

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA  
FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la  
VIE  
DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

**MASTER en Mathématiques**

Option : **Statistique**

Par

**HAIF KHAIF Fatna**

Titre :

**Estimation de la fonction de distribution**

Membres du Comité d'Examen :

Dr. <b>CHERFAOUI Mouloud</b>	UMKB	Président
Dr. <b>SAYAH Abdallah</b>	UMKB	Encadreur
Dr. <b>OUANOUGHY Yasmina</b>	UMKB	Examinatrice

Juin 2021

## Dédicace

Je dédie ce travail à :

A mon père **Abdelkader** et ma mère **Hakima**, je vous remercie pour tout ce que vous avez fait pour moi. Que dieu vous préserve une longue vie heureuse.

Avec toute ma tendresse.

A mon frère **Mahdi** et ma sœur **Féryal**.

A ma famille, grands et petits.

A mes amis.

A tous les enseignants qui ont contribué à mes études.

## REMERCIEMENTS

Tout d'abord je tiens à remercier **Dieu** le tout puissant de m'avoir donné le courage,  
la morale et la santé  
pour mener à bien ce travail.

Je tiens à remercier et mon plus grand respect et appréciation mon encadreur : **Dr. Sayah Abdellah.**

Mes remerciements vont aussi aux membres du jury **Dr. Ouanoughi Yasmina** et **Dr. Cherfaoui Mouloud** qui nous font le grand honneur d'accepté d'examiner et d'évaluer notre travail et de votre grande patience et de vos conseils forts judicieux.

Naturellement, je ne peux oublier **Almi Nassima** qui m'ont apporté leur aide.

Une spéciale remerciement pour **Pr. NECIR Abdelhakim.**

Je tiens à remercier mon ami **Boubeche Noureddine.**

Un grand merci pour **Azzeddine Haif.**

Finalement un grand merci à mes familles, mes chères amie et à toutes les personnes qui nous ont encouragées pendant la réalisation de mon mémoire.

# Table des matières

<b>Remerciements</b>	<b>ii</b>
<b>Table des matières</b>	<b>ii</b>
<b>Table des figures</b>	<b>v</b>
<b>Liste des tables</b>	<b>vi</b>
<b>Introduction</b>	<b>1</b>
<b>1 Variables aléatoires et estimations paramétriques</b>	<b>3</b>
<b>1.1 Variables aléatoires</b> . . . . .	3
<b>1.1.1 Variables aléatoires discrètes</b> . . . . .	4
<b>1.1.2 Variables aléatoires continues</b> . . . . .	5
<b>1.1.3 Fonction de distribution</b> . . . . .	6
<b>1.1.4 Caractéristiques d'une variable aléatoire :</b> . . . . .	7
<b>1.1.5 Les lois de probabilités usuelles</b> . . . . .	9
<b>1.2 Convergence d'une suite de variables aléatoires</b> . . . . .	15
<b>1.2.1 Convergence en probabilité</b> . . . . .	16
<b>1.2.2 Convergence en loi</b> . . . . .	17

1.2.3	Convergence presque-sûre	17
1.2.4	Convergence en moyenne d'ordre $p$	18
1.2.5	Théorème central limite	19
1.3	<b>Estimation de la fonction de distribution</b>	19
1.3.1	Définitions de base	19
1.3.2	Estimation paramétrique	23
<b>2</b>	<b>Estimation non paramétrique de la fonction de distribution</b>	<b>29</b>
2.1	<b>Fonction de distribution empirique</b>	29
2.1.1	Propriétés statistiques	30
2.2	<b>Estimation de la fonction de distribution par la méthode du</b>	
	<b>noyau</b>	32
2.2.1	Erreur quadratique moyenne	37
2.2.2	Erreur quadratique moyenne intégrée asymptotique	38
2.3	<b>Simulation</b>	40
2.3.1	Exemples sur le noyau Quartique	41
2.3.2	Exemples sur le noyau d'Epanechnikov	42
	<b>Conclusion</b>	<b>43</b>
	<b>Bibliographie</b>	<b>44</b>
	<b>Annexe A : Logiciel R</b>	<b>46</b>
2.4	Qu'est-ce-que le langage R ?	46
	<b>Annexe B : Abréviations et Notations</b>	<b>47</b>

# Table des figures

1.1 fonction de distribution de loi uniforme et Binomiale. . . . .	7
2.1 Fonction de répartition empirique. . . . .	31
2.2 Les courbes des noyaux. . . . .	34
2.3 la fonction de distribution, noyau Quartique. . . . .	41
2.4 la fonction de distribution, noyau d'Epanechnikov. . . . .	42

# Liste des tableaux

1.1	Tableau représente les valeurs et les probabilités . . . . .	8
2.1	Tableau représente quelques exemples de noyaux . . . . .	34

# Introduction

La théorie de l'estimation est l'une des branches les plus basiques de la statistique mathématique. Cette théorie est habituellement divisée en deux composantes principales à savoir, l'estimation paramétrique et l'estimation non paramétrique. L'estimation paramétrique a comme inconvénient principal de nécessiter une connaissance préalable du phénomène aléatoire considéré. L'estimation non paramétrique estime la distribution directement à partir de l'information disponible sur l'ensemble d'observations. Il existe plusieurs méthodes non paramétriques pour l'estimation, on peut citer la méthode de l'histogramme, la méthode de séries orthogonales, la méthode splines et la méthode du noyau. Cette dernière est la plus utilisée et la plus populaire parmi les autres. Cette popularité peut s'expliquer au moins trois raisons : la simplicité de sa forme, ses modes de convergence multiples et sa flexibilité. Dans ce mémoire nous étudions l'estimation de la distribution à partir d'un échantillon indépendante et identiquement distribué (i.i.d).

Ce mémoire est partagé en deux chapitres :

**Chapitre 1 : "Variables aléatoires et estimations paramétriques"** : Dans le premier chapitre on donne quelques définitions et notations sur les variables aléatoires et les types de convergence de variables aléatoires et on étudie la méthode d'estimation paramétrique.

**Chapitre 2 : "Estimation non paramétrique de la fonction de distribution"** : Le deuxième chapitre est consacré à l'établissement de certaines propriétés

statistiques et les propriétés asymptotiques de l'estimateur de la fonction de distribution et on donne une étude générale par la méthode de noyau. Nous terminons ce chapitre par une simulation où nous donnons des exemples de simulation par le logiciel R qui expriment l'importance du choix de paramètre de lissage et du noyau.

# Chapitre 1

## Variables aléatoires et estimations paramétriques

Dans ce chapitre, nous allons présenter la variable aléatoire et ses types et les Caractéristiques et on donne quelques exemples sur les lois de probabilités usuelles et les types de convergence de suites de variable aléatoire. Ensuite, nous étudions la méthode d'estimation ( estimation paramétrique ).

### 1.1 Variables aléatoires

Soit  $(\Omega, \mathcal{F}, \mathcal{P})$  un espace probabilisé, une variable aléatoire (v. a.)  $X$  est une application définie sur l'ensemble fondamental  $\Omega$  et à valeurs réelles par :

$$\begin{aligned} X : \Omega &\rightarrow \mathbb{R} \\ \omega &\rightarrow X(\omega) \end{aligned}$$

Il y a deux types de variables aléatoires : discrètes et continues.

### 1.1.1 Variables aléatoires discrètes

**Définition 1.1.1** On appelle v. a. discrète définie sur  $(\Omega, \mathcal{F})$  une application  $X : \Omega \rightarrow \mathbb{R}$  telle que  $X(\Omega)$  est finie ou infinie dénombrable, et pour tout  $x \in \mathbb{N}$  :

$$X^{-1}(x) = \{\omega \in \Omega / X(\omega) = x\} \in \mathcal{F}.$$

telle que :

1)  $X(\Omega) = \{x_i, i \in I\}$  avec  $I \subseteq \mathbb{N}$ .

2) pour tout  $x \in X(\Omega)$ ,  $X^{-1}(x) \in \mathcal{F}$ .

-  $X(\omega)$  est l'ensemble des valeurs prises par  $X$  :

- Si  $I$  est finie  $X(\omega)$  est un ensemble finie et on dit que  $X$  est une variable aléatoire réelle discrète finie.

- Sinon on dit que  $X$  est une variable aléatoire réelle discrète infinie.

#### Exemple 1.1.1

Lors du lancer d'un dé :  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .

#### Loi de probabilité :

La loi de probabilité d'une v.a. réelle discrète est définie par les probabilités individuelles :

$$p_i = P(X = x_i) = P(X^{-1}(x_i)), \quad i \in \mathbb{N}.$$

telle que  $P$  est la probabilité définie sur  $X(\Omega)$ .

On appelle alors distribution ou loi de probabilité de  $X$  l'ensemble des couples  $(x_i, p_i)_{i \in I}$ .

Les probabilités  $p_i$  doivent satisfaire les deux conditions suivantes :

1. Toutes les  $p_i$  sont positives :  $p_i \geq 0$ .

2. La somme de toutes les  $p_i$  est égale à 1 : 
$$\begin{cases} \sum_{i=1}^n p_i = 1 & \text{si l'ensemble fini} \\ \sum_{i=1}^{\infty} p_i = 1 & \text{si l'ensemble infini} \end{cases}$$

### 1.1.2 Variables aléatoires continues

**Définition 1.1.2** On appelle v. a. réelle continue définie sur  $(\Omega, \mathcal{F})$  une application

$X : \Omega \rightarrow \mathbb{R}$  telle que pour tout intervalle  $I \subseteq \mathbb{R}$  on a :

$$X^{-1}(I) = \{\omega \in \Omega / X(\omega) \in I\} \in \mathcal{F}.$$

Cette condition est vérifiée si pour tout  $x \in \mathbb{R}$  on a  $X^{-1}(-\infty, x] \in \mathcal{F}$ .

#### Loi de probabilité :

La loi de probabilité est déterminée par la fonction de distribution  $F$ , qui est définie pour tout  $x$  réel par :

$$F(x) = P(X \leq x) = P\{X^{-1}(-\infty, x]\} = P\{\omega \in \Omega / X(\omega) \leq x\}.$$

#### Densité d'une variable aléatoire continue :

La densité de probabilité d'une (v. a.) continue est la dérivée première par rapport à  $x$  de la fonction de distribution. Cette dérivée prend le nom de fonction de densité, notée  $f(x) = \frac{dF(x)}{dx}$ .

La fonction de densité  $f(x)$  doit satisfaire les deux conditions suivantes :

1. La fonction de densité  $f(x)$  est positive :  $f(x) \geq 0$ .

2. L'intégrale de la fonction de densité  $f(x)$  est égale à 1 :  $\int_{\mathbb{R}} f(x) dx = 1$ .

La probabilité que  $X$  appartienne à l'intervalle  $[a, b]$  est donnée par l'intégrale :

$$P(a \leq X \leq b) = \int_a^b f(x) dx.$$

### 1.1.3 Fonction de distribution

La fonction de distribution est l'instrument de référence pour définir de façon unifiée la loi de probabilité d'une variable aléatoire qu'elle soit discrète ou continue. Si cette fonction est connue, il est possible de calculer la probabilité de tout intervalle et donc, en pratique, de tout événement, c'est pourquoi c'est elle qui est donnée dans les tables des lois de probabilité.

**Définition 1.1.3** Soit  $X$  une variable aléatoire, on appelle **fonction de distribution** de  $X$ , que l'on note  $F$  la fonction définie sur  $\mathbb{R}$  par :

$$F(x) = P(X \leq x).$$

$$F(x) = \begin{cases} \sum_{x_i \leq x} P(X = x_i) & \text{si } X \text{ est discrète} \\ \int_{-\infty}^x f(t) dt & \text{si } X \text{ est continue} \end{cases}$$

La fonction de distribution  $F$  d'une v. a r  $X$  satisfait les propriétés suivantes :

1.  $F$  est définie sur  $\mathbb{R}$  et à valeurs dans  $[0, 1]$ .
2. La fonction  $F$  est croissante.
3. Une fonction de distribution est continue à gauche, c'est à dire  $F(x) = F(x^-)$ .
4. Si la v. a. est continue, la fonction  $F$  est continue à droite et dérivable.
5.  $\lim_{x \rightarrow -\infty} F(x) = 0$  et  $\lim_{x \rightarrow +\infty} F(x) = 1$ .

Le graphe suivant représente la fonction de distribution dans les deux cas (discrét et continu) :

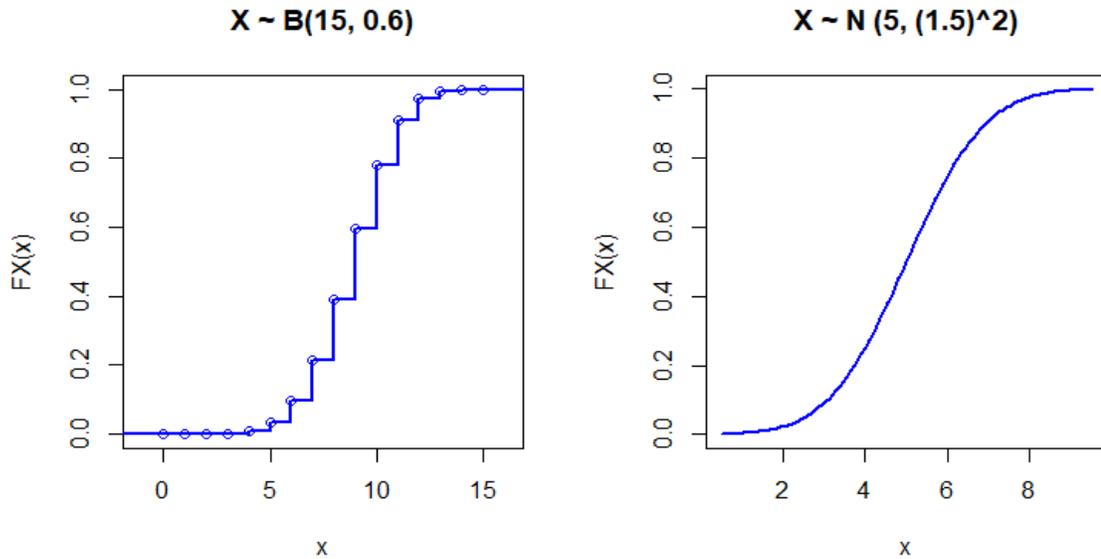


FIG. 1.1 – fonction de distribution de loi uniforme et Binomiale.

### 1.1.4 Caractéristiques d'une variable aléatoire :

Nous en définissons deux qui sont les plus utiles, mais il y en a évidemment beaucoup d'autres.

#### Espérance :

**Définition 1.1.4** On appelle *espérance mathématique* de  $X$ , si elle existe, la valeur notée  $E(X)$  telle que :

$$E(X) = \begin{cases} \sum_{i=1}^n x_i P(X = x_i) & \text{dans le cas discrèt,} \\ \int_{-\infty}^{+\infty} x f(x) dx & \text{dans le cas continu.} \end{cases}$$

Aussi  $E(X)$  sera appelée la moyenne de  $X$ .

**Remarque 1.1.1** Si l'intégrale  $\int_{-\infty}^{+\infty} xf(x) dx$  n'est pas convergente, alors l'espérance de  $X$  n'est pas définie.

**Variance :**

**Définition 1.1.5** On appelle **variance** de  $X$ , la valeur (si elle existe) notée  $Var(X)$  définie par :

$$\begin{aligned} Var(X) &= E[X - E(X)]^2 \\ &= E(X^2) - [E(X)]^2. \end{aligned}$$

La racine carrée de  $Var(X)$ , notée  $\sigma(X)$  est appelée écart-type de  $X$ .

$$\sigma(X) = \sqrt{Var(X)}.$$

**Exemple 1.1.2** Cas discret :

Soit  $\Omega = \{0, 1, 2\}$  nous rappelons que  $X$  représente le nombre de piles obtenus en lançant 2 pièces de monnaie. Il s'agit d'une variable aléatoire discrète. Ses valeurs et les probabilités associées dans le tableau suivant :

$x_i$	0	1	2
$p_i$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$

TAB. 1.1 – Tableau représente les valeurs et les probabilités .

$$\sum_{i=1}^n p_i = \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1.$$

$$E(X) = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1$$

$$Var(X) = \left( (0)^2 \times \frac{1}{4} + (1)^2 \times \frac{1}{2} + (2)^2 \times \frac{1}{4} \right) - [E(X)]^2.$$

$$Var(X) = \frac{3}{2} - 1 = \frac{1}{2}$$

$\sigma(X) \simeq 0.707$ .

**Exemple 1.1.3** *Cas continue :*

Soit  $X$  une variable aléatoire de densité de probabilité définie par :

$$f(x) = \begin{cases} \frac{1}{4} & \text{si } x \in [0, 4] \\ 0 & \text{sinon} \end{cases}$$

La fonction de distribution est définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x}{4} & \text{si } 0 \leq x < 4 \\ 1 & \text{si } x \geq 4 \end{cases}$$

$X$  une variable aléatoire et l'on vérifie que :

1.  $f(x) \geq 0$ .
2.  $\int_{-\infty}^{+\infty} f(x) dx = \int_0^4 \frac{1}{4} dx = 1$ .

Alors :  $E(X) = \int_0^4 x \frac{1}{4} dx = 2$ ,  $Var(X) = \frac{4}{3}$ ,  $\sigma(X) = \frac{2}{\sqrt{3}}$ .

### 1.1.5 Les lois de probabilités usuelles

Nous abordons ici une sorte de catalogue des lois les plus utilisées dans la modélisation statistique. Nous nous efforcerons de justifier l'utilité de ces lois en précisant le type de situations où elles sont appropriées. De façon générique et sauf mention expresse on notera  $X$  une v. a. qui suit la loi décrite. Chaque loi fera l'objet d'un symbole spécifique. Les tables de lois de probabilité fournies donnent notamment, pour les lois les plus usuelles, les probabilités élémentaires ou la densité, l'espérance, la variance. On présente dans cette section quelques propriétés supplémentaires de quelques-unes de ces lois.

## Les lois discrètes

### Loi de Bernoulli $\mathcal{B}(p)$

**Définition 1.1.6** *La loi de Bernoulli est utilisée lorsqu'une expérience aléatoire n'a que deux résultats possibles : le succès avec une probabilité de  $p$ , et l'échec avec une probabilité  $1 - p$ . Les applications de cette loi sont nombreuses. On dit que  $X$  suit la loi de Bernoulli de paramètre  $p$  et on note  $X \sim \mathcal{B}(p)$ , si sa fonction de masse est définie par :*

$$P(X = 1) = p \text{ et } P(X = 0) = 1 - p.$$

On a alors :

$$\forall x \in \{0, 1\}, \quad P(X = x) = p^x (1 - p)^{1-x}.$$

### Caractéristiques de la loi de Bernoulli :

Espérance, variance et l'écart-type :

$$E(X) = p, \quad \text{Var}(X) = p(1 - p) \quad \text{et} \quad \sigma(X) = \sqrt{p(1 - p)}.$$

### Loi de Binomiale $\mathcal{B}(n, p)$

**Définition 1.1.7** *On réalise  $n$  fois successivement et d'une manière indépendante une expérience aléatoire qui a deux résultats possibles, le succès ( associé au résultat pour lequel nous voulons déterminer la probabilité) qui a une probabilité  $p$  de se réaliser et l'échec qui a une probabilité  $1 - p$  de se réaliser. La v. a  $X$  ( nombre de succès obtenus au cours des  $n$  épreuves ).  $X$  suit la loi binomiale notée  $X \sim \mathcal{B}(n, p)$ , telle que  $n \in \mathbb{N}^*$  et  $p \in [0, 1]$ . La loi est donnée par :*

$$P(X = x) = C_n^x p^x (1 - p)^{n-x} \quad \text{pour } 0 \leq x \leq n.$$

il suffit de dénombrer combien il y a de choix de  $x$  positions parmi  $n$  positions. C'est le nombre de combinaisons à  $x$  éléments que l'on peut former à partir de  $n$  éléments :

$$C_n^x = \frac{n!}{x!(n-x)!}.$$

**Caractéristiques de la loi de Binomiale :**

Espérance, variance et écart-type :

$$E(X) = np, \quad \text{Var}(X) = np(1-p) \quad \text{et} \quad \sigma(X) = \sqrt{np(1-p)}.$$

**Exemple 1.1.4** *On lance une pièce de monnaie 8 fois de suite, alors la probabilité d'obtenir exactement 3 piles est :*

$$P(X = 3) = C_8^3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^{8-3} \simeq 0,219.$$

**Loi de Géométrique  $\mathcal{G}(p)$**

**Définition 1.1.8** *Répétons de façon indépendante une épreuve de Bernoulli de paramètre  $p$  jusqu'à ce que le premier succès ait lieu. La variable aléatoire  $X$  correspondant au temps du premier succès suit la loi géométrique de paramètre  $p$  ;  $X \sim \mathcal{G}(p)$ .*

*La loi de  $X$  est donné par :*

$$P(X = k) = p(1-p)^{k-1} \quad \text{avec } k \in \mathbb{N}, k \geq 1.$$

**Caractéristiques de la loi de Géométrie :**

Espérance, variance et écart-type :

$$E(X) = \frac{1}{p}, \quad \text{Var}(X) = \frac{1-p}{p^2} \quad \text{et} \quad \sigma(X) = \sqrt{\frac{1-p}{p^2}}.$$

**Loi de Poisson  $\mathcal{P}(\lambda)$**

**Définition 1.1.9** *Loi de Poisson convient à la description d'évènements dont les chances de réalisation sont faibles. La probabilité d'observer exactement  $\lambda$  occurrences d'un certain évènement dans une unité de temps. La loi de  $X$  est donnée par :*

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad \text{avec} \quad k \in \mathbb{N}, \quad \lambda > 0.$$

**Caractéristiques de la loi de Poisson :**

Espérance, variance et écart-type :

$$E(X) = \lambda, \quad \text{Var}(X) = \lambda \quad \text{et} \quad \sigma(X) = \sqrt{\lambda}.$$

**Les lois continues**

**La loi uniforme  $\mathcal{U}_{[a, b]}$**

**Définition 1.1.10** *Une variable aléatoire réelle  $X$ , suit une loi uniforme sur l'intervalle  $[a, b]$ , On note  $X \sim \mathcal{U}_{[a, b]}$  si la loi de probabilité admet une densité  $f$  définie par :*

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon} \end{cases}$$

Sa fonction de distribution  $F$  est définie par :

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x < b \\ 1 & \text{si } x \geq b \end{cases}$$

**Caractéristiques de la loi de uniforme :**

Espérance, variance et :

$$E(X) = \frac{a+b}{2}, \quad \text{Var}(X) = \frac{(b-a)^2}{12}.$$

**La loi exponentielle  $\text{exp}(\lambda)$**

**Définition 1.1.11** On dit qu'une variable aléatoire réelle positive  $X$  suit la loi exponentielle de paramètre  $\lambda$  positif notée  $X \sim \text{exp}(\lambda)$ , si sa densité de probabilité est donnée par :

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Sa fonction de distribution est donnée par :

$$F(t) = P(X \leq t) = \int_0^t \lambda e^{-\lambda x} dx = 1 - e^{-\lambda t}.$$

**Caractéristiques de la loi de exponentielle :**

Espérance, variance et :

$$E(X) = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

**Exemple 1.1.5** Soit la variable aléatoire  $X$  qui représente la durée de vie mesurée en heures d'un composant électronique de type donné. Supposons que la densité de probabilité de  $X$  est donnée par :

$$f(x) = \begin{cases} \frac{1}{100} e^{-\frac{x}{100}} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

Sa fonction de distribution est :

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{100}} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$$

La probabilité que la durée de vie d'un composant électronique de ce type soit inférieure à 200 heures est donnée par :

$$\begin{aligned} F(200) &\simeq 1 - 0,135 \\ &= 0,864. \end{aligned}$$

L'espérance de vie d'un tel composant est égale à 100 heures.

**La loi normale  $\mathcal{N}(\mu, \sigma^2)$**

**Définition 1.1.12** Soit  $X$  une variable aléatoire normale (ou gaussienne) de paramètre  $\mu$  et  $\sigma$  notée  $X \sim \mathcal{N}(\mu, \sigma^2)$ , définie sur  $\mathbb{R}$ , alors la fonction de densité est donnée par :

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), x \in \mathbb{R}.$$

**Caractéristiques de la loi de normale :**

Espérance, variance et :

$$E(X) = \mu, \quad \text{Var}(X) = \sigma^2.$$

$$\text{Var}(X) = \sigma^2.$$

**La loi normale centrée réduite**  $\mathcal{N}(0, 1)$

**Définition 1.1.13** *On dit que  $X$  suit une loi normale centrée réduite et on note  $X \sim \mathcal{N}(0, 1)$  si sa loi admet pour densité la fonction :*

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

**Caractéristiques de la loi de normale centrée réduite :**

Espérance, variance et :

$$E(X) = 0 \text{ et } \text{Var}(X) = 1.$$

## 1.2 Convergence d'une suite de variables aléatoires

Une suite  $(X_n)_{n \in \mathbb{N}^*}$  de variables aléatoires réelles est une suite de fonctions mesurables de  $\Omega$  dans  $\mathbb{R}$ . Il existe diverses façons de définir la convergence de  $(X_n)$ . On considère dans la suite, une suite  $(X_n)$  de variables aléatoires réelles et une variable aléatoire  $X$  définie sur le même espace probabilisé  $(\Omega, \mathcal{F}, P)$ .

### 1.2.1 Convergence en probabilité

**Définition 1.2.1** On dit que la suite  $(X_n)$  converge en probabilité (ou converge faiblement) vers  $X$  si :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0.$$

où, de façon équivalente :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| \leq \varepsilon) = 1.$$

On note :  $X_n \xrightarrow[n \rightarrow \infty]{P} X$ .

Pour ce mode de convergence comme pour les suivants la convergence vers une constante  $a$  s'explique naturellement en remplaçant  $X$  par  $a$ .

**Exemple 1.2.1** Soit  $(X_n)_{n \geq 1}$ , la suite de variables aléatoires réelles définies sur un espace de probabilité  $(\Omega, F, P)$  par :  $P(X_n = 0) = 1 - \frac{1}{n}$ ,  $P(X_n = 1) = \frac{1}{n}$ .

Cette suite converge en probabilité vers 0, en effet :

$$\forall \varepsilon > 0, P(|X_n - X| > \varepsilon) = P(|X_n| > \varepsilon) = P(X_n = 1) = \frac{1}{n} \xrightarrow[n \rightarrow \infty]{P} 0.$$

#### **Théorème 1.2.1 (Loi faible des grands nombres)**

Soit  $(X_n)_{n \in \mathbb{N}^*}$ , une suite de variables aléatoires indépendantes 2 à 2, identiquement distribuées, de moyenne  $\mu$  et d'écart-type  $\sigma$ . Alors :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{P} \mu,$$

i.e :  $\forall \varepsilon > 0, P(|X_n - \mu| > \varepsilon) \rightarrow 0$  quand  $n \rightarrow \infty$ .

### 1.2.2 Convergence en loi

Bien qu'elle soit la plus faible, elle est la plus utilisée en pratique car elle permet d'approximer la loi de  $X_n$  par celle de  $X$ .

**Définition 1.2.2** Soient  $(X_n)$  et  $X$  des variables aléatoires sur un même espace probabilisé  $(\Omega, \mathcal{F}, P)$ , de fonctions de distribution respectives  $F_n$  et  $F$ , on dit que les  $(X_n)$  convergent vers  $X$  en loi (et on note  $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X$ ) si en tout point  $x$  où  $F$  est continue, les  $F_n(x)$  convergent vers  $F(x)$ .

**Théorème 1.2.2** La suite  $(X_n)$  de variables aléatoires à valeurs dans  $\mathbb{R}^n$  converge en loi vers la variable aléatoire  $X$  à valeurs dans  $\mathbb{R}^m$  si et seulement si la fonction caractéristique de  $X_n$  converge ponctuellement vers la fonction caractéristique de  $X$ .

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \Leftrightarrow \forall t \in \mathbb{R}^n \Phi_{X_n}(t) \rightarrow \Phi_X(t).$$

**Proposition 1.2.1** (Convergence de la loi binomiale vers une loi de Poisson) :

Soit  $(X_n)$  une suite de variables aléatoires binomiales sur un même espace probabilisé : pour tout  $n$ ,  $X_n \sim \mathcal{B}(n, p_n)$ . On suppose que  $\lim_{n \rightarrow \infty} p_n = 0$  et  $\lim_{n \rightarrow \infty} np_n = \lambda$ . Alors  $(X_n)$  converge en loi, quand  $n$  tend vers l'infini, vers une loi de poisson de paramètre  $\lambda$  ( $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{P}(\lambda)$ ).

### 1.2.3 Convergence presque-sûre

**Définition 1.2.3** On dit que  $(X_n)$  converge presque sûrement vers  $X$  (ou avec une probabilité égale à 1) si :

$$P\left(\lim_{n \rightarrow \infty} X_n = X\right) = 1 \quad \text{ou bien} \quad P\left(\lim_{n \rightarrow \infty} X_n \neq X\right) = 0$$

et on note  $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ .

**Proposition 1.2.2**

a) Si  $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ , alors  $X_n \xrightarrow[n \rightarrow \infty]{p} X$ .

b)  $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ , si et seulement si, pour tout  $\delta > 0$ ,  $P\left(\limsup_{n \rightarrow \infty} \{|X_n - X| > \delta\}\right) = 0$ .

En particulier, si  $\sum_{n \geq 0} (|X_n - X| > \delta) < \infty$  pour tout  $\delta > 0$ , alors  $X_n \xrightarrow[n \rightarrow \infty]{p.s.} X$ .

c) Si  $X_n \xrightarrow[n \rightarrow \infty]{p} X$ , alors il existe une sous-suite  $(X_{\varphi(x)})$  qui converge vers  $X$  p. s.

**Théorème 1.2.3 (Loi forte des grands nombres)**

Soit  $(X_n)$  une suite de variables aléatoires, indépendantes de moyenne  $\mu$  et de variance  $\sigma^2$ , alors la suite des moyennes  $\bar{X}_n$  converge presque sûrement vers  $\mu$  :

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow[n \rightarrow \infty]{p.s.} \mu.$$

**1.2.4 Convergence en moyenne d'ordre  $p$**

Si  $E(|X_n - X|^p)$  existe alors on a :

**Définition 1.2.4** On dit que la suite de v. a.  $(X_n)$  converge en moyenne d'ordre  $p$ , avec  $0 < p < \infty$ , vers la v. a.  $X$  si :

$$E(|X_n - X|^p) \rightarrow 0 \quad \text{quand } n \rightarrow \infty,$$

on note :  $X_n \xrightarrow[n \rightarrow \infty]{L^p} X$ .

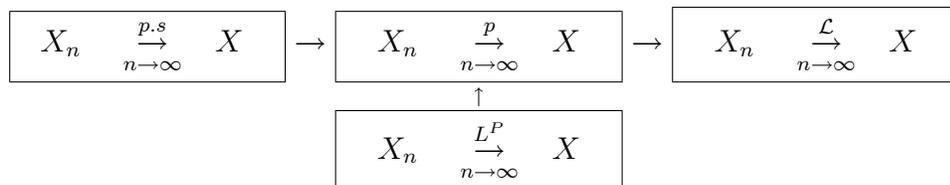
Dans le cas particulier  $p = 2$ , la convergence en moyenne d'ordre 2 s'appelle convergence en moyenne quadratique (m. q.).

### 1.2.5 Théorème central limite

**Théorème 1.2.4** *Si  $(X_n)$  une suite de variables aléatoires indépendantes et de même loi telles que  $E(X_n) = \mu$  et  $Var(X_n) = \sigma^2$  pour tout entier  $n$ , alors :*

$$\sqrt{n} \frac{\bar{X}_n - \mu}{\sigma} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

**Remarque 1.2.1** *Le schéma suivant résume les implications entre les différents types de convergence :*



## 1.3 Estimation de la fonction de distribution

### 1.3.1 Définitions de base

Nous présentons tout d'abord quelques rappels et définitions essentiels dans notre étude.

#### Population

la population est l'ensemble des individus d'intérêt d'une étude, que ce soient des patients, des plantes, des insectes ou différents lancers d'une pièce de monnaie, avant d'entreprendre une étude ou une expérience, il s'agit de définir autant précisément que possible qui nous intéresse.

## Echantillon

**Définition 1.3.1** *Un échantillon  $X_1, X_2, \dots, X_n$  est un  $n$ -uplet  $(X_1, X_2, \dots, X_n)$  de variables aléatoires  $X_i$  indépendantes et identiquement distribuées (même loi).*

Par simplicité nous utiliserons régulièrement le terme échantillon pour signifier à la fois l'échantillon d'observations  $x_1, \dots, x_n$  et le  $n$ -uplet aléatoire  $(X_1, X_2, \dots, X_n)$ .

## Statistique d'ordre d'un échantillon

**Définition 1.3.2** *La statistique d'ordre de l'échantillon  $(X_1, X_2, \dots, X_n)$  est le ré-arrangement croissant (aléatoire) de  $(X_1, X_2, \dots, X_n)$  telle que  $X_1 \leq X_2 \leq \dots \leq X_n$ , que l'on note  $(X_{(1, n)}, X_{(2, n)}, \dots, X_{(n, n)})$ . En particulier, on note :  $X_{(1, n)} = \min_{1 \leq i \leq n} (X_i)$  et  $M_n = X_{(n, n)} = \max_{1 \leq i \leq n} (X_i)$ .*

On va s'intéresser à cette variable aléatoire  $M_n$  dans le suivant.  $M_n$  représente la plus grande observée sur les  $n$  observées. Comme les variables aléatoires sont indépendantes et identiquement distribuées, on obtient :

1. La distribution du maximum  $F_{X_{(n, n)}}$  de la statistique d'ordre extrême  $M_n$  est donnée par :  $\forall x \in \mathbb{R}$

$$\begin{cases} F_{M_n}(x) = P(X_{(n, n)} \leq x) = [F(x)]^n, \\ f_{X_{(n, n)}}(x) = n[F(x)]^{n-1} f(x). \end{cases}$$

2. La distribution du minimum  $F_{X_{(1, n)}}$  de la statistique d'ordre extrême  $X_{(1, n)}$  est donnée par :  $\forall x \in \mathbb{R}$

$$\begin{cases} P(X_{(1, n)} \leq x) = 1 - [1 - F(x)]^n, \\ f_{X_{(1, n)}}(x) = n[1 - F(x)]^{n-1} f(x). \end{cases}$$

**3.** La distribution  $F_{X_{(i,n)}}$  de la statistique d'ordre extrême  $X_{(i,n)}$  est donnée par :

$$\begin{cases} F(x)_{X_{(i,n)}} = P(X_{(i,n)} \leq x) = \sum_{r=i}^n C_n^r [F(x)]^r [1 - F(x)]^{n-r}, \\ f_{X_{(i,n)}}(x) = \frac{n!}{(i-1)!(n-i)!} [F(x)]^{i-1} [1 - F(x)]^{n-i} f(x), \end{cases}$$

où  $F(x)$  est la fonction de distribution des  $X_i$  et  $f$  est la densité des  $X_i$ .

### estimateur

**Définition 1.3.3** Soit  $X$  une v. a. dont sa loi dépend d'un paramètre inconnu  $\theta$ .

Soit  $(X_1, \dots, X_n)$  un n-échantillon de  $X$  et  $(x_1, \dots, x_n)$  sa réalisation. Il s'agit d'estimer le paramètre  $\theta$ . un **estimateur** de  $\theta$  sera une statistique :

$$T = f(X_1, \dots, X_n),$$

et sa réalisation sera notée :

$$\hat{\theta} = (x_1, \dots, x_n).$$

### Qualité d'un estimateur

1. **estimateur avec biais** : On appelle **biais** de  $T$  pour  $\theta$  la valeur :

$$E(T) = b_\theta(T) + \theta.$$

2. **estimateur sans biais** : Un estimateur  $T$  de  $\theta$  est dit **sans biais** si :

$$E(T) = \theta, \quad (\text{ou bien } b_\theta(T) = 0).$$

**Exemple 1.3.1** La moyenne empirique  $\bar{X}$  est un estimateur sans biais du

paramètre  $\lambda$  d'une loi de poisson :

$$E(\bar{X}) = \lambda.$$

La variance empirique  $S_n^2$  est estimateur biaisé du même paramètre  $\lambda$  :

$$E(S_n^2) = \frac{n-1}{n}\lambda,$$

où  $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2$ .

3. **estimateur asymptotiquement sans biais** : Un estimateur  $T$  de  $\theta$  est dit **asymptotiquement sans biais** si :

$$\lim_{n \rightarrow \infty} E(T) = \theta.$$

4. **estimateur convergent** :

**Définition 1.3.4** Un estimateur  $T$  est dit **convergent** si  $E(T)$  tend vers  $\theta$  lorsque  $n$  tend vers l'infini. Il sera dit **consistant** si  $T$  converge vers  $\theta$  lorsque  $n$  tend vers l'infini.

**Théorème 1.3.1** si  $T$  convergent et de variance tendant vers 0 lorsque  $n$  tend vers l'infini alors  $T$  est consistant.

5. **Erreur quadratique moyenne** :

**Définition 1.3.5** La qualité d'un estimateur se mesure également par **l'erreur quadratique moyenne** (ou **risque quadratique**) qui est définie par :

$$E((T - \theta)^2).$$

**Théorème 1.3.2** Soit  $T$  Un estimateur du paramètre  $\theta$  à étudier. On a :

$$E((T - \theta)^2) = \text{Var}(T) + [E(T) - \theta]^2.$$

### 1.3.2 Estimation paramétrique

Dans les situations où il n'y a pas d'estimateur évident, on est amené à recourir à une méthode de construction d'un estimateur, les trois méthodes que nous présenterons ici étant celles du maximum de vraisemblance et des moments et la méthode d'estimation par intervalle de confiance.

#### La méthode des moments

Soit  $(X_1, X_2, \dots, X_n)$  un  $n$ -échantillon issu d'une variable aléatoire  $X$  de densité  $f(x, \theta_1, \theta_2, \dots, \theta_k)$  où :  $\theta_1, \theta_2, \dots, \theta_k$  sont es paramètres inconnus.

La méthode des moments consiste à estimer les paramètres  $\theta_1, \theta_2, \dots, \theta_k$ , en égalisant les moments empiriques calculés à partir de l'échantillon avec les moments théoriques de même ordre.

Soit  $\mu_r = E(X^r)$ ,  $r = 1, 2, \dots, k$  moments d'ordre  $r$  de la population (théorique) et on not  $\bar{X}_r = \frac{1}{n} \sum_{i=1}^n X_i^r$  moment empirique d'ordre  $r$  de l'échantillon.

La solution du système  $\bar{X}_r = \mu_r$ ,  $r = \overline{1, k}$  nous donne les estimateurs de  $\theta_1, \theta_2, \dots, \theta_k$ .

$$\begin{cases} \bar{X}_1 = \mu_1, \\ \bar{X}_2 = \mu_2, \\ \vdots \\ \bar{X}_k = \mu_k, \end{cases} \quad k \text{ équations à } k \text{ inconnus.}$$

**Remarque 1.3.1** Dans la plupart des cas, les estimateurs obtenus par la méthode des moments sont consistants, convergents, asymptotiquement normaux mais en gé-

*néral ne sont pas efficaces.*

**Exemple 1.3.2** *loi exponentielle*

Si  $X_1, X_2, \dots, X_n$  sont indépendantes et de même loi exponentielle  $\text{exp}(\lambda)$ ,  $E(X) = \frac{1}{\lambda}$ .

Donc l'estimateur de  $\lambda$  par la méthode des moments est  $\hat{\lambda} = \frac{1}{\bar{X}}$ .

**La méthode du maximum de vraisemblance**

La méthode du maximum de vraisemblance permet de trouver des estimateurs dans toutes les situations, même les plus compliquées. C'est une des méthodes d'estimation les plus utilisées. Cette méthode consiste à rechercher le paramètre  $\theta$  qui maximise la fonction de vraisemblance  $L(x_1, \dots, x_n; \theta)$ , c'est-à-dire pour lequel la densité de l'échantillon est la plus grande.

**Définition 1.3.6** *Quand les observations sont toutes discrètes ou toutes continues, on appelle **fonction de vraisemblance** (ou plus simplement vraisemblance) pour l'échantillon  $x_1, \dots, x_n$ , la fonction du paramètre  $\theta$  :*

$$L(x_1, \dots, x_n; \theta) = \begin{cases} P(X_1 = x_1, \dots, X_n = x_n; \theta) & \text{si les } X_i \text{ sont discrètes} \\ f_{X_1, \dots, X_n}(x_1, \dots, x_n; \theta) & \text{si les } X_i \text{ sont continues} \end{cases}$$

*Dans tous les exemples que nous traiterons ici, les  $X_i$  sont indépendantes et de même loi. Dans ce cas, la fonction de vraisemblance s'écrit :*

$$L(x_1, \dots, x_n; \theta) = \begin{cases} \prod_{i=1}^n P(X_i = x_i; \theta) = \prod_{i=1}^n P(X = x_i; \theta) & \text{si les } X_i \text{ sont discrètes} \\ \prod_{i=1}^n f_{X_i}(x_i; \theta) = \prod_{i=1}^n f(x_i; \theta) & \text{si les } X_i \text{ sont continues} \end{cases}$$

**Remarque 1.3.2** *La probabilité et la densité utilisées dans cette définition sont des fonctions des observations  $x_1, \dots, x_n$ , dépendant du paramètre  $\theta$ . A l'inverse, la fonc-*

tion de vraisemblance est considérée comme une fonction de  $\theta$  dépendant des observations  $x_1, \dots, x_n$ , ce qui permet, par exemple, de dériver cette fonction par rapport à  $\theta$ .

**Définition 1.3.7** *L'estimation de maximum de vraisemblance (EMV) de  $\theta$  est la valeur  $\hat{\theta}$  de  $\theta$  qui rend maximale la fonction de vraisemblance  $L(x_1, \dots, x_n; \theta)$ .*

L'estimateur de maximum de vraisemblance de  $\theta$  est la variable aléatoire correspondante.

La fonction de vraisemblance s'exprime comme un produit. Donc  $\hat{\theta}$  sera en général calculé en maximisant la log-vraisemblance :

$$\hat{\theta} = \arg \max_{\theta} \ln L(x_1, \dots, x_n; \theta).$$

Quand  $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$  et que toutes les dérivées partielles ci-dessous existent,  $\hat{\theta}$  est solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in \{1, \dots, d\}, \quad \frac{\partial}{\partial \theta_j} \ln L(x_1, \dots, x_n; \theta) = 0.$$

### Exemple 1.3.3

#### 1. Avec une loi discrète :

On souhaite estimer le paramètre  $\lambda$  d'une loi de Poisson  $\mathcal{P}(\lambda)$ , telle que :  $P(X = x) = e^{-\lambda} \frac{\lambda^x}{x!}$ . La fonction de vraisemblance est :

$$L(x_1, \dots, x_n; \lambda) = \prod_{i=1}^n e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} = e^{-\lambda n} \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!}.$$

$$\ln L(x_1, \dots, x_n; \lambda) = \ln e^{-\lambda n} + \ln \prod_{i=1}^n \frac{\lambda^{x_i}}{x_i!} = -\lambda n + \sum_{i=1}^n \ln \frac{\lambda^{x_i}}{x_i!} = -\lambda n + \ln \lambda \cdot \sum_{i=1}^n x_i - \sum_{i=1}^n \ln(x_i!).$$

$$\frac{\partial}{\partial \lambda} \ln L(x_1, \dots, x_n; \lambda) = -n + \frac{\sum_{i=1}^n x_i}{\lambda}, \text{ qui s'annule pour } \lambda = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}.$$

Par conséquent, l'EMV de  $\lambda$  est  $\hat{\lambda} = \bar{X}$ .

## 2. Avec une loi continue :

On souhaite estimer les paramètres  $\mu$  et  $\sigma$  d'une loi normale  $\mathcal{N}(\mu, \sigma^2)$ , telle que :

$f(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ . la fonction de vraisemblance est :

$$L(x_1, \dots, x_n; \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right) = \frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

D'où  $\ln L(x_1, \dots, x_n; \mu, \sigma^2) = -\frac{n}{2} \ln \sigma^2 - \frac{n}{2} \ln 2\pi - \frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2$ .

On doit annuler les dérivées partielles de ce logarithme par rapport à  $\mu$  et  $\sigma^2$ . On a :

$$-\frac{\partial}{\partial \mu} \ln L(x_1, \dots, x_n; \mu, \sigma^2) = -\frac{1}{2\sigma^2} \sum_{i=1}^n -2(x_i - \mu) = \frac{1}{\sigma^2} \left( \sum_{i=1}^n x_i - n\mu \right), \text{ qui s'annule}$$

pour  $\mu = \frac{1}{n} \sum_{i=1}^n x_i = \bar{x}$ .

$$-\frac{\partial}{\partial \sigma^2} \ln L(x_1, \dots, x_n; \mu, \sigma^2) = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^n (x_i - \mu)^2, \text{ qui s'annule pour } \sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2.$$

On a donc  $\hat{\mu} = \bar{X}_n$  et  $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n (X - \bar{X}_n)^2 = S_n^2$ .

## Estimation par intervalles de confiance

Soit  $X$  une variable aléatoire dont la loi dépend d'un paramètre  $\theta$  inconnu. Soit  $(X_1, X_2, \dots, X_n)$  un n-échantillon issu de  $X$  et  $\alpha \in ]0, 1[$ .

**Définition 1.3.8** *Un intervalle de confiance (IC) pour le paramètre  $\theta$ , de niveau de confiance  $1 - \alpha \in ]0, 1[$ , est l'intervalle qui a la probabilité  $1 - \alpha$  de contenir la vraie valeur du paramètre  $\theta$ .*

### Remarque 1.3.3

- Un intervalle de confiance indique la précision d'une estimation car pour un risque  $\alpha$  donné, l'intervalle est d'autant plus grand que la précision est faible.

- Les niveaux usuels sont 90%, 95% et 99% et correspondent respectivement à  $\alpha = 10%$ ,  $\alpha = 5%$  et  $\alpha = 1%$ .

**Intervalle de confiance pour l'espérance d'une loi normale  $\mathcal{N}(\mu, \sigma^2)$  :**

**a) Cas où variance est connue :**

Lorsque  $\sigma^2$  est inconnu un intervalle de confiance au niveau  $1 - \alpha$  de  $\mu$  est :

$$IC(\mu) = \left[ \bar{X} - q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{X} + q_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \right]$$

où  $q_{1-\frac{\alpha}{2}}$  est le quantile d'ordre  $1 - \frac{\alpha}{2}$  ( $F(q_{1-\frac{\alpha}{2}}) = 1 - \frac{\alpha}{2}$ ) de la loi normale centrée réduite avec  $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ .

**b) Cas où variance est inconnue :**

Lorsque  $\sigma^2$  est inconnu un intervalle de confiance au niveau  $1 - \alpha$  de  $\mu$  est :

$$IC(\mu) = \left[ \bar{X} - t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}}, \bar{X} + t_{n-1, 1-\frac{\alpha}{2}} \frac{S}{\sqrt{n}} \right]$$

où  $t_{n-1, 1-\frac{\alpha}{2}}$  est le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi de Student à  $n - 1$  degrés de liberté avec

$$S'^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

**Exemple 1.3.4** *Sur les cinq dernières années, on a observé les températures moyennes suivantes en degrés celsius au mois d'août à Biskra :*

$X = (X_1, \dots, X_5) = (45, 42, 40, 47, 41)$  telles que  $\bar{X}_5 = 43$  et  $S_5 = 2.92$ .

D'après la table de statistique de la loi de student,  $t_{0.975}^4 = 2.776$  et  $\frac{t_{4,0.975} \times S_5}{\sqrt{5}} = 3.63$

On en déduit que  $IC(\mu) = [39.37, 46.63]$  est un intervalle de confiance à 95% pour l'espérance de la températures moyennes au mois d'août.

**Intervalle de confiance pour la variance d'une loi normale  $\mathcal{N}(\mu, \sigma^2)$  :**

**a) Cas où moyenne est connue :**

Lorsque  $\mu$  est connu un intervalle de confiance au niveau  $1 - \alpha$  de  $\sigma^2$  est :

$$IC(\sigma^2) = \left[ \frac{1}{u_1} \sum_{i=1}^n (X_i - \mu)^2, \frac{1}{u_2} \sum_{i=1}^n (X_i - \mu)^2 \right],$$

où  $u_1$  et  $u_2$  sont les quantiles d'ordre  $\frac{\alpha}{2}$  et  $1 - \frac{\alpha}{2}$  de la loi du  $\mathcal{X}^2$  à  $n$  degrés de liberté.

**b) Cas où moyenne est inconnue :**

Lorsque  $\mu$  est inconnu un intervalle de confiance au niveau  $1 - \alpha$  de  $\sigma^2$  est :

$$IC(\sigma^2) = \left[ \frac{1}{u_1} \sum_{i=1}^n (X_i - \bar{X})^2, \frac{1}{u_2} \sum_{i=1}^n (X_i - \bar{X})^2 \right],$$

où  $u_1$  et  $u_2$  sont les quantiles d'ordre  $\frac{\alpha}{2}$  et  $1 - \frac{\alpha}{2}$  de la loi du  $\mathcal{X}^2$  à  $n - 1$  degrés de liberté.

**Intervalle de confiance sur une proportion :** L'intervalle de confiance de la proportion  $p$  avec un coefficient de confiance de  $1 - \alpha$  est :

$$IC(\sigma^2) = \left[ f - q_{1-\frac{\alpha}{2}} \alpha \sqrt{\frac{f(1-f)}{n}}, f + q_{1-\frac{\alpha}{2}} \alpha \sqrt{\frac{f(1-f)}{n}} \right],$$

où  $q_{1-\frac{\alpha}{2}}$  est le quantile d'ordre  $1 - \frac{\alpha}{2}$  de la loi  $\mathcal{N}(0, 1)$ .

# Chapitre 2

## Estimation non paramétrique de la fonction de distribution

La fonction de distribution  $F$  caractérise la loi de probabilité d'une variable aléatoire. Elle permet d'avoir un aperçu des principales caractéristiques de la distribution et c'est en terme de comportement locale de la fonction de distribution que s'explique le plus facilement. Le comportement des estimateurs fonctionnels (vitesse de convergence, normalité, asymptotique) et c'est finalement par un estimateur de la fonction de distribution que l'on passe pour estimer les probabilités d'ensembles. l'objectif de ce chapitre est de présenté la méthode d'estimation non paramétrique de fonction de distribution, dont nous allons présenter une étude de l'estimateur par la méthode du noyau ainsi que ses propriétés statistiques.

### 2.1 Fonction de distribution empirique

Soit  $X \sim F$ , avec  $F(x) = P\{X \leq x\}$  la fonction de distribution de  $X$ .

Soit  $X_1, X_2, \dots, X_n$  un échantillon iid de  $F$  (iid= indépendantes et identiquement distribuées) et

$X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$  les observations ordonnées.

Supposons que  $F$  soit complètement inconnue.

Un bon estimateur pour  $F$  est la fonction de distribution empirique, notée  $F_n$ , et définie par :

$$\begin{aligned}
 F_n(x) &= \frac{\text{nombre d'observations} \leq x}{n} \\
 &= \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}} = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_{(i)} \leq x\}} \tag{2.1} \\
 &= \begin{cases} 0 & \text{si } x < X_{(1)} \\ \frac{i}{n} & \text{si } X_{(i)} \leq x < X_{(i+1)} \quad k = 1, \dots, n-1 \\ 1 & \text{si } x \geq X_{(n)} \end{cases}
 \end{aligned}$$

où

$$\mathbb{I}_{\{X_i \leq x\}} = \begin{cases} 1 & \text{si } X_i \leq x \\ 0 & \text{si } X_i > x. \end{cases}$$

### 2.1.1 Propriétés statistiques

1. Biais de l'estimateur  $F_n(x)$  :

$$E(F_n(x)) = E\left(\frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}}\right) = P(X \leq x) = F(x).$$

Alors  $F_n(x)$  est un estimateur sans biais de  $F(x)$ .

2. Variance de l'estimateur  $F_n(x)$  :

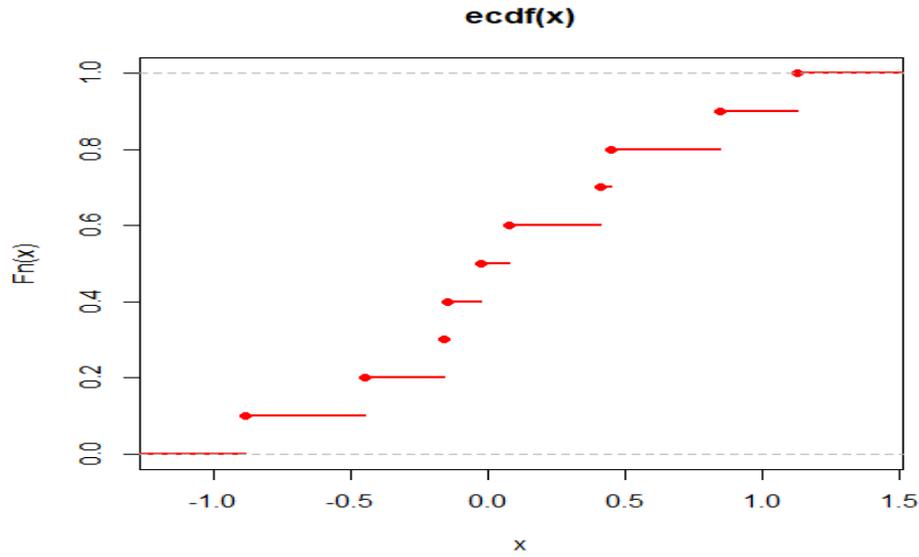


FIG. 2.1 – Fonction de répartition empirique.

$$\begin{aligned} E((F_n(x))^2) &= E\left(\frac{1}{n}\sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}}\right)^2 = \frac{1}{n^2} E\left(\frac{1}{n}\sum_{i=1}^n \mathbb{I}_{\{X_i \leq x\}}^2 + \sum_{i \neq j} \mathbb{I}_{\{X_i \leq x\}} \mathbb{I}_{\{X_j \leq x\}}\right) \\ &= \frac{1}{n} F(x) + \frac{n-1}{n} F^2(x). \end{aligned}$$

D’ou

$$\text{Var}(F_n(x)) = E((F_n(x))^2) - E^2(F_n(x)) = \frac{1}{n} F(x)(1 - F(x)).$$

**3.** Convergence de  $F_n(x)$  vers  $F(x)$  :

**Théorème 2.1.1** Soit  $(X_1, X_2, \dots, X_n)$  un  $n$ -échantillon de fonction de distribution empirique  $F_n(x)$  et  $F(x)$  fonction de distribution de  $X$  (variable aléatoire parente). Alors :  $F_n(x) \xrightarrow[n \rightarrow \infty]{p-s} F(x)$ .

**Preuve.**  $F_n(x)$  étant une moyenne empirique de variable aléatoire réelle indépendante (puisque les  $X_i$  le sont), d’après la loi forte des grands nombres : ■

$$F_n(x) \xrightarrow{p.s} E [\mathbb{I}_{\{X_i \leq x\}}] = E [\mathbb{I}_{\{X \leq x\}}] = F(x).$$

4. Convergence en loi de  $F_n(x)$  :

On a :  $\mathbb{I}_{\{X_i \leq x\}}$  est une variable aléatoire de Bernoulli  $\mathcal{B}(F(x))$ ,

$$P(F_n(x) = \frac{i}{n}) = P(nF_n(x) = i) = C_n^i (F(x))^i (1 - F(x))^{n-i},$$

donc  $nF_n(x) \sim \mathcal{B}(n, F(x))$ . Alors d'après le théorème central limite :

$$\frac{\sqrt{n}(F_n(x) - F(x))}{\sqrt{F(x)(1 - F(x))}} \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(0, 1).$$

**Théorème 2.1.2** (*Glivenko-Contelli*)

La convergence de  $F_n$  vers  $F$  est presque sûrement uniforme, i.e :

$$D_n = \sup_x |F_n(x) - F(x)| \xrightarrow[n \rightarrow \infty]{p.s} 0.$$

**Théorème 2.1.3** (*Kolmogorov-Smirnov*)

$$\lim_{n \rightarrow \infty} P(\sqrt{n}D_n < \lambda) = 2 \sum_{-\infty}^{\infty} (-1)^n \exp(-2n^2 \lambda^2),$$

Ce théorème signifie que la distribution asymptotique de la variable aléatoire  $D_n$  est connue et ne dépend pas de la variable de départ  $X$ , et permet de calculer des limites pour les valeur de  $D_n$ . La loi exacte de la variable  $D_n$  a été tabulée.

## 2.2 Estimation de la fonction de distribution par la méthode du noyau

La méthode du noyau est la plus populaire parmi les multiples méthodes d'estimation non paramétriques de la densité. Cet estimateur est une fonction de deux paramètres : Le noyau  $k$  et le paramètre de lissage  $h$ .

Soient  $X_1, \dots, X_n$  des variables aléatoires (iid) de fonction de distribution continue  $F$  et de densité  $f$ .

L'estimateur à noyau de la densité s'écrit sous la forme :

$$\hat{f}_n(x) = \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x - X_i}{h}\right), \quad (2.2)$$

avec un noyau  $k$  intégrable et d'intégrale 1 et une fenêtre  $h > 0$ , est un estimateur non paramétrique bien connu de  $f(x)$  introduit par **Rosenblatt** (1956) et a été généralisé par **Parzen** (1962). La littérature qu'il a suscitée est considérable. Dans le présent paragraphe nous nous limitons à ses applications orientées vers l'estimation de la fonction de distribution. Pour obtenir un estimateur non paramétrique de  $F$ , nous intégrons l'estimateur de la densité, cet estimateur de  $F$  est basé sur l'estimateur à noyau de type **Nadaraya** (1964), qui est défini par :

$$\begin{aligned} \hat{F}_n(x) &= \int_{-\infty}^x \hat{f}_n(t) dt = \int_{-\infty}^x \frac{1}{nh} \sum_{i=1}^n k\left(\frac{x - X_i}{h}\right) dt \\ &= \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right), \end{aligned}$$

où la fonction  $K$  est définie par :

$$K(x) = \int_{-\infty}^x k(t) dt.$$

$k : \mathbb{R} \rightarrow \mathbb{R}^+$  est une fonction intégrable dite noyau telle que  $\int_{-\infty}^{+\infty} k(t) dt = 1$  et vérifiant généralement les conditions suivantes :

$$\int_{-\infty}^{+\infty} t^2 |k(t)| dt = 1, \quad \int_{-\infty}^{+\infty} tk(t) dt = 0 \quad \text{et} \quad \int_{-\infty}^{+\infty} t^2 k(t) dt < \infty.$$

- Le paramètre  $h = h_n > 0$  est une fenêtre ou (paramètre de lissage) qui satisfait  $h = h_n \rightarrow 0$  quand  $n \rightarrow \infty$ .

Le tableau suivant présente exemple de quelque noyau continus symétrique et leurs formes sont présentées dans la figure 2.2 :

Noyau	$k(t)$
Uniforme	$\frac{1}{2}\mathbb{I}_{( t \leq 1)}$
Triangulaire	$(1 -  t )\mathbb{I}_{( t \leq 1)}$
Epanechnikov	$\frac{3}{4}(1 - t^2)\mathbb{I}_{ t \leq 1}$
Quartique	$\frac{15}{16}(1 - t^2)\mathbb{I}_{( t \leq 1)}$
Gaussien	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right), t \in \mathbb{R}$
Cosinus	$\frac{\pi}{4} \cos\left(\frac{\pi}{2}t\right)\mathbb{I}_{( t \leq 1)}$

TAB. 2.1 – Tableau représente quelques exemples de noyaux

voici quelques courbes de noyaux usuels présentées ci-dessous :

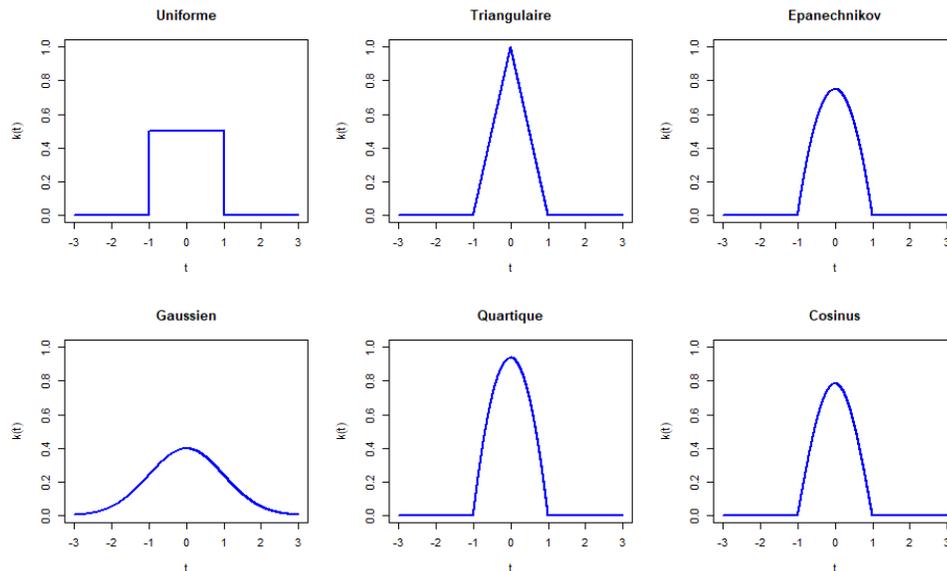


FIG. 2.2 – Les courbes des noyaux.

Ses propriétés sont connues depuis longtemps, par exemple sa convergence uniforme

vers  $F$  avec  $f$  continue **Nadaraya** (1964), **Winter** (1973), **Yamato** (1973) puis sans conditions sur  $f$  **Singh** (1983) ou sa normalité asymptotique ( **Watson** et **Leadbetter** ). **Winter** démontre aussi qu'il vérifie la propriété de **Chung-Smirnov**, c'est-à-dire :

$$\lim_{n \rightarrow \infty} \sup \left\{ \left( \frac{2n}{\log \log n} \right)^{\frac{1}{2}} \sup_{x \in \mathbb{R}} \left| \hat{F}_n(x) - F_n(x) \right| \right\} \leq 1,$$

avec probabilité 1. **Azzalini** trouve une expression asymptotique pour l'erreur quadratique moyenne (MSE) et détermine la fenêtre asymptotiquement optimale permettant d'avoir une (MSE) plus faible que pour  $F_n$ , **Reiss** prouve que l'inefficacité relative asymptotique de  $F_n$  par rapport à  $\hat{F}_n$  tend rapidement vers l'infini quand la taille de l'échantillon augmente avec un choix approprié de noyau, par exemple :

$$k(x) = \frac{9}{8} \left( 1 - \frac{5}{3} x^2 \right) \mathbb{I}_{|x| \leq 1}.$$

Et certaines conditions vérifiées notamment lorsque le support de  $k$  est borné et :

$$\int_{-\infty}^{+\infty} x k(x) K(x) > 0. \tag{2.3}$$

**Falk** (1983) donne ensuite une solution complète à ce problème en établissant la représentation de l'inefficacité relative de  $F_n$  par  $\hat{F}_n$  rapport à sous les conditions ci-dessus notamment lorsque le support de  $k$  est borné. Le nombre  $\varphi(k) = \int 2xk(x)K(x)dx$  est introduit par **Falk** (1984) comme une mesure de la performance asymptotique du noyau  $k$ . Mais il démontre qu'aucun noyau de carré intégrable ne minimise  $\varphi$ . Il utilise alors le nombre  $\rho(k) = \int k(x)^2 dx$  défini par **Epanechnikov** (1969) comme une mesure de la performance du noyau en estimation de la densité. Au sens de  $\rho$ , le

noyau d'**Epanechnikov** suivant :

$$k(x) = \frac{3}{4}(1 - x^2)\mathbb{I}_{|x| \leq 1},$$

est le meilleur mais les noyaux gaussiens ou uniformes ont des performances très proches. En utilisant le critère  $\varphi$  le noyau d'Epanechnikov est alors de loin le meilleur des trois.

Au sens de l'erreur quadratique moyenne intégrée (IMSE), le meilleur noyau est le noyau uniforme bien que les performances d'autres noyaux (Epanechnikov, normal, triangulaire) ne soient, en pratique, que légèrement moins bonnes **Jones** (1990). Il est intéressant de noter que ce ne sera pas le meilleur noyau dans le cadre d'estimation de la densité.

L'expression asymptotique de l'IMSE. est également étudiée par Swanepoel (1988). Pour une fonction  $f$  continue, il prouve que le meilleur noyau est le noyau uniforme  $k(x) = \left(\frac{1}{2\omega}\right) \mathbb{I}_{[-\omega, +\omega]}(x)$  pour une constante arbitraire  $\omega > 0$  (ce qui démontre que les critères de **Falk** pour définir un noyau optimal ne sont en fait pas adaptés à la fonction de distribution) alors que, pour  $f$  discontinue en un nombre fini de points, c'est le noyau exponentiel  $k(x) = \frac{c}{2} \exp(-c|x|)$  pour une constante arbitraire  $c > 0$ .  $\hat{F}_n$  est ici aussi plus efficace que  $F_n$  pour  $h_n = o(n^{-1/2})$  sous la condition (2.3).

Néanmoins,  $\hat{F}_n$  ne fournit pas toujours une meilleure estimation que  $F_n$ . En effet, dans le cas d'une fonction  $F$  uniformément lipchitzienne **Fernholz** (1991) obtient que :

$$\sqrt{n} \left\| \hat{F}_n - F_n \right\|_{\infty} \rightarrow 0 \text{ p. s.,}$$

et que  $\sqrt{n} \left\| \hat{F}_n - F \right\|_{\infty}$  et  $\sqrt{n} \|F_n - F\|_{\infty}$  ont la même distribution asymptotique. De plus, **Shirahata** et **Chu** (1992) démontrent que sous certaines hypothèses sur  $F$ , l'erreur quadratique intégrée (MISE)  $(\int_{-\infty}^{+\infty} (\hat{F}_n(x) - F_n(x))^2 dF(x))$  de  $\hat{F}_n$  est presque sûrement supérieure à celle de  $F_n$ .

### 2.2.1 Erreur quadratique moyenne

Nous obtenons d'abord la MSE. L'hypothèse utilisée par **Azzalini** (1981) qui sont que  $f$  est continu et différentiable avec la moyenne finie et les dérivées intégrables carrées,  $h \rightarrow 0$  et  $nh \rightarrow \infty$  comme  $n \rightarrow \infty$ , et le noyau satisfaisant les conditions ci-dessus. Nous avons :

$$E\left(\hat{F}_n(x)\right) = F(x) + \frac{h^2 f'(x) \mu_2(k)}{2} + o(h^2),$$

$$Biais\left(\hat{F}_n(x)\right) = E\left(\hat{F}_n(x)\right) - F(x) = \frac{h^2 f'(x) \mu_2(k)}{2} + o(h^2). \quad (2.4)$$

Et

$$Var\left(\hat{F}_n(x)\right) = n^{-1}F(x)(1-F(x)) - n^{-1}hf(x)\varphi(k) + o\left(\frac{h}{n}\right), \quad (2.5)$$

où

$$\mu_2(k) = \int_{-\infty}^{\infty} x^2 k(x) dx, \quad \varphi(k) = 2 \int_{-\infty}^{\infty} xk(x) K(x) dx.$$

De (2.4) et (2.5) l'erreur quadratique moyenne de  $\hat{F}_n$  est donne par :

$$MSE\left(\hat{F}_n(x)\right) = Biais^2\left(\hat{F}_n(x)\right) + Var\left(\hat{F}_n(x)\right).$$

Donc

$$MSE\left(\hat{F}_n(x)\right) = n^{-1}F(x)(1-F(x)) - n^{-1}hf(x)\varphi(k) + \frac{h^4 f'^2(x) \mu_2^2(k)}{4} + o\left(h^4 + \frac{h}{n}\right). \quad (2.6)$$

L'erreur quadratique moyenne asymptotique AMSE de  $\hat{F}_n$  est :

$$AMSE \left( \hat{F}_n(x) \right) = n^{-1} F(x) (1 - F(x)) - n^{-1} h f(x) \varphi(k) + \frac{h^4 f'^2(x) \mu_2^2(k)}{4}. \quad (2.7)$$

Facilement trouver que  $\hat{h}_{opt}$  qui minimise asymptotique de MSE (2.7) est :

$$\frac{\partial}{\partial h} AMSE \left( \hat{F}_n(x) \right) = h^3 f'^2(x) \mu_2^2(k) - n^{-1} f(x) \varphi(k). \quad (2.8)$$

D'où (2.8), nous obtenons :

$$\begin{aligned} h^3 f'^2(x) \mu_2^2(k) - n^{-1} f(x) \varphi(k) &= 0 \\ \Leftrightarrow h^3 f'^2(x) \mu_2^2(k) &= n^{-1} f(x) \varphi(k) \\ \Leftrightarrow \hat{h}_{opt} &= \left( \frac{f(x) \varphi(x)}{n f'^2(x) \mu_2^2(k)} \right)^{\frac{1}{3}}. \end{aligned}$$

L'erreur quadratique moyenne asymptotique associée à  $\hat{h}_{opt}$  est donnée par :

$$AMSE_{\hat{h}_{opt}} \left( \hat{F}_n(x) \right) = n^{-1} \left[ F(x) (1 - F(x)) - \frac{3}{4} \left( \frac{f^4(x) \varphi^4(x)}{n f'^2(x) \mu_2^2(k)} \right)^{\frac{1}{3}} \right].$$

### 2.2.2 Erreur quadratique moyenne intégré asymptotique

L'approximation asymptotique de la MISE de  $\hat{F}_n$  est formé en intégrant (2.7) est donné par :

$$\begin{aligned} AMISE \left( \hat{F}_n(x) \right) &= \int AMSE \left( \hat{F}_n(x) \right) dx \\ AMISE \left( \hat{F}_n(x) \right) &= \int \left( n^{-1} F(x) (1 - F(x)) - n^{-1} h f(x) \varphi(x) + \frac{h^4 f'^2(x) \mu_2^2(k)}{4} \right) dx. \end{aligned} \quad (2.9)$$

Facilement trouver que  $h_{opt}^*$  qui minimise asymptotique de MISE (2.9) est :

$$\frac{\partial}{\partial h} AMISE \left( \hat{F}_n(x) \right) = h^3 \mu_2^2(k) \int f'^2(x) dx - n^{-1} \varphi(k) \quad (2.10)$$

D'où (2.10), nous obtenons :

$$\begin{aligned} h^3 \mu_2^2(k) \int f'^2(x) dx - n^{-1} \varphi(k) &= 0 \\ \Leftrightarrow h^3 \mu_2^2(k) \int f'^2(x) dx &= n^{-1} \varphi(k) \\ \Leftrightarrow h_{opt}^* &= \left( \frac{\varphi(k)}{n \mu_2^2(k) \int f'^2(x) dx} \right)^{\frac{1}{3}}. \end{aligned}$$

L'erreur quadratique moyenne intégrée asymptotique associée à  $h_{opt}^*$  est :

$$AMSE_{h_{opt}^*} \left( \hat{F}_n(x) \right) = n^{-1} \left[ F(x)(1 - F(x)) - \frac{3}{4} \left( \frac{\varphi^4(k)}{n \mu_2^2(k) \int f'^2(x) dx} \right)^{\frac{1}{3}} \right].$$

- 1) Alors que l'amélioration par rapport à la  $\hat{F}_n(x)$  disparaît aussi  $n \rightarrow \infty$ , il le fait le taux lent de  $n^{-\frac{1}{3}}$ , ce qui suggère que l'  $\hat{F}_n(x)$  beaucoup ont un gain d'échantillon fini significatif sur le  $F_n(x)$ .
- 2) L'amélioration de  $\hat{F}_n(x)$  est inversement proportionnelle à  $\rho(f') = \int f'^2(x) dx$ . Ainsi, nous attendons les gains minimales lorsque la densité est raide.
- 3) Le choix du noyau  $k$  n'affecte que le AMISE à travers  $\varphi$  ( des valeurs plus élevées réduisent le AMISE ).
- 4) L'estimateur  $\hat{F}_n(x)$  est asymptotiquement plus efficace que le  $F_n(x)$  voir (**Swanepoel** 1988 ).

## 2.3 Simulation

L'objectif de ce chapitre, est de vérifier par simulation l'innocence du choix de paramètre de lissage et du noyau pour l'estimateur à noyau de la distribution. Les simulations sont réalisées par le logiciel **R**. On a vu que l'estimateur à noyau de la densité dépend de la taille de l'échantillon  $n$  mais aussi du noyau  $k$ , nous allons donc étudier le "choix du noyau".

Pour cela, on choisit deux différents noyaux "**Epanechnikov** et **Quartique**" qui sont mentionnées dans le tableau 2.1, dans les taille  $n = 200$  et  $n = 500$ . Pour l'estimation de la distribution nous avons choisi de simuler les densités de probabilité de deux types de lois, la loi **Beta** qui a une fonction de densité ( $B1$ ) définie sur  $[0, 1]$  et la deuxième est la loi **exponentielle** dont la densité ( $B2$ ) est définie sur  $\mathbb{R}_+$ . Les deux fonctions de densité sont définies comme suit :

( $B1$ ). la loi Beta  $\mathcal{B}e(\alpha, \beta)$  :

$$f(x; \alpha, \beta) = \begin{cases} \int_0^1 \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mu^{\alpha-1}(1-\mu)^{\beta-1}d\mu} & \text{pour } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$$

( $B2$ ). la loi exponentielle  $\text{exp}(\lambda)$  :

$$f(x) = \lambda e^{-\lambda x} \quad \text{si } x \geq 0.$$

On suppose que l'on a observé un échantillon  $X_1, \dots, X_n$ .

En utilisant l'estimateur à noyau de la distribution :

$$\hat{F}_n(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right)$$

Dans les résultats graphiques de cette section :

- 1) Le graphe bleu représente la fonction de distribution  $F(x)$ .
- 2) La droite en rouge représente l'estimateur à noyau de la distribution  $\hat{F}_n(x)$ .

### 2.3.1 Exemples sur le noyau Quartique

kernel

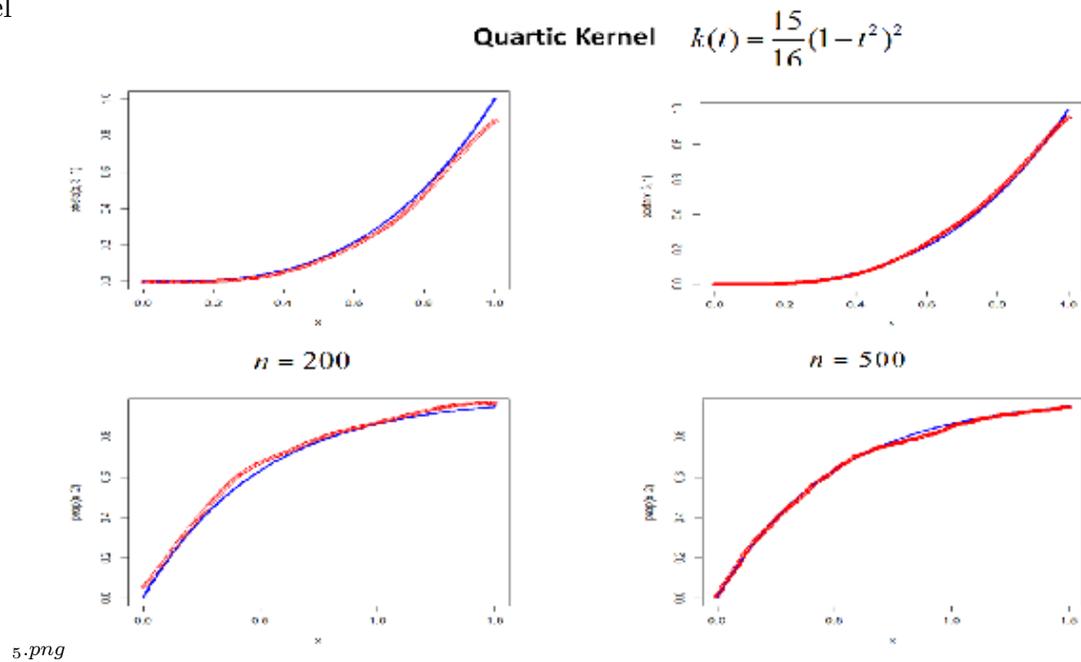


FIG. 2.3 – la fonction de distribution, noyau Quartique.

### 2.3.2 Exemples sur le noyau d'Epanechnikov

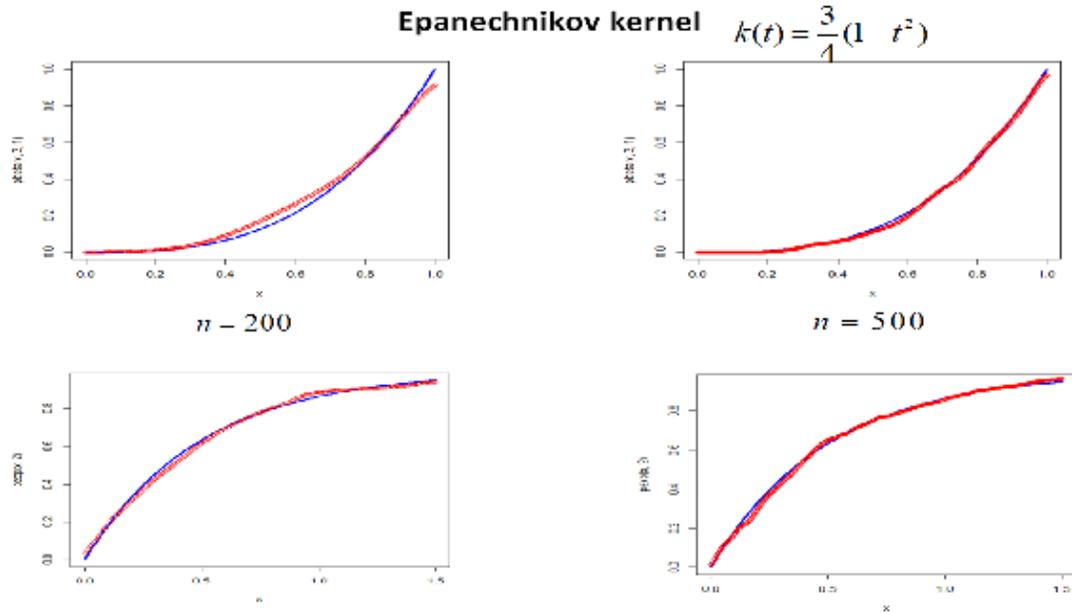


FIG. 2.4 – la fonction de distribution, noyau d'Epanechnikov.

Nous remarquons ici que l'estimateur lisse dès que  $n$  est grand pour chaque noyau utilisé, et sur le graphe ci-dessus que quand  $n$  est grand l'estimateur  $\hat{F}_n$  est proche de la distribution théorique.

Grâce à ses résultats, on peut conclure que l'influence du paramètre de lissage  $h$  est plus forte que l'influence du noyau  $k$ , par ailleurs, nous permettons de comparer les deux différents estimateurs on remarque que l'estimateur à noyau est toujours meilleur surtout quand  $t$  est proche de 1.

# Conclusion

L'estimation de la fonction de distribution un rôle centrale dans l'estimation fonctionnelle, nous avons donné les différentes méthodes d'estimation de la fonction de distribution à savoir : l'estimateur empirique, l'estimateur à noyau l'estimateur le plus simple, la fonction de distribution empirique a de bonnes propriétés de convergence mais possède certains inconvénients comme celui de ne pas prendre en compte une éventuelle information supplémentaire ou bien le fait d'être une fonction en escalier.

Enfin on conclut que l'estimateur à noyau est très utilisé en estimation non-paramétrique et les performances des estimateurs s'améliorent lorsque la taille de l'échantillon augmente.

# Bibliographie

- [1] Akakpo, N. (2018). Éléments de statistique Estimation non-paramétrique 1/2.
- [2] Dodge, Y., & Melfi, G. (2008). Premiers pas en simulation. Springer.
- [3] DUSART, P. (2018). Cours de statistique inferentielle, Licence 2-S4 SI-MASS, [https://www.unilim.fr/pages\\_perso/pierre.dusart/Probas/cours\\_stat\\_S4.pdf](https://www.unilim.fr/pages_perso/pierre.dusart/Probas/cours_stat_S4.pdf).
- [4] Gaudoin, O., & BÉGUIN, M. (2009). Principes et méthodes statistiques. Ensimag-2ème Année, INP Grenoble.
- [5] Haddou, M. (2007). Estimation non-paramétrique de la fonction de répartition et de la densité.
- [6] Hanebeck, A. (2020). Nonparametric Distribution Function Estimation : Nicht-parametrische Schätzung Von Verteilungsfunktionen (Doctoral dissertation, Karlsruher Institut für Technologie (KIT)).
- [7] JACQUES, J. Statistiques inférentielles.
- [8] JOURDAIN, B. (2018). Probabilités et statistique pour l'ingénieur.
- [9] Kouder, M. R. (2012). Estimation de moyenne des distributions à queues lourdes (Doctoral dissertation).
- [10] Lejeune, M. (2004). Statistique : La théorie et ses applications. Springer Science & Business Media.
- [11] MAAZ, A. (2016). Estimation de la fonction de densité.
- [12] MIHI, K. (2020). Estimation non paramétrique de la densité.

- [13] Mugdadi, A. R., & Munthali, E. (2003). Relative efficiency in kernel estimation of the distribution function. *Journal of Statistical Research*, 37(2), 203-218.
- [14] Rousson, V. (2013). *Statistique appliquée aux sciences de la vie*. Paris : Springer.
- [15] Ruch, J. J. (2013). *Statistique : estimation*. Préparation l'Agrégation Bordeaux, 11.
- [16] Servien, R. (2009). Estimation de la fonction de répartition : revue bibliographique. *Journal de la Société Française de Statistique*, 150(2), 84-104.

# Annexe A : Logiciel R

## 2.4 Qu'est-ce-que le langage R ?

- Le langage R est un langage de programmation et un environnement mathématique utilisés pour le traitement de données. Il permet de faire des analyses statistiques aussi bien simples que complexes comme des modèles linéaires ou non-linéaires, des tests d'hypothèse, de la modélisation de séries chronologiques, de la classification, etc. Il dispose également de nombreuses fonctions graphiques très utiles et de qualité professionnelle.

- R a été créé par Ross Ihaka et Robert Gentleman en 1993 à l'Université d'Auckland, Nouvelle Zélande, et est maintenant développé par la R Development Core Team. L'origine du nom du langage provient, d'une part, des initiales des prénoms des deux auteurs (Ross Ihaka et Robert Gentleman) et, d'autre part, d'un jeu de mots sur le nom du langage S auquel il est apparenté.

# Annexe B : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

- $\Omega$  : L'ensemble sur lequel porte notre étude statistique.
- $v.a.$  : Variable aléatoire.
- $E(x)$  : Espérance du v.a  $X$ .
- $Var(x)$  : Variance du v.a  $X$ .
- $f(x)$  : Fonction de densité.
- $F(x)$  : Fonction de distribution.
- $\xrightarrow{p}$  : Convergence en probabilité.
- $\xrightarrow{\mathcal{L}}$  : .Convergence en loi.
- $\xrightarrow{p.s}$  : Convergence presque sûre.
- $\xrightarrow{L^p}$  : Convergence en moyenne d'ordre  $p$
- $IC$  : Intervalle de confiance.
- $i.i.d$  : Indépendantes identiquement distribuées.
- $X_n$  : Suite de variables aléatoires.
- $F_n(x)$  : Fonction de distribution empirique.

- $k(t)$  : noyau.
- $\hat{f}_n(x)$  : estimateur de la densité  $f$  par noyau.
- $\hat{F}_n(x)$  : estimateur de la distribution  $F$  par noyau.
- $MSE$  : L'erreur quadratique moyenne.
- $MISE$  : L'erreur quadratique moyenne intégrée.
- $AMSE$  : L'erreur quadratique moyenne asymptotique.
- $AMISE$  : L'erreur quadratique moyenne intégrée asymptotique.

## Résumé

Dans ce travail, nous nous intéressons à l'estimation non paramétrique de la de distribution par la méthode du noyau. Nous rappelons les principales définitions et notions de base dans l'estimation paramétrique et non-paramétrique et nous déterminons la construction de l'estimateur à noyau et ses propriétés fondamentales : les théorèmes de convergence forte et faible, vitesse de convergence, et le choix du noyau .

Nous réalisons des simulations en utilisant le logiciel R qui nous permet d'observer l'influence de la fenêtre et la taille de l'échantillon.

**Mots clé**: Estimation non paramétrique, Fonction de distribution, Estimateur à noyau.

## Abstract

In this work, we are interested in the non-parametric estimation of the distribution by the kernel method. We recall the main definitions and basic notions in parametric and non-parametric estimation and we determine the construction of the kernel estimator and its fundamental properties: the theorems of strong and weak convergence, speed of convergence, and the choice of the kernel.

We realize simulations using R software which allows us to observe the influence of the bandwidth and the sample size.

**Key words**: Non-parametric estimation, Distribution function, Kernel estimateur.

## ملخص

في هذا العمل، نحن مهتمون بالتقدير غير البارامتري لدالة التوزيع بطريقة النواة. وقمنا بإعطاء التعاريف والمفاهيم الأساسية في التقدير البارامتري وغير البارامتري. وسنحدد بناء مقدر النواة هذا والخصائص الأساسية: نظريات التقارب القوي والضعيف، سرعة التقارب، واختيار النواة.

نقوم بإجراء عمليات المحاكاة باستخدام برنامج R الذي يسمح لنا بمراقبة تأثير النافذة وحجم العينة.

**الكلمات المفتاحية**: تقدير غير البارامتري ، دالة التوزيع، تقدير النواة.