Donc le problème est : Étant donné π connue plus ou moins bien, comment échantillonn selon π.

cours 1 : simulations sous π

cours 2-3 : simulations d'une cm dont loi limite $= \pi$: MCMC

Rq : LGN, TCL aussi vrais pour des cm (dont les propriétés théoriques sont vérifiables).

La compréhension des méthodes MCMC repose sur la compréhension des algorithmes de simulati de base, en particulier le rejet et l'échantillonnage préférentiel que nous allons voir mainten

Remarque : Tous les algos sont basé sur le fait que l'on sait simuler selon $\mathcal{U}([0,1])$

On crée des simulateurs sous π à partir de celui sous $\mathcal{U}([0,1])$

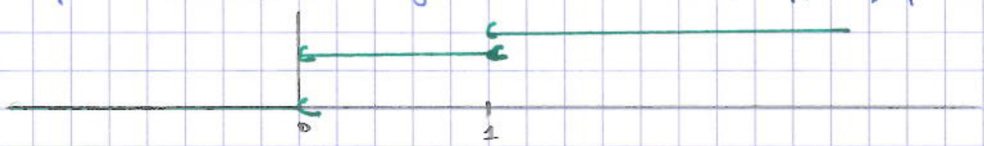
- Méthode d'inversion :

⚠ pour des v-a à valeur dans $\mathbb{R}$ seulement.

définition : (Fonction de répartition)
$$F : \mathbb{R} \to [0,1]$$
$$x \mapsto F(x) = \mathbb{P}(X \leq x) = \mathbb{P}_X(]-\infty, x])$$

Proposition : F est croissante ; $\lim\limits_{x \to -\infty} F(x) = 0$; $\lim\limits_{x \to +\infty} F(x) = 1$; F est continue à droit et a une limite à gauche.

⚠ F n'est pas nécessairement bijective : ex : $X = (1-p)\delta_0 + p\delta_1$



On définit son inverse généralisée

déf :
$$F^{-1} :]0,1[\to \mathbb{R}$$
$$u \mapsto \inf\{x \in \mathbb{R} \text{ tq } F(x) \geq u\}$$

Propriétés de F^{-1}
(a) F^{-1} est croissante
(b) F^{-1} est continue à gauche
(c) $F(x) \geq u \iff x \geq F^{-1}(u)$
(d) $F(F^{-1}(u)) \geq u$ avec égalité si F est continue en u.

Preuve : (a) comme F est $\nearrow$, F^{-1} l'est aussi

(c) Soit $x \in \mathbb{R}$, $p \in]0,1)$. Par déf de $F^{-1}(p)$, si $F(x) \geq p$ alors $x \geq F^{-1}(p)$ car c'est l'i
Réciproquement : si $x \geq F^{-1}(p)$ alors $\forall \varepsilon > 0$ $x + \varepsilon > F^{-1}(p)$
$\Rightarrow F(x+\varepsilon) \geq p$ par déf de F croissante
Comme F est C^0 à droite, en faisant $\varepsilon \to 0$ on obtient $F(x) \geq p$

(b) $F^{-1}(p-\varepsilon) \leq x$ $\forall x \in]0,p]$ $\iff$ $(p-\varepsilon) \leq F(x)$ $\forall \varepsilon \in]0,p]$
par ce qu'on vient de montrer
$\iff$ $p \leq F(x)$ $\iff$ $F^{-1}(p) \leq x$ par (c) à nouveau $\to$ continuité à gauc

(d) On rappelle les conventions $F(-\infty) = \lim_{-\infty} F(x)$ et $F(+\infty) = 1$

- Soit $x = F^{-1}(p)$ alors par (c) $F(F^{-1}(p)) \geq p$
- Cette inégalité est trivialement vraie si $F^{-1}(p) = +\infty$
- Si $F^{-1}(p) = -\infty$ alors par (c) $F(x) \geq p$ $\forall x \in \mathbb{R}$ donc par convention $p = -\infty$

$F(-\infty) \geq p$ donc $F(F^{-1}(p)) \geq p$ dans tous les cas.

Théorème : (a) Soit $\mathcal{U}([0,1])$ la loi uniforme sur $[0,1]$ et $U \sim \mathcal{U}([0,1])$

Alors $X = F^{-1}(U) \sim F$

(b) Si $X \sim F$ et F est C^0 alors $F(X) \sim \mathcal{U}([0,1])$

Preuve : (a) $\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(F(x) \geq U)$ par (c)

$= F(x)$ car U est uniforme

(b) $\mathbb{P}(F(X) < U) = 1 - \mathbb{P}(F(X) \geq u)$

$= 1 - \mathbb{P}(X \geq F^{-1}(u))$ par (c)

$= \mathbb{P}(X < F^{-1}(U)) = F(F^{-1}(u))$ car $X \sim F$ et F C^0 $\Rightarrow \leq \Leftrightarrow <$

$= u$ car F C^0.

Algorithme d'inversion : Pour simuler $X \sim F$: 1 – simuler $U \sim \mathcal{U}([0,1])$
2 – $X = F^{-1}(U)$

ex : (a) $X \sim \mathcal{E}xp(\lambda)$ $F(x) = 1 - e^{-\lambda x}$ (C^0 et inversible)
$F^{-1}(u) = -\frac{1}{\lambda}\log(1-u)$

Rq $U \sim \mathcal{U}[0,1] \Rightarrow 1-U$ aussi donc $F^{-1}(u) = x = -\frac{1}{\lambda}\log u$

(b) Pareto : $f(x) = \frac{a b^a}{x^{a+1}} \mathbb{1}_{\{x \geq b\}}$ avec $a, b > 0$

$F(x) = \int_{-\infty}^{x} f(t)\,dt = 1 - \left(\frac{b}{x}\right)^a$

$F^{-1}(u) = \frac{b}{(1-u)^{1/a}}$ ou $\frac{b}{u^{1/a}}$

Comment calculer F^{-1} :
- explicite (cf ex)
- Résolution numérique de $F(x) = u$
  → bissection : si $x \in [a,b]$, on compare $F(\frac{a+b}{2})$ à U et on cherche dans l'intervalle qu'il faut de ½ longueur et on itère
- Newton – Raphson (et dérivés)
  → ⚠ conditions d'arrêt.
- Approximation de F^{-1}

ex : pour $\mathcal{N}(0,1)$: $g(u) = (-2\ln u)^{1/2} + \frac{A[(-2\ln(u))^{1/2}]}{B[(-2\ln u)^{1/2}]}$, A, B 2 polynôme de deg pour $10^{-\ldots} \le u \le \frac{1}{2}$ si $\frac{1}{2} \le u \le 1-10$ $\to -g(1-u)$ sinon rien ...

I. Méthode d'acceptation - rejet : Valable pour toute v.a $\in \mathbb{R}^d$

→ Pour comprendre : 2 théorèmes

Théorème 1 : (a) Soient X de densité f sur $\mathbb{R}^d$ et $U \sim \mathcal{U}([0,1])$ indépendante de X et $c > 0$

Alors le couple $(X, cUf(X)) \sim \mathcal{U}(\mathcal{C})$ où $\mathcal{C} = \{(x,y) \text{ tq } 0 \le y \le cf(x)\}$

(b) Réciproquement, si $(X,Y) \sim \mathcal{U}(\mathcal{C})$ alors X a pour densité f.

Preuve : Comme $X \perp\!\!\!\perp U$ la loi du couple est $f(x)\, \mathbb{1}_{[0,1]}(u)\, du\, dx$

a) $$\mathbb{E}[h(X, cUf(X))] = \iint h(x, cuf(x))\, f(x)\, \mathbb{1}_{[0,1]}(u)\, du\, dx$$

$$= \iint h(x,y)\, f(x)\, \mathbb{1}_{\mathcal{C}}(x,y)\, \frac{dy}{cf(y)}\, dx \qquad \text{CHAR } (x,u) \to (x,y) \quad \mathcal{C} = \{0 \le y \le cf(x)\}$$

$$= \iint h(x,y)\, \frac{1}{c}\, \mathbb{1}_{\mathcal{C}}(x,y)\, dx\, dy = \mathbb{E}_{\mathbb{P}}[h(X,Y)]$$

où $\mathbb{P} = \mathcal{U}(\mathcal{C})$

Rq : $\int_{\mathcal{C}} dx\, dy = \int \int_0^{cf(x)} dy\, dx = c \int f(x)\, dx = c$ → bonne renormalisation

(b) $$f_X(x) = \int_y f_{(X,Y)}(x,y)\, dy = \frac{1}{c} \int_0^{cf(x)} dy = f(x)$$

Algorithme pour simuler une loi uniforme sur $\mathcal{C}$: càd sous la courbe $g(x)$

1. tirer $X \sim \dfrac{g(x)}{\int g(x)\, dx}$
2. tirer $U \sim \mathcal{U}([0,1])$ → poser $Y = g(X) \times U$
3. Retourner (X, Y)

Théorème 2 : Soit $(X_k)_{k \ge 0}$ une suite de v.a iid à valeur dans $\mathbb{R}^d$ de même loi que X. Soit A tq $\mathbb{P}(X \in A) = p > 0$. Soit $Y = X_{\tau_A}$ où $\tau_A = \inf\{k \ge 1 \text{ tq } X_k \in A\}$ (1^{er} instant dans A). Soit $B \subset \mathcal{X}$

Alors $$\mathbb{P}(Y \in B) = \frac{\mathbb{P}(X \in A \cap B)}{\mathbb{P}(A)} = \mathbb{P}(X \in B \mid X \in A)$$

loi conditionnelle sachant A.

Corollaire : Si $X \sim \mathcal{U}(\mathcal{C})$ où $\mathcal{C} \subset \mathbb{R}^d$. Soit $B \subset \mathcal{C}$. Alors

$$Y = X_{\tau_B} \sim \mathcal{U}[B]$$

Preuve : $\mathbb{E}[h(Y)] = \sum_{n\geq 1} \mathbb{E}[h(X_n) \mathbb{1}_{Z=n}]$

$= \sum_{n\geq 1} \mathbb{E}[h(X_n) \mathbb{1}_{\{X_1 \notin A\}} \cdots \mathbb{1}_{\{X_{n-1} \notin A\}} \mathbb{1}_{\{X_n \in A\}}]$

$\overset{iid}{=} \sum_{n\geq 1} \mathbb{P}(X_1 \notin A) \cdots \mathbb{P}(X_{n-1} \notin A)\, \mathbb{E}[h(X_n) \mathbb{1}_{\{X_n \in A\}}]$

$\overset{\hat{m}\text{ loi}}{=} \sum_{n\geq 1} (\mathbb{P}(X \notin A))^{n-1}\, \mathbb{E}[h(X) \mathbb{1}_{\{X \in A\}}]$

$= \left(\sum_{n\geq 1} (1-p)^{n-1}\right) \mathbb{E}[h(X) \mathbb{1}_{\{X\in A\}}]$

Si $h(x) = \mathbb{1}_{x\in B}$ on obtient : $\mathbb{P}(Y\in B) = \frac{1}{p}\mathbb{P}(X\in A\cap B) = \frac{\mathbb{P}(X\in A\cap B)}{\mathbb{P}(X\in A)}$

Algorithme du rejet pour simuler selon f sur $\mathbb{R}^d$:

Soit g une densité tq $\exists c>0$ tq $\forall x \in \mathbb{R}^d$ $f(x) \leq c\,g(x)$

Par thm 1 : tirer une uniforme selon $\mathcal{C}$ où $\mathcal{C} = \{(x,y),\ 0\leq y\leq g(x)\}$

Par thm 2 : Tant qu'on n'est pas dans $B = \{(x,y),\ 0\leq y\leq f(x)\}$ on continue. Quand on est dans B, par le corollaire et (b) du thm 1 on sait que l'abscisse $(X) \sim f$.

Pseudo code :
- $X \sim g(x)$
- $U \sim \mathcal{U}([0,1])$ et $Y = c\,g(x)\,U$
- Si $Y \leq f(x)$ → sortir X
- sinon reprendre.

⚠ Il faut savoir simuler facilement selon g et le choix de g est crucial !

En effet : Temps d'attente : Quelle est la loi de Z_A ?

- $\mathbb{P}(Z_A = n) = \mathbb{P}(X_1\notin A) \cdots \mathbb{P}(X_{n-1}\notin A)\,\mathbb{P}(X_n\in A)$ par $\perp\!\!\!\perp$ce

$= (1-p)^{n-1} p$ par iid $\Rightarrow Z_A \sim \text{Geom}(p)$: $\mathbb{E} = p^{-1}$, $\text{Var} = \frac{1-p}{p^2}$

où $p = \mathbb{P}(X\in A)$

- Dans le rejet :

on reformule Algo : Tant que $U > \frac{f(x)}{c\,g(x)}$ tirer $X\sim g$, $U\sim \mathcal{U}[0,1]$ et

et donc $p = \mathbb{P}\left(U \leq \frac{f(x)}{c\,g(x)}\right) = \int \mathbb{1}_{\{u\,c\,g(x)\leq f(x)\}} \mathbb{1}_{[0,1]}(u)\, g(x)\, dx\, d$

$= \frac{1}{c}\int \frac{f(x)}{g(x)} g(x)\, dx = \frac{1}{c}$

Donc plus c est grand, plus p est petit donc le nombre de tirages moyen avant succès $\mathbb{E}[Z_A] = \frac{1}{p} = c$ est grand → algorithme moins efficace

- Choix de (c,g) : • $f>0 \Rightarrow g>0$

  • $c = \sup_{x \text{ tq } f(x)>0} \frac{f(x)}{g(x)}$

  • $g \in$ famille paramétrique et on optimise le paramètre.

- Evaluation de f peut être coûteuse : majorer et minorer f par des fonctions facilement évaluable pour faire les comparaisons.

Ex 1: Simuler $\mathcal{N}(0,1)$ à partir d'une Laplace.

- Soit $g(x) = \frac{1}{2} e^{-|x|}$ sur $\mathbb{R}$

Simuler selon g : $\left.\begin{array}{l} X \sim \mathcal{E}xp(1) \\ \varepsilon \sim \mathcal{B}(1/2) \quad \perp\!\!\!\perp X \end{array}\right\} \varepsilon X \sim g$

$\forall h \in C_b^0 \quad \mathbb{E}[h(\varepsilon X)] = \frac{1}{2}\mathbb{E}[h(X)] + \frac{1}{2}\mathbb{E}[h(-X)]$ par $\perp\!\!\!\perp$ ce

$$= \frac{1}{2}\int_0^\infty h(x)e^{-x}\,dx + \frac{1}{2}\int_0^\infty h(-x)e^{-x}\,dx$$

$$= \frac{1}{2}\int_0^\infty h(x)e^{-|x|}\,dx + \frac{1}{2}\int_{-\infty}^0 h(y)e^{-|y|}\,dy$$

$$= \frac{1}{2}\int_{\mathbb{R}} h(y)e^{-|y|}\,dy$$

On cherche c tq $f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right) \leq \frac{c\,e^{-|x|}}{2}$

or $e^{-\frac{1}{2}x^2+|x|} = e^{-\frac{1}{2}(x^2-2|x|)} = e^{-1/2(|x|-1)^2+1/2}$

$$c = \sup_{\mathbb{R}} \frac{2}{\sqrt{2\pi}} e^{1/2} e^{-1/2(|x|-1)^2} = \frac{2}{\sqrt{2\pi}} e^{+1/2}$$

Ex 2: $\mathcal{N}(0,1)$ à partir de Cauchy:

$g_\theta(x) = \frac{\theta}{\pi}\frac{1}{\theta^2+x^2}$ sur $\mathbb{R}$ $\quad \theta>0$

($X \sim g \rightsquigarrow$ inversion + chgt d'échelle + symétrie)

$$C_\theta = \sup_{\mathbb{R}} \frac{f(x)}{g_\theta(x)} = \sup_{\mathbb{R}} \left(\log f(x) - \log g_\theta(x)\right)$$

$h(x) = \log f(x) - \log g_\theta(x) = -\log\sqrt{2\pi} - \frac{1}{2}x^2 - \log\theta + \log\pi + \log(x^2+$

$$h'(x) = -x + \frac{2x}{\theta^2+x^2} = x\left(-1+\frac{2}{\theta^2+x^2}\right)$$

$x=0$ ou $\theta^2+x^2=2 \Rightarrow x = \pm\sqrt{2-\theta^2}$ avec $\theta^2 \leq 2$

$$C_\theta = \begin{cases} \sqrt{\frac{2\pi}{e\theta}}\exp\left(\frac{\theta^2}{2}\right) & \text{si } \theta<\sqrt{2} \\ \theta\sqrt{\frac{\pi}{2}} & \text{si } \theta \geq \sqrt{2} \end{cases}$$

C_θ min pour $\theta=1$ et $C_1 = e^{-1/2}\sqrt{2\pi}$