04&05 Estimation Ponctuelle

mardi 27 septembre 2022

1. Estimation

- -1- Cagre : $X_i, X_n \sim^{iid} X$ de distribution F_X inconnue
- -2- Hypothèse (hyp) : F_X appartient à $\mathcal{F}$, une famille de CDF^{-1} indexée par un paramètre,
  - 1- $F_X \in \mathcal{F} = \{F_\theta, \theta \in \Theta\}, \Theta \in R^d$
- -3- $\Theta \stackrel{\text{def}}{=}$ *emsemble des valeurs possibles de* θ
  - 1- $\exists \theta_0 \in \Theta, F_X = F_{\theta_0}$
- -4- Deux problèmes d'estimation ponctuelle
  - 1- Fournir un estimation de θ
    - 1> $\hat{\theta}_n = \hat{\theta}_n(X_1, \dots, X_n)$
    - 2> Sa réalisation, appelée estimation, est une approximation de θ_0
  - 2- S'assurer de la "performance" de l'estimateur $\hat{\theta}_n$
- -5- Estimation est une valeur, estimateur est une formulation/fonction.

2. Performance

- -1- Estimateur (asymptotiquement) sans biais
  - 1- Biais de l'estimateur
    - 1> $b(n,\theta) = E(\hat{\theta}_n) - \theta$
  - 2- Estimateur (asymptotiquement) sans biais
    - 1> $\lim\limits_{n\to+\infty} b(n,\theta) = 0$
- -2- Estimateur convergent (2 façons)
  - 1- $\lim\limits_{n\to+\infty} Var(\hat{\theta}_n) = 0 \Rightarrow \lim\limits_{n\to+\infty} \hat{\theta}_n = \theta$
  - 2- $\begin{cases} \lim\limits_{n\to+\infty} \hat{\theta}_n = \theta \\ \lim\limits_{n\to+\infty} a_n = 1 \end{cases} \Rightarrow \lim\limits_{n\to+\infty} a_n\hat{\theta}_n = \theta$
- -3- Risque quadratique moyenne/erreur quadratique moyenne : précision d'un estimateur
  - 1- $R(\hat{\theta}_n, \cdot): \theta \in \Theta \to R(\hat{\theta}_n, \theta) = E\left((\hat{\theta}_n - \theta)^2\right)$
  - 2- $R(\hat{\theta}_n, \theta) = Var(\hat{\theta}_n) + [b(n,\theta)]^2$
    - 1> Acddttention: un estimateur biaisé (de biais non nul) de θ peut-être plus précis qu'un estimateur sans biais de θ

3. Moments

- -1- Expression de θ en fonction de $m_{k,\theta}$ ou $m_{k,c,\theta}$

1- $\begin{cases} m_{k,\theta} = E(X^k) \\ m_{k,c,\theta} = E\left(\left(X - E(X)\right)^2\right) \end{cases}$

2- $\theta = g(m_{k,\theta}) = m_{k,\theta}^{-1}$ *ou* $\theta = g(m_{k,c,\theta}), k = 1, \dots, d$

3- $g: (m_{1,\theta}, \dots, m_{d,\theta}) \to \theta = (\theta_1, \dots, \theta_d) \in \Theta$

-2- Propriétés de l'estimateur des moments

1- Convergence des estimateurs des moments ($n \to \infty$)

2- Asymptotiquement normalité des estimateurs des moments

1> $\dfrac{\sqrt{n}(\hat{\theta}_m - \theta)}{\sigma_m} \xrightarrow{\mathcal{L}} N(0,1), n \to +\infty$

4. Maximum de vraisemblance

-1- Vraisemblance 可靠性

1- $\Theta \to R_+, \theta \to L(\theta; x_1, \dots, x_n) =^{iid} \begin{cases} \prod f_X(x_i; \theta), continue \\ \prod P_X(x_i; \theta), discret \end{cases}$

-2- Estimateur du maximum de vraisemblance (EMV)

1- $\hat{\theta}_{n,MV} = \hat{\theta}_{n,MV}(X_1, \dots, X_n)$ *EMV pour* θ *si* $\forall \theta \in \Theta, L(\hat{\theta}_{n,MV}; X_1, \dots, X_n) \geq L(\theta; X_1, \dots, X_n)$

1> $L(\hat{\theta}_{n,MV}; X_1, \dots, X_n) = max\, L(\theta_{n,MV}; X_1, \dots, X_n)$

2> Pas toujours d'existence d'EMV

-3- Log-vraisemblance

1- $l: \theta \in \Theta \mapsto l(\theta; x_1, \dots, x_n) = \ln\left(L(\theta; x_1, \dots, x_n)\right)$

a) c'est pour la vraisemblance identiques

2- Schéma de résolution

1> $\begin{cases} l'(\theta^*; x_1, \dots, x_n) = 0 \\ l''(\theta^*; x_1, \dots, x_n) < 0 \end{cases}$

Ex. A. "Logo favori"

2>

- $Y \sim B(n, \theta)$, $\theta \in \Theta =]0, 1[$; obs: $y = 48$
- Support de Y: $V_Y = \{0, \dots, n\}$ ne dépend pas de θ
- Vraisemblance $\theta \to L(\theta, y) = C_n^y \theta^y (1-\theta)^{n-y}$
- Log-vraisemblance : $\theta \to \ell(\theta, y) = \ln(C_n^y) + y\ln(\theta) + (n-y)\ln(1-\theta)$
- Equation de vraisemblance : $\ell'(\theta^\star, y) = 0 \Leftrightarrow \frac{y}{\theta^\star} - \frac{n-y}{1-\theta^\star} = 0 \Leftrightarrow \theta^\star = \frac{y}{n}$
- On montre que $\ell''(\theta^\star(y), y) < 0$
- $\to \hat{\theta}_{n,MV} = \frac{Y}{n}$

-4- Équations de vraisemblance : paramètre vectoriel

1- $$\begin{cases} \dfrac{\partial l(\theta_1,\dots,\theta_d;x_1,\dots,x_n)}{\partial\theta_j}=0, j=1,\dots,d \\ \left[\left(\dfrac{\partial^2 l(\theta_1,\dots,\theta_d;x_1,\dots,x_n)}{\partial\theta_i\partial\theta_j}\right)_{1\le i,j\le d}\right]_{|\theta} \ matrice\ déf.négative \end{cases}$$

-5- Propriétés

1- Soit $u:\theta\rightarrow\eta=u(\theta)$ une application $\Rightarrow \hat{\eta}_{n,MV}=u(\hat{\theta}_{n,MV})$ est un EMV pour le paramètre $\eta=u(\theta)$

-6- Information de Fisher du modèle $I_n(\theta)$

1- $I_n(\theta)=E\left[\left(l'(\theta;X_1,\dots,X_n)\right)^2\right]\overset{\text{def}}{=} variance\ du\ l'(\theta;X)$

2- Propriétés de l'information de Fisher

1> $I_1(\theta)=-E[l''(\theta;X_1)]$

2> $I_n(\theta)=-E[l''(\theta;X_1,\dots,X_n)]=^{iid} nI_1(\theta)$

3- EMV convergent et asymp. normal

1> l'estimateur du maximum de vraisemblance est convergent

$$\hat{\theta}_{n,MV}\overset{P}{\rightarrow}\theta, quand\ n\rightarrow+\infty$$

2> l'estimateur du maximum de vraisemblance est asymptotiquement gaussien

$$\frac{\hat{\theta}_{n,MV}-\theta}{\sqrt{\frac{1}{I_{n(\theta)}}}}\overset{\mathcal{L}}{\rightarrow}N(0,1), quand\ n\rightarrow+\infty$$

5. Optimalité

-1- Borne de Fréchet-Darmois-Cramer-Rao (FDCR)

1- Conditions de Cramer-Rao : V_X ne dépend pas de θ et conditions de régularité sur la distribution F_X de X

2- Théorème (FDCR) : si $\hat{\theta}_n$ est un estimateur sans biais de et si les conditions de Cramer-Rao sont satisfaites, alors

1> $$Var(\hat{\theta}_n)\ge\frac{1}{I_n(\theta)}\overset{\text{def}}{=}B_F(\theta)=\frac{1}{E\left[\left(l'(\theta;X_1,\dots,X_n)\right)^2\right]}$$

-2- Estimateur efficace

1- Efficacité

1> $Var(\hat{\theta}_n)=\dfrac{1}{I_n(\theta)}, alors\ \hat{\theta}_n\ est\ un\ estimateur\ efficace\ de\ \theta$