Fonction de répartition, fonction génératrice des moments, échantillonnage

1. Fonction de répartition d'une V.A.R.

Soit $\boldsymbol{X}$ une variable aléatoire réelle. On définit sa **fonction de répartition** par

$$\begin{aligned} \mathbb{F}_{\boldsymbol{X}} : \mathbb{R} &\longrightarrow [0,1] \\ t &\longmapsto \mathbb{p}[\boldsymbol{X} \leqslant t] = \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, t]\big) \end{aligned}$$

Propriétés :

- $\mathbb{F}_{\boldsymbol{X}}$ croissante : $\forall a \leqslant b \in \mathbb{R}$, $]-\infty, a] \subset]-\infty, b]$ donc $\mathbb{F}_{\boldsymbol{X}}(a) = \mathbb{p}[\boldsymbol{X} \leqslant a] \leqslant \mathbb{p}[\boldsymbol{X} \leqslant b] = \mathbb{F}_{\boldsymbol{X}}(b)$

- $\mathbb{F}_{\boldsymbol{X}}$ continue à droite : $\forall t \in \mathbb{R}, \quad \mathbb{F}_{\boldsymbol{X}}(t) = \lim_{t^+} \mathbb{F}_{\boldsymbol{X}} = \mathbb{F}_{\boldsymbol{X}}(t^+)$

  > Par caractérisation séquentielle de la limite à droite pour une fonction croissante, il suffit de vérifier que
  > $$\mathbb{F}_{\boldsymbol{X}}\big(t + 1/n\big) \xrightarrow[n \to \infty]{\searrow} \mathbb{F}_{\boldsymbol{X}}(t)$$
  > ce qui est vrai puisque $\big(]-\infty, t + 1/n]\big)_{n \in \mathbb{N}}$ est une suite décroissante dénombrable de boréliens, donc par limite décroissante,
  > $$\mathbb{F}_{\boldsymbol{X}}(t) = \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, t]\big) = \mathbb{p}_{\boldsymbol{X}}\Big(\bigcap_{n \in \mathbb{N}}^{\infty}]-\infty, t + 1/n]\Big) = \lim_{n \to \infty}^{\searrow} \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, t + 1/n]\big) = \lim_{n \to \infty}^{\searrow} \mathbb{F}_{\boldsymbol{X}}\big(t + 1/n\big)$$

- $\lim\limits_{-\infty} \mathbb{F}_{\boldsymbol{X}} = 0$ et $\lim\limits_{+\infty} \mathbb{F}_{\boldsymbol{X}} = 1$

  > Par caractérisation séquentielle de la limite pour une fonction croissante, il suffit de vérifier que
  > $$\mathbb{F}_{\boldsymbol{X}}(-n) \xrightarrow[n \to \infty]{\searrow} 0 \quad \text{et} \quad \mathbb{F}_{\boldsymbol{X}}(+n) \xrightarrow[n \to \infty]{\nearrow} 1$$
  > ce qui est vrai puisque $\big(]-\infty, +n]\big)_{n \in \mathbb{N}}$ est une suite croissante dénombrable de boréliens, donc par limite croissante,
  > $$1 = \mathbb{p}_{\boldsymbol{X}}(\mathbb{R}) = \mathbb{p}_{\boldsymbol{X}}\Big(\bigcup_{n \in \mathbb{N}}^{\infty}]-\infty, n]\Big) = \lim_{n \to \infty}^{\nearrow} \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, n]\big) = \lim_{n \to \infty}^{\nearrow} \mathbb{F}_{\boldsymbol{X}}(+n)$$
  > De même pour $-\infty$.

$\rightarrow$ On appelle **fonction de répartition** toute fonction $F : \mathbb{R} \longrightarrow [0,1]$ vérifiant les 4 propriétés précédentes (croissance, continuité à droite, $\lim_{-\infty} = 0$ et $\lim_{+\infty} = 1$).

- Les sauts de $\mathbb{F}_{\boldsymbol{X}}$ correspondent aux atomes de la loi de $\boldsymbol{X}$: $\quad \forall t \in \mathbb{R}$,
  $$\mathbb{F}_{\boldsymbol{X}}(t^-) = \lim_{t^-} \mathbb{F}_{\boldsymbol{X}} = \mathbb{F}_{\boldsymbol{X}}(t) - \mathbb{p}[\boldsymbol{X} = t]$$

  > Par caractérisation séquentielle de la limite à gauche pour une fonction croissante, il suffit de vérifier que
  > $$\mathbb{F}_{\boldsymbol{X}}\big(t - 1/n\big) \xrightarrow[n \to \infty]{\nearrow} \mathbb{F}_{\boldsymbol{X}}(t) - \mathbb{p}[\boldsymbol{X} = t]$$
  > ce qui est vrai puisque $\big(]-\infty, t - 1/n]\big)_{n \in \mathbb{N}}$ est une suite croissante dénombrable, donc par limite croissante
  > $$\begin{aligned} \mathbb{p}_{\boldsymbol{X}}\Big(\bigcup_{n \in \mathbb{N}}^{\infty}]-\infty, t - 1/n]\Big) &= \lim_{n \to \infty}^{\nearrow} \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, t - 1/n]\big) = \lim_{n \to \infty}^{\nearrow} \mathbb{F}_{\boldsymbol{X}}\big(t - 1/n\big) \\ &= \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, t[\big) = \mathbb{F}_{\boldsymbol{X}}(t) - \mathbb{p}\{\boldsymbol{X} = t\} \end{aligned}$$

- Donc $\mathbb{F}_{\boldsymbol{X}}$ continue à gauche (donc continue) en t si et seulement si t n'est pas un atome de $\mathbb{p}_{\boldsymbol{X}}$.

- $\forall a \leqslant b \in \mathbb{R}$,

$$\begin{aligned} \mathbb{p}\big[a < \boldsymbol{X} \leqslant b\big] &= \mathbb{F}_{\boldsymbol{X}}(b) - \mathbb{F}_{\boldsymbol{X}}(a) \\ \mathbb{p}\big[a < \boldsymbol{X} < b\big] &= \mathbb{F}_{\boldsymbol{X}}(b^-) - \mathbb{F}_{\boldsymbol{X}}(a) \quad \text{si } a < b \\ \mathbb{p}\big[a \leqslant \boldsymbol{X} \leqslant b\big] &= \mathbb{F}_{\boldsymbol{X}}(b) - \mathbb{F}_{\boldsymbol{X}}(a^-) \end{aligned}$$

- On sait que le nombre de points de discontinuité d'une fonction $\mathbb{R} \longrightarrow \mathbb{R}$ monotone, en particulier une fonction de répartition, est dénombrable [argument classique], donc

$$\big\{x \in \mathbb{R} : \mathbb{F}_{\boldsymbol{X}}(x^-) \neq \mathbb{F}_{\boldsymbol{X}}(x)\big\} = \big\{x \in \mathbb{R} : \mathbb{p}[\boldsymbol{X} = x] \neq 0\big\} = \big\{\text{atomes de } \boldsymbol{X}\big\} \quad \textit{dénombrable}$$

  Si cet ensemble est presque-sûr, $\boldsymbol{X}$ est une variable aléatoire réelle discrète.

Caractérisation de la loi par la fonction de répartition :

Soient $\boldsymbol{X}$ et $\boldsymbol{Y}$ deux V.A.R. Elles suivent la même loi si leurs fonctions de répartition coïncident :

$$\mathbb{F}_{\boldsymbol{X}} = \mathbb{F}_{\boldsymbol{Y}} \quad \Longrightarrow \quad \mathbb{p}_{\boldsymbol{X}} = \mathbb{p}_{\boldsymbol{Y}}$$

> C'est simplement le lemme d'unicité des probabilités (conséquence du lemme des classes monotones). En effet, on sait que la tribu de l'espace mesurable de $\mathbb{p}_{\boldsymbol{X}}$ et $\mathbb{p}_{\boldsymbol{Y}}$, l'espace des boréliens, est engendré par les $]-\infty, a]$:
> $$\mathfrak{B}(\mathbb{R}) = \sigma(\{]-\infty, a] : a \in \mathbb{R}\})$$
> et que $\{]-\infty, a]\}_{a \in \mathbb{R}}$ est stable par intersection finie : $\forall a, b \in \mathbb{R}$, $]-\infty, a] \cap]-\infty, b] = \big]-\infty, \min\{a, b\}\big]$. Or
> $$\forall a \in \mathbb{R}, \quad \mathbb{p}_{\boldsymbol{X}}\big(]-\infty, a]\big) = \mathbb{F}_{\boldsymbol{X}}(a) \overset{\text{hyp}}{=} \mathbb{F}_{\boldsymbol{Y}} = \mathbb{p}_{\boldsymbol{Y}}\big(]-\infty, a]\big)$$
> donc $\mathbb{p}_{\boldsymbol{X}}$ et $\mathbb{p}_{\boldsymbol{Y}}$ coïncident sur $\{]-\infty, a]\}_{a \in \mathbb{R}}$ donc par le lemme d'unicité des probabilités, $\mathbb{p}_{\boldsymbol{X}} = \mathbb{p}_{\boldsymbol{Y}}$.

Théorème de la fonction de répartition réciproque :

La réciproque généralisée d'une fonction de répartition F, ou **fonction quantile**, donnée par

$$\begin{aligned} F^{-1} :]0,1[&\longrightarrow \mathbb{R} \\ \alpha &\longmapsto \inf\{x \in \mathbb{R} : F(x) \geqslant \alpha\} \end{aligned}$$

est une variable aléatoire réelle sur l'espace $\big(]0,1[, \mathfrak{B}(]0,1[), \lambda_{]0,1[}\big)$ de Lebesgue, de fonction de répartition

$$\mathbb{F}_{F^{-1}} = F$$

Dit autrement, dans $(\Omega, \mathcal{A}, \mathbb{p})$, si $\boldsymbol{X}$ est une variable alétoire réelle et si $\boldsymbol{U} \rightsquigarrow \mathrm{Unif}(]0,1[)$,

$$\mathbb{F}_{\boldsymbol{X}}^{-1}(\boldsymbol{U}) \leftrightsquigarrow \boldsymbol{X}$$

puisque $\mathbb{p}_{\boldsymbol{U}} = \lambda_{]0,1[}$. C'est le *méthode de la transformée inverse*.

> Preuve : `https://fr.wikipedia.org/wiki/Fonction_de_répartition#Théorème_de_la_réciproque`

Convergence en loi et fonction de répartition :

Critère général de convergence en loi : convergence simple des fonctions de répartition en tout point où la fonction de répartition limite est continue (points de continuité de la limite)

> `https://fr.wikipedia.org/wiki/`
> `Fonction_de_répartition#Convergence_en_loi_et_fonction_de_répartition`

C'est grâce à cette propriété que l'on vérifie numériquement la convergence en loi.

Fonction de répartition de lois à densité :

Si $\boldsymbol{X}$ est une variable aléatoire diffuse, c'est à dire que $\forall x \in \mathbb{R},\ \mathbb{p}[\boldsymbol{X} = x] = 0$, alors $\mathbb{F}_{\boldsymbol{X}}$ est continue.

⚠ Il existe des variables aléatoires diffuses mais ne possédant pas pour autant une densité de probabilité. Par exemple, la variable aléatoire ayant pour fonction de répartition l'escalier de Cantor.

Lorsque $\boldsymbol{X}$ est de densité $\rho_{\boldsymbol{X}}$,

$$\forall t \in \mathbb{R}, \quad \mathbb{F}_{\boldsymbol{X}}(t) = \mathbb{F}_{\boldsymbol{X}}(t^-) = \mathbb{p}\big[\boldsymbol{X} \leqslant t\big] = \int_{-\infty}^{t} \rho_{\boldsymbol{X}}\, \mathrm{d}\lambda$$

Toutefois, ce n'est pas, en toute généralité, une primitive au sens strict du terme : on peut seulement affirmer : qu'une fonction de répartition est dérivable presque partout (pour la mesure de Lebesgue); et que si $\boldsymbol{X}$ est à densité, alors la dérivée de $\mathbb{F}_{\boldsymbol{X}}$ est presque partout égale à sa densité. Mais il y a beaucoup de « contre-exemples » : la fonction de répartition de la loi uniforme sur un intervalle, ou encore celle de la loi exponentielle, ne sont pas dérivables sur tout $\mathbb{R}$, et ne sont donc pas, au sens strict, des primitives de densités de probabilités.

1.1. Fonction génératrice des moments

Soit $\boldsymbol{X}$ une variable aléatoire réelle. On pose la **fonction génératrice des moments** par

$$\begin{array}{rccl} \mathrm{M}_{\boldsymbol{X}} : & \mathbb{R} & \longrightarrow & \overline{\mathbb{R}_+} \\ & t & \longmapsto & \mathbb{E}(\mathrm{e}^{t\boldsymbol{X}}) \end{array}$$

Contrairement à la fonction caractéristique, elle n'est pas toujours finie sur tout $\mathbb{R}^d$. On dit que la fonction génératrice des moments existe si $\exists r > 0 : \forall t \in]-r, +r[,\ \mathrm{M}_{\boldsymbol{X}}(t) < +\infty$. On a toujours $\mathrm{M}_{\boldsymbol{X}}(0) = 1$.

Lorsque $\boldsymbol{X}$ possède une densité, il s'agit d'une *transformée de Laplace bilatère* :

$$\mathrm{M}_{\boldsymbol{X}}(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{tx}\, \rho_{\boldsymbol{X}}(x)\, \mathrm{d}x$$

Si la fonction génératrice des moments existe ($\mathrm{M}_{\boldsymbol{X}} < +\infty$ sur un voisinage de 0), alors

- Tous les moments de $\boldsymbol{X}$ sont définis : $\forall k \in \mathbb{N},\ \mathbb{E}(|\boldsymbol{X}|^k) < +\infty$
- $\mathrm{M}_{\boldsymbol{X}}$ est $\mathcal{C}^\infty$ sur son rayon de convergence et $\forall k \in \mathbb{N}, \quad \mathbb{E}(\boldsymbol{X}^k) = \partial^k \mathrm{M}_{\boldsymbol{X}}(0)$
- On a même $\displaystyle \mathrm{M}_{\boldsymbol{X}}(t) = \sum_{k=0}^{\infty} \mathbb{E}(\boldsymbol{X}^k)\, \frac{t^k}{k!}$ ($\mathrm{M}_{\boldsymbol{X}}$ est analytique en 0)

Caractérisation de la loi par la fonction génératrice des moments :

Attention, deux distributions peuvent avoir les mêmes moments ou cumulants (même tous finis) tout en étant distinctes. Par contre, l'existence d'une fonction génératrice (rayon de convergence non nul) suffit à assurer que la distribution est uniquement déterminée par ses moments ou cumulants :

Soient $\boldsymbol{X}$ et $\boldsymbol{Y}$ deux variables aléatoires réelles. Si leurs fonctions génératrices existent et coïncident :

$$\exists r > 0 : \forall t \in]-r, +r[, \quad \mathrm{M}_{\boldsymbol{X}}(t) = \mathrm{M}_{\boldsymbol{Y}}(t) < +\infty$$

alors $\boldsymbol{X}$ et $\boldsymbol{Y}$ suivent la même loi : $\mathbb{p}_{\boldsymbol{X}} = \mathbb{p}_{\boldsymbol{Y}}$.

1.2. Échantillons, moyenne et fonction de répartition empirique

Soit $\boldsymbol{X}$ une variable aléatoire réelle *intégrable*. Soit $(\boldsymbol{X}_k)_{k\in\mathbb{N}}$ un échantillon de $\boldsymbol{X}$, c'est à dire une famille indépendante identiquement distribuée $\boldsymbol{X}_i \leftrightsquigarrow \boldsymbol{X}$.

La **moyenne empirique** de l'échantillon $(\boldsymbol{X}_k)_{1\leqslant k\leqslant n}$ est la variable aléatoire réelle

$$\hat{\boldsymbol{X}}_n := \frac{1}{n}\sum_{k=0}^{n} \boldsymbol{X}_n = \frac{\boldsymbol{S}_n}{n} \quad \text{avec} \quad \boldsymbol{S}_n := \sum_{k=0}^{n} \boldsymbol{X}_n$$

C'est un estimateur non biaisé de l'espérence $\mathbb{E}(\boldsymbol{X})$:

$$\forall n\in\mathbb{N}, \quad \mathbb{E}\big(\hat{\boldsymbol{X}}_n\big) = \mathbb{E}(\boldsymbol{X}) \quad \text{et} \quad \mathbb{V}\mathrm{ar}\big(\hat{\boldsymbol{X}}_n\big) = \tfrac{1}{n}\,\mathbb{V}\mathrm{ar}(\boldsymbol{X})$$

La **fonction de répartition empirique** de l'échantillon $(\boldsymbol{X}_k)_{1\leqslant k\leqslant n}$ est la fonction-variable aléatoire

$$\forall t\in\mathbb{R}, \quad \hat{\boldsymbol{F}}_n(t) := \frac{1}{n}\sum_{k=0}^{n} \mathbb{1}_{[\boldsymbol{X}_i\leqslant t]} = \frac{\#\text{ d'élém. } \leqslant t \text{ dans l'ech.}}{n} \quad \text{à valeurs dans } \left\{\frac{k}{n} : k\in[\![0,n]\!]\right\}$$

c'est à dire la variable aléatoire discrète sur $(\Omega,\mathcal{A})$ définie par $\hat{\boldsymbol{F}}_n(t)(\omega) = \frac{1}{n}\sum_{k=0}^{n}\mathbb{1}_{\boldsymbol{X}_i(\omega)\leqslant t}$.

C'est, à $t\in\mathbb{R}$ fixé, un estimateur non biaisé de la fonction de répartition $\mathbb{F}(t) := \mathbb{F}_{\boldsymbol{X}}(t)$ en t :

$$\forall n\in\mathbb{N}, \quad \mathbb{E}\big(\hat{\boldsymbol{F}}_n(t)\big) = \mathbb{F}(t)$$

$\mathbb{E}\big(\hat{\boldsymbol{F}}_n(t)\big) = \mathbb{E}\Big(\frac{1}{n}\sum_{k=0}^{n}\mathbb{1}_{[\boldsymbol{X}_i\leqslant t]}\Big) = \frac{1}{n}\sum_{k=0}^{n}\mathbb{E}(\mathbb{1}_{[\boldsymbol{X}_i\leqslant t]}) = \frac{n}{n}\,\mathbb{E}(\mathbb{1}_{[\boldsymbol{X}\leqslant t]}) = \mathbb{p}[\boldsymbol{X}\leqslant t] = \mathbb{F}_{\boldsymbol{X}}(t)$

Note : $\forall k\in\mathbb{N}$, $\forall t\in\mathbb{R}$, $\mathbb{1}_{[\boldsymbol{X}_i\leqslant t]}$ suit la loi de Bernoulli de paramètre $p = \mathbb{p}[\boldsymbol{X}_i\leqslant t] = \mathbb{F}(t)$, donc $\forall n\in\mathbb{N}$, $n\,\hat{\boldsymbol{F}}_n(t)$ est une somme de v.a. de Bernoulli, donc suit une loi binomiale $n\,\hat{\boldsymbol{F}}_n(t) \rightsquigarrow \mathcal{B}in(n,p)$ d'espérance $n\,\mathbb{F}(t)$.

C'est, à $t\in\mathbb{R}$ fixé, un estimateur consistant de $\mathbb{F}(t)$ en t par la *loi forte des grands nombres* :

$$\hat{\boldsymbol{F}}_n(t) \xrightarrow[n\to\infty]{\mathrm{L}^1,\text{ p.s.}} \mathbb{F}(t) \ \text{(V.A. constante)}$$

car $\mathbb{E}\big(\hat{\boldsymbol{F}}_n(t)\big) = \mathbb{F}_{\boldsymbol{X}}(t)$, toujours sous condition d'intégrabilité $\mathbb{E}(\boldsymbol{X})<\infty$. C'est à dire qu'on a, Ω-presque-sûrement, convergence simple (sur $\mathbb{R}$) vers la fonction de répartition $\mathbb{F}_{\boldsymbol{X}}$.

Théorème de Glivenko-Cantelli : toujours sous condition d'intégrabilité $\mathbb{E}(\boldsymbol{X})<\infty$,

$$\big\|\hat{\boldsymbol{F}}_n - \mathbb{F}\big\|_\infty = \sup_{t\in\mathbb{R}}\big\|\hat{\boldsymbol{F}}_n(t) - \mathbb{F}(t)\big\| \xrightarrow[n\to\infty]{\text{p.s.}} 0$$

Note : $\big\|\hat{\boldsymbol{F}}_n - \mathbb{F}\big\|_\infty$ est la variable aléatoire réelle $(\Omega,\mathcal{A},\mathbb{p}) \to (\mathbb{R},\mathfrak{B}(\mathbb{R}))$ définie par

$$\omega \longmapsto \sup_{t\in\mathbb{R}}\big\|\hat{\boldsymbol{F}}_n(t)(\omega) - \mathbb{F}(t)\big\| = \sup_{t\in\mathbb{R}}\|{-\mathbb{F}(t)}\|$$