

Introduction aux variables aléatoires à densité

Jusqu'à présent, toutes nos expériences aléatoires mettaient en évidence des variables aléatoires à valeurs discrètes (souvent à valeurs dans $\mathbb{N}$).

Mais si l'on étudie X la taille en mètres des individus adultes, ou Y la durée de vie d'un appareil électronique en heures, on voit que $X(\Omega)$ comme $Y(\Omega)$ sont des intervalles de réels (ou réunion d'intervalles) : plus précisément, on pourrait prendre par exemple $X(\Omega) = [1, 2.5]$ et $Y(\Omega) = \mathbb{R}_+$.

Il faut donc développer un cadre qui permette d'étudier de telles variables aléatoires à valeurs réelles.

Pour introduire ces variables réelles, on va partir de la loi uniforme discrète :

$X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$ si

$$\text{et } \sum_{k=1}^n P(X = k) =$$

→ le diagramme en bâtons de cette loi est :

Quel sens pourrait-on donner à $X \hookrightarrow \mathcal{U}([0, 1])$?

→ "Qu'est-ce qui doit faire 1" ?

→ Que pourrait valoir $P(X = x)$?

Proposition Caractérisation des densités

Toute fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ vérifiant les propriétés suivantes :

1. f est positive sur $\mathbb{R}$
2. f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points
3. L'intégrale $\int_{-\infty}^{+\infty} f(t)dt$ converge et vaut 1.

est la densité d'une variable aléatoire.

→ Vérifier que la fonction définie dans l'exemple introductif est bien une densité.

Mais qu'est-ce qu'une variable à densité ?

DÉFINITION

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une variable aléatoire réelle. On dit que X est une variable à densité, si sa fonction de répartition F_X est :

- (i) continue sur $\mathbb{R}$
- (ii) de classe C^1 sur $\mathbb{R}$, sauf éventuellement en un nombre fini de points.

Toute fonction $f_X : \mathbb{R} \rightarrow \mathbb{R}$ à valeurs positives telle que $f_X(x) = F'_X(x)$ pour tout réel x sauf éventuellement en un nombre fini de points, est appelée densité de X .

→ en pratique, pour définir f_X sur $\mathbb{R}$ tout entier, on choisit des valeurs arbitraires (≥ 0) là où F_X n'est pas C^1 .

Propriétés fondamentales : lien entre densité, fonction de répartition, calcul de probabilités.

Soit X une variable à densité de fonction de répartition F_X et de densité f_X . Alors

1. $\forall x \in \mathbb{R}, F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t)dt.$
2. $P(X > x) = \int_x^{+\infty} f_X(t)dt.$
3. $\forall x \in \mathbb{R}, P(X = x) = 0.$
4. pour tout $(x, y) \in \mathbb{R}^2$ tels que $x \leq y$,
 $P(x \leq X \leq y) = P(x < X \leq y) = P(x < X < y) = P(x \leq X < y) = F_X(y) - F_X(x) = \int_x^y f_X(t)dt$

→ idée de la preuve lorsque F_X est C^1 sur $\mathbb{R}$, et $f_X = F'_X$ est continue sur $\mathbb{R}$.

Retour exemple d'introduction : Soit $X \hookrightarrow \mathcal{U}([0, 1])$. Déterminer sa fonction de répartition, et vérifier qu'elle est bien continue sur $\mathbb{R}$ et de classe C^1 sur $\mathbb{R}$ sauf en un nombre fini de points (à déterminer).

Détermination de $X(\Omega) = \{X(\omega), \omega \in \Omega\}$, ensemble des valeurs prises par X :

La question n'est pas simple car si X est une variable à densité de densité f_X , pour tout $x \in \mathbb{R}$, $P(X = x) = 0$.
On ne peut donc pas prendre le même "critère" que pour une variable discrète.

→ *Illustration graphique à l'aide de la propriété 4. précédente*

$$\text{pour tout } h \geq 0 \text{ petit, } \underbrace{P(x-h \leq X \leq x+h)}_{\text{probabilité que } X \text{ prenne une valeur proche de } x} = \underbrace{\int_{x-h}^{x+h} f_X(t) dt}_{\text{aire sous la courbe de } f_X \text{ entre } x-h \text{ et } x+h}$$

Ainsi, plus la densité f_X est élevée autour de x , plus l'aire précédente sera grande, et plus la probabilité d'obtenir une valeur proche de x est forte. A contrario, si la densité est nulle au voisinage de x , l'aire le sera aussi, et la probabilité que X prenne une valeur au voisinage de x est nulle : x ne sera pas valeur prise.

Vu les lois à votre programme, on peut ainsi simplifier la définition : presque sûrement $X(\Omega)$ $= \{t \in \mathbb{R} / f_X(t) \neq 0\}$.

En pratique, on pourra écrire $X(\Omega)$ à la lecture d'une densité f_X , sans le justifier, mais avec un presque sûrement devant (une densité n'est pas unique!).

Exemple :

Soit $\lambda > 0$, et soit X une variable aléatoire à densité de densité $f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases}$.

1. Vérifier que f ainsi définie est bien une densité de probabilité. La dessiner.
2. Préciser $X(\Omega)$ ainsi que l'intervalle de longueur 1/2 contenant les valeurs les plus probables pour X .
3. Déterminer la fonction de répartition de X .

Espérance des variables à densité

DÉFINITION

Soit X une variable à densité, de densité f_X . On dit que X admet une espérance si l'intégrale $\int_{-\infty}^{+\infty} t f_X(t) dt$ converge absolument et dans ce cas on note $E(X) = \int_{-\infty}^{+\infty} t f_X(t) dt$ son espérance.

Point méthode : "Montrer que l'espérance existe et déterminer sa valeur"

Pour ne pas faire deux fois l'étude (1^{ère} fois avec des valeurs absolues pour l'existence, 2^e fois sans pour le calcul), découper l'intégrale $\int_{-\infty}^{+\infty} t f(t) dt$ en deux. L'intégrale $\int_0^{+\infty} t f(t) dt$ est d'intérieur positif donc la cv absolue revient à la cv, et l'intégrale $\int_{-\infty}^0 t f(t) dt$ est d'intérieur négatif donc par linéarité, la cv absolue revient à la convergence.

→ On verra également un raccourci lorsque X est presque sûrement bornée

Point méthode : "Montrer que l'espérance existe"

Dans ce cas, une seule étude, donc pas de problème avec les valeurs absolues. Penser aux 3 critères de cv.

Exemples Reprendre l'exemple d'introduction ainsi que l'exemple ci-dessus et déterminer si l'espérance existe. Préciser le cas échéant sa valeur.

Cas particulier important : si X est presque sûrement bornée (autrement dit si $X(\Omega) \subset [a, b]$, avec a, b deux réels), alors X admet une espérance.

→ idée de la preuve lorsque X est à valeurs positives.

Comme f est nulle en-dehors de $[a, b]$, $E(X)$ existe dès que $\int_a^b t f(t) dt$ cv absolument, donc converge puisque X est à valeurs positives donc $[a, b] \subset \mathbb{R}^+$.

Or pour tout $t \in [a, b]$, $0 \leq t f(t) \leq b f(t)$ puisque $t \leq b$. Comme f est une densité $\int_{-\infty}^{+\infty} f$ cv et vaut 1, donc $\int_a^b f(t) dt$ converge. Par linéarité, l'intégrale $\int_a^b b f(t) dt$ converge, donc par critère de comparaison, l'intégrale $\int_a^b t f(t) dt$ converge.

Récapitulatif de la loi uniforme sur $[0,1]$ notée $\mathcal{U}([0,1])$:

$X \hookrightarrow \mathcal{U}([0,1])$ si X est une variable à densité de densité $f(x) = \begin{cases} 1 & \text{si } x \in [0,1] \\ 0 & \text{sinon} \end{cases}$

La fonction de répartition d'une loi $\mathcal{U}([0,1])$ est la fonction F définie sur $\mathbb{R}$ par $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$

Si $X \hookrightarrow \mathcal{U}([0,1])$, alors X admet une espérance et $E(X) = \frac{1}{2}$.

Récapitulatif de la loi exponentielle de paramètre $\lambda > 0$ notée $\mathcal{E}(\lambda)$

Dans la définition suivante, j'ai choisi $X(\Omega) = \mathbb{R}^+$, mais on aurait pu prendre $X(\Omega) = \mathbb{R}_+^*$...

$X \hookrightarrow \mathcal{E}(\lambda)$, si X est une variable à densité de densité $f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$.

La fonction de répartition d'une loi $\mathcal{E}(\lambda)$ est la fonction F définie sur $\mathbb{R}$ par $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$

Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors X admet une espérance et $E(X) = \frac{1}{\lambda}$.

Feuille d'exercices 25 : Variables aléatoires à densité**2022/2023****Exercice 1:**

Montrer que les fonctions suivantes, définies sur $\mathbb{R}$, sont des densités de probabilité et les dessiner.

En notant X une variable aléatoire de densité f_i , préciser $X(\Omega)$ puis déterminer la fonction de répartition associée F_i , et étudier l'espérance de X .

$$f_1(x) = \begin{cases} \cos x & \text{si } x \in [0, \frac{\pi}{2}] \\ 0 & \text{sinon} \end{cases} \quad f_2(x) = \begin{cases} 1+x & \text{si } x \in [-1, 0] \\ 1-x & \text{si } x \in]0, 1] \\ 0 & \text{sinon} \end{cases} \quad f_3(x) = \begin{cases} \frac{1}{|x|^3} & \text{si } |x| \geq 1 \\ 0 & \text{sinon} \end{cases}$$

$$f_4(x) = \frac{e^x}{(e^x + 1)^2} \text{ pour } x \in \mathbb{R} \quad (\text{pour l'esp, on pourra remarquer que } f_4 \text{ est paire ...})$$

Exercice 2:

Soit la fonction f définie sur $\mathbb{R}$ par $f(x) = \frac{a}{x^2 + 1}$.

1. Dessiner f .
2. Déterminer le réel a pour que f soit une densité de probabilité.
On introduit alors une variable aléatoire X de densité f : on dit que X suit la loi de Cauchy.
3. Préciser $X(\Omega)$ puis montrer que X n'admet pas d'espérance.
4. Montrer que la fonction de répartition de X est donnée par : $\forall x \in \mathbb{R}, F(x) = \frac{1}{\pi} \arctan(x) + \frac{1}{2}$.

Exercice 3: pour s'entraîner

Soit la fonction f définie sur $\mathbb{R}$ par $f(x) = \begin{cases} \frac{a}{x\sqrt{x}} & \text{si } x \geq 1 \\ 0 & \text{sinon} \end{cases}$.

1. Déterminer le réel a pour que f soit une densité de probabilité d'une certaine variable aléatoire X .
2. Déterminer la fonction de répartition associée à X .
3. Montrer que X n'admet pas d'espérance.

Exercice 4:

Soit la fonction f définie sur $\mathbb{R}$ par $f(x) = \begin{cases} e^x & \text{si } x < 0 \\ 0 & \text{sinon} \end{cases}$

1. Montrer que f est une densité de probabilité d'une certaine variable aléatoire, que l'on notera X . On dessinera f , et on précisera $X(\Omega)$.
2. Déterminer la fonction de répartition associée.
3. Montrer que X admet une espérance et la calculer.
4. On pose $Y = 2X + 1$ et on admet que Y est bien une variable aléatoire.
 - (a) Déterminer la fonction de répartition notée G de Y
 - (b) Montrer alors que Y est une variable à densité et déterminer une densité g de Y .
 - (c) Mêmes questions avec $Z = X^2$.

Exercice 5: pour s'entraîner

Reprendre l'énoncé de l'exercice précédent avec $f(x) = \begin{cases} \ln x & \text{si } x \in [1, e] \\ 0 & \text{sinon} \end{cases}$ et $Y = \ln X$.

Exercice 6:

$$\text{Soit } f(x) = \begin{cases} e^{-|x|} & \text{si } -\ln 2 \leq x \leq \ln 2 \\ 0 & \text{sinon} \end{cases}$$

1. Montrer que f est une densité de probabilité d'une certaine variable aléatoire, que l'on notera X .
2. Déterminer la fonction de répartition associée.
3. Montrer que X admet une espérance et la calculer.
4. On pose $Y = |X|$ et on admet que Y est bien une variable aléatoire.
 - (a) Déterminer la fonction de répartition de G de Y
 - (b) Montrer alors que Y est une variable à densité et déterminer une densité g de Y .

Exercice 7: *pour aller plus loin*

$$\text{Soit } a \in \mathbb{R} \text{ et } f \text{ la fonction définie sur } \mathbb{R} \text{ par } f(x) = \begin{cases} \frac{a}{x+1} & \text{si } x \in]0, 1[\\ 0 & \text{sinon} \end{cases}$$

1. Déterminer le réel a pour que f soit une densité de probabilité d'une certaine variable aléatoire X .
2. Déterminer la fonction de répartition de X et son espérance, si elle existe.
3. ** On pose $Y = \frac{1}{X}$ et $N = \lfloor \frac{1}{X} \rfloor = \lfloor Y \rfloor$ et on admet que Y et N sont des variables aléatoires.
 - (a) Vérifier que Y est bien une variable à densité et préciser une densité. Y admet-elle une espérance?
 - (b) Déterminer la loi de N . N admet-elle une espérance?

Exercice 8:

Soit $X \hookrightarrow \mathcal{U}([0, 1])$. Reconnaître la loi de $T = -\frac{1}{\lambda} \ln(1 - X)$ (où $\lambda > 0$).

Application python : à l'aide de la syntaxe `rd.random()`, simuler une loi exponentielle de paramètre 5.

Exercice 9:

Soit $X \hookrightarrow \mathcal{E}(\lambda)$ avec $\lambda > 0$. On admet que $T = X^2$, et $U = \lfloor X \rfloor + 1$ sont bien des variables aléatoires définies sur le même espace probabilisé que X .

1. Rappeler une densité de X , sa fonction de répartition, ainsi que la valeur de son espérance.
2. Vérifier que T est bien une variable aléatoire à densité, et en déterminer une densité.
3. Reconnaître la loi de U . En déduire son espérance.

Exercice 10:

Soit la fonction f définie sur $\mathbb{R}$ par $f(x) = \frac{1}{2(1+|x|)^2}$.

1. Vérifier que f peut être considérée comme la densité de probabilité d'une certaine variable aléatoire X .
On note F sa fonction de répartition. (On ne calculera pas F).
2. On pose $Y = \ln(1 + |X|)$ et on admet que Y est une variable aléatoire.
 - (a) Exprimer la fonction de répartition de Y , notée G en fonction de F .
 - (b) En déduire que Y est une variable à densité, et exprimer sa densité en fonction de f .
 - (c) Reconnaître la loi de Y .

Exercice 11: *pour s'entraîner*

Reprendre les questions de l'exercice précédent avec $f(x) = \begin{cases} \frac{1}{2x^2} & \text{si } |x| \geq 1 \\ 0 & \text{sinon} \end{cases}$ et $Y = \ln(|X|)$.

bonus : déterminer F .

Exercice 12:

On désigne par a un réel et on considère une variable aléatoire X , de densité f strictement positive et continue sur $\mathbb{R}$, dont la fonction de répartition est notée F .

On pose alors pour tout $x \in \mathbb{R}$, $g(x) = \begin{cases} \frac{f(x)}{F(a)} & \text{si } x \leq a \\ 0 & \text{si } x > a \end{cases}$

1. Montrer que g est bien définie et peut être considérée comme la densité d'une certaine variable aléatoire Y .
2. On note G la fonction de répartition de Y .
 - (a) Exprimer, pour tout réel x , $G(x)$ à l'aide de F .
 - (b) Vérifier que l'on a pour tout réel x : $G(x) = P_{(X \leq a)}(X \leq x)$.