3.4.3. VA discrètes et continues, lois classiques

VA discrètes

VA discrète : $X : \Omega \to F \subset \mathbb{R}$ $\quad$ F fini ou dénombrable.

$F = \{x_i\}_{i \in I}$, $\quad I \subset \mathbb{N}$.

Ex : $F = \{1, \dots, N\}$ ou $F = \mathbb{N}$.

P_X est caractérisée par $p_i = P(X = x_i)$.

$$\boxed{P_X = \sum_{i \in I} p_i \, \delta_{x_i}}$$

δ_x : mesure de Dirac.

Condition : $\sum_i p_i = 1$

* Espérance $E(f(X)) = \int f(x)\, dP_X(x)$

$$= \sum_i f(x_i)\, p_i .$$

* Si $F \subset \mathbb{N}$, on utilise pour la fonction caractéristique
$x_i = n$

$$G_X(s) = E(s^X) = \sum_{n \in \mathbb{N}} p_n s^n$$

On avait défini $\phi_X(k) = E(e^{ikX}) = G_X(e^{ik})$.

$G_X(1) = 1$.

$G_X^{(n)}(1) = E(X(X-1)\cdots(X-n+1))$

$G_{X+Y}(s) = G_X(s)\, G_Y(s)$ si X, Y indépendants.

VA continues ou "à densité"

$$\exists \; f_X : \mathbb{R} \to \mathbb{R}_+ \text{ tq } F_X(x) = \int_{-\infty}^{x} f_X(y)\, dy$$

f_X → densité (de probabilité) de X.

De façon équivalente : $P_X = f_X \, \lambda$ ← mesure de Lebesgue.

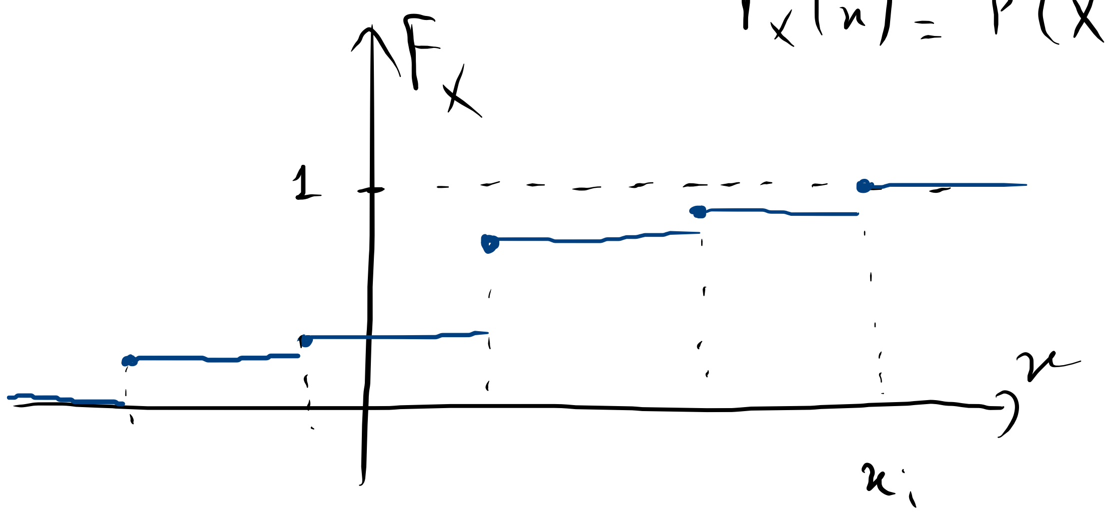
Normalisation : $1 = \lim_{x \to \infty} F_X(x) = \int_{-\infty}^{\infty} f_X(y)\, dy$.

$$\boxed{E(g(X)) = \int g(x)\, f_X(x)\, dx}$$

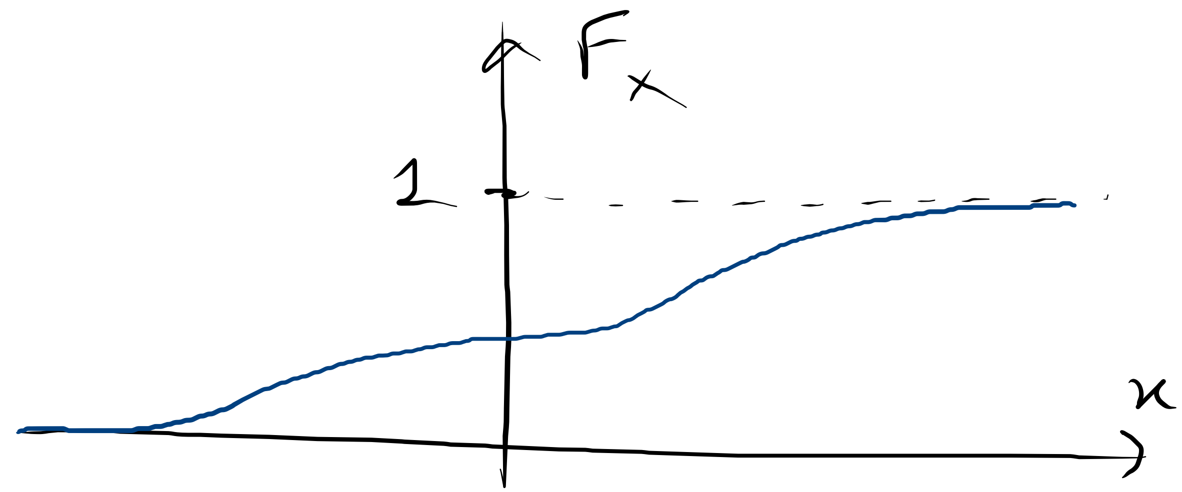
* Fonctions de répartition

Discret

$F_X(x) = P(X \le x)$

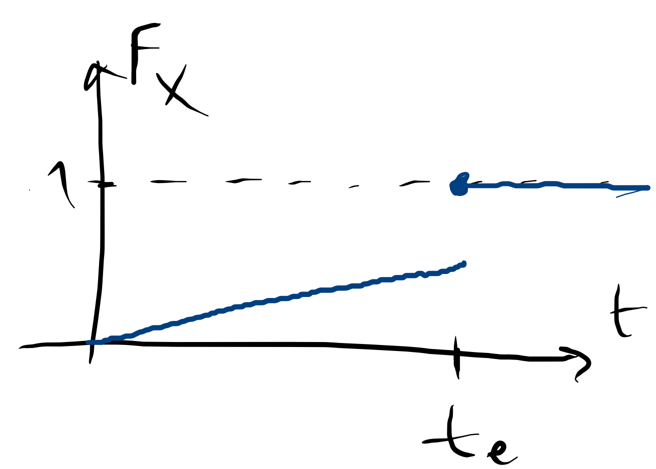


Continu



* V.A mixtes. Ex : temps T passé dans une salle d'examen de durée t_e.

$$P_T = f \lambda_{[0,t_e]} + p_{t_e} \delta_{t_e}$$



Normalisation : $\int_0^{t_e} f(t)\,dt + p_{t_e} = 1$.

Lois classiques

* VA uniforme $\quad F = \{1, \dots, n\} \quad p_i = \frac{1}{n}$.

En général $\quad \mathbb{P}(X = \omega) = \dfrac{1}{\operatorname{Card}(\Omega)}$.
$\omega \in \Omega$

* VA de Bernoulli $\quad F = \{0, 1\} \quad \mathbb{P}(X = 1) = p$

échec $\quad$ succès

$E(X^n) = p \quad \text{Var}(X) = p(1-p)$.
$n \geq 1$

* Loi binomiale $F = \{0, \dots, n\}$

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\left(= C_n^k\right)$$

Proba de k succès pour n essais.

$$E(X) = np \qquad Var(X) = np(1-p).$$

* Loi géométrique $F = \mathbb{N}_+$.

$$P(X=n) = (1-p)\, p^{n-1}$$

loi du moment du 1er échec.

$$E(X) = \frac{1}{1-p}, \quad Var(X) = \frac{p}{(1-p)^2}.$$

* Loi de Poisson de paramètre λ :

$F = \mathbb{N}$ $\quad P(X = n) = e^{-\lambda} \dfrac{\lambda^n}{n!}$

λL points répartis uniformément sur $[0, L]$:

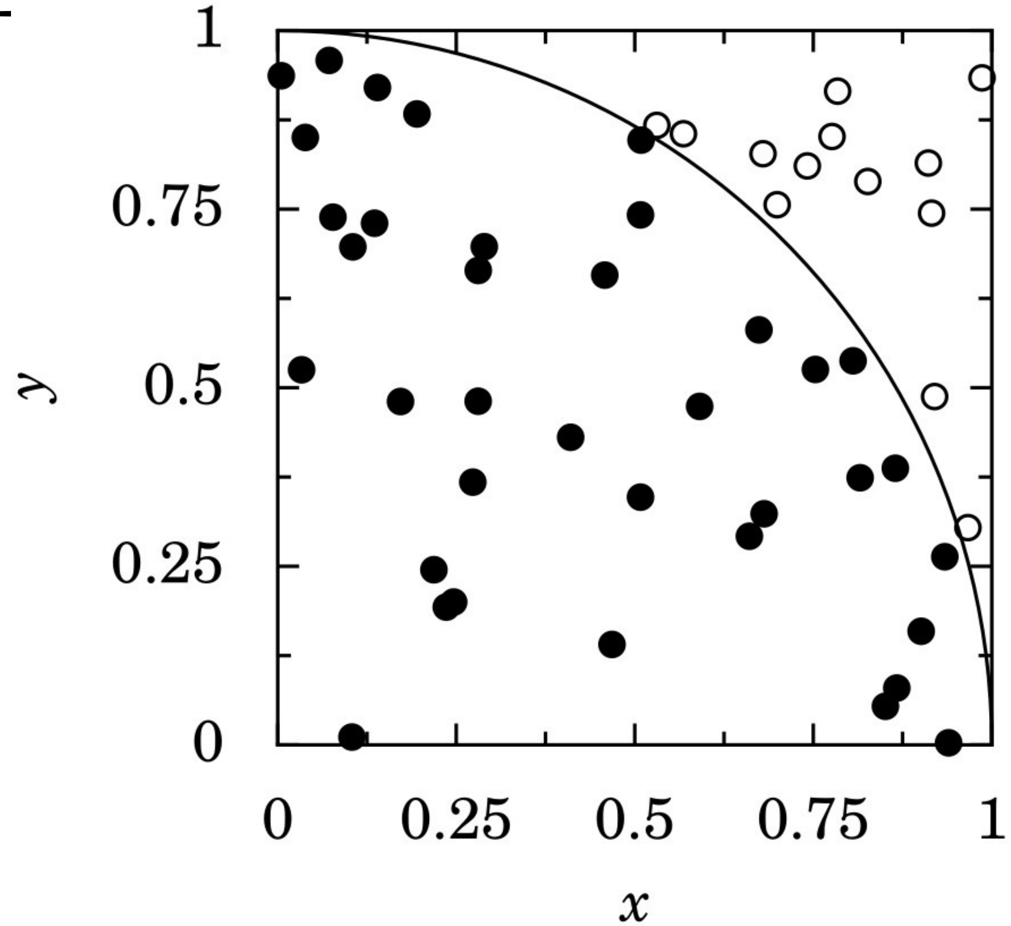
X : nombre de points sur $[0,1]$ quand $L \to \infty$.

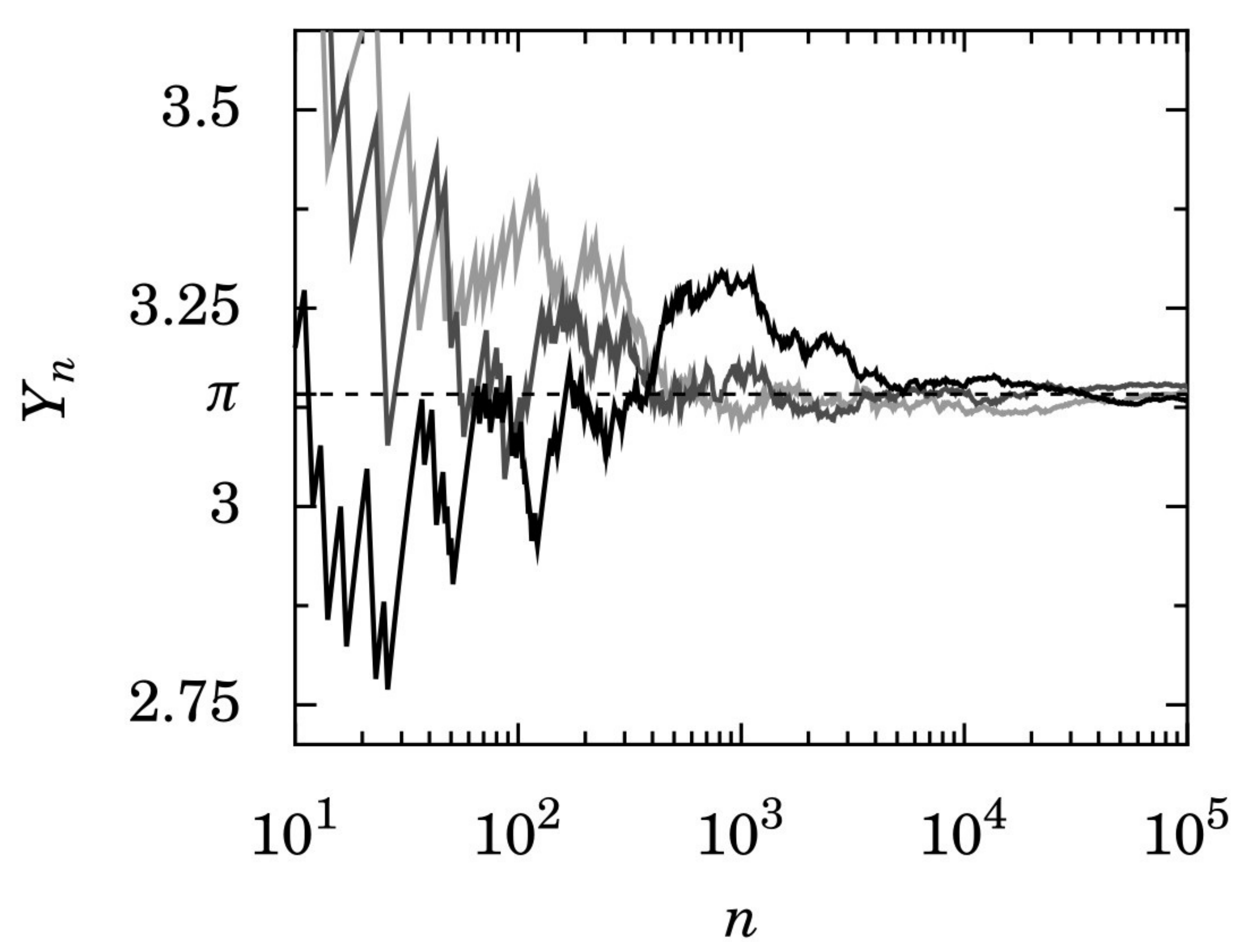
$E(X) = \lambda$, $\quad Var(X) = \lambda$.

* Loi uniforme continue : $F = [a, b] \subset \mathbb{R}$.

$$\mathbb{P}_X = \frac{1}{b-a} \lambda_{[a,b]}$$

Ex:





$$P\left(X^2+Y^2 \leq 1\right) = \frac{\pi}{4}.$$

* Loi exponentielle de paramètre λ :

$$P(X \geq x) = e^{-\lambda x} = 1 - F_X(x).$$

$$P(X \geq x_1 \mid X \geq x_0) = \frac{P(\overbrace{X \geq x_1 \cap X \geq x_0}^{X \geq x_1})}{P(X \geq x_0)}$$

$x_1 > x_0$

$$= \frac{e^{-\lambda x_1}}{e^{-\lambda x_0}} = e^{-\lambda(x_1 - x_0)}$$

$$= P(X \geq x_1 - x_0).$$

* Loi normale d'espérance μ et variance σ^2 :

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

Loi normale centrée réduite $\mathcal{N}$: $\mu = 0$, $\sigma^2 = 1$.

$$f_{\mathcal{N}}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

$$F_{\mathcal{N}}(x) = \int_{-\infty}^{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx = \Pi(x) = \frac{1}{2} + \mathrm{erf}\left(\frac{x}{\sqrt{2}}\right)$$

Fonction erreur : $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$

3.5. Suites de VA

Égalité de V.A

X et Y sont égales :

- presque sûrement si $P(X \neq Y) = 0$.
- en moyenne si $E(|X - Y|) = 0$
- en loi si $P_X = P_Y$.

Ex : lancer de dé $X : \Omega \to \{1, \dots, 6\}$.

$Y = 7 - X$. $P_X = P_Y$. $P(X \neq Y) = 1$.

Convergence de VA

$$X_n \xrightarrow[n\to\infty]{} X :$$

- presque sûrement si pour presque tout $\omega \in \Omega$

$$X_n(\omega) \xrightarrow[n\to\infty]{} X(\omega). \qquad X_n \xrightarrow{ps} X.$$

analogue : convergence simple.

- en probabilité si $\forall \epsilon > 0$ $P(|X_n - X| > \epsilon) \xrightarrow[n\to\infty]{} 0$.

$$X_n \xrightarrow{P} X.$$

- en moyenne si $E(|X_n - X|) \xrightarrow[n\to\infty]{} 0$

$$E(|X_n - X|) = \|X_n - X\|_1 \qquad X_n \xrightarrow{L^1} X.$$

- en loi : $\forall f(x)$ continue bornée

$$E(f(X_n)) \xrightarrow[n\to\infty]{} E(f(X)) . \qquad X_n \xrightarrow{\mathcal{L}} X .$$

$$\begin{tikzcd}
X_n \xrightarrow{p.s} X \arrow[rr, Rightarrow, "+\text{ domination}"] \arrow[dr, Rightarrow] & & X_n \xrightarrow{L^1} X \arrow[dl, Rightarrow] \\
& X_n \xrightarrow{P} X \arrow[d, Rightarrow] & \\
& X_n \xrightarrow{\mathcal{L}} X &
\end{tikzcd}$$

* $L^1 \Rightarrow \mathbb{P}$: $E(|X_n - X|) \xrightarrow[n\to\infty]{} 0$.

Inégalité de Markov : $\forall \epsilon > 0$

$$\mathbb{P}(|X_n - X| > \epsilon) \leq \frac{E(|X_n - X|)}{\epsilon}$$

* p.s. $\Rightarrow \mathbb{P}$: $A_{n,\epsilon} = (|X_n - X| > \epsilon)$

$Y_{n,\epsilon} = \chi_{A_{n,\epsilon}}$.

p.s. $\Rightarrow$ $Y_{n,\epsilon}(\omega) \xrightarrow[n\to\infty]{} 0 \quad \forall \omega$: $Y_{n,\epsilon} \xrightarrow{ps} 0$.

et $Y_{n,\epsilon}$ dominée donc $Y_{n,\epsilon} \xrightarrow{L^1} 0$.

$E(Y_{n,\epsilon}) \xrightarrow[n\to\infty]{} 0$. Or $E(Y_{n,\epsilon}) = \mathbb{P}(A_{n,\epsilon}) = \mathbb{P}(|X_n - X| > \epsilon)$

Convergence en loi

$\Leftrightarrow$ convergence simple $F_{X_n} \longrightarrow F_X$

en tout point où F_X est continue.

$\Leftrightarrow$ convergence des fonctions caractéristiques

(thm de Lévy).

Loi des grands nombres

$(X_n)_{n \in \mathbb{N}^*}$ VA indépendantes et identiquement distribuées.

$$\boxed{M_n = \frac{1}{n} \sum_{k=1}^{n} X_k .}$$

$E(X_k) = \mu$

$Var(X_k) = \sigma^2 .$

$E(M_n) = \mu .$

$$Var(M_n) = \frac{1}{n^2} Var\left(\sum_{k=1}^{n} X_k\right) = \frac{\sigma^2}{n} .$$

$$M_n \xrightarrow{?} \mu .$$

- Inégalité de Bienaymé-Tchebychev :

$$\mathbb{P}(|M_n - \mu| > \epsilon) \leq \frac{\mathrm{Var}(M_n)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \xrightarrow[n\to\infty]{} 0$$

donc $M_n \xrightarrow{\mathbb{P}} \mu$.

- $$E(|M_n - \mu|)^2 \leq E(|M_n - \mu|^2) = \mathrm{Var}(M_n) \xrightarrow[n\to\infty]{} 0$$

inégalité de Jensen

$$\Rightarrow E(|M_n - \mu|) \xrightarrow[n\to\infty]{} 0 \ : \ \boxed{M_n \xrightarrow{L^1} \mu}.$$

Loi faible des grands nombres.

Thm de la limite centrale

$$\begin{cases} E(M_n) = \mu \\ Var(M_n) = \dfrac{\sigma^2}{n} \end{cases} \qquad Y_n = \frac{M_n - \mu}{\sigma/\sqrt{n}}.$$

$$E(Y_n) = 0, \quad Var(Y_n) = 1.$$

Loi de Y_n quand $n \to \infty$.

$$\varphi_{Y_n}(k) = E\left(e^{ikY_n}\right) = E\left(e^{ik\frac{\sum_{j=0}^{n} X_j - n\mu}{\sqrt{n}\,\sigma}}\right)$$

$$= e^{-\frac{ik\sqrt{n}\mu}{\sigma}} \underbrace{E\left(e^{\frac{ik}{\sqrt{n}\sigma}\sum_{j=0}^{n} X_j}\right)}_{E\left(e^{\frac{ikX_1}{\sqrt{n}\sigma}}\right)^n}$$

$$\varphi_{Y_n}(k) = E\left(e^{\frac{ik(X_1-\mu)}{\sqrt{n}\,\sigma}}\right)^n$$

$$= \left[E\left(1 + \frac{ik(X_1-\mu)}{\sqrt{n}\,\sigma} - \frac{k^2(X_1-\mu)^2}{2n\sigma^2} + O\left(\frac{1}{n}\right)\right)\right]^n$$

$$= \left[1 - \frac{k^2}{2n} + o\left(\frac{1}{n}\right)\right]^n$$

$$\xrightarrow[n\to\infty]{} e^{-k^2/2}$$

fonction caractéristique de $\mathcal{N}$.

$$\boxed{Y_n = \frac{\sum_{k=1}^{n} X_k - n\mu}{\sqrt{n}\,\sigma} \xrightarrow{\mathcal{L}} \mathcal{N}}$$

Thm de la limite centrale.

Intervalle de confiance

$n = 2000$ personnes 48 % pour A

52 % pour B.

X_k : vote de la personne k : Bernoulli de proba de succès p. Estimation $\hat{p} = 48\%$.

On cherche ϵ tq $\mathbb{P}(|p - \hat{p}| \leq \epsilon) = \alpha = 0,95$.

$E(\hat{p}) = p$, $Var(\hat{p}) = \frac{\sigma^2}{n} = \frac{p(1-p)}{n} \leq \frac{1}{4n}$.

On sait que $\sqrt{4n}(\hat{p} - p) \overset{\mathcal{L}}{\simeq} \mathcal{N}$.

$$\sqrt{4n}\,(\hat{p} - p) \overset{\mathcal{L}}{\simeq} \mathcal{N}$$

$$\alpha = \mathbb{P}\left(|\hat{p} - p| \leq \epsilon\right) = \mathbb{P}\left(\sqrt{4n}\,|\hat{p} - p| \leq \sqrt{4n}\,\epsilon\right)$$

$$= \mathbb{P}\left(|\mathcal{N}| \leq \sqrt{4n}\,\epsilon\right)$$

$$= \Pi\left(\sqrt{4n}\,\epsilon\right) - \Pi\left(-\sqrt{4n}\,\epsilon\right)$$

$$= 2\Pi\left(\sqrt{4n}\,\epsilon\right) - 1$$

$$\epsilon = \frac{1}{\sqrt{4n}}\,\Pi^{-1}\left(\frac{1+\alpha}{2}\right)$$

Numériquement,

$$I_{2000;\,0{,}95} = [0{,}458\,;\ 0{,}502]$$

$$I_{2000;\,0{,}99} = [0{,}451\,;\ 0{,}509]$$