Chapitre 10a

Variables aléatoires à densité

Compétences attendues.

- ✓ Prouver qu'une variable aléatoire X est à densité.
- ✓ Montrer qu'une fonction f est une densité de probabilité.
- ✓ Déterminer la fonction de répartition d'une variable aléatoire $Y = g(X)$.
- ✓ Déterminer une densité d'une variable aléatoire X.
- ✓ Prouver l'existence et calculer l'espérance ou la variance d'une variable à densité.
- ✓ Utiliser le théorème de transfert pour calculer une espérance.

Mathieu Mansuy

Professeur en ECG deuxième année spécialité mathématiques approfondies au Lycée Louis Pergaud (Besançon)

Page personnelle : `mathieu-mansuy.fr/`

E-mail : `mathieu.mansuy@ac-besancon.fr`

1 Généralités sur les variables aléatoires à densité

Dans tout ce chapitre, $(\Omega, \mathscr{A}, P)$ désigne un espace probabilisé.

1.1 Définition

Rappels.

- On appelle *fonction de répartition* d'une variable aléatoire réelle X la fonction F_X définie sur $\mathbb{R}$ par :
$$\forall x \in \mathbb{R}, \quad F_X(x) = P(X \leq x).$$
- Une application $F : \mathbb{R} \to \mathbb{R}$ est la fonction de répartition d'une variable aléatoire réelle X si, et seulement si, elle satisfait les points suivants :
  (i) F est croissante sur $\mathbb{R}$;
  (ii) F est continue à droite en tout point de $\mathbb{R}$;
  (iii) $\lim\limits_{x \to -\infty} F(x) = 0$ et $\lim\limits_{x \to +\infty} F(x) = 1$.

  De plus, on dispose de l'égalité :
$$\forall x \in \mathbb{R}, \quad P(X = x) = P(X \leq x) - P(X < x) = F(x) - \lim_{t \to x^-} F(t).$$

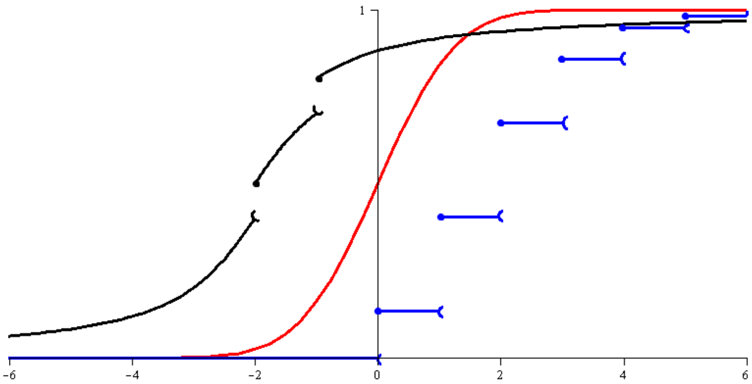
Définition.

Soit X une variable aléatoire réelle sur $(\Omega, \mathscr{A}, P)$. On dit que X est une *variable aléatoire à densité* (ou *variable aléatoire continue*) si sa fonction de répartition F_X est de plus :

(iv) continue sur $\mathbb{R}$;

(v) de classe $\mathscr{C}^1$ sur $\mathbb{R}$ éventuellement privé d'un nombre **fini** de points.

Remarque. Les trois courbes suivantes sont celles de fonctions de répartition de variables aléatoires réelles.



- La courbe bleue est en escalier. C'est donc la fonction de répartition d'une variable discrète X. Rappelons qu'elle est discontinue à droite en tout point de son ensemble image $X(\Omega)$. En particulier, une variable discrète n'est donc pas à densité.
- La courbe rouge représente une fonction de répartition (celle de la loi normale $\mathscr{N}(0,1)$) de classe $\mathscr{C}^1$ sur $\mathbb{R}$. C'est donc la fonction de répartition d'une variable continue.
- La fonction représentée par la courbe noire est discontinue en certains points, et n'est pas en escalier. C'est la fonction de répartition d'une variable aléatoire qui n'est ni discrète, ni continue. Cette année, nous n'étudierons pas ce type de variables aléatoires. On se focalisera sur les cas discret et continu.

Exercice. Soient $\lambda > 0$ et la fonction $F : \mathbb{R} \to \mathbb{R}$ définie pour tout $x \in \mathbb{R}$ par $F(x) = \begin{cases} 0 \text{ si } x < 0, \\ 1 - e^{-\lambda x} \text{ si } x \geq 0. \end{cases}$

1. Montrer qu'il existe une variable aléatoire X définie sur l'espace probabilisé $(\Omega, \mathscr{A}, P)$ de fonction de répartition F.

   On dit qu'une telle variable aléatoire suit une loi exponentielle de paramètre λ, ce qu'on note $X \hookrightarrow \mathscr{E}(\lambda)$.

2. Montrer que X est une variable aléatoire continue.

1.2 Densité de probabilité

Définition.

Soit X une variable aléatoire à densité.

On appelle *densité de probabilité de X* toute fonction f satisfaisant les deux points suivants :

(i) f est positive sur $\mathbb{R}$;

(ii) $f(x) = F_X'(x)$ pour tout x appartenant à $\mathbb{R}$ éventuellement privé d'un nombre fini de points.

Mise en garde.

Une densité d'une variable aléatoire n'est pas unique : si f est une densité de X et si l'on change la valeur de f en un nombre fini de points (en prenant pour nouvelles valeurs des réels positifs), alors on obtient une autre densité de f. Ainsi, on parlera pour f d'**une** densité de X et non de **la** densité de X.

Exercice. Soient $\lambda > 0$ et $X \hookrightarrow \mathscr{E}(\lambda)$. Déterminer une densité de X.

Pour aller plus loin.

Soit X une variable à densité, f une densité de X. Pour tout $x \in \mathbb{R}$ en lequel $f(x) = F_X'(x) \neq 0$ et pour tout $h > 0$ « petit » :

$$P(x < X \leq x+h) = F_X(x+h) - F_X(x) \approx f(x) \cdot h.$$

La quantité $f(x) \cdot h$ correspond donc approximativement à la probabilité que X soit proche de x par valeurs supérieures à la précision h. Ainsi, plus la densité prend une valeur élevée au point x, plus la probabilité d'obtenir une valeur proche de x est importante.

Propriété 1 (Lien fonction de répartition/densité)

Soit X une variable aléatoire à densité, f_X une densité de X. Pour tout $x \in \mathbb{R}$, l'intégrale $\displaystyle\int_{-\infty}^{x} f_X(t)\,\mathrm{d}t$ converge, et :

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\,\mathrm{d}t.$$

Remarque. La donnée d'une densité d'une variable aléatoire caractérise donc sa loi. En particulier si X et Y sont à densité de densités respectives f_X et f_Y, elles ont même loi si, et seulement si, $f_X = f_Y$ sauf éventuellement en un nombre fini de points.

Propriété 2 (Lien probabilités/densité)

Soit X une variable aléatoire à densité, f_X une densité de X. Pour tout $(a,b) \in \mathbb{R}^2$:

- $P(X = a) = 0$.
- $P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = \displaystyle\int_a^b f_X(t)\,\mathrm{d}t.$

Preuve. F_X étant continue sur $\mathbb{R}$, on obtient pour tout $a \in \mathbb{R}$:

$$P(X = a) = F_X(a) - \lim_{x \to a^-} F_X(x) = 0.$$

Il suit que, pour tout $a, b \in \mathbb{R}$, $P(a \leq X \leq b) = P(X = a) + P(a < X \leq b) = P(a < X \leq b)$. On démontre de même les autres égalités. Notons enfin que :

$$P(a \leq X \leq b) = P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F_X(b) - F_X(a) = \int_a^b f_X(t)\,\mathrm{d}t.$$

□

Mise en garde.

Ainsi si X est continue, $P(X = a) = 0$ pour tout $a \in \mathbb{R}$. Ceci est bien évidemment **faux** lorsque X est une variable aléatoire **discrète**.

Propriété 3

Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$. f est une densité d'une variable aléatoire si, et seulement si, elle satisfait les points suivants :

(i) f est continue sur $\mathbb{R}$ sauf éventuellement en un nombre **fini** de points ;

(ii) f est positive sur $\mathbb{R}$;

(iii) $\displaystyle\int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t$ converge et vaut 1.

Exercice. Soit $\lambda \in \mathbb{R}$. Pour tout $t \in \mathbb{R}$, on pose $f(t) = \dfrac{\lambda}{1+t^2}$.

1. Déterminer la valeur de λ pour que f soit une densité d'une variable aléatoire X.

   On dit que X suit une loi de Cauchy.

2. Déterminer sa fonction de répartition F_X.

3. Tracer les courbes représentatives de f et F_X et représenter graphiquement les probabilités $P(-1 \leq X \leq 1)$ et $P(X \geq 1)$. Calculer ces probabilités. Que vaut $P(X \leq -1)$?

Notation.

Soit X une variable à densité, et f_X une densité de X. Pour $I = \{t \in \mathbb{R},\ f_X(t) > 0\}$:

$$P(X \in I) = \int_I f_X(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} f_X(t)\,\mathrm{d}t = 1.$$

On notera dans la suite $X(\Omega) = \{t \in \mathbb{R},\ f_X(t) > 0\}$, qu'on appellera l'*ensemble image* de la variable X. Cet ensemble ainsi défini, dépend de la densité f_X choisie pour X.

Remarque. En notant $m = \inf(X(\Omega))$ et $M = \sup(X(\Omega))$ (dans le cas où l'ensemble $X(\Omega)$ est minoré et majoré), on a :

$$\forall x \leq m, \quad F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{x} 0\,\mathrm{d}t = 0,$$

$$\forall x \geq M, \quad F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t = \int_{-\infty}^{+\infty} f(t)\,\mathrm{d}t = 1.$$

1.3 Variable aléatoire fonction d'une variable aléatoire à densité

Propriété 4

Soit X une variable aléatoire réelle sur $(\Omega, \mathscr{A}, P)$, et soit $g : X(\Omega) \to \mathbb{R}$ une fonction continue. Alors $g(X)$ est une variable aléatoire réelle sur $(\Omega, \mathscr{A}, P)$.

Exemple. Si X est une variable aléatoire réelle, alors X^2, e^X, $aX+b$, … sont des variables aléatoires réelles.

Question. Si X est à densité, est-ce que $g(X)$ est encore à densité ? La réponse est non en général, et il nous faudra étudier chaque cas pour pouvoir conclure. Pour cela, nous procéderons comme suit.

Méthode. Comment montrer que $Y = g(X)$ est à densité et en donner une densité ?

Nous procèderons en quatre étapes :

Étape 1 : Recherche de $Y(\Omega)$.
On fixe une densité de X et on précise $X(\Omega)$. On détermine alors $Y(\Omega)$, image de $X(\Omega)$ par g, afin de trouver sans calcul sur quels intervalles F_Y vaut 0 ou 1.

Étape 2 : Détermination de F_Y.
On revient à la définition $F_Y(x) = P(Y \leq x)$, puis on résout cette inéquation (en prenant bien garde au sens des inégalités) pour l'exprimer à l'aide de F_X.

Étape 3 : Y est-elle continue ?
Une fois F_Y exprimée à l'aide de F_X, on étudie sa continuité sur $\mathbb{R}$ et son caractère $\mathscr{C}^1$ sur $\mathbb{R}$ privé éventuellement d'un nombre fini de points. On conclut ainsi que Y est (ou non) à densité.

Étape 4 : Détermination d'une densité de Y.
On dérive F_Y aux points où c'est possible pour obtenir une densité de Y. On prendra des valeurs arbitraires (positives) aux points où F_Y n'est pas dérivable.

Propriété 5 (Transformation affine d'une variable à densité)

Soit X une variable aléatoire à densité, f_X une densité de X. Pour tout $a \in \mathbb{R}^*$ et $b \in \mathbb{R}$, la variable $Y = aX + b$ est à densité, et une densité de Y est donnée par :

$$t \mapsto \frac{1}{|a|} f_X\left(\frac{t-b}{a}\right).$$

Preuve.

$\square$

Exercice. Supposons que $X \hookrightarrow \mathscr{E}(2)$. Déterminer une densité de probabilité de $Y = -2X + 3$.

Exercice. Soit X une variable aléatoire suivant une loi de Cauchy.

1. Montrer que $Z = e^X$ est une variable à densité, puis en déterminer une densité.
2. Même question pour $T = X^2$.

2 Moments d'une variable aléatoire à densité

2.1 Espérance

Définition.

Soit X une variable aléatoire à densité sur $(\Omega, \mathscr{A}, P)$, f_X une densité de X.
On dit que X *admet une espérance* lorsque l'intégrale $\displaystyle\int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t$ converge absolument.
Dans ce cas, on appelle *espérance de* X, et on note $E(X)$, le réel défini par :

$$E(X) = \int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t.$$

Exercice. Déterminer l'espérance de X, si elle existe, dans les cas suivants :

1. X suit une loi exponentielle $\mathscr{E}(\lambda)$;
2. X suit une loi de Cauchy.

Propriété 6 (Linéarité de l'espérance)

Soient X et Y des variables aléatoires à densité admettant une espérance. Alors pour tout $\lambda, \mu \in \mathbb{R}$, la variable aléatoire $\lambda X + \mu Y$ admet une espérance, et :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y).$$

Propriété 7 (Existence de l'espérance par domination)

Soient X, Y deux variables aléatoires à densité telles que $0 \leq |X| \leq Y$ presque sûrement. Si Y admet une espérance, alors X admet aussi une espérance, et $|E(X)| \leq E(Y)$.

Propriété 8 (Variables bornées et espérance)

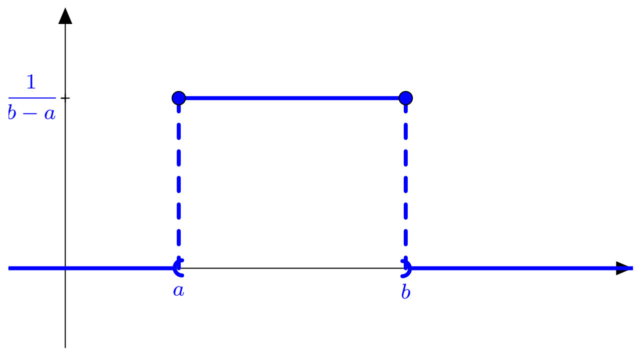
Soit X une variable aléatoire à densité.
Si X est bornée, c'est-à-dire si $X(\Omega)$ est bornée[a], alors X admet une espérance.

[a] Bien que $X(\Omega)$ dépende de la densité de X choisie, son caractère borné est lui indépendant de ce choix puisque deux densités de X diffèrent au plus sur un nombre fini de points.

Exemple. Une variable X suit une loi uniforme sur un intervalle $[a, b]$ si elle admet pour densité la fonction :

$$f : t \in \mathbb{R} \mapsto \begin{cases} \dfrac{1}{b-a} & \text{si } a \leq t \leq b, \\ 0 & \text{sinon.} \end{cases}$$

Puisque $X(\Omega) = [a, b]$, X est bornée et admet donc une espérance.



Densité de la loi uniforme sur $[a, b]$.

Propriété 9

Soient X, Y des variables aléatoires à densité admettant une espérance.

- *Positivité de l'espérance :* si $X \geq 0$ presque sûrement, alors $E(X) \geq 0$.
- *Croissance de l'espérance :* si $X \leq Y$ presque sûrement, alors $E(X) \leq E(Y)$.

2.2 Théorème de transfert

Théorème 10 (*de transfert*)

Soit X une variable aléatoire admettant une densité f_X nulle en dehors de $]a, b[$ ($a, b \in \mathbb{R} \cup \{\pm\infty\}$), et soit $\varphi :]a, b[\to \mathbb{R}$ une fonction continue sauf éventuellement en un nombre fini de points.

$E(\varphi(X))$ existe si, et seulement si, $\displaystyle\int_a^b \varphi(t) f_X(t)\,\mathrm{d}t$ converge **absolument**. Et en cas de convergence absolue :

$$E(\varphi(X)) = \int_a^b \varphi(t) f_X(t)\,\mathrm{d}t.$$

Corollaire 11

Si X est une variable aléatoire à densité admettant une espérance, alors pour tout $a, b \in \mathbb{R}$, la variable $aX + b$ admet une espérance, et :

$$E(aX + b) = aE(X) + b.$$

Preuve. Soit $\varphi : \mathbb{R} \to \mathbb{R}$ définie par $\varphi(t) = at + b$. φ est continue sur $\mathbb{R}$, et pour tout $t \in \mathbb{R}$:

$$|\varphi(t) f_X(t)| = |at f_X(t) + b f_X(t)| \leq |a||t f_X(t)| + |b||f_X(t)| \text{ par inégalité triangulaire.}$$

Or $\displaystyle\int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t$ converge absolument car X admet une espérance, et $\displaystyle\int_{-\infty}^{+\infty} f_X(t)\,\mathrm{d}t$ converge (absolument). Par théorème de comparaison, $\displaystyle\int_{-\infty}^{+\infty} \varphi(t) f_X(t)\,\mathrm{d}t$ converge absolument. Par le théorème de transfert, $E(aX + b)$ existe donc bien, et :

$$E(aX + b) = \int_{-\infty}^{+\infty} (at + b) f_X(t)\,\mathrm{d}t \overset{\text{tout converge}}{=} a \int_{-\infty}^{+\infty} t f_X(t)\,\mathrm{d}t + b \int_{-\infty}^{+\infty} f_X(t)\,\mathrm{d}t = aE(X) + b.$$

□

Exercice. Soit X une variable aléatoire suivant une loi exponentielle $\mathscr{E}(1)$.

1. La variable aléatoire $Y = e^X$ admet-elle une espérance ?
2. Montrer que $Z = X^2$ admet une espérance qu'on calculera.

2.3 Moments d'ordre supérieur, variance

Définition.

Soit X une variable aléatoire à densité, f_X une densité de X, et $r \in \mathbb{N}^*$.
On dit que X *admet un moment d'ordre* r si X^r admet une espérance. Par le théorème de transfert, c'est le cas si, et seulement si :

$$\int_{-\infty}^{+\infty} t^r f_X(t)\,\mathrm{d}t \quad \text{converge absolument.}$$

Dans ce cas, on note $m_r(X) = E(X^r)$ le *moment d'ordre* r *de* X.

Propriété 12

Soit X une variable aléatoire à densité, et soit $q, r \in \mathbb{N}^*$ tels que $q \leq r$.
Si X admet un moment d'ordre r, alors X admet un moment d'ordre q.

Preuve. La preuve est semblable au cas où X est discrète, et laissée en exercice. □

Remarque. Si une variable à densité X admet un moment d'ordre 2, alors X admet une espérance.

Définition.

Soit X une variable aléatoire à densité admettant une espérance, et f_X une densité de X.
On dit que X *admet une variance* si $(X - E(X))^2$ admet une espérance. D'après le théorème de transfert, c'est le cas si, et seulement si :

$$\int_{-\infty}^{+\infty} (t - E(X))^2 f_X(t)\,\mathrm{d}t \quad \text{converge } \textbf{absolument}.$$

Dans ce cas, la *variance de* X, notée $V(X)$, est :

$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f_X(t)\,\mathrm{d}t.$$

Propriété 13

Si X est une variable aléatoire à densité admettant une variance, alors $V(X) > 0$.

On appelle *écart-type de* X, et on note $\sigma(X)$, le réel défini par $\sigma(X) = \sqrt{V(X)}$.

Preuve. Supposons que X admette une variance, et notons f_X une densité de X. La fonction $g : t \mapsto (t - E(X))^2 f_X(t)$ est positive sur $\mathbb{R}$, d'où par positivité de l'intégrale :

$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f_X(t)\,\mathrm{d}t \geq 0.$$

Supposons $V(X) = 0$. Notons $a_1 < a_2 < \cdots < a_n$ les points de discontinuité (éventuels) de f_X. Pour tout $i \in [\![1, n-1]\!]$, la fonction g est continue et positive sur $]a_i, a_{i+1}[$, d'intégrale nulle sur cet intervalle. Par théorème de nullité de l'intégrale, g est nulle sur $]a_i, a_{i+1}[$. Puisque de plus $(t - E(X))^2$ s'annule uniquement pour $t = E(X)$, f_X est donc nulle sur $\mathbb{R}$ sauf éventuellement en un nombre fini de points $E(X), a_1, \ldots, a_n$. Mais alors $\displaystyle\int_{-\infty}^{+\infty} f_X(t)\,\mathrm{d}t = 0$, ce qui est faux puisque cette intégrale est égale à 1. Ainsi $V(X)$ est bien strictement positive. □

Propriété 14

Soit X une variable aléatoire à densité admettant une variance, et soient $a, b \in \mathbb{R}$. Alors $aX + b$ admet une variance, et :

$$V(aX + b) = a^2 V(X).$$

Propriété 15 (*Formule de Huygens*)

Soit X une variable aléatoire à densité admettant une espérance.

Alors X admet une variance si, et seulement si, X admet un moment d'ordre deux. Et lorsque c'est le cas :

$$V(X) = E(X^2) - E(X)^2.$$

Définition.

Soit X une variable aléatoire à densité.

- X est dite *centrée* si X admet une espérance et si $E(X) = 0$.
- X est dite *centrée réduite* si X admet une variance, et $E(X) = 0, V(X) = 1$.

Propriété 16

Soit X une variable aléatoire à densité admettant une variance. Alors :

- $X - E(X)$ est centrée ;
- $X^* = \dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite. On l'appelle *la variable aléatoire centrée réduite associée à* X.