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

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

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

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

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 :

$$X_n \xrightarrow{p} X$$

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 :

$$X_n \xrightarrow{loi} X$$

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

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

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

Processus de Monte Carlo

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