$$\Lambda_x = \sum_{t=1}^{\infty} \frac{1}{(1+i)^t} 1\!\!1_{]t;\infty[}(T_x)$$

ISFA
GRANDE ÉCOLE D'ACTUARIAT ET DE GESTION DES RISQUES
INSTITUT DE SCIENCE FINANCIÈRE ET D'ASSURANCES

MODÈLES DE DURÉE

Support de cours 2024-2025

Introduction

Frédéric PLANCHET

Version 4.9

Août 2024

SOMMAIRE

PRÉAMBULE

Les modèles de durée constituent un outil utilisé dans de nombreux domaines de l'assurance : durée de la vie humaine, durée de l'arrêt de travail, durée de chômage, mais aussi durée d'attente entre 2 sinistres, durée avant la ruine, *etc*.

Le domaine d'application de ces modèles est donc large.

L'objectif de cours est de présenter les principaux outils pour construire des modèles de durée ainsi que leur utilisation en assurance vie et non-vie.

Le présent document constitue une introduction à ce type de modèles et présente de manière succincte les outils de base qui seront développés dans la suite du cours.

Les aspects statistiques des modèles de durées (estimation et tests) ne sont pas abordés de manière détaillée dans cette première partie, ils seront développés dans la suite du cours. La documentation associée au cours se compose des supports suivants :

- Introduction - concepts de base (english version) ;
- Les modèles de durée paramétriques et semi-paramétriques (english version),
- Les tables de mortalité (english version) ;
- L'estimation non paramétrique (english version) ;
- Les méthodes de lissage et d'ajustements (english version).

Les supports de ce cours sont repris sous forme de livre dans PLANCHET et THÉROND [2011]. Des illustrations numériques avec le logiciel R sont également disponibles.

D'autres éléments sont disponibles sur la page du cours, https://www.ressources-actuarielles.net/md.

1. Introduction

1.1. Points de repères historiques[1]

L'analyse formalisée des données de durée remonte à l'école anglaise d'arithmétique politique, avec notamment les travaux de John GRAUNT (1620-1674) et William PETTY (1623-1687) à l'occasion des premières études sur la mortalité en Angleterre au 17ème siècle (*cf.* LE BRAS [2000]). Les notions d'espérance de vie et d'espérance de vie résiduelle sont alors définies.

La recherche de lois sous-jacentes pour ces phénomènes commence au 19ème siècle avec notamment la formule proposée par Benjamin GOMPERTZ en 1825 pour modéliser la probabilité de décéder à l'âge x :

$$h(x) = a \times b^x$$

Ce modèle (qui est en fait une progression géométrique des taux de décès de raison *b*) sera complété par William MAKEHAM en 1860 :

$$h(x) = c + a \times b^x$$

L'étude des durées de vie restera longtemps un problème étudié par les démographes et les actuaires, jusqu'à l'apparition de la théorie de la « fiabilité » pour les systèmes physiques. Ainsi W. WEIBULL publie en 1951 dans un journal de mécanique un article où il propose la forme suivante pour la fonction de hasard :

$$h(t) = \lambda \alpha t^{\alpha - 1}$$

L'article de WEIBULL aborde notamment l'une des particularités importantes des données de durée, la présence de données tronquées ou censurées.

Deux autres dates importantes doivent être citées : l'article d'E. KAPLAN et P. MEIER en 1958 dans lequel ils proposent d'utiliser dans le domaine médical un estimateur non paramétrique permettant d'intégrer les données censurées introduit en 1912 par P. BÖHMER, l'estimateur « PL » de la fonction de survie.

En 1972 David COX publie un article posant les bases d'un cas particulier important de modèle à « hasard proportionnel » faisant intervenir des variables explicatives (exogènes) en spécifiant :

$$h(x) = e^{\beta z} h_0(x)$$

avec β un vecteur de paramètres (inconnu) et h_0 la fonction de hasard de base inconnue ; il s'agit donc d'un modèle semi-paramétrique. Ce modèle de référence a donné lieu à de nombreux développements et variantes : introduction d'une évolution temporelle, prise en compte de dépendance entre les variables observées, stratification de l'effet des covariables, *etc.*

[1] Ce rappel est largement repris de DROESBEKE et *al.* [1989].

Enfin, pour clore ce bref panorama, on peut mentionner deux évolutions des modèles de durées :

- La problématique des tables prospectives et des modèles bi-dimensionnels « âge x année », dont la référence fondatrice est LEE et CARTER [1992].
- La quantification de la part non mutualisable du risque de mortalité, *via* les modèles stochastiques de mortalité (*cf.* CAIRNS et al. [2004]).

1.2. Les particularités des données de durée

La première particularité des données de durée est d'être générées par des variables aléatoires positives ; même si on peut imaginer de ramener toute variable aléatoire réelle sur $[0,+\infty[$ par une transformation bien choisie (la fonction exponentielle par exemple), il n'en demeure pas moins que cette caractéristique induit que la loi de référence des modèles de durée ne saurait être la loi normale.

L'interprétation en termes de durée des variables aléatoires étudiées va par ailleurs conduire à définir des représentations de la loi non plus au travers de la fonction de répartition, mais au travers de la fonction de survie et de la fonction de hasard.

Par ailleurs, on pourra noter comme troisième particularité à prendre en compte le fait que la situation de référence soit celle de données incomplètes. Ceci peut être la conséquence :

- Du fait que la variable aléatoire n'est observable que sur une sous partie de $[0,+\infty[$; le modèle est alors dit tronqué.
- Du fait que pour certains individus le résultat de l'expérience n'est observé que partiellement : par exemple l'expérience a une durée limitée T et pour les individus vivants en T on ne connaît pas la durée de vie, mais on sait seulement qu'elle est supérieure à T ; le modèle est alors dit censuré.

Enfin, les données de durée utilisent en général des variables explicatives exogènes : par exemple l'espérance de vie dépend du sexe, du niveau socioprofessionnel, de la région d'habitation, *etc.*

1.3. Les modèles statistiques

Les différents modèles usuels de la statistique se retrouvent dans la description des données de durée :

- Modèles paramétriques : par exemple le modèle de MAKEHAM.
- Modèles non paramétriques : c'est par exemple le cas de l'estimateur de KAPLAN-MEIER.
- Modèles semi-paramétriques : le modèle de COX est une illustration de ce type de modèle.

On peut également ajouter à cette typologie les modèles stochastiques, qui ont une place un peu à part (« sur couche » à l'un des modèles ci-dessus).

Les principaux modèles qui seront examinés dans le cours pour une population homogène sont résumés dans le schéma suivant :

Tab. 1. Synthèse des principaux modèles en l'absence d'hétérogénéité

Estimation par maximum de vraisemblance ⇐ Données tronquées à gauche et censurées à droite (e_i,t_i,d_i) ⇒ Modèle non paramétrique de Whittaker-Henderson

$$l(\theta) \propto \sum_{i=1}^{n} \left[d_i \ln\left(\mu_\theta(t_i)\right) + \ln \frac{S_\theta(t_i)}{S_\theta(e_i)} \right]$$

$$\varphi(q) = \sum_{i=1}^{p} E_i (q_i - \hat{q}_i)^2 + h \sum_{i=1}^{p-z} (\Delta^z q_i)^2$$

⇓

Discrétisation : on utilise les q_x en supposant la fonction de hasard constante entre x et x+1 ⇒ Maximum de vraisemblance discrétisé ⇒ Modèle semi-paramétrique de Brass

$$q_x = \frac{S(s) - S(x+1)}{S(x)} = 1 - \exp\left(-\int_x^{x+1} \mu(t)\,dt \right) \approx 1 - e^{-\mu_x}$$

$$\varphi(\theta) = \sum_{x=x_{Min}}^{x_{Max}} \frac{E_x}{\hat{q}_x(1-\hat{q}_x)} \left(q_x(\theta) - \hat{q}_x\right)^2 \qquad \hat{q}_x = \frac{D_x}{E_x}$$

$$\ln\left(\frac{q_x}{1-q_x}\right) = \alpha \times \ln\left(\frac{q_x^{ref}}{1-q_x^{ref}}\right) + \beta$$

$$\varphi(\alpha,\beta) = \sum_{x=x_{Min}}^{x_{Max}} \frac{E_x}{\hat{q}_x(1-\hat{q}_x)} \left(q_x(\alpha,\beta) - \hat{q}_x\right)^2$$

2. Représentation d'une distribution de survie

On considère une variable aléatoire *T* à valeurs dans $[0,+\infty[$, et on note dans la suite $F(t) = P(T \le t)$ sa fonction de répartition (continue à droite). Lorsque la densité de *T* existe, on la notera $f(t) = \frac{d}{dt}F(t) = \lim_{h \to 0} \frac{P(t \le T \le t+h)}{h}$.

2.1. La fonction de survie

La fonction de survie est par définition le complément à un de la fonction de répartition :

$$S(t) = 1 - F(t) = P(T > t)$$

S est donc une fonction décroissante telle que $S(0) = 1$ (si $P(T=0) = 0$, ce que nous supposerons) et $\lim_{t \to \infty} S(t) = 0$. Si la durée moyenne de survie existe alors elle s'exprime simplement à l'aide de S :

$$E(T) = \int_0^\infty t\,dF(t) = -\int_0^\infty t\,dS(t) = \int_0^\infty S(t)\,dt$$

<u>Démonstration</u> : On suppose que l'espérance existe. On écrit que $\int_0^\infty t\,dF(t) = \lim_{u \to +\infty} \int_0^u t\,dF(t)$; en intégrant par parties on peut écrire $\int_0^u t\,dF(t) = -\int_0^u t\,dS(t) = -uS(u) + \int_0^u S(t)\,dt$; l'inégalité de Markov assure alors que $tS(t) \le E(T)$ et donc le terme $uS(u)$ est borné. On en déduit que

l'intégrale $\int_0^\infty S(t)dt$ converge, ce qui implique que $\lim_{t\to\infty} tS(t) = 0$ et en passant à la limite on obtient le résultat attendu.

On peut démontrer également ce résultat de la manière suivante en observant que $\int_0^\infty S(t)dt = \int_0^\infty E\left(1_{\{T>t\}}\right)dt$ et par Fubini $\int_0^\infty E\left(1_{\{T>t\}}\right)dt = E\left(\int_0^\infty 1_{\{T>t\}}dt\right) = E\left(\int_0^T dt\right) = E(T)$.

On montre de la même manière que :

$$V(T) = 2\int_0^\infty tS(t)dt - E(T)^2 .$$

2.2. Survie conditionnelle

On pose tout d'abord $S_u(t) = P(T > u + t | T > u)$ la fonction de survie conditionnelle ; on s'intéresse donc à la survie d'un élément après un instant $u+t$, sachant qu'il a déjà fonctionné correctement jusqu'en *u*. En revenant à la définition de la probabilité conditionnelle on peut écrire :

$$S_u(t) = P(T > u + t | T > u) = \frac{P(T > t + u)}{P(T > u)} = \frac{S(u+t)}{S(u)}.$$

La fonction de survie conditionnelle s'exprime donc simplement à l'aide de la fonction de survie.

2.3. La fonction de hasard

La fonction de hasard[2] (ou taux de panne, taux de défaillance, taux de décès, risque instantané, etc.) est par définition :

$$h(t) = \frac{f(t)}{S(t)} = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}\ln S(t).$$

Il en résulte directement que la fonction de hasard détermine entièrement la loi de *T* et qu'on a la relation suivante :

$$S(t) = \exp\left(-\int_0^t h(s)ds\right).$$

On note en général $H(t) = \int_0^t h(s)ds$ la « fonction de hasard cumulée », qui est telle que $S(t) = \exp(-H(t))$ et qui est évidemment croissante. On utilise dans certains tests

[2] Cette expression est un anglicisme, en français on dirait « fonction de risque ».

d'adéquation le fait que $H(T)$ suit une loi exponentielle de paramètre 1. Cette propriété découle de :

$$P(H(T) > x) = P(T > H^{-1}(x)) = S(H^{-1}(x)) = \mathbf{exp}(-H(H^{-1}(x))) = \mathbf{exp}(-x).$$

D'après la définition de la fonction de survie conditionnelle et la formule ci-dessus on obtient :

$$S_u(t) = \mathbf{exp}\left(-\int_u^{u+t} h(s)\,ds\right).$$

Cela revient à dire que la fonction de hasard de la survie conditionnelle au fait d'être en fonctionnement à la date *u* est $t \to h(u+t)$. On en déduit en particulier que la fonction de hasard est croissante si et seulement si la durée de vie résiduelle après *u* est stochastiquement décroissante[3] comme fonction de *u*.

C'est souvent la fonction de hasard qui est utilisée pour spécifier un modèle de durée. Elle a en effet une interprétation « physique » ; en utilisant la définition de la fonction de hasard et de la fonction de survie on peut écrire :

$$h(t) = \lim_{u \to 0} \frac{P(t < T \le t+u \mid T > t)}{u} = \lim_{u \to 0} \frac{P(t < T \le t+u)}{uS(t)}$$

ce qui signifie que pour de « petites » valeurs de *u*, *h(t)u* est approximativement la probabilité que le composant tombe en panne entre *t* et *t+u*, sachant qu'il est en fonctionnement en *t*. En d'autres termes :

$$P(t < T \le T + dt \mid T > t) = h(t)\,dt.$$

2.4. Cas des variables discrètes

Si la variable aléatoire *T* prend des valeurs entières, sa distribution est décrite par les $p_k = P(T = k)$, pour $k \ge 0$. La fonction de survie s'écrit simplement $S(k) = \sum_{m \ge k+1} p_m$.

L'interprétation de la fonction de hasard donnée en 2.3 ci-dessus conduit naturellement à poser dans le cas discret :

$$h(k) = P(T = k \mid T > k-1) = \frac{p_k}{S(k-1)}.$$

La fonction de hasard au point *k* s'interprète donc comme le taux de décès à l'âge *k*. De l'expression ci-dessus on tire que $1 - h(k) = \frac{S(k)}{S(k-1)}$, puis, par récurrence :

[3] Par définition *X* est stochastiquement plus grande que *Y* si $S_X(t) \ge S_Y(t)$.

$$S(k) = \prod_{m=1}^{k} (1 - h(m)).$$

La fonction de survie s'assimile donc aux (L_x) d'une table de mortalité, la fonction de hasard s'assimilant quant à elle aux (q_x).

On notera qu'à toute représentation continue d'une durée de vie *T*, on peut associer une représentation discrète en posant $X = k, k \leq T < k+1$ (autrement dit $X = [T]$).

En pratique toutefois la problématique est en général inverse : on estime une loi discrète et on veut ensuite calculer les taux de décès à n'importe quel âge. Il est pour cela indispensable de formuler une hypothèse qui permette de passer d'une expression discrète de la loi à une expression continue ; trois approches sont classiquement utilisées :

- La linéarisation de la fonction de survie, qui revient à supposer une répartition uniforme des sorties sur $[k, k+1[$ (hypothèse DUD) ;
- L'hypothèse de constance de la fonction de hasard sur $[k, k+1[$, qui conduit à une forme exponentielle ;
- L'« hypothèse de Balducci », qui conduit à une forme hyperbolique.

Ces 3 approches sont résumées dans le tableau ci-dessous[4] :

Tab. 2. Synthèse des interpolations

Fonction	Linéaire (DUD)	Exponentielle (Force constante)	Hyperbolique (Balducci)
l_{x+t}	$l_x - t \cdot d_x$	$l_x \cdot (p_x)^t = (l_{x+1})^t \cdot (l_x)^{1-t}$	$\left[(1-t) \cdot \frac{1}{l_x} + t \cdot \frac{1}{l_{x+1}}\right]^{-1}$
${}_t p_x$	$1 - t \cdot q_x$	$(p_x)^t = e^{-\mu \cdot t}$	$\frac{px}{1-(1-t) \cdot q_x}$
${}_t q_x$	$t \cdot q_x$	$1-(1-q_x)^t$	$\frac{t \cdot q_x}{1-(1-t) \cdot q_x}$
${}_{1-t} p_{x+t}$	$\frac{p_x}{1-t \cdot q_x}$	$(p_x)^{1-t} = e^{-\mu \cdot (1-t)}$	$1-(1-t) \cdot q_x$
${}_{1-t} q_{x+t}$	$\frac{(1-t) \cdot q_x}{1-t \cdot q_x}$	$1-(1-q_x)^{1-t}$	$(1-t) \cdot q_x$
μ_{x+t}	$\frac{q_x}{1-t \cdot q_x}$	$\mu = -\ln p_x$	$\frac{q_x}{1-(1-t) \cdot q_x}$
${}_t p_x \cdot \mu_{x+t}$	q_x	$\mu \cdot e^{-\mu \cdot t}$	$\frac{q_x \cdot (1-q_x)}{[1-(1-t) \cdot q_x]^2}$
L_x	$l_x - \frac{1}{2} \cdot d_x = l_{x+1} + \frac{1}{2} \cdot d_x$	$\frac{d_x}{\mu}$	$-l_{x+1} \cdot \frac{\ln p_x}{q_x}$
m_x	$\frac{q_x}{1-\frac{1}{2} \cdot q_x}$	μ	$\frac{(q_x)^2}{-p_x \cdot \ln p_x}$

[4] Tableau extrait de LANGMEIER [2000].

Dans la suite on utilisera en général la forme exponentielle (force constante). Dans certains cas particuliers, notamment paramétriques, il n'est pas nécessaire de formuler une hypothèse, le modèle impose la forme à retenir (*cf.* le modèle de Makeham par exemple).

3. Les lois paramétriques usuelles

On ne reprend ci-après que les modèles les plus courants ; d'une manière générale, toutes les distributions utilisées pour modéliser des variables positives (log-normale, Pareto, logistique, etc.) peuvent être utilisées dans des modèles de survie[5]. Toutefois, la distribution de base des modèles paramétriques de durée est la distribution exponentielle, et ses diverses généralisations, pour des raisons qui seront développées *infra*[6].

Le choix du modèle détermine en particulier la forme de la fonction de hasard ; on distinguera notamment les modèles à fonction de hasard monotone des modèles permettant d'obtenir des fonctions de hasard « en cloche » ou en « U »; ces derniers modèles sont peu usités en assurance, la situation de référence étant un taux de hasard croissant (au sens large) avec le temps.

3.1. Le modèle exponentiel

La spécification la plus simple consiste à poser $h(t)=\lambda$, avec $\lambda>0$. On en déduit immédiatement que $S(t)=e^{-\lambda t}$.

Le modèle exponentiel est caractérisé par le fait que les fonctions de survie conditionnelles $\{S_u(.), u>0\}$ sont exponentielles de même paramètre, $\lambda>0$. Cela signifie que le comportement de la variable aléatoire *T* après l'instant *u* de dépend pas de ce qui est survenu jusqu'en *u*. Il est également caractérisé par le fait que la fonction de survie est multiplicative, au sens où $S(u+t)=S(u)S(t)$. Ces propriétés découlent aisément de l'expression de la fonction de survie conditionnelle présentée en 2.2 ci-dessus.

On vérifie aisément par un calcul direct que $E(T)=\frac{1}{\lambda}$ et $V(T)=\frac{1}{\lambda^2}$. L'estimation du paramètre λ est classique, à partir de l'expression $L(\lambda)=\lambda^n \mathbf{exp}\left(-\lambda\sum_{i=1}^{n} T_i\right)$ qui conduit facilement à $\hat{\lambda}=\frac{n}{\sum_{i=1}^{n} T_i}=\frac{1}{\overline{T}}$.

3.2. Le modèle de Weibull

On suppose ici que la fonction de hasard est de la forme :

[5] Pour les propriétés des distributions usuelles, voir par exemple PARTRAT et BESSON [2004].

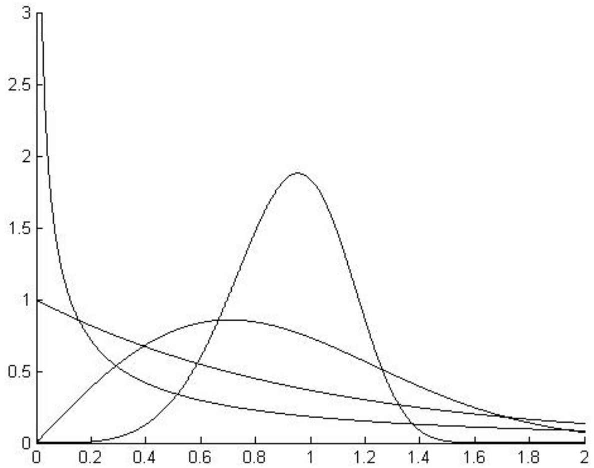
[6] Voir les notes de cours « Processus poissoniens et files d'attente ».

$$h(t) = \lambda \alpha t^{\alpha-1},\ \alpha, \lambda > 0$$

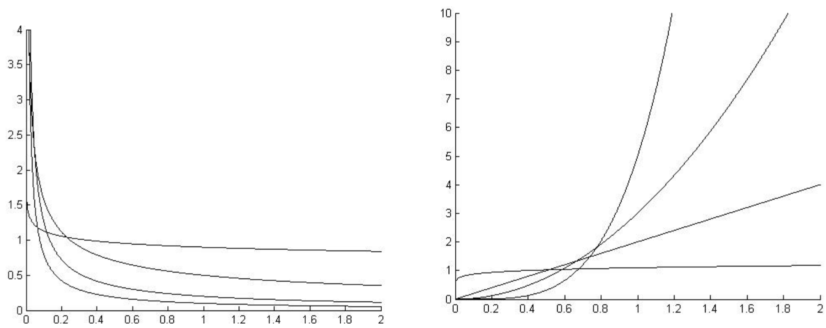
λ est un paramètre d'échelle et α un paramètre de forme. Il s'agit d'une généralisation simple du modèle exponentiel, permettant d'obtenir des fonctions de hasard croissantes avec *t* si $\alpha > 1$(il y a alors « usure ») et décroissantes avec *t* si $\alpha < 1$ (il y a « rodage »). Lorsque $\alpha = 2$ et $\lambda = \frac{1}{2}$ ce modèle porte le nom de « modèle de RAYLEIGH » ; il est utilisé en physique pour modéliser la durée de vie de certaines particules ou le bruit en sortie de certains récepteurs de transmissions[7].

La distribution de *T* est alors la distribution de Weibull $W(\alpha, \lambda)$, dont la fonction de survie s'écrit $S(t) = e^{-\lambda t^{\alpha}}$, $t > 0$. On peut notamment remarquer que si la variable *T* est distribuée selon une loi exponentielle de paramètre $\lambda > 0$ alors $T^{1/\alpha}$ suit $W(\alpha, \lambda)$[8].

En fonction de $\alpha > 0$ on peut obtenir des formes très différentes de la densité[9] :



La fonction de hasard est quant à elle monotone, avec l'allure suivante :



$\alpha < 1$ $\qquad\qquad$ $\alpha > 1$

[7] La loi de Rayleigh est également celle de la norme d'un vecteur gaussien centré réduit.
[8] Cela donne une méthode simple pour simuler des réalisations de la loi de Weibull.
[9] Le graphe est construit avec $\lambda = 1$.

Les moments s'obtiennent en observant que $E(T^k) = \lambda^{-k/\alpha}\Gamma\left(\frac{k}{\alpha}+1\right)$ avec $\Gamma(x) = \int_0^{+\infty} u^{x-1}e^{-u}du$. Pour démontrer cette égalité on écrit la densité de la loi de Weibull en utilisant $f(t) = S(t)h(t)$, ce qui donne $f(t) = \lambda\alpha t^{\alpha-1}\mathbf{exp}(-\lambda t^{\alpha})$. On en tire que :

$$E(T^k) = \int_0^{+\infty} \lambda\alpha t^{k+\alpha-1}\mathbf{exp}(-\lambda t^{\alpha})dt$$

Le changement de variable $u = \lambda t^{\alpha}$ permet de conclure. On a donc en particulier :

$$E(T) = \lambda^{-1/\alpha}\Gamma\left(\frac{\alpha+1}{\alpha}\right) \text{ et } V(T) = \lambda^{-2/\alpha}\left(\Gamma\left(\frac{2}{\alpha}+1\right) - \Gamma^2\left(\frac{1}{\alpha}+1\right)\right)$$

L'expression de la variance est la conséquence directe de $V(T) = E(T^2) - E(T)^2$. On déduit de ces expressions une propriété remarquable de la loi de Weibull, qui est que le coefficient de variation $\frac{\sigma(T)}{E(T)}$ ne dépend pas du facteur d'échelle λ.

Si on pose $X = \mathbf{ln}(T)$ alors $P(X \le x) = P(T \le e^x) = 1 - \exp(-\lambda e^{\alpha x})$, ce que l'on peut écrire :

$$P(X \le x) = 1 - \left(\mathbf{exp}\left(-\mathbf{exp}\left(\frac{x-\mu}{\sigma}\right)\right)\right)$$

en posant $\mu = \frac{1}{\alpha}\ln\left(1/\lambda\right)$ et $\sigma = 1/\alpha$. On reconnaît la loi de GUMBEL (ou double exponentielle), qui est l'une des 3 lois possibles comme loi limite du maximum d'un échantillon iid[10].

La loi de Weibull apparaît naturellement dans l'étude de la distribution limite du minimum d'un échantillon iid. En effet, si $(X_1,\ldots,X_n)$ est un échantillon d'une loi de fonction de répartition G sur $]0,+\infty[$ dont le comportement à l'origine vérifie :

$$\lim_{x\to 0^+} \frac{G(x)}{\lambda x^{\alpha}} = 1$$

alors $n^{1/\alpha}X_{(1)}$ converge en loi lorsque *n* tend vers l'infini vers une distribution $W(\alpha,\lambda)$.

<u>Démonstration</u> : On a $P\left(n^{1/\alpha}X_{(1)} > x\right) = \left[1 - G\left(\frac{x}{n^{1/\alpha}}\right)\right]^n$ et donc :

[10] Avec les distributions de Fréchet et Weibull; voir PLANCHET et *al.* [2011].

$$\ln\left(P\left(n^{1/\alpha}X_{(1)} > x\right)\right) = n\ln\left[1-G\left(\frac{x}{n^{1/\alpha}}\right)\right]$$

$$= n\ln\left[1-\lambda\left(\frac{x}{n^{1/\alpha}}\right)^{\alpha}+o\left(\frac{1}{n}\right)\right] = n\left[-\lambda\left(\frac{x}{n^{1/\alpha}}\right)^{\alpha}+o\left(\frac{1}{n}\right)\right] = -\lambda x^{\alpha}+o(1)$$

d'où l'on tire que $\lim_{n\to\infty} P\left(n^{1/\alpha}X_{(1)} > x\right) = e^{-\lambda x^{\alpha}}$, ce qui achève la démonstration.

En fait, cette propriété est à l'origine de la forme de la loi proposée par W. WEIBULL dans son article de 1951. Il se propose en effet de résoudre des problèmes de rigidité des matériaux. L'exemple qu'il utilise pour illustrer sa distribution est celui d'une chaîne. Comment peut-on établir la probabilité qu'une chaîne se brise ? Son raisonnement est que la chaîne va se briser si le plus faible des maillons se brise. Ceci revient donc à trouver la distribution du minimum d'un grand nombre d'objets. En théorie des valeurs extrêmes, on établit que la distribution du minimum ne dépend pas de la fonction de probabilité de chaque objet si le nombre d'objets est suffisamment grand (GALAMBOS [1978], GUMBEL [1958]). LOGAN [1992] a utilisé cette distribution dans l'optique d'une course (race model). Imaginons par exemple un grand nombre de neurones en compétition pour émettre un signal. Le signal émis sera produit par le neurone le plus rapide.

On peut enfin observer que comme $S(t) = e^{-\lambda t^{\alpha}}$, on a $\ln(-\ln(S(t))) = \ln(\lambda) + \alpha\ln(t)$; si $\hat{S}(t)$ désigne un estimateur non paramétrique de la fonction de survie empirique, les points $\left(\ln(t), \ln\left(-\ln\left(\hat{S}(t)\right)\right)\right)$ doivent donc être approximativement alignés. Cela fournit un moyen simple de vérifier si des données de durées peuvent être modélisées par une loi de Weibull.

On utilise parfois une paramétrisation différente de la loi de Weibull en posant $S(x) = \exp\left\{-\left(\frac{x}{l}\right)^{\alpha}\right\}$, ce qui revient à faire le changement de paramètre $\lambda = l^{-\alpha}$. Cela revient également à faire le changement de variable $y = \frac{x}{l}$ et donc à modifier l'unité de temps utilisée.

L'estimation des paramètres du modèle se fait en observant que la vraisemblance $L(\alpha, l) = \prod_{i=1}^{n} f(t_i)$ s'écrit :

$$L(\alpha, l) \propto \left(\frac{\alpha}{l^{\alpha}}\right)^{n} \prod_{i=1}^{n} t_i^{(\alpha-1)} \exp\left\{-\left(\frac{t_i}{l}\right)^{\alpha}\right\}$$

$$L(\alpha, l) \propto \left(\frac{\alpha}{l^{\alpha}}\right)^{n} \exp\left\{-l^{-\alpha}\sum_{i=1}^{n} t_i^{\alpha}\right\} \exp\left\{(\alpha-1)\sum_{i=1}^{n}\ln t_i\right\}$$

On en déduit l'expression suivante de la log-vraisemblance :

$$\ln L(\alpha,l) = \ln k + n(\ln \alpha - \alpha \ln l) - l^{-\alpha} \sum_{i=1}^{n} t_i^{\alpha} + (\alpha - 1) \sum_{i=1}^{n} \ln t_i$$

Les équations aux dérivés partielles s'écrivent donc :

$$\begin{cases} \dfrac{\partial}{\partial l} \ln L(\alpha,l) = -\dfrac{n\alpha}{l} + \alpha l^{-\alpha-1} \displaystyle\sum_{i=1}^{n} t_i^{\alpha} \\ \dfrac{\partial}{\partial \alpha} \ln L(\alpha,l) = n\left(\dfrac{1}{\alpha} - \ln l\right) + l^{-\alpha}\left(\ln l \displaystyle\sum_{i=1}^{n} t_i^{\alpha} - \sum_{i=1}^{n} t_i^{\alpha} \ln t_i\right) + \sum_{i=1}^{n} \ln t_i \end{cases}$$

On cherche donc les solutions du système suivant :

$$\begin{cases} l = \left(\dfrac{1}{n} \displaystyle\sum_{i=1}^{n} t_i^{\alpha}\right)^{1/\alpha} \\ \dfrac{1}{\alpha} = \dfrac{\displaystyle\sum_{i=1}^{n} t_i^{\alpha} \ln t_i}{\displaystyle\sum_{i=1}^{n} t_i^{\alpha}} - \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} \ln t_i \end{cases}$$

La deuxième équation peut être résolue numériquement par un algorithme de type Newton-Raphson qui converge vers $\hat{\alpha}$ pour autant qu'on lui fournisse une valeur initiale pas trop éloignée. Ainsi en notant $\varphi(\alpha) = \dfrac{\sum_{i=1}^{n} t_i^{\alpha} \ln t_i}{\sum_{i=1}^{n} t_i^{\alpha}} - \dfrac{1}{n} \sum_{i=1}^{n} \ln t_i - \dfrac{1}{\alpha}$, on utilisera la relation de récurrence :

$$\alpha_{i+1} = \alpha_i - \frac{\varphi(\alpha_i)}{\varphi'(\alpha_i)}$$

En pratique, cette valeur pourra être l'estimateur obtenu par la méthode des quantiles sur l'ensemble des observations complètes en observant que :

$$\breve{\alpha} = \frac{\ln\left[\dfrac{-\ln(1-p_2)}{-\ln(1-p_1)}\right]}{\ln(Q_{p_2}) - \ln(Q_{p_1})}$$

avec $Q_p = F^{-1}(p)$ la fonction quantile au point *p*. On rappelle que toute fonction de répartition admet une fonction inverse généralisée définie par :

$$F^{-1}(p) = \inf\{x\,; F(x) \geq p\}.$$

Dans le cas de la loi de Weibull, on vérifie aisément que :

$$F^{-1}(p) = l\left[-\ln(1-p)\right]^{1/\alpha}$$

Une fois $\hat{\alpha}$ obtenu, $\hat{l}$ s'en déduit grâce à la première équation.

3.3. Le modèle Gamma

Le modèle Gamma est une autre généralisation naturelle du modèle exponentiel : supposons que la durée T_r soit la durée d'attente de la réalisation d'un service dans une file d'attente et que la file d'attente soit composée de *r* serveurs indépendants et identiques qui traitent chacun une partie du service (ils sont donc montés en série). On fait l'hypothèse que la durée de réalisation du traitement de chacun des serveurs est une loi exponentielle de paramètre $\lambda > 0$.

Alors la durée globale de service est la somme de *r* variables exponentielles de même paramètre ; on en déduit que la durée de service est distribuée selon une loi Gamma de paramètre (r, λ) :

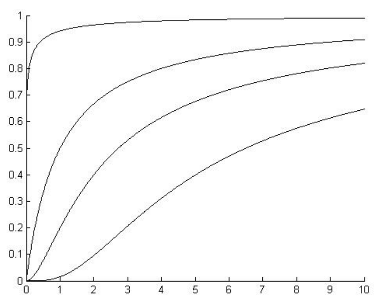
$$S_r(t) = \int_t^{+\infty} \frac{\lambda^r u^{r-1}}{(r-1)!} e^{-\lambda u} du .$$

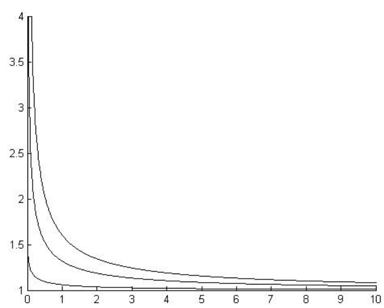
<u>Démonstration</u> : on utilise le fait que si $L(a) = E\left(e^{-aT}\right)$ désigne la transformée de Laplace au point *a* d'une loi exponentielle on a $L(a) = \frac{\lambda}{\lambda + a}$ et donc la transformée de Laplace de T_r est égale à $L_{T_r}(a) = \left(\frac{\lambda}{\lambda + a}\right)^r$; on reconnaît la transformée de Laplace au point *a* d'une loi Gamma[11].

Cette loi s'appelle, lorsque r est entier, la loi d'Erlang ; on peut définir de même un modèle de durée avec une loi Gamma dont le paramètre *r* n'est pas entier[12]. On a l'expression suivante pour la fonction de hasard :

$$h(t) = \frac{t^{r-1} e^{-\lambda t}}{\int_t^{+\infty} x^{r-1} e^{-\lambda x} dx}$$

Le sens de variation de cette fonction est déterminé par la position de *r* par rapport à 1 :



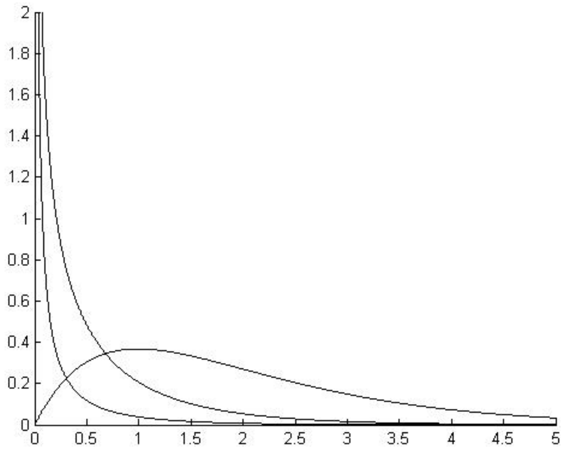


[11] Cela se vérifie aisément par un changement de variable dans l'intégrale.

[12] Avec $r = \frac{n}{2}$ et $\lambda = \frac{1}{2}$ on obtient la loi du Khi-deux à *n* degrés de liberté.

$r > 1$ $r < 1$

Les graphes ci-dessus mettent notamment en évidence le fait que la loi Gamma n'est *a priori* pas adaptée pour la modélisation de la mortalité humaine. La décroissance très rapide du taux de sortie lorsque $r < 1$ peut en revanche s'avérer en phase avec le comportement du maintien en arrêt de travail. L'allure de cette distribution est déterminée par la valeur de *r* ; en fonction de différentes valeurs de *r* on obtient le graphe ci-dessous[13] :



L'espérance et la variance d'une loi Gamma sont données par :

$$E(T) = \frac{r}{\lambda} \text{ et } V(T) = \frac{r}{\lambda^2}.$$

On déduit de ces expressions que le coefficient de variation d'une distribution gamma est :

$$cv = \frac{\sigma(T)}{E(T)} = \frac{1}{\sqrt{r}}.$$

On peut ainsi obtenir très simplement une estimation grossière du paramètre de forme *r* en calculant l'inverse du carré du coefficient de variation.

On peut également vérifier que la fonction de hasard $h_{r,\lambda}$ est croissante si $r > 1$ et décroissante si $r < 1$; de plus $\lim_{t \to +\infty} h_{r,\lambda}(t) = \lambda$, ce qui signifie qu'asymptotiquement on retrouve le modèle exponentiel.

<u>Démonstration</u> : en effectuant le changement de variable *u=x-t* dans l'expression de l'inverse la fonction de hasard, on met celle-ci sous la forme :

$$\frac{1}{h(t)} = \int_0^{+\infty} g(t,u) e^{-\lambda u} du$$

[13] Le graphe représente la densité de la loi Gamma.

avec $g(t,u) = \left(1+\frac{u}{t}\right)^{t-1}$. Le résultat découle de l'étude du signe de $\frac{\partial g}{\partial t}(t,u)$.

3.4. Le modèle de Gompertz-Makeham

Il s'agit du modèle de référence pour la construction de tables de mortalité et, dans une moindre mesure, de tables de maintien en arrêt de travail. Il est défini par la fonction de hasard suivante :

$$h(t) = \alpha + \beta \times \gamma^t$$

En démographie, la forme de cette fonction s'interprète de la manière suivante : le paramètre α représente un taux de décès accidentel (indépendant de l'âge), le terme en $\beta \times \gamma^t$ modélise quant à lui un vieillissement exponentiel (si $\gamma > 1$). Incidemment on retrouve le modèle exponentiel si $\beta = 0$. Par rapport à d'autres modèles, la fonction de Makeham a donc une ambition « explicative », ou « physique », en intégrant explicitement deux causes de décès clairement identifiées.

De manière plus précise, si on considère que le décès peut survenir de deux causes « concurrentes », l'accident et le vieillissement, la date de décès est de la forme $T = T_A \wedge T_V$, T_A (resp. T_V) représentant le décès accidentel (resp. dû au vieillissement). On suppose le décès accidentel modélisé par une loi exponentielle de paramètre *a*, et le décès associé au vieillissement modélisé par la fonction de hasard de Gompertz $h(t) = \beta \times \gamma^t$; alors *T* suit une loi de Makeham. Cela découle immédiatement du fait que la fonction de survie de *T* est le produit des fonctions de survies de T_A et T_V, et donc les fonctions de hasard s'ajoutent.

Un calcul direct conduit aisément à l'expression de la fonction de survie :

$$S_\theta(t) = \mathbf{exp}\left\{-\alpha t - \frac{\beta}{\mathbf{ln}(\gamma)}\left(\gamma^t - 1\right)\right\}.$$

Le calcul de l'espérance de *T* est en revanche complexe : $E(T) = \int_0^{+\infty} e^{-\alpha t - \frac{\beta}{\mathbf{ln}(\gamma)}\left(\gamma^t - 1\right)} dt$. Mais $S(t) = e^{\frac{\beta}{\mathbf{ln}(\gamma)}} \times e^{-\alpha t} \times e^{-\frac{\beta}{\mathbf{ln}(\gamma)}\gamma^t}$; on effectue alors le changement de variable :

$$u = \frac{\beta}{\mathbf{ln}(\gamma)}\gamma^t = \frac{\beta}{\mathbf{ln}(\gamma)} e^{t \times \mathbf{ln}(\gamma)}, \quad \frac{du}{\mathbf{ln}(\gamma) u} = dt$$

qui implique $\left(\frac{\mathbf{ln}(\gamma)}{\beta} u\right)^{1/\mathbf{ln}(\gamma)} = e^t$ puis :

$$E(T_\theta) = e^{\frac{\beta}{\ln(\gamma)}} \times \int_{\frac{\beta}{\ln(\gamma)}}^{+\infty} \left(\frac{\ln(\gamma)}{\beta} u \right)^{-\alpha/\ln(\gamma)} \times e^{-u} \frac{du}{\ln(\gamma) \times u}$$

$$= \frac{1}{\ln(\gamma)} \left(\frac{\ln(\gamma)}{\beta} \right)^{-\alpha/\ln(\gamma)} e^{\frac{\beta}{\ln(\gamma)}} \times \int_{\frac{\beta}{\ln(\gamma)}}^{+\infty} u^{-(1+\alpha/\ln(\gamma))} e^{-u} du$$

Avec le changement de variable $v = u - \frac{\beta}{\ln(\gamma)}$ on trouve

$$E(T) = \frac{1}{\beta} \times \int_0^{+\infty} \left(\frac{\ln(\gamma)}{\beta} v + 1 \right)^{-(1+\alpha/\ln(\gamma))} \times e^{-v} dv$$

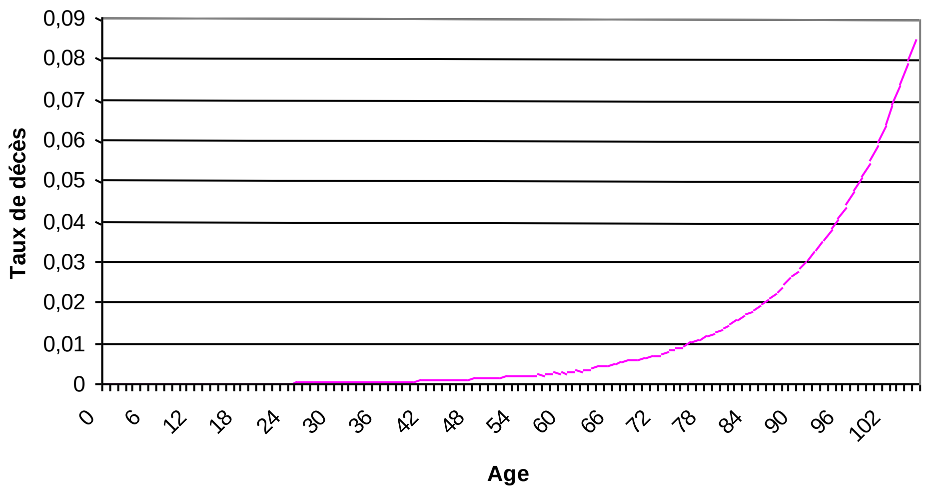
L'expression ci-dessus est complexe et on peut utiliser l'expression simplifiée suivante :

$$E_\theta(T) \approx e(\alpha, \beta, \gamma) = \sum_{t>0} \exp\left\{ -\alpha t - \frac{\beta}{\ln(\gamma)} \left(\gamma^t - 1 \right) \right\}.$$

Avec les valeurs « standards » des paramètres utilisés en mortalité humaine :

α	β	γ
8,81E-06	3,83E-05	1,076207

On trouve l'allure suivante des taux de hasard en fonction de l'âge :



On peut noter graphiquement la croissance plus rapide du taux instantané de décès avec l'âge que dans le cas d'une loi de Weibull (*cf.* 3.2 ci-dessus), qui est en général mieux adapté à la mortalité humaine.

On peut enfin observer que ce modèle possède une propriété géométrique permettant, comme dans le cas d'un modèle de Weibull, de valider graphiquement son adéquation aux données. En effet, en posant $s = \mathbf{exp}(-\alpha)$ et $g = \mathbf{exp}\left(-\frac{\beta}{\mathbf{ln}(\gamma)}\right)$ et en observant que $-q_x \approx \mathbf{ln}(1-q_x) = \mathbf{ln}(s) + \gamma^x(\gamma-1)\mathbf{ln}(g)$, on obtient que :

$$\mathbf{ln}(q_{x+1} - q_x) \approx x\,\mathbf{ln}(\gamma) + \mathbf{ln}\left((\gamma-1)^2\,\mathbf{ln}(g)\right).$$

Sous l'hypothèse que les taux de mortalité suivent une loi de Makeham, les points $(x, y = \mathbf{ln}(q_{x+1} - q_x))$ sont donc alignés sur une droite de pente $\mathbf{ln}(\gamma)$. L'utilisation pratique de cette remarque sera développée ultérieurement.

4. Les modèles composites

L'objet de cette section est de décrire les principales caractéristiques des modèles de base couramment utilisés dans un cadre paramétrique ou semi-paramétrique, et faisant appel à un degré de sophistication supérieur à la simple analyse d'un échantillon iid de loi paramétrique fixée *a priori*. Il s'agit de modèles que l'on rencontre en général lorsque l'on est confronté à une population hétérogène, composées d'individus avec des lois de survie différentes ; on a donc choisi de désigner ces modèles sous le nom générique de « modèles composites », et ils diffèrent par la manière dont l'hétérogénéité est prise en compte.

Les modèles purement non paramétriques seront étudiés par ailleurs ; ils ne sont pas évoqués ici.

4.1. Les mélanges de lois

4.1.1. Exemple introductif

On considère un système composé de deux éléments indépendants montés en parallèle, chacun des éléments ayant une durée de vie de loi exponentielle, avec des paramètres λ_1 et λ_2.

La durée de vie de l'équipement est mesurée par $T = T_1 \vee T_2$; la loi de *T* s'obtient facilement en observant que $1 - S(t) = \left(1 - e^{-\lambda_1 t}\right)\left(1 - e^{-\lambda_2 t}\right)$. On en déduit que dans le cas général la fonction de hasard est d'abord croissante, puis décroissante ; si $\lambda_1 = \lambda_2$, la fonction de hasard est croissante. L'indépendance temporelle est donc une propriété peu stable et elle se perd rapidement.

On va voir qu'elle se perd également dans le cas de l'agrégation de lois.

4.1.2. Agrégation de lois

Il arrive souvent en pratique que les durées que l'on observe résultent de l'agrégation de sous-populations ayant chacune un comportement spécifique, souvent inobservable. On parle alors d'hétérogénéité.

On suppose ici que la fonction de survie dépend d'un paramètre aléatoire ν, ce paramètre étant distribué selon une loi π. D'un point de vue heuristique, on se trouve en présence de sous-populations à l'intérieur desquelles la loi de survie est homogène et décrite par la loi de survie conditionnelle au fait que la valeur du paramètre soit ν, $S(t,\nu)$, la loi π décrivant le poids respectif de chaque sous-population dans la population totale.

On a donc la forme suivante pour la fonction de survie initiale de la population totale : $S(t)=\int S(t,\nu)\pi(d\nu)$:

$$S(t)=P(T>t)=E_{\nu}\left[P(T>t|\nu)\right]=\int S(t,\nu)\pi(d\nu).$$

La distribution d'hétérogénéité dépend *a priori* de t, puisque les individus des différentes sous-populations ne sortent pas du groupe à la même vitesse. À la date t, et en supposant la taille de la population infinie, on a ainsi :

$$\pi_t(d\nu)=\frac{S(t,\nu)}{S(t)}\pi(d\nu).$$

La fonction de hasard à la date t s'écrit alors $h(t)=\int h(t,\nu)\pi_t(d\nu)$. En effet, il suffit de remarquer que :

$$u^{-1}P(T\le t+u|T>t)=\int u^{-1}P(T\le t+u|T>t,\nu)\pi_t(d\nu),$$

puis de faire tendre u vers 0. Dans le cas particulier où $S(t,\nu)=\mathbf{exp}(-\lambda(\nu)t)$, c'est à dire où chaque sous-population est décrite par une loi exponentielle de paramètre $h(t,\nu)=\lambda(\nu)$, la fonction de survie agrégée s'écrit :

$$S(t)=\int_0^{+\infty}\mathbf{exp}(-\lambda(\nu)t)\pi(d\nu)$$

D'après l'expression ci-dessus de la fonction de hasard s'écrit donc $h(t)=\int \lambda(\nu)\pi_t(d\nu)$ et on en déduit que :

$$\frac{dh(t)}{dt}=-\int \lambda^2(\nu)\pi_t(d\nu)+\left(\int \lambda(\nu)\pi_t(d\nu)\right)^2.$$

En effet, de l'expression de $\pi_t(d\nu)=\frac{S(t,\nu)}{S(t)}\pi(d\nu)$ il découle :

$$\frac{\partial}{\partial t}\pi_t(d\nu)=\frac{\frac{\partial}{\partial t}S(t,\nu)\times S(t)-S(t,\nu)\times\frac{d}{dt}S(t)}{S(t)^2}\pi(d\nu)$$

avec $\frac{\partial}{\partial t} S(t,v) = -\lambda(v) S(t,v)$ et $\frac{S'(t)}{S(t)} = -h(t) = -\int \lambda(v) \pi_t(dv)$. On en déduit :

$$\frac{\partial}{\partial t}\pi_t(dv) = \frac{-\lambda(v)\times S(t,v)}{S(t)}\pi(dv) + \frac{S(t,v)\times h(t)}{S(t)}\pi(dv) = -\lambda(v)\pi_t(dv) + h(t)\pi_t(dv)$$

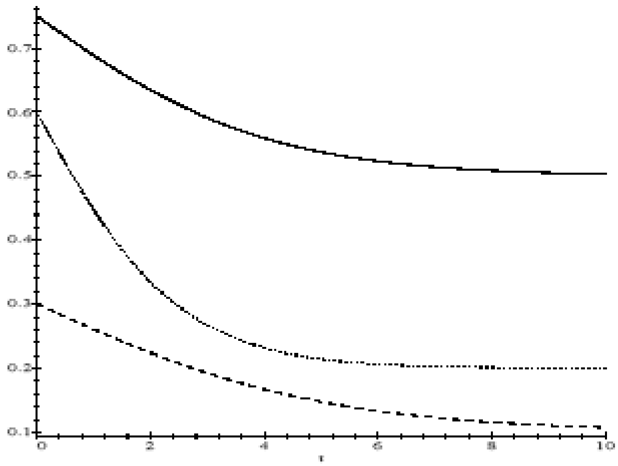
En écrivant $\frac{d}{dt}h(t) = \int \lambda(v) \frac{\partial}{\partial t}\pi_t(dv)$ on trouve donc finalement :

$$\frac{dh(t)}{dt} = -\int \lambda^2(v)\pi_t(dv) + h(t)^2$$

Ce qui est le résultat attendu. Cette égalité implique par l'inégalité de Schwarz (ou en remarquant que $\frac{dh(t)}{dt} = -V_{\pi_t}(\lambda(v))$) que $\frac{dh(t)}{dt} \leq 0$; l'agrégation de fonctions de hasard constantes conduit donc à une fonction de hasard globale décroissante. Ce phénomène s'explique par le fait que les individus ayant une valeur élevée de $\lambda(v)$ sortent en premier et il reste donc proportionnellement plus d'individus à $\lambda(v)$ faible lorsque le temps s'écoule. Le taux de sortie est donc logiquement décroissant. Ce phénomène porte le nom de « biais d'hétérogénéité », ou « mobile-stable ».

Exemple : mélange de 2 lois exponentielles

La durée est ici une variable exponentielle de paramètre λ_1 avec la probabilité *p* et λ_2 avec la probabilité 1-*p*, soit $S(t) = pe^{-\lambda_1 t} + (1-p)e^{-\lambda_2 t}$. La fonction de hasard a alors l'allure suivante :



On voit que le risque instantané peut être rapidement décroissant, alors même que les 2 fonctions d'origine sont à risque constant.

4.2. Les modèles à hasard proportionnel

Il s'agit d'un modèle semi-paramétrique dans lequel on se donne une fonction de survie de base, $B(t)$ et on fait l'hypothèse que la fonction de survie du phénomène observé est de la forme $S_\theta(t) = B(t)^\theta$, pour un paramètre $\theta > 0$ inconnu. Il est immédiat que la densité sous-jacente s'écrit $f_\theta(t) = \theta B(t)^{\theta-1} f(t)$, et la fonction de hasard est donc de la forme :

$$h_\theta(t) = \frac{f_\theta(t)}{S_\theta(t)} = \theta \frac{f(t)}{B(t)} = \theta h(t)$$

La fonction de hasard est ainsi proportionnelle à la fonction de hasard de base associée à $\theta = 1$, d'où la dénomination de « modèle à hasard proportionnel ». Le modèle exponentiel constitue un cas particulier de modèle à hasard proportionnel dans lequel la fonction de hasard de base est constante égale à l'unité.

On peut remarquer que ces modèles satisfont la propriété suivante : si la variable aléatoire T_θ est associée à la fonction de survie $S_\theta(t) = B(t)^\theta$, alors $E(T_\theta) = \int_0^{+\infty} S_\theta(t)dt = \int_0^{+\infty} B(t)^\theta dt$; or on reconnaît dans $\rho_\theta(T) = \int_0^{+\infty} B(t)^\theta dt$ la mesure de risque de Wang[14] associée à la fonction de distorsion $g_\theta(x) = x^\theta$ (appelée *PH-transform* de paramètre $\frac{1}{\theta}$).

En spécifiant différentes formes pour le coefficient de proportionnalité, on est conduit à définir différentes classes de modèles.

4.2.1. Le modèle de Cox

Ce modèle peut intégrer des variables explicatives utilisées pour définir le paramètre $\theta > 0$; pour cela on écrit $\theta = e^{z'\beta}$ avec $z = (z_1, \ldots, z_p)$ un vecteur de *p* variables explicatives et $\beta = (\beta_1, \ldots, \beta_p)$ le vecteur de paramètres ; avec cette formulation on a :

$$\ln h(t|Z = z) = \ln h(t) + \sum_{i=1}^{p} z_i \beta_i$$

et donc un modèle de régression linéaire. Ce modèle s'appelle le modèle de Cox. Il peut être appréhendé de deux manières différentes, selon que la fonction de hasard de base *h* est supposée connue (par exemple en supposant qu'il s'agit d'une d'un modèle de Weibull) ou qu'elle est inconnue. Dans ce dernier cas, elle devient un paramètre de nuisance de dimension infinie qui complique l'estimation des autres paramètres.

[14] Voir par exemple PLANCHET et *al.* [2011] pour une présentation plus générale des mesures de risque.

4.2.2. Les modèles de fragilité

Dans le modèle de Cox on cherche à modéliser l'effet de variables explicatives connues sur le niveau de la fonction de risque ; dans certaines situations, ces variables sont inobservables, et on souhaite tout de même évaluer les conséquences de ces variables inobservables sur la forme de la fonction de survie.

On repart de la formulation $S_\theta(t) = S(t|\theta) = B(t)^\theta$ ou, de manière équivalente, $h_\theta(t) = \theta h(t)$, d'un modèle à hasard proportionnel, et on considère que le paramètre θ est une variable aléatoire ; en d'autres termes on se donne la loi de survie conditionnelle au paramètre, et la loi globale s'obtient donc par intégration :

$$S(t) = E\left[B(t)^\theta\right]$$

l'espérance étant calculée par rapport à la loi de θ. Cette expression est analogue à l'expression $S(t) = \int S(t,v)\pi(dv)$ obtenue à la section 4.1.2. Le paramètre θ s'appelle la « fragilité ». Ces modèles sont également parfois appelés « modèles à effets aléatoires ».

Approche classique

Les modèles de fragilité ont été introduits par VAUPEL et al. [1979] pour rendre compte de l'hétérogénéité individuelle dans un contexte de mortalité. Le paramètre de fragilité permet en pratique d'introduire des différences de niveau de mortalité entre les individus, en supposant que l'évolution de la mortalité avec l'âge est identique pour tous les individus. L'hétérogénéité est alors modélisée *via* la distribution du paramètre θ. Dans VAUPEL et al. [1979] il est fait l'hypothèse d'une distribution $\gamma(r,\lambda)$:

$$\pi(d\theta) = f_{r,\lambda}(\theta) = \frac{\lambda^r \theta^{r-1}}{\Gamma(r)} \mathbf{exp}(-\lambda\theta)$$

que l'on choisit d'espérance 1, en imposant $r = \lambda$ et en considérant comme paramètre de contrôle la variance $\sigma^2 = \lambda^{-1}$. Dans ce cas, et pour une population observée depuis la naissance, on peut montrer que la fonction de hasard moyenne de la population à l'âge *t* est de la forme :

$$\bar{h}(t) = h(t) S(t)^{\sigma^2}$$

avec $S(t)$ la fonction de survie agrégée en t. Dans cette expression on a $\bar{h}(t) = \int h_\theta(t)\pi_t(d\theta) = h(t)\int \theta\pi_t(d\theta)$ avec $\pi_t(d\theta) = \frac{S(t,\theta)}{S(t)}\pi(d\theta)$. Par ailleurs, $S(t) = \int S(t,\theta)\pi(d\theta)$.

Ce modèle a été généralisé par BARBI [1999] qui a proposé, en supposant toujours une fragilité proportionnelle initialement distribuée selon une loi Gamma, un modèle d'hétérogénéité appelé « fragilité combinée », dans lequel en plus du paramètre θ, on se donne une distribution discrète τ indépendante de θ telle que :

$$h_{\theta,\tau}(t) = \theta h(t,\tau).$$

Cela revient à subdiviser la population initiale en sous-groupes chacun décrit, conditionnellement au facteur de fragilité proportionnel θ par une fonction de risque qui lui est propre. Ce modèle est notamment utilisé dans BARBI et *al.* [2003] pour étudier l'âge extrême de survie. Ces auteurs posent :

$$h(t,\tau_i) = a \times \mathbf{exp}(b_i \times x) + c$$

ce qui revient à faire l'hypothèse que la mortalité globale observée est un mélange de lois de Makeham (avec toujours l'hypothèse Gamma pour la distribution de fragilité proportionnelle). La fonction de risque agrégée est alors de la forme :

$$\bar{h}(x) = \sum \pi_i(x) h(x,\tau_i) s_{x|\tau_i}(x,\tau_i)^{\sigma^2}$$

avec $\pi_i(x)$ la proportion d'individus du groupe *i* survivant à l'âge x.

Approche alternative

Cette modélisation est également utile pour introduire de la dépendance entre différentes durées de vie. On suppose pour cela que les durées observées, $T_1,\ldots,T_n$ sont indépendantes conditionnellement à θ et que les marginales (conditionnelles) sont de la forme $S_i(t|\theta) = B_i(t)^\theta$; on en déduit directement l'expression de la fonction de survie conjointe :

$$S(t_1,\ldots,t_n) = E\left[\left(B_1(t_1)\ldots B_n(t_n)\right)^\theta\right]$$

Dans ce cas le paramètre de fragilité s'interprète comme un élément exogène qui modifie le comportement de l'ensemble des individus. En général la fonction de survie de base est identique pour tous les individus et on a

$$S(t_1,\ldots,t_n) = E\left[\prod_{i=1}^{n} B(t_i)^\theta\right]$$

Mais comme $B(t) = \mathbf{exp}(-H(t))$ où *H* est la fonction de hasard cumulée de référence, cette expression se met sous la forme :

$$S(t_1,\ldots,t_n) = E\left[\mathbf{exp}\left(-\theta\sum_{i=1}^{n} H(t_i)\right)\right]$$

On reconnaît, dans le membre de droite, la transformée de Laplace de la variable θ au point $\sum_{i=1}^{n} H(t_i)$. Lorsque θ est distribué selon une loi stable de paramètre α (c'est à dire que la transformée de Laplace de θ est $E\,\mathbf{exp}(-x\theta) = \mathbf{exp}(-x^\alpha)$) on obtient le modèle de

Hougaard (*cf.* HOUGAARD [2000]) avec la fonction de survie $S(t_1,\ldots,t_n) = \exp\left\{-\left[\sum_{i=1}^{n}\left(-\ln S(t_i)\right)^{1/\alpha}\right]^{\alpha}\right\}$; on peut remarquer que la loi conjointe étant de la forme $C(S_1,\ldots,S_n)$ on définit ainsi une copule, dite copule de Hougaard.

4.3. Les transformations croissantes de la durée (modèles AFT[15])

Il s'agit d'un modèle semi-paramétriques dans lequel on se donne une fonction de survie de base, $S(t)$ et on fait l'hypothèse que la fonction de survie du phénomène observé est de la forme $S_\theta(t) = S(\theta t)$, pour un paramètre $\theta > 0$. La fonction de hasard s'écrit ici :

$$h_\theta(t) = \frac{f_\theta(t)}{S_\theta(t)} = \theta \frac{f(\theta t)}{S(\theta t)} = \theta h(\theta t)$$

et cette expression ne se simplifie pas comme dans le cas du modèle à hasard proportionnel.

On peut toutefois remarquer que les deux approches sont équivalentes si et seulement si la fonction de hasard est de type Weibull[16].

Cette démarche peut être généralisée dès lors que l'on se donne une fonction croissante ψ_θ en considérant les fonctions de survie $S_\theta(t) = S\left(\psi_\theta^{-1}(t)\right)$; cela revient à étudier les variables $\psi_\theta(T)$, où *T* est la variable de base. La loi de Weibull en fournit un exemple avec $\psi_\alpha(t) = t^{1/\alpha}$ et une loi exponentielle (voir la section 3.2).

4.4. Les modèles à causes de sortie multiples

Dans certaines situations on est amené à distinguer entre différentes causes de sortie ; par exemple en décès on s'intéresse à la cause du décès, en arrêt de travail au motif de la sortie d'incapacité (retour au travail ou passage en invalidité), etc. C'est typiquement ce qu'on fait lorsqu'on interprète le modèle de Makeham (voir 3.4 ci-dessus)

Si on note $T_1,..,T_n$ les variables de durée associées à chacune des causes étudiées, la survie globale est simplement $T = T_1 \wedge .. \wedge T_n$; sous l'hypothèse d'indépendance des différentes composantes le modèle est simple et la fonction de hasard globale est la somme des fonctions de hasard. Mais l'hypothèse d'indépendance peut être parfois restrictive, et les modèles de fragilité fournissent un moyen simple de la relâcher. Cette approche a été proposée initialement par OAKES [1989].

On suppose donc que les durées associées à chaque cause, $T_1,\ldots,T_n$ sont indépendantes conditionnellement à θ et que les marginales (conditionnelles) sont de la forme

[15] AFT pour Accelerated Failure Time.

[16] Voir ce sujet d'examen pour une justification.

$S_i(t|\theta) = B_i(t)^\theta$. On est alors ramené à des calculs proches de la section 4.2.2 ci-dessus et on trouve :

$$S(t_1,\ldots,t_n) = E\left[\prod_{i=1}^{n} B_i(t_i)^\theta\right]$$

Exemple : avec deux causes de sortie distribuées chacune suivant une loi de Weibull et une distribution du paramètre de mélange selon une loi stable de paramètre *a*, on trouve $S(t) = \mathbf{exp}\left\{-\left(\lambda_1 t^{\alpha_1} + \lambda_2 t^{\alpha_2}\right)^a\right\}$, qui est une conséquence immédiate de $E(\mathbf{exp}(-x\theta)) = \mathbf{exp}(-x^a)$ et de l'expression de la fonction de survie de la loi de Weibull, $S(t) = \mathbf{exp}(-\lambda t^\alpha)$.

4.5. Les modèles à choc commun

L'idée est ici que la durée de survie dépend de deux facteurs, l'un propre à l'individu et l'autre affectant la population dans son ensemble. Ce second facteur peut être un facteur accidentel ou environnemental. On considère le modèle :

$$T_i = X_i \wedge Z$$

avec S_i la fonction de survie de X_i et S_Z la fonction de survie de *Z*. La loi conjointe du vecteur $(T_1,\ldots,T_n)$ s'obtient en observant que l'événement $\{X_i \wedge Z > t\}$ est égal à $\{X_i > t\} \cap \{Z > t\}$, ce qui conduit à :

$$S(t_1,\ldots,t_n) = \prod_{i=1}^{n} S_i(t_i) \times S_Z\left(\mathbf{max}(t_1,\ldots,t_n)\right).$$

MARSHALL et OLKIN [1967] proposent par exemple une distribution exponentielle pour *Z*.

5. Références

BARBI E.[1999] « Eterogeneità della popolazione e sopravvivenza umana : prospettive metodologiche ed applicazioni alle generazioni italiane 1870-1895 », Florence, Dipartimento Statistico (Ph.D. thesis), 91 p.

BARBI E., CASELLI G. et VALLIN J. [2003] « Hétérogénéité des générations et âge extrême de la vie », *Population* 2003/1, Vol. 58, p. 45-68.

BÖHMER P.E. [1912], « Theorie der unabhängigen Wahrscheinlichkeiten », Rapports, Mém. et Procès-verbaux 7e Congrès Int. Act. Amsterdam 2, 327–343.

CAIRNS A., BLAKE D., DOWD K. [2004], « *Pricing Frameworks for Securitization of Mortality Risk* », AFIR

COX D.R. [1972] « Regression models and life-tables (with discussion) ». *J. R. Statist. Soc. Ser. B*, pages 187-220.DROESBEKE J.J., FICHET B., TASSI P. [1989], *Analyse statistique des durées de vie.* Paris : Economica

GALAMBOS J. [1978] *The Asymptotic Theory of Extreme Order Statistics*. New York : John Wiley and Sons.

GUMBEL E. J. [1958] *The Statistics of Extremes*. New York: Columbia University Press.

HOUGAARD P. [2000] *Analysis of Multivariate Survival Data*, New York : Springer-Verlag.

KAPLAN E.L., MEIER P. [1958] « Non-parametric estimation from incomplete observations ». *Journal of the American Statistical Association*, 53, 457-481.

LANGMEIER R. [2000], « Etude de différentes méthodes d'ajustement de tables de mortalité : application aux données d'une compagnie d'assurance », HEC Lausanne

Le BRAS H. [2000] *Naissance de la mortalité*, Col. Hautes Études, Gallimard – Le Seuil

LEE, R.D., CARTER, L. [1992] « Modelling and forecasting the time series of US mortality ». *Journal of the American Statistical Association* 87,659-671.LOGAN G. D. [1992] « Attention and preattention in theories of automaticity ». *American Journal of Psychology*, 105, 317-339.

MAKEHAM W. [1860] « On the law of mortality. » Journal of the Institute of Actuaries, 13, 325-358.

MARSHALL AW., OLKIN, I. [1967] « A generalized bivariate exponential distribution ». *J. Appl. Probability* 4, 291-302

OAKES D. [1989] « Bivariate survival models induced by frailties ». *Journal of the American Statistical Association* 84, 487-493.

PARTRAT C., BESSON J.L. [2004], *Assurance non-vie - modélisation, simulation.* Paris : Economica

PLANCHET F., THÉROND P.E. [2011] *Modélisation statistique des phénomènes de durée – applications actuarielles*, Paris : Economica.

VAUPEL J. W., MANTON K., STALLARD E., 1979, « The impact of heterogeneity in individual frailty on the dynamics of mortality », *Demography*, 16, p. 439-454.

WEIBULL W. [1951] « A Statistical Distribution Function of Wide applicability ». *Journal of Applied Mechanic*, 18, p. 292-297.

6. Annexes : transformées de Laplace usuelles

Tab. 3. Table des transformées de Laplace

Fonctions		Transformée de Laplace
dérivée	$\frac{d}{dt}f(t)$	$s\widehat{f}(s) - f(0)$
dérivée k^{ieme}	$\frac{d^k}{dt^k}f(t)$	$s^k\widehat{f}(s) - s^{k-1}f(0) - s^{k-2}f'(0) - \cdots - f^{(k-1)}(0)$
intégrale	$\int_0^t f(x)dx$	$\frac{1}{s}\widehat{f}(s)$
dilatation $(\lambda > 0)$	$f(\lambda t)$	$\frac{1}{\lambda}\widehat{f}(s/\lambda)$
translation	$f(t - t_0)$	$\exp(-st_0)\widehat{f}(s)$
facteur polynomial	$(-t)^k f(t)$	$\frac{d^k}{ds^k}\widehat{f}(s)$
facteur exponentiel	$\exp(-\lambda t)f(t)$	$\widehat{f}(s + \lambda)$
Heaviside	$Y(t)$	$1/s$
Dirac	$\delta(t)$	1
exponentielle	$\lambda e^{-\lambda t}Y(t)$	$\frac{\lambda}{s + \lambda}$
gamma	$\frac{\lambda^\beta t^{\beta-1}e^{-\lambda t}}{\Gamma(\beta)}Y(t)$	$\frac{\lambda^\beta}{(s + \lambda)^\beta}$
sinus	$\sin(\omega t)Y(t)$	$\frac{\omega}{s^2 + \omega^2}$
cosinus	$\cos(\omega t)Y(t)$	$\frac{s}{s^2 + \omega^2}$