

Variables aléatoires à densité

Jusqu'à présent, toutes les expériences aléatoires rencontrées 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 :

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

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

→ 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 $P(X = x)$?

1 Généralités

1.1 Définitions

DÉFINITION

On appelle densité de probabilité 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 généralisée $\int_{-\infty}^{+\infty} f(t)dt$ converge et vaut 1.

→ vérifier que la fonction définie dans l'exemple introductif est bien une densité de probabilité.

DÉFINITION

Soit X une variable aléatoire réelle définie sur (Ω, \mathcal{A}, P) et de fonction de répartition F_X . On dit que X est une variable à densité lorsqu'il existe une densité de probabilité f telle que pour tout $x \in \mathbb{R}$, $F_X(x) = \int_{-\infty}^x f(t)dt$.

Remarque 1 : il n'y a pas unicité d'une densité de X . Toute fonction g positive qui ne diffère de f qu'en un nombre fini de points, sera également une densité de X . Mais il y a unicité de F_X !

Remarque 2 : cette définition permet d'obtenir des propriétés sur les variables à densité, mais elle ne sera pas utilisée en pratique pour savoir si une variable est à densité. On lui préférera la caractérisation de la section 1.3..

Proposition admise

Soit f une densité de probabilité. Alors il existe un espace probabilisé (Ω, \mathcal{A}, P) et une variable aléatoire X définie sur Ω telle que X est une variable aléatoire de densité f .

Premières propriétés

Proposition

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

1. F_X est croissante sur \mathbb{R}
2. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
3. F_X est continue sur \mathbb{R} .
4. F_X est dérivable en tout point x de continuité de f et $F'_X(x) = f(x)$.

Rappel : les deux premières propriétés sont vraies pour toutes les variables aléatoires (même discrètes). En revanche, les deux dernières propriétés caractérisent les variables aléatoires à densité (cf section 1.3.)

Proposition

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

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

→ preuve

Exemple 1 :

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

1. Déterminer λ pour que f soit une densité d'une variable aléatoire X , puis représenter la courbe de f .
2. Déterminer la fonction de répartition F_X de X , puis représenter la courbe de F_X .

1.2 Support d'une variable à densité

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*

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

probabilité que X prenne une valeur proche de x

Ainsi, plus la densité f_X prend des valeurs élevées 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 (autrement dit avec probabilité 1) $X(\Omega)$ = $\{t \in \mathbb{R} / f_X(t) \neq 0\}$.

En pratique, on pourra écrire un $X(\Omega)$ à la lecture d'une densité f_X , sans le justifier.

1.3 Caractérisation des variables à densité

Théorème

Soit (Ω, \mathcal{A}, P) un espace probablisé, X une variable aléatoire réelle de fonction de répartition F_X . Alors X est une variable à densité ssi F_X est continue sur \mathbb{R} et de classe C^1 sur \mathbb{R} , sauf éventuellement en un nombre fini de points.

Le cas échéant, 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 une 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 .

Exemple 2 :

Soit X une variable aléatoire sur un espace probablisé (Ω, \mathcal{A}, P) dont la fonction de répartition est

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } 1 < x \end{cases} . \text{ Vérifier que } X \text{ est bien une variable à densité, et préciser une densité de } X.$$

1.4 Fonction d'une variable à densité

L'étude générale des variables de type $Y = f(X)$ où X est une variable à densité et f une fonction définie sur $X(\Omega)$, n'est pas au programme : remarquer que selon f , la variable aléatoire Y n'est pas toujours à densité. Plusieurs cas seront étudiés dans les exercices : $Y = X^2, Y = |X|, Y = \ln(X), Y = \lfloor X \rfloor \dots$

Méthode : après avoir précisé $Y(\Omega)$ déterminer pour tout $x \in \mathbb{R}, P(Y \leq x)$ en fonction de $P(X \leq x) \dots$

Exercice : Soit X une variable à densité, de densité f_X et de fonction de répartition F_X .

Montrer que pour tous réels a, b , avec $a \neq 0$, la variable $Y = aX + b$ est encore une variable à densité.

1.5 Indépendance : rappels et compléments

DÉFINITION

Soit X et Y deux variables aléatoires à densité définies sur un même espace (Ω, \mathcal{A}, P) .

Alors X et Y sont indépendantes si et seulement si

$$\text{pour tous intervalle } I, J \text{ de } \mathbb{R}, \quad P((X \in I) \cap (Y \in J)) = P(X \in I) \times P(Y \in J).$$

Soit une famille finie (X_1, X_2, \dots, X_n) de variables aléatoires à densité définies sur un même espace (Ω, \mathcal{A}, P) .

Ces variables sont mutuellement indépendantes ssi pour tous intervalles de \mathbb{R} I_1, I_2, \dots, I_n ,

$$P((X_1 \in I_1) \cap (X_2 \in I_2) \cap \dots \cap (X_n \in I_n)) = \prod_{k=1}^n P(X_k \in I_k).$$

Propriétés de l'indépendance mutuelle :

- Toute sous-famille d'une famille de variables mutuellement indépendantes est mutuellement indépendante.
- Lemme des coalitions :
Soit une famille $(X_1, \dots, X_n, X_{n+1}, \dots, X_{n+p})$ de variables aléatoires mutuellement indépendantes. Alors pour toutes fonctions u et v , définies sur \mathbb{R}^n et \mathbb{R}^p , les variables $u(X_1, \dots, X_n)$ et $v(X_{n+1}, \dots, X_{n+p})$ sont indépendantes.
- Le lemme des coalitions peut se généraliser à une partition plus grande des variables : en particulier si (X_1, \dots, X_n) est une famille de variables mutuellement indépendantes, pour toutes fonctions u_1, \dots, u_n définies sur \mathbb{R} , les variables $u_1(X_1), \dots, u_n(X_n)$ sont encore mutuellement indépendantes.

2 Moments d'une variable à densité

2.1 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.

Point méthode : "Montrer que l'espérance existe"

Dans ce cas, une seule étude, donc on peut laisser les valeurs absolues. Penser aux 2 critères de convergence.

Exemples Reprendre les *Exemples 1. et 2.* et déterminer si l'espérance existe. Préciser le cas échéant sa valeur.

Linéarité de l'espérance :

Si X admet une espérance, alors pour tout $(a, b) \in \mathbb{R}^2$, $aX + b$ en admet une et $E(aX + b) = aE(X) + b$.

Si X et Y admettent toutes deux une espérance, alors $X + Y$ en admet une et $E(X + Y) = E(X) + E(Y)$.

Remarque : la linéarité reste vraie que les variables soient discrètes ou à densité, et sans savoir si leur résultante est discrète ou à densité ...

Indépendance et espérance :

Si X et Y admettent une espérance et sont indépendantes, alors XY admet une espérance et $E(XY) = E(X)E(Y)$.

Théorème Formule de transfert

Soit X une variable à densité de densité f , et soit u une fonction définie sur un intervalle I contenant $X(\Omega)$ et continue sauf éventuellement en un nombre fini de points.

Alors $u(X)$ admet une espérance si, et seulement si, l'intégrale $\int_{-\infty}^{+\infty} u(x)f(x)dx$ est absolument convergente.

Dans ce cas, $E(u(X)) = \int_{-\infty}^{+\infty} u(x)f(x)dx$.

Remarque

L'intégrale $\int_{-\infty}^{+\infty} u(x)f(x)dx$ a bien un sens, même si u n'est pas définie sur \mathbb{R} car f est nulle en dehors de $X(\Omega)$.

Plus précisément, on a : $\int_{-\infty}^{+\infty} u(x)f(x)dx = \int_I u(x)f(x)dx$.

Cas particulier où $u : x \mapsto x^2$.

X^2 admet une espérance si et seulement si l'intégrale $\int_{-\infty}^{+\infty} x^2 f(x) dx$ converge et dans ce cas, $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$.

En effet, la convergence absolue revient ici à la convergence puique pour tout $x \in \mathbb{R}$, $x^2 f(x) \geq 0$.

2.2 Moment d'ordre 2 et variance

Section identique à celle pour les variables discrètes.

DÉFINITION

Soit X une variable à densité sur (Ω, \mathcal{A}, P) . Sous-réserve d'existence, on appelle

1. moment d'ordre 2 de X le réel $E(X^2)$.
2. variance de X le réel $V(X) = E[(X - E(X))^2]$ (mesure l'écart entre X et sa moyenne)
3. écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$

Remarque

X peut admettre une espérance sans admettre de moment d'ordre 2. Mais si X admet un moment d'ordre 2, alors X admet une espérance, et donc par contraposée, si X n'admet pas d'espérance alors X n'admet pas non plus de moment d'ordre 2.

Proposition Formule de Koenig-Huygens

X admet un moment d'ordre 2 ssi X admet une variance et dans ce cas, $V(X) = E(X^2) - (E(X))^2$.

Propriétés de la variance

Proposition

Soit X et Y deux variables à densité sur (Ω, \mathcal{A}, P) admettant une variance et a, b deux nombres réels. Alors

1. $V(X) \geq 0$.
2. La variable $aX + b$ admet une variance et $V(aX + b) = a^2 V(X)$.
3. Si de plus X et Y sont indépendantes, alors $V(X + Y) = V(X) + V(Y)$.

Le dernier point se généralise : soit (X_1, \dots, X_n) n variables à densité indépendantes et admettant toutes une variance.

Alors $X_1 + \dots + X_n$ admet une variance et $V(X_1 + X_2 + \dots + X_n) = V(X_1) + V(X_2) + \dots + V(X_n)$.

DÉFINITION

Une variable X est dite centrée lorsque $E(X) = 0$ et réduite lorsque $V(X) = 1$.

Proposition

Soit X une variable admettant une variance non-nulle (donc X n'est pas constante). Alors,

la variable $X^* = \frac{X - E(X)}{\sigma(X)}$ est une variable centrée réduite, appelée variable centrée réduite associée à X .

3 Lois usuelles à densité

3.1 Loi uniforme

DÉFINITION

On dit que X suit la loi uniforme sur l'intervalle $[a, b]$, avec $a < b$, et on note $X \hookrightarrow \mathcal{U}([a, b])$,

si X est une variable à densité de densité $f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b] \end{cases}$.

Remarquer qu'avec probabilité 1, $X(\Omega) = [a, b]$, mais l'intervalle qui apparait dans f aurait pu être $[a, b[$, $]a, b]$ ou encore $]a, b[$.

→ Vérifier que f ainsi définie est bien une densité de probabilité. La dessiner.

La fonction de répartition d'une loi $\mathcal{U}([a, b])$ est la fonction F définie sur \mathbb{R} par

$$F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x > b \end{cases}$$

→ preuve

Proposition

Si $X \hookrightarrow \mathcal{U}([a, b])$, alors X admet une espérance et une variance et $E(X) = \frac{a+b}{2}, V(X) = \frac{(b-a)^2}{12}$.

→ preuve

Cas particulier de la loi $\mathcal{U}([0, 1])$: $f(x) = \begin{cases} 1 & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases}$, $F(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$ et $E(X) = \frac{1}{2}$

Exercice : Soit $a < b$ deux réels. Montrer que $X \hookrightarrow \mathcal{U}([0, 1]) \Leftrightarrow Y = a + (b - a)X \hookrightarrow \mathcal{U}([a, b])$.

Remarque : la loi uniforme a un rôle central en modélisation, car elle permet de simuler beaucoup d'autres lois, ainsi que diverses expériences.

Rappel de la syntaxe python : `random()` simule une réalisation de la loi uniforme sur $[0, 1]$.

3.2 Loi exponentielle

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

DÉFINITION

Soit $\lambda > 0$. On dit que X suit la loi exponentielle de paramètre λ et on note $X \hookrightarrow \mathcal{E}(\lambda)$,

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

→ Vérifier que f ainsi définie est bien une densité de probabilité. La dessiner.

Proposition

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

→ preuve

Proposition

Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors X admet une espérance et une variance et $E(X) = \frac{1}{\lambda}, V(X) = \frac{1}{\lambda^2}$.

→ preuve

Exercice : Soit $\lambda > 0$. Montrer que $X \hookrightarrow \mathcal{E}(1) \Leftrightarrow Y = \frac{1}{\lambda}X \hookrightarrow \mathcal{E}(\lambda)$.

Remarque : cette loi modélise souvent une durée de vie (d'une ampoule, d'une machine...), ou un temps d'attente (d'une panne...), malgré la propriété d'absence de mémoire.

Proposition Propriété d'absence de mémoire :

La loi exponentielle est une loi sans mémoire.

Autrement dit, si $X \hookrightarrow \mathcal{E}(\lambda)$ alors $\forall (x, y) \in (\mathbb{R}^+)^2, P_{(X > y)}(X > y + x) = P(X > x)$.

→ preuve

Remarque : si X mesure la durée de vie d'un composant, alors l'absence de mémoire (=absence de vieillissement) signifie que la probabilité que la machine fonctionne encore x unités est la même après fabrication qu'après avoir déjà fonctionné y unités.

Rappel : dans le chapitre Variables Discrètes, on a vu que la loi Géométrique était une loi sans mémoire.

3.3 Loi normale

DÉFINITION

On dit que X suit la loi normale centrée réduite et on note $X \hookrightarrow \mathcal{N}(0, 1)$,

si X est une variable à densité de densité $\varphi : x \mapsto \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Sa fonction de répartition est souvent notée $\Phi : \forall x \in \mathbb{R}, \Phi(x) = P(X \leq x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt$.

Remarque

On ne connaît pas de primitive de $t \mapsto e^{-\frac{t^2}{2}}$ donc on ne peut pas exprimer Φ directement. Pour faire une application numérique, vous pouvez utiliser la calculatrice ou python. Un échantillon de valeurs utiles (ou une table) pourra aussi

être rappelé dans le sujet.

Remarquer toutefois que $\Phi(0) =$

Propriétés de Φ :

φ étant continue sur \mathbb{R} , on en déduit que Φ est de classe C^1 sur \mathbb{R} , et $\forall x \in \mathbb{R}, \Phi'(x) = \varphi(x) > 0$.

Donc Φ réalise une bijection strictement croissante de \mathbb{R} dans $]0, 1[$.

Exercice : représenter graphiquement la densité φ d'une loi normale centrée réduite.

Visualiser graphiquement la propriété suivante de Φ : $\forall x \in \mathbb{R}, \Phi(-x) = 1 - \Phi(x)$ puis en faire la preuve.

Théorème

Soit $X \hookrightarrow \mathcal{N}(0, 1)$. Alors X admet une espérance et une variance, et $E(X) = 0, V(X) = 1$.

→ preuve

Loi normale généralisée :

Soit $\mu \in \mathbb{R}$, et $\sigma > 0$. On dit que X suit la loi normale $\mathcal{N}(\mu, \sigma^2)$, si X est une variable à densité de densité

$x \mapsto \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. Alors X admet une espérance et une variance, et $E(X) = \mu, V(X) = \sigma^2$.

Exercice :

Soit $\mu \in \mathbb{R}$, et $\sigma > 0$. Alors $X \hookrightarrow \mathcal{N}(\mu, \sigma^2) \Leftrightarrow X^* = \frac{X - \mu}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

Plus généralement :

Stabilité de la loi Normale :

Soit $X \hookrightarrow \mathcal{N}(\mu, \sigma^2)$. Alors pour tout $(a, b) \in \mathbb{R}^2$ avec $a \neq 0$, $aX + b$ suit encore une loi normale :
 $aX + b \hookrightarrow \mathcal{N}(a\mu + b, a^2\sigma^2)$.

Loi normale et modélisation :

La loi normale se retrouve "partout" car elle modélise beaucoup de phénomènes dont les courbes de croissance (carnet de santé) etc. Elle sera également vue comme "loi limite" dans le dernier chapitre de probabilité.

4 Somme de variables aléatoires à densité indépendantes

Définition-théorème admis

Soit X et Y deux variables aléatoires indépendantes de densités respectives f et g .

Alors $X + Y$ est une variable aléatoire à densité dont une densité h est donnée par :

$$\forall x \in \mathbb{R}, h(x) = \int_{-\infty}^{+\infty} f(x-t)g(t)dt = \int_{-\infty}^{+\infty} f(t)g(x-t)dt$$

Remarque : d'après le programme officiel, la formule du produit de convolution ci-dessus sera rappelée dans un sujet, et la convergence de l'intégrale n'est pas un attendu. Enfin, la fonction h est parfois notée $f \star g$.

Exercice : Soit X et Y deux variables indépendantes de même loi exponentielle de paramètre 1.

Déterminer la loi de $S = X + Y$.

Théorème Stabilité de la loi normale par somme

Soit X et Y deux variables indépendantes suivant des lois normales : $X \hookrightarrow \mathcal{N}(m, s^2)$ et $Y \hookrightarrow \mathcal{N}(\mu, \sigma^2)$.

Alors $X + Y$ suit encore une loi normale : $X + Y \hookrightarrow \mathcal{N}(m + \mu, s^2 + \sigma^2)$.

Plus généralement, soit (X_1, X_2, \dots, X_n) n variables aléatoires mutuellement indépendantes telles que pour

tout $k \in \llbracket 1, n \rrbracket$, $X_k \hookrightarrow \mathcal{N}(\mu_k, \sigma_k^2)$. Alors $\sum_{k=1}^n X_k$ suit encore une loi normale : $\sum_{k=1}^n X_k \hookrightarrow \mathcal{N}(\sum_{k=1}^n \mu_k, \sum_{k=1}^n \sigma_k^2)$

→ moyen mnémotechnique pour retenir les paramètres :