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

HAIKEL AHLEM

Titre :

Méthodes D'estimation Paramétrique

Membres du Comité d'Examen :

Dr.	BENAMEUR Sana.................	UMKB	Président
Pr.	BENATIA Fatah....................................	UMKB	Encadreur
Dr.	TOUBA Sonia...........................	UMKB	Examinateur

Septembre 2020

Dédicace

Je remercie dieu "Allah" le plus puissant qui m'a éclairé le chemin et de ce que je suis aujourd'hui.

À qui m'a éduqué le don sans avoir attendre.

À qui je prend son nom avec grande fièrté, à qui m'a poussé obtenir à ma promière école dans la vie : "Mon père".

À mon ange dans la vie, à le sens de l'amour et tendresse, à celle qui sa prière étais le secret de mon succés à : "Mon mère".

À ceux qu'ont été prés de mois, qui ont grandis avec moi dans le bon et le mal mon frères et ma soeur : samir,djalal et hayet.

À mes amis : Romaissa Khadraoui, Bouthaina El Guerri, Ahlem Bouden, Somaia Hani.

REMERCIEMENTS

Je voudrais dans un premier temps remercier, mon directeur de mémoire **F.Benatia**, professeur de mathématique à l'université de Biskra, pour sa patience, sa disponibilité et surtout ses judicieux conseils, qui ont contribué alimenter ma réflexion.

Je tiens aussi à remercier Dr.BENAMEUR Sana pour avoir accepté de présider ce jury, ainsi que Dr.TOUBA Sonia pour avoir donner de son temps et juger ce modeste travail, qu'elles trouvent ici toutes mes considération.

J'adresse mes sincères remerciements à tous les professeurs, intervenants et toutes les personnes qui par leurs paroles, leurs écrits, leurs conseils et leurs critiques ont guidé mes réflexions et ont accepté de me rencontrer et de répondre mes questions durant mes études universitaires.

Table des matières

Table des figures

Liste des tableaux

Introduction

L'estimation paramétrique est une méthode statistique qui permet d'estimer les valeurs des paramètres fondée sur des données empirique mesurées et elle est utilisée pour estimer les paramètres de la population, le but est de déterminer la valeur "approchée" $\hat{\theta}$ du paramètre inconnu θ.

La procédure d'utilisation des informations obtenues à partir d'un échantillon qui permet de déduire des résultats concernant l'ensemble de la population est appelée estimation.

La valeur inconnue d'une population à estimer à partir d'un échantillon est appelée un paramètre qui peut être moyenne, une variance, un écate-type, où un pourcentage. Le paramètre de la population est estimé à partir d'une statistique calculer sur la base d'un échantillon. Soit X un certain caractère de distribution $f(x, \theta)$ connue et qui dépend d'un paramètre inconnue θ, on cherche à estimer θ. Pour cela, on tire un échantillon aléatoire de taille n dans la population, et on essaie à partir de l'information obtenue, de déterminer une valeur numérique précise qui sera prise comme valeur du paramètre θ inconnu.

Le but de notre travail est d'explorer l'estimation par la méthode des moments, la méthode du maximum de vraisemblance, et par intervalle de confiance qui sont des méthodes classiques pour estimer les paramètres de la loi de probabilité d'un échantillon donnée. Nous allons aussi donner les propriétés essentielles et étudier particulièrement la qualité de ces estimateurs.

Chapitre 1 : Dans ce chapitre, on va présenter des généralités sur la probabilité et les statistiques donnant la définition d'un espace probabilisé, variable aléatoire, fonction densité de répartition, moments et caractéristiques d'une variable aléatoire, les lois usuelles, et les types de convergence des variables aléatoires.

Chapitre 2 : Dans ce chapitre nous présentons la qualité d'un estimateur en donnant les définitions d'une statistique, d'un estimateur, de la moyenne et la variance empirique, ainsi que l'estimateur avec ou sans biais, l'estimateur convergent (consistant), l'estimateur asymptotiquement normale, l'estimateur efficace et l'estimateur robuste, et enfin la définition de l'information de Fisher.

Chapitre 3 : Nous sommes intéressés dans ce dernier chapitre par la présentation des méthodes d'estimations ponctuelle (méthode de moment et de maximum du vraisemblance) et l'estimation par intervalle de confiance. Leurs carracteristiques et leurs propriétés sont détaillés dans cette partie. En donnant les section : Un maximum de vraisemblance ; Dénition de statistique exhaustive, méthode de moment **(EMM)**, méthode du maximum de vraisemblance (**EMV**), et l'intervalle de confiance, les résultats et les représentations graphiques obtenus à l'aide du logiciel de traitement statistique R. Les codes de programmation sous **R** et les abréviations et notations utilisées sont rassemblés en deux annexes.

Chapitre 1

Notions usuelles de Probabilités et Statistiques

Nous commençons d'abords par présenté d'une manière simple les principaux outils probabilistes de base utilisé dans toute le suite de ce travail.

Parmi celles-ci l'espace probabilisé, les lois et distributions des variables aléatoires, les differentes formes de convergences stochastiques, les lois faibles et forts des grands nombres, ainsi que le théorème centrale limite.

1.1 Espace Probabilisé

A toute experience on associé l'ensemble de toutes les réalisations possible appellé ensemble fondamentale noté Ω et $\mathcal{P}(\Omega)$ l'ensemble de toutes les parties de Ω, le couple $(\Omega, \mathcal{P}(\Omega))$ est appelé espace probabilisable.

1.1.1 Définition de tribu

Soit Ω un ensemble, un sous-ensemble $\mathcal{F}$ de $\mathcal{P}(\Omega)$ est une tribu sur Ω si et seulement si :

1.$\phi \in \mathcal{F}$.

2.$\mathcal{F}$ est stable par réunion dénombrable :

Pour toute famille $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{F}$ on a :

$$\cup_{n\in\mathbb{N}} A_n \in \mathcal{F}.$$

3.$\mathcal{F}$ est stable par passage au complémentaire

Pour tout $A \in \mathcal{F}$:

$$A^c \in \mathcal{F}.$$

Donc $(\Omega, \mathcal{F})$ est un espace probabilisable.

1.1.2 Définition de la probabilité :

Une mesure de probabilité sur $(\Omega, \mathcal{F})$ est une application $\mathbb{P} : \mathcal{F} \to [0, 1]$ telle que :

1.$\mathbb{P}(\Omega) = 1$.

2.pour toute suite d'évenements $(A_n)_{n\in\mathbb{N}}$ de $\mathcal{F}$ deux à deux disjoints :

$$\mathbb{P}(\cup_{n=0}^{\infty} A_n) = \sum_{n=0}^{\infty} \mathbb{P}(A_n).$$

Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ est appellé espace probabilisé.

1.2 Variable aléatoire

Le concept d'une variable aléatoire (v.a) formalise une grandeur variant selon le résultat d'une expérience aléatoire il existe deux types de v.a : les v.a discrètes et les v.a continues, dont on va définir ci-dessus.

On dit que X défini sur $(\Omega, \mathcal{F})$ dans U est une v.a réelle si X est une application mesurable de

$(\Omega, \mathcal{F})$ dans $U \subseteq \mathbb{R}$ (une partie de $\mathbb{R}$)

$$\begin{array}{ccc} (\Omega, \mathcal{F}) & \xrightarrow{\mathbb{P}} & [0,1] \\ X & \searrow & \\ & & U \end{array}$$

1.2.1 Variable aléatoire discrète

On dit que X est une **v.a discrète** lorsque l'ensemble des valeurs qu'elle peut prendre est fini ou infini dénombrable.

Où

$$\{X = x_i\} = \{\omega \in \Omega : X(\omega) = x_i\} .i \in I$$

Définition 1.2.1 *Soit X une v.a.r discrète.On définit la probabilité de x_i par.*

$$p_i = \mathbb{P}(X = x_i), i \in I.$$

Les p_i sont des réels de $[0,1]$ et tels que $\sum\limits_{i \in I} p_i = 1$.

La donnée des p_i définit la **loi de la variable aléatoire** X .

1.2.2 Variable aléatoire continue

On dit que X est une v.a.r **absolument continue ou continue,** si elle admet une fonction f définie de $\mathbb{R}$ dans $\mathbb{R}$ satisfaisant les conditions suivants :

1. $f_X(x) \geq 0$.
2. $\int\limits_{\mathbb{R}} f_X(x)dx = 1$.

Sachant que la fonction f_X est appelé **fonction densité** de X.

et telleque son ensemble d'arrivé est $\mathbb{R}$.

1.3 Distributions d'une variable aléatoire

La loi $\mathbb{P}_X$ d'une v.a (variable aléatoire) est une mesure de probabilité sur$\mathbb{R}$

1.3.1 Définition de fonction de répartition

La fonction de répartition notée F_X d'une v.a X et définie de $(\mathbb{R}, \mathbf{B}_{\mathbb{R}})$ dans $[0, 1]$ par :

$$F_X(x) = \mathbb{P}_X(] - \infty, x]) = \mathbb{P}(X \leq x).$$

Telle que :

1. $\mathbb{P}(X > x) = 1 - F_X(x)$.
2. $\mathbb{P}(x < X \leq y) = F_X(y) - F_X(x)$.
3. $\mathbb{P}(X = x) = F_X(x) - \lim_{y \to x} F_X(y)$.
4. Elle est croissante,continue à droite et admet une limite à gauche

1.3.2 Définition de fonction densité

Une v.a X est à densité s'il existe une fonction de Lebesgue intégrable positive f_X telle que :

$$\mathbb{P}(X \in A) = \int_A f_X(x)dx.$$

pour tout $A \in \mathcal{F}$, f_X est la densité(de probabilité)de X

Propriété Soit X une v-a de densité f_X et de fonction de répartitionF_X, alors :

1. $\int_{\mathbb{R}} f_X(x)dx = 1$.
2. $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$.
3. $\mathbb{P}(X = x) = 0$ pour tout $x \in \mathbb{R}$.

1.3.3 Moments et caractéristiques d'une variable aléatoire

L'espérance ou la moyenne

Définition 1.3.1 *Soit X une variable aléatoire sur un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$. L'espérance mathématique(ou premier **moment initial**) de X est la valeur moyenne des valeurs prises par X, pondérées par leur probabilité de réalisation.Noté $E(X)$.*

$$E[X] = \begin{cases} \sum\limits_{i \in I} x_i \mathbb{P}[X = x_i]. & Si\ X\ est\ discrète \\ \int\limits_{\mathbb{R}} x f(x) dx. & Si\ X\ est\ à\ densité f \end{cases}$$

à condition que la série et l'intégrale soient **absolument convergente.**

Définition 1.3.2 *La v.a X est dite **centrée** si son espérance mathématique est **nulle** $(E(X) = 0)$.*

Théorème 1.3.1 *Pour toute fonction g*

$$E[g(X)] = \begin{cases} \sum\limits_{i \in I} g(x_i) \mathbb{P}[X = x_i] & si\ X\ est\ discrèt \\ \int\limits_{\mathbb{R}} g(x) f(x) dx & si\ X\ est\ continue \end{cases}$$

La variance

Soit X une variable aléatoire admettant une espérance.

On appelle **variance (**ou deuxième **moment centré**) de X le nombre réel, noté $Var(X)$ telleque

$$Var(X) = E\left[|X - E[X]|^2\right] < \infty$$

Remarque 1.3.1 *Si X admet une variance, on appelle **écart-type** de X le nombre*

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

Moments initiaux et centrés d'une variable aléatoire

Définition 1.3.3 *Soit X variable aléatoire réelle, $k \in \mathbb{N}^*$, on dit que X admet un* ***moment d'ordre*** *k si $E(X^k) < \infty$, noté μ_k.*

$$\mu_k = E\left[X^k\right].$$

Définition 1.3.4 *Si X admet un espérance, on appelle* ***moment centré d'ordre*** **k** *de X, noté m_k.*

$$m_k = E\left[(X - E(X))^k\right]. \tag{1.1}$$

Corollaire 1.3.1 *Si X admet un moment d'ordre p, alors pour tout $q \leq p$, X admet un moment d'ordre q.*

Transformées classiques

Définition 1.3.5 ***(Fonction génératrice)***

Soit X une variable aléatoire à valeurs dans $\mathbb{R}$.On définit sa fonction génératrice

$$G_X(t) = E[\exp(tX)] = \begin{cases} \sum\limits_x e^{tx}\mathbb{P}(X = x) & \text{si } X \text{ discrèt} \\ \int e^{tx} f(x)\, dx & \text{si } X \text{ continue} \end{cases}.$$

1. Elle est continue sur $[0, 1]$ et C^∞ sur $[0, 1[$.
2. On dit que la v.a X admet un moment d'ordre p, si et seulement si, la fonction G_X est k fois dérivable, et alors

$$G_X^{(k)}(t) = \mu_k$$

Définition 1.3.6 ***(Fonction caractéristique)***

Soit X une variable aléatoire réelle.On définit sa fonction caractéristique $\varphi_X : \mathbb{R} \to \mathbb{C}$ par :

$$\varphi_X(t) = E[e^{itX}]$$

-Cas discrèt :$\varphi_X(t) = \sum_{x \in X(\Omega)} e^{itx} \mathbb{P}(X = x)$.

-Cas continue :$\varphi_X(t) = \int_{\mathbb{R}} e^{itx} f_X(x) dx$.

La fonction caractéristique d'une v.a.r X est la **transformation de Fourier** de la loi de probabilité.

Proposition 1.3.1 *Soit X une v.a.r (continue ou discrète) et a, λ sont des constantes*:

1. $\varphi_{\lambda X}(t) = \varphi_X(\lambda t)$.
2. $\varphi_{X+a}(t) = \exp(ita)\varphi_X(t)$.
3. φ_X est continue, de module inférieur à 1.
4. φ_X caractérise la loi de$X$.
5. Si X, Y sont deux variables indépendantes, $\varphi_{X+Y}(t) = \varphi_X(t)\varphi_Y(t)$.

Théorème 1.3.2 ***(Lévy)***

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires.

– Si $(X_n)_n$ converge en loi vers X, alors $(\varphi_{X_n})_n$ converge simplement vers φ_X.

– Si $(\varphi_{X_n})_n$ converge simplement vers φ continue en $\mathbf{0}$, alors φ est la fonction caractéristique d'une variable X, et X_n converge en loi vers X.

Pour les besoins du prochain chapitre nous allons rappelés certains types de convergence stochastique.

Théorème 1.3.3 *Si $E\left(X^k\right) < \infty$, $k = 1, 2, ...,$alors*

$$E(X^k) = i^{-k} \frac{d^k}{dt^k} \varphi_X(0).$$

Parmi les lois des variables aléatoires usuelles on peut citée :

1.3.4 Les lois discrètes usuelles

Dans cette partie on v.a cité quelques lois de v.as discrètes comme la loi de Bernoulli et les lois Binomiale, Poisson, Géometrique et Uniforme discrète et continue.

loi de Bernoulli $(B(p))$ **:**

Le plus simple des variables aléatoires discrètes est l' indicatrice d'évènement.Si A est un évènement de probabilité p, la variable aléatoire $\mathbf{1}_A$ prend la valeur 1 si A est réalisé, et 0 sinon.sa loi est **la loi de Bernoulli** de paramètre p.

$$X = \mathbf{1}_A : \mathbb{P}(A) = p$$

$$\begin{cases} \mathbb{P}(\mathbf{1}_A = 0) = q \\ \mathbb{P}(\mathbf{1}_A = 1) = p \end{cases} \quad p + q = 1.$$

Avec

1. $E(X) = p.$
2. $Var(X) = pq.$
3. $\varphi_X(t) = q + p\exp(it).$
4. $G_X(t) = q + p\exp(t).$

Loi Binomiale $(B(n,p))$ **:**

Si on répète la même expérience de Bernoulli n fois indépendamment et on note par X le nombre de fois où l'évènement A se produit.

La v.a $X = \#succés$ est dite v.a de $B(n,p)$, et donc sa loi est définie par :

$$\mathbb{P}(X = x) = C_n^x p^x (1-p)^{n-x}.$$

Avec

1. $E(X) = np.$
2. $Var(X) = npq.$
3. $\varphi_X(t) = (q + p\exp(it))^n.$
4. $G_X(t) = (q + p\exp(t))^n.$

Loi de poisson $(P(\lambda))$ **:**

De nombreuses variables aléatoires discrètes correspondent à des comptages d'objets possédant un caractère relativement rare dans un grand ensemble : atomes d'un isotope, virus, ...

On utilise souvent une loi de poisson comme modèle pour ces comptages.

Une v.a suit la loi de **Poisson** de paramètre $\lambda > 0$ si elle prend ses valeurs dans $\mathbb{N}$, et si pour tout $k \in \mathbb{N}$:

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

La loi de Poisson est considérée comme étant la loi limite d'une loi Binomiale.

Avec

1. $E(X) = \lambda$.
2. $Var(X) = \lambda$.
3. $\varphi_X(t) = \exp(\lambda(\exp(it) + 1))$.
4. $G_X(t) = \exp(\lambda(\exp(t) + 1))$.

1.3.5 Les lois continues usuelles

Dans cette partie, nous examinerons certains concepts des lois usuelles continues telleque les lois (exponentielle, normale, uniforme, ...) et décrire leurs divers caractéristiques.

On note dans toute la suite la fonctin de répartition de X par F_X et f_X par densité.

Loi Normal $(\mathcal{N}(\mu, \sigma^2))$

Soient $\mu \in \mathbb{R}$ et $\sigma > 0$, on dit que la v.a X suit la loi normale (ou gaussienne) si sa fonction densité f est donnée par :

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

pour $\mu = 0$ et $\sigma^2 = 1$ on a $X \backsim \mathcal{N}(0,1)$ loi normale **centrée réduite.**

Donc la fonction de répartition :

$$F_Z(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-x^2/2\right) dx.$$

Avec

1. $E(X) = \mu$.
2. $Var(X) = \sigma^2$.
3. $\varphi_X(t) = \exp\left(i\mu t + t^2/2\sigma^2\right)$.
4. $G_X(t) = \exp\left(\mu t + t^2/2\sigma^2\right)$.

Loi uniforme $(U([a,b]))$

On dit que la v.a X suit la loi **uniforme** sur l'intervalle **[a,b]** qu'on note par $X \backsim U([a,b])$ si sa fonction densité f est donnée par

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

Sa fonction de répartition :

$$F_X(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(x)dx = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

Avec

1. $E(X) = \frac{a+b}{2}$.
2. $Var(X) = \frac{(b-a)^2}{12}$.
3. $\varphi_X(t) = \frac{\exp(itb)-\exp(ita)}{(b-a)it}$ avec $t \neq 0$.
4. $G_X(t) = \frac{\exp(tb)-\exp(ta)}{(b-a)t}$ avec $t \neq 0$.

Loi exponentielle $(\zeta(\theta))$

Soit $\theta > 0$, on dit que la v.a X suit la loi **exponentielle** si sa fonction densité f est

$$f(x) = \theta e^{-\theta x}.$$

Donc la fonction de répartition :

$$F_X(x) = 1 - e^{-\theta x}.$$

Avec

1. $E(X) = 1/\theta.$
2. $Var(X) = 1/\theta^2.$
3. $\varphi_X(t) = \theta/(\theta - it).$
4. $G_X(t) = \theta/(\theta - t).$

1.4 Les types de convergence des variables aléatoires :

Parmi celles-ci la convergence simple,en probabilité et surtout la convergence en loi.

Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v-a et X une autre v-a toutes construites sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{R})$.

1.4.1 La convergence simple :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge simplement vers X pour n assez grand (et on écrit $\lim\limits_{n\to\infty} X_n = X$) si et seulement si :

$$\forall\varepsilon > 0, \exists N_\varepsilon \geq 0, \text{telque } \forall n \geq N_\varepsilon \text{ on a } |X_n - X| < \varepsilon.$$

1.4.2 La convergence en probabilité :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X si :

$$\forall\varepsilon > 0 \lim \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

et on note :$X_n \underset{n\to\infty}{\overset{p}{\to}} X$

Théorème 1.4.1 *Si* $\lim\limits_{n\to+\infty} E(X_n) = m$ *et* $V(X_n) = 0$,*alors* X_n *converge en probabilité vers* m.

1.4.3 La convergence en moyenne d'ordre p :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en moyenne d'ordre p vers X si :

$$\lim_{n\to+\infty} E\left[|X_n - X|^p\right] = 0.$$

et on note :$X_n \overset{L^p}{\to} X$.

La plus utilisée est la convergence en moyenne quadratique si $p = 2$ et on note $X_n \underset{n\to\infty}{\overset{m.q}{\to}} X$.

On di que $X_n$converge en moyenne quadratique vers X si et seulement si :

$$\lim_{n\to\infty} E[X_n] = E[X]$$

et

$$\lim_{n\to\infty} E\left((X_n - X)^2\right) = 0 \iff \lim_{n\to\infty} Var[X_n] = 0.$$

1.4.4 La convergence en loi :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge en loi vers X de fonction de répartition F_X (resp fonction de caractéristique $\varphi_X(t)$) si, en tout point de continuité de F_X (resp $\varphi_X(t)$), la suite (F_{X_n}) des fonctions de répartition de X_n (resp $\varphi_{X_n}(t)$ des fonction caractéristique de X_n) convergent vers

F_X (resp $\varphi_X(t)$) c-à-d :

$$\forall x \in \mathbb{R} \lim_{n\to+\infty} F_{X_n}(x) = F_X(x),$$

$$\forall t \in \mathbb{R} \lim_{n\to+\infty} \varphi_{X_n}(t) = \varphi_X(t).$$

et on note : $X_n \xrightarrow{\mathcal{L}} X$. On encore $X_n \xrightarrow{\mathcal{D}} X$, pour $n \to \infty$.

Théorème 1.4.2 ***de slutsky***

$\forall c \in \mathbb{R}, X_n \xrightarrow{\mathcal{L}} X$ et $Y \xrightarrow{\mathcal{L}} c$ implique :

1)$X_n + Y_n \xrightarrow{\mathcal{L}} X + c$.

2)$X_n Y_n \xrightarrow{\mathcal{L}} Xc$.

3)$\frac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \frac{X}{c}$.

1.4.5 la convergence presque sûre :

On dit que la suite $(X_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers X si :

$$\mathbb{P}(\forall \omega : X_n(\omega) \underset{n\to+\infty}{\rightarrow} X(\omega)) = 1 \Longleftrightarrow \mathbb{P}(\forall \omega : |\lim_{n\to+\infty} X_n(\omega) = X(\omega)|) = 1.$$

et on note :$X_n \underset{n\to+\infty}{\overset{p.s}{\rightarrow}} X$.

D'autre part on a :

Si $\forall \varepsilon > 0$, la suite$\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon)$ converge.

Alors $X_n \xrightarrow{p.s} X$, quand $n \to \infty$.

Théorème 1.4.3 *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a et X une autre v-a,alors les propriétés suivantes sont équivalentes :*

1)$X_n \xrightarrow{p.s} X$.

2)$\forall \varepsilon > 0 \ \lim_n \mathbb{P}(\sup_{m\geq n} |X_m - X| > \varepsilon) = 0$.

3)$\forall\varepsilon > 0 \ \ \lim\limits_n \mathbb{P}(\cup_{m\geq n}|X_m - X| > \varepsilon) = 0$

4)$\forall\varepsilon > 0 \ \ \lim\limits_n \mathbb{P}(\cap_{m\geq n}|X_m - X| > \varepsilon) = 1.$

Théorèmes Limites :

Loi des grands nombres : Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes et de même loi, on considère $\overline{X_n} = \frac{X_1+...+X_n}{n}$ la moyenne arithmétique de $X_1, ..., X_n$.

On a alors un premier énoncé

Théorème 1.4.4 ***(loi faible des grands nombres)***

Si les X_i admettent les moments dordre 1 et 2, on note $m = E(X_i)$ et $\sigma^2 = Var(X_i)$ alors $(X_n)_{n\geq 1}$ converge en **probabilités** vers la constante m quand n tends vers l'infini .

Et on a

Théorème 1.4.5 ***(loi forte des grands nombres)***

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires de L^1, centrées, indépendantes et identiquement distribuées(iid). Alors, en posant $S_n = X_1 + ... + X_n$,

$$\frac{S_n}{n} \to 0 \ \ ps, \text{ quand } n \to \infty.$$

pour tout $n \geq 1$, et tout $\varepsilon > 0$, on a :

$$\varepsilon\mathbb{P}\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \leq \left\|\frac{S_n}{n}\right\|_1.$$

autrement dit, si $\frac{S_n}{n} \xrightarrow{L^1} 0$, alors $\frac{S_n}{n} \xrightarrow{ps} 0$.

Remarque 1.4.1 *La loi des grands nombres nous dit que la moyenne empirique $\overline{X_n}$ converge en probabilités vers l'espérance m, Mais elle ne nous dit rien en ce qui concerne la vitesse de cette convergence.*

Théorème 1.4.6 *(théorème central limite)*

Soit $(X_n)_{n\in\mathbb{N}}$ une suite des variables aléatoires indépendantes et de même loi admettant une espérance m et une variance σ^2, et posant $\overline{X_n} = \frac{X_1+.....+X_n}{n}$.

On a vu que $E\left[\overline{X_n}\right] = m$ et $Var\left[\overline{X_n}\right] = \sigma^2$.

La suite des v.a

$$Y_n = \frac{\sqrt{n}}{\sigma}(\overline{X_n} - m) \quad \forall n \geq 1.$$

converge en loi vers la loi $\mathcal{N}(0,1)$, et on note

$$\frac{(X_n - m)}{\sigma/\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1). \tag{1.2}$$

La relation entre les défférents types de convergence :

$$\begin{array}{ccccc} & & L^p & & \\ & & \Downarrow & & \\ p.s & \Longrightarrow & p & \Longrightarrow & \mathcal{L} \end{array}$$

Chapitre 2

Qualité d'un estimateur

Nous allons dans ce chapitre présentés quelques estimateurs, certaines de leurs propriétés et caractéristiques, qui permettent de les distingues et de les comparer pour choisir les milleurs estimateurs.

2.1 Définition d'un estimateur

Soit X une v.a dont la la loi de probabilité $\mathbb{P}_\theta$ dépend du paramètre $\theta \in \Theta$.

Soit $(x_1, ..., x_n)$ un échantillon aléatoire.

Définition 2.1.1 ***(Statistique)***

Une statistique est une fonction mesurable T des variables aléatoires X_i :

$$(X_1, ..., X_n) \to T(X_1, X_2, ..., X_n).$$

Définition 2.1.2 ***(Un estimateur)***

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

2.2 Moyenne et Variance empirique

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

2.2.1 Moyenne empirique $\left(\overline{X_n}\right)$

Définition 2.2.1 *On appelle **moyenne empirique** (échantillonnale, expérimentale) la statistique notée $\overline{X_n}$ définie par :*

$$\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i. \tag{2.1}$$

Caractéristique de la statistique $\overline{X_n}$

1. Soit X une v.a d'espérance μ et de variance σ^2, alors

$$E\left(\overline{X_n}\right) = \mu.$$
$$Var\left(\overline{X_n}\right) = \frac{\sigma^2}{n}.$$

2. Les moment centrée (1.1) de $\overline{X_n}$ d'ordre 3 et 4 sont respectivement

$$\mu_3^c\left(\overline{X_n}\right) = \frac{\mu_3^c}{n^2}.$$
$$\mu_4^c\left(\overline{X_n}\right) = \frac{\mu_4^c + 3\sigma^4(n-1)}{n^3}.$$

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

Remarque 2.2.1 *Lorsque n est grand, la moyenne empirique est **symétrique et d'aplatissement normal (égale à 3)**.Ce qui fait penser à sa normalité asymptotique qui sera confirmer par le **(TCL)**.*

2.2.2 Variance empirique$\left(S_n^2\right)$

Définition 2.2.2 *On appelle* ***Variance empirique,*** *la statistique*

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}_n\right)^2. \tag{2.2}$$

Caractéristique de la statistique S_n^2

Si X est une v.a de variance σ^2 alors

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

Remarque 2.2.2 *Lorsque n tend vers l'infini on a*

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

2.2.3 Cas d'échantillon Gaussiens

Lorsque la population est Gaussienne, on a les résulats suivants

Soit $X \backsim \mathcal{N}\left(\mu, \sigma^2\right)$ on a :

1. La moyenne empirique est Gaussienne

$$\overline{X}_n \backsim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

2. La variance empirique vérifie

$$\frac{nS_n^2}{\sigma^2} \backsim \chi_{n-1}^2. \tag{2.3}$$

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

2.2.4 Représentation graphique de la moyenne et la variance empirique d'observations uniformes et normales

On prélève des échantillon de taille $n = 500$, d'une population uniforme et d'une population normale standard respectivement. La figure (2.1) suivante représenté la moyenne et la variance de ces population :

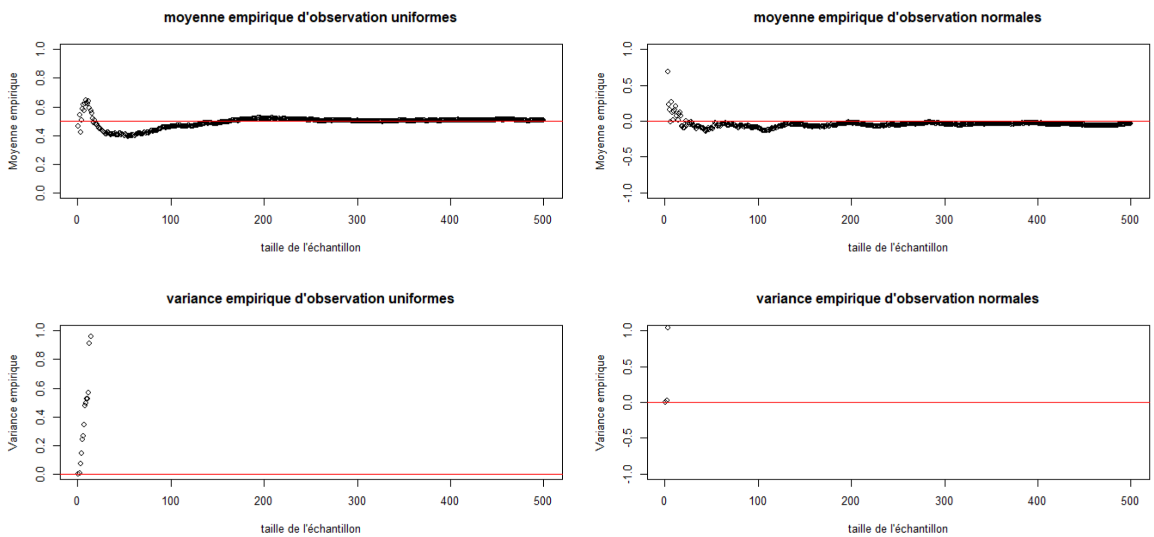

FIG. 2.1 – Représentation graphique de la Moyenne et de la Variance empirique d'observations uniformes et Gaussiennes standard

2.3 Les propriétés et caractéristiques d'un estimateur :

2.3.1 Le biais d'un estimateur

Estimateur sans biais

Un estimateur T est une v.a dont on suppose connue. La loi de probabilité pour une valeur de $\theta \in \Theta$ est fixé. L'erreur d'estimation est la v.a $(T_n - \theta)$ que l'on peut décomposer en

$$T_n - \theta = T_n - E(T_n) + E(T_n) - \theta.$$

où $E(T)$ est l'espérence de T.

Le premier terme $T_n - E(T_n)$ représente les variations aléatoire de T_n par rapport à sa valeur moyenne, la deuxième la partie $E(T_n) - \theta$ représente l'erreur systématique due au fait que T_n varie autour de sa valeur $E(T_n)$ et non autour de θ, sauf si $E(T_n) = \theta$.

Définition 2.3.1 *On appelle* ***biais*** *d'un estimateur T_n au point θ la fonction $b_{\sigma_n}(\theta)$ définie par :*

$$\theta \to b_{T_n}(\theta) = E(T_n) - \theta \qquad \forall\theta \in \Theta.$$

On dit que l'estimateur est :

-Un estimateur **sans biais** si :

$$b(T_n, \theta) = 0 \implies E(T_n) = \theta \qquad \forall\theta \in \Theta.$$

-Un estimateur **asymptotiquement sans biais** si

$$\lim_{n\to\infty} E(T_n) = \theta.$$

Exemple 2.3.1 *Soit $(X_1, ..., X_n)$ des variables aléatoires iid et $\forall i = 1, ..., n$ on a $E(X_i) = m$, $Var(X_i) = \sigma^2$.*

La statistique $\overline{X}$ est un **estimateur sans biais** pour l'espérance mathématique m. On obtient l'espérance et la variance de $\overline{X}$

En effet

$$\begin{aligned} b_{\overline{X}}(m) &= E\left[\overline{X}\right] - m \\ &= m - m \\ &= 0. \end{aligned}$$

D'autre part on définit la statistique :

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

Comme un estimateur sans biais de la variance.

Estimateur avec biais

On dit que T_n est un estimateur de biais $b(T_n, \theta)$ si :

$$E(T_n) \neq \theta.$$

$$b(T_n, \theta) = E(T_n) - \theta \neq 0.$$

Exemple 2.3.2 *La statistique S^2 est un* ***estimateur avec biais*** *pour la variance mathématique $Var(X)$*

$$E\left(S_n^2\right) = \frac{n-1}{n}\sigma^2 \neq \sigma^2 \Longrightarrow b_{S_n^2}\left(\sigma^2\right) = -\frac{\sigma^2}{n}.$$

Mais quand n tend vers l'infini dans ce cas S_n^2 est dit estimateur ***asymptotiquement sans biais.***

2.3.2 Estimateur convergent :

Une statistique T_n est convergente si sa distribution tend à se concentrer autour de la valeur inconnue du paramètre θ quand la taille n d'échantillon tend vers l'infini.

Définition 2.3.2 *On dit que l'estimateur est* ***convergent(ou consistant)*** *si la suite T_n converge en probabilité vers θ $\left(T_n \xrightarrow{p} \theta\right)$.*

C'est-à-dire :

$$\forall \alpha \in \mathbb{R}^+,\ \exists \varepsilon > 0\ \mathbb{P}\left(|T_n - \theta| \leq \alpha\right) \geq 1 - \varepsilon.$$

où $\lim\limits_{n\to\infty} \mathbb{P}\left(|T_n - \theta| \leq \alpha\right) = 1.$

-On dit que l'estimateur est **fortement convergent** lorsqu'on a la convergence **presque sûre(p.s)**.

Théorème 2.3.1 *Un estimateur T_n, dont l'espérance mathématique tend vers θ, et la variance tend vers **zéro**, est convergent pour θ c'est-à-dire :*

$$\left\{ \begin{array}{c} E(T_n) \underset{n\to\infty}{\rightarrow} \theta \\ et\ Var(T_n) \underset{n\to\infty}{\rightarrow} 0 \end{array} \right\} \Longrightarrow T_n \underset{n\to\infty}{\rightarrow} \theta.$$

Exemple 2.3.3 *L'estimateur $\overline{X}$ est un estimateur convergent .*

En effet :

$$\begin{aligned} E(\overline{X}) &= m \\ et \qquad Var(\overline{X}) &= \frac{\sigma^2}{n} \underset{n\to\infty}{\rightarrow} 0. \end{aligned}$$

2.3.3 Estimateur asymptotiquement normale

Définition 2.3.3 *Soit T_n un estimateur du paramètre θ de la loi $\mathbb{P}_\theta$ d'une v.a X. Nous supposons qu'il existe deux fonctions $a = a(\theta, n)$ et $b = b(\theta, n)$ telle que :*

$$\lim_{n\to\infty} \mathcal{L}\left(\frac{T_n - a}{b}\right) = \mathcal{N}(0,1).$$

Nous disons alors que T_n est un **estimateur asymptotiquement normale (a.n).**

Exemple 2.3.4 *Soit une v.a centrée et possèdant un moment d'ordre 4 :*

$$E\left(X^4\right) = \mu_4 < \infty.$$

Soit $(X_1, ..., X_n)$ *n-échantillon telque*

$$E(X) = \mu = 0 \text{ et } \sigma^2 = E\left(X^2\right) < \infty.$$

Nous considérons l'estimateur

$$T_n = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

en effet

$$E(T_n) = \sigma^2,$$
$$Var(T_n) = \frac{\mu_4 - \sigma^2}{n}.$$

Ainsi T_n est un estimateur convergent et sans biais de σ^2.De plus, il satisfait au condition du T.C.L. Nous pouvons écrire :

$$\lim_{n\to\infty} \mathcal{L}\left(\frac{\sqrt{n}\left(T_n - \sigma^2\right)}{\sqrt{\mu_4 - \sigma^4}}\right) = \mathcal{N}(0, 1),$$

et en déduire que T_n est un estimateur asymptotiquement normale pour $a = \sigma^2$ et $b = \sqrt{\frac{\mu_4 - \sigma^2}{n}}$.

2.3.4 Robustesse et efficacité d'un estimateur

Estimateur efficace

Définition 2.3.4 *On mesure la précision d'un estimateur* T *du paramètre* θ *par* ***l'erreure quadratique moyenne*** *définie par :*

$$EQM(T) = E\left[(T - \theta)^2\right].$$

Théorème 2.3.2 *Si* T *est un estimateur du paramètre* θ *alors*

$$EQM(T) = Var(T) + (b_T(\theta))^2. \tag{2.4}$$

Pour rendre l'erreur quadratique moyenne la plus petite possible, il faut que :

-$E(T_n) = \theta$, donc choisir un estimateur sans biais.

-$Var(T_n)$ soit petite.

Parmi les estimateurs sans biais, on choisira donc celui qui a la variance la plus petite, cette propriété traduit **l'efficacité** de l'estimateur.

Estimateur Robuste

Soit X une v.a. à valeurs sur (L, B, P), $(X_1, ..., X_n)$ les v.a. associées au modèle d'échantillonnage $(L, B, P)^n$, T_n une statistique définie sur $(L, B, P)^n$ et à valeurs dans $\left(\mathbb{R}^k, \mathbb{R}^k\right)$, P étant muni de la distance de Prohorov d_p.

Définition 2.3.5 *La suite $T_n = T(X_1, ..., X_n)$ est dite robuste en P si :*

$$\forall \varepsilon > 0; \exists \alpha > 0; \forall \mathrm{Q} \in P; \forall n \in \mathbb{N}^*$$
$$d_p(P, \mathrm{Q}) < \alpha \Longrightarrow d_p\left(P^{T_n}, \mathrm{Q}^{T_n}\right) \leq \varepsilon.$$

2.4 Information de Fisher

Définition 2.4.1 *On appelle quantité d'information de fisher $I_n(\theta)$ apportée par un échantillon sur le paramètre θ la quantité suivante :*

$$I_n(\theta) = E\left[\left(\frac{\partial(\ln(L(x;\theta)))}{\partial\theta}\right)^2\right].$$

Théorème 2.4.1 *Si le domaine de définition de X ne dépend pas de θ et si la vraisemblance est dérivable deux fois, alors :*

$$I_n(\theta) = E\left[\left(\frac{\partial^2(\ln(L(x;\theta)))}{\partial\theta^2}\right)\right].$$

Si cette quantité existe.

Chapitre 3

Méthodes d'estimation paramétrique

L'estimation permet de donner des valeurs approximatives aux paramètres (μ, $\sigma^2, ...$) inconnus, d'une population en utilisant un échantillon de taille n extrait de cette population d'origine.

Le problème général de l'estimation est le suivant :

Soit X une variable dont la densité $f(x, \theta)$ dépend d'un paramètre inconnu θ ; comment trouver la **"meilleure"** valeur possible de θ à partir d'un échantillon observés $x_1, ..., x_n$?

3.1 Méthodes d'estimation paramétrique (ponctuelle)

Il existe de différents types d'estimation, le cas où la loi de probabilité est **connue** et les paramètres sont **inconnus** est appelé **l'estimation paramétrique**. Quand on prend un échantillon de v.a X_i i.i.d $(i = 1, ..., n)$ selon la loi de probabilité du paramètre θ, alors l'estimation serait appliquée sur ce paramètre par deux méthodes **(La méthode des Moments (EMM) et la méthode de Maximum de Vraisemblance (EMV).**

Mais il faut d'abord définir précisement ce que sont une estimation et surtout un maximum de vraisemblance.

3.1.1 Estimation par la méthode du maximum de vraisemblance(EMV)

Définition du maximum de vraisemblance

Définition 3.1.1 *(fonction de vraisemblance)*

On appelle fonction de vraisemblance de θ *pour une réalisation* $(x_1, ..., x_n)$ *d'un échantillon, l'application* $L : \Theta \to \mathbb{R}$*, telle que*

$$L(x_1, ..., x_n; \theta) = f(x_1, ..., x_n; \theta) = \begin{cases} \prod_{i=1}^{n} f(x_i; \theta), \text{ si } X \text{ est continue} \\ \prod_{i=1}^{n} \mathbb{P}(X = x_i/\theta), \text{ si } X \text{ est discrèt} \end{cases}.$$

où f *est la densité.*

Définition 3.1.2 *Soit* $L(x; \theta)$ *la fonction de vraisemblance (définition .....) au point* $\theta \in \Theta$ Θ *quelconque.*

On dit que l'**estimateur du maximum de vraisemblance (EMV)** pour la statistique, donnée est telque :

$$L\left(x; \hat{\theta}^{MV}\right) \geq L(x; \theta) \qquad \forall \theta \in \Theta \quad ps.$$

Définition de statistique exhaustive

Définition 3.1.3 *La statistique* T *sera dite exhaustive si la loi conditionnelle de* X *sachant* $T(x) = t$ *et indépendante du paramètre* θ*. C'est-à-dire :*

$$\mathbb{P}(X/T(x) = t) \text{ ne dépend pas de } \theta.$$

Théorème 3.1.1 ***de factorisation (Fisher-Neyman)***

Soit le modèle $(X, \mathbb{P}_\theta)$,T une statistique exhaustive pour θ si et seulement si pour tout X, il existe deux fonctions mesurables positives h, g telle que g dépend de θ et t, h dépendant de X, telle que :

$$f(x;\theta) = g(t;\theta)h(x)$$

3.1.2 Équation de vraisemblance et les propriétés l'EMV

Soit un échantillon des valereurs $X = (X_1, X_2 \ldots\ldots X_n)$ de lois de probabilité inconnues mais identiques. Nous cherchons à estimer cette loi $\mathbb{P}$ inconnue à partir des observations $x_1, x_2, ..., x_n$, La méthode d'estimation du maximum de vraisemblance (EMV) est basée sur la vraisemblance, qui est la probabilité conjointe de la série $x_1, x_2, ..., x_n$

$$L(x_1, ..., x_n; \theta) = \prod_{i=1}^{n} f_X(X = x_i).$$

Donc on cherche à la maximiser. La maximisation de $L(\theta)$ est identique à la maximisation de son logarithme $(\ln L(\theta))$.

$$\hat{\theta}^{MV} = \arg\max_{\theta} L(x_1, ..., x_n; \theta) \Leftrightarrow \hat{\theta}^{MV} = \arg\max_{\theta} l(x_1, ..., x_n; \theta).$$

où $l(x_1, ..., x_n; \theta) = \ln(L(x_1, ..., x_n; \theta))$.

Cas d'estimation d'un seul paramètre

L'estimateur qui maximise la vraisemblance est celui qui satisfait les conditions suivantes

$$\begin{cases} \frac{\partial L(x_1,...,x_n;\theta)}{\partial \theta} = 0 \\ \frac{\partial^2 L(x_1,...,x_n;\theta)}{\partial \theta^2} < \infty \end{cases} \text{où} \begin{cases} \frac{\partial l(x_1,...,x_n;\theta)}{\partial \theta} = 0 \\ \frac{\partial^2 l(x_1,...,x_n;\theta)}{\partial \theta^2} < \infty \end{cases}.$$

On prend comme estimateur de θ la solution de l'équation $\frac{\partial l(x_1,...,x_n;\theta)}{\partial \theta} = 0$ et qui vérifie $\frac{\partial^2 l(x_1,...,x_n;\theta)}{\partial \theta^2} < \infty$.

Cas d'estimation simultanée de plusieurs paramètres

Soit X est une v.a de loi de probabilité $\mathbb{P}_\theta$ dépendant d'un paramètre qui peut être un ensemble de plusieurs variables réels $\theta = (\theta_1, ..., \theta_d)$, alors la vraisemblance du paramètre est : $L(x_1, ..., x_n; \theta_1, ..., \theta_d)$.

La méthode du maximum de vraisemblance conduit en général à la résolution du système d'équation, qui vérifié les deux conditions suivantes :

1. 1^{er} cdt est :

$$\begin{cases} \frac{\partial L(x_1,...,x_n;\theta_1)}{\partial \theta_1} = 0, \\ \frac{\partial L(x_1,...,x_n;\theta_2)}{\partial \theta_2} = 0, \\ . \\ . \\ \frac{\partial L(x_1,...,x_n;\theta_d)}{\partial \theta_d} = 0 \end{cases} \quad \text{où} \quad \begin{cases} \frac{\partial l(x_1,...,x_n;\theta_1)}{\partial \theta_1} = 0, \\ \frac{\partial l(x_1,...,x_n;\theta_2)}{\partial \theta_2} = 0, \\ . \\ . \\ \frac{\partial l(x_1,...,x_n;\theta_d)}{\partial \theta_d} = 0 \end{cases}$$

   avec $l(x_1, ..., x_n; \theta_1) = \ln L(x_1, ..., x_n; \theta_1)$.

2. $2^{\text{ème}}$ cdt est :

$$\left[\frac{\partial^2 L(x_1, ..., x_n; \theta)}{\partial \theta_i \partial \theta_j} < 0\right]_{i,j=1,...,d}, \text{ où } \left[\frac{\partial^2 l(x_1, ..., x_n; \theta_d)}{\partial \theta_i \partial \theta_j} < 0\right]_{i,j=1,...,d}$$

   la matrice obtenu doit être définie positive.

Proposition 3.1.1 ***1.****Soit T une statistique exhaustive,toute fonction de x solution de l'équation de la vraisemblance :*$\frac{\partial \ln L(X;\theta)}{\partial \theta} = 0$.

Est solution de l'équation :

$$\frac{\partial L(t;\theta)}{\partial \theta} = 0.$$

Alors pour réaliser un maximum de $L(X;\theta)$,il suffit de réaliser un maximum de $g(t;\theta)$.

2.*Si $\hat{\theta}$ est un estimateur de maximum de vraisemblance de θ, $f\left(\hat{\theta}\right)$ est l'estimateur de maximum de vraisemblance de $f(\theta)$.*

Caractéristique

1. L'EMV peut ne pas existe.
2. La vraisemblance n'est pas a priori dérivable en tout point $\theta \in \Theta$.
3. Il n'y aucune raison pour que l'EMV soit sans biais.
4. L'EMV n'a aucune raison d'être unique.

Remarque 3.1.1 ***1.****S'il existe un estimateur efficace de θ, alors il est égal à l'unique EMV de θ. Mais la réciproque est fausse.*

2.*Soit T une statistique exhaustive,toute fonction de x solution de l'équation de la vraisemblance :*

$$\frac{\partial \ln L(X;\theta)}{\partial \theta} = 0.$$

Est solution de l'équation :

$$\frac{\partial L(t;\theta)}{\partial \theta} = 0.$$

Alors pour réaliser un maximum de $L(X;\theta)$,il suffit de réaliser un maximum de $g(t;\theta)$.

3.*Si $\hat{\theta}$ est un estimateur de maximum de vraisemblance de θ, $f\left(\hat{\theta}\right)$ est l'estimateur de maximum de vraisemblance de $f(\theta)$.*

4.*Si la taille de l'échantillon est grande, l'EMV devient unique, et tend (en probabilité) vers la vraie valeur du paramètre θ, alors c'est un estimateur convergent*

$$\hat{\theta}_n \xrightarrow{p} \theta.$$

5.*L'EMV est asymptotiquement normale (**a.n**) et asymptotiquement efficace (pour $n \geq 30$).*

Exemple de l'EMV

Exemple 3.1.1 *(variable aléatoire de loi normale)*

Soient $X_1, X_2, \ldots\ldots X_n$, n-variables aléatoires de lois normales et indépendantes, $X_i \rightarrow \mathcal{N}(\mu, \sigma^2)$.

La fonction densité de probabilité est donnée par :

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

La vraisemblance de la loi normale est :

$$\begin{aligned}
L\left(x_1, ..., x_n; \mu, \sigma^2\right) &= \prod_{i=1}^{n} f(x_i; \theta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right\}, \\
\ln L\left(x_1, ..., x_n; \mu, \sigma^2\right) &= -n \ln\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2. \\
&= \frac{-n}{2}\ln\left(\sigma^2\right) - \frac{n}{2}\ln(2\pi) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2.
\end{aligned}$$

1)l'estimateur $\hat{\mu}$ de l'espérance μ est alors $\hat{\mu}$ tel que :

$$\begin{aligned}
\frac{\partial \ln L\left(x_1, ..., x_n; \mu, \sigma^2\right)}{\partial \mu} &= 0 \qquad (3.1) \\
&\Longleftrightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) = 0 \\
&\Longleftrightarrow \sum_{i=1}^{n}(x_i - \mu) = 0 \\
&\Longleftrightarrow \sum_{i=1}^{n} x_i - n\mu = 0.
\end{aligned}$$

La solution de l'équation (3.1) est :

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

On a

$$\frac{\partial^2 \ln L\left(x_1, ..., x_n; \mu, \sigma^2\right)}{\partial \mu^2} = \frac{-n}{\sigma^2}.$$

Alors

$$\frac{\partial^2 \ln L\left(x_1, ..., x_n; \mu, \sigma^2\right)}{\partial \mu^2} < 0.$$

car σ^2 et n sont toujours **positifs**, la vraisemblance atteint sont maximum au point $\hat{\mu}$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{X}$$

2)l'estimateur $\hat{\sigma}^2$ de la variance σ^2 vérifie :

$$\frac{\partial \ln L\left(x_1, ..., x_n; \mu, \sigma^2\right)}{\partial \sigma^2} = 0 \tag{3.2}$$

alors

$$\frac{\partial}{\partial \sigma^2}[\frac{-n}{2}\ln\left(\sigma^2\right) - \frac{n}{2}\ln\left(2\pi\right) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - \mu\right)^2] = 0.$$
$$\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n}\left(x_i - \mu\right)^2 = 0.$$

La solution de l'équation (3.2) est :

si μ connue

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \mu\right)^2.$$

si μ inconnue

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2. \tag{3.3}$$

3.1.3 Estimateurs obtenus par la méthodes du MV

Loi	Estimateur
$X \backsim U\left([0, b]\right)$	$\hat{b} = \sup\left(x_i\right)$
$X \backsim B\left(n, p\right)$	$\hat{p} = \frac{k}{n}$
$X \backsim P\left(\lambda\right)$	$\hat{\lambda} = \overline{X}$
$X \backsim N\left(\mu, \sigma^2\right)$	$\hat{\mu} = \overline{X}; \hat{\sigma}^2 = \frac{n}{n-1}S^2$
$X \backsim Exp\left(\lambda\right)$	$\hat{\lambda} = \frac{1}{\overline{X}}$

TAB. 3.1 – Estimateur obtenu par la méthode du MV

3.1.4 Méthode des moment (EMM) :

C'est la méthode la plus naturelle, que nous avons déjà utilisée sans la formalisée

Principe :

Cette méthode repose sur la propriété de convergence presque sûre des moments empiriques d'un échantillon iid. $(X_1, ..., X_n)$, extrait de X vers les moments théoriques correspondants de X .

Soit $\Theta \subset \mathbb{R}^p$, $\theta = (\theta_1, ..., \theta_p)$, nous noterons par $m_k(\theta)$, le moment théorique d'ordre k de X ; qu'il soit centré ou non, et par $m_k(e, n)$ le moment empirique d'ordre k .

$$\begin{cases} E(X) = m_1(\theta) = f_1(\theta_1, ..., \theta_p), \\ E(X^2) = m_2(\theta) = f_2(\theta_1, ..., \theta_p), \\ . \\ . \\ . \\ E(X^p) = m_p(\theta) = f_p(\theta_1, ..., \theta_p). \end{cases} \Longleftrightarrow \begin{cases} \theta_1 = g_1(m_1, ..., m_p), \\ \theta_2 = g_2(m_1, ..., m_p), \\ . \\ . \\ . \\ \theta_p = g_p(m_1, ..., m_p). \end{cases}$$

Définition 3.1.4 *On appelle estimateur de θ obtenu par la méthode des moments **(EMM)**, la solution θ_n^* du système :*

$$\begin{cases} \hat{m}_1(\theta) = m_1(e, \theta) = \frac{1}{n}\sum_{i=1}^{n} X_i, \\ \hat{m}_2(\theta) = m_2(e, \theta) = \frac{1}{n}\sum_{i=1}^{n} X_i^2, \\ . \\ . \\ . \\ \hat{m}_p(\theta) = m_p(e, \theta) = \frac{1}{n}\sum_{i=1}^{n} X_i^p. \end{cases} \Longleftrightarrow \begin{cases} \hat{\theta}_1 = g_1(\hat{m}_1, ..., \hat{m}_p), \\ \hat{\theta}_2 = g_2(\hat{m}_1, ..., \hat{m}_p), \\ . \\ . \\ . \\ \hat{\theta}_p = g_p(\hat{m}_1, ..., \hat{m}_p). \end{cases}$$

Remarque 3.1.2 *Si les fonction g_i sont continues en $(m_1, m_2, ..., m_p)$, alors les estimateurs obtenus sont des estimateurs convergents en presque sûre.*

3.1.5 Les propriétes de l'EMV

1. La méthode des moments fournit des estimateurs convergents et asymptotiquement sans biais.
2. La méthode des moments est conceptuellement plus simple que la méthode du maximum de vraisemblance.
3. La méthode des moments fournit des estimateurs peu précis lorsque n est modéré.

Exemples de l'EMM

Exemple 3.1.2 ***(loi géométrique)***

Soint X une v.a suit la loi géométrique$(G(p))$ telle que :

$$P(X=x)=p(1-p)^{x}$$

$$E(X)=\frac{1}{p};Var(X)=\frac{(1-p)}{p^{2}}.$$

En effet :

pour $k=1$on a

$$g_{1}(X)=X^{1}=X.$$

$$\overline{g_{1}(X)}=\frac{1}{n}\sum_{i=1}^{n}g_{1}(X_{i})=\frac{1}{n}\sum_{i=1}^{n}X_{i}=\overline{X}.$$

$$E(g_{1}(X))=E(X)=\frac{1}{p}$$

$$E(X)=m_{1}(p)=\frac{1}{p}=t.$$

$$m_{1}^{-1}(t)=\frac{1}{t}.$$

$$m_1^{-1}(\overline{g_1(X)}) = m_1^{-1}(\overline{X}) = \frac{1}{\overline{\overline{X}}} = \widehat{p}.$$

Exemple 3.1.3 ***(loi normale)***

Soit X variable aléatoire de loi normale de paramètres μ et σ^2

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}.\text{avec } \mu \in \mathbb{R} \text{ et } \sigma \in \mathbb{R}$$
$$E(X) = \mu; Var(X) = \sigma^2.$$

en effet

1)pour $k = 1$ on a

d'aprés l'exemple précédent on a

$$m_1^{-1}(t) = t.$$
$$m_1^{-1}(\overline{g_1(X)}) = m_1^{-1}(\overline{X.}) = \overline{X} = \widehat{\mu}.$$

2)pour $k = 2$ on a

$$g_2(X) = X^2,$$
$$\overline{g_2(X)} = \frac{1}{n}\sum_{i=1}^{n} g_2(X_i) = \frac{1}{n}\sum_{i=1}^{n} X_i^2,$$
$$E(g_2(X)) = E(X^2) = Var(X) + [E(X)]^2 = \sigma^2 - \overline{X}^2,$$
$$E(X^2) = m_2(\sigma^2) = \sigma^2 + \overline{X}^2 = t,$$
$$m_2^{-1}(t) = t - \overline{X}^2,$$
$$m_2^{-1}(\overline{g_2(X)}) = m_2^{-1}(\overline{X^2.}) = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \overline{X}^2 = \hat{\sigma}^2.$$

donc

$$\hat{\sigma}^2 = S^2.$$

3.1.6 Estimateurs obtenus par la méthodes de moment

Pour les lois de probabilité les plus courantes, les estimateurs obtenus par la méthode des moment sont donnés ci-dessous :

Loi	Estimateur
$X \backsim B(n,p)$	$\hat{p} = \frac{\overline{X}}{n}$
$X \backsim G(p)$	$\hat{p} = \frac{1}{\overline{X}}$
$X \backsim P(\lambda)$	$\hat{\lambda} = \overline{X}$
$X \backsim U([a,b])$	$\hat{a} = \overline{X} - \sqrt{3\frac{n-1}{n}S^2}; \hat{b} = \overline{X} + \sqrt{3\frac{n-1}{n}S^2}$
$X \backsim U([0,b])$	$\hat{b} = 2\overline{X}$
$X \backsim Exp(\lambda)$	$\hat{\lambda} = \frac{1}{\overline{X}}$
$X \backsim N(0,\sigma^2)$	$\hat{\sigma}^2 = \hat{E}(X^2)$
$X \backsim N(\mu,\sigma^2)$	$\hat{\mu} = \overline{X}; \hat{\sigma}^2 = \frac{n}{n-1}S^2$

Tab. 3.2 – Estimateur obtenu par la méthode du moment

Remarque 3.1.3 *Les deux méthodes d'estimation n'est pas nécéssarement donner le même estimateur. Par exemple, pour le paramètre $\theta > 0$ de la loi uniforme sur l'intervalle $[0,\theta]$, l'estimateur des moments est $2\overline{X_n}$ alors que celui du MV est $\max\limits_{1\leq i\leq n} X_i$.*

Jusqu'à maintenant nous avons déterminé les estimateurs ponctuels c'est-à-dire l'estimateur **EMV** et l'estimateur par la méthode des **moments**. Nous allons traiter dans cette section la méthode **d'estimation par intervalle de confiance**. Cette méthode est fréquemment utilisée et cela dans plusieurs domaines : la biologie, l'épidémiologie, la finance ...

Nous présentons ici les intervalles de confiances les plus utilisés.

Etant certain que $\hat{\theta}_n \neq \theta$, il est plus judicieux de choisir un intervalle pour θ en fonction de $\hat{\theta}_n$.

3.2 Estimation par intervalle de confiance

Nous avons vu à la section précédente comment estimer une valeur inconnue, c'est-à-dire comment proposer une valeur plausible pour cette grandeur inconnue. Mais nous commettons nécessairement une erreur : l'aléatoire fait que nous ne donnons pas exactement la valeur théorique, mais une valeur approchée. Le but est donc maintenant de donner cette marge d'erreur. Plus précisément nous allons construire un intervalle (ou une fourchette) dans lequel la grandeur recherchée a une probabilitée forte de se trouver.

Définition 3.2.1 *Soit X une v.a dont la loi dépend d'un paramètre réel θ inconnu et $\alpha \in [0,1]$ un nombre donné. On appelle* ***«intervalle de confiance »*** *pour le paramètre θ, de niveau de confiance $1-\alpha$, un intervalle qui a la probabilité $1-\alpha$ de contenir la vraie valeur du paramètre θ (généralement $\alpha = 0.01, \alpha = 0.05, \alpha = 0.10$).*

Remarque 3.2.1

1. *Le nombre α représente le risque que la vraie valeur de θ n'appartienne pas à $IC_{1-\alpha}(\theta)$.C'est la probabilité de l'erreur qu'on commet en affirmant que $\theta \in IC_{1-\alpha}(\theta)$.*
2. *Le nombre $1-\alpha$ représente la confiance qu'on a en disant que $\theta \in IC_{1-\alpha}(\theta)$.*
3. *Il est clair que plus le risque est petit, plus l'IC est large, c'est à dire moins la précision est bonne.*
4. *Un bon IC est un intervalle dont les bornes $\theta_{\min}$ et $\theta_{\max}$ sont des v.a qui dépendent d'un estimateur perfomant $\hat{\theta}_n$ de θ.*

3.2.1 Intervalle de confiance pour la moyenne et la variance dans le cas d'un échantillon gaussien

Soit $(X_1, ..., X_n)$ un n-échantillon de v.a iid de loi normale $\mathcal{N}(\mu, \sigma^2)$. Nous avons vu précédemment comment estimer les paramètres μ et σ^2 à l'aide du théorème de Fisher (2.4.1), nous pouvons déterminer la précision des estimateurs que nous avons construits et nous en déduirons ainsi des intervalles de confiance.

Estimation de la moyenne μ

1.lorsque la variance σ^2 est connue Le proplème dans cette partie est : d'encadrer μ (la moyenne), c'est à dire on chercher μ_1 et μ_2 tellesque :

$$\mathbb{P}\left(\mu_1 \leq \mu \leq \mu_2\right) = 1 - \alpha.$$

Pour estimer μ, on utilise la moyenne empirique $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ qui suit la loi $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$
alors d'aprés le **TCL** (1.2) on a

$$\sqrt{n}\left(\frac{\overline{X_n} - \mu}{\sigma}\right) \to \mathcal{N}\left(0, 1\right). \tag{3.4}$$

Cas d'un intervalle bilatéral

En utilisant la relation (3.4), on peut écrire :

$$\mathbb{P}\left(z_{\alpha_1} \leq \sqrt{n}\left(\frac{\overline{X_n} - \mu}{\sigma}\right) \leq z_{1-\alpha_2}\right) = 1 - \alpha$$
$$\Leftrightarrow \mathbb{P}\left(\overline{X_n} - z_{1-\alpha_2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X_n} + z_{\alpha_1}\frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

où z_{α_1} et $z_{1-\alpha_2}$ sont des valeurs lue dans la table de la loi normale centrée réduite, telle que $z_{1-\frac{\alpha}{2}}$ est le fractile d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{N}\left(0, 1\right)$.

Donc

$$IC_{1-\alpha}\left(\mu\right) = \left[\overline{X_n} - z_{1-\alpha_2}\frac{\sigma}{\sqrt{n}}; \overline{X_n} + z_{\alpha_1}\frac{\sigma}{\sqrt{n}}\right].$$

Cas particulier : intervalle bilatérale symétrique

Intervalle bilatérale symétrique si $\alpha_1 = \alpha_2 = \alpha/2$

et comme $z_{1-\frac{\alpha}{2}} = -z_{\frac{\alpha}{2}}$ on a

$$\mathbb{P}\left(-z_{1-\frac{\alpha}{2}} \leq \sqrt{n}\left(\frac{\overline{X_n} - \mu}{\sigma}\right) \leq z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

Ceci équivaut à

$$\mathbb{P}\left(\overline{X_n} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \overline{X_n} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

On obtient donc un **IC** pour l'espérance μ avec coefficient de sécurité $1-\alpha$ dans le cas où σ est connu, il s'agit de l'intervalle aléatoire.

$$IC_{1-\alpha}(\mu) = \left[\overline{x_n} - z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}, \overline{x_n} + z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right] \tag{3.5}$$

Cas d'un intervalle unilatéral à droite

Dans ce cas : $\alpha_1 = 0$ et $\alpha_2 = \alpha$

$$\mathbb{P}\left(\sqrt{n}\left(\frac{\overline{X_n}-\mu}{\sigma}\right) \leq z_{1-\alpha}\right) = 1-\alpha,$$
$$\mathbb{P}\left(\mu \geq \overline{X_n} - z_{1-\alpha}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Cas d'un intervalle unilatéral à gauche

Dans ce cas : $\alpha_1 = \alpha$ et $\alpha_2 = 0$

$$\mathbb{P}\left(\sqrt{n}\left(\frac{\overline{X_n}-\mu}{\sigma}\right) \geq -z_{\alpha}\right) = 1-\alpha,$$
$$\mathbb{P}\left(\mu \leq \overline{X_n} + z_{\alpha}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

Remarque 3.2.2 *Si la taille de l'échantillon n est grande alors l'IC devient petit, par conséquent, pour avoir une bonne précision il faut travailler avec des échantillons de grandes tailles.*

2.lorsque la variance σ^2 est inconnue : Lorsque la variance σ^2 est inconnue, il est alors nécessaire de remplacer dans les formules précédentes cette quantité par la **variance empirique**, qui en est un estimateur convergent. Il faut donc considérer non plus la quantité $\sqrt{n}\left(\frac{\overline{X_n}-\mu}{\sigma}\right)$ mais plutôt

$$\sqrt{n}\left(\frac{\overline{X_n}-\mu}{S_n}\right) \backsim \mathcal{T}_{n-1} \tag{3.6}$$

Alors d'aprés la relation (3.6) on a :

$$\mathbb{P}\left(-t_{1-\frac{\alpha}{2}} \leq \sqrt{n}\left(\frac{\overline{X_n}-\mu}{S_n}\right) \leq t_{1-\frac{\alpha}{2}}\right)=1-\alpha$$

où $t_{1-\frac{\alpha}{2}}$ est le fractile d'ordre ${}_{1-\frac{\alpha}{2}}$ de la loi du student.

Ceci équivaut à

$$\mathbb{P}\left(\overline{X_n}-t_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}} \leq \mu \leq \overline{X_n}+t_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}}\right)=1-\alpha$$

On obtient donc un **IC** pour l'espérance μ avec coefficient de sécurité $1-\alpha$ dans le cas où σ^2 est inconnu, il s'agit de l'intervalle aléatoire

$$IC_{1-\alpha}(\mu)=\left[\overline{x_n}-t_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}}, \overline{x_n}+t_{1-\frac{\alpha}{2}}\frac{S_n}{\sqrt{n}}\right] \tag{3.7}$$

Remarque 3.2.3 *Lorsque n est grand, les quantiles de Student pouvent être remplacés par ceux de Gauss.*

Remarque 3.2.4 *Lorsque X suit loi quelconque (ou inconnue), les intervalles définis dans* (3.5) *et* (3.7) *deviennent asymptotiques pour des échantillon de grandes tailles.*

Estimation de la variance σ^2

Pour estimer la variance σ^2, on cherche deux valeurs σ_1^2 et σ_2^2 telle que : $\mathbb{P}\left(\sigma_1^2 \leq \sigma \leq \sigma_2^2\right)=1-\alpha$.

1.Si la moyenne μ inconnue La **variance empirique** S_n^2, définie par (2.2) représente l'estimateur naturelle de la variance σ^2,l'estimateur S_n^2 est obtenu par les méthodes du **MV** et des **moment** pour les lois de probabilités usuelles.

Alors la statistique qui conduit à la construction de l'intervalle de confiance de la variance au niveau $1-\alpha$ est $\frac{n-1}{\sigma^2}S_n^2 \backsim \chi^2_{1-\alpha}(n-1)$.

De plus, on lit dans des tables les quantiles d'ordre $\alpha/2$ et $1-\alpha/2$ de la loi du $\chi^2_{(n-1)}$,respectivement notés $\upsilon_{\alpha/2}$ et $\upsilon_{1-\alpha/2}$.

On obtient alors

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

Ceci équivaut à

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

On obtient donc un **IC** pour la variance σ^2 avec coefficient de sécurité $1-\alpha$:

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

2.Si la moyenne μ connue On a la statistique

$$R_n^2 := \frac{1}{n}\sum_{i=1}^{n}\left(X_i-\mu\right)^2.$$

constitue un estimateur de σ^2plus précis que S_n^2.

De plus

$$\frac{n}{\sigma^2} R_n^2 \backsim \chi_{1-\alpha}^2\left(n\right).$$

$$\begin{aligned} &\mathbb{P}\left(\omega_{\alpha/2} \leq \frac{n}{\sigma^2} R_n^2 \leq \omega_{1-\alpha/2}\right) = 1-\alpha \\ &\qquad\qquad \Leftrightarrow \mathbb{P}\left(\frac{nR_n^2}{\omega_{1-\alpha/2}} \leq \sigma^2 \leq \frac{nR_n^2}{\omega_{\alpha/2}}\right) = 1-\alpha. \end{aligned}$$

On obtient finalement l'intervalle de confiance pour la variance au niveau $1-\alpha$:

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

$\omega_{\alpha/2}$ et $\omega_{1-\alpha/2}$ étant les fractiles d'ordre respectifs $\alpha/2$ et $1-\alpha/2$ de la loi du khi-deux à n degré de liberté.

Remarque 3.2.5 *Soient $X_1, X_2, ..., X_n$, n variables aléatoires de lois normales et indépendantes :*

$X_i \backsim \mathcal{N}(\mu, \sigma^2)$ *et* $Z_i = \frac{X_i - \mu}{\sigma} \backsim \mathcal{N}(0,1)$ *et on définit* $\sum_{i=1}^{n} Z_i^2 \backsim \chi^2(n)$

alors

$$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \backsim \chi^2(n)$$

$$\Longleftrightarrow \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2 \backsim \chi^2(n)$$

$$\Leftrightarrow \frac{n}{\sigma^2}\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2 \backsim \chi^2(n)$$

$$\Longleftrightarrow \frac{n}{\sigma^2}R_n^2 \backsim \chi^2(n).$$

avec

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

3.2.2 Intervalle de confiance pour la proportion

Pour construire l'intervalle de confiance d'une proportion p (inconnue) des individus possédant un certain caractère appartenant à une population infinie (ou finie si le tirage s'effectue avec remise), on utilise F_n, l'estimateur naturel de P, $F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{(X_i \leq x)}$ appellé la **fréquence empirique,** calculée dans un échantillon de taille $n \geq 1$.

1. Si n est faible on utilise les tables de la **loi binomiale** puisque : $nF_n \backsim B(n,p)$.
2. Si $n \geq 30$ on a : $F_n \backsim \mathcal{N}\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$

où $F_n(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{(X_i \leq x)}$ avec $X_i \backsim B(p), \forall i = 1, ..., n$ et sont indépendantes.

$$E(F_n) = p$$
$$Var(F_n) = \frac{p(1-p)}{n}.$$

alors d'aprés le **TCL** on a : $\frac{F_n - p}{\sqrt{\frac{p(1-p)}{n}}} \backsim \mathcal{N}(0,1)$

On retient alors un intervalle symétrique :

$$\mathbb{P}\left(-z_{1-\alpha/2} \leq \frac{F_n - p}{\sqrt{\frac{p(1-p)}{n}}} \leq z_{1-\alpha/2}\right) = 1-\alpha$$

$$\Leftrightarrow \mathbb{P}\left(F_n - z_{1-\alpha/2}\sqrt{\frac{p\,(1-p)}{n}} \leq p \leq F_n + z_{1-\alpha/2}\sqrt{\frac{p\,(1-p)}{n}}\right) = 1-\alpha$$

où la valeur $z_{1-\alpha/2}$ lue dans la table de la loi normale centrée réduite.

Alors l'IC de la proportion p est donné comme :

$$IC_{1-\alpha}\,(p) = \left[F_n \underset{-}{+} z_{1-\alpha/2}\sqrt{\frac{F_n\,(1-f_n)}{n}}\right].$$

3.2.3 Exemples de l'IC

Dans ce partie, on applique les résultats de la section précédent sur des données.

Les codes sont présentées dans l'annexe A.

Illustration numérique

Paramètres d'une population normale

On tire un échantillon da taille $n = 500$ d'une population normale standard. Nous comptons sur le niveau de confiance $1-\alpha = 0.975; 0.90; 0.95$. Les résulats sont regroupés dans le tableau suivant :

$1-\alpha$	b_inf	$\hat{\mu}$	b-sup	b_inf	$\hat{\sigma}$	b-sup
0.975	-0.06239912	0.03435849	0.13111611	0.729	1.073375	1.403
0.90	-0.03656449	0.03435849	0.10528148	0.826	1.073375	1.339
0.95	-0.05019966	0.03435849	0.11891664	0.770	1.073375	1.381

TAB. 3.3 – Intervalle de confiance, de différents niveau pour la moyenne et la variance d'une population normale standard basés sur 500 observations

Proportion de succès

On répète, 500 fois de façon indépendante, une expérience de Bernoulli de probabilté de succès $p = 0.5$.Avec les même niveaux de confiance.

Alors on note que les intervalles de confiance dans le tableau suivant :

$1-\alpha$	b_inf	$\hat{p}$	b-sup
0.975	0.2429479	0.282	0.3236520
0.90	0.3567614	0.4	0.4444282
0.95	0.5555718	0.6	0.6432386

TAB. 3.4 – Intervalle de confiance, de différents niveaux pour la proportion de succès lors de 500 expériences de Bernoulli de paramètre p=0.5

Illustration graphique

Paramètres d'un population normale

On prélève des échantillon de taille $n = 500$, d'une population normale standard et on fixe le niveau de confiance 95%. Les intervalles de confiance relatifs à la moyenne et à la variance sont illustrés par la figure 3.1.

Proportion de succès

Des échantillon de taille $n = 500$, d'une v.a de Bernoulli de paramètre $p = 0.5$.

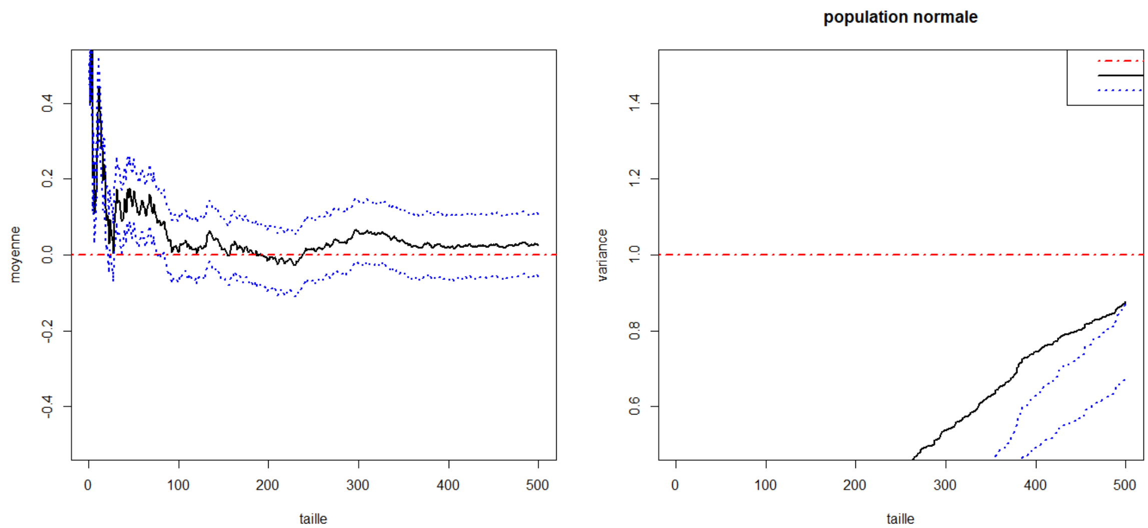

FIG. 3.1 – Intervalle de confiance de niveau 95%, pour la moyenne empirique(paneau gauche) et variance empirique (paneau droit) d'une population normale standard.

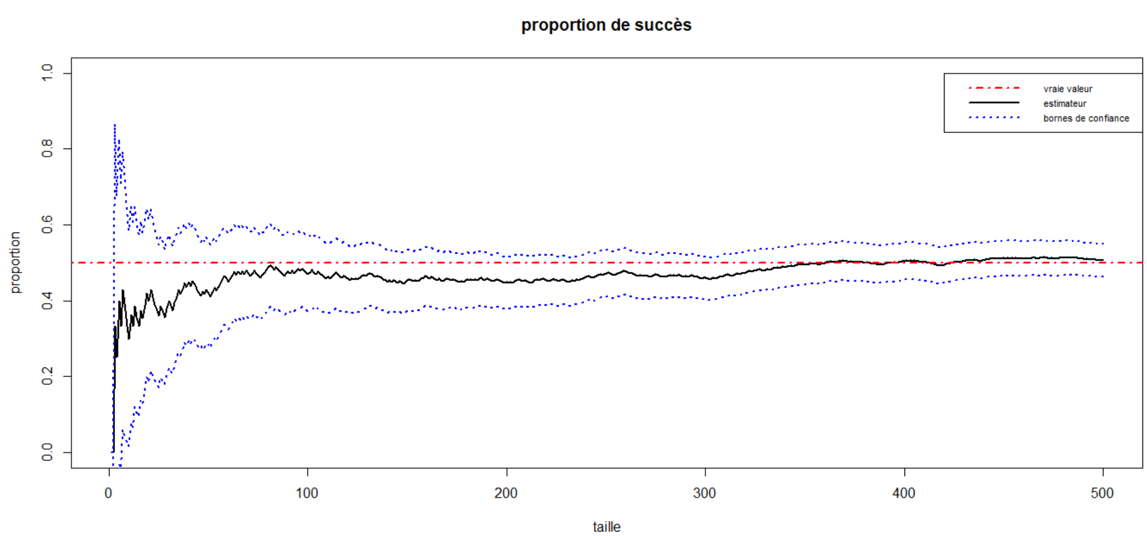

FIG. 3.2 – Intervalle de confiance, de niveau 95%, pour la proportion de succés lors d'expériences de Bernoulli de paramètre 0.5.

Conclusion

Dans ce travail, nous avons étudié les méthodes d'estimation paramétrique (ponctuelle et par intervalle de confiance). Elles sont les plus utilisées en statistiques.

Les estimateurs obtenues par ces méthodes possèdent de très bonnes propriétés, notamment : la convergence, l'efficacité, la normalité asymptotiquement, des illustrations graphiques et des tableaux de résultats nemuriques sont présentés à la fin de ce mémoire.

Bibliographie

[1] BOUZIANE, N. (2005). Efficacité-Robustesse des estimateurs des modèles paramétriques (Doctoral dissertation, UNIVERSITE DE MOHAMED KHIDER BISKRA).

[2] De Micheaux, P. L., Drouilhet, R., & Liquet, B. (2014). Le logiciel R : maîtriser le langage, effectuer des analyses (bio) statistiques. Springer.

[3] Dusart, P. (2015). Cours de statistiques inférentielle.

[4] Jean-Yves DAUXOIS(2011-2012). CTU, Master Enseignement des Mathématiques Statistique Inférentielle.

[5] Hachani Nadia(2019), Estimation du maximum de vraisemblance. Mémoire.

[6] HERVé, M. (2014). Aide-mémoire de statistique appliquée à la biologie. Constr Son étude Anal Résultats à l'aide du logiciel R. Version, 5.

[7] GAGUI Abdelmalek(2016), Théorie d'estimation : Types d'estimateurs et Applications. Mémoire.

[8] Lipshutz, S. Probabilités (cours et problèmes), série Schaum.

[9] Meraghni Brahim(2019), Estimation par intervalle de confiance. Mémoire.

[10] Nicolas Marie(2018), Introduction à la modélisation probabiliste et statistique

[11] Perrut, A. Cours de probabilités et statistiques.

[12] Saporta, G. (2006). Probabilités, analyse des données et statistique. Editions Technip.

[13] Velenik, Y. (2011). Probabilités et statistique. Université de Geneve.

Annexe A : Codes *R*

La simulation est une partie pratique dans certain logiciel, dans ce dernier chapitre nous allons présonte quelque propriétés de certain loi usuelle (moyenne empirique, et variance empirique), et les déffirentes méthodes d'estimation paramétrique de deuxième chapitre à l'aide du logicile d'analyse statistique **R 3.6.1.**

3.3 Ilustration numérique

3.3.1 Paramètres d'une population normale

L'intervalle de confiance de la moyenne

> n<-500 ;

> x<-rnorm(n)

>

One Sample t-test

data : x

t = 0.49001, df = 199, p-value = 0.6247

alternative hypothesis : true mean is not equal to 0

90 percent confidence interval :

-0.03656449 0.10528148

sample estimates :

```
mean of x

0.03435849
```

L'intervalle de confiance de la variance

```
> n<-500;

> x<-rnorm(n)

> require("boot")

> variance<-function(x,indice) var(x[indice])

> x.boot<-boot(x,variance,R=999,stype="i",sim="ordinary")

> boot.ci(x.boot,conf=0.95,type=c("norme","basic","perc","bca")

> boot.ci(x.boot,conf=0.95,type=c("norme","basic","perc","bca"))
```

3.3.2 Proportion de succès

```
> n<-500;

> binom.test(141,n)

Exact binomial test

data : 141 and n

number of successes = 141, number of trials = 500, p-value < 2.2e-16

alternative hypothesis : true probability of success is not equal to 0.5

95 percent confidence interval :

0.2429479 0.3236520

sample estimates :

probability of success

0.282
```

3.4 Ilustration graphique

3.4.1 La Moyenne et Variance empirique d’observations uniformes et Gaussiennes standard $(figure\,(2.1))$

```
> n<-500 ;N<-seq(1,n,by=1) ;X1<-runif(n) ;X1bar<-cumsum(X1)/N ;
> S1<-cumsum((X1-X1bar)^2)
> S2<-cumsum((X2-X2bar)^2)
> op<-par(mfrow=c(2,2))
> plot(N,X1bar,ylim=c(0,1),main="moyenne empirique d’observation 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’observation normales",xlab="taille de l’échantillon",ylab="Moyenne empirique") ;abline(h=0,col="red")
> plot(N,S1,ylim=c(0,1),main="variance empirique d’observation uniformes",xlab="taille de l’échantillon",ylab="Variance empirique") ;abline(h=0,col="red")
> plot(N,S2,ylim=c(-1,1),main="variance empirique d’observation normales",xlab="taille de l’échantillon",ylab="Variance empirique") ;abline(h=0,col="red")
> par(op)
```

3.4.2 L’intervalle de confiance

Paramètres d’une population normale $(figure\,(3.1))$

```
> n=500 ;x=rnorm(n) ;mu=(cumsum(x))/(1 :n) ;ss=cumsum((x-mu)^2) ;
> s=c(ss[1],ss[2 :n]/(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,lty=3,col="blue",lwd=2)

> lines(icm2,lty=3,col="blue",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,lty=3,col="blue",lwd=2)

> lines(icv2,lty=3,col="blue",lwd=2)

> 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"

Proportion de succès($figure$ (3.2))

> 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",lwd=2,ylab="proportion",xlab="taille",ylim=c(0,1))

> abline(h=0.5,lty=4,col="red",lwd=2)

> lines(ic1,lty=3,col="blue",lwd=2)

> lines(ic2,lty=3,col="blue",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","b

Annexe B : Abréviations et Notations

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

Notation	Signification
Θ	Ensemble des paramètres
$E(X)$	Espérance matimatique du v.a X
$Var(X)$	Variance du X
$\mathbb{P}_\theta$	Loi de paramètre θ
$F(x;\theta)$	Fonction de répartition
$f(x;\theta)$	Fonction densité
φ_X	Fonction caractéristique
G_X	Fonction génératrice
$\overline{X_n}$	Moyenne empirique
S_n^2	Variance empirique
v.a	Variable aléatoire
EMM	Estimateur de la méthode des moments
EMV	Estimateur de la méthode du maximum de vraisemblance
a.n	Asymptotiquement normale
a.s.b	Asymptotiquement sans biais
TCL	Théorème central limite
IC	Intervalle de confiance

$\xrightarrow{\mathcal{L}}$	Convergence en loi
$\xrightarrow{L^p}$	Convergence au moyenne d'ordre p
$\xrightarrow{p.s}$	Convergence presque sûre