الجمهورية الجزائرية الديمقراطية الشعبية
REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
وزارة التعليم العالي و البحث العلمي
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
جامعة عمار الثليجي بالأغواط
UNIVERSITE AMAR TELIDJI LAGHOUAT
كلية العلوم
FACULTÉ DES SCIENCES
DÉPARTEMENT DE MATHEMATIQUES

Mémoire de MASTER

Domaine : Mathématiques et Informatique
Filière : Mathématiques
Option : Analyse Mathématique

Par : **RAHMOUN Fatiha**

THÈME

Application sur la théorie de l'estimation paramétrique

Soutenu publiquement devant le jury composé de:

- **Mr. A.Rahmoune.** M.C.A **Président**
- **Mr. F.Yazid.** M.C.B **Examinateur**
- **Mr. B.Merrad.** M.C.B **Encadreur**

Année Universitaire 2020/2021

Dédicaces

Tous le

re

cette thè

À Mon trè

ex

toujours pour vous.

Rien au monde ne vaut le

mon éducation et mon bien être. Ce travail et le fruit de te

sacrifice

formation le long de ce

À Ma tendre Mère Khadidja : Tu re

source de tendre

ce

pour que se

leurs étude

A ma adorable grand-mère : Oum elkhir.

À ma adorable grand-père : Ahmed.

A ma chère soeur et profe

A me

Mouhamed.

À me

Mouhamed, Aissa, Ameur.

À ma chère tante : Abla Nakhla, D
À me
À me
Kholoud, Amina, Dounia, rania, nada, Houria Zakhrouf, Iman,
Masseouda.
À tous le
de
chers et que j'ai omis de citer.

Remerciements

En tout pre- mier lieu, je remercie mon Dieu, tout puissant, de m'avoir donné la force pour survivre, ainsi que l'audace pour dépasser toutes les difficultés et Je tiens à exprimer ma reconnaissance à Dr. Boulerbah Merrad pour avoir accepté de m'encadrer dans cette étude. Mes remerciements iront également aux membres du jury pour avoir accepté d'évaluer mon travail de recherche. A tous ceux qui m'ont aidée de prés ou de loin , par un geste, une parole ou un conseil, je leur dis merci. Sans oublier tous nos enseignants qui nous ont assurés des études de haut niveau et qui nous permis d'acquérir des connaissances.

Afin de n'oublier personne, mes vifs remerciements s'adressent à tous ceux qui m'ont aidée à la réalisation de ce modeste mémoire.

♡

❖ *Fatiha*

Résumé

Dans ce mémoire, Nous avons d'abord rappelé les principes fondamentaux de la théorie de l'estimation.
Nous nous sommes intéressés aux différentes méthodes d'estimation (la méthode du maximum de vraisemblance, la méthode des moments et estimation par intervalle de confiance).
à titre illustratif en fin un exemple d'application sur des données simulées traitées par le logiciel R.

Mots-Clés : Estimation ; Méthode du moment ; Méthode du maximum de vraisemblance ; Intervalle de confiance ; Simulation.

ملخص

في هذه المذكرة قمنا بالتذكير بالمبادئ الاساسية لنظرية التقدير.
وكان اهتمامنا بمختلف طرق التقدير (طريقة المعقولية العظمى، طريقة اللحظات، طريقة تقدير بفاصل ثقة).
وفي النهاية ،من أجل توضيح اخذنا مثال تطبيقي على معطيات محاكاة تمت معالجتها ببرنامج R.

الكلمات المفتاحية: ا التقدير ، طريقة اللحظة ، طريقة الاحتمالية القصوى ، فاصل الثقة ، المحاكاة.

Abstract

In this paper, we first recalled the fundamental principles of the theory of estimation. We looked at the different estimation methods (maximum likelihood method, moment method and confidence interval estimation).
As an illustration at the end an example of an application on simulated data processed by the R software.

Key words : Estimation ; Moment Method ; Maximum Likelihood Method ; Confidence Interval ; Simulation

Table des matières

Introduction

La théorie de l'estimation étudie les propriétés des estimateurs et des méthodes générales d'estimation.

La statistique mathématique se compose de deux parties : l'estimation et les tests d'hypothèses. Le premier aspect permet d'évaluer des quantités théoriques inconnues relatives à une population toute entière. Le second est une procédure de prise de décision amenant,sur la base de mesures expérimentales,à l'acceptation ou au rejet d'une hypothèse faite à priori sur une caractéristique inconnue d'une variable aléatoire. Dans ce travail,on s'intéresse à la théorie de l'estimation et plus particulièrement à l'estimation par intervalles de confiance.L'estimation ponctuelle permet de construire les estimateurs puis établir et étudier leurs propriétés.Pour cela,il existe une variété de méthodes comme celle des moments ou du maximum de vraisemblance. L'estimation par intervalles de confiance est complémentaire de l'estimation ponctuelle. En effet, pour évaluer la confiance d'une estimation, il est nécessaire de déterminer un intervalle contenant,avec une probabilité donnée, la vraie valeur de la quantité inconnue. Ce mémoire est rédigé en deux chapitres :

Chapitre 1 : Dans ce chapitre, on va présenté des généralités sur les statistiques, en donnant la définition d'une statistique, fonction de vraisemblance, statistique exhaustive et l'information de Fisher.

Chapitre 2 : Nous sommes intéressés dans ce chapitre à présenter l'estimation par les méthodes du maximum de vraisemblance, moment et par intervalle de confiance.

Chapitre 1

Notion de base sur l'estimation

L'estimation consiste à donner des valeurs approximatives aux paramètres d'une population à l'aide d'un échantillon de n observations issues de cette population. On peut se tromper sur la valeur exacte, mais on donne la " meilleure valeur " possible que l'on peut supposer.

Définitions

Echantillon

Définition 1

Soit X une variable aléatoire sur un référentiel Ω. ***Un échantillon*** *de X de taille n est un n-uplet (X_1 ;... ;X_n) de variables aléatoires indépendantes de même loi que X. La loi de X sera appelée loi mère. Une réalisation de cet échantillon est un n-uplet de réels (x_1 ;... ;x_n) où*

$X_i(w) = x_i$.

Statistique

Définition 2

On appelle statistique sur un n-échantillon une fonction de (X_1 ;... ;X_n).

Estimateur

Définition 3

*Un **estimateur** de θ sera une statistique $T = f(X_1, ..., X_n)$ et sa réalisation sera notée $t = f(x_1, ..., x_n)$*

Pour un même paramètre, il peut y avoir plusieurs estimateurs possibles. Pour pouvoir choisir, il faut définir les qualités qui font qu'un estimateur sera meilleur.

Remarque 1

*On appelle **erreur d'estimation** : $T - \theta$.*

Celle-ci peut se décomposer de la façon suivante :

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

Le terme $T - E(T)$ traduit la fluctuation de T autour de son espérance et le terme

$E(T) - \theta = B(T)$

*représente l'erreur systématique et s'appelle **Biais de l'Estimateur***

1.1 Qualité d'un estimateur

Définition 4

*On appelle **biais** de T pour θ la valeur*

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

*Un estimateur T est dit **sans biais** si $E(T) = \theta$.*

Exemple 1

Soit $(X_1, X_2, \ldots, X_n)$ un n échantillon telles que $\mathbb{E}(X_i) = m, \forall i = 1, \ldots, n$ $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ (la moyenne empirique) est un estimateur sans biais de m, en effet :

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} nE(X_1) = \frac{1}{n} nm = m$$

Définition 5

Un estimateur T de θ est dit ***asymptotiquement sans biais*** *si $E(T) \longrightarrow \theta$ ur $n \to \infty$.*

Exemple 2

Soit $(X_1, X_2, \ldots, X_n)$ un n échantillon telles que $\mathbb{V}(X_i) = \sigma^2 < \infty, \forall i = 1, \ldots, n$ $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2$, la variance empirique est un estimateur asympto- tiquement sans biais de σ^2, en effet :

$$\begin{aligned}\mathbb{E}\left(S^2\right) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \bar{X}^2\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}X_i^2 - \mathbb{E}\bar{X}^2 = \mathbb{E}X^2 - \mathbb{E}\bar{X}^2, \\ &= \left(\mathbb{V}(X) + \mathbb{E}^2(X)\right) - \left(\mathbb{V}(\bar{X}) + \mathbb{E}^2(\bar{X})\right) = \sigma^2 + m^2 - \frac{\sigma^2}{n} - m^2 = \sigma^2 - \frac{\sigma^2}{n},\end{aligned}$$

$\lim_{n\to\infty} \mathbb{E}\left(S^2\right) = \lim_{n\to\infty}\left(\sigma^2 - \frac{\sigma^2}{n}\right) = \sigma^2.$

Définition 6

L'estimateur sans biais et de variance minimale est appelé estimateur efficace.

Définition 7

Soient T et T' deux estimateurs sans biais de $\theta.T$ est dit plus efficace que T' si

$$V(T) \leq V\left(T'\right).$$

Définition 8

Un estimateur T est dit convergent si E(T) tend vers θ lorsque n tend vers l'infini. Il sera dit consistant si T converge en probabilité vers θ lorsque n tend vers l'infini.

Théorème 1.1.1

Si T est convergent et de variance tendant vers 0 lorsque n tend vers l'infini alors T est consistant.

Démonstration : On a, pour tous réels θ et $\alpha > 0$

$$|T - \theta| > \alpha \Rightarrow |T - E(T)| > \alpha - |\theta - E(T)|$$

Si lim $E(T) = \alpha$, alors à partir d'un certain rang N, on a

$$|\theta - E(T)| < \frac{\alpha}{2}$$

Ainsi

$$\begin{aligned} p(|T - \theta| > \alpha) &\leq p(|T - E(T)| > \alpha - |\theta - E(T)|) \\ &\leq p(|T - E(T)| > \frac{\alpha}{2} \\ &\leq \frac{4}{\alpha^2} Var(T) \end{aligned}$$

borne supérieure qui tend vers 0 lorsque n tend vers l'infini.

De façon générale, on peut écrire

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

ainsi

— La grandeur $T - E(T)$ représente les fluctuations de T autour de sa moyenne

— et $|E(T) - \theta|$ représente l'erreur systématique (biais). ■

1.2 Statistique exhaustive

Afin de faire de l'inférence statistique, le statisticien va devoir extraire de l'information de la suite de variables aléatoiresX_1 ;... ;X_n dont il dispose. Lorsque la taille de l'échantillon n est grande, il est naturel de tenter de réduire l'échantillon et et de résumer l'information qui y est contenue.

Lorsque il est possible de "remplacer" (X_1 ;... ;X_n) par une statistique T = T (X_1 ;... ;X_n) on optera bien sûr pour cette solution. Cependant, une question se pose :

Comment savoir si la réduction des données opérée par la statistique T ne conduit pas à une perte d'information ? C'est ce type de problèmes que cherche à résoudre la notion d'exhaustivité.

L'idée est basée sur la remarque suivante : si la loi conditionnelle de X sachant T ne dépend pas de la loi P_θ de X, alors T est suffisamment informative pour P_θ. En effet, dans cette situation la loi conditionnelle de X sachant S peut être spécifiée indépendamment de P_θ, lorsqu'on donne S, on peut générer une variable aléatoire X_0 de même loi que X. Donc les informations données par X = $(X_1\,;...\,;X_n)$ ne donnent pas plus sur P_θ que T ne le fait. T est dite alors statistique exhaustive (ou suffisante).

Définition 9

On considère un n échantillon $(X_1\,;...\,;X_n)$. On dit q'une statistique T est exhaustive pour θ si la loi conditionnelle de l'échantillon X = $(X_1\,;...\,;X_n)$ sachant $T = t$ n'est pas une fonction de θ c-à-d

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

Donc, on peut dire qu'une fois T connu, nous n'obtenons plus d'autre information de l'échantillon concernant θ et donc que T porte toute l'information disponible sur θ.

1.2.1 Théorème de factorisation

Supposons que X = $(X_1\,;...\,;X_n)$ admet une densité jointe f(x ; θ) pour $\theta \in \Theta$. Alors, T = T(X) est une statistique exhaustive pour θ si et seulement s'il existe deux fonctions mesurables $g : \mathbb{R}^p \times \Theta \to \mathbb{R}^+$ et $h : E \to \mathbb{R}^+$ telles que f(x ;θ) se met sous la forme

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

(T et θ peuvent être des vecteurs).

Exemple 3

Supposons que $X_1\,;...\,;X_n$ sont des variables. avec

$$f(x;\theta) = \frac{1}{\theta} \text{ pour } 0 \leqslant x \leqslant \infty$$

avec $\theta > 0$. La densité jointe de $X = (X_1;\ldots;X_n)$ est

$$\begin{aligned} f(x;\theta) &= \frac{1}{\theta^n} \quad pour \quad 0 \leqslant X_1;\ldots;X_n \leqslant \infty \\ &= \frac{1}{\theta^n} I(0 \leqslant X_1;\ldots;X_n \leqslant \infty) \\ &= \frac{1}{\theta^n} I(\max_{1\leq i\leq n} x_i \leq \theta) I(\min_{1\leq i\leq n} x_i \geq 0 \\ &= g(\max_i x_i;\theta)h(x) \end{aligned}$$

et$X_{(n)} = \max\limits_{1,\ldots,n}(X_i)$est une statistique exhaustive pour θ

1.3 Information de Fisher

Soit X une variable aléatoire réelle de loi p_θ ou $\theta \in \Theta$ et Θ désigne un ouvert de $\mathbb{R}$. On supposera dans cette section que les hypothèses suivantes sont satisfaites.
H_1 :$\forall\theta \in \Theta$, $\forall x \in \chi$, $f(x;\theta) > 0$.
H_2 :$\forall\theta \in \Theta$, $\forall x \in \chi$, $\dfrac{\partial f}{\partial\theta}(x,\theta)$ et $\dfrac{\partial^2 f}{\partial\theta^2}(x,\theta)$ existent.
H_3 :$\forall\theta \in \Theta$,$\forall B \in \beta$ on peut dériver 2 fois $\int_B f(x;\theta)dx$ par rapport à θ sous le signe d'intégration ; autrement dit on peut échanger les opérateurs d'intégration et de dérivation

Définition 10

Soit le modèle $((\chi,\beta)\,;\, p_\theta,\ \forall\theta \in \Theta)$, La quantité

$$I(\theta) = E(\frac{\partial}{\partial\theta} log f(X,\theta)^2)$$

s'appelle l'information de Fisher au point θ

A noter que l'espérance est prise par rapport à la loi p_θ de X.

1.3.1 Autres expressions de l'information de Fisher

On appelle score la quantité S(x ;θ) =$\frac{\partial}{\partial\theta}logf(x,\theta)$

$$\text{I}(\theta)\text{=E}[S^2]$$

Sous les hypothèsesH_1, H_2 et H_3 on a : I(θ) = -E($\frac{\partial^2}{\partial\theta^2}logf(x,\theta)$) L'intérèt de cette proposition est que cette expression de I(θ) est plus facile à manipuler que celle de la définition.

Proposition 1.3.1

Sous les hypothèses H_1, H_2 etH_3 on a :

$$I(\theta) = \text{-}E(\frac{\partial^2}{\partial\theta^2}logf(x,\theta))$$

L'intérét de cette proposition est que cette expression de I(θ) est plus facile à manipuler que celle de la définition.

Proposition 1.3.2

$I(\theta) \geqslant 0$ pour tout θ

Proposition 1.3.3

Soient x et y deux v.a.r. à valeurs dans X et Y, indépendantes, et de loi respective p_θ et Q_θ. On note $I_{(x;y)}(\theta)$, $I_x(\theta)$ et $I_y(\theta)$ les informations de Fisher au point θ respectivement apportées par x, y , et (x ; y)
Dans ces conditions on a :

$$I_{(x;y)}(\theta) = I_x(\theta) + I_y(\theta)$$

Conséquence. Soit X une variable aléatoire à valeurs dans ($\mathcal{X}$; $\mathcal{B}$) de loi p_θ. L'information de Fisher (au point θ fourni par le modèle d'échantillonnage ($\mathcal{X}$; $\mathcal{B}$; p_θ ; $\theta \in \Theta)^n$ vérifie :

$$I_{(X_1,\ldots,X_n)}\text{=n}I_X(\theta).$$

1.4 Relation entre l'information de Fisher et la statistique exhaustive

On considère le modèle d'échantillonnage $(\mathcal{X}\,;\,\mathcal{B}\,;\,p_\theta\,;\,\theta\in\Theta)^n$. Soit T une statistique définie sur $\mathcal{X}^n$ à valeurs dans un espace mesurable Y. On a l'equivalence suivante qui relie les notions d'exhaustivité et d'information de Fisher :

$$I_{(X_1\dots,X_i,\dots,X_n)}(\theta)=I_T(\theta)\Leftrightarrow \text{T est exhaustive}$$

$I_T(\theta)$ représente l'information de Fisher au point θ

Théorème 1.4.1

Pour toute statistique T on a :

$$I_T(\theta)\leqslant I(\theta)$$

où $I_T(\theta)$ est l'information de Fisher au point θ

Exemple 4

Soient X_1 ;... ;X_n un échantillon qui suit $\aleph(m,\delta^2)$: Considérons la statistique

$$T=S'^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

on définit

$$Y_n=(n-1)\,\frac{T}{\delta^2}$$

telles que $Y_n\sim(\chi_{n-1})^2$, d'après le théorème de Fisher.
Alors

$$f_n(y)=\frac{1}{2^{\frac{n-1}{2}}T(\frac{n-1}{2})}y^{\frac{n-3}{2}}exp(-\frac{y}{2})I_{y\geq 0}$$

On fait un changement de variable, $Y_n\to T$, on obtient :

$$f_T(t,\delta^2)=\frac{1}{2^{\frac{n-1}{2}}T(\frac{n-1}{2})}\frac{n-1}{\delta^2}t^{\frac{n-3}{2}}exp(-\frac{(n-1)t}{2\delta^2})I_{t\geq 0}$$

On calcule l'information de Fisher

$$lnf_T(t,\delta^2) = c(t) - \frac{(n-1)t}{2\delta^2} + \frac{n-1}{2} ln(\frac{n-1}{\delta^2})$$

où c(t) est une constante qui ne dépend pas de δ

$$\frac{\partial^2}{\partial\delta^2} lnf_T(t,\delta^2)(t,\delta^2) = \frac{n-1}{2\delta^4} - \frac{(n-1)t}{\delta^6}$$

donc

$$I_T(\delta^2) = \frac{n-1}{\delta^6}\mathbb{E}(T) - \frac{n-1}{2\delta^4} = \frac{n-1}{\delta^4} - \frac{n-1}{2\delta^4} = \frac{n-1}{2\delta^4}$$

D'autre part, on a :

$$I_n(\delta^2) = nI(\delta^2) = \frac{n}{2\delta^4}$$

Il s'en suit que pour une taille d'échantillon finie n, la variance empirique, T n'est pas exhaustive pour δ^2 puisque $I_T(\delta^2) < I_n(\delta^2)$

1.5 Estimateur sans biais de variance minimale

Il existe plusieurs théorèmes qui montrent que l'estimateur de variance minimale est lié à l'existence d'une statistique exhaustive.

Théorème 1.5.1 (*unicité*)

S'il existe un estimateur de θ sans biais de variance minimale, il est unique presque sûrement.

Théorème 1.5.2 (*Rao-Blackwell*)

Soit T un estimateur sans biais de θ quelconque et U une statistique exhaustive pour θ. Alors $T^* = \mathbb{E}(T/U)$ est un estimateur sans biais de θ au moins aussi bon que T.

Théorème 1.5.3

S'il existe une statistique exhaustive U de θ, alors l'estimateur T sans biais de θ de variance minimale (unique d'après le théorème précédent) ne dépend que de U

Définition 11

On dit q'une statistique exhaustive U d'un modèle statistique paramétrique est complète si pour toute fonction borélienne $h : \mathbb{R}^d \to \mathbb{R}$ telles que h est intégrable et

$$\mathbb{E}[h(U)] = 0 \quad \forall\theta \in \Theta \Rightarrow h = 0 \ p.s$$

Théorème 1.5.4 (*Lehmann-Scheffé*)

Si T^* est un estimateur sans biais de θ dépendant d'une statistique exhaustive complète U alors T^* est l'unique estimateur sans biais de variance minimale. En particulier si l'on dispose déja de T estimateur sans biais de θ.

$$T^* = \mathbb{E}(T/U)$$

Inégalité de Fréchet-Darmois-Cramer-Rao(FDCR)

Le résultat suivant nous indique qua 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 :

$$\mathbb{V}(T) \geqslant \frac{1}{I_n(\theta)}$$

et si T est un estimateur sans biais de k(θ) :

$$\mathbb{V}(T) \geq \frac{[k'(\theta)]^2}{I_n(\theta)}$$

où k est une fonction dérivable.

Démonstration : Voir [1], page 302 . ■

Définition 12

On dit que T est un estimateur efficace si :

1. T est un estimateur sans biais de θ alors

$$\mathbb{V}(T) = \frac{1}{I_n(\theta)}$$

2. T est un estimateur sans biais de $k(\theta)$ alors

$$\mathbb{V}(T) = \frac{[k'(\theta)]^2}{I_n(\theta)}$$

Théorème sur l'efficacité

-La borne de Cramer-Rao ne peut être atteinte que si la loi de X est de forme exponentielle :

$$f(x, \theta) = \exp[a(x)\alpha(\theta) + b(x) + \beta(\theta)]$$

car T est nécessairement exhaustive pour θ. donc, il n'existe qu'une seule fonction $k(\theta)$ qui puisse être estimée efficacement :

$$k(\theta) = -\frac{\beta'(\theta)}{\alpha'(\theta)}$$

l'estimateur de $k(\theta)$ est

$$T_n = \frac{1}{n}\sum_{i=1}^{n} a(X_i)$$

la variance minimale est :

$$\mathbb{V}(T_n) = -\frac{1}{n\alpha'(\theta)}\frac{\partial}{\partial\theta}\left(\frac{\beta'(\theta)}{\alpha'(\theta)}\right) = \frac{k'(\theta)}{n\alpha'(\theta)}$$

Démonstration : Voir [1], page 303-304. ■

Exemple 5

Loi gamma de paramètre θ On a vu que : $T = \sum_{i=1}^{n} \ln X_i$ est exhaustive On a

$$f(x, \theta) = \exp[(\theta - 1)\ln x - x - \ln \Gamma(\theta)]$$

donc la loi gamma appartient à la famille exponentielle telles que $a(x) = \ln x, \alpha(\theta) = \theta - 1, \beta(\theta) = -\ln \Gamma(\theta), b(x) = -x$.
D'après le théorème de Darmois $T = \sum_{i=1}^{n} a(X_i) = \sum_{i=1}^{n} \ln X_i = \ln\left(\prod_{i=1}^{n} X_i\right)$ est exhaustive.
Donc d'après le théorème précédent la seule fonction qui puisse être estimée efficacement est

$$k(\theta) = -\frac{\beta'(\theta)}{\alpha'(\theta)} = \frac{\Gamma'(\theta)}{\Gamma(\theta)} = \frac{\partial}{\partial \theta} \ln \Gamma(\theta)$$

Chapitre 2

Les méthodes d'estimation paramétrique

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

2.1 La méthode du maximum de vraisemblance

Un des estimateurs les plus utilisés en statistique est l'estimateur du maximum de vraisemblance. La vraisemblance est une fonction qui contient toute l'information des données sur un paramètre inconnu. Elle joue un rôle important dans de nombreuses méthodes statistiques. Et l'estimateur du maximum de vraisemblance a de très bonnes propriétés d'optimalité.

Définition 13

Soient $X_1, \cdots, X_n$ des variables aléatoires de densité $f(\mathbf{x};\theta)$ avec $\theta \in \Theta$. Et considérons $\mathbf{x} = (x_1, \cdots, x_n)$ une réalisation du vecteur $\mathbf{X} = (X_1, \cdots, X_n)$. La fonction de vraisemblance [1] est définie par

$$\mathcal{L}(\theta) = f(\mathbf{x};\theta)$$

C'est une fonction réelle définie sur l'espace des paramètres Θ. La vraisemblance donne, en quelque sorte, la probabilité que la réalisation $\mathbf{x} = (x_1, \cdots, x_n)$ soit émise par le modèle associé à la valeur θ du paramètre. Ainsi plus il est probable que le modèle émette cette réalisation pour la valeur θ, plus la vraisemblance sera grande.

Définition 14

Soit $\mathbf{X} = (X_1, \cdots, X_n)$ *un vecteur aléatoire suivant le modèle* $F_\theta, \theta \in \Theta$. *Pour une réalisation* $\mathbf{x} = (x_1, \cdots, x_n)$, *l'estimateur du maximum de vraisemblance (EMV) est l'estimateur* $\hat{\theta} = S((X))$ *avec* S *telle que*

$$\mathcal{L}(S(x)) \geq \mathcal{L}(\theta) \text{ pour tout } \theta \in \Theta$$

Exemple 6

pour une loi discrète - Supposons que $X_1, \cdots, X_n$ *sont des variables aléatoires i.i.d de loi de Bernouilli de paramètre inconnu* $\theta \in]0, 1[$. *Et notons* $f(\cdot; \theta)$ *la densité de* X_i *pour tout* $i = 1, \cdots, n$. *On a* $f(x; \theta) = \theta^x (1-\theta)^{1-x}$

Soit $(x_1, \cdots, x_n)$ *une réalisation de* $(X_1, \cdots, X_n)$. *Par définition, la vraisemblance est*

$$\mathcal{L}(\theta; x_1, \cdots, x_n) = \prod_{i=1}^{n} f_\theta(x_i)$$

car les variables X_i *sont indépendantes. D'où*

$$\mathcal{L}(\theta; x_1, \cdots, x_n) = \prod_{i=1}^{n} \theta^{x_i} (1-\theta)^{1-x_i} = \theta^{n_{i=1}^{n} x_i} (1-\theta)^{n-\sum_{i=1}^{n} x_i}$$

On en déduit la log-vraisemblance :

$$\ln \mathcal{L}(\theta; \mathbf{x}) = \sum_{i=1}^{n} x_i \ln \theta + \left(n - \sum_{i=1}^{n} x_i\right) \ln(1-\theta)$$

Le maximum est la racine de la dérivée de cette expression dont la dérivée seconde est négative.

$$\frac{\partial \ln \mathcal{L}(\theta; \mathbf{x})}{\partial \theta} = \frac{\sum_{i=1}^{n} x_i}{\theta} + \frac{n - \sum_{i=1}^{n} x_i}{1-\theta}$$

s'annule en $\theta^* = \dfrac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}_n$. *On vérifie que*

$$\frac{\partial^2 \ln \mathcal{L}(\theta; \mathbf{x})}{\partial \theta^2} < 0$$

au voisinage de $\theta^* = \bar{x}_n$. *Et on conclut que l'estimateur du maximum de vraisemblance pour le paramètre de la loi de Bernoulli est* $\hat{\theta}_n = \bar{X}_n$.

Proposition 2.1.1

Soit ϕ telle que $\phi = g(\theta)$ pour une fonction bijective g. Si $\hat{\theta}$ est un EMV pour θ alors $\hat{\phi} = g(\hat{\theta})$ est un EMV pour ϕ.

2.2 Propriétés asymptotiques de l'EMV

Nous allons voir dans cette section, que sous des conditions faibles de régularité, on peut montrer que l'EMV est consistant et asymptotiquement normal. Dans la suite de cette section, nous supposons que $X_1, \cdots, X_n$ sont des v.a.i.i.d. de densité $f(x;\theta)$ avec $\theta \in \Theta \subset \mathbb{R}$ avec $\ell(x;\theta) = \ln f(x;\theta)$ trois fois différentiable par rapport à θ. Et nous ajoutons les hypothèses suivantes :

(A1) Θ est un sous ensemble ouvert de $\mathbb{R}$.

(A2) L'ensemble $A = \{x : f(x;\theta) > 0\}$ ne dépend pas de θ.

(A3) $f(x;\theta)$ est trois fois continument différentiable par rapport à θ sur A.

(A4) $E_\theta\left[\ell'\left(X_i;\theta\right)\right] = 0$ pour tout θ et $\mathrm{Var}_\theta\left[\ell'\left(X_i;\theta\right)\right] = I(\theta)$ où $0 < I(\theta) < \infty$ pour tout θ.

(A5) $E_\theta\left[\ell''\left(X_i;\theta\right)\right] = -J(\theta)$ où $0 < J(\theta) < \infty$ pour tout θ

(A6) Pour θ et $\delta > 0, |\ell'''(x;t)| < M$ pour $|\theta - t| \leq \delta$ où $E_\theta\left[M\left(X_i\right)\right] < \infty$.

Définition 15

Si le modèle (E, P_θ) vérifie les hypothèses (A1)à(A3) et que l'intégrale $\int_B f(x;\theta)d\mu(x)$ est au moins deux fois dérivable sous le signe d'intégration pour tout borélien B, aors on dit que le modèle est un modèle régulier.

Exemple 7

soient $X_1, \ldots X_n$ sont de loi Géométrique de paramètre p, alors la fonction de vraisemblance du modèle est

$$\mathcal{L}\left(p, \{x_1, \ldots, x_n\}\right) = \prod_{i=1}^{n} p(1-p)^{x_i - 1} = p^n(1-p)^{\sum x_i - n}.$$

La log-vraisemblance vaut

$$\log \mathcal{L}\left(p, \{x_1, \ldots, x_n\}\right) = n\log(p) + \left(\sum x_i - n\right)\log(1-p)$$

L'équation du premier ordre s'écrit

$$0 = \frac{\partial \log \mathcal{L}\left(p, \{x_1, \ldots, x_n\}\right)}{\partial p} = \frac{n}{p} - \frac{\sum x_i - n}{1-p}$$

ou encore $1/p = \left(\overline{x_n} - 1\right)/(1-p)$ *dont la solution est* $p = 1/\overline{x_n}$. *L'estimateur du maximum de vraisemblance de* p *est donc* $\widehat{p}_n = 1/\overline{X_n}$.

Exemple 8

Estimation de l'esperance d'une loi normale $\mathcal{N}\left(m, \sigma^2\right)$; σ *connu*

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

d'où

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

alors

$$\frac{\partial \ln L\left(x_1, \ldots, x_n; m\right)}{\partial m} = \sum_{i=1}^{n} \frac{\left(x_i - m\right)}{\sigma^2}$$
$$\frac{\partial \ln L\left(x_1, \ldots, x_n; m\right)}{\partial m} = 0 \Longrightarrow \hat{m} = \bar{x}$$

et

$$\frac{\partial^2 \ln L\left(x_1, \ldots, x_n; m\right)}{\partial m^2} = -\frac{n}{\sigma^2}$$
$$\left.\frac{\partial^2 \ln L\left(x_1, \ldots, x_n; m\right)}{\partial^2 m}\right|_{m=\hat{m}} = -\frac{n}{\sigma^2} < 0,$$

alors l'EMV de m *est :*

$$\hat{m} = \bar{X}$$

2.3 La méthode des moments

La méthode des moments est relativement intuitive. Il semble naturel d'estimer les moments par leur version empirique :

— L'esperance $\mathbb{E}(X)$ correspond a une moyenne théorique et peut être estimée par la moyenne observée $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$

– La variance $\mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}\left(X^2\right) - (\mathbb{E}X)^2$ peut être estimée par la variance observée $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X_n}\right)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\overline{X_n}\right)^2$

Lorsque l'on souhaite estimer un paramètre θ, on essaie de l'exprimer en fonction de $\mathbb{E}X$ et de $\mathrm{Var}(X)$: $\theta = g(\mathbb{E}X, \mathrm{Var}(X))$. On sait ensuite que $\mathbb{E}X$ et de $\mathrm{Var}(X)$ peuvent être estimés respectivement par $\overline{X_n}$ et S_n^2. On propose donc d'estimer θ par $\widehat{\theta}_n = g\left(\overline{X_n}, S_n^2\right)$ Ce principe se généralise avec l'ensemble des moments d'ordre $p, p \in \mathbb{N}^*$.

Exemple 9

- Si $X_1, \ldots X_n$ sont de loi Géométrique de paramètre p, alors $\mathbb{E}(X_i) = \frac{1}{p}$. L'estimateur des moments de p est donc $\hat{p}_n = 1/\overline{X_n}$.

- Si $X_1, \ldots X_n$ sont de loi uniforme sur l'intervalle $[a; b]$ alors, $\mathbb{E}(X_i) = (a + b)/2$ et $\mathrm{Var}\,(X_i) = (b - a)^2/12$.

Par conséquent $a = \mathbb{E}(X_i) - \sqrt{3\,\mathrm{Var}\,(X_i)}$ et $b = \mathbb{E}(X_i) + \sqrt{3\,\mathrm{Var}\,(X_i)}$. Les estimateurs des moments de a et b sont donc $\widehat{a}_n = \overline{X_n} - \sqrt{3}S_n$ et $\widehat{b}_n = \overline{X_n} + \sqrt{3}S_n$

2.4 La méthode d'estimation par intervalle de confiance

Nous avons vu a 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 probabilité forte de se trouver.

Définition 16

Soit $(X_1, \ldots, X_n)$ un échantillon de loi $\mathbb{P}$. On appelle intervalle de confiance (IC) de niveau de confiance $1 - \alpha$ telles que $\alpha \in [0, 1]$ donné, un intervalle aléatoire $[\theta_1, \theta_2]$ où $\theta_1 \leq \theta_2$ sont deux statistiques, fonction de l'échantillon, telles que.

$$\mathbb{P}\left(\theta_1 \leq \theta \leq \theta_2\right) = 1 - \alpha.$$

Remarque 2

α est donc la probabilité que $[\theta_1, \theta_2]$ ne recouvre pas la vraie valeur du paramètre.

2.4.1 Estimation par intervalle de confiance pour les paramètres de la loi normale

Soient $X_1, \ldots, X_n$ indépendantes et de même loi $\mathcal{N}\left(m, \sigma^2\right)$. Nous avons vu précédemment comment estimer m et σ^2. A l'aide du théorème de Fisher, nous pouvons déterminer la précision des estimateurs que nous avons construits et nous en déduirons ainsi des intervalles de confiance.

Théorème 2.4.1

Si $(X_1, \ldots, X_n)$ est un échantillon de la loi $\mathcal{N}\left(m, \sigma^2\right)$, alors :

1. $\dfrac{\bar{X} - m}{\frac{\sigma}{\sqrt{n}}}$ *suit la loi normale $\mathcal{N}(0,1)$ (d'après le théorème centrale limite (TCL)).*
2. $\dfrac{\bar{X} - m}{\frac{S}{\sqrt{n-1}}}$ *suit une loi de Student $\mathcal{T}_{n-1}$ de degré de liberté $(n-1)$.*
3. $\dfrac{nS^2}{\sigma^2}$ *suit la loi du khi-deux $\mathcal{X}^2_{n-1}$, degré de liberté $(n-1)$*

Estimation par intervalle de confiance de la moyenne lorsque la variance est connue

On sait que $\bar{X} \sim \mathcal{N}\left(m, \dfrac{\sigma^2}{n}\right)$, on a aussi $\bar{X}$ est le meilleur estimateur de m, et $\dfrac{\sqrt{n}(\bar{X} - m)}{\sigma} \sim \mathcal{N}(0,1)$ alors

$$\mathbb{P}\left(-u_{1-\frac{\alpha}{2}} \leq \frac{\sqrt{n}(\bar{X} - m)}{\sigma} \leq u_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$$

$$\mathbb{P}\left(-u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq (\bar{X} - m) \leq u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

$$\mathbb{P}\left(\bar{X} - u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X} + u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha$$

donc l'intervalle de confiance est :

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

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

Exemple 10

Après des essais antérieurs, on peut supposer que la résistance à l'écla- tement d'un certain type de réservoirs est une variable aléatoire suivant une loi normale $\mathcal{N}\left(m,\sigma^2\right)$ *où* $\sigma = 4$ kg/cm^2. *Des essais sur un échantillon de 25 réservoirs donnent une résistance moyenne à l'éclatement égale à* 300 kg/cm^2

Donc $n = 25, \bar{X} = 300$ kg/cm$^2, \sigma = 4$ kg/cm^2, *le niveau de confiance :* $1-\alpha = 0.95, u_{1-\frac{\alpha}{2}} = 1.\ 96$.

L'intervalle de confiance

$$\mathbb{P}\left(-1.96 \leq \frac{5(\bar{X}-m)}{4} \leq 1.96\right) = 0.95$$

donc

$$\mathbb{P}\left(300 - 1.96\frac{4}{5} \leq m \leq 300 + 1.96\frac{4}{5}\right) = \mathbb{P}(298.432 \leq m \leq 301.568) = 0.95$$

$$IC(m) = [298.432, 301.568]$$

l'intervalle $[298.432, 301.568]$ *a une probabilité égale à* 0.95 *de contenir la vraie valeur de la résistance à l'éclatement de ce type de réservoirs.*

Remarque 3

Si $n \geq 30$ *et* σ *est inconnu en remplace* σ *par* S' *telles que* $S' = \sqrt{\frac{n}{n-1}}S$.

Estimation par intervalle de confiance de la moyenne lorsque la variance 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 $\sqrt{n}\left(\frac{\bar{X}_n-\mu}{\sigma}\right)$ mais plutôt

$$\sqrt{n}\left(\frac{\bar{X}_n-\mu}{S_n}\right)$$

qui ne suit plus une loi normale mais une loi dite de Student à $n-1$ degrés de liberté, que l'on note $\mathcal{T}_{n-1}$. La densité de la loi de Student est une fonction paire, comme la loi normale $\mathcal{N}(0,1)$. On dispose de tables pour obtenir les quantiles de cette loi. On en déduit donc que

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

ce qui équivaut à

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

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

$$\left[\bar{X}_n - t_{1-\alpha/2}\frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{1-\alpha/2}\frac{S_n}{\sqrt{n}}\right]$$

Ainsi, dans les calculs, l'IC est donné par

$$IC_{1-\alpha}(\mu) = \left[\bar{x}_n - t_{1-\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{1-\alpha/2}\frac{s_n}{\sqrt{n}}\right]$$

où $\bar{x}_n$ et s_n^2 sont les estimations ponctuelles respectives de la moyenne μ et de la variance σ^2.

Exemple 11

Sur un échantillon de taille $n = 20$ *durées de vie d'un certain modèle de lampe on a obtenu comme moments empiriques* $\bar{X} = 2000$ h *et* $S = 300$ h *L'intervalle de confiance de niveau* 0.95 *pour la durée de vie moyenne* m *est donc*

$$\mathbb{P}\left(2000 - t_{1-\frac{a}{2}}\frac{300}{\sqrt{19}} \leq m \leq 2000 + t_{1-\frac{\alpha}{2}}\frac{300}{\sqrt{19}}\right) = 0.95$$

où $t_{1-\frac{a}{2}} = 2.093$ *d'où l'intervalle de confiance de* m *au niveau de confiance* 0.95 *est :*

$$\mathbb{P}(1855.95 \leq m \leq 2144.05) = 0.95$$

donc

$$IC(m) = [1855.95, 2144.05]$$

Si σ^2 est connue, soit $\sigma = 300$ alors l'intervalle de confiance est définie par :

$$\mathbb{P}\left(2000 - u_{1-\frac{a}{2}} \frac{300}{\sqrt{20}} \leq m \leq 2000 + u_{1-\frac{a}{2}} \frac{300}{\sqrt{20}}\right) = 0.95$$

avec $u_{1-\frac{a}{2}} = 1.96$ donc l'intervalle de confiance de m au niveau de confiance 0.95 est :

$$\mathbb{P}(1868.52 \leq m \leq 2131.48) = 0.95$$

$$IC(m) = [1868.52, 2131.48]$$

L'intervalle obtenu est plus grand ; la connaissance du paramètre σ conduit logiquement à un intervalle plus précis.

Estimation par intervalle de confiance de la variance Si la moyenne est connue

$T = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2$ est le meilleur estimateur de la variance σ^2 et on a $\frac{nT}{\sigma^2} \sim \mathcal{X}_n^2$ comme somme de n carrées de $\mathcal{N}(0,1)$ indépendantes, alors :

$$\mathbb{P}\left(k_1 \leq \frac{nT}{\sigma^2} \leq k_2\right) = 1 - \alpha$$

$$\mathbb{P}\left(\frac{k_1}{nT} \leq \frac{1}{\sigma^2} \leq \frac{k_2}{nT}\right) = 1 - \alpha$$

$$\mathbb{P}\left(\frac{nT}{k_2} \leq \sigma^2 \leq \frac{nT}{k_1}\right) = 1 - \alpha$$

donc l'intervalle de confiance de σ^2 est :

$$IC\left(\sigma^2\right) = \left[\frac{nt}{k_2}, \frac{nt}{k_1}\right]$$

où k_1 est le quantile d'ordre $\frac{\alpha}{2}$ et k_2 est le quantile d'ordre $\left(1 - \frac{\alpha}{2}\right)$ de la loi $\mathcal{X}_n^2$.

Exemple 12

Soit X une v.a suivant la loi normale $\mathcal{N}(30, \sigma)$, on prélève un échan- tillon de taille $n = 25$. Cherchons un intervalle de confiance de la variance au niveau de confiance $1 - \alpha = 0.95, n = 25$ alors $\alpha = 0.05; \frac{\alpha}{2} = \frac{0.05}{2} = 0.025$ donc $k_1 = \mathcal{X}_{0.025}^2 = 13.120; k_2 = \mathcal{X}_{0.975}^2 = 40.644$ et $t = 15$

$$\mathbb{P}\left(\frac{25 \times 15}{40.644} \leq \sigma^2 \leq \frac{25 \times 15}{13.120}\right) = \mathbb{P}\left(9.23 \leq \sigma^2 \leq 28.58\right) = 0.95$$

donc

$$IC\left(\sigma^2\right) = [9.23, 28.58]$$

Estimation par intervalle de confiance de la variance Si la moyenne est inconnue

On utilise $S^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$ et on sait que $\frac{nS^2}{\sigma^2} \sim \mathcal{X}^2_{n-1}$,

$$\mathbb{P}\left(l_1 \leq \frac{nS^2}{\sigma^2} \leq l_2\right) = 1 - \alpha$$
$$\mathbb{P}\left(\frac{l_1}{nS^2} \leq \frac{1}{\sigma^2} \leq \frac{l_2}{nS^2}\right) = 1 - \alpha$$
$$\mathbb{P}\left(\frac{nS^2}{l_2} \leq \sigma^2 \leq \frac{nS^2}{l_1}\right) = 1 - \alpha$$

donc l'intervalle de confiance de σ^2 est :

$$IC\left(\sigma^2\right) = \left[\frac{nS^2}{l_2}, \frac{nS^2}{l_1}\right]$$

où l_1 est le quantille d'ordre $\frac{\alpha}{2}$ et l_2 d'ordre $1 - \frac{\alpha}{2}$ de la loi $\mathcal{X}^2_{n-1}$

Exemple 13

$n = 30; S^2 = 12; 1 - \alpha = 0.90$ *alors* $l_1 = \mathcal{X}^2_{0.05} = 17.7$ *et* $l_2 = \mathcal{X}^2_{0.95} = 42.6$ *donc*

$$8.46 \leq \sigma^2 \leq 20.33$$

Remarque 4

Si $n > 30$ *donc on a les deux approximations suivantes :*

$$\sqrt{2}\mathcal{X}^2_p - \sqrt{2p-1} \rightarrow \mathcal{N}(0,1) \text{ si } p > 30$$

approximation de Fisher et

$$\mathcal{X}^2_p = p\left(u\sqrt{\frac{2}{9p}} + 1 - \frac{2}{9p}\right)^3$$

approximation de wilson Hilferty valable même pour les valeurs faible de p.

Exemple 14

L'ecart type de la durée du vie d'un échantillon de 200 ampoules élec- triques est égale à 100 h, *calculons les limites de confiance à* 95% *pour l'écart type de toute la population. On a :*

$$\frac{200S^2}{\sigma^2} \sim \mathcal{X}^2_{199}$$

$$\mathbb{P}\left(l_1 \leq \frac{nS^2}{\sigma^2} \leq l_2\right) = 0.95$$

on utilisant l'approximation de Fisher : $l_1 = \mathcal{X}^2_{0.025} = \frac{1}{2}(-1.96 + \sqrt{2 \times (199) - 1})^2 =$ 162 *et* $l_2 = \mathcal{X}^2_{0.975} = \frac{1}{2}(1.96 + \sqrt{2 \times (199) - 1})^2 =$ *239*

$$\mathbb{P}\left(\frac{nS^2}{l_2} \leq \sigma^2 \leq \frac{nS^2}{l_1}\right) = 0.95$$

donc les limites de confiance à 95% *pour l'écart type sont :*

$$2 \leq \sigma \leq 111.3$$

2.5 Simulation avec le logiciel R :

Dans cette section,on utilise le logiciel R pour simulées des données numériques et éstimé les moyennes de deux populations dans différents niveaux de confiances, plus une illustration graphiques.

2.5.1 Moyenne d'une population

On préléve des échantillons de taille $n = 100$ de deux populations,l'une suit la loi exponentielle de paramètre θ et l'autre suit la loi de Poisson (λ), on prend $\lambda = \theta = 1$. On fait varier le niveau de confiance $1 - \alpha = 0.90$; 0.95 et 0.99. Les résultats sont résumés dans les tableaux :

Exponentielle θ			
$1-\alpha$	Borne Inf	$\frac{1}{\widehat{\overline{\theta}}}$	Borne Sup
0.90	0.778	0.957	1.135
0.95	0.744	0.957	1.170
0.99	0.675	0.957	1.239

Table 2.1 – **Intervalles de confiance de diffŕents niveaux,pour la moyenne d'une population exponentielle**

Poisson (λ)			
$1-\alpha$	Borne Inf	$\widehat{\lambda}$	Borne Sup
0.90	0.864	1.050	1.236
0.95	0.827	1.050	1.273
0.99	0.675	1.050	1.345

Table 2.2 – **Intervalles de confiance de diffŕents niveaux,pour la moyenne d'une population de poisson**

2.5.2 Graphe de moyenne d'une population :

On prélève des échantillons de taille n variant de 1 à 500 de deux populations,l'une exponentielle et l'autre de Poisson,de paramètres respectifs θ et λ égaux à 1. Le comportement des estimateurs des deux moyennes ainsi que celui des bornes de confiance,au niveau 0.95 ; est décrit par la figure :

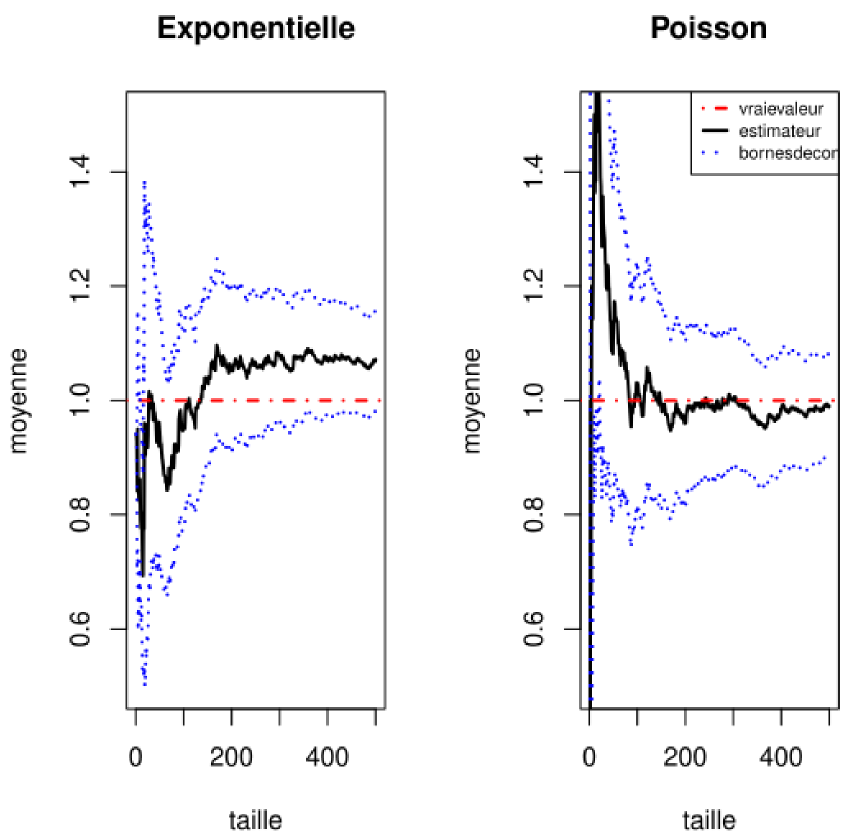

Figure 2.1 – **Intervalles de confiance,de niveau 95% ; pour les moyennes d'une population exponentielle(à gauche)et d'une v.a de Poisson(à droite)**

Annexe A : Logiciel R

Le logiciel R est un logiciel de statistique crée par Ross Ihaka et Robert Gentleman. Il est à la fois un langage informatique et un environnement de travail : les commandes sont exécutées grâce a des instructions codées dans un langage relativement simple, les résultats sont affiches sous forme de texte et les graphiques sont visualises directement dans une fenêtre qui leur est propre. C'est un clone du logiciel S-plus qui est fondé sur le langage de programmation orienté objet S, développée par AT T Bell Laboratories en 1988. Ce logiciel sert à manipuler des données, à tracer des graphiques et à faire des analyses statistiques sur ces données.

Annexe B : Codes R

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

Annexe C : Abréviations et Notations

Abréviations et Notations	Signification
v.a	Variable aléatoire.
$\mathbb{E}(X)$	Espérance mathématique ou moyenne du v.a X.
$\mathbb{V}(X)$	Variance du v.a X.
i.i.d	Indépendantes identiquement distribuées.
C-à-d	C'est à dire.
B (n ; θ)	Le biais de l'estimateur T_n.
$I_n(\theta)$	Information de Fisher.
EMV	Estimateur du maximum de vraisemblance.
IC	Intervalle de confiance.

Table des figures

Liste des tableaux

Bibliographie

[1] p.Dusart *Cours de Statistiques inférentielles*, Licence 2-S4 SI-MASS,Année 2018.

[2] L.Grammont, *COURS DE STATISTIQUES INFERENTIELLES*, Licence d'économie et de gestion, September 19, 2003.

[3] V. Monbet, *Master Statistique et econométrie Notes de cours*, Master 1 - 2011.

[4] G. Saporta, *analyse des données et statistique*, Probabilités, Technip, Paris.