Chapitre 2

Estimation ponctuelle

1- Généralités:

On cherche à déduire à partir des caractéristiques d'un échantillon, les caractéristiques de la population dont il est issu. L'estimation ponctuelle consiste à donner des valeurs approximatives aux paramètres d'une population à l'aide d'un échantillon de n observations. On peut se tromper sur les valeurs exactes à estimer mais on donne les meilleures valeurs possibles que l'on peut supposer.

Soit $(\mathbb{E}, \mathcal{E}, \{\mathbb{P}_\theta, \theta \in \Theta \subseteq \mathbb{R}^k\})$ un modèle et $\tilde{X} = (X_1, \ldots, X_n)$ un échantillon induit par X de X.

Définition:

On appelle estimateur de $g(\theta)$ toute statistique $T_n = T(\tilde{X})$ (mesurable) défini de $\mathbb{E}^n \longrightarrow g(\Theta)$.

Une réalisation de T_n est appelée estimation et elle est notée par $t_n = T(\tilde{x})$: c'est à dire la valeur prise par la statistique T_n au point $\tilde{x} = (x_1, \ldots, x_n)$.

2- Méthodes d'estimation

2-1- Méthode des moments (MM)

Soit $r \in \mathbb{N}^*$; notons par :

$$\mu_r(\theta) = \mathbb{E}(X^r)$$

moment non centré ~~d'ordre~~ théorique d'ordre r

$$m_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$$

moment non centré empirique d'ordre r

Définition :

Si $\forall \theta \in \Theta$, $\mu_k(\theta) < +\infty$ et que

$$\begin{cases} \mu_1(\theta) = m_1 \\ \mu_2(\theta) = m_2 \\ \vdots \\ \mu_k(\theta) = m_k \end{cases}$$

admet une unique solution, celle-ci est appelée l'estimateur de θ par la MM.

Exemple :

1) Modèle de Poisson $\{\mathcal{P}(\lambda), \lambda > 0\}$

Résoudre $\mathbb{E}(X) = \overline{X} \Rightarrow \lambda = \overline{X}$

$\Rightarrow \overline{X}$ est l'estimateur de λ par la MM.

2) Modèle Uniforme $\{\mathcal{U}([0,\theta]), \theta > 0\}$

On a : $\mathbb{E}(X) = \overline{X} \Rightarrow \dfrac{\theta}{2} = \overline{X} \Rightarrow \theta = 2\overline{X}$

Donc $\hat{\theta} = 2\overline{X}$ est l'estimateur de θ par la MM.

③ Modèle gamma $\{ \gamma(\lambda, k) \mid k>0, \lambda>0 \}$

Résolvons $\begin{cases} \mathbb{E}(X) = \overline{X} \\ \mathbb{E}(X^2) = \overline{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \end{cases}$

$$\Rightarrow \begin{cases} \mathbb{E}(X) = \overline{X} = \frac{k}{\lambda} \\ \mathbb{E}(X^2) = \frac{k}{\lambda^2} + \frac{k^2}{\lambda^2} \end{cases} \left(\mathbb{E}(X^2) = \mathrm{Var}(X) + [\mathbb{E}(X)]^2 \right)$$

$$\Rightarrow \begin{cases} \overline{X} = \frac{k}{\lambda} \\ \overline{X^2} = \frac{k}{\lambda^2} + \frac{k^2}{\lambda^2} \end{cases} \Rightarrow \begin{cases} k = \lambda \overline{X} \\ \overline{X^2} = \frac{\overline{X}}{\lambda} + (\overline{X})^2 \end{cases}$$

$$\Rightarrow \begin{cases} k = \lambda \overline{X} \\ \frac{\overline{X}}{\lambda} = \overline{X^2} - \overline{X}^2 = S^2 \end{cases} \Rightarrow \begin{cases} k = \lambda \overline{X} \\ \lambda = \frac{\overline{X}}{S^2} \end{cases}$$

$$\Rightarrow \begin{cases} k = \frac{(\overline{X})^2}{S^2} \\ \lambda = \frac{\overline{X}}{S^2} \end{cases} \Rightarrow \cancel{\hat{\theta} = (\hat{\lambda}, \hat{k}) = \left(\frac{\overline{X}}{S^2}, \frac{\overline{X}^2}{S^2}\right)}$$

$$\Rightarrow \hat{\theta} = (\hat{k}, \hat{\lambda}) = \left(\frac{\overline{X}^2}{S^2}, \frac{\overline{X}}{S^2} \right)$$

Attention : $\overline{X}^2 = (\overline{X})^2 \neq \overline{X^2} = \frac{1}{n}\sum X_i^2$

$\hat{\theta}$ est l'estimateur de θ par la MM.

2-2- Méthode du maximum de vraisemblance (MMV)

Définition :

On appelle estimateur du paramètre θ obtenu par la MMV la ~~fonction~~ $\cancel{\theta}$ statistique $\hat{\theta}$ qui maximise la fonction de vraisemblance c'est à dire

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \mathbb{L}(\tilde{X};\theta)$$

On utilise souvent $\hat{\theta}$ qui maximise la fonction $\log[\mathbb{L}(\tilde{X},\theta)]$ appelée log-vraisemblance.

Exemple :

Considerons le modèle uniforme $(\mathcal{U}([0,\theta]), \theta>0)$

$$\mathbb{L}(\tilde{x},\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta}, \; x_i \in [0,\theta]$$

$$\Rightarrow \mathbb{L}(\tilde{x},\theta) = \frac{1}{\theta^n} \; ; \; x_i \in [0,\theta]$$

Comme la vraisemblance est une fonction décroissante en θ et que toutes les v.a X_i sont dans $[0,\theta]$ alors la valeur qui maximise la vraisemblance est

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

Propriété

Si le support de la loi de X ne dépend pas de θ et que le score est dérivable par rapport à θ alors l'estimateur $\hat{\theta}$ de θ obtenu par la MMV est solution du système suivant :

$$\begin{cases} S(\tilde{X};\theta) = 0 \\ \frac{\partial}{\partial\theta} S(\tilde{X};\theta)\Big/_{\hat{\theta}} < 0 \end{cases}$$

Exemple:

$$X \rightsquigarrow \mathcal{E}(1/\theta)$$

$$\Rightarrow f(x;\theta) = \frac{1}{\theta} e^{-\frac{1}{\theta}x}, \quad x \geq 0 \text{ (ne dépend pas de } \theta)$$

$$* \; L(\tilde{x},\theta) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum x_i}$$

$$\Rightarrow \log[L(\tilde{x},\theta)] = -n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

$$* \; S(\tilde{x},\theta) = \frac{\partial}{\partial\theta}\log[L(\tilde{x},\theta)] = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i$$

$$S(\tilde{x},\theta) = 0 \Leftrightarrow -n + \frac{1}{\theta}\sum_{i=1}^{n} x_i = 0$$

$$\Leftrightarrow \theta = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Donc $\hat{\theta} = \overline{X}$ est candidat.

* Vérifions la seconde condition

$$\frac{\partial}{\partial\theta} S(\tilde{x},\theta) = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} x_i$$

$$\frac{\partial}{\partial\theta} S(\tilde{x},\theta) \Big/ \hat{\theta} = \frac{n}{\hat{\theta}^2} - \frac{2}{\hat{\theta}^3}\sum x_i = \frac{n}{\hat{\theta}^2} - \frac{2n\hat{\theta}}{\hat{\theta}^3}$$

$$\Rightarrow \frac{\partial}{\partial\theta} S(\tilde{x},\theta) \Big/ \hat{\theta} = \frac{n}{\hat{\theta}^2} - \frac{2n}{\hat{\theta}^2} = -\frac{n}{\hat{\theta}^2} < 0$$

D'où $\hat{\theta} = \bar{X}$ est l'EMV

Propriété [d'invariance du EMVS]

Soit $\hat{\theta}$ l'EMV de θ. Soit g une fonction mesu-rable de $\Theta \longrightarrow g(\Theta) = \Theta'$ alors l'EMV de $g(\theta)$ $\{\mathbb{P}_{g(\theta)}, g(\theta) \in \Theta\}$ est donnée par $g(\hat{\theta}) = \widehat{g(\theta)}$

On appelle cette propriété aussi lemme de Zehna.

3- Qualité d'un estimateur :

3-1- Biais :

Soit T_n estimateur $g(\theta)$. On appelle biais de l'esti-mateur

$$b(T_n; g(\theta)) = E(T_n) - g(\theta)$$

* Si $b(T_n; g(\theta)) = 0$ alors on dit l'estimateur est sans biais sinon elle est dite biaisé.

* Si $b(T_n; g(\theta)) \xrightarrow[n \to +\infty]{} 0$ alors l'estimateur

est asymptotiquement sans biais.

Exemple:

* La moyenne empirique $\overline{X}$ est un estimateur sans biais de la moyenne théorique μ.
* La variance empirique S^2 est un estimateur biaisé de la variance théorique mais asymptotiquement sans biais.

Remarque:

Le biais mesure l'erreur d'estimation lorsqu'il est négatif, l'estimateur sous estime le paramètre lorsqu'il est positif, on fait une sur-estimation.

3-2- Convergence

Définition

L'estimateur T_n de $g(\theta)$ est dit convergente si T_n converge en probabilité vers $g(\theta)$ c'est à dire

$$\lim_{n \to +\infty} \mathbb{P}\left(|T_n - g(\theta)| \geq \varepsilon\right) = 0$$

Remarque: [condition suffisante]

Si la variance d'un estimateur asymptotiquement sans biais tend vers 0 alors l'estimateur est consistant.

$$\mathbb{P}(|T_n - g(\theta)| \geq \varepsilon) \leq \frac{\mathbb{E}[(T_n - g(\theta))^2]}{\varepsilon^2}$$

$$\Rightarrow \mathbb{P}(|T_n - g(\theta)| \geq \varepsilon) \leq \frac{\mathbb{V}ar(T_n) + b^2(T_n ; g(\theta))}{\varepsilon^2}$$

3-3- Risque quadratique moyen

Définition :

Le risque quadratique moyen ou "Mean square Error MSE" mesure la précision de l'estimation. Elle est définie par :

$$\mathbb{R}(T_n ; g(\theta)) = \mathbb{E}[(T_n - g(\theta))^2]$$

Remarque :

$$\mathbb{R}(T_n ; g(\theta)) = \mathbb{V}ar(T_n) + b^2(T_n ; g(\theta))$$

Définition

Soient T_n et T_n' deux estimateurs de $g(\theta)$. On dit que T_n est préferable à T_n' si :

$$\mathbb{R}(T_n ; g(\theta)) \leq \mathbb{R}(T_n' ; g(\theta))$$

4- Estimateur optimal

Définition :

On dit que l'estimateur T_n de $g(\theta)$ est optimal s'il est préferable à tous les estimateurs de $g(\theta)$.

4-1- Information de Fischer :

Définition :

Soit $\{\mathbb{P}_\theta ; \theta \in \Theta\}$ où Θ est un ouvert un modèle paramètrique. Il est dit régulier s'il vérifie les conditions suivantes :

(H_1) le support des lois est indépendant de θ

$$\mathbb{P}_\theta \left(\Delta = \{x \in E \,/\, f_\theta(x) > 0\} \right)$$

(H_2) $\forall (x;\theta) \in E \times \Theta$, $\frac{\partial}{\partial \theta} f(x,\theta)$ et $\frac{\partial^2}{\partial \theta^2} f(x,\theta)$ existent.

(H_3) $\forall A \in \mathcal{E}$, on a :

$$\frac{\partial}{\partial \theta} \int_A f(x;\theta)\,dx = \int_A \frac{\partial}{\partial \theta} f(x;\theta)\,dx < +\infty$$

$$\frac{\partial^2}{\partial \theta^2} \int_A f(x;\theta)\,dx = \int_A \frac{\partial^2}{\partial \theta^2} f(x;\theta)\,dx < +\infty$$

Définition :

On appelle information de fischer $I_X(\theta)$ apportée par la variable aléatoire X sur le paramètre θ, la variance de $S(X;\theta)$ (si elle existe). On a :

$$I_X(\theta) = I(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial \theta} \log [f(X,\theta)] \right)^2 \right]$$

Exemple :

Soit $X \rightsquigarrow \mathcal{E}(\theta)$; donc $f(x,\theta) = \theta e^{-\theta x}$, $x \geq 0$

* $\log[f(X;\theta)] = \log(\theta) - \theta X$

* $\frac{\partial}{\partial\theta} \log[f(X;\theta)] = \frac{1}{\theta} - X$

* $\left(\frac{\partial}{\partial\theta} \log[f(X;\theta)]\right)^2 = \left(\frac{1}{\theta} - X\right)^2 = \frac{1}{\theta^2} - \frac{2}{\theta}X + X^2$

* $\mathbb{I}_X(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta} \log[f(X;\theta)]\right)^2\right] = \mathbb{E}\left(\frac{1}{\theta^2} - \frac{2}{\theta}X + X^2\right)$

$$= \frac{1}{\theta^2} + \mathbb{E}(X^2) - \frac{2}{\theta}\mathbb{E}(X) = \frac{1}{\theta^2} + \mathbb{V}ar(X) + \mathbb{E}(X)^2 - \frac{2}{\theta}\mathbb{E}(X)$$

$$\mathbb{I}_X(\theta) = \frac{1}{\theta^2} + \frac{1}{\theta^2} + \frac{1}{\theta^2} - \frac{2}{\theta} \times \frac{1}{\theta} = \frac{1}{\theta^2}$$

$$\Rightarrow \boxed{\mathbb{I}_X(\theta) = \frac{1}{\theta^2}}$$

Théorème :

Sous les conditions de régularité, l'information de fischer rapportée par X sur le paramètre θ est donnée par :

$$\mathbb{I}_X(\theta) = -\mathbb{E}\left[\frac{\partial}{\partial\theta} S(X;\theta)\right] = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2} \log[f(X;\theta)]\right]$$

$$I_X(\theta) = -\mathbb{E}\left[\frac{\partial}{\partial\theta} S(X;\theta)\right] = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left[f(X;\theta)\right]\right]$$

<u>Preuve :</u>

On a : $S(X;\theta) = \frac{\partial}{\partial\theta}\log\left[f(X;\theta)\right] = \frac{f'(X;\theta)}{f(X;\theta)}$

$$\frac{\partial}{\partial\theta} S(X;\theta) = \frac{f''(X;\theta)f(X;\theta) - \left[f'(X;\theta)\right]^2}{\left[f(X;\theta)\right]^2}$$

$$= \frac{f''(X;\theta)}{f(X;\theta)} - \left[\frac{f'(X;\theta)}{f(X;\theta)}\right]^2$$

$$= \frac{f''(X;\theta)}{f(X;\theta)} - \left(\frac{\partial}{\partial\theta}\log\left[f(X;\theta)\right]\right)^2$$

$$\Rightarrow \frac{\partial}{\partial\theta} S(X;\theta) = \frac{f''(X;\theta)}{f(X;\theta)} - \left[S(X;\theta)\right]^2$$

$$\Rightarrow \mathbb{E}\left(\frac{\partial}{\partial\theta} S(X;\theta)\right) = \underbrace{\mathbb{E}\left[\frac{f''(X;\theta)}{f(X;\theta)}\right]}_{\Delta} - \mathbb{E}\left[\left(S(X;\theta)\right)^2\right]$$

Montrons que $\Delta = 0$

$$\mathbb{E}\left[\frac{f''(X;\theta)}{f(X;\theta)}\right] = \int_E \frac{f''(x;\theta)}{f(x;\theta)} \cdot f(x,\theta)\,dx$$

$$= \int_E f''(x;\theta)\,dx = \int_E \frac{\partial^2}{\partial\theta^2} f(x,\theta)\,dx$$

$$\Rightarrow \mathbb{E}\left[\frac{f''(X;\theta)}{f(X;\theta)}\right] = \frac{\partial^2}{\partial\theta^2}\int_E f(x;\theta)\,dx \quad (\text{sous } H_3)$$

or par définition $\int_{E} f(x;\theta)\,dx = 1$

$$\Rightarrow \mathbb{E}\left[\frac{f''(X;\theta)}{f(X;\theta)}\right] = \frac{\partial^2}{\partial\theta^2}(1) = 0$$

$$\Rightarrow \boxed{\mathbb{E}\left([S(X;\theta)]^2\right) = -\mathbb{E}\left(\frac{\partial}{\partial\theta}S(X;\theta)\right)}$$

<u>Exemple:</u>

$X \rightsquigarrow \mathcal{E}(\theta)$; $f(x,\theta) = \theta e^{-\theta x}$, $x \geq 0$

$$\mathbb{I}_X(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left[f(X;\theta)\right]\right]$$

$$= -\mathbb{E}\left(\frac{\partial}{\partial\theta}\left(\frac{1}{\theta}-X\right)\right] = -\mathbb{E}\left(-\frac{1}{\theta^2}\right)$$

$$\Rightarrow \boxed{\mathbb{I}_X(\theta) = -\left(-\frac{1}{\theta^2}\right) = \frac{1}{\theta^2}}$$

<u>Définition:</u>

On appelle quantité d'information apportée par l'échan-tillon $\tilde{X} = (X_1, \ldots, X_n)$ sur le paramètre θ, la variance (si elle existe) de $S(\tilde{X};\theta)$.

$$\mathbb{I}_{\tilde{X}}(\theta) = \mathbb{I}_n(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\log\left[L(\tilde{X};\theta)\right]\right)^2\right]$$

<u>Théorème:</u>

Si le modèle est régulier alors :

$$\mathbb{I}_{\underset{\sim}{X}}(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left[\mathbb{L}(\underset{\sim}{X};\theta)\right]\right]$$

__Exemple:__

Modèle exponentielle de paramètre θ $\{\mathcal{E}(\theta), \theta > 0\}$.

Si $X \rightsquigarrow \mathcal{E}(\theta)$; $f(x,\theta) = \theta e^{-\theta x}$, $x \geq 0$

Le modèle ast régulier alors

$$\mathbb{I}_{\underset{\sim}{X}}(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta^2}\log\left[\mathbb{L}(\underset{\sim}{X};\theta)\right]\right]$$

avec $\mathbb{L}(\tilde{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n}\left(\theta e^{-\theta x_i}\right)$

$\Rightarrow \log\left[\mathbb{L}(\tilde{x};\theta)\right] = \sum_{i=1}^{n}\log\left(\theta e^{-\theta x_i}\right) = \sum_{i=1}^{n}\left(\log(\theta) - \theta x_i\right)$

$\Rightarrow \log\left[\mathbb{L}(\tilde{x};\theta)\right] = n\log(\theta) - \theta\sum_{i=1}^{n} x_i$

$\Rightarrow \frac{\partial}{\partial\theta}\log\left[\mathbb{L}(\tilde{x};\theta)\right] = \frac{n}{\theta} - \sum_{i=1}^{n} x_i$

$\Rightarrow \frac{\partial^2}{\partial\theta^2}\log\left[\mathbb{L}(\tilde{x};\theta)\right] = -\frac{n}{\theta^2}$

$$\Rightarrow \boxed{\mathbb{I}_{\underset{\sim}{X}}(\theta) = -\mathbb{E}\left(\frac{-n}{\theta^2}\right) = \frac{n}{\theta^2}}$$

__Remarque:__

Si le modèle ast régulier alors $\mathbb{I}_{\underset{\sim}{X}}(\theta) = n\,\mathbb{I}_{X}(\theta)$.

On parle d'additivité de l'information.

Preuve: (à faire)

Propriétés:

① Pour toute statistique T_n on a : $I_{T_n}(\theta) \leq I_n(\theta)$

avec $I_{T_n}(\theta) = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta} \log\left[f_{T_n}(T_n;\theta)\right]\right)^2\right]$

② Si le support est indépendant de θ et T_n un esti-mateur de $g(\theta)$ alors si $I_{T_n}(\theta) = I_n(\theta)$ alors T_n est une statistique exhaustive.

4-2- Borne de Cramer-Rao (CR)

ou Frechet-Cramer-Darmois-Rao (FCDR)

Théorème:

Considerons un modèle régulier et soit $T_n = T(\tilde{X})$ un estimateur sans biais de $g(\theta) \in \mathbb{R}$; $0 < T_n < +\infty$ et

$$\frac{\partial}{\partial\theta}\int_{E^n} T(\tilde{x})L(\tilde{x};\theta)d\tilde{x} = \int_{E^n} T(\tilde{x})\frac{\partial}{\partial\theta}L(\tilde{x};\theta)d\tilde{x} < +\infty$$

alors g est dérivable et nous avons l'inégalité suivante :

$$Var(T_n) \geq \frac{[g'(\theta)]^2}{I_n(\theta)}$$

Preuve:

$$\mathbb{E}(T_n) = \int_{E^n} T(\tilde{x})L(\tilde{x};\theta)\,d\tilde{x} = g(\theta)$$

$$\Rightarrow g'(\theta) = \int_{\mathbb{E}^n} \mathbb{T}(\tilde{x}) \frac{\partial}{\partial\theta} \mathbb{L}(\tilde{x};\theta)\, d\tilde{x}$$

$$= \int_{\mathbb{E}^n} \mathbb{T}(\tilde{x}) \frac{\mathbb{L}'(\tilde{x};\theta)}{\mathbb{L}(\tilde{x};\theta)} \cdot \mathbb{L}(\tilde{x};\theta)\, d\tilde{x}$$

$$= \int_{\mathbb{E}^n} \mathbb{T}(\tilde{x}) \frac{\partial}{\partial\theta} \log\left[\mathbb{L}(\tilde{x};\theta)\right] \cdot \mathbb{L}(\tilde{x};\theta)\, d\tilde{x}$$

$$\Rightarrow g'(\theta) = \int_{\mathbb{E}^n} \mathbb{T}(\tilde{x}) \cdot S(\tilde{x};\theta) \cdot \mathbb{L}(\tilde{x};\theta)\, d\tilde{x}$$

donc

$$g'(\theta) = \mathbb{E}\left[\mathbb{T}(\tilde{X}) \cdot S(\tilde{X};\theta)\right]$$

or

$$\operatorname{cov}\left[\mathbb{T}(\tilde{X}), S(\tilde{X};\theta)\right] = \mathbb{E}\left(\mathbb{T}(\tilde{X}) \cdot S(\tilde{X};\theta)\right] - \mathbb{E}(\mathbb{T}(\tilde{X})) \cdot \underbrace{\mathbb{E}(S(\tilde{X};\theta))}_{0}$$

$$\Rightarrow \operatorname{cov}\left[\mathbb{T}(\tilde{X}), S(\tilde{X};\theta)\right] = \mathbb{E}\left[\mathbb{T}(\tilde{X}) \cdot S(\tilde{X};\theta)\right]$$

$$\Rightarrow \operatorname{cov}\left[\mathbb{T}(\tilde{X}), S(\tilde{X};\theta)\right] = g'(\theta)$$

D'après l'inégalité de Cauchy-Schwartz

$$\left[\operatorname{cov}\left(\mathbb{T}(\tilde{X}); S(\tilde{X};\theta)\right)\right]^2 \leq \operatorname{Var}(\mathbb{T}_n) \cdot \operatorname{Var}\left[S(\tilde{X};\theta)\right]$$

$$\Rightarrow \left[g'(\theta)\right]^2 \leq \operatorname{Var}(\mathbb{T}_n) \cdot \mathbb{I}_n(\theta)$$

$$\Rightarrow \boxed{Var(T_n) \geq \frac{[g'(\theta)]^2}{I_n(\theta)}}$$

Définition :

$$B_{inf}(\theta) = \frac{[g'(\theta)]^2}{I_n(\theta)}$$

est appelé borne de Cramer - Rao.

Un estimateur atteignant cette borne est dit efficace.

Autrement dit, on ne peut pas trouver un autre estimateur sans biais de variance faible. Un estimateur efficace est le meilleur estimateur.

Exemple :

Soit $X_1, \ldots, X_n \xrightarrow{iid} \mathcal{E}(1/\theta)$

* Déterminer l'EMV de θ.
* Etudier : biais, consistance, efficacité de $\hat{\theta}$

Solution :

* Déterminons l'EMV de θ :

$$L(\tilde{x};\theta) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum x_i} \Rightarrow \log[L(\tilde{x};\theta)] = -n\log(\theta) - \frac{1}{\theta}\sum x_i$$

$$\frac{\partial}{\partial\theta} \log[L(\tilde{x},\theta)] = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum x_i = 0 \Rightarrow \theta = \bar{x}$$

$$\Rightarrow \hat{\theta} = \bar{X} \text{ est candidat.}$$

$$\frac{\partial^2}{\partial\theta^2} \log[L(\tilde{x};\theta)] = \frac{n}{\theta^2} - \frac{2}{\theta^3}\sum x_i$$

$$\Rightarrow \frac{\partial^2}{\partial\theta^2}\log[L(\tilde{x};\theta)] \Big/ \hat{\theta} = -\frac{n}{\hat{\theta}^2} < 0$$

$$\Rightarrow \boxed{\hat{\theta} = \bar{X} \text{ est l'EMV de } \theta}$$

* Etude des propriétés de $\hat{\theta}$

- biais :

$$b(\hat{\theta};\theta) = E(\hat{\theta}) - \theta = E(\bar{X}) - \theta = E(X) - \theta$$

$$\Rightarrow b(\hat{\theta};\theta) = \theta - \theta = 0 \Rightarrow \boxed{\hat{\theta} \text{ est sans biais}}$$

- consistance

$$Var(\hat{\theta}) = Var(\bar{X}) = \frac{Var(X)}{n} = \frac{\theta^2}{n}$$

$$\begin{cases} b(\hat{\theta};\theta) = 0 \\ Var(\hat{\theta}) \xrightarrow[n\to+\infty]{} 0 \end{cases} \Rightarrow \boxed{\hat{\theta} \text{ est consistant}}$$

- efficacité

$$I_n(\theta) = -E\left[\frac{\partial^2}{\partial\theta^2}\log[L(\tilde{X};\theta)]\right]$$

$$= -E\left[\frac{n}{\theta^2} - \frac{2}{\theta^3}\sum X_i\right]$$

$$= -\frac{n}{\theta^2} + \frac{2}{\theta^3}\sum_{i=1}^{n} E(X) = -\frac{n}{\theta^2} + \frac{2n\theta}{\theta^3}$$

$$\Rightarrow I_n(\theta) = \frac{n}{\theta^2} \text{ or } B_{inf}(\theta) = \frac{[g'(\theta)]^2}{I_n(\theta)} = \frac{1}{n/\theta^2} = \frac{\theta^2}{n}$$

$$\Rightarrow B_{inf}(\theta) = Var(\hat{\theta}) \Rightarrow \boxed{\hat{\theta} \text{ est efficace}}$$

$$g(\theta) = \theta \Rightarrow g'(\theta) = 1$$

Exercice : TAF.

$X_1, \ldots, X_n$ extrait de X de $f(x;\theta) = \theta x^{-(\theta+1)}$, $x \geq 1$

où $\theta > 0$.

1) Déterminer l'EMV : $\hat{\theta}$ de θ.

2) Déduire un estimateur sans biais de θ et calculer sa variance.

3) $g :]0;+\infty[\longrightarrow \mathbb{R}_+^*$

$\theta \longmapsto g(\theta) = 1/\theta$

Montrer que $T_n = g(\hat{\theta})$ est un estimateur sans biais et efficace de $g(\theta)$

Solution :

① Déterminons l'EMV : $\hat{\theta}$ de θ :

On a : $L(\tilde{x};\theta) = \prod_{i=1}^{n} f(x_i;\theta)$

$$\Rightarrow \log[L(\tilde{x};\theta)] = \sum_{i=1}^{n} \log[f(x_i;\theta)]$$

$$= \sum_{i=1}^{n} \log(\theta . x^{-(\theta+1)})$$

$$= \sum_{i=1}^{n} \log(\theta) - (\theta+1) \sum_{i=1}^{n} \log(x_i)$$

$$\Rightarrow \log[L(\tilde{x};\theta)] = n\log(\theta) - (\theta+1)\sum \log(x_i)$$

$$S(\tilde{x};\theta) = \frac{\partial}{\partial\theta} \log[L(\tilde{x};\theta)] = \frac{n}{\theta} - \sum_{i=1}^{n} \log(x_i)$$

$$\Rightarrow S(\tilde{x};\theta) \Rightarrow \theta = n / \sum \log(x_i)$$

$\Rightarrow \hat{\theta} = \dfrac{1}{\frac{1}{n}\sum \log(X_i)}$ est candidat

$\Rightarrow \dfrac{\partial}{\partial\theta} S(\vec{x};\theta) = -\dfrac{n}{\theta^2} \Rightarrow \dfrac{\partial}{\partial\theta} S(\vec{x};\theta)\Big|_{\hat{\theta}} = -\dfrac{n}{\hat{\theta}^2} < 0$

donc $\hat{\theta}$ est un EMV de θ.

② En déduisons un estimateur sans biais de θ

a) cherchons la loi de $\log(X_i)$

posons $Y_i = \log(X_i)$

$\mathbb{P}(Y_i \leq y) = \mathbb{P}(\log(X_i) \leq y) \; ; \forall y \geq 0$

$= \mathbb{P}[X_i \leq e^y]$

$\Rightarrow F_{Y_i}(y) = \mathbb{P}(Y_i \leq y) = F_{X_i}(e^y)$

$\Rightarrow f_{Y_i}(y) = e^y f_{X_i}(e^y)$

$= \theta e^y . e^{-y(\theta+1)}$

$f_{Y_i}(y) = \theta e^{-\theta y} \; ; \forall y \geq 0$

$\Rightarrow Y_i = \log(X_i) \rightsquigarrow \mathcal{E}(\theta)$.

b) Cherchons la loi de $\bar{Y} = \frac{1}{n}\sum \log(X_i) = \frac{1}{n}\sum Y_i$

$Y_i \rightsquigarrow \mathcal{E}(\theta) = \gamma(1;\theta)$

$\Rightarrow \sum Y_i \rightsquigarrow \gamma(n;\theta)$

$$\Rightarrow \overline{Y} \rightsquigarrow \gamma(n; n\theta)$$

$$\Rightarrow f_{\overline{Y}}(y) = \frac{(n\theta)^n}{\Gamma(n)} y^{n-1} e^{-n\theta y}, \quad y \geq 0$$

$$\Rightarrow \mathbb{E}\left(\frac{1}{\overline{Y}}\right) = \int_0^{+\infty} \frac{1}{y} \times \frac{(n\theta)^n}{\Gamma(n)} y^{n-1} e^{-n\theta y} \, dy$$

$$\Rightarrow \mathbb{E}\left(\frac{1}{\overline{Y}}\right) = \int_0^{+\infty} \frac{(n\theta)^n}{\Gamma(n)} y^{n-2} \cdot e^{-n\theta y} \, dy$$

posons $u = n\theta y \Rightarrow y = \frac{u}{n\theta} \Rightarrow dy = \frac{1}{n\theta} du$

$\Rightarrow y > 0 \Rightarrow u > 0$

$$\Rightarrow \mathbb{E}\left(\frac{1}{\overline{Y}}\right) = \int_0^{+\infty} \frac{(n\theta)^n}{\Gamma(n)} \times \frac{1}{(n\theta)} \times (n\theta)^{-n+2} \cdot u^{n-2} e^{-u} \, du$$

$$= \int_0^{+\infty} \frac{(n\theta)^{n-1-n+2}}{\Gamma(n)} u^{n-2} e^{-u} \, du$$

$$= \frac{(n\theta)}{\Gamma(n)} \int_0^{+\infty} u^{(n-1)-1} e^{-u} \, du$$

$$= \frac{(n\theta)}{\Gamma(n)} \times \Gamma(n-1) = \frac{n\theta}{(n-1)!} (n-2)!$$

$$= n\theta \times \frac{1}{(n-1)} \times \frac{1}{(n-2)!} \times (n-2)!$$

$$\boxed{\mathbb{E}\left(\frac{1}{\overline{Y}}\right) = \frac{n}{n-1} \theta}$$

2-20

c) déduisons en un estimateur sans biais de θ.

$$\mathbb{E}(\hat{\theta}) = \frac{n}{n-1}\theta \Rightarrow \frac{n-1}{n}\mathbb{E}(\hat{\theta}) = \theta$$

$$\Rightarrow \mathbb{E}\left(\frac{n-1}{n}\hat{\theta}\right) = \theta$$

d'où $\hat{\theta}' = \frac{n-1}{n}\hat{\theta}$ est un estimateur sans biais de θ.

d) Calculons sa variance

$$Var(\hat{\theta}') = Var\left(\frac{n-1}{n}\hat{\theta}\right) = \frac{(n-1)^2}{n^2} Var(\hat{\theta})$$

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

et $$\mathbb{E}(\hat{\theta}^2) = \int_0^{+\infty} \left(\frac{1}{y}\right)^2 \cdot \frac{(n\theta)^n}{\Gamma(n)} y^{n-1} e^{-n\theta y}\, dy$$

$$= \frac{(n\theta)^n}{\Gamma(n)} \int_0^{+\infty} y^{n-3} \cdot e^{-n\theta y}\, dy$$

$$= \frac{(n\theta)^n}{\Gamma(n)} \cancel{\int_0^{+\infty}} \times \frac{\Gamma(n-2)}{(n\theta)^{n-2}} \underbrace{\int_0^{+\infty} \frac{(n\theta)^{n-2}}{\Gamma(n-2)} y^{n-2-1} e^{-n\theta y}\, dy}_{1}$$

$$\mathbb{E}(\hat{\theta}^2) = \frac{(n\theta)^{n-n+2}}{\Gamma(n)} \times \frac{\Gamma(n-2)}{1}$$

or $\Gamma(n) = (n-1)(n-2)\Gamma(n-2)$

$$\Rightarrow \mathbb{E}(\hat{\theta}^2) = \frac{n^2}{(n-1)(n-2)}\theta^2$$

2-21

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

$$= \frac{(n\theta)^2}{(n-1)(n-2)} - \frac{(n\theta)^2}{(n-1)^2}$$

$$\mathrm{Var}(\hat{\theta}) = \frac{(n\theta)^2}{n-1}\left(\frac{1}{n-2} - \frac{1}{n-1}\right)$$

$$\Rightarrow \mathrm{Var}(\hat{\theta}') = \frac{(n-1)^2}{n^2}\mathrm{Var}(\hat{\theta})$$

$$= \frac{(n-1)^2}{n^2} \times \frac{n^2\theta^2}{n-1}\left(\frac{1}{n-2} - \frac{1}{n-1}\right)$$

$$= (n-1)\theta^2\left(\frac{1}{n-2} - \frac{1}{n-1}\right)$$

$$= \theta^2\left(\frac{n-1}{n-2} - 1\right) = \theta^2\left(\frac{n-1-n+2}{n-2}\right)$$

$$\Rightarrow \boxed{\mathrm{Var}(\hat{\theta}') = \frac{1}{n-2}\theta^2}$$

(3) $g :]0;+\infty[\longrightarrow \mathbb{R}_+^*$

$\theta \longmapsto g(\theta) = \frac{1}{\theta}$

Montrons que $T_n = g(\hat{\theta})$ est un estimateur sans biais et efficace :

$* \; b(T_n, g(\theta)) = \mathbb{E}(T_n) - g(\theta) = \mathbb{E}[g(\hat{\theta})] - g(\theta)$

Calculons $g(\hat{\theta})$.

$\hat{\theta} = \frac{1}{\overline{Y}} \Rightarrow g(\hat{\theta}) = \overline{Y}$ ~~car~~ , $g(\theta) = \frac{1}{\theta}$

$b(T_n, g(\theta)) = \mathbb{E}(\overline{Y}) - \frac{1}{\theta}$ or $\overline{Y} \rightsquigarrow \gamma(n, n\theta)$

$\Rightarrow \mathbb{E}(\overline{Y}) = \frac{n}{n\theta} = \frac{1}{\theta}$

d'où $b(T_n ; g(\theta)) = \frac{1}{\theta} - \frac{1}{\theta} = 0$

donc T_n est un estimateur sans biais de $g(\theta) = 1/\theta$.

* efficacité

$\mathbb{V}ar(T_n) = \mathbb{V}ar(\overline{Y}) = \frac{n}{(n\theta)^2} = \frac{1}{n\theta^2}$

$B_{inf} = \frac{[g'(\theta)]^2}{I_{\underset{\sim}{X}}(\theta)}$ avec $g(\theta) = \frac{1}{\theta} \Rightarrow g'(\theta) = -\frac{1}{\theta^2}$

et $I_{\underset{\sim}{X}}(\theta) = -\mathbb{E}\left[\frac{\partial}{\partial\theta} S(\underset{\sim}{X};\theta)\right] = \frac{n}{\theta^2}$

$\Rightarrow B_{inf} = \frac{1/\theta^4}{n/\theta^2} = \frac{1}{n\theta^2} = \mathbb{V}ar(T_n)$

$\Rightarrow T_n$ est un estimateur efficace.

Remarque :

$$I_{ij}(\theta) = -E\left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log\left[f(X;\theta)\right]\right]$$

avec $1 \leq i, j \leq k$

Pour tout modèle régulier.

Exemple :

$$\{\mathcal{N}(\mu, \sigma^2),\ \theta^t = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*\}$$

induit par X.

① Calculer le vecteur score et vérifier qu'il est centré.

② Calculer l'information de Fischer apportée par X sur θ.

Solution :

① $$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2},\ x \in \mathbb{R}$$

$$\Rightarrow f(x, \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2},\ x \in \mathbb{R}$$

$$\Rightarrow f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \times \frac{(x-m)^2}{\sigma^2}},\ x \in \mathbb{R}$$

$$\log\left[f(x;\mu,\sigma^2)\right] = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2} \times \frac{(x-m)^2}{\sigma^2}$$

$$\forall x \in \mathbb{R}$$

$$\frac{\partial}{\partial \mu} \log[f(x;\mu,\sigma^2)] = \frac{x-\mu}{\sigma^2}$$

$$\frac{\partial}{\partial \sigma^2} \log[f(x;\mu,\sigma^2)] = -\frac{1}{2\sigma^2} + \frac{(x-\mu)^2}{2(\sigma^2)^2}$$

or $S(X;\mu,\sigma^2) = \nabla \log[f(X;\mu,\sigma^2)]$

$$\Rightarrow S(X;\mu,\sigma^2) = \begin{pmatrix} \frac{X-\mu}{\sigma^2} \\ -\frac{1}{2\sigma^2} + \frac{(X-\mu)^2}{2\sigma^4} \end{pmatrix}$$

① Verifions qu'il est centré :

$$\mathbb{E}[S(X;\mu,\sigma^2)] = \begin{pmatrix} \mathbb{E}\left(\frac{X-\mu}{\sigma^2}\right) \\ \mathbb{E}\left(-\frac{1}{2\sigma^2} + \frac{(X-\mu)^2}{2\sigma^4}\right) \end{pmatrix}$$

$$= \begin{pmatrix} 0 \\ -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}\mathbb{E}[(X-\mu)^2] \end{pmatrix}$$

$$\mathbb{E}[S(X;\mu,\sigma^2)] = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$ car $\underbrace{\mathbb{E}[(X-\mu)^2]}_{\mathbb{V}ar(X)} = \sigma^2$

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

② L'information de fischer

• Cherchons I_{11}

$$\left(\frac{\partial}{\partial\mu}\log\left[f(X;\mu,\sigma^2)\right]\right)^2 = \left(\frac{X-\mu}{\sigma^2}\right)^2$$

$$\Rightarrow I_{11} = E\left[\left(\frac{X-\mu}{\sigma^2}\right)^2\right] = \frac{E\left[(X-\mu)^2\right]}{\sigma^4} = \frac{1}{\sigma^2}$$

• Cherchons I_{12}

$$\frac{\partial}{\partial\mu}\log\left[f(X;\mu,\sigma^2)\right]\cdot\frac{\partial}{\partial\sigma^2}\log\left[f(X;\mu,\sigma^2)\right] = A \text{ (On pose)}$$

$$\Rightarrow A = \frac{(X-\mu)}{\sigma^2}\left(-\frac{1}{2\sigma^2}+\frac{(X-\mu)^2}{2\sigma^4}\right)$$

$$A = -\frac{1}{2\sigma^4}(X-\mu)+\frac{1}{2\sigma^6}(X-\mu)^3$$

$$I_{12} = E(A) = -\frac{1}{2\sigma^4}\underbrace{E(X-\mu)}_{0}+\frac{1}{2\sigma^6}\underbrace{E\left[(X-\mu)^3\right]}_{0} = 0$$

car la loi est symétrique

Le moment centré d'ordre k (impair) est égal à 0.

$$\mu_k = \begin{cases} 0 & \text{si } k \text{ impair} \\ \dfrac{(2p)!}{(p!)\,2^p}\sigma^{2p}, & k = 2p. \end{cases}$$

• $I_{12} = I_{21} = 0$

• Cherchons $\mathbb{I}_{22}$

$$\left[\frac{\partial}{\partial\sigma^2}\log f(X;\mu,\sigma^2)\right]^2 = \frac{1}{4\sigma^4} + \frac{(X-\mu)^4}{4\sigma^8} - \frac{(X-\mu)^2}{2\sigma^6}$$

$$\Rightarrow \quad \mathbb{I}_{22} = \frac{1}{4\sigma^4} + \frac{\mu_4}{4\sigma^8} - \frac{\sigma^2}{2\sigma^6}$$

$$\mu_4 = \frac{(2\times 2)!}{2^2\, 2!}\sigma^{2\times 2} = 3\sigma^4$$

$$\Rightarrow \mathbb{I}_{22} = \frac{1}{4\sigma^4} + \frac{3\sigma^4}{4\sigma^8} - \frac{1}{2\sigma^4} = \frac{1}{\sigma^4}\left(\frac{1}{4} + \frac{3}{4} - \frac{1}{2}\right)$$

$$\Rightarrow \quad \mathbb{I}_{22} = \frac{1}{2\sigma^4}$$

$$\boxed{\mathbb{I}_X(\theta) = \begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{pmatrix}}$$

Puisque le modèle est régulier on pouvait utiliser la formule suivante :

$$\mathbb{I}_{ij} = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta_i\partial\theta_j}\log[f(X;\theta)]\right)$$

Remarque!

Si $\underline{X} = (X_1, \ldots, X_n)$ échantillon de X alors

On a !

* Vecteur Score

$$S(\tilde{X};\theta) = \nabla \log\left[L(\tilde{X};\theta)\right]$$

* L'information de fischer apportée par l'échantillon sur θ :

$$\mathbb{I}_n(\theta) = \mathbb{E}\left[\left(S(\tilde{X};\theta)\right)\left(S(\tilde{X};\theta)\right)^t\right]$$

* $\mathbb{I}_n(\theta) = n\,\mathbb{I}(\theta)$ $\left[\mathbb{I}_{\tilde{X}}(\theta) = n\,\mathbb{I}_{X}(\theta)\right]$

Théorème :

(1) $\frac{1}{n}\sum_{i=1}^{n} S(X_i;\theta) \xrightarrow{p.s} 0$

(D'après la loi forte des grands nombres)

(2) Si les conditions de régularité sont satisfaites et l'information de fischer existe, nous avons la conver-gence, grâce au TCL, suivante

$$\frac{1}{\sqrt{n}} \times S(\tilde{X};\theta) \xrightarrow{\mathcal{L}} \mathcal{N}_{0_k}\left(0;\mathbb{I}(\theta)\right)$$

Preuve (2)

On a :

On a $\dfrac{\overline{\mathbb{V}} - \mathbb{E}(\overline{\mathbb{V}})}{\sqrt{Var(\overline{\mathbb{V}})}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

grâce à la stabilité de la loi normale

$$\overline{\mathbb{V}} - \mathbb{E}(\overline{\mathbb{V}}) \xrightarrow{\mathcal{L}} \mathcal{N}(0, Var(\overline{\mathbb{V}}))$$

$$Var(\overline{\mathbb{V}}) = \frac{\sigma^2}{n} \; ; \; \mathbb{V} \rightsquigarrow \mathcal{N}(\mu, \sigma^2)$$

Par analogie : $\mathbb{V}_i = S(X_i, \theta)$

$$\frac{1}{n} \sum S(X_i, \theta) \xrightarrow{\mathcal{L}} \mathcal{N}_k\left(0_k \; ; \; \frac{V[S(\tilde{X}, \theta)]}{n^2}\right)$$

$$\mathcal{N}_k\left(0_k \; ; \; \frac{\mathbb{I}(\theta)}{n}\right)$$

$$\sqrt{n} \times \frac{1}{n} \sum S(X_i, \theta) \longrightarrow \mathcal{N}_k(0_k \; ; \; \mathbb{I}(\theta))$$

$$\sqrt{n} \times \frac{1}{(\sqrt{n})^2} \sum S(X_i, \theta) \longrightarrow \mathcal{N}_k(0_k \; ; \; \mathbb{I}(\theta))$$

$$\Rightarrow \frac{1}{\sqrt{n}} \sum S(X_i, \theta) \longrightarrow \mathcal{N}_k(0_k, \mathbb{I}(\theta))$$

Théorème :

Si T_n est l'EMV de $g(\theta)$; $g \in C^1$; $I(\theta)$ existe alors :

$$\sqrt{n}\left(T_n - g(\theta)\right) \xrightarrow{\mathcal{L}} \mathcal{N}\left(0 ; \frac{\partial g(\theta)}{\partial \theta} \times I^{-1}(\theta) \times \frac{\partial g(\theta)^t}{\partial \theta}\right)$$

On dit que l'estimateur du MV est asymptotiquement efficace mais aussi asymptotiquement normal.

$$b_{inf}(\theta) = \frac{[g'(\theta)]^2}{I_n(\theta)}$$

$$b_{inf}(\theta) = \frac{\partial g(\theta)}{\partial \theta} I_n^{-1}(\theta) \cdot \frac{\partial g(\theta)^t}{\partial \theta}$$

Dans le cas multidimentionnel

$$V(T_n) \geq b_{inf}(\theta)$$