République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté en vue de l'obtention du Diplôme **:**

MASTER en Mathématiques

Option : **Statistique**

Par

Hachani Nadia

Titre :

Estimation par la méthode du maximum de vraisemblance

Membres du Comité d'Examen :

Dr.	**Ouanoughi Yasmina**	UMKB	Encadreur
Dr.	**Djaber Ibtissem**	UMKB	Président
Dr.	**Berkane Hassiba**	UMKB	Examinateur

Juin 2019

Dédicace

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

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

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

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

"Ma mère".

À ceux qu'ont été prés de moi, qui ont grandis avec moi dans le bon et le mal mon frère et ma soeur : Samir et Siham.

À Mes amis : Amina Djebba, Salma Boudib et Randa slatnia.

À mon présent et mon future, à mon bonheur : Mon Marie.

À ma vie, mon âme et mon coeur :

À mon fils.

REMERCIEMENTS

Je tient à remercier toute personne qui a constribué au succes de mon travail.

Je présente mes sincères remerciments à mon encadreur "Ouanoughi Yasmina" qui a m'exprimé toute ma

reconnaissance pour ses précieux conseils et son soutien constant pendant la réalisationde cette recherche.

J'adresse mes remerciments aux président et membres des jury qui ont accepté d'examiner ce mémoire.

À tous les enseignants et le personnel du départiment de mathématique.

Je veux remercier Mes collègue de la promotion 2019 pour leur encouragement et leur sentiment, je vous souhaite que le succés.

À tous MERCI

Table des matières

Liste des tableaux

Introduction

L'estimation paramétrique est une méthode statistique qui permet d'estimet les valeurs des paramètres basés sur des données empiriques mesurées et elle est utilisé pour estimer les paramètres de la société, l'objectif est de déterminer la valeur du paramètre inconnu. La procédure d'utilisation des informations obtenues à partir d'un échantillon qui permet de déduire des résultats concernant l'ensemble de la population est appelée estimation. La valeur inconnue d'une population à estimer à partir d'un échantillon est appelée un paramètre qui peut être une moyenne, un total, un pourcentage, un écart-type ou une variance. Le paramètre de la population est estimé à partir d'une statistique calculée sur la base d'un échantillon. Soit X *un certain caractère de distribution* $f(x, \theta)$ *connue et qui dépend d'un paramètre inconnue* θ*, on cherche à estimer* θ*. Pour cela, on tire un échantillon aléatoire de taille* n *dans la population, et on essaie à partir de l'information obtenue, de déterminer une valeur numérique précise qui sera prise comme valeur du paramètre* θ *inconnu.*

L'objectif de notre travail est d'étudier l'estimation par la méthode du maximum de vraisemblance qui est une méthode pour inférer les paramètres de la loi de probabilité d'un échantillon donnée pour chercher l'estimateur inconnu, dont nous allons donner les propriétés essentielles puis nous étudierons principalement les statistiques exhaustives, la quantité d'information apporté par un échantillon de taille n*. Pour cela on divise ce mémoire en trois 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 et leurs caractéristiques.*

Chapitre 2 : *Ce chapitre présente la qualité d'un estimateur qui étudie quelques estimateurs, en donnant la définition d'un estimateur, estimateur sans biais, estimateur convergent, estimateur asymptotiquement normale et estimateur efficace.*

Chapitre 3 : *Nous sommes intéressés dans le dernier chapitre à présenter l'estimation par la méthode du maximum de vraisemblance. On a étudié la méthode pour estimer le paramètre inconnu θ et l'on a quelques propriétés. En donnant les sections : Définition du maximum de vraisemblance ; Méthode du maximum de vraisemblance ; Exemple de l'EMV, Maximum de vraisemblance et exhaustivité ;* Estimateurs obtenus par la méthode du MV *;* EMV et efficacité et EMV et asymptotiquement normale.

Chapitre 1

Généralité

L'estimation consiste à donner des valeurs approchées aux paramètres d'une population $(m; \sigma, etc.)$ à l'aide d'un échantillon de n observation issues de cette population.

1.1 Définition d'une statistique

Soit X une variable aléatoire dont la loi P_θ dépent du paramètre $\theta \in \Theta$

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

$$T\left(X_1, X_2, ..., X_n\right)$$

A un échantillon, on peut associer différentes statistique.

La théorie de l'estimation consiste à définir des statistiques particulières, appellées estimateurs. Une fois l'échantillon effectivement réalisé, l'estimateur prend une valeur numérique appelée **estimation de paramètre** θ . On netra $\widehat{\theta}$ l'estimateur du paramètre θ.

Exemple 1.1.1 *Soit $\overline{X}$ la statistique :*

$$\overline{X} = T\left(X_1, ..., X_n\right) = \frac{1}{n}\sum_{i=n}^{n} X_i$$

C'est-à-dire la fonction moyenne arithmétique des observation d'un échantillon. Cette statistique peut être considérée comme un estimateur, a priori raisonnable, de l'espérence mathématique.

Exemple 1.1.2 *La variace de l'échantillon aléatoire est une statistique, qui est défini comme :*

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}X_i^2 - n\overline{X}^2\right]$$

et un estimateur de la variance théorique σ^2.

Caractéristique de la statistique $\overline{X}$

- *Espérence mathématique*

L'espérence mathématique de $\overline{X}$ est égale à la moyenne m de la population, et est défini comme suit :

$$\begin{aligned}
\mathbb{E}\left(\overline{X}\right) &= \mathbb{E}\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right) \\
&= \frac{1}{n}\mathbb{E}\left(X_1 + ... + X_n\right) \\
&= \frac{1}{n}\left(\mathbb{E}\left(X_1\right) + ... + \mathbb{E}\left(X_n\right)\right) && \text{(n fois)} \\
&= \frac{1}{n}\left[m + ... + m\right] && \text{(n fois)} \\
&= m
\end{aligned}$$

- *Variance de $\overline{X}$*

La variance de $\overline{X}$ est égale à la variance σ^2 de la population divisée par la taille n de l'échantillon et est définit comme suit :

$$\begin{aligned}
Var\left(\overline{X}\right) &= Var\left[\frac{1}{n}\left(X_1+...+X_n\right)\right] \\
&= \frac{1}{n^2}Var\left[X_1+...+X_n\right] \\
&= \frac{1}{n^2}\left[Var\left(X_1\right)+...+Var\left(X_n\right)\right] \\
&= \frac{1}{n^2}\left[\sigma^2+...+\sigma^2\right] && \text{(n fois)} \\
&= \frac{\sigma^2}{n}
\end{aligned}$$

Caractéristique de la statistique de S^2

- *Espérence mathématique de*

L'espérence mathématique de S^2 est égale à la variance σ^2 de la population. En effet :

$$\begin{aligned}
\mathbb{E}\left(S^2\right) &= \mathbb{E}(\frac{1}{n-1}\left(\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right) \\
&= \frac{1}{n-1}\mathbb{E}\left[\sum_{i=1}^{n}X_i^2-n\overline{X}^2\right] \\
&= \frac{1}{n-1}\left[\sum_{i=1}^{n}\mathbb{E}\left(X_i^2\right)-n\mathbb{E}\left(\overline{X}^2\right)\right] \\
&= \frac{1}{n-1}\left[n\mathbb{E}\left(X_1^2\right)-n\mathbb{E}\left(\overline{X}^2\right)\right] \\
&= \frac{1}{n-1}\left[n\left(\sigma^2+m^2\right)-n\left(\frac{\sigma^2}{n}+m^2\right)\right] \\
&= \frac{(n-1)\,\sigma^2}{n-1} \\
&= \sigma^2
\end{aligned}$$

1.2 Fonction vraisemblance

Comme un échantillon aléatoire simple $(X_1, X_2, ..., X_n)$ est une variable aléatoire à n dimensions, on peut parler de la distribution de probabilité jointe de $(X_1, X_2, ..., X_n)$ Cette distribution est appellée **vraisemblance** de l'échantillon et elle est notée par $L(x_1, ..., x_n; \theta)$ $\forall \theta \in \Theta$.

Soit

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

s'il s'agit des variables continues où $f(x_i; \theta)$ est la densité de probabilité de X_i.

s'il s'agit des variables discrètes alors la vraisemblance devient :

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

1.3 Statistique exhaustive

Une statistique T apporte des informations du paramètre θ si sa loi de probabilité P_θ dépend de ce paramètre. Si la loi conditionnelle de l'échantillon, on suppose que la statistique $T = t$ est connue, mais elle ne dépend pas du paramètre θ, et cet échantillon ne donne plus information sur θ. Donc T apporte tout informations possible du paramètre. Cette statistique s'appelle statistique exhaustive du paramètre θ.

Définition 1.3.1 (une statistique exhaustive) *Les variables aléatoires (X_i) étant indépendantes, la densité de l'échantillon X est :*

$$L(x, \theta) \qquad \text{(Fonction vraisemblance)}$$

où $x = (x_1, ..., x_n)$ est une réalisation de l'échantillon X.

Cette densité $L(x; \theta)$ est une fonction de θ appellée vraisemblence de l'échantillon peut se mettre sou la forme :

Cas discrète

$$L(x, \theta) = g(t; \theta) . P_\theta (X = x/T = t)$$

$g(t; \theta)$ est la probabilité de la statistique T.

$P_\theta (X = x/T = t)$ *est la probabilité conditionnelle de l'échantillon sachant* $T = t$.

Cas continue

$$L(x, \theta) = g(t; \theta) . h(x; \theta/T = t)$$

$g(t; \theta)$ *est la densité de la statistique* T.

$h(x; \theta/T = t)$ *est la densité conditionnelle de l'échantillon sachant* $T = t$.

La statistique T *est une statistique exhaustive si la densité conditionnelle (resp probabilité conditionnelle) de l'échantillon sachant* T *ne dépend pas de* θ. *C'est-à-dire :*

$$L(x, \theta) = g(t; \theta) . h(x), (resp\ L(x, \theta) = g(t; \theta) . P(X = x/T = t))$$

Quand la valeur t *de la statistique est connue, l'échantillon n'apporte plus aucune information sur le paramètre* θ.

1.3.1 Propriétés d'une statistique exhaustive

1. La propriété d'exhaustivité pour une statistique est intérissante si elle ne dépend pas de la taille de l'échantillon.
2. Soit T une statistique exhaustive pour le paramètre θ et φ une fonction strictement monotone de T. Alors :
   la statistique $S = \varphi(T)$ est une statistique exhaustive pour le paramètre θ.

1.3.2 Exemple sur la statistique exhaustive

Cas discrètes

Exemple 1.3.1 *Soit* X *est une v.a suivant une loi de Poisson d'un paramètre* $\lambda \in \Theta = \mathbb{R}$ *inconnu,* $T = \sum_{i=1}^{n} X_i$ *est une statistique exaustive pour* λ.
la probabilité de X *est :*

$$P(X = k) = \frac{\lambda^k}{k!} \exp(-\lambda)$$

En effet T suit la loi de Poisson de paramètre $n\lambda$ donc la loi de la statistique T :

$$g(t;\theta)=\exp(-n\lambda)\frac{(n\lambda)^t}{t!}=\exp(-n\lambda)\frac{(n\lambda)^{\sum\limits_{i=1}^{n}x_i}}{\left(\sum\limits_{i=1}^{n}x_i\right)!}$$

et

$$\begin{aligned}L(x;\lambda)&=\prod_{i=1}^{n}P(X=x_i)=\prod_{i=1}^{n}\frac{\lambda^{x_i}}{x_i!}\exp(-\lambda)\\&=\exp(-n\lambda)\frac{\lambda^t}{\prod\limits_{i=1}^{n}x_i!}\end{aligned}$$

donc

$$h(x)=\frac{L(x;\lambda)}{g(x;\lambda)}=\frac{t!}{n^t.\prod\limits_{i=1}^{n}x_i!}$$

ne dépend pas de λ.

Exemple 1.3.2 *Soit X une variable aléatoire de la loi de Bernoulli de paramétre p. Alors $T=\sum\limits_{i=1}^{n}X_i$ est une statistique exaustive pour p*
la probabilité de X est :

$$\left\{\begin{array}{c}P(X=0)=1-p\\P(X=1)=p\end{array}\right.,p\in[0,1]$$

On a :

$$\begin{aligned}L(x,p)&=\prod_{i=1}^{n}P(X=x_i)=\prod_{i=1}^{n}p^{x_i}(1-p)^{1-x_i}\\&=p^{\sum\limits_{i=1}^{n}x_i}(1-p)^{n-\sum\limits_{i=1}^{n}x_i}\\&=p^t(1-p)^{n-t}\end{aligned}$$

et T *suit la loi Binomiale de paramètre* (n, p) *:*

$$g(x,p) = P(T = t) = C_n^t p^t (1-p)^{n-t}$$

alors :

$$P(X = x/T = t) = \begin{cases} \frac{1}{C_n^t} & si \sum_{i=1}^{n} x_i = t \\ 0 & sinon \end{cases}$$

donc $P(X = x/T = t)$ *ne dépend pas de* p.

Cas continues

Exemple 1.3.3 *Soit* X *une v.a suivant une loi uniforme sur* $[0, \theta], \forall\theta > 0$*. Elle a pour densité :*

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

La statistique :

$$T = \sup\{X_i\} \qquad i \in \{1, ..., n\}$$

est résumé exhaustif de l'échantillon $X = (X_1, ..., X_n)$ *pour le paramètre* θ*. En effet :*

Vraisemblence de l'échantillon :

$$L(x;\theta) = \begin{cases} \frac{1}{\theta^n} & si\ x \in [0;\theta] \\ 0 & sinon \end{cases}$$

Loi de statistique T *:*

Fonction de répartition :

$$G(t;\theta) = \begin{cases} 0 & si\ t < 0 \\ P(T \leq t) = \left(\frac{t}{\theta}\right)^n & si\ 0 \leq t \leq \theta \\ 1 & si\ t > \theta \end{cases}$$

la densité

$$g(t;\theta)=\begin{cases}\frac{nt^{n-1}}{\theta^{n}} & si\ 0\leq t\leq\theta\\ 0 & sinon\end{cases}$$

D'où la vraisemblance :

$$L(x;\theta)=\frac{1}{\theta^{n}}=\frac{nt^{n-1}}{\theta^{n}}\cdot\frac{1}{nt^{n-1}}$$

on trouve $h(t)=\frac{1}{nt^{n-1}}$ *ne dépond pas de* θ*. alors la factorisation set une statistique exhaustive.*

Exemple 1.3.4 *Soient* $X_1,...,X_n$ *des variables aléatoires i.i.d de loi exponentielle de paramètre* θ *. Alors* $T=\sum\limits_{i=1}^{n}X_i$ *est une statistique exhaustive.*
La densité de X_i *:*

$$f_\theta(x_i)=\frac{1}{\theta}\exp\left\{-\frac{x_i}{\theta}\right\}=\theta^{-1}\exp\left\{-\frac{x_i}{\theta}\right\}$$

Alors La fonction vraisemblance :

$$\begin{aligned}L(x;\theta)&=\prod_{i=1}^{n}f_\theta(x_i)=\prod_{i=1}^{n}\theta^{-1}\exp\left\{-\frac{x_i}{\theta}\right\}\\&=\theta^{-n}\exp\left\{-\frac{\sum\limits_{i=1}^{n}x_i}{\theta}\right\}=\theta^{-n}\exp\left\{-\frac{t}{\theta}\right\}\end{aligned}$$

et on a $L(x;\theta)=g(x;\theta).h(x;\theta/T=t)$ *où* $g(x;\theta)=\theta^{-n}\exp\left\{-\frac{t}{\theta}\right\}$ *Alors* $h(x;\theta/T=t)=h(x)=1$ *mais n'est pas toujours.*

Remarque 1.3.1 *Le principe de factorisation nous donne donc un moyen de reconnaître si une statistique est exhaustive, mais permet difficilement ou même de savoir s'il en existe une.*

Le théorème suivant répond à ces deux préoccupactions :

1.4 Loi permettant une statistique exhaustive

Théorème 1.4.1 (de Darmois) *Soit X une variable aléatoire dont le domaine de définition ne dépend pas de θ une condition nécessaire et suffisante pour que l'échantillon $(X_1, ..., X_n)$ admette une statistique exhaustive est que la forme de la densité soit*

$$f(x;\theta) = \exp\left[a(x).\alpha(\theta) + b(x) + \beta(\theta)\right] \qquad \text{(famille exponentielle)}$$

Si la densité est de cette forme et, si de plus l'application $x \longmapsto \sum_{i=1}^{n} a(x_i)$ et bijective et continûement différentiable pour le paramètre θ, alors

$$T = \sum_{i=1}^{n} a(x_i)$$

est une statistique exhaustive.

Exemple 1.4.1 *Soit $X_1, ..., X_n$ un échantillon i.i.d et on a $f(x;\theta)$ une famille exponentielle.*

$$\begin{aligned}
L(x;\theta) &= \prod_{i=1}^{n} f(x_i;\theta) \\
&= \prod_{i=1}^{n} \exp\left[a(x_i).\alpha(\theta) + b(x_i) + \beta(\theta)\right] \\
&= \exp\left[\sum_{i=1}^{n} \left(a(x_i).\alpha(\theta) + b(x_i) + \beta(\theta)\right)\right] \\
&= \exp\left[\sum_{i=1}^{n} a(x_i).\alpha(\theta) + n\beta(\theta)\right].\exp\left[\sum_{i=1}^{n} b(x_i)\right] \\
&= g(x;\theta).h(x)
\end{aligned}$$

a partir de ce résultat, on trouve que $T = \sum_{i=1}^{n} a(x_i)$ est une statistique exhaustive pour θ.

1.5 Information de Fisher

Définition 1.5.1 *On appelle quantité d'information de Fisher $I_n(\theta)$ apportée par un échantillon sur le paramètre θ la quantité suivante positive ou nulle il est de la forme :*

$$I_n(\theta) = \mathbb{E}\left[\left(\frac{\partial\left(\ln\left(L(x;\theta)\right)\right)}{\partial\theta}\right)^2\right] \tag{1.1}$$

Théorème 1.5.1 *Si le domaine de définition de X ne dépend pas de θ et si la vraisemblance est deux fois dérivable, alors :*

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

si cette quantité existe.

Démonstration

L est une densité de probabilité donc :

$$\int_{\mathbb{R}^n} L(x;\theta)\,dx = 1$$

En dérivant les deux fois par rapport à θ et en remarquant que :

$$\frac{\partial\left(L(x;\theta)\right)}{\partial\theta} = L(x;\theta)\,\frac{\partial\left(\ln\left(L(x;\theta)\right)\right)}{\partial\theta} \tag{1.2}$$

il vient :

$$\int_{\mathbb{R}^n} L(x;\theta)\,\frac{\partial\left(\ln\left(L(x;\theta)\right)\right)}{\partial\theta}dx = 0$$

Ce qui prouve que la variable aléatoire $\frac{\partial(\ln(L(x;\theta)))}{\partial\theta}$ est centrée et que :

$$I_n(\theta) = Var\left(\frac{\partial\left(L(x;\theta)\right)}{\partial\theta}\right)$$

Dérivons une deuxième fois :

$$\int\limits_{\mathbb{R}^n} L(x;\theta) \frac{\partial^2 (\ln (L(x;\theta)))}{\partial\theta^2} dx + \int\limits_{\mathbb{R}^n} \frac{\partial (\ln (L(x;\theta)))}{\partial\theta} . \frac{\partial (\ln (L(x;\theta)))}{\partial\theta} dx = 0$$

D'aprés 1.2

$$\int\limits_{\mathbb{R}^n} L(x;\theta) \frac{\partial^2 (\ln (L(x;\theta)))}{\partial\theta^2} dx + \int\limits_{\mathbb{R}^n} \left[\frac{\partial (\ln (L(x;\theta)))}{\partial\theta} \right]^2 dx = 0$$

$$\Updownarrow$$

$$\mathbb{E}\left[\left(\frac{\partial (\ln (L(x;\theta)))}{\partial\theta} \right)^2 \right] + \mathbb{E}\left[\frac{\partial^2 (\ln (L(x;\theta)))}{\partial\theta^2} \right] = 0$$

ce qui démontre le théorème.

Propriété 1.5.1 *Si le domaine de définition ne dépend pas de paramètre θ, on a les propriétés suivantes :*

- ***Additivité :****Si le domaine de définition ne dépend pas de paramètre θ, on a*

$$I_n(\theta) = nI_1(\theta)$$

En effet les operateurs espérence et dérivée seconde sont linéaires.

Ceci veut dire que chaque observation a la même importence, ce qui n'est pas le cas pour la loi uniforme sur $[0, \theta]$ où la plus grande observation est la plus intéressante.

-***Dégradation de l'information*** *: Montrons que l'information portés par une statistique est inférieure ou égale à celle apportée par l'echantillon. Soit T de densité $g(x;\theta)$ la statistique que l'on substitue à l'échantillon, on a :*

$$L(x;\theta) = g(x;\theta).h(x;\theta/T = t)$$

où $h(x;\theta/T = t)$ est la densité conditionnelle de l'échantillon. On a donc , en prenant *l'espérence des dérivées secondes :*

$$I_n(\theta) = I_T(\theta) - \mathbb{E}\left[\left(\frac{\partial^2 (\ln (h(x;\theta)))}{\partial\theta^2}\right)\right]$$

le dernier terme est la quantité d'information conditionnelle $I_{n/T}(\theta)$ (ou information supplémentaire) ; elle est positive ou nulle, donc :

$$I_T(\theta) \leq I_n(\theta)$$

on voit donc que si T est exhaustive $I_T(\theta) = I_n(\theta)$ et que la réciproque est vraie si le domaine de X est indépendante de θ.

Exemple 1.5.1 *Soit X une v.a de loi exponentielle de paramètre α, et dispose d'un échantillon indépendant $(x_1, ..., x_n)$*
Calculer $I_n(\alpha)$.

$$f(x) = \alpha \exp(-\alpha x)$$

$$\begin{aligned} L(x;\alpha) &= \prod_{i=1}^{n} \alpha \exp(-\alpha x_i) \\ &= \alpha^n \exp\left(-\alpha \sum_{i=1}^{n} x_i\right) \end{aligned}$$

$$\ln(L(x;\alpha)) = n\ln(\alpha) - \alpha\sum_{i=1}^{n} x_i$$

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

$$\begin{aligned} \frac{\partial^2(\ln(L(x;\alpha)))}{\partial\alpha^2} &= \frac{\partial}{\partial\alpha}\left(\frac{\partial(\ln(L(x;\alpha)))}{\partial\alpha}\right) \\ &= n\left(-\frac{1}{\alpha^2}\right) \end{aligned}$$

Calculons $I_n(\alpha)$

$$\begin{aligned} I_n(\alpha) &= \mathbb{E}\left(-\frac{\partial^2 (\ln (L(x;\alpha)))}{\partial \alpha^2}\right) \\ &= \frac{n}{\alpha^2} \end{aligned}$$

d'où

$$I_n(\alpha) = \frac{n}{\alpha^2} \ \textit{et}\ I_1(\alpha) = \frac{1}{\alpha^2}$$

Chapitre 2

Qualité d'un estimateur

Dans ce chapitre, on a étudié quelque estimateurs qui fournis la puisse des bonnes estimateurs parmis toutes les statistiques possibles, il doit posséder le meilleur estimateur.

2.1 Définition d'un estimateur

Soit X une variable aléatoire dont la loi de probabilité P_θ dépond d'un seul paramètre θ. Soit $(X_1, ..., X_n)$ un échantillon aléatoire.

Définition 2.1.1 *Un estimateur est une statistique qui a des propriétés bien définies.* $T_n = T(X_1, ..., X_n)$ *est une suite de statistique définie :*

$$T_n = \varphi(X_1, ..., X_n)$$

est appelée estimateur du paramètre θ*. Si* T_n *tend vers* θ *quand* n *tend vers l'infini, la convergence étant une convergence en probabilité, presque sûre ou moyenne quadratique. En général, on note* $\hat{\theta}(X_1, ..., X_n)$ *ou* $\hat{\theta}_n$*, ou plus simplement,* $\hat{\theta}$ *l'estimateur du paramètre* θ.

2.2 Estimateur sans biais

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

$$T_n - \theta = T_n - \mathbb{E}(T_n) + \mathbb{E}(T_n) - \theta$$

où $\mathbb{E}(T)$ est l'espérence de T.

Le premier terme $T_n - \mathbb{E}(T_n)$ représente les fluctuations aléatoires de T_n par rapport à sa valeur moyenne, tendis que deuxième terme $\mathbb{E}(T_n) - \theta$ représente un erreur systématique due au fait que T_n varie autour de sa valeur $\mathbb{E}(T_n)$ et non autour de θ, sauf si $\mathbb{E}(T_n) = \theta$.

Définition 2.2.1 *Nous appellons **biais** d'un estimateur T_n au point θ la fonction*

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

On dit que l'estimateur est :

*- Un estimateur **sans biais** si*

$$\forall \theta \in \Theta \qquad b_{T_n}(\theta) = 0 \Rightarrow \mathbb{E}(T_n) = \theta$$

*- Un estimateur **asymptotiquement sans biais (a.s.b)** si*

$$\lim_{n \to \infty} b_{T_n}(\theta) = 0 \qquad \forall \theta \in \Theta$$

Exemple 2.2.1 *1. Soit $(X_1, ..., X_n)$ des variables aléatoires i.i.d ie*

$$\forall i = 1, ..., n \ \mathbb{E}(X_i) = m \ et \ Var(X_i) = \sigma^2$$

On a l'estimateur $\overline{X}$ *est un estimateur sans biais de* m*.On obtient l'espérence et la variance de* $\overline{X}$

$$\mathbb{E}\left(\overline{X}\right) = m$$

$$Var\left(\overline{X}\right) = \frac{\sigma^2}{n}$$

(voir l'exemple 1.1.2). On calcule le biais

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

qui est vérifié.

2. *L'estimateur de* S_n^2

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

on va calculer l'espérence de S_n^2 *:*

$$\begin{aligned} \mathbb{E}\left(S_n^2\right) &= \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}\left(X_i^2 - \overline{X}^2\right)\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[\mathbb{E}\left(X_i^2\right) - \mathbb{E}\left(\overline{X}^2\right)\right] \\ &= \frac{1}{n}\sum_{i=1}^{n}\left[Var\left(X_i\right) + \mathbb{E}^2\left(X\right) - \left(Var\left(\overline{X}\right) + \mathbb{E}^2\left(\overline{X}\right)\right)\right] \\ &= \sigma^2 + m^2 - \left(\frac{\sigma^2}{n} + m\right) = \sigma^2 - \frac{\sigma^2}{n} \\ &= \frac{n-1}{n}\sigma^2 \end{aligned}$$

Calculons le biais de S_n^2

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

alors S_n^2 est un estimateur biaisé mais quand n tend vers l'infini dans ce cas S_n^2 est dit un estimateur ***a.s.b.***

2.3 Estimateur convergent

Dans cette paragraphe, nous considérons un estimateur T_n dont la loi dépond de paramètre θ qu'on s'intéresse un bon estimateur T_n. S'il est proche de la vraie valeur de θ. Pour savoir la dernière, on peut l'interpréter, par exemple : la convergence en probabilité de T_n vers la vraie valeur du paramètre θ.

Définition 2.3.1 *L'estimateur est dit* ***convergent (ou consistant)*** *si la suite T_n converge en probabilité vers θ $\left(T_n \xrightarrow{P} \theta\right)$*

$$\forall \varepsilon > 0,\ \forall \theta \in \Theta,\ \lim_{n\to\infty} P\left(|T_n - \theta| > \varepsilon\right) = 0$$

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

Propriété 2.3.1 *Pour valider la convergence d'un estimateur, il suffit de montrer que :*

$$\forall \theta \in \Theta,\ \lim_{n\to\infty} \mathbb{E}\left(T_n\right) = \theta \quad et \quad \lim_{n\to\infty} Var\left(T_n\right) = 0$$

Pour vérifier ce critère, il faut connaître la loi de X et celle de T_n pour calculer l'espérence et la variance.

Exemple 2.3.1 *Nous avons montrer que $\overline{X}$ est un estimateur sans biais (voir l'exemple 1 partie 1). On a :*

$$\mathbb{E}\left(\overline{X}\right) = m \quad et \quad Var\left(\overline{X}\right) = \frac{\sigma^2}{n} \underset{n\to\infty}{\longrightarrow} 0$$

Alors $\overline{X}$ est un estimateur convergent.

Exemple 2.3.2 *Lorsque m est connu, l'expression de S_n^2 est :*

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2$$

Son espérence est

$$\mathbb{E}\left(S_n^2\right) = \mathbb{E}\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - m)^2\right] = \frac{1}{n}\sum_{i=1}^{n}Var\left(X_i\right) = \sigma^2$$

Par ailleurs, les variables $(X_i - m)$ sont indépendantes :

$$\begin{aligned} Var\left(S_n^2\right) &= \frac{1}{n^2}\sum_{i=1}^{n}Var\left((X_i - m)^2\right) \\ &= \frac{1}{n}\left(\mathbb{E}\left[(X - m)^4\right] - \left[\mathbb{E}\left[(X - m)^2\right]\right]^2\right) \\ &= \frac{1}{n}\left(\mu_4 - \sigma^4\right) \end{aligned}$$

avec $\mu_k = \mathbb{E}\left[(X - m)^k\right]$. Alors S_n^2 est un estimateur sans biais et convergent.

2.4 Estimateur asymptotiquement normale

Définition 2.4.1 *Soit T_n un estimateur du paramètre de la loi P_θ d'une variable aléatoire observée X. Nous supposons qu'il existe deux fonction $a = a(\theta, n)$ et $b = b(\theta, n)$ telle que :*

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

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

Remarque 2.4.1 *La loi asymptotique sera d'autant plus utile que les fonctions a et b ont une expression simple.*

Exemple 2.4.1 *Soit une v.a centrée et possèdant un moment d'ordre 4 :*

$$\mathbb{E}\left(X^4\right) = \mu_4 < \infty$$

Dans ces conditions, nous avons

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

Soit n-échantillions $(X_1, ..., X_n)$ *pour estimer* σ^2*, nous considérons l'estimateur suivant :*

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

Il est facile de montrer que :

$$\mathbb{E}\left(T_n\right) = \sigma^2$$

On calcule la variance

$$\begin{aligned}
Var\left(T_n\right) &= \mathbb{E}\left(T_n^2\right) - \left(\mathbb{E}\left(T_n\right)\right)^2 \\
&= \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i^2\right)^2\right] - \sigma^4 \\
&= \frac{1}{n^2}\mathbb{E}\left[\sum_{i=1}^{n} X_i^4 + \sum_i\sum_j X_i^2 X_j^2\right] - \sigma^4, \qquad i \neq j \\
&= \frac{1}{n^2}\left[\mathbb{E}\left[\sum_{i=1}^{n} X_i^4\right] + \mathbb{E}\left[\sum_i\sum_j X_i^2 X_j^2\right]\right] - \sigma^4, \qquad i \neq j \\
&= \frac{1}{n^2}\left[n\mu_4 + n\left(n-1\right)\sigma^4\right] - \sigma^4 \\
&= \frac{\mu_4 - \sigma^4}{n}
\end{aligned}$$

Ainsi T_n *est un estimateur convergent (d'aprés propriété 2.3.1) et sans biais de* σ^2*. De* *plus, il satisfait au condition du théorème centrale limite (T.C.L). Nous pouvons écrire :*

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

et on a : $a=\sigma^2$ *et* $b=\sqrt{\frac{\mu_4-\sigma^4}{n}}$. *En déduirant que* T_n *est un estimateur* ***a.N.***

Exemple 2.4.2 *Soit* X *une v.a de loi uniforme* $\mathcal{U}([0,\theta])$ *avec* $\theta>0$*.Nous savons alors que :*

$$f_X(t)=\begin{cases} 0 & si\ t\notin[0,\theta] \\ \frac{1}{\theta} & si\ t\in[0,\theta] \end{cases} \quad et\ F_X(t)=\begin{cases} 0 & si\ t<0 \\ \frac{t}{\theta} & si\ 0\le t\le\theta \\ 1 & si\ t>\theta \end{cases}$$

Soit $(X_1,...,X_n)$ *n-échantillons, et on a l'estimateur*

$$T_n=X_{n,n}=\max_{1\le i\le n} X_i$$

des propriétés de la statistique d'ordre :

$$F_{X_{n,n}}(t)=\begin{cases} 0 & si\ t<0 \\ \left(\frac{t}{\theta}\right)^n & si\ 0\le t\le\theta \\ 1 & si\ t>\theta \end{cases} \quad et\ f_{X_{n,n}}(t)=\begin{cases} 0 & si\ t\notin[0,\theta] \\ \frac{nt^{n-1}}{\theta^n} & si\ t\in[0,\theta] \end{cases}$$

- Calculons l'espérence

$$\begin{aligned}\mathbb{E}(T_n)&=\int_{\mathbb{R}} t f_{X_{n,n}}(t)\,dt=\frac{n}{\theta^n}\int_0^\theta t^n dt\\ &=\frac{n}{\theta^n}\left.\frac{t^{n+1}}{n+1}\right|_0^\theta\\ &=\frac{n\theta}{n+1}\underset{n\to\infty}{\longrightarrow}\theta\end{aligned}$$

et la variance

$$Var(T_n)=\mathbb{E}\left(T_n^2\right)-\left(\mathbb{E}(T_n)\right)^2$$

On a : $\mathbb{E}\left(T_n^2\right)=\int\limits_{\mathbb{R}} t^2 f_{X_{n,n}}(t)\,dt=\frac{n}{\theta^n}\int\limits_0^\theta t^{n+1}dt=\frac{n\theta^2}{n+2}$ *alors*

$$\begin{aligned} Var\left(T_n\right) &= \frac{n\theta^2}{n+2}-\left(\frac{n\theta}{n+1}\right)^2 \\ &= \left(\frac{n\left(n+1\right)^2-n\left(n+2\right)}{\left(n+1\right)^2\left(n+2\right)}\right)\theta^2 \\ &= \frac{n\theta^2}{\left(n+1\right)^2\left(n+2\right)} \underset{n\to\infty}{\longrightarrow} 0 \end{aligned}$$

A partir des fonctions de répartition, avec un calcul sur les probabilités, nous pouvons montrer que :

$$\lim_{n\to\infty}\mathcal{L}\left(n\left(T_n-\theta\right)\right)=\mathcal{GA}\left(1;\frac{1}{\theta}\right)$$

En déduirant que T_n *est un estimateur convergent et a.s.b, mais n'est pas asymptotiquement Normale. On sait que les paramètres de loi gamma dans le cas général sont* (α,λ) *et son espérence est :* $\mathbb{E}\left(X\right)=\frac{\alpha}{\lambda}$*. Dans cette exemple, on obtient* $\alpha=\frac{n}{n+1}\underset{n\to\infty}{\longrightarrow}1$ *et* $\lambda=\frac{1}{\theta}$

2.5 Estimateur efficace

2.5.1 Erreur quadratique moyenne

On mesure la précision d'un estimateur T du paramètre θ par l'erreur quadratique moyenne définie par :

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

Théorème 2.5.1 *Si* T *est un estimateur du paramètre* θ *alors :*

$$EQM\left(T\right)=Var\left(T\right)+\left(b_T\left(\theta\right)\right)^2$$

Proof. On a :

$$
\begin{aligned}
EQM(T) &= \mathbb{E}\left[(T-\theta)^2\right] \\
&= \mathbb{E}\left[(T-\mathbb{E}(T)+\mathbb{E}(T)-\theta)^2\right] \\
&= \mathbb{E}\left[(T-\mathbb{E}(T))^2 + 2\left[(T-\mathbb{E}(T))(\mathbb{E}(T)-\theta)\right] + (\mathbb{E}(T)-\theta)^2\right]
\end{aligned}
$$

Comme $\mathbb{E}(T)-\theta$ est une constante et que $\mathbb{E}[T-\mathbb{E}(T)]=0$, il vient :

$$EQM(T) = \mathbb{E}\left[(T-\mathbb{E}(T))^2\right] + [\mathbb{E}(T)-\theta]^2$$

Donc

$$EQM(T) = Var(T) + (b_T(\theta))^2$$

■

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

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

- $Var(T_n)$ soit petite.

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

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

En statistique, **l'inégalité de Fréchet-Darmois-Cramer-Rao (ou borne FDCR)** exprime une borne inférieure sur la variance d'un estimateur sans biais, basée sur l'information de Fisher. Elle est aussi appelée la **borne de Cramér-Rao**.

Si X prend ses valeurs dans un ensemble ne dépend pas de θ, si la densité $f(x;\theta)$ est eux fois continument dérivable par rapport à θ, et sous certaines condition de réguliarité tout estimateur T_n sans biais de θ dont la variance existe vérifier la **borne Cramer-Rao**

$$Var\left(T_n\right) \geq \frac{1}{I_n\left(\theta\right)}\left[\frac{\partial \mathbb{E}\left(T_n\right)}{\partial \theta}\right]^2 \qquad \forall \theta \in \Theta$$

Où $I_n\left(\theta\right)$ est la quantité d'information de Fisher apportée par l'échantillon sur θ.

Cas particulier

-Si T_n est un estimateur sans biais d'une fonction $h\left(\theta\right)$, c-à-d $\mathbb{E}\left(T_n\right) = h\left(\theta\right)$. L'inégalité de FDCR s'écrit :

$$Var\left(T_n\right) \geq \frac{\left[h'\left(\theta\right)\right]^2}{I_n\left(\theta\right)} \qquad \forall \theta \in \Theta$$

-Si T_n est un estimateur sans biais de θ, c-à-d $\mathbb{E}\left(T\right) = \theta$. La borne de FDCR s'écrit :

$$Var\left(T_n\right) \geq \frac{1}{I_n\left(\theta\right)} \qquad \forall \theta \in \Theta$$

Définition 2.5.1 *Un estimateur T_n sans biais de θ est dit* ***efficace*** *si sa variance est égale à la borne de FDCR* **:**

$$Var\left(T_n\right) = \frac{1}{I_n\left(\theta\right)} \qquad \forall \theta \in \Theta$$

Exemple 2.5.1 *Soit $X = \left(X_1, ..., X_n\right)$ des v.a iid suit la loi normale (ie $X_i \backsim \mathcal{N}\left(m, \sigma^2\right)$) et on a la variance $Var\left(\overline{X}\right) = \frac{\sigma^2}{n}$*

la fonction de vraisemblance est :

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

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

On va calculer la première dérivée de la fonction $\ln\left(L\left(x;m\right)\right)$ *par rapport à* m

$$\frac{\partial \ln\left(L\left(x;m\right)\right)}{\partial m}=\frac{1}{\sigma^2}\sum_{i=1}^{n}\left(x_i-m\right)$$

La deuxième dérivée est :

$$\frac{\partial^2 \ln\left(L\left(x;m\right)\right)}{\partial m^2}=-\frac{n}{\sigma^2}$$

On a la quantité d'information est :

$$I_n\left(m\right)=-\mathbb{E}\left(\frac{\partial^2 \ln\left(L\left(x;m\right)\right)}{\partial m^2}\right)=-\left(-\frac{n}{\sigma^2}\right)=\frac{n}{\sigma^2}$$

d'où

$$Var\left(\overline{X}\right)=\frac{1}{I_n\left(m\right)}$$

c'est à dire que $\overline{X}$ *est un estimateur efficace.*

Chapitre 3

Estimation par la méthode du maximum de vraisemblance

Dans ce chapitre, on a étudier l'estimation par la méthode du maximum de vraisemblance de paramètre θ qui consiste à associer à un échantillon et en donnant quelque propriété.

3.1 Définition du maximum de vraisemblance

Définition 3.1.1 *Soit $L(x;\theta)$ la fonction de vraisemblance au point $\theta \in \Theta$, Θ quelconque. On appelle **estimateur du maximum de vraisemblance (emv)** pour la statistique :*

$$\hat{\theta}^{MV} : \mathcal{L} \longrightarrow \Theta$$
$$(x_1, \cdots, x_n) \overset{\hat{\theta}^{MV}}{\longrightarrow} \hat{\theta}^{MV}(x_1, \cdots, x_n)$$

Telle que :

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

3.2 Méthode du maximum de vraisemblance

Le principe de l'estimation par maximum de vraisemblance est de dire que plus la probabilité d'avoir obtenu les observations est forte, plus le modèle est proche de réalité.
Ainsi, on retient le modèle pour le quel la vraisemblance de notre échantillon est la plus élevée :

$$\hat{\theta}^{MV} = \arg\max_{\theta} L\left(x_1, \cdots, x_n; \theta\right) \tag{3.1}$$

En pratique, On va savoir un problème de résoudre directement en raison de la présence du produit mais il suffit de prendre le logarithme de la vraisemblance :

$$\hat{\theta}^{MV} = \arg\max_{\theta} l\left(x_1, \cdots, x_n; \theta\right)$$

où $l\left(x_1, \cdots, x_n; \theta\right) = \log\left(L\left(x_1, \cdots, x_n; \theta\right)\right)$
Pour trouver le maximum, on résoud l'équation du premier ordre :

$$\left.\frac{\partial l\left(x_1, \cdots, x_n; \theta\right)}{\partial \theta}\right|_{\theta = \hat{\theta}^{MV}} = 0$$

La théorie nous dit que la solution de cette équation nous donne toujours le maximum.On obtient $\hat{\theta}^{MV}$ sous la forme $\hat{\theta}^{MV} = g\left(x_1, \cdots, x_n\right)$.

Propriété 3.2.1 (Invariance fonctionnelle) *Si $\hat{\theta}^{MV}$ est l'estimateur de θ par la méthode du maximum de vraisemblance, $f\left(\hat{\theta}^{MV}\right)$ est l'estimateur de $f\left(\theta\right)$ par la méthode du maximum de vraisemblance.*

3.3 Exemple de l'EMV

Exemple 3.3.1 (variable aléatoire de loi Bernoulli) *On suppose que $X \backsim \mathcal{B}(p)$ (p est la vrai paramètre inconnu). La probabilité de X est :*

$$P(X = x) = p^x (1-p)^{1-x} \quad où\ x, p \in [0, 1]$$

On tire un échantillon $X_1, \cdots, X_n$ iid de taille n.

La vraisemblance de cette échantillon est donnée par :

$$L(x_1, \cdots, x_n; p) = \prod_{i=1}^{n} P(X_i = x_i; p)$$

soit

$$\begin{aligned} L(x_1, \cdots, x_n; p) &= \prod_{i=1}^{n} p^{x_i} (1-p)^{1-x_i} \\ &= p^{\sum\limits_{i=1}^{n} x_i} (1-p)^{1-\sum\limits_{i=1}^{n} x_i} \\ &= p^s (1-p)^{1-s} \end{aligned}$$

L'EMV $\hat{p}$ maximise $L(x_1, \cdots, x_n; p)$ par rapport à p.

Une condition nécessiare pour obtenir le maximum de $L(x_1, \cdots, x_n; p)$ est que la dérivée par rapport à p est nulle.

puisque

$$\frac{\partial L(x_1, \cdots, x_n; p)}{\partial p} = sp^{s-1} (1-p)^{n-s} - p^s (n-s) (1-p)^{n-s-1}$$

Cette condition implique :

$$sp^{s-1} (1-p)^{n-s} - p^s (n-s) (1-p)^{n-s-1} = 0$$

En dérivant par $p^{s-1}(1-p)^{n-s-1}$, *on obtient*

$$s(1-\hat{p}) - \hat{p}(n-s) = 0$$

et alors

$$s - \hat{p}n = 0 \Longrightarrow \hat{p} = \frac{s}{n}$$

ou encore

$$\sum_{i=1}^{n} x_i - \hat{p}n = 0 \Longrightarrow \hat{p} = \frac{\sum_{i=1}^{n} x_i}{n} = \bar{X}$$

On peut facilement vérifier (dérivée seconde de $L(x_1, \cdots, x_n; p)$ *par rapport à* p *est négative) donc* $p = \bar{X}$ *est le maximum unique de* $L(x_1, \cdots, x_n; p)$.

Exemple 3.3.2 (Variable aléatoire de loi normale) *On suppose que* $X \backsim \mathcal{N}(m, \sigma^2)$ *et sa densité est :*

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

Où la vraie moyenne m *et la vraie variance* σ^2 *sont des inconnues, On tire un échantillon* $X_1, \cdots, X_n$ *iid de taille* n. *Les valeurs observées sont* $x_1, \cdots, x_n$.

On essaye de trouver les estimateurs du maximum de vraisemblance de la moyenne $\hat{m}^{MV}$ *et de la variance* $\hat{\sigma}^{2MV}$.

- Considérons d'abord l'estimation de m *:*

Puisque X *est une v.a continue,la fonction de vraisemblance doit être définit à partir de la densité.*

Soit

$$\begin{aligned} L(x_1, \cdots, x_n; m) &= f(x_1, m) \cdots f(x_n, m) \\ &= \prod_{i=1}^{n} f(x_i, m) \end{aligned}$$

La vraisemblance qu'une population de la moyenne m *et de la variance* σ^2 *généré l'échan-*

tillon observé s'avère être :

$$\begin{aligned} L\left(x_1, \cdots, x_n; m\right) &= \prod_{i=1}^{n}\left[\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{1}{2}\left(\frac{x_i-m}{\sigma}\right)^2\right)\right] \\ &= \frac{1}{\left(\sqrt{2\pi\sigma^2}\right)^n}\exp\left(-\frac{1}{2}\sum_{i=1}^{n}\left(\frac{x_i-m}{\sigma}\right)^2\right) \end{aligned}$$

Puis, on a le logarithme de la vraisemblance :

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

On a l'équation de vraisemblance :

$$\frac{\partial \ln L\left(X; m\right)}{\partial m} = \sum_{i=1}^{n}\frac{x_i-m}{\sigma^2}$$

Ainsi, $\hat{m}^{MV}$ *est déterminé par l'équation suivante :*

$$\begin{aligned} \sum_{i=1}^{n}\frac{x_i-m}{\sigma^2} = 0 &\Longrightarrow \sum_{i=1}^{n}\left(x_i-m\right) = 0 \\ &\Longrightarrow \hat{m}^{MV} = \sum_{i=1}^{n}\frac{x_i}{n} = \bar{X} \end{aligned}$$

- *Considérons maintenant l'EMV de* σ^2 :

Soit toujours

$$\begin{aligned} \ln L\left(x_1, \cdots, x_n; \sigma^2\right) &= -n\ln\left(\sigma\sqrt{2\pi}\right) - \frac{1}{2}\sum_{i=1}^{n}\frac{\left(x_i-m\right)^2}{\sigma^2} \\ &= -\frac{n}{2}\ln\left(\sigma^2\right) - \ln\left(\sqrt{2\pi}\right) - \frac{1}{2}\sum_{i=1}^{n}\frac{\left(x_i-m\right)^2}{\sigma^2} \end{aligned}$$

L'équation de la vraisemblance est :

$$\frac{\partial \ln L\left(X, m, \sigma^2\right)}{\partial \sigma^2} = \frac{-n}{2\sigma^2} + \sum_{i=1}^{n}\frac{\left(x_i-m\right)^2}{2\sigma^4}$$

L'EMV $\hat{\sigma}^2_{MV}$ est déterminé, ainsi, par l'équation suivante :

$$\frac{-n}{2\sigma^2} + \sum_{i=1}^{n} \frac{(x_i - m)^2}{2\sigma^4} = 0 \Longrightarrow \frac{-n\sigma^2 + \sum\limits_{i=1}^{n} (x_i - m)^2}{2\sigma^4} = 0$$
$$\Longrightarrow -n\sigma^2 + \sum_{i=1}^{n} (x_i - m)^2 = 0$$

Alors

$$\hat{\sigma}^{2^{MV}} = \frac{1}{n} \sum_{i=1}^{n} (x_i - m)^2$$

Donc on remplace par sa valeur devient

$$\hat{\sigma}^{2^{MV}} = \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \bar{X}\right)^2 = S^2$$

3.4 Maximum de vraisemblance et exhaustivité

Soit T une statistique exhaustive :

$$L(x_1, \cdots, x_n; \theta) = g(t, \theta)\, h(x_1, \cdots, x_n)$$

Toute fonction de $(x_1, \cdots, x_n)$ solution de l'équation de la vraisemblance :

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

est solution de l'équation :

$$\frac{\partial \ln g(t, \theta)}{\partial \theta} = 0$$

Pour réaliser un maximum de $L(x_1, \cdots, x_n; \theta)$, il suffit de réaliser un maximum de $g(t, \theta)$. Donc, toute estimation de θ par la méthode du maximum de vraisemble est une fonction de T mais n'est pas nécessairement une statistique exhaustive.

Distribution	Parmètre à estimer	Estimateur
Loi uniforme sur $[0,\theta]$	θ	$\hat{\theta}=\sup(x_i)$
Loi de binomiale	p	$\hat{p}=\frac{k}{n}$
k nombre de succés en épreuves		
Loi de poisson	λ	$\hat{\lambda}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{X}$
Loi normale	m	$\hat{m}=\frac{1}{n}\sum_{i=1}^{n}x_i=\bar{X}$
	σ^2	$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-\bar{X}\right)^2$

Tab. 3.1 – Estimateurs obtenus par la méthode du MV

3.5 Estimateurs obtenus par la méthode du MV

Dans cette paragraphe, on a obtenu quelque loi avec leurs estimateurs :3.1

3.6 EMV et efficacité

Théorème 3.6.1 (Sur l'efficacité) *S'il existe un réel $a_n(\theta)$ dépendant de θ tels que l'emv $\hat{\theta}^{MV}$ de θ vérifie la factorisation*

$$\frac{\partial \ln L\left(x_1,\cdots,x_n;\theta\right)}{\partial\theta}=a_n(\theta)\left(\hat{\theta}^{MV}-\theta\right)$$

alors :

- $\hat{\theta}^{MV}$ est un estimateur sans biais : $E(\hat{\theta}^{MV})=\theta$.
- On a :

$$Var\left(\hat{\theta}^{MV}\right)=\frac{1}{a_n(\theta)}$$

- On a $I_n(\theta)$ une quantité d'information :

$$I_n(\theta)=\frac{1}{Var\left(\hat{\theta}^{MV}\right)}$$

En particulier, le premier et le troisième points entraînent que $\hat{\theta}^{MV}$ est un estimateur efficace de θ.

Corollaire 3.6.1 *S'il existe un estimateur efficace de θ, alors c'est l'emv $\hat{\theta}^{MV}$ de θ (en revanche, un emv n'est pas nécessairement efficace).*

3.7 EMV et asymptotiquement normale

Proposition 3.7.1 *On possède un échantillon indépendant $(X_1, \cdots, X_n)$ et on note $\hat{\theta}^{MV}$ l'emv de θ.*
l'emv $\hat{\theta}_n^{MV}$ de θ est asymptotiquement normale ; la suite de variable :

$$\sqrt{I_n(\theta)}\left(\hat{\theta}_n^{MV} - \theta\right) \longrightarrow \mathcal{N}(0,1)$$

converge en loi normale centré réduite quand n tend vers l'infini.

Conclusion

Dans ce travail, nous avons étudié la méthode du maximum de vraisemblance. Elle est l'une des méthodes la plus utilisée en statistique. L'estimateur obtenue par la méthode du maximum de vraisemblance possède des propriétés intéressantes, notamment : la convergence, l'efficacité, la normalité asymptotique.

Bibliographie

[1] Besma Belhadj-Kaabi, Inférence statistique, Tunisien, 2001.

[2] Christophe Chesneau, Sur l'estimateur du maximum de vraisemblance, Disponible dans : (https ://cel.archives-ouvertes.fr/cel-01430435/document)

[3] François Cottet-Emard, Probabilités et tests d'hypothèse, Université de Paris, France, 2014.

[4] Gilbert Saporta, Probabilités Analyse des données et Statistique, Université de Paris, France, mars 2006.

[5] Gannaz, Introduction à la statistique. Disponible dans : (http ://math.univ-lyon1.fr/~gannaz/Cours/cours_stat.pdf).

[6] Michel Lejeune, Statistique la théorie et ses applications, Université de Grenoble 2, France, avril 2004.

[7] Philipe Tassi, Méthode statistique, Université de Paris, France, 1989.

[8] Renée Veysseyre, Aide mémoird Statistique et Probabilité pour l'ingénieur, Université de Paris, France, 2001.

[9] Thérèse Phan, Probabilités et statistiques (Cours 1ère année). 2-Ecole Centrale Paris, 2005

[10] William Hines, Probabilités et statistiques pour l'ingénieur 2èmes édition, 2005.

[11] Yadolah Dodge auth, premier pas en statistique, Université de Neuchâtel, Septempre 1999.

Annexe B : Abréviations et Notations

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

Θ	Ensemble des paramètres
$\mathbb{E}[T]$	Espérance mathématique ou moyenne du v.a. T.
$Var\,[T]$	Variance du v.a.
$\exp$	Exponentiel.
P_θ	Loi de paramètre θ
$F\,(x;\theta)$	Fonction de répartition.
$f\,(x;\theta)$	Densité de probabilité de X_i.
$i.i.d$	Indépendantes et identiquement distribuées.
$v.a$	Variable aléatoire.
$L\,(x;\theta)$	Fonction de vraisemblance.
I_n	Information de Fisher.
$\bar{X}$	La moyenne empirique.
S^2	La variance empirique
EQM	Erreur quadratique moyenne.
$a.s.b$	Asymptotiquement sans biais
$a.n$	Asymptotiquement normale

$FDCR$	Fréchet-Darmois-Cramer-Rao
$\mathcal{GA}(\alpha, \lambda)$	loi Gamma de paramètre (α, λ)
$\mathcal{N}(0,1)$	loi normale centré réduite
$\mathcal{L}$	loi
$\mathcal{N}(m, \sigma^2)$	loi normale
$\hat{\theta} = \hat{\theta}_n = \hat{\theta}(x_1, \cdots, x_n)$	estimateur
$\mathcal{B}(p)$	loi Bernoulli de paramètre p.
ln	Le logarithme néparien
EMV	Estimateur du maximum de vraisemblance
$\hat{\theta}^{MV}$	Estimateur du maximum de vraisemblance
$p.s$	presque sûre