

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique  
Université Abderrahmane Mira de Béjaia



Faculté des Sciences Exactes  
Département des Mathématiques

**MÉMOIRE DE FIN D'ÉTUDE**

En vue de l'obtention du Diplôme de Master en Mathématiques

Option : Statistique et Analyse Décisionnelle

**Thème**

---

**Estimation non paramétrique des paramètres de la fiabilité**

---

Réalisé par :

*M<sup>elle</sup>* MOKRAOUI Dalila

*M<sup>elle</sup>* TOULOUM Sihem

Soutenu devant le jury composé de :

Présidente : *M<sup>me</sup>* HOCINE Safia M.A.A Université de Béjaia.

Encadreur : *M<sup>me</sup>* LAGHA Karima M.C.B Université de Béjaia.

Examinatrice : *M<sup>elle</sup>* SAADI Nora M.A.A Université de Béjaia.

Promotion : 2015/2016

# Remerciements

*Tout d'abord, nous remercions DIEU le tout-puissant qui nous a donné la force et le savoir afin d'accomplir ce travail.*

*Un grand merci pour nos familles, surtout nos parents qui nous ont épaulés, soutenus et suivis tout au long de ce projet.*

*A nos chères amis qui ont toujours été présents et fidèles.*

*A notre encadreur M<sup>me</sup> **LAGHA** pour son encouragement et son suivi attentif pour la réalisation de ce travail.*

*Au membres de l'unité de recherche LaMOS particulièrement M<sup>elle</sup> **ZIANE Yasmina** pour leurs aides et leurs disponibilités pendant toute la réalisation de notre travail.*

*Nous tenons aussi à remercier également tous les membres de Jury pour avoir accepté d'évaluer notre travail.*

*Enfin, pour toute personne qui a contribué, de près ou de loin, à l'élaboration de ce mémoire. Veuillez bien trouver ici l'expression de nos sincères remerciements.*

# *Dédicaces*

*Ce modeste travail est dédié :*

*A nos chers parents.*

*A nos frères et sœurs.*

*A nos familles.*

*A nos amis et collègues et tous ceux qui nous ont aidés.*

*Dalila et Sihem*

# Table des matières

Liste des figures	vi
Liste des tableaux	vii
Introduction générale	1
<b>1 Concepts généraux de la fiabilité</b>	<b>4</b>
1.1 Introduction	4
1.2 Objectifs et intérêts de la fiabilité	5
1.3 Défaillance d'un système	5
1.3.1 Paramètres d'identification des défaillances	5
1.3.2 Densité de défaillance	8
1.3.3 Fonction de défaillance cumulée	8
1.4 Fiabilité et probabilité de survie	9
1.4.1 Définition du concept de fiabilité	9
1.4.2 Fonction de fiabilité ou fonction de survie	10
1.5 Indices de fiabilité	11
1.5.1 Moyenne des temps jusqu'à la première défaillance (MTTF)	12
1.5.2 Taux de défaillance $\lambda$	12
1.5.3 Fonction de hasard (Fonction de risque)	13
1.5.4 La courbe en baignoire (Bathtub Curve)	13
1.5.5 Moyenne des temps entre les défaillances (MTBF)	15
1.6 Loi de survie	15
1.7 Fiabilité des systèmes	16
1.7.1 Système en série	16
1.7.2 Système en parallèle	17
1.7.3 Système $K$ parmi $n$	17
1.8 Principales lois de probabilité utilisées en fiabilité	18
1.8.1 La loi exponentielle $exp(\lambda)$	18
1.8.2 La loi de weibull $\mathcal{W}(\gamma, \beta, \eta)$	20
1.8.3 Loi Gamma $\mathcal{G}(\alpha, \beta)$	21
1.8.4 Loi log-normale $Log\mathcal{N}(\mu, \sigma^2)$	22
1.9 Conclusion	22

<b>2</b>	<b>Estimation non paramétrique de la densité par la méthode du noyau</b>	<b>23</b>
2.1	Introduction . . . . .	23
2.2	Critères d'erreur et définitions . . . . .	24
2.2.1	Les différents critères d'erreur . . . . .	24
2.2.2	Quelques définitions . . . . .	24
2.3	Estimateur de Parzen-Rosenblatt . . . . .	25
2.4	Propriétés de l'estimateur de Parzen-Rosenblatt . . . . .	26
2.4.1	Propriétés statistiques de l'estimateur à noyau . . . . .	26
2.4.2	Propriétés asymptotiques de l'estimateur à noyau . . . . .	29
2.4.2.1	Comportement asymptotique du biais . . . . .	29
2.4.2.2	Comportement asymptotique de la variance . . . . .	29
2.4.2.3	Convergence en moyenne quadratique . . . . .	30
2.4.2.4	Convergence en moyenne quadratique intégrée . . . . .	30
2.4.2.5	Convergence uniforme en probabilité . . . . .	30
2.4.2.6	Convergence uniforme presque complète . . . . .	31
2.4.2.7	Convergence en loi . . . . .	31
2.4.2.8	Convergence $\mathbb{L}^1$ presque complète . . . . .	32
2.5	Le Choix pratique de $(K, h)$ . . . . .	33
2.5.1	Choix du noyau $K$ . . . . .	33
2.5.1.1	Noyaux symétriques . . . . .	33
2.5.1.2	Noyaux asymétriques . . . . .	34
2.5.2	Choix du paramètre de lissage $h$ . . . . .	36
2.5.2.1	Méthodes plug-in . . . . .	37
2.5.2.2	Méthode Validation Croisée (Cross-Validation CV) . . . . .	38
2.6	Estimation de la densité de probabilité basée sur un nombre aléatoire d'observations . . . . .	41
2.6.1	Propriétés statistiques de l'estimateur . . . . .	42
2.6.2	Propriétés asymptotiques de l'estimateur . . . . .	44
2.7	Conclusion . . . . .	45
<b>3</b>	<b>Estimation des paramètres de la fiabilité et simulation</b>	<b>47</b>
3.1	Introduction . . . . .	47
3.2	Estimation empirique de la fonction de répartition . . . . .	47
3.2.1	Propriétés statistiques de l'estimateur $\hat{F}_n(x)$ . . . . .	48
3.2.2	Propriétés asymptotiques de l'estimateur $\hat{F}_n(x)$ . . . . .	49
3.3	Estimation de la fiabilité . . . . .	50
3.4	Estimation du taux de défaillance . . . . .	50
3.5	Estimation dans le cas d'une taille d'échantillon aléatoire . . . . .	51
3.5.1	Fonction de répartition . . . . .	51
3.5.2	Fonction de fiabilité . . . . .	51
3.5.3	Fonction du taux de défaillance . . . . .	52
3.6	Simulation . . . . .	52
3.6.1	Plan de simulation . . . . .	53
3.6.2	Algorithme de simulation . . . . .	53
3.6.3	Résultats de simulation : cas où la taille de l'échantillon est fixe . . . . .	54
3.6.4	Résultats de simulation : cas où la taille de l'échantillon est aléatoire . . . . .	61

**Table des matières** **iii**

---

3.6.5 Étude comparative . . . . .	69
3.7 Conclusion . . . . .	75
<b>Conclusion générale</b>	<b>75</b>
<b>Bibliographie</b>	<b>77</b>

# Table des figures

1.1	Fonction de fiabilité . . . . .	11
1.2	Courbe en baignoire . . . . .	14
1.3	Configuration série. . . . .	17
1.4	Configuration parallèle. . . . .	17
1.5	Propriété sans mémoire de la loi exponentielle de paramètre $\lambda$ . . . . .	19
2.1	Noyau gaussien pour $h = 1.2$ , $x = 1$ , $x = 3$ et $x = 5$ . . . . .	32
3.1	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n=50$ . . . . .	55
3.2	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n=200$ . . . . .	55
3.3	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n=2000$ . . . . .	55
3.4	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n=50$ . . . . .	56
3.5	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n=200$ . . . . .	56
3.6	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n=2000$ . . . . .	57
3.7	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n=50$ . . . . .	57
3.8	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n=200$ . . . . .	58
3.9	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n=2000$ . . . . .	58
3.10	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n=50$ . . . . .	59

3.11	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n=200$ . . . . .	59
3.12	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n=2000$ . . . . .	59
3.13	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n=50$ . . . . .	60
3.14	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n=200$ . . . . .	60
3.15	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n=2000$ . . . . .	61
3.16	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n = 27$ sur l'intervalle $[0, 50]$ . . . . .	61
3.17	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n = 105$ sur l'intervalle $[0, 200]$ . . . . .	62
3.18	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour $n = 1017$ sur l'intervalle $[0, 2000]$ . . . . .	62
3.19	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n = 23$ sur l'intervalle $[0, 50]$ . . . . .	63
3.20	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n = 100$ sur l'intervalle $[0, 200]$ . . . . .	63
3.21	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour $n = 1018$ sur l'intervalle $[0, 2000]$ . . . . .	64
3.22	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n = 30$ sur l'intervalle $[0, 50]$ . . . . .	64
3.23	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n = 107$ sur l'intervalle $[0, 200]$ . . . . .	65
3.24	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour $n = 1017$ sur l'intervalle $[0, 2000]$ . . . . .	65
3.25	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n = 24$ sur l'intervalle $[0, 50]$ . . . . .	66
3.26	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n = 103$ sur l'intervalle $[0, 200]$ . . . . .	66
3.27	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour $n = 1029$ sur l'intervalle $[0, 2000]$ . . . . .	66

---

3.28	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n = 20$ sur l'intervalle $[0, 50]$ . . . . .	67
3.29	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n = 87$ sur l'intervalle $[0, 200]$ . . . . .	67
3.30	Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour $n = 1040$ sur l'intervalle $[0, 2000]$ . . . . .	68
3.31	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle $[0, 50]$ . . . . .	70
3.32	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle $[0, 200]$ . . . . .	70
3.33	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle $[0, 2000]$ . . . . .	70
3.34	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l'intervalle $[0, 50]$ . . . . .	71
3.35	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l'intervalle $[0, 200]$ . . . . .	71
3.36	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l'intervalle $[0, 2000]$ . . . . .	72
3.37	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l'intervalle $[0, 50]$ . . . . .	72
3.38	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l'intervalle $[0, 200]$ . . . . .	73
3.39	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l'intervalle $[0, 2000]$ . . . . .	73
3.40	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) sur l'intervalle $[0, 50]$ . . . . .	74
3.41	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) sur l'intervalle $[0, 200]$ . . . . .	74
3.42	Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) sur l'intervalle $[0, 2000]$ . . . . .	74

# Liste des tableaux

2.1	Efficacité du noyau. . . . .	34
3.1	Résultats de simulation de la loi Gamma de paramètre $\mathcal{G}(9, 2.9)$ (cas fixe). . . . .	54
3.2	Résultats de simulation de la loi Log-normale de paramètre $Log\mathcal{N}(0, 1)$ (cas fixe). . . . .	56
3.3	Résultats de simulation de la loi Weibull de paramètre $\beta = 0.5$ et $\alpha = 5$ (cas fixe). . . . .	57
3.4	Résultats de simulation de la loi Weibull de paramètre $\beta = 1$ et $\alpha = 5$ (cas fixe). . . . .	58
3.5	Résultats de simulation de la loi Weibull de paramètre $\beta = 7$ et $\alpha = 5$ (cas fixe). . . . .	60
3.6	Résultats de simulation de la loi Gamma de paramètre $\mathcal{G}(9, 2.9)$ (cas aléatoire). . . . .	61
3.7	Résultats de simulation de la loi Log-normale de paramètre $Log\mathcal{N}(0, 1)$ (cas aléatoire). . . . .	63
3.8	Résultats de simulation de la loi Weibull de paramètre $\beta = 0.5$ , $\alpha = 5$ (cas aléatoire). . . . .	64
3.9	Résultats de simulation de la loi Weibull de paramètre $\beta = 1$ , $\alpha = 5$ (cas aléatoire). . . . .	65
3.10	Résultats de simulation de la loi Weibull de paramètre $\beta = 7$ , $\alpha = 5$ (cas aléatoire). . . . .	67
3.11	Résultats de simulation et comparaison de la loi Weibull (0.5, 5) . . . . .	69
3.12	Résultats de simulation et comparaison de la loi Weibull (1, 5) . . . . .	71
3.13	Résultats de simulation et comparaison de la loi Weibull (7, 5) . . . . .	72
3.14	Résultats de simulation et comparaison de la loi Log-normale (0, 1) . . . . .	73

# Introduction générale

La théorie de l'estimation est l'une des branches les plus basiques de la statistique. Cette théorie est habituellement divisée en deux composantes principales, à savoir, l'estimation paramétrique et l'estimation non paramétrique. L'approche paramétrique a comme inconvénient principal de nécessiter une connaissance préalable du phénomène aléatoire considéré. L'approche non paramétrique estime différentes fonctions directement à partir de l'information disponible sur l'ensemble d'observations, on dit que souvent dans cette approche les données parlent d'elles mêmes. On s'intéresse dans le cadre de ce travail à l'approche non paramétrique.

Dans la littérature, plusieurs méthodes non paramétriques ont été dédiées à l'estimation de la fonction de densité de probabilité. Nous citons la méthode d'estimation par histogramme dont l'origine est attribuée à *John Grant* au XVII<sup>ème</sup> siècle. Les propriétés des estimateurs par histogramme ont été initialement étudiées par *Geffroy (1974)* [22] et *Abou – Jaoude (1976)* [3]. Plusieurs méthodes ont été proposées et utilisées au cours de ces dernières années pour estimer la densité  $f$ . Elles comprennent : la méthode d'estimation par des séries orthogonales pour les densités à supports bornés, développée à partir des travaux de *Cencov (1962)* [13], et étudiée ensuite par plusieurs auteurs dont *Kronmal and Tarter (1968)*, *Wahba (1981)* [58], *Bosq (2005)* [7] et récemment *Saadi and Adjabi (2009)* [49]. Les méthodes d'interpolation par les fonction splines (*Wahba (1975)*) [59]. Une autre méthode basée sur les développements en ondelettes a été proposée. Les premiers résultats sur ces estimateurs sont donnés dans *Walter (1992)* [55] et *Kerkyacharian and Picard (1993)* [24]. Et la méthode du noyau, qui est la plus utilisée vu sa simplicité et la qualité de l'estimation qu'elle assure. Nous nous intéressons, dans notre travail à l'estimation par la méthode du noyau.

*Rosenblatt (1956)* [41], et *Parzen (1962)* [38] sont les premiers à proposer une classe d'estimateurs à noyau d'une densité univariée. Cet estimateur est une fonction de deux paramètres, le noyau  $K$  et le paramètre de lissage  $h$  (appelé aussi la largeur de la fenêtre).

Cette méthode est une généralisation de la méthode d'estimation par histogramme.

Les noyaux symétriques fournissent de bons résultats lorsque la densité à estimer est définie sur  $\mathbb{R}$  (voir [38, 41, 45]). Cependant, ils s'avèrent mal appropriés au cas de l'estimation des densités définies sur  $[0, +\infty[$ ; ceci est le cas dans l'estimation de la fonction fiabilité et c'est dû au poids qu'ils assignent en dehors du support positif de la densité à estimer pour cela les noyaux asymétriques sont introduit [9, 11, 43, 8]. Par ailleurs, le choix de la fenêtre est d'une ultime importance pour la qualité du lissage. Dans le cas de l'estimation d'une densité, il s'est développé une importante bibliographie sur la sélection de ce paramètre. Plusieurs travaux ont montré que les estimateurs peuvent changer dramatiquement suite à une petite variation du paramètre de lissage. Dans la littérature, des méthodes diverses ont été proposées pour le choix de ce paramètre; parmi elles, on peut citer les méthodes reposant sur la validation croisée, biaisée [47] et non biaisée [40, 5], la règle du pouce proposée par *Silvermann* (1986) [45], la méthode de ré-injection proposée par *Scott, Tapia* et *Thomson* [48] et les méthodes bootstrap proposées par *Taylor* (1989)[52], *Faraway* et *Jhun* (1990)[20].

Il existe plusieurs situations en pratique où le nombre d'observations récoltées est aléatoire, donc la densité de probabilité doit être estimée sur la base de cette échantillon. Pour cela, on utilise l'estimateur par la méthode du noyau de *Parzen – Rosenblatt* adapté au cas où la taille de l'échantillon est aléatoire qui a été proposé par *Srivastava* [44]. Les résultats asymptotiques de cet estimateur sont étudiés par *Srivastava* [44] et *Carroll* [15].

La théorie de la fiabilité a pour objectif d'étudier l'aptitude de dispositifs techniques (machines, équipements,...) à accomplir une fonction requise, dans des conditions données, durant un temps donné. Actuellement, c'est une discipline à part entière. Prévoir la fiabilité d'un système est essentielle pour des problèmes de sécurité (systèmes de freinage, systèmes nucléaires, systèmes informatiques...). La quasi-impossibilité de réparer certains matériels (satellites) et les problèmes économiques (coût de défaillances, gestion du personnel de maintenance, maintenance des stocks des pièces de rechange...) rendent nécessaire la connaissance de la fiabilité des systèmes utilisés. Toutefois la détermination de la probabilité de tomber en panne après un certain temps de bon fonctionnement de tels systèmes est complexe, car les lois d'apparition des pannes sont générales (quelconques).

L'objectif de ce travail est d'appliquer la méthode du noyau pour estimer la fiabilité et le taux de défaillance. Pour cela, on utilise l'estimateur empirique de la fonction de

répartition adapté au cas où la taille de l'échantillon est aléatoire, pour estimer la fiabilité et l'estimateur de *Parzen – Rosenblatt* adapté au cas où la taille de l'échantillon est aléatoire pour estimer la densité. Ces deux estimateurs seront combinés pour estimer la fonction du taux de hasard.

Ce mémoire est composé d'une introduction générale, de trois chapitres, d'une conclusion générale et d'une liste de références bibliographiques.

Dans le premier chapitre, nous présentons les principaux résultats sur la fiabilité et ces paramètres.

Le second chapitre comprend les principaux résultats sur l'estimation de la densité par la méthode du noyau, dans le cas où la taille de l'échantillon est fixe et aléatoire.

Dans le dernier chapitre, nous présentons les résultats de calcul par simulation de la fiabilité et le taux de hasard.

Enfin, nous terminons par une conclusion générale et quelques perspectives.

# Chapitre 1

## Concepts généraux de la fiabilité

### 1.1 Introduction

Très tôt, de grandes entreprises ont montré un grand intérêt pour la fiabilité : General Motors (depuis les années 1940), la NASA (le Département de la Défense au Etats-Unis, depuis les années 1950), Airbus, Air Force, Bell Telephone Laboratoires (depuis les années 1960), Thomson, Philips, Kodak, Citroën,... (depuis les années 1970) *Villemeur* [1988]. La recherche de la diminution du coût des défaillances en exploitation a entraîné une augmentation des exigences de fiabilité sur les systèmes. Ainsi, en 1995, General Electric a estimé que les coûts de non-fiabilité représentaient de 8 à 12 milliards dollars *Pyzdek* [2001] et a décidé d'augmenter le niveau de qualité de ses produits. La société Renault, quant à elle, estime qu'une modification effectuée au démarrage d'une série coûte de 10 à 100 fois plus cher que ce qu'elle aurait coûté deux ans auparavant *Rieuneau* [1993].

La maîtrise de la fiabilité d'un système représente un enjeu économique important pour toute entreprise. La mesure de cette grandeur est un premier pas indispensable vers sa concrétisation *DeMarco* [1982].

La fiabilité recouvre de multiples aspects : l'analyse de défaillance des systèmes, la fiabilité prévisionnelle, les banques de données de fiabilité, les essais de fiabilité, la fiabilité opérationnelle, les méthodes prévisionnelles de fiabilité et de sécurité, l'assurance de la fiabilité et de la qualité *Villemeur* [1988].

Il est impératif de faire appel aux calculs des probabilités afin d'étudier les problèmes de défaillance des systèmes qui se produisent de façon aléatoire, et pour lesquels la performance de l'entreprise en dépend énormément. Ainsi, nous définissons la fiabilité d'un dispositif comme étant sa probabilité de fonctionner correctement pendant une durée déterminée, autrement dit la probabilité qu'aucune défaillance ne se produise pendant cette

durée. Dans ce chapitre, nous allons rappeler les éléments de base sur la fiabilité.

## 1.2 Objectifs et intérêts de la fiabilité

L'analyse de la fiabilité constitue une phase indispensable dans toute étude de sûreté de fonctionnement. A l'origine, la fiabilité concernait les systèmes à haute technologie (centrales nucléaires et aérospatial). Aujourd'hui, la fiabilité est devenue un paramètre clé de la qualité et d'aide à la décision, dans l'étude de la plupart des composants, produits et processus "grand public" : Transport, énergie, bâtiments, composants électroniques, composants mécaniques, etc.

De nombreux industriels travaillent dans l'évaluation et l'amélioration de la fiabilité de leurs produits au cours de leur cycle de développement, de la conception à la mise en service (conception, fabrication et exploitation) afin de développer leurs connaissances sur le rapport Coût/Fiabilité et maîtriser les sources de défaillance [53].

## 1.3 Défaillance d'un système

La défaillance est une perte partielle ou totale des propriétés d'un élément qui diminuent significativement et entraîne la perte totale de sa capacité de fonctionnement, voir ([1, 36]).

Un dispositif peut se trouver dans l'un des états suivants :

- ★ Apte à fonctionner correctement, c'est à dire en état de service.
- ★ Inapte à fonctionner correctement, c'est à dire en panne ou hors de service.

### 1.3.1 Paramètres d'identification des défaillances

On pourra identifier les défaillances et les classer à partir d'une combinaison de paramètres répartis en six groupes en fonction de :

#### a. Causes fondamentales d'apparition :

- *La faiblesse inhérente :*

Défaillance attribuable à une faiblesse inhérente au matériel lui même lorsque les contraintes dépassent les capacités du matériel au-delà des possibilités données de celui-ci.

- *Défaillance première :*

Défaillance d'un matériel dont la cause directe ou indirecte n'est pas due à la défaillance d'un autre équipement.

- *Défaillance seconde :*

Défaillance d'un matériel dont la cause directe ou indirecte est due à la défaillance d'un autre équipement.

**b. La vitesse de manifestation des défaillances :**

- *Les défaillances graduelles :*

Elles se caractérisent par une variation progressive des paramètres déterminant la fiabilité du système. Ce type de défaillance est prépondérant pour le matériel mécanique, pour lequel se manifeste l'usure des composants après une certaine durée de bon fonctionnement.

- *Les défaillances subites :*

Elles sont dues à une perte totale des propriétés du système entraînant l'incapacité de fonctionnement. Ce type de défaillance caractérise spécialement un matériel électronique.

**c. En fonction de leur amplitude :**

- *Les défaillances partielles :*

Elles résultent de déviation d'une ou des caractéristiques au delà des limites spécifiées, mais telles qu'elles n'entraînent pas une disparition complète de la fonction requise.

- *Les défaillances complètes :*

Elles résultent de variations d'une ou des caractéristiques, telles qu'elles entraînent une disparition complète de la fonction requise.

**d. L'âge :**

- ***Les défaillances précoces :***

Elles surviennent durant la période de jeunesse de l'équipement, c'est-à-dire juste après sa mise en fonctionnement. Ces défaillances peuvent être attribuées à la conception, la fabrication ou la mauvaise utilisation de l'équipement. Généralement, un contrôle rigoureux est effectué avant la mise en service pour détecter les éléments présentant ce type de défaillance.

- ***Les défaillances aléatoires :***

Elles apparaissent durant la période de vie utile de l'équipement. Ce sont des défaillances accidentelles. Elles ont la même probabilité d'apparition et sont généralement catalectiques (c'est-à-dire à la fois soudaines et complètes).

- ***Les défaillances par usures :***

Ce sont des pannes progressives et elles apparaissent suite à la fatigue, l'usure et la dégradation de l'équipement après une longue période de fonctionnement, de processus de détérioration ou de corrosion.

**e. L'aptitude à être constatée :**

- ***Défaillance fugitive (systématique) :***

Défaillance de courte durée difficilement constatable et non reproductible à volonté.

- ***Défaillance intermittente :***

Défaillance non permanente, mais plus au moins facilement reproductible.

- ***Défaillance permanente :***

Défaillance constatable à tout moment.

## f. Conséquence :

• *Défaillance critiques :*

Défaillance qui risque de causer des blessures à des personnes ou des dégâts importants au matériel.

• *Défaillances majeure :*

Défaillance qui n'est pas critique mais qui risque de réduire l'aptitude d'un ensemble à accomplir sa fonction requise.

• *Défaillance mineure :*

Défaillance qui n'est pas critique et qui ne réduit pas l'aptitude d'un ensemble à accomplir sa fonction requise.

## 1.3.2 Densité de défaillance

Soit  $T$  la *v.a* représentant la durée de vie d'un dispositif (i.e, le temps qui sépare la mise en service et la défaillance). On note par  $f(t)$  la densité de probabilité de la *v.a*  $T$  et  $F_T(t)$  sa fonction de répartition. La densité de probabilité de la *v.a*  $T$  est appelée densité de défaillance ou probabilité de panne de l'élément à l'instant  $t$  [28].

$$f(t) = F'_T(t) = \lim_{dt \rightarrow 0} \left[ \frac{F_T(t + dt) - F_T(t)}{dt} \right].$$

## 1.3.3 Fonction de défaillance cumulée

La fonction de défaillance cumulée, notée  $F_T(t)$  représente, la fonction de répartition de la *v.a*  $T$ . C'est la probabilité de panne ou la probabilité de défaillance dans l'intervalle  $[0, t]$  [28].

$$F_T(t) = \mathbb{P}(T \leq t) = \int_0^t f(x)dx, \quad F'_T(t) = f(t), \quad t > 0.$$

Tels que :  $0 \leq F_T(t) \leq 1, \quad t \in [0, +\infty[.$

## 1.4 Fiabilité et probabilité de survie

### 1.4.1 Définition du concept de fiabilité

Le terme "fiabilité" est un néologisme. En effet, ce terme traduit du mot anglo-saxon "reability", a été introduit aux U.S.A en 1960 par la Commission Electrotechnique Internationale.

D'après l'association française de normalisation (AFNOR), la fiabilité est la caractéristique d'un dispositif exprimée par la probabilité que ce dispositif accomplisse une fonction requise dans les conditions d'utilisation et pour une période de temps déterminée. Toutes les défaillances de nature accidentelle sont prises en compte sans aucune discrimination vis-à-vis de leurs sévérités (Voir [1, 27]). L'objectif de l'étude de la fiabilité est de répondre à des questions telles que :

- ✓ Quelle est la durée de vie d'un produit pour fonctionner sans panne ?.
- ✓ Quelle proportion du produit va faire défaillance avant la fin de la garantie ?.
- ✓ Quelle période de garantie peut-on donner à un nouveau produit ?.
- ✓ Quand doit-on effectuer les travaux de maintenance ou remplacer une pièce ?.

La fiabilité fait partie intégrante du concept de qualité d'un produit ou du service qui repose sur :

- ✓ Sa conformité à l'instant de l'achat, ( $t = 0$ ).
- ✓ Sa fiabilité, notée  $R(t)$ .

Dans la définition de la fiabilité, on retient 4 notions fondamentales :

- **Notion de bien :**

Ces biens peuvent être classés en biens d'équipements, biens intermédiaires et biens de consommation. En outre, ils peuvent être classés en fonction de leur durabilité.

- **Notion "accomplir une fonction" :**

Exigence d'un niveau minimum de performances en deçà duquel le bien est considéré comme défaillant (critère de défaillance, taux de défaillance, mode de défaillance,...).

- **Notion de probabilité :**

La probabilité  $\mathbb{P}$  de réalisation d'un événement  $A$  est définie par le rapport entre le

nombre de cas favorables  $m$  et le nombre de cas possibles  $n$  :

$$\mathbb{P} = \text{Prob}\{A\} = m/n.$$

• **Notion de "temps donné" :**

La fiabilité est une fonction probabiliste associée à une variable aléatoire  $X$  dépendante du temps (Nombre de tours, de cycles,...).

La notion de fiabilité se distingue selon l'étape étudiée de la vie du système :

- a) **Fiabilité prévisionnelle** : c'est celle qui est déterminée durant la phase de conception sur la base d'un modèle mathématique, défini à partir des données de fiabilité de ses composants.
- b) **Fiabilité estimée** : c'est la fiabilité mesurée après avoir conçu le système et ceci à l'aide d'un ensemble d'essais.
- c) **Fiabilité opérationnelle** : c'est la fiabilité mesurée sur un matériel en exploitation. Elle dépend des conditions réelles d'exploitation et du support logistique.

### 1.4.2 Fonction de fiabilité ou fonction de survie

La fiabilité s'exprime par la fonction  $R(t)$  telle que :

$R(t) = \mathbb{P}(\text{qu'une entité } E \text{ soit non défailante sur la durée } [0, t], \text{ en supposant qu'elle n'est pas défailante à l'instant } t = 0)$ . C'est à dire,

$$R(t) = \mathbb{P}(T > t) = 1 - \mathbb{P}(T \leq t) = 1 - F(t) = 1 - \int_0^t f(x)dx = \int_t^{+\infty} f(x)dx.$$

Où,

$$f(t) = F'(t) = -R'(t),$$

avec

$$0 \leq R(t) \leq 1$$

La fiabilité  $R(t)$  est la probabilité que le système remplisse sa mission entre l'instant 0 et l'instant  $t$ . C'est une fonction décroissante du temps qui tend vers 0. La durée de vie peut s'exprimer en heures, semaines, cycles, etc.

En général, la fonction de fiabilité a la forme suivante :

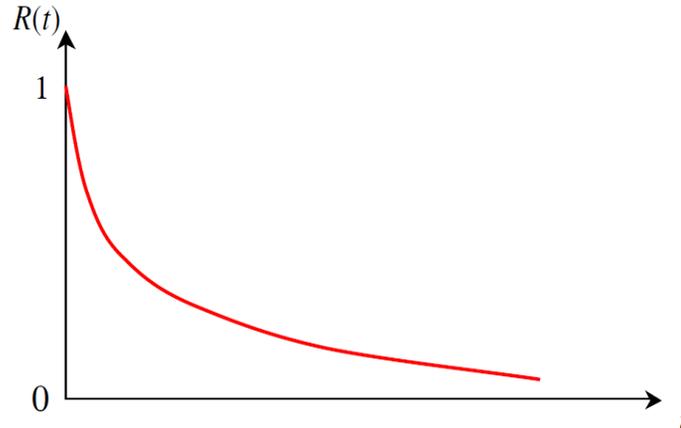


FIG. 1.1 – Fonction de fiabilité

**Propriété 1.4.1.** [28]

1.  $R(0) = 1$ .
2.  $\lim_{t \rightarrow +\infty} R(t) = 0$ .
3.  $\mathbb{P}(t_1 < T \leq t_2) = F_T(t_2) - F_T(t_1) = R(t_1) - R(t_2)$ .

## 1.5 Indices de fiabilité

$\lambda$  et le MTBF sont les principaux indicateurs de fiabilité. On distingue deux types de dispositif [28] :

- **Dispositif réparable** : ce sont les dispositifs pour lesquels il y'a une possibilité de remplacer un ou plusieurs composants. La fiabilité est caractérisée par :
  - ★  $\lambda$  : taux de défaillance.
  - ★  $t$  : temps de fonctionnement.
  - ★ **MTBF** : temps moyenne entre les défaillances ( Mean Time Between Failure).
  - ★ **MTTF** : temps moyenne jusqu'à la première défaillance (Mean Time To First Failure).
- **Dispositif non réparable** : ce sont les dispositifs pour lesquels il n'y a pas de possibilité de remplacement, c'est à dire arrêt total de l'élément. La fiabilité est caractérisée par :
  - ★  $\lambda$  : taux de défaillance.

- ★ **t** : temps de fonctionnement.
- ★ **MTBF** : paramètre inexistant.
- ★ **MTTF** : représente la durée de vie moyenne du système.

### 1.5.1 Moyenne des temps jusqu'à la première défaillance (MTTF)

C'est la durée de vie moyenne du système, notée  $\bar{T}$ , telle que,

$$\bar{T} = E(T).$$

- **Cas discret :**

$$\bar{T} = E(T) = \sum_{i=0}^{+\infty} i\mathbb{P}(T = i) = \sum_{i=0}^{+\infty} R(i)$$

- **Cas continu :**

$$\bar{T} = E(T) = \int_0^{+\infty} tf(t)dt = \int_0^{+\infty} R(t)dt.$$

### 1.5.2 Taux de défaillance $\lambda$

Le taux de défaillance (appelé aussi taux d'avarie)  $\lambda(t)$  est la probabilité que l'équipement fonctionne encore après  $t$  unités de temps et tombe en panne durant les  $dt$  unités supplémentaires. C'est donc la probabilité conditionnelle que l'équipement tombe en panne entre  $t$  et  $t + dt$ , sachant qu'il a fonctionné sans défaillance entre 0 et  $t$ .

$$\lambda(t) = \lim_{dt \rightarrow 0} \left[ \frac{\mathbb{P}(t < T \leq t + dt / T > t)}{dt} \right].$$

On a :

$$f(t)dt \approx \mathbb{P}(t < T \leq t + dt).$$

D'où,

$$\lambda(t)dt \approx \mathbb{P}(t < T \leq t + dt / T > t).$$

- **Lien avec la fiabilité**

$$\begin{aligned} \lambda(t) &= \lim_{dt \rightarrow 0} \left[ \frac{1}{dt} \mathbb{P}(t < T \leq t + dt / T > t) \right] \\ &= \lim_{dt \rightarrow 0} \left[ \frac{1}{dt} \frac{\mathbb{P}(t < T \leq t + dt)}{\mathbb{P}(T > t)} \right] \\ &= \frac{1}{\mathbb{P}(T > t)} \lim_{dt \rightarrow 0} \left[ \frac{1}{dt} \mathbb{P}(t < T \leq t + dt) \right] \end{aligned}$$

$$\Rightarrow \lambda(t) = \frac{f(t)}{\mathbb{P}(T > t)} = \frac{f(t)}{R(t)} = \frac{-R'(t)}{R(t)} = \frac{F'(t)}{1 - F(t)}.$$

En utilisant le fait que  $R(0) = 1$ , on obtient :

$$R(t) = \exp\left(-\int_0^t \lambda(x)dx\right). \quad t > 0$$

Le taux de défaillance caractérise la vitesse de variation de la fiabilité au cours du temps. Pour une période donnée,

$$\lambda = \frac{\text{Nombre de défaillances pendant le service}}{\text{Durée de bon fonctionnement}}$$

### 1.5.3 Fonction de hasard (Fonction de risque)

La fonction de hasard ou taux de défaillance cumulé, noté  $\Lambda(t)$ , est définie par :

$$\Lambda(t) = \int_0^t \lambda(x)dx = -\log(R(t)).$$

On obtient :

- **La fiabilité** :  $R(t) = e^{-\Lambda(t)}$ ,  $t > 0$ .
- **La fonction de répartition** :  $F(t) = 1 - e^{-\Lambda(t)}$ ,  $t > 0$ .
- **La fonction de densité** :  $f(t) = \lambda(t)e^{-\Lambda(t)}$ ,  $t > 0$ .

### 1.5.4 La courbe en baignoire (Bathtub Curve)

La courbe en baignoire (voir la figure (1.2) ci dessous), donne l'évolution du taux de défaillance  $\lambda(t)$  en fonction de l'âge du matériel. Elle comprend ( en particulier pour un matériel mécanique) trois phases, chacune avec un sens de variation différent [4].

**Période de jeunesse** : elle est caractérisée par des pannes précoces (défauts de conception et surproduction),  $\lambda(t)$  est décroissant.

**Période de maturité** : correspond à la vie utile de l'élément, c'est la zone de

maturité ou de pleine activité (défauts occasionnels),  $\lambda(t)$  est pratiquement constant.

**Période de vieillesse** : correspond à la période d'usure caractérisée par les défauts dus à l'âge (détérioration, corrosion,...). C'est la période de fin de vie de l'élément,  $\lambda(t)$  est croissant.

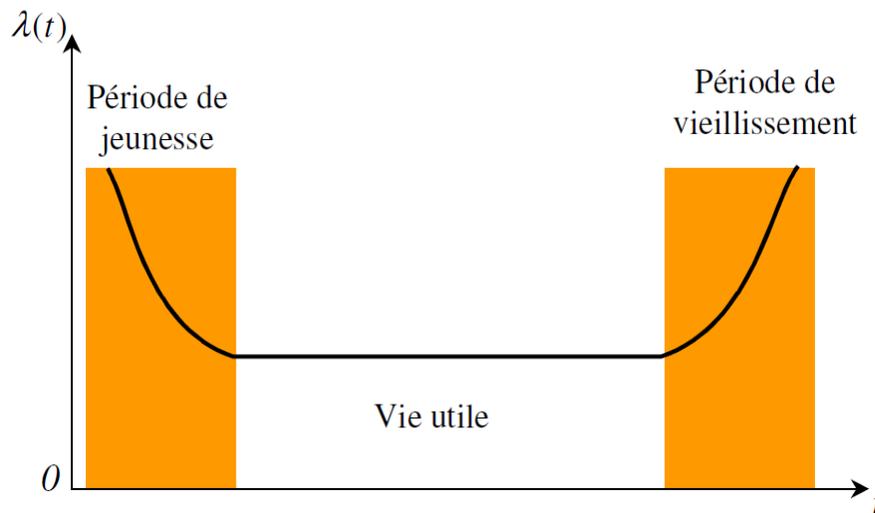


FIG. 1.2 – Courbe en baignoire

#### ✘ Taux de défaillance décroissant

Est caractéristique d'un matériel qui s'améliore (phase de déverminage ou de jeunesse, mortalité infantile). On retrouve ce type de défaillance sur des équipements, comme par exemple, des logiciels (dont on règle les problèmes au fur et à mesure) et des structures (dont la conception est révisée jusqu'à ce que la résistance soit mise au point).

#### ✘ Taux de défaillance constant

Est caractéristique d'un matériel qui est dans la période de la vie utile (phase d'exploitation ou de vie utile).

#### ✘ Taux de défaillance croissant

Qui est la caractéristique d'un matériel qui se détériore (phase d'obsolescence ou de vieillissement). On retrouve ce type de défaillance dans les pièces d'usure, les freins, les amortisseurs d'impact, les caoutchoucs et toute pièce sujette à la dégradation.

### 1.5.5 Moyenne des temps entre les défaillances (MTBF)

Le MTBF représente le temps moyen de bon fonctionnement ou temps moyen entre les défaillances.

$$MTBF = \frac{\text{Somme des temps de bon fonctionnement}}{\text{Nombre des temps de bon fonctionnement}}$$

Si  $\lambda$  est constant alors :

$$MTBF = \frac{1}{\lambda}.$$

## 1.6 Loi de survie

L'analyse de survie est une branche des statistiques qui cherche à modéliser les temps restants avant la mort pour les organismes biologiques, ou les temps restants avant la panne des systèmes artificiels. On parle aussi d'analyse de fiabilité en ingénierie [28].

**Notation :**

$N_0$  : nombre d'éléments mis en service à la date  $t = 0$ .

$N(t)$  : nombre de survivants à la date  $t$  (les éléments qui sont encore en service).

$\Delta(dt) = N(t) - N(t + dt)$  : nombre de survivant dans l'intervalle  $[t, t + dt]$ .

- **Fonction de survie** : c'est la fréquence relative des survivants ou probabilité de survie, donnée par :

$$R(t) = \frac{N(t)}{N_0}.$$

- **Fonction de répartition** : c'est la probabilité d'observer au moins un décès avant  $t$ .

$$F(t) = 1 - \frac{N(t)}{N_0} = 1 - R(t).$$

- **Fonction de densité** :

C'est la proportion d'éléments défailants dans l'intervalle  $[t, t + dt]$ .

$$f(t) = \frac{\Delta(dt)}{N_0} = \frac{N(t) - N(t + dt)}{N_0}.$$

- **Le taux de défaillance** : c'est la proportion d'éléments défailants depuis le dernier inventaire par rapport au nombre précédent.

※ *Premier cas* : les éléments défaillants sont remplacés dans l'intervalle  $dt$ .

$$\lambda(t) = \frac{\Delta(dt)}{N_0 dt}.$$

※ *Deuxième cas* : les éléments défaillants ne sont pas remplacés dans l'intervalle  $dt$ .

$$\lambda(t) = \frac{\Delta(dt)}{N(t)dt}.$$

## 1.7 Fiabilité des systèmes

Nous pouvons calculer la fiabilité d'un système par le biais de sa structure (ou architecture). Le principe est de considérer qu'un système est constitué de composants élémentaires, et que sa fiabilité dépend à la fois de la fiabilité de ses composants et de la façon dont le bon fonctionnement ou la panne de chaque composant influe sur le bon fonctionnement ou la panne du système tout entier [51].

Selon l'incidence de la fiabilité de l'un de ses composants sur l'état du système, on distingue les configurations suivantes :

### 1.7.1 Système en série

Pour une telle configuration, la panne de l'un des  $n$  éléments  $E_1, E_2, \dots, E_n$  placés en série entraîne la panne de tout le système, (voir la figure (1.3)).

La durée de vie du système est alors :

$$X = \min_{i \in \{1, 2, 3, \dots, n\}} (X_1, X_2, \dots, X_n),$$

où  $X_i$  est la durée de vie de l'élément  $i$ ,  $i = \overline{1..n}$ .

La fiabilité du système est donnée par :

$$R(t) = \prod_{i=1}^n R_i(t) = \exp \left( - \int_0^t \sum_{i=0}^n \lambda_i(x) dx \right), t > 0,$$

où  $R_i$  et  $\lambda_i$  sont la fiabilité et le taux de défaillance respectifs de l'élément  $i$ ,  $i = \overline{1..n}$ .

Dans ce cas  $\sum_{i=0}^n \lambda_i(x)$  est le taux de défaillance du système en série constitué de  $n$  composants.

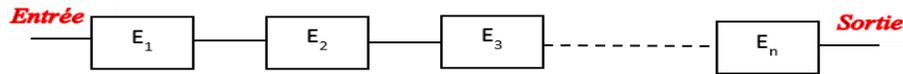


FIG. 1.3 – Configuration série.

### 1.7.2 Système en parallèle

Le système reste en état de fonctionnement si au moins un de ses  $n$  composants  $E_1, E_2, \dots, E_n$  placés en parallèle fonctionnent, (voir la figure (1.4)).

La durée de vie du système est alors :

$$X = \max_{i \in \{1, 2, 3, \dots, n\}} (X_1, X_2, \dots, X_n),$$

où  $X_i$  est la durée de vie de l'élément  $i$ ,  $i = \overline{1.n}$ .

La fiabilité du système est donnée par :

$$R(t) = 1 - \prod_{i=1}^n (1 - R_i(t)) = 1 - \prod_{i=1}^n \left[ 1 - \exp \left( - \int_0^t \sum_{i=0}^n \lambda_i(x) dx \right) \right], \quad t > 0,$$

où  $R_i$  et  $\lambda_i$  sont la fiabilité et le taux de défaillance respectifs de l'élément  $i$ ,  $i = \overline{1.n}$ .

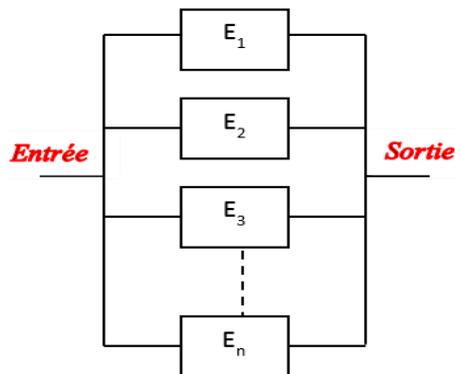


FIG. 1.4 – Configuration parallèle.

### 1.7.3 Système $K$ parmi $n$

Ce système fonctionne si au moins " $K$ " éléments parmi " $n$ " fonctionnent.

On suppose que tous les éléments sont identiques et indépendants et soit  $N_t$  le nombre d'éléments qui fonctionnent à l'instant  $t$ , alors :

$$N_t \rightsquigarrow \text{Binomiale}(n, R(t))$$

$$\begin{aligned} R(t) &= \mathbb{P}(N_t \geq K) \\ &= \sum_{m \geq K} \mathbb{P}(N_t = m) \\ &= \sum_{m \geq K}^n C_n^m (R(t))^m (1 - R(t))^{n-m}. \end{aligned}$$

### Cas particuliers

- Si  $K=1$  on obtient un système en parallèle.
- Si  $K=n$  on obtient un système en série.

#### *Remarque 1.7.1.*

Il existe d'autres configurations. Citons par exemple le système à configuration mixtes (qui sont obtenus en combinant les systèmes séries et les systèmes parallèles), le système en pont, le système à éléments dépendants, etc.

## 1.8 Principales lois de probabilité utilisées en fiabilité

Dans ce paragraphe, nous présentons quelques distributions de vie qui interviennent le plus fréquemment dans l'analyse des données de survie et de fiabilité qui sont communes à plusieurs disciplines. Nous énonçons les principales caractéristiques de ces lois (densité de probabilité, fonctions fiabilité et taux de défaillance) ainsi que leur application en fiabilité [33].

### 1.8.1 La loi exponentielle $\exp(\lambda)$

C'est la plus couramment utilisée en fiabilité électronique pour décrire la période durant laquelle le taux de défaillance des équipements est considéré comme constant.

Elle décrit le temps écoulé jusqu'à une défaillance ou l'intervalle de temps entre deux défaillances successives. Elle est définie en fonction du taux de défaillance  $\lambda$  ou en fonction de la moyenne des temps de bon fonctionnement (MTBF) qui est égale à  $1/\lambda$ .

- Densité de probabilité :

$$f(t) = \lambda e^{-\lambda t} \quad t \geq 0, \quad \lambda > 0.$$

- **Fonction de répartition :**

$$F(t) = 1 - e^{-\lambda t} \quad t > 0.$$

- **Fonction de fiabilité :**

$$R(t) = 1 - F(t) = e^{-\lambda t} \quad t \geq 0 \quad \lambda > 0.$$

- **Taux de défaillance :**

$$\lambda(t) = \lambda = cste \quad \forall t > 0.$$

- **Espérance :**

$$E(T) = \frac{1}{\lambda}.$$

- **Variance :**

$$V(T) = \frac{1}{\lambda^2}.$$

**Proposition 1.8.1.**

*Une propriété importante de la distribution exponentielle est la perte de mémoire ou absence de mémoire. Cette propriété se traduit mathématiquement par l'équation suivante :*

$$\mathbb{P}(T > s + t \mid T > t) = \mathbb{P}(T > s) \quad \forall s, t \geq 0.$$

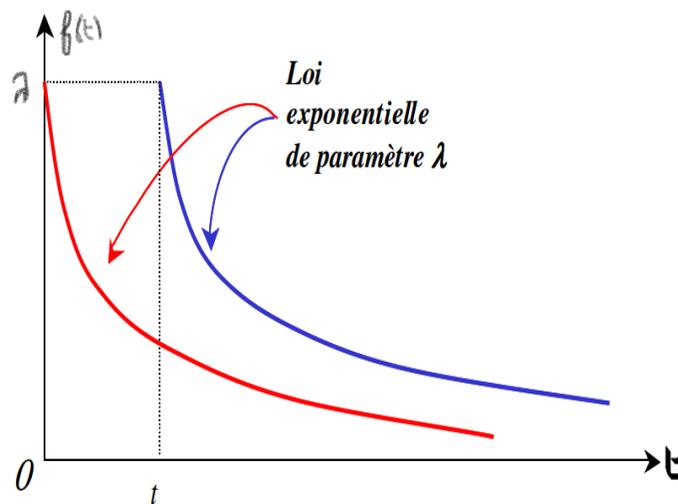


FIG. 1.5 – Propriété sans mémoire de la loi exponentielle de paramètre  $\lambda$ .

### 1.8.2 La loi de weibull $\mathcal{W}(\gamma, \beta, \eta)$

C'est la plus populaire des lois utilisées dans plusieurs domaines (électronique, mécanique,..). Elle permet de modéliser en particulier de nombreuses situations d'usure de matériel. Elle caractérise le comportement du système dans les trois phases de vie : période de jeunesse, période de vie utile et période d'usure ou vieillissement. Dans sa forme la plus générale, la distribution de Weibull dépend des trois paramètres suivants :  $\gamma$ ,  $\eta$  et  $\beta$  [1].

- **Le paramètre de position  $\gamma$  :**

Il représente le décalage existant entre le début de l'observation et le début du processus que l'on observe.

- **Le paramètre d'échelle  $\eta$ ,  $\eta > 0$  :**

Ce paramètre nous renseigne sur l'étendue de la distribution.

- **Le paramètre de forme  $\beta$ ,  $\beta > 0$  :**

Il est associé à la cinétique du processus observé.

- Si  $\beta < 1$  : Le taux de défaillance décroît dans le temps, ce qui correspond à la période de jeunesse des matériels dans la courbe en baignoire.
- Si  $\beta = 1$  : Le taux de défaillance est constant, ce qui correspond à la période de vie utile.
- Si  $\beta > 1$  : Le taux de défaillance est croissant, ce qui correspond à la période de vieillesse ou d'usure.

**Remarque 1.8.1.**

Pour  $\beta > 3$ , la loi de Weibull se rapproche de la loi normale.

- **Densité de probabilité :**

$$f(t) = \frac{\beta}{\eta} \left( \frac{t - \gamma}{\eta} \right)^{\beta-1} \exp \left[ - \left( \frac{t - \gamma}{\eta} \right)^\beta \right] \quad t > \gamma.$$

- **Fonction de répartition :**

$$F(t) = 1 - \exp \left( - \left( \frac{t - \gamma}{\eta} \right)^\beta \right) \quad t > \gamma.$$

- **Fonction de fiabilité :**

$$R(t) = \exp\left(-\left(\frac{t-\gamma}{\eta}\right)^\beta\right) \quad t > \gamma.$$

- **Espérance :**

$$E(T) = \int_0^{+\infty} R(t)dt = \gamma + \eta\Gamma\left(1 + \frac{1}{\beta}\right).$$

Où,

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt.$$

- **Taux de défaillance :**

$$\lambda(t) = \frac{f(t)}{R(t)} = \frac{\beta}{\eta} \left(\frac{t-\gamma}{\eta}\right)^{\beta-1}, t > 0.$$

**Remarque 1.8.2.**

Lorsque  $\gamma = 0$ , on considère la loi de Weibull à deux paramètres  $\beta, \eta$ .

### 1.8.3 Loi Gamma $\mathcal{G}(\alpha, \beta)$

La loi Gamma est définie par un paramètre de forme  $\alpha > 0$  et un paramètre d'échelle  $\beta > 0$ .

Elle est utilisée pour modéliser le temps de défaillance d'un élément.

- **Fonction de densité :**

$$f(t) = \frac{e^{-\frac{t}{\beta}} t^{\alpha-1}}{\beta^\alpha \Gamma(\alpha)} \quad t \geq 0.$$

- **Taux de défaillance :**

$$\lambda(t) = \frac{t^{\alpha-1} e^{-\left(\frac{t}{\beta}\right)}}{\beta^\alpha \int_t^{+\infty} \Gamma(\alpha) f(x) dx} \quad t \geq 0.$$

**Remarque 1.8.3.**

1- La fonction de répartition de la loi Gamma n'a pas d'expression explicite, ce qui est de même pour la fiabilité.

2- Si  $X$  suit la loi Gamma pour ( $\alpha = 1$  et  $\beta = \frac{1}{\lambda}$ ) on obtient la loi exponentielle ( $exp(\lambda)$ ).

### 1.8.4 Loi log-normale $Log\mathcal{N}(\mu, \sigma^2)$

Une variable aléatoire continue et positive est distribuée selon une loi log-normal si son logarithme est distribuée suivant une loi normale. Cette distribution est utilisée en fiabilité pour modéliser les défaillances par fatigue.

- **Fonction de densité :**

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} e^{-\frac{(\ln(t)-\mu)^2}{2\sigma^2}} \quad t \geq 0.$$

- **Fiabilité :**

$$R(t) = 1 - \Phi\left(\frac{\ln(t) - \mu}{\sigma}\right) \quad t \geq 0.$$

- **Taux de défaillance :**

$$\lambda(t) = \frac{e^{-\frac{(\ln(t)-\mu)^2}{2\sigma^2}}}{t \int_0^{+\infty} \sigma\sqrt{2\pi} f(t) dt}.$$

Où  $\Phi$  est la fonction de répartition de la loi normale centrée réduite.

Le domaine de définition n'étant jamais négatif, il n'y a aucune limitation à l'emploi de la distribution log-normale en fiabilité. Le taux de défaillance est croissant dans le début de vie puis décroissant en tendant vers zéro et la distribution est très dissymétrique.

#### *Remarque 1.8.4.*

Nous avons cité dans cette section les lois les plus utilisées en fiabilité, mais il existe d'autres lois comme la loi Birnbaum-Saunders, la loi Binômiale, la loi de Poisson, la loi Bêta, la loi Uniforme, etc.

## 1.9 Conclusion

Dans ce chapitre, nous avons présenté les notions fondamentales de la théorie de la fiabilité et de ses paramètres ainsi que les différentes lois usuelles de fiabilités susceptible, selon l'expérience, de représenter des distributions de durée de vie qui interviennent le plus fréquemment dans l'analyse de la fiabilité.

# Chapitre 2

## Estimation non paramétrique de la densité par la méthode du noyau

### 2.1 Introduction

La méthode du noyau est l'une des méthodes d'estimation non paramétrique la plus utilisée. *Rosenblatt* (1956) suivi de *Parzen* (1962), ont proposé une classe d'estimateurs à noyau d'une densité de probabilité, cet estimateur est une fonction de deux paramètres : le noyau  $K$  et le paramètre de lissage  $h$ . Le succès rencontré par cet estimateur s'explique par sa simplicité, sa flexibilité et aussi ses propriétés de convergence, il laisse à l'utilisateur une grande latitude non seulement dans le choix du noyau  $K$ , mais aussi dans le choix du paramètre de lissage  $h$ .

L'utilisation de cet estimateur ne se limite pas à l'estimation de la fonction de densité dans le cas où la taille de l'échantillon est fixe. En 1973, *Strivastava* utilise cet estimateur pour l'estimation de la densité dans le cas où le nombre d'observations est aléatoire. En effet, il y a plusieurs situations en pratique où l'on voudrait estimer la densité de probabilité sur la base d'un nombre aléatoire de données. Par exemple, si on veut estimer la fonction  $f(x)$  de la distribution de la vitesse des véhicules passant par un endroit particulier ou sur une route donnée, on peut enregistrer les vitesses de  $N_t$  véhicules qui passeront par cet endroit dans l'intervalle  $[0, t]$ . Dans ce cas  $N_t$  est une variable aléatoire à valeurs entières. Il s'agit donc d'un problème d'estimation séquentielle de la densité.

L'objectif de ce chapitre est de présenter l'estimateur de la fonction de densité par la méthode du noyau et ses propriétés dans le cas où la taille de l'échantillon est fixe ensuite dans le cas où la taille de l'échantillon est aléatoire.

## 2.2 Critères d'erreur et définitions

### 2.2.1 Les différents critères d'erreur

Pour mesurer les performances théoriques des estimateurs et identifier le meilleur, il est nécessaire de spécifier un critère d'erreur. Nous considérons la densité de probabilité  $f$  et son estimateur  $\hat{f}$  [61].

- **L'erreur quadratique intégrée ISE (Integrated Square Error) :**

$$ISE(f, \hat{f}) = \int |f(x) - \hat{f}(x)|^2 dx. \tag{2.1}$$

- **L'erreur quadratique moyenne MSE (Mean Square Error) :**

$$\begin{aligned} MSE(f(x), \hat{f}(x)) &= \mathbb{E}(f(x) - \hat{f}(x))^2 \\ &= \left( \text{biais}(\hat{f}(x)) \right)^2 + \mathbb{V}(\hat{f}(x)). \end{aligned}$$

En effet,

$$\begin{aligned} MSE(f(x), \hat{f}(x)) &= \mathbb{E}(f(x) - \hat{f}(x))^2 \\ &= \mathbb{E}(f(x)^2) - 2\mathbb{E}[f(x)\hat{f}(x)] + \mathbb{E}(\hat{f}(x)^2) \\ &= \mathbb{E}(f(x)^2) - 2\mathbb{E}[f(x)\hat{f}(x)] + \mathbb{E}(\hat{f}(x)^2) + [\mathbb{E}\hat{f}(x)]^2 - [\mathbb{E}\hat{f}(x)]^2 \\ &= f(x)^2 - 2f(x)\mathbb{E}(\hat{f}(x)) + \mathbb{E}(\hat{f}(x)^2) + [\mathbb{E}\hat{f}(x)]^2 - [\mathbb{E}\hat{f}(x)]^2 \\ &= [\mathbb{E}(\hat{f}(x)) - f(x)]^2 + \mathbb{E}(\hat{f}(x)^2) - [\mathbb{E}\hat{f}(x)]^2. \end{aligned}$$

Où  $\mathbb{E}(X)$  désigne l'espérance mathématique et  $\mathbb{V}(X)$  la variance de la v.a  $X$ .

- **L'erreur quadratique moyenne intégrée MISE (Mean Integrated Square Error) :**

$$\begin{aligned} MISE(f, \hat{f}) &= \int MSE(f(x), \hat{f}(x))dx = \int \mathbb{E} \left( f(x) - \hat{f}(x) \right)^2 dx \\ &= \int \left[ \left( \text{biais}(\hat{f}(x)) \right)^2 + \mathbb{V}(\hat{f}(x)) \right] dx. \end{aligned} \tag{2.2}$$

### 2.2.2 Quelques définitions

**Définition 2.2.1.** On dit qu'un estimateur  $\hat{f}$  de  $f$  est sans biais si :

$$\mathbb{E}(\hat{f}(x)) = f(x).$$

Il est dit asymptotiquement sans biais si :

$$\lim_{n \rightarrow \infty} \mathbb{E}(\hat{f}(x)) = f(x), \text{ en tout point } x \text{ pour lequel la densité } f \text{ est continue.}$$

Un estimateur  $\hat{f}$  de  $f$  est dit asymptotiquement uniformément sans biais si :

$$\limsup_{n \rightarrow \infty} \sup_x |\mathbb{E}[\hat{f}(x) - f(x)]| = 0.$$

**Définition 2.2.2.** On dit qu'un estimateur  $\hat{f}$  de  $f$  est ponctuellement consistant en moyenne quadratique si :

$$\lim_{n \rightarrow \infty} MSE(f(x), \hat{f}(x)) = 0, \text{ en tout point } x \text{ pour lequel la densité } f \text{ est continue.}$$

**Définition 2.2.3.** On dit qu'un estimateur  $\hat{f}$  de  $f$  est uniformément consistant en moyenne quadratique intégrée si :

$$\lim_{n \rightarrow \infty} MISE(f, \hat{f}) = 0.$$

**Définition 2.2.4.** On dit qu'un estimateur  $\hat{f}$  de  $f$  est asymptotiquement normal si :

$$\hat{f} \xrightarrow{\text{en loi}} \mathcal{N}(\mathbb{E}(\hat{f}), \mathbb{V}(\hat{f})).$$

## 2.3 Estimateur de Parzen-Rosenblatt

En statistique, l'estimation par noyau donnée par *Parzen* en (1962) et *Rosenblatt* en (1956) est une méthode non paramétrique d'estimation de la densité de probabilité d'une variable aléatoire. Elle est basée sur un échantillon d'une population statistique et permet d'estimer la densité en tout point du support. En ce sens, cette méthode généralise astucieusement la méthode d'estimation par histogramme. L'idée de l'estimation par la méthode du noyau consiste à évaluer la densité  $f(x)$  au point  $x$  en comptant le nombre d'observations figurant dans un certain voisinage de  $x$  sur  $\mathbb{R}$ .

Soit un n-échantillons  $X_1, X_2, \dots, X_n$  de variables aléatoires indépendantes et de même densité  $f$  inconnue. L'estimateur de  $f$  de *Parzen – Rosenblatt* est défini par :

$$\hat{f}_h(x) = \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right), \quad (2.3)$$

où  $h_n = h(n)$  est le paramètre de lissage vérifiant  $h_n \rightarrow 0$  et  $nh_n \rightarrow +\infty$  et  $K$  est la fonction noyau définie sur  $\mathbb{R}$  vérifiant les propriétés suivantes :

$$K(y) \geq 0, K(y) = K(-y), \int_{\mathbb{R}} K(y) dy = 1, \int_{\mathbb{R}} yK(y) dy = 0, \int_{\mathbb{R}} y^2 K(y) dy = \sigma_K^2 < \infty. \quad (2.4)$$

L'estimateur donné par (2.3) est bien une densité de probabilité. En effet :

$$\begin{aligned}
 \int_{\mathbb{R}} \hat{f}_h(x) dx &= \int_{\mathbb{R}} \frac{1}{nh_n} \sum_{i=1}^n K\left(\frac{x - X_i}{h_n}\right) dx \\
 &= \frac{1}{nh_n} \sum_{i=1}^n \int_{\mathbb{R}} K\left(\frac{x - X_i}{h_n}\right) dx \\
 &= \frac{1}{n} \sum_{i=1}^n \int_{\mathbb{R}} K(u) du \\
 &= 1.
 \end{aligned}$$

## 2.4 Propriétés de l'estimateur de Parzen-Rosenblatt

Il existe de nombreuses notions de convergence pour l'estimateur à noyau. Elle sont essentielles pour les applications et elles servent surtout à présenter et identifier certaines propriétés. Dans cette section, nous donnons deux théorèmes sur le comportement asymptotique du biais et de la variance qui ont été établis par *Parzen* [38]. Ensuite, nous donnons des résultats sur la convergence en moyenne quadratique, en moyenne quadratique intégrée, la convergence uniforme (en probabilité et presque complète) et le comportement asymptotique de l'estimateur. Le premier résultat de convergence est dû à Bochner, dont le lemme est à la base des principaux théorèmes de convergence.

**Lemme 2.4.1.** (*Bochner*) :

1. Soit  $K$  un noyau de *Parzen – Rosenblatt* et  $g$  une fonction de  $\mathbb{L}^1$ . Alors, en tout point  $x$ , où  $g$  est continue :

$$\text{On note } K_h = \frac{1}{h_n} K\left(\frac{\cdot}{h_n}\right),$$

$$\lim_{h \rightarrow 0} (K_h * g)(x) = g(x).$$

2. Soit  $K$  un noyau quelconque et  $g$  une fonction de  $\mathbb{L}^1$  uniformément continue, alors

$$\limsup_{h \rightarrow 0} \sup_x |(K_h * g)(x) - g(x)| = 0.$$

### 2.4.1 Propriétés statistiques de l'estimateur à noyau

#### ✂ L'espérance mathématique

L'espérance mathématique de  $\hat{f}_h(x)$  est :

$$\begin{aligned}
\mathbb{E}(\hat{f}_h(x)) &= \frac{1}{nh_n} \mathbb{E} \left[ \sum_{i=1}^n K \left( \frac{x - X_i}{h_n} \right) \right] \\
&= \frac{1}{nh_n} \sum_{i=1}^n \int_{\mathbb{R}} K \left( \frac{x - u}{h_n} \right) f(u) du \\
&= \frac{1}{h_n} \int_{\mathbb{R}} K \left( \frac{x - u}{h_n} \right) f(u) du.
\end{aligned}$$

En posant  $y = \frac{x-u}{h_n}$ , on obtient :

$$\mathbb{E}(\hat{f}_h(x)) = \int_{-\infty}^{+\infty} K(y) f(x - h_n y) dy.$$

En faisant le développement de Taylor à l'ordre 2 au point  $h_n = 0$  de  $f(x - h_n y)$ . On obtient :

$$f(x - h_n y) = f(x) - \frac{h_n y}{1!} f'(x) + \frac{h_n^2 y^2}{2!} f''(x) + o(h_n^2).$$

D'où,

$$\begin{aligned}
\mathbb{E}(\hat{f}_h(x)) &= \int_{-\infty}^{+\infty} K(y) f(x) - h_n y f'(x) + \frac{h_n^2 y^2}{2!} f''(x) dy + o(h_n^2) \\
&= f(x) \int_{-\infty}^{+\infty} K(y) dy - h_n f'(x) \int_{-\infty}^{+\infty} y K(y) dy + \frac{h_n^2}{2!} f''(x) \int_{-\infty}^{+\infty} y^2 K(y) dy + o(h_n^2).
\end{aligned}$$

En plus des propriétés (2.4), si le noyau  $K(y)$  est symétrique par rapport à 0, alors les expressions finales sont données par :

$$\mathbb{E}(\hat{f}_h(x)) = f(x) + \frac{h_n^2}{2} f''(x) \sigma_K^2 + o(h_n^2). \quad (2.5)$$

### ✂ Le biais

Le biais de  $\hat{f}_h(x)$  est :

$$biais(\hat{f}_h(x)) = \mathbb{E}(\hat{f}_h(x)) - f(x) = \int_{-\infty}^{+\infty} K(y) f(x - h_n y) dy - f(x).$$

D'où

$$biais(\hat{f}_h(x)) = \mathbb{E}(f(x)) - f(x) = \frac{h_n^2}{2} f''(x) \sigma_K^2 + o(h_n^2). \quad (2.6)$$

### ✂ La variance

La variance de  $\hat{f}_h(x)$  est :

$$\begin{aligned}
\mathbb{V}(\hat{f}_h(x)) &= \mathbb{V}\left[\sum_{i=1}^n \frac{1}{nh_n} K\left(\frac{x-X_i}{h_n}\right)\right] \\
&= \frac{1}{n^2 h_n^2} \sum_{i=1}^n \mathbb{V}\left[K\left(\frac{x-X_i}{h_n}\right)\right] \\
&= \frac{1}{n^2 h_n^2} \sum_{i=1}^n \left[\mathbb{E}\left(K\left(\frac{x-X_i}{h_n}\right)\right)^2\right] - \frac{1}{n^2 h_n^2} \sum_{i=1}^n \left[\mathbb{E}\left(K\left(\frac{x-X_i}{h_n}\right)\right)\right]^2 \\
&= \frac{1}{nh_n^2} \int_{-\infty}^{+\infty} \left[K\left(\frac{x-y}{h_n}\right)\right]^2 f(y) dy - \frac{1}{n} \left[\frac{1}{h_n} \int_{-\infty}^{+\infty} K\left(\frac{x-y}{h_n}\right) f(y) dy\right]^2.
\end{aligned}$$

Avec le changement de variable,  $u = \frac{x-y}{h_n}$ , on obtient :

$$\mathbb{V}(\hat{f}_h(x)) = \frac{1}{nh_n} \int_{-\infty}^{+\infty} (K(y))^2 f(x - h_n y) dy - \frac{1}{n} \left(\int_{-\infty}^{+\infty} K(y) f(x - h_n y) dy\right)^2.$$

D'où,

$$\mathbb{V}(\hat{f}_h(x)) = \frac{1}{nh_n} R(K) f(x) + o\left(\frac{1}{nh_n}\right). \quad (2.7)$$

Où  $R(g) = \int_{\mathbb{R}} g^2(y) dy$ , pour toute fonction  $g$  de carré intégrable.

### ✂ Le MSE de l'estimateur à noyau

$$\begin{aligned}
MSE(f(x), \hat{f}_h(x)) &= \mathbb{E}[(f(x) - \hat{f}_h(x))^2] \\
&= [\text{Biais}(\hat{f}_h(x))]^2 + \mathbb{V}(\hat{f}_h(x)) \\
&= \left[\frac{h_n^2}{2} f''(x) \sigma_K^2 + o(h_n^2)\right]^2 + \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(y) dy + o\left(\frac{1}{nh_n}\right) \\
&= \frac{h_n^4}{4} f''^2(x) \sigma_K^4 + \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(y) dy + o(h_n^4) + o\left(\frac{1}{nh_n}\right) \\
&= \frac{h_n^4}{4} f''^2(x) \sigma_K^4 + \frac{1}{nh_n} f(x) R(K) + o(h_n^4) + o\left(\frac{1}{nh_n}\right). \quad (2.8)
\end{aligned}$$

Où,  $\sigma_K^4 = \left(\int y^2 K(y) dy\right)^2$  est la variance carré du noyau  $K$ .

### ✂ Le MISE de l'estimateur à noyau

$$\begin{aligned}
MISE(f, \hat{f}_h) &= \int_{\mathbb{R}} MSE(f(x), \hat{f}_h(x)) dx \\
&= \int_{\mathbb{R}} \left\{ \frac{h_n^4}{4} (f''(x))^2 \sigma_K^4 + \frac{1}{nh_n} f(x) \int_{\mathbb{R}} K^2(y) dy + o(h_n^2) + o\left(\frac{1}{nh_n}\right) \right\} dx \\
&= \frac{1}{nh_n} \int_{\mathbb{R}} K^2(y) dy + \frac{h_n^4}{4} \sigma_K^4 \int_{\mathbb{R}} f''^2(x) dx + o\left(h_n^4 + \frac{1}{h_n}\right) \\
&= \frac{1}{nh_n} R(K) + \frac{h_n^4}{4} \sigma_K^4 R(f'') + o\left(h_n^4 + \frac{1}{h_n}\right). \tag{2.9}
\end{aligned}$$

## 2.4.2 Propriétés asymptotiques de l'estimateur à noyau

Dans ce qui suit, on suppose que les conditions suivantes sont vérifiées :

$$\lim_{n \rightarrow \infty} h_n = 0, \quad \lim_{y \rightarrow \infty} |yK(y)| = 0, \quad \sup_y |K(y)| < \infty, \quad \int_{-\infty}^{+\infty} K(y) dy = 1. \tag{2.10}$$

### 2.4.2.1 Comportement asymptotique du biais

**Théorème 2.1.** (*Parzen [38]*)

Si la fonction  $K$  satisfait les conditions (2.10) :

Alors, l'estimateur  $\hat{f}_h(x)$  est asymptotiquement sans biais c'est -à-dire :

$$\lim_{n \rightarrow \infty} \mathbb{E}(\hat{f}_h(x)) = f(x), \quad \text{en tout point } x \text{ pour lequel la densité } f \text{ est continue.}$$

### 2.4.2.2 Comportement asymptotique de la variance

**Théorème 2.2.** (*Parzen [38]*)

Si la fonction  $K$  satisfait les conditions (2.10) : alors

$$\lim_{n \rightarrow \infty} nh_n \mathbb{V}(\hat{f}_h(x)) = f(x) \int_{-\infty}^{+\infty} K^2(y) dy, \quad \text{en tout point } x \text{ pour lequel la densité } f$$

est continue.

Les convergences au sens de l'erreur quadratique moyenne et de l'erreur quadratique moyenne intégrée ont été établies respectivement par *Parzen [38]* et *Tiago de Oliveira [50]*, pour l'estimateur  $\hat{f}_h(x)$  défini dans (2.3).

### 2.4.2.3 Convergence en moyenne quadratique

En remplaçant les expressions finales du biais (2.6) et de la variance (2.7) dans l'équation (2.2), on obtient :

$$MSE(f(x), \hat{f}_h(x)) = \frac{f(x)}{nh_n} R(K) + \frac{h_n^4}{4} (f''(x))^2 \sigma_K^4 + o\left(\frac{1}{nh_n} + h_n^2\right). \quad (2.11)$$

**Théorème 2.3.** (*Parzen [38]*)

Si  $\lim_{n \rightarrow \infty} nh_n = \infty$  et  $K$  satisfait aux conditions (2.10), alors l'estimateur  $\hat{f}_h(x)$  est consistant en moyenne quadratique, c'est-à-dire :

$$\lim_{n \rightarrow \infty} MSE(f(x), \hat{f}_h(x)) = 0, \text{ en tout point } x \text{ pour lequel la densité } f \text{ est continue.}$$

### 2.4.2.4 Convergence en moyenne quadratique intégrée

**Théorème 2.4.** (*Parzen [38]*)

Si  $K$  est un noyau de *Parzen – Rosenblatt*, on a :

$$\lim_{n \rightarrow \infty} h_n = 0, \lim_{n \rightarrow \infty} nh_n = \infty.$$

Alors,

$$(\forall f \in \mathbb{L}^p), \lim_{n \rightarrow \infty} MISE((\hat{f}_h(x), f(x))) = 0.$$

Où  $\mathbb{L}^p$  est l'ensemble des fonctions réelles de puissance  $p^{\text{ième}}$  intégrable, c'est-à-dire l'ensemble des fonctions  $f$  définies sur  $\mathbb{R}$ , telles que  $\int |f(x)|^p < \infty$ .

### 2.4.2.5 Convergence uniforme en probabilité

La convergence uniforme en probabilité a été obtenue par *Parzen (1962)*.

**Théorème 2.5.** (*Parzen [38]*)

Si

$$\lim_{n \rightarrow \infty} nh_n^2 = \infty,$$

et si la fonction  $K$  satisfait les conditions (2.10). Et si la transformée de Fourier  $\tilde{K}(z) = \int_{-\infty}^{+\infty} \exp(-izy)K(y)dy$  est absolument intégrable, alors,  $\hat{f}_h(x)$  est un estimateur uniformément consistant en probabilité, c'est-à-dire :

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}(\sup_x |\hat{f}_h(x) - f(x)| < \epsilon) = 1.$$

### 2.4.2.6 Convergence uniforme presque complète

La convergence presque complète a été obtenue par [Nadaraya \(1965\)](#).

**Théorème 2.6.** ([Nadaraya \[35\]](#))

Si  $K$  est un noyau positif à variation bornée et  $f$  est uniformément continue et si,

$$\lim_{n \rightarrow \infty} h_n = 0, \quad \sum_{n=1}^{\infty} \exp(-\gamma n h_n^2) < \infty, \forall \gamma > 0.$$

Alors,

$$\sup_x |\hat{f}_h(x) - f(x)| \rightarrow 0 \text{ avec une probabilité 1.}$$

[Silvermann \[46\]](#) a donné le même théorème sur la convergence presque complète en remplaçant la condition,  $\sum_{n=1}^{\infty} \exp(-\gamma n h_n^2) < \infty$  par les deux conditions suivantes :

$$\lim_{n \rightarrow \infty} h_n = 0 \text{ et } \lim_{n \rightarrow \infty} \frac{\log n}{n h_n} = 0.$$

**Théorème 2.7.** ([Silvermann \[45\]](#))

Si on a :

$$\lim_{n \rightarrow \infty} h_n = 0 \text{ et } \lim_{n \rightarrow \infty} \frac{\log n}{n h_n} = 0.$$

et  $K$  satisfait aux conditions suivantes :

- ★  $K$  est uniformément continu et à variation bornée sur  $\mathbb{R}$ ,
- ★ supposons aussi que  $f$  est uniformément continu,
- ★  $\int_{-\infty}^{+\infty} \sqrt{|y \log(y)|} |dK(y)| < \infty$ ,
- ★  $\int_{-\infty}^{+\infty} K(y) dy = 1$ .

Alors,

$$\limsup_{n \rightarrow \infty} \sup_x [\hat{f}_h(x) - f(x)] = 0 \text{ p.s.}$$

### 2.4.2.7 Convergence en loi

La convergence en loi a été établie par [Parzen \[1962\]](#).

**Théorème 2.8.** ([Parzen \[38\]](#)).

$\lim_{n \rightarrow \infty} h_n = 0, \lim_{n \rightarrow \infty} n h_n = \infty$  et  $K$  satisfait les conditions du théorème (2.1), alors  $\hat{f}_h(x)$

est un estimateur asymptotiquement normal en tout point  $x$  pour lequel la densité  $f$  est continu, c'est à dire :

$$\frac{\hat{f}_h(x) - \mathbb{E}\{\hat{f}_h(x)\}}{\sqrt{\mathbb{V}\{\hat{f}_h(x)\}}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

où  $\xrightarrow{\mathcal{L}}$  désigne la convergence en loi et  $\mathcal{N}(0,1)$  est la loi normale standard.

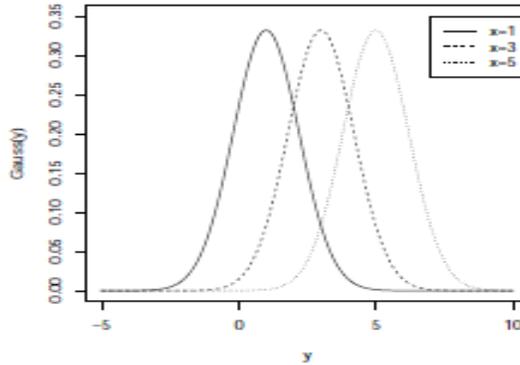


FIG. 2.1 – Noyau gaussien pour  $h = 1.2$ ,  $x = 1$ ,  $x = 3$  et  $x = 5$

### 2.4.2.8 Convergence $\mathbb{L}^1$ presque complète

**Théorème 2.9.** (*Devroye [16]*)

Soit  $K$  une fonction absolument intégrable sur  $\mathbb{R}^d$ , alors :

$$\lim_{n \rightarrow \infty} h_n \rightarrow 0, \lim_{n \rightarrow \infty} nh_n \rightarrow +\infty \Leftrightarrow (\forall f \in \mathbb{R}^d), \int |\hat{f}_h(x) - f(x)| dx \xrightarrow{p.co} 0.$$

Dans ces conditions, les convergences  $\mathbb{L}^1$ , probabilité, presque sûre et complète sont équivalentes.

Une version améliorée de ce théorème est donnée par *Devroye (1983)* [16] comme suit :

Si

$$(\exists f \in \mathbb{R}^d), \int |\hat{f}_h(x) - f(x)| dx \xrightarrow{p} 0.$$

Alors :

$$\lim_{n \rightarrow \infty} h_n \rightarrow 0, \lim_{n \rightarrow \infty} nh_n \rightarrow \infty \Leftrightarrow (\forall f \in \mathbb{R}^d), \int |\hat{f}_h(x) - f(x)| dx \xrightarrow{p.co} 0.$$

## 2.5 Le Choix pratique de $(K, h)$

L'estimateur de *Parzen – Rosenblatt* de la fonction de densité de probabilité (2.3) étant une fonction du noyau  $K$  et du paramètre de lissage  $h$ , il est nécessaire de faire un bon choix pour ces paramètres pour avoir un estimateur avec de bonnes propriétés statistiques et asymptotiques. Le noyau  $K$  peut être discret ou continu, et choisi selon la symétrie du domaine de définition. Il existe deux catégories principales de noyaux, les noyaux symétriques et les noyaux asymétriques. Quant au choix du paramètre de lissage  $h$ , qui est crucial pour la qualité de l'estimateur, plusieurs méthodes ont été développées. Dans ce qui suit, nous présentons quelques noyaux usuels, ensuite on donnera quelques méthodes du choix de la fenêtre  $h$ .

### 2.5.1 Choix du noyau $K$

Le choix du noyau semble être réglé lorsque on choisit le MISE comme critère d'optimisation. En réalité, on peut être amené à envisager d'autres noyaux pour des raisons de lissage ou de facilité de calcul. Dans ce qui suit, nous présentons quelques noyaux des deux catégories.

#### 2.5.1.1 Noyaux symétriques

##### Noyau uniforme (noyau de Rosenblatt) :

Ce noyau a été proposé par *Rosenblatt* (1956) [41]. L'avantage de ce noyau est la simplicité de sa forme, qui est donnée comme suit :

$$K(u) = \begin{cases} 1/2, & \text{si } |u| \leq 1; \\ 0, & \text{sinon.} \end{cases}$$

##### Noyau triangulaire :

Ce noyau a un avantage par rapport au noyau uniforme, il est continu partout, ce qui conduit à une estimation continue de  $f$ , sa forme est :

$$K(u) = \begin{cases} (1 - |u|), & \text{si } |u| \leq 1; \\ 0, & \text{sinon.} \end{cases}$$

##### Noyau Gaussien :

Ce noyau s'écrit sous la forme :

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right), \quad u \in \mathbb{R}.$$

**Noyau Epanechnikov :**

Ce noyau a été proposé par *Epanechnikov* en 1969 [19], il est défini par :

$$K(u) = \begin{cases} \frac{3}{4\sqrt{5}}(1 - \frac{u^2}{5}), & \text{si } |u| < \sqrt{5}; \\ 0, & \text{sinon.} \end{cases}$$

**noyau biweight :**

Le noyau biweight s’écrit sous la forme :

$$K(u) = \begin{cases} \frac{15}{16}(1 - u^2)^2, & \text{si } |u| \leq 1; \\ 0, & \text{sinon.} \end{cases}$$

Le tableau (2.1) donne l’efficacité relative de ces noyaux les plus utilisés. Comme ces efficacités sont très rapprochées, le choix du noyau influe peu sur la valeur du MISE asymptotique. En pratique on choisira  $K$  en tenant compte de la facilité des calculs plutôt que de l’efficacité relative.

$$Eff(K) = \frac{\int K^2(u)du}{\int K_E^2(u)du} = \frac{\int K^2(u)du}{\frac{3}{5\sqrt{5}}}.$$

noyau	Efficacité
Uniforme	1.0758
Triangulaire	1.0143
Gaussien	1.0513
noyau biweight	1.0061
Epanechnikov	1

TAB. 2.1 – Efficacité du noyau.  
Sihem Touloum 12 :50 Sihem Touloum

**2.5.1.2 Noyaux asymétriques**

Quoique les méthodes précédentes diminuent le biais aux bornes, elles restent peu efficaces, car le biais reste considérable si on le compare aux biais de l’intérieur du support [12]. Pour obtenir un biais aux bornes de même ordre que celui de l’intérieur, et dans l’objectif de trouver un équilibre, plusieurs auteurs ont proposé des estimateurs dans ce sens. Cependant, ces estimateurs attribuent des poids négatifs aux valeurs du voisinage des bornes. La solution la plus récente est d’utiliser des noyaux asymétriques adaptés qui n’assignent aucun poids à l’extérieur du support. *Chen* (1999,2000a) ([9] et [10]) propose, le noyau Beta pour les densités à support compact (exemple : [a,b]) et le noyau

Gamma pour les densités à support positif (c'est-à-dire sur  $[0, +\infty[$ ). Ces deux noyaux asymétriques seront présentés dans le paragraphe suivant :

**Noyau Gamma :**

Le noyau gamma a été introduit par *Chen (2000a)* [10] pour estimer des densités à support  $\mathbb{N} = [0, +\infty[$ . Il a utilisé la loi gamma pour construire des noyaux asymétriques. Deux classes de noyaux ont été proposées.

La première classe des noyaux gamma est :

$$K_{Gam(\frac{x}{h}+1,h)}(t) = \frac{t^{\frac{x}{h}} e^{-\frac{t}{h}}}{h^{\frac{x}{h}+1} \Gamma(\frac{x}{h} + 1)}, \quad t \in [0, +\infty[. \tag{2.12}$$

où  $\Gamma(\alpha) = \int_0^\infty \exp(-u)u^{\alpha-1}du$ ,  $\alpha > 0$  est la fonction Gamma et  $h$  est le paramètre de lissage satisfaisant les conditions  $h_n \rightarrow 0$  et  $nh_n \rightarrow \infty$  quand  $n \rightarrow \infty$ . L'estimateur de la densité associé au noyau gamma est donné par :

$$\hat{f}_{Gam}(x) = \frac{1}{n} \sum_{i=1}^n K_{Gam(\frac{x}{h}+1,h)}(X_i). \tag{2.13}$$

La deuxième classe a été proposée par le même auteur pour améliorer les performances (réduire le biais) de l'estimateur à noyau Gamma. La forme du noyau gamma modifié est :

$$K_{Gam(\rho_h(x),h)}(t) = \frac{t^{\frac{x}{h}} e^{-\frac{t}{h}}}{h^{\rho_h(x)} \Gamma(\rho_h(x))}, \tag{2.14}$$

avec :

$$\rho_h(x) \begin{cases} \frac{x}{h}, & \text{si } x \geq 2h; \\ \frac{1}{4}(x/h)^2 + 1, & \text{si } x \in [0, 2h]. \end{cases}$$

L'estimateur à noyau gamma modifié est alors

$$\hat{f}_{Gam-mod}(x) = \frac{1}{n} \sum_{i=1}^n K_{Gam-mod(\rho_h(x),h)}(X_i). \tag{2.15}$$

**Noyau Beta :**

Le noyau Beta a été proposé par *Brown* et *Chen (1999)* [9], *Chen (1999,2000b)*[9, 11] pour l'estimation non paramétrique de la courbe de régression et des densités unidimensionnelles définies sur un support compact.

L'idée de *Harrell* et *Davis (1982)*[23] et *Chen (1999)* [9] est d'utiliser le noyau Beta pour estimer la densité à support  $[0, 1]$  et de régler ainsi le problème du biais aux bornes. L'estimateur est de la forme :

$$\hat{f}_{Beta}(x) = \frac{1}{n} \sum_{i=1}^n K_{Beta} \left( X_i, \frac{x}{h} + 1, \frac{1-x}{h} + 1 \right), \quad (2.16)$$

où  $K_{Beta}(\cdot, \alpha, \beta)$  est la densité de la distribution Beta de paramètres  $\alpha$  et  $\beta$ ,

$$K_{Beta}(x, \alpha, \beta) = \frac{x^\alpha(1-x)^\beta}{\mathbf{B}(\alpha, \beta)}, \quad x \in [0, 1],$$

avec,

$$\mathbf{B}(\alpha, \beta) = \frac{\Gamma(\alpha, \beta)}{\Gamma(\alpha)\Gamma(\beta)} \text{ et } \Gamma(\alpha) = \int_0^{+\infty} e^{-x} x^{\alpha-1} dx, \quad \forall \alpha > 0.$$

Où  $\alpha = \frac{x}{h} + 1$  et  $\beta = \frac{1-x}{h} + 1$ .

Le noyau Beta a deux avantages, il peut parfaitement estimer les densités à support compact et il possède une forme flexible qui change le lissage dans le sens naturel quand on s'éloigne des bornes. Par conséquent, le noyau Beta élimine le biais aux bornes et fournit une réduction de la variance. *Charpentier, Fermanian et Scaillet* [14] ont montré par simulation que l'estimateur à noyau Beta est plus performant que les estimateurs avec des noyaux standards.

### 2.5.2 Choix du paramètre de lissage $h$

Si le choix du noyau ne pose pas un grand problème dans l'estimation de la densité, il n'en est pas de même pour le choix de la largeur de fenêtre qui ne dépend que de la taille  $n$  de l'échantillon. Plusieurs travaux ont montré que les estimateurs peuvent changer dramatiquement pour de petites variations du paramètre de lissage. Il existe plusieurs méthodes de sélection de ce paramètre qui reposent sur la minimisation de l'estimation de l'Erreur Quadratique Moyenne (MSE) ou l'Erreur Quadratique Moyen Intégrée (MISE). Ces méthodes restent incomplètes ou autrement dit, ces méthodes ont toujours des inconvénients, soit au sens de la qualité de l'estimateur par rapport à une norme d'erreur bien déterminée, soit par l'allure graphique de la courbe (lisse ou non). Parmi ces méthodes, nous citons :

- La méthode de validation croisée.
  - La méthode de validation croisée biaisée.
  - La méthode de validation croisée non biaisée.
  - La méthode validation croisée lissée.
- Les méthodes plug-in (ré-injection).
  - La règle du pouce (Rule Of Thumb).

- Surlissage (Oversmoothing).
- Plug-in itéré.
- Plug-in itéré moderne.
- La méthode de bootstrap.

Dans ce qui suit nous introduisons deux méthodes : la méthode de validation croisée (biaisée et non biaisée) et la méthodes plug-in donnée par La règle du pouce (Rule Of Thumb).

### 2.5.2.1 Méthodes plug-in

La décision d'un choix optimal pour le paramètre de lissage suppose la spécification d'un critère d'erreur qui puisse être optimisé. Dans les méthodes plug-in, Le critère à minimiser est l'Erreur Quadratique Moyenne Intégrée MISE. Dans ce cas, on obtient :

$$MISE(f, \hat{f}_h) = \int_{\mathbb{R}} \mathbb{E}[\hat{f}_h(x) - f(x)]^2 dx = \frac{h_n^4}{4} \sigma_K^4 R(f'') + \frac{R(K)}{nh_n} + o\left(h_n^4 + \frac{1}{n}\right).$$

L'erreur quadratique moyenne intégrée asymptotique est alors :

$$AMISE = \frac{h_n^4}{4} \sigma_K^4 R(f'') + \frac{R(K)}{nh_n}.$$

Le paramètre de lissage  $h^*$  qui minimise l'erreur quadratique moyenne integrée asymptotique est alors de la forme :

$$h^* = \left[ \frac{R(K)}{\sigma_K^4 R(f'')} \right]^{\frac{1}{5}} n^{-\frac{1}{5}}. \tag{2.17}$$

La valeur du AMISE optimale  $AMISE^* = AMISE(h^*)$  est alors de la forme :

$$AMISE^* = \frac{5}{4} [\sigma_K R^4(K) R(f'')]^{\frac{1}{5}} n^{-\frac{4}{5}}.$$

On constate que la largeur de la fenêtre optimale  $h^*$  dépend de la densité inconnue  $f$  à travers le paramètre  $R(f'')$ . Cette largeur de fenêtre 'idéale' (relativement au critère d'erreur retenu) n'est pas donc directement calculable. Une façon classique de remédier à ce problème consiste à remplacer la quantité  $R(f'')$  par un estimateur approprié.

#### ✂ La règle du pouce (Rule of thumb)

L'idée de cette méthode revient à [Deheuvels \(1977\)](#) [18] avant d'être publiée par [Silvermann \(1986\)](#) [45]. Le choix du paramètre de lissage par cette méthode consiste à remplacer la partie inconnue  $R(f'')$  dans l'expression de l'estimateur optimal  $h^*$ , donné

par l'équation (2.17), par une distribution connue afin d'obtenir un estimateur pour  $h$ . Si on choisit  $f$  comme étant la distribution normale de moyenne 0 et de variance  $\sigma^2$ , on obtient :

$$R(f'') = \int (f''(x))^2 dx = \frac{3}{2\sqrt{\pi}} \sigma^{-1/5}. \tag{2.18}$$

De plus, si on utilise un noyau gaussien, alors la valeur de  $h^*$ , notée  $h_{rot}$  est donnée par :

$$\begin{aligned} h_{rot} &= (4\pi)^{-1/10} \left[ \frac{3}{8} \pi^{-1/2} \sigma \right] n^{-1/5} \\ &= \left( \frac{4}{3} \right)^{1/5} \sigma n^{-1/5} \\ &= 1.06 \sigma n^{-1/5}. \end{aligned} \tag{2.19}$$

Il suffit donc d'estimer  $\sigma$  à partir des données et de le remplacer dans la formule (2.19) de  $h_{rot}$ .

• **Avantages et inconvénients des méthodes plug-in**

Pour les méthodes plug-in, plusieurs points importants peuvent être impérativement soulignés. D'abord, cette technique est très satisfaisante théoriquement puisque l'expression de  $h^*$  minimisant le  $MISE(h)$  qui est d'ordre  $O(n^{-1/5})$  (voir l'expression (2.17)). De plus, cette technique nous confine à l'étude de fonctions de distributions suffisamment régulières (d-fois continûment différentiables). Enfin, des difficultés importantes se posent en pratique : pour calculer  $h^*$  il faut, en effet, estimer  $[\sigma_K^4 R(f'')]^{1/5}$  et donc les dérivées de la fonction de densité  $f$  qui s'avère techniquement délicat. De plus, la sélection de la largeur de fenêtre par plug-in n'est valable qu'asymptotiquement. A taille d'échantillon fixée, l'analyse s'avère délicate. Ensuite, les formules qui sont au coeur des méthodes plug-in imposent des restrictions sur la densité inconnue  $f$  souvent difficiles à vérifier dans la pratique. Par exemple, l'expression (2.17) de la largeur de fenêtre asymptotique  $h^*$  n'est pas valide pour une densité cible exponentielle ou uniforme [61].

**2.5.2.2 Méthode Validation Croisée (Cross-Validation CV)**

L'idée de base des méthodes de validation croisée consiste à trouver une fonction de score  $CV(h)$  qui soit plus simple que  $MISE(h)$ .

✂ Méthode Validation Croisée non biaisée (Unbiased Cross- Validation UCV)

*Rundemo* en 1982 [40] et *Bowman* en 1984 [5] ont donné l'idée de cette méthode. En utilisant la formule développée de l'erreur quadratique intégrée  $ISE(h)$ , on choisit le paramètre de lissage  $h$  qui minimise cette erreur.

$$ISE(h) = \int_{\mathbb{R}} [\hat{f}_h(x) - f(x)]^2 dx = \int_{\mathbb{R}} \hat{f}_h^2(x) dx - 2 \int_{\mathbb{R}} \hat{f}_h(x) f(x) dx + \int_{\mathbb{R}} f^2(x) dx. \quad (2.20)$$

Puisque  $\int f^2(x) dx$  ne dépend pas du paramètre de lissage  $h$ . donc on peut choisir le  $h$  de façon à ce qu'il minimise le critère de la validation croisée défini par :

$$UCV(h) = ISE(h) - \int_{\mathbb{R}} f^2(x) dx = \int_{\mathbb{R}} \hat{f}_h^2(x) dx - 2 \int_{\mathbb{R}} \hat{f}_h(x) f(x) dx. \quad (2.21)$$

On doit donc trouver un estimateur pour  $\int_{\mathbb{R}} \hat{f}_h(x) f(x) dx$ . Remarquons que :

$$\int_{\mathbb{R}} \hat{f}_h(x) f(x) dx = \mathbb{E}(\hat{f}_h(x)).$$

Son estimateur empirique est alors :  $\frac{1}{n} \sum_{i=1}^n f_{h,i}(x)$ ,

avec :  $f_{h,i}(x) = \frac{1}{(n-1)h_n} \sum_{\substack{j=1 \\ j \neq i}}^n K\left(\frac{x - X_j}{h_n}\right)$  est l'estimateur de la densité construit à partir

de tous les points sauf le point  $X_i$ .

En remplaçant cet estimateur dans (2.21), pour un noyau  $K$ , le critère de la validation croisée non biaisée  $UCV(h)$  est donné par :

$$UCV(h) = \int_0^{+\infty} \left( \frac{1}{n} \sum_{i=1}^n K_{x,h}(X_i) \right)^2 dx - \frac{2}{n(n-1)} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n K_{X_i,h}(X_j). \quad (2.22)$$

La fenêtre optimale  $h_{UCV}$  obtenue par la méthode de validation croisée non biaisée est donnée par :

$$h_{UCV} = \arg \min_h UCV(h).$$

• **Inconvénients de la méthode validation croisée non biaisée**

Cette méthode présente deux problèmes majeurs (points faibles) : d'une part son manque de robustesse par rapport aux changements de taille de l'échantillon, c'est-à-dire le résultat de simulation peut se révéler extrêmement variable d'un échantillon à l'autre. D'autre part, la fonctionnelle à minimiser a souvent tendance à présenter plusieurs minimums locaux.

✂ **Validation croisée biaisée (Biased Cross-Validation BCV)**

La validation croisée biaisée, a été introduite par *Terrell (1987)* [47] pour remédier aux problèmes de validation croisée non biaisée. Il s'agit d'introduire un biais dans *UCV* afin de réduire sa variance. L'erreur quadratique moyenne intégrée asymptotique étant :

$$AMISE = \frac{h_n^4}{4} \sigma_k^4 R(f'') + \frac{R(K)}{nh_n},$$

le paramètre de lissage basé sur la méthode de validation croisée est la valeur  $h$  qui minimise un estimateur du *AMISE*. On peut estimer le *AMISE* si l'on estime  $f''$  par  $\hat{f}_h''$  où  $\hat{f}_h$  est l'estimateur à noyau de la densité  $f$ .

**Lemme 2.5.1.** (*Scott et Terrell [47]*)

Supposant que le noyau  $K$  satisfait aux conditions suivantes :

$$\int K''(u)du = 0, \int uK''(u)du = 0, \int u^2K''(u)du \neq 0.$$

On obtient le développement asymptotique :

$$\mathbb{E}[R(\hat{f}_h''(x))] = R(f'') + \frac{R(K'')}{nh_n^5} + o(h_n^2).$$

**Proposition 2.5.1.** (*Scott et Terrell [47]*)

Soit  $X_1, X_2, \dots, X_n$  un  $n$ -échantillon, issu d'une variable aléatoire  $X$  de fonction de densité  $f$ . Pour un noyau  $K$  on obtient :

$$BCV(h) = \frac{R(K)}{nh_n} + \frac{h_n^4 \sigma_K^4}{4} \left( R(f_h'') - \frac{R(K'')}{nh_n^5} \right). \tag{2.23}$$

La fenêtre optimale  $h_{BCV}$  obtenue par la méthode de validation croisée biaisée est donnée par :

$$h_{BCV} = \arg \min_h BCV(h).$$

• **Avantages et inconvénients de la méthode validation croisée**

Le principal inconvénient des méthodes validation croisée est que la largeur de fenêtre estimée par cette technique présente une grande variabilité, c'est-à-dire que pour deux échantillons distincts issus de la même distribution, les fenêtres obtenues seront très différentes. Ce problème a été étudié par *Hall* et *Marron* [32] et *Scott* et *Terrell* [47] qui ont montré que,  $h_{ucv}$  et  $h_{bcv}$  convergent très lentement vers  $h^*$ .

Cette méthode présente cependant de nombreux avantages : outre le fait qu'elle ne demande pour être applicable, que des hypothèses faibles sur le degré de différentiabilité de  $f$ , c'est une méthode automatique entièrement guidée par les données. Ce point est particulièrement satisfaisant dans le contexte de comparaison de modèles [61].

## 2.6 Estimation de la densité de probabilité basée sur un nombre aléatoire d'observations

En général, pour estimer une densité de probabilité avec la méthode du noyau, on utilise un échantillon de taille déterministe (non aléatoire). Dans la pratique, il existe des situations où nous devons observer un phénomène sur une durée  $[0, t]$ . Le nombre observations dans le cas est alors une variable aléatoire. Pour cela, l'estimation doit être adaptée à ce cas.

Soit un échantillon  $X_1, X_2, \dots, X_{N_t}$  indépendante issu d'une v.a  $X$  de fonction de densité  $f(x)$ .  $N_t$  représente le nombre d'observations obtenue dans l'intervalle  $[0, t]$ ,  $t > 0$ . Il est clair que  $N_t$  est une variable aléatoire discrète et non négative.

On considère l'estimateur de la densité de *Parzen – Rosenblatt* adapté au cas où la taille de l'échantillon est aléatoire, défini par *Srivastava* (1973) par :

$$\hat{f}_{N_t}(x) = \frac{1}{N_t h(N_t)} \sum_{j=1}^{N_t} K\left(\frac{x - X_j}{h(N_t)}\right). \tag{2.24}$$

où  $h(N_t)$  prend la valeur  $h(n)$  lorsque  $N_t$  prend la valeur  $n$  tel que :

$$h_n \rightarrow 0 \text{ quand } n \rightarrow \infty, \tag{2.25}$$

et  $K(y)$  vérifie :

$$K(y) \geq 0, \int_{-\infty}^{+\infty} K(y)dy = 1, \sup_{y \in \mathbb{R}} K(y) < \infty, K(y) = K(-y), \lim_{y \rightarrow \infty} |yK(y)| \rightarrow 0. \tag{2.26}$$

### 2.6.1 Propriétés statistiques de l'estimateur

#### Théorème 2.10.

Soit  $K$  un noyau de *Parzen – Rosenblatt* vérifiant les conditions (2.26) et  $h$  la fenêtre de lissage de l'estimateur (2.24) vérifiant la condition (2.25).

L'estimateur (2.24) vérifie les propriétés suivantes :

- $Biais(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} \left[ \frac{h_n^2}{2} f''(x) \sigma_K^2 + o(h_n^2) \right] \mathbb{P}(N_t = n).$
- $MSE(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} \left[ \frac{h_n^4}{4} f''^2(x) \sigma_K^4 + \frac{1}{nh_n} f(x) R(K) + o(h_n^4) + o\left(\frac{1}{nh_n}\right) \right] \mathbb{P}(N_t = n).$
- $MISE(\hat{f}_{N_t}) = \sum_{n \geq 0} \left[ \frac{1}{nh} R(K) + \frac{h^4}{4} \sigma_K^4 R(f'') \right] \mathbb{P}(N_t = n).$

Pour la démonstration on utilise le lemme suivant :

#### Lemme 2.6.1.

- $\mathbb{E}(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)) \mathbb{P}(N_t = n).$
- $\mathbb{V}(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)^2) \mathbb{P}(N_t = n) - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2.$
- $MSE(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} MSE(\hat{f}_n(x)) \mathbb{P}(N_t = n).$  (2.27)

- $MISE(\hat{f}_{N_t}) = \sum_{n \geq 0} \left( MISE(\hat{f}_n) \right) \mathbb{P}(N_t = n).$  (2.28)

**Preuve du lemme :**

✂ **L'espérance mathématique :**

$$\begin{aligned}
 \mathbb{E}(\hat{f}_{N_t}(x)) &= \mathbb{E}[\mathbb{E}(\hat{f}_{N_t}(x))/N_t] \\
 &= \sum_{n \geq 0} (\mathbb{E}(\hat{f}_{N_t}(x))/N_t = n) \mathbb{P}(N_t = n) \\
 &= \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)) \mathbb{P}(N_t = n).
 \end{aligned}$$

En utilisant la relation (2.5) on obtient :

$$\mathbb{E}(\hat{f}_{N_t}(x)) = \sum_{n \geq 0} \left[ f(x) + \frac{h_n^2}{2} f''(x) \sigma_K^2 + o(h_n^2) \right] \mathbb{P}(N_t = n).$$

✱ La variance :

$$\begin{aligned}
\mathbb{V}(\hat{f}_{N_t}(x)) &= \mathbb{E} \left[ \left( \hat{f}_{N_t}(x) - \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right] \\
&= \mathbb{E} \left[ \mathbb{E} \left( \left( \hat{f}_{N_t}(x) - \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right) / N_t \right] \\
&= \sum_{n \geq 0} \mathbb{E} \left[ \left( \hat{f}_n(x) - \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right] \mathbb{P}(N_t = n) \\
&= \sum_{n \geq 0} \mathbb{E} \left[ \hat{f}_n(x)^2 - 2\hat{f}_n(x)\mathbb{E}(\hat{f}_{N_t}(x)) + \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right] \mathbb{P}(N_t = n) \\
&= \sum_{n \geq 0} \left[ \mathbb{E}(\hat{f}_n(x)^2) - 2\mathbb{E}(\hat{f}_n(x))\mathbb{E}(\hat{f}_{N_t}(x)) + \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right] \mathbb{P}(N_t = n) \\
&= \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)^2)\mathbb{P}(N_t = n) - 2\mathbb{E}(\hat{f}_{N_t}(x)) \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x))\mathbb{P}(N_t = n) + \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \\
&= \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)^2)\mathbb{P}(N_t = n) - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2.
\end{aligned}$$

$$\begin{aligned}
\mathbb{V}(\hat{f}_{N_t}(x)) &= \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)^2) - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \\
&= \sum_{n \geq 0} \left[ \mathbb{V}(\hat{f}_n(x)) + \left( \mathbb{E}(\hat{f}_n(x)) \right)^2 \right] \mathbb{P}(N_t = n) - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \\
&= \sum_{n \geq 0} \mathbb{V}(\hat{f}_n(x))\mathbb{P}(N_t = n) + \sum_{n \geq 0} \left( \mathbb{E}(\hat{f}_n(x)) \right)^2 \mathbb{P}(N_t = n) - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \\
&= \sum_{n \geq 0} \mathbb{V}(\hat{f}_n(x))\mathbb{P}(N_t = n) + \sum_{n \geq 0} \left[ \left( \mathbb{E}(\hat{f}_n(x)) \right)^2 - \left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 \right] \mathbb{P}(N_t = n).
\end{aligned}$$

Où,

$$\begin{aligned}
\left( \mathbb{E}(\hat{f}_{N_t}(x)) \right)^2 &= \left( \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x))\mathbb{P}(N_t = n) \right)^2 = \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x))^2 \mathbb{P}^2(N_t = n) \\
&\quad + \sum_{n \neq m} \mathbb{E}(\hat{f}_n(x))\mathbb{E}(\hat{f}_m(x))\mathbb{P}(N_t = n)\mathbb{P}(N_t = m).
\end{aligned}$$

Démonstration du théorème :

✂ Le Biais :

$$\begin{aligned}
 \text{Biais}(\hat{f}_{N_t}(x)) &= \mathbb{E}(\hat{f}_{N_t}(x)) - f(x) \\
 &= \sum_{n \geq 0} \mathbb{E}(\hat{f}_n(x)) \mathbb{P}(N_t = n) - f(x) \\
 &= \sum_{n \geq 0} \left( \mathbb{E}(\hat{f}_n(x)) - f(x) \right) \mathbb{P}(N_t = n) \\
 &= \sum_{n \geq 0} \text{Biais}(\hat{f}_n(x)) \mathbb{P}(N_t = n). \tag{2.29}
 \end{aligned}$$

✂ Le MSE :

$$\begin{aligned}
 \text{MSE}(\hat{f}_{N_t}(x)) &= \mathbb{E}(\hat{f}_{N_t}(x) - f(x))^2 \\
 &= \mathbb{E} \left( \mathbb{E}(\hat{f}_{N_t}(x) - f(x))^2 / N_t \right) \\
 &= \sum_{n \geq 0} \left( \mathbb{E}(\hat{f}_{N_t}(x) - f(x))^2 / N_t = n \right) \mathbb{P}(N_t = n) \\
 &= \sum_{n \geq 0} \left( \mathbb{E}(\hat{f}_n(x) - f(x))^2 \right) \mathbb{P}(N_t = n) \\
 &= \sum_{n \geq 0} \text{MSE}(\hat{f}_n(x)) \mathbb{P}(N_t = n).
 \end{aligned}$$

✂ Le MISE :

$$\begin{aligned}
 \text{MISE}(\hat{f}_{N_t}) &= \int \text{MSE}(\hat{f}_{N_t}(x)) dx \\
 &= \int \left( \sum_{n \geq 0} \text{MSE}(\hat{f}_n(x)) \mathbb{P}(N_t = n) \right) dx \\
 &= \sum_{n \geq 0} \left( \int \text{MSE}(\hat{f}_n(x)) \right) \mathbb{P}(N_t = n) dx \\
 &= \sum_{n \geq 0} \text{MISE}(\hat{f}_n) \mathbb{P}(N_t = n).
 \end{aligned}$$

## 2.6.2 Propriétés asymptotiques de l'estimateur

Soit  $\{Y_n\}$  une suite de variables aléatoires. On suppose que des observations  $Y_n$  et :

(C1) : la variable aléatoire  $N_t$  est indépendante.

(C2) :  $N_t \xrightarrow{p} \infty$  quand  $t \rightarrow \infty$ .

**Lemme 2.6.2.** (*Srivastava [44]*)

Supposons que les conditions (C1) et (C2) sont vérifiées :

Si  $\mathbb{E}(Y_n) \rightarrow \theta$  quand  $n \rightarrow \infty$ , alors  $\mathbb{E}(Y_{N_t}) \rightarrow \theta$  quand  $t \rightarrow \infty$ .

En effet,

$$\mathbb{E}(Y_{N_t}) = \mathbb{E}(\mathbb{E}(Y_{N_t})/N_t) = \sum_{n \geq 0} \mathbb{E}(Y_{N_t}/N_t = n)P(N_t = n) = \sum_{n \geq 0} \mathbb{E}(Y_n)P(N_t = n).$$

Si  $\lim_{n \rightarrow +\infty} Y_n = \theta$ , alors pour  $N_t \xrightarrow{p} \infty$  lorsque  $t \rightarrow +\infty$ , on obtient que  $\mathbb{E}Y_{N_t} = \theta$ .

En appliquant ce lemme à l'estimateur  $\hat{f}_{N_t}(x)$  on obtient les résultats du théorème suivant :

**Théorème 2.11.**

Supposons que les conditions (C1) et (C2) sont vérifiées, pour des observations  $(X_n)$ . Les estimateurs (2.24) et (2.3) vérifiant :

- (i)  $\lim_{t \rightarrow \infty} \mathbb{E}(\hat{f}_{N_t}(x)) = \lim_{n \rightarrow \infty} \mathbb{E}(\hat{f}_n(x))$ .
- (ii)  $\lim_{t \rightarrow \infty} \mathbb{V}(\hat{f}_{N_t}(x)) = \lim_{n \rightarrow \infty} \mathbb{V}(\hat{f}_n(x))$ .
- (iii)  $\lim_{t \rightarrow \infty} \text{MSE}(\hat{f}_{N_t}(x)) = \lim_{n \rightarrow \infty} \text{MSE}(\hat{f}_n(x))$ .
- (iv)  $\lim_{t \rightarrow \infty} \text{MISE}(\hat{f}_{N_t}) = \lim_{n \rightarrow \infty} \text{MISE}(\hat{f}_n)$ .

D'après le théorème (2.11), l'estimateur (2.24) est asymptotiquement sans biais et ayant les mêmes propriétés asymptotiques que l'estimateur de (2.3), lorsque  $N_t \xrightarrow{p} \infty$ .

De plus, l'estimateur (2.24) est asymptotiquement normal (*Srivastava [44]*, *Carroll [15]*). Par conséquent,

$$\lim_{t \rightarrow \infty} \mathbb{P}[\sqrt{N_t h(N_t)}(\hat{f}_{N_t} - f(x)) \leq y] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{t^2}{2}} dt,$$

et vérifie la loi du logarithme itérée donnée dans le théorème suivant :

**Théorème 2.12.** (*K.Lagha et S.Adjabi [29]*)

On suppose que  $N_t \xrightarrow{p} \infty$  quand  $t \rightarrow \infty$ . On a alors la convergence en probabilité

$$\limsup_{t \rightarrow \infty} \pm \left\{ \frac{N_t h^2(N_t)}{2\mathbb{E}h(N_t) \log \log N_t} \right\}^{1/2} [\hat{f}_{N_t}(x) - \mathbb{E}(\hat{f}_{N_t}(x))] = \left[ f(x) \int K^2(Z) dz \right]^{1/2}$$

## 2.7 Conclusion

Dans ce chapitre, nous avons présenté l'estimation à noyau de la densité dans le cas d'une taille d'échantillon fixe (donné par *Parzen – Rosenblatt*) et aléatoire (donné par

*Srivastava*). Nous avons donné les propriétés statistiques et asymptotiques de ces estimateurs.

Ces estimateurs seront utilisés pour l'estimation de la fonction du taux de hasard (taux de défaillance), dans le chapitre suivant.

# Chapitre 3

## Estimation des paramètres de la fiabilité et simulation

### 3.1 Introduction

L'analyse statistique des durées de vie constitue un élément important des études sur la fiabilité. Une telle analyse est généralement développée à partir d'un modèle probabiliste (exponentiel, Weibull, Gamma,...), pour la variable temps  $T$ .

Les fonctions de densité  $f$  et de répartition  $F$ , tout comme la fonction caractéristique, décrivent complètement la loi de probabilité des observations et en connaître une estimation convenable permet de résoudre un nombre de problèmes statistiques. Cette estimation tient donc naturellement une place importante dans l'étude de nombreux phénomènes de nature aléatoire, ainsi il existe une relation entre la fiabilité et le taux de hasard à travers la fonction de densité et de répartition. La fiabilité étant  $R(t) = 1 - F(x)$  et le taux de hasard est :  $\lambda(t) = f(x)/R(x)$ ,  $t > 0$ .

Nous commençons par donner l'estimateur empirique de la fonction de répartition et nous l'adaptons au cas où la taille de l'échantillon est aléatoire, pour estimer la fonction de fiabilité. Ensuite, nous utilisons l'estimateur de *Parzen – Rosenblatt* adapté au cas où la taille de l'échantillon est aléatoire pour estimer la fonction du taux de défaillance.

Une application est effectuée par simulation dans le but de connaître le comportement des estimateurs. Pour cela nous avons utilisé le noyau asymétrique Gamma et l'approche classique validation croisée non biaisée pour la sélection du paramètre de lissage.

### 3.2 Estimation empirique de la fonction de répartition

L'estimation de la fonction de répartition d'une variable aléatoire est un volet important de l'estimation non paramétrique. De nombreuses méthodes ont été proposées

et étudiées afin de modifier efficacement l'outil brut qu'est la fonction de répartition. Dans cette partie nous abordons l'estimation empirique de la fonction de répartition et ses propriétés [31].

Soit  $X_1, X_2, \dots, X_n$  un échantillon de variables aléatoires indépendantes et identiquement distribuées à valeurs dans  $\mathbb{R}$ , avec pour fonction de répartition  $F(x)$  telle que :

$$F : x \longmapsto F(x) = \mathbb{P}(X \leq x).$$

L'estimateur naturel de la fonction de répartition  $F(x)$  est la fonction de répartition empirique, notée  $\hat{F}_n(x)$  et définie comme suit :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{]-\infty, x]}(X_i) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}. \quad (3.1)$$

Avec,

$$\mathbf{1}_{]-\infty, x]}(X_i) \begin{cases} 1 & \text{si } X_i \leq x; \\ 0 & \text{sinon.} \end{cases}$$

### 3.2.1 Propriétés statistiques de l'estimateur $\hat{F}_n(x)$

#### ✂ Biais de l'estimateur

$$\text{Biais}(\hat{F}_n(x)) = \mathbb{E}(\hat{F}_n(x)) - F(x) = 0.$$

En effet,

$$\begin{aligned} \mathbb{E}(\hat{F}_n(x)) &= \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{]-\infty, x]}(X_i)\right) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{E}(\mathbf{1}_{]-\infty, x]}(X_i)) \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{P}(X_i \leq x) \\ &= F(x). \end{aligned}$$

Donc, pour tout point  $x$ ,  $\hat{F}_n(x)$  est un estimateur sans biais de  $F(x)$ .

### ✂ Variance de l'estimateur

Il est facile de montrer que pour tout  $x$ , la variance de l'estimateur  $\hat{F}_n(x)$  est donnée par :

$$\begin{aligned}\mathbb{V}(\hat{F}_n(x)) &= \mathbb{V}\left(\frac{1}{n} \sum_{i=1}^n \mathbf{1}_{]-\infty, x]}(X_i)\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{V}(\mathbf{1}_{]-\infty, x]}(X_i)) \\ &= \frac{1}{n} \mathbb{V}(\mathbf{1}_{]-\infty, x]}(X)) \\ &= \frac{F(x) - F^2(x)}{n} \\ &= \frac{F(x)(1 - F(x))}{n}.\end{aligned}$$

On remarque que,  $\mathbb{V}(\hat{F}_n(x)) \longrightarrow 0$  lorsque  $n \longrightarrow +\infty$ .

### 3.2.2 Propriétés asymptotiques de l'estimateur $\hat{F}_n(x)$

#### ✂ L'erreur en moyenne quadratique (MSE)

$$MSE(\hat{F}_n(x)) = \mathbb{E}[(\hat{F}_n(x) - F(x))^2] = \text{biais}^2(\hat{F}_n(x)) + \mathbb{V}(\hat{F}_n(x)) = \mathbb{V}(\hat{F}_n(x)) \longrightarrow 0$$

quand  $n \rightarrow \infty$ .

#### ✂ Convergence en probabilité

$$\hat{F}_n(x) \xrightarrow{p} F(x) \quad \text{quand } n \rightarrow \infty.$$

En effet, d'après l'inégalité de Markov, la convergence en moyenne quadratique implique la convergence en probabilité. C'est à dire,

$$\forall \epsilon > 0, \mathbb{P}(|\hat{F}_n(x) - F(x)| \geq \epsilon) \leq \frac{\mathbb{V}(\hat{F}_n(x))}{\epsilon^2} \longrightarrow 0 \quad \text{quand } n \rightarrow \infty.$$

#### ✂ Convergence presque sûre

D'après la loi des grands nombres, on a :

$$\hat{F}_n(x) \xrightarrow{p.s} F(x) \quad \text{quand } n \rightarrow \infty.$$

### ✠ Convergence en loi

D'après le théorème central limite, on a :

$$\sqrt{n}(\hat{F}_n(x) - F(x)) \xrightarrow{\text{en loi}} \mathcal{N}(0, F(x)(1 - F(x))) \quad \text{quand } n \rightarrow \infty.$$

## 3.3 Estimation de la fiabilité

Étant donné que la fiabilité  $R(x)$  s'écrit en fonction de la fonction de répartition  $F(x)$  par :

$$R(x) = 1 - F(x), \quad x > 0.$$

On considère l'estimateur :

$$\hat{R}_n(x) = 1 - \hat{F}_n(x) = 1 - \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}, \quad x > 0.$$

$$\hat{R}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \geq x\}}, \quad x > 0. \quad (3.2)$$

Les propriétés de cet estimateur découlent de celles de l'estimateur empirique de la fonction de répartition (3.1). On obtient donc que  $\hat{R}_n(x)$  est sans biais, avec  $\mathbb{V}(\hat{R}_n(x)) = \mathbb{V}(\hat{F}_n(x))$ .

L'estimateur est alors convergent en moyenne quadratique, en probabilité, en loi et presque sûrement vers  $R(x)$ .

## 3.4 Estimation du taux de défaillance

Le taux de défaillance  $\lambda(x)$  étant défini par :

$$\lambda(x) = \frac{f(x)}{R(x)}, \quad x > 0.$$

On utilise l'estimateur :

$$\hat{\lambda}_n(x) = \frac{\hat{f}_n(x)}{\hat{R}_n(x)}, \quad x > 0. \quad (3.3)$$

Où  $\hat{f}_n(x)$  est l'estimateur de *Parzen – Rosenblatt* de la fonction de densité défini dans la formule (2.3) et  $\hat{R}_n(x)$  l'estimateur défini dans (3.2).

### 3.5 Estimation dans le cas d'une taille d'échantillon aléatoire

En considérant l'échantillon  $X_1, X_2, \dots, X_{N_t}$  issu de la *v. a*  $X$  de densité  $f$  et de fonction de répartition  $F$ , tel que  $N_t$  est une *v. a* à valeur entières, représentant le nombre d'observations enregistrés dans l'intervalle  $[0, t]$ . On considère les estimateurs suivants :

#### 3.5.1 Fonction de répartition

L'estimateur (3.1) peut être adapté au cas où la taille de l'échantillon est aléatoire et s'écrit alors [44] :

$$\begin{aligned}\hat{F}_{N_t}(x) &= \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{1}_{\{X_i \leq x\}} \\ &= \frac{[\text{Nombre d'observations } X_i \leq x]}{N_t}, \quad x > 0.\end{aligned}\tag{3.4}$$

L'estimateur (3.4) vérifie les propriétés suivantes :

##### Propriété 3.5.1.

Pour  $N_t \xrightarrow{P} \infty$ ,  $t \rightarrow \infty$ , l'estimateur (3.4) est :

- (i) sans biais,
- (ii) consistant,
- (iii) asymptotiquement normal.

#### 3.5.2 Fonction de fiabilité

L'estimateur (3.2) peut être adapté au cas où la taille de l'échantillon est aléatoire et s'écrit alors :

$$\begin{aligned}\hat{R}_{N_t}(x) &= 1 - \hat{F}_{N_t}(x) \\ &= \frac{1}{N_t} \sum_{i=1}^{N_t} \mathbf{1}_{\{X_i > x\}} \\ &= \frac{[\text{Nombre d'observations } X_i > x]}{N_t}, \quad x > 0.\end{aligned}\tag{3.5}$$

Dont les propriétés statistiques et asymptotiques découlent de celles de  $\hat{F}_{N_t}(x)$ .

### 3.5.3 Fonction du taux de défaillance

L'estimateur (3.3) peut être adapté au cas où la taille de l'échantillon est aléatoire et s'écrit :

$$\hat{\lambda}_{N_t}(x) = \frac{\hat{f}_{N_t}(x)}{\hat{R}_{N_t}(x)}, \quad x > 0. \quad (3.6)$$

Où  $\hat{f}_{N_t}(x)$  est l'estimateur de densité définie dans la formule (2.24) et  $\hat{R}_{N_t}(x)$  est l'estimateur de la fiabilité définie dans la formule (3.5).

## 3.6 Simulation

Dans le cadre de l'application, nous sommes intéressés, en pratique, à obtenir des estimations de la fiabilité et du taux de hasard pour les valeurs de  $x$  comprises entre 0 et  $t$ , où  $t$  est fixé (*Srivastava* [44]). Les estimateurs donnés par les formules (3.2) et (3.3) permettent d'estimer ces paramètres pour  $x$  dans  $\mathbb{R}$ . Ils seront alors modifiés de sorte que nous établissons des estimations pour  $x$  dans  $[0, t]$ .

On obtient donc des nouvelles fonctions,

$$R^*(x) = 1 - F^*(x),$$

et

$$\lambda^*(x) = \frac{f^*(x)}{1 - F^*(x)}.$$

Où,

$$F^*(x) = \frac{F(x)}{F(t)} \quad \text{et} \quad f^*(x) = \frac{f(x)}{F(t)} \quad \text{pour} \quad x \in [0, t].$$

D'où les estimateurs,

$$\hat{R}_{N_t}^*(x) = \frac{\hat{F}_{N_t}(t) - \hat{F}_{N_t}(x)}{\hat{F}_{N_t}(t)}, \quad x \in [0, t].$$

Et

$$\hat{\lambda}_{N_t}^*(x) = \frac{\hat{f}_{N_t}(x)}{\hat{F}_{N_t}(t) - \hat{F}_{N_t}(x)}, \quad x \in [0, t].$$

Où  $\hat{f}_{N_t}$  et  $\hat{F}_{N_t}$  sont les estimateurs (2.24) et (3.4) respectivement.

### 3.6.1 Plan de simulation

Nous allons utiliser quelques méthodes présentés dans les chapitres précédents afin de simuler différentes lois pour étudier le comportement et l'allure de la fiabilité et du taux de hasard :

**La fonction noyau :** Nous avons utilisé le noyau Gamma de paramètre  $(\frac{x}{h} + 1, h)$ , puisque le support des données est  $[0, \infty[$ .

**Le paramètre de lissage :** le paramètre de lissage est sélectionné par le biais d'une approche classique en utilisant la validation croisée non biaisée **UCV**.

**Lois de fiabilités :** nous construisons les échantillons simulés à partir des densités cibles suivantes :

- **La loi Log-normal :** de paramètre  $(\mu, \sigma^2)$ .
- **La loi Weibull :** de paramètre  $(\beta, \alpha)$  à différents stades de vie : Jeunesse ( $\beta < 1$ ), maturité ( $\beta = 1$ ) et Vieillesse ( $\beta > 1$ ).
- **La loi Gamma :** de paramètre  $(\alpha, \beta)$ .

Nous avons utilisé le logiciel *R* pour la programmation des estimateurs.

### 3.6.2 Algorithme de simulation

Afin de visualiser l'allure de la fiabilité et du taux de hasard des différentes lois usuelles souvent utilisées en fiabilité, nous suivons les étapes suivantes :

1. Donner l'horizon de simulation  $T$ .
2. Générer les durées de vies (les durées de bon fonctionnement jusqu'à la première panne) suivant les lois suivantes : Gamma, Log-normale et Weibull.
3. Générer  $N_t$  qui représente le nombre de pannes observées durant une période  $[0, T]$  suivant la loi de poisson de paramètre  $\lambda T$ .
4. Utiliser le noyau Gamma et la fenêtre de lissage  $h$  en utilisant la validation croisée non biaisée **UCV**.
5. Introduire les points cibles  $x$ .
6. Calculer la valeur théorique de la densité  $f(x)$  et ses estimateurs  $\hat{f}_n(x)$  et  $\hat{f}_{N_t}(x)$ .
7. Calculer la valeur théorique de la fonction de répartition  $F(x)$  et ses estimateurs  $\hat{F}_n(x)$  et  $\hat{F}_{N_t}(x)$ .
8. Calculer la valeur théorique de la fiabilité  $R(x)$  et ses estimateurs  $\hat{R}_n(x)$  et  $\hat{R}_{N_t}^*(x)$ .

9. Calculer la valeur théorique du taux de hasard  $\lambda(x)$  et ses estimateurs  $\hat{\lambda}_n(x)$  et  $\hat{\lambda}^*_{N_t}(x)$ .

10. Calculer le ISE de la fiabilité et du taux de hasard, tel que :

$$ISE = \int (\hat{f}(x) - f(x))^2 dx.$$

11. Tracer les graphes correspondant à  $f(x)$ ,  $R(x)$ ,  $\lambda(x)$  et leurs estimateurs  $\hat{f}_n(x)$ ,  $\hat{f}^*_{N_t}(x)$ ,  $\hat{R}_n(x)$ ,  $\hat{R}^*_{N_t}(x)$ ,  $\hat{\lambda}_n(x)$  et  $\hat{\lambda}^*_{N_t}(x)$ .

En se basant sur l’algorithme décrit ci-dessus, nous simulons les lois usuelles en fiabilité citées auparavant en variant l’horizon  $T = (50, 200, 2000)$ .

**Remarque 3.6.1.**

L’étape numéro trois concerne uniquement le cas où la taille de l’échantillon est aléatoire.

Les résultats obtenus par la simulation sont donnés sous forme de tableaux et graphiques. Ce qui nous permet de visualiser le comportement et l’allure de la densité de défaillance, la fiabilité et le taux de hasard qui diffère selon la loi retenue. Les graphiques permettent de faire une comparaison entre les résultats théoriques et les résultats obtenus par estimation. les résultats de simulation nous permettent aussi de faire une comparaison entre les résultats obtenus quand la taille de l’échantillon est fixe et celle quand elle est aléatoire.

**3.6.3 Résultats de simulation : cas où la taille de l’échantillon est fixe**

• **Loi Gamma  $\mathcal{G}(9, 2.9)$**

$n$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
		$R(x)$	$\hat{R}(x)$	ISE	$\lambda(x)$	$\hat{\lambda}(x)$	ISE	$f(x)$	$\hat{f}(x)$	ISE
50	2.963	0.812	0.72	$1.210.10^{-2}$	0.404	0.367	$2.429.10^{-2}$	0.329	0.264	$2.429.10^{-2}$
	3.020	0.487	0.44		0.802	0.668		0.391	0.294	
	3.444	0.334	0.30		0.979	0.908		0.327	0.272	
200	2.963	0.812	0.815	$1.077.10^{-3}$	0.404	0.350	$2.171.10^{-2}$	0.329	0.285	$5.159.10^{-3}$
	3.020	0.487	0.510		0.802	0.665		0.391	0.339	
	3.444	0.334	0.310		0.979	0.981		0.327	0.304	
2000	2.963	0.812	0.805	$7.478.10^{-5}$	0.404	0.396	$4.909.10^{-4}$	0.329	0.319	$2.558.10^{-4}$
	3.020	0.487	0.485		0.802	0.783		0.391	0.379	
	3.444	0.334	0.334		0.979	0.973		0.327	0.324	

TAB. 3.1 – Résultats de simulation de la loi Gamma de paramètre  $\mathcal{G}(9, 2.9)$  (cas fixe).

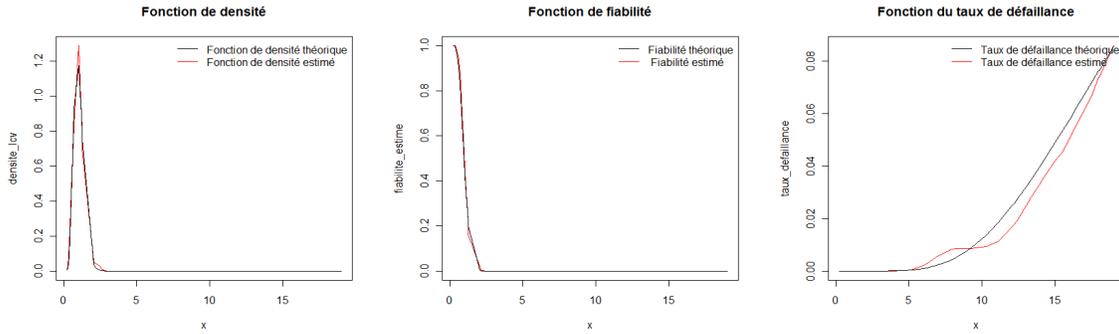


FIG. 3.1 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour n=50.

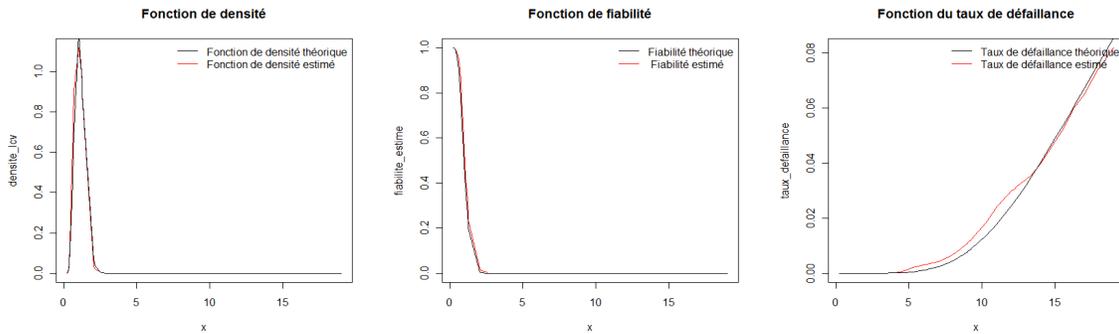


FIG. 3.2 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour n=200.

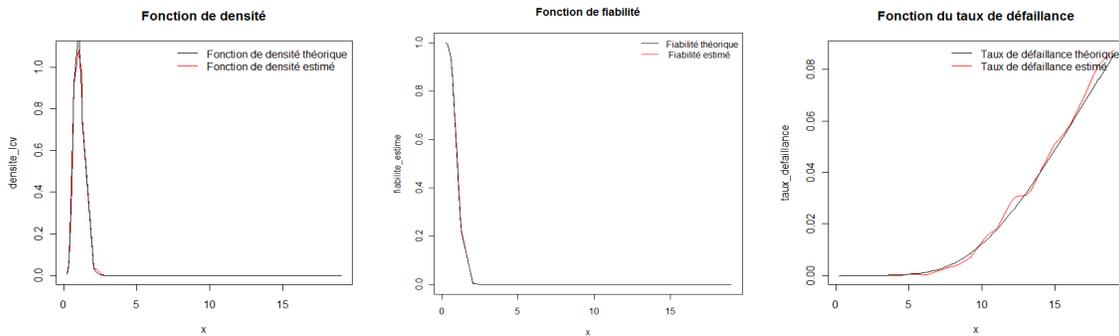


FIG. 3.3 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour n=2000.

• Loi Log-normale  $Log\mathcal{N}(0, 1)$

n	x	Fonction de fiabilité			Taux de hasard			Fonction de densité		
		R(x)	$\hat{R}(x)$	ISE	$\lambda(x)$	$\hat{\lambda}(x)$	ISE	f(x)	$\hat{f}(x)$	ISE
50	2.963	0.218	0.20	$1.120 \cdot 10^{-3}$	0.620	0.614	$5.962 \cdot 10^{-3}$	0.135	0.122	$2.123 \cdot 10^{-4}$
	3.020	0.134	0.16		0.533	0.457		0.071	0.073	
	3.444	0.108	0.12		0.498	0.513		0.053	0.061	
200	2.963	0.218	0.205	$4.359 \cdot 10^{-4}$	0.620	0.618	$2.696 \cdot 10^{-3}$	0.135	0.126	$7.951 \cdot 10^{-5}$
	3.020	0.134	0.125		0.533	0.562		0.071	0.070	
	3.444	0.108	0.095		0.498	0.541		0.053	0.051	
2000	2.963	0.218	0.218	$3.154 \cdot 10^{-5}$	0.620	0.625	$8.508 \cdot 10^{-4}$	0.135	0.136	$7.226 \cdot 10^{-6}$
	3.020	0.134	0.139		0.533	0.504		0.071	0.070	
	3.444	0.108	0.111		0.498	0.500		0.053	0.055	

TAB. 3.2 – Résultats de simulation de la loi Log-normale de paramètre  $Log\mathcal{N}(0, 1)$  (cas fixe).

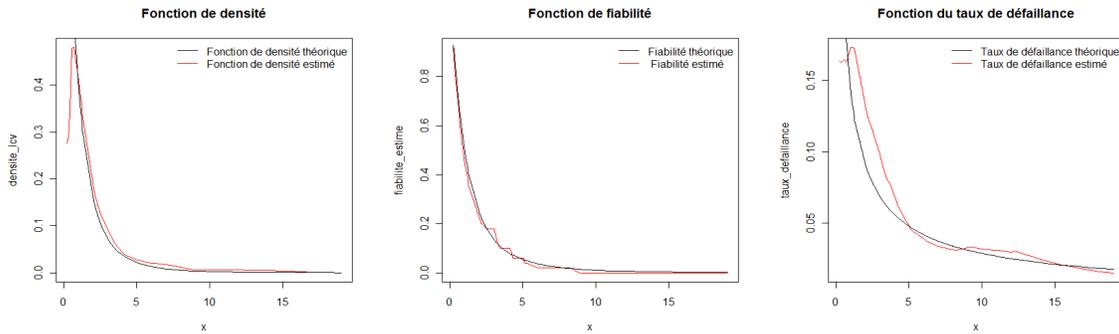


FIG. 3.4 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour n=50.

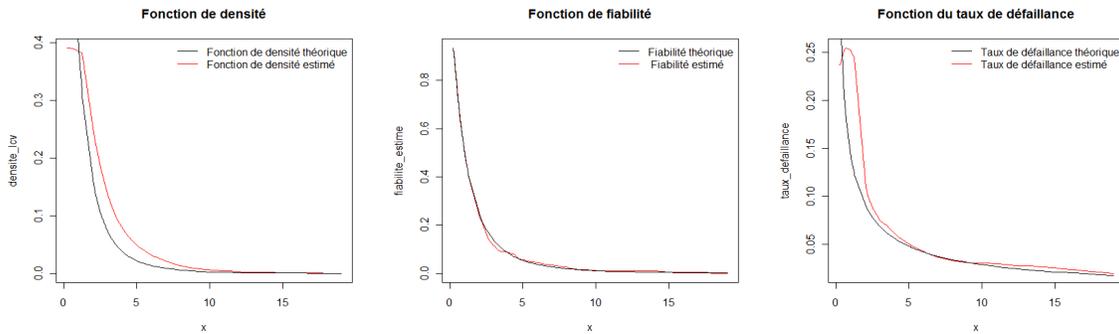


FIG. 3.5 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour n=200.

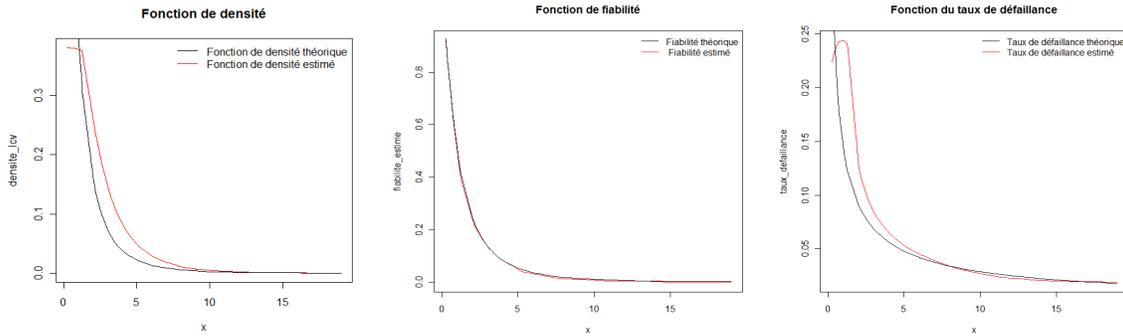


FIG. 3.6 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour n=2000.

• La loi Weibull  $\mathcal{W}(0.5, 5)$

n	x	Fonction de fiabilité			Taux de hasard			Fonction de densité		
		R(x)	$\hat{R}(x)$	ISE	$\lambda(x)$	$\hat{\lambda}(x)$	ISE	f(x)	$\hat{f}(x)$	ISE
50	2.963	0.516	0.54	$1.520 \cdot 10^{-3}$	0.151	0.188	$3.124 \cdot 10^{-3}$	0.078	0.101	$8.020 \cdot 10^{-4}$
	3.020	0.459	0.48		0.128	0.095		0.059	0.045	
	3.444	0.436	0.46		0.120	0.094		0.052	0.043	
200	2.963	0.516	0.485	$1.231 \cdot 10^{-3}$	0.151	0.146	$9.957 \cdot 10^{-4}$	0.078	0.071	$2.915 \cdot 10^{-4}$
	3.020	0.459	0.445		0.128	0.100		0.059	0.044	
	3.444	0.436	0.435		0.120	0.134		0.052	0.058	
2000	2.963	0.516	0.515	$1.605 \cdot 10^{-5}$	0.151	0.157	$8.074 \cdot 10^{-4}$	0.078	0.081	$1.440 \cdot 10^{-4}$
	3.020	0.459	0.458		0.128	0.139		0.059	0.063	
	3.444	0.436	0.432		0.120	0.146		0.052	0.063	

TAB. 3.3 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 0.5$  et  $\alpha = 5$  (cas fixe).

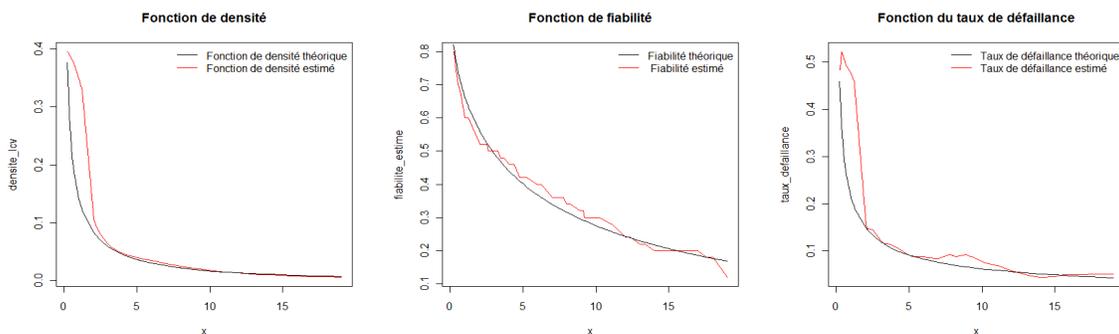


FIG. 3.7 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour n=50.

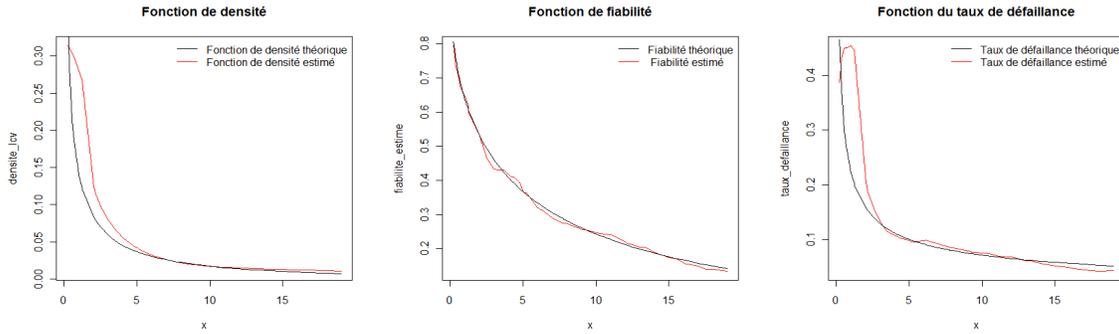


FIG. 3.8 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour n=200.

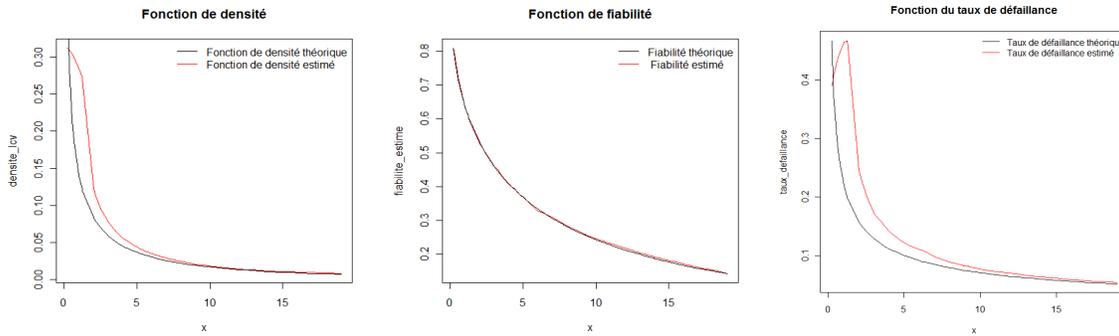


FIG. 3.9 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour n=2000.

• Loi Weibull  $\mathcal{W}(1, 5)$

n	x	Fonction de fiabilité			Taux de hasard			Fonction de densité		
		R(x)	$\hat{R}(x)$	ISE	$\lambda(x)$	$\hat{\lambda}(x)$	ISE	f(x)	$\hat{f}(x)$	ISE
50	2.963	0.646	0.660	5.102.10 <sup>-3</sup>	0.2	0.175	1.234.10 <sup>-3</sup>	0.129	0.116	1.817.10 <sup>-4</sup>
	3.020	0.546	0.480		0.2	0.222		0.109	0.107	
	3.444	0.502	0.480		0.2	0.211		0.100	0.101	
200	2.963	0.646	0.695	4.860.10 <sup>-3</sup>	0.2	0.187	1.715.10 <sup>-4</sup>	0.129	0.130	6.927.10 <sup>-5</sup>
	3.020	0.546	0.580		0.2	0.198		0.109	0.114	
	3.444	0.502	0.540		0.2	0.197		0.100	0.106	
2000	2.963	0.646	0.644	5.731.10 <sup>-5</sup>	0.2	0.204	8.162.10 <sup>-5</sup>	0.129	0.131	1.532.10 <sup>-5</sup>
	3.020	0.546	0.547		0.2	0.204		0.109	0.111	
	3.444	0.502	0.595		0.2	0.206		0.100	0.102	

TAB. 3.4 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 1$  et  $\alpha = 5$  (cas fixe).

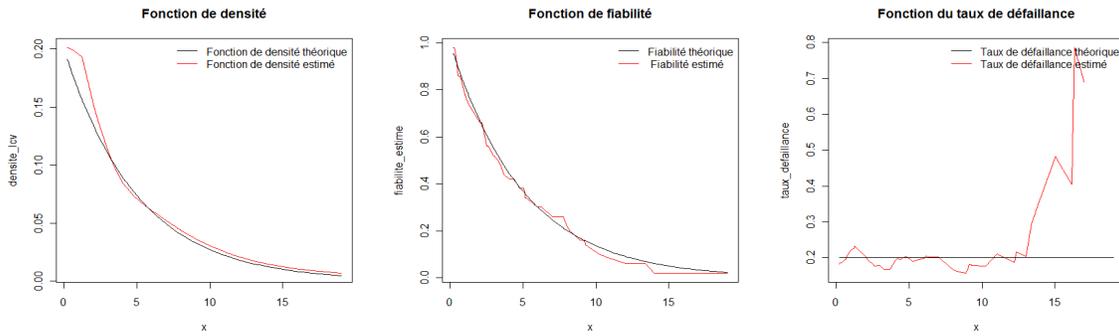


FIG. 3.10 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n=50$ .

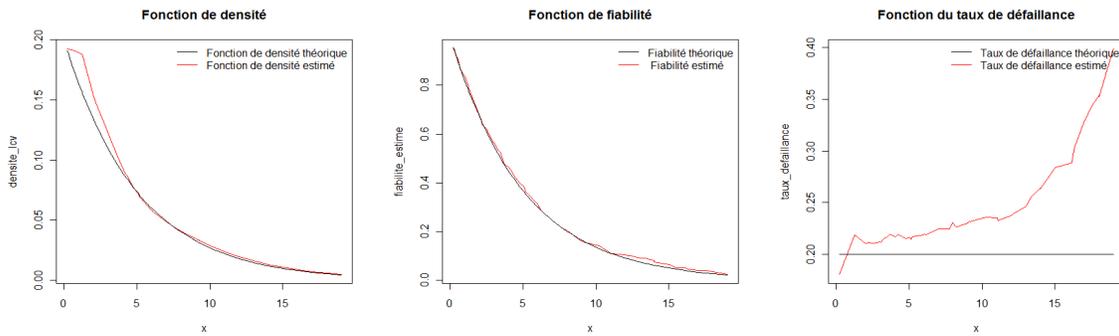


FIG. 3.11 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n=200$ .

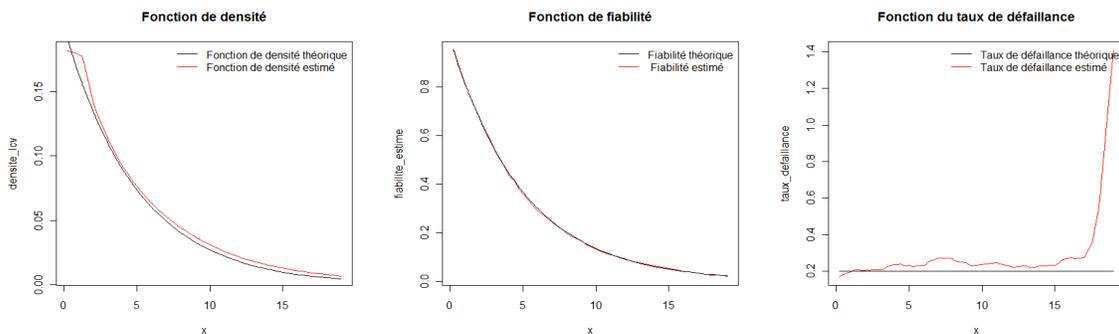


FIG. 3.12 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n=2000$ .

• La loi Weibull  $\mathcal{W}(7, 5)$

n	x	Fonction de fiabilité			Taux de hasard			Fonction de densité		
		R(x)	$\hat{R}(x)$	ISE	$\lambda(x)$	$\hat{\lambda}(x)$	ISE	f(x)	$\hat{f}(x)$	ISE
50	2.963	0.997	1	2.073.10 <sup>-4</sup>	0.009	0.0006	7.323.10 <sup>-4</sup>	0.009	0.0006	6.372.10 <sup>-4</sup>
	3.020	0.971	0.98		0.067	0.084		0.066	0.082	
	3.444	0.929	0.94		0.149	0.129		0.138	0.121	
200	2.963	0.997	0.995	2.017.10 <sup>-5</sup>	0.009	0.013	5.782.10 <sup>-4</sup>	0.009	0.133	5.190.10 <sup>-4</sup>
	3.020	0.971	0.975		0.067	0.076		0.066	0.075	
	3.444	0.929	0.930		0.149	0.171		0.138	0.159	
2000	2.963	0.997	0.997	4.699.10 <sup>-5</sup>	0.009	0.010	1.706.10 <sup>-5</sup>	0.009	0.010	9.241.10 <sup>-6</sup>
	3.020	0.971	0.973		0.067	0.068		0.066	0.067	
	3.444	0.929	0.935		0.149	0.145		0.138	0.136	

TAB. 3.5 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 7$  et  $\alpha = 5$  (cas fixe).

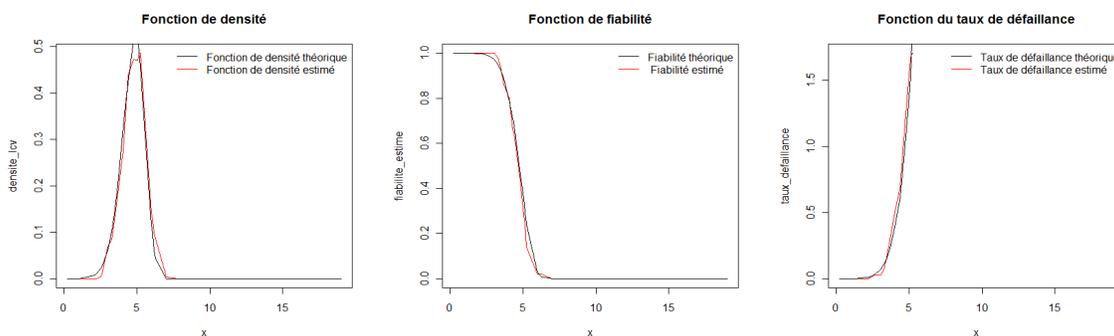


FIG. 3.13 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour n=50.

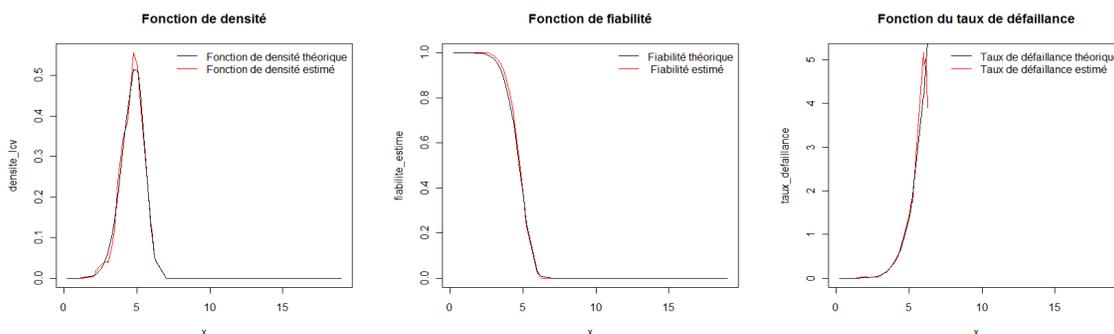


FIG. 3.14 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour n=200.

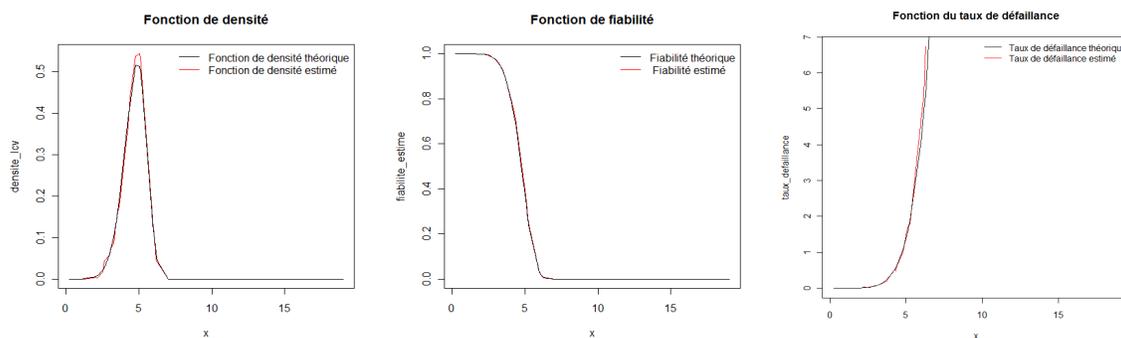


FIG. 3.15 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour  $n=2000$ .

### 3.6.4 Résultats de simulation : cas où la taille de l'échantillon est aléatoire

- La loi Gamma  $\mathcal{G}(9, 2.9)$ ,  $N_t \rightsquigarrow \mathcal{P}(0.5)$

Horizon	$N_t$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
			$R(x)$	$\hat{R}^*(x)$	ISE	$\lambda(x)$	$\lambda^*(x)$	ISE	$f(x)$	$f^*(x)$	ISE
[1, 50]	27	2.963	0.812	0.777	$3.128 \cdot 10^{-3}$	0.404	0.303	$6.932 \cdot 10^{-2}$	0.329	0.236	$2.402 \cdot 10^{-2}$
		3.020	0.487	0.444		0.802	0.636		0.391	0.282	
		3.444	0.334	0.333		0.979	0.801		0.327	0.267	
[1, 200]	105	2.963	0.812	0.857	$2.723 \cdot 10^{-3}$	0.404	0.315	$3.500 \cdot 10^{-2}$	0.329	0.270	$6.350 \cdot 10^{-3}$
		3.020	0.487	0.514		0.802	0.667		0.391	0.343	
		3.444	0.334	0.342		0.979	0.884		0.327	0.303	
[1, 2000]	1017	2.963	0.812	0.814	$6.757 \cdot 10^{-4}$	0.404	0.387	$2.221 \cdot 10^{-2}$	0.329	0.315	$6.971 \cdot 10^{-4}$
		3.020	0.487	0.476		0.802	0.831		0.391	0.396	
		3.444	0.334	0.310		0.979	1.124		0.327	0.349	

TAB. 3.6 – Résultats de simulation de la loi Gamma de paramètre  $\mathcal{G}(9, 2.9)$  (cas aléatoire).

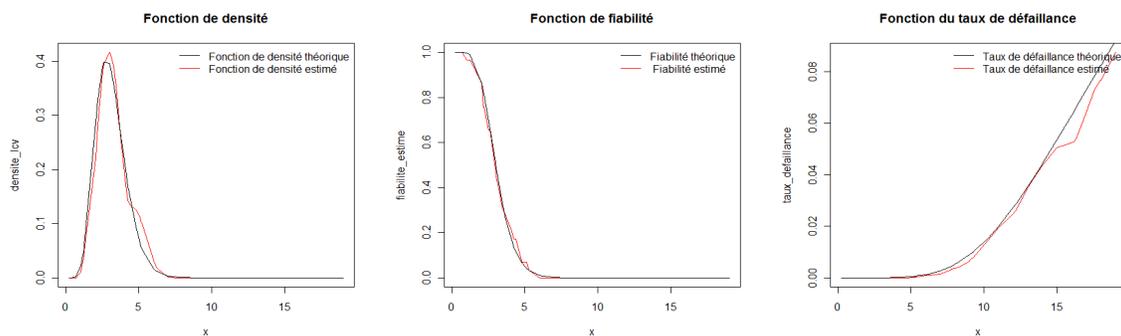


FIG. 3.16 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour  $n = 27$  sur l'intervalle  $[0, 50]$ .

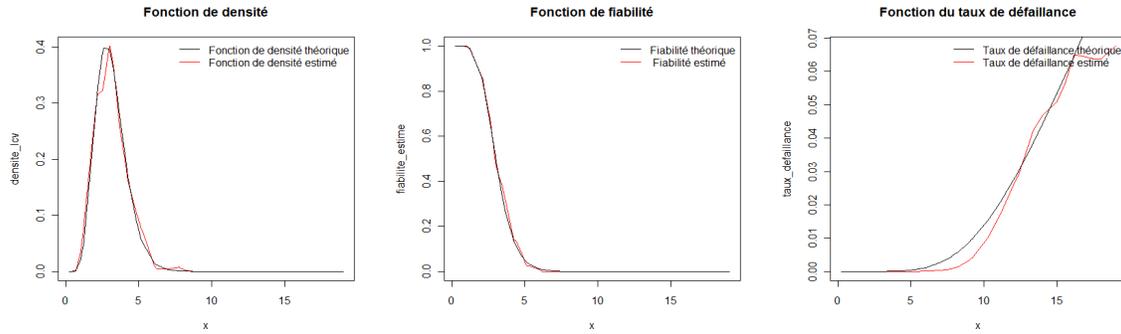


FIG. 3.17 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour  $n = 105$  sur l'intervalle  $[0, 200]$ .

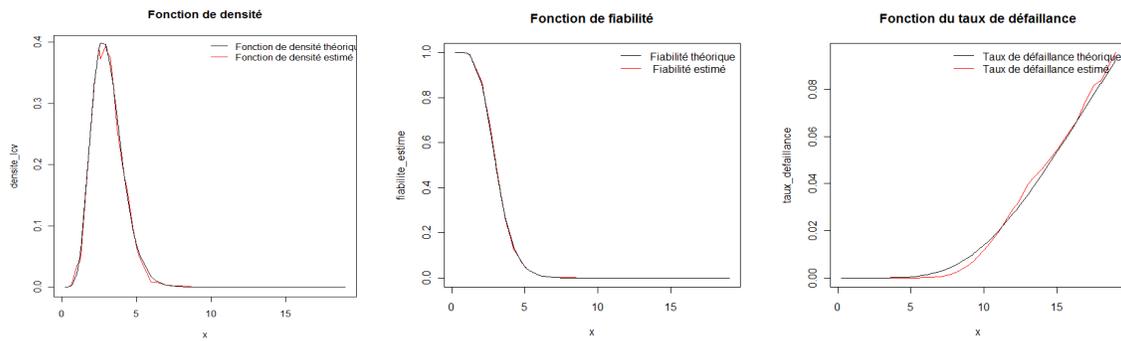


FIG. 3.18 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Gamma (9, 2.9) pour  $n = 1017$  sur l'intervalle  $[0, 2000]$ .

• La loi Log-normale  $\text{Log}\mathcal{N}(0, 1)$ ,  $N_t \rightsquigarrow \mathcal{P}(0.5)$

Horizon	$N_t$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
			$R(x)$	$\hat{R}^*(x)$	ISE	$\lambda(x)$	$\hat{\lambda}^*(x)$	ISE	$f(x)$	$\hat{f}^*(x)$	ISE
[1, 50]	23	2.963	0.218	0.173	$2.473 \cdot 10^{-3}$	0.620	0.703	$1.574 \cdot 10^{-2}$	0.135	0.122	$2.074 \cdot 10^{-4}$
		3.020	0.134	0.130		0.533	0.500		0.071	0.065	
		3.444	0.108	0.130		0.498	0.411		0.053	0.053	
[1, 200]	100	2.963	0.218	0.22	$3.691 \cdot 10^{-5}$	0.620	0.691	$6.129 \cdot 10^{-3}$	0.135	0.152	$2.863 \cdot 10^{-4}$
		3.020	0.134	0.14		0.533	0.501		0.071	0.070	
		3.444	0.108	0.11		0.498	0.492		0.053	0.054	
[1, 2000]	1018	2.963	0.218	0.214	$2.560 \cdot 10^{-5}$	0.620	0.598	$2.227 \cdot 10^{-3}$	0.135	0.128	$9.508 \cdot 10^{-5}$
		3.020	0.134	0.132		0.533	0.497		0.071	0.065	
		3.444	0.108	0.107		0.498	0.519		0.053	0.055	

TAB. 3.7 – Résultats de simulation de la loi Log-normale de paramètre  $\text{Log}\mathcal{N}(0, 1)$  (cas aléatoire).

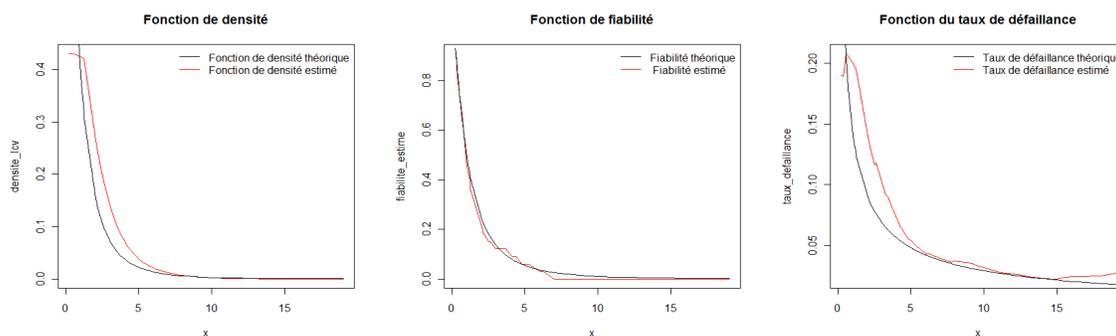


FIG. 3.19 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour  $n = 23$  sur l'intervalle  $[0, 50]$ .

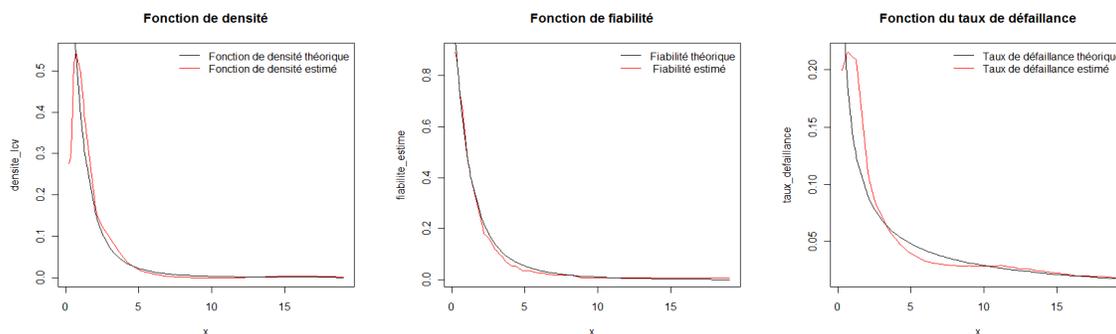


FIG. 3.20 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour  $n = 100$  sur l'intervalle  $[0, 200]$ .

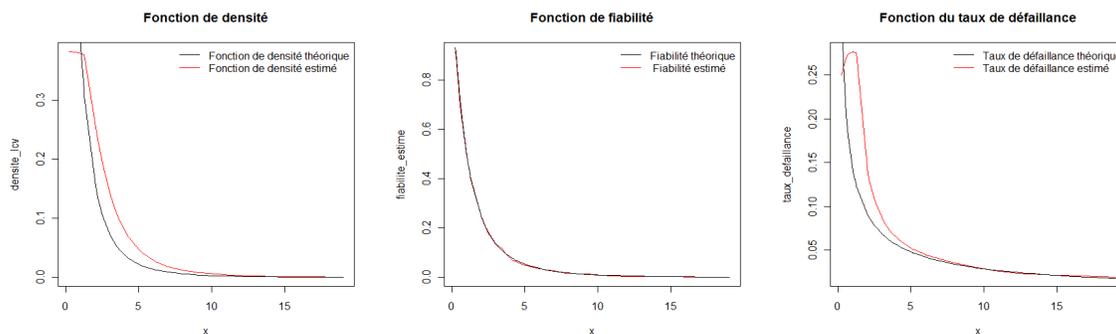


FIG. 3.21 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale (0, 1) pour  $n = 1018$  sur l’intervalle  $[0, 2000]$ .

• La loi Weibull  $\mathcal{W}(0.5, 5)$ ,  $N_t \rightsquigarrow \mathcal{P}(0.5)$

Horizon	$N_t$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
			$R(x)$	$\hat{R}^*(x)$	ISE	$\lambda(x)$	$\hat{\lambda}^*(x)$	ISE	$f(x)$	$\hat{f}^*(x)$	ISE
[1, 50]	30	2.963	0.516	0.566	$8.191 \cdot 10^{-3}$	0.151	0.163	$7.864 \cdot 10^{-4}$	0.078	0.092	$2.273 \cdot 10^{-4}$
		3.020	0.459	0.500		0.128	0.118		0.059	0.059	
		3.444	0.436	0.500		0.120	0.097		0.052	0.048	
[1, 200]	107	2.963	0.516	0.532	$2.494 \cdot 10^{-4}$	0.151	0.151	$1.060 \cdot 10^{-3}$	0.078	0.096	$7.779 \cdot 10^{-5}$
		3.020	0.459	0.495		0.128	0.125		0.059	0.059	
		3.444	0.436	0.467		0.120	0.120		0.052	0.049	
[1, 2000]	1017	2.963	0.516	0.521	$3.228 \cdot 10^{-5}$	0.151	0.157	$2.925 \cdot 10^{-4}$	0.078	0.081	$6.610 \cdot 10^{-5}$
		3.020	0.459	0.461		0.128	0.142		0.059	0.065	
		3.444	0.436	0.432		0.120	0.128		0.052	0.055	

TAB. 3.8 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 0.5$ ,  $\alpha = 5$  (cas aléatoire).

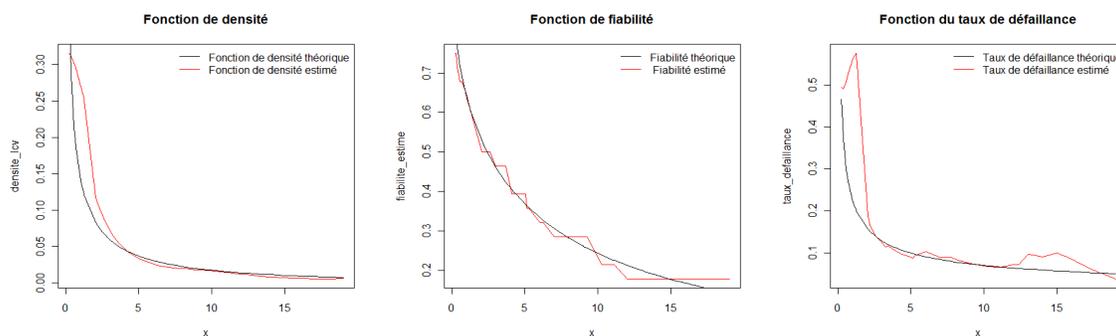


FIG. 3.22 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour  $n = 30$  sur l’intervalle  $[0, 50]$ .

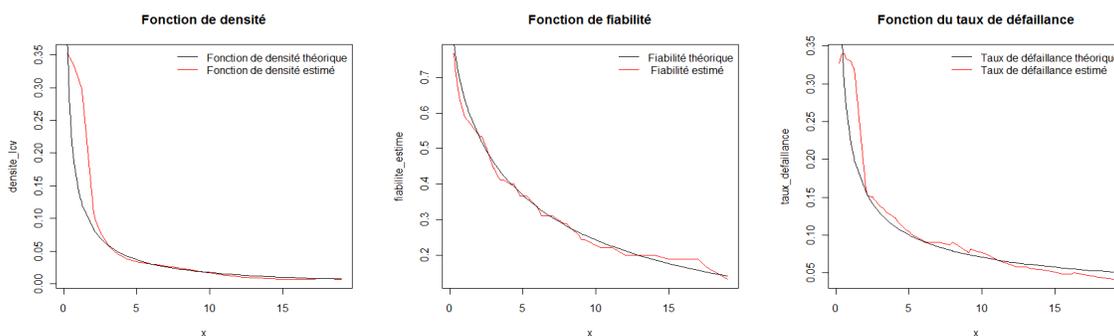


FIG. 3.23 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour  $n = 107$  sur l'intervalle  $[0, 200]$ .

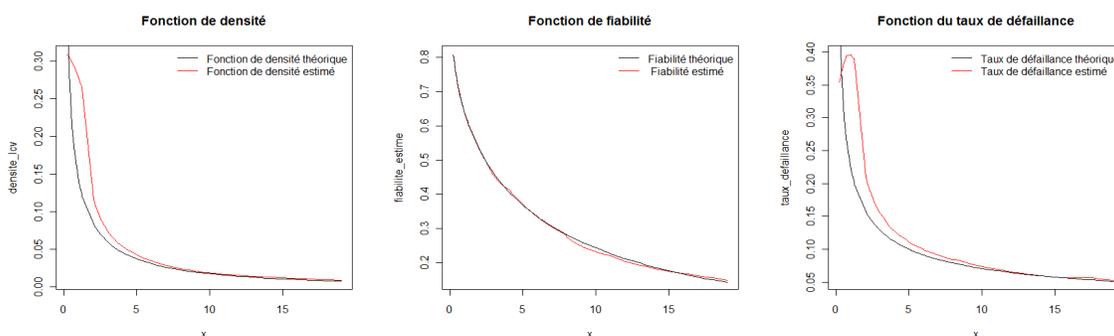


FIG. 3.24 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) pour  $n = 1017$  sur l'intervalle  $[0, 2000]$ .

• La loi Weibull  $\mathcal{W}(1, 5)$ ,  $N_t \rightsquigarrow \mathcal{P}(0.5)$

Horizon	$N_t$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
			$R(x)$	$\hat{R}^*(x)$	ISE	$\lambda(x)$	$\hat{\lambda}^*(x)$	ISE	$f(x)$	$\hat{f}^*(x)$	ISE
[1, 50]	24	2.963	0.646	0.667	$1.975 \cdot 10^{-3}$	0.2	0.176	$6.074 \cdot 10^{-4}$	0.129	0.117	$2.034 \cdot 10^{-4}$
		3.020	0.546	0.541		0.2	0.206		0.109	0.111	
		3.444	0.502	0.541		0.2	0.199		0.100	0.107	
[1, 200]	103	2.963	0.646	0.660	$2.304 \cdot 10^{-4}$	0.2	0.198	$2.426 \cdot 10^{-4}$	0.129	0.131	$9.003 \cdot 10^{-5}$
		3.020	0.546	0.553		0.2	0.208		0.109	0.115	
		3.444	0.502	0.505		0.2	0.212		0.100	0.107	
[1, 2000]	1029	2.963	0.646	0.634	$1.861 \cdot 10^{-4}$	0.2	0.198	$4.048 \cdot 10^{-5}$	0.129	0.129	$3.027 \cdot 10^{-5}$
		3.020	0.546	0.541		0.2	0.196		0.109	0.109	
		3.444	0.502	0.499		0.2	0.196		0.100	0.104	

TAB. 3.9 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 1$ ,  $\alpha = 5$  (cas aléatoire).

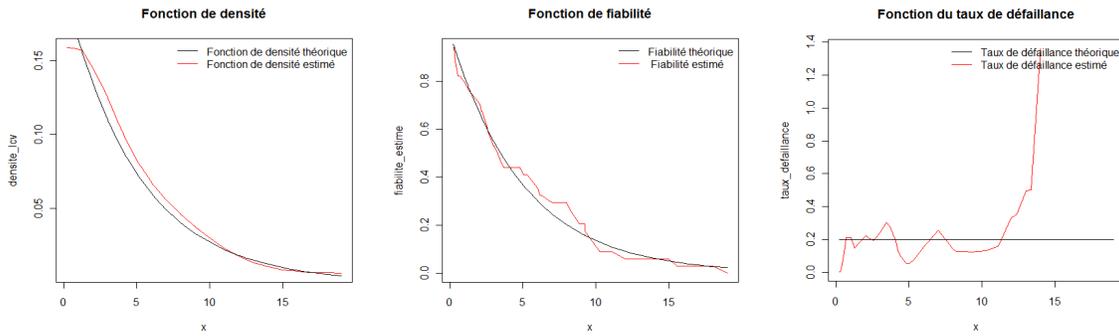


FIG. 3.25 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n = 24$  sur l'intervalle  $[0, 50]$ .

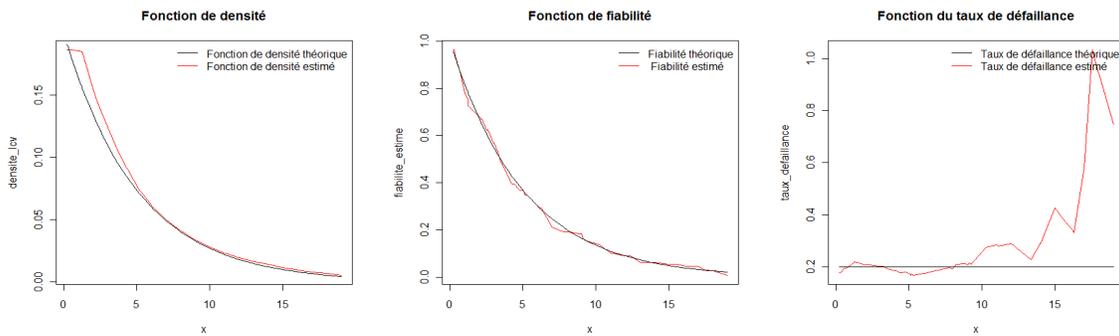


FIG. 3.26 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n = 103$  sur l'intervalle  $[0, 200]$ .

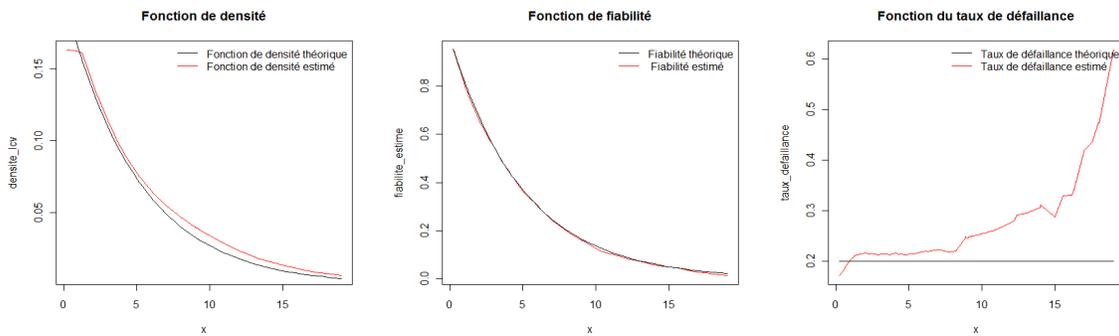


FIG. 3.27 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) pour  $n = 1029$  sur l'intervalle  $[0, 2000]$ .

• La loi Weibull  $\mathcal{W}(7, 5)$ ,  $N_t \rightsquigarrow \mathcal{P}(0.5)$

Horizon	$N_t$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
			$R(x)$	$\hat{R}^*(x)$	ISE	$\lambda(x)$	$\hat{\lambda}^*(x)$	ISE	$f(x)$	$\hat{f}^*(x)$	ISE
[1, 50]	20	2.963	0.997	1.000	$5.909.10^{-3}$	0.009	0.013	$2.677.10^{-3}$	0.009	0.013	$2.040.10^{-3}$
		3.020	0.971	0.9		0.067	0.112		0.066	0.100	
		3.444	0.929	0.9		0.149	0.123		0.138	0.110	
[1, 200]	87	2.963	0.997	0.988	$1.076.10^{-4}$	0.009	0.029	$6.403.10^{-4}$	0.009	0.029	$5.895.10^{-4}$
		3.020	0.971	0.965		0.067	0.076		0.066	0.074	
		3.444	0.929	0.931		0.149	0.137		0.138	0.127	
[1, 2000]	1040	2.963	0.997	0.998	$5.749.10^{-6}$	0.009	0.008	$1.055.10^{-4}$	0.009	0.013	$9.470.10^{-5}$
		3.020	0.971	0.971		0.067	0.076		0.066	0.054	
		3.444	0.929	0.926		0.149	0.154		0.138	0.125	

TAB. 3.10 – Résultats de simulation de la loi Weibull de paramètre  $\beta = 7$ ,  $\alpha = 5$  (cas aléatoire).

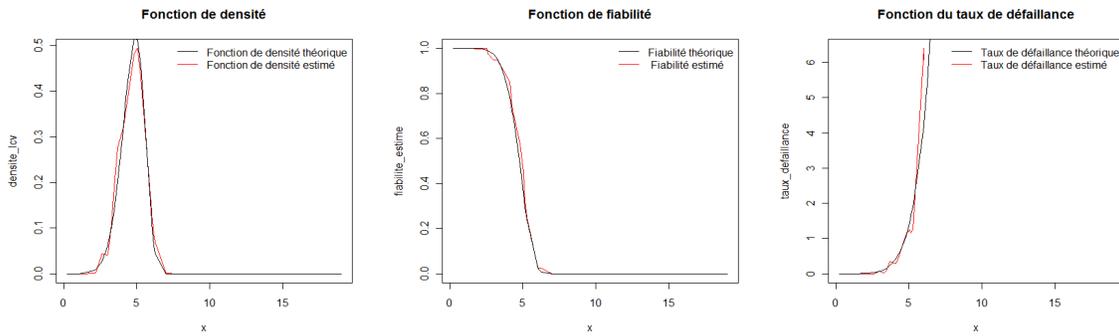


FIG. 3.28 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour  $n = 20$  sur l'intervalle  $[0, 50]$ .

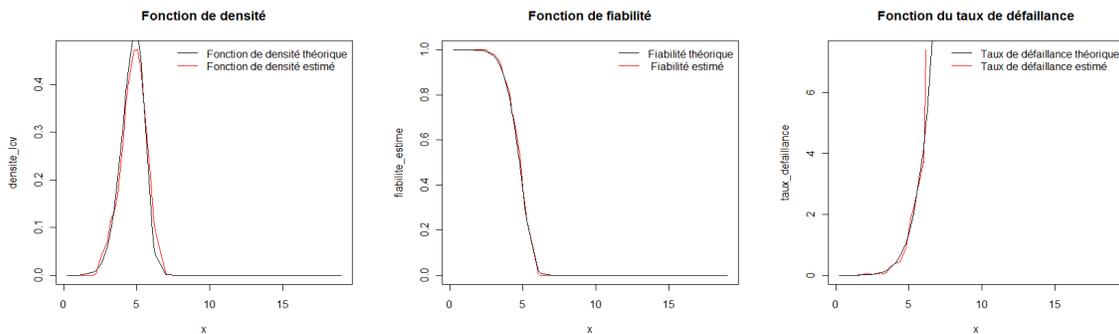


FIG. 3.29 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour  $n = 87$  sur l'intervalle  $[0, 200]$ .

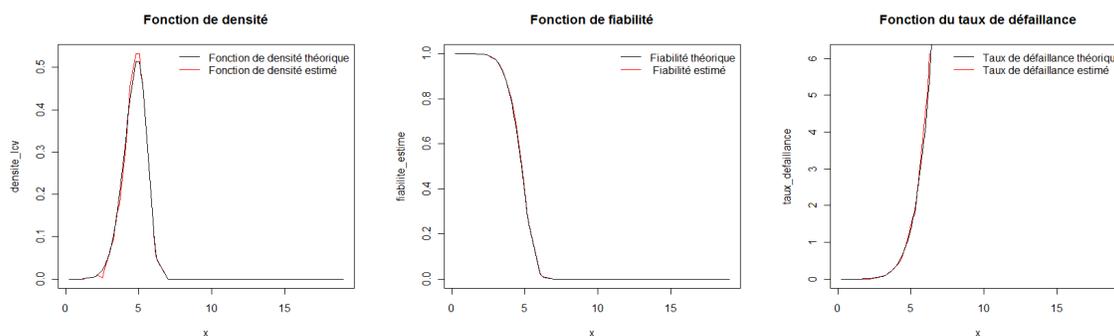


FIG. 3.30 – Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) pour  $n = 1040$  sur l'intervalle  $[0, 2000]$ .

### ✧ Interprétation des résultats

L'interprétation et l'analyse des différents résultats obtenus en simulant différentes lois sont résumées ci-dessous :

#### • Loi Gamma

Les résultats de simulation sont très appréciables, d'après les graphes (3.1), (3.2), (3.3), (3.16), (3.17) et (3.18), nous remarquons que lorsqu'on augmente la taille de l'échantillon dans le cas fixe ou l'horizon dans le cas aléatoire, nous obtenons une bonne estimation de la densité, de la fiabilité et du taux de hasard, ceci est confirmé par les résultats du ISE mentionnés dans les tableaux (3.1) et (3.6) qui diminuent à mesure que la taille de l'échantillon augmente.

#### • Loi Log-normale

Nous remarquons que l'allure des courbes estimées se rapproche de celles des courbes théoriques [voir (3.4), (3.5), (3.6), (3.19), (3.20) et (3.21)] à mesure que nous augmentons la taille de l'échantillon dans les deux cas d'étude.

Nous pouvons confirmer cela à partir des résultats obtenus du ISE dans les tableaux (3.2) et (3.7). Dans la plupart des résultats de simulations l'erreur (ou le ISE) diminue quand nous augmentons la taille de l'échantillon dans le cas fixe ou l'horizon dans le cas aléatoire.

• **Loi Weibull**

**Loi Weibull (0.5, 5)**

Nous constatons que l’allure des courbes (3.7), (3.8), (3.9), (3.22), (3.23) et (3.24) des fonctions théoriques se confondent avec celles estimées, ce qui est confirmé par les résultats obtenus dans les tableaux (3.3) et (3.8).

**Loi Weibull (1, 5)**

Pour un échantillon de petite taille dans le cas fixe ou aléatoire l’écart entre la fonction théorique et estimée est important. Pour la fonction du taux de défaillance, une légère amélioration est constatée avec l’augmentation de la taille de l’échantillon. Quant aux autres fonctions les résultats théoriques et estimées sont presque comparables à mesure que la taille de l’échantillon augmente dans les deux cas d’étude. Nous confirmons ceci avec les résultats du ISE qui diminue avec l’augmentation de la taille de l’échantillon dans la plupart des résultats obtenus [voir les tableaux (3.4) et (3.9) et les graphes (3.10), (3.11), (3.12), (3.25), (3.26) et (3.27)].

**Loi Weibull (7, 5)**

Les résultats des estimations et celles théoriques sont très proche [voir (3.13) et (3.14), (3.15), (3.28), (3.29), (3.30)], nous constatons ceci à partir des résultats des tableaux (3.5, 3.10). Nous remarquons que le ISE diminue quand nous augmentons la taille de l’échantillon dans les deux cas d’étude.

**3.6.5 Étude comparative**

• **La loi Weibull  $\mathcal{W}(0.5, 5)$**

Horizon	$N_t$	$n$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
				$R(x)$	$\tilde{R}_{N_t}^*(x)$	$\tilde{R}_n(x)$	$\lambda(x)$	$\tilde{\lambda}_{N_t}^*(x)$	$\tilde{\lambda}_n(x)$	$f(x)$	$\tilde{f}_{N_t}(x)$	$\tilde{f}_n(x)$
[0,50]	30	30	2.963	0.516	0.566	0.466	0.151	0.163	0.204	0.078	0.092	0.095
			3.020	0.459	0.500	0.433	0.128	0.118	0.167	0.059	0.059	0.072
			3.444	0.436	0.500	0.366	0.120	0.097	0.185	0.052	0.048	0.068
[0,200]	107	107	2.963	0.516	0.532	0.551	0.151	0.151	0.133	0.078	0.096	0.073
			3.020	0.459	0.495	0.504	0.128	0.125	0.126	0.059	0.059	0.064
			3.444	0.436	0.467	0.467	0.120	0.120	0.131	0.052	0.049	0.061
[0,2000]	1017	1017	2.963	0.516	0.521	0.521	0.151	0.157	0.135	0.078	0.081	0.070
			3.020	0.459	0.461	0.465	0.128	0.142	0.120	0.059	0.065	0.055
			3.444	0.436	0.432	0.441	0.120	0.128	0.121	0.052	0.055	0.053

TAB. 3.11 – Résultats de simulation et comparaison de la loi Weibull (0.5, 5)

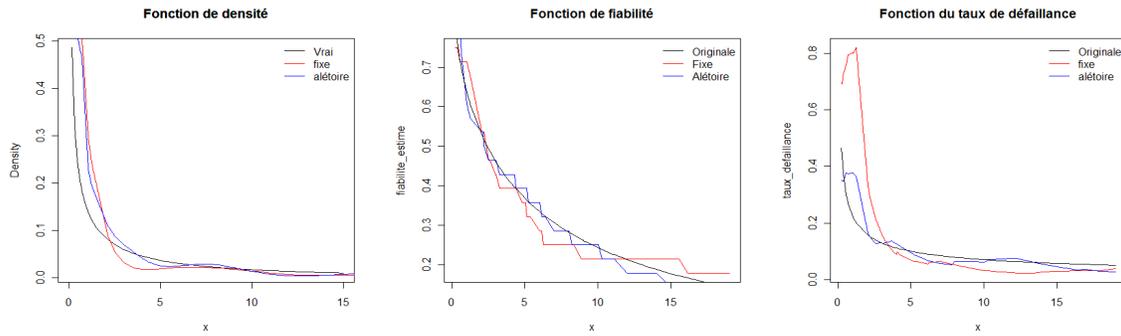


FIG. 3.31 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle [0, 50].

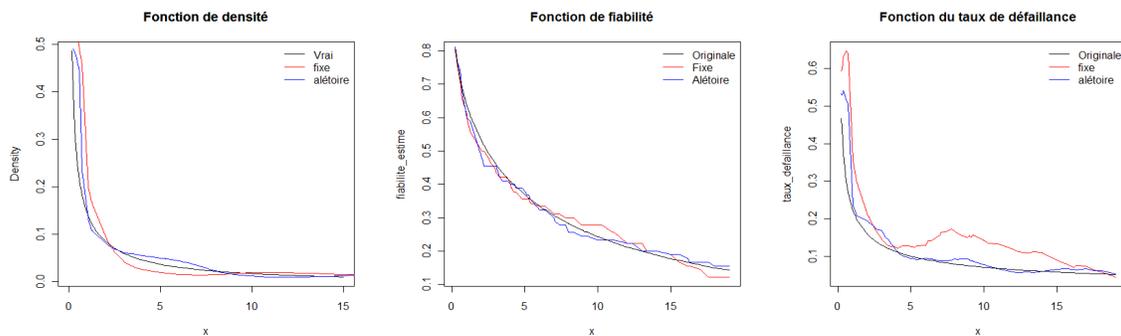


FIG. 3.32 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle [0, 200].

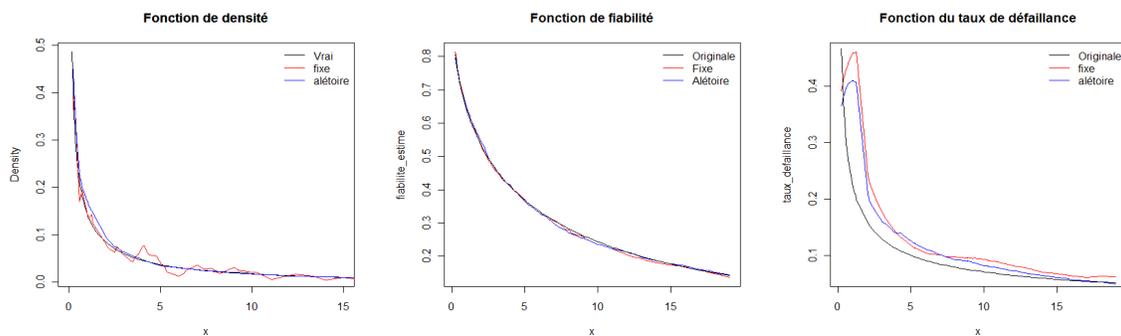


FIG. 3.33 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (0.5, 5) sur l'intervalle [0, 2000].

• La loi Weibull  $\mathcal{W}(1, 5)$

Horizon	$N_t$	$n$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
				$R(x)$	$\hat{R}_{N_t}^*(x)$	$\hat{R}_n(x)$	$\lambda(x)$	$\hat{\lambda}_{N_t}^*(x)$	$\hat{\lambda}_n(x)$	$f(x)$	$\hat{f}_{N_t}(x)$	$\hat{f}_n(x)$
[0,50]	24	24	2.963	0.646	0.667	0.667	0.2	0.176	0.232	0.129	0.117	0.154
			3.020	0.546	0.541	0.500	0.2	0.206	0.208	0.109	0.111	0.108
			3.444	0.502	0.541	0.500	0.2	0.199	0.166	0.100	0.107	0.166
[0,200]	103	103	2.963	0.646	0.660	0.640	0.2	0.198	0.190	0.129	0.131	0.121
			3.020	0.546	0.553	0.553	0.2	0.208	0.178	0.109	0.115	0.098
			3.444	0.502	0.505	0.524	0.2	0.212	0.187	0.100	0.107	0.098
[0,2000]	1029	1029	2.963	0.646	0.634	0.659	0.2	0.198	0.195	0.129	0.129	0.129
			3.020	0.546	0.541	0.551	0.2	0.196	0.204	0.109	0.109	0.113
			3.444	0.502	0.499	0.514	0.2	0.196	0.204	0.100	0.104	0.105

TAB. 3.12 – Résultats de simulation et comparaison de la loi Weibull (1, 5)

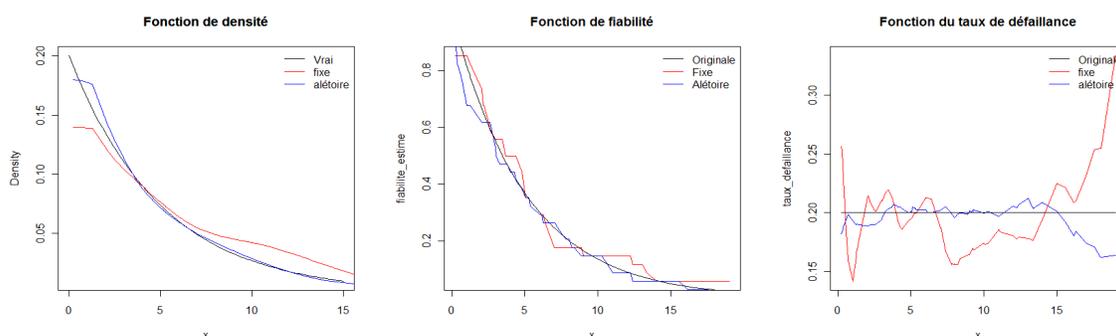


FIG. 3.34 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l'intervalle [0, 50].

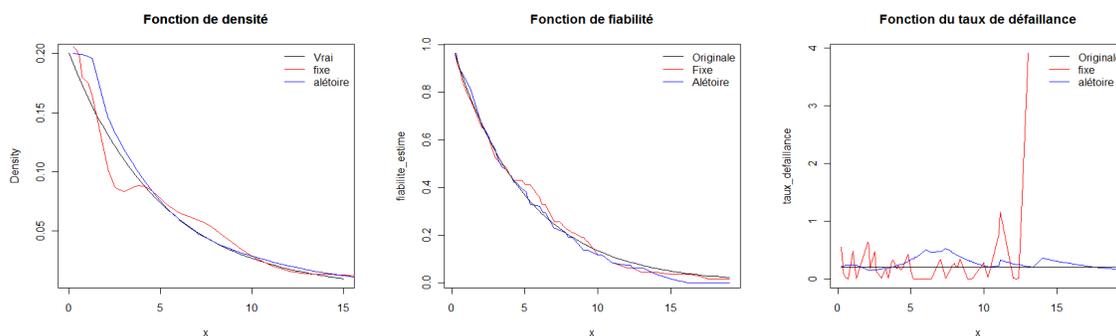


FIG. 3.35 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l'intervalle [0, 200].

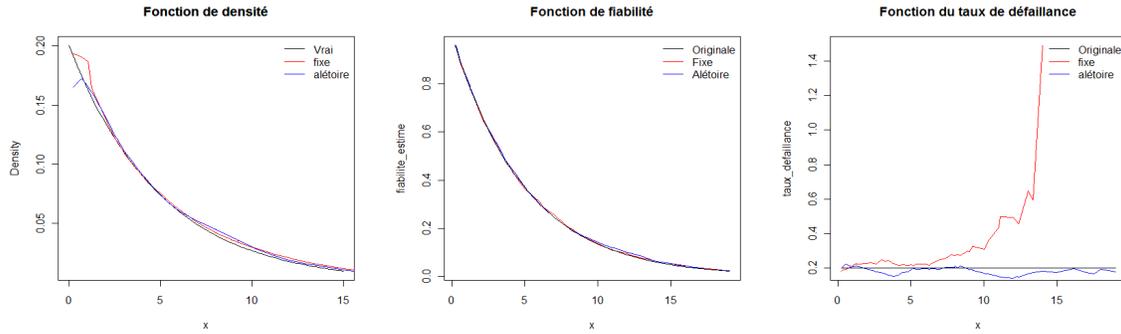


FIG. 3.36 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (1, 5) sur l’intervalle [0, 2000].

• La loi Weibull  $\mathcal{W}(7, 5)$

Horizon	$N_t$	$n$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
				$R(x)$	$\hat{R}_{N_t}^*(x)$	$\hat{R}_n(x)$	$\lambda(x)$	$\hat{\lambda}_{N_t}^*(x)$	$\hat{\lambda}_n(x)$	$f(x)$	$\hat{f}_{N_t}(x)$	$\hat{f}_n(x)$
[0,50]	20	20	2.963	0.997	1.000	1.00	0.009	0.013	0.008	0.009	0.013	0.008
			3.020	0.971	0.9	0.95	0.067	0.112	0.120	0.066	0.100	0.114
			3.444	0.929	0.9	0.85	0.149	0.123	0.177	0.138	0.110	0.150
[0,200]	87	87	2.963	0.997	0.988	1.000	0.009	0.029	0.005	0.009	0.029	0.005
			3.020	0.971	0.965	0.977	0.067	0.076	0.039	0.066	0.074	0.038
			3.444	0.929	0.931	0.954	0.149	0.137	0.158	0.138	0.127	0.151
[0,2000]	1040	1040	2.963	0.997	0.998	0.996	0.009	0.008	0.008	0.009	0.013	0.008
			3.020	0.971	0.971	0.971	0.067	0.076	0.068	0.066	0.054	0.066
			3.444	0.929	0.926	0.931	0.149	0.154	0.134	0.138	0.125	0.124

TAB. 3.13 – Résultats de simulation et comparaison de la loi Weibull (7, 5)

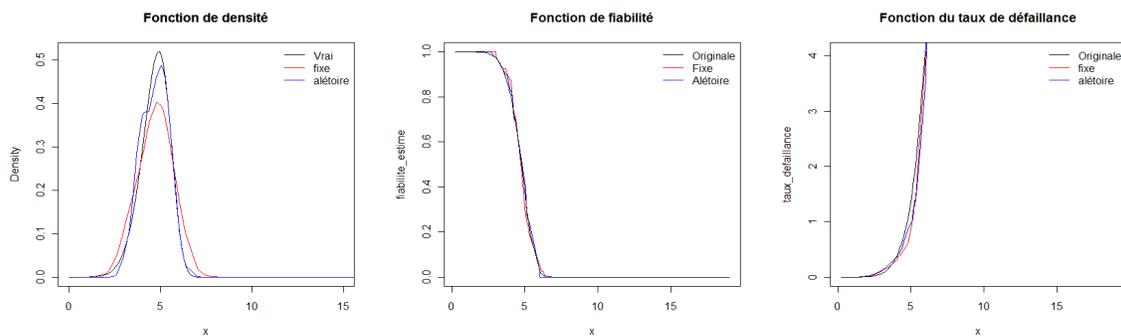


FIG. 3.37 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l’intervalle [0, 50].

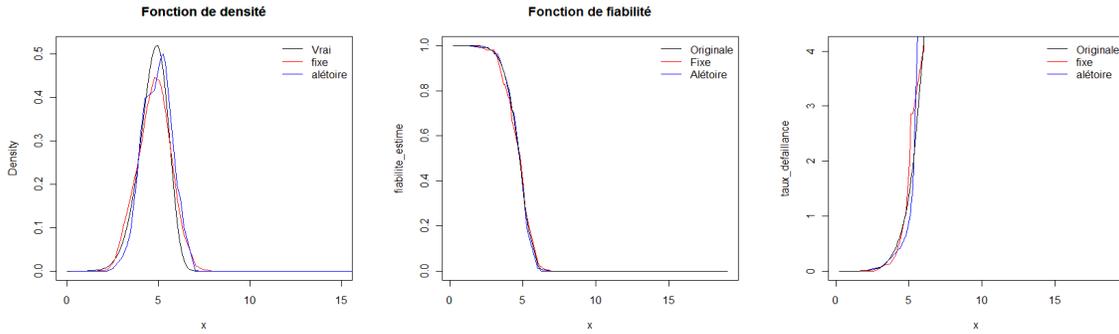


FIG. 3.38 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l’intervalle [0, 200].

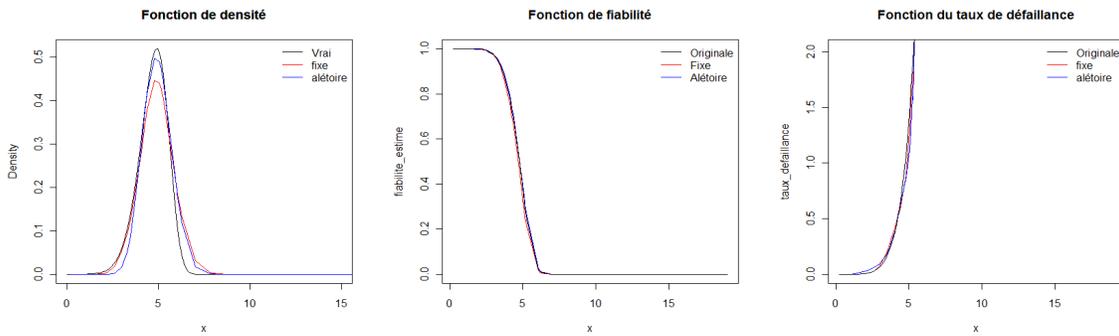


FIG. 3.39 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Weibull (7, 5) sur l’intervalle [0, 2000].

• La loi Log-normale  $Log\mathcal{N}(0, 1)$

Horizon	$N_t$	$n$	$x$	Fonction de fiabilité			Taux de hasard			Fonction de densité		
				$R(x)$	$\hat{R}_{N_t}^*(x)$	$\hat{R}_n(x)$	$\lambda(x)$	$\hat{\lambda}_{N_t}^*(x)$	$\hat{\lambda}_n(x)$	$f(x)$	$\hat{f}_{N_t}(x)$	$\hat{f}_n(x)$
[0,50]	23	23	2.963	0.218	0.173	0.260	0.620	0.703	0.557	0.135	0.122	0.145
			3.020	0.134	0.130	0.217	0.533	0.500	0.378	0.071	0.065	0.082
			3.444	0.108	0.130	0.173	0.498	0.411	0.738	0.053	0.053	0.128
[0,200]	100	100	2.963	0.218	0.22	0.20	0.620	0.691	0.555	0.135	0.152	0.111
			3.020	0.134	0.14	0.13	0.533	0.501	0.480	0.071	0.070	0.062
			3.444	0.108	0.11	0.12	0.498	0.492	0.468	0.053	0.054	0.056
[0,2000]	1018	1018	2.963	0.218	0.214	0.223	0.620	0.598	0.686	0.135	0.128	0.153
			3.020	0.134	0.132	0.139	0.533	0.497	0.504	0.071	0.065	0.070
			3.444	0.108	0.107	0.113	0.498	0.519	0.487	0.053	0.055	0.055

TAB. 3.14 – Résultats de simulation et comparaison de la loi Log-normale (0, 1)

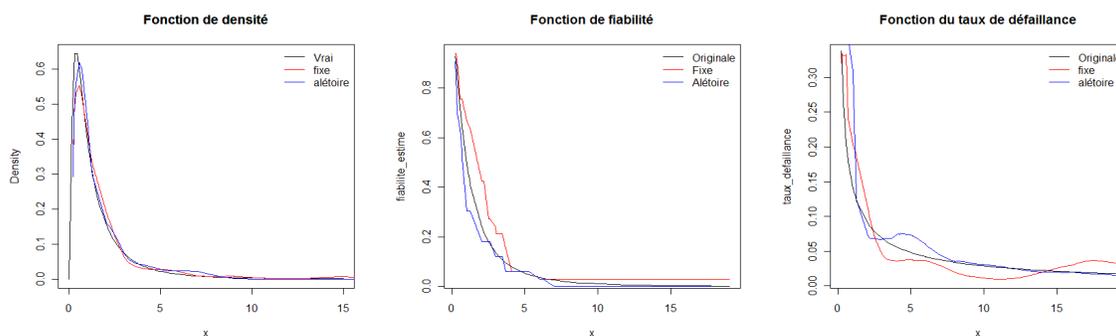


FIG. 3.40 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale  $(0, 1)$  sur l'intervalle  $[0, 50]$ .

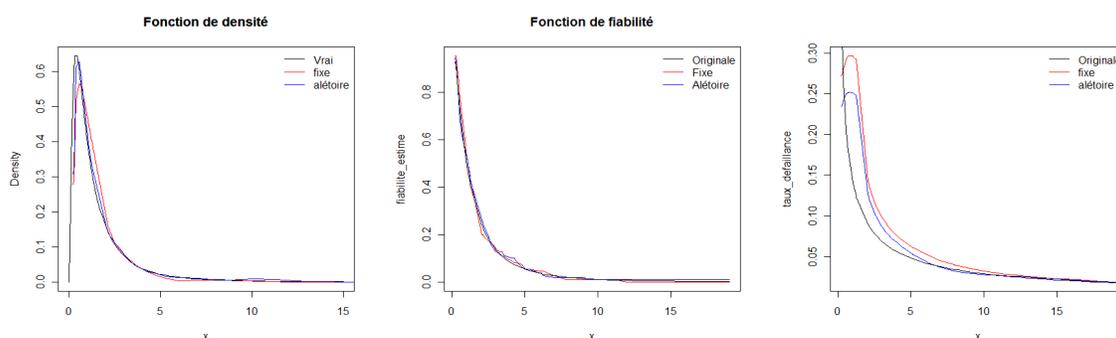


FIG. 3.41 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale  $(0, 1)$  sur l'intervalle  $[0, 200]$ .

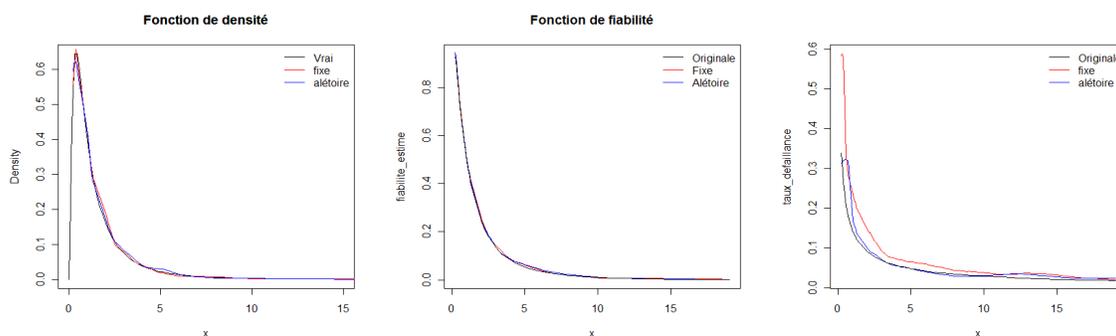


FIG. 3.42 – Étude comparative : Densité à noyau, Fiabilité et Taux de hasard associé à la loi Log-normale  $(0, 1)$  sur l'intervalle  $[0, 2000]$ .

### ✂ Interprétation des résultats

D'après les graphes (3.31), (3.32), (3.33), (3.34), (3.35), (3.36), (3.37), (3.38) et (3.39), (3.40), (3.41), (3.42) nous remarquons que les courbes correspondantes aux fonctions de densité, fiabilité et taux de hasard obtenues dans le cas d'une taille d'échantillon aléatoire

sont proches de celles obtenues dans le cas fixe, et ce pour un horizon grand. Cependant, les courbes correspondantes au cas aléatoire sont plus proches que celles du cas fixe des courbes correspondantes aux lois cibles, lorsque l'horizon est petit ou moyen.

Ceci est confirmé par les résultats numériques donnés dans les tableaux (3.11), (3.12), (3.13) et (3.14).

### 3.7 Conclusion

Dans ce chapitre, nous avons estimé la fonction de répartition en utilisant l'estimateur empirique et la fonction de densité estimée par la méthode du noyau de *Parzen – Rosenblatt* dans les cas où la taille de l'échantillon est fixe et aléatoire. Ces résultats estimés sont utilisés pour l'estimation de la fonction de fiabilité et le taux de hasard.

Les résultats de simulation montrent que la fiabilité et le taux de hasard estimés par la méthode du noyau, donnent des résultats appréciables pour les lois paramétriques utilisées, à savoir, Log-normale, Weibull et Gamma.

D'après les résultats de comparaison, les estimations obtenues dans le cas aléatoire sont meilleurs que celles obtenues dans le cas fixe pour un horizon petit ou moyen.

Les résultats sont meilleurs pour une taille d'échantillon grande (cas fixe) et pour un horizon grand (cas aléatoire).

# Conclusion générale

L'objectif principal de ce travail est d'évaluer par simulation la fiabilité et le taux de hasard dans le cas d'une taille d'échantillon aléatoire. Pour cela, nous avons utilisé l'estimateur de la densité de *Parzen–Rosenblatt* adapté au cas où la taille de l'échantillon est aléatoire et l'estimateur empirique pour la fonction de répartition.

Nous avons étudié les propriétés statistiques et asymptotiques de ces estimateurs que nous avons utilisés pour l'estimation des fonctions de fiabilité et de hasard.

Les nombreuses simulations que nous avons menés sur les différentes lois de fiabilité (Weibull, Gamma, Log-normale) nous ont montré que les valeurs de ces indices de fiabilité issues de ces simulations sont meilleures lorsqu'on augmente l'horizon (l'intervalle d'observation).

Dans le but de montrer que ces estimations sont bonnes, nous les avons comparé avec les estimations obtenues dans le cas d'une taille d'échantillon fixe.

Les résultats de comparaison se rapprochent lorsque l'horizon est grand, donc pour une taille d'échantillon grande. Cependant, les estimations obtenues dans le cas aléatoire sont meilleurs que celles obtenues dans le cas fixe pour un horizon petit ou moyen.

Pour l'obtention de ces résultats, nous avons utilisé le noyau asymétrique Gamma et la fenêtre de lissage UCV.

## Perspectives

Nous proposons d'étudier en perspectives :

- L'utilisation d'autres lois cibles, telles que : Normale, Birnbaum-Saunders et Pareto.
- Faire une application numérique sur des données réelles.
- Estimer la fonction de fiabilité à travers la fonction de répartition par la méthode du noyau.
- De faire une application numérique sur des données réelles que ce soit dans le cas où la taille de l'échantillon est fixe ou aléatoire.

- Utiliser d'autres méthodes pour la sélection de la fenêtre de lissage à savoir : plug-in moderne et l'approche bayésienne et une éventuelle étude comparative.

# Bibliographie

- [1] D. Aissani, *Modèles Stochastique de la théorie de fiabilité*. O.P.U,1992.
- [2] F.J. Anscome, *Large sample theory of sequentiel estimation*. Proc. Camb. Phlos. Soc, 48, pages 600-607,1952.
- [3] S. Abou-Jouadé, *Sur une condition nécessaire et suffisante de  $L^1$  convergence presque complète de l'estimateur de la partition fixe pour une densité*. C. R. Acad. Sci. Paris Sér A-B, 283(16), pages A1107–A1110, 1976.
- [4] A. Berlinet and L. Devroye, *A comparaison of Kernel density estimates*. Publication de l'Institut de Statistique de l'Université de Paris 38, pages 3-59,1994.
- [5] A.W. Bowman, *An alternative method of cross-validation for the smoothing density estimates*. Biometrika, 71, pags 353–360, 1984.
- [6] B.M. Brown, S. Chen, *Beta-bernstein smoothing for regression curves with compact supports*. Scandinavian Journal of Statistics, (26), pages 47-59, 1999.
- [7] D. Bosq, *Estimation suroptimale de la densité par projection*. Canadian Journal of Statistics, 33(1), pages 21–37, 2005.
- [8] T. Bouezmarni, O. Scaillet, *Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data*. Econometric Theory, (21), pages 390-412, 2003.
- [9] S X. Chen, *A Beta kernel estimation for density functions*. Computational Statistics and Data Analysis, 31, pages 131-145, 1999.
- [10] S X. Chen, *Gamma kernel estimators for density functions*. Annals of the Institute of Statistical Mathematics, 52, pages 471–480, 2000.
- [11] S X. Chen, *Beta kernel smoothers for regression curves*. Statistica Sinica, 10, pages 73-91, 2000.
- [12] M. Cherfaoui, *Bootstrap dans l'estimation non paramétrique de la densité de probabilité et la courbe de regression de la moyenne*. Mémoire de magister, Université de Béjaia, 2009.

- [13] N. Cencov, *Evaluation of an unknown distribution density from observations*. Sovet, Math, 3, pages 1559–1562, 1962.
- [14] A. Charpentier, J.D. Fermanian and O. Scaillet, *The estimation of copulas : Theory and practice*. Ensaë-Crest and Katholieke Universiteit Leuven, BNP-Paribas and Crest, HEC Genève and Swiss Finance Institute, 2006.
- [15] R. J. Carroll, “ *On Sequential Density Estimation*”, Z. Wahrscheinlichkeits theorie verw. Gebiete, 36, pages 137–151, 1976.
- [16] L. Devroye, *The equivalence of weak, strong and complete convergence in  $L^1$  for kernel density estimates*. The Annals of Statistics, 11, pages 896–904, 1983.
- [17] T. DeMarco, *Controlling Software Projects : Management, Measurement and Estimation*. Prentice Hall PTR. 1982.
- [18] P. Deheuvels, *Estimation non paramétrique de la densité par histogrammes généralisés*. Revue de Statistique Appliqué, 25, pages 5-42, 1977.
- [19] V.A. Epanechnikov, *Nonparametric estimation of a multidimensional probability density*. Theory Probab. Appl, 14, pages 153-158, 1969.
- [20] J.J. Faraway, M. Jhung. *Bootstrap choice of bandwidth for density estimation*. Journal of the American Statistical Association, (85), pages 1119-1122, 1990.
- [21] A. Gabriel Mihalache, *Modélisation et évaluation de la fiabilité des systèmes mécatroniques : application sur système embarqué*. Thèse de Doctorat Université d’Angers, À l’Institut des Sciences et Techniques de l’Ingénieur d’Angers, Décembre 2007.
- [22] Geffroy, *Sur l’estimateur de la densité dans un espace métrique*. C.R. Acad. Sci. Paris Sér A-B. 278, pages 1449-1452. 1973.
- [23] F.E. Harrell, C.E. Davis, *A new distribution-free quantile estimator*. Biometrika, 69, pages 635-640, 1982.
- [24] G. Kerkyacharian and D. Picard. *Density estimation by kernel and wavelets methods : optimality of besov spaces*. Statistics and Probability Letters, 18(4) :327–336, 1993.
- [25] R A. Kronmal and M E. Tarter, *The estimation of probability densities and cumulatives by fourie series method*. J. Amerc. Stat. Assoc, 63, pages 925–952, 1968.
- [26] J.C. Ligeron, P. Lionnet, *La fiabilité en exploitation, organisation et traitement des données*. Ed. Techniques et documentation (Lavoisier), 1993.
- [27] P. Lionnet, *La maintenance, mathématique et méthodes*. Ed. Techniques et documentation(Lavoisier), 1988.

- [28] K. Lagha, *Cours de fiabilité et analyse des données de survie*. Département de mathématique. Université de Béjaia, 2015.
- [29] K. Lagha, S.Adjabi, *Laws of the iterated logarithm for non parametric sequential density estimators* Journal of statistical theory and application. Vol-12, No.2, pages 145-151, July 2013.
- [30] M. Lejeune, *Statistique : la théorie et ses applications*. Springer, France, deuxième édition, 2004
- [31] C . Matias, *Introduction à la statistique non paramétrique*, Laboratoire Statistique et Génome, Évry, 2013-2014.
- [32] J. S. Marron, P. Hall, and B.U. Park, *Smoothed cross validation. Probability Theory and Related Fields*, 92, pages 1-20, 1992.
- [33] C. Marcovici, et Ligeron, J.C, *Utilisation des techniques de fiabilité en mécanique*. Technique et documentation, 1974.
- [34] V. K. Murthy, *Estimation of jumps, reliability and hazard rate*. Ann. Stat. 36, pages 1034-1040, 1965.
- [35] E.A. Nadaraya, *On nonparametric estimation of density function and regression. Theory. Probab. Appl*, 10, pages 186–190, 1965.
- [36] H. Procaccia, L. Piepszownik, *Fiabilité des équipements et théorie de la décision statistique fréquentielle et bayésienne*. Ed. Ayrolle, 1992.
- [37] T. Pyzdek, *The Six Sigma handbook. A complete guide for greenbelts blackbelts, managers at all levels*. McGraw-Hill Compagnies the United States of America, 2001.
- [38] E. Parzen, *On estimation of a probability density function and mode. Annals of Mathematical Statistics*. 33, pages 1065–1076, 1962.
- [39] F. Rieuneau, *Sûreté de fonctionnement en phase de developpement des systèmes embarqués automobiles. In Integrated Logistics Concurrent Engineering*, Montpellier, 1993.
- [40] M. Rudemo, *Empirical choice of histograms and kernel density estimators*. Scandinavian Journal of Statistics, 9, pages 65–78, 1982.
- [41] M. Rosenblatt, *Remarks in some nonparametric estimates of a density function. Annals of Mathematical Statistics*, 27, pages 832–837, 1956.
- [42] A. Renyi, *On the asymptotic distribution of the sum of a random number of independent random variables*. Acta. Math. Acad. Sci. Hung 8, pages 193-199, mars 1957.

- [43] E.F. Schuster, *Incorporating support constraints into nonparametric estimators of densities*. Communications in statistics-Theory and Methodes, 14, pages 1123-1136, 1985.
- [44] R.C. Srivastava, *Estimation of probability density function based on random number observation with applications*. Int.stat.Revue, Vol. 41, No. 1 pages 77-86, 1973.
- [45] B.W. Silvermann, *Density Estimation for Statistics and Data Analysis*. Published in Monographs on Statistics and Applied Probability, Chapman and Hall, New York, 1986.
- [46] B.W. Silvermann, *Weak and strong uniform consistency of the kernel estimate of density function and its derivatives*. Ann. Statist, 6, pages 177–184, 1978.
- [47] D.W. Scott and G.R. Terrel, *Biased and unbiased cross-validation in density estimation*. Journal of the American Statistical Association. 82, pages 1131–1146, 1987.
- [48] D.W. Scott, R.A. Tapia and J.R. Thomson, *Kernel density estimation revisited*. Non-linear Analysis, Theory, Methods and Applications, 1, pages 339-372, 1977.
- [49] N. Saadi and S. Adjabi, *On the estimation of the probability density by trigonometric series*. Communications in Statistics Theory and Methods, 38(19), pages 3583–3595, 2009.
- [50] J. Tiago de Oliveira. *Estatística de densidades. Resultados Assintoticos*. Rev. Fac. Ci. Univ. Lisboa, Ser A, 9, pages 111–206,1963.
- [51] M. Tahan clement, P. Thomas, *Fiabilité et maintenance industrielle*. Université de Québec, école de technologie supérieur , Departement de génie mécanique, 2003.
- [52] C. Taylor. *Bootstrap choice of smoothing parameter in kernel density estimation*. Biometrika,(76), pages 705-712, 1989.
- [53] O. Tebbi, *Estimation des lois de fiabilité en mécanique par les essais accélérés*. PHD thesis, Institut des sciences et techniques de l'ingénieur d'Angers, 2005.
- [54] A. Villemeur, *Sûreté de fonctionnement des systèmes industriels : fiabilité, facteurs humains, informatisation*. Eyrolles, Paris, 1988.
- [55] G. G. Walter. *Approximation of the delta function by wavelets*. Journal of Approximation Theory. 71(3) :329–343, 1992.
- [56] G. S. Watson and M. R. Leadbetter, *Hasard analysis I*. Biometrika, 51, pages 174-184. 1964a.
- [57] G. S. Watson and M. R. Leadbetter, *Hasard analysis II*. Sankhya, 26A, pages 110-116. 1964b.

- 
- [58] G. Wahba, *Data based optimal smoothing of orthogonal serie density estimates*. Ann. Statist, 9(1), pages 146–156, 1981.
- [59] G. Wahba, *Optimal convergence properties of variable knot, kernel, and orthogonal series methods for density estimation*. The Annals of Statistics, pages 15–29, 1975.
- [60] N. Zougab, *Etude comparative des méthodes de sélection du paramètre de lissage dans l'estimation de la densité de probabilité par la méthode du noyau*. Mémoire de de Magister, Université de Béjaia, Algérie, 2007.
- [61] N. Zougab, *Approche Bayésienne dans l'estimation non paramétrique de la densité de probabilité et la courbe de régression de la moyenne*. Thèse de Doctorat en Mathématiques Appliquées, Université de Béjaia, Algérie, 2013.

## Résumé

L'étude de la fiabilité des systèmes est un axe de recherche qui a toujours intéressé les chercheurs, pour prendre des mesures nécessaires afin d'éviter les pannes.

Notre travail est d'étudier la fiabilité et le taux de défaillance d'un système, lorsque celui-ci est observé sur une durée fixée  $(0, t)$ ,  $t > 0$ . Le nombre d'observations est alors aléatoire.

Les estimateurs que nous avons proposés sont basés sur l'estimateur à noyau de *Parzen–Rosenblatt* et l'estimateur empirique pour estimer la densité et la fonction de répartition, respectivement. Les simulations effectuées permettent de conclure sur la bonne qualité de ces estimateurs, comparés aux estimateurs classiques obtenus par les mêmes méthodes dans le cas fixe.

**Mots clés :** Densité, Estimateur à noyau de *Parzen – Rosenblatt*, Estimateur empirique, Fiabilité, Fonction de répartition, Taux de défaillance.

## Abstract

The study of system reliability is a line of research that has always interested researchers, to take necessary measures to prevent damage.

Our work is to study the reliability and failure rate of a system, when it is observed over a fixed period  $(0, t)$ ,  $t > 0$ . The number of observation is so random.

The estimators we have proposed are based on the kernel estimator of *Parzen – Rosenblatt* and empirical estimator to estimate the density and distribution function, respectively. Simulations can be concluded on the good quality of these estimators, compared to conventional estimators obtained by the same methods in the fixed case.

**Keywords :** Density, kernel estimator of *Parzen – Rosenblatt*, Empirical estimator, Reliability, Distribution function, Failure rate .