Vecteur Gaussiens, convergences et approximations

Plan :

Plan :

Matrice de variances-covariances

Définition

Soit X et Y deux v.a dans $L^2(\Omega, \mathcal{A}, P)$, (c-à-d : X et Y sont de carré intégrable), alors on définit la covariance entre X et Y par la formule :

$$Cov(X, Y) = E\Big((X - E(X))(Y - E(Y)\Big).$$

Matrice de variances-covariances

Propriétés

- $Cov(X, X) = V(X)$.
- $Cov(X, Y) = Cov(Y, X)$.
- **Formule de Koenig :** $Cov(X, Y) = E(XY) - E(X)E(Y)$.
- $Cov(aX, bY) = abCov(X, Y)$.
- $V(X + Y) = V(X) + V(Y) + 2Cov(X, Y)$.
- **Si** X **et** Y **sont indépendantes alors** $Cov(X, Y) = 0$.

Matrice de variances-covariances

Preuve

- $Cov(X,X) = E(X - E(X))^2 = V(X)$.
- Trivial.
- $Cov(X,Y) = E\Big[(X - E(X))(Y - E(Y))\Big] =$
  $E\Big[XY - XE(Y) - YE(X) + E(X)E(Y))\Big]$
  $= E(XY) - E(X)E(Y) - E(Y)E(X) + E(X)E(Y)) = E(XY) - E(X)E(Y)$.
- D'après la formule de Koenig, on a
  $Cov(aX, bY) = E(aXbY) - E(aX)E(bY)$
  $= abE(XY) - abE(X)E(Y) = abCov(X,Y)$.

Matrice de variances-covariances

- $$\begin{aligned} V(X+Y) &= E(X+Y)^2 - (E(X)+E(Y))^2 \\ &= [E(X^2)+E(Y^2)+2E(XY)] - [(E(X))^2+(E(Y))^2-2E(X)E(Y)] \\ &= [E(X^2)-(E(X))^2] + [E(Y^2)-(E(Y))^2] + 2[E(XY)-E(X)E(Y)] \\ &= V(X)+V(Y)+2Cov(X,Y). \end{aligned}$$
- **Il suffit de voir que si X et Y sont indépendantes alors** $E(XY) = E(X)E(Y)$.

Matrice de variances-covariances

Proposition

- **Si** (X, Y) **est un couple discret de loi** $\Big(P(X = x, Y = y)\Big)$**, alors**

$$Cov(X, Y) = \sum_{x \in E} \sum_{y \in F} (x - E(X))(y - E(Y))P(X = x, Y = y).$$

- **Si** (X, Y) **est un couple absolument continue de densité** $f_{(X,Y)}$ **à support** D**, alors**

$$Cov(X, Y) = \int_D \int_D (x - E(X))(y - E(Y))f_{(X,Y)}(x, y)dxdy.$$

Matrice de variances-covariances

Définition

- On appelle espérance mathématique du vecteur aléatoire $X = (X_1, ..., X_n)$ le vecteur $E(X)$ de $\mathbb{R}^n$ de composantes $E(X_1), ..., E(X_n)$.
- Soit $X = (X_1, ..., X_n)$ un vecteur aléatoire de composantes de carré intégrable. On appelle matrice de variances-covariances de X, la matrice carrée, notée Γ_X, d'ordre n et de terme général

$$Cov(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j), \;\; 1 \leq i, j \leq n.$$

Matrice de variances-covariances

Remarque

La matrice Γ_X est forcément symétrique et ses termes diagonaux sont les variances des composantes des vecteurs : $V(X_i) = Cov(X_i, X_i)$, $1 \leq i, j \leq n$.

Exemple

Pour le vecteur aléatoire $X = (X_1, X_2, X_3)$, on a

$$\Gamma_X = \begin{pmatrix} V(X_1) & Cov(X_1, X_2) & Cov(X_1, X_3) \\ Cov(X_2, X_1) & V(X_2) & Cov(X_2, X_3) \\ Cov(X_3, X_1) & Cov(X_3, X_2) & V(X_3) \end{pmatrix}$$

Plan :

Vecteurs Gaussiens

Définition

- **On dit que le vecteur $X = (X_1, ..., X_n)$ est gaussien ou normal dans $\mathbb{R}^n$ si toute combinaison linéaire des composantes du vecteur X est une v.a normale dans $\mathbb{R}$, c-à-d pour tout $\alpha = (\alpha_1, ..., \alpha_n) \in \mathbb{R}^n$, la v.a**

$$Z = \alpha' X = \alpha_1 X + ... + \alpha_n X_n$$

suit une loi normale dans $\mathbb{R}$.

Vecteurs Gaussiens

Proposition

- **Si** $X = (X_1, ..., X_n)$ **est un vecteur gaussien de moyenne** m **et de matrice de variances-covariances** Γ_X**, alors** $Z = \alpha' X$ **suit une loi normale** $\mathcal{N}(\alpha' m, \alpha' \Gamma_X \alpha)$.
- **La densité du vecteur gaussien** $X = (X_1, ..., X_n)$ **est donnée par :**

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt{det(\Gamma_X)}} \exp\left(-\frac{1}{2}\left[(x-m)' \Gamma_X^{-1} (x-m)\right]\right),$$

avec $x = (x_1, ..., x_n)$ **et** $m = E(X)$.
- **Si** X **est un vecteur gaussien et** A **une matrice non nulle, alors la transformation linéaire :** $Y = AX$ **est un vecteur gaussien de moyenne** $AE(X)$ **et de matrice de variances-covariances** $\Gamma_Y = A \Gamma_X A'$.

Vecteurs Gaussiens

Preuve

Nous montrons seulement le premier point.

- **Par définition du vecteur gaussien,** $Z = \alpha' X$ **suit une loi normale dans** $\mathbb{R}$ **de paramètres :**

$$E\left(\alpha' X\right) = \alpha' E(X)$$

$$V\left(\alpha' X\right) = \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j Cov(X_i, X_j) = \alpha' \Gamma_X \alpha.$$

Vecteurs Gaussiens

Proposition

- **Si X est un vecteur gaussien alors tous les v.a $X_1,...,X_n$ suivent des lois normales dans $\mathbb{R}$.**
- **Inversement, si $X_1,...,X_n$ sont indépendantes et suivent des lois normales dans $\mathbb{R}$, alors $X = (X_1, ..., X_n)$ est un vecteur gaussien. Sa matrice de variances-covariances est diagonale, c-à-d :** $a_{ij} = Cov(X_i, X_j) = 0$ **pour tous** $i \neq j$.

Vecteurs Gaussiens

Exemple

Si X **et** Y **sont deux v.a i.i.d de loi** $\mathcal{N}(0,1)$ **alors** $E(X) = E(Y) = 0$ **et** $V(X) = V(Y) = 1$ **et par suite** (X, Y) **est un vecteur gaussien de moyenne** $m = (0,0)$ **et de matrice de variances-covariances :**

$$\Gamma_X = I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

La densité du couple (X, Y) **est donnée par :**

$$f_{(X,Y)}(x,y) = \frac{1}{2\pi} \exp\left(-\frac{x^2+y^2}{2}\right) = f_X(x) f_Y(y).$$

Plan :

Fonction caractéristique

Définition

Soit $X = (X_1, ..., X_n)$ **un vecteur aléatoire à valeurs dans** $\mathbb{R}^n$**. La fonction caractéristique de** X **est définie pour tout** $t = (t_1, ..., t_n) \in \mathbb{R}^n$ **par :**

$$\varphi_X(t) = E(e^{it'X}) = E(e^{i(t_1 X_1 + ... + t_n X_n)}).$$

Fonction caractéristique

Propriétés

1. **Si on connaît la fonction caractéristique du vecteur X, on peut en déduire la fonction caractéristique de chaque composante X_i. On a :**

$$\forall\ i \in [1:n], \forall\ t_i \in \mathbb{R}:\ \varphi_{X_i}(t_i) = \varphi_X(0, ..., t_i, ..., 0).$$

2. **Si $X_1, ..., X_n$ sont indépendantes, alors pour tout $t = (t_1, ..., t_n) \in \mathbb{R}^n$, on a :**

$$\varphi_{X_1 + ... + X_n}(t) = \prod_{i=1}^{d} \varphi_{X_i}(t_i).$$

3. **Les v.a.r $X_1, ..., X_n$ sont indépendantes ssi pour tout $t = (t_1, ..., t_n) \in \mathbb{R}^n$, on a :**

$$\varphi_X(t) = \prod_{i=1}^{n} \varphi_{X_i}(t_i).$$

Fonction caractéristique

Proposition

$X = (X_1, ..., X_n)$ est vecteur gaussien dans $\mathbb{R}^n$ ssi il existe un vecteur m de $\mathbb{R}^n$ et une matrice Γ_X symétrique et positive d'ordre n tels que sa fonction caractéristique est donnée pour tout $t \in \mathbb{R}^n$ par :

$$\varphi_X(t) = \exp\left\{it'm - \frac{1}{2}t'\Gamma_X t\right\}.$$

Dans ce cas $E(X) = m$ et Γ_X est la matrice de variances-covariances de X.

Fonction caractéristique

Preuve

Sens direct :

Si $X = (X_1, ..., X_n)$ est un vecteur gaussien dans $\mathbb{R}^n$, alors $t'X$ est un v.a normale dans $\mathbb{R}$, de moyenne $t'm$ et de variance $t'\Gamma_X t$. Par suite : pour tout $u \in \mathbb{R}$, on a

$$\varphi_{t'X}(u) = E\left(e^{iut'X}\right) = exp\left\{iut'm - \frac{1}{2}u^2 t'\Gamma_X t\right\}.$$

Prenons $u = 1$, on trouve

$$\varphi_X(t) = E\left(e^{it'X}\right) = \exp\left\{it'm - \frac{1}{2}t'\Gamma_X t\right\}.$$

Fonction caractéristique

Sens inverse :

Supposons maintenant que $X = (X_1, ..., X_n)$ est un vecteur aléatoire dans $\mathbb{R}^n$ de fonction caractéristique :

$$\varphi_X(t) = \exp\left\{it'm - \frac{1}{2}t'\Gamma_X t\right\}.$$

Soit $Y = t'X$, alors

$$\varphi_Y(u) = E\left(e^{iuY}\right) = E\left(e^{iut'X}\right) = \exp\left\{it'm - \frac{1}{2}u^2 t'\Gamma_X t\right\} = \exp\left\{ia - \frac{1}{2}u^2 b\right\},$$

avec $a = t'm$ et $b = t'\Gamma_X t$. Par caractérisation, la v.a Y qui une combinaison linéaire de X est de loi normale dans $\mathbb{R}$. Ceci implique que $X = (X_1, ..., X_n)$ est un vecteur gaussien dans $\mathbb{R}^n$.

Vecteurs Gaussiens

Corollaire

Les composantes $X_1,\dots,X_n$ du vecteur gaussien X sont indépendantes ssi la matrice de variances-covariances de X est diagonale.

Preuve

Il suffit, bien sûr, de montrer la réciproque. Supposons donc que Γ_X soit diagonale telle que $V(X_i) = \sigma_i^2$, alors sa fonction caractéristique est de la forme :

$$\varphi_X(t) = \exp\left\{\sum_{i=1}^{n}\left(it_i m_i - \frac{1}{2}t_i^2\sigma_i^2\right)\right\} = \prod_{i=1}^{n}\varphi_{X_i}(t_i).$$

D'où le résultat.

Conséquence

Si (X,Y) est un vecteur gaussien, alors $X \perp Y$ ssi $Cov(X,Y) = 0$.

Plan :

Introduction

Dans cette section, nous allons introduire différentes notions de convergences pour une suite de v.a. Après avoir étudié les liens entre ces différentes notions de convergences, nous énoncerons deux théorèmes limites très importants des probabilités :

- La loi forte des grands nombres qui nous donne la convergence de la moyenne empirique $\overline{X_n} = \frac{1}{n} \sum_{i=1}^{n} X_i$ d'une suite $(X_n)_n$ de v.a.r i.i.d et intégrables vers $E(X_1)$ lorsque $n \to +\infty$.
- Le théorème de la limite centrale qui indique à quelle vitesse cette convergence a lieu sous l'hypothèse supplémentaire que les $(X_n)_n$ sont de carré intégrable.

Convergences et approximations

Soit $(X_n)_n$ une suite de v.a définies sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ et X une v.a définie sur le même espace.
Les modes de convergence les plus utilisés en probabilité sont les suivants :

- **la convergence presque sûre notée p.s,**
- **la convergence en probabilité notée p,**
- **la convergence en moyenne quadratique (ou dans L^2) notée m.q,**
- **la convergence dans L^1,**
- **la convergence en loi notée $\mathcal{L}$.**

Convergences et approximations

Définition

Lorsque $n \to +\infty$, on dit que :

1. $X_n \stackrel{p.s}{\to} X$ **si** $P\left(\left\{\omega, \lim_{n\to+\infty} X_n(\omega) = X(\omega)\right\}\right) = 1.$
2. $X_n \stackrel{p}{\to} X$ **si** $\forall\, \epsilon > 0, \quad \lim_{n\to+\infty} P(|X_n - X| > \epsilon) = 0.$
3. $X_n \stackrel{mq}{\to} X$ **si** $(X_n)_n$ **et** X **sont de carré intégrables et** $\lim_{n\to+\infty} E(|X_n - X|^2) = 0.$
4. $X_n \stackrel{L^1}{\to} X$ **si** $(X_n)_n$ **et** X **sont intégrables et** $\lim_{n\to+\infty} E(|X_n - X|) = 0.$
5. $X_n \stackrel{\mathcal{L}}{\to} X$ **si pour toute** f **continue et bornée,** $\lim_{n\to+\infty} E(f(X_n)) = E(f(X)).$

Convergences et approximations

Exemple :

Considérons la suite de v.a $(X_n)_n$ où chaque X_n prenant deux valeurs 0 et $\sqrt{n}$ avec les probabilités suivantes :

$$P(X_n = 0) = 1 - \frac{1}{n} \qquad P(X_n = \sqrt{n}) = \frac{1}{n}.$$

Alors lorsque $n \to +\infty$, on a :

1. $X_n \xrightarrow{L^1} 0$, **car** $E|X_n| = \frac{1}{\sqrt{n}} \longrightarrow 0.$
2. $X_n \xrightarrow{p} 0$ **car d'après l'inégalité de Markov on a :**

$$P(|X_n| > \epsilon) \leq \frac{E|X_n|}{\epsilon} = \frac{1}{\epsilon\sqrt{n}} \longrightarrow 0.$$

Convergences et approximations

Proposition

On note $(F_{X_n})_n$ la suite des fonctions de répartitions de $(X_n)_n$ et F_X de X.

- $X_n \xrightarrow{\mathcal{L}} X$ **ssi** $\lim\limits_{n\to+\infty} F_{X_n}(x) = F_X(x)$ **en tout point de continuité x de F.**

Convergences et approximations

Remarque

- **Dans le cas des v.a.d :** $X_n \xrightarrow{\mathcal{L}} X$ **ssi** $\lim\limits_{n \to +\infty} P(X_n = x) = P(X = x)$**.**
- **Dans le cas des v.a à densités :** $X_n \xrightarrow{\mathcal{L}} X$ **ssi** $\lim\limits_{n \to +\infty} f_{X_n}(x) = f_X(x)$**.**

Convergences et approximations

Exemple :

Considérons la suite de v.a $(X_n)_n$ où chaque X_n ait pour loi :

$$P\left(X_n = 2 + \frac{1}{n}\right) = 1 \quad \textbf{c-à-d} \quad P_{X_n} = \delta_{2+\frac{1}{n}}.$$

Puisque $2 + \frac{1}{n}$ converge vers 2, alors : $X_n \xrightarrow{\mathcal{L}} X$ où X est de loi δ_2 et on a $\lim\limits_{n \to +\infty} F_{X_n}(x) = F_X(x)$ en tout point $x \in \mathbb{R} \setminus \{2\}$.

Théorème de Paul Lévy

Théorème de Paul Lévy

On note $(\varphi_{X_n})_n$ la suite des fonctions caractéristiques de $(X_n)_n$ et φ_X de X.

- Si $X_n \xrightarrow{\mathcal{L}} X$, alors φ_{X_n} converge simplement vers φ_X.
- Inversement, si φ_{X_n} converge simplement vers une fonction φ continue en 0, alors $X_n \xrightarrow{\mathcal{L}} X$ dont la fonction caractéristique de X est φ.

Théorème de Slutsky

Théorème de Slutsky

Si $X_n \xrightarrow{\mathcal{L}} X$ et g est une application continue de $\mathbb{R}$ vers $\mathbb{R}$, alors $g(X_n) \xrightarrow{\mathcal{L}} g(X)$.

Convergences et approximations

Le résultat suivant est une simple application du lemme de Borel-Cantelli qui montre que si $(\Omega, \mathcal{A}, P)$ est un espace probabilisé et $(A_n)_n$ une suite des événements, alors

$$\sum_n P(A_n) < +\infty \Longrightarrow P(\limsup A_n) = 0.$$

Proposition

1. **Si la suite $(X_n)_n$ converge p.s vers X, alors il converge en probabilité vers X.**
2. **Inversement, si la suite $(X_n)_n$ converge en probabilité vers X, alors il existe une sous suite $(X_{n_k})_k$ qui converge p.s vers X.**

Convergences et approximations

Proposition :

L^2 :
$\mathbb{E}(|X_n - X|^2) \to 0$

$\Downarrow$

L^1 :
$\mathbb{E}|X_n - X| \to 0$

presque sûre :
$\mathbb{P}(X_n \to X) = 1$

$\searrow \qquad \swarrow$

Proba : $\forall \varepsilon > 0$,
$\mathbb{P}(|X_n - X| \geq \varepsilon) \to 0$

$\Downarrow$

Loi : $\forall f$ continue bornée,
$\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$

Convergences et approximations

Approximation de la loi de Poisson par la loi binomiale

Proposition

Soit $(X_n)_n$ **une suite de v.a de loi** $\mathcal{B}(n, p_n)$ **telles que** $np_n \to \lambda > 0$, **lorsque** $n \to +\infty$, **alors** $X_n \xrightarrow{\mathcal{L}} X$ **qui suit une loi de poisson** $\mathcal{P}(\lambda)$.

Convergences et approximations

Preuve

Il suffit de montrer que lorsque $n \to +\infty$:

$\varphi_{X_n}(t) \longrightarrow e^{\lambda(e^{it}-1)} = \varphi_X(t), \quad \forall\, t \in \mathbb{R}.$

$$\varphi_{X_n}(t) = \left(1 - p_n + p_n e^{it}\right)^n = \left(1 - \frac{np_n(1-e^{it})}{n}\right)^n$$

$$= e^{n \ln\left(1 - \frac{np_n(1-e^{it})}{n}\right)} \longrightarrow e^{\lambda(e^{it}-1)} = \varphi_X(t).$$

Convergences et approximations

Exemple

Soit X **une v.a de loi** $\mathcal{B}(100; 0,05)$. **Nous allons calculer :** $P(X=2)$.

- **Calcul exact :**

$$P(X=2) = C_{100}^{2}(0.05)^2(1-0.05)^{100-2} \approx 0,0812.$$

- **Calcul approché : On approche** $\mathcal{B}(100; 0,05)$ **par** $\mathcal{P}(\lambda=5)$, $(np_n = 100 \times 0,05 = 5)$.

$$P(X=2) \approx \frac{5^2}{2!}e^{-5} \approx 0,0843.$$

Lois des grands nombres

On rappelle le résultat suivant.

Inégalité de Bienaymé-Tchebychev

Soit X une v.a réelle définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ possédant une espérance $E(X)$ et une variance $V(X)$, alors

$$\forall \epsilon > 0, \qquad P\Big(|X - E(X)| \geq \epsilon\Big) \leq \frac{V(X)}{\epsilon^2}.$$

Loi faible des grands nombres

Théorème

Soit $(X_n)_n$ **une suite de v.a i.i.d et définies sur un espace probabilisé** $(\Omega, \mathcal{F}, P)$, **possédant une espérance** $m = E(X_1)$ **et une variance** $\sigma^2 = V(X_1)$. **On pose :** $\overline{X_n} = \frac{1}{n}\sum_{i=1}^{n} X_i$. **Alors** $\overline{X_n} \xrightarrow{p} m$, **c-à-d** $\forall \epsilon > 0 \quad \lim_{n \to +\infty} P(|\overline{X_n} - m| \geq \epsilon) = 0$.

Loi faible des grands nombres

Preuve

On a $E(\overline{X_n}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n} nm = m$ **et comme les v.a sont indépendants, on trouve que** $V(\overline{X_n}) = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n} n\sigma^2 = \frac{\sigma^2}{n}$.

D'après l'inégalité de Bienaymé-Tchebychev, on a

$$P\left(|\overline{X_n} - m| \geq \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2}.$$

Donc par encadrement de limits, $\lim_{n\to+\infty} P(|\overline{X_n} - m| \geq \epsilon) = 0$.

Convergences et approximations

Exemple

On considère une suite d'épreuves indépendantes, et un événement A, de probabilité p, qui peut ou non se réaliser au cours d'une des épreuves. On note X_i la v.a qui vaut 1 si A s'est réalisé à la i-ème épreuve et 0 sinon. Dans ce cas $\overline{X_n}$ représente la fréquence de réalisation de l'événement A au cours des n premières épreuves. La loi faible des grands nombres nous dit que cette fréquence (tend) vers p. Cela justifie a posteriori la façon d'introduire la notion de probabilité comme étant la limite de la fréquence d'apparition de l'événement donné.

Loi forte des grands nombres

Théorème

Soit $(X_n)_n$ **une suite de v.a réelles i.i.d définies sur un espace probabilisé** $(\Omega, \mathcal{F}, P)$**, possédant une espérance** $m = E(X_1)$ **et une variance** $\sigma^2 = V(X_1)$**.**
Alors

$$\overline{X_n} \overset{p.s}{\to} m.$$

Plan :

Approximation d'une fonction de répartition

Proposition

On définit pour n v.a i.i.d $X_1, \ldots, X_n$, la fonction de répartition empirique F_n définie pour tout $x \in \mathbb{R}$ par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{(X_i \leq x)}.$$

Alors, lorsque n tend vers l'infini, $F_n(x) \overset{p.s}{\to} F(x)$ qui est la fonction de répartition de X_1.

Approximation d'une fonction de répartition

Preuve

Ce résultat est une simple application de la lois forte des grands nombre, il suffit de prendre : $Y_i = 1_{(X_i \leq x)}$**, par suite**

$m = E(Y_1) = E(1_{(X_1 \leq x)}) = P(X_1 \leq x) = F(x).$

Méthodes de Monte Carlo

Les méthodes de Monte Carlo permettent d'estimer des intégrales à l'aide de la loi forte des grands nombres. Le principe de fonctionnement de ces méthodes est le suivant : Considérons une intégrale de la forme : $I = \int_{\mathbb{R}} g(x)dx$**, qu'on veut estimer sa valeur inconnue. Alors**

$$I = \int_{\mathbb{R}} \frac{g(x)}{f_X(x)} f_X(x)dx = \int_{\mathbb{R}} h(x)f_X(x)dx = E(h(X)),$$

où la fonction f_X **est supposée être une densité de probabilité de** X**.**

En général, une telle intégrale n'est pas calculable explicitement car on ne connaît pas de primitive de la fonction intégrée. Les méthodes de Monte Carlo proposent d'estimer I par la quantité :

$$I_n = \frac{1}{n} \sum_{i=1}^{n} h(X_i),$$

pour n v.a i.i.d $X_1, ..., X_n$ suivant la loi de densité f_X, (n assez grand).

En pratique on prend par exemple :

- $X \sim \mathcal{U}(0,1)$ **pour** $\int_0^1 g(x)dx$.
- $X \sim \mathcal{U}(a,b)$ **pour** $\int_a^b g(x)dx$.
- $X \sim \mathcal{E}(1)$ **pour** $\int_0^{+\infty} g(x)dx$.
- $X \sim \mathcal{N}(0,1)$ **pour** $\int_{-\infty}^{+\infty} g(x)dx$.
- **Pour** $\int_a^{+\infty} g(x)dx$, $a > 0$, **on fait le changement de variable** $t = \frac{a}{x}$ **pour retrouver** $\int_0^1 g(x)dx$.

Plan :

Exemple sous logiciel R

On veut estimer la valeur de l'intégrale :

$$I = \int_0^2 x^2 dx = \frac{8}{3} \simeq \mathbf{2.666667}.$$

$>$ $x = runif(5000, 0, 2)$
$>$ $y = x^2$
$>$ $2 * mean(y)$
$>$ **2.599169**

Que pensez-vous de la qualité de la valeur obtenue comparée avec la valeur exacte ?

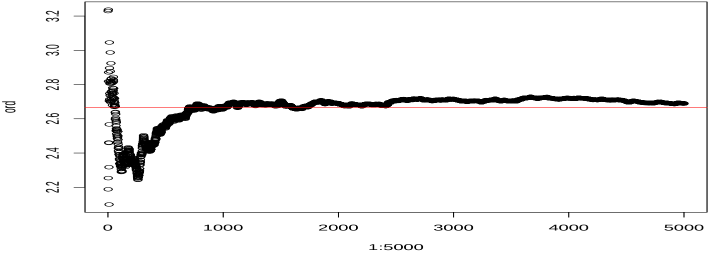
Exemple sous logiciel R

Pour observer l'évolution de la qualité de cette estimation lorsque n varie.

$> \quad ord = cumsum(y)/(1:5000)$

$> \quad plot(1:5000, ord)$

$> \quad abline(h = 8/3, col = "red")$



Plan :

Théorème de la limite centrale

La vitesse de convergence des méthodes de Monte Carlo est assurée par le théorème suivant.

Théorème

Soit $(X_n)_n$ **une suite de v.a i.i.d définies sur un espace probabilisé** $(\Omega, \mathcal{F}, P)$ **possédant une espérance** $m = E(X_1)$ **et une variance** $\sigma^2 = V(X_1)$**. Alors lorsque** $n \to +\infty$

$$Z_n = \frac{\overline{X_n} - E(\overline{X_n})}{\sigma(\overline{X_n})} = \frac{\sqrt{n}(\overline{X_n} - m)}{\sigma} \xrightarrow{\mathcal{L}} X \sim \mathcal{N}(0,1).$$

Théorème de la limite centrale

Preuve

D'après la Si Y est une v.a centré réduite, alors la fonction caractéristique de Y admet le développement limité suivant :

$$\varphi_Y(t) = 1 - \frac{t^2}{2} + o(t^2), \qquad t \to 0.$$

Posons $Y_i = \frac{X_i - m}{\sigma}$. Les $Y_1, \dots, Y_n$ sont n v.a i.i.d centrés réduites telles que

$$Z_n = \sum_{i=1}^{n} \frac{Y_i}{\sqrt{n}}.$$

$$\varphi_{Z_n}(t) = \left[\varphi_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right]^n \to e^{\frac{-t^2}{2}}, \quad \textbf{lorsque} \ \ n \to +\infty.$$

Mais cette limite est la fonction caractéristique de la loi normale centrée réduite $\mathcal{N}(0,1)$, d'où l'on déduit le théorème de la limite centrale grâce au théorème de convergence de Paul Lévy.

Théorème de la limite centrale

Approximation de la loi binomiale par la loi normale

Théorème de de Moivre-Laplace

Soit $(X_n)_n$ **une suite de v.a de loi** $\mathcal{B}(n,p)$. **Alors**

$$Z_n = \frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} X \sim \mathcal{N}(0,1).$$

Théorème de la limite centrale

Preuve

Ce résultat est une simple application du théorème de la limite centrale. Il suffit de prendre : $X_n = \sum_{i=1}^{n} Y_i$ **la somme de** n **v.a i.i.d de loi** $\mathcal{B}(p)$**. On a :** $E(Y_1) = p,\ V(Y_1) = p(1-p)$, **donc** $X_n = n\overline{Y_n},\ E(X_n) = np$ **et** $V(X_n) = np(1-p)$.