

Chapitre 3

Estimation Ponctuelle

3.1 Généralités (Cas réel $\theta \in \Theta \subset \mathbb{R}$)

Définition 1 Soit $\rho = (\mathcal{X}, \mathcal{B}, \{P_\theta, \theta \in \Theta\})$ un modèle statistique associé à une observation x d'une v.a X de loi P_θ . Soit :

$$g : \Theta \subset \mathbb{R}^p \rightarrow \mathbb{R}^k$$

$$\theta \rightsquigarrow g(\theta)$$

On appelle estimateur de $g(\theta)$ toute statistique :

$$T : \mathcal{X} \rightarrow g(\Theta) \subset \mathbb{R}^k$$

$$x \rightsquigarrow T(x)$$

$T(x)$ est l'estimateur de $g(\theta)$.

Définition 2 Un estimateur T de $g(\theta)$ est dit sans biais ssi :

$$E[T(X) - g(\theta)] = 0 \Leftrightarrow E[T(X)] = g(\theta) \cdot \forall \theta \in \Theta.$$

où, $E(T(X)) - g(\theta)$ est appelé le biais.

Définition 3 Une suite d'estimateurs $(T_n)_{n \in \mathbb{N}}$ de la fonction $g(\theta)$ est dite asymptotiquement sans biais ssi :

$$\lim_n E(T_n(X)) = g(\theta).$$

Exemple : Trouver un E.S.B apporté par n ech de $X \rightsquigarrow N(m, \sigma^2)$.

Remarques : i/ Si $E(T) = g(\theta) + a$: T n'est pas un ESB de $g(\theta) \Rightarrow T' = T - a$ est un ESB de $g(\theta)$.

ii/ Si $E(T) = ag(\theta)$: T n'est pas un ESB de $g(\theta) \Rightarrow T' = \frac{T}{a}$ est un ESB de $g(\theta)$.

Définition 4 Une suite d'estimateurs $(T_n)_{n \in \mathbb{N}}$ est dite convergente en probabilité (resp. PS, m.q) si :

$$T_n \xrightarrow[\text{P.S.}]{\text{proba}} g(\theta).$$

3.2 Méthodes d'estimation paramétrique

3.2.1 Méthode des moments

Elle consiste à évaluer les moments empiriques d'un n échantillon ou d'une v.a avec les moments théoriques du même ordre. (Si on a k paramètres à estimer, on obtient un système de k équations).

Exemple : Estimer les paramètres de la v.a X qui suit les lois suivantes :

1°. $X \rightsquigarrow \mathcal{B}(p)$.

2°. $X \rightsquigarrow \mathcal{E}(\lambda)$.

3°. $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$.

4°. $X \rightsquigarrow \mathcal{U}_{[\theta_1, \theta_2]}$.

3.2.2 Méthode de maximum de vraisemblance

a/ Statistique du maximum de vraisemblance

Définition 5 On appelle statistique du maximum de vraisemblance la statistique $\hat{\theta}(x_1, \dots, x_n)$ telle que :

$$\mathcal{L}(\hat{\theta}(x_1, \dots, x_n), \underline{x}) = \max_{\theta} \mathcal{L}(\theta, \underline{x})$$

La maximisation de cette fonction est réécrite comme suit :

$$\max_{\theta} \mathcal{L}(\theta, \underline{x}) = \mathcal{L}(\hat{\theta}, \underline{x}) \Leftrightarrow \max_{\theta} \log \mathcal{L}(\theta, \underline{x}) = \log \mathcal{L}(\hat{\theta}, \underline{x})$$

Cette dernière fournit les équations dites de maximum de vraisemblance :

$$\begin{cases} \frac{\partial}{\partial \theta} \mathcal{L}(\theta, \underline{x}) = 0 \text{ et } \frac{\partial^2}{\partial \theta^2} \mathcal{L}(\theta, \underline{x}) < 0 \Leftrightarrow \frac{\partial}{\partial \theta} \log \mathcal{L}(\theta, \underline{x}) = 0 \\ \frac{\partial^2}{\partial \theta^2} \log \mathcal{L}(\theta, \underline{x}) < 0 \end{cases}$$

Définition 6 On appelle Estimateur du maximum de vraisemblance de paramètre θ , un estimateur donné par la statistique du maximum de vraisemblance.

Exemple : Trouver l'e.m.v des paramètres des v.a suivantes :

1°. $X \rightsquigarrow \mathcal{B}(p)$.

2°. $X \rightsquigarrow \mathcal{U}_{[0, \theta]}$.

3°. $X \rightsquigarrow \mathcal{N}(m, \sigma^2)$.

Proposition 7 (Invariance fonctionnelle) Soit : $g : \Theta \rightarrow \Theta$

$$\theta \rightsquigarrow g(\theta).$$

Alors, l'e.m.v de $g(\theta)$ noté $\widehat{g(\theta)} = g(\hat{\theta})$ où $\hat{\theta}$ est un e.m.v de θ .

Exemple : Soit $X \rightsquigarrow \mathcal{P}(\lambda)$. Trouver l'e.m.v de $\lambda e^{-\lambda} = P(X = 1)$?

3.3 Estimation non paramétrique

Théorème 8 Soit X une v.a admettant des moments d'ordre r et (X_1, \dots, X_n) un n éch de $X \Rightarrow \frac{1}{n} \sum_{i=1}^n X_i^k$ est un ESB, convergent P.S, de $E(X^k) \forall k \leq r$.

Théorème 9 Si X admet un moment d'ordre 2. Alors : $S_n^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2$ est un ESB, cv p.s de $\sigma^2 = \text{var}(X)$.

Démonstration : en cours.

Remarque : S_n^2 n'est pas un ESB de σ^2 .

Théorème 10 Soit (X, Y) un couple de v.a admettant des moments d'ordre 2 et soit $(X_1, Y_1), \dots, (X_n, Y_n)$ un n éch de (X, Y) .

Alors : $S_{X,Y}^2 = \frac{1}{n-1} \sum (x_i - \bar{x})(y_i - \bar{y})$ est un ESB cv ps de $\text{cov}(X, Y) = C_{X,Y}$.

Démonstration : en cours.

Théorème 11 La fréquence $F_n(A) = \frac{1}{n} \sum_{i=1}^n 1_A(x_i)$ est un ESB cv ps de $P(X \in A)$.

Démonstration : en cours.

3.4 Estimation sans biais

3.4.1 Ordre sur les estimateurs

Soit $g : \Theta \subset \mathbb{R}^p \rightarrow \mathbb{R}$

$$\theta \rightsquigarrow g(\theta)$$

et T un estimateur de $g(\theta)$.

Définition 12 (Fonction perte) On définit la fonction perte quadratique par :

$$L(T(x), \theta) = \|T(x) - g(\theta)\|^2$$

Définition 13 (Fonction risque) On appelle fonction risque associée à l'estimateur T de $g(\theta)$ relativement à la fonction perte L , la fonction :

$$R : \Theta \rightarrow \mathbb{R}$$

$$\theta \rightsquigarrow R(T, \theta) = E_{\theta}[L(T(X), \theta)]$$

Par exemple, la fonction risque quadratique est donnée par :

$$R(T, \theta) = E_{\theta}((T(X) - g(\theta))^2)$$

Remarques : i/ La fonction risque est définie si T admet un moment d'ordre deux.

ii/ Si T est un estimateur sans biais de $g(\theta)$, $E(T(X)) = g(\theta)$, on a :

$$R(T, \theta) = E((T(X) - E(T(X)))^2) = \text{var}(T(X)).$$

Définition 14 Un estimateur T est dit meilleur qu'un autre estimateur T' relativement à la fonction risque R si $R(T, \theta) \leq R(T', \theta) \forall \theta \in \Theta$.

Si T et T' sont des ESB de $g(\theta)$. Alors : $\text{var}T \leq \text{var}T' \forall \theta \in \Theta$.

Remarque : Si $R(T, \theta) \leq R(T', \theta) \Leftrightarrow R(T, \theta) = \min_{T'} R(T', \theta) \forall \theta \in \Theta$ et $\forall T'$ un estimateur de $g(\theta)$.

3.4.2 Estimateur sans biais à variance minimum (E.S.B.V.U.M)

Théorème 15 (Rao-Blackwell) Soit T un ESB de $g(\theta)$ et soit S une statistique exhaustive pour θ . Alors :

i/ $T' = E_\theta(T/S)$ est un ESB de $g(\theta)$ (meilleur que T).

ii/ $\text{var}_\theta(T') \leq \text{var}_\theta(T), \forall \theta \in \Theta$.

Théorème 16 (Lehman-Sheffé) Soit T un ESB de $g(\theta)$ et soit S une statistique exhaustive et complète pour θ . Alors :

$E_\theta(T/S)$ est l'unique E.S.B.V.U.M de $g(\theta)$.

Remarque : L'estimateur est nécessairement une fonction de S .

Exemples : 1°. $X \rightsquigarrow N(m, \sigma^2)$: Donner l'ESBVUM de m .

2°. $X \rightsquigarrow U_{[0, \theta]}$: Donner l'ESBVUM de θ .

3.4.3 Estimateurs efficaces

Théorème 17 (Inégalité de Fréchet-Darmois-Cramer-Rao) Sous les hypothèses de régularité pour la définition de l'information de Fisher : i/ hyp de Fisher.

ii/ $E(T^2) \exists \forall T$.

iii/ $\frac{\partial}{\partial \theta} \int_{\mathcal{X}} T(x) f(\theta, x) dx = \int_{\mathcal{X}} \frac{\partial}{\partial \theta} T(x) f(\theta, x) dx$ et $\int_{\mathcal{X}} \left| T(x) \frac{\partial}{\partial \theta} f(\theta, x) \right| < +\infty$.

On a :

i/ g est dérivable sur Θ .

ii/ $\forall T$ un ESB de $g(\theta)$, $\text{var} T \geq \frac{(g'(\theta))^2}{I_n(\theta)}, \forall \theta \in \Theta$.

Remarque : Parmi toute les conditions citées précédemment, on ne vérifiera que la plus importante, à savoir :

le support de $f(\theta, x)$ est indépendant de θ .

Définition 18 $\frac{(g'(\theta))^2}{I_n(\theta)} = B.C.R$ "Borne de Cramer-Rao".

Définition 19 T ESB de $g(\theta)$ est dit efficace ssi : $\text{var}(T) = BCR$.

Théorème 20 T estimateur efficace de $g(\theta)$ ssi :

$$\frac{\partial}{\partial \theta} \log \mathcal{L}(\theta, \underline{x}) = k(\theta) [T(x) - g(\theta)]$$

Théorème 21 La BCR ne peut être atteinte que si la loi de X est de la forme exponentielle. Sous cette condition, il n'existe qu'une seule fonction $g(\theta)$ qui puissent être estimée efficacement

$g(\theta) = -\frac{\beta'(\theta)}{\alpha'(\theta)}$. L'estimateur de $g(\theta)$ est alors $\frac{1}{n} \sum_{i=1}^n a(x_i)$ et la variance minimale

est telle que : $V(T) = \frac{g'(\theta)}{n\alpha'(\theta)}$.
 $f(x, \theta) = c(\theta) \exp[a(x)\alpha(\theta)] h(x)$
 $\log f(x, \theta) = a(x)\alpha(\theta) + \log(c(\theta)) + \log(h(x))$
 $= a(x)\alpha(\theta) + \beta(\theta) + b(x)$

Exemple : n ech de $X \sim N(m, \sigma^2)$ avec σ^2 connu.

3.5 Cas vectoriel

• Pour les mêmes raisons dans le cas où θ est réel, on ne peut trouver un estimateur optimal pour $g(\theta)$. On se limitera alors aux estimateurs sans biais de $g(\theta)$. Càd : $E(T_i(x)) = g_i(\theta), \forall i = 1, \dots, p$. Au quelle cas : $R(T, \theta) = \sum_{i=1}^p \text{var}(T_i)$.

• Un estimateur T sera dit meilleur qu'un autre estimateur T' si :

$$R(T, \theta) \leq R(T', \theta) \Leftrightarrow \sum_{i=1}^p \text{var}(T_i) = \sum_{i=1}^p \text{var}(T'_i)$$

$$\Leftrightarrow \text{tr } v^T \leq \text{tr } v'^T, \forall \theta \in \Theta, \text{ où } v^T \text{ est la matrice de var-cov de } T.$$

Remarques : i/ Pour avoir les relations précédentes, il est nécessaire que la matrice $v'^T - v^T$ soit symétrique définie positive.

ii/ T est meilleur que T' ssi : $v'^T - v^T$ est semi définie positive.

• **E.S.B.V.U.M :**

i/ $T' = E_\theta(T/S)$ meilleur que $T : v^{E_\theta(T/S)} \leq v^T \Leftrightarrow v^T - v^{E_\theta(T/S)}$ est semi définie positive, $\forall \theta \in \Theta$.

ii/ T ESB de $g(\theta) \Leftrightarrow E(T_i) = g_i(\theta)$. Dans ce cas : $B.C.R. = J.I_n^{-1}(\theta)J'$, où J est la matrice Jacobienne de $g : J = \left(\frac{\partial}{\partial \theta_j} g_i(\theta) \right)_{1 \leq i \leq p, 1 \leq j \leq k}$.

• **Estimateur efficace :**

i/ $\forall T$ ESB de $g(\theta) : v^T > J.I_n^{-1}(\theta)J'$.

ii/ T efficace $\Leftrightarrow v^T = J.I_n^{-1}(\theta)J'$.

• **Cas particulier :**

Si $g(\theta) = \theta \Rightarrow J = I$. Dans ce cas : $B.C.R. = I_n^{-1}(\theta) = n^{-1}I^{-1}(\theta)$.

Exemple : n ech de $X \rightsquigarrow N(m, \sigma^2)$ avec m, σ^2 inconnus. Trouver l'E.S.B.V.U.M de (m, σ^2) .

• **Efficacité :**

Définition 22 On appelle efficacité d'un ESB de $g(\theta)$ le nombre noté e_T défini par :

$$e_T = \frac{BCR}{\text{var}(T)}.$$

Remarque : $0 < e_T \leq 1$.

Définition 23 Soit T et T' deux ESB de $g(\theta)$. On appelle efficacité de T par rapport à T' $e_{T/T'}$ la quantité :

$$e_{T/T'} = \frac{\text{var} T'}{\text{var} T} = \frac{e_T}{e_{T'}}.$$

Remarque : Si $e_{T/T'} < 1 \Leftrightarrow T'$ meilleur que T .