Computational Statistics

TP1 : Reminder on Markov Chain – Stochastic gradient descent

Exercise 1 : Box-Muller and Marsaglia-Bray algorithm

1. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une fonction mesurable bornée :

$$\mathbb{E}[h(X,Y)] = \mathbb{E}[h(R\cos\theta, R\sin\theta)]$$

$$= \int_0^{2\pi}\int_{\mathbb{R}^+} h(r\cos\theta, r\sin\theta)\, f_R(r)\, f_\theta(\theta)\, dr\, d\theta$$

$$= \int_0^{2\pi}\int_{\mathbb{R}^+} h(r\cos\theta, r\sin\theta)\, r e^{-\frac{r^2}{2}} \times \frac{1}{2\pi}\, dr\, d\theta$$

On effectue le changement de variable $\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$

$\varphi : \mathbb{R}^+ \times]0, 2\pi[\to \mathbb{R}^2$ est un $\mathcal{C}^1$-difféomorphisme, de jacobien :
$(r,\theta) \mapsto (r\cos\theta, r\sin\theta)$

$$J_\varphi(r,\theta) = \begin{vmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{vmatrix} = r\cos^2\theta + r\sin^2\theta = r$$

Alors $\mathbb{E}[h(X,Y)] = \frac{1}{2\pi}\int_{\mathbb{R}^2} h(x,y) \det J_{\varphi^{-1}}(\cancel{x},y)\, \cancel{r}\, e^{-\frac{x^2+y^2}{2}}\, dx\, dy$

$$\boxed{\mathbb{E}[h(X,Y)] = \frac{1}{2\pi}\int_{\mathbb{R}^2} h(x,y)\, e^{-\frac{x^2+y^2}{2}}\, dx\, dy}$$

Ainsi, d'après la méthode de la fonction muette, X et Y sont indépendantes et ont une distribution $\mathcal{N}(0,1)$.

2. Soit F_R la fonction de répartition de la variable aléatoire de Rayleigh R.

$$F_R(x) = \int_0^x f_R(r)\, dr$$

$$= \int_0^x r e^{-\frac{r^2}{2}} \mathbb{1}_{\mathbb{R}^+}(r)\, dr$$

$$F_R(x) = \left[-e^{-\frac{x^2}{2}} \right]_0^x$$

$$\boxed{F_R(x) = 1 - e^{-\frac{x^2}{2}}}$$

On inverse la fonction de répartition :

$$y = 1 - e^{-\frac{x^2}{2}} \iff -\frac{x^2}{2} = \log(1-y)$$
$$\iff x = \sqrt{-2\log(1-y)}$$

Alors : $\boxed{F_R^{-1}(x) = \sqrt{-2\log(1-x)}}$

Etant donné une variable U suivant une loi uniforme sur $[0,1]$:

$$F_R(x) = \mathbb{P}(R \leq x) = \mathbb{P}(F_R(U) \leq x) = \mathbb{P}(U \leq F_r^{-1}(x))$$

On peut donc sampler la variable aléatoire R à partir d'une variable aléatoire U de loi uniforme sur $[0,1]$. On peut également sampler une variable aléatoire $\theta = 2\pi U'$ où U' suit une loi uniforme sur $[0,1]$.

$X = R\cos\theta$ et $Y = R\sin\theta$ sont, d'après la question 1, indépendantes et suivent une loi normale $\mathcal{N}(0,1)$

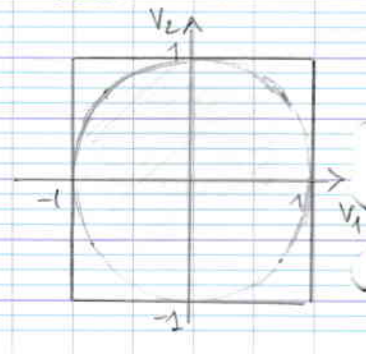
<u>Algorithme</u> :

take $U \sim \mathcal{U}([0,1])$
take $\theta \sim \mathcal{U}([0,2\pi])$
$R = \sqrt{-2\log(1-U)}$
$X = R\cos\theta$, $Y = R\sin\theta$
return X, Y.

3a) Soient $V_1^{(k)}, V_2^{(k)}$ les étapes intermédiaires de la boucle while.

$V_1^{(k)} \sim \mathcal{U}([-1,1])$
$V_2^{(k)} \sim \mathcal{U}([-1,1])$

On sort de la boucle si $V_1^{(k)2} + V_2^{(k)2} = 1$



Donc : $(V_1, V_2) \sim \mathcal{U}(\mathcal{B}(0,1))$

b) Soit T la variable aléatoire telle que $V_1^{(T)}, V_2^{(T)}$ terminent la boucle.

$$\mathbb{P}(T=m) = \mathbb{P}\left(V_1^{(1)2} + V_2^{(1)2} \geq 1\right) \times \mathbb{P}\left(V_1^{(2)2} + V_2^{(2)2} \geq 1\right) \times \dots$$
$$\times\ \mathbb{P}\left(V_1^{(T)2} + V_2^{(T)2} < 1\right).$$

$$\mathbb{P}(T=m) = \left(1 - P_{accept}\right)^{m-1} P_{accept}$$

où $P_{accept} = \dfrac{\mathcal{A}(\mathcal{B}(0,1))}{\mathcal{A}([-1,1]^2)} = \dfrac{\pi}{4}$

donc $\quad \mathbb{P}(T=m) = \dfrac{\pi}{4}\left(1 - \dfrac{\pi}{4}\right)^{m-1}$

Puis, $\quad \mathbb{E}[T] = \displaystyle\sum_{m=1}^{+\infty} m\,\mathbb{P}(T=m)$

$$= \sum_{m=1}^{+\infty} m\left(1 - \frac{\pi}{4}\right)^{m-1} \frac{\pi}{4}$$

$$\boxed{\mathbb{E}[T] = \frac{4}{\pi}}$$

Le nombre moyen d'étapes dans la boucle est donc de $\frac{4}{\pi}$.

c) Soit $h : \mathbb{R}^2 \to \mathbb{R}^2$ une fonction mesurable bornée :

$$\mathbb{E}\left[h(T_1, V)\right] = \int_{\mathbb{R}^2} h\left(\frac{v_1}{\sqrt{v_1^2+v_2^2}}, v_1^2+v_2^2\right) \times \frac{1}{4} \mathbb{1}_{\{v_1^2+v_2^2<1\}}(v_1, v_2)\, dv_1 d$$

$$\varphi : \mathbb{R}^2 \to \mathbb{R}^2$$
$$(v_1, v_2) \mapsto \left(\frac{v_1}{\sqrt{v_1^2+v_2^2}}, v_1^2+v_2^2\right)$$

On pose $\quad t = \dfrac{v_1}{\sqrt{v_1^2+v_2^2}}, \; v = v_1^2+v_2^2$

Alors : $\quad v_1 = t\sqrt{v} \quad$ et $\quad v_2 = \pm\sqrt{v(1-t^2)}$

La disjonction de v_2 empêche la bijectivité de φ.

On définit alors $D+ = \{(v_1, v_2) \mid v_1^2 + v_2^2 < 1 \text{ et } v_2 \geq 0\}$
$D- = \{(v_1, v_2) \mid v_1^2 + v_2^2 < 1 \text{ et } v_2 < 0\}$

φ induit une bijection de $D+$ vers $\varphi(D+)$ et de $D-$ vers $\varphi(D-)$ et :

$$\varphi^{-1}_{D+}(t, v) = \left(t\sqrt{v}, \sqrt{v(1-t^2)}\right)$$
$$\varphi^{-1}_{D-}(t, v) = \left(t\sqrt{v}, -\sqrt{v(1-t^2)}\right)$$

On calcule le jacobien de ces deux $\mathcal{C}^1$-difféomorphismes :

$$\left|J_{\varphi^{-1}_{D+}}(t,v)\right| = \begin{vmatrix} \sqrt{v} & \frac{t}{2\sqrt{v}} \\ -\frac{\cancel{2}t\sqrt{v}}{\cancel{2}\sqrt{1-t^2}} & \frac{\sqrt{1-t^2}}{2\sqrt{v}} \end{vmatrix} = \frac{\sqrt{1-t^2}}{2} + \frac{t^2}{2\sqrt{1-t^2}} = \frac{1}{2\sqrt{1-t^2}}$$

et $\left|J_{\varphi^{-1}_{D-}}(t,v)\right| = \frac{1}{2\sqrt{1-t^2}}$

Alors $$\mathbb{E}\left[h(T_1, V)\right] = \int_{D^-} h(\varphi_{D-}(v_1, v_2)) \frac{1}{\pi} dv_1 dv_2 + \int_{D^+} h(\varphi_{D+}(v_1, v_2)) \frac{1}{\pi} dv_1 dv_2$$
$$= \int_{\varphi(D^-)} h(t,v) \frac{1}{\pi} \frac{1}{2\sqrt{1-t^2}} dt dv + \int_{\varphi(D^+)} h(t,v) \frac{1}{\pi} \frac{1}{2\sqrt{1-t^2}} dt dv.$$

Or, $\varphi_{D^-}(D^-) = \varphi_{D^+}(D^+) = [-1,1] \times [0,1] = \varphi(D)$.

d'où : $$\mathbb{E}\left[h(T_1, V)\right] = \int_{\varphi(D)} h(t,v) \frac{1}{\pi} \times \frac{1}{\sqrt{1-t^2}} dt dv$$

Donc, d'après la méthode de la fonction muette, on a bien une densité à variables séparées, donc T_1 et V sont indépendantes, et $V \sim \mathcal{U}([0,1])$
et la densité de T_1 est : $[-1,1] \to \mathbb{R}$
$t \mapsto \frac{1}{\pi\sqrt{1-t^2}}$

On pose $T_2 = \dfrac{V_2}{\sqrt{V_1^2+V_2^2}}$.

$$\mathbb{E}\left[h(T_1,T_2)\right] = \int_D h\left(\frac{v_1}{\sqrt{v_1^2+v_2^2}}, \frac{v_2}{\sqrt{v_1^2+v_2^2}}\right) \frac{1}{\pi}\, dv_1\, dv_2.$$

On effectue le changement de variable $\begin{cases} \cos\theta = \dfrac{v_1}{\sqrt{v_1^2+v_2^2}} \\ \sin\theta = \dfrac{v_2}{\sqrt{v_1^2+v_2^2}} \end{cases}$

$$\mathbb{E}\left[h(T_1,T_2)\right] = \int_0^{2\pi}\int_0^1 h(\cos\theta, \sin\theta)\, \frac{1}{\pi}\, dr\, d\theta$$

$$= \int_0^{2\pi} h(\cos\theta, \sin\theta)\, \frac{1}{2\pi}\, d\theta$$

donc (T_1,T_2) ont la même loi que $(\cos\theta, \sin\theta)$ où $\theta \sim \mathcal{U}([0,2\pi])$

d) On a : $S = \sqrt{-2\log V}$ et $X = S T_1$, $Y = S T_2$.

D'après la question 3c), V et T_1 sont indépendantes (de même avec V et T_2), alors ST_1 et ST_2 sont indépendantes.

Comme $T_1 = \cos\theta$, $T_2 = \sin\theta$, $\theta \sim \mathcal{U}([0,2\pi])$ et $S = \sqrt{-2\log V}$, $V \sim \mathcal{U}([0,1])$, alors comme démontré dans la question 2 :
X et Y suivent une loi normale $\mathcal{N}(0,1)$.

Exercice 2 : Invariant distribution

1. Soit $A \subset \mathbb{R}^+$

$$P(x,A) = \int_{A\cap[0,1]} P(x,y)\,dy$$

$$P(x,A) = \mathbb{1}_{[0,1]\setminus\{\frac{1}{k},\,k\in\mathbb{N}^*\}}(x) \int_{A\cap[0,1]} \mathbb{1}_{[0,1]}(y)\,dy$$
$$+ \mathbb{1}_{\{\frac{1}{k},\,k\in\mathbb{N}^*\}}(x) \int_{A\cap[0,1]} \left((1-x^2)\,\delta_{\frac{1}{k+1}}(y) + \mathbb{1}_{[0,1]}(y)\,x^2\right) dy$$

$$P(x,A) = \begin{cases} (1-x^2)\,\delta_{\frac{1}{k+1}}(A) + \int_{A\cap[0,1]} x^2\,dy & \text{si } x = \frac{1}{k} \\ \int_{A\cap[0,1]} dy & \text{sinon} \end{cases}$$

2. Soit π la distribution uniforme sur $[0,1]$

Soient $y \in [0,1]$:

$$\pi P(y) = \int_0^1 P(x,y)\,\pi(x)\,dx$$
$$= \int_0^1 P(x,y)\,dx$$

car $\{\frac{1}{k},\,k\in\mathbb{N}^*\}$ est de mesure nulle

$$= \int_{[0,1]\setminus\{\frac{1}{k},\,k\in\mathbb{N}^*\}} P(x,y)\,dx$$
$$= \int_{[0,1]\setminus\{\frac{1}{k},\,k\in\mathbb{N}^*\}} \mathbb{1}_{[0,1]}(y)\,dx$$
$$= \mathbb{1}_{[0,1]}(y)$$

$$\boxed{\pi P(y) = \pi(y)}$$

Donc π est un invariant pour P.

3. Soit f une fonction bornée mesurable.

- $Pf(x) = \mathbb{E}[f(X_1) \mid X_0 = x]$

$$= \int_{\mathbb{R}} f(y)\, \mathbb{P}(X_1 = y \mid X_0 = x)\, dy$$

$$= \int_{\mathbb{R}} f(y)\, P(x,y)\, dy$$

Or, $x \notin \{\frac{1}{k}, k \in \mathbb{N}^*\}$, donc :

$$\boxed{Pf(x) = \int_{\mathbb{R}} f(y)\pi(y)\, dy = \int_0^1 f(y)\, dy}$$

- $P^2 f(x) = P(Pf)(x)$

$$= \int_{\mathbb{R}} Pf(y)\, P(x,y)\, dy$$

$$= \int_{\mathbb{R}} \left[\int_{\mathbb{R}} f(z)\pi(z)\, dz\right] \pi(y)\, dy$$

$$= \int_{\mathbb{R}} f(z)\pi(z) \underbrace{\left[\int_{\mathbb{R}} \pi(y)\, dy\right]}_{=1} dz$$

$$= Pf(x)$$

Puis par récurrence : $\boxed{P^n f(x) = Pf(x) \quad n \in \mathbb{N}^*}$

- On en déduit : $\boxed{\lim_{n \to +\infty} P^n f(x) = \int_0^1 f(y)\pi(y)\, dy}$

4a) Pour $n = 1$: $P^1\left(\frac{1}{k}, \frac{1}{k+1}\right) = \left(\frac{1}{k}\right)^2 \cancel{\int_{\{\frac{1}{k+1}\}} dt} + \left(1 - \left(\frac{1}{k}\right)^2\right) \underbrace{\delta_{\frac{1}{k+1}}\left(\left\{\frac{1}{k+1}\right\}\right)}_{=1}$

(d'après la question 1)

$\{\frac{1}{k+1}\}$ est de mesure nulle

donc : $P^1\left(\frac{1}{k}, \frac{1}{k+1}\right) = \left(1 - \left(\frac{1}{k}\right)^2\right)$.

<u>Hypothèse de récurrence</u> : $P^n\left(\frac{1}{k}, \frac{1}{k+n}\right) = \prod_{i=0}^{n-1} \left(1 - \left(\frac{1}{k+i}\right)^2\right)$

$$P^{n+1}\left(\frac{1}{k}, \frac{1}{k+n+1}\right) = \int_{\mathbb{R}} P\left(y, \frac{1}{k+n+1}\right) P^n\left(\frac{1}{k}, y\right) dy \quad \text{[Chapman Kolmogorov]}$$

$$= \int_{\mathbb{R}} \left(1 - \left(\frac{1}{k+n}\right)^2\right) \delta_{\frac{1}{k+n}}(y)\, P^n\left(\frac{1}{k}, y\right) dy$$

$$= \left(1 - \left(\frac{1}{k+n}\right)^2\right) P^n\left(\frac{1}{k}, \frac{1}{k+n}\right)$$

hypothèse de récurrence

$$= \prod_{i=0}^{n} \left(1 - \left(\frac{1}{k+i}\right)^2\right)$$

b) $A = \bigcup_{q \in \mathbb{N}} \left\{\frac{1}{k+1+q}\right\}$

$$P^n(x, A) = \sum_{q \in \mathbb{N}} P^n\left(x, \frac{1}{k+1+q}\right) \quad \text{(indépendance)}$$

$$= \sum_{q \in \mathbb{N}} \mathbb{1}_{\{n = q+1\}} P^n\left(\frac{1}{k}, \frac{1}{k+n}\right)$$

$$\boxed{P^n(x, A) = \prod_{i=0}^{n-1} \left(1 - \left(\frac{1}{k+i}\right)^2\right)}$$

$$\ln P^n(x, A) = \sum_{i=0}^{n-1} \ln\left(1 - \left(\frac{1}{k+i}\right)^2\right)$$

Or, $\ln\left(1 - \left(\frac{1}{k+i}\right)^2\right) \underset{i \to +\infty}{\sim} -\left(\frac{1}{k+i}\right)^2$ qui est le terme général d'une série convergente.

Donc : $\ln P^n(x, A)$ converge vers C et <u>$P^n(x, A)$ converge vers $e^C > 0$</u>

Or, $\pi(A) = \int_A \mathbb{1}_{[0,1]}(y)\, dy = 0$ car A est de mesure nulle.

Donc : $\boxed{\lim_{n \to +\infty} P^n(x, A) \neq \pi(A)}$