

Espérance mathématique et simulation

par P.L. HENNEQUIN

INTRODUCTION :

L'enseignement actuel du calcul des probabilités distingue trois théories : la théorie très élémentaire se place sur les espaces probabilisés finis, la théorie générale utilise la théorie des mesures bornées et la théorie correspondante de l'intégration, entre les deux on développe, en général dans le premier cycle de l'enseignement supérieur, une théorie élémentaire qui manipule essentiellement les variables discrètes par des calculs de séries et les variables absolument continues par des manipulations d'intégrales de Riemann, bien souvent dans des exposés disjoints.

Par ailleurs, la "simulation" repose sur l'utilisation systématique d'une suite "ayant les mêmes propriétés" que la suite des réalisations d'une variable uniforme sur $]0,1[$. Le rapprochement de ces deux situations nous conduit à une théorie élémentaire où on utilise systématiquement comme espace de base $]0,1[$, $\mathcal{B}$, L) où $\mathcal{B}$ désigne la tribu des boréliens de $]0,1[$ et L la mesure de Lebesgue (prolongement de la longueur).

Nous supposons définie la fonction de répartition d'une variable aléatoire réelle et connues ses propriétés, en particulier le théorème de prolongement (qui permet entre autres de définir L). La restriction à des intégrales de Riemann peut sembler alourdir le formalisme par rapport à la théorie générale de l'intégration ; il se justifie pour l'étude des suites équiréparties (cf. III).

- ESPERANCE MATHÉMATIQUE

PROPOSITION 1 : $\forall X$, v.a.r. sur $(\Omega, \mathcal{H}, P)$, il existe une fonction croissante sur $]0,1[$, $\bar{X}$ qui, considérée comme v.a.r. sur $]0,1[$, $\mathcal{B}$, L), a

même répartition que X .

Tant qu'on ne s'intéressera qu'à une variable X , on utilisera $\bar{X}$ pour simuler X .

DEMONSTRATION : Soit $F(x) = P\{\omega/X(\omega) \leq x\}$ la fonction de répartition de X .

Définissons $\bar{X}(x) = \{\inf y/F(y) \geq x\}$

a) $\bar{X}$ est croissante, car $x \leq x'$ et $F(y) \geq x' \implies F(y) \geq x$ définie sur $]0, 1[$ et continue à gauche (et c'est la seule ayant ses propriétés).

b) $\bar{X}$ n'est pas, en général une fonction réciproque de F mais l'on a

$$\bar{X}(x) \leq t \iff x \leq F(t)$$

En effet, $\bar{X}(x) \leq t \implies \forall \epsilon > 0, \exists y \leq t + \epsilon, F(y) \geq x$

$$\implies x \leq F(t + \epsilon) \text{ et } F \text{ continue à droite}$$

$$\implies x \leq F(t)$$

$x < F(t) \implies \bar{X}(x) \leq t$ par définition de $\bar{X}(x)$.

On a donc $L\{x/\bar{X}(x) \leq t\} = L\{x/x \leq F(t)\} = F(t)$ puisque $0 \leq F(t) \leq 1$.

DEFINITION 1 :

On dit que X possède une espérance mathématique EX si $\int_0^1 \bar{X}(x) dx$ existe comme intégrale de Riemann impropre, et on pose :

$$EX = \int_0^1 \bar{X}(x) dx$$

REMARQUES : 1. $\bar{X}$ est intégrable Riemann, puisque monotone, sur tout intervalle $[a, b]$ tel que $0 < a \leq b < 1$. La condition d'existence équivaut donc à l'existence de

$$\lim_{\substack{a \downarrow 0 \\ b \uparrow 1}} \int_a^b \bar{X}(x) dx$$

et porte sur le comportement de $\bar{X}$ au voisinage de 0 et de 1, donc sur le comportement de $F(x)$ quand $x \rightarrow +\infty$ ou $-\infty$.

2. Dans le cas très élémentaire où X prend un nombre fini de valeurs $X = \sum_{i=1}^n x_i \cdot 1_{A_i}$, les x_i étant supposés numérotés dans l'ordre croissant,

avec $P(A_1) = p_1$,

$\bar{X}$ prend la valeur x_i sur $\left[\sum_{k=0}^{i-1} p_k, \sum_{k=0}^i p_k \right]$ et on retrouve la formule

$$\int_0^1 \bar{X}(x) dx = \sum_{i=1}^n x_i p_i$$

PROPOSITION 2 : Soit X une v.a.r. sur $(\Omega, \mathcal{H}, P)$ et g une application ^{mesurable} de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\int_0^1 g(\bar{X}(x)) dx$ existe comme intégrale de Riemann impropre absolument convergente, alors $E(g(X))$ existe et on a :

$$E(g(X)) = \int_0^1 g(\bar{X}(x)) dx$$

Cette formule permet donc de calculer $E(g(X))$ par une intégrale sur $(0, 1)$ faisant intervenir $\bar{X}$ et non $\overline{g(X)}$ dont le calcul peut être plus compliqué.

DEMONSTRATION

a) $g(X)$ comme v.a. sur $(\Omega, \mathcal{H}, P)$ a même répartition que $g(\bar{X})$ sur $[0, 1[, \mathcal{B}, L)$.

$$\begin{aligned} \text{En effet, } P\{\omega/g(X(\omega)) \in B\} &= P\{\omega/X(\omega) \in g^{-1}(B)\} \\ &= L\{x/\bar{X}(x) \in g^{-1}(B)\} \\ &= L\{g(\bar{X}(x)) \in B\} \end{aligned}$$

$\forall B$ Borélien (on utilise ici le théorème de prolongement).

b) Il suffit de démontrer que

$$\int_0^1 g(\bar{X}(x)) dx = \int_0^1 \overline{g(X)}(x) dx$$

puisque le second membre est égal à

$$\int_0^1 \overline{g(X)}(x) dx = E(g(X)) \quad \text{par définition.}$$

Posons donc $g \circ \bar{X} = h$; nous devons montrer que :

$$\int_0^1 h(x) dx = \int_0^1 \overline{h}(x) dx$$

où le premier membre existe comme intégrale de Riemann impropre.

Supposons d'abord h étagée sur $]0, 1]$.

$$h(x) = \sum_{i=1}^n h_i \cdot \mathbb{I}_{[a_i, a_{i+1})}(x)$$

alors

$$\int_0^1 h(x) dx = \sum_{i=1}^n h_i (a_{i+1} - a_i)$$

et la fonction de répartition H de h est donnée par :

$$H(x) = \sum_{h_i \leq x} (a_{i+1} - a_i)$$

Soit $\{h'_j\}$ l'ensemble des valeurs de la suite $\{h_i\}$ numérotées dans l'ordre croissant et

$$p_j = \sum_{i: h_i = h'_j} (a_{i+1} - a_i)$$

alors

$$\bar{h}(x) = h'_j \text{ si } \sum_{k=1}^{j-1} p_k \leq x < \sum_{k=1}^j p_k$$

de sorte que

$$\int_0^1 \bar{h}(x) dx = \sum_j h'_j p_j = \sum_{i=1}^n h_i (a_{i+1} - a_i)$$

(associativité et commutativité pour une somme finie).

c) Si $h = g \circ \bar{X}$ est intégrable Riemann sur $[0, 1]$, $\forall \varepsilon > 0$, il existe ϕ et ψ en escalier telles que $\phi \leq h \leq \psi$ et

$$\int_0^1 \psi(x) dx \leq \int_0^1 \phi(x) dx + \varepsilon$$

Les fonctions de répartition correspondantes satisfont

$$\Psi \leq H \leq \Phi$$

donc

$$\bar{\phi} \leq \bar{h} \leq \bar{\psi}$$

Comme, par b),

$$\int_0^1 \bar{\phi}(x) dx = \int_0^1 \phi(x) dx \text{ et } \int_0^1 \bar{\psi}(x) dx = \int_0^1 \psi(x) dx,$$

on en déduit :

$$\forall \varepsilon > 0, \left| \int_0^1 h(x) dx - \int_0^1 \bar{h}(x) dx \right| \leq \varepsilon, \text{ d'où l'égalité.}$$

d) Supposons $h \geq 0$ et $\int_0^1 h(x) dx$ définie comme intégrale de Riemann impropre. Soit

$$h_N(x) = \begin{cases} h(x) & \text{si } h(x) < N \\ N & \text{si } h(x) \geq N \end{cases}$$

On a $H_N(t) = H(t)$ pour $t < N$ et $H_N(t) = 1$ pour $t \geq N$. On en déduit :

$$\bar{h}_N(x) = \bar{h}(x) \text{ si } \bar{h}(x) < N \text{ et } \bar{h}_N(x) = N \text{ si } \bar{h}(x) \geq N,$$

donc $\bar{h}_N = [\bar{h}]_N$

Comme $\int_0^1 h_N(x) dx \rightarrow \int_0^1 h(x) dx$ et $\int_0^1 \bar{h}_N(x) dx \rightarrow \int_0^1 \bar{h}(x) dx$ quand $N \rightarrow \infty$,

on déduit de c) :

$$\int_0^1 h_N(x) dx = \int_0^1 \bar{h}_N(x) dx$$

d'où

$$\int_0^1 h(x) dx = \int_0^1 \bar{h}(x) dx.$$

e) Si $\int_0^1 h(x) dx$ existe comme intégrale de Riemann impropre, posons :

$$h^+ = \sup(h, 0), \quad h^- = \sup(-h, 0)$$

alors

$$h = h^+ - h^-$$

et, comme en d) :

$$\bar{h}^+ = \bar{h}^+ \text{ et } \bar{h}^- = \bar{h}^-.$$

On déduit de d) :

$$\int_0^1 h^+(x) dx = \int_0^1 \bar{h}^+(x) dx \text{ et } \int_0^1 h^-(x) dx = \int_0^1 \bar{h}^-(x) dx$$

d'où

$$\int_0^1 h(x) dx = \int_0^1 (h^+(x) - h^-(x)) dx = \int_0^1 \bar{h}(x) dx.$$

CAS PARTICULIERS

1 - Cas discret $X = \sum_{i=1}^{\infty} x_i 1_{A_i} \quad P(A_i) = p_i$

Pour construire et manipuler $\bar{X}$, il faudrait pouvoir numéroté les x_i dans l'ordre croissant, ce qui n'est pas toujours possible (par exemple, si $\{x_i\} = 0$). La simulation nous suggère d'introduire $\bar{X}'$ qui a la même répartition que X et qui est ainsi définie :

Supposons les p_i numérotés dans l'ordre croissant, ce qui est loisible, car $\forall \varepsilon > 0$, il y a un nombre fini de $p_i \geq \varepsilon$ puisque $p_i > 0$ et que $\sum p_i = 1$.

Alors

$$\bar{X}'(x) = x_i \text{ si } \sum_{k=1}^{i-1} p_k < x \leq \sum_{k=1}^i p_k;$$

$\bar{X}'$ et $\bar{X}$ ont la même répartition, donc :

$$EX = \int_0^1 \bar{X}'(x) dx = \sum_1^{\infty} p_i x_i$$

si et seulement si cette série converge absolument, et comme

$$g(X) = \sum g(x_i) 1_{A_i},$$

$$E(g(X)) = \sum_1^{\infty} p_i g(x_i)$$

si cette série converge absolument.

2 - Cas absolument continu

La fonction de répartition F de X est définie par :

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{avec } f \geq 0 \text{ (Intégrale de Riemann impropre)}$$

On veut calculer $E(g(X))$ quand elle existe.

PROPOSITION 3 : Sous ces hypothèses, $E(g(X)) = \int_{-\infty}^{+\infty} g(t) f(t) dt$.

DEMONSTRATION : Les étapes sont les mêmes que dans la démonstration de la proposition 2.

$$a) \text{ Si } g(\bar{X}) = \sum_{i=1}^n h_i 1_{[a_i, a_{i+1}]} \\ E(g(X)) = \sum_{i=1}^n h_i (a_{i+1} - a_i)$$

La continuité de F implique que $\bar{X}$ est strictement croissante ; si $g \circ \bar{X}$ est étagée, la restriction de g à $\bar{X}([0, 1])$ est étagée :

$$t \in \bar{X}([0, 1]) \Rightarrow g(t) = \sum_{i=1}^n h_i 1_{[b_i, b_{i+1}]}(t) \quad \text{avec } b_i = \bar{X}(a_i).$$

Alors

$$\begin{aligned} \int_{-\infty}^{+\infty} g(t) f(t) dt &= \int_{\bar{X}([0, 1])} g(t) f(t) dt \\ &= \sum_{i=1}^n h_i \int_{b_i}^{b_{i+1}} f(t) dt \\ &= \sum_{i=1}^n h_i (F(b_{i+1}) - F(b_i)) \end{aligned}$$

$$= \sum_{i=1}^n h_i (a_{i+1} - a_i) = E(g(X))$$

b) Si $g(\bar{X})$ est intégrable Riemann sur $]0, 1[$, $\forall \varepsilon > 0$ la restriction de g à $\bar{X}(]0, 1[)$ est encadrée par deux fonctions étagées ϕ et ψ telles que :

$$\int_0^1 \phi(\bar{X}(x)) dx \geq \int_0^1 \psi(\bar{X}(x)) dx - \varepsilon$$

Comme $f \geq 0$, si $t \in \bar{X}(]0, 1[)$,

$$\phi(t) f(t) \leq g(t) f(t) \leq \psi(t) f(t)$$

et

$$\int_{-\infty}^{+\infty} \phi(t) f(t) dt = \int_{\bar{X}(]0, 1[)} \phi(t) f(t) dt = \int_0^1 \phi(\bar{X}(x)) f(x) dx$$

de même pour ψ .

On en déduit l'existence de $\int_{-\infty}^{+\infty} f(t) g(t) dt$ et son égalité à $E(g(X))$.

c) Si $g \geq 0$, on pose $g_N = \inf(g, N)$, et on déduit de b) :

$$E(g_N(X)) = \int_{-\infty}^{+\infty} f(t) g_N(t) dt$$

$$E(g_N(X)) = \int_{-\infty}^{+\infty} f(t) g(t) dt \quad \text{en faisant tendre } N \text{ vers } \infty.$$

Enfin, si g est absolument intégrable, on déduit l'égalité pour

$g = g^+ - g^-$ de l'égalité pour g^+ et g^- .

II - FAMILLES DE VARIABLES ALÉATOIRES REELLES

Jusqu'à maintenant, nous avons laissé sous silence deux propriétés fondamentales de l'espérance mathématique : la linéarité et la multiplicité pour des variables indépendantes. Leur énoncé introduit un couple (X, Y) de variables aléatoires réelles. La définition d'un tel couple comme application mesurable de $(\Omega, \mathcal{A})$ dans $(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2})$ ne soulève aucune difficulté, mais celles-ci apparaissent dès qu'on veut simuler un tel couple à partir par exemple de la fonction de répartition, car $\mathbb{R}^2$ n'est plus, comme $\mathbb{R}$, totalement ordonné. Ceci nous amène à donner une définition, d'apparence

restrictive et dont seule une théorie générale (Radon-Nikodym et probabilités conditionnelles régulières) peut montrer qu'elle recouvre le cas le plus général (en remplaçant les conditions de Riemann intégrabilité par des conditions d'intégrabilité).

DEFINITION 2 :

Nous dirons qu'une famille de v.a.r. $(X_1, X_2, \dots, X_n)$ sur $(\Omega, \mathcal{A}, P)$ est simulable s'il existe n applications $\bar{X}_i, \bar{X}_i$ de $\mathbb{R}^1$ dans $]0, 1[$ telles que

$\forall x_1, x_2, \dots, x_{i-1}, x_i \rightarrow \bar{X}_i(x_1, x_2, \dots, x_i)$ soit croissante,
et que $\forall a_i, 0 < a_i < 1,$

$$P(X_i \leq a_i, 1 \leq i \leq n) = \int_{\bar{X}_1 \leq a_1} dx_1 \int_{\bar{X}_2(x_1, x_2) \leq a_2} dx_2 \dots \\ \dots \int_{\bar{X}_n(x_1, x_2, \dots, x_n) \leq a_n} dx_n$$

le second membre existant comme intégrale de Riemann multiple.

Alors pour simuler $X_1, X_2, \dots, X_n$, on utilisera $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n$, définies à partir de $x_1, x_2, \dots, x_n$, résultat de n tirages indépendants d'une variable uniforme sur $]0, 1[$.

REMARQUE : Si $(X_1, X_2, \dots, X_n)$ est indépendante, elle est simulable (Prendre $\bar{X}_i(x_1, x_2, \dots, x_i) = \bar{X}_i(x_i)$).

PROPOSITION 4 : Si $\int_{]0,1[^n} g(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) dx$ existe comme intégrale de Riemann impropre absolument convergente, elle est égale à

$$E(g(X_1, X_2, \dots, X_n))$$

Ce théorème, extension à $\mathbb{R}^n$ de la proposition 2, se démontre de façon tout à fait analogue. Posant $g(\bar{X}_1, \bar{X}_2, \dots, \bar{X}_n) = h(x_1, x_2, \dots, x_n)$, on est ramené à démontrer que

$$\int_{]0,1[^n} h(x) dx = \int_0^1 \bar{h}(t) dt,$$

où $\bar{h}$ est la fonction croissante sur $]0, 1[$ de même répartition que h . On le montre successivement pour h en escalier, pour h intégrable Riemann sur $]0, 1[$, pour $h \geq 0$.

COROLLAIRE 1 : $E(X + Y) = EX + EY$

COROLLAIRE 2 : Si X et Y sont indépendants, $EXY = EX EY$

DEMONSTRATION : On déduit de la proposition 4, en prenant successivement

$g(x, y) = x + y, x, y, xy,$

$$E(X + Y) = \int_0^1 \int_0^1 (\bar{X} + \bar{Y}) dx dy = \int_0^1 \int_0^1 \bar{X} dx dy + \int_0^1 \int_0^1 \bar{Y} dx dy = EX + EY.$$

$$\begin{aligned} EXY &= \int_0^1 \int_0^1 \bar{X}(x) \bar{Y}(x, y) dx dy \\ &= \int_0^1 \bar{X}(x) dx \int_0^1 \bar{Y}(x, y) dy \end{aligned}$$

si X et Y sont indépendants, on peut choisir $\bar{Y}$ ne dépendant pas de x .

D'où

$$\begin{aligned} EXY &= \int_0^1 \bar{X}(x) dx \cdot \int_0^1 \bar{Y}(y) dy \\ &= \int_0^1 \int_0^1 \bar{X}(x) dx dy \cdot \int_0^1 \int_0^1 \bar{Y}(y) dx dy = EX EY. \end{aligned}$$

III - SIMULATION

DEFINITION 3 :

On dit qu'une suite de nombres réels $\{x_n\}$ possède une répartition F si

$$\forall a \in \mathbb{R} \quad \frac{\text{Card} \{k/x_k \leq a, 1 \leq k \leq n\}}{n}$$

converge vers une limite $F(a)$. On dit que $\{x_n\}$ est équirépartie sur $]0, 1[$ si $F(a) = a \forall a$ tel que $0 < a < 1$.

REMARQUES :

1 - Si F est une fonction de répartition sur $\mathbb{R}$ et si x_n est équiré-

partie sur $]0, 1[$, $\bar{X}$ associée à F par la proposition 1, alors $\bar{X}(x_n)$ a pour répartition F .

En effet, $\bar{X}(x_n) \leq a \iff x_n \leq F(a)$. On utilisera en particulier la suite $\bar{X}(x_n)$ pour simuler la v.a. X .

2 - Il peut se faire que, dans la définition 3, F ne soit pas une fonction de répartition sur $\mathbb{R}$, mais sur $\bar{\mathbb{R}}$: par exemple si $x_n \rightarrow \infty$, $F(a) = 0$, $\forall a \in \mathbb{R}$.

Il existe des suites équiréparties (par exemple, $x_n = na \pmod{1}$ avec a irrationnel) et en particulier on a la

PROPOSITION 5 : Soit $\{X_n\}$ une suite de variables aléatoires indépendantes uniformes sur $]0, 1[$; alors pour presque tout ω , la suite $\{X_n(\omega)\}$ est équirépartie.

DEMONSTRATION : Par la loi forte des grands nombres,

$\forall a, 0 \leq a \leq 1$

$$\frac{1}{n} \sum_{i=1}^n 1_{]-\infty, a]}(X_i(\omega)) \rightarrow a$$

sauf pour $\omega \in N(a)$ avec $P(N(a)) = 0$, car

$$E(1_{]-\infty, a]}(X_1)) = a$$

et les $1_{]-\infty, a]}(X_i)$ sont indépendantes.

On en déduit que la propriété est vraie quelque soit a rationnel de $[0, 1]$ si $\omega \notin N = \bigcup_{a \in \mathbb{Q} \cap [0, 1]} N(a)$; puisque $\mathbb{Q}$ est dénombrable, on a encore $P(N) = 0$.

Soit alors a réel quelconque de $[0, 1]$. $\forall \epsilon > 0$, il existe deux rationnels q et q' tels que $0 \leq q \leq a \leq q' \leq q + \epsilon \leq 1$ et

$$1_{]-\infty, q]} \leq 1_{]-\infty, a]} \leq 1_{]-\infty, q']}$$

Comme

$$\frac{1}{n} \sum_{i=1}^n 1_{]-\infty, q]}(X_k(\omega)) \rightarrow q \text{ et } \frac{1}{n} \sum_{i=1}^n 1_{]-\infty, q']}(X_k(\omega)) \rightarrow q',$$

si $\omega \notin N$, on en déduit que pour n assez grand

$$\left| \frac{1}{n} \sum_{k=1}^n 1_{]-\infty, a]}(X_k(\omega)) - a \right| \leq 2\varepsilon$$

La suite $\{X_n(\omega)\}$ est donc équirépartie pour $\omega \notin N$.

REMARQUE : Si $\{x_n\}$ est équirépartie, il résulte immédiatement de la défini-

tion que :

$$\frac{1}{n} \sum_{k=1}^n 1_B(x_k) \rightarrow L(B)$$

pour B intervalle ou réunion finie d'intervalles de $]0, 1[$, mais cette propriété peut être fausse pour une réunion dénombrable, a fortiori pour B Borélien quelconque de $[0, 1]$: Il suffit de prendre pour B le complémentaire de l'ensemble $\{x_n\}$, la limite est alors 0 et non $L(B) = 1$.

Ainsi, si les X_k sont indépendantes, uniformes sur $[0, 1]$, $\forall B$ borélien fixé,

$$\frac{1}{n} \sum_{k=1}^n 1_B(X_k(\omega)) \rightarrow L(B) \text{ presque sûrement,}$$

mais presque sûrement $\exists B$ tel que

$$\frac{1}{n} \sum_{k=1}^n 1_B(X_k(\omega)) \neq L(B)$$

L'utilisation des suites équiréparties permet le calcul d'intégrales de Riemann propres grâce à la proposition suivante.

PROPOSITION 6 : Si $\{x_n\}$ est équirépartie sur $]0, 1[$ et f Riemann intégrable sur $[0, 1]$ alors

$$\frac{1}{n} \sum_{k=1}^n f(x_k) \rightarrow \int_0^1 f(x) dx.$$

En effet, f peut être encadré par deux fonctions en escalier ϕ et ψ

telles que :

$$\int_0^1 \phi(x) dx - \int_0^1 \psi(x) dx \leq \varepsilon;$$

$\int_0^1 \phi(x) dx$ et $\int_0^1 \psi(x) dx$ encadrent $\int_0^1 f(x) dx$ et en diffèrent de moins de ε .

Alors $\frac{1}{n} \sum_1^n f(x_k)$ est encadrée par $\frac{1}{n} \sum_1^n \phi(x_k)$ et $\frac{1}{n} \sum_1^n \psi(x_k)$ qui convergent respectivement vers $\int_0^1 \phi(x) dx$ et $\int_0^1 \psi(x) dx$.

En effet, si

$$\phi = \sum_i \phi_i \cdot 1_{[a_i, a_{i+1}[}$$

$$\begin{aligned} \frac{1}{n} \sum_1^n \phi(x_k) &= \sum_i \phi_i \cdot \frac{1}{n} \sum_k 1_{[a_i, a_{i+1}[}(x_k) \rightarrow \sum_i \phi_i (a_{i+1} - a_i) \\ &= \int_0^1 \phi(x) dx \end{aligned}$$

REMARQUE : La proposition 6 peut, jointe aux propositions 2 et 3 permettre des calculs approchés d'intégrales, mais on prendra bien soin de vérifier qu'on se ramène à une intégrale de Riemann propre sur $[0, 1]$. En effet, on a le

CONTRE-EXEMPLE : Il peut se faire que $\int_0^1 f(x) dx$ existe comme intégrale de Riemann impropre absolument convergente et que $\frac{1}{n} \sum_1^n f(x_k)$ ne converge pas vers $\int_0^1 f(x) dx$. Toutefois, si $f \geq 0$,

$$\liminf \frac{1}{n} \sum_1^n f(x_k) \geq \int_0^1 f(x) dx.$$

Supposons en effet que nous nous donnions une suite $\{t_n\}$ équirépartie sur $[0, 1]$. Remarquons déjà que notre définition n'exclut pas $t_n = 0$ ou 1 et qu'il pourrait se faire que $f(0)$ ou $f(1)$ ne soient pas définies ; nous supposons que $\forall n, 0 < t_n < 1$ et que f est finie sur $]0, 1[$ de sorte que $\frac{1}{n} \sum_1^n f(t_k)$ est bien définie.

Considérons la suite x_n définie par :

$$\begin{aligned} x_n &= t_n & \text{si } n \neq 2^j \\ x_n &= \frac{1}{e^{2^n}} & \text{si } n = 2^j \end{aligned} \quad j = 1, 2, \dots$$

$(x_n)_{n=1}^\infty$ est équirépartie.

En effet, $\forall a, 0 \leq a \leq 1$,

$$1_{]-\infty, a]}(x_k) = 1_{]-\infty, a]}(t_k) \quad \text{si } k \neq 2^j$$

et

$$|1_{]-\infty, a]}(x_k) - 1_{]-\infty, a]}(t_k)| \leq 1 \quad \text{sinon.}$$

Alors $\frac{1}{n} \sum_{k=1}^n |1]_{-\infty, a}] (x_k) - 1]_{-\infty, a}] (t_k)| \leq \frac{1}{n} \log_2 n \rightarrow 0$ quand $n \rightarrow \infty$.

Donc $\frac{1}{n} \sum_{k=1}^n 1]_{-\infty, a}] (x_k)$ converge comme $\frac{1}{n} \sum_{k=1}^n 1]_{-\infty, a}] (t_k)$ vers a .

Prenons alors $f(x) = -\log x$ de sorte que $\int_0^1 -\log x \, dx = 1$,

tandis que

$$\begin{aligned} \frac{1}{n} \sum_{k=1}^n (-\log x_k) &\geq \frac{1}{n} \sum_{2^{k_0} \leq k \leq n} -\log x_{2^k} \\ &\geq \frac{1}{2^{k_0+1}} \cdot 2^{2^{k_0}} \rightarrow \infty \quad \text{quand } k_0 \rightarrow \infty \end{aligned}$$

où $2^{k_0} \leq n < 2^{k_0+1}$.

Si $f \cdot 1]_{\varepsilon, 1-\varepsilon}]$ est intégrable Riemann et $f \geq 0$,

$$\frac{1}{n} \sum_{k=1}^n f(x_k) \geq \frac{1}{n} \sum_{k=1}^n f(x_k) 1]_{\varepsilon, 1-\varepsilon}] (x_k)$$

et le second membre tend vers $\int_{\varepsilon}^{1-\varepsilon} f(t) \, dt$ qui est arbitrairement voisin de $\int_0^1 f(x) \, dx$.

CONSEQUENCE : Si on veut calculer à l'aide d'une suite équirépartie une intégrale de Riemann impropre absolument convergente, on procède en trois étapes:

- 1° On se ramène à $f \geq 0$
- 2° Pour $\varepsilon > 0$ donné, on détermine η tel que

$$\int_0^{\eta} f(x) \, dx + \int_{1-\eta}^1 f(x) \, dx \leq \varepsilon$$

- 3° On approche $\int_{\eta}^{1-\eta} f(x) \, dx$ par $\frac{1}{n} \sum_{k=1}^n f(x_k) 1]_{\eta, 1-\eta}] (x_k)$

En pratique, supposons qu'on ait $f \geq 0$ et qu'on ait établi la convergence de $\int_0^1 f(x) \, dx$ par majoration $|f(x)| \leq K |x|^{-(1+\alpha)}$, $\alpha < 0$, alors il suffira d'éliminer dans le calcul par un test convenable les valeurs de n telles que le calcul de $f(x_n)$ conduise à un dépassement de capacité.

IV - SIMULATION DE VECTEURS

Au paragraphe III, nous avons simulé une variable aléatoire X de répartition F par une suite de même répartition. On peut étendre sans difficulté ces définitions à des suites et variables aléatoires dans $\mathbb{R}^p$. Ainsi une suite $\{V_k\}$ de vecteurs de $\mathbb{R}^p$ sera équirépartie sur $[0, 1]^p$ si :

$$\frac{1}{n} \sum_{k=1}^n 1_C(V_k) \rightarrow v(C)$$

pour tout polyèdre C de $[0, 1]^p$ de volume $v(C)$.

La suite $\{V_n\}$ permettra, grâce à la définition 3 et la proposition 5, de simuler une famille $(X_1, X_2, \dots, X_p)$ de variables simulables, par la suite

$$(\bar{X}_1(v_n^1), \bar{X}_2(v_n^1, v_n^2), \dots, \bar{X}_n(v_n^1, v_n^2, \dots, v_n^n))$$

En fait, la suite $\{X_n\}$ de variables indépendantes uniformes sur $[0, 1]$ possède la propriété suivante (qui résulte immédiatement de la loi forte des grands nombres) :

$\forall p$ fixé, la suite de vecteurs $V_n^{(p)}$ de composantes

$$V_n^i = X_{np+i}(\omega) \quad 1 \leq i \leq p$$

est équirépartie sur $[0, 1]^p$ pour $\omega \notin N_p$ avec $P(N_p) = 0$.

Comme on a aussi $P(\cup N_p) = 0$, on en déduit que presque sûrement

$\forall p, V_n^{(p)}$ est équirépartie.

DEFINITION 4 :

Nous dirons qu'une suite $\{x_n\}$ de réels de $[0, 1]$ est parfaitement équirépartie si $\forall p$, la suite de vecteurs $\{v_n^{(p)}\}$ est équirépartie sur $[0, 1]^p$ où $v_n^{(p)}$ a pour composante $v_n^i = x_{np+i}$

Les suites parfaitement équiréparties permettent donc de simuler les vecteurs aléatoires simulables de dimension quelconque.

Faisons une dernière remarque dont l'intérêt pratique est nul : il y a bijection entre l'ensemble des variables uniformes sur $]0, 1[$ et l'ensemble des suites de variables uniformes indépendantes sur $]0, 1[$, autrement dit, ce n'est pas plus compliqué de trier une infinité de valeurs qu'une seule !

En effet, soit $X(\omega)$ uniforme sur $]0, 1[$. Choisissons une base de numération a . Associons à $X(\omega)$ la suite de ses "décimales" en base a (en excluant les suites qui comportent une infinité de fois consécutives le chiffre $a - 1$).

Soit $a^i(\omega)$ la $i^{\text{ième}}$.

$a^i(\omega)$ est uniforme sur $0, 1, 2, \dots, a-1$ et les a^i sont indépendants.

Prenons ensuite la suite des décimales de rang impair $2p + 1$. C'est la suite des décimales d'un nouveau nombre $Y_1(\omega)$ uniforme sur $]0, 1[$.

La suite des décimales de rang $2(2p + 1)$ est la suite des décimales d'un second nombre $Y_2(\omega)$ uniforme sur $]0, 1[$ indépendant de Y_1 .

La suite des décimales de rang $2^n(2p + 1)$ définit de même Y_{n-1} .

$\{Y_k\}$ est une suite de v.a. indépendantes uniformes sur $]0, 1[$. Inversement, les $Y_k(\omega)$ définissent $X(\omega)$ uniforme sur $]0, 1[$.