Méthodes numériques en probabilités et statistique

A. Gloter et A. Guilloux

Organisation

- Les documents (slides des cours, codes python, etc) sont sur ma pageweb http://www.math-evry.cnrs.fr/members/aguilloux/welcome
- Pour me joindre agathe.guilloux@univ-evry.fr
- Le cours est en 3 parties
  - Partie 1 (A. Guilloux, 2 séances) : Simulations de v.a. et méthode de Monte Carlo
  - Partie 2 (A. Gloter, 3 séances) : Chaînes de Markov et MCMC
  - Partie 3 (A. Guilloux, 1 séance) : Bootstrap
- Le contrôle continu se fera en 2 parties : un DM (avec code) sur la partie 1, un contrôle d'une heure pour la partie 2
- Vous devez installer python 3 avec anaconda distribution (> 4) voir https://www.anaconda.com/products/individual

Chapitre 1 : Introduction

Introduction générale : notations

- On note X une variable aléatoire (v.a.) , i.e. une application mesurable de $(\Omega, \mathcal{F}, \mathbb{P})$ dans $(E, \mathcal{E})$.
- On va considérer presque partout que $E = \mathbb{R}$ et $\mathcal{E} = \mathcal{B}_{\mathbb{R}}$
- Dans ce cas, on notera F la fonction de répartition (f.d.r.) de cette v.a. et, si elle existent, f sa densité, $\mathbb{E}(X)$ et $\mathbb{V}(X)$ ses espérance et variance.

Introduction générale : échantillon et réalisation

- En probabilités, comme en statistique, on considère souvent des échantillons $X_1, \ldots, X_n$ constitués de v.a. indépendantes et identiquement distribuées (i.i.d.) on dit aussi que ce sont des copies i.i.d. de X
- On va chercher dans cette partie à obtenir pour une f.d.r. F donnée une réalisation
$$x_1 = X_1(\omega), \ldots, x_n = X_n(\omega)$$
de $X_1, \ldots, X_n$
- cela servira à illustrer certains résultats de probabilités, faire des calculs numériques (intégrales, etc), cela sert aussi en cryptographie par exemple.

Vraie fonction de répartition

Définition

Une fonction F de $\mathbb{R}$ dans $[0, 1]$ est une fonction de répartition si elle est croissante, a pour limite 0 en $-\infty$, 1 en ∞ et est continue à droite avec limites à gauche.

X a pour f.d.r. F si et seulement si pour tout $x \in \mathbb{R}$: $\mathbb{P}(X \leq x) = F(x)$.

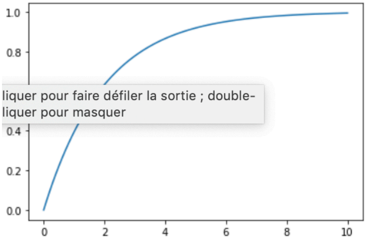


Figure 1: f.d.r de la loi $\mathcal{E}(2)$

Vraie fonction quantile

Définition

Soit F une fonction de répartition, on définit la fonction quantile associée F^{-} de $]0,1[$ dans $\mathbb{R}$ comme l'inverse généralisée de F, i.e. pour tout $u \in]0,1[$

$$F^{-}(u) = \inf\left\{x \in \mathbb{R}, F(x) \geq u\right\}.$$

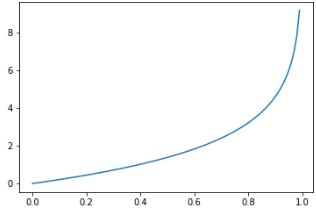


Figure 2: f.d.r de la loi $\mathcal{E}(2)$

Remarque : Si F^{-1} existe alors $F^{-} = F^{-1}$.

Lien entre la f.d.r. et la fonction quantile

Propriété

Soit F une f.d.r. et F^{-} sa fonction quantile, on a

- F^{-} est croissante et continue à gauche,
- et pour tous $x \in \mathbb{R}$ et $p \in]0,1[$:

$$F(x) \geq p \Longleftrightarrow x \geq F^{-}(p) \tag{1}$$

- pour tout $p \in]0,1[$

$$F\big(F^{-}(p)\big) \geq p$$

  avec égalité si F est continue en $F_{-}(p)$
- pour tout $x \in \mathbb{R}$

$$F^{-}\big(F(x)\big) \leq x$$

  avec égalité si F^{-} est continue en $F(x)$

Preuve de (1) $F(x) \geq p \Leftrightarrow x \geq F^-(p)$

Soient $p \in]0,1[$ et $x \in \mathbb{R}$

Si $F(x) \geq p$ alors $x \in \{y \in \mathbb{R}, F(y) \geq p\}$

par définition de F^-, x est + gd

que $\inf\{y \in \mathbb{R}, F(y) \geq p\} = F^-(p)$

donc $x \geq F^-(p)$.

Réciproque : on suppose que $x \geq F^-(p)$

$\forall \varepsilon > 0$ $x + \varepsilon > F^-(p)$ alors $F(x+\varepsilon) \geq p$

on laisse $\varepsilon \searrow 0$ comme F est càd

$$F(x+\varepsilon) \to F(x) \geq p.$$

Pour montrer $F(F^-(p)) \geq p$ on pose

$x = F^-(p)$ dans (1).

Fonction de répartition empirique (1)

On a maintenant un échantillon i.i.d. $X_1, \ldots, X_n$ de v.a. réelles

Définition

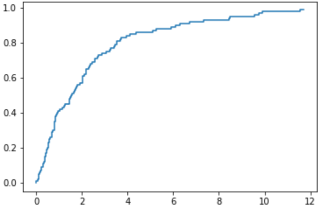
La fonction de répartition empirique F_n associée à l'échantillon $X_1, \ldots, X_n$ est définie pour tout $x \in \mathbb{R}$ par

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(X_i \le x)}. = \text{proportion aléatoire de } X_i \le x \text{ dans l'échantillon.}$$

$$\xrightarrow{p.s.} \mathbb{E}(\mathbb{1}(X_i \le x)) = \mathbb{P}(X_i \le x) = F(x)$$

Pour une réalisation $x_1 = X_1(\omega), \ldots, x_n = X_n(\omega)$ la fonction $F_n^\omega(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{(X_i(\omega) \le x)}$ est une fonction de répartition constante par morceaux.

elle est constante par morceaux et prend les valeurs $0, \frac{1}{n}, \frac{2}{n}, \ldots, 1$



Fonction de répartition empirique (2)

Théorème de Glivenko-Cantelli

Si $X_1, \ldots, X_n$ sont i.i.d. de f.d.r F alors

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \stackrel{p.s.}{\rightarrow} 0.$$

Preuve : cf. cours de Proba

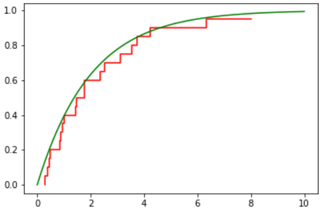


Figure 4: f.d.r vraie et empirique $n = 20$

Fonction de répartition empirique (3)

Théorème de Glivenko-Cantelli

Si $X_1, \ldots, X_n$ sont i.i.d. de f.d.r F alors

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \overset{p.s.}{\to} 0.$$

Preuve : cf. cours de Proba

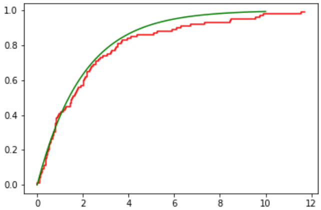


Figure 5: f.d.r vraie et empirique $n = 100$

Fonction quantile empirique (1)

Définition

La fonction quantile empirique F_n^- associée à l'échantillon $X_1, \ldots, X_n$ est définie comme l'inverse généralisée de F_n, i.e. pour tout $p \in]0,1[$ par

$$F_n^-(u) = \inf\Big\{x \in \mathbb{R}, F_n(x) \geq u\Big\}.$$

⚠ $F^- :]0,1[\to \mathbb{R}$

Propriété

Pour tout $p \in]0,1[$, on a

$$F_n^-(\overset{p}{\cancel{u}}) = X_{(\lceil n\overset{p}{\cancel{u}} \rceil)}$$

où $X_{(1)} \leq X_{(2)} \leq \ldots \leq X_{(n)}$ sont les statistiques d'ordre de l'échantillon $X_1, \ldots, X_n$.

Preuve

On veut calculer

$$F_n^-(u) = \inf\{x \in \mathbb{R}, F_n(x) \geq u\}$$

$$= \inf\{x \in \mathbb{R}, F_n(x) \geq \frac{k_u}{n}\}$$

$$= X_{(k_u)}$$

$\forall u \in]0,1[$

$\exists ! \, k_u \in \{1, \dots, n\}$

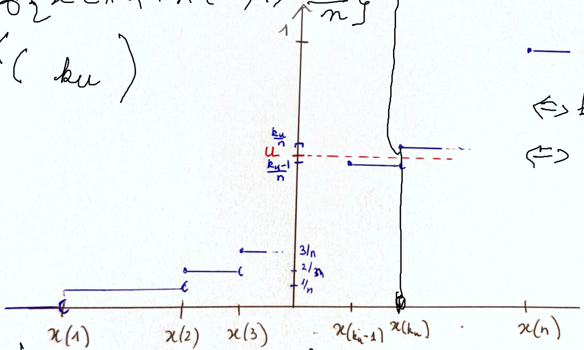
tel que

$$\frac{k_u - 1}{n} < u \leq \frac{k_u}{n}$$

$$\iff k_u - 1 < un \leq k_u$$

$$\iff k_u = \lceil un \rceil$$

ceiling.



donc $F_n^-(u) = X_{(\lceil nu \rceil)}$.

Figure 6: F_n^ω et quantiles empiriques

Preuve

Fonction quantile empirique (2)

Propriété

Pour tout $p \in]0,1[$, si F^{-} est continue en p,on a

$$F_n^{-}(p) \overset{p.s.}{\to} F^{-}(p).$$

Preuve via le théorème de Glivenko-Cantelli et la continuité de F^{-} en p.

Si $X_1, \ldots, X_n$ sont i.i.d. de f.d.r. F on doit donc avoir pour tout $k \in [\![1, n-1]\!]$

$$F_n^{-}(k/n) = X_{(k)} \simeq F^{-}(k/n).$$

Quantile-quantile plot - QQ-plot

Pour vérifier si $X_1, \ldots, X_n$ ont la f.d.r. H, on représente le nuage de points

$$\left(X_{(k)}, H^{-}(k/n)\right)_{k \in [\![1, n-1]\!]}$$

si H est la f.d.r., les points doivent s'aligner.

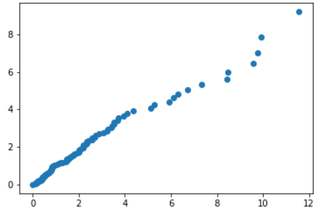


Figure 7: Réalisations d'une $\mathcal{E}(2)$ et quantiles de la $\mathcal{E}(2)$

Quantile-quantile plot - QQ-plot

Pour vérifier si $X_1, \ldots, X_n$ ont la f.d.r. H, on représente le nuage de points

$$\left(X_{(k)}, H^{-}(k/n)\right)_{k \in [\![1, n-1]\!]}$$

si H est la f.d.r., les points doivent s'aligner.

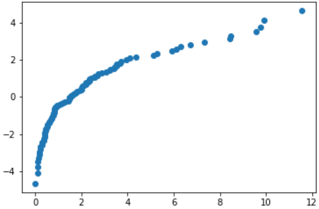


Figure 8: Réalisations d'une $\mathcal{E}(2)$ et quantiles de la $\mathcal{N}(0, 1)$

Générateurs congruentiels

Les générateurs de nombres alétoires programmés dans la plupart des librairies (c'est le cas dans `numpy`) sont de type congruentiel : ils renvoient les termes d'un suite $\left(z_n/m\right)_{n\geq 0}$ définie par

$$\begin{cases} z_0 & \text{seed ou racine} \\ z_{n+1} = az_n + c & \text{(modulo } m\text{)} \end{cases}$$

Cette suite imite très bien le hasard uniforme si la période est très (très très) grande.

Minimal standard

On prend $m = 2^{31} - 1$ et $a = 7^5$, la période est de $2^{31} - 1 \simeq 2 \times 10^9$

Mersenne twister - Matsumoto et Nishimura (1997)

L'algorithme a été amélioré pour atteindre la période de $2^{19937} - 1$.

Chapitre 2: Simulation de variables aléatoires

Introduction

On se donne une f.d.r. F (ou d'une loi) et on souhaite obtenir des réalisations $x_1, \ldots, x_n$ de $X_1, \ldots, X_n$ de f.d.r. F - on dira simuler selon F.

Suivant la forme de F, l'un ou l'autre des méthodes que nous allons voir sera la plus judicieuse.

Propriété fondamentale

Théorème

Soient U une v.a. de loi $\mathcal{U}_{]0,1[}$ et une f.d.r. F alors $F^{-}(U)$ a pour f.d.r. F

La méthode d'inversion marche si on dispose d'une forme explicite pour F^{-} : loi exponentielle, de Weibull, lois discrètes, etc

MAIS ne marche pas autrement : loi normale, loi de Student, loi du X^2, loi gamma...

Preuve $F^{-}(U)$ a pour fdr $F \iff \forall x \in \mathbb{R}$
on veut montrer que. $\mathbb{P}(F^{-}(U) \leq x) = F(x)$

mais on sait que $\forall x \in \mathbb{R}, \forall p \in]0,1[$

$$F(x) \geq p \iff x \geq F^{-}(\alpha)$$

$$\{F^{-}(U) \leq x\} = \{U \leq F(x)\}$$

donc $\mathbb{P}(F^{-}(U) \leq x) = \mathbb{P}(U \leq F(x))$

$U \sim \mathcal{U}_{[0,1]}$ donc la fdr de U est l'identité

sur $[0,1]$ donc $\mathbb{P}(U \leq F(x)) = F(x)$ $\square$.

Exercise

Loi exponentielle

Appliquer cette méthode à la loi exponentielle $\mathcal{E}(\lambda)$, pour $\lambda > 0$.

La fdr de $\mathcal{E}(\lambda)$ est $G : x \to 1 - e^{-\lambda x}$ sur $\mathbb{R}_+$

son inverse $y \mapsto -\frac{1}{\lambda}\log(1-y)$

donc la v.a. $-\frac{1}{\lambda}\log(1-\alpha)$

suit la loi $\mathcal{E}(\lambda)$.

Principe de la méthode de rejet

On se donne une f.d.r. F pour laquelle on souhaite obtenir des réalisations $x_1, \dots, x_n$ de $X_1, \dots, X_n$ de f.d.r. F et on suppose maintenant qu'elle admet une densité f

$$F(x) = \int_{-\infty}^{x} f(u)du \text{ pour tout } x \in \mathbb{R}.$$

On choisit une densité g (suivant laquelle on sait déjà simuler) telle qu'il existe une constante $c > 0$ et

$$f(x) \leq cg(x) \text{ pour tout } x \in \mathbb{R}. \quad \Rightarrow \int f \leq c \int g$$

$$\Rightarrow 1 \leq c$$

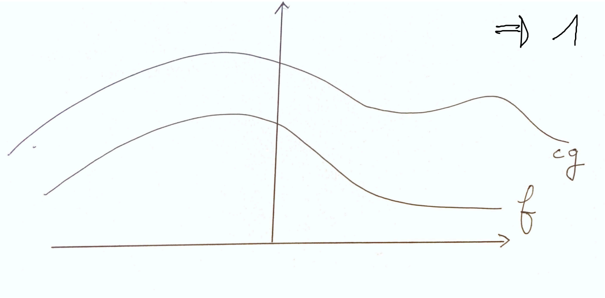


Figure 9: f et cg

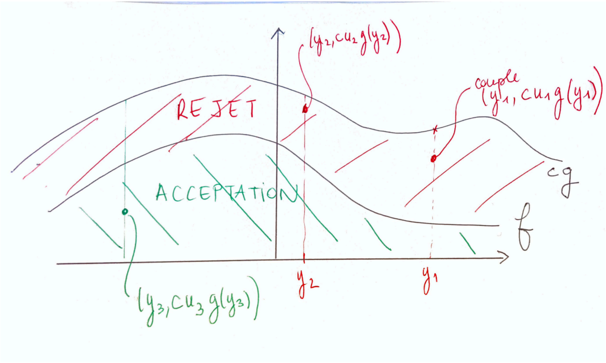
Algorithme

Théorème

Soient :

- $\left(Y_n\right)_{n\geq 0}$ suite de v.a. i.i.d. de densité g indépendante de
- $\left(U_n\right)_{n\geq 0}$ suite de v.a. i.i.d. de loi $\mathcal{U}_{]0,1[}$.

On note $\tau = \inf\left\{n \geq 1, U_n c g(Y_n) \leq f(Y_n)\right\}$ alors Y_τ a pour densité f.



On veut montrer que Y_τ a pour densité f
cela revient à montrer que

$$\forall x \in \mathbb{R} \quad \mathbb{P}(Y_\tau \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

⚠ les Y_k et τ sont des v.a $\quad \tau(\Omega) = \mathbb{N}^*$

On a $\mathbb{P}(Y_\tau \leq x) = \mathbb{P}\left(Y_\tau \leq x \cap \left(\bigcup_{k \geqslant 1} \tau = k\right)\right)$

$$= \mathbb{P}\left(\bigcup_{k \geqslant 1} \left\{(Y_\tau \leq x) \cap (\tau = k)\right\}\right) = \sum_{k \geqslant 1} \mathbb{P}(Y_\tau \leq x \cap \tau = k)$$

$$= \sum_{k \geqslant 1} \mathbb{P}(Y_k \leq x, \tau = k)$$

$$\{Y_k \le x \cap \tau = k\} = \underbrace{\bigcap_{l=1}^{k-1} \{(U_l \, c \, g(Y_l) > f(Y_l))\} \cap \{(U_k \, c \, g(Y_k) \le f(Y_k))}_{\{\tau = k\}} \cap (Y_k \le x)\}$$

$$\mathbb{P}(Y_k \le x \cap \tau = k) = \mathbb{P}(\_\_) = \prod_{l=1}^{k-1} \mathbb{P}(U_l \, c \, g(Y_l) > f(Y_l)) \cdot \, !$$

$$\mathbb{P}(U_k \, c \, g(Y_k) \le f(Y_k) \cap Y_k \le x)$$

On calcule :

$\mathbb{P}(U_e\, c\, g(Y_e) > f(Y_e))$ dépend de la loi de (U_e, Y_e).

dont la densité est $\mathbb{1}_{]0,1[}(u)\, g(y)$ car $U_e \perp\!\!\!\perp Y_e$

$$= \iint \mathbb{1}_{(u c g(y) > f(y))}\, \mathbb{1}_{]0,1[}\, g(y)\, du\, dy$$

$$= \int g(y) \int_0^1 \mathbb{1}_{\left(u > \frac{f(y)}{c g(y)}\right)}\, du\, dy$$

Par définition $f(y) \leq c\, g(y) \Leftrightarrow \frac{f(y)}{c\, g(y)} \leq 1$

$$= \int g(y) \left(1 - \frac{f(y)}{c\, g(y)}\right) dy$$

$$= \int g(y)\, dy - \int \frac{\cancel{g(y)}\, f(y)}{c\, \cancel{g(y)}}\, dy = 1 - \frac{1}{c} \int f = 1 - \frac{1}{c}.$$

$$\mathbb{P}(U_k\, c\, g(Y_k) \le f(Y_k),\ Y_k \le x)$$

$$= \iint \mathbb{1}_{(u c g(y) \le f(y))} \mathbb{1}(y \le x)\, \mathbb{1}_{]0,1[}(u)\, g(y)\, du\, dy$$

$$= \int \mathbb{1}_{(y \le x)}\, g(y) \int_0^1 \mathbb{1}\left(u \le \frac{f(y)}{c\, g(y)}\right) du \quad dy$$

$$= \int_{\mathbb{R}} \mathbb{1}_{(y \le x)}\, \cancel{g(y)}\, \frac{f(y)}{c\, \cancel{g(y)}}\, dy$$

$$= \frac{1}{c} \int_{-\infty}^{x} f(y)\, dy$$

$$= \left[\frac{1}{1-\left(1-\frac{1}{c}\right)} \right] \frac{1}{c} \int_{-\infty}^{x} f(y)\,dy$$

$$= \cancel{c}\ \frac{1}{\cancel{c}} \int_{-\infty}^{x} f(y)\,dy \quad \square$$

$$= \int_{\mathbb{R}} \mathbb{1}(y \leq x)\, \cancel{g(y)} \frac{f(y)}{c\,\cancel{g(y)}}\, dy$$

$$= \frac{1}{c} \int_{-\infty}^{x} f(y)\,dy$$

<u>Remarque</u> : A partir d'une loi $U_{]0,1[^d}$ on obtient les lois uniformes sur n'importe quel borélien $B \subset]0,1[^d$.

exemple

On a une loi uniforme sur $]0,1[^d$ à partir de (U_1, U_2) $U_1 \perp U_2$ de loi $\mathcal{U}\,]0,1[$

Preuve

Preuve

Simulation suivant la loi gaussienne

Algorithme de Box et Muller - 1958

Soient deux v.a. $U \sim \mathcal{U}_{]0,1[}$ et $V \sim \mathcal{E}(1)$ indépendantes, on définit

$$\begin{cases} X = & \sqrt{2V}\cos(2\pi U) \\ Y = & \sqrt{2V}\sin(2\pi U). \end{cases}$$

X et Y sont deux v.a. indépendantes de loi $\mathcal{N}(0,1)$

Remarque

On peut obtenir un couple gaussienne de loi $\mathcal{N}(\mu,\Sigma)$ en translatant (X,Y) de μ et en utilisant la décomposition de Cholesky : Σ est définie positive donc il existe une matrice triangulaire inférieur L telle que $LL^\top = \Sigma$.

Le vecteur $L(X,Y)^\top + \mu$ a la loi $\mathcal{N}(\mu,\Sigma)$.

I Il faut montrer que la densité du couple (X,Y)

$$\frac{1}{2\pi} \exp\left(-\frac{1}{2}(x^2+y^2)\right) \qquad \sim \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\right)$$

$$L\begin{pmatrix}X\\Y\end{pmatrix}+N \qquad \mathbb{E}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+N\right) = L\,\mathbb{E}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)+N = N.$$

$$\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+\mu\right)=\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}\right)=L\,\underbrace{\mathbb{V}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)}\,L^T$$

$$=L\begin{pmatrix}1&0\\0&1\end{pmatrix}L^T$$

$$=LL^T=\Sigma.$$

$$\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+\mu\right) = \mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}\right) = L\,\mathbb{V}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)L^T$$

$$= L\begin{pmatrix}1&0\\0&1\end{pmatrix}L^T$$

$$= LL^T = \Sigma.$$

→ Exercice sur la méthode de reject.

$\mathbb{E}(\tau)$. $\tau \sim \mathcal{G}(p)$ $p=?$

Preuve

Preuve

Méthode Monte-Carlo

C'est une méthode utilisant la loi des grands nombres qui sert à approximer une valeur numérique. Son nom est dû aux jeux des casinos de Monte-Carlo, elle a été introduite par N. Metropolis (1949).

Loi des grands nombres

Soit $\left(Z_n\right)_{n\geq 1}$ une suite de v.a. i.i.d. intégrables ($\mathbb{E}\left(|Z_1|\right) < \infty$) on a

$$\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i \overset{p.s.}{\to} \mathbb{E}(Z_1).$$

Inégalité de Bienaymé-Tchebychev

Soit $\left(Z_n\right)_{n\geq 1}$ une suite de v.a. i.i.d. admettant une variance($\mathbb{E}\left(|Z_1|^2\right) < \infty$) on a

$$\mathbb{P}\left(|\overline{Z}_n - \mathbb{E}(Z_1)| > \epsilon\right) \leq \frac{\mathbb{V}(Z_1)}{n\epsilon}.$$

Principe de la méthode (1)

On souhaite approximer l'intégrale

$$I = \int_\Delta g,$$

où g est une fonction de $\mathbb{R}^d$ dans $\mathbb{R}$ que l'on suppose intégrable sur Δ ($\int_\Delta |g| < \infty$).

On remarque pour cela que l'on peut toujours récrire l'intégrale I de la façon suivante :

$$I = \int_\Delta \frac{g}{f_X} f_X = \int \frac{g}{f_X} f_X$$

où f_x est une densité sur $\mathbb{R}^d$ de support Δ (pour $x \in \mathbb{R}^d$, $f_X(x) > 0 \Leftrightarrow x \in \Delta$).

Il faut montrer que la densité du couple (X,Y)

$$\frac{1}{2\pi}\exp\left(-\frac{1}{2}(x^2+y^2)\right)$$

$$\sim \mathcal{N}\left(\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1&0\\0&1\end{pmatrix}\right)$$

$$L\begin{pmatrix}X\\Y\end{pmatrix}+\mu$$

$$\mathbb{E}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+\mu\right) = L\,\mathbb{E}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)+\mu=\mu.$$

$$\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+\mu\right)=\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}\right)=L\,\underbrace{\mathbb{V}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)}\,L^T$$

$$=L\begin{pmatrix}1&0\\0&1\end{pmatrix}L^T$$

$$=LL^T=\Sigma.$$

$$\mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}+\mu\right) = \mathbb{V}\left(L\begin{pmatrix}X\\Y\end{pmatrix}\right) = L\,\underbrace{\mathbb{V}\left(\begin{pmatrix}X\\Y\end{pmatrix}\right)}\,L^T$$

$$= L\begin{pmatrix}1&0\\0&1\end{pmatrix}L^T$$

$$= LL^T = \Sigma.$$

→ Exercice sur la méthode de reject.

$\mathbb{E}(\tau)$ $\quad\tau \sim \mathcal{G}(p)\quad p = ?$

Principe de la méthode (1)

Méthode de Monte-Carlo

Si X est une v.a. de $\mathbb{R}^d$ de densité f_X, on peut alors écrire :

$$I = \mathbb{E}\Big(\frac{g}{f_X}(X)\Big).$$

Soient $\big(X_n\big)_{n\geq 1}$ une suite v.a. i.i.d. de même loi que X, on approxime I par

$$\bar{I}_n = \frac{1}{n}\sum_{i=1}^{n}\frac{g}{f_X}(X_i).$$

$\bar{I}_n$ converge p.s. vers I

Preuve

Illustration pour approximer π (1)

On pose

$$I = 4\int_{[-1,1]^2} \mathbb{1}_{(x^2+y^2\leq 1)} dxdy = \pi.$$

On considère U et V deux v.a. indépendantes de loi uniforme sur $[-1,1]$, on peut récrire I comme

$$I = \mathbb{E}(\mathbb{1}_{(U^2+V^2\leq 1)}).$$

On peut alors l'approximer à partir de copies i.i.d. de U et V

$$\bar{I}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(U_i^2+V_i^2\leq 1)}$$

Illustration pour π (2)

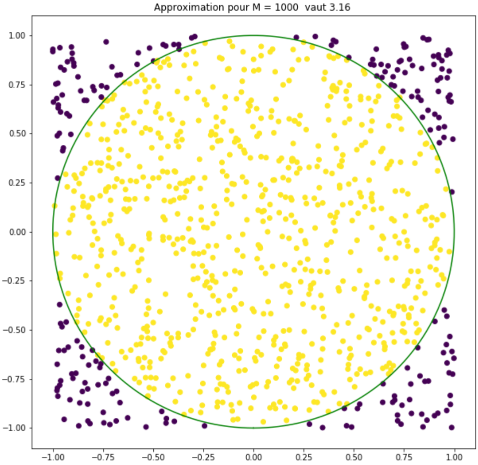
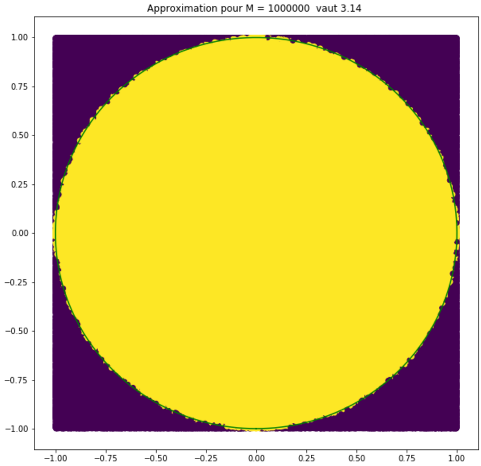


Illustration pour π (3)



Comment choisir f_X ?

On suppose maintenant que l'inégalité de Bienaymé-Tchebychev est vérifiée (i.e. $\int g^2/f_x < \infty$), on a alors

$$\mathbb{P}\Big(|\bar{I}_n - I| > \epsilon\Big) \leq \frac{\mathbb{V}\big(g/f_X(X)\big)}{n\epsilon}.$$

Donc, à ϵ fixé, plus $\mathbb{V}\big(g/f_X(X)\big)$ est petite, plus petite sera l'erreur d'approximation !

On calcule facilement

$$\mathbb{V}\big(g/f_X(X)\big) = \int_\Delta \frac{g^2(x)}{f_X(x)} dx - I^2$$

c'est le premier terme qu'on va chercher à rendre petit.

Exemple avec un extrême gaussien (1)

Soit $X \sim \mathcal{N}(0,1)$, on note φ sa densité. On veut approximer

$$I = \mathbb{P}(X > 5).$$

On récrit

$$I = \mathbb{P}(X > 5) = \mathbb{E}\left(\mathbb{1}_{(X>5)}\right) = \int \mathbb{1}_{(x>5)}\varphi(x)dx = \int \frac{\mathbb{1}_{(x>5)\varphi(x)}}{\varphi(x)}\varphi(x)dx.$$

Avec les notations précédentes, on a

$$g(x) = \mathbb{1}_{(x>5)}\varphi(x) \text{ et } f_X(x) = \varphi(x) \text{ soit } \frac{g}{f_X}(x) = \mathbb{1}_{(x>5)}$$

On peut donc approximer I par

$$\bar{I}_n = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{(X_i>5)} \text{ avec } \left(X_n\right)_{n\geq 1} \text{ i.i.d. } \mathcal{N}(0,1).$$

Est ce une bonne idée ?

Exemple avec un extrême gaussien (2)

Dans ce cas, on a

$$\int \frac{g^2(x)}{f_X(x)} dx = \int_5^{\infty} \frac{1}{\varphi(x)} dx$$

MAIS φ prend de très petites valeurs après 5 ($\varphi(5) \simeq 1.49e - 06$), cette intégrale est donc très grande !!