

Université Joseph Fourier - Institut National Polytechnique de Grenoble

Master Recherche en Mathématiques Appliquées

Fiabilité des Systèmes Réparables

Notes de cours

Olivier Gaudoin

Table des matières

1	Introduction	5
2	Modélisation stochastique du processus des défaillances d'un système réparable	9
2.1	Le processus de comptage des défaillances	9
2.2	Propriétés des processus ponctuels auto-excités	12
2.2.1	Loi du nombre de défaillances survenues	12
2.2.2	Loi des instants de défaillance	15
2.2.3	Loi des durées inter-défaillances	17
2.2.4	Fiabilité et MTTF	18
2.2.5	Fonctions de vraisemblance	19
2.3	Classification des modèles de fiabilité des systèmes réparables	20
2.3.1	Les processus de Poisson non homogènes (NHPP)	20
2.3.2	Les processus de renouvellement	21
2.3.3	Les modèles de réparation imparfaite	22
2.3.4	Les modèles à Durées Inter-Défaillances Exponentielles (DIDE)	23
3	Propriétés générales des NHPP	25
3.1	Introduction	25
3.2	Processus de comptage des défaillances	25
3.3	Lois des instants de défaillance et des durées inter-défaillances	27
3.4	Fiabilité et MTTF	30
3.5	Fonctions de vraisemblance	30
3.6	Quelques propriétés utiles	31
4	Les processus de Poisson homogènes	35
4.1	Propriétés probabilistes	35
4.1.1	Loi du nombre de défaillances survenues à chaque instant	35
4.1.2	Lois des instants de défaillances	36
4.1.3	Lois des durées inter-défaillances	36
4.1.4	Fiabilité et MTTF	36
4.1.5	Fonctions de vraisemblance	37
4.1.6	Quelques propriétés utiles	37
4.2	Lien entre processus de Poisson homogène et non homogène - Simulation des NHPP	38
4.3	Estimation des grandeurs de la fiabilité	39
4.3.1	Censure de type 2	39

4.3.2	Censure de type 1	43
4.4	Tests d'adéquation	45
4.4.1	Problématique	45
4.4.2	Censure de type 2	47
4.4.3	Censure de type 1	50
5	Le processus de puissance	53
5.1	Propriétés probabilistes	53
5.1.1	Loi du nombre de défaillances survenues à chaque instant	54
5.1.2	Lois des instants de défaillances	54
5.1.3	Lois des durées inter-défaillances	55
5.1.4	Fiabilité et MTTF	55
5.1.5	Fonctions de vraisemblance	55
5.1.6	Quelques propriétés utiles	56
5.2	Estimation des grandeurs de la fiabilité	57
5.2.1	Censure de type 2	57
5.2.2	Censure de type 1	60
5.3	Tests d'adéquation	61
5.4	Application	62
5.4.1	Jeu de données 1	62
5.4.2	Jeu de données 2	62
	Bibliographie	63

Chapitre 1

Introduction

De nombreux phénomènes aléatoires se présentent sous la forme d'occurrences d'évènements ponctuels successifs dans le temps. C'est le cas par exemple des arrivées de requêtes sur un serveur, de sinistres dans une compagnie d'assurances, de clients dans un magasin, de l'occurrence de tremblements de terre ou de cracks boursiers, de l'émission de particules par un système radioactif, etc...

Nous nous intéressons ici à la suite des défaillances (ou pannes) de systèmes réparables. Les tableaux 1.1 et 1.2 donnent deux exemples de données de ce type.

3	120	227	21	529	860	108	22
30	26	65	233	379	983	1	75
113	114	476	134	44	707	3109	482
81	325	58	357	129	33	1247	5509
115	55	457	193	810	868	943	100
9	242	300	236	290	724	700	10
2	68	97	31	300	2323	875	1071
91	422	263	369	529	2930	245	371
112	180	452	748	281	1461	729	790
15	10	255	1	160	843	4897	6150
138	1146	197	231	828	12	447	3321
50	600	193	330	1011	261	386	1045
77	15	6	365	445	1800	446	648
24	36	79	1222	296	865	122	5485
108	4	816	543	1755	1435	990	1160
88	1	1351	10	1064	30	948	1864
670	7	148	16	1783	143	1082	4116

TAB. 1.1 – *Jeu de données 1 : durées inter-défaillances successives d'un logiciel*

18 12 52 31 8 5

TAB. 1.2 – *Jeu de données 2 : nombre de mois entre défaillances successives d'une pompe*

Le jeu de données 1 est constitué de 136 durées entre défaillances successives (en secondes CPU) d'un logiciel temps réel de contrôle-commande de l'armée américaine. A chaque défaillance, le logiciel est corrigé, puis relancé. Le tableau est à lire en colonnes.

Le jeu de données 2 est constitué des nombres de mois entre les 6 premières défaillances d'une pompe. A chaque défaillance, la pompe est réparée (on dit aussi que l'on fait une maintenance corrective) puis remise en fonctionnement. On dispose en plus de l'information selon laquelle il n'y a pas eu de défaillance dans les 3 mois suivant la dernière panne.

Dans les deux cas, les systèmes n'ont pas fonctionné pendant qu'on les corrigeait ou réparait. Les durées de réparation ne sont pas données, donc on ne les prendra pas en compte ici. Cependant, il peut arriver que l'information dont on dispose soit une suite alternée de durées de bon fonctionnement et de réparations successives. On dispose de méthodes pour traiter ce type de données, mais elles ne relèvent pas de ce cours.

On constate que, manifestement, les durées entre les défaillances successives du logiciel ont tendance à augmenter, ce qui signifie que le logiciel est de plus en plus fiable. C'est logique car les corrections apportées au logiciel ont pour but de supprimer des erreurs (bugs) dans le programme et doivent donc augmenter sa fiabilité. Inversement, la pompe a tendance à tomber de plus en plus souvent en panne : elle s'use, vieillit, et est de moins en moins fiable.

Au vu de ces données, on peut se poser un certain nombre de questions, essentiellement de deux types :

- **Prévision.** Que va-t-il se passer dans le futur ? Quand en moyenne se produira la prochaine défaillance ? Combien de défaillances peut-on s'attendre à observer la prochaine semaine ou la prochaine année ? Ces évaluations sont importantes en particulier pour décider d'une durée de garantie ou mettre en place une politique de maintenance adaptée.
- **Décision.** Peut-on certifier que la fiabilité du logiciel a augmenté ? Que celle de la pompe a diminué ? Le logiciel est-il devenu suffisamment fiable pour pouvoir être mis en vente ? La pompe est-elle devenue suffisamment peu fiable pour devoir être remplacée par une neuve ?

Pour répondre à ces questions, il faut procéder en deux étapes :

- **Première étape probabiliste.** Il faut proposer une modélisation stochastique (aléatoire) du processus ayant engendré ces observations. Autrement dit, il faut proposer un modèle probabiliste pour le processus des durées inter-défaillances successives. Ayant proposé un modèle, on peut calculer toutes les quantités intéressantes comme l'espérance mathématique de la durée d'attente de la prochaine défaillance ou la probabilité que cette prochaine défaillance survienne après une durée donnée.
- **Deuxième étape statistique.** Les modèles issus de la première étape ont des paramètres qu'il faut estimer. Ces estimations permettront de donner une valeur numérique aux quantités calculées plus haut (espérance, probabilité, ...). On peut aussi adopter une démarche non-paramétrique en estimant les quantités en question sans utiliser de modèle probabiliste particulier pour les observations. Enfin, la réponse aux problèmes de décision cités plus haut se fait à l'aide de tests d'hypothèses statistiques.

Le plan du cours suit ces deux étapes. On commence par proposer une **modélisation stochastique** très générale du phénomène étudié, en utilisant le cadre des **processus aléatoires ponctuels auto-excités**. Puis nous étudierons dans le détail une famille particulière de modèles, les **processus de Poisson non homogènes**, en abordant à la fois les aspects probabilistes et statistiques. En particulier, on s'intéressera à deux types de tests d'hypothèses, les **tests de tendance**, dont le but est de déterminer le sens de variation de la fiabilité, et les **tests d'adéquation**, dont le but est de choisir un modèle adapté à un jeu de données particulier.

Chapitre 2

Modélisation stochastique du processus des défaillances d'un système réparable

2.1 Le processus de comptage des défaillances

On s'intéresse au fonctionnement d'un système à partir d'un instant $T_0 = 0$. Des défaillances se produisent à des instants que l'on note $\{t_i\}_{i \geq 1}$. Après chaque défaillance, le système est réparé ou non, puis remis en fonctionnement. Les durées de réparation sont considérées comme négligeables ou ne sont pas comptabilisées.

Le **processus des défaillances** est alors défini de manière équivalente par l'un des 3 processus aléatoires suivants :

- $\{T_i\}_{i \geq 1}$, suite des instants de défaillance du système, avec $T_0 = 0$.
- $\{X_i\}_{i \geq 1}$, suite des durées inter-défaillances successives où $\forall i \geq 1, X_i = T_i - T_{i-1}$ est la durée entre la $(i-1)^{\text{ème}}$ et la $i^{\text{ème}}$ défaillance.
- $\{N_t\}_{t \geq 0}$, processus de comptage des défaillances, où $\forall t \in \mathbb{R}^+, N_t = \sum_{i=1}^{+\infty} \mathbb{1}_{\{T_i \leq t\}} = \max\{i \in \mathbb{N}; T_i \leq t\}$ est le nombre cumulé de défaillances survenues entre 0 et t .

$\{N_t\}_{t \geq 0}$ est un **processus ponctuel de comptage**, ce qui signifie que [2]:

- $N_0 = 0$ p.s.
- $\{N_t\}_{t \geq 0}$ est à valeurs entières.
- Les trajectoires de $\{N_t\}_{t \geq 0}$ sont constantes par morceaux, continues à droite avec des limites à gauche (cadlag) et croissantes.

On se contentera ici de supposer que le processus $\{N_t\}_{t \geq 0}$ est simple, c'est-à-dire qu'il ne peut y avoir plus d'une défaillance à la fois :

$$\forall t \geq 0, \forall \Delta t \geq 0, \mathbb{P}(N_{t+\Delta t} - N_t \geq 2) = o(\Delta t) \quad (2.1)$$

Dans ce cas, les sauts du processus sont d'une unité à la fois. Alors, la trajectoire du processus des défaillances est du type de celle présentée dans la figure 2.1.

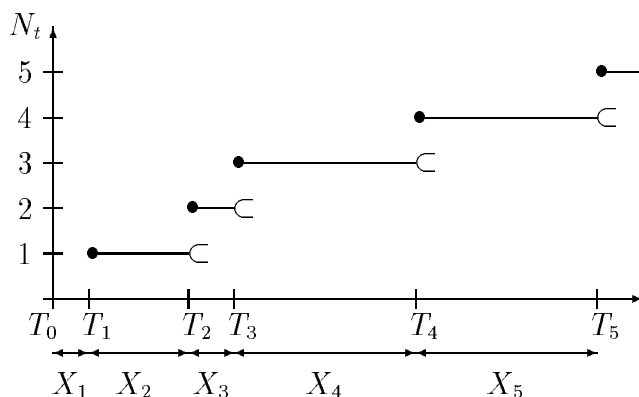


FIG. 2.1 – Une trajectoire du processus des défaillances

Il est tout à fait possible d'envisager des cas où le processus des défaillances n'est pas simple. Dans la pratique, cela arrive plutôt quand les séries d'événements ne sont pas des défaillances (par exemple, arrivée simultanée de plusieurs clients à un guichet ou de plusieurs sinistres dans une compagnie d'assurances). Nous ne traiterons pas ce cas dans ce cours.

Il est clair que si on connaît la loi du processus $\{T_i\}_{i \geq 1}$, on connaît celle du processus $\{X_i\}_{i \geq 1}$ et inversement, puisque $\forall i \in \mathbb{N}, X_i = T_i - T_{i-1}$ et $T_i = \sum_{j=1}^i X_j$. En fait c'est aussi le cas pour le processus $\{N_t\}_{t \geq 0}$ puisque :

$$\forall i \in \mathbb{N}, \forall t \in \mathbb{R}^+, [T_i \leq t] \iff [N_t \geq i] \text{ et } [T_{i+1} > t] \iff [N_t \leq i] \quad (2.2)$$

Par conséquent, le processus des défaillances est bien défini de manière équivalente par les processus aléatoires $\{T_i\}_{i \geq 1}$, $\{X_i\}_{i \geq 1}$ ou $\{N_t\}_{t \geq 0}$.

En complément de 2.2, on peut remarquer que :

$$\text{Pour } t > t_i, \text{ sachant } [T_1 = t_1, \dots, T_i = t_i], \text{ on a } [T_{i+1} > t] \iff [N_t = i] \quad (2.3)$$

Cette remarque sera utile ultérieurement.

Pour pouvoir prédire le futur du processus, on a besoin du passé. Plus formellement, il faut introduire la notion de filtration. Nous reprenons pour cela le formalisme de Coccozza-Thivent [2].

Tout d'abord, on suppose que les variables aléatoires $N_t, t > 0$, sont toutes définies sur le même espace probabilisé $(\Omega, \mathcal{A}, \mathcal{P})$.

Une **filtration** $\mathcal{F} = \{\mathcal{F}_t\}_{t \geq 0}$ est une suite croissante de sous-tribus de \mathcal{A} : $s < t \implies \mathcal{F}_s \subset \mathcal{F}_t$. Le processus $\{N_t\}_{t \geq 0}$ est dit **\mathcal{F} -adapté** ssi $\forall t \geq 0, N_t$ est \mathcal{F}_t -mesurable. En clair, cela signifie que \mathcal{F}_t contient l'ensemble des informations du passé de l'instant t susceptibles d'influencer la réalisation de la variable aléatoire N_t .

La loi de probabilité du processus $\{N_t\}_{t \geq 0}$ est entièrement déterminée par son **intensité de défaillances** :

$$\lambda_t = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1 | \mathcal{F}_t) \quad (2.4)$$

On peut aussi écrire cette définition sous la forme :

$$\mathbb{P}(N_{t+\Delta t} - N_t = 1 | \mathcal{F}_t) = \lambda_t \Delta t + o(\Delta t) \quad (2.5)$$

L'intensité exprime la propension du système à avoir une défaillance juste après l'instant t , sachant toutes les informations disponibles à l'instant t . Mais ce n'est pas une probabilité.

La **fiabilité** du système à l'instant t donne la probabilité que le système fonctionne sans défaillance pendant n'importe quelle durée à partir de t . C'est donc la fonction du temps R_t définie par :

$$\forall \tau \in \mathbb{R}^+, R_t(\tau) = \mathbb{P}(N_{t+\tau} - N_t = 0 | \mathcal{F}_t) = \mathbb{P}(T_{N_t+1} - t > \tau | \mathcal{F}_t) \quad (2.6)$$

On conçoit bien la définition de la fiabilité à l'aide de la figure 2.2.

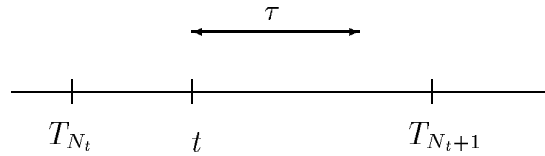


FIG. 2.2 – Fiabilité pour une durée τ à partir de t

Dans le cadre de cours, on se contentera de supposer que le futur du processus des défaillances ne dépend que de son propre passé. Mathématiquement, cela signifie que \mathcal{F}_t est la tribu engendrée par le passé du processus à l'instant t :

$$\mathcal{F}_t = \sigma(\{N_s\}_{s \leq t}) \quad (2.7)$$

Comme le processus $\{N_t\}_{t \geq 0}$ est constant par morceaux et ne change de valeurs qu'aux instants $\{T_i\}_{i \geq 1}$, on a aussi :

$$\mathcal{F}_t = \sigma(N_t, T_1, \dots, T_{N_t}) \quad (2.8)$$

et l'intensité du processus peut s'écrire :

$$\lambda_t = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1 | N_t, T_1, \dots, T_{N_t}) \quad (2.9)$$

On voit que, dans ce cas, l'intensité exprime la propension du système à avoir une défaillance juste après l'instant t , sachant combien il y a eu de défaillances avant t et quand elles se sont produites. Alors le processus $\{N_t\}_{t \geq 0}$ est appelé un **processus ponctuel auto-excité** et on note pour simplifier l'intensité :

$$\lambda_t = \lambda_t(N_t, T_1, \dots, T_{N_t}) \quad (2.10)$$

L'intensité λ_t est aux systèmes réparables ce que le taux de défaillance $\lambda(t)$ est aux systèmes non réparables. Une différence notable est que $\lambda(t)$ est une quantité déterministe alors que λ_t est une variable aléatoire, fonction de N_t, T_1, \dots, T_{N_t} . Sur un jeu de données,

l'observation sera une réalisation n, t_1, \dots, t_n des variables aléatoires N_t, T_1, \dots, T_{N_t} , et l'intensité observée $\lambda_t(n, t_1, \dots, t_n)$ sera une réalisation de la variable aléatoire $\lambda_t = \lambda_t(N_t, T_1, \dots, T_{N_t})$.

En général, le futur du processus des défaillances dépend non seulement de son passé mais aussi de variables environnementales ou stress. Dans ce cas, il faut prendre en compte ces variables dans la filtration et le processus n'est plus auto-excité.

2.2 Propriétés des processus ponctuels auto-excités

On montre (voir par exemple Snyder-Miller [5]) que le processus des défaillances est entièrement déterminé par l'intensité de défaillance, ce qui signifie que les lois de probabilité de $\{T_i\}_{i \geq 1}$, $\{X_i\}_{i \geq 1}$ et $\{N_t\}_{t \geq 0}$, et toutes les grandeurs caractéristiques de la fiabilité s'expriment à l'aide de λ_t . C'est ce que nous allons voir dans cette section.

2.2.1 Loi du nombre de défaillances survenues

On appelle **intensité de comptage conditionnelle** la fonction $\hat{\lambda}_t$ de \mathbb{N} dans \mathbb{R}^+ définie par :

$$\hat{\lambda}_t(n) = \mathbb{E}[\lambda_t | N_t = n] = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} = n + 1 | N_t = n) \quad (2.11)$$

Alors que λ_t est aléatoire, $\hat{\lambda}_t$ est déterministe. Cette fonction exprime la propension du système à tomber en panne juste après t sachant qu'on a observé n défaillances jusque là. L'intérêt de cette fonction est qu'elle permet de déterminer la loi de probabilité de la variable aléatoire N_t .

Proposition 1 . *La loi de probabilité du nombre de défaillances survenues à chaque instant N_t est donnée par :*

$$\mathbb{P}(N_t = 0) = \exp\left(-\int_0^t \hat{\lambda}_s(0) ds\right) \quad (2.12)$$

et $\forall n \geq 1$,

$$\mathbb{P}(N_t = n) = \int_{0 < t_1 < \dots < t_n} \left[\prod_{i=1}^n \hat{\lambda}_{t_i}(i-1) \right] \exp\left(-\sum_{i=0}^n \int_{t_i}^{t_{i+1}} \hat{\lambda}_s(i) ds\right) dt_1, \dots, dt_n \quad (2.13)$$

avec la convention $t_0 = 0$ et $t_{n+1} = t$.

Démonstration: On pose $p_n(t) = \mathbb{P}(N_t = n)$. La définition de $\hat{\lambda}_t(n)$ entraîne que :

$$\forall t \in \mathbb{R}^+, \forall n \in \mathbb{N}, \mathbb{P}(N_{t+\Delta t} = n + 1 | N_t = n) = \hat{\lambda}_t(n)\Delta t + o(\Delta t) \quad (2.14)$$

En particulier, pour $n = 0$, $\mathbb{P}(N_{t+\Delta t} = 1 | N_t = 0) = \hat{\lambda}_t(0)\Delta t + o(\Delta t)$. Comme on a fait l'hypothèse qu'il n'y a pas de défaillances simultanées, on a $\mathbb{P}(N_{t+\Delta t} \geq 2 | N_t = 0) = o(\Delta t)$, d'où :

$$\mathbb{P}(N_{t+\Delta t} = 0 | N_t = 0) = 1 - \hat{\lambda}_t(0)\Delta t + o(\Delta t) = \frac{\mathbb{P}(N_{t+\Delta t} = 0)}{\mathbb{P}(N_t = 0)} = \frac{p_0(t + \Delta t)}{p_0(t)} \quad (2.15)$$

$$\begin{aligned}
 \text{D'où } p'_0(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [p_0(t + \Delta t) - p_0(t)] = p_0(t) \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \left[\frac{p_0(t + \Delta t)}{p_0(t)} - 1 \right] \\
 &= p_0(t) \lim_{\Delta t \rightarrow 0} \left[-\hat{\lambda}_t(0) + \frac{o(\Delta t)}{\Delta t} \right] = -p_0(t) \hat{\lambda}_t(0)
 \end{aligned} \tag{2.16}$$

Alors $\frac{p'_0(t)}{p_0(t)} = -\hat{\lambda}_t(0) = \frac{d}{dt} \ln p_0(t)$ d'où $\ln p_0(t) = -\int_0^t \hat{\lambda}_s(0) ds + cste$.

Or $\ln p_0(0) = \ln \mathbb{P}(N_0 = 0) = \ln 1 = 0 = cste$, ce qui prouve que :

$$p_0(t) = \exp \left(-\int_0^t \hat{\lambda}_s(0) ds \right) \tag{2.17}$$

Pour $n \geq 1$, on utilise le théorème des probabilités totales et le fait qu'il ne peut pas y avoir plus d'une défaillance à la fois pour écrire :

$$\begin{aligned}
 \mathbb{P}(N_{t+\Delta t} = n) &= \mathbb{P}(N_{t+\Delta t} = n | N_t = n) \mathbb{P}(N_t = n) \\
 &\quad + \mathbb{P}(N_{t+\Delta t} = n | N_t = n-1) \mathbb{P}(N_t = n-1) + o(\Delta t) \\
 &= \left[1 - \hat{\lambda}_t(n) \Delta t + o(\Delta t) \right] \mathbb{P}(N_t = n) \\
 &\quad + \left[\hat{\lambda}_t(n-1) \Delta t + o(\Delta t) \right] \mathbb{P}(N_t = n-1) + o(\Delta t)
 \end{aligned} \tag{2.18}$$

Alors

$$p'_n(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [p_n(t + \Delta t) - p_n(t)] = -\hat{\lambda}_t(n) p_n(t) + \hat{\lambda}_t(n-1) p_{n-1}(t) \tag{2.19}$$

Posons $q_n(t) = p_n(t) \exp \left(\int_0^t \hat{\lambda}_s(n) ds \right)$. On a :

$$q'_n(t) = \exp \left(\int_0^t \hat{\lambda}_s(n) ds \right) \left[p'_n(t) + \hat{\lambda}_t(n) p_n(t) \right] = \exp \left(\int_0^t \hat{\lambda}_s(n) ds \right) \hat{\lambda}_t(n-1) p_{n-1}(t) \tag{2.20}$$

D'où :

$$q_n(t) = \int_0^t \hat{\lambda}_u(n-1) p_{n-1}(u) \exp \left(\int_0^u \hat{\lambda}_s(n) ds \right) du = p_n(t) \exp \left(\int_0^t \hat{\lambda}_s(n) ds \right) \tag{2.21}$$

et :

$$p_n(t) = \int_0^t \hat{\lambda}_u(n-1) p_{n-1}(u) \exp \left(-\int_u^t \hat{\lambda}_s(n) ds \right) du \tag{2.22}$$

On peut recommencer en exprimant $p_{n-1}(u)$ à l'aide de p_{n-2} et ainsi de suite. On obtient :

$$\begin{aligned}
 p_n(t) &= \int_0^t \int_0^{t_n} \hat{\lambda}_{t_n}(n-1) \hat{\lambda}_{t_{n-1}}(n-2) p_{n-2}(t_{n-1}) e^{-\int_{t_{n-1}}^{t_n} \hat{\lambda}_s(n-1) ds - \int_{t_n}^t \hat{\lambda}_s(n) ds} dt_{n-1} dt_n \\
 &= \dots \\
 &= \int_0^t \int_0^{t_n} \dots \int_0^{t_2} \left[\prod_{i=1}^n \hat{\lambda}_{t_i}(i-1) \right] p_0(t_1) \exp \left(-\sum_{i=1}^n \int_{t_i}^{t_{i+1}} \hat{\lambda}_s(i) ds \right) dt_1, \dots, dt_n
 \end{aligned} \tag{2.23}$$

en posant $t_{n+1} = t$.

Comme $p_0(t_1) = \exp\left(-\int_0^{t_1} \hat{\lambda}_s(0) ds\right)$, on obtient le résultat final :

$$\mathbb{P}(N_t = n) = \int_{0 < t_1 < \dots < t_n} \left[\prod_{i=1}^n \hat{\lambda}_{t_i}(i-1) \right] \exp\left(-\sum_{i=0}^n \int_{t_i}^{t_{i+1}} \hat{\lambda}_s(i) ds\right) dt_1, \dots, dt_n \quad (2.24)$$

■

Liées à la loi de N_t , on introduit maintenant les notions suivantes :

- La **fonction d'accumulation des défaillances** est $m(t) = \mathbb{E}(N_t)$.
- L'**intensité cumulée** du processus des défaillances est $\Lambda_t = \int_0^t \lambda_s ds$.
- Le **taux d'occurrence des défaillances**, noté *ROCOF* pour rate of occurrence of failures, est $ROCOF_t = m'(t) = \frac{d}{dt} \mathbb{E}(N_t)$.

Le *ROCOF* est en quelque sorte le nombre moyen de défaillances par unité de temps. On peut en fait le relier très simplement à l'intensité du processus :

Proposition 2 . $\forall t \in \mathbb{R}^+, ROCOF_t = \mathbb{E}(\lambda_t)$.

Démonstration : $\mathbb{E}(N_t) = \sum_{n=0}^{+\infty} n \mathbb{P}(N_t = n)$. D'où :

$$\begin{aligned} ROCOF_t &= \frac{d}{dt} \mathbb{E}(N_t) = \sum_{n=0}^{+\infty} n \frac{d}{dt} \mathbb{P}(N_t = n) = \sum_{n=0}^{+\infty} n p'_n(t) \\ &= \sum_{n=1}^{+\infty} n \left[-\hat{\lambda}_t(n) p_n(t) + \hat{\lambda}_t(n-1) p_{n-1}(t) \right] \\ &= \sum_{n=0}^{+\infty} \left[-n \hat{\lambda}_t(n) p_n(t) + (n+1) \hat{\lambda}_t(n) p_n(t) \right] \quad (2.25) \\ &= \sum_{n=0}^{+\infty} \hat{\lambda}_t(n) p_n(t) = \sum_{n=0}^{+\infty} \mathbb{E}[\lambda_t | N_t = n] \mathbb{P}(N_t = n) \\ &= \mathbb{E}[\mathbb{E}[\lambda_t | N_t]] = \mathbb{E}(\lambda_t) \end{aligned}$$

■

On en déduit immédiatement que si l'intensité de défaillance n'est pas aléatoire, c'est à dire si $\lambda_t = \lambda(t)$ est une fonction déterministe du temps, alors $ROCOF_t = \lambda(t)$ et $m(t) = \mathbb{E}(N_t) = \int_0^t \lambda(s) ds = \Lambda(t)$.

Proposition 3 . $\forall t \in \mathbb{R}^+, m(t) = \sum_{i=1}^{+\infty} F_{T_n}(t)$. Si T_n admet une densité f_{T_n} , alors $\forall t \in \mathbb{R}^+, ROCOF_t = \sum_{i=1}^{+\infty} f_{T_n}(t)$.

Démonstration: Puisque $[N_t \geq n] \iff [T_n \leq t]$, on a $F_{T_n}(t) = P(T_n \leq t) = P(N_t \geq n)$. D'où :

$$\begin{aligned} m(t) &= \mathbb{E}(N_t) = \sum_{n=0}^{+\infty} n \mathbb{P}(N_t = n) = \sum_{n=1}^{+\infty} n [P(N_t \geq n) - P(N_t \geq n+1)] \quad (2.26) \\ &= \sum_{n=1}^{+\infty} n [F_{T_n}(t) - F_{T_{n+1}}(t)] = \sum_{n=1}^{+\infty} [n F_{T_n}(t) - (n-1) F_{T_n}(t)] = \sum_{n=1}^{+\infty} F_{T_n}(t) \end{aligned}$$

Evidemment, si m et F_{T_n} sont dérivables, on a :

$$m'(t) = ROCOF_t = \sum_{i=1}^{+\infty} f_{T_n}(t) \quad (2.27)$$

■

2.2.2 Loi des instants de défaillance

Proposition 4 . *La loi de T_1 a pour densité*

$$f_{T_1}(t) = \lambda_t(0) \exp\left(-\int_0^t \lambda_s(0) ds\right) \quad (2.28)$$

La loi conditionnelle de T_{n+1} sachant $[T_1 = t_1, \dots, T_n = t_n]$ a pour densité

$$f_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = \lambda_t(n, t_1, \dots, t_n) \exp\left(-\int_{t_n}^t \lambda_s(n, t_1, \dots, t_n) ds\right) \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \quad (2.29)$$

Démonstration: Tout d'abord, on a évidemment :

$$\mathbb{P}(T_1 > t) = \mathbb{P}(N_t = 0) = \exp\left(-\int_0^t \hat{\lambda}_u(0) du\right) \quad (2.30)$$

$$\begin{aligned} \text{Or } \hat{\lambda}_t(0) &= \mathbb{E}[\lambda_t | N_t = 0] = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} = 1 | N_t = 0) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1 | N_t = 0) = \lambda_t(0) \end{aligned} \quad (2.31)$$

On a donc bien :

$$f_{T_1}(t) = \lambda_t(0) \exp\left(-\int_0^t \lambda_s(0) ds\right) \quad (2.32)$$

Pour $n \geq 1$ et $0 < t_1 < \dots < t_n < t$, on écrit :

$$\begin{aligned} \lambda_t(n, t_1, \dots, t_n) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1 | N_t = n, T_1 = t_1, \dots, T_n = t_n) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(t < T_{n+1} \leq t + \Delta t | N_t = n, T_1 = t_1, \dots, T_n = t_n) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{\mathbb{P}(t < T_{n+1} \leq t + \Delta t \cap N_t = n | T_1 = t_1, \dots, T_n = t_n)}{\mathbb{P}(N_t = n | T_1 = t_1, \dots, T_n = t_n)} \end{aligned} \quad (2.33)$$

Or, d'après (2.3), sachant $[T_1 = t_1, \dots, T_n = t_n]$, $[N_t = n] \iff [T_{n+1} > t]$. D'où :

$$\begin{aligned} \lambda_t(n, t_1, \dots, t_n) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \frac{\mathbb{P}(t < T_{n+1} \leq t + \Delta t | T_1 = t_1, \dots, T_n = t_n)}{\mathbb{P}(T_{n+1} > t | T_1 = t_1, \dots, T_n = t_n)} \\ &= \frac{f_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t)}{1 - F_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t)} \\ &= -\frac{d}{dt} \ln [1 - F_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t)] \end{aligned} \quad (2.34)$$

Par conséquent :

$$F_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = 1 - \exp \left(- \int_0^t \lambda_s(n, t_1, \dots, t_n) ds + cste \right) \quad (2.35)$$

Puisque $T_{n+1} > t_n$, on a $F_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t_n) = 0$, donc $cste = \int_0^{t_n} \lambda_s(n, t_1, \dots, t_n) ds$, et on obtient :

$$F_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = 1 - \exp \left(- \int_{t_n}^t \lambda_s(n, t_1, \dots, t_n) ds \right) \text{ pour } 0 < t_1 < \dots < t_n < t \quad (2.36)$$

Il ne reste plus qu'à dériver pour obtenir la densité conditionnelle de T_{n+1} sachant $[T_1 = t_1, \dots, T_n = t_n]$:

$$f_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = \lambda_t(n, t_1, \dots, t_n) \exp \left(- \int_{t_n}^t \lambda_s(n, t_1, \dots, t_n) ds \right) \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \quad (2.37)$$

■

On en déduit immédiatement la densité conjointe du vecteur des n premiers instants de défaillance (T_1, \dots, T_n) :

Proposition 5 .

$$f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) = \left[\prod_{i=1}^n \lambda_{t_i}(i-1, t_1, \dots, t_{i-1}) \right] \exp \left(- \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \lambda_s(i-1, t_1, \dots, t_{i-1}) ds \right) \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \quad (2.38)$$

Démonstration :

$$\begin{aligned} f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) &= f_{T_n|T_1=t_1, \dots, T_{n-1}=t_{n-1}}(t) f_{(T_1, \dots, T_{n-1})}(t_1, \dots, t_{n-1}) \\ &= \dots = \left[\prod_{i=2}^n f_{T_i|T_1=t_1, \dots, T_{i-1}=t_{i-1}}(t_i) \right] f_{T_1}(t_1) \end{aligned} \quad (2.39)$$

$$\begin{aligned} &= \left[\prod_{i=2}^n \lambda_{t_i}(i-1, t_1, \dots, t_{i-1}) \exp \left(- \int_{t_{i-1}}^{t_i} \lambda_s(i-1, t_1, \dots, t_{i-1}) ds \right) \right] \\ &\quad \mathbb{1}_{\{0 < t_1 < \dots < t_{i-1}\}} \lambda_{t_1}(0) \exp \left(- \int_0^{t_1} \lambda_s(0) ds \right) \end{aligned} \quad (2.40)$$

ce qui donne bien la formule annoncée. ■

2.2.3 Loi des durées inter-défaillances

Les relations liant les instants de défaillance T_n aux durées inter-défaillances X_n permettent d'établir les caractéristiques de la loi des X_n .

Proposition 6 . *La loi de X_1 a pour densité*

$$f_{X_1}(x) = \lambda_x(0) \exp\left(-\int_0^x \lambda_s(0) ds\right) \quad (2.41)$$

La loi conditionnelle de X_{n+1} sachant $[T_1 = t_1, \dots, T_n = t_n]$ a pour densité :

$$f_{X_{n+1}|T_1=t_1, \dots, T_n=t_n}(x) = \lambda_{t_n+x}(n, t_1, \dots, t_n) \exp\left(-\int_0^x \lambda_{t_n+s}(n, t_1, \dots, t_n) ds\right) \quad (2.42)$$

Démonstration: Il suffit d'écrire :

$$\begin{aligned} F_{X_{n+1}|T_1=t_1, \dots, T_n=t_n}(x) &= \mathbb{P}(X_{n+1} \leq x | T_1 = t_1, \dots, T_n = t_n) \\ &= \mathbb{P}(T_{n+1} \leq t_n + x | T_1 = t_1, \dots, T_n = t_n) \\ &= 1 - \exp\left(-\int_{t_n}^{t_n+x} \lambda_s(n, t_1, \dots, t_n) ds\right) \\ &= 1 - \exp\left(-\int_0^x \lambda_{t_n+s}(n, t_1, \dots, t_n) ds\right) \end{aligned} \quad (2.43)$$

D'où le résultat, par dérivation. ■

On peut interpréter ce résultat en disant que le taux de défaillance (ou taux de hasard) de la loi de X_{n+1} sachant $[T_1 = t_1, \dots, T_n = t_n]$ est $\lambda_{t_n+x}(n, t_1, \dots, t_n)$.

Par ailleurs, on peut aussi écrire que la loi conditionnelle de X_{n+1} sachant $[X_1 = x_1, \dots, X_n = x_n]$ a pour densité :

$$f_{X_{n+1}|X_1=x_1, \dots, X_n=x_n}(x) = \lambda_{\sum_{i=1}^n x_i+x}(n, x_1, \dots, \sum_{i=1}^n x_i) \exp\left(-\int_0^x \lambda_{\sum_{i=1}^n x_i+s}(n, x_1, \dots, \sum_{i=1}^n x_i) ds\right) \quad (2.44)$$

Corollaire 1 . *Les durées inter-défaillances X_i sont indépendantes si et seulement si l'intensité du processus de défaillance est de la forme :*

$$\lambda_t = h(N_t, t - T_{N_t}) \quad (2.45)$$

Dans ce cas, le processus de défaillance est un processus de renouvellement : tout se passe comme si, après chaque défaillance, le système était remis à neuf.

Démonstration: Les X_i sont indépendantes ssi $\forall n \geq 1, f_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = \prod_{i=1}^n f_{X_i}(x_i)$ ou bien $f_{X_{n+1}|X_1=x_1, \dots, X_n=x_n}(x)$ ne dépend pas de x_1, \dots, x_n .

Compte-tenu de (2.42), les X_i seront donc indépendants ssi $\lambda_{t_n+x}(n, t_1, \dots, t_n)$ ne dépend pas de t_1, \dots, t_n , c'est-à-dire s'il existe une fonction h telle que $\lambda_{t_n+x}(n, t_1, \dots, t_n) = h(n, x)$, ou, pour $t > t_n$, $\lambda_t(n, t_1, \dots, t_n) = h(n, t - t_n)$.

Donc l'intensité du processus doit être de la forme $\lambda_t = h(N_t, t - T_{N_t})$.

Par ailleurs, on voit que, dans ce cas, la densité de X_i est :

$$f_{X_i}(x) = f_{X_i|X_1=x_1, \dots, X_{i-1}=x_{i-1}}(x) = h(i-1, x) \exp\left(-\int_0^x h(i-1, s) ds\right) \quad (2.46)$$

■

2.2.4 Fiabilité et MTTF

La fiabilité à l'instant t est une fonction du temps qui s'exprime comme une probabilité conditionnelle au passé du processus à l'instant t :

$$R_t(\tau; N_t, T_1, \dots, T_{N_t}) = \mathbb{P}(N_{t+\tau} - N_t = 0 | N_t, T_1, \dots, T_{N_t}) = \mathbb{P}(T_{N_t+1} > \tau | N_t, T_1, \dots, T_{N_t}). \quad (2.47)$$

Par commodité, on la notera simplement $R_t(\tau)$.

Proposition 7 . *La fiabilité à l'instant t est définie par :*

$$R_t(\tau) = \exp\left(-\int_t^{t+\tau} \lambda_s(N_t, T_1, \dots, T_{N_t}) ds\right) \quad (2.48)$$

Démonstration: On a, pour $0 < t_1 < \dots < t_n < t$,

$$\mathbb{P}(T_{N_t+1} - t > \tau | N_t = n, T_1 = t_1, \dots, T_n = t_n) = \mathbb{P}(T_{n+1} > t + \tau | N_t = n, T_1 = t_1, \dots, T_n = t_n)$$

$$= \frac{\mathbb{P}(T_{n+1} > t + \tau \cap N_t = n | T_1 = t_1, \dots, T_n = t_n)}{\mathbb{P}(N_t = n | T_1 = t_1, \dots, T_n = t_n)}$$

$$= \frac{\mathbb{P}(T_{n+1} > t + \tau | T_1 = t_1, \dots, T_n = t_n)}{\mathbb{P}(T_{n+1} > t | T_1 = t_1, \dots, T_n = t_n)} \quad (2.49)$$

$$= \frac{\exp\left(-\int_{t_n}^{t_n+\tau} \lambda_s(n, t_1, \dots, t_n) ds\right)}{\exp\left(-\int_{t_n}^t \lambda_s(n, t_1, \dots, t_n) ds\right)}$$

$$= \exp\left(-\int_t^{t+\tau} \lambda_s(n, t_1, \dots, t_n) ds\right)$$

d'où le résultat. ■

Corollaire 2 . *La fiabilité à l'instant de la $n^{\text{ème}}$ défaillance est :*

$$R_{T_n}(\tau) = \exp\left(-\int_{T_n}^{T_n+\tau} \lambda_s(n, t_1, \dots, t_n) ds\right) \quad (2.50)$$

Ce résultat peut se voir aussi bien comme une application de la proposition 6 ou de la proposition 7.

Corollaire 3 . *Le MTTF à l'instant t est l'espérance (conditionnelle) de la durée d'attente de la prochaine défaillance à partir de l'instant t . Il vaut :*

$$MTTF_t = \mathbb{E}[T_{N_t+1} - t | N_t, T_1, \dots, T_{N_t}] = \int_0^{+\infty} R_t(\tau) d\tau \quad (2.51)$$

Démonstration: Il suffit d'utiliser le résultat qui dit que si T est une variable aléatoire positive absolument continue par rapport à la mesure de Lebesgue et de fonction de répartition F , alors son espérance peut s'écrire $E(T) = \int_0^{+\infty} [1 - F(t)] dt$. ■

2.2.5 Fonctions de vraisemblance

Dans une perspective statistique où il faudra estimer des paramètres de modèles, il est important de calculer la fonction de vraisemblance associée à l'observation du processus des défaillances. Or il y a deux types d'observation, correspondant aux deux exemples de jeux de données. Comme on ne peut pas observer le processus des défaillances jusqu'à l'infini, l'observation est une troncature ou une censure du processus global. Les types d'observation correspondent à deux types différents de censure.

- **Censure de type 1.** Le processus des défaillances est observé sur un intervalle $[0, t]$. L'observation est alors le nombre N_t de défaillances sur cet intervalle et les instants de ces défaillances T_1, \dots, T_{N_t} . C'est le cas des données de la pompe. La fonction de vraisemblance associée à cette observation est la fonction des paramètres θ du modèle sous-jacent définie par [5] :

$$\mathcal{L}(\theta; n, t_1, \dots, t_n) = \mathbb{P}(N_t = n | T_1 = t_1, \dots, T_n = t_n; \theta) f_{(T_1, \dots, T_n)}(t_1, \dots, t_n; \theta) \quad (2.52)$$

En appliquant les résultats des propositions 4 et 5, on obtient :

$$\begin{aligned} \mathcal{L}(\theta; n, t_1, \dots, t_n) &= \mathbb{P}(T_{n+1} > t | T_1 = t_1, \dots, T_n = t_n; \theta) f_{(T_1, \dots, T_n)}(t_1, \dots, t_n; \theta) \\ &= \exp\left(-\int_{t_n}^t \lambda_s(n, t_1, \dots, t_n) ds\right) \mathbb{1}_{\{t_n < t\}} \left[\prod_{i=1}^n \lambda_{t_i}(i-1, t_1, \dots, t_{i-1}) \right] \\ &\quad \exp\left(-\sum_{i=1}^n \int_{t_{i-1}}^{t_i} \lambda_s(i-1, t_1, \dots, t_{i-1}) ds\right) \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \\ &= \left[\prod_{i=1}^n \lambda_{t_i}(i-1, t_1, \dots, t_{i-1}) \right] \exp\left(-\sum_{i=1}^{n+1} \int_{t_{i-1}}^{t_i} \lambda_s(i-1, t_1, \dots, t_{i-1}) ds\right) \\ &\quad \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \end{aligned} \quad (2.53)$$

- **Censure de type 2.** Le processus est observé jusqu'à la $n^{\text{ème}}$ défaillance. Les observations sont alors les instants de ces défaillances T_1, \dots, T_n . C'est le cas des données du logiciel. La fonction de vraisemblance est alors simplement :

$$\mathcal{L}(\theta; n, t_1, \dots, t_n) = f_{(T_1, \dots, T_n)}(t_1, \dots, t_n; \theta) \tag{2.54}$$

2.3 Classification des modèles de fiabilité des systèmes réparables

Puisque la caractéristique essentielle du processus est son intensité, on peut proposer une classification des modèles de fiabilité des systèmes réparables en fonction de la forme de leur intensité de défaillance. On ne signale ici que les trois classes les plus usuelles, mais il en existe bien d'autres.

2.3.1 Les processus de Poisson non homogènes (NHPP)

Le processus des défaillances est un **processus de Poisson non homogène** (NHPP pour Non Homogeneous Poisson Process) ssi son intensité est une fonction déterministe du temps :

$$\lambda_t = \lambda(t) \tag{2.55}$$

Il n'est pas a priori indispensable que λ soit une fonction continue, mais c'est toujours le cas dans la pratique. Cela signifie qu'après la réparation, le système est dans le même état qu'avant : la réparation n'a fait que remettre en fonctionnement le système, mais sans le rendre plus fiable qu'avant. On dit que la réparation est minimale ou que le système est "aussi mauvais que vieux" : **As Bad As Old (ABAO)**.

La forme de λ caractérise l'usure du système. Quand λ est croissante (figure 2.3), le système s'use, vieillit, se dégrade. C'est a priori le cas des systèmes matériels, comme la pompe. Quand λ est décroissante (figure 2.4), le système s'améliore avec le temps. C'est a priori le cas des logiciels, pour lesquels la correction des bugs améliore leur fiabilité.

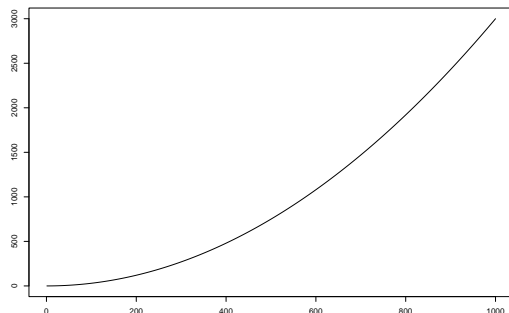


FIG. 2.3 – NHPP avec intensité croissante

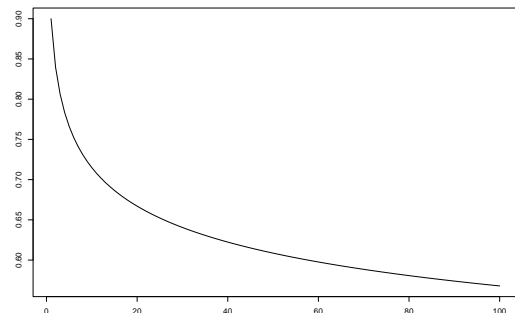


FIG. 2.4 – NHPP avec intensité décroissante

En pratique, on observe souvent pour les matériels une intensité dite “en forme de baignoire” (bathtub-shaped, figure 2.5). Dans ce cas, λ se décompose en 3 parties :

- La **période de jeunesse** : quand un système est neuf, on observe souvent des défaillances précoces, dues à des fautes de conception, des problèmes de montage ou autre. L’intensité de défaillance est donc assez forte au tout début de la vie du système. Ensuite, elle diminue car, s’il y a des défauts initiaux, ils vont se manifester tôt. λ est donc d’abord décroissante. C’est le rodage pour les matériels mécaniques et le déverminage pour les matériels électromécaniques.
- La **vie utile** : pendant cette période, l’intensité de défaillance est pratiquement constante, les défaillances sont purement accidentelles.
- Le **vieillessement** : λ se met à croître sur la fin à cause de l’usure du système.

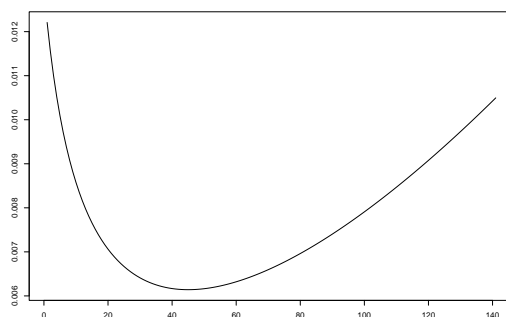


FIG. 2.5 – NHPP avec intensité en forme de baignoire

2.3.2 Les processus de renouvellement

Le processus des défaillances est un **processus de renouvellement** ssi les durées inter-défaillances X_i sont indépendantes et de même loi. Or d’après le corollaire 1, les X_i sont indépendantes ssi l’intensité du processus de défaillance est de la forme $\lambda_t = h(N_t, t - T_{N_t})$, et dans ce cas, la densité de X_i est $f_{X_i}(x) = h(i - 1, x) \exp(-\int_0^x h(i - 1, s) ds)$.

Si on veut que les X_i soient de même loi, on voit qu’il faut que $h(i - 1, s)$ ne dépende que de s . Autrement dit, l’intensité est de la forme (figure 2.6) :

$$\lambda_t = h(t - T_{N_t}) \tag{2.56}$$

Alors la loi des durées inter-défaillances a pour densité $f(x) = h(x) \exp(-\int_0^x h(s) ds)$, ce qui signifie que h est le taux de hasard de cette loi.

Après chaque instant de défaillance T_i , l’intensité du processus repart comme à l’origine. Par conséquent, cela signifie que le système est comme neuf après chaque réparation, d’où le terme “renouvellement”. On dit aussi que le système est “aussi bon que neuf” : **As Good As New (AGAN)**, ou que la réparation est parfaite.

Dans le cas où l’intensité est de la forme

$$\lambda_t = h(N_t, t - T_{N_t}) \tag{2.57}$$

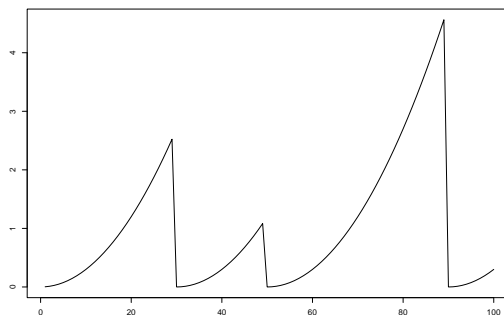


FIG. 2.6 – Intensité d'un processus de renouvellement

les X_i sont indépendantes mais pas de même loi. Les processus correspondant sont parfois appelés **processus de quasi-renouvellement**.

2.3.3 Les modèles de réparation imparfaite

Dans la pratique, pour les matériels, on se situe souvent entre les deux extrêmes ABAO et AGAN : la réparation ralentit la dégradation du système (elle est mieux que minimale), mais elle ne va pas jusqu'à remettre le système à neuf. Il est donc intéressant de proposer des modèles intermédiaires entre les cas ABAO et AGAN.

C'est le cas du modèle à réduction arithmétique de l'âge, pour lequel l'intensité de défaillance est de la forme

$$\lambda_t = h(t - \rho T_{N_t}) \tag{2.58}$$

ρ est un paramètre caractérisant l'efficacité de la réparation :

- $\rho = 0$: on retrouve le cas ABAO, la réparation est minimale.
- $\rho = 1$: on retrouve le cas AGAN, la réparation est parfaite.
- $0 < \rho < 1$: la réparation est imparfaite, mais pas minimale.

La figure 2.7 représente l'intensité d'un de ces modèles pour $\rho = 0.5$.

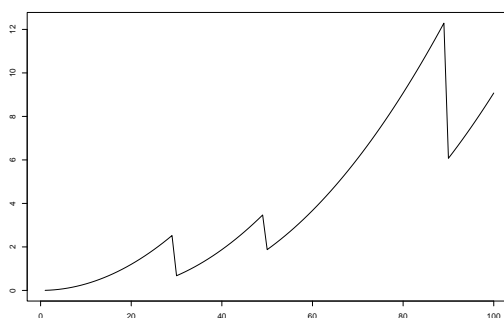


FIG. 2.7 – Intensité d'un modèle de réparation imparfaite

2.3.4 Les modèles à Durées Inter-Défaillances Exponentielles (DIDE)

Comme leur nom l'indique, les modèles à Durées Inter-Défaillances Exponentielles (DIDE) sont tels que les X_i sont indépendantes et de lois exponentielles de paramètres respectifs λ_i .

En utilisant le corollaire 1, on voit que cela implique que l'intensité est de la forme $\lambda_t = h(N_t, t - T_{N_t})$ et la densité de X_i est :

$$f_{X_i}(x) = \lambda_i \exp(-\lambda_i x) = h(i - 1, x) \exp\left(-\int_0^x h(i - 1, s) ds\right) \tag{2.59}$$

Alors on a $h(i - 1, x) = \lambda_i$ et $\lambda_t = \lambda_{N_t+1}$.

Autrement dit, l'intensité de défaillances ne dépend que du nombre de défaillances survenues. On l'écrit plutôt :

$$\lambda_t = h(N_t) \tag{2.60}$$

Alors les X_i sont indépendantes et de lois respectives $\exp(h(i - 1))$.

Pour ces modèles, l'intensité de défaillances est constante entre deux réparations, ce qui signifie une absence d'usure ou d'amélioration sur ces périodes. Ces modèles sont donc en général peu utilisés pour les matériels. Mais ils sont particulièrement adaptés aux logiciels qui, par définition, sont des systèmes qui ne s'usent pas.

Les figures 2.8 et 2.9 donnent des exemples d'intensités de modèles DIDE pour un système qui se dégrade et un système qui s'améliore.

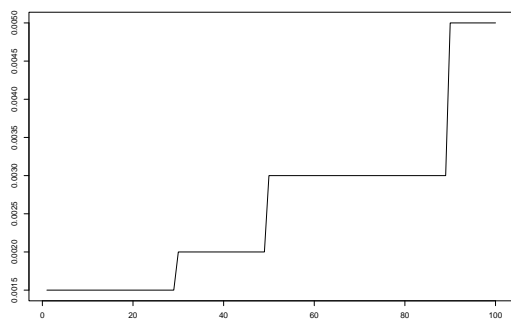


FIG. 2.8 – *Modèle DIDE avec intensité croissante*

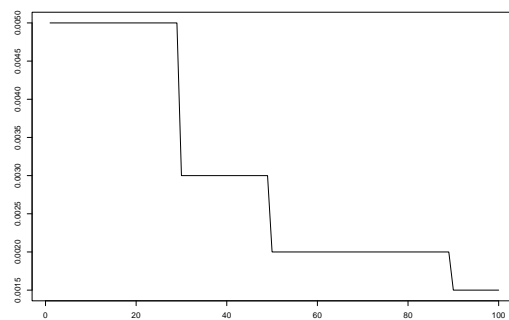


FIG. 2.9 – *Modèle DIDE avec intensité décroissante*

Les propriétés générales des processus ponctuels auto-excités développées dans ce chapitre permettent a priori de déterminer les caractéristiques de toutes ces familles de modèles. Faut de temps, nous n'étudierons dans la suite de ce cours que la plus utilisée de ces familles, celle des processus de Poisson non homogènes.

Chapitre 3

Propriétés générales des NHPP

3.1 Introduction

Rappelons que le processus des défaillances est un processus de Poisson non homogène (NHPP) ssi son intensité est une fonction déterministe du temps: $\lambda_t = \lambda(t)$, que l'on supposera continue. Cela revient à supposer que la réparation est minimale ou que le système après réparation est As Bad As Old (ABAO).

Dans ce chapitre, nous allons établir les propriétés générales des NHPP en appliquant les résultats du chapitre 2 sur les processus ponctuels auto-excités au cas où $\lambda_t = \lambda(t)$. On s'intéressera en particulier à certaines propriétés qui s'avèreront particulièrement utiles dans une perspective statistique.

Les résultats présentés dans ce chapitre peuvent se retrouver dans Snyder-Miller [5], Coccozza-Thivent [2] ou Rigdon-Basu [4].

3.2 Processus de comptage des défaillances

Proposition 8 . *Le processus de comptage des défaillances $\{N_t\}_{t \geq 0}$ vérifie les propriétés suivantes :*

1. *Il est à accroissements indépendants.*
2. $\forall t \geq 0$, N_t est de loi de Poisson $\mathcal{P}(m(t))$ (d'où le nom du processus), avec $m(t) = \mathbb{E}(N_t) = \int_0^t \lambda(u) du = \Lambda(t)$.
3. $\forall (s, t), 0 \leq s \leq t$, $N_t - N_s$ est de loi de Poisson $\mathcal{P}(m(t) - m(s)) = \mathcal{P}\left(\int_s^t \lambda(u) du\right)$.

Démonstration :

1. L'intensité du processus est définie par $\lambda_t = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1 | N_t, T_1, \dots, T_{N_t})$. Pour un NHPP, elle ne dépend pas de N_t, T_1, \dots, T_{N_t} donc elle peut s'écrire $\lambda_t = \lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_{t+\Delta t} - N_t = 1)$.

Par conséquent, $N_{t+\Delta t}$ et N_t sont indépendants, ce qui prouve que $\{N_t\}_{t \geq 0}$ est à accroissements indépendants. Plus généralement, pour $s_1 < t_1 < s_2 < t_2$, $N_{t_1} - N_{s_1}$ et $N_{t_2} - N_{s_2}$ sont des variables aléatoires indépendantes.

2. Par définition, $m(t) = \mathbb{E}(N_t)$, $\Lambda(t) = \int_0^t \lambda(u) du$ et $ROCOF_t = m'(t) = \frac{d}{dt} \mathbb{E}(N_t) = \mathbb{E}(\lambda_t)$.

Ici, $\mathbb{E}(\lambda_t) = \lambda_t = \lambda(t) = m'(t)$. Donc $\forall t$, $m(t) = \Lambda(t)$ et on sait déjà que $\mathbb{E}(N_t) = \int_0^t \lambda(u) du$.

Pour déterminer la loi de N_t , on applique la proposition 1. Comme λ_t n'est pas aléatoire, l'intensité de comptage conditionnelle est $\hat{\lambda}_t(n) = \mathbb{E}[\lambda_t | N_t = n] = \lambda(t)$. Donc on obtient directement :

$$\mathbb{P}(N_t = 0) = \exp\left(-\int_0^t \hat{\lambda}_s(0) ds\right) = \exp\left(-\int_0^t \lambda(s) ds\right) \quad (3.1)$$

Pour $n \geq 1$,

$$\begin{aligned} \mathbb{P}(N_t = n) &= \int_{0 < t_1 < \dots < t_n} \left[\prod_{i=1}^n \hat{\lambda}_{t_i}(i-1) \right] \exp\left(-\sum_{i=0}^n \int_{t_i}^{t_{i+1}} \hat{\lambda}_s(i) ds\right) dt_1, \dots, dt_n \\ &= \int_{0 < t_1 < \dots < t_n} \left[\prod_{i=1}^n \lambda(t_i) \right] \exp\left(-\sum_{i=0}^n \int_{t_i}^{t_{i+1}} \lambda(s) ds\right) dt_1, \dots, dt_n \quad (3.2) \\ &= \exp\left(-\int_0^t \lambda(s) ds\right) \int_{0 < t_1 < \dots < t_n} \left[\prod_{i=1}^n \lambda(t_i) \right] dt_1, \dots, dt_n \end{aligned}$$

On a :

$$\begin{aligned} \int_0^{t_2} \lambda(t_1) dt_1 &= \Lambda(t_2) \\ \int_0^{t_3} \Lambda(t_2) \lambda(t_2) dt_2 &= \frac{\Lambda(t_3)^2}{2} \quad (3.3) \\ \int_0^{t_4} \frac{\Lambda(t_3)^2}{2} \lambda(t_3) dt_3 &= \frac{\Lambda(t_4)^3}{3!} \end{aligned}$$

Et ainsi de suite. On obtient finalement :

$$\mathbb{P}(N_t = n) = \exp(-\Lambda(t)) \frac{\Lambda(t)^n}{n!} \quad (3.4)$$

ce qui prouve que N_t est de loi de Poisson $\mathcal{P}(\Lambda(t)) = \mathcal{P}(m(t))$.

3. Pour $s < t$, N_s est de loi $\mathcal{P}(\Lambda(s))$, N_t est de loi $\mathcal{P}(\Lambda(t))$ et $N_t - N_s$ est indépendant de N_s . La fonction génératrice de N_t est :

$$G_{N_t}(u) = E(u^{N_t}) = e^{m(t)(u-1)} \quad (3.5)$$

Grâce à l'indépendance des accroissements, on a :

$$G_{N_t}(u) = E(u^{N_t}) = E(u^{N_t - N_s + N_s}) = E(u^{N_t - N_s}) E(u^{N_s}) = G_{N_t - N_s}(u) G_{N_s}(u) \quad (3.6)$$

D'où

$$G_{N_t - N_s}(u) = \frac{G_{N_t}(u)}{G_{N_s}(u)} = \frac{e^{m(t)(u-1)}}{e^{m(s)(u-1)}} = e^{[m(t) - m(s)](u-1)} \quad (3.7)$$

ce qui prouve que $N_t - N_s$ est de loi $\mathcal{P}(m(t) - m(s)) = \mathcal{P}\left(\int_s^t \lambda(u) du\right)$. ■

En fait, on peut montrer que $\{N_t\}_{t \geq 0}$ est un NHPP d'intensité $\lambda(t)$ ssi :

- $\{N_t\}_{t \geq 0}$ est simple.
- $\{N_t\}_{t \geq 0}$ est à accroissements indépendants.
- $N_0 = 0$ p.s.
- $\forall t > 0$, N_t est de loi $\mathcal{P}\left(\int_0^t \lambda(u) du\right)$.

Proposition 9 . Si $\lim_{t \rightarrow +\infty} m(t) = +\infty$, alors :

- $\frac{N_t}{m(t)} \xrightarrow[t \rightarrow +\infty]{p.s.} 1$.
- $\frac{N_t - m(t)}{\sqrt{m(t)}} \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$.

Démonstration : Il suffit d'utiliser les propriétés connues des lois de Poisson : Si K est de loi $\mathcal{P}(\lambda)$, $\frac{K}{\lambda} \xrightarrow[\lambda \rightarrow +\infty]{p.s.} 1$ (loi des grands nombres) et $\frac{K - \lambda}{\sqrt{\lambda}} \xrightarrow[\lambda \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$ (théorème central-limite). ■

Remarquons que $\lim_{t \rightarrow +\infty} m(t) = \lim_{t \rightarrow +\infty} \mathbb{E}(N_t) = +\infty$ signifie qu'on suppose que le nombre de défaillances du système n'est pas borné. Or ce n'est pas toujours le cas. Par exemple, pour un logiciel, quand toutes les fautes sont corrigées, le logiciel n'aura plus jamais de défaillance.

3.3 Lois des instants de défaillance et des durées inter-défaillances

En appliquant les résultats généraux des propositions 4 et 5, on obtient directement les résultats suivants sur les lois des instants de défaillance T_i .

Proposition 10 .

$$f_{T_1}(t) = \lambda(t) e^{-m(t)} \tag{3.8}$$

$$f_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = \lambda(t) e^{-[m(t) - m(t_n)]} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \tag{3.9}$$

$$f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) = \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t_n)} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \tag{3.10}$$

Le second résultat prouve que la loi de T_{n+1} sachant $[T_1 = t_1, \dots, T_n = t_n]$ ne dépend que de t_n . Donc le processus $\{T_i\}_{i \geq 1}$ est markovien. On peut alors écrire cette densité sous

la forme :

$$f_{T_{n+1}|T_n=t_n}(t) = \lambda(t) e^{-[m(t)-m(t_n)]} \mathbb{1}_{\{0 < t_n < t\}} \quad (3.11)$$

Corollaire 4 . Si $\lim_{t \rightarrow +\infty} m(t) = +\infty$, la loi de T_n a pour densité

$$f_{T_n}(t) = \lambda(t) e^{-m(t)} \frac{[m(t)]^{n-1}}{(n-1)!} \quad (3.12)$$

Démonstration :

$$\begin{aligned} f_{T_n}(t) &= \int_{t_1, \dots, t_{n-1}} f_{(T_1, \dots, T_{n-1}, T_n)}(t_1, \dots, t_{n-1}, t) dt_1 \dots dt_{n-1} \\ &= \int_{t_1, \dots, t_{n-1}} \left[\prod_{i=1}^{n-1} \lambda(t_i) \right] \lambda(t) e^{-m(t)} \mathbb{1}_{\{0 < t_1 < \dots < t_{n-1} < t\}} dt_1 \dots dt_{n-1} \quad (3.13) \\ &= \lambda(t) e^{-m(t)} \int_{0 < t_1 < \dots < t_{n-1} < t} \left[\prod_{i=1}^{n-1} \lambda(t_i) \right] dt_1 \dots dt_{n-1} \quad (3.14) \end{aligned}$$

Le calcul de l'intégrale est similaire à celui qui a déjà été effectué dans la détermination de la loi de N_t . Elle vaut $\frac{[m(t)]^{n-1}}{(n-1)!}$, d'où le résultat.

On peut aussi obtenir ce résultat en écrivant :

$$\begin{aligned} f_{T_n}(t) &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [F_{T_n}(t + \Delta t) - F_{T_n}(t)] = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [\mathbb{P}(N_{t+\Delta t} \geq n) - \mathbb{P}(N_t \geq n)] \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_t = n-1 \cap N_{t+\Delta t} - N_t = 1) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \mathbb{P}(N_t = n-1) \mathbb{P}(N_{t+\Delta t} - N_t = 1) \quad (3.15) \\ &= e^{-m(t)} \frac{[m(t)]^{n-1}}{(n-1)!} \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} e^{-[m(t+\Delta t)-m(t)]} \frac{[m(t+\Delta t) - m(t)]^1}{1!} \\ &= e^{-m(t)} \frac{[m(t)]^{n-1}}{(n-1)!} m'(t) = e^{-m(t)} \frac{[m(t)]^{n-1}}{(n-1)!} \lambda(t) \end{aligned}$$

■

Corollaire 5 . Sachant $[T_n = t_n]$, le vecteur (T_1, \dots, T_{n-1}) a la même loi que la statistique d'ordre d'un échantillon de taille $n-1$ de la loi de densité $\frac{\lambda(s)}{m(t_n)} \mathbb{1}_{[0, t_n]}(s)$.

Démonstration: Si Y_1, \dots, Y_n sont n variables aléatoires indépendantes et de densité f , si $Y_1^* \leq \dots \leq Y_n^*$ sont les statistiques d'ordre de cet échantillon, alors la densité de la statistique d'ordre (Y_1^*, \dots, Y_n^*) est :

$$f_{(Y_1^*, \dots, Y_n^*)}(y_1^*, \dots, y_n^*) = n! \prod_{i=1}^n f(y_i^*) \mathbb{1}_{\{y_1^* \leq \dots \leq y_n^*\}} \quad (3.16)$$

La densité de la loi conditionnelle de (T_1, \dots, T_{n-1}) sachant $[T_n = t_n]$ est :

$$\begin{aligned} f_{(T_1, \dots, T_{n-1})|T_n=t_n}(t_1, \dots, t_{n-1}) &= \frac{f_{(T_1, \dots, T_n)}(t_1, \dots, t_n)}{f_{T_n}(t_n)} \\ &= \frac{\left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t_n)} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}}}{e^{-m(t_n)} \frac{[m(t_n)]^{n-1}}{(n-1)!} \lambda(t_n)} \\ &= (n-1)! \prod_{i=1}^{n-1} \frac{\lambda(t_i)}{m(t_n)} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \quad (3.17) \\ &= (n-1)! \prod_{i=1}^{n-1} \left[\frac{\lambda(t_i)}{m(t_n)} \mathbb{1}_{[0, t_n]}(t_i) \right] \mathbb{1}_{\{t_1 < \dots < t_{n-1}\}} \end{aligned}$$

ce qui prouve le résultat. ■

Passons maintenant à la loi des durées inter-défaillances X_i .

Proposition 11 . *La densité de la loi conditionnelle de X_{n+1} sachant $[T_n = t_n]$ est :*

$$f_{X_{n+1}|T_n=t_n}(x) = \lambda(t_n + x) \exp \left(- \int_0^x \lambda(t_n + s) ds \right) \quad (3.18)$$

Démonstration: Il suffit d'appliquer la proposition 6 :

$$\begin{aligned} f_{X_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) &= \lambda_{t_n+x}(n, t_1, \dots, t_n) \exp \left(- \int_0^x \lambda_{t_n+s}(n, t_1, \dots, t_n) ds \right) \\ &= \lambda(t_n + x) \exp \left(- \int_0^x \lambda(t_n + s) ds \right) \quad (3.19) \end{aligned}$$

Comme seul t_n subsiste dans cette formule, le conditionnement se réduit à un conditionnement par $[T_n = t_n]$. ■

On peut interpréter ce résultat en disant que le taux de défaillance de la loi de X_{n+1} sachant $[T_n = t_n]$ est λ_{t_n+x} .

Puisque le résultat dépend de t_n , cela montre que les durées inter-défaillances ne sont pas indépendantes. C'est normal, en vertu du corollaire 1, car l'intensité n'est pas de la forme $h(N_t, t - T_{N_t})$.

Il n'y a pas d'expression simple pour la loi (marginale) de chaque X_i .

La fonction de répartition de la loi conditionnelle de X_{n+1} sachant $[T_n = t_n]$ est :

$$\begin{aligned} F_{X_{n+1}|T_n=t_n}(x) &= 1 - \exp\left(-\int_0^x \lambda(t_n + s) ds\right) = 1 - \exp\left(-\int_{t_n}^{t_n+x} \lambda(s) ds\right) \\ &= 1 - \frac{\exp\left(-\int_0^{t_n+x} \lambda(s) ds\right)}{\exp\left(-\int_0^{t_n} \lambda(s) ds\right)} \end{aligned} \quad (3.20)$$

Comme λ est le taux de défaillance de l'instant de la première défaillance X_1 , on a $\exp\left(-\int_0^x \lambda(s) ds\right) = \mathbb{P}(X_1 > x)$. Par conséquent, on obtient :

$$F_{X_{n+1}|T_n=t_n}(x) = 1 - \frac{\mathbb{P}(X_1 > t_n + x)}{\mathbb{P}(X_1 > t_n)} = 1 - \mathbb{P}(X_1 > t_n + x | X_1 > t_n) \quad (3.21)$$

ce qui s'écrit aussi :

$$\mathbb{P}(X_{n+1} > x | T_n = t_n) = \mathbb{P}(X_1 > t_n + x | X_1 > t_n) \quad (3.22)$$

Cela correspond bien à une hypothèse As Bad As Old : sachant que la $n^{\text{ème}}$ réparation a eu lieu à l'instant t_n , la probabilité que le système fonctionne bien pendant au moins une durée x après t_n est égale à la probabilité qu'un système neuf ayant fonctionné sans défaillances jusqu'à t_n fonctionne encore pendant une durée au moins égale à x .

3.4 Fiabilité et MTTF

Une simple application de la proposition 7 et du corollaire 3 amène :

$$R_t(\tau) = \exp\left(-\int_t^{t+\tau} \lambda(s) ds\right) = e^{-[m(t+\tau)-m(t)]} \quad (3.23)$$

$$MTTF_t = \int_0^{+\infty} e^{-[m(t+\tau)-m(t)]} d\tau = e^{m(t)} \int_t^{+\infty} e^{-m(s)} ds \quad (3.24)$$

3.5 Fonctions de vraisemblance

D'après (2.54), le cas de la censure de type 2 (observation des n premières défaillances) est directement réglé à l'aide de la loi conjointe de (T_1, \dots, T_n) :

$$\mathcal{L}(\theta; n, t_1, \dots, t_n) = f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) = \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t_n)} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \quad (3.25)$$

Pour la censure de type 1 (observation sur $[0, t]$), il faut appliquer (2.53) :

$$\begin{aligned} \mathcal{L}(\theta; n, t_1, \dots, t_n) &= \left[\prod_{i=1}^n \lambda(t_i) \right] \exp\left(-\sum_{i=1}^{n+1} \int_{t_{i-1}}^{t_i} \lambda(s) ds\right) \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \\ &= \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t)} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \end{aligned} \quad (3.26)$$

Les expressions des vraisemblances pour les deux types de censure sont donc très semblables. Si on a des résultats pour la censure de type 1, il suffira de remplacer t par t_n pour obtenir ceux de la censure de type 2.

Il est possible d'obtenir ces fonctions de vraisemblances par un raisonnement tout à fait différent, qui a le mérite de faire comprendre d'où viennent chacun des termes des expressions obtenues. Par exemple, pour la censure de type 2, on a :

$$\begin{aligned}
& f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) \\
&= \lim_{\Delta t_1 \rightarrow 0, \dots, \Delta t_n \rightarrow 0} \frac{1}{\Delta t_1 \dots \Delta t_n} \mathbb{P}(T_1 \in]t_1, t_1 + \Delta t_1], \dots, T_n \in]t_n, t_n + \Delta t_n]) \\
&= \lim_{\Delta t_1 \rightarrow 0, \dots, \Delta t_n \rightarrow 0} \frac{1}{\Delta t_1 \dots \Delta t_n} \mathbb{P}(N_{t_1} = 0, N_{t_1 + \Delta t_1} - N_{t_1} = 1, N_{t_2} - N_{t_2 + \Delta t_1} = 0, \dots \\
&\quad \dots, N_{t_n} - N_{t_{n-1} + \Delta t_{n-1}} = 0, N_{t_n + \Delta t_n} - N_{t_n} = 1) \\
&= \lim_{\Delta t_1 \rightarrow 0, \dots, \Delta t_n \rightarrow 0} \frac{1}{\Delta t_1 \dots \Delta t_n} e^{-m(t_1)} e^{-[m(t_1 + \Delta t_1) - m(t_1)]} [m(t_1 + \Delta t_1) - m(t_1)] e^{-[m(t_2) - m(t_1 + \Delta t_1)]} \\
&\quad \dots e^{-[m(t_n) - m(t_{n-1} + \Delta t_{n-1})]} e^{-[m(t_n + \Delta t_n) - m(t_n)]} [m(t_n + \Delta t_n) - m(t_n)] \\
&= e^{-m(t_n)} \prod_{i=1}^n \lim_{\Delta t_i \rightarrow 0} \frac{m(t_i + \Delta t_i) - m(t_i)}{\Delta t_i} \\
&= \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t_n)} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \tag{3.27}
\end{aligned}$$

en incluant le fait que les t_i doivent être ordonnés.

Pour la censure de type 1, il faut rajouter l'information qu'il n'y a pas de défaillance entre t_n et t sous la forme :

$$\mathbb{P}(N_t - N_{t_n + \Delta t_n} = 0) = e^{-[m(t) - m(t_n + \Delta t_n)]} \tag{3.28}$$

3.6 Quelques propriétés utiles

Pour terminer ce chapitre, nous donnons quelques propriétés des NHPP qui seront utiles dans une perspective statistique. Ces propriétés concernent essentiellement des lois conditionnelles déterminées à partir du vecteur des instants de défaillance. Elles sont de même nature que le corollaire 5.

Corollaire 6 . Sachant $[N_t = n]$, le vecteur (T_1, \dots, T_n) a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi de densité $\frac{\lambda(s)}{m(t)} \mathbb{1}_{[0, t]}(s)$.

Démonstration: En appliquant la formule de Bayes, on a :

$$\begin{aligned}
f_{(T_1, \dots, T_n) | N_t = n}(t_1, \dots, t_n) &= \frac{\mathbb{P}(N_t = n | T_1 = t_1, \dots, T_n = t_n) f_{(T_1, \dots, T_n)}(t_1, \dots, t_n)}{\mathbb{P}(N_t = n)} \\
&= \frac{\mathcal{L}(\theta; n, t_1, \dots, t_n)}{\mathbb{P}(N_t = n)} \\
&= \frac{\left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t)} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}}}{e^{-m(t)} \frac{[m(t)]^n}{n!}} \\
&= n! \left[\prod_{i=1}^n \frac{\lambda(t_i)}{m(t)} \mathbb{1}_{[0, t]}(t_i) \right] \mathbb{1}_{\{< t_1 < \dots < t_n\}}
\end{aligned} \tag{3.29}$$

ce qui prouve le résultat. ■

Corollaire 7 . Sachant $[T_n = t_n]$, le vecteur $\left(\frac{m(T_1)}{m(T_n)}, \dots, \frac{m(T_{n-1})}{m(T_n)} \right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi uniforme sur $[0, 1]$, $\mathcal{U}[0, 1]$.

Démonstration: On a vu que, sachant $[T_n = t_n]$, le vecteur (T_1, \dots, T_{n-1}) a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi de densité $\frac{\lambda(s)}{m(t_n)} \mathbb{1}_{[0, t_n]}(s)$. La fonction de répartition de cette loi est :

$$G(s) = \begin{cases} 0 & \text{si } s < 0 \\ \frac{m(s)}{m(t_n)} & \text{si } 0 \leq s \leq t_n \\ 1 & \text{si } s > t_n \end{cases} \tag{3.30}$$

On utilise alors le résultat qui dit que si X est une variable aléatoire de fonction de répartition F continue, alors $F(X)$ est de loi $\mathcal{U}[0, 1]$.

Donc le vecteur $(G(T_1), \dots, G(T_{n-1}))$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi $\mathcal{U}[0, 1]$. ■

Corollaire 8 . Le vecteur $\left(\frac{m(T_1)}{m(T_n)}, \dots, \frac{m(T_{n-1})}{m(T_n)} \right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi $\mathcal{U}[0, 1]$ et est indépendant de T_n .

Démonstration: Il suffit de remarquer que, dans le résultat précédent, la loi du vecteur $\left(\frac{m(T_1)}{m(T_n)}, \dots, \frac{m(T_{n-1})}{m(T_n)} \right)$ sachant $[T_n = t_n]$ ne dépend pas de t_n . ■

Corollaire 9 . $-2 \sum_{i=1}^{n-1} \ln \frac{m(T_i)}{m(T_n)}$ est indépendant de T_n et est de loi $\chi_{2(n-1)}^2$.

Démonstration : Si U est de loi $\mathcal{U}[0, 1]$, $-\ln U$ est de loi $\exp(1)$. Donc le vecteur $\left(-\ln \frac{m(T_{n-1})}{m(T_n)}, \dots, -\ln \frac{m(T_1)}{m(T_n)}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n-1$ de la loi $\exp(1)$ (notons que l'ordre des T_i est inversé à cause du signe moins).

Si (Y_1^*, \dots, Y_n^*) est la statistique d'ordre de l'échantillon (Y_1, \dots, Y_n) , $\sum_{i=1}^n Y_i^* = \sum_{i=1}^n Y_i$.

Donc $-\sum_{i=1}^{n-1} \ln \frac{m(T_i)}{m(T_n)}$ a même loi que la somme de $n-1$ variables aléatoires indépendantes et de loi $\exp(1)$, donc est de loi gamma $G(n-1, 1)$.

Par ailleurs, si X est de loi $G(a, b)$, alors pour $c > 0$, cX est de loi $G\left(a, \frac{b}{c}\right)$. Donc $-2 \sum_{i=1}^{n-1} \ln \frac{m(T_i)}{m(T_n)}$ est de loi $G\left(n-1, \frac{1}{2}\right)$, c'est-à-dire de loi $\chi_{2(n-1)}^2$. ■

Symétriquement, on a aussi les résultats suivants :

Corollaire 10 . Sachant $[N_t = n]$, le vecteur $\left(\frac{m(T_1)}{m(t)}, \dots, \frac{m(T_n)}{m(t)}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{U}[0, 1]$.

Cette fois, on ne peut pas se passer du conditionnement car la loi conditionnelle dépend de la variable de conditionnement (N_t) à travers la dimension du vecteur.

Corollaire 11 . Sachant $[N_t = n]$, $-2 \sum_{i=1}^n \ln \frac{m(T_i)}{m(t)}$ est de loi χ_{2n}^2 .

Notons que ce n'est pas le même nombre de degrés de liberté dans les deux corollaires 9 et 11.

Nous avons dans ce chapitre donné les propriétés générales des NHPP. Dans les deux chapitres suivants, nous allons étudier en détail les deux NHPP les plus utilisés :

- Le processus de Poisson homogène, pour lequel $\forall t \geq 0$, $\lambda(t) = \lambda$.
- Le processus de puissance, pour lequel $\forall t \geq 0$, $\lambda(t) = \alpha \beta t^{\beta-1}$.

Chapitre 4

Les processus de Poisson homogènes

4.1 Propriétés probabilistes

Les **Processus de Poisson Homogènes** (HPP pour Homogeneous Poisson Process) sont les plus simples de tous les processus ponctuels. Ils sont définis par une intensité constante :

$$\forall t \geq 0, \lambda_t = \lambda. \quad (4.1)$$

où λ est un paramètre réel positif.

La forme de cette intensité montre que c'est à la fois un NHPP, un processus de renouvellement, un modèle de réparation imparfaite et un modèle DIDE.

La constance de l'intensité fait qu'à tout instant la probabilité de défaillance instantanée est la même. Par conséquent, il n'y a ni dégradation ni amélioration, ni vieillissement ni rajeunissement du système. On dit aussi que la fiabilité est stable ou stabilisée.

Les propriétés des HPP se déduisent très simplement des propriétés générales des NHPP. Passons les en revue, dans l'ordre du chapitre précédent.

4.1.1 Loi du nombre de défaillances survenues à chaque instant

- $\forall t \geq 0, m(t) = \mathbb{E}(N_t) = \int_0^t \lambda ds = \lambda t.$
- $\{N_t\}_{t \geq 0}$ est à accroissements indépendants.
- $\forall t \geq 0, N_t$ est de loi de Poisson $\mathcal{P}(\lambda t)$. $\forall (s, t), 0 \leq s \leq t, N_t - N_s$ est de loi $\mathcal{P}(\lambda(t - s))$.
- $\lim_{t \rightarrow +\infty} m(t) = \lim_{t \rightarrow +\infty} \lambda t = +\infty$, donc

- $\frac{N_t}{\lambda t} \xrightarrow[t \rightarrow +\infty]{p.s.} 1$, ce qui peut s'interpréter en disant que le nombre de défaillances par unité de temps N_t/t tend vers la constante λ .

- $\frac{N_t - \lambda t}{\sqrt{\lambda t}} = \frac{\sqrt{t}}{\sqrt{\lambda}} \left(\frac{N_t}{t} - \lambda \right) \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$, ce qui donne la vitesse de convergence de N_t/t vers λ .

4.1.2 Lois des instants de défaillances

- $\forall n \geq 1$, $f_{T_n}(t) = \lambda e^{-\lambda t} \frac{[\lambda t]^{n-1}}{(n-1)!} = \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1}$. On reconnaît que T_n est de loi gamma $G(n, \lambda)$. En particulier, la durée moyenne d'attente de la $n^{\text{ème}}$ défaillance est donc $\mathbb{E}(T_n) = \frac{n}{\lambda}$.
- $f_{T_{n+1}|T_1=t_1, \dots, T_n=t_n}(t) = f_{T_{n+1}|T_n=t_n}(t) = \lambda e^{-\lambda(t-t_n)} \mathbb{1}_{\{t_n < t\}}$. C'est la densité d'une loi exponentielle décalée.

4.1.3 Lois des durées inter-défaillances

$f_{X_{n+1}|T_n=t_n}(x) = \lambda e^{-\lambda x}$. Comme cela ne dépend pas de t_n , on obtient le résultat fondamental suivant, qui peut servir de définition au processus de Poisson homogène :

Proposition 12 *Le processus des défaillances est un processus de Poisson homogène d'intensité λ si et seulement si les durées inter-défaillances X_i sont indépendantes et de même loi exponentielle $\exp(\lambda)$:*

$$\forall i \geq 1, \forall x \in \mathbb{R}^+, f_{X_i}(x) = \lambda e^{-\lambda x} \quad (4.2)$$

On en déduit que $\forall i \geq 1$, $\mathbb{E}(X_i) = \frac{1}{\lambda}$. λ s'interprète donc comme l'inverse de la durée moyenne entre deux défaillances.

La propriété classique sur les sommes de variables aléatoires indépendantes et de même loi exponentielle permet de retrouver que $T_n = \sum_{i=1}^n X_i$ est de loi $G(n, \lambda)$.

Puisque les X_i sont indépendants et de même loi, le HPP est aussi un processus de renouvellement. On peut aussi dire que λ_t est de la forme $h(t - T_{N_t})$, avec $\forall x, h(x) = \lambda$.

4.1.4 Fiabilité et MTTF

$$\forall t \geq 0, R_t(\tau) = e^{-[\lambda(t+\tau) - \lambda t]} = e^{-\lambda \tau} \quad (4.3)$$

La fiabilité est donc indépendante de l'instant auquel on la calcule, ce qui confirme la stabilité de la fiabilité au cours du temps.

C'est la propriété d'absence de mémoire de la loi exponentielle qui explique ce phénomène : calculer la fiabilité en t , c'est comme la calculer en T_{N_t} . Or $T_{N_{t+1}} - T_{N_t} = X_{N_t+1}$ est de loi $\exp(\lambda)$, d'où le résultat.

$$\forall t \geq 0, MTTF_t = \int_0^{+\infty} R_t(\tau) d\tau = \int_0^{+\infty} e^{-\lambda \tau} d\tau = \frac{1}{\lambda} \quad (4.4)$$

$MTTF_t$ est indépendant de t pour les mêmes raisons que précédemment. Par conséquent, quel que soit l'instant auquel on se place, la durée moyenne d'attente de la prochaine défaillance est $\frac{1}{\lambda}$. Ce phénomène est parfois connu sous le nom de **paradoxe du bus**.

Plutôt que la durée moyenne d'attente de la prochaine défaillance, il est souvent intéressant de considérer la durée médiane d'attente de la prochaine défaillance. A l'instant t , cette durée médiane est notée $MedTTF_t$ et est définie pour un HPP par :

$$\mathbb{P}(T_{N_{t+1}} - t \leq MedTTF_t) = \frac{1}{2} \quad (4.5)$$

Autrement dit, il y a une chance sur deux que la prochaine défaillance survienne avant $MedTTF_t$ et une chance sur deux qu'elle survienne après.

La propriété d'absence de mémoire de la loi exponentielle fait qu'ici :

$$\begin{aligned} \mathbb{P}(X_{N_{t+1}} \leq MedTTF_t) &= \mathbb{P}(T_{N_{t+1}} - t \leq MedTTF_t) = \frac{1}{2} = 1 - e^{-\lambda MedTTF_t} \\ \Rightarrow e^{-\lambda MedTTF_t} &= \frac{1}{2} \Rightarrow -\lambda MedTTF_t = \ln \frac{1}{2} = -\ln 2 \\ \Rightarrow MedTTF_t &= \frac{\ln 2}{\lambda} \approx 0.69 MTF_t \end{aligned} \quad (4.6)$$

C'est un phénomène caractéristique de la loi exponentielle: les pannes surviennent dans l'ensemble assez tôt, mais certaines se font attendre très longtemps, ce qui tire la moyenne vers le haut. Il est donc utile de compléter l'information sur la moyenne par une information sur la médiane.

4.1.5 Fonctions de vraisemblance

- Censure de type 1 (observation sur $[0, t]$):

$$\mathcal{L}(\lambda; n, t_1, \dots, t_n) = \lambda^n e^{-\lambda t} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \quad (4.7)$$

- Censure de type 2 (observation des n premières défaillances):

$$\mathcal{L}(\lambda; t_1, \dots, t_n) = \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \quad (4.8)$$

4.1.6 Quelques propriétés utiles

$\frac{\lambda(s)}{m(t)} = \frac{\lambda}{\lambda t} = \frac{1}{t}$ et $\frac{m(s)}{m(t)} = \frac{\lambda s}{\lambda t} = \frac{s}{t} \cdot \frac{1}{t} \mathbb{1}_{[0, t]}(s)$ est la densité de la loi uniforme sur $[0, t]$.

Donc les corollaires 5 à 11 s'écrivent pour les HPP de la façon suivante :

- Sachant $[T_n = t_n]$, le vecteur (T_1, \dots, T_{n-1}) a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi $\mathcal{U}[0, t_n]$.
- Sachant $[N_t = n]$, le vecteur (T_1, \dots, T_n) a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{U}[0, t]$.
- Le vecteur $\left(\frac{T_1}{T_n}, \dots, \frac{T_{n-1}}{T_n}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi uniforme $\mathcal{U}[0, 1]$ et est indépendant de T_n .

- $-2 \sum_{i=1}^{n-1} \ln \frac{T_i}{T_n}$ est indépendant de T_n et est de loi $\chi_{2(n-1)}^2$.
- Sachant $[N_t = n]$, le vecteur $\left(\frac{T_1}{t}, \dots, \frac{T_n}{t}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{U}[0, 1]$.
- Sachant $[N_t = n]$, $-2 \sum_{i=1}^n \ln \frac{T_i}{t}$ est de loi χ_{2n}^2 .

4.2 Lien entre processus de Poisson homogène et non homogène - Simulation des NHPP

Proposition 13 . Soit $\{N_t\}_{t \geq 0}$ un NHPP d'intensité $\lambda(t)$ et d'intensité cumulée $m(t) = \int_0^t \lambda(s) ds$. On note $\{M_t\}_{t \geq 0}$ le processus de comptage d'événements survenant aux instants $0 \leq m(T_1) \leq m(T_2) \leq \dots$. Alors $\{M_t\}_{t \geq 0}$ est un processus de Poisson homogène d'intensité 1.

Démonstration: Par définition, $M_t = \sum_{i=1}^{+\infty} \mathbb{1}_{\{m(T_i) \leq t\}}$. m est croissante mais pas forcément inversible. On considère donc l'inverse généralisée ou pseudo-inverse de m , la fonction m^- définie par :

$$m^-(u) = \text{Inf}\{t; m(t) \geq u\} \quad (4.9)$$

Si m est inversible, $m^- = m^{-1}$.

Alors $\{m(T_i) \leq t\} \iff \{T_i \leq m^-(t)\}$. Donc $M_t = \sum_{i=1}^{+\infty} \mathbb{1}_{\{T_i \leq m^-(t)\}} = N_{m^-(t)}$.

Par conséquent, $\forall t > 0$, M_t est de loi de Poisson $\mathcal{P}(m[m^-(t)]) = \mathcal{P}(t)$.

Il est facile de voir que les autres propriétés du HPP d'intensité 1 sont vérifiées : $\{M_t\}_{t \geq 0}$ est simple, à accroissements indépendants et $M_0 = 0$ p.s. Comme $t = \int_0^t 1 ds$, $\{M_t\}_{t \geq 0}$ est bien un processus de Poisson homogène d'intensité 1. ■

En inversant le point de vue, on en déduit immédiatement le corollaire suivant :

Corollaire 12 . Si Y_1, \dots, Y_n sont les instants d'occurrence d'un processus de Poisson homogène d'intensité 1, alors $m^-(Y_1), \dots, m^-(Y_n)$ sont les instants d'occurrence d'un NHPP d'intensité $\lambda(t) = m'(t)$.

Par conséquent, pour simuler un NHPP d'intensité $\lambda(t)$, on simule une suite de variables aléatoires $\{Z_i\}_{i \geq 1}$ indépendantes et de même loi $\exp(1)$. On prend alors $Y_i = \sum_{j=1}^i Z_j$ et les instants d'occurrence du NHPP sont les $T_i = m^-(Y_i)$.

Cette propriété permet d'obtenir de nouveaux résultats généraux sur les NHPP :

Proposition 14 . Si $\{N_t\}_{t \geq 0}$ est un NHPP d'intensité $\lambda(t)$ alors :

1. Les $m(T_i) - m(T_{i-1})$ sont indépendants et de même loi $\exp(1)$.
2. $m(T_n)$ est de loi $G(n, 1)$ et $2m(T_n)$ est de loi χ_{2n}^2 .

Démonstration: 1. est obtenu directement en disant que les $m(T_i)$ sont les instants d'occurrence d'un HPP d'intensité 1.

2. $m(T_n) = \sum_{i=1}^n (m(T_i) - m(T_{i-1}))$ est de loi $G(n, 1)$ comme somme de variables aléatoires indépendantes et de même loi $\exp(1)$, et $2m(T_n)$ est de loi $G(n, 1/2) = \chi_{2n}^2$. ■

On peut de la même façon faire le lien entre deux processus de Poisson non homogènes quelconques :

Proposition 15 . Si $\{N_t\}_{t \geq 0}$ est un NHPP d'intensité cumulée $m_1(t)$, alors $\{N_{m_1^{-1} \circ m_2(t)}\}_{t \geq 0}$ est un NHPP d'intensité cumulée $m_2(t)$.

4.3 Estimation des grandeurs de la fiabilité

Le problème qui se pose est l'estimation de λ au vu des observations, ainsi que du $MTTF = 1/\lambda$ et de la fiabilité $R(\tau) = e^{-\lambda\tau}$. On peut faire une estimation ponctuelle ou ensembliste par intervalles de confiance. Enfin, on peut faire des tests portant sur λ . Comme la procédure d'estimation dépend du type de données, c'est-à-dire du type de censure, on va distinguer les deux cas.

4.3.1 Censure de type 2

On observe les instants T_1, \dots, T_n des n premières défaillances. La fonction de vraisemblance est donnée par (4.8) :

$$\mathcal{L}(\lambda; t_1, \dots, t_n) = \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0 < t_1 < \dots < t_n\}} \quad (4.10)$$

L'estimateur de maximum de vraisemblance de λ est la valeur de λ qui maximise la fonction de vraisemblance. Comme λ n'intervient pas dans l'indicatrice, on peut ne pas en tenir compte dans la maximisation.

Alors $\ln \mathcal{L}(\lambda; t_1, \dots, t_n) = n \ln \lambda - \lambda t_n$.

$$\frac{\partial}{\partial \lambda} \ln \mathcal{L}(\lambda; t_1, \dots, t_n) = \frac{n}{\lambda} - t_n, \text{ qui s'annule pour } \lambda = \frac{n}{t_n}.$$

Par conséquent, l'estimateur de maximum de vraisemblance de λ est

$$\hat{\lambda}_n = \frac{n}{T_n} \quad (4.11)$$

Il faut déterminer la qualité de cet estimateur : est-il sans biais, convergent, efficace?

On sait que T_n est de loi $G(n, \lambda)$. Donc on a :

$$\begin{aligned}\mathbb{E}[\hat{\lambda}_n] &= \mathbb{E}\left[\frac{n}{T_n}\right] = n \int_0^{+\infty} \frac{1}{t} \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} dt \\ &= n \int_0^{+\infty} \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-2} dt = \frac{n}{n-1} \lambda \int_0^{+\infty} \frac{\lambda^{n-1}}{(n-2)!} e^{-\lambda t} t^{n-2} dt \quad (4.12) \\ &= \frac{n}{n-1} \lambda \int_0^{+\infty} f_{G(n-1, \lambda)}(t) dt = \frac{n}{n-1} \lambda\end{aligned}$$

Comme $\mathbb{E}[\hat{\lambda}_n] \neq \lambda$, $\hat{\lambda}_n$ est un estimateur biaisé de λ . Mais on a évidemment $\mathbb{E}\left[\frac{n-1}{n} \hat{\lambda}_n\right] = \lambda$.

Donc $\hat{\lambda}'_n = \frac{n-1}{n} \hat{\lambda}_n = \frac{n-1}{T_n}$ est un estimateur sans biais de λ . Calculons sa variance.

$$\begin{aligned}Var[\hat{\lambda}'_n] &= \mathbb{E}[\hat{\lambda}'_n{}^2] - \left(\mathbb{E}[\hat{\lambda}'_n]\right)^2 = (n-1)^2 \mathbb{E}\left[\frac{1}{T_n^2}\right] - \lambda^2 \\ &= (n-1)^2 \int_0^{+\infty} \frac{1}{t^2} \frac{\lambda^n}{(n-1)!} e^{-\lambda t} t^{n-1} dt - \lambda^2 \quad (4.13)\end{aligned}$$

$$\begin{aligned}&= (n-1)^2 \frac{\lambda^2}{(n-1)(n-2)} \int_0^{+\infty} \frac{\lambda^{n-2}}{(n-3)!} e^{-\lambda t} t^{n-3} dt - \lambda^2 \\ &= \frac{n-1}{n-2} \lambda^2 \int_0^{+\infty} f_{G(n-2, \lambda)}(t) dt - \lambda^2 \quad (4.14) \\ &= \frac{n-1}{n-2} \lambda^2 - \lambda^2 = \frac{\lambda^2}{n-2}\end{aligned}$$

$\lim_{n \rightarrow +\infty} Var[\hat{\lambda}'_n] = 0$, donc $\hat{\lambda}'_n$ est un estimateur convergent en moyenne quadratique (en fait ce n'est pas la peine de faire ce calcul car c'est une propriété générale des estimateurs de maximum de vraisemblance).

La quantité d'information associée à λ est :

$$\mathcal{I}_n(\lambda) = Var\left[\frac{\partial}{\partial \lambda} \ln \mathcal{L}(\lambda; T_1, \dots, T_n)\right] = Var\left[\frac{n}{\lambda} - T_n\right] = Var(T_n) = \frac{n}{\lambda^2} \quad (4.15)$$

Donc l'efficacité de l'estimateur proposé est :

$$Eff[\hat{\lambda}'_n] = \frac{\left[\frac{\partial}{\partial \lambda} \mathbb{E}[\hat{\lambda}'_n]\right]^2}{\mathcal{I}_n(\lambda) Var[\hat{\lambda}'_n]} = \frac{1}{\frac{n}{\lambda^2} \frac{\lambda^2}{n-2}} = \frac{n-2}{n} \quad (4.16)$$

$Eff[\hat{\lambda}'_n] \neq 1$, donc $\hat{\lambda}'_n$ n'est pas un estimateur efficace de λ , mais $\lim_{n \rightarrow +\infty} Eff[\hat{\lambda}'_n] = 1$, donc $\hat{\lambda}'_n$ est asymptotiquement efficace.

D'après le théorème de Rao-Blackwell, l'estimateur sans biais et de variance minimale (ESBVM) de λ est la projection d'un estimateur sans biais sur une statistique exhaustive et complète. La forme de la fonction de vraisemblance montre que T_n est une statistique

exhaustive. Comme T_n est de loi $G(n, \lambda)$, il est facile de montrer que T_n est complète. Donc l'ESBVM de λ est $\mathbb{E}[\hat{\lambda}'_n | T_n] = \hat{\lambda}'_n$. D'où finalement :

Proposition 16 . $\hat{\lambda}'_n = \frac{n-1}{T_n}$ est l'estimateur sans biais et de variance minimale de λ .

Pour déterminer des intervalles de confiance pour λ , il faut trouver une fonction pivotale, c'est-à-dire une fonction de λ et des observations T_1, \dots, T_n dont la loi de probabilité est indépendante de λ .

En appliquant la partie 2. de la proposition (14) au HPP, pour lequel $m(t) = \lambda t$, on obtient que $2m(T_n) = 2\lambda T_n$ est de loi χ_{2n}^2 . Donc $2\lambda T_n$ est une fonction pivotale.

En fiabilité, on est souvent intéressé par des intervalles de confiance unilatéraux : on veut garantir que le taux de défaillance λ n'est pas plus grand qu'une certaine quantité, donc on souhaite avoir un intervalle de confiance de la forme $[0, \lambda_{sup}]$. On écrit alors :

$$\mathbb{P}(\lambda \in [0, \lambda_{sup}]) = \mathbb{P}(\lambda \leq \lambda_{sup}) = \mathbb{P}(2\lambda T_n \leq 2\lambda_{sup} T_n) = F_{\chi_{2n}^2}(2\lambda_{sup} T_n) \quad (4.17)$$

Pour que cette probabilité soit égale à $1 - \alpha$, il suffit de prendre $\lambda_{sup} = \frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}(1 - \alpha)$. On détermine de la même façon des intervalles de confiance de la forme $[\lambda_{inf}, +\infty[$ et $[\lambda_{inf}, \lambda_{sup}]$, et on obtient :

Proposition 17 . Des intervalles de confiance de seuil α pour λ sont :

$$\left[0, \frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}(1 - \alpha)\right], \quad \left[\frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}(\alpha), +\infty\right], \quad \left[\frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}\left(\frac{\alpha}{2}\right), \frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}\left(1 - \frac{\alpha}{2}\right)\right]$$

A partir des intervalles de confiance, il est facile de construire des tests d'hypothèses portant sur λ . Par exemple, il est usuel de considérer que le système étudié répond bien aux normes si son intensité de défaillance λ est inférieure à une quantité fixée λ_0 . Pour savoir au vu de données si on peut admettre avec un bon niveau de confiance que le système est aux normes, il faut tester $H_0 : \lambda \geq \lambda_0$ contre $H_1 : \lambda < \lambda_0$.

En utilisant la dualité entre tests d'hypothèses et intervalles de confiance, on rejettera l'hypothèse H_0 au profit de H_1 au seuil α si et seulement si λ_0 n'appartient pas à l'intervalle de confiance de seuil α de la forme $[0, \lambda_{sup}]$. Cela se produit si $\lambda_0 > \frac{1}{2T_n} F_{\chi_{2n}^2}^{-1}(1 - \alpha)$. Donc un tel test aura pour région critique

$$W = \left\{2\lambda_0 T_n > F_{\chi_{2n}^2}^{-1}(1 - \alpha)\right\} = \left\{\hat{\lambda}'_n < \frac{2(n-1)\lambda_0}{F_{\chi_{2n}^2}^{-1}(1 - \alpha)}\right\} \quad (4.19)$$

On sait que l'estimateur de maximum de vraisemblance est convergent presque sûrement. C'est évidemment aussi le cas de $\hat{\lambda}'_n$. Enfin, l'estimateur de maximum de vraisemblance est asymptotiquement gaussien et efficace. Plus précisément :

$$\sqrt{\mathcal{I}_n(\lambda)} (\hat{\lambda}_n - \lambda) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad (4.20)$$

ce qui se traduit ici par :

$$\sqrt{n} \left(\frac{\hat{\lambda}_n}{\lambda} - 1 \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{et} \quad \sqrt{n} \left(\frac{\hat{\lambda}'_n}{\lambda} - 1 \right) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad (4.21)$$

Plutôt qu'estimer λ , on peut être amenés à estimer une fonction de λ , par exemple la fiabilité ou le *MTTF*.

Comme $MTTF = \frac{1}{\lambda}$, l'estimateur de maximum de vraisemblance du *MTTF* est $\frac{1}{\hat{\lambda}_n} = \frac{T_n}{n} = \frac{1}{n} \sum_{i=1}^n X_i$. Il est logique d'estimer la durée moyenne entre défaillances par la moyenne empirique des durées inter-défaillances et on sait que c'est un excellent estimateur.

La fiabilité à l'instant τ est $R(\tau) = e^{-\lambda\tau}$. Donc l'estimateur de maximum de vraisemblance de $R(\tau)$ est $\hat{R}_n(\tau) = e^{-\hat{\lambda}_n\tau}$. Pour déterminer le biais de cet estimateur, il faut écrire :

$$\mathbb{E} \left[e^{-\hat{\lambda}_n\tau} \right] = \mathbb{E} \left[e^{-\frac{n\tau}{T_n}} \right] = \int_0^{+\infty} e^{-\frac{n\tau}{u}} \frac{\lambda^n}{(n-1)!} e^{-\lambda u} u^{n-1} du \quad (4.22)$$

Il n'est vraiment pas facile de faire ce calcul, donc on ne peut pas calculer le biais de l'estimateur de maximum de vraisemblance de la fiabilité. Mais ce n'est pas un problème car on peut en fait trouver directement l'estimateur optimal de $R(\tau)$ en utilisant le théorème de Rao-Blackwell : il faut projeter un estimateur sans biais sur la statistique exhaustive et complète T_n .

Or on a très facilement un estimateur sans biais, $\mathbb{1}_{\{T_1 > \tau\}}$. En effet, $\mathbb{E} [\mathbb{1}_{\{T_1 > \tau\}}] = \mathbb{P}(T_1 > \tau) = e^{-\lambda\tau} = R(\tau)$.

Donc l'ESBVM de $R(\tau)$ est $\tilde{R}_n(\tau) = \mathbb{E} [\mathbb{1}_{\{T_1 > \tau\}} | T_n] = \mathbb{P}(T_1 > \tau | T_n)$.

Or on a vu que la loi conditionnelle du vecteur (T_1, \dots, T_{n-1}) sachant $[T_n = t_n]$ est la loi de la statistique d'ordre d'un échantillon de taille $n-1$ de la loi $\mathcal{U}[0, t_n]$. Donc la loi de T_1 sachant $[T_n = t_n]$ est la loi du minimum d'un tel échantillon.

La fonction de répartition de la loi $\mathcal{U}[0, t_n]$ est :

$$G(\tau) = \begin{cases} 0 & \text{si } \tau < 0 \\ \frac{\tau}{t_n} & \text{si } \tau \in [0, t_n] \\ 1 & \text{si } \tau > t_n \end{cases} \quad (4.23)$$

La fonction de répartition de la loi du minimum d'un échantillon de taille $n-1$ de cette loi est $1 - [1 - G(\tau)]^{n-1}$. Comme $\tau > 0$, on en déduit que :

$$\mathbb{P}(T_1 > \tau | T_n = t_n) = \begin{cases} \left(1 - \frac{\tau}{t_n}\right)^{n-1} & \text{si } \tau \in [0, t_n] \\ 0 & \text{si } \tau > t_n \end{cases} \quad (4.24)$$

Et on obtient finalement :

Proposition 18 . $\forall \tau \geq 0$, l'estimateur sans biais et de variance minimale de $R(\tau) = e^{-\lambda\tau}$ est :

$$\tilde{R}_n(\tau) = \begin{cases} \left(1 - \frac{\tau}{T_n}\right)^{n-1} & \text{si } T_n > \tau \\ 0 & \text{si } T_n \leq \tau \end{cases} \quad (4.25)$$

A partir d'un intervalle de confiance pour λ , on a immédiatement un intervalle de confiance pour $e^{-\lambda\tau}$ puisque

$$\mathbb{P}(a \leq \lambda \leq b) = 1 - \alpha \Rightarrow \mathbb{P}(e^{-b\tau} \leq e^{-\lambda\tau} \leq e^{-a\tau}) = 1 - \alpha \quad (4.26)$$

Par exemple, un intervalle de confiance unilatéral de seuil α pour $R(\tau)$ est :

$$\left[e^{-\frac{\tau}{2T_n} F_{\chi_{2n}^2}^{-1}(1-\alpha)}, 1 \right] \quad (4.27)$$

4.3.2 Censure de type 1

On observe le processus sur un intervalle de temps fixé $[0, t]$. L'observation est alors le nombre N_t de défaillances observées entre 0 et t et les instants T_1, \dots, T_n de ces défaillances. La fonction de vraisemblance est donnée par (4.7) :

$$\mathcal{L}(\lambda; n, t_1, \dots, t_n) = \lambda^n e^{-\lambda t} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} \quad (4.28)$$

Là encore, on peut ne pas tenir compte de l'indicatrice dans les calculs.

Alors $\ln \mathcal{L}(\lambda; n, t_1, \dots, t_n) = n \ln \lambda - \lambda t$.

$\frac{\partial}{\partial \lambda} \ln \mathcal{L}(\lambda; n, t_1, \dots, t_n) = \frac{n}{\lambda} - t$, qui s'annule pour $\lambda = \frac{n}{t}$.

Par conséquent, l'estimateur de maximum de vraisemblance de λ est

$$\hat{\lambda}_t = \frac{N_t}{t} \quad (4.29)$$

Les calculs sont très proches des précédents, mais pour une censure de type 2 le nombre de défaillances observées n est déterministe et l'instant de fin d'observation T_n est aléatoire, alors que pour une censure de type 1, le nombre de défaillances observées N_t est aléatoire et l'instant de fin d'observation t est déterministe. Cela change radicalement les propriétés des estimateurs.

On sait que N_t est de loi de Poisson $\mathcal{P}(\lambda t)$, donc :

$$\mathbb{E}[\hat{\lambda}_t] = \frac{\lambda t}{t} = \lambda \quad (4.30)$$

$$\text{Var}[\hat{\lambda}_t] = \frac{\lambda t}{t^2} = \frac{\lambda}{t} \quad (4.31)$$

$$\begin{aligned} \mathcal{I}_t(\lambda) &= \text{Var} \left[\frac{\partial}{\partial \lambda} \ln \mathcal{L}(\lambda; N_t, T_1, \dots, T_n) \right] = \text{Var} \left[\frac{N_t}{\lambda} - t \right] \\ &= \frac{\text{Var}(N_t)}{\lambda^2} = \frac{\lambda t}{t^2} = \frac{\lambda}{t} = \text{Var}[\hat{\lambda}_t] \end{aligned} \quad (4.32)$$

$$\text{d'où } \text{Eff}[\hat{\lambda}_t] = 1.$$

Par conséquent :

Proposition 19 . $\hat{\lambda}_t = \frac{N_t}{t}$ est l'estimateur sans biais et de variance minimale de λ . Il est efficace.

Par rapport à la censure de type 2, on voit que l'estimateur de maximum de vraisemblance est directement l'ESBVM, et qu'en plus il est efficace, alors qu'il n'existait pas d'estimateur efficace dans le cas précédent.

Les résultats asymptotiques sont aussi immédiats :

- $\frac{N_t}{t} \xrightarrow[t \rightarrow +\infty]{p.s.} \lambda$, donc $\hat{\lambda}_t \xrightarrow[t \rightarrow +\infty]{p.s.} \lambda$.
- $\frac{\sqrt{t}}{\sqrt{\lambda}} \left(\frac{N_t}{t} - \lambda \right) \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$, donc $\frac{\sqrt{t}}{\sqrt{\lambda}} (\hat{\lambda}_t - \lambda) \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1)$.

Les choses se compliquent quand on s'intéresse aux intervalles de confiance. En effet, il n'y a pas de fonction pivotale facile à trouver sur la base de N_t de loi $\mathcal{P}(t)$, parce que la loi de Poisson est une loi discrète.

Il s'avère que la façon la plus simple de construire un intervalle de confiance est d'utiliser la dualité entre intervalle de confiance et tests d'hypothèses. Par exemple, un intervalle de confiance de seuil α du type $[0, \lambda_{sup}]$ est constitué de toutes les valeurs de λ_0 pour lesquelles on ne rejettera pas, au seuil α , l'hypothèse $H_0 : \lambda \geq \lambda_0$ au profit de $H_1 : \lambda < \lambda_0$.

Un tel test consiste à rejeter H_0 si $\hat{\lambda}_t$ est "trop petit" ou, ce qui revient au même, si N_t est "trop petit". La région critique est donc de la forme $W = \{N_t \leq k_\alpha\}$. Comme N_t est entier, on peut également prendre k_α entier.

On détermine k_α en disant que le seuil du test est :

$$\begin{aligned} \alpha &= \sup_{H_0} \mathbb{P}(N_t \in W) = \sup_{\lambda \geq \lambda_0} \mathbb{P}(N_t \leq k_\alpha) = \sup_{\lambda \geq \lambda_0} \mathbb{P}(T_{k_\alpha+1} > t) \\ &= \sup_{\lambda \geq \lambda_0} [1 - \mathbb{P}(2\lambda T_{k_\alpha+1} \leq 2\lambda t)] = \sup_{\lambda \geq \lambda_0} \left[1 - F_{\chi_{2(k_\alpha+1)}^2}(2\lambda t) \right] \\ &= 1 - F_{\chi_{2(k_\alpha+1)}^2}(2\lambda_0 t) \end{aligned} \quad (4.33)$$

car le terme entre crochets est une fonction décroissante de λ .

A λ_0 fixé, on rejette H_0 ssi $N_t \leq k_\alpha$, où k_α est tel que $F_{\chi^2_{2(k_\alpha+1)}}(2\lambda_0 t) = 1 - \alpha$. Notons qu'il n'existe pas forcément de k_α entier qui vérifie cette égalité. On prendra alors pour k_α le plus petit entier tel que $F_{\chi^2_{2(k_\alpha+1)}}(2\lambda_0 t) \geq 1 - \alpha$.

Maintenant, renversons le point de vue. A N_t fixé, les valeurs de λ_0 pour lesquelles on ne rejettera pas H_0 sont les valeurs de λ_0 pour lesquelles $N_t > k_\alpha$. Or $F_{\chi^2_{2(n+1)}}(2\lambda t) = 1 - \mathbb{P}(N_t \leq n)$ est une fonction décroissante de n , donc

$$N_t > k_\alpha \iff F_{\chi^2_{2(N_t+1)}}(2\lambda_0 t) < F_{\chi^2_{2(k_\alpha+1)}}(2\lambda_0 t) = 1 - \alpha \iff \lambda_0 < \frac{1}{2t} F_{\chi^2_{2(N_t+1)}}^{-1}(1 - \alpha) \quad (4.34)$$

ce qui fournit la borne supérieure de l'intervalle de confiance cherché. Comme on n'a pas forcément $F_{\chi^2_{2(k_\alpha+1)}}(2\lambda_0 t) = 1 - \alpha$, le test et l'intervalle de confiance proposés ne sont pas de seuil exactement égal à α , mais inférieur ou égal à α (on dit qu'ils sont conservatifs).

Pour le second intervalle de confiance unilatéral, on procède de la même manière en utilisant le fait que $\mathbb{P}(N_t \leq n) = P(T_n \geq t)$. Pour l'intervalle de confiance bilatéral, on part du test de " $\lambda = \lambda_0$ " contre " $\lambda \neq \lambda_0$ ", donc la région critique est de la forme $W = \{N_t \leq k_{1,\alpha} \text{ ou } N_t \geq k_{2,\alpha}\}$. Finalement, on a les résultats suivants :

Proposition 20 . Des intervalles de confiance de seuil inférieur ou égal à α pour λ sont :

$$\left[0, \frac{1}{2t} F_{\chi^2_{2(N_t+1)}}^{-1}(1 - \alpha) \right], \quad \left[\frac{1}{2t} F_{\chi^2_{2N_t}}^{-1}(\alpha), +\infty \right], \quad \left[\frac{1}{2t} F_{\chi^2_{2N_t}}^{-1}\left(\frac{\alpha}{2}\right), \frac{1}{2t} F_{\chi^2_{2(N_t+1)}}^{-1}\left(1 - \frac{\alpha}{2}\right) \right]$$

Il reste enfin à estimer la fiabilité $R(\tau) = e^{-\lambda\tau}$. L'estimateur de maximum de vraisemblance est $\hat{R}_t(\tau) = e^{-\hat{\lambda}_t\tau} = e^{-N_t\tau/t}$. L'estimateur optimal est la projection de l'estimateur sans biais $\mathbb{1}_{\{T_1 > \tau\}}$ sur la statistique exhaustive et complète N_t : $\tilde{R}_t(\tau) = \mathbb{E}[\mathbb{1}_{\{T_1 > \tau\}} | N_t] = \mathbb{P}(T_1 > \tau | N_t)$.

Or on sait que la loi conditionnelle du vecteur (T_1, \dots, T_n) sachant $[N_t = n]$ est la loi de la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{U}[0, t]$. Donc la loi de T_1 sachant $[N_t = n]$ est la loi du minimum d'un tel échantillon. Un calcul analogue à celui effectué dans la section précédente permet d'obtenir le résultat suivant :

Proposition 21 . $\forall \tau \geq 0$, l'estimateur sans biais et de variance minimale de $R(\tau) = e^{-\lambda\tau}$ est :

$$\tilde{R}_t(\tau) = \begin{cases} \left(1 - \frac{\tau}{t}\right)^{N_t} & \text{si } t > \tau \\ 0 & \text{si } t \leq \tau \end{cases} \quad (4.36)$$

4.4 Tests d'adéquation

4.4.1 Problématique

Toutes les estimations effectuées dans la section précédente permettent de faire des évaluations et des prévisions de fiabilité à partir de données. Ainsi, sur les données des

exemples, on obtient les résultats suivants :

- **jeu de données 1** : il s'agit d'une censure de type 2, à $n = 136$ défaillances. Les données sont les durées inter-défaillances x_1, \dots, x_n . L'instant de la $n^{\text{ème}}$ défaillance est $t_n = 91982$ s.

L'ESBVM de λ est $\hat{\lambda}'_n = \frac{n-1}{t_n} = 1.468 \cdot 10^{-3}$.

L'ESBVM du *MTTF* est $\bar{x}_n = 676$ s.

L'ESBVM de la fiabilité à 1000 s est $\tilde{R}_n(1000) = \left(1 - \frac{1000}{91982}\right)^{135} = 22.9\%$.

On estime donc que la durée moyenne d'attente de la prochaine défaillance du logiciel est de 676 secondes CPU et que la probabilité qu'elle survienne après 1000 secondes CPU n'est que de 22.9%.

- **jeu de données 2** : il s'agit d'une censure de type 1, à $t = 129$ mois. Les données sont le nombre $n = 6$ de défaillances observées sur cette période et les durées inter-défaillances x_1, \dots, x_n .

L'ESBVM de λ est $\hat{\lambda}_t = \frac{n}{t} = 4.65 \cdot 10^{-2}$.

L'ESBVM du *MTTF* est $\bar{x}_n = 21$ mois.

L'ESBVM de la fiabilité à 6 mois est $\tilde{R}_t(6) = \left(1 - \frac{6}{129}\right)^6 = 75.1\%$.

On estime donc que la durée moyenne d'attente de la prochaine défaillance de la pompe est de 21 mois et qu'il y a 75.1% de chances que la pompe fonctionne sans panne pendant encore au moins 6 mois.

Mais ces résultats numériques ne sont valides que si les données sont bien issues d'un processus de Poisson homogène, c'est-à-dire si les systèmes observés ne vieillissent pas et ne s'améliorent pas. Or a priori, le logiciel s'améliore et la pompe s'use. On voit pratiquement à vue d'oeil que le *MTTF* du logiciel est sûrement supérieur à 676 s et que celui de la pompe est sûrement inférieur à 21 mois.

Il est donc important de déterminer si des observations peuvent être considérées comme étant issues d'un processus de Poisson homogène. La méthode pour y parvenir est un **test d'adéquation**.

Il s'agit d'un test de l'hypothèse nulle

$$H_0 : \text{“les observations sont issues d'un HPP”}$$

contre l'hypothèse alternative

$$H_1 : \text{“les observations ne sont pas issues d'un HPP”}.$$

Parfois, on s'intéresse à des alternatives particulières. Par exemple, un **test de tendance** consiste à tester H_0 contre :

$$H_2 : \text{“il y a croissance de fiabilité”}$$

ou

$$H_3 : \text{“il y a décroissance de fiabilité”}.$$

On s'attend à rejeter H_0 au profit de H_2 pour le logiciel et au profit de H_3 pour la pompe.

Quand la censure est de type 2, les observations sont les durées inter-défaillances X_1, \dots, X_n qui, sous H_0 , sont indépendantes et de même loi exponentielle. Donc l'hypothèse nulle peut s'écrire :

$$H_0: "X_1, \dots, X_n \text{ sont indépendantes et de même loi exponentielle}"$$

Ainsi, un test d'adéquation au HPP est identique à un test d'adéquation d'un échantillon à la loi exponentielle. Il existe quantité de méthodes pour cela : test de Kolmogorov-Smirnov, test du χ^2 , etc... Mais l'inconvénient de ces méthodes est qu'elles ne prennent pas en compte la spécificité des observations, qui est leur séquentialité : les X_i sont observées les unes après les autres, donc leur ordre d'apparition est important.

Pour illustrer ce point, considérons le cas d'école suivant, extrait d'Ascher-Feingold [1] : on a 3 systèmes réparables dont les durées inter-défaillance sont exactement identiques, mais ne sont pas dans le même ordre :

Système 1 : 15 27 32 43 51 65 177

Système 2 : 177 65 51 43 32 27 15

Système 3 : 51 43 27 177 15 65 32

Un test qui ne prend pas en compte l'ordre des durées inter-défaillances, comme celui de Kolmogorov-Smirnov, considèrera que ces 3 cas de figure sont identiques, alors que manifestement, il faut conclure à la croissance de fiabilité pour le système 1, la décroissance de fiabilité pour le système 2, et la situation n'est pas claire pour le système 3.

Aussi allons-nous nous intéresser ici à deux tests qui permettent de discriminer ces situations : le test de Laplace et le test logarithmique. On va pour cela distinguer les deux types de censure.

4.4.2 Censure de type 2

Les deux tests de tendance que nous allons traiter sont basés sur la même remarque, qui se visualise bien sur les systèmes 1, 2 et 3. Ces 3 systèmes ont la même valeur de t_n .

Pour le système 1 (croissance de fiabilité), les x_i sont de plus en plus grands, donc les $t_i, i \in \{1, \dots, n-1\}$, sont plutôt petits par rapport à t_n , donc $\sum_{i=1}^{n-1} t_i$ est "plutôt petit". Pour le système 2 (décroissance de fiabilité), les x_i sont de plus en plus petits, donc les t_i sont plutôt grands, donc $\sum_{i=1}^{n-1} t_i$ est "plutôt grand".

4.4.2.1. Le test de Laplace

Le test de Laplace consiste à rejeter H_0 au profit de H_2 si $\sum_{i=1}^{n-1} t_i$ est "trop petit", au profit de H_3 si $\sum_{i=1}^{n-1} t_i$ est "trop grand", ou au profit de H_1 si $\sum_{i=1}^{n-1} t_i$ est "soit trop petit, soit

trop grand”.

Pour pouvoir construire la région critique du test, il faut connaître la loi de $\sum_{i=1}^{n-1} T_i$ sous H_0 . On ne la connaît pas directement, mais on sait que le vecteur $\left(\frac{T_1}{T_n}, \dots, \frac{T_{n-1}}{T_n}\right)$ a, sous H_0 , la même loi que la statistique d'ordre d'un échantillon de taille $n-1$ de la loi uniforme $\mathcal{U}[0, 1]$.

Par conséquent, $\sum_{i=1}^{n-1} \frac{T_i}{T_n}$ a la loi de la somme de $n-1$ variables aléatoires indépendantes et de même loi $\mathcal{U}[0, 1]$. Cette loi n'a pas d'expression simple. Aussi on utilise un résultat asymptotique.

Comme l'espérance et la variance de la loi $\mathcal{U}[0, 1]$ sont respectivement $1/2$ et $1/12$, le théorème central-limite permet d'écrire que, sous H_0 :

$$U = \frac{\sum_{i=1}^{n-1} \frac{T_i}{T_n} - (n-1)\frac{1}{2}}{\sqrt{(n-1)\frac{1}{12}}} = \sqrt{\frac{12}{(n-1)T_n^2}} \left[\sum_{i=1}^{n-1} T_i - (n-1)\frac{T_n}{2} \right] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad (4.37)$$

$\sum_{i=1}^{n-1} T_i$ est “trop petit” si et seulement si U est “trop petit”. Donc un test de H_0 contre H_2 aura une région critique de la forme $W = \{U < k_\alpha\}$. Pour calculer k_α , on écrit que $\alpha = \mathbb{P}_{H_0}(U < k_\alpha)$, qui tend, quand n tend vers l'infini, vers $\phi(k_\alpha)$, où ϕ est la fonction de répartition de la loi normale centrée-réduite. Donc, asymptotiquement, $k_\alpha = \phi^{-1}(\alpha)$ est le quantile d'ordre α de la loi $\mathcal{N}(0, 1)$.

De même, un test de H_0 contre H_3 aura une région critique de la forme $W = \{U > l_\alpha\}$, d'où, asymptotiquement, $\alpha = \mathbb{P}_{H_0}(U > l_\alpha) = 1 - \phi(l_\alpha)$, et $l_\alpha = \phi^{-1}(1 - \alpha)$.

Pour le test de H_0 contre H_1 , on écrit $W = \{U < k_{1,\alpha} \text{ ou } U > k_{2,\alpha}\}$. On a alors asymptotiquement $\alpha = \mathbb{P}_{H_0}(U < k_{1,\alpha} \text{ ou } U > k_{2,\alpha}) = \mathbb{P}_{H_0}(U < k_{1,\alpha}) + \mathbb{P}_{H_0}(U > k_{2,\alpha}) = \phi(k_{1,\alpha}) + 1 - \phi(k_{2,\alpha})$.

Il y a une infinité de façons possibles de choisir $k_{1,\alpha}$ et $k_{2,\alpha}$ vérifiant cette propriété. Usuellement, on choisit de prendre $\phi(k_{1,\alpha}) = \alpha/2$ et $\phi(k_{2,\alpha}) = 1 - \alpha/2$, d'où $k_{1,\alpha} = \phi^{-1}(\alpha/2)$ et $k_{2,\alpha} = \phi^{-1}(1 - \alpha/2)$.

Mais la symétrie de la loi normale fait que $\phi^{-1}(\alpha/2) = -\phi^{-1}(1 - \alpha/2)$. Donc $W = \{U < -\phi^{-1}(1 - \alpha/2) \text{ ou } U > \phi^{-1}(1 - \alpha/2)\} = \{|U| > \phi^{-1}(1 - \alpha/2)\}$.

On obtient au bout du compte le test suivant :

Proposition 22 . On appelle **test de Laplace** le test basé sur la statistique $U = \sqrt{\frac{12}{(n-1)T_n^2}} \left[\sum_{i=1}^{n-1} T_i - (n-1)\frac{T_n}{2} \right]$. L'hypothèse nulle est H_0 : “les observations sont issues d'un HPP”. Les trois hypothèses alternatives ci-dessous correspondent à trois formes différentes du test, définies par leur région critique. Dans les trois cas, le test est asymptotiquement de seuil α .

- H_1 : “les observations ne sont pas issues d'un HPP”. $W = \{|U| > \phi^{-1}(1 - \alpha/2)\}$.

- H_2 : “il y a croissance de fiabilité”. $W = \{U < \phi^{-1}(\alpha)\}$.
- H_3 : “il y a décroissance de fiabilité”. $W = \{U > \phi^{-1}(1 - \alpha)\}$.

Pour les 3 systèmes, on obtient :

Système 1 : $u = -2.00$.

Système 2 : $u = +2.00$.

Système 3 : $u = 0.09$.

Au seuil $\alpha = 5\%$, les quantiles de la loi $\mathcal{N}(0, 1)$ sont $\phi^{-1}(0.05) = -1.645$, $\phi^{-1}(0.95) = +1.645$ et $\phi^{-1}(0.975) = +1.96$

- $-2 < -1.645$, donc on rejette l'hypothèse HPP et on conclut à la croissance de fiabilité pour le système 1.
- $+2 > +1.645$, donc on rejette l'hypothèse HPP et on conclut à la décroissance de fiabilité pour le système 2.
- $|0.09| < 1.96$, donc on ne peut pas rejeter l'hypothèse que les observations proviennent d'un HPP pour le système 3.

Ce sont bien les conclusions attendues.

Pour le jeu de données 1, on trouve $u = -9.17$, qui est très largement inférieur à -1.645 . Donc on n'hésite pas à rejeter l'hypothèse HPP et à conclure que les corrections apportées au logiciel ont augmenté sa fiabilité.

4.4.2.2. Le test logarithmique

Le test logarithmique consiste à utiliser le résultat qui dit que, sous H_0 , $V = -2 \sum_{i=1}^{n-1} \ln \frac{T_i}{T_n}$ est de loi $\chi_{2(n-1)}^2$. Sous H_0 , les t_i seront “petits” donc v sera “grand”. On en déduit donc le test suivant :

Proposition 23 . On appelle **test logarithmique** le test basé sur la statistique $V = -2 \sum_{i=1}^{n-1} \ln \frac{T_i}{T_n}$. L'hypothèse nulle est H_0 : “les observations sont issues d'un HPP”. Les trois hypothèses alternatives ci-dessous correspondent à trois formes différentes du test, définies par leur région critique. Dans les trois cas, le test est exactement de seuil α .

- H_1 : “les observations ne sont pas issues d'un HPP”.
 $W = \left\{ V < F_{\chi_{2(n-1)}^2}^{-1}(\alpha/2) \text{ ou } V > F_{\chi_{2(n-1)}^2}^{-1}(1 - \alpha/2) \right\}$.

- H_2 : “il y a croissance de fiabilité”. $W = \left\{ V > F_{\chi_{2(n-1)}^2}^{-1}(1 - \alpha) \right\}$.
- H_3 : “il y a décroissance de fiabilité”. $W = \left\{ V < F_{\chi_{2(n-1)}^2}^{-1}(\alpha) \right\}$.

On remarque que le test logarithmique est exact alors que le test de Laplace n’était qu’asymptotique. Sur les données, on obtient :

Système 1 : $v = 22.02$.

Système 2 : $v = 4.10$.

Système 3 : $v = 10.90$.

Jeu de données 1 : $v = 570.13$.

Au seuil $\alpha = 5\%$ et pour $n = 7$, les quantiles de la loi $\chi_{2(n-1)}^2$ sont $F_{\chi_{12}^2}^{-1}(0.05) = 5.23$, $F_{\chi_{12}^2}^{-1}(0.95) = 21.03$, $F_{\chi_{12}^2}^{-1}(0.025) = 4.40$ et $F_{\chi_{12}^2}^{-1}(0.975) = 23.34$. Pour $n = 136$, $F_{\chi_{270}^2}^{-1}(0.95) = 309.3$

- $22.02 > 21.03$, donc on rejette l’hypothèse HPP et on conclut à la croissance de fiabilité pour le système 1.
- $4.10 < 5.23$, donc on rejette l’hypothèse HPP et on conclut à la décroissance de fiabilité pour le système 2.
- $4.40 < 10.90 < 23.34$, donc on ne peut pas rejeter l’hypothèse que les observations proviennent d’un HPP pour le système 3.
- $570.13 > 309.3$, donc on rejette l’hypothèse HPP et on conclut à la croissance de fiabilité pour le logiciel.

Ce sont bien les mêmes conclusions qu’avec le test de Laplace..

4.4.3 Censure de type 1

Le principe est le même que pour la censure de type 2. Mais cette fois, au lieu de comparer (T_1, \dots, T_{n-1}) à T_n , on va comparer (T_1, \dots, T_n) à t , en utilisant les deux propriétés suivantes des HPP :

- Sachant $[N_t = n]$, le vecteur $\left(\frac{T_1}{t}, \dots, \frac{T_n}{t} \right)$ a la même loi que la statistique d’ordre d’un échantillon de taille n de la loi $\mathcal{U}[0, 1]$.
- Sachant $[N_t = n]$, $-2 \sum_{i=1}^n \ln \frac{T_i}{t}$ est de loi χ_{2n}^2 .

Du premier résultat, on déduit que, sachant $[N_t = n]$, on a sous H_0 :

$$U = \frac{\sum_{i=1}^n \frac{T_i}{t} - n\frac{1}{2}}{\sqrt{n\frac{1}{12}}} = \sqrt{\frac{12}{nt^2}} \left[\sum_{i=1}^n T_i - n\frac{t}{2} \right] \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1) \quad (4.38)$$

Contrairement au cas de la censure de type 2, n n'est pas une donnée du problème, mais l'observation de la variable aléatoire N_t . Donc, dans les statistiques de test, c'est N_t qui va intervenir à la place de n . Heureusement, on peut montrer que cela ne change rien sur la construction des tests et on obtient les résultats suivants.

Proposition 24 . On appelle **test de Laplace** le test basé sur la statistique $U = \sqrt{\frac{12}{N_t t^2}} \left[\sum_{i=1}^{N_t} T_i - N_t \frac{t}{2} \right]$. L'hypothèse nulle est H_0 : "les observations sont issues d'un HPP". Les trois hypothèses alternatives ci-dessous correspondent à trois formes différentes du test, définies par leur région critique. Dans les trois cas, le test est asymptotiquement de seuil α .

- H_1 : "les observations ne sont pas issues d'un HPP". $W = \{|U| > \phi^{-1}(1 - \alpha/2)\}$.
- H_2 : "il y a croissance de fiabilité". $W = \{U < \phi^{-1}(\alpha)\}$.
- H_3 : "il y a décroissance de fiabilité". $W = \{U > \phi^{-1}(1 - \alpha)\}$.

Proposition 25 . On appelle **test logarithmique** le test basé sur la statistique $V = -2 \sum_{i=1}^{N_t} \ln \frac{T_i}{t}$. L'hypothèse nulle est H_0 : "les observations sont issues d'un HPP". Les trois hypothèses alternatives ci-dessous correspondent à trois formes différentes du test, définies par leur région critique. Dans les trois cas, le test est exactement de seuil α .

- H_1 : "les observations ne sont pas issues d'un HPP".
 $W = \left\{ V < F_{\chi_{2N_t}^2}^{-1}(\alpha/2) \text{ ou } V > F_{\chi_{2N_t}^2}^{-1}(1 - \alpha/2) \right\}$.
- H_2 : "il y a croissance de fiabilité". $W = \left\{ V > F_{\chi_{2N_t}^2}^{-1}(1 - \alpha) \right\}$.
- H_3 : "il y a décroissance de fiabilité". $W = \left\{ V < F_{\chi_{2N_t}^2}^{-1}(\alpha) \right\}$.

Pour le jeu de données 2, on a $n = 6$, $u = 1.13$ et $v = 8.20$. Pour $\alpha = 5\%$, $\phi^{-1}(0.95) = 1.645$, $\phi^{-1}(0.975) = 1.96$, $F_{\chi_{12}^2}^{-1}(0.05) = 5.23$, $F_{\chi_{12}^2}^{-1}(0.025) = 4.40$ et $F_{\chi_{12}^2}^{-1}(0.975) = 23.34$.

On obtient donc $1.13 < 1.645 < 1.96$, $8.20 > 5.23$, $4.10 < 8.20 < 23.34$. Par conséquent, les deux tests s'accordent à conclure qu'on ne peut pas rejeter l'hypothèse HPP, ni au profit de H_3 , ni à celui de H_1 . Contrairement à ce qu'on pourrait croire, on ne peut donc pas conclure que la pompe s'use. En fait, c'est parce qu'il y a très peu de données que l'on ne peut pas détecter statistiquement le vieillissement de la pompe.

Chapitre 5

Le processus de puissance

5.1 Propriétés probabilistes

Le **processus de puissance** (noté PLP pour Power-Law Process) est le plus connu et le plus utilisé des modèles de fiabilité des systèmes réparables. Il porte aussi le nom de **modèle de Duane** car Duane [3] a observé empiriquement que ce modèle était bien adapté à des matériels électriques. C'est le NHPP dont l'intensité est une puissance du temps :

$$\forall t \geq 0, \lambda_t = \lambda(t) = \alpha \beta t^{\beta-1} \quad \alpha \in \mathbb{R}^+, \beta \in \mathbb{R}^+ \quad (5.1)$$

Une des premières qualités de ce modèle est sa faculté à modéliser des situations très différentes suivant la valeur du paramètre β (voir figure 5.1) :

- $\beta > 1 \implies \lambda(t)$ est croissante \implies décroissance de fiabilité
- $\beta < 1 \implies \lambda(t)$ est décroissante \implies croissance de fiabilité
- $\beta = 1 \implies \lambda(t)$ est constante \implies fiabilité stabilisée: on retrouve le processus de Poisson homogène

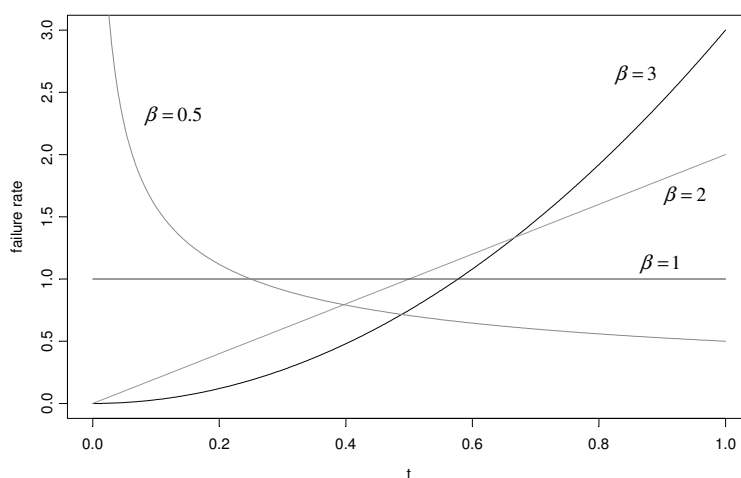


FIG. 5.1 – Intensité du PLP pour différentes valeurs de β

Le paramètre β s'interprète facilement comme le degré de dégradation ou d'amélioration du système. Pour un matériel qui s'use, on s'attend donc à avoir un paramètre β supérieur à 1, tandis que pour un logiciel, on s'attend à avoir β inférieur à 1. α est un paramètre d'échelle.

Comme pour le HPP, les propriétés du PLP se déduisent facilement des propriétés générales des NHPP.

5.1.1 Loi du nombre de défaillances survenues à chaque instant

- $\forall t \geq 0, m(t) = \mathbb{E}(N_t) = \int_0^t \lambda(s) ds = \alpha t^\beta.$
- $\{N_t\}_{t \geq 0}$ est à accroissements indépendants.
- $\forall t \geq 0, N_t$ est de loi de Poisson $\mathcal{P}(\alpha t^\beta)$. $\forall (s, t), 0 \leq s \leq t, N_t - N_s$ est de loi $\mathcal{P}(\alpha(t^\beta - s^\beta))$.
- $\lim_{t \rightarrow +\infty} m(t) = +\infty$, donc

$$\begin{aligned} & - \frac{N_t}{\alpha t^\beta} \xrightarrow[t \rightarrow +\infty]{p.s.} 1. \\ & - \frac{N_t - \alpha t^\beta}{\sqrt{\alpha t^\beta}} \xrightarrow[t \rightarrow +\infty]{\mathcal{L}} \mathcal{N}(0, 1). \end{aligned}$$

5.1.2 Lois des instants de défaillances

D'après le corollaire 4, la loi de T_n a pour densité :

$$f_{T_n}(t) = \lambda(t) e^{-m(t)} \frac{[m(t)]^{n-1}}{(n-1)!} = \alpha \beta t^{\beta-1} e^{-\alpha t^\beta} \frac{[\alpha t^\beta]^{n-1}}{(n-1)!} = \frac{\alpha^n \beta}{(n-1)!} t^{n\beta-1} e^{-\alpha t^\beta} \quad (5.2)$$

Cette loi de probabilité n'est pas une loi classique. Pour $\beta = 1$, on retrouve bien sûr la loi gamma $G(n, \lambda)$. Pour $n = 1$, on a :

$$f_{T_1}(x) = \alpha \beta x^{\beta-1} e^{-\alpha x^\beta} \quad (5.3)$$

ce qui prouve que T_1 (ou X_1) est de loi de Weibull.

En effet, la loi de Weibull $\mathcal{W}(\eta, \beta)$ est définie par sa fonction de répartition $F(t) = 1 - e^{-(t/\eta)^\beta}$ et par sa densité $f(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1} e^{-(t/\eta)^\beta}$. On voit qu'ici $\alpha = \eta^{-\beta}$, donc T_1 est de loi $\mathcal{W}(\alpha^{-1/\beta}, \beta)$. Par conséquent, le PLP est aux systèmes réparables ce que la loi de Weibull est aux systèmes non réparables.

La loi de T_n généralise à la fois la loi gamma et la loi de Weibull. On l'appelle parfois **loi gamma-Weibull**. On sait que si X est de loi $\exp(\alpha)$, alors $X^{1/\beta}$ est de loi $\mathcal{W}(\alpha^{-1/\beta}, \beta)$. De la même façon, si X est de loi $G(n, \alpha)$, alors $X^{1/\beta}$ a même loi que T_n .

5.1.3 Lois des durées inter-défaillances

L'intensité du PLP n'est pas de la forme $h(N_t, t - T_{N_t})$, donc, d'après le corollaire 1, les X_i ne sont pas indépendantes. D'après la proposition 11, la densité de la loi conditionnelle de X_{n+1} sachant $[T_n = t_n]$ est :

$$f_{X_{n+1}|T_n=t_n}(x) = \lambda(t_n + x) e^{-[m(t_n+x)-m(t_n)]} = \alpha\beta(t_n + x)^{\beta-1} e^{-\alpha[(t_n+x)^\beta - t_n^\beta]} \quad (5.4)$$

On constate que, à part pour $n = 0$, cette loi n'est pas une loi de Weibull. De même, X_1 est de loi de Weibull, mais pas les autres X_i . En fait, on peut concevoir le PLP comme un modèle de fiabilité pour un système dont la durée de vie initiale est de loi de Weibull et qui subit de petites réparations.

5.1.4 Fiabilité et MTTF

La fiabilité est :

$$R_t(\tau) = e^{-[m(t+\tau)-m(t)]} = e^{-\alpha[(t+\tau)^\beta - t^\beta]} \quad (5.5)$$

Le *MTTF* est :

$$MTTF_t = \int_0^{+\infty} R_t(\tau) d\tau = e^{\alpha t^\beta} \int_0^{+\infty} e^{-\alpha(t+\tau)^\beta} d\tau \quad (5.6)$$

A l'aide d'un changement de variables $u = s^\beta$, on montre que :

$$MTTF_t = e^{\alpha t^\beta} \Gamma\left(\frac{1}{\beta} + 1\right) \alpha^{-1/\beta} [1 - F_{G(1/\beta, \alpha)}(t^\beta)] \quad (5.7)$$

Et comme la fonction de répartition de la loi gamma peut s'écrire sous forme d'une fraction continue, on obtient également :

$$MTTF_t = \frac{t}{\beta} \frac{1}{\alpha t^\beta + \frac{1 - 1/\beta}{2}} \frac{1}{1 + \frac{2 - 1/\beta}{\alpha t^\beta + \frac{1}{1 + \dots}}} \quad (5.8)$$

5.1.5 Fonctions de vraisemblance

• Censure de type 2

$$\mathcal{L}(\alpha, \beta; t_1, \dots, t_n) = \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t_n)} = \left[\prod_{i=1}^n \alpha\beta t_i^{\beta-1} \right] e^{-\alpha t_n^\beta} = \alpha^n \beta^n \left[\prod_{i=1}^n t_i^{\beta-1} \right] e^{-\alpha t_n^\beta} \quad (5.9)$$

• Censure de type 1

$$\mathcal{L}(\alpha, \beta, n; t_1, \dots, t_n) = \left[\prod_{i=1}^n \lambda(t_i) \right] e^{-m(t)} = \alpha^n \beta^n \left[\prod_{i=1}^n t_i^{\beta-1} \right] e^{-\alpha t^\beta} \quad (5.10)$$

5.1.6 Quelques propriétés utiles

$\frac{\lambda(s)}{m(t)} = \frac{\alpha\beta s^{\beta-1}}{\alpha t^\beta} = \frac{\beta}{t} \left(\frac{s}{t}\right)^{\beta-1}$ La loi de probabilité de densité $f(t) = \frac{\beta}{t} \left(\frac{s}{t}\right)^{\beta-1} \mathbb{1}_{[0,t]}(s)$ est appelée **loi puissance** $\mathcal{P}u(t, \beta)$. On remarque que cette loi ne dépend pas de α .

$$\text{Par ailleurs, } \frac{m(T_i)}{m(T_n)} = \frac{\alpha T_i^\beta}{\alpha T_n^\beta} = \left(\frac{T_i}{T_n}\right)^\beta.$$

Les corollaires 5 à 11 s'écrivent donc pour le PLP de la façon suivante :

- Sachant $[T_n = t_n]$, le vecteur (T_1, \dots, T_{n-1}) a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi $\mathcal{P}u(t_n, \beta)$.
- Sachant $[N_t = n]$, le vecteur (T_1, \dots, T_n) a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{P}u(t, \beta)$.
- Le vecteur $\left(\left(\frac{T_1}{T_n}\right)^\beta, \dots, \left(\frac{T_{n-1}}{T_n}\right)^\beta\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi uniforme $\mathcal{U}[0, 1]$ et est indépendant de T_n .
- $-2 \sum_{i=1}^{n-1} \ln \left(\frac{T_i}{T_n}\right)^\beta = -2\beta \sum_{i=1}^{n-1} \ln \frac{T_i}{T_n}$ est indépendant de T_n et est de loi $\chi_{2(n-1)}^2$.
- Sachant $[N_t = n]$, le vecteur $\left(\left(\frac{T_1}{t}\right)^\beta, \dots, \left(\frac{T_n}{t}\right)^\beta\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\mathcal{U}[0, 1]$.
- Sachant $[N_t = n]$, $-2\beta \sum_{i=1}^n \ln \frac{T_i}{t}$ est de loi χ_{2n}^2 .

Une propriété classique dit que si U est une variable aléatoire de loi $\mathcal{U}[0, 1]$, alors $-\ln U$ est de loi $\exp(1)$ et $-\frac{1}{\lambda} \ln U$ est de loi $\exp(\lambda)$. Ici, $-\frac{1}{\beta} \ln \left(\frac{T_i}{T_n}\right)^\beta = -\ln \frac{T_i}{T_n} = \ln \frac{T_n}{T_i}$. On en déduit que :

- Le vecteur $\left(\ln \frac{T_n}{T_{n-1}}, \dots, \ln \frac{T_n}{T_1}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n - 1$ de la loi $\exp(\beta)$ et est indépendant de T_n .
L'intérêt de cette propriété est que le vecteur en question ne dépend pas des paramètres du modèle. De même, on a :
- Sachant $[N_t = n]$, le vecteur $\left(\ln \frac{t}{T_n}, \dots, \ln \frac{t}{T_1}\right)$ a la même loi que la statistique d'ordre d'un échantillon de taille n de la loi $\exp(\beta)$.

Enfin, la relation entre NHPP et HPP et les propriétés qui en découlent s'écrivent ici :

- Le processus de comptage d'évènements survenant aux instants αT_i^β est un HPP

d'intensité 1. De manière équivalente, le processus de comptage d'évènements survenant aux instants T_i^β est un HPP d'intensité α .

- Les $m(T_i) - m(T_{i-1}) = \alpha (T_i^\beta - T_{i-1}^\beta)$ sont indépendants et de même loi $\exp(1)$.
- $m(T_n) = \alpha T_n^\beta$ est de loi $G(n, 1)$ et $2\alpha T_n^\beta$ est de loi χ_{2n}^2 .

5.2 Estimation des grandeurs de la fiabilité

On s'intéresse ici à l'estimation de α et de β . On pourra utiliser ces estimations pour estimer la fiabilité, le *MTTF*, etc...

5.2.1 Censure de type 2

On observe les instants des n premières défaillances. La fonction de vraisemblance est :

$$\mathcal{L}(\alpha, \beta; t_1, \dots, t_n) = \alpha^n \beta^n \left[\prod_{i=1}^n t_i^{\beta-1} \right] e^{-\alpha t_n^\beta} \quad (5.11)$$

On a :

$$\ln \mathcal{L}(\alpha, \beta; t_1, \dots, t_n) = n \ln \alpha + n \ln \beta + (\beta - 1) \sum_{i=1}^n \ln t_i - \alpha t_n^\beta \quad (5.12)$$

D'où $\frac{\partial \ln \mathcal{L}}{\partial \alpha} = \frac{n}{\alpha} - t_n^\beta$, qui vaut 0 pour $\alpha = \frac{n}{t_n^\beta}$,

et $\frac{\partial \ln \mathcal{L}}{\partial \beta} = \frac{n}{\beta} + \sum_{i=1}^n \ln t_i - \alpha t_n^\beta \ln t_n$

En annulant cette quantité et en remplaçant α par $\frac{n}{t_n^\beta}$, on obtient :

$$\frac{n}{\beta} + \sum_{i=1}^n \ln t_i - n \ln t_n = \frac{n}{\beta} + \sum_{i=1}^n \ln \frac{t_i}{t_n} = 0 \quad (5.13)$$

D'où

$$\beta = \frac{n}{-\sum_{i=1}^n \ln \frac{t_i}{t_n}} = \frac{n}{\sum_{i=1}^n \ln \frac{t_n}{t_i}} \quad (5.14)$$

On obtient finalement :

Proposition 26 Les estimateurs de maximum de vraisemblance des paramètres du PLP sont :

$$\hat{\beta}_n = \frac{n}{\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}} \quad \text{et} \quad \hat{\alpha}_n = \frac{n}{T_n^{\hat{\beta}_n}} \quad (5.15)$$

Ces estimateurs ont une forme explicite, ce qui est très rare en fiabilité des systèmes réparables. C'est en grande partie l'explication de la popularité de ce modèle.

En remplaçant α et β par $\hat{\alpha}_n$ et $\hat{\beta}_n$ dans les expressions de $R_t(\tau)$ et $MTTF_t$, on peut estimer la fiabilité et le MTTF à n'importe quel instant.

Proposition 27 . $2n \frac{\beta}{\hat{\beta}_n}$ est indépendant de T_n et est de loi $\chi^2_{2(n-1)}$.

Démonstration: Il suffit de remarquer que $2n \frac{\beta}{\hat{\beta}_n} = -2\beta \sum_{i=1}^{n-1} \ln \frac{T_i}{T_n}$ et d'utiliser le résultat sur cette variable aléatoire vu dans le chapitre précédent. ■

Cette propriété permet de construire très facilement des intervalles de confiance pour β . Par exemple (en notant γ le seuil de l'intervalle car α désigne l'autre paramètre du PLP):

$$\begin{aligned} \mathbb{P}(\beta \in [0, \beta_{sup}]) &= \mathbb{P}(\beta \leq \beta_{sup}) = \mathbb{P}\left(2n \frac{\beta}{\hat{\beta}_n} \leq 2n \frac{\beta_{sup}}{\hat{\beta}_n}\right) = F_{\chi^2_{2(n-1)}}\left(2n \frac{\beta_{sup}}{\hat{\beta}_n}\right) = 1 - \gamma \\ &\implies \beta_{sup} = \frac{\hat{\beta}_n}{2n} F_{\chi^2_{2(n-1)}}^{-1}(1 - \gamma) \end{aligned} \quad (5.17)$$

D'où :

Proposition 28 . Des intervalles de confiance de seuil γ pour β sont :

$$\left[0, \frac{\hat{\beta}_n}{2n} F_{\chi^2_{2(n-1)}}^{-1}(1 - \gamma)\right], \quad \left[\frac{\hat{\beta}_n}{2n} F_{\chi^2_{2(n-1)}}^{-1}(\gamma), +\infty\right], \quad \left[\frac{\hat{\beta}_n}{2n} F_{\chi^2_{2(n-1)}}^{-1}\left(\frac{\gamma}{2}\right), \frac{\hat{\beta}_n}{2n} F_{\chi^2_{2(n-1)}}^{-1}\left(\frac{1 - \gamma}{2}\right)\right]$$

Le calcul de $\mathbb{E}[\hat{\lambda}_n]$ effectué dans la démonstration de la proposition 16 pour le HPP montre que, si X est de loi $G(n, \lambda)$, $\mathbb{E}\left(\frac{1}{X}\right) = \frac{\lambda}{n-1}$. Ici on obtient :

$2n \frac{\beta}{\hat{\beta}_n}$ est de loi $\chi^2_{2(n-1)} = G(n-1, 1/2)$. Donc $\mathbb{E}\left(\frac{\hat{\beta}_n}{2n\beta}\right) = \frac{1}{2n\beta} \mathbb{E}(\hat{\beta}_n) = \frac{1/2}{n-2}$. D'où $\mathbb{E}(\hat{\beta}_n) = \frac{n}{n-2} \beta$. Cela signifie que $\hat{\beta}_n$ est un estimateur biaisé de β mais que :

Proposition 29 . $\hat{\beta}'_n = \frac{n-2}{n} \hat{\beta}_n = \frac{n-2}{\sum_{i=1}^{n-1} \ln \frac{T_n}{T_i}}$ est un estimateur sans biais de β .

On n'a pas de résultats simples sur le biais de $\hat{\alpha}_n$, ni d'intervalles de confiance exacts pour α , car il n'y a pas de fonction pivotale simple.

En revanche, si on reparamétrise le modèle de façon à ce que son intensité soit de la forme $\lambda(t) = \frac{\beta}{\eta} \left(\frac{t}{\eta}\right)^{\beta-1}$, alors $\hat{\beta}_n$ a la même expression que précédemment et $\hat{\eta}_n = \frac{T_n}{n^{1/\hat{\beta}_n}}$.

De plus, $2m(T_n) = 2\left(\frac{T_n}{\eta}\right)^\beta$ est de loi χ_{2n}^2 , ce qu'on peut écrire :

$$2\left(\frac{\hat{\eta}_n n^{1/\hat{\beta}_n}}{\eta}\right)^\beta = 2\left(\frac{\hat{\eta}_n}{\eta}\right)^\beta n^{\beta/\hat{\beta}_n} = 2\left[n\left(\frac{\hat{\eta}_n}{\eta}\right)^{\hat{\beta}_n}\right]^{\beta/\hat{\beta}_n} \text{ est de loi } \chi_{2n}^2. \quad (5.19)$$

Comme la loi de $\frac{\beta}{\hat{\beta}_n}$ ne dépend ni de α ni de β , c'est aussi le cas de la loi de $\left(\frac{\hat{\eta}_n}{\eta}\right)^{\hat{\beta}_n}$, ce qui fournit une fonction pivotale. On construit alors des intervalles de confiance pour η à l'aide des quantiles de cette loi, que l'on peut obtenir numériquement.

Proposition 30 . *Propriétés asymptotiques des estimateurs de maximum de vraisemblance :*

- $\hat{\beta}_n \xrightarrow{ps} \beta$
- $\sqrt{n}\left(\frac{\hat{\beta}_n}{\beta} - 1\right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$
- $\hat{\alpha}_n \xrightarrow{ps} \alpha$
- $\frac{\sqrt{n}}{\ln n} \ln \frac{\hat{\alpha}_n}{\alpha} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

Démonstration: Les résultats sur β peuvent se déduire de la convergence de la loi $\chi_{2(n-1)}^2$ vers la loi $\mathcal{N}(0,1)$, ou de la propriété sur les $\ln \frac{T_n}{T_i}$. Les résultats sur α sont plus complexes à obtenir. ■

Le résultat de convergence en loi sur α permet de construire des intervalles de confiance asymptotiques pour α :

Proposition 31 . *Des intervalles de confiance asymptotiques de seuil γ pour α sont :*

$$\left[0, \hat{\alpha} n^{-\phi^{-1}(\gamma)/\sqrt{n}}\right], \quad \left[\hat{\alpha} n^{-\phi^{-1}(1-\gamma)/\sqrt{n}}, +\infty\right], \quad \left[\hat{\alpha} n^{-\phi^{-1}(1-\gamma/2)/\sqrt{n}}, \hat{\alpha} n^{-\phi^{-1}(\gamma/2)/\sqrt{n}}\right]$$

où ϕ est la fonction de répartition de la loi $\mathcal{N}(0,1)$.

La dualité entre intervalles de confiance et tests d'hypothèses fait que l'on peut facilement construire des tests portant sur la valeur de β . En particulier, comme " $\beta = 1$ " correspond au HPP, on peut tester l'hypothèse qu'un PLP est en fait un HPP en testant " $\beta = 1$ " contre " $\beta \neq 1$ ".

Le test consiste à rejeter " $\beta = 1$ " au seuil γ si la valeur 1 n'appartient pas à l'intervalle de confiance bilatéral pour β de seuil γ . Donc la région critique de ce test est :

$$\begin{aligned}
W &= \left\{ 1 \notin \left[\frac{\hat{\beta}_n}{2n} F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{\gamma}{2} \right), \frac{\hat{\beta}_n}{2n} F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{1-\gamma}{2} \right) \right] \right\} \\
&= \left\{ \frac{2n}{\hat{\beta}_n} < F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{\gamma}{2} \right) \text{ ou } \frac{2n}{\hat{\beta}_n} > F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{1-\gamma}{2} \right) \right\} \\
&= \left\{ 2 \sum_{i=1}^{n-1} \ln \frac{T_n}{T_i} < F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{\gamma}{2} \right) \text{ ou } 2 \sum_{i=1}^{n-1} \ln \frac{T_n}{T_i} > F_{\chi_{2(n-1)}^2}^{-1} \left(\frac{1-\gamma}{2} \right) \right\}
\end{aligned} \tag{5.21}$$

On voit que l'on retrouve exactement le test logarithmique d'adéquation au HPP. De la même façon, les tests logarithmiques de croissance et de décroissance de fiabilité sont les tests de " $\beta \geq 1$ " contre " $\beta < 1$ " et de " $\beta \leq 1$ " contre " $\beta > 1$ ". On peut montrer que ces tests sont optimaux, c'est-à-dire uniformément les plus puissants parmi les tests sans biais.

5.2.2 Censure de type 1

On observe le processus sur l'intervalle $[0, t]$. La fonction de vraisemblance est :

$$\mathcal{L}(\alpha, \beta, n; t_1, \dots, t_n) = \alpha^n \beta^n \left[\prod_{i=1}^n t_i^{\beta-1} \right] e^{-\alpha t^\beta} \tag{5.22}$$

Il est clair que le calcul est exactement le même que pour la censure de type 2, en remplaçant le t_n du dernier terme par t . En prenant en compte le fait que n est la réalisation de la variable aléatoire N_t , on obtient :

Proposition 32 *Les estimateurs de maximum de vraisemblance des paramètres du PLP sont :*

$$\hat{\beta}_t = \frac{N_t}{\sum_{i=1}^{N_t} \ln \frac{t}{T_i}} \quad \text{et} \quad \hat{\alpha}_t = \frac{N_t}{t \hat{\beta}_t} \tag{5.23}$$

On a vu que sachant $[N_t = n]$, $-2\beta \sum_{i=1}^n \ln \frac{T_i}{t}$ est de loi χ_{2n}^2 . On en déduit que :

Proposition 33 . *Sachant $[N_t = n]$, $2n \frac{\beta}{\hat{\beta}_t}$ est de loi χ_{2n}^2 .*

On peut alors en déduire des intervalles de confiance conditionnels pour β :

Proposition 34 . *Si on a observé n défaillances entre 0 et t , les intervalles suivants sont des intervalles de confiance conditionnels de seuil γ pour β , au sens où $\mathbb{P}(\beta \in I | N_t = n) = 1 - \gamma$.*

$$\left[0, \frac{\hat{\beta}_t}{2n} F_{\chi_{2n}^2}^{-1}(1-\gamma) \right], \quad \left[\frac{\hat{\beta}_t}{2n} F_{\chi_{2n}^2}^{-1}(\gamma), +\infty \right], \quad \left[\frac{\hat{\beta}_t}{2n} F_{\chi_{2n}^2}^{-1} \left(\frac{\gamma}{2} \right), \frac{\hat{\beta}_t}{2n} F_{\chi_{2n}^2}^{-1} \left(\frac{1-\gamma}{2} \right) \right]$$

Comme précédemment, on peut en déduire des tests portant sur la valeur de β , qui s'avèrent être les tests logarithmiques d'adéquation au HPP ou de tendance.

Par ailleurs, $\mathbb{E} \left[\frac{\hat{\beta}_t}{2n\beta} \mid N_t = n \right] = \frac{1/2}{n-1}$. On en déduit que $\mathbb{E} \left[\frac{n-1}{n} \hat{\beta}_t \mid N_t = n \right] = \beta$.

Donc, conditionnellement à $[N_t = n]$, $\frac{n-1}{n} \hat{\beta}_t$ est un estimateur sans biais de β . Mais cela ne permet pas d'en déduire un estimateur directement sans biais de β .

5.3 Tests d'adéquation

Tester l'adéquation au PLP, c'est faire un test de l'hypothèse nulle

$$H_0 : \text{“les observations sont issues d'un PLP”}$$

contre l'hypothèse alternative

$$H_1 : \text{“les observations ne sont pas issues d'un PLP”}.$$

On sait bien faire des tests d'adéquation quand les observations sont indépendantes sous H_0 . Mais ce n'est pas le cas pour le PLP. Une solution est de se ramener à des variables aléatoires indépendantes (ou presque).

Pour cela, on utilise la propriété qui dit que le vecteur $\left(\ln \frac{T_n}{T_{n-1}}, \dots, \ln \frac{T_2}{T_1} \right)$ a la même loi que la statistique d'ordre d'un échantillon de taille $n-1$ de la loi $\exp(\beta)$. Tester l'adéquation des T_i au PLP revient donc à tester l'adéquation des $\ln \frac{T_n}{T_i}$ à un échantillon ordonné de loi exponentielle, ce que l'on sait faire par des méthodes classiques. Le test le plus efficace est le test d'Anderson-Darling.

En pratique, on utilise souvent la procédure très simple d'adéquation graphique suivante.

Rappelons que, pour le PLP, $m(t) = \alpha t^\beta$. On a donc $\ln m(t) = \ln \alpha + \beta \ln t$. Pour $t = t_i$, on obtient $\ln m(t_i) = \ln \alpha + \beta \ln t_i$.

Or $m(t) = \mathbb{E}[N_t]$ et $N_{T_i} = i$, donc $m(t_i)$ devrait être proche de i . Par conséquent, si le PLP est un modèle adapté aux données, on devrait avoir pour tout i , $\ln i \approx \ln \alpha + \beta \ln t_i$.

Un test d'adéquation graphique consiste à tracer le nuage des points $(\ln t_i, \ln i)$ pour $i \in \{1, \dots, n\}$. Si les points sont approximativement alignés, on admettra que le PLP est un modèle acceptable pour les données. De plus, la pente et l'ordonnée à l'origine de cette droite fourniront de nouveaux estimateurs de α et β . Si les points ne sont manifestement pas alignés, on rejettera le modèle PLP.

L'alignement des points peut être mesuré à l'aide du coefficient de détermination R^2 usuel des procédures de régression linéaire. On peut même construire un test d'adéquation statistique en déterminant la loi de R^2 sous l'hypothèse PLP.

En conclusion, on est tout à fait capables de déterminer si le PLP est un modèle adapté ou pas à un jeu de données.

5.4 Application

5.4.1 Jeu de données 1

On obtient les résultats suivants :

- $\hat{\alpha}_n = 0.583$.
- $\hat{\beta}_n = 0.477$. $0.477 < 1$ confirme la croissance de fiabilité pour ces données.
- $\widehat{MTTF}_n = 1429$ s. C'est bien plus vraisemblable que la valeur 676 s obtenue sous l'hypothèse HPP.
- $\hat{R}_n(1000) = 49.5\%$. Là encore, c'est bien plus vraisemblable que la valeur 22.9% obtenue sous l'hypothèse HPP.

Quand on applique les tests d'adéquation, l'hypothèse PLP n'est pas rejetée au seuil 5%, mais elle l'est au seuil 10%. Donc le PLP n'est quand même pas un modèle très adapté à ces données.

5.4.2 Jeu de données 2

On obtient les résultats suivants :

- $\hat{\alpha}_t = 0.0049$.
- $\hat{\beta}_n = 1.463$. $1.463 > 1$ confirme la décroissance de fiabilité pour ces données.
- $\widehat{MTTF}_t = 10$ mois. C'est bien plus vraisemblable que la valeur 21 mois obtenue sous l'hypothèse HPP.
- $\hat{R}_t(6) = 66.2\%$. Là encore, c'est plus vraisemblable que la valeur 75% obtenue sous l'hypothèse HPP (mais la différence est moins frappante que pour les autres critères).

L'hypothèse PLP n'est pas rejetée par les tests d'adéquation. Mais ce n'est pas forcément très significatif, car il est difficile pour un test de rejeter l'hypothèse nulle avec si peu de données.

En conclusion, dans ces deux cas, il apparait que le PLP est un modèle plus vraisemblable que le HPP. Mais il existe peut-être des modèles qui soient encore plus pertinents. Il faut donc envisager des modèles concurrents. Il en existe de très nombreux.

Bibliographie

- [1] ASCHER H. ET FEINGOLD H., *Repairable systems reliability: modeling, inference, misconceptions and their causes*, Marcel Dekker, 1984.
- [2] COCOZZA-THIVENT C., *Processus stochastiques et fiabilité des systèmes*, Springer, 1997.
- [3] DUANE J.T., Learning curve approach to reliability monitoring, *IEEE Transactions on Aerospace*, AS-2, **2**, 563-566, 1964.
- [4] RIGDON S.E. ET BASU A.P., *Statistical Methods for the Reliability of Repairable Systems*, Wiley, 2000.
- [5] SNYDER D.L. ET MILLER M.I., *Random point processes in time and space*, Springer, 1991.