Examen LM346 "Processus et simulations", 1ère session 2016–17, sans document, ni calculatrice.

Rappels. Une variable aléatoire est de loi exponentielle de paramètre $\lambda > 0$ si sa densité est $f(x) = \lambda \exp(-\lambda x) 1_{x>0}$.

Une variable aléatoire Gaussienne centrée réduite est de densité $\frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$.

Une variable aléatoire Gaussienne de l'espérance a et de la variance σ^2 est de densité $\frac{1}{\sqrt{2\pi\sigma^2}} \exp(-(x-a)^2/2\sigma^2)$.

Une variable aléatoire de loi de Poisson de paramètre $\lambda > 0$ prend des valeurs $k = 0, 1, 2, \ldots$ avec des probabilités $\exp(-\lambda)\frac{(\lambda)^k}{k!}$ respectivement.

(1) Calculer la fonction de répartition et l'espérance de la loi exponentielle de paramètre 1.

(2) Soient $U_1, \ldots, U_i, \ldots$ une suite de variables aléatoires indépendantes de loi uniforme sur $[0, 1]$.

Proposer une méthode de simulation d'une suite de variables aléatoires indépendantes $X_1, X_2, \ldots, X_i, \ldots$ de loi exponentielle de paramètre 1 à partir de $U_1, U_2, \ldots$.

(3) Les variables aléatoires $X_1, \ldots, X_6$ de la question (2) prennent les valeurs $1, 1/2, 1/3, 1/4, 1/5, 1/6$. Simuler une valeur d'une variable aléatoire Y de loi de Poisson de paramètre 2.

(4) Calculer $P\Big((X_1 \in [a, b]) \cap ((1 - X_1)^2 < 2X_2)\Big)$ pour $a < b$, $a, b \in \mathbf{R}$.

(5) Déduire $P\Big((1 - X_1)^2 < 2X_2\Big)$

(6) Soit $\nu = \min\{i \in \{1, 2, 3, \ldots\} : (1 - X_{2i})^2 < 2X_{2i-1}\}$. Calculer $P(\nu = k)$ pour $k = 1, 2, 3, \ldots$.

(7) Calculer $P(X_\nu \in [a, b])$. En déduire la densité de X_ν.

(8) Comment, à partir de la variable aléatoire X_ν, simuler une variable aléatoire Gaussienne centrée réduite Z?

(9) Comment à partir d'une variable aléatoire Gaussienne centrée réduite Z, simuler une variable aléatoire Gaussienne dont l'espérance est a et la variance σ^2 ?

Voici les données pour les questions (10)–(12) en Statistiques. Les astronomes ont découvert un objet dans l'espace et ils essaient d'étudier les températures sur lui. Ils mésurent les températures pendant $N = 100$ jours. La somme de ces températures pendant $N = 100$ jours s'avère $S_N = 5 \times 10^6$ degrés ! La somme des carrés de ces températures pendant $N = 100$ jours est calculée, c'est $C_N^2 = 8 \times 10^{12}$. Ils essaient de construire un intervalle de confiance pour la température moyenne sur cet objet de niveau de fiabilité 0.99.

Pour cela, ils supposent qu' à chaque mesure la valeur obtenue est indépendante des autres.

On note q_α le nombre tel que $\frac{1}{\sqrt{2\pi}} \int_{q_\alpha}^{\infty} \exp(-x^2/2)dx = \alpha$.

(10) On suppose de plus qu' à chaque mesure la valeur est une réalisation d'une variable aléatoire Gaussienne de variance 1. Construire l'intervalle de confiance de niveau de fiabilité 0.99 pour la température moyenne en terme de N, S_N et q_α avec α proprement choisi.

(11) On suppose que les températures sont les réalisations de variables aléatoires de même loi quelconque de variance 1. Construire l'intervalle de confiance asympototique(!) pour la température moyenne de niveau de fiabilité 0.99 en terme de N, S_N et q_α avec α proprement choisi.

Enoncer proprement le Théorème surlequel s'appuie votre résultat.

(12) On suppose que les températures sont les réalisations de variables aléatoires de même loi quelconque de variance inconnue. Construire l'intervalle de confiance asymptotique(!) pour la température moyenne de niveau de fiabilité 0.99 en terme de N, S_N, C_N^2 et q_α avec α proprement choisi. (On pourrait penser à un estimateur pour la variance.)

(13) On considère une chaine de Markov $(X_n)_{n\geq 0}$ de matrice de transition suivante

$$\begin{pmatrix} 1/3 & 1/3 & 0 & 0 & 0 & 0 & 1/3 \\ 1/2 & 1/4 & 0 & 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1/6 & 0 & 0 & 5/6 & 0 \\ 1/4 & 0 & 0 & 1/6 & 2/6 & 0 & 1/4 \\ 0 & 0 & 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 0 & 1/3 & 0 & 0 & 2/3 & 0 \\ 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 \end{pmatrix}$$

Donner ses classes d'états. Lesquelles sont fermées ?

Pour un sous-ensemble $A \subset \{1, 2, \ldots, 7\}$:

on note $T^A = \inf\{n > 0 : X_n \in A\}$ le premier instant positif lorsque la CM se retrouve dans A,

on note $h_i^A = P(T^A < \infty \mid X_0 = i)$ la probabilité d'atteindre l'ensemble A à partir de l'état i.

(14) Donner $\lim_{n\to\infty}(P(T^1 < \infty \mid X_0 = 1))^n$, $\lim_{n\to\infty}(P(T^4 < \infty \mid X_0 = 4))^n$.

(15) Calculer $h_4^{\{1,2,7\}}$, $h_5^{\{1,2,7\}}$, $h_4^{\{3,6\}}$, $h_5^{\{3,6\}}$.

(16) Donner $h_i^{\{1\}}$ pour $i = 2, 3, 4, 5, 6, 7$.

(17) Calculer l'ensemble de mesures invariantes de probabilté pour cette CM.

(18) Donner $\lim_{n\to\infty} P(X_n = 1 \mid X_0 = i)$ pour $i = 1, 2, 3, 4, 5, 6, 7$ et argumentez votre réponse.

(19) Calculer $E(T^{\{1\}} \mid X_0 = i)$ pour $i = 1, 2, 3, 4, 5, 6, 7$ (ceci ne demande pas de calculs pour certains i).

(20) Soit $(3/21, 4/63, 4/21, 0, 0, 10/21, 8/63)$ une mesure initiale. Donner $P(X_4 = 2)$ sous cette mesure.

Correction

(1) La fonction de répartition est $(1-\exp(-x))1_{x>0}$, l'espérance est 1 (les étudiants devaient faire une intégration par parties pour cela.)

(2) Par inversion de la fonction de répartition $X_i = \ln(1-U_i)^{-1}$, $i = 1, 2, \ldots$. Ou encore comme $1-U$ est de même loi que U, on pourrait prendre $X_i = \ln U_i^{-1}$.

(3) On cherche $Y = \max\{k \geq 1 : X_1 + \cdots X_k \leq 2\}$ (et $Y = 0$ si $X_1 > 2$), alors Y est de loi de Poisson de paramètre $1 \times 2 = 2$. Dans notre cas $1 < 2$, $1 + 1/2 < 2$, $1 + 1/2 + 1/3 < 2$, $1 + 1/2 + 1/3 + 1/4 > 2$, donc $Y = 3$.

(4) C'est $r_{a,b} = \int_a^b \left(\int_{1/2(1-x_1)^2}^{\infty} \exp(-x_1 - x_2)dx_2\right)dx_1 = (-1)\int_a^b(-1)\exp(-x_1 - (1/2)(1-x_1)^2)dx_1 = \exp(-1/2)\int_a^b \exp(-x_1^2/2)dx_1$.

(5) $p = \int_0^{\infty}\left(\int_{1/2(1-x_1)^2}^{\infty} \exp(-x_1 - x_2)dx_2\right)dx_1 = \exp(-1/2)\int_0^{\infty}\exp(-x_1^2/2)dx_1 = \frac{2\exp(-1/2)}{\sqrt{2\pi}}$

(6) $P(\nu = k) = (1-p)^{k-1}p$ avec p ci-dessus.

(7) $P(X_\nu \in [a,b]) = \sum_{k\geq 1}^{\infty} P(X_\nu \in [a,b], \nu = k) = \sum_{k=1}^{\infty}(1-p)^{k-1}r_{a,b} = \frac{r_{a,b}}{p} = \frac{2}{\sqrt{2\pi}}\int_a^b \exp(-x^2/2)dx$.

La densité $\frac{2}{\sqrt{2\pi}}\exp(-x^2/2)1_{x>0}$.

(8) Soit V une v.a. indép. de X_ν à valeurs ± 1 avec probabilité 1/2. Alors $Z = V \times X_\nu$ est de loi Gaussienne centrée réduite. On peut simuler $V = 1_{U<1/2} - 1_{U>1/2}$, U étant indépendante de v.a. précédentes.

(9) On prend $\sigma Z + a$.

(10) Soit T cette température moyenne. $(S_N - TN)/\sqrt{N \times 1}$ est une variable aléatoire Gaussienne de param. $(0,1)$. Soit $q = q_{0.005}$ la quantile de cette loi Gaussienne. Alors $P(-q_{0.005} < (S_N - TN)/\sqrt{N} < q_{0.005}) = 0.99$. $T \in [(S_N - q_{0.005}\sqrt{N})/N; (S_N + q_{0.005}\sqrt{N})/N] = [5 \times 10^4 - 10^{-1}q_{0.005}, 5 \times 10^4 + 10^{-1}q_{0.005}]$.

(11) $[(S_N - q_{0.005}\sqrt{N})/N\ ;\ (S_N + q_{0.005}\sqrt{N})/N]$ par le Thm de la limite Centrale qu'il fallait énoncer.

(12) Comme l'estimateur $V_N = C_N^2/N - (S_N/N)^2$ converge p.s. vers la variance de ces variables aléatoires par la loi de grands nombres, alors $(S_N - TN)/\sqrt{N \times V_N}$ converge en loi vers la loi Gaussienne par le Thm de la limite centrale. L'intervalle de confiance asymptotique $[(S_N - q_{0.005}\sqrt{NV_N})/N; (S_N + q_{0.005}\sqrt{NV_N})/N]$.

On a $V_N = 8 \times 10^{10} - 25 \times 10^8$, $S_N = 5 \times 10^6$, $N = 100$, il reste a substituter.

(13) Les classes $\{1,2,7\}$ et $\{3,6\}$ sont fermées, la classe $\{4,5\}$ n'est pas fermée.

(14) $\lim_{n\to\infty}(P(T^1 < \infty \mid X_0 = 1))^n = 1$, $\lim_{n\to\infty}(P(T^4 < \infty \mid X_0 = 4))^n = 0$. En effet, l'état 1 est récurrent, donc $P(T^1 < \infty \mid X_0 = 1) = 1$, l'état 4 est transient, donc $P(T^4 < \infty \mid X_0 = 4) < 1$.

(15) Si on note $a_4 = h_4^{\{1,2,7\}}$, $a_5 = h_5^{\{1,2,7\}}$, on a $a_4 = (1/6)a_4 + (2/6)a_5 + 1/2$, $a_5 = (1/4)a_4 + (1/4)a_5$. Alors $a_4 = 9/13$, $a_5 = 3/13$.

Comme les états $4, 5$ sont transients, la CM va finir dans une des deux classes récurrentes, donc on en déduit sans calcul $h_4^{\{3,6\}} = 4/13$, $h_5^{\{3,6\}} = 10/13$.

(16) Comme dans une classe récurrente tous les états sont atteints p.s., alors $h_1^{\{1\}} = h_2^{\{1\}} = h_7^{\{1\}} = 1$.

Pour la même raison $h_4^{\{1,2,7\}} = h_4^{\{1\}} = 9/13$, $h_5^{\{1,2,7\}} = h_5^{\{1\}} = 3/13$.

Evidemment $h_3^{\{1\}} = h_6^{\{1\}} = 0$.

(17) On résout $\vec{\pi} = \vec{\pi}P$ avec $\vec{\pi}$ vecteur ligne. Comme $3, 6$ sont transients, on peut poser des le début $\pi_3 = \pi_6 = 0$ ce qui simplifie le calcul.

On a alors $\vec{\pi} = (9c/21, 4c/21, 2(1-c)/7, 0, 0, 5(1-c)/7, 8c/21)$ avec $c \in [0,1]$.

(18) La classe $\{1,2,7\}$ est récurrente positive et apériodique car $p_{11} > 0$. Donc pour $i = 1,2,7$: $\lim_{n\to\infty} P(X_n = 1 \mid X_0 = i) = 9/21$ la mesure invariante de probabilté de l'état 1 dans la CM restreinte à trois états $\{1,2,7\}$.

On a aussi pour $i = 4,5$: $\lim_{n\to\infty} P(X_n = 1 \mid X_0 = i) = 9/21 \times h_i^{\{1,2,7\}}$, où les valeurs $h_4^{\{1,2,7\}} = 9/13$, $h_5^{\{1,2,7\}} = 3/13$.

Finalement pour $i = 3,6$: $\lim_{n\to\infty} P(X_n = 1 \mid X_0 = i) = 0$.

(19) Pour $i = 4,5$: $P(T^{\{1\}} = \infty \mid X_0 = i) = h_i^{\{3,6\}} > 0$, donc $E(T^{\{1\}} \mid X_0 = i) = \infty$.

Pour $i = 3,6$: $P(T^{\{1\}} = \infty \mid X_0 = i) = 1$, donc $E(T^{\{1\}} \mid X_0 = i) = \infty$.

Pour $i = 1$, $E(T^{\{1\}} \mid X_0 = 1) = 9/2$ – l'inverse de la mesure de proba invariante de l'état 1 dans la CM restreinte à $\{1,2,7\}$.

On note $k_i = E(T^{\{1\}} \mid X_0 = i)$ pour $i = 2,7$. On a $k_2 = 1 + 1/3k_2 + 1/3k_7$, $k_7 = 1 + 1/2k_7$, d'où $k_7 = 2$, $k_2 = 5/2$.

(20) C'est une des mesures invariantes, celle avec $c = 1/3$. Donc $P(X_4 = 2) = 4/63$.