République Algérienne Démocratique et Populaire
Ministère de l'enseignement supérieure et de la recherche scientifique

N° Attribué par la bibliothèque

Année univ.: 2017/2018

Introduction to Parametric and Nonparametric Families

Mémoire présenté en vue de l'obtention du diplôme de

Master Académique

Université de Saida - Dr Moulay Tahar
Discipline : MATHEMATIQUES
Spécialité : Analyse Stochastique, Statistique des Processus et Applications (ASSPA)

par

Wafaa Zaoui [1]

Sous la direction de

Dr. Fethi MADANI

Soutenue le 24/06/2018 devant le jury composé de

Mr. M. LAOUNI	Université Dr Tahar Moulay - Saïda	Président
Dr. F. MADANI	Université Dr Tahar Moulay - Saïda	Encadreur
Dr. F. MOKHTARI	Université Dr Tahar Moulay - Saïda	Examinatrice
Dr. N. HACHEMI	Université Dr Tahar Moulay - Saïda	Examinatrice

1. e-mail : zaouiwafaa.zw@gmail.com

Remerciements

En tout premier lieu, je remercie le bon **Dieu**, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés.

Le travail présenté dans ce mémoire a été réalisé au laboratoire de l'Analyse des Modèles Stochastiques, statistique et Application de la faculté des sciences de l'Université de Saida, sous la direction du Monsieur Fethi Madani, ma plus grande gratitude va à mon encadreur, pour sa disponibilité et la confiance qu'elle m'a accordée. J'ai profité pendant longtemps du savoir et du savoir-faire dont j'ai pu bénéficier au cours de nombreuses discussions. J'aimerais aussi la remercier pour l'autonomie qu'elle m'a accordée, et ses précieux conseils qui m'ont permis de mener à bien ce travail.

Je remercie **Mr. Laouni.** d'avoir accepté à présider le jury de mon mémoire.

Je souhaite également remercier **Dr. F. Mokhtari** et **Mademoiselle. N. Hachemi** . de m'avoir fait l'honneur d'accepter d'être examinatrices de mon mémoire. Leurs remarques et précieux conseils ont abouti à une amélioration du document de mémoire.

Je remercie le Professeur **A.Kandouci** et tout les membres du laboratoire **LMSSA** pour m'avoir accueilli au sein du labo. Je tiens également à leur remercier pour leur disponibilité, la confiance qu'ils m'ont accordé et pour leurs conseils et leurs commentaires fort utiles qui ont fortement enrichi ma formation.

Ma gratitude va également aux enseignants du Département de Mathématiques que j'ai eu durant mon cursus.
J'adresse mes sincères remerciements à tous les professeurs qui m'ont enseigné, et tous les professeurs de notre laboratoire **LMSSA**.
je remercie toute ma famille pour leur soutien qu'ils m'ont apporté pendant ces années d'études.

Afin de n'oublier personne, mes vifs remerciements s'adressent à tous ceux qui m'ont aidée à la réalisation de ce modeste mémoire.

Dédicaces

Je dédie ce modeste travail en signe de respect, reconnaissance et de remerciement :

A mes chers parents, ma familles et à tous ce qui me sont proches.

Table des matières

Notation

IHR : Increasing Hazard Rate (Taux de Risque Croissant).
DHR : Decreasing Hazard Rate (Taux de Risque Décroissant).
IHRA : Increasing Hazard Rate Average (La moyenne de Taux de Risque Croissant).
DHRA : Decreasing Hazard Rate Average (La moyenne de Taux de Risque Décroissant).

Résumé

Pour décrire mathématiquement la distribution d'une variable aléatoire, plusieurs fonctions sont utilisées. Ces fonctions comprennent des fonctions de distribution, fonctions de survie, des densités et la fonction de hazard,....etc. Quand ils existent. L'un de ces fonctions peuvent être obtenues, au moins théoriquement. Mais il y a de bonnes raisons pour s'intéresser à toutes ces fonctions sachant qu'aucune n'est uniformément meilleur que l'autre.
Parfois, l'une a une forme particulièrement simple alors que d'autres ont des formes complexes. Il est important de dire que certains aspects d'une distribution sont révélés plus clairement par une autre fonction. Différentes personnes peuvent avoir différentes préférences, en fonction de l'intuition qu'ils ont développée. Aussi, certaines de ces fonctions peuvent être plus faciles à estimer que d'autres.

Pour l'analyse des données, plusieurs approches statistiques sont utilisées, à savoir, Les familles paramétriques, les familles non-paramétrique et les familles semi-paramétrique.

Chapitre 1

Introduction générale

Les variables aléatoires non négatives apparaissent dans une grande variété d'applications. Les longueurs de vie de dispositifs artificiels ou d'organismes biologiques sont respectivement l'accent de la fiabilité et de l'analyse de survie. Mais d'autres types de les temps d'attente se produisent également dans les applications ; ceux-ci peuvent être des temps d'attente pour les retards dans la circulation, les intervalles entre les séismes ou les inondations ou les périodes nécessaire pour apprendre une tâche. Variables aléatoires non-négatives apparaissent comme des grandeurs liées à des objets physiques. Longueurs de fissures, diamètres ou hauteurs des arbres, les vitesses du vent, les forces des matériaux, les débits des cours d'eau, les précipitations, l'usure des pneus ou les produits chimiques. L'économie est un autre domaine d'applications où des variables aléatoires non-négatives apparaissent ; le revenu, la taille de l'entreprise, les prix et l'actuariat En revanche, la distribution normale, qui a longtemps joué un rôle central dans les statistiques, permet aux variables aléatoires correspondantes de prendre sur toutes les valeurs réelles, à la fois négatives et positives. C'est le cas pour erreurs de mesure, le contexte dans lequel la distribution normale d'abord s'est s'élever. Pour les variables aléatoires non négatives avec des écarts-types pertes sont par nature non-négatives, et elles sont petites par rapport à leurs moyens, la distribution normale a été largement utilisé et peut souvent fournir d'excellentes approximations. En d'autre cas, la distribution normale peut être un modèle inapproprié, et les alternatives doivent être considérées.

Pour les variables aléatoires non négatives, il n'y a pas de distribution aussi présente comme la distribution normale, avec sa fondation dans le théorème centrale limite. Cela signifie qu'une grande variété de distributions partagent une importance relative. Le but de ce sujet est d'étudier les origines et propriétés des différentes distributions pour une variable aléatoire non-négative.

Une motivation statistique

Pour l'analyse des données, plusieurs approches statistiques sont communes utilisation. Ces approches forment la hiérarchie suivante :

Les méthodes (non paramétriques)

Méthodes statistiques qui ne dépendent pas d'hypothèses sur le sous-jacent la distribution sont attrayants parce qu'il n'y a pas d'hypothèses à question. Bien qu'attrayant de ce point de vue, les conclusions peut être plus faible que ce qui pourrait être possible avec certains possible hypothèses.

Les méthodes conditionnellement qualifiées

Familiarité avec les origines de les données peuvent rendre les Hypothèses qualitatives raisonnables. Par exemple, un praticien peut savoir que les données proviennent d'une distribution avec une densité décroissante, une densité unimodale, ou que la densité a un queue lourde de la main droite. Ils peuvent suspecter des considérations physiques que le taux de risque est monotone ou qu'il décroissant initialement et croissant éventuellement. Un certain nombre de procédures statistiques sont connues pour tester la validité de telles hypothèses. D'autres sont basés sur de tels hypothèses, et ils peuvent conduire à des idées difficiles à obtenir avec méthodes sans connaitres distribution.

Les méthodes semi-paramétriques

Il y a un certain nombre d'hybride possible méthodes qui impliquent une Distribution non spécifiée et un paramètre paramétrique spécifique structure. Peut-être le plus connu est-il le modèle proportionnel modèle des risques, dans lequel la fonction de survie de l'espèce non spécifiée la distribution est souevée à un pouvoir positif ; ce pouvoir est alors un paramètre. Des modèles comme celui-ci sont appelés modèles semi-paramétriques.

Les méthodes paramétriques

Enfin, On peut être prêt à supposer que les données proviennent d'une famille paramétrique spécifiée. C'est peut-être l'approche la plus connue des problèmes statistiques, et bien sûr, l'hypothèse de normalité est la plus familière.

Etat de l'arts

Sauf pour les méthodes sans distribution, toutes les approches statistiques décrit ci-dessus dépendent de ce qu'on appelle parfois "Modèles". Parce que l'hypothèse d'un modèle inapproprié peut conduire à des conclusions erronées, pourquoi les modèles sont-ils jamais utilisés ?
Dans certaines applications, l'hypothèse d'un modèle implique peu ou aucun risque. Par exemple, le théorème cenrale limite peut faire le normal la distribution d'un choix clair.
Les modèles paramétriques ont été introduits au début des études de la vie humaine . Dans ce contexte, les tables de mortalité sont fondamentales. Idéalement, une table de mortalité commence par un groupe fixe d'individus, tous né en même temps, et enregistre le nombre de personnes vivant à la fin de chaque année successive jusqu'à ce que tous sont décédés. Une telle table représente un dossier empirique avec des données regroupées par années.
On a souvent supposé que les décès d'une année se produisaient uniformément pendant l'année ; Aux fins de l'obtention du diplôme, De Moivre (1724) [33] a fait l'hypothèse d'uniformité sur des périodes encore plus longues, bien qu'il reconnu que cette approximation n'est pas strictement vraie. Encore pour le fins de l'obtention du diplôme, Gompertz (1825) [18] introduit le modèle paramétrique maintenant connu sous le nom de la distribution Gompertz comme une approximation à la "vraie loi de la mortalité." Gompertz a travaillé en collaboration avec les données, puis développé une base théorique pour sa distribution, mais à obtenir un bon ajustement aux données, il a jugé nécessaire de diviser les âges en trois groupes, en utilisant différentes valeurs

de paramètres dans chaque groupe d'âge.
Dans un certain nombre d'autres applications, les modèles paramétriques ont joué un rôle historique important dans la théorie statistique, et ont longtemps été étudiés sous la rubrique plus ancienne de "ajustement de la courbe." Certains des plus tôt connus travail est celui de Pearson (1895) [37], qui a créé un ensemble de "courbes" ou "fréquence distributions "qui pourraient convenir à l'adaptation aux données dans une variété de contextes. D'autres ensembles de distributions ont été construits par Charlier [9], Thiele [48] et d'autres. Elderton Johnson (1969, p.2) [12] indiquent que "les avantages de tout système des courbes dépendent de la simplicité des formules et du nombre des classes d'observations qui peuvent être traitées de manière satisfaisante,. . . " Voir aussi Elderton (1906, 1934) [10], Elderton et Johnson (1969) [12], Sarndal (1971) [44], Cramér (1972) [8]. Ainsi, il y avait un accent sur la richesse à la fois d'applications et de simplicité mathématique.

1.1 Descriptions probabilistique

Pour décrire mathématiquement la distribution d'une variable aléatoire, divers les fonctions alternatives sont couramment utilisées. Ces fonctions comprennent fonctions de distribution, fonctions de survie, densités, taux de risque (Taux de hazard), moyenne vies résiduelles, et temps total sur les transformations de test....etc. Quand ils existent, l'un des ces fonctions peuvent être obtenues, au moins théoriquement, à partir de n'importe quelle autre.
Mais il y a de bonnes raisons de s'intéresser à toutes ces fonctions ; aucun n'est pas uniformément le meilleur. Parfois, on a une forme particulièrement simple alors que d'autres sont difficiles à travailler.

La fonction de distribution et la fonctions de survie

Définition 1.1.0.1. *La fonction F définie sur l'intervalle* $(-\infty, \infty)$ *par :*

$$F(x) = P\{X \leq x\}$$

est appelée la fonction de distribution de X.
Les fonctions de distribution sont parfois appelées "Distribution cumulative".

Définition 1.1.0.2. *La fonction $\bar{F}$ définie sur l'intervalle $(-\infty, \infty)$ par :*

$$\bar{F}(x) = P\{X > x\}$$

est appelée la fonction de survie de X.

Bien sûr, $\bar{F} = 1 - F$ et il peut donc sembler superflu d'introduire la fonction de survie. Mais c'est souvent le cas que pour non négatif variables aléatoires, la fonction de survie est plus significative et prend une forme plus pratique que la distribution mieux connue une fonction.
La fonction de survie est parfois appelée la "fonction de survivant" ou la "fiabilité". Diverses notations pour cette fonction ont été utilisées dans la littérature ; la notation "bar" a été introduite par Frank Proschan et a été utilisé par Barlow et Proschan (1963) [3].

La fonction de masse de probabilité et la fonctions de densité

Pour toute variable aléatoire, la fonction de distribution et la fonction de survie toujours exister. Cet avantage n'est pas apprécié par la masse de probabilité fonctions ou fonctions de densité, mais ces fonctions ont parfois d'autres avantages.
Supposons d'abord que la variable aléatoire X ne peut pas prendre qu'une valeur finie ou nombre dénombrable de valeurs. Par exemple, X peut être le nombre de essais nécessaires pour obtenir des "têtes" dans les lancers répétés d'une pièce de monnaie. Puis X (et F) est dit être discret. Les fonctions de distribution discrète sont l'étape des fonctions.

Définition 1.1.0.3. *Si $x_1, x_2, x_3, \ldots$ est l'ensemble des valeurs possibles de X et $P(x_i) = P\{X = x_i\}$, i=1,2,... , puis*

$$F(x) = \sum_{x_i \leq x} P(x_i)$$

tel que P est appelée la fonction de masse de probabilité de X.

Lorsque X prend toutes les valeurs dans un intervalle (éventuellement infini) de la ligne réelle, il est souvent possible d'écrire F comme une intégrale.

Définition 1.1.0.4. *Si f est une fonction non-négative pour laquelle*

$$F(x) = \int_{-\infty}^{x} f(z)dz, \quad \textit{pour tout rel } x$$

alors f est appelée une densité de probabilité de X (ou F).

Quand une densité existe, X (ou F) est dit absolument continu. Lorsque les densités existent, elles ne sont pas uniques car elles peuvent être modifiées arbitrairement à des points isolés sans changer l'intégrale.

Définition 1.1.0.5. *On dit qu'une distribution F est concentrée sur intervalle fermé [a, b] si F(x)=0 pour tout x<a, et F(x)=1 pour tout x> b. Le support de la distribution F est l'ensemble de tous les points x tels que $F(x+\epsilon) - F(x-\epsilon) > 0$, pour tout $\epsilon > 0$.*

On peut montrer que le support d'une distribution est un ensemble fermé, et dans la plupart des exemples discutés, c'est un intervalle. Si F est concentré sur l'intervalle [a, b], alors le support de F est un sous-ensemble de cet intervalle. Quand F est absolument continu et concentré sur l'intervalle fermé [a,b], alors il y a une version naturelle f de la correspondante densité qui satisfait f(x)=0, pour tout $x \in [a, b]$. Parfois, le support de F est défini comme la fermeture de l'ensemble de tous les points x tels que f(x)> 0, mais cela suppose qu'une version appropriée du la densité a été choisie.

Unimodalité

L'idée de l'unimodalité est mieux comprise en termes de densité, mais est peut-être le mieux défini en termes de fonction de distribution.

Définition 1.1.0.6. *Une fonction de distribution F est dite unimodale avec mode à m si F(x) est convexe dans x<m et concave dans x>m.*
Lorsque F est unimodal et a une densité continue f, alors f(x) croissant en x<m et décroissant en x> m pour que f(m) soit un maximum de f.

La fonctions de hazard et le taux de hazard

Quelques aspects d'une distribution absolument continue, importante ou même essentiel dans certains contextes, peut être vu plus clairement de taux de risque que de la fonction de distribution ou de la fonction de densité.

Définition 1.1.0.7. *La fonction R définie sur $(-\infty, \infty)$ par :*

$$R(x) = -\log \bar{F}(x) \tag{1.1}$$

est appelée la fonction de hazard de F, ou de X.

Pour une variable aléatoire non négative, R(0-)=0, R croissant, et $\lim_{x\to\infty} R(x) = \infty$; Les fonctions de hazard sont parfois utile en mathématiques applliquées , mais contrairement aux taux de risque, ils ont peu d'intérêt.

Définition 1.1.0.8. *Si F est une fonction de distribution absolument continue avec la densité f, alors la fonction r définie sur $(-\infty, \infty)$ par :*

$$\begin{aligned} r(x) &= \frac{f(x)}{\bar{F}(x)}, \quad si \quad \bar{F}(x) > 0 \\ &= \infty, \quad si \quad \bar{F}(x) = 0 \end{aligned} \tag{1.2}$$

est appelé un taux de risque de F, ou de X.

Pour x tel que $\bar{F}(x) = 0$, le rapport $f(x)/\bar{F}(x)$ est indéterminé ; la valeur ∞ est arbitraire mais parfois commode. Parce que les densités sont pas unique.
Bien que les taux de risque ne soient pas uniques, le plus souvent est fait pour "le taux de hazard." Lorsque F est absolument continu et la fonction de hazard R est dérivable, alors sa dérivée est un taux de hazard.
Pour mieux comprendre les taux de risque, il est utile de noter que ;

$$r(x) = \lim_{\Delta \downarrow 0} \frac{P\{x < X \leq x + \Delta | X > x\}}{\Delta}$$

Ainsi,

$$\Delta r(x) = P\{x < X \leq x + \Delta | X > x\}$$

Par conséquent, $\Delta r(x)$ peut être considéré comme la probabilité conditionnelle de la fonction de survie à temps x, de la mort ou de l'échec dans la petite croissante suivante Δ de temps. De (1.1) et (1.2), on peut voir que si F(0)=0, alors

$$\bar{F}(x) = \exp\{-R(x)\} = \exp\{-\int_0^x r(z)dz\}. \tag{1.3}$$

Cette formule clé montre comment récupérer la fonction de survie à partir du taux de risque Mais la seconde égalité de (1.3) n'est pas valable que si F est absolument continu.
Lorsque F est absolument continu, on peut le voir à partir de la seconde l'égalité de (1.3)

$$r(x) = \frac{dR(x)}{dx}.$$

De (1.2) et (1.3) il s'ensuit qu'une fonction r est le taux de hazard de certaine distribution sur $(0, \infty)$ si et seulement si ;
(i) $r(x) \geq 0$, pour tous $x > 0$.
(ii) $\int_0^x r(t)dt < \infty$, pour certains $x > 0$,
(iii) $\int_0^\infty r(t)dt = \infty$,
(iv) $\int_0^x r(t)dt = \infty$ implique $r(z) = \infty$, pour tous $z > x$.
La condition (ii) nécessite quelques explications. Si la distribution F correspondante à r vérifie $F(x) < 1$ pour tout x, alors l'intégrale de (ii) est fini pour tout $x < \infty$. Mais s'il y a un nombre $a < \infty$ tel que F(x)<1 pour x<a et F(a)=1, alors l'intégrale (ii) n'est finie que pour x<a. La condition (iv) résulte de la façon dont le taux de risque est défini pour x>a.

Proposition 1.1.0.1. *Si F est concentré sur [0,a] et posséde un risque notez r, alors* $\limsup_{x \uparrow a} r(x) = \infty$. *En effet ; Ceci est une conséquence de (1.3) qui montre que r est intégrable sur [0, x] pour x<a, mais l'intégrale doit diverger sur [0, a].*

Remarque 1.1.1. *Si le taux de risque r décroissant à* $x = x_0$*, alors la densité f correspondante est également décroissante à* x_0*. Cela suit directement à*

partir de (1.2).
Il est montré que si $Z = \min[X, Y]$*, alors le taux de risque de Z est la somme des taux de hazard de X et Y. Ce mécanisme produit une variété de formes de taux de risque.*

La fonctions de hazard inversé et le taux de hazard inversé

La fonction de risque inverse est définie d'une manière similaire au fonction de risque $R(x) = \log \bar{F}(x)$, mais avec la fonction de distribution F remplacer la fonction de survie $\bar{F}$. En autre, le signe moins est omis pour le rendre, comme la fonction de risque, est croissant.

Définition 1.1.0.9. *La fonction S définie sur* $(-\infty, \infty)$ *par :*

$$S(x) = \log F(x), \tag{1.4}$$

est appelée la fonction de risque inverse de F, ou de X. Si F est absolument fonction de distribution continue avec densité f, puis une fonction s défini sur $(-\infty, \infty)$ *par :*

$$s(x) = f(x)/F(x), \tag{1.5}$$

est appelé un taux de risque inverse de F, ou de X.

Notez que $S(x) = \log F(x)$, alors que $R(x) = -\log \bar{F}(x)$; avec ces définitions, les deux fonction de risque inversé et le taux de risque inversé de F n'ont pas joué un rôle important dans la littérature. Le taux de risque inverse était introduit par von Mises (1936) [32] et a été discuté brièvement par Barlow, Marshall et Proschan (1963)[41], qui notent que s(-x) est le taux de hasard de -X, donc la terminologie ici. Plus récemment, il a été discuté en détail par Block, Savits et Singh (1998)[4] ; voir aussi Shaked et Shanthikumar (1994, page 24)[45]. sont croissant .
De (1.4) et (1.5), on peut voir que si F(0)=0,

$$F(x) = \exp\{S(x)\} = \exp\{\int_0^x s(z)dz\}. \tag{1.6}$$

Le taux de risque inverse a été largement ignoré dans la littérature principalement parce qu'il n'a pas le fort contenu intuitif le taux de risque.

Proposition 1.1.0.2. *Le taux de risque r et le taux de risque inversé s ont les propriétés de monotonie suivantes :*

$$\begin{aligned} s \quad est \quad croissant \Rightarrow & \quad r \quad est \quad croissant \\ r \quad est \quad decroissant \Rightarrow & \quad s \quad est \quad decroissant \end{aligned}$$

Ces résultats sont faibles et les implications ne s'inversent pas. Ils peuvent être facilement obtenu à partir de la relation $r(x) = s(x)[F(x)/\bar{F}(x)]$.

La distribution de la survie résiduelle

La répartition de la durée de survie restante pour un élément de temps t non atteint est souvent d'intérêt et joue un rôle récurrent dans ce qui suit.

Définition 1.1.0.10. *Soit F une fonction de distribution telle que F(0)=0. La distribution de vie résiduelle* F_t *de F et t est définie pour tous t=0 tel que* $\bar{F}(t) > 0$ *par :*

$$\bar{F}_t(x) = \frac{\bar{F}(x+t)}{\bar{F}(t)}, \quad x \geq 0 \tag{1.7}$$

Si F a une densité f, alors F_t *a une densité* f_t *et un taux de risque* r_t *donné par :*

$$f_t(x) = \frac{f(x+t)}{\bar{F}(t)}, \quad x \geq 0 \tag{1.8}$$

$$\begin{aligned} r_t(x) &= \frac{f(x+t)}{\bar{F}(x+t)} \\ &= r(x+t), \quad x \geq 0 \end{aligned} \tag{1.9}$$

Clairement, la distribution de survie résiduelle F_t est une distribution conditionnelle de la survie restante donné la survie à temps t. Cette distribution est d'un intérêt pratique considérable, car la durée de survie restante des appareils (voitures d'occasion, etc.) ou d'entités biologiques (les gens, par exemple) est souvent d'intérêt.

Proposition 1.1.0.3. *Si le taux de risque r de F a une limite positive finie,* $\lim_{t\to\infty} r(t) = \lambda$, *alors* F_t *converge dans la distribution vers une distribution exponentielle avec le paramètre* λ *comme* $t \to \infty$. *En effet : De (1.3), il s'ensuit que*

$$-\log \bar{F}_t(x) = -\log \bar{F}(x+t) + \log \bar{F}(t) = \int_t^{t+x} r(z)dz \to \lambda x \quad quandt \to \infty.$$

La fonction de durée résiduelle moyenne

Afin d'introduire le concept de durée de vie résiduelle moyenne, il est nécessaire de donner ces définitions.

Définition 1.1.0.11. *Supposons que la variable aléatoire X a la distribution de fonction F et que l'intégrale*

$$\int_{-\infty}^{\infty} |x| dF(x),$$

existe (est fini). Ensuite, la valeur attendue EX de X existe et est donnée par l'intégrale,

$$EX = \int_{-\infty}^{\infty} x dF(x).$$

La valeur attendue de X est aussi appelée la moyenne de X, ou l'attente de X, et est souvent désigné par μ.

Proposition 1.1.0.4.

$$EX = \int_{0}^{\infty} \bar{F}(x) dx - \int_{-\infty}^{0} F(x) dx; \tag{1.10}$$

pour les variables aléatoires non-négatives, c'est-à-dire pour des distributions telles que F(x)=0 pour $x < 0$,

$$EX = \int_{0}^{\infty} \bar{F}(x) dx. \tag{1.11}$$

En effet ; Pour obtenir (1.10), on utilise le théorème de Fubini pour calculer ;

$$\begin{aligned} EX = \int_{-\infty}^{\infty} x dF(x) &= \int_{x=0}^{\infty} \int_{z=0}^{x} dz dF(x) - \int_{x=-\infty}^{0} \int_{z=x}^{0} dz dF(x) \\ &= \int_{z=0}^{\infty} \int_{x=z}^{\infty} dF(x) dz - \int_{z=-\infty}^{0} \int_{x=-\infty}^{z} dF(x) dz \\ &= \int_{0}^{\infty} \bar{F}(x) dx - \int_{-\infty}^{0} F(x) dx. \end{aligned}$$

Voir la Figure 1.1

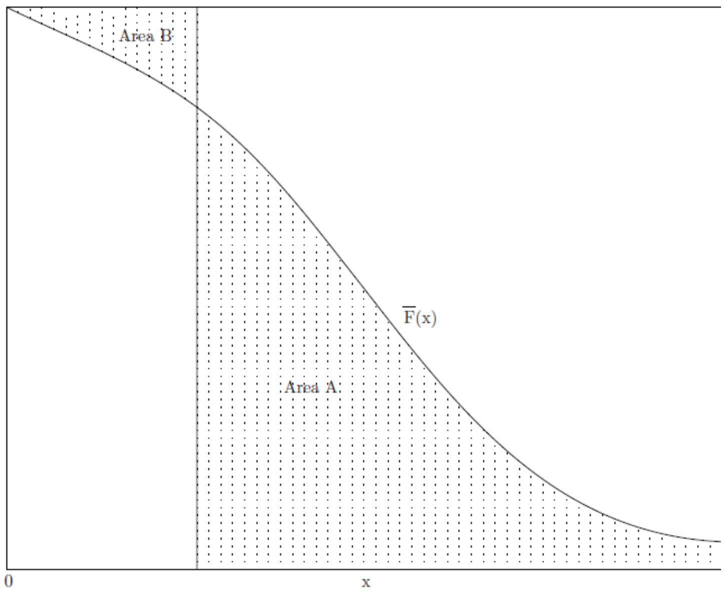

figure 1.1 L'attente en termes de superficie : $EX = ZoneA - ZoneB$

Définition 1.1.0.12. *La fonction de survie résiduelle moyenne $m(t)$ est la moyenne de la distribution de la survie résiduelle F_t en fonction de t. Plus explicitement, lorsque F a une moyenne finie μ et $F(x) = 0$, pour $x < 0$, le résidu moyen de la fonction de survie est donnée par :*

$$m(t) = \begin{cases} \int_0^{\infty} \frac{\bar{F}(x+t)}{\bar{F}(t)} dx = \int_t^{\infty} \frac{\bar{F}(z)}{\bar{F}(t)dz} = \int_t^{\infty} \frac{(t-z)}{\bar{F}(t)} dF(z) \quad si \quad \bar{F}(t) > 0. \\ \hfill 0 \qquad si \quad \bar{F}(t) = 0 \end{cases}$$

La fonctions de survie résiduelle moyenne fournit une autre façon de décrire une distribution. Pour voir qu'il détermine la fonction de survie, d'abord calculer directement que pour t tel que $\bar{F}(t) > 0$,

$$\frac{d}{dt} \log \int_t^{\infty} \bar{F}(z)dz = -\frac{1}{m(t)}.$$

Maintenant,on intégré les deux côtés de cette équation de 0 à x et faites usage de (1.11) pour obtenir ;

$$\begin{aligned} \int_0^x \frac{dt}{m(t)} &= \log \int_0^{\infty} \bar{F}(t)dt - \log \int_x^{\infty} \bar{F}(t)dt \\ &= \log \mu - \log m(x) - \log \bar{F}(x). \end{aligned}$$

Cela donne, pour t tel que $\bar{F}(t) > 0$.

$$\bar{F}(t) = \frac{\mu}{m(t)} \exp\{-\int_0^t \frac{dz}{m(z)}\}. \tag{1.12}$$

Voici la fonction de survie en fonction de la fonction de survie résiduelle moyenne ; ce résultat est donné par Cox(1962)[7], Muth(1977) [35] et Gupta(1979) [20]. Le taux de hazard peut également être directement obtenu de la fonction de survie résiduelle moyenne à travers l'équation.

$$r(t) = \frac{m'(t) + 1}{m(t)}. \tag{1.13}$$

Une excellente revue de la théorie et des applications du résidu moyen la fonction de survie est fournie par Guess et Proschan(1988) [19]. Voir aussi Hall et Wellner(1981) [23], Kupka et Loo (1989) [29] et Ghai jhet Mi (1999) [15] pour d'autres discussions sur la survie résiduelle moyenne.

La distributions d'équilibre

Soit F une fonction de distribution avec une moyenne finie μ telle que F(x)=0 pour x <0, on a :

$$f_{(1)}(x) = \begin{cases} \frac{\bar{F}(x)}{\mu} = \frac{\bar{F}(x)}{\int_0^\infty \bar{F}(z)dz}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

Pour toute distribution F telle que F(0)=0, $f_{(1)}(0) = 1/\mu$, nous aurons, F peut être récupéré de $f_{(1)}$. Clairement, chaque densité d'équilibre $f_{(1)}$ est décroissant. Une densité g est une densité d'équilibre si et seulement si (i) g(x)=0, x<0, (ii) g est décroissant sur $[0, \infty)$ et (iii) $g(0) < \infty$.
Il est simple de montrer que le taux de risque $r_{(1)}$ la distribution d'équilibre est l'inverse de la durée résiduelle moyenne, c'est-à-dire ;

$$r_{(1)}(t) = \frac{1}{m(t)} \tag{1.14}$$

Définition 1.1.0.13. *La fonction ϕ^+, définie pour x telle que F(x)> 0 par :*

$$\phi^+(x) = \bar{F}(x)/F(x), \tag{1.15}$$

est appelé rapport de chance de survivre au-delà du temps x. La fonction ϕ^- *défini pour x tel que* $\bar{F}(x) > 0$ *par :*

$$\phi^-(x) = F(x)/\bar{F}(x), \tag{1.16}$$

est appelé le rapport d'échec par le temps x.

Parfois, le terme «chance» est utilisé à la place de «odds ratio». Si à la fois F(x)$>$ 0 et $\bar{F}(x) > 0$, alors $\phi^+(x) = 1/\phi^-(x)$. Ces chances ratios peuvent être définis pour tous x avec la convention qu'ils prennent la valeur ∞ quand un dénominateur est 0.
Lorsque les rapports de cotes existent,

$$\bar{F}(x) = \phi^+(x)/[1 + \phi^+(x)] = 1/[1 + \phi^-(x)], \tag{1.17}$$

lorsque les deux odds ratios existent et F a une densité,

$$\phi^+(x)r(x) = s(x) \quad et \quad \phi^-(x)s(x) = r(x), \tag{1.18}$$

où r et s sont le taux de hazard et le taux de hazard inverse, respectivement. Notez que l'odds ratio $\phi^+(x)$ décroissant en x et $\phi^-(x)$ est croissant en x. On peut vérifier que le rapport de chance ϕ_t^+ le résiduel la distribution de la survie à t est donnée par :

$$\phi_t^+ = [\bar{F}(t)/\bar{F}(x + t)] - 1, \tag{1.19}$$

Une autre interpretation est donnée par :

$$\bar{H}_-(x) = e^{-F(x)/\bar{F}(x)} = e^{-\phi^-(x)}, \tag{1.20}$$

définit par une fonction de survie. De même, $-\phi^+(x)$ a toutes les propriétés d'un taux de risque inverse tel que ;

$$H_+(x) = e^{-\bar{F}(x)/F(x)} = e^{-\phi^+(x)}, \tag{1.21}$$

les équations (1.2O) et (1.21) peuvent être résolues pour que $\bar{F}$ donnée par :

$$\bar{F}(x) = \frac{1}{1 - \log \bar{H}_-(x)} = \frac{-\log H_+(x)}{1 - \log H_+(x)}.$$

Les calculs directs montrent que si F a le taux de hazard r, alors H_- le taux de risque $r(.)/\bar{F}(.)$. Ainsi, si r est croissant, alors le taux de hazard de H_- est croissant.

Résumé

On résumé, le tableau B.1. présente diverses fonctions qui définissent une fonction de survie.

Tableau B.1. Alternatives pour la détermination d'une fonction de survie

Fonction	*Fonction de survie*
Densitè :	
$f(x) = F'(x)$	$\bar{F}(x) = \int_x^\infty f(z)dz$
Fonction de hazard :	
$R(x) = -\log \bar{F}(x)$	$\bar{F}(x) = \exp\{-R(x)\}$
Taux de hazard :	
$r(x) = R'(x) = f(x)/\bar{F}(x)$	$\bar{F}(x) = \exp\{-\int_0^x r(x)dz\}$
Fonction de hazard inversé :	
$S(x) = \log F(x)$	$\bar{F}(x) = 1 - \exp\{S(x)\}$
Taux de hazard inversé :	
$s(x) = S'(x) = f(x)/F(x)$	$\bar{F}(x) = 1 - \exp\{\int_0^x s(x)dz\}$
Répartition de la vie résiduelle :	
$m(t) = \int_0^\infty \bar{F}(x+t)/F(t)dx$	$\bar{F}(t)\frac{\mu}{m(t)} \exp\{-\int_0^t \frac{dz}{m()t}\}$
Odds ratios :	
$\phi^+(x) = \bar{F}(x)/F(x)$	$\bar{F}(x) = \phi^+(x)/[1+\phi^+(x)]$
$\phi^-(x) = F(x)/\bar{F}(x)$	$\bar{F}(x) = 1/[1+\phi^-(x)]$
Distribution d'équilibre :	
$f_{(1)}(x) = \bar{F}(x)/\mu, \ x \geq 0$	$\bar{F}(x) = [f_1(x)]/[f_1(0)], \ quand$ F(0)=0.
Fonction de distribution inverse :	
$F^{-1}(p) = \sup\{z : F(z) \leq p\} \quad 0 \leq p \leq 1$	$\bar{F}(x) = 1 - \inf\{p : F^{-1}(p) \geq x\}$
$= \sup\{z : F(z) < 1\}$, p=1	

Les types de moments

Pour les variables aléatoires discrètes, la valeur attendue de X peut être écrite par :

$$EX = \sum_{n=1}^{\infty} x_i p(x_i);$$

et pour des variables aléatoires absolument continues avec la densité f,

$$EX = \int_{-\infty}^{\infty} x f(x) dx.$$

Bien entendu, pour les variables aléatoires non négatives, la limite inférieure de l'intégrale peut être remplacé par 0.
Pour une variable aléatoire avec fonction de distribution F telle que F(t)=0 pour t <0, il est bien connu que si $Y = \psi(X)$ et si l'attente EY existe, alors

$$EY = E\psi(X) = \int_{0}^{\infty} \psi(x) dF(x), \tag{1.22}$$

un cas d'intérêt particulier est $\psi(x) = x^r$, et alors ;

$$\mu_r = EX^r = \int_{0}^{\infty} x^r dF(x) \tag{1.23}$$

est appelé le rième moment de X.

Les fonctions génératrices de moments

Diverses fonctions génératrices peuvent être utilisées pour trouver les moments μ_r quand r est un nombre entier positif.

Definition 1.1.0.0.1. *La fonction*

$$mgf(s) = Ee^{sx} = \sum_{j=0}^{\infty} \frac{s^j EX^j}{j!},$$

et le moment de la fonction génératrices de x.

$$EX^r = \frac{d^r}{ds^r} mgf(s)|_{s=0} = \mu_r.$$

Par la suite on a, les relations suivantes ;

$$\lambda_r = \mu_r / \Gamma(r+1), \tag{1.24}$$

$$EX^r = r\int_0^{\infty} \bar{F}(x)x^{r-1}dx \tag{1.25}$$

$$EX^r = \int_0^1 [F^{-1}(p)]^r dp. \tag{1.26}$$

1.2 Les familles de distributions

Familles de distributions indexées par un nombre réel ou par plusieurs réels les nombres sont appelés familles paramétriques, et les variables d'indexation sont appelés paramètres. Pour une distribution F ayant un paramètre θ, les notations $F(\cdot|\theta)$ et F_θ sont utilisés de manière interchangeable dans ce qui suit.
Les paramètres les plus connus sont les paramètres de localisation et d'échelle, qui sont mieux introduites au moyen de la définition suivante.

Definition 1.2.1. *Les distributions F et G sont dites de le même type si, pour un certain nombre réel b et certains a> 0,*

$$F(x) = G(ax + b), \quad \textit{pour tous } x;$$

ou bien par cette formule :

$$F(\frac{x-b}{a}) = G(x).$$

Definition 1.2.2. *Une famille paramétrique* $\{F(\cdot|\lambda), \lambda > 0\}$ *de la forme* $F(x|\lambda) = F(\lambda x|1)$ *est une famille de paramètres d'échelle et* λ *est appelé un paramètre d'échelle.*

1.3 La fonction de distribution avec poids (Mixtures of distribution)

Si F_1 et F_2 sont des fonctions de distribution et $0 < \pi < 1$,

$$F = \pi F_1 + (1-\pi)F_2; \tag{1.27}$$

ou équivalent,

$$\bar{F} = \pi\bar{F}_1 + (1-\pi)\bar{F}_2; \tag{1.28}$$

alors F est dit être un mélange de F_1 et F_2. Sauf si F est la distribution de fonction d'une variable aléatoire constante, il peut apparaître comme un mélange de deux différentes distributions de façon infiniment différentes de sorte qu'une décomposition ou la représentation en mélange de la forme (1.31) n'est pas unique.

Definition 1.3.1. *Soit $F = F_\theta | \theta \in \Theta$ être une famille de distributions et Soit G une distribution sur Θ. Alors,*

$$F(x) = \int_\Theta F_\theta(x) dG(\theta)$$

est le mélange de F *par rapport à* G.

Les mélanges ont parfois été appelés distributions composées.
Il est facile de voir que les densités (si elles existent) et les fonctions de survie de mélanges ont des mélanges des densités correspondantes et leur fonctions de survie . C'est,

$$f(x) = \int_\Theta f_\theta(x) dG(\theta),$$
$$\bar{F}(x) = \int_\Theta \bar{F}_\theta(x) dG(\theta),$$

mais les formules correspondantes ne sont pas vraies pour les fonctions de hazard ou taux de risque Lorsque les distributions F_θ de (1.32) ont des densités f_θ, F et le taux de hazard ;

$$r(x) = \frac{\int_\Theta f_\theta dG(\theta)}{\int_\Theta \bar{F}_\theta(x) dG(\theta)}, \tag{1.29}$$

un cas particulier intéressant est le mélange de deux distributions, disons F_1 et F_2, comme dans (1.31). Soit $\bar{\pi} = 1 - \pi$. Ensuite, quand les densités existent, (1.32) et (1.33) prendre la forme ;

$$F(x) = \pi F_1(x) + \bar{\pi} F_2(x), \tag{1.30}$$

et ;

$$r(x) = \frac{\pi f_1(x) + \bar{\pi} f_2(x)}{\pi \bar{F}_1(x) + \bar{\pi} \bar{F}_2(x)}. \tag{1.31}$$

Dans ce cas particulier, $\bar{F}(x) = \pi \bar{F}_1(x) + \bar{\pi} \bar{F}_2(x)$.

La fonction de hazard dans le cas d'une mixture

La fonction de hazard d'un mélange, déterminée par les fonctions de hazard ρ_θ des composants $F(\cdot|\theta)$ avec la distribution de mélange G, est donné explicitement au moyen de la transformée de hazard par :

$$\bar{F}(x) = \int_\Theta \bar{F}(x|\theta)dG(\theta)$$

1.4 Les familles paramétriques : exemples de base

Dans cette section, les distributions exponentielles, Weibull, gamma et lognormales sont brièvement présentés pour les rendre disponibles à titre illustratif fins.

La distribution exponentielle

La famille de paramètres des distributions exponentielles, souvent référée simplement comme "la distribution exponentielle", est sans concurrence pour la position de la famille de base la plus fondamentale des distributions de la survie. Pour cette distribution, le paramètre $\lambda > 0$ est un paramètre d'échelle et ;

$$\bar{F}(x) = e^{-\lambda x}, \quad x \geq 0, \tag{1.32}$$

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0, et \tag{1.33}$$

$$r(x) = \lambda, \quad x \geq 0. \tag{1.34}$$

Pour la distribution exponentielle, il est facile de voir que,

$$\frac{\bar{F}(x+t)}{\bar{F}(t)} = \bar{F}(x); \tag{1.35}$$

C'est-à-dire la probabilité conditionnelle de survivre à une période supplémentaire de temps x,

$$\mu_r = EX^r = \int_0^\infty x^r \lambda e^{-\lambda x} dx = \Gamma(r+1)/\lambda^r. \tag{1.36}$$

La distribution gamma

La famille de deux paramètres de distributions gamma comprend la disttibution exponentielle en tant que un cas particulier. Alors que la densité de distribution du gamma a une belle forme, la fonction de survie et le taux de risque peuvent être écrit sous forme fermée uniquement pour certaines valeurs de paramètres. Encore une fois, le paramètre d'échelle $\lambda > 0$. De plus, il existe un paramètre de forme $\upsilon > 0$ et ;

$$f(x|\lambda, \upsilon) = \lambda^{\upsilon} x^{\upsilon-1} e^{-\lambda x}/\Gamma(\upsilon), \quad x \geq 0, \tag{1.37}$$

avec le paramètre de forme $\upsilon = 1$, il s'agit juste d'une distribution exponentielle. Le fait que cette densité s'intègre à l'unité est une conséquence directe de la définition habituelle de la fonction gamma comme une intégrale.
Lorsque le paramètre de forme υ est un entier,

$$\bar{F}(x|\lambda, \upsilon) = \sum_{k=0}^{\upsilon-1} e^{-\lambda x}(\lambda x)^k/k!, \quad x \geq 0. \tag{1.38}$$

La distribution de weibull

La distribution de Weibull est une autre famille de deux paramètres qui comprend la distribution exponentielle. Cette famille a un paramètre d'échelle λ et un paramètre de forme α, les deux positifs. Contrairement à la distribution du gamma, le fonction de survie ici est une forme simple, en particulier ;

$$\overline{F}(x) = \exp\{-(\lambda x)^{\alpha}\}, \quad x \geq 0, \tag{1.39}$$

La différenciation de cette fonction de survie donne la fonction de densité,

$$f(x) = \alpha\lambda(\lambda x)^{\alpha-1} \exp\{-(\lambda x)^{\alpha}\}, \quad x \geq 0, \tag{1.40}$$

Clairement, et le taux de risque

$$r(x) = \alpha\lambda(\lambda x)^{\alpha-1}, \quad x \geq 0. \tag{1.41}$$

La distribution de lognormale

Si Y est une variable aléatoire ayant une distribution normale et $X = e^Y$, alors X est dit avoir une distribution lognormale. Parce que le normal fonction de distribution n'a pas de forme fermée, ni la distribution fonction ni le taux de risque de la distribution lognormale peut être exprimé sous forme fermée. Mais la densité peut être obtenue à partir du densité de la distribution normale.
La distribution log-normale peut être utilement paramétrée dans plusieurs façons, dont trois sont notés ici. Supposer que $-\infty < \mu$, $\beta < \infty$, $0 < \lambda$, σ, α, et laissez ;

$$\mu = -\log \lambda, \sigma = 1/\alpha et \beta = \alpha^2 \mu = \mu/\sigma^2. \tag{1.42}$$

Pour $x > 0$, la densité est donnée par :

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\{\frac{-(\log x - \mu)^2}{2\sigma^2}\} \tag{1.43}$$

$$= \frac{\alpha}{\sqrt{2\pi}x} \exp\{\frac{-[\log(\lambda x)^\alpha]^2}{2}\} \tag{1.44}$$

$$= x^\beta \exp\{-\beta^2/2\alpha^2\} \frac{\alpha}{\sqrt{2\pi}x} \exp\{\frac{-[\log x^\alpha]^2}{2}\} \tag{1.45}$$

Ainsi, la densité peut être exprimée en termes de paramètres (μ, σ), (λ, α) ou (α, β).

1.5 Les familles non paramétriques : exemples de base

Un certain nombre de familles non paramétriques de distributions de la survie sont discutées au chapitre 3 ; ici, quelques exemples de base sont brièvement présentés.

Les densités Log-Concave et Log-Convexe

Le logarithme d'un certain nombre de densités standard est soit convexe soit concave.

(i) La densité normale est un exemple bien connu d'une densité log-concave.
(ii) On peut voir et la concavité de la fonction logarithmique aussi que la densité du gamma est log concave pour $\upsilon \geq 1$ et log convexe pour $\upsilon \leq 1$.
(iii) La densité de Weibull est log concave pour $\alpha \geq 1$ et log convexe pour $\alpha \leq 1$.

Chapitre 2

Les familles paramétriques

2.1 Intod無ction

La famille paramètrique de survie la plus importante est la famille de distributions exponentielles. Cette importance est en partie due du fait que plusieurs des familles les plus couramment utilisées de distributions de la survie sont des extensions à deux ou trois paramètres de la distributions exponentielles. Mais les distributions exponentielles, avec leur constant taux de risque, forment une base de référence pour évaluer d'autres familles. Car ils ont un seul paramètre, ils sont assez simples à décrire et sont exceptionnellement favorables aux analyses statistiques.
En raison de leurs propriétés remarquables, les distributions exponentielles se poser naturellement dans les paramètres théoriques. Ils ont beaucoup de caractérisations d'importance à la fois théorique et pratique.

2.2 La distribution exponentielle

Definition 2.2.1. *On note $\overline{F}$ la fonction de survie définie par :*

$$\overline{F}(x) = e^{-\lambda x} \quad \forall x \geq 0,$$

on définit la fonction de densité par :

$$f(x) = \lambda e^{-\lambda x} \quad \forall x \geq 0,$$

Taux de hazard :

$$r(x) = \lambda.$$

Cette fonction est aussi appeleé fonction de risque instantané de décés . le taux de hasard ,note λ. est définit par $r(x) = \lambda \quad \forall x \geq 0$. La distribution exponentielle ;

$$f_x(x) = \lambda e^{-\lambda x} \quad x \geq 0. \tag{2.1}$$

$$\begin{aligned} F_x(t) &= \int_0^t \lambda \exp^{-\lambda s} ds \\ &= 1 - e^{-\lambda t}, \end{aligned}$$

$$\begin{aligned} \overline{F}(x) &= 1 - F_x(x) \qquad (2.2) \\ &= 1 - (1 - e^{-\lambda x}) \\ &= e^{-\lambda x} \quad (la\, fonction\, de\, survie), \end{aligned}$$

$$\begin{aligned} r(x) &= \frac{f_x(x)}{\overline{F}(x)} \qquad (2.3) \\ &= \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} \\ &= \lambda \qquad (Taux\, de\, hasard). \end{aligned}$$

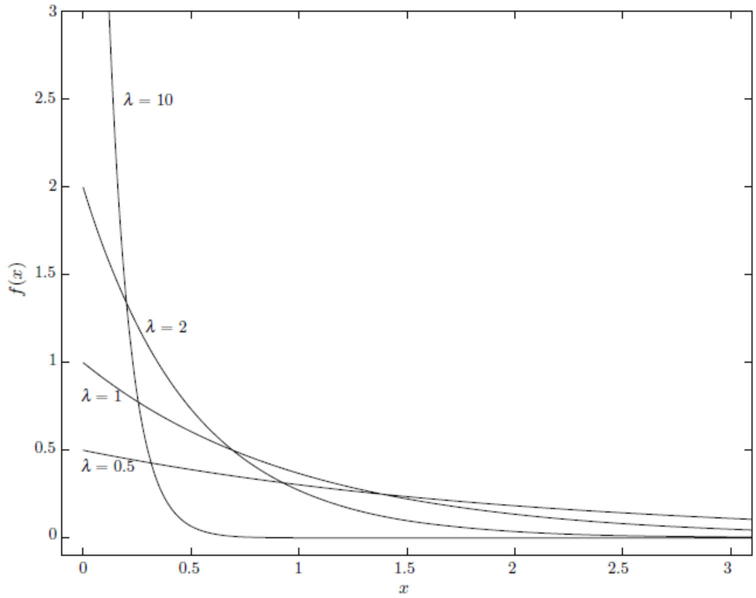

figure 2.1 : Densités de la distribution exponentielle

Lorsque le paramètre $\lambda > 0$, agit à la fois comme un paramètre d'échelle et d'un paramètre de taux de risque. On a beaucoup de distribution de la vie sont lieés la distribution exponentielle .
Les fonctions de survie prennent souvent une forme plus pratique que les fonctions de distribution ona :

$$F(x) = 1 - e^{-\lambda t}. \tag{2.4}$$

Pour les distributions exponentielle ,il est facile de vérifier que la durée de survie risuduelle à la distribution t est independante de t,en fait , ceci caractérise la distribution exponentielle .
La moyenne de survie résiduelle de la distribution exponentielle est indépendante de $l'\hat{a}ge$ t, une autre proprieté caractérisante la distribution G avec la fonction de hasard R.

$$\overline{G}(x) = \exp\{-R(x) = P(x > R(x))\}.$$

où X a une distibution exponentielle de paramêtre 1,de ce point de vue , il est à prévoir que la distribution exponentielle va jouer un rôle centre . Parce que la fonction de survie a une forme simple, il est facile d'écrire le rapport :

$$\begin{aligned}\phi(x) &= \frac{\overline{F}(x)}{F(x)} \\ &= \frac{e^{-\lambda x}}{1-e^{-\lambda x}} \\ &= \frac{1}{e^{\lambda x}-1}.\end{aligned}$$

Pour trouver le temps total sur test transformer,il faut d'abord vérifier que ;

$$F^{-1}(P) = [-log(1-p)]/\lambda, \quad 0 \leq p \leq 1, \tag{2.5}$$

À partir de (2.5) il s'ensuit que le temps total sur le transformation définition est donnée par cette forme :

$$\psi(p) = p/\lambda \quad 0 \leq p \leq 1, \tag{2.6}$$

Il est verifier que la moyenne de la distribution 1 est $\frac{1}{\lambda}$, de sorte que le temps total d'essai normalisé sur transformer ;

$$\widetilde{\psi} = \frac{\psi}{\mu},$$

est donnée par :

$$\psi(p) = p \quad 0 \leq p \leq 1,$$

pour les distributions exponentielle est donnée par :

$$L(p) = p + (1-p)\log(1-p) \quad 0 \leq p \leq 1.$$

Et que l'indice de Gini est de $\frac{1}{2}$.
Les deux graphes representent l'inverse et courbes de lorenz pour la distribution exponentielle.

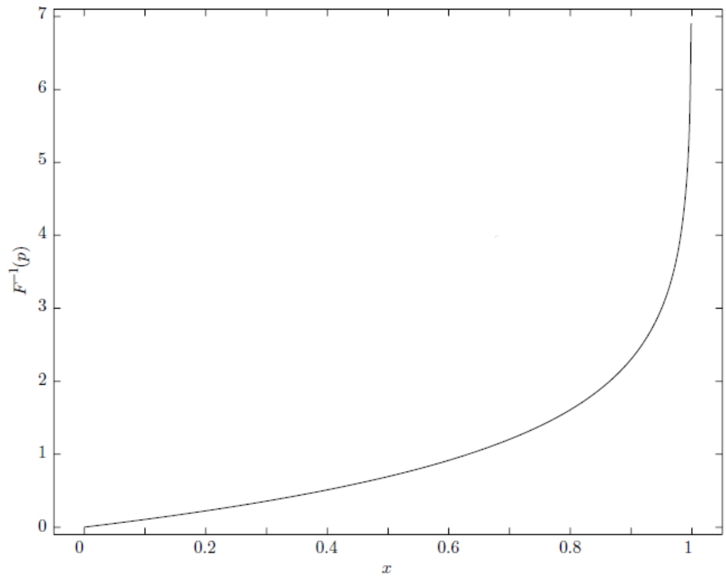

figure 2.2 :la fonction de distribution exponentielle inverse

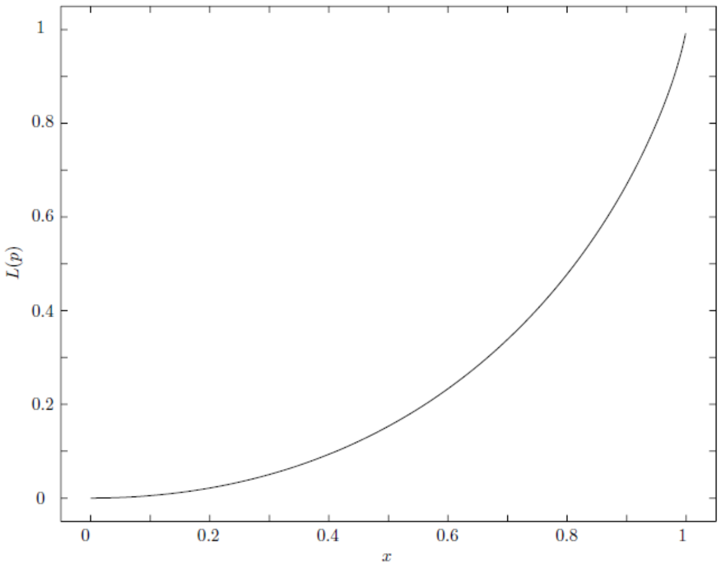

figure 2.3 : La courbe de Lorenz pour la distribution exponentielle.

Les moments

Si X a une distribution exponentielle de paramètre λ. Puis pour $r > -1$;

$$\mu_r = Ex^r = \int_0^\infty x^r \lambda e^{-\lambda x} dx = \frac{\Gamma(r+1)}{\lambda^r}, \tag{2.7}$$

ou Γ est la fonction gamma d'habitude discutée, ainsi les distributions exponentielle ont des moments de tous ordres supérieure à (-1) et ils prennent une forme simple. Les moments normalisés,

$$\lambda_r = \frac{\mu_r}{\Gamma(r+1)} = \frac{1}{\lambda^r}, \tag{2.8}$$

avoir même une forme plus simple. Rappelons que le moment normalisé joue un rôle spécial, il s'ensuit que x a l'écart type ;

$$\delta^2 = var(x) = \frac{1}{\lambda^2} = \mu_1^2, \tag{2.9}$$

cela signifie que x est le coefficient de variation.

$$cv(x) = \frac{\mu}{\delta} = 1. \tag{2.10}$$

Ainsi, pour la distribution exponentielle, l'écart type et les moyennes sont égales ; ceci est en contraste frappant avec la distribution normale, où ces quantités sont sans rapport. Cela pourrait être attendu le fait que la distribution normale a plus d'un paramètre.
Un calcul direct montre que la transformée de laplace ϕ de x est donné par :

$$\phi(s) = E(e^{-sx}) = \frac{\lambda}{\lambda+s}, \quad s > -\lambda. \tag{2.11}$$

2.3 Caractérisation de la distribution exponentielle

Une distribution F est exponentielle si et seulement si ;

$$\overline{F}(x+t) = \overline{F}(x)\overline{F}(t) \quad pour\ tous\ x, t \geq 0, \tag{2.12}$$

ou bien de la manière suivante :

$$\overline{F}_t(x) = \frac{\overline{F}(x+t)}{\overline{F}(t)} = \overline{F}(x) \quad pour\, tous\, x, t \geq 0. \tag{2.13}$$

L'interprétation évidente de (2.13) est que la probabilité conditionnelle de survie pendant un temps supplémentaire x donné la survie à temps t est indépendant de l'âge t,

Proposition 2.3.1. *Si*

$$\lim_{t\to\infty} \overline{F}_t(x) = \lim_{t\to\infty} \frac{\overline{F}(x+t)}{\overline{F}(t)} = \psi(x),$$

existe pour tout $x \geq 0$, alors pour certains λ, $0 \leq \lambda \leq \infty$,

$$\psi(x) = \exp\{-\lambda x\}.$$

En effet ; Pour $x, y \geq 0$

$$\begin{aligned}
\psi(x+y) &= \lim_{t\to\infty} \overline{F}(t)(x+y) = \lim_{t\to\infty} \frac{\overline{F}(x+y+t)}{\overline{F}(t)} \\
&= \lim_{t\to\infty} \frac{\overline{F}(x+y+t)}{\overline{F}(y+t)} \frac{\overline{F}(y+t)}{\overline{F}(t)} \\
&= \lim_{t\to\infty} \frac{\overline{F}(x+y+t)}{\overline{F}(y+t)} \lim_{t\to\infty} \frac{\overline{F}(y+t)}{\overline{F}(t)} \\
&= \psi(x)\psi(y)
\end{aligned} \tag{2.14}$$

Proposition 2.3.2. *Si le taux de risque r de F a une limite positive $\lim_{t\to\infty} r(t) = \lambda$, alors F_t converge dans la distribution vers une exponentielle distribution avec le paramètre λ comme $t \to \infty$.*

Proposition 2.3.3. *Si*

$$\lim_{t\to\infty} \overline{F}_t(x) = \lim_{t\to\infty} \frac{\overline{F}(t+x)}{\overline{F}(t)}.$$

Existe sur un ensemble dense, puis sur $(0, \infty)$, la limite tend vers 0, tend vers 1, ou est la fonction de survie d'une distribution exponentielle.

Proposition 2.3.4. *Une distribution F est une distribution exponentielle si et seulement si l'un des éléments suivants et vérifié.*
$(i)F$ a le support $[0, \infty[$ et à une densité qui est à la fois log concave et log convexe sur son support.
(ii) F est à l'intersection des classes de distributions IHR et DHR.
(iii) est à l'intersection des classes de distributions IHRA and DHRA distributions.

2.4 Extensions paramétriques de Distribution exponentielle

La distribution exponentielle a un seul paramètre qui sert à la fois comme paramètre d'echelle ou comme un paramètre de fiabilité.

La distribution gamma

La distribution gamma a la de densité ;

$$f(x|\lambda,\upsilon)=\frac{\lambda^{\upsilon}x^{\upsilon-1}e^{-\lambda x}}{\Gamma(\upsilon)},\quad x\geq 0, \tag{2.15}$$

dans ce cas, la fonction de survie peut être donnée sous forme fermée seulement lorsque υ est un entier.

$$\overline{F}(x|\lambda,\upsilon)=\frac{\sum_{k=0}^{\upsilon-1}e^{-\lambda x}(\lambda x)^{k}}{k!}. \tag{2.16}$$

Clairement, le taux de hazard de la distribution gamma ne prend pas forme même lorsque υ est un nombre entier. Le paramètre est parfois appelé un «paramètre de forme», mais il est également connu comme l '«indice» de la Distribution. Comme on peut le voir à partir de (2.15) ou (2.16), λ est un paramètre d'échelle.

Les moments

Proposition 2.4.0.5. *Si F est une distribution gamma avec densité (2.15), alors*

$$EX^{r}=\frac{\Gamma(r+\upsilon)}{\Gamma(\upsilon)\lambda^{r}},\quad r>-\upsilon, \tag{2.17}$$

ce résultat peut être obtenu directement à partir de la définition du fonction gamma comme une intégrale. De (2.16), il s'ensuit que ;

$$EX=\frac{\upsilon}{\lambda}\quad et\quad VarX=\frac{\upsilon}{\lambda^{2}}, \tag{2.18}$$

de sorte que le coefficient de variation est $CV(X)=\frac{\sqrt{VarX}}{EX}=\frac{1}{\sqrt{\upsilon}}$. *D'une manière similaire, on peut vérifier que F a la transformation de Laplace ;*

$$Ee^{-sx}=[\frac{\lambda}{(\lambda+s)}]^{\upsilon},\quad s>-\lambda. \tag{2.19}$$

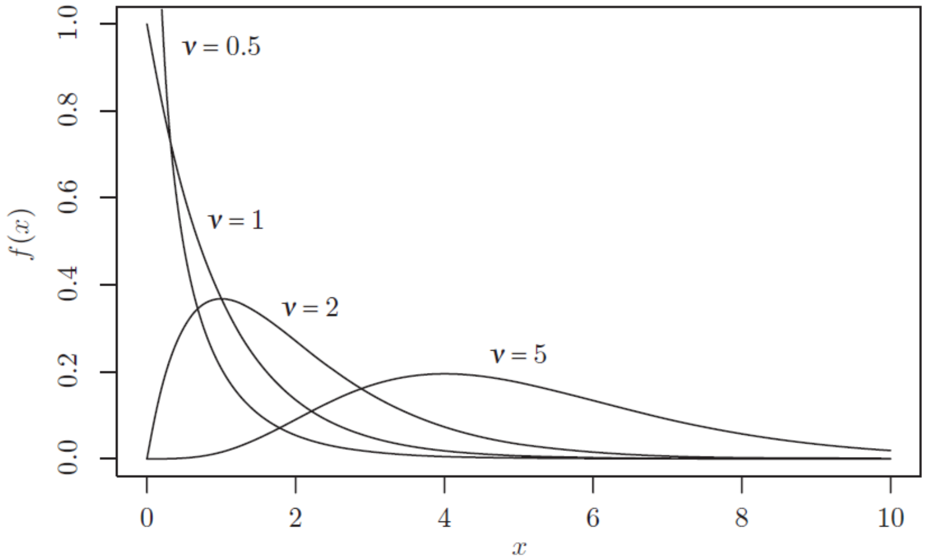

figure 2.4 : Densités de la distribution gamma ($\lambda = 1$)

La dérivations de la distribution gamma

Il y a plusieurs façons de dériver la distribution gamma, et chacune est instructif à sa manière. Tous ceux décrits ici sont basés dans certains manière sur la distribution exponentielle.

À partir d'un processus de poisson

Une dérivation standard de la distribution gamma est via un processus de Poisson. Soit X le temps d'attente pour le υth saut dans le processus de Poisson $N(t), t = 0$ avec le paramètre λ. La fonction de survie de X peut être obtenue directement :

$$P\{X > x\} = P\{N(x) < \upsilon\} = \frac{\sum_{k=0}^{\upsilon-1} e^{-\lambda x}(\lambda x)^k}{k!};$$

c'est-à-dire que X a la fonction de survie (2.16).

Par l'introduction d'un paramètre de moment

Le plus simple dérivation de la distribution gamma via la distribution exponentielle semble être par l'introduction d'un paramètre de moment .

C'est simple parce que les moments de la distribution exponentielle, ont la forme simple $\mu_r = \frac{\Gamma(r+1)}{\lambda^r}$. Lorsque cela est fait, la densité résultante est :

$$f(x|\lambda, \beta) = \frac{\lambda^{\beta+1} x^{\beta} e^{-\lambda x}}{\Gamma(\beta+1)}, \quad x \geq 0, \quad \lambda > 0, \quad \beta > -1 \tag{2.20}$$

Par l'introduction d'un paramètre de convolution

La transformation de la place ϕ de la distribution exponentielle est donnée par :

$$\phi(s) = \frac{\lambda}{\lambda + s}, \quad s > -\lambda,$$

parce que la distribution exponentielle est infiniment divisible, il s'ensuit que ;

$$\phi(s|\upsilon) = [\phi(s)]^{\upsilon} = (\frac{\lambda}{\lambda + s})^{\upsilon}, \quad s > -\lambda. \tag{2.21}$$

Est une transformation de Laplace, pour tout $\upsilon > 0$.

En présence d'une mixture

. Pour $\upsilon < 1$, la densité (2.15) apparaît comme un mélange de distributions exponentielles. La représentation suivante est due à Gleser (1989) [17] :

$$f(x|\lambda, \upsilon) = \int \theta e^{-\theta x} g(\theta|\lambda, \upsilon) d\theta, 0 < \upsilon < 1, \tag{2.22}$$

ou,

$$\begin{aligned} g(\theta|\lambda, \upsilon) = & \frac{(\theta - \lambda)^{-\upsilon} \lambda^{\upsilon}}{\theta \Gamma(1-\upsilon) \Gamma(\upsilon)}, \theta \geq \lambda, \\ = & \quad 0, \quad autrement. \end{aligned} \tag{2.23}$$

La densité du gamma (2.15) a été obtenue par Pearson (1895) [37] et est connu sous le nom de courbe de Pearson de type III. Pearson a dérivé la densité d'un équation différentielle. Si X est la densité du gamma (2.15), alors on peut montrer que $\frac{1}{X}$ a la densité g donnée par :

$$g(x) = \frac{\lambda^{\upsilon}}{\Gamma(\upsilon) x^{\upsilon-1}} e^{-\frac{\lambda}{x}}, \quad x > 0 \tag{2.24}$$

Cette densité a également été obtenue par Pearson (1895)[37] et est connue Courbe de Pearson de type V. Pour plus de détails sur les courbes de Pearson, voir aussi Elderton et Johnson (1969) [12] ou Kendall et Stuart (1963)[28].

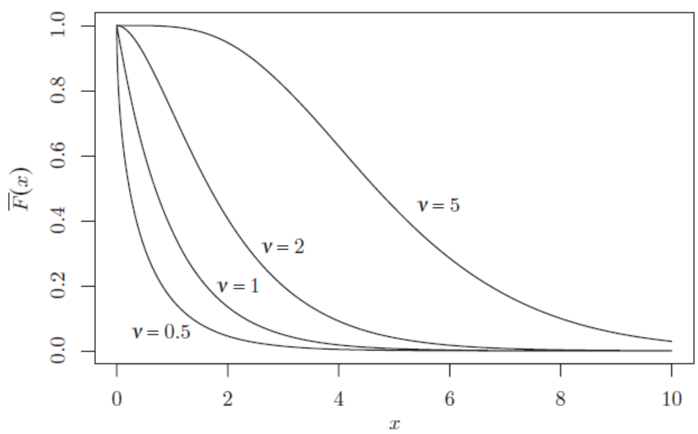

figure 2.5 : Fonctions de survie de la distribution gamma ($\lambda = 1$)

Propriétés de la distribution et de la fonction de survie

En général, ni la fonction de distribution ni la fonction de survie de la distribution gamma a une expression simple ; ils peuvent être donnés seulement en termes de fonction gamma incomplète lorsque υ n'est pas un nombre entier. Cependant, les propriétés de ces fonctions peuvent encore être déterminées.

Proposition 2.4.0.6. *La fonction de survie $\overline{F}$ correspondant à la densité gamma (2.15) est log concave, pour $\upsilon \geq 1$, et log convexe sur $[0, \infty)$, pour $\upsilon \leq 1$. La fonction de distribution gamma F est log concave, pour tout υ.*

Le taux de risque

Parce que la fonction de survie de la distribution du gamma peut être donnée seulement en termes de fonction gamma incomplète lorsque υ n'est pas entier, ni le taux de risque ni le taux de risque inverse ne peuvent être exprimé sous forme fermée. Même ainsi, une forme du taux de risque peut être étant donné que cela est utile pour l'identification de ses propriétés. Suivant l'approche de Barlow et Proschan (1975)[3], écrire la fonction de survie comme une partie intégrante de la densité pour obtenir ;

$$\frac{1}{r(x)} = \int_x^\infty (\frac{z}{x})^{\upsilon-1} e^{-\lambda(z-x)} dx = \int_0^\infty [1 + \frac{u}{x}]^{\upsilon-1} e^{-\lambda u} du. \tag{2.25}$$

où la seconde intégrale est obtenue à partir du premier par le changement de variable $u = z - x$. Maintenant, il est clair que ;

$$\begin{array}{ll} r(x) & \text{est croissant en } x \geq 0 \text{ quand } \upsilon > 1 \\ r(x) & \text{est la constante } \lambda \text{ quand } \upsilon = 1, et \\ r(x) & \text{est decroissant en } x \geq 0 \text{ quand } 0 < \upsilon < 1. \end{array} \tag{2.26}$$

Les résultats du taux de risque (2.26) découlent également directement de la concavité logarithmique et log-convexité de la fonction de survie ;

$$\begin{aligned} \lim_{x \to 0} r(x) &= 0, \text{ pour } \upsilon > 1 \\ &= \lambda, \text{ pour } \upsilon = 1 \\ &= \infty, \text{ pour } 0 < \upsilon < 1 \end{aligned} \tag{2.27}$$

De la deuxième forme de $r(x)$ dans (2.28), il s'ensuit directement que :

$$\lim_{x \to \infty} r(x) = \lambda \text{ pour tous } \upsilon > 0 \tag{2.28}$$

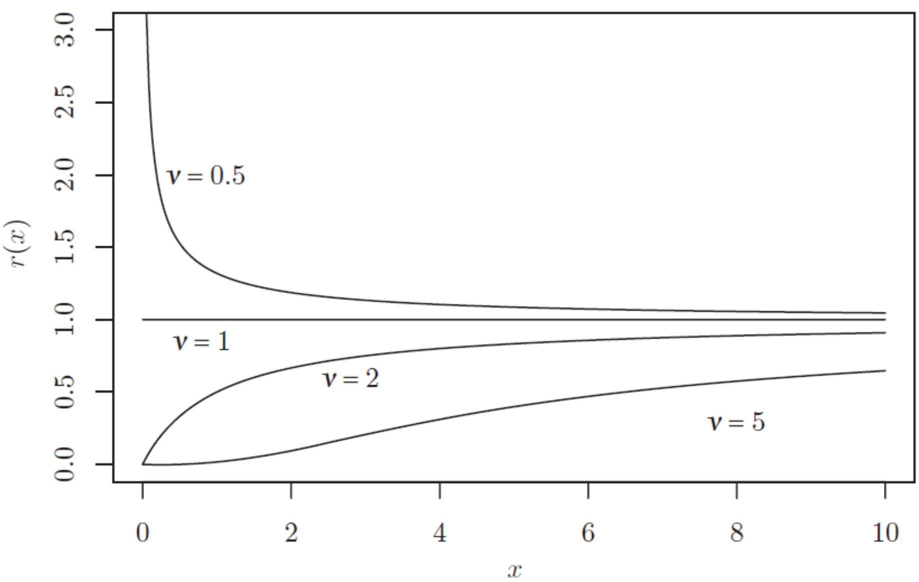

figure 2.6 : Taux de hazard de la distribution gamma ($\lambda = 1$)

Notez que pour λ fixe, le taux de risque à tout x fixe est décroissant υ ; ceci est une conséquence du fait que υ est un paramètre de convolution (voir la figure 2.6).

Parce que la fonction de distribution du gamma F est log concave, la dérivée du log F est décroissante ; cela signifie que le taux risque inverse décroissant pour toutes les valeurs de υ. Ce fait contraste avec les propriétés du taux de risque qui dépendent de si $\upsilon \geq 1$ où $\upsilon \leq 1$.

La distribution résiduelle de la survie

Proposition 2.4.0.7. *Si F est une distribution du gamma avec densité (2.15), alors pour tout υ, la fonction de survie résiduelle $\overline{F}_\tau(x) = \frac{\overline{F}(x+\tau)}{\overline{F}(\tau)}$ satisfait,*

$$\lim_{\tau\to\infty} \overline{F}_\tau(x) = e^{-\lambda x},$$

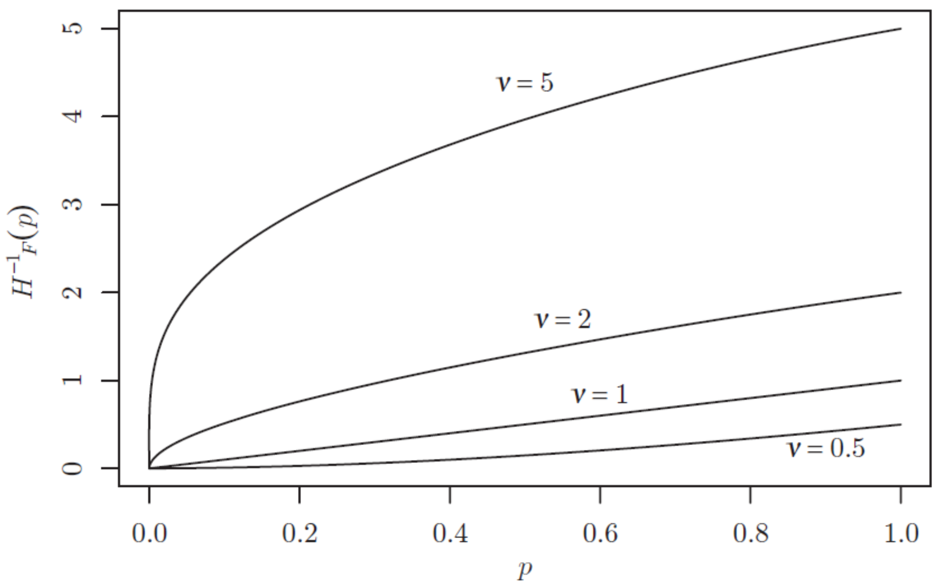

figure 2.7 : Temps total sur la transformée de test pour la distribution gamma ($\lambda = 1$)

La distribution de Weibull

Avec l'introduction d'un paramètre de puissance α, la fonction survie exponentielle devient ;

$$\overline{F}(x|\lambda, \alpha) = \exp\{-(\lambda x)^\alpha\}, \lambda, \alpha, x \geq 0, \tag{2.29}$$

c'est la fonction de survie de la distribution de Weibull.De (2.29), il résulte directement que la densité f et le taux de risque r de la distribution de Weibull sont donnés par :

$$f(x|\lambda,\alpha) = \alpha\lambda(\lambda x)^{\alpha-1}\exp\{-(\lambda x)^{\alpha}\}, x \geq 0, \tag{2.30}$$

$$r(x|\lambda,\alpha) = \alpha\lambda(\lambda x)^{\alpha-1}, \alpha \geq 0. \tag{2.31}$$

Voir les figures 2.5, 2.6 et 2.7 pour les graphiques de la densité, la fonction de survie et le taux de risque. Pour cette distribution, le paramètre de puissance α est généralement appelé «paramètre de forme». Malheureusement, aucun paramétrage de la famille Weibull a été standardisé ; certains auteurs écrivent λ à la place de λ^{α}. Divers paramétrages de la distribution de Weibull sont discutés dans un article de revue par Hallinan (1993) [22].
Une mention précoce de la distribution de Weibull a été faite par Thiele (1872) [48] citant Oppermann du Insurance Record du 11 février 1870. a proposé d'utiliser ;

$$r(x) = r(x|\lambda_1, 1/2) + r(x|\lambda_2, 1) + r(x|\lambda_3, 3/2),$$

taux de risque des jeunes jusqu'à l'âge de 20 ans. Cette somme de trois les taux de hasard de Weibull produisent une fonction de survie qui est le produit de trois fonctions de survie de la forme (2.29). Dans un contexte similaire, Makeham (1890)[31] a dérivé la distribution de Weibull avec le paramètre de puissance 2.
La distribution de Weibull a été rencontrée par Fisher et Tippett (1928) [13] comme une distribution de valeur extrême, une dérivation discutée. Dans une étude de la distribution de la taille des particules dans charbon en poudre, il a été encore dérivé par Rosin et Rammler (1933)[39] en utilisant méthodes entièrement différentes. La distribution de Weibull a été proposée par Weibull (1939a, b) [51] [52] dans le contexte de la résistance des matériaux et des ruptures dans les solides. Mais plus tard, dans des documents plus accessibles (Weibull, 1951, 1952)[53] [54], il promu la distribution pour l'utilisation dans l'analyse d'une plus grande variété qu'un la fonction de hazard peut être non-négative.
La distribution de Weibull a été largement utilisée dans les études médicales,

et il a gagné une popularité considérable en tant que modèle dans plusieurs autres les champs, y compris la foresterie et l'ingénierie ; un certain nombre de documents employant la distribution de Weibull dans une grande variété d'applications est listée par Murthy. En raison de sa popularité plutôt omniprésente, des articles mettant en garde contre la discrimination aveugle.

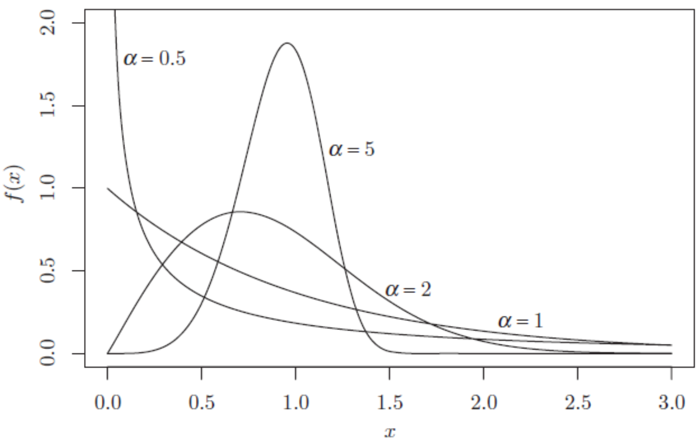

figure 2.8 : Densités de la distribution de Weibull ($\lambda = 1$)

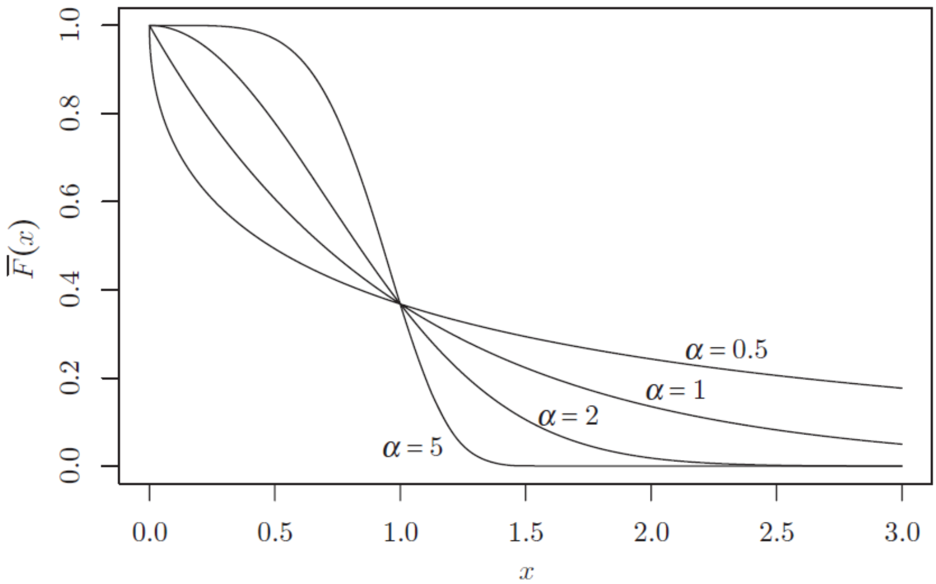

figure 2.9 : Fonctions de survie de la distribution de Weibull ($\lambda = 1$)

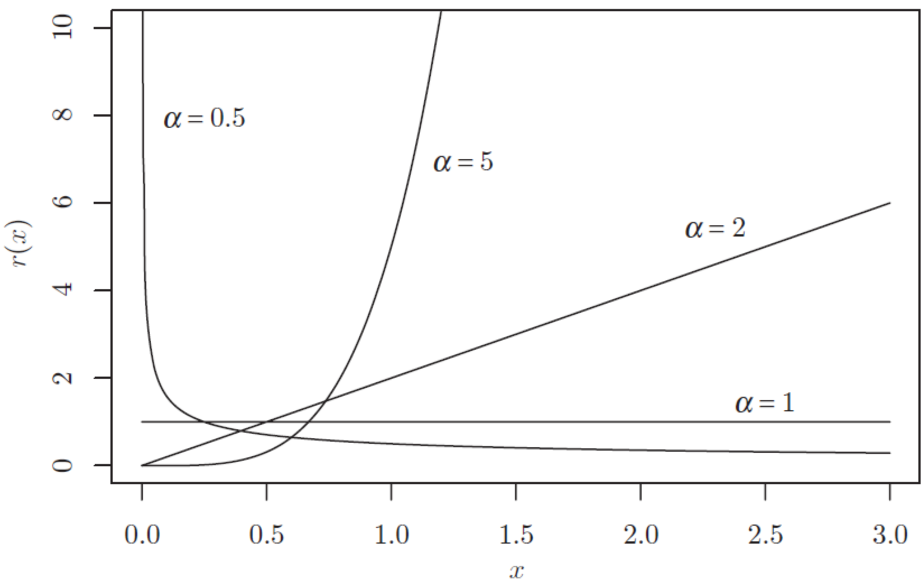

figure 2.10 Taux de risque pour la distribution de Weibull ($\lambda = 1$)

Caractérisations de la distribution de Weibull

Si X a une distribution de Weibull avec le paramètre de puissance α, alors X^α a une distribution exponentielle. Ce fait peut être utilisé pour traduire toute caractérisation de la distribution exponentielle à une caractérisation de la distribution de Weibull avec paramètre de puissance fixe. Par exemple, Wang (1976) [49] a montré qu'une variable aléatoire X satisfait ;

$$PX > \sqrt[\alpha]{x^\alpha + y^\alpha}|X > y = PX > x, x, y > 0, \tag{2.32}$$

si et seulement si X a une distribution de Weibull avec le paramètre de puissance α. Pour le prouver, réécrivez (2.32) dans le formulaire ;

$$PX^\alpha > x^\alpha + y^\alpha|X^\alpha > y^\alpha = PX^\alpha > x^\alpha, x, y > 0.$$

Cette forme de la propriété "manque de mémoire" que X^α est une distribution exponentielle. Pour une autre caractérisation de la distribution de Weibull, voir Arnold et Isaacson (1976) [1].

Une représentation de mélange

Walker et Stephens (1999)[50] ont observé que si ;

$$g(x|u) = \frac{\alpha x^{\alpha-1}}{u}, \quad 0 < x < u^{1/\alpha},$$

est la densité d'une distribution uniforme sur $[0, u^{1/\alpha}]$ avec un paramètre de puissance α, et si

$$h(u) = c^2 u e^{-cu}, u > 0,$$

est une densité du gamma avec 2 paramètre, puis le mélange ;

$$f(x) = \int_{x^\alpha}^{\infty} g(x|u)h(u)du.$$

Est une densité de Weibull ; avec $c = \lambda^\alpha$; cette densité est donnée par (2.30).

Mélange des distributions de Weibull

Des mélanges de distributions de Weibull ont été utilisés dans diverses applications, un grand nombre d'entre eux sont répertoriés par Murthy. Les

taux de risque de même le mélange de deux distributions de Weibull peut prendre une variété de formes ; ces taux de risque peuvent avoir jusqu'à quatre maxima et minima relatifs. Pour un catalogue de ces taux de risque,voir Jiang et Murthy (1998)[25].
Si X est une distribution de Weibull avec $\alpha = 2$ et Y a une distribution exponentielle, alors min (X,Y) a un taux de risque linéaire croissant.Les mélanges de distributions avec de tels le taux de risque peuvent avoir une variété de formes de taux de risque ; pour une étude détaillée de ces mélanges, voir Block [4].

Les distributions de Weibull inverses et généralisées

Si X est une distribution de Weibull, alors la distribution de $1/X$ est parfois appelé une distribution de Weibull inverse. Un calcul direct montre que si X a la fonction de survie (2.33), alors $1/X$ a la fonction de distribution ;

$$w_2^*(x) = \exp\{-(\theta x)^{-\alpha}\}, \theta, \alpha > 0, x > 0.$$

Cette distribution de valeur extrême pour les maxima, a déjà été trouvé par Fréchet (1927) [14]. La densité (2.30) de la distribution de Weibull et la densité de W_2^* combiner bien pour former ce qu'on pourrait appeler la densité d'un étendu la famille de Weibull. A cet effet, il est pratique de reparamétrer W_2^* en remplaçant θ par λ (pas $1/\lambda$), et remplacer α par $-\alpha$, de sorte que $\alpha < 0$. Ensuite, W_2^* prend la forme ;

$$F(x|\lambda, \alpha) = \exp\{-(\lambda x)^{\alpha}\}, a < 0, \lambda > 0, x > 0, \tag{2.33}$$

la distribution de cette densité, disons f, qui se combine avec (2.30) pour devenir ;

$$f(x|\lambda, \alpha) = |\alpha|\lambda(\lambda x)^{\alpha-1} \exp\{-(\lambda x)^{\alpha}\}, \lambda > 0, \alpha \neq 0, x > 0. \tag{2.34}$$

Pour $\alpha < 0$, cette densité est unimodale, avec un mode au point ;

$$x = \frac{1}{\lambda}(\frac{\alpha - 1}{\alpha})^{1/\alpha};$$

Cela peut facilement être vérifié en définissant la dérivée du logarithme de la densité est égale à 0.

Quand $\alpha < 0$ est le taux de hazard ;

$$r(x|\lambda,\alpha) = \frac{|\alpha|\lambda(\lambda x)^{\alpha-1}\exp -(\lambda x)^{\alpha}}{1-\exp -(\lambda x)^{\alpha}}, x > 0. \tag{2.35}$$

Satisfait ;

$$\lim_{x\to\infty} r(x|\lambda,\alpha) = \lim_{x\to 0} r(x|\lambda,\alpha) = 0.$$

Lorsque $\alpha < 0$, le taux de hasard r de (2.35) est un mode unique ; c'est-à-dire que r est une forme d'une fonction de distribution avec poids inversée. Cette distribution a été étudiée par Jiang, Murthy et Ji (2001, p.115) [24], qui déterminent que le mode du taux de hazard est donné par le solution de l'équation ;

$$z(t) = (1-\frac{1}{\alpha})(1-e^{-z(t))}, \tag{2.36}$$

où $z(t) = (\lambda t)^{\alpha}$. Parce que le côté gauche de (2.39) est linéaire en z(t) et le côté droit est croissant et convexe en z(t), il n'y a qu'un positif Solution.

La distribution de la survie résiduelle de Weibull

La distribution de la survie résiduelle de Weibull est une fonction de survie ;

$$\overline{F}_t(x) = \exp -[\lambda(x+t)]^{\alpha} + (\lambda t)^{\alpha}, x, \alpha, \lambda > 0, \tag{2.37}$$

et taux de risque ;

$$r_t(x) = \alpha\lambda(x+t)^{\alpha-1}, x, \alpha, \lambda > 0. \tag{2.38}$$

Notez que ce taux de risque satisfait (2.32), mais que contrairement au taux de hazard de la distribution de Weibull, peut $r_t(0) = \alpha\lambda t^{\alpha-1}$ prendre d'autres valeurs que 0 et ∞.
Notez que pour $\alpha > 0$, $\overline{F}_t(x) < \overline{F}(x)$, C'est ;

$$\overline{F}(x+t) < \overline{F}(x)\overline{F}(t), x, t > 0.$$

De sorte que F est NBU (New Better than Used). Cette inégalité est inversée si $\alpha < 1$, et alors F est NWU (new worse then used). Ces résultats peuvent être obtenus facilement du fait que pour positif numéros, $a_1, ..., a_n$, $(\sum_{i=1}^{n} a_i^r)^{1/r}$ décroissant dans $r > 0$.

La Durée de survie résiduelle de la distribution de Weibull

Si F est une distribution de Weibull, la distribution de la survie résiduelle est une fonction de survie, la densité et le taux de risque donné par :

$$\bar{F}_t(x) = \exp\{-[\lambda(x+t)]^\alpha + [\lambda t]^\alpha\}, \quad (2.39)$$

$$f_t(x) = \bar{F}_t(x)\lambda[\alpha\lambda(x+t)]^{\alpha-1}, \quad (2.40)$$

$$r_t(x) = \lambda[\alpha\lambda(x+t)]^{\alpha-1}. \quad (2.41)$$

Cette distribution n'a pas trouvé sa place dans la littérature, et son l'importance n'est pas claire. Mais ici, il y a une nouvelle flexibilité dans le choix le taux de risque à 0, donc dans certaines circonstances, la distribution peut être un concurrent à d'autres extensions à trois paramètres de la distribution de Weibull. Ces distributions de vie résiduelle ont un taux risque monotone, et bien sûr, ces taux de risque sont seulement ceux de la distribution de Weibull, mais avec l'origine déplacé à t.
Si la variable aléatoire X est la fonction de survie de (2.29), alors $Y = \log[1 + (X/t)]$ a une fonction de survie Gompertz ;

$$\bar{G}(y) = \exp\{-\xi(e^{\alpha y} - 1)\}, y \geq 0,$$

où $\xi = (\lambda t)^\alpha$. La distribution de Gompertz .

Proposition 2.4.0.8. *Dans l'ordre de transformation convexe, de distribution résiduelle de la vie (2.29) decroissant dans t pour $\alpha \leq 1$, et croissant dans t pour $\alpha \geq 1$, les paramètres α et λ étant fixes.*
En utilisant le fait que

$$\bar{F}_t^{-1}(z) = \{[(\lambda t)^\alpha - \log z]^{1/\alpha}/\lambda\} - t,$$

La convexité requise peut être vérifiée par différenciation.

La distribution de Weibull avec un paramètre d'inclinaison

Lorsque F est une fonction de distribution de Weibull, alors extension géométrique-extrême stable de la distribution de Weibull avec la fonction de survie ;

$$\bar{F}(x|\lambda,\gamma,\alpha) = \frac{\gamma \exp[-(\lambda x)^{\alpha}]}{1-\bar{\gamma}\exp[-(-\lambda x)^{\alpha}]}, x \geq 0, \lambda, \alpha, \gamma > 0, \tag{2.42}$$

la densité et le taux de hazard de la distribution donnée par (2.29). En particulier, le taux de risque est ;

$$r(x|\lambda,\alpha,\gamma) = \frac{\lambda\alpha(\lambda x)^{\alpha-1}}{1-\bar{\gamma}\exp[-(\lambda x)^{\alpha}]}, x \geq 0, \lambda, \alpha, \gamma > 0. \tag{2.43}$$

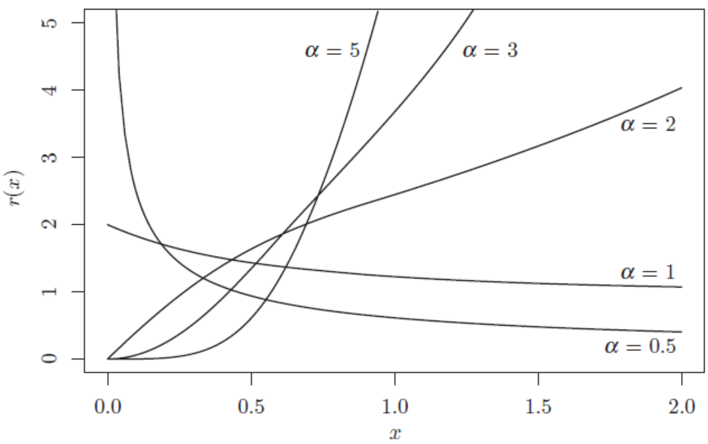

figure 2.11 : Taux de risque de la distribution de Weibull avec paramètre d'inclinaison ($\lambda = 1, \gamma = 0.5$)

Cette fonction est représentée graphiquement sur les figures 2.11, 2.12 et 2.13. Il peut être vérifié en utilisant des méthodes élémentaires que le taux de risque est croissant si $\gamma \geq 1$, $\alpha \geq 1$, et décroissant si $\gamma \leq 1$, $\alpha \leq 1$. Si $\alpha > 1$ et $\gamma < 1$, alors le taux de risque croissant initialement et finalement croissant, mais il peut y avoir un intervalle où il décroissant. De même, si $\alpha < 1$ et $\gamma > 1$, alors le taux de risque est initialement décroissant et finalement décroissant, mais il peut y avoir un intervalle où il est croissant.

Si Y a une distribution exponentielle avec le paramètre 1, alors $X = Y^{1/\alpha}/\lambda$ a la fonction de survie de la distribution de Weibull. De même, si Y a la fonction de survie avec $\lambda = 1$, alors $X = Y^{1/\alpha}/\lambda$ a la fonction de survie (2.29). Par conséquent, les moments.

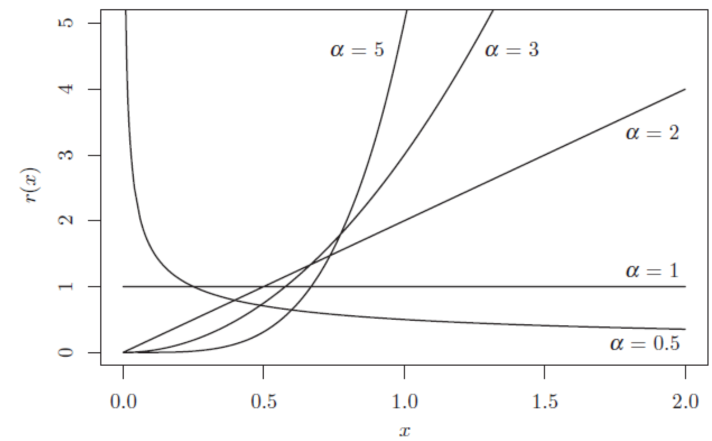

figure 2.12 : Taux de risque de la distribution de Weibull avec paramètre d'inclinaison $(\lambda = 1, \gamma = 1)$

En utilisant $EX^r = [EY^{r/\alpha}]/\lambda^r$. Ainsi, si X a la fonction de survie, puis ;

$$EX^r = \frac{r\gamma}{\alpha\lambda^r}\sum_{j=0}^{\infty}\frac{\bar{\gamma}^j}{(j+1)^{r/\alpha}}\Gamma(\frac{r}{\alpha}), |\bar{\gamma}| \leq 1. \tag{2.44}$$

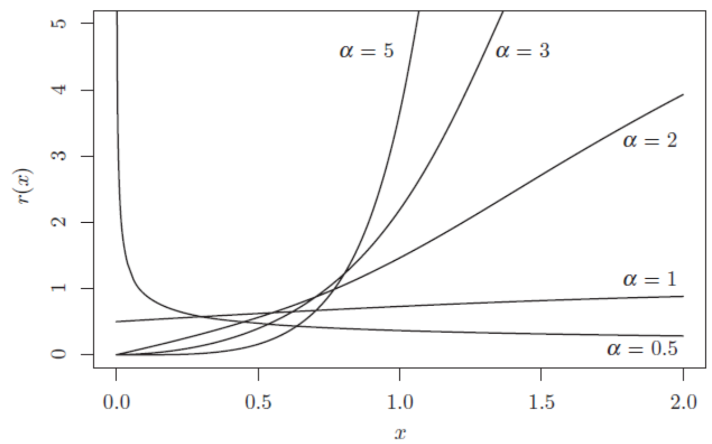

figure 2.13 : Taux de risque de la distribution de Weibull avec paramètre d'inclinaison $(\lambda = 1, \gamma = 2)$

$\alpha \backslash \gamma$	0.2	0.5	1.0	1.5	2.0	10	20
0.2	24.68	61.01	120.0	177.4	233.3	1014	1828
0.3	2.046	4.897	9.261	13.28	17.03	61.14	99.29
0.4	0.807	1.844	3.323	4.609	5.764	17.54	26.38
0.5	0.537	1.164	2.000	2.690	3.290	8.779	12.48
0.6	0.446	0.915	1.505	1.971	2.365	5.681	7.739
0.7	0.410	0.801	1.266	1.621	1.913	4.222	5.563
0.8	0.398	0.742	1.133	1.422	1.657	3.408	4.373
0.9	0.397	0.710	1.052	1.299	1.495	2.900	3.641
1	0.402	0.693	1.000	1.216	1.386	2.558	3.153
2	0.518	0.714	0.886	0.994	1.072	1.522	1.710
5	0.725	0.835	0.918	0.966	0.999	1.165	1.225
10	0.842	0.906	0.951	0.977	0.994	1.076	1.105
100	0.982	0.989	0.994	0.997	0.999	1.007	1.010

Tableau 2.14 : Valeurs attendues de X lorsque X a une distribution Weibull avec inclinaison paramètre γ et $\lambda = 1$.

Cependant, les moments ne peuvent pas être donnés sous une forme fermée ; donc, même le premier moment de doit être obtenu numériquement. Pour $\lambda = 1$, tableau 2.14 montre les valeurs du premier moment pour diverses combinaisons de α et γ.

Notez que pour α fixe, la valeur attendue croissant dans γ, ceci est une conséquence fait qu'un commande stochastique implique la commande d'attentes.
En écrivant

$$EX^s = \int_0^\infty Sx^{s-1}\bar{F}(x|\lambda, \gamma, \alpha)dx, \quad s > 0$$

Il peut être montré avec le changement de variables $y = (\lambda x)^\alpha$ un l'intégration par des parties.

$$\lim_{\alpha \to \infty} EX^{-s} = \lambda^{-s}, \quad s > 0,$$

Ce sont les moments d'une variable aléatoire dégénérée à $1/\lambda$. Ça peut vérifier aussi que la limite de la fonction de survie comme $\alpha \to \infty$ est dégénérer.

Chapitre 3

Les familles non paramétriques : Densités et les taux de hazard

3.1 Introduction

Un certain nombre de familles non paramétriques de distributions de la vie ont été étudiées, en cas particulier dans le contexte de la théorie de la fiabilité ; le plus important de ces familles sont discutées dans ce chapitre. Parce que les familles sont défini par des propriétés qui ont des interprétations physiquement significatives, une hypothèse selon laquelle une distribution réside dans une famille particulière parfois peut être justifiée par une compréhension physique du mécanisme de défaillance. Pour ces familles non paramétriques, de nombreuses procédures statistiques sont disponibles qui peut être considéré comme debout entre paramétrique et standard analyses sans distribution.
Les familles non paramétriques ont reçu peu d'attention dans des contextes médicaux, où les analyses ont tendance à être basée soit sur paramétrique modèles ou ont utilisé des méthodes non paramétriques standard.
En dehors de l'utilité des familles non paramétriques pour les données analyses, une compréhension de leurs propriétés est très utile étudier des familles paramétriques particulières parce que les membres de ces familles appartiennent généralement à au moins une des familles non paramétriques. Un compréhension des propriétés définissant les familles non paramétriques peut

jouer un rôle important dans le choix de la famille paramétrique.

3.2 Les densités Log-Concave et Log-Convexe

Les densités log-concaves et log-convexes sont importantes en partie parce qu'elles sont souvent rencontrés et d'une partie parce qu'ils ont des propriétés intéressantes.
L'importance des densités log-concaves dans les statistiques était peut-être d'abord reconnu par Karlin et Rubin (1956) [27]. Plus récemment, le concept a joué un rôle dans la théorie économique.

Definition 3.2.1. *Si F est une fonction de distribution absolument continue, avec une version f de la densité ayant la propriété que logf est concave, alors F est dit avoir une densité log-concave. Si* $f(x) = 0$*,* $x < 0$*, et log f est convexe sur* $[0, \infty)$*, alors F est dit avoir un log-convexe densité.*

Des exemples de densités log-concaves et log-convexes sont familiers ; effectivement, il est bien connu et facilement vérifié que :
(i)Les densités normales sont log concaves,
(ii)Les densités exponentielles sont à la fois log concave et log convexe,
(iii)Les densités gamma sont log concaves si $\upsilon \geq 1$ et log convexe si $0 < \upsilon \leq 1$,
(iv) Les densités de weibull sont log concave si $\alpha \geq 1$ et si log convexe $0 < \alpha \leq 1$.

Quelques Propriétés

Proposition 3.2.0.9. *Les densités log-concaves sont unimodales, c'est-à-dire sont non décroissantes jusqu'à un certain point et non-croissantes au-dela de ce point. En effet ; Pour montrer qu'une densité est unimodale, il suffit de montrer que pour toute constante positive c.* $f(x) - c$*, change de signe au plus deux fois, et l'ordre* $-, +, -$ *s'il y a deux changements de signe. Mais cela est vrai si et seulement si ;* $\log f(x) - d$*, est pour une constante d, au plus deux changements de signe, dans l'ordre* $-, +, -$ *s'il y a deux signes changements. Ce modéle de changement de signe est valide car le* $\log f$ *est concave.*

Proposition 3.2.0.10. *Soit $\{f_\theta, \theta \in A\}$ une famille des densités log-convexes concentré sur l'intervalle I et supposons que A est un convexe ouvert ensemble. Si $f_\theta(x)$ est une fonction mesurable de θ pour chaque x fixe dans I, et si G est une fonction de distribution de probabilité sur A, alors le mélange $f(x) = \int f_\theta(x)dG(\theta)$ est une densité log-convexe.*

Les densités complétement monotones

Les densités complétement monotones forment une sous-classe importante des densités log-convex est une fonction ϕ définie sur $(0, \infty)$ est dite complétement monotone s'il posséde les dérivées $\phi^{(n)}$ de tous les ordres n, et ;

$$(-1)^n \phi^{(n)}(x) \geq 0, x > 0.$$

Il résulte de résultats bien connus sur les transformations de laplace qu'une fonction de survie est complétement monotone si et seulement si pour une fonction de distribution H, il a la forme ;

$$\overline{F}(x) = \int_0^\infty e^{-\lambda x} dH(\lambda), x \geq 0, \tag{3.1}$$

ainsi, les fonctions de survie complétement monotones sont exactement celles qui sont des mélanges des fonctions de survie exponentielles. Une fonction de survie est complètement monotone si et seulement si elle est une densité f complètement monotone c-à-d :

$$f(x) = \int_0^\infty \lambda e^{-\lambda x} dH(\lambda), x > 0. \tag{3.2}$$

Exemple 1. *Quand H est une distribution exponentielle, disons, avec paramètre $1/\theta$, puis(3.1) devient ;*

$$\overline{F}(x) = \int_0^\infty e^{-\lambda x} \theta^{-1} e^{-\lambda/\theta} d\lambda = 1/(1 + \theta x), x > 0.$$

C'est un type spécial de la fonction de survie de Pareto [35].

Proposition 3.2.0.11. *(Steutel,1967,1969) [46] [47]. Si $\overline{F}$ est complètement monotone (ou de manière équivalente, si la densité correspondante est complètement monotone), alors il est infiniment divisible.*

Proposition 3.2.0.12. *Si $F(0) = 0$ et $\overline{F}$ est complètement monotone, alors f est log convexe sur $[0, \infty)$.*

Proposition 3.2.0.13. *Si f est log concave, alors le taux de risque $r = f/\overline{F}$ de F croissant, et la taux de risque inverse $S = f/F$ de F décroissant. Si f est log concave sur $[0, \infty)$, alors le taux de risque r décroissant sur $[0, \infty)$.*

Exemple 2. *Supposons que $r(x) = x + (1 + x)^{-1}$, $x \geq 0$ Ça peut être vérifiée par différenciation que r est croissante, et donc $R(x) = \int_0^x r(t)dt = \frac{1}{2}x^2 + \log(1 + x)$ est convexe. $\log \overline{F}(x) = -R(x)$, et par conséquent $\overline{F}$ est log concave.*
Pour montrer que f n'est ni log concave, ni log convexe, utilisez la formule $\log f(x) = \log r(x) + \log F(x)$ pour vérifier que la dérivée seconde de $\log f(x)$ est positif à 0 et devient finalement négatif.
Un calcul un peu plus impliqué mais similaire montre également que $\log F(x)$ ni convexe ni concave.

Proposition 3.2.0.14. *Si $f(x) = 0$, $x < 0$ et le $\log f$ est convexe sur $[0, \infty)$, alors le $\log F$ est concave $(-\infty, \infty)$.*

Proposition 3.2.0.15. *Si f est log concave [convexe], alors la vie résiduelle la densité $f_t(x) = f(x + t)/\overline{F}(t)$ est log concave [convexe] dans x, pour tout $t > 0$.*

3.3 Le taux de risque monotone

Il ne devrait pas être surprenant que le concept d'un "risque monotone" taux découle de la condition que le taux de risque $r = f/\overline{F}$ est monotone, mais cette idée nécessite l'existence d'une densité. le concept peut être défini de diverses manières, dont les meilleures exiger l'existence d'une densité.

Definition 3.3.1. *Soit F une distribution telle que $F(x) = 0$, $x < 0$. Alors F est dit avoir un taux de risque décroissant [croissant] si pour tous $x \geq 0$ et tout t tel que $F(t) < 1$.*

$$\frac{[F(t + x) - F(t)]}{1 - F(t)} = \frac{[\overline{F}(t) - \overline{F}(t + x)]}{\overline{F}(t)}, \tag{3.3}$$

Est croissant en t et $F(0+) = 0$ [décroissant en t et $F(0-) = 0$].
Cette condition indique que la probabilité de survivre pendant un intervalle de temps de longueur x décroissant [croissant] à l'âge initial.
Le concept d'un taux de risque croissant peut être étendu utilement à des distributions telles que la distribution normale qui ont une masse positive sur l'intervalle $(-\infty, 0)$, mais une distribution ne peut pas avoir le taux de hazard dans tout l'intervalle $(-\infty, \infty)$; le soutien de ces distributions doit avoir un point limite .

Les conditions équivalentes

La proposition suivante donne une condition pour le taux de hazard monotone cela est généralement plus pratique à vérifier que l'état de la définition.

Proposition 3.3.0.16. *Une distribution F a une croissante [décroissante] taux de risque si et seulement si pour tout t tel que* $\overline{F}(t) > 0$*,*

$$\overline{F}_t(x) = \frac{\overline{F}(t+x)}{\overline{F}(t)}, dcroissant[croissant]entpourtoutx = 0, \tag{3.4}$$

c'est-à-dire que les distributions de durée de survie résiduelle décroissant stochastiquement [croissant] en t.

En effet, pour tout $x \geq 0$ et tout t tel que $F(t) < 1$,

$$\frac{\overline{F}(t) - \overline{F}(t+x)}{\overline{F}(t)} = 1 - \frac{\overline{F}(t+x)}{\overline{F}(t)},$$

est croissant en t si et seulement si $\overline{F}(t+x)/\overline{F}(t)$ est décroissant.

Proposition 3.3.0.17. *Si F est une distribution telle que* $F(x) = 0$ *pour* $x < 0$*, alors F a un taux de risque décroissant [croissant] si et seulement si fonction de hazard* $R = -\log \overline{F}$ *est convexe si finie [concave sur* $[0, \infty)$].

Proposition 3.3.0.18. *Supposons que* $F(0-) = 0$ *et F a une densité. Alors F a un taux de risque décroissant [croissant] si et seulement si est une version f de la densité telle que le taux de risque correspondant* $r = f/\overline{F}$ *croissant [décroissant] sur* $[0, \infty)$.

Proposition 3.3.0.19. *La fonction de distribution F a une croissante de taux de risque si et seulement si le déterminant ;*

$$\begin{vmatrix} \overline{F}(t_1 - s_1) & \overline{F}(t_1 - s_2) \\ \overline{F}(t_2 - s_1) & \overline{F}(t_2 - s_2) \end{vmatrix} \geq 0,$$

où $s_1 \leq s_2$, $t_1 \leq t_2$, c'est-à-dire, $\overline{F}$ est une fonction de fréquence de Pólya d'ordre 2.
La fonction de distribution F a un taux de risque décroissant si et seulement si le support de F est $[0, \infty)$ et le déterminant ;

$$\begin{vmatrix} \overline{F}(t_1 + s_1) & \overline{F}(t_1 + s_2) \\ \overline{F}(t_2 + s_1) & \overline{F}(t_2 + s_2) \end{vmatrix} \geq 0,$$

chaque fois que $s_1 \leq s_2$, $t_1 \leq t_2$ et $s_1 + t_1 \geq 0$, c'est-à-dire, $\overline{F}(x + y)$ est totalement positif en x et y, $x + y \geq 0$.

Proposition 3.3.0.20. *Une distribution F a une croissante [décroissante] de taux de risque si et seulement si pour tout $x \geq 0$, et tout t tel que $\overline{F}(x+t) > 0$, le rapport de cotes $\phi_t^-(x) = [F(x + t) - F(t)]/[\overline{F}(x + t)]$ de la vie résiduelle la distribution F_t croissant [décroissant] en t.*

Proposition 3.3.0.21. *Soit X une variable aléatoire positive avec la fonction de survie $\overline{F}$, et notons la distribution de $1/X$ par G. Si $\log \overline{F}$ est concave (F est IHR), alors le $\log G$ est concave (G est un taux de risque décroissant inverse).*

En effet, car $G(x) = P\{1/X \leq x\} = P\{X \geq 1/x\} = \overline{F}(1/x)$, $x > 0$, $\log G(x) = \log \overline{F}(1/x)$ a le dérivé,

$$\frac{d}{dx} \log G(x) = \frac{d}{dx} \log \overline{F}(1/x) = \frac{f(1/x)}{\overline{F}(1/x)} \frac{1}{x^2}.$$

Si $r(x) = f(x)/\overline{F}(x)$ croissant, alors $r(1/X)$ décroissant. Par conséquent, $r(1/X)/x^2$ est décroissant, et par conséquent, $\log G$ a la une valeur décroissante dérivé.

3.4 Les taux de risque croissant

Le contenu intuitif d'un taux de risque croissant découle de l'interprétation de r(t)dt comme la probabilité conditionnelle de défaillance dans le intervalle [t,t+dt] étant donné la survie jusqu'à l'instant t. Ainsi, avec une croissante de taux de risque, la probabilité de défaillance dans l'instant suivant est croissante lorsque le dispositif ou l'organisme vieillit. Dans un sens très réel, c'est un traduction mathématique du concept intuitif de «vieillissement adverse» mais il serait injuste de prétendre que c'est la seule traduction mathématique de ce concept.

Les propriétés des distributions avec des taux de risque croissants

Proposition 3.4.0.22. *Si F est un taux de risque croissant, alors F a un densité sauf éventuellement à l'extrémité droite de son support, où il peut avoir une masse positive.*
Ce résultat est une conséquence de la définition de la concavité logarithmique. Il montre que l'utilisation d'une définition n'exigeant pas l'existence d'une densité n'a pas étendu l'idée très loin. Mais encore, c'est une extension utile.

Proposition 3.4.0.23. *Supposons que F a un taux de risque croissant. Alors F a des moments finis de tous les ordres positifs finis.*

Le théorème de préservation d'un taux de risque croissant

Le théorème suivant indique que la classe de distributions avec croissante les taux de risque sont fermés sous convolutions. La preuve élégante donné ici et par Barlow, Marshall et Proschan (1963)[2] est due à Frank Proschan ; cela dépend de la théorie de la positivité totale, qui est connectée au comportement du taux de hazard. Il y a aussi intégration critique par parties. Aucune vraie preuve élémentaire n'est pas connue.
L'interprétation intuitive du théorème suivant est : Si chacun des deux appareils s'usent en ce sens qu'ils ont un taux risque croissant, et si l'un des

dispositifs est utilisé comme une pièce de rechange pour l'autre et mettre en service au moment où le premier échoue, alors ce système (un dispositif et sa réserve) s'use dans le sens où il est un taux de risque croissant (en supposant que les deux longueurs de vie sont indépendantes). Le lecteur peut souhaite interpréter ce théorème dans d'autres contextes tels que la médecine ou la biologie.

Théorème 3.4.0.1. *(Barlow, Marshall et Proschan, 1963)* [2]. *Si F et G avoir des taux de risque croissants, alors la convolution H=F*G a un taux de risque croissant.*

En effet, Supposons que F est une densité de f et G est une densité de g. la condition croissant du taux de hazard peut être écrit en termes d'un déterminant. Pour $t_1 < t_2$, et $u_1 < u_2$, forme le déterminant ;

$$D = \begin{vmatrix} \overline{H}(t_1 - u_1) & \overline{H}(t_1 - u_2) \\ \overline{H}(t_2 - u_1) & \overline{H}(t_2 - u_2) \end{vmatrix}$$

$$= \begin{vmatrix} \int \overline{F}(t_1 - s)g(s - u_1)ds & \int \overline{F}(t_1 - s)g(s - u_2)ds \\ \int \overline{F}(t_2 - s)g(s - u_1)ds & \int \overline{F}(t_2 - s)g(s - u_2)ds \end{vmatrix}$$

$$= \int\int_{s_1<s_2} \begin{vmatrix} \overline{F}(t_1 - s_1) & \overline{F}(t_1 - s_2) \\ \overline{F}(t_2 - s_1) & \overline{F}(t_2 - s_2) \end{vmatrix} \begin{vmatrix} g(s_1 - u_1) & g(s_1 - u_2) \\ g(s_2 - u_1) & g(s_2 - u_2) \end{vmatrix} ds_2 ds_1;$$

la dernière égalité est une application de la formule de composition de base. Maintenant, intégrer l'intégrale interne par parties pour obtenir ;

$$D = \int\int_{s_1<s_2} \begin{vmatrix} \overline{F}(t_1 - s_1) & f(t_1 - s_2) \\ \overline{F}(t_2 - s_1) & f(t_2 - s_2) \end{vmatrix} \begin{vmatrix} g(s_1 - u_1) & g(s_1 - u_2) \\ \overline{G}(s_2 - u_1) & \overline{G}(s_2 - u_2) \end{vmatrix} ds_2 ds_1.$$

Le signe du premier déterminant est le même que celui de,

$$\frac{f(t_2 - s_2)\overline{F}(t_2 - s_2)}{\overline{F}(t_2 - s_2)\overline{F}(t_2 - s_1)} - \frac{f(t_1 - s_2)\overline{F}(t_1 - s_2)}{\overline{F}(t_1 - s_2)\overline{F}(t_1 - s_1)},$$

en supposant que les dénominateurs sont non-nulles. Mais ;

$$\frac{f(t_2 - s_2)}{\overline{F}(t_2 - s_2)} \geq \frac{f(t_1 - s_2)}{\overline{F}(t_1 - s_2)},$$

par hypothèse et ;

$$\frac{\overline{F}(t_2 - s_2)}{\overline{F}(t_2 - s_1)} \geq \frac{\overline{F}(t_1 - s_2)}{\overline{F}(t_1 - s_1)}.$$

Proposition 3.4.0.24. *(Lynch, 1999)* [29]. *Soit* $\{F(x|\theta), \theta \in \Theta\}$ *une famille des fonctions de distribution, indexées par un paramètre* $\theta \in \Theta \subset R^n$, *pour certaines n. Si G a une densité log-concave g, et* $\log \overline{F}(x|\theta)$ *est concave dans le paire* $(x|\theta)$, *puis le mélange.*

$$\overline{F}(x) = \int_{\Theta} \overline{F}(x|\theta)dG(\theta) = \int_{\Theta} \overline{F}(x|\theta)g(\theta)d\theta,$$

est log concave ; c'est-à-dire que F a un taux de risque croissant.

En effet ; Cette proposition est une application immédiate du théorème de Prékopa21.A.14 [40].
L'exigence selon laquelle G a une densité log-concave est remplacé dans la proposition suivante par la condition plus faible que G a un taux de risque croissant.

Proposition 3.4.0.25. *(Lynch, 1999)*[30]. *Soit* $\{F(x|\theta), \theta > 0\}$ *une famille des fonctions de distribution telles que*

$$\log \overline{F}(x|\theta) \text{ est concave dans } (x,\theta), \text{ et croissant dans } \theta \text{ pour chaque } x \text{ fixe.} \tag{3.5}$$

Si G est une distribution avec un taux de risque croissant et G(0)=0, alors Le mélange

$$\overline{H}(x) = \int_0^{\infty} \overline{F}(x|\theta)dG(\theta), \tag{3.6}$$

est un taux de risque croissant. Inversement, si H est IHR (increasing hazard rate) chaque fois que (3.4) est satisfait, alors G est IHR.

En effet ; (Block, Li et Savits, 2003a)[5]. Soit G la distribution de Θ ; par hypothèse, G est un IHR. Dans la proposition C.1.i, prenez $h(t|\theta) = \overline{F}(t|\theta)$ (avec θ à la place de X) pour conclure que $Eh(t,\Theta) = \overline{H}(t)$ est log concave.

Les propriétés des distributions avec des taux de risque décroissants

Dans cette section, il est montré que les distributions avec risque décroissant les taux de risque ont des densités décroissantes, et peuvent avoir des queues assez lourdes pour exclure l'existence de moments d'ordre positif.

Proposition 3.4.0.26. *Supposons que $F(0-) = 0$ et que F ait une valeur décroissante taux de risque Alors, F a une densité sauf éventuellement positive masse à l'origine. Il y a une version f de la densité qui diminue et vérifie $f(x) > 0$, pour tout $x > 0$.*

En effet ; Une distribution F a un taux de risque décroissant lorsque $R = -\log \overline{F}$ est concave sur $[0, \infty)$. Par conséquent, R a un dérivé continu sauf pour beaucoup de points . Où ce dérivé $f = F'$ existe, il sert de densité ; ailleurs, la densité peut être défini en utilisant la continuité. Cela signifie que $R' = f/\overline{F}$ est en baisse, et par conséquent, f est décroissant.
Devrait-il y avoir un point $a < \infty$ tel que $\overline{F}(a) = 0$, alors $R(a) = \infty$ et la concavité de R serait violée. De plus, s'il y avait un point $b > 0$ tel que $\overline{F}(b) = 1$, alors R(b) = 0, et encore la concavité de R serait violé sauf si R(x)=0, pour tout x> 0, mais alors F n'est pas une distribution fonction.

Proposition 3.4.0.27. *Supposons que F a un taux de risque décroissant. Alors, F n'a pas besoin d'avoir des moments finis de n'importe quel ordre positif.*

Exemple 3. *Si $\overline{F}(x) = 1/(1 + x)$, $x \geq 0$, alors aussi $r(x) = 1/(1 + x)$, $x \geq 0$. Cette distribution est une distribution spéciale de Pareto avec taux de risque décroissant, et le premier moment n'existe pas. Le danger le taux d'une distribution de Pareto plus générale, à savoir*

$$r(x) = \frac{\lambda\alpha\xi(\lambda x)^{\alpha-1}}{1 + (\lambda x)^{\alpha}}.$$

Ce taux de risque est décroissant pour a<1. Avec $\xi = 1$, le moment n'existe finiment que pour $-\alpha < r < \alpha$.

Le théorème de préservation d'un taux de risque décroissant

Proposition 3.4.0.28. *Si F a un taux de risque décroissant, alors La distribution de la vie résiduel F_t a également un taux de risque décroissant.*
La généralisation, due à Frank Proschan, est fondamental à la compréhension des origines des distributions avec taux de risque décroissant.

Théorème 3.4.0.2. *(Barlow, Marshall et Proschan, 1963) [2]. La famille des distributions avec le taux de risque décroissant est fermé sous la formation de mélanges.*

Proposition 3.4.0.29. *Supposons que F a un taux de risque décroissant et premier moment μ. Ensuite, la distribution avec densité $f_{(1)}(x) = \overline{F}/\mu$, $x \geq 0$, a un taux de risque décroissant.*
Ce fait, parallèle, est un autre cas particulier. Cependant, voici l'hypothèse d'un premier moment fini est critique parce que les distributions de DHR n'ont pas besoin de moments finis.

Proposition 3.4.0.30. *Soit X une variable aléatoire avec distribution F et soit Y de densité $f_{(1)}$. Alors, F est un taux de risque décroissant si et seulement si $X \leq Y$.*

Résumé

Les mélanges de distributions DHR sont toujours DHR, mais les mélanges de distributions IHR n'ont pas besoin d'être IHR et peuvent même être DRH ; convolutions de RSI sont toujours IHR, mais les convolutions de Les distributions DHR peuvent être IHR

3.5 Détermination de la forme du taux de risque

Parfois, une distribution est définie en termes de densité, mais la distribution fonction et la fonction de survie ne peut pas être donnée sous forme fermée. En plus de la distribution normale, des exemples incluent la distributions de gamma lognormales introduites au chapitre 1. Lorsque ni la fonction de survie ou le taux de risque peut être donné sous forme fermée, l'étude directe du comportement du taux de risque ne peut pas être particulièrement facile. le but de cette section est de fournir quelques méthodes pour la détermination du comportement du taux de risque utile dans les chapitres ultérieurs où la distributions de les familles paramétriques sont étudiées.
On peut le voir directement à partir de la définition du taux de risque ;
(i) Si la densité f est croissant à $x = x_0$, alors le taux de risque r est croissant

à $x = x_0$;
(ii) Si le taux de risque r est décroissant à $x = x_0$, alors la densité f est décroissant à $x = x_0$.
Ces observations peuvent souvent être utiles.
Le fait essentiel à exploiter dans ce qui suit est que le comportement du taux de risque est liée au comportement de la dérivée de la logarithme de la densité, à savoir ;

$$\rho(x) = -\frac{d \log f(x)}{dx} = -\frac{f'(x)}{f(x)}, x > x_0. \tag{3.7}$$

Bien sûr, cela nécessaire que la densité soit différentiable, implicite hypothèse dans ce qui suit.

Lemme 3.5.1. *Soit f une densité strictement positive et différentiable sur* (x_0, ∞) *tel que* $\lim_{x\to\infty} f(x) = 0$. *Indiquer le taux de hazard de f par r et supposons que* $\rho(x) - c$ *a K changements de signe dans* $x > x_0$. *Alors,*
(i) r(x)-c a au plus K changements de signe dans $x > x_0$ *et*
(ii) Si r(x)-c est K changements de signe dans $x > x_0$, *alors ils sont dans le même ordre comme ceux de* $\rho(x) - c$.

Théorème 3.5.1. *Soit f une fonction de densité satisfaisant les conditions du lemme 3.6.1 avec* $x_0 = 0$.
(a) Si ρ *croissant, alors r croissant.*
(b) Si ρ *décroissant, alors r décroissant.*
(c) S'il existe x_1 *pour lequel* ρ *décroissant dans* $x \leq x_1$ *et croissant en* $x \geq x_1$, *alors il existe* x_2, $0 \leq x_2 < x_1$, *de sorte que r décroissant en* $x \leq x_2$ *et croissant en* $x \geq x_2$.
(d) S'il existe x_1 *pour lequel* ρ *croissant dans* $x \leq x_1$ *et décroissant enx* $\geq x_1$, *alors il existe* x_2, $0 \leq x_2 \leq x_1$, *de sorte que r augmente dans* $x \leq x_2$ *et décroissant en* $x \geq x_2$.

Proposition 3.5.1. *(Glaser, 1980)*[16].
(i) Supposons que ρ *satisfasse aux conditions de (c) du théorème 3.6.1. Si* $\lim_{x\to 0} f(x) = \infty$, *alors* $x_2 > 0$; *si,* $\lim_{x\to 0} f(x) = 0$, *alors* $x_2 = 0$.
(ii) Supposons que ρ *satisfasse aux conditions de (d) du théorème 3.6.1. Si* $\lim_{x\to 0} f(x) = \infty$, *alors* $x_2 = \infty$; *si,* $\lim_{x\to 0} f(x) = 0$, *alors* $x_2 < \infty$.

Proposition 3.5.2. *(Gupta et Warren, 2001)* [*21*]. *Soit f un double différentiable densité concentrée sur l'intervalle* $(0, \infty)$ *telle que f(x)> 0, pour tout x> 0. Conserver la notation introduite ci-dessus et supposer que l'équation* $\rho'(x) = 0$ *a n solutions, disons* $x_1, ..., x_n$*, où* $0 = x_0 < x_1 < ... < x_n$*. Ensuite, l'équation* $r'(X) = 0$ *a au plus une solution dans l'intervalle fermé* $[x_{k-1}, x_k]$*, k=1, 2,..., n.*

Chapitre 4

Conclusion

Pendant de nombreuses années, les auteurs ont été intéressés par le développement des méthodes pour générer des distributions multi variées. Notre motivation était de présenter des modéles qui seraient utiles dans la fiabilité et l'analyse de survie. Cela nous a conduits à l'idée d'écrire un manuscrit sur la structure des familles paramétriques et non-paramétrique. Nous tenons à dire que ce travail est limité aux aspects probabilistes portant sur des modèles de l'anayle de survie.
Comme perspective, on peut se poser quelques questions :

1. Comment se présente les familles semi-paramétriques ?
2. Comment inclure l'aspect statistique sur les différentes familles ?

Bibliographie

[1] Arnold, Barry C. and Dean Isaacson (1976). On solutions. Zeitschrift fur Warscheinlichkeitstheorie und verwandte Gebiete 35, 115-119.

[2] Barlow, Richard E., Albert W. Marshall, and Frank Proschan (1963). Properties of probability distributions with monotone hazard rate. The Annals of Mathematical Statistics 34, 375-389.

[3] Barlow, Richard E. and Frank Proschan (1975). Statistical Theory of Reliability and Life Testing. New York : Holt, Rinehart and Winston.

[4] Block, Henry W., Thomas H. Savits, and Harshinder Singh (1998). The reversed hazard rate function. Probability in the Engineering and Informational Sciences 12, 69-90.

[5] Block, Henry W., Yulin Li, and Thomas H. Savits (2003a). Preservation of properties under mixture. Probability in the Engineering and Informational Sciences 17, 205-212.

[6] Block, Henry W., Yulin Li, and Thomas H. Savits (2003b). Initial and finalbehaviour of failure rate functions for mixtures and systems. Journal of Applied Probability 40, 721-740.

[7] Cox, D.R. (1962). Renewal Theory. London, U.K. : Methuen.

[8] Cramér, Harald (1972). On the history of certain expansions used in mathematical statistics. Biometrika 59, 205-207.

[9] Charlier, C.V.L. (1906). ¨Uber das Fehlergesetz. Arkiv f¨ur Mathematik, Astronomi och Fysik 2(8). 9 pp.

[10] Elderton, W. Palin (1906). Frequency Curves and Correlation. London, U.K. : Layton.

[11] Elderton, W.P. (1934). An approximate law of survivorship and other notes on the use of frequency curves in actuarial statistics. Journal of the Institute of Actuaries 65, 1-37.

[12] Elderton, William Palin and Norman Lloyd Johnson (1969). Systems of Frequency Curves. London : Cambridge Univ. Press.

[13] Fisher, R.A. and L.H.C. Tippett (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24, 180-190.

[14] Fréchet, Maurice (1927). Sur la loi de probabilité de l'écart maximum. Annales de la Sociétè Polonaise de Mathématique, Cracovie 6, 93-116.

[15] Ghai, Gauri L. and Jie Mi (1999). Mean residual life and its association with failure rate. IEEE Transactions on Reliability 48, 262-265.

[16] Glaser, Ronald E. (1980). Bathtub and related failure rate characterizations. Journal of the American Statistical Association 75, 667-672.

[17] Gleser, Leon Jay (1989). The gamma distribution as a mixture of exponential distributions. The American Statistician 43, 115-117.

[18] Gompertz, B. (1825). On the nature of the function expressive on the law of human mortality. Philosophical Transactions of the Royal Society of London 115, 513-583.

[19] Guess, Frank and Frank Proschan (1988). Mean Residual Life : Theory and Applications. Handbook of Statistics, Vol. 7 : Quality Control and Reliability, pp. 215-224, P.R. Krishnaiah and C.R. Rao, (eds.). Amsterdam, The Netherlands : North Holland.

[20] Gupta, Ramesh C. (1979). On the characterization of survival distributions in reliability by properties of their renewal densities. Communications in Statistics-Theory and Methods A8, 685-697.

[21] Gupta, Ramesh C. and Robin Warren (2001). Determination of change points of non-monotonic failure rates. Communications in Statistics-Theory and Methods 30, 1903-1920.

[22] Hallinan, Arthur J. Jr. (1993). A review of the Weibull distribution. Journal of Quality Technology 25, 85-93.

[23] Hall, W.J. and J.A. Wellner (1981). Mean residual life. Statistics and Related Topics, pp. 169-184, M. Csorgo, D.A. Dawson, J.N.K. Rao, and A.K. Saleh,(eds.). Amsterdam, The Netherlands : North-Holland.

[24] Jiang, R., D.N.P. Murthy, and P. Ji (2001). n-fold Weibull multiplicative model. Reliability Engineering and System Safety 74, 211-219.

[25] Jiang, R. and D.N.P. Murthy (1998). Mixture of Weibull distributions-Parametric characterization of failure rate function. Stochastic Models and Data Analysis 14, 47-65.

[26] Johnson, Norman L., Samuel Kotz and N. Balakrishnan (1994). Continuous Univariate Distributions. Vol. 1, 2nd edn. New York : John Wiley et Sons.

[27] Karlin, Samuel and Herman Rubin (1956). The theory of decision procedures for distributions with monotone likelihood ratio. The Annals of Mathematical Statistics 27, 272-299.

[28] Kendall, Maurice G. and Alan Stuart (1963). The Advanced Theory of Statistics, Vol. 1, 2nd edn. New York : Hafner Publishing Co.

[29] Kupka, Joseph and Sonny Loo (1989). The hazard and vitality measures of ageing. Journal of Applied Probability 26, 532-542.

[30] Lynch, James D. (1999). On conditions for mixtures of increasing failure rate distributions to have an increasing failure rate. Probability in the Engineering and Informational Sciences 13, 33-36.

[31] Makeham, William Matthew (1890). On further development of Gompertzs law. The Assurance Magazine and Journal of the Institute of Actuaries (London) 28, 152-159, 185-192, 316-332.

[32] Mises, R. (1936). La distribution de la plus grande de n valeurs. Revue Math´ematique de l'Union Interbalkanique 1, 141-160. Selected Papers of Richard von Mises, 1964, Vol. 2, pp. 271-294. Providence, Rhode Island : American Mathematical Society.

[33] Moivre, A. (1724). Treatise of Annuities Upon Lives. London.

[34] Murthy, D.N. Prabhakar, Min Xie, and Renyan Jiang (2004). Weibull Models. Hoboken, New Jersey : John Wiley et Sons.

[35] Muth, Eginhard J. (1977). Reliability models with positive memory derived from the mean residual life function. The Theory and Applications of Reliability, Vol. II, pp. 401-435, C.P. Tsokos and I.N. Shimi (eds.). New York : Academic Press.

[36] Pareto, V. (1897). Cours d'economie Politique, Vol. II. Lausanne, Switzerland : F. Rouge.

[37] Pearson, Karl (1895). Contributions to the mathematical theory of evolution. II. Skew variations in homogeneous material. Philosophical Transactions of the Royal Society of London A186, 343-414.

[38] Pòlya, G. and G. Szegö (1972). Problems and Theorems in Analysis, Vol. I. (Translated by D. Aeppli), Springer-Verlag, Berlin and New York (1st ed.in German, 1925).

[39] Rosin, P. and E. Rammler (1933). The laws governing the fineness of coal. Journal of the Institute of Fuels 6, 29-36.

[40] Prékopa, A. (1971) Logarithmic concave measures with application to stochastic programming. Acta Scientiarum Mathematicarum. (Szeged.) 32, 301-315.

[41] Proschan, Frank (1963). Theoretical explanation of observed decreasing failure rate. Technometrics 5, 375-384.

[42] Proschan, Frank and Myles Hollander (1984). Nonparametric concepts and methods in reliability. Handbook of Statistics, Vol. 4, P. R. Krishnaiah and P.K. Sen (eds.), pp. 613-655. Amsterdam, The Netherlands : Elsevier.

[43] Rosin, P. and E. Rammler (1933). The laws governing the fineness of coal. Journal of the Institute of Fuels 6, 29-36.

[44] Sarndal, Carl-Erik (1971). The hypothesis of elementary errors and the Scandinavian school of statistical theory. Biometrika 58, 375-207.

[45] Shaked, Moshe and J. George Shanthikumar (1994). Stochastic Orders andTheir Applications. San Diego : California : Academic Press.

[46] Steutel, F.W. (1967). Note on the infinite divisibility of exponential mixtures. The Annals of Mathematical Statistics 38, 1303-1305.

[47] Steutel, F.W. (1969). Note on completely monotone densities. The Annals of Mathematical Statistics 40, 1130-1131.

[48] Thiele, T.N. (1872). On a mathematical formula to express the rate of mortality throughout the whole of life, tested by a series of observations made use of by the Danish Life Insurance Company of 1871. Journal of the Institute of Actuaries 16, 313-329.

[49] Wang, Y.H. (1976). A functional equation and its application to the characterization ofWeibull and stable distributions. Journal of Applied Probability 13, 385-391.

[50] Walker, Stephen G. and David A. Stephens (1999). A multivariate family of distributions. Biometrika 86, 703-709.

[51] Weibull, W. (1939a). A statistical theory of the strength of materials. Ingeniors Vetenskaps Akademiens Handligar,No. 151, Stockholm, Sweden.

[52] Weibull, W. (1939b). The phenomenon of rupture in solids. Ingeni¨ors Vetenskaps Akademiens Handligar, No. 153, Stockholm, Sweden.

[53] Weibull, W. (1951). A statistical distribution function of wide applicability. Journal of Applied Mechanics 18, 293-297.

[54] Weibull, W. (1952). Discussion : A statistical distribution function of wide applicability. Journal of Applied Mechanics 19, 233-234.