Statistique 3
2019-2020. Exercices d'entraînement : Approche Asymptotique (1)

Exercice 1 Comportements asymptotiques de suites de variables aléatoires

Soit $(T_n)_{n\geq 1}$ une suite de vecteurs aléatoires dans $\mathbb{R}^d$, $d \geq 1$. On dit que la suite T_n est *bornée en probabilité* (ou encore qu'elle est *tendue*), et on écrit $T_n = O_{\mathbb{P}}(1)$, si et seulement si pour tout $\varepsilon > 0$, il existe $A > 0$ et $n_0 \geq 1$ tels que pour tout $n \geq n_0$, $\mathbb{P}[\|T_n\|_2 \geq A] \leq \varepsilon$. Plus géneralement, si $(s_n)_{n\geq 1}$ est une suite de variables aléatoires réelles, on écrit que $T_n = O_{\mathbb{P}}(s_n)$ si pour tout $\varepsilon > 0$, il existe $A > 0$ et $n_0 \geq 1$ tels que pour tout $n \geq n_0$, $\mathbb{P}[\|T_n\|_2 \geq A s_n] \leq \varepsilon$.

1. Montrer que si $T_n = o_{\mathbb{P}}(1)$ (i.e., T_n tend vers zéro en probabilité), alors $T_n = O_{\mathbb{P}}(1)$.
2. Montrer que si T_n converge en probabilité, alors elle est tendue.
3. Montrer que si T_n converge en loi, alors elle est tendue.
4. Montrer que si $(\rho_n)_{n\geq 1}$ est une suite tendant vers $+\infty$, et si $\rho_n T_n$ converge en loi, alors $T_n = o_{\mathbb{P}}(1)$.
5. Supposons que T_n converge vers zéro en probabilité. Soit $g : \mathbb{R}^d \to \mathbb{R}$ une fonction telle que $g(x) = o(\|x\|_2^p)$ lorsque $x \to 0$. Montrer que $g(T_n) = o_{\mathbb{P}}(\|T_n\|_2^p)$, i.e., pour tout $\varepsilon > 0$, $\mathbb{P}[|g(T_n)| \geq \varepsilon \|T_n\|_2^p] \to 0$ quand $n \to \infty$.
6. Supposons que T_n converge vers zéro en probabilité. Soit $g : \mathbb{R}^d \to \mathbb{R}$ une fonction telle que $g(x) = O(\|x\|_2^p)$ lorsque $x \to 0$. Montrer que $g(T_n) = O_{\mathbb{P}}(\|T_n\|_2^p)$.
7. Pour $n \geq 1$, soit X_n une variable aléatoire de Poisson de paramètre $1/n$.
   a) Montrer que $X_n = o_{\mathbb{P}}(1)$.
   b) Montrer (le fait étonnant) que pour toute suite de réels strictements positifs $(u_n)_{n\geq 1}$, $X_n = o_{\mathbb{P}}(u_n)$.
   c) X_n tend-elle vers zéro presque sûrement ?

Exercice 2 Modes de convergence

Soit $(T_n)_{n\geq 1}$ une suite de vecteurs aléatoires dans $\mathbb{R}^d$ ($d \geq 1$) et T un vecteur aléatoire dans $\mathbb{R}^d$.

1. Montrer que si T_n converge presque sûrement vers T, alors T_n converge aussi en probabilité vers T.
2. Montrer que si T_n converge en probabilité vers T, alors T_n converge en loi vers T.
3. Si T est déterministe, montrer que la convergence en loi vers T implique la convergence en probabilité vers T.

Exercice 3 La divergence de Kullback-Leibler

Soit $(\mathcal{X},\mathcal{F})$ un espace mesurable quelconque. On définit la divergence de Kullback-Leibler entre deux probabilités P et Q sur $(\mathcal{X},\mathcal{F})$ de la manière suivante:

$$D(P\|Q) = \begin{cases} \int_{\mathcal{X}} \log\left(\frac{\mathrm{d}P}{\mathrm{d}Q}(x)\right) \mathrm{d}P(x) \text{ si } P \ll Q \\ \infty \text{ sinon.} \end{cases}$$

1. Montrer que $D(P\|Q)$ est toujours positive.
2. Calculer $D(P\|Q)$ lorsque P et Q sont, respectivement, les lois de Bernoulli de paramètres $p \in (0,1)$ et $q \in (0,1)$. A-t-on $D(P\|Q) = D(Q\|P)$?
3. Montrer que $D(P\|Q) = 0$ si et seulement si $P = Q$.
4. Montrer que pour toutes probabilités P, Q sur $(\mathcal{X},\mathcal{F})$, $D(\cdot\|Q)$ et $D(P\|\cdot)$ sont des fonctions convexes (ces fonctions étant définies sur l'ensemble des mesures de probabilité sur $(\mathcal{X},\mathcal{F})$).

Exercice 4 Médianes

Soit X une variable aléatoire réelle. Pour tout $\theta \in \mathbb{R}$, on pose $\Phi(\theta) = \mathbb{E}[|X-\theta| - |X|]$. On rappelle qu'une médiane est un nombre réel m tel que $\mathbb{P}[X \leq m] \geq 1/2$ et $\mathbb{P}[X \geq m] \geq 1/2$

1. Calculer l'ensemble des médianes de chacune des lois suivantes:
   a) La loi de Bernoulli de paramètre $p \in [0,1]$.
   b) La loi géométrique de paramètre $p \in (0,1)$ ($\mathbb{P}[X = k] = (1-p)^{k-1}p, \forall k \in \{1,2,\ldots\}$).
   c) La loi uniforme sur un intervalle $[a,b]$, où $a < b$.
   d) La loi exponentielle de paramètre $\lambda > 0$.
   e) La loi normale de paramètres μ, σ^2, où $\mu \in \mathbb{R}$, $\sigma^2 > 0$.
   f) $\frac{1}{2}\mathcal{U}([-2,-1]) + \frac{1}{2}\mathcal{U}([1,2])$, où $\mathcal{U}[a,b]$ signifie la loi uniforme sur l'intervalle $[a,b]$, pour $a < b$.
2. Montrer que si la loi de X est symétrique (i.e., X et $-X$ ont la même loi), alors 0 est une médiane.
3. Montrer que si la loi de X admet une densité strictement positive sur un intervalle et nulle partout ailleurs, alors elle admet une unique médiane.
4. Montrer que Φ est convexe. Donner un exemple de loi pour laquelle Φ n'est pas strictement convexe.
5. Montrer que Φ admet au moins un minimum.
6. Dans ce qui suit, on suppose que X admet une densité f.

a) Montrer que Φ est deux fois dérivable sur $\mathbb{R}$, et calculer ses deux premières dérivées *(Indice: écrire* $\Phi(\theta)$ *sous forme d'une intégrale qu'on découpera en deux intégrales dont les bornes dépendent de* θ*)*

b) En déduire qu'un réel minimise Φ si et seulement s'il est une médiane de la loi de X.

c) Montrer que si la densité f est strictement positive sur $\mathbb{R}$, alors Φ est strictement convexe. Peut-elle être fortement convexe sur $\mathbb{R}$?

7. Dans le cas général, montrer qu'un réel minimise Φ si et seulement s'il est une médiane de la loi de X.