261 LOI D'UNE VARIABLE ALÉATOIRE : CARACTÉRISATIONS, EXEMPLES, APPLICATIONS.

Soit $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, et $(E, \mathcal{E})$ un espace probabilisable. Soit $d \in \mathbb{N}^*$.

I. Loi d'une variable aléatoire et premières caractérisations

I. A. Définitions et caractérisation des variables aléatoires discrètes [Ouv08, Ch1, p14–20] [BL07, §III.4/5, p52–71]

Définition 1. **[Variable aléatoire, loi d'une variable aléatoire]**
- Une variable aléatoire X est une fonction mesurable $X : (\Omega, \mathcal{A}) \longrightarrow (E, \mathcal{E})$.
- On note $\mathbb{P}_X$ la mesure image de $\mathbb{P}$ par X et on l'appelle loi de X.
- Si $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on dit que X est une variable aléatoire réelle.
- Si $X(\Omega)$ est dénombrable presque sûrement, on dit que X est discrète.
- Si $\mathbb{P}_X$ est absolument continue par rapport à la mesure de Lebesgue, on dit que X est à densité, de densité $d\mathbb{P}_X/d\,\mathrm{Leb}$.

Exemple 2. Variable aléatoire qui associe le résultat du lancer d'une pièce : $E = \{\mathrm{P}, \mathrm{F}\}$, le résultat de la couleur d'une boule tirée dans une urne : $E = \{\text{couleurs}\}$.
$\mathcal{B}(p)$ est discrète, $\mathcal{E}(\lambda)$ est continue.

Soit X une variable aléatoire discrète à valeurs dans E.

Proposition 3. *Enumérons $E' = (e_i)_{i\in I} \subset E$ où I est au plus dénombrable de sorte que $X(\Omega) \subset E'$ p.s.. Alors $\mathbb{P}_X$ est caractérisée par la famille $(p_i)_{i\in I}$ où $p_i = \mathbb{P}(X = e_i)$ et $\sum_{i\in I} p_i = 1$. De plus, on a $\mathbb{P}_X = \sum_{i\in I} p_i \delta_{e_i}$.*

Remarque 4. Réciproquement, si $f : E' \longrightarrow [0,1]$ est telle que $\sum_{i\in I} f(e_i) = 1$, alors il existe une unique probabilité $\mathbb{Q}$ sur $(E, \mathcal{E})$ telle que $\forall i \in I, \mathbb{Q}(e_i) = f(e_i)$. Une variable aléatoire de loi Q est alors discrète.

Exemple 5. $\mathcal{B}(p), \mathcal{B}(n,p), \mathcal{G}(p), \mathcal{P}(\lambda), \ldots$ et leur loi $\mathbb{P}_X$ associée **[voir annexe]**

I. B. Caractérisation par l'espérance de classes de fonctions [BL07, ChIII, p41]

Soit X une variable aléatoire à valeurs dans $E = \mathbb{R}^d$.

Remarque 6. Il est clair que $\mathbb{P}_X$ est entièrement déterminée par la donnée de $\mathbb{P}_X(A)$ pour $A \in \mathcal{B}(\mathbb{R}^d)$, ou pour $A \in \Lambda$ où Λ est un π-système générateur de $\mathcal{B}(\mathbb{R}^d)$.

Définition 7. **[Espérance]** Si X est réelle, on définit lorsque $\int_\Omega |X(\omega)|\ d\mathbb{P}(\omega) < +\infty$ l'espérance de X par $\mathbb{E}[X] = \int_\Omega X(\omega)\ d\mathbb{P}(\omega)$. Sinon, on définit l'espérance de $X = (X_1, \ldots, X_d)$ lorsque $\mathbb{E}[\|X\|] < +\infty$ par $\mathbb{E}[X] = (\mathbb{E}[X_1], \ldots, \mathbb{E}[X_d])$.

Exemple 8. Espérance de $\mathcal{B}(n,p), \mathcal{N}(0,1), \ldots$ **[voir annexe]**
Les lois de Cauchy n'admettent pas d'espérance.

Remarque 9. L'espérance ne dépend que de la loi de la variable aléatoire.

Proposition 10. **[Lemme de transfert]**
Soit $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ une fonction borélienne.
Alors $f(X) \in L^1(\Omega, \mathcal{A}, \mathbb{P}) \iff f \in L^1(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d), \mathbb{P}_X)$ et dans ce cas :

$$\mathbb{E}[f(X)] = \int_\Omega f(X(\omega))d\mathbb{P}(\omega) = \int_{\mathbb{R}^d} f(x)d\mathbb{P}_X(x)$$

Application 11. Si $X \sim \mathcal{E}(\lambda)$ avec $\lambda > 0$ et $\theta > 0$, alors $\theta X \sim \mathcal{E}(\lambda/\theta)$.

Application 12. **[Théorème de Weierstrass]**
$\mathbb{R}[X]$ est dense dans $\mathcal{C}([a,b], \mathbb{C}, \|.\|_\infty)$ pour $a < b \in \mathbb{R}$.

Proposition 13. *La donnée de $\mathbb{E}[f(X)]$ pour $f \in \mathcal{C}_c^+(\mathbb{R}^d)$ l'ensemble des fonctions continues positives à support compact caractérise la loi de X. [Ouv09, p10]*

Remarque 14. On peut aussi prendre pour famille de fonctions $\mathcal{C}_b(\mathbb{R}^d), \mathcal{C}_0(\mathbb{R}^d), \mathcal{C}_c^\infty(\mathbb{R}^d) \ldots$

Application 15. Si $U_1, U_2 \sim \mathcal{N}(0,1)$ et sont indépendantes, alors $U_1/U_2 \sim \mathcal{C}(0,1)$.

II. Caractérisations par des fonctions

II. A. Fonction de répartition [BL07, §III.2, p45]

Soit X une variable aléatoire réelle.

Définition 16. On définit F_X la fonction de répartition de X par $F_X(x) = \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x)$ pour $x \in \mathbb{R}$.

Exemple 17. Fonction de répartition de $\mathcal{B}(p), \mathcal{E}(\lambda)$.

PROPOSITION 18.
(i) F_X est croissante, continue à droite et $\lim_{-\infty} F_X = 0, \lim_{+\infty} F_X = 1$.
(ii) Réciproquement, toute fonction F satisfaisant ces hypothèses est la fonction de répartition d'une variable aléatoire.
(iii) Si f est une densité de probabilité, alors $F : x \longmapsto \int_{-\infty}^{x} f(t)dt$ est la fonction de répartition d'une variable aléatoire Y telle que $\mathbb{P}_Y$ admette pour densité f.

APPLICATION 19. On définit l'inverse généralisée de F_X par $F^{\leftarrow}(u) = \inf\{t \mid F_X(t) \geq u\}$ pour $u \in [0,1]$. Alors si $U \sim \mathcal{U}([0,1])$, on a que $F^{\leftarrow}(U) \sim \mathbb{P}_X$. Par exemple $-\ln(1-U) \sim \mathcal{E}(1)$.

PROPOSITION 20. *F_X caractérise $\mathbb{P}_X$.*

REMARQUE 21. On peut généraliser la notion de fonction de répartition pour une variable aléatoire à valeurs dans $\mathbb{R}^d$. Le résultat reste alors valable.

II. B. Fonction caractéristique [BL07, §III.5, p61]

Soit X une variable aléatoire à valeurs dans $\mathbb{R}^d$.

DÉFINITION 22. [FONCTION CARACTÉRISTIQUE]
On définit ϕ_X la fonction caractéristique de X par $\phi_X(\lambda) = \mathbb{E}\left[\mathrm{e}^{i\langle\lambda|X\rangle}\right]$ pour $\lambda \in \mathbb{R}^d$.

EXEMPLE 23. Fonctions caractéristiques de $\mathcal{B}(n,p), \mathcal{E}(\lambda)\ldots$ **[VOIR ANNEXE]**

PROPOSITION 24. *ϕ_X caractérise $\mathbb{P}_X$.*

PROPOSITION 25. *Si X est réelle et admet un moment d'ordre p, alors ϕ_X est p fois dérivable et on a $\phi_X^{(p)}(\lambda) = i^p\mathbb{E}[X^p\,\mathrm{e}^{i\lambda X}]$. En particulier, $\phi^{(p)}(0) = i^p\mathbb{E}[X^p]$.*

REMARQUE 26. [Ouv09, p205] Réciproquement, si p est pair est ϕ_X est p fois dérivable en 0, alors X admet un moment d'ordre p. On peut construire X n'admettant pas d'espérance mais telle que ϕ_X soit dérivable en 0 (admis).

II. C. Fonction génératrice [Ouv08, §5.3, p138]

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ presque sûrement.

DÉFINITION 27. [FONCTION GÉNÉRATRICE]
On définit g_X la fonction génératrice de X par $g_X(s) = \mathbb{E}[s^X] = \sum_{k\geq 0}\mathbb{P}(X = k)s^k$ (lorsque la série converge).

REMARQUE 28. Le rayon de convergence de cette série est au moins égal à 1 et $G_X(1) = 1$.

EXEMPLE 29. Fonctions génératrices de $\mathcal{B}(p), \mathcal{B}(n,p), \mathcal{P}(\lambda), \mathcal{G}(p)\ldots$ **[VOIR ANNEXE]**

PROPOSITION 30. *G_X caractérise $\mathbb{P}_X$. En particulier, $\forall k \in \mathbb{N}, \mathbb{P}(X = k) = \frac{G^{(k)}(0)}{k!}$.*

PROPOSITION 31. *X admet un moment d'ordre $p \in \mathbb{N}^*$ si et seulement si g_X est p fois dérivable en 1, et dans ce cas $G_X^{(p)}(1) = \mathbb{E}[X(X-1)\ldots(X-p)]$.*

APPLICATION 32. Expression de l'espérance, la variance, ou plus généralement d'un moment en fonction des dérivées en 0 de G_X.

APPLICATION 33. Calcul de la loi d'une variable aléatoire multinomiale dont le paramètre n n'est pas déterministe mais suit une loi $\mathcal{P}(\lambda)$.

II. D. Moments et fonction génératrice des moments [BL07, §III.5, p66]

Soit X une variable aléatoire réelle.

DÉFINITION 34. [MOMENT ET MOMENT CENTRÉ D'ORDRE p]
Si $\mathbb{E}[|X|^p] < +\infty$, on dit que X admet un moment d'ordre p (ou $X \in L^p$) et alors on appelle $\mathbb{E}[X^p]$ le moment d'ordre p et $\mathbb{E}[(X - \mathbb{E}[X])^p]$ le moment centré d'ordre p de X.

PROPOSITION 35. *Si $X \in L^p$ alors $X \in L^q$ pour tout $1 \leq q \leq p$.*

DÉFINITION 36. [FONCTION GÉNÉRATRICE DES MOMENTS]
On définit M_X la fonction génératrice des moments de X par $M_X(\lambda) = \mathbb{E}[\mathrm{e}^{\lambda X}]$ pour $\lambda \in \mathbb{R}$ tel que cette quantité soit définie.

EXEMPLE 37. $M_{\mathcal{N}(0,1)} : \lambda \longmapsto \exp(\lambda^2/2)$ et $M_{\mathcal{R}(1/2)} : \lambda \longmapsto \mathrm{ch}(\lambda)$.

PROPOSITION 38. *Si M_X est définie sur un voisinage de 0, alors M_X caractérise $\mathbb{P}_X$ et alors on a $M_X(\lambda) = \sum_{k\geq 0}\frac{t^k}{k!}\mathbb{E}[X^k]$ au voisinage de 0, en particulier X admet des moments de tout ordre et $\mathbb{E}[X^p] = M_X^{(p)}(0)$ pour tout p.*

THÉORÈME 39. [THÉORÈME DES MOMENTS] *(admis)*
Si X et Y sont bornées p.s. et si leurs moments sont égaux à tout ordre, alors $X \sim Y$.

REMARQUE 40. C'est faux si X et Y ne sont pas bornées : les considérer de densités $f_X : x \longmapsto \frac{1}{\sqrt{2\pi}x} \mathrm{e}^{-\ln(x)^2/2} \mathbb{1}_{\mathbb{R}^+}(x)$ et $f_Y : x \longmapsto f_X(x)(1+a\sin(2\pi\ln(x)))$ pour un $a \in [-1,1]$.

III. Vecteurs aléatoires et indépendance

[BL07, ChIV, p73]

Soient X, Y deux variables aléatoires à valeurs dans E, F des espaces probabilisables.

DÉFINITION 41. [LOIS CONJOINTES, LOIS MARGINALES]
La loi du couple (X,Y) est appelée loi conjointe.
Les lois de X et de Y sont appelées lois marginales.

REMARQUE 42. La donnée de la loi conjointe permet de retrouver les lois marginales. La réciproque est fausse : si $X \perp\!\!\!\perp Y \sim \mathcal{B}(1/2)$, alors (X,X) et (X,Y) ont même lois marginales mais pas même loi conjointe.

PROPOSITION 43. *X et Y sont indépendantes si et seulement si :*
(i) $\mathbb{P}_{(X,Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$,
et lorsque $E = \mathbb{R}^d$ et $F = \mathbb{R}^\ell$:
(ii) $\mathbb{E}[f_1(X)f_2(Y)] = \mathbb{E}[f_1(X)]\mathbb{E}[f_2(X)]$ *pour toutes fonctions boréliennes f_1, f_2 définies sur $\mathbb{R}^d$ et $\mathbb{R}^\ell$ respectivement,*
(iii) $\forall(\lambda_1,\lambda_2) \in \mathbb{R}^d \times \mathbb{R}^\ell, \phi_{(X,Y)}(\lambda_1,\lambda_2) = \phi_X(\lambda_1)\phi_Y(\lambda_2)$.
Si X et Y sont indépendantes, alors $\mathrm{Cov}(X,Y) = 0$, $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$ et $G_{X+Y} = G_X G_Y$ lorsque ces quantités sont définies.

REMARQUE 44. On peut généraliser ces propositions à des familles quelconques de v.a..

APPLICATION 45. Un vecteur gaussien est une variable aléatoire à valeurs dans $\mathbb{R}^d$ telle que toute combinaison linéaire de ses composantes suit une loi normale. Les composantes d'un vecteur gaussien sont indépendantes si et seulement si sa matrice de covariance est diagonale.

IV. Convergence en loi

[BL07, §V.4, p121]

Soient $(X_n)_{n\in\mathbb{N}}$ et X des variables aléatoires à valeurs dans $\mathbb{R}^d$.

DÉFINITION 46. [CONVERGENCE EN LOI]
On dit que $(X_n)_{n\in\mathbb{N}}$ converge en loi vers X si pour tout fonction $f \in \mathcal{C}_b(\mathbb{R}^d)$, on a $\mathbb{E}[f(X_n)] \underset{n\to+\infty}{\longrightarrow} \mathbb{E}[f(X)]$.

EXEMPLE 47. $X_n \sim \mathcal{N}(0,\sigma_n^2)$ converge en loi vers $X \sim \mathcal{N}(0,\sigma^2)$ si $\sigma_n \to \sigma$.

THÉORÈME 48. [THÉORÈME DE LÉVY (FAIBLE)]
$X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X$ si et seulement si $\phi_{X_n} \underset{n\to+\infty}{\overset{s}{\longrightarrow}} \phi_X$.

Supposons maintenant $d = 1$ et les variables aléatoires réelles.

PROPOSITION 49. *Si les variables sont à valeurs dans $\mathbb{N}$, $X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X$ si et seulement si $\forall k \in \mathbb{N}, \mathbb{P}(X_n = k) \underset{n\to+\infty}{\longrightarrow} \mathbb{P}(X = k)$.*

PROPOSITION 50. *$X_n \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} X$ si et seulement si $F_{X_n} \underset{n\to+\infty}{\longrightarrow} F_X$ en tout point de continuité de F_X.*

THÉORÈME 51. [THÉORÈME CENTRAL LIMITE]
Si les $(X_i)_{i\in\mathbb{N}^*}$ sont i.i.d. de carré intégrable, alors :

$$\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - \mathbb{E}[X_1]) \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} Z \sim \mathcal{N}(0, \mathrm{Var}(X_1))$$

APPLICATION 52. [INTERVALLE DE CONFIANCE ASYMPTOTIQUE] [RS12, §3.4, p25]
Soient $(X_i)_{i\in\mathbb{N}^*}$ des réalisations i.i.d. de $\mathcal{B}(p)$ pour un $p \in [0,1]$ inconnu. Le théorème central limite donne un intervalle de confiance asymptotique de niveau α pour p en fonction de la moyenne empirique $\hat{p}_n = \frac{1}{n}\sum_{k=1}^n X_i$: $\mathrm{IC}_\alpha(p) = \left[\hat{p}_n \pm \frac{q_{1-\alpha/2}}{2}\frac{1}{\sqrt{n}}\right]$, où q_t est le quantile d'ordre t de $\mathcal{N}(0,1)$.

APPLICATION 53. [MÉTHODE DE MONTE-CARLO] [Mé11]
Soit $f : [0,1] \longrightarrow \mathbb{R}$ intégrable par rapport à la mesure de LEBESGUE et $(X_i)_{n\in\mathbb{N}^*}$ une suite de variables aléatoiresi.i.d. de loi $\mathcal{U}([0,1])$. Alors $I_n = \frac{1}{n}\sum_{i=1}^n f(X_i) \underset{n\to+\infty}{\longrightarrow} \int_0^1 f(x)dx$ p.s..

THÉORÈME 54. [THÉORÈME DE SLUTSKY]
Reprenons les $(X_n)_n$ et X dans $\mathbb{R}^d$ et soient $(Y_n)_{n\in\mathbb{N}}$ des variables aléatoires de $\mathbb{R}^\ell$ ($\ell \in \mathbb{N}^*$) convergeant en probabilités vers une constante c. Alors $(X_n, Y_n) \underset{n\to+\infty}{\overset{\mathcal{L}}{\longrightarrow}} (X, c)$.

APPLICATION 55. [INTERVALLE DE CONFIANCE ASYMPTOTIQUE 2]
Dans le même cadre on a l'intervalle de confiance asymptotique : $\widetilde{\mathrm{IC}}_\alpha(p) = \left[\hat{p}_n \pm q_{1-\alpha/2}\frac{\sqrt{\hat{p}_n(1-\hat{p}_n)}}{\sqrt{n}}\right]$.

ANNEXE

[BL07, Appendice, p229]

Tableau des principales lois discrètes

Soit X une loi discrète et $p \in [0,1], q \in]0,1[, n \in \mathbb{N}, \lambda > 0$.

$\mathcal{L}(X)$	$\mathbb{P}_X$	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$\phi_X(t)$	$G_X(t)$
$\mathcal{U}([\![1,n]\!])$	$\sum_{k=1}^{n} \frac{1}{n}\delta_k$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$	$\frac{1}{n}\sum_{k=1}^{n} \mathrm{e}^{ikt}$	$\frac{1}{n}\frac{t-t^{n+1}}{1-t} \quad (t \neq 1)$
$\mathcal{B}(p)$	$p\delta_1 + (1-p)\delta_0$	p	$p(1-p)$	$p\,\mathrm{e}^{it} + (1-p)$	$1-p+pt$
$\mathcal{B}(n,p)$	$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}\delta_k$	np	$np(1-p)$	$p\,\mathrm{e}^{it} + (1-p)^n$	$(1-p+pt)^n$
$\mathcal{P}(\lambda)$	$\sum_{k=0}^{\infty} \frac{\lambda^k \mathrm{e}^{-\lambda}}{k!}\delta_k$	λ	λ	$\mathrm{e}^{\lambda(\mathrm{e}^{it}-1)}$	$\mathrm{e}^{(\lambda-1)t}$
$\mathcal{G}(q)$	$\sum_{k=1}^{\infty} q(1-q)^{k-1}\delta_k$	$\frac{1}{q}$	$\frac{1-q}{q^2}$	$\frac{q\,\mathrm{e}^{it}}{1-(1-q)\,\mathrm{e}^{it}}$	$\frac{pt}{1-(1-p)t}$

Tableau des principales lois à densité

Soit X une loi admettant une densité f par rapport à la mesure de LEBESGUE. Soit $a < b \in \mathbb{R}, m \in \mathbb{R}, \sigma > 0, \lambda > 0, \alpha, \beta > 0$.

$\mathcal{L}(X)$	$f(t)$	$\mathbb{E}[X]$	$\mathrm{Var}(X)$	$\phi_X(t)$
$\mathcal{U}([a,b])$	$\frac{\mathbb{1}_{[a,b]}(t)}{b-a}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	$\frac{\mathrm{e}^{itb}-\mathrm{e}^{ita}}{it(b-a)}$
$\mathcal{N}(m,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2	$\mathrm{e}^{imt-\frac{t^2\sigma^2}{2}}$
$\mathcal{C}(m,\sigma^2)$	$\frac{1}{\pi\sigma}\frac{1}{1+(\frac{x-m}{\sigma})^2}$	/	/	$\mathrm{e}^{-\lvert t\rvert}$
$\mathcal{E}(\lambda)$	$\lambda \exp(-\lambda t)\mathbb{1}_{[0,+\infty[}(t)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$	$\frac{1}{1-\frac{it}{\lambda}}$
$\Gamma(\alpha,\beta)$	$\frac{\beta^\alpha}{\Gamma(\alpha)}\mathrm{e}^{-\beta t}\, t^{\alpha-1}\mathbb{1}_{[0,+\infty[}(t)$	$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\left(1-\frac{it}{\beta}\right)^{-\alpha}$

COMMENTAIRES

Point de vue historique : la théorie des probabilités a été formalisée dans les années 30 par KOLMOGOROV. Théorème de DE MOIVRE-LAPLACE $\to$ TCL dans le cadre d'une loi binomiale.
Aujourd'hui point de vue fonctionnel : les variables aléatoires ... il est plus intéressant d'étudier la loi d'une variable aléatoire plutôt que la variable aléatoire en elle-même.

QUESTIONS

Q Une variable aléatoire sans atomes admet-elle nécessairement une densité ?

R Non ! Penser par exemple à l'escalier de CANTOR ... C'est la fonction de répartition d'une certaine variable aléatoire X ! On a alors $\mathbb{P}(X \in K) = 1$ où K est l'ensemble de CANTOR. La loi est sans atome et n'admet pas de densité car $\mathrm{Leb}(K) = 0$.

BIBLIOGRAPHIE

[BL07] P. BARBE et M. LEDOUX : *Probabilité*. EDP Sciences, 2007.

[Mé11] S. MÉLÉARD : *Aléatoire : introduction à la théorie et au calcul des probabilités*. Les éditions de l'École Polytechnique, 2011.

[Ouv08] J.-Y. OUVRARD : *Probabilités : Tome 1*. Cassini, 2ème édition, 2008.

[Ouv09] J.-Y. OUVRARD : *Probabilités : Tome 2*. Cassini, 3ème édition, 2009.

[RS12] V. RIVOIRARD et G. STOLTZ : *Statistique mathématique en action*. Vuibert, 2 ème édition, 2012.