

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

**MERAGHNI Brahim**

Titre :

**Estimation par Intervalle de Confiance**

Membres du Comité d'Examen :

Dr. Yahia Djabrane	UMKB	Président
Pr. Brahim Brahim	UMKB	Encadreur
Dr. Sayah Abdallah	UMKB	Examineur

Juin 2019

## DÉDICACE

A mes chers parents :

Ma mère, symbole de tendresse, qui s'est sacrifiée pour mon bonheur et ma réussite.

Mon père, école de mon enfance, qui a veillé à me protéger, m'aider et m'encourager.

A mes soeurs.

A mes neveux Bachir et Adam.

A toute ma grande famille.

Et à la mémoire des mes grands-pères et mon ami Ayoub.

*Brahim.*

## REMERCIEMENTS

Tout d'abord, je remercie "Allah Le Tout Puissant" de m'avoir aidé et donné la santé et volonté pour arriver à ce stade.

Mes vifs remerciements sont adressés à mon encadreur Pr. Brahim Brahimi pour ses précieux conseils, ses orientations pertinentes et sa patience tout au long de la réalisation de ce mémoire.

Je tiens à remercier Messieurs Djabrane Yahia et Abdallah Sayah qui m'ont fait l'honneur de faire partie du jury de soutenance.

Je remercie tous les enseignants qui ont contribué à ma formation, ainsi que tous les employés du département de Mathématiques.

Je remercie tout particulièrement mes parents, pour leur encouragement et soutien sur tous les aspects, ainsi que toute ma famille.

Je n'oublie pas l'ensemble de mes amis proches et aussi mes collègues d'études.

A ceux qui ont contribué, de près ou de loin, à la réalisation de ce modeste travail : un grand merci à vous tous.

# Table des matières

<b>Dédicace</b>	i
<b>Remerciements</b>	ii
<b>Table des matières</b>	iii
<b>Liste des figures</b>	v
<b>Liste des tableaux</b>	vi
<b>Introduction</b>	1
<b>1 Généralités</b>	<b>3</b>
1.1 Lois de probabilité usuelles . . . . .	3
1.1.1 Lois discrètes . . . . .	3
1.1.2 Lois continues . . . . .	6
1.2 Moyenne et variance empirique . . . . .	9
1.2.1 Moyenne empirique . . . . .	10
1.2.2 Variance empirique . . . . .	11
1.2.3 Cas d'échantillons Gaussiens . . . . .	13
1.3 Convergence de variables aléatoires . . . . .	14
<b>2 Intervalles de confiance</b>	<b>19</b>
2.1 Estimation ponctuelle . . . . .	20

2.1.1 Définitions et propriétés . . . . .	20
2.1.2 Estimation de la moyenne . . . . .	20
2.1.3 Estimation de la variance . . . . .	22
2.1.4 Estimation d'une proportion . . . . .	23
2.2 Principales méthodes d'estimation . . . . .	24
2.2.1 Méthode des moments . . . . .	24
2.2.2 Méthode du maximum de vraisemblance (MV) . . . . .	25
2.3 Construction d'un intervalle de confiance . . . . .	28
2.3.1 Intervalles de confiance pour les paramètres gaussiens . . . . .	29
2.4 Application sous R . . . . .	33
2.4.1 Illustration numérique . . . . .	33
2.4.2 Illustration graphique . . . . .	34
<b>Conclusion</b>	<b>37</b>
<b>Bibliographie</b>	<b>38</b>
<b>Annexe A : Codes R</b>	<b>39</b>
<b>Annexe B : Abréviations et Notations</b>	<b>44</b>

# Table des figures

1.1	Graphes de probabilité de la loi Binomiale de paramètres 10 et 0.5 (panneau de gauche) et de la loi de Poisson de paramètre 1 (panneau de droite) . . . . .	5
1.2	Densité de probabilité de la loi normale centrée avec différentes variances . . . . .	7
1.3	Densité de probabilité des lois du Khi-deux (panneau de gauche) et de Student (panneau de droite) avec différents degrés de liberté . . . . .	9
1.4	Approximation de la loi de Student, avec différents ddl, par la distribution standard de Gauss . . . . .	10
1.5	Illustration de la loi des grands nombres sur la moyenne empirique d'observations uniformes sur $[0, 1]$ (panneau gauche) et Gaussiennes standard (panneau droit). . . . .	17
1.6	Normalité asymptotique de la somme de 50 observations Poissonniennes (panneau gauche) et exponentielles (panneau droit) de paramètres égaux à 1. . . . .	18
2.1	Intervalles de confiance, de niveau 95%, pour la moyenne (panneau gauche) et la variance (panneau droit) d'une population normale standard. . . . .	35
2.2	Intervalles de confiance, de niveau 95%, pour les moyennes d'une population exponentielle (panneau gauche) et d'une v.a de Poisson (panneau droit) de paramètres égaux à l'unité. . . . .	36
2.3	Intervalles de confiance, de niveau 95%, pour la proportion de succès lors d'expériences de Bernoulli de paramètre 0.5. . . . .	36

# Liste des tableaux

1.1 Moyennes et variances de quelques v.a discrètes usuelles . . . . .	5
1.2 Moyennes et variances de quelques v.a continues usuelles . . . . .	8
1.3 Approximation de quelques lois usuelles par la distribution normale . . . . .	9
2.1 Intervalles de confiance, de différents niveaux, pour la moyenne (à gauche) et la variance (à droite) d'une population normale standard basés sur 100 observations.	33
2.2 Intervalles de confiance, de différents niveaux, pour la moyenne d'une population exponentielle (à gauche) et de Poisson (à droite), de paramètre 1, basés sur 100 observations.	33
2.3 Intervalles de confiance (exacts à gauche et approchés à droite), de différents niveaux, pour la proportion de succès lors de 100 expériences de Bernouilli de paramètre 0.5.	34

# Introduction

Les concepts et les méthodes statistiques ne sont pas seulement utiles, mais ils sont souvent indispensables pour comprendre le monde qui nous entoure. Ils fournissent des moyens de gagner de nouvelles idées dans la perspective de nombreux phénomènes qu'on aborde dans tous les domaines de la vie socio-économique.

La statistique mathématique ou inférentielle se compose de deux parties : l'estimation et les tests d'hypothèses. Le premier aspect permet, au vu d'un échantillon, d'évaluer des quantités théoriques inconnues relatives à une population toute entière. Le second est une procédure de prise de décision amenant, sur la base de mesures expérimentales, à l'acceptation ou au rejet d'une hypothèse faite à priori sur une caractéristique inconnue d'une variable aléatoire.

Dans ce travail, on s'intéresse à la théorie de l'estimation et plus particulièrement à l'estimation par intervalles de confiance. L'estimation ponctuelle permet de construire les estimateurs puis établir et étudier leurs propriétés. Pour cela, il existe une variété de méthodes comme celle des moments et/ou celle dite du maximum de vraisemblance, ...

L'estimation par intervalles de confiance est complémentaire de l'estimation ponctuelle qui a l'inconvénient de ne pas apporter d'information sur la précision des résultats obtenus. En effet, pour évaluer la confiance que l'on peut avoir en une estimation, il est nécessaire de déterminer un intervalle contenant, avec une probabilité donnée, fixée au préalable, la vraie valeur de la quantité inconnue.

En plus de cette introduction et d'une conclusion, ce mémoire est rédigé en deux chapitres :

- **Chapitre 1** : ce chapitre est consacré à un rappel des lois de probabilité usuelles et leurs caractéristiques ainsi qu'aux propriétés fondamentales de l'échantillonnage relatives à la

moyenne et variance empiriques. Des résultats importants, tels les lois des grands nombres et le théorème central limite, sont introduits à la fin du chapitre.

- **Chapitre 2** : on commence ce chapitre par un rappel des définitions et propriétés d'un estimateur, puis on introduit les deux méthodes les plus connues qui permettent d'obtenir un estimateur ponctuel dans le cas paramétrique. On prend comme exemples de paramètres à estimer la moyenne, la variance et la proportion. Dans la deuxième partie du chapitre, on décrit la construction d'un intervalle de confiance, basée sur la distribution de probabilité des estimateurs. On termine par une étude de simulation pour illustrer les résultats théoriques obtenus.

Enfin, les résultats numériques et les représentations graphiques sont obtenus à l'aide du logiciel de traitement statistique R [4]. Les codes de programmation sous R et les abréviations et notations utilisées sont rassemblés en deux annexes.

# Chapitre 1

## Généralités

Ce chapitre est consacré à un **rappel des notions de base** de la statistique mathématique. On passe en revue les lois de probabilités classiques aussi bien discrètes que continues. On donne les principaux résultats qui seront utiles pour la deuxième partie de ce travail.

Une variable aléatoire (v.a) réelle  $X$  est une fonction définie sur un espace de probabilité  $(\Omega, \mathcal{F}, P)$  à valeurs dans  $\mathbb{R}$  et mesurable par rapport aux tribus  $\mathcal{F}$  et  $\mathcal{B}_{\mathbb{R}}$  (tribu borélienne de  $\mathbb{R}$ ).

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

### 1.1 Lois de probabilité usuelles

#### 1.1.1 Lois discrètes

**Définition 1.1.1 (Loi uniforme)** *On dit qu'une v.a discrète  $X$  est **uniforme** sur un ensemble de points réels  $x_1, x_2, \dots, x_n$ ,  $n \geq 1$ , si elle admet pour loi de probabilité*

$$P(X = k) = \frac{1}{n} \mathbf{1}_{\{x_1, x_2, \dots, x_n\}}(k).$$

Dans le cas particulier où  $x_i = i, i = 1, \dots, n$ , on écrit  $X \sim \mathcal{U}(n)$  et on a

$$E(X) = \frac{n+1}{2} \text{ et } Var(X) = \frac{n^2-1}{12}.$$

**Définition 1.1.2 (Loi de Bernoulli)** On dit qu'une v.a  $X$  suit la loi de **Bernoulli** de paramètre  $0 < p < 1$ , et on écrit  $X \sim \mathcal{B}(p)$ , si elle ne prend que deux valeurs possibles 0 ou 1 avec des probabilités respectives  $1-p$  et  $p$ , c-à-d

$$P(X = k) = \begin{cases} 1-p & \text{si } k = 0, \\ p & \text{si } k = 1. \end{cases}$$

Le paramètre  $p$  est dit probabilité de succès. On a

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

**Définition 1.1.3 (Loi binomiale)** Une v.a  $X$  est dite **binomiale** de paramètres  $n \in \mathbb{N}$  et  $0 < p < 1$  (on écrit  $X \sim \mathcal{B}(n, p)$ ) si sa loi de probabilité s'écrit de la forme suivante :

$$P(X = k) = C_n^k p^k (1-p)^{n-k} \mathbf{1}_{\{0, \dots, n\}}(k).$$

Elle représente le nombre de succès lors de  $n$  épreuves indépendantes de Bernoulli de probabilité de succès  $p$ . En d'autres termes, on a  $X = \sum_{i=1}^n X_i$  avec  $X_i \sim \mathcal{B}(p)$  **indépendantes** et de **même paramètre**  $p$ . On a

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

**Définition 1.1.4 (Loi de Poisson)** Une v.a.  $X$  est dite de **Poisson** de paramètre  $\lambda > 0$  (on écrit  $X \sim \mathcal{P}(\lambda)$ ) si sa loi de probabilité est

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \mathbf{1}_{\mathbb{N}}(k).$$

Elle modélise le nombre d'arrivées pendant une certaine période de temps. On a

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

Le paramètre  $\lambda$  représente le nombre moyen des arrivées.

On résume les valeurs des espérances et variances des lois discrètes ci-dessus dans le tableau [1.1](#).

La figure [1.1](#) montre le comportement de la loi binomiale et celle de Poisson pour des valeurs particulières de leurs paramètres.

Loi	$E(X)$	$Var(X)$
$\mathcal{U}(n)$	$(n + 1) / 2$	$(n^2 - 1) / 12$
$\mathcal{B}(p)$	$p$	$p(1 - p)$
$\mathcal{B}(n, p)$	$p$	$np(1 - p)$
$\mathcal{P}(\lambda)$	$\lambda$	$\lambda$

TAB. 1.1 – Moyennes et variances de quelques v.a discrètes usuelles

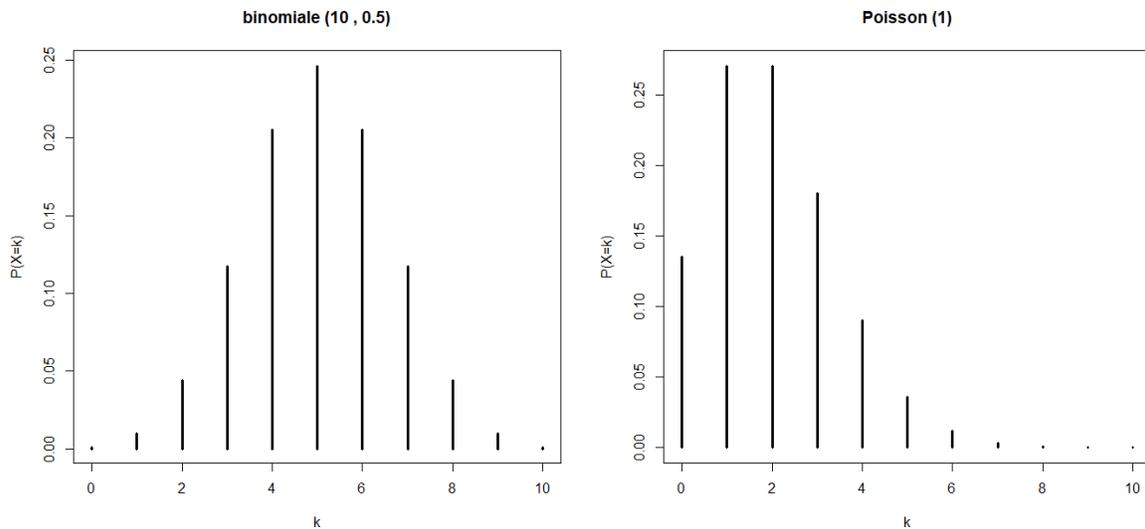


FIG. 1.1 – Graphes de probabilité de la loi Binomiale de paramètres 10 et 0.5 (panneau de gauche) et de la loi de Poisson de paramètre 1 (panneau de droite)

### 1.1.2 Lois continues

**Définition 1.1.5 (Loi uniforme)** On dit que une v.a continue  $X$  suit une loi **uniforme** sur un intervalle  $[a, b]$ , où  $a$  et  $b$  sont deux réels tels que  $a < b$ , si elle admet pour densité de probabilité

$$f(x) = \frac{1}{b-a} \mathbf{1}_{[a,b]}(x).$$

On écrit  $X \sim \mathcal{U}([a, b])$  et on a

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

Le cas particulier  $a = 0$  et  $b = 1$  donne naissance à la loi uniforme standard, pour laquelle on a  $E(X) = 1/2$  et  $Var(X) = 1/12$ .

**Définition 1.1.6 (Loi exponentielle)** On dit qu'une v.a  $X$  suit la loi **exponentielle** de paramètre  $\theta > 0$ , et on écrit  $X \sim \mathcal{E}(\theta)$ , si sa densité de probabilité est définie par

$$f(x) = \theta e^{-\theta x} \mathbf{1}_{[0, \infty[}(x).$$

Elle modélise le temps entre deux arrivées de Poisson appelé aussi temps d'attente. Par intégration, on trouve sans difficulté que

$$E(X) = \frac{1}{\theta} \text{ et } Var(X) = \frac{1}{\theta^2}.$$

Le temps moyen d'attente est égal à l'inverse du paramètre  $\theta$ .

**Définition 1.1.7 (Loi normale ou gaussienne)** Une v.a  $X$  suit une loi **normale** ou **gaussienne** (loi de Laplace-Gauss) de paramètres  $\mu \in \mathbb{R}$  et  $\sigma^2 > 0$  si elle admet pour densité de probabilité la fonction  $f$  définie, pour tout nombre réel  $x$ , par

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

Cette distribution, notée  $\mathcal{N}(\mu, \sigma^2)$ , représente un bon modèle pour beaucoup de phénomènes naturels. De plus, le TCL (théorème [1.3.2](#)), la rend la plus importante en statistique et donc la plus populaire. On a

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

Lorsque l'espérance est nulle,  $X$  est dite v.a normale centrée. Sa densité est illustrée par la figure [1.2](#). Si de plus la variance vaut 1, on parle alors de ce que l'on appelle loi normale **standard** ou **centrée réduite**. Dans ce cas, la fonction de densité de probabilité est paire et a la forme simple suivante :

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

Ainsi, sa représentation graphique, appelée cloche de Gauss, est symétrique par rapport à l'axe des ordonnées.

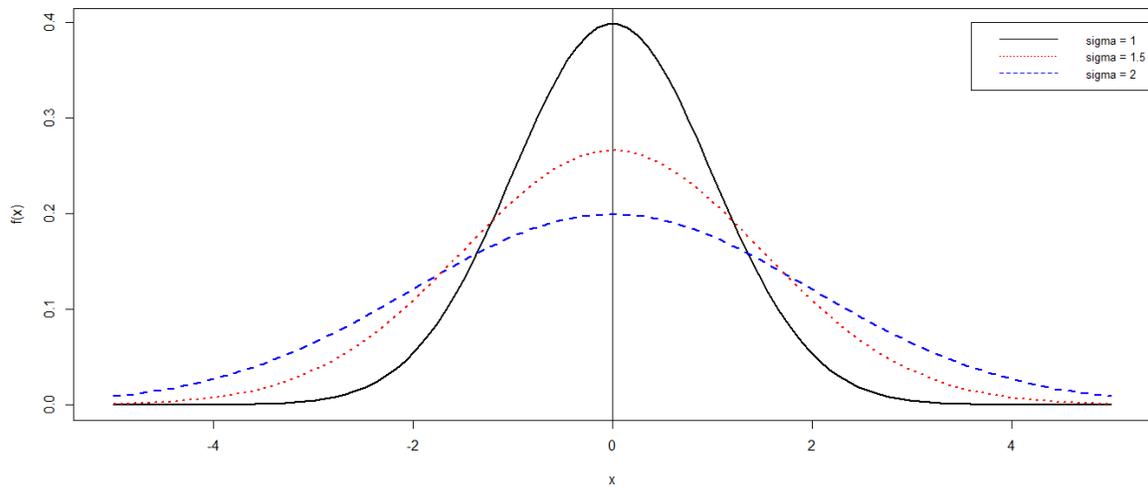


FIG. 1.2 – Densité de probabilité de la loi normale centrée avec différentes variances

Le passage du cas général au cas particulier se fait par un simple changement de variable. En effet, on a

$$X \sim \mathcal{N}(\mu, \sigma^2) \Leftrightarrow Z := \frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1).$$

**Définition 1.1.8 (Loi du Khi-Deux)** On dit qu'une v.a  $X$  suit une loi du **khi-deux** ou loi de **Pearson** et on écrit  $X \sim \chi^2_n$ , si elle admet pour densité

$$f(x) = \frac{2^{-n/2}}{\Gamma(n/2)} e^{-x/2} x^{n/2-1} \mathbf{1}_{[0, \infty[}(x),$$

où  $\Gamma(t) := \int_0^\infty e^{-u} u^{t-1} du$ ,  $t > 0$ , désigne la fonction Gamma.

**Remarque 1.1.1**

1. C'est la loi de la somme des carrés de  $n$  v.a normales centrées réduites indépendantes.
2. Son espérance et variance valent respectivement  $n$  et  $2n$ .
3. Si  $Z \sim \mathcal{N}(0, 1)$ , alors  $Z^2 \sim \chi^2_1$ .

**Définition 1.1.9 (Loi de Student)** Une v.a  $X$  est dite variable de **Student** de degré de liberté  $n \geq 1$  (on note  $X \sim t_n$ ) si elle s'écrit sous la forme suivante :

$$X = \frac{Z}{\sqrt{Y/n}},$$

où  $Z \sim \mathcal{N}(0, 1)$  et  $Y \sim \chi^2_n$  sont indépendantes.

D'après la proposition 5.9 de [7], on a  $E(X) = 0$ , si  $n \geq 2$  et  $Var(X) = n/(n-2)$  si  $n \geq 3$ . Il est à noter que pour un ddl égal à l'unité, l'espérance et la variance sont toutes les deux infinies. Une illustration graphique des densités des v.a de Pearson et de Student, pour des valeurs distinctes des ddl, est donnée par la figure 1.3. Un résumé des espérances et variances des lois continues ci-dessus est présenté dans le tableau 1.2.

Loi	$\mathbf{E}(X)$	$\mathbf{Var}(X)$
$\mathcal{U}([a, b])$	$(a + b) / 2$	$(b - a)^2 / 12$
$\mathcal{E}(\theta)$	$1 / \theta$	$1 / \theta^2$
$\mathcal{N}(\mu, \sigma^2)$	$\mu$	$\sigma^2$
$\chi^2_n$	$n$	$2n$

TAB. 1.2 – Moyennes et variances de quelques v.a continues usuelles

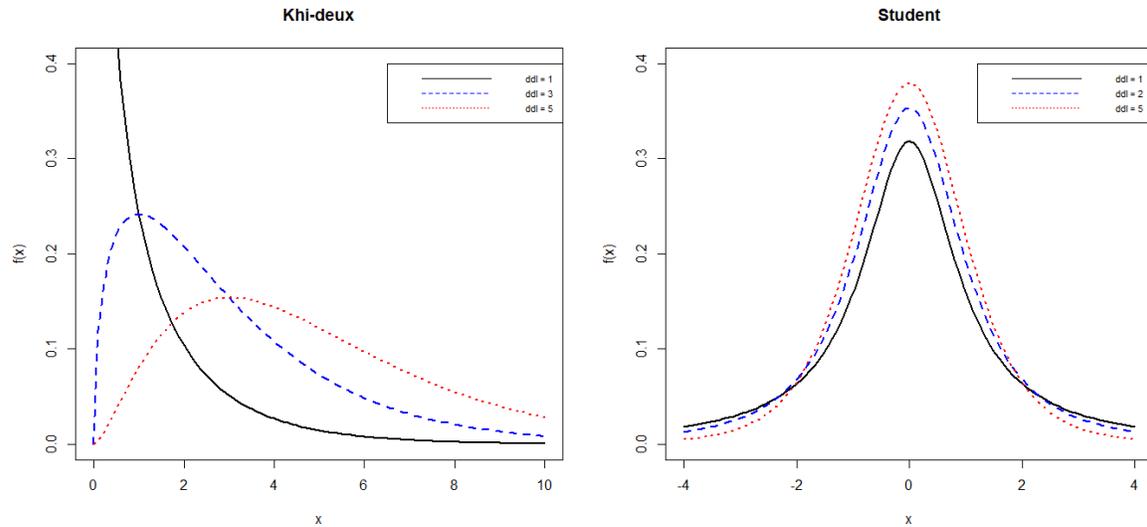


FIG. 1.3 – Densité de probabilité des lois du Khi-deux (panneau de gauche) et de Student (panneau de droite) avec différents degrés de liberté

La loi de Gauss peut être, sous certaines conditions, une bonne approximation pour des lois de probabilités comme le montre le tableau 1.3. A titre d'exemple, la figure 1.4 illustre l'assimilation d'une v.a de Student, de ddl  $n > 30$ , à une v.a normale centrée réduite.

Loi exacte	Loi approchée	Condition
$\mathcal{B}(n, p)$	$\mathcal{N}(np, np(1-p))$	$np > 30, np > 10, n(1-p) > 10$
$\mathcal{P}(\lambda)$	$\mathcal{N}(\lambda, \lambda)$	$\lambda > 10$
$t_n$	$\mathcal{N}(0, 1)$	$n > 30$
$\sqrt{\chi^2_n}$	$\mathcal{N}(\sqrt{(2n-1)/2}, 1/2)$	$n > 30$

TAB. 1.3 – Approximation de quelques lois usuelles par la distribution normale

Pour plus de détails sur ces lois de probabilité et sur d'autres, on peut consulter les ouvrages [8, Chapitre 2] et [7, Chapitre 4] ainsi que les notes de cours [1].

## 1.2 Moyenne et variance empirique

**Définition 1.2.1** *Un  $n$ -échantillon aléatoire issu d'une v.a  $X$  est une suite  $(X_1, \dots, X_n)$  de  $n \geq 1$  v.a indépendantes et identiquement distribuées (iid) ayant la même loi que  $X$ . Le nombre  $n$  est appelé taille de l'échantillon.*

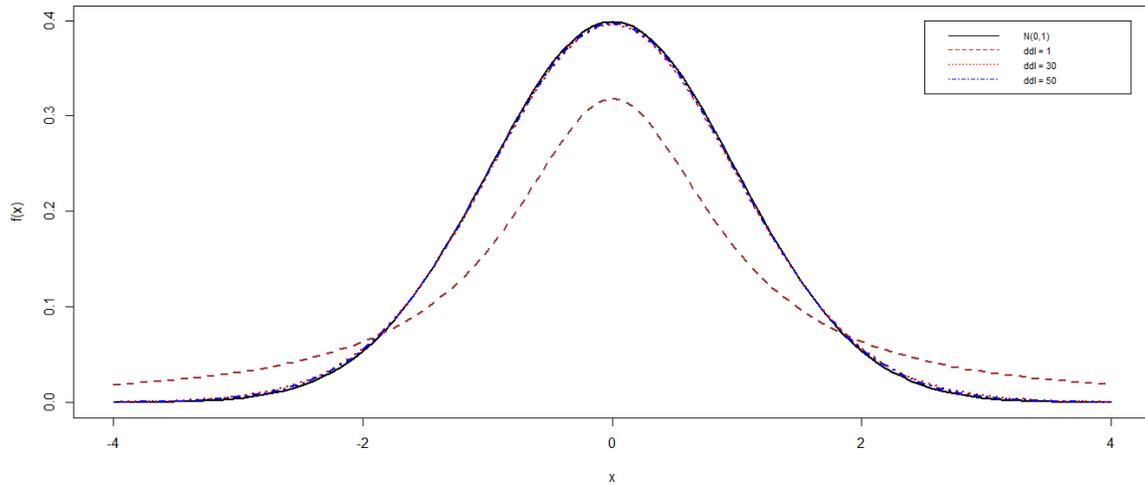


FIG. 1.4 – Approximation de la loi de Student, avec différents ddl, par la distribution standard de Gauss

**Définition 1.2.2** On appelle *statistique* sur un  $n$ -échantillon  $(X_1, \dots, X_n)$  toute fonction mesurable des  $X_i$ ,  $i = 1, \dots, n$ .

### 1.2.1 Moyenne empirique

**Définition 1.2.3** On appelle *moyenne empirique* (échantillonnale, expérimentale) la statistique notée  $\bar{X}_n$  définie par

$$\bar{X}_n := \frac{1}{n} \sum_{i=1}^n X_i. \quad (1.1)$$

#### Proposition 1.2.1

1. Soit  $X$  une v.a de moyenne  $\mu$  et de variance  $\sigma^2$ . Alors, on a

$$E(\bar{X}_n) = \mu \text{ et } \text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}.$$

2. Les moments centrés de  $\bar{X}_n$  d'ordre 3 et 4 sont respectivement

$$\mu_3^c(\bar{X}_n) = \frac{\mu_3^c}{n^2} \text{ et } \mu_4^c(\bar{X}_n) = \frac{\mu_4^c + 3\sigma^4(n-1)}{n^3},$$

où  $\mu_k^c := E(X - \mu)^k$ ,  $k \geq 1$ , désigne le  $k^{\text{ème}}$  moment centré de  $X$ . On en déduit que

$$\gamma_1(\bar{X}_n) = \frac{\gamma_1}{\sqrt{n}} \text{ et } \gamma_2(\bar{X}_n) = 3 + \frac{\gamma_2 - 3}{n},$$

où  $\gamma_1 := \mu_3^c/\sigma^3$  et  $\gamma_2 := \mu_4^c/\sigma^4$  sont respectivement les coefficients d'asymétrie et d'aplatissement de  $X$ .

**Preuve.**

1. Pour  $E(\bar{X}_n)$ , on utilise la linéarité de l'espérance et pour  $Var(\bar{X}_n)$ , on exploite l'indépendance des  $X_i$ .

2. On commence par montrer les résultats lorsque  $X$  est une v.a centrée ( $\mu = 0$ ), puis on généralise. Si  $X$  est centrée, alors

$$\mu_k(\bar{X}_n) = E(\bar{X}_n)^k, \quad k = 3, 4.$$

On applique le développement de Newton pour arriver aux formules voulues.

Enfin, les coefficients d'asymétrie et d'aplatissement de  $\bar{X}_n$  sont obtenus par application directe des définitions. ■

**Remarque 1.2.1** Lorsque  $n$  est grand, la moyenne empirique est symétrique et d'aplatissement normal (égal à 3). Ce qui fait penser à sa normalité asymptotique qui sera confirmée par le TCL (théorème [1.3.2](#)).

## 1.2.2 Variance empirique

**Définition 1.2.4** On appelle *variance empirique*, la statistique

$$S_n^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2. \tag{1.2}$$

**Proposition 1.2.2** *Si  $X$  est une v.a de variance  $\sigma^2$ , alors*

$$E(S_n^2) = \frac{n-1}{n}\sigma^2.$$

**Preuve.** On suppose que  $E(X) = \mu$ , alors on a

$$\begin{aligned} E(S_n^2) &= \frac{1}{n} \sum_{i=1}^n E(X_i - \bar{X}_n)^2 = \frac{1}{n} \sum_{i=1}^n E(X_i - \mu + \mu - \bar{X}_n)^2 \\ &= \frac{1}{n} \sum_{i=1}^n E(X_i - \mu)^2 - E(\bar{X}_n - \mu)^2 \\ &= \frac{1}{n} \sum_{i=1}^n \text{Var}(X_i) - \text{Var}(\bar{X}_n) \\ &= \sigma^2 - \frac{\sigma^2}{n}. \end{aligned}$$

D'où le résultat. ■

**Remarque 1.2.2** *Dans [8, page 280], il est dit que lorsque  $n$  tend vers l'infini,  $\text{Var}(S_n^2) \rightarrow (\mu_4 - \sigma^4)/n$  et*

$$\sqrt{n} \frac{S_n^2 - \sigma^2}{\sqrt{\mu_4 - \sigma^4}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1).$$

**Proposition 1.2.3 (Covariance entre  $\bar{X}_n$  et  $S_n^2$ )** *On a*

$$\text{cov}(\bar{X}_n, S_n^2) = \frac{n-1}{n^2} \mu_3.$$

**Preuve.** Par définition de la covariance, on a

$$\text{cov}(\bar{X}_n, S_n^2) = E[(\bar{X}_n - E(\bar{X}_n))(S_n^2 - E(S_n^2))],$$

qui devient après développement

$$\text{cov}(\bar{X}_n, S_n^2) = E(\bar{X}_n S_n^2) - E(\bar{X}_n) E(S_n^2).$$

Puisque la covariance est insensible à des changements par translation sur un des termes, alors

on peut supposer (sans nuire à la généralité) que  $\mu = 0$ . Dans ce cas, on a

$$\begin{aligned} \text{cov}(\bar{X}_n, S_n^2) &= E(\bar{X}_n S_n^2) \\ E(\bar{X}_n S_n^2) &= E\left[\left(\frac{1}{n} \sum_{i=1}^n X_i\right) \left(\frac{1}{n} \sum_{j=1}^n X_j^2 - \bar{X}_n^2\right)\right] \\ &= \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n E(X_i X_j^2) - E(\bar{X}_n)^3. \end{aligned}$$

En vertu l'indépendance des  $X_i$ , on a  $E(X_i X_j^2) = E(X_i) E(X_j^2) = \mu E(X_j^2) = 0$ , pour  $i \neq j$ .

Alors,

$$\begin{aligned} \text{cov}(\bar{X}_n, S_n^2) &= \frac{1}{n^2} \sum_{i=1}^n E(X_i^3) - \frac{1}{n^3} E\left(\sum_{i=1}^n X_i\right)^3 \\ &= \frac{\mu_3}{n} - \frac{\mu_3}{n^2}. \end{aligned}$$

La réduction au même dénominateur permet d'obtenir le résultat espéré. ■

### Remarque 1.2.3

1. Dans la cas où la distribution de la population  $X$  est symétrique ( $\mu_3 = 0$ ), les deux v.a  $\bar{X}_n$  et  $S_n^2$  sont non corrélées. Mais elles ne sont pas nécessairement indépendantes.
2. Pour un échantillon de grande taille d'une population quelconque,  $\bar{X}_n$  et  $S_n^2$  sont asymptotiquement non corrélées

### 1.2.3 Cas d'échantillons Gaussiens

Lorsque la population mère est Gaussienne, on a les résultats suivants.

**Proposition 1.2.4** Soit  $X \sim \mathcal{N}(\mu, \sigma^2)$ . Alors on a

1. La moyenne empirique est Gaussienne.

$$\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right). \tag{1.3}$$

2. La variance empirique vérifie

$$\frac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1}. \quad (1.4)$$

3.  $\bar{X}_n$  et  $S_n^2$  sont indépendantes.

**Preuve.**

1. Comme  $\bar{X}_n$  est une combinaison linéaire de variables normales, alors elle est Gaussienne.

2. On a

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2 &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n + \bar{X}_n - \mu)^2 \\ &= \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X}_n)^2 + (\bar{X}_n - \mu)^2, \end{aligned}$$

d'où

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

En multipliant par  $n/\sigma^2$ , on obtient

$$\frac{nS_n^2}{\sigma^2} = \sum_{i=1}^n Z_i^2 - Z^2,$$

où  $Z_i := (X_i - \mu)/\sigma$  et  $Z := \sqrt{n}(\bar{X}_n - \mu)/\sigma$  sont normales centrées réduites. D'après le premier point de la remarque [1.1.1](#), on conclut que  $nS_n^2/\sigma^2 \sim \chi^2_{n-1}$ .

3. En plus de la normalité, la non corrélation entre  $\bar{X}_n$  et  $S_n^2$ , due à la symétrie de  $X$  (voir la remarque [1.2.3](#)), implique l'indépendance. ■

## 1.3 Convergence de variables aléatoires

Vu l'utilité de la notion de convergence dans l'estimation statistique, on rappelle, dans cette section, les définitions et résultats essentiels relatifs aux différents modes de convergence. Soit  $(X_n)_{n \geq 1}$  une suite de v.a.

**Définition 1.3.1 (Convergence en probabilité)** On dit que  $(X_n)_{n \geq 1}$  converge en probabilité

vers une v.a.  $X$ , et on écrit  $X_n \xrightarrow{P} X$ , si

$$\forall \epsilon > 0, \lim_{n \rightarrow \infty} P(|X_n - X| > \epsilon) = 0. \quad (1.5)$$

**Définition 1.3.2 (Convergence presque sûre)** On dit que  $(X_n)_{n \geq 1}$  converge presque sûrement vers une v.a.  $X$ , et on écrit  $X_n \xrightarrow{ps} X$ , si

$$P\left(\lim_{n \rightarrow \infty} X_n \neq X\right) = 0. \quad (1.6)$$

**Définition 1.3.3 (Convergence en loi)** On dit que  $(X_n)_{n \geq 1}$  converge en loi vers une v.a.  $X$ , et on écrit  $X_n \xrightarrow{\mathcal{L}} X$ , si

$$\lim_{n \rightarrow \infty} F_n(x) = F(x). \quad (1.7)$$

où  $F_n$  et  $F$  désignent les fonctions de répartition de  $X_n$  et  $X$  respectivement.

**Définition 1.3.4 (Convergence en moyenne quadratique)** On dit que  $(X_n)_{n \geq 1}$  converge en moyenne quadratique vers une v.a.  $X$ , et on écrit  $X_n \xrightarrow{mq} X$ , si

$$\lim_{n \rightarrow \infty} E(X_n - X)^2 = 0. \quad (1.8)$$

**Remarque 1.3.1**

1. Les relations (1.5), (1.6), (1.8), et (1.7) restent valables si on remplace la v.a.  $X$  par une constante réelle  $\ell$ .
2. La convergence en moyenne quadratique se généralise à la convergence en moyenne d'ordre  $k \geq 2$ .
3. Les implications suivantes permettent le passage entre certains types de convergence.

$$\begin{aligned} X_n \xrightarrow{P} X &\implies X_n \xrightarrow{\mathcal{L}} X \\ X_n \xrightarrow{ps} X &\implies X_n \xrightarrow{P} X \\ X_n \xrightarrow{mq} X &\implies X_n \xrightarrow{P} X \end{aligned}$$

Avant d'énoncer les deux théorèmes limites fondamentaux en statistique (théorèmes [1.3.1](#) et [1.3.2](#)), on commence par introduire, dans la proposition [1.3.1](#), deux inégalités remarquables.

**Proposition 1.3.1**

- **Inégalité de Markov** : Pour tout réel  $a > 0$  et toute fonction mesurable positive sur  $\mathbb{R}$  on a

$$P(g(X) \geq a) \leq \frac{1}{a} E(g(X)). \quad (1.9)$$

- **Inégalité de Chebychev** : Soit  $Y$  une v.a de variance finie  $\sigma^2$ . Alors

$$\forall \epsilon > 0, P(|Y - E(Y)| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}. \quad (1.10)$$

*Preuve.* On a

$$E(g(X)) = \int_{\mathbb{R}} g(x)f(x)dx = \int_{g(x) < a} g(x)f(x)dx + \int_{g(x) \geq a} g(x)f(x)dx.$$

Or la fonction  $g$  est positive, d'où

$$E(g(X)) \geq \int_{g(x) \geq a} g(x)f(x)dx \geq a \int_{g(x) \geq a} f(x)dx = aP(g(X) \geq a),$$

ce qui donne [\(1.9\)](#). Pour montrer l'inégalité [\(1.10\)](#), il suffit de prendre  $g(X) = (Y - E(Y))^2$ ,  $a = \epsilon^2$  dans [\(1.9\)](#). ■

**Théorème 1.3.1 (Lois des grands nombres)** Soient  $X_1, \dots, X_n$  des v.a iid d'espérance  $\mu$  et de variance  $\sigma^2$  (finies). Alors, on a :

- **Loi faible** :

$$\bar{X}_n \xrightarrow{P} \mu.$$

- **Loi forte** :

$$\bar{X}_n \xrightarrow{ps} \mu.$$

**Preuve.** Pour la loi faible, on applique l'inégalité de Chebychev (1.10) à la v.a  $Y = \bar{X}_n$ , avec  $E(\bar{X}_n) = \mu$  et  $Var(\bar{X}_n) = \sigma^2/n$ . Pour  $\epsilon > 0$ , on a

$$P(|\bar{X}_n - \mu| \geq \epsilon) \leq \frac{\sigma^2}{n\epsilon^2} \rightarrow 0, \text{ lorsque } n \rightarrow \infty,$$

c'est à dire  $\bar{X}_n \xrightarrow{P} \mu$ .

Pour la preuve de loi forte, qui est plus technique, on réfère à [5], page 77]. ■

Un exposé détaillé sur les LGN, leurs démonstrations et leurs applications se trouve dans le chapitre 1 de [9]. La figure 1.5 permet de voir la convergence de la moyenne empirique vers la moyenne théorique, pour deux exemples de lois de probabilités.

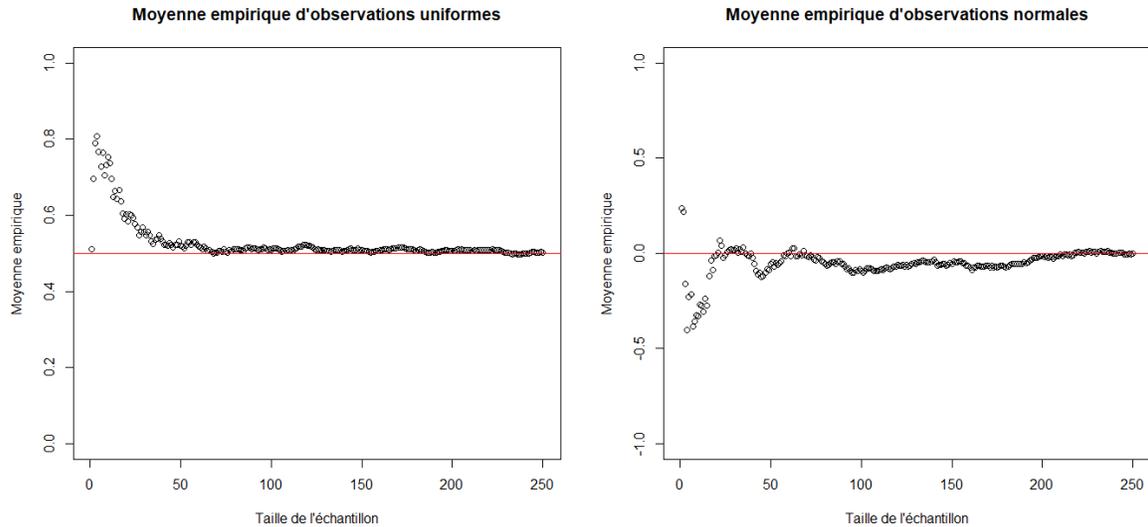


FIG. 1.5 – Illustration de la loi des grands nombres sur la moyenne empirique d'observations uniformes sur  $[0, 1]$  (panneau gauche) et Gaussiennes standard (panneau droit).

**Théorème 1.3.2 (Théorème centrale limite)** Soit  $(X_n)_{n \geq 1}$  une suite de v.a iid, d'espérance  $\mu$  et de variance  $\sigma^2$  finies. Alors

$$\frac{U_n - n\mu}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1),$$

où  $U_n := \sum_{i=1}^n X_i$  désigne la somme des  $n$  premières v.a.

**Preuve.** La démonstration est basée sur la fonction caractéristique. Pour les détails, voir, par exemple, [1, page 14] et [5, page 88]. ■

Une illustration graphique du TCL est donnée dans la figure 1.6. Un résultat immédiat du théorème 1.3.2, donné dans le corollaire 1.3.1, assure la normalité asymptotique de la moyenne empirique.

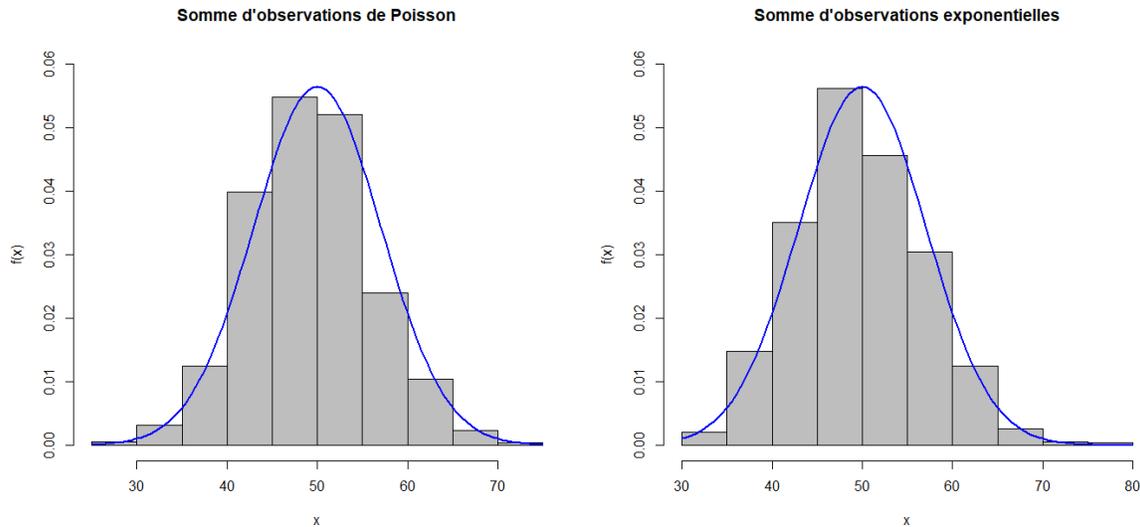


FIG. 1.6 – Normalité asymptotique de la somme de 50 observations Poissonniennes (panneau gauche) et exponentielles (panneau droit) de paramètres égaux à 1.

**Corollaire 1.3.1** *La moyenne empirique  $\bar{X}_n$  définie par (1.1) est asymptotiquement normale.*

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

Les LGN et le TCL constituent des outils de base pour l'étude des propriétés asymptotiques des estimateurs, à savoir la consistance et la normalité asymptotique. On note que ces théorèmes existent sous d'autres versions avec des conditions moins restrictives sur les v.a que celles données dans les théorèmes 1.5 et 1.3.2 (voir [1, page 11] et [2, page 366]).

# Chapitre 2

## Intervalles de confiance

**E**n statistique, l'**estimation** est un processus qui permet d'évaluer une quantité théorique inconnue (espérance, variance, densité de probabilité, fonction de répartition,...) relative à une v.a. Elle sert à approximer certaines caractéristiques d'une population à partir de données obtenues sur un échantillon (à l'issue d'un sondage par exemple). Elle constitue une des deux composantes (l'autre étant les tests d'hypothèses) de la statistique inférentielle. On distingue trois types d'estimation selon que la distribution de la population est connue, partiellement connue ou inconnue. Il s'agit respectivement de l'estimation paramétrique, semi-paramétrique ou non-paramétrique.

Dans ce chapitre, on s'intéresse à l'**estimation par intervalles de confiance** qui sont construits autour de valeurs obtenues par l'une des diverses méthodes d'**estimation ponctuelle**. Parmi ces dernières, on peut citer la très populaire méthode du maximum de vraisemblance, la méthode des moments et la méthode des moindres carrés surtout appliquée dans les problèmes de régression.

## 2.1 Estimation ponctuelle

### 2.1.1 Définitions et propriétés

Soit  $X$  une v.a dont la distribution dépend d'un (ou plusieurs) paramètre  $\theta$ . La donnée d'un modèle statistique c'est la donnée d'une famille de probabilités  $\{P_\theta, \theta \in \Theta\}$ , où  $\Theta$  représente l'espace des valeurs du paramètre inconnu. On dit alors que la loi de  $X$  appartient au modèle  $\{P_\theta, \theta \in \Theta\}$ . On souhaite estimer ce paramètre à partir de l'observation d'un échantillon aléatoire  $(X_1, \dots, X_n)$ , de taille  $n \geq 1$ , extrait de la population  $X$ .

**Définition 2.1.1** *Un **estimateur** du paramètre  $\theta$  est une statistique (voir la définition [1.2.2](#)), généralement notée par  $\hat{\theta}_n$ , contenant le plus d'information possible sur  $\theta$ .*

La qualité d'un estimateur s'exprime par sa convergence ou consistance, son biais, son efficacité et/ou sa robustesse.

**Définition 2.1.2 (Consistance)** *Un estimateur  $\hat{\theta}_n$  est dit **convergent** ou **consistant** s'il est "proche" de  $\theta$  au sens de la convergence en probabilité : pour tout  $\epsilon > 0$ ,*

$$P\left(\left|\hat{\theta}_n - \theta\right| > \epsilon\right) \xrightarrow{n \rightarrow +\infty} 0.$$

**Définition 2.1.3 (Biais)** *On appelle **biais** d'un estimateur  $\hat{\theta}_n$  la quantité  $E(\hat{\theta}_n) - \theta$ . L'estimateur  $\hat{\theta}_n$  est dit sans biais si  $E(\hat{\theta}_n) = \theta$ , sinon il est dit biaisé.*

Dans ce qui suit, on va s'intéresser à l'estimation de trois paramètres très connus, à savoir la moyenne, la proportion et la variance. On considère un  $n$ -échantillon  $(X_1, \dots, X_n)$  issu d'une loi de moyenne  $\mu$  et de variance  $\sigma^2$ , toutes deux inconnues.

### 2.1.2 Estimation de la moyenne

L'estimateur naturel (non-paramétrique) de l'espérance  $\mu$  est la **moyenne empirique**  $\bar{X}_n$  définie par [\(1.1\)](#). Dans le cas des lois de probabilités usuelles, il est également obtenu par les méthodes du MV et des moments.

Plus généralement, le  $k^{\text{ème}}$  moment théorique

$$\mu_k := E(X^k), \quad k \geq 1, \quad (2.1)$$

est naturellement estimé par le moment empirique d'ordre  $k$

$$m_k := \frac{1}{n} \sum_{i=1}^n X_i^k, \quad k \geq 1. \quad (2.2)$$

**Proposition 2.1.1 (Propriétés de  $\bar{X}_n$ )**

1.  $\bar{X}_n$  est un estimateur convergent (faiblement et fortement) de  $\mu$ .
2.  $\bar{X}_n$  est un estimateur sans biais.

**Preuve.**

1. C'est les LGN [1.3.1](#)
2. D'après , on a  $E(\bar{X}_n) = \mu$ , d'où un biais nul. ■

**Proposition 2.1.2 (Loi de  $\bar{X}_n$ )**

1. Si  $X \sim \mathcal{N}(\mu, \sigma^2)$ , alors
  - si la variance  $\sigma^2$  est connue, on a

$$\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1), \quad (2.3)$$

- si la variance  $\sigma^2$  est inconnue, on a

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n-1}} \sim t_{n-1}. \quad (2.4)$$

2. Si  $X$  est de distribution quelconque ou même inconnue, la loi de  $\bar{X}_n$  est approchée par une loi normale d'espérance  $\mu$  et de variance  $\sigma^2/n$ , lorsque  $n$  est grand (en pratique  $n > 30$ ).

**Preuve.**

1. La loi (2.3) n'est autre que le résultat (1.3).

Pour la distribution (2.4), on a

$$\frac{\bar{X}_n - \mu}{S_n/\sqrt{n-1}} = \frac{Z}{\sqrt{Y/(n-1)}},$$

où  $Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \sim \mathcal{N}(0, 1)$  et  $Y = \frac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1}$  indépendantes. D'où la loi de Student.

2. C'est le corollaire 1.3.1. ■

### 2.1.3 Estimation de la variance

Lorsque la moyenne  $\mu$  de  $X$  est inconnue, la **variance empirique**  $S_n^2$ , définie par (1.2), représente l'estimateur naturel de la variance  $\sigma^2$ . Comme  $\bar{X}_n$ , l'estimateur  $S_n^2$  est obtenu par les méthodes du MV et des moments pour les lois de probabilités usuelles. Dans le cas où  $\mu$  est connue, la statistique

$$R_n^2 := \frac{1}{n} \sum_{i=1}^n (X_i - \mu)^2, \quad (2.5)$$

constitue un estimateur de  $\sigma^2$  plus précis que  $S_n^2$ . En effet, cette dernière est calculée sur la base de  $\bar{X}_n$  qui est une valeur approchée de  $\mu$ , alors que  $R_n^2$  est exprimée en termes de la valeur exacte de  $\mu$ .

#### Proposition 2.1.3 (Propriétés de $S_n^2$ )

1.  $S_n^2$  est un estimateur convergent (faiblement et fortement) de la variance  $\sigma^2$ .
2.  $S_n^2$  est un estimateur biaisé pour  $\sigma^2$ . On construit un estimateur sans biais par

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

3.  $\tilde{S}_n^2$  est un estimateur asymptotiquement sans biais.
4. Si  $X \sim \mathcal{N}(\mu, \sigma^2)$ , alors

$$\frac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1}. \quad (2.6)$$

**Preuve.**

1. On réécrit la variance empirique comme

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

On applique deux fois les LGN [1.3.1](#), on obtient une limite égale à  $\sigma^2$ .

2. D'après la proposition [1.2.2](#), on a  $E(S_n^2) \neq \sigma^2$ . D'autre part, on a

$$E(\tilde{S}_n^2) = \frac{n}{n-1} E(S_n^2) = \sigma^2.$$

3. On a  $\lim_{n \rightarrow \infty} E(S_n^2) = \lim_{n \rightarrow \infty} \frac{n-1}{n} \sigma^2 = \sigma^2$ .

4. C'est le résultat [\(1.4\)](#). ■

**Remarque 2.1.1**  $R_n^2$  est un estimateur convergent (faiblement et fortement), sans biais de  $\sigma^2$  et vérifiant

$$\frac{nR_n^2}{\sigma^2} \sim \chi_n^2. \tag{2.7}$$

En effet,  $nR_n^2/\sigma^2$  est égal à la somme des carrés de  $n$  v.a normales centrées réduites indépendantes.

### 2.1.4 Estimation d'une proportion

Soit  $0 < p < 1$  la proportion (pourcentage) d'individus, dans une population, possédant un certain caractère. C'est le paramètre de succès dans une expérience de Bernoulli. L'estimateur naturel de  $p$  est ce que l'on appelle la **fréquence empirique**, notée  $K_n$ , calculée sur un échantillon de taille  $n \geq 1$ . Elle peut être considérée comme une moyenne empirique particulière, où les v.a  $X_i$  sont des variables (indépendantes) de Bernoulli de paramètre  $p$ . Elle est alors définie par

$$K_n := \frac{1}{n} \sum_{i=1}^n X_i. \tag{2.8}$$

**Proposition 2.1.4 (Propriétés de  $K_n$ )**

1.  $K_n$  est un estimateur convergent (faiblement et fortement) de  $p$ .
2.  $K_n$  est un estimateur sans biais.
3. La distribution de  $K_n$  est donnée par

$$nK_n \sim \mathcal{B}(n, p) \text{ et } \frac{K_n - p}{\sqrt{p(1-p)/n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (2.9)$$

**Preuve.**

1. Application directe des LGN [1.3.1](#).
2. Calcul direct de l'espérance.
3. Pour le premier résultat, on note que  $nK_n$  est une somme de v.a  $X_i \sim \mathcal{B}(p)$  indépendantes. Le deuxième résultat est une conséquence directe du TCL [1.3.2](#) ■

## 2.2 Principales méthodes d'estimation

Il existe plusieurs façons de construire un estimateur pour un paramètre donné. Les plus populaires sont la méthode des moments et celle du maximum de vraisemblance (MV).

### 2.2.1 Méthode des moments

**Définition 2.2.1** On appelle *estimateur des moments* de  $\theta$ , toute solution  $\hat{\theta}_n$  des équations des moments

$$\mu_k = m_k, \quad k \geq 1,$$

où  $\mu_k$  et  $m_k$  désignent les moments théorique et empirique respectivement définis par [\(2.1\)](#) et [\(2.2\)](#). En d'autres termes, si le paramètre  $\theta$  s'exprime comme fonction  $\varphi(\mu_1, \mu_2, \dots)$  des premiers moments, alors son estimateur  $\hat{\theta}_n$  est de la forme  $\varphi(m_1, m_2, \dots)$ .

**Exemple 2.2.1**

1. **Loi de Bernoulli**  $\mathcal{B}(p)$  : On a  $p = \mu_1$ , donc l'estimateur de  $p$  par la méthode des moments est  $\hat{p} = m_1 = K_n$ . Cet estimateur n'est autre que la proportion de succès lors de  $n$  expériences indépendantes de Bernoulli. On retrouve ainsi le principe d'estimation d'une probabilité par une proportion empirique.
2. **Loi exponentielle** : Si  $X$  est une v.a de loi  $\mathcal{E}(\theta)$ , où  $E(X) = 1/\theta$ . Donc l'estimateur de  $\theta$  par la méthode des moments est  $\hat{\theta}_n = 1/m_1 = 1/\bar{X}_n$ .

**2.2.2 Méthode du maximum de vraisemblance (MV)**

Cette méthode consiste à maximiser une fonction  $L$ , appelée fonction de vraisemblance, définie ci-dessous.

**Définition 2.2.2** On appelle **fonction de vraisemblance** la fonction définie sur  $\Theta$  par

$$L(x_1, \dots, x_n; \theta) := f(x_1, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

où  $(x_1, \dots, x_n)$  est une réalisation de l'échantillon  $(X_1, \dots, X_n)$  et

$$f(x; \theta) := \begin{cases} f_\theta(x) & \text{si } X \text{ est continue de densité } f_\theta, \\ P_\theta(x) = P(X = x) & \text{si } X \text{ est discrète.} \end{cases}$$

**Définition 2.2.3** L'estimateur du MV du paramètre  $\theta$  correspond à la valeur  $\hat{\theta}_n$  où la fonction  $L$  atteint son maximum, c'est à dire

$$\hat{\theta}_n := \arg \max_{\theta \in \Theta} L(\theta).$$

La recherche de l'estimateur du MV revient donc à un problème d'optimisation. En pratique, on traite le logarithme de la fonction de vraisemblance afin de simplifier la résolution.

**Remarque 2.2.1** En général, l'estimateur du MV  $\hat{\theta}_n$  est solution de l'équation (appelée équation de vraisemblance)

$$\frac{\partial L(x_1, \dots, x_n; \theta)}{\partial \theta} = 0 \text{ ou } \frac{\partial \ln L(x_1, \dots, x_n; \theta)}{\partial \theta} = 0,$$

avec

$$\frac{\partial^2 L(x_1, \dots, x_n; \hat{\theta}_n)}{\partial \theta^2} < 0 \text{ ou } \frac{\partial^2 \ln L(x_1, \dots, x_n; \hat{\theta}_n)}{\partial \theta^2} < 0.$$

**Exemple 2.2.2 (loi discrète)** On applique cette méthode à la loi de Poisson de paramètre  $\lambda > 0$ . On a

$$f(x; \lambda) = P_\lambda(x) = e^{-\lambda} \frac{\lambda^x}{x!}.$$

La fonction de vraisemblance s'écrit

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

Il est clair que la dérivation de  $L$  n'est pas simple, alors on utilise la fonction log-vraisemblance

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

dont la dérivée première

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

s'annule pour  $\lambda = \bar{x}_n$ . Sa dérivée seconde est

$$\frac{\partial^2 \ln L(x_1, \dots, x_n; \lambda)}{\partial \lambda^2} = -\frac{n}{\bar{x}_n} < 0.$$

Par conséquent, l'estimateur du MV de  $\lambda$  est  $\hat{\lambda}_n = \bar{X}_n$ .

**Exemple 2.2.3 (loi continue)** On souhaite estimer les paramètres  $\mu$  et  $\sigma^2$  d'une loi normale

à partir d'un  $n$ -échantillon. La fonction densité est

$$f(x) = f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}.$$

La vraisemblance et le log-vraisemblance sont respectivement

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

et

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

On a

$$\frac{\partial}{\partial \mu} \ln L(x_1, \dots, x_n; \mu, \sigma^2) = \frac{n(\bar{x}_n - \mu)}{\sigma^2},$$

qui s'annule en  $\mu = \bar{x}_n$ . La dérivée seconde

$$\frac{\partial^2}{\partial \mu^2} L(x_1, \dots, x_n; \bar{x}_n) = -\frac{n}{\sigma^2} < 0.$$

Ceci conduit à un estimateur du MV  $\hat{\mu}_n = \bar{X}_n$ .

Pour le deuxième paramètre, on a

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

qui s'annule pour  $\sigma^2 = \sum_{i=1}^n (x_i - \mu)^2 / n$ . De plus pour cette valeur de  $\sigma^2$ , on a

$$\frac{\partial^2}{\partial \sigma^2} \ln L(\sigma^2) = \frac{n}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=1}^n (x_i - \mu)^2 = -\frac{n}{2\sigma^4} < 0.$$

En tenant compte de l'estimation de  $\mu$  par  $\bar{x}_n$ , on prend

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

comme l'estimateur du MV de  $\sigma^2$ .

On note que l'un des estimateurs obtenus est sans biais ( $\hat{\mu}_n$ ) et que l'autre est biaisé ( $\hat{\sigma}_n^2$ ).

**Remarque 2.2.2** Les deux méthodes d'estimation ne donnent pas nécessairement le même estimateur. Par exemple, pour le paramètre  $\theta > 0$  de la distribution uniforme sur l'intervalle  $[0, \theta]$ , l'estimateur des moments est  $2\bar{X}_n$  alors que celui du MV est  $\max_{1 \leq i \leq n} X_i$ .

## 2.3 Construction d'un intervalle de confiance

Le problème avec l'estimation ponctuelle est qu'elle n'apporte pas d'information sur la précision des résultats, c'est-à-dire qu'elle ne tient pas compte des erreurs dues aux fluctuations d'échantillonnage. Pour évaluer la confiance que l'on peut avoir en une valeur estimée d'un paramètre, il est nécessaire de déterminer un intervalle contenant, avec une certaine probabilité fixée à priori, la vraie valeur du paramètre. C'est l'estimation par intervalles de confiance.

**Définition 2.3.1** S'il existe des v.a  $\theta_{\min} = \theta_{\min}(X_1, \dots, X_n) < \theta_{\max} = \theta_{\max}(X_1, \dots, X_n)$  telles que

$$P(\theta \in [\theta_{\min}, \theta_{\max}]) = 1 - \alpha,$$

où  $0 < \alpha < 1$ , on dit que l'intervalle  $[\theta_{\min}, \theta_{\max}]$  est un **intervalle de confiance** pour  $\theta$ , de niveau  $1 - \alpha$ . On le note par  $IC_{1-\alpha}(\theta)$ .

### Remarque 2.3.1

1. Le nombre  $\alpha$  représente le risque que la vraie valeur de  $\theta$  n'appartienne pas à  $IC_{1-\alpha}(\theta)$ . C'est la probabilité de l'erreur qu'on commet en affirmant que  $\theta \in IC_{1-\alpha}(\theta)$ . Par contre, le nombre  $1 - \alpha$  représente la confiance qu'on a en disant que  $\theta \in IC_{1-\alpha}(\theta)$ . En d'autres termes, il y a  $(1 - \alpha) \times 100\%$  de chances que la valeur inconnue de  $\theta$  soit comprise entre  $\theta_{\min}$  et  $\theta_{\max}$ . Dans les problèmes pratiques (économie, agronomie, sociologie, sciences biomédicales,...), les valeurs usuelles du risque  $\alpha$  sont 10%, 5% et 1% correspondant à des niveaux de confiance 90%, 95% et 99% respectivement.

2. Il est clair que plus le risque est petit, plus l'IC est large, c'est à dire moins la précision est bonne.
3. Un bon IC est un intervalle dont les bornes  $\theta_{\min}$  et  $\theta_{\max}$  sont des v.a qui dépendent d'un estimateur performant  $\hat{\theta}_n$  de  $\theta$ .

Il est à noter que  $IC_{1-\alpha}(\theta)$  est un intervalle aléatoire. Dans les calculs, on considère les bornes (non aléatoires) associées à une réalisation  $(x_1, \dots, x_n)$  de l'échantillon  $(X_1, \dots, X_n)$ .

En résumé, la construction d'un IC revient à la détermination de  $\theta_{\min}$  et  $\theta_{\max}$  qui vérifient

$$P(\theta < \theta_{\min}) = \alpha_1 \text{ et } P(\theta > \theta_{\max}) = \alpha_2,$$

avec  $\alpha_1 + \alpha_2 = \alpha$ . En général, on choisit  $\alpha_1 = \alpha_2 = \alpha/2$  et on parle alors d'un IC à risques symétriques. D'autre part, un IC de niveau  $1 - \alpha$  est dit unilatéral s'il est défini par

$$P(\theta > \theta_{\min}) = 1 - \alpha \text{ ou } P(\theta < \theta_{\max}) = 1 - \alpha.$$

La connaissance de la distribution de probabilité de  $\hat{\theta}_n$  permet d'obtenir les bornes de l'IC comme on va le voir dans les cas usuels qui suivent.

### 2.3.1 Intervalles de confiance pour les paramètres gaussiens

Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon d'une v.a normale de paramètres  $\mu$  et  $\sigma^2$ . On souhaite construire des IC pour ces derniers.

#### Estimation de la moyenne

L'IC de  $\mu$  est construit autour de son meilleur estimateur qui est la moyenne empirique  $\bar{X}_n$  définie par (2.2).

#### Cas où $\sigma^2$ est connue

En utilisant la relation (2.3), on peut écrire

$$P\left(-z_{1-\alpha/2} \leq \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \leq z_{1-\alpha/2}\right) = 1 - \alpha$$

où  $z_{1-\alpha/2}$  représente le quantile d'ordre  $1 - \alpha/2$  de la loi normale standard, c'est à dire

$$P(\mathcal{N}(0, 1) \leq z_{1-\alpha/2}) = 1 - \alpha/2.$$

Ceci équivaut à

$$P\left(\bar{X}_n - \frac{\sigma}{\sqrt{n}}z_{1-\alpha/2} \leq \mu \leq \bar{X}_n + \frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}\right) = 1 - \alpha.$$

Par conséquent, on a

$$IC_{1-\alpha}(\mu) = \left[ \bar{X}_n - \frac{\sigma}{\sqrt{n}}z_{1-\alpha/2}, \bar{X}_n + \frac{\sigma}{\sqrt{n}}z_{1-\alpha/2} \right]. \quad (2.10)$$

**Remarque 2.3.2** *Il est clair que lorsque  $n$  est grand l'IC devient petit. Par conséquent, pour avoir une bonne précision il faut travailler avec des échantillons de grandes tailles.*

### Cas où $\sigma^2$ est inconnue

Cette fois, on utilise la relation (2.4) pour écrire

$$P\left(-t_{1-\alpha/2} \leq \frac{\bar{X}_n - \mu}{S_n/\sqrt{n-1}} \leq t_{1-\alpha/2}\right) = 1 - \alpha$$

où  $t_{1-\alpha/2}$  représente le quantile d'ordre  $1 - \alpha/2$  de la loi de Student à  $n - 1$  ddl, c'est à dire

$$P(t_{n-1} \leq t_{1-\alpha/2}) = 1 - \alpha/2.$$

De façon équivalente, on a

$$P\left(\bar{X}_n - \frac{S_n}{\sqrt{n-1}}t_{1-\alpha/2} \leq \mu \leq \bar{X}_n + \frac{S_n}{\sqrt{n-1}}t_{1-\alpha/2}\right) = 1 - \alpha.$$

D'où

$$IC_{1-\alpha}(\mu) = \left[ \bar{X}_n - \frac{S_n}{\sqrt{n-1}}t_{1-\alpha/2}, \bar{X}_n + \frac{S_n}{\sqrt{n-1}}t_{1-\alpha/2} \right]. \quad (2.11)$$

On note que lorsque  $n$  est grand, les quantiles de Student peuvent être remplacés par ceux de Gauss.

**Remarque 2.3.3** En vertu du corollaire [1.3.1](#), lorsque  $X$  est une population de loi quelconque ou inconnue, les intervalles définis dans [\(2.10\)](#) et [\(2.11\)](#) deviennent asymptotiques pour des échantillons de grandes tailles ( $n > 30$  en pratique).

## Estimation de la variance

### Cas où $\mu$ est connue

La meilleure estimation de  $\sigma^2$  est donnée par la variable  $R_n^2$  définie par [\(2.5\)](#). En utilisant la distribution [\(2.7\)](#), on écrit

$$P\left(\nu_{\alpha/2} \leq \frac{nR_n^2}{\sigma^2} \leq \nu_{1-\alpha/2}\right) = 1 - \alpha,$$

où  $\nu_{\alpha/2}$  et  $\nu_{1-\alpha/2}$  sont les quantiles de la loi du Khi-deux à  $n$  ddl, d'ordres respectifs  $\alpha/2$  et  $1 - \alpha/2$ , c'est à dire

$$P(\chi^2_n \leq \nu_{\alpha/2}) = P(\chi^2_n \geq \nu_{1-\alpha/2}) = \frac{\alpha}{2}.$$

On a donc

$$P\left(\frac{nR_n^2}{\nu_{1-\alpha/2}} \leq \sigma^2 \leq \frac{nR_n^2}{\nu_{\alpha/2}}\right) = 1 - \alpha,$$

ce qui entraîne que

$$IC_{1-\alpha}(\sigma^2) = \left[ \frac{nR_n^2}{\nu_{1-\alpha/2}}, \frac{nR_n^2}{\nu_{\alpha/2}} \right].$$

### Cas où $\mu$ est inconnue

Dans ce cas, on estime  $\sigma^2$  par la variance empirique  $S_n^2$  définie par [\(1.2\)](#). D'après la distribution [\(2.6\)](#), on peut écrire

$$P\left(\tau_{\alpha/2} \leq \frac{n}{\sigma^2} S_n^2 \leq \tau_{1-\alpha/2}\right) = 1 - \alpha,$$

où  $\tau_{\alpha/2}$  et  $\tau_{1-\alpha/2}$  sont les quantiles d'ordres  $\alpha/2$  et  $1 - \alpha/2$  de la loi du Khi-deux de ddl égal à

$n - 1$ , c'est à dire

$$P(\chi^2_{n-1} \leq \tau_{\alpha/2}) = P(\chi^2_{n-1} \geq \tau_{1-\alpha/2}) = \frac{\alpha}{2}.$$

De la même manière que précédemment on obtient

$$IC_{1-\alpha}(\sigma^2) = \left[ \frac{n}{\tau_{1-\alpha/2}} S_n^2, \frac{n}{\tau_{\alpha/2}} S_n^2 \right].$$

### Estimation d'une proportion

On désire construire un IC pour la proportion dont on a parlé dans le paragraphe 2.1.4. On rappelle que le meilleur estimateur de  $p$  est la fréquence empirique  $K_n$  définie par (2.8). D'après (2.9), on peut déterminer un IC exact basé sur la loi binomiale qui est une distribution exacte pour tout  $n$ , et/ou un IC approximatif basé sur la loi de Gauss qui est une distribution asymptotique. Les bornes de l'IC exact s'expriment en fonction des quantiles de la loi de Fisher-Snédecor. Les détails sur la construction d'un tel IC se trouvent, par exemple, dans [3, page 57]. Pour l'IC approximatif, on utilise la loi asymptotique de (2.9) pour avoir

$$P\left(-z_{1-\alpha/2} \leq \frac{K_n - p}{\sqrt{p(1-p)/n}} \leq z_{1-\alpha/2}\right) \xrightarrow{n \rightarrow +\infty} P(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha,$$

où  $Z$  est une v.a normale centrée réduite, c'est à dire

$$P\left(K_n - z_{1-\alpha/2} \sqrt{p(1-p)/n} \leq p \leq K_n + z_{1-\alpha/2} \sqrt{p(1-p)/n}\right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha.$$

Ceci ne fournit pas un IC car les bornes dépendent de l'inconnue  $p$ . Cependant, on a le même résultat de convergence, en remplaçant  $p$  (dans les bornes de l'intervalle) par son estimateur convergent  $K_n$  (voir détails dans [8, page 310]). On obtient alors

$$P\left(K_n - z_{1-\alpha/2} \sqrt{K_n(1-K_n)/n} \leq \pi \leq K_n + z_{1-\alpha/2} \sqrt{K_n(1-K_n)/n}\right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha.$$

En conclusion, on a

$$IC_{1-\alpha}(p) \approx \left[ K_n - z_{1-\alpha/2} \sqrt{K_n(1-K_n)/n}, K_n + z_{1-\alpha/2} \sqrt{K_n(1-K_n)/n} \right].$$

## 2.4 Application sous $\mathbf{R}$

Dans cette section, on applique les résultats de la section précédente sur des données simulées, dans le but de les vérifier. Les codes sont regroupés dans l'annexe A.

### 2.4.1 Illustration numérique

#### Paramètres d'une population normale

On prélève des échantillons de taille  $n = 100$  d'une population normale standard. On fait varier le niveau de confiance  $1 - \alpha = 0.90, 0.95$  et  $0.99$ . Les résultats sont résumés dans le tableau [2.1](#).

$1 - \alpha$	<b>b. inf</b>	$\hat{\mu}$	<b>b. sup</b>	<b>b. inf</b>	$\hat{\sigma}^2$	<b>b. sup</b>
0.90	-0.183	-0.029	0.125	0.918	1.131	1.469
0.95	-0.213	-0.029	0.156	0.881	1.131	1.542
0.99	-0.272	-0.029	0.215	0.814	1.131	1.701

TAB. 2.1 – Intervalles de confiance, de différents niveaux, pour la moyenne (à gauche) et la variance (à droite) d'une population normale standard basés sur 100 observations.

#### Moyenne d'une population quelconque

On prélève des échantillons de taille  $n = 100$  de deux populations, l'une exponentielle et l'autre de Poisson, de paramètres respectifs  $\theta$  et  $\lambda$  égaux à 1. On utilise les mêmes niveaux de confiance que précédemment et on résume les résultats dans le tableau [2.2](#).

$1 - \alpha$	<b>b. inf</b>	$1/\hat{\theta}$	<b>b. sup</b>	<b>b. inf</b>	$\hat{\lambda}$	<b>b. sup</b>
0.90	0.778	0.957	1.135	0.864	1.050	1.236
0.95	0.744	0.957	1.170	0.827	1.050	1.273
0.99	0.675	0.957	1.239	0.755	1.050	1.345

TAB. 2.2 – Intervalles de confiance, de différents niveaux, pour la moyenne d'une population exponentielle (à gauche) et de Poisson (à droite), de paramètre 1, basés sur 100 observations.

### Proportion de succès

On répète, 100 fois de façon indépendante, une expérience de Bernoulli de probabilité de succès  $p = 0.5$ . Avec les mêmes niveaux de confiance, on obtient les résultats qu'on résume dans le tableau [2.3](#). On note que les intervalles exacts sont obtenus grâce à la fonction `binom.test` de R.

$1 - \alpha$	<b>b. inf</b>	$\hat{p}$	<b>b. sup</b>	<b>b. inf</b>	$\hat{p}$	<b>b. sup</b>
0.90	0.423	0.51	0.596	0.427	0.51	0.593
0.95	0.408	0.51	0.611	0.412	0.51	0.608
0.99	0.378	0.51	0.641	0.381	0.51	0.639

TAB. 2.3 – Intervalles de confiance (exacts à gauche et approchés à droite), de différents niveaux, pour la proportion de succès lors de 100 expériences de Bernoulli de paramètre 0.5.

On constate que, dans tous les tableaux, plus la confiance augmente, plus la précision de l'estimation diminue comme mentionné dans la remarque [2.3.1](#).

## 2.4.2 Illustration graphique

### Paramètres d'une population normale

On fixe le niveau de confiance à 95% et on prélève des échantillons, de taille  $n$  variant de 1 à 500, d'une population normale standard. Les intervalles de confiance relatifs à la moyenne et à la variance sont illustrés par la figure [2.1](#).

### Moyenne d'une population quelconque

On prélève des échantillons de taille  $n$  variant de 1 à 500 de deux populations, l'une exponentielle et l'autre de Poisson, de paramètres respectifs  $\theta$  et  $\lambda$  égaux à 1. Le comportement des estimateurs des deux moyennes ainsi que celui des bornes de confiance, au niveau égal à 0.95, est décrit par la figure [2.2](#).

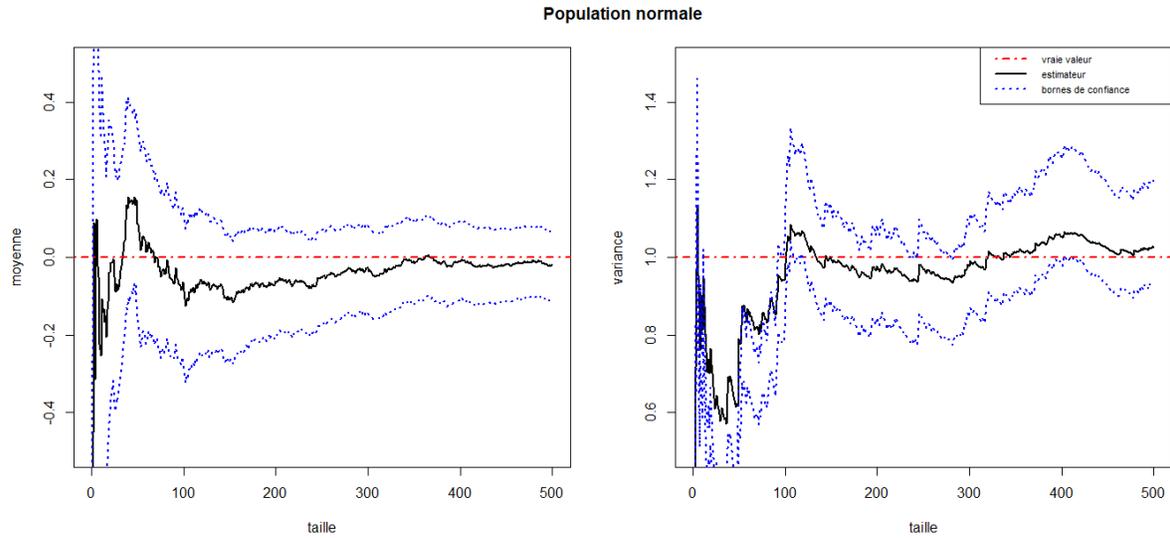


FIG. 2.1 – Intervalles de confiance, de niveau 95%, pour la moyenne (panneau gauche) et la variance (panneau droit) d’une population normale standard.

### Proportion de succès

Des échantillons, de taille variant de 1 à 500, d’une v.a de Bernoulli de paramètre  $p = 0.5$ , sont prélevés dans le but de construire des intervalles d’estimation, à 95% de confiance, pour ce dernier. Les résultats de simulation sont indiqués dans la figure [2.3](#).

On constate que, sur toutes les figures, les intervalles de confiance se rétrécissent au fur et à mesure que les tailles des échantillons grandissent. Ceci confirme donc la remarque [2.3.1](#).

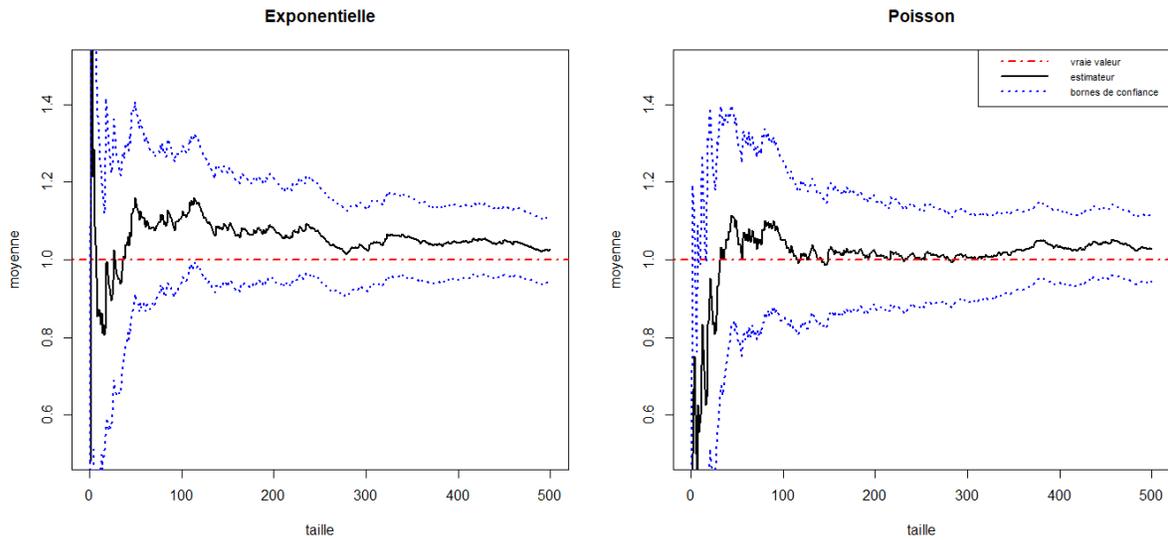


FIG. 2.2 – Intervalles de confiance, de niveau 95%, pour les moyennes d’une population exponentielle (panneau gauche) et d’une v.a de Poisson (panneau droit) de paramètres égaux à l’unité.

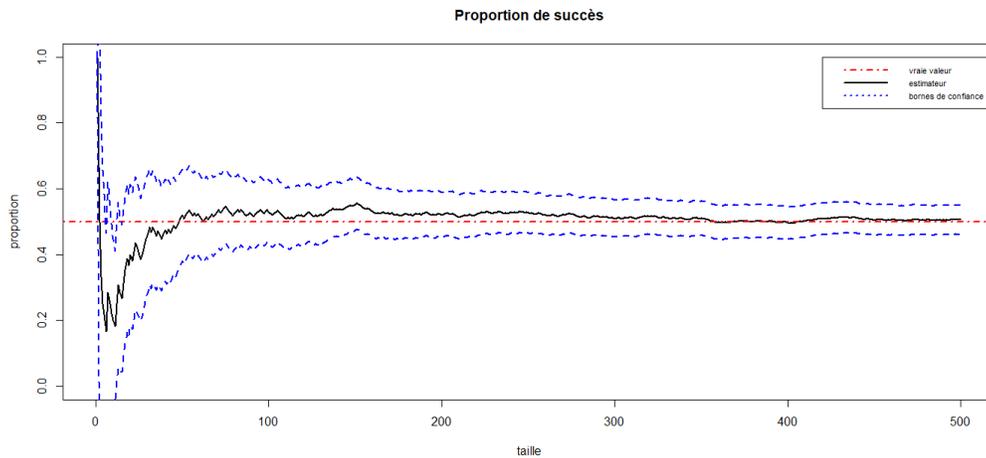


FIG. 2.3 – Intervalles de confiance, de niveau 95%, pour la proportion de succès lors d’expériences de Bernoulli de paramètre 0.5.

# Conclusion

L'estimation par intervalles de confiance a l'avantage de permettre l'évaluation du degré de précision des quantités estimées ponctuellement.

Dans ce travail, on s'est particulièrement intéressé à la construction d'intervalles de confiance pour les paramètres usuels, à savoir la moyenne, la variance et la proportion. Une étude de simulation, sous le logiciel d'analyse statistique R, a permis de vérifier et illustrer numériquement et graphiquement les résultats théoriques obtenus.

D'autres quantités, telles la densité de probabilité et la fonction de répartition (estimation non-paramétrique) et l'indice des queues et les grands quantiles en théorie des valeurs extrêmes (estimation semi-paramétrique) peuvent faire l'objet d'étude similaire.

# Bibliographie

- [1] Dusart, P. (2015). Cours de statistiques inférentielles. Université de Limoges.
- [2] DeGroot, M., Shervish, M. (2012). Probability and statistics. Addison-Wesley.
- [3] Gaudoin, O. (2017). Principes et Méthodes Statistiques. Notes de cours. INP, Grenoble.
- [4] Ihaka, R., Gentleman, R. (1996). R : A language for data analysis and graphics. Journal of Computational and Graphical Statistics 5, 299-314.
- [5] Jourdain, B. (2018). Probabilités et Statistique pour l'Ingénieur. Ecole des Ponts, Paris Tech.  
Site web : <http://cermics.enpc.fr/~jourdain/probastat/poly.pdf>.
- [6] Le Coutre, J.P. (2016). Statistique et Probabilités. Dunod.
- [7] Lejeune, M. (2010). Statistique La Théorie et ses Applications. Springer.
- [8] Saporta, G. (2006). Probabilités, Analyse des données et Statistique. Technip.
- [9] Toulouse, P. (1999). Thèmes de Probabilités et Statistique. Dunod.

# Annexe A : Codes R

Les différents programmes ayant conduit aux résultats numériques et représentations graphiques sont rassemblés ci-dessous.

**Graphes de probabilité de la loi Binomiale de paramètres 10 et 0.5 et de la loi de Poisson de paramètre 1 (figure [1.1](#))**

```
n<-10; p<-0.5
l=2; op<-par(mfrow=c(1,2))
plot(0:n,dbinom(0:10,n,p),type="h",ylab="P(X=k)",xlab="k",lwd=3,main="binomiale (10 ,
0.5)")
plot(0:n,dpois(0:10,1),type="h",ylab="P(X=k)",xlab="k",lwd=3,main="Poisson (1)")
par(op)
```

**Densité de probabilité de la loi normale centrée avec différentes variances (figure [1.2](#))**

```
x<-seq(-5,5,length=100); f<-dnorm(x)
plot(x,f,type="l",ylab="f(x)",lwd=2,); abline(v=0)
curve(dnorm(x,0,2),lty=2,col="blue",lwd=2,add=T)
curve(dnorm(x,0,1.5),lty=3,col="red",lwd=2,add=T)
legend(3.6,0.4,c("sigma = 1","sigma = 1.5","sigma = 2"),
lty=c(1,3,2),col=c("black","red","blue"),cex=0.8)
```

**Densité de probabilité des lois du Khi-deux et de Student avec différents degrés de liberté (figure [1.3](#))**

```
x<-seq(-5,5,length=100);f<-dnorm(x)
plot(x,f,type="l",ylab="f(x)",lwd=2);abline(v=0)
curve(dnorm(x,0,2),lty=2,col="blue",lwd=2,add=T)
curve(dnorm(x,0,1.5),lty=3,col="red",lwd=2,add=T)
legend(3.6,0.4,c("sigma = 1","sigma = 1.5","sigma = 2"),
lty=c(1,3,2),col=c("black","red","blue"),cex=0.8)
```

### Approximation de la loi de Student, avec différents ddl, par la distribution standard de Gauss (figure [1.4](#))

```
curve(dnorm(x),ylab = "f(x)",main="",xlim=c(-4,4),lwd=2)
curve(dt(x,1),ylab = "f(x)",col="brown",lwd=2,lty=2,add=T)
curve(dt(x,30),ylab = "f(x)",col="red",lwd=2,lty=3,add=T)
curve(dt(x,50),ylab = "f(x)",col="blue",lwd=2,lty=4,add=T)
legend(2.5,0.4,c("N(0,1)","ddl = 1","ddl = 30","ddl = 50"),
lty=c(1,2,3,4),col=c("black","brown","red","blue"),cex=0.7)
```

### Illustration de la LGN sur la moyenne empirique d'observations uniformes sur et Gaussiennes standard (figure [1.5](#))

```
n<-250; N<-seq(1,n, by=1); X1<-runif(n); X1bar<-cumsum(X1)/N;
X2<-rnorm(n); X2bar<-cumsum(X2)/N
op<-par(mfrow=c(1,2))
plot(N,X1bar,ylim=c(0,1),main="Moyenne empirique d'observations uniformes",xlab="Taille
de l'échantillon",ylab="Moyenne empirique");abline(h=0.5,col="red")
plot(N,X2bar,ylim=c(-1,1),main="Moyenne empirique d'observations normales",xlab="Taille
de l'échantillon",ylab="Moyenne empirique");abline(h=0,col="red")
par(op)
```

### Normalité asymptotique de la somme de 50 observations Poissonniennes et exponentielles (figure [1.6](#))

```
n=1000; N=50; X=rep(0,n); Y=rep(0,n)
```

```
#n itérations
for (i in 1 :n)
  {X[i]=sum(rpois(N,1))
  Y[i]=sum(rexp(N))}
mu=N ; sigma=sqrt(N)
op=par(mfrow=c(1,2))

#histogramme
hist(X,col="grey",probability=T,ylim=c(0,0.06),xlab="x",ylab="f(x)",main="Somme d'observations de Poisson")
curve(dnorm(x,mu,sigma),col="blue",add=T,lwd=2)
for (i in 1 :n) X[i]=sum(runif(N))

#histogramme
hist(Y,col="grey",probability=T,ylim=c(0,0.06),xlab="x",ylab="f(x)",main="Somme d'observations exponentielles")
curve(dnorm(x,mu,sigma),col="blue",add=T,lwd=2)
par(op)
```

### Illustration graphique

#### Paramètres d'une population normale (figure [2.1](#))

```
n=500 ; x=rnorm(n) ; mu=(cumsum(x))/(1 :n) ; ss=cumsum((x-mu)^2) ;
s=c(ss[1],ss[2 :n]/(1 :(n-1))) ; sd<-sqrt(s)
q=qt(0.975,n-1) ; icm1=mu-q*sd/sqrt(1 :n) ; icm2=mu+q*sd/sqrt(1 :n) ;
l1=qchisq(0.025,n-1) ; l2=qchisq(0.975,n-1)
icv1=((n-1)*s^2)/l2 ; icv2=((n-1)*s^2)/l1
op<-par(mfrow=c(1,2))
plot(mu,type="l",lwd=2,ylab="moyenne",xlab="taille", ylim=c(-0.5,0.5))
abline(h=0,,lty=4,col="red",lwd=2)
lines(icm1,col="blue",lty=3,lwd=2)
lines(icm2,col="blue",lty=3,lwd=2)
```

```
plot(s,type="l",lwd=2,ylab="variance",xlab="taille", ylim=c(0.5,1.5))
abline(h=1,lty=4,col="red",lwd=2)
lines(icv1,col="blue",lty=3,lwd=2)
lines(icv2,col="blue",lty=3,lwd=2)
par(op)
title(main="Population normale")
legend(435,1.55,c("vraie valeur","estimateur","bornes de confiance"),
lty=c(4,1,3),lwd=c(2,2,2),col=c("red","black","blue"),cex=0.7)
```

### **Moyenne d'une population quelconque** (figure [2.2](#))

```
n=500; x1=rexp(n); x2=rpois(n,1); mu1=(cumsum(x1))/(1:n);
mu2=(cumsum(x2))/(1:n); ss1=cumsum((x1-mu1)^2)
s1=c(ss1[1],ss1[2:n]/(1:(n-1))); sd1<-sqrt(s1);
ss2=cumsum((x2-mu2)^2); s2=c(ss2[1],ss2[2:n]/(1:(n-1)))
sd2<-sqrt(s2); q=qt(0.975,n-1);
ice1=mu1-q*sd1/sqrt(1:n); ice2=mu1+q*sd1/sqrt(1:n)
icp1=mu2-q*sd2/sqrt(1:n); icp2=mu2+q*sd2/sqrt(1:n)
op<-par(mfrow=c(1,2))
plot(mu1,type="l",ylab="moyenne",xlab="taille",ylim=c(0.5,1.5),lwd=2,main="Exponentielle")
abline(h=1,lty=4,col="red",lwd=2)
lines(ice1,col="blue",lty=3,lwd=2)
lines(ice2,col="blue",lty=3,lwd=2)
plot(mu2,type="l",ylab="moyenne",xlab="taille",ylim=c(0.5,1.5),lwd=2,main="Poisson")
abline(h=1,lty=4,col="red",lwd=2)
lines(icp1,col="blue",lty=3,lwd=2)
lines(icp2,col="blue",lty=3,lwd=2)
par(op)
legend(435,1.55,c("vraie valeur","estimateur","bornes de confiance"),
lwd=c(2,2,2),lty=c(4,1,3),col=c("red","black","blue"),cex=0.7)
```

**Proportion de succès** (figure 2.3)

```
n=500; x=rbinom(n,1,0.5); kn=cumsum(x)/(1:n); q=qnorm(0.975)
ic1=kn-q*sqrt(kn*(1-kn))/sqrt(1:n); ic2=kn+q*sqrt(kn*(1-kn))/sqrt(1:n)
plot(kn,type="l",ylab="proportion",xlab="taille",ylim=c(0,1),lwd=2)
abline(h=0.5,lty=4,lwd=2,col="red")
lines(ic1,col="blue",lty=2,lwd=2)
lines(ic2,col="blue",lty=2,lwd=2)
title(main="Proportion de succès")
legend(420,1,c("vraie valeur","estimateur","bornes de confiance"),
lty=c(4,1,3),lwd=c(2,2,2),col=c("red","black","blue"),cex=0.7)
```

# Annexe B : Abréviations et Notations

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

<b>Notation</b>	<b>Signification</b>
b. inf	borne inférieure
b. sup	borne supérieure
LGN	loi des grands nombres.
TCL	théorème centrale limite.
v.a	variable aléatoire.
iid	indépendantes identiquement distribuées.
$IC$	intervalle de confiance.
MV	maximum de vraisemblance.
ssi	si et seulement si.
$\mathcal{B}(p)$	loi de Bernoulli
$\mathcal{B}(n, p)$	loi binomiale
$\mathcal{P}(\lambda)$	loi de Poisson
$\mathcal{U}(n)$	loi uniforme discrète
$\mathcal{E}(\theta)$	loi exponentielle
$\chi^2_n$	loi de Khi-deux
$\mathcal{N}(\mu, \sigma^2)$	loi normale
$\mathcal{N}(0, 1)$	loi normale centrée réduite

<b>Notation</b>	<b>Signification</b>
$t_n$	loi de Student
$\mathcal{U}([a, b])$	loi uniforme continue
$E(X)$	espérance de la v.a $X$
$Var(X)$	variance de la v.a $X$
$\approx$	approximativement égal à
$\sim$	suit la loi
$\rightarrow$	converge vers
$\xrightarrow{P}$	convergence en probabilité
$\xrightarrow{ps}$	convergence presque sûre
$\xrightarrow{mq}$	convergence en moyenne quadratique
$\xrightarrow{\mathcal{L}}$	convergence en loi

ميزة تقدير مجال الثقة هي أنه يجعل من الممكن قياس درجة دقة المقدر. لا يمكن ضمان ذلك من خلال تقدير النقطة الذي هو فقط لتوفير تقريب القيمة الحقيقية. يتم بناء مجال الثقة حول مقدر، ويعمل بشكل جيد من خلال استغلال توزيع الاحتمال للمقدر.

**الكلمات المفتاحية :** التقدير الاحصائي؛ المحاكاة؛ توزيع غوص؛ توزيع بيرسون؛ توزيع ستودنت؛ قانون الاعداد الكبيرة؛ مجال الثقة؛ معامل؛ التقارب؛ توزيع الاحتمال؛ مبرهنة النهاية المركزية؛ مستوى الثقة.

## Abstract

The advantage of confidence interval estimation is that it makes it possible to measure the degree of precision of an estimator. This cannot be ensured by the point estimate which settles only to provide an approximation of the true value. A confidence interval is built around an estimator, well performing, by exploiting the probability distribution of the estimator.

**Keywords :** Central limit theorem ; Confidence interval ; Confidence level ; Convergence ; Gaussian distribution ; Law of large numbers ; Parameter ; Pearson distribution ; Probability distribution ; Simulation ; Statistical estimation ; Student distribution.

## Résumé

L'intérêt de l'estimation par intervalles de confiance est qu'elle permet de mesurer le degré de précision d'un estimateur. Ceci ne peut être assuré par l'estimation ponctuelle qui se contente seulement de fournir une approximation de la vraie valeur. Un intervalle de confiance est construit, autour d'un estimateur, performant en exploitant la distribution de probabilité de ce dernier.

**Mots clés :** Convergence ; Distribution de probabilité ; Estimation statistique ; Intervalle de confiance ; Loi de Gauss ; Loi de Pearson ; Loi de Student ; Loi des grands nombres ; Niveau de confiance ; Paramètre ; Simulation ; Théorème central limite.