Leçon 18

Convergence en probabilité
de suites de variables aléatoires

La convergence en probabilité (ou convergence en mesure) est une forme affaiblie à la fois de la convergence presque sûre et de la convergence dans les espaces L^p, qui est bien adaptée aux inégalités classiques comme celles de Markov et Tchebychev.

Dans cette leçon, X_n, $n \in \mathbb{N}$, et X sont des variables aléatoires sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ à valeurs dans $\mathbb{R}$, éventuellement $\mathbb{R}^d$ (munis des tribus boréliennes).

1 Convergence en probabilité

Définition 1 (Convergence en probabilité)**.** *Une suite X_n, $n \in \mathbb{N}$, de variables aléatoires réelles converge en probabilité vers une variable aléatoire X si pour tout $\varepsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}\big(|X_n - X| \geq \varepsilon\big) \, = \, 0.$$

Si $|\cdot|$ désigne le module sur $\mathbb{C}$, ou la norme euclidienne sur $\mathbb{R}^d$ (ou tout autre norme), la définition est donc identique pour des vecteurs aléatoires. En considérant la distance $d(X_n, X)$ la définition s'applique de la même façon à des variables aléatoires à valeurs dans un espace métrique (E, d).

L'inégalité de Markov montre que pour tout $p > 0$, et tout $\varepsilon > 0$,

$$\mathbb{P}\big(|X_n - X| \geq \varepsilon\big) \, \leq \, \frac{1}{\varepsilon^p} \, \mathbb{E}\big(|X_n - X|^p\big).$$

Il est clair ainsi que si la suite X_n, $n \in \mathbb{N}$, converge vers X dans L^p, alors elle converge en probabilité. L'implication ayant lieu pour tout $p > 0$, cette observation conduit parfois à qualifier la convergence en probabilité comme la convergence dans L^0. Mais la convergence en probabilité est strictement plus faible que la convergence dans L^p, $p > 0$, comme le montre l'exemple de la suite de variables aléatoires X_n, $n \in \mathbb{N}$, de lois de Bernoulli $\mathbb{P}(X_n = n^{\frac{1}{p}}) = \frac{1}{n}$,

$\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$, pour laquelle $\mathbb{E}(|X_n|^p) = 1$ pour tout $n \in \mathbb{N}$, mais convergeant toutefois en probabilité vers la variable aléatoire $X = 0$ car $\mathbb{P}(|X_n| \geq \varepsilon) = \frac{1}{n}$ pour tout $n \geq \varepsilon^p > 0$.

La convergence en probabilité est donc une conséquence de la convergence dans L^p (pour un $p > 0$). Elle est aussi plus faible que la convergence presque sûre puisque, comme étudié dans la Leçon 17, X_n, $n \in \mathbb{N}$, converge presque sûrement vers X équivaut à dire que pour tout $\varepsilon > 0$,

$$\lim_{m \to \infty} \mathbb{P}\Big(\sup_{n \geq m} |X_n - X| \geq \varepsilon \Big) = 0$$

(et $\{|X_m - X| \geq \varepsilon\} \subset \{\sup_{n \geq m} |X_n - X| \geq \varepsilon\}$). Il sera mis en évidence dans le Paragraphe 3 que l'implication est stricte.

Les remarques élémentaires suivantes sont souvent utiles.

D'abord, dans beaucoup d'exemples d'illustration, il sera question de suites X_n, $n \in \mathbb{N}$, convergeant en probabilité vers 0. Il s'agit donc tout simplement de la convergence vers la variable aléatoire $X = 0$ (ou seulement presque sûrement égale à 0), exprimée par $\lim_{n \to \infty} \mathbb{P}(|X_n| \geq \varepsilon) = 0$ pour tout $\varepsilon > 0$.

Par décroissance, il suffit de vérifier la propriété de la définition pour tout $0 < \varepsilon \leq \varepsilon_0$ pour un $\varepsilon_0 > 0$, par exemple $\varepsilon_0 = 1$. Il est équivalent de considérer $\mathbb{P}(|X_n - X| \geq \varepsilon)$ ou $\mathbb{P}(|X_n - X| > \varepsilon)$ (puisque $\varepsilon > 0$ est arbitraire), et il suffit de démontrer que

$$\limsup_{n \to \infty} \mathbb{P}\big(|X_n - X| \geq \varepsilon \big) \ = \ 0$$

(puisque les probabilités sont positives).

D'après la définition de la limite d'une suite de nombres réels, dire que X_n converge en probabilité vers X signifie que, pour chaque $\varepsilon > 0$, pour tout $\eta > 0$, il existe n_0 $(= n_0(\varepsilon, \eta))$ tel que si $n \geq n_0$,

$$\mathbb{P}\big(|X_n - X| \geq \varepsilon \big) \ \leq \ \eta.$$

Il peut être commode pour des illustrations d'observer qu'il est possible de choisir $\eta = \varepsilon$ dans cette description. Autrement dit, X_n converge en probabilité vers X si pour tout $\varepsilon > 0$, il existe n_0 $(= n_0(\varepsilon))$ tel que pour tout $n \geq n_0$,

$$\mathbb{P}\big(|X_n - X| \geq \varepsilon\big) \;\leq\; \varepsilon.$$

En effet, si $\eta > 0$ est un autre paramètre positif, soit $\eta \geq \varepsilon$ ce qui fournit la première formulation, soit $\eta \leq \varepsilon$, et alors

$$\mathbb{P}\big(|X_n - X| \geq \varepsilon\big) \;\leq\; \mathbb{P}\big(|X_n - X| \geq \eta\big) \;\leq\; \eta$$

pour tout $n \geq n_0(\eta)$. D'où l'affirmation pour $n \geq \max(n_0(\varepsilon), n_0(\eta))$.

2 Distance en probabilité

La convergence en probabilité peut être décrite à travers une distance sur l'espace des variables aléatoires.

Pour X et Y deux variables aléatoires réelles sur $(\Omega, \mathcal{A}, \mathbb{P})$, poser

$$D(X, Y) \;=\; \mathbb{E}\big(\min\big(|X - Y|, 1\big)\big).$$

Cette expression a une similitude avec la distance induite par la norme L^1 : le minimum avec la constante 1 assure l'intégrabilité (toute autre valeur strictement positive ferait l'affaire). À l'image de la définition de convergence en probabilité, $D(X, Y)$ se définit de la même manière pour des vecteurs aléatoires (voire des variables aléatoires à valeurs dans un espace métrique). D'autres choix sont possibles pour une définition de distance en probabilité à travers une distance bornée sur $\mathbb{R}$ équivalente à la distance usuelle, comme par exemple $\mathbb{E}\big(\frac{|X-Y|}{1+|X-Y|}\big)$.

La fonctionnelle D vérifie l'inégalité triangulaire. En effet, si a, b, c sont des réels,

$$\min\big(|a - b|, 1\big) \;\leq\; \min\big(|a - c|, 1\big) + \min\big(|c - b|, 1\big).$$

Il n'y a rien à démontrer si $|a-c| \geq 1$ ou $|c-b| \geq 1$ (car alors le membre de droite est toujours plus grand que 1). Dans le cas contraire, l'inégalité triangulaire habituelle exprime que

$$\min\big(|a-b|, 1\big) \;\leq\; |a-b| \;\leq\; |a-c| + |c-b|$$

ce qui fournit le résultat. Cette inégalité appliquée à $a = X(\omega)$, $b = Y(\omega)$, $c = Z(\omega)$ pour tout $\omega \in \Omega$, où X, Y, Z sont trois variables aléatoires réelles, puis intégrée par rapport à $\mathbb{P}$, établit l'inégalité triangulaire annoncée

$$D(X,Y) \;\leq\; D(X,Z) + D(Z,Y).$$

Par ailleurs, $D(X,Y) \geq 0$, $D(X,Y) = D(Y,X)$ et $D(X,X) = 0$. Enfin, si $D(X,Y) = 0$, alors $\min(|X-Y|, 1) = 0$ presque sûrement, soit $X = Y$ presque sûrement.

Ainsi D définit une distance sur l'espace des variables aléatoires quotienté par la relation d'équivalence $X = Y$ presque sûrement (comme dans la définition des espaces L^p). Cette distance décrit la convergence en probabilité.

Proposition 2. *Une suite X_n, $n \in \mathbb{N}$, converge en probabilité vers X si et seulement si*

$$\lim_{n\to\infty} D(X_n, X) \;=\; 0.$$

Démonstration. Une implication reproduit la comparaison avec la convergence dans les espaces L^p. En effet, si $0 < \varepsilon \leq 1$, pour tout $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\big(|X_n - X| \geq \varepsilon\big) \;&=\; \mathbb{P}\Big(\min\big(|X_n - X|, 1\big) \geq \varepsilon\Big) \\
&\leq\; \frac{1}{\varepsilon}\,\mathbb{E}\Big(\min\big(|X_n - X|, 1\big)\Big)
\end{aligned}$$

d'après l'inégalité de Markov. Ainsi, si $\lim_{n\to\infty} D(X_n, X) = 0$, alors X_n converge en probabilité vers X.

La réciproque est plus intéressante. Soit $\varepsilon > 0$; en décomposant l'espérance par la relation de Chasles, pour tout $n \in \mathbb{N}$,

$$\begin{aligned} \mathbb{E}\big(\min(|X_n - X|, 1)\big) \; &= \; \mathbb{E}\big(\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}} \min(|X_n - X|, 1)\big) \\ &\quad + \mathbb{E}\big(\mathbb{1}_{\{|X_n - X| < \varepsilon\}} \min(|X_n - X|, 1)\big). \end{aligned}$$

La première espérance peut être majorée par $\mathbb{P}(|X_n - X| \geq \varepsilon)$ (en utilisant que $\min(|X_n - X|, 1) \leq 1$), la seconde par ε (en utilisant que $\mathbb{P}(|X_n - X| < \varepsilon) \leq 1$). Ainsi

$$\mathbb{E}\big(\min(|X_n - X|, 1)\big) \; \leq \; \mathbb{P}\big(|X_n - X| \geq \varepsilon\big) + \varepsilon.$$

Sous l'hypothèse de convergence en probabilité de X_n vers X, pour tout $\eta > 0$, il existe n_0 ($= n_0(\varepsilon, \eta)$) tel que si $n \geq n_0$, $\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \eta$. Ainsi, si $n \geq n_0$,

$$\mathbb{E}\big(\min\big(|X_n - X|, 1\big)\big) \; \leq \; \eta + \varepsilon.$$

Comme $\varepsilon > 0$ et $\eta > 0$ sont arbitraires (choisir $\eta = \varepsilon$ éventuellement), la conclusion s'ensuit. La proposition est donc établie. □

3 Comparaison avec la convergence presque sûre

Comme il en a déjà été fait état, si une suite X_n, $n \in \mathbb{N}$, (de variables aléatoires réelles) converge presque sûrement vers (une variable aléatoire réelle) X, c'est-à-dire

$$\lim_{m \to \infty} \mathbb{P}\Big(\sup_{n \geq m} |X_n - X| \geq \varepsilon\Big) \; = \; 0$$

pour tout $\varepsilon > 0$, elle converge également en probabilité puisque

$$\mathbb{P}\big(|X_m - X| \geq \varepsilon\big) \; \leq \; \mathbb{P}\Big(\sup_{n \geq m} |X_n - X| \geq \varepsilon\Big).$$

Bien entendu la réciproque n'est pas vraie en général. En fait, la différence entre les deux modes de convergence est similaire à la différence entre la convergence d'une série et la convergence vers 0 de son terme général. L'exemple suivant, déjà étudié dans la Leçon 17, le démontre clairement. Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires de loi de Bernoulli sur $\{0,1\}$ de paramètres de succès respectifs p_n, $n \in \mathbb{N}$, $p_n \in]0,1[$. Autrement dit $\mathbb{P}(X_n = 1) = p_n$, $\mathbb{P}(X_n = 0) = 1 - p_n$, $n \in \mathbb{N}$. En supposant que $p_n \to 0$, il est naturel d'envisager la convergence de la suite X_n, $n \in \mathbb{N}$, vers la variable aléatoire $X = 0$ (de loi de Dirac en 0). Pour tout $0 < \varepsilon \leq 1$,

$$\mathbb{P}\big(|X_n| \geq \varepsilon\big) \,=\, \mathbb{P}(X_n = 1) \,=\, p_n$$

de sorte que X_n converge en probabilité vers la variable aléatoire $X = 0$ sous l'hypothèse $\lim_{n\to\infty} p_n = 0$. Maintenant, si les variables X_n, $n \in \mathbb{N}$, sont indépendantes, le critère basé sur le lemme de Borel-Cantelli examiné dans la Leçon 17 indique que la suite X_n, $n \in \mathbb{N}$, converge presque sûrement vers 0 si et seulement si, pour tout $0 < \varepsilon \leq 1$

$$\sum_{n\in\mathbb{N}} \mathbb{P}\big(|X_n| \geq \varepsilon\big) \,=\, \sum_{n\in\mathbb{N}} p_n \,<\, \infty.$$

Il est donc facile de discriminer sur cet exemple la convergence en probabilité et la convergence presque sûre.

Alors que les convergences en probabilité et dans L^p sont métriques, il peut être observé que la convergence presque sûre ne peut pas être décrite par une distance (sinon l'Exercice 4 montrerait qu'elle est équivalente à la convergence en probabilité).

4 Bilan de convergences

La comparaison entre les trois modes de convergence, convergence presque sûre, convergence dans L^p $(p > 0)$ et convergence en probabilité, pour une suite X_n, $n \in \mathbb{N}$, de variables aléatoires (réelles) vers une variable aléatoire X, s'établit comme suit, à l'aide des implications et des exemples divulgués dans les leçons et paragraphes précédents :

Convergence presque sûre et convergence dans L^p ne sont pas comparables. Le théorème de convergence dominée fournit toutefois une condition pour qu'une suite convergeant presque sûrement converge aussi dans L^p.

La convergence presque sûre entraîne la convergence en probabilité (la réciproque est fausse en général).

La convergence dans L^p ($p > 0$) entraîne la convergence en probabilité (la réciproque est fausse en général).

5 Complément : Intégrabilité uniforme

Le théorème suivant est une version simplifiée d'un énoncé relatif à la notion d'*intégrabilité uniforme*. Il fournit une forme améliorée du théorème de convergence dominée en remplaçant l'hypothèse de domination par une condition plus souple (accessoirement aussi en remplaçant la convergence presque sûre par la convergence en probabilité).

Théorème 3. *Sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires réelles convergeant en probabilité vers une variable aléatoire réelle X ; soit également $p > 0$; si, pour un $r > p$,*

$$\sup_{n\in\mathbb{N}} \mathbb{E}\big(|X_n|^r\big) \,<\, \infty,$$

alors X_n, $n \in \mathbb{N}$, converge vers X dans L^p, autrement dit

$$\lim_{n\to\infty} \mathbb{E}\big(|X_n - X|^p\big) \,=\, 0.$$

Démonstration. La première étape consiste à s'assurer que $\mathbb{E}(|X|^r) < \infty$. C'est une conséquence du lemme de Fatou (Lemme 2, Leçon 3) ; en effet, d'après l'Exercice 4, il existe une sous-suite d'entiers n_k, $k \in \mathbb{N}$, le long de laquelle X_{n_k} converge vers X presque sûrement, et donc aussi, par continuité, $|X_{n_k}|^r \to |X|^r$ presque sûrement. Mais alors, par le lemme de Fatou donc,

$$\mathbb{E}\big(|X|^r\big) \,=\, \mathbb{E}\big(\liminf_{k\to\infty} |X_{n_k}|^r\big) \,\leq\, \liminf_{k\to\infty} \mathbb{E}\big(|X_{n_k}|^r\big)$$

et le membre de droite est fini par hypothèse. Grâce à cette affirmation et l'inégalité triangulaire, il découle de l'hypothèse qu'il existe une constante $C > 0$ telle que

$$\sup_{n\in\mathbb{N}} \mathbb{E}\big(|X_n - X|^r\big) \,\leq\, C.$$

Soit alors $\varepsilon > 0$ fixé ; par la relation de Chasles, pour tout entier n,

$$\begin{aligned}
\mathbb{E}\big(|X_n - X|^p\big) \;&=\; \mathbb{E}\big(\mathbb{1}_{\{|X_n - X| \leq \varepsilon\}}\,|X_n - X|^p\big) + \mathbb{E}\big(\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}}\,|X_n - X|^p\big) \\
&\leq\; \varepsilon^p + \mathbb{E}\big(\mathbb{1}_{\{|X_n - X| \geq \varepsilon\}}\,|X_n - X|^p\big) \\
&\leq\; \varepsilon^p + \big[\mathbb{P}\big(|X_n - X| \geq \varepsilon\big)\big]^{1-\frac{p}{r}}\,\big[\mathbb{E}\big(|X_n - X|^r\big)\big]^{\frac{p}{r}}
\end{aligned}$$

où la dernière inégalité résulte d'une application de l'inégalité de Hölder (au couple d'exposants conjugués $(\frac{r}{p}, \frac{r}{r-p})$). Avec la première étape, pour tout $n \in \mathbb{N}$,

$$\mathbb{E}\big(|X_n - X|^p\big) \;\leq\; \varepsilon^p + C^{\frac{p}{r}}\big[\mathbb{P}\big(|X_n - X| \geq \varepsilon\big)\big]^{1-\frac{p}{r}}.$$

Puisque la suite X_n, $n \in \mathbb{N}$, converge en probabilité vers X, il existe $n_0 = n_0(\varepsilon)$ tel que si $n \geq n_0$, alors $\mathbb{P}(|X_n - X| \geq \varepsilon) \leq \varepsilon$. Ainsi, si $n \geq n_0$,

$$\mathbb{E}\big(|X_n - X|^p\big) \;\leq\; \varepsilon^p + C^{\frac{p}{r}}\varepsilon^{1-\frac{p}{r}}.$$

Si $\eta > 0$, $\varepsilon > 0$ peut être choisi tel que $\varepsilon^p + C^{\frac{p}{r}}\varepsilon^{1-\frac{p}{r}} \leq \eta$, de sorte que $\mathbb{E}(|X_n - X|^p) \leq \eta$ pour tout $n \geq n_0$, ce qui assure la convergence dans L^p annoncée. La démonstration est terminée. $\square$

L'exemple suivant est instructif. Soient X_n, $n \in \mathbb{N}$, des variables aléatoires indépendantes telles que $\mathbb{P}(X_n = n^{\frac{3}{2}}) = \frac{1}{n^2}$, $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n^2}$, $n \in \mathbb{N}$. D'après le critère de Borel-Cantelli, la suite X_n, $n \in \mathbb{N}$ converge presque sûrement vers 0. Comme $\mathbb{E}(X_n) = \frac{1}{\sqrt{n}}$, $n \in \mathbb{N}$, il n'est pas nécessaire d'invoquer un théorème pour assurer que $\mathbb{E}(X_n) \to 0$. Il peut être toutefois observer que $\sup_{n\in\mathbb{N}} \mathbb{E}(|X_n|^{\frac{4}{3}}) = 1 < \infty$, alors que la suite X_n, $n \in \mathbb{N}$, ne peut être dominée par une variable intégrable. En effet, pour tout $m \in \mathbb{N}$,

$$\mathbb{E}\big(\sup_{n\in\mathbb{N}} |X_n|\big) \;\geq\; \mathbb{E}\big(\sup_{n\geq m} X_n\big) \;\geq\; m^{\frac{3}{2}}\,\mathbb{P}\big(\sup_{n\geq m} X_n \neq 0\big)$$

(car si $\sup_{n\geq m} X_n$ n'est pas nul, il est plus grand que $m^{\frac{3}{2}}$). Mais, par indépendance des variables X_n, $n \in \mathbb{N}$,

$$\begin{aligned}
\mathbb{P}\big(\sup_{n \geq m} X_n \neq 0\big) \;&=\; 1 - \mathbb{P}\big(\sup_{n \geq m} X_n = 0\big) \\
&=\; 1 - \prod_{n \geq m} \Big(1 - \frac{1}{n^2}\Big) \\
&\geq\; 1 - e^{-\frac{1}{m-1}}
\end{aligned}$$

$(m \geq 2)$ par prise du logarithme, de sorte que

$$\lim_{m \to \infty} m^{\frac{3}{2}}\, \mathbb{P}(\sup_{n \geq m} X_n \neq 0) \;=\; \infty.$$

Donc $\mathbb{E}(\sup_{n \in \mathbb{N}} |X_n|) = \infty$ et le théorème de convergence dominée ne s'applique pas, au contraire du Théorème 3.

Comme mentionné en introduction, l'énoncé précédent est une forme affaiblie d'un théorème d'intégrabilité uniforme. Pour $p = 1$, la condition

$$\sup_{n \in \mathbb{N}} \mathbb{E}\big(|X_n|^r\big) \;<\; \infty,$$

pour un $r > 1$ peut en effet être remplacée par l'existence d'une fonction (borélienne) $G : \mathbb{R}^+ \to \mathbb{R}^+$ vérifiant $\lim_{t \to \infty} \frac{G(t)}{t} = \infty$ telle que

$$\sup_{n \in \mathbb{N}} \mathbb{E}\big(G\big(|X_n|\big)\big) \;<\; \infty.$$

La démonstration est en fait à peine modifiée. Mais le point important est que cette dernière condition devient en fait aussi nécessaire (à la convergence de X_n, $n \in \mathbb{N}$, dans L^1). Il s'agit du *Théorème de de La Vallée-Poussin*[1]

1. Charles-Jean de La Vallée Poussin, mathématicien belge (1866–1962).

Exercices

(*Une étoile * désignera une question de difficulté supérieure.*)

Exercice 1. Reprendre l'Exercice 1, Leçon 17, en incluant la convergence en probabilité dans la discussion.

Exercice 2. Soit X_n, $n \in \mathbb{N}$, une suite de variables aléatoires de même loi uniforme $\mathcal{U}(0,1)$ sur $[0,1]$; démontrer que

a) $\frac{1}{nX_n} \to 0$ en probabilité ;

b) $\frac{X_n}{n} \to 0$ presque sûrement ;

c) $\frac{1}{n^2 X_n} \to 0$ presque sûrement.

On suppose en outre les variables X_n, $n \in \mathbb{N}$, mutuellement indépendantes. Établir que

d) $\frac{1}{nX_n}$ ne converge pas presque sûrement.

Exercice 3. Soient X_n, Y_n, $n \in \mathbb{N}$, des suites de variables aléatoires réelles, ainsi que des variables aléatoires réelles X et Y, définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

a) Si les suites X_n, $n \in \mathbb{N}$, et Y_n, $n \in \mathbb{N}$, convergent dans L^p, $p \geq 1$, vers des variables aléatoires X et Y respectivement, démontrer que la suite $X_n + Y_n$, $n \in \mathbb{N}$, converge dans L^p vers $X + Y$.

La suite de l'exercice a pour but de démontrer que la propriété de stabilité précédente est encore vérifiée pour la convergence en probabilité.

b) Si U et V sont des variables aléatoires et $t \in \mathbb{R}$, démontrer que

$$\mathbb{P}(U + V \geq 2t) \;\leq\; \mathbb{P}(U \geq t) + \mathbb{P}(V \geq t).$$

c) Si les suites X_n, $n \in \mathbb{N}$, et Y_n, $n \in \mathbb{N}$, convergent en probabilité vers X et Y

respectivement, démontrer que la suite X_n+Y_n, $n \in \mathbb{N}$, converge en probabilité vers $X+Y$.

d*) Dans le cadre de la question précédente, démontrer que la suite X_nY_n, $n \in \mathbb{N}$, converge en probabilité vers XY.

Exercice 4. Soit une suite X_n, $n \in \mathbb{N}$, de variables aléatoires convergeant en probabilité vers une variable aléatoire X.

a) Construire par récurrence une suite strictement croissante d'entiers n_k, $k \in \mathbb{N}$, telle que, pour tout $k \geq 1$,

$$\mathbb{P}\big(|X_{n_k} - X| \geq \tfrac{1}{k}\big) \;\leq\; \tfrac{1}{2^k}\,.$$

b) Démontrer que pour tout $\varepsilon > 0$, $\sum_{k\geq 1} \mathbb{P}(|X_{n_k} - X| \geq \varepsilon) < \infty$. En conclure que la suite X_{n_k}, $k \in \mathbb{N}$, converge vers X presque sûrement.

Exercice 5. Pour toute variable aléatoire $Z \geq 0$, démontrer que pour tout $t > 0$,

$$\mathbb{P}(Z \geq t) \;\leq\; \frac{1}{t}\,\mathbb{E}\big(Z\,\mathbb{1}_{\{Z\geq t\}}\big).$$

Soit à présent X une variable aléatoire réelle de carré intégrable.

a) Démontrer que

$$\lim_{t\to\infty} \mathbb{E}\big(X^2\mathbb{1}_{\{|X|\geq t\}}\big) \;=\; 0.$$

Déduire de la question préliminaire que pour tout $\varepsilon > 0$,

$$\lim_{n\to\infty} n\,\mathbb{P}\big(|X| \geq \varepsilon\sqrt{n}\,\big) \;=\; 0.$$

b) Soient X_n, $n \geq 1$, de même loi que X. En utilisant le question précédente, que peut-on dire de la convergence de la suite de variables aléatoires $Y_n = \frac{1}{\sqrt{n}}\max(X_1,\ldots,X_n)$, $n \geq 1$, lorsque $n \to \infty$?

Exercice 6 (*Polynômes de Bernstein.*) Soit une $f : [0,1] \to \mathbb{R}$ une fonction continue et soit X_n une variable aléatoire sur $(\Omega, \mathcal{A}, \mathbb{P})$ suivant une loi binomiale $\mathcal{B}(n,x)$ de taille $n \geq 1$ et de paramètre $x \in [0,1]$. On note $T_n = \frac{X_n}{n}$.

a) Montrer que $Q_n(x) = \mathbb{E}(f(T_n))$ est un polynôme (en x), appelé polynôme de Bernstein (de degré n).

b) Calculer $\mathbb{E}(T_n)$ et $\mathrm{Var}(T_n)$.

c) Établir que $|f(x) - Q_n(x)| \leq \mathbb{E}(|f(T_n) - f(x)|)$.

d) Pour $\eta > 0$, poser $A = \{|T_n - x| \leq \eta\}$ et montrer que $\mathbb{P}(A^c) \leq \frac{1}{\eta^2 n}$.

e) Invoquer un théorème pour affirmer que pour tout $\varepsilon > 0$, il existe $\eta = \eta(\varepsilon) > 0$ tel que si $u, v \in [0,1]$ et $|u - v| \leq \eta$, alors $|f(u) - f(v)| \leq \varepsilon$.

f) Pour $\varepsilon > 0$ et $\eta = \eta(\varepsilon) > 0$ comme dans la question précédente, décomposer, pour tout $x \in [0,1]$, $\mathbb{E}(|f(T_n) - f(x)|)$ suivant A et A^c pour obtenir que

$$\big|f(x) - Q_n(x)\big| \leq \varepsilon + \frac{2M(f)}{\eta^2 n}$$

où $M(f) = \sup_{t \in [0,1]} |f(t)|$.

g) Déduire de la question précédente que pour tout $\varepsilon > 0$, il existe un entier $n = n(\varepsilon)$ tel que

$$\sup_{x \in [0,1]} \big|f(x) - Q_n(x)\big| \leq 2\varepsilon.$$

h) Quel théorème vient-il d'être démontré ?