

Pour rappel, dans un modèle statistique $(\mathcal{X}^n, \otimes \mathcal{F}, \otimes P_\theta)$ où l'on observe n réalisations $(x_i)_{1 \leq i \leq n}$ de v.a iid $(X_i)_{1 \leq i \leq n}$ de loi P_θ , la vraisemblance $L(x_1, \dots, x_n, \theta)$ est donnée par les formules

$$L(x_1, \dots, x_n, \theta) = \begin{cases} \prod_{i=1}^n P_\theta(X_i = x_i) & \text{si les } X_i \text{ sont des v.a discrètes} \\ \prod_{i=1}^n f_\theta(x_i) & \text{si les } X_i \text{ sont à densité, de densité connue } f_\theta. \end{cases}$$

L'estimateur du maximum de vraisemblance $\hat{\theta}_n$ de θ est alors défini comme le maximiseur de la vraisemblance:

$$\hat{\theta}_n := \operatorname{argmax}_{\theta \in \Theta} L(x_1, \dots, x_n, \theta)$$

Par définition, c'est une fonction de seules données $(x_i)_{1 \leq i \leq n}$.

Exemple : si $X_i \sim \mathcal{E}(\theta)$ où $\theta > 0$ est inconnu ie

$$f_\theta(x) = \lambda e^{-\lambda x} \mathbb{1}_{x \geq 0}, \text{ alors}$$

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n (\theta e^{-\theta x_i} \mathbb{1}_{x_i \geq 0}) = \theta^n e^{-\theta \sum_{i=1}^n x_i} \mathbb{1}_{x_1 \geq 0, \dots, x_n \geq 0}$$

Si la fonction $\theta \mapsto L(x_1, \dots, x_n, \theta)$ est maximale en $\hat{\theta}_n$ alors sa dérivée s'annule en $\hat{\theta}_n$ or

$$\frac{\partial}{\partial \theta} L(x_1, \dots, x_n, \theta) = L(x_1, \dots, x_n, \theta) \left(\frac{n}{\theta} - \sum_{i=1}^n x_i \right)$$

$$\text{et } \frac{\partial}{\partial \theta} L(x_1, \dots, x_n, \theta) = 0 \Leftrightarrow \theta = \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^{-1}$$

Ainsi $\hat{\theta}_n = \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^{-1}$ qui est bien un maximiseur de $L(x_1, \dots, x_n, \theta)$

Exercice 1

1) On observe $X_i \sim \mathcal{S}(\lambda)$ où $\lambda > 0$ est inconnu.

Ce sont des variables discrètes donc la vraisemblance est donnée par

$$L(x_1, \dots, x_n, \lambda) = \prod_{i=1}^n P(X_i = x_i) = \prod_{i=1}^n \left(e^{-\lambda} \frac{\lambda^{x_i}}{x_i!} \right) = e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}$$

Si $\hat{\lambda}_n$ est le maximiseur de L alors $\frac{\partial}{\partial \lambda} L|_{\lambda=\hat{\lambda}_n} = 0$ or

$$\frac{\partial}{\partial \lambda} L(x_1, \dots, x_n, \lambda) = L(x_1, \dots, x_n, \lambda) \left(-n + \frac{\sum_{i=1}^n x_i}{\lambda} \right).$$

de sorte que $\frac{\partial}{\partial \lambda} L(x_1, \dots, x_n, \lambda) = 0 \Leftrightarrow \lambda = \frac{\sum_{i=1}^n x_i}{n}$.

Ainsi $\hat{\lambda}_n = \frac{1}{n} \sum_{i=1}^n x_i$ qui est bien un maximiseur.

2) On observe une variable $X \sim \beta(m, p)$ avec

m connu et p inconnue. La vraisemblance est

$$L(x, p) = P(X=x) = \binom{m}{x} p^x (1-p)^{m-x}$$

On cherche à maximiser L en fonction de $p \in [0, 1]$.

$$\frac{\partial L(x, p)}{\partial p} = \left(\frac{x}{p} - \frac{m-x}{1-p} \right) L(x, p) \quad \text{de sorte que}$$

$$\frac{\partial L}{\partial p} = 0 \Leftrightarrow \frac{x}{m} \cdot \frac{1}{p} = \frac{1-x/m}{1-p} \Leftrightarrow \frac{1-p}{p} \Leftrightarrow \frac{1-x/m}{x/m}$$

$$\Leftrightarrow p = \frac{x}{m}$$

Ainsi $\hat{p}_n = \frac{x}{m}$ qui est bien un maximiseur de $L(x, p)$.

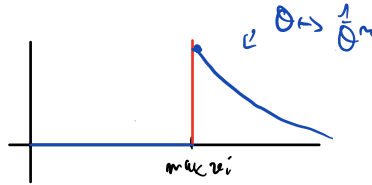
3) On observe n réalisations (x_i) de va iid (X_i) de loi uniforme $U[0, \theta]$ ie de densité connue $f_\theta(u) = \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(u)$. La vraisemblance s'écrit

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f_{\theta}(x_i) = \prod_{i=1}^n \frac{1}{\theta} \mathbb{1}_{[0, \theta]}(x_i)$$

$$= \frac{1}{\theta^n} \mathbb{1}_{0 \leq \min_{1 \leq i \leq n} (x_i) \leq \max_{1 \leq i \leq n} (x_i) \leq \theta}$$

La fonction $\theta \mapsto L(u_1, \dots, u_n, \theta)$ est maximale en

$$\theta_m^* = \max_{1 \dots m} (x_i)$$



Exercice 2

On observe une réalisation x d'un X de loi $\mathcal{P}(\theta)$ avec $\theta \in \left\{ \frac{1}{10}, \frac{3}{10} \right\}$. La variable aléatoire

$$L(\theta, x) = P_\theta(X=x) = \begin{cases} \theta & \text{if } x=1 \\ 1-\theta & \text{if } x=0 \end{cases}$$

alternativ d.h. $L(x, \theta) = \theta^x (1-\theta)^{1-x}$.

Dans ce modèle, on cherche donc le maximum en θ

$$\left(\frac{1}{10}\right)^u \left(\frac{9}{10}\right)^{1-u} \quad \text{or} \quad \left(\frac{8}{10}\right)^u \left(\frac{2}{10}\right)^{1-u}$$

S: $\begin{cases} x=1 \\ x=0 \end{cases}$ le max est $\frac{\partial}{\partial x}$ i.e. $\Theta = \frac{\partial}{\partial x}$
 $\frac{\partial}{\partial y}$ i.e. $\Theta = \frac{1}{10}$

Exercice 3

On peut écrire le modèle sous forme linéaire

$$Y = \theta + \sum \varepsilon$$

$$\text{avec } \Sigma = \sigma^2 \begin{pmatrix} \frac{1}{g(i/n)} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{g(1/n)} \end{pmatrix} \text{ et } \varepsilon \sim \mathcal{N}(0, Id)$$

En particulier $Y \sim \mathcal{N}_n(\theta, \Sigma)$ avec $\Sigma^{-1} = \begin{pmatrix} \sigma^2 g(i/n) & & 0 \\ & \ddots & \\ 0 & & \sigma^2 g(1/n) \end{pmatrix}$

La densité de Y_i est donc

$$f_i(x) = \frac{1}{\sqrt{2\pi \frac{\sigma^2}{g(i/n)}}} e^{-\frac{(x-\theta)^2 g(i/n)}{2\sigma^2}}$$

et la vraisemblance du modèle est donnée par

$$L(x_1, \dots, x_n, \theta) = \prod_i f_i(x_i) = \frac{e^{-\sum \frac{(x_i - \theta)^2}{\sigma^2} g(i/n)}}{\sqrt{2\pi \sigma^2}^n \sqrt{\prod_i g(i/n)}}$$

Maximiser L revient à maximiser son $\log$ avec

$$\log L(x_1, \dots, x_n, \theta) = -\sum_i \frac{(x_i - \theta)^2}{\sigma^2} g(i/n) - \frac{n}{2} \log(2\pi \sigma^2) - \frac{1}{2} \log\left(\prod_i g(i/n)\right).$$

$$\text{On a alors } \frac{\partial}{\partial \theta} \log L = -\sum_i 2 \frac{(x_i - \theta)}{\sigma^2} g(i/n)$$

$$\text{et } \frac{\partial}{\partial \theta} \log L = 0 \Leftrightarrow \sum_i x_i g(i/n) = \theta \sum_i g(i/n)$$

$$\text{ie } \hat{\theta}_n = \frac{\sum_i x_i g(i/n)}{\sum_i g(i/n)}.$$

On a alors, com $\mathbb{E}[Y_i] = \theta$

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\left[\frac{\sum_1^m Y_i g(i/n)}{\sum_1^m g(i/n)}\right] = \theta.$$

$$\text{et } \text{var}(\hat{\theta}_n - \theta) = \mathbb{E}[(\hat{\theta}_n - \theta)^2]$$

$$= \mathbb{E}\left[\left(\frac{\sum_1^m (Y_i - \theta) g(i/n)}{\sum_1^m g(i/n)}\right)^2\right]$$

$$= \frac{1}{\left(\sum_1^m g(i/n)\right)^2} \sum_1^m \text{var}(Y_i) g^2(i/n)$$

$$= \frac{1}{\left(\sum_1^m g(i/n)\right)^2} \cdot \sum_1^m \frac{\sigma^2}{g(i/n)} g^2(i/n)$$

$$= \frac{\sigma^2}{\sum_1^m g(i/n)}$$

Exercice 4

On considère des v.a. i.i.d. (X_i) de loi $\Gamma(k, \frac{1}{\theta})$ i.e.

de densité $f_{\theta}(x) = \left(\frac{1}{\theta}\right)^k \frac{1}{\Gamma(k)} x^{k-1} e^{-\frac{1}{\theta}x} \mathbb{1}_{x \geq 0}$

où k est connu et θ est inconnu.

Le modèle est donc

$$(\mathbb{R}^m, \otimes_1 P(\mathbb{R}^+), \otimes_1 P_{\theta}) \quad \text{où } P_{\theta} \text{ de densité } f_{\theta}.$$

La vraisemblance est donnée par

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f_\theta(x_i) \\ = \left(\frac{1}{\theta}\right)^{m_k} \frac{1}{\Gamma(k)^m} \left(\prod_{i=1}^n x_i\right)^{k-1} e^{-\frac{1}{\theta} \sum_{i=1}^n x_i} \mathbb{1}_{x_i \geq 0}.$$

Maximiser L revient à maximiser son log

$$\log L(x_1, \dots, x_n, \theta) = -m_k \log(\theta) - m \log(\Gamma(k)) \\ + (k-1) \sum_{i=1}^n \log(x_i) - \frac{1}{\theta} \sum_{i=1}^n x_i$$

On dérive par rapport à θ

$$\frac{\partial}{\partial \theta} \log L(x_1, \dots, x_n, \theta) = -\frac{m_k}{\theta} + \frac{1}{\theta^2} \sum x_i$$

$$\frac{\partial}{\partial \theta} \log L = 0 \Leftrightarrow \hat{\theta}_n = \frac{1}{m_k} \sum_{i=1}^n x_i$$

qui est bien un maximum.

On rappelle que si $X \sim \Gamma(k, \frac{1}{\theta})$ alors

$$E[X] = \left(\frac{1}{\theta}\right)^k \cdot \frac{1}{\Gamma(k)} \int_0^{\infty} x^k e^{-\frac{1}{\theta} x} dx = \left(\frac{1}{\theta}\right)^k \cdot \frac{1}{\Gamma(k)} \theta^{k+1} \Gamma(k+1)$$

$$= k \cdot \theta.$$

$$E[X^2] = \left(\frac{1}{\theta}\right)^k \frac{1}{\Gamma(k)} \int_0^{\infty} x^{k+1} e^{-\frac{1}{\theta} x} dx = \left(\frac{1}{\theta}\right)^k \frac{1}{\Gamma(k)} \Gamma(k+2) \cdot \theta^{k+2}$$

$$= \theta^2 k(k+1)$$

$$\text{donc } \text{var}(X) = \theta^2 k^2 + \theta^2 k - k^2 \theta^2 = \theta^2 k.$$

Par suite, comme les (X_i) sont indépendantes

$$E(\hat{\Theta}_n) = \frac{1}{n\ell} \sum_1^n E(X_i) = \Theta.$$

$$\text{var}(\hat{\Theta}_n) = \frac{1}{n^2\ell^2} \sum_1^n \text{var}(X_i) = \frac{\Theta^2}{n\ell}$$

i.e. $\hat{\Theta}_n$ est sans biais et $E[(\hat{\Theta}_n - \Theta)^2] \xrightarrow[n \rightarrow \infty]{} 0$.

On rappelle que l'information de Fisher $I_n(\Theta)$ est définie comme

$$I_n(\Theta) = E \left[\left(\frac{\partial}{\partial \Theta} \log L(X_1, \dots, X_n, \Theta) \right)^2 \right]$$

$$I_n: \frac{\partial}{\partial \Theta} \log L(X_1, \dots, X_n, \Theta) = -\frac{n\ell}{\Theta} + \frac{1}{\Theta^2} \sum_1^n X_i$$

$$= -\frac{n\ell}{\Theta^2} \left(\Theta - \frac{1}{n\ell} \sum_1^n X_i \right)$$

$$= -\frac{n\ell}{\Theta^2} (\Theta - \hat{\Theta}_n)$$

$$\text{et } I_n(\Theta) = \frac{n^2\ell^2}{\Theta^4} \text{var}(\hat{\Theta}_n) = \frac{n^2\ell^2}{\Theta^4} \frac{\Theta^2}{n\ell} = \frac{n\ell}{\Theta^2}$$

de sorte que

$$\text{Eff}(\hat{\Theta}_n) = \frac{1}{\text{var}(\hat{\Theta}_n) I_n(\Theta)} = 1.$$

Autrement dit, $\hat{\Theta}_n$ est un estimateur efficace.

Exercice 5

- On observe des variables (X_i) de loi $\Gamma(2, \theta)$ i.e. de densité $f_\theta(x) = \theta^2 x e^{-\theta x} \mathbb{1}_{x \geq 0}$.

On a alors $E[X] = \theta^2 \int_{\mathbb{R}^+} x^2 e^{-\theta x} dx = 2 \frac{\theta^2}{\theta^3} = \frac{2}{\theta}$

Ainsi d'après la LGV, l'estimateur des moindres carrés

est $\tilde{\theta}_n = 2 \left(\frac{\sum_{i=1}^n X_i}{n} \right)^{-1}$

- La vraisemblance est donnée par

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f_\theta(x_i) = \theta^{2n} \left(\prod_{i=1}^n x_i \right) e^{-\theta \sum_{i=1}^n x_i} \mathbb{1}_{x_i \geq 0 \text{ } \forall i}$$

$$\log L(x_1, \dots, x_n, \theta) = 2n \log \theta + \sum_{i=1}^n \log x_i - \theta \sum_{i=1}^n x_i$$

$$\frac{\partial \log L(x_1, \dots, x_n, \theta)}{\partial \theta} = \frac{2n}{\theta} - \sum_{i=1}^n x_i$$

$$\text{et } \hat{\theta}_n = \frac{2n}{\sum_{i=1}^n x_i} = \tilde{\theta}_n.$$

- Si $X_i \sim \Gamma(2, \theta)$ alors $\sum_{i=1}^n X_i \sim \Gamma(2n, \theta)$

$$\text{donc } E[\hat{\theta}_n] = 2n E\left[\frac{1}{\sum_{i=1}^n X_i}\right]$$

$$= 2n \cdot \frac{\theta^{2n}}{\Gamma(2n)} \int_{\mathbb{R}^+} \frac{1}{x} \cdot x^{2n-1} e^{-\theta x} dx$$

$$= 2n \frac{\theta^{2n}}{\Gamma(2n)} \int x^{2n-1-1} e^{-\theta x} dx$$

$$\begin{aligned} \text{ie } E[\hat{\theta}_n] &= 2n \frac{\theta^{2n}}{\Gamma(2n)} \frac{\Gamma(2n-1)}{\theta^{2n-1}} = \frac{2n\theta}{2n-1} \\ &= \theta + \frac{\theta}{2n-1} \end{aligned}$$

$$\begin{aligned} E[\hat{\theta}_n^2] &= h_n^2 \frac{\theta^{2n}}{\Gamma(2n)} \int \frac{1}{x^2} x^{2n-1} e^{-\theta x} dx \\ &= h_n^2 \frac{\theta^{2n}}{\Gamma(2n)} \int x^{(2n-2)-1} e^{-\theta x} dx \\ &= h_n^2 \frac{\theta^{2n}}{\Gamma(2n)} \frac{\Gamma(2n-2)}{\theta^{2n-2}} = \frac{h_n^2 \theta^2}{(2n-1)(2n-2)} \end{aligned}$$

$$\begin{aligned} \text{var}(\hat{\theta}_n) &= h_n^2 \theta^2 \left[\frac{1}{(2n-1)(2n-2)} - \frac{1}{(2n-1)^2} \right] \\ &= \frac{h_n^2 \theta^2}{(2n-1)} \left(\frac{1}{2n-2} - \frac{1}{2n-1} \right) \\ &= \frac{h_n^2 \theta^2}{(2n-1)^2 (2n-2)} = O\left(\frac{1}{n}\right) \end{aligned}$$

Exercice 6

On a vu dans l'exo 1 que l'EMV de $S(\lambda)$ est $\hat{\lambda}_n = \frac{1}{n} \sum_{i=1}^n X_i$
 On a alors $E[\hat{\lambda}_n] = \lambda$ et $\text{var}(\hat{\lambda}_n) = \frac{\lambda}{n}$.

On rappelle que la vraisemblance est donnée par

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

$$= e^{-n\lambda} \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n (x_i!)}$$

de sorte que

$$\log L(x_1, \dots, x_n, \lambda) = -n\lambda + \left(\sum_{i=1}^n x_i\right) \log \lambda - \sum_{i=1}^n \log(x_i!)$$

$$\frac{\partial}{\partial \lambda} \log L(x_1, \dots, x_n, \lambda) = -n + \sum_{i=1}^n x_i \cdot \frac{1}{\lambda}$$

$$= \frac{-n}{\lambda} \left(\lambda - \sum_{i=1}^n x_i \right)$$

de sorte que

$$I_n(\lambda) = \frac{n^2}{\lambda^2} \text{var}(\bar{\lambda}_n) = \frac{n^2}{\lambda^2} \frac{\lambda}{n} = \frac{n}{\lambda}$$

Autrement dit

$$\text{Eff}(\hat{\lambda}_n) = \frac{1}{\text{var}(\hat{\lambda}_n) I_n(\lambda)} \stackrel{\text{sans biais}}{=} 1$$

ie $\bar{\lambda}_n$ est efficace.

Exercice 7

Le modèle est $(\mathbb{R}^{+m}, \otimes_{i=1}^m \mathcal{P}(\mathbb{R}^+), \otimes_{i=1}^m P_\theta)$ où P_θ de densité $f_\theta(x) = \frac{1}{2\sqrt{x\theta}} \mathbb{1}_{0 < x \leq \theta}$.

La vraisemblance s'écrit

$$L(x_1, \dots, x_m, \theta) = \prod_{i=1}^m f_\theta(x_i) = \frac{1}{2^m \theta^{m/2}} \left(\prod_{i=1}^m x_i \right)^{-\frac{1}{2}} \mathbb{1}_{0 < \min(x_i) \leq \max(x_i) \leq \theta}$$

Cette fonction est maximale en $\hat{\theta}_0 = \max(x_i)$

La fonction de répartition des X_i est

$$\begin{aligned} F_X(x) &= \int_{-\infty}^x f_\theta(u) du \\ &= \int_0^{x \wedge \theta} \frac{1}{2\sqrt{u\theta}} du = \left[\frac{\sqrt{u}}{\sqrt{\theta}} \right]_0^{x \wedge \theta} \\ &= \sqrt{\frac{x \wedge \theta}{\theta}} \end{aligned}$$

La fonction de répartition de $\hat{\theta}_0 = \max(X_i)$ est alors

$$\begin{aligned} F_{\hat{\theta}_0}(x) &= P(\max(X_i) \leq x) \\ &= P(\forall i, X_i \leq x) \\ &\stackrel{i.i.d.}{=} P(X_1 \leq x)^m = \left(\frac{x \wedge \theta}{\theta} \right)^{\frac{m}{2}}. \end{aligned}$$

Comme $\bar{\theta}_0 \geq 0$, on peut utiliser la formule

$$\begin{aligned} E[\bar{\theta}_0] &= \int_{\mathbb{R}^+} P(\bar{\theta}_0 > u) du \\ &= \int_0^{\theta} \left(1 - \left(\frac{u}{\theta}\right)^{\frac{m}{2}}\right) du \\ &= \theta \int_0^1 (1 - u^{\frac{m}{2}}) du \\ &= \theta \left[1 - \left[\frac{2}{\frac{m}{2}+1} u^{\frac{m}{2}+1}\right]_0^1\right] \\ &= \theta \left(1 - \frac{2}{m+2}\right). \end{aligned}$$

$$\begin{aligned} \bullet \quad O_1 \sim E[X_1] &= \int_0^{\theta} \frac{u}{2\sqrt{u}\theta} du = \frac{1}{2\sqrt{\theta}} \int_0^{\theta} \sqrt{u} du \\ &= \frac{1}{2\sqrt{\theta}} \left[\frac{2}{3} u^{3/2}\right]_0^{\theta} = \frac{1}{3} \theta. \end{aligned}$$

L'estimateur des moindres carrés est donc

$$\hat{\theta}_2 = 3 \left(\frac{1}{n} \sum_{i=1}^n X_i \right).$$