Solution Fiche TD n°1

Exercice 1. Pour tout nombre réel θ, on définit l'application p_θ sur $\mathbb{R}$ par :

$$p_\theta(x) = \begin{cases} 0 & \text{si } x < \theta \\ e^{-(x-\theta)} & \text{si } x \geq \theta \end{cases}$$

1. Montrer que, pour tout nombre réel θ, l'application p_θ est une densité de probabilité sur $\mathbb{R}$.
2. Soit X_0 une variable aléatoire de densité p_θ. Montrer que X_0 admet un moment d'ordre deux, et expliciter son espérance et sa variance.
3. Déterminer la fonction de répartition F_0 de X_0.
4. Soit $n \geq 1$ et $(X_1, \dots, X_n)$ un échantillon de taille n de la loi de X_0. On pose $U_n = \min(X_1, \dots, X_n)$ et $M_n = \frac{1}{n}\sum_{i=1}^n X_i$.
   (a) Pour tout nombre réel t, calculer $\mathbb{P}[U_n > t]$. En déduire la loi de U_n.
   (b) Montrer que $M_n - 1$ et $U_n - \frac{1}{n}$ sont des estimateurs sans biais de θ.

Exercice 2. Soit $(X_1, \dots, X_n)$ un n-éch antillon de loi de Bernoulli de paramètre θ

$$P(X_i = x_i) = \theta^{x_i}(1-\theta)^{1-x_i} \quad \forall x_i \in \{0,1\}, \theta \in]0,1[$$

1. Ecrire la vraisemblance de l'échantillon
2. Montrer que $T(X_1, \dots, X_n) = \sum_{i=1}^n X_i$ est une statistique exhaustive pour θ.

Solution :

L'ensemble des valeurs possibles est $\{0,1\}$. Le paramètre inconnu est p. Si $(x_1, \dots, x_n) \in \{0,1\}^n$ est un échantillon, la vraisemblance vaut :

$$L(x_1, \dots, x_n, p) = p^{\sum x_i}(1-p)^{n-\sum x_i} .$$

Son logarithme est :

$$\log(L(x_1, \dots, x_n, p)) = (\sum x_i) \log p + (n - \sum x_i) \log(1-p) .$$

La dérivée par rapport à p est :

$$\frac{\partial \log(L(x_1, \dots, x_n, p))}{\partial p} = (\sum x_i)\frac{1}{p} - (n - \sum x_i)\frac{1}{1-p} .$$

Elle s'annule pour :

$$\widehat{p} = \frac{\sum x_i}{n} .$$

La dérivée seconde est :

$$\frac{\partial^2 \log(L(x_1, \dots, x_n, p))}{\partial p^2} = -(\sum x_i)\frac{1}{p^2} - (n - \sum x_i)\frac{1}{(1-p)^2} .$$

Elle est strictement négative, la valeur $\widehat{p}$ est bien un maximum. Si $(X_1, \dots, X_n)$ est un échantillon de la loi de Binomiale (Bernoulli) de paramètre p, l'estimateur du maximum de vraisemblance de p est :

$$\frac{\sum X_i}{n} ,$$

à savoir la fréquence empirique.

Exercice 3. Pour chacune des lois suivantes, écrire la vraisemblance de l'échantillon $(X_1, \dots, X_n)$ et donner une statistique exhaustive.

1. Loi de Poisson de paramètre θ,
2. Loi de Uniforme sur $[0, \theta]$.

Solution :

5. Pour $X \sim \text{Poisson}(\lambda)$, et $x \in \mathbb{N}$

$$\mathbb{P}(X = x) = \frac{\lambda^x}{x!} e^{-\lambda}$$

Donc pour $X_1, \ldots, X_n$ i.i.d $\sim$ Poisson (λ), (il s'git d'une loi discrète ici) pour un échantillor $(x_1, \cdots, x_n)$ donné, la vraisemblance s'écrit

$$\begin{aligned} L(\lambda; x_1, \cdots, x_n) &= \prod_{i=1}^{n} \mathbb{P}(X_i = x_i) \\ &= \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} e^{-\lambda} \\ &= \frac{e^{-n\lambda}}{x_1! \cdots x_n!} \lambda^{x_1 + \cdots + x_n} \end{aligned}$$

La log-vraisemblance s'écrit

$$\begin{aligned} \ell(\lambda; x_1, \cdots, x_n) &= \log L(\lambda; x_1, \cdots, x_n) \\ &= -n\lambda - \log(x_1! \cdots x_n!) + (x_1 + \cdots + x_n) \log(\lambda) \end{aligned}$$

On cherche a maximiser $\ell(\lambda; x_1, \ldots, x_n)$ pour $\lambda > \mathbb{C}$. On remarque que $\ell \to -\infty$ quand $\lambda \longrightarrow 0$, et $\ell \to -\infty$ quand $\lambda \longrightarrow +\infty$. En plus,

$$\frac{\partial \ell}{\partial \lambda} = 0 \to -n + (x_1 + \cdots + x_n) \frac{1}{\lambda} = 0$$

On obtient l'estimation de maximum de vraisemblance

$$\hat{\lambda}^{\mathrm{MV}} = \frac{1}{n}(x_1 + \cdots + x_n)$$

$$\hat{\lambda}^{\mathrm{MV}} = \frac{1}{\mathrm{n}}(X_1 + \cdots + X_{\mathrm{n}})$$

et nous retrouvons donc

$$\hat{\lambda}^{\mathrm{MV}} = \bar{X}_{\mathrm{n}}.$$

$\boxed{\text{Exercice 4.}}$ Soit $(X_1, \ldots, X_n)$ un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{1}{\theta}(x-\gamma)} \text{ si } x > \gamma \\ 0 \text{ sinon} \end{cases}$$

où $\theta > 0$. Déterminez une estimation des paramètres θ et γ par la méthode du MV.

$\boxed{\text{Exercice 5.}}$

Soit $X_1, \ldots, X_n$ un échantillon aléatoire simple issu d'une population de densité

$$f_\theta(x) = \begin{cases} \frac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$$

où $1/2 < \theta < 1$. Déterminez le MV pour θ.

Solution :

$$\begin{aligned} L_{\theta,\gamma}(\mathbf{X}) &= \prod_{i=1}^{n} \frac{1}{\theta} \exp\left(-\theta^{-1}(X_i - \gamma)\right) \mathbb{I}_{[\gamma,+\infty[}(X_i) \\ &= \frac{1}{\theta^n} \exp\left(-\theta^{-1} \sum (X_i - \gamma)\right) \mathbb{I}_{[\gamma,+\infty}\left[X_{(1)}\right). \end{aligned}$$

En se limitant à $X_{(1)} > \gamma$,

$$\log L_{\theta,\gamma}(\underline{\mathbf{X}}) = -n \log \theta - \frac{1}{\theta} \sum_i (X_i - \gamma)$$

$$\partial_\theta \log L_{\theta,\gamma}(\underline{\mathbf{X}}) = -n\frac{1}{\theta} + \frac{1}{\theta^2} \sum_i (X_i - \gamma)$$

$$\begin{aligned} \partial_\theta \log L_{\theta,\gamma}(\underline{\mathbf{X}}) = 0 &\Leftrightarrow -n\theta + \sum_i X_i - n\gamma = 0 \\ &\Leftrightarrow \theta = \bar{X} - \gamma \end{aligned}$$

$$\partial_\gamma \log L_{\theta,\gamma}(\underline{\mathbf{X}}) = \frac{n}{\theta}.$$

Cette dernière quantité n'est jamais nulle. Souhaitant maximiser la vraisemblance, on remarque qu'à θ fixé, la vraisemblance est une fonction croissante de γ. Quand γ prend sa valeur maximale, la vraisemblance sera maximale. Or, $\gamma \leq X_{(1)}$.

On trouve alors
$$\hat{\gamma} = X_{(1)} = X_{\min} \text{ et } \hat{\theta} = \bar{X} - X_{(1)}.$$

Exercice 6.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité
$$f_\theta(x) = \frac{\theta+1}{2}(1-|x|)^\theta, x \in (-1,1)$$
où on suppose le paramètre $\theta > -1$. On en extrait un échantillon simple $X_1, \ldots, X_n$. Déterminez l'estimateur du maximum de vraisemblance $\hat{\theta}$ de θ.

Solution :

Immédiatement, en supposant $X_1 \ldots, X_n$ dans le support de f_θ
$$\begin{aligned} L_\theta(\underline{\mathbf{X}}) &= \prod_{i=1}^n f_\theta\left(X_i\right) \\ &= \prod_{i=1}^n \frac{\theta}{1-\theta} X_i^{\frac{2\theta-1}{1-\theta}} \mathbb{I}_{0<X<1}\left(X_i\right) \\ &= \left(\frac{\theta}{1-\theta}\right)^n \left(\prod_{i=1}^n X_i\right)^{\frac{2\theta-1}{1-\theta}} \prod_{i=1}^n \mathbb{I}_{0<X<1}\left(X_i\right) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 \Leftrightarrow \theta &= \frac{1}{1-\frac{1}{n}\sum_i \log\left(X_i\right)} \end{aligned}$$

Exercice 7.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité
$$f_\theta(x) = \frac{\sqrt{\theta}}{\sqrt{2\pi}} e^{-\theta x^2/2}$$
où $\theta > 0$. Pour étudier le paramètre θ, on a effectué une suite de n expériences indépendantes qui ont donné les réalisations $x_1, \ldots, x_n$ de n v.a. $X_1, \ldots, X_n$ i.i.d. de même loi que X.

1. Déterminez un estimateur $\hat{\theta}$ du paramètre θ par la méthode du maximum de vraisemblance.
2. $\hat{\theta}$ est-il exhaustif ?
3. Calculez la moyenne et la variance de $\hat{\theta}$. Déduisez-en un estimateur $\hat{\theta}_1$ de θ non biaisé. Quelle est la variance de $\hat{\theta}_1$? Est-il convergent ?

Solution :

Comme d'habitude, en supposant que les indicatrices sont vérifiées (le support ne dépendant pas de θ),
$$\begin{aligned} L_\theta(\underline{\mathbf{X}}) &= \prod_{i=1}^n f_\theta\left(X_i\right) \\ &= \left(\frac{\theta+1}{2}\right)^n \left(\prod_{i=1}^n \left(1-|X_i|\right)\right)^\theta \mathbb{I}_{\{X \geq -1\}}\left(X_{(1)}\right) \mathbb{I}_{\{X \leq 1\}}\left(X_{(n)}\right) \\ \log L_\theta(\underline{\mathbf{X}}) &= n \log\left(\frac{\theta+1}{2}\right) + \theta \sum_{i=1}^n \log\left(1-|X_i|\right) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) &= \frac{n}{\theta+1} + \sum_{i=1}^n \log\left(1-|X_i|\right) \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 \Leftrightarrow \theta &= -\frac{n}{\sum \ln\left(1-|X_i|\right)} - 1 \end{aligned}$$

Exercice 8.

Les éléments d'une population possèdent un caractère X qui suit une loi de densité
$$f_\theta(x) = \frac{2}{\sqrt{\pi}\theta^{3/2}} x^2 e^{-x^2/\theta}$$
où $\theta > 0$. Une suite de n expériences indépendantes a donné les valeurs $x_1, \ldots, x_n$.

1. Déterminez un estimateur $\hat{\theta}$ du paramètre θ par la méthode du maximum de vraisemblance.
2. Examinez les qualités suivantes de $\hat{\theta}$: efficacité, biais, convergence, exhaustivité.

Exercice 9.

Les éléments d'une population possèdent un caractère X qui suit une loi de probabilité dont la densité est

$$f_{a,\theta}(x) = \begin{cases} \theta e^{-\theta(x-a)} \text{ si } x \geq a \\ 0 \text{ sinon} \end{cases}$$

où $\theta, a > 0$. Une suite de n expériences indépendantes a donné les valeurs $x_1, \ldots, x_n$.

Inférence sur a, en supposons que θ est connu. Proposez un estimateur $\hat{a}$ de a par la méthode du maximum de vraisemblance.