PROBABILITÉS ET STATISTIQUES

DUREE : 6 Heures

DÉFINITIONS, NOTATIONS, 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 $\mathbb{N}$. La tribu $\mathcal{B}_\infty$ est la plus petite tribu sur $\mathbb{R}^{\mathbb{N}}$ qui pour tout n rend mesurable la projection canonique de $\mathbb{R}^{\mathbb{N}}$ sur $\mathbb{R}^{\{0, 1, \ldots, n\}}$.

3° On note Γ la fonction de $\mathbb{R}^{+*} \to \mathbb{R}$ définie par :

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$$

et pour $a > 0$, $b > 0$ on appelle loi $\beta(a, b)$ la probabilité de densité

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} t^{a-1}(1-t)^{b-1} 1_{]0,1[}(t)$$

4° Si $(\mathcal{F}_n)_{n \in \mathbb{N}}$ est une suite croissante de tribus, on note $\mathcal{F}_\infty$ la plus petite tribu contenant l'anneau $\bigcup_{n \in \mathbb{N}} \mathcal{F}_n$.

5° On convient de poser $\inf \varnothing = +\infty$; on notera par ailleurs $\underline{x}$ la suite $(x_n)_{n \in \mathbb{N}}$.

6° Toutes les variables aléatoires (en abrégé v.a.) introduites dans ce problème sont supposées définies sur un même espace de probabilité $(\Omega, \mathcal{F}, P)$. Une v.a. à valeurs dans $\mathbb{R}$ est appelée v.a. réelle. Le symbole $E(X)$ désigne, quand elle existe, l'espérance mathématique de la v.a. réelle X; $\sigma^2(X)$ désigne la variance de X. « Presque sûrement » est noté en abrégé p.s.

7° Si U est une v.a. réelle intégrable et V une v.a. à valeurs dans $\mathbb{R}^n$ on note $E(U \mid V)$ la (classe de) v.a. réelles caractérisée par l'égalité

$$E(E(U \mid V)\, h(V)) = E(U\, h(V))$$

où h parcourt l'ensemble des fonctions boréliennes bornées de $\mathbb{R}^n$ dans $\mathbb{R}$.

8° On note $\underline{X}$ la suite de v.a. réelles $(X_n)_{n \in \mathbb{N}}$. On appelle loi de $\underline{X}$ la probabilité sur $(\mathbb{R}^{\mathbb{N}}, \mathcal{B}_\infty)$ image de P par $\underline{X}$; elle est caractérisée par la suite des lois des v.a. $(X_0, X_1, \ldots, X_n)$, à valeurs dans $\mathbb{R}^{n+1}$, pour n parcourant $\mathbb{N}$.

9° *Processus d'urne :* Soit x_0 un réel ($0 \leqslant x_0 \leqslant 1$), m un entier positif ou nul, f une fonction de [0, 1] dans [0, 1]. On appelle processus d'urne, associé à f, de composition initiale (x_0, m) une suite $\underline{X}$ pouvant être définie par les équations de récurrence :

$$\left\{\begin{array}{l} X_0 = x_0 \\ (m + k + 1)\, X_{k+1} = (m + k)\, X_k + 1_{A_{k+1}}, \end{array}\right. \qquad k \in \mathbb{N}$$

et où la suite d'événements $(A_k)_{k \in \mathbb{N}^*}$ satisfait pour tout $k \in \mathbb{N}$, à :

$$E\,(1_{A_{k+1}} \mid X_0, X_1, \ldots, X_k) = f(X_k) \qquad \text{p.s.}$$

On se propose d'en étudier les propriétés asymptotiques.

10° Pour B élément de $\mathcal{B}_\infty$ et $\underline{X}$ processus d'urne de composition initiale (x_0, m), associé à la fonction f on notera :

$$Q^f_{x_0, m}(B) = P(\underline{X} \in B)$$

et, lorsqu'une seule fonction f est considérée, on abrégera

$$Q^f_{x_0, m} \text{ en } Q_{x_0, m}.$$

11° On admettra que pour un processus d'urne $\underline{X}$ de composition initiale (x_0, m) associé à une fonction f, on a pour toute fonction borélienne bornée h de $\mathbb{R}^{\mathbb{N}}$ dans $\mathbb{R}$

$$\begin{aligned} & E\,(h\,(X_{n+1}, X_{n+2}, \ldots, X_{n+k}, \ldots) \mid X_1, \ldots X_n) \\ = {} & E\,(h\,(X_{n+1}, X_{n+2}, \ldots, X_{n+k}, \ldots) \mid X_n) \qquad \text{p.s.} \\ = {} & \int_{\mathbb{R}^{\mathbb{N}}} h\,(\underline{x})\, d\, Q^f_{X_n, n+m}(\underline{x}) \qquad \text{p.s.} \end{aligned}$$

12° On posera $S_n = \sum_{i=1}^{n} 1_{A_i}$.

PRÉLIMINAIRES

1° Soit U une v.a. de loi $\beta\,(a, b)$. Calculer :

a. Pour $0 \leqslant k \leqslant n$ $\qquad E\,(U^k\,(1 - U)^{n-k})$

b. La variance $\sigma^2\,(U)$.

2° Soit $\underline{X}$ un processus d'urne associé à la fonction f et h une fonction borélienne de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que, si elle existe,

$$E\,(h\,(X_{n+1}) \mid X_n) = f\,(X_n)\, h\left(\frac{(m+n)\,X_n + 1}{m+n+1}\right) + (1 - f\,(X_n))\, h\left(\frac{(m+n)\,X_n}{m+n+1}\right) \qquad \text{p.s.}$$

3° Montrer que si $\underline{X}$ est un processus d'urne associé à la fonction f, la suite $\underline{Y}$ définie par :

$$Y_n = 1 - X_n$$

est un processus d'urne associé à une fonction qu'on précisera.

4° Montrer qu'on a pour tout $\underline{X}$ processus d'urne

a. $\qquad \lim_{n \to \infty} \left(X_n - \dfrac{S_n}{n}\right) = 0, \qquad$ et si $\qquad m = 0,\ X_n = \dfrac{S_n}{n}$

b. $\qquad | X_{n+1} - X_n | \leqslant \dfrac{1}{n+1}.$

I. URNE DE BERNOULLI ET APPLICATIONS

Soit $\underline{X}$ un processus d'urne de composition initiale $(x_0, 0)$ et associé à

$$t \rightsquigarrow f(t) = p \qquad (0 < p < 1).$$

1° Quelle est la loi de S_n ?

2° Étudier :

a. $\lim_{n \to \infty} X_n$

b. pour $t \neq p$ $\lim_{n \to \infty} \sum_{k < tn} C_n^k \, p^k (1 - p)^{n-k}$.

3° Soit U une v.a. comprise entre 0 et 1, de fonction de répartition F continue. Étudier :

$$\lim_{n \to \infty} \sum_{k < tn} C_n^k \, E\left(U^k (1 - U)^{n-k}\right).$$

II. MARTINGALES ET SOUS-MARTINGALES

Une suite $\underline{M}$ de v.a. réelles intégrables est une martingale (resp. sous-martingale) s'il existe une suite croissante de sous-tribus $(\mathcal{F}_n)_{n \in \mathbb{N}}$ telle que pour tout $n \in \mathbb{N}$

$$\begin{cases} . \ M_n \text{ est } \mathcal{F}_n \text{ mesurable} \\ .. \ E(M_{n+1} \mid \mathcal{F}_n) = M_n \qquad \text{p.s.} \\ \quad (\text{resp} \geqslant) \end{cases}$$

Si de plus $E(M_n^2) < +\infty$ pour tout entier n, $\underline{M}$ est dite de carré intégrable.

1° Soit $\underline{M}$ une martingale (resp. une sous-martingale); comparer $E(M_n)$ et $E(M_{n+1})$.

2° Montrer que si $\underline{M}$ est une martingale de carré intégrable, $\underline{M}^2$ est une sous-martingale et que

$$E[(M_{n+1} - M_n)^2] = E(M_{n+1}^2) - E(M_n^2)$$

3° Soit $\underline{M}$ une sous-martingale positive de carré intégrable.

a. Montrer que la suite $\underline{M}^2$ est une sous-martingale.

b. Montrer en considérant la suite $\underline{M}'$ définie par

$$M'_n = \sum_{j=0}^{n} M_j - \sum_{j=0}^{n-1} E(M_{j+1} \mid \mathcal{F}_j)$$

que $\underline{M}$ est la somme d'une martingale de carré intégrable et d'une suite croissante.

c. Montrer que pour tout $n \in \mathbb{N}$, on a :

$$E(M'^2_{n+1}) - E(M'^2_n) \leqslant E(M^2_{n+1}) - E(M^2_n)$$

4° Soit $\underline{M}$ une sous-martingale positive et x un nombre strictement positif.

On pose $K = \inf \{ n \in \mathbb{N} : M_n > x \}$

Montrer que pour tout $n \in \mathbb{N}$:

a. $$\mathrm{E}(\mathrm{M}_n) \geqslant \mathrm{E}\left(\sum_{k=0}^{n} \mathrm{M}_k 1_{\{\mathrm{K}=k\}}\right)$$

b. $$\mathrm{P}(\mathrm{Max}(\mathrm{M}_0, \mathrm{M}_1, \ldots, \mathrm{M}_n) > x) \leqslant \frac{\mathrm{E}(\mathrm{M}_n)}{x}$$

5° Soit $\varepsilon > 0$ et $\underline{\mathrm{M}}$ une martingale de carré intégrable. Montrer que pour tous n et k entiers positifs :

$$\mathrm{P}(\max_{j=1,2,\ldots,k} |\mathrm{M}_{n+j} - \mathrm{M}_n| > \varepsilon) \leqslant \frac{\mathrm{E}(\mathrm{M}^2_{n+k}) - \mathrm{E}(\mathrm{M}^2_n)}{\varepsilon^2}$$

6° En déduire que si $\underline{\mathrm{M}}$ est une martingale pour laquelle la suite $(\mathrm{E}(\mathrm{M}^2_n))_{n\in\mathbb{N}}$ est bornée, $\underline{\mathrm{M}}$ est une suite presque sûrement convergente.

7° Déduire de 3° et 6° qu'une sous-martingale positive $\underline{\mathrm{M}}$ converge presque sûrement vers une v.a. réelle si la suite $(\mathrm{E}(\mathrm{M}^2_n))_{n\in\mathbb{N}}$ est bornée.

8° Soit U une v.a. réelle bornée et $(\mathcal{F}_n)$ une suite croissante de sous-tribus.

a. Calculer :

$$\mathrm{E}(\lim_{k\to\infty} \mathrm{E}(\mathrm{U} \mid \mathcal{F}_k) / \mathcal{F}_n)$$

b. Soit Z une v.a. réelle de carré intégrable, $\mathcal{F}_n$ - mesurable.

En déduire la valeur de :

$$\mathrm{E}[\mathrm{Z}(\mathrm{E}(\mathrm{U} \mid \mathcal{F}_\infty) - \lim_{k\to\infty} \mathrm{E}(\mathrm{U} \mid \mathcal{F}_k))]$$

c. Montrer que :

$$\mathrm{E}(\mathrm{U} \mid \mathcal{F}_\infty) = \lim_{k\to\infty} \mathrm{E}(\mathrm{U} \mid \mathcal{F}_k) \quad \text{p.s.}$$

III. DEUX APPLICATIONS AUX PROCESSUS D'URNE

(On n'oubliera pas Préliminaires 4)

1° On considère l'élément $\mathrm{E}_{a,b}$ de $\mathcal{B}_\infty$

$$\mathrm{E}_{a,b} = \left\{ \underline{x} \in \mathbb{R}^{\mathbb{N}} : \liminf_{n\to\infty} x_n < a < b < \limsup_{n\to\infty} x_n \right\}.$$

Soit $\underline{\mathrm{X}}$ un processus d'urne de composition initiale (x_0, m), associé à la fonction f.

a. Montrer que :

$$\lim_{n\to\infty} \mathrm{E}(1_{\mathrm{E}_{a,b}} \circ \underline{\mathrm{X}} \mid \mathrm{X}_n) = 1_{\mathrm{E}_{a,b}} \circ \underline{\mathrm{X}} \quad \text{p.s.}$$

b. On suppose que $\mathrm{Q}^f_{x_0,m}(\mathrm{E}_{a,b}) > 0$. Montrer que pour tout $\varepsilon > 0$, pour tous c et d tels que $a < c < d < b$, il existe, pour une infinité d'entiers n, des compositions initiales (y_n, n), avec $c < y_n < d$, et telles que

$$\mathrm{Q}^f_{y_n,n}(\mathrm{E}_{a,b}) > 1 - \varepsilon.$$

2° Soit f une fonction de [0, 1] dans [0, 1], telle que, pour un p_0 de [0, 1], f vérifie

$$(f(t) - t)(t - p_0) \geqslant 0, \qquad \text{quel que soit } t \in [0, 1].$$

a. Montrer que si $\underline{\mathrm{X}}$ est un processus d'urne associé à f, la suite $(|\mathrm{X}_n - p_0|)_{n\in\mathbb{N}}$ est une sous-martingale.

b. Montrer que $\underline{\mathrm{X}}$ converge presque sûrement.

IV. PROCESSUS DE POLYA

On étudie ici un processus $\underline{X}$, de composition initiale (x_0, m), avec $m > 0$, associé à la fonction $f : t \rightsquigarrow f(t) = t$ (processus de Polya).

1° Montrer que pour $0 \leqslant k \leqslant n$:

$$P(S_n = k) = C_n^k \frac{\prod_{j=0}^{k-1} (mx_0 + j) \prod_{j=0}^{n-k-1} (m(1 - x_0) + j)}{\prod_{j=0}^{n-1} (m + j)}$$

et écrire une expression de

$$P(S_n < nt).$$

2° Montrer, à l'aide des préliminaires, de I, et de II ou (III + II), que $\underline{X}$ converge presque sûrement vers une v.a. X de loi $\beta(mx_0, m(1 - x_0))$.

3° Montrer que pour tout $\varepsilon > 0$

$$P\left(\sup_{n \in \mathbb{N}} | X_n - x_0 | > \varepsilon\right) \leqslant \frac{\sigma^2(X)}{\varepsilon^2}.$$

4° On donne quatre réels a, b, c et d vérifiant $a < c < d < b$, $0 \leqslant c < d \leqslant 1$ et on pose pour $\underline{x} \in \mathbb{R}^{\mathbb{N}}$:

$$\tau(\underline{x}) = \inf \{ n \in \mathbb{N} : x_n \notin [a, b] \}.$$

a. Montrer que :

$$\lim_{m \to \infty} \sup_{c \leqslant y \leqslant d} Q_{y,m} \{ \underline{x} : \tau(\underline{x}) < \infty \} = 0.$$

b. Soit $\underline{X}'$ un processus d'urne associé à une fonction g telle que $g(t) = t$ si $a \leqslant t \leqslant b$ et $\underline{X}$ un processus de Polya. Montrer que si $\underline{X}$ et $\underline{X}'$ ont même composition initiale (y, m) et si $c < y < d$, les v.a. $\tau(\underline{X})$ et $\tau(\underline{X}')$ ont même loi.

V. THÉORÈMES DE CONVERGENCE

1° Soit m un entier positif ou nul, y un réel compris entre 0 et 1, h_1 et h_2 deux fonctions définies sur [0, 1] et vérifiant, pour tout t dans [0, 1], $0 \leqslant h_1(t) \leqslant h_2(t) \leqslant 1$. Soit $(U_n)_{n \in \mathbb{N}}$ une suite indépendante de v.a. uniformes sur [0, 1]. On définit deux processus $\underline{V}$ et $\underline{W}$, par les relations de récurrence :

$$\begin{cases} V_0 = W_0 = y \\ (m + n + 1) V_{n+1} = (m + n) V_n + 1_{\{U_{n+1} \leqslant h_1(V_n)\}} \\ (m + n + 1) W_{n+1} = (m + n) W_n + 1_{\{U_{n+1} \leqslant h_2(W_n)\}} \end{cases} \quad n \in \mathbb{N}$$

Montrer que les processus $\underline{V}$ et $\underline{W}$ sont deux processus d'urnes de compositions initiales (y, m) associés respectivement à h_1 et h_2 et satisfont à :

$$V_n \leqslant W_n \qquad \text{quel que soit } n \in \mathbb{N}$$

2° Soit f une fonction de $[0, 1]$ dans $[0, 1]$ telle que pour tout intervalle non vide $]\alpha, \beta[$ inclus dans $[0, 1]$, il existe un intervalle $[a, b]$, (avec $\alpha < a < b < \beta$) sur lequel $f(t) - t$ ne change pas de signe. $\underline{X}$ désigne un processus d'urne associé à f, de composition initiale (x_0, m). On définit $E_{a, b}$ comme en III.1°, et τ comme en IV.4°; on pose :

$$\begin{aligned} g(t) &= f(t) \quad \text{si} \quad t \notin [a, b] \\ &= t \quad \text{si} \quad t \in [a, b] \, . \end{aligned}$$

a. Soit $\varepsilon > 0$. Montrer que si $Q^f_{x_0, m}(E_{\alpha, \beta}) > 0$ il existe pour $a < c < d < b$, pour une infinité d'entiers n, des compositions initiales (y_n, n) satisfaisant à $c < y_n < d$ et à

$$Q^g_{y_n, n}(\underline{x} \in \mathbb{R}^{\mathbb{N}} : \tau(\underline{x}) < +\infty) \geqslant Q^f_{y_n, n}(E_{a, b}) > 1 - \varepsilon.$$

b. Montrer que pour tous $\alpha < \beta$, $Q^f_{x_0, m}(E_{\alpha, \beta}) = 0$ et en déduire que tout processus d'urne $\underline{X}$ associé à f, de composition initiale (x_0, m) converge presque sûrement.

3° On suppose dans ce 3° que f est une fonction continue de $[0, 1]$ dans $[0, 1]$ et $\underline{X}$ un processus d'urne de composition initiale (x_0, m) associé à f.

a. Montrer que la suite $\underline{X}$ converge presque sûrement.

On pose pour $\varepsilon > 0$ et pour $\underline{x}$ suite de $[0, 1]^{\mathbb{N}}$:

$$\begin{aligned} A_\varepsilon &= \{ t \in [0, 1] : f(t) > t + \varepsilon \} \\ T(\underline{x}) &= \inf \{ n \geqslant 0 : x_n \notin A_\varepsilon \} \\ T_k(\underline{x}) &= \operatorname{Min}(T(\underline{x}), k) \qquad k \in \mathbb{N}^*. \end{aligned}$$

b. Montrer que si $x_0 \in A_\varepsilon$, pour tout $j \in \mathbb{N}^*$:

$$E\left((X_j - X_{j-1}) 1_{\{j \leqslant T_k(\underline{X})\}} / X_0, X_1, \ldots, X_{j-1}\right) \geqslant \frac{\varepsilon}{m + j} 1_{\{j \leqslant T_k(\underline{X})\}} \qquad \text{p.s.}$$

c. En déduire que pour tout m dans $\mathbb{N}$ et tout x_0 dans A_ε,

$$P(T(\underline{X}) = +\infty) = 0.$$

$$\left(\text{on pourra considérer } \sum_{j=1}^{\infty} (X_j - X_{j-1}) 1_{j \leqslant T_k(\underline{X})}\right).$$

d. Soit $X = \lim_{n \to \infty} X_n$ p.s. Montrer que pour toute composition initiale (x_0, m)

$$P(X \in A_\varepsilon) = 0;$$

puis en se souvenant des préliminaires 3°, établir que :

$$P(X \in \{ t : t = f(t) \} = 1.$$

4° On suppose dans ce 4° que f est une fonction de $[0, 1]$ dans $[0, 1]$ telle qu'il existe g, fonction continue de $[0, 1]$ dans $[0, 1]$, et p_0 $(0 < p_0 < 1)$ satisfaisant à

. $(f(t) - g(t))(p_0 - t) \geqslant 0$ pour tout $t \in [0, 1]$

.. $\{p_0\} = \{ t : g(t) = t \}$.

Soit $\underline{X}$ un processus d'urne associé à f et de composition initiale (x_0, m).

a. Montrer que pour tout $\delta > 0$:

$$P(\liminf_{n \to \infty} X_n > p_0 - \delta) = 1$$

(Utiliser la fonction Inf (g, h_δ), où h_δ est la fonction affine par morceaux, continue, et égale à 1 si $t < p_0 - \delta$, égale à 0 si $t > p_0$).

b. Montrer que la suite $\underline{X}$ converge presque sûrement vers p_0.

5° On suppose dans ce 5° que f est une fonction de $[0, 1]$ dans $[0, 1]$ telle que pour un p_0, $(0 < p_0 < 1)$, et pour tout $t \in [0, 1]$:

$$(f(t) - t)(t - p_0) \geqslant 0.$$

Soit X la limite d'un processus d'urne $\underline{X}$ associé à f et de composition initiale (x_0, m) avec $0 < x_0 < 1$ et $m > 0$. On pose pour $\lambda > 0$ et $z \in]0, 1[$:

$$\varphi_n(z) = \int_0^1 \frac{\Gamma(\lambda + 2)\,\Gamma(n)\, t^{\lambda p_0 + nz - 1}\,(1 - t)^{\lambda(1 - p_0) + n(1 - z) - 1}}{\Gamma(\lambda p_0 + 1)\,\Gamma(\lambda(1 - p_0) + 1)\,\Gamma(nz)\,\Gamma(n(1 - z))}\,dt.$$

a. Montrer que :

$$\mathrm{E}(\varphi_{n+m+1}(\mathrm{X}_{n+1}) / \mathrm{X}_n)$$

$$= \frac{\varphi_{n+m}(\mathrm{X}_n)}{\lambda + m + n}\left[\frac{f(\mathrm{X}_n)}{\mathrm{X}_n}(\lambda p_0 + (m + n)\mathrm{X}_n) + \frac{1 - f(\mathrm{X}_n)}{1 - \mathrm{X}_n}(\lambda(1 - p_0) + (m + n)(1 - \mathrm{X}_n))\right]$$

b. En déduire que $[\mathrm{E}(\varphi_{n+m}(\mathrm{X}_n))]_{n \in \mathbb{N}}$ est une suite décroissante.

c. Montrer que si $\lambda_n + \mu_n \to +\infty$ et si $\dfrac{\lambda_n}{\lambda_n + \mu_n} \to q$, la loi $\beta(\lambda_n, \mu_n)$ converge étroitement vers la mesure de Dirac en q, et que le maximum de sa densité est atteint en $\dfrac{\lambda_n - 1}{\lambda_n + \mu_n - 2}$, si $0 < q < 1$ et n assez grand.

d. Montrer que :

$$\lim_{n \to \infty} \mathrm{E}(\varphi_{n+m}(\mathrm{X}_n)) \geqslant \mathrm{E}\left(\frac{\Gamma(\lambda + 2)\,\mathrm{X}^{\lambda p_0}(1 - \mathrm{X})^{\lambda(1 - p_0)}}{\Gamma(\lambda p_0 + 1)\,\Gamma(\lambda(1 - p_0) + 1)}\right).$$

e. Montrer que :

$$\lim_{\lambda \to \infty} \varphi_m(x_0) = \frac{\Gamma(m)\, p_0^{\,m x_0 - 1}(1 - p_0)^{m(1 - x_0) - 1}}{\Gamma(m x_0)\,\Gamma(m(1 - x_0))}$$

en déduire que :

$$\mathrm{P}(\mathrm{X} = p_0) = 0.$$