261

Loi d'une variable aléatoire. Caractérisation, exemples, applications.

Cadre : On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

I. Loi d'une variable aléatoire. Définitions, premiers exemples.

A) Loi, fonction de répartition.

Def 1. On appelle variable aléatoire toute application mesurable définie sur un espace probabilisé.

Def 2. Soit $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (E, \mathcal{B})$ une v.a. On appelle loi de X sous la probabilité $\mathbb{P}$ la mesure de probabilité image $\mathbb{P}_X$ sur $(E, \mathcal{B})$.
$\forall B \in \mathcal{B}$, $\mathbb{P}_X(B) = \mathbb{P}(\{\omega \mid X(\omega) \in B\}) = \mathbb{P}(X \in B)$.
Pour la suite, on prendra $(E, \mathcal{B}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. On aura des v.a. réelles.

Def 3. On appelle fonction de répartition de X, ou de la loi $\mathbb{P}_X$, et on note F^X, la fonction sur $\mathbb{R}$ définie par : $F^X(t) = \mathbb{P}_X(]-\infty, t]) = \mathbb{P}(\{X \leq t\})$, $t \in \mathbb{R}$.

Prop 1. Si F est une fonction de répartition, alors :
* $0 \leq F \leq 1$
* F est croissante, continue à droite, avec une limite à gauche en tout point.
* $\lim_{t \to -\infty} F(t) = 0$ et $\lim_{t \to +\infty} F(t) = 1$.

Réciproquement, une fonction vérifiant ces 3 propriétés est une fonction de répartition d'une v.a. réelle.

Prop 2. La fonction de répartition caractérise la loi : $F^X = F^Y \iff \mathbb{P}_X = \mathbb{P}_Y$.

Prop 3. Une fonction de répartition admet au plus un nombre dénombrable de points de discontinuité.

Prop 4. $\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x) = 0 \iff F^X$ est continue en x.

Ex 1. $F = \mathbb{1}_{[a, +\infty[}$ est la fonction de répartition de la masse de Dirac δ_a.

Def 4. Soit X une v.a. réelle définie sur $(\Omega, \mathcal{A}, \mathbb{P})$. Si X est intégrable, on appelle espérance de X (sous la probabilité $\mathbb{P}$) le réel $\int_\Omega X \, d\mathbb{P} = \mathbb{E}(X)$.

Def 5. X est dite centrée si elle est intégrable et $\mathbb{E}(X) = 0$.

Def 6. Si $X \in L^p$, $p \geq 1$, on définit le moment d'ordre p : $\mathbb{E}(X^p)$.

Def 7. Si $X \in L^2$, sa variance est définie par $Var(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$.
$\sigma_X = \sqrt{Var(X)}$ s'appelle l'écart-type de X.

Rmq 1. σ_X mesure l'écart des valeurs de X à la moyenne.

B) Lois discrètes : exemples et utilisations.

Def 8. Une loi est dite discrète si c'est une combinaison linéaire finie ou dénombrable de masses de Dirac. Une v.a. de loi discrète $\sum_{i \in I} \alpha_i \delta_{x_i}$ ne prend presque-sûrement qu'un nombre fini ou dénombrable de valeurs.

Prop 5. La loi d'une variable aléatoire X à valeurs dans un espace dénombrable F est caractérisée par $\{(x_i, p_i^X), x_i \in F$, avec $p_i^X = \mathbb{P}(X = x_i)\}$.

Ex 2. Lois discrètes usuelles.

Nom de la loi de X	Espace dans lequel X prend ses valeurs	Loi	Espérance ($X \in L^1$)	Variance ($X \in L^2$)	Utilisation en modélisation
Loi discrète	F dénombrable	$(x_i, p_i^X)_{i \in F}$	$\sum_{i \in F} x_i p_i^X$	$\sum_{i \in F} (x_i - \mathbb{E}(X))^2 p_i^X$	/
Loi de Bernoulli $\mathcal{B}(1, p)$, $p \in [0,1]$	$\{0, 1\}$	$\mathbb{P}(X=0) = 1-p$, $\mathbb{P}(X=1) = p$	p	$p(1-p)$	"Pile ou face"
Loi binomiale $\mathcal{B}(n, p)$, $n \geq 1$	$[\![0, n]\!]$	$\mathbb{P}(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$	Répétition de n "pile ou face"
Loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$	$\mathbb{N}$	$\mathbb{P}(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$	λ	λ	Nombre d'apparitions d'un événement rare dans une suite infinie d'événements
Loi géométrique $\mathcal{G}(p)$, $p \in]0,1[$	$\mathbb{N}^*$	$\mathbb{P}(X=k) = (1-p)^{k-1} p$	$1/p$	$\frac{1-p}{p^2}$	Obtention du 1er succès de la répétition d'épreuves ind. de Bernoulli
Loi uniforme	Ensemble fini F	$(x_i, p_i^X)_{i \in F}$, avec $p_i^X = \frac{1}{\mathrm{Card}\, F}$	$\frac{\sum_{i \in F} x_i}{\mathrm{Card}\, F}$	$\frac{1}{\mathrm{Card}\, F} \left(\sum_{i \in F} (x_i - \mathbb{E}(X))^2 \right)$	Toutes les valeurs ont même proba

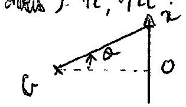
C) Lois à densité sur $\mathbb{R}$.

Def 9. Si une loi est absolument continue par rapport à la mesure de Lebesgue λ sur $\mathbb{R}$, et si X est de loi $\mathbb{P}_X$, on dira par abus de langage que X admet la densité f par rapport à λ.
Cela signifie que $\forall x \in \mathbb{R}$, $\mathbb{P}_X(]-\infty, x]) = \int_{-\infty}^x f(t)\,dt$.

Prop 6. f définie sur $\mathbb{R}$ à valeurs réelles est une densité de probabilité si elle est positive, intégrable et vérifie $\int_{\mathbb{R}} f(t)\,dt = 1$.

Ex 3. Lois à densité usuelles.

Nom de la loi de X	Densité de la loi de X	Espérance ($X \in L^1$)	Variance ($X \in L^2$)	Utilisation en modélisation
Loi uniforme sur $[a, b]$, $\mathcal{U}_{[a,b]}$	$x \mapsto \frac{1}{b-a} \mathbb{1}_{[a,b]}(x)$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	Point au hasard sur un segment, simulation d'autres lois
Loi exponentielle $\mathcal{E}(\lambda)$, $\lambda > 0$	$x \mapsto \lambda e^{-\lambda x} \mathbb{1}_{[0, +\infty[}(x)$	$1/\lambda$	$1/\lambda^2$	Durée de vie ou temps d'attente (loi "sans mémoire")
Loi normale $\mathcal{N}(m, \sigma^2)$	$x \mapsto \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2	Très courante, voir suite

Loi de Cauchy	$x \mapsto \frac{1}{\pi(1+x^2)}$	Non intégrable	Distribution de l'abscisse du point d'impact x du rayon avec θ uniforme dans $]-\pi/2, \pi/2[$ 

Prop 7: Soit X de densité f.

(i) F^X est continue, de sorte que $\forall x \in \mathbb{R}$, $\mathbb{P}(X=x)=0$.

(ii) F est dérivable en tout point x où f est continue, et $F'(x)=f(x)$

(iii) Si la fonction de répartition F^X d'une v.a.r X est dérivable, ou seulement continue partout et dérivable par morceaux, alors X admet la densité $F'(x)=f(x)$.

Ex. 4: Il existe des cas "mixtes" de loi n'admettant pas de densité et n'étant pas non plus discrètes. Si $f \geq 0$, intégrable tq $\int_{\mathbb{R}} f > 0$ et I fini ou dénombrable, avec des $q_i > 0$ tels que : $\int_{\mathbb{R}} f + \sum_{i \in I} q_i = 1$, alors $F : x \mapsto \int_{-\infty}^{x} f(t)dt + \sum_{i \in I : i \leq x} q_i$ est une fonction de répartition

II Caractérisation et utilisations de lois :

A) Méthodes usuelles de caractérisation :

Théorème 1 (de transfert): Soit X une v.a. réelle et ϕ borélienne de $\mathbb{R}$ dans $\mathbb{R}$.

Si $\phi \geq 0$, $\mathbb{E}[\phi(X)] = \int_{\Omega} \phi \circ X(\omega)\, d\mathbb{P}(\omega) = \int_{\mathbb{R}} \phi(x)\, d\mathbb{P}_X(x)$

Si ϕ à valeurs quelconques, $\phi(X) \in L^1(\Omega, \mathcal{A}, \mathbb{P}) \iff \phi \in L^1(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$.

Dans ce cas, l'égalité précédente a lieu.

Ex 5. Si X de loi $\frac{1}{2}\delta_0 + \frac{1}{2}\delta_2$, $\mathbb{E}(X) = 1$ et $\mathbb{E}(X^2) = 2$

Prop 8. La donnée de $\mathbb{E}(\phi(X))$ pour toute fonction borélienne bornée caractérise la loi de X. On peut se restreindre aux fonctions continues bornées, voire aux fonctions $\mathcal{C}^\infty$ bornées.

Def 10: Soit X une v.a. à valeurs discrètes (dans $\mathbb{N}$). On définit la fonction génératrice G_X de X par $G_X : [0,1] \longrightarrow \mathbb{R}$, $s \longmapsto \mathbb{E}[s^X] = \sum_{n=0}^{+\infty} p_n s^n$

Prop 9. La fonction génératrice est continue sur $[0,1]$ et infiniment dérivable sur $[0,1[$. Elle caractérise la loi de X.

Prop 10. Soit X une v.a à valeurs dans $\mathbb{N}$, de fct génératrice G_X.

X intégrable $\iff$ G_X dérivable à gauche en $s=1$, et $\mathbb{E}(X) = G_X'(1)$ dans ce cas.

Ex 6: • $X \sim \mathcal{B}(n,p)$: $G_X(s) = (1-p+ps)^n$.

• $X \sim \mathcal{P}(\theta)$: $G_X(s) = e^{\theta(s-1)}$.

Def 11: Si X une v.a. réelle, on définit sa fonction caractéristique ϕ_X par :

$\phi_X : \mathbb{R} \longrightarrow \mathbb{C}$, $t \longmapsto \mathbb{E}(e^{itX})$. Si X à valeurs dans $\mathbb{R}^n$, $\phi_X(t) = \mathbb{E}(e^{i\langle t, X\rangle})$.

Rmq 2: ϕ_X est la "transformée de Fourier" de la loi $\mathbb{P}_X$ (avec une convention différente de celle de l'analyse).

Prop 11: ϕ_X est continue, de module ≤ 1, et $\phi_X(0) = 1$, $\phi_X(-u) = \overline{\phi_X(u)}$ $\forall u \in \mathbb{R}$.

$\forall (a,b) \in \mathbb{R}^2$, $\phi_{a+bX}(t) = e^{iat}\phi_X(bt)$.

Ex 7: • Si $X \sim \mathcal{N}(0,1)$, $\phi_X(t) = e^{-t^2/2}$.

• Si $X \sim \mathcal{U}([a,b])$, $\phi_X(t) = \frac{e^{itb} - e^{ita}}{it(b-a)}$

Théorème 2: ϕ_X caractérise la loi de la variable aléatoire X.

Si $\phi_X = \phi_Y$, alors $X \sim Y$.

B) Calculs et simulation de lois :

Prop 12 (calculs de moments). Soit X une v.a.r, de fonction caractéristique ϕ_X et de loi $\mathbb{P}_X$.

(i) Si $\mathbb{E}(|X|^n) < +\infty$, ϕ est n fois dérivable, de dérivée k-ième ($k \leq n$) $\phi^{(k)}(t) = i^k \mathbb{E}(X^k e^{itX})$

En particulier, $\phi^{(k)}(0) = i^k \mathbb{E}(X^k)$.

(ii) Réciproquement, si n est pair et si ϕ est n-fois dérivable en 0, alors X admet tout moment d'ordre $\leq n$.

Ex 8. Si $X \sim \mathcal{N}(0,1)$, elle admet des moments à tout ordre, ~~et [illegible]~~ $\mathbb{E}(X^2) = 1$.

Prop 13. Si pour tout h continue bornée, avec X v.a.r, $\mathbb{E}(h(X)) = \int_{\mathbb{R}} h(x)\mu(dx)$, alors la loi de X est égale à μ.

Appli 1 (recherche de densité). Si g fct mesurable tq $Y = g(X)$, avec X v.a.r de densité f_X, on va essayer de mettre, pour h fct bornée, $\mathbb{E}[h(Y)] = \mathbb{E}[h \circ g(X)]$ sous la forme $\int h(y) f_Y(y) dy$. f_Y sera alors la densité recherchée.

Ex 9. Si $Y = aX + b$, $(a,b) \in \mathbb{R}^2$: • Si $a=0$, $Y \sim \delta_b$ (sans densité)

• Si $a \neq 0$, $\mathbb{E}(h(Y)) = \int_{\mathbb{R}} h(y) f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}dy$

Donc $f_Y(y) = f_X\left(\frac{y-b}{a}\right)\frac{1}{|a|}$.

$\Rightarrow$ Si $X \sim \mathcal{N}(m, \sigma^2)$, alors $\frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$

Def 12: Une famille $(X_i)_{i \in I} : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est mutuellement indépendante si $\forall J \subset I$ fini, $\forall B_j \in \mathcal{B}(\mathbb{R})$, $j \in J$, $\mathbb{P}\left(\bigcap_{j \in J}\{X_j \in B_j\}\right) = \prod_{j \in J}\mathbb{P}(X_j \in B_j)$.

Prop 14: Soit $X = (X_1, \dots, X_n)$ à valeurs dans $\mathbb{R}^n$.

Les X_i sont indépendantes $\iff$ $\forall (u_1, \dots, u_n) \in \mathbb{R}^n$, $\phi_X(u_1, \dots, u_n) = \prod_{j=1}^{n}\phi_{X_j}(u_j)$

• la loi $\mathbb{P}_{(X_1, \dots, X_n)}$ est égale à $\mathbb{P}_{X_1} \otimes \dots \otimes \mathbb{P}_{X_n}$.

Prop 15. • Si X et Y 2 v.a.r indépendantes, alors $\phi_{X+Y} = \phi_X \phi_Y$.

• La loi de $X+Y$ est donnée par le produit de convolution $\mathbb{P}_X * \mathbb{P}_Y$ défini par, $\forall \phi$ borélienne bornée, $\int_{\mathbb{R}} \phi\, d(\mathbb{P}_X * \mathbb{P}_Y) = \int_{\mathbb{R}}\int_{\mathbb{R}} \phi(x+y)\, d\mathbb{P}_X(x)\, d\mathbb{P}_Y(y)$

Développement 1

Ex 10: • Si $X \sim \mathcal{N}(m, \sigma^2)$ et $Y \sim \mathcal{N}(m', (\sigma')^2)$, $X+Y \sim \mathcal{N}(m+m, \sigma^2+\sigma'^2)$.

• Si $X \sim \mathcal{P}(\theta)$, $Y \sim \mathcal{P}(\theta')$, $X+Y \sim \mathcal{P}(\theta+\theta')$.

• Si $X \sim \mathcal{B}(n,p)$, $Y \sim \mathcal{B}(m,p)$, alors $X+Y \sim \mathcal{B}(n+m, p)$.

Prop 16: La loi de $\max(X_1, \dots, X_n)$, pour $X_1, \dots, X_n$ n var indépendantes i.i.d, est donnée par :

$\forall t \in \mathbb{R}$, $F^{\max(X_1,\dots,X_n)}(t) = \prod_{i=1}^{n} F^{X_i}(t) = \prod_{i=1}^{n} \mathbb{P}(X_i \leq t)$.

Prop 17 (simulation de lois par méthode d'inversion). On peut simuler une v.a.r X de loi quelconque à partir d'une v.a. de loi $\mathcal{U}_{[0,1]}$. On définit $G: u \mapsto \inf\{x, \mathbb{P}(x) \geq u\}$ (*) sur $]0,1[$, avec F la fonction de répartition de la v.a qu'on veut simuler.

Soit $U \sim \mathcal{U}_{[0,1]}$, F une fonction de répartition et G définie par (*). Alors $Y = G(U)$ est une v.a. de fonction de répartition F.

C) Utilisation de lois usuelles en analyse:

Théorème 3 (Weierstrass): Soit $f: [0,1] \to \mathbb{C}$ une fonction continue ω son module de continuité uniforme, ie $\omega(h) = \sup\{|f(u) - f(v)|, |u-v| \leq h\}$. Pour $n \geq 1$, on considère le polynôme $B_n(f, x) = B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f(\frac{k}{n})$, le n-ième polynôme de Bernstein de f. Alors :

(i) B_n converge vers f uniformément sur $[0,1]$

(ii) $\|f - B_n\|_\infty \leq C\, \omega(\frac{1}{\sqrt{n}})$, C constante.

(iii) L'estimation de (ii) est optimale : il existe f lipschitzienne pour laquelle $\|f - B_n\|_\infty \geq \delta\, \omega(\frac{1}{\sqrt{n}})$, où $\delta > 0$.

III. Convergence en loi et applications:

A) Notion de convergence en loi:

Def 13: (X_n) suite de var, X var. On dit que (X_n) converge en loi vers X, $X_n \xrightarrow{\mathcal{L}} X$ si : pour toute fonction f continue bornée sur $\mathbb{R}$:

$$\mathbb{E}[f(X_n)] \xrightarrow[n \to +\infty]{} \mathbb{E}[f(X)].$$

Ex 11: Si les (X_n) ont un nombre fini de valeurs $\{x_i, 1 \leq i \leq N\}$, $(X_n)_n$ converge en loi vers X ssi $\lim_{n \to +\infty} \mathbb{P}(X_n = x_i) = \mathbb{P}(X = x_i)$ $\forall i \in [1, N]$.

• Soit $X_n \sim \mathcal{N}(0, \sigma_n^2)$ et $X \sim \mathcal{N}(0, \sigma^2)$. Si $\sigma_n \xrightarrow[n \to +\infty]{} \sigma > 0$, $X_n \xrightarrow{\mathcal{L}} X$.

Prop 18: Si $(X_n)_n$ et X var de fonctions de répartition respectives F_n et F.

$X_n \xrightarrow{\mathcal{L}} X \iff \forall x$ en lequel F est continue, $F_n(x) \xrightarrow[n \to +\infty]{} F(x)$

Coro 1: Si $X_n \xrightarrow{\mathcal{L}} X$, $(X_n)_n$ var et si X a une loi à densité, alors

$\forall a < b$, $\mathbb{P}(X_n \in]a, b]) \xrightarrow[n \to +\infty]{} \mathbb{P}(X \in]a, b])$.

Théorème 4 (Lévy). Soit $(X_n)_n$ une suite de var.

(i) Si $X_n \xrightarrow{\mathcal{L}} X$, alors ϕ_{X_n} converge simplement vers ϕ_X

(admis) (ii) Si les ϕ_{X_n} convergent simplement vers une fonction complexe ϕ sur $\mathbb{R}$, et si ϕ est continue en 0, alors ϕ est une fonction caractéristique d'une v.a.r X et $X_n \xrightarrow{\mathcal{L}} X$.

B) Le théorème de la limite centrale:

Dvlp 2

Théorème 5 (TCL):

Soit X une var tq $\mathbb{E}(X^2) < +\infty$, $\mu = \mathbb{E}[X]$ et $\sigma^2 = Var(X) > 0$.

Soit (X_n) une suite de var i.i.d tq $X_n \sim X$. On note $S_n = \sum_{i=1}^{n} X_i$. Alors on a :

$$\frac{S_n - n\mu}{\sqrt{n\sigma^2}} \xrightarrow{\mathcal{L}} \mathcal{N}(0,1).$$

Ex 12. Si $X_i \sim \mathcal{B}(1,p)$,

$\forall a < b$, $\lim_{n \to +\infty} \mathbb{P}\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) = \int_a^b e^{-t^2/2} dt$.

Cela peut permettre d'approximer p, et est utilisé pour construire des intervalles de confiance.

Théorème 6 (Limite central poissonien). Soit $S_n \sim \mathcal{B}(n, p_n)$.

Si $\lim_{n \to +\infty} np_n = \lambda > 0$, S_n converge en loi vers une variable aléatoire de Poisson de paramètre λ.

Rmq 3. Cela montre que la loi binomiale peut dans certains cas, pour n grand, être approximée par une loi de Poisson.