AI09

Abréviations :

- v.a. : variable aléatoire
- i.i.d. : indépendant et identiquement distribuées
- p.s. : presque sûr

Simulation stochastique et méthode de Monte-Carlo

Révisions de statistiques et de probabilités

Probabilités

Soit Ω l'ensemble des issues possibles d'une expérience; $\mathcal{F} := \{A : A \subset \Omega\}$.

Une application $\mathbb{P} : \mathcal{F} \to [0, 1]$ est appelée probabilité si les deux propriétés suivantes sont vérifiées :

1. $\mathbb{P}(\Omega) = 1$

2. Pour toute séquence $(A_n, n \geq 1) \subset \mathcal{F}$ d'évènements disjoints, c'est à dire $A_i \cap A_j = \emptyset$ pour tout $i \neq j$, on a :

$$\mathbb{P}(\cup_{n\geq 1} A_n) = \sum_{n\geq 1} \mathbb{P}(A_n)$$

Cette propriété est appelée σ-additivité de $\mathbb{P}$

Variable aléatoire

Une variable aléatoire est une modélisation :

$$X : \Omega \to E$$

- Si E est un ensemble fini ou dénombrable, X est une variable aléatoire discrète.
- Si $E = \mathbb{R}$ (or $E = \mathbb{C}$), X est une variable aléatoire réelle (ou complexe).
- Si $E = \mathbb{R}^d, d \geq 2$, X est un vecteur aléatoire.
- Si $E = \mathbb{R}^{\mathbb{N}}$, X est une séquence aléatoire.
- Si $E = \mathbb{R}^{\mathbb{R}_+}$, X est un process aléatoire.

Loi de probabilité, fonction de répartition, densité

Loi de probabilité d'une variable aléatoire X

$$P_X(B) = \mathbb{P}(X \in B) \\ \text{pour } B \subset \mathbb{R}$$

Fonction de répartition d'une variable aléatoire X:

$$F(x) := \mathbb{P}(X \leq x) = P_X(]-\infty, x]) \\ \text{pour tout } x \in \mathbb{R}$$

Densité de probabilité d'une variable aléatoire X (si elle existe) :

$$f : \mathbb{R} \to \mathbb{R}_+ \text{ telle que} \\ F(x) = \int_{-\infty}^{x} f(u)du \\ \text{pour tout } x \in \mathbb{R}$$

Espérance et moments

$$\mathbb{E}(X) = \int_{\mathbb{R}} xdF(x) = \begin{cases} \sum_{i\geq 1} x_i \mathbb{P}(X = x_i) & \text{si } X \text{ est une v.a. discrete} \\ \int_{\mathbb{R}} xf(x)dx & \text{si X est absolument continue} \end{cases}$$

$$\mathbb{E}X^k = \int_{\mathbb{R}} x^k dF(x) \\ \text{avec } k = 1, 2, \ldots$$

$$Var(X) = \mathbb{E}[(X - \mathbb{E}X)^2] = \mathbb{E}[X^2] - (\mathbb{E}X)^2$$

Propriétés de l'espérance :

1. Si $\mathbb{E}|X| < \infty$ et $\mathbb{E}|Y| < \infty$ alors $\mathbb{E}(aX = bY) = a\mathbb{E}X + b\mathbb{E}Y$ pour toutes constantes a et b
2. Si $X \leq Y$ alors $\mathbb{E}X \leq \mathbb{E}Y$
3. Si $\mathbb{E}|X| < \infty$ alors $\mathbb{E}X < \mathbb{E}|X|$
4. Si X quasiment égal à 0, alors $\mathbb{E}X = 0$
5. Si $\mathbb{E}|X| < \infty$ alors $\mathbb{E}[1_A|$ existe pour tout $A \in \mathcal{F}$
6. Si X quasiment égal à Y et $\mathbb{E}|X| < \infty$, alors $\mathbb{E}|Y| < \infty$ et $\mathbb{E}X = \mathbb{E}Y$
7. Si $X \geq 0$ et $\mathbb{E}X = 0$ alors X quasiment égal à 0

Théorème (fondamental)

Soit X une variable aléatoire dans $\mathbb{R}^d$. On dit que X a une densité f si et seulement si pour toute fonction mesurable et bornée $h : \mathbb{R}^d \to \mathbb{R}^m$ on a :

$$\mathbb{E}[h(X)] = \int_{\mathbb{R}d} h(x)f(x)dx$$

ou a une loi P_X si et seulement si :

$$\mathbb{E}[h(X)] = \int_{\mathbb{R}^d} h(x)dP_X$$

Covariance

Pour toutes variables aléatoires X et Y dans L^2, on définit leur covariance comme suit :

$$Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$$

Propriétés :

1. $Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$
2. $Cov(X, X) = Var(X)$
3. $Cov(X, Y) = Cov(Y, X)$
4. $Cov(aX + bY, Z) = aCov(X, Z) + bCov(Y, Z)$, $(a, b \in \mathbb{R})$

Inégalités

Inégalité de Bienaymé-Tchebychev

Si $\mathbb{E}(X^2) < \infty$ alors, pour tout $\epsilon > 0$:

$$\mathbb{P}(|X - \mathbb{E}(X)| < \epsilon) \leq \frac{Var(X)}{\epsilon^2}$$

Inégalité de Jensen

Soit ϕ une fonction réelle convexe définie sur un ensemble de réalisations d'une variable aléatoire X. Si X et $\phi(X)$ sont intégrables, alors :

$$\phi(\mathbb{E}(X)) \leq \mathbb{E}(\phi(X))$$

Théorème de Cauchy-Schwarz

Si $\mathbb{E}(X^2) < \infty$ et $\mathbb{E}(Y^2) < \infty$, alors

$$\mathbb{E}(XY) \leq \sqrt{\mathbb{E}(X^2)}\sqrt{\mathbb{E}(Y^2)}$$

Types de convergence

On considère une séquence de variables aléatoires réelles $X_n, n = 1, 2, \ldots$ et une variable aléatoire X, toutes définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$

Définition : convergence presque sûre

On dit que X_n converge presque sûrement (ou avec une probabilité 1) vers la variable aléatoire X quand $n \to \infty$, si $\mathbb{P}(\{\omega : X_n(\omega) \to X(\omega) \text{ quand } n \to \infty\}) = 1$. On l'écrit comme suit :

$$X_n \xrightarrow{\text{presque sur}} X$$

Définition : convergence dans L^p

On dit que X_n converge dans L^p, $p \geq 1$ vers la variable aléatoire X quand $n \to \infty$, si $\mathbb{E}[|X_n|^p] < \infty$ et $\mathbb{E}[|X_n - X|^p] \to 0$ quand $n \to \infty$. On l'écrit comme suit :

$$X_n \xrightarrow{Lp} X$$

Définition : convergence en probabilité

On dit que la séquence de variables aléatoires X_n converge en probabilité vers la variable aléatoire X quand $n \to \infty$, si pour tout $\epsilon > 0$ on a :

$$\mathbb{P}(|X_n - X| > \epsilon) \to 0 \text{ quand } n \to \infty$$

On l'écrit comme suit :

$$X_n \overset{p}{\to} X$$

Définition : convergence en loi

Soit F_n la fonction de répartition de X_n et F la fonction de répartition de X.

On dit qu'une séquence de variables aléatoires X_n converge en loi ou en distribution vers la variable aléatoire X quand $n \to \infty$ si $F_n(x) \to F(x)$ pour tous les points de continuité x de F. On l'écrit comme suit :

$$X_n \xrightarrow{loi} X$$

$$\frac{\sqrt{n}}{\sigma}(\frac{S_n}{n} - \mu) \overset{d}{\to} N(0,1)$$

Distributions conditionnelles dans $\mathbb{R}^2$

Considérons un vecteur aléatoire (X, Y) dans $\mathbb{R}^2$ de fonction de répartition $F(x, y)$ et de densité de probabilité $f(x, y)$ si elle existe. Nous avons les propriétés suivantes :

1. Densité de probabilité marginale de X et de Y :
$$f_X(x) = \int_{\mathbb{R}} f(x,y)dy, \quad f_Y(y) = \int_{\mathbb{R}} f(x,y)dx$$
2. Densité de probabilité de X avec $\{Y = y\}$:
$$f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$$
3. Fonction de répartition marginale :
$$F_X(x) = F(x,\infty), \quad F_Y(y) = F(\infty, y)$$
4. Fonction de répartition marginale de X avec $\{Y = y\}$:
$$F_{X|Y}(x|y) = \mathbb{P}(X \leq x|Y = y) = \int_{-\infty}^{x} f_{X|Y}(u|y)du$$

Espérances conditionnelles dans $\mathbb{R}^2$

Considérons un vecteur aléatoire (X, Y) dans $\mathbb{R}^2$ de fonction de répartition $F(x, y)$ et de densité de probabilité $f(x, y)$ si elle existe. Nous avons les propriétés suivantes :

1. Espérance conditionnelle de X avec $\{Y = y\}$:
$$\mathbb{E}(X|Y = y) = \int_{\mathbb{R}} x d_x F_{X|Y}(x|y) = \int_{\mathbb{R}} x f_{X|Y}(x|Y)dx$$

2. Espérance conditionnelle de $h(X)$ avec $\{Y=y\}$:

$$\mathbb{E}(h(X)|Y=y) = \int_{\mathbb{R}} h(x) d_x F_{X|Y}(x|y) = \int_{\mathbb{R}} h(x) f_{X|Y}(x|y) dx$$

3. Espérance conditionnelle de $h(X)$ sur Y :

$$\mathbb{E}(h(x)|Y) = \int_{\mathbb{R}} h(x) f_{X|Y}(x|Y) dx$$

4. Variance conditionnelle de X avec $\{Y=y\}$:

$$Var(X|Y=y) = \mathbb{E}[(X - \mathbb{E}(X|Y=y)^2 | Y=y)]$$

Monte Carlo

$$\begin{aligned} h &: [0,1] \to \mathbb{R} \\ I &= \int_0^1 h(x) d(x) \\ &= \int_{\mathbb{R}} h(x) 1_{[0,1]} d(x) \\ &= \int_{\mathbb{R}} h(x) f(x) dx \\ &= \underbrace{\mathbb{E}[h(X)]}_{\text{probl̀eme stochastique}} \end{aligned}$$

Processus de Monte Carlo

1. Réalisation de variables aléatoires (sampling)
2. Précision
3. Calcul de n_min
4. Accélération

Simuler des variables aléatoires

Donner des réalisations d'une variable aléatoire dont on connaît la Loi. Par exemple : loi exponentielle $X \sim \mathcal{E}(x)$

Densité :

$$f(x) = \begin{cases} \lambda \exp^{-\lambda t}, & t \geq 0 \\ 0, & t \leq 0 \end{cases}$$

Types de variables aléatoires :

- variable aléatoire de loi uniforme sur $[0,1]$
- variable aléatoire discrète
- variable aléatoire réelle
- variable aléatoire vectorielle

Variable aléatoire de loi uniforme sur $[0,1]$

$X \sim \text{unif}(0,1)$:

$$X : \Omega \to [0,1]$$
$$f(x) = 1_{[0,1]}(x) = \begin{cases} 1, & \text{si } x \in [0,1] \\ 0, & \text{si } x \notin [0,1] \end{cases}$$

$$\begin{gathered}x_0 = \text{racine}\\ x_n = ax_{n-1} + b \mod m\\ a, b, m \text{ constantes}(\mathbb{N})\\ x \in \{1, 2, \ldots, m-1\}\\ \frac{x_n}{m} \in [0, 1]\end{gathered}$$

Test de χ^2 (chi-deux)

$N =$ Nombre de boules (fixe)

$h =$ Longueur des intervalles

$r =$ Nombre d'intervalles

$I_i =](i-1)h, ih]$, $i = 1, 2, \ldots, r$

N_i = Nombre de points (boules) dans $I_i \rightarrow$ variable aléatoire

Pour N grand, on a :

$$\begin{gathered}K_N = \sum_{i=1}^{r} \frac{(N_i - N/r)^2}{N/r}\\ K_N \sim \chi^2_{r-1}\end{gathered}$$

Tests :

- H_0 : la loi (générateur) est uniforme
- H_1 : la loi n'est pas uniforme

On a sous H_0 : $N_i \sim \mathcal{B}(N, p = \frac{1}{r})$

$$\mathbb{E}(N_i) = \frac{N}{r} = N * \frac{1}{r} = \text{nombre moyen de points dans } I_i$$

Variables aléatoires discrètes

$X : \Omega \rightarrow E = \{x_1, x_2, \ldots, x_n, \ldots\}$ de loi $p = (P_1, P_2, \ldots, P_n, \ldots)$

Par exemple : $\mathbb{P}(X = x_k) = p_k,\ k \geq 1$

Algorithme :

1. $u = RAND$, $u \sim \text{unif}(0, 1)$
2. $\tilde{X}(\omega) = \sum_{k \geq 1} x_k 1_{]P_0 + P_1 + \cdots + P_{k-1}; P_0 + P_A + \cdots + P_{k-1} + P_k]}(u(\omega))$ $\tilde{X}(\omega) = \sum_{k \geq 1} x_k 1_{]a_{k-1}, a_k]}(u(\omega))$

Démonstration

Montrer que : $\mathbb{P}(\tilde{X} = x_k) = p_k\ \forall k \geq 1$

Soit F_u la fonction de répartition de $u \sim \text{unif}(0, 1)$

$$\begin{aligned}\mathbb{P}(\tilde{X} = x_k) &= P(\sum_{k \geq 1} x_k 1_{]P_0 + P_1 + \cdots + P_{k-1}; P_0 + P_A + \cdots + P_{k-1} + P_k]}(u(\omega)) = x_k)\\ &= \mathbb{P}(u \in]P_0 + \cdots + P_{k-1}; P_0 + \cdots + P_k])\\ &= F_u(P_0 + \cdots + P_k) - F_u(P_0 + \cdots + P_{k-1})\\ &= P_0 + P_1 + \cdots + P_k - (P_0 + P_1 + \cdots + P_k - 1)\\ &= P_k\end{aligned}$$

Conclusion : $\tilde{X} \sim X(P)$

Exemple

$X : \Omega \to E\{-1, 0, 1, 3\}$ de loi $p = (0.1, 0.2, 0.2, 0.5)$

1. $u(\omega) = 0.59$
2. $x_1 1_{]0,0.1]}(u) + x_2 1_{]0.1,0.3]}(u) + x_3 1_{]0.3,0.5]}(u) + x_4 1_{]0.5,1]}(u) = x_4 = 3$

Variables aléatoires réelles

$X : \Omega \to \mathbb{R}$

Fonction de répartition de la variable aléatoire X : F

Hypothèse : F est strictement croissante $\Rightarrow$ F est inversible (F^{-1} existe).

Théorème

Si $u \sim \text{unif}(0,1)$ alors $F^{-1}(u) \sim X$

Exemple :

$y = F(x) = 1 - \exp^{-\lambda x}$, $x \geq O$

$-\lambda x = \ln(1-y) \Rightarrow F^{-1} : x = -\frac{1}{n}\ln(1-y)$

Démonstration

On note $Y = F^{-1}(u)$:

$$\begin{aligned} F_Y(x) &= \mathbb{P}(Y \leq x) \\ &= \mathbb{P}(F^{-1}(u) \leq x) \\ &= \mathbb{P}(F(F^{-1}(u)) \leq F(x)) \\ &= \mathbb{P}(u \leq F(x)) \\ &= F_u(F(x)) \\ &= F(x) \end{aligned}$$

$\forall x$ on a : $F_Y(x) = F(x)$

Problème : donner des réalisation de X

1. $u = RAND$, $u \sim \text{unif}(0,1)$
2. Calculer $F^{-1}(u)$
3. $\tilde{X}(\omega) = F^{-1}(u(\omega))$

Exemple :

$X \sim \mathcal{E}(\lambda)$, $\lambda > 0$

1. $u = \text{RAND}$
2. Fonction inverse :

$$F(x) = 1 - \exp^{-\lambda x} = y$$
$$F^{-1}(u) = -\frac{1}{\lambda}\ln(1-u) = \tilde{X}$$

3. $X(\omega) = -\frac{1}{\lambda}\ln(1 - u(\omega))$

Comme $u \sim 1 - u$: $X(\omega) = -\frac{1}{\lambda}ln(u(\omega))$

Application :

$X \sim \mathcal{E}(\lambda)$, $\lambda = 0.01$

1. $u(\omega) = 0.2$
2. $X(\omega) = -\frac{1}{0.01}\ln(0.26) = 134.7$

On va estime/calculer approximativement la fonction de répartition par Monte Carlo

$X_1, \ldots, X_n$ i.i.d $\sim X$

$\frac{X_1+\cdots+X_n}{n} \cong \mathbb{E}(X)$

n = 1e6

Résultat = Somme / n

Variables aléatoires normales ou gaussiennes

$$Z : \Omega \to \mathbb{R}$$

On a : $Z \sim \mathcal{N}(0,1)$, $\begin{cases} \mathbb{E}(Z) = 0 \\ \mathrm{Var}(Z) = 1 \end{cases}$

Densité de probabilité

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp^{-x^2/2},\ x \in \mathbb{R}$$

Fonction de répartition

$$\Phi(x) = \int_{-\infty}^{x} \phi(u)du$$

Méthode de Box-Muller

Soient u_1, u_2 i.i.d $\sim \mathrm{unif}(0,1)$. Alors, les variables aléatoires

- $Z_1 = \sqrt{-2\ln u_1}\cos(2\pi u_2)$
- $Z_2 = \sqrt{-2\ln u_1}\sin(2\pi u_2)$

Sont i.i.d de loi $\mathcal{N}(0,1)$

Vecteur aléatoire Gaussien

$$Z : \Omega \to \mathbb{R}^d$$
$$\omega \mapsto Z(\omega) = \begin{pmatrix} Z_1(\omega) \\ \cdots \\ Z_d(\omega) \end{pmatrix}$$

Fonction de densité

$$f_z(x_1, \ldots, x_d) = (2\pi)^{-d/2}\exp(-\frac{1}{2}x^T x) = (2\pi)^{-d/2}\exp(-\frac{1}{2}(x|x))$$
$$(x|x) = x^T x = x_1^2 + \cdots + x_d^2$$

On note :

$$Z \sim N_d(0, I)$$
$$I = \begin{pmatrix} 1 & & \Large 0 \\ & \ddots & \\ \Large 0 & & 1 \end{pmatrix}$$

Exercice

Soit $Z \sim N_3(0, I)$. Donner une réalisation de Z. $Z(\omega) = \begin{pmatrix} -1.06 \\ 1.49 \\ -1.84 \end{pmatrix}$

Covariance et corrélation

Soient deux variables aléatoires $X_1, X_2 : \Omega \to \mathbb{R}$

Covariance de deux variables aléatoires X_1, X_2

$$\begin{aligned} Cov(X_1, X_2) &= \mathbb{E}[(X_1 - \mathbb{E}(X_1))(X_2 - \mathbb{E}(X_2))] \\ &= \mathbb{E}[X_1 X_2] - \mathbb{E}(X_1)\mathbb{E}(X_2) \end{aligned}$$

Propriétés de la covariance:

1. Si X_1 et X_2 sont indépendantes, alors $Cov(X_1, X_2) = 0$, la réciproque est fausse (sauf pour les variables aléatoires Gaussiennes)
2. $Cov(X_1, X_1) = \mathbb{E}[(X_1 - \mathbb{E}(X_1))^2] = Var(X_1)$
3. Linéaire
4. Symétrique

Coefficient de corrélation linéaire

$$-1 \leq p_{X_1, X_2} = \frac{Cov(X_1, X_2)}{\sigma_1 \sigma_2} \leq 1$$

$$X = \begin{pmatrix} X_1 \\ \vdots \\ X_d \end{pmatrix} : \Omega \to \mathbb{R}^d$$

Matrice de variances-covariances (ou de dispersion) du vecteur aléatoire X, notée K_X (matrice carrée)

$$K_X \in \mathbb{R}^{\mathrm{d*d}}$$
$$K_X = \mathbb{E}[(X - \mathbb{E}(X)\ (X - \mathbb{E}(\mathbb{X}))^T]$$

$$\mathbb{E}(X) = \begin{pmatrix} \mathbb{E}(X_1) \\ \vdots \\ \mathbb{E}(X_d) \end{pmatrix} \in \mathbb{R}^d$$

$$\mathbb{E}(X - \mathbb{E}(X)) = \begin{pmatrix} X_1 - \mathbb{E}(X_1) \\ \vdots \\ X_d - \mathbb{E}(X_d) \end{pmatrix}$$

$$K_X = \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \dots & Cov(X_1, X_D) \\ Cov(X_2, X_1) & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ Cov(X_d, X_1) & \dots & \dots & Var(X_d) \end{pmatrix}$$

Propriétés de K_X

1. Symétrique
2. Définie positive

$$\forall a \in \mathbb{R}^d, a^T K_X \geq 0$$
$$\sum_{i=1}^{d} \sum_{j=1}^{d} a_i K_X(i,j) a_j \geq 0$$

Théorème

Si $X \sim N_k(m_v, K_v)$ et $W = BV + b$ avec $B \in \mathbb{R}^{n*n}$ et $b \in \mathbb{R}^n$

Alors

$$\begin{cases} K_W = BK_V B^T \\ W \sim N_V(B * m + b; K_W) \end{cases}$$

Problème

Donner des réalisation de X connaissant m et K

$$X : \Omega \to \mathbb{R}^d$$
$$X \sim N_d(m, K), \begin{cases} m \in \mathbb{R}^d \\ K = \mathbb{E}[(X - \mathbb{E}(X))(X - \mathbb{E}(X))^T] \end{cases}$$

1. $Z \sim N_d(0, I)$
2. $X = CZ + m \Rightarrow K_X = CK_Z C^T = CIC^T = K$ où $CC^T = K$ Méthode de Cholesky : $C = chol(K)$

Exercice : Réalisation d'un vecteur aléatoire Gaussien

Soient $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \sim N_3(m, \Lambda)$

avec $m = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\Lambda = \begin{pmatrix} 2 & 1 & 3 \\ 1 & 4 & -2 \\ 3 & -2 & 8 \end{pmatrix}$

et $Y_1 = X_1 + X_2 + X_3$

$Y_2 = 2X_2 + X_3$

1. Donner la loi de $Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix}$

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix} = BX + b = BX + 0$$

$$m_Y = B * m = \begin{pmatrix} 6 \\ 7 \end{pmatrix}$$

$$K_Y = BK_X B^T = B\Lambda B^T = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 3 \\ 1 & 4 & -2 \\ 3 & -2 & 8 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} 18 & 15 \\ 15 & 16 \end{pmatrix}$$

2. Donner une réalisation de $Y(\omega)$
    1. $Z \sim N_2(0, I)$ *(Box-Muller)*
    2. $C = chol(K_Y)$
    3. $Y = CZ + m_Y$

Accélération

Schéma de Monte-Carlo

$$X : \Omega \to \mathbb{R} \overset{h}{\to} R$$

Problème : $\mathbb{E}[h(X)]$, $\mathbb{E}|H(X)| < \infty$

$$X_1, \ldots, X_n \text{ i.i.d} \sim X$$
$$\frac{h(X_1) + \cdots + h(X_n)}{n} \xrightarrow[n\to\infty]{p.s.} \mathbb{E}[h(X)]$$

Pour n assez grand, on a :

$$R_n = \frac{h(X_1) + \cdots + h(X_n)}{n} - \mathbb{E}[h(X)] \cong 0$$

Précision : (ϵ, θ)

$$P(|R_n| > \epsilon) \leq \theta$$
$$\Leftrightarrow P(|R_n| < \epsilon) > 1 - \theta$$

Inégalité de Bienaymé-Tchebychev

Soit Y tel que $\mathbb{E}(Y^2) < \infty$, on a $\forall \epsilon > 0$:

$$P(|Y - \mathbb{E}(Y)| > \epsilon) \leq \frac{Var(Y)}{\epsilon^2}$$

1. $\mathbb{E}[\frac{1}{n}\sum_{i=1}^n h(X_i)] = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[h(X)] = \mathbb{E}[h(X)]$
2. $Var[\frac{1}{n}\sum_{i=1}^n h(X_i)] = \frac{1}{n^2}\sum_{i=1}^n Var(h(X_i)) = \frac{Var(h(X))}{n}$

On prend pour $Y = \frac{h(X_1)+\cdots+h(X_n)}{n}$

$$P\left(\left|\frac{h(X_1) + \cdots + h(X_n)}{n} - \mathbb{E}|[h(X)]\right| > \epsilon\right) \leq \frac{Var(h(X))}{n\epsilon^2}$$

On a :

$$P(|R_n| > \epsilon) \leq \frac{Var(h(X))}{n\epsilon^2} \leq \theta$$
$$\Rightarrow n \geq \frac{Var(h(X))}{\theta\epsilon^2}$$

On estime : $Var(h(X))$ sur $X_1, \ldots, X_m$, $(m << n)$

$$S^2 = \frac{1}{m-n}\sum_{i=1}^m (h(X_i) - \overline{h(X_i)^2})$$
$$\bar{h} = \frac{h(X_1) + \cdots + h(X_m)}{m}$$

Problème : Borner $Var(h(X)) \leq C$

$$n \geq \frac{C}{\theta\epsilon^2}$$

Exemple :

$$Y_i \sim \mathcal{B}(p)$$
$$Var(Y) = p(1-p)$$
$$C = 1/4$$

Calculer

$$I=\int_5^{10}\exp^{x^2}\,dx=\int_{\mathbb{R}}\exp^{x^2}\,1_{[5,10]}(x)dx=\int_{\mathbb{R}}5\exp^{x^2}\,\frac{1}{5}1_{5,10}(x)dx=\int_{\mathbb{R}}h(x)f_x(x)dx$$

Avec une précision $(\epsilon,\theta)=(10^{-2},10^{-2})$

$$I=\mathbb{E}[h(X)],\ X\sim unif(5,10)$$
$$h(x)=5\exp^{x^2}$$

1. $U\sim unif(0,1):u=RAND$
2. $X=5+(10-5)U$

Estimer $Var(h(X))$ par S^2 sur $(X_1,\ldots,X_{100})$, $m=100$

$$S^2=\frac{1}{100-1}\sum_{i=1}^{100}(h(X_i)-\overline{h(X)})^2$$
$$\overline{h(X)}=\frac{h(X_1)+\cdots+h(X_{100})}{100}$$

Algorithme : Calcul de S^2

Pour i = 1 à 100

U(i) = RAND

X(i) = 5 + (10-5)U(i)

$\overline{h}=\overline{h}+h(X(i))$

$\overline{h}=\overline{h}/100$

Pour i = 1 à 100

$S^2(i)=S^2(i)+(h(X_i)-\overline{h})^2$

$S^2=\frac{S_2(100)}{99}$

Formules récursives pour $\overline{X}$ et $\overline{S^2}$

$$\overline{X_j}=\frac{1}{j}\sum_{i=1}^{j}X_i\text{ et }\overline{S_j^2}=\frac{1}{j-1}\sum_{i=1}^{j}(X_i-\overline{X_j})^2$$

On pose : $\begin{cases}S_I^2=0\\ \overline{X_0}=0\end{cases}$

- $\overline{X_{j+1}}=\overline{X_j}+\frac{X_{j+1}-\overline{X_j}}{j+1}$
- $S_{j+1}^2=(1-\frac{1}{j}S_j^2)+(j+1)(\overline{X_{J+1}}-\overline{X_j})^2$

Accélération

Etant donnée la précision (ϵ,θ), comment l'atteindre avec n petit ?

$$\gamma=\mathbb{E}_f[h(X)]=\int_{\mathbb{R}}h(x)f(x)dx=\int\left[\frac{h(x)f(x)}{g(x)}g(x)dx\right]=\mathbb{E}_g[h'(Y)]$$
$$Y\sim g$$

$$Y_1,\ldots,Y_n,\ \text{i.i.d}\ \sim g\ (\text{fonction de r\'epartition})$$
$$\hat{\gamma}_n=\frac{1}{n}\sum_{i=1}^{n}h'(Y_i)=\frac{1}{n}\sum_{i=1}^{n}\left[\frac{h(Y_i)f(Y_i)}{g(Y_i)}\right]$$

La bonne fonction de densité est celle qui minimise $Var(\hat{\gamma}_n)$