# PROBABILITÉS ET STATISTIQUES

Durée : 6 heures

## NOTATIONS ET RAPPELS

1° On note $1_A$ la fonction indicatrice d'une partie A d'un ensemble $\mathcal{X}$.

2° L'ensemble des entiers naturels est désigné par $\mathbb{N}$. On note $\mathcal{B}_n$ la tribu borélienne de $\mathbb{R}^n$ et on écrit $\mathcal{B}$ à la place de $\mathcal{B}_1$. Enfin $\mathcal{B}_\infty$ désigne la plus petite tribu sur $\mathbb{R}^{\mathbb{N}}$ qui, pour toute partie finie $J \subset \mathbb{N}$, rend mesurable la projection canonique $\Pi_J$ de $\mathbb{R}^{\mathbb{N}}$ sur $\mathbb{R}^J$.

3° Toutes les variables aléatoires considérées sont prises sur un même espace probabilisé $(\Omega, \mathcal{F}, P)$. Une variable aléatoire à valeurs dans $\mathbb{R}$ est appelée variable aléatoire réelle (en abrégé v.a.r.). Le symbole $E(X)$ désigne, quand elle existe, l'espérance mathématique de la v.a.r. X, relativement à P.

Si X est une variable aléatoire à valeurs dans $\mathbb{R}^n$ ou dans $\mathbb{R}^{\mathbb{N}}$, on note $\mathcal{C}(X)$ la sous-tribu de $\mathcal{F}$ engendrée par X et par $P_X$ la loi de X, mesure image de P par X.

4° On rappelle que toute suite $(X_n\ ;\ n \in \mathbb{N})$ de v.a.r. définit une variable aléatoire $\underline{X}$ à valeurs dans $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}_\infty)$ et que la loi $P_{\underline{X}}$ de cette suite $\underline{X} = (X_n\ ;\ n \in \mathbb{N})$ est déterminée de façon unique par ses valeurs sur l'algèbre des cylindres $\Pi_J^{-1}(B)$ où J décrit l'ensemble des parties finies de $\mathbb{N}$ et B l'ensemble des boréliens de $\mathbb{R}^J$.

5° Soit $L^1(\Omega, \mathcal{F}, P)$ l'espace des (P-classes de) v.a.r. intégrables sur $(\Omega, \mathcal{F}, P)$. Pour toute sous-tribu $\mathcal{G}$ de $\mathcal{F}$ et pour toute v.a.r. $X \in L^1(\Omega, \mathcal{F}, P)$, on désigne par $E(X/\mathcal{G})$ l'espérance mathématique conditionnelle de X par rapport à $\mathcal{G}$.

Si Z est une variable aléatoire à valeurs dans $\mathbb{R}^n$, on note plus simplement $E(X/Z)$ au lieu de $E(X/\mathcal{C}(Z))$ et on désigne par $E(X/Z = \cdot)$ l'unique élément de $L^1(\mathbb{R}^n, \mathcal{B}_n, P_Z)$ tel que pour toute $g : \mathbb{R}^n \to \mathbb{R}$ borélienne et bornée, on ait :

$$E(X\,g(Z)) = \int_{\mathbb{R}^n} E(X/Z = z)\,.\,g(z)\,P_Z(dz).$$

6° On désigne par $\delta_0$ la mesure de Dirac sur $\mathbb{R}$ au point zéro. Si $\mu$ est une probabilité borélienne sur $\mathbb{R}$, on note $\mu_n$ la puissance $n^{\text{ième}}$ de convolution de $\mu$, c'est-à-dire $\mu_0 = \delta_0$, $\mu_1 = \mu$, $\mu_{n+1} = \mu_n * \mu$ pour $n \geqslant 1$.

## PRÉLIMINAIRE

Soient X et Y deux variables aléatoires indépendantes à valeurs respectivement dans $\mathbb{R}^n$ et $\mathbb{R}^m$ et $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ une fonction borélienne et bornée.

Montrer que

$$E(f(X, Y)/X = x) = \int_{\mathbb{R}^m} f(x, y)\,P_Y(dy) \qquad P_X \text{ presque-sûrement}$$

## PARTIE I

Soit $\underline{X} = (X_n \; ; \; n \in \mathbb{N})$ une suite de v.a.r. positives ou nulles. On suppose que :

*a.* la suite $(X_n \; ; \; n \in \mathbb{N})$ est une suite indépendante de v.a.r.;

*b.* les v.a.r. $X_1$, $X_2$, ... ont toutes la même loi $\mu$ supposée différente de $\delta_0$.

On note $\nu$ la loi de $X_0$ et on pose pour tout $n \in \mathbb{N}$,

$$S_n = \sum_{i=0}^{n} X_i$$

On désigne d'autre part pour tout réel $t > 0$, par $N_t$ la variable aléatoire à valeurs dans $\overline{\mathbb{N}}$, définie par

$$N_t = \sum_{n \geq 0} 1_{[0,\, t]} \circ S_n \qquad \text{(nombre des } S_n \in [0, t])$$

1° *a.* Montrer qu'il existe $\varepsilon > 0$ tel que $P\left(\bigcap_{n \geq 0} \bigcup_{i \geq n} \{X_i \geq \varepsilon\}\right) = 1$.

*b.* En déduire que pour tout $t > 0$, $P(N_t = +\infty) = 0$.

*c.* Montrer de plus que

$$P\left(\bigcap_{t>0} \{N_t < +\infty\}\right) = 1 \quad \text{alors que } \{\lim_{t \to +\infty} N_t = +\infty\} = \Omega.$$

On suppose dans la suite que $0 < \int x \mu(dx) = m < +\infty$.

2° Montrer que $\dfrac{N_t}{t}$ converge presque-sûrement vers $\dfrac{1}{m}$, quand $t \to +\infty$ (on remarquera que $S_{N_t - 1} \leq t < S_{N_t}$)

3° On suppose dans cette question, que $\nu = \delta_0$ et que $\mu$ est la loi de Bernoulli de paramètre $p$ $(0 < p < 1)$ :

$$\mu(\{1\}) = p, \qquad \mu(\{0\}) = 1 - p = q.$$

*a.* Calculer $P(N_t = k)$ pour $k$ entier $\geq 1$.

*b.* Calculer $E(N_t)$ puis étudier $\dfrac{E(N_t)}{t}$ quand $t \to +\infty$.

*c.* Calculer $E(N_t^2)$ puis en déduire que $\operatorname*{Sup}_{t \geq 1} \dfrac{E(N_t^2)}{t^2} < +\infty$.

4° On revient au cas général.

*a.* Utiliser 3° pour démontrer que $N_t$ admet des moments de tous les ordres et que

$$\operatorname*{Sup}_{t \geq 1} \frac{E(N_t^2)}{t^2} < +\infty.$$

*b.* Déduire de ce qui précède et de 2°, que $\dfrac{E(N_t)}{t} \to \dfrac{1}{m}$ quand $t \to +\infty$.

*c.* Qu'arrive-t-il si $\int x \mu(dx) = +\infty$ ?

5°

*a.* Montrer que $f$ définie par $f(t) = \dfrac{\mu(]t, +\infty[)}{m} 1_{[0,\, +\infty[}(t)$ est une densité de probabilité sur $\mathbb{R}$. Quelles sont les probabilités $\mu$ pour lesquelles la probabilité de densité $f$ coïncide avec $\mu$ ?

*b.* Démontrer que si $\nu$ est la probabilité de densité $f$, alors

$$E(N_t) = \frac{t}{m} \qquad \text{quel que soit } t > 0.$$

(On pourra, soit effectuer un calcul direct, par exemple en utilisant la densité de $\nu * \mu_n$ ($\mu_n$ puissance $n^{\text{ième}}$ de convolution de $\mu$), soit utiliser la transformée de Laplace).

Nous admettrons, ce qui pourra être utile pour la question suivante, que pour $\mu$ fixée, la probabilité de densité $f$ est la seule loi $\nu$ telle que $E(N_t) = \frac{t}{m}$, quel que soit $t > 0$. (Ce résultat est obtenu facilement lorsque l'on utilise la deuxième méthode dans la question ci-dessus.)

6° On désigne par T une v.a.r. positive ou nulle indépendante de la suite $\underline{X} = (X_n \,;\, n \in \mathbb{N})$ et on pose

$$\alpha_T = \operatorname{Inf}\{ n \in \mathbb{N} : S_n > T \}$$

(avec la convention usuelle que $\alpha_T = +\infty$ si $S_n \leqslant T$ quel que soit $n \in \mathbb{N}$, convention qui sera encore utilisée dans la suite).

On pose également pour tout $n \in \mathbb{N}$

$$S_n^T = S_{\alpha_T + n} - T$$

(avec la convention $S_{+\infty} = +\infty$).

*a.* Montrer que $\alpha_T$ est une variable aléatoire à valeurs dans $\mathbb{N}$ telle que $P(\alpha_T = +\infty) = 0$ et $\{ \alpha_T = k \} \in \mathscr{C}(X_0, X_1, \ldots, X_k, T)$ quel que soit $k \in \mathbb{N}$.

*b.* Montrer que pour tout $n \in \mathbb{N}$, $S_n^T$ est une variable aléatoire positive.

*c.* En étudiant pour tout $n \in \mathbb{N}$, la loi conjointe de $S_0^T,\ S_1^T - S_0^T,\ S_2^T - S_1^T,\ \ldots,\ S_{n+1}^T - S_n^T$, démontrer que la suite $(S_0^T,\ S_1^T - S_0^T,\ \ldots,\ S_{n+1}^T - S_n^T,\ \ldots)$ est une suite indépendante de v.a.r. et que pour tout $n \in \mathbb{N}$, $S_{n+1}^T - S_n^T$ a la loi $\mu$.

*d.* Démontrer que si $\nu$ est la probabilité de densité $f$ définie en 5°, *a*, alors $S_0^T$ a pour densité de probabilité $f$.

7° Le but de la partie III est d'établir que pour une certaine classe de probabilités $\mu$, on a pour tout réel $h > 0$

$$\frac{E(N_{t+h} - N_t)}{h} \to \frac{1}{m} \qquad \text{quand } t \to +\infty.$$

*a.* Montrer que si pour $h > 0$, la limite quand $t \to +\infty$ de $\frac{E(N_{t+h} - N_t)}{h}$ existe, elle est nécessairement égale à $\frac{1}{m}$.

*b.* Montrer qu'il suffit d'établir l'existence de $\lim\limits_{t \to +\infty} \frac{E(N_{t+h} - N_t)}{h}$, pour $\nu = \delta_0$.

## PARTIE II

*(Cette partie est indépendante de la partie I.)*

On désigne par $\Sigma_n$ ($n$ entier $\geqslant 1$) l'ensemble des permutations de $\mathbb{N}^* = \mathbb{N} \setminus \{0\}$ qui laissent invariants les entiers $k$ tels que $k > n$. Et on pose $\Sigma = \bigcup\limits_{n \geqslant 1} \Sigma_n$, ensemble des permutations finies de $\mathbb{N}^*$.

Soit $X = (X_n \,;\, n \geqslant 1)$ une suite indépendante de v.a.r. ayant toutes la même loi. Pour $\sigma \in \Sigma$, on note $X_\sigma$ la suite $(X_{\sigma(n)} \,;\, n \geqslant 1)$.

Un événement $A \in \mathscr{C}(X)$ est dit symétrique relativement à X, si pour tout $\sigma \in \Sigma$, il existe $B \in \mathscr{B}_\infty$ tel que

$$A = \{ X \in B \} = \{ X_\sigma \in B \}.$$

1° *a.* Comparer $P_X$ et $P_{X_\sigma}$ pour $\sigma \in \Sigma$.

*b.* Soit $A_n$ un événement de la forme $A_n = \{ (X_1, \ldots, X_n) \in B_n \}$ avec $B_n \in \mathscr{B}_n$. Démontrer que si $A'_n = \{ (X_{2n}, \ldots, X_{n+1}) \in B_n \}$, alors $P(A_n \cap A'_n) = (P(A_n))^2$.

*c.* Démontrer alors que pour tout événement $A \in \mathscr{C}(X)$, symétrique relativement à X, on a $P(A) = 0$ ou 1.

2° Comparer le résultat précédent avec la loi du tout ou rien de Kolmogorov qui concerne les événements asymptotiques, c'est-à-dire appartenant à $\bigcap_{n \geqslant 1} \mathscr{C}(X_k \,;\, k \geqslant n)$. Donner un exemple d'un événement symétrique qui n'est pas asymptotique.

3° Soit $X_0$ une v.a.r. indépendante de la suite $X = (X_n \,;\, n \geqslant 1)$. On note $\underline{X}$ la suite $(X_n \,;\, n \geqslant 0)$ et on désigne par $\Sigma'$ l'ensemble des permutations finies de $\mathbb{N}$ qui laissent invariant 0.

Soit A un événement appartenant à $\mathscr{C}(\underline{X})$ qui est symétrique relativement à X, c'est-à-dire tel que pour tout $\sigma \in \Sigma'$, il existe $B \in \mathscr{B}_\infty$ avec $A = \{ \underline{X} \in B \} = \{ \underline{X}_\sigma \in B \}$.

Adapter ce qui a été fait en 1° pour montrer que $E(1_A / X_0)$ ne prend presque-sûrement que les valeurs 0 ou 1.

## PARTIE III

On rappelle que le support d'une probabilité $\mu$ borélienne sur $\mathbb{R}$ est le plus petit fermé F qui porte $\mu$, c'est-à-dire tel que $\mu(F) = 1$ ; on le note supp $\mu$. D'autre part on appelle symétrisée de $\mu$, la loi $\mu^s$ de la différence de deux v.a.r. indépendantes de loi $\mu$.

Soient $\underline{X} = (X_n \,;\, n \in \mathbb{N})$ et $\underline{X}' = (X'_n \,;\, n \in \mathbb{N})$ deux suites de v.a.r. positives ou nulles. On suppose que

*a.* la tribu $\mathscr{C}(\underline{X})$ est indépendante de la tribu $\mathscr{C}(\underline{X}')$.

*b.* la suite $(\mathscr{C}(X_n) \,;\, n \in \mathbb{N})$ est une suite indépendante de sous-tribus de $\mathscr{F}$, de même que la suite $(\mathscr{C}(X'_n) \,;\, n \in \mathbb{N})$.

*c.* toutes les $X_n$ et $X'_n$ pour $n \geqslant 1$, ont la même loi $\mu$ ayant une espérance mathématique $m$ telle que $0 < \int x\, \mu(dx) = m < +\infty$.

*d.* l'ensemble $\bigcup_{n \geqslant 1}$ supp $\mu_n^s$ est dense dans $\mathbb{R}$, où $\mu_n^s$ désigne la puissance $n^{\text{ième}}$ de convolution de $\mu^s$ symétrisée de $\mu$ (condition qui est remplie s'il n'existe pas de réel $d \geqslant 0$ tel que $\{ nd : n \in \mathbb{Z} \}$ porte $\mu^s$).

*e.* $X_0 = 0$ alors que $X'_0$ a pour densité $f$ définie en I, 5°.

On pose pour $n \in \mathbb{N}$, $S_n = \sum_{i=0}^{n} X_i$ et $S'_n = \sum_{i=0}^{n} X'_i$ et pour tout $t > 0$, $N_t = \sum_{n \geqslant 0} 1_{[0,t]} \circ S_n$ et $N'_t = \sum_{n \geqslant 0} 1_{[0,t]} \circ S'_n$.

Pour $j \in \mathbb{N}$, soit $g_j$ la fonction définie sur les couples de suites croissantes de réels par : si $\underline{s} = (s_n \,;\, n \in \mathbb{N})$ et $\underline{s}' = (s'_n \,;\, n \in \mathbb{N})$

$$g_j(\underline{s}, \underline{s}') = \operatorname{Inf} \{ s'_n - s_j : n \in \mathbb{N}, s'_n - s_j > 0 \}$$

On considère les variables aléatoires $Z_j$ définies par $Z_j = g_j(\underline{S}, \underline{S}')$ où $j \in \mathbb{N}$, $\underline{S} = (S_n \,;\, n \in \mathbb{N})$ et $\underline{S}' = (S'_n \,;\, n \in \mathbb{N})$.

Soit un réel $\delta > 0$ fixé. On pose pour $i \geqslant 0$,

$$A_i = \bigcup_{j \geqslant i} \{ Z_j < \delta \}$$

1° Soit $i \in \mathbb{N}$. Posons comme dans I. 6° $\alpha_{S_i} = \operatorname{Inf} \{ n \in \mathbb{N} : S'_n - S_i > 0 \}$, et pour $n \in \mathbb{N}$, $S_n^{S_i} = S_{n+i} - S_i$ et $S'^{S_i}_n = S'_{\alpha_{S_i}+n} - S_i$.

Vérifier que pour $i$ et $k \in \mathbb{N}$, on a

$$Z_{i+k} = g_k(\underline{S}^{S_i}, \underline{S}'^{S_i})$$

et en déduire que

$$P(A_0) = P(A_i) = P(A_\infty) \text{ où } A_\infty = \bigcap_{i=0}^{+\infty} A_i$$

2° *a.* En considérant la suite $X'_0, X_1, X'_1, X_2, X'_2, \ldots$ que l'on désignera par $(Y_n\,;\, n \in \mathbb{N})$, démontrer que $E(1_{A_\infty}/X'_0)$ ne prend presque-sûrement que les valeurs 0 ou 1.

*b.* Démontrer que $E(1_{A_1}/X'_0)$ est presque-sûrement strictement positif (pour cela on pourra comparer $A_1$ avec $\bigcup_{n \geqslant 1} \{0 < S'_n - S_n < \delta\}$).

*c.* Déduire de ce qui précède, la valeur commune des $P(A_t)$.

3° On considère les variables aléatoires presque-sûrement finies, définies par

$$K = \operatorname{Inf}\{\, i \in \mathbb{N} : Z_i < \delta \,\}, \qquad K' = \operatorname{Inf}\{\, j \in \mathbb{N} : S'_j > S_K \,\}.$$

*a.* Montrer que, quels que soient $k$ et $k' \in \mathbb{N}$,
$\{K = k\} \cap \{K' = k'\} \in \mathscr{T}(X_0, \ldots, X_k, X'_0, \ldots, X'_{k'})$. Puis établir que les variables aléatoires $(S_K\,;\, S_{K+n} - S_K)$ et $(S_K\,;\, S'_{K'+n} - S'_{K'})$ ont même loi ($n$ entier $\geqslant 1$).

*b.* Démontrer alors que pour tous réels $t > 0$ et $h > \delta$

$$E\Big(\sum_{n>0} 1_{]t+\delta,\, t+h]} \circ S'_{K'+n}\Big) \leqslant E\Big(\sum_{n>0} 1_{]t,\, t+h]} \circ S_{K+n}\Big) \leqslant E\Big(\sum_{n>0} 1_{]t,\, t+h+\delta]} \circ S'_{K'+n}\Big)$$

4° Soient $t$ et $h$ réels $> 0$.

*a.* Montrer que $N'_{t+h} - N'_t$ et $N'_h$ ont même loi puis démontrer que

$$E\Big(\sum_{k \leqslant K'} 1_{]t,\, t+h]} \circ S'_k\Big) \to 0 \qquad \text{quand} \quad t \to +\infty$$

*b.* Montrer que pour tout $n \in \mathbb{N}$, on a $P(N_{t+h} - N_t > n) \leqslant P(N_h > n)$. En déduire le comportement quand $t \to +\infty$ de $E\Big(\sum_{k \leqslant K} 1_{]t,\, t+h]} \circ S_k\Big)$.

5° Déduire de tout ce qui précède que pour tout réel $h > 0$

$$\frac{E(N_{t+h} - N_t)}{h} \to \frac{1}{m} \qquad \text{quand} \quad t \to +\infty$$

6° On suppose maintenant que $\int x\mu(dx) = +\infty$ et que le support de $\mu$ est $[0, +\infty[$. On désigne par $\lambda$ la mesure borélienne sur $[0, +\infty[$ telle que $\lambda([0, t]) = E(N_t)$ quel que soit $t > 0$.

*a.* Soient $\beta = \limsup_{t \to +\infty} E(N_{t+1} - N_t) = \limsup_{t \to +\infty} \lambda(]t, t+1])$ et $(t_k)$ une suite tendant vers l'infini telle que $\lim_{k \to +\infty} \lambda(]t_k, t_k + 1]) = \beta$.

En étudiant $\int_{[0,\, t+1]} \lambda(]t - y,\, t + 1 - y])\, \mu(dy) - \beta$, montrer que pour tout $j$ entier $\geqslant 1$, $\liminf_{k \to +\infty} \lambda(]t_k - j,\, t_k - j + 2]) \geqslant \beta$.

*b.* En étudiant alors l'intégrale $\int_{[0,\, t_k]} \mu(]t_k - y, +\infty[)\, \lambda(dy)$ et en tenant compte de la nature de la série $\sum_{i=1}^{+\infty} \mu(]2i, +\infty[)$, démontrer que $\beta = 0$ et donc que pour tout réel $h > 0$, $\lim_{t \to +\infty} E(N_{t+h} - N_t) = 0$.