Variables aléatoires absolument continues 2018-2019

Exercice 1. α est un réel strictement positif. Soit la fonction f définie par : $f(x) = \dfrac{\alpha}{x^{\alpha+1}}$ pour $x \geqslant 1$ et $f(x) = 0$ sinon.

1. Démontrer que f est une densité de probabilité. Déterminer la fonction de répartition correspondante.
2. Soit X de densité f. Quelle est la loi de $X^{-\alpha}$? En déduire une méthode pour simuler la loi de X.
3. Calculer l'espérance de X, quand elle existe.
4. On désigne par Y la partie entière de $1+\ln X$. Quelle est la loi de Y ? En déduire son espérance et sa variance.

Exercice 2. Soit a un réel strictement positif. On considère la fonction définie sur $\mathbb{R}$ par : $f(x) = \dfrac{c}{\pi(x^2+a^2)}$.

1. Déterminer la valeur de c pour laquelle f est une densité de probabilité. On désignera dans la suite par X une variable aléatoire qui admet cette densité de probabilité. Montrer que X n'a pas d'espérance.
2. Préciser la fonction de répartition de X. Donner un équivalent simple à $+\infty$ de $P(X > x)$. Quelle est la loi de $\arctan\left(\dfrac{X}{a}\right)$?
3. On considère la variable aléatoire $Y = \ln|X|$. Déterminer une densité de Y. Montrer sans la calculer que Y admet une espérance.

Exercice 3. On définit une fonction g sur $\mathbb{R}$ en posant : $\begin{cases} g(t) = 0 & \text{si } t < 1 \\ g(t) = \dfrac{b}{2^t} & \text{si } t \geqslant 1. \end{cases}$ où b est une constante.

1. Déterminer la constante b pour que g soit une densité d'une variable aléatoire X.
2. Reconnaître la loi de $Y = X-1$. En déduire que X possède une variance et donner cette variance.
3. On note Z la variable aléatoire égale à la partie entière de X. Déterminer la loi de Z.

Exercice 4. La variable aléatoire X suit une loi normale centrée réduite. La fonction `normal(m,s)` du module `numpy.random` permet de simuler une loi normale de paramètres `(m,s)`.

1. Déterminer une densité de X^2.
2. (a) Rappeler la valeur de $E(X^k)$ pour $k \in \{0,1,2\}$.
   (b) Écrire une fonction python qui prend en entrée l'entier k et qui sort une valeur approchée de $E(X^k)$.
   (c) Montrer que pour $k \geqslant 2$, $E(X^k) = (k-1)E(X^{k-2})$.
   (d) Donner la valeur de $E(X^{2k+1})$ et de $E(X^{2k})$ pour $k \in \mathbb{N}$.
3. Pour $n \in \mathbb{N}$ on pose $u_n = \ln(E(X^{2n}))$.
   (a) Montrer que $u_{n+1} - u_n = \ln(2n+1)$. La suite (u_n) peut-elle converger ?
   (b) Écrire une fonction python qui prend en entrée l'entier naturel n et renvoie la valeur de (u_n) et de $\left(\dfrac{u_n}{n\ln(n)}\right)$.
   Conjecturer la limite de la suite $\left(\dfrac{u_n}{n\ln(n)}\right)$.
   (c) Montrer que $\forall k \geqslant 2$, $\displaystyle\int_{k-1}^{k} \ln(2t-1)\,dt \leqslant \ln(2k-1) \leqslant \int_{k}^{k+1} \ln(2t-1)\,dt$

(d) Déterminer la limite de la suite $\left(\dfrac{u_n}{n\ln(n)}\right)$ et un équivalent à u_n.

Exercice 5. Soit $a>0$. On pose : $h(x)=\dfrac{1}{ax\sqrt{2\pi}}\mathrm{e}^{-\frac{(\ln x)^2}{2a^2}}$ si $x>0$ et $h(x)=0$ si $x\leqslant 0$

1. Montrer que h est une densité de probabilité (loi Log-normale).
2. X suit une loi de densité de probabilité h. Calculer $E(X)$. Donner la loi de $Y=\dfrac{\ln X}{a}$.

Exercice 6. Soit X une variable aléatoire à valeurs positives admettant une densité continue sur $\mathbb{R}+$ et ayant une espérance. Montrer que $E(X)=\displaystyle\int_0^{+\infty}P(X>t)dt$.

Exercice 7. (INA 2009) Soient $X_1,\ldots,X_n$ des variables aléatoires réelles indépendantes uniformes sur $[0,a]$, $(a>0)$ $a\in\mathbb{R}^*$.
On pose $U=\max(X_1,\ldots,X_n)$ et $V=\min(X_1,\ldots,X_n)$.

1. Déterminer une densité f_U de U. U admet elle une espérance et une variance ? Si oui les calculer.
2. Mêmes questions pour V.

Exercice 8. Soit f la fonction définie par : $\begin{cases} f(t)=\dfrac{1}{t^2} & \text{si } t\in[1,+\infty[\\ f(t)=0 & \text{sinon}\end{cases}$

1. Montrer que f est une densité de probabilité. Déterminer la fonction de répartition correspondante.
2. Que peut-on dire de l'espérance et de la variance de X ?
3. On pose $Y=\ln X$. Démontrer que Y suit une loi exponentielle dont on précisera le paramètre.

   Soient $X_1,\ldots,X_n$ n variables aléatoires indépendantes de densité f. On définit deux variables aléatoires U et V par : $U=\min(X_1,\ldots,X_n)$ et $V=\max(X_1,\ldots,X_n)$.
4. Pour tout réel t, exprimer l'événement $[U>t]$ à l'aide des événements $[X_1>t],\ldots,[X_n>t]$. En déduire la probabilité $\mathbf{P}([U>t])$.

   Montrer que U est una variable aléatoire dont on donnera une densité g.

   Montrer que U admet une espérance qu'on calculera.
5. Pour tout réel t, exprimer l'événement $[V\leqslant t]$ à l'aide des événements $[X_1\leqslant t]$ ... $[X_n\leqslant t]$. En déduire la probabilité $\mathbf{P}([V\leqslant t])$.

   En déduire que la variable aléatoire V admet pour densité donner une densité h de V.

   Montrer que pour tout réel $t\geqslant 2$, $h(t)\geqslant \dfrac{n}{2^{n-1}t^2}$. En déduire que V n'a pas d'espérance.

Exercice 9. X_1 et X_2 sont deux variables aléatoires indépendantes, suivant la même loi uniforme sur $[0,1]$. On pose $Z=X_1+X_2$.

1. (a) Déterminer, à l'aide de Python, une valeur approchée de $P(Z\leqslant 1)$.
   (b) Montrer que $P(X_1\leqslant X_2)+P(X_2\leqslant X_1)=1$.
   (c) Montrer que $P(X_1\leqslant X_2)=P(X_2\leqslant X_1)$.
   (d) Montrer que $1-X_2$ et X_2 suivent la même loi. En déduire la valeur exacte de $P(Z\leqslant 1)$.
   (e) Représenter l'ensemble $[0,1]\times[0,1]$. Quelle est la signification géométrique de $P(Z\leqslant 1)$?

2. On pose $T = X_1^2 + X_2^2$ et on note f_T une densité de T.

   (a) Montrer que la variable aléatoire X_1^2 est à densité et en déterminer une densité.

   (b) Déterminer $T(\Omega)$ et une expression de $f_T(x)$ pour $x \in]0,1]$. On laissera le résultat sous forme intégrale.

   (c) Soit $x \in]0,1]$. On note $I_x = \displaystyle\int_0^x \frac{1}{\sqrt{t(x-t)}}\,dt$, dont on admet la convergence. Montrer que $\forall x \in]0,1]$, $I_x = I_1$.

   (d) Exprimer $P(T \leqslant 1)$ en fonction de I_x. En déduire à l'aide de Python une valeur approchée de I_1.

Exercice 10. Rappel : Si X et Y sont deux variables aléatoires indépendantes admettant les densités respectives f_X et f_Y alors $S = X + Y$ admet la densité $f_S(s) = \displaystyle\int_{-\infty}^{+\infty} f_X(x) f_Y(s-x)\,dx$.

1. On considère la matrice $A = \begin{pmatrix} 0 & -1 \\ y & 2x \end{pmatrix}$, élément de $\mathcal{M}_2(\mathbb{R})$. Déterminer une condition nécessaire et suffisante portant sur les réels x et y pour que la matrice A soit diagonalisable dans $\mathcal{M}_2(\mathbb{R})$.

   Dans la suite X et Y sont des variables aléatoires réelles, définies sur le même espace de probabilité $(\Omega, \mathcal{A}, P)$, indépendantes et qui suivent toutes les deux la loi uniforme sur $[0,1]$.

2. Écrire un programme informatique pour déterminer une valeur approchée de la probabilité que $X^2 - Y$ soit positif.

3. Montrer que la variable aléatoire $X^2 - Y$ admet pour densité la fonction h définie par $\begin{cases} \sqrt{x+1} \text{ si } -1 \leqslant x < 0 \\ 1 - \sqrt{x} \text{ si } 0 \leqslant x \leqslant 1 \\ 0 \text{ sinon} \end{cases}$

4. Déterminer la probabilité que la matrice $M = \begin{pmatrix} 0 & -1 \\ Y & 2X \end{pmatrix}$ soit diagonalisable dans $\mathbb{R}$.

Exercice 11. La formule du produit de convolution est donnée. Soit une variable aléatoire U de loi uniforme sur] 0,1 [.

1. Déterminer la loi de $V = -\ln U$.

2. Soit X et Y des variables aléatoires indépendantes de loi exponentielle de paramètre 1.

   (a) $-Y$ admet-elle une densité? Si oui, en donner une.

   (b) Écrire une fonction qui simule $-Y$.

3. On pose $Z = X - Y$

   (a) Z admet-elle une densité? Si oui, en donner une. Déterminer l'espérance et la variance de Z.

   (b) Écrire un programme qui simule Z.

4. Déterminer la loi de $|X - Y|$.

Exercice 12. Soit a un réel strictement positif. On pose $\forall x \in \mathbb{R} \quad f_a(x) = \dfrac{a}{\pi(a^2 + x^2)}$.

1. Montrer que f_a est une densité de probabilité. On dit qu'une variable aléatoire admettant une telle densité de probabilité suit une loi de Cauchy de paramètre a.

2. Soit X suivant une loi de Cauchy de paramètre a. Donner la fonction de répartition de X . X a-t-elle une espérance ? Si oui, donner sa valeur.

3. Soit Y suivant une loi de Cauchy de paramètre 1 et a un réel strictement positif. On pose $X = aY$. Donner une densité de X.

4. Soit U suivant une loi uniforme sur $]0,1[$. On pose $Z = \tan(\pi U - \frac{\pi}{2})$. Montrer que Z est bien définie et a pour densité f_1.

5. En utilisant la fonction *random* , écrire une fonction Cauchy(a) permettant de simuler une variable aléatoire suivant une loi de Cauchy de paramètre a.

6. On souhaite maintenant examiner le comportement de la somme de deux variables aléatoires indépendantes suivant des lois de Cauchy de paramètres respectifs a et b.

   (a) La fonction `hist` de matplotlib.pyplot permet de construire un histogramme :
   `hist(L,bins=20,range=(-10*a,10*a))`.
   Écrire une fonction permettant de construire l'histogramme de $X+Y2$ et le comparer avec l'histogramme d'une variable de loi de Cauchy de paramètre $a+b$.

   (b) Faire afficher sur le même graphique la courbe représentative de f_{a+b}.

   (c) On veut démontrer que la somme de deux variables aléatoires à densité indépendantes suivant des lois de Cauchy de paramètres respectifs $a>0$ et $b>0$ suit une loi de Cauchy de paramètre $a+b$.

   (d) Pour x et t réels, on admet qu'on peut écrire $\dfrac{1}{(a^2+t^2)(b^2+(x-t)^2)} = \dfrac{\alpha t+\beta}{a^2+t^2} + \dfrac{\gamma(x-t)+\delta}{b^2+(x-t)^2}$
   avec $\beta = \dfrac{x^2+b^2-a^2}{(x^2+(a+b)^2)(x^2+(a-b)^2)}$ et $\delta = \dfrac{x^2+a^2-b^2}{(x^2+(a+b)^2)(x^2+(a-b)^2)}$ et $\alpha=\gamma$.
   En déduire le résultat voulu. La loi faible des grands nombres est-elle vérifiée ?

Convergences

Exercice 13. Soit Φ la fonction de répartition de la loi normale centrée réduite. Utiliser l'inégalité de Bienaymé Tchebychev pour montrer que : $\forall x>0,\ 1-\Phi(x) \leqslant \dfrac{1}{2x^2}$

Exercice 14.

1. X suit une loi de Poisson de paramètre λ. Calculer les valeurs de k pour lesquelles $P(X=k)$ est maximale.

2. Démontrer que X est plus souvent paire qu'impaire.

3. Soit $(X_k)_{k\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes telle que pour tout $k\in\mathbb{N}^*$, la variable X_k suit une loi de Poisson de paramètre $\lambda>0$. Soit $n\in\mathbb{N}^*$. Quelle est la loi de $S_n = X_1+\ldots+X_n$? Donner son espérance et sa variance. Énoncer le théorème de la central limite pour la suite $(X_k)_{k\in\mathbb{N}^*}$.

4. Démontrer que $\displaystyle\lim_{n\to+\infty}\sum_{k=0}^{n} \mathrm{e}^{-n}\frac{n^k}{k!} = \frac{1}{2}$.

Exercice 15. Soit θ un réel strictement positif. On considère n variables aléatoires indépendantes $X_1,\ldots,X_n$ suivant une loi uniforme sur $[0,\theta]$, θ étant un réel inconnu qu'on veut estimer.

1. On pose $M_n = \max(X_1,\ldots,X_n)$ et $M_n' = \dfrac{n+1}{n}M_n$. Démontrer que M_n est une variable aléatoire ayant une densité qu'on déterminera. Calculer $E(M_n), V(M_n), E(M_n'), V(M_n')$.
   Montrer que $\forall\varepsilon>0,\ \displaystyle\lim_{n\to+\infty} P(|M_n-\theta|>\varepsilon)=0$ et $\forall\varepsilon>0,\ \displaystyle\lim_{n\to+\infty} P(|M_n'-\theta|>\varepsilon)=0$.

2. On pose $Z_n = n\left(1-\dfrac{M_n}{\theta}\right)$. Déterminer $\displaystyle\lim_{n\to+\infty} F_{Z_n}(x)$ pour $x>0$.

3. On pose $Y_n = \dfrac{2}{n}\sum_{k=1}^{n} X_k$. Déterminer l'espérance et la variance de Y_n et montrer que $\forall \varepsilon > 0$, $\lim\limits_{n\to+\infty} P(|Y_n - \theta| > \varepsilon) = 0$.

4. Entre M'_n et Y_n, quel estimateur de θ choisissez-vous ?

Exercice 16. On considère n variables aléatoires indépendantes suivant la même loi de Poisson de paramètre inconnu λ.

1. On pose $S_n = \sum_{k=1}^{n} X_k$ et $\overline{X_n} = \dfrac{1}{n} S_n$. Rappeler la loi de S_n. Montrer que $\overline{X_n}$ est un estimateur de λ sans biais et que $\lim\limits_{n\to+\infty} V(\overline{X_n}) = 0$.

2. On pose $\theta = P(X = 0) = \mathrm{e}^{-\lambda}$ et on cherche un estimateur sans biais de θ. Calculer l'espérance de $T_n = \mathrm{e}^{-\overline{X_n}}$. T_n est-il un estimateur sans biais de θ ?

3. On considère la variable aléatoire $\widehat{\theta} = \left(1 - \dfrac{1}{n}\right)^{S_n}$. Montrer que $\widehat{\theta}$ est un estimateur sans biais de θ. Calculer sa variance et montrer que $\lim\limits_{n\to+\infty} V(\widehat{\theta}) = 0$.

Exercice 17. Pour estimer la proportion p de malades dans une population de N individus, on effectue un sondage anonyme suivant le protocole suivant : chaque personne interrogée rentre dans un isoloir avec une pièce de monnaie équilibrée. La personne lance la pièce :
• si elle obtient pile, elle sort de l'isoloir et répond "oui" si elle est malade et "non" sinon
• si elle obtient face, elle relance la pièce une seconde fois et répond "oui" si elle a obtenu pile et "non" si elle a obtenu face.
Pour $i = 1, 2, \ldots, N$, on note X_i la variable aléatoire qui prend la valeur 1 si la personne i s'est déclarée malade et 0 sinon.

1. Écrire un programme Python prenant p en paramètre, simulant le sondage d'une personne et renvoyant 1 si la personne se déclare malade et 0 sinon.

   Écrire un programme Python prenant p et N en paramètres, simulant le sondage des N individus et renvoyant la proportion de personnes malades dans la population estimée par cette méthode.

2. Soit$F_N = \dfrac{1}{N}\sum_{k=1}^{N} X_k$. Soit $\pi = \dfrac{p}{2} + \dfrac{1}{4}$. Exprimer l'espérance et la variance de F_N en fonction de π et N.

   Montrer que $\pi(1-\pi) \leqslant \dfrac{1}{4}$ puis : $P(|F_N - \pi| \geqslant \varepsilon|) \leqslant \dfrac{1}{4N\varepsilon^2}$.

3. Déterminer ε_0 tel que $\dfrac{1}{4N\varepsilon_0^2} = 0,05$ puis montrer que $P\left(2F_N - 2\varepsilon_0 - \dfrac{1}{2} \leqslant p \leqslant 2F_N + 2\varepsilon_0 - \dfrac{1}{2}\right) \geqslant 0,95$.

4. Écrire un programme Python prenant p et N en paramètres et renvoyant les bornes du segment $\left[2F_N - 2\varepsilon_0 - \dfrac{1}{2}, 2F_N \right.$

   Comment s'appelle un tel intervalle ?

Exercice 18. Dans tout l'exercice, l'unité de temps est l'année. Une entreprise utilise N centrifugeuses. Pour poursuivre son activité, l'entreprise a besoin de N ou $N-1$ centrifugeuses qui fonctionnent. La durée de fonctionnement en année d'une centrifugeuse suit une loi exponentielle et la durée moyenne de vie est de $\tau = 2$ ans. On note :
T_k : la durée de fonctionnement de la k ième machine pour $k \in \{1, \ldots, n\}$
X_t : le nombre de centrifugeuses qui fonctionnent à l'instant t pour $t \in [0, +\infty[$
U : la durée de fonctionnement de l'entreprise en année
On considère pour $N \in \mathbb{N}^*$ et $x \in [0,1]$ $G_N(x) = 1 - Nx^{N-1} + (N-1)x^N$.

1. (a) Montrer que pour tout $y \in [0,1]$, il existe un unique $x \in [0,1]$ tel que $G_N(x) = y$.

   (b) Réaliser une fonction Python qui prend en entrée y, N et qui renvoie en utilisant l'algorithme de dichotomie rappelé une valeur de x à 10^{-3} près vérifiant $y = G_N(x)$.

   (c) Pour $N = 11$ et $\varepsilon = 0.025$ déterminer t_1 et t_2 vérifiant $G_N\left(\mathrm{e}^{-\frac{t_1}{\tau}}\right) = \varepsilon$ et $G_N\left(\mathrm{e}^{-\frac{t_2}{\tau}}\right) = 1 - \varepsilon$.

   On donnera des valeurs approchées de t_1 et t_2 à 10^{-3} près.

2. Déterminer la probabilité pour qu'une machine fonctionne à l'instant t pour $t \in [0, +\infty[$.

3. En supposant les T_k indépendantes, donner la loi de X_t.

4. En déterminant la fonction de répartition de U, montrer que U est à densité.

5. Sachant $N = 11$, donner un intervalle de confiance de U à 95%.

Exercice 19. On cherche à évaluer le nombre N de lions d'Asie, espèce en voie de disparition, encore en vie dans la forêt de Gir. Pour cela, on capture d'abord, en une seule fois, m lions (avec $m \in \mathbb{N}^*$) que l'on tatoue avant de les relâcher dans la nature, et on admet que pendant toute la durée de l'étude il n'y a ni décès ni naissance, puis on utilise l'une des deux méthodes suivantes

Première méthode

On capture successivement, au hasard (donc avec équiprobabilité) et avec remise en liberté après l'observation du sujet, n lions. Soit Y_n le nombre de lions tatoués parmi eux.

1. Déterminer la loi de Y_n. En déduire que $\dfrac{1}{nm} Y_n$ est un estimateur sans biais de $\dfrac{1}{N}$.

   Justifier que $\forall \varepsilon > 0, \ \lim\limits_{n \to +\infty} P\left(\left|\dfrac{1}{nm} Y_n - \dfrac{1}{N}\right| > \varepsilon\right) = 0$

2. Pourquoi ne peut-on pas prendre $\dfrac{nm}{Y_n}$ comme estimateur de N ?

3. On pose $B_n = \dfrac{m(n+1)}{Y_n + 1}$. Calculer l'espérance de B_n et sa limite quand $n \to +\infty$.

Seconde méthode

On se donne $n \in \mathbb{N}^*$. On capture également, un par un, des lions de Gir au hasard et avec remise en liberté après l'observation du sujet. On note X_n, la variable aléatoire égale au nombre de lions qu'il a été juste nécessaire de capturer pour en obtenir n tatoués. On pose $D_1 = X_1$ et pour tout $i \in [\![2, n]\!]$, $D_i = X_i - X_{i-1}$. On admet que les D_i sont des variables indépendantes deux-à-deux.

1. (a) Pour tout $i \in [\![2, n]\!]$, que représente concrètement D_i ?

   (b) Déterminer, pour tout $i \in [\![2, n]\!]$, la loi de D_i, son espérance et sa variance. En déduire l'espérance et la variance deX_n.

   (c) On pose $A_n = \dfrac{m}{n} X_n$. Montrer que A_n est un estimateur sans biais et convergent de N et déterminer sa variance.

2. (a) Pour n assez grand, par quelle loi peut-on approcher la loi de la variable aléatoire $\widetilde{X_n} = \dfrac{X_n}{n}$?

   (b) On a tatoué $m = 200$ lions, puis capturé 450 lions, pour obtenir $n = 50$ lions marqués. On note σ l'écart-type de A_{50}. On a pu prouver que $\sigma \leqslant 100$. Déterminer en fonction de σ, un intervalle de confiance pour N au seuil de confiance 0,9 (on rappelle que $\Phi(1,64) \approx 0,95$).