Grauler Thomas

Computational Statistics TP 1

Exercice n°1

1) R et Θ sont indépendant, la loi jointe $P_{(R,\Theta)}$ est

$$f_{(R,\Theta)} = \frac{r}{2\pi} \exp\left(\frac{-r^2}{2}\right) \mathbb{1}_{\mathbb{R}^+}(r)\ \mathbb{1}_{[0;2\pi]}(\theta)$$

Prenons $h : \mathbb{R}^2 \to \mathbb{R}^2$ mesurable et borné

Soit $X = R\cos(\Theta)$, $Y = R\sin(\Theta)$

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

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} h(r\cos\theta, r\sin\theta)\, f_{r,\theta}\, dr\, d\theta$$

On définit $g(r,\theta) = (r\cos\theta, r\sin\theta)$, $\forall (r,\theta) \in \mathbb{R}^+ \times [0;2\pi]$

On applique le changement de variables g qui est un C^1 difféomorphisme

On peut calculer le jacobien J_g :

$$J_g = \begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} \Rightarrow \det J_g = r$$

$g(r,\theta) = (x,y) \Leftrightarrow r = \sqrt{x^2+y^2}$

$\theta = \arctan\left(\frac{y}{x}\right)$

Donc $\mathbb{E}[h(X,Y)] = \displaystyle\iint_{\mathbb{R}\,\mathbb{R}} h(x,y) \frac{\cancel{\sqrt{x^2+y^2}}}{2\pi\,\cancel{\sqrt{x^2+y^2}}}\, e^{-(x^2+y^2)/2}\, dx\, dy$

On reconnait la loi jointe de 2 indépendantes loi $\mathcal{N}(0;1)$

Ainsi on a $X \sim \mathcal{N}(0;1)$, $Y \sim \mathcal{N}(0,1)$ et $X \perp\!\!\!\perp Y$.

2) Soit F_R la fonction de répartition de $R \sim$ Rayleigh (1)

Pour $r \in \mathbb{R}^+$ on a $F_r(r) = \mathbb{P}(R \leq r) = \int_0^r x e^{-x^2/2} dx = \left[-e^{-x^2/2}\right]_0^r$

$= 1 - e^{-r^2/2}$

pour $u \in [0,1]$ on a

$u = 1 - e^{-r^2/2} \Leftrightarrow r = \sqrt{-2\ln(1-u)} = F_r^{-1}(u)$

Ainsi $U \sim \mathcal{U}([0,1])$, on a $R = F_R^{-1}(U) = \sqrt{-2\ln(U)}$

car avec $U \sim \mathcal{U}([0,1])$ on a $U \overset{\mathcal{L}}{=} 1-U$

De plus pour $U \sim \mathcal{U}([0;1])$ on a $2\pi U \sim \mathcal{U}([0;2\pi])$

L'algorithme est
- $U_1, U_2 \overset{iid}{\sim} \mathcal{U}([0,1])$
- $R \leftarrow \sqrt{-2\ln(U_1)}$
- $\Theta \leftarrow 2\pi U_2$
- $X \leftarrow R\cos\Theta$
- $Y \leftarrow Y\sin\theta$
- return X, Y

3) a) Pendant la boucle while V_1 et V_2 suivent une loi uniforme $[-1;1]$ et son indépendant.

Sortir de la condition implique $V_1^2 + V_2^2 \leq 1 \Leftrightarrow \sqrt{V_1^2+V_2^2} \leq 1$

Donc on conclue que V_1, V_2 sont dans le disque unité et suivent une loi uniforme. Avec $D(0,1) = \{(x,y) \in \mathbb{R}^2,\ x^2+y^2 \leq 1\}$

On a $(V_1, V_2) \sim \mathcal{U}(D(0,1))$

b) Le nombre de répétition de la boucle while peut être vue comme le nombre d'insuccès avant un succès. Elle suit une loi géométrique $\mathcal{G}$ de paramètre p. $p = \mathbb{P}(\text{"succès"}) = \mathbb{P}(V_1^2+V_2^2 \leq 1)$

avec $U_1, U_2 \overset{iid}{\sim} \mathcal{U}([-1,1])$

$\mathbb{E}(G) = \frac{1}{p}$

On calcul $p = \mathbb{P}(V_1^2+V_2^2 \leq 1) = \mathbb{E}[\mathbb{1}_{V_1^2+V_2^2 \leq 1})$

$$= \iint_{\mathbb{R}\,\mathbb{R}} \mathbb{1}_{\{u^2+v^2\leq 1\}} \frac{1}{2}\mathbb{1}_{[-1;1]}(u) \frac{1}{2} \mathbb{1}_{[-1,1]}(v)\, du\, dv$$

$$= \frac{1}{4}\int_{-1}^{1}\int_{-\sqrt{1-v^2}}^{\sqrt{1-v^2}} 1\, du\, dv$$

$$= \frac{1}{4}\int_{-1}^{1} 2\sqrt{1-v^2}\, dv$$

avec le changement de variable

$$\left| \begin{array}{l} v = \cos(x) \iff x = \arccos(x) \\ -1 \leq v \leq 1 \iff -\pi \leq x \leq 0 \\ dv = -\sin(x)\, dx \end{array} \right.$$

$$= \frac{1}{2}\int_{-\pi}^{0} \sqrt{1-\cos^2(x)}\,(-\sin(x))\,dx$$

$$= \frac{1}{2}\int_{-\pi}^{0} \sin^2(x)\,dx$$

$$= \boxed{\frac{\pi}{4}}$$

Donc le nombre d'itération est $\mathbb{E}(G) = \frac{1}{p} = \frac{4}{\pi}$

On aurait aussi pu le voir comme le rapport de mesure des aires entre le carré $[-1,1]^2$ et le disque $D(0,1)$

c) Soit $T_1 = \frac{V_1}{\sqrt{V_1^2+V_2^2}}$, $V = V_1^2+V_2^2$

Sachant que $(V_1,V_2) \sim \mathcal{U}(D(0,1))$, la loi jointe est

$$f_{V_1,V_2}(u,v) = \frac{1}{\pi}\mathbb{1}_{\{u^2+v^2 \leq 1\}}$$

Prenons φ une fonction borné mesurable.

$$\mathbb{E}(\varphi(T_1,V)) = \mathbb{E}\left(\varphi\left(\frac{V_1}{\sqrt{V_1^2+V_2^2}}, V_1^2+V_2^2\right)\right)$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi\left(\frac{u}{\sqrt{u^2+v^2}}, u^2+v^2\right) \frac{1}{\pi}\mathbb{1}_{\{u^2+v^2\leq 1\}}\,du\,dv$$

Avec le même changement de variable polaire de la question 1

$$\mathbb{E}[\varphi(T_1,V)] = \iint_{\mathbb{R}\mathbb{R}} \varphi(\cos\theta, r^2)\frac{r}{\pi}\mathbb{1}_{r\in[0,1]}\mathbb{1}_{\{\theta\in[0,2\pi[}\,dr\,d\theta$$

$$\overset{dr = \frac{1}{2r}dy}{=} \iint_{\mathbb{R}\mathbb{R}} \varphi(\cos\theta, y)\frac{1}{2\pi}\mathbb{1}_{\theta\in[0,2\pi[}\mathbb{1}_{y\in[0,1]}\,dy\,d\theta$$

On reconnait la loi jointe séparable d'une densité $\mathcal{U}([0,1]) \times \mathcal{U}([0,2\pi])$. De plus on a aussi $T_1 \perp\!\!\!\perp V$, avec $T_1 \overset{\mathcal{L}}{=} \cos(\theta)$ avec $\theta \sim \mathcal{U}([0,2\pi])$ et $V \sim \mathcal{U}([0,1])$

d) ~~Retour~~ Si on utilise la question précédente dans l'algorithme
Avec la fin des itérations la boucle while donne $(V_1,V_2) \sim \mathcal{U}(D(0,1))$
Avec $S = \sqrt{-2\log(V_1^2+V_2^2)}$ on a $S = \sqrt{-2\log(V)}$ avec $V \sim \mathcal{U}([0,1])$

Avec la question précédente et la question 2 on sait que $S \sim \text{Rayleigh}(1)$. De plus $X = S\frac{V_1}{\sqrt{V_1^2+V_2^2}} = S\cos(\theta)$ et $Y = S\frac{V_2}{\sqrt{V_1^2+V_2^2}} = S\sin(\theta)$

avec $\theta \sim \mathcal{U}([0,2\pi])$

Avec la question 1 on déduit $(X,Y) \sim \mathcal{N}(0,1)^2$

Exercice n°2 :

Soit $n \in \mathbb{N}$ on a $P(x,A) = \mathbb{P}(X_{n+1} \in A \mid X_n = x)$

1) Il y 2 cas, prenons $x \in [0,1]$ et A un état mesurable

- Si il existe $k \in \mathbb{N}^*$ tel que $x = \frac{1}{k}$, on pose
$B_0 = \{X_{n+1} \in A \mid X_n = \frac{1}{h}\}$ et $A_0 = \{X_{n+1} = \frac{1}{h+1}\}$; $A_1 = \{X_{n+1} \sim \mathcal{U}([0,1])$

Donc $P(x,A) = \mathbb{P}(B_0) = P(B_0 \cap A_0) + \mathbb{P}(B_0 \cap A_1)$ système complet d'évènement

car $\left.\begin{array}{l}\mathbb{P}(A_0 \mid X_n = x = \frac{1}{h}) = 1-x^2 \\ \mathbb{P}(A_1 \mid X_n = x = \frac{1}{h}) = x^2\end{array}\right| = (1-x^2)\delta_{\frac{1}{h+1}}(A) + x^2\int_{A\cap[0,1]} 1\,dt$

- Avec $B_1 = \{X_{n+1} \in A \mid \exists h\ X_n = \frac{1}{h}\}$
$A_2 = \{X_{n+1} \sim \mathcal{U}[0,1]\}$
$P(x,A) = \mathbb{P}(B_1) = P(B_1 \cap A_2)$
$= \int 1\,dt$

On a donc caractérisé la chaîne de markov $(X_n)_{n \geq 0}$

2) Avec la loi uniforme π sur $[0,1]$
On sait que π est invariant pour $P \Leftrightarrow \forall A \ \pi(A) = \int_{\mathbb{R}} \pi(x) P(x,A)$

On a $\int_{\mathbb{R}} \pi(x) P(x,A) \, dx = \int_{\mathbb{R}} \mathbb{1}_{[0,1]}(x) P(x,A) \, dx$

$$= \int_{\mathbb{R}} \underbrace{\mathbb{1}_{[0,1]}(x)}_{1} \left(\int_{A \cap [0,1]} dt \right) dx$$

$$= \int_{A \cap [0,1]} dt = \pi(A)$$ mesure positive + fubini car dénombrable.

Donc $\boxed{\pi P = \pi}$

3) Soit $x \notin \{ \frac{1}{k}, k \in \mathbb{N}^* \}$ et f une fonction borné mesurable

$Pf(x) = \mathbb{E}[f(X_1) | X_0 = x]$ avec $X_1 | X_0 = x \sim \mathcal{U}([0,1])$

$= \int_{\mathbb{R}} f(y) \mathbb{1}_{[0,1]}(y) dy$

Sachant que si $X_n \neq \frac{1}{k}$ alors X_{n+1} est $\mathcal{U}([0,1])$

Calculons $P^2 f(x) = P(Pf(x)) = \mathbb{E}[Pf(X_1) | X_0 = x]$

~~P^2~~ Étant donné que $\mathcal{U}[0,1]$ est invariant pour P,

on a $P^2 f(x) = Pf(x) = \int_{\mathbb{R}} f(y) \mathbb{1}_{[0,1]}(y) dy$.

Par récurrence immédiate $P^n f(x) = Pf(x)$.

Donc $P^n f(x) = \int f(y) \mathbb{1}_{[0,1]}(y) dy$

$\forall n \geq 1 \quad P^n f(x) = \int_0^1 f(y) dy$

On a donc $\lim_{n \to +\infty} P^n f(x) = \int_0^1 f(y) dy$

Pour cette chaîne de markov $\int_0^1 f(y) dy = \int_{\mathbb{R}} f(x) \pi(x) dx$

On obtient $P^n f(x) \xrightarrow{+\infty} \int_{\mathbb{R}} f(x) \pi(x) dx$

4) a) Soit $x = \frac{1}{h}$, $h \geq 2$. Soit $n \in \mathbb{N}^*$. Pour se déplacer de $\frac{1}{h}$ vers $\frac{1}{h+n}$ après n étapes, il existe une seule façon.

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

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

On obtient $P^n \left(x, \frac{1}{n+h} \right) = P^n \left(\frac{1}{h}, \frac{1}{n+h} \right) = \prod \left(1 - \frac{1}{(h+i)^2} \right)$

b) Si $A = \bigcup_{q \in \mathbb{N}} \left\{ \frac{1}{h+1+q} \right\}$ A est l'ensemble des point tel que

$\pi(A) = \int_{A \cap [0,1]} dt = 0$

De plus $P^n \left(x, \frac{1}{n+h} \right) \xrightarrow[+\infty]{} 0$ comme le produit de valeur entre 0 et 1.

Donc on a $\lim_{n \to +\infty} P^n (x, A) = \pi(A) = 0$

Exercice n°3

1) Une façon de minimiser $R_n(w)$ est d'utiliser est descente de gradient stochastique. Contrairement à la descente de gradient classique, la SGD utilise un petit batch d'estimation pour calculer le gradient.

1 - on initialise w_0 (le paramètre à estimer)
2 - Pour chaque itération :
- on sélectionne un échantillon aléatoire (x_i, y_i)
- calcul le gradient $\nabla R_n(w) = -2x_i(y_i - x_i^T w)$
- on met à jour les paramètres $w_{k+1} = w_k - \eta \nabla_w R_n(w_k)$

On arrête quand on répète 3 fois 2 jusqu'à la convergence
et que la valeur ne change pas on arrête (ε près)
eps