République Algériennne 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

Départementde Mathématiques

Mémoire présenté pour obtenir le diplôme de

Master en "**Mathematiques Appliquées**"

Option :Statistique

Par : Herzallah Nour Elhouda

Titre :

Estimation paramétrique et applications

Devant le Jury :

Dr.	Abdelli Jihane	U. Biskra	Président
Dr.	Roubi Affef	U. Biskra	Encadreur
Dr.	Khereddine Souraya	U. Biskra	Examinatrice

Soutenu Publiquement le 27/06/2022

Dédicace

Je dédié ce modeste travail :

À mes chers parents qui ont toujours été présent pour moi,

Que le grand Dieu me les protèges,

À mes chers frères et sœurs, ma source de joie et de bonheur,

À mes oncles, mes tantes, mes cousins et cousines,

À tous mes amis et camarades,

À tous mes professeurs et tous ceux qui ont contribué à mon éducation

À tous ceux qui me sont chers,

À tous ceux qui m'aiment,

Et a tous ceux que j'aime.

Remerciements

Tout d'abord je tiens à remercier Dieu de m'avoir donné le courage, la volonté et la santé

pour mener à bien ce travail.

Je remercie ma mère, mon père mercie pour vos efforts, votre aide et vos sacrifices et mes chères sœurs et frères Madjda, Zahra, Habiba, Khaoula, Amar, Ahmed et Islam pour leur soutien et leur encouragement.

Je voudrais remercier également les membres de jury de m'avoir fait l'honneur en acceptant de juger ce travail.

Merci également a la doctorante : ROUBI AFFEF pour son encouragement, ses conseils et sa précieuse aide.

Enfin, je tiens à exprimer ma profonde gratitude à mes amis et toutes les personnes qui

m'ont apportée une aide pour la réalisation de ce travail.

Notations et symbols

$(X_1, ..., X_n)$	:	échantillon de taille n.
F	:	fonction de répartition.
f	:	densité de probabilité.
$v.a.d$	:	variable aléatoire discrète.
$v.a.c$	:	variable aléatoire continue.
$c-\grave{a}-d$	:	c'est à dire.
$i.i.d$	:	indépendantes et identiquement distribuées.
$E(.)$	:	espérance mathématique.
$Var(.)$	:	variance.
$\overset{loi}{\rightarrow}$	:	convergence en loi.
$\overset{p}{\rightarrow}$	:	convergence en probabilité.
$\overset{p.s}{\rightarrow}$	:	convergence presque sûre.
$MSE(.)$	:	erreur quadratique moyenne.
MV	:	maximum de vraisemblance.
$\hat{\theta}$	:	estimateur de θ.
$\mathcal{N}(\mu, \sigma^2)$	:	loi normale de paramètre μ et σ^2.
$\mathcal{N}(0, 1)$	:	loi normale centrée réduite.
$\mathcal{X}_n^2$	:	loi de khi-deux de n degrés de liberté.

t_n : loi de Student de n degrés de liberté.

X : la moyenne empirique.

$\tilde{S}^2$: variance empirique.

$\tilde{S}^*$: variance empirique corrigée.

$IC_{1-\alpha}$: intervalle de confiance (au niveau de confiance $1-\alpha$).

$I_n(.)$: information de Fisher.

Table des matières

Table des figures

Liste des tableaux

Introduction

L'estimation statistique est un domaine très important de la statistique mathématique est divisée en deux composantes principaux, à savoir l'estimation paramétrique et non paramétrique. L'approche paramétrique qui considère que les modèles sont connues. La loi de la variable étudiée est supposée appartenir à une famille de lois peuvent être caractérisées par une forme fonctionnelle connue, le plus souvent la famille paramétrique à laquelle la loi de X est réputée appartenir sera d'écrite par la famille de densités de probabilité (respectivement de fonctions de probabilité) $\{f(x;\theta);\theta \in \Theta\}$. Par opposition en l'approche non paramétrique, le modèle n'est pas décrit par un nombre fini de paramètres et estime la densité directement à partir de l'information disponible sur l'ensemble d'observations.

Dans ce travail, on s'intéresse à l'estimation paramétrique et plus particulièrement à l'estimation par différentes méthodes notamment : la méthode des moments et la méthode du maximum de vraisemblance et la méthode des intervalles de confiance puis établir et étudier leurs propriétés et essayez trouver quel méthode est plus effective que l'autre. Ce mémoire est composé de deux chapitres.

Le premier chapitre **Généralités sur l'estimation paramétrique :** ce chapitre est consacré aux quelques notions et définitions de base en statistique, ensuite nous présentons les types des convergences. Dans la deuxième partie du chapitre on parle sur les estimateurs et leurs propriétés.

Le deuxième chapitre **Méthodes d'estimation** : dans ce chapitre, on présente deux méthodes d'estimation ponctuelles : la méthode des moments et celle du maximum de vraisemblance puis on présente la méthode d'estimation par intervalle de confiance.

Chapitre 1

Généralités sur l'estimation paramétrique

1.1 Notions et généralités

Ce chapitre est consacré à un rappel des notions de base de la statistique mathématique comme : l'estimateur et leurs propriétés, ensuite, nous étudions quelques théorèmes et types de convergences de variables aléatoires.

1.1.1 Fonction de répartition

Définition 1.1.1 *Soit X une variable aléatoire, on appelle* ***fonction de répartition*** *de X, que l'on note F, la fonction définie sur $\mathbb{R}$ par*

$$\forall\ x \in \mathbb{R},\ \ F\,(x) = P(X \leq x) = P(]-\infty, x]).$$

Remarque 1.1.1 *Une fonction de répartition doit vérifier un certain nombre de* *propriétés suivantes*

$$\forall\, x \in \mathbb{R}, \qquad 0 \leq F(x) \leq 1.$$

- *Si* $x_1 < x_2$, *alors* $F(x_1) \leq F(x_2)$, *c'est à dire* F *est une fonction croissante·*

- *Pour tout nombre* a, $b \in \mathbb{R}$ *tels que* $a < b$

$$P(a \leq X \leq b) = F(b) - F(a).$$

- $\forall a \in \mathbb{R}$, *on a*

$$\forall a \in \mathbb{R}, P(x > a) = 1 - P(X \leq a) = 1 - F(a).$$

-F *est continue à droite en tout point* x *de* $\mathbb{R}$

- *La limite de* $F(x)$ *quand* x *tend vers* $+\infty$ *égale à 1,* $\left(\lim\limits_{x \to +\infty} F(x) = 1\right)$.

- *La limite de* $F(x)$ *quand* x *tend vers* $-\infty$ *égale à 0,* $\left(\lim\limits_{x \to -\infty} F(x) = 0\right)$.

1.1.2 Densité de probabilité

Définition 1.1.2 *Une fonction* $f : \mathbb{R} \longrightarrow \mathbb{R}$ *est* ***appelée densité*** *de probabilité, si elle est positive (en tout* $x \in \mathbb{R}$ *où elle est définie,* $f(x) \geq 0$ *), intégrable sur* $\mathbb{R}$ *et si*

$$\int_{\mathbb{R}} f(x)dx = 1.$$

– Pour tous ensemble $\mathcal{B} \subset \mathbb{R}$, on a alors

$$P(X \in \mathcal{B}) = \int_{\mathcal{B}} f(x)dx.$$

– Lorsque $\mathcal{B}$ est un intervalle de la forme $\mathcal{B} = [a,b]$, la probabilité

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

1.1.3 Variables aléatoires discrètes et continues

Une v.a X est dite discrète $(v.a.d)$ ssi elle est à valeurs dans un ensemble E fini ou dénombrable on note $E = \{x_1, x_2, ..., x_n\}$, la loi de probabilité d'une $(v.a.d)$ est

$$F(x) = P(X \leq x) = \sum_{x_i \in E} P(X = x_i).$$

Une v.a X est dite continue $(v.a.c)$ ssi sa fonction de répartition est continue et presque partout dérivable, la loi de probabilité d'une $v.a.c$ est

$$P(X \leq x) = \int_{-\infty}^{x} f(x)\,dx.$$

De plus

$$\forall\ \mathcal{B} \in \mathbb{R},\ P(X \in \mathcal{B}) = \int_{-\infty}^{x} f(x)\,dx,$$

où $f(x)$ est une densité de probabilité de X.

1.1.4 Espérance mathématique

Pour une variable aléatoire discrète, on définit l'espérance $E(X)$ par la formule

$$E(X) = \sum_{x_i \in E} x_i P(X = x_i).$$

(si cette expression a un sens), $E(X)$ est la moyenne arithmétique des différentes valeurs de X pondérées par leurs probabilités.

Pour une variable continue, on définit l'espérance $E(X)$ par la formule

$$E(X) = \int_{\mathbb{R}} x f_X(x) \; dx.$$

Définition 1.1.3 (*Le moment d'ordre k*)

Le moment absolu d'ordre k par rapport à un point a, est égal à, sous réserve de l'existence de l'intégrale

$$E(X-a)^k = \int_{\mathbb{R}} (x-a)^k f(x) \; dx.$$

1.1.5 Variance et écart-type

On appelle variance de X notée $V(X)$ ou σ^2 la quantité définie par

$$\begin{aligned} \sigma^2 &= E(X-\mu)^2 \\ &= \int_{\mathbb{R}} (x-\mu)^2 \, dP_X(x), \end{aligned}$$

où $\mu = E(X)$.

La variance est donc le moment centré d'ordre 2 de la distribution et est une mesure de la dispersion de X autour de μ.

1.1.6 Convergences des suites de variables aléatoires

Dans cette section on va rappeler les notions de convergence des suites des variables aléatoires, aussi deux théorèmes fondamentaux en statistique on va les données.

Soit $\{X_n\}_{n\geq 1}$ une suite de $v.a.$

Convergence en loi

On dit que $\{X_n\}_{n\geq 1}$ converge en loi vers une $v.a.X$, et on écrit $X_n \overset{loi}{\rightarrow} X$, si $\lim\limits_{n\to\infty} F_n \to F$, en tout point de continuité de F.

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

Convergence presque sûre

On dit que $\{X_n\}_{n\geq 1}$ converge presque sûrement vers la $v.a.$ réelle X si on a

$$P\left(\lim_{n\to\infty} X_n \neq X\right) = 0. \tag{1.1}$$

Convergence en probabilité

On dit que $\{X_n\}_{n\geq 1}$converge en probabilité vers une $v.a$

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

Convergence en moyenne quadratique

Soit $\{X_n\}_{n\geq 1}$ une suite de $v.a.$ réelles de moyennes et de variances finies. On dit que la suite X_n converge en moyenne quadratique vers $\{X_n\}$ si

$$E\left((X_n - X)^2\right) \to 0, \text{ quand } (n \to +\infty)\cdot \tag{1.2}$$

Remarque 1.1.2 - *Les relations (1·1) , (1·2) ,(1·3) et (1·4) restent valables si on remplace la v.a. X par une constante réelle a.*

- Les implications suivantes permettent le passage entre certains types de conver-

gence.

$$X_n \overset{p.s}{\to} X \Longrightarrow X_n \overset{p}{\to} X.$$
$$X_n \overset{p}{\to} X \Longrightarrow X_n \overset{loi}{\to} X.$$

1.1.7 Théorème central limite et loi des grands nombres

L'étude de sommes de variables indépendantes et de même loi joue un rôle capital en statistique. Le théorème suivant connu sous le nom de théorème central-limite (il vaudrait mieux dire théorème de la limite centrée) établit la convergence vers la loi de Gauss sous des hypothèses peu contraignantes.

Théor■me 1.1.1 ***(Central limite)*** *Soit* $\{X_n\}_{n\geq 1}$*une suite de variables aléatoires indépendantes de même loi d'espérance* μ *et d'écart-type* σ*. Alors*

$$\frac{1}{\sqrt{n}}\left(\frac{X_1 + X_2 + ...X_n - n\mu}{\sigma}\right) \overset{loi}{\to} \mathcal{N}(0,1), \text{ quand } (n \to \infty).$$

Preuve. Voir ([6]) ■

Théor■me 1.1.2 ***(Loi des grands nombres)*** *Soit* $\{X_n\}_{n\geq 1}$ *une suite des v.a (iid), supposons que* μ *et* σ^2 *finies, notons* $\bar{X}_n = \frac{1}{n}\sum_i^n X_i$*, alors*

$$\bar{X}_n \overset{p.s}{\to} X, \quad \text{quand } (n \to \infty).$$
$$\bar{X}_n \overset{p}{\to} X, \text{ quand } (n \to \infty).$$

Preuve. Voir ([6]) ■

Remarque 1.1.3 $\bar{X}_n$ *converge presque sûrement (loi forte) et en probabilité (loi faible) vers* $E(X)$*, cela signifie que quand on fait un très grand nombre d'expé-*

riences identiques et indépendantes, la moyenne des réalisations de la v.a tend vers l'espérance de sa loi.

Définition 1.1.4 *On appelle modèle statistique, la donnée d'un espace des observations E, d'une tribu ε d'évènements sur E et d'une famille de probabilités P sur l'espace probabilisable (E, ε). On le note (E, ε, P).*

Définition 1.1.5 *Un modèle paramétrique est un modèle où l'on suppose le type de loi de X est connu, mais qu'il dépend d'un paramètre θ inconnu, de dimension n. Alors, la famille de lois de probabilité possibles pour X peut s'écrire P=$\{P_\theta,\ \theta \in \Theta \subset \mathbb{R}^n\}$.*

– Le modèle gaussien $\{N(\mu, \sigma), \mu \in \mathbb{R}, \sigma^2 > 0\}$.
– Le modèle de Poisson$\{P(\lambda),\ \lambda > 0\}$.
– Le modèle exponentiel $\{\varepsilon(\lambda),\ \lambda > 0\}$.

Exemple 1.1.1 *Soit X une v.a qui suit la loi de Poisson $\mathcal{P}(\theta)$, soit $(X_1, X_2, ..., X_n)$ un échantillon, et soit la statistique*

$$\hat{\theta} = \sum_{i=1}^{n} X_i$$

On montre que la statistique $\hat{\theta}$ est une statistique exhaustive pour θ.

On a

$$\mathbb{P}(X_1, X_2, ..., X_n/\hat{\theta} = t) = \frac{\mathbb{P}(X_1, X_2, ..., X_n; \hat{\theta} = t)}{\mathbb{P}(\hat{\theta} = t)}.$$

La statistique $\hat{\theta}$ suit la $\mathcal{P}(n\theta)$, alors

$$\mathbb{P}\left(\hat{\theta} = t\right) = \frac{e^{(-n\theta)}\,(n\theta)^t}{t!},$$

et

$$\mathbb{P}(X_1, X_2, ..., X_n, \hat{\theta} = t) = \mathbb{P}(X_1, X_2, ..., X_{n-1}, \hat{\theta} = \sum_{i=1}^{n-1} = X_i),$$

$$\mathbb{P}(X_1, X_2, ..., X_{n-1}, \hat{\theta} = \textstyle\sum_{i=1}^{n-1} = X_i) \quad = \bigcap_{i=1}^{n-1} \left(\frac{e^{-\theta}\theta^{x_i}}{(x_i)!}\right) \frac{e^{-\theta}\theta^{t-\sum_{i=1}^{n-1} x_i}}{\left(t-\sum_{i=1}^{n-1} x_i\right)!}$$
$$= \frac{e^{-n\theta}\theta^t}{(x_1)!(x_2)!...(x_{n-1})!\left(t-\sum_{i=1}^{n-1} x_i\right)!},$$

d'où

$$\mathbb{P}(X_1, X_2, ..., X_n / \hat{\theta} = t) = \frac{t!}{n^t\,(x_1)!\,(x_2)!...\,(x_{n-1})!\,\left(t - \sum_{i=1}^{n-1} x_i\right)!}.$$

La probabilité conditionnelle ne dépend pas de θ, donc $\hat{\theta}$ est une statistique exhaustive pour θ.

1.2 Estimation

L'estimation consiste à donner des valeurs approchées aux paramètres d'un modèle ou plus précisement à évaluer ses paramètres à l'aide d'un échantillon de taille n.

1.2.1 Définition d'un estimateur

Soit X une variable aléatoire dont la loi de probabilité $f(x, \theta)$ dépend d'un seul paramètre θ. Soit $(X_1, ..., X_n)$ un échantillon de taille n de cette variable (variable parente).

Définition 1.2.1 ***(Statistique)*** *On appelle statistique sur un échantillon $(X_1, ..., X_n)$ toute fonction mesurable des X_i, $i = 1, ..., n$.*

Définition 1.2.2 ***(Estimateur)*** *Un estimateur du paramètre θ est une statistique c-à-d, une fonction des X_i $(i = 1, ..., n)$ à valeur dans un domaine acceptable* *pour θ on le note généralement par T_n, $\hat{\theta}_n$ ou $\hat{\theta}$.*

1.2.2 Propriétés d'un estimateur

Soit T_n l'estimateur de θ.

L'erreur d'estimation est mesurée par la quantité $T - \theta$ qui peut s'écrire

$$T - \theta = T - E(T) + E(T) - \theta,$$

où

• $T - E(T)$ représente les fluctuations de l'estimateur T autour de sa valeur moyenne $E(T)$ (espérance mathématique).

• $E(T) - \theta$ est une erreur systématique car l'estimateur T varie autour de son espérance mathématique $E(T)$ et non autour de la valeur θ du paramètre sauf si $E(T) = \theta$.

La quantité $b(\theta) := E(T) - \theta$ est le biais de l'estimateur.

Alors, on peut écrire les deux définitions suivantes.

Estimateur avec biais

Définition 1.2.3 *Un estimateur T_n de θ est dit biaisé si pour tout $\theta \in \Theta$ et tout entier positif n*

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

Exemple 1.2.1 *Soit $(X_1, ..., X_n)$ un échantillon issu d'une v.a X telles que $E(X_i) = \mu$ et $V(X_i) = \sigma^2$, $\forall i = 1, ..., n$.*

La variance empirique

$$\tilde{S}^2 := \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \bar{X} \right)^2 \tag{1}$$

est un estimateur biaisé de σ^2.

En effet

$$\begin{aligned}
E(\tilde{S}^2) &= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} E\left(X_i - \bar{X}\right)^2 \\
&= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} E\left(X_i - \mu + \mu - \bar{X}\right)^2 \\
&= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} E(X_i - \mu)^2 - E(\bar{X} - \mu)^2 \\
&= \tfrac{1}{n}\textstyle\sum_{i=1}^{n} V(X_i) - V(\bar{X}) \\
&= \sigma^2 - \frac{\sigma^2}{n} \\
&= \frac{n-1}{n}\sigma^2 \neq \sigma^2.
\end{aligned}$$

Estimateur sans biais

Définition 1.2.4 *Un estimateur* T_n *de* θ *est dit sans biais si pour tout* $\theta \in \Theta$

$$E(T_n) = \theta \text{ ou } (b(\theta) = 0).$$

Exemple 1.2.2 ***1-*** *Soit* $(X_1, ..., X_n)$ *un échantillon issu d'une v.a* X *telles que* $E(X_i) = \mu$ *et* $V(X_i) = \sigma^2$, $\forall i = 1, ..., n$.

La moyenne empirique $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ *est un estimateur sans biais de l'espérance mathématique.*

En effet

$$E\left(\bar{X}\right) = E\left(\tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i\right) = \tfrac{1}{n}\textstyle\sum_{i=1}^{n} E\left(X_i\right) = \tfrac{1}{n} n E\left(X_i\right) = E\left(X_i\right) = \mu.$$

2- *La variance empirique corrigée* $\tilde{S}^{*2}$ *définie par*

$$\tilde{S}^{*2} := \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \tag{2}$$

est un estimateur sans biais de σ^2*.*

On a $E\left(\tilde{S}^{*2}\right) = \frac{n}{n-1} E\left(\tilde{S}^2\right) = \frac{n}{n-1} E\left(\tilde{S}^2\right)$
$= \frac{n}{n-1} (\frac{n-1}{n} \sigma^2) = \sigma^2.$

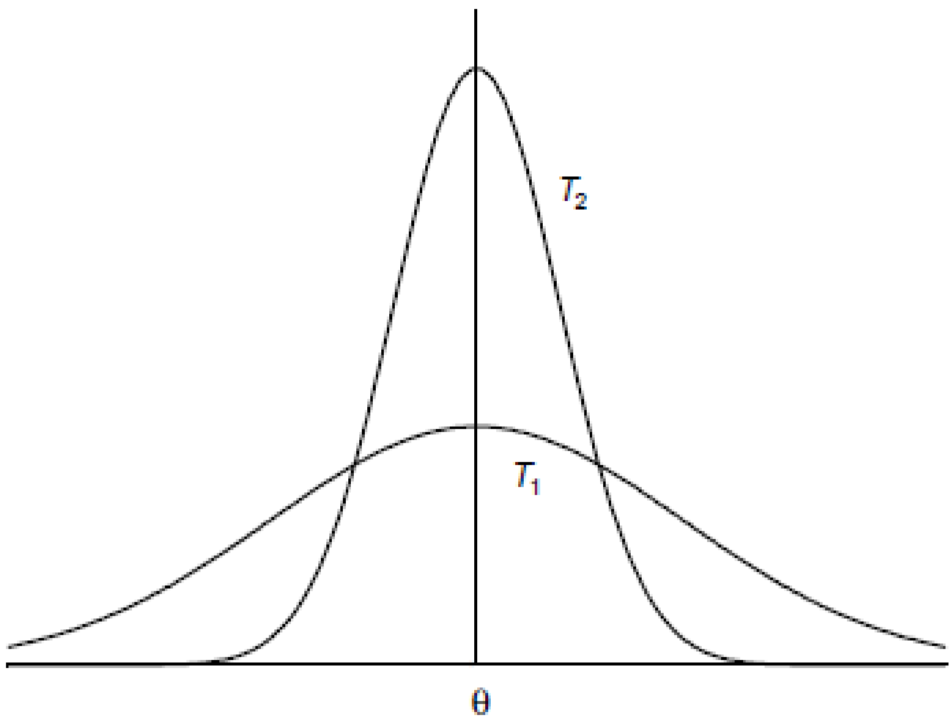

FIG. 1.1 – Comparaison d'estimateur avec $E(T_1)$ =$E(T_2)$ et $Var(T_1) > Var(T_2)$.

Estimateur asymptotiquement sans biais Un estimateur T_n de θ est dit asymptotiquement sans biais si pour tout $\theta \in \Theta$

$$E(T_n) = \theta, \text{ quand } (n \to \infty).$$

Exemple 1.2.3 *La variance empirique $\tilde{S}^2$ est un estimateur asymptotiquement sans biais de σ^2.*

En effet $E\left(\tilde{S}^2\right) = \frac{n-1}{n}\sigma^2$

$$\lim_{n\to+\infty} E\left(\tilde{S}^2\right) = \lim_{n\to+\infty}\left(\tfrac{n-1}{n}\sigma^2\right) = \sigma^2.$$

Estimateur convergent

– Un estimateur T_n de θ si T_n converge en probabilité vers θ lorsque n tend vers l'infini.

– Un estimateur T_n de θ est dit asymptotiquement normale ssi

$$\frac{E(T_n)-\theta}{\sqrt{Var(\theta)}} \overset{loi}{\to} N\left(0,1\right), \text{ quand } (n\to\infty).$$

Erreur (Risque) quadratique moyenne

L'erreur quadratique moyenne de T_n est défini comme

$$\begin{aligned} MSE(n) &= E\left[(T_n-\theta)^2\right] \\ &= Biais^2(T_n) + Var(T_n). \end{aligned}$$

Comparaison d'estimateurs

Définition 1.2.5 *Soit T_1, T_2 deux estimateurs sans biais de θ, T_1 est dit plus efficace que T_2 si*

$$Var\left(T_1\right) < Var\left(T_2\right).$$

Cependant que T_1 est « meilleur » que T_2. En effet

– il tient compte de toute l'information apportée par l'échantillon,

– sa variance est la plus petite : $Var\left(T_1\right) < Var\left(T_2\right)$.

1.3 Statistique exhaustive

Définition 1.3.1 *Soit $(X_1, ..., X_n)$ un échantillon. On dit qu'une statistique T est une statistique exhaustive pour θ si la loi conditionnelle de l'échantillon eqt indépendante du θ. Ceci veut dire qu'un fois T connu, aucune valeur de l'échantillon ni aucune autre statistique ne nous apportera de renseignements supplémentaires sur θ.*

Exemple 1.3.1 *Dans la loi normale $\mathcal{N}(\mu, \sigma^2)$*

- *si σ est connu, $T = \sum_{i=1}^{n} X_i$ est une statistique exhaustive pour μ ;*
- *si μ est connu, $T = \sum_{i=1}^{n}(X_i - \mu)^2$ est une statistique exhaustive pour σ^2 ;*
- *si μ et σ sont tous les deux inconnus, le couple $(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n}(X_i - \bar{X})^2)$ ou $(\bar{X}, \tilde{S}^2)$ est exhaustive pour le couple (μ, σ^2).nformation de Fisher*

1.3.1 Information de Fisher

Soit le modèle paramétrique $(E, \mathcal{E}, P_\theta)$, $\theta \in \Theta$, de v.a X.

Définition 1.3.2 *On appelle quantité d'information de Fisher $I_n(\theta)$ apportée par un échantillon sur le paramètre θ, la quantité positive où nulle suivante*

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

où

$$L\,(X_1, ..., X_n, \theta) = \prod_{i=1}^{n} f\,(X_i, \theta)\,.$$

Théor■me 1.3.1 *Si le domaine de définition de X ne dépend pas du paramètre θ alors*

$$I_n(\theta) = -E\left(\frac{\partial^2 \ln L\,(X, \theta)}{\partial \theta^2}\right).$$

Preuve. Voir([5]). ■

Inégalité de Frechet-Darmois-Cramer-Rao (FDRS)

Le résultat suivant nous indique que la variance d'un estimateur ne peut être inférieure à une certaine borne, qui dépend de la quantité d'information de Fisher apportée par l'échantillon sur le paramètre θ.

Si le domaine de définition de X ne dépend pas de θ, on a pour tout estimateur T sans biais de θ

1.
$$Var(T) \geq \frac{1}{I_n(\theta)},$$

et si T est un estimateur sans biais de $h(\theta)$

$$Var(T) \geq \frac{(h'(\theta))^2}{I_n(\theta)}.$$

Estimateur sans biais de variance minimale

La question se pose de savoir si on peut o la borne minimale de la variance, un tel estimateur sera qualifié d'efficace.

L'efficacité n'est donc définie que dans les conditions de régularité suivantes qui sont celles de FDCR

1) Le domaine de définition E est indépendant de θ.

2) $\frac{\partial L}{\partial \theta}$ existe et est continue par rapport à θ.

3) $I_n(\theta)$ est finie.

4) $\frac{\partial L}{\partial \theta}$,$T\frac{\partial L}{\partial \theta}$ sont intégrables par apport à θ.

Définition 1.3.3 *On dit que T est un estimateur efficace si sous les conditions précédentes*

- *T est un estimateur sans biais de θ*

$$Var(T) = \frac{1}{I_n(\theta)}.$$

- *T est un estimateur sans biais de $h(\theta)$*

$$Var(T) = \frac{(h'(\theta))^2}{I_n(\theta)}.$$

Chapitre 2

Méthodes d'estimations paramétrique

L'estimation forme une grande branche de la statistique inférentielle, dont on trouve qu'il existe plusieurs méthodes d'estimation ou de construction d'un estimateur d'un paramètre. Alors l'objectif de ce chapitre est de présenter celles qui sert à donner une valeur unique pour un paramètre inconnu θ qu'on l'appelées méthodes d'estimation ponctuelles comme la méthode des moments et la méthode du maximum de vraisemblance aussi celles qui sert à donner un ensemble des valeurs plausibles pour θ essentiellement sous forme d'un intervalle (méthode d'estimation par intervalle de confiance).

2.1 Estimation par la méthode des moments

Bien que cette méthode ne soit pas toujours satisfaisante nous l'introduisons dès maintenant en raison de son côté intuitif.

Soit $\mathcal{M} = (E, \mathcal{E}, \{P_\theta; \theta \in \Theta\})$ un modèle paramétré tel que Θ est un ouvert de

$\mathbb{R}^n$.

2.1.1 Méthode classique

La méthode dite **des moments** a été introduite par K. Pearson et fut la première méthode de construction d'estimateurs.

L'idée est simple, elle s'appuie sur le fait que les moments empiriques $\mathrm{M}_k(n)$ sont des estimateurs acceptables des moments théoriques m_k d'une loi, puisque, d'après la loi faible des grands nombres

$$\mathrm{M}_k(n) = \frac{1}{n}\sum_{i=1}^{n} X_i^k \xrightarrow{P_\theta} E(X_1^k) = m_k(\theta) \text{ si } E(X_1^k) \text{ existe.}$$

Cette idée a par ailleurs un fondement asymptotique.

Par conséquent, puisque $\Theta \subset \mathbb{R}^n$, une valeur de θ pour laquelle n moments empiriques sont égaux aux n moments théoriques correspondants semble être une estimation raisonnable du paramètre.

Définition 2.1.1 *Soit $\mathcal{M}$ un modèle tel que pour tout $\theta \in \Theta$, l'observation X possède tous les P_θ -moments d'order inférieur ou égal à n. Une solution mesurable, lorsqu'elle existe, du système (d'inconnue θ)*

$$\begin{cases} M_1(n) & = & m_1(\theta) \\ \quad . & & \\ \quad . & & \\ \quad . & & \\ M_n(n) & = & m_n(\theta) \end{cases}$$

est appelée estimateur obtenu par la méthode des moments.

2.1.2 Généralisation

Dans la pratique, il est possible qu'en utilisant les n premiers moments, le système associé n'admette pas de solution. On est alors amené à choisir d'autre moments $m_{i_1}(\theta)$,...,$m_{i_n}(\theta)$. Plus généralement, il est possible de choisir des applications fonctions par X, autre que les fonctions puissances et de suivre la même démarche que celle développée précédemment.

soit $\varphi = (\varphi_1, ..., \varphi_n)$ un vecteur composé de fonction $\mathcal{P}$-intégrables.

Posons

$$\psi(\theta) = E_\theta\varphi(X) et \varphi_n(X) = \frac{1}{n}\sum_{i=1}^{n}\varphi(X_i) = (\phi_{1,n}(X), ..., \phi_{n,n}(X)).$$

On appellera estimateur obtenu par la méthode des moments une solution mesurable, lorsqu'elle existe du système suivant

$$\left\{\begin{array}{ccccc} \phi_{1,n}(X) & = & E_\theta\varphi_1(X) & = & \psi_1(\theta) \\ . & & & & \\ . & & & & \\ . & & & & \\ \phi_{n,n}(X) & = & E_\theta\varphi_n(X) & = & \psi_n(\theta) \end{array}\right. .$$

Exemple 2.1.1 ***(Cas d'un paramètre à une dimension)***

Pour une réalisation $x_1, ..., x_n$ de l'échantillon la méthode consiste alors à choisir pour estimation de θ la valeur telle que la moyenne théorique m_1 (ou premier moment de la loi) coïncide avec la moyenne empirique $\bar{x}$.

- Pour la loi $\mathcal{E}(\lambda)$, par exemple, l'estimation de λ sera $\hat{\lambda}$ telle que $\frac{1}{\hat{\lambda}} = \bar{x}$, soit $\hat{\lambda} = \frac{1}{\bar{x}}$.

- Pour la loi $\mathcal{BN}(r,p)$ avec r connu, l'estimation de p sera $\hat{p}$ telle que

$$\frac{r\,(1-\hat{p})}{\hat{p}} = \bar{x},\ \ d\text{'où } \hat{p} = \frac{r}{r+\bar{x}}.$$

- Pour la loi de Poisson la solution est $\bar{x}$ puisque le paramètre λ est lui-même la moyenne de la loi.

Exemple 2.1.2 *- Pour la loi de Bernoulli p est estimée par la moyenne qui est la fréquence relative observée.*

Exemple 2.1.3 (Cas d'un paramètre à deux dimensions)

L'estimation résulte de la résolution de deux équations, l'une reposant sur le premier moment, l'autre sur le moment d'ordre 2.

- Prenons le cas de la loi de Gauss avec (μ,σ^2) comme paramètre, dont le premier moment est μ lui-même et le moment d'ordre 2 est $E(X^2) = \mu^2+\sigma^2$.

On résout donc (en passant directement aux v.a.)

$$\begin{cases} \mu = \bar{x}, \\ \mu^2+\sigma^2 = \frac{1}{n}\sum_{i=1}^n x_i^2, \end{cases}$$

d'où $\hat{\mu} = \bar{x}$ et $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \bar{x}^2 = \tilde{S}^2$. La moyenne et la variance théoriques sont donc estimées naturellement par la moyenne et la variance empiriques.

- Prenons maintenant le cas moins intuitif de la loi de Gumbel de paramètre (α,β), dont la moyenne est $\alpha+\gamma\beta$, où γ est la constante d'Euler, et la variance est $\frac{\pi^2\beta^2}{6}$.

On résout

$$\begin{cases} \alpha+\gamma\beta = \bar{x}, \\ (\alpha+\gamma\beta)^2 + \frac{\pi^2\beta^2}{6} = \frac{1}{n}\sum_{i=1}^n x_i^2, \end{cases}$$

où, de façon équivalente

$$\begin{cases} \alpha + \gamma\beta & = \bar{x}, \\ \frac{\pi^2\beta^2}{6} & = \tilde{S}^2, \end{cases}$$

ce qui donne la solution $\hat{\beta} = \frac{\sqrt{6}}{\pi}\tilde{S}$ *et* $\hat{\alpha} = \bar{x} - \frac{\gamma\sqrt{6}}{\pi}\tilde{S}$

D'une façon générale l'estimation de θ *de dimension* 2 *par la méthode des moments est la solution (supposée exister et être unique pour toute réalisation du n-échantillon aléatoire) du système*

$$\begin{cases} m_1(\theta) & = \bar{x}, \\ m_2(\theta) & = \frac{1}{n}\sum_{i=1}^{n} x_i^2, \end{cases}$$

Cette solution appliquée à $\bar{x}$ *et* $\frac{1}{n}\sum_{i=1}^{n} x_i^2$ *donne l'estimateur de* θ *correspondant. Du fait de la correspondance des formules de décentrage pour les moments empiriques et pour les moments théoriques il est équivalent de résoudre*

$$\begin{cases} m_1(\theta) = \bar{x}, \\ \sigma^2(\theta) = \tilde{S}^2. \end{cases}$$

Remarque 2.1.1 *Cette méthode ne convient pas dans toutes les applications pratiques et dans certains cas, elle ne fournit pas les meilleures estimations. C'est pour cette raison que le statisien **Ronald Fisher** proposa dans les années **1920, la méthode du maximum de vraisemblance.***

2.2 Estimation par la méthode du maximum de vraisemblance (MV)

Le maximum de vraisemblance est une méthode générale pour estimer les paramètres d'un modèle statistique. Par exemple, supposons que nous avons une série d'observations d'une variable aléatoire X et un modèle statistique pour cette variable. En général, un tel modèle contient différents paramètres inconnus qui doivent être ajustés aux données observées. Selon le maximum de vraisemblance, les meilleurs estimés des paramètres d'un modèle sont ceux qui maximisent la probabilité des valeurs observées de la variable. Cette méthode peut être appliquée peu importe la forme mathématique du modèle, ce qui permet de choisir les modèles les plus compatibles avec notre compréhension.

2.2.1 Principe

Cette méthode consiste, étant donné un échantillon de valeurs $(X_1, ..., X_n)$ à prendre comme estimation de θ la valeur de $\hat{\theta}$ qui rend maximale la vraisemblance

$$L\left(x_1, ..., x_n\ ;\ \theta\right).$$

En pratique on prend comme estimation de θ une solution de l'équation $\frac{\partial}{\partial\theta}\ ln$ $L(X;\theta) = 0$, dite "**équation de la vraisemblance**". Intuitivement, puisque L représente une densité de probabilité, cela revient à supposer que l'événement qui s'est produit était le plus « probable ».

Définition 2.2.1 *On appelle* ***fonction de vraisemblance*** *pour une réalisation*

$(x_1, ..., x_n)$ *d'un échantillon* $(X_1, ..., X_n)$, *la fonction de* θ

$$L(x_1, ..., x_n; \theta) = \begin{cases} \prod_{i=1}^{n} f(x_i, \theta) & si\ X\ est\ vac, \\ \prod_{i=1}^{n} P_\theta(x_i) & si\ X\ est\ vad. \end{cases}$$

Définition 2.2.2 *On appelle estimateur du maximum de vraisemblance* (MV) *de* $\theta \in \Theta \subset \mathbb{R}^m$ *avec* $m = 1, 2, 3...$ *la statistique définie par*

$$\hat{\theta} = \underset{\theta \in \Theta}{argmax} L(x_1, ..., x_n; \theta).$$

Le problème ci-dessus est compliqué à résoudre en raison de la présence du produit mais il suffit de prendre le logarithme.

Fonction de Log-vraisemblance On appelle fonction de log-vraisemblance pour les donnée $x_1, ..., x_n$ la fonction de définie par

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

Pour trouver le maximum on résoudre l'équation de vraisemblance définie par

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

on obtient l'MV $\hat{\theta} = \varphi(x_1, ..., x_n)$.

Remarque 2.2.1 *Il faut vérifier que* $\hat{\theta}$ *est bien un maximum pour* $L(x_1, ..., x_n; \theta)$, *c'est à dire, il faut qu'il vérifie*

$$\frac{\partial^2}{\partial \theta^2} L(x_1, ..., x_n; \theta) \leq 0, \ ou\ \frac{\partial^2}{\partial \theta^2} l(x_1, ..., x_n; \theta) \leq 0.$$

2.2.2 Propriétés des estimateurs du maximum de vraisemblance

Propriété 2.2.1 $(Biais)$

$\hat{\theta}$ est sans biais ou asymptotiquement sans biais
$E\left(\hat{\theta}\right) = \theta$ ou $E\left(\hat{\theta}\right) \to \theta$ quand $n \to \theta$.

Propriété 2.2.2 $(Convergence)$

$\hat{\theta}$ converge en probabilité vers θ : $\hat{\theta} \underset{proba}{\to} \theta$.

Propriété 2.2.3 $(Efficacité)$

La loi de probabilité de $\hat{\theta}$ tend vers une loi normale de moyenne θ et de variance $\frac{1}{I_n(\theta)}$. $I_n(\theta)$ est la quantité d'information de Fisher.
$Var\left(\hat{\theta}\right) \to \frac{1}{I_n(\theta)}$ exprime que l'estimateur est asymptotiquement efficace.
- L'estimateur de maximum de vraisemblance peut ne pas exister.
- Si un estimateur de maximum de vraisemblance existe, il n'est pas toujours unique, par exemple : modèle de Cauchy et modèle de Laplace.

Exemple 2.2.1 ***(Cas discrèt)***
Soit X une variable aléatoire binomiale d'ordre n et de paramètre p. $\forall\ x \in \{0, 1, ..., n\}$, la probabilité élémentaire $P(X = x)$ de x est donnée par

$$P(X = x) = C(n, x)\ p^x(1 - p)^{n-x},$$

de plus

$$\begin{cases} E[X] = & np \\ Var[X] = & np(1-p) \end{cases},$$

(a) **Recherche du maximum de vraisemblance** Considérons un r-échantillon de cette structure. Sa fonction de vraisemblance est définie pour tout $p \in [0,1]$ et tout $(x_1,...,x_r) \in \{0,1,...,n\}^r$ par

$$\begin{aligned} L(x_1,...,x_n,p) &= \prod_{i=1}^{r} P(x_i), \\ &= \left[\prod_{i=1}^{r} C(n,x)\right] p^{\sum_{i=1}^{r} x_i}(1-p)^{nr-\sum_{i=1}^{r} x_i}, \end{aligned}$$

d'où

$$\ln L(x_1,...,x_r,p) = \ln\left[\prod_{i=1}^{r} C(n,x)\right] + \sum_{i=1}^{r} x_i \ln(p) + \left(nr - \sum_{i=1}^{r} x_i\right)\ln(1-p),$$

Il en résulte que

$$\begin{aligned} \frac{\partial}{\partial p}\ln L(x_1,...,x_r,p) &= \frac{\sum_{i=1}^{r} x_i}{p} - \frac{(\ nr - \sum_{i=1}^{r} x_i)}{1-p}, \\ &= \frac{\sum_{i=1}^{r} x_i - rnp}{p(1-p)} \end{aligned}$$

d'où

$$\frac{\partial}{\partial p}\ln L(x_1,...,x_r,p) = 0 \Rightarrow p = \frac{1}{rn}\sum_{i=1}^{r} x_i.$$

et comme

$$\frac{\partial^2}{\partial p^2}\ln L(x_1,...,x_r,p) \leq 0$$

donc, l'estimateur du maximum de vraisemblance d'un r-échantillon d'une structure binomiale est définie par

$$\hat{p} = \frac{1}{rn}\sum_{i=1}^{r} x_i,$$

(b) **Etude des propriétés de $\hat{p}$**

$$E[\hat{p}] = \frac{1}{n}\ E[X],$$

et

$$\begin{aligned} Var[\hat{p}] &= \frac{Var[X]}{rn^2}, \\ &= \frac{p(1-p)}{rn}. \end{aligned}$$

on en déduit que $\hat{p}$ est un estimateur sans biais et convergent de p.

Exemple 2.2.2 ***(Cas continu)***

Soit X une variable aléatoire normale de paramètres μ et σ^2. Sa densité de probabilité f est définie pour tout $x \in \mathbb{R}$ par

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

de plus

$$\begin{cases} E[X] & = \mu \\ Var[X] & = \sigma^2 \end{cases}.$$

Considérons un n-échantillon de cette structure. Sa fonction de vraisemblance est définie pour tout $\mu \in \mathbb{R}$ tout $\sigma > 0$, et tout $(x_1, ..., x_n)$ par

$$\begin{aligned} L(x_1, ..., x_n; \mu, \sigma^2) &= \prod_{i=1}^{n} f\ \ (x_i), \\ &= \frac{1}{\left(\sigma\sqrt{2\pi}\right)^n} \exp\left(\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\mu)^2\right), \end{aligned} \tag{2.1}$$

d'où

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

Il en résulte que

$$\begin{cases} \frac{\partial}{\partial\mu}\ln L(x_1, ..., x_n; \mu, \sigma^2) & = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu), \\ \frac{\partial}{\partial\sigma}\ln L(x_1, ..., x_n; \mu, \sigma^2) & -\frac{n}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2 . \end{cases}$$

d'où

$$\begin{cases} \frac{\partial}{\partial\mu}\ln L(x_1, ..., x_n; \mu, \sigma^2) & = 0, \\ \frac{\partial}{\partial\sigma}\ln L(x_1, ..., x_n; \mu, \sigma^2) & = 0. \end{cases} \Longrightarrow \begin{cases} \mu & = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \sigma^2 & = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \end{cases} .$$

pour ces deux valeurs, on a

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

et

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

Donc les estimateurs du maximum de vraisemblance d'un n-échantillon d'une structure normale sont

$$\begin{cases} \hat{\mu} & = \frac{1}{n}\sum_{i=1}^{n} x_i \\ \hat{\sigma}^2 & = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 \end{cases} .$$

$b-$ **Etude des propriétés de $\hat{\mu}$ et $\hat{\sigma}^2$**

$$E[\hat{\mu}] = E[X] = \mu,$$

Loi de probablité	Estimateur du maximum de vraisemblance
Bernoulli	$\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$
Géométrique	$\hat{p} = \frac{n}{\sum_{i=1}^{n} X_i}$
Poisson	$\hat{\alpha} = \frac{1}{n}\sum_{i=1}^{n} X_i$
Exponentielle	$\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} X_i}$
Uniforme	$\hat{\theta} = \max\left(X_1, ..., X_n\right)$

TAB. 2.1 – Estimation par la méthode du maximum de vraisemblance des paramètres de quelques lois de probablités.

et

$$E\left[\hat{\sigma}^2\right] = \frac{n-1}{n}\sigma^2.$$

On en déduit que $\hat{\mu}$ est un estimateur sans biais et convergent de μ, mais $\hat{\sigma}$ est un estimateur biaisé de σ.

Donc

2.3 Estimation par la méthode des intervalles de confiance

2.3.1 Définition d'un intervalle de confiance

L'estimation par intervalle de confiance d'un paramètre θ consiste donc à associer à un échantillon, un intervalle aléatoire I, choisi de telle façon que la probabilité pour qu'il contienne la valeur inconnue du paramètre soit égale à un nombre fixé à l'avance, aussi grand que l'on veut. On écrit

$$P\left(\theta \in I\right) = 1 - \alpha,$$

où $(1-\alpha)$ est la probabilité associée à l'intervalle d'encadrer la vraie valeur du paramètre, c'est le seuil de confiance ou la quasi-certitude.

2.3.2 Construction d'un intervalle de confiance

Soit X une variable aléatoire dont la densité $f(x,\theta)$ dépend du paramètre θ et $(X_1,...,X_n)$ un échantillon de taille n de cette variable. Soit $\hat{\theta}=\varphi(X)$ un estimateur du paramètre θ et $g(t;\theta)$ la loi de probabilité de cet estimateur.
Étant donnée une probabilité α, on peut à partir de cette loi et si on suppose que le paramètre θ connu, construire un intervalle de probabilité pour la variable aléatoire θ

$$P(\theta-\mathfrak{h}_1<\hat{\theta}<\theta+\mathfrak{h}_2)=1-\alpha \tag{2.2}$$

Les bornes de l'intervalle sont définies par (avec $\alpha_1+\alpha_2=\alpha$)

$$P(\hat{\theta}<\theta-\mathfrak{h}_1)=\int_{-\infty}^{\theta-\mathfrak{h}_1} g(t;\theta)dt=\alpha_1 P(\hat{\theta}>\theta+\mathfrak{h}_2)=\int_{\theta+\mathfrak{h}_2}^{+\infty} g(t;\theta)dt=\alpha_2.$$

Si l'égalité (1) est vérifiée, on obtient l'égalité suivante

$$P(t-\mathfrak{h}_2<\theta<t+\mathfrak{h}_2)=1-\alpha$$

où t est la valeur de la statistique $\hat{\theta}$ donnée par l'échantillon est également vérifiée.
L'intervalle $I=[t-\mathfrak{h}_2,t+\mathfrak{h}_1]$ a une probabilité égale à $(1-\alpha)$ de contenir le paramètre θ c'est un intervalle de confiance au niveau de confiance $(1-\alpha)$.

2.3.3 Principe

La méthode des intervalles de confiance est la suivante : Soit $\hat{\theta}$ un estimateur de θ (on prendra évidemment le meilleur estimateur possible), dont on connaît la loi de probabilité pour chaque valeur de θ. Étant donné une valeur θ_0 de θ, on détermine un intervalle de probabilité de niveau $(1-\alpha)$ pour $\hat{\theta}$, c'est -à-dire, on détermine deux bornes t_1 et t_2 telles que

$$P(t_1 < \hat{\theta} < t_2/\theta = \theta_0) = 1 - \alpha.$$

Ces bornes dépendent évidemment de θ. On choisit dans la plupart des cas un intervalle de probabilité à risques symétriques $\frac{\alpha}{2}$ et $\frac{\alpha}{2}$.

On adopte alors la règle de décision suivante : soit t la valeur observée de $\hat{\theta}$

- si $t \in [t_1, t_2]$ on conserve θ_0 comme valeur possible de θ,
- si $t \notin [t_1, t_2]$ on élimine θ_0.

On répète cette opération pour toutes les valeurs de θ.

2.3.4 Intervalle de confiance de la moyenne d'une loi normale

Soit $(X_1, ..., X_n)$ un échantillon issu d'une variable aléatoire X de loi Normale $\mathcal{N}(\mu, \sigma^2)$.

Il existe une expression approchée pour l'IC de niveau $(1-\alpha)$ d'une moyenne μ, ou l'intervalle est fondé sur la valeur observée $\hat{\mu}$ obtenue à partir de l'échantillon. Mais cet intervalle dépend de l'écart type s'il est : connu ou inconnu. Dans ce qui suit on va traiter ces deux situations.

Cas σ est connu

On sait que $\bar{X}$ est le meilleur estimateur de μ et qu'il suit une loi $\mathcal{N}(\mu, \frac{\sigma^2}{n})$, ce qui donne que la variable aléatoire

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

On peut déterminer la valeur $u_{\frac{\alpha}{2}}$ du quantile d'ordre $(1-\frac{\alpha}{2})$ de la loi $\mathcal{N}(0,1)$, tel que

$$P(-u_{\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\bar{X}-\mu)}{\sigma} \leq u_{\frac{\alpha}{2}}) = 1-\alpha,$$

ce qui conduit à

$$\begin{aligned} P(-u_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \leq (\bar{X}-\mu) \leq u_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}}) &= 1-\alpha \\ P(\bar{X}-u_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}+u_{\frac{\alpha}{2}}\tfrac{\sigma}{\sqrt{n}}) &= 1-\alpha \end{aligned}$$

d'où l'intervalle de confiance pour μ au niveau de confiance $(1-\alpha)$ est

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

Exemple 2.3.1 (σ ***connu***)

Après des essais antérieurs, on peut supposer que la résistance à l'éclatement d'un certain type de réservoirs est une variable aléatoire suivant une loi normale de moyenne μ inconnue et d'écart-type égal à $4kg/cm^2$. Des essais sur un échantillon de 9 réservoirs donnent une résistance moyenne à l'éclatement égale à $215kg/cm^2$. On a donc

* une estimation ponctuelle de la moyenne donnée par l'échantillon : $\bar{X} = 215kg/cm^2$,
* la loi suivie par la moyenne d'un échantillon de taille $n = 9$ (avec l'hypothèse

admise sur la loi suivie par la résistance) : la loi Normale $\mathcal{N}(215, 16/9)$,

* le niveau de confiance : $(1-\alpha) = 0.95$.

l'intervalle de confiance pour la résistance moyenne sera donné par

$$P(-1.96 \leq \frac{\bar{X}-\mu}{4/3} \leq 1.96) = 0.95, \quad \bar{X} = 215,$$

alors

$$P(215 - 4/3 \times 1.96 < \mu < 21514/331.96) = P(212.386 < \mu < 217.613) = 0.95.$$

D'où $IC_{1-\alpha}(\mu) = [212.386, 217.613]$, ce qui désigne que l'intervalle $[212.386, 217.613]$ a une probabilité égale à 0.95 de contenir la vraie valeur de la résistance à l'éclatement de ce type de réservoirs.

Remarque 2.3.1 *Cet exemple montre que si la taille n de l'échantillon augmente, α et σ restant constants, l'étendue de l'intervalle diminue en revanche, si le seuil σ diminue l'étendue de l'intervalle augmente.*

Cas σ est inconnu

On utilise le fait que la statistique

$$T = \frac{\bar{X}-\mu}{\tilde{S}}\sqrt{n-1}$$

suit une loi de Student à $(n-1)$ degrés de liberté, où $\tilde{S}$ est l'estimateur ponctuelle de σ basé sur la variance empirique définie par 2.2.
A cet effet, la détermination d'un IC de la moyenne, dans cette situation, consiste à déterminer la valeur du fractile d'ordre $(1-\frac{\alpha}{2})$ d'une loi de Student de degré de

liberté $(n-1)$ $(t_{\frac{\alpha}{2}})$, c'est à dire

$$P(-t_{\frac{\alpha}{2}} \leq \frac{\bar{X}-\mu}{\tilde{S}}\sqrt{n-1} \leq t_{\frac{\alpha}{2}}) = 1-\alpha,$$

d'où l'intervalle de confiance de μ est de la forme

$$\bar{X} - t_{\frac{\alpha}{2}}\frac{\tilde{S}}{\sqrt{n-1}} \leq \mu \leq \bar{X} + t_{\frac{\alpha}{2}}\frac{\tilde{S}}{\sqrt{n-1}},$$

ou encore

$$IC_{1-\alpha}(\mu) = [\bar{X} - t_{\frac{\alpha}{2}}\frac{\tilde{S}}{\sqrt{n-1}}, \bar{X} + t_{\frac{\alpha}{2}}\frac{\tilde{S}}{\sqrt{n-1}}.$$

On peut aussi réécrire cet intervalle à l'aide de la variance empirique corrigée $(\tilde{S}^{*2})$ définie par **??** comme suit

$$IC_{1-\alpha}(\mu) = [\bar{X} - t_{\frac{\alpha}{2}}\frac{\tilde{S}^*}{\sqrt{n}}, \bar{X} + t_{\frac{\alpha}{2}}\frac{\tilde{S}^*}{\sqrt{n}}].$$

Remarque 2.3.2 *Le théorème central-limite a pour conséquence que les intervalles précédents sont valables pour estimer μ d'une loi quelconque que n est assez grand.*

Exemple 2.3.2 *(σ **estimé**)*

Afin d'étudier le salaire journalier, en euros, des ouvriers d'un secteur d'activité, on procède à un tirage non exhaustif, d'un échantillon de taille $n = 16$. On a obtenu les résultats suivants

41	40	45	50	41	41	49	43
45	52	40	48	50	49	47	56

On suppose que la loi suivie par la variable aléatoire « salaire journalier » est une loi normale de moyenne μ et d'écart-type σ inconnus, alors

- l'estimateur ponctuel de μ est

$$\bar{X} = 45.4375,$$

- l'estimateur ponctuel de la variance sera donnée par

$\tilde{S}^2 = 15.2460 = (3.9046)^2$ (estimateur biaisé),

$\tilde{S}^{*2} = 16.2625 = (4.0326)^2$ (estimateur non biaisé).

L'intervalle de confiance du salaire journalier moyen μ, au seuil de confiance $0,95$ est obtenu en suivant les étapes précédentes, donc la variable aléatoire $T_{(15)} = \frac{\bar{X}-\mu}{\tilde{S}/\sqrt{n-1}}$ suit une loi de Student à $(16-1=15)$ degrés de liberté, d'où

$$P(-t_{\frac{\alpha}{2}} < T_{(15)} < t_{\frac{\alpha}{2}}) = 0,95 \qquad \text{où } (t_{\frac{\alpha}{2}} = t_{0,025} = 2.131) \cdot$$

En tenant compte les résultats donnés par l'échantillon, on trouve

$$P(-2.131 \leq \frac{45.4375-\mu}{3.9046/\sqrt{15}} \leq 2.131) = 0.95$$

$$P(45.4375 - 2.131 \times 3.9046/\sqrt{15} < \mu < 45.4375 + 2.131 \times 3.9046/\sqrt{15}) = 0.95$$

alors

$$P(43.2895 < \mu < 47.5859) = 0.95.$$

L'intervalle $IC_{1-\alpha}(\mu) = [43.2895, 47.5859]$ a une probabilité égale à 0.95 de contenir la vraie valeur du salaire moyen journalier des ouvriers de ce secteur d'activité.

2.3.5 Intervalle de confiance de la variance d'une loi normale

Comme dans le cas de la moyenne, l'intervalle de confiance de la variance σ^2 nécessite une information sur le paramètre μ. Alors, deux constructions d'un intervalle de confiance de σ^2 peut être obtenues, et cela selon μ est connu ou non.

Cas la moyenne μ est connue

On a $S^2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ est le meilleur estimateur de σ^2 dans le cas où la vraie moyenne est connue et la variable aléatoire $\dfrac{nS^2}{\sigma^2}$ suit une loi de Khi deux à n degrés de liberté comme somme de n carrés de loi $\mathcal{N}(0,1)$ indépendantes.
Soit k_1 et k_2 les bornes de l'intervalle de probabilité d'un $\mathcal{X}_n^2$, alors

$$P(k_1 \leq \frac{nS^2}{\sigma^2} \leq k_2) = 1-\alpha,$$

donc

$$P(\frac{k_1}{nS^2} \leq \frac{1}{\sigma^2} \leq \frac{k_2}{nS^2}) = 1-\alpha,$$
$$P(\frac{nS^2}{k_2} \leq \sigma^2 \leq \frac{nS^2}{k_1}) = 1-\alpha,$$

ce qui équivalent à

$$IC_{1-\alpha}(\sigma^2) = [\frac{nS^2}{k_2}, \frac{nS^2}{k_1}],$$

où k_1 et k_2 sont les quantiles d'ordre $\frac{\alpha}{2}$ et $(1-\frac{\alpha}{2})$ de loi $\mathcal{X}_n^2$ respectivement.

Exemple 2.3.3 *Soit une v.a X de loi $\mathcal{N}(12,\sigma^2)$, $n = 20$, $\sum_{i=1}^{20} X_i^2 = 2960$, on veut déterminer un intervalle de confiance pour σ^2 au niveau de confiance $1-\alpha = 0.9$.*

Le niveau de confiance est 90%*, ce qui donne des risques symétriques* $\frac{\alpha}{2} = 5\%$*.*

Les valeurs k_1 *et* k_2 *lues dans les tables d'une* $\mathcal{X}_n^2$ *sont* $k_1 = 10.85$ *et* $k_2 = 31.41$*.*

$$S^2 = \frac{\sum_{i=1}^{20}(X_i - \mu)^2}{20} = \frac{\sum_{i=1}^{20} X_i^2}{20} - \mu^2 = \frac{2960}{20} - 144 = 4.$$

Alors, un intervalle de confiance pour σ^2 *sera donné par*

$$IC_{0.90}(\sigma^2) = [\frac{20 \times 4}{31.41}, \frac{20 \times 4}{10.85}] = [2.55, 7.37]$$

La variance est inconnue, l'échantillon observé aboutit à une estimation ponctuelle égale à 4*, mais sa valeur exacte à* 90% *de chances de se situer entre* 2.55 *et* 7.37*.*

Cas la moyenne μ est inconnue

On utilise dans ce cas la statistique $\tilde{S}^2 = \frac{1}{n}\sum\limits_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ comme le meilleur estimateur de σ^2 et on sait que $\dfrac{n\tilde{S}^2}{\sigma^2}$ suit une loi de $\mathcal{X}_{n-1}^2$.

On peut déterminer l_1 et l_2 de telle sort que

$$P(l_1 \leq \frac{n\tilde{S}^2}{\sigma^2} \leq l_2) = 1 - \alpha,$$

ce qui donne

$$P(\frac{l_1}{n\tilde{S}^2} \leq \frac{1}{\sigma^2} \leq \frac{l_2}{n\tilde{S}^2}) = 1 - \alpha,$$

$$P(\frac{n\tilde{S}^2}{l_2} \leq \sigma^2 \leq \frac{n\tilde{S}^2}{l_1}) = 1 - \alpha,$$

c'est à dire

$$IC_{1-\alpha}(\sigma^2) = [\frac{n\tilde{S}^2}{l_2}, \frac{n\tilde{S}^2}{l_1}].$$

Les valeurs de l_1 et l_2 sont en fait déterminer par

$$P(\mathcal{X}^2_{n-1} < l_1) = P(\mathcal{X}^2_{n-1} > l_2) = \frac{\alpha}{2}.$$

Remarque 2.3.3 *Ces formules ne sont valables que si X suit une loi normale.*

Exemple 2.3.4 *μ est estimée*

On prend l'exemple 2. .

On a la variable $\frac{n\tilde{S}^2}{\sigma^2}$ suit une loi du Khi-deux à $16-1=15$ degrés de liberté.

D'où la suite des calculs :

$$\begin{aligned}
&P\left(l_1 \leq \frac{n\tilde{S}^2}{\sigma^2} \leq l_2\right)\\
&= P\left(6.262 < \frac{n\tilde{S}^2}{\sigma^2} < 27.488\right)\\
&= P\left(6.262 \leq \tfrac{16\times 15.246}{\sigma^2} \leq 27.488\right)\\
&= P\left(\tfrac{16\times 15.246}{27.488} \leq \sigma^2 \leq \tfrac{16\times 15.246}{6.262}\right)\\
&= P\left(8.874 \leq \sigma^2 \leq 38.955\right)\\
&P\left(6.262 \leq \tfrac{16\times 15.246}{\sigma^2} \leq 27.488\right)\\
&P\left(\tfrac{16\times 15.246}{27.488} \leq \sigma^2 \leq \tfrac{16\times 15.246}{6.262}\right)\\
&= P\left(8.874 \leq \sigma^2 \leq 38.955\right)\\
&= 0.95.
\end{aligned}$$

L'intervalle $IC_{1-\alpha}(\sigma^2) = [8.874, 38.955]$ a une probabilité égale à 0.95 de contenir la vraie valeur de la variance du salaire horaire des ouvriers de ce secteur d'activité.

2.3.6 Intervalle de confiance pour une proportion p

Étant donné une population infinie (ou finie si le tirage s'effectue avec remise) où une proportion p des individus possède un certain caractère A, il s'agit de trouver un intervalle de confiance pour p à partir de f_n qui représente la fréquence empirique ou la proportion des individus ayant le caractère A dans un échantillon de taille n.

On sait que nf_n suit une loi binomiale $\mathcal{B}(n,p)$; si n est faible on utilisera les tables de la loi binomiale et si n est grand on utilise le fait que $nf_n \sim N(np; np\,(1-p))$ donc que :

$$f_n \sim \mathcal{N}\left(p, \sqrt{\frac{p\,(1-p)}{n}}\right).$$

Soit $\alpha = [0,1]$, alors on a

$$P\left(-u_{\frac{\alpha}{2}} \leq \frac{f_n - p}{\sqrt{\frac{p(1-p)}{n}}} \leq u_{\frac{\alpha}{2}}\right) = 1-\alpha,$$

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

Une bonne approximation de l'intervalle de confiance de niveau $(1-\alpha)$ de p, fondé sur la fréquence empirique f_n est donné par l'intervalle ci-dessous

$$IC_{1-\alpha}\,(\mu) = \left[f_n - u_{\frac{\alpha}{2}}\sqrt{\frac{f_n\,(1-f_n)}{n}}, f_n + u_{\frac{\alpha}{2}}\sqrt{\frac{f_n\,(1-f_n)}{n}}\right].$$

2.3.7 Intervalle de confiance pour le paramètre λ d'une loi de Poisson

Soit $\bar{X}$ la moyenne d'un échantillon issu d'une variable aléatoire X de loi de Poisson : $\mathcal{P}(\lambda)$.

Comme on sait que $P(X \leq k) = P\left(X^2_{2(k+1)} > 2\lambda\right)$, on en déduit que l'intervalle de confiance pour λ à risques symétriques de niveau $(1-\alpha)$ est de la forme

$$\frac{1}{2n}a \leq \lambda \leq \frac{1}{2n}b$$

où a est le quantile d'ordre $(\alpha/2)$ de loi Khi deux à $(2n\bar{X})$ degrés de liberté, tandis que b est le quantile d'ordre $(\alpha/2)$ de loi Khi deux à $(2(n\bar{X}+1))$ degrés de liberté.

Exemple 2.3.5 *Soit* $n = l5; n\bar{X} = \sum_{i=1}^{n} X_i = 20; \alpha = 0.1.$

On a a et b sont les quantiles d'ordre 0.05 *de loi Khi deux à* 40 *et* 42 *degrés de liberté respectivement, donc* $a = 26.5$ *et* $b = 58.1$*.*

Un intervalle de confiance pour λ *est donnée par*

$\frac{1}{30}a \leq \lambda \leq \frac{1}{30}b$*, donc*

$\frac{26.5}{30} \leq \lambda \leq \frac{58.1}{30}$ *soit* $0.88 \leq \lambda \leq 1.94$*.*

Pour les grandes valeurs de n, lorsque $2n\bar{X}$ *dépasse les possibilités des tables de* $\mathcal{X}^2_n$*, on utilisera une des approximations normales de la loi du* $\mathcal{X}^2_n$*. Si l'on utilise l'approximation de Wilson et Hilferty, qui est de loin la plus précise, on a*

$$\bar{X}\left(1 - \frac{u}{3\sqrt{n\bar{X}}} - \frac{1}{9n\bar{X}}\right)^3 \leq \lambda \leq \left(\bar{X} + \frac{1}{n}\right)\left(\frac{u}{3\sqrt{n\bar{X}+1}} + 1 - \frac{1}{9\left(n\bar{X}+1\right)}\right)^3.$$

Conclusion

En conclusion, il y a deux types d'estimation proposés : l'estimation ponctuelle et l'estimation par intervalle de confiance. L'estimation ponctuelle donne une seule valeur. En effet dans l'estimation par intervalle on détermine deux valeurs et on donne la probabilité que le paramètre de la population se trouve à l'intérieur de cet intervalle de confiance, et cette estimation a l'avantage de montrer la précision dans l'estimation du paramètre et ceci facilite la compréhension des résultats si l'on ne connait pas les méthodes statistiques.

Les deux estimations sont efficaces chacun dans certain cas, nous pouvons donner des estimateurs égaux, mais là où ces deux cas il est asymptotiquement sans biais, et sa distribution asymptotique est normale.

Finalement, le meilleur estimateur sans biais à variance minimale.

Bibliographie

[1] Cherfaoui.M,(2016/2017). Cours statistique à l'expérimentation en science biologique. Univ-biskra.

[2] Deguilhem, Thibaud. "Tests statistique."

[3] Goldfarb, B. & paradoux, C. (2011) . Introuduction à la méthode statistique, 6° édition. Dunod.

[4] Lejeune, Michel.(2004). Statistique : La théorie et ses applications. Springer Science & Business Media.

[5] Saporta, G.(2011). Probabilités, analyse des données et statistique. Technip, Paris.

[6] Veysseyre,R. (2006). Aide mémoire, statistique et probabilités pour l'ingénieur. Dunod, Paris.