الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

جامعة سعد دحلب البليدة-1-
Université Saad Dahlab Blida -1-

Mémoire de fin D'études

En vue de l'obtention du diplôme Master

Faculté de sciences
Département : Mathématique

Spécialité : Modélisation stochastique et statistique

THEME

Modèles de régression flexibles de Tweedie pour des données continues, Application aux assurances d'automobiles

Présenté par :
MOUMENI Ouahiba
DOUM Sihem

Soutenu le : 24-09-2020

Devant le Jury :

Président :	O.TAMI	MAA	**Université de Blida 1**
Promoteur :	A. RASSOUL	MCA	**ENSH de Blida**
Examinateur :	R.FRIHI	MAA	**Université de Blida 1**

Promotion : 2019/2020

DEDICACE

Tous d'abord, je remercie le ***Allah*** *de m'avoir donné le courage pour réaliser ce travail et la patience pour aller jusqu'au bout du parcours de mes études.*

Je dédie le fruit de ma patience, de ma persévérance :

Aux êtres les plus chers au monde,

à ma raison de vivre et ma fleur de vie, ***ma chère mère*** *qui m'a apporté beaucoup d'amour et d'affection, je la remercie de sa présence dans les meilleurs moments comme les mauvais.*

Mon ***cher père*** *qui n'a jamais cessé de combattre pour me voir réussir un jour.*

Que dieu leurs accorde une longue vie.

A toute ***ma famille MOUMENI****.*

A tous mes enseignants, je leurs exprime ma profonde gratitude.

À toutes les personnes qui m'ont aidé, soutenu et contribué de près ou de loin pour la réalisation de ce travail.

Ouahiba

DEDICACE

Tous d'abord, je remercie le ***Allah*** *de m'avoir donné le courage pour réaliser ce travail et la patience pour aller jusqu'au bout du parcours de mes études.*

Je dédie le fruit de ma patience, de ma persévérance :

Aux êtres les plus chers au monde,

à ma raison de vivre et ma fleur de vie, ***ma chère mère*** *qui m'a apporté beaucoup d'amour et d'affection, je la remercie de sa présence dans les meilleurs moments comme les mauvais.*

Mon ***cher père*** *qui n'a jamais cessé de combattre pour me voir réussir un jour.*

Que dieu leurs accorde une longue vie.

A toute ***ma famille DOUM***

A tous mes enseignants, je leurs exprime ma profonde gratitude.

À toutes les personnes qui m'ont aidé, soutenu et contribué de près ou de loin pour la réalisation de ce travail.

Sihem

REMERCIEME

Au terme de ce mémoire je tiens à remercier, en premier lieu Dieu le tout puissant qui m'a donne la force, le courage et la patience pour terminer ce modeste travail.

On tient à remercier vivement note promoteur A.RASSOUL, pour ses encouragements et ses précieux conseils Durant le déroulement de ce travail.

Mes remerciements s'adressent au president ainsi que les members du jury pour avoir de juger ce travail.

Nous tenons à exprimer toute nos reconnaissances à toutes mes amies et collègues de 2éme année master MSS.

Enfin, il remercie toutes les personnes qui ont contribute de prés ou de loin, à la realization de ce modeste travail.

ملخص

توفر نماذج الانحدارالتويد مجموعة مرنة من التوزيعات التي يمكنها معالجة البيانات المستمرة الغير سلبية بكتلة احتمالية على أساس طريقة الاحتمال القصوى من خلال وجود مبلغ لا حصر له في البيانات التي تساوي الصفر .و يتم الطعن في تقدير و استدلال دالة الاحتمال و القيود الغير مرتبطة بالعمل على مساحة معلمة الطاقة. في هذا العمل نقدم النموذج الخطي المعمم و المتغيرات العشوائية لعائلة التويد التي تعتبر نماذج للتشتت الأسي. نطبق هذا نموذج تأمين على السيارات لنمذجة خسائر شركة التامين.

Résumé

Les Modèles de Régression Tweedie (TRM) fournissent une famille de distribution flexible qui peuvent traiter des données continues non négatives avec une masse de probabilité des données à zéro. L'estimation et inférence des TRM basée sur la méthode du maximum de vraisemblances ont défis par la présence d'une somme infinie dans la fonction de probabilité et les restrictions non liées du travail sur l'espace des paramètres de puissance. Dans ce travail nous présentons le modèle linéaire généralisé englobant et les variables aléatoires de la famille de Tweedie qui sont des modèles de dispersion exponentielle. Nous appliquons ce modèle sur l'assurance d'automobiles pour modéliser les pertes d'une compagnie d'assurance et calcul de la prime d'assurance.

Abstract

Tweedie Regression Models (TRM) provide a flexible family of distributions that can process non-negative continuous data with zero data probability mass. The estimation and inference of MRTs based on the maximum likelihood method are challenged by the presence of an infinite sum in the probability function and the unrelated restrictions of the work on the power parameter space. In this work we present the generalized linear model encompassing and the random variables of the Tweedie family which are models of exponential dispersion. We apply this model to auto insurance to model the losses of an insurance company and to calculate the insurance premium.

TABLE DES MATIÈRES

TABLE DES FIGURES

LISTE DES TABLEAUX

INTRODUCTION GÉNÉRALE

La modélisation statistique est l'une des domaines le plus importants de la statistique appliquée avec des applications dans de nombreux domaines de la recherche scientifique, tels que la sociologie, l'argronomie, les assurances et la médecine pour n'en citer que quelques-uns. Il existe une infinité de cadres de modélisation statistique, mais la classe existe des modèles linéaires généralisées (GLM)(cf. John Nelder et Wedderburn, (1972)) est la plus utilisée au cours des quatre dernières décennies. Le succès de cette approche est dû à sa capacité à traiter différents types de variables de réponse, telles que binaire, comptage et continu dans un cadre général avec un système puissant d'estimation et d'inférence basé sur le paradigme de vraisemblance. Les distribution de Tweedie sont largement utilisées dans la modélisation statatistique, motivent ainsi l'étude de leur estimation dans un cadre plus général. L'objectif de ce mémoire porte sur les modèles de régression tweedie flexible pour des données continues .Ce travail est divisée par trois chapitres :

Dans le premier chapitre on a présente les cas particuliers de la classe (GLM) incluent le modèle linéaire gaussien pour traiter des données continues, les modèles de régression gamma et gaussienne inverse pour traiter des données continues positives. Modèles de régréssion logistique et de poisson pour traiter respectivement les données binaires ou binomiales et de comptage.

Le deuxiéme chapitre présente les distributions de Tweedie qui appartiennent à la classe des modèles de dispersion exponentielle qui sont caractérés par une fonction de variance d'unité de puissance, sont infiniment divisibles et sont fermées sous les traductions et les transformations d'échelle. Notamment, une variables aléaitoire de Tweedie a un paramètre d'indexation / puissance qui essentiel pour décrire sa distribution. Généralement les densités des distributions Tweedie n'ont pas une forme fermée, c'est-à-dire que nous ne pouvons pas calculer la densité directement. Quelques distributions particulières, telles que la variable aléaitoire discrète (poisson), une variable aléaitoire mixte (poisson-gamma), des variables aléaitoires continues, et des

variables stables. Peuvent être exprimées sous forme fermée. Puisque la fonction de vraisemblance de la distribution de Tweedie ne peut pas être écrite sous forme fermée, les techniques d 'estimation du maximum de vraisemblance sont difficiles à utiliser. L'objectif est d'étudier les méthodes d'estimation de la densité de Tweedie et de montrer l'utilité de cette densité dans le domaine actuariel, plus précisément en assurance automobile.

Nous terminerons ce mémoire par le chapitre 3, qui est consacré à l'application de la distribution Tweedie en assurance à l'aide d'une base de données réelles qui représentent les totaux des réclamations individuelles d'assurance. Les résultats obtenus montrent que la distribution Tweedie s'ajuste bien aux réclamations d 'assurance. Par ailleurs, la densité de Tweedie estimée nous a aidé également à calculer la transformée et la prime d'Esscher.

CHAPITRE 1

MODÈLES LINÉAIRES GÉNÉRALISÉES

Sommaire

1.1 Introduction

En statistiques, le modèle linéaire généralisé (GLM) est une généralisation de la régression linéaire. Les modèles linéaires généralisés ont été comme un moyen d'unifier les modèles statistiques compris la régression linéaire, la régression logistique et la régression de poisson.

1.2 Généralités sur les modèles linéaire généralisés

Définition 1 *En statistiques, le modèle linéaire généralisé (MLG) souvent connu les initailes anglaises GLM (general linear models) est une généralisation de la régression linéaire. Il permet d'étudier la liaison entre une variable réponse Y et une ensemble des variables explicatives ou prédicteures* $(X_1, X_2, \dots, X_n)$.

Les modéle linéaire généralisés ont été formulés par (cf. John Nelder et Robet wedderburn, (1972)) comme un moyen d'unifier les auters modéles statistiques comparis la régression linéaire la régression logistique et la régression de poisson.

1.3 Présentation des modèles linéaires généralés

Les modèle linéaire généralisés sont formés de trois composantes :

- La variable de réponse Y, composante aléatoire à laquelle est associée une loi de probabilité.
- La variable explicatives $(X_1, X_2..., X_n)$ utilisées comme prédicteurs dans le modèle définissent sous forme d'une combinaision linéaire la composante déteministe.
- La fonction lien décrit la relation fonctionnelle entre la combinaision linéaire des variable $(X_1, X_2, ..., X_n)$ et l'espérance mathématique de la variable de réponse Y.

Composante aléatoire est définie par la distribution de probabilité de la variable Y. Elle peut être choisié dans la famille exponentielle à la quelle appartiennent de lois normale, binomiales, poisson, gamma, etc... une propriété de ces lois est pour chacune d'elle, il existe une relation spécifique entre l'espérance $E(Y) = \mu$ et la $Var(Y) = \sigma^2$.

composante déterministe la composante déterministe, exprimée sous forme d'une combinaison linéaire $a_0 + a_1X_1 + ... + a_nX_n$ (appelée aussi prédicteur linéaire) précise quels sont les prédicteurs.

Certaines des variables X_i peuvent se déduire de variables intailes utilisées dans le modèle, par exemple :

- $X_3 = X_1 * X_2$ de façon à étuduier l'interaction entre X_1 et X_2.
- Ou encore $X_4 + X_1^2$ de façon à prendre en compte un effet non linéaire de la variable X_1.

fonction lien la troisième composante d'un modèle linéaire généralisé est lien entre la composante aléatoire et la composante déterministe.

Le modèle linéaire sous sa forme générale Pour le modèle théorique, la forme générale du modèle linéaire est données par :

$$Y = a_0 + a_1 X_1 + a_2 X_2 + ... + a_p X_p + \varepsilon_i = \sum_{i=0}^{p} a_i X_i + \varepsilon_i \qquad \text{et } i = 1, ..., p,\ X_0 = 1.$$

Y une variable réponse et $X_1, X_2, ..., X_p$ les variables explicatives et ε_i erreur de spécification et a_i des paramètres naturelle.

— Pour n observation :

$$Y_i = a_0 + a_1 X_{1i} + a_2 X_{2i} + ... + a_p X_{pi} + \varepsilon_i = \sum_{j=0}^{p} a_j X_{ij} + \varepsilon_i,\ i = 1, ..., n.$$

— Pour le modèle estimé :

$$Y_i = \hat{a}_0 + \hat{a}_1 X_i + ... + \hat{a}_p X_p + e_i\ ,\ i = 1, ..., p.$$

$\hat{a}_i$ sont des paramétres estimés et e erreur résidu.

1.4 Les modèles linéaires

1.4.1 Régression linéaire simple

Définition 2 *Le modèle linéaire à 2 variables ou le modèle linéaire simple est utilisé pour étudier la relation entre deux variables : une variable dépendante (réponse ou endogène) Y et une variable indépendante (exogène ou explicative). Pour analyser la relation entre X et Y, nous commençons par la représentation graphique : si le nuage de points peut être ajusté par une droite $\rightsquigarrow$ proposer le modèle linéaire.*

— Le modèle théorique :

$$Y_i = a_0 + a_1 X_i + \varepsilon_i,\ \ i = 1, 2, ..., n. \tag{1.1}$$

ε_i : erreur de spécification.

— Le modèle estimé :

$$Y_i = \hat{a}_0 + \hat{a}_1 X_i + e_i,\ \ i = 1, 2, ..., n. \tag{1.2}$$

e_i : résidu, on a

$$e_i = Y_i - \hat{Y}_i. \tag{1.3}$$

et

$$\hat{Y} = \hat{a}_0 + \hat{a}_1 X_i.$$

le modèle théorique :

$$Y_i = a_0 + a_1 X_i + \varepsilon_i$$

l'erreur de spécification ε_i est supposée vérifier les hypothèses suivantes :

$H_1 : \varepsilon_i$ est normalement distribué

$$\varepsilon_i \rightsquigarrow \mathcal{N}(0, \sigma_\varepsilon^2).$$

H_2 : Les erreurs sont indépendantes entre elle c-à-d

$$E(\varepsilon_i \varepsilon_j) = 0 \quad , \forall i \neq j.$$

$H_3 : \varepsilon_i$ est indépendantes de la variable explicative X_i c-à-d

$$E(\varepsilon_i X_i) = 0 \ ; \forall i = 1, 2, ..., n.$$

1.4.2 Estimation des paramètres

— Le résidu

$$e_i = Y_i - \hat{Y}_i = Y_i - (\hat{a}_0 + \hat{a}_1 X_i)$$

— Le carré de résidu

$$e_i^2 = (Y_i - (\hat{a}_0 + \hat{a}_1 X_i))^2$$

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} (Y_i - (\hat{a}_0 + \hat{a}_1 X_i))^2 = \varphi(\hat{a}_0, \hat{a}_1) \tag{1.4}$$

$\rightsquigarrow$ erreur quadratique pour estimes les coefficient, il suffet de chercher $\hat{a}_0$ et $\hat{a}_1$ qui minimise $\varphi(\hat{a}_0, \hat{a}_1)$ $\rightsquigarrow$ Méthode des moindre carrées ordinaire (M.C.O).

$$\frac{d\varphi(\hat{a}_0, \hat{a}_1)}{d\hat{a}_0} = 0 \tag{1.5}$$

$$\frac{d\varphi(\hat{a}_0, \hat{a}_1)}{d\hat{a}_1} = 0 \tag{1.6}$$

$$\min \sum_{i=1}^{n} [Y_i - (\hat{a}_0 + \hat{a}_1 X_i)]^2 = \min \varphi(\hat{a}_0, \hat{a}_1) \tag{1.7}$$

$$\min \sum_{i=1}^{n} (Y_i^2 + \hat{a}_0^2 + \hat{a}_1^2 X_i^2 + 2\hat{a}_0 \hat{a}_1 X_i - 2\hat{a}_0 Y_i - 2\hat{a}_1 Y_i X_i)$$

$$\min(\sum_{i=1}^{n} Y_i^2 + \sum_{i=1}^{n} \hat{a}_0^2 + \sum_{i=1}^{n} \hat{a}_0^2 X_i^2 + 2\hat{a}_0 \hat{a}_1 \sum_{i=1}^{n} X_i - 2\hat{a}_0 \sum_{i=1}^{n} Y_i - 2\hat{a}_1 \sum_{i=1}^{n} Y_i X_i)$$

on a :

$$\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

et

$$\bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$$

$$\frac{d\varphi(\hat{a}_0, \hat{a}_1)}{d\hat{a}_0} = 2n\hat{a}_0 + 2\hat{a}_1 \sum_{i=1}^{n} X_i - 2 \sum_{i=1}^{n} Y_i = 0$$

$$\frac{d\varphi(\hat{a}_0, \hat{a}_1)}{d\hat{a}_1} = 2\hat{a}_1 \sum_{i=1}^{n} X_i^2 + 2\hat{a}_0 \sum_{i=1}^{n} X_i - 2 \sum_{i=1}^{n} Y_i X_i = 0$$

pour $\hat{a}_0$:

$$2n\hat{a}_0 + 2\hat{a}_1 \sum_{i=1}^{n} X_i - 2 \sum_{i=1}^{n} Y_i = 0$$

$$2\hat{a}_1 \sum_{i=1}^{n} X_i - 2 \sum_{i=1}^{n} Y_i = -2n\hat{a}_0$$

$$2nE(Y) - 2n\hat{a}_1 E(X_i) = 2n\hat{a}_0$$

donc

$$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X} \tag{1.8}$$

Pour $\hat{a}_1$:

$$\hat{a}_1 = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} (X_i Y_i - X_i \bar{Y} + Y_i \bar{X} - \bar{X}\bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} X_i Y_i - n\,\bar{X}\,\bar{Y}}{\sum_{i=1}^{n} (X_i - \bar{X})^2}.$$

donc

$$\hat{a}_1 = \frac{\sum_{i=1}^{n} X_i Y_i - n\,\bar{X}\,\bar{Y}}{\sum_{i=1}^{n} (X_i - \bar{X})^2}. \tag{1.9}$$

1.4.3 Propriétés des estimateurs

— $\hat{a}_1$ est l'estimateur de a_1.

— $\hat{a}_0$ est l'estimateur de a_0.

— Si $E(\hat{a}_1) = a_1$ alors l'estimateur $\hat{a}_1$ est estimateur sans biais de a_1.

— Si $E(\hat{a}_0) = a_0$ alors l'estimateur $\hat{a}_0$ est estimateur sans biais de a_0.

Sont ils converegents ?

— Si l'estimateur $\hat{a}_1$ est sans biais, pour qu'il sont convergent il suffit de montrer que $Var(\hat{a}_1) \rightarrow_{n\rightarrow\infty} 0$.

Calcul l'esperance $E(\hat{a}_1)$:

Nous avons

$$\hat{a}_1 = \frac{\sum_{i=1}^{n} (\mathbf{X}_i - \bar{X})(\mathbf{Y}_i - \bar{\mathbf{Y}})}{\sum_{i=1}^{n} (\mathbf{X}_i - \bar{X})^2}$$

on pose

$$x_i = X_i - \bar{X} \tag{1.10}$$

$$\hat{a}_1 = \frac{\sum_{i=1}^{n} x_i (Y_i - \bar{Y})}{\sum_{i=1}^{n} x_i^2}$$

on pose

$$C_i = \frac{x_i}{\sum_{i=1}^{n} x_i^2} \tag{1.11}$$

$$\hat{a}_1 = \sum_{i=1}^{n} C_i (Y_i - \bar{Y}) = \sum_{i=1}^{n} C_i Y_i - \sum_{i=1}^{n} C_i \bar{Y} \tag{1.12}$$

sachant que

$$Y_i = a_0 + a_1 X_i + \varepsilon_i, i = 1, ..., n$$

$$\hat{a}_1 = \sum_{i=1}^{n} C_i (a_0 + a_1 X_i + \varepsilon_i) - \sum_{i=1}^{n} C_i \bar{Y} = a_0 \sum_{i=1}^{n} C_i + a_1 \sum_{i=1}^{n} C_i X_i + \sum_{i=1}^{n} C_i \varepsilon_i$$

$$\sum_{i=1}^{n} C_i = \frac{\sum_{i=1}^{n} x_i}{\sum_{i=1}^{n} x_i^2} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} X_i - \sum_{i=1}^{n} \bar{X}}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{n\bar{X} - n\bar{X}}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = 0$$

$$\sum_{i=1}^{n} C_i X_i = \frac{\sum_{i=1}^{n} (X_i - \bar{X}) X_i}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} X_i^2 - \bar{X} \sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = 1$$

sachant que

$$\sum_{i=1}^{n}(X_i-\bar{X})^2=\sum_{i=1}^{n}X_i^2-n\bar{X}^2$$

$$\hat{a}_1=a_1+\sum_{i=1}^{n}C_i\varepsilon_i=a_1+\frac{\sum\limits_{i=1}^{n}x_i\varepsilon_i}{\sum\limits_{i=1}^{n}x_i^2}$$

$$E(\hat{a}_1)=E(a_1)+\frac{E(\sum\limits_{i=1}^{n}x_i\varepsilon_i)}{\sum\limits_{i=1}^{n}x_i^2}=a_1+\frac{1}{\sum\limits_{i=1}^{n}x_i^2}E(\sum_{i=1}^{n}x_i\varepsilon_i)=a_1+[\frac{1}{\sum\limits_{i=1}^{n}x_i^2}\sum_{i=1}^{n}E(x_i\varepsilon_i)=0]$$

$$E(\hat{a}_1)=a_1 \tag{1.13}$$

$\hat{a}_1$ est un estimateur sans biais de a_1.

Remarque 3 *espérance de constante est un constante et la variance de constante est 0.*

Calcul la variance de $Var(\hat{a}_1)$

$$\begin{aligned} Var(\hat{a}_1) &= (Var(a_1)=0)+Var(\sum_{i=1}^{n}C_i\varepsilon_i) \\ &= \sum_{i=1}^{n}Var(C_i\varepsilon_i)=\sum_{i=1}^{n}C_i^2Var(\varepsilon_i)=\sum_{i=1}^{n}C_i^2\sigma_\varepsilon^2=\frac{\sum\limits_{i=1}^{n}x_i^2}{(\sum\limits_{i=1}^{n}x_i^2)^2}\sigma_\varepsilon^2 \end{aligned}$$

$$Var(\hat{a}_1)=\frac{\sigma_\varepsilon^2}{\sum\limits_{i=1}^{n}(X_i-\bar{X})^2} \tag{1.14}$$

$$\lim_{n\longrightarrow+\infty}Var(\hat{a}_1)=\lim_{n\longrightarrow+\infty}\frac{\sigma_\varepsilon^2}{\sum\limits_{i=1}^{n}(X_i-\bar{X})^2}=0$$

donc $\hat{a}_1$ est un estimateur convergent.

Pour la construction des testes :

Nous avons besoin de calculer une estimateur de σ_ε^2

$$\sigma_\varepsilon^2=\frac{\sum\limits_{i=1}^{n}e_i^2}{n-k}=\frac{\sum\limits_{i=1}^{n}e_i^2}{n-2} \tag{1.15}$$

k : nombre de paramètre estimées.

$\hat{\sigma}_\varepsilon^2$: est estimateur sans biais de σ_ε^2.

L'estimateur sans biais de la variance de $\hat{a}_1$ c'est $\hat{\sigma}^2_{\hat{a}_1}$

$$\sigma^2_{\hat{a}_1} = Var(\hat{a}_1) = \frac{\sigma^2_\varepsilon}{\sum_{i=1}^{n}(X_i - \bar{X})^2}$$

$$\hat{\sigma}^2_{\hat{a}_1} = \frac{\sigma^2_\varepsilon}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\sum_{i=1}^{n} e_i}{(n-2)(\sum_{i=1}^{n}(X_i - \bar{X})^2)} \tag{1.16}$$

$$\hat{\sigma}^2_{\hat{a}_0} = \frac{\sigma^2_\varepsilon \sum_{i=1}^{n} X_i^2}{n(\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{1.17}$$

1.4.4 Test de la qualité d'un ajustement

1.4.4.1 Equation d'analyse de la variance :

Nous montrons les deux égalités suivantes :

a) La somme des résidus est nulle

$$\sum_{i=1}^{n} e_i = 0 \tag{1.18}$$

ona :

$$e_i = Y_i - \hat{Y}_i = Y_i - (\hat{a}_0 + \hat{a}_1 X_i)$$

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} Y_i - \sum_{i=1}^{n} \hat{a}_0 - \hat{a}_1 \sum_{i=1}^{n} X_i$$

$$= n\bar{Y} - n\hat{a}_0 - \hat{a}_1 n\bar{X}$$

$$\frac{1}{n}\sum_{i=1}^{n} e_i = \bar{Y} - \hat{a}_0 - \hat{a}_1 \bar{X}$$

on sait que

$$\hat{a}_0 = \bar{Y} - \hat{a}_1 \bar{X}$$

donc

$$\frac{1}{n}\sum_{i=1}^{n} e_i = \bar{Y} - (\bar{Y} - \hat{a}_1 \bar{X}) - \hat{a}_1 \bar{X} = 0$$

$\Rightarrow$

$$\sum_{i=1}^{n} e_i = 0$$

b)

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \hat{Y}_i \tag{1.19}$$

ona

$$\sum_{i=1}^{n} e_i = 0 = \sum (Y_i - \hat{Y}_i) = 0$$

⇒

$$\sum_{i=1}^{n} Y_i = \sum_{i=1}^{n} \hat{Y}_i$$

à partir de ces deux égalités nous avons l'équation d'analyse de la variance suivante

$$\begin{aligned} \sum_{i=1}^{n} (Y_i - \bar{Y})^2 &= \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y}_i)^2 + \sum_{i=1}^{n} e_i^2 \\ SCT &= SCE + SCR \end{aligned} \tag{1.20}$$

Variabilité totale(SCT) = Variabilité expliquer (SCE) + Variabilité résidus(SCR).

1.4.4.2 Coefficient de détermination

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = \frac{SCE}{SCT} = 1 - \frac{\sum_{i=1}^{n} e_i^2}{SCT} \tag{1.21}$$

TABLE 1.1 – **Tableau d'analyse de la variance**

Source de variation	Somme des carrées	degrés de la liberté (d.d.l)	carrés moyen
X	$SCE = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$	$k-1$	$SCE/1$
Residu	$SCR = \sum_{i=1}^{n} e_i^2$	$n-2$	$SCR/n-2$
Total	$SCT = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$	$n-1$	/

ddl : le nombre de valeurs qu'on peut choisir de maniere arbitraire.

1.4.5 Tests

1.4.5.1 Test de Student

Nous allons tester l'hypothése suivante

$$H_0 : a_1 = 0 \& H_1 : a_1 \neq 0$$

— Calcules

$$t^* = \frac{|\hat{a}_1 - a_1|}{\hat{\sigma}_{\hat{a}_1}} \tag{1.22}$$

$\rightsquigarrow$ suit la loi de Student à $(n-2)$ d.d.l.

— Selon le seuil $\alpha = 5\%$ lu la valeur tableur $t_{n-2}^{\alpha/2}$. Si

$$t^* > t_{n-2}^{\alpha/2}$$

on rejette l'hypothèse H_0. a_1 est significativement différent de 0, d'où la variable X est une variable explicative de la variable Y.

Test de signification globale

Hypothése : $H_0 : \ a_1 = 0 \& H_1 : \ a_1 \neq 0$

Calcul :

$$F^* = \frac{S.C.E/1}{S.C.R/n-2} = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} e_i^2 / n-2} \tag{1.23}$$

$\rightsquigarrow$suit la loi de Fisher à 1 et $(n-2)$ d.d.l. Selon le seuil $\alpha = 5\%$ lu sur le table de Fisher $F_{1,n-2}^{\alpha}$.

— Si $F^* > F_{1,n-2}^{\alpha}$ on rejette l'hypothése H_0 (on accepte l'hypothése H_1).

La régression est globalement significative.

1.5 Régression linéaire multiple

Le modèle

$$Y = a_0 + a_1 X_1 + ... + a_p X_p + \varepsilon \tag{1.24}$$

pour n observation :

$$Y_i = a_0 + a_1 X_{1i} + ... + a_p X_{pi} + \varepsilon_i$$

$i = 1, 2, ..., n$

$$\begin{aligned}
Y_1 &= a_0 + a_1 X_{11} + a_2 X_{21} + ... + a_p X_{P1} + \varepsilon_1 \\
Y_2 &= a_0 + a_1 X_{12} + a_2 X_{22} + ... + a_p X_{p2} + \varepsilon_2 \\
&\ \ \vdots \\
Y_n &= a_0 + a_1 X_{1n} + a_2 X_{2n} + ... + a_p X_{pn} + \varepsilon_n
\end{aligned}$$

Ecriture matricielle :

$$Y = aX + \varepsilon$$

$$Y_{(n,1)} = \begin{pmatrix} Y_1 \\ Y_2 \\ . \\ . \\ . \\ Y_n \end{pmatrix} , X_{(n,\, p+1)} = \begin{pmatrix} 1 & X_{11} & X_{21} & . & . & X_{p1} \\ 1 & X_{12} & X_{22} & . & . & X_{p2} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 1 & X_{1n} & X_{2n} & . & . & X_{pn} \end{pmatrix}, a_{(p+1,1)} = \begin{pmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_p \end{pmatrix}, \varepsilon_{(n,1)} = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{pmatrix}$$

Hypothèses :

— H_1 : les valeurs X_{it} sont observés sans erreurs (non aléatoires).

— H_2 : la moyenne le modèle est bien spécifié $E(\varepsilon_t) = 0, \ \forall t = 1,2,...,n.$

— H_3 : la variance de l'erreur est constante

$$Var(\varepsilon_t) = \sigma_\varepsilon^2 \quad ; \forall t = 1,2,...,n.$$

— H_4 : les erreurs sont non-corrélées

$$E(\varepsilon_t \varepsilon_{\grave{t}}) = 0 \qquad ; \forall t \neq \grave{t}$$

— H_5 : l'erreur est indépandanté des variables explicatives

$$E(\varepsilon_t X_{it}) = 0 \ ; \quad \forall t = 1,2,...,n$$

— H_6 : absence de colineairité entre les variables explicative.

— H_7 : $({}^tXX)/n \rightsquigarrow$matrice finie.

H_8 : $n > p$ le nombre d'absorvations et suprieur au nombre des variables éxogènes.

1.5.1 Estimation des coefficients

Pour estimer les coefficients $a_0, a_1, .., a_p$, nous allons utiliser le crétere des moidres carrés ordinaire qui consiste à minimiser $(\sum_{i=1}^{n} \varepsilon_i^2)$.

$$Y = Xa + \varepsilon \Rightarrow \varepsilon = Y - Xa \tag{1.25}$$

$$\min \sum_{i=1}^{n} \varepsilon_i^2 = \min({}^t\varepsilon\varepsilon) = \min(\underbrace{{}^tYY - 2{}^ta{}^tXY + {}^ta{}^tXXa}) = \min\varphi(a)$$

$$\begin{aligned} {}^t\varepsilon\varepsilon &= {}^t(Y-Xa)(Y-Xa) = {}^tYY - {}^tYXa - {}^ta{}^tXY + {}^ta{}^tXXa \\ &= {}^tYY - 2{}^ta{}^tXY + {}^ta{}^tXXa \end{aligned}$$

on a

$${}^tYXa = {}^ta{}^tXY$$

condition nécessaire

$$\frac{d\varphi(a)}{da} = -2{}^tXY + 2{}^tXXa = 0$$

$$({}^tXX)a = {}^tXY \Rightarrow ({}^tXX)^{-1}({}^tXY)\hat{a}$$

$$\hat{a} = ({}^tXX)^{-1t}XY \tag{1.26}$$

1.5.2 Propriétés des estimateur

Le modèle sous forme matricialle

$$Y = aX + \varepsilon$$

et ona

$$\hat{Y} = \hat{a}X \tag{1.27}$$

$$\begin{aligned} \hat{a} &= ({}^tXX)^{-1t}XY = ({}^tXX)^{-1t}X(Xa+\varepsilon) \\ &= (X{}^tX)^{-1}XXa + ({}^tXX)^{-1t}X\varepsilon \\ &= a + ({}^tXX)^{-1t}X\varepsilon \end{aligned}$$

$$\begin{aligned} E(\hat{a}) &= E(a) + E(({}^tXX)^{-1t}X\varepsilon \\ &= a + (({}^tXX)^{-1t}XE(\varepsilon) = 0) \\ &= a \end{aligned} \tag{1.28}$$

$\hat{a}$ est un estimateur sans biais de a.

Matrice des variance covariance $\Omega_{\hat{a}}$ des coefficients :

$$\begin{aligned}\Omega_{\hat{a}} &= E[{}^t(\hat{a}-a)(\hat{a}-a)] \\ &= E[(({}^tXX)^{-1t}X\varepsilon)^t(({}^tXX)^{-1t}X\varepsilon) \\ &= E(({}^tXX)^{-1t}X\varepsilon^t\varepsilon X({}^tXX)^{-1}) \\ &= ({}^tXX)^{-1t}XE(\varepsilon^t\varepsilon)X({}^tXX)^{-1}\end{aligned}$$

Ω_ε matrice des variances covariance de l'erreur ε.

$$\Omega_\varepsilon = E(\varepsilon\varepsilon^t) = \begin{pmatrix} E(\varepsilon_1^2) & E(\varepsilon_1\varepsilon_2) & . & . & E(\varepsilon_1\varepsilon_n) \\ E(\varepsilon_1\varepsilon_2) & E(\varepsilon_2^2) & . & . & E(\varepsilon_2\varepsilon_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ E(\varepsilon_1\varepsilon_n) & E(\varepsilon_2\varepsilon_n) & . & . & E(\varepsilon_n^2) \end{pmatrix} = \begin{pmatrix} Var(\varepsilon_1) & Cov(\varepsilon_1\varepsilon_2) & . & . & Cov(\varepsilon_1\varepsilon_n) \\ Cov(\varepsilon_1\varepsilon_2) & Var(\varepsilon_2) & . & . & Cov(\varepsilon_2\varepsilon_n) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ Cov(\varepsilon_1\varepsilon_n) & Cov(\varepsilon_2\varepsilon_{n)} & . & . & Var(\varepsilon_n) \end{pmatrix}$$

$$Var(\varepsilon_i) = \sigma_\varepsilon^2 \qquad , \forall i = 1,2,...,n \ .....H_3$$

$$Cov(\varepsilon_i\varepsilon_j) = 0 \qquad , \forall i \neq j \ .....H_4$$

$$\Omega_\varepsilon = \begin{pmatrix} \sigma_\varepsilon^2 & 0 & 0 & . & . & 0 \\ 0 & \sigma_\varepsilon^2 & 0 & . & . & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 & . & 0 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ 0 & 0 & . & . & 0 & \sigma_\varepsilon^2 \end{pmatrix} = \sigma_\varepsilon^2 I$$

$$\begin{aligned}\Omega_{\hat{a}} &= ({}^tXX)^{-1t}X\Omega_\varepsilon X({}^tXX)^{-1} \\ &= \sigma_\varepsilon^2 I({}^tXX)^{-1}\end{aligned} \tag{1.29}$$

L'estimateur sans biais de σ_ε^2 est

$$\hat{\sigma}_\varepsilon^2 = \sum_{i=1}^{n} e_i(n-(p+1)) = \frac{{}^tee}{n-(p+1)} \tag{1.30}$$

donc

$$\hat{\Omega}_{\hat{a}} = \hat{\sigma}_{\varepsilon}^2 ({}^tXX)^{-1} = \frac{{}^tee}{n-p-1} ({}^tXX)^{-1} \rightarrow_{n\rightarrow+\infty} 0$$

$\hat{a}$ est un estimateur convergent.

1.5.3 Analyse de la variance

$$\sum_{i=1}^{n} e_i = 0$$

et

$$\sum_{i=1}^{n} \hat{Y}_i = \sum_{i=1}^{n} Y_i$$

L'equation analyse de variance

$$\begin{aligned} \sum_{i=1}^{n} (Y_t - \bar{Y})^2 &= \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^{n} e_i^2 \\ SCT &= SCE + SCR \end{aligned}$$

Coefficient de détermination :

$$R^2 = \frac{\sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2} = 1 - \frac{\sum_{i=1}^{n} e_i^2}{\sum_{i=1}^{n} (Y_i - \bar{Y})^2}$$

TABLE 1.2 – Tableau d'analyse de la variance pour modèle de régression multiples

La source de variation	Somme des carrées	d.d.l	moyen
X	$SCE = \sum_{i=1}^{n} (\hat{Y}_i - \bar{Y})^2$	p	SCE/p
résidu	$SCR = \sum_{i=1}^{n} e_i^2$	$n-(p+1)$	$SCR/n-(p-1)$
total	$SCT = \sum_{i=1}^{n} (Y_i - \bar{Y})^2$	$n-1$	/

1.5.4 Construction des tests

Le test de signification globale d'une régression tester si l'ensemble des p variables exogènes a une influence sur la variable Y.

Hypothèse

$H_0 : a_1 = a_2 = ... = a_p = 0$ (tous les coefficient sont nulles)

H_1 : (au moins un des coefficient est différent de 0.)

$$F^* = \frac{SCE/p}{SCR/n-(p+1)} \tag{1.31}$$

$\rightsquigarrow$ sont la loi de Fisher a p et $n-(p+1)$.

Si

$$F^* > F^{\alpha}_{p,n-p-1}$$

(au seuil de α) alors nous rejetons l'hypothése H_0 donc la régression est globalement significative.

1.5.5 Comparaison d'un coefficient de régression a une valeur fixe

Tester l'hypothèse suivante :

$$H_0 : a_i = \bar{a} \& H_1 : a_i \neq \bar{a}$$

$$t^*_{a_i} = \frac{|a_i - \bar{a}|}{\hat{\sigma}_{a_i}}$$

$\rightsquigarrow$ suit la loi de Student a $(n-p-1)$d.d.l.

Si

$$t^*_{a_i} > t^{\alpha/2}_{n-p-1}$$

(au seuilα) nous rejetons l'hypothèse H_0.

Donc a_i est significativement .

1.5.5.1 Comparaison d'un ensemble de coefficient a un ensemble de valeur fixe

tester l'hypothèse suivante :

a_q : vecteur de coefficient de dimension.

q : le nombre de variable a tester .

$$H_0 : a_q = \bar{a}_q \quad \& H_1 : a_q \neq \bar{a}_q.$$

$$F^* = \frac{1}{q}\,{}^t(\hat{a}_q - \bar{a}_q)\hat{\Omega}_{\hat{a}_q}\hat{a}_q^{-1}(\hat{a}_q - \bar{a}_q) \tag{1.32}$$

$\rightsquigarrow$suit la loi de fisher à q et $(n-p-1)$ ddl.

Exemple 4 $q = 2$, $\hat{a}_q = \begin{pmatrix}\hat{a}_1\\ \hat{a}_2\end{pmatrix}$, $\bar{a}_q = \begin{pmatrix}1.5\\ 2\end{pmatrix}$

$$\hat{a}_q - \bar{a}_q = \begin{pmatrix}\hat{a}_1 - 1.5\\ \hat{a}_2 - 2\end{pmatrix}$$

$\hat{\Omega}_{\hat{a}_q}$ sous matrice de la variance covariance des coefficients qui correspondes .
Si

$$F^* > F^{\alpha}_{q,(n-p-1)}$$

alors rejetons l'hypothése H_0 les coefficients testes soit significativement $\neq$ de $\bar{a}_0$.
Coefficient de détermination corrigé :

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-p-1} \tag{1.33}$$

où

$$R^2 = \frac{SCE}{SCT}. \tag{1.34}$$

1.6 Régression logistique

Définition 5 *la régression logistique ou modèle logit est un modèle de régression binomaile. Il s'agit de modèliser mieux un modèle mathématique simple à des observations réeles nombreusses.*

Application :

★ En médecine, elle permet par exemple de trouver les facteurs qui caractérisient un groupe des sujets malades par rapport à des sains.

★ Dans le domaine des assurances, elle permet de cibler une fraction de clientéle qui sera sensible à une police d'assurance sur tel risque particulier.

★ Dans le domaine bancaire, pour détecter les groupe à risque lors de la souscription d'un crédit.

★ En économétrie, pour expliquer une variable discréte. Par exemple, les intentions de vote aux élections.

1.6.1 Critère d'information d'Akaike

Le critère d'information d'akaike en anglais Akaike information criterion ou (AIC) et proposée par (cf. Hirotugu Akaike, (1973)) s'écrit comme suit :

$$AIC = 2k - 2\ln(L)$$

oū K est le nombre de paramètres à estimer du modèle et L est le maximum de la fonction de vraisemblances du modèle.

Si l'on concideré un ensemble de modéle candidats, le modèle choisi est celui qui aura la plus faible valeur d'AIC.

Utilisation paratique de l'AIC : supposons disposer d'un ensemble de modèle - candidats, dont on calcule les valeurs d'AIC associées. Il y aura toujours une perte d'information, du fait qu'on utilise un modèle pour représenter le processus générant les données réelles, et nous cherchons donc à sélectionner le modèle qui minimise cette perte d'information.

Notion les diverses valeurs d'AIC des différents modèles AIC_1, AIC_2, AIC_3,...,AIC_R et $AIC_{\min}$ le minimum de ces valeurs.

Dans ce cas

$$\exp((AIC_{\min} - AIC_i)/2)$$

est proportionnel à la probabilité que iéme candidat modèle minimise l'estimation de la perte d'information.

par exemple, supposons qu'il y ait trois modèle candidats avec pour AIC respectives : 100, 102 et 110. Dans ce cas, la probabilité pour que le second modèle soit celui qui minimise la perte d'information est de

$$\exp((100-102)/2) = 0.368$$

fois la probabilité pour que ce soit le premier modèle. De le même maniére, la probabilité pour que troisième modèle.

AIC_c est une correction de l'AICpour le cas d'échantillons de petite taille :

$$AIC_C = AIC + \frac{2k(k+1)}{n-k-1}$$

oũ désigne la taille de l'échontillon.

Comparaison au BIC il existe de nombreux critéres d'informations inspirés du critére d'Akaike. Parmi ceux -ci, le critére d'information bayésien est l'un des plus populaires. Il se définit comme suit :

$$BIC = k\ln(n) - 2\ln(L)$$

avec n le nombre d'observations dans l'échontillon étudié et k le nombre de paramètres.

L'AIC pénalise le nombre de paramètres moins fortement que le BIC (cf. Burnhamet Anderson, (2004)) proposent une comparaison de l'AIC au BIC. Les auteurs montrent que l'AIC et l'AIC$_C$ peuvent être construits dans le même contexte bayésien que le BIC, mais avec des hypothéses différentes.

1.7 Régression de poisson

Est un modèle linéaire généralisé utilisé pour les données de comptage et les tableaux de contingence.

1.7.1 Distibution de la loi de poisson

$$P(Y = k) = \frac{\exp(-\lambda)\lambda^k}{k!}, k = 0, 1, 2, ..., n \tag{1.35}$$

λ est le paramètre de la loi de poisson, où

$$E(Y) = \lambda \text{ et } var(y) = \lambda$$

1.7.2 Modèle de régression de Poisson

$$\ln(Y) = \alpha + \beta_1 X_1 + ... + \beta_i X_i$$

α est estimateur et β vecteur des coefficients et $i = 1, 2, ..., n$.
Exemples d'applications :

- Nombre de naissance par césarienne public/privé.
- Nombre de PV de stationnement Paris/ Rennes.
- Lien entre cylindré d'un véhicule et son nombre de PV.

CHAPITRE 2

DISTRIBUTION DE TWEEDIE

Sommaire

2.1 INTRODUCTION

Les distributions de Tweedie sont très connues et utiles dans plusieurs domaines de recherche tels que l'analyse de survie, les études de dépenses et de consommation, de l'écologie et de la météorologie. La distribution de Tweedie est particulièrement utilisée dans la recherche actuarielle, plus précisément dans la modélisation des réclamations d'assurance.

Ce chapitre, porte sur les distributions de Tweedie qui appartiennent à la famille de modèles de dispersion exponentielle, qui ont des fonctions de variance d'unité de puissance, sont infiniment divisibles et sont fermées sous les trudictions et les transformations d'échelle. Notamment, une variable aléaitoire de Tweedie a un paramètre d'indexation / puissance qui essentiel pour d'écrire sa distribution. Les actuaires définissent généralement ce paramètre sur une valeur par défaut, tandis que le package tweedie R fournit des outils pour estimer la puissance de Tweedie via une estimation du maximum de vraisemblance, cette estimation et testée sur des simulations et appliquée à un ensemble de données sur le coût des pertes d'habitations. Les modèles construits avec une puissance de Tweedie estimée respectent le critère d'information AKaiKe inférieur par rapport aux modèles construits avec les pouvoirs Tweedie par défaut. Cependant ce réglage de paramètre ne modifie que marginalement les coefficients de régression et les prédictions du modèle. Compte tenu des contraintes de temps, nous recommandons aux actuaires d'utiliser les pouvoirs de Tweedie par défaut et d'envisager une autre ingénierie des fonctionnalités.

2.2 Les modèles de dispersion exponentielle (EDM)

La famille de Tweedie est un sous ensemble d'une classe de variables aléaitoires décrite par (cf. Bent Jorgensen, (1997)) dans The Theory of Dispersion Modele. Par conséquent, nous devons d'abord couvrir les modèles de dispersion expenentielle (EDM) avant de discuter la famille Tweedie. Jorgensen présente deux descriptions d'EDM dans sa monographie :

Une axiomatique et une constructif. La version axiomatique définit les EDM sans justifier les origines des distributions ; La version constructive commence par une fonction cumulative et construit la théorie à partir de là (plus tard, Jorgensen prouve que la définition axiomatique correspond à la définition constructive). Ici nous fournissons la définition axiomatique pour les modèles de dispersion exponentielle. Les idées inspirent plus d'idées. Cet adage s'applique au développement de modèles de dispersion expenentielle. Les EDM maintiennent la structure de la distribution normale. Afin de

parler de cette structure en termes abstraites, Nous devons établir quelques définitions.

2.2.1 Définition axiomatique Pour les EDM

Définition 6 *Soit f une fonction de densité réelle pour la variables aléaitoire X. Le support de X est l'ensemble des éléments du domaine de f qui ne correspondent pas à zéro. Le support convexe de X est la petit intervalle contenant le support.*

Définition 7 *Soit $\mathbb{C}$ un support convexe et soit $\mathbb{D}$ l'intérieur de $\mathbb{C}$, $\mathbb{D}$ et $\mathbb{C}$ sont des intervalles statisfiant que $\mathbb{D} \subset \mathbb{C} \subset \mathbb{R}$. Une déviance unitaire est une fonction $d : \mathbb{C} \times \mathbb{D} \longrightarrow \mathbb{R}$ qui statisfait les éléments suivants :*

- ◇ $d(y, y) = 0, \forall\, y \in \mathbb{D}$.
- ◇ $d(y, z) > 0, \forall\, y \neq z$.

Remarque 8 *Cette définition semble familière à la définition d'une métrique, sauf sans l'inégalité triangulaire. Essentiellement, la déviance unitaire est un outil pour mesurer la distance. Equipé de ces deux définitions, Nous sommes prêts à présenter la définition d'un modèle de dispersion exponentielle. Considérons la fonction de densité pour une variable aléatoire normale $\mathcal{N}(\mu, \sigma^2)$:*

$$f(y, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(y-\mu)^2\right\} . \tag{2.1}$$

Les fonctions de densité pour les modèles de dispersion exponentielle partagent ce format .

Définition 9 *Un modèle de dispersion exponentielle $EDM(\mu, \sigma^2)$ est une distribution de porobabilité dont la fonction de densité par rapport à une mesure appropriée a la forme*

$$f(y, \mu, \sigma^2) = a(y; \sigma^2) \exp\left\{-\frac{1}{2\sigma^2} d(y, \mu)\right\},\ y \in \mathbb{C} . \tag{2.2}$$

où $a \geqslant 0$ est une fonction appropriée, d est une déviance unitaire de la forme $d(y, \mu) = yg(\mu) + h(\mu) + k(y)$, $\mathbb{C}$ est le support convexe, $\mu \in \mathbb{D} = int(\mathbb{C})$ et $\mathbb{D}$ est un intervalle

Remarque 10 *Nous appelons μ le paramètre de position et σ^2 le paramètre de dispersion. Ce langage et cette notation s'inspirent de la théorie normale. La figure 2.1 montre la courbe en cloche normale, μ est l'endroit où se trouve le centre de masse de la distribution, σ^2 décrit la répartition de la masse de la distribution.*

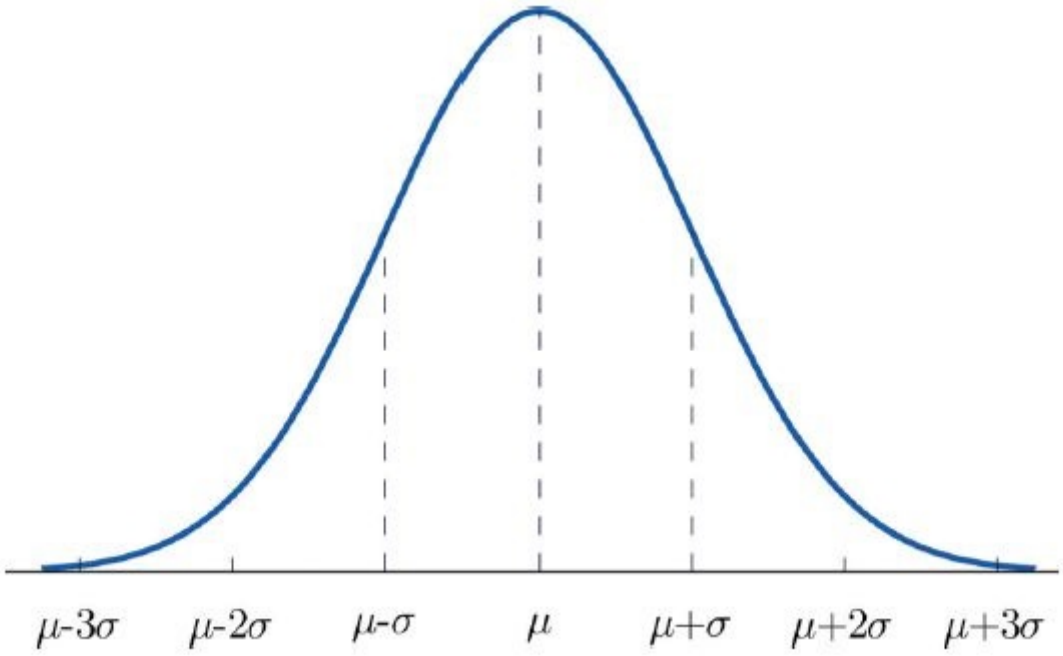

FIGURE 2.1 – Interprétation de la courbe en cloche de μ et σ^2

Proposition 11 *Les distributions suivantes sont des modèles de dispersion exponentielle :*

- Distribution Normale.
- Distribution de Poisson.
- Distribution Binomiale.
- Distribution Gamma.

Preuve. Pour montrer qu'une distribution est un EDM, nous proposons d'abord $a(y,\sigma^2)$ et $d(y,\mu)$. Ensuite, nous soutenons que la déviance unitaire d de la forme correcte. Enfin, nous faisons l'algèbre nécessaire pour montrer que $f(y,\mu,\sigma^2)$ correspond à la fonction de densité habituelle de la distribution. Consédirons une variable aléatoire $\mathcal{N}(\mu,\sigma^2)$ avec $\mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}_+$ et $y \in \mathbb{R}$. Soit $a(y,\sigma^2) = \frac{1}{\sqrt[2]{2\pi\sigma^2}}$ et $d(y,\mu) = (y-\mu)^2$.

Remarquez que

$$d(y,\mu) = (y-\mu)^2 = y^2 - 2y\mu + \mu^2$$

Définissez $g(x) = -2x$, $h(x) = x^2$ et $k(x) = x^2$. La déviance unitaire est de la bonne forme et $f(y,\mu,\sigma^2)$ correspond à la densité normale habituell .

Considérons une variable aléatoire de poisson qui définie par $\mu \in \mathbb{Z}_+$ et $y \in \mathbb{Z}_{0+}$. Soit $a(y,\sigma^2) = \frac{y^y}{y!}\exp(-y)$ et $d(y,\sigma^2) = 2(y\log\frac{y}{\mu} - y + \mu)$. Ici, la fonction $a(y,\sigma^2)$ est indépendante du paramètre de dispersion σ^2. Pour une variable aléatoire de poisson, $\sigma^2 = 1$. Notez que

$$d(y,\mu) = 2\left(y\log(\frac{y}{\mu}) - y + \mu\right) = 2y\log(y) - 2y\log(\mu) - 2y + 2\mu\,.$$

Définissez

$$g(x) = -2\log x\,,\, h(x) = 2x$$

et

$$K(x) = 2x\log(x) - 2x$$

$$\begin{aligned} f(y,\mu,1) &= a(y)\exp\left\{\frac{-1}{2}2\left(y\log(\frac{y}{\mu}) - y + \mu\right)\right\} = a(y)\exp\left\{y\log(\mu) - y\log(y) + y - \mu\right\} \\ &= a(y)e^{y\log\mu}e^{-y\log y}e^{y}e^{-\mu} = a(y)y^{\mu}y^{-y}e^{y}e^{-\mu} = \frac{1}{y!}\ y^{\mu}e^{-\mu}y^{-y}y^{y}e^{y}e^{-y} = \frac{1}{y!}y^{\mu}e^{-\mu}. \end{aligned}$$

ainsi, $f(y,\mu,\sigma^2)$ correspond à la fonction de masse de probabilité de poisson habituelle.

Pour prouver qu'une variable aléatoire binomiale est un EDM. Nous ajustons légèrement notre notation. En pratique, les variables aléatoires binomiales se produisent lorsque nous avons un événment avec deux résultats possibles et que nous voulons savoir la probabilité qu'un résultat se produise m fois sur n tentatives indépendantes. Supposons que l'espace de probabilité de la variable aléatoire X soit $\{0,1\}$ et définissons $pr(x=1) = p$. Remlacez y par m, μ par p et σ^2par n. De plus, nous connaissons le paramètre n car c'est le nombre de fois que nous exécutons l'expérience. Fixe n comme un entier positif . Considérons une variable aléatoire binomiale avec $p \in (0\ ,\ 1)$, $m \in \mathbb{Z}_{0+}$ et $n \in \mathbb{Z}_{+}$. Laisser $a(m,n) = \binom{n}{m}$ et

$$d(m,p) = -2n\left(m\log\left(\frac{p}{1-p}\right) + n\log\left(1-p\right)\right)\,.$$

Remarquer que

$$d(m,p) = -2n\left(m\log\left(\frac{p}{1-p}\right) + n\log\left(1-p\right)\right)\ = -2nm\log(\frac{p}{1-p}) - 2n^2\log(1-p).$$

Définissez $g(x) = -2n\log\left(\frac{x}{1-x}\right)$, $h(x) = -2n^2\log(1-x)$, et $k(x) = 0$. La déviance unitaire et de la bonne forme. Maintenant,

$$\begin{aligned} f(m,p,n) &= a(m,n)\exp\left\{\frac{-1}{2n}.(2n)\left(m\log\left(\frac{p}{1-p}\right) + n\log\left(1-p\right)\right)\right\} \\ &= \binom{n}{m}\exp\left\{\log\left(\frac{p}{1-p}\right) + n\log\left(1-p\right)\right\} \\ &= \binom{n}{m}\exp\left\{m\log p + (n-m)\log\left(1-p\right)\right\}\ = \binom{n}{m}p^{m}\left(1-p\right)^{n-m}. \end{aligned}$$

Ainsi, $f(y,\mu,\sigma^2)$ correspond à la fonction de masse de probabilité binomiale habituelle. Nous ajustons à nouveau la notation pour prouver qu'une variable aléatoire

gamma est un EDM. Conventionnellement, une variable gamma aléaitoire est paramètrée par le paramètre de forme α et le paramètre de vitesse β. La densité est

$$f(y,\alpha,\beta) = \frac{\beta^{\alpha} y^{\alpha-1} e^{\beta y}}{\Gamma(\alpha)}. \tag{2.3}$$

où fonction gamma est définie comme

$$\Gamma(x) = \int_0^{\infty} y^{x-1} e^{-y} dy. \tag{2.4}$$

Définissez $\mu = \alpha/\beta$ et $\sigma^2 = 1/\alpha$. $\alpha > 0$ et $\beta > 0$ donc $\mu > 0$. et $\sigma^2 > 0$. y prend les valeur dans $\mathbb{R}_+$. Soit $a(y,\sigma^2) = a(y,1/\alpha) = \frac{1}{\Gamma(\alpha)}\alpha^{\alpha} e^{-\alpha} y^{-1}$ et $d(y,\mu) = 2\left(\frac{y}{\mu} - \log\left(\frac{y}{\mu}\right) - 1\right)$. Nous omettons l'argument selon lequel la déviance unitaire d est sous la forme correcte. Ilimte précident arguments. Maintenant,

$$\begin{aligned} f(y,\mu,\sigma^2) &= f(y,\frac{\alpha}{\beta},\frac{1}{\alpha}) = a(y,\frac{1}{\alpha})\exp\left\{\frac{-\alpha}{2}.2\left(y\frac{\beta}{\alpha} - \log\left(\frac{y\beta}{\alpha}\right) - 1\right)\right\} \\ &= a(y,\frac{1}{\alpha})\exp\left\{-y\beta + \alpha\log(y\beta) - \alpha\log(\alpha) + \alpha\right\} \\ &= \frac{1}{\Gamma(\alpha)}\alpha^{\alpha} e^{-\alpha} e^{-\alpha}(y\beta)^{\alpha} y^{-1} e^{-y\beta} = \frac{1}{\Gamma(\alpha)}\beta^{\alpha} y^{\alpha-1} e^{-y\beta}. \end{aligned}$$

Ainsi, $f(y,\mu,\sigma^2)$ correspond à la fonction de densité Gamma habittuelle. ■

Notation 12 *Les équations $\mu = \frac{\alpha}{\beta}$ et $\sigma^2 = \frac{1}{\alpha}$ sont utiles si vous utilisez à la fois les fonctions de Tweedie et Gamma dans les logiciel statistique R. Les fonctions Tweedie attendent les paramétres α et β.*

Remarque 13 *La preuve de cette proposition met en évidence la flexibilité de la fonction $a(y,\sigma^2)$ dans le cadre EDM. Cette flexibilité est néccessaire pour relier des distributions déffirentes dans le même cadre .*

IL est important de souligner que le paramétre de dispersion σ^2 n'est pas, en génival, égal à la variance de la variable aléatoire.

Soit $X \sim EDM(\mu,\sigma^2)$. Ce qui suit est vrai :

— $E[X] = \mu$.

— IL existe une fonction $V(\mu)$ telle que $Var(X) = \sigma^2.V(\mu)$.

$V(.)$ est appelée la fonction de variance unitaire. En supposant que nous savons ce qu'est une fonction de variance unitaire. Nous fournissons une définition formelle cidessous.

Définition 14 *La déviance untaire d est réguléire si $d(y,\mu)$ est deux foix continuelle-*

ment défférenciable par rapport à (y,μ) *sur* $\mathbb{D}\times\mathbb{D}$ *est satisfait*

$$\frac{\partial^2 d}{\partial \mu^2}(\mu,\mu) > 0, \forall \mu \in \mathbb{D}\ .$$

Définition 15 *La fonction de variance unitaire* $V : \mathbb{D} \to \mathbb{R}_+$ *d'une déviance unitaire régulière est*

$$\frac{2}{\frac{\partial^2 d}{\partial \mu^2}(\mu,\mu)}.$$

Exemple 16 $d(y,\mu) = (y-\mu)^2$ *est une déviance d'unité réguliers. L'expansion* $y^2 - 2y\mu + \mu^2$ *est clairement deux fois continuellement défférenciable par rapport à* (y,μ)*. Calculez les dérivées partielles :*

$$\begin{aligned}\frac{\partial d}{\partial \mu} &= -2y + 2\mu,\\ \frac{\partial^2 d}{\partial \mu^2} &= 2\end{aligned}$$

Par définition, $V(\mu) = 1$. Rappelons que la déviance unitaire $(y-\mu)^2$ appartient à la variable aléaitoire normale. Par conséquent, la variance d'une variable aléatoire normale est égale à la valeur de son paramètre de dispersion.

Exemple 17 *La déviance de l'unité de poisson est de* $2y\log(y) - 2y\log(\mu) - 2y + 2\mu$*. Cette unité peut être défférentiée deux fois par rapport à* (y,μ)*,* $\frac{\partial^2 d}{\partial \mu^2} = \frac{2y}{\mu^2}$*, donc*

$$\frac{\partial^2 d}{\partial \mu^2}(\mu,\mu) = \frac{2}{\mu}.$$

La fonction de variance est μ. si $\sigma^2 = 1$, comme c'est le cas lorsque la variable aléaitoire de poisson n'est pas sur-dispersée, la variance est égale à l'espérance .

IL existe un joli lemme qui offre la flixibilité du mathématicien dans le calcul de la fonction de variance.

Proposition 18 *Soit* d *une déviance unitaire régulière.*

$$\frac{\partial^2 d}{\partial y^2}(\mu,\mu) = \frac{\partial^2 d}{\partial \mu^2}(\mu,\mu) = -\frac{\partial^2 d}{\partial y \partial \mu}(\mu,\mu), \forall \mu \in \mathbb{D}. \tag{2.5}$$

Preuve. Par définition, nous savons que $d(\mu,\mu) = 0$ et $d(y,\mu) > 0\ \forall y \neq \mu$. Cela suffit pour affirmer que $d(\ .,\mu)$ a un minimum local à μ. Donc, $\frac{\partial d}{\partial \mu}(\mu,\mu) = 0$ et $\frac{\partial d}{\partial y}(\mu,\mu) = 0$.

L'ajout de "0"donne que

$$\frac{\partial d}{\partial \mu}(\mu,\mu)+\frac{\partial d}{\partial y}(\mu,\mu)=0\,.$$

Puisque d est régulier. Prenez des dérivées partielles :

$$\frac{\partial^2 d}{\partial \mu^2}(\mu,\mu)+\frac{\partial^2 d}{\partial \mu \partial y}(\mu,\mu)=0$$

$$\frac{\partial^2 d}{\partial y^2}(\mu,\mu)+\frac{\partial^2 d}{\partial y \partial \mu}(\mu,\mu)=0$$

Soustrez par $\frac{\partial^2 d}{\partial \mu \partial y}(\mu,\mu)$ Donc

$$\frac{\partial^2 d}{\partial y^2}(\mu,\mu)=\frac{\partial^2 d}{\partial \mu^2}(\mu,\mu)=-\frac{\partial^2 d}{\partial y \partial \mu}(\mu,\mu),\forall \mu \in \mathbb{D}.$$

■

2.2.2 Définition constructive d'une EDM

Cette sous-section réitère une grande partie de ce que (cf.Jorgensen, (1997)) formalise dans the Theory of Dispersion Modèls. Je sélectionne les définitions et les théorème pour la discussion. Nous commençons avec des modèles à un paramètre. Ensuite, On complique les choses en introduisant un second paramètre.

Définition 19 *Soit ν une mesure σ-finie sur $\mathbb{R}$. Une définez la fonction cumulative $K(\theta)$ pour $\theta \in \mathbb{R}$ comme*

$$K(\theta)=\log \int e^{\theta y}\nu(dy). \tag{2.6}$$

Le domaine de la fonction cumulant $K(\theta)$ est

$$\Theta=\left\{\theta \in R\ /\ \left(\int e^{\theta y}\nu dy\right)<\infty\right\}.$$

Cette définition demande à beaucoup de lecteurs à la fois. En pratique, nous calculons rarement la fonction cumulante $K(\theta)$. Les preuves et les résultats que nous donnons comme substitut en $K(\theta)$ sont un moyen pratique d'en-capsuler cette expression analytique. Ce qui est suspect dans cette définition, c'est la mesure σ-finie ν. Nous aimerions savoir ce qu'est ν avant de continuer. Rappeler les fonctions $a(y,\sigma^2)$ présentes dans les densités EDM. Aprés la proposition, j'ai remarqué que ces fonctions $a(y,\sigma^2)$ maintiennent une flexibilité incroyable. Nous avons proposé des fonctions $a(y,\sigma^2)$ appropriées pour les cas normaux, de poisson, binomiaux et gamma. Ces candidats ne se ressemblaient pas. En géniral, ces fonctions $a(y,\sigma^2), b(y,\sigma^2)$ n'ont pas de formes fermés. Soit $b(y,\sigma^2)$ comme $a(y,\sigma^2)$ qui prend l'entrée (y,σ^2). Pour l'instant,

soit $\sigma^2 = 1$.

$$\nu dy = b(y,1)dy\ ,$$

Où dy la mesure de Lebesgue. Similaire à la fonction $a(y,\sigma^2)$, cette fonction $b(y,\sigma^{\,2})$ a beaucoup de liberté. Pour une variable aléatoire Y paramètrée par θ et définie sur un ensemble mesurable A, la distribution cumulée est

$$pr_\theta(y \in A) = \int_A \exp\left\{y\theta - K(\theta)\right\} \nu(dy). \tag{2.7}$$

Remarquez la fonction de densité $\exp\left\{y\theta - K(\theta)\right\} b(y,1)$ à l'intérieur de cette expression. Comparez cette densité avec la densité axiomatique :

$$f(y,\mu,1) = a(y,1)\exp\left\{-\frac{1}{2}d(y,\mu)\right\}\ . \tag{2.8}$$

Nous allons bientôt faire valoir que ces deux formulations sont les même. Avec cette expression pour la distribution, il est facile de déterminer la fonction de géniration de moment et la fonction de géniration de cumul pour la variable Y.

$$\begin{aligned} M_Y(t,\theta) &= \int \exp(yt)\exp\left\{y\theta - K(\theta)\right\}\nu(dy) = \exp\{-K(\theta)\}\int \exp\left\{y\theta + yt\right\}\nu(dy) \\ &= \exp\{-K(\theta)\}\exp\{K(\theta+t)\} = \exp\{K(\theta+t) - K(\theta)\}\ . \end{aligned} \tag{2.9}$$

Alors

$$K_Y(t,\theta) = K(\theta+t) - K(\theta)\ . \tag{2.10}$$

En utilisant ces expressions simples, on peut calculer le jième cumulant et le jième moment par rapport à t. Cette valuation justifie le jargon ≪ fonction cumulative ≫ $K(\theta)$.

$$\begin{aligned} K^{(j)}(t,\theta) &= \frac{\partial^{(j)}K(t,\theta)}{\partial t^j} = K^{(j)}(\theta+t) \\ K^{(j)}(0,\theta) &= K^{(j)}(\theta)\ . \end{aligned}$$

Rappelons que le premier cumulant est le μ moyen. Observe ceci

$$K^{'}(\theta) = \mu. \tag{2.11}$$

Définir une fonction $\tau : \Theta \to \mathbb{D}$ où Θ est l'espace des paramètres, $\mathbb{D}$ est l'espace moyen des paramètres, et $\tau(\theta) = K'(\theta)$. En d'autres termes, nous avons une fonction qui associe le paramètre θ au paramètre de position μ. Soit $\tau^{-1} : \mathbb{D} \to \Theta$ la fonction inverse de τ. Cette fonction τ^{-1} associe un paramètre de position μ au paramètre θ. Avec les fonctions

τ et τ^{-1}, Nous montrons bientôt que la définition axiomatique et la définition d'un

modèle de dispersion expriment la même idée. Tout d'abord, nous formalisons la définition constructive d'un modèle de dispersion exponentielle .

Définition 20 *Nous appelons* $\{pr : \theta \in \Theta\}$ *une famille exponentielle à un paramètre si*

- ♣ *Les fonctions de distribution ne correspondent pas à une valeur constante à* 1,
- ♣ Θ *contient plus d'éléments que juste 0 .*

Ces deux conditions indiquent que les fonctions de distribution décrivent un comportement aléatoire et que la famille comprend au moins deux membres. Nous généralisons cette famille de paramètres à une famille de deux paramètres :

Modèles de dispersion exponentielle. Soit $\sum$ un ensemble contenant des éléments $\sigma^2 > 0$. Étant donné une famille exponentielle d'un paramètre,

$$\frac{K(\theta)}{\sigma^2} = \log \int \exp\left\{\theta \frac{y}{\sigma^2}\right\} \nu_{\frac{1}{\sigma^2}}(dy) .$$

Contient une mesure $\sigma-$finie $\nu_{\frac{1}{\sigma^2}}$. Essentiellement, nous avons mis à l'échelle la famille d'origine à un seul paramètre en fonction du second paramètre (le paramètre de dispersion). La fonction de distribution pour un membre de la famille exponentielle à deux paramètres a une forme familière :

$$pr_{(\theta,\sigma^2)}\left(y \in A\right) = \int_A \exp\left\{\frac{y\theta - K(\theta)}{\sigma^2}\right\} \nu_{\frac{1}{\sigma^2}}(dy) .$$

Nous évaluons la fonction de génération de moment et la fonction de géniration de cumul avec cette fonction de distribution .

$$\begin{aligned} M_Y(t,\theta,\sigma^2) &= \int \exp\{yt\} \exp\left\{\frac{y\theta - K(\theta)}{\sigma^2}\right\} \nu(dy) \qquad (2.12) \\ &= \exp\left\{-\frac{K(\theta)}{\sigma^2}\right\} \int \exp\left\{\frac{y}{\sigma^2}(\theta + t\sigma^2)\right\} \nu(dy) \\ &= \exp\left\{-\frac{K(\theta)}{\sigma^2}\right\} \exp\left\{\frac{K\left(\theta + t\sigma^2\right)}{\sigma^2}\right\} = \exp\left\{\frac{K\left(\theta + t\sigma^2\right) - K(\theta)}{\sigma^2}\right\} \end{aligned}$$

$$K_Y(t,\theta,\sigma^2) = \log\left(\exp\left\{\frac{K(\theta + t\sigma^2) - K(\theta)}{\sigma 2}\right\}\right) = \frac{K(\theta + t\sigma^2) - K(\theta)}{\sigma^2}.$$

Trouvez les premier et deuxième cumulants en différenciant par rapport à $t = 0$.

$$\frac{\sigma^2 K^{'}(\theta)}{\sigma^2} = K^{'}(\theta) = \tau(\theta) = \mu \ .$$

$$\frac{\left(\sigma^2\right)^2 K^{''}(\theta)}{\sigma^2} = \sigma^2 K^{''}(\theta) = \sigma^2 \tau^{'}(\theta) \ .$$

Rappelons que le deuxième cumulant $K^{''}(0)$ donne la variance pour la variable aléatoire. De plus, $\sigma^2.V(\mu)$ est la variance d'un EDM. Par conséquent, $\tau^{'}(\theta) = K^{''}(\theta) = V(\mu)$. Nous avons trois façons de décrire la fonction de variance pour un modèle de dispersion exponentielle. Revenons à la fonction de distribution $pr_{(\theta,\sigma^2)}(y \in A)$. Nous avons défini cette fonction avec une mesure $\sigma-$finie $\nu_{\frac{1}{\sigma^2}}(y)$, mais nous n'avons pas expliqué ce qu'est la mesure . Soit $\nu_{\frac{1}{\sigma^2}}(y)$ égal $b(y;\sigma^2)dy$. IL s'ensuit que la densité est

$$f(y,\mu,\sigma^2) = b(y,\sigma^2)\exp\left\{\frac{y\theta - K(\theta)}{\sigma^2}\right\} = b(y,\sigma^2)\exp\left\{\frac{y\tau^{-1}(\mu) - K\left(\tau^{-1}(\mu)\right)}{\sigma^2}\right\} \quad (2.13)$$

Comparer avec la densité axiomatique

$$f(y,\mu,\sigma^2) = a(y;\sigma^2)\exp\left\{-\frac{1}{2\sigma^2}d(y,\mu)\right\}.$$

Dans la théorie de modèles de dispersion, Jorgensen donne que la déviance unitaire $d(y,\mu)$ pour un modèle de dispersion exponentielle comme

$$2\left[\sup_{\theta\in\Theta}\left(y\theta - K(\theta)\right) - y\tau^{-1}(\mu) + K\left(\tau^{-1}(\mu)\right)\right]. \quad (2.14)$$

Utilisez le calcul pour trouver θ qui maximise $y\theta - K(\theta)$.

$$\frac{\partial}{\partial\theta}\left(y\theta - K(\theta)\right) = y - K^{'}(\theta) = y - \tau(\theta) = q\,.$$

Supposons que y soit dans l'espace moyen des paramètres $\mathbb{D}$. Observe que $\tau^{-1}(y)$ maximise $y\theta - K(\theta)$. Ainsi, lorsque $y \in \mathbb{D}$, nous obtenons la déviance unitaire $d(y,\mu)$ comme

$$2\left[y\tau^{-1}(y) - K\left(\tau^{-1}(y)\right) - y\tau^{-1}(\mu) + K\left(\tau^{-1}(\mu)\right)\right]. \quad (2.15)$$

Rappelons que l'espace paramétrique moyen $\mathbb{D}$ est l'intérieur du support convexe C. Le Tableau 1 précédent enregistre la prise en charge convexe des modèles Tweedie. La plupart de temps, le support convexe est un intervalle ouvert comme $\mathbb{R}$. Par conséquent, y est presque toujours dans l'espace des paramètres moyen. Ce n'est que pour $T_{W_P}(\mu,\sigma^2)$ avec $1 < p < 2$ que le support convexe est l'intervalle semi-ouvert $[0,\infty)$. Jorgensen gère ce cas particulier lorsque $y = 0$ dans son livre . Pour garder les choses simples, nous supposons que $y \in \mathbb{D}$. La proposition suivante établit le lien entre la définition axiomatique et la définition constructive .

Proposition 21 *Soit f_a la fonction de densité de la définition axiomatique et f_c la fonction de densité de la définition constructive. Supposons que $a(y,\sigma^2) = f_c(y,y,\sigma^2)$. En-*

suite, pour les modèles Tweedie,

$$f_a(y,\mu,\sigma^2) = f_c(y,\mu,\sigma^2)\ . \tag{2.16}$$

Autrement dit,

$$a(y,\sigma^2)\exp\left\{-\frac{1}{2\sigma^2}d(y,\mu)\right\} = b(y,\sigma^2)\exp\left\{\frac{y\theta - K(\theta)}{\sigma^2}\right\}\ . \tag{2.17}$$

Preuve. Par hypothèse, nous avons un modèle de Tweedie. Par conséquent, $y \in \mathbb{D}$ est la déviance unitaire est

$$2[y\tau^{-1}(y) - K(\tau^{-1}(y)) - y\tau^{-1}(\mu) + K(\tau^{-1}(\mu))].$$

Remplacer dans $a(y,\sigma^2) = f_c(y,y,\sigma^2)$;

$$\begin{aligned} f_a(y,\mu,\sigma^2) &= f_c(y,y,\sigma^2)\exp\left\{-\frac{1}{2\sigma^2}d(y,\mu)\right\} \\ &= b(y,\sigma^2)\exp\left\{\frac{y\tau^{-1}(y) - K(\tau^{-1}(y))}{\sigma^2}\right\}\exp\left\{-\frac{1}{2\sigma^2}d(y,\mu)\right\} \\ &= b(y,\sigma^2)\exp\left\{\frac{y\theta - K(\theta)}{\sigma^2}\right\} = f_c(y,\mu,\sigma^2) \end{aligned}$$

■

Remarque 22 *Nous discutons que les supports convexes des modèles Tweedie. En général, les définitions axiomatiques et constructives des modèles de dispersion exponentielle sont les mêmes. Voir Théorie des modèles de dispersion. Pour résumer, nous avons maintenant une façons différente de penser les densités EDM (et Tweedie) en termes de fonctions cumulantes, de fonctions géniratrices de moments et de fonctions géniratrices de cumulant. Cette version est identique à ce que nous avons discuté précédent. En plus d'étendre notre boîte à outils avec des moments et des cumulant, nous avons révélé comment les EDM sont construits en mettant à l'échelle des familles exponentielles à un paramètre.*

2.3 La famille de Tweedie

Les modèles deTweedie sont des modèles de dispersion exponentielle fermés par des transformations d'échelle et des traductions. Nous désignons un modèle Tweedie comme $T_{W_p}(\mu,\sigma^2)$. Remarquez comment cette notation inclut un indice p .

Tout au long de ce texte, Nous désignerons p comme paramètre de puissance ou comme paramètre d'indexation .

Le paramétre p peut être vu comme un indice qui identifie le type de variable aléaitoire de Tweedie. Son rôle dans les fonctions de variance des modèles de Tweedie lui vaut la désignation comme paramètre de puissance. Les modèles de Tweedie ont des fonctions de variance de la forme $V(\mu) = \mu^p$, d'où le paramétre de puissance de la langue. Ci-dessous un tableau des modèles Tweedie et de leurs pouvoirs.

TABLE 2.1 – Tableau des Modèles Tweedie basés sur le paramètre d'indexation p

Distribution	Domaine	valeur p
$\alpha-$Stable	R	$P<0$
Normal	R	0
Poisson	N	1
composé poisson-gamma	R_{0+}	$1<P<2$
Gamma	R_+	2
α_Stable*	R_+	$2<P<3$
Gaussien inverse	R_+	3
α_Stable *	R_+	$P>3$

* n'est pas stable.

Avant de donner un théorème qui spécifie ce qui différencie les modèles de Tweedie des autres EDM, nous devons clarifier une définition et argumenter deux lemmes .

Définition 23 *Soit X une variable aléatoire. On dit que X est divisible à l'infini si $\forall\, n \in \mathbb{N}$ on peut écrire X Comme la somme de n variables aléaitoires indépendantes, distribuées de manière identique $(i.i.d)$.*

Lemme 24 *Soit X une variable aléaitoire. Alors,*

$$Var(cX) = c^2 Var(X). \tag{2.18}$$

Preuve. De (cf. Ross, (2009))

nous utilisons cela

$$Var(X) = E[X^2] - E[X]^2$$

$$\begin{aligned} Var(cX) &= E[(cX)^2] - (E[cX])^2 = c^2 E[X^2] - c^2 E[X]^2 \\ &= c^2(E[X^2] - E[X]^2) = c^2 Var(X). \end{aligned}$$

■

Lemme 25 *Soit g une fonction qui est au moins une fois défférentiables, prend des entrées réelles positives, et pour la quelle $g(xy) = g(x)g(y)$ est vrai. Alors $g(x) = x^a$ où, a est une constante*

Preuve. Considérons une fonction f telle que

$$f(x) = \log g(e^x).$$

Donc,

$$f(x+y) = \log g(e^{x+y}) = \log g\left(e^x e^y\right) = \log g(e^x)g(e^y) = \log g(e^x) + \log g(e^y) = f(x) + f(y)\,.$$

De plus,

$$f(c.x) = c.f(x)\,.$$

Ainsi, f est linéaire. Si f est linéaire, alors g doit avoir été exponentiel. ■

2.3.1 Invariace au changementd'échelle

Parmi les propriétés importantes de la distribution de Tweedie, nous avons la propriété d'invariance d'échelle, comme le montre le théorème suivant

Théorème 26 *Soit X un $EDM(\mu,\sigma^2)$ tel que $V(1) = 1$et V soit au moins une fois défférentiable. S'il existe une fonction $f : \mathbb{R}_+ \times \sum \to \sum$ pour laquelle*

$$cX = EDM(c\mu, f(c,\sigma^2))\ \forall c > 0 \tag{2.19}$$

détient, Alors :

- ♡ *X est un modèle Tweedie.*
- ♡ *$f(c,\sigma^2) = c^{2-p}\sigma^2$.*
- ♡ *X est infiniment divisible.*

Preuve. Selon le 2.18,

$$Var(cX) = c^2 Var(X) = c^2\sigma^2 V(\mu).$$

Notre hypothèse donne que $Var(cX)$ est également égal à $f(c,\sigma^2)V(c\mu)$. Définissez ces deux expressions pour qu'elles soient égales à un autre .

$$c^2\sigma^2 V(\mu) = f(c,\sigma^2)V(c\mu)\,.$$

Diviser par $f(c,\sigma^2)$. Cette division est légale car le codomaine de f est $\mathbb{R}_+$. Maintenant,

$$\frac{C^2\sigma^2}{f(C^2,1)}V(\mu) = V(c\mu)$$

prenez $\sigma^2 = 1$. Alors

$$\frac{C^2}{f(C^2,1)} = V(c)$$

car $V(1) = 1$ à partir de nos hypothèses initiales. Cette manœvre suggère que $\frac{C^2\sigma^2}{f(C^2,1)}$ est $V(C)$, bien que nous ne sachions pas encore ce qu'est cette fonction $V(.)$. En tout état de cause, nous avons la relation

$$V(C\mu) = V(C)V(\mu)$$

Utiliser le lemme pour affirmer que $V(\mu) = \mu^p$ pour certains $p \in \mathbb{R}_+$. Ensuite, nous comparons

$$\frac{C^2\sigma^2}{f(C^2,\sigma^2)} = V(c) = c^p$$

et conclure que $f(c,\sigma^2) = c^{2-p}\sigma^2$. Pour $p \neq 2$, $f(c,\sigma^2)$ varie en $\mathbb{R}_+$ car σ^2et c varient en $\mathbb{R}_+$. Cette évolutivité nous permet de construire X comme une somme de $idd.\ CX$ variables aléaitoires; c'ast à dire que X est divisible à l'infini pour 2. Lorsque $p \neq 2$, X est une variable aléatoire Gamma. IL bien connu que les variables aléatoires Gamma sont divisibles à l'infini. Ainsi, X est également divisible à l'infini dans le cas $P = 2$. ■

Considérons la transformation $Z = cY$ pour $c > 0$ où Y suit une distribution Tweedie de densité $f_Y(z,\mu,\sigma^2)$ avec une moyenne μ et fonction de variance $V(\mu) = \mu^p$. À partir de résultat $f(c,\sigma^2) = c^{2-p}\sigma^2$,nous pouvons constater que Z suit une loi de distribution de Tweedie avec le même paramètre de distribution p, de moyenne $c\mu$ et dispersion $c^{2-p}\sigma^2$.Plus exactement, nous avons

$$f_Z(z,\mu,\sigma^2) = c f_Y(cz, c\mu, c^{2-p}\sigma^2) \text{ pour tout } z > 0 \text{ et } c > 0\,. \tag{2.20}$$

Les modèles de Tweedie sont fermés par des transformations d'échelle. Pour une constante réelle positive $c > 0$ et un modèles Tweedie $T_{w_p}(\mu,\sigma^2)$,

$$C.T_{W_P}(\mu,\sigma^2) = T_{W_P}(c\mu, c^{2-p}\sigma^2)\,. \tag{2.21}$$

Preuve. Ce corollaire découle immédiatement du théorème. ■

Le corollaire précédent fournit une astuce utile pour la mise à l' échelle des modèles Tweedie. Escaladant $X \sim \mathcal{N}(\mu,\sigma^2)$ par C conduit à $CX \sim \mathcal{N}(C\mu, C^2\sigma^2)$. De même, $\mathcal{G}\dashv(\alpha,\beta)$ donne une variable aléaitoire $\mathcal{G}\dashv(\alpha,\frac{\beta}{C})$ car $\mathcal{T}(\frac{\alpha}{\beta},\frac{1}{\alpha}) = \mathcal{G}\dashv(\alpha,\beta)$. Ces exemples montrent à quel point l est facile de mettre à l'échelle des modèles Tweedie. Une autre propriété utile des modèles Tweedie est que nous pouvons les traduire, c'est-à-dire. Ajouter ou soustraire par une valeur constante.

Théorème 27 *Soit X un $EDM(\mu,\sigma^2)$ fermé sous translation et avec une fonction de variance unitaire différenciable. Cette fermeture signifie qu'il existe une fonction $h(C,\sigma^2)$ telle que*

$$C + X = EDM(C + \mu, h(C,\sigma^2)),\ \forall c \in \mathbb{R}.$$

Un tel EDM est divisible à l'infini et possède une fonction de variance unitaire exponentielle.

Preuve. Évaluer les écarts des deux côtés

$$\sigma^2 V(\mu) = h(C,\sigma^2) V(C+\mu) .$$

Ainsi $V(C\mu) = g(c)V(\mu)$ où $g(c) = \sigma^2 / h(c,\sigma^2)$ Parce que $V(.)$ est différentiable et positif, g est différentiable et positif. C varie en R . Afin que nous puissions nous différencier à son sujet. Différencier par rapport à c à 0. On obtient $V'(\mu) = g'(0)V(\mu)$. Ensuite, nous résolvons l'équation différentielle.

$$\int \frac{1}{V(\mu)} V'(\mu) = \int g'(0)$$

$$\log V(\mu) = \mu g'(0) + C_0$$

$$V(\mu) = C_0 \exp\{\mu g'(0)\}$$

Ici $C_0 > 0$ représente une constante. Cette constante est une conséquence de l'intégration indéfinie. Clairement, $V(\mu)$ est une fonction exponentielle. Enfin, utilisez la substitution pour résoudre $h(C,\sigma^2)$ dans l'équation.

$$h(C,\sigma^2) V(C+\mu) = \sigma^2 V(\mu).$$

$$h(c,\sigma^2) = \sigma^2 \exp\{-g'(0)c\}.$$

Lorsque $g(0) = 0$, la fonction de variance unitaire correspond à celle d'une variable aléaitoire normale IL est bien connu qu'une somme de iid. les variables aléaitoires normales sont une variable aléaitoire normale. Lorsque $g'(0) \neq 0$, $h(c,\sigma^2)$varie avec $C \in \mathbb{R}$. Nous pouvons donc construire une somme de iid. Variables aléaitoires. Ainsi X est infiniment divisible. ■

Remarque 28 *Techniquement parlent, la fonction de variance unitaire dans le théorème précédent n'a pas la forme $V(\mu) = \mu^p$ pour certains $p \in \mathbb{R}$. Décrire les modèles Tweedie comme des EDM qui ont une variance unitaire $V(\mu) = \mu^p$ facilite la mémorisation, mais nous doit assouplir cette dénonciation pour inclure la fermeture sous traduction. Les modèles de Tweedie sont des EDM qui sont divisibles à l'infini et fermés par translation et transformations d'échelle.*

Jusqu'à présent, nous avons déclaré que les modèles Tweedie ont une variance unitaire $V(\mu) = \mu^p$ pour certains $p \in \mathbb{R}$, et nous avons fourni quelques exemples de valeurs possibles pour le paramètre d'indexation. Mais, la droite réelle est infiniment grande. IL doit sûrement y avoir des valeurs pour le paramètre de puissance qui ne fonctionnent pas .

Proposition 29 *IL n'ya pas de modèles Tweedie avec le paramètre d'index* $0 < p < 1$.

La preuve de cette proposition implique des fonctions génératrices de moments et des fonctions générant des cumulant. Parce que le corps principal de ce texte cible un public de premier cycle, nous réservons la preuve pour l'annexe. L'argument synthétise de nombreuses pages du livre de (cf. jorgensen, (1997)) The Theory of Dispersion Modeles[.

Étant donné que les modèles Tweedie sont des EDM. IL ont des déviations d'unité. La propositionnelle .1 donne les déviations unitaires pour les cas normaux, Poisson et Gamma. Mais qu'en est-il lorsque $p \notin \{0,1,2\}$? pour justifier la déviance unitaire dans le cas géniral. Nous avons encore besoin d'installations avancées avec EDM.

Proposition 30 *Pour un modèle de Tweedie avec une puissance d'index* $p \notin \{0,1,2\}$, *la déviance unitaire* $d(y,\mu)$ *est*

$$2\left\{\frac{\max(y,0)^{2-p}}{(1-p)(2-p)} - \frac{y\mu^{1-p}}{1-p} + \frac{\mu^{2-p}}{2-p}\right\} \tag{2.22}$$

2.3.2 Preuves de La Famille Tweedie

Nous avons fait valoir certaines propositions de modèles Tweedie précédent sans justification. De plus, nous avons suggéré dans une note de bas de page que certains paramètres d'indice p correspondent à des distributions stables. Je suis heureux de donner maintenant des arguments pour ces résultats. Rappelons que la fonction de variance pour modèle Tweedie est $V(\mu) = \mu^p$. Nous déterminons une expression pour le paramètre θ et la fonction cumulative $K(\theta)$ en terme de μ et p. Observer ceci

$$K''(\theta) = \frac{\partial \tau(\theta)}{\partial \theta} = \frac{\partial \mu}{\partial \theta} = \mu^p.$$

Ignorer les constantes arbitraires, nous obtenons dans une intégration définitive,

$$\theta = \frac{\mu^{1-p}}{1-p}, \quad p \neq 1 : \log \mu, \quad p = 1. \tag{2.23}$$

Jorgensen écrit μ intervalles de θ et pas bien. Il introduit un paramètres α qui lié à p en ce que

$$\alpha = \frac{p-2}{p-1}.$$

La relation inverse dit que

$$p = \frac{\alpha - 2}{\alpha - 1}$$

Considère que

$$p-1=\frac{\alpha-2}{\alpha-1}-1=\frac{\alpha-2}{\alpha-1}-\frac{a-1}{\alpha-1}=-\frac{1}{\alpha-1}.$$

Maintenant, calculez μ en termes de θ et α en trouvant l'inverse de θ en terme de μ et p .

$$\mu=\begin{cases}\frac{\theta}{\alpha-1})^{\alpha-1}, & p\neq 1\\ \theta, & p=1.\end{cases} \tag{2.24}$$

Ensuite, nous trouvons la fonction cumulative $K(\theta)$ en résolvant l'équation différentielle $K'(\theta)=\tau(\theta)=\mu$. Pour $p\neq 1,2$, intégrer pour obtenir

$$K(\theta)=\frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^{\alpha},$$

pour $p=1$, nous intégrons pour $K(\theta)=e^{\theta}$. Le cas délicat est pour $p=2$. Lorsque $p=2$, $\alpha=0$. Ainsi, nous obtenons

$$K'(\theta)=-\frac{1}{\theta}\,.$$

L'anti dérivé est $-\log(-\theta)$. En résumé

$$K(\theta)=\begin{array}{cc}\frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^{\alpha} & p\neq 1,2\\ -\log(-\theta) & p=2\\ e^{\theta} & p=1\end{array}\quad. \tag{2.25}$$

Je comprends qu'il est difficile de se souvenir de ces nombreuses relations entre p, α, μ, θ, $K(\theta)$, et $V(\mu)$. Les mathématique essaient parfois parce qu'elles défient vous facultés mentales, et parfois il essaie parce qu'il vous oblige réguriter and synthéstiser une masse d'information. La tâche à accomplir s'adresse à ce dernier. Pour les preuves procédurales, nous utiliserons ces expressions. Reportez-vous à cette page et à la page précédente lorsque vous devez rappeler les relations entre les paramètres et les fonctions Tweedie.

Proposition 31 *IL n'y a pas de modèles Tweedie avec le paramètre d'index* $0<p<1$.

Preuve. *Supposons qu'il existe un* $T_{W_p}(\mu,\sigma^2)$ *où* $p\in(0,1)$. *Nous savons que*

$$K(\theta)=\frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^{\alpha}$$

Défférencier deux fois par rapport à θ *trouver que*

$$K''(\theta)=\left(\frac{\theta}{\alpha-1}\right)^{\alpha-2}.$$

■

Avant de commencer la prochaine étape, nous devons considérer ce que c'est α. La relation entre α et p est mieux représentée graphiquement. Observe ceci $\lim_{p\to 0^+} \frac{p-2}{p-1}$ et que la cartographie à partir de la concentration d'isomonotone augmente. Par conséquent, $\alpha - 2 > 0$. Rappelons que $0 \in \Theta$ pour tous les modèles de dispersion exponentielle. Pour $\theta = 0$, La variance de la variable aléatoire Tweedie est de 0. Autrement dit, la variable aéatoire Tweedie n'est pas du tout stochastique. Ce résultat conterdit le fait que les modèles Tweedie sont des objets aléatoires. Nous concluons qu'il n'existe aucun modèle de Tweedie indexé par $p \in (0,1)$. Nous pouvons également dériver la déviance générale des ubités pour les modèles Tweedie. Précédemment, nous avons vu les déviations unitaires pour les cas normaux, de Poisson et de Gamma. Ces modèles Tweedie sont des cas spéciaux .

Proposition 32 *Pour un modèle Tweedie avec puissance d'indexation $p \notin \{0,1,2\}$. La déviance unitaire $d(y,\mu)$ est*

$$2\left\{\frac{\max(y,0)^{2-p}}{(1-p)(2-p)} - \frac{y\mu^{1-p}}{1-p} + \frac{\mu^{2-p}}{2-p}\right\}. \tag{2.26}$$

Preuve. Rappelons que les modèles de Tweedie ont une dèviance unitaire 2.14

$$2\left\{\sup_{\theta\in\Theta}\left[y\theta - K(\theta)\right] - \left(y\tau^{-1}(\mu)\right) + K\left(\tau^{-1}(\mu)\right)\right\}.$$

Nous savons que $\tau^{-1}(\mu) = \theta$ et $K\left(\tau^{-1}(\mu)\right) = K(\theta)$. Vérifier que

$$y\tau^{-1}(\mu) = y\theta = \frac{y\mu^{1-p}}{1-p}.$$

Cette égalité découle directement de la façons dont nous avons décrit θ en termes de μ et p. L'écture de $K(\theta)$ en termes de μ et p nécessite plus de manipulation algébrique que simple substitution. Vérifiez l'algèbre ci-dessous. Nous commençons par simplifier ce qui est à l'interieur des parenthèse, puis nous simplifions le multiplicateur à l'extérieur des parenthèse, puis en simplifions le multiplicateur en dehors des parenthèse. Observez que nous appliquons la formule $(p-1)(\alpha-1) = -1$ plusieurs fois.

$$\begin{aligned}
k(\theta) &= \frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^{\alpha} = \frac{\alpha-1}{\alpha}\left(\frac{\mu^{1-p}}{(1-p)(\alpha-1)}\right)^{\alpha} \\
&= \frac{\alpha-1}{\alpha}\left(\mu^{1-p}\right)^{\alpha} = \frac{\alpha-1}{\alpha}(\mu^{1-p})^{\frac{p-2}{p-1}} \\
&= \frac{\alpha-1}{\alpha}\mu^{\frac{(p-2)(1-p)}{p-1}} = \frac{\alpha-1}{\alpha}\mu^{2-p} = \frac{(\alpha-1)(p-1)}{p-2} = -\frac{\mu^{2-p}}{p-2} = \frac{\mu^{2-p}}{2-p}.
\end{aligned}$$

Ensuite, nous nous préoccupons de $\sup_{\theta\in\Theta}\left[y\theta - K(\theta)\right]$. Du calcul, nous voyons que le maximum se produit lorsque

$$y = \left(\frac{\theta}{\alpha-1}\right)^{\alpha-1} = \left(\frac{\mu}{(\alpha-1)(1-p)}\right)^{(\alpha-1)(1-p)} = \mu^{(\alpha-1)(1-p)} = \mu$$

Ce calcul signifie que nous n'obtenons une solution que si $y \geqslant 0$. Substitut dans $\max(y,0)$ pour y. Calculer

$$\begin{aligned}\sup_{\theta\in\Theta}\left[y\theta - K(\theta)\right] &= \max(y,0).\frac{(y,0)^{1-p}}{1-p} - \frac{\max(y,0)^{2-p}}{2-p} \\ &= \frac{\max(y,0)^{2-p}}{1-p} - \frac{\max(y,0)^{2-p}}{2-p} = \frac{\max(y,0)^{2-p}}{(1-p)(2-p)}\end{aligned}$$

Rassemblez toutes les pièces. Pour $T_{W_P}\left(\mu,\sigma^2\right)$ avec $p \in \{0,1,2\}$, Nous concluons que la déviance unitaire est

$$2\left\{\frac{\max(y,0)^{2-p}}{(1-p)(2-p)} - \frac{y\mu^{1-p}}{1-p} + \frac{\mu^{2-p}}{2-p}\right\}.$$

La dernière chose intèressante que nous allons dire à propos des modèles de Tweedie renvoie à un commentaire que nous avons fait précédemment. Nous avons remarqué que certains modèles de Tweedie sont des distributions stables pour certains choix de paramètres θ. Les distribution stables se connectent aux distributions limites et trouvent leur application dans de nombreux paramètres financiers. Par exemple,(cf. James Weatherall , (2013)) parle à un niveau élevédes distributions stables dans son livre de vulgarisation scientifique The Physics of Wall Street. Pour l'instant, nous introduisons des distributions stables dans le contexte des modèles de Tweedie.

Définition 33 *Soit $X_1,....,X_n$ des variables aléatoires indépendante, de distribution identique avec la distribution F. On dit que X est une variable aléatoire stable, si, pour tout $n \in \mathbb{N}$, il existe une une constante b et des non constantes c_i telles que :*

$$aX + b = \sum_{i=1}^{n} c_i X_i.$$

A également la distribution F. Si $b = 0$, On dit que X est une variable aléatoire strictement stable.

■

Théorème 34 *Soit X un modèle de Tweedie avec $p \in (-\infty,0]\cup(2,\infty)$ et avec $\tau^{-1}\left(\mu\right) = 0$. X est une variable aléatoire strictement stable.*

Preuve. Observe ceci $\tau^{-1}(\mu) = \theta$, donc $\theta = 0$. La fonction génératrice de cumulant du modèles de Tweedie est

$$K_X(t,0,\sigma^2) = \frac{K(t\sigma^2) - K(0)}{\sigma^2}.$$

Rappeler que $K(\theta) = \frac{\alpha-1}{\alpha}\left(\frac{\theta}{\alpha-1}\right)^{\alpha}$ pour le p donné pour $p \leq 0$, $\alpha \in (1,2]$; pour $p > 2$, $\alpha \in (0,1)$. Nous voyons ces ma pages dans les figures 2.2 et 2.3. Par conséquent, $|\alpha - 1| > 0$ et $K(0) = 0$.

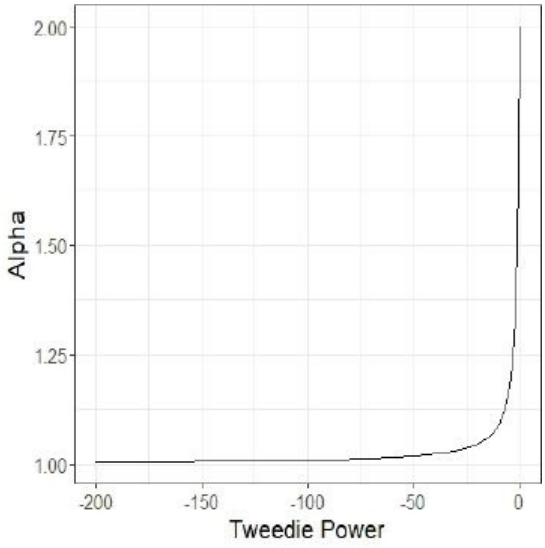

FIGURE 2.2 – Relation entre α et β pour $p \leq 0$

■

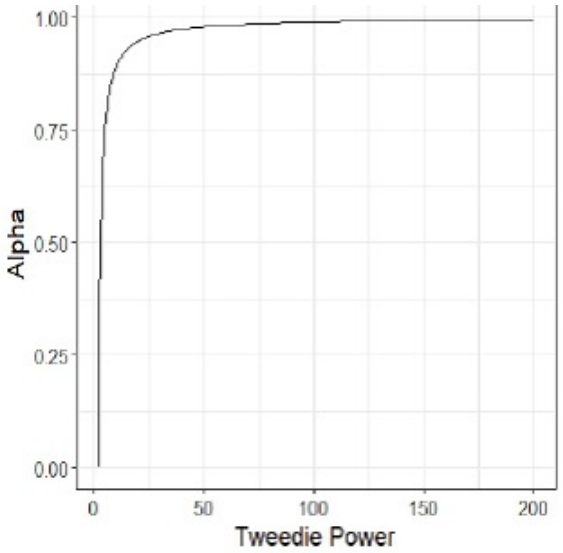

FIGURE 2.3 – Relation entre α et β pour $p > 2$

Considérez $X_1, ..., X_n$ copies de X identiques et distribuées de manière identique. Évaluez la fonction génératrice de cumulant de $\sum_{i=1}^{n} X_i$. C'est à dire,

$$\sum_{i=1}^{n} K_{X_i}\left(t,0,\sigma^2\right) = n.K_X\left(t,0,\sigma^2\right) = \frac{n(\alpha-1)}{\alpha\sigma^2}\left(\frac{t\sigma^2}{\alpha-1}\right)^{\alpha} = \frac{\alpha-1}{\alpha\sigma^2}\left(\frac{n^{\frac{1}{\alpha}}t\sigma^2}{\alpha-1}\right)^{\alpha} = K_X\left(n^{\frac{1}{\alpha}}t,0,\sigma^2\right).$$

$X_1,...,X_n$ a la même distribution que $n^{\frac{1}{\alpha}}X$. Par définition, X est une distribution strictement stable.

Outre le fait que certaines variables aléatoires de Tweedie sont stables dans des cas particuliers, (cf. Jorgensen,(1997)) justifie le jargon "$stable$" en déclarant que les variables aléatoire Tweedie ont des propriétés similaires aux distributions stables. Par exemple, les variables aléatoire de Tweedie et les variables aléatoire stables sont divisible à l'infini.

De plus, les distributions de Tweedie apparaissent comme des distributions limites dans une sorte de théorème de limite centrale généralisée . Voir Théorie des modèles de dispersion pour étudier cette idée tangentielle.

2.3.3 Famille de Tweedie et les Modèles Linéaires

Les actuaires utilisent les statistiques, les mathématiques finanncières et l'intelligence d'affaires pour évaluer les polices d'assurances et réserver de l'argent pour le paiement des sinistres. Ces professionnelles de l'assurance construisent souvent des modèles statistiques pour résoudre les problèmes de régression. L'analyse de régression examine la relation entre une variable cible dépendante et des variables explicatives indépendantes.

La variable dépendante d'un modèle d'assurance est généralement soit la fréquence des sinistres. Gravité de réclamation, ou les coût de perte. La fréquence des réclamations mesure le nombre de réclamations déposées par un titulaire de police. Les actuaires construisent des modèles de poisson pour prédire la fréquence des sinistres. La gravité d'une réclamation mesure le coût monétaire d'une réclamation. Un ensemble de données sur la gravité des revendications ne contient que des observations de revendications des déclarants. Les actuaires créent des modèles gamma ou gaussiens inverses pour prédire la gravité des sinistres. Les coûts des pertes mesurent le montant qu'un assureur paie pour indemniser un pareur d'assurance. La plupart du temps, les assurés ne déposent pas de réclamation. Les actuaires utilisent des modèles de poisson-gamma pour prédire les coûts de perte .

Maurice Charles Kenneth(cf. MCk Tweedie, (1947)) a proposé un cadre qui englobe toutes ces variables aléaitoires dans une classe. Nous appelons cette classe de variables aléaitoires de la famille Tweedie et les distributions qu'elle contient. La famille Tweedie est une famille robuste de distributions de probabilité. IL comprend une variable aléaitoire discrète (poisson), une variable aléaitoire mixte (poisson-gamma), des variables

aléaitoires continues, et des variables stables. Les modèles de Tweedie sont souvent mis en œvre dans l'analyse de régression, mais ils sont rarement compris au-delà d'un niveau superficiel. En fournissant plus de contexte, j'espère inspirer les modélistes actuariels à être plus précis et créatifs dans leur application des modèles Tweedie. Outre la science actuarielle, d'autres disciplines utilisent des modèles Tweedie. Des modèles de Tweedie ont été utilisés pour décrire les précipitations mensuelles en Australie et pour effectuer des analyses de capture par unité d'effort pour la recherche halieutique.(cf. MCK tweedie, (1984)) se cite lui-même et d'autres statisticiens qui mettent en œvre des modèles tweedie dans les sciences biologiques et médicales. L'étude de cette famille plus en détail sera fructueuse pour de nombreux chercheurs, pas seulement pour les actuaires .

2.4 Méthode d'approximation de la densité de Tweedie

La distribution de Tweedie n'a pas une densité avec une forme fermée, ce qui signifie que nous ne pouvons pas l'exprimer facilement. Or, cette distribution est très utile en actuariat, d'où la nécessité de trouver des méthodes pour approximer sa densité.

Nous étudierons les méthodes d'approximation de la densité de la distribution de Tweedie. L'approximation de la densité est basée sur les méthodes suivantes : la méthode d'inversion de Fourier, la méthode de développement en séries infinies et la méthode de point-selle.

Afin de simplifier les opérations mathématiques d'approximation de la densité de Tweedie par des méthodes qui seront discutées (cf. Dunn et Smyth, (2008)) ont choisi $k(\theta) = 0$ et $\mu = 1$ à $\theta = 0$, sans perte de généralité, ce qui donne

$$\theta = \begin{cases} \frac{\mu^{1-p}-1}{1-p}, & \text{if} p \neq 1 \\ \log\mu, & p = 1. \end{cases} \tag{2.27}$$

$$\mu = \left[\theta\left(1-p\right)+1\right]^{\frac{1}{1-p}} \text{ ,pour } p \neq 1 \tag{2.28}$$

et

$$K(\theta) = \begin{cases} \frac{\mu^{2-p}-1}{2-p}, & p \neq 2 \\ \log\mu, p = 2 \end{cases} \tag{2.29}$$

Dans la formule [2.23] θ n'est pas continue en $p = 1$ et $K(\theta)$, dans l'équation [2.25] nous avons $\theta \to \log\mu$ lorsque $p \to 1 \text{et} K(\theta) \to \log\mu$ lorsque $p \to 2$. Ces nouvelles formes de θ et $K(\theta)$ sont donc

continues en p aussi bien qu'en θ. Dans ce cas, la densité s'écrit sous la forme suivante

$$f\left(z,\mu,\sigma^2\right)=a'(z,\sigma^2)\exp\left\{\frac{1}{\sigma^2}\left(z\frac{\mu^{1-p}-1}{1-p}-\frac{\mu^{2-p}-1}{2-p}\right)\right\}. \tag{2.30}$$

Notons que la fonction $a'(z,\sigma^2)$,est défirente de la fonction $a(z,\sigma^2)$del'equation 2.2,puisque

$$\begin{aligned} f\left(z,\mu,\sigma^2\right) &= a'\left(z,\sigma^2\right)\exp\left\{\frac{1}{\sigma^2}\left(z\frac{\mu^{1-p}-1}{1-p}-\frac{\mu^{2-p}-1}{2-p}\right)\right\} \\ &= a'\left(z,\sigma^2\right)\exp\left\{\frac{1}{\sigma^2}\left(z\frac{\mu^{1-p}}{1-p}-z\frac{1}{1-p}-\frac{\mu^{2-p}}{2-p}+\frac{1}{2-p}\right)\right\} \\ &= a'\left(z,\sigma^2\right)\exp\left\{\frac{1}{\sigma^2}\left(-z\frac{1}{1-p}+\frac{1}{2-p}\right)\right\}\exp\left\{\frac{1}{\sigma^2}\left(\frac{\mu^{1-p}}{1-p}\right)-\frac{\mu^{2-p}}{2-p}\right\} \end{aligned} \tag{2.31}$$

où $a'(z,\sigma^2)\exp\left\{\frac{1}{\sigma^2}\left(-z\frac{1}{1-p}+\frac{1}{2-p}\right)\right\}$représente la fonction $b\left(z,\sigma^2\right)$ de l'équation[2.13]

2.4.1 Approximation de la densité Tweedie par la méthode d'inversion de Fourier

L'approximation de la densité de Tweedie par la méthode d'inversion de Fourier se base sur l'inversion de la fonction génératrice des cumulant

$$K_Y(t)=\frac{1}{\sigma^2}\left[k(\theta+t\theta)-k(\theta)\right].$$

Pour approximer la densité de Tweedie par cette méthode, nous avons besoin de la formule

$$f(z,\mu,\sigma^2)=a'(y,\sigma^2)\exp\left\{\frac{1}{\sigma^2}\left(z\frac{\mu^{1-p}-1}{1-p}-\frac{\mu^{2-p}-1}{2-p}\right)\right\}.$$

et la formule

$$f\left(z,\mu,\sigma^2\right)=a(z,\sigma^2)\exp\left\{\frac{-d(z,\mu)}{2\sigma^2}\right\}.$$

de la densité de Tweedie. L'expression de $K(\theta)$

$$K(\theta)=\begin{cases}\frac{\mu^{2-p}-1}{2-p}, & p\neq 2\\ \log\mu, & p=2.\end{cases}$$

On montre que la fonction génératrice des cumulant a une forme analytique simple cependant les fonctions $a(.)$ et $b(.)$ n'ont pas de forme fermée,Pour évaluer ces fonctions, nous utilisons alors l'approche d'inversion de Fourier.

2.4.1.1 Inversion de Fourier de la fonction caractéristique

La fonction de densité d'une distribution continue appartenant aux EDM peut être représentée sous la forme d'une intégrale en utilisant la fonction caractéristique de cette distribution, tel que le montre le théorème suivant

Théorème 35 *(cf. Peters et Shevchenko, (2015)) Considérons la variable aléatoire Y avec fonction de répartition $F_Y(y)$, fonction de densité $f_Y(y)$, et fonction caractéristique $\varphi_Y(y)$, telle que : $f_Y(y)$ et $\varphi_Y(y)$ intégrables. Nous pouvons obtenir, par l'inversion de Fourier de la fonction caractéristique σ^2*

$$f_Y(y) = \frac{1}{2\pi}\int \varphi_Y(y)\exp(-ity)\,dt. \tag{2.32}$$

Afin de simplifier les calculs dans l'approximation de la densité Tweedie par l'inversion de Fourier, (cf. Dunn et Smyth, (2008)) ont effectué un changement d'échelle dans la fonction. $f_z(y,\mu,\sigma^2)$ Par la propriété d 'invariance au changement d'échelle de la distribution Tweedie, nous pouvons écrire la densité de Y sous la forme suivante

$$f_Z(y,\mu,\sigma^2) = c\,f_Y(cz, c\mu, c^{2-p}\sigma^2),\ \textit{pour tout}\ y>0\ \textit{et}\ c>0$$

Cette transformation nous permettra de simplifier l'expression de la densité, $f_Z(z,\mu,c)$,A cet effet,(cf. Dunn et Smyth ,(2008)) ont proposé trois façons pour évaluer la densité Tweedie par l'inversion de Fourier.

2.4.1.2 Méthode 1

Dans cette première méthode, (cf. Dunn et Smyth, (2008)) ont posé $\mu = 1$ directement dans la formule de la fonction de densité

$$f_Y(z,\mu,\sigma^2) = f_z(z,1,\sigma^2) = a'(y,\sigma^2)\exp\left[z\frac{1^{1-p}}{1-p} - \frac{1^{2-p}}{2-p}\right] = a'(z,\sigma^2)$$

Dans ce cas, nous n'avons pas besoin d'utiliser le changement d'échelle et nous pouvons appliquer l'inversion de Fourier sur la fonction$a'(z;\sigma^2)$de l'équation [2.30]au lieu de$f_Z(z,\mu,\sigma^2)$ afin de simplifier les calculs.

2.4.1.3 Méthode 2

Cette méthode utilise également la formule de densité de Tweedie de l'équation [2.30]

$$f(z;\mu,\sigma^2) = a'(z,\sigma^2)\exp\left\{\frac{1}{\sigma^2}\left(z\frac{\mu^{1-p}-1}{1-p} - \frac{\mu^{2-p}-1}{2-p}\right)\right\}.$$

Nous utilisons encore $\mu = 1$. (cf. Dunn et Smyth, (2008)) ont choisi $c = \frac{1}{\mu}$ dans la densité transformée de l'équation[2.20[]

$$f_z(z,\mu,\sigma^2) = c f_Y(cz, c\mu, c^{2-p}\sigma^2)\ , z = cY, c > 0\ .$$

En remplaçant c par $c = \frac{1}{\mu}$ dans $c f_Y(cz, c\mu, c^{2-p}\sigma^2)$, nous trouvons

$$f_Z(z,\mu,\sigma^2) = \frac{1}{\mu} f_Y(z/\mu, 1, \sigma^2/\mu^{2-p}).$$

Nous aurons alors

$$f_Y\left(z/\mu, 1, \sigma^2\right) = a'\left(z/\mu, \sigma^2/\mu^{2-p}\right).$$

Ce qui implique

$$f_Z\left(z, 1, \sigma^2/\mu^{2-p}\right) = \frac{1}{\mu} a'\left(z/\mu, \sigma^2/\mu^{2-p}\right).$$

Afin de simplifier les calculs, nous appliquons l'inversion de Fourier, dans ce cas, sur la fonction $a'(z/\mu, \sigma^2/\mu^{2-p})$ à la place de $f_Z(z,\mu,\sigma^2)$.

2.4.1.4 Méthode 3

Cette méthode utilise la formule de la densité de Tweedie de l'équation

$$f_Z(z,\mu,\sigma^2) = a(z,\mu)\exp\left(\frac{-d(z,\mu)}{2\sigma^2}\right)$$

Posons $z = \mu$, donc ,

$$f_Z(z,z,\sigma^2) = a(z,\sigma^2) = c f_Y(z, cz, c^{2-p}\sigma^2)$$

Dunn et Smyth(2008) ont choisi $c = \frac{1}{z}$, nous aurons alors $z = \mu = 1$, ce qui signifie que ,

$$f_z(1, 1, \frac{\sigma^2}{z^{2-p}}) = a(z, \frac{\sigma^2}{\mu^{2-p}}) = a(1, \xi)$$

avec $\xi = \frac{\sigma^2}{\mu^{2-p}}$ et $a(z,\sigma^2) = \frac{1}{z} a(1,\xi)$

Dans ce cas, nous appliquons l'inversion de Fourier sur la fonction $b(1,\xi)$, ce qui signifie que la performance de la méthode 3 dépend de c seulement.

2.4.2 Approximation de la densité de Tweedie par la méthode de développement en séries infinies

Une de ces méthodes consiste à approximer une expression en série infinie, (cf. Joregensen, (1997)). Nous pouvons écrire $a(y;\sigma^2)$ fonctioner en série (cf. Petre et smyth, (2005)). Pour un modèle Poisson-Gamma

$$a(y,\sigma^2) = \frac{1}{y} W(y,\sigma^2,p),$$

où $W(y,\sigma^2,p) = \sum_{j=1}^{\infty} W_j$. Pour $p > 2$ et $y > 0$

$$a(y,\sigma^2) = \frac{1}{\pi y} V(y,\sigma^2,p)$$

avec $V(y,\sigma^2,p) = \sum_{k=1}^{\infty} V_k$.

Les deux Wj et V_k sont des fractions compliquées avec de nombreux paramètres, (voir l'article de (cf. Dunn et Smyth, (1996)) pour l'expression exacte)Néanmoins, nous voulons évaluer W et V. (cf. Dunn et Smyth,(1996)) déterminent les indices $j$et k pour lesquels W_j et V_k atteignent des maximums. Ils accomplissent cette tâche en traitant j et k comme continus, en se différenciant par rapport à eux et en mettant les dérivées à zéro. En d'autres termes, ils trouvent le maximum de la manèire habituelle enseignée dans le calcul universitaire. (cf. Dunn et Smyth, (1996)) trouvent des limites supérieures et inferieures pour j et k autour de $j$max et $k$max. Ces limites ont trouvées par calcul en trouvant W_j et V_k suffisamment petits par rapport à Wmax et Vmax. En fin de compte leur algorithme additionne un nombre fini de W_j et V_k pour servir d'approximation pour W et V.

Cette approche par séries vise à additionner uniquement les termes de la série qui contribuent de manière significative. Nous trouvons le terme qui contribue le plus, nous l'incluons ainsi que ses voisins dans le calcul. Comme l'approche d'inversion de Fourier, ce algorithme pourrait être intensif en calculs ou inexact en fonction de paramètres y, σ^2et p.

2.4.3 Approximation de la densité de Tweedie par la méthode de point-selle

La troisième façon dont (cf. Dunn et Smyth, 1996)) approchent $a(y,\sigma^2)$ est par l'approximation de point-selle. Cette technique est bien connue dans la communauté statistique.Vous pouvez trouver un traitement plus approfondi de l'approximation du point -selle dans littérature mathématique. Dans tous les cas, l'approximation du point-selle donne que $a(y,\sigma^2) \approx \left(2\pi\sigma^2 y p\right)^{1/2}$, où p est la puissance de Tweedie (cf. Peter et Smyth, (2001)).

2.4.3.1 Algorithme de la Méthode du point-selle.

1. Si $y = 0$. Alors, $f_Y(y) = exp(-\lambda)$, avec $\lambda = \frac{\mu^{2-p}}{\sigma^2(2-p)}$,
2. Si $z > 0$. Alors

$$f_Z(z) = (2\pi\sigma^2 y^p)^{-1/2} \exp\left(-\frac{d(y,\mu)}{2\sigma^2}\right),$$

avec,

$$d(y,\mu) = \frac{\left[\max(y,0)\right]^{(2-p)}}{(1-p)(2-p)} - y\frac{\mu^{1-p}}{1-p}.$$

CHAPITRE 3

APPLICATION EN ASSURANCE

Sommaire

3.1 INTRODUCTION

L'utilisation des distributions de Tweedie dans la modélisation des réclamations d'assurance est particulièrement intéressante. Les méthodes vues au chapitre 2 permettent une estimation efficace des paramètres p et σ^2, en utilisant la fonction de vraisemblance, ainsi qu'une vérification de l'adéquation du modèle. Toutefois, comme la distribution Tweedie fait partie des distributions de la famille exponentielle, nous savons que $\hat{\mu} = \bar{X}$ par l'estimateur du maximum de vraisemblance (MLE). Dans ce chapitre, nous effectuons une application en assurance à l'aide d'une base de données qui représente les coûts totaux individuels d'assurance automobile. Nous utilisons également la densité de Tweedie estimée pour calculer la transformée d'Esscher.

(cf. Dunn, (2004)), (cf. Hasan et Dunn, (2011)),

(cf. Hasan et Dunn, (2010)), (cf. Hasan et Dunn, (2015)) ont utilisé des distributions de la famille de Tweedie pour modéliser l'apparition et la quantité de précipitations. Ils ont trouvé que le modèle de Tweedie ajuste bien les précipitations de pluie.

3.2 Distribution de Poisson composée

Soit Z une variable aléatoire suivant la loi de Poisson composée telle que

$$Z = \sum_{k=1}^{N} Y_k. \tag{3.1}$$

où $(Y_k)_k$ est une suite de variables aléatoires indépendantes et identiquement distribuées. Où N est de loi de Poisson de paramètre λ. Notons que N est indépendante des Y_k. Par convention, $Z = 0$ si $N = 0$.

Les distributions de Poisson composée sont très utiles en actuariat. Généralement, N représente le nombre de sinistres ou de réclamations et les variables aléatoires positives Y_k, $k = 1, .., N$, correspondent aux montants des N sinistres. Pour construire le modèle, nous supposons que le nombre de sinistres n'a pas d'influence sur les montants des sinistres. De plus, les montants de chaque sinistre ont le même comportement aléatoire. En fait, le montant du premier sinistre n'a pas d'incidence sur le montant du deuxième sinistre et ainsi de suite.

3.2.1 Fonction de répartition et densité de distribution de Poisson composée

La fonction de répartition de la distribution de Poisson composée s'écrit en conditionnant sur N comme suit

$$\begin{aligned} F_Z(z) &= p(Z<z) = \sum_{k\geqslant 0}^{\infty} p(N=k)P(Z<z/N=k) \\ &= \sum_{k\geqslant 0}^{\infty} P(N=k)F_Y^{(k)}(z) \text{ (car } N\text{et } Y \text{ sont indépendantes), pour } k=0,1,2, \end{aligned}$$

où $F_Y^{(k)}(z)$ représente la k iême convolution de la fonction de répartition $F_Y(z)$ de$Y$, en effet

$$F_Y^{(k)}(z) = P(y_1+y_2+...+y_k \leqslant z) = F_{Y_1+Y_2+....+Y_k}(z), \text{ pour } k \geqslant 1.$$

Donc la fonction de répartition de la somme aléatoire Z est donnée par

$$F_Z(z) = p(Z<z) = \sum \frac{\lambda^k}{k!} e^{\lambda} F_Y^{(K)}(z). \tag{3.2}$$

De même, sa fonction de densité est donnée par

$$f_Z(z) = \sum_{k\geqslant 0} p(Z=k)\, f_Y^{(k)}(z) = \sum \frac{\lambda^k}{k!} e^{\lambda} f_Y^{(k)}(z). \tag{3.3}$$

où $f_Y^{(k)}(z)$ représente la k iême convolution de la fonction de densité. $f_Y(z)$ de Y, pour$k=0,1,2,$

3.2.2 Fonction génératrice de probabilité d'une distribution de Poisson composée

La fonction génératrice de Z est obtenue en conditionnant sur la variable N de la manière suivante

$$G_Z(t) = E\left(t^Z\right) = E\left[E\left(t^{\sum_{i=1}^{i=k} Y_i} \backslash N=k\right)\right], \forall\, i=1,2,...,k,$$

pour t réel tel que $E[t^{Y_i}]$ existe. Puisque les variables Y. Sont indépendantes, il s'avère que

$$G_Z(t) = E\left[\prod_{i=1}^{i=k} E\left(t^{Y_i} \backslash N=k\right)\right]$$

Comme les Y_i sont identiquement distribuées et indépendantes de N, il en résulte

$$G_Z(t) = E\left[\prod_{i=1}^{k} (G_Y(t))^N\right] \tag{3.4}$$

3.2.3 Fonction génératrice des moments d'une distribution de Poisson composée

Nous avons

$$\begin{aligned}
M_Z(t) &= E\left(e^{tz}\right) = E\left[E\left(e^{tz} \backslash N = n\right)\right] \text{ (par idépendance des } Y_i) \\
&= E\left([M_Y(t)]^N\right) \text{ (par indépendznce des } Y_i \text{ avec } N) \\
&= \sum_{n=0}^{\infty} [M_Y(t)]^n P(N=n) = \sum^{\infty} [M_Y(t)]^n e^{-\lambda} \frac{\lambda^{(n)}}{n!} \text{car } N \sim \mathscr{P}(\lambda) \\
&= e^{-\lambda} \sum_{n=0}^{\infty} \frac{(M_Y(t)\lambda)}{n!} = e^{-\lambda} e^{M_Y(t)\lambda} \left(\text{car } e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \forall x \in \mathbb{C}\right) \\
&= e^{\lambda[M_Y(t)-1]}
\end{aligned} \tag{3.5}$$

3.2.4 La Fonction génératrice des cumulants d'une distribution de Poisson composée

Elle s'écrit comme suit

$$K_Z(t) = \log[M_Z(t)] = \log\left[e^{\lambda[M_Y(t)-1]}\right] = \lambda[M_Y(t)-1] \tag{3.6}$$

3.2.5 Fonction caractéristique d 'une distribution de Poisson composée

nous pouvons calculer que la fonction caractéristique de la distribution de Poisson composée Z, en conditionnant

sur N, s 'exprime comme suit

$$\begin{aligned}
\varphi_Y(t) &= E(e^{itZ}) = E\left[E\left(e^{itZ}|N=n\right)\right] = E\left[\prod_{i=0}^{n} E\left(e^{itY_i}|N=n\right)\right] \text{ (par indépendance des } Y_i) \\
&= \sum_{n=0}^{\infty} \left[\varphi_Y(t)\right]^n p[N=n] = \sum_{n=0}^{\infty} \left[\varphi_Y(t)\right]^n e^{-\lambda} \frac{\lambda^n}{n!} \text{ (car } N \sim \mathscr{P}(\lambda) \text{)} \\
&= e^{-\lambda} \sum_{n=0}^{\infty} \frac{\left[\varphi_Y(t)\lambda\right]^n}{n!} = e^{-\lambda} e^{\varphi_Y(t)}.
\end{aligned} \tag{3.7}$$

où φ_Y est la fonction caractéristique des variables Y_k, alors que λ est l'intensité de N .

3.2.6 Espérance et variance d'une distribution de Poisson composée

Nous identifions l'expression de l'espérance de Z en conditionnant sur N et en utilisant la formule de l'espérance totale.

$$E(Z) = E[E(Z|N)].$$

Ainsi, l'expression de $E(Z)$ est donnée par

$$\begin{aligned} E(Z) &= E[\sum_{i=1}^{N} Y_i] = E\left[\sum_{i=1}^{N} Y_i | N = n\right] = E(Y)\,E(N) \text{ (par indépendance des } Y_i \text{ avec } N) \\ &= \lambda E(Y) \text{ (car} E(N) = \lambda). \end{aligned} \tag{3.8}$$

En actuariat, l'espérance des coûts pour un risque, $E(Z)$, correspond au produit du nombre espéré de sinistres $E(N)$ et du montant espéré d'un sinistre $E(Y)$ ou, en d'autres termes, la fréquence multipliée par la sévérité. Dans le contexte de l'assurance, l'espérance de la somme Z correspond à ce qui s'appelle la prime pure. De même, pour l'expression de la variance de Z, nous conditionnons à nouveau sur la variable aléatoire N, tout en utilisant la formule de la variance totale

$$Var(Z) = Var[E(Z|N)] + E[Var(Z|N)] \tag{3.9}$$

$$= Var\left[E\left(\sum_{i=1}^{n} Y_i | N = n\right)\right] + Var\left[E\left(\sum_{i=1}^{n} Y_i | N = n\right)\right] \tag{3.10}$$

$$\begin{aligned} &= Var\,[NE(Y)] + E[NVar\,(Y)] = [E(Y)]^2\,Var\,(N) + E(N)\,Var\,(Y) \\ &= \lambda\,(Var\,(Y)) + \lambda Var\,(Y) \quad (\text{car } E(N) = \lambda \text{et } Var\,(N)) \\ &= \lambda\left(Var\,(Y) + [E(Y)]^2\right) = \lambda\left[E\left(Y^2\right)\right] \ (\text{car } Var\,Y) = E\left(Y^2\right) - E(Y)^2) \end{aligned}$$

3.3 Distribution Poisson-Gamma

Lorsque les variables aléatoires Y_k, $k = 1, ..., n$, dans l'équation $Z = \sum_{k=0}^{N} Y_k$ sont distribuées selon une loi $Gamma(-\lambda, \gamma)$, le modèle s'appelle Poisson-Gamma. Nous posons α négatif, afin de simplifier la paramétrisation dans les formules. La distribution Poisson-Gamma est donc un cas particulier d 'une distribution de Poisson composée. Elle est courante dans le domaine de l'assurance.

Soit la densité de Y

$$f_{Y_k}(y) = \frac{1}{\gamma^{(-\alpha)}\Gamma(-\alpha)}(y)^{-\alpha-1}\exp\left(\frac{-y}{\gamma}\right) \text{ pour } y \in \mathbb{R}_+ \tag{3.11}$$

où $-\alpha < 0$ et $\gamma = \frac{-m}{\alpha}$ sont les paramètres de forme et d'échelle respectivement. De façon

générale, la fonction de densité de la somme des variables aléatoires indépendantes Y_k, pour $k = 1, 2, ..., n$, s'écrit

$$f_Y(y) = \frac{(-\alpha/m)^{\alpha n}}{\Gamma(-\alpha n)} y^{(-\alpha n-1)} \exp\left(-y(-\alpha/m)\right), \tag{3.12}$$

où $Y = Y_1 + Y_2 + ... + Y_n$. Nous savons que si dy est un nombre réel positif infiniment petit, alors la probabilité que Y soit inclus dans l'intervalle $[y, y+dy]$ est égale à $f(y)dy$, donc $p(y < y < y + dy) = f(y)dy$. Alors la distribution conjointe de N et Y est donnée, en conditionnant sur N, par la formule suivante, pour tout $\sigma^2 > 0, n > 0$ et $y > 0$.

$$\begin{aligned} f_{N,Y}(n, y, \lambda, m, \alpha)dy &= P(y < Y < y + dy | N = n)P(N = n) \\ &= \frac{(-\alpha/m)^{-\alpha n}}{\Gamma(-\alpha n)} y^{-\alpha n-1} \exp\left(-\frac{-\alpha}{m} y\right) e^{-\lambda} \frac{\lambda^n}{n!} dy. \end{aligned} \tag{3.13}$$

Cette distribution appartient à la famille de dispersion exponentielle.

3.4 Application sur l'assurance d'automobiles

3.4.1 Contexte et données

Les données utilisées dans cet exemple sont formées de 753828 observations qui représentent des données d'assurance individuelle, pour les années 2003, 2004, 2005, 2006 et 2007, où nous pouvons observer des contrats d'assurances. Parmi ces observations, il y a 703075 observations (ou 92.05%) qui ont une réponse zéro. Ces zéros exacts représentent l'absence d'accident ou de remboursement par véhicule assuré pour une année.

Le tableau 3.1 représente les différentes statistiques : moyenne, médiane, coefficient de variation (CV), asymétrie, 95ême percentile et 99ême percentile, pour l'ensemble de données.

TABLE 3.1 – Tableau représente les déffirents Statistiques sur les coûts totaux individuels d'assurance.

moyenne	médiane	cv	0%	asymétrie	95%	99%
529.67	0	1916.436	92.05	128.38	1598.20	9118.18

La figure 3.1 illustre la répartition des observations en fonction des coûts. Nous remarquons une grande masse de probabilité à 0, étant donné que la majorité des observations sont des zéros.

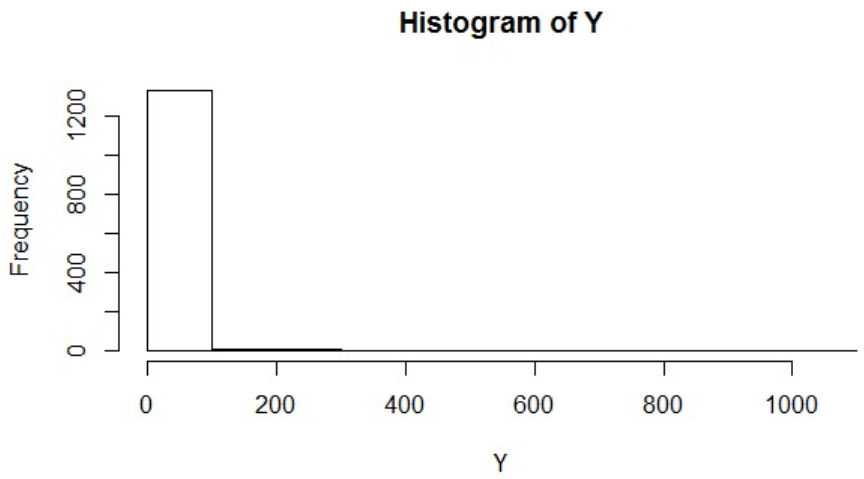

FIGURE 3.1 – Histogramme de distribution de perte

3.4.2 Analyse empirique

Nous effectuons plusieurs permutations aléatoires des observations. A chaque permutation,nous répartissons les observations en 760 échantillons de 1000 observations. La figure 3.2 illustre la relation log-variance échantillonnale et log-moyenne échantillonnale. Les parcelles sont presque linéaires, pour toutes les permutations, ce qui suggère que la variance est proportionnelle à une certaine puissance de la moyenne ; c'est-à-dire $V(Z) = \sigma^2 \mu^p$, donc la distribution Tweedie est suggérée dans ce cas.

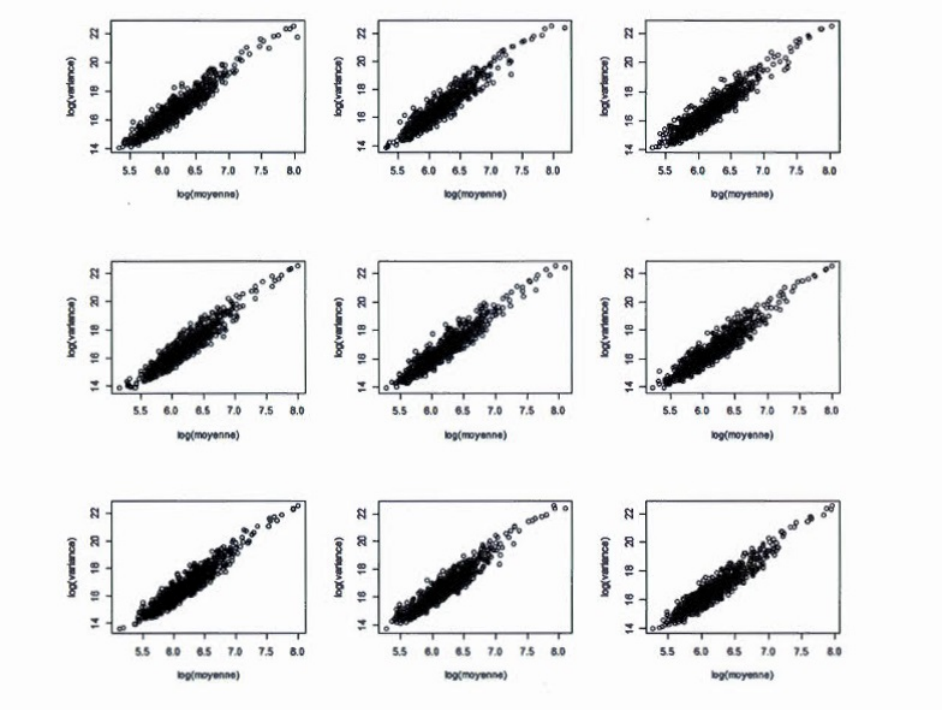

FIGURE 3.2 – Graphique qui montrent la relation moyenne_variance pour plusieurs permutations

Dans ce cas, les coûts peuvent êtres exprimés comme suit

$$Z = \begin{cases} \sum_{i=1}^{N} Y_i, & pour N > 0 \\ 0, & sinon. \end{cases} \tag{3.14}$$

où Z est le total des réclamations, N représente le nombre de réclamations et Y_i représente le montant du ième sinistre ou réclamation. Les données sont continues avec des zéros exacts, cela signifie que la distribution mixte Poisson-Gamma peut être utilisée pour modéliser les coûts totaux d'assurance individuels.

3.5 Ajustement de la distribution Poisson-Gamma

Afin d'ajuster le modèle Poisson-Gamma aux données, nous devons estimer les paramètres inconnus. Nous avons trois paramètres à estimer, la moyenne μ, le paramètre de dispersion σ^2 et l'indice de la distribution p.

Pour simplifier les calculs, nous ignorons la variabilité d'échantillonnage dans les paramètres estimés, et nous supposons aussi que les observations z_i sont indépendantes. Comme la densité du modèle Poisson-Gamma est approximée avec précision par lesméthodes vues précédemment, nous pouvons utiliser cette densité approximée par

l'une de ces méthodes pour calculer les estimateurs du maximum de vraisemblance

des paramètres inconnus dans le modèle.

La fonction tweedie.profile, dans R(cf. Dunn, (2013)), donne une estimation des paramètres σ^2 et p utilisant la fonction de log-vraisemblance de la densité Tweedie.

La fonction de log-vraisemblance est calculée par l'une des méthodes vues au chapitre 2.

La fonction de log-vraisemblance pour estimer les paramètres du modèle est donnée par

$$\log L = \sum_{i=1}^{n} \log f\left(z_i, \mu_i, \sigma^2\right)$$

Pour différentes valeurs de p, nous estimons σ^2 par la méthode du maximum de vraisemblance. À partir de l'estimation de σ^2, nous pouvons obtenir une estimation précise de p en utilisant une fonction profil de vraisemblance (cf. Smyth, (1996)). Nous calculons la fonction de log-vraisemblance pour les valeurs de σ^2 estimées. Les paramètres p et σ^2 pour lesquels la log-vraisemblance est un maximum sont choisis.

3.6 Adéquation du modèle

Pour illustration, nous utilisons un échantillon aléatoire de 1000 observations, l'estimation du maximum de vraisemblance de pour le modèle ajusté est $\hat{\sigma}^2 = 262.18$ et l'estimateur du maximum de vraisemblance de σ^2 est $\hat{P} = 1.64$. Le fait que $-2logL$ soit asymptotiquement de loi χ_2 Krzanowski, (1998) peut nous aider à construire un intervalle de confiance pour p. Pour cet échantillon, l'intervalle de confiance à 95% de p est donné par $IC = (1.59, 1.69)$ qui est calculé par la fonction tweedie.profile dans R. La figure 3.3 illustre la fonction de log-vraisemblance profil en fonction de p, elle nous permet de voir la valeur de p qui maximise la log-vraisemblance.

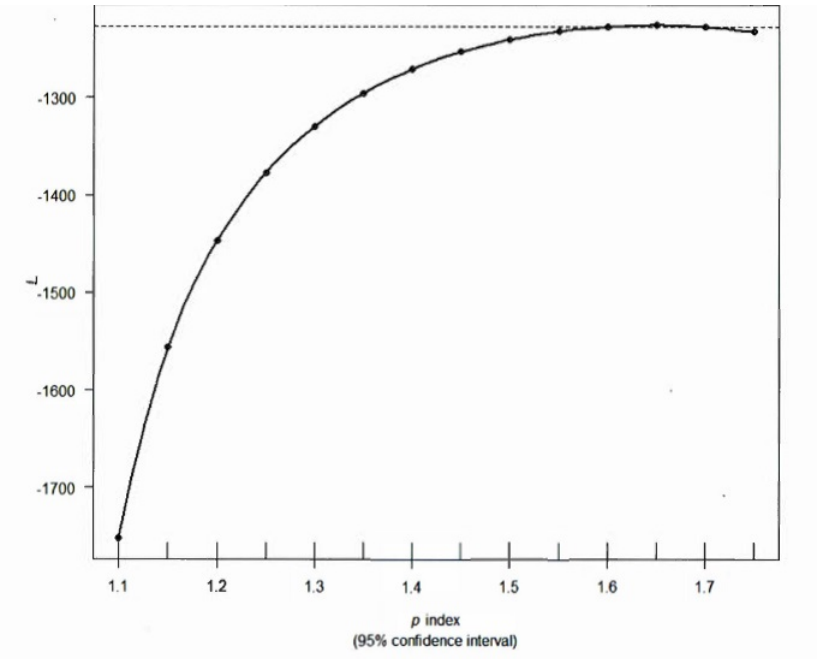

FIGURE 3.3 – La fonction log vraisemblance profil pour les réclamation individuelles d'assurance. L'estimation du maximum de vraisemblance de p pour un échantillon de 1000 observations est de 1.64, avec un intervalle de confiance à 95 pour cent qui est $[1.59, 1.69]$

3.6.1 Étude des résidus quantiles du modèle ajusté

Afin de juger l'adéquation du modèle Tweedie ajusté, nous étudions les résidus de ce modèle. (cf. Dunn et Smyth, (1996)) ont trouvé utile de calculer les résidus quantiles pour vérifier l'adéquation du modèle Tweedie. À cet effet, les résidus quantiles sont définis comme suit

$$r_{i,q} = \Phi^{-1}\left\{F\left(z_i, \hat{\mu}, \hat{\sigma}^2\right)\right\}. \tag{3.15}$$

L'estimation du maximum de vraisemblance de p pour un échantillon de 1000 observations est de 1.64, avec un intervalle de confiance à 95% qui est $[1.59, 1.69]$. Où F est la fonction de répartition d 'une distribution continue quelconque et σ^2est la fonction de répartition cumulative de la distribution normale standard. L'hypothèse de normalité des résidus inclut les distributions appartenant aux EDM (cf. Dunn et Smyth, (1996)), ce qui signifie que les résidus de la distribution Poisson-Gamma sont normalement distribués. Dans notre cas, les résidus quantiles sont calculés en inversant la fonction de répartition ajustée de la distribution de Tweedie à chaque valeur z_i, puis en trouvant le quantile normal standard équivalent.

Par ailleurs, les observations sont continues avec des zéros exacts. Dans ce cas, les résidus quantiles utilisent la randomisation L seulement pour les observations pour lesquelles la réponse est exactement zéro (cf. Dunn et Smyth, (1996)), c'est-à-dire que la randomisation est utilisée pour produire des résidus distribués en continu.

Les résidus quantiles utilisent le minimum nécessaire de randomisation de sorte qu'aucune granularité

(la plus petite valeur possible) ne reste dans la distribution des résidus (cf. Dunn et Smyth, (1996)). Dans ce cas, les résidus quantiles sont définiscomme suit : soit $a_i = \lim_{z \uparrow z_i} F\left(z_i; \hat{\mu}_i, \hat{\sigma}^2\right)$ et $b_i = F\left(z_i; \hat{\mu}_i, \hat{\sigma}^2\right)$, nous avons alors

$$r_{q,i} = \Phi^{-1}\{u_i\}.$$

où μ_i est une variable aléatoire uniforme sur l'intervalle $(a_i, b_i]$. Dans ce cas, les résidus $r_{q,i}$ sont normalement distribués. Lorsque tous les points sont situés sur la ligne droite, cela indique un ajustement parfait, tandis que les points situés loin de la ligne indiquent un mauvais ajustement du modèle. Dans ces figures, les statistiques d'ordre de l'échantillon sont tracées en fonction de l'ordre attendu des statistiques d'ordre normal. Le tracé des résidus quantiles de la distribution normale à la figure 3.4 montre que tous les résidus se trouvent sur ou à proximité de la ligne à part quatre valeurs extrêmes. Cela signifie que la distribution Poisson-Gamma avec $\hat{p} = 1.64$ ajuste bien les totaux des coûts individuels, et que le modèle suggère un ajustement adéquat des données.

Les quatre points qui sont loin de la ligne peuvent représenter des sinistres graves en assurance automobile qui mériteraient des analyses plus poussées, voir (cf. Doucet, (2014)) pour plus de détails.

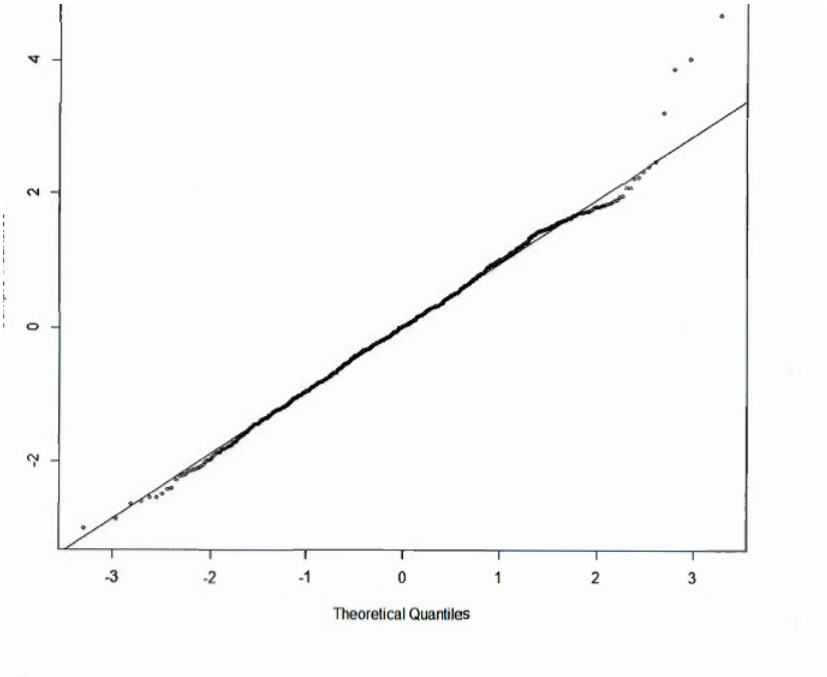

FIGURE 3.4 – Résidus quantiles de distribution normale appliqués sur un échantillon de 1000 observations des réclamations individuelles d'assurance.

3.6.2 Étude des percentiles

Afin de tester la performance du modèle ajusté, nous pouvons vérifier si le modèle ajusté simule des données ayant des caractéristiques similaires aux données observées.Nous estimons les paramètres de la distribution Poisson-Gamma pour un échantillon aléatoire de 1000 observations. Soit de nouveau $\hat{p} = 1.64$, $\hat{\sigma}^2 = 262.18$ et $\hat{\mu} = 536.72$. Comme le modèle est linéaire, alors la moyenne estimée il n'est que la moyenne de l'échantillon X. Nous calculons la médiane, le 90ême percentile, le 95ême percentile et le ggême percentile, de cet échantillon, et les résultats sont donnés dans le tableau 3.2

D'autre part, nous simulons 1000 échantillons, de taille 1000 chacun, à partir de la distribution Poisson-Gamma ajustée, avec les paramètres estimés $\hat{P}$, $\hat{\mu}$ et $\hat{\sigma}^2$. Ensuite, nous calculons les percentiles simulés pour chaque échantillon : nous aurons 1000 valeurs pour chaque percentile différent. Puis, nous calculons la médiane des 1000 percentiles simulés : nous aurons une seule valeur pour chaque percentile différent.

Nous construisons des intervalles de confiance empiriques à 95%, à partir des statistiques estimées par les données simulées selon le modèle de Tweedie ajusté.

Ces statistiques sont considérées comme des représentants des événements faibles, moyens et élevés des coûts totaux individuels. Les résultats de la simulation sont représentés dans le tableau 3.2.

TABLE 3.2 – Tableau de Comparaison entre les percentiles observés et les percentiles similaires des données simulées pour un échantillon de 1000.

	médiane	90êmepercentile	95ême percentile	99ême percentile
observé	0	0	1920.44	9815.80
Simulé	0	0	2467.38	14121.42

Le tableau 3.2 montre que le 90 ême percentile observé est nul et égal au percentile similaire des données simulées ; ceci est dû au fait que la majorité des observations sont des zéros exacts. Pour le 95ême percentile, nous voyons que la valeur du percentile observé est proche de celle simulée, contrairement au 99 ême percentile, la valeur observée est un peu loin de la valeur simulée.

Nous estimons un intervalle de confiance empirique à 95% pour le 95ême percentile simulé afin de vérifier si le 95 ême percentile observé appartient à cet intervalle. Étant donné un intervalle de confiance $IC = [498.61, 7036.7]$, nous remarquons que le 95 ême percentile observé (1920.44) appartient à l'intervalle de confiance, ce qui indique que la distribution Poisson-Gamma simule des données avec des propriétés très similaires aux données observées.

Nous répétons le test précédent pour 100 échantillons aléatoires de taille 1000 observations chacun. Nous vérifions si les 95êmes percentiles observés appartiennent aux intervalles de confiance empiriques. La figure 3.5 illustre les résultats des 95êmes

percentiles observés et les intervalles de confiance empiriques estimés à 95% correspondants.

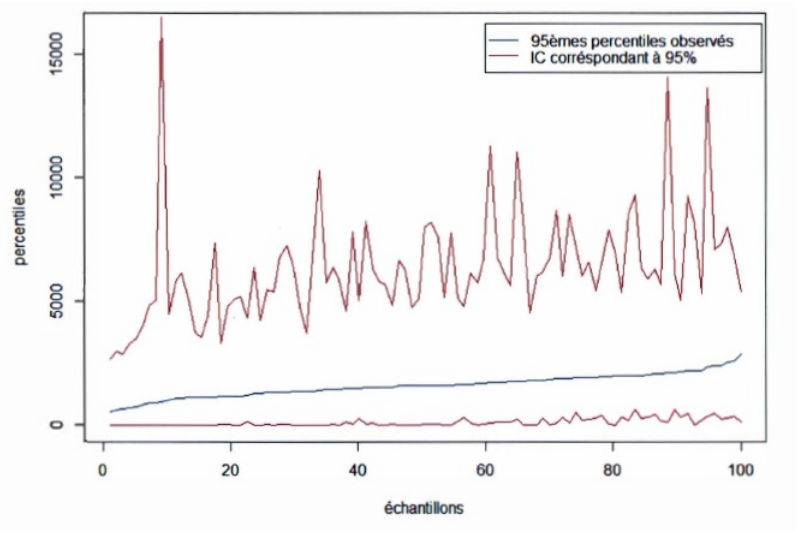

FIGURE 3.5 – Résidus quantiles de distribution normale appliqués sur échantillon de 1000 des réclamations individuelles d'assurance.

Nous remarquons que tous les percentiles observés se trouvent dans l'intervalle empirique correspondant. Donc, l'intervalle de confiance de 95% des statistiques simulées contient les statistiques observées respectivement à 100% des cas, cequi signifie que le modèle est bien ajusté et que la distribution Poisson-Gamma pourrait être suggérée pour modéliser les coûts individuels d'assurance.

3.6.3 Comparaison des estimateurs obtenus par les méthodes d'approximation de la densité de Tweedie

Pour comparer les paramètres estimés, $\hat{p}$ et $\hat{\sigma}^2$, obtenus par l'estimation du maximum de vraisemblance des quatre méthodes d'estimation de densité : inversion de Fourier, séries infinies, interpolation et point-selle, nous tirons un échantillon aléatoire de taille 1000 observations et nous estimons les paramètres par les méthodes précédentes. Le tableau ≥ illustre les résultats de l'estimation.

TABLE 3.3 – Tableau de Comparaison des paramètres $\hat{p}$ et $\hat{\sigma}^2$, estimés par la méthode du maximum de vraisemblance pour les quatre approches : inversion de Fourier, séries infinies, interpolation et point-selle.

Estimateurs	Inversion de Fourier	Séries infinies	Interpolation	Point-selle
$\hat{p}$	1.678571	1.678571	1.676531	1.75
$\hat{\sigma}^2$	288.4715	288.4715	284.4813	157.2945

Dans le tableau 3.3, nous remarquons que les estimateurs de p et σ^2; sont identiques pour la méthode d'inversion de Fourier et la méthode des séries infinies. Pour la

méthode d'interpolation, l'estimateur de p change à partir du 3ême chiffre après la virgule, par rapport aux 2 méthodes précédentes. L'estimateur de σ^2, par cette méthode, est plus petit de 3.9902, par rapport aux 2 méthodes précédentes. En ce qui concerne la méthode de point-selle, elle donne des estimateurs qui sont très différents par rapport aux estimateurs des autres méthodes, étant donné quela méthode de point-selle est seulement une méthode approximative qui n'estime pas la densité avec précision, comme nous l'avons déjà signalé précédemment.

3.7 Utilisation de la distribution Tweedie en assurance

3.7.1 Primes et principes de prime

La prime est un paiement chargé par un assureur qui permet aux assurés de bénéficier d'une couverture d'assurance, complète ou partielle, contre un risque financier (cf. Keller, (2008)). Les risques financiers sont considérés comme des variables aléatoires non négatives avec fonction de densité f et fonction génératrice des moments $M(t)$.

En pratique, le montant de la prime dépend de nombreux facteurs : les assureurs doivent tenir compte non seulement des caractéristiques des risques qu'ils assurent, mais aussi d'autres facteurs tels que les primes considérées par leurs concurrents. .

Nous désignons par Π_X la prime chargée par l'assureur pour couvrir un risque X. Dans notre cas, la variable aléatoire X représente les réclamations individuelles d'assurance automobile. La règle qui attribue une valeur numérique à Π_X est appelée le principe de prime, qui est de la forme $\Pi_X = \varrho(X)$, où $\varrho(X)$ est une fonction quelconque.

3.7.2 Propriétés des principes de prime

Il existe de nombreuses propriétés sur les principes de calcul des primes, parmi les propriétés de base (cf. Dickson., (2005), (cf. Teugels.,(2004)), nous trouvons

1. Marge de sécurité positive : $\Pi_X \geqslant E(X)$, c'est-à-dire que la prime ne doit pas être inférieure au montant espéré des réclamations.
2. Additivité : Si X_1et X_2 sont deux risques indépendants, alors la prime pour le risque combiné est égale à la somme des deux primes individuelles, il $X_1 + X_2 = \Pi_{X_1} + \Pi_{X_2}$.
3. Proportionnalité : Pour toute constante $a > 0$, si $Z = aX$, alors $Z = a\Pi_X$, c'est-à-dire qu'en cas de changement d'échelle du montant des réclamations, cette propriété garantit à l'assureur le même niveau de bénéfice.

4. Cohérenc : Si $Y = X + c$, où $c > 0$, nous aurons alors $\Pi_Y = \Pi_X + c$. Ainsi, si la distribution de Y est décalée de c unités de la distribution de X, alors la prime pour le risque Y est celle du risque X augmentée de c unités.
5. Plafonnement : Soit x_m le montant de sinistre maximal pour un risque X, nous aurons alors $\Pi_X \geqslant \Pi_m$. Si cette condition n'est pas satisfaite, il n'y aurait aucune raison pour l'assuré de souscrire à la couverture d'assurance.

Les principes de calcul des primes d'assurance sont des méthodes qui autorisent une compagnie d'assurances de calculer la prime en fonction du risque X. Il existe de nombreuses méthodes différentes pour calculer les primes d'assurances, mais dans cette section nous nous concentrons sur le principe d'Esscher. Nous présentons d 'abord quelques exemples importants du principe de prime, ensuite nous appliquons le principe d'Esscher sur la densité Tweedie en utilisant un échantillon aléatoire de 1000 observations des coûts totaux individuels d 'assurance.

Exemple 36 *Principe de la prime nette :* $\Pi_X = E(X)$ *, la prime est égale au montant espéré des réclamations. Ce principe respecte toutes les propriétés debase, mais il n 'inclut pas les marges de sécurité et ne peut donc être utilisé pour établir la prime finale.*

Exemple 37 *Principe de la valeur espérée :* $\Pi_X = \delta(1+)E(X)$*, pour* $\delta > O$*. Le montant espéré des réclamations est majoré d'une marge de sécurité* $\delta E(X)$*. Enpratique, l'inconvénient majeur de ce principe est qu'il accorde la même marge aux risques qui ont le même montant espéré. Les risques avec des moyennes identiques mais des variances différentes devraient avoir des primes différentes.*

Exemple 38 *Principe de la variance :* $\Pi_X = E(X) + \delta Var(X)$*, pour* $\delta > 0$*. Ce principe est différent du principe de la valeur espérée, car la marge de cette prime est proportionnelle à la variance du risque. Cette modification permet de donner une prime différente à des risques qui n'ont pas la même variance. Comme* $\delta > 0$*, alors le principe de la variance a une marge non négative.*

Exemple 39 *Principe de l'écart-type :* $\Pi_X = E(X) + \delta\sqrt{Var(X)}$*, pour* $\delta > 0$*. La marge dans ce cas est proportionnelle à l'écart-type du risque* X *.*

3.8 Principe de la prime d'Esscher

Le principe d'Esscher produit également une pondération du risque X. Il a été introduit par (cf. Bühlmann, (1980)) et (cf Gerber, (1980)). Ce principe découle du principe de l'utilité exponentielle. Il survient lorsque l'assureur vise à optimiser son utilité selon le principe de l'utilité équivalente. La fonction d 'utilité exponentielle de la prime d 'Esscher est de la forme

$$u_x = \frac{1 - \exp(-cx)}{c} \tag{3.16}$$

où c une constante supérieure à 0 qui est égale à $-\frac{d}{d}\ln u$, et mesure l'aversion au dx risque[2] pour une compagnie d'assurance (c.f. Denuit et al., (2006)). La fonction d'utilité exponentielle (3.16) donne une aversion au risque constant.

Pour couvrir un risque X, la prime fix est déterminée en maximisant l'utilité du contrat, $E[u(\Pi_X - X)]$ sur toutes les fonctions croissantes continues w, telles que $E[w(X)] = 1$, ce qui donne

$$w(x) = \frac{\exp(cx)}{M_X(c)} \tag{3.17}$$

le poids $w(x)$ dans l'equation(3.17) donne une prime égale à

$$E[Xw(X)] = E[X\frac{\exp(cX)}{M_X(c)}] = \int_0^{+\infty} x\frac{\exp(cx)f_X(x)}{M_X(c)}dx \tag{3.18}$$

Par ailleurs, Goovaerts et al ., (1984) présentent la prime d 'Esscher comme la valeur attendue du risque X après avoir multiplié la densité de X par une fonction de pondération croissante, ce qui rend le risque moins attrayant pour l'assureur.

Nous pouvons définir le principe d'Esscher comme une mesure de risque, qui est donnée par

$$\begin{aligned} \Pi_X &= \frac{E\left[X\, e^{hX}\right]}{E\left[e^{hX}\right]} = \frac{\int_0^{\infty} xe^{hx}f(x)dx}{\int_0^{\infty} e^{hx}f(x)dx} \\ &= \frac{\frac{d}{dh}M_X(h)}{M_X(h)} = \frac{d}{dh}\log M_X(h) = E[\tilde{X}] \end{aligned} \tag{3.19}$$

où $h > 0$ et $M_X(h)$ est la fonction génératrice des moments de la variable X. Π_X est considérée ici comme la valeur espérée de la transformée d'Esscher$\tilde{X}$ de X : c'est la prime pure de la transformée d'Esscher du risque initial X.

Dans le domaine actuariel, la transformée d'Esscher (cf. Gerber et al., (1994)) est une transformation qui prend une densité de probabilité. $f(x)$ et la transforme en une nouvelle densité de probabilité $g(x)$ avec un paramètre h, tel que $h > 0$. Le but de cette transformation est d 'approximer la distribution du montant des réclamations globales autour d'un point d'intérêt x_0, le paramètre h étant choisi de telle sorte que la nouvelle moyenne soit égale à x_0. Soit $f(x)$ la densité d'une variable aléatoire non négative X, la transformée d'Esscher de. $f(x)$ est donnée par la fonction $g(x)$ qui est supposée être la fonction de densité de la variable aléatoire $\tilde{X}$. Soit g une fonction définie par

$$g(x) = \frac{e^{hx}f(x)}{\int_0^{\infty} e^{hx}f(x)dx} \tag{3.20}$$

oú

$$\int_0^{\infty} e^{hx}f(x)dx = M_X(h) \tag{3.21}$$

La densité $g(x)$ est une fonction pondérée de la densité. $f(x)$, car nous pouvons l'écrire

sous la forme $g(x) = w(x)f(x)$, avec

$w(x) = e^{hx}/M_X(h)$, c'est-à-dire que nous attribuons un poids $w(x)$ à chaque valeur x. Puisque $h > 0$ et $w\prime(x) > 0$, alors le poids augmente lorsque x augmente. La fonction de répartition de la variable aléatoire X est donnée par

$$G_{\tilde{X}}(x) = \int_0^\infty w(y)f(y)dy = \frac{\int_0^\infty e^{hy}f(y)dy}{M_X(h)} \tag{3.22}$$

Cette fonction est appelée la transformée d'Esscher de la fonction de répartition de X avec le paramètre h. Nous pouvons calculer la fonction génératrice des moments de la variable X, à partir de la fonction génératrice des moments de la variable X, comme suit

$$M_{\tilde{X}}(t) = \int_0^\infty e^{tx}g(x)dx \tag{3.23}$$

De l'equation (3.20) nous avons

$$M_{\tilde{X}}(t) = \int_0^\infty e^{tx}g(x)dx = \frac{\int_0^\infty e^{tx}e^{hx}f(x)dx}{\int_0^\infty e^{hx}f(x)dx} = \frac{M_X(t+h)}{M_X(h)} \tag{3.24}$$

La prime d'Esscher Π_X est calculée à partir de la relation dans l'équation (3.24). Pour $h = 0$, nous avons $M_{\tilde{X}}(t) = M_X(t)$, donc

Exemple 40 $E[\tilde{X}] = \Pi_X = E[X]$. *De façon générale,*

$$E[\tilde{X}^n] = \frac{d^n}{dt^n}M_{\tilde{X}}(t)|_{t=0} = \frac{d^n}{dt^n}\frac{M_X(t+h)}{M_X(h)}|_{t=0} = \frac{M_X^n(h)}{M_X(h)} \tag{3.25}$$

Le principe d'Esscher satisfait la propriété de marge de sécurité positive. À partir de la formule (3.25), nous avons

$$\frac{d}{dh}\Pi_X = \frac{d}{dh}E[\tilde{X}] = \frac{d}{dh}\frac{M_X^{'}(h)}{M_X(h)} = \frac{M_X^{(2)}(h)M_X(h)}{M_X(h)^2} - \frac{M_X^{'}(h)^2}{M_X(h)^2} = E[\tilde{X}^2] - E[\tilde{X}]^2 \geq 0 \tag{3.26}$$

Par conséquent, Π_X est une fonction non décroissante de h, donc Π_X ~$E[X]$, pour tout h~O. D'autre part, le principe d'Esscher satisfait la propriété de cohérence. Soit $Y = X + c$, la prime d'Esscher de la variable Y s'écrit

$$\begin{aligned}
\Pi_Y &= \frac{E\left[Ye^{hY}\right]}{E\left[e^{hY}\right]} = \frac{E\left[(X+c)\,e^{h(X+c)}\right]}{E\left[e^{h(X+c)}\right]} \\
&= \frac{E\left[Xe^{h(X)}\right]e^{hc} + cE\left[e^{h(X)}\right]e^{hc}}{E\left[e^{hX}\right]} \\
&= \frac{E\left[Xe^{hX}\right]}{E\left[e^{hX}\right]} + c = \Pi_X + c.
\end{aligned} \tag{3.27}$$

Le principe d 'Esscher satisfait également la propriété d 'additivité

$$\begin{aligned}\Pi_{X_1+X_2} &= \frac{E\left[(X_1+X_2)\,e^{hX_1+X_2}\right]}{E\left[e^{hX_1+X_2}\right]} \\ &= \frac{E\left[(X_1)\,e^{h(X_1)}\right]E\left[e^{h(X_2)}\right]+E\left[(X_2)\,e^{h(X_2)}\right]E\left[e^{h()}\right]}{E\left[e^{h(X1)}\right]E\left[e^{h(X_2)}\right]} \\ &= \frac{E\left[X_1e^{hX_1}\right]}{E\left[e^{hX_2}\right]}+\frac{E\left[X_2e^{hX_2}\right]}{E\left[e^{X_2}\right]}=\Pi_{X_1}+\Pi_{X_2}.\end{aligned}$$

La propriété de plafonnement est aussi satisfaite puisque si X_m est le plus grand montant de réclamation possible, de sorte que $P(X<x)=1$, alors

$$Xe^{hX}\leq x_me^{hX}$$

$$\Pi_X=\frac{E\left[Xe^{hX}\right]}{E\left[e^{hX}\right]}\leq\frac{E\left[x_me^{hx_m}\right]}{E\left[e^{hx_m}\right]}=x_m$$

Cependant, le principe d'Esscher n'est pas invariant à l'échelle. Soit $Z=aX$, la prime d 'Esscher

$$\Pi_Z(h)=\frac{E\left[Ze^{hz}\right]}{E\left[e^{hz}\right]}\Pi_Y=\frac{E\left[aXe^{haX}\right]}{E\left[e^{hax}\right]}=a\Pi_X(ah)\neq\Pi_X(ah)$$

La transformée d 'Esscher peut transformer les distributions de famille de dispersion exponentielle, à cet effet, voici quelques exemples qui sont donnés dans Kaas et al., (2008)

1. $\mathscr{N}(0,1)$ en $\mathscr{N}\left(\mu,1\right)$, pour $h=\mu$;
2. $\mathscr{P}(1)$ en $\mathscr{P}\left(\mu\right)$, lorsque $h=\log\mu$;
3. $\mathscr{B}(m,1/2)$ en $\mathscr{B}\left(m,p\right)$, lorsque $p=1/\left(1+e^{-h}\right)$, donc $h=\log\left[p/\left(1-p\right)\right]$;
4. $\mathscr{BN}(r,1/2)$ en $\mathscr{BN}\left(r,p\right)$, lorsque $p=1-\frac{1}{2}e^h$, donc $h=\log\left[p/\left(1-p\right)\right]$;
5. $Gamma(1,1)$en $Gamma\left(1,\beta\right)$, lorsque $p=1-\frac{1}{2}e^h$, donc $h=\log\left[p/\left(1-p\right)\right]$;
6. $\mathscr{IG}(1,1)$en $\mathscr{IG}(1,\beta)$, losque $\alpha=(1-2h)^{1/2}$, donc $h=(1-2\alpha)/2$.

Dans le cas d'une distribution de Tweedie, la fonction génératrice des moments est donnée par

$$M_X(h)=\exp\left\{\lambda\left[\left(1-h\gamma\right)^{\alpha}-1\right]\right\}.$$

avec $\alpha=\left(2-p\right)/\left(1-p\right)$, $\lambda=\mu^{(2-p)}/\phi\left(2-p\right)$ et $\gamma=\phi\left(p-1\right)\mu^{(p-1)}$. Donc la fonction $g(x)$ dans la formule (3.20)devient

$$g(x)=\frac{e^{hx}f(x)}{\exp\left\{\lambda\left[\left(1-h\gamma\right)^{\alpha}-1\right]\right\}}. \tag{3.28}$$

À partir de la formule (3.24), la fonction génératrice des moments, d'une distribution Tweedie, pour un risque X s'écrit comme suit

$$\begin{aligned} M_{\tilde{X}}(t) &= \frac{M_X(t+h)}{M_X(h)} = \frac{\exp\left\{\lambda\left[(1-(t+h)\gamma)^{\alpha}-1\right]\right\}}{\exp\left\{\lambda\left[\left(1-h\gamma\right)^{\alpha}\right]-1\right\}} \\ &= \exp\left\{\lambda[\left(1-(t+h)\gamma\right)^{\alpha}-\left(1-h\gamma\right)^{\alpha}]\right\}. \end{aligned} \tag{3.29}$$

La figure 3.6 illustre les courbes des deux fonctions, $f(x)$ et $g(x)$, où f représente la fonction de densité de la distribution Tweedie pour un échantillon de 1000 observations des réclamations individuelles d'assurance automobile, alors que g est la fonction pondérée de la fonction f, obtenue par la transformée d'Esscher de la fonction f.

La fonction de poids $w(x)$ dans $g(x)$ met plus de masse sur les grandes valeurs de X, ce qui implique un chargement de sécurité. Nous observons sur la figure 3.6 que la densité $g(x)$ est inférieure à la densité. $f(x)$; cependant, la densité $g(x)$ est supérieure à la densité $f(x)$ dans la queue, de sorte que la transformation se traduit par une densité avec une queue plus grosse, ceci rend le risque moins attrayant pour l'assureur.

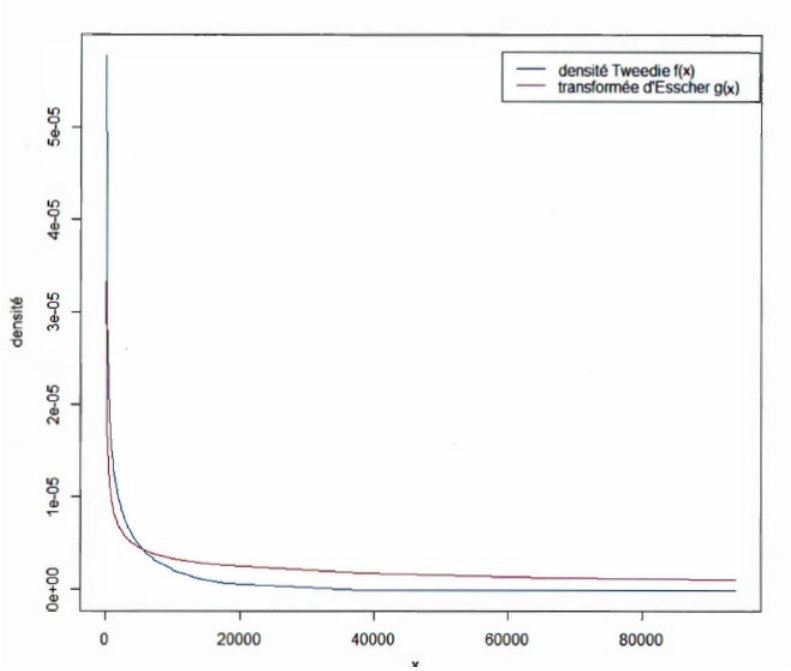

FIGURE 3.6 – Graphique qui représente la densité de la distribution $f(x)$ et la transformée d'Esscher de cette densité $g(x)$, pour $h = 0.0001$.

À partir d'un échantillon aléatoire de 1000 observations des réclamations individuelles d'assurance automobile, nous estimons les paramètres de la densité Tweedie, $\hat{p} = 1.64$, $\hat{\mu} = 536.73$, ($\hat{\sigma}^2 = 262.18$. Pour mieux voir le rôle de la transformée d'Esscher, les observations qui sont des zéros sont enlevées sur la figure , et la masse de ces ob-

servations est de 0.9044.

Les valeurs des paramètres de la distribution Tweedie, selon cet échantillon, sont $\alpha = -0.553$, $\lambda = 0.1004463$ et

$\gamma = 9661.078$. Donc, la fonction génératrice des moments $M_Y(h) = 1.7377$, pour $h = 0.0001$.

En utilisant l'équation 3.19, la prime d'Esscher pour la distribution Tweedie est comme suit

$$\Pi_X = \frac{d}{dh}\log M_X(h) = \frac{d}{dh}\left\{\lambda\left[(1-h\gamma)^{\alpha-1}\right]\right\} = \lambda\left[\alpha(-\gamma)(1-h\gamma)^{\alpha-1}\right].$$

Dans le cas de notre échantillon, le calcul de la prime donne $\Pi_X = 102954.3$, pour $h = 0.0001$. Le paramètre h est choisi selon l'intérêt de l'assureur.

TABLE 3.4 – Tableau Calcul de la prime pour différentes valeurs de h.

h	10^{-4}	10^{-5}	10^{-6}	10^{-7}	10^{8}	10^{-9}	10^{-10}	10^{-11}
Π_X	102954.3	628.47	544.88	537.54	536.81	536.74	536.73	536.73

Le paramètre h d 'Esscher reflète le degré d'aversion du risque pour l'assureur. Le tableau 3.4 représente les valeurs de la prime Π_X en fonction du paramètre h. Nous observons dans le tableau 3.4 que les valeurs de la prime deviennent stables et égales à la moyenne à partir de $h = 10^{-10}$ (h proche de zero). La figure 3.7 montre comment la prime est stable pour les très petites valeurs de h, ce qui confirme que $E[\tilde{X}] = \Pi_X = E[X]$ pour $h = 0$.

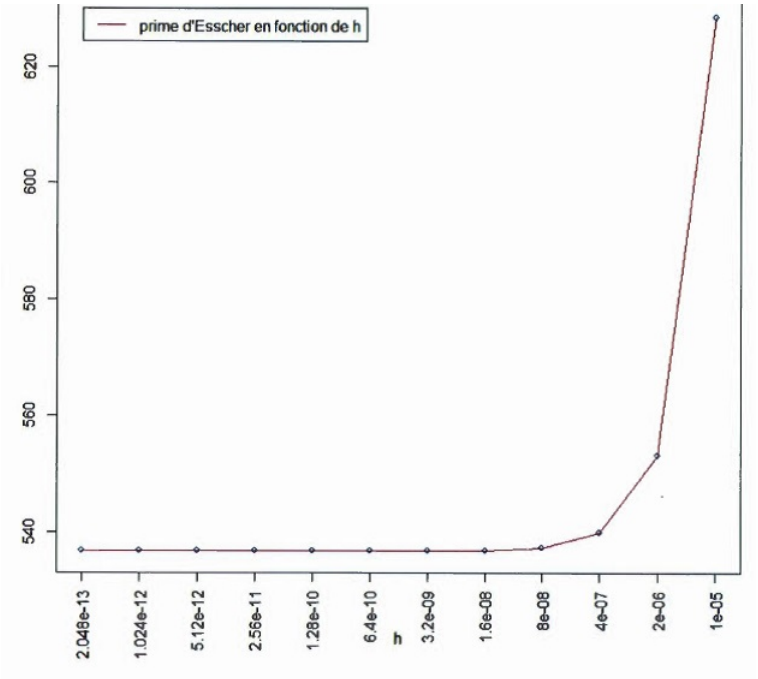

FIGURE 3.7 – Graphique représentant la prime d'Esscher Π_X en fonction de paramètre h.

3.9 Étude de cas :Réclamation pour dommages corporels à l'automobile

Cette section présente le travail appliqué que j'ai effectué pour modéliser la gravité avec une puissance de Tweedie optimisé. Pour les ensembles de données de gravité, les actuaires utilisent généralement une distribution gamma ou gaussienne inverse pour décrire la distribution de la variable cible. Les modèles gamma correspondent à une puissance Tweedie de 2 et les modèles gaussiens inverses correspondent à une puissance Tweedie 3. Je veux évaluer si nous pouvons améliorer la vraisemblance d'un modèle de gravité en estimant la vraie puissance de Tweedie .

Nos données de gravité ont été collectées par le conseil de la recherche en assurance en 2002. Nous accédons aux données du package R InsuranceData. L'ensemble de données est intitulé AutoBi, ce qui signifie blessures corporelles automobiles. La couverture des dommages coroporelles couvre les frais des personnes blessées dans un accident de voiture. Ces frais peuvent impliquer des frais médicaux, des frais juridiques, des frais funéraires et une perte de revenu. Étant donné que le soins médicaux et les services juridiques coûtent cher en Amérique, vous pouvez imaginer que les réclamations pour blesseurs corporelles ont des paiements élevés. Outre le données de la colonne sur les frais de réclamation, AutoBi contient six autres variables catégorielles. Ces variables expliquent si le demandeur s'est fait représenter par un avocat, si le demandeur portait une ceinture de sécurité, si le demandeur avait une assurance, l'âge du demandeur, le sexe de demandeur et l'état matrimonial du demandeur. Autotal, l'ensemble de données contient 1340 systèmes.

Le modèle que nous construisons utilise AGE, AVOCAT et CENTURE DE SÉCURITÉ comme variables explicatives AVOCAT est une variable binaire, CENTURE DE SÉCURITÉ a des catégories oui, Non,sans objet et AGE est divisé en gros entre 10 et 20 ans. Seules ces trois variables ont diminué le critère d'information d'Akaike. Nous nous attendons à ce que AVOCAT soit une covariable fortement prédictive dans mon modèle, car elle capture les coût associés à la représentation jurdique. L'utilisation d'une ceinture de sécurité devrait réduire le risque de décès et blessures graves, il est donc logique que la CEINTURE DE SÉCURITÉ soit prédictive .Je suis moins confiant en mettant AGE dans notre modèle. ll fournit moins d'importance et seulement une diminution marfinale de l'AIC. Néanmoins,nous voulons avoir au moins quelques covariables dans le modèle pour différencier les assurés. La modélisation avec seulement trois variables explicatives n'est pas idéale. Ce manque de variables ne parvient pas idéaile.ne parvient pas capturer suffisamment de signal sur la cible.

Je voudrais modéliser avec des variables plus explicatives pour les produits d'assurance réels. En tout état de cause, cette étude vise à évaluer l'utilité de régler la puissance de Tweedie. Ci-dessous, nous affichons le graphique de profil log-vraisemblance

réalisé pour éstimer la puissance de Tweedie. Selon cette technique, la puissance de Tweedie est trés probablement 2,3 et le paramètre de dispersion est trés probablement 1,110.Nous créons un modèle de Tweedie avec ces choix de paramètres et un modèle gamma basé sur 75 pourcent de données d'entrainent.

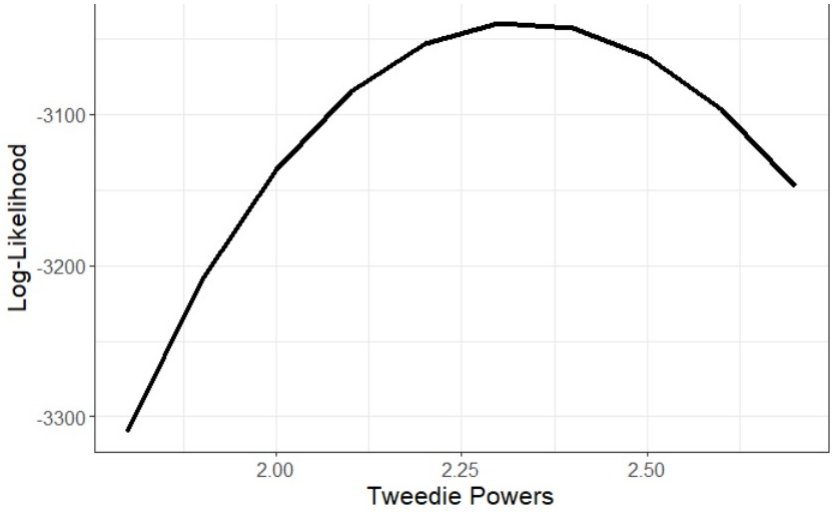

FIGURE 3.8 – Tracé du profil log-vraissemblance pour AutoBi.

Nous comparons les deux modèles. Le critère d'information Akaike pour le modèle de Tweedie évalue à 4611,25 pour le modèle gamma. Ces résultat indiquent que le modèle Tweedie s'adapte mieux aux données que le modèle gmma. Nous soulignons que l'AIC parle de la qualité d'un modèle par rapport à un autre modèle. ll est possible que les deux modèles correspondent mal aux données .

Sur la base métrique AIC, nous obtenons un meilleur modèle lorsque nous obtenons un meilleur modèle lorsque nous estimons la puissance de Tweedie .Cependant,nous voulons quantifier cette amélioration. Le tableau 3.5 présente les pondérations des deux modèles. Observez que les coefficients changent avec un ordre de grandeur 10-2. Le tableau 3.6 enregistre les prédactions pour les5 premiers cas de l'ensemble d'apprentissage .

Nous voyons des différences dans les dizaines et les centaines de dollars .

TABLE 3.5 – comparaison des poids linéaires entre les deux modèles de gravité AutoBi

variable explicative	Modèle Tweedie	Gamma Modèle
Intercepet	1.8699	1.8382
Avocat	-1.4757	-1.4772
Centure de sécurité 1	0.4490	0.4887
Centure de sécurité 2	1.3231	1.4061
Age 1	-1.4032	-1.3988
Age 2	-0.6246	-0.6544
Age 3	-0.1750	-0.1814
Age 4	-0.0666	-0.0669
Age 5	-0.6338	-0.6296

TABLE 3.6 – Prédictions de perte pour les 5 premières lignes des données de test $AoutoBi$.

Idex des assurés	Perte réelle	Modèle Tweedie	Gamma Modèle
7	3538	5443.06	5325.41
9	874	1232.93	1257.50
22	230	1955.63	1951.10
23	26262	8554.29	8546.74
33	603	1244.36	1215.72

L'ensemble de formation représente environ 5,1 millions de dollars de pertes. Nous calculons que le modèle de Tweedie prédit environ 100000 dollars de moins en pertes que le modèle gamma. Je attribue cette déffirence à la façon dont le modèle de Tweedie se situe entre les cas gamma et gaussiens inverse. Les modèles gaussiens inverses ont des pics plus petits et des queues plus larges que les modèles gamm. Comme la plupart des réclamations résident dans la zone de pointe, je soupçonne que le plus petit pic du modèle Tweedie explique la réduction. Ces informations suggèrent que l'estimation de la puissance de Tweedie pourrait améliorer la précision d'envoin 2 pourcent. Nous recevons des résultats moins encourageants lorsque nous examinons l'ensemble de tests. Ces pertes représentent environ 2,8 millions de dollars de pertes. Nous prédisons les pertes à partir des deux modèles et observons que la diffirénce des pertes prévues ne varie que d'environ deux mille dollars. Ce calcul indique que le modèle entraîne ne se traduit pas également par les données de test. De plus les deux modèles ne parviennent pas à prédire un montant énorme de pertes totales. Le ratio des pertes globales réelles est d'environ 57 pourcent. Si nous avions accès à plus de variables explicatives, je pense que nous pourrions amélorer mes modèles et résoudre ce problème. Ensuite, nous traçons les prédictions du modèle et la distribution empirique de la cible. Notamment, on voit de multiples bosses et des queues minces sur les figures3.9 et3.10, alors que la distributionréelle a une bosse et une queue plus épaisse. Aucun des deux modèles ne prèdit très bien les pertes importantes. De plus,

nous voyons peu de différence entre les figure 3.9 et 3.10. Les formes des deux distributions semblent presque identiques. Cette observation remet en question la différence entre deux modèles Tweedie lorsqu'ils appartiennent tous deux à la même sous-classe.

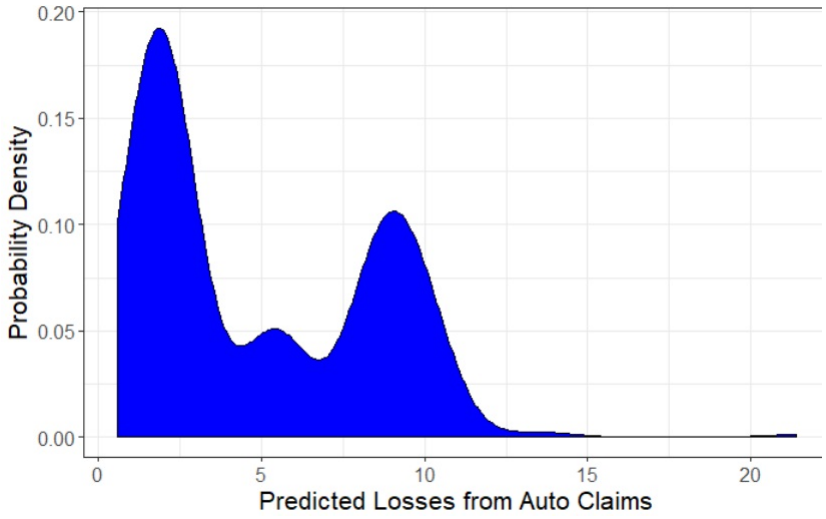

FIGURE 3.9 – Distribution $T_{W_{2,3}}$ du modèle de sévérité.

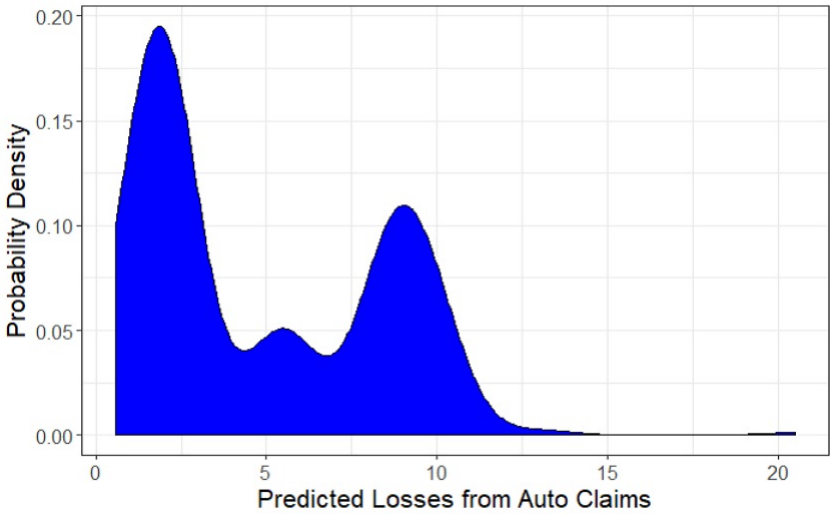

FIGURE 3.10 – Répartition de la gravité des gamma.

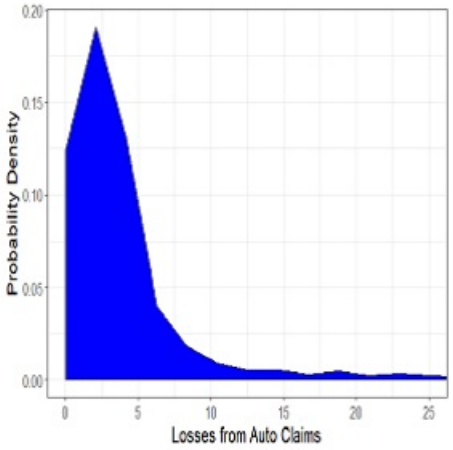

FIGURE 3.11 – Répartition des pertes AutoBi.

Cette expériences a considéré la modélisation de la gravité avec une puissance Tweedie $p \in (2,3)$. Les actuaires rencontrent normalement le nom Tweedie en termes de variables aléaitoires composées de Poisson-Gamma et de modélisation des coût de perte. En ce sens, cette étude de cas apporte une certaine nouveauté. Nous avons créé un modèle Tweedie pour les allégations de gravité et nous avons fait valoir ses avantages par rapport à un modèles gamma.

Malheureusement, notre réglage des paramètres de la puissance Tweedie n'a guère changé les prédictions du modéle.

TABLE 3.7 – Distribution des pertes

Min	1st Qu	Median	Mean	3rd Qu	Max
0.005	0.640	2.331	5.954	3.995	1067.697

TABLE 3.8 – Tableau de Modèle linéaire générale de la famile de Tweedie pour $p = 2$

	Estimate	Std. Error	t value	Pr(T>Z)
intercept	1.871e+00	6.688e-01	2.797	0.00525
CASENUM	-1.131e-05	5.918e-06	-1.911	0.05632
ATTORNEY	-1.438e+00	1.199e-01	-12.001	<2e-16
CLMSEX	5.674e-03	1.203e-01	0.047	0.96238
MARITAL	-9.550e-02	9.464e-02	-1.009	0.31315
CLMINSUR	1.718e-01	2.026e-01	0.848	0.39667
SEATBELT	1.109e+00	4.430e-01	2.504	0.01243 *
CLMAGE	1.522e-02	3.588e-03	4.242	2.4e-05 ***

TABLE 3.9 – Tableau de Modèle linéaire générale de la famile de Tweedie pour $p = 1.5$

	Estimate	Std. Error	t value	Pr(T>Z)
intercept	1.806e+00	6.455e-01	2.798	0.005237
CASENUM	-1.262e-05	6.416e-06	-1.968	0.049371
ATTORNEY	-1.465e+00	1.407e-01	-10.409	< 2e-16 ***
CLMSEX	-2.907e-02	1.293e-01	-0.225	0.822105
MARITAL	-1.045e-01	1.016e-01	-1.029	0.303934
CLMINSUR	2.777e-01	2.238e-01	1.241	0.215024
SEATBELT	1.174e+00	3.407e-01	3.444	0.000594 ***
CLMAGE	1.295e-02	3.832e-03	3.379	0.000754 ***

TABLE 3.10 – Tableau de Modèle linéaire générale de la famile de Tweedie pour $p = 0.5$

	Estimate	Std. Error	t value	Pr(T>Z)
intercept	1.806e+00	6.455e-01	2.798	0.005237 **
CASENUM	-1.262e-05	6.416e-06	-1.968	0.049371 *
ATTORNEY	-1.465e+00	1.407e-01	-10.409	< 2e-16 ***
CLMSEX	-2.907e-02	1.293e-01	-0.225	0.822105
MARITAL	-1.045e-01	1.016e-01	-1.029	0.303934
CLMINSUR	2.777e-01	2.238e-01	1.241	0.215024
SEATBELT	1.174e+00	3.407e-01	3.444	0.000594 ***
CLMAGE	1.295e-02	3.832e-03	3.379	0.000754 ***

TABLE 3.11 – Tableau de Modèle linéaire générale pour la famille de poison

Null deviance	13771 on 1090 degrees of freedom
Residual deviance	10460 on 1083 degrees of freedom
AIC	Inf

TABLE 3.12 – Tableau de Modèle linéaire générale pour la famille Binomiale négative

Null deviance	13771 on 1090 degrees of freedom
Residual deviance	10460 on 1083 degrees of freedom
AIC	5501.1

CONCLUSION

En probabilité et en statistiques, les distributions de Tweedie appartiennent à la classe des modèles de dispersion exponentielle, célèbres pour leur rôle dans les modèles linéaire généralisé. C'est une famille de distribution de probabilité qui comprend des distributions continue telles que la distribution Normale et Gamma, la distribution de poisson exclusivement discrète, et la classe de distributions composées mixtes poisson-Gamma qui ont une quantité importante zéro. Les distributions de Tweedie sont trés connues et utiles dans plusieurs domaines de recherche tels que l'analyse de survie, les études dépenses et de consommation, de l'écologie et de la météorologie. La distribution de Tweedie est particulièrement utilisée dans la recherche actuarielle, plus précisément dans la modélisation des réclamations d'assurance.

La modélisation des réclamations d'assurance automobile, montre que le modèle s'ajuste bien aux données observées, et par conséquent, la distribution de Tweedie est suggérée pour modéliser les coûts individuels d'assurance.

Enfin, la densité de la distribution de Tweedie nous a permis de faire une application en assurance, plus précisément de calculer la transformée d'Esscher, qui rend le risque moins attrayant pour l'assureur, et de calculer la prime d'assurance en fonction du paramètre d'Esscher h, qui reflète le degré d'aversion du risque pour l'assureur.

BIBLIOGRAPHIE

[Alicja et Michal, (2014)] Alicja Wolny-Dominiak and Michal Trzesiok. insuranceData : A Collection of Insurance Datasets Useful in Risk Classification in Non-life Insurance. R package version 1.0. 2014. URL : https ://CRAN.R-project.org / package=insuranceData.

[Bent Jorgensen, (1997)] Bent Jorgensen. The Theory of Dispersion Models. Chapman and Hall, 1997.

[Beginning, (2018)] Beginning Statistics : Continuous Random Variables. https//2012books.lardbucket.org/books/beginning–statistics/s09-continuousrandom-variables.html. Accessed : 2018-05-03.

[Burnham et Anderso, (2004)] K.P. Burnham et D. R. Anderso, ≪Multimode inference : understanding AIC and BIC in Model Selection≫, Sociological Methodes and Research, 2004, p. 216-304.Bühlmann, H. (1980). An economie premium principe. Astin Bulletin, 11(11), 52-60.

[Dunn, (2004)] Dunn, P. K. (2004) . Occurrence and quantity of precipitation can be modelled simultaneously. International Journal of Climatology, 24(10) , 1231-1239.

[Dunn et Smyth, (2008)] Dunn, P. K. et Smyth, G. K. (2008). Evaluation of tweedie exponential dispersion madel densities by fourier inversion. Statistics and Computing, 18(1) , 73- 86.

[Dunn et Smyth, (1996)] Dunn, P. K. et Smyth, G. K. (1996). Randomize quantile residuals. Journal of computational and Graphical Statistics, 5(3), 236-244.

[Dunn, (2013)] Dunn, Peter K, M. P. K. (2013). Package 'tweedie'. R package version, 2(7).

[Doucet, (2014)] Doucet, E. (2014). Estimateurs à noyeau et théorie des valeurs extrêmes : comparaison de leur pou voir prédictif dans l'analyse du coûts des réclamation en assurance automobile.

[Dickson, (2005)] Dickson, D. C. (2005). Insurance risk and ruin. Cambridge university Press.

[Denuit et all., (2006)] Denuit, M., Dhaene, J., Goovaerts, M. et Kass, R. (2006). Actuarial theory for dependent risks : measures, orders and models. John Wiley & sous.

[Gerber, (1980)] Gerber, H. U. (1980). A characterization of certain families of distributions via esscher transfors and independence. Journal of the American Statistical Association, 75(372) , 1015- 1018.

[Gerber et al., (1984)] Gerder , H. U., Shiu , E. s. et al. (1994). Option pricing by esscher transforms, Transations of the Society of Actuaries, 46(99), 140.

[Goovaerts et all., (1984)] Goovaerts, M. De Vylder, F . et Haezendonck, J. (1984). Insurance premiums .

[Keller, (2008)] Keller, A. (2008). Applying Robust S cale M-Estimators to Compute Credibility Premiums in the Large Claim Case. Logos Verlag Berlin GmbH.

[Krzanowski, (1998)] Krzanowski, W.J. (1998). An Introduction to Statistical Modelling. Arnold, a meber of the Hodder Headli Group.

[Kass et al., (2008)] Kass, R Goovaerts, M., Dhaaene, J, et Denuit , M. (2008). Moder actuarit risk theory : using R, volume 128. Springer Science & Business. Media.

[Hassan et Dunn, (2010)] Hassan, M. M. et Dunn, P. k. (2010). A simplepoisson-gamma for modelling rainfall occurrence and amout simultaneously. Agricultural and forest meteorology, 150(10), 1319-1330.

[Hassan et Dunn, (2015)] Hassan, M. M. et Dunn, P. k. (2015) Seasonal rainfall totals of australian stationcan be modelled with distributions from the tweedie family. International Journal of Climatology, 35(10), 3093-3101.

[Hassan et Dunn, (2011)] Hasan, M. M. et Dunn, P. K. (2011). Two tweedie distributions that are nearoptimal for modelling monthly rainfall in australia. International Journal of Climatology, 31(9), 1389- 1397.

[Hiroshi, (2008)] Hiroshi Shono. "Application of the Tweedie Distribution to Zero-catch Data in CPUEAnalysis". In : Fisheries Research 93.1 (2008), pp. 154–162.

[Hirotugu Akaike, (1973)] Hirotugu Akaike, ≪ Information theory and an extension of the maximum likelihood principe ≫, dans Second International Symposium on Information Theory, 1973, 267.-281 .

[James, (2013)] James Owen Weatherall. The Physics of Wall Street : A Brief History of Predictingthe Unpredictable. Houghton Mifflin.

[John et Wedderburn, (1972)] John Nelder rt Robet Wedderburn, ≪Generalized Linear Models≫, Journal of Rayal Statistical Society. Series A (General), Blackawell Publishing, vol.135, n°3, 1972, p. 370-384.

[Md Masud et peter K,(2011)] Md Masud Hasan and Peter K Dunn. “Two Tweedie distributions that are Nearoptimal for Modelling Monthly Rainfall in Australia”. In : International Journal ofClimatology 31.9 (2011), pp. 1389–1397.

[Roos, (2009)] Sheldon M Ross. A First Course in Probability. Pearson Education International, 2009.

[Peter et Smyth, (2005)] Peter K Dunn and Gordon K Smyth. "Series Evaluation of Tweedie Exponential Dispersion Model Densitiees". In : Statistics and Computing 15.4 (2005), pp.267-280.

[Peterset Shevchenko, (2015)] Peters, G.W. et Shevchenko, P.V .(2015). Advances un Heavy Tailed Risk Modeling : A Handbook of Operational Risk. John Wiley.

[Peter et Smyth, (2001)] Peter K Dunn and Gordon K Smyth. “Tweedie family densities : methods of evaluation”.In : Proceedings of the 16 th International Workshop on Statistical Modelling, Odense, Denmark. 2001, pp. 2–6.

[Tweedie, (1984)] MCK Tweedie. “An Index which Distinguishes between Some Important Exponential Families”. In : Statistics : Applications and new directions : Proc. Indian statistical institute golden Jubilee International conference. Vol. 579. 1984, p. 604..

[Tweedie, (1947)] MCK Tweedie. “Functions of a Statistical Variate with Given Means, with SpecialReference to Laplacian Distributions”. In : Mathematical Proceedings of the Cambridge Philosophical Society. Vol. 43. 1. Cambridge University Press. 1947, pp. 41–49.

[Teugels, (2004)] Teugels, B. S. (2004). Encyclopedia of Actuariat Science. Weliy.