Chapitre 3 : suite

Exemple de la méthode des maux de vraisemblance

$(X_1, \ldots, X_n)$ un n-échantillon de loi exponentielle de paramètre $\frac{1}{\theta}$ où $\theta > 0$ inconu

➔ Quel est un estimateur de θ par la méthode du maximum de vraisemblance

Soit X une v.a de loi exponentielle de paramètre $\frac{1}{\theta}$

Sa fonction de densité f est donnér par

$$f(x) = \left| \begin{array}{l} \frac{1}{\theta} e^{-\frac{1}{\theta}x} \ \ si\ x \geq 0 \\ \ \ 0\ si\ x < 0 \end{array} \right.$$

La vraisemblance associé est alors

$L(x_1, \ldots, x_n;\ \theta) = \prod_{i=1}^{n} f(x_i) = f(x_1) * f(x_2) * \ldots * f(x_n)$

$L(x_1, \ldots, x_n;\ \theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-\frac{1}{\theta}x} \ \ ssi\ \forall i, x_i \geq 0 \ = \ \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$

Remarque : si il existe un x_i tel que $x_i < 0 \rightarrow L(x_1, \ldots, x_n;\ \theta) = 0$

<u>Principe du max de vraisemblance :</u>

On considère la fonction $g(\theta) = L(x_1, \ldots, x_n;\ \theta)$

On cherche la solution de θ qui maximise g

Puisque g est nulle si un des x_i est négatif et strictement positive ailleurs, la maximisation se produit là où g est strictement positive

$$g(\theta) = \frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

A la plage de g, on va considerer $\ln g$

$$\ln g(\theta) = \ln(\frac{1}{\theta^n} e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i})$$
$$= \ln\left(\frac{1}{\theta^n}\right) = \ln\left(e^{\frac{1}{\theta}\sum x_i}\right) = \ln(1) - \ln(\theta^n) - \frac{1}{\theta}\sum x_i = -n\ln\theta - \frac{1}{\theta}\sum x_i = h(\theta)$$

Calculons $h'(\theta)$

$$h'(\theta) = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum x_i$$

On résout $h^{'}(\theta) = 0$ (cela va donner les point extremaux)

$$h^{'}(\theta) = 0 \rightarrow \frac{n}{\theta} = \frac{1}{\theta^2}\sum x_i \rightarrow \frac{\sum x_i}{n} = \overline{x_n}$$

Il faut regarder si $\overline{x_n}$ est un maximum

On regarde alors si $h^{''}(\overline{x_n}) < 0$

$$h^{''}(\theta) = \frac{n}{\theta^2} - 2/\theta^3 \sum x_i$$

D'où $h^{''}(\overline{x_n}) = \frac{n}{\overline{x_n}^2} - \frac{2}{\overline{x_n}^3}\sum x_i = \frac{n}{\overline{x_n}^2} - \frac{2}{\overline{x_n}^3} * n\overline{x_n} = \frac{n}{\overline{x_n}^2} - \frac{2n}{\overline{x_n}^2} = -\frac{n}{\overline{x_n}^2} < 0$

- ➔ $\overline{x_n}$ est un maximum

Donc $\overline{x_n}$ est un estimateur de θ, l'estimateur associé est $\overline{X_n} = \frac{1}{n}\sum_{i=i}^{n} X_i$

Remarque : Si on applique la méthode des moments

X$\sim \mathcal{E}\left(\frac{1}{\theta}\right) \rightarrow E[X] = \theta$ et V[X] = θ^2

E[X] a un estimateur $\overline{x_n}$ et V[X] un estimateur $\sqrt{S_n^2}$ avec $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x_n})^2$

- ➔ Pour un même paramètre à utiliser, il peut y avoir plusieurs estimateurs, comment les comparer ?

II. Propriétés d'un estimateur

Soit θ un paramètre inconnu à estimer et $T_n(X_1, \dots, X_n)$ un estimateur de θ

Soit $(x_1, \dots, x_n)$ une observation de $(X_1, \dots, X_n)$

Pour que $T_n(X_1, \dots, X_n)$ soit un bon estimateur de θ, on a envie que $T_n(X_1, \dots, X_n)$ soitproche de la valeur fixe θ

Puisque $T_n(X_1, \dots, X_n)$ est une v.a, on ne va pas pouvoir comparer directement $T_n(X_1, \dots, X_n)$ à θ. On va le faire en moyenne

<u>Def :</u> On appelle biais de $T_n(X_1, \dots, X_n)$ la quantité $b_n(\theta) = E_\theta(T_n(X_1, \dots, X_n)) - \theta$

Un estimateur est dit sans biais si son biais est nul

Ex : Si $\theta = E[\lambda]$, alors un estimateur classique de θ $est\ \overline{x_n} = \frac{1}{n}\sum X_i$ où $(X_1, \dots, X_n)$ est un n-échantillon de même loi que X

$E_\theta(\overline{x_n})$ = $E_\theta\left(\frac{1}{n\sum x_i}\right) = 1/n\ \sum_{i=1}^{n} E_\theta(X_i) = 1/n\ \sum_{i=1}^{n} E_\theta(X) = E_\theta[X] = \theta$

➔ $b_n(\theta) = 0$

Donc $\overline{x_n}$ est toujours un estimateur sans biais de E[X]

Si $\theta = V[X], un\ estimateur\ classique\ est\ S_n^2$ avec $S_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{x_n})^2$

On veut regarder si S_n^2 est un estimateur sans biais de V[X], donc on doit calculer $E_\theta[S_n^2]$

$$\sum(X_i - \overline{x_n})^2 = \sum((X_i - E[X]] - (E[X] - \overline{x_n}))^2$$
$$= \sum(X_i - E[X]]^2 - \sum(E[X] - \overline{x_n}))^2 + 2\sum(X_i - E[X]] - (E[X] - \overline{x_n}))$$

Si $\sum\left((X_i - E[X]] - (E[X] - \overline{x_n})\right) = (E[X] - \overline{x_n}))\ \sum(X_i - E[X]]$ = $(E[X] - \overline{x_n})(\sum X_i - n\,E[X] = (E[X] - \overline{x_n})(n\overline{x_n} - nE[X] = -n(E[X] - \overline{x_n})^2$

D'où $\sum\ (X_i - \overline{x_n})^2$ = $\sum((X_i - E[X]]^2 + n(E[X] - \overline{x_n})^2 + 2(-n(E[X] - \overline{x_n})^2) = \sum(X_i - E[X]]^2 - n(E[X] - \overline{x_n})^2$

Donc $E_\theta(S_n^2)$ = $E_\theta\left(\frac{1}{n}\sum\ (X_i - \overline{x_n}\right)^2 = \frac{1}{n}E_\theta[\sum(X_i - E[X]]^2 - n(E[X] - \overline{x_n})^2]$

$E_\theta(S_n^2)$ = 1/n $\sum_{i=1}^{n} E_\theta[(X_{i-E[X])^2}]$

= $E_\theta[(E[X] - \overline{x_n})^2]$

A $E_\theta(X_i - E[X]^2] = V[X_i]$

$E_\theta((\overline{x_n} - E[X])^2)$ = V$[\overline{x_n}]$

$E_\theta[S_n^2]$ = 1/n $\sum V[X_i] - V[\overline{x_n}]$

=$\frac{1}{n} * nV[X] - V[\overline{x_n}]$

V$[\overline{x_n}] = V[\frac{1}{n}\sum X_i] = \frac{1}{n^2}V[\sum X_i] = \frac{1}{n^2}\sum V[X] <- \ car\ indépendant$

= $\frac{nV[X]}{n^2} car\ de\ meme\ loi\ que\ C = V[X]/n$

D'où $E_\theta(S_n^2) = V[X] - \frac{V[X]}{n} = \frac{n-1}{n}V[X]$

$b_n(\theta) = E_\theta(S_n^2) - \theta = -\frac{V[X]}{n} \neq O\ puique\ ici\ \theta = V[X]$

Comment modifier S_n^2 pour obtenir un estimateur sans biais de V[X] ?

Posons $T_n(X_1, \dots, X_n) = S_n^2 + \frac{V[X]}{n}$

- Est-ce que T_n est un estimateur sans biais de V[X] ?

Ceci ne peut pas être un estimateur sans biais de V[X]

Certes $E_\theta(T_n(X_1, \dots, X_n))$ = E[S_n^2] $+ \frac{V[\overline{x_n}]}{n} = V[X]$

Pb ici : Tn dépend de la quantité à donner -> pas un estimateur

Ici $E_\theta(S_n^2)$ = (n-1)/n V[X]

- ➔ $E_\theta(\frac{n}{n-1} S_n^2)$ = V[X]

n/(n-1) S_n^2 est un estimateur sans biais de V[X]

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

Remarque : S_n^2 n'est pas un estimateur sans biais mais il est asymptotiquement sans biais

Def : $T_n(X_1, \dots, X_n)$ est une stimateur asymptotiqement sans biais de θ si $\lim_{n \to \infty} b_n(\theta) = 0$

Convergence d'un estimateur

Intuitivement, plus n est grand , plus l'estimateur Tn devrait se rapprocher de la vrai valeur θ

Def : Un estimateur Tn de θ es dit convergent si la suite (Tn)n converge en probabilité vers θ, autrement dit

$\forall \varepsilon > 0, Pn(|Tn - \theta| > \varepsilon)\overrightarrow{n \to +\infty}$ 0

Th : toute estimateur sans biais dont la variance tend vers 0 lorsque n tend vers ∞ est convergent

Preuve : Par l'inégalité de Bienaymé-Tchebychev

P(|Tn – E[Tn]| > $\varepsilon) \leq \frac{V[Tn]}{\varepsilon^2}$

Si Tn estimateur sans biais de θ, alors E[Tn]= θ

P(|Tn – $\theta)|> \varepsilon) \leq \frac{V[Tn]}{\varepsilon^2}$

Qualité d'un estimateur ?

Soit Tn estimateur de θ

L'erreur quadratique moyenne EQ (Tn) = $E_\theta((Tn - \theta)^2)$

Th : EQ(Tn) = V(Tn) + ${b_n}^2\,(\theta)$

Preuve

EQ (Tn) = $E_\theta((Tn-\theta)^2) = E_\theta([(Tn - E[Tn]) + (E[Tn] - \theta)]^2)$