II.4 Moment d'ordre 2 et variance d'une variable réelle

Dans ce paragraphe et le suivant, les variables aléatoires sont discrètes et réelles.
Comme dans le cas des variables finies, la valeur moyenne n'est pas une donnée suffisante pour avoir une bonne compréhension d'une variable aléatoire. En effet, l'espérance ne mesure pas la dispersion de la variable aléatoire : les valeurs prises peuvent tout aussi bien rester toujours très proches de $\mathbb{E}(X)$, ou alors s'en éloigner beaucoup, de part et d'autre, de façon à ce qu'en moyenne cela se compense.
Pour des raisons techniques, il est plus facile de travailler avec la moyenne quadratique $\mathbb{E}((X - \mathbb{E}(X))^2)$ de la distance à $\mathbb{E}(X)$, plutôt que la moyenne arithmétique $\mathbb{E}(|X - \mathbb{E}(X)|)$. En effet, ce choix nous ramène à des méthodes bilinéaires.
On a déjà défini les moments d'ordre k. Les seuls au programme sont les moments d'ordre 1 (l'espérance) et 2.

Définition II.23 (Moment d'ordre 2)**.** *Soit X une variable aléatoire réelle discrète.*
- Le moment d'ordre 2 de X est l'espérance de X^2. Ce moment d'ordre 2 est toujours bien défini dans $\mathbb{R}_+$.
- Ainsi, d'après la formule de transfert,

$$\mathbb{E}(X^2) = \sum_{x \in X(\Omega)} x^2 \mathbb{P}(X = x).$$

- On dit que X admet un moment d'ordre 2 fini (et on note $X \in L^2$) si $\mathbb{E}(X^2) < +\infty$.

Proposition II.24 (Espace des variables L^2 sur Ω)**.** *1. Soit X et Y deux variables aléatoires réelles discrètes sur Ω. Si X et Y sont dans L^2, alors $XY \in L^1$.*
2. L'ensemble $L^2(\Omega, \mathbb{R})$ des variables aléatoires définies sur Ω et admettant un moment d'ordre 2 fini est un espace vectoriel sur $\mathbb{R}$.

Corollaire II.25 (Espérance d'une variable L^2)**.** *Soit X une variable aléatoire réelle discrète. Si $X \in L^2$, alors $X \in L^1$.*

Éléments de preuve. Prendre $Y = 1$. □

Théorème II.26 (Inégalité de Cauchy-Schwarz pour l'espérance)**.** *- Pour toutes variables aléatoires réelles discrètes X et Y dans L^2,*

$$\mathbb{E}(XY)^2 \leqslant \mathbb{E}(X^2)\mathbb{E}(Y^2).$$

- Dans l'inégalité précédente, le cas d'égalité est obtenu si et seulement si X et Y sont colinéaires presque sûrement, donc s'il existe des constantes λ et μ réelles telles que $\lambda X + \mu Y \sim 0$.

Éléments de preuve. C'est l'inégalité de Cauchy-Schwarz pour la forme bilinéaire symétrique positive $(X, Y) \mapsto \mathbb{E}(XY)$ définie sur $L^2(\Omega, \mathbb{R})$.
On rappelle que l'inégalité de Cauchy-Schwarz pour une forme bilinéaire symétrique positive ϕ peut être obtenue sans l'hypothèse de séparation, et que le cas d'égalité est alors obtenu pour le couple (x, y) si et seulement s'il existe λ et μ tels que $\phi(\lambda x + \mu y, \lambda x + \mu y) = 0$. □

Afin d'alléger les énoncés, nous ne rappelons plus par la suite à chaque fois que les variables considérées sont réelles discrètes (ce que nous sous-entendrons par l'écriture $X \in L^2$ que nous n'avons définie que dans ce cas).

Définition II.27 (Variance et écart-type d'une variable aléatoire)**.** *Soit $X \in L^2$. Alors, en vertu de ce qui précède, $X - \mathbb{E}(X)$ est bien défini et L^2. On définit alors :*
- la variance de X :

$$\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \sum_{x \in X(\Omega)} (x - \mathbb{E}(X))^2 \mathbb{P}(X = x).$$

- l'écart type de X :

$$\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{\mathbb{E}((X - \mathbb{E}(X))^2)}.$$

Proposition II.28 (Positivité et stricte positivité de $\mathbb{V}$)**.** *Soit $X \in L^2$. Alors $\mathbb{V}(X) \geqslant 0$, l'égalité étant obtenue si et seulement si X est constante presque sûrement.*

Dans la pratique, on calcule souvent la variance avec le résultat suivant.

Théorème II.29 (Formule de König-Huygens)**.** *Soit $X \in L^2$. Alors*

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

Proposition II.30 (Quadraticité et invariance par translation de la variance)**.** *Soit a et b deux réels et $X \in L^2$. Alors $aX + b \in L^2$, et*

$$\mathbb{V}(aX + b) = a^2 \mathbb{V}(X).$$

Il arrive souvent de normaliser les variables aléatoires, afin de pouvoir les comparer. Il faut dans ce cas à la fois normaliser la valeur moyenne et la dispersion.

Définition II.31 (Variable centrée, réduite)**.** *Une variable X est dite :*
- centrée si $X \in L^1$, et $\mathbb{E}(X) = 0$ - réduite si $X \in L^2$, et $\mathbb{V}(X) = 1$ - centrée réduite si elle est à la fois centrée et réduite.

Proposition II.32 (Variable centrée réduite associée à X)**.** *Soit $X \in L^2$ telle que $\sigma(X) > 0$. Il existe d'uniques réels $a > 0$ et b tels que la variable X^* définie par $X^* = aX + b$ soit centrée réduite. Plus précisément,*

$$X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}.$$

On dit que X^* est la variable aléatoire centrée réduite associée à X.

Éléments de preuve. Analyse synthèse par linéarité de $\mathbb{E}$ et quadraticité de $\mathbb{V}$. □

II.5 Covariance

Pour l'instant, contrairement au cas de l'espérance, nous n'avons pas donné de moyen de calculer la variance d'une somme. Nous avons en effet besoin pour cela d'un outil supplémentaire, qui nous permet de mesurer la dépendance entre deux variables.
En effet on comprend bien que la variance de $X + Y$ va fortement dépendre de la corrélation entre X et Y. Pour prendre deux cas extrêmes, supposons que X et Y suivent une même loi symétrique par rapport à 0 :
- si $X = Y$, on aura alors $\mathbb{V}(X+Y) = \mathbb{V}(2X) = 4\mathbb{V}(X)$; - si $X = -Y$, on aura alors $\mathbb{V}(X+Y) = 0$.
Ainsi, la seule donnée des lois de X et Y ne peut pas suffire à déterminer $\mathbb{V}(X+Y)$. La covariance $\mathrm{cov}(X, Y)$ va nous permettre de mesurer la propension de deux variables à évoluer de façon parallèle ou opposée.

Définition II.33 (Covariance)**.** *Soit X et Y deux variables L^2. Alors $(X - \mathbb{E}(X))(Y - \mathbb{E}(Y)) \in L^1$ et on définit la covariance de X et Y par :*

$$\mathrm{cov}(X, Y) = \mathbb{E}\big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big).$$