

# وزارة التعليم العالي والبحث العلمي

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

BADJI MOKHTAR-ANNABA

UNIVERSITY

UNIVERSITE BADJI MOKHTAR

ANNABA



جامعة باجي مختار

- عنابة -

Faculté des Sciences

Département de Mathématiques

Année : 2024/2025



## THÈSE

Présentée en vue de l'obtention du diplôme de Doctorat

### Modèles de provisionnement en assurance non vie

Filière

Mathématiques Appliquées

Spécialité

Actuariat

Par

BECHIRI Sarra

DIRECTEUR DE THÈSE: REMITA Mohamed Riad

Prof. ENSIA - Alger

Devant le jury

PRESIDENT: ZEGHDOUDI Halim

Prof. U.B.M. - Annaba

EXAMINATEUR : HADDARI Allaeddine

MCA. U. - Batna 2

EXAMINATEUR: METIRI Farouk

MCA. U.B.M. - Annaba

EXAMINATEUR: EZZEBSA Abdelali

MCA. U. - Guelma

# Dédicace

À mes parents, pur leur amour inconditionnel, leur soutien indéfectible et leurs encouragements constants tout au long de ce parcours académique. Leur sagesse et leur patience ont été ma source d'inspiration et de persévérance.

À mon fils "Yahia", pour ta présence lumineuse dans les jours sombres et ton souffle de vie dans mon aventure intellectuelle. Et pour ton amour qui a porté chacun de mes pas.

À mon partenaire, pour sa compréhension, sa patience et son soutien tout au long des longues nuits et des moments difficiles. Merci de croire en moi, même quand je doutais.

À ma sœur et à mon frère, pour votre soutien et votre présence, sources de force et de réconfort.

À mes professeurs et mentors, pour leur guidance, leurs conseils avisés et leur dévouement. Leur passion pour la connaissance et leur engagement envers l'excellence m'ont poussé à donner le meilleur de moi-même.

À mes amis, pour leur soutien moral, leurs encouragements.

Enfin, à tous ceux qui ont contribué, de près ou de loin, à l'accomplissement de ce travail.

# Remerciements

Je souhaite exprimer ma profonde gratitude à toutes les personnes qui ont contribué, de près ou de loin, à l'aboutissement de cette thèse de doctorat.

Tout d'abord, je tiens à remercier mon directeur de thèse, le Professeur **REMITA Mohamed Riad**, pour son encadrement, sa patience et ses précieux conseils. Son expertise et son soutien constant ont été déterminants pour la réalisation de ce travail.

Je remercie également les membres du jury, Mr **ZEGHDOUDI Halim**, Mr **HADDARI Allaeddine**, Mr **METIRI Farouk** et Mr **EZZEBSA Abdelali** pour avoir accepté d'évaluer cette thèse et pour leurs commentaires constructifs qui ont permis d'améliorer ce manuscrit.

Je souhaite exprimer ma reconnaissance à mes collègues et amis du laboratoire **LaPS**, pour leur soutien amical et leurs échanges stimulants. En particulier, je remercie Mme **BELHAMRA Thara** pour ses discussions enrichissantes et son aide précieuse tout au long de ce parcours.

Un grand merci à ma famille, pour leur amour inconditionnel, leur soutien moral et leur patience. À mes parents, merci de m'avoir inculqué les valeurs de travail et de persévérance qui m'ont guidé jusqu'ici. À mes amis, pour leur présence réconfortante et leurs encouragements, merci d'avoir toujours cru en moi.

Enfin, je remercie toutes les personnes, institutions et organismes qui, par leur soutien financier ou logistique, ont contribué à la réalisation de cette thèse.

Cette thèse est le fruit d'un travail collectif et c'est avec une profonde gratitude que je remercie tous ceux qui ont rendu cette aventure possible.

# Résumé

Dans le domaine de l'assurance non vie, les provisions pour sinistres à payer sont généralement les plus importantes en termes de coût et demandent donc une évaluation d'une grande précision. La modélisation des montants se repose sur deux types de distributions distinctes -à queue lourde et à queue légère- qui sont associées à deux catégories de risque.

Cette thèse propose une nouvelle distribution nommée la distribution mixte Dagum-Exponentielle, cette dernière permet d'obtenir des caractéristiques importantes telles que la fonction quantile, le coefficient d'asymétrie, le coefficient d'aplatissement, la fonction de fiabilité ainsi que la fonction de hasard. De ailleurs, les paramètres du modèle sont estimés via la méthode du maximum de vraisemblance, garantissant ainsi une précision optimale. Enfin l'application du modèle sur un ensemble de données réelles met en évidence l'efficacité et la signification du nouveau modèle.

**Mots clés** : Distribution mixte Dagum-Exponentielle, Maximum de vraisemblance, Provision technique, Simulation.

# Abstract

In the field of non-life insurance, the provisions for claims payable are generally the most important in terms of cost and therefore require a very precise assessment. The modelling of amounts is based on two distinct types of distributions - heavy-tailed and light-tailed- which are associated with two categories of risk.

This thesis proposes a new distribution called the Dagum-Exponential mixed distribution, which allows to obtain important characteristics such as quantile function, asymmetry coefficient, flattening coefficient, the reliability function as well as the hazard function. Moreover, the parameters of the model are estimated via the maximum likelihood method, thus ensuring optimal accuracy. Finally, the application of the model to a real data set highlights the effectiveness and significance of the new model.

**Keywords** : Dagum-Exponential distribution, Maximum likelihood estimation, Simulation, Technical provisions.

## مُلخّص

في مجال التأمين على غير الحياة، تعتبر أحكام المطالبات المستحقة عمومًا الأكثر أهمية من حيث التكلفة وبالتالي تتطلب تقييمًا دقيقًا للغاية. تعتمد نمذجة المبالغ على نوعين متميزين من التوزيعات - الثقيلة والخفيفة - والتي ترتبط بفتتين من المخاطر.

تقترح هذه الأطروحة توزيعًا جديدًا يسمى توزيع داغوم الآسي المختلط، والذي يسمح بالحصول على خصائص مهمة مثل دالة الكمية، ومعامل عدم التماثل، ومعامل التسطیح، ودالة الوثوقية، بالإضافة إلى دالة الفرصة. علاوة على ذلك، يتم تقدير معلمات النموذج عبر طريقة الاحتمالية القصوى، وبالتالي ضمان الدقة المثلى. وأخيرًا، عبر طريقة الاحتمالية القصوى، وبالتالي ضمان الدقة المثلى. وأخيرًا، فإن تطبيق النموذج على مجموعة بيانات حقيقية يسلط الضوء على فعالية وأهمية النموذج الجديد.

**الكلمات المفتاحية :** الإمكان الأعظم، التوزيع المختلط داغوم - آسي، المحاكاة، المخصصات الفنية.

# Table des matières

Dédicace	II
Remerciements	III
Résumé	IV
Abstract	V
Introduction	IX
<b>1 Les modèles de provisionnement en assurance non-vie</b>	<b>1</b>
1.1 Points généraux concernant l'assurance . . . . .	1
1.1.1 Types d'assurance . . . . .	2
1.1.2 Composantes de l'assurance . . . . .	4
1.1.3 Principes de base . . . . .	5
1.1.4 Fonctionnement de l'assurance . . . . .	5
1.2 Les modèles de provisionnement . . . . .	7
1.2.1 Provisions techniques . . . . .	7
1.2.2 Problème de provisionnement . . . . .	7
1.2.3 Les modèles déterministes . . . . .	10
1.2.4 Les modèles stochastiques . . . . .	13
<b>2 Modélisation des risques</b>	<b>17</b>
2.1 Identification des risques . . . . .	18
2.2 Analyse des risques . . . . .	18
2.2.1 Fréquence de sinistre . . . . .	19
2.2.2 Coût de sinistre . . . . .	22
2.3 Evaluation du risque . . . . .	27
2.4 Traitement du risque . . . . .	28

<b>3</b>	<b>Le nouveau modèle Dagum-Exponentielle</b>	<b>29</b>
3.1	Introduction . . . . .	29
3.2	La distribution mixte Dagum-Exponentielle . . . . .	30
3.2.1	Les différentes fonctions du modèle . . . . .	31
3.2.2	Propriétés statistiques de la distribution . . . . .	34
3.2.3	Fonction quantile . . . . .	38
3.3	Estimation . . . . .	39
3.3.1	Estimations des paramètres avec la méthode maximum de vraisemblance . . . . .	39
3.3.2	Estimation du paramètre contrainte-résistance . . . . .	40
3.4	Mesures actuarielles . . . . .	42
3.5	Application . . . . .	44
3.5.1	Étude de simulation . . . . .	44
3.5.2	Application réelle . . . . .	46
	<b>Conclusion</b>	<b>47</b>
	<b>Bibliographie</b>	<b>48</b>

# Introduction

Dans un monde où les incertitudes économiques, environnementales et sociales sont de plus en plus marquées, la protection contre les risques devient essentielle. L'assurance non-vie occupe une place centrale dans ce dispositif. Elle vise à couvrir les assurés contre les conséquences financières d'événements imprévus n'ayant pas de lien direct avec la durée de vie humaine, tels que les accidents, les catastrophes naturelles, les incendies ou encore les dommages causés à autrui.

Au cœur de l'efficacité de ce système se trouve la gestion des risques, une démarche stratégique qui permet d'identifier, d'évaluer, de prévenir et de transférer les aléas potentiels. Cette gestion rigoureuse permet à l'assureur de proposer une couverture adaptée aux besoins de ses clients tout en maîtrisant sa propre exposition aux sinistres.

Par ailleurs, pour assurer sa solidité financière et garantir le paiement des indemnités, l'assureur doit procéder à un provisionnement précis. Ce mécanisme consiste à constituer des réserves techniques destinées à faire face aux engagements futurs, que ce soit pour les sinistres déjà survenus mais non encore réglés, ou pour ceux encore inconnus.

Le provisionnement est donc un pilier essentiel de la viabilité économique et de la solvabilité des compagnies d'assurance. Ainsi, l'assurance non-vie ne se limite pas à indemniser des dommages : elle repose sur une organisation complexe, combinant gestion proactive des risques et rigueur financière, afin de garantir sécurité et stabilité à l'ensemble des parties prenantes.

L'assurance non-vie s'impose comme un instrument fondamental de protection contre les risques du quotidien. Sa performance repose sur une gestion proactive et structurée des risques, appuyée par un provisionnement rigoureux. Ces mécanismes assurent non seulement la sécurité des assurés, mais garantissent aussi la stabilité financière des compagnies d'assurance face aux incertitudes de demain.

L'objectif de cette thèse est d'établir un nouveau modèle pour modéliser les risques plus précisément le coût de sinistre, et le comparer avec d'autres modèles. Parmi ces modèles Dagum introduit en 2017 par Dey et al [12] et le modèle Log-Normal introduit en 2002 par Antoniou et al [5]. Aussi pour étudier l'influence de

ce nouveau modèle sur les provisions techniques.

Notre modèle mixte Dagum-Exponentielle est une distribution issue de la composition de deux lois, utilisée principalement pour modéliser des données économiques, actuarielle ou financière. Dans ce modèle, le paramètre de taux de l'Exponentielle est lui-même une variable aléatoire suivant une loi Dagum, ce qui permet d'introduire une hétérogénéité dans les sinistres modélisés. Dagum-Exponentielle offre un cadre robuste pour l'analyse actuarielle des sinistres, notamment dans les domaines de la tarification, de l'évaluation des risques et du calcul des provisions techniques.

La structure de cette thèse est comme suit :

Cette thèse est composée de trois chapitres

Dans le premier chapitre, On commence par une généralité sur l'assurance et les méthodes de provisionnement en assurance non vie.

Dans le deuxième chapitre, on concentre sur la modélisation des risques, plus précisément sur le coût de sinistre.

Dans le dernier chapitre, on termine par notre nouveau modèle Dagum-Exponentielle et l'application avec des données réelles.

# Chapitre 1

## Les modèles de provisionnement en assurance non-vie

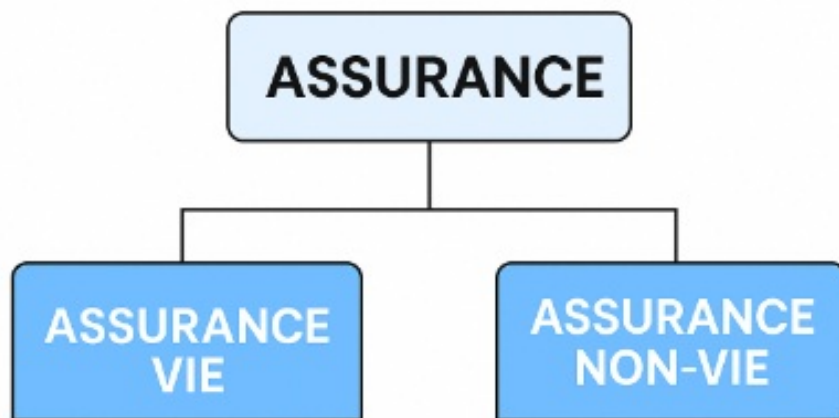
### 1.1 Points généraux concernant l'assurance

#### Définition de l'assurance :

L'assurance est un contrat par lequel une personne (l'assuré) paie une prime à une compagnie d'assurance (l'assureur) en échange de la promesse de cette dernière de l'indemniser ou de l'aider financièrement en cas de survenance d'un risque spécifique. Ce risque peut être un dommage, une perte ou un événement particulier, comme un accident, une maladie, un décès ou des dégâts matériels.

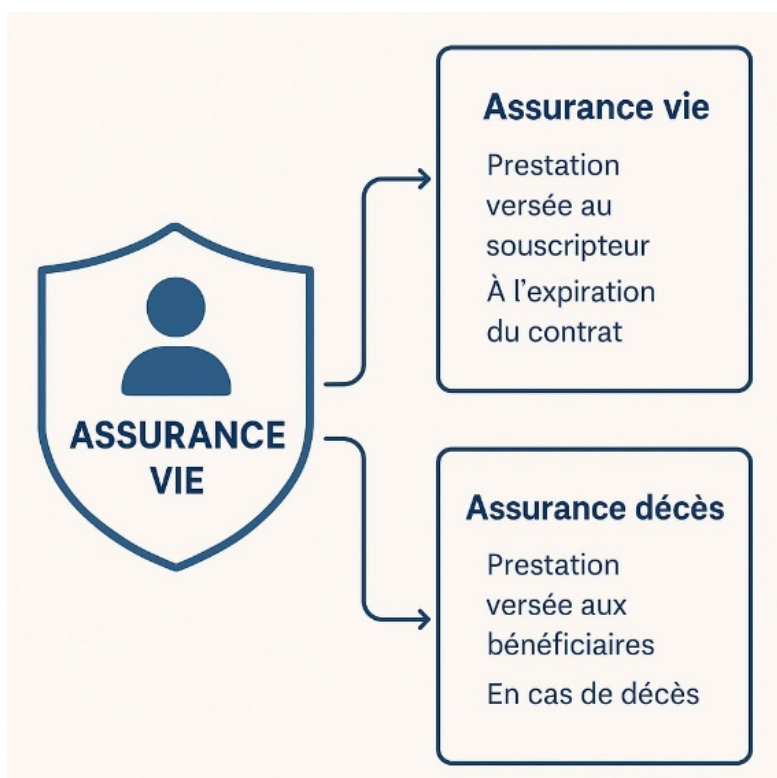
L'assurance est un secteur clé dans la gestion des risques, offrant une protection financière contre divers types de pertes ou de dommages. Voici quelques points généraux à considérer :

### 1.1.1 Types d'assurance



**Figure 01** : Assurance vie et non vie.

**Assurance vie** : L'assurance vie est un contrat par lequel l'assureur s'engage à verser un capital ou une rente à un bénéficiaire désigné en cas de décès de l'assuré ou à l'assuré lui-même en cas de survie à une date déterminée. Elle est principalement utilisée pour protéger financièrement les proches de l'assuré ou pour constituer une épargne.



**Figure 02 :** Assurance vie.

**Assurance non vie :** L'assurance non vie, aussi appelée assurance dommages ou assurance générale, couvre les risques autres que ceux liés à la vie de l'assuré. Elle inclut la protection contre les pertes ou dommages matériels, les responsabilités civiles, et autres risques financiers non liés au décès.

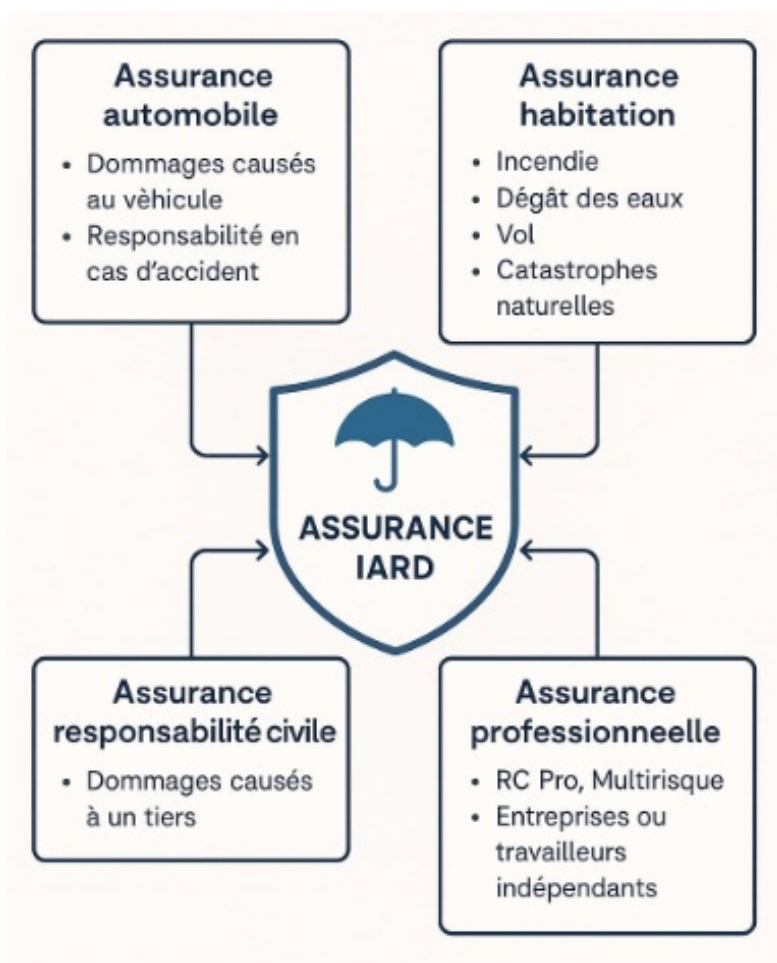


Figure 03 : Assurance non vie.

### 1.1.2 Composantes de l'assurance

**Contrat d'assurance (Police d'assurance) :** Document légal qui précise les termes, conditions, couvertures, exclusions, et obligations des deux parties (assuré et assureur).

**Prime d'assurance :** Montant payé périodiquement (mensuellement, annuellement, etc.) par l'assuré à l'assureur pour maintenir la couverture.

**Franchise :** Montant que l'assuré doit payer avant que l'assurance ne commence à couvrir les coûts.

**Indemnisation :** Somme versée par l'assureur à l'assuré ou au bénéficiaire en cas de sinistre, conformément aux termes du contrat.

**Bénéficiaire :** Personne ou entité désignée pour recevoir le bénéfice de l'assurance en cas de sinistre.

### 1.1.3 Principes de base

**Mutualisation des risques :** Les primes des assurés sont regroupées pour couvrir les pertes de quelques-uns.

**Prime d'assurance :** Montant payé par l'assuré pour obtenir la couverture.

**Franchise :** Montant que l'assuré doit payer avant que l'assurance ne prenne en charge.

**Indemnisation :** Paiement effectué par l'assureur pour couvrir une perte couverte par la police.

### 1.1.4 Fonctionnement de l'assurance

**Évaluation des risques :** L'assureur évalue les risques associés à l'assuré pour déterminer la prime adéquate. Cette évaluation repose sur des facteurs comme l'âge, l'état de santé, le type de bien assuré, et l'historique des sinistres.

**Versement de la prime :** L'assuré paie une prime régulière pour maintenir la couverture d'assurance active.

**Survenance d'un sinistre :** En cas de survenance d'un risque couvert, l'assuré doit notifier l'assureur et soumettre une réclamation.

**Réclamation et indemnisation :** Après vérification, l'assureur verse une indemnisation à l'assuré ou au bénéficiaire pour couvrir les pertes conformément aux termes de la police.

#### \* Considérations importantes

**Conditions de la police :** Comprendre les termes et les conditions, y compris les exclusions.

**Montant de la couverture :** S'assurer que la couverture est suffisante pour protéger contre les risques spécifiques.

**Historique des réclamations :** L'historique des réclamations peut affecter les primes futures.

**Comparaison des offres :** Comparer différentes offres pour obtenir la meilleure couverture au meilleur prix.

**\* Régulation et conformité**

**Réglementation :** Les assurances sont régulées par des organismes gouvernementaux pour protéger les consommateurs.

**Solvabilité :** Les compagnies d'assurance doivent maintenir une certaine solvabilité pour garantir qu'elles peuvent payer les réclamations.

**Transparence :** Obligation pour les assureurs de divulguer clairement les termes des polices et les frais.

**\* Objectif de l'assurance**

L'assurance vise à :  
Protéger contre les conséquences financières imprévues d'événements défavorables.

**Mutualiser les risques :** En regroupant les contributions de nombreux assurés, les assureurs peuvent répartir les risques et fournir une aide financière en cas de sinistres.

**Apporter la tranquillité d'esprit :** Permettre aux individus et aux entreprises de fonctionner avec moins d'incertitude et de peur des pertes financières majeures.

En résumé, l'assurance est un outil financier crucial pour gérer les risques, offrir une protection économique et garantir une certaine stabilité financière en cas de sinistres imprévus.

## 1.2 Les modèles de provisionnement

### 1.2.1 Provisions techniques

#### Définition

Les provisions techniques, également appelées réserves techniques, sont des montants que les compagnies d'assurances mettent de côté pour couvrir leurs engagements futurs envers les assurés.

Ces provisions sont essentielles pour garantir que l'assureur dispose des fonds nécessaires pour payer les sinistres et autres obligations financières.

#### Catégories de provisions techniques

**1- Provision pour sinistre à payer :** montants destinés à couvrir les sinistres qui ont déjà eu lieu mais qui n'ont pas encore été réglés ou entièrement payés.

**2- Provision pour risque en cours :** réserves correspondant à la part des primes payées qui couvrent des périodes futures.

**3- Provision pour primes non acquises :** montants correspondant à la part des primes perçues qui n'est pas encore acquise, c'est-à-dire la part des primes qui couvre la période restante des contrats d'assurance.

### 1.2.2 Problème de provisionnement

#### Triangle de liquidation

Le triangle de liquidation est un outil utilisé en actuariat et en gestion des provisions techniques pour les assurances, afin d'estimer les provisions nécessaires pour faire face aux sinistres futurs. Il permet de visualiser et d'analyser les paiements de sinistres au fil du temps, en regroupant ces paiements selon l'année d'occurrence du sinistre et l'année de règlement.

Voici une explication des éléments clés d'un triangle de liquidation en provisionnement :

**1- Année d'occurrence :** C'est l'année où le sinistre a eu lieu. Chaque ligne du triangle représente une année d'occurrence.

**2- Année de développement :** C'est l'année suivant l'année d'occurrence, durant laquelle les paiements sont effectués. Chaque colonne du triangle représente une année de développement. Par exemple, la colonne "0" contient les paiements effectués l'année même du sinistre, la colonne "1" contient les paiements effectués la première année après l'occurrence du sinistre, et ainsi de suite.

**3- Montants payés :** Les cellules du triangle contiennent les montants cumulés des paiements effectués pour les sinistres survenus durant l'année d'occurrence correspondante et payés durant l'année de développement correspondante.

**4- Facteurs de développement :** Ces facteurs sont calculés pour estimer les paiements futurs en se basant sur les paiements passés. Ils représentent le ratio des paiements cumulés d'une année de développement à l'autre. Ces facteurs sont utilisés pour projeter les paiements futurs et estimer les provisions nécessaires.

**Exemple de triangle de liquidation ou Représentation des données**  
Imaginons un triangle simple avec 3 années d'occurrence et jusqu'à 3 années de développement :

Année d'occurrence	0 (Année de survenance)	1 (Année suivante)	2 (Deux ans après)
2021	23000	28000	33000
2022	25000	30000	
2023	27000		

- Pour les sinistres survenus en 2021, 23000 ont été payées la même année, 28000 au total après un an, et 33000 au total après deux ans.

- Pour les sinistres survenus en 2022, 25000 ont été payées la même année, et 30000 au total après un an.

- Pour les sinistres survenus en 2023, 27000 ont été payées la même année.

### Calcul des facteurs de développement

Les facteurs de développement sont calculés comme suit :

- Facteur entre année 0 et année 1 =  $\frac{\text{Total année 1}}{\text{Total année 0}}$

- Facteur entre année 1 et année 2 =  $\frac{\text{Total année 2}}{\text{Total année 1}}$

Utilisons les données ci-dessus pour les calculer :

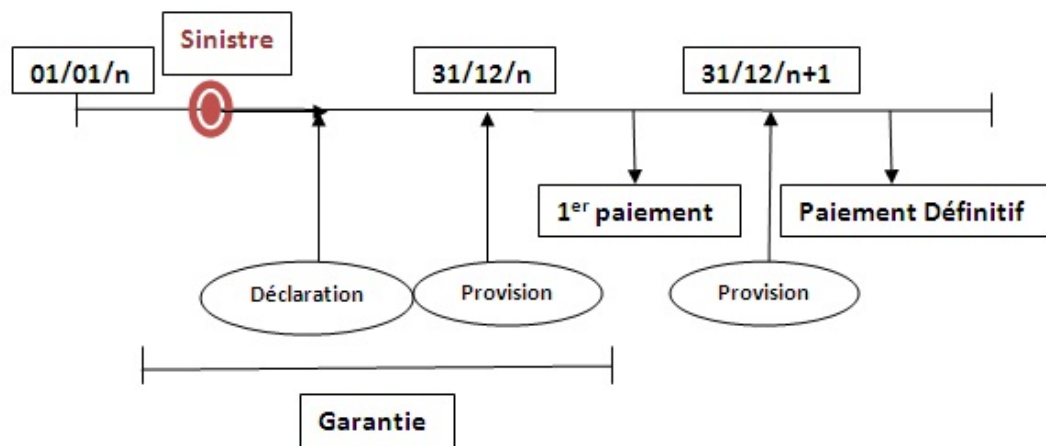
- Facteur 0-1 :  $\frac{28000}{23000} = 1.22$

- Facteur 1-2 :  $\frac{33000}{28000} = 1.18$

Ces facteurs peuvent ensuite être appliqués aux paiements cumulés pour projeter les paiements futurs et estimer les provisions nécessaires.

En utilisant ces facteurs, les actuaires peuvent estimer les montants des paiements futurs pour chaque année d'occurrence et ainsi déterminer les provisions à constituer pour couvrir les sinistres non encore totalement réglés.

Le triangle de liquidation est un outil puissant pour analyser les paiements de sinistres et estimer les provisions nécessaires. Il aide les assureurs à assurer la solvabilité et à gérer les risques de manière efficace.



**Figure 04 :** Provision et paiement d'un sinistre.

Les modèles de provisionnement en assurance non vie sont des outils et des méthodes utilisés pour estimer les montants des provisions techniques que les compagnies d'assurance doivent constituer afin de couvrir les engagements futurs liés aux sinistres survenus. Ces modèles permettent de garantir la solvabilité des compagnies d'assurance et de protéger les assurés. Voici un aperçu des principaux modèles de provisionnement utilisés en assurance non vie :

### 1.2.3 Les modèles déterministes

Les modèles déterministes de provisionnement sont des méthodes qui utilisent des données historiques et des règles fixes pour estimer les provisions techniques.

Voici quelques exemples de ces modèles :

#### Méthode de Chain-Ladder

C'est une technique actuarielle utilisée principalement pour estimer les provisions techniques, en particulier les réserves pour sinistres non réglés dans le domaine de l'assurance.

La méthode de Chain Ladder est donc une technique puissante pour estimer les réserves de sinistres, en s'appuyant sur l'historique des données de développement.

#### Hypothèses de la méthode :

$H_1$  : Nous supposons que les années de survenance sont indépendantes c-à-d  $C_{ij}$  et  $C_{kj}$  sont indépendants si  $i \neq k$   
Cette hypothèse est en général vérifiée dans la plupart des domaines d'assurance.

$$H_2 : \forall i = 1, 2, \dots, n, \forall j = 1, 2, \dots, n - 1 \\ E(C_{i,j+1}/C_{i1}, \dots, C_{ij}) = \lambda_i C_{ij}$$

#### Schéma du raisonnement

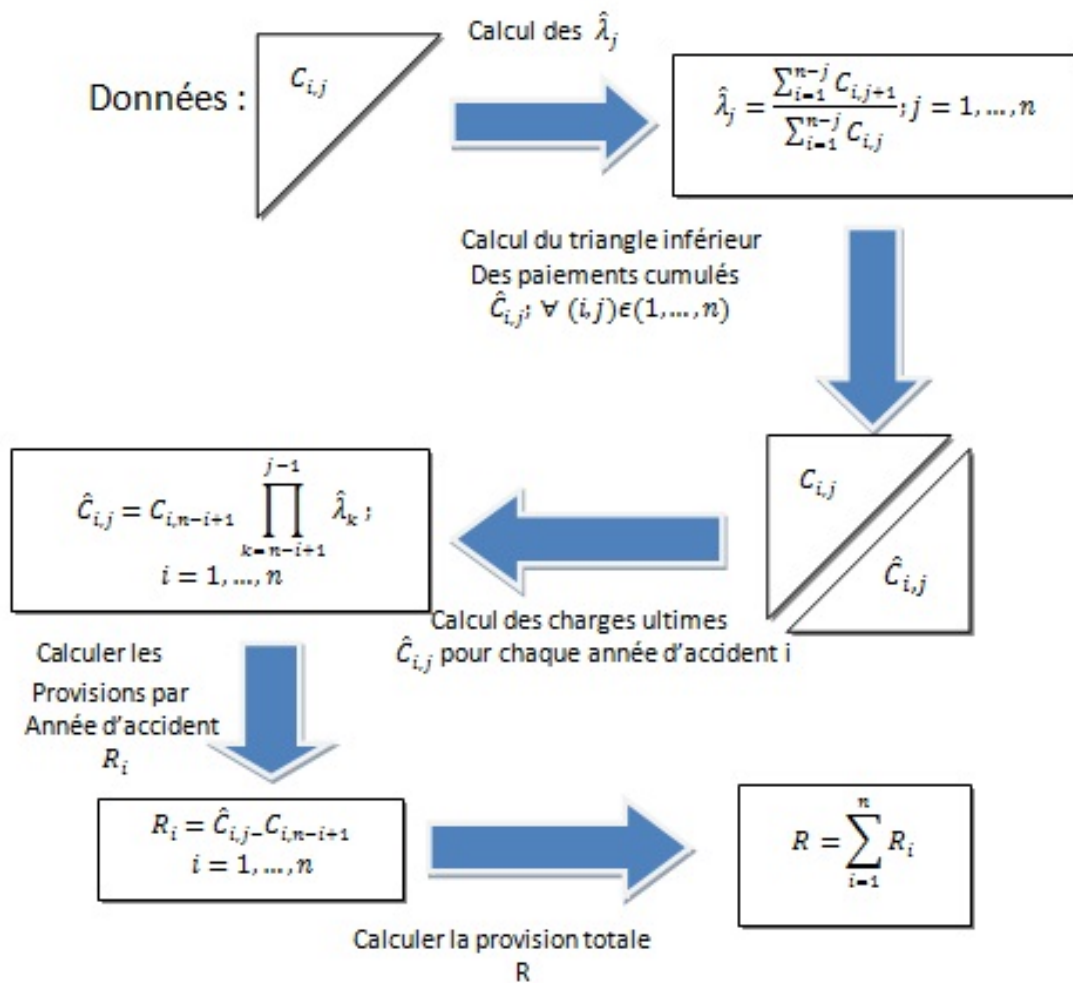


Figure 05 : Méthode de Chain Ladder.

### La variante de Chain-Ladder

La variante de Chain-Ladder, souvent appelée "London Chain", combine à la fois des aspects de la méthode Chain-Ladder traditionnelle et des techniques statistiques plus avancées pour améliorer la précision des estimations de provisions techniques en assurance non vie.

La variante de Chain-Ladder est l'une des fonctions appelées autorégressives, qui reposent sur une hypothèse moins contraignante que Chain-Ladder.

Cette variante offre généralement des estimations plus précises en tenant compte des aspects statistiques des données historiques, mais elle peut nécessiter des compétences et des ressources supplémentaires pour sa mise en oeuvre par rapport à la

méthode Chain-Ladder traditionnelle.

### Hypothèses de la méthode :

On reprend les hypothèses de la méthode de Chain-Ladder.  
La dynamique des  $(C_{i,j})$  est supposée être de la forme suivante :

$$C_{i,j+1} = \lambda_j C_{i,j} + \alpha_j \quad \forall i \in ]0, n - j - 1[$$

### Schéma du raisonnement

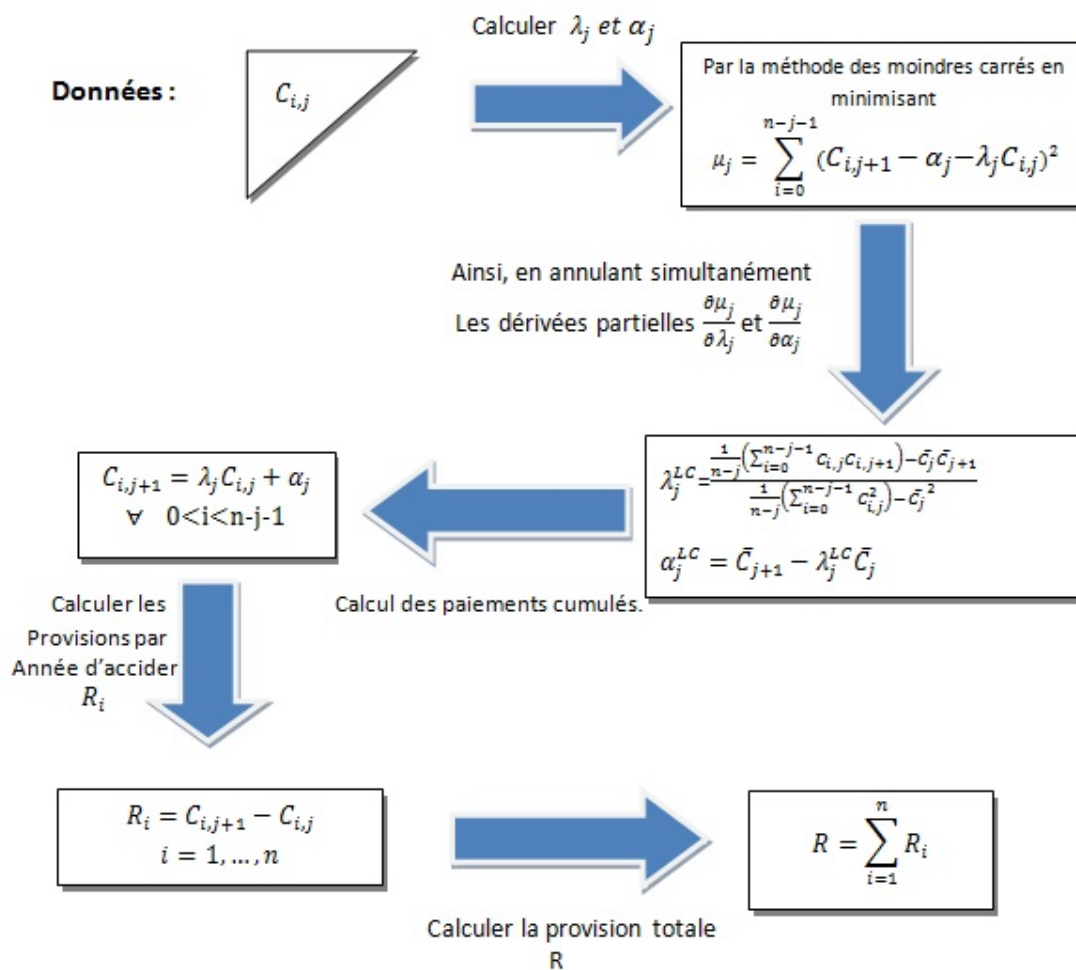


Figure 06 : Méthode de Variante Chain Ladder.

## 1.2.4 Les modèles stochastiques

Les méthodes stochastiques de provisionnement en assurance non vie sont des approches qui utilisent des modèles probabilistes pour estimer les réserves nécessaires afin de couvrir les sinistres futurs.

### Modèle de Mack

Le modèle de Mack est une généralisation du modèle Chain-Ladder. il permet de calculer non seulement les provisions nécessaires mais aussi l'estimation de l'incertitude associée à ces provisions.

Ce modèle est basé sur des hypothèses moins restrictives pour estimer les facteurs de développement,

**H1 :** Les années de survenance des sinistres sont indépendantes les unes des autres,  $C_{ij}$  et  $C_{kj}$  sont indépendants si  $i \neq k$

**H2 :**  $\forall i = 1, 2, \dots, n, \forall j = 1, 2, \dots, n - 1$   
 $E(C_{i,j+1}/C_{i1}, \dots, C_{ij}) = \lambda_i C_{ij}$  C-à-d L'espérance conditionnelle est sans biais

**H3 :**  $\forall i = 1, 2, \dots, n, \forall j = 1, 2, \dots, n - 1, \exists \tau_j$  tel que :  
 $Var(C_{i,j+1}/C_{i1}, \dots, C_{ij}) = \tau_j^2 C_{ij}$

### Schéma du raisonnement

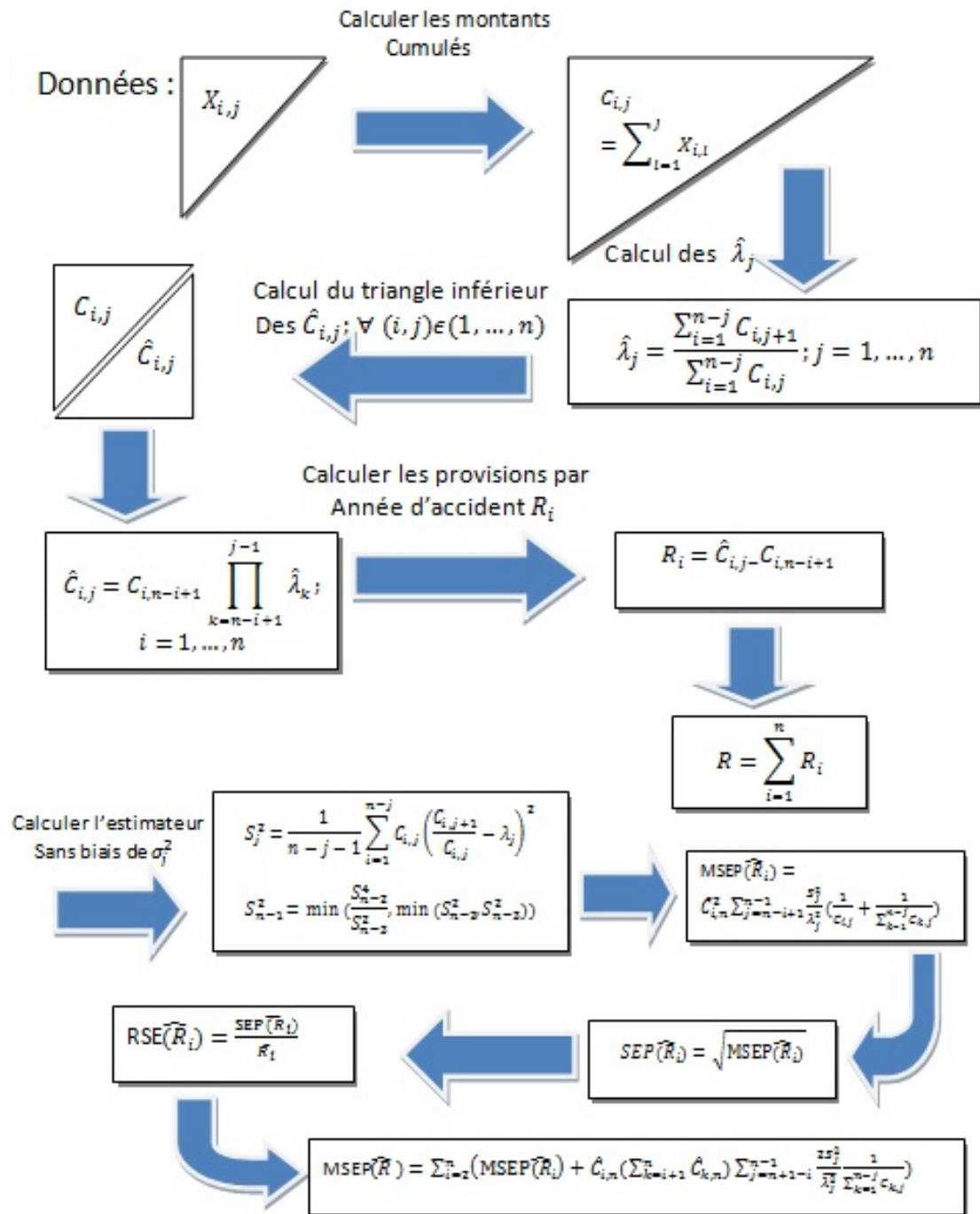


Figure 07 : Modèle de Mack.

Le modèle de Mack est apprécié pour sa capacité à prendre en compte à la fois les aspects déterministes et stochastiques des sinistres, offrant ainsi des estimations

plus robustes des provisions techniques.

### **Bootstrap variante chain ladder**

La méthode Bootstrap variante CL est une technique actuarielle utilisée pour estimer la distribution de la réserve des sinistres dans le cadre du développement des sinistres. Elle combine :

- Le modèle Chain-Ladder traditionnel, qui permet d'estimer les sinistres futurs à partir des données historiques de triangles de développement.
- Avec la méthode Bootstrap, qui permet d'évaluer l'incertitude de ces estimations en générant des simulations aléatoires.

### **Schéma du raisonnement**

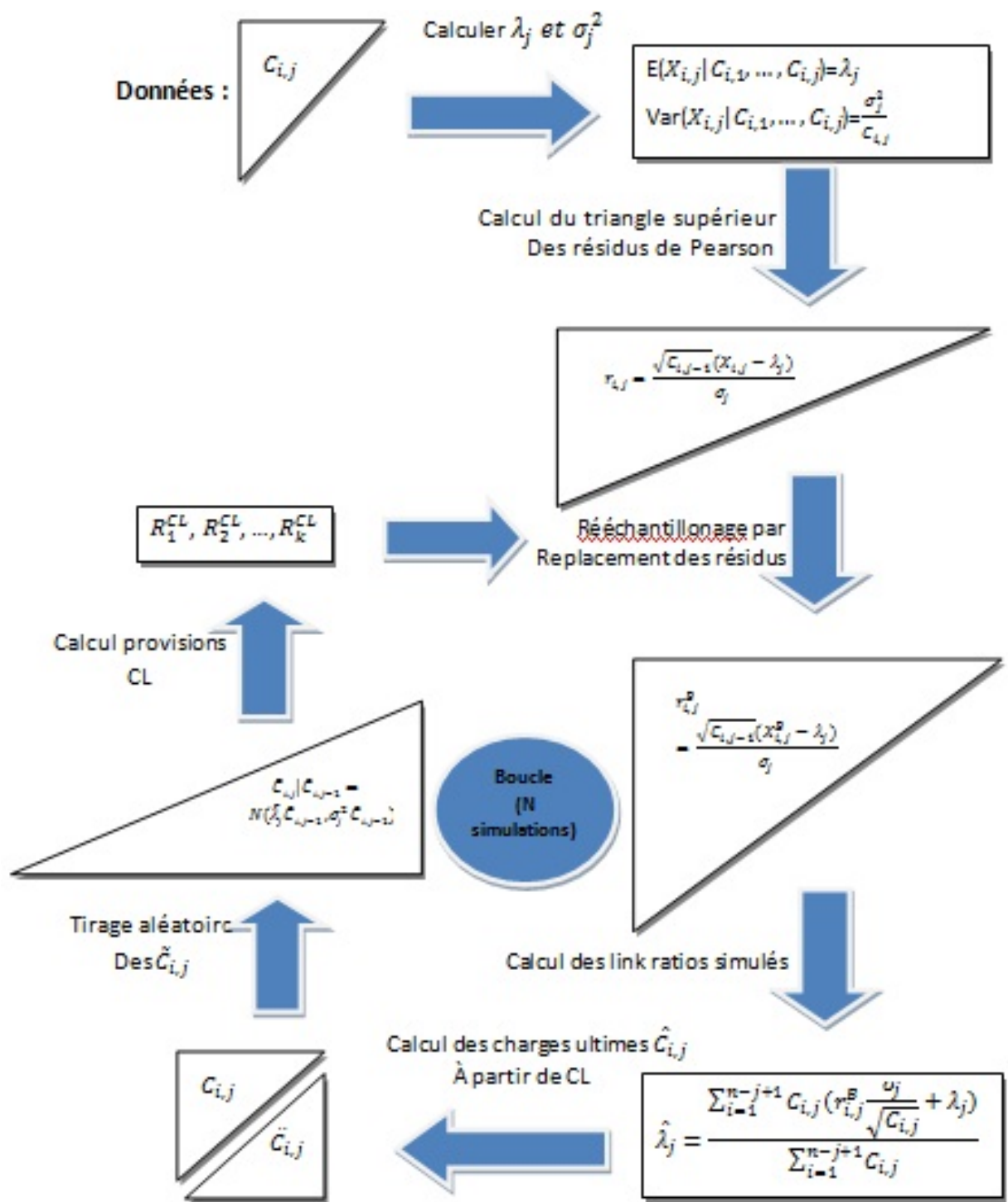


Figure 08 : Méthode Bootstrap variante Chain Ladder.

# Chapitre 2

## Modélisation des risques

### Définition d'un sinistre :

Un sinistre fait référence à un événement prévu et couvert par le contrat d'assurance, entraînant une perte ou un dommage matériel, et pour lequel une indemnisation est demandée à l'assureur. Cela peut inclure des accidents, des catastrophes naturelles, etc.

La **modélisation des risques** est un processus qui consiste à identifier, analyser et quantifier les risques potentiels auxquels une organisation, un projet ou un système peut être exposé. Cela permet d'anticiper les impacts négatifs et de mettre en place des stratégies de gestion adaptées.

Voici les grandes étapes de la modélisation des risques :

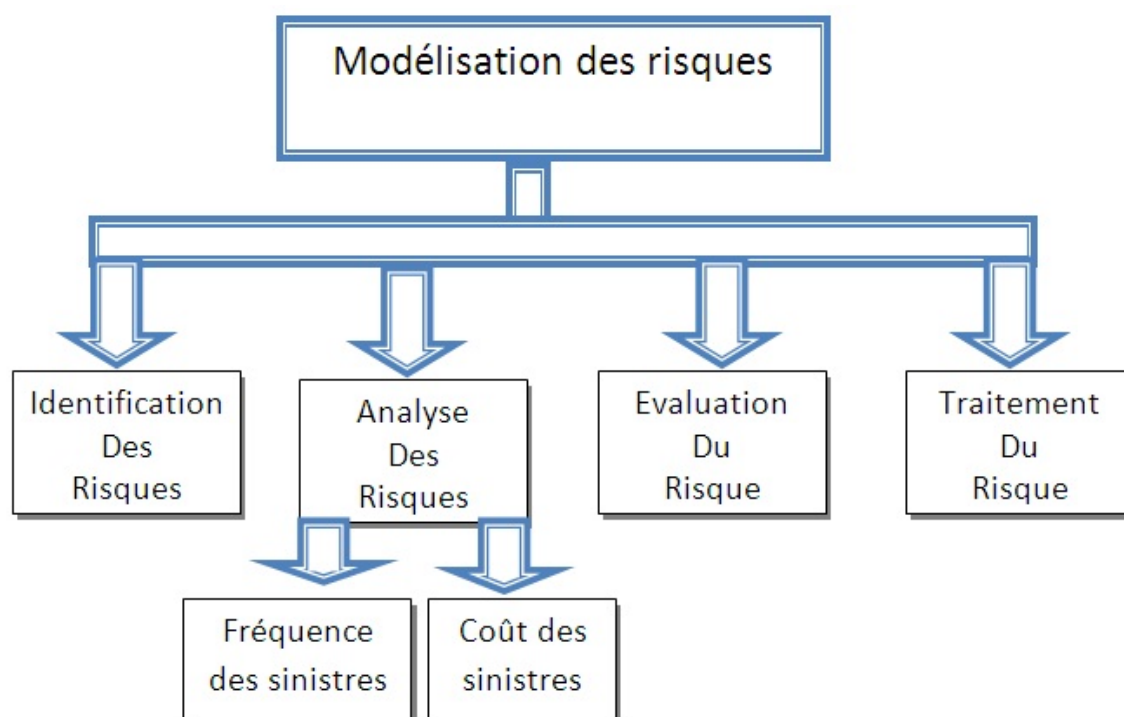


Figure 09 : Modélisation des risques.

## 2.1 Identification des risques

Est une étape fondamentale pour anticiper les sources potentielles de sinistres et adapter les stratégies de tarification, de provisionnement et de gestion du capital. C'est un processus itérative et dynamique, nécessitant une mise à jour régulière pour intégrer les évolutions du marché, de l'environnement et des comportements assurés. Une approche combinant **data science** et **expertise actuarielle** est indispensable pour une couverture optimale.

## 2.2 Analyse des risques

L'analyse des risques est le processus qui consiste à identifier, évaluer, quantifier et classer les risques liées aux contrats d'assurance couvrant les biens, la responsabilité civile, les accidents, les catastrophes naturelles ou les événements imprévus, à l'exclusion des risques liés à la vie humaine (Assurance vie).

Elle permet à l'assureur de :

- Comprendre la nature des risques souscrits.

- Estimer la probabilité de sinistres et leurs coûts potentiels.
- Définir des primes adéquates pour couvrir ces risques.
- Constituer des réserves suffisantes pour faire face aux engagements futurs.
- Respecter les exigences de solvabilité.
- Préparer des stratégies de réassurance et de diversification.

## 2.2.1 Fréquence de sinistre

### Définition

La fréquence des sinistres en assurance non-vie fait référence au nombre de réclamations ou de sinistres signalés par les assurés au cours d'une période donnée, généralement exprimée en pourcentage. Cette mesure est importante pour les compagnies d'assurance afin d'évaluer les risques et de fixer les primes. Elle peut varier en fonction de divers facteurs tels que le type de couverture, la région géographique et les caractéristiques des assurés.

Les fréquences de sinistres sont un indicateur clé en assurance, car elles permettent d'évaluer et de maîtriser les risques. Leur importance se manifeste à plusieurs niveaux :

**1. Évaluation des risques :** La fréquence des sinistres mesure le niveau de risque associé à un portefeuille d'assurance ou à une catégorie spécifique de contrats. Elle permet aux assureurs de :

- Identifier les profils, biens ou zones à risque élevé.
- Comprendre les tendances et l'évolution des sinistres.

### 2. Calcul des primes d'assurance :

La fréquence des sinistres est un facteur déterminant pour fixer les primes :  
 Une fréquence élevée dans une catégorie de risque entraîne des primes plus élevées.

Une fréquence faible permet de proposer des primes compétitives pour attirer les clients.

### 3. Rentabilité de l'assureur :

Contrôle des coûts : Les assureurs surveillent les fréquences pour éviter qu'elles n'excèdent les prévisions, ce qui pourrait affecter leur rentabilité.

Gestion de la sinistralité : Une fréquence bien maîtrisée réduit la charge des indemnisations et améliore le ratio combiné (rapport entre les charges et les revenus).

### 4. Segmentation et tarification différenciée :

L'analyse des fréquences permet d'adopter une tarification différenciée en fonction des caractéristiques des assurés, comme :

L'âge ou l'expérience du conducteur.

La localisation géographique.

L'usage du bien assuré (usage personnel ou professionnel).

### **5. Prévention et réduction des risques :**

Les fréquences permettent d'identifier les causes récurrentes des sinistres et de mettre en place des mesures préventives :

Sensibilisation des assurés (campagnes de sécurité routière, incitations à l'entretien des biens).

Promotion d'équipements ou services pour réduire les risques (alarme antivol, détecteurs de fumée, systèmes d'assistance à la conduite).

### **6. Stratégie commerciale et fidélisation :**

Ciblage des segments : Les assureurs peuvent concentrer leurs efforts commerciaux sur des segments à faible fréquence pour limiter leur exposition.

Incitations à la fidélité : Les assurés avec un faible taux de sinistralité peuvent bénéficier de réductions ou d'avantages, comme des bonus.

### **7. Indicateur de performance sectorielle :**

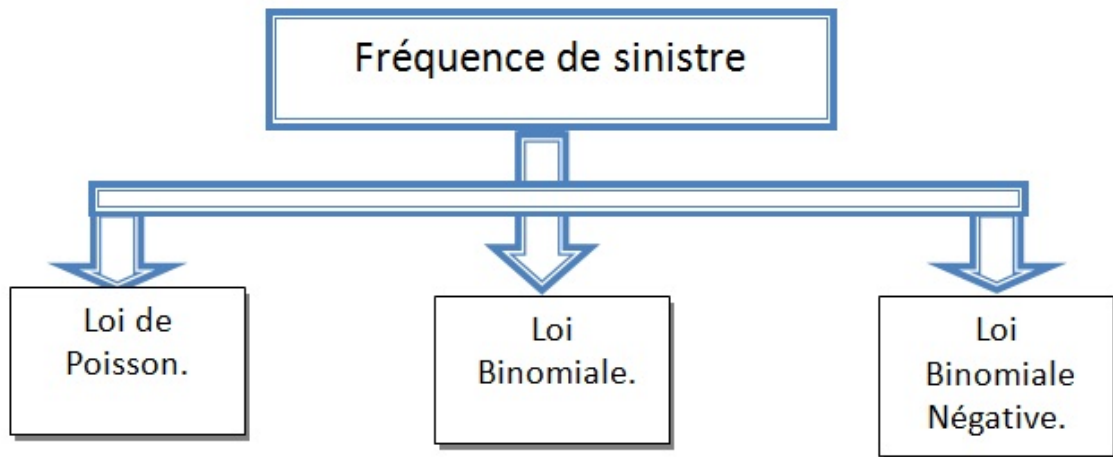
Dans le secteur de l'assurance, la fréquence des sinistres est un indicateur clé utilisé pour :

- Comparer la performance entre compagnies.
- Évaluer l'impact de nouvelles réglementations ou tendances, comme l'impact des véhicules autonomes sur les sinistres autos.

Dans le domaine de l'actuariat, on utilise principalement les lois de Poisson, Binomiale, Binomiale-Négative pour modéliser la variable aléatoire de fréquence. Ces 3 lois font partie de la famille de Panjer, qui se distingue par la présence d'une relation de récurrence spécifique entre leurs masses de probabilité.

La relation de récurrence :

$$\exists x < 1, y \in R, \forall l \in N^*, P_N(l) = (x + \frac{y}{l})P_N(l-1)$$



**Figure 10 :** Fréquence de sinistre.

### Loi de Poisson

$N \sim P(\Theta)$ ,  $\Theta > 0$  avec

$$P_N(l) = \frac{e^{-\Theta} \Theta^l}{l!}, \quad l = 0, 1, 2, \dots$$

### Loi Binomiale

$N \sim \text{Bin}(m, p)$ ,  $0 < p < 1$ ,  $m \in \mathbb{N}^*$  avec

$$P_N(l) = \binom{m}{l} p^l q^{m-l}, \quad l = 1, 2, \dots, m.$$

Si  $m = 1 \Rightarrow N \sim \text{Bernoulli}(p)$ .

### Loi Binomiale Négative

$N \sim \text{Bin-Neg}(\beta, p)$ ,  $0 < p < 1$ ,  $\beta > 0$  avec

$$P_N(l) = \frac{\Gamma(\beta+l)}{\Gamma(l+1)\Gamma(\beta)} p^l q^\beta, \quad l = 1, 2, \dots$$

Si  $\beta = 1 \Rightarrow N \sim \text{Géométrique}(p)$ .

La fréquence des sinistres est un levier stratégique pour les assureurs. Elle joue un rôle central dans la gestion des risques, la fixation des primes, la maîtrise

des coûts et l'amélioration des services proposés. Une analyse régulière et précise de cet indicateur permet de rester compétitif tout en assurant la pérennité financière.

## 2.2.2 Coût de sinistre

### Définition

Les montants de sinistres se réfèrent aux coûts financiers associés aux réclamations d'assurance suite à des incidents ou des événements couverts par une police d'assurance. Ces sinistres peuvent inclure des dommages matériels, des frais médicaux, des pertes d'exploitation, et d'autres types de pertes financières. Le montant d'un sinistre peut varier considérablement en fonction de la nature et de la gravité de l'incident, ainsi que des conditions spécifiques de la police d'assurance.

### Principaux éléments qui influencent les montants de sinistres :

**1. Type de sinistre :** Les sinistres peuvent être liés à des accidents de voiture, des dommages causés par des catastrophes naturelles, des incidents de responsabilité civile, des vols, etc. Chaque type de sinistre a des coûts spécifiques associés.

**2. Étendue des dommages :** Plus les dommages sont importants, plus le montant du sinistre sera élevé. Cela inclut la réparation ou le remplacement des biens endommagés, ainsi que les frais médicaux et autres coûts indirects.

**3. Conditions de la police d'assurance :** La couverture offerte par la police d'assurance, les franchises, les plafonds de couverture et les exclusions influencent directement le montant du sinistre.

**4. Frais de gestion de sinistres :** Les coûts administratifs associés à l'évaluation et au règlement des sinistres, y compris les frais juridiques et les honoraires des experts, peuvent également affecter le montant total.

**5. Inflation et coûts de reconstruction :** Les augmentations des coûts des matériaux et de la main-d'œuvre peuvent faire augmenter les montants des sinistres, en particulier pour les dommages matériels importants.

Les montants de sinistres jouent un rôle crucial dans le secteur des assurances pour plusieurs raisons :

**1. Évaluation des Risques :** Les assureurs utilisent les données sur les montants des sinistres pour évaluer les risques associés à différents types de polices

d'assurance. Des sinistres élevés peuvent indiquer des risques plus importants, influençant ainsi les primes d'assurance.

**2. Tarification :** Les montants de sinistres aident à déterminer les primes d'assurance. En analysant les sinistres passés, les assureurs peuvent estimer les coûts futurs et ajuster les tarifs en conséquence pour s'assurer qu'ils restent rentables tout en offrant une couverture adéquate.

**3. Réserves Techniques :** Les compagnies d'assurance doivent constituer des réserves pour payer les sinistres futurs. Les montants des sinistres influencent la taille de ces réserves, garantissant que l'assureur dispose des fonds nécessaires pour honorer les réclamations.

**4. Prévention et Gestion des Risques :** En identifiant les causes fréquentes et coûteuses des sinistres, les assureurs peuvent proposer des mesures de prévention des risques à leurs assurés. Cela peut inclure des recommandations de sécurité, des inspections régulières, et des programmes de sensibilisation.

**5. Réglementation et Conformité :** Les régulateurs exigent des informations détaillées sur les montants des sinistres pour s'assurer que les compagnies d'assurance maintiennent une stabilité financière et respectent les exigences de solvabilité.

**6. Évaluation de la Performance :** Les montants de sinistres sont un indicateur clé de la performance d'une compagnie d'assurance. Un montant de sinistres élevé peut signaler des problèmes dans la souscription des polices ou dans la gestion des risques, tandis qu'un montant bas peut indiquer une gestion efficace des risques et une sélection prudente des clients.

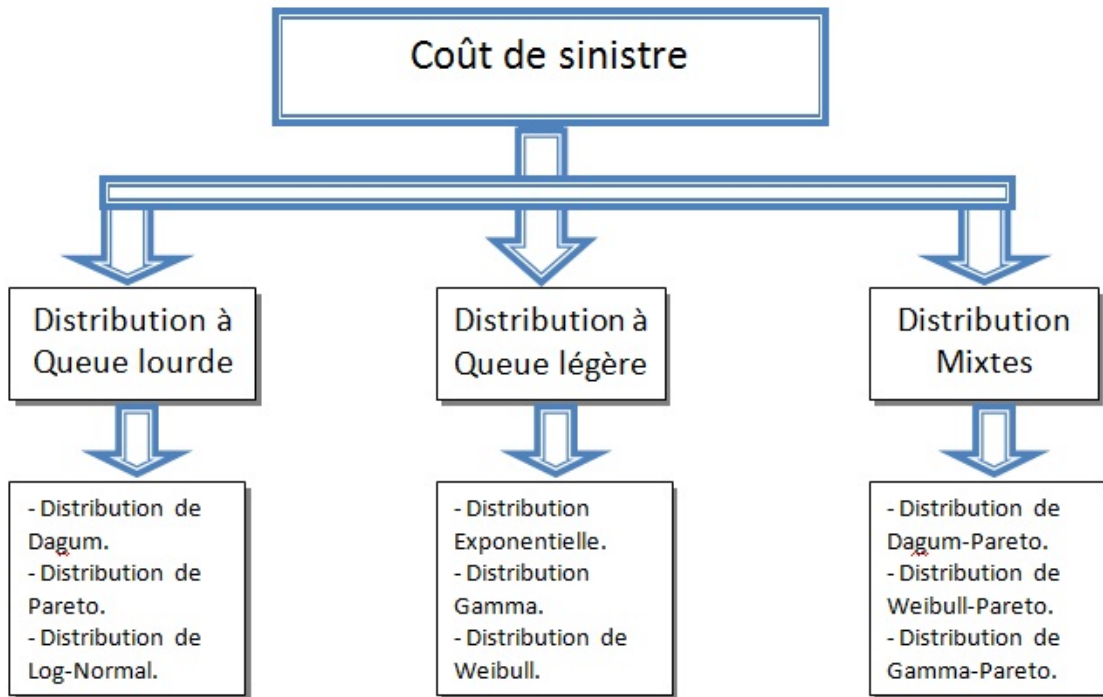


Figure 11 : Coût de sinistre.

### Distribution à queue lourde

Une distribution à queue lourde se réfère à une distribution de probabilité où les événements extrêmes ont une probabilité plus élevée que ce que prédit une distribution normale. Cela signifie que les pertes importantes, comme celles causées par des catastrophes naturelles ou des événements catastrophiques, peuvent survenir plus fréquemment que ce qui est traditionnellement anticipé. Les assureurs doivent prendre en compte cette caractéristique lors de la modélisation des risques et de la tarification des polices pour assurer leur solvabilité face à de tels événements.

On dit qu'une distribution à queue lourde si, pour tout  $i > 0$ ,  $j > 0$  :

$$\bar{G}_Y(y) > i \exp^{-jy} \Leftrightarrow \forall z > 0 M_Y(z) = \infty$$

$M_Y(z) = E(\exp^{yz})$ , représente la fonction génératrice des moments liée à la variable aléatoire  $Y$ .

#### 1- Distribution de Dagum

Pour tout  $\beta > 0, \phi > 0, \delta > 0$

$$g_Y(y) = \frac{\beta \delta \phi^\beta y^{\delta-1}}{(\phi + y^\delta)^{\beta+1}}$$

Et

$$G_Y(y) = 1 - \left(\frac{\lambda}{\lambda + y^\delta}\right)^\beta; y \in R^+$$

### 2- Distribution de Pareto

Pour tout  $\beta > 0, \phi > 0$

$$g_Y(y) = \frac{\beta}{\phi + y} \left(\frac{\phi}{\phi + y}\right)^\beta$$

Et

$$G_Y(y) = 1 - \left(\frac{\phi}{\phi + y}\right)^\beta$$

### 3- Distribution de Log-Normal

Pour tout  $\nu \in R, \tau > 0$

$$g_Y(y) = \frac{1}{\sqrt{2\pi\tau y}} \exp^{-\frac{(\ln y - \nu)^2}{2\tau^2}}$$

Et

$$G_Y(y) = \frac{1}{2} \left[ 1 + \operatorname{erf}\left(\frac{\ln y - \nu}{\sqrt{2}\tau}\right) \right]$$

Avec  $\operatorname{erf}(t) = \frac{2}{\sqrt{\pi}} \int_0^t e^{-u^2} du$ .

### Distribution à queue légère

Une distribution à queue légère est une loi de probabilité dont la queue de distribution (c'est à dire la probabilité des valeurs extrêmes) décroît rapidement lorsque les valeurs augmentent. Cela signifie que les événements rares et extrêmes ont une probabilité très faible, contrairement aux distribution à queue lourde, où les événements extrêmes sont plus probables.

On dit qu'une distribution à queue légère si,  $\exists i > 0, j > 0$ , telle que :

$$\overline{G}_Y(y) = 1 - G_Y(y) < i \exp^{-jy} \Leftrightarrow \exists z > 0 \text{ telle que } M_Y(z) < \infty$$

### 1- Distribution Exponentielle

Il existe  $\alpha > 0, y > 0$ , telle que :

$$g_Y(y) = \alpha \exp^{-\alpha y}$$

Et

$$G_Y(y) = 1 - \exp^{-\alpha y}$$

### 2- Distribution Gamma

Il existe  $\beta > 0, \delta > 0$ , telle que :

$$g_Y(y) = \frac{\delta^\beta}{\Gamma(\beta)} y^{\beta-1} \exp^{-\delta y}$$

Et

$$G_Y(y) = \frac{\gamma(\beta, \delta y)}{\Gamma(\beta)}$$

### 3- Distribution de Weibull

Il existe  $\delta > 0, \omega \geq 1$ , telle que :

$$g_Y(y) = \delta \omega y^{\omega-1} \exp^{-\delta y^\omega}$$

Et

$$G_Y(y) = 1 - \exp^{-\left(\frac{y}{\delta}\right)^\omega}$$

## Distribution mixte

Les distributions mixtes sont des modèles statistiques qui combinent plusieurs distributions pour représenter un phénomène complexe. En assurance, cela peut signifier combiner des distributions pour représenter à la fois la fréquence et la gravité des sinistres, ou même pour modéliser différents types de risques dans un portefeuille d'assurance. Ces modèles permettent de mieux comprendre la variabilité et les interactions entre différents aspects des données, ce qui est crucial pour la tarification des polices et la gestion des risques.

Parmi les distributions mixtes, on a :

### 1- Distribution de Dagum-Pareto

Pour tout,  $\beta > 0, \gamma > 0, \delta > 0$  et  $y > 0$

$$g(y) = \frac{\beta \gamma^\beta}{1 + \gamma} \left[ \frac{\gamma}{(\gamma + y)^{\beta+1}} + \frac{\delta y^{\delta-1}}{(\gamma + y^\delta)^{\beta+1}} \right]$$

Et

$$G_Y(y) = \frac{1}{1+\gamma} \left[ \gamma - 1 - \frac{\gamma^{\beta+1}}{(\gamma+y)^\beta} + \frac{\gamma^\beta}{(\gamma+y^\delta)^\beta} \right]$$

### 2- Distribution de Weibull-Pareto

Pour tout,  $l, w, \delta > 0$  et  $y > \lambda$

$$g(y) = \frac{lw}{\delta y} \left( \frac{l}{\delta} \log\left(\frac{y}{\lambda}\right) \right)^{w-1} \exp\left(-\left(\frac{l}{\delta} \log\left(\frac{y}{\lambda}\right)\right)^w\right)$$

Et

$$G_Y(y) = 1 - \exp\left(-\left(\frac{l}{\delta} \log\left(\frac{y}{\lambda}\right)\right)^w\right)$$

### 3- Distribution de Gamma-Pareto

Pour tout,  $\alpha, \beta, b, \lambda > 0$  et  $y > 0$

$$g(y) = \frac{\lambda^{\frac{b}{\alpha}}}{\left(\frac{\alpha}{b}\right)^\beta \Gamma(\beta)} (y + \lambda)^{-(1+\frac{b}{\alpha})} \left(\log\left(1 + \frac{y}{\lambda}\right)\right)^{\beta-1}$$

Et

$$G_Y(y) = 1 - \left[ y^{1+\frac{b}{\alpha}} \left(\frac{\alpha}{b}\right)^\beta \left(\Gamma(\beta - \gamma(\beta, (\frac{\alpha}{b})^{-1} \log(\frac{y}{\lambda})))\right) \right]; y > \lambda$$

En résumé, les montants de sinistres sont essentiels pour la viabilité et la gestion efficace des compagnies d'assurance, influençant tout, de la tarification et des réserves aux stratégies de gestion des risques et à la conformité réglementaire.

L'influence de la distribution du coût de sinistre sur les provisions techniques est un aspect fondamental en actuariat.

## 2.3 Evaluation du risque

L'évaluation du risque consiste à mesurer et quantifier l'impact financier des sinistres potentiels sur un portefeuille d'assurance. Elle permet de :

- Déterminer de primes adéquates (Tarification).
- Constituer des réserves suffisantes (Provisions).
- Calculer le capital nécessaire (Solvabilité 2).

L'évaluation du risque en assurance non vie repose sur :

- Des modèles statistiques solides (Fréquence-sévérité, extrêmes).

- Des outils modernes (Machine Learning, simulation).
  - Une approche dynamique (Adaptation aux nouveaux risques).
- Une bonne évaluation permet de minimiser les surprises et d'optimiser la rentabilité.

## 2.4 Traitement du risque

Le traitement du risque consiste à maîtriser l'impact financier des sinistres grâce à des stratégies adaptées. En assurance non vie, cela passe par :

- La réduction du risque (Prévention).
- Le transfert du risque (Réassurance, dérivés).
- La rétention contrôlée (Provisions, capital économique).

Une gestion optimale permet de protéger la rentabilité tout en garantissant la solvabilité.

# Chapitre 3

## Le nouveau modèle Dagum-Exponentielle

### 3.1 Introduction

En actuariat, il existe deux types de distributions liées à deux types de risques. Contrairement aux distributions à queue lourde (Log-normale, Pareto, etc.), définies par l'absence de définition de la fonction génératrice des moments, nous discutons ici des distributions à queue légère (Exponentielle, de Type phase, etc.).

La modélisation du montant des sinistres est l'un des principaux domaines de la théorie et de l'application actuarielle.

Il a été démontré que dans certaines circonstances, la combinaison de distributions plus étroitement adaptées aux réclamations donne lieu à des distributions différentes.

De nombreuses distributions contemporaines résultent de la combinaison de deux distributions distinctes. L'objectif principal de cette étude est de présenter une nouvelle distribution, issue de la fusion des distributions Exponentielle et de Dagum. Les autres distributions, nouvelles et classiques, sont représentées par ce nouveau modèle, à la fois comme cas général et comme cas particulier. Il fournit un cadre essentiel pour la modélisation actuarielle et biologique des données de durée de vie. Ce travail définit la distribution mixte Dagum-Exponentielle, une nouvelle distribution. Chacune de ces distributions possède une caractéristique unique qui la rend unique, et de nombreux chercheurs ont introduit ce processus.

La fonction de densité de probabilité (pdf) et la fonction de distribution cumulative (CDF) de la distribution Exponentielle sont les suivantes :

$$g(x) = \lambda e^{-\lambda y}; y > 0, \lambda > 0 \quad (3.1.1)$$

Et

$$G(x) = 1 - e^{-\lambda y}; y > 0, \lambda > 0 \quad (3.1.2)$$

Une variable aléatoire X est de la distribution Dagum avec trois paramètres comme suit :

$$g(x) = \frac{\beta \delta \lambda^\beta x^{\delta-1}}{(\lambda + x^\delta)^{\beta+1}}; y > 0, \beta > 0, \lambda > 0, \delta > 0 \quad (3.1.3)$$

Où  $\lambda > 0$  est le paramètre d'échelle,  $\beta$  et  $\delta$  sont des paramètres de forme positifs.

$$G(x) = 1 - \left( \frac{\lambda}{\lambda + x^\delta} \right)^\beta; y > 0, \beta > 0, \lambda, \delta > 0 \quad (3.1.4)$$

Créer une nouvelle distribution avec un comportement plus flexible que les distributions de base est l'objectif des distributions mixtes.

Cette étude présente la structure organisationnelle suivante. Dans la section 3.2 nous établissons la nouvelle distribution et nous examinons ces propriétés statistiques. L'approche du maximum de vraisemblance pour estimer les paramètres inconnus du modèle et l'estimation du paramètre contrainte-résistance sont abordées dans la section 3.3. Tandis que la section 3.4 fournit des mesures actuarielles et la section 3.5 propose des simulations ainsi qu'une application sur des données réelles.

## 3.2 La distribution mixte Dagum-Exponentielle

Soient  $g_1(y)$ ,  $g_2(y)$  les fonctions de densité (pdf) et  $G_1(y)$ ,  $G_2(y)$  les fonctions de répartition (CDF) de la loi Exponentielle et Dagum respectivement d'une variable aléatoire Y, alors la pdf et la CDF de la distribution proposée de la variable aléatoire Y sont respectivement de la forme :

$$g(y) = pg_1(y) + (1 - p)g_2(y) \quad (3.2.1)$$

Et

$$G(y) = pG_1(y) + (1 - p)G_2(y) \quad (3.2.2)$$

Où  $0 \leq p \leq 1$  tel que  $p = \frac{1}{1 + \lambda}$ , est nommé la proportion de mélange.

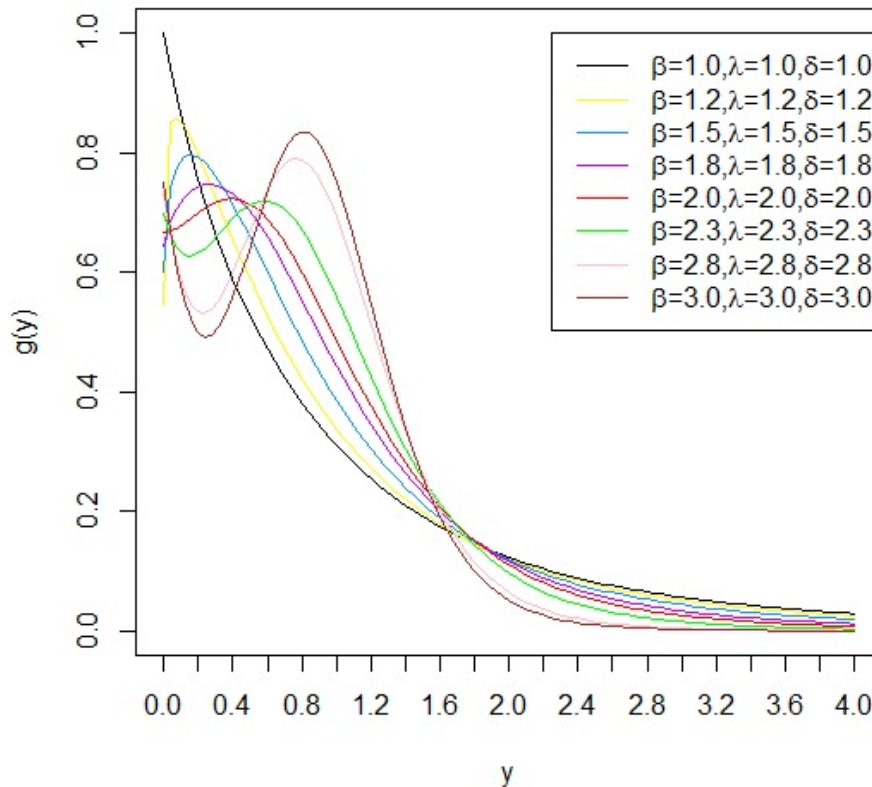
### 3.2.1 Les différentes fonctions du modèle

#### Fonction de densité

En insérant (3.1.1) et (3.1.3) dans (3.2.1), la pdf de la distribution mixte proposée est obtenue comme suit :

$$g(y, \beta, \lambda, \delta) = \frac{1}{1+\lambda} \lambda e^{-\lambda y} + \frac{\lambda}{1+\lambda} \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda + y^\delta)^{\beta+1}}$$
$$= \frac{\lambda}{1+\lambda} \left[ e^{-\lambda y} + \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda + y^\delta)^{\beta+1}} \right]; y > 0, \beta > 0, \lambda > 0, \delta > 0 \quad (3.2.3)$$

Nous traçons la fonction de densité de la distribution mixte Dagum-Exponentielle pour les valeurs de paramètres sélectionnées.



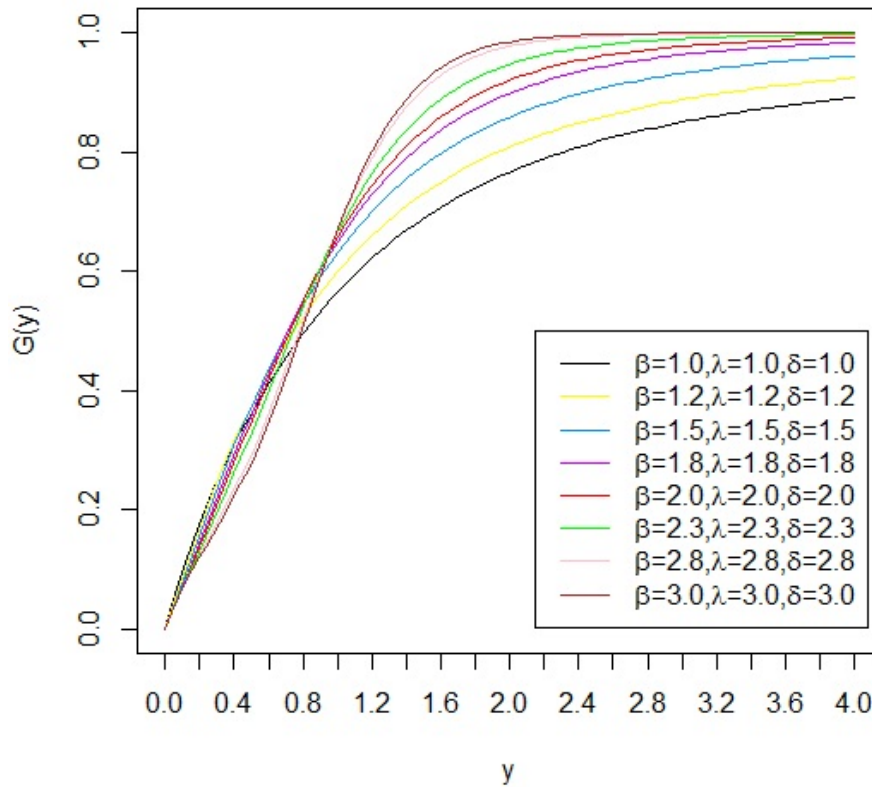
**Figure 01** : La fonction de densité de la distribution mixte Dagum-Exponentielle pour différentes valeurs de paramètres.

## Fonction de répartition

Et la fonction de répartition correspondante est obtenue par :

$$\begin{aligned}
 G(y, \beta, \lambda, \delta) &= \int_{-\infty}^y f(z, \beta, \lambda, \delta) dz = \int_0^y \frac{\lambda}{(1+\lambda)} \left[ e^{-\lambda z} + \frac{\beta \delta \lambda^\beta z^{\delta-1}}{(\lambda+z^\delta)^{\beta+1}} \right] dz \\
 &= \frac{\lambda}{(1+\lambda)} \left[ \int_0^x e^{-\lambda z} dz + \int_0^x \frac{\beta \delta \lambda^\beta z^{\delta-1}}{(\lambda+z^\delta)^{\beta+1}} dz \right] \\
 &= \frac{\lambda}{1+\lambda} \left[ \left. \frac{-e^{-\lambda z}}{\lambda} \right|_0^x - \left. \frac{\lambda^\beta}{(\lambda+z^\delta)^{\beta+1}} \right|_0^x \right] \\
 G(y, \beta, \lambda, \delta) &= \frac{1}{1+\lambda} \left[ 1 + \lambda - e^{-\lambda y} - \frac{\lambda^{\beta+1}}{(\lambda+y^\delta)^\beta} \right]; \\
 & \qquad \qquad \qquad y > 0, \beta > 0, \lambda > 0, \delta > 0 \qquad (3.2.4)
 \end{aligned}$$

Nous traçons la fonction de répartition cumulative (CDF) de la distribution Dagum-Exponentielle.



**Figure 02 :** La fonction de répartition cumulative de la distribution mixte Dagum-Exponentielle pour différentes valeurs de paramètres.

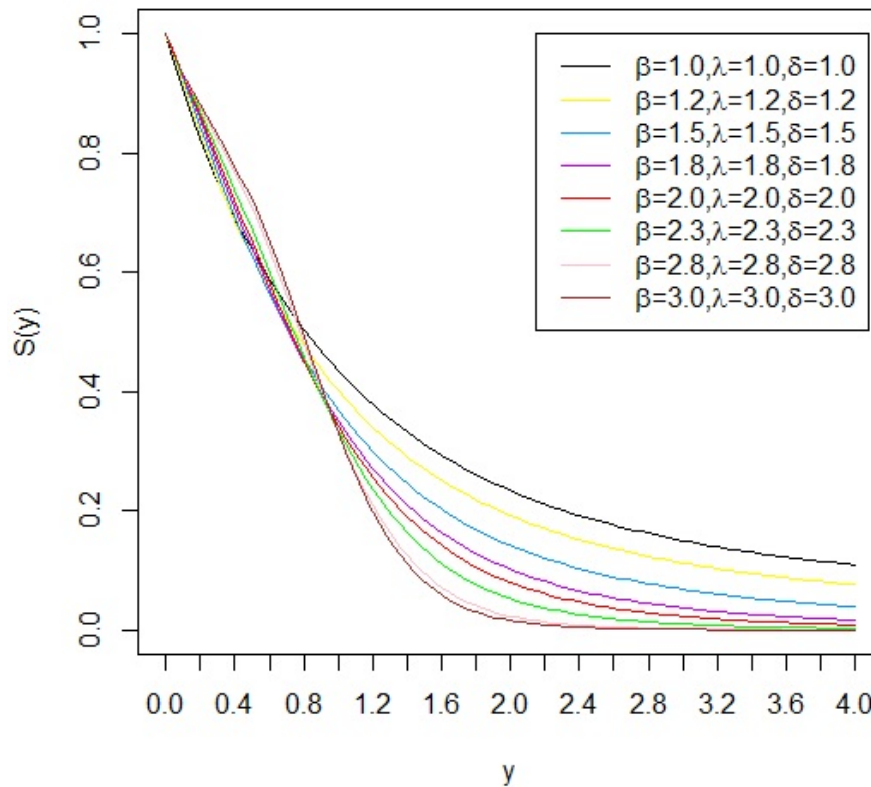
## Fonction de survie

L'expression mathématique de la fonction de fiabilité (survie) est :

$$S(x) = \overline{G(x)} = 1 - G(x) = 1 - \left[ \frac{1}{1+\lambda} (1 + \lambda - e^{-\lambda y} + \frac{\lambda^\beta + 1}{(\lambda + y^\delta)^\beta}) \right]$$

$$S(x) = \frac{1}{1+\lambda} \left[ e^{-\lambda y} - \frac{\lambda^\beta + 1}{(\lambda + y^\delta)^\beta} \right]; y > 0, \beta > 0, \lambda > 0, \delta > 0 \quad (3.2.5)$$

Nous traçons la fonction de fiabilité de la distribution Dagum-Exponentielle.



**Figure 03 :** La fonction de fiabilité de la distribution mixte Dagum-Exponentielle.

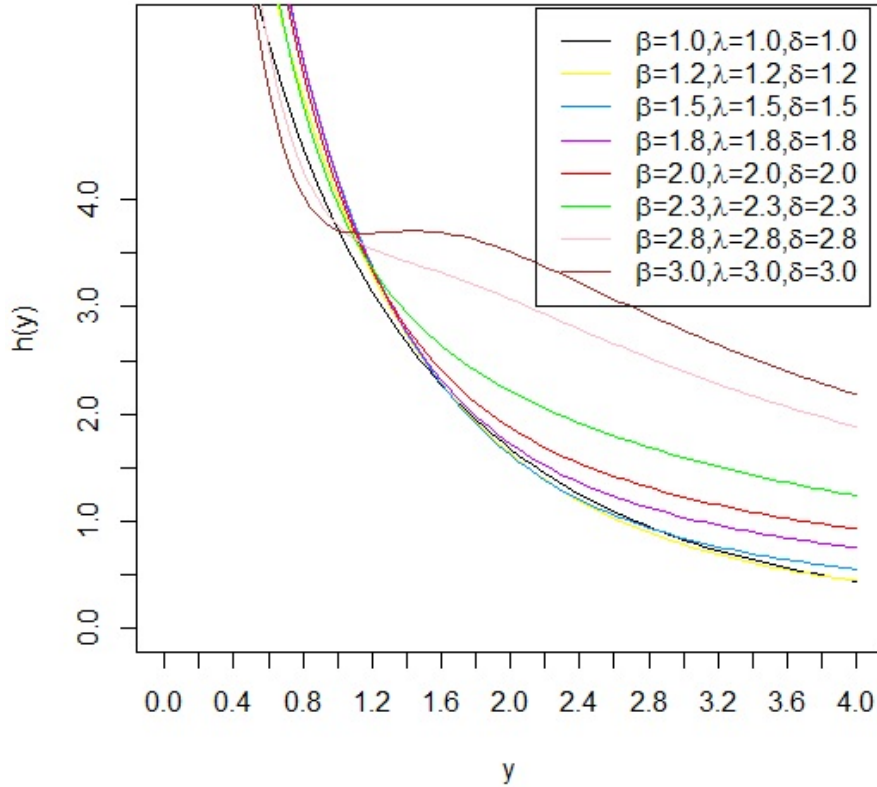
## Fonction de hasard

La fonction de hasard de la distribution mixte Dagum-Exponentielle est définie par :

$$h(y) = \frac{g(y)}{1 - G(y)} = \frac{g(y)}{S(y)} = \frac{\frac{\lambda}{1+\lambda} (e^{-\lambda y} + \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda + y^\delta)^{\beta+1}})}{\frac{1}{1+\lambda} (e^{-\lambda y} + \frac{\lambda^{\beta+1}}{(\lambda + y^\delta)^\beta})}$$

$$= \frac{\lambda}{1 + \lambda^{\beta+1}(\lambda + y^\delta)^{-\beta}} + \frac{\beta\delta y^{\delta-1}(\lambda + y^\delta)^{-1}}{1 + e^{-\lambda y}}; y > 0, \beta > 0, \lambda > 0, \delta > 0 \quad (3.2.6)$$

Nous traçons la fonction de hasard de la distribution Dagum-Exponentielle.



**Figure 04 :** La fonction de hasard de la distribution mixte Dagum-Exponentielle pour différentes valeurs des paramètres.

### 3.2.2 Propriétés statistiques de la distribution

#### $k^{eme}$ moments

Le  $k^{ème}$  moment brut d'une variable aléatoire continue Y, noté par  $\nu_k$ , est défini par :

$$\nu_l = E(Y^l) = \int_{-\infty}^{+\infty} y^l g(y) dy$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} y^l \frac{\lambda}{1+\lambda} \left[ e^{-\lambda y} + \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda + y^\delta)^{\beta+1}} \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \int_{-\infty}^{+\infty} y^l e^{-\lambda y} + \int_{-\infty}^{+\infty} \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda + y^\delta)^{\beta+1}} \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \frac{l!}{\lambda^{l+1}} + \frac{\lambda^{\frac{l}{\delta}}}{\Gamma(\beta)} \Gamma\left(1 + \frac{l}{\delta}\right) \Gamma\left(\beta - \frac{l}{\delta}\right) \right] \\
E(Y^l) &= \frac{1}{1+\lambda} \left[ \frac{l!}{\lambda^l} + \frac{\lambda^{\frac{l}{\delta}+1}}{\Gamma(\beta)} \Gamma\left(1 + \frac{l}{\delta}\right) \Gamma\left(\beta - \frac{l}{\delta}\right) \right] \tag{3.2.7}
\end{aligned}$$

Où  $\Gamma(\cdot)$  est la fonction gamma.  
Il devient facile d'obtenir les quatre premiers moments bruts de la distribution mixte Dagum-Exponentielle.

$$\begin{aligned}
\nu_1 = E(Y) &= \frac{1}{1+\lambda} \left[ \frac{1}{\lambda} + \frac{\lambda^{\frac{1}{\delta}+1}}{\Gamma(\beta)} \Gamma\left(1 + \frac{1}{\delta}\right) \Gamma\left(\beta - \frac{1}{\delta}\right) \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \frac{1}{\lambda^2} + \lambda^{\frac{1}{\delta}} \frac{\Gamma\left(1 + \frac{1}{\delta}\right) \Gamma\left(\beta - \frac{1}{\delta}\right)}{\Gamma(\beta)} \right] \tag{3.2.8}
\end{aligned}$$

$$\begin{aligned}
\nu_2 = E(Y^2) &= \frac{1}{1+\lambda} \left[ \frac{2!}{\lambda^2} + \frac{\lambda^{\frac{2}{\delta}+1}}{\Gamma(\beta)} \Gamma\left(1 + \frac{2}{\delta}\right) \Gamma\left(\beta - \frac{2}{\delta}\right) \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \frac{2}{\lambda^3} + \lambda^{\frac{2}{\delta}} \frac{\Gamma\left(1 + \frac{2}{\delta}\right) \Gamma\left(\beta - \frac{2}{\delta}\right)}{\Gamma(\beta)} \right] \tag{3.2.9}
\end{aligned}$$

$$\begin{aligned}
\nu_3 = E(Y^3) &= \frac{1}{1+\lambda} \left[ \frac{3!}{\lambda^3} + \frac{\lambda^{\frac{3}{\delta}+1}}{\Gamma(\beta)} \Gamma\left(1 + \frac{3}{\delta}\right) \Gamma\left(\beta - \frac{3}{\delta}\right) \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \frac{6}{\lambda^4} + \lambda^{\frac{3}{\delta}} \frac{\Gamma\left(1 + \frac{3}{\delta}\right) \Gamma\left(\beta - \frac{3}{\delta}\right)}{\Gamma(\beta)} \right] \tag{3.2.10}
\end{aligned}$$

$$\begin{aligned}
\nu_4 = E(Y^4) &= \frac{1}{1+\lambda} \left[ \frac{4!}{\lambda^4} + \frac{\lambda^{\frac{4}{\delta}+1}}{\Gamma(\beta)} \Gamma\left(1 + \frac{4}{\delta}\right) \Gamma\left(\beta - \frac{4}{\delta}\right) \right] \\
&= \frac{\lambda}{1+\lambda} \left[ \frac{24}{\lambda^5} + \lambda^{\frac{4}{\delta}} \frac{\Gamma\left(1 + \frac{4}{\delta}\right) \Gamma\left(\beta - \frac{4}{\delta}\right)}{\Gamma(\beta)} \right] \tag{3.2.11}
\end{aligned}$$

Aussi, le moment central est donné par l'expression :

$$\begin{aligned}
\nu'_l &= E(Y - \nu)^l = E\left\{ \sum_{n=1}^l \binom{l}{n} Y^{l-n} (-\nu)^n \right\} \\
&= \sum_{n=1}^l \binom{l}{n} (-1)^n \nu_{l-n} \nu^n
\end{aligned} \tag{3.2.12}$$

Les 2ème, 3ème et 4ème moments centraux peuvent être obtenus :

$$\nu'_2 = \nu_2 - \nu_1^2$$

$$\nu'_3 = \nu_3 - 3\nu_2\nu_1 + 2\nu_1^3$$

Et,

$$\nu'_4 = \nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4$$

### Asymétrie et aplatissement

L'asymétrie est définie par :

$$S_k = \frac{\nu'_3}{(\nu'_2)^{\frac{3}{2}}} = \frac{\nu_3 - 3\nu_2\nu_1 + 2\nu_1^3}{(\nu_2 - \nu_1^2)^{\frac{3}{2}}} \tag{3.2.13}$$

L'aplatissement est définie par :

$$K_s = \frac{\nu'_4}{(\nu'_2)^2} = \frac{\nu_4 - 4\nu_3\nu_1 + 6\nu_2\nu_1^2 - 3\nu_1^4}{(\nu_2 - \nu_1^2)^2} \tag{3.2.14}$$

Les valeurs numériques de la moyenne ( $\nu_1$ ), de la variance ( $\sigma^2$ ), du coefficient de variation ( $\epsilon$ ), de l'asymétrie ( $S_k$ ) et de l'aplatissement ( $K_s$ ) pour la distribution mixte Dagum-Exponentielle pour les valeurs sélectionnées de  $\beta, \lambda, \delta$  sont répertoriées dans le tableau 01.

$\beta$	$\lambda$	$\delta$	$\nu_1$	$\sigma^2$	$\epsilon$	$S_k$	$K_s$	
0.1	0.3	0.3	2.234545	12.11642	1.557751	1.862607	6.170821	
0.5			3.252231	6.583494	0.7889453	4.280486	1.262418	
1.2			3.666999	3.560506	0.5145707	12.2754	-31.28079	
1.7			2.149001	13.75354	1.725721	1.418681	5.194416	
3.3			2.583389	10.21779	1.237339	2.283026	6.134248	
0.3	0.1	0.3	9.063421	99.67259	1.101528	2.009764	1.014444	
		0.2	3.65872	28.28166	1.453528	1.765916	4.663778	
0.3	0.3	0.1	2.562834	10.5259	1.265928	2.142873	5.99421	
		0.4	3.279311	6.453348	0.7746579	4.402207	0.6918331	
		0.6	2.027931	10.26434	1.579837	3.049043	9.642333	
		0.9	3.295122	4.341839	0.6323614	9.89535	0.4820928	
		1.5	2.809848	4.807161	0.7802999	10.21272	16.97129	
0.5	0.5	0.5	-0.6666667	2.222222	2.236068	10.55424	65.69874	
			8	1.340262	3.533154	1.400374	2.312207	9.406382
			9	1.342278	3.537816	1.403372	2.309462	9.400535
0.5	0.5	3	1.691475	2.971228	1.019066	2.576232	6.124965	
		4	1.677818	2.924721	1.019289	2.624937	8.476268	
		5	1.672253	2.910544	1.020199	2.653701	8.658063	
		8	1.667181	2.898324	1.021152	2.682739	8.83027	
0.5	0.1	0.5	0.6885714	7.518212	3.982065	2.459821	9.173544	
			0.2	2.065641	12.58788	1.717598	1.882239	6.446879
			0.3	4.006667	25.57916	1.262292	1.957307	4.143562
			0.4	9.069091	99.56861	1.100265	2.012646	0.9995889

**Tableau 01** : Quantités  $(\nu_1)$ ,  $(\sigma^2)$ ,  $(\epsilon)$ ,  $(S_k)$  et  $(K_s)$  de la distribution mixte Dagum-Exponentielle.

### Interprétation :

- La moyenne et la variance diminuent lorsque  $\beta$  et  $\delta$  sont constants mais que  $\lambda$  varie, les mêmes résultats ont également été observés lorsque les valeurs de  $\beta$  et  $\lambda$  sont fixées mais que  $\delta$  varie.
- La moyenne et la variance diminuent lorsque  $\lambda$  et  $\delta$  sont constants mais que  $\beta$  varie, les mêmes résultats ont également été observés lorsque les valeurs de  $\beta$  et  $\delta$  sont fixées mais que  $\lambda$  varie.

En conclusion : une variance élevée indique que les données sont plus dispersées, tandis qu'une variance faible indique que les données sont plus concentrées autour de la moyenne.

- Coefficient de variation proche de 0 indique une dispersion très faible, les données sont très proches de la moyenne, et la variation est minimal.
- Coefficient de variation égal à 1 indique une dispersion égale à la moyenne, l'écart-

type et la moyenne sont égaux.

- Coefficient de variation supérieur à 1 indique une dispersion relativement forte, les données sont plus dispersées autour de la moyenne, et la variation est plus importante.

En conclusion : Plus la valeur du coefficient de variation est élevée, plus la dispersion autour de la moyenne est grande.

- L'asymétrie égale à 0 indique que la distribution est symétrique.

- L'asymétrie positive indique que la distribution est étalée vers la droite (queue plus longue à droite).

- L'asymétrie négative indique que la distribution est étalée vers la gauche (queue plus longue à gauche).

- Une distribution avec un aplatissement égal à 0 est considérée comme mesokurtique, et ses queues sont similaires à celles d'une distribution normale.

- Une distribution avec un aplatissement positif est considérée comme lépochurtique et a des queues plus légères que la distribution normale.

- Une distribution avec un aplatissement négatif est considérée comme platykurtique et a des queues plus lourdes que la distribution normale.

### 3.2.3 Fonction quantile

La fonction quantile d'une distribution de probabilité avec une fonction de répartition  $G(x)$  est définie par  $q = G^{-1}(y_q)$  où  $0 < q < 1$ . La fonction quantile de la distribution mixte Dagum-Exponentielle est définie par :

$$q = G^{-1}(y_q), 0 < q < 1$$

$$q = \frac{1}{1 + \lambda} \left( 1 - e^{-\lambda y_q} + \lambda - \frac{\lambda^{\beta+1}}{(\lambda + y_q^\delta)^\beta} \right)$$

$$q(1 + \lambda) = \frac{\lambda(\lambda + y_q^\delta)^\beta - \lambda^{\beta+1}}{(\lambda + y_q^\delta)^\beta} + T(y_q)$$

Où  $T(y)$  est la fonction de répartition de la distribution Exponentielle qui est l'une des densités mixte de la distribution mixte Dagum-Exponentielle.

$$q(1 + \lambda) = \lambda - \frac{\lambda^{\beta+1}}{(\lambda + y_q^\delta)^\beta} + T(y_q)$$

$$y_q = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{q(1 + \lambda)} + \lambda^{\beta-1} T^{-1}(y_q) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}}$$

Mais  $T^{-1}(y_q) = \frac{-\log(1-q)}{\lambda}$ , qui est la fonction quantile de la distribution Exponentielle. Alors, la fonction quantile de la distribution mixte Dagum-Exponentielle devient :

$$y_q = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{q(1+\lambda)} - \lambda^{\beta-2} \log(1-q) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}} \quad (3.2.15)$$

$$Q_q = G^{-1}(y_q)$$

Les 1er, 2ème et 3ème quartiles de la distribution mixte Dagum-Exponentielle sont donnés par :

$$Q_1 = G^{-1}\left(\frac{1}{4}\right) = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{\frac{1}{4}(1+\lambda)} - \lambda^{\beta-2} \log\left(\frac{3}{4}\right) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}}$$

$$Q_2 = G^{-1}\left(\frac{1}{2}\right) = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{\frac{1}{2}(1+\lambda)} - \lambda^{\beta-2} \log\left(\frac{1}{2}\right) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}}$$

$$Q_3 = G^{-1}\left(\frac{3}{4}\right) = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{\frac{3}{4}(1+\lambda)} - \lambda^{\beta-2} \log\left(\frac{1}{4}\right) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}}$$

### 3.3 Estimation

#### 3.3.1 Estimations des paramètres avec la méthode maximum de vraisemblance

Soit  $Y_1, Y_2, \dots, Y_m$  un échantillon aléatoire de taille  $m$  de la distribution Dagum-Exponentielle avec des paramètres  $(\beta, \lambda, \delta)$ . La fonction de vraisemblance est définie par :

$$L(y_1, y_2, \dots, y_m; \beta, \lambda, \delta) = \prod_{k=1}^m g(y_k; \beta, \lambda, \delta) = \prod_{k=1}^m \frac{\lambda}{1+\lambda} \left( e^{-\lambda y_k} + \frac{\beta \delta \lambda^\beta y_k^{\delta-1}}{(\lambda + y_k^\delta)^{\beta+1}} \right) \quad (3.3.1)$$

La fonction log-vraisemblance est :

$$l = \ln L(y_1, y_2, \dots, y_m; \beta, \lambda, \delta) = \frac{\lambda^m}{(1 + \lambda)^m} \left( e^{-\lambda \sum_{k=1}^m y_k} + \frac{\beta^m \delta^m \lambda^{m\beta} \prod_{k=1}^m y_k^{\delta-1}}{\prod_{k=1}^m (\lambda + y_k^\delta)^{\beta+1}} \right)$$

$$l = m \ln \lambda - m \ln(1 + \lambda) - \lambda \sum_{k=1}^m y_k + m \ln \beta + m \ln \delta + m \delta \ln \lambda + (\delta - 1) \sum_{k=1}^m \ln y_k - (\beta + 1) \sum_{k=1}^m \ln(\lambda + y_k^\delta) \quad (3.3.2)$$

Différenciez (3.3.2) par rapport à  $\beta$ ,  $\lambda$  and  $\delta$ , respectivement, pour obtenir les résultats suivants :

$$\frac{\partial l}{\partial \beta} = \frac{m}{\beta} + m \ln \lambda - \sum_{k=1}^m \ln(\lambda + y_k^\delta) \quad (3.3.3)$$

$$\frac{\partial l}{\partial \lambda} = \frac{m}{\lambda} - \frac{m}{1 + \lambda} - \sum_{k=1}^m y_k + \frac{m\beta}{\lambda} - (\beta + 1) \sum_{k=1}^m (\lambda + y_k^\delta)^{-1} \quad (3.3.4)$$

$$\frac{\partial l}{\partial \delta} = \frac{m}{\delta} + \sum_{k=1}^m \ln y_k - (\beta + 1) \frac{\sum_{k=1}^m y_k^\delta \ln y_k}{(\lambda + y_k^\delta)} \quad (3.3.5)$$

Égaliser les équations ci-dessus de (3.3.3) à (3.3.5) à zéro et puis résoudre ces équations simultanément pour y, on obtient les estimations du maximum de vraisemblance  $(\hat{\beta}, \hat{\lambda}, \hat{\delta})$  des paramètres  $(\beta, \lambda, \delta)$ . Il est beaucoup plus facile de résoudre ces équations en utilisant des algorithmes dans les logiciels statistiques comme R, lorsque des ensembles de données sont disponibles.

### 3.3.2 Estimation du paramètre contrainte-résistance

Le modèle contrainte-résistance en fiabilité caractérise la durée de vie d'un composant ayant une résistance aléatoire X sous une contrainte aléatoire Y. Lorsque  $X > Y$ , le composant fonctionnera bien. Le composant tombe en panne lorsque la contrainte qui lui est appliquée dépasse la résistance. Nous examinons la question de l'estimation de  $R = P(X > Y)$  dans cette section.

La fiabilité contrainte-résistance R de la distribution peut être obtenue comme suit :

$$\begin{aligned}
R &= P(X, Y) = \int_0^{+\infty} P(X > Y | Y = y) g_Y(y) dy = \int_0^{+\infty} [1 - G_X(y)] g_Y(y) dy = \\
&\int_0^{+\infty} S_X(y) g_Y(y) dy \\
&= \int_0^{+\infty} \frac{1}{1 + \lambda_1} \left( e^{-\lambda_1 y} + \frac{\lambda_1^{\beta_1+1}}{(\lambda_1 + y^{\delta_1})^{\beta_1}} \right) \frac{\lambda_2}{1 + \lambda_2} \left( e^{-\lambda_2 y} + \frac{\beta_2 r_2 \lambda_2^{\beta_2} y^{\delta_2-1}}{(\lambda_2 + y^{\delta_2})^{\beta_2+1}} \right) dy \\
&= \frac{\lambda_2}{(1 + \lambda_1)(1 + \lambda_2)} \int_0^{+\infty} e^{(\lambda_1 + \lambda_2)y} + \frac{\lambda_1 e^{-\lambda_2 y}}{(\lambda_1 + y^{\delta_1})^{\beta_1}} + \left( e^{-\lambda_1 y} + \frac{\lambda_1^{\beta_1+1}}{(\lambda_1 + y^{\delta_1})^{\beta_1}} \right) \frac{\beta_2 r_2 \lambda_2^{\beta_2} y^{\delta_2-1}}{(\lambda_2 + y^{\delta_2})^{\beta_2+1}} dy \\
R &= \frac{\lambda_2}{(1 + \lambda_1)(1 + \lambda_2)} \left[ \int_0^{+\infty} e^{(\lambda_1 + \lambda_2)y} dy + \int_0^{+\infty} \frac{\lambda_1 e^{-\lambda_2 y}}{(\lambda_1 + y^{\delta_1})^{\beta_1}} dy + \int_0^{+\infty} \left( e^{-\lambda_1 y} + \frac{\lambda_1^{\beta_1+1}}{(\lambda_1 + y^{\delta_1})^{\beta_1}} \right) \right. \\
&\quad \left. \times \frac{\beta_2 \delta_2 \lambda_2^{\beta_2} y^{\delta_2-1}}{(\lambda_2 + y^{\delta_2})^{\beta_2+1}} dy \right] \tag{3.3.6}
\end{aligned}$$

$$= H(\beta_1, \lambda_1, \delta_1, \beta_2, \lambda_2, \delta_2) \tag{3.3.7}$$

Commençons d'abord par obtenir les estimateurs du maximum de vraisemblance (EMV) de  $\beta_1, \lambda_1, \delta_1, \beta_2, \lambda_2$  et  $\delta_2$  avant de calculer l'EMV de R. Par conséquent, la fonction log-vraisemblance de  $\beta_1, \lambda_1, \delta_1, \beta_2, \lambda_2$  et  $\delta_2$  peut être trouvée par

$$\begin{aligned}
L(x_a, y_b; \beta_1, \lambda_1, \delta_1, \beta_2, \lambda_2, \delta_2) &= \prod_{k=1}^{m_1} g_X(x_k) \prod_{l=1}^{m_2} g_X(y_l) \\
&= \prod_{k=1}^{m_1} \frac{\lambda_1}{1 + \lambda_1} \left( e^{-\lambda_1 x_k} + \frac{\beta_1 \delta_1 \lambda_1^{\beta_1} x_k^{\delta_1-1}}{(\lambda_1 + x_k^{\delta_1})^{\beta_1+1}} \right) \prod_{l=1}^{m_2} \frac{\lambda_2}{1 + \lambda_2} \left( e^{-\lambda_2 y_l} + \frac{\beta_2 \delta_2 \lambda_2^{\beta_2} y_l^{\delta_2-1}}{(\lambda_2 + y_l^{\delta_2})^{\beta_2+1}} \right) \\
&= \frac{\lambda_1^{m_1}}{(1 + \lambda_1)^{m_1}} \left( e^{-\lambda_1 \sum_{k=1}^{m_1} x_k} + \frac{\beta_1 \delta_1 \lambda_1^{\beta_1} \prod_{k=1}^{m_1} x_k^{\delta_1-1}}{\prod_{k=1}^{m_1} (\lambda_1 + x_k^{\delta_1})^{\beta_1+1}} \right) + \frac{\lambda_2}{1 + \lambda_2} \left( e^{-\lambda_2 \sum_{l=1}^{m_2} y_l} + \frac{\beta_2 \delta_2 \lambda_2^{\beta_2} \prod_{l=1}^{m_2} y_l^{\delta_2-1}}{\prod_{l=1}^{m_2} (\lambda_2 + y_l^{\delta_2})^{\beta_2+1}} \right) \tag{3.3.8}
\end{aligned}$$

$$\begin{aligned}
\ln L &= \sum_{k=1}^{m_1} g_X(x_k) \sum_{l=1}^{m_2} g_X(y_l) \\
&= m_1 \ln \lambda_1 - m_1 \ln(1 + \lambda_1) - \lambda_1 \sum_{k=1}^{m_1} x_k + m_1 \ln \beta_1 + m_1 \ln \delta_1 + m_1 \beta_1 \ln \lambda_1 + (\delta_1 - 1) \sum_{k=1}^{m_1} \ln x_k \\
&\quad - (\beta_1 + 1) \sum_{k=1}^{m_1} \ln(\lambda_1 + x_k^{\delta_1}) + m_2 \ln \lambda_2 - m_2 \ln(1 + \lambda_2) - \lambda_2 \sum_{l=1}^{m_2} y_l + m_2 \ln \beta_2 + m_2 \ln \delta_2 \\
&\quad + m_2 \beta_2 \ln \lambda_1 + (\delta_2 - 1) \sum_{l=1}^{m_2} \ln y_l - (\beta_2 + 1) \sum_{l=1}^{m_2} \ln(\lambda_2 + y_l^{\delta_2}) = 0 \tag{3.3.9}
\end{aligned}$$

Différenciez (3.3.9) par rapport à  $\beta_1, \lambda_1, \delta_1, \beta_2, \lambda_2$  et  $\delta_2$ , respectivement, pour obtenir les résultats suivants :

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{m_1}{\beta_1} + m_1 \ln \lambda_1 - \sum_{k=1}^{m_1} \ln(\lambda_1 + x_k^{\delta_1}) = 0 \quad (3.3.10)$$

$$\frac{\partial \ln L}{\partial \lambda_1} = \frac{m_1}{\lambda_1} - \frac{m_1}{1 + \lambda_1} - \sum_{k=1}^{m_1} x_k + \frac{m_1 \beta_1}{\lambda_1} - (\beta_1 + 1) \sum_{k=1}^{m_1} \frac{1}{\lambda_1 + x_k^{\delta_1}} = 0 \quad (3.3.11)$$

$$\frac{\partial \ln L}{\partial \delta_1} = \frac{m_1}{\delta_1} + \sum_{k=1}^{m_1} x_k - (\beta_1 + 1) \sum_{k=1}^{m_1} \frac{x_k^{\delta_1} \ln x_k}{\lambda_1 + x_k^{\delta_1}} = 0 \quad (3.3.12)$$

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{m_2}{\beta_2} + m_2 \ln \lambda_2 - \sum_{l=1}^{m_2} \ln(\lambda_2 + y_l^{\delta_2}) = 0 \quad (3.3.13)$$

$$\frac{\partial \ln L}{\partial \lambda_2} = \frac{m_2}{\lambda_2} - \frac{m_2}{1 + \lambda_2} - \sum_{l=1}^{m_2} y_l + \frac{m_2 \beta_2}{\lambda_2} - (\beta_2 + 1) \sum_{l=1}^{m_2} \frac{1}{\lambda_2 + y_l^{\delta_2}} = 0 \quad (3.3.14)$$

$$\frac{\partial \ln L}{\partial \delta_2} = \frac{m_2}{\delta_2} + \sum_{l=1}^{m_2} y_l - (\beta_2 + 1) \sum_{l=1}^{m_2} \frac{y_l^{\delta_2} \ln y_l}{\lambda_2 + y_l^{\delta_2}} = 0 \quad (3.3.15)$$

### 3.4 Mesures actuarielles

Plusieurs caractéristiques actuarielles de la distribution Dagum-Exponentielle sont abordées dans cette section.

#### La valeur à risque ou la fonction quantile

La valeur à risque de la distribution Dagum-Exponentielle est définie comme suit :

$$VaR = Q_p = y_p = \left( \left[ \lambda^\beta - \frac{\lambda^{\beta-1}}{(1-p)(1+\lambda)} - \lambda^{\beta-2} \log(p) \right]^{\frac{1}{\beta}} - \lambda \right)^{\frac{1}{\delta}} \quad (3.4.1)$$

## La fonction de durée de vie

La fonction de durée de vie résiduelle ou excédentaire moyenne pour une variable aléatoire de montant de sinistre  $Y$  est le paiement prévu par sinistre sur une police avec une franchise fixe de  $y$ , où les sinistres dont les montants sont inférieurs ou égaux à  $y$  sont ignorés. Elle est définie pour la distribution Dagum-Exponentielle comme suit :

$$\begin{aligned}
 e(y) &= E(Y - y | Y > y) = \frac{1}{1-G(y)} \int_y^{+\infty} [1 - G(t)] dt \\
 &= \frac{1}{G(y)} \int_y^{+\infty} [G(y)] dt \\
 &= \frac{1}{\frac{1}{1+\lambda} [e^{-\lambda y} - \frac{\lambda^{\beta+1}}{(\lambda+y^\delta)^\beta}]} \times \frac{1}{1+\lambda} \int_y^{+\infty} [e^{-\lambda u} - \frac{\lambda^{\beta+1}}{(\lambda+u^\delta)^\beta}] du \\
 &= \frac{1}{e^{-\lambda y} - \frac{\lambda^{\beta+1}}{(\lambda+y^\delta)^\beta}} \left[ \frac{e^{-\lambda y}}{\lambda} + \frac{(\lambda+y^\delta)^{2\beta} \Gamma(1+\frac{1}{\delta}) \Gamma(\beta-\frac{1}{\delta})}{\lambda^{2\beta-\frac{1}{\delta}} \Gamma(\beta)} \overline{B}\left(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \alpha-\frac{1}{\delta}\right) - \frac{y(\lambda+y^\delta)^\beta}{\lambda^{\beta-1}} \right]
 \end{aligned} \tag{3.4.2}$$

Où  $\overline{B}\left(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \beta-\frac{1}{\delta}\right) = 1 - B\left(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \beta-\frac{1}{\delta}\right)$ , et  $B\left(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \beta-\frac{1}{\delta}\right)$  est la fonction bêta incomplète.

## La fonction de valeur espérée limitée

Voici la définition de la fonction de valeur espérée limitée  $L$  d'une variable de taille de sinistre  $Y$ , ou de la fonction de répartition associée  $G(y)$ .

$$\begin{aligned}
 L(t) &= E\{\min(Y, t)\} = \int_0^t y dG(y) + t(1 - G(t)) \quad , \text{où } t > 0. \\
 &= \int_0^t y f(y) dx + t\overline{G}(t) \\
 &= m_1(t) + t\overline{G}(t)
 \end{aligned}$$

Où

$$\begin{aligned}
 m_1(t) &= \int_0^t y g(y) dy = \left[ \int_0^t y \frac{\lambda}{1+\lambda} [e^{-\lambda y} + \frac{\beta \delta \lambda^\beta y^{\delta-1}}{(\lambda+y^\delta)^{\beta+1}}] dy \right] \\
 &= \frac{\lambda}{1+\lambda} \left[ \frac{e^{-\lambda t}(t-1)}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda^{\frac{1}{\delta}} \Gamma(1+\frac{1}{\delta}) \Gamma(\beta-\frac{1}{\delta}) \overline{B}\left(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \alpha-\frac{1}{\delta}\right)}{\Gamma(\beta)} \right]
 \end{aligned}$$

Et,

$$\overline{tG}(t) = \frac{t}{1+\lambda} \left[ e^{-\lambda t} - \frac{\lambda^{\beta+1}}{(\lambda+t^\delta)^\beta} \right]$$

Donc,

$$L(t) = \frac{\lambda}{1+\lambda} \left[ \frac{e^{-\lambda t}(t-1)}{\lambda} + \frac{1}{\lambda^2} + \frac{\lambda^{\frac{1}{\delta}} \Gamma(1+\frac{1}{\delta}) \Gamma(\beta-\frac{1}{\delta}) \bar{B}(\frac{y^\delta}{\lambda+y^\delta}; 1+\frac{1}{\delta}; \alpha-\frac{1}{\delta})}{\Gamma(\beta)} \right] + \frac{t}{1+\lambda} \\ \times \left[ e^{-\lambda t} - \frac{\lambda^{\beta+1}}{(\lambda + t^\delta)^\beta} \right] \quad (3.4.3)$$

## La valeur de queue à risque

Une mesure du risque liée à la valeur globale en risque est la valeur de queue à risque (TVaR), parfois appelée attente conditionnelle de queue. L'espérance de pertes au-delà de la VaR est mesurée par TVaR. Pour la distribution Dagum-Exponentielle, la TVaR est définie comme suit :

$$TVaR = E(Y|Y > VaR) = \frac{1}{1-p} \int_{VaR}^{+\infty} yg(y)dy \\ = \frac{1}{(1-p)(1+\lambda)} \int_{VaR}^{+\infty} \left[ \lambda ye^{\lambda y} + \frac{\beta \delta \lambda^{\beta+1} y^\delta}{(\lambda+y^\delta)^{\beta+1}} \right] dy \\ TVaR = \frac{1}{(1-p)(1+\lambda)} \left( \frac{1}{\lambda} - \frac{\ln(1-p)}{\lambda} + \frac{\lambda^{\frac{1}{\delta}+1} \Gamma(1+\frac{1}{\delta}) \Gamma(\beta-\frac{1}{\delta})}{(1-p)\Gamma(\beta)} \right. \\ \left. \times \bar{B}\left(\frac{VaR_p(Y)^\delta}{\lambda + VaR_p(Y)^\delta}; 1 + \frac{1}{\delta}; \alpha - \frac{1}{\delta}\right) \right) \quad (3.4.4)$$

## 3.5 Application

### 3.5.1 Étude de simulation

Dans cette section, nous effectuons les études de simulation en utilisant la distribution mixte Dagum-Exponentielle pour voir les performances des estimateurs ci-dessus correspondant à cette distribution. La génération de la distribution mixte Dagum-Exponentielle peut être facilement obtenue par sa fonction quantile. Nous générons  $N = 1000$  échantillons de taille  $n = 30, 50, \dots, 300$ , et  $(\beta, \lambda, \delta) = (0,3, 0,5), (0,7, 0,3, 1,9), (0,5, 1,4, 0,8), (1,2, 1,2, 1,2)$ .

Pour chaque échantillon, les MLE sont obtenus, ceux-ci sont utilisés pour calculer les valeurs des quantités suivantes à l'aide du package R.

- 1) Biais moyen du MLE  $\hat{\theta}$  du paramètre  $\theta = \beta, \lambda, \delta$  :

$$\frac{1}{N} \sum_{k=1}^N (\hat{\theta} - \theta)$$

2) Erreur quadratique moyenne (MSE)  $\hat{\theta}$  du paramètre  $\theta = \beta, \lambda, \delta$  :

$$\frac{1}{N} \sum_{k=1}^N (\hat{\theta} - \theta)^2$$

Paramètres $\beta, \lambda, \delta$	n	Biais moyen $(\hat{\beta})$	Erreur quadratique moyenne $(\hat{\beta})$	Biais moyen $(\hat{\lambda})$	Erreur quadratique moyenne $(\hat{\lambda})$	Biais moyen $(\hat{\delta})$	Erreur quadratique moyenne $(\hat{\delta})$
0.3, 0.5, 0.5	30	0.00141627	0.0200583	0.00502043	0.0252047	0.01288778	0.166095
	50	0.00105241	0.00110758	0.00423498	0.0179351	0.01082644	0.1172117
	100	0.00092550	0.00085656	0.00344188	0.0118466	0.00887452	0.0787571
	150	0.00072632	0.00052755	0.00263798	0.0069589	0.00696255	0.0484772
	200	0.00046960	0.00022052	0.00175304	0.0030731	0.00447595	0.0200342
	300	0.00042262	0.00017860	0.00110075	0.0012116	0.00267971	0.0071807
0.7, 0.3, 1.9	30	0.00131534	0.00173012	0.04580596	2.098186	0.08356383	6.982913
	50	0.00074201	0.00055058	0.04250207	1.806426	0.07772872	6.041754
	100	0.00057333	0.00032871	0.03606934	1.300997	0.06573282	4.320804
	150	0.00042025	0.00017661	0.02720505	0.7401149	0.04934979	2.435402
	200	0.00014794	0.00011887	0.00689226	0.0475033	0.01127154	0.1270477
	300	0.00009620	0.00009254	0.00528147	0.0278939	0.00846321	0.071626
0.5, 1.4, 0.8	30	0.00154194	0.00237759	0.05305677	2.815021	0.0961573	9.246225
	50	0.00123384	0.00152237	0.04186695	1.752841	0.07532416	5.67373
	100	0.00082067	0.00067351	0.04111728	1.69063	0.07415485	5.498941
	150	0.00055826	0.00031166	0.03426607	1.174163	0.06156011	3.789647
	200	0.00012827	0.00016455	0.01099087	0.1207992	0.01770881	0.313602
	300	0.00012317	0.00015170	0.00691773	0.0478550	0.01063283	0.1130572
1.2, 1.2, 1.2	30	0.00683663	0.00767395	0.08278980	0.7983013	0.05384835	2.0289964
	50	0.00526521	0.00403378	0.07181435	0.7720555	0.05380267	2.0289472
	100	0.00397107	0.00388511	0.067821183	0.7559072	0.05298244	2.0280714
	150	0.00107997	0.00116633	0.02931559	0.4843542	0.0390933	1.528286
	200	0.00092171	0.00084957	0.02447384	0.4711257	0.03893505	1.515938
	300	0.00078030	0.00060882	0.02161525	0.4672192	0.03879364	1.504946

**Tableau 02** : Le biais et l'erreur quadratique moyenne des paramètres  $\theta = \beta, \lambda, \delta$

**Interprétation :**

- L'erreur de biais moyenne est une mesure de l'écart systématique entre les résultats d'une méthode de mesure et la valeur réelle. Les valeurs proches de 0 sont optimales, les valeurs négatives indiquent une sous-estimation et les valeurs positives une surestimation.

- L'erreur quadratique moyenne combine la dispersion des erreurs individuelles. C'est le deuxième moment (autour de l'origine) de l'erreur et intègre donc à la fois la variance de l'estimateur (la dispersion des estimations d'un échantillon de données à un autre) et son biais (la distance entre la valeur moyenne estimée et la valeur réelle). MSE combine la dispersion des erreurs individuelles.

### 3.5.2 Application réelle

**Ensemble de données :** L'analyse d'un ensemble de données d'enregistrement obligatoire des assurances automobiles semble être un projet de recherche passionnant. L'ensemble de données fourni par le Centre d'information et de surveillance des assurances routières fait l'objet de cette analyse.

L'ensemble de données comprend les données relatives aux sinistres, la police expirant en décembre 2018 et débutant en janvier 2016 (voir annexe).

Pour confirmer l'adéquation de notre modèle à ces données, nous examinons plusieurs tests d'ajustement avec des distributions connues, fréquemment utilisées pour l'ajustement d'ensemble de données. Parmi ces tests figurent la distribution mixte Dagum-Exponentielle, la distribution Weibull, la distribution Exponentielle, la distribution Log-normale et la distribution de Dagum. Le processus itératif de l'algorithme est implémenté en code R.

Distribution	-2 log L	AIC	BIC	CAIC
Dagum-Exponentielle	<b>3489,82</b>	<b>3500,64</b>	<b>3505,15</b>	<b>3501,34</b>
Dagum	3892,78	3887,58	3892,52	3888,45
Exponentielle	68588,8	68624,22	68612,44	68611,77
Log-Normal	65599,5	65604,54	65608,98	65604,81
Weibull	3825,08	3828,24	3831,46	3828,45

**Tableau 03 :** Statistiques d'ajustement du modèle Dagum-Exponentielle.

Les valeurs des statistiques AIC, BIC, CAIC, -2 log L dans le tableau 03 indiquent que le modèle Dagum-Exponentielle est un concurrent sérieux des autres distributions. C'est le meilleur ajustement mesurant la qualité des statistiques d'ajustement précédentes.

# Conclusion

Dans cette thèse, nous avons présenté une nouvelle distribution à trois paramètres, appelée Dagum- Exponentielle, dérivée de la combinaison des distributions Dagum et Exponentielle. Initialement nous avons parlé globalement sur la modélisation des risques, après nous avons concentré sur les coûts de sinistres dans le cas des distributions à queue lourde, à queue légère et distribution mixtes.

Notre nouvelle distribution mixte possède des propriétés utiles dans diverses situations.

La méthode du maximum de vraisemblance est utilisée pour estimer les paramètres du nouveau modèle Dagum-Exponentielle. Une approche de simulation a été utilisée pour évaluer les performances de l'estimateur proposé selon la technique d'estimation décrite. À l'aide de statistiques d'ajustement, l'efficacité du nouveau modèle, comparée aux dernières itérations des distributions Dagum, weibull, Log-Normale et Exponentielle, est illustrée à l'aide d'un ensemble de données réelle relatif aux sinistres, où la police expire dans deux ans.

La distribution mixte Dagum-Exponentielle influence les provisions techniques en offrant un compromis entre la modélisation de sinistres fréquents de faible à moyenne intensité et la prise en compte modérée des sinistres extrêmes. Grâce à sa flexibilité, cette distribution permet de représenter de manière réaliste les coûts de sinistres dans des portefeuilles où la sévérité est relativement contrôlée, mais où une certaine variabilité subsiste. Elle conduit généralement à des provisions moins volatiles que celles issues de distributions à queue lourde comme la Pareto, tout en restant plus prudente qu'une simple loi exponentielle. Son utilisation permet ainsi une estimation des réserves techniques plus équilibrée, adaptée aux exigences de solvabilité et à la gestion des risques dans les branches non-vie, en particulier pour les produits où les sinistres élevés existent mais ne dominent pas le portefeuille.

**Perspective :** Afin de résoudre le problème des moments infinis, les recherches futures pourraient se concentrer sur l'étude de l'estimation bayésienne des paramètres du modèle Dagum-Exponentielle et présenter une version abrégée de ce modèle.

# Bibliographie

[1] Abdi, M., Asgharzadeh, A., Bakouch, H.S. & Alipour, Z. A new compound Gamma and Lindley distribution with application to failure data, *Austrian Journal of Statistics*, 48, 2019.

[2] Affify, A.Z., Gemeay, A.M. & Akma Ibrahim, N. The heavy-tailed Exponential distribution : risk measures, estimation, and application to actuarial data, *Mathematics*, 1276 (8), 2020.

[3] Alzaatreh, A., Famoye, F. & Lee, C. Gamma-Pareto distribution and its applications, *Journal of Modern Applied Statistical Methods*, 11(1), 78-94, 2012.

[4] Alzaatreh, A., Famoye, F. & Lee, C. Weibull-Pareto distribution and its applications, *communications in statistics- theory and methods*, 42(9), 1673-1691, 2013.

[5] Antoniou, I., Ivanov, V.V., Ivanov, Valery.V. & Zrelov, P.V. On the Log-normal distribution of network traffic, *Physica D Nonlinear Phenomena*, 167, 72-85, 2002.

[6] Badr, M. & Ijaz, M. The Exponentiated Exponential Burr XII distribution : theory and application to lifetime and simulated data, *PLOS ONE*, 16(3), 2021.

[7] Bechiri, S. & Remita, M. R. On Dagum-Pareto mixture distribution : properties, simulation and application in insurance, *Studies in Engineering and Exact Sciences*, 5(2), 1-21, 2024.

[8] Bechiri, S. & Remita, M. R. On Dagum-Exponentielle mixture distribution : properties, simulation and application in insurance, *Utilitas Mathematica*, 122(1), 176-194, 2025.

[9] Beghriche, Abdelfateh, H. Zeghdoudi, Vinoth Raman & Sarra Chouia. On size-biased approach in generalizing Lindley distribution : properties and applica-

tions. *Int. J. Agricult. Stat. Sci.*, 16(2), 519-526, 2022.

[10] Belhamra, T., Zeghdoudi, H., & Raman, V. Reliability for Zeghdoudi distribution with an outlier, fuzzy reliability and application. *Statistics in Transition new series*, 25(1), 167-177, 2024.

[11] Belhamra, T., Zeghdoudi, H., & Raman, V. A new compound Exponential-Lindley distribution : application and comparison. *International Journal of Agricultural and Statistical Science*, 18(2), 755-766, 2022.

[12] Bhatt, F.A., Hamedani, G.G., Korkmaz, M.C., Sheng, W. & Ali, A. On the Burr XII moment Exponential distribution, *PLOS ONE*, 16(2), 2021.

[13] Chouia, S. & Zeghdoudi, H. The x lindley distribution : properties and application. *J. Stat. Theory Appl.* 20(2), 318-327, 2021.

[14] Dey, S., Al-Zahrani, B. & Basloom, S. Dagum distribution : properties and different methods of estimation. *International Journal of Statistics and Probability*, 6(2), 2017.

[15] Djemoui, N. H. Estimation de la prime chargée et des caractéristiques de fiabilité dans des modèles généralisés, *Université Badji Mokhtar Annaba*, 2024.

[16] El Adlouni, S., Bobée, B. & Ouarda, T. Caractérisation des distributions à queue lourde pour l'analyse des crues, *Université du Québec*, 2007.

[17] Ekhosuehi, N., Nzei, L.C. & Opone, F.A new mixture of Exponential-Gamma distribution . *Journal of Science*, 33(2), 2020.

[18] Goffard, P. O. Modélisation Charge Sinistre, *Université de Lyon 1 ISFA*, 2019.

[19] Kerdali, A. Modélisation stochastique des provisions techniques en assurance non-vie Modèles récursifs de provisionnement, *Ecole Nationale Supérieure de Statistique et d'Economie Appliquée*, 2017.

[20] Khodja, N., Gemeay, A.M., Zeghdoudi, H., Karakaya, K., Alsahangitis, A.M., Bakr, M.E., Oluwafemi Samson Balogun., MUSE, A.H. & Hussam.E. Modeling voltage real data set by a new version of Lindley distribution, *IEEE Access*, 2023.

[21] Khordj, M., Haffar, A. & Teulon, F. Provisionnement et mesure de risque en assurance dommage dans le cadre de Solvabilité II, *Cairn.info*, 34(3), 137-168, 2017.

[22] Le Tesson, A., Lenain, A., Samba, S. & Ung, J. Estimation de l'erreur de prediction dans ' le cas de l'utilisation d'une combinaison de methodes pour le calcul de provisions ' en assurance IARD, *Euro Institute of Actuarial Studies*, 2014.

[23] Marceau, E. Modélisation et évaluation quantitative des risques en actuariat modèles sur une période, *Springer*, 2013.

[24] Pelumi E. Oguntunde, Member, IAENG, Adebowale O. Adejumo, Enahoro A. Owoloko, Manoj K. Rastogi, & Oluwole A. Odetunmibi. The Burr X-Exponential distribution : theory and applications . *World Congress on Engineering*, 1, 2017.

[25] Planchet, F. Assurance non vie. ISFA – Support de cours - 1, 2003.

[26] Saaidia, N., Belhamra, T., & Zeghdoudi, H. On ZLindley distribution : statistical properties and applications. *Studies in Engineering and Exact Sciences*, 5(1), 3078-3097, 2024.

## Annexe

### Montants réclamés entre 2016 et 2018 (en DA).

Mois-Année	Montants réclamés
January, 2016	4477181800.5
February, 2016	5006102503.1
March, 2016	4121727487.9
April, 2016	4288391254.7
May, 2016	2344716296.1
June, 2016	33892121.2
July, 2016	877209622.3
August, 2016	4757585745.8
September, 2016	68463532.2
October, 2016	1250647142.5
November, 2016	8544193298.1
December, 2016	4967526209.9
January, 2017	4808504924.4
February, 2017	4278134665.8
March, 2017	1238844907.3
April, 2017	4244801650.2
May, 2017	4946063021.3
June, 2017	4766793446.4
July, 2017	4705322787.4
August, 2017	4804505758.3
September, 2017	4856654476.3
October, 2017	8463219700.8
November, 2017	2288724944.2
December, 2017	6287162470.5
January, 2018	61299958.4
February, 2018	84208270.1
March, 2018	8646708121.8
April, 2018	6423014566.1
May, 2018	5129211945.7
June, 2018	18453338.4
July, 2018	74026741.9
August, 2018	5283458942.7
September, 2018	1254577870.6
October, 2018	10735988.2
November, 2018	5164514853.5
December, 2018	61615231.0