

RÉSUMÉ DE COURS : VARIABLES ALÉATOIRES À DENSITÉ

1 Généralités sur les variables aléatoires

1.1 Définitions et propriétés de base

Définition 1.1. Soit (Ω, \mathcal{A}) un espace probabilisable. Une variable aléatoire (réelle) sur (Ω, \mathcal{A}) est une application $X : \Omega \rightarrow \mathbb{R}$ telle que, pour tout $x \in \mathbb{R}$, l'ensemble $X^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid X(\omega) \leq x\}$ est un événement, c'est-à-dire appartient à \mathcal{A} .

Théorème 1.2. Toute combinaison linéaire, tout produit, tout maximum et tout minimum de variables aléatoires définies sur un même espace probabilisable (Ω, \mathcal{A}) est une variable aléatoire sur (Ω, \mathcal{A}) .

Définition 1.3. Soit (Ω, \mathcal{A}, P) un espace probabilisé, et soit X une variable aléatoire sur (Ω, \mathcal{A}, P) . Alors la loi de X est la donnée des probabilités $P([X \in I])$, où I est un intervalle de \mathbb{R} .

Définition 1.4. Soit X une variable aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) . Alors la fonction de répartition de X est l'application :

$$F_X : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & P([X \leq x]) \end{cases} .$$

Théorème 1.5. Soit X une variable aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) . Alors sa fonction de répartition F_X vérifie les propriétés suivantes :

1. La fonction F_X est croissante sur \mathbb{R} .
2. La fonction F_X est continue à droite en tout point de \mathbb{R} .
3. On a les limites suivantes : $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
4. Pour tous réels a, b tels que $a < b$, on a : $P([a < X \leq b]) = F_X(b) - F_X(a)$.

De plus, on montre (et nous admettrons) que, pour tout $x \in \mathbb{R}$:

$$\lim_{t \rightarrow x^-} F_X(t) = F_X(x) - P([X = x]).$$

Dès lors, on voit que l'on peut toujours retrouver la loi de X à partir de sa fonction de répartition. En effet, si l'on connaît la fonction de répartition de X , alors les résultats rassemblés dans le théorème 1.5 nous permettent de calculer $P([X \in I])$ pour tout intervalle I de \mathbb{R} . En particulier, on a le :

Théorème 1.6. Deux variables aléatoires sur un même espace probabilisé, dont les fonctions de répartition sont égales, suivent la même loi.

En d'autres termes, on dira de façon plus synthétique (et on retiendra) que :

$$\boxed{\text{la fonction de répartition de } X \text{ caractérise complètement la loi de } X.}$$

1.2 Notion de couple de variables aléatoires

Définition 1.7. Soit (Ω, \mathcal{A}) un espace probabilisable. Un couple de variables aléatoires (réelles) sur (Ω, \mathcal{A}) est une application :

$$Z : \begin{cases} \Omega & \longrightarrow & \mathbb{R}^2 \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{cases} ,$$

où X, Y sont 2 variables aléatoires sur (Ω, \mathcal{A}) . Elle est notée $Z = (X, Y)$.

Définition 1.8. Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) . La loi conjointe de (X, Y) est l'application :

$$F_{(X,Y)} : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & P([X \leq x] \cap [Y \leq y]) \end{cases} .$$

A noter que, si (X, Y) est un couple de variables aléatoires et si g est une application continue de \mathbb{R}^2 dans \mathbb{R} , alors $g(X, Y)$ est une variable aléatoire. En outre, on montre (et nous admettrons) le :

Théorème 1.9. Soient (X_1, Y_1) et (X_2, Y_2) deux couples de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) , qui suivent tous deux la même loi. Si g est une application continue de \mathbb{R}^2 dans \mathbb{R} , alors les variables aléatoires $g(X_1, Y_1)$ et $g(X_2, Y_2)$ suivent la même loi.

Définition 1.10. Soient X, Y des variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) . On dit que X et Y sont indépendantes si, pour tout $(x, y) \in \mathbb{R}^2$, on a :

$$F_{(X,Y)}(x, y) = F_X(x)F_Y(y).$$

En particulier, des variables aléatoires X et Y sont indépendantes si et seulement si les événements $[X \leq x]$ et $[Y \leq y]$ sont indépendants pour tout $(x, y) \in \mathbb{R}^2$. Grosso modo, les variables aléatoires X et Y sont indépendantes si et seulement si X et Y ne s'influencent pas, c'est-à-dire si les valeurs prises par X ne dépendent pas des valeurs prises par Y (et inversement).

Théorème 1.11. Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) . Alors les conditions suivantes sont équivalentes :

1. Les variables aléatoires X et Y sont indépendantes,
2. Pour tous intervalles I, J de \mathbb{R} , on a : $P([X \in I] \cap [Y \in J]) = P([X \in I])P([Y \in J])$.

Théorème 1.12. (Lemme des coalitions) Soit (X, Y) un couple de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) . Si X et Y sont indépendantes, alors toute variable aléatoire fonction de X est indépendante de toute variable aléatoire fonction de Y .

1.3 Notion de vecteur aléatoire

Définition 1.13. Un vecteur aléatoire sur un espace probabilisable (Ω, \mathcal{A}) est une application :

$$Z : \begin{cases} \Omega & \longrightarrow & \mathbb{R}^n \\ \omega & \longmapsto & (X_1(\omega), \dots, X_n(\omega)) \end{cases} ,$$

où X_1, \dots, X_n sont des variables aléatoires sur (Ω, \mathcal{A}) . On le note $Z = (X_1, \dots, X_n)$.

Théorème 1.14. Soit (X_1, \dots, X_n) un vecteur aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) . La loi conjointe de (X_1, \dots, X_n) est l'application :

$$F_{(X_1, \dots, X_n)} : \begin{cases} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ (x_1, \dots, x_n) & \longmapsto & P([X_1 \leq x_1] \cap \dots \cap [X_n \leq x_n]) \end{cases} .$$

A noter que, si (X_1, \dots, X_n) est un vecteur aléatoire et si g est une application continue de \mathbb{R}^n dans \mathbb{R} , alors $g(X_1, \dots, X_n)$ est une variable aléatoire. En outre, on montre (et nous admettrons) le :

Théorème 1.15. Soient (X_1, \dots, X_n) et (Y_1, \dots, Y_n) deux vecteurs aléatoires sur un même espace probabilisé (Ω, \mathcal{A}, P) , qui suivent tous deux la même loi. Si g est une application continue de \mathbb{R}^n dans \mathbb{R} , alors les variables aléatoires $g(X_1, \dots, X_n)$ et $g(Y_1, \dots, Y_n)$ suivent la même loi.

Définition 1.16. Des variables aléatoires X_1, \dots, X_n sur un espace probabilisé (Ω, \mathcal{A}, P) sont dites (mutuellement) indépendantes si, pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$, on a :

$$F_{(X_1, \dots, X_n)}(x_1, \dots, x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n).$$

En particulier, des variables aléatoires X_1, \dots, X_n sont indépendantes si et seulement si les événements $[X_1 \leq x_1], \dots, [X_n \leq x_n]$ sont (mutuellement) indépendants pour tout $(x_1, \dots, x_n) \in \mathbb{R}^n$. Grosso modo, les variables aléatoires X_1, \dots, X_n sont indépendantes si et seulement si X_1, \dots, X_n ne s'influencent pas, c'est-à-dire si les valeurs prises par certaines d'entre elles ne dépendent pas des valeurs prises par les autres.

Théorème 1.17. Soit (X_1, \dots, X_n) un vecteur aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) . Alors les conditions suivantes sont équivalentes :

1. Les variables aléatoires X_1, \dots, X_n sont indépendantes,
2. Pour tous intervalles I_1, \dots, I_n de \mathbb{R} , on a : $P(\bigcap_{i=1}^n [X_i \in I_i]) = \prod_{i=1}^n P([X_i \in I_i])$.

Définition 1.18. On dit qu'une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) est une suite de variables aléatoires indépendantes si, pour toute partie finie I de \mathbb{N} , les variables aléatoires X_i (avec $i \in I$) sont (mutuellement) indépendantes.

Théorème 1.19. (Lemme des coalitions) Soient X_1, \dots, X_n des variables aléatoires sur un espace probabilisé (Ω, \mathcal{A}, P) . Si les variables aléatoires X_1, \dots, X_n sont indépendantes, alors toute variable aléatoire fonction de p d'entre elles est indépendante de toute variable aléatoire fonction des $(n - p)$ restantes.

2 Notion de variable aléatoire à densité

2.1 Définitions et propriétés de base

Définition 2.1. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est une densité de probabilité si :

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

Théorème - définition 2.2. Soit X une variable aléatoire sur un espace probabilisé (Ω, \mathcal{A}, P) . On dit que X est une variable aléatoire à densité si sa fonction de répartition F_X est continue sur \mathbb{R} et de classe \mathcal{C}^1 sur \mathbb{R} privé d'un nombre fini (éventuellement vide) de points. Dans ce cas, toute fonction f positive sur \mathbb{R} , qui ne diffère de F'_X qu'en un nombre fini de points, est une densité de probabilité, et l'on dira que f est une densité de X .

En particulier, le résultat ci-dessus entraîne que, si X est une variable aléatoire à densité, de densité f_X , alors l'intégrale ci-dessous converge et de plus :

$$\int_{-\infty}^{+\infty} f_X(t)dt = 1.$$

A noter qu'avec le théorème-définition 2.2, aucune variable aléatoire discrète n'est à densité, puisque la fonction de répartition d'une variable aléatoire discrète est discontinue en au moins un point. A noter aussi qu'une densité n'est définie qu'à un nombre fini de points près. En effet, on peut choisir un nombre fini de points de \mathbb{R} et attribuer différentes valeurs à une densité de X sans pour autant changer le fait que ce soit une densité de X . C'est pourquoi on ne parlera pas de la densité de X , mais d'une densité de X . En d'autres termes, l'ensemble des densités de X est l'ensemble des fonctions f positives sur \mathbb{R} et continues sur \mathbb{R} sauf en un nombre fini de points, telles que $f(x) = F'_X(x)$ pour tout réel x sauf peut-être en un nombre fini de points. S'il n'y a pas d'ambiguïté, on notera f_X une densité de X . A noter qu'une densité est forcément positive par définition, ce qui fait sens car elle est (sauf en un nombre fini de points) la dérivée d'une fonction de répartition, qui est toujours croissante.

Dans tout ce qui suit, toutes les variables aléatoires que l'on rencontrera seront supposées définies sur un même espace probabilisé (Ω, \mathcal{A}, P) .

Théorème 2.3. Soit X une variable aléatoire à densité. Alors, pour tout $x \in \mathbb{R}$, on a : $P([X = x]) = 0$.

En d'autres termes, on dira qu'une variable aléatoire à densité ne charge aucun des points de \mathbb{R} , c'est-à-dire qu'elle leur attribue à tous une probabilité nulle d'arriver. C'est l'une des différences notables avec les variables aléatoires discrètes.

Théorème 2.4. Soit X une variable aléatoire à densité. Alors, pour tout $x \in \mathbb{R}$, l'intégrale ci-dessous converge et de plus :

$$F_X(x) = P([X \leq x]) = \int_{-\infty}^x f_X(t)dt.$$

En particulier, dans le cas d'une variable aléatoire à densité X , on voit que, si l'on connaît une densité de X , alors on peut retrouver la fonction de répartition de X , et donc la loi de X . En d'autres termes :

Corollaire 2.5. La donnée d'une densité f_X d'une variable aléatoire à densité X caractérise la loi de X .

De façon équivalente, ceci signifie que deux variables aléatoires à densité X et Y , qui admettent des densités f_X et f_Y égales sur \mathbb{R} privé d'un nombre fini (éventuellement vide) de points, suivent en fait la même loi. En particulier, si l'on demande la loi d'une variable à densité X , on pourra indifféremment en donner la fonction de répartition ou une densité.

Corollaire 2.6. Soit X une variable aléatoire à densité. Alors, pour tous $a, b \in \mathbb{R}$ tels que $a < b$, l'intégrale ci-dessous converge et de plus :

$$P([a \leq X \leq b]) = P([a \leq X < b]) = P([a < X \leq b]) = P([a < X < b]) = \int_a^b f_X(t)dt.$$

Corollaire 2.7. Soit X une variable aléatoire à densité. Alors, pour tout $x \in \mathbb{R}$, l'intégrale ci-dessous converge et de plus :

$$P([X \geq x]) = 1 - F_X(x) = \int_x^{+\infty} f_X(t)dt.$$

A noter que l'ambiguïté qui règne sur la notion de densité ne pose aucun problème, puisque deux fonctions qui ne diffèrent qu'en un nombre fini de points ont toujours les mêmes intégrales. A présent, nous allons voir comment caractériser les fonctions qui sont des densités de variables aléatoires. Plus précisément :

Théorème 2.8. Une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est une densité d'une variable aléatoire X si et seulement si f est une densité de probabilité, c'est-à-dire si elle vérifie les trois conditions suivantes :

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

2.2 Fonctions d'une variable à densité

Définition 2.9. Soit X une variable aléatoire à densité. Le support de X est l'ensemble $X(\Omega)$.

A noter qu'une variable aléatoire Y de la forme $Y = \varphi(X)$, où X est une variable à densité et φ une application de $X(\Omega)$ dans \mathbb{R} , n'est pas nécessairement une variable à densité. Cependant, dans le cas où φ est strictement monotone de classe \mathcal{C}^1 , on peut montrer que $\varphi(X)$ est une variable à densité dont on peut calculer explicitement une densité. Ce résultat, connu sous le nom de *premier théorème de transfert*, est hors-programme et rarement utilisé dans la pratique. En général, on cherchera plutôt à calculer explicitement la fonction de répartition de $\varphi(X)$ en fonction de celle de X , c'est-à-dire à exprimer $P([\varphi(X) \leq y])$ en fonction de probabilités d'événements liés à X . Si l'on montre que la fonction de répartition de $\varphi(X)$ est continue sur \mathbb{R} et de classe \mathcal{C}^1 sur \mathbb{R} sauf en un nombre fini de points, alors $\varphi(X)$ sera une variable à densité. Il suffira alors de dériver la fonction de répartition de $\varphi(X)$ pour obtenir une densité de $\varphi(X)$. A noter qu'on ne donnera pas ici l'énoncé général du premier théorème de transfert, mais on en retiendra les deux cas particuliers suivants :

Théorème 2.10. Soit X une variable aléatoire à densité, de densité f et soit $(a, b) \in \mathbb{R}^2$ tel que $a \neq 0$. Alors $Y = aX + b$ est une variable aléatoire à densité, dont une densité g est donnée pour tout $y \in \mathbb{R}$ par :

$$g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right).$$

Théorème 2.11. Soit X une variable aléatoire à densité, de densité f . Alors $Y = X^2$ est une variable aléatoire à densité, dont une densité g est nulle sur \mathbb{R}_- et donnée pour tout $y > 0$ par :

$$g(y) = \frac{1}{2\sqrt{y}} [f(\sqrt{y}) + f(-\sqrt{y})].$$

A noter que, conformément au programme, les étudiants doivent savoir retrouver ces résultats par eux-mêmes, c'est-à-dire sans utiliser les théorèmes ci-dessus.

3 Moments d'une variable aléatoire à densité

3.1 Espérance d'une variable aléatoire à densité

Définition 3.1. Soit X une variable aléatoire à densité, de densité f_X . On dit que X admet une espérance si l'intégrale $\int_{-\infty}^{+\infty} tf_X(t)dt$ converge absolument. Dans ce cas, l'espérance de X est définie par :

$$E(X) = \int_{-\infty}^{+\infty} tf_X(t)dt.$$

En particulier, on voit que $E(1) = 1$. A noter qu'il existe des variables aléatoires à densité qui ne possèdent pas d'espérance. Comme pour les variables aléatoires discrètes, l'espérance de X (si elle existe) est la moyenne des valeurs de X pondérée par la densité de probabilité qu'a la variable aléatoire X de prendre ces valeurs. A noter que, comme toute densité est une fonction positive, X admet une espérance si et seulement si les intégrales $\int_{-\infty}^0 tf_X(t)dt$ et $\int_0^{+\infty} tf_X(t)dt$ convergent, c'est-à-dire si l'intégrale $\int_{-\infty}^{+\infty} tf_X(t)dt$ converge. Dès lors, la convergence absolue équivaut ici à la convergence.

Théorème 3.2. (Transfert) Soit X une variable aléatoire à densité, de densité f_X nulle en dehors d'un intervalle $]a, b[$ avec $-\infty \leq a < b \leq +\infty$. Soit φ une application de $]a, b[$ dans \mathbb{R} , continue sur $]a, b[$ sauf peut-être en un nombre fini de points. Alors $\varphi(X)$ admet une espérance si et seulement si l'intégrale $\int_a^b \varphi(t)f_X(t)dt$ converge absolument. Dans ce cas, l'espérance de $\varphi(X)$ est donnée par :

$$E(\varphi(X)) = \int_a^b \varphi(t)f_X(t)dt.$$

A noter que la transformée $\varphi(X)$ d'une variable aléatoire X à densité n'est pas toujours une variable aléatoire à densité. On fera donc bien attention à respecter les hypothèses du théorème de transfert ! De façon générale, le théorème de transfert est particulièrement utile pour calculer l'espérance (si elle existe) d'une variable aléatoire de la forme $\varphi(X)$, où X est une variable à densité, directement à partir d'une densité de X et sans avoir à calculer la loi de $\varphi(X)$ au préalable. Bien utilisé, il permet de réduire considérablement les calculs.

Théorème 3.3. (Existence de l'espérance par domination) Soient X, Y deux variables aléatoires à densité, telles que $|X| \leq Y$ presque sûrement. Si Y admet une espérance, alors X admet aussi une espérance et de plus : $|E(X)| \leq E(Y)$.

Rappelons que la condition " $|X| \leq Y$ presque sûrement" signifie que $P(|X| > Y) = 0$. Dans la pratique, on utilise le théorème d'existence de l'espérance par domination comme pour les variables aléatoires discrètes, et ce comme suit. Pour montrer qu'une variable à densité X admet une espérance, il suffit de majorer $|X|$ par une variable à densité Y plus simple, dont on sait qu'elle admet une espérance.

Théorème 3.4. (Linéarité de l'espérance) Soient X, Y deux variables aléatoires à densité admettant une espérance et soient $\lambda, \mu \in \mathbb{R}$. Alors $\lambda X + \mu Y$ admet une espérance, et de plus :

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

A noter que la combinaison linéaire $\lambda X + \mu Y$ est bien une variable aléatoire d'après le théorème 1.2. A noter aussi que le théorème 3.4 permet non seulement d'établir l'existence de l'espérance pour une combinaison linéaire de variables aléatoires, mais aussi de la calculer plus facilement en réduisant considérablement les calculs (puisque l'on n'a pas à déterminer la loi de $\lambda X + \mu Y$).

Théorème 3.5. (Positivité de l'espérance) Soit X une variable aléatoire à densité admettant une espérance. Si X est positive presque sûrement, c'est-à-dire si $P([X \geq 0]) = 1$, alors on a : $E(X) \geq 0$.

Directement à partir de la linéarité et de la positivité de l'espérance, on montre alors le :

Théorème 3.6. (Croissance de l'espérance) Soient X, Y deux variables aléatoires à densité admettant une espérance. Si $X \leq Y$ presque sûrement, c'est-à-dire si $P([X \leq Y]) = 1$, alors : $E(X) \leq E(Y)$.

A noter que les théorèmes 3.3, 3.4, 3.5, 3.6 restent toujours vrais pour des variables aléatoires quelconques (discrètes et/ou à densité).

Définition 3.7. Une variable aléatoire à densité X est dite centrée si elle admet une espérance nulle.

Théorème - définition 3.8. Soit X une variable aléatoire à densité admettant une espérance. Alors la variable aléatoire $X - E(X)$ est centrée : on l'appelle la variable aléatoire centrée associée à X .

3.2 Variance d'une variable aléatoire à densité

Tout comme pour les variables aléatoires discrètes, nous allons introduire la notion de variance pour les variables aléatoires à densité, et en donner les principales propriétés.

Définition 3.9. Soit X une variable aléatoire à densité et soit $r \in \mathbb{N}^*$. Si X^r admet une espérance, alors on dit que X admet un moment d'ordre r . Dans ce cas, l'espérance de X^r est appelé le moment d'ordre r de X , et noté $E(X^r)$ ou $m_r(X)$.

A noter que, d'après le théorème de transfert, une variable aléatoire à densité X , de densité f_X , admet un moment d'ordre r si et seulement si l'intégrale $\int_{-\infty}^{+\infty} t^r f_X(t)dt$ converge absolument. A noter aussi que, comme toute densité est une fonction positive, on voit que X admet un moment d'ordre r si et seulement si les intégrales $\int_{-\infty}^0 t^r f_X(t)dt$ et $\int_0^{+\infty} t^r f_X(t)dt$ convergent, c'est-à-dire si $\int_{-\infty}^{+\infty} t^r f_X(t)dt$ converge. Dès lors, la convergence absolue équivaut ici à la convergence.

Théorème 3.10. Soit X une variable aléatoire à densité. Si X admet un moment d'ordre r pour un certain entier $r \geq 1$, alors X admet un moment d'ordre k pour tout entier naturel $k \leq r$.

En particulier, une telle variable aléatoire admet une espérance puisque, sous réserve d'existence, l'espérance de X n'est ni plus ni moins que son moment d'ordre 1.

Théorème - définition 3.11. Soit X une variable aléatoire à densité. Si X admet un moment d'ordre 2, alors $X - E(X)$ admet aussi un moment d'ordre 2. Dans ce cas, on dit que X admet une variance, et la variance de X est le moment d'ordre 2 de $X - E(X)$, c'est-à-dire : $V(X) = E((X - E(X))^2)$.

En particulier, si f_X est une densité de X , alors on voit avec le théorème de transfert que la variance de X (si elle existe) est donnée par :

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

En particulier, $V(X)$ est un réel positif, ce qui justifie la :

Définition 3.12. Soit X une variable aléatoire à densité admettant un moment d'ordre 2. Alors l'écart-type de X est le réel $\sigma(X) = \sqrt{V(X)}$.

En d'autres termes, on peut interpréter l'écart-type de X (s'il existe) comme la racine carrée de la moyenne des carrés des écarts de la variable aléatoire X avec son espérance. De façon plus mathématique et comme pour les variables aléatoires discrètes, on dira que la variance de X mesure la dispersion de X par rapport à son espérance. A noter que les variables aléatoires à densité présentent une différence remarquable par rapport aux variables aléatoires discrètes, à savoir le :

Théorème 3.13. Soit X une variable aléatoire à densité admettant une variance. Alors : $V(X) > 0$.

Grosso modo, on peut interpréter ce résultat de la façon suivante. Si une variable aléatoire admet une variance nulle, alors cela signifie qu'elle est égale presque sûrement à son espérance, mais ceci est impossible pour une variable à densité puisqu'une telle variable ne charge aucun des points de \mathbb{R} , au contraire d'une variable aléatoire discrète.

Théorème 3.14. (Formule de Koenig-Huygens) Soit X une variable aléatoire à densité admettant une variance. Alors :

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

La formule de Koenig-Huygens est notamment très utile pour faire des calculs de variance. Dans la pratique, on s'en servira systématiquement pour les calculs.

Théorème 3.15. Soit X une variable aléatoire à densité admettant une variance. Alors, pour tous réels a et b , la variable aléatoire $aX + b$ admet une variance, et de plus :

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

Définition 3.16. Soit X une variable aléatoire à densité admettant une variance. On dit que X est centrée réduite si $E(X) = 0$ et $\sigma(X) = 1$.

Théorème - définition 3.17. Soit X une variable aléatoire à densité admettant une variance non nulle. Alors la variable aléatoire $X^* = \frac{X - E(X)}{\sigma(X)}$ est une variable aléatoire à densité, centrée et réduite. On l'appelle la variable aléatoire centrée réduite associée à X .

3.3 Somme de variables aléatoires à densité

Dans ce paragraphe, nous allons donner (sans démonstration) quelques résultats concernant les sommes de variables aléatoires à densité. Plus précisément, on montre les :

Théorème 3.18. (Produit de convolution) Soient X, Y deux variables aléatoires indépendantes à densité, de densités respectives f_X et f_Y . Si la fonction $h : x \mapsto \int_{-\infty}^{+\infty} f_X(t) f_Y(x - t) dt$ est définie et continue sur \mathbb{R} sauf peut-être en un nombre fini de points, alors h est une densité de $X + Y$.

A noter que la fonction h (si elle existe et est bien définie) est appelée le produit de convolution de f_X et f_Y , et notée $f_X * f_Y$. A noter aussi que $f_X * f_Y$ est bien défini et continu sur \mathbb{R} si l'une des densités f_X, f_Y est bornée. L'intérêt du produit de convolution vient évidemment de ce qu'il permet de calculer une densité pour une somme de variables aléatoires indépendantes. De plus, par une récurrence facile, la propriété de linéarité de l'espérance se généralise à n variables aléatoires. Plus précisément :

Théorème 3.19. (Linéarité de l'espérance) Soient X_1, \dots, X_n des variables aléatoires quelconques qui admettent toutes une espérance. Alors, pour tous réels $\lambda_1, \dots, \lambda_n$, la variable aléatoire $X = \lambda_1 X_1 + \dots + \lambda_n X_n$ admet aussi une espérance et de plus :

$$E(\lambda_1 X_1 + \dots + \lambda_n X_n) = \lambda_1 E(X_1) + \dots + \lambda_n E(X_n).$$

Enfin, comme pour les variables aléatoires discrètes, on a le :

Théorème 3.20. Soient X, Y deux variables aléatoires indépendantes à densité, qui admettent toutes deux une variance. Alors :

1. XY admet une espérance et de plus : $E(XY) = E(X)E(Y)$.
2. $X + Y$ admet une variance et de plus : $V(X + Y) = V(X) + V(Y)$.

A noter que ce résultat est encore valable pour des variables aléatoires quelconques (discrètes et/ou à densité). Plus généralement, on montre par récurrence et à l'aide du lemme des coalitions que, pour toutes variables aléatoires X_1, \dots, X_n , indépendantes et admettant toutes une espérance, la variable aléatoire $X_1 \dots X_n$ admet aussi une espérance et :

$$E(X_1 \dots X_n) = E(X_1) \dots E(X_n).$$

De la même façon, on montre que, pour toutes variables aléatoires X_1, \dots, X_n , indépendantes et admettant toutes une variance, la variable aléatoire $X_1 + \dots + X_n$ admet aussi une variance et :

$$V(X_1 + \dots + X_n) = V(X_1) + \dots + V(X_n).$$

4 Les lois usuelles

Nous terminerons ce chapitre par la présentation de quelques lois continues classiques, c'est-à-dire de lois de probabilité parmi les plus connues pour les variables aléatoires à densité. Pour chacune d'entre elles, nous en donnerons les caractéristiques principales.

4.1 La loi uniforme sur $[a, b]$

Définition 4.1. Soient $a, b \in \mathbb{R}$ tels que $a < b$. On dit qu'une variable aléatoire X suit la loi uniforme sur $[a, b]$, et on note $X \hookrightarrow \mathcal{U}([a, b])$, si X admet une densité donnée par :

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b] \end{cases}.$$

A noter qu'on peut choisir le support de X comme étant l'intervalle $[a, b]$, c'est-à-dire : $X(\Omega) = [a, b]$.

Théorème 4.2. Si $X \hookrightarrow \mathcal{U}([a, b])$, alors sa fonction de répartition est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}.$$

Théorème 4.3. Soit X une variable aléatoire et soient $a, b \in \mathbb{R}$ tels que $a < b$. Alors $X \hookrightarrow \mathcal{U}([0, 1])$ si et seulement si $(b - a)X + a \hookrightarrow \mathcal{U}([a, b])$.

Théorème 4.4. Si $X \hookrightarrow \mathcal{U}([0, 1])$, alors X admet une espérance et une variance données par :

$$E(X) = \frac{1}{2} \quad \text{et} \quad V(X) = \frac{1}{12}.$$

Théorème 4.5. Si $X \hookrightarrow \mathcal{U}([a, b])$, alors X admet une espérance et une variance données par :

$$E(X) = \frac{a+b}{2} \quad \text{et} \quad V(X) = \frac{(b-a)^2}{12}.$$

4.2 La loi exponentielle de paramètre λ

Définition 4.6. Soit λ un réel > 0 . On dit qu'une variable aléatoire X suit la loi exponentielle de paramètre λ , et on note $X \hookrightarrow \mathcal{E}(\lambda)$, si X admet une densité donnée par :

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} .$$

A noter qu'on peut choisir le support comme étant l'intervalle \mathbb{R}_+ , c'est-à-dire : $X(\Omega) = \mathbb{R}_+$.

Théorème 4.7. Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors sa fonction de répartition est donnée par :

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases} .$$

Théorème 4.8. Soit X une variable aléatoire et soit $\lambda \in \mathbb{R}_+^*$. Alors : $X \hookrightarrow \mathcal{E}(\lambda) \iff \frac{1}{\lambda}X \hookrightarrow \mathcal{E}(1)$.

Théorème 4.9. Si $X \hookrightarrow \mathcal{E}(\lambda)$, alors X admet une espérance et une variance données par :

$$E(X) = \frac{1}{\lambda} \quad \text{et} \quad V(X) = \frac{1}{\lambda^2} .$$

Définition 4.10. Soit X une variable aléatoire. On dit que X suit une loi sans mémoire si X ne prend que des valeurs ≥ 0 , et si pour tous $(x, y) \in \mathbb{R}_+^2$, on a : $P([X > x + y]) = P([X > x])P([X > y])$.

Théorème 4.11. Soit X une variable aléatoire. Si X suit une loi sans mémoire, alors soit X est nulle presque sûrement, soit X suit une loi exponentielle.

4.3 La loi gamma γ

Définition 4.12. La fonction Γ est l'application définie pour tout $\nu \in \mathbb{R}_+^*$ par : $\Gamma(\nu) = \int_0^{+\infty} t^{\nu-1} e^{-t} dt$.

Rappelons que la convergence de cette intégrale a été établie au chapitre "Intégrales impropres".

Définition 4.13. Soit $\nu \in \mathbb{R}_+^*$. On dit qu'une variable aléatoire X suit la loi $\gamma(\nu)$, et on note $X \hookrightarrow \gamma(\nu)$, si X admet une densité définie par :

$$f_X(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ \frac{1}{\Gamma(\nu)} t^{\nu-1} e^{-t} & \text{si } t > 0 \end{cases} .$$

A noter qu'on peut choisir le support comme étant l'intervalle \mathbb{R}_+ , c'est-à-dire : $X(\Omega) = \mathbb{R}_+$. A noter aussi qu'en général (c'est-à-dire si $\nu \notin \mathbb{N}^*$), la fonction de répartition d'une loi gamma de paramètre ν ne peut être explicitement décrite à l'aide des fonctions usuelles, et c'est pourquoi on n'en donnera pas d'expression pour tout $\nu > 0$! A noter enfin que la loi $\mathcal{E}(1)$ correspond à la loi $\gamma(1)$.

Théorème 4.14. Si $X \hookrightarrow \gamma(\nu)$, alors X admet une espérance et une variance données par :

$$E(X) = \nu \quad \text{et} \quad V(X) = \nu .$$

A noter que les lois gamma vérifient des propriétés remarquables vis-à-vis de l'addition, à savoir :

Théorème 4.15. (Stabilité de la loi gamma par addition) Si X_1, \dots, X_n sont des variables aléatoires indépendantes telles que $X_i \hookrightarrow \gamma(\nu_i)$ pour tout $i \in \llbracket 1, n \rrbracket$, alors : $\sum_{i=1}^n X_i \hookrightarrow \gamma(\sum_{i=1}^n \nu_i)$.

Corollaire 4.16. Si X_1, \dots, X_n sont des variables aléatoires indépendantes telles que $X_i \hookrightarrow \mathcal{E}(1)$ pour tout $i \in \llbracket 1, n \rrbracket$, alors : $\sum_{i=1}^n X_i \hookrightarrow \gamma(n)$.

4.4 La loi normale ou de Laplace-Gauss

Définition 4.17. Soient $m, \sigma \in \mathbb{R}$ tels que $\sigma > 0$. On dit qu'une variable aléatoire X suit la loi normale de paramètres m, σ^2 , et on note $X \hookrightarrow \mathcal{N}(m, \sigma^2)$, si X admet une densité donnée par :

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

A noter que $X(\Omega) = \mathbb{R}$. A noter aussi que cette loi porte aussi le nom de *loi gaussienne*. A noter enfin que la fonction de répartition d'une loi normale ne peut être explicitement décrite à l'aide des fonctions usuelles, et c'est pourquoi on ne la donnera pas ici !

Définition 4.18. On dit qu'une variable aléatoire X suit la loi normale centrée réduite, et on note $X \hookrightarrow \mathcal{N}(0, 1)$, si X admet une densité donnée par :

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

A noter que la fonction de répartition Φ de la loi normale centrée réduite est définie pour tout $x \in \mathbb{R}$ par :

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

A noter aussi que, comme une densité d'une variable normale centrée réduite X est une fonction paire, alors on montre facilement que :

$$\boxed{\forall x \in \mathbb{R}, \quad \Phi(x) = 1 - \Phi(-x) \text{ et } \Phi(0) = \frac{1}{2}.}$$

Théorème 4.19. Soit X une variable aléatoire et soient $a, b \in \mathbb{R}$ tels que $a \neq 0$. Si $X \hookrightarrow \mathcal{N}(m, \sigma^2)$, alors $aX + b \hookrightarrow \mathcal{N}(am + b, a^2\sigma^2)$.

Corollaire 4.20. Soit X une variable aléatoire et soient $m, \sigma \in \mathbb{R}$ tels que $\sigma \neq 0$. Alors $X \hookrightarrow \mathcal{N}(m, \sigma^2)$ si et seulement si $\frac{X-m}{\sigma} \hookrightarrow \mathcal{N}(0, 1)$.

Théorème 4.21. Si $X \hookrightarrow \mathcal{N}(0, 1)$, alors X admet une espérance et une variance données par :

$$E(X) = 0 \quad \text{et} \quad V(X) = 1.$$

Théorème 4.22. Si $X \hookrightarrow \mathcal{N}(m, \sigma^2)$, alors X admet une espérance et une variance données par :

$$E(X) = m \quad \text{et} \quad V(X) = \sigma^2.$$

A noter enfin que la loi normale vérifie des propriétés remarquables vis-à-vis de l'addition, à savoir :

Théorème 4.23. (Stabilité de la loi normale par addition) Si X_1, \dots, X_n sont des variables aléatoires indépendantes telles que $X_i \hookrightarrow \mathcal{N}(m_i, \sigma_i^2)$ pour tout $i \in \llbracket 1, n \rrbracket$, alors : $\sum_{i=1}^n X_i \hookrightarrow \mathcal{N}\left(\sum_{i=1}^n m_i, \sum_{i=1}^n \sigma_i^2\right)$.