Leçon 264 · Va discrètes

I) Espaces probabilisés discrets

1) Définition

Déf 1 On appelle espace probabilisé discret un espace probabilisable $(\Omega, \mathcal{F})$ où Ω est dénombrable et $\mathcal{F} = \mathcal{P}(\Omega)$

Lemme 2: Ω ensemble dénombrable.

(i) g une application de Ω dans $\mathbb{R}_+$ telle que $\sum_{\omega \in \Omega} g(\omega) = 1$. Pour $A \in \mathcal{P}(\Omega)$, on pose

$$\mathbb{P}(A) = \sum_{\omega \in A} g(\omega)$$

(ii) Tte probabilité sur $(\Omega, \mathcal{P}(\Omega))$ peut être obtenue de cette façon, à partir de

$$g : \Omega \to [0,1], \quad \omega \mapsto \mathbb{P}(\{\omega\}).$$

2) Exemples de lois de probas discrètes

a) Loi uniforme

Déf 3: Supposons $X(\Omega) = [\![1, n]\!]$ pour n dans $\mathbb{N}^*$. X est de loi uniforme si $\mathbb{P}(X = k) = \frac{1}{n}$ pour tout k dans $[\![1, n]\!]$

Exemple 4: Soit un jet d'un dé à six faces, X la variable aléatoire qui renvoie le résultat du jet. $\mathbb{P}(X = k) = \frac{1}{6}$ pour tout k dans $[\![1, 6]\!]$

b) Modèles géométriques

Déf 5: X suit une loi géométrique de paramètre $p \in]0,1[$ si $\forall k \in \mathbb{N}^*$ $\mathbb{P}(X = k) = pq^{k-1}$ où $q = 1 - p$

Exemple 6: On lance une pièce qui tombe sur face avec probabilité p. Si X est le nombre de lancers jusqu'au premier "face", on a $\mathbb{P}(X = k) = pq^{k-1}$

c) Modèle de Bernoulli et loi binomiale

Déf 7 Une variable de Bernoulli de paramètre p est une variable aléatoire X telle que $\mathbb{P}(X = 0) = 1 - p$, $\mathbb{P}(X = 1) = p$.

Exemple 8: On lance une pièce équilibrée, alors si $X = \begin{cases} 0 & \text{si la pièce tombe sur face} \\ 1 & \text{si la pièce tombe sur pile} \end{cases}$ suit une loi de Bernoulli

Déf 9: Une variable aléatoire X suit une loi binomiale de paramètre (n, p) notée $\mathcal{B}(n, p)$ où $n \in \mathbb{N}^*$, $p \in]0,1[$ si $\forall k \in [\![0, n]\!]$, $\mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}$ où $q = 1 - p$

Ex 10 On lance n pièces équilibrées, si Y représente le nombre de faces obtenues après les n lancers, alors $Y \sim \mathcal{B}(n, \frac{1}{2})$

d) Loi de Poisson

Déf 11: Soit $\lambda > 0$. Une variable aléatoire X suit une loi de Poisson de paramètre λ, notée $\mathcal{P}(\lambda)$, si $\forall k \in \mathbb{N}$, $\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

II) Moments d'une variable aléatoire discrète (v.a.d)

1) Espérance

Def 12. X v.a.d. Soit la famille de réels $(x\,\mathbb{P}(X=x))_{x\in\mathbb{R}}$. Si $\sum_{x\in\mathbb{R}} |x|\,\mathbb{P}(X=x)$ est finie, on dit que X possède une moyenne (ou espérance), notée $\mathbb{E}[X] = \sum_{x\in\mathbb{R}} x\,\mathbb{P}(X=x)$

Exemple 13. • X suit la loi uniforme sur $\{x_1,\dots,x_n\}$ où $x_i \in \mathbb{R}$ pour tout $i\in[\![1,n]\!]$

$$\mathbb{E}[X] = \sum_{i=1}^{n} x_i \frac{1}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

• X suit une loi de Bernoulli de paramètre p

$\mathbb{E}[X] = p$

Théorème 14 : (Thm de transfert)

Soit X v.a.d, $X(\omega)\in E$ et $f : E \to \mathbb{R}$ mesurable

$Y = f\circ X$ est une v.a.d.

Pour que Y admette une moyenne il faut et il suffit que

$\mathbb{E}[|f(X)|] = \sum_{x\in E} |f(x)|\,\mathbb{P}(X=x)$ soit finie

et $\mathbb{E}[f(X)] = \sum_{x\in E} f(x)\,\mathbb{P}(X=x)$

Application 15 : (Théorème de Weierstraß) $\boxed{\text{DVLPT 1}}$

Soient $f : [0,1] \to \mathbb{R}$ fonction continue sur $[0,1]$, $(B_n)_{n\in\mathbb{N}}$ la suite de fonctions définies par

$$B_n : [0,1] \to \mathbb{R}, \quad x \mapsto \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}$$

alors $(B_n)_{n\in\mathbb{N}}$ converge uniformément vers f

Application 16 : Soit X une v.a.d d'espérance finie

(i) $\mathbb{E}[aX+b] = a\,\mathbb{E}[X] + b$

(ii) $(\mathbb{P}(X=b)=1) \Rightarrow \mathbb{E}[X] = b$

(iii) $(\mathbb{P}(a<X<b)=1) \Rightarrow a < \mathbb{E}[X] < b$

(iv) si $g(X)$ et $h(X)$ d'espérance finie, alors

$$\mathbb{E}[g(X)+h(X)] = \mathbb{E}[g(X)] + \mathbb{E}[h(X)]$$

2) Moments d'ordres supérieurs

Def 17 : p réel ≥ 1, si $\sum_{x\in E} |x|^p\,\mathbb{P}(X=x)$ est finie, alors X admet un moment d'ordre p noté

$$\mathbb{E}[X^p] = \sum_{x\in E} x^p\,\mathbb{P}(X=x)$$

Def 18. $\mathbb{E}[(X-\mathbb{E}[X])^2]$ est appelée la variance de X, notée $Var(X)$

L'écart type est $\sqrt{Var(X)}$

III) Couples et sommes de v.a.d.

1) Indépendance

Def 19 : deux v.a.d X et Y sont dites indépendantes si pour tout x dans $X(\Omega)$ et y dans $Y(\Omega)$

$$\mathbb{P}(X=x, Y=y) = \mathbb{P}(X=x)\,\mathbb{P}(Y=y)$$

Prop 20 : Si $X : \Omega \to E_1$ et $Y : \Omega \to E_2$ sont deux v.a.d indépendantes et si $f : E_1 \to \mathbb{R}$ et $g : E_2 \to \mathbb{R}$ sont deux fonctions réelles, alors $f(X)$ et $g(X)$ sont indépendantes.

2) Sommes de v.a.d indépendantes

Prop 21 : Soient X et Y deux v.a.d indépendantes, alors $X+Y$ est une v.a.d et sa loi est donnée par

$\forall x \in E \quad \mathbb{P}(X+Y=x) = \sum_{x_1 \in E} \mathbb{P}(X=x_1)\,\mathbb{P}(Y=x-x_1)$

Exemple 22 : • Si X_1 suit une loi de Poisson de paramètre λ_1 et X_2 suit une loi de Poisson de paramètre λ_2, alors X_1+X_2 suit une loi de Poisson de paramètre $\lambda_1+\lambda_2$

3) Espérance d'une somme

Théorème 23. Soit g une fonction réelle et X et Y deux v.a.d, si le membre de droite de l'équation ci-dessous converge, alors

$$\mathbb{E}[g(X,Y)] = \sum_{\substack{x\in X(\Omega)\\ y \in Y(\Omega)}} g(x,y)\,\mathbb{P}(X=x, Y=y)$$

Corollaire 24 : • $\forall a,b$ constantes, $\mathbb{E}[aX+bY] = a\mathbb{E}[X]+b\mathbb{E}[Y]$ lorsque ces termes existent

• $Var(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$

Exemples 25 : • $X \sim \mathcal{B}(n,p)$, $\mathbb{E}[X] = np$ donc $Var(X) = np(1-p)$

• si $X \sim \mathcal{P}(\lambda)$, $Var(X) = \lambda$

III) Théorèmes limites

1) Approximation de poisson

$\boxed{\text{DVLPT 2}}$

Théorème 26 : Soit, pour tout $n \in \mathbb{N}^*$, une famille finie ([illegible])$_{1\le j\le m_n}$ de v.a.d de loi de Bernoulli de paramètre $p_{n,j}$

$\mathbb{P}(X_{n,j}=1) = p_{n,j}$, $\mathbb{P}(X_{n,j}=0) = 1-p_{n,j}$

On suppose que :

(i) m_n tend en croissant vers $+\infty$

(ii) $\sum_{j=1}^{m_n} p_{n,j} \to \lambda > 0$ quand n tend vers $+\infty$

(iii) $\max_{1\le j\le m_n} p_{n,j}$ tend vers 0

Alors, si $S_n = X_{n,1} + \dots + X_{n,m_n}$, alors $S_n \xrightarrow[n\to+\infty]{\mathcal{L}} Z$ où Z suit une loi de poisson de paramètre λ.

2) Loi faible et forte des grands nombres

Def 27 : Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de v.a.d. On dit que $(X_n)_{n\in\mathbb{N}^*}$ converge en probabilité vers X si $\forall \varepsilon > 0$

$\mathbb{P}(|X_n - X| > \varepsilon) \xrightarrow[n\to+\infty]{} 0$.

On note $X_n \xrightarrow{\mathbb{P}} X$

• On dit que $(X_n)_{n\in\mathbb{N}^*}$ converge presque sûrement vers X noté $X_n \xrightarrow[n\to+\infty]{p.s} X$ si $\mathbb{P}(X_n \xrightarrow[n\to+\infty]{} X) = 1$

Théorème 28. Soit (X_n) une suite de v.a.d i.i.d, de variance et d'espérance finie, $\mu = \mathbb{E}[X_1]$, $\sigma_2 = Var(X_1)$

$\frac{1}{n}(X_1 + \dots + X_n) \xrightarrow{\mathbb{P}} \mu$

Ex 29 : Si on lance une pièce n fois, plus n est grand plus on a de chance d'avoir obtenu autant de pile que de face.

Théorème 30 : Soit (X_n) suite de v.a.d i.i.d avec [illegible] $< +\infty$, alors

$\frac{1}{n}(X_1 + \dots + X_n) \xrightarrow{p.s} \mathbb{E}[X_1]$

Ex 31 : Soit $\varepsilon > 0$, la moyenne du nombre de face obtenues sur n lancers de pièces sera à distance $> \varepsilon$ de $1/2$ seulement un nombre fini de fois, avec probabilité 1.

<u>Références</u> :

Ouvrard, Probabilités 1
Ouvrard, Probabilités 2
Barbe - Ledoux, Probabilité

Développements de la leçon 264

Théorème de Bernstein

[Référence : Zuily-Queffélec, Analyse pour l'agrégation, 4ème édition, p518]

Théorème. *Soit $f : [0,1] \to \mathbb{C}$ une fonction continue, ω son module de continuité ($\omega(h) = \sup\{|f(u) - f(v)|; |u - v| \leq h\}$). Pour $n \geq 1$, on considère le polynôme $B_n(f,x) = B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right)$, le n-ième polynôme de Bernstein de f. Alors :*

1. *B_n converge vers f uniformément sur $[0,1]$*
2. *Plus précisément on a $\|f - B_n\|_\infty \leq C\omega\left(\frac{1}{\sqrt{n}}\right)$, où C est une constante numérique.*
3. *L'estimation de (2) est optimale : il existe une fonction lipschitzienne f pour laquelle $\|f - B_n\|_\infty \geq \delta\omega\left(\frac{1}{\sqrt{n}}\right)$, où δ est une constante numérique.*

Démonstration. (1) Soit $x \in [0,1]$, $X_1, \dots, X_n$ n variables aléatoires i.i.d de Bernoulli de paramètre x et $S_n = \sum_{i=1}^{n} X_i$.

S_n suit une loi binomiale de paramètres n et x donc :
$E[f(\frac{S_n}{n})] = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right) = B_n(x)$.
Soit $\delta \in]0,1[$, on a :

$$f(x) - B_n(x) = E[f(x) - f(\tfrac{S_n}{n})]$$

d'où :

$$|f(x) - B_n(x)| \leq E[|f(x) - f(\tfrac{S_n}{n})|]$$

Si $|x - \frac{S_n}{n}| \leq \delta$, alors $|f(x) - f(\frac{S_n}{n})| \leq \omega(\delta)$ car d'après le théorème de Heine f est uniformément continue sur $[0,1]$.
Si $|x - \frac{S_n}{n}| > \delta$, alors $|f(x) - f(\frac{S_n}{n})| \leq 2\|f\|_\infty$.
D'où :

$$\begin{aligned} E[|f(x) - f(\tfrac{S_n}{n})|] &\leq E[|f(x) - f(\tfrac{S_n}{n})| \mathbb{1}_{|x - \frac{S_n}{n}| \leq \delta}] + E[|f(x) - f(\tfrac{S_n}{n})| \mathbb{1}_{|x - \frac{S_n}{n}| > \delta}] \\ &\leq \omega(\delta) + 2\|f\|_\infty P(\|\tfrac{S_n}{n} - x| > \delta) \\ &\leq \omega(\delta) + \tfrac{\|f\|_\infty}{2n\delta^2} \end{aligned}$$

d'après l'inégalité de Tchebychev appliquée à $\frac{S_n}{n}$.
Donc :

$$|f(x) - B_n(x)| \leq \omega(\delta) + \tfrac{\|f\|_\infty}{2n\delta^2}$$

puis

$$\|f - B_n\|_\infty \leq \omega(\delta) + \tfrac{\|f\|_\infty}{2n\delta^2}$$

D'où

$$\limsup_{n\to\infty} \|f - B_n\|_\infty \leq \omega(\delta)$$

Or $\omega(\delta) \to 0$ quand $\delta \to 0^+$ donc la convergence uniforme sur $[0,1]$ est établie.

(2) Montrons qu'il existe C une constante telle que :

$$\|f - B_n\|_\infty \leq C\omega\left(\tfrac{1}{\sqrt{n}}\right)$$

Montrons d'abord que pour tout h, λ tels que $h, \lambda h \in [0,1]$, $\omega(\lambda h) \leq (\lambda+1)\omega(h)$.
ω est croissante et $\omega(h+k) \leq \omega(h) + \omega(k)$ donc par récurrence sur $n \in \mathbb{N}$, $\omega(nh) \leq n\omega(h)$.
On a alors, pour $\lambda \in \mathbb{R}, \omega(\lambda h) \leq \omega(\lceil \lambda \rceil h) \leq \lceil \lambda \rceil \omega(h) \leq (\lambda+1)\omega(h)$.

De plus on a :

$$|f(x) - B_n(x)| \leq E[|f(x) - f(\tfrac{S_n}{n})|] \leq E[\omega(x - \tfrac{S_n}{n})]$$

On en déduit

$$\begin{aligned}
|f(x) - B_n(x)| &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\mathbb{E}\left(\sqrt{n}\left|x - \frac{S_n}{n}\right| + 1\right) \\
&= \omega\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\left\|x - \frac{S_n}{n}\right\|_1\right) \\
&\leq \omega\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\left\|x - \frac{S_n}{n}\right\|_2\right)
\end{aligned}$$

par l'inégalité de Hölder. Or

$$\begin{aligned}
\left\|x - \frac{S_n}{n}\right\|_2^2 &= \mathbb{E}\left(\left|x - \frac{S_n}{n}\right|^2\right) \\
&= \mathrm{Var}\left(x - \frac{S_n}{n}\right) + \left(\mathbb{E}\left(x - \frac{S_n}{n}\right)\right)^2 \\
&= \frac{1}{n^2}nx(1-x) + \left(x - \frac{1}{n}nx\right)^2 \\
&= \frac{x(1-x)}{n}.
\end{aligned}$$

D'où

$$\begin{aligned}
|f(x) - B_n(x)| &\leq \omega\left(\frac{1}{\sqrt{n}}\right)\left(1 + \sqrt{n}\sqrt{\frac{x(1-x)}{n}}\right) \\
&\leq \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right).
\end{aligned}$$

D'où $\|f - B_n\|_\infty \leq \frac{3}{2}\omega\left(\frac{1}{\sqrt{n}}\right)$.

(3) On pose $f : x \mapsto \left|x - \frac{1}{2}\right|$, on a alors $\omega(h) \leq h$. Par ailleurs,

$$\begin{aligned}
\|f - B_n\|_\infty &\geq \left|f\left(\frac{1}{2}\right) - B_n\left(\frac{1}{2}\right)\right| \\
&= \left|B_n\left(\frac{1}{2}\right)\right| \\
&= \mathbb{E}\left|\frac{S_n}{n} - \frac{1}{2}\right| \\
&= \frac{1}{2n}\mathbb{E}|2S_n - n|.
\end{aligned}$$

D'où $\|f - B_n\|_\infty \geq \frac{1}{2n}\mathbb{E}|\varepsilon_1 + \cdots + \varepsilon_n|$ avec $\varepsilon_j := 2X_j - 1$ n variables de Rademacher i.i.d.

On va maintenant utiliser l'inégalité de Khintchine que l'on admettra.

Inégalité de Khintchine. *Soit $n \in \mathbb{N}$, $\epsilon_1, \ldots, \epsilon_n$ des variables aléatoires de Rademacher indépendantes, $a_1, \ldots, a_n \in \mathbb{R}$ et $X = \sum_{i=1}^n a_i \epsilon_i$. Alors on a :*

$$\|X\|_1 \geq \frac{1}{\sqrt{e}}\|X\|_2 \tag{1}$$

D'où

$$\begin{aligned}
\|f - B_n\|_\infty &\geq \frac{1}{2n}\|\varepsilon_1 + \cdots + \varepsilon_n\|_1 \\
&\geq \frac{1}{2n\sqrt{e}}\|\varepsilon_1 + \cdots + \varepsilon_n\|_2
\end{aligned}$$

d'après l'inégalité de Khintchine.

Or $\|\varepsilon_1 + \cdots + \varepsilon_n\|_2^2 = \mathrm{Var}(\varepsilon_1 + \cdots + \varepsilon_n) + (\mathbb{E}(\varepsilon_1 + \cdots + \varepsilon_n))^2 = n$. D'où

$$\|f - B_n\|_\infty \geq \frac{1}{2\sqrt{ne}} \geq \frac{1}{2\sqrt{e}}\omega\left(\frac{1}{\sqrt{n}}\right).$$

L'estimation obtenue en (2) est donc bien optimale.

□

Théorème des événements rares de Poisson

[Référence : Ouvrard, Probabilités 2, Théorème 14.20]

Théorème. *Soit, pour tout $n \in \mathbb{N}^*$, une famille finie $\{A_{n,j} \mid 1 \leq j \leq M_n\}$ d'événements indépendants définis sur un espace de probabilité $(\Omega, \mathcal{A}, P)$. On pose $P(A_{n,j}) = p_{n,j}$ et on note :*

$$S_n = \sum_{j=1}^{M_n} \mathbb{1}_{A_{n,j}}$$

On suppose que la suite de terme général M_n tend en croissant vers $+\infty$, que

$$\max_{1 \leq j \leq M_n} p_{n,j} \underset{n \to +\infty}{\longrightarrow} 0 \text{ et que } \sum_{j=1}^{M_n} p_{n,j} \underset{n \to +\infty}{\longrightarrow} \lambda$$

où $\lambda > 0$. Alors la suite $(S_n)_{n \in \mathbb{N}^*}$ converge en loi vers la loi de Poisson $\mathcal{P}(\lambda)$

Démonstration. Afin de démontrer la convergence en loi on va utiliser le théorème de Lévy. Par indépendance des $A_{n,j}, 1 \leq j \leq M_n$, on a pour tout $t \in \mathbb{R}$:

$$\begin{aligned}
\phi_{S_n}(t) &= \prod_{j=1}^{M_n} \phi_{X_{n,j}}(t) \\
&= \prod_{j=1}^{M_n} [p_{n,j} \exp(it) + (1 - p_{n,j})] \\
&= \prod_{j=1}^{M_n} [1 + p_{n,j}(\exp(it) - 1)]
\end{aligned}$$

Si log est la détermination principale du logarithme complexe, il résulte de la formule de Taylor avec reste intégral que, pour tout z tel que $|z| < 1$, on a :

$$\log(1+z) = z - z^2 \int_0^1 (1-u) \frac{1}{(1+uz)^2} du$$

Notons $z = \exp(it) - 1$. Puisque $\max\limits_{1 \leq j \leq M_n} p_{n,j} \underset{n \to +\infty}{\longrightarrow} 0$, il existe N tel que, pour tout $n \geq N$, on ait $\max\limits_{1 \leq j \leq M_n} p_{n,j} < 1/2$. Pour tout $n \geq N$, on a alors :

$$\log \phi_{S_n}(t) = z \sum_{j=1}^{M_n} p_{n,j} - z^2 \sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{(1+up_{n,j}z)^2} du$$

D'après l'inégalité triangulaire, on a, pour tout $n \geq N$, et pour tout $u \in [0,1]$,

$$|1 + up_{n,j}z| \geq 1 - p_{n,j}|z| \geq \frac{1}{2}$$

On a donc pour tout $n \geq N$:

$$| \sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{(1+up_{n,j}z)^2} du | \leq 2 \left[\max_{1 \leq j \leq M_n} p_{n,j} \right] \left[\sum_{j=1}^{M_n} p_{n,j} \right]$$

D'après les hypothèses $\max\limits_{1 \leq j \leq M_n} p_{n,j} \underset{n \to +\infty}{\longrightarrow} 0$ et $\sum_{j=1}^{M_n} p_{n,j} \underset{n \to +\infty}{\longrightarrow} \lambda$ donc :

$$\sum_{j=1}^{M_n} p_{n,j}^2 \int_0^1 \frac{1-u}{(1+up_{n,j}z)^2} du \underset{n \to +\infty}{\longrightarrow} 0$$

D'où :

$$\log \phi_{S_n}(t) \underset{n \to +\infty}{\longrightarrow} \lambda z$$

C'est-à-dire :

$$\forall t \in \mathbb{R}, \lim_{n \to \infty} \phi_{S_n}(t) = \exp[\lambda(\exp(it) - 1)]$$

On achève la démonstration en appliquant le théorème de Lévy.

□

Théorème de Lévy. *Soit pour tout $n \in \mathbb{N}$, une variable aléatoire X_n, définie sur un espace probabilisé (Ω_n, A_n, P_n), à valeurs dans $\mathbb{R}^d$, de fonction caractéristique ϕ_{X_n}.*

1. *Si la suite de variables aléatoires (X_n) converge en loi vers X, où X est une variable aléatoire définie sur un espace probabilisé (Ω, A, P), à valeurs dans $\mathbb{R}^d$, alors la suite (ϕ_{X_n}) des fonctions caractéristiques converge simplement vers la fonction caractéristique ϕ_X de X.*

2. *Inversement, si la suite ϕ_{X_n} des fonctions caractéristiques converge simplement vers une fonction ϕ continue en 0, alors ϕ est la transformée de Fourier d'une probabilité μ sur $\mathbb{R}^d$, et la suite des variables aléatoires (X_n) converge en loi vers μ. De plus, il existe une variable aléatoire (non unique) X définie sur un espace probabilisé (Ω, A, P), à valeurs dans $\mathbb{R}^d$, telle que la suite de variables aléatoires (X_n) converge en loi vers X.*