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 Π_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 δ_0 la mesure de Dirac sur $\mathbb{R}$ au point zéro. Si μ est une probabilité borélienne sur $\mathbb{R}$, on note μ_n la puissance $n^{\text{ième}}$ de convolution de μ, 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. $X_1, X_2, \ldots$ ont toutes la même loi μ supposée différente de δ_0.

On note ν 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 $\frac{N_t}{t}$ converge presque-sûrement vers $\frac{1}{m}$, 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 μ 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 ≥ 1.

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

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

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 $\frac{E(N_t)}{t} \to \frac{1}{m}$ quand $t \to +\infty$.

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

5°

a. Montrer que f définie par $f(t) = \frac{\mu(]t, +\infty[)}{m} 1_{[0, +\infty[}(t)$ est une densité de probabilité sur $\mathbb{R}$.

Quelles sont les probabilités μ pour lesquelles la probabilité de densité f coïncide avec μ ?

b. Démontrer que si ν 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 $\nu * \mu_n$ (μ_n puissance $n^{\text{ième}}$ de convolution de μ), soit utiliser la transformée de Laplace).

Nous admettrons, ce qui pourra être utile pour la question suivante, que pour μ fixée, la probabilité de densité f est la seule loi ν 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 α_T est une variable aléatoire à valeurs dans $\mathbb{N}$ telle que $P(\alpha_T = +\infty) = 0$ et $\{ \alpha_T = k \} \in \mathcal{T}(X_0, X_1, \ldots, X_k, T)$ 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, $S_1^T - S_0^T$, $S_2^T - S_1^T$, ..., $S_{n+1}^T - S_n^T$, 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 μ.

d. Démontrer que si ν 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 μ, 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 Σ_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_σ la suite $(X_{\sigma(n)} \,;\, n \geqslant 1)$.

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

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

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

b. Soit A_n un événement de la forme $A_n = \{ (X_1, \ldots, X_n) \in B_n \}$ avec $B_n \in \mathcal{B}_n$. 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 $A \in \mathcal{T}(X)$, 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 $\bigcap_{n \geq 1} \mathscr{T}(X_k \; ; k \geq n)$. Donner un exemple d'un événement symétrique qui n'est pas asymptotique.

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

Soit A un événement appartenant à $\mathscr{T}(\underline{X})$ 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 μ borélienne sur $\mathbb{R}$ est le plus petit fermé F qui porte μ, c'est-à-dire tel que $\mu(F) = 1$; on le note supp μ. D'autre part on appelle symétrisée de μ, la loi μ^s de la différence de deux v.a.r. indépendantes de loi μ.

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 $\mathscr{T}(\underline{X})$ est indépendante de la tribu $\mathscr{T}(\underline{X}')$.

b. la suite $(\mathscr{T}(X_n) \; ; n \in \mathbb{N})$ est une suite indépendante de sous-tribus de $\mathscr{F}$, de même que la suite $(\mathscr{T}(X'_n) \; ; n \in \mathbb{N})$.

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

d. l'ensemble $\bigcup_{n \geq 1}$ supp μ_n^s est dense dans $\mathbb{R}$, où μ_n^s désigne la puissance $n^{\text{ième}}$ de convolution de μ^s symétrisée de μ (condition qui est remplie s'il n'existe pas de réel $d \geq 0$ tel que $\{ nd : n \in \mathbb{Z} \}$ porte μ^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 $t > 0$, $N_t = \sum_{n \geq 0} 1_{[0,t]} \circ S_n$ et $N'_t = \sum_{n \geq 0} 1_{[0,t]} \circ S'_n$.

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 $i \geq 0$,

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

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 $k' \in \mathbb{N}$,
$\{K = k\} \cap \{K' = k'\} \in \mathscr{C}(X_0, \ldots, X_k, X'_0, \ldots, X'_{k'})$. 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 μ est $[0, +\infty[$. On désigne par λ 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 $h > 0$, $\lim_{t \to +\infty} E(N_{t+h} - N_t) = 0$.

RAPPORT SUR L'ÉPREUVE «PROBABILITÉS ET STATISTIQUE»

I. Analyse du sujet

Le problème concernait les processus de renouvellement. Il se proposait notamment de faire établir :

a). Le premier théorème du renouvellement (Feller 1941) : s. est différent de δ_o, alors $\frac{E(Nt)}{t} \longrightarrow \frac{1}{m}$ quand $t \longrightarrow +\infty$ (partie I).

b). Le théorème du renouvellement de Blackwell (1953) : si de plus μ n'est pas portée par un ensemble de la forme $\{nd \; ; \; n \in \mathbb{N}\}$ $(d > 0)$, alors pour tout $h > 0$, $\frac{E(N_{t+h} - N_t)}{h} \longrightarrow \frac{1}{m}$ quand $t \longrightarrow +\infty$. (partie III).

Pour faire établir ce dernier résultat, on se servait d'une méthode probabiliste récemment publiée par Lindvall (1977). Cette méthode utilise une idée ancienne de Doeblin, qui consiste à associer au processus étudié, un processus indépendant de celui-ci, de même loi de renouvellement mais stationnaire : les deux processus "finiront par avoir deux points voisins" et à partir de là, leurs évolutions probabilistes seront "semblables". La méthode ne nécessite que des calculs élémentaires en dehors de l'utilisation de la loi du zéro-un de Hewitt et Savage (partie II). La condition d de la partie III est destinée à éviter de faire établir des résultats assez classiques sur la convolution.

II. Remarques générales

Ce problème, certes long, ne contenait aucune difficulté spéciale pour les candidats ayant une expérience moyenne de la conduite de calculs classiques. Celle-ci semble manquer à bon nombre de candidats. Le rapport de l'an dernier notait que les candidats semblaient plus à l'aise dans des raisonnements de type "théorie de la mesure" que dans les arguments probabilistes. Le caractère fortement probabiliste du problème de cette année explique sans doute la relative difficulté que les candidats ont trouvée à ce problème. De toutes façons "le meilleur reste le meilleur" : le candidat qui a dominé les épreuves écrites communes, a aussi largement surpassé les autres candidats dans l'épreuve "Probabilités et Statistiques", conduisant les calculs avec beaucoup de sûreté et d'énergie.

Au chapitre des erreurs ou méconnaissances versons deux points :

a). Soit μ une mesure σ-finie sur un anneau $\mathscr{R}$ de parties de Ω. Le théorème de Carathéodory affirme notamment que l'unique prolongement de μ en une mesure sur la tribu engendrée $\mathscr{T} = \mathscr{T}(\mathscr{R})$ coïncide sur $\mathscr{T}$ avec la mesure extérieure correspondante :

$$\mu^*(E) = \mathrm{Inf}(\Sigma \; \mu(A_n) \; ; \; E \subset \cup A_n \; , \; A_n \in \mathscr{R})$$

Il en résulte que, le prolongement de μ à $\mathscr{T}$ étant encore désigné par μ, pour tout $\varepsilon > 0$, pour tout $A \in \mathscr{T}$ avec $\mu(A) < +\infty$, il existe $B \in \mathscr{R}$ tel que $\mu(A_\Delta B) < \varepsilon$.

Dans la partie II (question 1.c) il convenait d'appliquer ce résultat à une probabilité P. Cela a été très rarement fait.

b). Soient B un borélien de $\mathbb{R}^{\mathbb{N}}$ et B_n le cylindre de $\mathbb{R}^{\mathbb{N}}$ dont la base est la projection de B sur $\mathbb{R}^n$. Il a été souvent affirmé dans la même question II. 1.c que $B = \bigcap_n B_n$. Cela est faux comme on pourra par exemple s'en persuader en prenant $B = \{x = (x_i) : \Sigma x_i^2 = 1\}$; alors $\bigcap B_n = \{x : \Sigma x_i^2 \leqslant 1\}$.

Remarquons que le Jury a eu, dans cette option, à se plaindre de la qualité de la rédaction de beaucoup de copies. Notons que la concision toujours souhaitable n'empêche pas d'indiquer les étapes d'un raisonnement ou d'un calcul, que les calculs doivent être lisibles, qu'une version exacte d'un raisonnement ou d'un calcul suivant sans commentaire une version erronée de la même partie ne saurait effacer celle-ci (!).

Notons également :

c). Les indices aléatoires n'ont pas toujours été bien compris ni maniés correctement. Par exemple dans la question I.6, pour la partie b), ce qui nécessitait explication est la mesurabilité de S_n^T et non sa positivité, et pour la partie c) l'utilisation du point a) est fondamentale.

d). Certains candidats ont paru peu familiers avec la loi binômiale. Notamment une partie d'entre-eux a perdu du temps pour retrouver (avec plus ou moins de bonheur) la formule donnant $P(S_n = k)$.

III. Corrigé succint du problème

Préliminaire. - L'indépendance de X et Y est équivalente à l'égalité $P_{X,Y} = P_X \otimes P_Y$. La relation demandée résulte alors de l'application du théorème de Fubini.

I.1. - La condition $\mu \neq \delta_o$ implique l'existence de $\varepsilon > 0$ tel que $\mu([\varepsilon, +\infty[) > 0$. Le résultat de 1.a se déduit alors immédiatement du lemme de Borel-Cantelli ou peut être directement établi par passage au complémentaire.

Alors presque sûrement, la suite croissante (S_n) tend vers $+\infty$ quand $n \longrightarrow +\infty$. L'événement $\{N_t = +\infty\} = \{\lim_n S_n \leqslant t\}$ est de probabilité nulle et l'événement $\bigcap_{t>o} \{N_t < +\infty\} = \bigcap_{k \in \mathbb{N}^*} \{N_k < +\infty\}$ (croissance en t de N_t) est de probabilité 1.
On a aussi :

$$\{\lim_{t \to +\infty} N_t < +\infty\} = \bigcup_{n \geqslant 1} \{\lim_t N_t \leqslant n\} = \bigcup_{n \geqslant 1} \{S_n = +\infty\} = \phi.$$

I.2. - D'après la loi des grands nombres de Kolmogorov, $\frac{1}{n}\sum_{i=1}^{n} X_i$ converge presque sûrement vers m ; il en est de même de $\frac{S_n}{n} = \frac{X_o}{n} + \frac{1}{n}\sum_{i=1}^{n} X_i$ et $\frac{S_{n-1}}{n} = \frac{n-1}{n} \cdot \frac{S_{n-1}}{n-1}$.

Mais d'après I.1, presque sûrement $N_t \in \mathbb{N}$ pour tout $t > o$ et $\lim_{t \to +\infty} N_t = +\infty$.

Par définition, on a pour tout $n \geqslant 1$, $\{N_t = n\} = \{S_n > t\} \cap \{S_{n-1} \leqslant t\}$, donc $S_{N_t - 1} \leqslant t < S_{N_t}$. En divisant les termes de cette double inégalité par N_t et en tenant compte de ce qui précède, on obtient immédiatement le résultat.

I.3. - La v.a.r. S_n a la loi binômiale de paramètre n et p. Comme on a $\{N_t = k\} = \{S_k > t\} \cap \{S_{k-1} \leqslant t\}$ et que X_k ne prend que les valeurs 0 et 1, on obtient compte tenu de l'indépendance des X_n :

$$P(N_t = k) = P(X_k = 1).P(S_{k-1} = [t])$$

$$= C_{k-1}^{[t]}\ p^{[t]+1}\ q^{k-1-[t]} \qquad \text{si} \quad 0 < t < k$$

et naturellement $\quad P\ (N_t = k) = 0 \qquad$ si $\quad t \geqslant k.$

$$E(N_t) = \sum_{k \geqslant [t]+1} k\ C_{k-1}^{[t]}\ p^{[t]+1}\ q^{k-1-[t]} = \frac{p^{[t]+1}}{[t]\,!} \sum_{n \geqslant 1} (n+[t]) \ldots n\ q^{n-1}$$

On reconnaît ici la dérivée d'ordre $([t]+1)$ d'un développement en série entière classique (!) et on en déduit $E(N_t) = \frac{[t]+1}{p}$.

On peut aussi écrire que

$$E(N_t) = \sum_{k \geqslant [t]+1} \frac{[t]+1}{p}\ C_k^{[t]+1}\ p^{[t]+2}\ q^{k-1-[t]}$$

et noter que $\sum_{k \geqslant [t]+1} P(N_{t+1} = k+1) = 1.$

De façons semblables, on obtient aussi

$$E(N_t.(N_t+1)) = \frac{([t]+1).([t]+2)}{p^2}$$

d'où $E(N_t^2) = E(N_t.(N_t+1)) - E(N_t) = \frac{([t]+1)}{p} \left(\frac{[t]+2}{p} - 1 \right).$

I.4. - Soit $\varepsilon > 0$ tel que $P(X_1 \geqslant \varepsilon) = p > 0$. Si l'on pose $X'_o = 0$ et pour tout $i \geqslant 1$, $X'_i = 1_{\{X_i \geqslant \varepsilon\}}$, on a $\varepsilon X'_i \leqslant X_i$ quel que soit i et $N_t \leqslant N'_{\frac{t}{\varepsilon}} = \sum_{n \geqslant o} 1_{[0,\frac{t}{\varepsilon}]} \circ S'_n$.

Ce que l'on demande en a, se déduit alors de I.3 (on peut remarquer que les résultats de I.3 demeurent si $p = 1$, ce qui permet d'englober le cas $\mu = \delta_\varepsilon$(!)).

La condition $\sup_{t \geqslant 1} \frac{E(N_t^2)}{t^2} < +\infty$ entraîne l'uniforme intégrabilité de $\frac{N_t}{t}$ (mais non celle de $\frac{N_t^2}{t^2}$ (!)) ; celle-ci jointe à la convergence en probabilité de $\frac{N_t}{t}$ vers $\frac{1}{m}$ implique la partie b. On peut aussi, sans faire appel aux résultats généraux que nous venons d'utiliser, écrire

$$\int \left| \frac{N_t}{t} - \frac{1}{m} \right| dP \leqslant \varepsilon + \int \left| \frac{N_t}{t} - \frac{1}{m} \right| \cdot 1_{\{|\frac{N_t}{t} - \frac{1}{m}| > \varepsilon\}} dP$$

et appliquer Cauchy-Schwarz.

Si $\int x \, \mu(dx) = +\infty$, on a $\frac{E(N_t)}{t} \longrightarrow 0$. Ceci s'obtient en considérant pour $A > 0$, la suite (X_i^A) définie par $X_i^A = X_i$ si $X_i \leqslant A$, $X_i^A = A$ si $X_i > A$.

I.5. - On établit que f définie en a est une densité de probabilité en appliquant à la fonction borélienne $1_{\{t < u\}}(t,u)$ le théorème de Tonelli.

Si μ est densitable, $G(t) = \mu(]t, +\infty[)$ est continue. Alors $G(t) = \int_t^{+\infty} \frac{G(u)}{m} du$ quel que soit $t > 0$, entraîne que G est dérivable sur $]0, +\infty[$ et vérifie l'équation différentielle $G'(t) = -\frac{G(t)}{m}$. La seule probabilité possible est celle de densité $\frac{1}{m} \exp(-\frac{t}{m}) 1_{[0, +\infty[}(t)$ et elle vérifie bien la condition voulue.

Pour la partie b, on écrit :

$$E(N_t) = \sum_{n \geqslant o} E(1_{[0,t]} \circ S_n) = \sum_{n \geqslant o} \nu * \mu_n([0,t])$$

En remarquant que $\nu * \mu_n$ a pour densité $g_n(u) = \int_{[0,u]} f(u-v)\mu_n(dv)$ et en remplaçant $f(u-v)$ par sa valeur, on obtient en trois lignes le résultat. On peut aussi suivant l'autre méthode indiquée, considérer la transformée de Laplace de la mesure borélienne sur $\mathbb{R}_+$, λ, associée à la fonction croissante $t \longrightarrow E(N_t)$. On obtient :

$$\hat{\lambda}(s) = \hat{\nu}(s) \cdot \frac{1}{1-\hat{\mu}(s)} .$$

D'après l'injectivité de la transformée de Laplace, on a : $E(N_t) = \frac{t}{m}$ quel que soit $t > 0 \Longleftrightarrow \hat{\nu}(s) = \frac{1 - \hat{\mu}(s)}{m\,s}$ pour tout $s > 0 \Longleftrightarrow \nu$ de densité f.

I.6. - On a

$$\{\alpha_T = k\} = \{S_{k-1} \leqslant T < S_k\} \qquad k \in \mathbb{N}$$

et $$\{\alpha_T = +\infty\} = \{\lim S_n \leqslant T\} \subset \{\lim S_n < +\infty\}.$$

Il en résulte le point a.

On peut écrire :

$$S_{\alpha_T + n} = \sum_{k \geqslant o} 1_{\{\alpha_T = k\}} \cdot S_{k+n} + (+\infty) \cdot 1_{\{\alpha_T = +\infty\}}$$

d'où l'on déduit le point b.

Si $B_o, B_1, \ldots, B_{n+1}$ sont des boréliens de $\mathbb{R}_+$, on peut écrire

$\{S_o^T \in B_o\} \cap \{S_1^T - S_o^T \in B_1\} \cap \ldots \cap \{S_{n+1}^T - S_n^T \in B_{n+1}\}$

$= \bigcup_{k \geqslant o} \{\alpha_T = k\} \cap \{S_k - T \in B_o\} \cap \{X_{k+1} \in B_1\} \cap \ldots \cap \{X_{k+n+1} \in B_{n+1}\}.$

D'après le point a et l'indépendance des X_i, les événements $\{\alpha_T = k\} \cap \{S_k - T \in B_o\}$, $\{X_{k+1} \in B_1\}, \ldots, \{X_{k+n+1} \in B_{n+1}\}$ sont indépendants. On en déduit alors rapidement le point c.

Si l'on pose $N_t^T = \sum_{n \geqslant o} 1_{[0,t]} \circ S_n^T$, on a : $E(N_t^T) = \sum_{n \geqslant o} P(S_n - T \in]0,t])$

puis par indépendance de S_n et T

$$= \sum_{n \geqslant o} \int_o^{+\infty} P(S_n \in]u, t+u])P_T(du)$$

$$= \int_o^{+\infty} (E(N_{t+u}) - E(N_u))P_T(du).$$

Alors le point d résulte de 5.b (partie directe et partie réciproque).

I.7. - En écrivant $E(N_{kh}) = E(N_h) + \sum_{n=1}^{k-1} E(N_{(n+1)h} - N_{nh})$ il suffit pour établir le point a, de se rappeler que pour une suite numérique la convergence ordinaire implique la convergence au sens de Cesaro (!).

L'existence de la limite si $\nu = \delta_o$, s'écrit :

$$g(t) = \sum_{n \geqslant o} \mu_n(]t, t+h]) \longrightarrow \frac{h}{m} \quad \text{quand} \quad t \longrightarrow +\infty.$$

Sous cette condition g est de plus bornée sur $[0, +\infty[$, car bornée sur tout intervalle fini. Le point b est alors établi en appliquant le théorème de convergence dominée à

$$E(N_{t+h} - N_t) = \int g(t-u)\, 1_{[0,t+h]}(u)\, \nu(du).$$

II.1. - Pour $C = \Pi_J^{-1}(B)$ avec $B \in \mathscr{B}_n$, on a

$$P_X(C) = P_{X_\sigma}(C) = \mu^{n\otimes}(B)$$

Par Carathéodory, le théorème sur les π et λ systèmes ou 4$^{\text{e}}$ des rappels les probabilités P_X et P_{X_σ} coïncident sur tout $\mathscr{B}_\infty$.

On a donc $P(A_n) = P(A'_n)$ et comme par associativité de l'indépendance A_n et A'_n sont indépendants, on a

$$P(A_n \cap A'_n) = P(A_n).P(A'_n) = [P(A_n)]^2.$$

Comme nous l'avons rappelé dans les remarques générales, pour tout $A \in \mathscr{T}(X)$, il existe une suite (A_n) de la forme $A_n = \{(X_1,\ldots,X_n) \in B_n\}$ telle que $P(A_\Delta A_n) \longrightarrow 0$ quand $n \longrightarrow +\infty$. Soient σ_n la permutation de Σ_{2n} telle que $\sigma_n(1) = 2n$, $\sigma_n(2) = 2n-1,\ldots,\sigma_n(2n) = 1$, $\sigma_n(i) = i$ pour $i > 2n$, et A'_n l'événement $\{(X_{\sigma_n(1)},\ldots,X_{\sigma_n(n)}) \in B_n\}$. Si A est un événement symétrique, on a $P(A_\Delta A'_n) = P(A_\Delta A_n)$ et les suites (1_{A_n}) et $(1_{A'_n})$ convergent vers 1_A dans $L^1(\Omega, \mathscr{F}, P)$, et il en est de même de la suite $(1_{A_n \cap A'_n})$. Il suffit alors de passer à la limite dans la formule $P(A_n \cap A'_n) = [P(A_n)]^2$ pour obtenir le résultat de la partie c.

Remarquons que nous avons utilisé l'invariance de A seulement pour les permutations σ_n. On aurait pu utiliser d'autres suites de permutations, en particulier la suite définie par $\sigma'_n(1) = 2n+1$, $\sigma'_n(2) = 2n,\ldots,\sigma'_n(2n+1) = 1$, $\sigma'_n(i) = i$ pour $i > 2n+1$, qui respecte la parité. Cette remarque sera utile pour III 2$^{\text{e}}$.

II.2. - La loi du tout ou rien de Kolmogorov affirme sous l'hypothèse d'indépendance de la suite (X_n), que la probabilité d'un événement asymptotique ne peut être que 0 ou 1. Le résultat précédent (loi du tout ou rien de Hewitt et Savage) est établi sous la condition supplémentaire de l'équidistribution des X_n, mais concerne une famille plus vaste d'événements. Il est en effet immédiat de constater que tout événement asymptotique est symétrique ; par contre $A = \bigcap_{n\geqslant 1} \{X_n \in B\}$ est symétrique sans être (en général !) asymptotique.

II.3. - Soit $A \in \mathscr{T}(\underline{X})$ un événement symétrique relativement à X. De la même façon que dans 1º et avec les mêmes notations en remplaçant simplement X par $\underline{X}$ et en prenant $\sigma_n(0) = 0$, les suites (1_{A_n}), $(1_{A'_n})$ et $(1_{A_n \cap A'_n})$ convergent dans L^1 vers 1_A. D'après la continuité de l'espérance mathématique conditionnelle, les espérances mathématiques conditionnelles par rapport à X_o des trois suites précédentes convergent dans L^1 vers $E(1_A/X_o)$. En appliquant le résultat du préliminaire, on obtient $E(1_{A_n \cap A'_n}/X_o) = E(1_{A_n}/X_o).(1_{A'_n}/X_o)$ et cette suite converge dans L^1 à la fois vers $E(1_A/X_o)$ et vers $(E(1_A/X_o))^2$. D'où le résultat demandé.

III.1. - $\underline{S}^{S_i}$ (resp. $\underline{S}'^{S_i}$) a la même loi que $\underline{S}$ (resp. $\underline{S}'$) (cf. I.6). $\underline{S}^{S_i}$ et $\underline{S}'^{S_i}$ sont indépendantes comme $\underline{S}$ et $\underline{S}'$. Donc la suite $(Z_k, k \in \mathbb{N})$ a la même loi que la suite $(Z_{i+k}, k \in \mathbb{N})$. Il en résulte que $P(A_o) = P(A_i)$ quantité encore égale à $P(A_\infty)$ en remarquant la décroissance de la suite (A_i).

III.2. - L'événement A_∞ peut s'écrire : il y a une infinité de S_j qui sont suivis à moins de δ par un S'_n. Sur $\{\lim S_n = +\infty\}$, il a même trace que A'_∞ : "il existe une infinité de couples (i_k, j_k) telle que si $k > k'$, alors $i_k > i_{k'}$ et $j_k > j_{k'}$, et $0 < S'_{j_k} - S_{i_k} < \delta$. L'événement A'_∞ n'est pas symétrique par rapport à $\underline{Y}$ mais il est invariant pour les permutations qui laissent invariants 0 et globalement les pairs et les impairs. D'après la remarque faite en II.1 et applicable encore à II.3, on obtient la partie a.

On a $A_1 \supset \bigcup_{n\geqslant 1} \{0 < S'_n - S_n < \delta\}$. Alors presque-sûrement

$E(1_{A_1}/X'_o) \geqslant E(\sup_n 1_{\{0<S'_n - S_n<\delta\}}/X'_o) \geqslant \sup_n E(1_{]0,\delta[}(S'_n - S_n)/X'_o)$. Mais d'après le préliminaire, on a $P_{X'_o}$ presque-sûrement

$$E(1_{]0,\delta[}(S'_n - S_n)/X'_o = x'_o) = \mu^S_n(]-x'_o, \delta - x'_o[).$$

Alors l'utilisation de l'hypothèse d, donne le résultat du point b.

L'événement $A_1 \setminus A_\infty$ étant négligeable, on a $E(1_{A_1}/X'_o) = E(1_{A_\infty}/X'_o)$. D'après les parties b et a, $E(1_{A_1}/X'_o) = 1$ presque-sûrement ; il en résulte que $P(A_o) = P(A_i) = P(A_\infty) = 1$.

III.3. - L'événement $\{K = k\} \cap \{K' = k'\}$ s'exprime à l'aide de $S_o, S_1, \ldots, S_k, S'_o, S'_1, \ldots, S'_{k'}$; il appartient donc à la tribu $\mathcal{T}(X_o, \ldots, X_k, X'_o, \ldots, X'_{k'})$.

Alors pour toutes fonctions boréliennes bornées f et g, on a :

$$E(f(S_K).g(S_{K+n} - S_K)) = \sum_{k,k'} E(f(S_k).g(S_{k+n} - S_k)\, 1_{\{K=k\} \cap \{K'=k'\}}).$$

Mais $E(f(S_k).g(S_{k+n} - S_k) 1_{\{K=k\} \cap \{K'=k'\}} / \mathcal{T}(X_o, \ldots, X_k, X'_o, \ldots, X'_{k'}))$

$$= f(S_k) 1_{\{K=k\} \cap \{K'=k'\}} \cdot \int g(x) \mu_n(dx).$$

$$= E(f(S_k).g(S'_{k'+n} - S'_{k'}) 1_{\{K=k\} \cap \{K'=k'\}} / \mathcal{T}(X_o, \ldots, X_k, X'_o, \ldots, X'_{k'})).$$

En prenant l'espérance mathématique de l'espérance mathématique conditionnelle et en sommant en k et k', on obtient :

$$E(f(S_K).g(S_{K+n} - S_K)) = E(f(S_K).g(S'_{K'+n} - S'_{K'}))$$

d'où la partie a.

Il en résulte en particulier que S_{K+n} et $S'_{K'+n} - S'_{K'} + S_K$ ont même loi. Comme $0 < S'_{K'} - S_K < \delta$, on a

$$1_{]t+\delta, t+h[} \circ S'_{K'+n} \leqslant 1_{]t,t+h[} \circ (S'_{K'+n} - S'_{K'} + S_K) \leqslant 1_{]t,t+h+\delta[} \circ S'_{K'+n}.$$

En prenant l'espérance mathématique dans cette double inégalité puis en sommant en n, on obtient la double inégalité demandée dans la partie b.

III.4. - En utilisant les notations de I.6 pour la suite (X'_n) au lieu de (X_n), on a

$$N'_{t+h} - N'_t = \sum_{n \geqslant o} 1_{]t,t+h]} \circ S'_n = \sum_{n \geqslant o} 1_{]t,t+h]} \circ S'_{\alpha_t+n} = \sum_{n \geqslant o} 1_{]0,h]} \circ S'^{t}_n.$$

Comme les deux suites de v.a. $(S'^{t}_{n}, n \in \mathbb{N})$ et $(S'_{n}, n \in \mathbb{N})$ ont même loi, il en est de même pour les deux v.a. $N'_{t+h} - N'_{t}$ et N'_{h} ; $(P(S'_{n} = 0) = 0)$.

En partitionnant par l'événement $\{N'_{t+h} - N'_{t} \leqslant A\}$ et son contraire (A réel $\geqslant 1$) et tenant compte de ce qui précède, on obtient :

$$E\Big(\sum_{k \leqslant K'} 1_{]t,t+h]} \circ S'_{k}\Big) \leqslant A.P(S'_{K'} > t) + E(N'_{h} . 1_{\{N'_{h} > A\}}).$$

Il suffit alors de faire tendre t vers $+\infty$, puis A vers $+\infty$ pour obtenir la fin de la partie a.

Toujours avec les notations de I.6 mais cette fois appliquées à (X_n), on a :

$$P(N_{t+h} - N_{t} > n) = P(S_{\alpha_t + n} \leqslant t+h) = \int_{]t,t+h]} \mu_n([0,t+h-s])P_{S_{\alpha_t}}(ds) \leqslant \mu_n([0,h]) = P(N_h > n).$$

Alors la démonstration se conduit comme dans la partie a.

Pour conclure il faut remarquer que

$$E((N_{t+h} - N_{t}).1_{\{N_{t+h} - N_{t} > A\}}) \leqslant E(N_{h}.1_{\{N_{h} > A\}})$$

en se rappelant que pour une v.a. W à valeurs dans $\mathbb{N}$, on a $E(W) = \sum_{n \geqslant o} P(W > n)$.

III.5. - En utilisant ce que l'on vient d'obtenir en 3 et 4 et la question I.5, on obtient :

$$\frac{h-\delta}{m} = \lim_{t \to +\infty} E(N'_{t+h} - N'_{t+\delta}) = \lim_{t \to +\infty} E\Big(\sum_{n > o} 1_{]t+\delta,t+h]} \circ S'_{K'+n}\Big)$$

$$\leqslant \lim_{t \to +\infty} \inf E\Big(\sum_{n > o} 1_{]t,t+h]} \circ S_{K+n}\Big) = \lim_{t \to +\infty} \inf E(N_{t+h} - N_{t})$$

$$\leqslant \lim_{t \to +\infty} \sup E(N_{t+h} - N_{t}) \quad \text{et de façon analogue}$$

$$\leqslant \lim_{t \to +\infty} E(N'_{t+h+\delta} - N'_{t}) = \frac{h+\delta}{m}.$$

D'où le résultat demandé.

III.6. - Compte tenu que $\nu = \delta_o$, on a pour tout $t > 0$:

$$\lambda(]t,t+1]) - \beta = \sum_{n \geqslant 1} \mu_n(]t,t+1]) - \beta = \int_{[0,t+1]} \lambda(]t-y, t+1-y])\mu(dy) - \beta.$$

Pour $0 < j < A < t+1$, en se souvenant que $\lambda([0,1])$ majore $\lambda(]t-y,\ t+1-y])$, on obtient par une majoration grossière :

$$\lambda(]t,t+1]) - \beta \leqslant \lambda([0,1]).\mu([A,+\infty[)+(\lambda(]t-j,\ t+2-j]) - \beta).\mu([j-1,j[)$$
$$+ (\sup_{\theta \geqslant t-A} \lambda(]\theta,\theta+1]) - \beta).\mu([0,A[).$$

En appliquant ceci à une sous-suite (t_{k_n}) de t_k telle que $\lambda(]t_{k_n} - j,\ t_{k_n} - j+2]) \longrightarrow \beta'$, puis en passant à la limite en n et en faisant ensuite tendre A vers $+\infty$, on obtient :

$$(\beta' - \beta)\ \mu([j-1,\ j[) \geqslant 0.$$

D'où la partie a, compte tenu de l'hypothèse sur le support de μ (dont on peut en réalité se passer).

On a $\int_0^{+\infty} \mu(]t,+\infty[)dt = \int_{[0,+\infty[} x\ \mu(dx) = +\infty$; d'où l'on déduit rapidement que $\sum_1^{+\infty} \mu(]2i,+\infty[) = +\infty$.

On a $$\int_{[0,t_k]} \mu(]t_k-y,+\infty[)\ \lambda(dy) = 1 \geqslant \sum_{i=1}^{I_k} \int_{]t_k-2i,t_k - 2i+2]} \mu(]t_k-y,+\infty[)\lambda(dy)$$

($2\ I_k$ est le plus grand entier pair inférieur ou égal à t_k)

$$\geqslant \sum_{i=1}^{I_k} \mu(]2i,+\infty[)\lambda(]t_k-2i,t_k-2i+2]).$$

De là, on déduit aisément (par exemple par Fatou !) que

$$1 \geqslant \beta \cdot \sum_{i=1}^{+\infty} \mu(]2i,+\infty[)$$

et donc que $\beta = 0$.

IV. Statistique des résultats

Nombre de copies corrigées : 550

Notes	$\leqslant 1$	2 à 5	6 à 10	11 à 15	16 à 20	21 à 25	26 à 30	31 à 35	36 à 40
Nombre de copies	226	102	93	60	47	11	3	4	4