Examen LM346 "Processus et simulations", 2ème session 2015–2016, sans document, ni calculatrice.

Rappels. *La densité de la loi Gaussienne de paramètres* $(0,1)$*, c'est-à-dire centrée réduite, est* $(1/\sqrt{2\pi})\exp(-x^2/2)$.

(1) On considère une variable aléatoire Z de densité $g_Z(x) = D \times x\exp(-x^2/2) \times 1_{[0,1]}(x)$ où D est une constante. Calculer la constante D. Calculer la fonction de répartition $F_Z(x)$ de la variable aléatoire Z.

Suggestion pour faciliter le calcul : que vaut la dérivée de $\exp(-x^2/2)$ *?*

(2) Proposer une méthode de simulation de la loi de la variable aléatoire Z.

(3) On considère une variable aléatoire X de densité $f_X(x) = C \times \exp(-x^2/2) \times 1_{[0,1]}(x)$ où C est une constante. *(Faites attention, la densité est différente de* $g_Z(x)$*, il n'y a pas de* x*).* Que vaut la constante C ?

Proposer une méthode de calcul approché de cette constante (il s'agit d'une méthode apprise cours).

(4) Soient $X_1, X_2, X_3, \dots$ des variables aléatoires Gaussiennes centrées réduites. Proposer une ou des méthode(s) de leur simulation.

(5) Soit $\nu = \min\{n \geq 1 : X_n \in [0,1]\}$ ($X_1, X_2, \dots$ sont comme dans la question (4)). Donner $P(\nu = k)$ pour $k = 1, 2, 3, \dots$. Exprimer cette probabilité en terme de la constante C de la question (3).

(6) Calculer la probabilité $P(X_\nu \in I)$ où un ensemble Borélien $I \subset [0,1]$. *(On pensera à calculer d'abord les probabilités* $P(X_\nu \in I, \nu = k)$ *pour* $k = 1, 2, \dots$.*)*

(7) Proposer une ou des méthode(s) de simulation d'une variable aléatoire de même loi que X de la question (3). *(On pourra penser au résultat de la question (6) et aussi à une méthode apprise en cours).*

(8) Soient $Y_1, \dots, Y_n, \dots$ des variables aléatoires indépendantes, de même loi, de second moment fini, $\mathbf{E}Y_1 = 0$. Pour $\beta \in]0, 1/2[$, on note $a(\beta)$ tel que $(\sqrt{2\pi})^{-1}\int_{a(\beta)}^{\infty} \exp(-t^2/2)dt = \beta$. Donner la valeur de la limite suivante : $\lim_{n\to\infty} \mathbf{P}\Big(-a(\beta) < \frac{Y_1+\cdots+Y_n}{\sqrt{n\mathrm{Var}\,(Y_1)}} < a(\beta)\Big)$.

(9) Soit $\beta \in]0, 1/2[$. Construire un intervalle de confiance asymptotique $[I_n(\beta), \infty[$ de niveau de fiabilité $1 - 2\beta$ pour la variance $\mathrm{Var}\,(Y_1)$, c'est-à-dire tel que

$$\lim_{n\to\infty} \mathbf{P}(\mathrm{Var}\,(Y_1) \in [I_n(\beta), \infty[) \geq 1 - 2\beta.$$

(10) Un referendum s'annonce dans un pays, avec une seule question à laquelle chaque citoyen doit répondre "oui" ou "non". Une proportion p de la populatiion est déjà determinée dans sa réponse : exactement la moitié compte répondre "oui" à la question du referendum et l'autre moitié "non". La proportion $(1 - p)$ de la population n'est pas encore determinée dans sa réponse.

On veut estimer p. Pour cela on fait un sondage de 100000 individus : 47000 se disent indécis, 27000 vont répondre "oui" et 26000 vont répondre "non". Construire un intervalle de confiance pour la proportion p de niveau $1 - 2\beta$.

Suggestion : on pourrait essayer de modéliser la réponse de chaque individu par une variable aléatoire à trois valeurs $1, -1, 0$ *(si l'individu répond "oui", "non" ou indécis respectivement) qui sont prises avec des probabilités que vous préciserez (en fonction de ce que vous avez lu ci-dessus). On suposera que les réponses des individus sont indépendantes. On pourrait ensuite calculer la variance de ces variables aléatoires en terme de p et utiliser le résultat de la question (9).*

TOURENEZ LA PAGE SVP

(11) On considère une Chaine de Markov $(X_n)_{n\geq 0}$ de matrice de transition

$$\begin{pmatrix} 1/2 & 0 & 0 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 1/3 & 0 & 2/3 & 0 & 0 \\ 1/4 & 0 & 0 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 2/3 & 0 & 0 \end{pmatrix}$$

Identifier les classes d'états, distinguer les classes fermées et non fermées.

(12) Trouver l'ensemble de mesures invariantes de cette Chaine de Markov.

(13) Soit $T^i = \inf\{n > 0 : X_n = i\}$, $h_j^{(i)} = P(T^i < \infty \mid X_0 = j)$ la probabilité que l'état i sera visité à un instant du temps fini poistif au départ de l'état initial j. Calculer $h_3^{(1)}$ et $h_6^{(1)}$.

(14) En déduire $h_3^{(5)}$ et $h_6^{(5)}$.

(15) En déduire $h_3^{(2)}$ et $h_6^{(2)}$.

(16) Donner $\lim_{n\to\infty} P(X_n = i \mid X_0 = 1)$ pour $i = 1, 2, 3, 4, 5, 6$ et jusitifer votre réponse.

(17) Donner $\lim_{n\to\infty} P(X_n = i \mid X_0 = 3)$ pour $i = 1, 2, 3, 4, 5, 6$ et jusitifer votre réponse.

(18) On considère maintenant une Chaine de Markov avec ϵ qui vérifie l'inégalité $0 \leq \epsilon < 1/2$.

$$\begin{pmatrix} 1/2 & 0 & \epsilon & 0 & 1/2-\epsilon & 0 \\ 0 & 1/2 & 0 & 1/2-\epsilon & 0 & \epsilon \\ 0 & 0 & 0 & 0 & 1/4 & 3/4 \\ 0 & 1/3 & 0 & 2/3 & 0 & 0 \\ 1/4 & 0 & 0 & 0 & 3/4 & 0 \\ 0 & 0 & 1/3 & 2/3 & 0 & 0 \end{pmatrix}$$

Supposons dans la question (18) que $\epsilon > 0$ (ϵ strictement positif) Donner les classes de cette chaine de Markov. Donner $h_i^{(j)}$ pour de différentes pairs d'états i et j, $i, j \in \{1, 2, 3, 4, 5, 6\}$ et justifier votre réponse (ceci ne demande pas beaucoup de calculs)

(19), (20) Soit $\mu = (1/6, 2/10, 0, 3/10, 1/3, 0)$ la mesure initiale.

Que pouvez-vous dire de la suite $\alpha_n = P(X_n = 2)$ pour $n \geq 0$ sous cette mesure initiale si $\epsilon = 0$?

Que pouvez-vous dire de la suite $\alpha_n = P(X_n = 2)$ pour $n \geq 0$ sous cette mesure initiale si $\epsilon > 0$?

Correction de l'examen LM346 "Processus et simulations", 2ème session 2015–2016.

(1) $D = \Big(\int_0^1 x \exp(-x^2/2)dx \Big)^{-1} = (1-\exp(-1/2))^{-1}$.

$F_Z(x) = 0$ pour $x \le 0$, $F_Z(x) = D(1-\exp(-x^2/2))$ pour $x \in [0,1]$ et $F_Z(x) = 1$ pour $x \ge 1$.

(2) C'est la méthode d'inversion de $F_Z(x)$. On obtient $Z = \sqrt{-2\ln(1-U/D)}$.

(3) $C = \Big(\int_0^1 \exp(-x^2/2)dx \Big)^{-1}$. On peut calculer $n\Big(\sum_{i=1}^n \exp(-U_i^2/2) \Big)^{-1}$ avec n grand, $U_1, U_2, \ldots$ étant des v.a. de loi uniforme sur $[0,1]$. C'est la méthode de Monté-Carlo.

(4) Une méthode est celle de Box-Muller $\sqrt{-2\ln U_1}\cos(2\pi U_2)$, $\sqrt{-2\ln U_1}\sin(2\pi U_2)$, U_1, U_2 étant indép. de loi uniforme sur $[0,1]$.

L'autre –méthode de rejet. On simule (par la méthode d'inversion) des v.a. indép. de loi exponentielle de paramètre $(1/2)$, cad $X_1, Y_1, X_2, Y_2, \ldots$. On pose $\nu_0 = 0$, On pose $\nu_k = \min\{n > \nu_{k-1} : Y_n > (1/2)(1-X_n)^2\}$, $k = 1, 2, \ldots$. Soient $\xi_1, \xi_2, \ldots$ des v.a. indép. à valeurs ± 1 avec probabilités $1/2$. Alors $\xi_1 X_{\nu_1}, \xi_2 X_{\nu_2}, \ldots \xi_k X_{\nu_k}, \ldots$ sont indép. Gaussiennes centrées réduites.

(5)

$$P(\nu = k) = P(X_1, \ldots, X_{k-1} \notin [0,1], X_k \in [0,1]) = \Big(1-(1/\sqrt{2\pi})\int_0^1 \exp(-x^2/2)dx\Big)^{k-1}(1/\sqrt{2\pi})\int_0^1 \exp(-x^2/2)dx$$

$$= \Big(1 - \frac{1}{C\sqrt{2\pi}}\Big)^{k-1}\frac{1}{C\sqrt{2\pi}}.$$

(6)

$$\sum_{k=1}^{\infty} P(X_\nu \in I, \nu = k) = \sum_{k=1}^{\infty} P(X_1, \ldots, X_{k-1} \notin [0,1], X_k \in I)$$

$$= \sum_{k\ge 1} \Big(1-(1/\sqrt{2\pi})\int_0^1 \exp(-x^2/2)dx\Big)^{k-1}(1/\sqrt{2\pi})\int_I \exp(-x^2/2)dx$$

$$= \frac{(1/\sqrt{2\pi})\int_I \exp(-x^2/2)dx}{(1/\sqrt{2\pi})\int_0^1 \exp(-x^2/2)dx} = C\int_I \exp(-x^2/2)dx.$$

(7) On peut juste simuler X_ν d'après la question (6).

On peut aussi utiliser la methode de rejet par rapport à la densité sur un segment (méthode du cours) : $\theta = \min\{n \ge 1 : V_n < \exp(-U_n^2/2)\}$, $U_1, V_1, U_2, V_2, \ldots$ étant indép. de loi uniforme sur $[0,1]$. On simule U_θ.

(8) La limite vaut $(\sqrt{2\pi})^{-1}\int_{-a(\beta)}^{a(\beta)} \exp(-t^2/2)dt = 1-2\beta$ par le Thm de la limite centrale.

(9) L'inégalité $-a(\beta) < \frac{Y_1+\cdots+Y_n}{\sqrt{n\mathrm{Var}\,(Y_1)}} < a(\beta)$ est équivalente à $\sqrt{\mathrm{Var}(Y_1)} > |(Y_1+\cdots+Y_n)/(\sqrt{n}a(\beta))|$.

Alors $I_n(\beta) = (Y_1+\cdots+Y_n)^2/(n(a(\beta))^2)$.

(10) On pose $Y_i = 0$ si le $i$ème sondé n'est pas determiné dans son choix et $Y_i = 1$ (resp. $Y_i = -1$) si le $i$ème sondé est pour "oui" (resp. "non"). Alors $Y_1, \ldots, Y_{100000}$ sont des v.a. indép., de même loi, à valeurs dans 0, 1, -1 avec probabilités $1-p$, $p/2$, $p/2$ respectivement. On remarque que $\mathbf{E}Y_1 = 0$, $\mathrm{Var}Y_1 = p$. Le résultat du sondage nous donne : $Y_1+\cdots+Y_{100000} = 27000-26000 = 1000$. Alors $I_{100000}(\beta) = (Y_1+\cdots+Y_{100000})^2/(100000(a(\beta))^2) = 10/(a(\beta))^2$. L'intervalle pour p est $[10/(a(\beta))^2, 1]$.

(11) $\{1,5\}$, $\{2,4\}$ sont les classes fermées, $\{3,6\}$ la classe non fermée.

(12) C'est l'ensemble $(c_1/3, 2c_2/5, 0, 3c_1/5, 2c_2/3, 0)$, $c_1, c_2 \ge 0$, $c_1+c_2 = 1$.

(13) On a $h_3^{(1)} = h_3^{(1,5)} = h_3^{(5)}$ et $h_6^{(1)} = h_6^{(1,5)} = h_6^{(5)}$ car les états 1 et 5 sont dans la même classe. On a $h_3^{(1,5)} = 1/4 + 3/4h_6^{(1,5)}$, $h_6^{(1,5)} = 1/3h_3^{(1,5)}$.

Donc $h_3^{(1,5)} = 1/3 = h_3^{(1)}$ et $h_6^{(1,5)} = 1/9 = h_6^{(1)}$.

(14) $h_3^{(1,5)} = 1/3 = h_3^{(5)}$ et $h_6^{(1,5)} = 1/9 = h_6^{(5)}$.

(15) $h_3^{(2,4)} = 1 - 1/3 = 2/3 = h_3^{(2)}$ et $h_6^{(2,4)} = 1 - 1/9 = 8/9 = h_6^{(2)}$.

(16) Soit $i = 1, 5$. La classe $\{1, 5\}$ étant récurrente apériodique, la limite vaut la mesure de probabilité invariante de l'état i pour la CM restreinte à cette classe, donc $1/3$ pour $i = 1$ et $2/3$ pour $i = 5$. Elle vaut 0 pour $i = 2, 3, 4, 6$.

(17) Les classes $\{1, 5\}$ et $\{2, 4\}$ sont récurrentes apériodiques. La limite vaut $1/3 \times h_3^{(1,5)} = 1/9$ pour $i = 1$ et $2/3 \times h_3^{(1,5)} = 2/9$ pour $i = 5$.

La limite vaut $2/5 \times h_3^{(2,4)} = 4/15$ pour $i = 2$ et $3/5 \times h_3^{(2,4)} = 6/15$ pour $i = 5$.

La limite vaut 0 pour $i = 3, 6$ car ces états son trabsients.

(18) Pour $\epsilon > 0$ cette CM est irréductible, donc $h_j^{(i)} = 1$ pour tous i, j.

(19), (20) Pour $\epsilon = 0$, la mesure initiale $\vec{\mu}$ est une des mesures invarriantes d'après l'une des questions précédentes. Donc c'est $\alpha_n = \nu_2 = 2/10$ pour tout $n \geq 0$.

Pour $\epsilon > 0$, la mesure intiale μ n'est pas une mesure invariOr la CM est irréductible apériodique finie. Donc cette suite $\alpha_n \to \pi_2$ où $\pi = (\pi_1, \pi_2, \ldots, \pi_6)$, est l'unique mesure de probabilité invariante de cette CM avec $\epsilon > 0$.