Variables aléatoires

1. Définitions et notations

Une **variable aléatoire** à valeurs dans un espace mesurable $(E,\mathcal{E})$ est une application mesurable

$$\boldsymbol{X} : \Omega \longrightarrow E \quad \in \mathcal{M}\big((\Omega,\mathcal{A}),(E,\mathcal{E})\big)$$

Souvent en probabilités, on ne s'intéresse pas à l'espace de probabilité $(\Omega,\mathcal{A},\mathbb{p})$, mais plutôt aux propriétés de variables aléatoires sur cet espace. On note $\mathbf{0}$ et $\mathbf{1}$ les variables déterministes constantes/ω à 0 et 1.

La **loi d'une variable aléatoire $\boldsymbol{X}$** à valeurs dans E est la mesure-image $\mathbb{p}_{\boldsymbol{X}} = \boldsymbol{X}_\star\mathbb{p}$ de $\mathbb{p}$ par $\boldsymbol{X}$:

$$\forall X \in \mathcal{E}, \qquad \mathbb{p}_{\boldsymbol{X}}(X) = \mathbb{p}\big(\boldsymbol{X}^{\leftarrow}(X)\big) = \mathbb{p}[\boldsymbol{X} \in X] = \mathbb{p}\big(\{\omega \in \Omega : \boldsymbol{X}(\omega) \in X\}\big)$$

$\mathbb{p}_{\boldsymbol{X}}(X)$ est la probabilité que la valeur de la variable aléatoire $\boldsymbol{X}$ tombe dans X lors d'une expérience modélisé par l'espace $(\Omega,\mathcal{A},\mathbb{p})$. Ainsi $(E,\mathcal{E},\mathbb{p}_{\boldsymbol{X}})$ est un espace probabilisé, et on oublie souvent Ω.

Si des variables aléatoires $\boldsymbol{X}$ et $\boldsymbol{Y}$ ont la même loi $\mathbb{p}_{\boldsymbol{X}} = \mathbb{p}_{\boldsymbol{Y}}$, elles sont dites **identiquement distribuées**, et on écrit $\boldsymbol{X} \leftrightsquigarrow \boldsymbol{Y}$. En particulier, $\boldsymbol{X} \overset{\text{p.s.}}{=} \boldsymbol{Y} \implies \boldsymbol{X} \leftrightsquigarrow \boldsymbol{Y}$.

⚠ Ce n'est pas parce qu'elles ont la même loi qu'elles sont égales, ni même égales presque sûrement.

Si $\forall i \in [\![1,n]\!]$, $\boldsymbol{X}_i$ est une varibale aléatoire sur $(E_i,\mathcal{E}_i)$, on considère la variable aléatoire conjointe

$$\vec{\boldsymbol{X}} = \overrightarrow{(\boldsymbol{X}_i)}_{1\leqslant i\leqslant n} : \begin{array}{rcl} \Omega & \longrightarrow & \bigtimes_{i=1}^n \mathcal{E}_i \\ \omega & \longmapsto & \big(\boldsymbol{X}_i(\omega)\big)_{1\leqslant i\leqslant n} \end{array} \quad \text{sur l'espace produit} \quad \left(\bigtimes_{i=1}^n E_i, \bigotimes_{i=1}^n \mathcal{E}_i\right)$$

La loi de $\vec{\boldsymbol{X}}$, dite **loi conjointe** de $(\boldsymbol{X}_i)_{1\leqslant i\leqslant n}$, est déterminée par sa valeur sur les *pavés* :

$$\forall (X_i)_{1\leqslant i\leqslant n} \in \bigtimes_{i=1}^n \mathcal{E}_i, \quad \mathbb{p}_{\vec{\boldsymbol{X}}}\big(\bigtimes_{i=1}^n X_i\big) = \mathbb{p}\big(\bigcap_{i=1}^n [\boldsymbol{X}_i \in X_i]\big) = \mathbb{p}\big[\bigwedge_{i=1}^n \boldsymbol{X}_i \in X_i\big]$$

par le *théorème d'unicité des probabilités* puisque la classe des pavés $\mathcal{P}av\big((\mathcal{E}_i)_{1\leqslant i\leqslant n}\big)$ est stable par intersection finie et engendre la tribu produit $\bigotimes_{i=1}^n \mathcal{E}_i$.

→ Lorsque $\boldsymbol{X}$ est à valeurs dans $(\mathbb{R},\mathfrak{B}(\mathbb{R}))$, on dit que $\boldsymbol{X}$ est une **variable aléatoire réelle** (V.A.R.). Lorsque $\vec{\boldsymbol{X}}$ est à valeurs dans $(\mathbb{R}^d,\mathfrak{B}(\mathbb{R}^d))$, on parle de **vecteur aléatoire** (réel) et on considère ses composantes $(\boldsymbol{X}_i)_{1\leqslant i\leqslant d}$.

La somme, le produit, le quotient (lorsque définit) de deux V.A.R. est une V.A.R., le supremum et l'infimum, la limite inférieur/supérieure, la limite (lorsqu'elle existe) d'une suite de V.A.R. est une V.A.R., et une fonction indicatrice est une V.A.R.

→ Lorsque E est *dénombrable* et $\mathcal{E} = \wp(E)$, on dit que $\boldsymbol{X}$ est une **variable aléatoire discrète**. Et alors

$$\mathbb{p}_{\boldsymbol{X}} = \sum_{x\in E} p_x\,\delta_x \quad \text{avec} \quad \forall x \in E,\ p_x := \mathbb{p}[\boldsymbol{X} = x] = \mathbb{p}\big(\boldsymbol{X}^{\leftarrow}(\{x\})\big)$$

[comme pour les probablités sur Ω dénombrable], la famille $(p_x)_{x\in E} \in [0,1]^E$ étant sommable de somme :

$$\sum_{x\in E} p_x = 1$$

Lorsque $p_x \neq 0$, alors $\mathbb{p}_{\boldsymbol{X}}(x) = p_x \neq 0$, c'est à dire que x est un **atome** de la loi de $\boldsymbol{X}$.

On étend cette notion aux variables aléatoires telles que l'image $\boldsymbol{X}(\Omega)$ est dénombrable, même si E ne l'est pas : il suffit de se ramener à $E = \boldsymbol{X}(\Omega)$ en retreignant à $\boldsymbol{X} : \Omega \longrightarrow \boldsymbol{X}(\Omega)$. En résumé,

$$\forall A \subset \boldsymbol{X}(\Omega), \quad \mathbb{p}[\boldsymbol{X} \in A] = \sum_{a\in A} \mathbb{p}[\boldsymbol{X} = a]$$

$\rightarrow$ Lorsque $\boldsymbol{X}$ est à valeurs dans $(\mathbb{R}^d, \mathfrak{B}(\mathbb{R}^d))$ et que sa loi $\mathbb{p}_{\boldsymbol{X}}$ est une mesure *absolument continue par rapport à la mesure de Lebesgue* λ, c'est à dire que

$$\forall X \in \mathcal{E}, \quad \lambda(X) = 0 \implies \mathbb{p}_{\boldsymbol{X}}(X) = 0$$

on dit que $\boldsymbol{X}$ est une **variable aléatoire à densité**. En particulier, $\mathbb{p}_{\boldsymbol{X}}$ ne possède pas d'atomes : elle est **diffuse**. Par la *théorème de Radon-Nikodym*, $\exists!_{\lambda\text{-pp}}\, \rho_{\boldsymbol{X}} : \mathbb{R}^d \longrightarrow \mathbb{R}_+ \in \mathcal{M}(\mathbb{R}^d)$ borélienne telle que

$$\forall B \in \mathfrak{B}(\mathbb{R}^d), \quad \mathbb{p}_{\boldsymbol{X}}(B) = \mathbb{p}[\boldsymbol{X} \in B] = \int_B \rho_{\boldsymbol{X}} \,\mathrm{d}\lambda = \int_B \rho_{\boldsymbol{X}}(x) \,\mathrm{d}x$$

On appelle cette fonction $\rho = \rho_{\boldsymbol{X}} = \frac{\mathrm{d}\mathbb{p}_{\boldsymbol{X}}}{\mathrm{d}\lambda}$, unique à λ-pp près, la **densité** de $\boldsymbol{X}$. Elle caractérise $\boldsymbol{X}$. Par extension, on parle de **densité de la loi** de $\boldsymbol{X}$. Puisque $\mathbb{p}_{\boldsymbol{X}}(\mathbb{R}^d) = 1$, on a la normalisation

$$\int_{\mathbb{R}^d} \rho \,\mathrm{d}\lambda = 1$$

En dimension 1 ($\boldsymbol{X}$ V.A.R.), $\forall a \leqslant b \in \mathbb{R}$,

$$\mathbb{p}\big[\boldsymbol{X} \in [a,b]\big] = \mathbb{p}\big[a \leqslant \boldsymbol{X} \leqslant b\big] = \int_a^b \rho_{\boldsymbol{X}}(x) \,\mathrm{d}x$$

2. Espérance et moments d'un variable aléatoire *réelle*

Soit $\boldsymbol{X} : \Omega \longrightarrow \mathbb{R}$ une variable aléatoire réelle (donc borélienne). On note son **espérance**

$$\mathbb{E}(\boldsymbol{X}) := \int_\Omega \boldsymbol{X} \,\mathrm{d}\mathbb{p} = \int_\Omega \boldsymbol{X}(\omega) \,\mathrm{d}\mathbb{p}(\omega)$$

qui est bien définie lorsque $\boldsymbol{X} \geqslant 0$ positive ou lorsque $\boldsymbol{X}$ est **intégrable** (ou d'espérance finie) :
$\|\boldsymbol{X}\|_1 = \mathbb{E}(|\boldsymbol{X}|) = \int |\boldsymbol{X}| \,\mathrm{d}\mathbb{p} < +\infty$, c'est à dire $\boldsymbol{X} \in \mathscr{L}^1((\Omega, \mathcal{A}, \mathbb{p}), \mathbb{R})$.

Lorsque $\boldsymbol{X}$ est à valeurs complexes ou que $\boldsymbol{X}$ est un vecteur aléatoire, on généralise l'espérance et les espaces $\mathscr{L}^p$ de manière usuelle, en particulier $\mathbb{E}((\boldsymbol{X}_i)_{1 \leqslant i \leqslant d}) = (\mathbb{E}(\boldsymbol{X}_i))_{1 \leqslant i \leqslant d}$.

Puisque l'espérance est simplement l'intégrale de $\boldsymbol{X}$ contre la mesure $\mathbb{p}$, on a

- par définition, $\forall A \in \mathcal{A}$, $\mathbb{E}(\mathbb{1}_A) = \mathbb{p}(A)$
- linéarité de l'espérance : $\mathbb{E}(\alpha \boldsymbol{X} + \boldsymbol{Y}) = \alpha\, \mathbb{E}(\boldsymbol{X}) + \mathbb{E}(\boldsymbol{Y})$ pour $\alpha \in \mathbb{R}$ (ou $\mathbb{R}_+$) et $\boldsymbol{X}$,$\boldsymbol{Y}$ V.A.R. intégrables (ou positives)
- positivité de l'espérance : si $\boldsymbol{X} \geqslant 0$, $\mathbb{E}(\boldsymbol{X}) \geqslant 0$ et alors $\mathbb{E}(\boldsymbol{X}) = 0 \implies \boldsymbol{X} \overset{\text{p.s.}}{=} \boldsymbol{0}$
- intégrale d'une constante : si $a \in \mathbb{R}$ et $\boldsymbol{X} \overset{\text{p.s.}}{=} a\,\boldsymbol{1}$, on a $\mathbb{E}(\boldsymbol{X}) = a$, car $\mathbb{E}(\boldsymbol{1}) = 1$.
- inégalité triangulaire : $|\mathbb{E}(\boldsymbol{X})| \leqslant \mathbb{E}(|\boldsymbol{X}|)$ si $\boldsymbol{X}$ intégrable
- intégrabilité contre une mesure de probabilité : si $\exists c > 0$: $|\boldsymbol{X}| \overset{\text{p.s.}}{\leqslant} c$, alors $\boldsymbol{X}$ est intégrable
- tous les théorèmes généraux sur des suites $(\boldsymbol{X}_n)$ de V.A.R. sur Ω :
    - convergence monotone : si $\boldsymbol{X}_n \geqslant 0$ suite croissante, $\quad \mathbb{E}(\boldsymbol{X}_n) \xrightarrow[n \to \infty]{\nearrow} \mathbb{E}\Big(\lim\limits_{n \to \infty} \boldsymbol{X}_n\Big)$
    - inégalité de Fatou : si $\boldsymbol{X}_n \geqslant 0$, $\quad \mathbb{E}\Big(\liminf\limits_{n \to \infty} \boldsymbol{X}_n\Big) \leqslant \liminf\limits_{n \to \infty} \mathbb{E}(\boldsymbol{X}_n)$

- convergence dominée : si $\exists \boldsymbol{Z} \in \mathscr{L}^1$: $\forall n,\ |\boldsymbol{X}_n| \leqslant \boldsymbol{Z}$ et $\boldsymbol{X}_n \xrightarrow[n\to\infty]{\text{p.s.}} \boldsymbol{X}$,
$$\lim_{n\to\infty} \mathbb{E}(\boldsymbol{X}_n) = \mathbb{E}\Big(\lim_{n\to\infty} \boldsymbol{X}_n\Big) \qquad \text{(en particulier lorsque les } \boldsymbol{X}_n \text{ sont unif. bornées)}$$

- inégalité de Hölder : $\mathbb{E}(|\boldsymbol{X}\boldsymbol{Y}|) \leqslant \mathbb{E}\big(|\boldsymbol{X}|^p\big)^{1/p} \cdot \mathbb{E}\big(|\boldsymbol{Y}|^q\big)^{1/q}$ pour $1 \leqslant p, q \leqslant +\infty$ si $\frac{1}{p} + \frac{1}{q} = 1$

- inégalité de Cauchy-Schwarz : $\mathbb{E}(|\boldsymbol{X}\boldsymbol{Y}|)^2 \leqslant \mathbb{E}\big(\boldsymbol{X}^2\big) \cdot \mathbb{E}\big(\boldsymbol{Y}^2\big)$
  et en particulier, $\mathbb{E}(|\boldsymbol{X}|)^2 \leqslant \mathbb{E}\big(\boldsymbol{X}^2\big)$

- inégalité de Jensen : si $f : I \to \mathbb{R}$ est une fonction *convexe* sur un intervalle $I \supset \boldsymbol{X}(\Omega)$
  alors $f\big(\mathbb{E}(\boldsymbol{X})\big) \leqslant \mathbb{E}\big(f(\boldsymbol{X})\big)$

On dit que $\boldsymbol{X}$ possède un **moment d'ordre p** si $\boldsymbol{X} \in \mathscr{L}^p((\Omega, \mathcal{A}, \mathbb{p}), \mathbb{R})$, c'est à dire $\mathbb{E}(|\boldsymbol{X}|^p) < +\infty$.

3. Fonction de variable alétoire et Théorème de transfert

Soit $\boldsymbol{X}$ une variable aléatoire à valeurs dans $(E, \mathcal{E})$ et $f : E \longrightarrow \mathbb{R}\ \in \mathcal{M}\big((E,\mathcal{E}), \mathbb{R}\big)$ (ou $\mathbb{C}$, ou $\mathbb{R}^d$) une fonction borélienne. Alors si $f \geqslant 0$ positive ou si $f(\boldsymbol{X})$ intégrable ($\mathbb{E}(|f(\boldsymbol{X})|) < +\infty$), on a

$$\mathbb{E}\big(f(\boldsymbol{X})\big) = \int_\Omega f(\boldsymbol{X}(\omega))\,\mathrm{d}\mathbb{p}(\omega) = \int_E f\,\mathrm{d}\mathbb{p}_{\boldsymbol{X}} = \int_E f(x)\,\mathrm{d}\mathbb{p}_{\boldsymbol{X}}(x) \quad \text{avec } f(\boldsymbol{X}) := f \circ \boldsymbol{X}$$

Rappel : $f(\boldsymbol{X})$ est bien une variable aléatoire réelle par composition d'une borélienne avec une mesurable. Ce théorème est une propriété générale d'une mesure-image : $\phi \leftrightarrow \boldsymbol{X}$, $E \leftrightarrow \Omega$, $(F, \mathcal{B}) \leftrightarrow (E, \mathcal{E})$, $\mu \leftrightarrow \mathbb{p}$, $\phi_\star\mu \leftrightarrow \boldsymbol{X}_\star\mathbb{p} = \mathbb{p}_{\boldsymbol{X}}$, $\int_E \mathrm{d}\mu \leftrightarrow \mathbb{E}$. On redémontre ce fait dans le cas des variables aléatoires réelles discrètes et continues :

Dans le cas où $\boldsymbol{X}$ est une variable aléatoire *réelle* (et donc $f : \mathbb{R} \longrightarrow \mathbb{R}$ borélienne) (ou dans $\mathbb{C}$ ou $\mathbb{R}^d$) :

↳ Lorsque $\boldsymbol{X}$ est une V.A.R. à densité $\rho_{\boldsymbol{X}} : \mathbb{R} \longrightarrow \mathbb{R}_+$, on a
$f(\boldsymbol{X})$ intégrable $\iff$ $f \cdot \rho_{\boldsymbol{X}} \in \mathscr{L}^1(\mathbb{R})$ Lebesgue-intégrable, et alors, ou si $f \geqslant 0$,

$$\mathbb{E}\big(f(\boldsymbol{X})\big) = \int_\mathbb{R} f \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda = \int_\mathbb{R} f(x)\,\rho_{\boldsymbol{X}}(x)\,\mathrm{d}x$$

En particulier, $\boldsymbol{X}$ intégrable $\Leftrightarrow x \mapsto x\,\rho_{\boldsymbol{X}}(x) \in \mathscr{L}^1(\mathbb{R})$, et alors $\mathbb{E}(\boldsymbol{X}) = \displaystyle\int_\mathbb{R} x\,\rho_{\boldsymbol{X}}(x)\,\mathrm{d}x$.

Si $f = \mathbb{1}_B$ avec $B \in \mathfrak{B}(\mathbb{R})$, on a $f(\boldsymbol{X}) = \mathbb{1}_B \circ \boldsymbol{X} = \mathbb{1}_{\boldsymbol{X}^\leftarrow(B)}$ donc on a bien

$$\mathbb{E}\big(f(\boldsymbol{X})\big) \overset{\text{def.}}{=} \int_\Omega f(\boldsymbol{X})\,\mathrm{d}\mathbb{p} = \int_\Omega \mathbb{1}_{\boldsymbol{X}^\leftarrow(B)}\,\mathrm{d}\mathbb{p} \overset{\text{def.}}{=} \mathbb{p}[\boldsymbol{X} \in B] \underset{\text{Radon-Nikodym}}{\overset{\text{à densité}}{=}} \int_B \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda = \int_\mathbb{R} \mathbb{1}_B \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda$$

Donc par linéarité de cette formule, le théorème est vrai pour toute fonction étagée positive. Maintenant, soit $f \in \mathcal{M}(\mathbb{R}, \overline{\mathbb{R}_+})$ une fonction borélienne positive. Par le théorème d'approximation, soit $(f_n)_{n\in\mathbb{N}} \in (\text{Étag}_+)^\mathbb{N}_\nearrow$ suite croissante de fonctions étagées positives telle que $f_n \xrightarrow[n\to\infty]{\nearrow} f$. Alors par convergence monotone $\times 2$,

$$\mathbb{E}(f(\boldsymbol{X})) \xleftarrow[n\to\infty]{\nearrow} \mathbb{E}\big(f_n(\boldsymbol{X})\big) \underset{\text{Étag}_+}{\overset{\text{vrai pour}}{=}} \int_\mathbb{R} f_n \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda \xrightarrow[n\to\infty]{\nearrow} \int_\mathbb{R} f \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda$$

puisque $(f_n \cdot \rho_{\boldsymbol{X}})_{n\in\mathbb{N}}$ est une suite croissante de fonctions boréliennes positives telle que $f_n \cdot \rho_{\boldsymbol{X}} \xrightarrow[n\to\infty]{\nearrow} f \cdot \rho_{\boldsymbol{X}}$ car $\rho_{\boldsymbol{X}} \geqslant 0$ constante/n. Donc $\mathbb{E}(f(\boldsymbol{X})) = \int_\mathbb{R} f \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda$ pour $f \in \mathcal{M}(\mathbb{R}, \overline{\mathbb{R}_+})$.
Maintenant, soit $f \in \mathcal{M}(\mathbb{R}, \mathbb{R})$ borélienne quelconque. Par le point précédent, on a

$$f(\boldsymbol{X})\ \mathbb{p}\text{-intégrable} \iff \mathbb{E}(|f(\boldsymbol{X})|) = \int_\mathbb{R} |f| \cdot \rho_{\boldsymbol{X}}\,\mathrm{d}\lambda < +\infty \iff f \cdot \rho_{\boldsymbol{X}}\ \lambda\text{-intégrable}$$

et alors en décomposant en partie positive/négative $f = f_+ - f_-$ comme en général, on obtient le résultat.

↪ Lorsque $\boldsymbol{X}$ est une v.a.r. discrète de loi $\mathbb{p}_{\boldsymbol{X}} = \sum_{k\in\mathbb{N}} p_k\,\delta_{x_k}$ déterminée par $\big(p_k = \mathbb{p}[\boldsymbol{X}=x_k]\big)_{k\in\mathbb{N}}$

$$f(\boldsymbol{X}) \text{ intégrable } \iff \big(|f(x_k)|\,p_k\big)_{k\in\mathbb{N}} \text{ sommable,} \qquad \text{et alors}$$

$$\mathbb{E}\big(f(\boldsymbol{X})\big) = \sum_{k\in\mathbb{N}} f(x_k)\,p_k = \sum_{k\in\mathbb{N}} f(x)\,\mathbb{p}\{\boldsymbol{X}=x_k\}$$

En particulier, $\boldsymbol{X}$ intégrable $\Leftrightarrow$ $\big(|x_k|\,p_k\big)_{k\in\mathbb{N}}$ sommable, et alors $\mathbb{E}(\boldsymbol{X}) = \sum_{k\in\mathbb{N}} x_k\,p_k$.

Si $f = \mathbb{1}_B$ avec $B\in\mathfrak{B}(\mathbb{R})$, on a comme précédemment

$$\mathbb{E}\big(f(\boldsymbol{X})\big) \overset{\text{def.}}{=} \int_\Omega \mathbb{1}_{\boldsymbol{X}^{\leftarrow}(B)}\,\mathrm{d}\mathbb{p} \overset{\text{def.}}{=} \mathbb{p}[\boldsymbol{X}\in B] \underset{\text{discrète}}{\overset{\text{loi}}{=}} \sum_{k\in\mathbb{N}} p_k\,\delta_{x_k}(B) \overset{\text{def.}}{=} \sum_{k\in\mathbb{N}} \mathbb{1}_B(x_k)\,p_k = \int_E \mathbb{1}_B(x)\,\mathbb{p}[\boldsymbol{X}=x]\,\mathrm{d}\mu_{\mathsf{c}}(x)$$

si on pose $E = \boldsymbol{X}(\Omega)$ et μ_{c} la mesure de comptage sur $\mathbb{N}$. Alors, comme précédemment, par linéarité de cette formule puis approximation par des fonctions étagées et convergence monotone, $\forall f\in\mathcal{M}(\mathbb{R},\overline{\mathbb{R}_+})$,

$$\mathbb{E}\big(f(\boldsymbol{X})\big) = \int_E f(x)\,\mathbb{p}[\boldsymbol{X}=x]\,\mathrm{d}\mu_{\mathsf{c}}(x) = \sum_{k\in\mathbb{N}} f(x_k)\,p_k$$

Maintenant, soit $f\in\mathcal{M}(\mathbb{R},\mathbb{R})$ borélienne quelconque. Par le point précédent, on a

$$f(\boldsymbol{X}) \text{ intégrable } \iff \mathbb{E}(|f(\boldsymbol{X})|) = \sum_{k\in\mathbb{N}} |f(x_k)|\,p_k < +\infty \iff \big(|f(x_k)|\,p_k\big)_{k\in\mathbb{N}} \text{ sommable}$$

et alors en décomposant en partie positive/négative $f = f_+ - f_-$ comme précédemment, on obtient le résultat.

Méthode de la fonction muette : (réciproque du théorème de transfert)

Soit $\boldsymbol{X}$ une variable aléatoire à valeurs dans $(E,\mathcal{E})$ et μ une mesure sur $(E,\mathcal{E})$. Alors si la formule de transfert est vérifiée avec l'intégrale contre μ pour une classe suffisemment générale de fonctions f, par exemple les fonctions boréliennes positives $\mathcal{M}(E,\mathbb{R}_+) := \mathcal{M}((E,\mathcal{E}),\mathbb{R}_+)$, alors on la loi de $\boldsymbol{X}$ est μ :

$$\left(\forall f\in\mathcal{M}(E,\mathbb{R}_+),\quad \mathbb{E}\big(f(\boldsymbol{X})\big) = \int_E f\,\mathrm{d}\mu\right) \implies \mathbb{p}_{\boldsymbol{X}} = \mu$$

Dans le cas particulier où $\boldsymbol{X}$ est une v.a.r. à densité, si $\rho\colon \mathbb{R}\longrightarrow\mathbb{R}_+$ est une fonction borélienne,

$$\left(\forall f\in\mathcal{M}(E,\mathbb{R}_+),\quad \mathbb{E}\big(f(\boldsymbol{X})\big) = \int_{\mathbb{R}} f\cdot\rho\,\mathrm{d}\lambda\right) \implies \rho_{\boldsymbol{X}} \overset{\lambda\text{-pp}}{=} \rho$$

Dans le cas particulier où $\boldsymbol{X}$ est une v.a.r. discrète dans $\{x_k\}_{k\in\mathbb{N}}$, si $\{p_k\}_{k\in\mathbb{N}}\in\mathbb{R}_+^{\mathbb{N}}$,

$$\left(\forall f\in\mathcal{M}(E,\mathbb{R}_+),\quad \mathbb{E}\big(f(\boldsymbol{X})\big) = \sum_{k\in\mathbb{N}} f(x_k)\,p_k\right) \implies \mathbb{p}_{\boldsymbol{X}} = \sum_{k\in\mathbb{N}} p_k\,\delta_{x_k}$$

Lorsque l'on test contre les fonctions mesurables positives, il suffit simplement de prendre, $\forall X\in\mathcal{E}$, $f = \mathbb{1}_X$, et

$$\mathbb{p}_{\boldsymbol{X}}(X) = \mathbb{p}[\boldsymbol{X}\in X] = \int_\Omega \mathbb{1}_{\boldsymbol{X}^{\leftarrow}(X)}\,\mathrm{d}\mathbb{p} = \int_\Omega \mathbb{1}_X\circ\boldsymbol{X}\,\mathrm{d}\mathbb{p} = \mathbb{E}\big(\mathbb{1}_X(\boldsymbol{X})\big) = \int_E \mathbb{1}_X\,\mathrm{d}\mu = \mu(X)$$

donc $\mathbb{p}_{\boldsymbol{X}} = \mu$. De même pour les cas particuliers des v.a.r.

⇒ Cette méthode est utile pour construire une intuition de l'influence des paramètres de loi, ainsi que pour identifier la loi/densité d'une transformée $f(\boldsymbol{X})$ d'une variable aléatoire en utilisant tout la puissance des changements de variables, ou encore pour déterminer la loi/densité d'une variable aléatoire à partir de $\mathbb{p}$.