Séance 8 : Variables aléatoires réelles

Nous avons vu au chapitre sur les lois discrètes la définition générale d'une variable aléatoire X. Dans ce chapitre nous avons abordé le cas - facile - quand lorsque l'espace $\mathcal{X}$ des valeurs de X était fini ou dénombrable. Dans le présent chapitre nous allons nous contenter de résoudre, sans démonstrations complètes, le cas plus général (et beaucoup plus difficile) de $\mathcal{X} = \mathbb{R}$.

Fonction de répartition et densité

Définition 1 *La fonction de répartition (f.d.r.) de la variable aléatoire X sur $\mathbb{R}$ est la fonction suivante :*

$$F_X(x) = \mathbb{P}(X \in]-\infty, x]) = \mathbb{P}(X \leqslant x).$$

Propriétés :

1. la fonction $F_X(x)$ est croissante, continue à droite,
$$\lim_{x\to-\infty} F_X(x) = 0; \quad \lim_{x\to+\infty} F_X(x) = 1.$$
2. Comme F_X est croissante, elle admet une limite à gauche en chaque point, limite qu'on notera $F_X(x^-)$. Nous avons :
$$P(X \in]x;y]) = F_X(y) - F_X(x); \qquad P(X \in]x;y[) = F_X(y^-) - F_X(x);$$
$$P(X \in [x;y]) = F_X(y) - F_X(x^-); \qquad P(X \in [x;y[) = F_X(y^-) - F(x).$$
En particulier, $P(X = x) = F_X(x) - F_X(x^-)$ est le "saut" de la fonction F_X au point x. On a donc $P(X = x) = 0$ pour tout x si et seulement si la fonction F_X est continue en tout point.

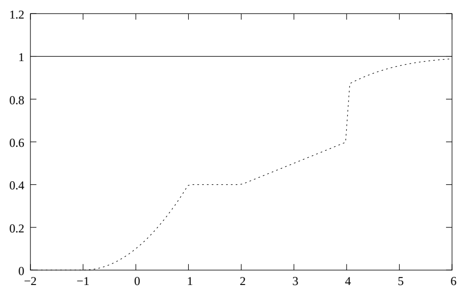

FIG. 1 – Fonction de répartition d'une variable aléatoire à valeur dans $[-1,1] \cup [2,+\infty[$ et $\mathbb{P}(X = 4) = 0.25$.

Dans la suite on suppose qu'il existe une fonction $f_X(x)$ bornée sur $\mathbb{R}$ telle que

$$F_X(x) = \int_{-\infty}^{x} f(t)dt.$$

On appelle une telle fonction **la densite** de X et on dit que $F_X(x)$ admet une densité, bien évidemment, f_X vérifie $f_X(x) = F'_X(x)$ et

$$\mathbb{P}(X \in [x,y]) = \mathbb{P}(X \in]x,y]) = \int_x^y f_X(t)dt, \quad \int_{-\infty}^{+\infty} f(t)dt = 1.$$

Remarque : il existe bien-sûr des probabilités sur $\mathbb{R}$ qui n'ont pas de densité : c'est le cas des probabilités discrètes. Voici une interprétation "intuitive" de la densité f_X de F_X. Si dx est un "petit" accroissement de la variable x, on a (si du moins f est continue en x) :

$$f_X(x) \sim \frac{\mathbb{P}(X \in [x, x+dx[)}{dx}.$$

En théorie des probabilités on montre que si X_1, X_2, ..., X_m sont des variables aléatoires et $g(x_1, ..., x_m)$ une fonction continue, alors $Y = g(X_1, ..., X_m)$ est, elle aussi, une variable aléatoire. Comme application de cet énoncé, on vérifie que si (X_i), $i = 1, ..., m$, X et Y sont des variables aléatoires réelles, alors

$$X + Y, \quad \frac{X}{Y} \text{ si } Y \neq 0 \quad \text{sont des v.a.}$$
$$\max_{1 \leqslant i \leqslant m} X_i, \quad \min_{1 \leqslant i \leqslant m} X_i \quad \text{sont des v.a.}$$

Lois usuelles

Loi uniforme $\mathcal{U}_{[a,b]}$

La loi uniforme sur $[a, b]$: on a ici deux réels $a < b$, et c'est la f.d.r. admettant la densité

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si} \quad a \leqslant x \leqslant b, \\ 0 & \text{sinon} \quad . \end{cases}$$

En vertu de la première remarque ci-dessus, on aurait aussi bien pu choisir $f(a) = 0$ ou $f(b) = 0$. Au vu de l'interprétation, le fait que f soit constante sur $[a, b]$ correspond au fait que si on choisit une point selon cette loi, on a "autant de chances" de tomber au voisinage de chaque point de l'intervalle $[a, b]$, ce qui explique le nom "uniforme". Remarquer aussi que $P(X = x) = 0$ pour tout x (comme pour toutes les lois avec densité) : on a donc une probabilité nulle de tomber exactement en un point x fixé à l'avance.

Loi exponentielle $\mathcal{E}(\lambda)$

La loi exponentielle de paramètre $\lambda > 0$ est la loi de densité

$$f(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ \lambda e^{-\lambda x} & \text{sinon} \quad . \end{cases}$$

et de fonction de répartition

$$F(x) = \begin{cases} 0 & \text{si} \quad x < 0, \\ (1 - e^{-\lambda x}) & \text{sinon} \quad . \end{cases}$$

Notons que la loi exponentielle jouit aussi d'une propriété importante pour les applications (propriété de "non-vieillissement") : soit X une variable aléatoire positive, telle que $P(X > s) > 0$ pour tout $s \in \mathbb{R}$. On a

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s)$$

pour tous s et $t > 0$ si et seulement si X suit une loi exponentielle.

Loi normale $\mathcal{N}(0,1)$

La loi normale centrée réduite (ou loi de Gauss) : c'est la loi de densité

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

Pour vérifier que cette fonction est d'intégrale 1, on remarque que $I = \int_{-\infty}^{+\infty} f(t)dt$ vérifie

$$I^2 = \int\int_{\mathbb{R}^2} f(x)f(y)dxdy = \int_0^{2\pi} \frac{1}{2\pi} d\theta \int_0^{+\infty} e^{-\rho^2/2}\rho d\rho.$$

En passant en coordonnées polaires dans l'intégrale double), et un calcul simple montre alors que $I^2 = 1$.

Espérance de variables aléatoires de loi continue

Définition 2 (Espérance) *Si l'intégrale $\int_{-\infty}^{+\infty} |x| f_X(x)dx$ est finie, on appelle la valeur*

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

l'espérance mathématique de la v.a. X.

Remarque : on démontre dans la théorie de probabilités que si $g(x)$ est une fonction telle que $\int |g(x)| f_X(x)dx$ est finie, alors l'espérance de la v.a. $Y = g(X)$ est

$$\mathbb{E}Y = \int g(x) f_X(x)dx$$

Soit X et Y variables aléatoires d'espérance bien définie, et a, b des réels alors

$$\mathbb{E}(aX+b) = a\mathbb{E}(X) + b \text{ et } \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$$

("linéarité" de l'esperance). Ceci implique que si $\mathbb{E}(X) = \mu$ et $Y = X - \mu$, alors

$$\mathbb{E}Y = \mathbb{E}(X-\mu) = \mathbb{E}X - \mu = 0.$$

On dit que la variable aléatoire Y est **centrée**.

Si la v.a. X est "carré-integrable", i.e. $\int x^2 f_X(dx) < \infty$, alors

$$\mathbb{E}(X^2) = \int x^2 f_X(x)dx.$$

Définition 3 (Variance) *Si la v.a. X est carré-integrable, sa variance est l'espérance de la variable aléatoire $[X - \mathbb{E}(X)]^2$, et on la note $Var(X)$ ou σ_X^2. Elle est positive, et sa racine carrée positive s'appelle l'écart-type de X, qu'on note σ_X.*

Notons que, comme dans le cas des v.a. discrètes,

$$\mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2X\mathbb{E}(X) + \mathbb{E}^2(X)) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}^2(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X),$$

et $Var(X) = \mathbb{E}(X^2) - \mathbb{E}^2(X)$.

Notons que si $Y = aX + b$,

$$Var(Y) = \mathbb{E}((Y-\mathbb{E}(Y))^2) = \mathbb{E}((aX+b-[a\mathbb{E}(X)+_b])^2) = \mathbb{E}(a^2(X-\mathbb{E}(X))^2) = a^2 Var(X).$$

Donc, si $\mathbb{E}(X) = \mu$ et $Var(X) = \sigma^2$ et $Y = \frac{X-\mu}{\sigma}$,

$$\mathbb{E}Y = 0 \text{ et } Var(Y) = \sigma^{-2} Var(X - \mathbb{E}(X)) = \sigma^{-2} Var(X) = 1.$$

On dit que Y est **centrée-réduite**.

1. Soit X est une v.a. de loi uniforme sur $[a, b]$ Alors
$$\mathbb{E}(X) = \frac{b-a}{2} \ \text{ et } \ Var(X) = \frac{(b-a)^2}{12}.$$
2. Si X suit la loi exponentielle de paramètre λ, alors
$$\mathbb{E}X = \frac{1}{\lambda} \ \text{ et } \ Var(X) = \frac{1}{\lambda^2}.$$
3. Pour une v.a. normale centrée-réduite Y l'espérance $\mathbb{E}Y = 0$ et $Var(Y) = 1$. Notons que si $X = \sigma Y + \mu$, $\mathbb{E}X = \mu$ et $Var(X) = \sigma^2$. La densité de la loi de X est
$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$
Cette loi est notée $\mathcal{N}(\mu, \sigma^2)$ (loi normale avec la moyenne μ et la variance σ^2).

Variables aléatoires indépendantes

Soit $Z = (X_1, ..., X_n)$ un **vecteur aléatoire** dans $\mathbb{R}^n$. Comme dans le cas d'une v.a. on définit la fonction de repartition **jointe** de Z en tout point $z = (z_1, ..., z_n)$ de $\mathbb{R}^n$:

$$F_Z(z) = P(X_1 \le z_1,\ X_2 \le z_2,\ ..., X_n \le z_n).$$

Dans la suite on suppose que la fonction de repartition jointe possède la **densité jointe** :

$$f_Z(z) = \frac{\partial^n}{\partial z_1 \ldots \partial z_n} F_z(z).$$

Comme dans le cas des v.a discrètes nous avons

Définition 4 (Indépendance) *Soit $Z = (X_i)$ $i = 1, ..., n$ un vecteur aléatoire, composé de variables aléatoires réelles (scalaires), i.e. chaque X_i est à valeurs dans $\mathbb{R}$. Les v.a. $X_1, ..., X_n$ sont indépendantes si pour tous $z = (z_1, ..., z_n)$*

$$\mathbb{P}(X_1 \leqslant z_1,\ ...,\ X_n \leqslant z_n) = \mathbb{P}(X_1 \leqslant z_1)...\mathbb{P}(X_n \leqslant z_n)$$

ou bien

$$f_Z(z) = f_{X_1}(z_1)...f_{X_n}(z_n)$$

Si, de plus, les densités $f_{X_1}, ..., f_{X_n}$ sont les mêmes, on dit que les v.a. $X_1, ..., X_n$ sont indépendantes et identiquement distribuées (i.i.d.).

Cette definition s'étend sans peine à une suite infinie de v.a..

Proprietes :
- Soit X et Y à valeurs dans $\mathbb{R}$. Soit aussi $g(x)$ et $h(x)$ deux fonctions continues, telles que $g(X)$ et $h(Y)$ soient aussi des variables aléatoires. Si X et Y sont indépendantes, les variables aléatoires $g(X)$ et $h(Y)$ sont aussi indépendantes. Si de plus $\mathbb{E}g(X)$ et $\mathbb{E}h(Y)$ existent, on a
$$\mathbb{E}g(X)h(Y) = \mathbb{E}g(X)\mathbb{E}h(Y).$$
- soit (X_i), $i = 1, 2, ..., n$ une suite de v.a. indépendantes. Alors (comme dans le cas général) $\mathbb{E}((X_1 + ... + X_n) = \mathbb{E}X_1 + ... + \mathbb{E}X_n$ et $Var(X_1 + ... + X_n) = Var(X_1) + ... + Var(X_n)$.