Chapitre 23 : Espérance et variance

Dans tout ce chapitre (Ω, P) est un espace probabilisé fini et les variables aléatoires sont à valeurs réelles ou complexes.

I Espérance

1.1 Définition

Définition 1

Soit X une variable aléatoire définie sur Ω, on appelle espérance de X et on note $E(X)$ le réel

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

On dit que la variable aléatoire X est centrée ssi $E(X) = 0$.

Remarque :

- L'espérance est la moyenne des valeurs prises par X, chacune étant pondérée par sa probabilité. Il s'agit d'un indicateur de position.
- L'espérance d'une variable aléatoire ne dépend que de sa loi : deux variables aléatoires réelles de même loi ont même espérance.
- Si $A \in \mathcal{P}(\Omega)$, alors $E(\mathbb{1}_A) = P(A)$.

Proposition 1

Soit X une variable aléatoire sur Ω, son espérance est donnée par :

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

➪ **Exemple 1 :** On considère un dé truqué qui donne le résultat 6 avec la probabilité $\frac{1}{2}$ et tous les autres résultats de façon équiprobable. On note X la variable aléatoire donnant le résultat obtenu. Déterminer $E(X)$.

1.2 Propriétés de l'espérance

Proposition 2 : Linéarité

Soient X et Y deux variables aléatoires et $\lambda, \mu \in \mathbb{R}$ ou $\mathbb{C}$. Alors :

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

Proposition 3 : Positivité

Soit X une variable aléatoire positive. Alors :

$$E(X) \geq 0.$$

Proposition 4 : Croissance

Soient X et Y deux variables aléatoires réelles telles que $X \leq Y$. Alors :

$$E(X) \leq E(Y).$$

Proposition 5 : Inégalité triangulaire

Soit X une variable aléatoire. On a :

$$|E(X)| \leq E(|X|).$$

1.3 Cas particuliers

Proposition 6

Soient $a \in \mathbb{C}$ et X la variable constante égale à a. Alors :

$$E(X) = a$$

Proposition 7

Si X suit la loi de Bernoulli de paramètre $p \in [0,1]$, alors :

$$E(X) = p$$

Proposition 8

Si X suit la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$, on a :

$$E(X) = np$$

Proposition 9

Si X suit la loi uniforme sur $\{x_1, \dots, x_n\}$, on a :

$$E(X) = \frac{1}{n}\sum_{k=1}^{n} x_k.$$

En particulier, si X suit la loi uniforme sur $[\![1,n]\!]$, on a :

$$E(X) = \frac{n+1}{2}$$

➪ **Exemple 2 :** Une personne écrit à une autre pendant un an (365 jours) selon la règle suivante :

- le jour de l'an, il écrit de façon certaine,
- s'il a écrit le jour i, il écrira le jour suivant avec une probabilité $\frac{1}{2}$,
- s'il n'a pas écrit le jour i, il écrira le jour suivant de façon certaine.

Soit X_i la variable aléatoire valant 1 si une lettre a été écrite le jour i et 0 sinon.

1. Exprimer $P(X_{i+1} = 1)$ en fonction de $P(X_i = 1)$.
2. En déduire la loi de X_i pour $i \in [\![1,365]\!]$.
3. Soit X la variable aléatoire égale au nombre de lettres écrites dans l'année. Déterminer l'espérance de X.

1.4 Formule de transfert

Théorème 1 : Formule de transfert

Soit $X : \Omega \to E$ une variable aléatoire (non nécessairement à valeurs réelles ou complexes) et $f : E \to \mathbb{R}$ ou $\mathbb{C}$. Alors

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

Remarque :

- Pour déterminer l'espérance de $f(X)$, il n'est pas nécessaire de connaître la loi de $f(X)$, il suffit de connaître la loi de X.
- Ce résultat s'applique également aux couples de variables aléatoires :

$$E(f(X,Y)) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} f(x,y)P((X = x) \cap (Y = y)).$$

- Il se généralise également aux n-uplets de variables aléatoires.

➪ **Exemple 3 :** On considère un dé truqué qui donne le résultat 6 avec la probabilité $\frac{1}{2}$ et tous les autres résultats de façon équiprobable. On note X la variable aléatoire donnant le résultat obtenu. Déterminer $E(X^2)$.

1.5 Indépendance

Proposition 10

Soient X et Y deux variables aléatoires.
Si X et Y sont **indépendantes**, alors :
$$E(XY) = E(X)E(Y).$$

Remarque : La réciproque est fausse en général. On va voir un contre-exemple.

➪ **Exemple 4 :** Soient U et V deux variables aléatoires indépendantes de loi $\mathcal{B}(\frac{1}{2})$.
Posons : $X = 2U + V$ et $Y = \frac{1}{2}U - V$.
Montrons que $E(XY) = E(X)E(Y)$ et que X et Y ne sont pas indépendantes.

II Variance, écart type et covariance

2.1 Définitions

Définition 2

Soit X une variable aléatoire réelle. On appelle variance de X et on note $V(X)$ le réel défini par
$$V(X) = E\left((X - E(X))^2\right).$$
On appelle écart-type de X le réel noté $\sigma(X)$ défini par :
$$\sigma(X) = \sqrt{V(X)}.$$
On dit que X est réduite ssi : $V(X) = 1$.

Remarque :
- L'écart type est bien défini car $(X - E(X))^2 \geq 0$ donc par positivité de l'espérance, on a $V(X) \geq 0$.
- La variance représente la moyenne de $(X - E(X))^2$, il s'agit donc d'un indicateur de dispersion de X par rapport à sa valeur moyenne $E(X)$.
- D'après le théorème du transfert, on a :
$$V(X) = \sum_{x \in X(\Omega)} (x - E(X))^2 P(X = x).$$

2.2 Propriétés de la variance

Proposition 11 : Formule de Kœnig Huygens

Soit X une variable aléatoire réelle. On a :
$$V(X) = E(X^2) - E(X)^2.$$

➪ **Exemple 5 :** On lance simultanément 2 dés à 6 faces. On note X la variable aléatoire donnant la valeur absolue de la différence des deux numéros obtenus. Déterminer l'espérance et la variance de X.

Lemme 1

Soit $n \in \mathbb{N}$, on a :
$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6} \quad \text{et} \quad \sum_{k=1}^{n} k^3 = \left(\frac{n(n+1)}{2}\right)^2.$$

➪ **Exemple 6 :** Soit $X \sim \mathcal{U}(\llbracket 1, n \rrbracket)$. Calculer $V(X)$.

Proposition 12

Soit X une variable aléatoire réelle, soient $a, b \in \mathbb{R}$, on a :
$$V(aX + b) = a^2 V(X).$$

Corollaire 1

Soit X une variable aléatoire réelle telle que $\sigma(X) > 0$. La variable aléatoire : $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Proposition 13

Si X suit la loi de Bernoulli de paramètre $p \in [0,1]$, alors :

$$V(X) = p(1-p)$$

2.3 Covariance de deux variables aléatoires

Définition 3

Soient X et Y des variables aléatoires réelles. On appelle covariance de X et Y et on note $\operatorname{cov}(X)$ le réel défini par

$$\operatorname{cov}(X,Y) = E\left((X - E(X))(Y - E(Y))\right).$$

On dit que X et Y sont décorrélées ssi : $\operatorname{cov}(X,Y) = 0$.

Remarque : On a : $\operatorname{cov}(X,X) = V(X)$.

Proposition 14

Soient X et Y des variables aléatoires réelles. On a :

$$\operatorname{cov}(X,Y) = E(XY) - E(X)E(Y).$$

Corollaire 2

Deux variables aléatoires indépendantes sont décorrélées, autrement dit : soient X et Y des variables aléatoires indépendantes, on a :

$$\operatorname{cov}(X,Y) = 0.$$

➪ **Exemple 7 :** Soients X et Y des variables aléatoires indépendantes de loi $\mathcal{B}(p)$ avec $p \in]0,1[$. On pose : $U = X + Y$ et $V = X - Y$. Déterminer la covariance de U et V. Les variables U et V sont-elles indépendantes ?

2.4 Variance d'une somme

Proposition 15

Si X et Y sont deux variables aléatoires, alors :

$$V(X+Y) = V(X) + 2\operatorname{cov}(X,Y) + V(Y).$$

Corollaire 3

Si X et Y sont deux variables aléatoires **indépendantes**, alors :

$$V(X+Y) = V(X) + V(Y).$$

Proposition 16

Si X suit la loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$, on a :

$$V(X) = np(1-p)$$

➪ **Exemple 8 :** A un péage autoroutier n voitures franchissent au hasard et indépendamment l'une des trois barrières de péage mises à leur disposition. On note X_1 (resp. X_2, X_3) les variables aléatoires donnant le nombre de voitures ayant franchi la barrière 1 (resp. 2, 3).

1. Déterminer la loi de X_1.
2. Calculer les variances de X_1, X_2 et de $X_1 + X_2$.
3. En déduire la covariance de X_1 et X_2.

III Inégalités probabilistes

Proposition 17 : Inégalité de Markov

Soit X une variable aléatoire réelle et soit $a > 0$. Alors

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

➪ **Exemple 9 :** Soit X une variable aléatoire à valeurs dans $[\![0, n]\!]$. Montrer que, pour tout $c \in]\!]1, n[\![$:

$$\frac{E(X) - c}{n - c} \leq P(X \geq c) \leq \frac{E(X) - 1}{c - 1}.$$

Proposition 18 : Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle et soit $a > 0$. Alors

$$P(|X - E(X)| \geq a) \leq \frac{V(X)}{a^2}.$$

➪ **Exemple 10 :** On souhaite estimer l'équilibre d'une pièce. On note p la probabilité (inconnue) que la pièce tombe sur pile. On lance n fois la pièce et on note S_n le nombre de lancers ayant donné pile.
A partir de combien de lancers peut-on supposer que $\frac{S_n}{n}$ est une approximation de p à 0.01 près avec une probabilité supérieure à 95% ?

➪ **Exemple 11 :** Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires deux à deux indépendantes avec X_n suivant une loi de Bernoulli de paramètre p_n. Montrer que pour tout $\varepsilon > 0$:

$$\lim_{n\to+\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \frac{1}{n}\sum_{i=1}^{n} p_i\right| < \varepsilon\right) = 1.$$