

Séance 2 : Estimateurs convergents, non biaisés et exhaustifs.

Exercice 1

Soient les variables aléatoires X_i ($i = 1, \dots, n$) i.i.d. Montrez que $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ est un estimateur non biaisé de σ^2 , où $\sigma^2 = \text{Var}[X_1]$.

Solution :

On utilise ici le fait que, pour toute valeur de i ,

$$\mathbb{E}[X_i] = \mu \text{ et } \text{Var}(X_i) = \sigma^2.$$

Les calculs sont immédiats : On trouve

$$\begin{aligned} \mathbb{E}[X_i^2] &= \text{Var}(X_i) + (\mathbb{E}[X_i])^2 = \sigma^2 + \mu^2 \\ \mathbb{E}[\bar{X}^2] &= \frac{1}{n^2} \mathbb{E}\left[\left(\sum_{i=1}^n X_i\right)\left(\sum_{j=1}^n X_j\right)\right] = \frac{1}{n^2} \sum_{i \neq j} \mathbb{E}[X_i X_j] + \frac{1}{n^2} \sum_{i=1}^n \mathbb{E}[X_i^2] \\ &= \frac{1}{n^2} n(n-1)\mu^2 + \frac{n}{n^2}(\sigma^2 + \mu^2) = \mu^2 + \frac{\sigma^2}{n} \\ \mathbb{E}[\bar{X} X_i] &= \frac{1}{n} \sum_{j=1}^n \mathbb{E}[X_j X_i] = \frac{n-1}{n}\mu^2 + \frac{1}{n}(\sigma^2 + \mu^2) = \mu^2 + \frac{\sigma^2}{n} \\ \mathbb{E}[S^2] &= \frac{1}{n-1} \sum_{i=1}^n \mathbb{E}[(X_i^2 - 2\bar{X}X_i + \bar{X}^2)] \\ &= \frac{n}{n-1}(\sigma^2 + \mu^2) - \frac{2n}{n-1}\left(\mu^2 + \frac{\sigma^2}{n}\right) + \frac{n}{n-1}\left(\mu^2 + \frac{\sigma^2}{n}\right) \\ &= \sigma^2 \end{aligned}$$

Remarquons que ce long développement peut être considérablement raccourci en voyant que

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n X_i^2 - \frac{n}{n-1} \bar{X}^2.$$

Exercice 2

Soient $s_1^2, s_2^2, \dots, s_k^2$ les k variances-échantillons obtenues à partir de k échantillons aléatoires simples indépendants de tailles respectives n_1, n_2, \dots, n_k .

1. $U = \frac{n_1 s_1^2 + n_2 s_2^2 + \dots + n_k s_k^2}{n_1 + n_2 + \dots + n_k}$ est-il un estimateur non biaisé de σ^2 ?
2. Si ce n'est pas le cas, comment le biais peut-il être enlevé ?

Solution :

Des calculs immédiats (rappelez-vous que $\mathbb{E}[s^2] = (n-1)\sigma^2/n$ pour un échantillon aléatoire simple de taille n) montrent que

$$\mathbb{E}[U] = \frac{n_1 \frac{(n_1-1)\sigma^2}{n_1} + n_2 \frac{(n_2-1)\sigma^2}{n_2} + \dots + n_k \frac{(n_k-1)\sigma^2}{n_k}}{n_1 + n_2 + \dots + n_k} = \sigma^2 - \frac{k\sigma^2}{n_1 + n_2 + \dots + n_k}.$$

Cet estimateur est clairement biaisé. Un estimateur non biaisé de σ^2 est donné par

$$\tilde{U} = \frac{U}{1 - \frac{k}{n_1 + n_2 + \dots + n_k}}.$$

Exercice 3

Considérons une population distribuée selon une loi de Poisson de paramètre $\lambda > 0$. On extrait de cette population une observation X_1 . Si notre but est d'estimer $e^{-3\lambda}$, que pouvez-vous dire de la statistique $(-2)^{X_1}$?

Solution :

Pour rappel, la loi de probabilité d'une variable aléatoire $X \sim \mathcal{P}(\lambda)$ est donné par, pour k naturel,

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Ainsi,

$$\mathbb{E}[(-2)^{X_1}] = \sum_{k=1}^{\infty} (-2)^k \mathbb{P}[X = k] = \sum_{k=1}^{\infty} e^{-\lambda} \frac{(-2\lambda)^k}{k!} = e^{-3\lambda}.$$

Cet "estimateur" est donc sans biais. Malgré cette propriété agréable, il serait extrêmement déraisonnable de prendre $(-2)^{X_1}$ comme estimateur pour la simple et bonne raison que ce n'en est pas un. En effet, un estimateur est une statistique à valeurs dans l'espace des paramètres (ici, l'ensemble des valeurs de $e^{-3\lambda}$, c'est-à-dire \mathbb{R}_0^+). Or, $(-2)^{X_1}$ peut prendre des valeurs négatives. Ce n'est donc pas un estimateur, mais une statistique sans biais.

Exercice 4

Soient les variables aléatoires indépendantes et de Bernoulli X_1, \dots, X_n . Nous avons donc, pour tout $i = 1, \dots, n$,

$$\begin{cases} P[X_i = 1] = p \\ P[X_i = 0] = 1 - p; \end{cases}$$

1. Montrez, à partir de la définition d'exhaustivité, que $T = \sum_{i=1}^n X_i$ est une statistique exhaustive pour le paramètre p .
2. Pour $n = 3$, la statistique $\tilde{T} = e^{X_1+X_2+X_3}$ est-elle aussi exhaustive ? Même question pour $S = X_1 + 2X_2 + X_3$.

Solution :

Par définition, une statistique $T(\underline{\mathbf{X}})$ est exhaustive si

$$\mathbb{P}[\underline{\mathbf{X}} = (x_1, \dots, x_n) | T(\underline{\mathbf{X}}) = y]$$

est indépendant de la valeur du paramètre. Calculons donc

$$\begin{aligned} \mathbb{P}[\underline{\mathbf{X}} = \underline{\mathbf{x}} | T(\underline{\mathbf{X}}) = y] &= \mathbb{P}[(X_1, \dots, X_n) = (x_1, \dots, x_n) | T(\underline{\mathbf{X}}) = y] \\ &= \mathbb{P}\left[X_1 = x_1 \text{ et } X_2 = x_2 \text{ et } \dots \text{ et } X_n = x_n \left| \sum_{i=1}^n X_i = y \right.\right] \\ &= \frac{\mathbb{P}[X_1 = x_1 \text{ et } X_2 = x_2 \text{ et } \dots \text{ et } X_n = x_n \text{ et } \sum_{i=1}^n X_i = y]}{\mathbb{P}[\sum_{i=1}^n X_i = y]} \\ &= \frac{\mathbb{P}[X_2 = x_2 \text{ et } \dots \text{ et } X_n = x_n \text{ et } X_1 = y - x_2 - x_3 - \dots - x_n]}{\mathbb{P}[\sum_{i=1}^n X_i = y]} \end{aligned}$$

Par indépendance, on obtient alors (en se rappelant que la somme de n v.a. de Bernoulli est une binomiale de paramètres (n, p) , pour $x_1 = y - x_2 - \dots - x_n$ (le cas $x_1 \neq \dots$ vaut trivialement 0),

$$\begin{aligned} \mathbb{P}[\underline{\mathbf{X}} = \underline{\mathbf{x}} | T(\underline{\mathbf{X}}) = y] &= \frac{\mathbb{P}[X_2 = x_2] \mathbb{P}[X_3 = x_3] \dots \mathbb{P}[X_n = x_n] \mathbb{P}[X_1 = y - x_2 - x_3 - \dots - x_n]}{\mathbb{P}[\sum_{i=1}^n X_i = y]} \\ &= \frac{p^{x_2}(1-p)^{1-x_2} \dots p^{x_n}(1-p)^{1-x_n} p^{y-x_2-\dots-x_n}(1-p)^{1-(y-x_2-\dots-x_n)}}{\binom{n}{y} p^y (1-p)^{(n-y)}} \\ &= \frac{1}{\binom{n}{y}} \end{aligned}$$

Cette quantité est clairement indépendante de p . La statistique $T(\underline{\mathbf{X}})$ est donc exhaustive. Remarquons que le calcul effectué ci-dessus revient exactement à calculer la vraisemblance conditionnelle et à vérifier que celle-ci ne dépend de p que via une fonction de y (apparaissant dans la condition). Pour $n = 3$, la statistique \hat{T} est en bijection avec T . Elle est donc exhaustive. Ce n'est pas le cas de S .

Exercice 5

Soient les variables aléatoires X_1, \dots, X_n i.i.d. $\mathcal{P}(\lambda)$. Montrez, à partir de la définition d'exhaustivité, que $T = \sum_{i=1}^n X_i$ est une statistique exhaustive pour le paramètre λ .

Solution :

On applique ici la même construction qu'à l'exercice précédent...et on obtient (la somme de n v.a. de Poisson de paramètre λ est une v.a. de Poisson de paramètre $n\lambda$) :

$$\begin{aligned} \mathbb{P}[\underline{\mathbf{X}} = \underline{\mathbf{x}} | T(\underline{\mathbf{X}}) = y] &= \frac{\mathbb{P}[X_2 = x_2] \mathbb{P}[X_3 = x_3] \dots \mathbb{P}[X_n = x_n] \mathbb{P}[X_1 = y - x_2 - x_3 - \dots - x_n]}{\mathbb{P}[\sum_{i=1}^n X_i = y]} \\ &= \frac{e^{-\lambda} \lambda^{x_2} (x_2!)^{-1} \dots e^{-\lambda} \lambda^{x_n} (x_n!)^{-1} e^{-\lambda} \lambda^{y-x_2-\dots-x_n} ((y-x_2-\dots-x_n)!)^{-1}}{e^{-n\lambda} \frac{(n\lambda)^y}{y!}} \\ &= \frac{(x_2!)^{-1} \dots (x_n!)^{-1} ((y-x_2-\dots-x_n)!)^{-1}}{\frac{n^y}{y!}} \end{aligned}$$

qui est indépendant de la valeur de λ . La statistique est bien exhaustive.

Exercice 6

Soit une population $\mathcal{N}(\mu, \sigma^2)$. On extrait de cette population un échantillon aléatoire simple X_1, \dots, X_n . Construisez à partir de cet échantillon une statistique exhaustive

1. pour μ, σ^2 étant connu.
2. pour σ^2, μ étant connu.
3. pour $\underline{\theta} = (\mu, \sigma^2)^T$.

Solution :

On utilise ici le critère de factorisation. La vraisemblance s'écrit

$$\begin{aligned} L_{\mu, \sigma^2} &= \prod_{i=1}^n \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{1}{2} \left(\frac{x_i - \mu}{\sigma} \right)^2 \right) \\ &= \frac{1}{(\sigma \sqrt{2\pi})^n} \exp \left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2 \right) \end{aligned}$$

Il est immédiat qu'il est possible de réécrire cette vraisemblance sous les formes suivantes :

$$L_{\mu, \sigma^2} = \underbrace{\frac{1}{(\sigma\sqrt{2\pi})^n} \exp\left(-\frac{1}{2\sigma^2} \sum x_i^2\right)}_{h(\underline{\mathbf{x}})} \underbrace{\exp\left(-\frac{1}{2\sigma^2} \mu \sum x_i\right) \exp\left(-\frac{1}{2\sigma^2} n\mu^2\right)}_{g_{\mu}(T_1(\underline{\mathbf{x}}))},$$

où $T_1(\underline{\mathbf{X}}) = \sum X_i$,

$$L_{\mu, \sigma^2} = \underbrace{\frac{1}{(\sqrt{2\pi})^n} \frac{1}{\sigma^n}}_{h(\underline{\mathbf{x}})} \underbrace{\exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)}_{g_{\sigma^2}(T_2(\underline{\mathbf{x}}))},$$

où $T_2(\underline{\mathbf{X}}) = \sum (X_i - \mu)^2$,

$$L_{\mu, \sigma^2} = \underbrace{\frac{1}{(\sqrt{2\pi})^n} \frac{1}{\sigma^n}}_{h(\underline{\mathbf{x}})} \underbrace{\exp\left(-\frac{1}{2\sigma^2} \left(\sum_{i=1}^n x_i^2 - 2\mu \sum_{i=1}^n x_i + n\mu^2\right)\right)}_{g_{\mu, \sigma^2}(\underline{\mathbf{T}}_3(\underline{\mathbf{x}}))},$$

où $\underline{\mathbf{T}}_3(\underline{\mathbf{X}}) = (\bar{X}, m'_2)^T$, où l'on a posé $m'_2 = n^{-1} \sum X_i^2$. Ainsi, T_1 est une statistique exhaustive pour l'estimation de μ , T_2 est une statistique exhaustive pour l'estimation de σ^2 et $\underline{\mathbf{T}}_3$ est une statistique exhaustive pour le vecteur $(\mu, \sigma^2)^T$.

Exercice 7

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

$$f_{\theta}(x) = \begin{cases} \frac{4x^3}{\theta^4} & \text{si } 0 < x < \theta \\ 0 & \text{sinon.} \end{cases}$$

Une suite de n expériences indépendantes a donné les valeurs x_1, \dots, x_n . Proposer un estimateur non biaisé $T(\underline{\mathbf{X}})$ du paramètre θ lié simplement à la moyenne $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ de l'échantillon.

Solution :

Nous cherchons un estimateur non biaisé basé sur la moyenne. Afin de déterminer la transformation à appliquer à celle-ci, il faut commencer par calculer son biais :

$$\mathbb{E}[\bar{X}] = \mathbb{E}[X] = \int_0^{\theta} \frac{4x^4}{\theta^4} dx = \left[\frac{4x^5}{5\theta^4} \right]_0^{\theta} = \frac{4\theta}{5}.$$

Ainsi, $\mathbb{E}[\bar{X}]$ est une fonction linéaire de θ et l'on peut trouver un estimateur sans biais en inversant cette relation :

$$\theta = \frac{5}{4} \mathbb{E}[\bar{X}] = \mathbb{E}\left[\frac{5\bar{X}}{4}\right].$$

L'estimateur convenant ici est donc $\frac{5}{4} \bar{X}$.

Exercice 8

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

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

où a et θ sont deux constantes strictement positives. Une suite de n expériences indépendantes a donné les valeurs x_1, \dots, x_n . Supposons que la valeur du paramètre θ soit connue. Proposer un estimateur non biaisé $T(\underline{\mathbf{X}})$ du paramètre a .

Solution :

Nous utilisons ici le même raisonnement qu'à l'exercice précédent. On trouve :

$$\mathbb{E}[\bar{X}] = \mathbb{E}[X] = \int_0^a \frac{1}{\theta a^{1/\theta}} x^{1/\theta} dx = \left[\frac{x^{\frac{1}{\theta}+1}}{(1+\theta)a^{1/\theta}} \right]_0^a = \frac{a}{1+\theta}.$$

La valeur de θ étant supposée connue, une simple transformation linéaire nous permet à nouveau d'obtenir un estimateur sans biais :

$$\mathbb{E}[(1+\theta)\bar{X}] = a.$$

L'estimateur à prendre est donc $(1+\theta)\bar{X}$.

Exercice 9

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

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

où $\theta > 0$. Déterminez une statistique exhaustive pour le paramètre θ .

Solution :

La méthode la plus simple pour calculer une statistique exhaustive passe par la condition nécessaire et suffisante. Il faut pour cela calculer la vraisemblance :

$$L_\theta(\underline{x}) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \left(\theta x_i^{\theta-1} \mathbb{I}_{[0 < x_i < 1]} \right) = \theta^n \left(\prod_{i=1}^n x_i \right)^{\theta-1} \prod_{i=1}^n \mathbb{I}_{[0 < x_i < 1]}.$$

Celle-ci peut s'écrire $g_\theta(T(\underline{x}))h(\underline{x})$ pour la statistique (exhaustive donc)

$$T(\underline{X}) = \prod_{i=1}^n X_i.$$

Exercice 10

Soient les variables aléatoires X_1, \dots, X_n iid $U[0, \theta]$. Déterminez une statistique exhaustive pour le paramètre θ .

Solution :

Pour rappel, la fonction de densité d'une variable aléatoire uniforme $[0, \theta]$ est

$$f_\theta(x) = \frac{1}{\theta} \mathbb{I}_{[0 \leq x \leq \theta]}.$$

La vraisemblance est donnée par

$$L_\theta(\underline{x}) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \left(\frac{1}{\theta} \mathbb{I}_{[0 \leq x_i \leq \theta]} \right) = \left(\frac{1}{\theta^n} \mathbb{I}_{[x_{(n)} \leq \theta]} \right) \mathbb{I}_{[0 \leq x_{(1)}]}.$$

Une statistique exhaustive est obtenue via le CNS d'exhaustivité et est donnée par

$$T(\underline{X}) = X_{(n)}.$$

Exercice 11

Soient les variables aléatoires X_1, \dots, X_n iid $U[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$. Déterminez une statistique exhaustive pour le paramètre θ .

Solution :

La fonction de densité de chacune de nos variables aléatoires est donnée par

$$f_\theta(x) = \frac{1}{\theta} \mathbb{I}_{[\theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2}]}.$$

La vraisemblance est

$$L_\theta(\underline{\mathbf{x}}) = \prod_{i=1}^n f_\theta(x_i) = \prod_{i=1}^n \left(\frac{1}{\theta} \mathbb{I}_{[\theta - \frac{1}{2} \leq x_i \leq \theta + \frac{1}{2}]} \right) = \left(\frac{1}{\theta^n} \mathbb{I}_{[x_{(1)} \geq \theta - \frac{1}{2}]} \mathbb{I}_{[x_{(n)} \leq \theta + \frac{1}{2}]} \right) \cdot 1$$

Une statistique exhaustive est donnée par

$$T(\underline{\mathbf{X}}) = (X_{(1)}, X_{(n)}).$$

Exercice 12

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

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

où $\theta > 0$. Déterminez une statistique exhaustive pour le paramètre θ .

Solution :

La vraisemblance est donnée par

$$L_\theta(\underline{\mathbf{x}}) = \prod_{i=1}^n f_\theta(x_i) = \theta^n \prod_{i=1}^n e^{-x_i+\theta} \mathbb{I}_{[x_i \geq \theta]} = \underbrace{\exp\left(-\sum_{i=1}^n x_i\right)}_{h(\underline{\mathbf{x}})} \underbrace{\theta^n \exp(n\theta) \mathbb{I}_{[x_{(1)} \geq \theta]}}_{g_\theta(T(\underline{\mathbf{x}}))}.$$

Une statistique exhaustive est donc donnée par

$$T(\underline{\mathbf{X}}) = X_{(1)}.$$

Exercice 13

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

$$f_\theta(x) = \theta e^{-\theta x} I_{[0, \infty[}(x).$$

Déterminez une statistique exhaustive pour le paramètre θ .

Solution :

Il est immédiat que $T(\underline{\mathbf{X}}) = \sum X_i$ est une statistique exhaustive (toujours par le même critère).

Exercice 14

Soient les variables aléatoires X_1, \dots, X_n iid telles que $E[X_i] = \mu$ et $\text{Var}[X_i] = \sigma^2$ pour tout $i = 1, \dots, n$. Montrer que $\bar{X}^{(n)} = \frac{1}{n} \sum_{i=1}^n X_i$ est un estimateur convergent de μ .

Solution :

Nous allons montrer pour cela (comme très souvent) que

$$\mathbb{E}[\bar{X}] = \mu \text{ et } \text{Var}(\bar{X}) \rightarrow 0.$$

Le premier résultat est immédiat. De plus, on a $\text{Var}(\bar{X}) = \sigma^2/n$. Ceci a été montré (indirectement) à l'exercice 1, puisque $\text{Var}(\bar{X}) = \mathbb{E}[\bar{X}^2] - \mathbb{E}[\bar{X}]^2$.

Exercice 15

Soient les variables aléatoires X_1, \dots, X_n iid $\text{Bin}(1, p)$. Considérons les estimateurs du paramètre p suivants :

$$\hat{p}_1 = \frac{\sum_{i=1}^n X_i - 1}{n}, \hat{p}_2 = \frac{1}{\lfloor n/2 \rfloor} \sum_{i=1}^{\lfloor n/2 \rfloor} X_{2i}, \hat{p}_3 = \frac{1}{10} \sum_{i=1}^{10} X_i.$$

Sont-ils biaisés, convergents, exhaustifs ?

Solution :

Nous avons vu (exercice 4) que l'estimateur $T(\mathbf{X}) = \sum X_i$ est un estimateur exhaustif du paramètre p . Par bijection, on conclut immédiatement que \hat{p}_1 est exhaustif, au contraire de \hat{p}_2 et \hat{p}_3 , qui ne sont pas en bijection avec T . Des calculs rapides mettent en évidence que

$$\begin{aligned} \mathbb{E}[\hat{p}_1] &= \frac{\sum \mathbb{E}[X_i] - 1}{n} = \frac{np - 1}{n} = p - \frac{1}{n}. \\ \mathbb{E}[\hat{p}_2] &= \frac{\sum_{i=1}^{\lfloor n/2 \rfloor} \mathbb{E}[X_{2i}]}{\lfloor n/2 \rfloor} = p. \\ \mathbb{E}[\hat{p}_3] &= \frac{\sum_{i=1}^{10} \mathbb{E}[X_i]}{10} = p. \\ \text{Var}(\hat{p}_1) &= \frac{\text{Var}(\sum_i X_i - 1)}{n^2} \stackrel{\text{ind.}}{=} \frac{\sum_i \text{Var}(X_i)}{n^2} = \frac{np(1-p)}{n^2}. \\ \text{Var}(\hat{p}_2) &= \frac{\sum_{i=1}^{\lfloor n/2 \rfloor} \text{Var}(X_{2i})}{(\lfloor n/2 \rfloor)^2} = \frac{\lfloor n/2 \rfloor p(1-p)}{(\lfloor n/2 \rfloor)^2}. \\ \text{Var}(\hat{p}_3) &= \dots = \frac{p(1-p)}{10}. \end{aligned}$$

On conclut que \hat{p}_1 est un estimateur biaisé mais asymptotiquement non biaisé. Sa variance tendant vers zéro, il est convergent. Les deux autres estimateurs (\hat{p}_2 et \hat{p}_3) sont non biaisés. \hat{p}_2 est convergent. La variance de \hat{p}_3 ne convergeant pas vers 0, cet estimateur n'est pas convergent.

Exercice 16

Soient les variables aléatoires $X_i (i = 1; 2)$ i.i.d. de moyenne μ et de variance σ^2 . Lequel des deux estimateurs de μ non biaisés suivants conseilleriez-vous :

$$M_1 = \frac{X_1 + X_2}{2} \text{ ou } M_2 = \frac{aX_1 + bX_2}{a + b},$$

où $a, b \in \mathbb{R}$?

Solution :

On trouve rapidement que

$$\text{Var}(M_1) = \frac{\sigma^2}{2} \text{ et } \text{Var}(M_2) = \frac{a^2 + b^2}{(a + b)^2} \sigma^2.$$

La question qui se pose alors est de savoir quel estimateur dispose de la plus petite variance (nous avons alors deux estimateurs non biaisés, celui possédant la plus petite variance sera considéré comme "meilleur"). Il est facile de voir que

$$\frac{a^2 + b^2}{(a + b)^2} \geq \frac{1}{2},$$

pour tout a et b avec égalité $\Leftrightarrow a = b$. Il est relativement aisé de généraliser ce résultat au cas de n v.a. i.i.d. en considérant les deux estimateurs que sont la moyenne (classique) et la moyenne pondérée de coefficients a_1, \dots, a_n .

On peut donc conclure que le meilleur estimateur en terme de variance (critère d'efficacité, voir tp suivant) dans la classe des estimateurs sous la forme de moyenne pondérée est la moyenne "classique".

Exercice 17

Soient les variables aléatoires X_1, \dots, X_n indépendantes identiquement distribuées de loi $U[0, \theta]$. L'estimateur

$$\hat{\theta} = (n+1)X_{(1)}$$

est-il exhaustif? non biaisé? convergent?

Solution :

Nous avons vu à l'exercice 10 que $T(\underline{X}) = X_{(n)}$ est une statistique exhaustive pour le paramètre θ . Ce n'est pas le cas de notre estimateur $\hat{\theta}(\underline{X})$, celui-ci n'étant pas en bijection avec $T(\underline{X})$. Nous allons maintenant étudier de deux manières différentes la convergence. Dans un premier temps, nous allons utiliser uniquement la définition :

Pour rappel, la fonction de répartition de $\hat{\theta} = X_{(1)}$ est donnée par

$$F_{X_{(1)}}(x) := \mathbb{P}_\theta [X_{(1)} \leq x] = 1 - (\mathbb{P}_\theta [X_1 > x])^n = \left(1 - \left(1 - \frac{x}{\theta}\right)^n\right) \mathbb{I}_{[0 < x < \theta]} + \mathbb{I}_{[x \geq \theta]}.$$

On trouve immédiatement que la fonction de densité est

$$f_{X_{(1)}}(x) := F'_{X_{(1)}}(x) = \frac{n}{\theta} \left(1 - \frac{x}{\theta}\right)^{(n-1)} \mathbb{I}_{[0 < x < \theta]}.$$

Regardons la convergence faible :

$$\begin{aligned} \mathbb{P}_\theta [|(n+1)X_{(1)} - \theta| > \epsilon] &= \mathbb{P}_\theta [-\epsilon < (n+1)X_{(1)} - \theta < \epsilon] \\ &= \mathbb{P}_\theta \left[\frac{\theta - \epsilon}{(n+1)} < X_{(1)} < \frac{\theta + \epsilon}{(n+1)} \right] \\ &= F_{X_{(1)}} \left(\frac{\theta + \epsilon}{(n+1)} \right) - F_{X_{(1)}} \left(\frac{\theta - \epsilon}{(n+1)} \right) \\ &= \left(1 - \left(1 - \frac{\theta + \epsilon}{(n+1)\theta} \right)^n \right) - \left(1 - \left(1 - \frac{\theta - \epsilon}{(n+1)\theta} \right)^n \right) \\ &\rightarrow -\exp\left(-\frac{\theta + \epsilon}{\theta}\right) + \exp\left(-\frac{\theta - \epsilon}{\theta}\right) \neq 0. \end{aligned}$$

L'estimateur n'est ainsi pas faiblement convergent. On conclut immédiatement qu'il n'est pas fortement convergent.

Une autre méthode passe par le calcul de l'espérance et de la variance de notre estimateur (régulant au passage la question du biais). Disposant de la fonction de densité, on trouve

$$\mathbb{E}_\theta [X_{(1)}] = \int_{\mathbb{R}} x f_{X_{(1)}} dx = \int_0^\theta x \left(1 - \frac{x}{\theta}\right)^{(n-1)} \frac{n}{\theta} dx = \dots = \frac{\theta}{(n+1)}.$$

L'estimateur est donc sans biais puisque

$$\mathbb{E}_\theta [\hat{\theta}] = \mathbb{E}_\theta [(n+1)X_{(1)}] = (n+1) \frac{\theta}{(n+1)} = \theta.$$

La variance est donnée par

$$\text{Var}(X_{(1)}) = \mathbb{E}_\theta [X_{(1)}^2] - \frac{\theta^2}{(n+1)^2} = \cdots = \frac{n\theta^2}{(n+1)^2(n+2)}.$$

Ainsi,

$$\text{Var}(\hat{\theta}) = \text{Var}((n+1)X_{(1)}) = (n+1)^2 \text{Var}(X_{(1)}) = \frac{n\theta^2}{n+2} \rightarrow 0.$$

On conclut à nouveau à la non-convergence faible.

Exercice 18

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

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

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

1. Montrez que l'estimateur $\hat{\theta} = \frac{n}{\sum_{i=1}^n X_i^2}$ de θ est une statistique exhaustive pour ce paramètre.

2. Cet estimateur est-il biaisé? Convergent?

Aide : commencez par déterminer la loi de probabilité de $Y = \theta X^2$, puis celle de $Z = \sum_{i=1}^n Y_i$, où les Y_i sont des variables aléatoires mutuellement indépendantes et de même loi que Y .

Solution :

1. Le support de la distribution ne dépendant pas du paramètre inconnu θ , nous supposons que tous les X_i sont strictement positifs. Ecrivons la vraisemblance du modèle :

$$L_\theta(\underline{X}) = (2^n \prod_{i=1}^n X_i) \theta^n e^{-\theta \sum_{i=1}^n X_i^2}$$

Le critère de factorisation nous dit alors que $\sum_{i=1}^n X_i^2$ est une statistique exhaustive pour θ . Par ailleurs, la fonction $x \rightarrow n/x$ est une fonction bijective sur \mathbb{R}_0^+ , et $\sum_{i=1}^n X_i^2$ est toujours dans \mathbb{R}_0^+ . Etant en bijection avec une statistique exhaustive, la statistique $\hat{\theta}$ proposée comme estimateur est également exhaustive.

2. On effectue tout d'abord le changement de variable $y = \theta x^2$, ce qui implique $dy = 2\theta x dx$, et donc en réécrivant la fonction de répartition :

$$\begin{aligned} P(Y \leq u) &= 0 \text{ si } u \leq 0 \text{ et,} \\ P(Y \leq u) &= P(X \leq \sqrt{\frac{u}{\theta}}) = \int_0^{\sqrt{\frac{u}{\theta}}} 2\theta x e^{-\theta x^2} dx = \int_0^u e^{-y} dy \text{ si } u \geq 0. \end{aligned}$$

On constate ainsi que la densité de Y est

$$f_\theta(y) = \begin{cases} e^{-y} & \text{si } y > 0 \\ 0 & \text{sinon,} \end{cases}$$

soit la densité d'une loi $\Gamma(1, 1)$. Puisque les X_i sont indépendants, les Y_i le sont aussi et grâce aux propriétés de la loi gamma nous savons donc que

$$Z = \sum_{i=1}^n Y_i \sim \Gamma(1, n).$$

Ceci nous permet de calculer le biais de $\hat{\theta}$.

$$\mathbb{E}_{\theta} [\hat{\theta}] = \mathbb{E} \left[\frac{n\theta}{Z} \right] = \int_{\mathbb{R}} \frac{n\theta}{z} f(z) dz$$

avec

$$f(z) = \begin{cases} \frac{1}{\Gamma(n)} z^{n-1} e^{-z} & z > 0 \\ 0 & z \leq 0 \end{cases}$$

i.e.

$$\mathbb{E}_{\theta} [\hat{\theta}] = \int_0^{\infty} \frac{n\theta}{\Gamma(n)} z^{n-2} e^{-z} dz.$$

La présence de z^{n-2} dans l'expression de l'intégrande, par ailleurs fort proche de la densité d'une loi gamma, nous suggère de faire apparaître la densité d'une $\Gamma(1, n-1)$. En utilisant les propriétés de la fonction Γ :

$$\mathbb{E}_{\theta} [\hat{\theta}] = \frac{n\theta}{n-1} \int_0^{\infty} \frac{1}{\Gamma(n-1)} z^{n-2} e^{-z} dz$$

l'intégrale ci dessus étant simplement l'intégrale d'une densité $\Gamma(1, n-1)$ sur son domaine, elle vaut exactement 1 et par conséquent

$$\mathbb{E}_{\theta} [\hat{\theta}] = \frac{n}{n-1} \theta.$$

L'estimateur est donc biaisé, mais asymptotiquement sans biais.

Pour étudier la convergence de $\hat{\theta}$ il faut encore calculer la variance de cet estimateur : si celle-ci tend vers 0 lorsque n tend vers l'infini, on aura démontré sa convergence. On a

$$\mathbb{E}_{\theta} [\hat{\theta}^2] = \int_0^{\infty} \frac{n^2 \theta^2}{z^2} f(z) dz$$

Le raisonnement est le même que pour l'espérance, mais cette fois la puissance de z dans l'intégrale est abaissée de deux unités, et l'on fait donc apparaître la densité d'une loi $\Gamma(1, n-2)$

$$\mathbb{E}_{\theta} [\hat{\theta}^2] = \frac{n^2 \theta^2}{(n-1)(n-2)} \int_0^{\infty} \frac{1}{\Gamma(n-2)} z^{n-3} e^{-z} dz = \frac{n^2 \theta^2}{(n-1)(n-2)} \theta^2$$

ce qui implique

$$\begin{aligned} \text{Var}_{\theta} (\hat{\theta}) &= \frac{n^2 \theta^2}{(n-1)(n-2)} \theta^2 - \left(\frac{n}{n-1} \right)^2 \theta^2 \\ &= \frac{n^2 \theta^2}{(n-1)^2 (n-2)} \end{aligned}$$

cette quantité tend bien vers 0, ainsi l'estimateur proposé est convergent.

Exercice 19

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

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

où a et θ sont deux constantes strictement positives.

1. Montrez que l'estimateur $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n \log \frac{X_i}{a}$ de θ est une statistique exhaustive pour ce paramètre.

2. Cet estimateur est-il biaisé? Convergent?

Aide : commencez par déterminer la loi de probabilité de $Y = -\frac{1}{\theta} \log \frac{X}{a}$, puis celle de $Z = \sum_{i=1}^n Y_i$, où les Y_i sont des variables aléatoires mutuellement indépendantes et de même loi que Y .

Solution :

1. Ecrivons la vraisemblance du modèle

$$L_{\theta}(\underline{X}) = \left(\frac{1}{\theta a^{1/\theta}} \right)^n e^{(1/\theta-1) \sum_{i=1}^n \log X_i}$$

Ceci implique que la somme $\sum \log X_i$ ainsi que toute fonction bijective de cette dernière statistique est une statistique exhaustive pour θ . En particulier l'estimateur proposé.

2. Effectuons le changement de variable proposé

$$Y = g(X) = -\frac{1}{\theta} \log \frac{X}{a}.$$

La densité de Y est alors donnée par

$$f_Y(y) = f_X(g^{-1}(y)) \frac{1}{|g'(g^{-1}(y))|}.$$

(ceci est la formule du changement de variables dans les intégrales; comme quoi le cours de cdi peut servir à quelque chose). Ici on a

$$g^{-1}(y) = ae^{-\theta y} \text{ et } g'(x) = -\frac{1}{\theta x}$$

et donc

$$f_Y(y) = \frac{1}{\theta a} e^{-(1-\theta)y} \theta a e^{-\theta y} = e^{-y}$$

pour y positif et $x \in (0, a)$. On reconnaît donc la densité d'une loi $\Gamma(1, 1)$. Comme précédemment, la loi de $Z = \sum Y_i$ est donc une $\Gamma(1, n)$. Par ailleurs, l'estimateur $\hat{\theta}$ peut s'écrire

$$\hat{\theta} = \frac{\theta}{n} Z$$

et on calcule donc l'espérance de cet estimateur grâce à la formule donnée pour l'espérance de Z :

$$\mathbb{E}_{\theta} [\hat{\theta}] = \frac{\theta}{n} \mathbb{E}(Z) = \theta$$

et l'estimateur est sans biais. La formule nous permettant de calculer le moment d'ordre deux également, on en déduit la variance de notre estimateur

$$\text{Var}_{\theta} (\hat{\theta}) = \frac{\theta^2}{n^2} (n-1)$$

Cette quantité tend vers zéro lorsque n tend vers l'infini, ainsi, l'estimateur $\hat{\theta}$ est convergent.

Exercice 20

Prouvez la propriété : "Si $\hat{\theta}^{(n)}(\underline{\mathbf{X}})$ est un estimateur asymptotiquement sans biais et tel que

$$\text{Var} \left(\hat{\theta}^{(n)}(\underline{\mathbf{X}}) \right) \rightarrow 0,$$

alors $\hat{\theta}^{(n)}(\underline{\mathbf{X}})$ est faiblement convergent."

Aide : Utilisez pour cela l'inégalité de Chebychev : Si X est une variable aléatoire de moyenne μ et de variance σ^2 , alors, pour tout α positif,

$$\mathbb{P} [|X - \mu| > \alpha] \leq \frac{\sigma^2}{\alpha^2}.$$

Solution :

Il nous faut montrer ici que, pour tout ϵ ,

$$\mathbb{P} \left[\left| \hat{\theta}^{(n)} - \theta \right| > \epsilon \right]$$

tend vers 0 pour n tendant vers l'infini. On est très tenté d'appliquer Chebychev...mais pour cela nous avons besoin de l'espérance de notre variable aléatoire $\hat{\theta}^{(n)}$. On écrit donc

$$\mathbb{P} \left[\left| \hat{\theta}^{(n)} - \theta \right| > \epsilon \right] = \mathbb{P} \left[\left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] + \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| > \epsilon \right].$$

Or,

$$\left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] + \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| \leq \left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] \right| + \left| \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right|.$$

Ainsi, l'événement

$$A = \left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] + \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| > \epsilon$$

implique l'événement

$$B = \left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] \right| + \left| \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| > \epsilon.$$

De plus, $\hat{\theta}^{(n)}$ étant asymptotiquement sans biais, on sait qu'il existe N tel que $\forall n \geq N$,

$$\left| \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| < \frac{\epsilon}{2}.$$

D'où, si n est suffisamment grand,

$$\begin{aligned} \mathbb{P} \left[\left| \hat{\theta}^{(n)} - \theta \right| > \epsilon \right] &\leq \mathbb{P} \left[\left| \hat{\theta}^{(n)} - \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] \right| > \epsilon - \left| \mathbb{E}_{\theta} \left[\hat{\theta}^{(n)} \right] - \theta \right| > \frac{\epsilon}{2} \right] \\ &\leq \frac{\text{Var} \left(\hat{\theta}^{(n)} \right)}{(\epsilon/2)^2} \end{aligned}$$

par Chebychev. ϵ étant fixé et la variance tendant vers zéro par hypothèse, on a bien le résultat.