PSI* 2024/25 DS 5

Grandes déviations

Origine : Centrale 2017 - PSI1, mais bon c'est digne d'un sujet X à plusieurs endroits...

Toutes les variables aléatoires mentionnées dans ce sujet sont supposées discrètes.
La partie **1** est composée de trois sous-parties mutuellement indépendantes, toutes trois utilisées dans la partie **2**.

Notations et rappels.

Soient X une variable aléatoire discrète réelle et $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires réelles, mutuellement indépendantes, définies sur un même espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, suivant toutes la loi de X. On pose $S_0 = 0$ et, pour $n \in \mathbb{N}^*$,

$$S_n = \sum_{k=1}^{n} X_k$$

Si Y est une variable aléatoire réelle admettant un moment d'ordre 1, on note $\mathbb{E}[Y]$ l'espérance de Y.
Si Y est une variable aléatoire réelle admettant un moment d'ordre 2, on note $\mathbb{V}[Y]$ la variance de Y.
Si Y est une variable aléatoire à valeurs dans $\mathbb{R}^+$, on abrège "Y est d'espérance finie" en "$\mathbb{E}[Y] < +\infty$".
Si $\tau \in \mathbb{R}^{+*}$, on dit que X vérifie (C_τ) si $\mathbb{E}\left[e^{\tau|X|}\right] < +\infty$.
On pourra utiliser la propriété suivante :

$$\mathcal{P} : \text{si } Y, Z \text{ sont des variables aléatoires réelles telles que } 0 \leqslant Y \leqslant Z,\ \mathbb{E}[Z] < +\infty \Rightarrow \mathbb{E}[Y] < +\infty.$$

Etant données deux variables aléatoires réelles Y et Z sur $(\Omega, \mathcal{A}, \mathbb{P})$, on dit que Y *est presque sûrement égale à* Z lorsque $\mathbb{P}(Y = Z) = 1$.
On admet le résultat suivant (lemme des coalitions) : soit $(Y_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes. Soient A et B des sous-ensembles de $\mathbb{N}^*$ disjoints. Alors, toute variable aléatoire fonction de Y_n, $n \in A$ est indépendante de toute variable aléatoire fonction des Y_n, $n \in B$.

I Premiers résultats

A Une classe de variables aléatoires

1. (a) Soit V une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ possédant un moment d'ordre 2. Montrer que $\mathbb{E}\left[V^2\right] = 0$ si et seulement si V est nulle presque sûrement.

   (b) Soient U et V deux variables aléatoires sur $(\Omega, \mathcal{A}, \mathbb{P})$ possédant un moment d'ordre 2 et telles que V n'est pas une variable nulle presque sûrement.
   Montrer que $\mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right] - \mathbb{E}[UV]^2 \geqslant 0$ et que $\mathbb{E}\left[U^2\right]\mathbb{E}\left[V^2\right] - \mathbb{E}[UV]^2 = 0$ si et seulement si il existe $\lambda \in \mathbb{R}$ tel que $\lambda V + U$ est nulle presque sûrement.

2. (a) On suppose que X est bornée. Justifier que X vérifie (C_τ) pour tout τ dans $\mathbb{R}^{+*}$.

   (b) On suppose que X suit la loi géométrique de paramètre $p \in]0, 1[$

   $$\forall k \in \mathbb{N}^*,\ \mathbb{P}(X = k) = p(1-p)^{k-1}$$

   Quels sont les réels t tels que $\mathbb{E}\left[e^{tX}\right] < +\infty$? Pour ces t, donner une expression simple de $\mathbb{E}\left[e^{tX}\right]$.

(c) On suppose que X suit la loi de Poisson de paramètre λ :

$$\forall k \in \mathbb{N},\ \mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!} \quad \text{où} \quad \lambda \in \mathbb{R}^{+*}$$

Quels sont les réels t tels que $\mathbb{E}\left[e^{tX}\right] < +\infty$? Pour ces t, donner une expression simple de $\mathbb{E}\left[e^{tX}\right]$.

3. Soient a et b deux réels tels que $a < b$. On suppose que $\mathbb{E}\left[e^{aX}\right] < +\infty$ et $\mathbb{E}\left[e^{bX}\right] < +\infty$.
   (a) Montrer que $\forall t \in [a, b],\ e^{tX} \leqslant e^{aX} + e^{bX}$. En déduire $\mathbb{E}\left[e^{tX}\right] < +\infty$.
   Que peut-on en déduire sur l'ensemble $\left\{t \in \mathbb{R} / \ \mathbb{E}\left[e^{tX}\right] < +\infty\right\}$?
   (b) Soient $k \in \mathbb{N}$, t dans $]a, b[$. On note $\theta_{k,t,a,b}$ la fonction $y \in \mathbb{R} \mapsto \frac{y^k e^{ty}}{e^{ay} + e^{by}}$.
   Déterminer les limites de $\theta_{k,t,a,b}$ en $+\infty$ et $-\infty$. Montrer que cette fonction est bornée sur $\mathbb{R}$.
   (c) Montrer que $\mathbb{E}\left[|X|^k e^{tX}\right] < +\infty$.
   (d) On reprend les notations de la question (b). Soient $k \in \mathbb{N}$, c et d deux réels tels que $a < c < d < b$. Montrer qu'il existe $M_{k,a,b,c,d} \in \mathbb{R}^+$ tel que pour tout $t \in [c, d]$ et pour tout $y \in \mathbb{R}$: $|\theta_{k,t,a,b}(y)| \leqslant M_{k,a,b,c,d}$.
4. Dans cette question, τ est un élément de $\mathbb{R}^{+*}$ et X vérifie (C_τ).
   (a) Montrer que l'ensemble des réels t tels que $\mathbb{E}\left[e^{tX}\right] < +\infty$ est un intervalle I contenant $[-\tau, \tau]$.
   Pour tout t dans I, on note $\varphi_X(t) = \mathbb{E}\left[e^{tX}\right]$.
   (b) Montrer que si $X(\Omega)$ est fini, φ_X est continue sur I et de classe C^∞ sur l'intérieur de I.
   (c) On suppose maintenant que $X(\Omega)$ est un ensemble infini dénombrable. On note $X(\Omega) = \{x_n / \ n \in \mathbb{N}^*\}$ où $(x_n)_{n\in\mathbb{N}^*}$ est une suite de réels deux à deux distincts et on pose pour tout $n \in \mathbb{N}^*$, $p_n = \mathbb{P}(X = x_n)$.
   En utilisant les résultats établis en **I.A.3** et deux théorèmes relatifs aux séries de fonctions que l'on énoncera complétement, montrer que φ_X est continue sur I et de classe C^∞ sur l'intérieur de I.
   (d) Vérifier que pour t dans l'intérieur de I et $k \in \mathbb{N}$, $\varphi_X^{(k)}(t) = \mathbb{E}\left[X^k e^{tX}\right]$.
   (e) Soit $\psi_X = \frac{(\varphi_X)'}{\varphi_X}$.
   Montrer que ψ_X est croissante sur l'intérieur de I et que, si X n'est pas presque sûrement égale à une constante, ψ_X est strictement croissante sur l'intérieur de I.

B Inégalité de Bienaymé-Tchebychev

On suppose que X admet un moment d'ordre 2.

1. Soit δ un élément de $\mathbb{R}^{+*}$. Montrer que, pour n dans $\mathbb{N}^*$,

$$\mathbb{P}(|S_n - n\mathbb{E}[X]| \geqslant n\delta) \leqslant \frac{\mathbb{V}[X]}{n\delta^2}$$

2. Si u et v sont deux nombres réels tels que $u < \mathbb{E}[X] < v$, déterminer la limite de la suite $(\pi_n)_{n\in\mathbb{N}^*}$ définie par

$$\forall n \in \mathbb{N}^*,\ \pi_n = \mathbb{P}(nu \leqslant S_n \leqslant nv)$$

C Suites sur-additives

Soit $(u_n)_{n\geqslant 0}$ une suite réelle telle que : $\forall (m, n) \in \mathbb{N}^2,\ u_{m+n} \geqslant u_m + u_n$.
On suppose que l'ensemble $\left\{\frac{u_n}{n},\ n \in \mathbb{N}^*\right\}$ est majoré et on note s sa borne supérieure.

1. Soient m, q, r des éléments de $\mathbb{N}$. On note $n = mq + r$. Comparer les deux nombres réels u_n et $qu_m + u_r$ et montrer que $u_n - ns \geqslant q(u_m - ms) + u_r - rs$.
2. On fixe m dans $\mathbb{N}^*$ et ε dans $\mathbb{R}^{+*}$. En utilisant la division euclidienne de n par m, montrer qu'il existe un entier N tel que pour tout $n > N$,

$$\frac{u_n}{n} \geqslant \frac{u_m}{m} - \varepsilon$$

3. Montrer que $\lim\limits_{n\to+\infty} \frac{u_n}{n} = s$.

II Le théorème des grandes déviations

Soit a un nombre réel.

A Exposant des grandes déviations

1. Montrer que $\mathbb{P}(X \geqslant a) = 0 \iff \forall n \in \mathbb{N}^*,\ \mathbb{P}(S_n \geqslant na) = 0$.
2. Soient $m, n \in \mathbb{N}$.
   (a) Montrer que $S_{m+n} - S_m$ et S_n ont même loi.
   (b) Soit b un nombre réel. Montrer que $\mathbb{P}\Big(S_{m+n} \geqslant (n+m)b\Big) \geqslant \mathbb{P}(S_n \geqslant nb)\mathbb{P}(S_m \geqslant mb)$.

On suppose dans toute la suite du problème que $\mathbb{P}(X \geqslant a) > 0$.

3. Montrer que la suite $\left(\frac{\ln(\mathbb{P}(S_n \geqslant na))}{n}\right)_{n \geqslant 1}$ est bien définie et admet une limite $\gamma_a \leqslant 0$ vérifiant

$$\forall n \in \mathbb{N}^*,\ \mathbb{P}(S_n \geqslant na) \leqslant e^{n\gamma_a}$$

Dans toute la suite du problème, on suppose que X vérifie (C_τ) pour un certain $\tau > 0$ et n'est pas presque sûrement constante. On suppose également que $a > \mathbb{E}\left[X\right]$.

On se propose d'établir que $\gamma_a < 0$ (ce qui montre que la suite $(\mathbb{P}(S_n \geqslant na))_{n \geqslant 1}$ converge géométriquement vers 0) puis de déterminer γ_a.

B Majoration des grandes déviations

L'intervalle I et la fonction φ_X sont définis comme dans la question **1.A.4**.

1. Montrer que, pour $n \in \mathbb{N}^*$ et $t \in I \cap \mathbb{R}^+$,

$$\mathbb{E}\left[e^{tS_n}\right] = (\varphi_X(t))^n,\ \mathbb{P}(S_n \geqslant na) \leqslant \frac{\varphi_X(t)^n}{e^{nta}}$$

2. On définit la fonction $\chi : \begin{array}{ccc} I & \to & \mathbb{R} \\ t & \mapsto & \ln(\varphi_X(t)) - ta \end{array}$
   (a) Montrer que la fonction χ est minorée sur $I \cap \mathbb{R}^+$.
   On note η_a la borne inférieure de χ sur $I \cap \mathbb{R}^+$.
   (b) Donner un équivalent de $\chi(t)$ lorsque t tend vers 0. En déduire $\eta_a < 0$.
   (c) Montrer que $\forall n \in \mathbb{N}^*,\ \mathbb{P}(S_n \geqslant na) \leqslant e^{n\eta_a}$. En déduire que $\gamma_a < 0$.
   (d) Dans chacun des deux cas suivants, déterminer l'ensemble des nombres réels a vérfiant les conditions $\mathbb{P}(X \geqslant a) > 0$ et $a > \mathbb{E}\left[X\right]$; puis, pour a vérifiant ces conditions, calculer η_a.
   i. X suit la loi de Bernoulli $\mathcal{B}(p)$ avec $0 < p < 1$.
   ii. X suit la loi de Poisson $\mathcal{P}(\lambda)$ avec $\lambda > 0$.

C Le théorème de Cramer

On suppose ici que la borne inférieure η_a de la fonction χ sur $I \cap \mathbb{R}^+$ est atteinte en un point σ intérieur à $I \cap \mathbb{R}^+$. Soient t un nombre réel intérieur à I et tel que $t > \sigma$, b un nombre réel tel que $b > \frac{(\varphi_X)'(t)}{\varphi_X(t)}$.

1. (a) Calculer $\sum\limits_{x \in X(\Omega)} \frac{e^{tx}}{\mathbb{E}[e^{tX}]}\mathbb{P}(X = x)$.

On admet alors (quitte à modifier $(\Omega, \mathcal{A}, \mathbb{P})$)

- qu'il existe une variable aléatoire X' sur $(\Omega, \mathcal{A})$ telle que $X'(\Omega) = X(\Omega)$ et dont la loi de probabilité est donnée par

$$\forall x \in X(\Omega),\ \mathbb{P}(X' = x) = \frac{e^{tx}}{\mathbb{E}\left[e^{tX}\right]}\mathbb{P}(X = x)$$

- qu'il existe une suite $(X'_n)_{n \in \mathbb{N}^*}$ de variables aléatoires mutuellement indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ suivant toutes la même loi que X'.

(b) Montrer que

$$\mathbb{E}\left[X'\right] = \frac{(\varphi_X)'(t)}{\varphi_X(t)}, \qquad \mathbb{E}\left[X'\right] > a$$

2. On admet que si $n \in \mathbb{N}^*$ et si f est une application de $X(\Omega)^n$ dans $\mathbb{R}^+$, on a

$$\mathbb{E}\left[f(X_1', \dots, X_n')\right] = \frac{\mathbb{E}\left[f(X_1, \dots, X_n\right] e^{tS_n})}{\varphi_X(t)^n}$$

(a) Pour $n \in \mathbb{N}^*$, on pose $S_n' = \sum\limits_{k=1}^{n} X_k'$. Montrer que $\mathbb{P}(na \leqslant S_n' \leqslant nb) \leqslant \mathbb{P}(S_n \geqslant na)\frac{e^{ntb}}{\varphi_X(t)^n}$.

On pourra introduire l'application f définie de $X(\Omega)^n$ dans $\mathbb{R}$ par

$$f(x_1, \dots, x_n) = \begin{cases} 1 & \text{si } \; na \leqslant \sum\limits_{i=1}^{n} x_i \leqslant nb \\ 0 & \text{sinon} \end{cases}$$

(b) En utilisant les questions **I.B.2**, **II.B.2c** et le (a) ci-dessus, montrer finalement que $\eta_a = \gamma_a$.

3. Dans cette question, on pourra utiliser les résultats de **II.B.2d**.

(a) Soit α dans $]0, 1/2[$. Pour $n \in \mathbb{N}^*$, on pose

$$A_n = \{k \in \{0, \dots, n\},\ |k - \frac{n}{2}| \geqslant \alpha n\},\ U_n = \sum_{k \in A_n} \binom{n}{k}$$

Déterminer la limite de la suite $(U_n^{1/n})_{n \geqslant 1}$.

(b) Soit $\lambda > 0$, $\alpha > \lambda$. Pour $n \in \mathbb{N}^*$, on pose

$$T_n = \sum_{\substack{k \in \mathbb{N} \\ k \geqslant \alpha n}} \frac{n^k \lambda^k}{k!}$$

Déterminer la limite de la suite $(T_n^{1/n})_{n \geqslant 1}$.