Chapitre 24

Probabilités 3 – Espérance, variance

Plan du chapitre

Hypothèse

Dans tout ce chapitre, $(\Omega, \mathbb{P})$ désigne un espace probabilisé fini (i.e. Ω désigne un univers fini et $\mathbb{P}$ est une probabilité définie sur Ω).
E, F sont des ensembles quelconques.

1 Espérance

1.1 Définition

On rappelle qu'une v.a. est dite réelle (resp. complexe) si elle est à valeurs dans un ensemble $E \subset \mathbb{R}$ (resp. $E \subset \mathbb{C}$).

Définition 24.1

Soit $X : \Omega \to E$ une v.a. réelle (resp. complexe). On définit l'espérance de X comme étant le réel (resp. le complexe) noté

$$\mathbb{E}(X) := \sum_{x \in E} x \mathbb{P}(X = x)$$

Remarque.

- L'espérance d'une variable aléatoire ne dépend ainsi que de sa loi. Si $X \sim Y$, alors $\mathbb{E}[X] = \mathbb{E}[Y]$.
- Bien que E soit infini, la somme ci-dessus est finie : on a plus précisément

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

  où $X(\Omega)$ est fini.

L'espérance d'une v.a. représente sa valeur moyenne, comme l'illustrent les exemples suivants :

Exemple 1. On lance un dé à six faces. On note X la v.a. qui correspond au chiffre obtenu. On a donc

$$\forall i \in [\![1,6]\!] \qquad \mathbb{P}(X = i) = \frac{1}{6}$$

Alors l'espérance de X vaut :

1.2 Espérances des lois usuelles

Proposition 24.2

Si $X \sim \mathcal{U}([\![a,b]\!])$ avec $a, b \in \mathbb{R}$, alors $\mathbb{E}(X) = \dfrac{a+b}{2}$.

Démonstration. Similaire à l'exemple précédent. □

Exemple 2. En particulier, si $a \in E$ et X est la v.a. constante égale à a (on peut alors noter $X = a$), alors

$$\mathbb{E}(X) = \mathbb{E}(a) = a$$

Proposition 24.3

Soit $p \in [0,1]$ et X une v.a. telle que $X \sim \mathcal{B}(p)$. Alors $\mathbb{E}(X) = p$.

Démonstration.

□

Exemple 3. Soit $n \in \mathbb{N}^*$, $p \in [0,1]$ et X une v.a. telle que $X \sim \mathcal{B}(n,p)$. Alors $\mathbb{E}(X) = np$.

Démonstration. Cf Exemple 4. □

1.3 Propriétés de l'espérance

Le lemme qui suit n'a que peu d'intérêt en pratique. Il est néanmoins utile pour les démonstrations des propriétés qui suivent.

Lemme 24.4

Soit $X : \Omega \to E$ une v.a. Alors
$$\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$$

Proposition 24.5 (Linéarité)

Soit X, Y deux v.a. complexes et $\lambda \in \mathbb{C}$.

1. Linéarité :
$$\mathbb{E}(X+Y) = \mathbb{E}(X)+\mathbb{E}(Y)$$
$$\mathbb{E}(\lambda X) = \lambda\mathbb{E}(X)$$
2. Positivité : si X est à valeurs dans $\mathbb{R}_+$ alors
$$\mathbb{E}(X) \geq 0$$
3. Croissance : si X, Y sont des v.a.r. telles que $X \leq Y$, alors
$$\mathbb{E}(X) \leq \mathbb{E}(Y)$$
4. Inégalité triangulaire :
$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$$

Démonstration. On ne montre que la première et la dernière assertion. On a

$$\begin{aligned}
\mathbb{E}(X+Y) &= \sum_{\omega\in\Omega} (X+Y)(\omega)\,\mathbb{P}(\{\omega\}) \\
&= \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\}) + \sum_{\omega\in\Omega} Y(\omega)\,\mathbb{P}(\{\omega\}) \\
&= \mathbb{E}(X)+\mathbb{E}(Y)
\end{aligned}$$

De plus, par l'inégalité triangulaire “classique” :

$$\begin{aligned}
|\mathbb{E}(X)| &= \left|\sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\{\omega\})\right| \\
&\leq \sum_{\omega\in\Omega} |X(\omega)\,\mathbb{P}(\{\omega\})| \\
&= \sum_{\omega\in\Omega} |X(\omega)|\,\mathbb{P}(\{\omega\}) \qquad \text{car } \mathbb{P}(\{\omega\}) \geq 0 \\
&= \mathbb{E}(|X|)
\end{aligned}$$

□

Remarque. Ces propriétés ressemblent énormément à celles des intégrales. C'est tout à fait normal, l'espérance est en fait une “intégrale déguisée”. Une autre propriété connue des intégrales est la suivante : si $f \in C^0([a,b],\mathbb{R}_+)$ est $\boxed{\text{positive}}$, alors

$$\int_a^b f = 0 \qquad \Longleftrightarrow \qquad f \equiv 0$$

L'adaptation aux v.a. donne le résultat suivant : si X est une v.a. à valeurs dans $\boxed{\mathbb{R}_+}$, alors

$$\mathbb{E}(X) = 0 \qquad \Longleftrightarrow \qquad \mathbb{P}(X = 0) = 1$$

On dit alors que X est presque sûrement nulle.

A noter : $\mathbb{E}(X)$ est une constante qui ne dépend que de la loi de X. Ainsi, par linéarité,

$$\mathbb{E}\left(\mathbb{E}(X)\right) = \mathbb{E}(X)\mathbb{E}(1) = \mathbb{E}(X)$$

Exemple 4. Si $X \sim \mathcal{B}(n, p)$, alors $\mathbb{E}(X) = np$.

Définition 24.6

Une v.a. X telle que $\mathbb{E}(X) = 0$ est dite une v.a. centrée.

Exemple 5. Si X est une v.a. quelconque, $X - \mathbb{E}(X)$ est ainsi une v.a. centrée.

1.4 Formule de transfert

Théorème 24.7 (Formule de transfert)

Soit $X : \Omega \to E$ une v.a. et $f : E \to \mathbb{C}$ une application. Alors

$$\mathbb{E}(f(X)) = \sum_{x \in E} f(x)\mathbb{P}(X = x)$$

Remarque. Cette formule est incontournable. Elle permet de calculer l'espérance de $f(X)$ uniquement à partir de la loi de X (et non celle de $f(X)$).

Exemple 6. On suppose que $X \sim \mathcal{B}(n, p)$ avec $p > 0$. Déterminer la loi de $Y = \dfrac{1}{1+X}$.

1.5 Espérance d'un produit de v.a. indépendantes

Proposition 24.8

Soit X, Y deux v.a. complexes *indépendantes*. Alors

$$\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$$

Plus généralement, si $X_1, \cdots, X_n$ sont n v.a. indépendantes, alors

$$\mathbb{E}(X_1 \cdots X_n) = \mathbb{E}(X_1) \cdots \mathbb{E}(X_n)$$

Exemple 7. On considère $X_1, \cdots, X_n$ des v.a. indépendantes suivant la même loi de Bernoulli $p \in [0,1]$. Alors

$$\mathbb{E}(X_1 \cdots X_n) = \prod_{i=1}^{n} \mathbb{E}(X_i) = \prod_{i=1}^{n} p = p^n$$

2 Variance et covariance

2.1 Définition

Définition 24.9 (Variance)

Soit X une v.a.r. définie sur Ω. On appelle variance de X le réel positif :

$$\mathbb{V}(X) := \mathbb{E}\left((X - \mathbb{E}(X))^2 \right) \qquad \geq 0$$

Autrement dit, on considère la v.a. centrée $X - \mathbb{E}(X)$, et on calcule l'espérance de son carré. Plus $\mathbb{V}(X)$ est grand, plus X prend des valeurs dispersées autour de sa moyenne.

Proposition 24.10

Soit X une v.a.r. définie sur Ω. Alors

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

Démonstration. On a

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}\left((X - \mathbb{E}(X))^2 \right) \\ &= \mathbb{E}\left(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2\right) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}\left(\mathbb{E}(X)X\right) + \mathbb{E}\left(\mathbb{E}(X)^2\right)\end{aligned}$$

Or $\mathbb{E}(X)$ est une (v.a.) constante : on peut donc la sortir de l'espérance :

$$\begin{aligned}\mathbb{V}(X) &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\ &= \mathbb{E}(X^2) - \mathbb{E}(X)^2\end{aligned}$$

□

C'est en général avec la formule de la Proposition ci-dessus qu'on calcule la variance d'une v.a.r.

Exemple 8. Si X est une v.a.r. constante égale à $a \in \mathbb{R}$, alors $\mathbb{V}(X) = 0$.

Voici des exemples de v.a.r avec une forte variance et une faible variance :

2.2 Variance des lois usuelles

Proposition 24.11

Si $X \sim \mathcal{B}(p)$, alors $\mathbb{V}(X) = p(1-p)$.

Démonstration. □

Proposition 24.12

Si $X \sim \mathcal{B}(n,p)$, alors $\mathbb{V}(X) = np(1-p)$.

Démonstration. Vue ultérieurement □

2.3 Propriétés de la variance

Proposition 24.13

Soit $a, b \in \mathbb{R}$ et X une v.a.r. définie sur Ω. Alors

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

Démonstration. On sait que $\mathbb{E}(aX+b) = a\mathbb{E}(X)+b$, de sorte que

$$\begin{aligned}\mathbb{V}(aX+b) &= \mathbb{E}\left((aX+b-\mathbb{E}(aX+b))^2\right)\\ &= \mathbb{E}\left((aX+b-a\mathbb{E}(X)-b)^2\right)\\ &= \mathbb{E}\left(a^2(X-\mathbb{E}(X))^2\right)\\ &= a^2\mathbb{E}\left((X-\mathbb{E}(X))^2\right)\\ &= a^2\mathbb{V}(X)\end{aligned}$$

□

Attention cependant : en général, $\mathbb{V}(X+Y) \neq \mathbb{V}(X)+\mathbb{V}(Y)$! Pour traiter la variance d'une somme, on a besoin d'une autre notion : la covariance (cf partie suivante).

2.4 Écart-type

Définition 24.14 (Écart-type)

Soit X une v.a.r. définie sur Ω. On appelle écart-type de X le réel positif :

$$\sigma(X) := \sqrt{\mathbb{V}(X)}$$

Définition 24.15

On dit qu'une v.a.r. X est centrée réduite si $\mathbb{E}(X) = 0$ et $\mathbb{V}(X) = 1$ (ou encore $\sigma(X) = 1$).

Proposition 24.16

Si X est une v.a.r. telle que $\sigma(X) > 0$, alors

$$Y = \frac{X - \mathbb{E}(X)}{\sigma(X)}$$

est une v.a.r. centrée réduite.

Démonstration.

□

3 Covariance

3.1 Définition

Définition 24.17

Soit X, Y deux v.a.r. définies sur Ω. On appelle covariance de X, Y le réel

$$\mathrm{Cov}(X,Y) := \mathbb{E}\big((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\big)$$

En d'autres termes, on considère les deux v.a.r. centrées $X - \mathbb{E}(X)$ et $Y - \mathbb{E}(Y)$ et on calcule l'espérance de leur produit.

- Lorsque $\mathrm{Cov}(X,Y) > 0$, cela signifie que ces v.a. centrées prennent, lorsque ω parcourt Ω, des valeurs de même signe : elles sont alors dites positivement corrélées. Ce phénomène est d'autant plus marqué que $\mathrm{Cov}(X,Y)$ est grand.

- À l'inverse, si $\text{Cov}(X,Y) < 0$, cela signifie que ces v.a. centrées prennent, lorsque ω parcourt Ω, des valeurs de signes opposés : elles sont alors dites négativement corrélées. Ce phénomène est d'autant plus marqué que $\text{Cov}(X,Y)$ est petit.

Par exemple, la température extérieure et la consommation de glaces sont des v.a. positivement corrélées. À l'inverse, la température extérieure et la consommation de chauffage sont des v.a. négativement corrélées.

Proposition 24.18

Soit X, Y deux v.a.r. définies sur Ω. Alors :

$$\text{Cov}(X,Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$$

Démonstration. On a

$$\begin{aligned}
\text{Cov}(X,Y) &= \mathbb{E}\left((X-\mathbb{E}(X))(Y-\mathbb{E}(Y))\right)\\
&= \mathbb{E}\left(XY - X\mathbb{E}(Y) - Y\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y)\right)\\
&= \mathbb{E}(XY) - \mathbb{E}\left(X\mathbb{E}(Y)\right) - \mathbb{E}\left(Y\mathbb{E}(X)\right) + \mathbb{E}\left(\mathbb{E}(X)\mathbb{E}(Y)\right)
\end{aligned}$$

Comme $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ sont des constantes, on peut les sortir de l'espérance :

$$\begin{aligned}
\text{Cov}(X,Y) &= \mathbb{E}(XY) - 2\mathbb{E}(X)\mathbb{E}(Y) + \mathbb{E}(X)\mathbb{E}(Y)\\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}$$

□

Exemple 9. Si $X = a$ est une v.a. constante, alors $\text{Cov}(X,Y) = 0$.

Exemple 10. Soit $p \in]0,1[$ et X une v.a. telle que $X \sim \mathcal{B}(p)$. On pose $Y = 1 - X$. Calculer $\text{Cov}(X,Y)$.

3.2 Propriétés de la covariance

Proposition 24.19

Soit X, X', Y, Y' des v.a.r. définies sur Ω.

1. $\text{Cov}(X, X) = \mathbb{V}(X) \geq 0$
2. $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
3. Cov est une forme bilinéaire : pour tout $\lambda \in \mathbb{R}$,

$$\text{Cov}(\lambda X + X', Y) = \lambda \text{Cov}(X, Y) + \text{Cov}(X', Y)$$

$$\text{Cov}(X, \lambda Y + Y') = \lambda \text{Cov}(X, Y) + \text{Cov}(X, Y')$$

Plus précisément, Cov : $\mathcal{E} \times \mathcal{E} \to \mathbb{R}$ où $\mathcal{E}$ est l'e.v. de v.a.r. définies sur $(\Omega, \mathbb{P})$.

Proposition 24.20 ("Identité remarquable")

Soit X, Y des v.a.r. définies sur Ω. Alors

$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{Cov}(X, Y)$$

Démonstration. On a

$$\begin{aligned}
\mathbb{V}(X+Y) &= \mathbb{E}\left((X+Y)^2\right) - \mathbb{E}(X+Y)^2 \\
&= \mathbb{E}\left(X^2 + Y^2 + 2XY\right) - (\mathbb{E}(X) + \mathbb{E}(Y))^2 \\
&= \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY) - \mathbb{E}(X)^2 - \mathbb{E}(Y)^2 - 2\mathbb{E}(X)\mathbb{E}(Y) \\
&= \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{Cov}(X, Y)
\end{aligned}$$

□

On peut généraliser la Proposition précédente à n variables :

Proposition 24.21

Pour toute famille $(X_1, \cdots, X_n)$ de v.a.r. définies sur Ω, on a

$$\mathbb{V}(X_1 + \cdots + X_n) = \sum_{i=1}^{n} \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

Exemple 11. Voir exemple 12 plus loin.

3.3 Variables aléatoires décorrélées

Définition 24.22

Deux v.a.r. X, Y définies sur Ω sont dites décorrélées si $\mathrm{Cov}(X, Y) = 0$.
Dans ce cas, on a
$$\mathbb{V}(X+Y) = \mathbb{V}(X) + \mathbb{V}(Y)$$

Proposition 24.23

Si X, Y sont des v.a.r. indépendantes, alors elles sont décorrélées.

Exemple 12. Soit X une v.a.r. telle que $X \sim \mathcal{B}(n,p)$. Montrer que $\mathbb{V}(X) = np(1-p)$.

Attention, la réciproque de la Proposition 24.23 est fausse.

Exemple 13. Soit X, Y deux v.a. telles que $X \sim \mathcal{U}(\{-1,0,1\})$ et $Y = 1_{\{X=0\}}$. Alors X, Y sont décorrélées, mais ne sont pas indépendantes.

4 Inégalités probabilistes

Proposition 24.24 (Inégalité de Markov)

Soit X une v.a. à valeurs dans $\mathbb{R}_+$. Alors

$$\forall a > 0 \qquad \mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

Démonstration. Soit $a > 0$. On affirme que les v.a. $1_{\{X \geq a\}}$ et $\dfrac{X}{a}$ vérifient :

$$1_{\{X \geq a\}} \leq \frac{X}{a}$$

Pour le prouver, on raisonne par disjonction de cas :

- Si $\omega \in \Omega$ est tel que $X(\omega) < a$, alors cette inégalité devient $0 \leq \dfrac{X(\omega)}{a}$, qui est vraie car X est positive.
- Si $\omega \in \Omega$ est tel que $X(\omega) \geq a$, alors cette inégalité devient $1 \leq \dfrac{X(\omega)}{a}$, qui est vraie car $X(\omega) \geq a$.

Alors, en passant à l'espérance, on a

$$\mathbb{E}\left(1_{\{X \geq a\}}\right) \leq \mathbb{E}\left(\frac{X}{a}\right)$$

donc

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

□

Proposition 24.25 (Inégalité de Bienaymé-Tchebychev)

Soit X une v.a.r. Alors

$$\forall \varepsilon > 0 \qquad \mathbb{P}\left(|X - \mathbb{E}(X)| \geq \varepsilon\right) \quad \leq \quad \frac{\mathbb{V}(X)}{\varepsilon^2}$$

En passant au complémentaire, on a donc

$$\mathbb{P}\left(|X - \mathbb{E}(X)| < \varepsilon\right) \quad \geq \quad 1 - \frac{\mathbb{V}(X)}{\varepsilon^2}$$

Démonstration. On pose la v.a. positive

$$Y := (X - \mathbb{E}(X))^2$$

ainsi que $a := \varepsilon^2 > 0$. On applique l'inégalité de Markov à Y en a : on a donc

$$\begin{aligned}
&\mathbb{P}(Y \geq a) \leq \frac{\mathbb{E}(Y)}{a} \\
\iff &\mathbb{P}\left((X - \mathbb{E}(X))^2 \geq \varepsilon^2 \right) \leq \frac{\mathbb{E}\left((X - \mathbb{E}(X))^2 \right)}{\varepsilon^2} \\
\iff &\mathbb{P}\left(\sqrt{(X - \mathbb{E}(X))^2} \geq \varepsilon \right) \leq \frac{\mathbb{V}(X)}{\varepsilon^2} \\
\iff &\mathbb{P}\left(|X - \mathbb{E}(X)| \geq \varepsilon \right) \leq \frac{\mathbb{V}(X)}{\varepsilon^2}
\end{aligned}$$

□

Remarque. L'inégalité de Bienaymé-Tchebicheff est une *inégalité de concentration* : elle fournit une majoration de la probabilité que X n'appartienne pas à l'intervalle $[\mathbb{E}(X) - \varepsilon, \mathbb{E}(X) + \varepsilon]$, autrement dit que X s'écarte à plus de ε de son espérance.

Exemple 14. Soit $X_1, \cdots, X_n$ des v.a.r. indépendantes de même loi : on suppose que leur espérance (commune) est $m \in \mathbb{R}$ et leur variance (commune) est $v \in \mathbb{R}$. On pose

$$S_n = \frac{1}{n} \sum_{k=1}^{n} X_k$$

ce qui correspond à la moyenne des valeurs prises par les X_k. Alors

$$\mathbb{E}(S_n) = \ldots$$

$$\mathbb{V}(S_n) = \ldots$$

L'inégalité de Bienaymé-Tchebychev appliquée à S_n donne :

Exemple 15. On lance n fois une pièce équilibrée. Donner une minoration de la probabilité d'obtenir entre 40% et 60% de faces.

Note : si par exemple $n = 100$, la réponse exacte est

$$\sum_{k=40}^{60} \binom{100}{k} 0.5^{100} \approx 0.965$$