261. Loi d'une variable aléatoire : caractérisations. Exemples, applications.

Refs : [B-L] Barbe Ledoux Probabilités. [F-F] Foata Fuchs Calcul des probas [Can] Candelpergher Théorie des probas.

Devts : TCL, Stirling.

Cadre : On se fixe $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, par défaut X désignera une variable aléatoire $\Omega \to \mathbb{R}^d$, où $\mathbb{R}^d$ est muni de sa structure euclidienne usuelle et de sa tribu borélienne (on distingue d>1 où on parle de vecteur aléatoire).

I. Loi d'une variable aléatoire.

1) Définitions et premiers exemples.

[B-L] 44

Def 1 : On appelle loi (ou loi de probabilité) de la variable X la mesure image de $\mathbb{P}$ par X sur $\mathcal{B}(\mathbb{R}^d)$, définie par
$$P(X\in B) := \mathbb{P}_X(B) := \mathbb{P}(X^{-1}(B))$$

Rq 2 : La loi de X est bien une mesure sur $\mathbb{R}^d$ l'espace d'arrivée. La connaissance de Ω et a fortiori de X est "superflue" en comparaison à celle de la loi de X.

77 Prop Def 3 : Soit $X=(X_1,\dots,X_d)$ un vecteur aléatoire, on appelle i-ème marginale de X la loi de X_i. Les variables $X_1\dots X_d$ sont mutuellement indépendantes si et seulement si la loi de X est la mesure produit de ses lois marginales.

Rq 4 : Cette caractérisation de l'indépendance nous permettra d'utiliser le théorème de Fubini.

[B-L] 53

Théo 5 Transfert. Soit $X:(\Omega,\mathcal{A},\mathbb{P})\to(\mathbb{R}^d,\mathcal{B}(\mathbb{R}^d))$ un vecteur aléatoire et $\phi:\mathbb{R}^d\to\mathbb{R}$ une fonction borélienne, alors $\phi\circ X\in L^1(\mathbb{P})$ si et seulement si $\phi\in L^1(\mathbb{P}_X)$, avec $\int_\Omega \phi\circ X\, d\mathbb{P} = \int_{\mathbb{R}^d}\phi(x)\,d\mathbb{P}_X(x)$ (*)

2) Variables aléatoires discrètes

[F-F]

Def 6 : On dit que X est discrète si $X(\Omega)$ est fini ou dénombrable. Pour $i\in X(\Omega)$, on notera $P(X=i)$ pour $\mathbb{P}_X(\{i\})$.

Rq 7. Dans ce contexte, le théorème de transfert se reformule par $\int_\Omega \phi\circ X\, d\mathbb{P} = \sum_{i\in X(\Omega)}\phi(i)P(X=i)$.

Ex 8 : Pour $x_0\in\mathbb{R}^d$, on peut considérer la variable constante égale à x_0, avec $P(X=x_0)=1$ et $P(X=y)=0$ sinon.

A) Loi de Bernoulli, loi binomiale

La loi de Bernoulli sur $\mathbb{R}$ est la mesure $\mu \overset{\mathcal{B}(p)}{=} p\delta_1+(1-p)\delta_0$ où $p\in[0,1]$ et n un entier non nul. (On notera $X\rightsquigarrow\mu$ pour dire que la loi de X est μ). Cette loi modélise un tirage avec probabilité p de réussite.

[F-F]

La loi binomiale de paramètres n, p est définie par $\mathcal{B}(n,p)=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}\delta_k$, il s'agit de la loi suivie par la somme de n variables indep de loi de Bernoulli $\mathcal{B}(p)$. (Compte le nombre de succès dans une répétition d'expériences de Bernoulli indépendantes).

B) Loi géométrique.

Pour $p\in]0,1]$, on pose $\mathcal{G}(p)=\sum_{k\geq 1}(1-p)^{k-1}p\,\delta_k$. Modélise le premier succès dans une répétition indépendante d'expériences de Bernoulli $\mathcal{B}(p)$.

C) Loi de Poisson.

Pour $\lambda>0$, on pose $\mathcal{P}(\lambda)=\sum_{k\in\mathbb{N}}e^{-\lambda}\frac{\lambda^k}{k!}\delta_k$, modélise les "évènements" rares, approxime la loi binomiale pour "n grand et p petit" (convergence en loi, cf 34).

3) Variables aléatoires à densité.

[B-L]

Def 9 : On dit qu'une mesure μ sur $\mathbb{R}^d$ est à densité si il existe $f_\mu:\mathbb{R}^d\to\mathbb{R}$ mesurable telle que $\forall B\in\mathcal{B}(\mathbb{R}^d)$, $\mu(B)=\int_B f_\mu\, d\lambda$.

Dans le cas où μ est une mesure de probabilité, on remarque que f doit être positive et d'intégrale totale égale à 1.
Réciproquement, toute fonction mesurable positive d'intégrale 1 induit une unique loi de probabilités.

Prop 10. Si $X=(X_1,\dots,X_d)$ est un vecteur aléatoire de densité $f_X:\mathbb{R}^d\to\mathbb{R}$, alors X_i est à densité avec $f_{X_i}: x\mapsto\int_{\mathbb{R}^{d-1}}f(x_1\dots x\dots x_d)\,dx_1\dots dx_d$.

Rq 11 : La réciproque est fausse : si X est à densité, (X,X) ne charge que la diagonale et ne peut être à densité.

Prop 12 : Si $X_1\dots X_d$ sont à densités $f_1\dots f_d$ indep, alors $X=(X_1\dots X_d)$ est à densité donnée par $f_X=f_{X_1}f_{X_2}\dots f_{X_d}$. Réciproquement, si la densité du vecteur est le produit des marginales, celles-ci sont indeps.

(*) Théo 55 Si X_1 et X_2 sont à valeurs dans $\mathbb{R}^d$ et indépendantes, alors $\mathbb{P}_{X_1+X_2}=\mathbb{P}_{X_1}*\mathbb{P}_{X_2}$, si de plus X_1 et X_2 sont à densité, la densité de X_1+X_2 est la convolution de celle de X_1 et de X_2.

A) Loi uniforme.

(sur $K\subseteq\mathbb{R}^d$ compact) notée $U(K)$, elle est donnée par la densité $f : x \mapsto \frac{1}{\lambda(K)} \mathbb{1}_K(x)$. Cette loi modélise un "choix continu".

B) (Loi normale) Soient $m \in \mathbb{R}$ et $\sigma^2 > 0$, la loi normale $\mathcal{N}(m, \sigma^2)$ est définie par la densité $x \mapsto \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-m)^2}{2\sigma}}$. Elle apparaît comme une 'loi limite' (cf TCL) et donc utilisée fréquemment en statistiques. On dira que $\mathcal{N}(0,1)$ est centrée réduite.

C) (Loi exponentielle) Pour $a > 0$, la loi exponentielle $\mathcal{E}(a)$ est définie par la densité $x \mapsto \mathbb{1}_{\mathbb{R}_+} a e^{-ax}$. Elle modélise des phénomènes "sans vieillissement" c'est-à-dire tels que $P(X > t+s \mid X > t) = P(X > s)$ quand $X \sim \mathcal{E}(a)$. C'est par exemple le cas des désintégrations atomiques.

D) Loi de Cauchy $a \in \mathbb{R}$, $b > 0$. La loi de Cauchy $C(a,b)$ est donnée par la densité $x \mapsto \frac{1}{\pi} \frac{b}{(x-a)^2 + b^2}$.

Cette loi nous sera utile comme contre-exemple pour des questions de moments (cvg, v.a.d). Elle modélise des phénomènes d'optique. $\exists$ va ni disc ni à densité.

II Espérance et caractérisation des lois

Identifier la loi d'une variable fixée, ou montrer l'identité des lois de deux variables peut être assez difficile à faire sans développer d'outils plus sophistiqués.

[BL] 68.65-66

Def 13 : On dit que X admet un moment d'ordre p ($1 \le p < \infty$) si $X \in L^p(P)$. Par le théorème de transfert ceci est équivalent à demander $\int_{\mathbb{R}} |X|^p dP = \int_{\mathbb{R}^d} |x|^p dP_X < \infty$. Si X admet un moment d'ordre 1, on note $E(X) = \int_\Omega X dP$ l'espérance de X. Si X admet un moment d'ordre 2, on note $Var(X) = E[(X - EX)^2]$ la variance de X.

Rq 14 : Comme $(\Omega, \mathcal{A}, P)$ est probabilisé, on a les inclusions $L^p \subset L^q$ pour $p > q$. Donc la définition de la variance est bien posée.

Ex 15 : La connaissance des moments ne suffit pas à caractériser une loi. Si $X \sim \mathcal{N}(0,1)$ et $Z := e^X$ de densité $f_Z(z) = \frac{1}{\sqrt{2\pi}} \frac{1}{z} e^{-\frac{(\ln z)^2}{2}} \mathbb{1}_{\mathbb{R}_+^*}(z)$, $a \in [-1,1]$. Z_a de densité $f_a(x) = f_Z(x)(1 + a\sin(2\pi \ln x))$ $x > 0$, Z et Z_a ont mêmes moments mais pas même loi.

Comme $E(\mathbb{1}_B(X)) = P_X(B)$, la connaissance de $E(f(X))$ pour toute f mesurable positive caractérise la loi, on cherche à prendre une classe de f facile d'utilisation et qui caractérise la loi.

[Cou] 77-78

1) Fonction de répartition

Def 16 : Soit X une variable aléatoire réelle, on définit $F_X(t) = P(X \le t) = E(\mathbb{1}_{]-\infty, t]})$ la fonction de répartition de la loi P_X.

Prop 17 : La fonction de répartition F_X de X respecte les propriétés suivantes.
- $F_X \ge 0$, $F_X(t) \in [0,1]$, $\lim_{-\infty} F_X = 0$, $\lim_{+\infty} F_X = 1$, F_X croissante
- F_X est continue à gauche et limitée à droite
- $P(X=a) = F_X(a) - \lim_{t \to a^-} F_X(t)$. ($F_X$ est C^0 en a ssi $P(X=a)=0$).

Réciproquement, une fonction respectant les deux premières conditions est la fonction de répartition d'une certaine loi.

Prop 18 : Deux lois de probabilité ayant la même fonction de répartition sont égales : la fonction de répartition caractérise la loi.

Prop 19 : Si F_X est continue et C^1 par morceaux, alors P_X est à densité et sa densité est dérivée de sa fonction de répartition.

Exple 20 : Fonction de répartition d'une loi discrète ou d'une loi continue. Fig 1

Tableau des fonctions de répartition des lois classiques.

[Cou] 218 221.

2) Fonction génératrice.

Def 21 : Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, on appelle fonction génératrice de X la fonction G_X définie par $G_X(z) = E(z^X) = \sum_{n \ge 0} P(X=n) z^n$. Cette fonction est une série entière qui converge au moins sur $D(0,1)$.

Prop 22 : Si X et Y sont indépendantes alors $G_{X+Y} = G_X G_Y$. De plus la fonction G_X est infiniment dérivable en 0, avec $P(X=n) = \frac{G^{(n)}(0)}{n!}$, la fonction génératrice de X caractérise donc sa loi.

Théo 23 : La variable X admet une espérance si et seulement si la fonction G_X est dérivable à gauche en 1, on a alors $E(X) = G_X'(1)$.

Rq 24 : Cette formule peut être itérée.

Ex 25 Calculs des fonctions génératrices des lois à valeurs dans $\mathbb{N}$ usuelles. (La fonction gén de $B(n,p)$ s'obtient par indépendance)

3) Fonction caractéristique.

Def 26 On appelle fonction caractéristique de $\mathbb{P}_X$ la fonction complexe φ_X définie par

$$\varphi_X(t) := E(e^{i\langle t,X\rangle}). \quad \forall t \in \mathbb{R}^d.$$

où $\langle . \rangle$ est le produit scalaire usuel

Rq 27: Si $\mathbb{P}_X$ admet une densité f_X, alors $\varphi_X(t) = \hat{f}(-t)$ où $\hat{f}$ est la transformée de Fourier de f.

Rq 28: Si X est à valeurs dans $\mathbb{N}$, on a $\varphi_X(t) = G_X(e^{it})$.

Prop 29: Si X est une va, alors

- $\varphi_X(t)$ est uniformément continue
- $\varphi_X(0) = 1$, $\varphi_X(-t) = \overline{\varphi_X(t)}$ $|\varphi_X(t)| \leq 1$
- Si X et Y sont indep, alors $\varphi_{X+Y} = \varphi_X \varphi_Y$

Théo 30: Soit X une v.a réelle et φ_X sa fonction caractéristique. Si X admet un moment d'ordre k, alors φ_X est de classe C^k, avec $\varphi_X^{(k)}(t) = E((iX)^k e^{itX})$.
Réciproquement, si φ_X est de classe C^k, alors X a un moment d'ordre $2\left[\frac{k}{2}\right]$.

Application 31 Calcul des fonctions caractéristiques des lois normales et de Cauchy par l'analyse complexe.

Exemple 32. Calcul des fonctions caractéristiques des lois usuelles.

III Applications.

1) Convergence en loi

Def 33 Soient $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires et X une variable aléatoire.
On dit que X_n converge en loi vers X si pour toute fonction continue bornée de $\mathbb{R}^d$ dans $\mathbb{R}$, on a $\lim_{n\to\infty} E(f(X_n)) = E(f(X))$. (convergence étroite)

Ex 34: Si X_n suit $B(n,p_n)$ avec $np_n \to \lambda > 0$, alors X_n converge en loi vers une variable suivant $P(\lambda)$.

Rq 35 La convergence en loi ne fait pas appel à l'espace de départ : il peut changer pour les X_n ou pour X.

Théo 36 La convergence en loi est équivalente à la convergence simple des fonctions de répartition.

Théo 37 (Lévy). La convergence en loi est équivalente à la convergence simple des fonctions caractéristique.

2) Théorème central limite

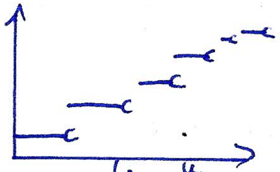
Théo 38 (Central limite). Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de var aléatoire indépendantes et identiquement distribuées admettant des moments d'ordre 2. On définit $Y_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - E(X_i))$
Cette suite variable converge en loi vers Y suivant $\mathcal{N}(0, Var(X_1))$.

Appli 39: Approximation d'une loi binomiale par une loi normale : si $X_n \sim B(n,p)$. Alors $\frac{X_n - np}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, p(1-p))$.

Appli 40 Calcul des intervalles de confiance en statistiques.

Appli 41 Formule de Stirling : $n! \sim \sqrt{2\pi}\, n^{n+\frac{1}{2}} e^{-n}$

Fig 1



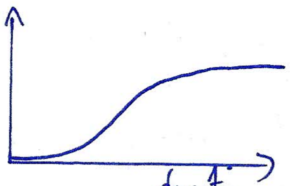


Fig 2:

$B(p)$	p	$1-p+pt$
$B(n,p)$	np	$(1-p+pt)^n$
$G(p)$	$1/p$	$(pt)/(1-(1-p)t)$
$P(\lambda)$	λ	$e^{\lambda(t-1)}$

$U[a,b]$	$\frac{a+b}{2}$	$\frac{1}{b-a}\,\frac{e^{itb}-e^{ita}}{it}$
$\mathcal{N}(m,\sigma^2)$	m	$e^{imt}\,e^{-\sigma^2t^2/2}$
$\mathcal{E}(\lambda)$	$1/\lambda$	$1/(1-it/\lambda)$
$C(a,b)$	$\emptyset$	$e^{imt}\,e^{-a\lvert t\rvert}$.