

Université de Montréal

Remplacement préventif périodique: une approche
non paramétrique

par

Patrice Mathieu

Département de mathématiques et de statistique
Faculté des arts et des sciences

Mémoire présenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en statistique

juillet 1999



2.844.11MS

QA

3

U54

1999

V.010

Université de Montréal

Remplacement préventif périodique: une approche

des paramétriques

par

Patrick Mathieu

Édition et distribution par l'Université de Montréal
Révisé par les soins de l'auteur

Édition par les soins de l'auteur

Le prix de l'édition est de \$15,00

Le prix de l'édition est de \$15,00

© 1999

1999



1999

Université de Montréal

Faculté des études supérieures

Ce mémoire intitulé

**Remplacement préventif périodique: une approche
non paramétrique**

présenté par

Patrice Mathieu

a été évalué par un jury composé des personnes suivantes :

M. Roch Roy

(président-rapporteur)

M. Robert Cléroux

(directeur de recherche)

M. Guy Lapalme

(co-directeur)

M. Yves Lepage

(membre du jury)

Mémoire accepté le :

1 juillet 1999

SOMMAIRE

Dans ce mémoire, nous nous intéressons aux politiques de remplacement préventif. Le développement de stratégies qui prévoient le remplacement d'une pièce avant qu'elle ne tombe en panne permet souvent de limiter les dépenses occasionnées par une telle panne. La politique de remplacement que l'on privilégiera sera celle qui minimisera la fonction de coût, donnée par le coût moyen par unité de temps sur un horizon infini. Nous présenterons plus en détail cette fonction et nous insisterons sur son application dans un contexte de remplacement préventif périodique. Ce type de remplacement consiste à remplacer systématiquement chaque pièce par une neuve aux instants kT , $k = 1, 2, 3, \dots$ et ce peu importe l'état ou l'âge de la pièce. Nous mettrons une emphase particulière sur une approche non paramétrique pour estimer cette fonction pour un tel type de stratégie préventive. Notons que la littérature sur le sujet est relativement peu développée et qu'une telle approche n'a jamais été utilisée pour ce type de remplacement.

Pour l'estimation de la fonction de coût, nous aurons recours à quelques notions de la théorie du renouvellement. Un survol de cette théorie sera donc effectué en insistant sur son élément le plus important: la fonction de renouvellement. Elle est définie comme le nombre moyen de remplacements dans un intervalle $[0, t]$.

Ce mémoire présentera donc, dans un premier temps, une étude détaillée sur la fonction de renouvellement. Différentes techniques d'estimation de cette fonction, principalement des méthodes non paramétriques, seront alors énumérées. Un résumé de la littérature et des simulations permettront de vérifier la performance de chaque estimateur. Nous observerons que, si la distribution de laquelle proviennent les durées de vie observés est connue, les méthodes paramétriques performant le mieux. Dans le cas contraire, les estimateurs non paramétriques sont tout de même assez efficaces. Nous verrons aussi que l'hypothèse de non décroissance du taux de panne (distribution IFR) ne permet pas une meilleure estimation de la fonction de renouvellement.

Par la suite, ces résultats seront utilisés pour estimer la fonction de coût d'une politique de remplacement préventif périodique et en particulier son minimum. Des comparaisons et des simulations seront à nouveau effectuées. L'estimation paramétrique est toujours préférable, mais cette fois, l'hypothèse d'une distribution IFR semble procurer des résultats légèrement supérieurs (plus petits biais moyens, plus petites erreurs quadratiques moyennes et plus petites erreurs relatives moyennes).

REMERCIEMENTS

Je tiens à remercier particulièrement mon directeur de recherche, M. Robert Cléroux pour son aide, ses idées et sa grande disponibilité durant toutes les étapes de la réalisation de ce mémoire. Je lui suis aussi très reconnaissant de m'avoir donné l'opportunité de présenter les résultats de ce mémoire dans une conférence à Marrakech au Maroc. Son soutien financier m'a aussi été très utile. Je remercie également mon co-directeur, M. Guy Lapalme, pour son aide au niveau informatique.

D'un point de vue plus personnel, je souhaite remercier ma famille et mes amis pour leurs encouragements continuels. Un merci tout spécial à mes parents et mon frère pour l'intérêt qu'ils ont toujours manifesté envers mes études.

Table des matières

Sommaire	iii
Remerciements	v
Table des figures	viii
Liste des tableaux	ix
Introduction	1
Chapitre 1. La fonction de renouvellement	5
1.1. Hypothèses de départ	6
1.2. La fonction de renouvellement	8
1.3. Estimation de la fonction de renouvellement	9
1.3.1. Premier estimateur de $M(t)$	10
1.3.2. Deuxième estimateur de $M(t)$	14
1.3.3. Troisième estimateur de $M(t)$	16
1.3.4. Quatrième estimateur de $M(t)$	20
1.3.5. Cinquième estimateur de $M(t)$	23
Chapitre 2. Comparaison des estimateurs	29
2.1. Résumé de la littérature	30
2.2. Simulations	39

2.2.1. Plan de simulation	39
2.2.2. Résultats	41
2.2.3. Analyse et conclusion	51
Chapitre 3. Politiques de remplacement préventif.....	61
3.1. Remplacement préventif périodique	64
3.1.1. Estimation de la fonction de coût	67
3.1.2. Simulations	68
3.1.3. Analyse et conclusion	70
3.2. Remplacement préventif périodique avec réparation minimale à la panne	75
3.2.1. Estimation de la fonction de coût	77
3.2.2. Simulations	79
3.2.3. Analyse et conclusion	80
Conclusion	85
Annexe A. Analyse de variance.....	88
A.1. Distribution de Weibull	88
A.2. Distribution normale tronquée.....	91
Annexe B. Programmes	94
B.1. section 1: S-Plus.....	94
B.2. section 2: Fortran	95
Bibliographie	144

Table des figures

2.2.1	Histogrammes des distributions gamma simulées	42
2.2.2	Histogrammes des distributions de Weibull simulées.....	43
2.2.3	Histogrammes des distributions normales tronquées simulées	44
2.2.4	Fonction de renouvellement et ses estimateurs.....	52
2.2.5	Interaction entre le temps et les estimateurs (distribution gamma) ...	59
3.1.1	Schéma d'un remplacement préventif périodique	64
3.1.2	Fonction de coût théorique (remplacement périodique)	66
3.1.3	Fonction de coût et ses estimateurs (remplacement périodique).....	73
3.2.1	Fonction de coût et ses estimateurs (réparation minimale à la panne).	83
A.1.1	Interaction entre le temps et les estimateurs (distribution de Weibull)	90
A.2.1	Interaction entre le temps et les estimateurs (distribution normale tronquée).....	93

Liste des tableaux

2.1.1	Estimation de $M(t)$ pour l'exemple de Juran & Gryna	31
2.1.2	Estimation de $M(t)$ pour la loi Gamma (littérature)	33
2.1.3	Estimation de $M(t)$ pour la loi de Weibull (littérature)	34
2.1.4	Estimation de $M(t)$ pour la loi normale tronquée (littérature)	35
2.1.5	Estimation de $M(t)$ pour la loi log-normale (littérature)	36
2.1.6	Temps en secondes du calcul de $M_{3n}(t)$ et de $M_{4n}(t)$	38
2.2.1	Paramètres des distributions simulées	40
2.2.2	Estimation de $M(t)$ pour la loi gamma	45
2.2.3	Estimation de $M(t)$ pour la loi de Weibull	47
2.2.4	Estimation de $M(t)$ pour la loi normale tronquée	49
2.2.5	Table d'anova pour la loi Gamma ($\alpha = 0.05$)	55
2.2.6	Contrastes entre les estimateurs ($\alpha = 0.05$) (distribution gamma)	56
3.1.1	Estimation de $B(t)$ minimal pour la loi gamma	70
3.1.2	Estimation de $B(t)$ minimal pour la loi de Weibull	71
3.1.3	Estimation de $B(t)$ minimal pour la loi normale tronquée	72
3.2.1	Estimation de $D(t)$ minimal pour la loi gamma	80
3.2.2	Estimation de $D(t)$ minimal pour la loi de Weibull	81
3.2.3	Estimation de $D(t)$ minimal pour la loi normale tronquée	82

A.1.1	Table d'anova pour la loi de Weibull	88
A.1.2	Contrastes entre les estimateurs (distribution de Weibull)	89
A.2.1	Table d'anova pour la loi normale tronquée	91
A.2.2	Contrastes entre les estimateurs (distribution normale tronquée)	92

INTRODUCTION

Tout objet, appareil ou pièce d'équipement ne peut fonctionner ou être utilisé indéfiniment. Ces derniers s'usent, se brisent ou tombent en panne et il faut les réparer ou les remplacer. Puisque ces réparations et ces remplacements impliquent des coûts (prix de la pièce, main-d'oeuvre, etc.), il est intéressant d'en prévoir la fréquence. Ceci permettra notamment de calculer le coût d'une machine ou d'un objet pour une certaine période ou de déterminer la durée optimale d'une garantie qu'un manufacturier doit accorder à un client. Il est aussi de plus en plus important d'instaurer des politiques de remplacement préventif (remplacer une pièce qui fonctionne bien avant qu'elle ne tombe en panne), afin de minimiser les coûts d'installation ou les frais encourus lors d'une panne. Les stratégies de maintenance et de remplacement préventif sont parmi les applications principales de la théorie du renouvellement et seront au coeur de ce mémoire.

Avant tout, nous traiterons de la fonction de renouvellement. Elle est la base de cette théorie du renouvellement et joue un rôle primordial dans le développement de politiques de remplacement préventif. Elle sera donc essentielle pour étudier ces politiques. La fonction de renouvellement est en fait un bon outil pour obtenir la fréquence de remplacement d'une pièce. Elle est notée $M(t)$ et est définie par le nombre moyen de remplacements dans un intervalle de temps $[0,t]$. Il sera donc abondamment question de cette fonction dans ce travail. Les deux premiers chapitres de ce mémoire y sont d'ailleurs consacrés.

Au premier chapitre, nous présenterons quelques notions de base sur la théorie du renouvellement et en particulier sur la fonction de renouvellement. Comme nous le constaterons, il est en général difficile de calculer cette fonction, puisqu'il n'existe aucune formule explicite permettant de le faire. D'autre part, le développement de bonnes méthodes d'estimation est primordial. Plusieurs estimateurs ont été développés depuis les années soixante, mais c'est surtout au cours des 15 dernières années que l'estimation de la fonction de renouvellement a suscité le plus d'intérêt. La plupart des estimateurs employés aujourd'hui, principalement des estimateurs non paramétriques, ont été introduits à partir de ce moment. Dans ce qui suit, nous énoncerons cinq estimateurs de cette fonction et présenterons en détail leurs principales propriétés et caractéristiques (propriétés asymptotiques, paramétriques ou non, hypothèses de départ, ...).

Au chapitre suivant, des comparaisons entre ces 5 estimateurs de $M(t)$ seront effectuées. Nous ferons une synthèse des différents résultats disponibles dans la littérature et nous développerons nos propres simulations, afin de vérifier la performance de chaque estimateur.

Le troisième chapitre permettra d'introduire la notion de remplacement préventif. Que ce soit en remplaçant une pièce lorsqu'elle atteint un certain âge ou en remplaçant toutes les pièces simultanément après une certaine période de temps fixée, il existe plusieurs façons d'effectuer des remplacements préventifs. Nous ferons un bref survol de différents types de remplacement préventif en insistant sur les remplacements de types *périodique* et *périodique avec réparation minimale à la panne*.

Dans le cadre de ces types spécifiques de remplacement préventif, nous nous intéresserons en particulier à la fonction de coût. Puisque le but de l'implantation de stratégies de remplacement est de minimiser les coûts liés à l'entretien, à la réparation ou au remplacement d'une pièce ou d'un système, la fonction de coût prend une importance capitale. Nous définirons donc cette fonction et en présenterons les différentes caractéristiques pour les deux types de remplacement mentionnés ci-haut. Dans un contexte de remplacement périodique, la fonction de coût est intimement liée à la fonction de renouvellement. Nous appliquerons donc les notions développées au premier chapitre pour l'estimer. Nous aurons aussi recours au taux de panne, que nous définirons, pour estimer la fonction de coût, lorsque les pièces en panne subissent des réparations minimales. A nouveau, des simulations et des comparaisons seront effectuées. Nous insisterons alors principalement sur le minimum de cette fonction de coût. En effet, cette valeur, et le moment où elle est obtenue, sont importants, puisqu'ils déterminent l'instant idéal pour effectuer le remplacement systématique, c'est-à-dire les stratégies optimales de remplacement préventif.

La question de l'estimation de la fonction de coût et de son minimum dans un tel contexte est un problème relativement nouveau qui n'a pas été souvent soulevé dans la littérature. En particulier, l'utilisation d'estimateurs non paramétriques est à toute fin pratique inexistante. Une attention spéciale y sera donc portée dans ce mémoire. Nous présenterons et vérifierons la performance et la qualité de certains estimateurs de la fonction de coût (surtout des estimateurs non paramétriques). Nous ferons aussi un parallèle avec l'estimation de la fonction de renouvellement. Il sera intéressant de mesurer jusqu'à quel point les meilleurs

estimateurs de $M(t)$ le sont aussi pour estimer la fonction de coût. Nous verrons également si les deux types de remplacement que nous étudierons en particulier conduisent aux mêmes résultats et conclusions.

Chapitre 1

LA FONCTION DE RENOUVELLEMENT

Avant d'aborder la question de remplacement préventif, le point culminant de ce mémoire, attardons-nous quelques instants sur la fonction de renouvellement. Rappelons que nous l'avons définie comme étant le nombre moyen de remplacements que l'on doit effectuer dans l'intervalle $[0, t]$ et qu'elle sera très utile pour développer des stratégies préventives de remplacement.

La fonction de renouvellement a plusieurs applications dans l'étude de modèles probabilistes. Que ce soit en théorie de la fiabilité, en théorie de l'inventaire, dans des plans d'échantillonnage continu ou pour faire de l'analyse séquentielle, on utilise fréquemment la fonction de renouvellement. En théorie de la fiabilité, elle est notamment utilisée à des fins plus spécifiques, comme pour l'étude du coût d'une garantie (voir Blischke et Scheuer [5]) ou pour la détermination de politiques de remplacement. Par exemple, une entreprise pourra utiliser cette fonction pour prévoir les dépenses associées à l'entretien, à la réparation ou au remplacement de certaines pièces de leur équipement. Une meilleure planification des dépenses permettra de les optimiser et fournira une idée plus précise des profits que l'entreprise réalisera. La fonction de renouvellement permettra aussi de mesurer la fiabilité et la durabilité des produits qu'elle mettra sur le marché.

Dans ce chapitre, nous introduirons la fonction de renouvellement. Nous énoncerons d'abord les hypothèses que l'on doit établir et donnerons ensuite les principales propriétés de cette fonction. Dans un deuxième temps, nous présenterons cinq estimateurs différents de la fonction de renouvellement. Rappelons que ces estimateurs seront comparés au chapitre suivant en citant quelques exemples disponibles dans la littérature. Nous effectuerons aussi des simulations pour vérifier et compléter ces résultats.

1.1. HYPOTHÈSES DE DÉPART

Dans ce qui suit, nous considérerons les cas où les hypothèses suivantes sont respectées. Ces hypothèses sont des hypothèses standards en théorie de la fiabilité et en particulier en théorie du renouvellement. Elles nous permettront de simplifier et de mieux cibler la question de l'estimation de $M(t)$.

- i. Le système considéré ne possède qu'une seule composante. C'est-à-dire que le système fonctionne si et seulement si l'unique pièce qui le compose fonctionne. Les termes pièce, composante et système seront donc des synonymes dans le texte ci-dessous.
- ii. Le système commence à fonctionner en $t = 0$.
- iii. Chaque item qui tombe en panne est remplacé par un item identique et neuf. (Quand il est question de coûts, ils sont les mêmes pour chaque item.)
- iv. Le remplacement d'un item se fait instantanément lorsqu'il tombe en panne. (Le temps d'installation est négligeable par rapport au temps total.)
- v. La distribution de durée de vie de chaque item de remplacement est la même, c'est-à-dire X_1, X_2, \dots, X_n sont i.i.d. avec fonction de distribution

sous-jacente $F(t)$ où X_i est la durée de fonctionnement entre la $(i - 1)^e$ et la i^e panne. (En fait, X_i est la durée de fonctionnement du i^e item utilisé.) De plus, les X_i sont des variables aléatoires non toutes nulles à la fois avec probabilité 1.

Considérons maintenant S_k , l'instant du k^e renouvellement. Nous avons alors:

$$S_k = X_1 + X_2 + \dots + X_k \quad k = 1, \dots, n$$

Nous pouvons donc obtenir:

$$\begin{aligned} \mathcal{P}(\text{au moins } k \text{ renouvellements dans } [0, t]) \\ &= \mathcal{P}(S_k \leq t) \\ &= \mathcal{P}(X_1 + X_2 + \dots + X_k \leq t) \\ &= F^{(k)}(t) \end{aligned}$$

où $F^{(k)}(t)$ est la convolution de $F(t)$ k fois avec elle-même.

Nous pouvons alors définir $N(t)$, le nombre de renouvellements dans $[0, t]$. La distribution de $N(t)$ est donnée par:

$$\begin{aligned} \mathcal{P}(N(t) = n) &= \mathcal{P}(S_n \leq t < S_{n+1}) \\ &= \mathcal{P}(S_n \leq t) - \mathcal{P}(S_{n+1} \leq t) \\ &= F^{(n)}(t) - F^{(n+1)}(t) \end{aligned}$$

1.2. LA FONCTION DE RENOUVELLEMENT

Nous possédons maintenant tous les outils nécessaires pour définir la fonction de renouvellement. Elle est notée $M(t)$ et est donnée par

$$M(t) = \text{nombre moyen de renouvellements dans } [0, t]$$

Plus précisément, on a:

$$\begin{aligned} M(t) &= E[N(t)] \\ &= \sum_{n=0}^{\infty} n \mathcal{P}(N(t) = n) \\ &= \sum_{n=0}^{\infty} n [F^{(n)}(t) - F^{(n+1)}(t)] \\ &= \sum_{n=1}^{\infty} F^{(n)}(t) \end{aligned}$$

Mentionnons que la fonction de répartition $F(t)$ détermine $M(t)$ et, inversement, $M(t)$ détermine $F(t)$. Il existe aussi une autre représentation de la fonction de renouvellement, soit l'*équation fondamentale du renouvellement* (voir Cléroux [6]). Elle est donnée par:

$$M(t) = F(t) + \int_0^t M(t-x) dF(x)$$

Comme nous l'avons vu ci-haut, la fonction de renouvellement est obtenue à partir des convolutions de $F(t)$ n fois avec elle-même ($n = 1, \dots, \infty$). A l'exception de quelques distributions, telles l'exponentielle ou la gamma (avec $\alpha = 2$), il n'est pas possible d'obtenir une forme explicite pour ces convolutions. Par conséquent,

il sera aussi impossible d'en obtenir une pour la fonction de renouvellement. Nous devons donc nous contenter de résultats asymptotiques, d'approximations ou nous devons avoir recours à des algorithmes pour calculer de telles fonctions et, en particulier, pour estimer $M(t)$.

1.3. ESTIMATION DE LA FONCTION DE RENOUVELLEMENT

Les premiers efforts à être déployés dans le but d'estimer la fonction de renouvellement l'ont été dans les années soixante. Les personnes qui se sont penchées sur le sujet au cours de ces années ont principalement développé des méthodes paramétriques ou ont utilisé un estimateur asymptotique et non paramétrique simple. Un des premiers à utiliser une approche non paramétrique plus complexe a été Edward W. Frees, au milieu de la décennie 80. Ses travaux, relativement récents donc, ont été repris par de nombreux auteurs. Ces derniers ont précisé et peaufiné certains résultats de Frees et ont développé différents algorithmes pour calculer le plus efficacement possible les estimateurs qu'il a proposés.

Il existe donc plusieurs estimateurs de la fonction de renouvellement. Nous en avons retenu cinq que nous énoncerons dans cette section. Nous discuterons des principaux avantages et inconvénients de chacun d'eux et étudierons leurs principales propriétés. Nous verrons aussi comment ils se comportent en présence de données censurées. Des comparaisons seront effectuées au chapitre suivant en mentionnant les différentes conclusions disponibles dans la littérature. Nous effectuerons aussi des simulations pour les vérifier ou pour tirer nos propres conclusions.

Afin de choisir un bon estimateur, plusieurs facteurs doivent être pris en considération. Certaines situations permettent le choix d'un estimateur paramétrique alors que pour d'autres cas, le choix d'un estimateur non paramétrique est plus approprié. De plus, la connaissance de certaines caractéristiques telles la croissance ou la décroissance du taux de panne (que nous introduirons plus loin) permettent l'utilisation d'estimateurs plus performants. Aussi, la capacité de l'équipement informatique dont on dispose doit être prise en considération. Ce problème se pose moins à l'aube du 21^e siècle, alors que la performance des ordinateurs est sans cesse grandissante. Ceci n'a pas toujours été le cas, d'où l'intérêt par le passé à développer des estimateurs nécessitant un temps de calcul relativement court. L'utilisation et les besoins spécifiques de chaque situation influence aussi le choix d'un estimateur. Par exemple, un estimateur basé sur un résultat asymptotique risque d'être moins, voire complètement inefficace pour de petites valeurs de t .

1.3.1. Premier estimateur de $M(t)$

Le premier estimateur de la fonction de renouvellement que nous allons considérer est basé sur le comportement asymptotique d'un processus de renouvellement. C'est au début des années soixante qu'une telle approche a été considérée pour la première fois. Cet estimateur asymptotique découle en fait des deux théorèmes suivants (dont les preuves se trouvent dans Cléroux [6]).

Théorème 1.3.1 (Théorème élémentaire du renouvellement). *Si $F(t)$ a une moyenne μ , alors $\lim_{t \rightarrow \infty} \frac{M(t)}{t} = \frac{1}{\mu}$*

Théorème 1.3.2. *Si $F(t)$ est non arithmétique avec moyenne μ et variance σ^2 , alors $\lim_{t \rightarrow \infty} [M(t) - \frac{t}{\mu}] = \frac{\sigma^2}{2\mu^2} - \frac{1}{2}$*

Rappelons qu'une variable aléatoire est dite *arithmétique* si elle prend ses valeurs parmi $\{0, \pm d, \pm 2d, \dots\}$ où d est une constante. Dans le cas contraire, la variable aléatoire est dite *non arithmétique*. Par conséquent, la fonction de répartition d'une telle variable aléatoire est dite arithmétique ou non arithmétique.

Il s'ensuit que, pour t grand:

$$M(t) \approx \frac{t}{\mu} + \frac{\sigma^2}{2\mu^2} - \frac{1}{2} \quad (1.3.1)$$

Soulignons que pour la distribution exponentielle, $M(t) = \frac{t}{\mu} \forall t$. Dans de telles circonstances, l'approximation donnée par (1.3.1) est exacte, peu importe la valeur de t .

En remplaçant μ et σ^2 par des estimateurs convergents $\hat{\mu}$ et $\hat{\sigma}^2$ respectivement, on peut estimer la fonction $M(t)$ par la droite donnée en (1.3.1). (μ et σ^2 sont obtenus à partir des n durées de vie échantillonnables x_1, \dots, x_n .) On obtient ainsi notre premier estimateur de $M(t)$, soit:

$$M_{1n}(t) = \frac{t}{\hat{\mu}} + \frac{\hat{\sigma}^2}{2\hat{\mu}^2} - \frac{1}{2} \quad (1.3.2)$$

Ne nécessitant que la connaissance des deux premiers moments échantillonnables, l'estimateur $M_{1n}(t)$ est non paramétrique. Il s'obtient très facilement et nécessite un temps de calcul très court. Cependant, à l'exception de la distribution exponentielle, $M_{1n}(t)$ est basé sur un résultat limite et n'est donc valable que

pour de grandes valeurs de t . Il a d'ailleurs été montré que cet estimateur donnait de bons résultats pour t grand. Cependant, en pratique, on rencontre souvent de petites valeurs de t , ce qui rend $M_{1n}(t)$ inefficace.

Attardons-nous quelques instants sur la signification de t *grand*. Qu'entendons-nous par une valeur de t suffisamment grande? Cette valeur critique de t dépend évidemment de la précision souhaitée, mais aussi de la distribution des X_i . Certains auteurs, tels Baker [2] considèrent que t doit être supérieur à 2μ . Pour sa part, Sahin [16] a étudié la vitesse de convergence de $M_{1n}(t)$ pour certaines distributions spécifiques en portant une attention particulière à cet aspect.

Sahin affirme que t doit être supérieur à $k(\epsilon)\mu$ où ϵ détermine la précision désirée. Il a montré que la valeur de k dépend principalement de c , le coefficient de variation de la distribution considérée ($c = \frac{\sigma}{\mu}$). De façon générale, l'auteur a noté que la valeur de k diminue à mesure que c augmente lorsque $c < 1$. Pour $c > 1$, k augmente au lieu de diminuer. De plus, à quelques exceptions près (en particulier pour la distribution log-normale), les valeurs de k varient entre 1 et 4 dont la majorité se situent entre 1 et 2 (avec $\epsilon = 0.01$). Sahin a également développé une équation polynomiale du second degré de la forme $\hat{k}(\epsilon) = a_0(\epsilon) + a_1(\epsilon)c + a_2(\epsilon)c^2$, afin d'estimer la valeur de la constante k en fonction du coefficient de variation. Des coefficients a_i ont été obtenus pour chacune des distributions et les résultats se sont avérés très satisfaisants.

A la lumière des travaux de Sahin, il semble que l'estimateur $M_{1n}(t)$ performe assez bien pour des valeurs de t relativement modérées et que la borne $t > 2\mu$ utilisée fréquemment soit assez bonne. Cependant, en pratique, on rencontre souvent

des valeurs de t beaucoup plus petites. Aussi, pour la distribution log-normale, l'estimation s'est moins bien comportée et semble nécessiter une barre plus haute pour les valeurs de t permises. Notons enfin que l'auteur n'a considéré que des coefficients de variation inférieurs à 2 et que les distributions en cause ont des densités unimodales. (avec le mode atteint en $t > 0$). Pour les autres cas, les conclusions pourraient être complètement fausses.

Par ailleurs, des approximations asymptotiques d'ordre supérieur ont également été développées (Sahin [16]). Par exemple, sous certaines conditions à propos de $F(t)$, on peut estimer $M(t)$ par:

$$M'_{1n}(t) = \frac{t}{\hat{\mu}} + \frac{\hat{\sigma}^2}{2\hat{\mu}^2} - \frac{1}{2} + \frac{S(t)}{\hat{\mu}^2} + \frac{Q * Q(t)}{\hat{\mu}^3} \quad (1.3.3)$$

où $S(t) = -\int_t^\infty Q(u)du$ et $Q(t) = \int_t^\infty [1 - F(u)]du$ (elles doivent être estimées car $F(u)$ est inconnue) et $*$ représente la convolution. $\hat{\mu}$ et $\hat{\sigma}$ sont des estimateurs obtenus à partir d'un échantillon de durées de vie.

Les quatrième et cinquième termes de (1.3.3) impliquent que le taux de convergence de l'estimateur asymptotique vers $M(t)$ dépend de la forme de la fonction de densité $f(t)$. Cette estimateur ne sera pas considéré dans notre étude.

Les contraintes sur la valeur de t par rapport à μ ne sont pas les seuls désavantages de $M_{1n}(t)$. Un deuxième inconvénient de cet estimateur est qu'il est, en général, non convergent. De plus, rappelons que, pour respecter l'hypothèse du théorème 1.3.2, $F(t)$ doit être non arithmétique. Pour le cas où $F(t)$ est arithmétique, on utilise plutôt l'estimateur proposé par Yang [21]

$$M_{1n}^*(t) = \frac{t}{\hat{\mu}} + \frac{\hat{\sigma}^2 + \hat{\mu} + \hat{\mu}^2}{2\hat{\mu}^2} - 1$$

Prenons quelques lignes pour aborder le cas où certaines données sont censurées. Schneider, Lin et Tang [18] se sont intéressés à cette situation particulière. Rappelons que nous sommes en présence de données censurées lorsque, après un certain temps T' , il est impossible de connaître la durée de vie d'un item. Le temps observé est alors posé égal à ce temps T' . La valeur de T' peut être différente pour chaque durée de vie et peut être une constante ou elle-même une variable aléatoire. Cet aspect est important en théorie de la fiabilité et du renouvellement, car souvent, après un certain temps T' , l'information n'est pas disponible. Par exemple, un manufacturier ne sait pas quand un objet se brise si la garantie est échuë. Si les variables aléatoires sont censurées, les moyennes et variances échantillonnales ne sont pas de bons estimateurs à employer pour obtenir $M_{1n}(t)$. Un estimateur plus complexe pour la moyenne a été développé par Sursala, Tsai et Van Ryzin [20] tandis que Schneider et Weissfeld [19] ont proposé un estimateur pour la variance σ^2 .

1.3.2. Deuxième estimateur de $M(t)$

Le deuxième estimateur que nous considérerons est un estimateur paramétrique. Si on connaît la famille paramétrique de la distribution F (gamma, Weibull, normale, ...), on peut trouver facilement un estimateur convergent de $M(t)$. Cet estimateur que nous allons présenter est un des premiers à avoir été employé, il y a plus de trente ans. Comme nous l'avons signalé précédemment, l'approche paramétrique a longtemps été la seule à être utilisée pour estimer la fonction de renouvellement.

Considérons la fonction de distribution $F(x; \alpha_1, \alpha_2, \dots, \alpha_p)$. On connaît la forme de F , mais pas les paramètres α_i , $i = 1, \dots, p$. On n'a qu'à estimer $\alpha_1, \alpha_2, \dots, \alpha_p$ par des estimateurs convergents $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$. On estime finalement $F^{(k)}(x; \alpha_1, \alpha_2, \dots, \alpha_p)$ par $\hat{F}_n^{(k)}(x; \hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p)$

Nous pouvons obtenir notre deuxième estimateur à l'aide de la proposition suivante (voir Frees [8]):

Proposition 1.3.1. *Soit $f = f(x; \alpha_1, \alpha_2, \dots, \alpha_p)$ la fonction de densité de $F = F(x; \alpha_1, \alpha_2, \dots, \alpha_p)$ où F est absolument continue. Supposons que f est continue en chaque paramètre α_i , $i = 1, \dots, p$. Soit $F_n^{(k)}$ la k^e convolution de F avec elle-même et soit $M_{2n}(t) = \sum_{k=1}^{\infty} \hat{F}_n^{(k)}(t)$, alors, $M_{2n}(t)$ est un estimateur convergent de $M(t)$.*

Notre deuxième estimateur est donc:

$$M_{2n}(t) = \sum_{k=1}^{\infty} \hat{F}_n^{(k)}(t) \quad (1.3.4)$$

Dans certains cas, cet estimateur s'obtient facilement, mais pour d'autres, il en est tout autrement. Par exemple, si $F(t)$ est Weibull, il peut être difficile d'obtenir les estimateurs du maximum de vraisemblance des paramètres. Le calcul de $F_n^{(k)}(t)$ est aussi difficile, mais des algorithmes existent pour le faire. Le programme CONVOL de Cléroux et McConalogue [7] en est un.

Le principal inconvénient de cet estimateur paramétrique est qu'il dépend des hypothèses de départ sur la distribution. Lorsqu'on connaît la distribution de laquelle proviennent les temps observés, des gains appréciables sont obtenus.

Si ce n'est pas le cas, $M_{2n}(t)$ est alors inefficace. La robustesse et la qualité des estimateurs des paramètres doivent aussi être prises en considération. Le choix d'un estimateur non paramétrique peut donc s'avérer plus judicieux dans certains cas.

Finalement, tout comme pour $M_{1n}(t)$, soulignons que l'estimateur paramétrique que nous venons de présenter peut être adapté si l'on est en présence de données censurées.

1.3.3. Troisième estimateur de $M(t)$

L'estimateur de la fonction de renouvellement le plus fréquemment utilisé dans la littérature a été proposé par Frees [8, 9] en 1986. Cet estimateur, le troisième que nous allons étudier, est non paramétrique et est basé sur le fait que le nombre de remplacements au temps t peut être estimé à partir de n durées de vie par

$$N_n(t) = \sum_{k=1}^n I(S_k \leq t)$$

où I est la fonction indicatrice.

Puisque $E[N_n(t)] = \sum_{k=1}^n \mathcal{P}(S_k \leq t)$, $[N_n(t)]$ défini ci-haut est un estimateur de $\sum_{k=1}^n \mathcal{P}(S_k \leq t)$, il serait naturel de prendre $I(S_k \leq t)$ comme estimateur sans biais de $\mathcal{P}(S_k \leq t)$ (pour $k = 1, \dots, n$). (Rappelons que $\mathcal{P}(S_k \leq t) = F^{(k)}(t)$.) Cependant, un meilleur estimateur existe et est basé sur le raisonnement suivant:

Pour estimer $\mathcal{P}(S_1 \leq t)$, un meilleur estimateur que $I(X_1 \leq t)$ est donné par

$$\frac{1}{n} \sum_{i=1}^n I(X_i \leq t)$$

De la même façon, pour $k = 2$, un meilleur estimateur de $\mathcal{P}(S_2 \leq t)$ est

$$\binom{n}{2}^{(-1)} \sum_{1 \leq i < j \leq n} I(X_i + X_j \leq t)$$

au lieu de $I(X_1 + X_2 \leq t)$.

On peut continuer ainsi pour obtenir un estimateur sans biais de $\mathcal{P}(S_k \leq t)$, soit

$$F_n^{(k)}(t) = \binom{n}{k}^{-1} \sum_c I(X_{i_1} + X_{i_2} + \dots + X_{i_k} \leq t) \quad (1.3.5)$$

où \sum_c est la somme sur toutes les combinaisons distinctes de (i_1, \dots, i_k) .

Tout comme $I(S_k \leq t)$, l'estimateur donné en (1.3.5) est un estimateur sans biais de $\mathcal{P}(S_k \leq t)$, mais le deuxième a une plus petite variance que le premier. En fait, $F_n^{(k)}(t)$ est une statistique U. $F_n^{(k)}(t)$ est sans biais et a plusieurs propriétés désirables. Entre autres, il est l'estimateur sans biais à variance minimale de $F^{(k)}(t)$. On peut alors considérer $\sum_{k=1}^n F_n^{(k)}(t)$ comme estimateur de la fonction de renouvellement.

La convergence très rapide de cet estimateur nous permet d'introduire le paramètre m . Ce paramètre est un entier positif qui dépend de n et est tel que $m \leq n$ et $m \rightarrow \infty$ quand $n \rightarrow \infty$. Il est alors suffisant de sommer jusqu'à m (plutôt que n). Ceci s'avérera très utile lors du calcul de l'estimateur, car beaucoup

de temps sera sauvé. En général, il sera possible de prendre un m beaucoup plus petit que n .

Notre troisième estimateur est donc:

$$M_{3n}(t) = \sum_{k=1}^m F_n^{(k)}(t) \quad (1.3.6)$$

où $F_n^{(k)}(t)$ est donné par l'équation (1.3.5).

Soulignons que $M_{3n}(t)$ n'est pas une statistique U. Ses propriétés asymptotiques ne sont donc pas immédiates. De nombreux efforts ont été accomplis, principalement par Frees [9], afin de vérifier les propriétés asymptotiques de cet estimateur non paramétrique. Dans ce qui suit, nous présenterons sans preuve les principaux résultats.

Le théorème suivant montre la convergence presque sûre de $M_{3n}(t)$ vers $M(t)$. Ce théorème est valable autant pour des valeurs de X positives que négatives. Cependant, en pratique, les variables aléatoires X représentent des durées de vie et seront positives. (Voir Frees [9] pour la preuve du théorème.)

Théorème 1.3.3. *Posons $X^- = \min(0, X)$. Supposons que F a une moyenne positive μ et une variance finie σ^2 . Considérons m , un entier positif qui dépend de n et est tel que $m \leq n$ et $m \rightarrow \infty$ quand $n \rightarrow \infty$. Supposons aussi qu'au moins une des deux propriétés suivantes soit vraie:*

$$\text{Pour } r \geq 2, E|X^-|^r < \infty \text{ et } n = O(m^{r-2})$$

ou

Pour un $\theta_1 > 0$ et $\forall |\theta| < \theta_1$, $E[\exp(-\theta X^-)] < \infty$ et $\log(n) = o(m)$.

Alors, $\forall t \geq 0$,

$$M_{3n}(t) \rightarrow M(t) \text{ presque sûrement.}$$

Par ailleurs, on peut montrer la normalité asymptotique de $M_{3n}(t)$ en utilisant le théorème 1.3.4. (voir Frees [9] pour la preuve)

Théorème 1.3.4. *Supposons que $\sigma^2 < \infty$. Considérons m , un entier positif qui dépend de n et est tel que $m \leq n$ et $m \rightarrow \infty$ quand $n \rightarrow \infty$. Supposons qu'au moins une des deux propriétés suivantes soit vraie:*

$$\text{Pour } r \geq 2, E|X^-|^r < \infty \text{ et } n = O(m^{2r-4})$$

ou

$$\text{Pour un } \theta_1 > 0 \text{ et } \forall |\theta| < \theta_1, E[\exp(-\theta X^-)] < \infty \text{ et } \log(n) = o(m).$$

Alors, $\forall t \geq 0$,

$$\sqrt{n} (M_{3n}(t) - M(t)) \rightarrow N(0, \sigma_1^2) \text{ en loi}$$

$$\text{où } \sigma_1^2 = \sum_{s=1}^{\infty} \sum_{r=1}^{\infty} r s \xi_{rs}(1) < \infty$$

$$\text{et } \xi_{rs}(c) = \text{Cov} (F^{(r-c)}(t - (X_1 + \dots + X_c)), F^{(s-c)}(t - (X_1 + \dots + X_c)))$$

Il est aussi possible d'estimer σ_1^2 en utilisant le théorème suivant (voir Frees [9] pour la preuve):

Théorème 1.3.5. Soit $m_1 = m_1(n)$ un entier positif dépendant de n tel que $m_1 \leq n$ et $m_1 \rightarrow \infty$ quand $n \rightarrow \infty$.

Posons

$$\sigma_n^2 = \sum_{r,s=1}^{m_1} r s \hat{\xi}_{rs} - \left(\sum_{k=1}^{m_1} k F_n^{(k)}(t) \right)^2$$

Alors,

si $\sigma_1^2 < \infty$, on a $\sigma_n^2 \rightarrow \sigma_1^2$, en probabilité.

On obtient facilement (voir Frees [9]) que $\frac{\sqrt{n}}{\sigma_n} (M_{3n}(t) - M(t)) \rightarrow N(0, 1)$ en loi. Nous pouvons alors obtenir aisément un intervalle de confiance asymptotique pour $M_{3n}(t)$.

Notons finalement que $M_{3n}(t)$ est avantageux si nous sommes en présence de données censurées. En effet, si dans $M_{3n}(t)$ on remplace X_i par $X'_i = \min(X_i, T)$ pour obtenir $M'_{3n}(T)$, alors $M'_{3n}(T)$ est convergent pour $M(T)$. Cet estimateur est beaucoup plus facile à obtenir que lorsqu'on utilise plutôt les méthodes asymptotiques ou paramétriques. L'inconvénient majeur de $M'_{3n}(T)$, est qu'il n'est valable que lorsque le temps où la fonction de renouvellement est évaluée est inférieur au temps de censure. Dans le cas contraire, par exemple quand les temps de censure sont aléatoires, l'estimateur de Frees adapté aux données censurées est inadéquat.

1.3.4. Quatrième estimateur de $M(t)$

Il est assez facile de constater que l'estimateur de Frees ($M_{3n}(t)$) décrit ci-haut pourra s'avérer fastidieux à obtenir, en particulier si n ou t sont grands. Une valeur de m le moins élevée augmentera aussi considérablement le

nombre de calculs à effectuer et, par conséquent, le temps de calcul de l'estimation. Il pourrait donc être très avantageux de développer un autre estimateur en employant une autre approche non paramétrique. La quatrième méthode d'estimation de $M(t)$ que nous allons considérer a initialement été introduite par Frees [8, 9], mais principalement développée par Lin, O'Connell et Schneider [13]. Leur but: obtenir des résultats comparables à ceux de l'estimateur de Frees en réduisant considérablement le temps de calcul.

L'estimateur que ces auteurs ont proposé est basé sur la fonction de distribution empirique $F_n^{(1)}(t)$, que nous noterons $\tilde{F}_n^{(1)}(t)$. Le quatrième estimateur sera donc appelé estimateur empirique dans ce qui suit. L'estimation de $F^{(k)}(t)$ est donnée récursivement par la relation

$$\tilde{F}_n^{(k)}(t) = \int_0^t \tilde{F}_n^{(k-1)}(t-u) d\tilde{F}_n^{(1)}(u) \quad (1.3.7)$$

Bien que $\tilde{F}_n^{(k)}(t)$ soit un estimateur biaisé de $F^{(k)}(t)$ (pour $k \geq 2$), il a l'avantage d'être l'estimateur du maximum de vraisemblance non paramétrique (voir Lin, O'Connell et Schneider [13]). De plus, $\tilde{F}_n^{(k)}(t)$ est une statistique V et est relié de près à la statistique U $F_n^{(k)}(t)$ (1.3.5).

Notre quatrième estimateur de la fonction de renouvellement est donc:

$$M_{4n}(t) = \sum_{k=1}^n \tilde{F}_n^{(k)}(t) \quad (1.3.8)$$

où $\tilde{F}_n^{(k)}(t)$ est donné par (1.3.7).

Par ailleurs, on peut montrer que $M_{4n}(t)$ est convergent et a la même distribution asymptotique que $M_{3n}(t)$. Encore une fois, la convergence très rapide de cet estimateur nous permet de ne pas sommer tous les n termes de (1.3.8), mais seulement les m premiers où m est défini comme pour $M_{3n}(t)$.

L'estimateur de $F^{(k)}(t)$ donné par (1.3.7) n'est pas nécessairement simple à obtenir. Cependant, des algorithmes existent afin d'y parvenir. Lin, O'Kinneide et Schneider [13] en ont développé un. Dans le but d'obtenir $M_{4n}(t)$, on peut considérer l'équation fondamentale du renouvellement énoncée précédemment et donnée par:

$$M(t) = F(t) + \int_0^t M(t-x)dF(x) \quad (1.3.9)$$

Pour résoudre cette équation, on peut discrétiser l'intervalle $[0, t]$. L'intégrale pourra alors être remplacée par une somme finie, ce qui facilitera le calcul de $M_{4n}(t)$. Afin de discrétiser l'intervalle $[0, t]$, on pourrait simplement arrondir les temps observés à l'unité la plus près. De la précision risque alors d'être perdue. Une autre technique consiste à déplacer la position de la virgule vers la droite en multipliant par une constante c où c vaut 10, 100, etc. Pour évaluer $M_{4n}(t)$, il suffit donc d'évaluer $M_{4n}^*(ct)$. Il ne reste alors plus qu'à résoudre:

$$M_{4n}^*(t) = \tilde{F}^*(t) + \sum_{k=1}^t \hat{M}_{4n}^*(t-k)\tilde{F}^*(k) - \tilde{F}^*(k) \quad (1.3.10)$$

L'avantage de cet algorithme proposé par Lin, O'Kinneide et Schneider est qu'il s'obtient relativement rapidement et qu'il ne nécessite pas l'utilisation du

paramètre m . Par contre, pour obtenir $M_{4n}(t)$, il faut estimer la fonction de renouvellement pour toutes les valeurs X_i inférieures à t . Ceci requiert plus de temps pour de grande valeurs de t .

Notre quatrième estimateur devient donc, en appliquant cet algorithme,

$$M_{4n}(t) = M_{4n}^*(ct) \tag{1.3.11}$$

où $M_{4n}^*(t)$ est défini par (1.3.10). (Nous n'utiliserons donc plus (1.3.8).)

A nouveau, comme pour les estimateurs précédents, $M_{4n}(t)$ peut être adapté au cas où certaines variables sont censurées. Il suffit de prendre l'estimateur de Kaplan-Meier pour estimer la fonction de répartition. Cet estimateur est fréquemment utilisé en fiabilité et surtout en analyse de survie, lorsque des données sont censurées.

1.3.5. Cinquième estimateur de $M(t)$

Le cinquième estimateur que nous allons considérer est semi-paramétrique. Il est un cas intermédiaire entre l'estimateur paramétrique pour lequel on possède de bonnes connaissances à propos de la fonction de répartition des temps observés et les estimateurs non paramétriques où aucune information concernant $F(t)$ n'est disponible. Pour cet estimateur, on ne connaît pas, a priori, la famille de distribution de F , mais on sait que le taux de panne est monotone (Il est plus fréquent d'avoir un taux de panne non décroissant, mais il existe aussi un estimateur pour le cas où il est non croissant.). Cette hypothèse sur $F(t)$ est beaucoup moins contraignante que sa connaissance précise et complète. En théorie de la

fiabilité, il est souvent possible, en pratique, d'affirmer que le taux de panne est non décroissant (ou plus rarement non croissant).

Avant d'aller plus loin, rappelons quelques notions concernant le taux de panne. Le taux de panne est en fait la probabilité conditionnelle qu'une pièce ou un système tombe en panne à un instant précis donné, étant donné qu'il a survécu jusqu'alors. Voici quelques définitions et propriétés relatives au taux de panne (issues de Marshall et Proschan [14] et Prakasa Rao [15]).

Définition 1.3.1. *Le taux de panne est donné par $r(x) = \frac{f(x)}{1-F(x)}$*

Définition 1.3.2. *F est IFR si le taux de panne de F est non décroissant. Inversement, F est DFR si le taux de panne de F est non croissant.*

Proposition 1.3.2. *Si F est IFR, $R(x) = \log(1 - F(x))$ est convexe.*

Proposition 1.3.3. *Si F est IFR et $F(z) < 1$, alors, F est absolument continue sur $(-\infty, z)$.*

Proposition 1.3.4. *Pour toute distribution F et pour tout x tel que r est fini sur $[-\infty, x]$, on a:*

$$1 - F(x) = \exp[-R(x)] = \exp\left[-\int_{-\infty}^x r(z)dz\right]$$

Soulignons enfin que la connaissance du taux de panne $r(t)$ implique aussi la connaissance de la fonction de distribution $F(t)$ et, par conséquent, la connaissance de $M(t)$, la fonction de renouvellement.

La démarche pour obtenir le cinquième estimateur consiste donc à estimer la fonction de répartition en estimant dans un premier temps le taux de panne. Pour ce faire, nous utiliserons les résultats proposés par Grenander [10] et repris par Marshall et Proschan [14] et par Prakasa Rao [15]. Soulignons que tous ces travaux ont porté uniquement sur l'estimation du taux de panne et de la fonction de répartition et non sur la fonction de renouvellement. La relation de $r(t)$ et $F(t)$ avec $M(t)$ nous permettra d'appliquer leurs résultats à l'estimation de la fonction de renouvellement. Notons que cette façon de procéder pour estimer $M(t)$ n'a jamais été employée par les auteurs que nous avons cités et n'a donc jamais été comparée à l'aide de simulations.

L'estimateur de $F(t)$ que Marshall et Proschan [14] et Prakasa Rao [15] ont obtenu est, à toute fin pratique, l'estimateur du maximum de vraisemblance. Voici les principales étapes de leur raisonnement.

Premièrement, considérons \mathcal{F} , la classe des distributions IFR et $X_1 \leq X_2 \leq \dots \leq X_n$, les statistiques d'ordre obtenues en ordonnant les différentes durées de vie observées. Marshall et Proschan ont soulevé le point qu'il est impossible d'obtenir l'estimateur du maximum de vraisemblance de F (et par conséquent de r) en maximisant simplement $\prod_{i=1}^n f(x_i)$. Ceci s'explique par le fait que $f(X_n)$ peut être arbitrairement grand. Ils ont donc plutôt considéré la sous-classe \mathcal{F}^M des distributions de \mathcal{F} . Les distributions constituant \mathcal{F}^M sont celles pour lesquelles le taux de panne peut être borné par une certaine constante M . En procédant ainsi, la fonction de vraisemblance pourra être maximisée (pour une distribution $F \in \mathcal{F}^M$) et cette distribution pour laquelle la fonction de vraisemblance est maximale est unique. De plus, on peut montrer que l'estimateur \hat{F}_n^M ($\hat{F}_n^M \in \mathcal{F}^M$)

converge en distribution vers \hat{F}_n ($\hat{F}_n \in \mathcal{F}$) lorsque $M \rightarrow \infty$. Des estimateurs du maximum de vraisemblance pour la fonction de densité f et pour le taux de panne r peuvent être obtenus en suivant le même raisonnement.

En utilisant la proposition 1.3.4, on peut montrer (voir Grenander [10]) que la fonction log-vraisemblance est donnée par:

$$L(F) = \sum_{i=1}^n \log(r(X_i)) - \sum_{i=1}^n \int_{-\infty}^{X_i} r(z) dz$$

On peut maximiser la fonction $L(F)$ sur \mathcal{F}^M de la façon suivante proposée par Grenander [10]. Considérons une distribution F ($F \in \mathcal{F}^M$) avec taux de panne r . On considère ensuite la distribution F^* , une distribution avec un taux de panne constant entre chaque observation. Ce taux de panne est donné par:

$$r^*(x) = \begin{cases} 0 & \text{si } x < X_1 \\ r(X_i) & \text{si } X_i \leq x < X_{i+1} \\ r(X_n) & \text{si } x \geq X_n \end{cases}$$

On obtient alors que $F^* \in \mathcal{F}^M$ et que $r(x) \geq r^*(x)$. Il s'ensuit alors que $\int_{-\infty}^x r(z) dz \geq \int_{-\infty}^x r^*(z) dz$. Ceci a pour effet que $L(F) \leq L(F^*)$. Cette relation est la base de leur estimateur.

La fonction log-vraisemblance (L) peut donc être remplacée par

$$L = \sum_{i=1}^n \log(r(X_i)) - \sum_{i=1}^{n-1} (n-i)(X_{i+1} - X_i)r(X_i)$$

En maximisant cette fonction sous la contrainte $r(X_1) \leq r(X_2) \leq \dots \leq r(X_n)$ Grenander [10] a montré que l'estimateur de r (pour une distribution F de la sous-classe \mathcal{F}^M) est donné par:

$$\hat{r}_n^M(X_i) = \min\left(\min_{v \geq i+1} \max_{u \leq i} \left\{ \frac{1}{v-u} [r_u^{-1} + \dots + r_{v-1}^{-1}] \right\}^{-1}, M\right)$$

où $r_j = [(n-j)(X_{j+1} - X_j)]^{-1}$ pour $j = 1, 2, \dots, n-1$ et $r_n = M$.

Notons que si $r_i > r_{i+1}$ pour certains i , on remplace r_i et r_{i+1} par leur moyenne harmonique $((r_i^{-1} + r_{i+1}^{-1})^{-1})$. Ceci permet d'obtenir un estimateur non décroissant. L'estimateur recherché (pour $F \in \mathcal{F}^M$) est donné par

$$\hat{r}_n(x) = \begin{cases} 0 & \text{si } x < X_1 \\ \min(r_{n_i+1, n_{i+1}}, M) & \text{si } X_{n_i+1} \leq x < X_{n_{i+1}+1} \\ M & \text{si } x \geq X_n \end{cases}$$

L'estimateur pour une distribution $F \in \mathcal{F}^M$ est alors obtenu en prenant la limite quand $M \rightarrow \infty$. On obtient donc que l'estimateur du maximum de vraisemblance du taux de panne $r(x)$ (dans le cas IFR) est donné par:

$$\hat{r}_n(x) = \begin{cases} 0 & \text{si } x < X_1 \\ \hat{r}_n(X_i) & \text{si } X_i \leq x < X_{i+1} \\ \infty & \text{si } x \geq X_n \end{cases} \quad (1.3.12)$$

où

$$\hat{r}_n(X_i) = \min_{v \geq i+1} \max_{u \leq i} \left[(v-u) \sum_{j=u}^{v-1} (n-j)(X_{j+1} - X_j) \right]^{-1} \quad (1.3.13)$$

En utilisant la définition 1.3.1 et la proposition 1.3.4, on obtient des estimateurs de $f(t)$ et de $F(t)$. Le théorème suivant prouvé par Marshall et Proschan [14] montre la convergence de $\hat{F}_n(t)$.

Théorème 1.3.6. *Si le taux de panne r est non décroissant, alors $\forall t, \lim_{n \rightarrow \infty} \hat{F}_n(t) = F(t)$.*

On peut ensuite trouver $\hat{F}_n^{(k)}(t)$ ($k = 1, \dots, m$) pour obtenir, en sommant sur $k = 1, \dots, m$ les convolutions, un estimateur de la fonction de renouvellement.

Notre cinquième estimateur est donc :

$$M_{5n}(t) = \sum_{k=1}^m \hat{F}_n^{(k)}(t) \quad (1.3.14)$$

où $\hat{F}_n^{(k)}(t)$ est obtenu en estimant d'abord le taux de panne par (1.3.12). Le paramètre m est à nouveau celui présenté pour $M_{3n}(t)$.

Rappelons qu'un estimateur semblable est aussi disponible pour le cas où le taux de panne est non croissant (DFR). Cependant, dans cette situation, contrairement au cas IFR, l'estimateur $\hat{r}_n(x)$ n'est pas unique. Nous n'insisterons pas davantage sur ce cas, puisqu'il est plus rare et moins utile en pratique.

Chapitre 2

COMPARAISON DES ESTIMATEURS

Au chapitre précédent, nous avons présenté les principales propriétés et caractéristiques de cinq estimateurs de la fonction de renouvellement fréquemment utilisés. Le présent chapitre est consacré à leur comparaison. Nous discuterons des avantages et inconvénients de chacun d'eux et les comparerons en résumant la littérature existante sur ce sujet. Des simulations seront effectuées plus loin dans ce chapitre.

Comme nous l'avons vu au premier chapitre, nous pouvons comparer les cinq estimateurs sur plusieurs aspects. Par exemple, leur simplicité et leur temps de calcul, s'ils nécessitent des hypothèses sur la distribution, leurs propriétés asymptotiques, etc.

L'estimateur le plus simple à obtenir est sans contredit l'estimateur asymptotique de $M(t)$ ($M_{1n}(t)$). Cet estimateur nécessite uniquement la connaissance des deux premiers moments de la distribution de durée de vie échantillonnale et est le seul qui n'implique pas le calcul de convolutions. Etant basé sur un résultat asymptotique, il n'est cependant valable que pour d'assez grandes valeurs de t (par rapport à μ), ce qui n'est souvent pas le cas. De plus, il est généralement non convergent. Il y a donc un prix à payer pour la simplicité de $M_{1n}(t)$. Bien que cet

estimateur donne relativement de bons résultats, il est souvent inadéquat et il est préférable d'employer un des quatre autres. Le choix d'un estimateur parmi ces derniers dépend, entre autres, des connaissances que l'on possède a priori sur la distribution. Si la loi exacte de distribution est connue, l'estimateur paramétrique ($M_{2n}(t)$) est le plus approprié, alors que si le taux de panne est IFR, $M_{5n}(t)$ peut donner de bons résultats.

2.1. RÉSUMÉ DE LA LITTÉRATURE

Commençons notre survol de la littérature en présentant un exemple fréquemment utilisé par plusieurs auteurs [2, 8, 9, 13]. Cet exemple est celui des 105 observations de durée de vie d'une composante électronique provenant de Juran & Gryna [11, p.171]. La moyenne de la distribution de cet échantillon est $\bar{x} = 55.603$ unités. Selon les caractéristiques de la fonction de taux de panne, il est justifié d'estimer la fonction de renouvellement aux instants $t = 20$ et $t = 100$.

A la lumière de ce jeu de données, Frees[8] a montré l'efficacité de $M_{1n}(t)$ pour $t = 100$ (t grand par rapport à \bar{x}). Pour $t = 20$, comme prévu, les résultats obtenus sont nettement différents et certainement moins satisfaisants, comme nous l'avons justifié précédemment en discutant du cas des petites valeurs de t . Pour sa part, l'estimateur paramétrique $M_{2n}(t)$ a été calculé en utilisant la distribution de Weibull. Ce dernier a mieux performé que $M_{1n}(t)$, pour un petit t . Pour $t = 100$, le taux de panne un peu particulier (voir Frees [9]) met en doute l'hypothèse de la distribution de Weibull. Le nombre moyen de renouvellements est alors surévalué.

L'estimateur $M_{3n}(t)$ s'avère donc être, dans ce cas-ci, le meilleur estimateur. Le résultat obtenu pour $M_{3n}(100)$ est comparable à celui obtenu pour $M_{1n}(100)$. Pour $t = 20$, on sait que $M_{3n}(t)$ est meilleur que $M_{1n}(t)$. Pour cet estimateur, il a suffi de calculer uniquement les cinq premières convolutions (i.e. prendre $m = 5$).

Par ailleurs, Lin, O'Conneide et Schneider [13] ont calculé $M_{4n}(t)$ ($t = 20$ et $t = 100$) pour ce jeu de données. Les résultats obtenus sont très semblables à ceux obtenus par Frees pour $M_{3n}(t)$. Rappelons cependant que le nombre considérable de calculs pour obtenir $M_{3n}(t)$ et le temps nécessaire pour les effectuer avaient poussé les auteurs à développer l'estimateur empirique $M_{4n}(t)$, une approche initialement proposée par Frees [8, 9]. Ils ont atteint leur objectif. En effet, en estimant la fonction de renouvellement pour l'exemple de Juran et Gryna, les temps de calculs (en CPU) ont été réduits de plus de 1000 fois. Comme on peut le constater dans le tableau 2.1.1, la valeur des estimations se rapproche beaucoup de celles de Frees.

Le tableau suivant regroupe les quatre estimateurs de $M(t)$ pour ces données, obtenues pour $t = 20$ et $t = 100$.

TAB. 2.1.1. Estimation de $M(t)$ pour l'exemple de Juran & Gryna

	$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$
$t = 20$	0,1717	0,3456	0,4619	0,46
$t = 100$	1,6105	1,7769	1,4798	1,44

L'exemple précédent permet d'avoir certaines indications sur chacun des estimateurs, mais il n'est qu'un exemple. Afin d'avoir une meilleure idée sur la performance des différents estimateurs de $M(t)$, des simulations peuvent être effectuées. Dans la littérature, on compare les estimateurs et on mesure leur performance en calculant le biais et l'erreur quadratique moyenne.

$$\text{Biais moyen (BM)} = \frac{1}{r} \sum_{i=1}^r (\hat{M}_i(t) - M(t)) \quad (2.1.1)$$

$$\text{Err. quad. moyenne (EQM)} = \frac{1}{r} \sum_{i=1}^r (\hat{M}_i(t) - M(t))^2 \quad (2.1.2)$$

A l'aide de simulations de Monte Carlo, Frees [8] a simulé tour à tour des échantillons provenant des distributions les plus fréquemment utilisées en théorie du renouvellement. Que ce soit pour la distribution gamma, Weibull, normale tronquée ou log-normale, 500 échantillons ont été simulés. Dans cet article, il a comparé les comportements de $M_{1n}(t)$ et de $M_{3n}(t)$ (et de $M_{2n}(t)$ pour la distribution de Weibull). Les résultats qu'il a obtenus sont regroupés dans les tableaux 2.1.2 à 2.1.5.

De façon générale, l'estimateur $M_{3n}(t)$ (de Frees) performe mieux que $M_{1n}(t)$, (l'estimateur asymptotique). Le biais est généralement beaucoup plus petit (souvent 10 à 20 fois plus petit). Cependant, un résultat plus surprenant est la plus faible différence entre les erreurs quadratiques moyennes (EQM) des deux estimateurs. Bien que l'EQM de $M_{3n}(t)$ soit tout de même le meilleur, ils sont souvent plus semblables. De plus, pour $M_{1n}(t)$ et à l'exception de la distribution gamma, le biais a tendance à décroître avec l'augmentation de t . Le phénomène inverse

TAB. 2.1.2. Estimation de $M(t)$ pour la loi Gamma (littérature)

Gamma($\alpha = 0,55, \beta = 1,0, \mu = 0,55$)					
n	t	$M_{1n}(t)^1$		$M_{3n}(t)^1$	
		Biais	EQM	Biais	EQM
10	0,25	-0,0064	0,2000	-0,0157	0,1346
	0,50	0,0347	0,4157	-0,0169	0,3158
	0,75	0,1054	0,7825	-0,0048	0,5864
	1,00	0,1877	1,3060	-0,0061	1,0510
	1,25	0,2758	1,9890	-0,0334	1,5210
15	0,25	0,0346	0,1642	0,0006	0,0980
	0,50	0,0440	0,2688	0,0033	0,2110
	0,75	0,0831	0,4444	0,0091	0,3722
	1,00	0,1337	0,6930	0,0135	0,7033
	1,25	0,1901	1,0160	-0,0099	1,0730
20	0,25	0,0332	0,1248	-0,0005	0,0667
	0,50	0,0269	0,1918	-0,0068	0,1468
	0,75	0,0501	0,3068	-0,0022	0,2672
	1,00	0,0850	0,4705	0,0022	0,5201
	1,25	0,1256	0,6839	-0,0208	0,8142
25	0,25	0,0561	0,1080	0,0033	0,0573
	0,50	0,0407	0,1608	0,0000	0,1226
	0,75	0,0548	0,2528	0,0038	0,2218
	1,00	0,0806	0,3833	0,0088	0,4555
	1,25	0,1122	0,5528	-0,0131	0,7273
30	0,25	0,0660	0,1011	0,0005	0,0497
	0,50	0,0432	0,1403	-0,0027	0,1032
	0,75	0,0501	0,2098	-0,0006	0,1845
	1,00	0,0685	0,3081	0,0042	0,3926
	1,25	0,0927	0,4355	-0,0170	0,6500

¹: Frees [7]

TAB. 2.1.3. Estimation de $M(t)$ pour la loi de Weibull (littérature)

		Weibull($\alpha = 3, 0, \beta = 1, 0, \mu = 0, 89$)							
n	t	$M_{1n}(t)^1$		$M_{2n}(t)^1$		$M_{3n}(t)^1$		$M_{4n}(t)^2$	
		Biais	EQM	Biais	EQM	Biais	EQM	Biais	EQM
10	0,25	-0,1716	0,0319	0,0011	0,0004	-0,0012	0,0015	0,002	0,002
	0,50	0,0091	0,0006	-0,0071	0,0056	0,0009	0,0098	0,011	0,012
	0,75	0,0585	0,0157	-0,0099	0,0194	0,0016	0,0238	0,023	0,025
	1,00	0,0217	0,0280	0,0121	0,0279	-0,0008	0,0318	0,029	0,031
	1,25	-0,0071	0,0306	0,0103	0,0320	-0,0035	0,0307	0,028	0,034
15	0,25	-0,1708	0,0310	0,0010	0,0003	-0,0017	0,0009		
	0,50	0,0090	0,0047	-0,0049	0,0043	-0,0013	0,0071		
	0,75	0,0575	0,0122	-0,0071	0,0137	0,0036	0,0177		
	1,00	0,0199	0,0150	0,0090	0,0190	0,0009	0,0211		
	1,25	-0,0098	0,0219	0,0118	0,0208	-0,0020	0,0219		
20	0,25	-0,1700	0,0303	0,0005	0,0001	-0,0012	0,0007	0,002	0,001
	0,50	0,0102	0,0036	-0,0032	0,0031	0,0001	0,0053	0,006	0,006
	0,75	0,0591	0,0103	-0,0031	0,0104	0,0043	0,0130	0,016	0,013
	1,00	0,0219	0,0118	0,0106	0,0151	0,0034	0,0174	0,022	0,017
	1,25	-0,0075	0,0170	0,0159	0,0151	0,0035	0,0171	0,023	0,018
25	0,25	-0,1708	0,0303	0,0002	0,0001	-0,0013	0,0006		
	0,50	0,0081	0,0027	-0,0038	0,0024	-0,0012	0,0038		
	0,75	0,0556	0,0082	-0,0054	0,0079	0,0001	0,0104		
	1,00	0,0171	0,0087	0,0048	0,0110	-0,0012	0,0130		
	1,25	-0,0135	0,0127	0,0096	0,0111	-0,0013	0,0125		
30	0,25	-0,1720	0,0305	-0,0001	0,0001	-0,0014	0,0005	0,002	0,001
	0,50	0,0058	0,0024	-0,0051	0,0021	-0,0039	0,0031	0,007	0,004
	0,75	0,0522	0,0072	-0,0085	0,0069	-0,0035	0,0090	0,015	0,008
	1,00	0,0125	0,0075	-0,0006	0,0096	-0,0046	0,0114	0,018	0,012
	1,25	-0,0193	0,0113	0,0036	0,0096	-0,0047	0,0111	0,018	0,012

¹: Frees [7], ²: Schneider, Lin et O'Kinneide [13]

TAB. 2.1.4. Estimation de $M(t)$ pour la loi normale tronquée (littérature)

Normale tronquée($\mu = 0,78, \sigma = 1,0$)					
n	t	$M_{1n}(t)^1$		$M_{3n}(t)^1$	
		Biais	EQM	Biais	EQM
10	0,25	-0,1160	0,0471	0,0130	0,0277
	0,50	-0,0325	0,0669	0,0084	0,0628
	0,75	0,0250	0,1147	0,0108	0,1079
	1,00	0,0649	0,1823	0,0101	0,1666
	1,25	0,0940	0,2670	0,0113	0,2290
15	0,25	-0,1218	0,0375	0,0110	0,0184
	0,50	-0,0444	0,0472	0,0009	0,0440
	0,75	0,0071	0,0788	0,0024	0,0786
	1,00	0,0490	0,1251	0,0069	0,1223
	1,25	0,0639	0,1833	0,0117	0,1658
20	0,25	-0,1214	0,0297	0,0075	0,0126
	0,50	-0,0478	0,0314	0,0004	0,0289
	0,75	-0,0001	0,0502	0,0044	0,0517
	1,00	0,0299	0,0791	0,0069	0,0823
	1,25	0,0491	0,1157	0,0104	0,1106
25	0,25	-0,1186	0,0260	0,0090	0,0103
	0,50	-0,0471	0,0249	0,0042	0,0248
	0,75	-0,0015	0,0386	0,0062	0,0406
	1,00	0,0264	0,0607	0,0120	0,0636
	1,25	0,0435	0,0885	0,0132	0,0850
30	0,25	-0,1196	0,0242	0,0088	0,0085
	0,50	-0,0497	0,0210	0,0028	0,0201
	0,75	-0,0057	0,0316	0,0039	0,0334
	1,00	0,0205	0,0495	0,0086	0,0524
	1,25	0,0360	0,0721	0,0111	0,0708

¹: Frees [7]

TAB. 2.1.5. Estimation de $M(t)$ pour la loi log-normale (littérature)

		Log-normale ($\mu = 1,6, \sigma = 1,0$)					
n	t	$M_{1n}(t)^1$		$M_{3n}(t)^1$		$M_{4n}(t)^2$	
		Biais	EQM	Biais	EQM	Biais	EQM
10	0,25	0,0642	0,1045	-0,0001	0,0077	0,008	0,009
	0,50	0,0654	0,0995	0,0149	0,0255	0,015	0,028
	0,75	0,0579	0,1018	0,0128	0,0422	0,027	0,053
	1,00	0,0534	0,1129	0,0163	0,0659	0,036	0,086
	1,25	0,0528	0,1329	0,0117	0,0918	0,050	0,121
15	0,25	0,1351	0,1537	0,0001	0,0054		
	0,50	0,1279	0,1427	0,0123	0,0179		
	0,75	0,1121	0,1353	0,0099	0,0296		
	1,00	0,0992	0,1346	0,0133	0,0443		
	1,25	0,0904	0,1405	0,0093	0,0619		
20	0,25	0,1750	0,1706	-0,0010	0,0038	0,005	0,004
	0,50	0,1646	0,1582	0,0094	0,0124	0,007	0,012
	0,75	0,1456	0,1472	0,0095	0,0218	0,010	0,023
	1,00	0,1296	0,1415	0,0141	0,0329	0,013	0,035
	1,25	0,1176	0,1411	0,0126	0,0452	0,018	0,052
25	0,25	0,2024	0,1838	0,0016	0,0029		
	0,50	0,1909	0,1693	0,0111	0,0101		
	0,75	0,1707	0,1555	0,0123	0,0188		
	1,00	0,1535	0,1467	0,0179	0,0277		
	1,25	0,1403	0,1430	0,0185	0,0384		
30	0,25	0,2160	0,1857	0,0009	0,0023	0,007	0,003
	0,50	0,2015	0,1698	0,0079	0,0080	0,011	0,008
	0,75	0,1783	0,1537	0,0067	0,0151	0,014	0,016
	1,00	0,1581	0,1422	0,0096	0,0218	0,017	0,024
	1,25	0,1419	0,1352	0,0112	0,0305	0,017	0,034

¹: Frees [7], ²: Schneider, Lin et O'Connell [13]

se produit par contre pour l'EQM qui croît en même temps que t (autant pour $M_{1n}(t)$ que pour $M_{3n}(t)$).

Par ailleurs, l'estimateur $M_{3n}(t)$ a été obtenu à partir des cinq, puis des dix premières convolutions de la fonction de répartition avec elle-même ($m = 5$ et $m = 10$). En raison de la convergence très rapide de $M_{3n}(t)$, nous pouvons noter que très peu de différences entre les deux estimateurs. En fait, de légères variations surviennent lorsque t est petit par rapport à la moyenne μ . De façon générale, plus t est petit par rapport à la moyenne, plus on peut se permettre une petite valeur de m . Les tableaux précédents regroupent les résultats avec $m = 5$.

L'estimateur paramétrique $M_{2n}(t)$ a été calculé pour la distribution Weibull. Sans surprise, cette estimation de $M(t)$ s'est avérée de beaucoup supérieure à $M_{1n}(t)$. La différence avec $M_{3n}(t)$ est cependant moins grande. A l'exception de $t = 0.25$, le biais de ce dernier est généralement plus petit que celui de $M_{2n}(t)$. L'erreur quadratique moyenne de l'estimateur paramétrique est toutefois plus petite. Ceci est plus rassurant, en raison des fortes hypothèses de départ sur la distribution.

Lin, O'Conneide et Schneider [13] ont, pour leur part, comparé l'efficacité de $M_{4n}(t)$ par rapport à $M_{3n}(t)$. Des simulations ont aussi été accomplies pour les distributions de Weibull et log-normales. Les résultats sont disponibles dans les tableaux 2.1.3 et 2.1.5. Alors que les biais obtenus pour $M_{4n}(t)$ sont légèrement plus élevés que ceux de $M_{3n}(t)$, les erreurs quadratiques moyennes se sont avérées plutôt semblables. Le gros avantage du quatrième estimateur est sans contredit son temps de calcul relativement court, comme l'illustre le tableau 2.1.6 ci-dessous.

TAB. 2.1.6. Temps en secondes du calcul de $M_{3n}(t)$ et de $M_{4n}(t)$

n	$M_{3n}(t)$	$M_{4n}(t)$
10	1,703	0,013
20	3,315	0,014
30	4,909	0,015

Soulignons cependant que, en raison de la puissance sans cesse croissante des ordinateurs, cet avantage de $M_{4n}(t)$ prend de moins en moins d'importance. Un aspect positif demeure toutefois; on ne doit pas spécifier le paramètre m .

En résumé, voici donc quelques règles d'utilisation, issues de la littérature, des quatre premiers estimateurs que nous venons de présenter. Premièrement, lorsque la distribution de durée de vie $F(t)$ est connue, l'estimateur paramétrique $M_{2n}(t)$ est l'estimateur le plus approprié. Par contre, si $F(t)$ est inconnue, le choix d'un autre estimateur s'impose. Si t est suffisamment grand (par rapport à μ), $M_{1n}(t)$ est un choix intéressant. Il est le plus simple à calculer et a donné de bons résultats dans les exemples disponibles dans la littérature. Cependant, si t est trop petit (très fréquent en pratique), $M_{3n}(t)$ ou $M_{4n}(t)$ sont des choix plus judicieux. Ces deux estimateurs ont été comparés par Lin, O'Conneide et Schneider [13]. Ils ont montré que le biais de $M_{4n}(t)$ est plus grand que celui de $M_{3n}(t)$, tandis que leurs écarts quadratiques moyens sont comparables. Par contre, $M_{4n}(t)$ requiert un plus petit temps de calcul que $M_{3n}(t)$ et son utilisation ne nécessite pas la détermination du paramètre m . Cet estimateur est donc plus avantageux à plusieurs égards. Cependant, dans certains cas extrêmes, tel un taux de panne rapidement

décroissant, $M_{4n}(t)$ performe moins bien. Il est alors préférable d'utiliser l'estimateur $M_{3n}(t)$. Rappelons enfin que, à notre connaissance, le cinquième estimateur présenté dans ce mémoire n'a jamais été comparé avec les quatre premiers.

2.2. SIMULATIONS

Les simulations et comparaisons disponibles dans la littérature sont représentatives, mais incomplètes. En plus de notre estimateur pour le cas IFR ($M_{5n}(t)$) que personne n'a étudié, aucun auteur n'a comparé tous les estimateurs simultanément. C'est ce que nous faisons maintenant.

Dans cette section, nous comparons, à notre tour, l'efficacité des estimateurs décrits précédemment. Des simulations ont été effectuées pour trois des distributions les plus fréquemment utilisées en fiabilité, soit la gamma, la Weibull et la normale tronquée.

2.2.1. Plan de simulation

Pour les trois distributions considérées, les paramètres ont été déterminés de telle sorte que la moyenne de l'échantillon (durée de vie moyenne) soit de 9080 heures et que l'écart-type soit de 3027 heures. Les paramètres sont disponibles au tableau 2.2.1.

Pour chaque simulation, 1000 échantillons ont été simulés. Dans le but de mesurer l'impact de la taille de l'échantillon sur les résultats, nous avons considéré

TAB. 2.2.1. Paramètres des distributions simulées

Gamma	$\alpha = 9$ et $\lambda = 0,000990971$
Weibull	$\alpha = 3,3031$ et $\lambda = 0,588990E - 13$
Normale tronquée	Valeurs positives d'une $N(9065, 3049^2)$

tour à tour des échantillons de taille 50, 100 et 200. Tous ces échantillons ont été obtenus à l'aide du logiciel S-Plus et soulignons qu'ils sont tous indépendants les uns des autres.

Le logiciel S-Plus a aussi servi pour calculer l'estimateur asymptotique, soit $M_{1n}(t)$. Le langage Fortran a, quant à lui, été utilisé pour calculer les autres estimateurs de $M(t)$. Le programme CONVOL, développé par Cléroux et McConalogue [6] a, pour sa part, permis de calculer les vraies valeurs de la fonction $M(t)$. Tous les programmes utilisés constituent l'annexe B. Ces programmes sont facilement utilisables et il est possible de les adapter à plusieurs distributions en ne changeant que quelques lignes.

Afin de mesurer la précision de chaque estimateur, nous avons, tout comme pour les simulations disponibles dans la littérature, calculé le biais et l'erreur quadratique moyenne des estimateurs pour différentes valeurs de t (équations 2.1.1 et 2.1.2). Par ailleurs, nous avons en plus évalué l'erreur relative de ces estimateurs. Elle est donnée par:

$$\text{Erreur relative moyenne (ERM)} = \frac{1}{r} \sum_{i=1}^r \frac{|\hat{M}_i(t_k) - M(t_k)|}{M(t_k)} \quad (2.2.1)$$

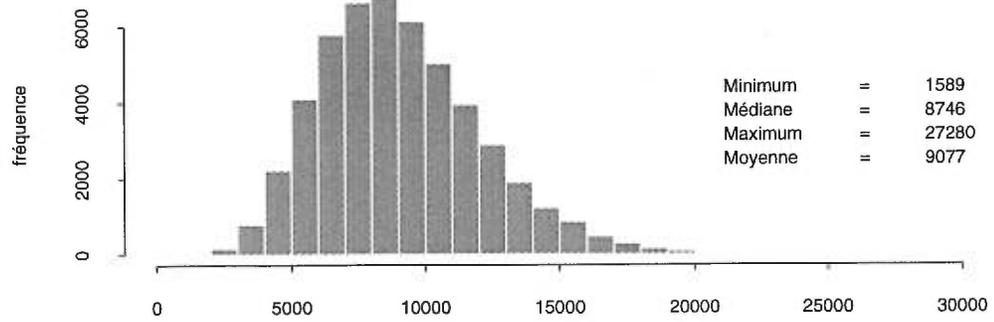
La moyenne et l'écart-type des distributions utilisées étant respectivement de 9080 et 3027, les valeurs de t considérées, pour les simulations, ont été de 3000, 4000, 5000, 7000, 10 000 et 15 000. Pour donner une meilleure idée de la répartition des temps simulés, des histogrammes sont disponibles aux figures 2.2.1 à 2.2.3.

Mentionnons enfin que pour satisfaire aux exigences de l'estimateur basé sur l'estimation du taux de panne, les distributions considérées sont toutes à taux de panne non décroissant (IFR).

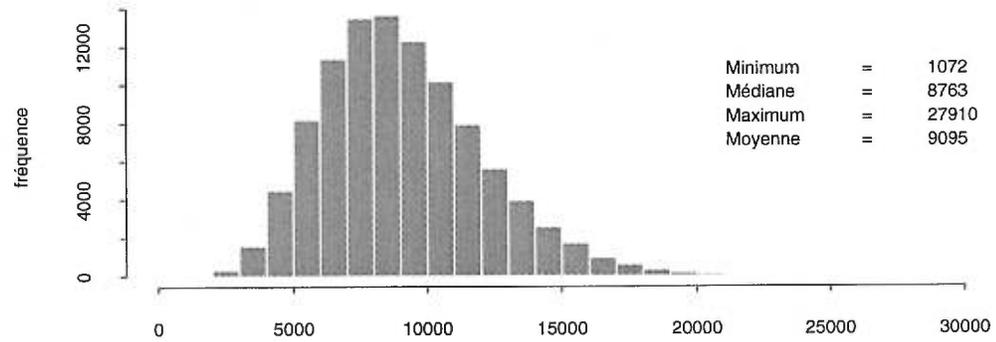
2.2.2. Résultats

Les tableaux 2.2.2 à 2.2.4 résument les résultats obtenus pour les différentes valeurs de t . La figure 2.2.4 présente graphiquement la fonction de renouvellement théorique et un exemple de chaque estimateur (1 échantillon choisi aléatoirement parmi les 1000) pour un échantillon de taille 100 provenant d'une distribution gamma. Notons qu'il a été vérifié que les autres distributions et les autres tailles échantillonnales présentaient des courbes semblables.

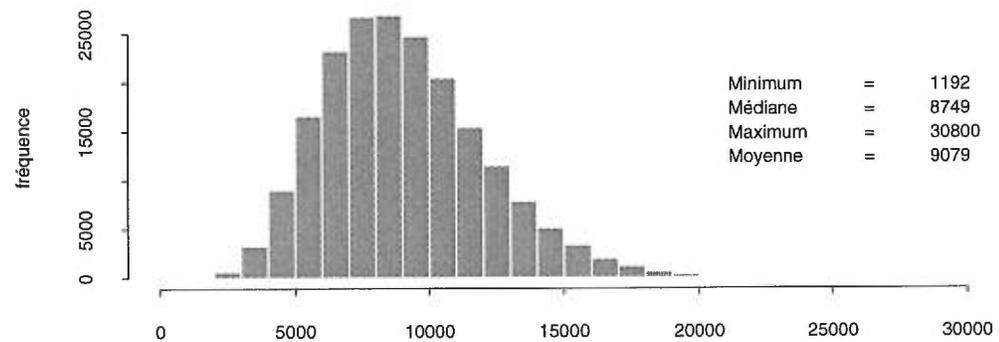
Distribution Gamma



n = 50



n = 100



n = 200

FIG. 2.2.1. *Histogrammes des distributions gamma simulées*

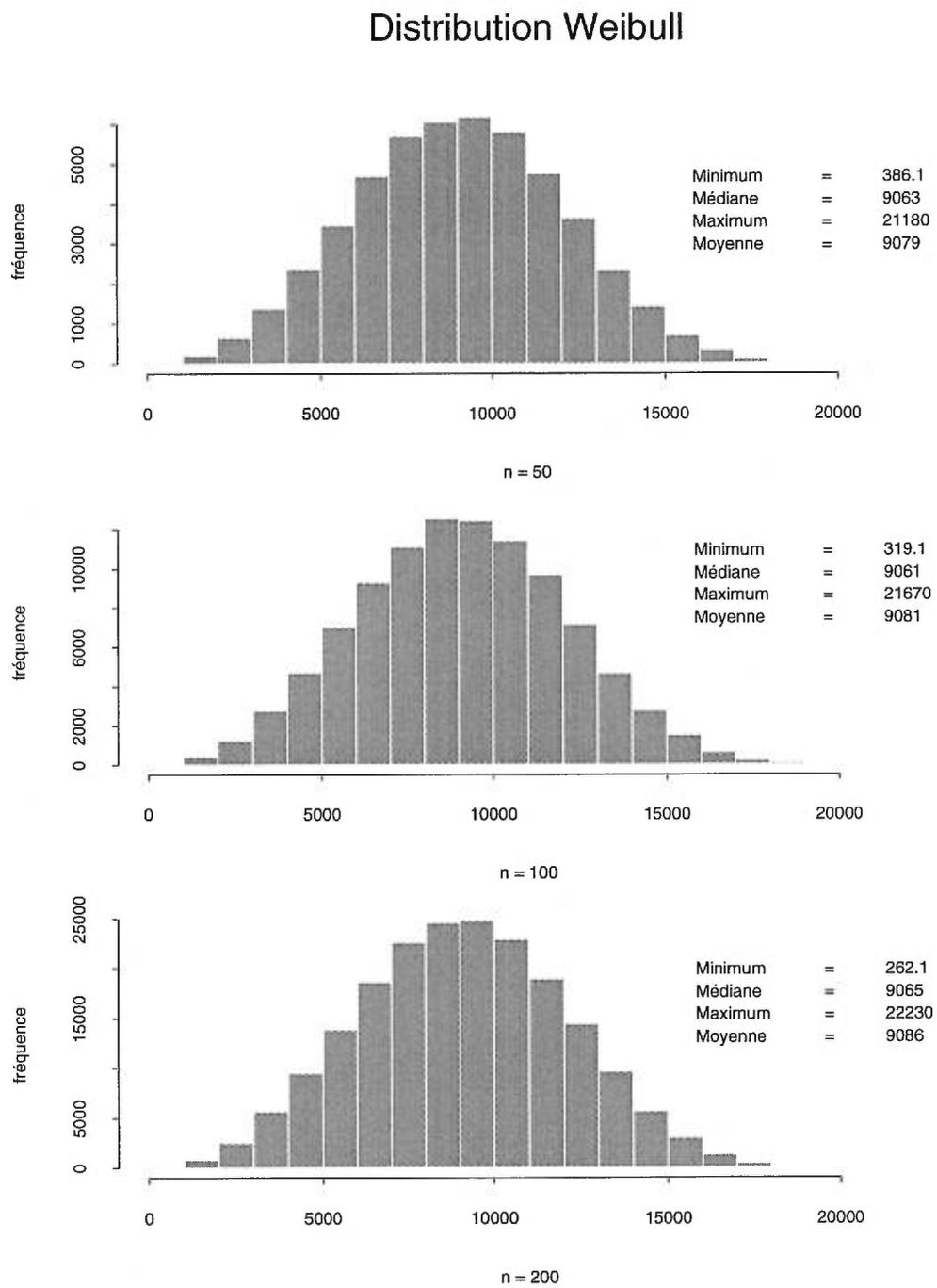


FIG. 2.2.2. *Histogrammes des distributions de Weibull simulées*

Distribution Normale Tronquée

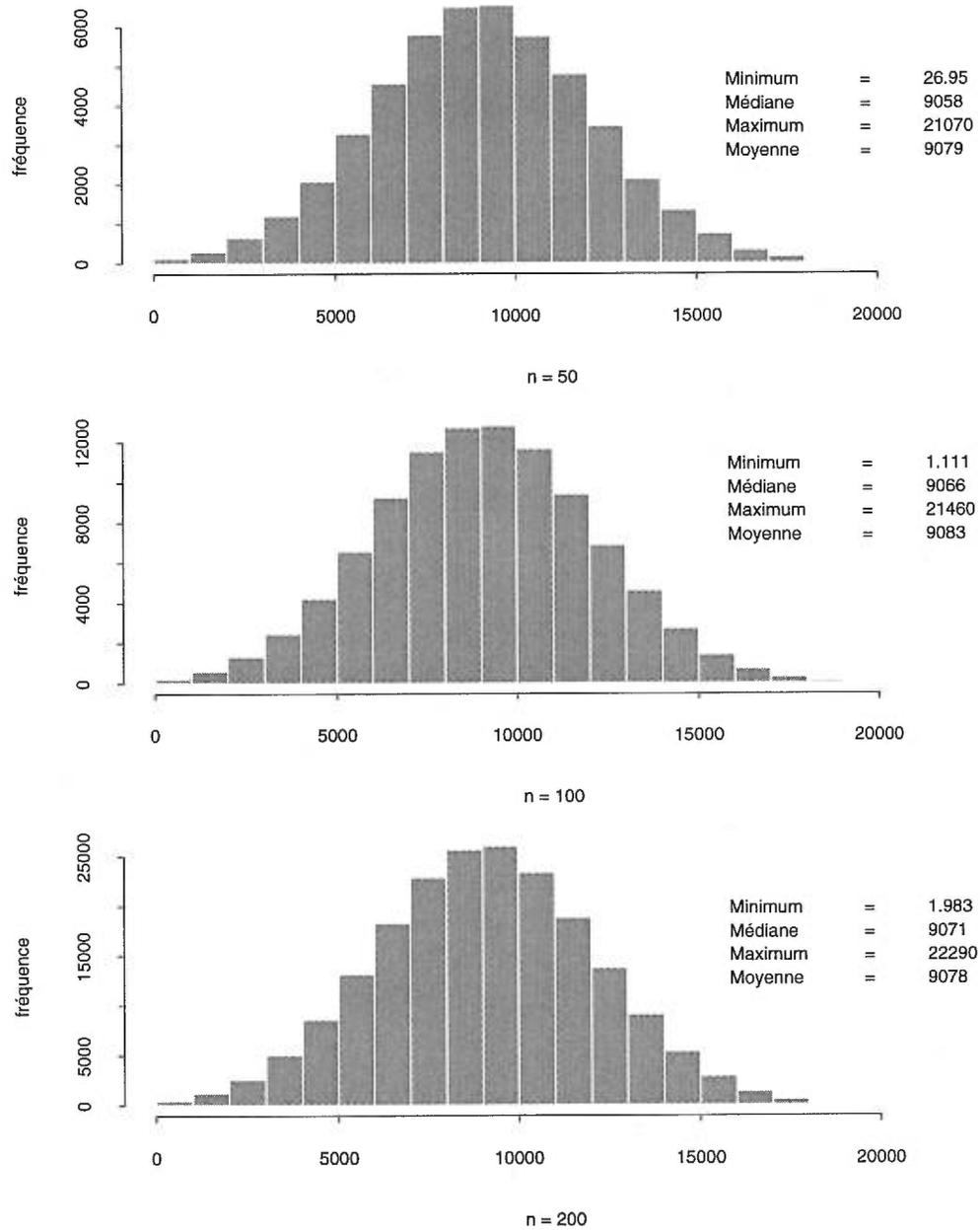


FIG. 2.2.3. *Histogrammes des distributions normales tronquées simulées*

TAB. 2.2.2. Estimation de $M(t)$ pour la loi gamma

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
3000	50	BM	1,1727E-01	8,5845E-04	-3,5884E-03	-6,8451E-05	-2,4756E-03
		EQM	1,4133E-02	1,8037E-05	1,2877E-05	7,1614E-05	1,8176E-05
		ERM	3,26815	0,78654	0,99999	1,66094	1,04991
	100	BM	-1,1795E-01	4,5142E-04	-3,0584E-03	4,1750E-05	-1,7392E-03
		EQM	1,4107E-02	7,8541E-06	1,5573E-05	3,4927E-05	1,8239E-05
		ERM	3,61939	0,56789	1,05169	1,39164	1,02956
	200	BM	-1,17220E-01	3,4792E-04	-2,4734E-03	1,2172E-04	-1,2148E-03
		EQM	1,3844E-02	3,97119E-06	1,3749E-05	1,9252E-05	1,4053E-05
		ERM	3,26661	0,41021	0,96871	0,99591	0,88601
4000	50	BM	-2,3579E-02	1,3315E-03	-1,3427E-02	-7,2619E-04	-6,2062E-03
		EQM	1,1286E-03	1,7519E-04	3,5217E-04	3,6560E-04	3,2050E-04
		ERM	1,37831	0,48930	0,82075	0,70552	0,75756
	100	BM	-2,4596E-02	6,3081E-04	-9,1278E-03	-5,8579E-04	-1,8733E-03
		EQM	8,9732E-04	8,8051E-05	2,4152E-04	2,0367E-04	2,0547E-04
		ERM	1,26790	0,36193	0,65795	0,54518	0,57554
	200	BM	-2,3736E-02	7,4175E-04	-5,0128E-03	-8,1499E-05	5,8138E-04
		EQM	7,1724E-04	4,7616E-05	1,2661E-04	1,0434E-04	1,0560E-04
		ERM	1,17938	0,26356	0,45634	0,39680	0,39970
5000	50	BM	4,1952E-02	6,7793E-04	-2,0131E-02	-6,9419E-04	-1,4640E-03
		EQM	2,5822E-03	7,3195E-04	1,4744E-03	1,1370E-03	1,1654E-03
		ERM	0,66517	0,32653	0,49224	0,41411	0,40759
	100	BM	4,0594E-02	2,5188E-05	-1,0491E-02	-4,5649E-04	2,5274E-03
		EQM	2,0650E-03	3,8223E-04	7,4857E-04	6,4031E-04	6,0314E-04
		ERM	0,62737	0,24003	0,34319	0,30768	0,29748
	200	BM	4,1584E-02	8,9459E-04	-4,8411E-03	2,1392E-04	3,5672E-03
		EQM	1,9473E-03	2,0840E-04	3,4157E-04	3,1825E-04	3,2117E-04
		ERM	0,63791	0,17463	0,22614	0,21628	0,21723

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
7000	50	BM	6,4943E-02	-1,2471E-03	-1,9993E-02	4,4099E-04	1,3568E-02
		EQM	5,7075E-03	2,6789E-03	4,5288E-03	4,1436E-03	4,1269E-03
		ERM	0,25073	0,15671	0,20667	0,19307	0,19019
	100	BM	6,2901E-02	-2,1371E-03	-1,2348E-02	-1,9406E-03	7,4606E-03
		EQM	4,7070E-03	1,3633E-03	2,2112E-03	2,0677E-03	2,1243E-03
		ERM	0,23971	0,11275	0,14573	0,14040	0,13967
	200	BM	6,4151E-02	3,1708E-04	-3,3177E-03	1,9960E-03	7,2186E-03
		EQM	4,5039E-03	7,2079E-04	1,0098E-03	1,0079E-03	1,0151E-03
		ERM	2,4389E-01	8,1482E-02	9,7014E-02	9,7053E-02	9,7197E-02
10000	50	BM	-1,0866E-02	4,4786E-03	-2,5981E-02	1,7742E-03	4,0996E-02
		EQM	3,0312E-03	3,6922E-03	5,7011E-03	5,2171E-03	7,5776E-03
		ERM	6,5528e-02	7,1818E-02	9,0859E-02	8,7018E-02	1,0206E-01
	100	BM	-1,3932E-02	1,5731E-04	-1,6199E-02	-1,1898E-03	2,3580E-02
		EQM	1,6530E-03	1,8371E-03	2,6548E-03	2,4618E-03	3,3003E-03
		ERM	4,8778E-02	5,1404E-02	6,1357E-02	5,9178E-02	6,9169E-02
	200	BM	-1,2293E-02	1,3910E-03	-7,4708E-03	5,1519E-04	1,5584E-02
		EQM	9,0048E-04	9,1993E-04	1,2997E-03	1,2612E-03	1,5273E-03
		ERM	3,5778e-02	3,6362E-02	4,3042E-02	4,2214E-02	4,6450E-02
15000	50	BM	8,7507E-03	2,6748E-03	-4,5904E-02	-6,6116E-03	1,4930E-01
		EQM	6,4863E-03	6,0659E-03	9,4497E-03	5,7643E-03	1,2367E+00
		ERM	5,2690E-02	5,1164E-02	6,6303E-02	5,0607E-02	1,6627E-01
	100	BM	3,9778E-03	-1,2622E-03	-2,9989E-02	-2,2882E-03	6,8178E-02
		EQM	3,2121E-03	3,0052E-03	4,0661E-03	3,2289E-03	1,5245E-02
		ERM	3,7858E-02	3,6592E-02	4,2941E-02	3,7998E-02	6,7138E-02
	200	BM	6,2651E-03	1,3878E-03	-1,3539E-02	1,4063E-03	4,0151E-02
		EQM	1,6700E-03	1,5493E-03	1,7802E-03	1,6357E-03	3,6598E-03
		ERM	2,7398E-02	2,6355E-02	2,8209E-02	2,7290E-02	4,0966E-02

TAB. 2.2.3. Estimation de $M(t)$ pour la loi de Weibull

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
3000	50	BM	-1,3090E-01	1,2393E-03	-1,7257E-02	2,2539E-04	-6,1472E-03
		EQM	1,7634E-02	9,3183E-05	3,1106E-04	3,6058E-04	2,5384E-04
		ERM	7,33026	0,41629	0,97760	0,81661	0,77635
	100	BM	-1,3175E-01	4,1082E-04	-8,7972E-03	-5,9723E-04	-3,1878E-03
		EQM	1,7618E-02	4,5320E-05	2,0091E-04	1,7285E-04	1,6420E-04
		ERM	7,37818	0,29414	0,70849	0,60324	0,59596
	200	BM	-1,3177E-01	3,3336E-04	-5,1369E-03	-3,0326E-04	-4,7470E-04
		EQM	1,7496E-02	2,1657E-05	1,1020E-04	8,9949E-05	8,5186E-05
		ERM	7,37904	0,20825	0,49643	0,43128	0,41247
4000	50	BM	-4,8247E-02	1,5205E-03	-1,8075E-02	2,9323E-04	-4,6874E-03
		EQM	3,0525E-03	3,4936E-04	1,0721E-03	8,8735E-04	7,9906E-04
		ERM	1,08115	0,32444	0,62028	0,52150	0,50078
	100	BM	-4,9240E-02	3,3051E-04	-1,0541E-02	-5,3637E-04	-2,3289E-04
		EQM	2,8022E-03	1,7697E-04	5,5157E-04	4,5404E-04	4,3372E-04
		ERM	1,08496	0,23111	0,43690	0,38049	0,35788
	200	BM	-4,93707E-02	3,6839E-04	-4,7948E-03	2,9994E-04	1,7786E-03
		EQM	2,6310E-03	8,6862E-05	2,3118E-04	2,1012E-04	2,0209E-04
		ERM	1,08317	0,16451	0,26896	0,25379	0,25305
5000	50	BM	1,4644E-02	1,1618E-03	-1,9759E-02	5,9557E-04	1,4459E-03
		EQM	1,2153E-03	8,7614E-04	1,9706E-03	1,6270E-03	1,5677E-03
		ERM	0,29862	0,25486	0,39809	0,35102	0,339801
	100	BM	1,3513E-02	-1,7960E-04	-1,0858E-02	-4,5276E-04	3,4845E-03
		EQM	7,0576E-04	4,5389E-04	9,7952E-04	8,7105E-04	8,2778E-04
		ERM	0,22652	0,18263	0,27226	0,25171	0,24272
	200	BM	1,3269E-02	7,1530E-05	-4,6468E-03	6,6350E-04	3,6042E-03
		EQM	4,4489E-04	2,2672E-04	4,7909E-04	4,6119E-04	4,1540E-04
		ERM	0,18269	0,13043	0,19061	0,18508	0,17340

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
7000	50	BM	6,9594E-02	-2,5374E-03	-2,1709E-02	8,4597E-04	1,5172E-02
		EQM	6,5507E-03	2,4734E-03	4,1200E-03	3,7667E-03	3,8861E-03
		ERM	0,27262	0,15456	0,19964	0,18852	0,1904
	100	BM	6,8189E-02	-3,5030E-03	-1,1682E-02	1,8047E-04	8,4260E-03
		EQM	5,5451E-03	1,3028E-03	2,1558E-03	2,0625E-03	1,9220E-03
		ERM	0,26353	0,11148	0,14472	0,14044	0,13535
	200	BM	6,7715E-02	-2,9906E-03	-6,2186E-03	7,6696E-06	3,6012E-03
		EQM	5,0462E-03	6,7043E-04	1,0336E-03	1,0113E-03	9,2049E-04
		ERM	2,6158E-01	7,9719E-02	9,8950E-02	9,7492E-02	9,2725E-02
10000	50	BM	1,5808E-02	3,3615E-03	-3,3245E-02	-7,3460E-04	4,6951E-02
		EQM	3,4020E-03	4,4099E-03	6,3453E-03	5,6050E-03	9,0524E-03
		ERM	7,2549E-02	8,3213E-02	9,9617E-02	9,3233E-02	1,1698E-01
	100	BM	1,3990E-02	1,2853E-03	-1,7605E-02	-3,5409E-05	2,7814E-02
		EQM	1,8523E-03	2,2776E-03	3,1227E-03	2,9170E-03	3,7721E-03
		ERM	5,2834E-02	5,9017E-02	6,9981E-02	6,7081E-02	7,5685E-02
	200	BM	1,3173E-02	-6,4335E-05	-1,0072E-02	-6,5036E-04	1,5968E-02
		EQM	1,0262E-03	1,1712E-03	1,5613E-03	1,4938E-03	1,7961E-03
		ERM	3,9250E-02	4,1884E-02	4,8851E-02	4,7620E-02	5,1689E-02
15000	50	BM	-6,4170E-03	2,7938E-03	-4,0615E-02	-6,1771E-03	3,2462E-01
		EQM	6,6286E-03	5,4725E-03	8,6231E-03	5,5360E-03	9,5957E+00
		ERM	5,3981E-02	4,8846E-02	6,2016E-02	4,9129E-02	3,4684E-01
	100	BM	-8,9225E-03	8,2507E-04	-2,8943E-02	-2,7877E-03	8,4443E-02
		EQM	3,5460E-03	2,8888E-03	4,0832E-03	3,0241E-03	1,7996E-02
		ERM	3,9134E-02	3,4932E-02	4,2724E-02	3,5931E-02	7,5003E-02
	200	BM	-1,0311E-02	-2,4675E-05	-1,6350E-02	-7,0146E-05	4,5867E-02
		EQM	1,8895E-03	1,4799E-03	1,8102E-03	1,5896E-03	4,6459E-03
		ERM	2,8324E-02	2,4898E-02	2,8004E-02	2,5776E-02	4,2913E-02

TAB. 2.2.4. Estimation de $M(t)$ pour la loi normale tronquée

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
3000	50	BM	-1,3418E-01	8,5283E-04	-1,9287E-02	1,5796E-03	-3,3520E-03
		EQM	1,8548E-02	1,2176E-04	4,0917E-04	4,6281E-04	3,8622E-04
		ERM	6,39984	0,42276	0,93440	0,77989	0,79225
	100	BM	-1,3440E-01	6,3897E-04	-7,7256E-03	1,4372E-03	-3,1351E-04
		EQM	1,8329E-02	6,2133E-05	2,3680E-04	2,1727E-04	2,0010E-04
		ERM	3,74545	0,29922	0,62228	0,53559	0,54716
	200	BM	-1,3447E-01	4,4553E-04	-3,8656E-03	1,1556E-03	1,6441E-03
		EQM	1,8229E-02	3,4586E-05	1,3120E-04	1,2215E-04	1,1418E-04
		ERM	6,41333	0,22482	0,45387	0,42375	0,40197
4000	50	BM	-4,8204E-02	7,0400E-04	-1,6661E-02	1,7222E-03	-2,5534E-04
		EQM	3,1051E-03	3,6178E-04	1,0602E-03	9,6826E-04	1,0038E-03
		ERM	1,09020	0,34008	0,62469	0,54963	0,55909
	100	BM	-4,8579E-02	6,7070E-04	-8,2285E-03	2,0085E-03	4,3312E-03
		EQM	2,7444E-03	1,8351E-04	4,8099E-04	4,3053E-04	4,3417E-04
		ERM	2,40055	0,239500	0,39832	0,37153	0,35929
	200	BM	-4,8642E-02	5,3650E-04	-2,8845E-03	2,3368E-03	4,8208E-03
		EQM	2,5784E-03	1,0313E-04	2,4091E-04	2,4148E-04	2,5370E-04
		ERM	1,07237	0,18052	0,27334	0,26581	0,27660
5000	50	BM	1,9916E-02	1,3419E-04	-1,9052E-02	1,6502E-03	6,2503E-03
		EQM	1,4698E-03	8,3635E-04	1,9391E-03	1,6679E-03	1,8109E-03
		ERM	0,34638	0,26802	0,41644	0,37607	0,38214
	100	BM	1,9382E-02	4,7516E-04	-8,2413E-03	2,4393E-03	9,2160E-03
		EQM	9,0594E-04	4,1778E-04	8,7732E-04	8,3611E-04	8,7249E-04
		ERM	0,37105	0,18733	0,27627	0,26255	0,26262
	200	BM	1,9322E-02	5,1867E-04	-1,7618E-03	3,7377E-03	7,4446E-03
		EQM	6,6516E-04	2,3521E-04	4,3768E-04	4,5604E-04	4,5570E-04
		ERM	0,24080	0,14112	0,19002	0,19099	0,19076

t	n		$M_{1n}(t)$	$M_{2n}(t)$	$M_{3n}(t)$	$M_{4n}(t)$	$M_{5n}(t)$
7000	50	BM	7,9648E-02	-4,8401E-03	-2,2122E-02	1,3703E-03	2,0434E-02
		EQM	8,1580E-03	2,3478E-03	4,3663E-03	4,0636E-03	4,4198E-03
		ERM	0,32275	0,15631	0,21571	0,20562	0,21103
	100	BM	7,8797E-02	-3,8984E-03	-8,8188E-03	3,6344E-03	1,5096E-02
		EQM	7,1116E-03	1,1439E-03	2,0527E-03	2,0454E-03	2,0935E-03
		ERM	0,29966	0,10844	0,14851	0,14698	0,14653
	200	BM	7,8743E-02	-3,3898E-03	-3,2278E-03	3,4213E-03	7,7934E-03
		EQM	6,6941E-03	6,3603E-04	9,8931E-04	1,0133E-03	9,6808E-04
		ERM	3,1670E-01	8,0715E-02	1,0036E-01	1,0165E-01	9,8525E-02
10000	50	BM	1,3746E-02	2,5125E-03	-2,7977E-02	6,0921E-03	5,9604E-02
		EQM	3,5086E-03	4,2684E-03	6,1760E-03	5,9407E-03	1,1041E-02
		ERM	7,2486E-02	7,9902E-02	9,6650E-02	9,5110E-02	1,2899E-01
	100	BM	1,2417E-02	8,7046E-04	-1,6057E-02	2,8058E-03	3,6790E-02
		EQM	1,8175E-03	2,1741E-03	3,1235E-03	3,0076E-03	4,7091E-03
		ERM	5,0272E-02	5,7670E-02	6,7862E-02	6,7495E-02	8,5084E-02
	200	BM	1,2373E-02	9,6646E-04	-6,3477E-03	3,8486E-03	2,1860E-02
		EQM	1,0569E-03	1,1489E-03	1,5950E-03	1,6138E-03	2,0906E-03
		ERM	3,8945E-02	4,0918E-02	4,9544E-02	4,9310E-02	5,5993E-02
15000	50	BM	-3,1577E-04	1,5456E-03	-4,3107E-02	-7,9374E-03	2,5254E-01
		EQM	6,8802E-03	5,5249E-03	9,2840E-03	6,5196E-03	3,5392E+00
		ERM	5,4271E-02	4,8646E-02	6,4927E-02	5,3690E-02	2,6027E-01
	100	BM	-2,4386E-03	5,3940E-04	-2,6163E-02	1,6629E-03	1,0538E-01
		EQM	3,4758E-03	2,7816E-03	3,8960E-03	2,9525E-03	8,6329E-02
		ERM	3,8947E-02	3,4509E-02	4,1611E-02	3,5577E-02	9,3878E-02
	200	BM	-2,4670E-03	9,8972E-04	-1,2767E-02	4,2918E-03	5,3128E-02
		EQM	1,8799E-03	1,5149E-03	1,7960E-03	1,7080E-03	5,1105E-03
		ERM	2,8056E-02	2,5067E-02	2,8013E-02	2,6638E-02	4,7465E-02

La figure 2.2.4 permet de comparer graphiquement les cinq estimateurs avec la fonction de renouvellement théorique. On remarque sur ce graphique la bonne performance de $M_{1n}(t)$ pour t grand, alors que pour t petit, l'approximation est grossière et les valeurs peuvent même être négatives (puisque l'estimateur est donné par une droite). On constate aussi que les troisième et quatrième estimateurs sont très semblables. Par ailleurs, ces derniers oscillent davantage que $M_{5n}(t)$ (et que $M_{1n}(t)$ évidemment).

2.2.3. Analyse et conclusion

Après une première observation des résultats obtenus, il est clair que le meilleur estimateur est l'estimateur paramétrique ($M_{2n}(t)$). En effet, il est, dans tous les cas, supérieur aux autres (autant par rapport au biais, à l'erreur quadratique moyenne ou à l'erreur relative). Cette bonne performance est attribuable à l'hypothèse de départ formulée à propos de la distribution. Evidemment, puisque les temps ont été simulés selon des distributions précises et connues, ces hypothèses étaient justes. L'estimateur paramétrique semble donc se distinguer positivement des autres, lorsqu'on est en mesure de connaître la distribution d'origine des observations. Dans le cas où ceci est impossible, $M_{2n}(t)$ risque de fournir des résultats erronés.

Par la suite, la compétition entre les quatre autres estimateurs est plus corsée. Malgré ce que l'on pourrait croire, l'hypothèse de départ concernant la croissance du taux de panne qui caractérise l'estimateur $M_{5n}(t)$ ne semble pas procurer

Fonction de renouvellement et ses estimateurs

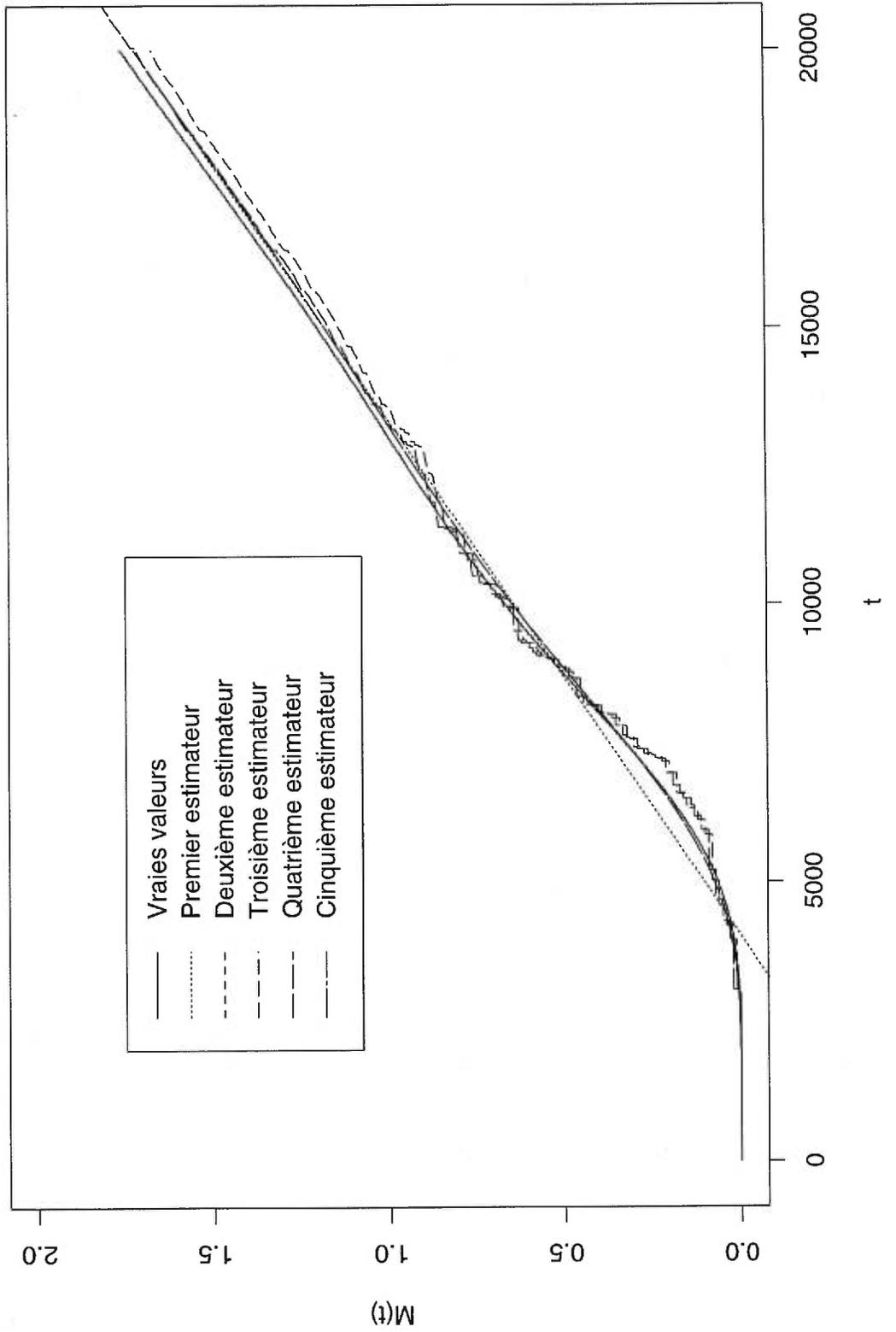


FIG. 2.2.4. *Fonction de renouvellement et ses estimateurs*

des résultats plus précis. Contrairement à l'estimateur paramétrique, l'estimateur pour le cas IFR ($M_{5n}(t)$) ne se démarque pas des autres. On remarque même que pour de grandes valeurs de t (10 000 et 15 000), cet estimateur est le moins performant. Il est même battu par l'estimateur asymptotique, plutôt trivial. Cependant pour de plus petites valeurs de t , $M_{5n}(t)$ se comporte un peu mieux. Il est le meilleur lorsque la distribution est de Weibull et se comporte assez bien pour la distribution gamma.

Soulignons que la contre-performance de $M_{5n}(t)$ pour de grandes valeurs de t (principalement pour $n = 50$) est due à un problème qui survient lors de l'estimation du taux de panne. En utilisant l'équation (1.3.12) pour l'estimer, on le pose égal à l'infini pour des temps plus grands que la dernière statistique d'ordre. Dans notre cas, il arrive quelquefois, particulièrement pour les plus petits échantillons, que la valeur $t = 15000$ soit supérieure au plus grand temps simulé. Il semble donc que l'estimation du taux de panne par l'infini ne soit pas l'idéal. Le calcul des fonctions de densité, de répartition et ultimement de renouvellement sera aussi affectée, d'où les moins bons résultats obtenus pour $t = 15000$.

Un autre résultat quelque peu surprenant est l'excellent rendement de l'estimateur empirique $M_{4n}(t)$. Pour 48 des 54 simulations effectuées, $M_{4n}(t)$ surpasse l'estimateur de Frees ($M_{3n}(t)$). Les erreurs quadratiques moyennes et erreurs relatives moyennes des deux estimateurs se ressemblent, mais celles de $M_{4n}(t)$ sont presque toujours légèrement inférieures. Le biais de $M_{3n}(t)$ est toutefois nettement supérieur (en valeur absolue) dans la plupart des cas. Ceci est un peu inattendu

et contredit la littérature à ce sujet. En effet, rappelons que $M_{4n}(t)$ a été introduit dans le but de réduire le temps de calcul de $M_{3n}(t)$, en acceptant une légère augmentation du biais. C'est l'inverse qui s'est produit dans notre cas.

Comme Frees le notait aussi en réalisant ses simulations, l'estimateur asymptotique ($M_{1n}(t)$) se comporte plutôt bien dans certaines situations. Il est cependant inadéquat pour de petites valeurs de t . Ceci était attendu, puisqu'il est basé sur un résultat limite valable pour de grandes valeurs de t . Il commence toutefois à mieux se comporter à partir de $t = 5000$ (bien que $\mu = 9080$). De façon générale, $M_{1n}(t)$ demeure moins bon que ses compétiteurs, mais se situe tout de même très souvent dans le peloton. Tel que prévu donc, pour de grandes valeurs de t , l'estimateur asymptotique est assez fiable. Il est même le meilleur estimateur de $M(t)$ pour $t = 10000$. En $t = 15000$, il constitue souvent le deuxième estimateur le plus performant. On peut cependant souligner que pour $n = 200$, $M_{1n}(t)$ se comporte un peu moins bien par rapport aux autres. un peu En fait, la progression de cet estimateur lorsque n augmente semble moins marquée que dans le cas des quatre autres estimateurs. Ceci est dû au fait que, comme il est mentionné au chapitre 2, l'estimateur asymptotique est généralement non convergent.

Ceci nous amène à analyser l'impact qu'a la taille de l'échantillon sur les résultats. Sans surprise, on note que le fait de prendre des échantillons plus grands améliore sensiblement la qualité des estimations réalisées. De plus, la plupart du temps, l'ordre des estimateurs selon leur rendement reste inchangé. Bref, le seul aspect intéressant à soulever concernant la taille des échantillons est la non convergence de $M_{1n}(t)$, qui implique moins d'amélioration de cet estimateur en fonction de l'augmentation de n .

Prenons quelques instants pour confirmer ou préciser ces observations en ayant recours à l'analyse de variance. Trois facteurs seront pris en considération: la taille échantillonnale, le temps et l'estimateur. Seules les interactions d'ordre 2 ont été considérées, l'interaction d'ordre 3 étant assimilée à l'erreur. Pour cette analyse, nous avons laissé tombé l'estimateur asymptotique en raison de son comportement particulier (mauvaise performance pour t petit et non convergent). Il ressort de cette anova une grande similitude entre chacune des trois distributions (pour un niveau de confiance de $\alpha = 0.05$). Nous présenterons donc ici les résultats pour la distribution gamma et indiquerons les particularités de chacune s'il y a lieu. La table d'anova pour la gamma est résumée dans le tableau 2.2.5 et les résultats complets pour la Weibull et la normale tronquée sont à l'annexe A.

TAB. 2.2.5. Table d'anova pour la loi Gamma ($\alpha = 0.05$)

		Biais moy.		Eq. quad. moy.		Err. rel. moy.	
Facteur	d. lib.	Stat. F	Pr. > F	Stat. F	Pr. > F	Stat. F	Pr. > F
n	2	0,23	0,7987	1,08	0,3532	48,84	0,0001
t	5	2,92	0,0289	1,09	0,3848	389,02	0,0001
estimateur	3	21,34	0,0001	1,04	0,3878	26,02	0,0001
n*t	10	0,52	0,8612	1,02	0,4491	3,91	0,0018
n*estimateur	6	2,20	0,0702	1,00	0,4408	0,54	0,7748
t*estimateur	15	6,76	0,0001	1,03	0,4512	11,70	0,0001

TAB. 2.2.6. *Contrastes entre les estimateurs ($\alpha = 0.05$) (distribution gamma)*

Biais moyen				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{3n}(t)$	0.001834	0.014368	0.026901	Oui
$M_{2n}(t) - M_{4n}(t)$	-0.011431	0.001103	0.013637	Non
$M_{2n}(t) - M_{5n}(t)$	-0.031757	-0.019224	-0.006690	Oui
$M_{3n}(t) - M_{4n}(t)$	-0.025799	-0.013265	-0.000731	Oui
$M_{3n}(t) - M_{5n}(t)$	-0.046125	-0.033591	-0.021057	Oui
$M_{4n}(t) - M_{5n}(t)$	-0.032860	-0.020326	-0.007793	Oui
Erreur relative moyenne				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{3n}(t)$	-0.20450	-0.14188	-0.07927	Oui
$M_{2n}(t) - M_{4n}(t)$	-0.23615	-0.17353	-0.11092	Oui
$M_{2n}(t) - M_{5n}(t)$	-0.19008	-0.12747	-0.06485	Oui
$M_{3n}(t) - M_{4n}(t)$	-0.09427	-0.03165	0.03096	Non
$M_{3n}(t) - M_{5n}(t)$	-0.04820	0.01441	0.07703	Non
$M_{4n}(t) - M_{5n}(t)$	-0.01655	0.04607	0.10868	Non

Dans un premier temps, soulignons que nous n'avons décelé aucune différence significative au niveau des différents écarts quadratiques moyens. Aucun des trois facteurs, en particulier le facteur estimateur n'a d'effet sur ces derniers. Aussi, il n'y a pas d'interaction entre les trois facteurs.

Concernant les biais moyens, on note que les facteurs temps et estimateur ont un effet significatif (seulement le facteur estimateur pour la Weibull). Il y a

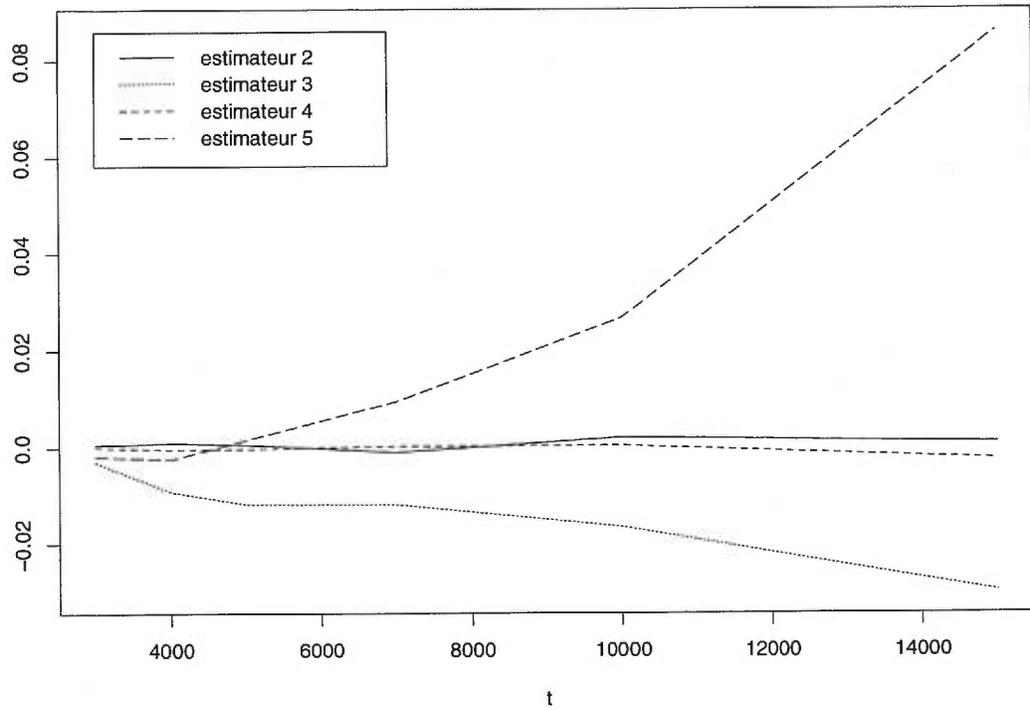
aussi présence d'interaction entre ces deux facteurs. Une comparaison des estimateurs deux à deux à l'aide d'un test de Scheffé montre bien la différence entre chacun d'eux. Comme le confirme le tableau 2.2.6, on obtient que, en moyenne, le biais de $M_{3n}(t)$ est plus petit que ceux de $M_{2n}(t)$ et de $M_{4n}(t)$, qui ne sont pas significativement différents. Le biais de $M_{5n}(t)$ est le plus grand de tous. La figure 2.2.5, qui illustre l'interaction des estimateurs avec le temps confirme ces différences et apporte un peu plus de précision. Par exemple, on remarque que le biais de $M_{3n}(t)$ est toujours négatif. Les deuxième et quatrième s'avèrent donc être, en moyenne, les meilleurs estimateurs en ce qui concerne le biais moyen. Mentionons que le biais de $M_{3n}(t)$ n'est pas significativement différent de ceux de $M_{2n}(t)$ et de $M_{4n}(t)$ pour les deux autres distributions. Cependant, les graphiques représentant l'interaction sont à toute fin pratique identiques à la figure 2.2.5 (voir annexe A). Notons que ces résultats sont en accord avec les observations soulevées précédemment dans cette section, notamment la meilleure performance de $M_{4n}(t)$ par rapport à $M_{3n}(t)$ et le moins bon rendement de $M_{5n}(t)$, en particulier pour t grand.

Par ailleurs, les trois facteurs considérés semblent avoir un effet sur l'erreur relative moyenne. C'est ce qui découle de l'analyse de variance effectuée sur cette statistique. Il y a aussi présence d'interaction entre la taille n et le temps t et entre t et les estimateurs. Nous n'insisterons pas sur l'interaction entre n et t , puisque le but de ce mémoire est de comparer chaque estimateur. L'interaction entre t et les estimateurs est aussi illustrée à la figure 2.2.5. On remarque la meilleure performance de $M_{2n}(t)$ pour tout t , qui est confirmée par les contrastes disponibles au tableau 2.2.6. Aussi, comme prévu, $M_{5n}(t)$ performe un peu moins bien pour de grandes valeurs de t .

En résumé, on constate que plusieurs faits observés en consultant les tableaux des résultats ont été confirmés statistiquement par l'analyse de variance. Ceci ne signifie pas que ceux qui ne l'ont pas été n'existent pas. Les différences observées précédemment qui n'ont pas été appuyées par l'anova existent probablement, mais sont plus modérées.

Dans un autre ordre d'idée, discutons maintenant des différents temps nécessaires au calcul de chaque estimateur. Soulignons que les simulations ont été réalisées à partir d'un *Sun Ultra60* à 768 megs de mémoire vive et à 2 CPUS. L'estimateur le plus rapide à obtenir est, sans surprise, l'estimateur asymptotique $M_{1n}(t)$. Par ailleurs, contrairement à nos attentes, nous n'avons pas noté une nette différence entre le temps de calcul des estimateurs de Frees et empirique. La puissance de l'ordinateur employé y est sûrement pour quelque chose. Pour $n = 50$, $M_{3n}(t)$ à même été obtenu plus rapidement que $M_{4n}(t)$. Ce n'est que pour $n = 200$ que $M_{4n}(t)$ a véritablement représenté un avantage. En fait, le temps nécessaire pour obtenir $M_{3n}(t)$ augmente considérablement avec l'augmentation de n et augmente proportionnellement avec le nombre de points pour lesquels son évaluation est désirée. Si la taille échantillonnale est raisonnable et que l'on souhaite estimer $M(t)$ pour un petit nombre de valeurs de t , l'estimateur $M_{3n}(t)$ est plus efficace, par rapport au temps de calcul, que $M_{4n}(t)$. En raison de l'amélioration de l'équipement informatique disponible, l'avantage que possédait $M_{4n}(t)$ (et la raison pour laquelle il a été développé) est moins significatif. La bonne performance de $M_{4n}(t)$ lors de nos simulations nous incite cependant à le conserver dans notre étude. Mentionnons finalement que $M_{5n}(t)$ s'obtient moins

Gamma – Biais



Gamma – erreur relative

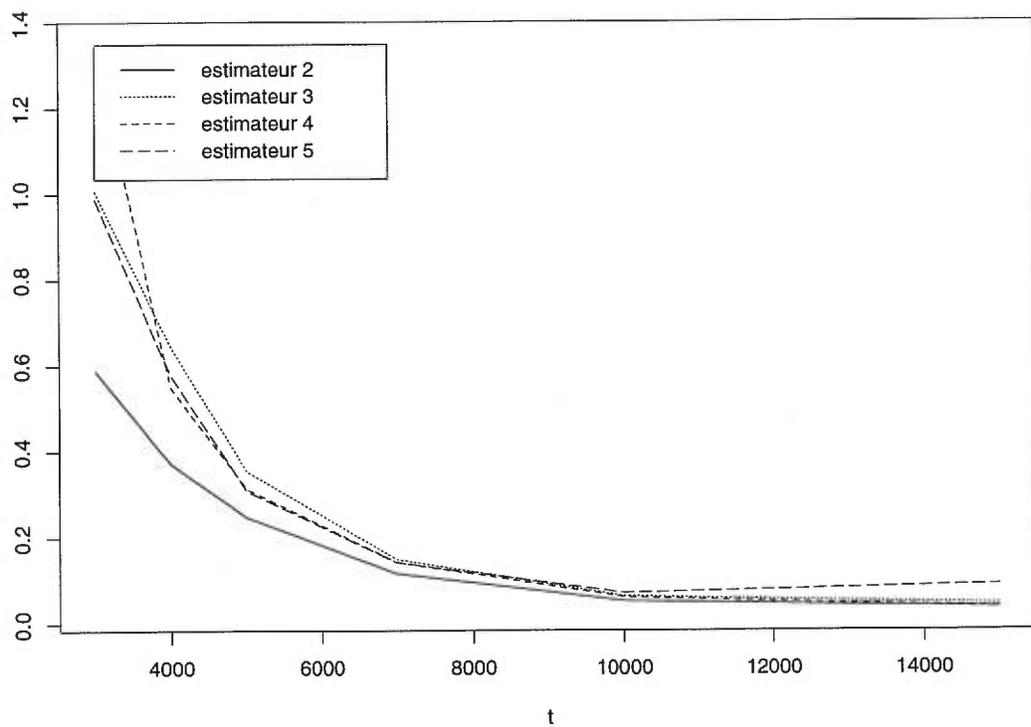


FIG. 2.2.5. Interaction entre le temps et les estimateurs (distribution gamma)

rapidement que les autres, puisqu'il nécessite, dans un premier temps, l'estimation du taux de panne et qu'il implique l'utilisation d'algorithmes, tels CONVOL, pour calculer des convolutions.

En résumé, l'estimateur paramétrique s'est avéré, comme prévu, le plus performant lors de nos simulations. Puisque les distributions desquelles provenaient les données étaient connues, ceci était prévisible. Un résultat plus inattendu est la performance plutôt ordinaire de $M_{5n}(t)$, qui implique pourtant des hypothèses de départ relativement fortes. Aussi, l'estimateur de Frees ($M_{3n}(t)$) a souvent été battu par l'estimateur empirique ($M_{4n}(t)$), ce qui n'était pas prévu et qui est quelque peu différent des simulations effectuées par Lin, O'Conneide et Schneider [13]. Notons que nous avons effectué un plus grand nombre de simulations que ces derniers et que notre taille échantillonnale était plus grande. Mentionnons cependant que les différences ne sont toutefois pas très grandes dans les deux cas. De plus, l'estimateur asymptotique ($M_{1n}(t)$) ne s'est pas fait surclasser par les autres estimateurs, un fait qui avait été soulevé par plusieurs auteurs et qui les avait quelque peu surpris.

Chapitre 3

POLITIQUES DE REMPLACEMENT PRÉVENTIF

Les notions développées aux chapitres précédents vont maintenant nous permettre de traiter d'une des applications principales de la fonction de renouvellement: les politiques de remplacement préventif périodique.

En général, les coûts d'opération d'un appareil croissent à mesure que son âge augmente. Lorsqu'une pièce prend de l'âge, il est nécessaire d'effectuer des réparations plus fréquemment et ces dernières sont souvent plus coûteuses. Aussi, une pièce en moins bon état risque davantage de tomber en panne, ce qui occasionne des frais parfois élevés en raison d'une certaine période d'inactivité. Il est alors avantageux et surtout économiquement rentable d'instaurer des politiques de remplacement préventif. Ceci consiste à remplacer une pièce par une neuve, et ce, même si elle est encore en état de fonctionner. Au cours des années, de nombreuses stratégies ont été développées, afin d'effectuer de tels remplacements. Elles ont pour but la planification de méthodes d'entretien ou de réparations qui minimiseront le coût global d'utilisation et de maintenance du système. Ces programmes de maintenance permettent une meilleure planification à long terme des coûts liés à ce système et une meilleure gestion des ressources humaines ou

matérielles dont dispose l'entreprise. Dans les lignes qui suivent, nous énonçons quelques-unes des stratégies préventives les plus fréquemment utilisées.

Il existe deux grandes catégories de remplacement préventif. Les remplacements de type âge et ceux de type périodique. Le premier type est donc un remplacement basé sur l'âge de la pièce. Il s'agit de remplacer la pièce lorsqu'elle atteint un certain âge T (ou lorsqu'elle tombe en panne). Cet âge T est déterminé de façon optimale, afin de minimiser le coût d'utilisation et d'entretien du système. Le moment où les remplacements ont lieu sont donc dépendants de celui où les précédents ont été effectués. De telles politiques de remplacement préventif ont été fréquemment considérées dans la littérature à ce sujet. Autant l'approche paramétrique que non paramétrique ont été employées pour l'estimation des coûts minimaux d'entretien et de maintenance d'un système. C'est pour cette raison que nous n'insisterons pas davantage sur cette catégorie de remplacement préventif. Arunkumar [1], Scheaffer [17] ou Léger et Cléroux [12] sont de bonnes références à ce propos.

La deuxième catégorie de stratégies préventives, les remplacements périodiques, consiste à remplacer la pièce aux instants $T, 2T, 3T, \dots$, et ce peu importe son âge. Cette stratégie a été moins étudiée par le passé et l'approche non paramétrique pour estimer le coût minimal n'a pas vraiment été prise en considération. Il existe plusieurs variantes au remplacement de type périodique. Entre autres, on peut choisir de remplacer ou non une pièce qui tombe en panne entre deux remplacements systématiques. Si l'instant où survient la panne est près d'un temps de remplacement, il peut être plus avantageux de ne pas la remplacer (les coûts de remplacement peuvent être plus élevés que les frais causés par la non

disponibilité de la pièce). On peut aussi choisir de réparer la pièce défectueuse si son coût de réparation n'est pas trop élevé (réparation minimale à la panne) ou on peut la remplacer par une pièce usagée, mais fonctionnelle. Encore une fois, les temps de remplacement sont obtenus de façon optimale, tout comme les bornes pour les coûts de réparation ou de remplacement qui déterminent si on répare ou si on remplace une pièce, lorsqu'elle tombe en panne entre deux remplacements périodiques.

Les politiques de remplacement préventif décrites ci-haut sont cependant peu réalistes, étant plutôt simplistes. En pratique, on doit tenir compte de nombreux autres facteurs, tels l'augmentation du prix d'une composante du système due à l'inflation, la dépréciation du système ou de nouvelles techniques d'entretien ou de réparation. Les modèles mathématiques sous-jacents à de telles stratégies plus complexes le seront eux aussi. Des algorithmes puissants seront nécessaires. Pour notre part, nous nous contenterons des modèles plus simples, puisqu'il sont suffisants pour étudier le comportement général de la fonction de coût et surtout la qualité de ses estimateurs. A nouveau, les hypothèses posées à la section 1.1 s'appliqueront.

Le principal intérêt de l'étude de la fonction de coût se situe au niveau de l'estimation de son minimum et de l'instant où il est atteint. Ce sont ces derniers qui déterminent les politiques optimales de remplacement préventif. Ces politiques s'obtiennent, comme nous le verrons plus loin, en minimisant le coût moyen par unité de temps sur un horizon infini. Evidemment, ces stratégies préventives optimales diffèrent d'une situation à l'autre. Dans ce qui suit, nous considérerons

le cas du remplacement préventif de type périodique et de type périodique avec réparation minimale à la panne.

3.1. REMPLACEMENT PRÉVENTIF PÉRIODIQUE

Comme mentionné ci-haut, une politique de remplacement préventif est dite périodique, lorsque des remplacements des composantes d'un système sont effectués systématiquement à la fin d'une période d'une durée prédéfinie. Ces remplacements surviennent, même si une pièce est encore en bon état et ils sont indépendants de l'âge de celle-ci. Lorsqu'une panne survient, la pièce défectueuse est immédiatement remplacée par une neuve. Nous supposons à nouveau que le temps d'installation de la nouvelle pièce est négligeable et que cette installation est effectuée parfaitement.

On peut représenter schématiquement un remplacement préventif de type périodique à l'aide de la figure ci-dessous.

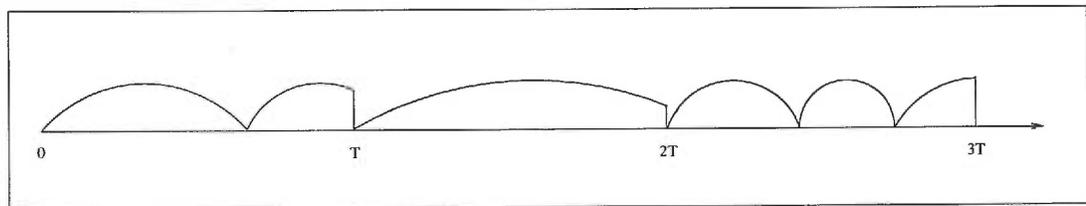


FIG. 3.1.1. Schéma d'un remplacement préventif périodique

Afin de déterminer la politique optimale, c'est-à-dire le moment le plus propice (plus économique) pour effectuer les remplacements préventifs, nous devons

considérer la fonction de coût. Pour ce faire, nous devons utiliser les constantes et fonctions suivantes

c_1 = coût d'un remplacement à la panne

c_2 = coût d'un remplacement préventif

$N_1(t)$ = nombre de pannes dans $[0, t]$

$N_2(t)$ = nombre de remplacements préventifs dans $[0, t]$

Soulignons qu'il est mathématiquement nécessaire et certainement plus logique que le coût d'un remplacement à la panne soit supérieur au coût d'un remplacement préventif ($c_1 > c_2$).

De plus, posons:

$$\begin{aligned} C(t) &= \text{coût total dans } [0, t] \\ &= c_1 N_1(t) + c_2 N_2(t) \end{aligned}$$

La fonction de coût que nous utiliserons est en fait le coût moyen par unité de temps sur un horizon infini et elle est donnée par:

$$\begin{aligned} B(T) &= \lim_{t \rightarrow \infty} \frac{E[C(t)]}{t} \\ &= c_1 \lim_{t \rightarrow \infty} \frac{E[N_1(t)]}{t} + c_2 \lim_{t \rightarrow \infty} \frac{E[N_2(t)]}{t} \\ &= c_1 \frac{M(t)}{T} + c_2 \frac{1}{T} \end{aligned} \tag{3.1.1}$$

On minimise donc la fonction de coût en minimisant l'équation (3.1.1).

Prenons quelques instants pour présenter les principales propriétés de la fonction de coût d'une stratégie de remplacement préventif périodique. L'allure générale de $B(t)$ est illustrée par la figure 3.1.2 ci-dessous. Pour l'obtenir, la distribution de Weibull avec paramètres $\alpha = 5$ et $\lambda = 1$ a été utilisée. La constante c_1 a été fixée à 5 et c_2 à 1.

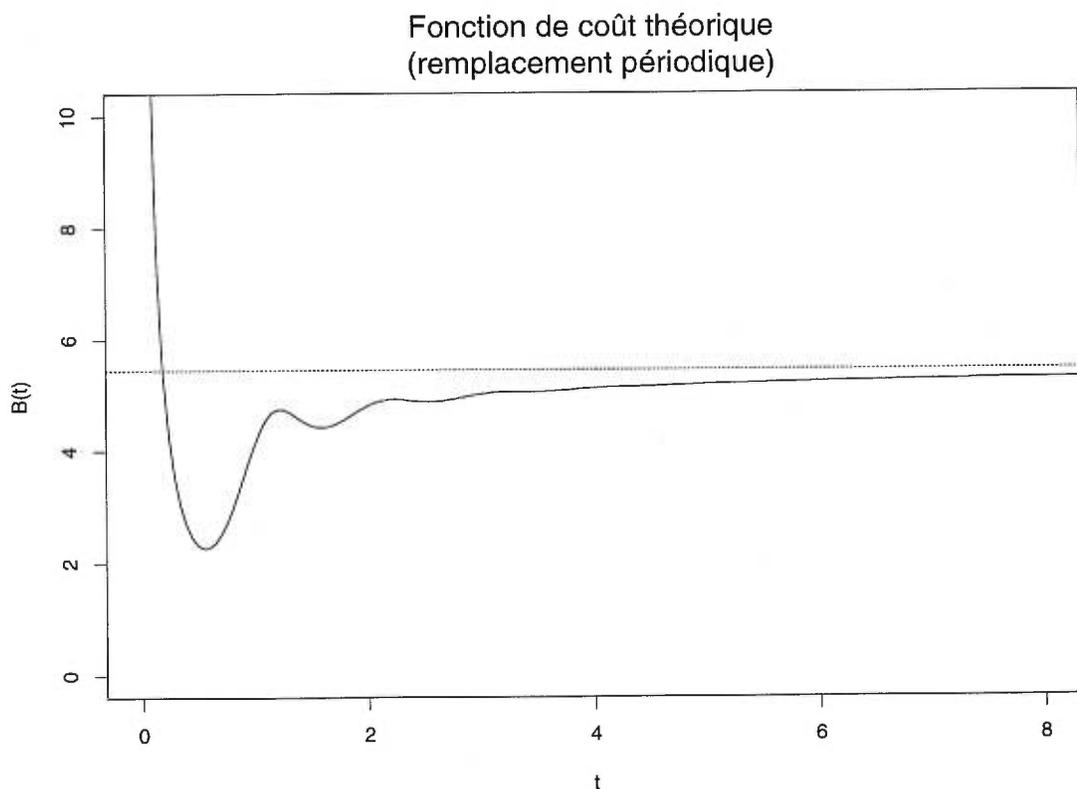


FIG. 3.1.2. *Fonction de coût théorique (remplacement périodique)*

On remarque que $B(0) = \infty$ et que la courbe décroît rapidement jusqu'au point où elle atteint son minimum. Par la suite, la fonction croît en douceur en subissant quelques oscillations dont l'amplitude diminue graduellement. La

limite quand $t \rightarrow \infty$ de la fonction de coût est $\frac{c_1}{\mu}$ (5,45 dans l'exemple). Le cas où $t = \infty$ peut être interprété comme étant la politique où un remplacement a lieu uniquement à la panne. Il permet donc d'affirmer qu'une politique optimale existe toujours. Par ailleurs, on peut montrer que si $F(t)$ est continue, $M(t)$ l'est aussi et par conséquent, $B(t)$ est également continue.

Voici quelques propriétés intéressantes de la fonction de coût (voir Cléroux [6] pour les preuves).

Proposition 3.1.1. *Si $F'(t) = f(t)$ et $M'(t) = m(t)$, $B'(t) = 0 \iff Tm(t) - M(t) = \frac{c_2}{c_1}$ et la politique optimale T^* est la solution de cette équation.*

Proposition 3.1.2. *Si T^* est la solution optimale, $B(T^*) = c_1 m(T^*)$.*

Des comparaisons peuvent être effectuées avec les politiques de remplacement préventif de type âge. Berg [4] a montré qu'une politique utilisant l'âge d'une pièce est plus efficace (coût minimal plus petit) qu'une politique de type périodique. Cependant, en employant une stratégie périodique, on remplace généralement plus souvent les composantes d'un système, ce qui a pour effet d'améliorer sa fiabilité.

3.1.1. Estimation de la fonction de coût

La littérature à propos du remplacement préventif de type périodique dans un contexte non paramétrique est relativement peu développée. Plus d'efforts ont été déployés par le passé pour trouver de bons estimateurs de $M(t)$ que de $B(t)$. Plusieurs auteurs ont considéré ce type de stratégie préventive, mais principalement dans un contexte paramétrique. La fonction de coût pour un tel type de prévention étant définie à partir de la fonction de renouvellement $M(t)$, les résultats concernant l'estimation de cette dernière s'appliquent évidemment à l'estimation de $B(t)$. Le reste de cette section est consacré à l'estimation du coût moyen par unité de temps ($B(t)$) en employant les estimateurs de $M(t)$ étudiés précédemment. Nous comparerons les estimateurs de $B(t)$ ainsi obtenus, afin de déterminer si certains sont supérieurs aux autres. Nous vérifierons ensuite si ces résultats sont en accord avec ceux obtenus lors de l'estimation de $M(t)$. Encore une fois, une attention particulière sera portée aux estimateurs non paramétriques.

3.1.2. Simulations

Dans cette section, des simulations sont effectuées, afin de comparer la performance des estimateurs de $B(t)$. Pour ce faire, nous utiliserons l'exemple présenté au chapitre précédent et nous emploierons les mêmes temps simulés. De nouveau, les distributions gamma, Weibull et normale tronquée seront considérées (pour des tailles échantillonales de 50, 100 et 200). Rappelons que la moyenne de chaque distribution était de 9080, l'écart-type de 3027 et que 1000 échantillons ont été simulés pour chaque simulation. Pour calculer la fonction de coût, les constantes c_1 et c_2 ont été fixées à 1100 et 100 respectivement. La programmation de ces différents estimateurs a de nouveau été faite en langage Fortran.

Dans le but de mesurer la performance relative de chaque estimateur, nous calculerons encore le biais, l'erreur quadratique moyenne et l'erreur relative moyenne de chacun d'eux (donnés par (2.1.1), (2.1.2) et (2.2.1)). Ces différentes mesures seront obtenues pour l'estimateur du coût minimal. Nous les obtiendrons aussi pour l'estimation du temps optimal que nous définissons par le temps où le coût minimal est atteint.

Les résultats obtenus pour différentes valeurs de t sont regroupés dans les tableaux 3.1.1 à 3.1.3. et une représentation graphique de la fonction de coût théorique et des exemples de ses estimateurs est disponible à la figure 3.1.3. Ces graphiques proviennent d'un des 1000 échantillons de taille 100 (choisi aléatoirement) issu de la distribution gamma.

TAB. 3.1.1. Estimation de $B(t)$ minimal pour la loi gamma

	n		$B_{2n}(t)$	$B_{3n}(t)$	$B_{4n}(t)$	$B_{5n}(t)$
cout min.	50	BM	-5,7371E-05	-7,6453E-03	-4,5820E-03	-3,4784E-03
		EQM	1,2224E-05	6,6626E-05	3,3557E-05	3,2287E-05
		ERM	8,9330E-02	2,5123E-01	1,6569E-01	1,5814E-01
	100	BM	-5,0408E-05	-5,2935E-03	-3,1468E-03	-1,5023E-03
		EQM	6,3620E-06	3,5237E-05	1,8673E-05	1,6110E-05
		ERM	6,5760E-02	1,7566E-01	1,2075E-01	1,0680E-01
	200	BM	8,1893E-05	-3,3448E-03	-2,0576E-03	-4,4914E-04
		EQM	3,4001E-06	1,6198E-05	9,5439E-06	7,0048E-06
		ERM	4,7329E-02	1,1469E-01	8,4637E-02	6,9059E-02
temps opt.	50	BM	40,626	584,134	339,344	-12,951
		EQM	115793	660132	516492	273177
		ERM	6,7313E-02	1,6676E-01	1,4168E-01	1,0527E-01
	100	BM	24,609	272,414	224,572	-76,659
		EQM	60182	292560	314871	181472
		ERM	4,9078E-02	1,0802E-01	1,1107E-01	8,6893E-02
	200	BM	1,419	129,680	136,802	-51,204
		EQM	31529	175194	192642	140537
		ERM	3,5185E-02	8,3116E-02	8,6717E-02	7,5763E-02

3.1.3. Analyse et conclusion

Soulignons tout d'abord que $B_{1n}(t)$, la fonction de coût estimée à l'aide de l'estimateur asymptotique de $M(t)$ ($M_{1n}(t)$) ne possède pas de minimum fini. Ceci s'explique par le fait que le coût minimal est atteint à un instant t nettement plus petit que la moyenne. En fait, la moyenne de chaque distribution est de 9080, alors

TAB. 3.1.2. Estimation de $B(t)$ minimal pour la loi de Weibull

	n		$B_{2n}(t)$	$B_{3n}(t)$	$B_{4n}(t)$	$B_{5n}(t)$
cout min.	50	BM	1,6907E-04	-1,0230E-02	-5,9204E-03	-3,2380E-03
		EQM	2,4201E-05	1,2959E-04	7,0105E-05	6,2448E-05
		ERM	0,1045468	0,279703	0,189805	0,1713354
	100	BM	-3,3314E-05	-6,9373E-03	-4,3187E-03	-1,3937E-03
		EQM	1,2160E-05	6,6752E-05	3,9212E-05	2,9740E-05
		ERM	0,07415	0,19342	0,14175	0,11587
	200	BM	3,7750E-05	-4,1580E-03	-2,7431E-03	-3,4115E-04
		EQM	5,9959E-06	2,8089E-05	1,8624E-05	1,3668E-05
		ERM	0,05265	0,12053	0,09512	0,08037
temps opt.	50	BM	34,319	381,861	299,721	-372,948
		EQM	102114	687685	796060	595936
		ERM	0,06563	0,17126	0,18376	0,15800
	100	BM	26,117	162,612	183,434	-80,906
		EQM	51477	411505	493436	381742
		ERM	0,04696	0,13142	0,14448	0,13012
	200	BM	0,119	76,919	106,660	-6,112
		EQM	26268	288841	314230	264366
		ERM	0,03352	0,11331	0,11800	0,10892

que le coût minimal théorique est atteint en $t = 3979$ pour la distribution gamma, en $t = 3858$ pour la Weibull et en $t = 4096$ pour la normale tronquée.

Dans un deuxième temps, c'est toujours sans surprise que $B_{2n}(t)$, l'estimateur paramétrique, est le plus performant. Il est donc évident que lorsque la distribution selon laquelle les temps observés sont distribués est connue, $B_{2n}(t)$ est le meilleur choix. Dans le cas contraire, cet estimateur est inefficace.

TAB. 3.1.3. Estimation de $B(t)$ minimal pour la loi normale tronquée

	n		$B_{2n}(t)$	$B_{3n}(t)$	$B_{4n}(t)$	$B_{5n}(t)$
cout min.	50	BM	-9,0522E-06	-9,8806E-03	-5,4099E-03	-1,6772E-03
		EQM	2,9420E-05	1,3062E-04	7,3470E-05	7,6413E-05
		ERM	0,11727	0,27592	0,19436	0,18937
	100	BM	9,0059E-05	-6,1434E-03	-3,5132E-03	2,0479E-04
		EQM	1,4627E-05	5,8904E-05	3,5267E-05	3,0307E-05
		ERM	0,08195	0,17790	0,13103	0,11549
	200	BM	9,7451E-05	-3,3688E-03	-1,9547E-03	6,8180E-04
		EQM	8,1506E-06	2,5321E-05	1,8171E-05	1,7154E-05
		ERM	0,06155	0,11236	0,09307	0,08754
temps opt.	50	BM	47,086	327,728	305,940	-143,287
		EQM	79500	812725	888412	580460
		ERM	0,05343	0,17466	0,18250	0,14890
	100	BM	22,464	66,123	118,118	-94,832
		EQM	35972	447740	519948	428539
		ERM	0,03655	0,13233	0,14052	0,12843
	200	BM	10,358	-7,180	59,055	23,490
		EQM	19214	314538	343678	304289
		ERM	0,02670	0,11267	0,11572	0,10682

La lutte entre les trois autres estimateurs est toutefois plus corsée. Les résultats obtenus sont davantage rapprochés les uns des autres. Une constante ressort cependant; $B_{5n}(t)$, l'estimateur basé sur l'estimation du taux de panne pour le cas IFR est toujours le meilleur des trois. En effet, que ce soit pour le coût minimal ou pour le temps optimal et autant pour chaque distribution que pour chaque taille échantillonnale, $B_{5n}(t)$ performe mieux que ses compétiteurs. Les gains les plus appréciables sont situés au niveau du biais. Des améliorations sont aussi notées concernant les deux autres statistiques calculées, mais elles sont généralement

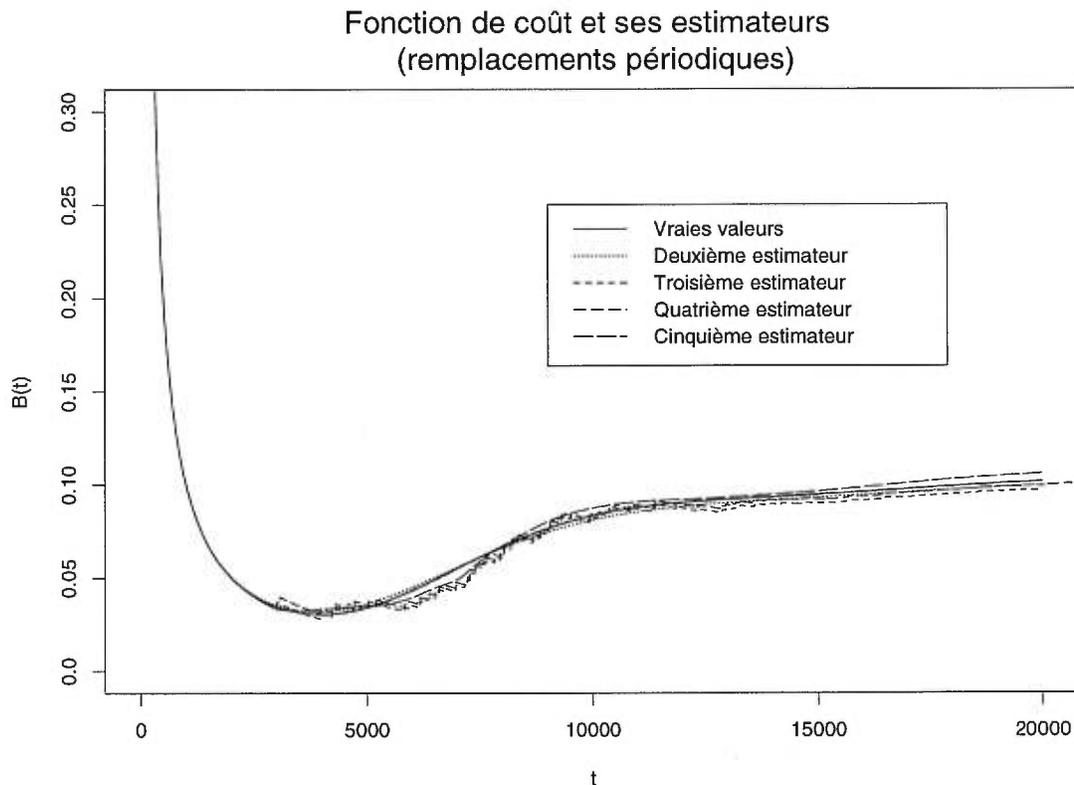


FIG. 3.1.3. *Fonction de coût et ses estimateurs (remplacement périodique)*

plus modérées. Par contre, nous sommes un peu ambivalent quant à la surprise d'obtenir de tels résultats. D'un côté, il est prévisible que $B_{5n}(t)$ soit meilleur que les deux autres en raison des hypothèses de départ sur la distribution (distribution à taux de panne non décroissant). Cependant, lors de l'estimation de $M(t)$, cette hypothèse ne permettait pas une meilleure estimation et $M_{5n}(t)$ était souvent dépassé par les deux autres estimateurs non paramétriques. Il semble donc que, contrairement à l'estimation de la fonction de renouvellement, l'estimateur le plus approprié pour estimer le coût minimal (et le temps optimal) soit $B_{5n}(t)$, l'estimateur pour le cas IFR (après $B_{2n}(t)$ bien sûr).

Par ailleurs, il apparaît que, pour estimer le coût minimal, $B_{4n}(t)$ soit un peu plus efficace que $B_{3n}(t)$. Le premier performe toujours mieux que le deuxième lors de nos simulations. Quant à l'estimateur du temps optimal, il semble plutôt que la situation soit inversée. Pour huit des neuf simulations effectuées, on obtient une meilleure estimation en employant l'estimateur de Frees ($M_{3n}(t)$) au lieu de l'estimateur empirique ($M_{4n}(t)$) pour estimer la fonction de renouvellement. En pratique, il sera toutefois toujours plus intéressant de mieux estimer le coût minimal que le temps optimal.

Mentionnons également que, à l'exception de $B_{5n}(t)$ pour la distribution normale tronquée avec $n = 100$ et $n = 200$, le biais des estimateurs $B_{3n}(t)$ à $B_{5n}(t)$ est toujours négatif. Pour le temps optimal, le biais du cinquième estimateur est généralement négatif, alors que les deux autres estimateurs surestiment, dans la plupart des cas, la vraie valeur de ce temps.

La figure 3.1.3 illustre bien la bonne performance de $B_{5n}(t)$ (estimateur pour le cas IFR) par rapport à $B_{3n}(t)$ (estimateur de Frees) et $B_{4n}(t)$ (estimateur empirique). On remarque aussi que ces deux estimateurs sont, encore une fois, très semblables et oscillent un peu plus que l'estimateur pour le cas IFR.

Par ailleurs, nous avons effectué une analyse de variance à deux facteurs (taille échantillonnale et estimateur) afin de vérifier ces différentes conclusions. Les résultats détaillés ne sont pas inclus dans ce mémoire puisque cette analyse ne fait pas ressortir de différence significative entre les différents estimateurs ou les différentes tailles échantillonnales employées. Que ce soit en ne comparant que les estimateurs non paramétriques ou en incluant l'estimateur paramétrique,

aucune différence, au niveau $\alpha = 0,05$, n'est observée. Cependant, le fait que les différentes statistiques calculées soient toujours plus petites pour l'estimateur paramétrique n'est certainement pas dû au hasard. Il en est de même pour l'estimateur pour le cas IFR par rapport aux autres estimateurs non paramétriques. Les statistiques considérées sont systématiquement plus petites pour $B_{5n}(t)$. Il est donc raisonnable de conclure qu'il y a un avantage à employer $B_{5n}(t)$ plutôt qu'un des deux autres estimateurs non paramétriques. Cet avantage est toutefois modeste.

Soulignons enfin que les temps nécessaires pour le calcul de ces différents estimateurs est proportionnel à celui pour obtenir les estimateurs de $M(t)$. $B_{3n}(t)$ et $B_{4n}(t)$ sont plutôt semblables, tandis que $B_{5n}(t)$ requiert plus de temps.

3.2. REMPLACEMENT PRÉVENTIF PÉRIODIQUE AVEC RÉPARATION MINIMALE À LA PANNE

Ce type particulier de remplacement périodique ne diffère de celui considéré précédemment que par l'action posée lors d'une panne. On répare une pièce défectueuse plutôt que de la remplacer par une neuve. L'item en question redevient alors fonctionnel et se retrouve dans le même état qu'il était immédiatement avant la panne. L'état général du système est évidemment moins bon, puisque la pièce brisée demeure usagée. Cependant, on est parfois gagnant avec une telle politique étant donné qu'une réparation est souvent beaucoup moins coûteuse qu'un remplacement. La stratégie que l'on privilégiera dépendra donc des coûts de réparation et de remplacement. Pour définir la fonction de coût associée à

une politique de remplacement préventif avec réparation minimale à la panne, la constante c_3 sera nécessaire. Elle représente le coût d'une réparation minimale. Notons que, contrairement au remplacement périodique où c_1 devait être supérieur à c_2 , aucune condition n'est nécessaire sur c_3 . Il est cependant peu logique de considérer des cas où une réparation est plus coûteuse qu'un remplacement ($c_3 > c_2$). Rappelons également qu'il existe plusieurs variantes à un tel type de remplacement préventif. Par exemple, on peut remplacer une pièce en panne par une pièce usagée ou plus simplement, n'effectuer aucune réparation ou remplacement. La meilleure politique à utiliser sera déterminée par le coût de chacune des actions possibles lors d'une panne.

Dans le contexte de remplacement périodique avec réparation minimale à la panne, le coût moyen par unité de temps sur un horizon infini est donné par:

$$\begin{aligned}
 D(T) &= \lim_{t \rightarrow \infty} \frac{E[C(t)]}{t} \\
 &= c_3 \lim_{t \rightarrow \infty} \frac{E[N_1(t)]}{t} + c_2 \lim_{t \rightarrow \infty} \frac{E[N_2(t)]}{t} \\
 &= c_3 \frac{\int_0^T r(u) du}{T} + c_2 \frac{1}{T}
 \end{aligned}$$

où $C(t)$, $N_1(t)$, $N_2(t)$, et c_2 sont définis comme à la section précédente et $r(t)$ représente le taux de panne. $E[N_1(t)]$ est donné par $\int_0^T r(u) du$ parce qu'en réparant une pièce défectueuse, on rétablit l'état de la pièce dans le même qu'il était avant la panne, c'est-à-dire que la fonction du taux de panne demeure la même.

L'allure de cette fonction de coût ressemble beaucoup à celle d'une stratégie de remplacement périodique présentée à la section précédente. Dans cette section, nous nous intéresserons donc à $D(t)$ et, en particulier, à son minimum.

3.2.1. Estimation de la fonction de coût

Si l'estimation de la fonction de coût dans un contexte de remplacement périodique était relativement peu présente dans la littérature, celle pour le cas où on effectue une réparation minimale à la panne est à toute fin pratique inexistante. Nous y accorderons donc une attention particulière dans cette section. Autant des estimateurs paramétriques que non paramétriques seront étudiés, puis comparés. Un parallèle sera également fait avec les différentes estimations, simulations et conclusions présentées jusqu'à maintenant dans ce mémoire.

L'estimation de la fonction de coût donnée par $D(t)$ nécessite au préalable l'estimation du taux de panne. Cette fois, nous considérerons trois techniques différentes pour y parvenir. La première nous permettra d'obtenir un estimateur paramétrique. De la même façon que lors de l'obtention de $M_{2n}(t)$, la connaissance de la distribution des temps de durée de vie observés nous permet d'estimer $F(t)$. On trouve d'abord des estimateurs convergents des paramètres de cette distribution, puis l'estimation de $F(t)$ est directe. On définit ensuite $r_{1n}(t)$ en utilisant la relation

$$1 - F(x) = \exp[-R(x)] = \exp\left[-\int_{-\infty}^x r(z)dz\right] \quad (3.2.1)$$

Notre deuxième estimateur, $r_{2n}(t)$, est, quant à lui, non paramétrique. On estime d'abord la fonction de distribution théorique à l'aide de la fonction de distribution empirique et on utilise encore la relation (3.2.1).

Pour le troisième estimateur, nous aurons recours à l'estimateur proposé par Marshall et Proschan [14] et que nous avons introduit lors de la présentation de notre cinquième estimateur de $M(t)$ à la section 1.3.5. Rappelons que cette estimation du maximum de vraisemblance n'est valable que pour le cas où le taux de panne est non décroissant. Il est donné par:

$$\hat{r}_n(x) = \begin{cases} 0 & \text{si } x < X_1 \\ \hat{r}_n(X_i) & \text{si } X_i \leq x < X_{i+1} \\ \infty & \text{si } x \geq X_n \end{cases}$$

où

$$\hat{r}_n(X_i) = \min_{v \geq i+1} \max_{u \leq i} (v - u) \left[\sum_{j=u}^{v-1} (n - j)(X_{j+1} - X_j) \right]^{-1}$$

Ces trois méthodes différentes pour estimer le taux de panne conduisent donc à trois estimateurs différents de la fonction de coût pour une stratégie de remplacement périodique avec réparation minimale à la panne. La performance des trois estimateurs sera mesurée à l'aide de simulations que nous présentons à la section suivante.

3.2.2. Simulations

L'exemple du chapitre 2 et repris à la section précédente est à nouveau employé pour mesurer et comparer l'efficacité de nos estimateurs. Les mêmes statistiques (biais, erreur quadratique moyenne, erreur relative) sont toujours calculées et permettent de vérifier le rendement des estimateurs du coût minimal et du temps optimal. Pour effectuer nos simulations, nous avons fixé le coût d'une réparation minimale (c_3) à 50. Les programmes utilisés pour ces estimations sont disponibles aux parties 6 à 8 de la section 2 de l'annexe B.

Les résultats de nos simulations sont regroupés dans les tableaux 3.2.1 à 3.2.3 et une représentation graphique de la fonction de coût théorique et de ses estimateurs est fournie à la figure 3.2.1. De nouveau, un des échantillons de taille 100 provenant de la distribution gamma a été utilisé pour ces graphiques.

TAB. 3.2.1. Estimation de $D(t)$ minimal pour la loi gamma

t	n		$D_{1n}(t)$	$D_{2n}(t)$	$D_{3n}(t)$
cout min.	50	BM	-6.5567E-06	-4.8549E-04	-4.4603E-04
		EQM	5.4876E-07	8.6763E-07	8.3830E-07
		ERM	3.8203E-02	4.9307E-02	4.8290E-02
	100	BM	-3.0274E-05	-3.7976E-04	-3.1464E-04
		EQM	2.7116E-07	4.9098E-07	4.5576E-07
		ERM	2.7341E-02	3.7086E-02	3.5464E-02
	200	BM	-1.3911E-07	-2.5257E-04	-1.9607E-04
		EQM	1.3596E-07	2.3765E-07	2.1877E-07
		ERM	1.9399E-02	2.5671E-02	2.4642E-02
temps opt.	50	BM	29.216	796.247	299.147
		EQM	732506	5051570	2905220
		ERM	0.06985	0.16172	0.12994
	100	BM	36.137	679.379	201.413
		EQM	357928	3403760	1707190
		ERM	0.04950	0.126869	0.10086
	200	BM	23.174	364.704	132.653
		EQM	174080	1599780	967545
		ERM	0.03424	0.08926	0.07927

3.2.3. Analyse et conclusion

Encore une fois, l'estimateur paramétrique est celui qui offre les meilleurs résultats. Peu importe la distribution ou la taille échantillonnale, $D_{1n}(t)$ est sans contredit le plus performant des trois. Cependant, les fortes hypothèses de départ sur la distribution sont souvent un obstacle à son utilisation. Dans un tel cas, on doit se rabattre sur l'un des deux autres estimateurs. Ces derniers présentent des résultats plutôt semblables, mais de façon générale, $D_{3n}(t)$ (l'estimateur pour

TAB. 3.2.2. Estimation de $D(t)$ minimal pour la loi de Weibull

t	n		$D_{1n}(t)$	$D_{2n}(t)$	$D_{3n}(t)$
cout min.	50	BM	3.0953E-07	-3.7218E-04	-3.8606E-04
		EQM	4.5387E-07	6.9244E-07	7.2640E-07
		ERM	3.6656E-02	4.5618E-02	4.6724E-02
	100	BM	-1.3372E-06	-2.8104E-04	-2.5755E-04
		EQM	2.3927E-07	3.7122E-07	3.7049E-07
		ERM	2.6379E-02	3.3373E-02	3.3240E-02
	200	BM	-8.7703E-06	-2.0555E-04	-1.7264E-04
		EQM	1.2296E-07	1.9605E-07	1.8783E-07
		ERM	1.8717E-02	2.4166E-02	2.3627E-02
temps opt.	50	BM	26.291	396.241	174.301
		EQM	277776	1578700	1348690
		ERM	4.3269E-02	9.7120E-02	9.1656E-02
	100	BM	1.460	234.591	113.407
		EQM	136931	880875	755437
		ERM	3.0453E-02	7.5053E-02	7.0819E-02
	200	BM	16.157	180.724	64.999
		EQM	63790	558659	479062
		ERM	2.0837E-02	5.9444E-02	5.6022E-02

le cas IFR) est légèrement supérieur. La plus grande amélioration de $D_{3n}(t)$ par rapport à $D_{2n}(t)$ est la réduction de plus de la moitié du biais pour l'estimation du temps optimal. L'erreur quadratique moyenne et l'erreur relative, autant pour le coût minimal que pour le temps optimal, sont aussi inférieures. Cet avantage de $D_{3n}(t)$ sur $D_{2n}(t)$ est intéressant, d'autant plus que l'hypothèse d'une distribution à taux de panne non décroissant n'est pas vraiment restrictive, puisque la plupart des distributions considérées en théorie de la fiabilité le sont.

TAB. 3.2.3. Estimation de $D(t)$ minimal pour la loi normale tronquée

t	n		$D_{1n}(t)$	$D_{2n}(t)$	$D_{3n}(t)$
cout min.	50	BM	1.5273E-07	-2.9999E-04	-3.2965E-04
		EQM	4.3174E-07	6.3279E-07	6.6627E-07
		ERM	3.4921E-02	4.3214E-02	4.4430E-02
	100	BM	-2.6371E-05	-2.3574E-04	-2.1699E-04
		EQM	2.1292E-07	3.4581E-07	3.4686E-07
		ERM	3.6719E-02	3.2009E-02	3.2001E-02
	200	BM	1.5433E-06	-1.6882E-04	-1.4183E-04
		EQM	1.1631E-07	1.8669E-07	1.8277E-07
		ERM	1.8020E-02	2.3859E-02	2.3523E-02
temps opt.	50	BM	-8.918	300.999	137.968
		EQM	216334	1341370	1204220
		ERM	3.9078E-02	9.2745E-02	9.0075E-02
	100	BM	24.400	276.082	149.777
		EQM	123382	991506	820873
		ERM	2.9951E-02	7.8072E-02	7.3428E-02
	200	BM	4.204	169.015	77.458
		EQM	59474	519298	468794
		ERM	2.0523E-02	5.9995E-02	5.7512E-02

Par contre, mentionnons que $D_{2n}(t)$ a mieux performé que $D_{3n}(t)$ pour les distributions Weibull et normale tronquée lorsque $n = 50$. En fait, l'écart entre ces deux estimations semble s'accroître avec l'augmentation de la taille échantillonnale. Cette observation est importante, puisqu'en pratique il est fréquent de ne posséder qu'un nombre restreint de temps observés. Cette particularité n'avait pas été remarquée lors de l'estimation de $B(t)$ (la fonction de coût pour un remplacement périodique).

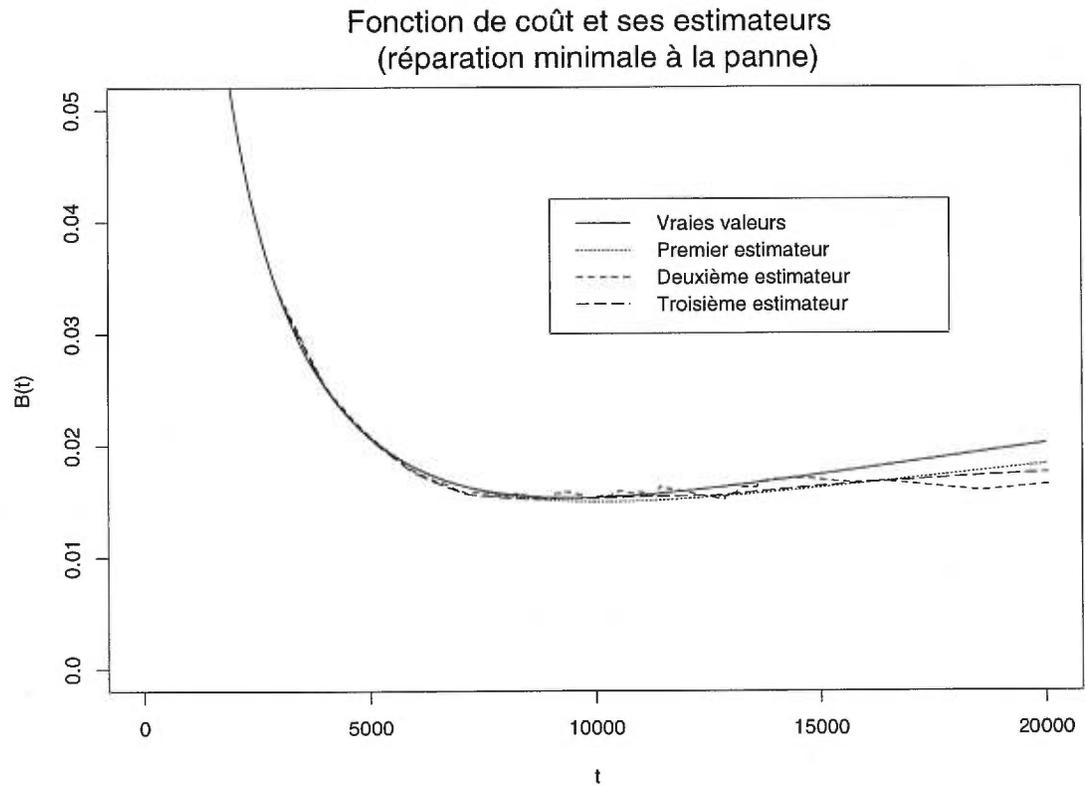


FIG. 3.2.1. *Fonction de coût et ses estimateurs (réparation minimale à la panne)*

Soulignons par ailleurs que le biais de tous les estimateurs du coût minimal est pratiquement toujours négatif, alors que le temps optimal est toujours surévalué.

A la lumière de la figure 3.2.1, on note à nouveau la bonne performance de l'estimateur paramétrique, alors que l'estimateur empirique est plus oscillant. Le graphique fait aussi ressortir la plus grande différence entre les estimateurs par rapport à la fonction théorique pour de plus grandes valeurs de t . Cet aspect était moins remarquable lors de l'estimation de $B(t)$, à la section précédente.

A nouveau, nous avons effectué une analyse de variance afin de vérifier si les facteurs taille échantillonnale et estimateur ont un effet sur les résultats obtenus. Encore une fois, aucune différence significative n'est décelée, au niveau $\alpha = 0,05$. Tout comme précédemment, nous concluons que la différence notée en observant les tableaux 3.2.1 à 3.2.3 existe, mais que cette dernière est plutôt modérée.

Enfin, en ce qui concerne le temps nécessaire pour le calcul des estimateurs, notons que les plus rapides à obtenir ont été $D_{2n}(t)$ et $D_{3n}(t)$. L'estimateur pour le cas IFR n'a pas été long à calculer puisque, cette fois-ci, il ne nécessitait pas l'estimation de la fonction de renouvellement.

CONCLUSION

Il est facile de se convaincre de l'importance des politiques de remplacement préventif pour une entreprise. Le développement de stratégies de maintenance, de réparation et de remplacement permet une meilleure planification des ressources humaines, matérielles et financières. Ceci s'avère primordial à une époque où les budgets sont de plus en plus serrés. La fonction de coût associée à une politique de remplacement, en particulier son minimum, est un bon outil pour obtenir la stratégie optimale à employer.

Dans ce mémoire, nous avons développé différentes façons d'estimer la fonction de coût pour une catégorie particulière de remplacement préventif: les remplacements périodiques. Nous n'avons pas beaucoup insisté sur les stratégies préventives basées sur l'âge d'une pièce, l'autre catégorie importante, puisque le sujet a été abordé plus fréquemment dans la littérature.

Afin d'estimer adéquatement la fonction de coût et son minimum, l'utilisation de la fonction de renouvellement, le nombre moyen de remplacements dans l'intervalle $[0, t]$, est essentielle. Notre étude exhaustive à propos de cette fonction effectuée dans les deux premiers chapitres nous a permis de comparer cinq estimateurs et de nous fournir une bonne base pour l'estimation de la fonction de coût. La suprématie de l'estimateur paramétrique de $M(t)$ ne fait plus aucun doute à la suite de nos simulations. Cependant, les estimateurs non paramétriques

de la fonction de renouvellement ont procuré des résultats très satisfaisants en l'absence d'information sur la distribution d'origine des temps observés.

Notre estimateur asymptotique ($M_{1n}(t)$) s'est comporté étonnamment bien pour d'assez grandes valeurs de t , mais il a été peu utile pour l'estimation de la fonction de coût. Pour l'utiliser, de lourdes contraintes sur la valeur des constantes de coût auraient été nécessaires pour que le coût minimal soit atteint à un instant t suffisamment grand. Par ailleurs, l'estimateur pour le cas IFR ($M_{5n}(t)$) ne s'est pas démarqué pour l'estimation de $M(t)$. Les estimateurs de Frees ($M_{3n}(t)$) et empirique ($M_{4n}(t)$) ont mené à des résultats relativement semblables, lors de nos simulations, avec un certain avantage parfois pour le deuxième. Cette conclusion est quelque peu surprenante, puisqu'elle est contradictoire avec les résultats des simulations issues de la littérature. Les différences entre nos deux estimateurs ne sont toutefois pas importantes au point où elles seraient inquiétantes.

Les méthodes non paramétriques utilisées pour l'estimation de la fonction de renouvellement ont naturellement été employées pour l'estimation de la fonction de coût d'une politique de remplacement préventif périodique. Cette approche est pratiquement inexistante dans la littérature; les auteurs ayant surtout développé des techniques paramétriques. En plus des stratégies de remplacement périodique, nous avons étudié une de ses variantes: le remplacement préventif périodique avec réparation minimale à la panne.

Hormis l'estimateur paramétrique qui est toujours supérieur aux autres, le meilleur estimateur du coût minimal et de l'instant où il est atteint s'est avéré être l'estimateur pour le cas IFR. Contrairement à l'estimation de $M(t)$, l'hypothèse

supplémentaire sur la non décroissance du taux de panne a permis l'obtention d'un estimateur un peu plus précis. Ceci est vrai pour les deux types de remplacement considérés. Puisque l'hypothèse d'une distribution IFR n'est pas très contraignante en théorie de la fiabilité, l'emploi de $B_{5n}(t)$ et de $D_{3n}(t)$ procure un avantage certain, à tout le moins d'après les résultats de nos simulations.

Les deux types de remplacement présentés en détail dans cette étude sont fréquents en théorie du renouvellement, mais sont très simplifiés. Pour respecter davantage la réalité, de nombreux facteurs, tels l'inflation et le développement de nouvelles méthodes d'entretien et de réparation devraient être pris en considération. Les différentes observations notées à la suite de nos simulations devraient s'appliquer malgré tout. Ces deux stratégies préventives ont conduit à des résultats comparables quant à l'efficacité des estimateurs et au bon rendement de l'estimation pour le cas IFR. Il pourrait être intéressant de vérifier si ces conclusions s'appliquent aussi à d'autres types de remplacement. Le cas où on remplace une pièce défectueuse par une pièce usagée, mais fonctionnelle pourrait être considéré, tout comme celui où aucune action n'est posée lors d'une panne. Aussi, il peut arriver que les coûts de remplacement ou de réparation soient aléatoires. Une combinaison de ces différentes méthodes de prévention en fonction du coût de chacune des actions possibles serait aussi un cas intéressant à étudier.

Annexe A

ANALYSE DE VARIANCE

A.1. DISTRIBUTION DE WEIBULL

Le tableau A.1.1 résume la table d'anova obtenue. Le tableau A.1.2 donne les résultats du test de Scheffé qui compare les estimateurs deux à deux. La figure A.1.1 illustre l'interaction entre le temps et les estimateurs. Ces résultats ont été obtenu à l'aide du logiciel SAS.

TAB. A.1.1. *Table d'anova pour la loi de Weibull*

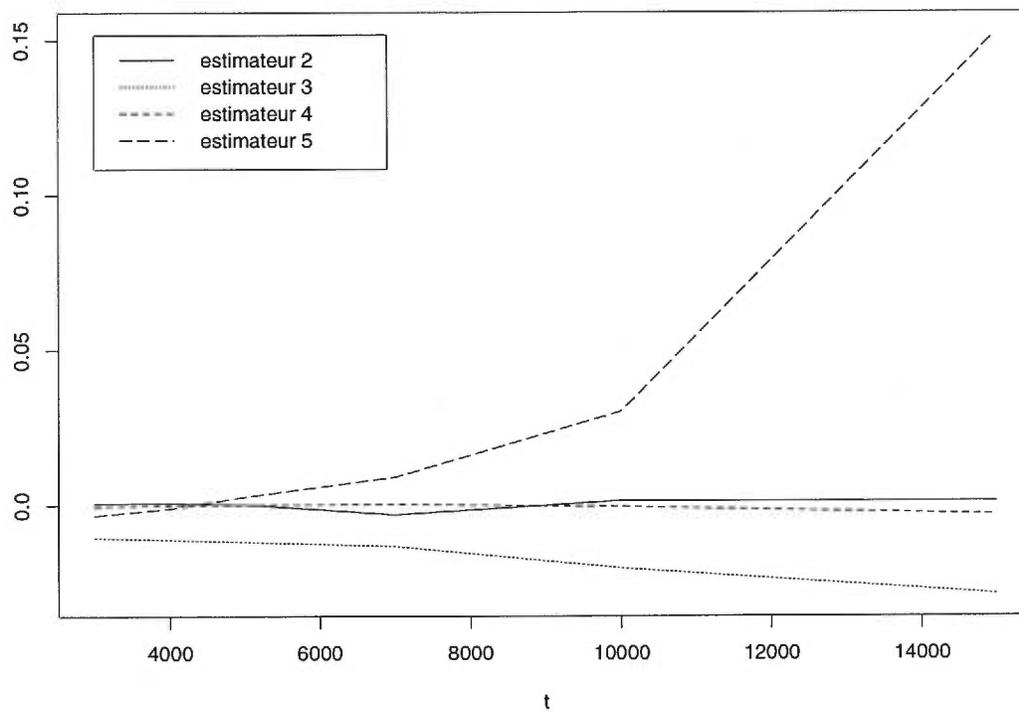
		Biais moy.		Eq. quad. moy.		Err. rel. moy.	
Facteur	d. lib.	Stat. F	Pr. > F	Stat. F	Pr. > F	Stat. F	Pr. > F
n	2	0,61	0,5491	1,01	0,3762	115,84	0,0001
t	5	2,16	0,0850	1,01	0,4281	295,95	0,0001
estimateur	3	7,47	0,0007	1,01	0,4033	38,54	0,0001
n*t	10	0,89	0,5498	1,00	0,4637	9,26	0,0001
n*estimateur	6	1,53	0,2030	1,00	0,4430	2,36	0,0552
t*estimateur	15	3,10	0,0040	1,01	0,4753	10,14	0,0001

TAB. A.1.2. Contrastes entre les estimateurs (distribution de Weibull)

Biais moyen				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{5n}(t)$	-0.01439	0.01615	0.04669	Non
$M_{2n}(t) - M_{4n}(t)$	-0.02978	0.00076	0.03130	Non
$M_{2n}(t) - M_{3n}(t)$	-0.06188	-0.03134	-0.00080	Oui
$M_{3n}(t) - M_{4n}(t)$	-0.04593	-0.01539	0.01515	Non
$M_{3n}(t) - M_{5n}(t)$	-0.07803	-0.04749	-0.01695	Oui
$M_{4n}(t) - M_{5n}(t)$	-0.06263	-0.03209	-0.00155	Oui

Erreur relative moyenne				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{3n}(t)$	-0.16721	-0.12883	-0.09045	Oui
$M_{2n}(t) - M_{4n}(t)$	-0.13253	-0.09415	-0.05577	Oui
$M_{2n}(t) - M_{5n}(t)$	-0.14587	-0.10749	-0.06911	Oui
$M_{3n}(t) - M_{4n}(t)$	-0.00370	0.03468	0.07305	Non
$M_{3n}(t) - M_{5n}(t)$	-0.01704	0.02133	0.05971	Non
$M_{4n}(t) - M_{5n}(t)$	-0.05172	-0.01334	0.02504	Non

Weibull – biais



Weibull – erreur relative

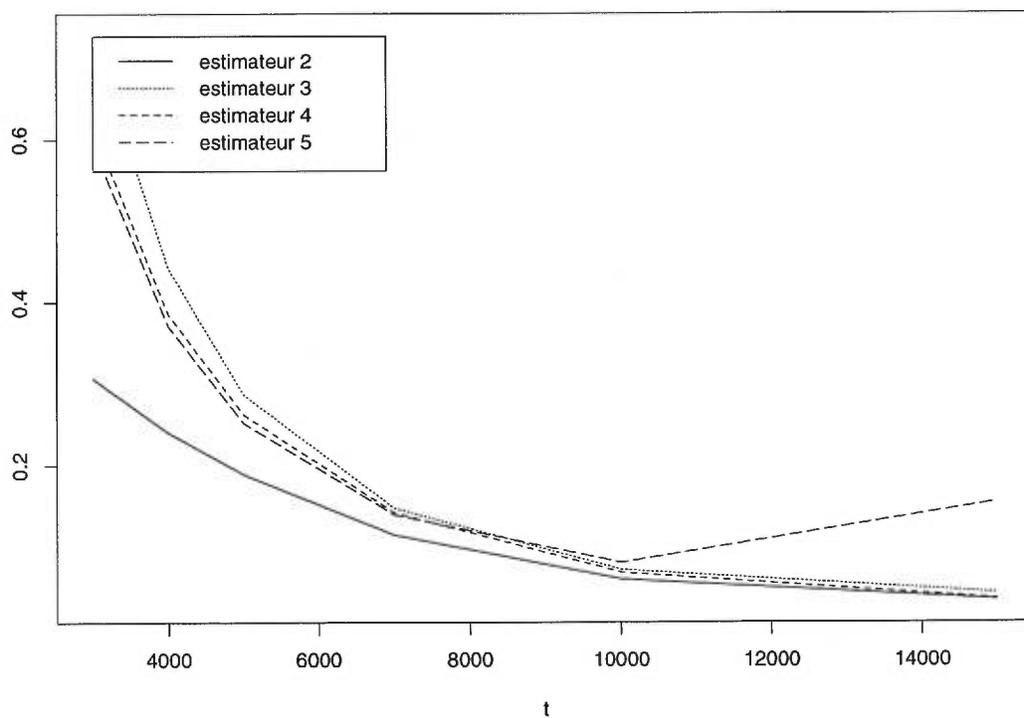


FIG. A.1.1. Interaction entre le temps et les estimateurs (distribution de Weibull)

A.2. DISTRIBUTION NORMALE TRONQUÉE

Le tableau A.2.1 résume la table d'anova obtenue. Le tableau A.2.2 donne les résultats du test de Scheffé qui compare les estimateurs deux à deux. La figure A.2.1 illustre l'interaction entre le temps et les estimateurs.

TAB. A.2.1. *Table d'anova pour la loi normale tronquée*

		Biais moy.		Eq. quad. moy.		Err. rel. moy.	
Facteur	d. lib.	Stat. F	Pr. > F	Stat. F	Pr. > F	Stat. F	Pr. > F
n	2	0,28	0,7543	1,03	0,3688	213,39	0,0001
t	5	3,28	0,0175	1,09	0,3842	483,97	0,0001
estimateur	3	15,49	0,0001	1,08	0,3716	59,42	0,0001
n*t	10	0,72	0,7001	1,01	0,4598	17,36	0,0001
n*estimateur	6	2,10	0,0832	1,00	0,4413	4,59	0,0020
t*estimateur	15	4,95	0,0001	1,07	0,4168	12,06	0,0001

TAB. A.2.2. Contrastes entre les estimateurs (distribution normale tronquée)

Biais moyen				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{3n}(t)$	-0.007219	0.014143	0.035504	Non
$M_{2n}(t) - M_{4n}(t)$	-0.023416	-0.002055	0.019307	Non
$M_{2n}(t) - M_{5n}(t)$	-0.007219	0.014143	0.035504	Non
$M_{3n}(t) - M_{4n}(t)$	-0.037559	-0.016198	0.005164	Non
$M_{3n}(t) - M_{5n}(t)$	-0.068958	-0.047596	-0.026234	Oui
$M_{4n}(t) - M_{5n}(t)$	-0.052760	-0.031398	-0.010037	Oui

Erreur relative moyenne				
Estimateurs	Borne inf.	Différence	Borne sup.	Diff. Significative
$M_{2n}(t) - M_{3n}(t)$	-0.143817	-0.114848	-0.085880	Oui
$M_{2n}(t) - M_{4n}(t)$	-0.117987	-0.089018	-0.060049	Oui
$M_{2n}(t) - M_{5n}(t)$	-0.138085	-0.109116	-0.080147	Oui
$M_{3n}(t) - M_{4n}(t)$	-0.003138	0.025830	0.054799	Non
$M_{3n}(t) - M_{5n}(t)$	-0.023236	0.005732	0.034701	Non
$M_{4n}(t) - M_{5n}(t)$	-0.049067	-0.020098	0.008871	Non

Normale tronquee – biais

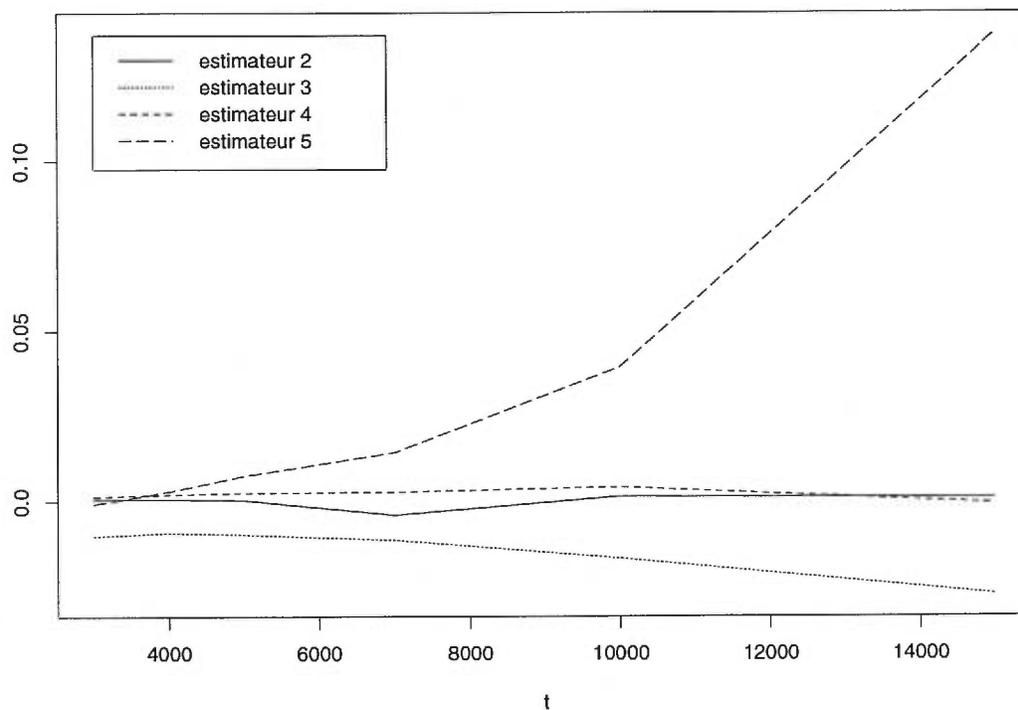
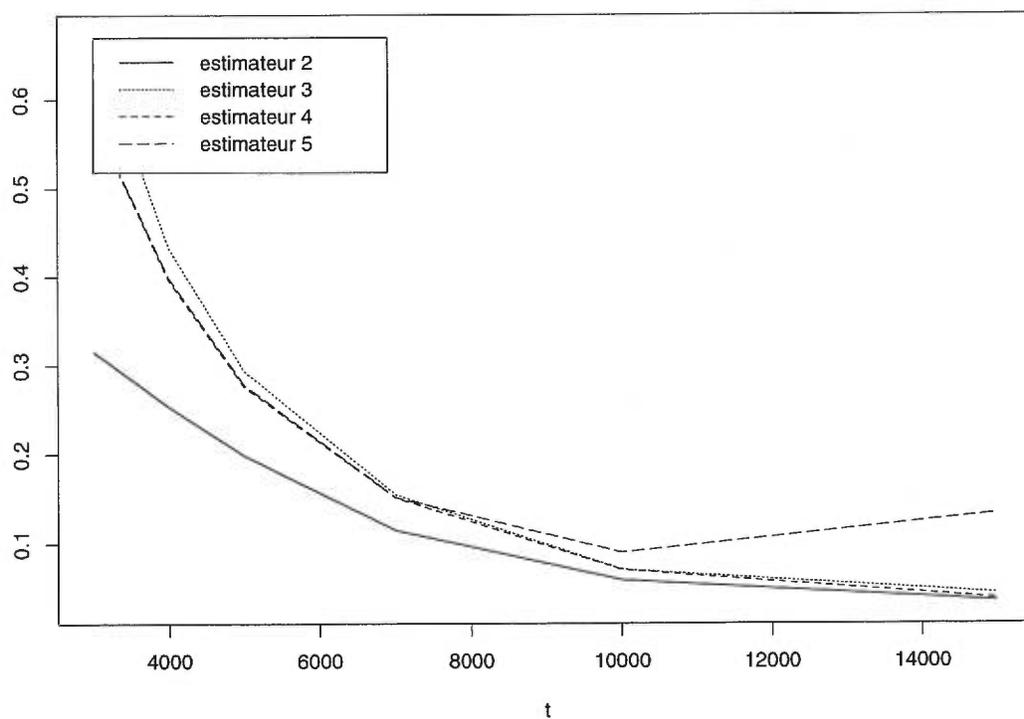
Normale tronquee – erreur relative – t^*est 

FIG. A.2.1. Interaction entre le temps et les estimateurs (distribution normale tronquée)

Annexe B

PROGRAMMES

Voici la liste des programmes qui ont été utilisés pour les calculs et les simulations. Ces programmes sont disponibles sur le réseau dms du département de mathématiques et de statistique de l'Université de Montréal. Ils sont situés dans le répertoire `"/home/cleroux/fiabilite/convolut"`. Des simulations ont été effectuées pour les distributions gamma, Weibull et normale tronquée avec des tailles échantillonales de 50, 100 et 200. Dans cet annexe, on retrouve les programmes adaptés pour la distribution gamma avec $n = 100$. Des changements mineurs permettent de les appliquer à d'autres distributions. Notons que chaque estimateur comporte son propre programme. L'ordre dans lequel ils sont exécutés n'a pas d'importance et il n'est pas obligatoire de tous les exécuter. La première section de cette annexe regroupe les commandes de S-Plus alors que la deuxième est consacrée aux programmes en langage Fortran.

B.1. SECTION 1: S-PLUS

Partie 1: Simulation des 1000 échantillons de tailles 50, 100 et 200 pour les trois distributions considérées.

Partie 2: Programme S-Plus qui a permis de calculer les résultats pour le premier estimateur (estimateur asymptotique).

B.2. SECTION 2: FORTRAN

- Partie 1: Calcul des paramètres alpha et lamda des 1000 distributions échantillonnales (gamma et weibull) et calcul des paramètres pour des distributions de moyenne 9080 et écart-type de 3027.
- Partie 2: Calcul des vraies valeurs de la fonction de renouvellement, du deuxième estimateur (estimateur paramétrique) et du cinquième estimateur (estimateur pour le cas IFR).
- Partie 3: Calcul du troisième estimateur (estimateur de Frees).
- Partie 4: Calcul du quatrième estimateur (estimateur empirique).
- Partie 5: Calcul de la fonction de densité échantillonnale pour les 1000 échantillons. Elles sont utilisées pour le calcul du cinquième estimateur.
- Partie 6: Estimation du coût minimum dans le cas d'un remplacement préventif avec réparation minimale à la panne - à partir du deuxième estimateur (estimateur paramétrique).
- Partie 7: Estimation du coût minimum dans le cas d'un remplacement préventif avec réparation minimale à la panne - à partir du troisième estimateur (estimateur de Frees).
- Partie 8: Estimation du coût minimum dans le cas d'un remplacement préventif avec réparation minimale à la panne - à partir du cinquième estimateur (estimateur pour le cas IFR).

Section 1: S-Plus

PARTIE 1:

=====

On simule 1000 échantillons de taille 100 provenant d'une distribution gamma de paramètre $\alpha=9$ et $\lambda=0.000990971$:

```
matrice.gamma_matrix(1:100000,ncol=100)
for (i in 1:1000){
  for(j in 1:100){
    matrice[i,j]_(1/0.000990971)*rgamma(1,9)}
    matrice[i,]_sort(matrice[i,])}
```

On simule 1000 échantillons de taille 100 provenant d'une distribution de Weibull de paramètre $\alpha=3.30312$ et $\lambda=0.588990E-13$:

```
matrice.weibull_matrix(1:100000,ncol=100)
for (i in 1:1000){
  for(j in 1:100){
    matrice.weibull[i,j]_rweibull(1,.330312E+01,1)/9.879355E-05}
    matrice.weibull[i,]_sort(matrice2.weibull[i,])}
```

On simule 1000 échantillons de taille 100 provenant d'une distribution normale tronquée. On prend les valeurs positives d'une normale($0.90653457E+04$, $0.30488999E$

```
norm_(rnorm(101000,.90653457E+04,.30488999E+04))
norm_norm[norm>0]
norm_norm[1:100000]
matrice.normale_matrix(norm,ncol=100)
for (i in 1:1000){
  matrice.normale[i,]_sort(matrice.normale[i,])}
```

PARTIE 2:

=====

On calcule le premier estimateur de $M(t)$ (estimateur asymptotique)

On observe la fonction aux temps suivants:

```
temps_c(200,1000,3000,4000,5000,7000,10000,15000)
```

```
sim.g.un_matrix(1:100000,ncol=100)
for (i in 1:1000){
  sim.g.un[i,]_as.matrix(m1((1:100)*200,matrice[i,])$M1)}
```

Dans la matrice `sim.g.un`, on retrouve les 1000 fonctions $M(t)$ (une par ligne).

Calculs effectués:

Biais:

```
biais.g.1_matrix(1:8000,ncol=8)
for (i in 1:1000){
  biais.g.1[i,]_sim.g.un[i,temps/200]-gamma1[temps/10,2]}
```

Les biais sont dans le vecteur suivant:

```
b.g.1.m_c(mean(biais.g.1[,1]),mean(biais.g.1[,2]),mean(biais.g.1[,3]),mean(biais
.g.1[,4]),mean(biais.g.1[,5]),mean(biais.g.1[,6]),mean(biais.g.1[,7]),mean(biais
.g.1[,8]))
```

Erreur quadratique moyenne:

```

-----

ce.g.1_matrix(1:8000,ncol=8)
for (i in 1:1000){
  ce.g.1[i,]_(sim.g.un[i,temps/200]-gamma1[temps/10,2])^2}

```

Les erreurs quadratiques moyennes sont dans le vecteur suivant:

```

ce.g.1.m_c(mean(ce.g.1[,1]),mean(ce.g.1[,2]),mean(ce.g.1[,3]),mean(ce.g.1[,4]),m
ean(ce.g.1[,5]),mean(ce.g.1[,6]),mean(ce.g.1[,7]),mean(ce.g.1[,8]))

```

Erreur relative:

```

-----

er.g.1_matrix(1:8000,ncol=8)
for (i in 1:1000){
  er.g.1[i,]_abs(sim.g.un[i,temps/200]-gamma1[temps/10,2])}

```

Les erreurs relatives sont dans le vecteur suivant:

```

er.g.1.m_c(mean(er.g.1[,1]),mean(er.g.1[,2]),mean(er.g.1[,3]),mean(er.g.1[,4]),m
ean(er.g.1[,5]),mean(er.g.1[,6]),mean(er.g.1[,7]),mean(er.g.1[,8]))

```

Fonction de cout:

```

-----

couts.g.1_matrix(1:100000,ncol=100)
  for (i in 1:1000){
    for (j in 1:100){
      couts.g.1[i,j]_(1100*sim.g.un[i,j]+100)/(200*j)}

```

En observant les fonctions de cout, on remarque qu'elles ne possèdent pas de minimum.

Section 2: Fortran

```

C *****
C *
C *   PROGRAMME   " Estimation de la fonction de renouvellement " *
C *
C *****
C *
C *   BUT:  Le but de ce programme est d'estimer la fonction de *
C *   ---  renouvellement a l'aide des differents estimateurs *
C *         enonces dans ce memoire.  Aussi, on calcule les vraies *
C *         valeurs de la fonction de renouvellement pour les trois *
C *         distributions considerees. *
C *
C *         On calcule aussi les couts minimaux (et les temps pour *
C *         lesquels on obtient ces couts) *
C *
C *   NOTE: Les programmes ci-joints sont adaptes pour la distri- *
C *         bution GAMMA, avec 1000 echantillons de taille 100. *
C *         Quelques modifications simples peuvent etre effectuees *
C *         facilement afin de calculer M(t) pour les deux autres *
C *         distributions considerees *
C *
C *****

C *****
C *
C *   PARTIE 1: *
C *   ----- *
C *   Cette partie calcule: *
C *
C *         Les parametres alpha et lambda des 1000 distributions *
C *         echantillonnales (pour gamma et weibull) *
C *         Cette partie provient de Beaucaire [3]. *
C *

```



```

WRITE(1,*)DD(5)
WRITE(2,*)1./BET

100 CONTINUE

DO 101 z=1,1000

moy=wmoy(z)
var=wvar(z)
DD(3) = moy/var
BET = var/moy
MOY2 = moy*moy
R = moy2/(moy2+var)
DD(5) = moy2/var

A(1)=0.0
A(2)=1.0
U=0.5

DO 20 J=1,26
  T=GAMMA(1.0+U)
  V=GAMMA(1.0+2.0*U)
  K=2

  IF (T*T .GT. R*V) K=1

  A(K)=U
  U=0.5*(A(1)+A(2))
20 CONTINUE

DD(5)=1.0/U
DD(6)=DD(5)*ALOG(T/MOY)
DD(4)=ALOG(DD(5))+DD(6)
U=EXP(DD(6))

c --- les parametres pour la weibull sont: -----

```

```
WRITE(3,*)DD(5)
WRITE(4,*)U

101 CONTINUE
    K=1
    A(2)=MOY
    T=MOY
    W=-0.1*MOY
31   T=T+W
36   R=MOY-T
    ELSQ=VAR+MOY*R
    EL=SQRT(ELSQ)
    V=0.70710678118655*T/EL
    DD(7)=1.2533141373155*EL*(1.0+SIGN(ERF(V),V))
    V=ELSQ-R*DD(7)*EXP(V*V)
    J=1

    IF (V.GT.0.0) J=2

    A(J)=T
    GOTO(30,35),K
30   GOTO(33,31),J

33   K=2
    L=0
35   T=0.5*(A(1)+A(2))
    L=L+1

    IF (L.LT.26) GOTO 36

    PRINT 49,T,EL
49   FORMAT (/////’0 Parametre pour NT... a = ',E13.8,' l = ',E13.8)

RETURN
END
```

```

C *****
C *
C *   Fonction GAMMA: Cette fonction permet de calculer la valeur *
C *                               de la fonction gamma(x) pour x reel *
C *
C *****

REAL FUNCTION GAMMA(X)
PARAMETER (MAX=8, EPS1=1.0E-12, EPS2=1.0E24)
REAL FACT
DIMENSION B(MAX)
DATA      B / 0.035868343, -0.193527818,  0.482199394,
+          -0.756704078,  0.918206857, -0.897056937,
+          0.988205891, -0.577191652      /

GAMMA=0.0
FACT=1.0
Y=X

      IF (Y-1.0) 2,6,1
2     IF (Y-EPS1) 7,7,8

7     GAMMA=EPS2

      RETURN

1     Y=Y-1.0
      IF (Y-1.0) 4,6,3

3     FACT=Y*FACT
      GOTO 1

8     FACT=1.0/Y

4     DO 5 I=1,8

```

```
      GAMMA=Y*(GAMMA+B(I))  
5     CONTINUE  
  
6     GAMMA=FACT*(1.0+GAMMA)  
  
      RETURN  
END
```

```

C *****
C *
C * PARTIE 2: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * LES VRAIES VALEURS DE LA FONCTION DE RENOUVELLEMENT *
C * LE DEUXIEME ESTIMATEUR (ESTIMATEUR PARAMETRIQUE) *
C * LE CINQUIEME ESTIMATEUR (ESTIMATEUR POUR LE CAS IFR) *
C * * *
C *****
C *
C * Pour calculer ces trois fonctions, on utilise, entre autre, *
C * la sous-routine CONVOL, developpee par R. Cleroux et D. *
C * McConalogue [6] *
C * * *
C * NB: Les fonctions de densites necessaires pour le calcul *
C * du cinquieme estimateur sont obtenues en utilisant la *
C * partie 5 de ce programme *
C * * *
C *****

```

PROGRAM APPEL

```

IMPLICIT REAL(A-H,O-Z)
DIMENSION alpha(1000),ala(1000)
COMMON/PARAM/DIST,ALFA,AL,X1,compte,compte2
DIMENSION Y1(2008),RE(2008)

```

```

C ---- Le fichier sim.t contient les valeurs de t pour lesquelles ---
C ---- on veut la valeur de M(t) ---
C ---- Les fichiers sim.alpha et sim.lamda contiennent les 1000 ----
C ---- estimations des parametres des 1000 dist. echantillonales ---

```

```

C ---- Le fichier sim.gamma contient les 1000 echantillons de ----
C ---- taille 100 ----

C ---- Le fichier sim.cinq.df contient les 1000 fonction de ----
C ---- densite echantillonnale (obtenues par la partie 5) ----

      OPEN(1,FILE='sim.t')
      OPEN(3,FILE='sim.concl')
      OPEN(4,FILE='sim.alpha')
      OPEN(5,FILE='sim.lamda')
      OPEN(7,FILE='sim.gamma')
      OPEN(8,FILE='sim.cinq.df')
      READ(1,*)Y1
      READ(4,*)alpha
      READ(5,*)ala
      READ(7,*)donnee1
      READ(8,*)donnee2

C      DETERMINATION DE LA DISTRIBUTION:
C      1=GAMMA
C      2=WEIBULL
C      3=NORMALE TRONQUEE
C      4=Distribution echantillonnale

      DIST=1

C ----- initialisation des constantes -----
      n=100
      compte=-(n-1)
      compte2=-(n-2)

C ---- On declare les valeurs theoriques de la fonction de ----
C ---- renouvellement pour differentes valeurs de t ainsi ----
C ---- que le cout minimal et le temps optimal theorique ----

```

```

vrai1=1.08807E-12
vrai2=1.04538E-06
vrai3=3.58845E-03
vrai4=2.03078E-02
vrai5=6.51914E-02
vrai6=0.263032
vrai7=6.70089e-01
vrai8=1.20255
cvrai=0.0305826
tvrai=3979

```

```

C --- La boucle 30 permet la repetition du programme 1000 fois -----
C --- ( pour les 1000 echantillons) -----

```

```

DO 30 I=1,1000
    ALFA=alpha(I)
    AL=ala(I)

aamin=1000000
compte=compte+100
compte2=compte2+99

```

```

C --- On calcule a l'interieur de cette boucle le cout minimal ---
C --- et le temps optimal. ---

```

```

DO 20 J=1,100
    z=20*J
    X1=Y1(z)
    CALL UN(RE(z))
    couts=(1100*RE(z)+100)/Y1(z)
    IF (couts.LE.aamin) THEN
        aamin=couts
        tmin=Y1(z)
        zmin=z
    ENDIF
20 CONTINUE

```

```

DO 40 J=zmin-20,zmin+19
X1=Y1(J)
CALL UN(RE(J))
  couts=(1100*RE(J)+100)/Y1(J)
  IF (couts.LE.aamin) THEN
    aamin=couts
    tmin=Y1(J)
  ENDIF
40 CONTINUE

```

C --- On calcule M(t) pour certains t fixes -----

```

X1=Y1(2001)
CALL UN(RE(2001))
X1=Y1(2002)
CALL UN(RE(2002))
X1=Y1(2003)
CALL UN(RE(2003))
X1=Y1(2004)
CALL UN(RE(2004))
X1=Y1(2005)
CALL UN(RE(2005))
X1=Y1(2006)
CALL UN(RE(2006))
X1=Y1(2007)
CALL DEUX(RE(2007))
X1=Y1(2008)
CALL DEUX(RE(2008))

```

C -- On calcule le biais, le carre moyen de l'erreur et l'erreur relative --

```

b0=b0+(RE(2001)-(vrai0))
b1=b1+(RE(2002)-(vrai1))
b2=b2+(RE(2003)-(vrai2))
b3=b3+(RE(2004)-(vrai3))
b4=b4+(RE(2005)-(vrai4))
b5=b5+(RE(2006)-(vrai5))

```

```

b6=b6+(RE(2007)-(vrai6))
b7=b7+(RE(2008)-(vrai7))

cm0=cm0+(RE(2001)-vrai0)**2
cm1=cm1+(RE(2002)-(vrai1))**2
cm2=cm2+(RE(2003)-(vrai2))**2
cm3=cm3+(RE(2004)-(vrai3))**2
cm4=cm4+(RE(2005)-(vrai4))**2
cm5=cm5+(RE(2006)-(vrai5))**2
cm6=cm6+(RE(2007)-(vrai6))**2
cm7=cm7+(RE(2008)-(vrai7))**2

er0=er0+(max(RE(2001),vrai0)-min(RE(2001),vrai0))
er1=er1+(max(RE(2002),(vrai1))-min(RE(2002),(vrai1)))
er2=er2+(max(RE(2003),vrai2)-min(RE(2003),vrai2))
er3=er3+(max(RE(2004),vrai3)-min(RE(2004),vrai3))
er4=er4+(max(RE(2005),vrai4)-min(RE(2005),vrai4))
er5=er5+(max(RE(2006),vrai5)-min(RE(2006),vrai5))
er6=er6+(max(RE(2007),vrai6)-min(RE(2007),vrai6))
er7=er7+(max(RE(2008),vrai7)-min(RE(2008),vrai7))

bco=bco+(aamin-cvrai)
cmco=cmco+(aamin-cvrai)**2
erco=erco+(max(aamin,cvrai)-min(aamin,cvrai))

bto=bto+(tmin-tvrai)
cmto=cmto+(tmin-tvrai)**2
erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))

30  CONTINUE

C  --  On conserve dans un fichier les resultats:  ---

WRITE(3,*)'t=200:',b0/1000,cm0/1000,er0/(1000*(vrai0))
WRITE(3,*)'t=1000:',b1/1000,cm1/1000,er1/(1000*(vrai1))

```

```

WRITE(3,*)'t=3000:',b2/1000,cm2/1000,er2/(1000*(vrai2))
WRITE(3,*)'t=4000:',b3/1000,cm3/1000,er3/(1000*(vrai2))
WRITE(3,*)'t=5000:',b4/1000,cm4/1000,er4/(1000*(vrai4))
WRITE(3,*)'t=7000:',b5/1000,cm5/1000,er5/(1000*(vrai5))
WRITE(3,*)'t=10000:',b6/1000,cm6/1000,er6/(1000*(vrai6))
WRITE(3,*)'t=15000:',b7/1000,cm7/1000,er7/(1000*(vrai7))
WRITE(3,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
WRITE(3,*)'le temps:',bto/1000,cmt0/1000,erto/(1000*tvrai)
TEMP=T
STOP
END

```

```

C *****
C *
C *   SOUS-ROUTINE UN:  cette sous-routine permet, en appelant la *
C *                   sous-routine CONVOL, de calculer la      *
C *                   fonction de renouvellement pour les t    *
C *                   plus petits (< 10 000 dans notre cas)    *
C *                   (la precision est plus grande (h=1))     *
C *
C *****

```

```

SUBROUTINE UN(REN)
DIMENSION CC(1000000),FN(1000000)
COMMON/PARAM/DIST,ALFA,AL,X1,compte,compte2
EXTERNAL DF

```

```

C
C   CALCUL DE LA FONCTION DE RENOUVELLEMENT REN
C

```

```

DIST=DIST
ALFA=ALFA
AL=AL
N=0
H=1
X1=X1
NPT=X1/H+1

```

```

REN=0
DO 100 I=1,5
CALL CONVOL(DF,H,NPT,FN,CC,N)
REN=REN+FN(NPT)
100 CONTINUE
T=H*(NPT-1)
RETURN
END

```

```

C *****
C *
C *   SOUS-ROUTINE DEUX: cette sous-routine permet, en appelant la *
C *                               sous-routine CONVOL, de calculer la *
C *                               fonction de renouvellement pour les t *
C *                               plus grands (10000 et 15000 dans notre cas)*
C *                               (la precision est plus faible (h=10) *
C *
C *****

```

```

SUBROUTINE DEUX(REN)
DIMENSION CC(1000000),FN(1000000)
COMMON/PARAM/DIST,ALFA,AL,X1,compte,compte2
EXTERNAL DF
DIST=DIST
ALFA=ALFA
AL=AL
N=0
H=10
X1=X1
NPT=X1/H+1
REN=0
DO 200 I=1,5
CALL CONVOL(DF,H,NPT,FN,CC,N)
REN=REN+FN(NPT)
200 CONTINUE
RETURN

```

END

C -- La fonction DF permet de specifier la fonction de densite ---

```

REAL FUNCTION DF(X)
IMPLICIT REAL(A-H,O-Z)
COMMON/PARAM/DIST,ALFA,AL,X1,compte,compte2
EXTERNAL GAMMA

IF(DIST.EQ.1) GO TO 101
IF(DIST.EQ.2) GO TO 102
IF(DIST.EQ.3) GO TO 103
IF(DIST.EQ.4) GO TO 104

101 DF=(AL*(AL*X)**(ALFA-1))*(EXP(-AL*X)*(1/GAMMA(ALFA)))
GO TO 104

102 DF=ALFA*(AL)*(X**(ALFA-1))*(EXP(-AL*(X**ALFA)))
GO TO 105

103 DF=CONST*EXP(-((X-ALFA)**2)/2*AL*2)
GO TO 105

104 IF (X.GE.0.00000) DF=0.0
DO 150 i=1,n-2
r=donnee2(i+compte2-1)-donnee2(i+1+compte2-1)
q=donnee1(i+compte-1)
d=donnee1(i+1+compte-1)-donnee1(i+compte-1)
IF(X.GT.q) THEN
DF=donnee2(i+compte2-1)-((X-q)/d)*(r)
ELSE
GO TO 151
ENDIF
150 CONTINUE
151 CONTINUE
GO TO 105

```

```

105 CONTINUE
      RETURN
      END

```

```

C *****
C *
C *   Fonction GAMMA: Cette fonction permet de calculer la valeur *
C *                   de la fonction gamma(x) pour x reel      *
C *
C *
C *****

```

```

REAL FUNCTION GAMMA(X)

```

```

PARAMETER (MAX=8, EPS1=1.0E-12, EPS2=1.0E24)

```

```

REAL FACT

```

```

DIMENSION B(MAX)

```

```

DATA      B / 0.035868343, -0.193527818,  0.482199394,
+          -0.756704078,  0.918206857, -0.897056937,
+          0.988205891, -0.577191652      /

```

```

GAMMA=0.0

```

```

FACT=1.0

```

```

Y=X

```

```

IF (Y-1.0) 2,6,1

```

```

2  IF (Y-EPS1) 7,7,8

```

```

7  GAMMA=EPS2

```

```

RETURN

```

```

1  Y=Y-1.0

```

```

IF (Y-1.0) 4,6,3

```

```

3  FACT=Y*FACT
   GOTO 1

8  FACT=1.0/Y

4  DO 5 I=1,8
   GAMMA=Y*(GAMMA+B(I))
5  CONTINUE

6  GAMMA=FACT*(1.0+GAMMA)

   RETURN
END

```

```

C-----
C  SUBROUTINE  SPSPL
C  PURPOSE
C      TO CALCULATE THE CUBIC SPLINE COFICIENTS FOR A SET OF
C      FUNCTION VALUES TABULATED AT A CONSTANT INTERVAL,WHERE THE
C      DERIVATIVE AT THE INITIAL POINT IS KNOWN
C
C  USAGE
C
C      CALL SPSPL(Y,H,NPT,C,D)
C
C  DESCRIPTION OF THE PARAMETERS
C
C  Y  VECTOR OF DIMENSION AT LEAST NPT.Y(1) TO Y(NPT) HOLD VALUES
C      OF FUNCTION IN THE CORRECT SEQUENCE ON ENTRY.
C  H  REAL.INTERVAL OF ARGUMENT FOR WHICH FUNCTION TABULATED.
C  NPT INTEGER.NUMBER OF DATA POINTS.MUST BE.GE.4.
C  C  VECTOR OF DIMENSION AT LEAST NPT.C(1) TO C(NPT) HOLD SPLINE
C      COEFFICIENTS ON EXIT
C  D  VECTOR OF DIMENSION AT LEAST NPT USED AS WORKSPACE.THE KNOWN
C      VALUE OF THE DERIVATIVE AT THE INITIAL POINT MUST BE SET IN

```

```

C          D(1) BEFORE SPSPL IS CALLED.
C
C      METHOD
C
C          SOLVES SIMULTANEOUS EQUATIONS BY GAUSSIAN ELIMINATION AND
C          BACK SUBSTITUTION. INITIAL DERIVATIVE IN D(1) IS INCORPORATED.
C          C(NPT-1) IS CALCULATED DIRECTLY FROM 3-POINT FINITE-
C          DIFFERENCE APPROXIMATION TO 2ND DERIVATIVE.
C-----

```

```

          SUBROUTINE SPSPL(Y,H,NPT,C,D)
          DIMENSION Y(NPT),C(NPT),D(NPT)
          M=NPT
          A=1./H
C
C      RIGHT HAND SIDES OF EQUATIONS INITIALLY PUT IN C,S
C
          C(1)=3.*A*(A*(Y(2)-Y(1))-D(1))
          A=1.5*A*A
          JM1=1
          DO 1 J=3,M
          JM2=JM1
          JM1=J-1
1 C(JM1)=A*(Y(JM2)-2.*Y(JM1)+Y(J))
          C(JM1)=2.*C(JM1)/3.
          C(JM2)=C(JM2)-.25*C(JM1)
C
C      GAUSSIAN ELIMINATION
C
          B=.5
          A=C(1)
          DO 2 J=2,JM2
          D(J)=B
          B=1./(4.-B)
          A=B*(4.*C(J)-A)
2 C(J)=A
C

```

C BUCK SUBSTITUTION

C

I=JM2

DO 3 J=2, JM2

IM1=I-1

A=C(IM1)-A*D(I)

C(IM1)=A

3 I=I-1

C(M)=2.*C(JM1)-C(JM2)

RETURN

END

C-----

C SUBROUTINE CONVOL

C

C PURPOSE

C

C FOR RECURSIVE CALCULATION AND TABULATION OF A SEQUENCE OF
 C CONVOLUTION INTEGRALS $F(N)(T)$ DEFINED RECURSIVELY BY $F(0)(T)$
 C $=1$, AND FOR $N.GE.1$, $F(N)(T)=\text{INTEGRAL FROM } X=0 \text{ TO } X=T \text{ OF}$
 C $F(N-1)(X)*DF(T-X)$ WHERE DF IS A CONTINUOUS PROBABILITY
 C DENSITY FUNCTION WITH CONTINUOUS FIRST DERIVATIVE

C

C USAGE

C

C CALL CONVOL(DF,H,NPT,FN,CC,N)

C

C DESCRIPTION OF PARAMETERS

C

C DF FUNCTION SUBROUTINE OF ONE ARGUMENT TO CALCULATE $DF(X)$,
 C APPEARING IN 'EXTERNAL'STATEMENT IN CALLING PROGRAM
 C H REAL STEP LENGTH, OR INTERVAL OF ARGUMENT FOR TABULATION
 C NPT INTEGER NUMBER OF POINTS. MUST BE $.GT.4$. MAXIMUM $T=H*(NPT-1)$
 C FN VECTOR OF DIMENSION AT LEAST NPT TO HOLD NPT VALUES OF
 C $F(N)(T)$ FROM $T=0$ TO $T=H*(NPT-1)$
 C CC VECTOR OF DIMENSION AT LEAST $8*NPT-6$, USED FOR WORKSPACE AND
 C STORAGE. AT EACH RETURN, $CC(1)$ TO $CC(NPT)$ HOLDS SPLINE COEFFS
 C CORRESPONDING TO VALUES IN $FN(1)$ TO $FN(NPT)$

C N NAME OF INTEGER REGISTER.MUST BE SET TO 0(ZERO) BY USER
 C BEFORE FIRST CALL,AND IS AUTOMATICALLY INCREASED BY1 AT EACH
 C SUBSEQUENT CALL
 C
 C OTHER SUBROUTINES USED
 C
 C CONVOL CALLS SPSPL WHICH MUST BE LOADED WITH IT.
 C
 C COMMENTS
 C
 C PARAMETERS DF,H AND NPT ARE NOT REFERRED TO AFTER
 C FIRST CALL
 C IF DF NEEDS OTHER PARAMETERS THAN THE ARGUMENT,THESE SHOULD
 C BE COMMUNICATED TO IT BY APPROPRIATE VERSION OF THE'COMMON'
 C STATEMENT.PRECISION INCREASES AS H DECREASES.RESONABLE
 C RESULTS CAN BE EXPECTED FOR H IN(.1,.5).SUGGESTED INITIAL
 C COMPROMISE,H=.25.CC(NPT+1) TO CC(4*NPT-3) CAN BE USED
 C BETWEEN SUCCESSIVE CALLS.SPLINE COEFFS IN CC(1) TO CC(NPT)
 C DEFINE A CONTINUOS APPROXIMATION TO F(N)(T).WHEN CONVOL IS
 C USED IN PARALLEL WITH DIFFERENT FUNCTIONS ASSIGNED TO DF,
 C SEPARATE REGISTERS MUST BE ASSIGNED TO FN,CC AND COUNTER N,
 C AND ALL MUST START WITH THEIR COUNTERS SET TO ZERO
 C
 C METHOD
 C
 C F(1)(T) AT THE NPT POINTS IS GOT BY PANEL-BY-PANEL QUADRATURE
 C OF DF USING 5-PT LOBATTO FORMULA.THE 4*NPT-3 WEIGHTED VALUES
 C OF DF AT LOBATTO NODES ARE STORED IN CC(4*NPT-2) TO
 C CC(8*NPT-6).TO GET F(N)(T),N.GT.1,SPLINE REPRESENTATION OF
 C F(N-1)(T) IS USED TO TABULATE ITS VALUES AT THE SAME PTS IN
 C CC(1) TO CC(4*NPT-3).TABULATED VALUES OF F(N)(T) ARE GIVEN AS
 C INNER PRODUCTS OF ELEMENTS OF THESE TWO TABLES TRAVERSED IN
 C OPPOSITE DIRECTIONS.
 C REFERENCE.CLEROUX AND MCCONALOGUE MANAGEMENT SCIENCE LATE,75
 C
 C-----
 C SUBROUTINE CONVOL(DF,H,NPT,FN,CC,N)

```

EXTERNAL DF
DIMENSION HH(4),A(3),FN(NPT),CC(8*NPT-6)
IF(N.EQ.0) GO TO 11
C
C ENTRY POINT FOR N.GE.1. USE SPLINE APPROXIMATION TO CALCULATE
C VALUES OF F(N)(T) AT NODES.STORE IN CC(1) TO CC(4*NPT-3).OMIT ZERO
C VALUES OF FN FROM CALCULATION
C
FK=FN(M)
CK=CC(M)
K=MT
CC(K)=FK
AK=G*FK-F*CK
J=M
DO 10 I=1,M1
J=J-1
CKP=CK
CK=CC(J)
FK=FN(J)
BK=AK
AK=G*FK-F*CK
DO 7 IP=1,3
L=4-IP
K=K-1
7 CC(K)=CK*A(IP)+CKP*A(L)+AK*HH(IP)+BK*HH(L)
K=K-1
10 CC(K)=FK
C
C CALCULATE CONVOLUTIONS FOR F(N+1)(T) AS INNER PRODUCTS
C
K=5
DO 8 J=2,M
L=K+MT
D=0.
DO 9 I=2,K
L=L-1
9 D=CC(I)*CC(L)+D

```

```

      IF(D.LT.0.0) D=0.
      FN(J)=D
8     K=K+4
      CC(MP1)=0.
1     N=N+1
C
C     CALCULATE SPLINE COEFFS INCORPORATING INITIAL DERIV IN CC(MP1)
C
      CALL SPSPL(FN,HH(4),M,CC,CC(MP1))
      RETURN
C
C     INITIAL ENTRY POINT FOR N=0.CALCULATE CONSTS FOR SUBSEQUENT USE
C
11    M=NPT
      M1=M-1
      MP1=M+1
      MT=4*M-3
      M2=M1-1
      HH(4)=H
      F=HH(4)/6.
      G=1./HH(4)
      BK=1.+SQRT(3./7.)
      A(2)=HH(4)*HH(4)/48.
      A(3)=A(2)*BK**3.
      A(1)=A(2)*(2.-BK)**3.
C
C     LOBATTO NODES IN HH(1) TO HH(3).
C
      HH(2)=.5*HH(4)
      HH(3)=BK*HH(2)
      HH(1)=HH(4)-HH(3)
C
C     LOBATTO WEIGHTS IN CC(1) TO CC(4)
C
      CC(1)=49.*HH(4)/180.
      CC(3)=CC(1)
      CC(2)=32.*HH(4)/90.

```

```
CC(4)=HH(4)/20.  
C  
C DERIV OF F(1)(T) AT T=0 SET FOR LATER CALC,N OF SPLINE COEFFS  
C  
CC(MP1)=DF(0.)  
C=CC(4)*CC(MP1)  
C  
C TABULATE DF AT LOBATTO NODES MULTIPLIED BY WEIGHTS.F(1)(T) IS GIVEN  
C AT SAME TIME BY PANEL-BY-PANEL SUMMATION  
C  
K=MT+1  
CC(K)=C  
FN(1)=0.  
X=0.  
DO 2 J=2,M  
DO 3 L=1,4  
XX=X+HH(L)  
D=CC(L)*DF(XX)  
C=C+D  
K=K+1  
3 CC(K)=D  
CC(K)=2.*D  
FN(J)=C  
C=C+D  
2 X=XX  
GO TO 1  
END
```

```

C *****
C * PARTIE 3: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * LE TROISIEME ESTIMATEUR (SOMME DES INDICATRICES) *
C * (estimateur de Frees) *
C * * *
C *****

C PROGRAM APPEL(INPUT,OUTPUT)
PROGRAM APPEL
CE,,RL
REAL donnees,x1,x2,x3,x4,c,somme,totl,minimum,b1,b2,b3,b4,b5
REAL cm1,cm2,cm3,cm4,cm5,er1,er2,er3,er4,er5,bco,cmco,erco
REAL bto,cmto,erto
INTEGER imax,k,compte,tmin
DIMENSION donnees(100000),totl(2687)

C ---Le fichier sim.gamma contient les 1000 echantillons de taille 100 ----

OPEN(2,FILE='sim.gamma')
OPEN(7,FILE='sim.concl3')
READ(2,*)donnees

C --- initialisation des constantes -----
n=100
compte=-(n-1)

C ---- On declare les valeurs theoriques de la fonction de -----
C ---- renouvellement pour differentes valeurs de t ainsi -----
C ---- que le cout minimal et le temps optimal theorique -----

vrai1=1.08807E-12

```

```

vrai2=1.04538E-06
vrai3=3.58845E-03
vrai4=2.03078E-02
vrai5=6.51914E-02
vrai6=0.263032
vrai7=6.70089e-01
vrai8=1.20255
cvrai=0.0305826
tvrai=3979

```

C --- La boucle 3 repete le programme 1000 fois (pour les 1000 echanti.) ---

```

DO 3 z=1,1000
  compte=compte+n
  somme=0
  c=0
  x1=0
  minimum=100000

```

C --- On specifie les valeurs de t pour lesquelles on estime M(t) ----

```

DO 2 a=1,2618
  IF (a.LE.15) THEN
    t=200*a
  ELSE
    IF (a.LE.2615) THEN
      t=3000+a-15
    ELSE
      IF (a.EQ.2616) t=7000
      IF (a.EQ.2617) t=10000
      IF (a.EQ.2618) t=15000
    ENDIF
  ENDIF

```

C ---- On specifie la valeur de k, qui est la valeur de m (le nombre de --
C ---- convolutions a calculer (le programme va jusqu'a 5) ----

```

        totl(a)=0
DO 5 k=1,5

imax=0
DO 1 i=compte,compte+n-1
    IF (donnees(i) .LE. t) THEN
        imax=imax+1
    ENDIF
1 CONTINUE

DO 10 i=1,min(imax,n-(k-1))
    x1=donnees(i+compte)
    somme=x1
    IF (k .GE. 2) THEN
        DO 20 j=i+1,min(imax,n-(k-2))
            x2=donnees(j+compte)
            somme=x1+x2
            IF (k .GE. 3) THEN

                DO 30 l=j+1,min(imax,n-(k-3))
                    x3=donnees(l+compte)
                    somme=x1+x2+x3

                IF (k .GE. 4) THEN

                    DO 40 m=l+1,min(imax,n-(k-4))
                        x4=donnees(m+compte)
                        somme=x1+x2+x3+x4

                    IF (k .GE. 5) THEN

                        DO 50 o=m+1,min(imax,n-(k-5))
                            x5=donnees(o+compte)
                            somme=x1+x2+x3+x4+x5
                            IF (somme .LE. t) THEN
                                c=c+1

```

```

                                                                    ENDIF
50          CONTINUE
          ELSE
            IF (somme .LE. t) THEN
              c=c+1
            ENDIF
          ENDIF

40          CONTINUE
          ELSE
            IF (somme .LE. t) THEN
              c=c+1
            ENDIF
          ENDIF

30          CONTINUE
          ELSE
            IF (somme .LE. t) THEN
              c=c+1
            ENDIF
          ENDIF

20          CONTINUE
          ELSE
            IF (somme .LE. t) THEN
              c=c+1
            ENDIF
          ENDIF

10  CONTINUE
```

C -- on obtient le nombre de combinaisons pour lesquelles la somme est $\leq t$ --

```

IF (k.EQ.1) THEN
  c=c/100
ELSE
  IF (k.EQ.2) THEN
    c=c/4950
  ELSE
    IF (k.EQ.3) THEN
      c=c/161700
    ELSE
      IF (k.EQ.4) THEN
        c=c/3921225
      ELSE
        IF (k.EQ.5) THEN
          c=c/75287520
        ENDIF
      ENDIF
    ENDIF
  ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
ENDIF
totl(a+compte-1)=totl(a+compte-1)+c

```

c=0

5 CONTINUE

C -- On calcule la fonction de cout, le cout minimum et le temps optimal --

```

couts=(1100*totl(a)+100)/t
IF (couts.LE.minimum) THEN
  minimum=couts
  tmin=t
ENDIF

```

2 CONTINUE

C -- On calcule le biais, le carre moyen de l'erreur et l'erreur relative --
 C -- Les valeurs specifiees sont les vraies valeurs de la fnct de renouvel. --

```

b0=b0+(totl(1)-(vrai0))
b1=b1+(totl(5)-(vrai1))
b2=b2+(totl(15)-(vrai2))
b3=b3+(totl(1015)-(vrai2))
b4=b4+(totl(2015)-(vrai4))
b5=b5+(totl(2616)-(vrai5))
b6=b6+(totl(2617)-(vrai6))
b7=b7+(totl(2618)-(vrai7))

cm0=cm0+(totl(1)-vrai0)**2
cm1=cm1+(totl(5)-(vrai1))**2
cm2=cm2+(totl(15)-(vrai2))**2
cm3=cm3+(totl(1015)-(vrai2))**2
cm4=cm4+(totl(2015)-(vrai4))**2
cm5=cm5+(totl(2616)-(vrai5))**2
cm6=cm6+(totl(2617)-(vrai6))**2
cm7=cm7+(totl(2618)-(vrai7))**2

er0=er0+(max(totl(1),vrai0)-min(totl(1),vrai0))
er1=er1+(max(totl(5),(vrai1))-min(totl(5),(vrai1)))
er2=er2+(max(totl(15),vrai2)-min(totl(15),vrai2))
er3=er3+(max(totl(1015),vrai2)-min(totl(1015),vrai2))
er4=er4+(max(totl(2015),vrai4)-min(totl(2015),vrai4))
er5=er5+(max(totl(2616),vrai5)-min(totl(2616),vrai5))
er6=er6+(max(totl(2617),vrai6)-min(totl(2617),vrai6))
er7=er7+(max(totl(2618),vrai7)-min(totl(2618),vrai7))

bco=bco+(minimum-cvrai)
cmco=cmco+(minimum-cvrai)**2
erco=erco+(max(minimum,cvrai)-min(minimum,cvrai))

bto=bto+(tmin-tvrai)

```

```
cmto=cmto+(tmin-tvrai)**2
erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))
```

```
3 CONTINUE
```

```
C -- On conserve les resultats dans un fichier ----
```

```
WRITE(7,*)'t=200:',b0/1000,cm0/1000,er0/(1000*(vrai0))
WRITE(7,*)'t=1000:',b1/1000,cm1/1000,er1/(1000*(vrai1))
WRITE(7,*)'t=3000:',b2/1000,cm2/1000,er2/(1000*(vrai2))
WRITE(7,*)'t=4000:',b3/1000,cm3/1000,er3/(1000*(vrai2))
WRITE(7,*)'t=5000:',b4/1000,cm4/1000,er4/(1000*(vrai4))
WRITE(7,*)'t=7000:',b5/1000,cm5/1000,er5/(1000*(vrai5))
WRITE(7,*)'t=10000:',b6/1000,cm6/1000,er6/(1000*(vrai6))
WRITE(7,*)'t=15000:',b7/1000,cm7/1000,er7/(1000*(vrai7))
WRITE(7,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
WRITE(7,*)'le temps:',bto/1000,cmto/1000,erto/(1000*tvrai)
END
```

```

C *****
C * PARTIE 4: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * LE QUATRIEME ESTIMATEUR (ESTIMATEUR EMPIRIQUE) *
C * * *
C *****

PROGRAM APPEL
REAL donnees,FN,HN,tempo,n,couts,minimum,b1,b2,b3,b4,b5
REAL cm1,cm2,cm3,cm4,cm5,er1,er2,er3,er4,er5,bco,cmco,erco
REAL bto,cmto,erto
DIMENSION FN(27913),HN(27913)
INTEGER donnees2,j,k,tmin,compte
DIMENSION donnees(100000),donnees2(100000)

C -- Le fichier sim.gamma.int contient les 1000 echantillons de ----
C -- taille 100 (les donnees sont arrondies a l' unite pres) ----

OPEN(2,FILE='sim.gamma.int')
OPEN(7,FILE='sim.concl4')
READ(2,*)donnees

C ---- initialisation des constantes -----

n=100
compte=-(n-1)

C ---- On declare les valeurs theoriques de la fonction de -----
C ---- renouvellement pour differentes valeurs de t ainsi -----
C ---- que le cout minimal et le temps optimal theorique -----

vrai1=1.08807E-12
vrai2=1.04538E-06

```

```

vrai3=3.58845E-03
vrai4=2.03078E-02
vrai5=6.51914E-02
vrai6=0.263032
vrai7=6.70089e-01
vrai8=1.20255
cvrai=0.0305826
tvrai=3979

```

C --- on repete le programme 1000 fois pour les 1000 echantillons ---

```

DO 1 z=1,1000
  compte=compte+n
  minimum=100000
  k=0
  j=1
  DO 15 i=1,donnees(compte+n-1)
    IF (i.LT.donnees(j+compte-1)) THEN
      FN(i)=k/n
    ELSE
      k=k+1
      j=j+1
      FN(i)=k/n
    ENDIF
  15 CONTINUE

  DO 20 i=1,donnees(compte+n-1)
    tempo=0
    somme=0

    DO 30 k=1,i
      tempo=HN(i-k)*(FN(k)-FN(k-1))
      somme=tempo+somme
    30 CONTINUE
    HN(i)=FN(i)+somme

```

20 CONTINUE

```

DO 100 t=1,donnees(compte+n-1)
  couts=(1100*HN(t)+100)/t
  IF (couts.LE.minimum) THEN
    minimum=couts
    tmin=t
  ENDIF

```

100 CONTINUE

C -- On calcule le biais, le carre moyen de l'erreur et l'erreur relative --
C -- Les valeurs specifiees sont les vraies valeurs de la fnct de renouvel. --

```

b0=b0+(HN(200)-(vrai0))
b1=b1+(HN(1000)-(vrai1))
b2=b2+(HN(3000)-(vrai2))
b3=b3+(HN(4000)-(vrai2))
b4=b4+(HN(5000)-(vrai4))
b5=b5+(HN(7000)-(vrai5))
b6=b6+(HN(10000)-(vrai6))
b7=b7+(HN(15000)-(vrai7))

```

```

cm0=cm0+(HN(200)-vrai0)**2
cm1=cm1+(HN(1000)-(vrai1))**2
cm2=cm2+(HN(3000)-(vrai2))**2
cm3=cm3+(HN(4000)-(vrai2))**2
cm4=cm4+(HN(5000)-(vrai4))**2
cm5=cm5+(HN(7000)-(vrai5))**2
cm6=cm6+(HN(10000)-(vrai6))**2
cm7=cm7+(HN(15000)-(vrai7))**2

```

```

er0=er0+(max(HN(200),vrai0)-min(HN(200),vrai0))
er1=er1+(max(HN(1000),(vrai1))-min(HN(1000),(vrai1)))
er2=er2+(max(HN(3000),vrai2)-min(HN(3000),vrai2))
er3=er3+(max(HN(4000),vrai2)-min(HN(4000),vrai2))
er4=er4+(max(HN(5000),vrai4)-min(HN(5000),vrai4))
er5=er5+(max(HN(7000),vrai5)-min(HN(7000),vrai5))

```

```
er6=er6+(max(HN(10000),vrai6)-min(HN(10000),vrai6))
er7=er7+(max(HN(15000),vrai7)-min(HN(15000),vrai7))
```

```
bco=bco+(minimum-cvrai)
cmco=cmco+(minimum-cvrai)**2
erco=erco+(max(minimum,cvrai)-min(minimum,cvrai))
```

```
bto=bto+(tmin-tvrai)
cmto=cmto+(tmin-tvrai)**2
erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))
```

```
1 CONTINUE
```

```
C -- On conserve les resultats dans un fichier ----
```

```
WRITE(7,*)'t=200:',b0/1000,cm0/1000,er0/(1000*(vrai0))
WRITE(7,*)'t=1000:',b1/1000,cm1/1000,er1/(1000*(vrai1))
WRITE(7,*)'t=3000:',b2/1000,cm2/1000,er2/(1000*(vrai2))
WRITE(7,*)'t=4000:',b3/1000,cm3/1000,er3/(1000*(vrai2))
WRITE(7,*)'t=5000:',b4/1000,cm4/1000,er4/(1000*(vrai4))
WRITE(7,*)'t=7000:',b5/1000,cm5/1000,er5/(1000*(vrai5))
WRITE(7,*)'t=10000:',b6/1000,cm6/1000,er6/(1000*(vrai6))
WRITE(7,*)'t=15000:',b7/1000,cm7/1000,er7/(1000*(vrai7))
WRITE(7,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
WRITE(7,*)'le temps:',bto/1000,cmto/1000,erto/(1000*tvrai)
```

```
END
```

```

C *****
C * PARTIE 5: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * LA FONCTION DE DENSITE ECHANTILLONNALE POUR LES *
C * 1000 ECHANTILLONS. ELLES SONT UTILISEES POUR *
C * LE CALCUL DU CINQUIEME ESTIMATEUR *
C * * *
C *****

```

```

PROGRAM APPEL
IMPLICIT REAL(A-H,O-Z)

```

```

INTEGER i,iu,iv
DIMENSION donnee(100000),aaint(100),aami(100),ptfn(100)

```

```

C --- Les donnees sont dans le fichier sim.gamma ----

```

```

OPEN(4,FILE='sim.g.cinq.df2')
OPEN(2,FILE='sim.gamma')
READ(2,*)donnee
somme=0

```

```

n=100.
compte=-(n-1)
aainte=0

```

```

C --- On repete 1000 fois pour les 1000 echantillons ----

```

```

DO 200 z=1,1000
  compte=compte+n

```

```

aainte=0

```

```
DO 150 i=1,n-1
rn=0
aami(i)=1000000
DO 151 iv=i+1,n
aama=0
DO 152 iu=1,i
    somme=0
    DO 153 j=iu,iv-1
        x=(n-j)*(donnee(compte-1+j+1)-donnee(compte-1+j))
        somme=somme+x
        rn=(iv-iu)*1/somme
153     CONTINUE
    IF (rn.GT.aama) THEN
        aama=rn
    ENDIF
152 CONTINUE
    IF (aama.LT.aami(i)) THEN
        aami(i)=aama
    ENDIF
151 CONTINUE
    aaint(i)=aami(i-1)*(donnee(compte-1+i)-donnee(compte-1+i-1))
    aainte=aainte+aaint(i)
    expo=exp(-1*(aainte))
    grandfn=-1*(expo-1)
    ptfn(i)=aami(i)*(1-grandfn)
150 CONTINUE
    DO 154 i=1,n-1
        r=ptfn(i)-ptfn(i+1)
        q=donnee(compte-1+i)
        d=donnee(compte-1+i+1)-donnee(compte-1+i)
        WRITE(4,*)ptfn(i)
154 CONTINUE
200 CONTINUE
END
```

```

C *****
C * PARTIE 6: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * L'estimation du cout minimal (et du temps optimal) *
C * dans le cas d'un remplacement préventif avec repa- *
C * ration minimale a la panne. On estime le taux de *
C * panne a l'aide de l'estimateur paramétrique de la *
C * fonction de repartition. *
C * *
C *****

C PROGRAM APPEL(INPUT,OUTPUT)
PROGRAM APPEL
IMPLICIT REAL(A-H,O-Z)
DIMENSION alpha(1000),ala(1000)
COMMON/PARAM/DIST,ALFA,AL,X1,compte,compte2
DIMENSION Y1(20000),RE(20000),aint(20000),couts(20000)

C ---- Les fichiers sim.alpha et sim.lamda contiennent les 1000 ---
C ---- estimations des parametres des 1000 dist. echantillonales ---

OPEN(2,FILE='r100.g.couts2')
OPEN(3,FILE='r100.g.concl2')
OPEN(4,FILE='sim.g.alpha')
OPEN(5,FILE='sim.g.lamda')
READ(4,*)alpha
READ(5,*)ala

C DETERMINATION DE LA DISTRIBUTION:
C 1=GAMMA
C 2=WEIBULL
C 3=NORMALE TRONQUEE

```

```

DIST=1

C -----  initialisation des constantes  -----
      n=100.
      compte=-(n-1)
      compte2=-(n-2)

C ----  On declare les valeurs theoriques de la fonction de  -----
C ----  renouvellement pour differentes valeurs de t ainsi  -----
C ----  que le cout minimal et le temps optimal theorique  -----

      r1vrai=0.
      r2vrai=7.71671E-06
      r3vrai=2.90219E-05
      r4vrai=6.73383E-05
      r5vrai=1.73658E-04
      r6vrai=3.30925E-04
      r7vrai=5.10693E-04
      cvrai=1.53166E-02
      tvrai=9483.00

C ---  La boucle 30 permet la repetition du programme 1000 fois  -----
C ---                               ( pour les 1000 echantillons)  -----

      DO 30 I=1,1000
          ALFA=alpha(I)
          AL=ala(I)

      aamin=1000000
      compte=compte+n
      compte2=compte2+n-1

C ---  On calcule a l'interieur de cette boucle le cout minimal  ---
C ---  et le temps optimal.  ---

```

```

DO 20 J=700,1500
X1=10*J
CALL DEUX(RE(J))
Y1(J)=-log(1-RE(J))
couts(J)=(50*Y1(J)+100)/X1
  IF (couts(J).LE.aamin) THEN
    aamin=couts(J)
    tmin=J*10
  ENDIF
  IF (aamin.EQ.couts(J-50)) THEN
    GO TO 24
  ENDIF
20 CONTINUE
24 CONTINUE
  aamin=1000000
  DO 27 d=tmin-10,tmin+10
    X1=d
    CALL UN(RE(d))
    Y1(d)=-log(1-RE(d))
    couts(d)=(50*Y1(d)+100)/d
    IF (couts(d).LE.aamin) THEN
      aamin=couts(d)
      tmin=d
    ENDIF
27 CONTINUE

```

C -- On calcule le biais, le carre moyen de l'erreur et l'erreur relative --

```

bco=bco+(aamin-cvrai)
cmco=cmco+(aamin-cvrai)**2
erco=erco+(max(aamin,cvrai)-min(aamin,cvrai))

bto=bto+(tmin-tvrai)
cmto=cmto+(tmin-tvrai)**2
erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))

```

30 CONTINUE

```
WRITE(3,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
WRITE(3,*)'le temps:',bto/1000,cmt0/1000,erto/(1000*tvrai)
END
```

NOTE:

Les fonctions et sous-routines appelées par ce programme sont les mêmes que dans la partie 2 de cette annexe. Nous ne les reproduirons donc pas ici. (L'unique différence est que l'on ne calcule que la première convolution (c.a.d la fonction de répartition). Les boucles 100 et 200 des sous-routines se terminent donc à I=1.

```

C *****
C * PARTIE 7: *
C * ----- *
C * Cette partie calcule: *
C * * *
C * L'estimation du cout minimal (et du temps optimal) *
C * dans le cas d'un remplacement préventif avec repa- *
C * ration minimale a la panne. On estime le taux de *
C * panne a l'aide de l'estimateur non-paramétrique *
C * de la fonction de repartition. *
C * * *
C *****

C PROGRAM APPEL(INPUT,OUTPUT)
PROGRAM APPEL
REAL donnees,x1,x2,x3,x4,c,somme,totl,minimum,b1,b2,b3,b4,b5
REAL cm1,cm2,cm3,cm4,cm5,er1,er2,er3,er4,er5,bco,cmco,erco
REAL bto,cmto,erto,diff
INTEGER imax,k,compte,tmin
DIMENSION donnees(100000),totl(2687),R(1000),t(1000),ptr(1000)

C ---- Le fichier sim.gamma contient les 1000 echantillons de ----
C ---- taille 100 ----

OPEN(2,FILE='sim.gamma')
OPEN(3,FILE='r100.g.res3')
OPEN(5,FILE='r100.g.couts3')
OPEN(7,FILE='r100.g.concl3')
READ(2,*)donnees

C ----- initialisation des constantes -----

n=100.
compte=-(n-1)

C ---- On declare les valeurs theoriques de la fonction de ----

```

```
C ---- renouvellement pour differentes valeurs de t ainsi ----
C ---- que le cout minimal et le temps optimal theorique ----
```

```
r1vrai=0.
r2vrai=7.71671E-06
r3vrai=2.90219E-05
r4vrai=6.73383E-05
r5vrai=1.73658E-04
r6vrai=3.30925E-04
r7vrai=5.10693E-04
cvrai=1.53166E-02
tvrai=9483.00
```

```
C --- La boucle 3 permet la repetition du programme 1000 fois ----
C --- ( pour les 1000 echantillons) ----
```

```
DO 3 z=1,1000
  compte=compte+n
  somme=0
  c=0
  x1=0
  minimum=100000
```

```
C ----- On doit specifier les valeur de t souhaitees -----
```

```
DO 2 a=1,n
  t(a)=donnees(a+compte-1)
```

```
DO 1 i=compte,compte+n-1
  IF (donnees(i) .LE. t(a)) THEN
    c=c+1
  ENDIF
```

```
1 CONTINUE
```

```
c=c/n
ENDIF
```

```

totl(a)=c

c=0

R(a)=-log(1-totl(a))
ptr(a)=(R(a)-R(a-1))/(t(a)-t(a-1))

couts=(50*R(a)+100)/t(a)
IF (couts.LE.minimum) THEN
    minimum=couts
    tmin=t(a)
    ttt=t(a+1)
    aaa=a
ENDIF

2  CONTINUE

bco=bco+(minimum-cvrai)
cmco=cmco+(minimum-cvrai)**2
erco=erco+(max(minimum,cvrai)-min(minimum,cvrai))

bto=bto+(tmin-tvrai)
cmto=cmto+(tmin-tvrai)**2
erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))

WRITE(5,*)tmin,minimum

3  CONTINUE
WRITE(7,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
WRITE(7,*)'le temps:',bto/1000,cmto/1000,erto/(1000*tvrai)
END

```

```

C *****
C *   PARTIE 8: *
C *   ----- *
C *   Cette partie calcule: *
C * *
C *   L'estimation du cout minimal (et du temps optimal) *
C *   dans le cas d'un remplacement préventif avec repa- *
C *   ration minimale a la panne. On estime le taux de *
C *   panne a l'aide de son estimateur dans le cas IFR. *
C * *
C *****

C   PROGRAM APPEL(INPUT,OUTPUT)
C   PROGRAM APPEL
CE, ,RL
C   IMPLICIT REAL(A-H,O-Z)

C   INTEGER i,iu,iv
C   DIMENSION donnee(100000),aaint(200),aami(200),R(210),couts(200)

C   ---- Le fichier sim.gamma contient les 1000 echantillons de   ---
C   ---- taille 100   ---

C   OPEN(2,FILE='sim.gamma')
C   OPEN(7,FILE='r100.g.concl5')
C   OPEN(3,FILE='r100.g.couts5')
C   READ(2,*)donnee

C   ----- initialisation des constantes -----

C   n=100.
C   compte=-(n-1)
C   somme=0

C   ---- On declare les valeurs theoriques de la fonction de   -----
C   ---- renouvellement pour differentes valeurs de t ainsi   -----

```

```

C ---- que le cout minimal et le temps optimal theorique ----

r1vrai=0.
r2vrai=7.71671E-06
r3vrai=2.90219E-05
r4vrai=6.73383E-05
r5vrai=1.73658E-04
r6vrai=3.30925E-04
r7vrai=5.10693E-04
cvrai=1.53166E-02
tvrai=9483.00

C --- La boucle 200 permet la repetition du programme 1000 fois ----
C --- ( pour les 1000 echantillons) ----

DO 200 z=1,1000
    compte=compte+n
    cmin=1000000

DO 150 i=1,n-1
    rn=0
    aami(i)=1000000
    DO 151 iv=i+1,n
        aama=0
        DO 152 iu=1,i
            somme=0
            DO 153 j=iu,iv-1
                x=(n-j)*(donnee(compte-1+j+1)-donnee(compte-1+j))
                somme=somme+x
                rn=(iv-iu)*1/somme
            ENDIF
        ENDIF
    ENDIF
    CONTINUE
153
    IF (rn.GT.aama) THEN
        aama=rn
    ENDIF
152
    CONTINUE

```

```

        IF (aama.LT.aami(i)) THEN
            aami(i)=aama
        ENDIF

151    CONTINUE

C --- On calcule l'integrale du taux de panne et le cout minimal ---

        aaint(i)=aami(i-1)*(donnee(compte-1+i)-donnee(compte-1+i-1))
        R(i)=R(i-1)+aaint(i)
        couts(i)=(50*R(i)+100)/donnee(compte-1+i)
        IF (couts(i).LE.cmin) THEN
            tmin=donnee(compte-1+i)
            cmin=couts(i)
        ENDIF

150    CONTINUE

        WRITE(3,*)tmin,cmin

        bco=bco+(cmin-cvrai)
        cmco=cmco+(cmin-cvrai)**2
        erco=erco+(max(cmin,cvrai)-min(cmin,cvrai))

        bto=bto+(tmin-tvrai)
        cmto=cmto+(tmin-tvrai)**2
        erto=erto+(max(tmin,tvrai)-min(tmin,tvrai))

200    CONTINUE
        WRITE(7,*)'les couts:',bco/1000,cmco/1000,erco/(1000*cvrai)
        WRITE(7,*)'le temps:',bto/1000,cmto/1000,erto/(1000*tvrai)
        END

```

BIBLIOGRAPHIE

- [1] ARUNKUMAR, *Non-parametric age replacement policy*, Sankhya, 1972, pp.251-256.
- [2] BAKER, R. D., *A nonparametric estimator of the renewal function*, Computers And Operations Research, vol 20, 1993, pp.167-178.
- [3] BEAUCAIRE, C., *Contribution au problème de remplacement préventif d'équipements*, Rapport de stage, Université de Montréal, 1987.
- [4] BERG, M., *A proof of optimality for age replacement policies*, Journal of Applied Probability, vol. 13, 1976, pp.751-759
- [5] BLISCHKE, W. R. ET SCHEUER, E. M., *Application of renewal theory in analysis of the free-replacement warranty*, Naval Research Logistics Quarterly, vol. 28, 1981, pp.193-205.
- [6] CLÉROUX, R. , *Notes du cours: Fiabilité, théorie et applications*, Université de Montréal.
- [7] CLÉROUX, R. ET MCCONALOGUE, D. J., *A numerical algorithm for recursively-defined convolution integrals involving distribution functions*, Management Science, vol. 22, no. 10, 1976, pp.1138-1146.
- [8] FREES, E. W., *Warranty analysis and renewal function estimation*, Naval Research Logistics Quarterly, vol. 33, 1986, pp.361-372.
- [9] FREES, E. W., *Nonparametric renewal function estimation*, The Annals of Statistics, vol. 14, no. 4, 1986, pp.1366-1378.
- [10] GRENANDER, U, *On the theory of mortality measurements, part II*, Skand. Aktuarietidskr, vol. 39, 1956, pp.125-153
- [11] JURAN, J. ET GRZYNA, F., *Quality planning and analysis*, McGraw-Hill, New York, 1980.
- [12] LÉGER, C. ET CLÉROUX, R., *Non-parametric age bootstrap confidence intervals for the optimal cost*, Operations Research, 1992, pp.1062-1073.
- [13] LIN, B. S., O'CONNOR, C. ET SCHNEIDER, H., *Comparison of nonparametric estimators for the renewal function*, Applied Statistics, vol. 39, no. 1, 1990, pp.55-61.

- [14] MARSHALL, A. W. ET PROSCHAN, F., *Maximum likelihood estimation for distribution with monotone failure rate*, The Annals of Mathematical Statistics, vol. 36, 1965, pp.69-77.
- [15] PRAKASA RAO, B.L.S., *Estimation for distribution with monotone failure rate*, The Annals of Mathematical Statistics, vol. 41, 1970, pp.507-519.
- [16] SAHIN, I., *On approximating the renewal function with its linear asymptot: how large is large enough?*, Operations Research Letters, vol 4, no. 5, 1986, pp.207-210.
- [17] SCHEAFFER, I., *On consistent estimation of age replacement interval*, Technometrics, 1976, pp.213-219.
- [18] SCHNEIDER, H., LIN, B. S. ET TANG, K., *Comparison of renewal function estimators for censored data.*, Journal of Statistical Computation and Simulation, vol 42, 1992, pp.125-135.
- [19] SCHNEIDER, H ET WEISSFELD, L, *Estimation in linear models with censored data*, Biometrika, vol. 73, 1986, pp.741-745.
- [20] SUSARLA, V., TSAI, W. Y. ET VAN RYZIN, J., *A Buckley-James-type estimator for the mean with censored data*, Biometrika, vol. 71, no. 3, 1984, pp.624-625.
- [21] YANG, G., *A renewal-process approach to continuous sampling plans*, Technometrics, vol 25, no. 1, 1983, pp.59-67.