Séance 5 : Exercices récapitulatifs sur l'estimation ponctuelle

Exercice 1

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 ?
4. Notons $B^{(n)}(\theta)$ la borne de Cramer-Rao pour l'estimation de θ. Dans l'ensemble des estimateurs non biaisés de θ, on définit l'efficacité relative de $\hat{\theta}_1$ comme étant
$$e(\hat{\theta}_1) := \frac{B^{(n)}(\theta)}{Var[\hat{\theta}_1]}.$$
Calculez $e(\hat{\theta}_1)$. Qu'en concluez-vous ?
5. Soit $s^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X})^2$. Peut-on définir un estimateur $\hat{\theta}_2$ non biaisé de θ, qui soit simplement lié à s^2 ? Quelle est son efficacité relative ? De $\hat{\theta}_1$ et $\hat{\theta}_2$, lequel préconiseriez-vous ?

Solution :

1. La vraisemblance s'écrit :
$$L_\theta(\underline{\mathbf{X}}) = \prod_{i=1}^n f_\theta(X_i) = \left(\frac{\theta}{2\pi}\right)^{n/2} e^{-\theta/2 \sum_{i=1}^n X_i^2}.$$
Cette fonction est C^∞ en θ, on peut donc chercher son maximum en utilisant les conditions du premier et second ordre. Comme elle est partout positive, on passe au logarithme pour simplifier les calculs :
$$\begin{aligned}\log L_\theta(\underline{\mathbf{X}}) &= \frac{n}{2}\log\left(\frac{\theta}{2\pi}\right) - \frac{\theta}{2}\sum_{i=1}^n X_i^2 \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) &= \frac{n}{2\theta} - \frac{1}{2}\sum_{i=1}^n X_i^2 \\ \partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 &\Leftrightarrow \hat{\theta} = \frac{n}{\sum_{i=1}^n X_i^2}\end{aligned}$$
On vérifie aisément que $\hat{\theta} = \frac{n}{\sum_{i=1}^n X_i^2}$ est un maximum et l'on conclut que $\hat{\theta}$ est l'estimateur du maximum de vraisemblance pour θ.
2. On utilise ici le critère de factorisation de la vraisemblance
$$L_\theta(\underline{\mathbf{X}}) = \left(\frac{\theta}{2\pi}\right)^{n/2} e^{-\theta/2 \sum_{i=1}^n X_i^2} = h(\underline{\mathbf{X}}) g_\theta(\hat{\theta}),$$
avec
$$h(t) = \left(\frac{1}{2\pi}\right)^{n/2} \quad \text{et} \quad g_\theta(x) = \theta e^{\frac{-\theta}{2nx}}.$$
La statistique $\hat{\theta}$ est donc exhaustive.

3. On ne veut pas intégrer directement $\hat{\theta}$ pour calculer son espérance, le changement de variables impliqué étant peu pratique. Il vaut mieux commencer par déterminer la loi de $\sum_{i=1}^{n} X_i^2$, qui est plus accessible. En effet, on remarque à la forme de la densité de f_θ que les X_i suivent en fait une loi normale'déguisée", si l'on pose $\mu = 0$ et $\sigma^2 = \theta^{-1}$. Ainsi,

$$X_i \sim \mathcal{N}(0, \frac{1}{\theta})$$

de façon indépendante. Si l'on se souvient que la somme de normales centrées réduites indépendantes est distribuée selon une loi chi-deux, on est proche de l'objectif. Il suffit de considérer les variables $\sqrt{\theta} X_i$ qui sont centrées réduites, et il vient que

$$Z = \theta \sum_{i=1}^{n} X_i^2 \sim \chi_n^2,$$

dont on connaît la densité. On a ainsi

$$\mathbb{E}\left[\hat{\theta}\right] = \mathbb{E}\left[\frac{\mathrm{n}\theta}{\mathrm{Z}}\right] = \mathrm{n}\theta\mathbb{E}\left[\mathrm{Z}^{-1}\right].$$

On va maintenant écrire l'espérance comme une intégrale et bricoler un peu à l'intérieur pour arriver au résultat.

$$\mathbb{E}\left[\frac{1}{\mathrm{Z}}\right] = \int_{\mathbb{R}^+} \frac{1}{\mathrm{z}} \frac{1}{2^{\mathrm{n}/2}\Gamma(\mathrm{n}/2)} \mathrm{z}^{\frac{\mathrm{n}}{2}-1} \mathrm{e}^{-\mathrm{z}/2} \mathrm{dz} = \int_{\mathbb{R}^+} \frac{1}{2^{\mathrm{n}/2}\Gamma(\mathrm{n}/2)} \mathrm{z}^{\frac{\mathrm{n}-2}{2}-1} \mathrm{e}^{-\mathrm{z}/2} \mathrm{dz}.$$

L'idée est de faire apparaître sous l'intégrale la densité d'une loi chi-deux en ajustant le nombre de degrés de libertés au fait que la puissance de z dans l'intégrande est diminuée de un. Il faut donc mettre les bonnes constantes. Entre autres, faire sortir un deux au dénominateur, ainsi qu'un facteur $(n-2)/2$ grâce aux propriétés de la fonction Γ. L'intégale restante est alors celle d'une densité sur son domaine et vaut donc un. D'où

$$\mathbb{E}\left[\frac{1}{\mathrm{Z}}\right] = \frac{1}{\mathrm{n}-2}.$$

On en déduit l'espérance de $\hat{\theta}$

$$\mathbb{E}\left[\hat{\theta}\right] = \mathrm{n}\theta\mathbb{E}\left[\frac{1}{\mathrm{Z}}\right] = \frac{\mathrm{n}}{\mathrm{n}-2}\theta.$$

Le raisonnement et les calculs sont similaires pour la variance. Connaissant l'espérance de $\hat{\theta}$, il nous reste pour obtenir la variance à calculer celle de son carré :

$$\hat{\theta}^2 = \frac{n^2\theta^2}{Z^2}.$$

On a alors, à l'aide des mêmes idées

$$\mathbb{E}\left[\hat{\theta}^2\right] = \frac{\mathrm{n}^2\theta^2}{(\mathrm{n}-2)(\mathrm{n}-4)}$$

et

$$\mathrm{Var}\left(\hat{\theta}\right) = \cdots = \frac{2n^2\theta^2}{(n-2)^2(n-4)}.$$

On note au passage que $\hat{\theta}$ est asymptotiquement sans biais et convergent. On en déduit un estimateur sans biais par une simple transformation linéaire :

$$\hat{\theta}_1 = \frac{n-2}{n}\hat{\theta} \Rightarrow \mathbb{E}\left[\hat{\theta}_1\right] = \theta.$$

On déduit aisément la variance du nouvel estimateur

$$\mathrm{Var}\left(\hat{\theta}_1\right) = \frac{2\theta^2}{n-4}.$$

qui montre que $\hat{\theta}_1$ est également convergent.

4. Comme la log-vraisemblance du modèle est C^∞, ne nous privons pas pour dériver deux fois avant de prendre l'espérance.

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

et donc,

$$I(\theta) = -\mathrm{I\!E}\left[\partial_\theta^2 \log \mathrm{L}_\theta(\underline{\mathbf{X}})\right] = \frac{\mathrm{n}}{2\theta^2} \quad \text{et} \quad \mathrm{B}^{(\mathrm{n})}(\theta) = \frac{2\theta^2}{\mathrm{n}}.$$

Remarquons que ces deux quantités peuvent bien sûr se calculer "à l'ancienne" en prenant l'espérance du carré de la première dérivée de la log-vraisemblance. Le fait que $I(\theta)$ soit un multiple de n est rassurant. On peut alors calculer l'efficacité relative de $\hat{\theta}_1$

$$e(\hat{\theta}_1) = \frac{2\theta^2}{n}\frac{n-4}{2\theta^2} = \frac{n-4}{n}.$$

On remarque que $e < 1$, donc que $\hat{\theta}_1$ n'est pas efficace, mais que $e \to 1$ et donc que $\hat{\theta}_1$ est asymptotiquement efficace.

5. Cette question fait appel à un résultat du cours. Le lemme de Fisher nous dit que dans un modèle i.i.d. gaussien, $\bar{X}$ et s^2 sont indépendants, et que $ns^2/\sigma^2 \sim \chi^2_{n-1}$, où n est la taille de l'échantillon. Comme notre échantillon simple est gaussien ; ce résultat s'applique, et les rappels sur la loi chi-deux nous donnent

$$\mathrm{I\!E}\left[\frac{\mathrm{ns}^2}{\sigma^2}\right] = \mathrm{I\!E}\left[\mathrm{ns}^2\theta\right] = (\mathrm{n}-1),$$

i.e.

$$\mathrm{I\!E}\left[\mathrm{s}^2\right] = (\mathrm{n}-1)/\mathrm{n}\theta.$$

Pour définir simplement un estimateur à partir de s^2, il faut considérer son inverse $1/s^2$. Calculons l'espérance de cette statistique, dans l'espoir qu'elle soit une fonction linéaire de θ.

$$\mathrm{I\!E}\left[1/\mathrm{s}^2\right] = \mathrm{I\!E}\left[\frac{\mathrm{n}\theta}{\mathrm{n}\theta\mathrm{s}^2}\right] = \mathrm{n}\theta\mathrm{I\!E}\left[\frac{1}{\mathrm{n}\theta\mathrm{s}^2}\right] = \cdots = \frac{\mathrm{n}\theta}{\mathrm{n}-3}.$$

Ainsi, la statistique $\hat{\theta}_2 = \frac{n-3}{ns^2}$ st un estimateur sans biais pour θ. Calculons sa variance, et pour ce faire, l'espérance de son carré :

$$\begin{aligned}
\mathrm{I\!E}\left[\frac{(\mathrm{n}-3)^2}{\mathrm{n}^2\mathrm{s}^4}\right] &= (n-3)^2\theta^2\mathrm{I\!E}\left[\frac{1}{\mathrm{n}^2\theta^2\mathrm{s}^4}\right] \\
&= (n-3)^2\theta^2\int_{\mathrm{I\!R}^+}\frac{1}{z^2}\frac{1}{2^{(n-1)/2}\Gamma((n-1)/2)}z^{\frac{(n-1)}{2}-1}e^{-z/2}dz \\
&= (n-3)^2\theta^2\frac{1}{(n-3)(n-5)} \\
&= \frac{(n-3)\theta^2}{(n-5)}
\end{aligned}$$

et donc

$$\mathrm{Var}\left(\hat{\theta}_2\right) = \frac{(n-3)\theta^2}{(n-5)} - \theta^2 = \frac{2\theta^2}{n-5}.$$

Cet estimateur est donc également convergent, et son efficacité relative est de

$$e(\hat{\theta}_2) = \frac{n-5}{n} < \frac{n-4}{n} = e(\hat{\theta}_1).$$

On lui préférera donc $\hat{\theta}_1$.

Exercice 2

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.

Solution :
On écrit la vraisemblance du modèle grâce à l'hypothèse i.i.d. et on obtient :

$$\begin{aligned}
L_\theta(\underline{\mathbf{X}}) &= \prod_{i=1}^{n} \frac{2}{\sqrt{\pi}\theta^{3/2}} X_i^2 e^{-X_i^2/\theta} \\
&= \frac{2^n}{\pi^{n/2}\theta^{3n/2}} \prod_{i=1}^{n} X_i^2 e^{-\frac{1}{\theta}\sum_i X_i^2} \\
\log L_\theta(\underline{\mathbf{X}}) &= C - \frac{3n}{2}\log\theta - 2\sum_i \log X_i - \frac{1}{\theta}\sum_{=1}^{n} X_i^2 \\
\partial_\theta \log L_\theta(\underline{\mathbf{X}}) &= -\frac{3n}{2\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} X_i^2 \\
\partial_\theta \log L_\theta(\underline{\mathbf{X}}) = 0 \quad &\Leftrightarrow \quad \hat{\theta} = \frac{2}{3n}\sum_{i=1}^{n} X_i^2.
\end{aligned}$$

Clairement, la vraisemblance (ou log-vraisemblance) peut se factoriser de la bonne manière et l'estimateur du maximum de vraisemblance $\hat{\theta}$ est donc exhaustif. Tentons de factoriser la dérivée de la log-vraisemblance :

$$\partial_\theta \log L_\theta(\underline{\mathbf{X}}) = -\frac{3n}{2\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} X_i^2 = \frac{3n}{2\theta^2}\left(\frac{2}{3n}\sum_{i=1}^{n} X_i^2 - \theta\right).$$

Ainsi, d'après le critère de factorisation, $\hat{\theta}$ est efficace (donc sans biais) pour θ. Reste à juger de sa convergence, en calculant sa variance. Pour ce faire, il nous manque l'espérance de son carré. On note pour ce faire que les variables $Y_i = X_i^2$ ont pour densité

$$g_\theta(y) = f_\theta(\sqrt{y})\frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{\pi}\theta^{3/2}} y^{1/2} e^{-y/\theta} \qquad y > 0,$$

où l'on reconnaît une loi Gamma de paramètres $a = 1/\theta$ et $b = 3/2$. La somme de n variables indépendantes de telle loi donne une variable Gamma de paramètres $a = 1/\theta$ et $b = 3n/2$. On connaît donc son moment d'ordre deux

$$\mathbb{E}\left[\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{Y}_\mathrm{i}^2\right] = \frac{\Gamma(3\mathrm{n}/2+2)}{\Gamma(3\mathrm{n}/2)}\theta^2 = (3\mathrm{n}/2+1)(3\mathrm{n}/2)\theta^2.$$

On en déduit la quantité recherchée

$$\mathbb{E}\left[\hat{\theta}^2\right] = \frac{4}{9\mathrm{n}^2}(3\mathrm{n}/2+1)(3\mathrm{n}/2)\theta^2.$$

Il est alors immédiat que

$$\mathrm{Var}\left(\hat{\theta}\right) = \frac{2}{3n}\theta^2.$$

L'estimateur est donc également convergent.

Exercice 3

En 1897, l'économiste suisse Vilfredo Pareto (1848-1923), professeur d'économie politique à l'université de Lausanne, eut l'idée de modéliser la loi des revenus en postulant que le nombre relatif de personnes dont le revenu dépasse une valeur x est inversément proportionnel à une puissance de x. La définition suivante fut adoptée : une variable aléatoire X, absolument continue, suit une loi de Pareto de paramètres α et θ si sa densité est donnée par

$$f_{\alpha,\theta}(x) = kx^{-\alpha} I_{[x\geq\theta]};$$

où $\theta > 0$ et $\alpha > 1$. Cet exercice vous propose d'étudier différentes caractéristiques de cette loi sur le plan probabiliste et statistique.

1. Analyse probabiliste

   a) Déterminez la constante de normalisation k.
   A ce stade vous êtes tous capable de calculer une intégrale définie... Réponse : $k = (\alpha-1)\theta^{\alpha-1}$.

   b) Ecrivez la fonction de répartition F_X de X.
   Même remarque pour une primitive... Réponse :

   $$F^X_{\alpha,\theta}(x) = \left(1 - (\theta/x)^{\alpha-1}\right) \mathbb{I}_{[x\geq 0]}.$$

   c) Calculez le moment non centré d'ordre $s(s \in \mathbb{N}_0)$. Précisez ses conditions d'existence. Déduisez-en $E[X]$ et $Var[X]$.
   De manière générale,

   $$\mathbb{E}\left[\mathrm{X^s}\right] = \int_\theta^{+\infty} \frac{(\alpha-1)\theta^{\alpha-1}}{\mathrm{x}^{\alpha-\mathrm{s}}}\mathrm{dx} = (\alpha-1)\theta^{\alpha-1}\left[\frac{\mathrm{x}^{\mathrm{s}-\alpha+1}}{\mathrm{s}-\alpha+1}\right]_\theta^{+\infty}.$$

   La condition d'existence de l'intégrale est $s-\alpha+1<0$, c'est-à-dire $s<\alpha-1$. On a alors

   $$\mathbb{E}\left[\mathrm{X^s}\right] = \frac{(\alpha-1)\theta^{\mathrm{s}}}{\alpha-1-\mathrm{s}}.$$

   Si $\alpha > 2$, l'espérance existe et vaut donc

   $$\mathbb{E}\left[\mathrm{X}\right] = \frac{(\alpha-1)\theta}{\alpha-2}.$$

   Si $\alpha > 3$, le moment d'ordre deux existe, et permet de calculer la variance

   $$\mathrm{Var}\,(X) = \mathbb{E}\left[\mathrm{X}^2\right] - (\mathbb{E}\left[\mathrm{X}\right])^2 = \frac{(\alpha-1)\theta^2}{\alpha-3} - \frac{(\alpha-1)^2\theta^2}{(\alpha-2)^2} = \frac{(\alpha-1)\theta^2}{(\alpha-3)(\alpha-2)^2}.$$

   d) Donnez la loi de la variable aléatoire U définie par $U = (\alpha-1)\ln(X/\theta)$.

   Sur le domaine de X, ce changement de variable est admissible et on utilise la formule habituelle

   $$U = G(X) \Rightarrow G'(X) = \frac{\alpha-1}{X}\ ,\ X = G^{-1}(U)$$
   $$\text{avec } G^{-1}(u) = \theta e^{u/(\alpha-1)}\ ,\ X \geq \theta \Leftrightarrow U \geq 0$$

   et donc

   $$f_U(u) = f_X(G^{-1}(u))\frac{1}{G'(G^{-1}(u))} = \frac{(\alpha-1)\theta^{\alpha-1}}{(\theta e^{u/(\alpha-1)})^\alpha}\frac{\theta e^{u/(\alpha-1)}}{\alpha-1} = e^{-u}.$$

   Ainsi, en combinant cette expression avec le domaine de u, on reconnaît la densité d'une loi exponentielle de paramètre 1 (ou $\Gamma(1,1)$).

e) Soient n variables aléatoires indépendantes $X_1, \ldots, X_n$ suivant toutes la même loi de Pareto de paramètres α et θ. Déterminez la loi de $Z = X_{(1)}$. Identifiez cette loi

Pour déterminer la loi du minimum de n v.a. i.i.d., on considère sa fonction de répartition

$$\mathbb{P}\left[Z \leq z\right] = 1 - \mathbb{P}\left[Z > z\right] = 1 - \prod_{i=1}^{n} \mathbb{P}\left[X_i > z\right] = 1 - (1 - F^X(z))^n,$$

et on réinjecte l'expression obtenue pour F^X à la question (b)

$$F^Z(z) = 1 - \left((\theta/z)^{\alpha-1}\right)^n = 1 - (\theta/z)^{n(\alpha-1)},$$

qui est de la forme $F(z) = 1 - (\theta/z)^{a-1}$ avec $a = n(\alpha - 1) - 1$. C'est donc une loi de Pareto de paramètres $n(\alpha - 1) - 1$ et θ.

2. Inférence sur α. Supposons que θ est connu. Soient les variables aléatoires $X_1, \ldots, X_n$ i.i.d. de loi de Pareto de paramètres α et θ.

a) Donnez une statistique exhaustive pour α.
Factorisons la vraisemblance. Comme θ est connu, on suppose que l'indicatrice vaut 1, sans quoi la vraisemblance serait identiquement nulle et toute statistique serait exhaustive.

$$L_\alpha(\underline{\mathbf{X}}) = (\alpha - 1)^n \theta^{n(\alpha-1)} \prod_{i=1}^{n} X_i^{-\alpha} = (\alpha - 1)^n \theta^{n(\alpha-1)} \left(\prod_{i=1}^{n} X_i\right)^{-\alpha}.$$

Ainsi, le produit des observations est exhaustif.

b) Proposez un estimateur $\hat{\alpha}$ de α par la méthode du maximum de vraisemblance.
On trouve

$$\begin{aligned}
\log L_\alpha(\underline{\mathbf{X}}) &= n \log(\alpha - 1) + n(\alpha - 1) \log \theta - \alpha \sum_{i=1}^{n} \log X_i \\
\partial_\alpha \log L &= \frac{n}{\alpha - 1} + n \log \theta - \sum_{i=1}^{n} \log X_i = \frac{n}{\alpha - 1} - \sum_{i=1}^{n} \log\left(\frac{X_i}{\theta}\right) \\
\partial_\alpha \log L = 0 &\Leftrightarrow \frac{n}{\hat{\alpha} - 1} = \sum_{i=1}^{n} \log\left(\frac{X_i}{\theta}\right) \\
&\Leftrightarrow \hat{\alpha} = 1 + \frac{n}{\sum_{i=1}^{n} \log\left(\frac{X_i}{\theta}\right)}
\end{aligned}$$

On vérifie que c'est bien un maximum en remarquant que

$$\partial_\alpha^2 \log L_\alpha(\underline{\mathbf{X}}) = -\frac{n}{(\alpha - 1)^2} < 0.$$

c) $\hat{\alpha}$ est-il sans biais pour α ? Si non, construisez un estimateur sans biais $\hat{\alpha}_1$ pour α et vérifiez s'il est convergent et efficace.
Il faut calculer l'espérance de $\hat{\alpha}$. On utilise pour cela de résultat montré précédemment sur la loi $(\alpha - 1)\log(X/\theta)$. En effet, en posant $Y_i = (\alpha - 1)\log(X_i/\theta)$, on a

$$Y_i \sim \Gamma(1, 1) \quad \text{et} \quad \hat{\alpha} = 1 + \frac{n(\alpha - 1)}{\sum_{i=1}^{n} Y_i}.$$

Les Y_i étant indépendants, leur somme W suit une loi $\Gamma(1, n)$. Ainsi,

$$\hat{\alpha} = 1 + \frac{(\alpha - 1)n}{W} \Rightarrow \mathbb{E}\left[\hat{\alpha}\right] = 1 + \mathrm{n}(\alpha - 1)\mathbb{E}\left[\mathrm{W}^{-1}\right].$$

Cette dernière espérance est facile à calculer comme une intégrale, en faisant apparaître dans l'intégrande une densité $\Gamma(1, n-1)$ et on trouve $\mathbb{E}\left[\mathrm{W}^{-1}\right] = (\mathrm{n}-1)^{-1}$. Ainsi,

$$\mathbb{E}[\hat{\alpha}] = 1 + \frac{\mathrm{n}(\alpha - 1)}{\mathrm{n} - 1} = \frac{\mathrm{n}\alpha - 1}{\mathrm{n} - 1}.$$

L'estimateur maximum de vraisemblance est donc biaisé (mais asymptotiquemen sans biais). Le biais est cependant aisément corrigé car l'espérance de $\hat{\alpha}$ est une fonction linéaire de α. Ainsi,

$$\hat{\alpha}_1 = \frac{(n-1)\hat{\alpha} + 1}{n}$$

est sans biais pour α. Pour juger de sa convergence et de son efficacité, on calcule sa variance. notons qu'avec les notations précédentes

$$\hat{\alpha}_1 = 1 + \frac{(n-1)(\alpha - 1)}{W}$$

et il nous faut l'espérance du carré de $\hat{\alpha}_1$, son espérance étant α. On trouve aisément que (en développant le carré et en calculant $\mathbb{E}\left[\mathrm{W}^{-1}\right]$ et $\mathbb{E}\left[\mathrm{W}^{-2}\right]$) que

$$\mathbb{E}\left[\hat{\alpha}_1^2\right] = \cdots = 2\alpha - 1 + \frac{(\mathrm{n}-1)^2(\alpha-1)^2}{(\mathrm{n}-1)(\mathrm{n}-2)}.$$

On en déduit alors la variance de l'estimateur :

$$\mathrm{Var}\left(\hat{\alpha}_1\right) = \mathbb{E}\left[\hat{\alpha}_1^2\right] - \left(\mathbb{E}\left[\hat{\alpha}_1\right]\right)^2 = \cdots = \frac{(\alpha-1)^2}{\mathrm{n}-2}.$$

On en déduit que l'estimateur est convergent. Calculons la borne de Cramer-Rao pour juger de son efficacité :

$$BCR = \frac{1}{I(\alpha)} \quad \text{avec} \quad I(\alpha) = -\mathbb{E}\left[\partial_\alpha^2 \log \mathrm{L}\right] = \frac{\mathrm{n}}{(\alpha-1)^2}.$$

Ainsi, $\mathrm{Var}\left(\hat{\alpha}_1\right) > BCR$ et l'estimateur n'est pas efficace.

3. Inférence sur θ. Supposons que α est connu. Soient les variables aléatoires $X_1, \ldots, X_n$ i.i.d. de loi de Pareto de paramètres α et θ.

   a) Donnez une statistique exhaustive pour θ.
   Bien entendu, la vraisemblance n'a pas foncièrement changé d'une fois à l'autre, mais il convient ici de ne pas oublier les indicatrices, qui dépendent de θ. Ainsi,

   $$L_\theta(\underline{\mathbf{X}}) = (\alpha-1)^n \theta^{n(\alpha-1)} \prod_{i=1}^n X_i^{-\alpha} \mathbb{I}_{[\theta,+\infty[}(X_i) = (\alpha-1)^n \theta^{n(\alpha-1)} \left(\prod_{i=1}^n X_i^{-\alpha}\right) \mathbb{I}_{[\theta,+\infty[}(X_{(1)}).$$

   En factorisant, on constate que $X_{(1)}$ est une statistique exhaustive.

   b) Proposez un estimateur $\hat{\theta}$ de θ par la méthode du maximum de vraisemblance.
   Il n'est pas question de passer au logarithme directement du fait de la présence de l'indicatrice, ni de dériver par rapport à θ. On note que la vraisemblance dépend de θ à travers l'indicatrice d'une part et d'autre part à travers le terme $\theta^{n(\alpha-1)}$, qui est une fonction croissante de θ puisque $\alpha > 1$. Ainsi, comme il existe des valeurs de θ pour lesquelles la vraisemblance n'est pas nulle, le maximum de vraisemblance est obtenu lorsque l'indicatrice vaut 1 ; ce qui correspond à $\theta \leq X_{(1)}$, et par croissance, le maximum est atteint pour la plus grande valeur de θ dans cet ensemble, soit

   $$\hat{\theta} = X_{(1)}.$$

c) $\hat{\theta}$ est-il convergent ?
Pour juger des caractéristiques d'optimalité de cet estimateur, on doit connaître son espérance et sa variance, mais les résultats obtenus dans la première partie de l'exercice nous garantissent que la loi de $\hat{\theta}$ est une loi de Pareto de paramètres θ et $n(\alpha-1)-1$. Ainsi,

$$\mathbb{E}\left[\hat{\theta}\right] = \frac{\mathrm{n}(\alpha-1)\theta}{\mathrm{n}(\alpha-1)-1} \text{ et } \mathrm{Var}\left(\hat{\theta}\right) = \frac{\mathrm{n}(\alpha-1)\theta^2}{(\mathrm{n}(\alpha-1)-2)(\mathrm{n}(\alpha-1)-1)^2}$$

pour peu que ces quantités existent, ce qui est vrai pour n suffisament grand. Cet estimateur est ainsi biaisé, mais asymptotiquement sans biais, et convergent.

Exercice 4

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 $x_1, \ldots, x_n$.

1. Inférence sur a. Supposons que θ est connu.

a) Proposez un estimateur $\hat{a}$ de a par la méthode du maximum de vraisemblance.
La vraisemblance est donnée par

$$L_{a,\theta}(\underline{\mathbf{X}}) = \theta^n e^{-\theta\sum_i X_i} e^{-na} \mathbb{I}_{[a,+\infty[}(X_{(1)}).$$

Un raisonnement identique à celui fait plus haut permet de conclure

$$\hat{a} = X_{(1)}$$

b) $\hat{a}$ est-il sans biais, convergent, exhaustif ?
Par factorisation immédiate, $\hat{a}$ est exhaustif. Calculons son espérance et sa variance. Pour ce faire, il faut obtenir la loi de $X_{(1)}$. On obtient aisément

$$F^{X_{(1)}}(x) = 1-(1-F^X(x))^n = \left(1-e^{-n\theta(x-a)}\right)\mathbb{I}_{[a,+\infty[}(x).$$

La densité est alors donnée par

$$f^{X_{(1)}}(x) = n\theta e^{-n\theta(x-a)}\mathbb{I}_{[a,+\infty[}(x).$$

Restent deux calculs faciles :

$$\begin{aligned}
\mathbb{E}\left[\hat{\mathrm{a}}\right] &= \int_a^\infty x n\theta e^{-n\theta(x-a)}dx \\
&= \left[xe^{-n\theta(x-a)}\right]_a^\infty + \int_a^\infty e^{-n\theta(x-a)}dx \\
&= a + \frac{1}{n\theta} \\
\mathbb{E}\left[\hat{\mathrm{a}}^2\right] &= \int_a^\infty x^2 n\theta e^{-n\theta(x-a)}dx \\
&= \left[x^2 e^{-n\theta(x-a)}\right]_a^\infty + \int_a^\infty 2xe^{-n\theta(x-a)}dx \\
&= a^2 + \frac{2}{n\theta}\mathbb{E}\left[\hat{\mathrm{a}}\right] \\
&= a^2 + \frac{2a}{n\theta} + \frac{2}{n^2\theta^2} \\
\mathrm{Var}\left(\hat{a}\right) &= a^2 + \frac{2a}{n\theta} + \frac{2}{n^2\theta^2} - a^2 - \frac{1}{n^2\theta^2} - \frac{2a}{n\theta} = \frac{1}{n^2\theta^2}.
\end{aligned}$$

L'estimateur maximum de vraisemblance est donc biaisé mais asymptotiquement sans biais, et convergent.

2. Inférence sur θ. Supposons que a est connu.

   a) Proposez un estimateur $\hat{\theta}$ de θ par la méthode du maximum de vraisemblance.
   On reprend la vraisemblance, mais cette fois-ci, a étant connu, on peut supposer les indicatrices toutes égales à un et s'en passer dans l'expression de L. On trouve alors

$$\partial_\theta \log L = \frac{n}{\theta} - \sum_{i=1}^{n} n(X_i - a).$$

   Ainsi (et car la dérivée seconde est partout négative), l'estimateur maximum de vraisemblance est

$$\hat{\theta} = \frac{n}{\sum_{i=1}^{n}(X_i - a)}$$

   b) Cherchez la densité de probabilité de la variable $Y = \theta \sum_{i=1}^{n}(X_i - a)$, où les X_i sont i.i.d. de même loi que X.
   On commence par la loi de $Y_i = \theta(X_i - a)$, dont la densité est de façon immédiate

$$f_Y(y) = e^{-y}\mathbb{I}_{[0,+\infty[}(y).$$

   Ainsi, les Y_i sont de loi $\Gamma(1,1)$. Y étant la somme des Y_i, par indépendance, sa loi est $\Gamma(1,n)$.

   c) $\hat{\theta}$ est-il non biaisé ? Si non, construisez un estimateur non biaisé $\hat{\theta}_1$ de θ et vérifiez s'il est convergent, efficace et exhaustif.
   On calcule l'espérance de $\hat{\theta}$ en se servant du résultat précédent (déjà vu plus haut)

$$\mathbb{E}\left[\hat{\theta}\right] = \mathrm{n}\theta\mathbb{E}\left[\mathrm{Y}^{-1}\right] = \frac{\mathrm{n}\theta}{\mathrm{n}-1}$$

   et l'on en déduit l'estimateur maximum de vraisemblance est biaisé mais asymptotiquement sans biais. On peut alors définir

$$\hat{\theta}_1 = \frac{(n-1)\hat{\theta}}{n}$$

   qui sera sans biais. On calcule l'espérance de son carré et on trouve

$$\mathbb{E}\left[\hat{\theta}_1^2\right] = \frac{(\mathrm{n}-1)\theta}{\mathrm{n}-2},$$

   d'où la variance

$$\mathrm{Var}\left(\hat{\theta}_1\right) = \frac{(n-1)\theta^2}{n-2} - \theta^2 = \frac{\theta^2}{n-2},$$

   ainsi $\hat{\theta}_1$ est convergent. La borne de CR est θ^2/n ce qui implique que l'estimateur $\hat{\theta}_1$ n'est pas efficace.