Chapitre 5

Espérance et variance

5.1 Espérance d'une variable aléatoire

5.1.1 Définitions et premières propriétés

L'idée intuitive de l'espérance puise son origine dans les jeux de hasard. Considérons le jeu suivant : on lance un dé équilibré et on gagne 1 euro si le résultat obtenu est pair, 2 euros si le résultat est 1 ou 3, et on perd 3 euros si le résultat est 5. Est-il intéressant de jouer à ce jeu ? Quel peut-être le gain moyen ?

Définition 5.1. *Soit X une variables aléatoire réelle définie sur un espace probabilisé fini* $(\Omega, \mathbb{P})$. *On appelle* ***espérance mathématique*** *(ou simplement espérance ou encore* ***moment d'ordre 1****) de X et l'on note $\mathbb{E}(X)$, la moyenne des valeurs prises par X pondérées par la probabilité d'obtenir ces valeurs. Plus précisément, si X est une variable à valeurs dans l'ensemble fini $\{x_1, \cdots, x_n\}$, son espérance est définie par $\mathbb{E}(X) = x_1\mathbb{P}([X = x_1]) + \cdots + x_n\mathbb{P}([X = x_n])$ soit $\mathbb{E}(X) = \sum_{k=1}^{n} x_k\mathbb{P}([X = x_k])$. Autrement dit :*

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

Remarques.

- L'espérance de X est caractérisée par la seule loi de X. En particulier, deux variables aléatoires de même loi ont même espérance.
- L'espérance est un indicateur de position. Si on décale toutes les valeurs de X, l'espérance est de même décalée.
- Si $x \notin X(\Omega)$, $[X = x] = \emptyset$ en sorte que $\mathbb{E}(X) = \sum_{x\in\mathbb{R}} x\mathbb{P}([X = x])$.

Exemple. Reprenons l'exemple introductif. Soit X la variable aléatoire correspondant au nombre d'euros gagnés ou perdus. La loi de X est

k	-3	1	2
$\mathbb{P}([X = k])$	$1/6$	$1/2$	$1/3$

L'espérance de gain, notée $\mathbb{E}(X)$, est alors $\mathbb{E}(X) = -3 \times 1/6 + 1 \times 1/2 + 2 \times 1/3 = 2/3$. Le joueur gagne donc en moyenne $2/3$ d'euros.

Proposition 5.2. *Soit X une variable aléatoire réelle définie sur un espace probabilisé fini $(\Omega, \mathbb{P})$. Alors :*

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

Démonstration : Pour tout réel x de $X(\omega)$ on a $[X = x] = \bigcup_{w\in\Omega,\, X(\omega)=x} \{\omega\}$ (réunion disjointe) donc $\mathbb{P}([X = x]) = \sum_{\omega\in\Omega,\, X(\omega)=x} \mathbb{P}(\{\omega\})$ puis $\mathbb{E}(X) = \sum_{x\in X(\Omega)} x \sum_{\omega\in\Omega,\, X(\omega)=x} \mathbb{P}(\{\omega\})$. On en déduit $\mathbb{E}(X) = \sum_{x\in X(\Omega)} \sum_{\omega\in[X=x]} X(\omega)\mathbb{P}(\{\omega\})$ et donc finalement, puisque $\Omega = \bigcup_{x\in X(\Omega)} [X = x]$ (réunion disjointe), $\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})$. □

Remarque. Cette nouvelle expression de l'espérance n'est pas très utile dans la pratique. Elle l'est par contre pour établir certains résultats théoriques...

Définition 5.3. *Lorsqu'une variable X vérifie $\mathbb{E}(X) = 0$, on dit que cette variable est* ***centrée****.*

Proposition 5.4. *On considère deux variables aléatoires réelles X et Y définies sur un même espace probabilisé fini $(\Omega, \mathbb{P})$. Alors :*

1. *$\forall (a, b) \in \mathbb{R}^2 \quad \mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ (L'espérance est linéaire)*
2. *Si $X \geqslant 0$ alors $E(X) \geqslant 0$ (Positivité de l'espérance).*
   (Une variable X est dite positive si la fonction X est à valeurs dans $\mathbb{R}_+$,c'est-à-dire $\forall\omega \in \Omega,\ X(\omega) \geqslant 0$.)
3. *Si $X \leqslant Y$ alors $E(X) \leqslant E(Y)$ (L'espérance est croissante).*
   (On dit que $X \leqslant Y$, si $\forall\omega \in \Omega,\ X(\omega) \leqslant Y(\omega)$.)
4. *$|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$ (Inégalité triangulaire)*

Démonstration :

1. Soit $(a, b) \in \mathbb{R}^2$. $\mathbb{E}(aX+bY) = \sum_{\omega\in\Omega} (aX+bY)(\omega)\mathbb{P}(\{\omega\}) = \sum_{\omega\in\Omega} (aX(\omega)\mathbb{P}(\{\omega\}) + bY(\omega)\mathbb{P}(\{\omega\}))$
   donc $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
2. Pour tout x de $X(\Omega)$ on a $x \geqslant 0$ donc $x\mathbb{P}([X = x]) \geqslant 0$ et par sommation $E(X) \geqslant 0$.
3. Si $X \leqslant Y$ alors $Y - X \geqslant 0$ donc $E(Y - X) \geqslant 0$ et la linéarité donne alors $\mathbb{E}(X) \leqslant E(Y)$.
4. On a $X \leqslant |X|$ donc $\mathbb{E}(X) \leqslant \mathbb{E}(|X|)$. De même $-X \leqslant |-X| = |X|$ et donc :
   $-\mathbb{E}(X) = \mathbb{E}(-X) \leqslant \mathbb{E}(|X|)$. Finalement, $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$.

□

Remarque. Pour toute variable aléatoire X, la variable aléatoire $X - \mathbb{E}(X)$ est centrée.

5.1.2 Exemples classiques

Variables constantes

Soit $a \in \mathbb{R}$. Si $X = a$ (variable aléatoire constante égale à a) alors $\mathbb{E}(X) = \mathbb{E}(a) = a$.

Démonstration : En effet, $X(\Omega) = \{a\}$ et $\mathbb{P}([X = a]) = 1$.
On a donc $\mathbb{E}(X) = \sum_{k\in X(\Omega)} k\mathbb{P}([X = k]) = a.\mathbb{P}([X = a])$ soit $\mathbb{E}(X) = a$. □

Loi Uniforme sur un ensemble fini

On rappelle que la variable aléatoire X suit la loi uniforme (équirépartie) sur l'ensemble à n éléments $\{x_1, \cdots, x_n\}$ si $X(\Omega) = \{x_1, \cdots, x_n\}$ et si $\mathbb{P}([X = x_i]\}) = 1/n$ pour tout $i \in \{1, \cdots, n\}$. Son espérance est alors $\mathbb{E}(X) = \dfrac{1}{n}\sum_{i=1}^{n} x_i$: c'est la moyenne arithmétique des x_i !

Remarque. Si X suit la loi uniforme sur $[\![a, b]\!]$ c'est-à-dire $X \sim \mathscr{U}([\![a, b]\!])$ alors $\mathbb{E}(X) = \dfrac{a + b}{2}$.

Loi de Bernoulli, $\mathscr{B}(p)$, avec $p \in]0,1[$

Soit X une variable aléatoire telle que $X \sim \mathscr{B}(p)$. Alors $\mathbb{E}(X) = p$.

Démonstration : En effet, $X(\Omega) = \{0,1\}$ et $\mathbb{P}([X=1]) = p \quad \mathbb{P}([X=0]) = 1-p$. On a donc $\mathbb{E}(X) = \sum_{k \in X(\Omega)} k\mathbb{P}([X=k]) = 0.\mathbb{P}([X=0]) + 1.\mathbb{P}([X=1])$ soit $\mathbb{E}(X) = p$. □

Loi Binomiale, $\mathscr{B}(n,p)$, avec $p \in]0,1[$, $n \in \mathbb{N}$, $n \geqslant 1$

Soit X une variable aléatoire telle que $X \sim \mathscr{B}(n,p)$. Alors $\mathbb{E}(X) = np$.

Démonstration : On a $X(\Omega) = \{0,1,\cdots,n\}$ et $\forall k \in \{0,1,\cdots,n\}$, $\mathbb{P}([X=k]) = \binom{n}{k}p^k(1-p)^{n-k}$. On a donc $\mathbb{E}(X) = \sum_{k \in X(\Omega)} k\mathbb{P}([X=k]) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k}$ soit

$$\mathbb{E}(X) = \sum_{k=1}^{n} k.\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k} = n\sum_{k=1}^{n} \frac{(n-1)!}{(k-1)!(n-1-(k-1))!}p.p^{k-1}(1-p)^{n-1-(k-1)}$$

ou encore $\mathbb{E}(X) = np\sum_{k=0}^{n-1}\binom{n-1}{k}p^k(1-p)^{n-k} = np(p+(1-p))^{n-1}$.

Autre méthode : On peut voir une variable aléatoire de loi binomiale $\mathscr{B}(n,p)$ comme une somme de variables aléatoires (indépendantes) de même loi de Bernoulli de paramètre p :

$$X = X_1 + \cdots + X_n$$

La linéarité de l'espérance entraîne immédiatement $\mathbb{E}(X) = \mathbb{E}(X_1) + \cdots + \mathbb{E}(X_n) = np$. □

5.1.3 Espérance d'une fonction de variable aléatoire

Proposition 5.5 (Formule de transfert). *Soit X une variable aléatoire réelle sur l'espace probabilisé fini $(\Omega, \mathbb{P})$. Si f est une fonction définie sur $X(\Omega)$ et à valeurs réelles alors la variable aléatoire $f(X) = f \circ X$ a pour espérance :*

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

Démonstration : On a déjà vu que la loi de probabilité $\mathbb{P}_Y$ de $Y = f \circ X$ est donnée par

$$\forall y \in f \circ X(\Omega) \quad \mathbb{P}_Y(\{y\}) = \mathbb{P}([Y=y]) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}([X=x])$$

on a donc $\mathbb{E}(f \circ X) = \sum_{y \in f \circ X(\Omega)} y\mathbb{P}([Y=y]) = \sum_{y \in f \circ X(\Omega)} y \sum_{x \in f^{-1}(\{y\})} \mathbb{P}([X=x])$ et en conséquence $\mathbb{E}(f \circ X) = \sum_{y \in f \circ X(\Omega)} \sum_{x \in f^{-1}(\{y\})} f(x)\mathbb{P}([X=x])$. Comme $f \circ X(\Omega) = \bigcup_{y \in f(X(\Omega))} f^{-1}(\{y\})$ (réunion disjointe), on a finalement $\mathbb{E}(f(X)) = \sum_{x \in X(\Omega)} f(x)\mathbb{P}([X=x])$. □

5.1.4 Espérance et indépendance

Proposition 5.6. *Si X et Y sont deux variables aléatoires réelles indépendantes définies sur un même espace probabilisé fini $(\Omega, \mathbb{P})$, alors $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.*

Démonstration : On a $\mathbb{E}(XY) = \sum_{z\in(XY)(\Omega)} z\mathbb{P}([XY = z])$ Soit alors $z \in (XY)(\Omega)$.
$[XY = z] = \bigcup_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ xy=z}} [X = x] \cap [Y = y]$. Cette réunion étant disjointe, on en déduit :
$\mathbb{P}([XY = z]) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ xy=z}} \mathbb{P}([X = x] \cap [Y = y]) = \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ xy=z}} \mathbb{P}([X = x]\mathbb{P}([Y = y])$ (par indépendance). Mais alors $\mathbb{E}(XY) = \sum_{z\in(XY)(\Omega)} \sum_{\substack{(x,y)\in X(\Omega)\times Y(\Omega)\\ xy=z}} xy\mathbb{P}([X = x]\mathbb{P}([Y = y])$ et donc $\mathbb{E}(XY) = \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} x\mathbb{P}([X = x].y\mathbb{P}([Y = y]) = \mathbb{E}(X)\mathbb{E}(Y)$. □

Remarque. Si on peut utiliser ce résultat pour montrer que deux variables aléatoires X et Y ne sont pas indépendantes (en montrant que $\mathbb{E}(XY) \neq \mathbb{E}(X)\mathbb{E}(Y)$), il faut prendre garde au fait que la réciproque est fausse : on peut avoir $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ sans que X et Y soient indépendantes. . .

5.2 Variance

5.2.1 Définitions et premières propriétés

Définition 5.7. *Soit X une variable aléatoire réelle définie sur un espace probabilisé fini $(\Omega, \mathbb{P})$. On appelle **variance** de X (que l'on note $V(X)$) la moyenne des « carrés des écarts à la moyenne des valeurs prises par X » pondérées par la probabilité d'obtenir ces valeurs. Plus précisément, $V(X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right)$.*
*On appelle **écart-type** de X (que l'on note $\sigma(X)$) la quantité $\sigma(X) = \sqrt{V(X)}$.*

Remarques.

- La variance de X est caractérisée par la seule loi de X. En particulier, deux variables aléatoires de même loi ont même variance.
- La variance est un indicateur de dispersion. Si on décale toutes les valeurs de X, la variance est inchangée.

Définition 5.8. *Lorsqu'une variable X vérifie $\sigma(X) = 1$, on dit que cette variable est **réduite**.*

Proposition 5.9 (Formule de König-Huygens)**.** *Soit X une variable aléatoire réelle X. Alors : $V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$.*

Démonstration : $V(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2)$ donc, par linéarité de l'espérance (en tenant compte que $\mathbb{E}(X)$ est une constante donc est égale à son espérance), $V(X) = \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2$. □

Proposition 5.10 (Variance d'une combinaison linéaire)**.** *On considère une variable aléatoire réelle X et deux réels a et b. Alors : $V(aX + b) = a^2V(X)$.*

Démonstration : On applique le résultat précédent : $V(aX + b) = \mathbb{E}((aX + b)^2) - \mathbb{E}(aX + b)^2$ soit $V(X) = \mathbb{E}(a^2X^2 + 2abX + b^2) - \mathbb{E}(aX + b)^2$ et donc, par linéarité de l'espérance,
$V(aX + b) = a^2\mathbb{E}(X^2) + 2ab\mathbb{E}(X) + b^2 - \left(a^2\mathbb{E}(X)^2 + 2ab\mathbb{E}(X) + b^2\right) = a^2\left(\mathbb{E}(X^2) - \mathbb{E}(X)^2\right)$.
□

Remarque. Pour toute variable aléatoire X de variance non nulle, la variable aléatoire $\dfrac{X-\mathbb{E}(X)}{\sigma(X)}$ est centrée réduite.

Exemple. Reprenons l'exemple introductif. On complète le tableau de la loi de X :

k	-3	1	2	Somme
$\mathbb{P}([X=k])$	$1/6$	$1/2$	$1/3$	1
$k\mathbb{P}([X=k])$	$-3/6$	$1/2$	$2/3$	$\mathbb{E}(X)=2/3$
$k^2\mathbb{P}([X=k])$	$9/6$	$1/2$	$4/3$	$\mathbb{E}(X^2)=10/3$

On a donc $V(X)=\mathbb{E}(X^2)-\mathbb{E}(X)^2=10/3-4/9$ soit $V(X)=26/9$.

Remarque. Soit X une variable aléatoire réelle sur un espace probabilisé fini. Alors $V(X)=0$ si et seulement si X est (presque sûrement) constante.

5.2.2 Exemples classiques

Variables constantes

Comme on vient de le voir, pour tout variable aléatoire constante X on a $V(X)=0$.

Loi Uniforme sur un ensemble fini

On a vu que si la variable aléatoire X suit la loi uniforme sur l'ensemble à n éléments $\{x_1,\cdots,x_n\}$ alors $\mathbb{E}(X)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i$ est la moyenne arithmétique des x_i. La formule de transfert appliquée à $f:x\mapsto x^2$ donne $\mathbb{E}(X^2)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^2$ et donc $V(X)=\dfrac{1}{n}\sum\limits_{i=1}^{n}x_i^2-\mathbb{E}(X)^2$: c'est la variance de la série statistique des x_i !

Remarque. Si X suit la loi uniforme sur $[\![1,n]\!]$ c'est-à-dire $X\sim\mathscr{U}([\![1,n]\!])$ alors $V(X)=\dfrac{n^2-1}{12}$.

Démonstration : En effet, on a ici $\mathbb{E}(X)=\dfrac{n+1}{2}$ et $\mathbb{E}(X^2)=\dfrac{1}{n}\sum\limits_{i=1}^{n}i^2=\dfrac{(n+1)(2n+1)}{6}$ donc $V(X)=\dfrac{(n+1)(2n+1)}{6}-\dfrac{(n+1)^2}{4}=\dfrac{(n+1)(4n+2-3n-3)}{12}$. □

Loi de Bernoulli, $\mathscr{B}(p)$, avec $p\in]0,1[$

Soit X une variable aléatoire telle que $X\sim\mathscr{B}(p)$. Alors $V(X)=p(1-p)$.

Démonstration : On sait déjà que $\mathbb{E}(X)=p$ et la formule de transfert donne d'autre part $\mathbb{E}(X^2)=0^2\mathbb{P}((X=0])+1^2\mathbb{P}([X=1]])=p$. On a donc $V(X)=p-p^2=p(1-p)$. □

Loi Binomiale, $\mathscr{B}(n,p)$, avec $p\in]0,1[$, $n\in\mathbb{N}$, $n\geqslant 1$

Soit X une variable aléatoire telle que $X\sim\mathscr{B}(n,p)$. Alors $V(X)=np(1-p)$.

Démonstration : Pour $n=1$, c'est le résultat précédent. Soit $n\geqslant 2$. Par la formule de transfert,

$$\mathbb{E}(X(X-1))=\sum_{k=0}^{n}k(k-1).\binom{n}{k}p^k(1-p)^{n-k}=\sum_{k=2}^{n}\frac{n!}{(k-2)!(n-k)!}p^k(1-p)^{n-k}\text{ d'où}$$

$$\mathbb{E}(X(X-1)) = p^2n(n-1)\sum_{j=0}^{n-2}\frac{(n-2)!}{j!(n-2-j)!}p^j(1-p)^{n-2-j} = n(n-1)p^2\sum_{j=0}^{n-2}\binom{n-2}{j}p^j(1-p)^{n-2-j}$$

On a donc $\mathbb{E}(X(X-1)) = p^2n(n-1)(p+(1-p))^{n-2} = p^2n(n-1)$ soit (par linéarité) $\mathbb{E}(X^2) = \mathbb{E}(X) + p^2n(n-1) = np(1+pn-p)$.
Finalement, $V(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = np(1-p+pn) - n^2p^2 = np(1-p)$. □

5.2.3 Variance d'une somme et indépendance

Proposition 5.11. *Si X et Y sont deux variables aléatoires réelles indépendantes définies sur un même espace probabilisé fini $(\Omega, \mathbb{P})$, alors $V(X+Y) = V(X) + V(Y)$.*

Démonstration : D'après la formule de König-Huygens,

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

soit (par linéarité) $V(X+Y) = V(X)+V(Y)+2\left(\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)\right)$. On conclut alors par 5.6. □

Remarque. Bien évidemment, là encore la réciproque est fausse ! On peut en effet avoir $V(X+Y) = V(X) + V(Y)$ sans que X et Y soient indépendantes...

Exemple. On considère deux variables aléatoires X et Y indépendantes et de même loi de Bernoulli $\mathscr{B}(p)$ et on pose $Z = X+Y$ et $T = X-Y$. Alors on a $V(Z+T) = V(2X) = 4V(X)$ et $V(Z)+V(T) = 2V(X)+2V(X)$ (par indépendance de X et Y) soit $V(Z+T) = V(Z)+V(T)$. Pourtant Z et T ne sont pas indépendantes.
En effet, Z suit la loi binomiale de paramètres 2 et p donc $\mathbb{P}([Z=2]) = p^2$. D'autre part, $[T=1] = [X=1]\cap[Y=0]$ donc $\mathbb{P}([T=1]) = p(1-p)$. Comme $[Z=2]\cap[T=1] = \emptyset$ on a finalement $\mathbb{P}([Z=2]\cap[T=1]) \neq \mathbb{P}([Z=2])\mathbb{P}([T=1])$.

Exercice. On considère deux variables aléatoires X_1 et X_2 de lois de Bernoulli de paramètres respectifs p_1 et p_2. Montrer que X_1 et X_2 sont indépendantes si et seulement si $V(X_1+X_2) = V(X_1)+V(X_2)$.

5.3 Inégalités probabilistes

5.3.1 Inégalité de Markov

Proposition 5.12 (Inégalité de Markov)**.** *On considère une variable aléatoire réelle X définie sur un espace probabilisé fini $(\Omega, \mathbb{P})$.*
Si X est positive alors, pour tout réel $a > 0$, on a : $\mathbb{P}([X \geqslant a]) \leqslant \dfrac{\mathbb{E}(X)}{a}$.

Démonstration : Comme X est positive, $\mathbb{E}(X) = \displaystyle\sum_{x\in X(\Omega)} x\mathbb{P}([X=x]) \geqslant \sum_{x\geqslant a} x\mathbb{P}([X=x])$.

On en déduit $E(X) \geqslant \displaystyle\sum_{x\geqslant a} a\mathbb{P}([X=x]) = a\mathbb{P}([X\geqslant a])$. □

5.3.2 Inégalité de Bienaymé-Tchebychev

Proposition 5.13 (Inégalité de Bienaymé-Tchebychev)**.** *On considère une variable aléatoire réelle X définie sur un espace probabilisé fini $(\Omega, \mathbb{P})$.*
Pour tout réel $\varepsilon > 0$, on a : $\mathbb{P}([|X-\mathbb{E}(X)| \geqslant \varepsilon]) \leqslant \dfrac{V(X)}{\varepsilon^2}$.

Démonstration : La variable $(X-\mathbb{E}(X))^2$ est positive et donc, d'après l'inégalité de Markov (avec $a = \varepsilon^2$), $\mathbb{P}([(X-\mathbb{E}(X))^2 \geqslant \varepsilon^2]) \leqslant \dfrac{\mathbb{E}((X-\mathbb{E}(X))^2)}{\varepsilon^2}$ □