Université de Toulon

M1 MEEF Mathématiques

TD 5 - Variables aléatoires à densité

$\boxed{\text{Le cours}}$

- **Variables aléatoires continues .** Soit $(\Omega, \mathfrak{F}, \mathbb{P})$ un espace mesurable, et soit X une variable aléatoire réelle sur $(\Omega, \mathfrak{F}, \mathbb{P})$. On dit que la loi de X est continue si pour tout $x \in \mathbb{R}$, $\mathbb{P}_X(\{x\}) = 0$.

  La fonction de répartition d'une variable aléatoire continue est continue sur $\mathbb{R}$ (On savait déjà que dans le cas d'une variable aléatoire quelconque, la fonction de répartition était continue à droite en tout point).

  Remarques : Dans certains ouvrages, la notion de variable aléatoire continue X est introduite à l'aide de l'ensemble des valeurs prises par X ; comme première approche intuitive c'est raisonnable, mais d'un point de vue rigoureux, ceci peut, au mieux, prêter à confusion, et au pire, aboutir à une définition fausse de variable aléatoire continue. Par exemple, il est faux de dire " X est une v.a. continue si $X(\Omega)$ est un intervalle". Imaginez par exemple qu'il y ait une chance sur deux que X prenne ses valeurs de façon uniforme dans $[0, 1]$ et une chance sur deux que X prenne la valeur 0. On a bien $X(\Omega) = [0, 1]$, mais $P(X=1)=P(X=0) = \frac{1}{2} \times \frac{1}{2} \neq 0$; et donc X n'est pas continue.

- **Variables aléatoires à densité.** Soit X une variable aléatoire réelle de loi P_X. On dit que X suit une loi à densité f_X si pour tout intervalle $I \subset \mathbb{R}$, on a

(1) $$\mathbb{P}_X(I) = \int_I f_X(x)\mathrm{d}x\,,$$

  ou, *de façon équivalente*, si pour tout $t \in \mathbb{R}$, la fonction de répartition de X est

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

  Une variable aléatoire à densité est une variable aléatoire continue. La réciproque est fausse en général, mais on ne verra aucun exemple de variable aléatoire continue qui ne soit pas à densité.

  Si X est une variable aléatoire à densité, de densité f_X, alors on montre que f_X vérifie :

  . f_X est positive (presque-partout au sens de Lebesgue).

  . $\displaystyle\int_{-\infty}^{+\infty} f_X(x)\mathrm{d}x = 1.$

  Réciproquement, si f vérifie les 3 propriétés suivantes :

  a) f est une fonction borélienne

  b) f est positive (presque-partout au sens de Lebesgue).

  c) $\displaystyle\int_{-\infty}^{+\infty} f(x)\mathrm{d}x = 1.$

  alors f est appelée *densité*, et toute variable aléatoire X dont la loi vérifie (1) avec f_X remplacée par f, est une variable aléatoire à densité.

  Remarque : Une condition suffisante pour obtenir a) est que f est continue sauf éventuellement en un nombre fini de points. Aussi, dans les ouvrages de lycée, la condition a) est omise. La condition b) est remplacée par : pour tout x dans $\mathbb{R}$, $f(x) \geq 0$.

  Soit X une variable aléatoire à densité, de densité f_X. On a alors :

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

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

et plus généralement

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

- **Fonction de répartition et densité d'une variable aléatoire à densité**
  Soit X une variable aléatoire réelle à densité f_X et de fonction de répartition F_X. Alors on a

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

et on a l'égalité suivante (presque-partout, i.e. partout sauf sur un ensemble de mesure de Lebesgue nulle) :

$$f_X(t) = F'_X(t) \,.$$

Attention ! La fonction de répartition d'une variable aléatoire continue, même si elle est à densité, n'est pas forcément dérivable partout. Par exemple, la fonction de répartition de la variable aléatoire X qui suit une loi uniforme sur $[0,\, 1]$ ne sera pas dérivable en 0 et en 1. Une façon de contourner ce problème (au lycée), est de parler de probabilité sur $[a,\, b]$.

- **Fonction caractéristique.** Soit X une variable aléatoire réelle. La fonction caractéristique de X est la transformée de Fourier φ_X de la loi de X, c'est à dire

$$\varphi_X(t) = \mathbb{E}(\mathrm{e}^{itX}) \,.$$

La fonction caractéristique caractérise la loi, i.e., deux variables aléatoires ayant même fonction caractéristique ont même loi.

Dans le cas d'une variable aléatoire X à densité f_X, la fonction caractéristique de X est donnée par

$$\varphi_X(t) = \int_{-\infty}^{+\infty} \mathrm{e}^{itx} f_X(x) \mathrm{d}x \,.$$

La fonction caractéristique a les propriétés usuelles d'une transformée de Fourier. En particulier, on a la formule d'inversion suivante : Si φ_X est intégrable, alors X est une variable aléatoire à densité, dont la densité est donnée (presque-partout) par

$$f_X(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathrm{e}^{-itx} \varphi_X(t) \mathrm{d}t \,.$$

La fonction caractéristique permet en particulier de calculer les moments d'une variable aléatoire :

Si $\mathbb{E}(|X|^n) < +\infty$, alors $\varphi_X \in C^n(\mathbb{R})$ et pour tout m compris entre 1 et n, on a

$$\varphi_X^{(m)}(t) = i^m \mathbb{E}(X^m \mathrm{e}^{itX}) \,, \quad \text{et} \quad \varphi_X^{(m)}(0) = i^m \mathbb{E}(X^m) \,.$$

Une application des fonctions caractéristiques : Le théorème de continuité de Paul Lévy, qui donne l'équivalence entre la convergence ponctuelle des fonctions caractéristiques et la convergence en loi des variables aléatoires associées, est utilisé pour démontrer le théorème de la limite centrale (c.f. un chapitre ultérieur).

- **Lois de probabilité à densité usuelles.**
  — Loi uniforme : Une variable aléatoire réelle X suit une loi uniforme sur $[a, b]$ $(-\infty < a < b < +\infty)$, notée $\mathcal{U}([a,\, b])$, si elle est à densité

$$f_X(x) = \frac{1}{b-a} 1\!\!1_{[a,b]}(x).$$

  (Ici, $1\!\!1_{[a,b]}(x)$ est la fonction indicatrice de l'intervalle $[a,\, b]$ qui vaut 1 si $x \in [a,\, b]$ et qui vaut 0 sinon).

  — Loi exponentielle de paramètre λ : Une variable aléatoire réelle X suit la loi exponentielle de paramètre $\lambda > 0$, notée $\mathcal{E}(\lambda)$, si elle est à densité

$$f_X(x) = \lambda \mathrm{e}^{-\lambda x} 1\!\!1_{]0,+\infty[}(x).$$

La loi exponentielle est *"sans mémoire"*, i.e., si $X \sim \mathcal{E}(\lambda)$, alors pour tout $a > 0$, $b > 0$ tels que $a < b$,

$$\mathbb{P}_{X \geq a}(X \geq b) = \mathbb{P}(X \geq b - a)$$

(démontrer cette propriété et expliquer l'expression "sans mémoire").

— Loi normale ou loi de Laplace-Gauss : Une variable aléatoire réelle X suit la loi normale de paramètres m et σ, notée $\mathcal{N}(m, \sigma)$, si elle est à densité

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right).$$

Dans le cas $m = 0$ et $\sigma = 1$, on parle de *loi normale centrée réduite.*
Si X_1 et X_2 sont deux variables aléatoires indépendantes telles que $X_1 \sim \mathcal{N}(m_1, \sigma_1)$ et $X_2 \sim \mathcal{N}(m_2, \sigma_2)$, alors $X_1 + X_2 \sim \mathcal{N} m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2})$.

— Loi Gamma : Une variable aléatoire X suit la loi gamma de paramètres p et λ si elle est à densité

$$f_X(x) = \frac{\lambda}{\Gamma(p)} \mathrm{e}^{-\lambda x} (\lambda x)^{p-1} \mathbb{1}_{[0,+\infty[}(x).$$

— Loi du Chi-Deux χ^2 : Soient X_1, X_2, ..., X_k, k variables aléatoires indépendantes qui suivent une loi normale centrée réduite. La variable aléatoire

$$X = \sum_{i=1}^{k} {X_i}^2\,,$$

suit une loi du chi-deux à k degrés de liberté, notée $\chi^2(k)$. C'est une variable aléatoire à densité, de densité

$$f_X(x) = \frac{1}{2^{\frac{k}{2}} \Gamma(\frac{k}{2})} x^{\frac{k}{2}-1} \mathrm{e}^{-\frac{x}{2}}\,.$$

Les exercices

Exercice 1. On considère la fonction suivante, définie sur $\mathbb{R}$:

$$f(t) = \begin{cases} kt^2 & \text{si } -1 \leq t \leq 1 \\ 0 & \text{sinon.} \end{cases}$$

(1) Déterminer k pour que f soit une densité de probabilité. On supposera désormais que k est égal à la valeur trouvée, et que f est la densité d'une variable aléatoire X.
(2) Déterminer la fonction de répartition F_X de X.
(3) Soit $a \in \mathbb{R}$. Déterminer, en fonction de a, les valeurs de x telles que $\mathbb{P}(X > x) = a$. Calculer $\mathbb{E}(X)$, $\mathbb{V}\mathrm{ar}(X)$.

Exercice 2. Lois à densité usuelles ... et autres
i) *Loi uniforme.* Pour $X \sim \mathcal{U}([a,\, b])$, calculer $\mathbb{E}(X)$, $\mathbb{V}\mathrm{ar}(X)$. Calculer la fonction caractéristique de X et retrouver les valeurs de $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.
ii) *Loi normale.* Calculer

$$\int_{\mathbb{R}^2} \mathrm{e}^{-(\frac{x^2}{2}+\frac{y^2}{2})} \mathrm{d}x\mathrm{d}y\,.$$

Pour $X \sim \mathcal{N}(0,\, 1)$, en déduire

$$\int_{\mathbb{R}} f_X(x)\mathrm{d}x = 1\,.$$

Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$. Généraliser pour $X \sim \mathcal{N}(m,\, \sigma)$.
Soit $Y = X^2$. Calculer la loi de X^2.
iii) *Loi Log-normale.* On considère une variable aléatoire X qui suit la loi normale $\mathcal{N}(m, \sigma)$. La variable aléatoire $Y = \mathrm{e}^X$ suit alors une loi appelée Log-normale. Considérons le cas particulier $m = 0$ et $\sigma = 1$, et calculer la fonction de répartition $F_Y(x) := \mathbb{P}\{Y \leq x\}$ en fonction de la fonction de répartition de X. En déduire que Y est une variable aléatoire à densité que l'on déterminera.
iv) *Loi exponentielle $\mathcal{E}(\lambda)$.* Soit $X \sim \mathcal{E}(\lambda)$. Calculer l'espérance $\mathbb{E}(X)$ et la variance $\mathbb{V}\mathrm{ar}(X)$.

Si X suit une loi $\mathcal{E}(\lambda)$, montrer que $Y = [X]$ (partie entière de X) suit une loi géométrique. Montrer que la variable aléatoire $U = \mathrm{e}^{-\lambda X}$ suit la loi uniforme sur $]0, 1[$.
v) *Première loi de Laplace.*
Soit une variable aléatoire X qui suit une loi à densité

$$f(x) = \frac{1}{2}\mathrm{e}^{-|x|},\ x \in \mathbb{R}.$$

Calculer $\mathbb{E}(X)$ et $\mathbb{V}\mathrm{ar}(X)$.
vi) *Loi de Cauchy.* Soit une variable aléatoire X qui suit une loi à densité

$$f(x) = \frac{1}{\pi}\frac{1}{1+x^2},\ x \in \mathbb{R}.$$

Peut-on calculer $\mathbb{E}(X)$? Soit V une variable aléatoire uniformément distribuée sur $]-\frac{\pi}{2}, \frac{\pi}{2}[$ et soient $Y = \tan V$, $Z = 1/\tan V$. Calculer $\mathbb{P}\{Y \leq x\}$ et en déduire que Y suit la loi de Cauchy. Montrer de même que Z suit la loi de Cauchy.
vii) *Loi gamma* $\Gamma(p, \lambda)$. Soit X une variable aléatoire qui suit une loi Γ de paramètres p et λ. Montrer que $\mathbb{E}(X) = \frac{p}{\lambda}$ et $\mathbb{V}\mathrm{ar}(X) = \frac{p}{\lambda^2}$.

Exercice 3. Soit X une variable aléatoire réelle qui suit une loi à densité uniforme sur $[0, n]$, $n \in \mathbb{N}^*$ étant fixé. Déterminer la loi de $Y = X - [X]$ où $[\,.\,]$ est la fonction partie entière.

Exercice 4. Dans le plan affine euclidien rapporté à un repère orthonormé $(O, \vec{i}, \vec{j})$, on note, respectivement, J et A les points de coordonnées $(0, 1)$ et $(1, 1)$. Soit X une v.a. de loi uniforme sur $[0, 1]$ et M le point du segment $[OA]$ d'abscisse X et N le point d'intersection entre la droite (JM) et l'axe des abscisses. Donner la loi de la variable aléatoire Y, abscisse de N.

Exercice 5. On se place dans un repère orthonormé $(O, \overrightarrow{i}, \overrightarrow{j})$. Soit M un point pris au hasard sur le quart de cercle unité qui est contenu dans l'ensemble des points d'abscisse et d'ordonnée positives. Soit Θ la variable aléatoire qui décrit l'angle $(\widehat{\overrightarrow{i}, \overrightarrow{OM}})$. On suppose que Θ suit une loi à densité uniforme sur $[0, \frac{\pi}{2}]$.
1) Déterminer l'espérance et la variance de Θ.
2) Soit Y la variable aléatoire égale à l'ordonnée du point M. Exprimer Y en fonction de Θ et déterminer la fonction de répartition de Y.
3) Déterminer la densité de Y.

Exercice 6 (extrait de MathX - terminale S)**.**

(1) À 23h, heure à laquelle finit la représentation de *Tosca*, la diva qui chante le rôle titre prends son temps pour quitter l'Opéra Bastille. Ce temps, en heures, peut être modélisé par une variable aléatoire T suivant la loi exponentielle de paramètre $\frac{1}{3}$.
Calculer la probabilité que la diva quitte l'opéra :
  (a) après une heure du matin.
  (b) avant 2h du matin sachant qu'à minuit elle n'est toujours pas sortie.

(2) Après chaque représentation, un admirateur attend la diva pour lui offrir des fleurs. Si elle sort avant 1h du matin, il rentre chez lui en métro. Sinon, il prend un taxi.
  (a) Calculer la probabilité p qu'il rentre en métro après une représentation
  (b) Cette année, *Tosca* est à l'affiche pour une série de 50 représentations. Soit X la variable aléatoire donnant le nombre de jours où l'admirateur rentre chez lui en métro. Déterminer la loi de X et son espérance $E(X)$.

(3) Le retour en métro revient à 1 €, mais celui en taxi coûte 10 €. Soit Y la variable aléatoire égale au cot de son retour à la maison durant la série de représentations. Déterminer la loi de Y et son espérance $E(Y)$.

Exercice 7. Soit Y une variable aléatoire de densité f_Y, et soit $m \in \mathbb{R}$. Montrer que si la fonction f_Y est symétrique par rapport à m, alors $\mathbb{E}(Y) = m$.

Exercice 8. Un contre-exemple sur la notion d'indépendance. Rappeler la définition d'indépendance de deux variables aléatoires X et Y, ainsi que la caractérisation de cette propriété à l'aide des espérances.

Soient X et U deux variables aléatoires indépendantes telles que X suit la loi normale $\mathcal{N}(0,1)$, et U suit une loi de type Bernouilli de paramètre $1/2$: $\mathbb{P}\{U=-1\} = \mathbb{P}\{U=1\} = 1/2$. On définit la variable aléatoire $Y = UX$. Montrer que $Y = \mathcal{N}(0,1)$. Montrer que $\mathbb{E}(XY) = 0 = \mathbb{E}(X)E(Y)$. Montrer que X et Y ne sont pas indépendantes

Exercice 9. La désintégration nucléaire.
A partir d'un instant 0, on s'intéresse à la désintégration nucléaire d'un atome, disons d'uranium 238. La variable aléatoire réelle "durée de vie" est notée T.
Pour $t > 0$, soit $G(t) := \mathbb{P}(T > t)$ la probabilité que l'atome soit encore en vie à la date t. On pose également $F(t) := 1 - G(t)$.
On admet que si on considère un atome radioactif à un instant t, la probabilité qu'il ne soit toujours pas désintégré à la date $t' > t$ ne dépend que de la durée $t' - t$. C'est à dire, on admet que la durée T de vie d'un atome est indépendant du temps déjà écoulé, i.e. l'atome ne vieillit pas. On suppose en outre que les atomes n'interagissent pas entre eux.
1- Comment interpréter $F(t)$?
2- Montrer que pour tout $t > 0$ et $t' > 0$, on a $G(t + t') = G(t)G(t')$.
3- En déduire qu'il existe $\lambda > 0$, caractéristique du type d'atome considéré, tel que pour tout $t > 0$, F est donnée par $F(t) = 1 - \mathrm{e}^{-\lambda t}$.
4- Montrer que T est une variable aléatoire à densité dont on donnera la densité f_T.
5- Calculer $\mathbb{E}(T)$ et $\mathbb{V}\mathrm{ar}(T)$.
6- Calculer la probabilité qu'un atome dépasse son espérance de vie.
7- Calculer la probabilité qu'un atome dépasse le double de son espérance de vie.
8- Calculer la demi-période τ définie par $P(T > \tau) = \frac{1}{2}$.
A l'instant 0, on a N atomes radioactifs. On s'intéresse au nombre X_N^t d'atomes désintégrés à la date $t > 0$.
9- Donner la loi de X_N^t.

Exercice 10. i) Soient X_1 et X_2 deux variables aléatoires réelles, indépendantes, à densité, de densités respectives f_{X_1} et f_{X_2}. Déterminer $\mathbb{P}\big((X_1, X_2) \in [a,b] \times [c,d]\big)$. En déduire la loi du couple $(X_1,\, X_2)$.

On admettra pour la suite que $X_1 + X_2$ est une variable aléatoire à densité, de densité

$$f_{X_1+X_2}(x) = f_{X_1} * f_{X_2}(x) := \int_{\mathbb{R}} f_{X_1}(x-y) f_{X_2}(y)\mathrm{d}y = \int_{\mathbb{R}} f_{X_2}(x-y) f_{X_1}(y)\mathrm{d}y\,.$$

ii) Soient X_1 et X_2 deux variables aléatoires indépendantes et qui suivent toutes les deux la loi uniforme sur $[0,1]$.
Quelle est la loi de $-X_2$? Déterminer la loi de $X_1 - X_2$.
iii) *Application :* Deux amis ont rendez-vous à midi ; Des causes de retard indépendantes font que chacun arrive entre 12h et 13h. La probabilité d'arrivée de chacun d'eux dans un intervalle de temps donné étant proportionnelle à la longueur de cet intervalle, quelle est la probabilité que les deux amis se rencontrent si chacun attend au plus 1/4 d'heure (On pourra construire deux variables aléatoires réelles continues indépendantes X_1 et X_2 qui donnent respectivement le temps de retard de chacun des deux amis).