à noyau.

L'estimateur de $f_{X,v}^K$ est donné par

$$\widehat{f}_{X,v}^K(x) = \sum_{k=1}^K \widehat{a}_k Q_k(x), \quad (2.126)$$

où les composantes du vecteur $\widehat{\mathbf{a}} = (\widehat{a}_1, \dots, \widehat{a}_K)$ sont estimées par

$$\widehat{a}_k = w_k \frac{1}{n} \sum_{i=1}^n Q_k(X_i), \quad \forall k \in \{1, \dots, K\}. \quad (2.127)$$

Le vecteur $\mathbf{w} = (w_1, \dots, w_K)$ est un modulateur dont les composantes vérifient $0 \leq w_k \leq 1$ pour $k = 1, \dots, K$. Le nombre de paramètres à estimer ne peut excéder le nombre d'observations, ce qui revient à imposer la condition $K \leq n$. Le problème d'estimation est non-paramétrique au sens où le nombre de paramètres à estimer augmente avec le nombre d'observations. L'estimation non-paramétrique de la densité via les séries orthogonales est un sujet classique dans la littérature statistique. Les avantages sont, entre autres, la facilité de l'implémentation et l'efficacité en dimension supérieure à 1. L'estimateur (2.126) est original car il fait intervenir une mesure de probabilité de référence. Les estimateurs non-paramétriques de la densité de ce type sont plus souvent définis comme des développements sur des bases de fonctions orthogonales. Ces estimateurs ont été initialement introduit dans les travaux de CENCOV [1962]. Les estimateurs basés sur les fonctions de Hermite, définies comme le produit de la densité de la loi normale et des polynômes d'Hermite, ont été étudiés dans les papiers de SCHWARTZ [1967] et WALTER [1977]. Une extension à l'estimation de la fonction de densité et de ses dérivées est proposée dans le travail de GREBLICKI et PAWLAK [1984]. Une étude de l'estimateur basé sur les fonctions de Laguerre, définis comme le produit de la densité de la loi gamma et des polynômes de Laguerre est proposée dans le papier de HALL [1980], Cet estimateur est comparé à l'estimateur basé sur les fonctions de Hermite.

L'optimisation du modulateur \mathbf{w} fait l'objet de nombreux travaux. Le critère proposé dans KRONMAL et TARTER [1968] revient à supposer \mathbf{w} de la forme $(1, \dots, 1, 0, \dots, 0)$. Cette idée a été reprise et affinée par exemple dans DIGGLE et HALL [1986]. Dans le travail de WATSON [1969], le modulateur \mathbf{w} est optimisé pour que la suite des composantes du vecteur soit monotone, décroissante vers 0. Le livre d'EFROMOVICH [1999] décrit de façon exhaustive les propriétés de ces estimateurs basés sur les fonctions orthogonales, l'accent est mis sur l'utilisation des fonctions trigonométriques. Une mesure de référence, sans l'aspect paramétrique, est utilisée dans le papier de ANDERSON et FIGUEIREDO [1980] avec un modulateur monotone.

Le problème d'estimation est rapproché du *Normal Mean problem* à l'image de ce qui est présenté dans les chapitres 7 et 8 de l'ouvrage de WASSERMAN [2006]. L'estimation des coefficients $\{\widehat{a}_k\}_{k \in \{1, \dots, n\}}$ est similaire à l'estimation d'une loi gaussienne multivariée. Le vecteur aléatoire $\mathbf{Z} = (Z_1, \dots, Z_n)$ est défini par

$$Z_k = \frac{1}{n} \sum_{i=1}^n Q_k(X_i), \quad k = 1, \dots, n. \quad (2.128)$$

Les composantes du vecteur \mathbf{Z} estiment sans biais les coefficients du développement, avec

$$\mathbb{E}(Z_k) = a_k, \quad k = 1, \dots, n. \quad (2.129)$$

La variance de Z_k est donnée par

$$\sigma_{k,n}^2 = \mathbb{V}(Z_k), \quad k = 1, \dots, n. \quad (2.130)$$

L'Erreur Quadratique Moyenne Intégrée (EQMI) issue de l'estimation de $f_{X,v}$ par l'estimateur (2.126) avec un ordre de troncature $K = n$ est donné par

$$R(f_{X,v}, \widehat{f_{X,v}}^n) = \mathbb{E} \left[L(f_{X,v}, \widehat{f_{X,v}}^n) \right] \quad (2.131)$$

$$= \sum_{k=1}^n (1 - w_k)^2 a_k^2 + \sum_{k=n+1}^{+\infty} a_k^2 \quad (2.132)$$

$$+ \sum_{k=1}^n w_k^2 \sigma_{k,n}^2. \quad (2.133)$$

Le risque (2.131) est la somme du biais au carré (2.132) et de la variance (2.133) de l'estimateur. Le but est de trouver le modulateur \mathbf{w} permettant d'optimiser ce risque. La quantité $\sum_{k=n+1}^{+\infty} a_k^2$ est inaccessible, elle est cependant positive et indépendante du modulateur. L'optimisation du risque (2.131) sur l'ensemble des modulateurs revient à optimiser

$$R(f_{X,v}^n, \widehat{f_{X,v}}^n) = \sum_{k=1}^n (1 - w_k)^2 a_k^2 + \sum_{k=1}^n w_k^2 \sigma_{k,n}^2. \quad (2.134)$$

De plus, il est classique d'obtenir un résultat du type

$$a_k = O\left(\frac{1}{k}\right), \quad k \rightarrow +\infty.$$

sous des hypothèses de régularité pour la fonction $f_{X,v}$ et ses dérivées. La quantité $\sum_{k=n+1}^{+\infty} a_k^2$ peut alors être négligée dans un contexte d'estimation non-paramétrique de la densité dans lequel la vitesse de convergence optimale est de l'ordre de $k^{-4/5}$.

Dans l'expression du risque (2.134), interviennent des quantités inconnues qu'il est nécessaire d'estimer statistiquement. le carré des coefficients de la représentation polynomiale a_k^2 sont estimés sans biais, via

$$\widehat{a}_k^2 = Z_k^2 - \sigma_{k,n}^2, \quad k = 1, \dots, n.$$

Afin de garantir la positivité de l'estimation \widehat{a}_k^2 cette estimateur est modifié, avec

$$\widehat{a}_k^2 = \left(Z_k^2 - \sigma_{k,n}^2 \right)_+.$$

L'estimateur du risque (2.134) à ce stade est donné par

$$\widehat{R}(f_{X,v}^n, \widehat{f_{X,v}}^n) = \sum_{k=1}^n (1 - w_k)^2 \left(Z_k^2 - \sigma_{k,n}^2 \right)_+ + \sum_{k=1}^n w_k^2 \sigma_{k,n}^2. \quad (2.135)$$

Remarque 2. L'approche Normal Means est justifiée par la convergence en loi du vecteur \mathbf{Z} vers une loi gaussienne multivariée. L'application du *Théorème Central Limite (TCL)* multivarié implique que

$$\sqrt{n}(\mathbf{Z} - \mathbf{a}) \underset{a.s.}{\sim} \mathcal{N}(0, \Sigma), \quad n \rightarrow +\infty, \quad (2.136)$$

où $\Sigma = \{\text{Cov}(Q_k(X), Q_l(X))\}_{k,l \in \{1, \dots, n\}}^2$ est la matrice de variance-covariance. Ce résultat peut permettre d'appliquer le théorème de Stein, voir [STEIN \[1981\]](#), qui donne une estimation sans biais du risque (2.134). Il est intéressant d'observer que l'estimation du risque de Stein, ou *Stein's Unbiased Risk Estimation (SURE)*, coïncide avec l'estimation obtenue dans l'équation (2.135)

Les termes de variance $\sigma_{k,n}^2$ sont estimés sans biais, avec

$$\hat{\sigma}_{k,n}^2 = \frac{1}{n(n-1)} \sum_{i=1}^n [Q_k(X_i) - Z_k]^2, \quad k = 1, \dots, n. \quad (2.137)$$

Le risque est finalement estimé par

$$\hat{\mathbb{R}}\left(f_{X,v}^n, \widehat{f}_{X,v}^n\right) = \sum_{k=1}^n (1 - w_k)^2 \left(Z_k^2 - \hat{\sigma}_{k,n}^2\right)_+ + \sum_{k=1}^n w_k^2 \hat{\sigma}_{k,n}^2. \quad (2.138)$$

La procédure [Risk Estimation and Adaptation after Coordinate Transformation \(REACT\)](#) proposée dans les travaux de [BERAN \[2000\]](#) pour estimer la fonction de régression à l'aide des fonctions orthogonales, est adaptée à l'estimation de la densité de probabilité par polynômes orthogonaux. Cette procédure permet d'optimiser le risque (2.138) sur certaines classes d'estimateurs. L'estimateur [REACT](#) est défini par :

Définition 15. Soit \mathcal{M} une classe de modulateurs. L'estimateur modulé de \mathbf{a} est $\hat{\mathbf{a}} = (\hat{w}_1 Z_1, \dots, \hat{w}_n Z_n)$ où $\hat{\mathbf{w}}$ minimise $\hat{\mathbb{R}}\left(f_{X,v}^n, \widehat{f}_{X,v}^n\right)$ sur l'ensemble des modulateurs de la classe \mathcal{M} . L'estimateur [REACT](#) de la densité de probabilité est

$$\widehat{f}_{X,v}^n = \sum_{k=0}^n \hat{a}_k Q_k(x).$$

Définition 16. Les 3 classes de modulateurs $\mathbf{w} = (w_1, \dots, w_n)$ sont :

1. La classe $\mathcal{M}_{\text{CONS}}$ qui désigne la classe des modulateurs constants tels que $w_i = w$ pour $i = 1, \dots, n$, où $0 \leq w \leq 1$,
2. La classe \mathcal{M}_{MON} qui désigne la classe des modulateurs monotones tels que $1 \geq w_1 \geq \dots \geq w_n \geq 0$ pour $i = 1, \dots, n$,
3. La classe \mathcal{M}_{SME} qui désigne la classe des modulateurs de [Sélection de Modèles Emboîtés \(SME\)](#) tels que $\mathbf{w} = (1, \dots, 1, 0, \dots, 0)$.

L'utilisation de ces 3 classes de modulateurs est justifiée dans le papier [BERAN et DUMBGEN \[1998\]](#) via le résultat suivant :

Théorème 8 (Beran et Dumbgen 1998). Soit \mathcal{M} une classe de modulateur parmi $\mathcal{M}_{\text{CONS}}$, \mathcal{M}_{MON} et \mathcal{M}_{SME} . Soit $\mathbb{R}(\mathbf{w})$ le risque de l'estimateur $(w_1 Z_1, \dots, w_n Z_n)$. Soit \mathbf{w}^* qui minimise le risque $\mathbb{R}(\mathbf{w})$ et $\hat{\mathbf{w}}$ qui minimise $\hat{\mathbb{R}}(\mathbf{w})$, alors

$$|\mathbb{R}(\mathbf{w}^*) - \mathbb{R}(\hat{\mathbf{w}})| \rightarrow 0, \quad n \rightarrow +\infty.$$

Il s'agit d'un résultat de consistance qui indique que les classes de modulateurs considérées optimisent asymptotiquement le vrai risque et non son estimation.

L'optimisation sur la classe de modulateur $\mathcal{M}_{\text{CONS}}$ consiste à trouver $\hat{w} = (w, \dots, w)$ tel que

$$\hat{w} = \underset{w \in [0,1]}{\operatorname{argmin}} (1-w)^2 \sum_{k=1}^n (Z_k^2 - \hat{\sigma}_{k,n}^2)_+ + w^2 \sum_{k=1}^n \hat{\sigma}_{k,n}^2.$$

La solution est donnée par

$$\hat{w} = \frac{\sum_{k=1}^n (Z_k^2 - \hat{\sigma}_{k,n}^2)_+}{\sum_{k=1}^n (Z_k^2 - \hat{\sigma}_{k,n}^2)_+ + \sum_{k=1}^n \hat{\sigma}_{k,n}^2}. \quad (2.139)$$

L'optimisation sur la classe \mathcal{M}_{MON} est un problème d'optimisation sous la contrainte de monotonie $w_1 > \dots > w_n$ qui se résout numériquement à l'aide d'algorithmes tel que l'algorithme **Pooled-Adjacent-Violators (PAV)**, décrit dans le papier de **ROBERTSON et collab. [1988]**.

L'optimisation sur la classe \mathcal{M}_{SME} consiste à trouver l'ordre de troncature $\hat{K} \leq n$ tel que

$$\hat{K} = \underset{K \in \{0, \dots, n\}}{\operatorname{argmin}} \sum_{k=K+1}^n (Z_k^2 - \hat{\sigma}_{k,n}^2)_+ + \sum_{k=1}^K \hat{\sigma}_{k,n}^2. \quad (2.140)$$

Le plus simple est de calculer le risque pour chaque valeur de $K = 0, \dots, n$, puis de sélectionner l'ordre de troncature \hat{K} associé à la plus petite valeur du risque.

L'analogie avec le *Normal Mean problem* présenté dans **WASSERMAN [2006]** admet ses limites. Le problème considéré ici est plus compliqué du fait de la corrélation entre les $\{Z_k\}_{k \in \{1, \dots, n\}}$ et la variance qui varie avec k . Il reste encore du travail afin de pouvoir établir des propriétés de minimaxité de l'estimateur via les **FENQ** et leurs polynômes orthogonaux, et appliquer le théorème de Pinsker. Un autre objectif est la mise en place d'une procédure pour construire un intervalle de confiance pour l'estimateur de la densité de probabilité.

2.5.2 Application aux distributions composées

Soit X une variable aléatoire de loi composée $(\mathbb{P}_N, \mathbb{P}_U)$. La variable aléatoire

$$X = \sum_{i=1}^N U_i, \quad (2.141)$$

modélise la charge totale de sinistres d'un portefeuille de contrats d'assurance non-vie sur une période d'exercice. Le nombre de sinistres pour la période d'exercice est modélisé par la variable aléatoire N et le montant unitaire des sinistres par la variable aléatoire U .

L'objet de cette section est de discuter de la meilleure façon d'exploiter des observations sur la sinistralité afin de proposer un estimateur pour la densité de probabilité de la variable aléatoire X . Soit n le nombre de périodes d'exercice observées. Un échantillon X_1, \dots, X_n de charges totales pour chaque période d'exercice est disponible, de la même façon un échantillon N_1, \dots, N_n de nombre de sinistres par période d'exercice est disponible. La taille de l'échantillon des montants de sinistres dépend du nombre de sinistres observés, et est égale à $n_U = \sum_{i=1}^n N_i$. Si la période d'exercice est de un an, il semble évident que n n'excédera jamais 50 ans. Cela fait déjà peu d'observations pour utiliser une méthode statistique non-paramétrique telle que le noyau de Parzen-Rosenblatt par exemple. A cela s'ajoute le problème de la non-stationnarité. La sinistralité peut évoluer suite à un changement dans le comportement des assurés, dans l'environnement financier ou dans la composition du portefeuille d'assurés. Le montant des sinistres peut admettre une tendance à la hausse au fur et à mesure des années à cause de l'inflation par exemple. Cette non-stationnarité nuit à l'exploitation de l'ensemble des données pour effectuer de l'inférence statistique. Trois approches sont considérées pour aborder le problème. Deux approches sont classiques l'une paramétrique et l'autre non-paramétrique. La troisième est une approche dite "semi-paramétrique".

L'approche paramétrique

La première approche consiste à proposer une loi paramétrique pour le nombre de sinistres N , et pour le montant des sinistres U . Cette approche est souvent privilégiée car elle permet d'inclure de l'information a priori basée sur de l'expérience ou de l'expertise, et peut se révéler pertinente avec peu d'observations dans le cas d'un bon choix de modèle. Bien entendu, un mauvais choix de modèle mène à des erreurs catastrophiques. Les paramètres des lois sont inférés statistiquement à l'aide des données via des procédures d'estimation habituelles telles que la méthode des moments ou du maximum de vraisemblance. Plusieurs modèles sont souvent en concurrence, la prise de décision se base sur le résultat de tests statistiques comme le test du χ^2 d'adéquation, le test de Kolmogorov-Smirnov ou encore le test du rapport de vraisemblance. Une fois le modèle calibré, une méthode d'approximation est choisie pour calculer la *f.d.s.* ou la *f.d.r.* de X . Le choix de la méthode d'approximation dépend des modèles sélectionnés pour la distribution du montant des sinistres et du nombre de sinistres. En effet, l'algorithme de Panjer est limité dans son application aux distributions de fréquence appartenant à la famille de Panjer. Les méthodes d'inversion de la transformée de Laplace et celles basées sur les moments se limitent aux distributions ayant une *f.g.m.* bien définie ou des moments définis jusqu'à un certain ordre. Certaines distributions pour les montants de sinistres, appartenant aux distributions des valeurs extrêmes notamment, rendent l'utilisation de ces méthodes difficile voire impossible.

L'approche non-paramétrique

Dans cette seconde approche, aucune hypothèse n'est faite sur la distribution de la fréquence des sinistres ni sur leur intensité. Dans ce contexte l'algorithme de Panjer est inutilisable.

La méthode des moments exponentiels s'adapte aisément à l'estimation statistique. Les formules d'approximation (2.105) et (1.46) se transforment en formule d'estimation en remplaçant les quantités $\mathcal{L}_X[-j \ln(b)]$ pour $j = 0, \dots, \alpha$ par leurs équivalents empiriques

$$\widehat{\mathcal{L}}_X[-j \ln(b)] = \frac{1}{n} \sum_{i=1}^n e^{-j \ln(b) X_i}, \quad j = 0, \dots, \alpha. \quad (2.142)$$

La réinjection de l'estimateur (2.142) dans les formules d'approximations (2.105) et (1.46) conduisent à la définition d'estimateur *plug in* de la f.d.r. et de la densité de probabilité.

L'application de la méthode d'estimation polynomiale conduit à proposer un estimateur de la densité défailante, avec

$$\widehat{g}_X(x) = \sum_{k=0}^n \widehat{a}_k Q_k(x) \widehat{f}_V(x), \quad (2.143)$$

où f_V est la densité d'une loi $\Gamma(\widehat{m}, \widehat{r})$ dont il faut estimer les paramètres. Le choix des paramètres de la mesure de référence dans le cadre du problème d'approximation s'est résumé à faire coïncider les moments de la mesure de référence avec les moments de la distribution des montants de sinistres. Cette conclusion contre-intuitive, car c'est la distribution de X et non celle de U qui est recherchée, l'est encore plus dans le cadre du problème d'estimation. L'exemple de la distribution Poisson composée avec une sévérité des sinistres de loi uniforme, indique qu'il est prudent de fixer $r = 1$ et de choisir m égal à la moyenne empirique du montant des sinistres. Les coefficients de l'estimateur polynomial (2.143) sont estimés par

$$\widehat{a}_k = \frac{w_k}{n} \sum_{i=1}^n Q_k(X_i), \quad k = 1, \dots, n. \quad (2.144)$$

où le modulateur $\mathbf{w} = (w_1, \dots, w_n)$ est le modulateur optimal au sein d'une des classes de modulateurs parmi $\mathcal{M}_{\text{CONS}}$, \mathcal{M}_{MON} et \mathcal{M}_{SME} .

Dans le cadre de l'approche non-paramétrique, il est possible d'avoir recours à l'estimateur à noyau de Parzen Rosenblatt pour estimer la densité et d'estimer la f.d.r. par la fonction de répartition empirique.

L'approche semi-paramétrique

Dans cette troisième approche, une loi paramétrique modélise le nombre de sinistres mais aucune hypothèse n'est faite en ce qui concerne la distribution du montant des prestations. Ce choix est justifié par le manque de données pour la fréquence des sinistres. L'inférence des paramètres et le choix de la distribution de N est effectué pour l'approche paramétrique. Soit U_1, \dots, U_{n_U} l'échantillon des montants de sinistres. L'objectif est de retrouver la distribution de X , à partir des observations de U et d'une loi fixée pour N .

Pour l'algorithme de Panjer, les procédures de discrétisation peuvent s'appliquer directement sur la base d'observations de U . Dans la méthode de discrétisation par l'arrondi, la f.d.s. est remplacée dans l'équation (1.25) par son équivalent empirique, avec

$$\tilde{f}_U(kh) = \begin{cases} \widehat{F}_U\left(\frac{h}{2}\right), & \text{pour } k = 0, \\ \widehat{F}_U\left(kh + \frac{h}{2}\right) - \widehat{F}_U\left(kh - \frac{h}{2}\right), & \text{pour } k \geq 1. \end{cases} \quad (2.145)$$

La f.d.r. empirique est définie par

$$\widehat{F}_U(u) = \frac{1}{n_U} \sum_{j=1}^{n_U} \mathbf{1}_{[0,u]}(U_j). \quad (2.146)$$

où

$$\mathbf{1}_{[0,u]}(U_j) = \begin{cases} 1, & \text{pour } U_j \in [0, u], \\ 0, & \text{sinon.} \end{cases} \quad (2.147)$$

Dans le cadre de la méthode LMM, les moments locaux définis dans l'équation (2.148) sont remplacés par leurs contre-parties empiriques, avec

$$\tilde{f}_U(knh + jh) = \frac{1}{\sum_{l=1}^{n_U} \mathbf{1}_{[nkh, un(k+1)h]}(U_l)} \sum_{l=1}^{n_U} \prod_{i \neq j} \frac{U_l - nkh - ih}{(j-i)h} \mathbf{1}_{[nkh, un(k+1)h]}(U_l), \quad (2.148)$$

où $j = 0, \dots, n$, et $k \in \mathbb{N}$. Une fois la discrétisation de la loi du montant des sinistres effectuée, l'algorithme de Panjer est appliqué de façon classique.

L'application de la méthode des moments exponentiels passe par l'estimation des quantités $\mathcal{L}_X[-j \ln(b)]$ pour $j = 0, \dots, \alpha$. Les observations U_1, \dots, U_{n_U} permettent l'estimation de $\mathbb{L}_U(s)$, avec

$$\widehat{\mathcal{L}}_U(s) = \frac{1}{n_U} \sum_{l=1}^{n_U} e^{sU_l}. \quad (2.149)$$

L'estimateur (2.149) est alors injecté dans l'expression de la transformée de Laplace de X , avec

$$\widehat{\mathcal{L}}_X(s) = \mathcal{G}_N \left[\widehat{\mathcal{L}}_U(s) \right]. \quad (2.150)$$

L'estimateur *plug-in* (2.150) est biaisé et doit pouvoir être amélioré. La réinjection de l'estimateur (2.150) dans les formules d'approximation (2.105) et (1.46) permet la définition d'estimateurs pour la f.d.r. et la densité de X .

En ce qui concerne la méthode d'approximation polynomiale, la mesure de référence ν est une mesure de probabilité gamma $\Gamma(\widehat{m}, \widehat{r})$ dont les paramètres sont évalués comme dans l'approche non-paramétrique. L'estimation des coefficients $\{a_k\}_{k \in \mathbb{N}}$ de la représentation polynomiale (2.143) est plus difficile en utilisant les observations de U plutôt que celles de X . Ces coefficients sont définis par une combinaison linéaire de moments de X . La formule

$$\widehat{a}_k = w_k Z_k, \quad \forall k \in \{1, \dots, K\}. \quad (2.151)$$

où $Z_k = \frac{1}{n} \sum_{i=1}^n Q_k(X_i)$, nécessite l'utilisation des observations de X pour estimer. L'ordre de troncature n'est pas connu à ce stade, même s'il n'excèdera pas n_U a priori. Le problème se situe dans l'évaluation de Z_k qui va être ici remplacé par un estimateur *plug in* basé sur l'estimation empirique des moments de U . Le polynôme Q_k est un polynôme de degré k , il admet donc une expression de la forme

$$Q_k(x) = \sum_{l=0}^k q_{l,k} x^l. \quad (2.152)$$

L'estimation des coefficients de la représentation polynomiale se réécrit

$$\begin{aligned} \widehat{a}_k &= w_k \frac{1}{n} \sum_{i=1}^n \sum_{l=0}^k q_{l,k} X_i^l \\ &= w_k \sum_{l=0}^k q_{l,k} \widehat{\mathbb{E}(X^l)} \end{aligned} \quad (2.153)$$

Si la loi de N appartient à la famille de Panjer alors il existe une relation de récurrence entre les moments successifs de X et les moments de U , avec

$$\widehat{\mathbb{E}(X^{k+1})} = \frac{1}{1-a} \sum_{j=0}^k \binom{k}{j} \left(\frac{k+1}{j+1} a + b \right) \widehat{\mathbb{E}(U^{j+1})} \widehat{\mathbb{E}(X^{k-j})}. \quad (2.154)$$

L'estimateur (2.154) est réinjecté dans l'estimation des coefficients de la représentation polynomiale (2.153). L'estimation des moments de X via (2.154) est biaisée, et l'estimation des coefficients de la représentation polynomiale aussi. C'est le défaut de cette procédure.

Les différentes approches présentées feront l'objet d'un futur travail de recherche. Il est intéressant de comparer les différentes méthodes d'estimation pour les distributions composées en faisant varier notamment le nombre d'observations et de répondre à des questions telles que

- Quel est l'impact d'une erreur de modèle dans l'approche paramétrique ?
- L'approche semi-paramétrique malgré son côté "bricolage" et ses estimateurs biaisés parvient-elle à renvoyer de bons résultats ?

2.6 Références

- ABATE, J., G. CHOUDHURY et W. WHITT. 1995, «On the Laguerre method for numerically inverting Laplace transforms», *INFORMS Journal on Computing*, vol. 8, n° 4, p. 413–427. [50](#)
- ABATE, J. et W. WHITT. 1992, «The Fourier-series method for inverting transforms of probability distributions.», *Queueing Systems*, vol. 10, p. 5–88. [61](#)
- ANDERSON, G. L. et R. J. FIGUEIREDO. 1980, «An adaptative orthogonal-series estimator for probability density functions», *The Annals of Statistics*, p. 347–376. [81](#)

- BALAKRIHSNAN, N. et C. D. LAI. 2009, *Continuous bivariate distributions*, Springer Science and Business Media. 52
- BARNDORFF-NIELSEN, O. 1978, *Information and exponential Families in Statistical Theory*, Wiley. 40
- BERAN, R. 2000, «React scatterplot smoothers : Superefficiency through basis economy», *Journal of the American Statistical Association*, vol. 95, n° 449, p. 155–171. 83
- BERAN, R. et L. DUMBGEN. 1998, «Modulation estimators and confidence sets», *The Annals of Statistics*, p. 1826–1856. 83
- BRUNK, H. D. 1978, «Univariate density estimation by orthogonal series», *Biometrika*, vol. 65, n° 3, p. 521–528. 80
- BUCKLAND, S. T. 1992, «Fitting density functions with polynomials», *Applied Statistics*, p. 63–76. 80
- CENCOV, N. N. 1962, «Estimation of an unknown density function from observations», *Dokl. Akad. Nauk, SSSR*, vol. 147, p. 45–48. 81
- DIACONIS, P. et R. GRIFFITHS. 2012, «Exchangeables pairs of Bernoulli random variables, Krawtchouk polynomials, and Ehrenfest urns», *Australian and New Zealand Journal of Statistics*, vol. 54, n° 1, p. 81–101. 57
- DIACONIS, P., K. KHARE et L. SALOFF-COSTE. 2008, «Gibbs sampling, exponential families and orthogonal polynomials», *Statistical Science*, vol. 23, n° 2, p. 151–178. 57
- DIGGLE, P. et P. HALL. 1986, «The selection of terms in an orthogonal series density estimator», *Journal of the American Statistical Association*, vol. 81, n° 393, p. 230–233. 81
- DUECK, J., D. EDELMANN et D. RICHARDS. 2015, «Distance correlation coefficients for lancaster distributions», *arXiv preprint*. 57
- EAGLESON, G. K. 1964, «Polynomial expansions of bivariate distributions», *The Annals of Mathematical Statistics*, p. 1208–1215. 57
- EFROMOVICH, S. 1999, *Nonparametric curve estimation : methods, theory and applications*, Springer Science and Business Media. 81
- EMBRECHTS, P., M. MAEJIMA et J. L. TEUGELS. 1985, «Asymptotic behavior of compound distributions», *Astin Bulletin*, vol. 15, n° 1, p. 45–48. 44
- GOFFARD, P.-O., S. LOISEL et D. POMMERET. 2015, «A polynomial expansion to approximate the ultimate ruin probability in the compound Poisson ruin model», *Journal of Computational and Applied Mathematics*. 46
- GREBLICKI, W. et M. PAWLAK. 1984, «Hermite series estimates of a probability density and its derivatives», *Journal of Multivariate Analysis*, vol. 15, n° 2, p. 174–182. 81

- HALL, P. 1980, «Estimating a density on the positive half line by the method of orthogonal series», *Annals of the institute of Statistical Mathematics*, vol. 32, n° 1, p. 351–362. [81](#)
- HJORT, N. L. et I. K. GLAD. 1995, «Nonparametric density estimation with a parametric start», *The Annals of Statistics*, p. 882–904. [80](#)
- JIN, T., S. B. PROVOST et J. REN. 2014, «Moment-based density approximations for aggregate losses», *Scandinavian Actuarial Journal*, vol. (ahead-of-print), p. 1–30. [44](#), [60](#)
- KOUDOU, A. E. 1995, *Problèmes de marges et familles exponentielles naturelles*, thèse de doctorat, Toulouse. [57](#)
- KOUDOU, A. E. 1996, «Probabilités de Lancaster», *Expositiones Mathematicae*, vol. 14, p. 247–276. [57](#)
- KOUDOU, A. E. 1998, «Lancaster bivariate probability distributions with Poisson, negative binomial and gamma margins», *Test*, vol. 7, n° 1, p. 95–110. [57](#)
- KRONMAL, R. et M. TARTER. 1968, «The estimation of probability densities and cumulatives by Fourier series methods», *Journal of the American Statistical Association*, p. 925–952. [81](#)
- LANCASTER, H. 1958, «The structure of bivariate distributions», *The Annals of Mathematical Statistics*, p. 719–736. [56](#)
- LANCASTER, H. et E. SENETA. 2005, *Chi-Square Distribution*, John Wiley And sons, Ltd. [56](#)
- MNATSAKANOV, R. M. et K. SARKISIAN. 2013, «A note on recovering the distribution from exponential moments», *Applied Mathematics and Computation*, vol. 219, p. 8730–8737. [61](#)
- MORRIS, C. N. 1982, «Natural Exponential Families with Quadratic Variance Functions», *The Annals of Mathematical Statistics*, vol. 10, n° 1, p. 65–80. [41](#)
- PAPUSH, D. E., G. S. PATRICK et F. PODGAITS. 2001, «Approximations of the aggregate loss distribution», *CAS Forum (winter)*, p. 175–186. [44](#)
- PROVOST, S. et M. JIANG. 2012, «Orthogonal polynomial density estimates : Alternative representation and degree selection», *International Journal of Computational and Mathematical Sciences*, vol. 6, p. 12–29. [80](#)
- PROVOST, S. B. 2005, «Moment-based density approximants», *Mathematica Journal*, vol. 9, n° 4, p. 727–756. [42](#), [44](#), [80](#)
- ROBERTSON, T., F. T. WRIGHT et R. L. DYKSTRA. 1988, *Order restricted statistical inference*, Wiley, New York. [84](#)

- ROLSKI, T., H. SCHMIDL, V. SCHMIDT et J. TEUGELS. 1999, *Stochastic Processes for Insurance and Finance*, Wiley series in probability and statistics. [61](#)
- SCHWARTZ, S. 1967, «Estimation of probability density by an orthogonal series», *The Annals of Mathematical Statistics*, p. 1261–1265. [81](#)
- STEIN, C. M. 1981, «Estimation of the mean of a multivariate normal distribution.», *The Annals of Statistics*, p. 1135–1151. [83](#)
- SUNDT, B. 1982, «Asymptotic behaviour of compound distributions and stop-loss premiums», *Astin Bulletin*, vol. 13, n° 2, p. 89–98. [44](#)
- SZEGÖ, G. 1939, *Orthogonal Polynomials*, vol. XXIII, American mathematical society Colloquium publications. [44](#), [47](#)
- SZÉKELY, G. J., M. L. RIZZO et N. K. BAKIROV. 2007, «Measuring and testing dependence by correlation of distances», *The Annals of Statistics*, vol. 35, n° 6, p. 2769–2794. [57](#)
- WALTER, G. 1977, «Properties of Hermite series estimation of probability density hermite series estimation of probability density», *The Annals of Statistics*, p. 1258–1264. [81](#)
- WASSERMAN, L. 2006, *All of nonparametric statistics*, Springer. [81](#), [84](#)
- WATSON, G. S. 1969, «Density estimation by orthogonal series», *The Annals of Statistics*, p. 1496–1498. [81](#)
- WHITTLE, P. 1958, «On the smoothing of probability density functions», *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 334–343. [80](#)

Chapitre 3

A polynomial expansion to approximate ruin probability in the compound Poisson ruin model

Sommaire

3.1 Introduction	94
3.2 Polynomial expansions of a probability density function	97
3.3 Application to the ruin problem	99
3.3.1 General formula	99
3.3.2 Approximation with Laguerre polynomials	99
3.3.3 Integrability condition	100
3.3.4 On the goodness of the approximation	102
3.3.5 Computation of the coefficients of the expansion	105
3.4 Numerical illustrations	105
3.5 Conclusion	112
3.6 Références	112

Abstract

A numerical method to approximate ruin probabilities is proposed within the frame of a compound Poisson ruin model. The defective density function associated to the ruin probability is projected in an orthogonal polynomial system. These polynomials are orthogonal with respect to a probability measure that belongs to Natural Exponential Family with Quadratic Variance Function (NEF-QVF). The method is convenient in at least four ways. Firstly, it leads to a simple analytical expression of the ultimate ruin probability. Secondly, the implementation does not require strong computer skills. Thirdly, our approximation method does not necessitate any preliminary discretisation step of the claim sizes distribution. Finally, the coefficients of our formula do not depend on initial reserves.

Keywords : compound Poisson model, ultimate ruin probability, natural exponential families with quadratic variance functions, orthogonal polynomials, gamma series expansion, Laplace transform inversion.

3.1 Introduction

A non-life insurance company is assumed to be able to follow the financial reserves' evolution associated with one of its portfolios in continuous time. The number of claims until time t is assumed to be an homogeneous Poisson process $\{N_t\}_{t \geq 0}$, with intensity λ . The successive claim amounts $(U_i)_{i \in \mathbb{N}^*}$, form a sequence of positive, **i.i.d.**, continuous random variables, and independent of $\{N_t\}_{t \geq 0}$, with density function f_U and mean μ . The initial reserves are of amounts $u \geq 0$, and the premium rate is constant and equal to $p \geq 0$. The risk reserve process is therefore defined as

$$R_t = u + pt - \sum_{i=1}^{N_t} U_i.$$

The associated claims surplus process is defined as $S_t = u - R_t$. In this work, we focus on the evaluation of ultimate ruin probabilities (or infinite-time ruin probabilities) defined as

$$\psi(u) = P\left(\inf_{t \geq 0} R_t < 0 \mid R_0 = u\right) = P\left(\sup_{t \geq 0} S_t > u \mid S_0 = 0\right). \quad (3.1)$$

This model is called a compound Poisson model (also known as Cramer-Lundberg ruin model) and has been widely studied in the risk theory literature. See, for example, **ASMUSSEN et ALBRECHER [2010]**; **ROLSKI et collab. [1999]**. We assume that the positive net profit condition holds for this model, namely $p > \lambda\mu$.

A useful technique in applied mathematics consists of determining a probability density function from the knowledge of its Laplace transform. We give here a brief review of the literature involving numerical inversion of Laplace transform and ruin probability approximations. In a few particular cases, the inversion of the Laplace transform associated with ruin probabilities is manageable analytically and leads to closed formula. But in most cases numerical methods are needed. The Laguerre method is an old established method based on the Tricomi-Widder Theorem of 1935. The recovered function takes the form of a sum of Laguerre functions derived through orthogonal

projections. The numerical inversion of Laplace transform using Laguerre series has been originally described in WEEKS [1966] and improved in ABATE et collab. [1995]. In the wake of Laguerre series method, we found attempts in the actuarial science literature to write probability density functions as sum of gamma densities. For instance the early work of Bowers BOWERS [1966] that gave rise to the so-called Beekman-Bowers approximation for the ultimate ruin probability, derived in BEEKMAN [1969]. The idea is to approximate the ultimate ruin probability by the survival function of a gamma distribution using moments fitting. Gamma series expansion has been employed in TAYLOR [1978] and later in ALBRECHER et collab. [2001]. The authors highlight that it is useful to carry out both analytical calculations and numerical approximations. They show that the direct injection of the gamma series expression into integro-differential equations leads to recurrence relations between the expansion's coefficients and therefore characterize them. They focus on the finite-time ruin probability but the results are valid in the infinite-time case by letting the time t tend to infinity. We also want to mention the Picard-Lefevre formula, see PICARD et LEFÈVRE [1997]; RUILLERE et LOISEL [2004], in which ruin probabilities are computed using the orthogonality property of the generalized Appell polynomials, defined in POMMERET [2001] for instance. We explain later that our method, within the frame of ruin probabilities approximation, is closely related to the Laguerre method and represents in fact an improvement.

The numerical inversion via Fourier-series techniques received a great deal of interest. These techniques have been presented for instance in ABATE et WHITT [1992] within a queueing theory setting. For an application within an actuarial framework, we refer to EMBRECHTS et collab. [1993] and ROLSKI et collab. [1999], Chapter 5, Section 5.5. There is also a great body of literature dealing with Laplace transform inversion linked to the Hausdorff moment problem. Probability density function are recovered from different kind of moments. The use of exponential moments and scaled Laplace transform is presented in MNATSAKANOV et SARKISIAN [2013] and has been performed for ruin probabilities computations in MNATSAKANOV et collab. [2014]. In the work of Gzyl et al GZYL et collab. [2013], the maximum entropy applied to fractional exponential moments is employed to determine the probability of ultimate ruin. Recently, Albrecher et al. ALBRECHER et collab. [2010] and Avram et al. AVRAM et collab. [2011] discuss different methods for computing the inverse Laplace transform. The first one consider a numerical inversion procedure based on a quadrature rule that uses as stepping stone a rational approximation of the exponential function in the complex plane. The second one implements and reviews much of the work done using Padé approximants to invert Laplace transform.

There are several usual techniques for calculation of ultimate ruin probabilities. We want to mention a classical iterative method that we will use for comparison purposes. The so-called Panjer's algorithm introduced in PANJER [1981], has been widely used in the actuarial field. One can find an application to the computation of the probability of ultimate ruin in DICKSON [1995].

The method that we propose here consists in an orthogonal projection of the defective probability density function, associated with the probability of ultimate ruin,

with respect to a reference probability measure that belongs to the Natural Exponential Families with Quadratic Variance Function. The desired defective PDF takes the form of an infinite serie of orthonormal polynomials (orthogonal with respect to the aforementioned reference probability measure). The coefficients of the expansion are defined by a scalar product and are computed from the Laplace transform or equivalently from the moments of the distribution. Ruin probability approximations are obtained through truncation of the infinite serie followed by integration. This method permits the recovery of functions from the knowledge of their Laplace transform. Once the set of coefficients of the expansion has been evaluated, ultimate ruin probabilities can be approximated for any initial reserve. It is easy to implement, does not necessitate large computation time and is competitive in terms of accuracy. The approximation of the ruin probability allows manipulations such as integration or reinjection in formulas to derive approximations of distributions that governed other quantities of interest in ruin theory. For instance, the probability density function of the surplus prior to ruin involves the ultimate ruin probability, we refer to [DICKSON \[1992\]](#); [GERBER et SHIU \[1998\]](#) for more details.

This work is also a theoretical background in view of a future statistical application. Many papers deal with statistical estimation of ruin probabilities when observations of the claim sizes are available. The use of a Laplace inversion formula as basis for a nonparametric estimator has already been employed for ruin probabilities estimation in [MNATSAKANOV et collab. \[2008\]](#); [ZHANG et collab. \[2014\]](#), scaled Laplace transform and the maximum of entropy may offer the same possibility and would be based on empirical estimation of the moments. The definition of the coefficients in our method are based on quantities that are well adapted to empirical estimations. By plugging in the estimators of the coefficients, we will obtain a nonparametric estimator taking the form of an orthogonal serie. The investigation of ruin probability statistical estimation will be at the center of a forthcoming paper, for now we only consider ruin probability approximations.

In Section 2, we introduce a density expansion formula based on orthogonal projection within the frame of NEF-QVF. Our main results are developed in Section 3 : the expansion for ultimate ruin probabilities is derived, a sufficient condition of applicability is given and the goodness of the approximation is discussed. Section 4 is devoted to numerical illustrations. Just like what is done in [GZYL et collab. \[2013\]](#), we compare our method to other existing methods, namely Panjer's algorithm, Fast Fourier Transform and scaled Laplace transform inversion.

3.2 Polynomial expansions of a probability density function

Let $F = \{F_\theta, \theta \in \Theta\}$ with $\Theta \subset \mathbb{R}$ be a Natural Exponential Family (NEF), see [BARNDORFF-NIELSEN \[1978\]](#), generated by a probability measure ν on \mathbb{R} such that

$$\begin{aligned} F_\theta(X \in A) &= \int_A \exp\{x\theta - \kappa(\theta)\} d\nu(x) \\ &= \int_A f(x, \theta) d\nu(x), \end{aligned}$$

where $A \subset \mathbb{R}$, $\kappa(\theta) = \log(\int_{\mathbb{R}} e^{\theta x} d\nu(x))$ is the Cumulant Generating Function (CGF) and $f(x, \theta)$ is the density of F_θ with respect to ν . Let X be a random variable F_θ -distributed. The mean of the random variable X , is given by

$$m = E_\theta(X) = \int x dF_\theta(x) = \kappa'(\theta),$$

and its variance is

$$V(m) = \text{Var}_\theta(X) = \int (x - m)^2 dF_\theta(x) = \kappa''(\theta). \quad (3.2)$$

The application $\theta \rightarrow \kappa'(\theta)$ is one to one. Its inverse function $m \rightarrow h(m)$ is defined on $\mathcal{M} = \kappa'(\Theta)$. With a slight change of notation, we can rewrite $F = \{F_m, m \in \mathcal{M}\}$, where F_m has mean m and density $f(x, m) = \exp\{h(m)x - \kappa(h(m))\}$ with respect to ν . A NEF has a Quadratic Variance Function (QVF) if there exists reals ν_0, ν_1, ν_2 such that

$$V(m) = \nu_0 + \nu_1 m + \nu_2 m^2. \quad (3.3)$$

The Natural Exponential Families with Quadratic Variance Function (NEF-QVF) include the normal, gamma, hyperbolic, Poisson, binomial and negative binomial distributions.

Define

$$P_n(x, m) = V^n(m) \left\{ \frac{\partial^n}{\partial m^n} f(x, m) \right\} / f(x, m), \quad (3.4)$$

for $n \in \mathbb{N}$. Each $P_n(x, m)$ is a polynomial of degree n in both m and x . Moreover, if ν is a NEF-QVF, $(P_n)_{n \in \mathbb{N}}$ is a family of orthogonal polynomials with respect to ν in the sense that

$$\langle P_n, P_m \rangle = \int P_n(x, m) P_m(x, m) f(x, m) d\nu(x) = \delta_{nm} \|P_n\|^2, \quad m, n \in \mathbb{N},$$

where δ_{mn} is the Kronecker symbol equal to 1 if $n = m$ and 0 otherwise. For the sake of simplicity, we choose $\nu = P_{m_0}$. Then $f(x, m_0) = 1$ and we write

$$P_n(x) = P_n(x, m_0) = V^n(m_0) \left\{ \frac{\partial^n}{\partial m^n} f(x, m) \right\}_{m=m_0}. \quad (3.5)$$

We also consider in the rest of the paper a normalized version of the polynomials defined in (4.4) with $Q_n(x) = P_n(x) / \|P_n\|$. For an exhaustive review regarding NEF-QVF and their properties, we refer to [MORRIS \[1982\]](#).

We denote by $L^2(\nu)$ the space of functions square integrable with respect to ν .

Proposition 12. *Let ν be a probability measure that generates a NEF-QVF, with associated orthonormal polynomials $\{Q_n\}_{n \in \mathbb{N}}$. Let X be a random variable with density function $\frac{dF_X}{d\nu}$ with respect to ν . If $\frac{dF_X}{d\nu} \in L^2(\nu)$ then we have the following expansion*

$$\frac{dF_X}{d\nu}(x) = \sum_{n=0}^{+\infty} E(Q_n(X))Q_n(x). \quad (3.6)$$

Démonstration. By construction $\{Q_n\}_{n \in \mathbb{N}}$ forms an orthonormal basis of $L^2(\nu)$, and by orthogonal projection we get

$$\frac{dF_X}{d\nu}(x) = \sum_{n=0}^{+\infty} \langle Q_n, \frac{dF_X}{d\nu} \rangle Q_n(x).$$

It follows that

$$\begin{aligned} \langle Q_n, \frac{dF_X}{d\nu} \rangle Q_n(x) &= \int Q_n(y) \frac{dF_X}{d\nu}(y) d\nu(y) \times Q_n(x) \\ &= \int Q_n(y) dF_X(y) \times Q_n(x) \\ &= E(Q_n(X))Q_n(x). \end{aligned}$$

□

Denote by f_ν and f_X the probability density functions of ν and X respectively. Proposition 12 gives an expansion of f_X that takes the following simple form

$$f_X(x) = \sum_{n=0}^{+\infty} a_n Q_n(x) f_\nu(x), \quad (3.7)$$

where $(a_n)_{n \in \mathbb{N}}$ is a sequence of real number called coefficients of the expansion in the rest of the paper, $(Q_n)_{n \in \mathbb{N}}$ is an orthonormal sequence of polynomials with respect to ν , and f_ν is the PDF of ν . The polynomials $(Q_n)_{n \in \mathbb{N}}$ are of degree n in x and can therefore be written as $Q_n(x) = \sum_{i=0}^n q_{i,n} x^i$. Using this last remark, we rewrite the coefficients of the expansion as

$$\begin{aligned} a_n &= E(Q_n(X)) \\ &= \sum_{i=1}^n q_{i,n} E(X^i) \\ &= \sum_{i=1}^n q_{i,n} (-1)^i \left[\frac{d^i \widehat{f_X}(s)}{ds^i} \right]_{s=0}, \end{aligned} \quad (3.8)$$

where $\widehat{f_X}(s) = \int e^{-sx} dF_X(x)$ is the Laplace transform of the random variable X . The approximation of the PDF of X is obtained by truncation of the infinite serie in (3.7). First, we choose a NEF-QVF, then we choose a member of this family characterized by its parameters. These choices are to be made wisely so as to ensure the validity of the expansion and to reach an acceptable level of accuracy that goes along with an order of truncation as small as possible. In the light of the expression of the coefficients (3.8), it seems natural to consider a statistical extension that would lead to a nonparametric estimator of the probability density function. However, it is of interest to start with the probabilistic problem as it is the theoretical basis for a statistical application. The next section shows how to use the orthogonal polynomials and NEF-QVF framework to approximate ruin probabilities.

3.3 Application to the ruin problem

3.3.1 General formula

The ultimate ruin probability in the compound Poisson ruin model is the survival function of a geometric compound distributed random variable

$$M = \sum_{i=1}^N U_i^I, \quad (3.9)$$

where N is a counting random variable having a geometric distribution with parameter $\rho = \frac{\lambda\mu}{p}$, and $(U_i^I)_{i \in \mathbb{N}^*}$ is a sequence of independent and identically distributed nonnegative random variables having PDF defined as $f_{U^I}(x) = \frac{\overline{F_U}(x)}{\mu}$. The distribution of M has an atom at 0 with probability mass $P(N = 0) = 1 - \rho$. The probability measure that governs M is

$$dF_M(x) = (1 - \rho) \delta_0(x) + dG_M(x), \quad (3.10)$$

where dG_M is the continuous part of the probability measure associated to M which admits a defective probability density function with respect to the Lebesgue measure. We denote by g_M the defective probability density function. The ultimate ruin probability follows from the integration of the the continuous part as the discrete part vanishes

$$\psi(u) = P(M > u) = \int_u^{+\infty} dG_M(x).$$

Theorème 9. *Let ν be an univariate distribution having a probability density function with respect to the Lebesgue measure that generates a NEF-QVE. Let $(Q_n)_{n \in \mathbb{N}}$ be the sequence of orthonormal polynomials with respect to ν . If $\frac{dG_M}{d\nu} \in L^2(\nu)$ then*

$$\psi(u) = \sum_{n=0}^{+\infty} a_n \int_u^{+\infty} Q_n(x) d\nu(x), \quad (3.11)$$

where $(a_n)_{n \in \mathbb{N}}$ is defined as in (3.8).

Démonstration. We apply Proposition 12 to get the result. □

3.3.2 Approximation with Laguerre polynomials

We derive an approximation for the ultimate ruin probability, using Theorem 9, combined with truncations of order K of the infinite series in (3.11). It yields

$$\Psi_K(u) = \sum_{n=0}^K a_n \int_u^{+\infty} Q_n(x) d\nu(x), \quad (3.12)$$

the approximated ruin probability with order of truncation K . As the distribution of M has support on \mathbb{R}^+ , we choose the Gamma distribution $\Gamma(r, m)$ with mean parameter m and shape parameter r , that is :

$$d\nu(x) = f_\nu(x) \mathbf{1}_{\mathbb{R}^+}(x) d\lambda(x) = \frac{x^{r-1} e^{-x/m}}{\Gamma(r) m^r} \mathbf{1}_{\mathbb{R}^+}(x) d\lambda(x).$$

The associated orthogonal polynomials are the generalized Laguerre polynomials. By definition, they satisfy the following orthogonality condition

$$\int_0^{+\infty} L_n^{r-1}(x)L_m^{r-1}(x)x^{r-1}e^{-x}dx = \binom{n+r-1}{n}\Gamma(r)\delta_{nm}.$$

The polynomials involved in the ruin probability approximation in (3.12) are the generalized Laguerre polynomials with a slight change in comparison to the definition given in SZEGÖ [1939], namely

$$Q_n(x) = (-1)^n \binom{n+r-1}{n}^{-1/2} L_n^{r-1}(x/m). \quad (3.13)$$

Remarque 3. The Laguerre functions are defined in ABATE *et collab.* [1995] as

$$l_n(x) = e^{-x/2}L_n(x), \quad x \geq 0. \quad (3.14)$$

The application of the Laguerre method consists in representing g_M as a Laguerre serie

$$g_M(x) = \sum_{n=0}^{+\infty} a_n l_n(x). \quad (3.15)$$

The representation (3.15) is close to the proposed expansion when choosing $r = 1$ and $m = 2$ as parameter of the reference measure. The difference lies in the possibility of changing the parameters in our expansion. We will see later that the parametrization is of prime importance.

The defective probability density function associated to G_M has the following expression

$$g_M(x) = \sum_{n=1}^{+\infty} (1-\rho)\rho^n f_{U^I}^{*n}(x). \quad (3.16)$$

Taking the Laplace transform of the defective probability density function defined as in (3.16) yields

$$\widehat{g}_M(s) = \frac{(1-\rho)\rho\widehat{f}_{U^I}(s)}{1-\rho\widehat{f}_{U^I}(s)}, \quad (3.17)$$

where $\widehat{f}_{U^I}(s) = \int e^{-sx}f_{U^I}(x)dx$ is the Laplace transform of f_{U^I} . The Laplace transform of the claim size distribution appears in the formula. This fact limits the application to claim sizes distributions that admit a well defined Laplace transform, namely light-tailed distributions. We want to mention that this problem might be reconsidered once the study of the statistical extension will be done. In MNATSAKOV *et collab.* [2014], approximations of the ruin probability in case of heavy tail claim amounts are derived from the associated nonparametric estimator computed with simulated data.

3.3.3 Integrability condition

We illustrate here how the applicability of the method is subject to the parametrization, namely the choice of m and r . The parametrization permits to ensure the

integrability condition. The adjustment coefficient γ is the unique positive solution of the so-called Cramer-Lundberg equation

$$\widehat{m}_{U^I}(s) = \frac{1}{\rho}, \quad (3.18)$$

where $\widehat{m}_{U^I}(s) = \int_0^{+\infty} e^{sx} f_{U^I}(x) dx$ is the moment generating function of U^I . The integrability condition $\frac{dG_M}{dv} \in L^2(dv)$ is equivalent to

$$\int_0^{+\infty} g_M^2(x) e^{x/m} x^{1-r} dx < \infty. \quad (3.19)$$

In order to ensure this condition, we need the following result.

Theorème 10. *Assume that the Cramer-Lundberg equation (3.18) admits a positive solution γ , then for all $x \geq 0$*

$$g_M(x) \leq C(s_0) e^{-s_0 x}, \quad (3.20)$$

with $s_0 \in [0, \gamma)$ and $C(s_0) \geq 0$.

Démonstration. In order to prove the theorem we need the following lemma regarding the survival function \overline{F}_U of the claim sizes distribution.

Lemme 1. *Let U be a non-negative random variable. Assume that there exists $s_0 > 0$ such that $\widehat{m}_U(s_0) < +\infty$. Then there exists $A(s_0) > 0$ such that for all $x \geq 0$*

$$\overline{F}_U(x) \leq A(s_0) e^{-s_0 x}. \quad (3.21)$$

Démonstration. As $\widehat{m}_U(s_0) < +\infty$, we have

$$\begin{aligned} \widehat{m}_U(s_0) - 1 &= \int_0^{+\infty} (e^{s_0 x} - 1) f_U(x) dx \\ &= s_0 \int_0^{+\infty} \int_0^x e^{s_0 y} f_U(x) dy dx \\ &= s_0 \int_0^{+\infty} e^{s_0 y} \overline{F}_U(y) dy \\ &\geq s_0 \int_0^x e^{s_0 y} \overline{F}_U(y) dy \\ &\geq \overline{F}_U(x) (e^{s_0 x} - 1). \end{aligned}$$

thus, we deduce that $\forall x \geq 0$

$$\overline{F}_U(x) \leq (\widehat{m}_U(s_0) - 1 + \overline{F}_U(x)) e^{-s_0 x}. \quad (3.22)$$

□

The equation (3.18) is equivalent to

$$\rho \widehat{m}_U(s) = 1 + s\mu. \quad (3.23)$$

The fact that γ is a solution of the equation (3.18) implies that $\widehat{m}_U(s) < +\infty, \forall s_0 \in [0, \gamma)$ and by application of Lemma 1, we get the following inequality upon the PDF of U^I

$$f_{U^I}(x) = \frac{\overline{F}_U(x)}{\mu} \leq B(s_0)e^{-s_0x}. \quad (3.24)$$

In view of (3.16), it is easily checked that g_M satisfies the defective renewal equation

$$g_M(x) = \rho(1 - \rho)f_{U^I}(x) + \rho \int_0^x f_{U^I}(x - y)g_M(y)dy. \quad (3.25)$$

We can therefore bound g_M as in (3.20)

$$\begin{aligned} g_M(x) &\leq \rho(1 - \rho)f_{U^I}(x) + \int_0^{+\infty} f_{U^I}(x - y)g_M(y)dy \\ &\leq \rho(1 - \rho)B(s_0)e^{-s_0x} + B(s_0)e^{-s_0x} \int_0^{+\infty} e^{s_0y}g_M(y)dy \\ &= (\rho(1 - \rho) + \widehat{g}_M(-s_0))B(s_0)e^{-s_0x} \\ &= C(s_0)e^{-s_0x}. \end{aligned}$$

□

Applying Theorem 10 yields a sufficient condition in order to use the polynomial expansion.

Corollaire 4. For $\frac{1}{m} < 2\gamma$ and $r = 1$, the integrability condition (3.19) is satisfied.

The choice of the parameter m is important. The Laguerre method, briefly described in Remark 3, does not offer the possibility of changing the parameter. In the next subsection, we shed light on another key aspect of parametrization.

3.3.4 On the goodness of the approximation

The approximation is obtained through the truncation of an infinite serie. The higher the order of truncation gets, the better the approximation is. Our goal is to work out an efficient numerical method that combines high accuracy and small computation time. We want to minimize the number of coefficients to compute. The sequence of coefficients of the expansion must decrease as fast as possible. we assume that $\frac{dG_M}{dv} \in L^2(v)$ in order to apply the method and consequently

$$\left\langle \frac{dG_M}{dv}, \frac{dG_M}{dv} \right\rangle = \sum_{n=0}^{+\infty} a_n^2 < +\infty. \quad (3.26)$$

The Parseval type relation (3.26) implies that the coefficients of the expansion decrease towards 0 as n tends to infinity. The question is how fast is the decay? We address this problem using generating function theory, just like what is done in [ABATE et collab. \[1995\]](#). Taking the Laplace transform of g_M defined as a polynomial expansion yields

$$\begin{aligned} \widehat{g}_M(s) &= \int_0^{+\infty} e^{-sx} \sum_{n=0}^{+\infty} a_n Q_n(x) f_v(x) dx \\ &= \sum_{n=0}^{+\infty} a_n \int_0^{+\infty} e^{-sx} Q_n(x) f_v(x) dx. \end{aligned}$$

The orthonormal polynomial system $(Q_n)_{n \in \mathbb{N}}$ are Laguerre's one, defined as

$$Q_n(x) = \frac{(-1)^n}{\sqrt{\binom{n+r-1}{n}}} L_n^{r-1}\left(\frac{x}{m}\right), \quad (3.27)$$

where $L_n^{r-1}(x) = \sum_{i=0}^n \binom{n+r-1}{n-i} \frac{(-x)^i}{i!}$. We refer to SZEGÖ [1939] for this last definition. We then have

$$\begin{aligned} \widehat{g}_M(s) &= \sum_{n=0}^{+\infty} a_n \int_0^{+\infty} e^{-sx} \frac{(-1)^n}{\sqrt{\binom{n+r-1}{n}}} L_n^{r-1}\left(\frac{x}{m}\right) \frac{x^{r-1} e^{-x/m}}{\Gamma(r) m^r} dx \\ &= \sum_{n=0}^{+\infty} a_n \frac{(-1)^n}{\sqrt{\binom{n+r-1}{n}}} \int_0^{+\infty} e^{-sx} \sum_{i=0}^n \binom{n+r-1}{n-i} \frac{(-x)^i}{i! m^r} \frac{x^{r-1} e^{-x/m}}{\Gamma(r) m^r} dx \\ &= \sum_{n=0}^{+\infty} a_n \frac{(-1)^n}{\sqrt{\binom{n+r-1}{n}}} \sum_{i=0}^n \frac{(n+r-1)!}{(n-i)!(r+i-1)!(r-1)!i!} (-1)^i \int_0^{+\infty} e^{-sx} \frac{x^{r+i-1} e^{-x/m}}{m^{r+i}} dx \\ &= \sum_{n=0}^{+\infty} a_n (-1)^n \sqrt{\binom{n+r-1}{n}} \sum_{i=0}^n \binom{n}{i} (-1)^i \int_0^{+\infty} e^{-sx} \frac{x^{r+i-1} e^{-x/m}}{m^{r+i} \Gamma(r+i)} dx \\ &= \sum_{n=0}^{+\infty} a_n (-1)^n \sqrt{\binom{n+r-1}{n}} \sum_{i=0}^n \binom{n}{i} (-1)^i \left(\frac{1}{sm+1}\right)^{r+i} \\ &= \sum_{n=0}^{+\infty} a_n (-1)^n \sqrt{\binom{n+r-1}{n}} \left(\frac{1}{sm+1}\right)^r \left(\frac{sm}{sm+1}\right)^n \\ &= \sum_{n=0}^{+\infty} a_n I_\nu(n) G_\nu(s) H_\nu(s)^n, \end{aligned}$$

where $I_\nu(n) = (-1)^n \sqrt{\binom{n+r-1}{n}}$, $G_\nu(s) = \left(\frac{1}{sm+1}\right)^r$ and $H_\nu(s) = \frac{sm}{sm+1}$. Define $b_n = I_\nu(n) a_n, \forall n \in \mathbb{N}$, we have

$$\widehat{g}_M(s) = G_\nu(s) B(H_\nu(s)), \quad (3.28)$$

where $B(z) = \sum_{n=0}^{+\infty} b_n z^n$ is the generating function of the sequence $(b_n)_{n \in \mathbb{N}}$. The generating function B is expressed as a function of the Laplace transform of g_M via a change of variable

$$\begin{aligned} B(z) &= \frac{\widehat{g}_M(H_\nu^{-1}(z))}{G_\nu(H_\nu^{-1}(z))} \\ &= (1-z)^{-r} \widehat{g}_M\left(\frac{z}{m(1-z)}\right). \end{aligned} \quad (3.29)$$

In generating function theory, it is possible to study the decay of a sequence by considering its radius of convergence. If the radius of convergence of $B(z)$ is greater than one then the sequence of coefficients admits a decay that is geometrically fast. Non-geometric convergence occurs when B admits a singularity on the unit circle $\{z : |z| = 1\}$. The parameters of the expansion permit to alter the form of the generating function in order to make it simpler. Sometimes it gets so simple that an exact formula is obtained. The two following examples may help convince the reader.

Example 1. In *ABATE et collab. [1995]*, several attempts are made to expand PDF when the convergence is not geometrically fast. One of those cases is the expansion of a Gamma PDF $\Gamma(\alpha, \beta)$, recall that the PDF is given by

$$f_X(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad (3.30)$$

with associated Laplace transform

$$\widehat{f}_X(s) = \left(\frac{1}{1 + \beta s} \right)^\alpha. \quad (3.31)$$

The generating function of the coefficient defined in (3.29) is

$$B(z) = \frac{m^\alpha (1-z)^{\alpha-r}}{(m-z(\beta-m))^\alpha}. \quad (3.32)$$

It is easily checked that taking $r = \alpha$ and $m = \beta$ yields $B(z) = 1$. Thus we deduce that $a_0 = 1$ and $a_n = 0, \forall n \geq 1$.

Example 2. The exact ultimate ruin probability for the classical ruin model is available in a closed form when the claim sizes are governed by an exponential distribution $\Gamma(1, \beta)$. In this particular case, the integrated tail distribution is also an exponential distribution with the same parameter as the claim amounts. Assume that the claim sizes are $\Gamma(1, \beta)$ – distributed. The Laplace transform of the defective PDF g_M is

$$\widehat{g}_M(s) = \frac{\rho}{1 + \frac{\beta}{(1-\rho)} s}.$$

The generating function of the coefficients is then

$$B(z) = \frac{\rho m (1-z)^{1-r}}{m - z \left(\frac{\beta}{1-\rho} - m \right)}.$$

It is easily checked that taking $r = 1$ and $m = \frac{\beta}{1-\rho}$ yields $B(z) = \rho$ and therefore $a_0 = \rho$ and $a_n = 0, \forall n \geq 1$. Thus $g_M(x) = \rho \frac{1-\rho}{\beta} e^{-\frac{1-\rho}{\beta} x}$ and $\psi(u) = \rho e^{-\frac{1-\rho}{\beta} u}$ which is indeed the exact ultimate probability in the studied case.

These two examples show how to tune the parameters to get a simpler generating function. This method could also be used as a mathematical tool to perform analytical Laplace transform inversion. The choice of the parameters is not automatic as one has to look at the generating function and choose smartly the parameters to have good results. Sometimes the generating function B takes a tedious form which leaves us with a difficult call regarding the parameters.

3.3.5 Computation of the coefficients of the expansion

The coefficients are obtained from the derivative of the generating function defined in (3.29)

$$a_n = \frac{1}{I_\nu(n)n!} \left[\frac{d^n}{dz^n} B(z) \right]_{z=0}. \quad (3.33)$$

Direct evaluation is doable using a computational software program. However if the expression of B is tedious then one might use an approximation procedure. The derivative in (3.33) can be expressed via Cauchy contour integral

$$a_n = \frac{1}{I_\nu(n)2\pi i} \int_{C_r} \frac{B(z)}{z^{n+1}} dz, \quad (3.34)$$

where C_r is a circle about the origin of radius $0 < r < 1$. Making the change of variable $z = re^{iw}$ yields

$$a_n = \frac{1}{I_\nu(n)2\pi r^n} \int_0^{2\pi} B(re^{iw}) e^{-iw} dw. \quad (3.35)$$

The integrals in (3.35) are approximated through a trapezoidal rule

$$\begin{aligned} a_n &\approx \bar{a}_n \\ &= \frac{1}{I_\nu(n)2\pi r^n} \sum_{j=1}^{2n} (-1)^j \Re(B(re^{\pi j i/n})) \\ &= \frac{1}{I_\nu(n)2\pi r^n} \left\{ B(r) + (-1)^n Q(-r) + 2 \sum_{j=1}^n (-1)^j \Re(B(re^{\pi j i/n})) \right\}, \end{aligned}$$

where $\Re(z)$ denotes the real part of some complex number z . The goodness of this approximation procedure is widely studied in [ABATE et collab. \[1995\]](#).

3.4 Numerical illustrations

We analyse the convergence of the sum in our method towards known exact values of ruin probabilities with gamma distributed claim sizes. Gamma distributions (with an integer-valued shape parameter) belongs to Phase-type distributions and we have explicit formulas for ruin probabilities that allow us to assess the accuracy of our approximations, see Chapter 8 of [ASMUSSEN et ALBRECHER \[2010\]](#). A comparison is done with the results obtained with the Fast Fourier Transform, the scaled Laplace transform inversion and Panjer's algorithm. Panjer's algorithm is applied in its basic form as we are trying to recover probabilities of a geometric compound distribution. The distribution of the random variable M, defined in (3.9), is approximated as follows

$$\begin{aligned} P(M = nh) &\approx g_n \\ &= \frac{\rho}{1 - \rho f_0} \sum_{j=1}^n f_j g_{n-j} \quad \forall n \geq 1, \end{aligned} \quad (3.36)$$

where h is the bandwidth and $f_j = F_{U^1} \left(jh + \frac{h}{2} \right) - F_{U^1} \left(jh - \frac{h}{2} \right)$, $\forall j \in \mathbb{N}$. This arithmetization design has been recommended in a recent paper, see [EMBRECHTS et FREI \[2009\]](#).

The algorithm is initialized with $g_0 = G_N(0)$, where $G_N(s) = E(s^N)$ is the probability generating function of N also defined in (3.9). The probability of ultimate ruin is approximated by

$$\tilde{\psi}(u) = 1 - \sum_{i=0}^{\lfloor u/h \rfloor} g_i.$$

The scaled Laplace transform technique has been presented in MNATSAKANOV et SARKISIAN [2013] and applied to the approximation of ruin probabilities recently in MNATSAKANOV et collab. [2014]. The ultimate ruin probability is approximated by

$$\psi_{\alpha,b}(u) = \frac{\ln(b)[\alpha b^{-u}]\Gamma(\alpha+2)}{\alpha\Gamma([\alpha b^{-u}]+1)} \sum_{j=0}^{\alpha-[\alpha b^{-u}]} \frac{(-1)^j \widehat{\psi}((j+[\alpha b^{-u}])\ln(b))}{j!(\alpha-[\alpha b^{-u}]-j)!}, \quad (3.37)$$

where

$$\widehat{\psi}(s) = \frac{1}{s} - \frac{1-\rho}{s - \frac{\lambda}{p} (1 - \widehat{f}_U(s))} \quad (3.38)$$

is the Laplace transform of the ruin probability. The authors of MNATSAKANOV et collab. [2014] recommend to consider the approximations of the ruin probabilities at given initial reserves defined by

$$u_i = \ln\left(\frac{\alpha}{\alpha-i+1}\right) / \ln(b), \quad i = 1, \dots, \alpha. \quad (3.39)$$

A linear interpolation at the points $\psi(u_i)$ yields the final approximation of $\psi(u)$. We are very thankful to the anonymous referee for asking Dr Hakobyan to provide us with approximated values of the ruin probabilities for the examples we have considered in this work. Regarding the Fourier series method, we apply exactly the procedure described in ROLSKI et collab. [1999], Chapter 5, Section 5.5. Throughout this section, we study the relative difference between the exact ruin probability value and its approximation

$$\Delta\psi(u) = \frac{\psi(u) - \psi_{Approx}(u)}{\psi(u)}. \quad (3.40)$$

For the first example, we suppose that the intensity of the Poisson process is $\lambda = 1$, and that premiums are collected at $p = 5$. We assume that the claim sizes are $\Gamma(2, 1)$ -distributed. In this case, the ultimate ruin probability is

$$\psi(u) = 0.461861e^{-0.441742u} - 0.0618615e^{-1.35826u}.$$

The adjustment coefficient is $\gamma = \frac{1}{24} (19 - \sqrt{265})$. According to Corollary 4, we have to set $r = 1$ and $m > \frac{1}{2\gamma}$. The expression of the generating function of the coefficient is tedious and does not help to choose the parameter.

Figure 3.1 displays the difference between the exact and the polynomial approximation of the ultimate ruin probability for different value of $m \in \left\{ \frac{2}{\gamma}, \frac{1}{\gamma}, \frac{4}{5\gamma}, \frac{4}{7\gamma} \right\}$. We set $K = 40$, which sounds like a good compromise between accuracy, computation time and numerical difficulties as $a_{40} \approx 10^{-9}$ when $m = 1/\gamma$. Intuitively, we thought that the best call would be $m = 1/\gamma$, however Figure 3.1 suggests that there exists an optimal

value for m between $1/\gamma$ and $4/7\gamma$. As we do not have justification for the optimal value of m , we set $m = 1/\gamma$ and $K = 40$ for the comparison to the other methods. We set $h = 0.1$ for the Panjer's algorithm and $\{\alpha = 400, b = 1.415\}$ for the scaled Laplace transform technique.

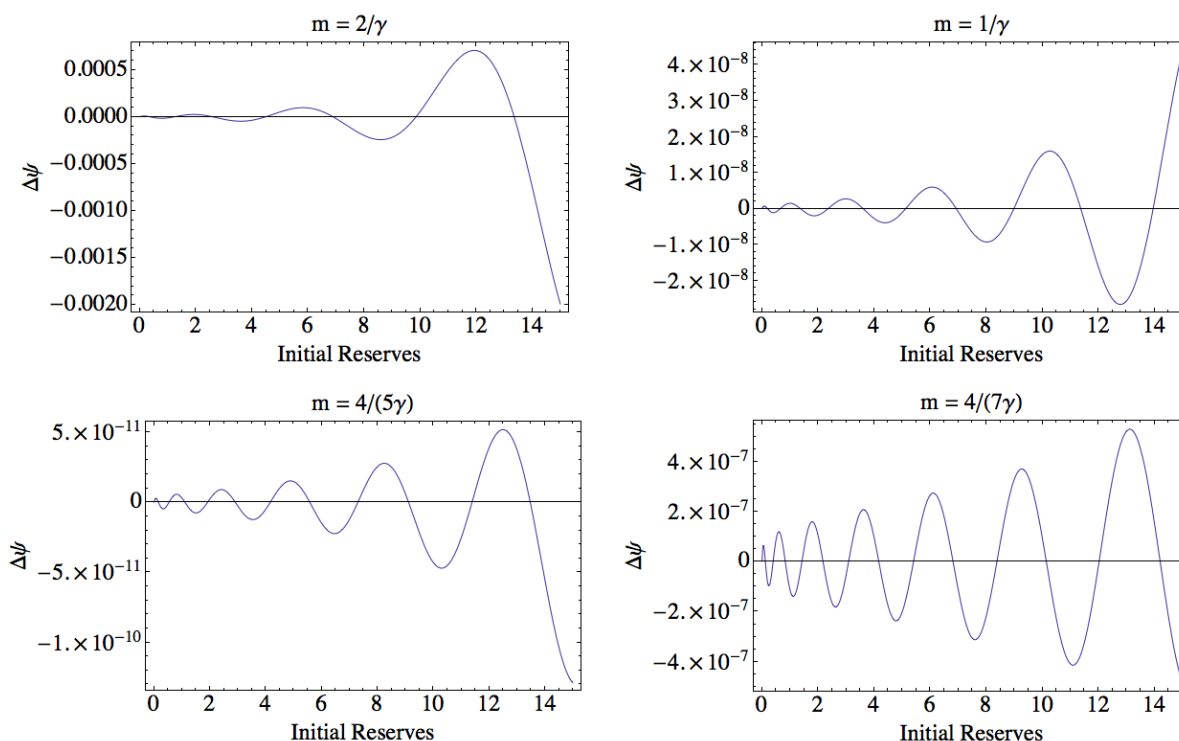


FIGURE 3.1 – Difference between exact and polynomial approximations of ruin probabilities for $\Gamma(2, 1)$ -distributed claim sizes using different parametrizations and an order of truncation $K = 40$.

Figure 3.2 shows the relative difference between exact and approximated ruin probabilities obtained via different numerical methods. Table 3.1 gives numerical values of the ultimate ruin probability computed via the different methods.

The polynomial expansion gives the best results, even better than the Fourier series based method in this case. In the second case, we assume that the claim amounts are $\Gamma(3, 1)$ -distributed and the premium are collected at $p = 3.6$. This setting implies a smaller safety loading than in the previous case and therefore greater ruin probabilities with respect to a given initial reserve. In this case, the exact probability of ultimate ruin is given by

$$\begin{aligned} \psi(u) &= 0.861024e^{-0.0859017u} - 0.0196231e^{-1.31816u} \sin(0.450173u) \\ &\quad - 0.0276908e^{-1.31816u} \cos(0.450173u) \end{aligned}$$

The adjustment coefficient is $\gamma \approx 0.086$.

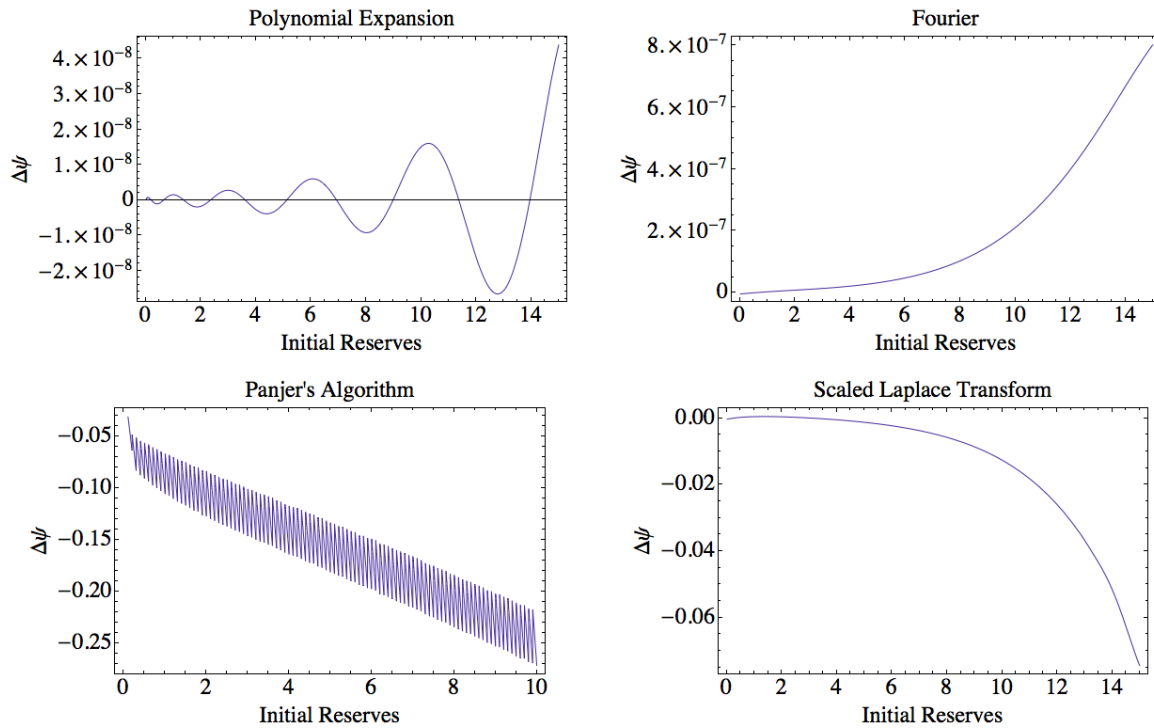


FIGURE 3.2 – Difference between exact and approximated ruin probabilities for $\Gamma(2,1)$ -distributed claim sizes.

Figure 3.3 displays the difference between exact and approximated ruin probabilities. We set $m = 1/\gamma$ and $K = 40$ for the polynomial expansion and $\{a = 400, b = 1.055\}$ for the scaled Laplace transform technique. Because we need to compute ruin probabilities for large initial reserves, we had to set $h = 1$ in order to get an acceptable computational time using Panjer's algorithm. The Fourier series based method performs better than the other method in this case. Polynomial expansion and scaled Laplace transform give close results. The performance of the polynomial expansion are influenced by the settings of the ruin model. We need an higher order of truncation to get the same results as in the previous example.

Our method enables us to approximate ruin probabilities in cases that are relevant for applications but where no formulas are currently available. In this last example, we suppose that the claim sizes are uniformly distributed between 0 and 100 and the premium are collected at $p = 80$. The adjustment coefficient is $\gamma \approx 0.013$. As the Fourier series based method did a good job in the previous cases, we take ruin probabilities approximated via the Fourier series technique as benchmark to assess the accuracy of the polynomial expansion. We set $K = 40$ and $m = 1/\gamma$ for the polynomial expansion and $\{\alpha = 200, b = 1,0045\}$ for the scaled Laplace transform inversion.

Figure 3.4 displays the relative difference between ruin probabilities approximated the polynomial expansion and the scaled Laplace transform inversion technique with respect to the approximation derived from the numerical inversion of the Fourier transform. Table 3.3 gives approximated ruin probabilities for some initial reserves obtained

CHAPITRE 3. A POLYNOMIAL EXPANSION TO APPROXIMATE RUIN PROBABILITY IN THE COMPOUND POISSON RUIN MODEL

i	u_i	Exact	Polynomial Expansion	Fourier transform inversion	Scaled Laplace transform inversion
40	0.295528	0.363928	0.363928	0.363928	0.363907
80	0.633837	0.322916	0.322916	0.322916	0.322815
120	1.01723	0.279149	0.279149	0.279149	0.279022
160	1.45959	0.233859	0.233859	0.233859	0.233746
200	1.98243	0.188205	0.188205	0.188205	0.188129
240	2.62167	0.143304	0.143304	0.143304	0.143275
280	3.44446	0.100282	0.100282	0.100282	0.100299
320	4.60063	0.0604002	0.0604002	0.0604002	0.0604562
360	6.56208	0.0254366	0.0254366	0.0254366	0.0255143
400	17.26	0.000225548	0.000225548	0.000225548	0.000258617

TABLEAU 3.1 – Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $\Gamma(2, 1)$ -distributed.

i	u_i	Exact	Polynomial Expansion	Fourier transform inversion	Scaled Laplace transform inversion
40	1.91605	0.727722	0.7277	0.727722	0.726683
80	4.10946	0.60488	0.604914	0.60488	0.604143
120	6.59516	0.488625	0.488556	0.488625	0.488154
160	9.46321	0.381922	0.381976	0.381922	0.381667
200	12.853	0.285439	0.285436	0.285439	0.285365
240	16.9975	0.199938	0.19991	0.199938	0.200009
280	22.332	0.126439	0.126464	0.126439	0.126612
320	29.828	0.0664098	0.0663945	0.0664098	0.0666329
360	42.545	0.0222743	0.0222812	0.0222743	0.0224793
400	111.905	0.0000575741	0.0000572446	0.0000575747	0.0000823883

TABLEAU 3.2 – Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $\Gamma(3, 1)$ -distributed.

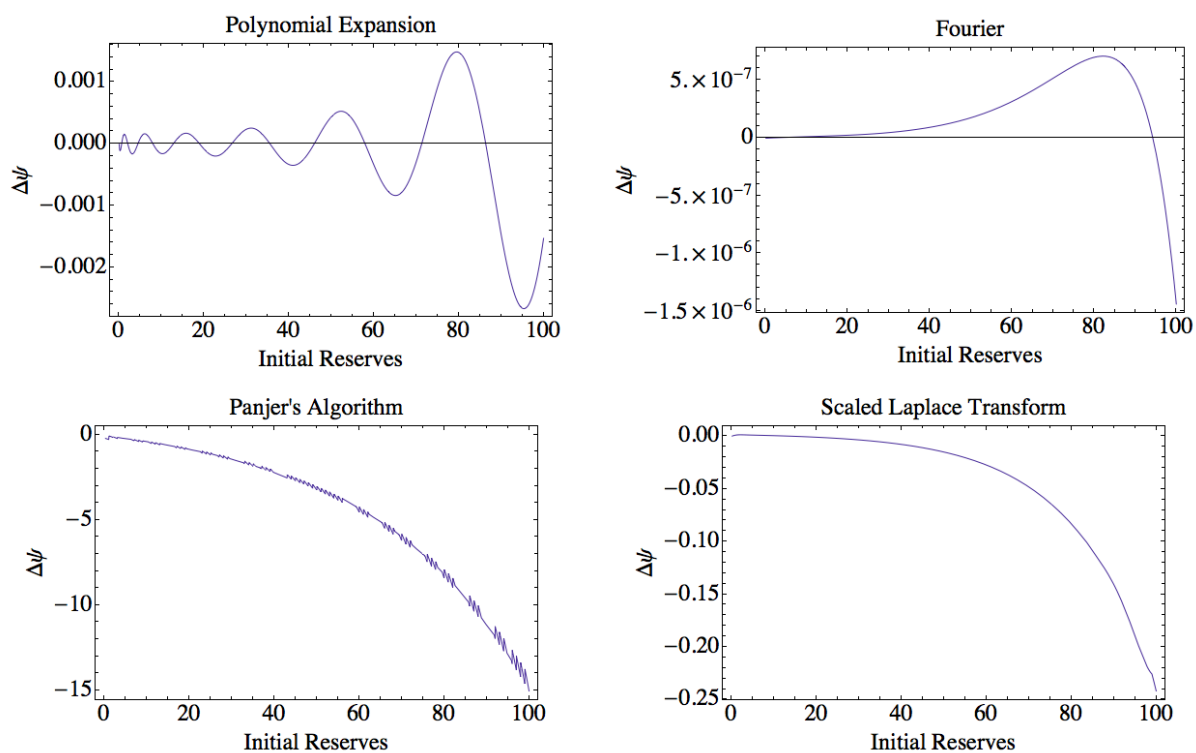


FIGURE 3.3 – Difference between exact and approximated ruin probabilities for $\Gamma(3,1)$ -distributed claim sizes.

via the Fourier transform inversion, the polynomial expansion, and the scaled Laplace transform inversion.

The difference between the polynomial approximation and the Fourier series approximation is small, and we see on Figure 3.4 that the ruin probability curves overlap.

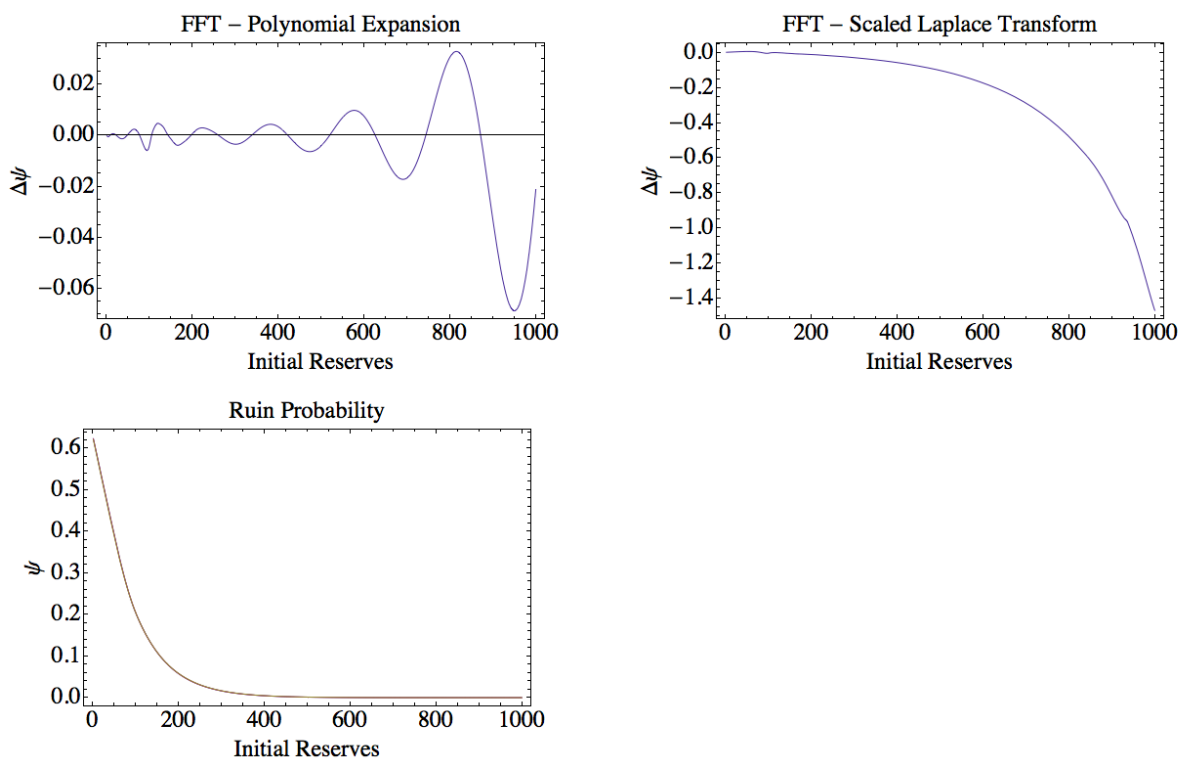


FIGURE 3.4 – (Top left) Difference between the Fourier series technique based and the polynomial approximation of the ultimate ruin probability for $U [0, 100]$ -distributed claim sizes. (Top right) Difference between the Fourier series technique and the scaled Laplace transform inversion approximation of the ultimate ruin probability for $U [0, 100]$ -distributed claim sizes. (Bottom left) Probabilities of ultimate ruin obtained via the Fourier series technique, the polynomial expansion and the scaled Laplace transform method.

i	u_i	Polynomial expansion	Fourier transform inversion	Scaled Laplace transform
20	22.2322	0.518736	0.518768	0.515691
40	48.3113	0.394639	0.394764	0.391591
60	77.8541	0.270484	0.27034	0.269242
80	111.924	0.173775	0.174403	0.174105
100	152.163	0.105036	0.10483	0.105196
120	201.311	0.0561043	0.0561651	0.0567006
140	264.47	0.0252093	0.0251985	0.0257021
160	352.957	0.00817819	0.00819723	0.00853082
180	501.969	0.0012417	0.00123727	0.00136279
200	1180.05	$2.479493055 \times 10^{-7}$	$2.28077799077 \times 10^{-7}$	$1.07450197444 \times 10^{-6}$

TABLEAU 3.3 – Probabilities of ultimate ruin approximated via the Fourier series method, the polynomial expansion and the scaled Laplace transform inversion when the claim amounts are $U [0, 100]$ -distributed.

3.5 Conclusion

Our method provides a very good approximation of the ruin probability when the claim sizes distribution is light-tailed. We obtained a theoretical result that ensures the validity of our expansions. As expected, the numerical results show the superiority of the Fourier series based method in term of accuracy. Nevertheless, our method provides an approximation of a simple form for the whole ruin function and allows reinjection to derive approximations of other quantities of interest in ruin theory. Another advantage is the possibility of a statistical extension that will lead to a nonparametric estimation of ruin probabilities just like the scaled Laplace transform inversion and maximum entropy methods. The great results in terms of accuracy are promising and it will be interesting to consider statistical application. It is also worth noting that this method can be easily adapted to a multivariate problem. The inversion of a bivariate Laplace transform will be at the center of a forthcoming paper.

Acknowledgement

The authors would like to thank X. Guerrault for helpful discussions. The authors are also very grateful to an anonymous referee for his comments which improve greatly the quality of this paper and for the data he has provided regarding the approximation via the scaled Laplace transform inversion. This work is partially funded by the Research Chair *Actuariat Durable* sponsored by Milliman, and AXA research fund sponsored by AXA.

3.6 Références

- ABATE, J., G. CHOUDHURY et W. WHITT. 1995, «On the Laguerre method for numerically inverting Laplace transforms», *INFORMS Journal on Computing*, vol. 8, n° 4, p. 413–427. [95](#), [100](#), [102](#), [104](#), [105](#)
- ABATE, J. et W. WHITT. 1992, «The Fourier-series method for inverting transforms of probability distributions.», *Queueing Systems*, vol. 10, p. 5–88. [95](#)
- ALBRECHER, H., F. AVRAM et D. KORTSCHAK. 2010, «On the efficient evaluation of ruin probabilities for completely monotone claim distributions», *Journal of Computational and Applied Mathematics*, vol. 233, n° 10, p. 2724–2736. [95](#)
- ALBRECHER, H., J. TEUGELS et R. TICHY. 2001, «On a gamma series expansion for the time dependent probability of collective ruin», *Insurance : Mathematics and Economics*, , n° 29, p. 345–355. [95](#)
- ASMUSSEN, S. et H. ALBRECHER. 2010, *Ruin Probabilities, Advanced Series on Statistical Science and Applied Probability*, vol. 14, World Scientific. [94](#), [105](#)
- AVRAM, F., D. CHEDOM et A. HORVARTH. 2011, «On moments based Padé approximations of ruin probabilities», *Journal of Computational and Applied Mathematics*, vol. 235, n° 10, p. 321–3228. [95](#)

- BARNDORFF-NIELSEN, O. 1978, *Information and exponential Families in Statistical Theory*, Wiley. 97
- BEEKMAN, J. 1969, «Ruin function approximation», *Transaction of Society of Actuaries*, vol. 21, n° 59 AB, p. 41–48. 95
- BOWERS, N. 1966, «Expansion of probability density functions as a sum of gamma densities with applications in risk theory», *Transaction of Society of Actuaries*, vol. 18, n° 52, p. 125–137. 95
- DICKSON, D. 1992, «On the distribution of the surplus prior to ruin», *Insurance : Mathematics and Economics*, vol. 11, n° 3, p. 191–207. 96
- DICKSON, D. C. M. 1995, «A review of Panjer’s recursion formula and it’s applications», *British Actuarial Journal*, vol. 1, n° 1, p. 107–124. 95
- EMBRECHTS, P. et M. FREI. 2009, «Panjer recursion versus FFT for compound distribution», *Mathematical Methods of Operations Research*, vol. 69, p. 497–508. 105
- EMBRECHTS, P., P. GRÜBEL et S. M. PITTS. 1993, «Some applications of the fast Fourier transform algorithm in insurance mathematics», *Statistica Neerlandica*, vol. 41, p. 59–75. 95
- GERBER, H. U. et E. SHIU. 1998, «On the time value of ruin», *North American Actuarial Journal*, vol. 2, n° 1, p. 48–72. 96
- GZYL, H., P. NOVI-INVERARDI et A. TAGLIANI. 2013, «Determination of the probability of ultimate ruin probability by maximum entropy applied to fractional moments», *Insurance : Mathematics and Economics*, vol. 53, n° 2, p. 457–463. 95, 96
- MNATSAKANOV, R., L. RUYMGAART et F. RUYMGAART. 2008, «Nonparametric estimation of ruin probabilities given a random sample of claims», *Mathematical Methods of Statistics*, vol. 17, n° 1, p. 35–43. 96
- MNATSAKANOV, R. M. et K. SARKISIAN. 2013, «A note on recovering the distribution from exponential moments», *Applied Mathematics and Computation*, vol. 219, p. 8730–8737. 95, 106
- MNATSAKANOV, R. M., K. SARKISIAN et A. HAKOBYAN. 2014, «Approximation of the ruin probability using the scaled Laplace transform inversion», *Working Paper*. 95, 100, 106
- MORRIS, C. N. 1982, «Natural Exponential Families with Quadratic Variance Functions», *The Annals of Mathematical Statistics*, vol. 10, n° 1, p. 65–80. 97
- PANJER, H. H. 1981, «Recursive evaluation of a family of compound distributions», *Astin Bulletin*, vol. 12, n° 1, p. 22–26. 95
- PICARD, P. et C. LEFÈVRE. 1997, «The probability of ruin in finite time with discrete claim size distribution», *Scandinavian Actuarial Journal*, p. 58–69. 95

- POMMERET, D. 2001, «Orthogonal and pseudo-orthogonal multi-dimensional Appell polynomials», *Applied Mathematics and Computation*, vol. 117, n° 2, p. 285–299. [95](#)
- ROLSKI, T., H. SCHMIDLI, V. SCHMIDT et J. TEUGELS. 1999, *Stochastic Processes for Insurance and Finance*, Wiley series in probability and statistics. [94](#), [95](#), [106](#)
- RUILLERE, D. et S. LOISEL. 2004, «Another look at the Picard-Lefèvre formula for finite-time ruin probability», *Insurance : Mathematics and Economics*, vol. 35, n° 2, p. 187–203. [95](#)
- SZEGÖ, G. 1939, *Orthogonal Polynomials*, vol. XXIII, American mathematical society Colloquium publications. [100](#), [103](#)
- TAYLOR, G. 1978, «Representation and explicit calculation of finite-time ruin probabilities», *Scandinavian Actuarial Journal*, p. 1–18. [95](#)
- WEEKS, W. T. 1966, «Numerical inversion of Laplace transforms using Laguerre functions», *Journal of the ACM*, vol. 13, n° 3, p. 419–429. [95](#)
- ZHANG, Z., H. YANG et H. YANG. 2014, «On a nonparametric estimator for ruin probability in the classical risk model», *Scandinavian Actuarial Journal*, vol. 2014, n° 4, p. 309–338. [96](#)

Chapitre 4

Polynomial approximations for bivariate aggregate claims amount probability distributions

Sommaire

4.1 Introduction	116
4.2 Expression of the joint density	118
4.2.1 Orthogonal polynomials associated with NEF-QVF	118
4.2.2 Bivariate probability measures and their Laplace transform . . .	119
4.3 Application to a bivariate aggregate claims amount distribution . . .	121
4.3.1 General formula	121
4.3.2 Integrability condition	123
4.4 Downton Bivariate Exponential distribution and Lancaster probabilities	125
4.5 Numerical illustrations	127
4.5.1 Approximation of common bivariate exponential distributions .	127
4.5.2 Approximation of the distribution of a bivariate claim amount distribution	130
4.5.3 Reinsurer's risk profile in presence of two correlated insurers . .	132
4.6 Conclusion	136
4.7 Références	137

Abstract

A numerical method to compute bivariate probability distributions from their Laplace transform is presented. The method consists in an orthogonal projection of the probability density function with respect to a probability measure that belongs to Natural Exponential Family with Quadratic Variance Function (NEF-QVF). The procedure allows a quick and accurate calculation of probabilities of interest and does not require strong coding skills. Numerical illustrations and comparisons with other methods are provided. This work is motivated by actuarial applications. We aim at recovering the joint distribution of two aggregate claims amount associated with two insurance policy portfolios that are closely related, and at computing survival functions for reinsurance losses in presence of two non-proportional reinsurance treaties.

Keywords : Bivariate aggregate claims model, bivariate distribution, bivariate Laplace transform, numerical inversion of Laplace transform, natural exponential families with quadratic variance functions, orthogonal polynomials.

4.1 Introduction

In insurance and in reinsurance, some common events may cause simultaneous, correlated claims in two lines of business. For example, in third-party liability motor insurance, an accident may cause corporal damage and material damage losses. These parts of the same claim are then handled by different claim managers and reserving is most often also done separately, which makes it necessary to study the joint distribution and not only the marginals or the distribution of the sum (except when one loss dominates the other one, which makes correlation of secondary importance). Similarly, a reinsurer who accepts two stop-loss treaties from two customers operating on the same market will be exposed to the sum of two excesses of aggregate losses, which contain in general some independent part and some common shock part, arising from events that generate claims for the two insurance companies at the same time. To compute the Value-at-Risk of the reinsurer's exposure, or to compute the probability that the reinsurer loses more than certain risk limits, we need the bivariate distribution of the aggregate claims amount of the two insurers. In addition to classical couples of independent aggregated losses drawn from compound distributions, one needs to study the common shock part of the bivariate loss vector. We define a bivariate collective model with common shock by

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ U_{2j} \end{pmatrix} + \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad (4.1)$$

where (S_1, S_2) denote the aggregate claims amount of two non-life insurance portfolios over a given time period. The first part of the right-hand side of (4.1) is the common shock part. A counting variable N models the number of claims, and $\{(U_{1j}, U_{2j})\}_{j \in \mathbb{N}^*}$ is a sequence of **i.i.d** couples of non-negative random variables that model severities incurred in the two portfolios. We assume that the sequence $\{(U_{1j}, U_{2j})\}_{j \in \mathbb{N}^*}$ is independent from N . We are interested in bivariate distributions that allow dependence between U_{1j} and U_{2j} . The correlation between the claim amounts are related

to the magnitude of the event causing the claims. The second part of the right-hand side of (4.1) represents specific additional claim amounts that each insurer has to pay. The random vector (X_1, X_2) is independent of the first part. The components are non-negative and mutually independent random variables. This modelization differs from what already exists in the literature. Correlation is usually specified upon the number of claims in two lines of business, see [AMBAGASPITIYA \[1998, 1999\]](#). Recursion based methods have been designed to cope with those kind of bivariate distributions, see [SUNDT \[1999\]](#); [VERNIC \[1999\]](#). Aggregate claims are governed by compound distributions. These distributions, in the univariate case, are already difficult to study because closed formulae for the probability density function are available in only a few cases. The cumbersome computation time induced by Monte Carlo simulations to obtain accurate values of probabilities motivates the use of numerical techniques.

In this paper, we develop an effective method to compute Bivariate Probability Density Functions (BPDFs) from the knowledge of their bivariate Laplace transform or equivalently from the moments of the distribution. First, we choose a reference distribution that belongs to a NEF-QVF. Then, we express the BPDF with respect to this reference measure as an expansion in terms of orthonormal polynomials. The polynomials are orthogonal with respect to the reference probability measure. When we aim at recovering a BPDF on the positive quadrant of the plane, as it is the case within the frame of aggregate claims amount distribution, our method looks like a well established method called the Laguerre method. The similarity is due to the fact that we choose a Gamma distribution as reference distribution associated with Laguerre polynomials to ensure the validity of the expansion. We refer to [ABATE et collab. \[1995\]](#) where an effective variant of the Laguerre method is presented. The same authors also work out a multivariate extension, see [ABATE et collab. \[1998\]](#). The main contribution of our work lies in the possibility of choosing the parameters of the reference distribution to enhance the performance of the method. A great body of the applied probability literature is dedicated to the numerical inversion of the Laplace transform of univariate PDF but to the best of our knowledge only a few attempts have been made in a multivariate context. A Fourier series based method is presented in [CHOUDHURY et collab. \[1994\]](#) motivated by applications in queueing theory. The desired function is recovered from the Fourier series coefficients of a periodic version of the function constructed by aliasing. The inversion formula involves integral computations that are completed using a trapezoidal rule justified by Poisson summation formula. It also implies truncation but not a simple one : Euler summation formula for alternating series is employed consequently to the neat replacement of the original infinite serie by an equivalent alternating one. Interesting applications are allowed because of the possibility of recovering MPDFs of couples of random variables formed by a discrete and a continuous one. We want to mention the work of Jin and Ren [JIN et REN \[2013\]](#). A model very similar to the one defined in 4.1 is considered and a comparison between Fourier series and recursion based method is performed. Recently, a moment-recovered approximation has been proposed in [MNATSAKOV \[2011\]](#). The stable approximants for the BPDF and the joint survival function is derived. A very interesting feature of this method is that the approximation can turned into a statistical estimation as there exist empirical counterparts of the aforementioned approximants. The expansion we propose is also well adapted to empirical estimation. The empirical estimation of BPDF when data are available will be at

the center of a forthcoming paper. Finally, one key aspect of the method is that once the parameters has been choosen and the coefficients computed, approximations for the joint probability density and survival functions are available in an explicit form. In Section 2, we introduce a BPDF expansion formula based on orthogonal projection within the frame of NEF-QVF. In Section 3, the expansion for bivariate aggregate claims amount distributions is derived along with conditions to ensure the validity of the polynomial expansion. Section 4 is devoted to numerical illustrations and examples of application. A comparison to the moment-recovered method is also proposed.

4.2 Expression of the joint density

4.2.1 Orthogonal polynomials associated with NEF-QVF

Let $F = \{F_\theta, \theta \in \Theta\}$ with $\Theta \subset \mathbb{R}$ be a Natural Exponential Family (NEF), see [BARNDORFF-NIELSEN \[1978\]](#), generated by a probability measure ν on \mathbb{R} such that

$$\begin{aligned} F_\theta(X \in A) &= \int_A \exp\{x\theta - \kappa(\theta)\} d\nu(x) \\ &= \int_A f(x, \theta) d\nu(x), \end{aligned}$$

where $A \subset \mathbb{R}$, $\kappa(\theta) = \log(\int_{\mathbb{R}} e^{\theta x} d\nu(x))$ is the Cumulant Generating Function (CGF), and $f(x, \theta)$ is the density of P_θ with respect to ν . Let X be a random variable F_θ distributed. The mean of X is

$$m = E_\theta(X) = \int x dF_\theta(x) = \kappa'(\theta),$$

and its variance is

$$V(m) = Var_\theta(X) = \int (x - m)^2 dF_\theta(x) = \kappa''(\theta).$$

The application $\theta \rightarrow \kappa'(\theta)$ is one to one. Its inverse function $m \rightarrow h(m)$ is defined on $\mathcal{M} = \kappa'(\Theta)$. With a slight change of notation, we can rewrite $F = \{F_m, m \in \mathcal{M}\}$, where F_m has mean m and density $f(x, m) = \exp\{h(m)x - \kappa(h(m))\}$ with respect to ν . A NEF has a Quadratic Variance Function (QVF) if there exist reals ν_0, ν_1, ν_2 such that

$$V(m) = \nu_0 + \nu_1 m + \nu_2 m^2. \quad (4.2)$$

The Natural Exponential Families with Quadratic Variance Function (NEF-QVF) include the normal, gamma, hyperbolic, Poisson, binomial and negative binomial distributions.

Define

$$P_k(x, m) = V^n(m) \left\{ \frac{\partial^n}{\partial m^n} f(x, m) \right\} / f(x, m), \quad (4.3)$$

for $n \in \mathbb{N}$. Each $P_n(x, m)$ is a polynomial of degree n in both m and x . Moreover, if F is NEF-QVF, $\{P_n\}_{n \in \mathbb{N}}$ is a family of orthogonal polynomials with respect to P_m in the sense that

$$\langle P_k, P_l \rangle = \int P_k(x, m) P_l(x, m) dF_m(x) = \delta_{kl} \|P_k\|^2, \quad k, l \in \mathbb{N},$$

where δ_{kl} is the Kronecker symbol equal to 1 if $k = l$ and 0 otherwise. For the sake of simplicity, we choose $\mathbf{v} = F_{m_0}$. Then, $f(x, m_0) = 1$ and we write

$$P_k(x) = P_k(x, m_0) = V^k(m_0) \left\{ \frac{\partial^k}{\partial m^k} f(x, m) \right\}_{m=m_0}. \quad (4.4)$$

We also consider in the rest of the paper a normalized version of the polynomials defined in (4.4) with $Q_k(x) = \frac{P_k(x)}{\|P_k\|}$. For an exhaustive review regarding NEF-QVF and their properties, we refer to MORRIS [1982].

4.2.2 Bivariate probability measures and their Laplace transform

Let (X, Y) be a couple of random variables, governed by a bivariate probability measure $F_{X,Y}$. We also assume that (X, Y) admits a joint probability density function with respect to a bivariate measure denoted by λ , such that

$$dF_{X,Y}(x, y) = f_{X,Y}(x, y) d\lambda(x, y). \quad (4.5)$$

Its bivariate Laplace transform is defined as

$$\widehat{f_{X,Y}}(s, t) = \int \int e^{sx+ty} f_{X,Y}(x, y) d\lambda(x, y). \quad (4.6)$$

We consider the bivariate NEF-QVF generated by

$$d\mathbf{v}(x, y) = d\mathbf{v}_1(x) \times d\mathbf{v}_2(y), \quad (4.7)$$

where \mathbf{v}_1 and \mathbf{v}_2 are two probability measures that belong to NEF-QVF absolutely continuous with respect to λ . With the notation of Section 2.1, the probability density function associated with \mathbf{v} is

$$f(x, y) = f(x, m_1) f(y, m_2). \quad (4.8)$$

Let $Q_{k,l}(x, y) = Q_k^1(x) Q_l^2(y)$ be a bivariate polynomial system that is orthogonal with respect to \mathbf{v} . We have

$$\begin{aligned} \langle Q_{k,l}, Q_{i,j} \rangle &= \int \int Q_{k,l}(x, y) Q_{i,j}(x, y) d\mathbf{v}(x, y) \\ \langle Q_{k,l}, Q_{i,j} \rangle &= \delta_{ki} \delta_{lj}. \end{aligned}$$

We denote by $L_2(\mathbf{v})$ the set of functions square integrable with respect to \mathbf{v} .

Proposition 13. Assume that $\frac{dF_{X,Y}}{d\mathbf{v}} \in L_2(\mathbf{v})$, we have the following expansion

$$\frac{dF_{X,Y}}{d\mathbf{v}}(x, y) = \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} \left\langle \frac{dF_{X,Y}}{d\mathbf{v}}, Q_{k,l} \right\rangle Q_{k,l}(x, y) \quad (4.9)$$

$$= \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} E(Q_{k,l}(X, Y)) Q_{k,l}(x, y). \quad (4.10)$$

Applying Proposition 13 yields an expansion of the probability density function

$$f_{X,Y}(x, y) = \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} \int \int (f_{X,Y}(x, y) Q_{k,l}(x, y) dx dy) \times Q_{k,l}(x, y) f(x, y). \quad (4.11)$$

We have

$$Q_{k,l}(x, y) = \sum_{i=0}^k \sum_{j=0}^l q_{i,j}(k, l) x^i y^j, \quad (4.12)$$

where $q_{i,j}(k, l)$ is expressed in terms of the coefficients of Q_k^1 and Q_l^2 . Reinjecting the expression of $Q_{k,l}$ given in (4.12) in the probability density function (4.11) gives the two following representations of $f_{X,Y}$

$$\begin{aligned} f_{X,Y}(x, y) &= \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} \sum_{i=0}^k \sum_{j=0}^l q_{i,j}(k, l) \int \int x^i y^j f_{X,Y}(x, y) dx dy \times Q_{k,l}(x, y) f(x, y) \\ &= \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} \sum_{i=0}^k \sum_{j=0}^l q_{i,j}(k, l) \left[\frac{\partial^{i+j}}{\partial s^i \partial t^j} \widehat{f_{X,Y}}(s, t) \right]_{s=0, t=0} Q_{k,l}(x, y) f(x, y) \\ &\doteq \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} a_{k,l} Q_{k,l}(x, y) f(x, y). \end{aligned}$$

Remarque 4. *The probability density function takes the form of an expansion. The coefficients are expressed in terms of the moments and equivalently in terms of the derivative of the bivariate Laplace transform. The expression of the probability density function is therefore a Laplace transform inversion formula.*

Approximation of the bivariate probability density function is obtained through truncation of the infinite series

$$f_{X,Y}^{K,L}(x, y) = \sum_{k=0}^K \sum_{l=0}^L a_{k,l} Q_{k,l}(x, y) f(x, y),$$

where K and L denote the order of truncation of the approximation. In view of future applications, we are interested in getting approximations of cumulative distribution functions, which are obtained through integration of the approximated density

$$F_{X,Y}^{K,L}(x, y) = \int_{-\infty}^x \int_{-\infty}^y f_{X,Y}^{K,L}(u, v) d\lambda(u, v) \quad (4.13)$$

$$= \sum_{k=0}^K \sum_{l=0}^L a_{k,l} \int_{-\infty}^x \int_{-\infty}^y Q_{k,l}(u, v) f(u, v) d\lambda(u, v). \quad (4.14)$$

The goodness of the approximation lies in the decay of the sequence defined by the coefficients of the expansion. Note that from $\frac{dF_{X,Y}}{dv} \in L_2(v)$, a Parseval type relation holds

$$\int \int f_{X,Y}^2(x, y) dx dy = \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} a_{k,l}^2 < +\infty. \quad (4.15)$$

Relation (4.15) implies that the sequence $\{a_{k,l}\}_{k,l \in \mathbb{N}}$ decreases towards 0 with respect to k and l . The choice of the NEF-QVF and its parameters depends on the problem, and are discussed in the next section.

4.3 Application to a bivariate aggregate claims amount distribution

4.3.1 General formula

Aggregate claims amount are modeled by non-negative random variables. The only NEF-QVF with support on \mathbb{R}^+ is generated by the Gamma distribution that admits a probability density function with respect to the Lebesgue measure defined as

$$f(x) = \frac{x^{r-1} e^{-x/m}}{\Gamma(r) m^r}. \quad (4.16)$$

The orthogonal polynomials associated with the Gamma distribution are the generalized Laguerre polynomials

$$Q_k(x) = \frac{(-1)^k L_k^{r-1}(x/m)}{\sqrt{\binom{k+r-1}{k}}}, \quad (4.17)$$

where $L_n^{r-1}(x) = \sum_{i=0}^n \binom{n+r-1}{n-i} \frac{(-x)^i}{i!}$, see SZEGÖ [1939]. The random couple (S_1, S_2) , defined in (4.1), is divided into two parts. We denote by

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ U_{2j} \end{pmatrix} \quad (4.18)$$

the first part which is governed by a bivariate compound distribution. The study of the distribution of (X_1, X_2) may be reduced to univariate problem that we can easily handle due to the independence of the components. The distribution of (Y_1, Y_2) is unknown and we apply polynomial expansion in order to recover its BPDF. The BPDF of (S_1, S_2) is obtained using a two-dimensional convolution procedure. The bivariate probability measure associated with (Y_1, Y_2) is divided into a discrete part with an atom at $(0, 0)$ with a probability mass equal to $p_0 = P(N = 0)$ and a continuous part that is absolutely continuous with respect to the bivariate Lebesgue measure

$$dF_{Y_1, Y_2}(x, y) = p_0 \delta_{\{x=0, y=0\}}(x, y) + dG_{Y_1, Y_2}(x, y), \quad (4.19)$$

where $\delta_{\{x=0, y=0\}}(x, y)$ is the Dirac measure equal to 1 at $(0, 0)$ and 0 otherwise. We aim at expanding the continuous part using Proposition 13. The defective density probability function of dG_{Y_1, Y_2} is denoted by g_{Y_1, Y_2} . We define a bivariate NEF-QVF probability measure as in Section 2 by

$$\nu(x, y) = \nu_1(x) \times \nu_2(y), \quad (4.20)$$

where ν_1 and ν_2 are Gamma distributions with parameter (m_1, r_1) and (m_2, r_2) respectively. Orthogonal polynomials with respect to ν are defined by

$$Q_{l,k}(x, y) = Q_k^1(x) \times Q_l^2(y), \quad (4.21)$$

where $\{Q_k^1\}_{k \in \mathbb{N}}$ and $\{Q_l^2\}_{l \in \mathbb{N}}$ are orthogonal polynomial systems with respect to ν_1 and ν_2 respectively, and defined as in (4.17). If $\frac{dG_{Y_1, Y_2}}{d\nu} \in L^2(\nu)$, applying Proposition 13 yields

$$g_{Y_1, Y_2}(x, y) = \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} a_{k,l} Q_{k,l}(x, y) f_\nu(x, y), \quad (4.22)$$

where f is the BPDF associated to the probability measure ν .

Remarque 5. The Laguerre series method in [ABATE et collab. \[1998\]](#) is related to our approach with a particular parametrization that is $m_1 = m_2 = 2$ and $r_1 = r_2 = 1$.

We know from the Parseval type relation (4.15) that the sequence of coefficients of the expansion is decreasing towards 0. But how fast does it actually decrease? We address this problem using generating function theory just like what is done in [ABATE et collab. \[1998\]](#). The interesting fact is that the generating function of the coefficients of the expansion is linked to the bivariate Laplace transform of the expanded probability density function. We start from the expression given in (4.22). By taking the bivariate Laplace transform we get

$$\begin{aligned} \widehat{g_{Y_1, Y_2}}(s_1, s_2) &= \int_0^{+\infty} \int_0^{+\infty} e^{s_1 x + s_2 y} g_{Y_1, Y_2}(x, y) dx dy \\ &= \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} a_{k,l} (-1)^{k+l} \sqrt{\binom{k+r_1-1}{k} \binom{l+r_2-1}{l}} \\ &\quad \times \left(\frac{1}{1-s_1 m_1} \right)^{r_1} \left(\frac{1}{1-s_2 m_2} \right)^{r_2} \left(\frac{s_1 m_1}{s_1 m_1 - 1} \right)^k \left(\frac{s_2 m_2}{s_2 m_2 - 1} \right)^l \\ &= \left(\frac{1}{1-s_1 m_1} \right)^{r_1} \left(\frac{1}{1-s_2 m_2} \right)^{r_2} B \left(\frac{s_1 m_1}{s_1 m_1 - 1}, \frac{s_2 m_2}{s_2 m_2 - 1} \right), \end{aligned} \quad (4.23)$$

where

$$B(z_1, z_2) = \sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} b_{k,l} z_1^k z_2^l \quad (4.24)$$

is the generating function of the two-dimensional sequence $\{b_{k,l}\}_{k,l \in \mathbb{N}}$. The coefficients of the expansion follow from

$$a_{k,l} = b_{k,l} \frac{(-1)^{k+l}}{\sqrt{\binom{k+r_1-1}{k} \binom{l+r_2-1}{l}}}. \quad (4.25)$$

The generating function B is derived from the Bivariate Laplace transform after a simple change of variable in Equation (4.23) :

$$B(z_1, z_2) = (1-z_1)^{-r_1} (1-z_2)^{-r_2} \widehat{g_{Y_1, Y_2}} \left(\frac{z_1}{m_1(1-z_1)}, \frac{z_2}{m_2(1-z_2)} \right). \quad (4.26)$$

In [ABATE et collab. \[1998\]](#), the authors compute the coefficients of the expansion through the derivative of their generating function. They use Cauchy contour integral to express the derivatives as integrals and approximate them via a trapezoidal rule. In some cases, when the Laplace transform expression is simple, the derivatives can be calculated directly with a computational software . The attractive feature of our method lies in the possibility of tuning the parameters of the Gamma distributions in order to modify the generating function. We illustrate this fact through the following example :

Exemple 3. Let (X, Y) be a couple of independent random variables exponentially distributed with parameters δ_1 and δ_2 . The bivariate Laplace transform associated with this couple of random variables is

$$\widehat{f_{X, Y}}(s_1, s_2) = \frac{\delta_1}{\delta_1 - s_1} \times \frac{\delta_2}{\delta_2 - s_2}. \quad (4.27)$$

We inject the bivariate Laplace transform expression (4.27) in the generating function defined in (4.26) to get

$$B(z_1, z_2) = \frac{(1 - z_1)(1 - z_2)}{(1 - z_1)^{r_1}(1 - z_2)^{r_2}} \times \frac{\delta_1 \delta_2}{[\delta_1 + z_1(1/m_1 - \delta_1)] [\delta_2 + z_2(1/m_2 - \delta_2)]}. \quad (4.28)$$

In this case, the parametrization is straightforward. We set $r_1 = r_2 = 1$, $m_1 = 1/\delta_1$ and $m_2 = 1/\delta_2$, so $B(z_1, z_2) = 1$ which implies that $a_{0,0} = 1$ and $a_{k,l} = 0$ for all $k, l \geq 1$.

The Laplace transform of the distributions that we consider in the rest of the paper is more complicated and it is less straightforward to choose the parameters. In the next Section, we show how to choose the parameters of the expansion in order to ensure the integrability condition on Proposition 13.

4.3.2 Integrability condition

The polynomial expansion is valid under the following condition :

$$\int_0^{+\infty} \int_0^{+\infty} \left(\frac{dG_{Y_1, Y_2}}{dv}(x, y) \right)^2 dv(x, y) < +\infty, \quad (4.29)$$

which is equivalent to

$$\int_0^{+\infty} \int_0^{+\infty} g_{Y_1, Y_2}(x, y)^2 x^{1-r_1} y^{1-r_2} e^{x/m_1 + y/m_2} dx dy < +\infty. \quad (4.30)$$

The parameters m_1 , m_2 , r_1 and r_2 of the reference distribution have a key role regarding the validity of the polynomial expansion. We define the set

$$E = \{(s, t) \in [0, +\infty) \times [0, +\infty); \widehat{f}(s, t) < +\infty\}, \quad (4.31)$$

and denote by E^c its complementary. The following result gives us insight on how to choose the parameters of the reference distribution.

Theorème 11. *Let (X, Y) be a couple of non-negative random variables with a joint PDF $f(x, y)$. We assume that $E \neq \{(0, 0)\}$. We assume additionally that there exist two real numbers $a, b \geq 0$ such that for all $(x, y) \in [a, +\infty) \times [b, +\infty)$, the applications $x \rightarrow f(x, y)$ and $y \rightarrow f(x, y)$ are strictly decreasing. Then we have*

$$f(x, y) \leq A(s, t) e^{-sx - ty}, \quad \forall (x, y) \in [a, +\infty) \times [b, +\infty), \quad (4.32)$$

where $A(s, t)$ is some real number independent of x and y .

Démonstration. Let $(x, y) \in [a, +\infty) \times [b, +\infty)$, we have

$$\begin{aligned}
 \widehat{f}(s, t) &> \int_a^x \int_b^y e^{su+tv} f(u, v) dudv \\
 &= \int_a^x e^{su} \frac{1}{t} (e^{tv} f_{X,Y}(u, y) - e^{tb} f(u, b)) dudv \\
 &\quad - \int_a^x e^{su} \int_b^y \frac{e^{tv}}{t} \frac{\partial f(u, v)}{\partial v} dudv \\
 &> \int_a^x e^{su} \frac{1}{t} (e^{tv} f(u, y) - e^{tb} f(u, b)) dudv \\
 &> \frac{e^{sx+ty}}{st} f(x, y) + \frac{e^{sa+tb}}{st} f(a, b) \\
 &\quad - \frac{1}{st} \left(e^{sa+ty} f(a, y) + e^{sx+tb} f(x, b) \right). \tag{4.33}
 \end{aligned}$$

In order to deal with the part (4.33), we need the following lemma.

Lemme 2. *Let f be a continuously differentiable function on \mathbb{R}^+ . Assume that there exist some $a > 0$ such that $x \rightarrow f(x)$ is strictly decreasing for $x > a$. Assume additionally that there exist $s > 0$ such that $\widehat{f}(s) < +\infty$. Then we have*

$$f(x) \leq A(s) e^{-sx} \quad \forall (x, y) \in [a, +\infty), \tag{4.34}$$

where $A(s)$ is some real number independent of x .

Démonstration. For $x \geq a$, we have

$$\begin{aligned}
 \widehat{f}(s) &> \int_a^x e^{sy} f(y) dy \\
 &= \left[\frac{e^{sy}}{s} f(y) \right]_a^x - \frac{1}{s} \int_a^x e^{sy} f'(y) dy. \tag{4.35}
 \end{aligned}$$

$$> \frac{1}{s} (f(x) e^{sx} - f(a) e^{sa}). \tag{4.36}$$

Thus, we deduce that $\forall x \geq a$,

$$f(x) < A(s) e^{-sx},$$

where $A(s) = (s\widehat{f}(s) + f(a) e^{sa})$. □

As $\widehat{f}(s, t) < +\infty$, we have

$$\widehat{f}(0, t) = \int_0^{+\infty} \int_0^{+\infty} e^{ty} f(x, y) < +\infty. \tag{4.37}$$

As $y \rightarrow e^{ty} f(x, y)$ is a non-negative function, applying Fubini-Tonelli's theorem yields $\int_0^{+\infty} e^{ty} f(x, y) < +\infty$ for all $x \in [0, +\infty)$, and for $x = a$ in particular. We can therefore apply Lemma 2 to the application $y \rightarrow f(a, y)$, which yields

$$f(a, y) < A_1(t) e^{-ty}. \tag{4.38}$$

Using the same arguments, we get that

$$f(x, b) < A_2(t) e^{-sx}. \tag{4.39}$$

Reinjecting inequalities (4.38) and (4.39) in (4.33) yields

$$f(x, y) < (st\widehat{f}(-s, -t) - e^{sa+tb}f(a, b) + A_2(t)e^{tb} + A_1(s)e^{sa})e^{-sx-ty}. \quad (4.40)$$

□

Theorem 11 allows us to bound continuous joint PDF when the bivariate Laplace transform is well defined for some positive arguments. We aim at recovering the distribution of (Y_1, Y_2) that admits a singular and a continuous part. We put aside the singular part and expand the continuous part represented through its defective bivariate defective probability density function g_{Y_1, Y_2} . Theorem 11 holds for g_{Y_1, Y_2} as for any common bivariate probability density function that satisfies the hypothesis of Theorem 11. We define $E^* = \overline{E} \cap \overline{E}^c$, where \overline{E} is the closure of E . One can note that for any $(s, t) \in [0, +\infty) \times [0, +\infty)$ if there exist (s^*, t^*) such that $s < s^*$ and $t < t^*$ then $(s, t) \in E$. The next result follows from that last remark.

Corollaire 5. *We consider a couple of random variables (Y_1, Y_2) defined as in (4.18). Assume that g_{Y_1, Y_2} satisfies the hypothesis of Theorem 11. For all $(s_1^*, s_2^*) \in E^*$, if we take $\{1/m_1 < 2s_1^*, 1/m_2 < 2s_2^*, r_1 = 1, r_2 = 1\}$, then the integrability condition (4.30) is satisfied.*

This corollary will help us greatly in next section in which the performances of the polynomial approximations are studied.

4.4 Downton Bivariate Exponential distribution and Lancaster probabilities

A classical problem, which has become known as the *Lancaster problem*, see LANCASTER [1958], is to characterize bivariate distributions with given margins and orthogonal eigenfunctions. Let ν_1 and ν_2 be two probability measures with associated orthonormal polynomials sequences $\{Q_n^1\}_{n \in \mathbb{N}}$ and $\{Q_n^2\}_{n \in \mathbb{N}}$ respectively. We denote by f_{ν_1} and f_{ν_2} the PDF of ν_1 and ν_2 . Exchangeable Lancaster bivariate densities have the form

$$f(x, y) = \left(\sum_{k=0}^{+\infty} a_k Q_k^1(x) Q_k^2(y) \right) f_{\nu_1}(x) f_{\nu_2}(y), \quad (4.41)$$

where $\{a_k\}_{k \in \mathbb{N}}$ is a sequence of real numbers with $a_0 = 1$. The Lancaster problem can be reduced to finding a proper sequence $\{a_k\}_{k \in \mathbb{N}}$ such that f is nonnegative and therefore a MPDF. The characterization of the Lancaster probabilities constructed via NEF-QVF has been widely studied by Koudou, see KOUDOU [1995, 1996]. He gave existence conditions and explicit forms of Lancaster sequences in various cases. In a recent publication, Diaconis et al. DIACONIS et GRIFFITHS [2012] studied the binomial-binomial case which means that the marginal distributions are both binomial. Characterizations with a more "probabilistic flavor" than the classical one are given. Recall that the classical way to construct a Lancaster probability consists in considering the distribution of $(X, Y) = (U + W, V + W)$, where U, V, W , are independent random variables governed by the same distribution. This "random element in common" way allows to construct

Lancaster probabilities based on any NEF-QVF distribution. To connect Lancaster probabilities with our approach, we start by considering our representation of the MPDF for a couple of nonnegative random variables (X, Y)

$$f_{(X,Y)}(x, y) = \left(\sum_{k=0}^{+\infty} \sum_{l=0}^{+\infty} E(Q_k^1(X)Q_l^2(Y)) Q_k^1(x)Q_l^2(y) \right) f_{v_1}(x) f_{v_2}(y), \quad (4.42)$$

where f_{v_i} is a Gamma PDF and $\{Q_k^i\}_{k \in \mathbb{N}}$ its orthonormal polynomials sequence. In order to get a Lancaster probability from the MPDF defined in (4.42), we need a bi-orthogonality condition, namely

$$E(Q_k^1(X)Q_l^2(Y)) = \delta_{kl} a_k. \quad (4.43)$$

Let (X, Y) be a couple of random variables governed by a Downton BiVariate Exponential distribution DBVE (μ_1, μ_2, ρ) introduced in DOWTON [1970]. We apply our method to recover the joint PDF of this distribution. The bivariate Laplace transform associated with the DBVE distribution is

$$\widehat{f_{X,Y}}(s_1, s_2) = \frac{\mu_1 \mu_2}{[(\mu_1 - s_1)(\mu_2 - s_2) - \rho s_1 s_2]}, \quad (4.44)$$

where $\mu_1, \mu_2 \geq 0$ and $0 \leq \rho \leq 1$. We inject the bivariate Laplace transform expression (4.52) in the generating function defined in (4.26) to get

$$B(z_1, z_2) = \frac{(1 - z_1)^{1-r_1} (1 - z_2)^{1-r_2} (\mu_1 \mu_2)}{\left([z_1(\mu_1 - 1/m_1) - \mu_1] [z_2(\mu_2 - 1/m_2) - \mu_2] - \frac{z_1 z_2 \rho}{m_1 m_2} \right)}. \quad (4.45)$$

We set $r_1 = r_2 = 1$, $m_1 = \frac{1}{\mu_1(1-\rho)}$ and $m_2 = \frac{1}{\mu_2(1-\rho)}$, so $B(z_1, z_2) = \left(\frac{1}{1 - z_1 z_2 \rho} \right)$ which implies that

$$a_k \doteq a_{k,l} = \rho^k \delta_{kl}. \quad (4.46)$$

The double sum turns into a simple one, leaving us with the same form as in (4.41) and yielding the following result :

Proposition 14. *The Downton Bivariate Exponential distribution belongs to Lancaster probabilities and its PDF admits an infinite series representation of the form*

$$f(x, y) = \left(\sum_{k=0}^{+\infty} a_k Q_k^1(x) Q_k^2(y) \right) f_{v_1}(x) f_{v_2}(y), \quad (4.47)$$

where $a_k = \rho^k \delta_{kl}$, f_{v_1} is the PDF of a $\Gamma\left(\frac{1}{\mu_1(1-\rho)}, 1\right)$ distribution and f_{v_2} is the PDF of a $\Gamma\left(\frac{1}{\mu_2(1-\rho)}, 1\right)$ distribution.

Historically, the DBVE distribution has been used to capture the joint lifetimes of two components for reliability purposes. These two components are assumed to collapse after a random numbers of shocks with exponential inter-arrival times. We therefore have

$$\begin{pmatrix} X \\ Y \end{pmatrix} = \sum_{j=1}^{N+1} \begin{pmatrix} U_{1j} \\ U_{2j} \end{pmatrix}, \quad (4.48)$$

where N is a counting random variable governed by a geometric distribution with parameters ρ . The $\{U_{ij}\}_{i \in \{1,2\}, j \in \mathbb{N}}$ are mutually independent and exponentially distributed with parameter μ_i and independent of N . We shed light here on a new way to construct Gamma-Gamma Lancaster probabilities through geometric compounding exponential random vectors.

4.5 Numerical illustrations

In this section, we aim at recovering the joint distribution of (S_1, S_2) defined in (4.1). As the joint probability density and survival functions are not available in a closed form, we assess the accuracy of the polynomial approximation on well known bivariate distributions. Then, we propose an example of practical applications by performing computations of quantities of interest for risk management purposes. Note that we consider a case where no closed formula of the BPDF is available.

4.5.1 Approximation of common bivariate exponential distributions

We define the approximation error as the relative difference between the exact function and its approximation. We propose a two-dimensional graphical visualization. The color of the surface indicates the level of error at each point located by its coordinates (x, y) . We compare the results of the polynomial approximation to the results of the moments-recovered method when approximation the joint PDF and the joint survival function of the considered distributions. The approximation of the joint PDF through the moment recovered method follows from

$$f_{X,Y}^{\alpha,\alpha'}(x, y) = \frac{e^{-x-y}\Gamma(\alpha+2)\Gamma(\alpha'+2)}{\Gamma(\lfloor \alpha e^{-x} \rfloor + 1)\Gamma(\lfloor \alpha' e^{-y} \rfloor + 1)} \times \sum_{k=0}^{\alpha - \lfloor \alpha e^{-x} \rfloor} \sum_{l=0}^{\alpha' - \lfloor \alpha' e^{-y} \rfloor} \frac{(-1)^{k+l} \widehat{f_{X,Y}}(\lfloor \alpha e^{-x} \rfloor + k, \lfloor \alpha' e^{-y} \rfloor + l)}{k!(\alpha - \lfloor \alpha e^{-x} \rfloor - k)l!(\alpha' - \lfloor \alpha' e^{-y} \rfloor - l)}. \quad (4.49)$$

This approximation tends towards the desired function when α and α' tend to infinity. The approximation of the joint survival function is given by

$$\overline{F_{X,Y}^{\alpha,\alpha'}}(x, y) = \sum_{k=0}^{\lfloor \alpha e^{-x} \rfloor} \sum_{l=0}^{\lfloor \alpha' e^{-y} \rfloor} \sum_{j=k}^{\alpha} \sum_{m=l}^{\alpha'} \binom{\alpha}{j} \binom{j}{k} \binom{\alpha'}{m} \binom{m}{l} (-1)^{j+m-k-l} \widehat{f_{X,Y}}(j, m). \quad (4.50)$$

The approximation formulas (4.49) and (4.50) are derived in [MNATSAKANOV \[2011\]](#).

Expansion of the Downton bivariate exponential distribution

The Downton BiVariate Exponential distribution - DBVE(μ_1, μ_2, ρ), introduced in [DOWTON \[1970\]](#), has been used to capture the joint lifetimes of two components for reliability purposes. These two components are assumed to collapse after a random numbers of shocks with exponential inter-arrival times. The joint PDF of a DBVE(μ_1, μ_2, ρ)-distributed random vector (X, Y) is given by

$$f_{X,Y}(x, y) = \frac{\mu_1 \mu_2}{1 - \rho} \exp\left(-\frac{\mu_1 x + \mu_2 y}{1 - \rho}\right) I_0 \left[\frac{1 - \sqrt{\rho \mu_1 \mu_2 x y}}{1 - \rho} \right], \quad x, y \in \mathbb{R}_+^2 \quad (4.51)$$

where I_0 denotes the modified Bessel, $\mu_1, \mu_2 \geq 0$ and $0 \leq \rho \leq 1$. The bivariate Laplace transform of (X, Y) is given by

$$\widehat{f_{X,Y}}(s_1, s_2) = \frac{\mu_1 \mu_2}{[(\mu_1 - s_1)(\mu_2 - s_2) - \rho s_1 s_2]}. \quad (4.52)$$

In this particular case, the study of the generating function of the coefficients is of interest to choose the parameters of the gamma measures. We inject the bivariate Laplace transform expression (4.52) in the generating function defined in (4.26) to get

$$B(z_1, z_2) = \frac{(1 - z_1)^{1-r_1} (1 - z_2)^{1-r_2} \mu_1 \mu_2}{[z_1(\mu_1 - 1/m_1 - \mu_1)] [z_2(\mu_2 - 1/m_2) - \mu_2] - \frac{z_1 z_2 \rho}{m_1 m_2}}. \quad (4.53)$$

We set $r_1 = r_2 = 1$, $m_1 = \frac{1}{\mu_1(1-\rho)}$ and $m_2 = \frac{1}{\mu_2(1-\rho)}$, so $B(z_1, z_2) = \left(\frac{1}{1-z_1 z_2 \rho}\right)$ which implies that

$$a_{k,l} = \rho^k \delta_{kl}, \quad (4.54)$$

along with a good accuracy. For the numerical illustration, we set $\{\mu_1 = 1/2, \mu_2 = 2, \rho = 1/4\}$. Figure 4.1 shows the approximation error associated to the polynomial approximation.

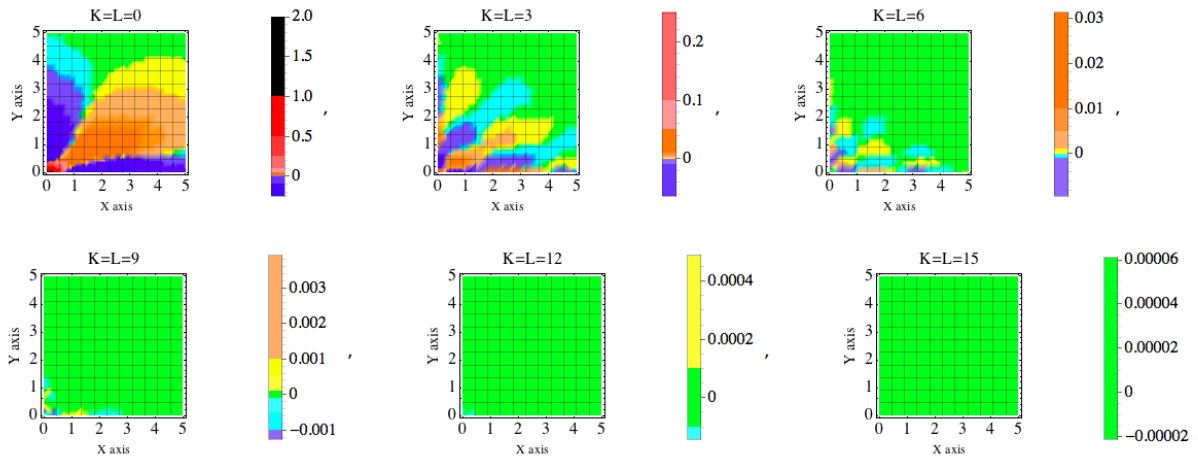


FIGURE 4.1 – Error map for the joint PDF of a DBVE(1/2, 2, 1/4) distribution approximated by polynomial expansion

An acceptable level of error is reached at almost every point of the bivariate probability density function with an order of truncation equal to 5. Figure 4.2 shows the approximation error of the moment-recovered based method. Increasing values of α and α' are considered.

The approximation of the moment-recovered based method works quite well too. However, the polynomial expansion seems to be better suited to the problem due to the geometrical decay of the coefficients of the expansion. In order to provide a fair comparison, we perform an approximation of the Marshall-Olkin bivariate exponential distribution in the next subsection.

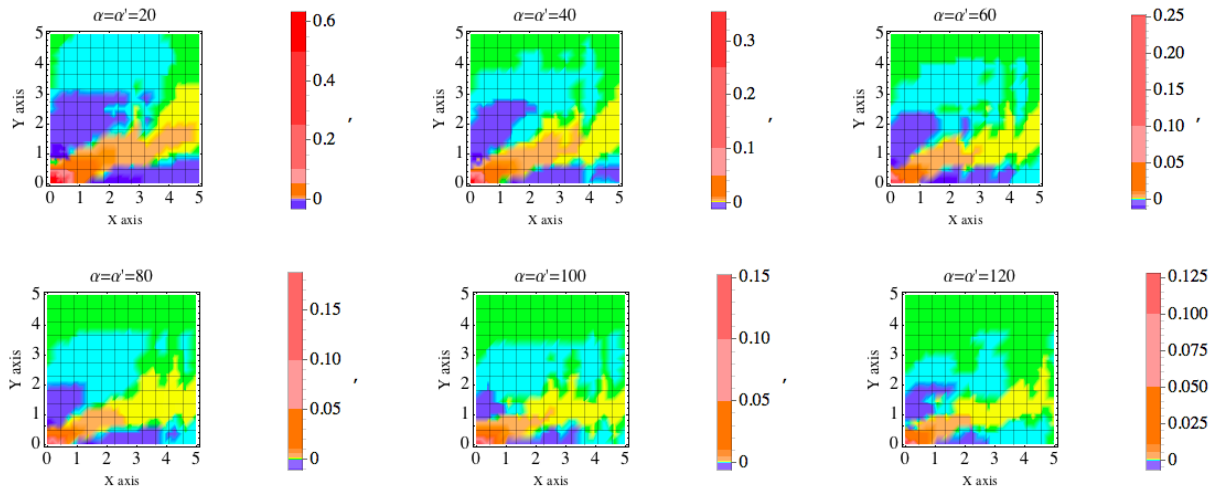


FIGURE 4.2 – Error map for the joint PDF of a DBVE(1/2, 2, 1/4) distribution approximated by the moment-recovered method

The Marshall-Olkin bivariate exponential distribution

The Marshall-Olkin bivariate exponential distribution - MOBVE($\lambda_1, \lambda_2, \lambda_{12}$) has been introduced in MARSHALL et OLKIN [1967] with reliability interpretations. It aims at modelling the time-to-failure of a two components system that receives shocks impacting one or both of the components. Each shock can cause the failure of none, one or both of the components. The definition implies the possibility that the two components fail at the same time, which means that the distribution admits a continuous and a singular part. This type of distribution might be pathological in a univariate context but arises naturally within a two-dimensional framework. The Marshall-Olkin BiVariate Exponential distribution MOBVE($\lambda_1, \lambda_2, \lambda_{12}$) has been originally characterized through its joint survival function

$$\overline{F_{X,Y}}(x, y) = \exp(-\lambda_1 x - \lambda_2 y - \lambda_{12} \max(x, y)). \quad (4.55)$$

The BPDF associated with this distribution is divided into the sum of an absolutely continuous part with respect to the bivariate Lebesgue measure and another part being singular, defined on the line $\{x = y\}$. The BPDF can be written as

$$f_{X,Y}(x, y) = \lambda_2(\lambda_1 + \lambda_{12})\overline{F_{X,Y}}(x, y)\mathbf{I}_{x>y}(x, y) \quad (4.56)$$

$$+ \lambda_1(\lambda_2 + \lambda_{12})\overline{F_{X,Y}}(x, y)\mathbf{I}_{x<y}(x, y) \quad (4.57)$$

$$+ \lambda_{12}e^{-\lambda \max(x,y)}\mathbf{I}_{x=y}(x, y), \quad (4.58)$$

where $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$. The Laplace transform associated to the MOBVE distribution is

$$\widehat{f_{X,Y}}(s_1, s_2) = \frac{(\lambda - s_1 - s_2)(\lambda_1 + \lambda_{12})(\lambda_2 + \lambda_{12}) + s_1 s_2 \lambda_{12}}{(\lambda - s_1 - s_2)(\lambda_1 + \lambda_{12} - s_1)(\lambda_2 + \lambda_{12} - s_2)}. \quad (4.59)$$

The distribution is not absolutely continuous with respect to the bivariate Lebesgue measure, the hypotheses of Proposition 13 are not satisfied. However, we manage to isolate the singular part. So we can expand the defective probability density function

associated with the continuous part of the distribution and add the singular part to the polynomial expansion in order to recover the entire distribution. The continuous part has a joint defective PDF $f_{X,Y}^C$, which is the sum of (4.56) and (4.57), and admits a Laplace transform of the form

$$\widehat{f_{X,Y}^C}(s_1, s_2) = \frac{s_2(\lambda_1 + s_1)}{(\lambda_1 + \lambda_2 + s_1)(\lambda_1 + \lambda_{12} + \lambda_2 + s_1 + s_2)} \quad (4.60)$$

$$+ \frac{s_1(\lambda_1 + s_2)}{(\lambda_1 + \lambda_{12} + s_2)(\lambda_1 + \lambda_{12} + \lambda_2 + s_1 + s_2)}. \quad (4.61)$$

The generating function of the coefficients of the expansion is tedious, making it difficult to choose relevantly the parameters of the expansion. Unlike in the case of the DBVE distribution, we cannot ensure that the chosen parametrization is the best. Approximations with different parametrization are compared in terms of accuracy.

- **Parametrization 1** : $\{m_1 = \frac{1}{\lambda_1 + \lambda_{12}}, m_2 = \frac{1}{\lambda_2 + \lambda_{12}}, \nu_1 = 1, \nu_2 = 1\}$
- **Parametrization 2** : $\{m_1 = \frac{1}{\lambda_1}, m_2 = \frac{1}{\lambda_2}, \nu_1 = 1, \nu_2 = 1\}$
- **Parametrization 3** : $\{m_1 = \frac{1}{\lambda}, m_2 = \frac{1}{\lambda}, \nu_1 = 1, \nu_2 = 1\}$

We set $\{\lambda_1 = 1/2, \lambda_2 = 2, \lambda_{12} = 1\}$. In view of a future probabilistic application, the error on the joint survival function is very interesting. Let us see how the polynomial method performs when it comes to approximating the joint survival function. We also apply the moment recovered method with the inversion formula dedicated to survival function approximation. On Figure 4.3, the relative difference between the exact value of the joint survival function and its polynomial approximation is plotted for different parametrizations.

The first parametrization seems to be the most suited to the problem as the convergence towards the exact value is quicker. On Figure 4.4, the difference between the exact value of the joint survival function and its moments approximation is plotted for increasing values of α and α' .

The results are satisfying for both methods, even if it looks like the moment recovered method does not perform well for $x, y \in [0, 0.5]$. One can note that the polynomial expansion with the first parametrization performs better than the moment-recovered based method. However, the moment-recovered based method has been enhanced in recent papers with the introduction of the scaled Laplace transform, see [MNATSAKANOV et SARKISIAN \[2013\]](#); [MNATSAKANOV et collab. \[2014\]](#). These improvements have been made in the univariate case but might be extended soon to the multivariate case and comparison would be interesting in the future.

4.5.2 Approximation of the distribution of a bivariate claim amount distribution

The main concern of this paper is to deal with the distribution of

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ U_{2j} \end{pmatrix}. \quad (4.62)$$

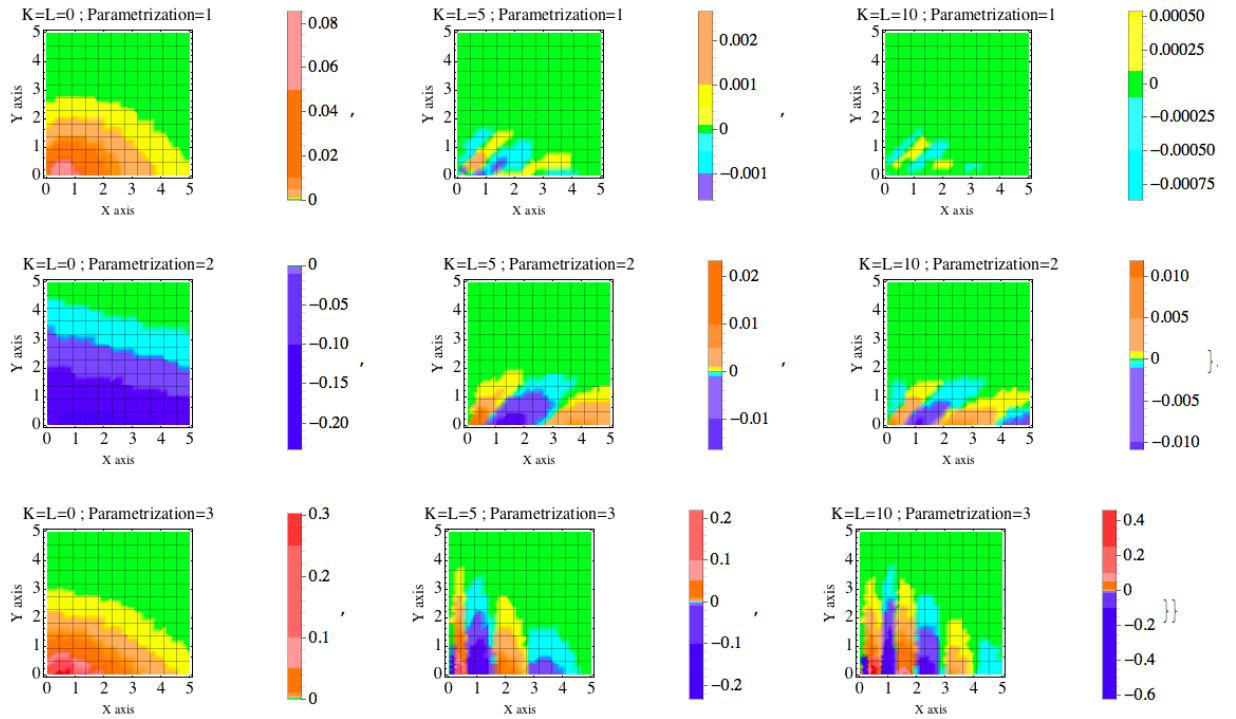


FIGURE 4.3 – Error map for the joint survival function of a MOBVE(1/2, 2, 1) distribution approximated by polynomial expansion.

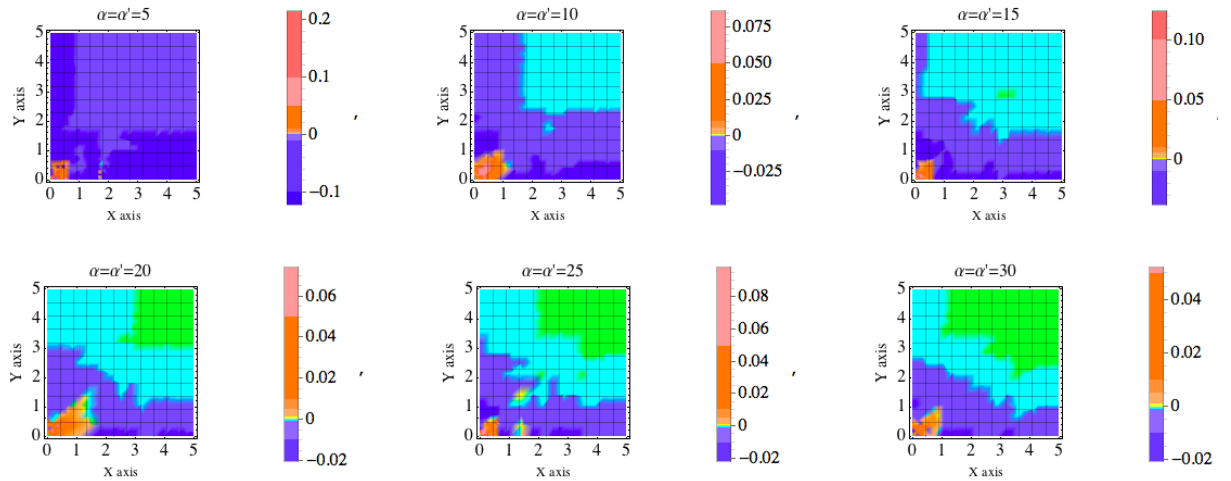


FIGURE 4.4 – Error map for the joint survival function of a MOBVE(1/2, 2, 1) distribution approximated by the moment recovered method.

We assume here that the sequence of random vectors (U_{1i}, U_{2i}) is **i.i.d.** and DBVE(μ_1, μ_2, ρ)-distributed. The number of claims is governed by a negative binomial distribution NEG-BIN(a, p). In this case, the corresponding bivariate Laplace transform

$$\widehat{g_{Y_1, Y_2}}(s_1, s_2) = \left(\frac{1-p}{1-p\widehat{f_{U_1, U_2}}(s_1, s_2)} \right)^a - (1-p)^a, \quad (4.63)$$

exists if $\widehat{f_{U_1, U_2}}(s_1, s_2) < p^{-1}$. In view of Corollary 5, the parameters must be chosen under the constraints

$$\{m_1 < 2s_1^*, m_2 < 2s_2^*, r_1 = r_2 = 1\},$$

where the couple (s_1^*, s_2^*) satisfies the equation $\widehat{f_{U_1, U_2}}(s_1^*, s_2^*) = p^{-1}$. We set $\{\mu_1 = 1, \mu_2 = 1, \rho = 1/4\}$ as parameters of the DBVE distribution and $\{a = 1, p = 3/4\}$ as parameters of the negative binomial distribution. In this particular case, if we set

$$\left\{ m_1 = \frac{1}{\mu_1(1-p)}, m_2 = \frac{1}{\mu_2(1-p)}, r_1 = 1, r_2 = 1 \right\}$$

as parameters of the reference distribution, the generating function of the coefficients of the expansion defined in (4.26) takes the form

$$B(z_1, z_2) = \frac{1}{1 + z_1 z_2 (p^2 - \rho(1-p)^2 - p)}. \quad (4.64)$$

The coefficients of the expansion are therefore $a_{kl} = (p^2 - \rho(1-p)^2 - p)^k \delta_{kl}$. The geometric decay of the coefficients implies a fast convergence of the approximation towards the desired function. It is easily checked that

$$\widehat{f_{U_1, U_2}}\left(\frac{1}{2\mu_1(1-p)}, \frac{1}{2\mu_2(1-p)}\right) < p,$$

making the chosen parametrization valid. The exact value of the survival function is unknown. A benchmark value is derived through Monte Carlo simulations using algorithms described in [HE et collab. \[2013\]](#). In order to get a good proxy via Monte-Carlo techniques, we produce 10^5 realizations of the couple (Y_1, Y_2) . Figure 4.7 shows the difference between the values of the survival function derived by Monte-Carlo and by polynomial expansion.

One can see that we get close to the benchmark with an order of truncation equal to 5. Figure 4.6 provides a 3D visualization of the joint probability density and survival function of the distribution.

4.5.3 Reinsurer's risk profile in presence of two correlated insurers

We now consider a reinsurer with two customers. We assume that the reinsurer accepted a stop-loss reinsurance treaty with limit b_i in excess of priority c_i from insurer i , where $i = 1, 2$. This means that if S_i is the aggregated claims amount for insurer i , then the reinsurer must pay $Z = \min([S_1 - c_1]^+, b_1) + \min([S_2 - c_2]^+, b_2)$. For risk management or pricing purposes, the reinsurer is interested in computing $P(Z > z)$ for different risk limits. To achieve that, one needs the joint distribution of (S_1, S_2) and it is not enough to know the distribution of $(S_1 + S_2)$, which would often be easier to obtain.

As an illustration, we assume that each insurer has to pay some specific claim amount, which correspond to two independent components. We also take into account common shocks arising from events which cause claims for both insurers, like hail or storm

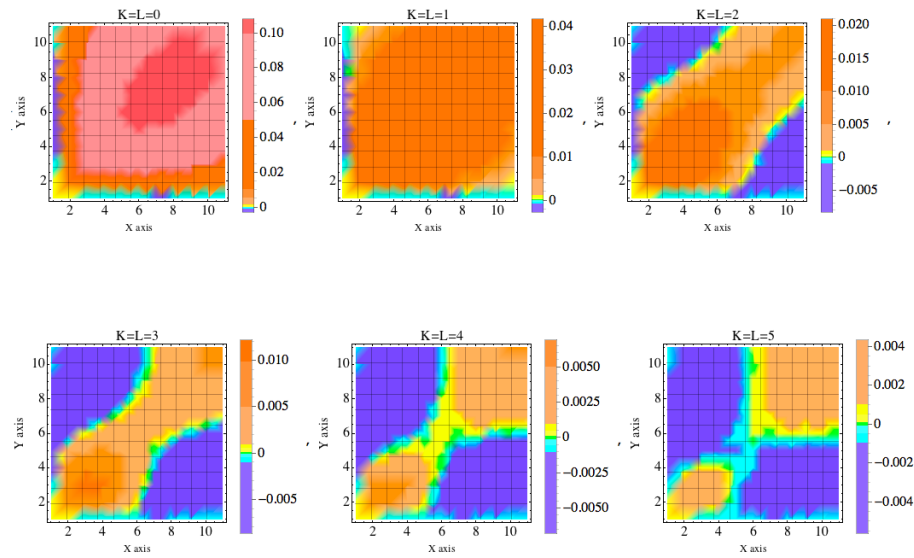


FIGURE 4.5 – Discrete error map for the joint survival function of compound NEG – BIN(1, 3/4) DBVE(1, 1, 1/4) distribution

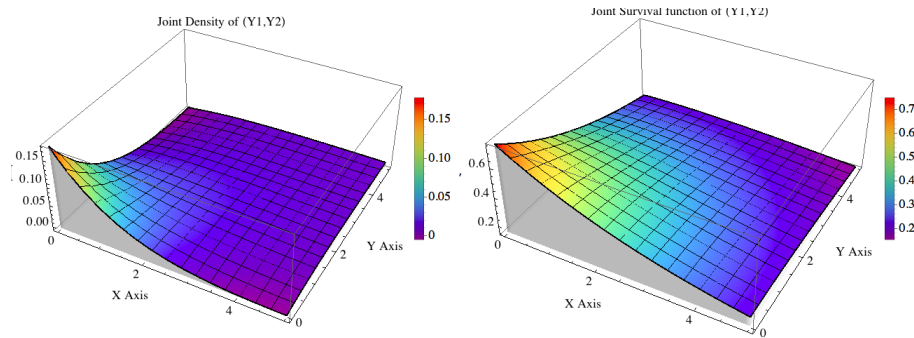


FIGURE 4.6 – Joint survival function of compound NEG – BIN(1, 3/4) DBVE(1, 1, 1/4) distribution with an order of truncation equal to 10

episodes : for example, in the case of a hail episode, it might happen that 1 windshield has to be replaced for insurer 1 and that 2 windshields have to be replaced for insurer 2. Of course, for one particular event (the k^{th} for instance) the two claims $U_{1,k}$ (incurred for insurer 1) and $U_{2,k}$ (incurred for insurer 2) are positively correlated, because they are influenced by the severity of the event causing the common shock.

We therefore consider that the vector (S_1, S_2) is described as

$$\begin{pmatrix} S_1 \\ S_2 \end{pmatrix} = \sum_{j=1}^N \begin{pmatrix} U_{1j} \\ U_{2j} \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^{N_1} V_i \\ \sum_{i=1}^{N_2} W_i \end{pmatrix}. \quad (4.65)$$

The first part of the right-hand side of 4.65 is exactly the same as the random vector described in the introduction. The number of claims that incurs in the two portfolios is a counting random variable denoted by N . The severities of those claims are modeled through a sequence of **i.i.d** random vectors $\{(U_{1,j}, U_{2,j})\}_{j \in \mathbb{N}}$ independent from

N. The second part of the right-hand side of 4.65 represents the claim amounts specific to each insurer. The numbers of claims are modeled by two independent counting random variables N_1 and N_2 . The sequences $\{V_i\}_{i \in \mathbb{N}}$ and $\{W_i\}_{i \in \mathbb{N}}$ are formed by **i.i.d** nonnegative random variables that represent the claim sizes. These sequences are mutually independent and independent from N_1 and N_2 . To recover the distribution of (S_1, S_2) using the polynomial technology, we use the method described in this paper to the first part of right-hand side of 4.65. Regarding the second part, we may apply a univariate polynomial expansion to the two components separately. We have already presented the polynomial expansion in a univariate context in a previous work, we refer the reader to ?. The joint probability density function of the second part is obtained by multiplying the marginal PDF due to the independence of the two components. The BPDF of (S_1, S_2) is derived by convoluting the polynomial approximations of the different parts. One must pay attention to the singularities that arise from the non-zero probability of having $\{N = 0\}$, $\{N_1 = 0\}$ and $\{N_2 = 0\}$.

For the common part, we assume that N is governed by a negative binomial distribution $\text{NEG} - \text{BIN}(a, p)$ and that (U_{1i}, U_{2i}) is $\text{DBVE}(\mu_1, \mu_2, \rho)$ -distributed. We set $\{a = 1, p = 3/4, \mu_1 = 1, \mu_2 = 1, \rho = 1/4\}$. For the specific part, we assume that N_1 and N_2 are governed by negative binomial distributions with parameters $\{a_1, p_1\}$ and $\{a_2, p_2\}$, the claim sizes are Gamma distributed with respective parameters $\{\alpha_1, \beta_1\}$ and $\{\alpha_2, \beta_2\}$. We set $\{a_1 = a_2 = 1, p_1 = p_2 = 3/4, \alpha_1 = \alpha_2 = 1, \beta_1 = \beta_2 = 1\}$. To expand the common part, we set $\left\{m_1 = \frac{1}{\mu_1(1-p)}, m_2 = \frac{1}{\mu_2(1-p)}, r_1 = 1, r_2 = 1\right\}$ as parameters of the polynomial expansion in view of the good results obtained in Subsection 4.5.2. The PDF of a geometric compound distribution with exponential claim sizes is available in a closed form, so we do not need to use univariate polynomial expansions. In order to ensure the goodness of our method, we need benchmark values. A sample of 10^5 Monte-Carlo simulations of (S_1, S_2) is produced. On Figure 4.7, the difference between the survival function obtained via Monte-Carlo and polynomial expansion is plotted.

The level of error is still acceptable. Figure 4.8 displays the joint probability density and survival functions.

As there exist a symmetry between the two portfolios, we assume that the same reinsurance treaty is proposed to the insurers with an equal priority and limit. We set $\{b_1 = b_2 = 4, c_1 = c_2 = 1\}$. Recall that the reinsurance cost is modeled by a random variable defined as

$$Z = \min([S_1 - c_1]^+, b_1) + \min([S_2 - c_2]^+, b_2). \quad (4.66)$$

Figure 4.1 displays two graphics. On the left one, we plot the survival function of Z delivered by the two methods. We can see that the two curves overlap. On the right one, we plot the difference between the two approximations to better appreciate their proximity. Some values of $P(Z > z)$ are given in Table 4.1.

The results are quite promising. The polynomial expansion seems to be well-suited to carry out numerical calculations. The operational advantage of numerical methods

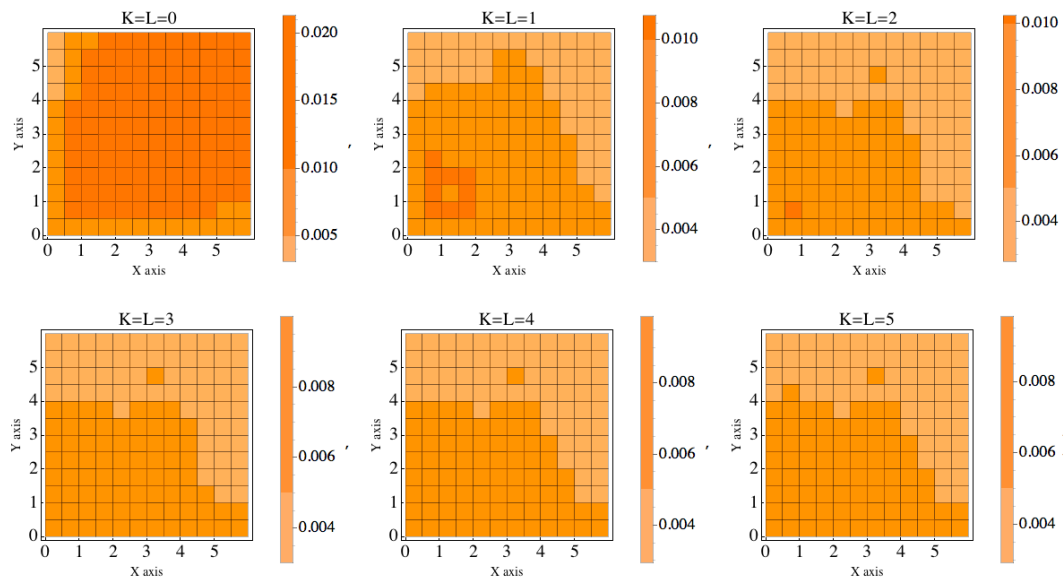


FIGURE 4.7 – Error map for the joint survival function of (S_1, S_2) obtained by polynomial expansion

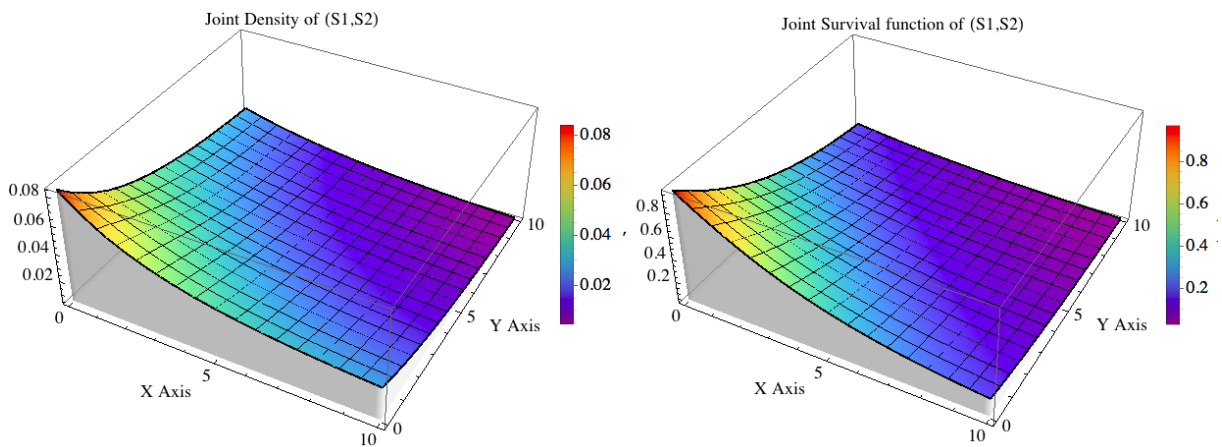


FIGURE 4.8 – Joint density and survival function of the aggregate claims in the two insurance portfolios

over Monte-Carlo techniques lies in the gain of computation time that they provide. It is difficult to quantify it as it varies given the software and the coding skills of the user. Note that, in the polynomial expansion case, the approximation of the PDF is almost immediate, although some computation time might be needed to derive survival function for instance.

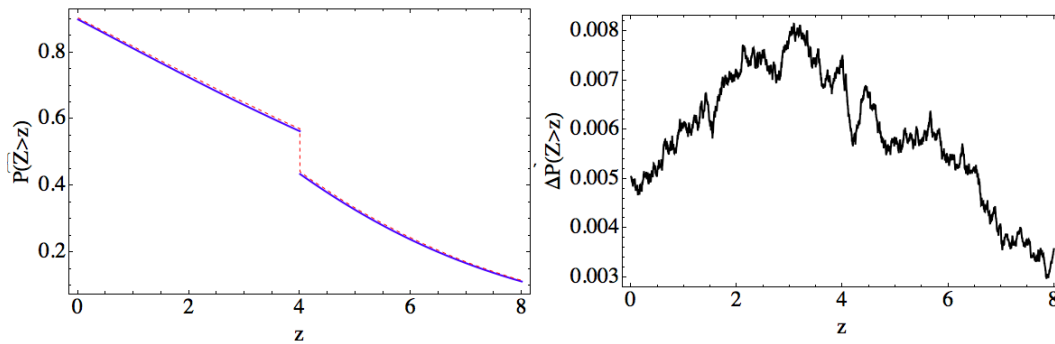


FIGURE 4.9 – Survival function of the reinsurance cost obtained by polynomial expansion (blue) and Monte-Carlo simulations (red) ; Difference between the survival function obtained with Monte Carlo and the one obtained by polynomial expansion

		P(Z>z)	
z	Monte Carlo approximation	Polynomial approximation	
0	0.90385	0.898808	
2	0.73193	0.724774	
4	0.44237	0.435013	
6	0.24296	0.237576	
8	0.	0.	

TABLEAU 4.1 – Survival function of the reinsurance cost obtained by polynomial expansion and Monte-Carlo simulations

4.6 Conclusion

We have considered polynomial expansions of bivariate densities with well defined Laplace transforms. We focused our attention on densities on $\mathbb{R}^+ \times \mathbb{R}^+$ and considered Laguerre polynomials expansions. MOBVE and DBVE distributions have been studied and numerical illustrations show great accuracy of the method. An application to a bivariate aggregate claims amount distribution with dependence has been proposed. This work is a numerical step that ensures the efficiency of the bivariate polynomial approach. Interesting applications follow from the fact that the approximants admit a tractable form. The approximation of the BPDF can turn into a nonparametric estimation of the BPDF. Indeed, the coefficients of the expansion can be replaced with their empirical counterparts when data are available. The statistical extension of this numerical method will be at the center of a forthcoming paper.

Acknowledgements

The authors would like to thank Pierre Miehe and Gilbert Macquart for their useful comments. This work is partially funded by the AXA research fund and the research chair *Actuariat Durable*, sponsored by Milliman.

4.7 Références

- ABATE, J., G. CHOUDHURY et W. WHITT. 1995, «On the Laguerre method for numerically inverting Laplace transforms», *INFORMS Journal on Computing*, vol. 8, n° 4, p. 413–427. [117](#)
- ABATE, J., G. CHOUDHURY et W. WHITT. 1998, «Numerical inversion of multidimensional Laplace transform by the Laguerre method», *Performance Evaluation*, vol. 31, n° 3, p. 229–243. [117](#), [122](#)
- AMBAGASPITIYA, R. 1998, «On the distribution of a sum of correlated aggregate claims», *Insurance : Mathematics and Economics*, vol. 23, n° 1, p. 15–19. [117](#)
- AMBAGASPITIYA, R. 1999, «On the distributions of two classes of correlated aggregate claims», *Insurance : Mathematics and Economics*, vol. 24, n° 3, p. 301–308. [117](#)
- BARNDORFF-NIELSEN, O. 1978, *Information and Exponential Families in Statistical Theory*, Wiley. [118](#)
- CHOUDHURY, G., D. LUCANTONI et W. WHITT. 1994, «Multidimensional transform inversion application to the transient M / G / 1 queue», *The Annals of Applied Probability*, vol. 4, n° 3, p. 719–740. [117](#)
- DIACONIS, P. et R. GRIFFITHS. 2012, «Exchangeables pairs of Bernoulli random variables, Krawtchouck polynomials, and Ehrenfest urns», *Australian and New Zealand Journal of Statistics*, vol. 54, n° 1, p. 81–101. [125](#)
- DOWTON, F. 1970, «Bivariate exponential distributions in reliability theory», *Journal of the Royal Statistical Society. Series B (Methodological)*, p. 408–417. [126](#), [127](#)
- HE, Q., H. NAGARAJA et C. WU. 2013, «Efficient simulation of complete and censored samples from common bivariate distributions», *Computational Statistics*, vol. 28, n° 6, p. 2479–2494. [132](#)
- JIN, T. et J. REN. 2013, «Recursions and Fast Fourier Transform for a new bivariate aggregate claims model», *Scandinavian Actuarial Journal*, , n° ahead-of-print, p. 1–24. [117](#)
- KOUDOU, A. E. 1995, *Problèmes de marges et familles exponentielles naturelles*, thèse de doctorat, Toulouse. [125](#)
- KOUDOU, A. E. 1996, «Probabilités de Lancaster», *Expositiones Mathematicae*, vol. 14, p. 247–276. [125](#)
- LANCASTER, H. 1958, «The structure of bivariate distributions», *The Annals of Mathematical Statistics*, p. 719–736. [125](#)
- MARSHALL, A. et L. OLKIN. 1967, «A multivariate exponential distribution», *Journal of the American Statistical Association*, vol. 62, n° 317, p. 30–44. [129](#)

- MNATSAKANOV, R. M. 2011, «Moment-recovered approximations of multivariate distributions : The Laplace transform inversion», *Statistics and Probability Letters*, vol. 81, n° 1, p. 1–7. [117](#), [127](#)
- MNATSAKANOV, R. M. et K. SARKISIAN. 2013, «A note on recovering the distribution from exponential moments», *Applied Mathematics and Computation*, vol. 219, p. 8730–8737. [130](#)
- MNATSAKANOV, R. M., K. SARKISIAN et A. HAKOBYAN. 2014, «Approximation of the ruin probability using the scaled Laplace transform inversion», *Working Paper*. [130](#)
- MORRIS, C. N. 1982, «Natural Exponential Families with Quadratic Variance Functions», *The Annals of Mathematical Statistics*, vol. 10, n° 1, p. 65–80. [119](#)
- SUNDT, B. 1999, «On multivariate Panjer recursions», *Astin Bulletin*, vol. 29, n° 1, p. 29–45. [117](#)
- SZEGÖ, G. 1939, *Orthogonal Polynomials*, vol. XXIII, American mathematical society Colloquium publications. [121](#)
- VERNIC, R. 1999, «Recursive evaluation of some bivariate compound distributions», *Astin Bulletin*. [117](#)

Chapitre 5

Is it optimal to group policyholders by age, gender, and seniority for BEL computations based on model points ?

Sommaire

5.1 Introduction	140
5.2 Cash flows projection models and best estimate liabilities assessment	141
5.3 Presentation of the aggregation procedure	143
5.3.1 The partitioning step	144
5.3.2 The aggregation step	147
5.4 Illustration on a real life insurance portfolio	147
5.5 Conclusion	154
5.6 Références	154

Abstract

An aggregation method adapted to life insurance portfolios is presented. We aim at optimizing the computation time when using Monte Carlo simulations for best estimate liability calculation. The method is a two-step procedure. The first step consists in using statistical partitioning methods in order to gather insurance policies. The second step is the construction of a representative policy for each aforementioned groups. The efficiency of the aggregation method is illustrated on a real saving contracts portfolio within the frame of a cash flow projection model used for best estimate liabilities and solvency capital requirements computations. The procedure is already part of AXA France valuation process.

Keywords : Solvency II, saving contracts claim reserving, statistical partitioning and classification, functional data analysis, longitudinal data classification.

5.1 Introduction

One of the challenges of stochastic asset/liability modeling for large portfolios of participating contracts is the running time. The implementation of such models requires a cumbersome volume of computations within the framework of Monte Carlo simulations. Indeed, the model has to capture the strong interactions existing between asset and liability through the lapse behavior and the redistribution of the financial and technical result. For an overview we refer to Planchet et al. [PLANCHET et collab. \[2011\]](#). Using a stochastic asset/liability model to analyze large blocks of business is often too time consuming to be practical. Practitioners make compromises to tackle this issue.

Grouping methods, which group policies into model cells and replace all policies in each groups with a representative policy, is the oldest form of modelling techniques. It has been advised in [EIOPA \[2010\]](#) and is commonly used in practice due to its availability in commercial actuarial softwares such as MG-ALFA, GGY-Axis and TRICAST suite. In a recent PhD thesis [SARUKKALI \[2013\]](#), representative portfolios have been drawn from a stratified sampling procedure over the initial portfolio. Computations are achieved on this smaller portfolio and the procedure is repeated to yield a final approximation. The idea remains to reduce the computation time by reducing the number of model cells.

Another idea consists in reducing the computation time for one policy. For instance one may think about reducing the number of economic scenarios. In [DEVINEAU et LOISEL \[2009\]](#) an algorithm, inspired by importance sampling, is presented to select the most relevant financial scenarios in order to estimate quantiles of interest. Applications to the estimation of remote (but not extreme) quantiles can be found in [CHAUVIGNY et collab. \[2011\]](#), in addition to consistency results regarding the estimator. However, we are interested in the calculation of the value of Best Estimate Liabilities (BEL) for participating contracts, which are the probability weighted average of future cash flows taking into account the time value of money. Therefore, Planchet et al. [NTEUKAM et PLANCHET \[2012\]](#) presented a way to aggregate all the scenarios in a set of cha-

racteristic trajectories associated to a probability of occurrence, well adapted to BEL computations. Another well-known modeling technique when dealing with financial sensitivities is the replicating portfolio method. As life insurance contracts share many features with derivative contracts, the idea of representing liabilities by a portfolio of financial products comes naturally. For an overview of the replicating portfolio method, we refer to [DEVINEAU et CHAUVIGNY \[2011\]](#); [SCHRAGER \[2008\]](#). We also want to mention the work of Planchet et al. [BONNIN et collab. \[2014\]](#), where an analytical approximation formula for best estimate valuation of saving contracts is derived. The approximation is based on the evaluation of a coefficient, which when applied to the mathematical reserve yields the BEL. We believe that the ways of dealing with computation time issues are not in competition, for instance a grouping method combined with an optimization of financial scenarios may yield great results.

In this paper, we present a new grouping method and we focus on the calculation of the value of Best Estimate Liabilities (BEL) for participating contracts via an asset/liability model. This work has been motivated by the revision of the valuation process in AXA France to be compliant with Solvency II. The aggregation method must meet many practical requirements such as easy implementation and understanding, applicability to the whole life insurance portfolio and good empirical and theoretical justification. We believe that many insurance companies face computation time problems, and this paper may help practionners to solve it. In Section 2, we give a description of the cash flow projection model being used. In Section 3, we describe a statistical based way to group policies, followed by procedures to construct the representative policy. Section 4 illustrates the efficiency of the method on a portfolio of saving contracts on the French market.

5.2 Cash flows projection models and best estimate liabilities assessment

We consider a classical asset-liability model to project the statutory balance sheet and compute mathematical reserves for saving contracts. The settings and notations are closed to those in [BONNIN et collab. \[2014\]](#). Consider a saving contract with surrender value $SV(0)$ at time $t = 0$, the surrender value at time t is defined as

$$SV(t) = SV(0) \times \exp\left(\int_0^t r_a(s) ds\right), \quad (5.1)$$

where r_a denotes the instantaneous accumulation rate (including any guaranteed rate) modeled by a stochastic process. We are interested in the present surrender value at time t , we need therefore to add a discount factor in order to define the Present Surrender Value

$$\begin{aligned} \text{PSV}(t) &= SV(t) \times \exp\left(-\int_0^t r_\delta(s) ds\right) \\ &= SV(0) \times \exp\left(\int_0^t (r_a(s) - r_\delta(s)) ds\right), \end{aligned} \quad (5.2)$$

where r_δ denotes the instantaneous discount rate, also modeled by a stochastic process. The spread between the accumulation and the discount rates in (5.2) is then a stochastic process. We assume that these stochastic processes are governed by a probability measure denoted by Q^f . We denote by \mathbf{F} a financial scenario, drawn from Q^f , which corresponds to a trajectory of r_δ and r_a . Let $\tau|\mathbf{F}$ be the early surrender time, modeled by a random variable having probability density function $f_{\tau|\mathbf{F}}$ on the positive half line, absolutely continuous with respect to the Lebesgue measure. The probability density function can be interpreted as an instantaneous surrender rate between times t and $t + dt$, which depends on the characteristics of the policyholder like for instance his age, his gender, the seniority of his contract and the financial environment that may influence policyholders behavior. More specifically, in the case of a saving contract, the payment of the surrender value occurs in case of early withdrawal due to lapse, death, or expiration of the contract. In the cash flow projection model definition, an horizon of projection is usually specified. There are also life insurance contracts with a fixed term. Both of these times are deterministic. We denote by T the minimum of the expiration date of the contract and the horizon of projection. The real surrender time $\tau \wedge T|\mathbf{F} = \min(\tau|\mathbf{F}, T)$ is a random variable associated with a probability measure divided into the sum of a singular part and a continuous part

$$dP_{\tau \wedge T|\mathbf{F}}(t) = f_{\tau|\mathbf{F}}(t)d\lambda(t) + \overline{F_{\tau|\mathbf{F}}}(T)\delta_T(t), \quad (5.3)$$

where λ is the Lebesgue measure on $[0, T]$, δ_T is the Dirac measure at T and $\overline{F_{\tau|\mathbf{F}}}(t)$ denotes the survival function associated to the random time $\tau|\mathbf{F}$. The BEL at time $t = 0$ associated to the financial scenario \mathbf{F} is defined as

$$\text{BEL}^{\mathbf{F}}(0, T) = E^{\mathbb{P}_{\tau \wedge T|\mathbf{F}}}(\text{PSV}(\tau \wedge T|\mathbf{F})). \quad (5.4)$$

We refer to GERBER [1990]; PLANCHET et collab. [2011] for this definition of best estimate liabilities. We write the BEL given a financial scenario as follows

$$\begin{aligned} \text{BEL}^{\mathbf{F}}(0, T) &= E^{\mathbb{P}_{\tau \wedge T|\mathbf{F}}}(\text{PSV}(\tau \wedge T|\mathbf{F})) \\ &= \int_0^{+\infty} \text{SV}(0) \times \exp\left(\int_0^t (r_a(s) - r_\delta(s)) ds\right) dP_{\tau \wedge T|\mathbf{F}}(t) \\ &= \int_0^T \text{SV}(0) \times \exp\left(\int_0^t (r_a(s) - r_\delta(s)) ds\right) f_{\tau|\mathbf{F}}(t) dt \\ &\quad + \overline{F_{\tau|\mathbf{F}}}(T) \times \text{SV}(0) \times \exp\left(\int_0^T (r_a(s) - r_\delta(s)) ds\right). \end{aligned} \quad (5.5)$$

In order to avoid tedious calculations of the integral in (5.5), time is often discretized. The BEL is therefore written as

$$\text{BEL}^{\mathbf{F}}(0, T) \approx \left[\sum_{t=0}^{T-1} p^{\mathbf{F}}(t, t+1) \prod_{k=0}^t \frac{1 + r_a(k, k+1)}{1 + r_\delta(k, k+1)} + p^{\mathbf{F}}(T) \prod_{k=0}^{T-1} \frac{1 + r_a(k, k+1)}{1 + r_\delta(k, k+1)} \right] \text{SV}(0), \quad (5.6)$$

where $p^{\mathbf{F}}(t, t+1)$ is the probability that surrender occurs between time t and $t+1$, and $r_a(t, t+1)$ and $r_\delta(t, t+1)$ are the accumulation and discount rates between time t and $t+1$. Monte Carlo methods for BEL evaluation consists in generating a set of financial scenarios under Q^f and compute the BEL for each one of them. The final estimation is

the mean over the set of all scenarios. This procedure is fast enough for one policy, it becomes time consuming for a large portfolio. The formula given in (5.6) is commonly used by practionners and compliant with Solvency II. However, we know that there is room for discussion and improvement regarding the "economic" scenario generation, the assumptions on the liability side or the link between assets and liabilities. Nevertheless, the efficiency of our grouping methodology is not affected by these aspects hence we choose to stick to AXA France assumptions without discussing it further.

5.3 Presentation of the aggregation procedure

The goal of aggregation procedures is to reduce the size of the input portfolio of the cash flow projection model. The first step consists in creating groups of policies that share similar features. The second one is the definition of an "average" policy, called Model Point (MP), that represents each group and forms the aggregated portfolio. The initial surrender value of the MP is the sum of the initial surrender values over the represented group. The method must be flexible so as to generate the best aggregated portfolio under the constraint of a given number of MPs. Let us consider two contracts having identical guarantees and characteristics (same age, same seniority,...). Therefore they have identical surrender probabilities. We build a contract having these exact characteristics and whose initial surrender value is the sum of the initial surrender values of the two aforementioned contracts. The BEL of this contract, given a financial scenario, is

$$\begin{aligned}
 \text{BEL}_{\text{MP}}^{\text{F}}(0, T) &= \left[\sum_{t=0}^{T-1} p^{\text{F}}(t, t+1) \prod_{k=0}^t \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} + p^{\text{F}}(T) \prod_{k=0}^{T-1} \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} \right] \\
 &\times \text{SV}_{\text{MP}}(0) \\
 &= \left[\sum_{t=0}^{T-1} p^{\text{F}}(t, t+1) \prod_{k=0}^t \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} + p^{\text{F}}(T) \prod_{k=0}^{T-1} \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} \right] \\
 &\times \sum_{i=1}^2 \text{SV}_{\text{C}_i}(0) \\
 &= \left[\sum_{t=0}^{T-1} p^{\text{F}}(t, t+1) \prod_{k=0}^t \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} \right] \text{SV}_{\text{C}_i}(0) \tag{5.7}
 \end{aligned}$$

$$\begin{aligned}
 &+ \sum_{i=1}^2 \left[p^{\text{F}}(T) \prod_{k=0}^{T-1} \frac{1+r_a(k, k+1)}{1+r_\delta(k, k+1)} \right] \text{SV}_{\text{C}_i}(0) \\
 &= \sum_{i=1}^2 \text{BEL}_{\text{C}_i}^{\text{F}}(0, T), \tag{5.8}
 \end{aligned}$$

where $\text{BEL}_{\text{C}_i}^{\text{F}}(0, T)$ and $\text{SV}_{\text{C}_i}(0)$ are the best estimate liability and initial surrender value of the contract $i \in \{1, 2\}$. The idea behind the grouping strategy lies in this linearity property of the BEL. The aggregation of contracts having the same surrender probabilities leads to an exact evaluation of the BEL of the portfolio. The creation of an aggregated portfolio by grouping the policies having identical characteristics leads to a portfolio that is usually still too big to perform multiple valuations (with different sets

of hypothesis). However one particular valuation might be doable using this aggregated portfolio in order to get a benchmark value for the BEL and assess the accuracy of the aggregation procedure in the validation phase. One may also note that the use of partitioning algorithms on the aggregated portfolio is faster because it is much smaller than the initial one. The AXA France portfolio of saving contracts contained different products associated to different guarantees. Obviously we cannot group together two contracts that belong to different lines of product. The variable PRODUCT divides the portfolio into sub-portfolios on which the aggregation procedure is used separately. We get aggregated sub-portfolios that are concatenated to yield the final aggregated portfolio. In the first subsection, we explain how to gather contracts having similar surrender probabilities and in the second subsection how to build a representative contract for each group.

5.3.1 The partitioning step

A portfolio is a set of contracts $\mathcal{P} = \{\mathbf{x}_i\}_{i \in 1, \dots, n}$, where n is the size of the portfolio. We aim at partitioning n contracts into k sets $\mathcal{C} = \{C_1, \dots, C_k\}$. We use clustering algorithms to identify sub-portfolios. A choice has to be made concerning the variables that characterize each observation and the metric that measures the dissimilarity between two observations. In order to get closer to the additivity of the BEL, every individuals in the portfolio is represented by its sequences of surrender probabilities

$$\mathbf{x}_i = \begin{pmatrix} p_i^{\mathbf{F}_1}(0, 1) & \dots & p_i^{\mathbf{F}_1}(T) \\ \vdots & \ddots & \vdots \\ p_i^{\mathbf{F}_N}(0, 1) & \dots & p_i^{\mathbf{F}_N}(T) \end{pmatrix} \quad (5.9)$$

where $\{\mathbf{F}_1, \dots, \mathbf{F}_N\}$ denotes set of N financial scenarios. Policies are therefore represented by matrices of size $N \times (T + 1)$. The determination of $p^{\mathbf{F}}(t, t + 1)$ necessitates the evaluation of a mortality and a lapse rate from one period to another using classical actuarial tools, see [PETAUTON et collab. \[2002\]](#). For saving contracts, lapse behavior is usually split between two components : a structural part linked to endogeneous factors and a dynamic part driven by the financial environment. the structural part is derived using historical data at product level. Figure 5.1 displays the lapse rate according to the seniorities of contracts in the portfolio we will consider later. Figure 5.1 clearly shows a step when reaching the 4th anniversary year and a peak at the 8th anniversary of the contract. The explanation lies in the french market specific tax rules that depend on the seniority.

Remarque 6. *The shape of the lapse rate curve in Figure 5.1 may entail questions regarding the use of a continuous probability density function to model surrendering time in (5.3). This is another practical advantage of using a discretized expression to compute best estimates liabilities. However, there are in the applied probability literature interesting probability distributions on \mathbb{R}^+ that allow great flexibility. they are called phase-type distributions and might be appropriate to model surrender time, an overview is given in Chapter 8 of [ASMUSSEN et ABRECHER \[2010\]](#).*

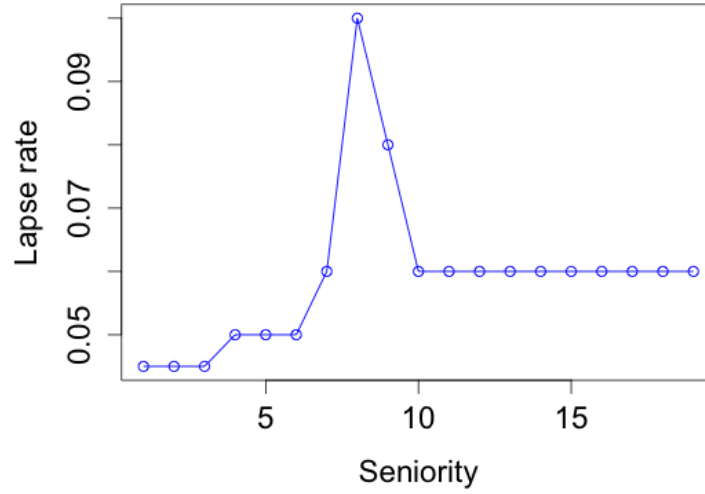


FIGURE 5.1 – Lapse rate curve depending on the seniority of contracts in a saving contract portfolio

The dynamic part introduces more lapses when the client is not satisfied with the annual performance of his contract either because it is less than expected given past performances or because competitors offer more. Both reasons are governed by the financial situation and modeled accordingly. We notice an interesting fact that allows to reduce the operational burden that occurs when working with matrices as in (5.9). The financial scenario has a very limited impact on the proximity of contracts. Namely, two contracts close to each other given a financial scenario are also close with respect to another financial scenario. If there exist $k \in \{1, \dots, N\}$ such that $\mathbf{x}_i^{\mathbf{F}^k} \approx \mathbf{x}_j^{\mathbf{F}^k}$ then we have approximately $\mathbf{x}_i^{\mathbf{F}^l} \approx \mathbf{x}_j^{\mathbf{F}^l}$ for all $l \in \{1, \dots, N\}$. This is due to the fact that we consider saving contracts that belong to the same line of product. Note also that it consumes less resources to store a vector for each policy than a matrix which is very important from an operational point of view. Thus we decided to consider for each contract a vector of surrender probabilities given one of the financial scenario. Every policy in the portfolio is now represented by a vector of surrender probabilities

$$\mathbf{x}_i = (p_i(0, 1), p_i(1, 2), \dots, p_i(T-1, T), p_i(T)). \quad (5.10)$$

The variables are quantitative and fall between 0 and 1. The natural dissimilarity measure between two observations is the euclidean distance defined as

$$\|\mathbf{x}_i - \mathbf{x}_j\|_2 = \sqrt{(p_i(0, 1) - p_j(0, 1))^2 + \dots + (p_i(T) - p_j(T))^2}. \quad (5.11)$$

Remarque 7. *It is worth noting that dealing with matrices instead of vectors is not a problem. Indeed, the techniques presented below still work by considering a dissimilarity measure between matrices.*

Partitioning algorithms are designed to find the k -sets partition that minimizes the Within Cluster Sum of Square (WCSS), which characterises the homogeneity in a

group,

$$\tilde{\mathcal{C}} = \underset{\mathcal{C}}{\operatorname{argmin}} \sum_{j=1}^k \sum_{\mathbf{x} \in C_j} \|\mathbf{x} - \mu_j\|_2^2, \quad (5.12)$$

where μ_j is the mean of the observations that belongs to the set C_j . The two foreseen methods are the so called KMEANS procedure and the agglomerative hierarchical clustering procedure with Ward criterion. These two standard methods are described in [EVERITT et collab. \[1995\]](#); [HARTIGAN \[1975\]](#), and Ward criterion has been introduced in [WARD \[1963\]](#).

The KMEANS algorithm starts with a random selection of k initial means or centers, and proceeds by alternating between two steps. The assignment step permits to assign each observation to the nearest mean and the update step consists in calculating the new means resulting from the previous step. The algorithm stops when the assignments no longer change. The algorithm that performs an agglomerative hierarchical clustering uses a bottom up strategy in the sense that each observation forms its own cluster and pairs of cluster are merged sequentially according to Ward criterion. At each step, the number of clusters decreases of one unit and the WCSS increases. This is due to Huygens theorem that divides the Total Sum of Square (TSS), into Between Cluster Sum of Square (BCSS) and WCSS,

$$\begin{aligned} \sum_{\mathbf{x} \in \mathcal{P}} \|\mathbf{x} - \mu\|_2^2 &= \sum_{j=1}^k \sum_{\mathbf{x} \in C_j} \|\mathbf{x} - \mu_j\|_2^2 + \sum_{j=1}^k \|\mu_j - \mu\|_2^2 \\ \text{TSS} &= \text{WCSS}(k) + \text{BCSS}(k) \end{aligned}$$

Note that TSS is constant and that BCSS and WCSS evolve in opposite directions. Applying Ward's criterion yields the aggregation of the two individuals that goes along with the smallest increase of WCSS at each step of the algorithm. The best way to visualize the data is to plot "surrender trajectories" associated with each policy as in [Figure 5.5](#). Our problem is analogous to the problem of clustering longitudinal data that arises in biostatistics and social sciences. Longitudinal data are obtained by doing repeated measurements on the same individual over time. We chose a nonparametric approach, also chosen in [GENOLINI et FALISSARD \[2010\]](#); [HEJBLUM et collab. \[2012\]](#). A parametric approach is also possible by assuming that the dataset comes from a mixture distribution with a finite number of components, see [NAGIN \[1999\]](#) for instance. We believe that the nonparametric approach is easier to implement and clearer from a practitioner point of view.

The KMEANS method, that takes the number of clusters as a parameter, seems to be more suited to the problem than the agglomerative hierarchical clustering method. In the Agglomerative Hierarchical Clustering (AHC) algorithm the partition in k sets depends on the previous partition. Furthermore, the KMEANS algorithm is less greedy in the sense that fewer distances need to be computed. However the KMEANS algorithm is an optimization algorithm and the common problem is the convergence to a local optimum due to bad initialization. To cope with this problem, we initialize the centers as the means of clusters builded by the AHC, thus we ensure that the centers are geometrically far from each other. The value of the BEL is highly correlated to the initial surrender value, thus a bad representation of policies having a significant initial

surrender value gives rise to a significant negative impact on the estimation error after aggregation. We decide to define a weight according to the initial surrender value as

$$w_{\mathbf{x}} = \frac{SV_{\mathbf{x}}(0)}{\sum_{\mathbf{x} \in \mathcal{P}}^n SV_{\mathbf{x}}(0)}, \quad (5.13)$$

and we define the Weighted Within Cluster Inertia - WWCI as

$$WWCI(k) = \sum_{j=1}^k \sum_{\mathbf{x} \in C_j} w_{\mathbf{x}} \|\mathbf{x} - \mu_j\|_2^2, \quad (5.14)$$

where μ_j becomes the weighted mean over the set C_j . Each surrender path is then stored in a group.

Remarque 8. *A time continuous approach within the cash flow projection model would leave us with probability density function to group. The problem would be analogous to the clustering of functional data that have been widely studied in the literature. A recent review of the different techniques has been done in JACQUES et PREDA [2013].*

5.3.2 The aggregation step

The aggregation step leads to the definition of a representative policy for each group resulting from the partitioning step. Probabilities of surrender depend on characteristics of the contracts and of the policyholders. The best choice as a representative under the least square criterion is the barycenter. Its surrender probabilities are defined through a mixture model

$$f_{\tau_C}(t) = \sum_{\mathbf{x} \in C} w_i f_{\tau_{\mathbf{x}}}(t), \quad (5.15)$$

where C is a group of policies and τ_C is the random early surrender time for every member of C . The equivalent within a discrete vision of time is a weighted average of the surrender probabilities with respect to each projection year. The probability density function of surrender of a given contract is associated to its age and seniority. The PDF defined in (5.15) is not associated to an age and a seniority. This fact might give rise to an operational problem if every MP in the aggregated portfolio must have an age and a seniority. The number of suitable combinations of age and seniority fall into a finite set given the possible features of policies. It is then possible to generate every "possible" surrender probability density functions in order to choose the closest to the barycenter. This optimal density function might be associated with a policy (or equivalently a combination of one age and one seniority) that does not exist in the initial portfolio.

5.4 Illustration on a real life insurance portfolio

The procedure is illustrated within the frame of a saving contracts portfolio extracted from AXA France portfolio. The mechanism of the product is quite simple. The policyholder makes a single payment when subscribing the contract. This initial capital is the initial surrender value that will evolve during the projection, depending on the investment strategy and the financial scenario. The policyholder is free to withdraw

at any time. In case of death, the surrender value is paid to the designated beneficiaries. The contractual agreement does not specify a fixed term. The projection horizon is equal to 30 years. Best estimate liabilities are obtained under a discrete vision of time as in (5.6). Mortality rates are computed using an unisex historical life table. Mortality depends therefore only on the age of the insured. Lapse probabilities are computed with respect to the observed withdrawal depending on the seniority and given a financial scenario that do not trigger any dynamic lapse. The main driver that explains the withdrawal behavior is the specific tax rules applied on French life insurance contracts, see Figure 5.1. Financial scenarios are generated through stochastic modeling. The instantaneous interest rates are stochastic processes and simulations are completed under a risk neutral probability. Table 5.1 gives the number of policies and the amount of the initial reserves in the portfolio. Surrender probabilities depend only on the age

Number of policies	Mathematical provision (euros)
140 790	2 632 880 918

TABLEAU 5.1 – Number of policies and amount of the initial surrender value of the portfolio

and the seniority and thus the heterogeneity of the trajectories depends on the distribution of ages and seniorities in the portfolio. Table 5.2 gives a statistical description of the variable AGE. The minimum is 1 because parents can subscribe life insurance contracts for their children. Table 5.3 gives a statistical description of the variable SENIORITY. The range of the variable AGE and SENIORITY are completely linked to the number of trajectories in the portfolio as there is a one to one correspondence between a trajectory and a combination of age and seniority. Figure 5.5 displays the surrender trajectories in the portfolio. An increase of the maximum of the variable SENIORITY would increase the number of trajectories in the portfolio. Partitioning algorithms are designed to cope as far as possible with an increase of the heterogeneity of the data. This makes our aggregation procedure quite robust from one portfolio to another that might have different AGE and SENIORITY distributions.

Variable : AGE			
Mean	Standard deviation	Minimum	Maximum
49.09	18.57	1	102

TABLEAU 5.2 – Statistical description of the variable AGE in the portfolio

Variable : SENIORITY			
Mean	Standard deviation	Minimum	Maximum
4.10	1.63	1	7

TABLEAU 5.3 – Statistical description of the variable SENIORITY in the portfolio

In Section 3, it has been pointed out that an exact evaluation of the BEL is obtained with a portfolio that aggregates policies having the same key characteristics. In our modeling, policies that have identical age and seniority are grouped together in order to get a

first aggregation of the portfolio that provides an exact value of BEL. Table 5.4 gives the number of policies and the BEL resulting from the valuation of this first aggregation of the portfolio. The error is defined as the relative difference between the exact BEL

Number of policies	BEL (euros)
664	2 608 515 602

TABLEAU 5.4 – Number of MP and best estimate liability of the aggregated portfolio

and the BEL obtained with an aggregated portfolio. We also compare the two aggregation ways discussed in Section 3.2. One corresponds to the exact barycenter of each group (METHOD=BARY), the other being the closest-to-barycenter policy associated with an age and a seniority (METHOD=PROXYBARY). These procedures are compared to a more "naive" grouping method (METHOD=NAIVE) that consists in grouping policies having the same seniority and belonging to a given class of age. The classes are defined by the quartiles of the distribution of ages in the portfolio. The ages and seniorities of the MP associated with each group are obtained through a weighted mean. The weights are defined as in (5.13). The "naive" method leads to an aggregated portfolio with 28 MPs. Table 5.5 reports the errors for the three methods with 28 MP. The

BEL error (euros) METHOD=BARY	BEL error (euros) METHOD=PROXYBARY	BEL error (euros) METHOD=Naive
-10 880	-199 734	1 074 983

TABLEAU 5.5 – Best estimate liabilities error with 28 model points depending on the aggregation method

two proposed methods outperform greatly the naive one. Figure 5.2 shows the error on BEL computation according to the number MPs created within the frame of BARY and PROXYBARY. The proposed methods permit a better accuracy even for a smaller number of MP. The use of BARY is the best choice. The question of the optimal choice of the number of clusters arises naturally. The optimal number of clusters has been widely discussed in the literature, and there exists many indicators. The main idea is to spot the number of clusters for which the WWCI reaches a sort of plateau when the number of clusters is increasing. Figure 5.3 displays the WWCI depending on the number of clusters. The optimal number of clusters seems to be 3 or 6. In order to automatize the choice, we can use indicators. The relevance of such indicators often depends on the partitioning problem. We need to choose the best suited to our problem. Among the indicators recommended in the literature, we find the index due to Calinsky and Harabasz CALINSKY et HARABASZ [1974] defined as

$$CH(k) = \frac{WBCI(k)/(k-1)}{WWCI(k)/(n-k)}, \quad (5.16)$$

where n is the number of observations, WBCI and WWCI denote the Weighted Between and Weighted Within Cluster Inertia. The idea is to find the number of clusters k that maximises CH. Note that CH(1) is not defined. This indicator is quite simple to understand, as a good partition is characterized by a large WBCI and a small WWCI. Another

indicator has been proposed in KRZANOWSKI et LAI [1988] as follows : First define the quantity

$$\text{DIFF}(k) = (k - 1)^{2/p} \times \text{WWCI}(k - 1) - k^{2/p} \times \text{WWCI}(k), \quad (5.17)$$

and choose k which maximises

$$\text{KL}(k) = \left| \frac{\text{DIFF}(k)}{\text{DIFF}(k + 1)} \right|. \quad (5.18)$$

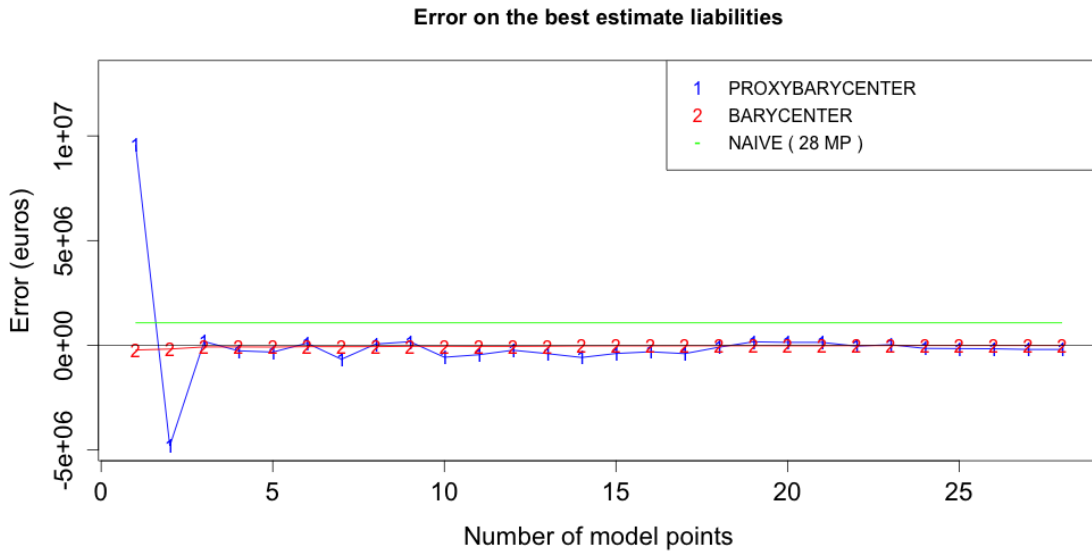


FIGURE 5.2 – Error on the BEL evaluation depending on the number of model points and the aggregation method

This permits to compare the decreasing of WWCI in the data with the decreasing of WWCI within data uniformly distributed through space. The silhouette statistic has been introduced in KAUFMAN et ROUSSEUW [1990] and is defined, for each observation i , as

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}}, \quad (5.19)$$

where $a(i)$ is the average distance between i and the others points in its cluster, and $b(i)$ is the average distance from i to the data points in the nearest cluster besides its own. A point is well clustered when $s(i)$ is large. The optimal number of clusters maximises the average of $s(i)$ over the data set. Figure 5.4 displays the different indicators computed for every partition ranging from one to twenty groups. The different indicators seems to retain 3 clusters (except for KL, that is maximized for 6 clusters but still have a large value for 3 clusters). Figure 5.5 offers a visualization of the 3-groups partition of the portfolio. Table 5.6 report the errors on the BEL, which have been normalized by its exact value and expressed in percentage. The presented methods outperform greatly the naive one with less model points. Our aggregation procedure does not only performed an accurate BEL evaluation but manages also to replicate the cash

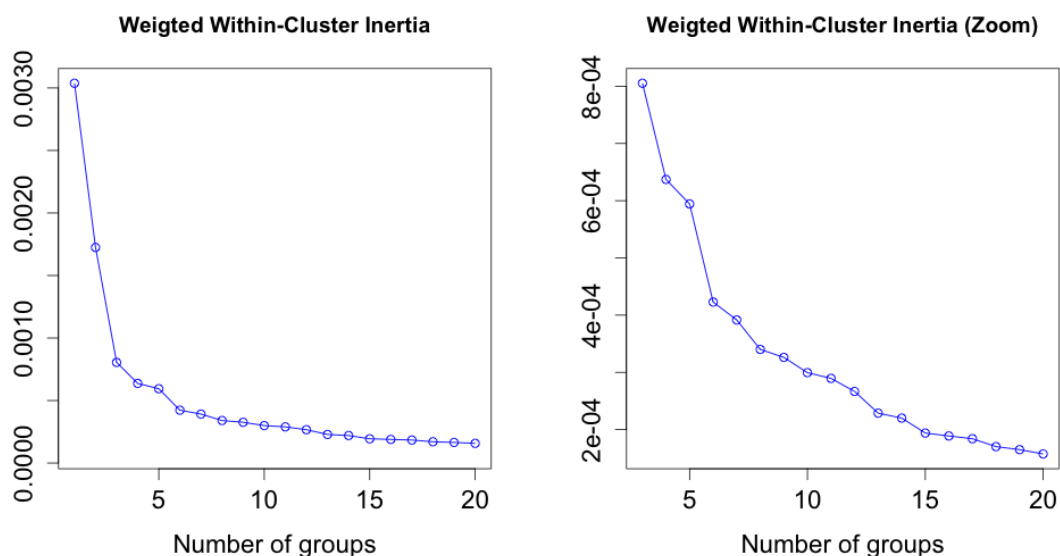


FIGURE 5.3 – WWCI evolution depending on the number of clusters

flows dynamics throughout the entire projection. Figure 5.6 shows the accuracy on the expected present value of the exiting cash flow associated with each projection year for the 3-MP aggregated portfolio.

BEL error % BARYCENTER	BEL error % PROXYBARYCENTER	BEL error % NAIVE (28 MP)
-0.003 %	0.007 %	0.0412 %

TABLEAU 5.6 – Best estimate liabilities error with 3 model points depending on the aggregation method

From a practical point of view, the optimal number of clusters should be associated with a level of error chosen by the user. We did not manage to establish a clear link between WWCI and the error on the BEL evaluations. Maybe it does not exist and we need to define another indicator, instead of WWCI, that we can compute from the probabilities of surrender and that is more linked to the evaluation error. This a topic of current research. We may also add that the optimal number of clusters is not necessarily a problem that needs much thoughts from a practitioner point of view. The number of model points in the aggregated portfolio is often constrained by the capability of computers to deal with the valuation of large portfolios. The number of model points is then a parameter of the aggregation process. Another technical requirement is to take into account the different lines of business. Two life insurance products cannot be grouped together because they may have characteristics, besides their probabilities of surrender, that impact BEL computations. A solution is to allocate a number of model points for each line of business in proportion to their mathematical provision.

CHAPITRE 5. IS IT OPTIMAL TO GROUP POLICYHOLDERS BY AGE, GENDER, AND SENIORITY FOR BEL COMPUTATIONS BASED ON MODEL POINTS?

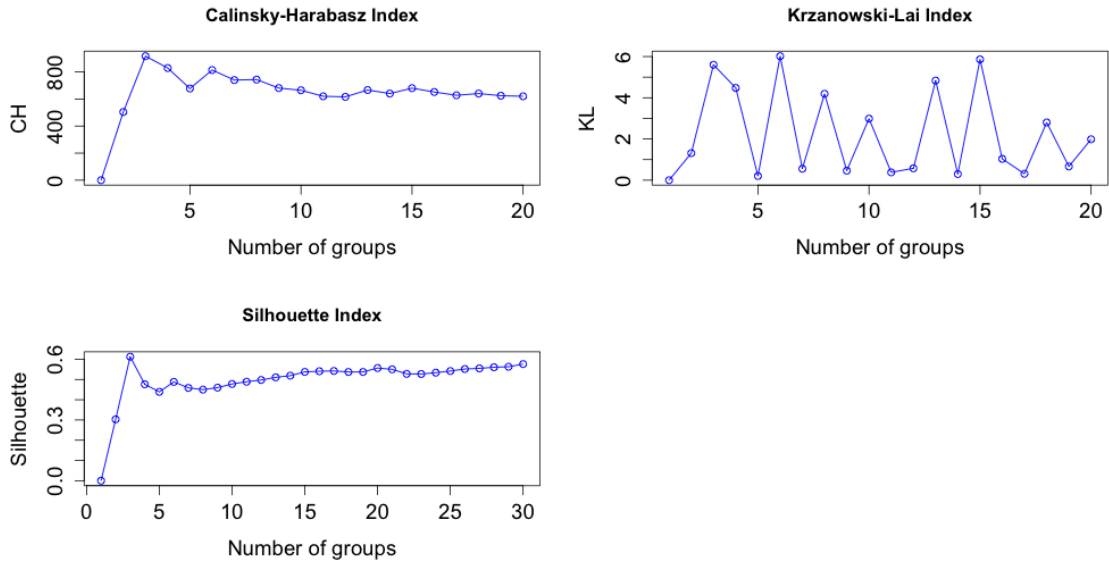


FIGURE 5.4 – Partitionning quality indicators variations depending on the number of clusters

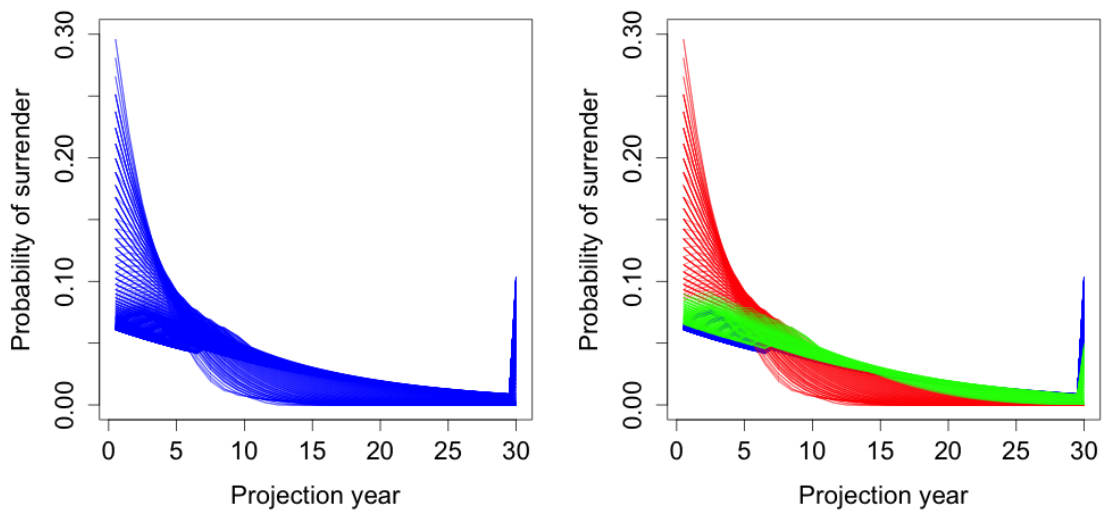


FIGURE 5.5 – Portfolio visualization through its trajectories of surrender

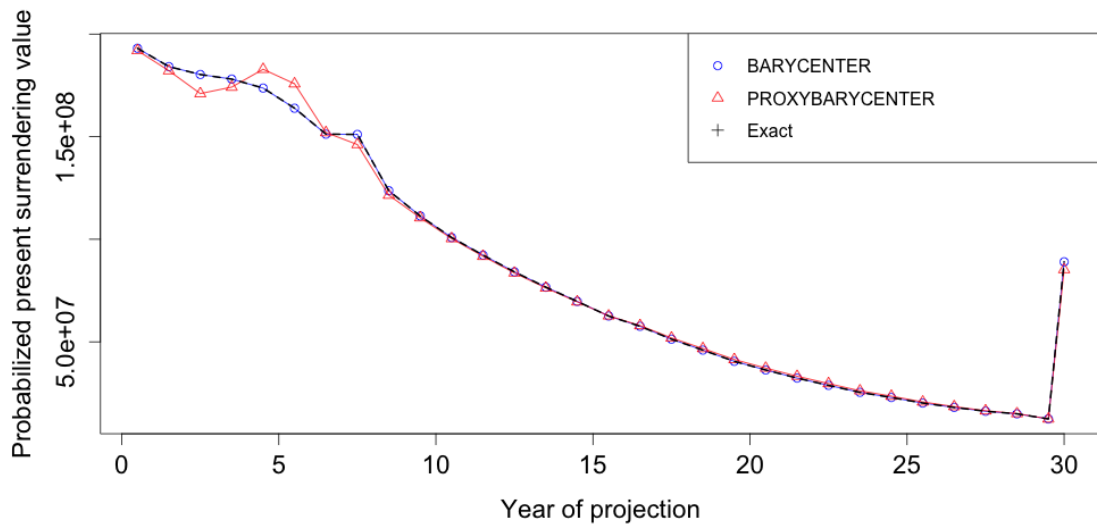


FIGURE 5.6 – Expected present value of surrender during the projection

5.5 Conclusion

The aggregation procedure permits a great computation time reduction that goes along with a limited loss of accuracy. The method is easy to understand and to implement as the statistical tools are available in most datamining softwares. The application may extend to the entire scope of life insurance business, and can be tailored to many cash flow projection model based on the general definition of best estimate liabilities given in the introduction. The production of aggregated portfolios can also be combined to optimization methods concerning economic scenarios. A reduction of the number of financial scenarios would permit to increase the number of model points. The optimal gain of accuracy would result from a trade-off between the number of model points and the number of financial scenarios.

This work represents the successful outcome of a Research and Development project in the industry. It is already implemented in AXA France, from a portfolio that contains millions of policies, the output is a portfolio of only a few thousands model points for a 0.005% error on the BEL. It remains many rooms for improvement, especially at the partitioning level where distance other than euclidean might be better suited to quantify the distance between trajectories. The definition of an indicator that gives insight on the error resulting from the aggregation would be as well a great improvement. For instance, the computation of this indicator may provide an optimal number of model points allocated to each line of business and therefore the whole portfolio.

5.6 Références

- ASMUSSEN, S. et H. ABRECHER. 2010, *Ruin probabilities*, vol. 14, World scientific. [144](#)
- BONNIN, F., F. PLANCHET et M. JULLIARD. 2014, «Best estimate calculations of saving contracts by closed formulas : Application to the ORSA», *European Actuarial Journal*, vol. 4, n° 1, p. 181–196. [141](#)
- CALINSKY, R. B. et J. HARABASZ. 1974, «A dendrite method for cluster analysis», *Communications in Statistics*, vol. 3, p. 1–27. [149](#)
- CHAUVIGNY, M., L. DEVINEAU, S. LOISEL et V. MAUME-DESCHAMPS. 2011, «Fast remote but not extreme quantiles with multiple factors : Application to Solvency II and Enterprise Risk Management», *European Actuarial Journal*, vol. 1, n° 1, p. 131–157. [140](#)
- DEVINEAU, L. et M. CHAUVIGNY. 2011, «Replicating portfolios : Calibration techniques for the calculation of the Solvency II economic capital», *Bulletin Français d'Actuariat*, vol. 21, p. 59–97. [141](#)
- DEVINEAU, L. et S. LOISEL. 2009, «Construction d'un algorithme d'accélération de la méthode des "simulations dans les simulations " pour le calcul du capital économique Solvabilité II», *Bulletin Français d'Actuariat*, vol. 10, n° 17, p. 188–221. [140](#)
- EIOPA. 2010, «Quantitative Impact Studies V : Technical Specifications», cahier de recherche, European comission, Brussels. [140](#)

- EVERITT, B. S., S. LANDAU et M. LEESE. 1995, *Cluster Analysis*, 4^e éd., A Hodder Arnold Publication. 146
- GENOLINI, C. et B. FALISSARD. 2010, «"KML : K-means for longitudinal data"», *Computational Statistics*, vol. 25, n^o 2, p. 317–328. 146
- GERBER, H. 1990, *Life insurance mathematics*, Springer-Verlag, Berlin. 142
- HARTIGAN, J. 1975, *Clustering Algorithms*, Wiley, New York. 146
- HEJBLUM, B., J. SKINNER et R. THIEBAUT. 2012, «Application of gene set analysis of time-course gene expression in a hiv vaccine trial», *33rd annual conference of international society for clinical biostatistics*. 146
- JACQUES, J. et C. PREDÀ. 2013, «Functional data clustering : a survey», *Advances in data analysis and classification*, p. 1–25. 147
- KAUFMAN, L. et P. J. ROUSSEEUW. 1990, *Finding groups in Data : An introduction to cluster analysis*, Wiley, New York. 150
- KRZANOWSKI, W. et Y. T. LAI. 1988, «A criterion for determining the number of groups in a data set using Sum-of-Squares clustering», *Biometrics*, vol. 44, p. 23–34. 150
- NAGIN, D. S. 1999, «Analysing developmental trajectories : A semiparametric group based approach», *Psychological Methods*, p. 139–157. 146
- NTEUKAM, T. et F. PLANCHET. 2012, «Stochastic evaluation of life insurance contracts : Model point on asset trajectories and measurements of the error related to aggregation», *Insurance : Mathematics and Economics*, vol. 51, n^o 3, p. 624–631. 140
- PETAUTON, P., D. KESSLER et J. BELLANDO. 2002, *Théorie et pratique de l'assurance vie : manuel et exercices corrigés*, Dunod. 144
- PLANCHET, F., P. THEROND et M. JUILLARD. 2011, *Modèles financiers en assurance. Analyses de risque dynamique.*, 2^e éd., Economica, Paris. 140, 142
- SARUKKALI, M. 2013, *Replicated Stratified Sampling for Sensitivity Analysis*, thèse de doctorat, University of Connecticut. 140
- SCHRAGER, D. 2008, «Replicating portfolios for insurance liabilities», *AENORM*, , n^o 59, p. 57–61. 141
- WARD, J. H. 1963, «Hierarchical grouping to optimize an objective function», *Journal of the American Statistical Association*, vol. 236–544, p. 236–544. 146

CHAPITRE 5. IS IT OPTIMAL TO GROUP POLICYHOLDERS BY AGE, GENDER, AND SENIORITY FOR BEL COMPUTATIONS BASED ON MODEL POINTS?

Conclusion

La méthode d'approximation polynomiale est une méthode efficace pour approcher la densité et la fonction de répartition d'une distribution composée, la probabilité de ruine ultime dans le modèle de ruine de Poisson composée ou encore la densité et la fonction de survie jointe d'une distribution composée bivariée. L'implémentation est facilitée par la présence des polynômes orthogonaux parmi les fonctions pré-définies dans les logiciels de calcul formel. L'obtention de l'approximation, en tout point, de la fonction nécessite simplement le calcul des coefficients de la représentation polynomiale définie par une combinaison linéaire de moments de la distribution. L'exécution de méthode est par conséquent rapide, et concurrence les autres méthodes numériques en termes de précision sous réserve de choisir un ordre de troncature suffisamment élevé. Une perspective de ce travail est l'application de la méthode d'approximation polynomiale à d'autres problèmes en actuariat mais aussi pourquoi pas liés à d'autres champs d'application des probabilités. Une autre perspective intéressante réside dans l'étude des performances de l'estimateur de la densité de probabilité qui résulte de la formule d'approximation.

La procédure de construction des *model points* fait partie intégrante du processus de valorisation des provisions *best estimate* depuis la clôture de 2013 chez AXA France et a été diffusée auprès des autres entités du groupe AXA.

Résumé

Cette thèse a pour objet d'étude les méthodes numériques d'approximation de la densité de probabilité associée à des variables aléatoires admettant des distributions composées. Ces variables aléatoires sont couramment utilisées en actuariat pour modéliser le risque supporté par un portefeuille de contrats. En théorie de la ruine, la probabilité de ruine ultime dans le modèle de Poisson composé est égale à la fonction de survie d'une distribution géométrique composée. La méthode numérique proposée consiste en une projection orthogonale de la densité sur une base de polynômes orthogonaux. Ces polynômes sont orthogonaux par rapport à une mesure de probabilité de référence appartenant aux Familles Exponentielles Naturelles Quadratiques. La méthode d'approximation polynomiale est comparée à d'autres méthodes d'approximation de la densité basées sur les moments et la transformée de Laplace de la distribution. L'extension de la méthode en dimension supérieure à 1 est présentée, ainsi que l'obtention d'un estimateur de la densité à partir de la formule d'approximation. Cette thèse comprend aussi la description d'une méthode d'agrégation adaptée aux portefeuilles de contrats d'assurance vie de type épargne individuelle. La procédure d'agrégation conduit à la construction de *model points* pour permettre l'évaluation des provisions *best estimate* dans des temps raisonnables et conformément à la directive européenne Solvabilité II.

Mots-clé : Distributions composées, théorie de la ruine, Familles Exponentielles Naturelles Quadratiques, polynômes orthogonaux, méthodes numériques d'approximation, Solvabilité II, provision *best estimate*, *model points*.

Abstract

This PhD thesis studies numerical methods to approximate the probability density function of random variables governed by compound distributions. These random variables are useful in actuarial science to model the risk of a portfolio of contracts. In ruin theory, the probability of ultimate ruin within the compound Poisson ruin model is the survival function of a geometric compound distribution. The proposed method consists in a projection of the probability density function onto an orthogonal polynomial system. These polynomials are orthogonal with respect to a probability measure that belongs to Natural Exponential Families with Quadratic Variance Function. The polynomial approximation is compared to other numerical methods that recover the probability density function from the knowledge of the moments or the Laplace transform of the distribution. The polynomial method is then extended in a multidimensional setting, along with the probability density estimator derived from the approximation formula.

An aggregation procedure adapted to life insurance portfolios is also described. The method aims at building a portfolio of model points in order to compute the best estimate liabilities in a timely manner and in a way that is compliant with the European directive Solvency II.

Keywords: compound distributions, ruin theory, Natural Exponential Families with Quadratic Variance Function, orthogonal polynomials, numerical methods, Solvency II, best estimate liabilities, model points.