13 - DEUX RÉSULTATS SUR LA CONVERGENCE EN PROBABILITÉ (*5 pages*)

Les X_n et X étant des variables aléatoires réelles sur $(\Omega,\mathcal{T})$, on supposera dans tout le texte que $X_n \to X$ pour une probabilité p sur $(\Omega,\mathcal{T})$: pour tout $\alpha > 0$, $\lim_n p(|X_n - X| \le \alpha) = 1$. Cette convergence est notée $X_n \xrightarrow{p} X$. La lettre h, indicée ou non, désignera toujours une fonction de $\mathbb{R}$ dans $\mathbb{R}$, au moins *borélienne*, c'est-à-dite mesurable pour la tribu des boréliens de $\mathbb{R}$, pour que les composés $h \circ X$ et les $h \circ X_n$ soient des variables aléatoire sur $(\Omega,\mathcal{T})$ (ou $\mathcal{T}$ -mesurables sur Ω).

Le but de cet article est de cerner l'ensemble non vide $\mathcal{H}$ des fonctions h pour lesquelles $h \circ X_n \xrightarrow{p} h \circ X$. Dans la partie **A**, on démontre que toutes les fonctions continues appartiennent à $\mathcal{H}$. Dans la partie **B**, on prouve, moyennant une hypothèse supplémentaire sur la suite (X_n), que toute fonction borélienne est dans $\mathcal{H}$ (ce résultat porte le nom de « lemme de Slutsky » dans la littérature probabiliste russe). Les démonstrations données ici ne sont pas les plus courtes, mais elles ne supposent pas de connaissance sur la convergence des variables aléatoires.

PARTIE A

1) $\mathcal{H}$ est un espace vectoriel réel

On vérifie la stabilité de $\mathcal{H}$ par les deux lois :

a) Si le réel k est non nul, $|k.h \circ X_n - k.h \circ X| \le \alpha = |h \circ X_n - h \circ X| \le \alpha / k$; la fonction $k.h$ est donc dans $\mathcal{H}$ si h l'est.

b) Comme $(h_1 + h_2) \circ X_n - (h_1 + h_2) \circ X = h_1 \circ X_n - h_1 \circ X + h_2 \circ X_n - h_2 \circ X$, on peut écrire $(|h_1 \circ X_n - h_1 \circ X| \le \alpha / 2) \cap (|h_2 \circ X_n - h_2 \circ X| \le \alpha / 2) \subset |(h_1 + h_2) \circ X_n - (h_1 + h_2) \circ X| \le \alpha$; reste à vérifier que si deux suite d'événements (A_n) et (B_n) vérifient $\lim_n p(A_n) = \lim_n p(B_n) = 1$, alors $\lim_n p(A_n \cap B_n) = 1$, or cette affirmation est une conséquence de l'inégalité $p(A_n \cap B_n) \ge p(A_n) + p(B_n) - 1$.

2) $h \circ X_n \to h \circ X$ pour la probabilité p si h est uniformément continue

D'abord, il existe $\beta > 0$ tel que $|X_n(\omega) - X(\omega)| \le \beta$ implique $|h \circ X_n(\omega) - h \circ X(\omega)| \le \alpha$ pour tout ω de Ω ; on en déduit $p(|X_n - X| \le \beta) \le p(|h \circ X_n - h \circ X| \le \alpha)$. Ensuite, quand *n* tend vers $+\infty$, $p(|X_n - X| \le \beta)$ tend vers 1, et il en est de même de $p(|h \circ X_n - h \circ X| \le \alpha)$ puisque $p(|X_n - X| \le \beta) \le p(|h \circ X_n - h \circ X| \le \alpha) \le 1$.

3) $h \circ X_n \xrightarrow{p} h \circ X$ pour la probabilité p si h est continue

On se fixe $\alpha > 0$ et $\varepsilon > 0$.

a) *Il existe N tel qu'à partir d'un certain rang* n_0, $p(|X| > N) + p(|X_n| > N) \le \varepsilon$

Pour relier $p(|X_n| > N)$ et $p(|X| > N)$, on observe que $|X| \le N$ et $|X_n - X| \le \alpha$ impliquent $|X_n| \le N + \alpha$, d'où on déduit $p(|X| \le N \text{ et } |X_n - X| \le \alpha) \le p(|X_n| \le N + \alpha)$. Avec l'inégalité $p(A \cap B) \ge p(A) + p(B) - 1$, on obtient

$$p(|X| \le N) + p(|X_n - X| \le \alpha) - 1 \le p(|X_n| \le N + \alpha).$$

Comme $X_n \xrightarrow{p} X$, à partir d'un certain rang,

$$p(|X| \leq N) \leq p(|X_n| \leq N+\alpha)+\varepsilon \text{ et donc } p(|X_n| > N) \leq p(|X| > N+\alpha)+\varepsilon .$$

Il suffit donc de choisir *N* tel que $p(|X| > N)+p(|X| > N+\alpha) \leq \varepsilon$, ce qui est possible puisque

$$\lim_{N\to+\infty} p(|X| > N) = 0 .$$

b) Sur $[-N,N]$, h est *uniformément continue*, donc il existe $\beta > 0$ tel que $|X_n(\omega)-X(\omega)| \leq \beta$ implique $|h\circ X_n(\omega)-h\circ X(\omega)| \leq \alpha$ **si** $X_n(\omega)$ **et** $X(\omega)$ **sont dans** $[-N,N]$. Donc, si $E_n = \{\omega : |X_n(\omega)-X(\omega)| \leq \beta\}$ et $E_{n,N} = \{\omega : |X_n(\omega)-X(\omega)| \leq \beta \text{ et } |X_n(\omega)| \leq N \text{ et } |X(\omega)| \leq N\}$, $p(E_{n,N}) \leq p(|h\circ X_n - h\circ X| \leq \alpha)$.

Or E_n est inclus dans la réunion (non disjointe) des trois événements :

$$\begin{aligned} E_{n,N} &= \{\omega : |X_n(\omega)-X(\omega)| \leq \beta \text{ et } |X_n(\omega)| \leq N \text{ et } |X(\omega)| \leq N\} \\ &\{\omega : |X_n(\omega)-X(\omega)| \leq \beta \text{ et } |X_n(\omega)| > N\} \\ &\{\omega : |X_n(\omega)-X(\omega)| \leq \beta \text{ et } |X(\omega)| > N\}. \end{aligned}$$

Pour $n \geq n_0$, il vient

$$p(E_n) \leq p(E_{n,N})+p(|X| > N)+p(|X_n| > N) \leq p(E_{n,N})+p(|X| > N)+p(|X| > N+\alpha)+\varepsilon \leq p(E_{n,N})+2\varepsilon .$$

A partir d'un rang n_1, $p(E_n) \geq 1-\varepsilon$ et donc $p(E_{n,N}) \geq 1-2\varepsilon$. Comme $p(E_{n,N}) \leq p(|h\circ X_n - h\circ X| \leq \alpha)$, $p(|h\circ X_n - h\circ X| \leq \alpha) \geq 1-3\varepsilon$ à partir du rang $\max(n_0,n_1)$, ce qui prouve que $p(|h\circ X_n - h\circ X| \leq \alpha)$ tend vers 1 quand *n* tend vers l'infini.

PARTIE B

Si la fonction h n'est pas continue sur $\mathbb{R}$, même si elle est étagée, il se peut que $h\circ X_n$ ne tende pas pour p vers $h\circ X$: si $h = 1_{]0,+\infty[}$ et si X_n est la variable aléatoire constante égale à 1/n, X_n tend en probabilité vers $X=0$ et cependant $h(X_n)$ ne tend pas vers $h(0)$. Pour établir que $h\circ X_n \xrightarrow{p} h\circ X$ pour toute h borélienne de $\mathbb{R}$ dans $\mathbb{R}$ nous ferons l'hypothèse suivante sur la suite (X_n) :

les X_n ont la même fonction de répartition F, $p(X_n \leq t) = F(t)$ pour tout n.

1) La fonction de répartition de X est F

On prouve séparément que, pour tout réel t, $F(t) \leq p(X \leq t)$, puis $p(X \leq t) \leq F(t)$.

Soit t un réel fixé et u un réel > 0.

Les inégalités $X_n \leq t$ et $|X_n - X| \leq u$ induisent $X \leq t+u$, d'où $p(X_n \leq t \text{ et } |X_n - X| \leq u) \leq p(X \leq t+u)$

On utilise $p(A\cap B) \geq p(A)+p(B)-1$, d'où $F(t)+p(|X_n - X| \leq u)-1 \leq p(X \leq t+u)$.

En faisant tendre n vers $+\infty$, il vient $F(t) \leq p(X \leq t+u)$; avec $u \to 0^+$, on aboutit à $F(t) \leq p(X \leq t)$ parce qu'une fonction de répartition est continue à droite.

Pour obtenir $p(X \leq t) \leq F(t)$, on recommence en permutant les rôles de X et X_n.

2) Les X_n et X ont même loi de probabilité (ou même distribution)

Il suffit d'établir que deux variables aléatoires X et Y ayant la même fonction de répartition ont la même loi de probabilité : $F_X = F_Y$ implique $p_X = p_Y$. Par hypothèse, les probabilités p_X et p_Y, définies sur la tribu des boréliens de $\mathbb{R}$, coïncident sur tous les intervalles fermés $]-\infty,t]$ et donc sur les intervalles ouverts $]t,+\infty[$. Comme $]-\infty,t[$ est la réunion dénombrable et croissante des $]-\infty,t-1/n]$, p_X et p_Y sont égales sur tous les intervalles $]-\infty,t[$, et donc sur les $[t,+\infty[$. L'intervalle $]a,b[$ étant la différence entre $]-\infty,a[$ et $]-\infty,b[$, $p_X(]a,b[) = p_Y(]a,b[)$: p_X et p_Y coïncident sur tous les intervalles ouverts de $\mathbb{R}$, donc sur tous les ouverts car tout ouvert de $\mathbb{R}$ est une réunion dénombrable disjointe d'intervalles ouverts. On en conclut (http://books.google.fr/books?id=kq69nXaoV-AC&printsec=frontcover#v=onepage&q&f=false, paragraphe **8.6**, page 132) que $p_X = p_Y$ sur tous les boréliens.

On désignera par p_X la distribution commune aux X_n et à X.

3) $h \circ X_n \xrightarrow{p} h \circ X$ si h est étagée sur les boréliens de $\mathbb{R}$

Soit $\lambda_1,\lambda_2,...,\lambda_k$ les valeurs *distinctes* prises par h : si B_i est le borélien $h^{-1}(\lambda_i)$, $\mathbb{R} = \sum_1^k B_i$. On remarque que $|h \circ X_n - h \circ X| \leq \alpha$ équivaut à $h \circ X_n - h \circ X = 0$ dès que $\alpha < \min\{|\lambda_i - \lambda_j| : i \neq j\}$, ce que l'on supposera.

Il en résulte que $|h \circ X_n - h \circ X| \leq \alpha$ est l'événement $\bigcup_{i=1}^k \left(X_n^{-1}(B_i) \cap X^{-1}(B_i)\right)$, réunion d'événements disjoints, donc $p(h \circ X_n - h \circ X = 0) = \sum_{i=1}^k p\left(X_n^{-1}(B_i) \cap X^{-1}(B_i)\right)$. Comme $\Omega = \sum_{i=1}^k X_n^{-1}(B_i) = \sum_{i=1}^k X^{-1}(B_i)$,

$\Omega = \sum_i \sum_j \left(X_n^{-1}(B_i) \cap X^{-1}(B_j)\right) = \sum_i \left(X_n^{-1}(B_i) \cap X^{-1}(B_i)\right) + \sum_{i \neq j} \left(X_n^{-1}(B_i) \cap X^{-1}(B_j)\right)$, reste donc à prouver que la somme finie $\sum_{i \neq j} p\left(X_n^{-1}(B_i) \cap X^{-1}(B_j)\right)$ a pour limite 0 quand n tend vers l'infini, ce qui équivaut à

$$\lim_n p\left(X_n^{-1}(B_i) \cap X^{-1}(B_j)\right) = 0 \text{ pour chaque i différent de j.}$$

Raisonnons sur $i = 1$ et $j = 2$. Par hypothèse, $p\left(X_n^{-1}(B_1)\right) = p_X(B_1)$ est indépendant de n : il existe donc un compact K_1 contenu dans B_1 tel que $p\left(X_n^{-1}(B_1)\right) \leq p\left(X_n^{-1}(K_1)\right) + \varepsilon$ *pour tout n* et K_2 un compact contenu K_2 dans B_2 tel que $p\left(X_n^{-1}(B_2)\right) \leq p\left(X_n^{-1}(K_2)\right) + \varepsilon$ pour tout n. Alors,

$$p\left(X_n^{-1}(B_1) \cap X^{-1}(B_2)\right) \leq p\left(X_n^{-1}(K_1) \cap X^{-1}(K_2)\right) + 3\varepsilon$$

car si $K \subset A$ et $L \subset B$, on peut poser $A = K + X$ et $B = L + Y$ et alors $A \cap B - K \cap L = K \cap Y + X \cap L + X \cap Y$ d'où $p(A \cap B - K \cap L) \leq p(Y) + p(X) + p(X \cap Y)$.

Comme K_1 et K_2 sont disjoints, leur distance $d(K_1,K_2)$ est strictement positive, d'où

$$\lim_n p\left(X_n^{-1}(K_1) \cap X^{-1}(K_2)\right) = 0.$$

En définitive, $p\left(X_n^{-1}(B_1) \cap X^{-1}(B_2)\right) \leq 4\varepsilon$ à partir d'un certain rang, CQFD.

4) Si, pour tout m, $h_m \circ X_n \xrightarrow{p} h_m \circ X$ quand n tend vers l'infini, et si $h_m \to h$ uniformément sur $\mathbb{R}$,

alors $h \circ X_n \xrightarrow{p} h \circ X$.

Par hypothèse, il existe un entier m tel que, *pour tout* n *et tout* ω,

$$|h_m \circ X_n(\omega) - h \circ X_n(\omega)| \leq \alpha/3 \text{ et } |h_m \circ X(\omega) - h \circ X(\omega)| \leq \alpha/3.$$

Comme

$|h \circ X_n(\omega) - h \circ X(\omega)| \leq |h \circ X_n(\omega) - h_m \circ X_n(\omega)| + |h_m \circ X_n(\omega) - h_m \circ X(\omega)| + |h_m \circ X(\omega) - h \circ X(\omega)|$,

pour cet entier m, $|h \circ X_n(\omega) - h \circ X(\omega)| \leq 2\alpha/3 + |h_m \circ X_n(\omega) - h_m \circ X(\omega)|$.

On en déduit $\{\omega : |h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha/3\} \subset \{\omega : |h \circ X_n(\omega) - h \circ X(\omega)| \leq \alpha\}$.

Avec ce m fixé, $\lim_n p(\{\omega : |h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha/3\}) = 1$ car $h_m \circ X_n \xrightarrow{p} h_m \circ X$, et donc

$$\lim_n p(\{\omega : |h \circ X_n(\omega) - h \circ X(\omega)| \leq \alpha/3\}) = 1.$$

5) $h \circ X_n \to h \circ X$ **pour la probabilité** p **dès que** h **est borélienne**

a) *La fonction borélienne* h *est positive et majorée*

On sait alors que h est limite uniforme sur Ω d'une suite de fonctions étagées h_m, à savoir $h_m = \sum_{k=1}^{m} \frac{(k-1)M}{m}.1_{A_k}$,

où M est un majorant *strict* de f sur Ω et A_k est l'événement $\{\omega \in \Omega : (k-1)/n \leq h(\omega) < kM/n\}$.

Sur Ω, $0 \leq f(\omega) - h_n(\omega) \leq M/n$: $h \circ X_n \xrightarrow{p} h \circ X$ en vertu de **3)** et **4)** .

b) *La fonction borélienne* h *est seulement positive*

Posons $h_m = \min(m,h)$: h_m est borélienne positive et bornée, aussi, pour tout m, $h_m \circ X_n \xrightarrow{p} h_m \circ X$ quand n tend vers l'infini. Les trois conditions $|h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha$, $h \circ X_n(\omega) \leq m$, $h \circ X(\omega) \leq m$ induisent $|h \circ X_n(\omega) - h \circ X(\omega)| \leq \alpha$, d'où

$$\begin{aligned} p\left(|h \circ X_n(\omega) - h \circ X(\omega)| \leq \alpha\right) &\geq p\left(|h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha \text{ et } h \circ X_n \leq m \text{ et } h \circ X \leq m\right) \\ &\geq p\left(|h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha\right) + p\left(h \circ X_n \leq m\right) + p\left(h \circ X \leq m\right) - 2. \end{aligned}$$

Il existe un entier m pour lequel $p\left(h \circ X_n \leq m\right) + p\left(h \circ X \leq m\right) \geq 2 - \varepsilon$: en effet,

$p\left(h \circ X \leq m\right) = p\left(X^{-1} \circ h^{-1}(]-\infty, m])\right) = p_X(h \leq m)$, de même $p\left(h \circ X_n \leq m\right) = p_X(h \leq m)$ et

$\lim_{m \to +\infty} p_X(h \leq m) = 1$.

Une fois m fixé, il existe un entier n_0 à partir duquel $p\left(|h_m \circ X_n(\omega) - h_m \circ X(\omega)| \leq \alpha\right) \geq 1 - \varepsilon$,

d'où $p\left(|h \circ X_n(\omega) - h \circ X(\omega)| \leq \alpha\right) \geq 1 - 2\varepsilon$, ce qui termine la démonstration.

c) *La fonction borélienne* h *est de signe quelconque*

On utilise la décomposition $f = f^+ - f^-$, où $f^+ = \frac{f + |f|}{2}$ et $f^- = \frac{|f| - f}{2}$ sont deux fonctions boréliennes positives : on conclut avec le **b)** et la structure d'espace vectoriel de $\mathcal{H}$.

En résumé,

Soit (X_n) une suite de variables aléatoires qui converge vers X pour une probabilité p :

a) $h \circ X_n \xrightarrow{p} h \circ X$ pour toute fonction **continue** h de $\mathbb{R}$ dans $\mathbb{R}$.

b) si les X_n sont identiquement distribuées, $h \circ X_n \xrightarrow{p} h \circ X$ pour toute fonction **borélienne** h de $\mathbb{R}$ dans $\mathbb{R}$.

FIN*