Chapitre 20 : Correction des tests

1 test

Test 1 (*Voir solution.*)

Soient $a \in \mathbb{R}_+^*$ et $n \in \mathbb{N}^*$. On considère $X_1, \dots, X_n$ une suite de variables aléatoires indépendantes de loi $\mathcal{U}([0, a])$. On pose $M_n = \max(X_1, \dots, X_n)$.

1. *Déterminer la fonction de répartition de M_n. En déduire que M_n est à densité et déterminer une densité.*
2. *Justifier que M_n est une estimateur du paramètre a et déterminer son biais.*
3. *Est-il asymptotiquement sans biais ?*
4. *Déterminer son risque quadratique.*
5. *Est-il convergent ?*

Test 2 (*Voir solution.*)

*Soit X une variable aléatoire possédant une espérance m et un moment d'ordre 2 noté m_2. Soit $(X_1, \dots, X_n)$ un n-échantillon de la loi de X. On appelle **variance empirique** de l'échantillon la variable :*

$$S_n^2 = \frac{1}{n} \sum_{k=1}^{n} (X_k - \overline{X}_n)^2.$$

1. *On considère S_n^2 comme un estimateur de $V(X)$. Déterminer son biais. Est-ce un estimateur sans biais ?*
2. *Montrer que $\frac{n}{n-1} S_n^2$ est un estimateur sans biais de $V(X)$.*

L'estimateur $\frac{n}{n-1} S_n^2 = \frac{1}{n-1} \sum_{k=1}^{n} (X_i - \overline{X}_n)^2$ est appelée la variance empirique modifiée.

Test 3 (*Voir solution.*)

Soit X une variable aléatoire suivant une loi géométrique de paramètre $p \in]0, 1[$ inconnue. Soit $(x_1, \dots, x_n)$ une réalisation d'un n-échantillon $(X_1, \dots, X_n)$ de la loi de X.

1. *On note $s = x_1 + \cdots + x_n$. Déterminer $\mathcal{L}_n(p)$ pour tout $p \in]0, 1[$.*
2. *Étudier les variations de $\ln \circ \mathcal{L}_n$ et un déduire l'estimateur du maximum de vraisemblance.*

2 Correction des tests

Correction du test 1 (*Retour à l'énoncé.*)

1. *Soit $x \in \mathbb{R}$. On a :*

$$\begin{aligned} F_{M_n}(x) &= P(M_n \leqslant x) = P(\max(X_1,\dots,X_n) \leqslant x) \\ &= P([X_1 \leqslant x] \cap \cdots \cap [X_n \leqslant x]) \\ &= P(X_1 \leqslant x) \times P(X_n \leqslant x) \quad \textit{par indépendance mutuelle} \\ &= P(X_1 \leqslant x)^n \quad \textit{car elles suivent toutes la loi de } X_1 \\ &= \begin{cases} 0 & \textit{si } x < 0 \\ \left(\frac{x}{a}\right)^n & \textit{si } x \in [0,a] \\ 1 & \textit{si } x > a \end{cases} . \end{aligned}$$

On constate que F_{M_n} est de classe C^1 (donc continue) sur $\mathbb{R}\backslash\{0,a\}$. De plus, on vérifie facilement que F_{M_n} est aussi continue en 0 et en a. Ainsi, elle est continue sur $\mathbb{R}$. On en déduit donc que M_n est à densité.
De plus pour tout $x \in \mathbb{R}\backslash\{0,a\}$ on a :

$$F'_{M_n}(x) = \begin{cases} 0 & \textit{si } x < 0 \\ \frac{n}{a^n}\frac{nx^{n-1}}{a^n} & \textit{si } x \in]0,a[\\ 0 & \textit{si } x > a \end{cases} .$$

Ainsi, la fonction f_n définie sur $\mathbb{R}$ par :

$$\forall x \in \mathbb{R}, \quad f_n(x) = \begin{cases} 0 & \textit{si } x < 0 \\ \frac{nx^{n-1}}{a^n} & \textit{si } x \in [0,a] \\ 0 & \textit{si } x > a \end{cases}$$

est une densité de M_n.

2. *La variable M_n est bien une fonction de $(X_1,\dots,X_n)$ indépendante de a. De plus, comme f_n est nulle en dehors de $[0,a]$ d'après le théorème de transfert, M_n possède une espérance si et seulement si l'intégrale $\int_0^a t f_n(t)\,dt$ converge absolument. Or la fonction $t \mapsto |t f_n(t)|$ est continue sur $[0,a]$ donc l'intégrale $\int_0^a |t f_n(t)|\,dt$ n'a pas d'impropreté donc converge. Ainsi M_n possède une espérance quelque soit la valeur de a et on a :*

$$E_a(M_n) = \int_0^a t f_n(t)\,dt = \frac{n}{a^n}\int_0^a t^n\,dt = \frac{n}{a^n}\left[\frac{t^{n+}}{n+1}\right]_0^a = \frac{n}{n+1}a.$$

Ainsi on obtient :

$$\forall a \in \mathbb{R}_+^*, \quad b_a(M_n) = \frac{n}{n+1}a - a = -\frac{a}{n+1}.$$

3. *On a vu que le biais de M_n est donné par :*

$$\forall a \in \mathbb{R}_+^*, \quad b_a(M_n) = \frac{n}{n+1}a - a = -\frac{a}{n+1}.$$

En particulier, on obtient :

$$\forall a \in \mathbb{R}_+^*, \quad \lim_{n\to+\infty} b_a(M_n) = \lim_{n\to+\infty} -\frac{a}{n+1} = 0.$$

L'estimateur M_n est donc un estimateur asymptotiquement sans biais de a.

4. *Comme f_n est nulle en dehors de $[0,a]$ d'après le théorème de transfert, M_n possède un moment d'ordre si et seulement si l'intégrale $\int_0^a t^2 f_n(t)\,dt$ converge absolument. Or la fonction $t \mapsto |t^2 f_n(t)| = t^2 f_n(t)$ est continue sur $[0,a]$ donc l'intégrale $\int_0^a |t^2 f_n(t)|\,dt$ n'a pas d'impropreté donc converge. Ainsi M_n possède un moment d'ordre 2 quelque soit la valeur de a. Déterminons le risque quadratique. D'après le théorème de*

transfert, on a pour tout $a \in \mathbb{R}_+^*$ *:*

$$\begin{aligned}
r_a(\mathrm{M}_n) = \mathrm{E}_a((\mathrm{M}_n - a)^2) = \int_0^a (t-a)^2 f_n(t)\,dt &= \frac{n}{a^n}\int_0^a (t-a)^2 t^{n-1}\,dt \\
&= \frac{n}{a^n}\int_0^a (t^2 - 2at + a^2)^2 t^{n-1}\,dt \\
&= \frac{n}{a^n}\left(\int_0^a t^{n+1}\,dt - 2a\int_0^a t^n\,dt + a^2\int_0^a t^{n-1}\,dt\right) \\
&= \frac{n}{a^n}\left(\left[\frac{t^{n+2}}{n+2}\right]_0^a - 2a\left[\frac{t^{n+1}}{n+1}\right]_0^a + a^2\left[\frac{t^n}{n}\right]_0^a\right) \\
&= \frac{n}{a^n}\left(\frac{a^{n+2}}{n+2} - 2\frac{a^{n+2}}{n+1} + \frac{a^{n+2}}{n}\right) \\
&= \frac{2a}{(n+1)(n+2)}.
\end{aligned}$$

5. *D'après la question précédente,* M_n *possède un moment d'ordre 2 pour tout* $a \in \mathbb{R}_+^*$ *et on a :*

$$\forall a \in \mathbb{R}_+^*, \quad \lim_{n\to+\infty} r_a(\mathrm{M}_n) = \lim_{n\to+\infty} \frac{2a}{(n+1)(n+2)} = 0.$$

La suite $(\mathrm{M}_n)_{n\in\mathbb{N}^*}$ *est donc un estimateur convergent de* a*.*

Correction du test 2 (*Retour à l'énoncé.*)

1. *On remarque que :*

$$\begin{aligned}
\mathrm{S}_n^2 &= \frac{1}{n}\left(\sum_{i=1}^n \mathrm{X}_i^2 - 2\overline{\mathrm{X}}_n \sum_{i=1}^n \mathrm{X}_i + n\overline{\mathrm{X}}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathrm{X}_i^2 - 2n\overline{\mathrm{X}}_n^2 + n\overline{\mathrm{X}}_n^2\right) \\
&= \frac{1}{n}\left(\sum_{i=1}^n \mathrm{X}_i^2 - n\overline{\mathrm{X}}_n^2\right).
\end{aligned}$$

Or, pour tout $i \in [\![1,n]\!]$, X_i *possède un moment d'ordre* 2 *donc* X_i^2 *possède une espérance. Par linéarité, on en déduit que* $\sum_{i=1}^n \mathrm{X}_i^2$ *possède une espérance donnée par :*

$$\mathrm{E}\left(\sum_{i=1}^n \mathrm{X}_i^2\right) = \sum_{i=1}^n \mathrm{E}(\mathrm{X}_i^2) = n m_2.$$

Par ailleurs, on a :

$$\overline{\mathrm{X}}_n^2 = \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{X}_i^2 + 2\sum_{i=1}^n\sum_{j=1}^{i-1} \mathrm{X}_i\mathrm{X}_j\right).$$

Par indépendance, pour tout $i \neq j$ *la variable* $\mathrm{X}_i\mathrm{X}_j$ *possède une espérance donnée par :*

$$\mathrm{E}(\mathrm{X}_i\mathrm{X}_j) = \mathrm{E}(\mathrm{X}_i)\mathrm{E}(\mathrm{X}_j) = m^2.$$

Par linéarité, on en déduit que $\overline{\mathrm{X}}_n^2$ *possède une espérance donnée par :*

$$\begin{aligned}
\mathrm{E}(\overline{\mathrm{X}}_n^2) &= \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{E}(\mathrm{X}_i^2) + 2\sum_{i=1}^n\sum_{j=1}^{i-1} \mathrm{E}(\mathrm{X}_i\mathrm{X}_j)\right) \\
&= \frac{1}{n^2}\left(n m_2 + 2\sum_{i=1}^n\sum_{j=1}^{i-1} m^2\right) \\
&= \frac{1}{n^2}\left(n m_2 + 2\sum_{i=1}^n (i-1)m^2\right) \\
&= \frac{1}{n^2}\left(n m_2 + n(n-1)m^2\right).
\end{aligned}$$

Finalement, par linéarité, S_n^2 *possède une espérance donnée par :*

$$\begin{aligned}
E(S_n^2) = \frac{1}{n}\left(E(\sum_{i=1}^{n} X_i^2) - nE(\overline{X}_n^2)\right) &= \frac{1}{n}\left(nm_2 - \frac{1}{n}(nm_2 + n(n-1)m^2)\right) \\
&= \frac{n-1}{n}\left(m_2 - m^2\right) \\
&= \frac{n-1}{n} V(X).
\end{aligned}$$

En particulier, on obtient le biais suivant :

$$b(S_n^2) = \frac{n-1}{n} V(X) - V(X) = -\frac{V(X)}{n}.$$

L'estimateur est donc biaisé.

2. *Par linéarité de l'espérance,* $\frac{n}{n-1} S_n^2$ *possède une espérance donnée par :*

$$E\left(\frac{n}{n-1} S_n^2\right) = \frac{n}{n-1} E(S_n^2) = \frac{n}{n-1}\frac{n-1}{n} V(X) = V(X).$$

L'estimateur $\frac{n}{n-1} S_n^2$ *est donc sans biais.*

Correction du test 3 (*Retour à l'énoncé.*)

Soit X *une variable aléatoire suivant une loi géométrique de paramètre* $p \in]0,1[$ *inconnue. Soit* $(x_1,\ldots,x_n)$ *une réalisation d'un* n*-échantillon* $(X_1,\ldots,X_n)$ *de la loi de* X.

1. *Par indépendance, on a :*

$$\begin{aligned}
\mathscr{L}_n(p) = P(X_1 = x_1,\ldots,X_n = x_n) &= P(X_1 = x_1) \times \cdots \times P(X_n = x_n) \\
&= p(1-p)^{x_1-1} \times \cdots \times p(1-p)^{x_n-1} \\
&= p^n (1-p)^{s-n}.
\end{aligned}$$

2. *Posons* $\ell_n = \ln \circ \mathscr{L}_n$. *On a, pour tout* $p \in]0,1[$ *:*

$$\ell_n(p) = n\ln(p) + (s-n)\ln(1-p).$$

La fonction ℓ_n *est dérivable et on a :*

$$\ell_n'(p) = \frac{n}{p} - \frac{s-n}{1-p} = \frac{(1-p)n - (s-n)p}{p(1-p)} = \frac{n-sp}{p(1-p)}.$$

On en déduit :

x	0		$\frac{s}{n}$		1
Signe de $\ell_n'(x)$		$+$	0	$-$	
Variations de ℓ_n		$\nearrow$		$\searrow$	

Comme $\ln$ *est strictement croissante, le maximum de la fonction* $\mathscr{L}_n$ *est le maximum de* ℓ_n. *Ainsi, l'estimateur du maximum de vraisemblance est* $\frac{s}{n}$, *c'est-à-dire* $\overline{X}_n$.