Étude probabiliste et statistique de quelques modèles principalement issus d'applications industrielles

Christian Paroissin

Mémoire d'habilitation à diriger les recherches

Étude probabiliste et statistique de quelques modèles principalement issus d'applications industrielles

Christian Paroissin

soutenu publiquement le 24 septembre 2015 devant le jury composé de :

Gilles Celeux	Directeur de Recherche, Inria Saclay
Anne Gégout-Petit	Professeure des Université, Université de Lorraine
Olivier Gaudoin	Professeur des Université, Grenoble INP - ENSIMAG
Ivan Kojadinovic	Professeur des Université, Université de Pau et des Pays de l'Adour
Bernard Ycart	Professeur des Université, Université Grenoble 1 - Joseph Fourrier

et après les avis de :

Gilles Celeux	Directeur de Recherche, Inria Saclay
Olivier Gaudoin	Professeur des Université, Grenoble INP - ENSIMAG
Gerardo Sanz	Profesor Titular, Universidad de Zaragoza

Remerciements

Tout d'abord, je remercie très sincèrement Gilles Celeux, Olivier Gaudoin et Gerardo Sanz d'avoir accepté d'être rapporteur pour mon mémoire d'habilitation et de s'être intéressés de près à tous mes travaux de recherche. Je remercie également Anne Gégout-Petit, Ivan Kojadinovic et Bernard Ycart d'avoir accepté de faire partie de ce jury. Je suis fier de ce jury qui réunit des personnes que j'admire et que j'apprécie fortement. Cette soutenance fut quelque peu épique grâce aux controleurs aériens... Merci à vous, membres du jury, d'avoir accepté de « venir » une seconde fois (mais rien n'est perdu, la fausse soutenance aura permis de démarrer une discussion avec Gilles et Jean-Christophe Turlot, à suivre donc...). Enfin, c'est l'occasion pour moi de renouveller encore une fois mes remerciements à Bernard pour m'avoir formé durant les années de thèse. Ce que je suis, c'est à toi que je le dois !

Je tiens à remercier également tous les collègues avec qui j'ai travaillé après ma thèse. A chaque fois, ce fut de belles aventures humaines, de belles expériences. Je ne vais pas pouvoir tous les citer ici, mais tous ont leur place ! J'aimerais néanmoins en remercier quelques uns tout particulièrement. Merci à Javiera Barrera et Thierry Huillet pour ces beaux travaux autour de l'heuristique move-to-front, des permutations biaisées par la taille et de la loi de Dirichlet. Merci à Laurent Bordes pour m'avoir proposé de co-encadrer/co-diriger trois étudiants en thèse. Je finirai donc par remercier mes trois étudiants pour avoir essuyer les platres comme on dit : dans l'ordre chronologique, Ali Salami, Aurélie Billon et Maïder Estécahandy.

Dix ans que je suis à Pau ! Je profite de ce mémoire pour saluer tous les collègues du Laboratoire de Mathématiques et de leurs Applications, en particulier ceux de l'équipe probabilité-statistique. Je voudrais remercier ici plus spécialement Sylvie Berton et Bruno Demoisy que j'embête bien trop souvent et qui me rendent de bien précieux services.

Pour terminer, un grand merci à Elise pour avoir encore relu ce que j'ai écrit (le prochain livre étant en anglais, elle évitera cette tâche pour une fois). Quel courage remarquable !

« Mais un ami, ce n'est pas quelqu'un qu'on voit, qu'on doit voir. Un ami c'est quelqu'un qu'on ne voit pas pendant dix ans mais, à chaque rencontre, on s'est vu hier, on se verra demain. On se tape dans le bide avec un sourire en se disant qu'on a grossi. On parle du futur plus que du passé. Parfois, on ne dit rien, on savoure l'instant. On se dit au revoir, confiant. On est amis. »

Blog d'un condamné, <http ://uncondamne.tumblr.com/>

A la mémoire de Béatrice Lachaud (ép. Blottière)...

Table des matières

Première partie

Modèles en fiabilité et problèmes connexes

Chapitre 1

Introduction

1.1 Problématiques générales

Comprendre le vieillissement d'un composant ou d'un système est un enjeu crucial pour un industriel, car cela lui permet de mieux maîtriser le(s) risque(s) associé(s) à son utilisation. L'élaboration et l'étude de modèles (tant les propriétés probabilistes que les problèmes d'inférence statistique) s'avèrent donc être un travail important. Deux grandes approches peuvent être distinguées : (1) les modèles de durée de vie ; (2) les modèles de dégradation. Le choix de l'une de ces deux approches dépend évidemment du contexte industriel.

Dans un modèle de durée de vie, on suppose que l'événement étudié (typiquement la panne du composant ou du système) survient au bout d'une durée qui est la réalisation d'une variable aléatoire positive (dans certaines situations, on peut aussi s'intéresser à la survenue d'un événement parmi plusieurs possibles - par exemple, un composant peut tomber en panne pour diverses causes : on parle alors de modèles à risques compétitifs). Des questions de nature probabiliste portent alors sur la caractérisation de la loi de cette variable aléatoire via les notions usuelles de vieillissement (IFR/DFR, NBU/NWU, etc.). Quant à l'inférence statistique, dans un cadre paramétrique, l'estimateur des paramètres de la loi dépend des données disponibles : on est rarement dans le cas classique de l'inférence où toutes les durées sont observées de manière exacte et pour certaines unités, on disposera d'une information partielle. Par exemple, on sait uniquement que l'événement se produira après une date donnée : on parle de durées censurées à droite. D'autres cas de censure (ou de troncature) sont également étudiés (censure à gauche ou par intervalle). Enfin, des covariables (ou facteurs environnementaux) peuvent influencer la durée de vie du composant/système. Ces éléments doivent être intégrés dans les modèles de durée de vie et plusieurs modèles paramétriques ou semi-paramétriques ont été proposés et étudiés (modèle de durée de vie accélérée, modèle à risques proportionnels, etc.). Enfin, on s'intéresse également à l'étude de politiques de remplacement des composants pour éviter des périodes d'indisponibilité, etc. Il s'agit de les optimiser afin de trouver un compromis entre un nombre trop élevé de remplacements et une éventuelle indisponibilité du composant.

A travers les grandes questions étudiées pour les modèles de durée de vie, on aura noté que l'observation de durées exactes est d'autant plus rare que le composant/système est "fiable" (dans le sens où la probabilité qu'il tombe en panne durant sa vie utile est faible). S'il est possible de mesurer le niveau de dégradation d'un composant/système, alors il peut être plus avantageux de modéliser la dégradation. Un autre avantage réside dans la possibilité de proposer des politiques de maintenance préventive plus fines. La mesure de dégradation peut être soit qualitative, soit quantitative. Dans le premier cas, on parle de modèles multi-états. Ces modèles sont également utilisés dans d'autres domaines, en particulier en biostatistique. Ce sont, en général, des processus markoviens ou semi-markoviens. Dans le second cas, on parle de modèle de dégradation continue et, la plupart du temps, des processus à accroissements indépendants (et, éventuellement, stationnaires) sont considérés. Dans tous les cas, il est possible de classer les problèmes étudiés en trois groupes : (1) l'inférence paramétrique ou semi-paramétrique des modèles ; (2) l'étude du temps d'atteinte d'un niveau de dégradation (considéré comme étant un seuil critique) vu comme le temps de panne du composant/système ; (3) l'étude d'une politique de maintenance ou d'un ensemble de politiques de maintenance (préventive et/ou corrective). Néanmoins, il possible de croiser les problèmes : étude de l'estimation du temps de panne (problèmes 1 et

2), estimation des paramètres du modèle de dégradation et maintenance optimale (problèmes 1 et 3), etc.

1.2 Contexte

Mes travaux de recherche dans ce domaine portent essentiellement sur les modèles de dégradation, que ce soit des modèles multi-états (chapitre 2) ou des modèles de dégradation continue (chapitre 3). La prise en compte des incertitudes est étudiée au chapitre 4 pour un modèle de dégradation, puis pour un modèle de durée de vie. Le chapitre 5 est consacrée à l'étude de problématiques liées aux durées de vie. Enfin, des perspectives sont présentées au chapitre 6.

J'ai travaillé avec Laurent Bordes (UPPA) via le co-encadrement de trois thèses, dont deux ont déjà été soutenues :

1. co-encadrement de la thèse d'Ali Salami (novembre 2007 - janvier 2011) [168], financée par une allocation de recherche du Conseil Régional d'Aquitaine. Actuellement, Ali est enseignant dans plusieurs universités au Liban ;
2. co-direction [1] de la thèse CIFRE d'Aurélie Billon (février 2009 - mai 2012) [64], en partenariat avec Turbomeca/Safran. Aurélie est actuellement ingénieure sûreté de fonctionnement (CDI) chez Safran Engineering Service.
3. co-direction [2] de la thèse CIFRE de Maïder Estécahandy (depuis mars 2013), en partenariat avec TOTAL. La thèse porte sur les méthodes de simulation d'événements rares pour les réseaux de Petri.

Enfin, les contrats industriels furent également des occasions de collaborer avec différents collègues palois. Il s'agit en général des problèmes liés à des données de dégradation, mais également sur des problèmes liés aux durées de vie.

Le tableau ci-dessous résume quantitativement mon activité de recherche sur ce thème.

Publications	[17, 18, 16, 6, 15, 5, 14, 7, 1]
Actes de congrès	[22, 20, 21, 24, 29, 30, 23, 27]
Pré-publications	[43]
Thèses soutenues co-encadrées	[168, 64]
Contrats industriels	Alstom EDF R&D Rivage Pro Tech Turbomeca TOTAL

1. La co-direction a été reconnue par le Président de l'Université de Pau et des Pays de l'Adour, le 3 avril 2012, sur proposition de la directrice de l'École Doctorale des Sciences Exactes et Applications de Pau (ED211) et après avis du Conseil Scientifique de l'établissement.

2. La co-direction a été reconnue par le Président de l'Université de Pau et des Pays de l'Adour, le 7 mars 2013, sur proposition de la directrice de l'École Doctorale des Sciences Exactes et Applications de Pau (ED211) et après avis du Conseil Scientifique de l'établissement.

Chapitre 2

Modèles de dégradation multi-états

Les modèles multi-états (markoviens ou semi-markoviens) sont utilisés pour des études en fiabilité lorsqu'on dispose de mesures qualitatives de la dégradation d'un composant (voir, par exemple, [112] et [113]). Ces modèles sont utilisés principalement en biostatistique et en épidémiologie [104, 79, 51], mais également dans d'autres domaines comme, par exemple, l'ingénierie [100, 138]. Dans [77], on trouvera une introduction claire et détaillée sur les processus markoviens et semi-markoviens avec leurs applications en fiabilité.

Trois parties distinctes constituent ce chapitre. La première partie correspond à mon travail de thèse (1999-2002), effectué sous la direction de B. Ycart (Université Paris 5, à l'époque). Il s'agissait d'étudier le comportement asymptotique de temps d'atteinte pour des systèmes monotones et redondants constitués de composants indépendants, identiques et réparables (l'asymptotique porte sur le nombre de composants). Les deux dernières parties portent sur des travaux dans le cadre de contrats industriels. Dans l'une des collaborations avec EDF R&D, on a considéré un composant dont la dégradation est décrite par une chaîne de Markov non-homogène, le caractère non-homogène étant lié au fait que les probabilités de transition peuvent dépendre de covariables qui évoluent dans le temps. La collaboration avec Turbomeca a porté sur la modélisation stochastique de dégradation de composant sur un turbomoteur (ces travaux s'inscrivent dans le cadre de la thèse CIFRE d'A. Billon [64]).

2.1 Systèmes markoviens redondants, monotones et réparables [18, 17]

On s'intéresse ici à un système constitué d'un grand nombre de composants indépendants et identiques. La qualité du fonctionnement de ces composants est mesurée sur une échelle ordinale : on parle de composants multi-états. Dans cette étude, nous formulons une hypothèse importante qui est la suivante : les composants du système sont partiellement interchangeables (on peut changer les positions de certains d'entre eux dans le système sans que cela ne change l'état du système global), ce qui traduit en général une certaine forme de redondance. En effet, il existe plusieurs systèmes de ce type parmi lesquels on peut citer : les systèmes k-sur-n, les systèmes k-consécutifs-sur-n, les systèmes série-parallèle et les systèmes parallèle-série. Le premier exemple est le cas extrême où tous les composants sont interchangeables. Les résultats obtenus sont des théorèmes asymptotiques portant soit sur la fiabilité du système, soit sur la disponibilité du système. On rappelle que la fiabilité $R(t)$ à l'instant t est la probabilité que le système ait toujours fonctionné entre 0 et t tandis que la disponibilité $A(t)$ à l'instant t est la probabilité que le système fonctionne à l'instant t. Il est immédiat de voir que, pour tout t, $R(t) \leqslant A(t)$ et que ces deux quantités coïncident si le système n'est pas réparable (i.e. si les composants ne sont pas réparables).

On note E l'espace d'états (discret) des composants. Un système à n composants est donc à valeurs dans E^n. Certains éléments de cet espace produit conduisent à un état de panne du système : on les notera par B_n. Le système est supposé être monotone ; en théorie des ensembles, ceci se traduit par la propriété de croissance de B_n :

Définition 2.1.1 Le sous-ensemble $B_n \subset E^n$ est dit croissant si et seulement si il vérifie :

$$(\eta \in B_n \quad \text{et} \quad \nu \preceq \eta) \Longrightarrow \nu \in B_n,$$

où $\preceq$ est une relation d'ordre partiel sur E^n.

Enfin, le caractère (partiellement ou totalement) échangeable des composants dans le système se traduit, quant à lui, par une propriété de symétrie de B_n basée sur les sous-groupes de permutation d'un ensemble à n éléments :

Définition 2.1.2 Le sous-ensemble $B_n \subset E^n$ est dit symétrique si et seulement si il existe un sous-groupe G de $\mathcal{S}_n$ (groupe des permutations) transitif sur $\{1, \ldots, n\}$ et tel que B_n soit invariant par l'action de G, i.e. tel que :

$$\forall g \in G\ , \forall \eta \in B_n\ , \quad g \cdot \eta = (\eta_{g^{-1}(1)}, \ldots, \eta_{g^{-1}(n)}) \in B_n.$$

On dit que B_n est totalement symétrique si et seulement si il est invariant par l'action de n'importe quelle permutation ($G = \mathcal{S}_n$).

Enfin, la dynamique de dégradation de chaque composant est modélisée par un processus markovien de sauts à valeurs dans E. De plus, on suppose que ceux-ci sont initialement dans l'état minimal $e_{min} \in E$ correspondant à l'état de parfait fonctionnement des composants (dont on supposera l'existence et l'unicité). On note $X_1, \ldots, X_n$ le n-échantillon de processus markoviens de sauts représentant l'état des composants (on rappelle que les composants du système sont supposés indépendants et identiques). Le générateur infinitésimal de ces processus est noté Λ et la loi de l'état d'un composant, sachant qu'il est initialement dans l'état e_{min}, $p(t)$.

On peut alors définir les deux grandeurs mentionnées précédemment : la disponibilité et la fiabilité. On rappelle que la disponibilité est la probabilité que le système soit en état de bon fonctionnement à un instant t donné :

$$A(t) = \mathbb{P}[(X_1(t), \ldots, X_n(t)) \notin B_n].$$

La fiabilité peut, quand à elle, être définie comme la fonction de survie du temps d'atteinte de l'état B_n :

$$R(t) = \mathbb{P}[T_n \geqslant t],$$

où :

$$T_n = \inf \{t \geqslant 0\ ;\ (X_1(t), \ldots, X_n(t)) \in B_n\}\,.$$

2.1.1 Composants totalement échangeables [18]

Dans le cas d'un système où les composants sont tous interchangeables, il est possible d'établir une loi des grands nombres (chapitre 3 de [46]) et un théorème central limite (chapitre 2 de [46] ; voir aussi [18]) pour le temps d'atteinte du sous-ensemble B_n. Autrement dit, ces résultats concernent la fiabilité du système.

Comme le sous-ensemble B_n est croissant et totalement symétrique, alors il doit être de la forme suivante :

$$B_n = \left\{ (x_1, \ldots, x_n)\ ;\ \sum_{i=1}^{n} f(x_i) \geqslant k(n) \right\},$$

où f est une fonction définie de E dans $\mathbb{R}_+$ et $k(n) \geqslant 0$ est un seuil. La fonction f peut être interprétée de la manière suivante dans notre contexte : elle mesure le niveau de dégradation des états possibles d'un composant. On notera $m(t)$ l'espérance et $v(t)$ la variance de la dégradation à l'instant t. L'exemple le plus simple est celui où $E = \{\text{marche}, \text{panne}\}$, $f(\text{marche}) = 0$ et $f(\text{panne}) = 1$: on obtient un système k-sur-n (i.e. le système tombe en panne dès qu'au moins k composants sont en panne). On reviendra plus loin sur cet exemple.

On considère alors la dégradation globale du système que l'on définit comme étant simplement la somme des dégradations individuelles de chaque composant :

$$\forall t \geqslant 0, \quad S_n(t) = \sum_{i=1}^{n} f(X_i(t)).$$

Ce processus n'est pas forcément markovien (cela dépend du choix de la fonction de dégradation f). Néanmoins, il vérifie le théorème central limite suivant :

Proposition 2.1.1 *Pour tout $n \in \mathbb{N}^*$, on définit le processus $Z_n = (Z_n(t))_{t\geqslant 0}$ par :*

$$\forall t \geqslant 0, \quad Z_n(t) = \frac{1}{\sqrt{n}}(S_n(t) - nm(t)).$$

La suite de processus $(Z_n)_{n\in\mathbb{N}^*}$ converge en loi vers le processus gaussien centré Z de fonction d'auto-covariance définie par :

$$\forall 0 \leqslant s \leqslant t\ , \quad Cov[Z_t, Z_s] = {}^t f K(s,t) f,$$

avec $f = (f(e))_{e\in E}$ et

$$K(s,t) = diag(p(s)) \exp(\Lambda(t-s)) - p(s)^t p(t).$$

où $diag(p(s))$ est une matrice diagonale constituée des éléments du vecteur $p(s)$. De plus, les trajectoires de Z sont presque sûrement à trajectoires continues.

La démonstration de ce résultat repose sur l'application conjointe d'un résultat de Hahn [102] et d'un résultat de Whitt [189]. Ce résultat a été étendu par Lachaud dans [123].

On s'intéresse ensuite au temps nécessaire pour que la dégradation totale dépasse un seuil intermédiaire $k(n) = \alpha n + o(\sqrt{n})$ (où α sera précisé plus bas) :

$$T_n = \inf\{t \geqslant 0\ ;\ S_n(t) \geqslant k(n)\}.$$

Théorème 2.1.1 *On suppose que m est croissante sur un intervalle $[0,\tau]$ avec $\tau \leqslant +\infty$. Soit $(k(n))_n$ une suite telle que $k(n) = \alpha n + o(\sqrt{n})$ avec $m(0) < \alpha < m(\tau)$. Alors,*

$$\sqrt{n}(T_n - t_\alpha) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\sigma^2),$$

avec :

$$t_\alpha = \inf\{t \geqslant 0\ ;\ m(t) = \alpha\} \quad et \quad \sigma^2 = \frac{v(t_\alpha)}{(m'(t_\alpha))^2}.$$

La démonstration figurant dans [46] diffère de celle proposée dans l'article [18]. La première démonstration repose sur une inégalité de la fonction de répartition de T_n :

Proposition 2.1.2 *Pour tout réel b, on pose $t_n(b) = \max\left\{0, -\frac{b}{\sqrt{n}}\right\}$. Il existe deux constantes c_1 et c_2 telles que pour $n > c_1$ et $b > c_2$ on ait l'inégalité suivante :*

$$\mathbb{P}[T_n \leqslant t_n(b)] \leqslant \exp\left(-\frac{1}{2}\left(\frac{b\mu(t_\alpha)}{\delta}\right)^2\right) + 2\exp\left(-\frac{b\sqrt{n}\mu(t_\alpha)}{16\delta}\right),$$

où $\delta = \max\{f(e) - f(e')\ e, e' \in E\}$ et $\mu(t_\alpha) = \inf\{m'(t)\ ;\ 0 \leqslant t \leqslant t_\alpha\}$.

La seconde démonstration, celle figurant dans [18], est plus directe : elle utilise à plusieurs reprises le théorème de représentation de Skorokhod [156].

Remarque 2.1.1

1. Il existe un cas assez général où l'hypothèse de monotonie sur m est satisfaite. En effet, si les processus markoviens de sauts sont stochastiquement monotones [178] et si la fonction de dégradation f est croissante, alors la fonction m est croissante.
2. Le théorème 2.1.1 peut être vu comme une généralisation du théorème central limite pour les valeurs centrales d'un n-échantillon ordonné (correspondant au cas de composants non réparables) [160]. En effet, si les composants ont des durées de vie de fonction de répartition G et de densité g, alors on a $m(t) = G(t)$ et donc, pour tout $\alpha \in]0,1[$, on a $t_\alpha = G^{-1}(\alpha)$ et

$$\sigma^2 = \frac{\alpha(1-\alpha)}{g(G^{-1}(\alpha))^2}.$$

Cela correspond au résultat classique sur les statistiques d'ordre. Cependant, on ne peut appliquer notre théorème qu'à des composants dont la loi de la durée de fonctionnement est de type phase (mais on rappelle que l'ensemble des lois de type phase est dense dans l'ensemble des lois portées par $\mathbb{R}_+$).

3. Le théorème central limite pour T_n a été depuis étendu au cas où les composants ne sont plus nécessairement indépendants. En effet, Doukhan *et al.* [86] ont démontré ce genre de résultat en présence de composants faiblement dépendants.

Pour conclure, présentons des exemples de modèles.

Exemple 2.1.1 L'exemple le plus simple est celui, déjà présenté, où les composants peuvent être dans deux états possibles (marche et panne) et où f est à valeurs dans $\{0,1\}$: $f(\text{marche}) = 0$ et $f(\text{panne}) = 1$. Dans ce cas, on est dans le cas d'un système k/n classique avec composants markoviens (i.e. avec des durées de fonctionnement et des durées de réparation exponentielles). On note λ le taux de défaillance et μ le taux de réparation. On peut calculer explicitement toutes les quantités qui interviennent dans les résultats énoncés précédemment. En particulier, on montre que :

$$m(t) = \frac{\lambda}{\lambda+\mu} - \frac{\lambda}{\lambda+\mu} e^{-(\lambda+\mu)t}.$$

C'est une fonction croissante sur $\mathbb{R}_+$: $\tau = \infty$ et $\alpha \in]0, \frac{\lambda}{\lambda+\mu}[$. Pour toute valeur de α dans cet intervalle, on a donc :

$$t_\alpha = -\frac{1}{\lambda+\mu} \log\left(1 - \frac{\alpha(\lambda+\mu)}{\lambda}\right).$$

De plus, la variance asymptotique est égale à :

$$\sigma^2 = \frac{\alpha(1-\alpha)}{(\lambda - \alpha(\lambda+\mu))^2}.$$

Exemple 2.1.2 On peut approcher le cas où les durées de fonctionnement et/ou de réparation ne sont pas exponentielles, en utilisant les lois de type phase [140]. Un exemple est proposé dans [46] et dans [18] où les durées de réparation sont déterministes. Dans ce cas, on peut montrer [59, 105] que la fonction m est donnée par :

$$m(t) = 1 - \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \left(t - \frac{k}{\mu}\right)^k e^{-\lambda(t-\frac{k}{\mu})} \mathbb{1}_{[\frac{k}{\mu},\infty[}.$$

Cette fonction est dans un premier temps croissante, puis présente des oscillations amorties autour de sa limite $\lambda/(\lambda+\mu)$. Les durées de réparation peuvent être approchées par des variables aléatoires de loi d'Erlang de paramètre $(r, r\mu)$. On introduit donc r états fictifs pour les états de panne. On a donc $E = \{\text{marche}, \text{panne}_1, \ldots, \text{panne}_r\}$ et le générateur infinitésimal suivant :

$$\Lambda = \begin{pmatrix} -\lambda & \lambda & 0 & \cdots & 0 \\ 0 & -r\mu & r\mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & & \ddots & -r\mu & r\mu \\ r\mu & 0 & \cdots & 0 & -r\mu \end{pmatrix}.$$

Naturellement, la fonction f est donnée par : $f(\text{marche}) = 0$ et $f(\text{panne}_j) = 1$, pour tout $j \in \{1, \ldots, r\}$. La transformée de Laplace de la loi marginale $p(t)$ peut se déduire facilement à partir des équations de Chapmann-Kolmogorov associées au générateur infinitésimal Λ du processus décrit ci-dessus. On peut en déduire la transformée de Laplace de m :

$$\int_0^\infty e^{-st} m(t) dt = \frac{\lambda[(s+r\mu)^r - (r\mu)^r]}{s[(s+\lambda)(s+r\mu)^r - (r\mu)^r]}.$$

Cependant, les autres quantités ne peuvent pas être calculées explicitement. On est alors obligé de se contenter de calculs numériques.

Exemple 2.1.3 Ce troisième exemple reprend une situation étudiée dans Pham *et al.* [153] avec des composants partiellement réparables. Le composant peut être dans un des trois états suivants : $E = \{\text{parfait état, partiellement dégradé, panne}\}$. On suppose qu'un composant en parfait état peut devenir partiellement dégradé ou tomber directement en panne. Un composant partiellement dégradé peut tomber en panne ou bien être réparé. Un composant en panne est remis à neuf. On a donc le générateur suivant :

$$\Lambda = \begin{pmatrix} -(\lambda_1+\lambda_2) & \lambda_1 & \lambda_2 \\ \mu_1 & -(\mu_1+\lambda_3) & \lambda_3 \\ \mu_2 & 0 & -\mu_2 \end{pmatrix}.$$

La loi marginale du processus (partant initialement du parfait état) peut être calculée explicitement. Avec nos notations, Pham *et al.* [153] choisissent la fonction f ainsi : $f(\text{parfait état}) = f(\text{partiellement dégradé}) = 0$ et $f(\text{panne}) = 1$. Dans ce cas, on obtient :

$$m(t) = \frac{\lambda}{\lambda+\mu_2} - \frac{\lambda}{\lambda+\mu_2} e^{-(\lambda+\mu_2)t}.$$

On peut alors appliquer le théorème 2.1.1 pour tout $\alpha \in]0, \frac{\lambda}{\lambda+\mu_2}[$. On aboutit à des expressions similaires à celles obtenues dans le cas d'un composant binaire markovien (voir le premier exemple ci-dessus).

On peut considérer un autre choix pour la fonction f qui prendrait plus en compte la différence entre les deux premiers états. Par exemple, on pourrait faire le choix suivant : $f(\text{parfait état}) = 0$, $f(\text{partiellement dégradé}) = \frac{1}{2}$ et $f(\text{panne}) = 1$. Dans ce cas, la fonction m n'est pas toujours croissante sur un intervalle $[0, \tau[$ (cela dépend de la valeur de μ_2) et notre résultat ne peut pas être toujours appliqué.

Exemple 2.1.4 On considère un composant pour lequel il y a deux types de panne possible (voir [77] par exemple). Afin de les distinguer, on introduit deux états distincts de panne. On a donc $E = \{\text{marche}, \text{panne}_1, \text{panne}_2\}$. L'espace d'états possibles d'un composant est donc partiellement ordonné (les deux types de panne ne peuvent pas être comparés). On a donc des transitions (dans les deux sens) entre l'état de marche et les états de panne (mais pas entre états de panne). On a donc le générateur suivant :

$$\Lambda = \begin{pmatrix} -\lambda & p\lambda & (1-p)\lambda \\ \mu_1 & -\mu_1 & 0 \\ \mu_2 & 0 & -\mu_2 \end{pmatrix}.$$

On choisit naturellement la fonction f suivante : $f(\text{marche}) = 0$ et $f(\text{panne}_1) = f(\text{panne}_2) = 1$. Malheureusement, on ne dispose pas d'expressions explicites pour les quantités intervenant dans le théorème 2.1.1. Des exemples numériques ont été étudiés dans [46].

2.1.2 Composants partiellement échangeables [17]

On s'intéresse maintenant au cas plus général où les composants du système ne sont que partiellement échangeables. On peut alors établir une loi du zéro-un (chapitre 4 de [46] ; voir aussi [17]), mais cette fois-ci pour la disponibilité. L'idée est la suivante : dans un système monotone constitué d'un grand nombre de composants, la disponibilité devrait être proche de 0 ou de 1 la plupart du temps.

On suppose ici que E est totalement ordonné (on notera $\leq$ l'ordre sur E) ; cela induit un ordre partiel sur E^n définie coordonnée par coordonnée (on le notera $\preceq$). On note e_{min} et e_{max} respectivement l'élément minimal et l'élément maximal de E. Soit $(p_t)_{t\geqslant 0}$ une famille de probabilités sur E représentant la loi de l'état d'un composant à tout instant. On suppose que p_t est une fonction dérivable du temps et que la famille (p_t) est stochastiquement monotone, i.e.

$$t_1 \leqslant t_2 \Longrightarrow \forall e \in E, F_{t_1}(e) \geqslant F_{t_2}(e),$$

où

$$\forall t \geqslant 0, \forall e \in E, \quad F_t(e) = \sum_{e' \leq e} p_t(e').$$

Cette hypothèse de monotonie stochastique traduit le fait que les composants ont une tendance à être de plus en plus dégradés avec le temps. L'indisponibilité à l'instant t sera notée :

$$\mu_t^n(B_n) = \mathbb{P}[(X_1(t), \ldots, X_n(t)) \in B_n].$$

On supposera que B_n est tel que $\mu_0^n(B_n) = 0$ (le système est initialement en état de marche) et que $\lim_{t\to\infty} \mu_t^n(B_n) = 1$ (le système est à long terme en panne, malgré les réparations). Pour tout $\varepsilon \in]0,1[$, on pose t_ε^n la valeur de t telle que $\mu_{t_\varepsilon}^n(B_n) = \varepsilon$. L'objectif est d'établir une loi du zéro-un pour t_ε^n, autrement dit que, pour tout $\varepsilon \in]0, \frac{1}{2}[$, t_ε^n et $t_{1-\varepsilon}^n$ sont proches.

Pour tout t, on note F_t^- le pseudo-inverse de F_t :

$$\forall u \in]0,1[, \quad F_t^- = \inf\{x \in E; F_t(x) \geqslant u\}$$

qui est bien défini puisque E est totalement ordonné. Notons π la loi uniforme sur $[0,1]$. Soit $U_1, \ldots, U_n$ des variables aléatoires indépendantes et de loi π. On a :

$$\begin{aligned} \mu_t^n(B_n) &= \mathbb{P}[(F_t^-(U_1), \ldots, F_t^-(U_n)) \in B_n] \\ &= \mathbb{P}[(U_1, \ldots, U_n) \in A_n^t] \\ &= \pi^{\otimes n}(A_n^t), \end{aligned}$$

où $\otimes$ désigne le produit tensoriel et A_n^t est le sous-ensemble de $[0,1]^n$ défini par :

$$A_n^t = \{(u_1, \ldots, u_n) \in [0,1]^n; (F_t^-(u_1), \ldots, F_t^-(u_n)) \in B_n\}.$$

Par construction de ce sous-ensemble, on a le résultat suivant :

Lemme 2.1.1

1. *Si B_n est un sous-ensemble croissant de E^n, alors A_n^t est un sous-ensemble croissant de $[0,1]^n$ et $\mu_t^n(B_n)$ est une fonction croissante de t.*
2. *Si B_n est un sous-ensemble symétrique de E^n, alors A_n^t est un sous-ensemble symétrique de $[0,1]^n$.*
3. *Si B_n est un sous-ensemble croissant de E^n, alors*

$$t_1 \geqslant t_2 \Longrightarrow A_n^{t_1} \subseteq A_n^{t_2}.$$

L'intérêt de transporter le problème à un sous-ensemble de l'hypercube en dimension n muni de la loi uniforme est de pouvoir utiliser les résultats de Bourgain *et al.* [69] (voir également [94]). On obtient alors la loi du zéro-un suivante :

Théorème 2.1.2 *Supposons que B_n est un sous-ensemble croissant et symétrique de E^n et supposons que, pour tout $\varepsilon \in]0,1[$, il existe $\Delta_\varepsilon > 0$ tel que, pour tout n,*

$$\forall t \leqslant t_n^\varepsilon, \quad \min_{e \in E \setminus e_{max}} \left| \frac{dF_t(e)}{dt} \right| \geqslant \Delta_\varepsilon.$$

Alors, il existe une constante C telle que, pour tout $\varepsilon \in]0, \frac{1}{2}[$,

$$t_{1-\varepsilon}^n - t_\varepsilon^n \leqslant \frac{C \log \frac{1}{2\varepsilon}}{\Delta_{1-\varepsilon} \log n}.$$

Le terme en $\log n$ dans cette inégalité peut être amélioré dans le cas particulier où B_n est totalement symétrique (voir la partie précédente) ; dans ce cas, on obtient du $\sqrt{n}$ à la place du $\log n$.

La démonstration repose sur la notion d'influence d'une coordonnée d'une fonction définie sur un espace produit telle qu'elle a été introduite dans la littérature. Dans le cas de composants binaires, cette notion coïncide avec le facteur d'influence proposé par Birnbaum [65] dans le domaine de la fiabilité.

Passons maintenant à quelques exemples de composants et de systèmes pour lesquels on a la loi du zéro-un énoncée ci-dessus. Commençons par présenter quelques modèles possibles pour les composants. L'hypothèse de monotonie stochastique, bien que naturelle, est assez difficile à vérifier pour un modèle donné. Certains processus markoviens de naissance et mort, englobant en particulier les processus binaires, vérifient cette hypothèse (voir [178] par exemple). C'est pourquoi nous avons considéré les deux types de composants suivants :

1. composants binaires : il s'agit du même type de composants que ceux vus dans le premier exemple de la partie précédente.
2. composants à dégradations progressives et réparations minimales : il s'agit d'une généralisation du cas précédent. Un composant dans l'état $e \in E \setminus \{e_{min}, e_{max}\}$ peut, soit passer dans l'état plus dégradé $e+1$ (état plus dégradé que e le plus proche de e), soit passer dans l'état moins dégradé $e-1$ (état moins dégradé que e le plus proche de e). On a donc le générateur suivant :

$$\Lambda = \begin{pmatrix} -\lambda_0 & \lambda_0 & & & & \\ -\mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & & & \\ & \ddots & \ddots & \ddots & & \\ & & -\mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & \\ & & & \mu_\gamma & -\mu_\gamma & \end{pmatrix}.$$

Pour terminer, nous allons présenter brièvement deux exemples de systèmes pour lesquels le sous-ensemble des états de panne est croissant et symétrique. En croisant ces exemples de systèmes avec les exemples de composants, on peut construire plusieurs exemples de couple système/composant. Dans certains cas, les calculs ne peuvent pas être faits explicitement. Les exemples ci-dessous utilisent la fonction f de dégradation introduite dans la partie précédente et proposent donc des versions multi-états de modèle classique pour des système.

1. Système k-sur-n : il s'agit du système étudié dans la partie précédente. Dans ce cas, le sous-ensemble B_n est totalement symétrique.
2. Système k-consécutifs-sur-n : les composants d'un tel système sont disposés en ligne ou en cercle (d'un point de vue asymptotique cela est équivalent). Ces systèmes sont en panne dès que la somme des dégradations de k composants consécutifs dépasse un certain seuil k.

Pour plus de détails et/ou pour d'autres exemples, on pourra consulter le chapitre 4 de [46] ou [17].

2.2 Modèle markovien avec covariables dépendantes du temps [15]

L'objectif de cette étude est de proposer un modèle de dégradation qui prenne en compte des covariables qui dépendent du temps. Pour une entreprise comme EDF, il est important de déterminer les facteurs physiques qui influencent la dégradation et les pannes des éléments constituant ses installations afin de pouvoir gérer leurs vieillissements. Ces facteurs sont souvent liés au mode d'exploitation des installations (par exemple, le nombre d'heures de fonctionnement, le nombre de démarrages à froid, intermédiaire ou à chaud, etc. d'une centrale). Cependant, il n'est pas toujours évident de savoir quels sont les facteurs les plus déterminants. Un modèle statistique peut donc aider les ingénieurs à les classifier par ordre d'influence sur la vitesse de dégradation et à déterminer un mode d'exploitation qui limite la cinétique de dégradation du mécanisme étudié.

La méthodologie présentée ici a été développée pour mieux comprendre le vieillissement de certains composants installés sur des centrales. Ces composants assurent une étanchéité. Si ce composant n'assure plus sa fonction d'étanchéité, des jets de vapeur peuvent survenir pouvant endommager des composants voisins. On considère alors que le composant est en panne s'il a perdu sa fonction d'étanchéité. Des inspections visuelles sont effectuées de temps en temps : on dispose donc d'informations qualitatives sur le niveau de dégradation (il n'y a pas de mesure quantitative de la dégradation). Lors des inspections, le niveau de dégradation des composants est relevé selon une échelle ordinale E (i.e. E est un ensemble fini). Sans perte de généralité, on note $E = \{0, \dots, q\}$ où 0 correspond au parfait état, q à l'état de défaillance et les autres états à des niveaux de dégradation intermédiaire. Par exemple, dans le cas de cette étude, $q = 3$ et on a deux états de dégradation intermédiaires : 1 est l'état de dégradation mineure et 2 est l'état de dégradation majeure. On supposera que les composants tombant en panne (i.e. dans l'état q) sont immédiatement décelés. On sait donc qu'au moment de la fin de l'étude les composants encore en usage ne sont pas défaillants. Une observation sera donc associée à une date d'inspection ou un instant de panne. De plus, une décision de remplacement est associée à chaque niveau de dégradation. Par exemple, dans le cas de notre étude, la

État	Niveau de dégradations	Remplacement ?
0	Parfait état	Non
1	Dégradation mineure	Non
2	Dégradation majeure	Oui
3	Panne	Oui

Table 2.1 – Niveaux de dégradation et actions de remplacement associées

politique de remplacement est décrite par le tableau 2.1. Ce tableau souligne la différence entres les états 1 et 2. Un composant classé dans l'état 1 signifie que, lors d'une inspection, une dégradation a été observée sur le composant mais qu'elle a été jugée suffisamment faible pour supposer que le composant ne tombera pas en panne avant la prochaine inspection. Un composant dans l'état 2 signifie que, lors d'une inspection, une dégradation assez importante a été observée sur le composant si bien qu'on a décidé de le remplacer par un composant neuf.

Dans une première partie, on présente le modèle de dégradation à temps discret sans, puis avec covariables. La seconde partie porte sur l'inférence statistique pour ces modèles. On trouvera dans [15] une application du modèle sur des données simulées (uniquement pour le modèle sans covariable), puis à un jeu de données EDF.

2.2.1 Modèle de dégradation à temps discret

On propose d'abord une chaîne de Markov à temps discret et à valeurs dans E pour décrire l'évolution de la dégradation d'un composant. Ce modèle dépend donc de q paramètres correspondant aux probabilités de transition à un pas de la chaîne. Cependant les probabilités de transition peuvent dépendre de quelques covariables qui évoluent dans le temps. Ces covariables correspondent au mode d'exploitation de la centrale et sont observées chaque mois. Puisqu'on dispose d'observations mensuelles des covariables, nous avons retenu un modèle à temps discret. L'influence des covariables est modélisée par des régressions logistiques pour chacune des probabilités de transition. On obtient donc un second modèle qui est non-homogène en temps.

Modèle sans covariable

Le modèle sans covariable est donc une simple chaîne de Markov sur l'ensemble des niveaux de dégradation possible. Sans action de maintenance préventive sur le composant (autre que remplacements par des composants neufs), il s'agit donc d'une chaîne de naissance pure (il n'y a pas de transition possible d'un état i vers un état $j < i$) et donc les probabilités de transition sont données par la matrice suivante :

$$P = \begin{pmatrix} 1-p_0 & p_0 & 0 & 0 \\ 0 & 1-p_1 & p_1 & 0 \\ 0 & 0 & 1-p_2 & p_2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Il s'agit donc d'un modèle à $q = 3$ paramètres. Pour tout $i \in \{0, \dots, q-1\}$, soit S_i le temps de séjour dans l'état i. Ce sont des variables aléatoires de loi géométrique dont les paramètres respectifs sont les probabilités de transition. Pour tout $i \in \{0, \dots, q-1\}$, le temps moyen de séjour dans l'état i est donc égal à $1/p_i$. Pour tout $i \in \{1, \dots, q\}$, soit T_i le temps d'atteinte de l'état i : $T_i = S_0 + \cdots + S_i$. Clairement, le temps de panne d'un composant (i.e. le temps d'atteinte de l'état q) est presque sûrement fini et le temps moyen jusqu'à la panne (MTTF) est égal à :

$$\text{MTTF} = \mathbb{E}[T_q] = \sum_{i=0}^{q-1} \frac{1}{p_i}.$$

Un cas simple apparaît quand toutes les probabilités de transition sont égales : $p_0 = \cdots = p_{q-1} = p$; dans ce cas, pour tout $i \in \{2, \ldots, q\}$, la loi marginale de T_i est la loi binomiale négative de paramètres (i, p). Pour $q = 3$, les lois de T_2 et T_3, dans le cas général, sont données par le lemme suivant.

Lemme 2.2.1 *La loi de T_2 est donnée par :*

$$\forall k \in \{2, 3, \ldots\}, \quad \mathbb{P}[T_2 = k] = \frac{p_0 p_1}{p_1 - p_0} \left[(1 - p_0)^{k-1} - (1 - p_1)^{k-1}\right]$$

et la loi de T_3 est donnée par :

$$\forall k \in \{3, 4, \ldots\}, \quad \mathbb{P}[T_3 = k] = \frac{p_0 p_1 p_2}{p_1 - p_0} \left[\frac{(1 - p_2)^{k-1} - (1 - p_0)^{k-1}}{p_0 - p_2} - \frac{(1 - p_2)^{k-1} - (1 - p_1)^{k-1}}{p_1 - p_2}\right].$$

Modèle avec covariables

On suppose que r covariables sont observées toutes les fins de mois (dans notre cas d'étude, $r = 2$). Soit $(z(t))_{t\in\mathbb{Z}}$ le vecteur des covariables disponibles : pour tout $t \in \mathbb{N}$, $z(t) = (z_1(t), \ldots, z_r(t)) \in \mathbb{R}^r$ avec $z_0(0) = 0$ en général. De plus, pour tout $j \in \{1, \ldots, r\}$, $t \mapsto z_j(t)$ est supposé être une fonction croissante du temps.

Pour tout $i \in \{0, \ldots, q-1\}$, on note $\Omega_i \subset \{1, \ldots, r\}$ l'ensemble des covariables influençant la probabilité de transition de l'état i vers $i + 1$. On considère que la relation entre les covariables et les probabilités de transition est décrite par une régression logistique :

$$\forall i \in \{0, \ldots, q-1\}, \forall t \in \mathbb{N}, \quad p_i(t) = \frac{\exp\left(\alpha_i + \sum_{j\in\Omega_i} \beta_{ij} z_j(t)\right)}{1 + \exp\left(\alpha_i + \sum_{j\in\Omega_i} \beta_{ij} z_j(t)\right)},$$

où $\alpha_i \in \mathbb{R}$ (intercept) et $\underline{\beta}_i = (\beta_{ij})_{j\in\Omega_i} \in \mathbb{R}_+^{|\Omega_i|}$. Pour tout $i \in \{0, \ldots, q-1\}$, soit $\theta_i = (\alpha_i, \underline{\beta}_i) \in \Theta_i = \mathbb{R} \times \mathbb{R}_+^{|\Omega_i|}$ et $\Theta = \Theta_0 \times \cdots \times \Theta_{q-1}$ (espace des paramètres). Le modèle complet correspond au cas où $\Omega_0 = \cdots = \Omega_{q-1} = \Omega = \{1, \ldots, r\}$ et est décrit par $q(r+1)$ paramètres (dans notre cas d'étude, le modèle complet contient 9 paramètres). Alors que l'intercept peut prendre n'importe quelle valeur réelle, il serait plutôt attendu que les paramètres associés aux covariables soient positifs. Plus la covariable est élevée, plus les probabilités de transitions seront grandes. Une covariable n'ayant pas d'influence sur une probabilité de transition donnée devrait donner un coefficient nul.

2.2.2 Inférence statistique

Après avoir décrit toutes les observations possibles, nous allons expliquer comment la fonction de vraisemblance peut être calculée. Afin de comprendre le vieillissement, nous proposons une estimation de la loi du temps de panne en considérant un comportement moyen des covariables.

Observations

Puisqu'on considère une chaîne de Markov de naissance pure (que des covariables soient incluses ou pas), il est possible de donner une liste exhaustive de toutes les observations possibles. Dans le cas de l'étude où $q = 3$, cela donne dix-huit cas différents recensés dans le tableau 2.2. La mention *cens* indique qu'à la fin de l'étude, le composant n'est pas en panne. Il y a donc trois états finaux possibles : *cens* (composant non encore en panne à la fin de l'étude, i.e. encore en service), 2 (composant fortement dégradé), 3 (composant en panne). Une telle énumération est possible pour un autre nombre d'états et/ou pour un autre choix de la politique de remplacement.

Fonction de vraisemblance

On note $H(E, \mathcal{R})$ le nombre de cas distincts qui dépend donc à la fois de l'espace d'états E et de la politique de remplacement $\mathcal{R}$ appliquée. Dans notre cas d'étude, on a $H(E, \mathcal{R}) = 18$ (voir plus haut). Pour

Cas	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
t_1	*cens*	2	3	0	1	0	1	0	1	0	0	0	1	1	1	0	0	0
t_2	-	-	-	*cens*	*cens*	2	2	3	3	1	1	1	1	1	1	1	1	1
t_3	-	-	-	-	-	-	-	-	-	*cens*	2	3	*cens*	2	3	1	1	1
t_4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	*cens*	2	3

Table 2.2 – Liste exhaustive des cas possibles avec $q = 3$

tout $h \in H(E, \mathcal{R})$, on note m_h le nombre de temps d'inspection ou de défaillance (ici, $m_h \in \{1, \ldots, 4\}$) et n_h le nombre de composants dans l'échantillon correspondant à ce cas. La fonction de vraisemblance s'écrit, en toute généralité, sous la forme suivante :

$$L(\underline{\theta}|\text{données}) = \prod_{h=1}^{H(E,\mathcal{R})} \prod_{i=1}^{n_h} \mathbb{P}\left[X_{t_{0,i}}^{(h,i)} = 0, X_{t_{1,i}}^{(h,i)} \in A_1, \ldots, X_{t_{m_h,i}}^{(h,i)} \in A_{m_h}\right],$$

où $X^{(h,i)}$ sont des copies i.i.d. de la chaîne de Markov X décrite précédemment et où $A_1, \ldots, A_{m_h}$ sont des sous-ensembles convenablement choisis pour décrire le h-ième cas. Par exemple, pour le cas 1 du tableau 2.2, on a $m_1 = 1$ et $A_1 = \{0, 1, 2\}$ et pour le cas 2, on a $m_2 = 1$ et $A_1 = \{2\}$, etc. Pour tout $h \in H(E, \mathcal{R})$ et pour tout $i \in \{1, \ldots, n_h\}$, on a :

$$\mathbb{P}\left[X_{t_{0,i}}^{(h,i)} = 0, X_{t_{1,i}}^{(h,i)} \in A_1, \ldots, X_{t_{m_h,i}}^{(h,i)} \in A_{m_h}\right] = \prod_{j=1}^{m_h} \mathbb{P}\left[X_{t_j}^{(h,i)} \in A_j | X_{t_{j-1}}^{(h,i)} \in A_{j-1}\right]$$

avec $A_0 = \{0\}$, parce qu'on a toujours $X_{t_0} = 0$. Chaque contribution à la fonction de vraisemblance correspond à la probabilité de la trajectoire observée pour chacun des composants.

Estimateur du maximum de vraisemblance

On obtient alors l'estimateur du maximum de vraisemblance en optimisant la fonction L :

$$\widehat{\underline{\theta}}_n = \text{argmax}_{\underline{\theta} \in \Theta} L(\underline{\theta}|\text{données}),$$

avec $n = \sum_{h \in H(E,\mathcal{R})} n_h$. La théorie du maximum de vraisemblance paramétrique permet de supposer raisonnablement l'approximation de la loi de $\widehat{\underline{\theta}}$ convenablement renormalisée par la loi normale centrée et de matrice de variance-covariance $I^{-1}(\underline{\theta})$:

$$\sqrt{n}\left(\widehat{\underline{\theta}}_n - \underline{\theta}\right) \overset{approx.}{\sim} \mathcal{N}\left(0, I^{-1}(\underline{\theta})\right)$$

la matrice d'information de Fisher $I(\underline{\theta})$ pouvant être estimée par la hessienne du logarithme de la vraisemblance évaluée en $\widehat{\underline{\theta}}_n$. Ainsi, si l'on dispose d'un échantillon de grande taille, on peut construire des intervalles de confiance asymptotiques pour chacun des paramètres (ou des ellipsoïdes de confiance pour plusieurs paramètres). En présence d'échantillon de taille petite ou moyenne, on pourra préférer l'obtention d'intervalles de confiance par bootstrap.

Comparaison et sélection de modèles

Il est bien connu que, plus on considère des covariables, plus la fonction de vraisemblance augmente. C'est pourquoi il est important d'utiliser un critère pénalisant la fonction de vraisemblance à l'optimum par la dimension du modèle. On propose de considérer le critère d'Akaike (AIC). Si l'ensemble des covariables pour un modèle donné est :

$$\Theta = \prod_{i=0}^{q-1} \mathbb{R} \times \mathbb{R}_+^{|\Omega_i|}$$

alors

$$AIC = -2\log L(\widehat{\theta}_n) + 2\sum_{i=0}^{q-1}(1 + |\Omega_i|).$$

On pourra donc chercher le sous-modèle donnant la plus petite valeur du critère d'Akaike. Cependant, une recherche exhaustive n'est pas toujours envisageable (d'autant plus que les calculs de l'estimation des paramètres d'un modèle sont assez longs). En effet, le nombre de sous-modèles croît exponentiellement avec le nombre de covariables. Par exemple, pour un modèle à quatre états et deux (resp. quatre) covariables, il existe 63 (resp. 4095) sous-modèles. Ainsi, excepté pour un petit nombre de covariables et/ou d'états, une recherche exhaustive n'est pas applicable en pratique. Une stratégie possible est la suivante :

1. *étapes globales* : une procédure ascendante de sélection de covariables est appliquée. Les covariables sont introduites une à une pour toutes les probabilités de transition tant que le critère d'Akaike diminue.
2. *étapes locales* : une procédure pas-à-pas peut alors être appliquée pour affiner le modèle obtenu à la fin de la première étape (qui peut être jugée acceptable). Des covariables peuvent être introduites ou retirées pour une probabilité de transition donnée.

Estimation de la loi du temps de panne

Une fois un modèle ajusté, on peut souhaiter estimer la loi du temps de panne. Pour le modèle sans covariable, on utilisera le lemme précédent. Dans le cas d'un modèle avec covariables, la situation est plus complexe. On peut considérer un comportement moyen des covariables pour avoir une idée de la loi du temps de panne. Pour cela, on pourra effectuer des simulations de Monte-Carlo afin d'obtenir un échantillon simulé de temps de panne pour ces covariables moyennes. Partant de cet échantillon, on peut estimer la loi du temps de panne à l'aide de la fonction de répartition empirique ou d'un estimateur à noyau de la densité.

2.3 Composant soumis à plusieurs modes d'endommagements [5]

Dans le cadre du suivi des moteurs en exploitation, on s'intéresse à la modélisation de la dégradation d'un moteur pouvant conduire à un événement redouté (événement étudié dans le cadre des analyses de sécurité). On considère par exemple l'événement redouté « arrêt en vol du moteur ». Cet événement peut résulter de différentes pannes sur différents composants, elles-mêmes dues à une ou plusieurs dégradations. L'objectif est alors de mettre en place, pour chaque composant, un modèle de dégradation dans lequel plusieurs mécanismes de défaillance sont en concurrence vis-à-vis de la défaillance de ce composant. Ces modèles permettront, par exemple, de vérifier que les exigences fixées de fiabilité et de sécurité du composant sont satisfaites, ou encore d'étudier l'impact d'un nouvel événement survenu en service sur la probabilité d'occurrence de l'événement redouté. Comme exemple de composants, on peut considérer le pignon qui est une roue d'engrenage permettant de transmettre un mouvement. On peut, par exemple, trouver ce genre de composant dans la boîte accessoires ou le réducteur. On considère les modes de dégradation qui peuvent conduire à la défaillance « perte de transmission du mouvement ». On suppose, pour simplifier, qu'il en existe deux : l'usure et la crique [1]. Ces deux dégradations ne peuvent être détectées que lors des inspections où différentes actions de maintenance peuvent être réalisées. Elles se différencient par le fait que la présence d'usure sur le pignon est tolérable dans une certaine limite, alors que la présence d'une crique n'est jamais acceptable. Lors de l'opération de maintenance, la pièce sera donc rebutée si l'usure a dépassé la tolérance fixée et/ou si une crique est présente (dans les autres cas, le composant sera remis en service en l'état). Le tableau 2.3 décrit la classification des différents niveaux de dégradation pour un endommagement. Les données disponibles (pour le pignon, mais aussi pour tous les autres composants d'un turbomoteur) dans le retour d'expériences (REX) correspondent aux heures de fonctionnement du composant lors de l'inspection ainsi que le niveau de dégradation du composant. Si une dégradation est observée lors d'une inspection, on ignore depuis quand celle-ci est présente. De manière générale, les différentes dégradations sont détectables

1. Fente à la surface d'un métal, de profil irrégulier, résultant de la séparation entre grains sous l'effet de contraintes anormales (source : Larousse)

Etat	Description
0	Pas d'endommagement
1	Endommagement dans les limites de tolérance
2	Endommagement hors des limites de tolérance
3	Défaillance

Table 2.3 – Classification des niveaux d'endommagement

dans seulement deux situations : (a) lors d'une inspection si la dégradation est présente mais n'a pas encore conduit à la défaillance du composant (effet non observable en exploitation) ; (b) par son effet sur le système si la dégradation a conduit à la défaillance du composant. L'état de panne du composant est observé seulement si la défaillance a lieu avant qu'une inspection ait révélé une dégradation hors tolérances. Donc, seuls les temps de défaillance sont exacts : on se trouve ainsi en présence de durées censurées. Pour chaque composant, on ne dispose de l'état de celui-ci qu'à une seule date. Le tableau 2.4 contient un exemple de données qui peuvent être extraites du REX (données modifiées pour des raisons de confidentialité), en utilisant la classification décrite dans le tableau 2.3. Si on s'intéresse seulement au temps jusqu'à l'état de

Heures de vol	Etat endommagement 1	Etat endommagement 2
1203	0	1
3201	3	0
908	0	2
2805	2	0

Table 2.4 – Exemple de données issues du REX (données modifiées)

défaillance pour chacun des modes d'endommagement, on peut essayer d'ajuster une loi paramétrique (par exemple, une loi de Weibull). Dans ce cas, le REX, comme montré dans le tableau ci-dessus, fournit de nombreuses durées censurées à droite et plus rarement des durées exactes. Durant une inspection, plusieurs actions sont possibles. Si aucune dégradation n'est présente ou si elle est dans les tolérances fixées, le composant sera remis en service en l'état. Si la dégradation détectée est en dehors des tolérances fixées, le composant sera rebuté, i.e. remplacé par un composant neuf.

Dans la sous-section 2.3.2, on propose une modélisation de la dégradation d'un composant sur la base du retour d'expériences. Deux cas sont distingués selon que les mécanismes de défaillance sont supposés indépendants ou non. Des estimateurs des paramètres de ce modèle sont proposés dans la sous-section 2.3.3. La sous-section 2.3.4 développe une méthode permettant d'évaluer les paramètres du modèle de la dégradation intrinsèque, i.e. de la dégradation du composant non maintenu. Ce modèle, une fois ajusté, permet par exemple d'optimiser la politique de maintenance préventive. Dans [64] (voir également [5], des études numériques sur des données simulées sont proposées.

2.3.1 Modélisation des mécanismes de défaillance d'un composant

En partant du cas d'étude, on propose un modèle où deux types de mécanismes de défaillance sont considérés.

Mécanisme de défaillance « avec tolérance »

Considérons dans un premier temps un mécanisme de défaillance qui évolue au cours du temps et dont la présence sur le composant est acceptable sous certaines conditions (par exemple, une usure sur un pignon). La tâche de maintenance associée à ce type de mécanisme de défaillance dépend d'un critère de tolérance (critère qualitatif ou quantitatif vérifié lors de l'inspection). Par conséquent, ce type de mécanisme de défaillance est caractérisé par le fait que l'on observe deux stades successifs de dégradation : le premier

stade correspond à un niveau de dégradation dans les tolérances, il est acceptable pour le maintien en service du composant ; le second correspond à un niveau de dégradation hors tolérances qui entraîne la remise à neuf du composant. Ce type de mécanisme de défaillance résulte de dégradations « avec tolérance ». Les différents états sont résumés sur la figure 2.1(a).

Ce modèle fait donc intervenir quatre états. Nous supposerons par la suite que les durées dans chacun des états (sauf le dernier état qui est absorbant) sont des variables aléatoires indépendantes. On peut décrire l'évolution de la dégradation à l'aide d'un processus stochastique $(Z_t^{(at)})_{t\geq 0}$ à valeurs dans $\{0,1,2,3\}$. On fait l'hypothèse que $Z_0^{(at)} = 0$ presque sûrement.

Compte tenu du REX disponible (taux de censure élevé), on impose des lois exponentielles aux temps de séjour. On obtient alors un modèle markovien dont les transitions sont décrites sur la figure 2.1(b). Pour tout couple d'états (u, v) avec $u < v$, on note $X_{u,v}^{(at)}$ la durée pour passer de l'état u à l'état v. En particulier, $X_{u,u+1}^{(at)}$ est le temps de séjour dans l'état u qui est donc de loi exponentielle dont le paramètre sera noté $\lambda_{u,u+1}^{(at)}$. Enfin, le temps de défaillance associé à ce mécanisme est la durée $X_{03}^{(at)}$ dont on peut calculer la loi explicitement. Par exemple, si tous les taux de transition sont différents, on obtient l'expression suivante pour la densité de la loi de $X_{03}^{(at)}$ (on peut aussi calculer les densités pour les temps d'atteinte des autres états) :

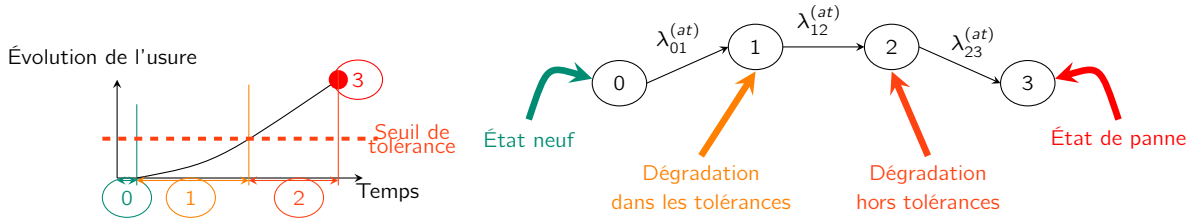
$$\forall t \geq 0, \quad f_{X_{03}^{(at)}}(t) = \frac{\lambda_{01}^{(at)}\lambda_{12}^{(at)}\lambda_{23}^{(at)}}{\lambda_{01}^{(at)} - \lambda_{12}^{(at)}} \left(\frac{e^{-\lambda_{23}^{(at)}t} - e^{-\lambda_{12}^{(at)}t}}{\lambda_{12}^{(at)} - \lambda_{23}^{(at)}} - \frac{e^{-\lambda_{23}^{(at)}t} - e^{-\lambda_{01}^{(at)}t}}{\lambda_{01}^{(at)} - \lambda_{23}^{(at)}} \right).$$

Il est possible de calculer la densité de $X_{03}^{(at)}$ lorsqu'exactement deux taux sont identiques ou lorsque tous les taux sont égaux (dans ce cas, on obtient une loi d'Erlang).

On peut relier le processus $(Z_t^{(at)})_{t\geq 0}$ aux durées de séjour décrites ci-dessus de la manière suivante :

$$\mathbb{P}\left[Z_t^{(at)} = u\right] = \mathbb{P}\left[X_{0,u}^{(at)} \leq t < X_{0,u+1}^{(at)}\right] \quad \text{pour } u \in \{0,1,2,3\},$$

en ajoutant les notations $X_{0,0} = 0$ et $X_{0,4} = +\infty$.



(a) Exemple d'évolution de dégradation « avec tolérance »

(b) Graphe de Markov de la dégradation « avec tolérance »

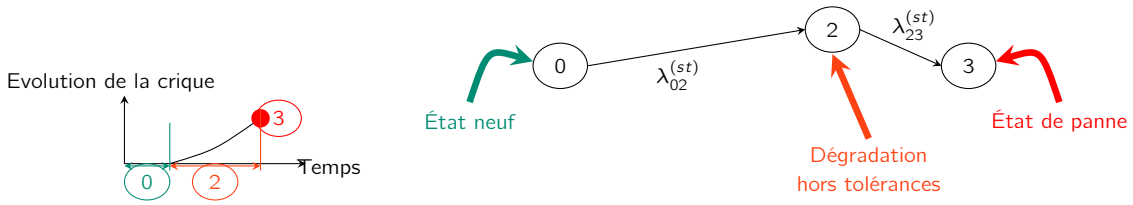
Figure 2.1 – Dégradation « avec tolérance »

Mécanisme de défaillance « sans tolérance »

Considérons maintenant un mécanisme de défaillance dont la présence n'est jamais tolérée sur le composant (par exemple, une crique sur un pignon). Dès leur apparition, ces dégradations sont hors tolérance. Ce type de mécanisme de défaillance résulte de dégradation « sans tolérance ». Les différents états sont résumés sur la figure 2.2(a). Ici également, on peut décrire l'évolution de cette dégradation à l'aide d'un processus stochastique $(Z_t^{(st)})_{t\geq 0}$ à valeurs dans $\{0,1,3\}$ (avec la même condition initiale $Z_0^{(st)} = 0$ presque sûrement).

De même que pour un mécanisme de défaillance « avec tolérance », on suppose que les durées dans chaque état sont des variables aléatoires indépendantes et de lois exponentielles. Le modèle markovien obtenu est décrit par la figure 2.2(b). On obtient la densité suivante pour le temps avant défaillance $X_{03}^{(st)}$ si les deux taux sont différents :

$$\forall t \geq 0, \quad f_{X_{03}^{(st)}}(t) = \frac{\lambda_{23}^{(st)}\lambda_{02}^{(st)}}{\lambda_{02}^{(st)} - \lambda_{23}^{(st)}} \left(e^{-\lambda_{23}^{(st)}t} - e^{-\lambda_{02}^{(st)}t} \right).$$



(a) Exemple de dégradation « sans tolérance »

(b) Graphe de Markov de la dégradation « sans tolérance »

Figure 2.2 – Dégradation « sans tolérance »

2.3.2 Modélisation de la dégradation d'un composant

Les mécanismes de défaillance sont mis en concurrence pour obtenir un modèle de la dégradation du composant. On commence par considérer que les deux mécanismes sont indépendants. Puis, on proposera un exemple où l'un des deux mécanismes a une influence sur l'autre

Indépendance des mécanismes de défaillance

Considérons maintenant un composant, comme le pignon, soumis à une dégradation avec tolérance et à une dégradation sans tolérance indépendantes l'une de l'autre. Ces deux modes de dégradation sont mis en concurrence, cela signifie que la défaillance du composant aura lieu dès que l'un des deux mécanismes de défaillance aura atteint le dernier état. Le composant peut donc se trouver dans les états (u, v) où u correspond à l'état de la dégradation avec tolérance et v à celui de la dégradation sans tolérance (figure 2.3).

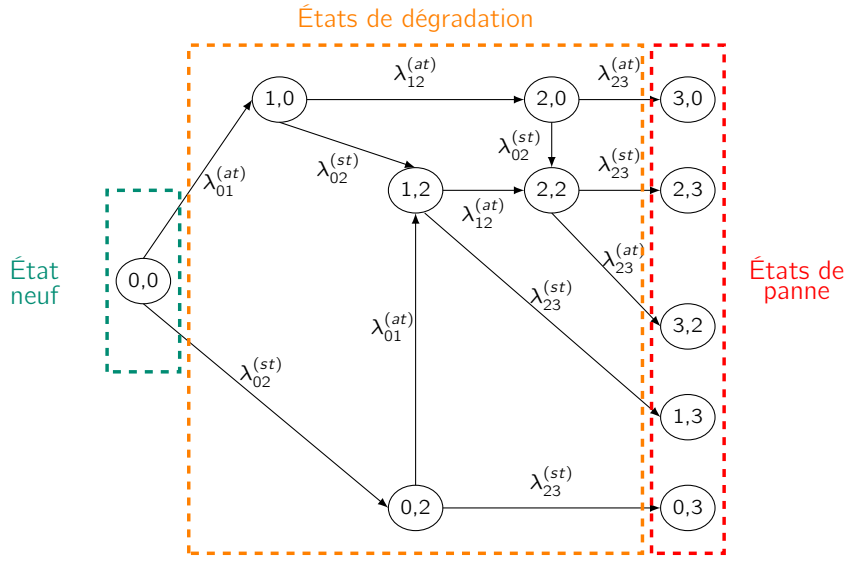


Figure 2.3 – Processus de Markov d'un système à deux mécanismes de défaillance

Les états pour lesquels une des dégradations est dans un état de panne sont des états absorbants. Une fois un état de panne atteint, on procède à une opération de maintenance pour remettre le composant en état de marche. Lorsque les deux dégradations sont markoviennes on vérifie facilement que la probabilité que les deux mécanismes de défaillance conduisent simultanément à la défaillance du composant est nulle, i.e. qu'il n'existe pas de cause commune de défaillance.

Un exemple avec dépendance des mécanismes de défaillance

Le modèle proposé ci-dessus repose donc sur l'hypothèse d'indépendance des mécanismes de défaillance en concurrence vis-à-vis de la défaillance du composant. Cette hypothèse peut être acceptée par avis d'expert qui peut confirmer que les mécanismes de dégradation considérés sont bien distincts et qu'ils n'interagissent pas. Cependant, il n'est pas toujours évident d'avoir un tel avis. On propose ici une solution pour évaluer la présence d'une dépendance entre deux mécanismes de défaillance.

On suppose que le niveau d'une dégradation a un impact sur l'évolution de la seconde dégradation. On va donc modifier le modèle précédent afin de prendre en compte cette dépendance. L'objectif sera de pouvoir diagnostiquer l'existence d'une dépendance entre deux mécanismes de défaillance mis en concurrence vis-à-vis d'un événement redouté.

Par exemple, on peut raisonnablement supposer qu'une usure en dehors des tolérances puisse avoir un impact sur la survenue d'une crique. Si aucune usure n'est présente sur le composant ou si l'usure présente est toujours dans des tolérances acceptables, l'usure n'a pas d'impact sur le risque d'apparition d'une crique. La probabilité d'occurrence d'une crique est donc la même que dans le cas de l'indépendance. Dans ce cas, on considère que le taux de transition $\lambda_{02}^{(st,dep)}$ entre l'état 0 (état neuf du composant) et l'état 2 (dégradation hors tolérance) de la crique est $\lambda_{02}^{(st)}$. Si une usure hors tolérance est présente sur le composant, on suppose qu'elle peut avoir un impact sur le risque d'apparition de crique. Pour traduire cet impact, on considère que le taux de transition $\lambda_{02}^{(st,dep)}$ entre les états 2 et 3 devient égal à $c\lambda_{02}^{(st)}$. Ce taux de transition dépend alors du temps et est constant par morceaux : conditionnellement à $X_{02}^{(at)} = x_{02}^{(at)}$

$$\forall t \geq 0, \quad \lambda_{02}^{(st,dep)}\left(t|x_{02}^{(at)}\right) = \lambda_{02}^{(st)} \ \mathbf{1}\left(t \in [0, x_{02}^{(at)}[\right) + c\lambda_{02}^{(st)} \ \mathbf{1}\left(t \in [x_{02}^{(at)}, +\infty[\right).$$

Cela revient à modifier le taux de transition de l'état $(2,0)$ vers l'état $(2,2)$ sur la figure 2.3 : on a donc effectué une modification locale du graphe de Markov. Le modèle bivarié reste donc markovien. Un modèle bivarié markovien « complètement » dépendant correspond à la situation où tous les taux de transition serait directement définis sur le modèle bivarié (et non pas à partir des processus marginaux comme pour le premier modèle). Un tel modèle ferait intervenir un grand nombre de paramètres. Or, dans les applications industrielles visées, il est préférable d'avoir un modèle parcimonieux, eu égard à des taux de censure élevés. Enfin, notons que le modèle retenu pour la transition de l'état 0 vers l'état 2 est un modèle paramétrique de Cox avec covariable dépendante du temps partiellement observée (rappelons que la durée $X_{02}^{(at)}$ est systématiquement censurée).

2.3.3 Inférence statistique

Afin d'estimer les paramètres du modèle proposé précédemment, on propose d'appliquer la méthode du maximum de vraisemblance. On va donc donner l'expression de la fonction de vraisemblance correspondant aux données disponibles dans le REX.

Modèle avec indépendance des mécanismes

On note n le nombre d'individus observés. Pour le i-ème individu, on note $(t_i, \delta_i^{(at)}, \delta_i^{(st)})$ les observations le concernant : à l'instant t_i, ce composant était dans l'état $\delta_i^{(at)}$ pour le mode d'endommagement avec tolérance et dans l'état $\delta_i^{(st)}$ vis-à-vis du mode d'endommagement sans tolérance. Les deux mécanismes d'endommagement étant supposés indépendants, on obtient la fonction de vraisemblance suivante :

$$L\left(\theta|\text{data}\right) = \prod_{i=1}^{n} f_{Z_{t_i}^{(at)}}\left(t_i, \delta_i^{(at)}|\theta^{(at)}\right) f_{Z_{t_i}^{(st)}}\left(t_i, \delta_i^{(st)}|\theta^{(st)}\right),$$

où $\theta = \left(\theta^{(at)}, \theta^{(st)}\right)$, avec $\theta^{(at)} = \left(\lambda_{01}^{(at)}, \lambda_{12}^{(at)}, \lambda_{23}^{(at)}\right)$ et $\theta^{(st)} = \left(\lambda_{02}^{(st)}, \lambda_{23}^{(st)}\right)$, et où

$$f_{Z_{t_i}^{(at)}}\left(t_i, \delta_i^{(at)}|\theta^{(at)}\right) = \begin{cases} \mathbb{P}[Z_{t_i}^{(at)} = \delta_i^{(at)}] & \text{si } \delta_i^{(at)} \in \{0,1,2\} \\ f_{X_{03}^{(at)}}(t_i) & \text{si } \delta_i^{(at)} = 3 \end{cases}$$

et de manière analogue :

$$f_{Z_{t_i}^{(st)}}\left(t_i, \delta_i^{(st)}|\theta^{(st)}\right) = \begin{cases} \mathbb{P}[Z_{t_i}^{(st)} = \delta_i^{(st)}] & \text{si } \delta_i^{(st)} \in \{0,2\} \\ f_{X_{03}^{(st)}}(t_i) & \text{si } \delta_i^{(st)} = 3 \end{cases}.$$

Détaillons le cas où $(Z_t^{(at)}, Z_t^{(st)}) = (2,0)$ pour un instant t donné (pour les autres cas, voir [64]) :

$$\begin{aligned} f_{Z_t^{(at)}}(2) &= \int_0^t \int_{t-x_{01}^{(at)}}^t \int_{t-x_{01}^{(at)}-x_{12}^{(at)}}^\infty f_{X_{23}^{(at)}}(x_{23}^{(at)}) f_{X_{12}^{(at)}}(x_{12}^{(at)}) f_{X_{01}^{(at)}}(x_{01}^{(at)}) dx_{23}^{(at)} dx_{12^{(at)}} dx_{01}^{(at)} \\ &= \frac{\lambda_{01}^{(at)}\lambda_{12}^{(at)}}{\lambda_{12}^{(at)} - \lambda_{23}^{(at)}} \left(\frac{e^{-\lambda_{12}^{(at)}t} - e^{-\lambda_{01}^{(at)}t}}{\lambda_{12}^{(at)} - \lambda_{01}^{(at)}} + \frac{e^{-\lambda_{23}^{(at)}t} - e^{-\lambda_{01}^{(at)}t}}{\lambda_{01}^{(at)} - \lambda_{23}^{(at)}} \right) \end{aligned}$$

et, de manière évidente,

$$f_{Z_t^{(st)}}(0) = e^{-\lambda_{02}^{(st)}t}.$$

On peut en déduire numériquement l'estimateur du maximum de vraisemblance :

$$\widehat{\theta} = \left(\widehat{\theta}^{(at)}, \widehat{\theta}^{(st)}\right) = \text{argmax}_{\theta \in \mathbb{R}_+^5} L\left(\theta|\text{data}\right).$$

La théorie du maximum de vraisemblance paramétrique permet de justifier l'approximation de la loi de $\sqrt{n}\left(\widehat{\theta} - \theta\right)$ par la loi normale centrée et de matrice de variance-covariance $I^{-1}(\theta)$, ce qui sera noté :

$$\sqrt{n}\left(\widehat{\theta} - \theta\right) \overset{approx.}{\sim} \mathcal{N}\left(0, I^{-1}(\theta)\right).$$

De plus, la matrice d'information de Fisher $I(\theta)$ est estimée par la hessienne de la log-vraisemblance évaluée en $\widehat{\theta}$. On peut alors construire des intervalles de confiance asymptotiques pour chacun des taux (ou des ellipsoïdes de confiance pour plusieurs taux). De manière alternative, on peut construire des intervalles de confiance par bootstrap, surtout en présence d'échantillons de taille petite ou moyenne.

Pour les applications industrielles, on peut utiliser la δ-méthode pour obtenir des intervalles de confiance pour la fiabilité à un instant donné. Pour tout $t \geq 0$, on note $R(t;\theta)$ la fiabilité à l'instant t :

$$\begin{aligned} R(t;\theta) &= \mathbb{P}\left[\min(X_{03}^{(at)}, X_{03}^{(st)}) \geq t\right] = \mathbb{P}\left[X_{03}^{(at)} \geq t\right] \mathbb{P}\left[X_{03}^{(st)} \geq t\right] \\ &= \mathbb{P}\left[\max(Z_t^{(at)}, Z_t^{(st)}) < 3\right]. \end{aligned}$$

Celle-ci est naturellement estimée par $R(t;\widehat{\theta})$. En appliquant la δ-méthode, on obtient que

$$\sqrt{n}\left(R(t;\widehat{\theta}) - R(t;\theta)\right) \overset{approx.}{\sim} \mathcal{N}\left(0, \frac{\partial R}{\partial \theta}(t;\theta) I^{-1}(\theta) \frac{\partial R'}{\partial \theta}(t;\theta)\right)$$

où $\frac{\partial R}{\partial \theta}(t;\theta)$ est estimée par $\frac{\partial R}{\partial \theta}(t;\widehat{\theta})$.

Modèle avec dépendance des mécanismes

En reprenant les mêmes notations, la fonction de vraisemblance est donnée par :

$$L\left(\theta, c|\text{data}\right) = \prod_{i=1}^n f_{(Z_{t_i}^{(at)}, Z_{t_i}^{(st)})}\left(t_i, \delta_i^{(at)}, \delta_i^{(st)}|\theta, c\right),$$

où, pour tout $t \geq 0$, $f_{(Z_t^{(at)}, Z_t^{(st)})}$ est la densité jointe du couple $(Z_t^{(at)}, Z_t^{(st)})$. Il est possible d'expliciter cette densité jointe :

$$f_{(Z_t^{(at)}, Z_t^{(st)})}(x,y) = \begin{cases} \mathbb{P}[Z_t^{(at)} = x, Z_t^{(st)} = y] & \text{si } (x,y) \in \{0,1,2\} \times \{0,2\} \\ \displaystyle\int_0^t f_{X_{02}^{(at)}}(u) f_{X_{23}^{(at)}}(t-u) \mathbb{P}[Z_t^{(st)} = y | X_{02}^{(at)} = u] \mathrm{d}u & \text{si } x = 3 \text{ et } y \in \{0,2\} \\ \mathbb{P}[Z_t^{(at)} = x] f_{X_{03}^{(st)}}(t) & \text{si } x \in \{0,1\} \text{ et } y = 3 \\ \displaystyle\int_0^t \mathbb{P}[X_{23}^{(at)} > t-u] f_{X_{03}^{(st)}|X_{02}^{(at)}=u}(t) f_{X_{02}^{(at)}}(u) \mathrm{d}u & \text{si } x = 2 \text{ et } y = 3. \end{cases}$$

En fait, un certain nombre de contributions restent inchangées par rapport au modèle avec indépendance. En effet, tant que l'état 2 n'est pas atteint pour le mécanisme avec tolérance, celui-ci n'a aucune influence sur le mécanisme sans tolérance et donc les contributions correspondantes sont les mêmes que précédemment. Comme pour le modèle avec indépendance, nous allons détailler le cas où $(Z_t^{(at)}, Z_t^{(st)}) = (2, 0)$ pour un instant t donné (pour les autres cas, voir [64]) :

$$\begin{aligned} f_{(Z_t^{(at)}, Z_t^{(st)})}(2,0) &= \int_0^t \int_{t-x_{02}^{(at)}}^{\infty} \int_t^{\infty} f_{X_{02}^{(at)}}(x_{02}^{(at)}) f_{X_{23}^{(at)}}(x_{23}^{(at)}) f_{X_{02}^{(st)}|X_{02}^{(at)}}(x_{02}^{(st)}|x_{02}^{(at)}) \mathrm{d}x_{02}^{(st)} \mathrm{d}x_{23}^{(at)} \mathrm{d}x_{02}^{(at)} \\ &= \frac{\lambda_{01}^{(at)} \lambda_{12}^{(at)}}{c\lambda_{02}^{(st)}(\lambda_{01}^{(at)} - \lambda_{12}^{(at)})} \left[\frac{e^{-\lambda_{23}^{(at)} t} - e^{-(\lambda_{12}^{(at)} + \lambda_{02}^{(st)})t}}{\lambda_{12}^{(at)} + \lambda_{02}^{(st)} - \lambda_{23}^{(at)}} + \frac{e^{-(\lambda_{01}^{(at)} + \lambda_{02}^{(st)})t} - e^{-\lambda_{01}^{(at)} t}}{\lambda_{12}^{(at)} + \lambda_{02}^{(st)} - \lambda_{23}^{(at)}} \right]. \end{aligned}$$

Ici également, on peut en déduire numériquement l'estimateur du maximum de vraisemblance :

$$(\widehat{\theta}, \widehat{c}) = \mathrm{argmax}_{(\theta,c)\in\mathbb{R}_+^6} L(\theta, c|\mathrm{data}).$$

De même que précédemment, la théorie du maximum de vraisemblance paramétrique permet de justifier l'approximation normale de la loi de cet estimateur. On pourra, en particulier, construire un intervalle de confiance de niveau $\alpha \in]0,1[$ pour le paramètre c : ainsi si 1 appartient à cet intervalle, on pourra accepter l'hypothèse d'indépendance des mécanismes de dégradation au risque de première espèce α.

2.3.4 Evaluation d'une politique de maintenance périodique

On dispose maintenant d'un modèle pour la dégradation d'un composant soumis à plusieurs modes de défaillance, qui tient compte de l'information disponible dans le REX. Mais les paramètres estimés via ce modèle dépendent des données de dégradation observées sur des composants soumis à une politique de maintenance. Autrement dit, à ce stade, on dispose d'un modèle pour la dégradation d'un composant pour la politique de maintenance en cours. Or, pour pouvoir envisager d'autres politiques de maintenance, il est nécessaire de disposer d'un modèle de dégradation « intrinsèque » du composant, c'est-à-dire d'un modèle de dégradation du composant non soumis à une politique de maintenance. C'est l'objectif de cette partie. On fera l'hypothèse que la dégradation intrinsèque peut être modélisée par le même processus bivariée.

Nous nous limitons ici au cas d'endommagements indépendants. En conséquence, pour simplifier la présentation de la méthodologie, nous ne considérons qu'un seul endommagement (« avec tolérance »). On commence par présenter un modèle de dégradation qui prend en compte la politique de maintenance périodique appliquée et qui est implicitement inclus dans les données du REX.

Modèle de dégradation avec maintenance périodique

On suppose désormais que les composants subissent une maintenance préventive toutes les M heures de service. La politique de remplacement appliquée reste la même que celle décrite précédemment : si lors de l'inspection un endommagement hors tolérance est constaté, le composant est remis à neuf. Le modèle suivant intègre donc cet aspect ignoré dans le premier modèle. Le modèle de dégradation devient donc un processus de renouvellement markovien. On peut donc calculer, de manière récursive, la fiabilité du composant pour ce modèle. On notera $\tilde{\theta} = (\tilde{\lambda}_{01}, \tilde{\lambda}_{12}, \tilde{\lambda}_{23})'$ l'ensemble des paramètres de ce second modèle.

Soit $P_t = (p_t(i,j))_{(i,j)\in\{0,1,2,3\}^2}$ la matrice de transition du processus markovien de sauts modélisant l'évolution de l'endommagement. En particulier, P_M indique les probabilités de transition entre la mise en service du composant et sa première inspection. La matrice P_t peut être calculée explicitement, soit de manière directe, soit en résolvant les équations de Chapman-Kolmogorov ($P_t' = AP_t$). En supposant tous les taux de transition distincts, nous obtenons ainsi les probabilités de transition dont les termes utiles au

calcul de la fiabilité du composant sont les suivants : pour tout $t \in \mathbb{R}^+$,

$$\begin{aligned}
p_t(0,0) &= e^{-\tilde{\lambda}_{01}t}, \qquad p_t(0,1) = \frac{\tilde{\lambda}_{01}}{\tilde{\lambda}_{12}-\tilde{\lambda}_{01}}\left(e^{-\tilde{\lambda}_{01}t} - e^{-\tilde{\lambda}_{12}t}\right), \\
p_t(0,2) &= \frac{\tilde{\lambda}_{01}\tilde{\lambda}_{12}}{\tilde{\lambda}_{23}-\tilde{\lambda}_{12}}\left(\frac{-e^{-\tilde{\lambda}_{12}t}+e^{-\tilde{\lambda}_{01}t}}{\tilde{\lambda}_{12}-\tilde{\lambda}_{01}} + \frac{e^{-\tilde{\lambda}_{23}t}-e^{-\tilde{\lambda}_{01}t}}{\tilde{\lambda}_{23}-\tilde{\lambda}_{01}}\right) \\
&\qquad -\frac{\tilde{\lambda}_{01}\tilde{\lambda}_{12}}{\tilde{\lambda}_{23}-\tilde{\lambda}_{12}}\left(\frac{-e^{-\tilde{\lambda}_{12}t}+e^{-\tilde{\lambda}_{01}t}}{\tilde{\lambda}_{12}-\tilde{\lambda}_{01}} + \frac{e^{-\tilde{\lambda}_{23}t}-e^{-\tilde{\lambda}_{01}t}}{\tilde{\lambda}_{23}-\tilde{\lambda}_{01}}\right), \\
p_t(1,1) &= e^{-\tilde{\lambda}_{12}t}, \qquad p_t(1,2) = \frac{\tilde{\lambda}_{12}}{\tilde{\lambda}_{23}-\tilde{\lambda}_{12}}\left(e^{-\tilde{\lambda}_{12}t} - e^{-\tilde{\lambda}_{23}t}\right).
\end{aligned}$$

On note $R_0(t;\tilde{\theta})$ (resp. $R_1(t;\tilde{\theta})$) la fiabilité du composant à l'instant t sachant qu'il était initialement dans l'état 0 (resp. dans l'état 1) :

$$R_0(t;\tilde{\theta}) = \mathbb{P}\left[Z_t^{(at)} \neq 3 \,\middle|\, Z_0^{(at)} = 0\right] \quad \text{et} \quad R_1(t;\tilde{\theta}) = \mathbb{P}\left[Z_t^{(at)} \neq 3 \,\middle|\, Z_0^{(at)} = 1\right].$$

Enfin, soit $R_{main}(t;\tilde{\theta},M)$ la fiabilité du composant à l'instant t. Comme on applique une maintenance préventive toutes les M heures, on a :

$$\forall t \in [Mm,(M+1)m[, \quad R_{main}(t;\tilde{\theta},M) = R_{main}^{(m)}(t;\tilde{\theta},M),$$

où $R_{main}^{(m)}(t;\tilde{\theta},M)$ correspond à la fiabilité du composant sachant qu'il a déjà eu m maintenances avant l'instant t. Dans le cas où $m = 0$, et sous l'hypothèse que le composant est initialement neuf, on a $R_{main}^{(0)}(t;\tilde{\theta},M) = R_0(t;\tilde{\theta})$. Par récurrence, nous obtenons l'expression suivante :

$$R_{main}^{(m)}(t;\tilde{\theta},M) = r_{00}^{(m)}(t;\tilde{\theta},M) + r_{01}^{(m)}(t;\tilde{\theta},M),$$

où

$$\begin{aligned}
r_{00}^{(1)}(t;\tilde{\theta},M) &= (p_M(0,0)+p_M(0,2))R_0(t-M;\tilde{\theta}), \\
r_{01}^{(1)}(t;\tilde{\theta},M) &= p_M(0,1)R_1(t-M;\tilde{\theta}), \\
r_{11}^{(1)}(t;\tilde{\theta},M) &= p_M(1,1)R_1(t-M;\tilde{\theta}), \\
r_{12}^{(1)}(t;\tilde{\theta},M) &= p_M(1,2)R_0(t-M;\tilde{\theta}).,
\end{aligned}$$

et pour $m \geq 2$

$$\begin{aligned}
r_{00}^{(m)}(t;\tilde{\theta},M) &= (p_M(0,0)+p_M(0,2))(r_{00}^{(m-1)}(t;\tilde{\theta},M) + r_{01}^{(m-1)}(t;\tilde{\theta},M)), \\
r_{01}^{(m)}(t;\tilde{\theta},M) &= p_M(0,1)(r_{11}^{(m-1)}(t;\tilde{\theta},M) + r_{12}^{(m-1)}(t;\tilde{\theta},M)), \\
r_{11}^{(m)}(t;\tilde{\theta},M) &= p_M(1,1)(r_{11}^{(m-1)}(t;\tilde{\theta},M) + r_{12}^{(m-1)}(t;\tilde{\theta},M)), \\
r_{12}^{(m)}(t;\tilde{\theta},M) &= p_M(1,2)(r_{00}^{(m-1)}(t;\tilde{\theta},M) + r_{01}^{(m-1)}(t;\tilde{\theta},M)).
\end{aligned}$$

Méthodologie d'évaluation de la dégradation intrinsèque

On décrit ici la méthodologie proposée pour estimer les paramètres du modèle de dégradation intrinsèque et étudier l'impact d'une autre périodicité de maintenance préventive. Cette méthodologie est schématisée sur la figure 2.4 pour le cas d'un composant soumis à deux mécanismes de dégradation.

Étape 1 On commence par estimer la fiabilité du composant à partir des données du REX à l'aide du premier modèle. Nous estimons ainsi la fiabilité du composant correspondant à la maintenance préventive appliquée en service (nous rappelons que les effets de la maintenance préventive sont implicitement inclus dans les données). On note $R_{rex}(\cdot;\widehat{\theta})$ l'estimation de la fiabilité à l'instant t.

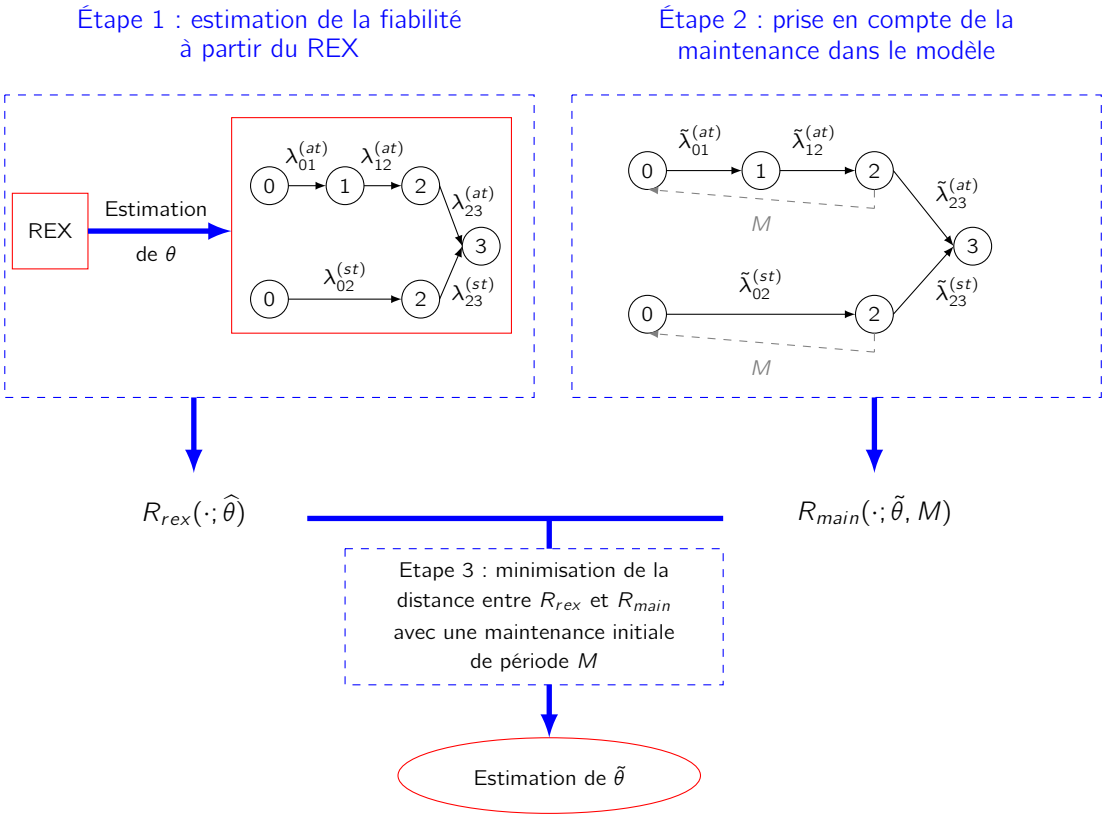


Figure 2.4 – Résumé de la méthodologie proposée

Étape 2 On estime ensuite les paramètres du second modèle en minimisant la distance (par exemple, celle de Cramér-von Mises) entre les fiabilités fournies par les deux modèles (pour la périodicité de maintenance appliquée en service). On va donc déterminer les paramètres $\tilde{\theta}$ tels que les fiabilités $R_{rex}(\cdot;\widehat{\theta})$ et $R_{main}(\cdot;\tilde{\theta},M)$ soient les plus proches, où M représente la périodicité de la maintenance préventive appliquée en service. On propose d'utiliser la distance de Cramér-von Mises (notée d_{CvM}) définie par :

$$d_{CvM}(\tilde{\theta},M) \quad = \quad \int_0^{+\infty} \left(R_{main}(t;\tilde{\theta},M) - R_{rex}(t;\widehat{\theta})\right)^2 f_{rex}(t;\widehat{\theta})\,\mathrm{d}t$$

où f_{rex} (resp. F_{rex}) représente l'estimation de la fonction de densité (resp. fonction de répartition) du temps de fonctionnement du composant estimé à partir des données de retour d'expérience. On obtient donc :

$$\widehat{\tilde{\theta}} = \mathrm{argmin}_{\tilde{\theta}\in\mathbb{R}_+^3}\, d_{CvM}(\tilde{\theta},M),$$

où, en pratique, la distance $d_{CvM}(\tilde{\theta},M)$ est estimée par la méthode de Monte-Carlo.

Ce choix de méthode d'estimation des paramètres du modèle de dégradation avec maintenance résulte de la nature du REX qui ne permet pas directement un ajustement du modèle incluant la maintenance à partir des données. En effet, les données disponibles sont obtenues suite à une dépose d'un moteur en centre de réparation. Les raisons de cette dépose peuvent être diverses, par exemple, suite à l'allumage d'un voyant d'alerte, à un événement en service ou encore lors d'une maintenance préventive. Cependant, lors d'une dépose pour une raison quelconque, une maintenance est effectuée sur le composant incriminé, mais un autre composant peut également être inspecté de part sa proximité ou parce qu'il semble endommagé. Il n'est donc pas toujours possible d'identifier la raison de la maintenance effectuée sur un composant : la dépose est-elle due à une maintenance programmée ou à une inspection opportuniste ? Nous ne pouvons donc pas estimer directement le paramètre $\tilde{\theta}$ sur les données issues du REX puisque nous ne savons pas quelles sont les données correspondant à un retour dû à une maintenance préventive.

Chapitre 3

Modèles de dégradation continue

La plupart des processus utilisés pour modéliser une dégradation continue appartiennent à la famille des processus de Lévy (processus à accroissements indépendants et stationnaires). Plus particulièrement, trois processus sont les plus fréquemment employés : le mouvement brownien avec tendance linéaire (positive), les processus de Poisson composés et le processus gamma. Seul le premier processus est à trajectoires continues presque sûrement, les deux autres étant des processus à sauts positifs (on parle de processus de Lévy spectralement positifs). Ceci justifie en partie que le mouvement brownien avec tendance linéaire a été utilisé pour modéliser une dégradation bien que n'étant pas à trajectoires monotones. Le processus gamma peut être interprété comme la limite d'un processus de Poisson composé. Pour un état de l'art sur l'utilisation du processus gamma (avec applications en maintenance), on pourra consulter l'article de van Noortwijk [184]. Les processus de Lévy ont tous une tendance linéaire. Cela constitue une hypothèse parfois trop restrictive quand on s'intéresse à la modélisation d'une dégradation. Donc, dans certains cas, on essaye de relâcher l'hypothèse de stationnarité des accroissements. Une solution possible consiste à dilater l'échelle de temps d'un processus de Lévy par une fonction déterministe, croissante et bijective sur $\mathbb{R}_+$.

Il est d'usage d'associer à ces processus de dégradation un temps de défaillance défini comme le temps de franchissement d'un seuil considéré comme critique (dans le sens, par exemple, où au-delà de ce seuil le niveau de sécurité n'est plus garanti). Ainsi, pour un modèle de dégradation donné, on peut distinguer trois types de question.

1. Loi du temps d'atteinte d'un seuil (ce seuil est généralement supposé constant et déterministe, mais on pourra aussi envisager le cas d'un seuil critique aléatoire indépendant du processus de dégradation) : on s'intéresse dans ce cas à la loi du temps de panne et à d'éventuelles propriétés de vieillissement (IFR, IFRA, NBU, etc).
2. Estimation des paramètres du modèle : il est assez naturel de s'intéresser à l'estimation des paramètres d'un modèle de dégradation. Il s'agit de problèmes d'inférence paramétrique, semi-paramétrique ou non-paramétrique selon le modèle considéré. Les observations correspondent, en général, à des mesures de dégradation d'un ou de plusieurs processus indépendants à un ou plusieurs instants.
3. Définition et optimisation d'une politique de maintenance : pour un modèle de dégradation donné et un temps de défaillance associé, il est possible de définir une politique de maintenance. Cette politique de maintenance dépend de coûts associés à chaque type d'actions possibles et de paramètres que l'on peut chercher à optimiser sur la base d'un critère (fonction de coût à minimiser). En général, on considère le coût unitaire asymptotique en utilisant la théorie asymptotique du renouvellement.

Dans ce chapitre, plusieurs modèles de dégradation continue sont étudiés. Dans la première section, on considère le processus gamma non-homogène. Dans le cadre de la thèse d'A. Salami, nous nous sommes intéressés à la loi du temps d'atteinte d'un seuil critique pour ce processus, le seuil critique pouvant être déterministe ou aléatoire (et indépendant du processus de dégradation). La deuxième section est consacrée à l'étude statistique du processus gamma perturbé par un mouvement brownien (supposé indépendant du processus gamma) sans ou avec covariables. Ce travail, en collaboration avec L. Bordes (UPPA), fait également partie des travaux de thèse d'A. Salami. La troisième section porte sur une généralisation du modèle

précédent (mais sans covariables). En effet, avec L. Rabehasaina (Université de Franche-Comté), nous avons considéré un subordinateur perturbé (par un mouvement brownien) comme modèle de dégradation. Pour ce modèle, nous avons étudié le problème du temps de passage d'un seuil critique déterministe. Comme les trajectoires d'un tel processus ne sont pas monotones, il y a deux choix possibles pour définir un temps de panne, suivant la discussion de Barker et Newby [58] : le premier ou le dernier temps de passage du niveau critique, le premier cas étant le choix classique. Les lois de ces deux temps de passage sont étudiées dans cette section. Dans la quatrième section, toujours en collaboration avec L. Rabehasaina, nous avons considéré un processus gamma évoluant dans un environnement aléatoire. Cet environnement est modélisé par un processus markovien de saut binaire et influence le processus gamma à travers sa fonction d'échelle linéaire par morceaux (la pente dépend du processus binaire). Enfin, dans la dernière section, nous étudions le cas d'un processus de Wiener retardé. En effet, dans certains cas, le composant ne commence pas à se dégrader tout de suite une fois en service, mais uniquement après une période de latence aléatoire.

3.1 Processus gamma non-homogène [16]

Pour commencer, on rappelle la définition du processus gamma afin de bien fixer les notations. Soit $\xi \in \mathbb{R}^+$ et $\eta = (\eta_t)_{t\geqslant 0}$ une fonction croissante telle que $\eta_0 = 0$ et $\lim_{t\to\infty} \eta_t = \infty$. On dit que (D_t) est un processus gamma s'il satisfait : (1) $D_0 = 0$; (2) ses accroissements sont indépendants ; (3) ses accroissements sont de loi gamma. Plus précisément, pour tout t et δ, $D_{t+\delta} - D_t$ est une variable aléatoire de loi gamma de paramètre $(\eta_{t+\delta} - \eta_t, \xi)$. En particulier, cela implique que toutes les marginales sont de loi gamma. Pour tout $t \geqslant 0$, la densité D_t est donnée par :

$$f_{D_t}(x) = \frac{1}{\xi\Gamma(\eta_t)} \left(\frac{x}{\xi}\right)^{\eta_t - 1} e^{-x/\xi} 1_{\mathbb{R}^+}(x),$$

où $\Gamma(\cdot)$ est la fonction gamma. L'espérance de D_t est donc égale à $\xi\eta_t$ et la variance de D_t à $\xi^2\eta_t$.

Un cas particulier souvent considéré dans la littérature est celui de la fonction de forme η linéaire ($\eta_t = \alpha t$). Dans ce cas très particulier, (D_t) est un processus à accroissements stationnaires : pour tous t et δ, $D_{t+\delta} - D_t$ et D_t ont la même loi. Alors, (D_t) devient un processus de Lévy. Quelques cas non linéaires ont également été étudiés. Le cas non linéaire le plus fréquemment étudié est celui où $\eta_t = \alpha t^\beta$, souvent sur la base d'études empiriques qui permettent de justifier ce choix. Cependant, pour des problèmes d'inférence statistique sur ce modèle, le paramètre β est souvent supposé connu et donné par un jugement d'expert. Récemment, Wang a proposé un estimateur semi-paramétrique pour un processus gamma non homogène, avec effets aléatoires [186] ou (éventuellement) en présence de covariables [185]. On considère ci-dessous la loi du temps d'atteinte d'un niveau critique par un tel processus. Le niveau sera supposé déterministe dans un premier temps, puis aléatoire dans un second temps.

3.1.1 Temps d'atteinte d'un seuil déterministe

Soit c une constante positive correspondant au seuil critique de dégradation. On considère alors le temps de panne T_c associé :

$$T_c = \inf \{t \geqslant 0 \ ; \ D_t \geqslant c\} .$$

Puisque le processus stochastique (D_t) est croissant, on a la relation suivante :

$$\forall t \geqslant 0 , \quad \mathbb{P}[T_c > t] = \mathbb{P}[D_t < c].$$

Le théorème ci-dessous généralise un résultat de Park et Padgett [145] (qui avaient uniquement considéré le cas linéaire) :

Théorème 3.1.1 *La fonction de répartition de T_c est donnée par :*

$$\forall t \geqslant 0 , \quad F_{T_c}(t) = \frac{\Gamma(\eta_t, c/\xi)}{\Gamma(\eta_t)},$$

où $\Gamma(\cdot,\cdot)$ est la fonction gamma incomplète supérieure. De plus, si η est dérivable, la densité de T_c est :

$$\forall t \geqslant 0\ , \quad f_{T_c}(t) = \eta_t' \left(\Psi(\eta_t) - \log\left(\frac{c}{\xi}\right)\right) \frac{\gamma(\eta_t, c/\xi)}{\Gamma(\eta_t)} + \frac{\eta_t'}{\eta_t^2 \Gamma(\eta_t)} \left(\frac{c}{\xi}\right)^{\eta_t} {}_2F_2(\eta_t, \eta_t; \eta_t+1, \eta_t+1; -c/\xi),$$

où Ψ est la fonction di-gamma (ou dérivée logarithmique de la fonction gamma), $\gamma(\cdot,\cdot)$ est la fonction gamma incomplète inférieure et ${}_2F_2$ est la fonction hypergéométrique généralisée d'ordre $(2,2)$.

D'autres expressions ont été proposées dans le cas stationnaire (voir l'état de l'art sur le processus gamma proposé par van Noortwijk [184] et les références dedans). Partant de remarques faites par Frenk et Nicolai [93], nous avons la relation suivante : $T_c \overset{(d)}{=} \eta^{-1}(\tilde{T}_c)$ où η^{-1} est la fonction réciproque de η et où $\tilde{T}_c$ est le temps de panne associé au processus gamma homogène de fonction de forme t et de paramètre d'échelle ξ. On en déduit alors une approximation un peu grossière du temps moyen avant panne (MTTF) :

$$\mathbb{E}[T_c] = \mathbb{E}[\eta^{-1}(\tilde{T}_c)] \simeq \eta^{-1}(\mathbb{E}[\tilde{T}_c]).$$

De plus, Bérenguer *et al.* [61] ont obtenu l'approximation suivante :

$$\mathbb{E}[\tilde{T}_c] \simeq \frac{c}{\xi} + \frac{1}{2}.$$

En utilisant ces deux approximations, on obtient :

$$\mathbb{E}[T_c] \simeq \eta^{-1}\left(\frac{c}{\xi} + \frac{1}{2}\right).$$

Comme remarqué par Park et Padgett [145], la loi de Birnbaum-Saunders peut être utilisée pour approcher la loi ci-dessus dans le cas linéaire. En effet, puisque la loi gamma est infiniment divisible, on peut utiliser le théorème de la limite centrale pour obtenir une approximation. Soit $t \geqslant 0$ fixé et $Y_i = D_{t_i} - D_{t_{i-1}}$ avec $t_i = it/n$ pour $i \in \{0, \ldots, n\}$: Y_i est de loi gamma de paramètres $(\xi, \eta_{t_i} - \eta_{t_{i-1}})$. En appliquant le théorème de la limite centrale de Lindeberg-Feller, on a :

$$\frac{D_t - \xi\eta_t}{\xi\sqrt{\eta_t}} = \frac{\sum_{i=1}^n Y_i - \xi\eta_t}{\xi\sqrt{\eta_t}} \xrightarrow[n\to\infty]{d} \mathcal{N}(0,1).$$

D'où il résulte l'approximation suivante :

$$\mathbb{P}[T_c \leqslant qt] = \mathbb{P}\left[\frac{\sum_{i=1}^n Y_i - \xi\eta_t}{\xi\sqrt{\eta_t}} \geqslant q\frac{c-\xi\eta_t}{\xi\sqrt{\eta_t}}\right] \simeq 1 - \Phi\left(\frac{c-\xi\eta_t}{\xi\sqrt{\eta_t}}\right) = \Phi\left(\sqrt{\frac{c}{\xi}}\left(\sqrt{\frac{\xi\eta_t}{c}} - \sqrt{\frac{c}{\xi\eta_t}}\right)\right),$$

où Φ est la fonction de répartition de la loi normale centrée réduite. On retrouve bien la loi de Birnbaum-Saunders dans le cas linéaire.

3.1.2 Temps d'atteinte d'un seuil aléatoire

Depuis un article d'Abdel-Hameed paru en 1975 [47], quelques auteurs ont étudié le problème de la loi du temps d'atteinte d'un niveau aléatoire. On supposera ici que le seuil critique C est une variable aléatoire indépendante du processus gamma (D_t).

Seuil de loi exponentielle

Partant de l'expression de la fonction de répartition donnée dans le théorème 3.1.1, on peut en déduire l'expression de celle de T_C lorsque C est une variable aléatoire de loi exponentielle.

Théorème 3.1.2 *Supposons que C est une variable aléatoire de loi exponentielle d'espérance λ. La fonction de répartition de T_C est donnée par :*

$$\forall t \geqslant 0\ , \quad F_{T_C}(t) = 1 - (1 + \xi/\lambda)^{-\eta_t},$$

De plus, si η est dérivable, la densité de T_C est :

$$\forall t \geqslant 0\ , \quad f_{T_C}(t) = \eta_t'(1 + \xi/\lambda)^{-\eta_t} \log(1 + \xi/\lambda).$$

On regarde maintenant brièvement quelques propriétés de vieillissement dans ce cas là. On commence par rappeler les notions classiques.

Définition 3.1.1 Soit X une variable aléatoire positive de fonction de répartition F. Soit R sa fonction de survie, $H = -\ln R$ son taux de risque cumulé et h son taux de risque instantané (s'il existe).

1. (IFR/DFR) F est à taux de risque croissant (resp. décroissant) si et seulement si H est convexe (resp. concave). Si h existe, cela équivaut à la croissance (resp. décroissance) de h.
2. (IFRA/DFRA) F est à taux de risque croissant (resp. décroissant) en moyenne si et seulement si H est étoilée (resp. anti-étoilée), i.e. si et seulement si $t \mapsto H(t)/t$ est croissante (resp. décroissante).
3. (NBU/NWU) F est meilleure (resp. pire) neuve que vieille si et seulement si H est sur-additive (resp. sous-additive), i.e. si et seulement si, pour tout $(t,s) \in \mathbb{R}_+^2$, $H(t+s) \geqslant H(t)+H(s)$ (resp. $H(t+s) \leqslant H(t)+H(s)$).

Ces notions sont reliées entre elles de la manière suivante. D'un coté, on a IFR $\Longrightarrow$ IFRA $\Longrightarrow$ NBU ; et, d'un autre coté, on a DFR $\Longrightarrow$ DFRA $\Longrightarrow$ NWU.

Dans le cas d'un seuil aléatoire de loi exponentielle, les propriétés de vieillissement découlent immédiatement du théorème précédent puisque le taux de risque cumulé est égal à la fonction de forme η, à une constante près.

Proposition 3.1.1 *La loi de T_C, lorsque C est de loi exponentielle, est*

1. *IFR (resp. DFR) si et seulement si η est convexe (resp. concave) ;*
2. *IFRA (resp. DFRA) si et seulement si η est étoilée (resp. anti-étoilée) ;*
3. *NBU (resp. NWU) si et seulement si η est sur-additive (resp. sous-additive) .*

Ces résultats peuvent en fait être aussi obtenus en appliquant le théorème 1 dans [47], puisque la loi exponentielle est à la fois IFR et DFR, IFRA et DFRA, etc.

Seuil de loi gamma

Le cas d'un seuil de loi gamma peut également être calculé de manière similaire.

Théorème 3.1.3 *Supposons que C est une variable aléatoire de loi gamma de paramètres (λ,μ). La fonction de répartition de T_C est donnée par :*

$$F_{T_C}(t) = \frac{\Gamma(\eta_t+\mu)}{\Gamma(\mu+1)\Gamma(\eta_t)}\left(1+\frac{\xi}{\lambda}\right)^{-\eta_t}\left(1+\frac{\lambda}{\xi}\right)^{-\mu}{}_2F_1\left(1,\eta_t+\mu;\mu+1;\frac{\xi}{\lambda+\xi}\right).$$

De plus, si η est dérivable, la densité de T_C est donnée, pour tout $t \geqslant 0$, par

$$f_{T_C}(t) = \frac{\eta_t'\Gamma(\eta_t+\mu)}{\Gamma(\mu+1)\Gamma(\eta_t)}\left(\frac{\xi}{\xi+\lambda}\right)^{\mu}\left(\frac{\lambda}{\xi+\lambda}\right)^{\eta_t}\left[\left(\psi(\eta_t+\mu)-\psi(\eta_t)\right.\right.$$
$$\left.\left.-\log\left(\frac{\xi+\lambda}{\lambda}\right)\right){}_2F_1\left(1,\eta_t+\mu;\mu+1;\frac{\xi}{\xi+\lambda}\right)+\frac{\xi}{\lambda+\xi}\frac{1}{\mu+1}F^{2:2,1}_{2:1,0}\left[\begin{array}{c}\eta_t+\mu+1,2:1,\eta_t;1;\\ 2,\mu+2:\eta_t+1;-;\end{array}\quad\frac{\xi}{\lambda+\xi},\frac{\xi}{\lambda+\xi}\right]\right],$$

où $F^{2:2,1}_{2:1,0}$ est la fonction de Kampé de Fériet [50].

Frenk et Nicolai [93] ont obtenu une expression différente (à une erreur près) pour la fonction de répartition T_C quand $\mu \in \mathbb{N}$:

$$F_{T_C}(t) = 1-\left(1+\frac{\xi}{\lambda}\right)^{-\eta_t}\sum_{j=0}^{\mu-1}\binom{\eta_t+j-1}{j}\left(1+\frac{\lambda}{\xi}\right)^{j}.$$

En appliquant le théorème 1 dans [47], on peut en déduire des propriétés de vieillissement.

Proposition 3.1.2 *La loi de T_C, lorsque C est de loi gamma de paramètres (λ,μ), est*

1. *IFR (resp. DFR) si et seulement si η est convexe (resp. concave) et $\mu > 1$ (resp. $\mu < 1$) ;*
2. *IFRA (resp. DFRA) si et seulement si η est étoilée (resp. anti-étoilée) et $\mu > 1$ (resp. $\mu < 1$) ;*
3. *NBU (resp. NWU) si et seulement si η est sur-additive (resp. sous-additive) et $\mu > 1$ (resp. $\mu < 1$).*

La loi gamma peut être utilisée pour approcher la loi de Dirac en un point donné. Donc, on peut utiliser la proposition ci-dessus pour obtenir des propriétés de vieillissement dans le cas d'un seuil constant déterministe. En effet, la limite d'une suite de fonctions convexes, si elle existe, est convexe [70].

Proposition 3.1.3 *Soit $c > 0$ un seuil fixé. La loi T_c est :*

1. *IFR si et seulement si η est convexe ;*
2. *IFRA si et seulement si η est étoilée ;*
3. *NBU si et seulement si η est sur-additive.*

Les calculs ne sont pas aussi aisés dans le cas général. Une solution consiste à utiliser les lois de type phase pour approcher une loi positive donnée pour C [140]. En effet, soit C une variable aléatoire positive et C_n une variable aléatoire de loi de type phase telle que C_n converge en loi vers C quand n tend vers l'infini. On en déduit que T_{C_n} converge en loi vers T_C quand n tend vers l'infini. Les deux propositions précédentes peuvent donc être utilisées à ces fins-là. En effet, il existe un certain nombre d'articles dans la littérature sur la façon dont les mélanges de lois exponentielles (appelés lois hyper-exponentielles) et les mélanges de loi gamma peuvent approcher une loi positive [68]. De nombreux logiciels ou packages de logiciel ont été développés pour ajuster une loi de type phase à une loi donnée. Par exemple, Asmussen *et al.* [54] ont proposé un algorithme EM. Malhotra et Reibman [129] ont étudié l'approximation par des lois de type phase pour des lois apparaissant plus spécialement dans les modèles semi-markoviens : loi de Dirac, loi log-normale et loi de Weibull. Ils affirment que la loi log-normale peut être bien approchée par un mélange de lois d'Erlang (chacune des lois d'Erlang ayant le même nombre de phase pour réduire le nombre de paramètres). Ils ont aussi considéré la loi d'Erlang pour approcher la loi de Weibull avec taux de risque croissant. Pour une loi de Weibull avec taux de risque décroissant, une loi hyper-exponentielle peut être considérée, voir [88] ou [174] (et leurs illustrations numériques).

3.2 Processus gamma perturbé [6, 43]

La principale partie de la thèse d'Ali Salami a porté sur l'inférence statistique du processus gamma perturbé par un mouvement brownien (indépendant du processus gamma). On considère le modèle de dégradation suivant :

$$\forall t \geqslant 0, \quad D(t) = G(t) + \sigma B(t),$$

où G est un processus gamma homogène de paramètre de forme α et de paramètre d'échelle ξ, c est un mouvement brownien indépendant de G et σ une constante (qu'on peut supposer être positive sans perte de généralité en utilisant la propriété de symétrie du mouvement brownien). Pour être plus précis, G_1 est une variable aléatoire de loi gamma de paramètre (α, ξ) dont la densité est :

$$f(x) = \frac{\xi^{\alpha t}}{\Gamma(\alpha t)} y^{\alpha t - 1} e^{-\xi y} \, \mathbb{I}_{\{y \geqslant 0\}}.$$

Par rapport au modèle précédent, on a donc une fonction d'échelle linéaire, i.e. $\eta_t = \alpha t$.

Ce modèle peut être justifié de deux manières distinctes selon le regard que l'on porte sur un modèle. Du point de vue de la modélisation même, on peut voir la perturbation brownienne soit comme reflétant les erreurs de mesure (la dégradation n'est pas mesurée parfaitement, mais elle est bruitée), soit comme reflétant des petites réparations (action de maintenance de bas niveau) ou des petites dégradations (en plus de la dégradation intrinsèque modélisée par le processus gamma). Il n'est pas rare d'être confronté à des données correspondant à des trajectoires de dégradations non toujours croissantes. Du point de vue de la statistique, ce modèle peut être utilisé à des fins de sélection de modèles. En effet, ce modèle englobe deux des approches les plus classiques dans le domaine : tout d'abord, on a évidemment le processus gamma

pur si $\sigma = 0$; ensuite, lorsque α/ξ tend vers une constante $\mu > 0$ et que α/ξ^2 tend vers zéro, alors le processus gamma est dégénéré en devenant une droite affine de pente μ et donc on obtient le mouvement brownien avec tendance linéaire (positive). Ainsi, on pourra construire des tests afin de choisir parmi un de ses modèles.

L'environnement dans lequel un système fonctionne influe en général sur la dégradation de celui-ci. Il est donc naturel d'intégrer des covariables dans un modèle de dégradation. Différentes manières de les inclure dans des modèles ont été proposés. Nous avons déjà vu (voir la partie précédente) l'approche proposée par Lawless et Crowder [125]. Ici, nous retenons l'approche proposée par Bagdonavičius et Nikulin [55] dans le cas du processus gamma pur. Nous nous sommes restreints au cas de covariables ne dépendant pas du temps et nous avons supposé que les covariables n'agissent que sur la dégradation liée au processus gamma. Si $x = \left(x^{(1)}, \ldots, x^{(p)}\right)^T$ est un vecteur de covariables, alors, conditionnellement à x, alors le modèle précédent devient :

$$\forall t \geqslant 0,\ D_x(t) = G\left(te^{\beta^T x}\right) + \sigma B(t),$$

où $\beta = (\beta_1, \ldots, \beta_p)^T$ est un vecteur de paramètres inconnus. Il en résulte que G est désormais un processus gamma stationnaire de paramètres d'échelle ξ et de forme $\alpha e^{\beta^T x}$. De plus, on suppose que le vecteur de covariables x est une observation de vecteur X de densité f_X par rapport à une mesure σ-finie μ_p de $\mathbb{R}^p$.

3.2.1 Inférence dans le modèle sans covariable [6]

On suppose qu'on observe la dégradation de n systèmes indépendants et identiques. Deux situations ont été étudiées. Dans le premier cas, on considère que les systèmes sont observés au(x) même(s) instant(s) et on notera N le nombre de ces instants. Dans ce cas, les accroissements sont des variables aléatoires indépendantes et identiquement distribuées. On peut donc facilement appliquer la méthodes des moments pour estimer les paramètres du modèle : $\theta = (\xi, \alpha, \tau^2)$. Deux régimes asymptotiques sont considérés : 1) n tend vers l'infini et N est fixé ; 2) n et N tendent vers l'infini. Dans la seconde situation, on considère que le cas général où le nombre d'observations par système ne sont pas nécessairement identiques et l'asymptotique porte uniquement sur n.

On suppose que l'on observe n processus gamma perturbés indépendants et de même loi, notés $D^{(1)}, \ldots, D^{(n)}$. Le i-ème processus est observé N_i fois aux instants $0 = t_{i,0} < t_{i,1} < \cdots < t_{i,N_i}$. On pose $\Delta t_{ij} = t_{i,j} - t_{i,j-1}$ et $\Delta D_{ij} = D^{(i)}_{t_{i,j}} - D^{(i)}_{t_{i,j-1}}$ qui sont des variables aléatoires indépendantes. On note $m^{(k)}_{ij}$ les moments non-centrés :

$$m^{(k)}_{ij} = \mathbb{E}\left[\Delta D^k_{ij}\right]$$

et $\overline{m}^{(k)}_{ij}$ les moments centrés :

$$\overline{m}^{(k)}_{ij} = \mathbb{E}\left[\left(\Delta D_{ij} - \mathbb{E}\left[\Delta D_{ij}\right]\right)^k\right].$$

On a trois moments linéaires en temps :

$$m^{(1)}_{ij} = \frac{\alpha}{\xi}\Delta t_{ij}, \quad \overline{m}^{(2)}_{ij} = \left(\frac{\alpha}{\xi^2} + \tau^2\right)\Delta t_{ij} \quad \text{et} \quad \overline{m}^{(3)}_{ij} = \frac{2\alpha}{\xi^3}\Delta t_{ij}.$$

On pose :

$$m = \begin{pmatrix} m^{(1)} \\ \overline{m}^{(2)} \\ \overline{m}^{(3)} \end{pmatrix} = \begin{pmatrix} m^{(1)}_{ij}/\Delta t_{ij} \\ \overline{m}^{(2)}_{ij}/\Delta t_{ij} \\ \overline{m}^{(3)}_{ij}/\Delta t_{ij} \end{pmatrix} = f(\theta).$$

Les moments empiriques $\widehat{m}_n$ sont :

$$\widehat{m}_n = \begin{pmatrix} \widehat{m}^{(1)}_n \\ \widehat{\overline{m}}^{(2)}_n \\ \widehat{\overline{m}}^{(3)}_n \end{pmatrix} = \left(\sum_{i=1}^n N_i\right)^{-1} \sum_{i=1}^n \sum_{j=1}^{N_i} \begin{pmatrix} \Delta D_{ij}/\Delta t_{ij} \\ \left(\Delta D_{ij} - \Delta t_{ij}\widehat{m}^{(1)}_n\right)^2/\Delta t_{ij} \\ \left(\Delta D_{ij} - \Delta t_{ij}\widehat{m}^{(1)}_n\right)^3/\Delta t_{ij} \end{pmatrix}.$$

On peut donc alors construire l'estimateur de θ par la méthode des moments :

$$\widehat{\theta}_n = f^{-1}(\widehat{m}_n).$$

Partant d'une loi des grands nombres pour des variables indépendantes [150], on montre la convergence de l'estimateur $\widehat{\theta}_n$:

Théorème 3.2.1 *Sous les hypothèses*

(H_1) $\sum_{n\geqslant q1} \sum_{j=1}^{N_n} (\Delta t_{nj})^{-1} \left(\sum_{i=1}^{n} N_i\right)^{-2} < \infty$

(H_2) $\exists\, d_u,\ \forall i \in \mathbb{N}^*,\ \forall j \in \{1,\ldots,N_i\},\ \Delta t_{ij} \leqslant q d_u$

$\widehat{\theta}_n$ converge p.s. vers θ quand n tend vers l'infini.

Le théorème central limite de Lindeberg-Feller permet de montrer la normalité asymptotique des moments empiriques. En appliquant la δ-méthode, on en déduit la normalité asymptotique de $\widehat{\theta}_n$:

Théorème 3.2.2 *Sous les hypothèses (H_1), (H_2) et*

(H_3) $\forall u \in \{0,1,3\}$, $\lim_{n\to\infty} \left(\sum_{i=1}^{n} N_i\right)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{N_i} \Delta t_{ij}^{u-2} = c_u < \infty$,

il existe une matrice symétrique définie positive Σ telle que :

$$\left(\sum_{i=1}^{n} N_i\right)^{1/2} \left(\widehat{\theta}_n - \theta\right) \xrightarrow[n\to\infty]{d} N(0,\Sigma).$$

Cette matrice Σ est décrite explicitement dans [6]. On peut donc utiliser ce résultat pour donner des intervalles de confiance asymptotiques. Nous avons également étudié l'interprétation des hypothèses requises dans les théorèmes précédents. En particulier, nous avons construit des situations où l'estimateur proposé est convergent et/ou asymptotiquement normal. Nous avons également étudié le comportement de cet estimateur sur la base de données simulées.

3.2.2 Inférence dans le modèle avec covariable [43]

On doit donc estimer un paramètre supplémentaire β. On suppose que l'on observe n processus gamma perturbés indépendants et de même loi, notés $D^{(1)},\ldots,D^{(n)}$. Ces processus sont observés aux mêmes instants sur l'intervalle de temps $[0,T]$: $0 = t_0 < t_1 < \cdots < t_N$ avec $t_j = jT/N$ (N est donc le nombre d'observations par processus). On observe aussi les n vecteurs de covariables $x_1,\ldots,x_n$ associées aux processus de dégradation.

On propose ici un estimateur de moindres carrés en deux étapes. On commence par estimer $\theta^{(1)} = (\gamma,\beta)$ avec $\gamma = \alpha/\xi$ en minimisant :

$$d_1^{(n)}\left(\theta^{(1)}\right) = \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\Delta D_{x_i}^{(i)}(t_j) - m_{\theta^{(1)}}^{(1)}(x_i)\right)^2 = \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\Delta D_{x_i}^{(i)}(t_j) - \frac{\gamma T}{N} e^{\beta^T x_i}\right)^2.$$

Puis on estime $\theta^{(2)} = (\alpha,\tau^2)$ en minimisant :

$$d_2^{(n)}\left(\theta^{(2)},\widehat{\theta}^{(1)}\right) = \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\left[\Delta D_{x_i}^{(i)}(t_j)\right]^2 - m_{\theta^{(2)}}^{(2)}\left(x_i,\widehat{\theta}^{(1)}\right)\right)^2$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{N}\left(\left[\Delta D_{x_i}^{(i)}(t_j)\right]^2 - m_{\widehat{\theta}^{(1)}}^{(1)}(x_i)\frac{\widehat{\gamma}}{\alpha} - m_{\widehat{\theta}^{(1)}}^{(1)}(x_i)^2 - \tau^2 T/N\right)^2.$$

On suppose que la covariable x est la réalisation d'un vecteur aléatoire de densité f_X par rapport à la mesure μ_p sur $\mathbb{R}^p$. On a donc une loi des grands nombres :

$$d_1^{(n)}\left(\theta^{(1)}\right) \xrightarrow[n\to\infty]{Pr} d_1\left(\theta^{(1)}\right) = \int \mathbb{E}_{\theta_0}\left[D_x\left(\frac{T}{N}\right) - m_{\theta^{(1)}}^{(1)}(x)\right]^2 f_X(x)\,\mathrm{d}\mu_p(x)$$

et :

$$d_2^{(n)}\left(\theta^{(2)},\theta^{(1)}\right) \xrightarrow[n\to\infty]{Pr} d_2\left(\theta^{(2)},\theta^{(1)}\right) = \int \mathbb{E}_{\theta_0}\left[D_x^2\left(\frac{T}{N}\right) - m_{\theta^{(2)}}^{(2)}\left(x,\theta^{(1)}\right)\right]^2 f_X(x)\,\mathrm{d}\mu_p(x).$$

Nous avons démontré un résultat général pour une procédure d'estimation de moindres carrés en deux étapes :

Lemme 3.2.1 *On suppose que :*

1. Θ *est un ensemble compact*
2. *Pour* $i \in \{1,2\}$, $\theta \longmapsto d_i(\theta)$ *est continue et vérifie :*
   — $d_i(\theta) \geqslant q0$,
   — $d_1(\theta)$ *a un unique minimum en* $\theta_0^{(1)} \in \mathring{\Theta}_1$,
   — $d_2(\theta)$ *a un unique minimum en* $\theta_0 \in \mathring{\Theta}$.
3. $\sup_{\theta\in\Theta} \| d^{(n)}(\theta) - d(\theta) \| \xrightarrow[n\to\infty]{Pr} 0$.

Alors $(\widehat{\theta}_n)_{n\geqslant q0}$ *converge en probabilité vers* θ_0 *quand* n *tend vers l'infini.*

On peut donc appliquer ce lemme à notre situation :

Proposition 3.2.1 *Sous les hypothèses suivantes :*

(A_1) Θ *est un ensemble compact tel que* $\theta_0 \in \mathring{\Theta}$ *et* $\beta_0 \neq 0$

(A_2) X *est un vecteur aléatoire uniformément borné*

(A_3) *Soit* $\mathcal{A} \subset \mathbb{R}^p$ *tel que* $\mu_p\left(\overline{\mathcal{A}}\right) = 0$. *Il existe* $x_1,\ldots,x_{p+1} \in \mathcal{A}$ *tel que :*
   (a) $\forall i \in \{1,\ldots,p+1\}$, $f_X(x_i) > 0$
   (b) pour tout $i \in \{1,\ldots,p+1\}$ *on pose* $\tilde{x}_i^T = \begin{pmatrix}1 & x_i^T\end{pmatrix}$. *Alors* $\tilde{x}_1,\ldots,\tilde{x}_{p+1}$ *sont linéairement indépendants,*

on a $\hat{\theta}_n \xrightarrow[n\to\infty]{Pr} \theta_0$.

Nous avons également montré la normalité asymptotique :

Théorème 3.2.3 *Sous les hypothèses* $(A_1)-(A_3)$ *et*

(A_4) *pour tout* $\theta \in \Theta$, $I_\infty(\theta)$ *est inversible*

(A_5) *il existe une constante* $\epsilon > 0$ *et des fonctions* E_{k_1,k_2} *tels que :*

$$\sup_{\beta\in B(\beta_0,\epsilon)} \left| \frac{1}{n}\sum_{i=1}^n x_i^{\otimes k_1} e^{k_2\beta^T x_i} - E_{k_1,k_2}(\beta) \right| \xrightarrow[n\longrightarrow\infty]{} 0,$$

où :

$$x_i^{\otimes k_1} = \begin{cases} 1 & if \quad k_1 = 0 \\ x_i & if \quad k_1 = 1 \\ x_i x_i^T & if \quad k_1 = 2 \\ x_i x_i^T x_i^{\otimes 2} & if \quad k_1 = 4 \end{cases}$$

il existe une matrice symétrique définie positive Σ_2 *telle que :*

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow[n\longrightarrow\infty]{d} N(0,\Sigma_2).$$

Cette matrice Σ_2 est décrite explicitement dans [43].

3.3 Subordinateurs perturbés [14]

On s'intéresse ici à une classe plus large de processus de dégradation que précédemment. En effet, on considère ici le modèle de dégradation suivant :

$$\forall t \geqslant 0\,,\ D_t = G_t + \sigma B_t \tag{3.3.1}$$

où (G_t) est un subordinateur, i.e. un processus de Lévy croissant (le processus gamma est donc un subordinateur). Puisque les sauts du processus (D_t) sont dus au subordinateur et sont positifs, on dit que (D_t) est spectralement positif. Ce processus est caractérisé par son exposant de Lévy ϕ_D défini par :

$$\forall u \in \mathbb{R}, \quad \mathbb{E}[e^{iuD_t}] = \exp(t\phi_D(u)) = \exp(t\phi_G(u))\exp(-\frac{1}{2}tu^2\sigma^2),$$

où ϕ_G est l'exposant de Lévy du processus (G_t) qui, en toute généralité, est caractérisé par un réel $\hat{\mu} \in \mathbb{R}$ et une mesure Q dont le support est inclus dans $\mathbb{R} \setminus \{0\}$ tels que :

$$\phi_G(u) = i\hat{\mu}u + \int_{\mathbb{R}\setminus\{0\}} \{e^{iux} - 1 - iux\mathbb{I}_{[-1,1]}(x)\}Q(\mathrm{d}x).$$

Comme (G_t) est un subordinateur, $\varphi_G(u)$ peut aussi s'écrire ainsi :

$$\varphi_G(u) = -\mu u + \int_0^\infty [e^{-ux} - 1]Q(\mathrm{d}x),$$

avec $\mu \geqslant 0$.

Le temps de panne d'un composant est traditionnellement obtenu à partir d'un modèle de dégradation en considérant le premier temps d'atteinte T_c d'un niveau critique $c > 0$. Le premier temps de passage dans le cas particulier de deux sous-modèles a déjà été étudié. Dans le cas d'un mouvement brownien avec tendance linéaire (on l'appellera aussi processus de Wiener), correspondant au cas où $G_t = \mu t$, $\mu > 0$, la loi de T_c est la loi gaussienne inverse, voir par exemple [75]. Le second cas est celui du processus gamma pur (i.e. $\sigma = 0$ et (G_t) est un processus gamma), cela a été étudié, comme on l'a vu précédemment, par exemple par Park et Padgett [145].

Un des objectifs de ce travail est de discuter de la notion de temps de panne pour cette famille de processus dont les trajectoires ne sont pas croissantes. En effet, récemment, Barker et Newby [58] ont proposé de considérer plutôt le dernier temps de passage du seuil critique comme instant de défaillance. Ce choix est justifié par le fait que, même si $\{D_t,\ t \geqslant 0\}$ passe au-dessus du seuil critique c impliquant que le composant est temporairement dégradé, il peut encore revenir en-dessous de c à moins que ce soit le dernier temps de passage du niveau c. Puisqu'après le dernier temps de passage, le processus ne reviendra plus en-dessous du seuil c, cet instant peut donc être légitimement considéré comme le temps de panne du système. La situation devient plus complexe car, à la différence du premier temps de passage, le dernier temps de passage n'est pas un temps d'arrêt (intuitivement, il faut connaître le futur pour pouvoir dire qu'un temps de passage est le dernier). Cette discussion n'a évidemment de sens que pour des processus dont les trajectoires ne sont pas croissantes.

Un des inconvénients de ces modèles est que, pour tout t, D_t peut être négatif avec une probabilité strictement positive. Ce problème peut être contourné en considérant la version réfléchie du processus (D_t) défini de la façon suivante :

$$\forall t \geqslant 0, \quad D_t^* := D_t - \inf_{0\leqslant s\leqslant t}(D_s \wedge 0). \tag{3.3.2}$$

Intuitivement, D_t^* est égal à D_t quand ce dernier est positif et est égal à zéro tant que D_t est négatif. Ce type de processus apparaît dans de nombreux domaines en probabilités appliquées, voir par exemple chapitre 4.1 of [120]. Le processus réfléchi constitue donc une approche intéressante pour modéliser une dégradation.

Dans la première partie, on considère la loi du temps de premier passage. Pour cela, via une approximation, on calcule la transformée de Laplace de T_c conjointement avec une fonction de pénalité dépendant de l'undershoot et de l'overshoot du processus :

$$\phi_w(\delta, b) = \mathbb{E}\left[e^{-\delta T_c} w(D_{T_c-}, D_{T_c})\right] \tag{3.3.3}$$

où $\delta \geqslant 0$ et w une fonction continue et bornée. On comparera cette approche avec celle récemment développée dans le cadre général des processus de Lévy, reposant sur les fonctions d'échelle. Dans la seconde partie, on considère donc le dernier temps de passage du seuil critique. Nous avons calculé la fonction de répartition du dernier temps de passage, ainsi que la transformée de Laplace du dernier temps de passage conjointement à l'overshoot et l'undershoot. La transformée de Laplace du dernier temps de passage pour le processus réfléchi est également obtenue. Dans [14], on a également proposé une application à un problème de maintenance inspiré de l'article de Barker et Newby [58] et pour lequel on peut calculer les quantités intervenant dans cette application.

3.3.1 Premier temps de passage comme instant de panne

On considère donc ici le premier temps de passage d'un seuil déterministe $b > 0$ par le processus perturbé $\{D_t,\ t \geqslant 0\}$:

$$T_c = \inf\{t \geqslant 0\ ;\ D_t \geqslant c\}$$

qui est presque sûrement fini (car $\lim_{t\to\infty} D_t = +\infty$ presque sûrement). On étudie la loi de (T_c, D_{T_c-}, D_{T_c}) en déterminant la quantité (3.3.3). Dans la suite, on notera $\phi(\delta, c)$ au lieu de $\phi_w(\delta, c)$ s'il n'y a pas d'ambiguïté. Dans [14], trois exemples sont traités afin d'illustrer nos résultats : processus de Wiener, processus gamma perturbé et processus de Poisson composé perturbé avec des sauts de loi de type phase.

Approche par les fonctions d'échelle

Il est également possible d'obtenir la loi jointe du temps de premier passage T_c et de D_{T_c} (overshoot) à l'aide des fonctions d'échelles. On rappelle la proposition suivante, voir par exemple Kyprianou et Palmowski [121], qui donne la transformée de Laplace de T_c :

Proposition 3.3.1 *Pour tout $\delta \geqslant 0$, on définit la fonction d'échelle $W^{(\delta)}$ via sa transformée de Laplace :*

$$\int_0^\infty e^{-\lambda x} W^{(\delta)}(x)dx = \frac{1}{\varphi_D(\lambda) - \delta}, \quad \lambda > \rho(\delta)$$

et la fonction $Z^{(\delta)}$ par :

$$Z^{(\delta)}(x) = 1 + \delta \int_0^x W^{(\delta)}(y)dy,$$

où $\rho(\delta)$ est la solution de l'équation de Lundberg $\varphi_D(\lambda) = \delta$. Alors,

$$\mathbb{E}[e^{-\delta T_c}] = Z^{(\delta)}(c) - \frac{\delta}{\rho(\delta)} W^{(\delta)}(c). \tag{3.3.4}$$

Précisons que, dans [121] (comme dans la plupart des articles), Kyprianou et Palmowski considèrent le cas de processus spectralement négatif alors qu'ici on a affaire à un processus spectralement positif. Il faut donc appliquer le résultat de Kyprianou et Palmowski au processus $\tilde{D}_t := -D_t$ avec $\tilde{D}_0 = c$. De manière similaire, le résultat obtenu par Biffis et Kyprianou [63] donne une expression de (3.3.3) à l'aide des fonctions d'échelle.

Théorème 3.3.1 *Soit $\overline{D}_{T_c-} := \sup_{t<T_c} D_t$ le dernier maximum avant le temps d'atteinte du niveau c. On a :*

$$\mathbb{E}\left[e^{-\delta T_c} w(D_{T_c-}, D_{T_c}, \overline{D}_{T_c-})\right] = \int_{(0,+\infty)^3} \mathbb{I}_{\{v\geq y\}} w(u + c, -v - c, -y - c) K_c^{(\delta)}(du, dv, dy),$$

où la fonction $w(\cdot,\cdot,\cdot)$ vérifie $w(\cdot, c, \cdot) = 0$ et

$$K_c^{(\delta)}(du, dv, dy) := e^{-\rho(\delta)(v-y)} \left[W^{(\delta)\prime}(c - y) - \rho(\delta) W^{(\delta)}(c - y)\right] Q(du + v)dydv.$$

En particulier,

$$\phi_w(\delta, c) = \int_{(0,+\infty)^2} w(u + c, -v - c) \tilde{K}_c^{(\delta)}(v) Q(du + v)dv \tag{3.3.5}$$

où $\tilde{K}_c^{(\delta)}(v) := \displaystyle\int_{y=0}^{v} e^{-\rho(\delta)(v-y)} \left[W^{(\delta)\prime}(c - y) - \rho(\delta) W^{(\delta)}(c - y)\right] dy.$

Les remarques suivantes seront utiles pour la suite.

Remarque 3.3.1 (Régularité des fonctions d'échelle) Une condition nécessaire pour la fonction $W^{(\delta)}$ définie dans la proposition 3.3.1 soit différentiable est que (D_t) ne soit pas à variation bornée, ce qui est le cas grâce à la composante gaussienne (i.e. $\sigma > 0$). En effet, Chan *et al.* [179] ont montré que $W^{(\delta)}$ est deux fois différentiable si $\sigma > 0$.

Remarque 3.3.2 (Valeur initiale de la fonction d'échelle) D'après le lemme 8.6 p.222 dans [120], comme (D_t) n'est pas à variation bornée, on a $W^{(\delta)}(0) = 0$.

L'approche donnée par la proposition 3.3.1 a tout de même un coût qui est l'inversion de la transformée de Laplace ci-dessus nécessaire pour calculer la fonction d'échelle. Cependant, récemment, des expressions explicites de $W^{(\delta)}$ ont été obtenues dans des cas particuliers, voir par exemple [106] et [87].

Approche par les équations de renouvellement

Une autre manière de calculer l'équation (3.3.3) consiste approcher la partie à sauts du processus (G_t) par un processus de Poisson composé. Cette technique a été déjà été proposée par Garrido et Morales dans le cadre de subordinateur non perturbé (pour plus de détails, voir l'appendice A.1 dans [95]). On peut alors appliquer des résultats existants en théorie de la ruine pour obtenir le résultat suivant. On rappelle que la convolution de deux fonctions f et g de $[0, +\infty)$ dans $\mathbb{R}$ est définie par $f \star g(z) = \int_0^z f(x)g(z-x)dx$.

Proposition 3.3.2 *Soit $\omega(x) := \int_x^\infty w(x, y-x)Q(\mathrm{d}y)$. La fonction $\phi(\delta, \cdot) = \phi_w(\delta, \cdot)$ satisfait l'équation de renouvellement :*

$$\phi(\delta, c) = \phi(\delta, \cdot) \star g(\delta, \cdot)(c) + h(\delta, c) \tag{3.3.6}$$

où les fonctions $g(\cdot, \cdot)$ et $h(\cdot, \cdot)$ sont définies par :

$$g(\delta, y) = \frac{2}{\sigma^2} \int_0^y e^{-[-2\mu/\sigma^2 + \rho(\delta)](y-s)} \int_s^\infty e^{-\rho(\delta)(x-s)} Q(dx)ds \tag{3.3.7}$$

$$h(\delta, y) = e^{-[-2\mu/\sigma^2 + \rho(\delta)]y} + \frac{2}{\sigma^2} \int_0^y e^{-[-2\mu/\sigma^2 + \rho(\delta)](y-s)} \int_s^\infty e^{-\rho(\delta)(x-s)} \omega(x)dxds. \tag{3.3.8}$$

Autrement dit, la fonction $\phi(\delta, c)$ est donnée par :

$$\phi(\delta, c) = \sum_{k=0}^{\infty} g^{\star k}(\delta, \cdot) \star h(\delta, \cdot)(\delta, c). \tag{3.3.9}$$

Ci-dessous, on explique comment il est possible d'approcher le processus (G_t). Comme affirmé précédemment, on peut l'approcher ponctuellement par une suite de processus de Poisson composés $((S(t, n))_{t \geqslant 0})_{n \in \mathbb{N}}$ telle que :

1. $(S(t, n))_{n \in \mathbb{N}}$ est croissant pour tout $t \geqslant 0$,
2. $\mu t + \lim_{n \to \infty} S(t, n) = G_t$ pour tout $t \geqslant 0$,
3. pour tout n, $(S(t, n))_{t \geqslant 0}$ est d'intensité $\lambda_n = \bar{Q}(1/n)$ et la fonction de répartition des sauts est $P_n(x)$ avec :
$$P_n(x) = \frac{\bar{Q}(1/n) - \bar{Q}(x)}{\bar{Q}(1/n)} \mathbb{I}_{\{x \geqslant 1/n\}}$$
où $\bar{Q}(x) := Q([x, +\infty))$.

Cette approximation est particulièrement intéressante lorsque $\lambda_n = \bar{Q}(1/n) \longrightarrow \bar{Q}(0) = Q([0, +\infty)) = +\infty$ quand $n \to \infty$, i.e. lorsque le processus a une infinité de (petits) sauts sur tout intervalle de temps (ce qui est le cas du processus gamma). Intuitivement, $\{S(t, n),\ t \geqslant 0\}$ est obtenu en supprimant tous les sauts de taille inférieur à $1/n$. Puisque $\{S(t, n),\ t \geqslant 0\}$ est croissant et tend vers $\{D_t,\ t \geqslant 0\}$, on a :

$$T_c^n \searrow T_c, \quad n \to \infty, \text{ p.s.}, \tag{3.3.10}$$

où T_c^n est le temps de premier passage du niveau c par le processus $\{D_t^n,\ t \geqslant 0\}$ défini par $D_t^n = \mu t + S(t,n) + \sigma B_t$ pour tout $t \geqslant 0$ et pour tout $n \in \mathbb{N}$. On rappelle que T_c^n est aussi presque sûrement fini et que T_c^n peut être interprété comme le temps de ruine dans un problème d'actuariat. On s'intéresse donc à la transformée de Laplace $\phi_n(\delta) := \mathbb{E}(e^{-\delta T_c^n} w(D^n_{T_c^n-}, D^n_{T_c^n}))$ de T_c^n avec la fonction de pénalité $w(\cdot,\cdot)$ pour tout $\delta \geqslant 0$. Soit $\rho_n = \rho_n(\delta)$ la solution positive de l'équation suivante :

$$\lambda_n \int_0^\infty e^{-\rho_n x} dP_n(x) = \lambda_n + \delta - \frac{\sigma^2}{2}\rho_n^2 + \mu\rho_n \tag{3.3.11}$$

appelée équation de Lundberg généralisée. On commence par montrer la convergence de ρ_n quand $n \to \infty$.

Proposition 3.3.3 *Quand $n \to \infty$, ρ_n converge vers l'unique solution $\rho > 0$ de l'équation de Lundberg généralisée :*

$$\delta - \frac{\sigma^2}{2}\rho^2 = \varphi_G(\rho) \quad \Longleftrightarrow \quad \delta = \varphi_D(\rho) \tag{3.3.12}$$

On peut alors appliquer à T_c^n un résultat de Tsai et Wilmott (voir le théorème 2 dans [182]) :

Théorème 3.3.2 *Soit w une fonction bornée et continue. On pose :*

$$\omega_n(x) = \int_x^\infty w(x, y-x) dP_n(y).$$

Alors, $b \mapsto \phi_n(\delta, c) := \mathbb{E}(e^{-\delta T_c^n} w(D^n_{T_c^n-}, D^n_{T_c^n}))$ satisfait l'équation de renouvellement suivante :

$$\phi_n(\delta, c) = \phi_n(\delta,\cdot) \star g_n(\delta,\cdot)(c) + h_n(\delta, c) \tag{3.3.13}$$

où les fonctions $g_n(\cdot,\cdot)$ et $h_n(\cdot,\cdot)$ sont définies par :

$$g_n(\delta, y) = \frac{2\lambda_n}{\sigma^2} \int_0^y e^{-[-2\mu/\sigma^2 + \rho_n(\delta)](y-s)} \int_s^\infty e^{-\rho_n(\delta)(x-s)} dP_n(x) ds \tag{3.3.14}$$

$$h_n(\delta, y) = e^{-[-2\mu/\sigma^2+\rho_n(\delta)]y} + \frac{2\lambda_n}{\sigma^2} \int_0^y e^{-[-2\mu/\sigma^2+\rho_n(\delta)](y-s)} \int_s^\infty e^{-\rho_n(\delta)(x-s)} \omega_n(x) dx ds. \tag{3.3.15}$$

En passant à la limite dans le théorème 3.3.2, on obtient la proposition 3.3.2.

On conclut cette partie par un résultat qui combine les propositions 3.3.2 et 3.3.1, donnant une expression explicite de la fonction d'échelle :

Proposition 3.3.4 *La fonction d'échelle $W^{(\delta)}$ est donnée par :*

$$W^{(\delta)}(x) = \int_0^x e^{-\rho(\delta)(x-y)} H(\delta, y) dy \tag{3.3.16}$$

avec :

$$H(\delta, x) := W^{(\delta)'}(x) - \rho(\delta) W^{(\delta)}(x) = -\frac{\rho(\delta)}{\delta} \sum_{k=0}^\infty \left[g^{\star k}(\delta,\cdot) + g^{\star k}(\delta,\cdot) \star h'(\delta,\cdot)(\delta, x) \right] \tag{3.3.17}$$

où $g(\delta,\cdot)$ est donnée par (3.3.7) et $h'(\delta,\cdot)$ est la dérivée de $h(\delta,\cdot)$ avec $w \equiv 1$, i.e.

$$\begin{aligned} h'(\delta, y) = &-[-2\mu/\sigma^2 + \rho(\delta)] e^{-[-2\mu/\sigma^2+\rho(\delta)]y} + \frac{2}{\sigma^2} \int_y^\infty e^{-[-2\mu/\sigma^2+\rho(\delta)](x-y)} \bar{Q}(x) dx \\ &- [-2\mu/\sigma^2 + \rho(\delta)] \frac{2}{\sigma^2} \int_0^y e^{-[-2\mu/\sigma^2+\rho(\delta)](y-s)} \int_s^\infty e^{-\rho(\delta)(x-s)} \bar{Q}(x) dx ds. \end{aligned} \tag{3.3.18}$$

On remarquera que l'équation (3.3.16) nécessite de calculer la série apparaissant dans l'équation (3.3.17), ce qui peut être difficile à calculer en pratique.

Remarque 3.3.3 Soit T_c^* le premier temps de passage du niveau c du processus réfléchi $\{D_t^*,\ t \geqslant 0\}$. Il est possible d'obtenir, à partir de la remarque 4 p.14 dans [84], une expression de la transformée de Laplace de T_c^* conjointement avec l'overshoot et de l'undershoot, i.e. $\mathbb{E}[e^{-\delta T_c^*} \mathbb{I}_{\{D^*_{T_c^*-} \in dy,\ D^*_{T_c^*} \in dz\}}]$. Cette expression repose sur la fonction d'échelle $W^{(\delta)}$ et la mesure de Lévy $\nu_D(\cdot)$ du processus non réfléchi (D_t).

3.3.2 Dernier temps de passage comme instant de panne

Soit L_c et L_c^* le dernier temps de passage respectivement par les processus (D_t) et $\{D_t^*,\ t \geqslant 0\}$ du niveau c :

$$L_c := \sup\{u \geqslant 0\ ;\ D_u \leqslant b\} \quad \text{et} \quad L_c^* := \sup\{u \geqslant 0\ ;\ D_u^* \leqslant b\}$$

qui sont bien définis car $\lim_{t\to+\infty} D_t = \lim_{t\to+\infty} D_t^* = +\infty$. On cherche ici à obtenir des résultats sur la loi de L_c et de L_c^*. En partant d'un résultat de Kyprianou *et al.* [122], on montre le théorème suivant :

Théorème 3.3.3 *Pour tout $t \geqslant 0$ et $a \in \mathbb{R}$,*

$$\mathbb{P}(L_c < t) = \int_c^\infty \mathbb{E}[D_1] W(a-b) f_{D_t}(a) da$$

et

$$\mathbb{P}(L_c \geqslant t,\ D_t \in da) = [1 - \mathbb{E}[D_1] W(a-b)] f_{D_t}(a) da$$

où $f_{D_t}(.)$ est la densité de D_t et $W(.) = W^{(\delta)}(.)$ avec $\delta = 0$. De plus, pour tout $\delta \geqslant 0$, et pour tout $b > y \geqslant 0$, $w > 0$, la transformée de Laplace de L_c conjointement avec l'undershoot et overshoot est donnée par :

$$\mathbb{E}[e^{-\delta L_c} \mathbb{I}_{\{b - D_{L_c^-} \in dy,\ D_{L_c} - b \in dw\}}] = \left[e^{\rho(\delta)(b-y)} \frac{1}{\varphi_D'(\rho(\delta))} - W^{(\delta)}(b-y) \right] dy . [1 - e^{-\rho(0)w}] Q(dw + y). \tag{3.3.19}$$

Pour le cas du processus réfléchi, on a montré le résultat suivant :

Théorème 3.3.4 *La transformée de Laplace de L_c^* est donnée par :*

$$\mathbb{E}\left[e^{-\delta L_c^*}\right] = \mathbb{E}[D_1] \int_c^\infty W'(a-b) \phi(\delta, a) da$$

où $\phi(\delta, a) = \mathbb{E}[e^{-\delta T_a}] = \phi_w(\delta, a)$ avec $w \equiv 1$.

3.4 Processus gamma modulé par un processus markovien [38]

Dans plusieurs articles, des covariables et/ou des effets aléatoires ont été introduits dans un processus gamma afin de prendre en compte l'environnement et/ou une hétérogénéité entre individus. Par exemple, un effet aléatoire a été considéré par Lawless et Crowder [125] qui supposent que le paramètre d'échelle est aléatoire. Quant à Bagdonavičius et Nikulin [55], ils ont proposé un modèle de dégradation accélérée en dilatant l'échelle de temps par des covariables dépendant du temps. Cependant, dans tous ces travaux, les covariables (dépendant ou pas du temps) sont supposées être déterministes. Dans cette section, on étudie un processus gamma qui intègre des covariables qui évoluent selon un processus markovien de sauts (supposé indépendant du processus gamma sous-jacent). Pour des raisons de simplicité, on se restreint au cas d'un processus markovien binaire, mais tout ce qui suit peut s'étendre facilement au cas multivarié. Ce processus markovien représente l'environnement dans lequel l'unité est utilisée. Par exemple, on suppose que le composant est utilisé alternativement sous une condition de stress nominal (état 0) et sous une condition de stress accéléré (état 1). Pour des modèles similaires, on pourra regarder [167, 195, 194, 175].

Dans la première sous-section, on définit le modèle pour lequel on donne des propriétés de base. Un tel modèle est appelé processus gamma modulé par un processus markovien ou encore processus gamma avec changement de régime markovien. Ensuite, on étudie la loi du temps d'atteinte d'un niveau critique par ce processus, ainsi que des propriétés de comparaison stochastique. Enfin, dans la dernière sous-section, on considère deux politiques de remplacement associées à ce processus.

3.4.1 Modèle de dégradation avec un environnement aléatoire

Dynamique de la covariable

La covariable (ou conditions de sollicitation) est modélisée par un processus markovien binaire $(J(t))$. On note λ (resp. μ) le taux de transition de l'état 0 vers l'état 1 (resp. de l'état 1 vers l'état 0). Alors, le générateur infinitésimal Q est donné par :

$$Q = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix}.$$

On rappelle qu'un processus markovien (sur un espace d'état discret) est caractérisé par son générateur infinitésimal et par sa loi initiale ν. Pour tout $t \geqslant 0$, la matrice de transition entre les instants 0 et t est égale à $P_t = \exp(tQ)$ et donc la loi de $J(t)$ est $\nu \exp(tQ)$. Dans le cas d'un processus binaire, pour tout $t \geqslant 0$,

$$P_t = \frac{1}{\lambda+\mu}\begin{pmatrix} \mu & \lambda \\ \mu & \lambda \end{pmatrix} + \frac{e^{-(\lambda+\mu)t}}{\lambda+\mu}\begin{pmatrix} \lambda & -\lambda \\ -\mu & \mu \end{pmatrix}.$$

Il en résulte que son unique loi stationnaire π est :

$$\pi = \begin{pmatrix} \frac{\mu}{\lambda+\mu} \\ \frac{\lambda}{\lambda+\mu} \end{pmatrix}.$$

Dans un cas plus général, on supposera que $(J(t))$ est un processus markovien homogène, irréductible et récurrent. Souvent, la matrice de transition ne peut pas être calculée explicitement.

Dynamique de la dégradation

Le processus de dégradation $(D(t))$ proposé ici est un processus gamma non-homogène de fonction de forme $(\eta(t))$ et de paramètre d'échelle ξ. La fonction de forme est linéaire par morceaux et aléatoire car dépendante du processus des covariables $(J(t))$ de la manière suivante :

$$\forall n \in \mathbb{N}, \quad \eta(t) = \begin{cases} \eta(\tau_{2n}) + \alpha_0(t-\tau_{2n}) & \text{if } \tau_{2n} < t \leqslant \tau_{2n+1} \\ \eta(\tau_{2n+1}) + \alpha_1(t-\tau_{2n+1}) & \text{if } \tau_{2n+1} < t \leqslant \tau_{2n+2} \end{cases}$$

où $(\tau_n)_{n\geqslant 0}$ sont les instants de sauts de $(J(t))$, avec par convention $\tau_0 = 0$ et $\eta(0) = 0$. La suite de variables aléatoires $(\tau_n)_{n\geqslant 0}$ est un processus de renouvellement alterné.

Ce processus peut aussi être décrit en considérant ses accroissements ainsi que ceux dans le cas non-modulé. Soit $(D^{(0)}(t))$ (resp. $(D^{(1)}(t))$) le processus gamma homogène de paramètre de forme α_0 (resp. α_1) et de paramètre d'échelle ξ. Alors, conditionnellement à $(\tau_n)_{n\geqslant 0}$, on a :

$$\forall t \in (\tau_{2n}, \tau_{2n+1}], \quad D(t) - D(\tau_{2n}) \overset{(d)}{=} D^{(0)}(t-\tau_{2n})$$

et

$$\forall t \in (\tau_{2n+1}, \tau_{2n+2}], \quad D(t) - D(\tau_{2n+1}) \overset{(d)}{=} D^{(0)}(\tau_{2n+1}-t)$$

où $\overset{(d)}{=}$ signifie "égalité en loi".

3.4.2 Loi du temps de panne

Comme de manière classique, le temps de panne associé à ce processus est défini comme le premier instant pour que le processus gamma modulé par un processus markovien atteigne un niveau fixe $c > 0$:

$$T_c = \inf\{t \geqslant 0 \,;\, D(t) \geqslant c\}.$$

Comme $(D(t))$ est aussi un processus à trajectoires croissantes, il en découle qu'on a :

$$\forall t \geqslant 0\,, \quad \mathbb{P}[T_c > t] = \mathbb{P}[D(t) < c].$$

Donc, il est suffisant d'étudier la loi de $D(t)$ pour tout $t \geqslant 0$. Dans la suite, on commence par rappeler des résultats connus pour le processus gamma non-modulé. Ensuite, on s'intéressera au cas modulé.

Cas d'un processus gamma non-modulé

Park e Padgett [145] ont obtenu l'expression suivante pour la fonction de répartition de T_c. Pour tout $t \geqslant 0$,

$$F_{T_c}(t) = \mathbb{P}[D(t) \geqslant c] = \frac{\Gamma(\alpha t, c/\xi)}{\Gamma(\alpha t)}, \tag{3.4.1}$$

où $\Gamma(\cdot,\cdot)$ est la fonction gamma incomplète supérieure. Ils ont également l'expression suivante pour la densité. Pour tout $t \geqslant 0$,

$$f_{T_c}(t) = \alpha\left(\Psi(\alpha t) - \log\left(\frac{c}{\xi}\right)\right)\frac{\gamma(\alpha t, c/\xi)}{\Gamma(\alpha t)} + \frac{1}{\alpha t^2\Gamma(\alpha t)}\left(\frac{c}{\xi}\right)^{\alpha t}{}_2F_2(\alpha t, \alpha t; \alpha t+1, \alpha t+1; -c/\xi),$$

où Ψ est la fonction di-gamma (ou dérivée logarithmique de la fonction gamma), $\gamma(\cdot,\cdot)$ est la fonction gamma incomplète inférieure et ${}_2F_2$ la fonction hypergéométrique généralisée d'ordre $(2,2)$. Voir [48] ou [98] pour plus de détails sur les fonctions spéciales.

Cas d'un processus gamma modulé

Considérons maintenant le cas d'un processus gamma modulé par un processus markovien. Soit $\Delta_{[0,t]}$ le temps d'occupation de l'état 0 sur l'intervalle de temps $[0,t]$ par le processus markovien binaire (partant initialement de l'état 0) :

$$\Delta_{[0,t]} = \int_0^t \mathbb{I}_{J(u)=0}\mathrm{d}u.$$

Ce type de variable aléatoire a été étudié dans la littérature, voir [119] et les références dedans. Dans ce papier, Kovchegov *et al.* [119] ont obtenu une expression explicite pour la transformée de Laplace de la densité de $\Delta_{[0,t]}$ pour tout processus markovien de sauts. Dans le cas particulier d'un processus binaire, la densité généralisée a été obtenue pour la première fois par Pedler [148] : pour tout $u \in [0,t]$,

$$f_{\Delta_{[0,t]}}(u)\mathrm{d}u = e^{-\lambda t}\delta_t(u) + e^{-\lambda u}e^{-\mu(t-u)}\left[\lambda I_0\left(2\sqrt{\lambda\mu u(t-u)}\right) + \sqrt{\frac{\lambda\mu u}{t-u}}I_1\left(2\sqrt{\lambda\mu u(t-u)}\right)\right]\mathrm{d}u,$$

où δ_t est la masse de Dirac au point t et I_r est la fonction de Bessel modifiée de la forme :

$$I_r(z) = \sum_{k=0}^{\infty}\frac{1}{k!\Gamma(k+r+1)}\left(\frac{z}{2}\right)^{2k+r}, \quad \text{for } r > -1.$$

La loi de $\Delta_{[0,t]}$ peut être vue comme un mélange entre une loi de Dirac (correspondant à l'événement qu'aucun saut ne s'est produit entre 0 et t) et une loi absolument continue.

La propriété d'indépendance des accroissements du processus gamma ainsi que l'hypothèse d'indépendance entre le processus gamma de base et le processus markovien de saut impliquent l'égalité en loi suivante :

$$D(t) \overset{(d)}{=} D^{(0)}(\Delta_{[0,t]}) + D^{(1)}(t - \Delta_{[0,t]}).$$

Ainsi, en conditionnant sur $\Delta_{[0,t]}$, on obtient l'expression suivante :

$$\mathbb{P}[T_c \geqslant t] = \int_0^t\left[1 - \frac{\Gamma(\alpha_0 u + \alpha_1(t-u), c/\xi)}{\Gamma(\alpha_0 u + \alpha_1(t-u))}\right]f_{\Delta_{[0,t]}}(u)\mathrm{d}u, \tag{3.4.2}$$

en utilisant le fait que la convolution de lois gamma avec le même paramètre d'échelle est de loi gamma. Il semble difficile d'obtenir une expression plus explicite. Cependant, des calculs numériques peuvent être menés facilement car les fonctions de Bessel sont déjà implémentées dans de nombreux langages ou logiciels scientifiques.

Comparaison stochastique

Soit $T_c^{(0)}$ (resp. $T_c^{(1)}$) le temps jusqu'à la panne pour le processus gamma $D^{(0)}$ (resp. $D^{(1)}$). Le résultat suivant, assez naturel et intuitif, établit un ordre stochastique entre les trois de temps de panne.

Théorème 3.4.1 *On a* $T_c^{(1)} \preceq_{st} T_c \preceq_{st} T_c^{(0)}$.

Comme il est bien connu, ce résultat implique un ordre entre les temps moyens jusqu'à la panne (MTTF) :

$$\mathbb{E}[T_c^{(1)}] \leqslant \mathbb{E}[T_c] \leqslant \mathbb{E}[T_c^{(0)}].$$

En utilisant une approximation du MTTF (dans le cas non-modulé) obtenu par Bérenguer *et al.* [61], on obtient des approximations pour les bornes inférieures et supérieures du MTTF dans le cas modulé (à condition que c/ξ soit assez grand) :

$$\frac{1}{\alpha_1}\left(\frac{c}{\xi}+\frac{1}{2}\right) \lessapprox \mathbb{E}[T_c] \lessapprox \frac{1}{\alpha_0}\left(\frac{c}{\xi}+\frac{1}{2}\right).$$

En fait, il est possible d'obtenir des relations explicites entre les fonctions de survie $S_{T_c^{(0)}}$ et $S_{T_c^{(1)}}$ respectivement de $T_c^{(0)}$ et $T_c^{(1)}$. En effet, en partant de leurs expressions (voir l'équation 3.4.1), on a :

$$\forall t \geqslant 0, \quad S_{T_c^{(1)}}(t) = S_{T_c^{(0)}}(\gamma t)$$

avec $\gamma = \alpha_1/\alpha_0 > 1$. Donc, on a $\mathbb{E}[T_c^{(1)}] = \gamma^{-1}\mathbb{E}[T_c^{(0)}]$. Il en découle que leurs densités $f_{T_c^{(0)}}$ e $f_{T_c^{(1)}}$ vérifient l'identité suivante :

$$\forall t \geqslant 0, \quad f_{T_c^{(1)}}(t) = \gamma f_{T_c^{(0)}}(\gamma t)$$

et, enfin, leurs fonctions de risque $h_{T_c^{(0)}}$ et $h_{T_c^{(1)}}$:

$$\forall t \geqslant 0, \quad h_{T_c^{(1)}}(t) = \gamma h_{T_c^{(0)}}(\gamma t).$$

Abdel-Hameed [47] a montré que $T_c^{(0)}$ (et donc aussi $T_c^{(1)}$) a un taux de risque croissant. Alors, pour tout $t > 0$, $h_{T_c^{(1)}}(t) > h_{T_c^{(0)}}(t)$. Ces relations seront utiles plus loin.

3.4.3 Étude de deux politiques de remplacement

Dans cette sous-section, on considère deux politiques de remplacement, la politique de remplacement par blocs (*block replacement policy*) et la politique de remplacement selon l'âge (age replacement policy). Pour plus de détails sur ces politiques de remplacement (ainsi que sur d'autres), on pourra consulter respectivement les chapitre 5 et 3 dans [136]. Ces politiques de remplacement dépendent d'un unique paramètre δ correspondant au délai entre deux inspections successives induisant un certain coût. On peut alors s'intéresser au délai inter-inspection optimal $\delta_\star$. Pour cela, on considère le coût asymptotique par unité de temps défini ainsi :

$$C(\delta) = \lim_{t\to\infty} \frac{C_t(\delta)}{t},$$

où $C_t(\delta)$ est le coût sur l'intervalle de temps $[0, t]$ quand l'unité est inspectée aux instants $(k\delta)_{k\in\mathbb{N}}$. Comme il est bien connu, en appliquant la théorie du renouvellement, on a :

$$C(\delta) = \frac{\text{coût moyen sur un cycle}}{\text{durée moyenne d'un cycle}}.$$

Politique de remplacement par blocs

On suppose que le niveau de dégradation d'une unité ne peut être mesurée que durant des inspections (i.e. il n'y a pas de mesure en continue de la dégradation) et que, à chaque remplacement, l'unité est remplacée par une unité neuve ou bien est parfaitement réparée (As Good As New - AGAN), la durée de remplacement/réparation étant supposée être négligeable. De plus, les remplacements ne peuvent avoir lieu

qu'après une inspection (en particulier, il n'y a pas de remplacement au moment de la panne et l'unité est indisponible entre la panne et la prochaine inspection.

Il existe donc deux coûts différents : le coût c_r de remplacement d'une unité et le coût c_u d'indisponibilité. On a :

$$C_{brp}(\delta) = \frac{\mathbb{E}[c_r + c_u(\delta - T_c)_+]}{\delta} = \frac{c_r + c_u \int_0^\delta F_{T_c}(u)\mathrm{d}u)}{\delta},$$

où F_{T_c} est la fonction de répartition du temps jusqu'à panne T_c. On a :

$$\lim_{\delta\to 0} C_{brp}(\delta) = +\infty \quad \text{e} \quad \lim_{\delta\to+\infty} C_{brp}(\delta) = c_u.$$

On note $\delta_\star := \mathrm{argmin}_{\delta>0} C_{brp}(\delta)$. En dérivant l'expression ci-dessus du coût, $\delta_\star$ est la solution de l'équation suivante par rapport à t :

$$\phi_{brp}(t) = \int_0^t u f_{T_c}(u)\mathrm{d}u = \mathbb{E}[T_c \mathbf{1}_{T_c\leqslant t}] = \frac{c_r}{c_u}(\leqslant 1).$$

La solution ne dépend donc des deux coûts c_r et c_u à travers leur rapport (et donc ne dépend pas de l'unité monétaire considérée, ce qui est très intuitif). Puisque ϕ_{brp} est croissante avec $\mathbb{E}[T_c]$ comme limite, il en résulte que $\delta_\star$ est finie si $\mathbb{E}[T_c] > c_r/c_u$ et infinie autrement.

La comparaison stochastique entre les trois temps jusqu'à défaillance induit une relation similaire entre les trois fonctions de coûts :

$$\forall \delta > 0, \quad C^{(0)}_{brp}(\delta) \leqslant C_{brp}(\delta) \leqslant C^{(1)}_{brp}(\delta),$$

où $C^{(0)}_{brp}$ et $C^{(1)}_{brp}$ sont les fonctions de coûts associées aux processus $D^{(0)}$ et $D^{(1)}$. Cependant, cet ordre n'est plus respecté pour les fonctions $\phi^{(0)}_{brp}$, ϕ_{brp} et $\phi^{(1)}_{brp}$. En fait, on peut au moins montrer que $\phi^{(0)}_{brp}$ et $\phi^{(1)}_{brp}$ se croisent. Cela vient du résultat suivant dans un cadre plus général avec des lois uni-modales (en utilisant des notations similaires à celles précédentes).

Théorème 3.4.2 *Soit X et Y deux variables aléatoires positives telles que $X \preceq_{st} Y$ et telles que les lois absolument continues de X et de Y sont uni modales de mode respectif $m_X \neq 0$ et m_Y avec $m_X < m_Y$. Alors, il existe un unique point $\hat\delta$ tel que $\phi_X(t) > \phi_Y(t)$ pour tout $t < \hat\delta$, $\phi_X(\hat\delta) = \phi_Y(\hat\delta)$ et $\phi_X(t) < \phi_Y(t)$ pour tout $t > \hat\delta$.*

On conjecture l'existence et l'unicité du mode pour $T^{(0)}_c$ (et donc aussi pour $T^{(1)}_c$), ce qui est supporté par plusieurs calculs numériques. Soit m_0 et m_1 les modes des lois de $T^{(0)}_c$ et de $T^{(1)}_c$. En partant de la relation entre les densités de $T^{(0)}_c$ et de $T^{(1)}_c$, on a :

$$\forall t \geqslant 0, \quad f'_{T^{(1)}_c}(t) = \gamma^2 f'_{T^{(0)}_c}(\gamma t),$$

et donc :

$$0 = f'_{T^{(1)}_c}(m_1) = \gamma^2 f'_{T^{(0)}_c}(\gamma m_1) = f'_{T^{(0)}_c}(\gamma m_0).$$

Ainsi, $m_0 = \gamma m_1 > m_1$. Alors, en utilisant le résultat ci-dessus, on obtient l'existence et l'unicité du point d'intersection $\hat\delta$ des fonctions $\phi^{(0)}_{brp}$ et $\phi^{(1)}_{brp}$. En conséquence, on conclut que :

$$\begin{cases} \hat\delta \leqslant \delta^{(0)}_\star \leqslant \delta^{(1)}_\star & \text{if } \phi^{(0)}_{brp}(\hat\delta) = \phi^{(1)}_{brp}(\hat\delta) \leqslant c_r/c_u \\ \delta^{(1)}_\star \leqslant \delta^{(0)}_\star \leqslant \hat\delta & \text{otherwise} \end{cases}$$

Politique de remplacement selon l'âge

On considère ici une autre politique de remplacement. Une unité est remplacée soit dès qu'elle tombe en panne, soit quand elle atteint un âge δ (à condition que la panne ne soit pas survenue entre les instants 0 et δ, $\delta = \infty$ signifiant que le remplacement n'a lieu que suite à une panne). Il y a donc, ici également, deux coûts : c_f correspondant à un remplacement d'une unité en panne et $c_{nf} < c_f$ correspondant à une unité encore en état de marche. On a :

$$C_{arp}(\delta) = \frac{c_f F_{T_c}(\delta) + c_{nf} S_{T_c}(\delta)}{\int_0^\delta S_{T_c}(u)\mathrm{d}u}.$$

Lorsque δ tend vers l'infini, puisque cela correspond au cas de remplacements uniquement suite à une panne, il est facile de voir que $C_{arp}(\delta)$ tend vers $\frac{c_f}{\mathbb{E}[T_c]} := C_{arp}(\infty)$.

D'après l'ordre stochastique entre les trois temps jusqu'à panne et puisqu'on a supposé que $c_{nf} < c_f$, il en découle qu'une relation similaire existe entre les trois fonctions de coûts :

$$\forall \delta > 0, \quad C_{arp}^{(0)}(\delta) \leqslant C_{arp}(\delta) \leqslant C_{arp}^{(1)}(\delta),$$

où $C_{arp}^{(0)}$ et $C_{arp}^{(1)}$ correspondent aux fonctions de coût associées aux processus $D^{(0)}$ et $D^{(1)}$. De plus, la comparaison entre les MTTF implique que l'ordre est aussi respecté à la limite :

$$C_{arp}^{(0)}(\infty) \leqslant C_{arp}(\infty) \leqslant C_{arp}^{(1)}(\infty).$$

Soit $\delta_\star := \text{argmin}_{\delta>0} C_{arp}(\delta)$. L'ordre entre les fonctions de coûts ne peut pas être utilisé directement pour essayer d'établir un ordre entre les délais optimaux $\delta_\star$, $\delta_\star^{(0)}$ et $\delta_\star^{(1)}$. Si T_c a une fonction de risque strictement croissante et continue et si $h_{T_c}(\infty) > c_f/(\mathbb{E}[T_c](c_f - c_{nf}))$, alors $\delta_\star$ est l'unique solution finie de l'équation (voir le théorème 3.2 dans [136]) :

$$\phi_{T_c}(t) := h_{T_c}(t) \int_0^t S_{T_c}(u)\mathrm{d}u - F_{T_c}(t) = \frac{c_{nf}}{c_f - c_{nf}}.$$

Nakagawa ([136], page 74) a démontré que ϕ_{T_c} est une fonction strictement croissante (pourvu que h_{T_c} soit continue et strictement croissante). Ce résultat implique que $\delta_\star^{(1)} < \delta_\star^{(0)}$ à condition que $h_{T_c^{(0)}}$ satisfasse les conditions ci-dessus (Abdel-Hameed [47] a montré la croissance de cette fonction). En effet, du lien entre $h_{T_c^{(0)}}$ e $h_{T_c^{(1)}}$, on en déduit que :

$$\forall t \geqslant 0, \quad \phi_{T_c^{(1)}}(t) = \phi_{T_c^{(0)}}(\gamma t) > \phi_{T_c^{(0)}}(t),$$

du fait de la croissance de $\phi_{T_c^{(0)}}$ et puisque $\gamma > 1$. En conséquence, $\delta_\star^{(1)} < \delta_\star^{(0)}$.

3.5 Processus de Wiener avec période d'initiation aléatoire

La plupart des modèles classiques supposent que le composant se dégrade dès qu'il est mis en service. Cependant, dans de nombreux cas, la dégradation ne débute qu'après une certaine période d'initiation (ou de latence). Le problème devient plus délicat car cette période est en général inconnue (aléatoire) et rarement observée exactement (durée censurée). Il existe seulement quelques articles traitant de ce type de modèle. Par exemple, Liao *et al.* [101] ainsi que Nelson [139] ont étudié un modèle de type trajectoire de dégradation (*degradation path*). Peng et Feng [149] ont considéré le modèle de croissance logistique de Verhulst avec un taux de croissance aléatoire démarrant à un instant aléatoire et à un niveau aléatoire, comme modèle de dégradation. Enfin, Zhang et Liao [193] ont considéré un processus gamma avec effet aléatoire et période d'initiation aléatoire.

Dans [45], on considère un autre modèle de dégradation ainsi qu'un autre modèle d'échantillonnage par rapport aux articles cités plus haut. Le modèle de dégradation $(X(t))_{t\geq 0}$ est la suivant :

$$\begin{cases} X(t) = 0 & \forall t \in [0, S) \\ X(t) = \mu(t - S) + \sigma B(t - S) & \forall t \in [S, \infty), \end{cases}$$

où S est une variable aléatoire positive absolument continue (de fonction de répartition F_S), indépendante du mouvement brownien $(B(t))_{t\geq 0}$: 0 correspond à l'instant où le composant est mis en service et, après un délai S, le composant commence à se dégrader. Autrement dit, $(X(t))_{t\geq 0}$ est un processus de Wiener de tendance μ et de volatilité σ^2, démarrant à l'instant aléatoire S.

Dans la première sous-section, on commence par introduire plusieurs notations utiles par la suite et on énonce un lemme important. Dans la sous-section suivante, on propose une méthodologie pour estimer les différents paramètres du modèle dans le cas d'une hypothèse paramétrique pour la loi de S. On supposera que n copies indépendantes $X_1, \ldots, X_n$ du processus stochastique X sont observées aux mêmes instants régulièrement espacés $0, \delta, 2\delta, \ldots, m\delta = \tau$. On s'intéressera enfin à l'estimation du temps moyen jusqu'à panne (MTTF).

3.5.1 Notations

L'instant où commence à se dégrader un composant n'est pas observé puisque le niveau de dégradation est connu seulement aux instants d'inspection $0, \delta, 2\delta, \ldots, m\delta = \tau$. On introduit alors le temps d'inspection aléatoire où la mesure de dégradation est non nulle pour la première fois. Pour le i-ème composant, on R_i la variable aléatoire à valeurs dans $\mathbb{N}^*$ tel que $(R_i - 1)\delta < S_i \leq R_i\delta$. La loi de R_i s'exprime facilement en fonction de celle de S_i :

$$\forall r \in \mathbb{N}^*, \quad \mathbb{P}[R_i = r] = \mathbb{P}[(r-1)\delta < S_i \leq r\delta] = \overline{F}_S((r-1)\delta) - \overline{F}_S(r\delta),$$

où $\overline{F}_S$ est la fonction de survie de S. Nous allons distinguer trois cas importants.

1. Si $R_i > m$, cela signifie que $S_i > m\delta = \tau$ et donc $X_i(j\delta) = 0$ pour tout $j \in \{0, \ldots, m\}$. Ces individus apportent de l'information uniquement sur la loi de S (censure à droite). On note $\mathcal{N}_0$ l'ensemble des individus dans cette situation et N_0 son cardinal.
2. Si $R_i = m$, cela signifie que $(m-1)\delta < S \leq m\delta$ et $X_i(j\delta) = 0$ pour tout $j \in \{0, \ldots, m-1\}$ mais $X_i(m\delta) \neq 0$. Une seule mesure de dégradation non nulle est observée, mais elle ne peut pas être (facilement) utilisée pour estimer μ et σ^2. Aussi, on ne considérera que ces individus apportent également de l'information uniquement sur la loi de S (censure par intervalle). On note $\mathcal{N}_1$ l'ensemble des individus dans cette situation et N_1 son cardinal.
3. Si $R_i < m$, alors au moins deux mesures non nulles de dégradation sont observées et ces individus apportent de l'information sur tous les paramètres du modèle. On note $\mathcal{N}_{2+}$ l'ensemble des individus dans cette situation et N_{2+} son cardinal.

Le vecteur aléatoire (N_0, N_1, N_{2+}) est de loi multinomiale de paramètres $(n; \overline{F}_S(\tau), \overline{F}_S(\tau-\delta) - \overline{F}_S(\tau), F_S(\tau-\delta))$. En particulier, on a :

$$\mathbb{E}[N_0] = n\overline{F}_S(\tau), \quad \mathbb{E}[N_1] = n[\overline{F}_S(\tau-\delta) - \overline{F}_S(\tau)] \quad \text{et} \quad \mathbb{E}[N_{2+}] = nF_S(\tau-\delta).$$

On considère ensuite le nombre aléatoire Q_n d'accroissements strictement positifs pour les individus de $\mathcal{N}_{2+}$:

$$Q_n = \sum_{i \in \mathcal{N}_{2+}} (m - R_i)$$

si $\mathcal{N}_{2+}$ est un sous-ensemble non vide et $Q_n = 0$ autrement. Si $\mathcal{N}_{2+}$ est vide, cela signifie qu'aucune dégradation n'est observée sur l'ensemble des n individus et, dans ce cas, aucune estimation ne peut être fournie. A partir de maintenant, on supposera que $\mathcal{N}_{2+}$ est non vide. Conditionnellement à $\mathcal{N}_{2+} \neq \emptyset$ (ce qui équivaut à $N_{2+} \neq 0$), Q_n est une variable aléatoire discrète à valeurs dans $\{1, \ldots, (m-1)n\}$.

Enfin, on introduit le vecteur aléatoire $\underline{K} = (K_1, \ldots, K_m)$ tel que, pour tout $r \in \{1, \ldots, m\}$,

$$K_r = \sum_{i=1}^{n} \mathbb{I}_{(r-1)\delta < S_i \leq r\delta} = \sum_{i=1}^{n} \mathbb{I}_{R_i = r},$$

qui est de loi binomiale de paramètres n et $\overline{F}_S((r-1)\delta) - \overline{F}_S(r\delta)$. Le nombre aléatoire Q_n d'accroissements strictement positifs peut aussi être exprimé à l'aide de $\underline{K}$:

$$Q_n = \sum_{j=1}^{m-1} (m-j) K_j. \tag{3.5.1}$$

Le lemme suivant donne d'importants résultats sur Q_n.

Lemme 3.5.1

1. *Pour tout $\alpha \in [0, 1)$,*

$$\frac{Q_n}{n^\alpha} \xrightarrow[n\to\infty]{Pr} \infty.$$

2. *Soit* $\alpha(m,\tau) = \frac{1}{m}\sum_{j=0}^{m-1} F_S(j\tau/m)$. *On a*

$$\frac{Q_n}{n} \xrightarrow[n\to\infty]{Pr} m\alpha(m,\tau).$$

3. *Conditionnellement à* $Q_n > 0$, *l'espérance de* Q_n^{-1} *tend vers zéro quand n tend vers l'infini :*

$$\mathbb{E}[Q_n^{-1}|Q_n > 0] \xrightarrow[n\to\infty]{} 0.$$

3.5.2 Estimation des paramètres

Les observations seront utilisées pour estimer soit les paramètres de la loi de S, soit les paramètres du processus de Wiener.

Estimation des paramètres de la loi de S

Comme expliqué précédemment, tous les individus apportent de l'information sur la loi de S, mais la nature de l'information peut être différente : les individus du sous-ensemble $\mathcal{N}_0$ correspondent à des durées censurées à droite, alors que les individus du sous-ensemble $\mathcal{N}_1 \cup \mathcal{N}_{2+}$ correspondent à des durées censurées par intervalle.

On suppose que la loi de S appartient à une famille paramétrique de loi de probabilités. On note par $\theta \in \Theta$ le paramètre inconnu, où Θ est un espace euclidien, et par $F_S(\cdot;\theta)$ la fonction de répartition de S (au lieu de $F_S(\cdot)$ comme précédemment). La fonction de log-vraisemblance est donnée par :

$$\ell(\theta|\text{data}) = N_0 \log \overline{F}_S(\tau;\theta) + \sum_{r=1}^{m} K_r \log\left(\overline{F}_S((r-1)\delta;\theta) - \overline{F}_S(r\delta;\theta)\right).$$

L'estimateur du maximum de vraisemblance de θ peut être calculé numériquement en maximisant la fonction ci-dessus (il n'y a pas d'expression explicite dans le cas général) :

$$\widehat{\theta}_n = \text{argmax}_{\theta\in\Theta}\ell(\theta|\text{data}).$$

Sous des hypothèses de régularité sur $\overline{F}_S$ (par rapport à θ), on peut montrer que l'estimateur du maximum de vraisemblance $\widehat{\theta}_n$ de θ est asymptotiquement normal. Pour des raisons de simplicité, on considère le cas d'un paramètre uni-dimensionnel pour la loi de S et on suppose que $\theta \in \mathbb{R}$. On a :

$$\ell'(\theta|\text{data}) = N_0 \frac{\partial_\theta \overline{F}_S(\tau;\theta)}{\overline{F}_S(\tau;\theta)} + \sum_{r=1}^{m} K_r \frac{\partial_\theta \overline{F}_S((r-1)\delta) - \partial_\theta \overline{F}_S(r\delta)}{\overline{F}_S((r-1)\delta) - \overline{F}_S(r\delta)},$$

où ∂_θ est la dérivée partielle par rapport à θ. Ainsi, $\widehat{\theta}_n$ est solution de l'équation $\ell'(\theta|\text{data}) = 0$. En appliquant une version de la δ-méthode pour des vecteurs aléatoires définis implicitement proposée par Benichou et Gail [60] et en utilisant la convergence en loi de $(K_1,\ldots,K_m)$ vers un vecteur gaussien, la normalité asymptotique de $\widehat{\theta}_n$ en découle. La variance asymptotique est obtenue en calculant l'information de Fisher. La dérivée seconde (par rapport à θ) de la fonction de log-vraisemblance peut être décomposée de la manière suivante :

$$\ell''(\theta|\text{data}) = N_0 g(\theta) + \sum_{r=1}^{m} K_r g_r(\theta),$$

où

$$g(\theta) = \frac{\left[\partial^2_{\theta^2}\overline{F}_S(\tau;\theta)\right]\overline{F}_S(\tau;\theta) - \left[\partial_\theta \overline{F}_S(\tau;\theta)\right]^2}{\left[\overline{F}_S(\tau;\theta)\right]^2}$$

avec $\partial^2_{\theta^2}$ désignant la dérivée partielle seconde par rapport à θ et où

$$g_r(\theta) = \frac{1}{\left[\overline{F}_S((r-1)\delta;\theta) - \overline{F}_S(r\delta;\theta)\right]^2}\Big\{ - \left[\partial_\theta \overline{F}_S((r-1)\delta;\theta) - \partial_\theta \overline{F}_S(r\delta;\theta)\right]^2 + \left[\partial^2_{\theta^2}\overline{F}_S((r-1)\delta;\theta) - \partial^2_{\theta^2}\overline{F}_S(r\delta;\theta)\right]\left[\overline{F}_S((r-1)\delta;\theta) - \overline{F}_S(r\delta;\theta)\right]\Big\}.$$

Alors, l'information de Fisher est égale à :

$$I(\theta) = -\overline{F}_S(\tau;\theta)g(\theta) - \sum_{r=1}^{m}\left[\overline{F}_S((r-1)\delta;\theta) - \overline{F}_S(r\delta;\theta)\right] g_r(\theta).$$

Exemple 3.5.1 Supposons que S est de loi exponentielle de paramètre λ. Dans ce cas, l'estimateur du maximum de vraisemblance $\widehat{\lambda}_n$ est calculable explicitement :

$$\widehat{\lambda}_n = \frac{1}{\delta}\log\left(\frac{N_0\tau + \delta\sum_{r=1}^{m} rK_r}{N_0\tau + \delta\sum_{r=1}^{m}(r-1)K_r}\right).$$

En appliquant les résultats précédents, on a :

$$\sqrt{n}\left(\widehat{\lambda}_n - \lambda\right) \xrightarrow[n\to\infty]{d} N(0,\rho^2),$$

où $\rho^2 = \frac{(e^{\lambda\delta}-1)^2}{\delta^2 e^{\lambda\delta}(1-e^{-\lambda\tau})}$. On notera que la variance asymptotique ρ^2 est décroissante avec δ décroissant. De plus, on peut facilement calculer sa limite quand δ tend vers zéro (dans ce cas, les temps d'initiation sont observés de manière exacte, mais tronqués à droite) :

$$\rho^2 \xrightarrow[\delta\to 0]{} \frac{\lambda^2}{1-e^{-\lambda\tau}}.$$

Ensuite, quand τ tend vers l'infini, cette limite tend vers λ^2 qui est la variance asymptotique de l'estimateur du maximum de vraisemblance lorsque les durées sont observées de manière exacte et sans troncature.

Estimation de μ et σ^2

Pour tout $i \in \mathcal{N}_{2+}$ et tout $j \in \{1,\ldots,m-R_i\}$, on pose $\Delta X_{i,j} = X_i((R_i+j)\delta) - X_i((R_i+j-1)\delta)$. Ces accroissements sont des variables aléatoires i.i.d. de loi normale de moyenne $\mu\delta$ et de variance $\sigma^2\delta$. On en déduit un estimateur naturel pour μ :

$$\widehat{\mu}_n = \frac{\displaystyle\sum_{i\in\mathcal{N}_{2+}}\sum_{j=1}^{m-R_i}\Delta X_{i,j}}{\delta\displaystyle\sum_{i\in\mathcal{N}_{2+}}(m-R_i)} = \frac{1}{\delta Q_n}\sum_{h=1}^{Q_n} Z_h,$$

où $Z_1,\ldots,Z_{Q_n}$ sont les accroissements rangés dans l'ordre lexico-graphique (par exemple). De même, un estimateur naturel pour σ^2 est le suivant :

$$\widehat{\sigma}_n^2 = \frac{1}{\delta(Q_n-1)}\sum_{h=1}^{Q_n}(Z_h - \delta\widehat{\mu}_n)^2.$$

Ces estimateurs s'expriment donc comme des sommes d'un nombre aléatoire de variables aléatoires. De nombreux articles ont étudiés des théorèmes de la limite centrale pour un nombre aléatoire d'observations. Ici, nous pouvons appliquer, par exemple, un résultat de Rényi [161] pour montrer la normalité asymptotique de $\widehat{\mu}_n$ et de $\widehat{\sigma}_n^2$, en utilisant le lemme 3.5.1.

Proposition 3.5.1

1. *$\widehat{\mu}_n$ est asymptotiquement normale :*

$$\sqrt{Q_n}\,(\widehat{\mu}_n - \mu) \xrightarrow[n\to\infty]{d} N\left(0, \frac{\sigma^2}{\delta}\right).$$

   et

$$\sqrt{n}\,(\widehat{\mu}_n - \mu) \xrightarrow[n\to\infty]{d} N\left(0, \frac{\sigma^2}{\tau\alpha(m,\tau)}\right),$$

   où $\alpha(m,\tau)$ est donné dans le lemme 3.5.1.

2. *$\widehat{\sigma}_n^2$ est asymptotiquement normale :*

$$\sqrt{Q_n}\left(\widehat{\sigma}_n^2 - \sigma^2\right) \xrightarrow[n\to\infty]{d} N\left(0, 2\sigma^4\right).$$

3.5.3 Estimation du temps moyen jusqu'à la panne

Comme pour tout modèle de dégradation, on peut définir le temps de panne comme T_c le premier temps d'atteinte d'un niveau critique c donné :

$$T_c = \inf\{t \geq 0; X(t) \geq c\}.$$

Soit $\tilde{T}_c$ le premier temps d'atteinte du niveau c par un processus de Wiener partant de 0 à l'instant 0. On rappelle que $\tilde{T}_c$ est de loi inverse gaussienne de paramètres c/μ et c/σ (voir [75] pour une référence générale sur cette loi). Pour le modèle étudié ici, T_c et $S + \tilde{T}_c$ sont égaux en loi, S et $\tilde{T}_c$ étant indépendants. Il en résulte que le temps moyen jusqu'à panne ($MTTF$) est tout simplement donné par :

$$MTTF = \mathbb{E}[S] + \frac{c}{\mu}.$$

Supposons que $\theta \in \Theta \subset \mathbb{R}$. Le temps moyen jusqu'à panne peut être estimé par :

$$\widehat{\mathbb{E}[S]} = \int_0^\infty \overline{F}_S(u; \widehat{\theta}_n)\mathrm{d}u.$$

Sous les mêmes hypothèses de régularité sur $\overline{F}_S$ (par rapport à θ) que précédemment, on en déduit que $\widehat{\mathbb{E}[S]}$ est un estimateur asymptotiquement normal de l'espérance de S en appliquant la δ-méthode :

$$\sqrt{n}\left(\widehat{\mathbb{E}[S]} - \mathbb{E}[S]\right) \xrightarrow[n\to\infty]{d} N\left(0, I(\theta)^{-1}\left(\int_0^\infty \partial_\theta \overline{F}_S(u;\theta)\mathrm{d}u\right)^2\right).$$

Il en découle que $\widehat{MTTF}_n = \widehat{\mathbb{E}[S]} + \frac{c}{\widehat{\mu}_n}$ est un estimateur asymptotiquement normal de $MTTF$ avec une variance asymptotique égale à :

$$I(\theta)^{-1}\left(\int_0^\infty \partial_\theta \overline{F}_S(u;\theta)\mathrm{d}u\right)^2 + \frac{c^2\sigma^2}{\mu^4\tau\alpha(m,\tau)}.$$

Chapitre 4

Prise en compte des incertitudes dans certains modèles en fiabilité

Une question importante, mais guère peu abordée dans la littérature, porte sur l'intégration et/ou l'impact des incertitudes sur une politique de remplacement ou de maintenance. Ce problème peut être abordé de deux manières possibles : soit une approche bayésienne, soit une approche fréquentiste. Dans la première section, avec N. Bousquet (EDF R&D), M. Fouladirad (Troyes) et A. Grall (Troyes) [7], nous avons proposé une procédure d'estimation bayésienne pour le processus gamma dans le but d'optimiser une politique de maintenance conditionnelle. Ensuite, dans la seconde section, en collaboration avec M. Fouladirad (Troyes) et A. Grall (Troyes), nous avons opté pour une approche fréquentiste pour l'étude d'une politique de remplacement. Les paramètres d'un modèle de durée de vie n'étant pas connus, ils sont estimés et donc on n'obtient au final qu'un estimateur du délai optimal et du coût optimal : quelles sont les propriétés de ces estimateurs ? On peut considérer ce problème comme une étude de sensibilité.

4.1 Inférence bayésienne pour le processus gamma [7]

Comme expliqué en introduction, un des principaux problèmes est celui de l'inférence statistique des modèles de dégradation. La plupart du temps, une approche fréquentiste est considérée. Quand on dispose d'avis d'experts, une démarche bayésienne constitue une alternative intéressante et peut améliorer la robustesse des estimateurs [23-25]. La capacité des approches bayésiennes à gérer les incertitudes dans des modèles de dégradation a été expliquée dans [26]. Mener une étude bayésienne informative induit une première étape d'explicitation de la loi *a priori* qui reflète bien l'information fournie par les experts. Cette tâche est difficile puisque soumise à des choix subjectifs. Il est donc fondamental que ce choix puisse être compris des experts, des ingénieurs fiabilistes et des preneurs de décision, comme cela a été indiqué dans [27]. Une seconde tâche est la proposition d'unz règle simple et claire pour l'estimation du délai optimal de maintenance. Souvent, un système est supposé être en panne dès que la mesure de dégradation est supérieure à un seuil critique pré-déterminé, ce qui conduit à une opération de maintenance corrective. Les fondements théoriques de la décision bayésienne permet de construire de tels estimateurs possédant des propriétés souhaitées [23-25], en minimisant une fonction de coût moyennée par la loi *a posteriori*.

Dans ce travail, nous proposons une approche permettant de répondre à ces deux objectifs dans le cadre du processus gamma homogène. On notera que des approches bayésiennes ont déjà été employées pour le processus de Wiener homogène ou non-homogène dans [28-30]. Dans la sous-section suivante, on commence par décrire le modèle d'échantillonnage, inspirée par un jeu de données disponibles chez EDF R&D. Puis, l'approche bayésienne est détaillée, y compris dans ses aspects calculatoires. Ensuite, plusieurs fonctions de coût sont proposées pour définir une règle appropriée de maintenance.

4.1.1 Modèle statistique

On considère un ensemble de M composants identiques soumis à un stress important pouvant provoquer une fissuration, par exemple. On suppose donc que, pour tout $k \in \{1, \ldots, M\}$, la détérioration du k-ème composant est caractérisée à l'instant t par une longueur de fissure $X_{k,t}$: $X_{k,t}$ est donc positif et croissant dans le temps. Une fois apparue, une fissure ne peut pas être observée avant qu'elle ne soit d'une longueur z à cause des limitations des appareils de mesure (par exemple, les ultra-sons). On suppose que le composant est en panne si la fissure est de longueur $\mu > z$. Notons t_{r_k} le dernier instant avant que la fissure soit observable.

$$t_{r_k+1} \quad = \quad \min_{i \in \{1,\ldots,n_k\}} \{t_i,\ X_{k,t_i} \geq z\}\,.$$

Pour tout $k \in \{1, \ldots, M\}$, la vraisemblance associée au k-ème composant va donc comporter différents termes de censure en fonction t_{r_k}. Un terme de censure à gauche est lié au fait que la fissure ne soit pas observable entre les instants $t_0 = 0$ et t_{r_k} (en supposant que $X_{k,t_0} = 0$) :

$$P\left(X_{k,t_1} \leq X_{k,t_2} \leq \ldots \leq X_{k,t_{r_k}} < z\right) = \frac{\gamma\left(\alpha t_{r_k}, \frac{z}{\beta}\right)}{\Gamma\left(\alpha t_{r_k}\right)}$$

où $\gamma(a,b) = \int_0^b x^{a-1}e^{-x}\,dx$ est la fonction gamma incomplète inférieure. Un terme de censure à droite est lié au fait que l'on observe $\tilde{Z}_{k,r_k+1} = X_{k,t_{r_k+1}} - z$, qui est une borne inférieure de Z_{k,r_k+1} :

$$P(Z_{k,r_k+1} > x_{k,t_{r_k+1}} - z) = \frac{\Gamma\left(\alpha(t_{r_k+1} - t_{r_k}), (x_{k,t_{r_k+1}} - z)/\beta\right)}{\Gamma\left(\alpha(t_{r_k+1} - t_{r_k})\right)}$$

où $\Gamma(a,b) = \int_b^\infty x^{a-1}e^{-x}\,dx$ est la fonction gamma incomplète supérieure. Au final, la contribution des observations $\mathbf{d}_k$ du k-ème composant à la fonction de vraisemblance correspond à l'un des trois cas suivants :

1. si aucune mesure exacte n'est observée :

$$\ell_k(\alpha, \beta|\mathbf{d}_k) = \frac{\gamma\left(\alpha t_{n_k}, \frac{z}{\beta}\right)}{\Gamma\left(\alpha t_{n_k}\right)};$$

2. si une seule mesure exacte est observée,

$$\ell_k(\alpha, \beta|\mathbf{d}_k) = \frac{\gamma\left(\alpha t_{n_k-1}, \frac{z}{\beta}\right)\Gamma\left(\alpha(t_{n_k} - t_{n_k-1}), (x_{k,t_{n_k}} - z)/\beta\right)}{\Gamma\left(\alpha t_{n_k-1}\right)\Gamma\left(\alpha(t_{n_k} - t_{n_k-1})\right)};$$

3. si plus de deux mesures exactes sont observées,

$$\ell_k(\alpha, \beta|\mathbf{d}_k) = \frac{\gamma\left(\alpha t_{r_k}, \frac{z}{\beta}\right)\Gamma\left(\alpha(t_{r_k+1} - t_{r_k}), (x_{k,t_{r_k+1}} - z)/\beta\right)}{\Gamma\left(\alpha t_{r_k}\right)\Gamma\left(\alpha(t_{r_k+1} - t_{r_k})\right)\prod_{i=r_k+2}^{n_k}\Gamma(\alpha(t_i - t_{i-1}))}$$
$$\times\ \beta^{-\alpha\left(t_{n_k} - t_{r_k+1}\right)}\exp\left\{-\frac{1}{\beta}\sum_{i=r_k+2}^{n_k} z_{k,i}\right\}\left[\prod_{i=r_k+2}^{n_k} z_{k,i}^{t_i - t_{i-1}}\right]^\alpha.$$

La fonction de vraisemblance $\ell(\alpha, \beta|\mathbf{d})$ associée aux observations $\mathbf{d}$ est le produit des contributions $\ell_1(\alpha, \beta|\mathbf{d}_1), \ldots, \ell_M(\alpha, \beta|\mathbf{d}_M)$ des M composants.

4.1.2 Estimation bayésienne

Un cadre bayésien est donc proposé ici pour estimer $\theta = (\alpha, \beta) \in \mathbb{R}_+^2$, mais également toute fonction des paramètres $h(\theta)$ (par exemple, un temps optimal de maintenance qui sera défini plus loin). Plus précisément, le vecteur des paramètres θ est considéré comme aléatoires. On lui adjoint une loi *a priori* de densité $\pi(\theta)$ et l'estimation consiste alors à obtenir la loi *a posteriori* $\pi(\theta|\mathbf{d}) \propto \pi(\theta)\ell(\theta|\mathbf{d})$ à partir des données $\mathbf{d}$ à travers la fonction de vraisemblance $\ell(\theta|\mathbf{d})$.

Cadre bayésien objectif

Pour éviter des problèmes bien connus portant sur le choix de la loi *a priori* [162], une démarche bayésienne objective peut être proposée [114, 191]. Particulièrement défendu par Clarke [76], une mesure *a priori*, non-informative et pertinente, $\pi(\theta)$ est celui de Jeffreys. Il est défini comme la racine carrée du déterminant de la matrice d'information de Fisher et il est donc intrinsèquement lié à la forme de la fonction de vraisemblance (voir [114] pour plus de détails).

Dans tout contexte bayésien objectif, la loi *a posteriori* est plus influencée par les données que par la loi *a priori*. Aussi, les contributions correspondant aux termes de censure dans la fonction de vraisemblance ne seront pas pris en compte ici afin de faciliter la construction de la loi *a posteriori*. De plus, afin de s'affranchir d'un effet d'échelle lié à l'hétérogénéité des pas de temps entre deux observations (pour une trajectoire mais aussi entre individus), la loi *a priori* est défini à partir de la loi gamma standard, c'est-à-dire pour la vraisemblance de la première occurrence du processus Z. D'après [192], on a :

$$\pi(\alpha,\beta)s \propto \frac{1}{\beta}\sqrt{\alpha\Psi_1(\alpha)-1}$$

où $\Psi_1(.)$ est la fonction tri-gamma. Il n'est pas possible de déterminer la loi *a posteriori* de manière explicite. Une alternative consiste donc à utiliser des simulations de la loi *a posteriori* en utilisant des techniques numériques spécifiques comme les méthodes de Monte Carlo par chaînes de Markov. Cependant, il est possible de prendre en compte la structure des données manquantes dans le modèle pour implémenter l'algorithme de Gibbs [163]. En effet, si toutes les statistiques manquantes $S_k = \sum_{i=1}^{r_k} Z_{k,i}$ pouvaient être connues (directement ou bien reconstituées) sous la contrainte que $S_k \leq z$), la loi *a posteriori* de β conditionnellement à α et $\mathbf{d}$ serait la loi inverse gamma (notée $\mathcal{IG}$) :

$$\beta|\alpha,\mathbf{d} \sim \mathcal{IG}\left(\alpha\sum_{k=1}^{M} t_{n_k,k}, \sum_{k=1}^{M}\left[S_k + \sum_{i=r_k+1}^{n_k} z_{k,i}\right]\right).$$

Une telle approche ne peut pas être menée pour la loi conditionnelle *a posteriori* de α. La méthode de Metropolis-Hastings peut alors être utilisée à chaque pas dans l'algorithme de Gibbs pour produire une chaîne de Markov ergodique [163]. L'algorithme complet de simulation est décrit dans [7].

Cadre bayésien subjectif (informatif)

Comme expliqué précédemment, l'explicitation d'une loi *a priori* $\pi(\theta)$ informative est plutôt délicate puisque cette approche introduit souvent un niveau de subjectivité lié au modèle. Cependant, des informations supplémentaires, telles que des avis d'experts, peuvent parfois être d'une plus grande importance que les observations qui peuvent être rares à cause de la fréquence des inspections et des limites des appareils de mesure. Pour cette raison, il est donc utile de fournir une méthodologie claire pour la construction de la loi *a priori*, sous la forme d'un modèle paramétrique pour $\pi(\theta)$, les paramètres devant avoir une interprétation concrète pour les utilisateurs (experts et ingénieurs fiabilistes). A notre connaissance, aucun travail n'a été publié dans ce sens pour le processus gamma, dans le contexte de la fiabilité. La plupart du temps (dans des articles en fiabilité ou pas), la loi gamma est utilisée comme loi *a priori* pour des modèles non-paramétriques ou semi-paramétriques d'intensité de défaillance [132, 108, 141]. Une approche bayésienne non-paramétrique pour le processus gamma a été étudiée dans [176]. L'algorithme complet de simulation est décrit dans [7].

4.1.3 Règle de décision pour la maintenance

Une fois les paramètres du modèle estimés, le temps de panne peut être prédit et une stratégie de maintenance peut être proposée. On rappelle que les fissures ne sont pas observées en temps continu et donc les données ne sont disponibles qu'aux temps d'inspection.

On note T_μ l'instant de panne, c'est-à-dire le premier temps d'atteinte du seuil μ :

$$T_\mu = \inf\{t > 0; X_t > \mu\}.$$

La fonction de répartition de T_μ est donnée par :

$$\mathbb{P}(T_\mu \leq t) = \iint \mathbb{P}(X_t \geq \mu|\alpha,\beta)\pi(\alpha,\beta|\mathbf{d})d\alpha d\beta = \iint \frac{\Gamma(\alpha t, \mu/\beta)}{\Gamma(\alpha t)}\pi(\alpha,\beta|\mathbf{d})d\alpha d\beta.$$

Pour une fonction de perte L donnée, l'instant du remplacement préventif est défini comme étant l'estimateur bayésien suivant :

$$\widehat{T}_\mu = \text{argmin}_x \iiint L(x,t)\frac{\partial\Gamma(\alpha t,\mu/\beta)/\Gamma(\alpha t)}{\partial t}\pi(\alpha,\beta|\mathbf{d})d\alpha d\beta dt \tag{4.1.1}$$

où, d'après Park et Padgett [145],

$$\begin{aligned}\forall t \geqslant 0\ , \quad \frac{\partial\Gamma(\alpha t,\mu/\beta)/\Gamma(\alpha t)}{\partial t} &= \alpha\left(\Psi(\alpha t) - \log\left(\frac{\mu}{\beta}\right)\right)\frac{\gamma(\alpha t,\mu/\beta)}{\Gamma(\alpha t)} \\ &\quad + \frac{1}{\alpha t^2\Gamma(\alpha t)}\left(\frac{\mu}{\beta}\right)^{\alpha t}{}_2F_2(\alpha t,\alpha t;\alpha t+1,\alpha t+1;-F/\beta)\end{aligned}$$

est la densité de T_μ, avec Ψ la fonction digamma et ${}_2F_2$ la fonction hypergéométrique généralisée d'ordre $(2,2)$. La méthode d'inversion de la fonction de répartition peut être utilisée pour simuler une telle variable aléatoire, pour un couple (α,β) donné. Il en résulte que toute fonction de T_μ peut être estimée par le biais de simulations de Monte Carlo.

Le choix d'une fonction de perte est donc d'une grande importance dans l'optique d'une stratégie de maintenance. Différents choix de L vont mener à différentes stratégies de maintenance, certaines étant plus conservatives que d'autres. Néanmoins, il n'est pas évident pour un ingénieur fiabiliste de déterminer la fonction de perte. Ainsi, des choix simples et faciles à comprendre doivent être faits pour des raisons pratiques. Trois différentes fonctions de perte vont être considérées ici pour répondre à ces objectifs. Leur différence réside principalement sur la manière dont les remplacements sont pénalisés.

La fonction de perte quadratique $L_1(x,t) = ||x-t||^2$ pénalise davantage les valeurs éloignées de celles attendues du temps de panne. Ainsi, cette fonction est adaptée s'il est important d'éviter des remplacements ni trop précoces, ni trop tardifs. Dans ce cas, selon la théorie classique bayésienne [162], l'instant de remplacement $\widehat{T}_\mu$ est l'espérance de la loi *a posteriori* définie par :

$$\widehat{T}_\mu = \iiint t\frac{\partial\Gamma(\alpha t,\mu/\beta)/\Gamma(\alpha t)}{\partial t}\pi(\alpha,\beta|\mathbf{d})d\alpha d\beta dt$$

qui peut donc être évaluée par le biais d'une méthode de Monte Carlo comme expliqué précédemment. Si $\beta\mu$ est assez grand, l'espérance conditionnelle de T_μ, $\mathbb{E}(T_\mu)$, est approximativement égale à $\frac{1}{\alpha}(\beta\mu+\frac{1}{2})$ [61], alors

$$\widehat{T}_\mu \simeq \iint \frac{1}{\alpha}(\beta\mu+\frac{1}{2})\pi(\alpha,\beta|\mathbf{d})d\alpha d\beta.$$

Il s'agit de l'estimateur bayésien le plus utilisé en pratique.

Une autre fonction de perte est celle de l'erreur absolue : $L_2(x,t) = |x-t|$. Elle traite de manière similaire que la fonction de perte précédente les petits et les grands écarts aux valeurs attendues. Pour cette fonction de perte, les instants de maintenance préventive ou corrective sont traités de la même manière que les écarts, ce qui implique que les pannes ne sont pas nécessairement plus coûteuses qu'un remplacement préventif. Ainsi, cette deuxième fonction de perte est plus adaptée à des systèmes pour lesquels toute action de maintenance est coûteuse et donc pour lesquels il n'y a aucun avantage à effectuer des maintenances trop tôt. Pour cette fonction de perte, l'instant de remplacement se calcule de la manière suivante :

ce qui implique que $\mathbb{P}(T_\mu \leq \widehat{T}_\mu) = \frac{1}{2}$. Alors, l'instant optimal de remplacement $\widehat{T}_\mu$ est la médiane de la loi marginale *a posteriori* de T_μ.

Enfin, considérons une troisième et dernière fonction de perte définie par : $L_3(x,t) = C_1|x-t|\mathbb{1}_{x\geq t} + C_2|x-t|\mathbb{1}_{x\leq t}$. Elle est conçue pour pénaliser les remplacements tardifs dans la mesure où la panne (dépassement

du seuil μ) induit un coût C_1 alors que l'action de remplacement préventif induit un coût $C_2 \leq C_1$. Cette fonction de perte va donc conduire à faire des remplacements avant la panne. Dans ce cas, l'instant optimal de remplacement est la solution de l'équation suivante :

$$\frac{\partial}{\partial x} \iiint L_3(x,t) \frac{\partial \Gamma(\alpha t, \mu/\beta)/\Gamma(\alpha t)}{\partial t} \pi(\alpha, \beta | \mathbf{d}) dt d\alpha d\beta = 0$$

ce qui donne :

$$\mathbb{P}(T_\mu \leq \widehat{T}) = \frac{C_1}{C_1 + C_2}.$$

Ainsi, l'instant optimal de remplacement $\widehat{T}_\mu$ est le quantile d'ordre $C_1/C_1 + C_2$ de la loi marginale *a posteriori* de T_μ (qui peut être, ici aussi, évaluée à l'aide d'une méthode de Monte Carlo). Plus l'écart entre C_1 et C_2 augmente, plus les actions de maintenance corrective sont coûteuses et plus l'instant de remplacement sera donc court.

4.2 Analyse de sensibilité pour une politique de remplacement

Pour un modèle paramétrique de durées de vie, on peut déterminer le délai optimal entre deux inspections pour la politique de remplacement par blocs. Cependant, ce délai optimal dépend des paramètres de la loi des durées de vie qui sont, en général, inconnus. Ces paramètres peuvent être estimés, par exemple, à partir d'un échantillon de durées de vie et cela fournit donc un estimateur du délai optimal. Il est alors légitime de se poser la question suivante : quelle est la variabilité induite par la phase d'estimation ?

Après avoir fait quelques rappels sur cette politique de remplacement, des résultats généraux sur son analyse de sensibilité sont établis, en se basant sur l'hypothèse de normalité asymptotique des estimateurs des paramètres. Cette hypothèse permet d'utiliser la δ-méthode et des variantes de celle-ci. Enfin, le cas de la loi exponentielle est détaillée.

4.2.1 Résultats généraux sur l'analyse de sensibilité

On commence par rappeler la définition de la politique de remplacement étudiée, ainsi qu'un certain nombre de résultats connus (voir le chapitre 5 dans [136], par exemple). Puis, en utilisant une généralisation de δ-méthode au cas de variables aléatoires définies implicitement, on établit des résultats généraux sur l'analyse de sensibilité.

Rappel sur la politique de remplacement

Soit T la durée de vie d'un composant. Supposons que la loi de T dépend de paramètres $\theta \in \Theta \subset \mathbb{R}^p$. On utilise les notations usuelles : $f_T(\cdot;\theta)$ pour la densité de T, $F_T(\cdot;\theta)$ pour sa fonction de répartition et $S_T(\cdot;\theta)$ pour sa fonction de survie. On suppose que l'état du composant n'est connu qu'au moment des inspections : la panne d'un composant n'est donc pas détectée immédiatement. Un composant décelé en panne lors d'une inspection est remplacé par un composant neuf, la durée de remplacement étant supposée négligeable. Il existe deux coûts associés à cette politique de remplacement par blocs : le coût c_r de remplacement d'un composant en panne par un composant neuf et le coût c_u d'indisponibilité. On suppose que $c_u > c_r$.

Cette politique ne dépend que d'un seul paramètre δ, le délai entre deux inspections consécutives. Le choix du délai mène à un coût. On cherche donc à déterminer le délai optimal $\delta^\star$ qui minimise ce coût. Afin de déterminer ce délai optimal, on considère le coût unitaire asymptotique défini de la manière suivante :

$$C(\delta;\theta) = \lim_{t\to\infty} \frac{C_t(\delta)}{t},$$

où $C_t(\delta)$ est le coût sur l'intervalle de temps $[0,t]$ quand le composant est inspecté aux instants $(k\delta)_{k\in\mathbb{N}}$. D'après la théorie du renouvellement [89], on a :

$$C(\delta;\theta) = \frac{\text{Coût moyen sur un cycle}}{\text{Longueur moyenne d'un cycle}}.$$

Pour la politique de remplacement par blocs, cela donne :

$$C(\delta;\theta) = \frac{\mathbb{E}[c_r + c_u(\delta - T)_+]}{\delta} = \frac{c_r + c_u \int_0^\delta F_T(u;\theta)\mathrm{d}u}{\delta}.$$

Alors,

$$\lim_{\delta\to 0} C(\delta;\theta) = +\infty \quad \text{et} \quad \lim_{\delta\to+\infty} C(\delta;\theta) = c_u.$$

Soit $\delta^\star := \mathrm{argmin}_{\delta>0} C(\delta;\theta)$. En dérivant cette fonction de coût, on obtient que $\delta^\star$ est la racine de la fonction suivante (par rapport à δ) :

$$\phi(\delta;\theta) = \mathbb{E}[T\mathbf{1}_{T\leqslant\delta}] - \frac{c_r}{c_u} \tag{4.2.1}$$

où

$$\mathbb{E}[T\mathbf{1}_{T\leqslant\delta}] = \int_0^\delta u f_T(u;\theta)\mathrm{d}u = -\delta S_T(\delta;\theta) + \int_0^\delta S_T(u;\theta)\mathrm{d}u \tag{4.2.2}$$

(en rappelant que $\frac{c_r}{c_u} \leqslant 1$). Cette solution ne dépend des deux coûts qu'à travers leur ratio (et donc ne dépend pas de l'unité monétaire considéré). Puisque, pour tout $\theta \in \Theta$, $\delta \mapsto \phi(\delta;\theta)$ est une fonction croissante tendant vers $\mathbb{E}[T] - c_r/c_u$ (qui dépend de θ) avec $\phi(0;\theta) = -c_r/c_u$, il en découle que $\delta^\star$ est fini si $\mathbb{E}[T] > c_r/c_u$ et infini autrement. Sous cette hypothèse de finitude de $\delta^\star$, il est bien connu que le coût optimal vérifie l'équation suivante :

$$C^\star = C(\delta^\star;\theta) = c_u F_T(\delta^\star;\theta).$$

Il en découle que le délai optimal est solution de l'équation :

$$\delta^\star = F_T^{-1}(C^\star/c_u;\theta),$$

où $F_T^{-1}(\cdot;\theta)$ est la fonction quantile de T. En remplaçant cela dans la fonction ϕ et après calculs, on obtient que $C^\star$ est défini implicitement par la fonction ψ suivante :

$$\psi(C^\star;\theta) = \int_0^{C^\star/c_u} F_T^{-1}(u;\theta)\mathrm{d}u - \frac{c_r}{c_u} = 0.$$

Analyse de sensibilité

Soit $\theta_0 \in \Theta$ la vraie valeur du paramètre pour la loi de la durée de vie. Comme expliqué précédemment, le délai optimal $\delta_0^\star$ entre deux inspections successives peut être calculé numériquement en cherchant le zéro de la fonction $\phi(\cdot;\theta_0)$ (s'il existe). La plupart du temps, le paramètre θ_0 est inconnu et peut être estimé à partir d'un échantillon de temps de panne. Soit $\widehat{\theta}_n$ un estimateur de θ, par exemple l'estimateur du maximum de vraisemblance. Supposons que cet estimateur possède de "bonnes" propriétés, comme la convergence :

$$\widehat{\theta}_n \xrightarrow[n\to\infty]{Pr} \theta_0, \tag{4.2.3}$$

et comme la normalité asymptotique :

$$\sqrt{n}\left(\widehat{\theta}_n - \theta_0\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0,\Sigma_0^2), \tag{4.2.4}$$

où la matrice de variance-covariance Σ_0^2 dépend de θ_0. Puisque θ_0 est inconnu, on peut remplacer θ_0 par $\widehat{\theta}_n$ afin d'obtenir une estimation de $\delta_0^\star$. Un problème naturel est alors le suivant : quelles sont les propriétés satisfaites par $\widehat{\delta}_n^\star$? De même, le coût optimal est inconnu mais peut être estimé par $\widehat{C}_n^\star$: quelles sont les propriétés satisfaites par $\widehat{C}_n^\star$? Comment est éloignée cette estimation de sa vraie valeur $C_0^\star$?

Un outil pratique pour répondre à ces questions est la δ-méthode qui peut être formulée de la manière suivante :

Théorème 4.2.1 *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de vecteurs aléatoires à valeurs dans $\mathbb{R}^p$. Supposons qu'il existe $\mu_X \in \mathbb{R}^p$ et Σ une matrice définie positive tels que*

$$\sqrt{n}\,(X_n - \mu_X) \xrightarrow[n\to\infty]{d} N(0, \Sigma).$$

Soient $f_1, \ldots, f_q$ des fonctions réelles admettant des dérivées partielles continues en μ_X, avec au moins une dérivée partielle non nulle en ce point. Pour $i \in \{1, \ldots, q\}$ et pour tout $n \in \mathbb{N}^*$, on pose $Y_{i,n} = f_i(X_n)$, $Y_n = (Y_{1,n}, \ldots, Y_{q,n})^T$ et $\mu_X = (f_1(\mu_X), \ldots, f_q(\mu_X))^T$. Alors, la suite $(Y_n)_{n\in\mathbb{N}^*}$ vérifie la propriété suivante :

$$\sqrt{n}\,(Y_n - \mu_Y) \xrightarrow[n\to\infty]{d} N(0, K\Sigma K^T),$$

où K est la matrice $q \times p$ avec $k_{i,j} = \partial f_i/\partial x_j$ pour $i \in \{1, \ldots, p\}$ et $j \in \{1, \ldots, q\}$.

Ce résultat permet, ici, de montrer que l'estimateur de la fonction de coût est ponctuellement asymptotiquement normal. Cependant, il ne permet pas d'en conclure que l'estimateur du délai optimal et l'estimateur du coût optimal sont aussi asymptotiquement normaux. En effet, $\delta_0^\star$ et $C_0^\star$ sont les solutions d'une équation implicite et donc la δ-méthode classique ne permet pas de traiter ce cas-là. Il nous faut donc une version adaptée à notre situation. Un tel résultat existe et a été montré par Benichou et Gail [60].

Théorème 4.2.2 *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite comme dans le théorème précédent. Soient $\mu_X \in \mathbb{R}^p$ et $\mu_Y \in \mathbb{R}^q$. Soient $g_1, \ldots, g_q$ un ensemble de q fonctions continues de $\mathbb{R}^p \times \mathbb{R}^q$ dans $\mathbb{R}$ avec des dérivées partielles continues sur un ouvert contenant (μ_X, μ_Y). Soit Y_n un vecteur aléatoire à valeurs dans $\mathbb{R}^q$ tel que $g_r(X_n, Y_n) = 0$ pour tout $r \in \{1, \ldots, q\}$. Soit $J_{x,y}$ une matrice $q \times q$ avec $\frac{\partial g_i}{\partial y_j}(x, y)$ et soit $H_{x,y}$ une matrice $q \times p$ avec $\frac{\partial g_i}{\partial x_j}(x, y)$. Si $|J_{\mu_X,\mu_Y}| \neq 0$ et si chaque ligne de $J^{-1}_{\mu_X,\mu_Y} H_{\mu_X,\mu_Y}$ contient au moins un élément non nul, alors*

$$\sqrt{n}\,(Y_n - \mu_Y) \xrightarrow[n\to\infty]{d} N(0, J^{-1}_{\mu_X,\mu_Y} H_{\mu_X,\mu_Y} \Sigma H^T_{\mu_X,\mu_Y} (J^{-1}_{\mu_X,\mu_Y})^T).$$

On commence par considérer un estimateur de la fonction de coût obtenu par plug-in :

$$\forall \delta > 0, \quad C(\delta; \widehat{\theta}_n) = \frac{c_r + c_u \int_0^\delta F_T(u; \widehat{\theta}_n)\mathrm{d}u}{\delta}.$$

En appliquant le théorème de la continuité séquentielle et en supposant que l'équation (4.2.3) est satisfaite, il s'agit d'un estimateur ponctuellement convergent :

$$\forall \delta > 0, \quad C(\delta; \widehat{\theta}_n) \xrightarrow[n\to\infty]{Pr} C(\delta; \theta_0).$$

La proposition suivante établit la normalité asymptotique (également de manière ponctuelle). Il s'agit d'une application directe de la δ-méthode (voir le théorème 4.2.1).

Théorème 4.2.3 *Supposons que l'équation (4.2.4) est satisfaite et que $\theta \mapsto C(\delta; \theta)$ est différentiable pour tout $\delta > 0$. Soit $\nabla_\theta C(\delta; \theta)$ le gradient (par rapport à θ) de la fonction de coût. Si $\nabla_\theta C(\delta; \theta)$ est continu et si $\nabla_\theta C(\delta; \theta_0) \neq 0_{\mathbb{R}^p}$, alors $C(\delta; \widehat{\theta}_n)$ est un estimateur (ponctuel) asymptotiquement normal de $C(\delta; \theta_0)$:*

$$\forall \delta > 0, \quad \sqrt{n}\left(C(\delta; \widehat{\theta}_n) - C(\delta; \theta_0)\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \sigma^2_{cost}),$$

avec $\sigma^2_{cost} = \nabla_\theta C(\delta; \theta_0) \Sigma_0 \nabla_\theta C(\delta; \theta_0)^T$.

En appliquant le théorème 4.2.2, on peut aussi montrer que $\widehat{\delta}_n^\star$ est également un estimateur asymptotiquement normal de $\delta_0^\star$ sous certaines conditions de régularité de la fonction ϕ.

Théorème 4.2.4 *Supposons que l'équation (4.2.4) est satisfaite et que $\theta \mapsto \phi(\delta; \theta)$ est différentiable pour tout $\delta > 0$. Soit $\nabla_\theta \phi(\delta; \theta)$ le gradient de ϕ (par rapport à θ). Si $\nabla_\theta \phi(\delta; \theta)$ est continue avec $\nabla_\theta \phi(\delta; \theta_0) \neq 0_{\mathbb{R}^p}$ et si $f_T(\delta_0^\star; \theta_0) \neq 0$, alors $\widehat{\delta}_n^\star$ est un estimateur asymptotiquement normal de $\delta_0^\star$:*

$$\sqrt{n}\left(\widehat{\delta}_n^\star - \delta_0^\star\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \sigma^2_{opt.delay}),$$

où $\sigma^2_{opt.delay}$ *est donné par :*

$$\sigma^2_{opt.delay} = \frac{\nabla_\theta \phi(\delta_0^\star;\theta_0)\Sigma_0 \nabla_\theta \phi(\delta_0^\star;\theta_0)^T}{\left[\delta_0^\star f_T(\delta_0^\star;\theta_0)\right]^2}.$$

De la même manière, le théorème 4.2.2 peut être utilisé pour montrer la normalité asymptotique du délai optimal $\widehat{C}_n^\star$.

Théorème 4.2.5 *Supposons que l'équation (4.2.4) est satisfaite et que* $\theta \mapsto \psi(C^\star;\theta)$ *est différentiable pour tout* $C^\star > 0$. *Soit* $\nabla_\theta \psi(C^\star;\theta)$ *le gradient de* ψ *(par rapport à* θ*). Si* $\nabla_\theta \psi(C^\star;\theta)$ *est continue avec* $\nabla_\theta \psi(C_0^\star;\theta_0) \neq 0_{\mathbb{R}^p}$, *alors* $\widehat{C}_n^\star$ *est un estimateur asymptotiquement normal de* $C_0^\star$ *:*

$$\sqrt{n}\left(\widehat{C}_n^\star - C_0^\star\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \sigma^2_{opt.cost}),$$

où $\sigma^2_{opt.cost}$ *est donné par :*

$$\sigma^2_{opt.cost} = \frac{c_u^2 \nabla_\theta \psi(C_0^\star;\theta_0)\Sigma_0 \nabla_\theta \psi(C_0^\star;\theta_0)^T}{\left[F_T^{-1}(C_0^\star/cu;\theta_0)\right]^2}.$$

Il est immédiat de voir que le gradient de ψ (par rapport à θ) est égal à :

$$\nabla_\theta \psi(C_0^\star;\theta_0) = \int_0^{C_0^\star/c_u} \nabla_\theta F_T^{-1}(u;\theta_0)\mathrm{d}u.$$

4.2.2 Exemple de la loi exponentielle

On suppose ici que le temps jusqu'à la panne T est une variable aléatoire de loi exponentielle de paramètre inconnu $\lambda_0 \in \mathbb{R}_+ = \Theta$:

$$\forall t \geq 0, \quad S_T(t) = e^{-\lambda_0 t}.$$

Supposons qu'on observe un n-échantillon $T_1, \ldots, T_n$ de même loi que T. L'estimateur du maximum de vraisemblance de λ_0 est donné par :

$$\widehat{\lambda}_n = \frac{n}{\sum_{i=1}^n T_i}.$$

Il est bien connu que c'est un estimateur asymptotiquement normal :

$$\sqrt{n}\left(\widehat{\lambda}_n - \lambda_0\right) \xrightarrow[n\to\infty]{d} \mathcal{N}(0, \lambda_0^2).$$

Pour un telle loi, la fonction de coût devient :

$$\begin{aligned} C(\delta;\lambda_0) &= \frac{1}{\delta}\left\{c_r + c_u \int_0^\delta \left(1 - \exp\left(-\lambda_0 u\right)\right)\mathrm{d}u\right\} \\ &= \frac{1}{\delta}\left\{c_r + c_u\delta - \frac{c_u}{\lambda_0}\left(1 - e^{-\lambda_0\delta}\right)\right\}. \end{aligned}$$

Pour toute valeur fixée de δ, la dérivée partielle de C par rapport à λ_0 est égale à :

$$\partial_\lambda C(\delta;\lambda_0) = \frac{c_u}{\delta}\left(\frac{1}{\lambda_0^2} - \left(\frac{\delta}{\lambda_0} + \frac{1}{\lambda_0^2}\right)e^{-\lambda_0\delta}\right).$$

Il en découle que, pour tout $\delta > 0$, la variance asymptotique dans le théorème 4.2.3 vaut :

$$\sigma^2_{cost} = c_u^2\left[\frac{1}{\lambda_0\delta} - \left(1 + \frac{1}{\lambda_0\delta}\right)e^{-\lambda_0\delta}\right]^2.$$

Maintenant, examinons la fonction ϕ introduite précédemment. Pour la loi exponentielle, on obtient :

$$\phi(\delta;\lambda_0) = \frac{1}{\lambda_0} - \left(\delta + \frac{1}{\lambda_0}\right)e^{-\lambda_0\delta} - \frac{c_r}{c_u}.$$

La dérivée partielle d'ordre 1 de ϕ est égale à :

$$\partial_\delta \phi(\delta; \lambda_0) = \lambda_0 \delta e^{-\lambda_0 \delta}$$

et

$$\partial_\lambda \phi(\delta; \lambda_0) = -\frac{1}{\lambda_0^2} + \left(1 + \lambda_0 \delta - \frac{1}{\lambda_0^2}\right) e^{-\lambda_0 \delta}.$$

Il en découle que la variance asymptotique dans le théorème 4.2.4 est égale à :

$$\sigma^2_{opt.delay} = \left[\frac{\delta_0^\star e^{-\lambda_0 \delta_0^\star}}{(1 + \lambda_0^2 + \delta_0^\star \lambda_0^3) e^{-\lambda_0 \delta_0^\star} - 1}\right]^2.$$

Enfin, en utilisant l'expression de la fonction quantile de la loi exponentielle, on obtient que le coût optimal $C_0^\star$ satisfait l'équation suivante :

$$\psi(C_0^\star; \lambda_0) = \left(1 - \frac{C_0^\star}{c_u}\right) \log\left(1 - \frac{C_0^\star}{c_u}\right) + \frac{C_0^\star}{c_u} - \lambda_0 \frac{c_r}{c_u} = 0.$$

Alors, la dérivée partielle d'ordre 1 de ϕ est égale à :

$$\partial_\delta \psi(C^\star; \lambda_0) = -\frac{1}{c_u} \log\left(1 - \frac{C_0^\star}{c_u}\right)$$

et

$$\partial_\lambda \psi(C^\star; \lambda_0) = -\frac{c_r}{c_u}.$$

D'où, on obtient l'expression suivante pour la variance asymptotique dans le théorème 4.2.5 :

$$\sigma^2_{opt.cost} = \left[\lambda_0 \log\left(1 - \frac{C_0^\star}{c_u}\right)\right]^2.$$

Chapitre 5

Comparaison d'échantillons de petite taille et applications

Un autre problème avec un fort enjeu industriel est celui de la surveillance de durées de vie. Par exemple, une entreprise manufacturière doit assurer la qualité et la fiabilité des biens qu'elle produit. Pour cela, elle dispose de durées de vie issues du retour d'expérience (REX=. Ces données permettent d'estimer le temps moyen jusqu'à panne (MTTF). Au cours du temps, des durées de fonctionnement sont donc observées et on s'intéresse ici à la détection d'un éventuel moment à partir duquel les durées de vie sont plus "courtes". Une telle modification dans les unités produites peut être due à un changement de fournisseur, à un changement dans le processus de fabrication, à un nouveau processus de maintenance, etc. Cependant, les produits étant de plus en plus fiables, ils ont des taux de panne faibles. Cela amène à travailler sur des échantillons de petites ou moyennes tailles. Un problème important est donc le suivi en ligne et la détection précoce d'une rupture dans les données.

Dans ce chapitre, deux approches connexes sont considérées afin de répondre à ces défis. La première consiste à proposer un test de détection de rupture. Dans ces modèles-là, on suppose que, s'il y a rupture, alors l'échantillon est divisée en deux sous-échantillons, chaque sous-échantillon étant constitué de variables aléatoires indépendantes et de même loi (mais les deux lois sous-jacentes ne sont pas les mêmes). La difficulté est que l'on ne sait pas s'il y a rupture, ni l'instant (le cas échéant) de rupture. Il s'agit d'un travail en collaboration avec N. Balakrishnan (Université de McMaster, Canada), L. Bordes (UPPA) et J.-C. Turlot (UPPA) et qui trouve son inspiration dans un contrat industriel avec Alstom Transport [24, 23]. Plus "dynamique", la seconde approche repose sur un outil issu de la maîtrise statistique des procédés : les cartes de contrôle. A partir d'un échantillon de référence et pour une statistique donnée, on détermine une zone de contrôle. Ensuite, à chaque nouvel échantillon, on teste si sa statistique est dans la zone de contrôle ou pas. Avec N. Balakrishnan et J.-C. Turlot [1], nous avons proposé des cartes de contrôles unilatérales pour le suivi de durées de vie.

En collaboration G. Verdier (UPPA), nous avons également travaillé sur la construction de cartes de contrôle pour le suivi de données physico-chimiques mesurant la qualité des eaux, dans le cadre d'un contrat de recherche avec Rivage Pro Tech, filiale de la Lyonnaise des Eaux.

5.1 Tests de détection d'une rupture dans un petit échantillon [24, 23]

Le problème de la détection de rupture dans une suite de variables aléatoires indépendantes est étudié depuis plusieurs décennies. Les premiers articles sur le sujet remontent au moins aux années 1950 [142, 143, 144]. Cependant, la plupart des études se situent dans un cadre asymptotique (voir le livre de Csörgö et Horváth [80], par exemple) et les vitesses de convergence sont en général très lentes. Ces résultats ne sont pas toujours utilisables quand on doit traiter des échantillons de petite taille. Nous avons ainsi développé

des approches permettant d'obtenir de meilleures puissances. La méthodologie générale est détaillée dans la première sous-section. Puis, dans les sous-sections suivantes, deux approches sont développées, l'une étant basée sur l'hypothèse de loi exponentielle pour les durées de vie, l'autre étant non-paramétrique. Ce choix est lié au fait qu'en pratique, la loi exponentielle est très souvent utilisée. Cependant, il est parfois difficile de valider ce choix (par un avis d'expert ou par un test statistique - qui serait peu puissant pour un petit échantillon). C'est pourquoi nous avons proposé un autre test non-paramétrique basé sur la statistique de Wilcoxon-Mann-Whitney (WMW).

5.1.1 Description de la méthodologie proposée

Soient $X_1, \ldots, X_n$ des variables aléatoires représentant des durées de vie, des durées inter-événements, etc. Elles sont supposées être indépendantes, mais on n'est pas certain qu'elles soient toutes de même loi. En particulier, on souhaite détecter un éventuel instant de rupture dans cette suite. On dit qu'il y a une rupture s'il existe un entier $k \in \{1, \ldots, n-1\}$ tel que $X_1, \ldots, X_k$ sont de même loi F et $X_{k+1}, \ldots, X_n$ sont de même loi G avec $F \neq G$ (ici, en fait, on s'intéresse au cas où $F < G$). La difficulté dans ce type de problème est qu'on ne sait pas s'il y a rupture et, s'il y a rupture, où elle se produit. On doit donc considérer tous les sous-échantillons possibles obtenus en séparant l'échantillon complet en deux parties consécutives. Dans [24, 23], la méthodologie proposée est basée sur le schéma suivant :

1. découpage de l'échantillon en deux sous-échantillons : $(X_1, \ldots, X_k)$ et $(X_{k+1}, \ldots, X_n)$, pour $k \in \{m, \ldots, n-m\}$ avec $m \geqslant 1$;
2. calcul de la statistique de test pour chaque découpage en deux sous-échantillons ;
3. utilisation de toutes ces statistiques pour construire une décision quant à la présence éventuelle d'une rupture.

La décision finale est, soit l'absence d'une rupture, soit la présence d'un instant de rupture $k^* \in \{m, \ldots, n-m\}$. Cependant, l'identification de k^* n'est pas un objectif prioritaire dans ce travail.

On se donne un test de comparaison des deux sous-échantillons $(X_1, \ldots, X_k)$ et $(X_{k+1}, \ldots, X_n)$. On note par $S_{n,k}$ la statistique qui mesure la "distance" entre les lois sous-jacentes aux deux sous-échantillons. En se basant sur les $n - 2m + 1$ statistiques possibles, on propose plusieurs statistiques de test globales :

1. statistique de type maximum :
$$M_n = \max_{m \leqslant k \leqslant n-m} \frac{S_{n,k}}{\sqrt{\text{var}(S_{n,k})}};$$
2. statistique de type χ^2 de première espèce :
$$\chi^2_n = \sum_{k=m}^{n-m} \frac{S^2_{n,k}}{\text{var}(S_{n,k})};$$
3. statistique de type χ^2 de seconde espèce :
$$\tilde{\chi}^2_n = \sum_{k=m}^{n-m} \left(\frac{S_{n,k}}{\text{var}(S_{n,k})} \right)^2;$$
4. statistique de type quadratique :
$$Q_n = \mathbf{S}_n^T \Sigma^{-1} \mathbf{S}_n,$$
où $\mathbf{S}_n^T = (S_{n,m}, \ldots, S_{n,n-m})$ et Σ est la matrice de variance-covariance de $\mathbf{S}_n$;
5. statistique linéaire :
$$U_n = \sum_{k=m}^{n-m} w_{k,n} S_{n,k},$$
où les poids $w_{m,n}, \ldots, w_{n-m,n}$ sont choisis afin de minimiser la variance de U_n sous la contrainte que $\mathbb{E}[U_n] = c$ pour une valeur de c donnée.

On observe qu'on a imposé que les sous-échantillons contiennent un nombre minimal d'observations (égal à m) afin d'assurer une estimation qui ne soit pas trop incorrecte des grandeurs définissant la statistique de test. Le choix de la valeur m sera précisé plus loin. Enfin, bien qu'on s'intéresse plus à la détection de rupture qu'à la localisation de celle-ci (le cas échéant), on notera tout de même que la statistique de type maximum permet de répondre à cette dernière question si on le souhaite.

5.1.2 Une approche basée sur la loi exponentielle

Nous supposons ici que les durées de vie $X_1, \ldots, X_n$ sont des variables aléatoires indépendantes et de loi exponentielle (mais donc pas nécessairement de même paramètre). Pour deux sous-échantillons $\{X_1, \ldots, X_k\}$ et $\{X_{k+1}, \ldots, X_n\}$, on souhaite donc tester si les deux paramètres sont égaux (H_0 : pas de rupture) ou pas (H_1 : une rupture). La statistique classiquement utilisée pour ce problème est le rapport de vraisemblance. Sous l'hypothèse nulle, le rapport de vraisemblance est égal à :

$$\Lambda_{k,n} = \frac{\left(\frac{n}{\sum_{i=1}^{n} X_i}\right)^n}{\left(\frac{k}{\sum_{i=1}^{k} X_i}\right)^k \left(\frac{n-k}{\sum_{i=k+1}^{n} X_i}\right)^{n-k}}.$$

Dans notre situation, on rejettera l'hypothèse nulle H_0 quand le maximum du logarithme du rapport de vraisemblance Z_n défini par

$$Z_n = \max_{m \leq k \leq n-m} \{-2 \log \Lambda_{k,n}\}$$

est trop grand. La valeur critique peut être obtenue, soit en utilisant des résultats asymptotiques (voir [80], par exemple), soit par le biais de simulation de Monte Carlo (surtout quand on dispose d'un échantillon de petite ou moyenne taille). La construction des statistiques globales basées sur $\Lambda_{k,n}$ s'avère être compliquée, à cause du calcul des variances et covariances. De plus, une étude numérique semble indiquer que les puissances de ces tests sont relativement faibles quand n est petit. C'est pour ces raisons que nous avons introduit une autre statistique de comparaison basée sur le rapport des moyennes empiriques des deux sous-échantillons :

$$T_{n,k} = \frac{n-k}{k} \frac{T_k}{T_n - T_k}$$

avec

$$T_k = \sum_{i=1}^{k} X_i \quad \text{et} \quad T_n - T_k = \sum_{i=k+1}^{n} X_i.$$

La statistique suivante est une statistique non biaisée du rapport des taux de risque et pourra donc naturellement servir de statistique de comparaison :

$$S_{n,k}^{(1)} = \frac{n-k-1}{k} \frac{T_k}{T_n - T_k}.$$

Sous l'hypothèse nulle, $\mathbb{E}\left(S_{n,k}^{(1)}\right) = 1$. Cette statistique tend à être plus grande que 1 lorsque la fréquence de pannes augmente. La matrice de variance-covariance de $\mathbf{S}_n^{(1)} = (S_{n,m}^{(1)}, \ldots, S_{n,n-m}^{(1)})^T$ est donnée dans la proposition suivante.

Proposition 5.1.1 *Pour tout $k \in \{m, \ldots, n-m\}$ avec $m \geqslant 2$,*

$$var\left(S_{n,k}^{(1)}\right) = \frac{n-1}{k(n-k-2)},$$

et, pour tout $(k, k') \in \{m, \ldots, n-m\}^2$ tel que $k' > k$ et $m \geqslant 3$,

$$cov\left(S_{n,k}^{(1)}, S_{n,k'}^{(1)}\right) = \frac{n-1}{k'(n-k-2)}.$$

L'inverse de la matrice de variance-covariance $\Sigma^{(1)} = (\sigma_{ij}^{(1)})_{m \leqslant i,j \leqslant n-m}$ de $\mathbf{S}_n^{(1)}$ peut être calculée explicitement. On peut montrer que, pour $i \leqslant j$,

$$(\Sigma^{(1)})_{ij}^{-1} = \begin{cases} \dfrac{m(m+1)(n-m-2)^2}{(n-1)(n-2)} & \text{si} \quad i=j=m, \\ \dfrac{2(n-2-i)^2 i^2}{(n-1)(n-2)} & \text{si} \quad i=j=m+1,\ldots,n-m-1, \\ \dfrac{(m-2)(m-1)(n-m)^2}{(n-1)(n-2)} & \text{si} \quad i=j=n-m, \\ -\dfrac{(n-i-3)(n-i-2)(i+1)i}{(n-1)(n-2)} & \text{si} \quad j=i+1 \quad \text{et} \quad i=m,\ldots n-m-1, \\ 0 & \text{si} \quad j>i+1. \end{cases}$$

La statistique $Q_n^{(1)}$ de type quadratique peut alors s'écrire sous la forme suivante :

$$Q_n^{(1)} = \frac{n-1}{n-2}\left[\sum_{k=m}^{n-m-1}\left(\frac{S_{n,k}^{(1)}}{\text{var}\left(S_{n,k}^{(1)}\right)} - \frac{S_{n,k+1}^{(1)}}{\text{var}\left(S_{n,k+1}^{(1)}\right)}\right)^2 + \frac{1}{m}\left(\frac{S_{n,m}^{(1)}}{\text{var}\left(S_{n,m}^{(1)}\right)}\right)^2 + \frac{1}{m-2}\left(\frac{S_{n,n-m}^{(1)}}{\text{var}\left(S_{n,n-m}^{(1)}\right)}\right)^2\right],$$

où on rappelle que $\text{var}\left(S_{n,k}^{(1)}\right) = \frac{n-1}{k(n-k-2)}$ pour $k \in \{m,\ldots,n-m\}$.

Considérons maintenant la statistique linéaire. Soit $\mathbf{w}_n^{(1)} = (w_{m,n}^{(1)},\ldots,w_{n-m,n}^{(1)})^T$ des poids tels que $U_n^{(1)} = \mathbf{w}_n^{(1)T}\mathbf{S}_n^{(1)}$. On souhaiter déterminer les poids $\mathbf{w}_n^{(1)}$ tels que $\text{var}(U_n^{(1)}) = \mathbf{w}_n^{(1)T}\Sigma^{(1)}\mathbf{w}_n^{(1)}$ soit minimale avec la contrainte que $\mathbb{E}[U_n^{(1)}] = c$. On montre que la meilleure statistique linéaire ne dépend pas de c et est que les poids sont donnés par :

$$\mathbf{w}_n^{(1)} = \frac{(\Sigma^{(1)})^{-1}\mathbf{1}}{\mathbf{1}^T(\Sigma^{(1)})^{-1}\mathbf{1}}.$$

On remarquera que $(\Sigma^{(1)})^{-1}\mathbf{1}$ est égal à la somme des lignes de $(\Sigma^{(1)})^{-1}$ alors que $\mathbf{1}^T(\Sigma^{(1)})^{-1}\mathbf{1}$ est tout simplement égal à la somme de tous les éléments de $(\Sigma^{(1)})^{-1}$. En utilisant ces expressions pour les éléments de $(\Sigma^{(1)})^{-1}$ et après quelques calculs, on obtient que la combinaison linéaire optimale de $\mathbf{S}_n^{(1)}$ est $U_n^{(1)} = \sum_{k=m}^{n-m} w_{k,n}^{(1)} S_{n,k}^{(1)}$, où $\mathbf{w}_n^{(1)} = (w_{n,m}^{(1)},\ldots,w_{n,n-m}^{(1)})^T$ est donné par :

$$\begin{array}{lcl} w_{n,m}^{(1)} & = & m(m+1)(n-m+2)/\Delta, \\ w_{n,k}^{(1)} & = & 2k((n-2-k)/\Delta, \quad \text{pour } k \in \{m+1,\ldots,n-m-1\}, \\ w_{n,n-m}^{(1)} & = & (m-2)(m-1)(n-m)/\Delta, \end{array}$$

et

$$\Delta = m(m+1)(n-m-2) + 2\sum_{k=m+1}^{n-m-1} k(n-2-k) + (m-2)(m-1)(n-m).$$

5.1.3 Une approche non-paramétrique

Dans cette partie, on suppose toujours que les variables aléatoires sont indépendantes, mais on ne formule plus d'hypothèse paramétrique sur leur loi. Une manière de comparer non-paramétriquement deux sous-échantillons est le test de Wilcoxon-Mann-Whitney. On rappelle que cette statistique, notée ici $S_{n,k}^{(2)}$, est définie par :

$$S_{n,k}^{(2)} = \sum_{i=1}^{k}\sum_{j=k+1}^{n} 1_{\{X_j < X_i\}}.$$

L'espérance et la variance de $S_{n,k}^{(2)}$ sont bien connues :

$$\mathbb{E}\left(S_{n,k}^{(2)}\right) = \frac{k(n-k)}{2}$$

et

$$\text{var}\left(S_{n,k}^{(2)}\right) = \frac{k(n-k)(n+1)}{12}.$$

La covariance entre $S_{n,k}^{(2)}$ et $S_{n,k'}^{(2)}$ peut aussi être calculée. On peut facilement montrer que, pour tout $(k,q) \in \{1,\ldots,n\}^2$ tel que $q > k$,

$$\text{cov}\left(S_{n,k}^{(2)}, S_{n,q}^{(2)}\right) = \frac{k(n-q)(n+1)}{12}$$

(ce résultat figure dans un article de Pettitt [151], mais sans preuve). La loi de $S_{n,k}^{(2)}$ n'est pas calculable explicitement, mais il est possible de trouver une récurrence.

Partant de ces statistiques, on peut en déduire des statistiques globales comme expliqué précédemment. D'autres statistiques globales basées sur ces statistiques de Wilcoxon-Mann-Whitney ont été proposées dans la littérature [169, 151, 152, 62, 128]. En utilisant les expressions données ci-dessus, on peut calculer l'inverse de $\Sigma^{(2)}$:

$$(\Sigma^{(2)})^{-1}_{ij} = \begin{cases} (m+1)/mn & \text{si} \quad i=j=1, \\ 2/n & \text{si} \quad i=j=2,\ldots,n-2m, \\ (m+1)/mn & \text{si} \quad i=j=n-2m+1, \\ -1/n & \text{si} \quad j=i+1 \quad \text{and} \quad i=1,\ldots n-2m, \\ 0 & \text{si} \quad j>i+1. \end{cases}$$

Alors, la statistique $Q_n^{(2)}$ se décompose comme la somme d'une statistique quadratique $Q_n^\star$ et d'un terme de correction :

$$Q_n^{(2)} = \frac{24}{n(n+1)} Q_n^\star - \frac{12(m-1)}{n(n+1)m}\left(\left(S_{n,m}^{(2)}\right)^2 + \left(S_{n,n-m}^{(2)}\right)^2\right),$$

où

$$Q_n^\star = \sum_{k=m}^{n-m}\left(S_{n,k}^{(2)}\right)^2 - \sum_{k=m}^{n-m-1} S_{n,k}^{(2)} S_{n,k+1}^{(2)}.$$

On peut également voir cette statistique comme la somme du carré de différences consécutives, à un facteur de correction près :

$$Q_n^{(2)} = \frac{12}{n(n+1)} \sum_{k=1}^{n-m-1}\left(S_{n,k}^{(2)} - S_{n,k+1}^{(2)}\right)^2 + \frac{12}{n(n+1)m}\left(\left(S_{n,m}^{(2)}\right)^2 + \left(S_{n,n-m}^{(2)}\right)^2\right).$$

Comme précédemment, on peut déterminer la statistique linéaire optimale de $\mathbf{S}_n^{(2)} = (S_{n,m}^{(2)},\ldots,S_{n,n-m}(2))$, notée par $U_n^{(2)} = {\mathbf{w}_n^{(2)}}^T \mathbf{S}_n^{(2)}$ selon le même critère. En utilisant les expressions pour les éléments de $(\Sigma^{(2)})^{-1}$ et après quelques calculs, on obtient que la combinaison linéaire optimale de $\mathbf{S}_n^{(2)}$ est $U_n^{(2)} = \sum_{k=m}^{n-m} w_{k,n}^{(2)} S_{n,k}^{(2)}$, où $\mathbf{w}_n = (w_{n,m}^{(2)},\ldots,w_{n,n-m}^{(2)})^T$ est donné par :

$$\begin{array}{lcl} w_{n,m}^{(2)} & = & (m+1)/\Delta, \\ w_{n,k}^{(2)} & = & 2/\Delta, \quad \text{pour } k \in \{m+1,\ldots,n-m-1\}, \\ w_{n,n-m}^{(2)} & = & (m+1)/\Delta, \end{array}$$

avec

$$\Delta = m(m+1)(n-m) + 2\sum_{k=m+1}^{n-m-1} k(n-k).$$

5.2 Cartes de contrôle unilatérale pour durées de vie [1]

On s'intéresse ici à un problème connexe à la détection de rupture : les cartes de contrôle. Les cartes de contrôle constituent un outil important dans le domaine de la maîtrise statistique des procédés (on parle aussi de contrôle de la qualité). On dispose d'un échantillon de référence (ou échantillon sous contrôle) que l'on compare successivement à des échantillons tests. La collecte des données permettant l'obtention de l'échantillon de référence et la détermination des limites de contrôle constituent la phase I, alors que la période de comparaisons d'échantillons est la phase II. On notera $X_1, \ldots, X_m$ (resp. $Y_1, \ldots, Y_n$) l'échantillon de référence (resp. un échantillon test) : ce sont des variables aléatoires i.i.d. de fonction de répartition F (resp. G). En général, les échantillons tests sont de plus petites tailles que l'échantillon de référence. Depuis quelques années, plusieurs cartes de contrôle non-paramétriques ont été étudiées, voir [1] pour un état de l'art ainsi que les autres articles dans le même numéro spécial.

Nous nous intéressons ici au suivi de durées de vie. Plus précisément, nous souhaitons détecter au plus tôt une "réduction" dans les durées de vie. Il est donc plus naturel de considérer des cartes de contrôles unilatérales, plutôt que des cartes bilatérales qui sont plus largement étudiées dans la littérature. De même, pour l'étude des performances des cartes de contrôle, l'hypothèse de translation est la plupart du temps formulée (c'est-à-dire qu'on suppose que $G(\cdot) = F(\cdot - \mu)$). Or, cette hypothèse n'est pas compatible avec la "réduction" des durées de vie. Aussi, nous avons choisi de considérer les deux hypothèses alternatives de Lehmann : (1) $G(\cdot) = F(\cdot)^\gamma$; (2) $G(\cdot) = 1 - (1 - F(\cdot))^\gamma$. Dans notre cas, on supposera que $\gamma \in]0, 1]$ pour la première alternative de Lehmann et que $\gamma \in [1, +\infty[$ pour la seconde alternative de Lehmann. Ces deux caractéristiques (cartes unilatérales et alternatives de Lehmann) sont des éléments originaux de ce travail.

Deux familles de cartes de contrôle sont proposées ci-dessous. La première repose sur la statistique des prédécesseurs et la seconde sur une version pondérée de cette statistique. Dans chacun des cas, on étudie les performances en terme de fausse alarme et de longueurs de run.

5.2.1 Carte de contrôle unilatérale basée sur la statistique des prédécesseurs

La statistique des prédécesseurs d'ordre b compte le nombre d'observations de l'échantillon test plus petite que la b-ème statistique d'ordre de l'échantillon de référence, notée $X_{b:m}$:

$$\forall b \in \{1, \ldots, m\}, \quad P_{(b)} = \sum_{i=1}^{n} \mathbb{I}_{]-\infty, X_{b:m}]}(Y_i) \in \{0, \ldots, n\}.$$

Le lemme suivante donne la loi de $P_{(b)}$ dans le cas général.

Lemme 5.2.1 *Pour toutes lois F et G absolument continues, on a, pour tout $c \in \{0, \ldots, n\}$,*

$$\mathbb{P}[P_{(b)} = c] = \frac{\binom{n}{c}}{B(b, m-b+1)} \int_0^1 \left[GF^{-1}(u)\right]^c \left[1 - GF^{-1}(u)\right]^{n-c} u^{b-1}(1-u)^{m-b} \mathrm{d}u,$$

et :

$$\mathbb{P}[P_{(b)} \geqslant c] = \frac{1}{B(c, n+1-c)B(b, m-b+1)} \sum_{h=0}^{n-c} \frac{(-1)^h}{c+h} \binom{n-c}{h} \int_0^1 \left[GF^{-1}(u)\right]^{c+h} u^{b-1}(1-u)^{m-b} \mathrm{d}u,$$

où $B(\cdot, \cdot)$ est la fonction beta.

On notera que la loi de $P_{(b)}$ dépend de F et G qu'à travers la composée de fonctions GF^{-1}. Cela permet d'obtenir des expressions très simples sous les hypothèses alternatives de Lehmann.

Calcul du taux de fausse alarme

Le résultat suivant est bien connu (voir [56], par exemple) et découle du lemme précédent.

Proposition 5.2.1 *Sous l'hypothèse $G = F$ (c'est-à-dire que le processus est sous contrôle),*

$$\forall c \in \{0, \dots, n\}, \quad \mathbb{P}[P_{(b)} = c] = \frac{\binom{b+c-1}{c}\binom{m+n-b-c}{n-c}}{\binom{m+n}{n}}.$$

Clairement, de grandes valeurs de $P_{(b)}$ traduisent le fait que le processus devient hors contrôle. Ainsi, la région critique est de la forme $\mathcal{W}_b = \{P_{(b)} \geqslant c_{b,\alpha}\} \Longleftrightarrow \{Y_{c_{b,\alpha}:n} < X_{b:m}\}$, où α est le taux de fausse alarme (FAR) nominal (fixé à l'avance). En utilisant la proposition précédente, il est possible de calculer la limite supérieure de contrôle (c'est-à-dire la valeur critique pour un test statistique) $c_{b,\alpha}$. Plus précisément, soit c tel que

$$\sum_{i=c+1}^{n} \frac{\binom{b+i-1}{i}\binom{m+n-b-i}{n-i}}{\binom{m+n}{n}} \leqslant \alpha < \sum_{i=c}^{n} \frac{\binom{b+i-1}{i}\binom{m+n-b-i}{n-i}}{\binom{m+n}{n}}.$$

Alors, la limite supérieure de contrôle associée à la statistique $P_{(b)}$ est donnée par :

$$c_{b,\alpha} = \begin{cases} c & \text{si } \alpha - \sum_{i=c+1}^{n} \mathbb{P}[P_{(b)} = i] \geqslant \sum_{i=c}^{n} \mathbb{P}[P_{(b)} = i] - \alpha \\ c+1 & \text{autrement.} \end{cases} \tag{5.2.1}$$

Une autre possibilité consiste à toujours choisir $c_{b,\alpha} = c + 1$ (valeur critique conservative). On remarquera que, pour tout $\alpha \in]0, 1[$ fixé, la limite de contrôle $c_{b,\alpha}$ n'existe pas toujours. Soit $b_{\alpha,max}$ tel que, pour tout $b \in \{1, \dots, b_{\alpha,max}\}$, la limite de contrôle $c_{b,\alpha}$ existe. En effet, pour tout $b > b_{\alpha,max}$, on a $\mathbb{P}[P_{(b)} \geqslant c] > \alpha$ pour $c \in \{1, \dots, n\}$, puisque ces variables aléatoires ont des queues à droite trop lourdes (et donc conduisent à un taux de fausse alarme nominal trop élevé). En prenant $c = n$ dans la proposition précédente, on obtient que $b_{\alpha,max}$ est le plus grand entier satisfaisant l'équation suivante :

$$\binom{b+n-1}{n} < \alpha \binom{m+n}{n}.$$

Calcul du taux d'alarme sous les hypothèses de Lehmann

On considère ici le taux d'alarme (AR) pour les deux hypothèses alternatives de Lehmann. On commence donc par déterminer la loi de $P_{(b)}$ sous ces hypothèses en utilisant le lemme précédent.

Proposition 5.2.2

1. *Sous la première hypothèse de Lehmann $G = F^\gamma$, pour tout $\gamma \in]0, 1[$ et pour tout $c \in \{0, \dots, n\}$, on a :*

$$\mathbb{P}[P_{(b)} = c] = \frac{\binom{n}{c}}{B(b, m-b+1)} \sum_{h=0}^{n-c} \binom{n-c}{h} (-1)^h B(b + (c+h)\gamma, m-b+1);$$

2. *Sous la seconde hypothèse de Lehmann $G = 1 - (1-F)^\gamma$, pour tout $\gamma \in]1, +\infty[$ et pour tout $c \in \{0, \dots, n\}$, on a :*

$$\mathbb{P}[P_{(b)} = c] = \frac{\binom{n}{c}}{B(b, m-b+1)} \sum_{h=0}^{c} \binom{c}{h} (-1)^h B(b, m-b+\gamma(n+h-c)+1).$$

Puisque $P_{(b)}$ est une variable aléatoire discrète, le taux de fausse alarme ne peut pas être exactement égal à la valeur nominale α. Pour pouvoir comparer les taux de fausse alarme pour différentes valeurs de b, on a considéré la procédure randomisée suivante :

$$\Phi = \begin{cases} 1 & \text{si } P_{(b)} \geqslant c_{b,\alpha} \quad \text{or} \quad \text{si } P_{(b)} = c_{b,\alpha} - 1 \text{ avec probabilité } \rho_{b,\alpha} \\ 0 & \text{si } P_{(b)} = c_{b,\alpha} - 1 \text{ avec probabilité } 1 - \rho_{b,\alpha} \quad \text{ou} \quad \text{si } P_{(b)} \leqslant c_{b,\alpha} - 2 \end{cases},$$

où $c_{b,\alpha}$ est le seuil conservatif précédemment défini et $\rho_{b,\alpha}$ est tel que :

$$\mathbb{P}[P_{(b)} \geqslant c_{b,\alpha}] + \rho_{b,\alpha} \mathbb{P}[P_{(b)} = c_{b,\alpha}] = \alpha.$$

Le taux d'alarme est alors égal à :

$$AR = \mathbb{P}[\mathcal{W}_b | G \neq F] + (1 - \rho_{b,\alpha}) \mathbb{P}[\partial \mathcal{W}_b | G \neq F],$$

où $\partial \mathcal{W}_b = \{P_{(b)} = c_{b,\alpha} - 1\}$.

Calcul de la période opérationnelle moyenne

Au lieu de considérer le taux d'alarme, la période opérationnelle moyenne ou longueur moyenne de run (ARL) est souvent utilisée pour évaluer la performance d'une carte de contrôle. Soit N la longueur de run de la carte de contrôle, c'est-à-dire le nombre d'échantillons tests avant d'en observer un hors contrôle. Pour la carte de contrôle étudiée ici, la loi de la longueur de run peut être calculée dans le cas général [166].

Proposition 5.2.3 *Pour toutes lois absolument continues F et G, la fonction de survie de N est donnée par :*

$$\forall k \in \mathbb{N}_*, \quad S(k) = \mathbb{P}[N > k] = \int_0^1 [p_{F,G}(u)]^k \frac{1}{B(b, m-b+1)} u^{b-1}(1-u)^{m-b} \mathrm{d}u,$$

où

$$p_{F,G}(u) = \frac{1}{B(c_{b,\alpha}, n - c_{b,\alpha} + 1)} \sum_{h=0}^{c_{b,\alpha}-1} \binom{c_{b,\alpha}-1}{h} (-1)^h \frac{[1 - GF^{-1}(u)]^{n+h-c_{b,\alpha}+1}}{n+h-c_{b,\alpha}+1}.$$

Puisque $ARL(F,G) = \mathbb{E}[N] = \sum_{k \geqslant 1} S(k)$, on peut facilement déduire du résultat précédent la longueur moyenne de run.

Corollaire 5.2.1 *Pour toutes lois absolument continues F et G, la longueur moyenne de run (ARL) pour la carte de contrôle unilatéral basée sur la statistique $P_{(b)}$ est donnée par :*

$$ARL(F,G) = \frac{1}{B(b, m-b+1)} \int_0^1 \frac{u^{b-1}(1-u)^{m-b}}{1 - p_{F,G}(u)} \mathrm{d}u,$$

où $p_{F,G}(u)$ est donné dans la proposition précédente.

Comme précédemment, il n'est pas possible d'atteindre de manière exacte une longueur moyenne de run fixée à l'avance. Aussi, nous avons utilisé une procédure randomisée pour déterminer la carte de contrôle la plus performante au regard de la longueur moyenne de run. Une telle approche est rarement considérée dans la littérature sur les cartes de contrôle. Notons ARL_0 la valeur nominale de la longueur moyenne de run. Pour m, n et b fixés, la longueur moyenne de run est une fonction croissante de c et donc on peut définir un seuil conservatif c_{b,ARL_0} de la manière suivante :

$$c_{b,ARL_0} = \min\{c \in \{1, \ldots, n\}; ARL_0(c) \geqslant ARL_0\},$$

où $ARL_0(c)$ est la longueur moyenne de run quand le processus est sous contrôle et pour une valeur critique égale à c. La décision randomisée est alors définie par :

$$\Phi = \begin{cases} 1 & \text{si } P_{(b)} \geqslant c_{b,ARL_0} \quad \text{ou} \quad \text{si } P_{(b)} = c_{b,ARL_0} - 1 \text{ avec probabilité } \rho_{b,ARL_0} \\ 0 & \text{si } P_{(b)} = c_{b,ARL_0} - 1 \text{ avec probabilité } 1 - \rho_{b,ARL_0} \quad \text{ou} \quad \text{si } P_{(b)} \leqslant c_{b,ARL_0} - 2 \end{cases},$$

pour une valeur $\rho_{b,ARL_0} \in]0,1[$ assurant que la longueur moyenne de run est exactement égale à ARL_0 quand le processus est sous contrôle (la constante ρ_{b,ARL_0} ne peut pas être déterminée de manière exacte, on doit se contenter d'un calcul numérique).

5.2.2 Carte de contrôle unilatéral basée sur la statistique pondérée des prédécesseurs

Comme expliqué dans [56], un effet masquant peut apparaître lorsqu'on considère la statistique des prédécesseurs. Un moyen de contourner ce problème est de considérer une version pondérée de cette statistique. La statistique des prédécesseurs peut être vue comme la somme des b premières statistiques de placement $N_1, \ldots, N_b$:

$$P_{(b)} = \sum_{j=1}^{b} N_j \,,$$

où N_j est le nombre d'observations dans l'échantillon test comprises $X_{(j-1:m)}$ et $X_{(j:m)}$:

$$\forall j \in \{1, \dots, m\}, \quad N_j = \sum_{i=1}^{n} \mathbb{I}_{]X_{j-1:m}, X_{j:m}]}(Y_i) \in \{0, \dots n\}$$

(avec, pour convention, que $X_{0:m} = 0$). On considère donc la statistique suivante :

$$P^{\star}_{(b)} = \sum_{j=1}^{b} (m - j + 1) N_j.$$

Les poids sont décroissants et accordent donc une plus grande importance aux premières statistiques de placement. Cette statistique est à valeurs dans $E_{m,n,b} = \{0\} \cup \{m - b + 1, \dots, mn\}$.

Calcul du taux de fausse alarme

Comme pour la statistique précédente, de grandes valeurs de $P^{\star}_{(b)}$ conduisent à croire que le processus est hors contrôle et donc la région critique est aussi de la forme $\{P^{\star}_{(b)} \geqslant c_{b,\alpha}\}$. La loi de $P^{\star}_{(b)}$ ne peut pas être calculé directement. Cependant, la loi jointe $(N_1, \dots, N_b)$, quand le processus est sous contrôle, est connue [56, 57].

Proposition 5.2.4 *Sous l'hypothèse que $G = F$, pour tout $(n_1, \dots, n_b) \in \{0, \dots, n\}^b$ avec $n_1 + \cdots + n_b \leqslant n$, on a :*

$$\mathbb{P}[N_1 = n_1, \dots, N_b = n_b] = \frac{\binom{m+n-\sum_{i=1}^{b} n_i - b}{m-b}}{\binom{m+n}{n}}.$$

On notera que ces probabilités ne dépendent que de $n_1 + \cdots + n_b$. Soit $\mathcal{A}_{b,c,s}$ l'ensemble suivant :

$$\mathcal{A}_{b,c,s} = \left\{ (n_1, \dots, n_b) \in \{0, \dots, n\}^b : \sum_{j=1}^{b} (m - j + 1) n_j = c \quad \text{et} \quad \sum_{j=1}^{b} n_j = s \right\}$$

pour tout $c \in E_{m,n,b}$ et pour tout $s \in \{0, \dots, n\}$. Dans [1], on a montré un lemme permettant de calculer par récurrence le cardinal de $\mathcal{A}_{b,c,s}$. On peut ainsi déterminer la loi de $P^{\star}_{(b)}$.

Proposition 5.2.5 *Sous l'hypothèse que $G = F$, pour tout $c \in E_{m,n,b}$, on a :*

$$\mathbb{P}[P^{\star}_{(b)} = c] = \sum_{s=0}^{n} \#\mathcal{A}_{b,c,s} \frac{\binom{m+n-s-b}{m-b}}{\binom{m+n}{n}}.$$

Ce résultat permet de calculer numérique la loi de $P^{\star}_{(b)}$ et donc d'en déduire la valeur critique. Soit c l'entier tel que $\mathbb{P}[P^{\star}_{(b)} \leqslant c] \leqslant \alpha < \mathbb{P}[P^{\star}_{(b)} \leqslant c + 1]$. Alors, le seuil critique $c_{b,\alpha}$ pour cette statistique est égale soit à c soit à $c + 1$ (voir la discussion pour la carte de contrôle précédente). Notons $\mathcal{N}_{b,\alpha}$ l'ensemble des b-uplets menant à une valeur dans la zone critique : $\mathcal{N}_{b,\alpha} = \cup_{c \geq c_{b,\alpha}} \cup_{s=0}^{n} \mathcal{A}_{b,c,s}$. La région critique est donc aussi de la forme : $\mathcal{W}_b = \{P_{(b)^{\star}} \geqslant c_{b,\alpha}\} = \{(N_1, \dots, N_b) \in \mathcal{N}_{b,\alpha}\}$.

Calcul du taux d'alarme sous les hypothèses de Lehmann

La loi du vecteur aléatoire $(N_1, \dots, N_b)$ peut être également facile à calculer pour les deux hypothèses alternatives de Lehmann

Proposition 5.2.6 *Soit κ la constante suivante :*

$$\kappa = \frac{n!m!}{\prod_{j=1}^{b} n_j! (n - \sum_{j=1}^{b} n_j)! (m - b)! \gamma^{b-1}}.$$

1. *Sous la première hypothèse de Lehmann* $G = F^\gamma$ *pour* $\gamma \in]0,1[$, *pour tout* $(n_1, \dots, n_b) \in \{0, \dots, n\}^b$ *avec* $n_1 + \cdots + n_b \leqslant n$,

$$\mathbb{P}[N_1 = n_1, \dots, N_b = n_b] = \kappa \prod_{j=1}^{b-1} B\left(\sum_{i=1}^{j} n_i + \frac{j}{\gamma}, n_{j+1} + 1\right) \times \sum_{h=0}^{n-\sum_{j=1}^{b} n_j} \binom{n - \sum_{j=1}^{b} n_j}{h} (-1)^h B\left(\gamma\left(\sum_{j=1}^{b} n_j + h\right) + b, m - b + 1\right).$$

2. *Sous la seconde hypothèse de Lehmann* $G = 1 - (1-F)^\gamma$ *pour* $\gamma \in]1, +\infty[$, *pour tout* $(n_1, \dots, n_b) \in \{0, \dots, n\}^b$ *avec* $n_1 + \cdots + n_b \leqslant n$,

$$\mathbb{P}[N_1 = n_1, \dots, N_b = n_b] = \kappa \prod_{j=1}^{b} B\left(n - \sum_{i=1}^{b-j+1} n_i + \frac{m - b + j}{\gamma}, n_{b-j+1} + 1\right) \times \sum_{h=0}^{n_1} \binom{n_1}{h} (-1)^h \left[\frac{1}{\gamma(n + h - n_1) + m}\right].$$

Calcul de la période opérationnelle moyenne

Le calcul de la longueur moyenne de run a déjà été considéré dans [57] pour une carte de contrôle bilatérale. Il est facile d'adapter ces calculs dans le cadre unilatéral :

$$ARL(F, G) = \int_{[0,1]^b} \frac{1}{1 - Q(GF^{-1}(u_1), \dots, GF^{-1}(u_p))} f_{1,\dots,b:m}(u_1, \dots, u_b) \mathrm{d}u_1 \cdots \mathrm{d}u_b,$$

où Q est la fonction suivante :

$$Q(v_1, \dots, v_p) = \sum_{c \geq c_{b,\alpha}} \sum_{s=0}^{n} \sum_{(n_1,\dots,n_b) \in \mathcal{A}_{b,c,s}} \frac{n!}{\prod_{j=1}^{b} n_j! (n - \sum_{j=1}^{b} n_j)!} v_1^{n_1} \prod_{j=2}^{b} (v_j - v_{j-1})^{n_j} (1 - v_b)^{n - \sum_{j=1}^{b} n_j},$$

et $f_{1,\dots,b:m}$ est la densité jointe de $(U_{1:m}, \dots, U_{b:m})$, le vecteur des b premières statistiques d'ordre de m variables aléatoires indépendantes et de loi uniforme sur $[0,1]$:

$$f_{1,\dots,b:m}(u_1, \dots, u_b) = \frac{m!}{(m-b)!} (1 - u_b)^{m-b} \mathbb{I}_{0 < u_1 < \cdots < u_b < 1}.$$

Ainsi, il est possible de calculer $ARL_0 = ARL(F, F)$ ainsi que la longueur moyenne de runs sous les deux hypothèses de Lehmann (aucune simplification dans les calculs n'apparaît ici).

Chapitre 6

Quelques perspectives de recherche

Dans ce chapitre, nous présentons quelques perspectives de recherche en lien avec les travaux exposés dans cette première partie. Ces problèmes portent tant sur des modèles de dégradation que sur des modèles liés aux durées de vie. Certains de ces projets sont déjà en cours d'étude et les premiers résultats ont parfois été présentés dans des congrès.

6.1 Modèles de dégradation

Cette section regroupe quelques travaux de recherche en cours ou à venir sur les modèles de dégradation à espace discret ou continu.

6.1.1 Estimation de la redondance dans des systèmes k-sur-n non réparables [30]

On considère ici un système constitué de n composants indépendants, identiques et non réparables. On suppose qu'un jugement d'expert sur le système indique que le système est en redondance de type k-sur-n mais que le nombre minimal k de composants en panne impliquant la panne du système est inconnue. De plus, on pourra également supposer que la loi de la durée de vie des composants est également inconnue.

On note $X_1, \ldots, X_n$ la durée de vie des n composants de densité commune f et de fonction de répartition commune F (on suppose que la loi des durées de vie est absolument continue). On propose ici une méthodologie pour estimer l'entier k et éventuellement f et/ou F, sur la base de l'observation des durées de plusieurs systèmes identiques et indépendants. Cette situation est proche de celle d'échantillons censurés de type II où les durées de vie sont observées jusqu'à la panne du système.

Puisque les composants ne sont pas réparables, on rappelle que la durée de vie T d'un système k-sur-n est alors la k-ième statistique d'ordre du n-échantillon $X_1, \ldots, X_n$. La fonction de répartition et la densité de T peuvent être écrite en fonction de F et de f :

Théorème 6.1.1 *Soit $X_1, \ldots, X_n$ un n-échantillon de fonction de répartition commune F et de densité commune f. Soit $T = X_{(k)}$ la k-ème statistique d'ordre associée à cet échantillon.*

1. *La fonction de répartition de T est égale à :*

$$\forall t \in \mathbb{R}, \quad F_T(t) = \sum_{i=k}^{n} \binom{n}{i} F(t)^i (1 - F(t))^{n-i}. \tag{6.1.1}$$

2. *La densité de T est égale à :*

$$\forall t \in \mathbb{R}, \quad f_T(t) = n \binom{n-1}{k-1} F(t)^{k-1} f(t) (1 - F(t))^{n-k}\ . \tag{6.1.2}$$

Les deux cas particuliers les plus courants sont évidemment celui d'un système en série ($k = 1$) et celui d'un système en parallèle ($k = n$). Dans ces deux cas, les expressions données ci-dessus se simplifient.

Comme l'a montré Mood [134] (voir également [73]), la fonction de répartition F_T de T peut être vue comme une transformation continue de F :

$$\begin{aligned} \forall t \geqslant 0, \quad F_T(t) &= \int_0^s n\binom{n-1}{k-1} F(s)^{k-1} f(s)(1-F(s))^{n-k} ds \\ &= \int_0^{F(t)} B(k, n+1-k) u^{k-1}(1-u)^{n-k} du \\ &:= h_k(F(t)) \end{aligned}$$

où $B(\cdot,\cdot)$ est la fonction beta. Il en résulte que $F_T(t)$ peut être vue comme la fonction de répartition d'une loi beta de paramètre $(k, n+1-k)$ au point $F(t)$. La fonction h_k est strictement croissante et uniformément continue sur l'intervalle unité. On a donc existence et unicité de la fonction réciproque h_k^{-1}.

On détaille ici une méthodologie pour estimer k. Une difficulté et une originalité est qu'il s'agit d'un paramètre à valeur entière. Trois cas sont distingués. Dans le premier cas, on suppose que la loi des durées des composants est complètement connue. Dans le deuxième cas, on suppose que la loi des durées de vie appartient à une famille paramétrique indexée par le paramètre θ inconnue. Il faudra donc estimer k et θ. Enfin, dans le dernier cas, on ne formulera pas d'hypothèses paramétriques sur la loi. En revanche, on il semblerait nécessaire de disposer d'une information supplémentaire, par exemple le temps moyen de fonctionnement (MTTF) des composants.

Dans tous les cas, on supposera que l'on observe les temps de panne de m systèmes indépendants et identiques à celui décrit précédemment. Soit $T_1, \ldots, T_m$ un m-échantillon de même loi que T. On rappelle que les durées de vie des composants ne sont pas observées.

Cas où f est complètement connue Dans ce cas, la vraisemblance associée à l'échantillon ne dépend que d'un paramètre entier :

$$L(k|t_1, \ldots, t_m) = \prod_{j=1}^m f_T(t_j) = \prod_{j=1}^m n\binom{n-1}{k-1} F(t_j)^{k-1} f(t_j)(1-F(t_j))^{n-k}.$$

L'estimateur du maximum de vraisemblance $\widehat{k}$ de k est donc simplement donnée par :

$$\widehat{k} = \text{argmax } L(k|t_1, \ldots, t_m).$$

On rappelle que $k \in \{1, \ldots, n\}$ (avec n connu). La plupart du temps, le nombre n de composants du système est suffisamment petit si bien que la log-vraisemblance peut être évaluée rapidement pour toutes les valeurs possibles de k. Donc l'estimateur du maximum de vraisemblance $\widehat{k}$ peut être déterminée en effectuant une recherche exhaustive. On peut montrer que cet estimateur converge en probabilité vers la vraie valeur.

Cas où f appartient à une famille paramétrique On suppose maintenant que f (et donc F aussi) dépend d'un paramètre euclidien θ. La vraisemblance de l'échantillon est donc désormais une fonction de k et de θ, le premier paramètre étant discret (toujours avec n valeurs possibles) et le second étant continu. Les paramètres peuvent être estimés dans une procédure en deux étapes. Conditionnellement à k, on peut calculer l'estimateur du maximum de vraisemblance $\widehat{\theta}_k$ de θ. Alors, l'estimateur du maximum de vraisemblance $\widehat{k}$ de k est donnée par :

$$\widehat{k} = \text{argmax } L(k, \widehat{\theta}_k|t_1, \ldots, t_m).$$

Au final, l'estimateur du maximum de vraisemblance de (k, θ) est $(\widehat{k}, \widehat{\theta}_{\widehat{k}})$. On peut voir cet estimateur comme un estimateur de vraisemblance profilé.

Cas où f est complètement inconnue Le problème devient alors beaucoup plus complexe. Pour le moment, il est nécessaire de supposer la disposition d'une certaine information sur la loi des durées de vie des composants.

La fonction de répartition empirique $\widehat{F}_T$ de la loi de la durée du système peut être calculée classiquement de la manière suivante :

$$\forall t \geqslant 0\ , \quad \widehat{F}_T(t) = \frac{1}{m}\sum_{j=1}^{m} \mathbb{I}_{T_j \leqslant t}.$$

D'après la relation fonctionnelle entre F et F_T, on en déduit une estimation non-paramétrique de F à partir de $\widehat{F}_T$:

$$\forall t \geqslant 0\ , \quad \widehat{F}(t) = h_k^{-1}(\widehat{F}_T(t)).$$

Cet estimateur est uniformément convergent et asymptotiquement normal (voir [73] pour plus de détails).

On suppose maintenant que l'on connaisse le MTTF des composants (par exemple, il s'agit d'une donnée fournie par le fournisseur). L'idée consiste alors à comparer ce MTTF à celui qui peut être estimée à partir de l'estimateur non-paramétrique $\widehat{F}$ construit ci-dessus (qui dépend de k). En effet, on a :

$$\widehat{MTTF}_k = \int_0^\infty (1 - \widehat{F}(t))\, dt\ .$$

Cette estimation dépend du paramètre k et peut être calculée pour tout $k \in \{1, \ldots, n\}$. On peut donc estimer k en retenant la valeur qui conduit à l'estimation la plus proche (en valeur absolue) du MTTF théorique :

$$\widehat{k} = \text{argmin}_{\ k}\{|\widehat{MTTF}_k - MTTF|, k = 1, \ldots, n\}.$$

6.1.2 Estimation semi-paramétrique pour quelques modèles non-homogènes de dégradation [32, 31]

Les processus de Lévy sont les processus les plus fréquemment employés comme modèles de dégradation. On rappelle que, parmi eux, les trois modèles les plus courants sont le mouvement brownien avec tendance linéaire (ou processus de Wiener), le processus gamma et les processus de Poisson composés. La dégradation moyenne de tous les processus de Lévy est linéaire en temps. Cela constitue une hypothèse restrictive, parfois inacceptable. Par exemple, la propagation d'une fissure est un phénomène non linéaire dans le sens où la longueur moyenne a une évolution plutôt exponentielle que linéaire. Il est donc nécessaire de disposer de modèles pour lesquels la dégradation moyenne n'est plus nécessairement linéaire. Une solution possible consiste à relâcher l'hypothèse de stationnarité des accroissements d'un processus de Lévy (on conserve en revanche l'hypothèse d'indépendance des accroissements) : on parle de processus non-homogènes de Lévy. On suppose que ces processus dépendent de deux paramètres, l'un dans un espace euclidien, l'autre dans un espace fonctionnel (en général, l'ensemble des fonctions réelles, positives et croissantes). Se pose alors naturellement le problème de l'inférence semi-paramétrique du modèle de dégradation sur la base de l'observation de plusieurs individus à plusieurs instants (non nécessairement identiques). Wang [185, 186, 187, 188] a étudié ce problème pour deux modèles non-homogènes, le processus gamma non-homogène et le processus de Wiener avec échelle de temps dilatée par une fonction croissante, avec présence éventuelle de covariables ou d'effets aléatoires. Pour les modèles sans covariables ni effets aléatoires, il a suggéré d'utiliser l'estimateur du maximum de pseudo-vraisemblance (obtenue en ignorant le fait qu'il y a dépendance entre les mesures pour un même individu). On peut alors adopter la même approche pour l'inférence d'autres modèles tels que le mouvement brownien avec tendance non linéaire ou le mouvement brownien géométrique non-homogène. Les premiers résultats obtenus sur ce sujet ont fait l'objet de présentation lors de deux congrès internationaux en 2013 [32, 31]. Outre l'estimation semi-paramétrique par maximum de pseudo-vraisemblance, on a étudié la loi du temps de panne pour ces modèles et son estimation. Il reste encore à établir les propriétés asymptotiques de ces estimateurs. On pourrait envisager d'autres méthodes d'estimation. Enfin, on peut se poser la question de la sélection du meilleur modèle pour un jeu de données.

6.1.3 Problème inverse pour le processus gamma non-homogène basée sur des durées de défaillance

On considère ici le processus gamma non-homogène comme modèle de dégradation. On note (λ_t) sa fonction de forme et μ son paramètre d'échelle. Une forme paramétrique peut être spécifiée pour (λ_t) (par

exemple, $\lambda_t = \alpha t^\beta$ pour tout $t \in \mathbb{R}_+$) ou pas (modèle semi-paramétrique). Comme on l'a vu, le temps de panne associé à un modèle peut être défini comme le premier temps de passage T_c d'un seuil critique c connu. Supposons que la loi du temps de panne est connue, par exemple à partir d'un avis d'expert, éventuellement à des paramètres près (par exemple, on sait que la durée est une variable aléatoire de loi de Weibull). Le problème est alors de déterminer (λ_t) et μ de sorte que la loi de T_c soit la plus proche possible de la loi donnée par les experts. Il s'agit d'un problème déterministe si la loi donnée par les experts est complètement connue.

Notons $F(\cdot;\theta)$, avec $\theta = c((\lambda_t), \mu)$, la fonction de répartition de T_c et $G(\cdot;\eta)$ la fonction de répartition donnée par les experts. Le vecteur de paramètre η est connue ou bien estimée à partir d'observations. On peut alors essayer de déterminer θ tel que F et G soit les plus proches possibles pour une distance donnée. Par exemple, on peut utiliser la distance de Cramer - von Mises :

$$d_{CvM}(\theta;\eta) = \int_{\mathbb{R}} [F(x;\theta) - G(x;\eta)]^2 \mathrm{d}G(x;\eta)$$

ou bien la distance de Kolmogorov-Smirnov :

$$d_{KS}(\theta;\eta) = \max_{x\in\mathbb{R}} |F(x;\theta) - G(x;\eta)|.$$

En supposant qu'il existe une solution, pour tout η, on pose :

$$\theta_{CvM} = \min_\theta d_{CvM}(\theta;\eta) \quad \text{et} \quad \theta_{KS} = \min_\theta d_{KS}(\theta;\eta).$$

Si η est inconnue, on le remplace dans les équations précédentes par une estimation $\widehat{\eta}$. On obtient alors des estimations pour θ_{CvM} et θ_{KS}. On peut alors étudier plusieurs problèmes. Tout d'abord, peut-on trouver des conditions d'existence aux problèmes de minimisation ? On sait que la loi de T_c a une fonction de risque croissante (IFR), cela impliquerait a priori qu'on ne puisse pas considérer n'importe quelle loi G de manière raisonnable. Ensuite, on peut étudier les propriétés des estimations pour θ_{CvM} et θ_{KS} quand η est inconnue. Ces problèmes s'apparentent aux estimateurs de la distance minimale [66, 146].

6.1.4 Temps de panne : premier ou dernier temps de passage ?

Pour un processus non-monotone $(X_t)_{t\geqslant 0}$ modélisant la dégradation d'un composant ou d'un système, l'instant de défaillance est classiquement défini comme le premier temps de passage d'un seuil critique (déterministe ou aléatoire) par le processus. Comme expliqué par Barker et Newby [58], le dernier temps de passage peut aussi être un candidat légitime pour définir l'instant de défaillance. c'est pourquoi, avec L. Rabehasaina, nous avions étudié sa loi dans le cas d'un subordinateur perturbé [14]. Cependant, le choix du dernier temps de passage comme instant de défaillance est difficilement acceptable pour l'industriel. En effet, il en résulte un trop grand risque lié au fait d'être pendant un moment (ou plusieurs) éventuellement long au dessus du seuil critique et il est donc peu souhaitable de laisser le composant/système fonctionner. Le premier temps de passage apparaît donc comme un choix plus prudent. En revanche, pour définir un seuil de maintenance préventive comme dans [61] et [99], on pourrait considérer le dernier temps de passage d'un seuil préventif à définir. Considérons les instants suivants :

$$T_c = \inf\{t; X_t \geqslant c\}$$

définissant l'instant de défaillance du composant/système,

$$T_m = \inf\{t; X_t \geqslant m\}$$

définissant l'instant de maintenance préventive basée sur le premier temps de passage du seuil de maintenance préventive m et

$$\tilde{T}_m = \sup\{t; X_t \leqslant m\}$$

définissant l'instant de maintenance préventive basée sur le dernier temps de passage du seuil de maintenance préventive m. Usuellement, les modèles considèrent le couple (T_m, T_c) alors qu'ici on propose de baser la

politique de maintenance sur le couple $(\tilde{T}_m, T_c)$. Il existe plusieurs difficultés. La première est que $\tilde{T}_m$ n'est pas un temps d'arrêt. La seconde est que $\tilde{T}_m$ n'est pas nécessairement presque sûrement plus petit que T_c (bien que cela n'apparaisse que dans des cas pathologiques peu cohérents avec le phénomène modélisé, les fluctuations devant être "petites"). En considérant $\tilde{T}_m$, on pourrait s'attendre à un niveau optimal de maintenance préventive inférieur à celui qu'on obtiendrait avec T_m. On pourrait donc étudier ce nouveau modèle de maintenance en reproduisant ce qui a été fait dans [61] et/ou [99] et comparer les deux approches.

6.2 Modèles de durées de vie

Cette section regroupe quelques travaux de recherche en cours ou à venir sur les cartes de contrôle pour le suivi de durées de vie.

6.2.1 Cartes de contrôle basées sur la statistique de prédécesseur et des excédants

Dans le travail avec N. Balakrishnan et J.-C. Turlot, nous avons montré que la carte de contrôle unilatérale basée sur la statistique des prédécesseurs est performante (tant du point de vue du taux d'alarme que de la période opérationnelle moyenne) sous la première hypothèse alternative de Lehmann. Pour la seconde hypothèse alternative de Lehmann, la performance est comparable à celle de la carte de contrôle basée sur la statistique de Wilcoxon-Mann-Whitney. Notre objectif est donc d'étudier d'autres statistiques, comme les statistiques des excédents, pouvant aboutir à des cartes plus performantes pour cette dernière alternative. On espère alors combiner l'ensemble de ces statistiques pour construire des cartes de contrôle qui soient performantes pour les deux hypothèses de Lehmann.

6.2.2 Autres cartes de contrôle

D'autres cartes de contrôle pour le suivi de durées de vie mériteraient d'être étudiées. En partiuclier, une carte de contrôle (non-paramétrique) basée sur la statistique de Savage pourrait se montrer performante. En effet, ce test a été élaboré spécifiquement pour la loi exponentielle. Ce test est localement plus puissant pour la loi exponentielle (c'est-à-dire lorsque l'alternative est proche de l'hypothèse nulle). Cela en fait donc une statistique pertinente pour l'élaboration d'une carte de contrôle pour le suivi de durées de vie.

Deuxième partie

Lois discrètes aléatoires et applications en informatique et en écologie

Chapitre 7

Introduction

7.1 Problématiques générales

Considérons l'ensemble des vecteurs aléatoires à valeurs dans Δ_n, le simplexe d'ordre n :

$$\Delta_n = \left\{ (u_1, \ldots, u_n) \in [0,1]^n \; ; \; \sum_{i=1}^{n} u_i = 1 \right\}.$$

La loi d'un tel vecteur aléatoire est appelée loi discrète aléatoire [116] et intervient dans de nombreux domaines : génétique, écologie, informatique théorique, statistique bayésienne, etc. [155].

Une construction possible des lois discrètes aléatoires $(p_1, \ldots, p_n)$ consiste à renormaliser un ensemble de n variables aléatoires positives et indépendantes (pas nécessairement de même loi) par leur somme. Autrement dit, pour toutes variables aléatoires $X_1, \ldots, X_n$ positives et indépendantes, on pose :

$$\forall i \in \{1, \ldots, n\}, p_i = \frac{X_i}{X_1 + \ldots + X_n}.$$

Kingman [116] avait proposé une construction différente, reposant sur l'utilisation d'un subordinateur défini sur $[0,1]$. On rappelle qu'un subordinateur $(\xi(t))_{t\in[0,1]}$ est un processus de Lévy, croissant et positif. Il s'agit donc d'un processus de sauts, à accroissements indépendants et stationnaires. On construit alors une variable aléatoire $(p_1, \ldots, p_n)$ sur le simplexe Δ_n en posant :

$$\forall j \in \{1, \ldots, n\}, \quad p_j = \frac{\xi(jn^{-1}) - \xi((j-1)n^{-1})}{\xi(1)}.$$

On a ainsi un vecteur aléatoire dont les lois marginales sont identiques. Deux exemples fondamentaux avaient été traités par Kingman [116] dans le cadre de l'algorithme move-to-front correspondant aux cas d'un subordinateur gamma et d'un subordinateur γ-stable. Dans le premier cas, la loi conjointe de $(p_1, \ldots, p_n)$ est connue explicitement et porte le nom de loi de Dirichlet [1].

De manière plus générale, Regazzini *et al.* [159] ont introduit les mesures aléatoires normalisées à accroissements indépendants dont la construction est la suivante. Soit $0 = t_0 < t_1 < \cdots < t_n = 1$ une partition de l'intervalle $[0,1]$. Alors, pour une partition $t_0, \ldots, t_n$ et un subordinateur ξ donnés, on pose :

$$\forall j \in \{1, \ldots, n\}, \quad p_j = \frac{\xi(t_j) - \xi(t_{j-1})}{\xi(1)}.$$

Les lois marginales de $(p_1, \ldots, p_n)$ ne sont donc plus nécessairement identiques.

1. Le nom de cette loi est, semble-t-il, dû à Wilks [190].

7.2 Contexte

Lors de mon séjour post-doctoral au Chili (CMM, Université du Chili, janvier à août 2003), j'ai travaillé, avec Javiera Barrera, sur l'analyse de l'heuristique auto-organisatrice move-to-front. Les paramètres de cet algorithme sont les probabilités de requête. Nous nous étions intéressés au cas où ces probabilités sont supposées être aléatoires, ce qui nous a amené à travailler sur les lois discrètes aléatoires. Ce séjour post-doctoral a été le point de départ d'une collaboration avec Javiera Barrera [2], qui s'est enrichie d'une collaboration avec Thierry Huillet. Ces questions-là constituaient ma principale activité de recherche jusqu'en 2007 (i.e. peu de temps après mon arrivée à Pau). Puis, j'ai de nouveau travaillé sur ce sujet en 2009 et 2010, suite à une discussion avec Fabrizio Leisen à Jaca.

Dans le chapitre suivant, on présente les différents résultats obtenus dans le cadre de l'utilisation des lois discrètes aléatoires pour deux algorithmes : move-to-front et move-to-root. Dans le second chapitre de cette partie, on présente quelques résultats sur des problèmes d'inférence statistique sur la loi de Dirichlet. Le tableau ci-dessous résume quantitativement mon activité de recherche sur ce thème.

Publications	[4, 2, 11, 3, 12, 13]
Acte de congrès	[19]

2. Javiera était ensuite venue trois ans en France dans le cadre de sa thèse en co-tutelle.

Chapitre 8

Utilisation en analyse d'algorithmes

8.1 Move-to-front [4, 2, 3]

8.1.1 Description de l'heuristique

Soit n objets stockés dans une liste (livres dans une bibliothèque, fichiers sur un disque dur, ...). La liste est une structure de données importante en informatique. On accède aux éléments d'une liste de manière séquentielle : de la tête de la liste à la queue. Par exemple, pour accèder au troisième élément de la liste, on est obligé de commencer par accéder au premier puis au deuxième élément. Aussi on souhaiterait arranger la liste de sorte à avoir les objets les plus populaires en tête de liste. Mais les popularités ou probabilités de requête de chacun des objets sont inconnues. Une solution a été proposée de manière indépendante par le russe Tsetlin en 1963 [183] et l'américain McCabe en 1965 [131] : l'heuristique move-to-front. A chaque unité de temps discret, un objet est demandé par un utilisateur et cet objet est mis en tête de liste : les objets avant celui demandé sont décalés d'une position vers la queue, les positions des autres objets étant inchangées.

On note $p_1, \ldots, p_n$ les probabilités de requête des n objets. L'heuristique proposée peut être vue comme une chaîne de Markov σ_n sur l'ensemble $\mathcal{S}_n$ des permutations à n éléments. Cette chaîne est ergodique de mesure stationnaire :

$$\forall s \in \mathcal{S}_n, \quad \mathbb{P}[\sigma_n = s] = p_{s_1} \prod_{i=2}^{n} \frac{p_{s_i}}{1 - p_{s_1} - \cdots - p_{s_{i-1}}},$$

appelée permutation biaisée par la taille [103, 85].

Une quantité intéressante est alors le coût de recherche $S_n(t)$ définie comme la position de l'objet demandé dans la liste au temps t. En effet, il s'agit du nombre de comparaisons nécessaires pour la requête à l'instant t. Par convention, la première position (tête de liste) est 0 : $S_n(t)$ est donc une variable aléatoire à valeurs dans $\{0, \ldots, n-1\}$. On peut la voir comme le mélange de n variables aléatoires :

$$S_n(t) = P_{n,t}(R),$$

où R est une v.a. discrète de loi $(p_1, \ldots, p_n)$ et

$$P_{n,t}(i) = \sum_{j \neq i} I_{ji}(t)$$

avec $I_{ji}(t)$ égale à 1 si j précède i dans la liste à l'instant t et 0 sinon. On a [91] :

$$\mathbb{P}[S_n(t) = k] = \sum_{i=1}^{n} p_i \mathbb{P}[P_{n,t}(i) = k].$$

Dans la suite, on ne s'intéressera qu'au coût de recherche à l'équilibre que l'on notera S_n.

8.1.2 Résultats de Kingman

Le coût moyen de recherche asymptotique est bien connu [131] :

$$\mu = \mathbb{E}[S_n] = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \frac{p_i p_j}{p_i + p_j}.$$

On dispose d'un majorant :

$$\mu \leqslant \frac{n-1}{2}.$$

De plus, si les popularités sont connues, alors la permutation optimale est $p_{(1)} \geqslant \cdots \geqslant p_{(n)}$ et le coût de recherche associé est égal à :

$$m = \sum_{i=1}^{n} (i-1) p_{(i)}.$$

On peut aussi montrer l'encadrement suivant :

$$m \leqslant \mu \leqslant 2m.$$

Supposons maintenant que les popularités sont un vecteur aléatoire de loi de Dirichlet $\mathcal{D}_n(\alpha)$ de paramètre α. La densité jointe de $(p_1, \ldots, p_n)$ est donc :

$$f(x_1, \ldots, x_n) = \frac{\Gamma(n\alpha)}{\Gamma(\alpha)^n} \prod_{i=1}^{n} x_i^{\alpha-1} \mathbb{I}_{(x_1,\ldots,x_n) \in \Delta_n}.$$

Kingman [116] a montré que l'espérance du coût moyen de recherche est égale à :

$$\mathbb{E}[\mu] = \frac{(n-1)\alpha}{1+2\alpha}$$

et que :

$$\mathbb{E}[m] = (n-1)\left(\frac{1}{2} - \frac{\Gamma(1+2\alpha)}{\Gamma(1+\alpha)^2 2^{1+2\alpha}}\right).$$

On peut s'intéresser à un régime asymptotique particulier, la limite Poisson-Dirichlet : $n \longrightarrow \infty$, $\alpha \longrightarrow 0$ et $n\alpha \longrightarrow \lambda$. On obtient alors :

$$\mathbb{E}[\mu] \longrightarrow \lambda \quad \text{et} \quad \mathbb{E}[m] \longrightarrow \lambda \log 2.$$

Plus généralement, pour toute fonction intégrable ϕ, on pose $s_\phi = \sum_{j=1}^{n} \phi(p_j)$. Alors, on a :

$$\mathbb{E}[s_\phi] = n\mathbb{E}[\phi(p_1)] \longrightarrow \lambda \int_0^1 \phi(x) x^{-1} (1-x)^{\lambda-1} dx$$

Par exemple, dans le cas de l'entropie h de $(p_1, \ldots, p_n)$, $\phi(x) = -x \log_2 x$ et :

$$\mathbb{E}[h(p_1, \ldots, p_n)] = \frac{1}{\log 2} \sum_{n=1}^{\infty} \frac{\lambda}{n(n+\lambda)}.$$

En fait, le cas de popularité Dirichlet apparaît comme un cas particulier de popularité construite à partir d'un subordinateur (en prenant le subordinateur gamma). On rappelle qu'un subordinateur ξ est un processus de Lévy croissant (donc à sauts) qui vérifie la formule de Lévy :

$$\mathbb{E}[e^{-\theta \xi(t)}] = e^{-t\psi(\theta)}$$

avec $\psi(\theta) = \int_0^\infty (1 - e^{-\theta x}) \mu(dx)$ sa mesure de Lévy. Dans le cas d'un subordinateur gamma, on a $\psi(\theta) = \log(1+\theta)$ et $\mu(dx) = x^{-1} e^{-x} dx$. Comme expliqué en introduction, on peut alors construire le vecteur des popularités $(p_1, \ldots, p_n)$ de la manière suivante :

$$\forall j \in \{1, \ldots, n\}, \quad p_j = \frac{\xi(jn^{-1}) - \xi((j-1)n^{-1})}{\xi(1)}.$$

Kingman [116] a alors donné des expressions générales pour $\mathbb{E}[\mu]$, $\mathbb{E}[m]$ et $\mathbb{E}[s_\phi]$. Il a également étudié un autre exemple, celui du subordinateur γ-stable :

$$\psi(\theta) = a\gamma^{-1}\Gamma(1-\gamma)\theta^\gamma \quad \text{et} \quad \mu(dx) = ax^{-\gamma-1}dx.$$

Il a montré [116] le comportement asymptotique de $\mathbb{E}[\mu]$:

$$\lim_{n\to\infty} \mathbb{E}[\mu] = \begin{cases} \frac{\gamma}{1-2\gamma} & \text{si } \gamma < \frac{1}{2} \\ \infty & \text{sinon} \end{cases}$$

et de $\mathbb{E}[m]$:

$$\lim_{n\to\infty} \mathbb{E}[m] = \begin{cases} \frac{\Gamma(1-2\gamma)}{\Gamma(1-\gamma)2^{1-2\gamma}} & \text{si } \gamma < \frac{1}{2} \\ \infty & \text{sinon} \end{cases}.$$

8.1.3 Généralisations

Le théorème ci-dessous donne la transformée de Laplace de S_n :

Théorème 8.1.1 *[91] Pour tout vecteur déterministe $(p_1, \ldots, p_n)$ de popularités,*

$$\mathbb{E}[e^{-uS_n}] = \int_0^\infty e^{-s} \sum_{i=1}^n p_i^2 \prod_{\substack{k=1 \\ k\neq i}}^n (1 + e^{-u}(e^{sp_k} - 1))ds.$$

On peut déduire de ce résultat la transformée de Laplace dans le cas de popularités aléatoires construites construites à partir de variables aléatoires indépendantes $X_1, \ldots, X_n$ normalisées par leur somme. Pour tout $i \in \{1, \ldots, n\}$, on notera f_i la densité de X_i, ϕ_i sa transformée de Laplace et μ_i son espérance (finie ou pas).

Théorème 8.1.2 *[4] Pour toute suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes,*

$$\forall s \geqslant 0, \quad \phi_{S_n}(s) = \sum_{i=1}^n \int_0^\infty \left(\int_t^\infty \phi_i''(r) \prod_{\substack{k=1 \\ k\neq i}}^n \psi_{t,s,k}(r)\, dr \right) dt,$$

où, pour tout $i \in \{1, \ldots, n\}$,

$$\forall t \geqslant 0, \forall r \geqslant t, \quad \psi_{t,s,i}(r) = \phi_i(r) + e^{-s}(\phi_i(r-t) - \phi_i(r))$$

On en déduit facilement les deux premiers moments dans le cas de v.a. indépendantes et de même loi :

Corollaire 8.1.1 *[4] Pour toute suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes et de même loi,*

$$\mathbb{E}[S_n] = n(n-1) \int_0^\infty \phi(r)^{n-2} \left(\int_r^\infty (\phi'(t))^2\, dt \right) dr$$

et

$$\begin{aligned} \mathbb{E}[S_n^2] = {} & n(n-1)(2n-3) \int_0^\infty \phi(r)^{n-2} \left(\int_r^\infty (\phi'(t))^2\, dt \right) dr \\ & -2n(n-1)(n-2) \int_0^\infty \phi(r)^{n-3} \left(\int_r^\infty \phi(t)(\phi'(t))^2\, dt \right) dr. \end{aligned}$$

On peut aussi en déduire un développement asymptotique de la transformée de Laplace de S_n dans le cas de v.a. indépendantes et de même loi :

Théorème 8.1.3 *[4] Pour toute suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes et de même loi,*

$$\forall s \geqslant 0, \quad \phi_{S_n}(s) \sim -\int_0^\infty \frac{\phi''(r)\psi_{r,s}(r)^n}{\psi_{r,s}'(r)}\, dr$$

quand n tend vers l'infini.

Toujours dans le cas de v.a. indépendantes et de même loi, on en déduit la limite pour les deux premiers moments sous une condition de finitude de μ :

Corollaire 8.1.2 *[4] Pour toute suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes et de même loi d'espérance μ finie,*

$$\frac{1}{n}\mathbb{E}[S_n] \xrightarrow[n\to\infty]{} \frac{1}{\mu}\int_0^\infty (\phi'(r))^2\, dr$$

et

$$\frac{1}{n^2}\mathbb{E}[S_n^2] \xrightarrow[n\to\infty]{} \frac{2}{\mu}\int_0^\infty (1-\phi(r))(\phi'(r))^2\, dr.$$

Plus précisément, on peut calculer la loi limite sous les mêmes hypothèses.

Théorème 8.1.4 *[3] Pour toute suite $(X_n)_{n\in\mathbb{N}^*}$ de v.a. indépendantes et de même loi d'espérance μ finie,*

$$\frac{S_n}{n} \xrightarrow[n\to\infty]{d} S,$$

où S est une v.a. aléatoire de densité f_S :

$$f_S(x) = -\frac{1}{\mu}\frac{\phi''\left(\phi^{-1}(1-x)\right)}{\phi'\left(\phi^{-1}(1-x)\right)}\mathbb{I}_{[0,1-p_0]}(x)$$

avec $p_0 = \mathbb{P}(X_1 = 0)$ et ϕ^{-1} la réciproque ϕ.

Il est assez surprenant de constater que la loi limite peut être calculée explicitement pour toutes les lois classiques.

Exemple 8.1.1 On suppose que $X_1, \ldots, X_n$ sont de loi de Dirac en 1. On obtient, à la limite, la loi uniforme sur $[0, 1]$. Ce résultat avait déjà été montré par Fill [90].

Exemple 8.1.2 On suppose que $X_1, \ldots, X_n$ sont de loi géométrique sur $\mathbb{N}$ de paramètre p. On obtient, à la limite, la densité suivante :

$$f_S(y) = \frac{p}{1-p}\left[\frac{2(1-y)}{p} - 1\right]\mathbb{I}_{[0,1-p]}(y).$$

Exemple 8.1.3 On suppose que $X_1, \ldots, X_n$ sont de loi de Poisson de paramètre λ. On obtient, à la limite, la densité suivante :

$$f_S(y) = \frac{\ln(1-y)+\lambda+1}{\lambda}\mathbb{I}_{[0,1-e^{-\lambda}]}(y).$$

Exemple 8.1.4 On suppose que $X_1, \ldots, X_n$ sont de loi de Pareto de paramètre $1/\alpha$. Dans ce cas, on obtient, à la limite, la loi de la v.a. $A(\alpha)$ obtenue et décrite par Fill [90] quand les popularités sont connues et de loi de Zipf généralisée.

Exemple 8.1.5 On suppose que $X_1, \ldots, X_n$ sont de loi béta de paramètre $(1/s, 1)$. Dans ce cas, on obtient, à la limite, la loi de la v.a. $B(s)$ obtenue et décrite par Fill [90] quand les popularités sont connues et de loi puissance généralisée.

Exemple 8.1.6 En dernier exemple, on considère le cas où $X_1, \ldots, X_n$ sont de loi gamma de paramètre α, autrement dit le cas de popularité Dirichlet. Les deux premiers moments sont égaux à :

$$\mathbb{E}[S_n] = \frac{\alpha(n-1)}{2\alpha+1} \quad \text{et} \quad \mathbb{E}[S_n^2] = \frac{\alpha(2\alpha n+1)(n-1)}{(2\alpha+1)(3\alpha+1)}.$$

On retrouve le résultat de Kingman [116] pour l'espérance. De plus, la loi limite est la loi béta de paramètre $(1, 1/(1+\alpha))$.

On peut aussi regarder le cas du régime Poisson-Dirichlet. En utilisant les propriétés de loi de Dirichlet, on a pu montrer directement que la loi limite de S_n sous ce régime est la loi géométrique sur $\mathbb{N}$ de paramètre γ [2] (attention : il n'y a pas de renormalisation à effectuer dans ce cas).

Pour finir, on considère le cas où les popularités $p_1, \dots, p_n$ sont construites à partir d'une mesure aléatoire normalisée à accroissements indépendants. Lijoi et Prünster [126] ont obtenu une généralisation du résultat de Kingman :

Proposition 8.1.1 *[126] Pour toute mesure aléatoire normalisée à accroissements indépendants* $(p_1, \dots, p_n)$,

$$\mathbb{E}[S_n] = \sum_{\substack{i,j=1 \\ i \neq j}} \frac{\bar{\alpha}_i \bar{\alpha}_j}{\bar{\alpha}_i + \bar{\alpha}_j}[1 + \mathcal{I}(\alpha_i, \alpha_j)],$$

où $\bar{\alpha}_i = \alpha_i / \sum_{i=1}^n \alpha_i$ *pour tout* $i \in \{1, \dots, n\}$ *et*

$$\mathcal{I}(\alpha_i, \alpha_j) = a \int_{(0,\infty)} e^{-a\phi(u)} \int_{(0,\infty)} \phi^{(2)}(u+v) e^{-(\alpha_i+\alpha_j)(\phi(u+v)-\phi(u))}\, dv du.$$

Le résultat de Kingman [116] apparaît donc comme un coromlaire du résultat ci-dessus :

Corollaire 8.1.3 *[126] Dans le cas particulier d'un subordinateur* γ*-stable,*

$$\mathbb{E}[S_n] = \sum_{\substack{i,j=1 \\ i \neq j}} \frac{\bar{\alpha}_i \bar{\alpha}_j}{\bar{\alpha}_i + \bar{\alpha}_j}[1 - (1-\gamma)_2 F_1(1, 1; 1 + \frac{1}{\gamma}; 1 - \bar{\alpha}_i - \bar{\alpha}_j)],$$

où ${}_2F_1$ *est la fonction hypergéométrique. De plus, si pour tout* $i \in \{1, \dots, n\}$, $\alpha_i = \frac{1}{n}$, *alors*

$$\lim_{n \to +\infty} \mathbb{E}[S_n] = \begin{cases} \gamma/(1-2\gamma) & si \quad \gamma < 1/2 \\ \infty & si \quad \gamma \geqslant 1/2. \end{cases}$$

En fait, la proposition de Lijoi et Prünster [126] peut aussi être vue comme un corollaire de nos résultats. On peut donc alors les utiliser pour montrer un résultat analogue à celui de Kingman pour les autres moments. Ainsi, Fabrizio Leisen, Antonio Lijoi et moi nous avons obtenu le résultat suivant :

Théorème 8.1.5 *[13] Si* ξ *est un subordinateur* γ*-stable et si* $\bar{\alpha}_i = 1/n$ *pour tout* $i \in \{1, \dots, n\}$, *alors*

$$\lim_{n \to \infty} \mathbb{E}(S_n^k) = \begin{cases} \sum_{l=1}^k \frac{(l!)^2}{(\frac{1}{\gamma} - l - 1)_l} a_l^{(k)} & si\ \gamma < \frac{1}{k+1} \\ \infty & sinon \end{cases}$$

avec

$$\begin{cases} a_1^{(k)} & = 1 \\ a_l^{(k)} & = a_{l-1}^{(k-1)} + l a_l^{(k-1)} \qquad l = 2, \dots, k-1 \\ a_k^{(k)} & = 1 \end{cases} .$$

8.2 Move-to-root [19]

Les arbres constituent une autre structure de données utilisée en informatique. Ici nous ne nous intéresserons qu'à une famille particulière d'arbres : les arbres binaires de recherche. Un arbre binaire de recherche est un arbre enraciné, orienté, étiqueté et dont les nœuds ont au plus deux descendants. L'étiquette (ou la clef) d'un nœud est plus grande que toutes celles de son sous-arbre gauche et plus petite que toutes celles de son sous-arbre droit. Cette structure de données permet d'effectuer une recherche efficace d'un élément.

On considère que n objets sont insérés dans un arbre binaire de recherche. Le coût de recherche d'un objet est donc égal au nombre d'ancêtres de celui-ci (ou encore la profondeur de l'objet dans l'arbre moins un). On suppose qu'à chaque unité de temps un objet de l'arbre est demandé par un utilisateur. Lorsque les popularités des objets sont connues, Knuth ([118], p.433-477) a proposé une construction optimale de l'arbre binaire de recherche (arbre de coût minimal).

Lorsque les popularités ne sont pas connues, afin d'espérer obtenir un arbre binaire de recherche optimal, Allen et Munro [49] ont proposé l'heuristique "move-to-root" sur la base d'échanges simples faisant remonter progressivement l'objet demandé à la racine. En effet, on ne peut pas remonter directement l'objet à la

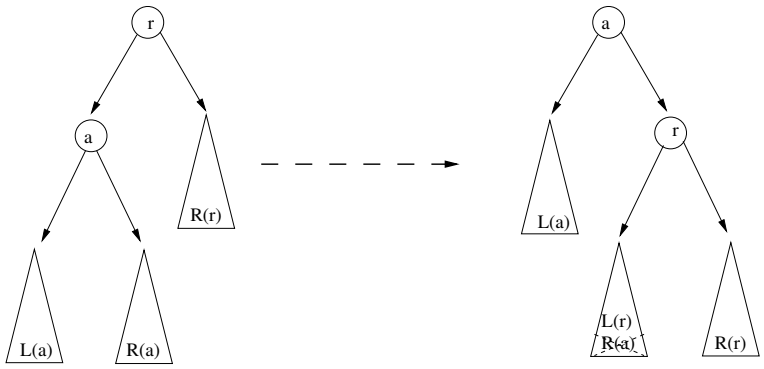


Figure 8.1 – Exemple d'un échange simple

racine sans perdre la propriété d'arbre binaire de recherche. On doit donc procéder pas à pas. Un échange simple consiste à déplacer un objet à la place de son ancêtre de la manière suivante :

Algorithme 8.2.1 Soit a l'élément demandé.

1. Si a est la racine, on ne fait rien.
2. Si a est le descendant gauche de l'objet r, on modifie le sous-arbre dont la racine est r de la manière suivante :
   — on bouge a à la place de r de sorte que a devienne la racine du nouveau sous-arbre ;
   — l'ancien sous-arbre gauche de a est inchangé ;
   — l'ancien sous-arbre droit de a devient le sous-arbre gauche de r ;
   — l'ancien sous-arbre droit de r est inchangé.
3. Si a est le descendant droit de l'objet r, on modifie le sous-arbre dont la racine est r de manière analogue au cas précédent.

La figure 8.1 donne un exemple d'échange simple. Cette manière de mettre à jour un arbre binaire de recherche est analogue à l'heuristique "move-to-front" pour une liste (car cette dernière heuristique peut être vue comme une suite de transposition jusqu'à ce que l'objet demandé arrive en tête de liste).

Il est clair que l'heuristique "move-to-root" induit une chaîne de Markov sur l'ensemble des arbres binaires de recherche à n éléments (on rappelle que le nombre d'arbres binaires de recherche à n éléments est égal au n-ième nombre de Catalan). Cette chaîne de Markov a été étudiée par Dobrow et Fill [82, 83]. Ils ont montré en particulier l'existence d'un lien entre les chaînes de Markov induites par les deux heuristiques. Il en résulte l'existence d'une unique mesure stationnaire.

Comme pour l'heuristique précédente, on considère que les popularités sont aléatoires. On note $S_n^{\mathcal{T}}$ (resp. $S_n^{\mathcal{L}}$) le coût de recherche lorsque la chaîne induite par l'heuristique "move-to-root" (resp. "move-to-front") est à l'état stationnaire.

Théorème 8.2.1 *[19] Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a. indépendantes et de même loi.*

1. *Le moment d'ordre 1 de $S_n^{\mathcal{T}}$ est égal à :*

$$\mathbb{E}[S_n^{\mathcal{T}}] = 2\sum_{i=1}^{n-1}\int_0^\infty\int_t^\infty (n-i)\phi'(u)^2\phi(u)^{i-1}\phi(t)^{n-i-1}\,dudt. \tag{8.2.1}$$

2. *Le moment d'ordre 2 de $S_n^{\mathcal{T}}$ est égal à :*

$$\begin{aligned}\mathbb{E}[(S_n^{\mathcal{T}})^2] &= \mathbb{E}[S_n^{\mathcal{T}}] \\ &\quad - 8\sum_{i=1}^{n-2}\sum_{j=1}^{n-i-1}(n-i-j)\int_0^\infty\int_t^\infty\int_u^\infty \phi'(v)^2\phi'(u)\phi(v)^{i-1}\phi(u)^{j-1}\phi(t)^{n-i-j-1}\,dvdudt.\end{aligned} \tag{8.2.2}$$

Remarque 8.2.1

1. On a la majoration suivante :
$$\mathbb{E}[S_n^{\mathcal{T}}] \leqslant 2\log n.$$
L'ordre de grandeur de cette borne n'est pas une surprise car la hauteur d'un arbre binaire de recherche à n éléments est de l'ordre de $\log n$.
2. On rappelle que la transformée de Laplace d'une v.a. positive est une fonction décroissante. On en déduit alors l'inégalité suivante :
$$\mathbb{E}[S_n^{\mathcal{T}}] \leqslant 2\sum_{1\leq i<j\leq n}\int_0^\infty \left(\int_t^\infty \phi_i'(r)\phi_j'(r)\,dr\right)\prod_{\substack{k=1\\k\neq i,j}}^{n}\phi_k(t)\,dt = \mathbb{E}[S_n^{\mathcal{L}}].$$
Le coût de recherche dans un arbre binaire de recherche mis à jour par "move-to-root" est donc toujours inférieur en moyenne à celui dans une liste mise à jour par "move-to-front".
3. Le cas de v.a. indépendantes mais pas de même loi est plus fastidieux à traiter. Pour l'espérance, on obtient :
$$\mathbb{E}[S_n^{\mathcal{T}}] = 2\sum_{1\leq i<j\leq n}\int_0^\infty\int_t^\infty \phi_i'(u)\phi_j'(u)\prod_{k=1}^{i-1}\phi_k(t)\prod_{k=i+1}^{j-1}\phi_k(u)\prod_{k=j+1}^{n}\phi_k(t)\,dudt.$$

Pour conclure, nous allons présenter deux exemples :

Exemple 8.2.1 On considère où les v.a. $X_1,\dots,X_n$ sont de loi de Dirac respectivement au point $a_1,\dots,a_n$ (popularités déterministes). Dans ce cas, on a :
$$\mathbb{E}[S_n^{\mathcal{T}}] = 2\sum_{1\leq i<j\leq n}\frac{a_i a_j}{A_n(a_i+\cdots+a_j)},$$
où $A_n = a_1+\cdots+a_n$. Dans le cas de popularités déterministes identiques (sans perte de généralité, on pose $a_i = 1$ pour tout $i\{1,\dots,n\}$), on obtient :
$$\mathbb{E}[S_n^{\mathcal{T}}] = 2\left(1+\frac{1}{n}\right)H_n - 4$$
où H_n est le nombre harmonique d'ordre n. Asymptotiquement, on obtient que $\mathbb{E}[(S_n^{\mathcal{T}})^2] \sim 2\log n$. On retrouve le résultat déjà montré par Allen et Munro [49]. Pour le moment d'ordre 2, on a :
$$\mathbb{E}[(S_n^{\mathcal{T}})^2] = 8\left(1+\frac{1}{n}\right)H_nH_{n-1} - 4\left(1+\frac{2}{n}\right)\left(H_{n-1}^2+H_{n-1}^{(2)}\right) - \left(14-\frac{10}{n}\right)H_n + 20 - \frac{16}{n}.$$
On obtient alors $\mathbb{E}[S_n^{\mathcal{T}}] \sim 4\log n^2$ quand n tend vers l'infini. En utilisant l'inégalité de Bienaymé-Tchebychev, il en résulte que $S_n^{\mathcal{T}}/\mathbb{E}[S_n^{\mathcal{T}}]$ converge vers 1 en probabilité. Autrement dit, la loi de $S_n^{\mathcal{T}}$ est de plus en plus concentrée autour de son espérance avec n croissant.

Exemple 8.2.2 On considère le cas de popularités $(p_1,\dots,p_n)$ de loi de Dirichlet non symétrique $\mathcal{D}(a_1,\dots,a_n)$, i.e. le cas où, pour tout $i\in\{1,\dots,n\}$, X_i est de loi gamma de paramètre $(a_i,1)$. Le premier moment de $S_n^{\mathcal{T}}$ est égal à :
$$\mathbb{E}[S_n^{\mathcal{T}}] = 2\sum_{1\leq i<j\leq n}\frac{a_i a_j}{A_n(a_i+\cdots+a_j+1)},$$
où $A_n = a_1+\cdots+a_n$. Dans le cas symétrique, i.e. dans le cas où pour tout $i\in\{1,\dots,n\}$, $a_i = a$, cette expression se simplifie ainsi :
$$\mathbb{E}[S_n^{\mathcal{T}}] = \frac{2a}{n}\sum_{i=1}^{n-1}\frac{n-i}{(i+1)a+1} = 2\left(a+\frac{a+1}{n}\right)\sum_{i=1}^{n-1}\frac{1}{(i+1)a+1} - \frac{2(n-1)}{n}.$$

On retrouve le même ordre de grandeur que dans l'exemple précédent, i.e. qu'asymptotiquement $\mathbb{E}[S_n^{\mathcal{T}}] \sim 2\log n$. Pour le moment d'ordre 2, on a :

$$\mathbb{E}[(S_n^{\mathcal{T}})^2] = \mathbb{E}[S_n^{\mathcal{T}}] + \frac{8a^2}{n}\sum_{i=1}^{n-2}\sum_{j=1}^{n-i-1}\frac{n-i-j}{(a(i+j+1)+1)(a(i+1)+1)}.$$

Si $a = 1$, les deux expressions précédentes se simplifient et on montre le même phénomène asymptotique de concentration de $S_n^{\mathcal{T}}$ autour de son espérance.

Chapitre 9

Problèmes d'inférence autour de la loi de Dirichlet

La loi de Dirichlet apparaît dans de nombreux domaines des probabilités appliquées et de la statistique, par exemple en écologie, génétique, analyse d'algorithme (cf. ci-dessus), statistique bayésienne, ... (voir [155] et ses références). On note par $\mathbf{S}_n = (S_1, \ldots, S_n)$ un vecteur aléatoire de loi de Dirichlet symétrique $D_n(\theta)$ de paramètre θ.

Avec Thierry Huillet, nous avons travaillé sur le problème de l'estimation pour la loi de Dirichlet symétrique. Deux cas ont été successivement considérés : le cas où seul le paramètre θ est inconnu, puis le cas où à la fois n et θ sont inconnus.

9.1 Estimation du paramètre θ [11]

Le problème de l'inférence statistique pour la loi de Dirichlet a été étudié par Ronning [165] dans le cas d'un m-échantillon de loi de Dirichlet non-symétrique (dans ce cas, il y a donc n paramètres à estimer). Narayanan a développé une procédure pour de petits échantillons [137]. Dans [11], nous avons regardé le problème de l'estimation du paramètre θ dans le cas de la loi de Dirichlet symétrique sur la base de l'observation d'une seule réalisation (ce qui est possible puisque nous avons alors un paramètre unidimensionnel à estimer). Nous avons proposé cinq estimateurs différents. Nous ne reprendrons ici que deux de ces estimateurs. Ils reposent sur la notion de permutations biaisées par la taille (PBT, en abrégé) du vecteur $\mathbf{S}_n$. On peut voir ces estimateurs comme obtenus par une méthode de bootstrap.

Dans une première étape d'un échantillonnage biaisé par la taille, on choisit au hasard un fragment. Soit U_1 une variable aléatoire uniforme sur $[0, 1]$. On note par M_1 l'intervalle auquel U_1 appartient et par L_1 la longueur de cet intervalle. La partition aléatoire $\mathbf{S}_n$ conduit à une nouvelle partition :

$$(L_1, S_1, \ldots, S_{M_1-1}, S_{M_1+1}, \ldots, S_n) = (L_1, (1 - L_1)\mathbf{S}_n^{(1)})$$

où

$$\mathbf{S}_n^{(1)} = (S_1^{(1)}, \ldots, S_{M_1-1}^{(1)}, S_{M_1+1}^{(1)}, \ldots, S_n^{(1)}),$$

est une partition aléatoire de l'intervalle unité en $n - 1$ fragments aléatoires[1]. On montre que $\mathbf{S}_n^{(1)}$ est indépendante de L_1 et suit la loi de Dirichlet $D_{n-1}(\theta)$. On a donc une propriété d'invariance et il s'agit d'une des propriétés fondamentales de cette famille de lois de probabilité. On peut itérer ce processus sur $\mathbf{S}_n^{(1)}$ (pour effectuer des tirages sans remise des fragments) et sur les suivants. A la fin, on obtient ce qu'on appelle une permutation biaisée par la taille, qui sera notée $\mathbf{L}_n$. Pour une référence générale sur les permutations biaisées par la taille, on pourra consulter le chapitre 9 de [117]. La permutation biaisée par la taille $\mathbf{L}_n$ peut se factoriser de la manière suivante (on parle alors de modèle d'allocation résiduelle) :

1. On pourra remarquer le lien entre ces étapes de l'échantillonnage et l'heuristique "move-to-front" étudiée précédemment.

Proposition 9.1.1 *La permutation biaisée par la taille* $\mathbf{L}_n$ *du vecteur aléatoire* $\mathbf{S}_n$ *de loi* $D_n(\theta)$ *vérifie la propriété suivante :* $V_1 \overset{(d)}{=} L_1$ *et*

$$L_k \overset{(d)}{=} V_k \prod_{i=1}^{k-1}(1-V_i),$$

où, pour tout $i \in \{1,\dots,n-1\}$, V_i *est une variable aléatoire de loi béta de paramètres* $(1+\theta,(n-i)\theta)$.

En utilisant ces propriétés-là, par simulation on génère un m-échantillon $\mathbf{L}_n^{(1)},\dots,fL_n^{(m)}$ de permutations biaisées par la taille de $\mathbf{S}_n$. Autrement dit, pour tout $k \in \{1,\dots,n-1\}$, on dispose d'un m-échantillon $V_k^{(1)},\dots,V_k^{(m)}$ de variables aléatoires i.i.d. de loi béta de paramètres $(1+\theta,(n-k)\theta)$. Nous avons donc proposé deux méthodes pour estimer le paramètre θ, la première par la méthode du maximum de vraisemblance et la seconde par la méthode des moments.

Méthode du maximum de vraisemblance

Pour tout $k \in \{1,\dots,n-1\}$, on note $\hat{\theta}_{k,n,m}^{(1)}$ l'estimateur du maximum de vraisemblance qui est la solution de l'équation suivante :

$$\Psi(\theta+1)+(n-k)\Psi((n-k)\theta)-(n-k+1)\Psi((n-k+1)\theta+1)=\frac{1}{m}\sum_{i=1}^{m}(\log x_i+(n-k)\log(1-x_i))$$

où Ψ est la fonction digamma ou dérivée logarithmique de la fonction gamma. Cet estimateur est asymptotiquement normal :

$$\sqrt{m}(\hat{\theta}_{k,n,m}^{(1)}-\theta) \xrightarrow[m\to\infty]{d} \mathcal{N}(0,\sigma_{ML,k}^2)$$

avec

$$\sigma_{ML,k}^2=\Psi_1(\theta+1)+(n-k)^2\Psi_1((n-k)\theta)-(n-k+1)^2\Psi_1((n-k+1)\theta+1)$$

où Ψ_1 est la fonction trigamma (i.e. la dérivée première de la fonction Ψ). On peut alors regarder la moyenne de ces estimateurs :

$$\bar{\hat{\theta}}_{n,m}^{(1)}=\frac{1}{n-1}\sum_{k=1}^{n-1}\hat{\theta}_{k,n,m}^{(1)}.$$

C'est aussi un estimateur asymptotiquement normal :

$$\sqrt{m}(\bar{\hat{\theta}}_{n,m}^{(1)}-\theta) \xrightarrow[m\to\infty]{d} \mathcal{N}(0,\sigma_{ML}^2)$$

avec :

$$\sigma_{ML}^2=\frac{1}{(n-1)^2}\left((n-1)\Psi_1(\theta+1)+\sum_{k=1}^{n-1}(k^2\Psi_1(k\theta)-(k+1)^2\Psi_1((k+1)\theta+1))\right).$$

Méthode des moments

Pour tout $k \in \{1,\dots,n-1\}$, on note $\hat{\theta}_{k,n,m}^{(2)}$ l'estimateur de la méthode des moments donné explicitement par :

$$\hat{\theta}_{k,n,m}^{(2)}=\frac{1-\bar{V}_k}{(n-k+1)\bar{V}_k-1}$$

où $\bar{V}_k$ est la moyenne empirique des variables aléatoires i.i.d. $V_k^{(1)},\dots,V_k^{(m)}$. Il s'agit également d'un estimateur asymptotiquement normal :

$$\sqrt{m}(\hat{\theta}_{k,n,m}^{(2)}-\theta) \xrightarrow[m\to\infty]{d} \mathcal{N}(0,\sigma_{MM,k}^2)$$

avec

$$\sigma_{MM,k}^2=\frac{\theta(\theta+1)(n-k)((n-k+1)\theta+1)^2}{(n-k)((n-k+1)\theta+2)}.$$

On peut ici également considérer l'estimateur moyen :

$$\bar{\hat{\theta}}^{(2)}_{n,m} = \frac{1}{n-1}\sum_{k=1}^{n-1} \hat{\theta}^{(2)}_{k,n,m}$$

qui est asymptotiquement normal :

$$\sqrt{m}(\bar{\hat{\theta}}^{(2)}_{n,m} - \theta) \xrightarrow[m\to\infty]{d} \mathcal{N}(0, \sigma^2_{MM})$$

où

$$\sigma^2_{MM} = \frac{\theta(\theta+1)}{(n-1)^2}\sum_{k=1}^{n-1} \frac{((k+1)\theta+1)^2}{k((k+1)\theta+2)}.$$

Dans [11], nous présentons une étude basée sur des données simulées, en particulier sur la comparaison des variances asymptotiques σ^2_{ML} et σ^2_{MM} (l'une n'est pas toujours plus petite que l'autre, cela dépend de la valeur de θ). Une application sur le nombre de naissances en 1999 dans trois pays (France, Israel et Malte) est également proposée.

9.2 Estimation des paramètres n et θ [12]

On s'intéresse au problème de l'estimation du nombre n d'espèces présentes sur une zone géographique donnée. Lors d'une campagne d'échantillonnage, k individus ont été prélevés. On note $A(q)$ le nombre d'espèces présentes q fois dans l'échantillon : $\sum_{q=0}^{k} A(q) = n$ et $\sum_{q=1}^{k} qA(q) = k$. De plus, le nombre $P_{n,k}$ d'espèces différentes observées dans l'échantillon s'exprime également en fonction de ces données : $P_{n,k} = \sum_{q=1}^{k} A(q)$. On cherche donc à estimer n sur la base de ces données synthétiques.

9.2.1 Modèle

On considère le modèle suivant : on suppose que les espèces sont présentes en proportion aléatoire $S_1, \ldots, S_n$, le vecteur $(S_1, \ldots, S_n)$ étant de loi de Dirichlet symétrique de paramètre θ. Ce modèle avait déjà été considéré par Keener *et al.* [115]. Il généralise quelques modèles classiques en écologie. Si $\theta = 1$, on obtient le modèle du bâton brisé utilisé par MacArthur [130] ; si $\theta \to \infty$, on obtient des proportions déterministes toutes égales ; enfin, si $\theta \to 0$, $n \to \infty$ et $n\theta \to \gamma > 0$, le vecteur des proportions est de loi Poisson-Dirichlet $PD(\gamma)$ et on obtient le modèle log-séries de Fisher [92].

En complément des notations précédemment introduites, on pose :

— pour tout $j \in \{1, \ldots, k\}$, M_j est l'espèce du j-ème individu de l'échantillon ;
— pour tout $m \in \{1, \ldots, n\}$, $\mathcal{K}_k(m) = \sum_{j=1}^{k} \mathbb{I}(M_j = m)$ est le nombre (éventuellement nul) d'occurrences de chaque espèce dans l'échantillon ;
— pour tout $q \in \{1, \ldots, p\}$, $\mathcal{B}_k(q) > 0$ est le nombre non nul d'individus de l'espèce q dans l'échantillon ;
— pour tout $i \in \{0, \ldots, k\}$, $\mathcal{A}_k(i)$ est le nombre d'espèces observées i fois dans l'échantillon :

$$\begin{aligned} \mathcal{A}_k(i) &= \#\{m \in \{1, \ldots, n\} : \mathcal{K}_k(m) = i\} \\ &= \sum_{m=1}^{n} \mathbb{I}(\mathcal{K}_k(m) = i). \end{aligned}$$

On dispose de formules d'échantillonnage de type Ewens pour la loi de Dirichlet [107] :

Théorème 9.2.1

1. *Pour tout $(b_1, \ldots, b_p)$ tel que $\forall q \in \{1, \ldots, p\}$, $b_q \geqslant 1$ et $\sum_{q=1}^{p} b_q = k$,*

$$\mathbb{P}(\forall j, \mathcal{B}_k(j) = b_j ; P_{n,k} = p) = \binom{n}{p} \frac{k!}{(n\theta)_k} \prod_{q=1}^{p} \frac{(\theta)_{b_q}}{b_q!}, \tag{9.2.1}$$

où $(x)_n$ est le symbole de Pochhammer (ou factorielle croissante) : $(x)_n = x(x+1)(x+2)\cdots(x+n-1) = \frac{\Gamma(x+n)}{\Gamma(x)}$.

2. Pour tout $(a_1, \ldots, a_k) \geqslant 0$ *tel que* $\sum_{i=1}^k i a_i = k$ *et* $\sum_{i=1}^k a_i = p$,

$$\mathbb{P}\left(\forall i, \mathcal{A}_k(i) = a_i; P_{n,k} = p\right) = \frac{n!k!}{(n-p)!(n\theta)_k} \prod_{i=1}^k \frac{(\theta)_i^{a_i}}{i!^{a_i} a_i!}. \tag{9.2.2}$$

Partant de ce résultat, on obtient la loi de la statistique $P_{n,k}$:

Théorème 9.2.2 *Pour tout* $p \geqslant 1$,

$$\mathbb{P}(P_{n,k} = p) = \frac{n!}{(n-p)!} \frac{1}{(n\theta)_k} B_{k,p}(\theta), \tag{9.2.3}$$

où $B_{k,p}(\cdot)$ *sont les polynômes de Bell définis par :*

$$B_{k,p}(\theta) = \frac{k!}{p!} \sum_{\substack{b_q \geqslant 1 \\ \sum_{q=1}^p b_q = k}} \prod_{q=1}^p \frac{(\theta)_{b_q}}{b_q!} = \sum_{\substack{a_i \geqslant 0 \\ \sum_{i=1}^k a_i = p \\ \sum_{i=1}^k i a_i = k}} k! \prod_{i=1}^k \frac{(\theta)_i^{a_i}}{i!^{a_i} a_i!}.$$

Keener *et al.* [115] avaient obtenu précédemment une autre expression de la loi de $P_{n,k}$:

Théorème 9.2.3 *[115] Pour tout* $p \geqslant 1$,

$$\mathbb{P}(P_{n,k} = p) = \sum_{q=1}^p (-1)^{p-q} \binom{n}{p} \binom{p}{q} \frac{(q\theta)_k}{(n\theta)_k}. \tag{9.2.4}$$

9.2.2 Estimation de n

On souhaite estimer n en supposant que le paramètre θ est connu. Les théorèmes 9.2.1 et 9.2.2 montrent que $P_{n,k}$ est une statistique suffisante. Le paramètre n étant entier, l'estimateur du maximum de vraisemblance $\widehat{n}_{hp}$ est définie par :

$$\frac{\mathbb{P}(P_{n,k} = p)}{\mathbb{P}(P_{n-1,k} = p)} = 1.$$

Partant de la loi de $P_{n,k}$, on obtient que $\widehat{n}_{hp}$ est la solution de l'équation suivante :

$$\widehat{n}_{hp} = p + \widehat{n}_{hp} \frac{((\widehat{n}_{hp} - 1)\theta)_k}{(\widehat{n}_{hp}\theta)_k}. \tag{9.2.5}$$

En comparaison, l'estimateur $\widehat{n}_{ksr}$ proposé par Keener *et al.* [115] est le suivant :

$$\widehat{n}_{ksr} = p + \frac{B_{k,p-1}(\theta)}{B_{k,p}(\theta)}.$$

Cet estimateur présente l'avantage d'être explicite. Cependant, il fait intervenir les polynômes de Bell dont le calcul numérique peut se révéler complexe. Si $\theta = 1$, alors les deux estimateurs sont :

$$\widehat{n}_{hp} = \frac{p(k-1)}{k-p} \qquad \text{et} \qquad \widehat{n}_{ksr} = \frac{pk}{k-p+1}$$

Si θ tend vers l'infini, $\widehat{n}_{hp}$ est alors la solution de l'équation suivante :

$$p = \widehat{n}_{hp} \left(1 - \left(1 - \frac{1}{\widehat{n}_{hp}}\right)^k\right)$$

et :

$$\widehat{n}_{ksr} = p + \frac{S_{k,p-1}}{S_{k,p}}$$

où $S_{k,p}$ sont les nombres de Stirling de seconde espèce.

9.2.3 Estimation de n et de θ

En général, le paramètre θ est également inconnu. On va donc devoir estimer conjointement n et θ. Cette situation n'est pas détaillée dans l'article de Keener *et al.* [115], bien que mentionnée dans les applications. Pour estimer ces deux paramètres, on va reprendre un des estimateurs étudiés dans [11]. Soit l_1 et l_2 deux animaux observés. On note δ_{l_1,l_2} la variable indicatrice de l'appartenance à la même espèce pour ces deux animaux :

$$\delta_{l_1,l_2} = \sum_{m=1}^{n} \mathbb{I}(M_{l_1} = m; M_{l_2} = m)$$

On peut alors définir l'homozygosité de l'échantillon [180] :

$$D = \frac{1}{k(k-1)} \sum_{l_1 \neq l_2 = 1}^{k} \delta_{l_1,l_2} = \frac{k}{k-1} \left(\sum_{q=1}^{P_{n,k}} \left(\frac{\mathcal{B}_k(q)}{k} \right)^2 - \frac{1}{k} \right).$$

L'espérance de D peut être calculée facilement :

$$\mathbb{E}(D) = \mathbb{E}(\delta_{l_1,l_2}) = \frac{1+\theta}{1+n\theta}$$

On en déduit un estimateur de θ par la méthode des moments :

$$\widehat{\theta} = \frac{1-D}{nD-1}.$$

Cet estimateur de θ dépend de n. On peut donc utiliser cet estimateur avec l'un des deux estimateurs de n introduits précédemment. En utilisant l'estimateur que nous avons proposé, on a :

$$\widehat{n}_{hp}^{(1)} = p + \widehat{n}_{hp}^{(1)} \frac{((\widehat{n}_{hp}^{(1)} - 1)\widehat{\theta}_{hp}^{(1)})_k}{(\widehat{n}_{hp}^{(1)}\widehat{\theta}_{hp}^{(1)})_k} \quad \text{et} \quad \widehat{\theta}_{hp}^{(1)} = \frac{1-D}{\widehat{n}_{hp}^{(1)} D - 1}.$$

De même, en utilisant l'estimateur que Keener *et al.* ont proposé, on a :

$$\widehat{n}_{ksr}^{(1)} = p + \frac{B_{k,p-1}(\widehat{\theta}_{ksr}^{(1)})}{B_{k,p}(\widehat{\theta}_{ksr}^{(1)})} \quad \text{et} \quad \widehat{\theta}_{ksr}^{(1)} = \frac{1-D}{\widehat{n}_{ksr}^{(1)} D - 1}.$$

Au lieu de considérer l'homozygosité, on peut considérer des fonctionnelles additives de la forme :

$$\phi_n = \sum_{m=1}^{n} h(S_m).$$

Par exemple, pour $h(s) = s^2$, on obtient l'indice de Simpson de biodiversité [173] ou bien pour $h(s) = -s \log s$, on obtient l'entropie de Shannon [154]. Un estimateur naturel de ϕ_n est le suivant :

$$\widehat{\phi}_k = \sum_{m=1}^{n} h\left(\frac{\mathcal{K}_k(m)}{k} \right)$$

On montre que :

$$\widehat{\phi}_k \xrightarrow[k\to\infty]{d} \phi_n$$

et

$$\mathbb{E}\left(\widehat{\phi}_k\right) \xrightarrow[k\to\infty]{} \mathbb{E}(\phi_n).$$

Cependant, $\widehat{\phi}_k$ dépend de n qui est inconnu. En revanche, on peut considérer l'estimateur suivant qui possède les mêmes propriétés asymptotiques :

$$\widehat{\psi}_k = \sum_{q=1}^{P_{n,k}} \frac{h\left(\frac{\mathcal{B}_k(q)}{k}\right)}{1 - \left(1 - \frac{\mathcal{B}_k(q)}{k}\right)^k}.$$

Dans le cas particulier où $h(s) = s^2$ (indice de Simpson), on a :

$$\mathbb{E}\left(\widehat{\psi}_k\right) \xrightarrow[k\to\infty]{} \frac{1+\theta}{1+n\theta}.$$

On peut donc obtenir un autre estimateur de θ en appliquant la méthode des moments à l'indice de Simpson. En utilisant $\widehat{n}_{hp}$, on obtient les estimateurs suivants :

$$\widehat{n}_{hp}^{(2)} = P + \widehat{n}_{hp}^{(2)} \frac{((\widehat{n}_{hp}^{(2)} - 1)\widehat{\theta}_{hp}^{(2)})_k}{(\widehat{n}_{hp}^{(2)}\widehat{\theta}_{hp}^{(2)})_k} \quad \text{et} \quad \widehat{\theta}_{hp}^{(2)} = \frac{1-\widehat{\psi}_k}{\widehat{n}_{hp}^{(2)}\widehat{\psi}_k - 1}.$$

De même, en utilisant $\widehat{n}_{ksr}$, on obtient les estimateurs suivants :

$$\widehat{n}_{ksr}^{(2)} = p + \frac{B_{k,p-1}(\widehat{\theta}_{ksr}^{(2)})}{B_{k,p}(\widehat{\theta}_{ksr}^{(2)})} \quad \text{et} \quad \widehat{\theta}_{ksr}^{(2)} = \frac{1-\widehat{\psi}_k}{\widehat{n}_{ksr}^{(2)}\widehat{\psi}_k - 1}.$$

Dans [12], nous avons considéré des questions connexes à ce problème. En particulier, nous avons regardé des critères d'arrêt d'échantillonnage. Trois règles d'arrêt ont été proposées, basées sur les questions suivantes : quand est-ce que l'espèce la plus rare est observée ? combien d'animaux faut-il prélever pour observer toutes les espèces ? quand est-ce que la probabilité d'observer une nouvelle espèce est inférieure à un certain seuil ? De plus, nous avons également étudié un critère d'ajustement.

9.2.4 Applications

Deux exemples d'applications sont donnés ici. Le premier exemple est une situation où, en fait, le paramètre à estimer est connu. Il avait été étudié par Keener *et al.* [115].

The Federalist Papers

Entre 1787 et 1788, une série de 77 articles, appelés *The Federalist Papers*, a été écrite pour soutenir la nouvelle constitution de l'État de New-York, publiée dans plusieurs journaux sous différents pseudonymes. On sait que ces articles ont été écrits par trois personnes (un seul auteur par article) : James Madison, Alexander Hamilton et John Jay. Pour la plupart des articles, l'auteur a été clairement identifié, mais Madison et Hamilton se disputent la paternité de douze articles. Afin de les départager, des chercheurs (voir, par exemple, le livre de Mosteller et Wallace [135]) ont étudié les textes où la paternité de Madison ou de Hamilton est indiscutable. Ils ont compté les occurrences des mots courants répertoriés de la liste Miller-Newman-Friedman. Keener *et al.* [115] n'ont utilisé que l'occurrence du mot "may" dans les textes attribués à Madison et l'occurrence du mot "can" dans les textes attribués à Hamilton. On connait donc le paramètre n que l'on cherche à estimer. On a donc appliqué notre estimateur à ces données et comparé avec les résultats obtenus par Keener *et al.* [115]. Pour les données relatives à Madison, on a : $n = 262$, $k = 172$ et $p = 106$. Les résultats sont les suivants.

Estimateur...	n	θ
basé sur l'homozygosité	274,6	1,09
basé sur l'indice de Simpson	274,6	1,09
de Keener *et al.*	217	1,998

Table 9.1 – Application aux données sur Madison

Pour les données relatives à Hamilton, on a : $n = 247$, $k = 90$ et $p = 139$. Les résultats sont les suivants. Les estimateurs que nous avons proposé, en particulier celui basé sur l'homozygosité, semblent donc plus performant, du moins sur ce jeu de données, que celui proposé par Keener *et al.* [115]. Malheureusement, Keener *et al.* n'expliquent pas comment ils procèdent pour l'estimation conjointe de n et de θ. L'estimateur basé sur l'homozygosité semble sur-estimer le paramètre n.

Estimateur...	n	θ
basé sur l'homozygosité	253,5	0,85
basé sur l'indice de Simpson	4526,3	0,01
de Keener *et al.*	10 000 001	$1,09 \times 10^{-5}$

Table 9.2 – Application aux données sur Hamilton

Insectes tropicaux

Janzen [109] a observé des scarabées sur ving-cinq sites au Costa-Rica et dans les Caraïbes. Son article contient toutes les données synthétisées comme dans le tableau ci-dessous (dans [12], nous avons utilisés deux autres jeux de données). Sur 835 scarabées observés, Janzen a donc observé 151 espèces différentes.

q	1	2	3	4	5	7	8	9	10	11	12
$A(q)$	61	24	13	12	5	6	5	2	4	2	3
q	13	15	17	18	19	26	30	33	40	44	62
$A(q)$	1	1	1	2	2	1	1	1	1	1	2

Table 9.3 – Osa secondary/night/dry/1967 data

En utilisant le premier estimateur, on obtient $\widehat{n}_{hp}^{(1)} = 184.1$ et $\widehat{\theta}_{hp}^{(1)} = 0.268$. En utilisant le second estimateur, on obtient $\widehat{n}_{hp}^{(2)} = 184.1$ et $\widehat{\theta}_{hp}^{(2)} = 0.252$.

Troisième partie

Études statistiques dans le domaine de la santé et de l'environnement

Chapitre 10

Introduction

La troisième partie de ce manuscrit rassemble plusieurs études pour lesquelles j'ai contribué à l'analyse statistique des données. Il s'agit donc plutôt d'applications de la statistique que des développements statistiques originaux et novateurs. Néanmoins, dans certains cas, un travail de modélisation a été effectué. Ces travaux ont en commun de concerner deux champs disciplinaires : la médecine et l'écologie.

En 2004, Jean-François Delmas (CERMICS, ENPC) m'a proposé de travailler sur un projet dans le cadre du CEMRACS [1], en collaboration avec des collègues biologistes (Université Paris 6 et IRD Montpellier). Durant l'été 2004, j'ai donc travaillé essentiellement avec Guillaume Constantin de Magny (alors doctorant à l'IRD) et Michel de Lara (CERMICS, ENPC) sur ce projet. La collaboration s'est prolongée au-delà du centre d'été et un article a été publié dans les actes du CEMRACS [26]. Une partie du travail a porté sur la modélisation conjointe de l'évolution de l'épidémie de choléra et de la dynamique du phytoplancton.

Dès mon arrivée à Pau en février 2005, Noëlle Bru (UPPA) m'a proposé de travailler sur des problématiques liées à l'environnement. En collaboration avec des collègues biologistes d'Arcachon (Université Bordeaux 1) et d'Anglet (IFREMER), un premier sujet a porté sur l'analyse de données de palourdes (Bassin d'Arcachon). L'objectif était double : étudier la dynamique de population ainsi que la croissance des palourdes. La modélisation de la dynamique de population a été réalisée en collaboration avec Laurent Bordes (UPPA) à l'aide de modèles de mélange de lois. Quant aux modèles de croissance, les problèmes étudiés provenaient de questions autour d'un logiciel gratuit spécialisé pour les données halieutiques. Une question a porté sur la comparaison des paramètres de croissance selon le site et le niveau hypsométrique (profondeur) [9]. Par ailleurs, en collaboration avec Laurence Després, une collègue biologiste de Grenoble, nous avons travaillé sur l'analyse de données sur la production annuelle de cônes de mélèze dans le Briançonnais. Nous avons alors étudié et comparé deux approches complémentaires : les modèles de régression pour données catégorielles et les modèles auto-régressifs discrets [44].

Depuis 2010, j'ai été sollicité par des médecins du CHU de Bordeaux pour les aider à mener l'analyse statistique de leurs données. Cela a été l'occasion de faire travailler des étudiants de première année de master. De cette collaboration, deux articles [8, 10] ont été publiés dans revues à comité de lecture et un article [35] dans les actes d'un congrès européen.

Enfin, plus récemment, je travaille avec Béatrice Lauga (UPPA), sur l'analyse statistique des données d'abondance de bactéries dans des drainages miniers acides. Afin d'étudier les interactions entre les bactéries et leurs environnements, nous utilisons les outils d'analyse que sont les graphes et les réseaux.

A quelques exceptions près, on aura donc constaté que ma participation à ces travaux de recherche est essentiellement du soutien méthodologique pour l'analyse de données. Les collègues des autres disciplines sont de plus en plus confrontés à l'acquisition de données (éventuellement massives) et se retrouvent parfois démunis sur le traitement statistique adapté à appliquer.

Le tableau ci-dessous résume quantitativement mon activité de recherche sur ce thème, ainsi que les collaborations trans-disciplinaires.

1. Centre d'Eté Mathématique de Recherche Avancée en Calcul Scientifique.

Publications	[9, 8, 10]
Actes de congrès	[26, 35]
Pré-publications	[44]
Collaborations trans-disciplinaires	AZTI [2]
	CHU Bordeaux, Université de Bordeaux
	IFREMER, Anglet
	IPREM-EEM, UPPA
	IRD, Montpellier
	EPOC, Université de Bordeaux
	Laboratoire d'Écologie Alpine (LECA), Université de Grenoble
	Laboratoire Écologie et Évolution, Université Paris 6

2. AZTI est un centre technique spécialisé dans la recherche marine et alimentaire, financé par la Communauté autonome du Pays basque.

Chapitre 11

Études dans le domaine de la santé

11.1 Épidémiologie du choléra et environnement [26]

Cette étude a été effectuée durant l'édition 2004 du CEMRACS. Durant quatre semaines, G. Constantin de Magny (alors étudiant en thèse) et moi-même, nous avons séjourné au CIRM pour travailler quotidiennement sur ce sujet, tout en étant en contact régulier avec les autres membres impliqués dans le projet. Cette collaboration s'est ensuite poursuivie durant toute la thèse de Guillaume (2006).

Ce projet relève du domaine émergent de l'écologie de la santé. Il s'agit d'une approche transdisciplinaire d'un problème de santé publique. L'étude en question a porté sur le choléra, maladie infectieuse qui demeure un grave problème de santé publique dans les pays en développement avec une mortalité importante et avec de nombreux phénomènes de ré-émergences ces dix dernières années. Les analyses antérieures ont évoqué des processus complexes à l'origine de la dynamique de la maladie et plus particulièrement des relations avec les écosystèmes aquatiques et les facteurs environnementaux associés. Les modèles mathématiques sont nécessaires pour améliorer la compréhension de la complexité de ces processus écologiques et épidémiologiques, et pour mieux appréhender la dynamique des épidémies de choléra. Après avoir fait un état de l'art des modèles épidémiologique sur le choléra, nous avons proposé un modèle modifié intégrant des forçages environnementaux. En effet, des analyses statistiques ayant révélé que la concentration en chlorophylle dans l'eau avait une influence significative sur les épidémies de choléra, notre modèle intègre cette association et nous avons proposé un modèle dans lequel l'épidémie est initiée par un bloom (pic) phytoplanctonique puis se propage dans la population humaine.

Le choléra est une maladie ancienne qui a disparu de la plupart des pays développés durant la seconde partie du vingtième siècle, mais reste présente dans différentes parties du monde, essentiellement dans des zones tropicales et cela malgré des efforts sanitaires (traitement des eaux, etc.). Cette maladie hautement contagieuse est due à de l'eau contaminée par la bactérie *vibrio cholerae* qui est présente dans l'écosystème de nombreuses zones estuariennes ou côtières. Cette bactérie est fortement liée aux phytoplanctons et aux zooplanctons. Ainsi une modification climatique et/ou environnementale peut être la cause de l'émergence d'épidémie dans des populations humaines. Le premier modèle mathématique intégrant des variables environnementales a été proposé par Capasso et Paveri-Fontana en 1979 [71]. Depuis ce modèle, seulement quelques autres modèles ont été proposés [78, 147, 157].

Le choléra est une maladie à déclarations obligatoires et les données sont publiées dans le Bulletin Épidémiologique Hebdomadaire de l'Organisation Mondiale de la Santé (OMS) disponible en ligne. Tous les pays sont censés déclarer à l'OMS le nombre d'infections et de morts liés à certaines maladies dont le choléra. Nous avons donc collecté toutes les données disponibles pour deux pays d'Afrique, le Mozambique et la Somalie. Par ailleurs, la Terre faisant l'objet d'une surveillance de plus en plus étroite, des images satellites de la NASA sont disponibles concernant la concentration en chlorophylle. Après un traitement des images, on retient la concentration maximale à 50 Kilomètres des côtes. Nous disposons ainsi de données épidémiologiques et environnementales sur la période allant de septembre 1997 à décembre 2002.

En se basant sur les modèles existants et sur les données disponibles, nous avons proposé un modèle reposant sur deux hypothèses qui diffèrent des modèles déjà étudiées. Tout d'abord, compte-tenu de la

période de l'étude (vingt-cinq années), nous devons intégrer la croissance de la population humaine. De plus, des études sur des volontaires ont montré une certaine immunité temporaire. Cela nous a conduit à considérer le modèle à compartiments ci-dessous (l'unité de temps est le mois) :

$$\begin{cases} \frac{dS_t}{dt} &= (b-d)S_t - \beta S_t I_t - \frac{\gamma S_t C_{t-\delta}}{k+C_{t-\delta}} + \rho' R_t \\ \frac{dI_t}{dt} &= \beta S_t I_t + \frac{\gamma S_t C_{t-\delta}}{k+C_{t-\delta}} - \tau I_t \\ \frac{dD_t}{dt} &= \lambda \tau I_t + d(S_t + R_t) \\ \frac{dR_t}{dt} &= (1-\lambda)\tau I_t + (b-d)R_t - \rho' R_t \end{cases} \tag{11.1.1}$$

Le tableau 11.1 résume les différentes variables et les différents paramètres.

Symboles	Description
Variables	
S	nombre d'individus susceptibles
I	nombre d'individus infectés
D	nombre d'individus morts
R	nombre d'individus en rémission
C	concentration en chlorophylle
Paramètres	
H	population initiale totale
b	taux de natalité (uniquement pour les susceptibles)
d	taux de mortalité naturelle (uniquement pour les susceptibles)
δ	paramètre de décalage (en mois^{-1})
β	taux d'infection par contact entre infectés et susceptibles
k	quantité de phytoplanctons donnant 50% de chance d'être contaminé
τ	taux de rémission
λ	taux de mortalité lié au choléra
ρ'	taux de perte d'immunité
γ	taux d'infection d'un individu susceptible par de l'eau contaminé

Table 11.1 – Modèle SIDR - liste des variables et des paramètres

Certains paramètres ont été fixés. Les autres paramètres de ce modèle ont été ajustés en minimisant la fonction non-linéaire suivante :

$$J(I_0, R_0, \beta, \gamma, k) = \sum_{k=0}^{N} \left(\Delta Y_k - \int_{(k-1)\Delta t}^{k\Delta t} \tau I_s ds \right)^2,$$

où ΔY_k est le nombre d'hospitalisations et de décès sur l'intervalle de temps $[(k-1)\Delta t, k\Delta t]$ et N le nombre de périodes observations. Les résultats sont présentés dans le tableau 11.2.

Paramètres fixés	Mozambique	Somalie
b	0.0033	0.0038
d	0.0013	0.0015
H	3 960 000	1 980 000
τ	15.25	15.25
λ	8 %	8 %
ρ'	0.167	0.167
δ	3.5	3.5
Paramètres estimés	**Mozambique**	**Somalie**
I_0	37.438	4.144
R_0	8.920	8.377
β	1.76×10^{-10}	7.22×10^{-8}
γ	2.18×10^{-3}	2.20×10^{-4}
k	0.79	0.26

Table 11.2 – Modèle SIDR - paramètres fixés et estimés

11.2 Applications de la statistique en médecine [35, 8, 10]

Depuis 2010, j'ai été régulièrement sollicité par des médecins et des chercheurs du CHU de Bordeaux pour les aider à analyser leurs données. Les méthodes statistiques utilisées restent relativement classiques, le plus difficile étant de bien comprendre les questions et de pouvoir y apporter des éléments de réponse.

11.2.1 Étude CIRS [35]

J. Youssef (alors chef de clinique en réanimation à l'Hôpital Saint-André, CHU de Bordeaux) et ses collègues s'intéressent à l'intérêt de mesurer le nombre de cellules dendritiques (un type de cellules immunitaires) chez des patients dans un état de choc. En effet, on sait que ce type de cellules est diminué lors d'une infection grave et cette diminution l'est d'autant plus que le pronostic est mauvais. Ils souhaitent comprendre le comportement de ces cellules dans un état de choc non infectieux (choc cardiogénique) et dans une infection sévère sans choc (sepsis sévère) afin de pouvoir déterminer la valeur diagnostique dans un état de choc (entre choc cardiogénique et choc septique) et la valeur pronostique dans le cas d'une infection sévère avec ou sans choc.

11.2.2 Étude NESAKI [8]

Cette étude a été en partie réalisée par deux étudiants palois en 1ère année de master. Leur travail arrive en soutien de la thèse de médecine de S. Oger sous la direction du Docteur F. Camou. L'étude prospective NESAKI (NGAL Evaluation In Septic Acute Kidney Injury) concerne l'intérêt du dosage du marqueur biologique NGAL comme marqueur prédictif de l'insuffisance rénale aiguë par nécrose tubulaire au cours du choc septique. Au cours des septicémies, les reins voient leur performance diminuer ce qui a pour conséquence l'apparition d'une insuffisance rénale aiguë sur quelques heures. Dans certain cas, cette insuffisance rénale n'est que transitoire et régresse avec la ré-hydratation, mais dans d'autres cas, elle s'aggrave et justifie le recours à une dialyse, il s'agit alors d'une insuffisance rénale organique. Dans ce dernier cas, il est connu que le retard à la mise en route de la dialyse a un impact sur le pronostic. L'objectif de l'étude statistique était de déterminer l'existence ou pas d'un seuil critique du marqueur NGAL pour la

mise sous dialyse d'un patient (si existence, quelle la meilleure date : dès l'arrivée dans le service ? à J+1 ? à J+2 ?).

11.2.3 Étude sur la croissance tumorale [10]

L'objectif de cette étude est de déterminer si certains lymphocytes humains peuvent inhiber la propagation des tumeurs *in vivo* dans le côlon humain et d'en évaluer son potentiel immunothérapeutique. Dans cette étude, des cellules humaines exprimant une activité bioluminescente ont été injectées dans des souris immunodéficientes. Pour analyser la relation entre le volume et la lumière émise par des tumeurs sous-cutanées bioluminescentes, nous avons utilisé le coefficient de corrélation de Spearman sur l'ensemble des données (ensemble des souris et à des instants d'observation). Puis une version du test de Wilcoxon-Mann-Whitney pour des durées censurées par intervalle a été appliqué pour comparer le temps d'apparition de métastases dans les deux groupes de souris (un groupe témoin et un groupe traité), mais également pour comparer la taille des tumeurs primaires lors de leur apparition.

En utilisant les techniques d'imagerie par bioluminescence pour suivre l'apparition de cellules cancéreuses dans l'intestin, cette étude montre qu'un traitement systémique par des cellules lymphatiques permet d'inhiber non seulement la croissance de la tumeur primaire de l'intestin, mais aussi l'émergence de tumeurs secondaires situées dans les poumons et le foie.

Chapitre 12

Études dans le domaine de l'environnement

12.1 Dynamique de croissance de palourdes [9]

Dans le cadre d'un projet en collaboration avec le LRHA de l'IFREMER, la Station Marine d'Arcachon du laboratoire EPOC et l'AZTI, avec N. Bru (UPPA), nous avons travaillé sur la dynamique de croissance de palourdes dans le Bassin d'Arcachon. Une contribution au projet a porté sur l'ajustement des paramètres dans un modèle (déterministe) de croissance de la longueur des palourdes. Le modèle de base est celui proposé par von Bertalanffy. On suppose que la longueur $(L_t)_{t\geqslant 0}$ est de la forme :

$$\forall t \geqslant 0\ , \quad L_t = L_\infty \left(1 - \exp\left(-k(t - t_0)\right)\right).$$

Les paramètres L_∞ et k représentent respectivement la longueur maximale pouvant être atteinte par une palourde et la vitesse à laquelle cette longueur maximale peut être atteinte. Le paramètre t_0 correspond à la date où la longueur est nulle. Ce paramètre, propre à chaque individu, n'est pas connu. On ne peut pas ajuster les paramètres à partir de données âges-longueurs. On se placera dans le cas de données de marquage-recapture, situation pour laquelle il existe des méthodes d'estimation spécifiques.

Une variante de ce modèle, le modèle de Somers, consiste à prendre en compte des effets saisonniers dans la croissance. Le modèle devient alors le suivant :

$$\forall t \geqslant 0\ , \quad L_t = L_\infty \left(1 - \exp\left(-k(t - t_0) - \frac{ck}{2\pi}\sin 2\pi(t - t_s) + \frac{ck}{2\pi}\sin 2\pi(t_0 - t_s)\right)\right).$$

où $t_s + 0.5$ est le point hivernal, i.e. le moment de l'année où la croissance est la plus lente et c est un paramètre modulant la saisonnalité dans la croissance :

— si $c = 0$, alors il n'y a pas de saisonnalité ;
— si $c = 1$, alors le taux de croissance est nulle exactement une fois par an ;
— si $0 < c < 1$, alors le taux de croissance baisse en période hivernale sans pourtant être nulle ;
— si $c > 1$, alors le taux de croissance est nul sur plusieurs semaines ou mois par année.

Il existe plusieurs méthodes d'estimation des paramètres dans le modèle de von Bertalanffy ou dans le modèle de Somers pour des données de marquage-recapture. Dans ces cas-là, on s'intéresse à tous les paramètres, excepté t_0 (qui varie d'un individu à un autre). La méthode de Gulland et Holt repose sur une équation différentielle satisfaite par $(L_t)_{t\geqslant 0}$ pour le modèle de von Bertalanffy, qui est ensuite approchée par leur accroissement (ce qui peut manquer de précisions). Toujours pour le modèle de von Bertalanffy, la méthode de Munro repose sur une expression de k en fonction des accroissements et de L_∞. Ainsi, pour une valeur de L_∞ fixée, on peut déterminer k^* en minimisant un critère des moindres carrés. On optimise alors le critère par rapport à L_∞ pour déterminer à la fin le couple optimal des paramètres (il s'agit donc d'une méthode d'estimation profilée). Enfin, la méthode d'Appeldoorn permet d'ajuster les paramètres du modèle

de Somers. Ce modèle peut être écrit de la manière suivante :

$$\frac{L_\infty - L_t}{L_\infty} = \exp(-k(t - t_0) - S_t + S_{t_0}),$$

où $(S_t)_{t\geqslant 0}$ est la composante saisonnière (composante sinusoïdale dans le modèle de Somers), i.e. pour tout $t \geqslant 0$, $S_t = \frac{ck}{2\pi} \sin 2\pi(t - t_s)$. D'où, pour $t_1 < t_2$, on obtient que :

$$\frac{L_\infty - L_{t_2}}{L_\infty - L_{t_1}} = \exp(-k(t_2 - t_1) + S_{t_1} - S_{t_2}). \tag{12.1.1}$$

D'où, on en déduit une expression de L_{t_2} en fonction de L_{t_1} :

$$L_{t_2} = L_\infty - (L_\infty - L_{t_1} \exp(-k(t_2 - t_1) + S_{t_1} - S_{t_2}).$$

En partant de cette équation, Appeldoorn (1987) a proposé d'estimer les paramètres du modèle de Somers par régression non-linéaire. Soriano et Pauly (1989) ont proposé une régression linéaire lorsque le paramètre L_∞ est fixé et on peut alors procéder comme avec la méthode de Munro.

Toutes ces méthodes d'ajustement sont implémentées dans le logiciel gratuit FiSAT II développé pour la FAO pour des données obtenues par marquage-recapture. La multiplicité des méthodes proposées peut laisser un collègue biologiste dubitatif sur laquelle choisir. C'est un des points sur lequel nous avons été sollicités dont la question peut se résumer ainsi : quel modèle faut-il choisir et quelle méthode d'estimation faut-il employer ? Pour répondre à cette question, nous avons étudié de près le logiciel de la FAO. De trop nombreux bugs ont alors été décelés mettant en cause sérieusement la validité de ce logiciel : erreur dans l'affichage des résultats (l'ordre des coefficients n'est pas celui indiqué), informations répétées à différents endroits mais avec des valeurs différentes, saturation des contraintes dans les méthodes d'optimisation numérique avec bornes, coefficients de détermination négatifs, etc. Ceci nous a obligé à implémenter ces méthodes d'estimation avec R.

Dans le cadre de ce projet, les données ont été recueillies sur quatre sites différents sur le bassin d'Arcachon (Lanton, Gujan, Andernos et Anguin) et sur quatre niveaux hypsométriques différents. On a donc ajusté les paramètres pour ces seize localisations. La question naturelle était alors de déterminer si le site ou le niveau hypsométrique ont une influence sur l'estimation des paramètres. Pour répondre à cette question, nous nous sommes basés sur le test de Kruskal-Wallis.

Ce projet a été soutenu par le Fond Commun de Coopérations Aquitaine-Euskadi qui a financé, entre autre, la thèse de C. Dang (biologie).

12.2 Étude de l'assemblage des communautés microbiennes

Une partie de la thèse de F. Javerliat a été consacrée à l'étude statistique de données pour mieux comprendre l'assemblage des communautés microbiennes. Les données faisant l'objet de cette étude avaient été produites grâce au séquençage massif d'amplicons ciblant les gènes 16S des communautés microbiennes de douze drainages miniers acides (DMA) répartis sur trois zones géographiques distinctes (France, Espagne et Bolivie). Dans un premier temps, la fouille de données publiques qui a permis de révéler les micro-organismes inféodés à ce milieu. Puis, nous avons construit des réseaux de cooccurrence permettant de révéler des interactions spécifiques entre certains de ces micro-organismes. Un article est en cours de rédaction. Cette collaboration a reçu le soutien de l'Université via un BQR en 2013.

Quatrième partie

Bibliographie

Liste de publications

Articles parus dans des revues avec comité de lecture

[1] N. Balakrishnan, C. Paroissin, and J.-C. Turlot. One-sided control charts based on precedence and weighted precedence statisticss. *Quality and Reliability Engineering International*, 31(1) :113–134, 2015. Special Issue on Nonparametric Statistical Process Control Charts (Guest editors : S. Chakraborti, P. Qiu and A. Mukherjee).

[2] J. Barrera, T. Huillet, and C. Paroissin. Size-biased permutation of Dirichlet partitions with application to search cost distribution. *Probab. Engrg. Inform. Sci.*, 19 :83–97, 2005.

[3] J. Barrera, T. Huillet, and C. Paroissin. Limiting search cost distribution for the move-to-front rule with random request probabilities. *Operat. Res. Letters*, 34(5) :557–563, 2006.

[4] J. Barrera and C. Paroissin. On the distribution of the stationary search cost for the move-to-front rule with random weights. *J. Appl. Prob.*, 41(1) :250–262, 2004.

[5] A. Billon, L. Bordes, P. Darfeuil, S. Humbert, and C. Paroissin. Modélisation de la dégradation d'un composant à partir du retour d'expérience. *Journal de la Société Française de Statistique*, 155(3) :26–47, 2014. Numéro spécial "Fiabilité" (Éditeur invité : O. Gaudoin).

[6] L. Bordes, C. Paroissin, and A. Salami. Parametric inference in a perturbed gamma degradation process. Accepted in *Commun. Statist. - Theor. Meth.*, 2014.

[7] N. Bousquet, M. Fouladirad, A. Grall, and C. Paroissin. Bayesian gamma processes for optimising condition-based maintenance under uncertainty. Accepted in *Appl. Stoch. Model Bus.* (special issue for MMR'13 conference), 2014.

[8] F. Camou, S. Oger, C. Paroissin, E. Guilhon, O. Guisset, G. Mourissoux, H. Pouyès, T. Lalanne, and C. Gabinski. Le dosage plasmatique de Neutrophil Gelatinase-Associated Lipocalin (NGAL) prédit la défaillance rénale au cours du choc septique dès l'admission en réanimation. *Ann. Fr. Anesth.*, 32(3) :157–164, 2013.

[9] C. Dang, X. de Montaudouin, M. Gam, C. Paroissin, N. Bru, and N. Caill-Milly. The Manila clam population in Arcachon Bay (SW France): can it be kept sustainable? *J. Sea Res.*, 63(2) :108–118, 2010.

[10] C. Devaud, B. Rousseau, V. Pitard, S. Netzer, C. Paroissin, P. Costet, J.-F. Moreau, F. Couillaud, J. Dechanet-Merville, and M. Capone. Anti-metastatic potential of human V$\delta 1^+$ $\gamma\delta$ T cells in an orthotopic mouse xenograft model of colon carcinoma. *Cancer Immunol. Immun.*, 62(7) :1199–1210, 2013.

[11] T. Huillet and C. Paroissin. Estimation of the parameter of a Dirichlet partition using residual allocation model representations and sampling properties. *Stat. Meth.*, 2(2) :95–110, 2005.

[12] T. Huillet and C. Paroissin. Sampling from Dirichlet populations: estimating the number of species. *Environmetrics*, 20(7) :853–876, 2009.

[13] F. Leisen, A. Lijoi, and C. Paroissin. Limiting behavior of the search cost distribution for the move-to-front rule in the stable case. *Stat. Probab. Letters*, 81(12) :1827–1832, 2011.

[14] C. Paroissin and L. Rabehasaina. First and last passage times of spectrally positive lévy processes with application to reliability. Accepted in *Methodol. Comput. Appl. Probab.*, 2013.

[15] C. Paroissin, E. Remy, and V. Verrier. Inférence statistique pour un modèle markovien de dégradation avec covariables dépendantes du temps. *Journal de la Société Française de Statistique*, 155(1) :99–116, 2014. Numéro spécial "Données longitudinales quantitatives, événementielles incomplètement observées" (Éditeurs invités : L. Bordes et D. Commenges).

[16] C. Paroissin and A. Salami. Failure time of non-homogeneous gamma processes. *Commun. Statist. - Theor. Meth.*, 43(15) :3148–3161, 2014.

[17] C. Paroissin and B. Ycart. Zero-one law for the non-availability of multistate repairable systems. *International Journal of Reliability, Quality and Safety Engineering*, 10(3) :311–322, 2003.

[18] C. Paroissin and B. Ycart. Central limit theorem for hitting times of functionals of Markov jump processes. *ESAIM - P&S*, 8 :66–75, 2004.

Articles parus dans des actes de conférence avec comité de lecture

[19] J. Barrera and C. Paroissin. On the stationary search cost for the move-to-root rule with random weights. In D. Gardy M. Drmota, P. Flajolet and B. Gittenberger, editors, *Mathematics and computer science III. Algorithms, trees, combinatorics and probabilities (Proceedings of the Third International Colloquium of Mathematics and Computer Sciences, Vienna, September 13-17, 2004)*, pages 147–148. Birkhäuser, Basel, 2004.

[20] A. Billon, S. Baysset, S. Humbert, P. Darfeuil, L. Bordes, and C. Paroissin. Modélisation statistique des dégradations de moteurs aéronautiques à partir de données de retour d'expériences. In *"Innovation et Maîtrise des risques", Actes du Congrès $\lambda - \mu$ 17, La Rochelle, 5-7 octobre 2010*. Institut pour la Maîtrise des Risques, 2010.

[21] A. Billon, S. Humbert, P. Darfeuil, L. Bordes, and C. Paroissin. Statistical modelling of aeronautical turboshaft engines ageing from field and repair data feedback including preventive maintenance. In C. Bérenguer, A. Grall, and C.G. Soares, editors, *"Risk, Reliability and Societal Safety", Proceedings of the European Safety and Reliability Conference 2011 (ESREL 2011), Troyes, France, 18-22 September 2011*. Taylor and Francis, 2011.

[22] E. Biritxinaga-Etxart, S. Baysset, L. Bordes, J.-M. Bosc, C. Paroissin, B. Puig, W. Tinsson, and J.-L. Vérit. Evaluation probabiliste du temps entre l'apparition et la constatation d'un endommagement sur un matériel aéronautique. In *"Les nouveaux défis de la Maîtrise des Risques", Actes du Congrès $\lambda - \mu$ 16, Avignon, 7-9 octobre 2008*. "Institut pour la Maîtrise des Risques, 2008.

[23] L. Bordes, C. Paroissin, and J.-C. Turlot. Early detection of change-point in occurrence rate with small sample size. In C. Bérenguer, A. Grall, and C.G. Soares, editors, *"Risk, Reliability and Societal Safety", Proceedings of the European Safety and Reliability Conference 2011 (ESREL 2011), Troyes, France, 18-22 September 2011*. Taylor and Francis, 2011.

[24] L. Bordes, C. Paroissin, J.-C. Turlot, and F. Ibled. Détections de rupture dans les taux de défaillance pour de petits échantillons. In *"Innovation et Maîtrise des risques", Actes du Congrès $\lambda - \mu$ 17, La Rochelle, 5-7 octobre 2010*. Institut pour la Maîtrise des Risques, 2010.

[25] L. Bordes, C. Paroissin, V. Verrier, and E. Remy. Statistical inference of a discrete-time markovian degradation model with time-dependent covariates. In R. Virolainen and T. Aven, editors, *Risk, Reliability and Societal Safety, Proceedings of the European Safety and Reliability Conference 2012 (ESREL 2012), Helsinki, Finland, 18-22 June 2012*. Taylor and Francis, 2012.

[26] G. Constantin de Magny, C. Paroissin, B. Cazelles, J.-F. Delmas, M. de Lara, and J.-F. Guégan. Modeling environmental impacts of plankton reservoirs on cholera population dynamics. In E. Cancès and J.-F. Gerbeau, editors, *ESAIM - Proc. (Actes du CEMRACS 2004, Mathématique et applications en biologie et médecine)*, volume 14, pages 156–173, 2005.

[27] M. Fouladirad, A. Grall, and C. Paroissin. Semi-parametric abrupt change detection and condition-based rmaintenance. In C. Bérenguer, A. Grall, and C.G. Soares, editors, *"Risk, Reliability and Societal Safety", Proceedings of the European Safety and Reliability Conference 2011 (ESREL 2011), Troyes, France, 18-22 September 2011*. Taylor and Francis, 2011.

[28] M. Fouladirad, C. Paroissin, N. Bousquet, and A. Grall. Bayesian optimization of condition-based maintenance under uncertainty. In M. Finkelstein, editor, *Eighth International Conference on Mathematical Methods in Reliability (MMR2013), Stellenbosch, South Africa, 1-4 July 2013*, 2013.

[29] M. Fouladirad, C. Paroissin, and A. Grall. Condition-based maintenance and non-homogenous gamma process. In B.J.M. Ale, I.A. Papazoglou, and E. Zio, editors, *"Risk, Reliability and Societal Safety", Proceedings of the European Safety and Reliability Conference 2010 (ESREL 2010), Rhodos, Greece, 7-10 September 2010*. CRC Press, 2010.

[30] C. Paroissin. About the estimation of the redundancy of a system. In B.J.M. Ale, I.A. Papazoglou, and E. Zio, editors, *"Risk, Reliability and Societal Safety", Proceedings of the European Safety and Reliability Conference 2010 (ESREL 2010), Rhodos, Greece, 7-10 September 2010*. CRC Press, 2010.

[31] C. Paroissin. Non-homogeneous degradation models : a review and some new results. In R. Steenbergen and P. van Gelder, editors, *Proceedings of the European Safety and Reliability Conference 2013 (ESREL 2013), Amsterdam, The Netherlands, 30 September - 2 October 2013*, 2013.

[32] C. Paroissin. Semi-parametric inference for wiener processes with non-linear drift. application to crack growth data. In M. Finkelstein, editor, *Eighth International Conference on Mathematical Methods in Reliability (MMR2013), Stellenbosch, South Africa, 1-4 July 2013*, 2013.

[33] C. Paroissin, M. Fouladirad, and A. Grall. Sensitivity analysis of the optimality of some maintenance policy for gamma degradation processes. In R. Virolainen and T. Aven, editors, *Risk, Reliability and Societal Safety, Proceedings of the European Safety and Reliability Conference 2012 (ESREL 2012), Helsinki, Finland, 18-22 June 2012*. Taylor and Francis, 2012.

[34] C. Paroissin and L. Rabehasaina. On the gamma process modulated by a markov jump process. In C. Bérenguer, A. Grall, and C.G. Soares, editors, *Risk, Reliability and Societal Safety, Proceedings of the European Safety and Reliability Conference 2011 (ESREL 2011), Troyes, France, 18-22 September 2011*. Taylor and Francis, 2011.

[35] J. Youssef, O. Guisset, F. Camou, C. Paroissin, M. Grenouillet-Delacre, A. Boyer, C. Gabinski, D. Gruson, and P. Blanco. Decrease of circulating dendritic cells in shock : a severe sepsis marker. In *Proceedings of the European Society of Intensive Care Medicine (ESICM) Annual Congress*, 2010.

Livres

[36] C. Paroissin. *Programmation et analyse statistique avec* R. À paraître chez Ellipses, 2015.

[37] S. Mercier W. Kahle and C. Paroissin. *Degradation processes in reliability*. À paraître chez ISTE-Wiley, 2016. (en cours de rédaction).

Chapitres de livre

[38] C. Paroissin and L. Rabehasaina. Time-to-failure of markov-modulated gamma process with application to replacement policies. In V. Couallier, L. Gerville-Réache, C. Huber-Carol, N. Limnios, and M. Mesbah, editors, *Statistical Models and Methods for Reliability and Survival Analysis. Book in honour of M. Nikulin*, chapter 24. ISTE-Wiley, 2013.

Co-direction d'ouvrages

[39] M.C. López de Silanes, M. Palacios, G. Sanz, J.J. Torrens, C. Paroissin M. Madaune-Tort, and D. Trujillo, editors. *Tenth International Conference Zaragoza-Pau on Applied Mathematics and Statistics*,

volume 35 of *Monografías del Seminario Matemático "García De Galdeano"*. Prensas Universitarias de Zaragoza, 2010. [ISBN 978-84-15031536].

[40] L.M. Esteban, B. Lacruz, F.J. López, P. Mateo, A. Pérez-Palomares, G. Sanz, and C. Paroissin, editors. *Pyrenees International Workshop on Statistics, Probability and Operations Resaearch (SPO'09)*, volume 36 of *Monografías del Seminario Matemático "García De Galdeano"*. Prensas Universitarias de Zaragoza, 2011. [ISBN 978-84-15031925].

[41] J. Giacomoni, M. Madaune-Tort, C. Paroissin, G. Vallet, M.C. López de Silanes, M. Palacios, G. Sanz, and J.J. Torrens, editors. *Eleventh International Conference Zaragoza-Pau on Applied Mathematics and Statistics*, volume 37 of *Monografías del Seminario Matemático "García De Galdeano"*. Prensas Universitarias de Zaragoza, 2012. [ISBN 978-84-15538158].

[42] B. Lacruz, F.J. López, P. Mateo, C. Paroissin, A. Pérez-Palomares, and G. Sanz, editors. *Pyrenees International Workshop on Statistics, Probability and Operations Resaearch (SPO'07)*, volume 34 of *Monografías del Seminario Matemático "García De Galdeano"*. Prensas Universitarias de Zaragoza, 2008. [ISBN 978-84-92521-18-0].

Prépublications

[43] L. Bordes, C. Paroissin, and A. Salami. Combining gamma and brownian processes for degradation modeling in presence of explanatory variables. 2010.

[44] N. Bru, L. Despres, and C. Paroissin. A comparison of statistical models for short categorical or ordinal time series with applications in ecology. 2007.

[45] C. Paroissin. Inference for the wiener process with random initiation time. 2015.

Thèse

[46] C. Paroissin. *Résultats asymptotiques pour des grands systèmes réparables monotones*. PhD thesis, Université Paris 7 Denis Diderot, 2002.

Autres références

[47] M. Abdel-Hameed. A gamma wear process. *IEEE Trans. Reliab.*, 24(2) :152–153, 1975.

[48] M. Abramowitz and I.A. Stegun, editors. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Number 55 in Applied Mathematics Series. National Bureau of Standards„ 1972.

[49] B. Allen and I. Munro. Self-organizing binary search trees. *J. Assoc. Comput. Mach.*, 25 :526–535, 1978.

[50] L.U. Ancarani and G. Gasaneo. Derivatives of any order of the gaussian hypergeometric function ${}_2f_1(a, b, c; z)$ with respect to the parameters a, b and c. *J. Phys. A : Math. Theor.*, 42 :1–10, 2009.

[51] P. K. Andersen and N. Keiding. Multi-state models for event history analysis. *Stat. Methods Med. Res.*, 11 :91–115, 2002.

[52] B. C. Arnold, N. Balakrishnan, and A. N. Nagaraja. *A first course in order statistics*. SIAM, Philadelphia, 2008.

[53] S. Asmussen and H. Albrecher. *Ruin Probabilities*. World Scientific, New Jersey, 2010.

[54] S. Asmussen, O. Nerman, and M. Olsson. Fitting phase-type distributions via the em algorithm. *Scand. J. Stat.*, 23(4) :419–441, 1996.

[55] V. Bagdonavičius and M. S. Nikulin. Estimation in degradation models with explanatory variables. *Lifetime Data Anal.*, 7(1) :85–103, 2001.

[56] N. Balakrishnan and H.K.T. Ng. *Precedence-type tests and applications*. Wiley, New York, 2006.

[57] N. Balakrishnan, I.S. Triantafyllou, and M.V. Koutras. Nonparametric control charts based on runs and wilcoxon-type rank-sum statistics. *J. Stat. Plan. Infer.*, 139 :3177–3192, 2009.

[58] C.T. Barker and M.J. Newby. Optimal non-periodic inspection for a multivariate degradation model. *Reliab. Eng. Syst. Safety*, 94 :33–43, 2009.

[59] R.E. Barlow and F. Proschan. *Mathematical reliability theory*. Wiley, New-York, 1965.

[60] J. Benichou and M.H. Gail. A delta method for implicitly defined random variables. *Am. Stat.*, 43(1) :41–44, 1989.

[61] C. Bérenguer, A. Grall, L. Dieulle, and M. Roussignol. Maintenance policy for a continuously monitored deteriorating system. *Probab. Engrg. Inform. Sci.*, 17(2) :235–250, 2003.

[62] P.K. Bhattacharya and R. Johnson. Nonparametric tests for shifts at an unknown time point. *Ann. Math. Stat.*, 39 :1731–1734, 1968.

[63] E. Biffis and Kyprianou A.E. A note on scale functions and the time value of ruin for lévy insurance risk. *Insur. Math. Econ.*, 46(1) :85–91, 2010.

[64] A. Billon. *Modélisation de la fiabilité de composants d'un moteur aéronautique basée sur les données des dégradations en fonction de la maintenance programmée*. PhD thesis, Université de Pau et des Pays de l'Adour, 2012.

[65] Z.W. Birnbaum. On the importance in a multicomponent system. In P.R. Krishnaiah, editor, *Multivariate Analysis II (1968)*, pages 581–592. Academic Press, 1969.

[66] C.R. Blyth. On the inference and decision models of statistics. *Ann. Math. Stat.*, 41(3) :1034–1058, 1970.

[67] L. Bordes and S. Mercier. Extended geometric processes : semiparametric estimation and application to reliability. *Journal of the Iranian Royal Statistical Society*, 12(1) :1–34, 2012.

[68] R.F. Botta and C.M. Harris. Approximation with generalized hyperexponential distributions : weak convergence results. *Queueing Syst.*, 1 :169–190, 1986.

[69] J. Bourgain, J. Kahn, G. Kalai, Y. Katznelson, and N. Linial. The influence of variables in product spaces. *Israel J. Math.*, 77 :55–64, 1992.

[70] A.M. Bruckner and E. Ostrow. Some function classes related to the class of convex functions. *Pacific J. Math.*, 12(4) :1203–1215, 1962.

[71] V. Capasso and S.L. Paveri-Fontana. A mathematical model for the 1973 cholera epidemic in the european mediterranean region. *Rev. Epidém. et Santé Pub.*, 27 :121–132, 1979.

[72] Ph. Capéraà and B. Van Cutsem. *Méthodes et modèles en statistique non paramétrique : exposé fondamental*. Presses de l'Université de Laval - Dunod, 1988.

[73] Z. Chen. Component reliability analysis of a *k*-out-of-*n* systems with censored data. *J. Stat. Plan. Infer.*, 116(1) :305–315, 2003.

[74] H. Chernoff and S. Zacks. Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Statist.*, 35 :999–1018, 1964.

[75] R.S. Chhikara and L. Folks. *The inverse Gaussian distribution : theory, methodology and applications*. Marcel Dekker, 1989.

[76] B.S Clarke. Implications of reference priors for prior information and for sample size. *J. Am. Stat. Assoc.*, 91 :173–184, 1996.

[77] C. Cocozza-Thivent. *Processus stochastiques et fiabilité des systèmes*. Springer, Paris, 1997.

[78] C.T. Codeço. Endemic and epidemic cholera : the role of the aquatic reservoir. *BMC Infect. Dis.*, 1(1), 2001.

[79] D. Commenges. Multi-state models in epidemiology. *Lifetime Data Anal.*, 5 :315–327, 1999.

[80] M. Csörgö and L. Horváth. *Limit theorems in change-point analysis*. Wiley, New York, 1997.

[81] E. de Souza e Silva and H.R. Gail. Calculating cumulative operational time distributions of repairable computer systems. *IEEE Trans. Computers*, C-35(4) :322–332, 1986.

[82] R.P. Dobrow and J.A. Fill. On the Markov chain for the move-to-root rule for binary search trees. *Ann. Appl. Probab.*, 5(1) :1–19, 1995.

[83] R.P. Dobrow and J.A. Fill. Rates of convergence for the move-to-root Markov chain for binary search trees. *Ann. Appl. Probab.*, 5(1) :20–36, 1995.

[84] R.A. Doney. Some excursion calculations for spectrally one-sided lévy processes. In M. Emery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXVIII*, pages 5–15. Springer, 2005.

[85] P. Donnelly. The heaps process, libraries and size-biased permutations. *J. Appl. Prob.*, 28 :321–335, 1991.

[86] P. Doukhan, G. Lang, S. Louhichi, and B. Ycart. A functional central limit theorem for interacting particle systems on transitive graphs. *Markov Proc. Rel. Fields*, 14 :79–114, 2008.

[87] M. Egami and K. Yamazaki. On scale functions of spectrally negative lévy processes with phase-type jumps. `arXiv:1005.0064v3`, 2010.

[88] A. Feldmann and W. Whitt. Fitting mixtures of exponentials to long-tail distributions to analyze network performance models. *IEEE Infocom'97*, pages 1098–1106, 1997.

[89] W. Feller. *An Introduction to probability theory and its applications. Volume 2*. Wiley, Hoboken, New Jersey, 1971.

[90] J.A. Fill. Limits and rates of convergence for the distribution of search cost under the move-to-front rule. *Theor. Comput. Sci.*, 164 :185–206, 1996.

[91] J.A. Fill and L. Holst. On the distribution of search cost for the move-to-front rule. *Random Struct. Algor.*, 8 :179–186, 1996.

[92] R.A. Fisher, A.S. Corbet, and C.B. Williams. The relation between the number of species and the number of individuals in a random sample of animal population. *J. Anim. Ecol.*, 12 :42–58, 1943.

[93] J.B. Frenk and R.P. Nicolai. Approximating the randomized hitting time distribution of a non-stationary gamma process. ERIM Report Series Reference No. ERS-2007-031-LIS, 2007.

[94] E. Friedgut. Influences on product spaces, KKL and BKKKL revisited. *Comb. Probab. Comput.*, 13(1) :17–29, 2004.

[95] J. Garrido and M. Morales. On the expected discounted penalty function for lévy risk processes. *North American Actuarial Journal*, 10(4) :196–218, 2006.

[96] J.D. Gibbons and S. Chakraborti. *Nonparametric statistical inference*. Marcel Dekker Inc., New York, 2003.

[97] A.G. Glen, L.M. Leemis, and J.H. Drew. Computing the distribution of the product of two continuous random variables. *Comput. Stat. Data An.*, 44 :451–464, 2004.

[98] I.S. Gradshteyn and I.M. Ryzhik. *Table of integrals, series and products*. Academic Press, 1965.

[99] A. Grall, L. Dieulle, C. Bérenguer, and M. Roussignol. Continuous-time predictive-maintenance scheduling for a deteriorating system. *IEEE Trans. Reliab.*, 51(2) :141–150, 2002.

[100] F. Guo, H. Rakha, and S. Park. Multistate model for travel time reliability. *Transport. Res. Rec.*, 2128 :46–54, 2010.

[101] G.H. Guo, A. Gerokostopoulos, H. Liao, and N. Pengying. Modeling and analysis for degradation with an initiation time. In *Reliability and Maintainability Symposium (RAMS)*, pages 1–6, 2013.

[102] M.G. Hahn. Central limit theorem in $D[0, 1]$. *Z. Wahrsch. Verw. Geb.*, 44 :89–101, 1978.

[103] W.J. Hendricks. The stationary distribution of an interesting Markov chain. *J. Appl. Prob.*, 9 :231–233, 1972.

[104] P. Hougaard. Multi-state models: A review. *Lifetime Data Anal.*, 5 :239–264, 1999.

[105] A. Hoyland and M. Rausand. *System reliability theory: models and statistical methods*. Wiley, New-York, 1994.

[106] F. Hubalek and A.E. Kyprianou. Old and new examples of scale functions for spectrally negative lévy processes. *Progress in Probability*, 63 :119–145, 2010.

[107] T. Huillet. Sampling formulae arising from random Dirichlet populations. *Commun. Statist. - Theor. Meth.*, 34(5) :1019–1040, 2005.

[108] L.F. James. Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhyä*, 65 :179–206, 2003.

[109] D.H. Janzen. Sweep samples of tropical foliage insects : description of study sites, with data on species abundances and size distributions. *Ecology*, 54 :659–686, 1973.

[110] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions. Volume 1.* Wiley, New York, 1995.

[111] N. L. Johnson, S. Kotz, and N. Balakrishnan. *Continuous univariate distributions. Volume 2.* Wiley, New York, 1995.

[112] M.J. Kallen. A comparison of statistical models for visual inspection data. In *H. Furuta, D.M. Frangopol and M. Shinozuka (Eds).* Safety, Reliability and Risk of Structures, Infrastructures and Engineering Systems, Proceedings of the Tenth International Conference on Structural Safety and Reliability (ICOSSAR'2009), pages 3235–3242. Taylor & Francis, London, 2009.

[113] M.J. Kallen and J.M. van Noortwijk. Statistical inference for Markov deterioration models of bridge conditions in the Netherlands. In *D.M. Frangopol P.J.S. Cruz and L.C. Neves (Eds),* Proceedings of the Third International Conference on Bridge Maintenance, Safety and Management (IABMAS), pages 16–19. Taylor & Francis Group, 2006.

[114] R.E. Kass and L. Wasserman. The selection of prior distributions by formal rules. *J. Am. Stat. Assoc.*, 91 :1343–1370, 1996.

[115] R. Keener, E. Rothman, and N. Starr. Distributions on partitions. *Ann. Stat.*, 15(4) :1466–1481, 1987.

[116] J.F.C. Kingman. Random discrete distributions. *J. R. Stat. Soc., Ser. B*, 37 :1–22, 1975.

[117] J.F.C. Kingman. *Poisson processes*. Oxford University Press Inc., New York, 1993.

[118] D.E. Knuth. *The Art of Computer Programming. Volume 3 : Sorting and Searching*. Addison-Wesley Publishing Co., Reading, 1973.

[119] Y. Kovchegov, N. Meredith, and E. Nir. Occupation times and bessel densities. *Stat. Probab. Letters*, 80 :104–110, 2010.

[120] A.E. Kyprianou. *Introductory lectures on fluctuations of Lévy processes with applications*. Springer, 2006.

[121] A.E. Kyprianou and Z. Palmowski. A martingale review of some fluctuation theory for spectrally negative lévy processes. In M. Emery, M. Ledoux, and M. Yor, editors, *Séminaire de Probabilités XXXVIII*, pages 16–29. Springer, 2005.

[122] A.E. Kyprianou, J.C. Pardo, and V. Rivero. Exact and asymptotic *n*-tuple laws at first and last passage. *Ann. Appl. Probab.*, 20(2) :522–564, 2010.

[123] B. Lachaud. *Détection de la convergence de processus de Markov*. PhD thesis, Université de Paris 5 - René Descartes, 2005.

[124] Y. Lam. *The geometric process and its applications*. World Scientific, Singapore, 2007.

[125] J. Lawless and M. Crowder. Covariates and random effects in a gamma process model with application to degradation and failure. *Lifetime Data Anal.*, 10 :213–227, 2004.

[126] A. Lijoi and I. Prünster. A note on the problem of heaps. *Sankhya*, 66 :232–240, 2004.

[127] J.K. Lindsey. *Statistical analysis of stochastic processes in time*. Cambridge University Press, Cambridge, UK, 2004.

[128] F. Lombard. Rank tests for change-point problems. *Biometrika*, 74(3) :615–624, 1987.

[129] M. Malhotra and A. Reibman. Selecting and implementing phase approximations for semi-markov models. *Commun. Statist. - Stochastic Models*, 9(4) :473–506, 1993.

[130] R.H. McArthur. On the relative abundance of bird species. *Proc. Nat. Acad. Sci. U.S.A.*, 43 :293–295, 1957.

[131] J. McCabe. On serial files with relocatable records. *Operations Res.*, 13 :609–618, 1965.

[132] M. Mezzetti and J.G. Ibrahim. Bayesian inference for the cox model using correlated gamma process priors. Technical report, Deparment of Biostatistics, Harvard School of Public Health, 200.

[133] M. Mitra and S.K. Basu. On some properties of the bathtub failure rate family of life distributions. *Microelectron. Reliab.*, 36(5) :679–684, 1996.

[134] A.M. Mood. *Introduction to the theory of statistics*. McGraw-Hill, 1974.

[135] F. Mosteller and D.L. Wallace. *Applied Bayesian and classical inference : the case of the Federalist papers*. Springer-Verlag, New-York, 1984.

[136] T. Nakagawa. *Maintenance theory of reliability*. Springer-Verlag, London, 2005.

[137] A. Narayanan. Small sample properties of parameter estimation in the Dirichlet distribution. *Comm. Statist. - Simulation Comput.*, 20(2-3) :647–666, 1991.

[138] B. Natvig and H.W. Mørch. An application of multistate reliability theory to an offshore gas pipeline network. *International Journal of Reliability, Quality and Safety Engineering*, 10 :361–381, 2003.

[139] W.B. Nelson. Defect initiation, growth, and failure - a general statistical model and data analyses. In M.S. Nikulin, N. Limnios, N. Balakrishnan, W. Kahle, and C. Huber-Carol, editors, *Advances in degradation modeling. Applications to reliability, survival analysis, and finance*. Birkhäuser-Basel, 2010.

[140] M.F. Neuts. *Matrix-geometric solutions in stochastic models: an algorithmic approach*. Dover Publisher, New-York, 1994.

[141] L.E. Nieto-Barajas and S.G. Walker. Markov beta and gamma processes for modelling hazard rates. *Scand. J. Stat.*, 29 :413–424, 2002.

[142] E.S. Page. Continuous inspection scheme. *Biometrika*, 41 :100–114, 1954.

[143] E.S. Page. A test for a change in a parameter occurring at an unknown point. *Biometrika*, 42 :523–527, 1955.

[144] E.S. Page. On problems in which a change in parameters occurs at an unknown point. *Biometrika*, 44 :248–252, 1957.

[145] C. Park and W.J. Padgett. Accelerated degradation models for failure based on geometric Brownian motion and gamma process. *Lifetime Data Anal.*, 11 :511–527, 2005.

[146] W.C. Parr and W.R. Schucany. Minimum distance and robust estimation. *J. Am. Stat. Assoc.*, 75(371) :616–624, 1980.

[147] M. Pascual, X. Rodó, S.E. Ellner, R. Colwell, and M.J. Bouma. Cholera dynamics and el niño-southern oscillation. *Science*, 289 :1766–1769, 2000.

[148] P.J. Pedler. Occupation times for two-state markov chains. *J. Appl. Prob.*, 8 :381–390, 1971.

[149] H. Peng and Q. Feng. Reliability modeling for ultrathin gate oxides subject to logistic degradation processes with random onset time. *Quality and Reliability Engineering International*, 29 :709–718, 2013.

[150] V.V. Petrov. *Limit theorems of probability theory - sums of independent random variables*. Oxford University Press Inc., New York, 1995.

[151] A.N. Pettitt. A non-parametric approach to the change-point problem. *Appl. Statist.*, 28(2) :126–135, 1979.

[152] A.N. Pettitt. Some results on estimating a change-point using non-parametric type statistics. *Appl. Statist.*, 11 :261–272, 1980.

[153] H. Pham, A. Suprasad, and R.B. Misra. Reliability analysis of k-out-of-n systems with partially repairable multi-state components. *Microelectron. Reliab.*, 36(10) :1407–1415, 1996.

[154] E.C. Piélou. *Ecological diversity*. John Wiley, New-York, 1975.

[155] J. Pitman and M. Yor. Random discrete distributions derived from self-similar random sets. *Electron. J. Probab.*, 1(4) :1–28, 1996.

[156] D. Pollard. *Convergence of stochastic processes*. Springer, New-York, 1984.

[157] E. Pourabbas, A. d'Onofrio, and M. Rafanelli. A method to estimate the incidence of communicable diseases under seasonal fluctuations with application to cholera. *Appl. Math. Comput.*, 118(2-3) :161–174, 2001.

[158] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5 :375–383, 1963.

[159] E. Regazzini, A. Lijoi, and I. Prünster. Distributional results for means of normalized random measures of independent increments. *Ann. Statist.*, 31 :560–585, 2003.

[160] R.D. Reiss. *Approximate distributions of order statistics, with application to non-parametric statistics*. Springer, New-York, 1989.

[161] A. Rényi. On the central limit theorem for the sum of a random number of independent random variables. *Acta Math. Acad. Sci. H.*, 11(1-2) :97–102, 1960.

[162] C.P. Robert. *The Bayesian Choice : A Decision Theoretic Motivation*. Springer-Verlag, New York, 2004.

[163] C.P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, 2004.

[164] V.K. Rohatgi. *An introduction to probability theory mathematical statistics*. Wiley, New-York, 1976.

[165] G. Ronning. Maximum likelihood estimation of a dirichlet distribution. *Comm. Statist. - Simulation Comput.*, 32 :215–221, 1989.

[166] P. van der Laan S. Chakraborti and M.A. van de Wiel. A class of distribution-free control charts. *Appl. Statist.*, 53 :443–462, 2004.

[167] B. Saassouh, L. Dieulle, and A. Grall. Online maintenance policy for a deteriorating system with random change of mode. *Reliab. Eng. Syst. Safety*, 92 :1677–1685, 2007.

[168] A. Salami. *Inférence statistique pour un modèle de dégradation en présence de variables explicatives*. PhD thesis, Université de Pau et des Pays de l'Adour, 2011.

[169] A. Sen and M.S. Srivastava. On tests for detecting change in mean. *Ann. Statist.*, 3(1) :98–108, 1975.

[170] B. Sericola. Interval-availability distribution of 2-states systems. *IEEE Trans. Reliab.*, 43(2) :335–343, 1994.

[171] B. Sericola. Occupation times in Markov processes. *Comm. Statist. - Stoch. Models*, 16(5) :479–510, 2000.

[172] M. Shaked and J.G. Shanthikumar. On the first-passage times of pure jump processes. *J. Appl. Probab.*, 25(3) :501–509, 1988.

[173] E.H. Simpson. Measurement of diversity. *Nature*, 163 :688, 1949.

[174] C. Singh, R. Billinton, and S.Y. Lee. The method of stages for non-markov models. *IEEE Trans. Reliab.*, 26(2) :135–137, 1977.

[175] N.D. Singpurwalla. Survival in dynamic environments. *Statistical Science*, 10(1) :96–103, 1995.

[176] D. Korkinof S.P. Chatzis and Y. Demiris. A spatially-constrained normalized gamma process for data clustering. *Artificial Intelligence Applications and Innovations. IFIP Advances in Information and Communication Technology*, 381 :337–346, 2012.

[177] M.D. Springer. *The algebra of random variables*. Wiley, New-York, 1979.

[178] D. Stoyan. *Comparison methods for queues and other stochastic models*. Wiley, Chichester, 1983.

[179] A.E. Kyprianou T. Chan and M. Savov. Smoothness of scale functions for spectrally negative lévy processes. *Probab. Theory Rel.*, 150 :691–708, 2010.

[180] S. Tavaré. Ancestral inference in population genetics. In J. Picard, editor, *Lectures on Probability Theory and Statistics (École d'été de Saint-Flour XXXI, 2001)*, volume 1837 of *Lectures Notes in Mathematics*. Springer, Berlin, 2004.

[181] C.-C. Tsai, S.-T. Tseng, and N. Balakrishnan. Mis-specification analyses of gamma and Wiener degradation processes. *J. Stat. Plan. Infer.*, 141 :3725–3735, 2011.

[182] C.C.L. Tsai and G.E. Willmot. A generalized defective renewal equation for the surplus process perturbed by diffusion. *Insur. Math. Econ.*, 30 :51–66, 2002.

[183] M.L. Tsetlin. Finite automata and models of simple forms of behavior. *Russian Math. Surveys*, 18(4) :1–27, 1963.

[184] J.M. van Noortwijk. A survey of the application of gamma processes in maintenance. *Reliab. Eng. Syst. Safety*, 94 :2–21, 2009.

[185] X. Wang. A pseudo-likelihood estimation method for nonhomogeneous gamma process model with random effects. *Stat. Sinica*, 18 :1153–1163, 2008.

[186] X. Wang. Nonparametric estimation of the shape function in a gamma process for degradation data. *Can. J. Stat.*, 37(1) :102–118, 2009.

[187] X. Wang. Semiparametric inference on a class of Wiener processes. *J. Time Ser. Anal.*, 30(2) :179–207, 2009.

[188] X. Wang. Wiener processes with random effects for degradation data. *J. Multivariate Anal.*, 101(2) :340–351, 2010.

[189] W. Whitt. Some useful functions for functionnal limit theorems. *Math. Oper. Res.*, 5(1) :67–85, 1980.

[190] S.S. Wilks. *Mathematical statistics*. Wiley, New-York, 1962.

[191] J. Pittman X. Lin and B. Clarke. Information conversion, effective samples, and parameter size. *IEEE T. Inform. Theory*, 53 :4438–4456, 2007.

[192] R. Yang and J.O. Berger. A catalog of non-informative priors. Duke University Research Report, 1998.

[193] Y. Zhang and H. Liao. Analysis of destructive degradation tests for a product with random degradation initiation time. *To appear in* IEEE Transaction on Reliability, 2014.

[194] X. Zhao, M. Fouladirad, and C. Bérenguer. Residual based inspection/replacement policy for a deteriorating system with markovian covariates. In *Proceeding of international conference on industrial engineering and engineering management, 2010 December 7-12, Macau*, pages 636–64, 2010.

[195] X. Zhao, M. Fouladirad, C. Bérenguer, and L. Bordes. Condition-based inspection/replacement policies for nonmonotone deteriorating systems with environmental covariates. *Reliab. Eng. Syst. Safety*, 95(8) :921–934, 2010.

Résumé

La première partie de ce mémoire porte sur l'étude de modèles en fiabilité, essentiellement des modèles de dégradation pour un composant ou un système, mais également des modèles de durées de vie. On s'intéresse tant à l'étude des propriétés probabilistes de ces modèles qu'à des problèmes d'estimation et qu'à l'optimisation d'une politique de remplacement ou de maintenance.

Au chapitre 2, différents modèles de dégradation multi-états sont étudiés. Il s'agit de situations où la détérioration se mesure sur une échelle qualitative. Le premier modèle étudié est un système monotone et redondant constitué de composants indépendants, identiques et réparables, pour lequel on étudie le comportement asymptotique du temps d'atteinte d'un seuil critique. Les deux autres modèles ont été développés dans le cadre de collaborations industriels. Dans le premier cas, on a considéré un composant dont la dégradation est décrite par une chaîne de Markov non-homogène, le caractère non-homogène étant lié au fait que les probabilités de transition peuvent dépendre de covariables qui évoluent dans le temps. Dans le second cas, on a modélisé la dégradation d'un composant soumis à plusieurs modes d'endommagement dans l'objectif de réduire éventuellement la périodicité des maintenances tout en garantissant un niveau acceptable de sécurité.

Au chapitre 3, on considère différents modèles de détérioration continue (la dégradation est mesurée quantitativement). On a d'abord étudié la loi du temps de panne pour un processus gamma non-homogène, défini comme le temps de franchissement d'un seuil critique (déterministe ou aléatoire). On a ensuite étudié le processus gamma (homogène) perturbé par un mouvement brownien. Pour ce modèle (sans ou avec covariables), on s'est intéressé à l'inférence statistique. Pour une classe de modèles plus large, on a ensuite déterminé la loi du temps de panne. Puis, on a étudié le cas d'un processus gamma modulé par un processus markovien (ce dernier représente l'environnement dynamique dans lequel évolue le composant dont on modélise la dégradation). Enfin, on s'est intéressé à l'inférence statistique pour un processus de Wiener avec une période d'initiation aléatoire (période pendant laquelle le composant ne se dégrade pas).

Aux chapitres 4 et 5, on étudie des problèmes connexes. D'abord, on s'intéresse à la prise en compte des incertitudes pour un modèle de dégradation ou pour un modèle de durée de vie. Ensuite, des problèmes de comparaison de deux échantillons de durée de vie sont étudiés, l'un avec sous l'angle de la détection de rupture, l'autre sous l'angle des cartes de contrôles. Le chapitre 6 présente quelques perspectives liés aux travaux précédemment menés.

La deuxième partie de ce mémoire est articulé autour des lois discrètes aléatoires (c'est-à-dire l'ensemble des lois de probabilités à valeurs dans un simplexe). Ce type de lois, en particulier la loi de Dirichlet, trouvent des applications dans divers domaines. Au chapitre 8, on les a utilisé pour étudier le coût de recherche asymptotique pour deux heuristiques auto-organisatrices pour lesquels les popularités des objets sont inconnues et aléatoires. Le chapitre 9 est consacré à des problèmes d'inférence statistique pour la loi de Dirichlet, avec des applications en écologie (estimation du nombre d'espèces à partir d'un échantillon).

Enfin, la troisième et dernière partie de ce mémoire est consacrée à un aperçu de travaux plus appliquées, en collaboration avec des biologistes, des médecins, etc.

Summary

The first part of this dissertation is devoted to the study of several models in dependability, essentially about degradation analysis for a component or a system, but also on lifetimes problems. We are both interested in probabilistic properties and statistical inference of these models, and in the optimization of some maintenance/replacement policy.

In Chapter 2, different multi-states models are studied. In such case, the deterioration is measured accordingly to a qualitative scale. The first model is about a monotone and redundant system made of independent, identical and repairable components, for which we study the asymptotic behaviour of first-passage time of a critical threshold. The next two models has been developed within industrial collaborations. In the first case, we have considered a component whose degradation is described by a non-homogeneous Markov chain (the transition probabilities are functions of time-dependent covariates). For the second case, we have studied a model of component subject to several types of degradations, the aim being here to reduce eventually the periodicity of the maintenance, but keeping an acceptable level of security.

In Chapter 3, we consider different continuous degradation models (here, the deterioration is measured quantitatively). First, we study the distribution of the time-to-failure for a non-homogeneous gamma process (defined as the hitting time of a deterministic or random critical threshold). Then, we consider an homogeneous gamma process perturbed by an independent Brownian motion, including eventually static covariates. For such kind of models, we propose estimation procedures for which we provide asymptotic results. For a wider class of models (the gamma process is replaced by any subordinator), we have studied the distribution of the time-to-failure. Next, we look the case of a gamma process modulated by a Markovian jump process (which represents the dynamical environments of the component). At least, we consider the statistical inference for the Wiener process with a random initiation period (the degradation of the component is assumed to start after a random delay).

In Chapters 4 and 5, we consider some related problems as those studied above. First, we are taking into account uncertainty for a degradation model or for a lifetime model. Then, two-sample comparison problems are studied by considering it as a change-point detection problem or by designing some control charts. Chapter 6 provides some future works related to those mentioned here.

The second part of this dissertation is guided around random discrete distributions (i.e. the set of all distributions taking values on a simplex). This kind of distributions, in particular the Dirichlet distribution, find applications in many different areas. In Chapter 8, we use it to study the asymptotic search cost for two self-organizing heuristics for which the popularities are unknown and random. Chapter 9 is devoted to some statistical inference problems for the Dirichlet distribution, with applications in ecology (estimation of the number of species from a sample).

At least, the third and last part of this dissertation gives an overview of some more applied studies, in collaboration with biologists, medical doctors, etc.