

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ KASDI MERBAH, OUARGLA

Faculté des Mathématiques et des Sciences de la Matière

DÉPARTEMENT DE MATHÉMATIQUES



Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Probabilités et Statistique**

Par :

Belabbas Rachida

Titre :

Estimation par la méthode du maximum de vraisemblance

Membres du Comité d'Examen :

Dr.	Ben brahim Radhia	UKM, OUARGLA	Président
Dr.	Mansoul Brahim	UKM, OUARGLA	Encadreur
Dr.	Saouli Mostapha abdelouahab	UKM, OUARGLA	Examineur

Juin 2021

DÉDICACE

Je dédie ce modeste travail : à ma chère mère, la bougie de ma vie.

à mon chère père, qui souffret pour mon éducation.

Allah vous protège.

à mes sœurs et frères, mon bras droit dans ma vie :

Miloud, Faiza, Aicha, Khadra, Doua, Tidjani, Nawal, Mohammed.

à mes fidèles amis : Imane et oumessaad.

à ma fiancée : Abdallatif. B.

à tous mes famille belabbas et tous mes amis.

REMERCIEMENTS

*Je tiens premièrement à me prosterner, remerciant «Allah» le tout
puissant de m'avoir donné la force et la volonté
pour terminer ce travail.*

Nous voudrions remercier le encadré Brahim Mansoul pour les orientations scientifiques
précises; nos profonds remerciements pour les membres de jury nous qui ont accepté
d'évaluer ce travail.

Merci également a tous les enseignants qui m'ont aidé pendant mon cursus, sans oublier
leurs conseils précieux.

Je remercie aussi toute personne de près ou de loin a contribué à la finalisation de ce travail

Table des matières

Dédicace	i
Remerciements	ii
Table des matières	iii
Introduction	1
1 Rappel sur les notions probabilités et statistiques	3
1.1 Tribu	3
1.1.1 Mesurabilité	4
1.2 Variable aléatoire	4
1.3 Loi d'une variable aléatoire	5
1.4 Lois de probabilité	5
1.4.1 Lois discrètes	5
1.4.2 Lois continues	8
1.5 Population et Caractère statistique	11
1.6 Échantillon	11
1.6.1 Échantillonnage aléatoire sans remise	12
1.6.2 Échantillonnage aléatoire avec remise	12
1.6.3 La distribution d'échantillonnage	12

1.6.4	Statistique	12
1.6.5	Moments empiriques	13
2	Estimation statistique paramétrique	16
2.1	Cadre de l'estimation paramétrique	16
2.2	Estimateur	16
2.3	Qualité d'un estimateur	17
2.3.1	Le biais d'un estimateur	17
2.3.2	Variance et erreur on moyenne quadratique d'un estimateur	18
2.3.3	Estimateur convergent	19
2.4	Exhaustivité	20
2.5	Information de Fisher	21
2.6	Inégalité de Fréchet-Darmonis-Cramer-Rao(FDCR)	22
3	Méthod du maximum de vraisemblance	24
3.1	fonction $L(x, \theta)$ de vraisemblance	24
3.2	Méthode du maximum de vraisemblance	25
3.3	Propriétés asymptotiques	26
3.4	Exemple de l'EMV	26
3.4.1	Loi discrète	26
3.4.2	Lois continues	27
3.5	Remarques	30
3.6	Application en R	32
3.6.1	Loi discrète	32
3.6.2	Loi continue	33
	Conclusion	35

Bibliographie	36
Annexe A : Logiciel R	38
Annexe B : Abréviations et Notations	40

Introduction

L'estimation est l'un des activités principales de statistique, et c'est la problématique inverse de l'échantillonnage, il s'agit d'estimer certaines caractéristiques statistiques de la loi (moyenne, variance, fonction de répartition, etc...) à travers d'une série d'observations. 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. La théorie de l'estimation étudie les méthodes générales d'estimations.

On a plusieurs méthodes pour estimer un paramètre comme la méthode des moindres carrés, et la méthode des moments, et la méthode du maximum de vraisemblance cette méthode a été développée par le statisticien Aylmer Ronald Fisher en 1922. Il utilisa la loi Binomiale pour illustrer son critère.

La méthode du maximum de vraisemblance a trouvé un ensemble des estimateurs des paramètres appelé la vraisemblance est d'avoir maximiser l'échantillon que nous avons utilisé, cet estimateur a été beaucoup étudié ainsi que ses propriétés asymptotiques (convergence, efficacité, normalité asymptotique,..) pour plusieurs modèle statistique.

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 .

Dans ce mémoire qui se compose de trois chapitre :

Première chapitre :

Dans ce chapitre, on introduit les notions générales des probabilités et statistiques, en donnant la définition d'une loi de probabilité , Population , Caractère, échantillon , statistique, les moments empiriques act....

Deuxième chapitre :

Dans ce chapitre on présente le concept de l'estimation,on parle d'estimateur, la qualité d'un estimateur ,(estimateur sans biais, estimateur convergent, variance et erreur quadratique moyenne d'un estimateur,...act), exhaustive, information de Fisher

Troisième chapitre

Dans le dernier chapitre, présente, définition du maximum de vraisemblance , la méthode maximum de vraisemblance et propriétés asymptotiques ,exemple de l'EMV, application par logiciel R.

Chapitre 1

Rappel sur les notions probabilités et statistiques

Dans ce chapitre on peut donner quelques concepts de base au notions statistiques

1.1 Tribu

[3]

Soit Ω un ensemble quelconque non vide.

Définition 1.1.1 : Soit A un ensemble de parties de Ω ($A \subset \mathbb{P}(\Omega)$), on dit que A est un tribu (ou δ – algèbre) si vérifie les conditions suivantes :

1. $\{\emptyset\} \in A$.
2. A est stable par passage au complémentaire ($\forall B \in A : B \in A \iff \bar{B} \in A$).
3. A est stable par union dénombrable ($\forall n \in \mathbb{N}, B_n \in A \implies \cup_{n \in \mathbb{N}} B_n \in A$).

Le couple (Ω, A) s'appelle espace mesurable.

Remarque 1.1.1 :

- Une intersection des tribus est une tribu.
- Une union des tribus n'est pas nécessairement une tribu.

- Tribu des boréliens de \mathbb{R} (on note $B_{\mathbb{R}}$), c'est la plus petit tribu contenant tous les intervalles ouverts.
- L'ensembles des parties de Ω (on note $\mathbb{P}(\Omega)$), c'est la plus grand tribu.
- Si Ω un ensemble des possibilités d'un expérience aléatoire, (Ω, A) dit espace probabilisable.

Définition 1.1.2 Une mesure sur (Ω, A) est une fonction .

$$\mu : A \longrightarrow \bar{\mathbb{R}}_+ = [0, +\infty].$$

Telle que :

1. $\mu(\emptyset) = 0$.
2. $\forall (A_i)_{i \in I \subseteq \mathbb{N}} \quad \mu(\cup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$ si $\forall i, j \in I \quad A_i \cap A_j = \{\emptyset\}$.
3. $\forall i, j \in I \quad \mu(\cup_{i \in I} A_i) \leq \sum_{i \in I} \mu(A_i)$ si $\forall i, j \in I \quad A_i \cap A_j \neq \{\emptyset\}$.

Le triplet (Ω, A, μ) s'appelle espace mesuré .

Remarque 1.1.2 Si $\mu(\Omega) = 1$; la **mésure** μ dite probabilité noté par P , l'espace (Ω, A, P) s'appelle espace de probabilité.

1.1.1 Mesurabilité

Définition 1.1.3 Soient (Ω, A) et (E, ξ) deux espaces mesurables, une application $f : \Omega \rightarrow E$ est dite mesurable par rapport à (E, ξ) : $\forall B \in \xi$, si $f^{-1}(B) \in A$.

1.2 Variable aléatoire

[6]

Définition 1.2.1 : Soit (Ω, A, \mathbb{P}) espace de probabilité on dit une variable aléatoire X tout fonction sur Ω dans \mathbb{R}

$$X : \Omega \rightarrow \mathbb{R}.$$

Variable aléatoire c'est une fonction mesurable.

Remarque 1.2.1 : Il y a deux types des variables aléatoires discrètes et continues :

1.3 Loi d'une variable aléatoire

La loi de la variable aléatoire X est la mesure image de \mathbb{P} par X . Notée \mathbb{P}_x , définie par :

$$\forall B \in \zeta; \mathbb{P}_x(B) = \mathbb{P}(X^{-1}(B)).$$

En pratique on écrit :

$$\mathbb{P}_x(B) = \mathbb{P}(X \in B) = \mathbb{P}(\omega : X(\omega) \in B).$$

Remarque 1.3.1 : La loi d'une variable aléatoire dans le cas continue donne par la fonction de densité f vérifier :

1. $\forall x \in \mathbb{R}; f(x) \geq 0$.
2. $\int_{\mathbb{R}} f(x)dx = 1$.

1.4 Lois de probabilité

[10]

1.4.1 Lois discrètes

Loi de Bernoulli :

Une variable aléatoire X suit la loi de Bernoulli $B(p), 0 \leq p \leq 1$ si :

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

Espérance et Variance :

L'espérance de la loi bernoulli :

$$\mathbb{E}(X) = p.$$

La variance de la loi bernoulli :

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

La fonction de répartition :

$$F_x(x) = (1 - p)1_{[0,1[}(x) + 1_{[1,\infty[}(x).$$

La fonction caractéristique :

$$\varphi_x(t) = 1 - p + pe^{it}.$$

Loi binomiale :

La loi binomiale, de paramètres $n \in \mathbb{N}^*$ et $0 \leq p \leq 1$, $B(n, p)$ est loi de probabilité d'une variable aléatoire X d'une répétition de n épreuves de Bernoulli, telle que :

$$P(X = x_i) = \begin{cases} C_n^{x_i} p^{x_i} (1 - p)^{n - x_i} & \text{si } x_i = 0, 1, \dots, n. \\ 0 & \text{sinon.} \end{cases}$$

Espérance et Variance :

L'espérance de la loi binomiale :

$$\mathbb{E}(X) = np.$$

La variance de la loi binomiale :

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

La fonction de répartition :

$$F_X(x) = \begin{cases} \sum_{k=0}^x c_k p (1-p)^{n-x} & \text{si } x = 0, 1, \dots, n. \\ 0 & \text{sinon..} \end{cases}$$

La fonction caractéristique :

$$\varphi_x(t) = (1 - p + pe^{it})^n.$$

Loi de Poisson :

Une variable aléatoire X à valeurs dans \mathbb{R} suit une loi de Poisson $P(\lambda)$ de paramètre λ ,

($\lambda > 0$) si :

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

Espérance et Variance :

L'espérance de la loi poisson :

$$\mathbb{E}(X) = \lambda.$$

La variance de la loi poisson :

$$Var(X) = \lambda.$$

La fonction de répartition :

$$F_x(x) = e^{-\lambda} \sum_{x=0}^n \frac{\lambda^x}{x!}.$$

La fonction caractéristique :

$$\varphi_x(t) = e^{\lambda(e^{it}-1)}.$$

Loi uniforme :

Une variable aléatoire X suit une loi uniforme lorsque toutes les valeurs prises par X équiprobables. telle que :

$$\forall i \in N, P(X = x_i) = \begin{cases} \frac{1}{n} & \text{si } x_i = 1, \dots, n. \\ 0 & \text{sinon.} \end{cases}$$

Espérance et Variance :

L'espérance de la loi uniforme :

$$\mathbb{E}(X) = \frac{n + 1}{2}.$$

La variance de la loi uniforme :

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

La fonction de répartition :

$$F_X(x) = \left(\frac{x}{n}\right)1_{[0,n]}(x) + 1_{]n,\infty[}(x).$$

1.4.2 Lois continues

Loi uniforme :

Définition 1.4.1 *La loi Uniforme est la loi exacte de phénomènes continus uniformément répartis sur un intervalle*

la variable aléatoire X suit une loi uniforme sur l'intervalle $[a, b]$ si :

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b]. \\ 0 & \text{sinon.} \end{cases}$$

Espérance et Variance :

L'espérance de la loi uniforme :

$$\mathbb{E}(X) = \frac{b + a}{2}.$$

La variance de la loi uniforme :

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

La fonction de répartition :

$$F_X(x) = \int_{-\infty}^x \frac{1}{b - a} 1_{[a,b]}(t) dt = \begin{cases} 0 & \text{si } x \leq a. \\ \frac{x-a}{b-a} & \text{si } x \in [a, b]. \\ 1 & \text{si } x \geq b. \end{cases}$$

La fonction caractéristique :

$$\varphi_X(t) = \frac{e^{ibt} - e^{iat}}{i(b - a)t}.$$

Loi normale ou loi de Laplace-Gauss :

Une variable aléatoire absolument continue X suit une loi normale de paramètres (μ, σ) si sa densité de probabilité est donnée par : $f : \mathbb{R} \rightarrow \mathbb{R}$

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

Notation : $X \rightarrow \mathcal{N}(\mu, \sigma)$

Espérance et Variance

L'espérance de la loi normale vaut :

$$\mathbb{E}(X) = \mu.$$

La variance de la loi normale vaut :

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

La fonction de répartition :

$$F_X(x) = \int_0^\infty \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt.$$

La fonction caractéristique :

$$\varphi_x(t) = e^{i\mu t - \frac{t^2\sigma^2}{2}}.$$

La loi exponentielle :

Un variable aléatoire X réelle positive suit un loi exponentielle de paramètre $\lambda > 0$ si sa densité de probabilité est définie par :

$$f(x) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0. \\ 0 & x < 0. \end{cases}$$

Espérance et Variance :

L'espérance de la loi exponentielle :

$$\mathbb{E}(X) = \frac{1}{\lambda}.$$

La variance de la loi exponentielle :

$$\text{Var}(X) = \frac{1}{\lambda^2}.$$

La fonction de répartition :

$$F_X(x) = \int_\infty^x \lambda e^{-\lambda t} \mathbf{1}_{\mathbb{R}^+}(t) dt = \begin{cases} 0 & \text{si } x \leq 0. \\ 1 - e^{-\lambda x} & \text{si } x > 0. \end{cases}$$

La fonction caractéristique :

$$\varphi_x(t) = \frac{\lambda}{\lambda - it}.$$

1.5 Population et Caractère statistique

[1]

Population :

Définition 1.5.1 *Une population est un ensemble d'objets sur lesquels une étude se porte. Ces objets sont appelés individus.*

Caractère statistique

Définition 1.5.2 *Toute propriété étudiée sur les individus d'une population est appelée caractère. Statistique noté X .*

Nature d'un caractère

Une caractère est dit :

- Quantitatif s'il mesure une quantité ou un nombre, les valeurs que prend X sont numériques (le nombre de personnes dans une salle, le nombre d'items...) dans ce cas le caractère s'appelle variable statistique .
- Qualitatif s'il mesure une catégorie, les valeurs que prend X sont des modalités (la couleur des yeux, la marque de voitures...).

1.6 Échantillon

[10]

Définition 1.6.1 : *Soit X une variable aléatoire sur un référentiel Ω . Un échantillon de X de taille n est un n -uplet (X_1, \dots, X_n) de variables aléatoires indépendantes et identiquement*

distribuées (i.i.d) 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, \dots, x_n) où $X_i(\omega) = x_i$.

1.6.1 Échantillonnage aléatoire sans remise

Définition 1.6.2 : Soit N la taille d'une population finie à échantillonner. Un échantillon aléatoire sans remise est un échantillon dans lequel la sélection des éléments a la même probabilité.

Cela équivaut à choisir les n unités les unes après les autres au hasard parmi les unités indexées par $\{1, 2, \dots, N\}$, et en les sélectionnant pour l'échantillon final seulement si elles n'ont pas encore été choisies.

1.6.2 Échantillonnage aléatoire avec remise

Définition 1.6.3 : Le plan d'échantillonnage avec remise consiste à former un échantillon comme une liste de n unités de telle sorte que les unités soient choisies les unes après les autres parmi toutes les unités possibles et de manière aléatoire, tout en gardant la possibilité de sélectionner éventuellement plusieurs fois la même unité dans l'échantillon.

1.6.3 La distribution d'échantillonnage

Définition 1.6.4 : Distribution de probabilité composée de toutes les valeurs possibles d'une statistique d'échantillon

1.6.4 Statistique

[7]

Définition 1.6.5 : Soit X une variable aléatoire dont la loi P_θ dépend du paramètre $\theta \in \Theta$

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

$$T = T(X_1, X_2, \dots, X_n).$$

Remarque 1.6.1 *A un échantillon, on peut associer différentes statistique.*

Proposition 1.6.1 *La théorie de l'estimation consiste à définir des statistiques particulières, appelées estimateurs. Une fois l'échantillon effectivement réalisé, l'estimateur prend une valeur numérique appelée **estimation de paramètre** θ notée $\hat{\theta}$.*

Exemple 1.6.1 : Soit \bar{X} la statistique :

$$\bar{X} = T(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n x_i.$$

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'espérance mathématique.

1.6.5 Moments empiriques

[5]

Définition 1.6.6 : *On appelle moment empirique d'ordre $r \in \mathbb{N}^*$ la statistique :*

$$M_r = \frac{1}{n} \sum_{i=1}^n (x_i)^r.$$

Définition 1.6.7 : *On appelle moment empirique centré d'ordre $r \in \mathbb{N}^*$ la statistique :*

$$\widetilde{M}_r = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^r.$$

Exemple 1.6.2 1. Le moment empirique d'ordre 1 est la moyenne empirique :

$$M_1 = \frac{1}{n} \sum_{i=1}^n x_i = \overline{X}_n.$$

2. Le moment empirique centré d'ordre 2 est appelé la variance empirique et noté S_n^2 est donnée par l'expression :

$$\widetilde{M}_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \overline{X}_n)^2 = S_n^2.$$

S_n^2 peut s'écrire encore :

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

Remarque 1.6.2 : Soit X une variable aléatoire de moyenne μ et d'écart-type σ on a :

$$\mathbb{E}(\overline{X}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(x_i) = \mu.$$

$$\text{Var}(\overline{X}_n) = \frac{1}{n} \sum_{i=1}^n \text{Var}(x_i) = \frac{\text{Var}(x)}{n^2} = \frac{\sigma^2}{n}.$$

D'après l'indépendance des X_i .

Remarque 1.6.3 : Soit X une variable aléatoire d'écart-type σ et de moment centré d'ordre 4 $\mu_4 = \mathbb{E}(X - \mu)^4$ on a :

$$\mathbb{E}(S_n^2) = \mathbb{E}(X - \mu)^2 - \mathbb{E}(\overline{X}_n - \mu)^2.$$

$$\text{var}(X) - \text{var}(\overline{X}_n) = \frac{n-1}{n} \sigma^2.$$

Remarque 1.6.4 : Les moments empiriques ont les propriétés asymptotiques suivantes :

Si $\mathbb{E}|X|^r < \infty$, alors :

$$M_r \xrightarrow{p.s} \mathbb{E}(x^r) = m_r.$$

$$\widetilde{M}_n \xrightarrow{p.s} \mathbb{E}[(X - \mathbb{E}(X))^r] = \mu_r.$$

m_r (resp μ_r) s'appellent *moments théoriques* (resp *centré*) d'ordre r :

Si $\mathbb{E}|X|^2 < \infty$, alors :

$$\sqrt{n}(\bar{X}_n - m_1) \xrightarrow{\text{loi}} N(0, \mu_2).$$

Si $\mathbb{E}|X|^4 < \infty$, alors :

$$\sqrt{n}(S_n^2 - \mu_2) \xrightarrow{\text{loi}} N(0, \mu_4 - \mu_2^2).$$

Chapitre 2

Estimation statistique paramétrique

2.1 Cadre de l'estimation paramétrique

[1]

L'estimation paramétrique consiste à estimer un ou plusieurs paramètres inconnus appartenir à X (valeur moyenne, variance proportion ...) à partir des données x_1, \dots, x_n . Naturellement, l'estimation du ou des paramètres inconnus doit être aussi précise que possible. Les trois grands types d'estimation paramétriques sont :

- L'estimation ponctuelle : on estime directement par un ou des réels le ou les paramètres inconnus
- L'estimation par intervalles de confiance : on détermine des intervalles de réels, les moins étendus possible, qui ont de fortes chances de contenir un paramètre inconnu
- Les tests statistiques : démarches qui consistent à accepter ou non une hypothèse mettant en jeu un ou plusieurs paramètres inconnus, avec un faible risque de se tromper.

2.2 Estimateur

[4]

Soit X une variable aléatoire dont la loi de probabilité P_θ dépend d'un seul paramètre θ .

Soit (X_1, \dots, X_n) un échantillon aléatoire.

Définition 2.2.1 *Un estimateur est une statistique qui a des propriétés bien définies.*

$T_n = T(X_1, \dots, X_n)$ est une suite de statistique définie :

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

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 sure ou moyenne quadratique.

Notation 2.2.1 *En général, on note $\hat{\theta}$ l'estimateur du paramètre θ .*

2.3 Qualité d'un estimateur

Quelque estimateurs qui fournis la puiisse des bonnes estimateurs parmi toutes les statistiques possibles, il doit posséder le meilleur estimateur.

2.3.1 Le biais d'un estimateur

Définition 2.3.1 : *On dit que $\hat{\theta}$ est un estimateur **avec biais** de θ si :*

$$\mathbb{E}(\hat{\theta}) \neq \theta \quad \text{c-à-d}$$

$$B(n, \theta) = \mathbb{E}(\hat{\theta} - \theta) = \mathbb{E}(\hat{\theta}) - \theta \neq 0.$$

Exemple 2.3.1 : *Soit (X_1, X_2, \dots, X_n) un n échantillon S^2 est un estimateur biaisé de σ^2 .*

$$S^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{X})^2 = \frac{1}{n} \sum_{i=1}^n x_i^2 - (\bar{X})^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 - (\bar{X} - \mu)^2.$$

$$\mathbb{E}(S^2) = \sigma^2 + m^2 - \left(\frac{\sigma^2}{n} - m^2 \right) = \sigma^2 - \frac{\sigma^2}{n} = \frac{(n-1)\sigma^2}{n} \neq \sigma^2.$$

L'estimateur est dit **sans biais** ou non biais de θ si: $\mathbb{E}(\hat{\theta}) = \theta$.

Exemple 2.3.2 : Soit (X_1, X_2, \dots, X_n) un n échantillon telles que

$$\mathbb{E}(x_i) = \mu, \forall i = 1, \dots, n, \quad \bar{X} = \frac{1}{n} \sum_{i=1}^n (x_i).$$

(La moyenne empirique) est un estimateur sans biais de μ , en effet :

$$\mathbb{E}(\bar{X}) = \mathbb{E}\left(\frac{1}{n} \sum_{i=1}^n x_i\right) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(x_i) = \frac{1}{n} n \mathbb{E}(x_i) = \frac{1}{n} n \mu = \mu$$

2.3.2 Variance et erreur on moyenne quadratique d'un estimateur

Définition 2.3.2 L'erreur quadratique moyenne mesure la précision d'un estimateur $\hat{\theta}$ du paramètre θ par

$$EQM(\hat{\theta}) = \mathbb{E}[(\hat{\theta} - \theta)^2].$$

Théorème 2.3.1 : Soit $\hat{\theta}$ un estimateur du paramètre θ alors :

$$EQM(\hat{\theta}) = \text{Var}(\hat{\theta}) + B(n, \hat{\theta})^2.$$

Preuve. : On a

$$\begin{aligned} EQM(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \theta)^2], \\ &= \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}) + \mathbb{E}(\hat{\theta}) - \theta)^2], \\ &= \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2 + 2(\hat{\theta} - \mathbb{E}(\hat{\theta}))(\mathbb{E}(\hat{\theta}) - \theta) + (\mathbb{E}(\hat{\theta}) - \theta)^2]. \end{aligned}$$

Comme $\mathbb{E}(\hat{\theta}) - \theta$ est une constante et que $\mathbb{E}(\hat{\theta} - \mathbb{E}(\hat{\theta})) = 0$, on effectue :

$$\begin{aligned} EQM(\hat{\theta}) &= \mathbb{E}[(\hat{\theta} - \mathbb{E}(\hat{\theta}))^2] + [\mathbb{E}(\hat{\theta}) - \theta]^2, \\ &= \text{Var}(\hat{\theta}) + B(n, \hat{\theta})^2. \end{aligned}$$

■

Remarque 2.3.1 *Pour réaliser la plus petite possible erreur quadratique moyenne, il faut que :*

1. $E(\hat{\theta}) = \theta$, donc choisir un estimateur sans biais.
2. $Var(\hat{\theta})$ 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.

2.3.3 Estimateur convergent

[8]

Dans ce paragraphe, nous considérons un estimateur $\hat{\theta}$ dont la loi dépend de paramètre θ qu'on s'intéresse un bon estimateur $\hat{\theta}$. 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 $\hat{\theta}$ vers la vraie valeur du paramètre θ .

Définition 2.3.3 : *L'estimateur est dit **convergent (ou consistant)** si la suite $\hat{\theta}$ converge en probabilité vers θ ($\hat{\theta} \xrightarrow{p} \theta$)*

$$\forall \varepsilon > 0, \forall \theta \in \Theta \lim_{n \rightarrow \infty} P(|\hat{\theta} - \theta| > \varepsilon) = 0.$$

Remarque 2.3.2 *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 :*

$$\forall \theta \in \Theta \lim_{n \rightarrow \infty} \mathbb{E}(\hat{\theta}) = \theta \text{ et } \lim_{n \rightarrow \infty} Var(\hat{\theta}) = 0.$$

Pour vérifier ce critère, il faut connaître la loi de X et celle de $\hat{\theta}$ pour calculer l'espérance et la variance.

Exemple 2.3.3 : On a vu que \bar{X} est un estimateur sans biais .

alors :

$$\mathbb{E}(\bar{X}) = m.$$

Et

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Donc \bar{X} est un estimateur convergent.

2.4 Exhaustivité

[2]

Définition 2.4.1 : On dit que $\hat{\theta}$ est une statistique exhaustive pour $\theta \in \Theta \subseteq \mathbb{R}^k$ si la loi conditionnelle de $X = (X_1, X_2, \dots, X_n)$ sachant $\hat{\theta}(x) = k$ ne dépend pas de θ .

$$P((X_1, X_2, \dots, X_n) / \hat{\theta}(x) = k).$$

Théorème 2.4.1 : *factorisation de Fisher-Neyman*

La statistique $\hat{\theta}(x) = k(X_1, X_2, \dots, X_n)$ est exhaustive pour θ si et seulement si la densité de probabilité (ou fonction de probabilité) conjointe s'écrit pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$ sous la forme :

$$f(x_1, \dots, x_n, \theta) = g((x_1, \dots, x_n); \theta)h(x_1, \dots, x_n).$$

Où $g(t(x_1, \dots, x_n))$ la densité de θ la statistique T_n .

Et $h(x_1, \dots, x_n)$ ne dépend de θ .

Exemple 2.4.1 : La loi de poisson du paramètre λ inconnu .

$$\begin{aligned} L(x_1, \dots, x_n; \lambda) &= \prod_{i=1}^n \exp(-\lambda) \frac{\lambda^{x_i}}{x_i!}, \\ &= \exp(-n\lambda) \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}. \end{aligned}$$

Alors

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

Et

$$h(x_1, \dots, x_n) = \sum_{i=1}^n x_i.$$

2.5 Information de Fisher

Soit $(\mathbf{E}, \mathcal{A}, P_\theta)$ un espace de probabilité tels que :

\mathbf{E} est l'ensemble fondamentale, \mathcal{A} est une tribu sur \mathbf{E} , et P_θ est une probabilité définie sur l'espace probabilisable $(\mathbf{E}, \mathcal{A})$.

Définition 2.5.1 : On appelle **quantité d'information de Fisher** $I_n(\theta)$ donnée par un n échantillon sur le paramètre θ , la valeur positive ou nulle suivante I .

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

Où

$$L(X_1, X_2, \dots, X_n; \theta) = \prod_{i=1}^n f(X; \theta).$$

Théorème 2.5.1 : Si le domaine de définition de X ne dépend pas du paramètre θ alors :

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

Exemple 2.5.1 : Soit X une v.a telle que $X \sim N(m, \sigma^2)$ alors l'information de Fisher pour cette v.a est :

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

$$\ln L(x; \theta) = \frac{-1}{2} \ln(2\pi) - \ln(\sigma) - \frac{1}{2} \left(\frac{x-m}{\sigma} \right)^2,$$

$$\frac{\partial}{\partial m} \ln L(x; \theta) = \frac{1}{\sigma^2} (x-m),$$

$$\frac{\partial^2}{\partial m^2} \ln L(x; \theta) = -\frac{1}{\sigma^2},$$

$$I(m) = -E \left[\frac{\partial^2}{\partial m^2} \ln L(x; \theta) \right] = \frac{1}{\sigma^2}.$$

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

Nous indiquons dans le suivant 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 $\hat{\theta}$ sans biais de θ :

$$V(\hat{\theta}) \geq \frac{1}{I_n(\theta)}.$$

et si $\hat{\theta}$ est un estimateur sans biais de $k(\theta)$:

$$V(\hat{\theta}) \geq \frac{[K'(\theta)]^2}{I_n(\theta)}.$$

où k est une fonction dérivable.

Définition 2.6.1 : On dit que T est un estimateur efficace si :

1. T est un estimateur sans biais de θ alors

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

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

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

Théorème sur l'efficacité

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

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

Car $\hat{\theta}$ 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 :

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

Chapitre 3

Méthod du maximum de vraisemblance

Dans ce chapitre, on a étudié l'estimation par la méthode du maximum de vraisemblance de paramètre qui consiste à associer à un échantillon et en donnant quelques propriétés.

3.1 fonction $L(x, \theta)$ de vraisemblance

[7]

Soit $(x_1, \dots, x_n; \theta)$, n échantillon de loi $f(x_i; \theta)$, on appelle $L(x; \theta)$ la fonction vraisemblance d'échantillon x_1, \dots, x_n la densité noté $L(x_1, \dots, x_n)$

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

Définition 3.1.1 x_1, \dots, x_n indépendents

Définition 3.1.2 Soient x_1, \dots, x_n n échantillon de la vraisemblance $L(x; \theta) = L(x_1, \dots, x_n; \theta)$ on dit que $\hat{\theta}$ estimateur de vraisemblance de θ si $\hat{\theta}$ est maximisé, la fonction $L(x; \theta)$.

Alors $\hat{\theta}$ est une solution de système suivant :

$$\begin{cases} \frac{\partial L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta}(\theta) = 0. \\ \frac{\partial^2 L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta^2}(\hat{\theta}) \leq 0. \end{cases}$$

Où

$$\begin{cases} \frac{\partial \ln L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta}(\theta) = 0. \\ \frac{\partial^2 \ln L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta^2}(\hat{\theta}) \leq 0. \end{cases}$$

3.2 Méthode du maximum de vraisemblance

[4]

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 :

$$\hat{\theta}^{MV} = \arg \max L(x_1, \dots, x_n; \theta).$$

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 :

$$\hat{\theta}^{MV} = \arg \max \ln L(x_1, \dots, x_n; \theta).$$

Pour trouver le maximum, on résoud l'équation du premier ordre :

$$\frac{\partial L(x_1, \dots, x_n; \theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}^{MV}} = 0.$$

La théorie nous dit que la solution de cette équation nous donne toujours le maximum, on obtient $\hat{\theta}^{MV}$ sous la forme

$$\hat{\theta}^{MV} = g(x_1, \dots, x_n).$$

3.3 Propriétés asymptotiques

L'estimateur obtenu par la méthode du maximum de vraisemblance n'est pas forcément unique (la vraisemblance peut avoir plusieurs maxima), ni sans biais, ni de variance minimale, ni efficace. présente des bonnes propriétés :

Propriété 3.3.1 :

Si les X_i sont indépendants et de même loi dépendant d'un paramètre réel on a :

1. $\hat{\theta}_n$ converge presque sûrement vers θ .
2. $\sqrt{I_n(\theta)}(\hat{\theta}_n - \theta) \xrightarrow{Loi} \mathcal{N}(0, 1)$ ce qui signifie que, quand n est grand $\hat{\theta}_n$, est approximativement de loi $\mathcal{N}(0, \frac{1}{I_n(\theta)})$, On déduit que $\hat{\theta}_n$ est asymptotiquement gaussien, sans biais et efficace. Cette propriété peut aussi s'écrire :

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{Loi} \mathcal{N}(0, \frac{1}{I_1(\theta)}).$$

3. Si $\hat{\theta}_n$ est l'EMV de θ , alors $\varphi(\hat{\theta}_n)$ est l'EMV de $\varphi(\theta)$. De plus, si φ est dérivable, on a :

$$\sqrt{n}(\varphi(\hat{\theta}_n) - \varphi(\theta)) \xrightarrow{Loi} \mathcal{N}(0, \frac{\varphi'(\theta)^2}{I_1(\theta)}).$$

Ce résultat est connu sous le nom de méthode delta. Quand n est grand, $\varphi(\hat{\theta}_n)$, est donc approximativement de loi $\mathcal{N}(0, \frac{\varphi'(\theta)^2}{I_n(\theta)})$.

3.4 Exemple de l'EMV

3.4.1 Loi discrète

Loi de poisson

On souhaite estimer le paramètre θ d'une loi de Poisson à partir d'un n échantillon. On a

$P(x, \theta) = P_\theta(X = x) = \frac{\theta^x e^{-\theta}}{x!}$. La fonction de vraisemblance s'écrit :

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

Il est plus simple d'utiliser le logarithme, la vraisemblance étant positive :

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

La dérivée première :

$$\frac{\partial \ln(X_1, X_2, \dots, X_n; \theta)}{\partial \theta} = -n + \frac{1}{\theta} \sum_{i=1}^n x_i.$$

S'annule pour $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i$. la dérivée seconde :

$$\frac{\partial^2 \ln(X_1, X_2, \dots, X_n; \theta)}{\partial \theta^2} = -\frac{1}{\theta^2} \sum_{i=1}^n x_i.$$

Est toujours négative ou nulle. ainsi l'estimation donnée par, $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i = \bar{X}$ conduit à un estimateur du maximum de vraisemblance égal à $\hat{\theta} = \bar{X}$. il est normal de retrouver la moyenne empirique qui est le meilleur estimateur possible pour le paramètre θ .

3.4.2 Lois continues

Variable aléatoire de loi normale :

On suppose que $X \sim \mathcal{N}(m, \sigma^2)$ et sa densité est :

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

de moyenne m et variance σ^2 sont des inconnues, on tire un échantillon X_1, \dots, X_n *iid* de taille n les valeurs observées sont x_1, \dots, x_n .

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

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

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

Soit :

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

Donc :

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

Puis, on a le logarithme de la vraisemblance :

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

On a l'équation de vraisemblance :

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

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

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

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

Soit toujours

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

L'équation de la vraisemblance est :

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

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

$$\begin{aligned} \frac{-n}{2\sigma^2} + \sum_{i=1}^n \frac{(x_i - m)^2}{2\sigma^4} = 0 &\implies \frac{-n\sigma^2 + \sum_{i=1}^n (x_i - m)^2}{2\sigma^4} = 0. \\ &\implies -n\sigma^2 + \sum_{i=1}^n (x_i - m)^2 = 0. \end{aligned}$$

Alors

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

Donc on remplace par sa valeur devient

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

3.5 Remarques

Cette méthode du maximum de vraisemblance parfois ne donne pas de résultat direct.

Exemple 3.5.1 Soit $X = (X_1, X_2, \dots, X_n) \sim U_{[0, \theta]}$, alors :

$$\begin{aligned} L(X; \theta) &= \prod_{i=1}^n \frac{1}{\theta} I_{[0, \theta]}(X_i), \\ &= \frac{1}{\theta^n} I_{\{0 \leq \inf_{1 \leq i \leq n} X_i \leq \sup_{1 \leq i \leq n} X_i \leq \theta\}}, \\ \ln L(X; \theta) &= -n \ln \theta I_{[0, \theta]}(X_i). \end{aligned}$$

Et

$$I_{[0, \theta]}(x_i) = \begin{cases} 1 & \text{si } x_i \in [0, \theta]. \\ 0 & \text{sinon.} \end{cases}$$

Alors

$$\begin{cases} \frac{\partial \ln L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta} = -\frac{n}{\theta} = 0 \dots \dots (*). \\ \frac{\partial^2 \ln L(X_1, X_2, \dots, X_n; \theta)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}^{MV}} = -\frac{n}{\theta^2} \leq 0. \end{cases}$$

Exemple 3.5.2 Mais l'équation (*) n'a pas de sens on peut écrire la fonction de vraisemblance en fonction de θ sous la forme

$$L(X; \theta) = \frac{1}{\theta^n} I_{[\max x_i; +\infty[}(\theta),$$

Exemple 3.5.3 Qui admet une maximum pour $\theta = \max_{1 \leq i \leq n} x_i X_i$.

donc

$$\hat{\theta} = \max_{1 \leq i \leq n} X_i.$$

Est l'estimateur du maximum de vraisemblance de θ :

Exemple 3.5.4 : Soit $(x_1, \dots, x_n; \theta)$, n échantillon de loi $\Gamma(a; b)$, la fonction de vraisemblance

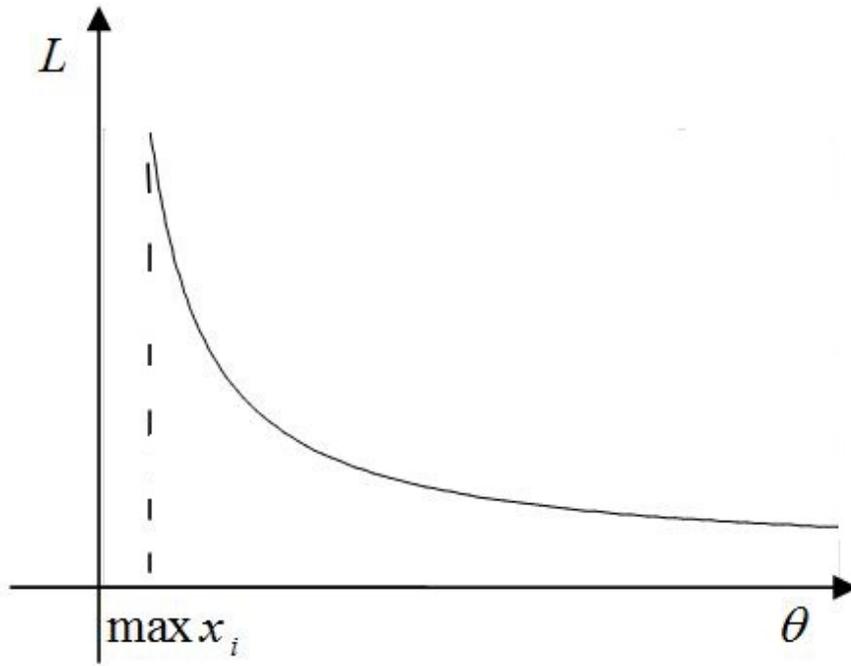


FIG. 3.1 – Figure : courbe de la fonction de vraisemblance

est donnée par :

$$L(x_1, \dots, x_n; a; b) = \prod_{1 \leq i \leq n} \frac{b^a}{\Gamma(a)} x_i^{a-1} e^{-bx_i} = \frac{b^{na}}{[\Gamma(a)]^n} e^{-b \sum_{1 \leq i \leq n} x_i} \prod_{1 \leq i \leq n} x_i^{a-1}$$

D'où

$$\ln L(x_1, \dots, x_n; a; b) = na \ln b - n \ln \Gamma(a) - b \sum_{1 \leq i \leq n} x_i + (a-1) \sum_{1 \leq i \leq n} \ln x_i$$

Donc

$$\left\{ \begin{array}{l} \frac{\partial \ln L(X_1, X_2, \dots, X_n; a; b)}{\partial a} = n \ln b - n \frac{\Gamma'(a)}{\Gamma(a)} + \sum_{1 \leq i \leq n} \ln x_i = 0 \dots \dots \dots (1) \\ \frac{\partial \ln L(X_1, X_2, \dots, X_n; a; b)}{\partial b} = \frac{na}{b} - \sum_{1 \leq i \leq n} x_i = 0 \dots \dots \dots (2) \end{array} \right.$$

D'après (2) on trouve $b = \frac{a}{\bar{x}}$ si on remplace ce résultat dans (1) on obtient que \hat{a} est solution de l'équation implicite

$$n \ln a - n \ln \bar{x} - n \frac{\Gamma'(a)}{\Gamma(a)} + \sum_{1 \leq i \leq n} \ln x_i = 0$$

cette équation est à résoudre par des méthodes numériques, si $\hat{\lambda}$ déterminé on en déduit $\hat{b} = \frac{\hat{\lambda}}{\bar{x}}$

3.6 Application en R

3.6.1 Loi discrète

Code pour la loi de poisson $P(\lambda)$

```
X=rpois(100,2)
```

Méthode du maximum de vraisemblance

```
##Fonction pour déterminer l'EMV numériquement
```

```
mv=function(Y){
```

```
  m=mean(Y)
```

```
  n=length(Y)
```

```
  logL=function(lamda){-(sum(Y)*log(lamda)-n*lamda)}
```

```
  mv=optim(m,logL)$par
```

```
  return(mv)}
```

```
mv(X)
```

Résultat numérique $\hat{\lambda}=1.88$

```
##Le biais de L'EMV
```

```
biaismv=function(n,m,lamda){
```

```
  col=numeric(m)
```

```
  for(i in 1 :m){col[i]=mv(rpois(n,lamda))}
```

```
  biais=mean(col)-lamda
```

```
  return(biais)}
```

```
biaismv(100,500,2)
```

Résultat numérique [1] -0.00062

```
#Erreur quadratique moyenne de L'EMV
EQMmv=function(n,m,lamda){
col=numeric(m)
for(i in 1 :m){col[i]=mv(rpois(n,lamda))}
biais=mean(col)-lamda
EQM=var(col)+biais^2
return(EQM)}
EQMmv(100,500,2)
Résultat numérique [1]0.019920840
#####
```

3.6.2 Loi continue

Code pour la loi normal $N(\mu; \sigma)$

```
x=rnorm(100,0,3)
Méthode du maximum de vraisemblance
##Fonction pour déterminer l'EMV numériquement
mv=function(Y){
logL=function(theta,ech=Y){
mu=theta[1]
sig2=theta[2]
n=length(ech)
a1=-(n/2)*log(2*pi)-(n/2)*log(sig2)
a2=-1/(2*sig2)
z=(ech-mu)^2
logv=(-(a1+a2*sum(z)))}
```

```
mv=optim(c(0,3),logL)$par
return(mv)}

mv(x)

Resultat numerique  $\hat{\mu} = -0.1710185$  et  $\hat{\sigma}^2 = 8.8084780$ 

#####Le biais de L'EMV

biaismv=function(n,m,mu,sig2){
col1=numeric(m)
col2=numeric(m)
ma=matrix(data=0,nrow=m,ncol=2)
for(i in 1 :m){
ma[i]=mv(rnorm(n,mu,sig2))}
for(i in 1 :m){
col1[i]=ma[i,1]
col2[i]=ma[i,2]}
biais1=mean(col1)-mu
biais2=mean(col2)-sig2
return(c(biais1,biais2))}

biaismv(100,500,0,3)

Résultat numérique[1] 0.00732572 -3.00000000

#####Erreur quadratique moyenne

EQMmv=function(n,m,mu,sig2){
col1=numeric(m)
col2=numeric(m)
ma=matrix(data=0,nrow=m,ncol=2)
for(i in 1 :m){
```

```
ma[i]=mv(rnorm(n,mu,sig2))}
for(i in 1 :m){
col1[i]=ma[i,1]
col2[i]=ma[i,2]}
biais1=mean(col1)-mu
biais2=mean(col2)-sig2
EQM1=var(col1)+biais1^2
EQM2=var(col2)+biais2^2
return(c(EQM1,EQM2))}
EQMmv(100,500,0,3)
Résultat numérique[1] 0.0951526 9.0000000
```

Conclusion

Dans ce travail, nous avons étudié l'une des méthodes d'estimation qui est la méthode du maximum de vraisemblance, c'est la plus utilisée, parce que la vraisemblance est une fonction qui contient toute l'information des données sur le paramètre inconnu.

Elle joue un rôle important dans de nombreuses méthodes statistiques, possède des propriétés intéressantes, notamment : la convergence, l'efficacité, c'est le meilleur des estimateurs, mais malgré la qualité d'estimation, il présente des inconvénients comme si l'ensemble de définition dépend du paramètre que l'on veut estimer.

Bibliographie

- [1] Christophe Chesneau, Caen (2018). cour sur l'Estimateur du Maximum de Vraisemblance
- [2] Gilbert Saporta, (2006) ,Probabilités Analyse des données et Statistique, Université de Paris, France
- [3] Jeanblanc, Monique, Master. (2006). Cours de calcul stochastique. 2IF EVRY
- [4] Michel Lejeune. Springer. (2010). Statistique. La théorie et ses applications.
- [5] Oliver Gaudoin, Note de cour deuxième. principes et Méthodes statiques
- [6] Philippe Huber, Ronchetti Elvezio. (2006). Maitriser l'aleatoire _ exercices resolus de probabilites et statistique
- [7] Pierre DUSART, (2018). Cour de Statistique inférentielles, Licence 2-S4-SI-MASS
- [8] Pierre Lecoutre , Statistique et Probabilités .DUNOD.
- [9] Thibault LAURENT, (2019) , Prise en main Du logiciel R.
- [10] Yadolah Dodge, Giuseppe Melfi. Springer (2008) . Premiers pas en simulation.

Annexe A : Logiciel R

[9]

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

R est un système, communément appelé langage et logiciel, qui permet de réaliser des analyses statistiques. Plus particulièrement, il comporte des moyens qui rendent possible la manipulation des données, les calculs et les représentations graphiques. *R* a aussi la possibilité d'exécuter des programmes stockés dans des fichiers textes et comporte un grand nombre de procédures statistiques appelées paquets. Ces derniers permettent de traiter assez rapidement des sujets aussi variés que les modèles linéaires (simples et généralisés), la régression (linéaire et non linéaire), les séries chronologiques, les tests paramétriques et non paramétriques classiques, les différentes méthodes d'analyse des données,...plusieurs paquets, tels *ade4*, *FactoMineR*, *MASS*, *multivariate*, *scatterplotdetrgl* entre autres sont destinés à l'analyse des données statistiques multidimensionnelles, il a été initialement créé, en 1996, par *Robert Gentleman* et *Ross Ihaka* du département de statistique de l'Université d'Auckland en Nouvelle Zélande. Depuis 1997, il s'est formé une équipe "*R Core Team*" qui développe *R*. Il est conçu pour pouvoir être utilisé avec les systèmes d'exploitation *Unix*, *Linux*, *Windows* et *MacOS*.

Un élément clé dans la mission de développement de *R* est le *ComprehensiveR Archive Network* (CRAN) qui est un ensemble de sites qui fournit tout ce qui est nécessaire à la distribution de *R*, ses extensions, sa documentation, ses fichiers sources et ses fichiers binaires. Le site maître du CRAN est situé en Autriche à Vienne, on peut y accéder par l'URL :

"[http : //cran.r – project.org/](http://cran.r-project.org/)". Les autres sites du CRAN, appelés sites miroirs, sont répandus partout dans le monde.

R est un logiciel libre distribué sous les termes de la "GNU Public Licence". Il fait partie intégrante du projet GNU et possède un site officiel à l'adresse "[http : //www.R – project.org](http://www.R-project.org/)".

Il est souvent présenté comme un clone de *S* qui est un langage de haut niveau développé par les AT&T Bell Laboratories et plus particulièrement par *RickBecker*, *JohnChambers* et *Allan,Wilks*. *S* est utilisable à travers le logiciel *S-Plus* qui est commercialisé par la société Insightful ([http : //www.splus.com/](http://www.splus.com/)).

Annexe B : Abréviations et Notations

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

$\hat{\theta}^{MV}$	Estimateur par la méthode maximum de vraisemblance.
Θ	Ensemble des paramètres.
<i>i.i.d</i>	Indépendantes et identiquement distribuées.
l	Fonction <i>log</i> vraisemblance.
$E(X)$	moyenne théorique X .
$V(X)$	Variance de la v.a X .
\bar{X}	La moyenne empirique.
$L(X; \theta)$	Fonction de vraisemblance.
P_θ	Loi de paramètre θ .
$F(x; \theta)$	Fonction de répartition.
$B(n, \theta)$	Le biais de l'estimateur T_n .
(X_1, \dots, X_n)	Échantillon aléatoire.
(x_1, \dots, x_n)	Réalisation de l'échantillon.
\xrightarrow{L}	Convergence en loi.
$\xrightarrow{p.s}$	Convergence presque sûre.
$I_n(\theta)$	Information de Fisher.

المخلص

و الاحتمالات ميدان في مهم رياضي مفهوم الى التطرق هو العمل هذا من الغرض مع القصوى الاحتمالية تقدير هو و اخرى علمية ميادين في وكذلك الاحصاء بالبرمجة مدروسة للأمتلة عددي تقدير اعطاء R. الكلمات المفتاحية: العينة, الاحصاء, التقدير, اللحظة التجريبية, الاحتمالية القصوى.

Résumé

Le but de ce travail est d'aborder un concept mathématique important dans les domaines des probabilités et des statistiques ainsi que dans d'autres domaines scientifiques, c'est une estimation du maximum de vraisemblance, donner une estimation numérique des exemples qui étudiés en logiciel : R.

Mots clés: échantillon, statistique, estimateur, moment empirique, maximum de vraisemblance.

Abstract

The purpose of this work is to address an important mathematical concept in the field of probability and statistics as well as in other scientific fields, it is an estimate of the maximum likelihood giving a numerical estimate of the examples studied using software : R.

Key words : sample , statistics, estimate, empirical moment, maximum likelihood.