Correction du TD 8 du proba: G8

Rédacteur / Correcteur

Bourrous Hani Anouar / Ali Kohen

Question de cours

- Rappeler le principe de la méthode de Monte-Carlo.

Soient $X_1, X_2, \ldots, X_n$ des variables indépendantes, de même loi et de carré intégrable ayant pour densité f_X. Soit ψ une fonction définie sur $[a,b]$ telle que l'intégrale I existe.

- Nous avons $\frac{1}{n}\sum_{i=1}^{n}\psi(X_i) \xrightarrow[n\to\infty]{} I = \int_a^b \psi(x)f(x)dx$

Exercice 1

Soient $(U_n)_n$ et $(V_n)_n$ deux suites de variables aléatoires suivant la loi uniforme sur l'intervalle $[0,1]$. On suppose que ces variables aléatoires sont indépendantes dans leur ensemble. On pose

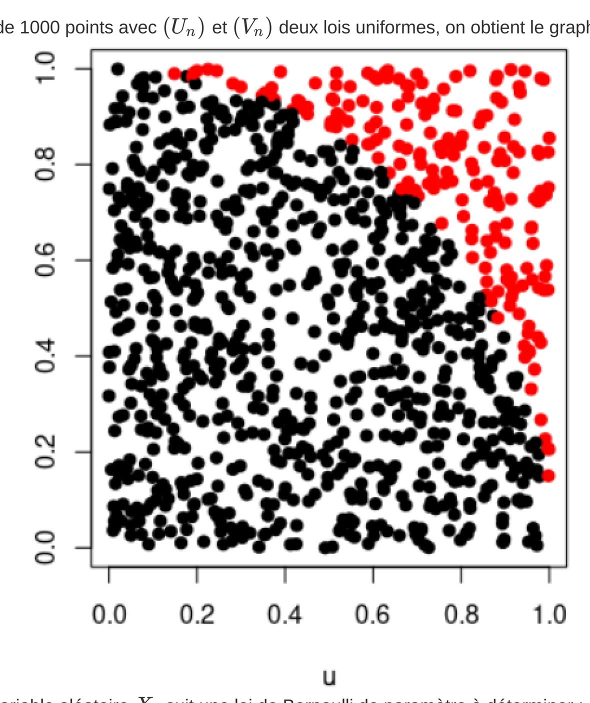
$$\forall n \geq 1, X_n = \begin{cases} 1 & \text{si } U_n^2 + V_n^2 \leq 1 \\ 0 & \text{sinon} \end{cases}$$

et $Z_n = 4 \times \frac{(X_1 + \cdots + X_n)}{n}$.

Question 1

- Déterminer la loi de la variable X_n.
- Calculer la variance de Z_n et montrer que la suite (Z_n) converge vers π.

Lorsque l'on fait la simulation du tirage de 1000 points avec (U_n) et (V_n) deux lois uniformes, on obtient le graphique suivant :



Pour tout $i \in \mathbb{N}^*$, on remarque que la variable aléatoire X_i suit une loi de Bernoulli de paramètre à déterminer :

Soit $n \geq 1$,

On considère la famille suivante $(U_n^2 = u)$ qui forme une famille totale, donc par la formule de probabilités totales dans le cas continue :

$$\begin{aligned} P(X_n = 1) &= P(U_n^2 + V_n^2 \leq 1) \\ &= \int_0^1 P(U_n^2 + V_n^2 \leq 1 \mid U_n^2 = u) f_U(u) du \\ &= \int_0^1 P(V_n^2 \leq 1 - u^2) du \\ &= \int_0^1 \sqrt{1-u^2} du \end{aligned}$$

On effectue le changement de variable $u = \cos(v)$ donc $du = -\sin(v)dv$

$$\begin{aligned} P(X_n = 1) &= \int_{\frac{\pi}{2}}^{0} -\sin(v)\sqrt{1-\cos^2(v)} dv \\ &= \int_0^{\frac{\pi}{2}} \sin^2(v) dv \\ &= \int_0^{\frac{\pi}{2}} \frac{1 - cos(2v)}{2} dv \\ &= \frac{\pi}{4} \end{aligned}$$

Pour tout $i \in \mathbb{N}^*$, on a $X_i \rightsquigarrow B(\frac{\pi}{4})$.

$\longmapsto$ Ainsi, par définition de Z_n, on a :

$\forall n \in \mathbb{N}^*$, $\frac{n}{4} Z_n \rightsquigarrow B(n, \frac{\pi}{4})$ et $\frac{\pi}{4}$

Grâce au cours on a la variance d'une loi Binomiale :

$$\begin{aligned} Var(Z_n) &= Var(\frac{n}{4} \times \frac{4}{n} Z_n) \\ &= \frac{16}{n^2} Var(\frac{n}{4} Z_n) \\ &= \frac{16}{n^2} \times n\frac{\pi}{4}(1 - \frac{\pi}{4}) \\ &= \frac{\pi(4-\pi)}{n} \end{aligned}$$

On va calculer l'espérance de la Z_n :

$$\begin{aligned} \mathbb{E}(Z_n) &= \mathbb{E}(\frac{n}{4} \times \frac{4}{n} Z_n) \\ &= \frac{4}{n}\mathbb{E}(\frac{n}{4} Z_n) \\ &= \frac{4}{n} \times n\frac{\pi}{4} \\ &= \pi \end{aligned}$$

Puis, d'après l'inégalité de Bienaymé-Tchebychev, on a :

$$\begin{aligned} P(|Z_n - \mathbb{E}(Z_n)| \geq \epsilon) &\leq \frac{V(Z_n)}{\epsilon^2} \\ &\leq \frac{\pi(4-\pi)}{n\epsilon^2} \end{aligned}$$

Or, $\lim_{n\to\infty} \frac{\pi(4-\pi)}{n\epsilon^2} = 0$ donc d'après la loi faible des grands nombres, (Zn) converge

Question 2

Soit $\alpha \in (0,1)$ et $\epsilon > 0$.

- A l'aide de l'inégalité de Chebishev, déterminer un entier n_0 tel que

$$\forall n \geq n_0, \quad \mathrm{P}(|Z_n - \pi| > \epsilon) \leq \alpha$$

- Ecrire un algorithme qui retourne une valeur approchée de π à 10^{-4} près, avec une probabilité supérieure à 0.95.

D'après l'inégalité de Tchebychev, on a :

$$\forall \epsilon > 0$$

$$P(|Y - \mathbb{E}[Y]| \geq \epsilon) \leq \frac{V[Y]}{\epsilon^2}$$

Soit $n \geq 1$,$Y = Z_n$ on obtient $\forall \epsilon > 0$

$$P(|Z_n - \mathbb{E}[Z_n]| \geq \epsilon) \leq \frac{V[Z_n]}{\epsilon^2} \Leftrightarrow P(|Z_n - \pi| \geq \epsilon) \leq \frac{V[Z_n]}{\epsilon^2}$$

Après avoir términer le premier bloc de la question,posant $\alpha = \frac{Var[Z_n]}{\epsilon^2}$ on obtient par équivalences succésives n$n_0 = \frac{V[X]}{\alpha\epsilon^2}$ On cherche une approximation de π à 10^{-4}.

D'après le calcul précédent, on a $\alpha = 0.05$ et $\epsilon = 10^{-4}$ on peut alors calculer n_0 :

$n_0 = \frac{\frac{\pi(4-\pi)}{n_0}}{0{,}05 * 10^{(-4)^2}} + 1$
$n_0^2 - n_0 - \frac{\pi(4-\pi)}{0{,}05 * 10^{(-4)^2}} = 0$

On se retrouve avec une équation du second degré que R peut résoudre seul avec les lignes suivantes :

```
a <- 0.05 #alpha
e <- 10^(-5) #epsilon
n <- (pi\*(4-pi)/n)/(a\*e^2) +1

u <- runif(n)
v <- runif(n)
4*mean(u^2 + v^2 < 1)
```

On obtient par exemple : `[1] 3.141632`

Question 3

Multipliant la variable Z_n par $\sqrt{n}$.

- Calculer la variance de la variable $\sqrt{n}(Z_n - \pi)$. Cette variance converge-t-elle vers 0 ? Vers une constante ?

On a $\sqrt{n}(Z_n - \pi)$:

$$\begin{aligned} Var(\sqrt{n}(Z_n - \pi)) &= n \times Var(Z_n) \\ &= \pi(4-\pi) \end{aligned}$$

Cette variance étant une constante, elle ne converge pas vers 0 mais vers une constante.

- Quelle loi connue fournit une bonne approximation de la loi de $\sqrt{n}(Z_n - \pi)$?

Z_n est une constante près, une somme de lois de Bernoulli indépendantes, On peut donc l'approximer par une loi Binomiale de paramètre $(n, \frac{\pi}{4})$ (d'après la loi des X_n)

Exercice 2

On considère une suite (U_n) de variables aléatoires indépendantes de loi uniforme sur $(0,1)$ et la fonction

$$\forall u \in (0,1), \quad \varphi(u) = \sqrt{(1-u)u^3}$$

Question 1

Pour tout $n \geq 1$, on pose

$$Y_n = \frac{1}{n}\sum_{i=1}^{n}\varphi(U_i)$$

- Montrer que la suite Y_n converge, au sens de la loi des grands nombres, vers la limite $\mathcal{I}$ définie ci-dessous

$$\mathcal{I} = \int_0^1 \varphi(u) du$$

On admettra que $\mathcal{I} = \frac{\pi}{16}$.

$\forall n, U_n \sim \mathcal{U}(0,1)$ donc $f_{U_n}(u) = 1$, ainsi par le théorème de transfert , $\mathcal{I} = \mathbb{E}[\varphi(U)]$.

D'après la loi forte des grands nombres, on a

$$\lim_{n\to\infty} Y_n = \mathbb{E}[\varphi(U)] = \mathcal{I}$$

Question 2

- Calculer la variance de la variable aléatoire Y_n.
- Soit $\epsilon = 10^{-3}$. Par l'inégalité de Chebyshev, donner une estimation du rang n à partir duquel on peut considérer que

$$\mathrm{P}(|Y_n - \mathcal{I}| < \epsilon) \geq 0.95$$

Y_n est une moyenne empirique de $\varphi(U)$. On a donc $V[Y_n] = \frac{V[\varphi(U)]}{n}$ .Par indépendance des V.a et le théorème de Koenig-Huygens on calcule cette variance.
Or par le théorème de transfert,

$$\begin{aligned} \mathbb{E}[\varphi(U)^2] &= \int_0^1 \varphi^2(u) du \\ &= \int_0^1 (1-u)u^3 du \\ &= [\frac{1}{4}u^4]_0^1 - [\frac{1}{5}u^5]_0^1 \\ &= (\frac{1}{4} - \frac{1}{5}) \\ &= \frac{1}{20} \end{aligned}$$

D'où

$$\begin{aligned} V[Y_n] &= \frac{V[\varphi(U)]}{n} \\ &= \frac{\mathbb{E}[\varphi(U)^2] - \mathbb{E}[\varphi(U)]^2}{n} \\ &= \frac{\frac{1}{20} - \mathcal{I}^2}{n} \\ &= \frac{\frac{1}{20} - \frac{\pi^2}{256}}{n} \end{aligned}$$

D'après l'inégalité de Tchebychev, on a $P(|Z_n - E[Z_n]| > \epsilon) \leq \frac{V[Z_n]}{\epsilon^2} = 0,95$

On estime ainsi $n = \frac{\frac{1}{20} - \frac{\pi^2}{256}}{0{,}95\epsilon^2} \approx 12050$

Question 3

On considère la loi f définie sur l'intervalle (0,1) de la manière suivante :

$\forall v \in (0,1)$, $\mathrm{f}(v) = 6v(1-v)$

Soit (V_n) une suite de variables aléatoires indépendantes, de loi de densité f. Pour tout n ≥ 1 , on pose

$$Z_n = \frac{1}{n}\sum_{i=1}^{n}\frac{\varphi(V_i)}{\mathrm{f}(V_i)}$$

- Montrer que la suite Z_n converge vers $\mathbb{I}$
- Comparer la variance de la variable aléatoire Z_n à celle de la variable Y_n.

On pose $\psi(x) = \frac{\varphi(x)}{\mathrm{f(x)}}$

$$Z_n = \frac{1}{n}\sum_{i=1}^{n}\psi(V_i)$$

D'après la loi forte des grands nombres , les $\psi(V_i)$ étant indépendantes et carrés intégrables :

$$\lim_{n\to+\infty} Z_n = \mathbb{E}[\psi(V_1)]$$

Or $\mathbb{E}[\psi(V_1)] = \int_0^1 \psi(x)f(x)dx = \int_0^1 \varphi(x)dx = I$

$\longrightarrow \lim_{n\to+\infty} Z_n = I$

$*$Comparons les variances :

D'après la quest 2 on a $Var(Y_n) = \frac{1}{n}(\frac{1}{20} - I^2)$

$Var(Z_n) = \frac{Var(\psi(V_i))}{n} = \frac{\mathbb{E}[\psi(V_i)^2] - \mathbb{E}[\psi(V_i)]^2}{n} = \frac{E[\psi(V_i)^2] - I^2}{n}$

Or on a $\mathbb{E}[\psi(V_i)^2] = \int_0^1 \psi(x)^2 f(x)dx = \int_0^1 \frac{\psi(x)^2}{f(x)} = \int_0^1 \frac{x^3}{6}dx = \frac{1}{18}$

D'où $Var(Z_n) = \frac{1}{n} \times (\frac{1}{18} - I^2)$

Ainsi :

$$Var(Y_n) < Var(Z_n)$$

Question 4

Le deuxième algorithme est moins efficace car $Var(Y_n) = 0,0114417 < Var(Z_n) = 0,0169908$ l'algorithme est plus précis pour une variable aléatoire de variance minimale".

Question 5

*On a $\int_{-\infty}^{+\infty} f_\alpha(x)dx = \int_0^1 c_\alpha x^\alpha (1-x)dx = c_\alpha(\frac{1}{\alpha+1} - \frac{1}{\alpha+2}) = c_\alpha \frac{1}{(\alpha+1)(\alpha+2)} = 1$

Alors :

$$c_\alpha = (\alpha+1)(\alpha+2)$$

*$Var(Z_n^\alpha)$ = $\frac{1}{n}[\mathbb{E}(\frac{\varphi(V_i)}{f_{\alpha(V_1)}})^2$ - $\mathbb{E}(\frac{\varphi(V_i)}{f_{\alpha(V_1)}}^2)]$

Donc :

$$Var(Z_n^\alpha) = \frac{1}{n}(\frac{1}{(\alpha+1)(\alpha+2)(3-\alpha+1)} - I^2)$$

L'algorithme est le plus précis lorsque $Var(Z_n^\alpha)$ est minimale. La valeur minimale est atteinte en $\alpha = 2$.

$$Var(Z_n^2) = \frac{1}{n}(\frac{1}{24} - I^2)$$

- Pour la question 1, le facteur multiplicatif de la variance était $\frac{1}{20} - (\frac{\pi}{16})^2 = 0,0011$ alors que celui de la quest 5 est $\frac{1}{24} - (\frac{\pi}{16})^2 = 0,0031$

La variance de la question 5 est inférieur à celle de la question 1 donc l'algorithme de la question 5 sera plus précis.