

Chapitre 14

Variables aléatoires discrètes

Au chapitre *Probabilités sur un univers fini*, on a vu comment modéliser une expérience aléatoire grâce à la notion d'espace probabilisé, noté (Ω, P) , où Ω désigne l'univers (fini) représentant l'ensemble des possibles et P est une probabilité sur Ω . Dans ce chapitre, on se propose d'aller plus loin dans la modélisation en étudiant le concept de variables aléatoires.

14.1 Notion de variables aléatoires

14.1.1 Variables aléatoires : formalisme probabiliste, loi et fonction de répartition

On commence naturellement ce chapitre en donnant la définition de ce qu'est une variable aléatoire.

Définition 1 (Variable aléatoire)

Soient Ω un univers fini et E un ensemble.

Une **variable aléatoire** X sur Ω est une application de Ω vers E .

Remarques :

- Si $E = \mathbb{R}$, on dit que la variable aléatoire X est *réelle*.
- Il est important d'avoir des exemples de variables aléatoires en tête afin de rendre cette notion la plus concrète possible.

Considérons par exemple $\Omega = \llbracket 1, 6 \rrbracket^2$. Soit $X : \Omega \rightarrow \llbracket 2, 12 \rrbracket$ la variable aléatoire définie par :

$$\forall \omega = (x, y) \in \llbracket 1, 6 \rrbracket^2, X(\omega) = x + y.$$

X est l'application qui, à chaque lancer de deux dés, associe la somme des points obtenus. Plus simplement, X représente le résultat de la somme du lancer de deux dés.

- Etudions un second exemple. Etant donné un univers fini Ω et $A \in \mathcal{P}(\Omega)$ un événement, on peut définir la variable aléatoire, appelée fonction indicatrice de A et notée 1_A , par : $\forall \omega \in \Omega, 1_A(\omega) = 1$ si $\omega \in A$ et $1_A(\omega) = 0$ si $\omega \notin A$. Ainsi, à tout événement A , on peut associer une variable aléatoire traduisant le fait que A se réalise ou non.

Bien qu'une variable aléatoire ne soit qu'une simple fonction, son étude diffère très largement de celle des fonctions étudiées dans un cours d'analyse par exemple. En probabilités, ce qui va nous intéresser, ce sont *les valeurs prises par une variable aléatoire* X et non les éléments en lesquels X atteint ces valeurs. Ceci justifie un formalisme propre aux probabilités et avec lequel il faut se familiariser le plus rapidement possible.

★ Formalisme probabiliste :

Soit X une variable aléatoire à valeurs dans E définie sur l'espace probabilisé fini (Ω, P) . Soit $A \subset E$. Considérons l'événement $X^{-1}(A)$, c'est-à-dire l'image réciproque de la partie A par la fonction X . Par définition, $X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\}$. Mais l'on a dit que ce qui nous intéresse, ce sont *les valeurs prises par* X , ici appartenant à A , et non les éléments en lesquels X atteint ses valeurs. On efface donc les ω dans la notation qui devient :

$$X^{-1}(A) = \{\omega \in \Omega, X(\omega) \in A\} = \{X \in A\}.$$

On lit ici la notation ensembliste à gauche et la notation probabiliste à droite.

Supposons que $E = \mathbb{R}$. On a par exemple :

$$\forall x \in \mathbb{R}, X^{-1}(\{x\}) = \{\omega \in \Omega, X(\omega) = x\} = \{X = x\} \text{ et}$$

$$\forall x \in \mathbb{R}, X^{-1}(]-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\} = \{X \leq x\}.$$

Soit X une variable aléatoire définie sur l'espace probabilisé fini (Ω, P) . On peut transférer la structure d'espace probabilisé de (Ω, P) à $X(\Omega)$ grâce à X en considérant une nouvelle probabilité, notée P_X .

Définition 2 (Loi de probabilité d'une variable aléatoire)

Soit X une variable aléatoire définie sur l'espace probabilisé fini (Ω, P) .

La loi de probabilité de X , notée P_X , est la loi image de P par X définie par :

$$\forall A \in \mathcal{P}(X(\Omega)), P_X(A) = P(X \in A).$$

P_X est une probabilité sur $X(\Omega)$.

Remarque : Précisons à nouveau le formalisme utilisé :

$$\forall A \in \mathcal{P}(X(\Omega)), P_X(A) = P(X \in A) = P(X^{-1}(A)) = P(\{\omega \in \Omega, X(\omega) \in A\}).$$

On souhaiterait à présent une formule pratique permettant de calculer $P_X(A)$ pour $A \in \mathcal{P}(X(\Omega))$.

Proposition 3 (Calcul de la probabilité d'un évènement à l'aide des événements $\{X = x\}, x \in X(\Omega)$)

Soit X une variable aléatoire définie sur l'espace probabilisé fini (Ω, P) . Alors,

$$\forall A \in \mathcal{P}(X(\Omega)), P_X(A) = \sum_{x \in A} P(X = x).$$

▷ Esquisse de preuve : Cette identité résulte de l'égalité :

$$\{X \in A\} = X^{-1}(A) = X^{-1}\left(\bigsqcup_{x \in A} \{x\}\right) = \bigsqcup_{x \in A} X^{-1}(\{x\}) = \bigsqcup_{x \in A} \{X = x\} \text{ et de l'additivité de } P.$$

Remarque : Si $A \in \mathcal{P}(E)$, on s'autorisera à écrire $P_X(A) = P(X \in A)$ qui est parfaitement bien défini. Par contre, l'égalité $P_X(A) = \sum_{x \in A} P(X = x)$ n'est pas toujours valable car A peut être infini. Cette dernière doit

être remplacée par : $P_X(A) = P(X \in A) = P(X \in A \cap X(\Omega)) + P(X \in A \cap \overline{X(\Omega)}) = P(X \in A \cap X(\Omega)) + 0 = \sum_{x \in A \cap X(\Omega)} P(X = x)$ avec $A \cap X(\Omega) \in \mathcal{P}(X(\Omega))$ qui est de cardinal fini.

On définit maintenant la notion de fonction de répartition. Néanmoins, aucune propriété de cet objet mathématique ne figure au programme.

Définition 4 (Fonction de répartition d'une variable aléatoire réelle)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) .

La fonction de répartition de X est la fonction $F_X : \mathbb{R} \rightarrow [0, 1]$ définie par :

$$\forall x \in \mathbb{R}, F_X(x) = P(X \leq x).$$

Remarques :

- D'après ce qui a été vu précédemment : $\forall x \in \mathbb{R}, F_X(x) = \sum_{k \in]-\infty, x] \cap X(\Omega)} P(X = k).$

* La fonction de répartition de X est une fonction croissante, continue à droite, qui admet une limite en tout point à gauche, de limite 0 en $-\infty$ et de limite 1 en $+\infty$. De plus, $\forall x \in \mathbb{R}, F_X(x) - F_X(x^-) = P(X = x)$.

* Deux variables aléatoires ont la même loi si et seulement si elles ont la même fonction de répartition.

14.1.2 Construction de variables aléatoires par composition de fonctions

Etant données une variable aléatoire X définie sur l'espace probabilisé fini (Ω, P) à valeurs dans un ensemble E et f une application de E vers un ensemble F , on peut considérer l'application composée $f \circ X : \Omega \rightarrow F$. Par définition, il s'agit d'une *variable aléatoire* définie sur Ω et à valeurs dans F . Par commodité, on choisit de la noter $f(X)$. On précise sa loi, qui est uniquement fonction de la loi de X (et de f), dans l'encadré qui suit.

Théorème 5 (Loi de probabilité de l'image d'une variable aléatoire par une application)

Soient E et F deux ensembles. Soient X une variable aléatoire définie sur l'espace probabilisé fini (Ω, P) puis $f : E \rightarrow F$ une application. Alors,

$$\forall y \in f(X)(\Omega), P(f(X) = y) = \sum_{x \in f^{-1}(\{y\})} P(X = x) = \sum_{x \in f^{-1}(\{y\})} P_X(\{x\}).$$

▷ Esquisse de preuve : La preuve résulte de la succession d'égalités ci-dessous combinée à l'additivité de P :

$$\{f(X) = y\} = \{\omega \in \Omega, f(X)(\omega) = y\} = \{\omega \in \Omega, X(\omega) \in f^{-1}(\{y\})\} = \bigsqcup_{x \in f^{-1}(\{y\})} \{X = x\}.$$

14.1.3 Couple de variables aléatoires : loi conjointe, lois marginales et lois conditionnelles

Soient E et F deux ensembles. On considère X et Y deux variables aléatoires définies sur l'espace probabilisé fini (Ω, P) et à valeurs dans E et F respectivement. On peut étudier l'application $(X, Y) : \Omega \rightarrow E \times F$ qui à tout $\omega \in \Omega$ associe $(X(\omega), Y(\omega))$. Il s'agit par définition d'une *variable aléatoire* à valeurs dans un produit cartésien. L'objet de cette partie est l'étude de la loi de ce couple, appelée *loi conjointe*, des lois de X et Y prises séparément, appelées *lois marginales* et de la loi d'une variable sachant l'autre, appelée *loi conditionnelle*.

Définition 6 (Loi conjointe d'un couple de variables aléatoires)

Soit (X, Y) un couple de variables aléatoires définies sur l'espace probabilisé fini (Ω, P) . La loi conjointe des variables X et Y est la loi de probabilité, notée $P_{(X, Y)}$, définie par :

$$\forall A \in \mathcal{P}((X, Y)(\Omega)), P_{(X, Y)}(A) = P((X, Y) \in A).$$

Remarques :

- Comme dans le cas d'une variable aléatoire, l'additivité de P assure que :

$$\forall A \in \mathcal{P}((X, Y)(\Omega)), P_{(X, Y)}(A) = \sum_{(x, y) \in A} P(\{X = x\} \cap \{Y = y\}).$$

Ainsi, pour déterminer la loi d'un couple de variables aléatoires (X, Y) , il suffit de préciser :

$$(X, Y)(\Omega) \quad \text{et} \quad (P(\{X = x\} \cap \{Y = y\}))_{(x, y) \in (X, Y)(\Omega)}.$$

- Si A est une partie de $(X, Y)(\Omega)$ de la forme $A = A_1 \times A_2$, alors :

$$P_{(X, Y)}(A) = P((X, Y) \in A) = P((X, Y) \in A_1 \times A_2) = P(\{X \in A_1\} \cap \{Y \in A_2\}).$$

Mais, attention! *Certaines parties de $(X, Y)(\Omega)$ peuvent ne pas s'écrire comme un produit cartésien.*

- On pourra trouver la notation suivante :

$$\forall (A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), P((X, Y) \in A \times B) = P(X \in A, Y \in B),$$

la virgule séparant les événements $\{X \in A\}$ et $\{Y \in B\}$ représentant une *intersection*.

Définition 7 (Lois marginales d'un couple de variables aléatoires)

Soit (X, Y) un couple de variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .

- La première loi marginale du couple (X, Y) est la loi de X .
- La seconde loi marginale du couple (X, Y) est la loi de Y .

Théorème 8 (Expressions des lois marginales en fonction de la loi conjointe)

Soit (X, Y) un couple de variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .

- Première marginale :

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

- Seconde marginale :

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

▷ Esquisse de preuve : Il faut savoir démontrer les égalités de ce théorème. Établissons la première.

$(\{Y = y\})_{y \in Y(\Omega)}$ est un système complet d'évènements. On peut donc écrire :

$$\{X = x\} = \{X = x\} \cap \Omega = \{X = x\} \cap \left(\bigsqcup_{y \in Y(\Omega)} \{Y = y\} \right) = \bigsqcup_{y \in Y(\Omega)} (\{X = x\} \cap \{Y = y\}).$$

On utilise alors l'additivité de P .

Remarque : La loi conjointe contient toutes les informations relatives au couple (X, Y) . Ainsi, il est naturel que cette dernière donne accès aux lois marginales. En revanche, *les lois marginales ne permettent pas en général de retrouver la loi conjointe.*

Définition 9 (Loi conditionnelle d'une variable aléatoire)

Soit (X, Y) un couple de variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .

Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) > 0$, la loi conditionnelle de X sachant $\{Y = y\}$ est la loi de X dans l'espace probabilisé fini $(\Omega, P(\cdot|Y = y))$. Pour la déterminer, il suffit de préciser :

$$X(\Omega) \quad \text{et} \quad (P(X = x|Y = y))_{x \in X(\Omega)}.$$

Remarques :

- Soit $y \in Y(\Omega)$ tel que $P(Y = y) > 0$. La connaissance de la loi conjointe de X et de Y donne accès à la loi conditionnelle de X sachant $\{Y = y\}$. Il suffit d'écrire :

$$\forall x \in X(\Omega), P(X = x|Y = y) = \frac{P((X, Y) = (x, y))}{P(Y = y)}.$$

- La connaissance de la loi de Y et de la loi conditionnelle de X sachant $\{Y = y\}$ pour tout $y \in Y(\Omega)$ tel que $P(Y = y) > 0$ donne accès à la loi conjointe de X et de Y . Soit $(x, y) \in (X, Y)(\Omega)$.

Si $P(Y = y) = 0$, alors $P((X, Y) = (x, y)) = 0$ (puisque $\{(X, Y) = (x, y)\} \subset \{Y = y\}$).

Si $P(Y = y) > 0$, alors :

$$P((X, Y) = (x, y)) = P(Y = y) \times P(X = x|Y = y).$$

- On peut définir de manière analogue la loi conditionnelle de Y sachant $\{X = x\}$ pour tout $x \in X(\Omega)$ tel que $P(X = x) > 0$ et établir des résultats similaires aux deux points précédents.

14.1.4 Variables aléatoires indépendantes

On a déjà défini la notion d'évènements indépendants dans le chapitre *Probabilités sur un univers fini*. L'objectif de cette sous-partie est d'étudier la notion de variables aléatoires indépendantes.

Définition 10 (Couple de variables aléatoires indépendantes)

Soient X et Y deux variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .
Les variables X et Y sont indépendantes si et seulement si :

$$\forall (A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega)), P(\{X \in A\} \cap \{Y \in B\}) = P(X \in A) \times P(Y \in B).$$

Remarques :

- On note $X \perp\!\!\!\perp Y$ lorsque X et Y sont indépendantes.
- On peut reformuler l'indépendance de deux variables aléatoires à l'aide de l'indépendance d'évènements. Précisément, les variables X et Y sont indépendantes si et seulement si pour tout $(A, B) \in \mathcal{P}(X(\Omega)) \times \mathcal{P}(Y(\Omega))$, les évènements $\{X \in A\}$ et $\{Y \in B\}$ sont indépendants.
- On suppose que X et Y sont à valeurs dans les ensembles E et F respectivement. On peut remplacer $A \in \mathcal{P}(X(\Omega))$ par $A \in \mathcal{P}(E)$ et $B \in \mathcal{P}(Y(\Omega))$ par $B \in \mathcal{P}(F)$. En effet, si $A \in \mathcal{P}(E)$, alors $A = (A \cap X(\Omega)) \sqcup (A \cap \overline{X(\Omega)})$ avec $A \cap X(\Omega) \in \mathcal{P}(X(\Omega))$ et $P(X \in A \cap \overline{X(\Omega)}) = 0$. On peut alors démontrer que l'égalité écrite en rouge reste valable.

On énonce à présent un critère permettant de diminuer le nombre de calculs à effectuer pour démontrer l'indépendance de deux variables aléatoires.

Théorème 11 (Critère d'indépendance de deux variables aléatoires)

Soient X et Y deux variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .
Les variables X et Y sont indépendantes si et seulement si :

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

Remarque : On peut reformuler l'indépendance de deux variables aléatoires à l'aide de leur loi conjointe. Précisément, les variables X et Y sont indépendantes si et seulement si leur loi conjointe est déterminée par la famille $(P(X = x) \times P(Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$. Dans ce cas,

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

Ainsi, en cas d'indépendance, les lois marginales permettent de retrouver la loi conjointe.

La définition de l'indépendance se généralise au cas de $n \geq 2$ variables aléatoires à travers l'indépendance mutuelle qui est une notion plus fine que l'indépendance deux à deux.

Définition 12 (Famille finie de variables aléatoires indépendantes)

Soient X_1, \dots, X_n , $n \geq 2$ variables aléatoires définies sur l'espace probabilisé fini (Ω, P) .

- X_1, \dots, X_n sont **deux à deux indépendantes** si et seulement si :

$$\forall (i, j) \in \llbracket 1, n \rrbracket^2, (i \neq j \implies X_i \text{ et } X_j \text{ sont indépendantes}).$$

- X_1, \dots, X_n sont **mutuellement indépendantes** (ou simplement indépendantes) si et seulement si :

$$\forall (A_i)_{i \in \llbracket 1, n \rrbracket} \in \prod_{i=1}^n \mathcal{P}(X_i(\Omega)), \forall I \in \mathcal{P}_f(\llbracket 1, n \rrbracket) \setminus \{\emptyset\}, P\left(\bigcap_{i \in I} \{X_i \in A_i\}\right) = \prod_{i \in I} P(X_i \in A_i).$$

Remarques :

- On constate que l'indépendance de variables aléatoires s'exprime à l'aide de l'indépendance d'évènements.
- Comme dans le cas de deux variables, on dispose d'un critère permettant de diminuer le nombre de vérifications. X_1, \dots, X_n sont indépendantes si et seulement si :

$$\forall (x_i)_{i \in [1, n]} \in \prod_{i=1}^n X_i(\Omega), \forall I \in \mathcal{P}_f([1, n]) \setminus \{\emptyset\}, P\left(\bigcap_{i \in I} \{X_i = x_i\}\right) = \prod_{i \in I} P(X_i = x_i).$$

- De même que pour les évènements, l'indépendance mutuelle de variables aléatoires implique l'indépendance deux à deux de ces variables mais la réciproque est fautive.

On termine cette partie par un énoncé assurant que l'indépendance est héritée par composition.

Théorème 13 (Lemme des coalitions)

Soient X_1, \dots, X_n , $n \geq 2$ variables aléatoires définies sur l'espace probabilisé fini (Ω, P) . Soient m un entier tel que $1 \leq m < n$ puis f et g deux applications.

Si les variables aléatoires X_1, \dots, X_n sont **indépendantes**, alors les variables aléatoires $f(X_1, \dots, X_m)$ et $g(X_{m+1}, \dots, X_n)$ sont **indépendantes**.

Remarque : On peut généraliser le résultat précédent au cas d'un nombre fini quelconque de coalitions. Par exemple, si X_1, \dots, X_7 sont des variables aléatoires indépendantes et si f , g et h sont des applications, alors les variables aléatoires $f(X_1, X_3, X_4)$, $g(X_2, X_7)$ et $h(X_5, X_6)$ sont indépendantes.

14.2 Espérance d'une variable aléatoire réelle

On étudie dans cette partie la notion d'espérance d'une variable aléatoire réelle que l'on peut interpréter comme un *indicateur de position*.

Définition 14 (Espérance d'une variable aléatoire réelle)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) .

L'espérance de la variable aléatoire X , notée $E(X)$, est définie par :

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

Remarques :

- Il faut retenir que l'espérance de X est la *moyenne pondérée des valeurs prises par X , les coefficients de pondération étant les probabilités des évènements $\{X = x\}$, $x \in X(\Omega)$.*
- Soient A un évènement et a un réel. On calcule facilement : $E(1_A) = P(A)$ et $E(a) = a$.
- On peut définir par la même formule l'espérance de variables aléatoires à valeurs complexes. Néanmoins, on ne manipule jamais de telles variables dans la pratique.
- On prendra garde au fait que l'espérance d'une variable aléatoire ne caractérise pas sa loi. Autrement dit, deux variables aléatoires ayant la même espérance n'ont pas nécessairement la même loi.

Théorème 15 (Espérance d'une variable aléatoire exprimée à l'aide des évènements élémentaires)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Alors,

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

Remarques :

- La démonstration du résultat précédent est instructive. La clé de la preuve est la décomposition :

$$\Omega = \bigsqcup_{x \in X(\Omega)} \{X = x\} = \bigsqcup_{x \in X(\Omega)} \bigsqcup_{\omega \in \{X=x\}} \{\omega\}.$$

- Cette réécriture de l'espérance permet de démontrer aisément la linéarité (voir ci-dessous) qui est un résultat important du cours. En revanche, elle *ne sert pas* en exercices et l'on se reportera à la définition de l'espérance pour effectuer des calculs.

Proposition 16 (Propriétés de l'espérance)

Soient X et Y deux variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) .

- Linéarité de l'espérance :

$$\forall (\alpha, \beta) \in \mathbb{R}^2, E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y).$$

- Positivité de l'espérance : Si X est à valeurs **positives**, alors $E(X) \geq 0$.
- Croissance de l'espérance : Si $X \leq Y$, alors $E(X) \leq E(Y)$.

Remarque : On pourrait également ajouter l'inégalité triangulaire : $|E(X)| \leq E(|X|)$.

Définition 17 (Variable aléatoire centrée)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) .

X est **centrée** si et seulement si $E(X) = 0$.

Remarque : On peut toujours construire de manière naturelle une variable aléatoire centrée à partir de X . Il suffit de considérer $X - E(X)$. Cette variable aléatoire est appelée la *variable centrée associée à X* .

Théorème 18 (Formule de transfert)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Soit $f : X(\Omega) \rightarrow \mathbb{R}$ une fonction. Alors,

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

Remarques :

- Ce théorème est le plus important de cette partie. On propose un moyen mnémotechnique permettant de le retenir. $E(f(X))$ est une somme indexée par les $x \in X(\Omega)$. Lorsque $X = x$, $f(X)$ prend la valeur $f(x)$. On écrit donc $f(x)$ que l'on pondère par la probabilité que X soit égal à x : $P(X = x)$.
- Ce résultat s'applique en particulier aux couples et aux n -uplets de variables aléatoires.
- Pour tout $k \geq 1$, $E(X^k) = \sum_{x \in X(\Omega)} x^k \times P(X = x)$ est appelé le **moment d'ordre k** de la variable aléatoire X .

Théorème 19 (Espérance d'un produit de variables aléatoires indépendantes)

Soient X et Y deux variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) .

Si X et Y sont **indépendantes**, alors $E(XY) = E(X)E(Y)$.

14.3 Covariance, variance et écart-type de variables aléatoires réelles

14.3.1 Covariance de deux variables aléatoires réelles

Définition 20 (Covariance d'un couple de variables aléatoires)

Soient X et Y deux variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) . La covariance du couple (X, Y) , notée $\text{Cov}(X, Y)$, est définie par :

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

Remarques :

- La covariance du couple (X, Y) est l'espérance du produit des variables centrées associées à X et Y .
- Deux variables aléatoires dont la covariance est nulle sont dites **décorrélées**.
- D'après la formule de transfert,

$$\text{Cov}(X, Y) = \sum_{(x,y) \in (X,Y)(\Omega)} (x - E(X))(y - E(Y))P((X, Y) = (x, y)).$$

Dans la pratique, on ne se sert jamais de cette formule pour effectuer des calculs de covariance. On lui préfère l'égalité du théorème qui suit.

Théorème 21 (Seconde expression de la covariance d'un couple de variables aléatoires)

Soient X et Y deux variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) . Alors,

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

▷ Esquisse de preuve : Il suffit de développer l'expression du produit apparaissant dans la définition de $\text{Cov}(X, Y)$ et d'utiliser la linéarité de l'espérance.

Proposition 22 (Nullité de la covariance d'un couple de variables aléatoires indépendantes)

Soient X et Y deux variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) . Si X et Y sont **indépendantes**, alors $\text{Cov}(X, Y) = 0$.

Remarques :

- Cette proposition n'est qu'une réécriture du théorème assurant que l'espérance du produit de deux variables aléatoires **indépendantes** est le produit de leurs espérances.
- On se sert souvent de la contraposée de ce résultat pour montrer que deux variables aléatoires ne sont pas indépendantes. Précisément, si $\text{Cov}(X, Y) \neq 0$, alors X et Y ne sont pas indépendantes.
- La réciproque est fautive : il existe des variables aléatoires X et Y qui ne sont pas indépendantes et qui vérifient $\text{Cov}(X, Y) = 0$. On peut par exemple considérer une variable X telle que $X(\Omega) = \{-1, 0, 1\}$ et P_X est la loi uniforme sur $X(\Omega)$ puis $Y = 1_{\{X=0\}}$.

Proposition 23 (Propriétés de la covariance)

- Linéarité à gauche : Pour toutes variables aléatoires X, Y, Z et pour tous réels a et b :

$$\text{Cov}(aX + bY, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z).$$

- Linéarité à droite : Pour toutes variables aléatoires X, Y, Z et pour tous réels a et b :

$$\text{Cov}(Z, aX + bY) = a\text{Cov}(Z, X) + b\text{Cov}(Z, Y).$$

- Symétrie : Pour toutes variables aléatoires X et Y , $\text{Cov}(X, Y) = \text{Cov}(Y, X)$.

14.3.2 Variance d'une variable aléatoire réelle

Etant donnée une variable aléatoire réelle X définie sur l'espace probabilisé fini (Ω, P) , on a déjà défini l'espérance de X , notée $E(X)$, qui représente la moyenne pondérée de ses valeurs. Néanmoins, on ne sait pas encore mesurer à quel point X fluctue autour de cette moyenne. On introduit à cette fin la notion de variance.

Définition 24 (Variance d'une variable aléatoire réelle)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) .
La variance de la variable aléatoire X , notée $V(X)$, est définie par :

$$V(X) = E((X - E(X))^2).$$

Remarques :

- La variance de X est la « moyenne du carré de l'écart à la moyenne ». En particulier, il s'agit d'un nombre positif. On retiendra que *la variance de X mesure la dispersion de X autour de sa moyenne*. « Plus $V(X)$ est proche de 0, moins X varie autour de sa moyenne ». Une variable aléatoire constante, donc égale à sa moyenne, admet une variance nulle. Réciproquement, si X est une variable aléatoire de variance nulle, alors $P(X = E(X)) = 1$. On dit que X est presque sûrement égale à sa moyenne.
- On observe que $V(X) = \text{Cov}(X, X)$.
- D'après la formule de transfert,

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

Dans la pratique, on ne se sert jamais de cette formule pour effectuer des calculs de variance. On lui préfère l'égalité du théorème qui suit.

Théorème 25 (Formule de König-Hyugens)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Alors,

$$V(X) = E(X^2) - E(X)^2.$$

Remarque : L'égalité est immédiate d'après la seconde expression de la covariance.

On énonce à présent les propriétés usuelles de la variance.

Proposition 26 (Propriétés de la variance)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) .

- $\forall a \in \mathbb{R}, V(aX) = a^2 V(X)$.
- $\forall a \in \mathbb{R}, V(X + a) = V(X)$.

Remarques :

- Au premier point, on prendra garde que les constantes multiplicatives sortent de la variance **élevées au carré**. Le second point traduit le fait que la variance ne « voit pas » les constantes additives.
- On peut résumer les deux propriétés de la manière suivante :

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

Proposition 27 (Variance d'une somme de variables aléatoires)

Soient X_1, \dots, X_n , $n \geq 2$ variables aléatoires réelles définies sur l'espace probabilisé fini (Ω, P) . Alors,

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

Remarques :

- Cette proposition résulte des propriétés de bilinéarité et de symétrie de la covariance.
- Dans le cas de deux variables X et Y , l'égalité précédente s'écrit :

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

Théorème 28 (Variance d'une somme de variables aléatoires indépendantes)

Soient X_1, \dots, X_n , $n \geq 2$ variables aléatoires réelles **deux à deux indépendantes** définies sur l'espace probabilisé fini (Ω, P) . Alors,

$$V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i).$$

Remarques :

- Ce résultat est immédiat puisque l'indépendance deux à deux implique que $\forall (i, j) \in \llbracket 1, n \rrbracket^2$, $\text{Cov}(X_i, X_j) = 0$.
- Le théorème reste vrai sous l'hypothèse plus faible où les variables aléatoires X_1, \dots, X_n sont deux à deux décorrélées.

14.3.3 Ecart-type d'une variable aléatoire réelle

On introduit dans cette sous-partie un second indicateur de la dispersion d'une variable aléatoire autour de sa moyenne, appelé écart-type, qui présente l'avantage de posséder la « même unité » que l'espérance.

Définition 29 (Ecart-type d'une variable aléatoire réelle)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . L'écart-type de la variable aléatoire X , notée $\sigma(X)$, est définie par :

$$\sigma(X) = \sqrt{V(X)}.$$

Remarque : D'après les propriétés de la variance, on a immédiatement : $\forall (a, b) \in \mathbb{R}^2$, $\sigma(aX + b) = |a|\sigma(X)$.

Définition 30 (Variable aléatoire réduite)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . X est **réduite** si et seulement si $\sigma(X) = 1$.

Remarque : Etant donnée une variable aléatoire X telle que $\sigma(X) > 0$, on peut toujours construire de manière naturelle une variable aléatoire centrée et réduite à partir de X . Il suffit de considérer $\frac{X - E(X)}{\sigma(X)}$. Cette variable aléatoire est appelée la *variable centrée réduite associée* à X .

14.4 Lois usuelles

Dans cette partie, on présente les trois lois usuelles au programme : la loi uniforme, la loi de Bernoulli et la loi binomiale. Etant donnée une variable aléatoire X suivant l'une de ces lois, on précise ce que peut modéliser X . On rappelle que la loi d'une variable aléatoire est connue dès lors qu'on donne $X(\Omega)$ et la famille des $P(X = x)$, $x \in X(\Omega)$. On explicite les espérances et variances de chacune de ces lois. Enfin, on termine cette section par l'étude d'une quatrième loi, la loi hypergéométrique, qui ne figure pas littéralement au programme mais qui apparaît régulièrement dans les exercices. Gardons en tête que si E est un ensemble fini et si $(p_e)_{e \in E}$ est une distribution de probabilité sur E , il existe un espace probabilisé fini (Ω, P) et une variable aléatoire X telle que $X(\Omega) = E$ et $\forall e \in E, P(X = e) = p_e$. On peut prendre $\Omega = E$ et définir P à partir des événements élémentaires : $\forall e \in E, P(\{e\}) = p_e$ puis poser : $\forall e \in E, X(e) = e$.

14.4.1 Etude de la loi uniforme

Définition 31 (Loi uniforme)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Soit $(a, b) \in \mathbb{Z}^2$ tel que $a \leq b$. X suit la loi uniforme sur $\llbracket a, b \rrbracket$, ce qui se note $X \sim \mathcal{U}(\llbracket a, b \rrbracket)$, si et seulement si :

- $X(\Omega) = \llbracket a, b \rrbracket$,
- $\forall k \in \llbracket a, b \rrbracket, P(X = k) = \frac{1}{b - a + 1}$.

Remarques :

- La variable \bar{X} représentant le résultat du lancer d'un dé équilibré à six faces suit la loi uniforme sur $\llbracket 1, 6 \rrbracket$.
- On observe que : $\forall k \in \llbracket a, b \rrbracket, P(X = k) = \frac{1}{b - a + 1} = \frac{1}{\text{card}(\llbracket a, b \rrbracket)}$. Ainsi, P_X est la loi uniforme sur $\llbracket a, b \rrbracket$.

Proposition 32 (Espérance et variance de la loi uniforme)

Soient $(a, b) \in \mathbb{Z}^2$ tel que $a \leq b$ puis X une variable aléatoire suivant la loi uniforme sur $\llbracket a, b \rrbracket$. Alors,

$$E(X) = \frac{a + b}{2} \quad \text{et} \quad V(X) = \frac{(b - a + 1)^2 - 1}{12}.$$

▷ Esquisse de preuve : On démontre d'abord ces égalités pour une variable aléatoire X suivant la loi uniforme sur $\llbracket 1, n \rrbracket$, $n \in \mathbb{N}^*$. $V(X)$ s'obtient par la formule de König-Huygens. On pose $n = b - a + 1$. Alors, la variable aléatoire $Y = X + a - 1$ suit la loi uniforme sur $\llbracket a, b \rrbracket$ et son espérance et sa variance se déduisent de celles de X .

14.4.2 Etude de la loi de Bernoulli

Définition 33 (Loi de Bernoulli)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Soit $p \in [0, 1]$. X suit la loi de Bernoulli de paramètre p , ce qui se note $X \sim \text{Ber}(p)$, si et seulement si :

- $X(\Omega) = \{0, 1\}$,
- $P(X = 1) = p$ et $P(X = 0) = 1 - p$.

Remarques :

- Une variable aléatoire suivant la loi de Bernoulli de paramètre p représente le résultat d'une expérience aléatoire à deux issues : le succès codé par 1 et de probabilité p et l'échec codé par 0 et de probabilité $1 - p$.
- Si (Ω, P) est un espace probabilisé fini et $A \in \mathcal{P}(\Omega)$ est un événement, alors la variable aléatoire 1_A suit la loi de Bernoulli de paramètre $p = P(A)$.

Proposition 34 (Espérance et variance de la loi de Bernoulli)

Soient $p \in [0, 1]$ puis X une variable aléatoire suivant la loi de Bernoulli de paramètre p . Alors,

$$E(X) = p \quad \text{et} \quad V(X) = p(1 - p).$$

14.4.3 Etude de la loi binomiale

Définition 35 (Loi binomiale)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Soit $(n, p) \in \mathbb{N}^* \times [0, 1]$. X suit la loi binomiale de paramètres n et p , ce qui se note $X \sim \mathcal{B}(n, p)$, si et seulement si :

- $X(\Omega) = \llbracket 0, n \rrbracket$,
- $\forall k \in \llbracket 0, n \rrbracket, P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$.

Remarques :

- Une variable aléatoire X suivant la loi binomiale de paramètres n et p représente le nombre de succès lors de la répétition de n expériences identiques et indépendantes ayant chacune la probabilité p de succès.
- L'existence de la loi binomiale résulte de la formule du binôme de Newton puisque $\sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = 1$.

Proposition 36 (Espérance et variance de la loi binomiale)

Soient $(n, p) \in \mathbb{N}^* \times [0, 1]$ puis X une variable aléatoire suivant la loi binomiale de paramètres n et p . Alors,

$$E(X) = np \quad \text{et} \quad V(X) = np(1 - p).$$

▷ Esquisse de preuve : Le calcul de l'espérance et de la variance d'une variable aléatoire suivant la loi binomiale de paramètres n et p n'est pas immédiat.

★ Pour calculer l'espérance, on utilise l'identité : $\forall n \geq 1, \forall k \in \llbracket 1, n \rrbracket, k \binom{n}{k} = n \binom{n-1}{k-1}$, appelée formule du capitaine.

★ Pour calculer la variance, on utilise la formule de König-Huygens, on écrit $E(X^2) = E(X(X-1)) + E(X)$ (artifice de calcul) et on invoque la formule du capitaine deux fois.

Remarque : Les deux méthodes de calcul précédentes seront réutilisées pour déterminer l'espérance et la variance d'une variable suivant la loi hypergéométrique (on y ajoutera seulement la formule de Vandermonde).

On s'intéresse à présent à la loi de la somme de deux variables aléatoires indépendantes suivant des lois binomiales de même paramètre $p \in [0, 1]$.

Théorème 37 (Somme de deux variables indépendantes suivant des lois binomiales de même paramètre p)

Soient $(n, m) \in (\mathbb{N}^*)^2$ et $p \in [0, 1]$. Soient X et Y deux variables aléatoires **indépendantes** telles que $X \sim \mathcal{B}(n, p)$ et $Y \sim \mathcal{B}(m, p)$. Alors,

$$X + Y \sim \mathcal{B}(n + m, p).$$

▷ Esquisse de preuve : Il faut savoir démontrer ce résultat. On mène le calcul sous-jacent de la manière suivante.

On écrit : $\forall k \in \llbracket 0, n+m \rrbracket, P(X+Y=k) = \sum_{i=0}^k P(X=i, Y=k-i)$ puis on utilise la formule de Vandermonde :

$$\forall (n, m) \in \mathbb{N}^2, \forall k \in \llbracket 0, n+m \rrbracket, \sum_{\substack{(i,j) \in \llbracket 0,n \rrbracket \times \llbracket 0,m \rrbracket \\ i+j=k}} \binom{n}{i} \binom{m}{j} = \sum_{i=0}^k \binom{n}{i} \binom{m}{k-i} = \binom{n+m}{k}.$$

Théorème 38 (Somme de n variables indépendantes suivant des lois de Bernoulli de même paramètre p)

Soit $p \in [0, 1]$. Soient X_1, \dots, X_n , $n \geq 2$ variables aléatoires **indépendantes** suivant la loi de Bernoulli de paramètre p . Alors,

$$X_1 + \dots + X_n \sim \mathcal{B}(n, p).$$

▷ Esquisse de preuve : Le résultat précédent se démontre par récurrence à l'aide du théorème sur la somme de deux variables aléatoires indépendantes suivant des lois binomiales de même paramètre p en notant que $\text{Ber}(p) = \mathcal{B}(1, p)$ et en utilisant le lemme des coalitions lors de la preuve de l'hérédité.

Remarque : Ce théorème permet de calculer très efficacement l'espérance et la variance d'une variable aléatoire X suivant la loi binomiale de paramètres n et p . En effet, X a même loi que $X_1 + \dots + X_n$ où les X_i , $1 \leq i \leq n$, sont n variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre p (qui ne vivent pas nécessairement sur le même espace probabilisé que X mais on ne détaille pas ce point).

14.4.4 Etude de la loi hypergéométrique

Définition 39 (Loi hypergéométrique)

Soit X une variable aléatoire réelle définie sur l'espace probabilisé fini (Ω, P) . Soit $(N, n, m) \in \mathbb{N}^3$ tel que $n \leq N$ et $m \leq N$.

X suit la loi hypergéométrique de paramètres N , n et m , ce qui se note $X \sim \mathcal{H}(N, n, m)$, si et seulement si :

- $X(\Omega) = \llbracket 0, \min(n, m) \rrbracket$,
- $\forall k \in \llbracket 0, \min(n, m) \rrbracket, P(X=k) = \frac{\binom{m}{k} \times \binom{N-m}{n-k}}{\binom{N}{n}}$.

Remarques :

- Voici un exemple. Une urne contient deux boules blanches et huit boules noires. Un joueur tire successivement et sans remise (ou de manière simultanée) cinq boules dans cette urne. La variable aléatoire X représentant le nombre de boules blanches tirées suit la loi hypergéométrique de paramètres $N = 10$, $n = 5$ et $m = 2$.
- L'existence de cette loi résulte de la formule de Vandermonde puisque $\sum_{k=0}^n \binom{m}{k} \binom{N-m}{n-k} = \binom{N}{n}$.

Proposition 40 (Espérance et variance de la loi hypergéométrique)

Soient $(N, n, m) \in \mathbb{N}^3$ tel que $n \leq N$ et $m \leq N$ puis X une variable aléatoire suivant la loi hypergéométrique de paramètres N , n et m . Alors,

$$E(X) = n \times \frac{m}{N} \quad \text{et} \quad V(X) = n \times \frac{m}{N} \times \left(1 - \frac{m}{N}\right) \times \frac{N-n}{N-1}.$$

Remarque : On note qu'une variable aléatoire suivant la loi hypergéométrique de paramètres N , n et m a même espérance qu'une variable aléatoire suivant la loi binomiale de paramètres n et $p = m/N$ et sa variance ne diffère que d'un facteur $(N-n)/(N-1)$.

14.5 Compléments : la loi géométrique et la loi de Poisson

Dans cette dernière section, on introduit deux nouvelles lois appelées loi géométrique et loi de Poisson permettant de modéliser des phénomènes aléatoires représentés par des grandeurs à valeurs dans \mathbb{N}^* ou \mathbb{N} . On sort donc du cadre des variables aléatoires définies sur un espace probabilisé fini. Néanmoins, toutes les notions vues précédemment s'étendent sans difficulté particulière.

Définition 41 (Variable aléatoire suivant une loi géométrique)

Soient $p \in]0, 1]$ et X une variable aléatoire.

On dit que X suit la loi géométrique de paramètre p et on note $X \sim \mathcal{G}(p)$ si et seulement si :

- $X(\Omega) = \mathbb{N}^*$,
- $\forall k \in \mathbb{N}^*, P(X = k) = (1 - p)^{k-1}p$.

\implies Comment reconnaître une variable aléatoire suivant la loi géométrique ?

On considère une succession d'épreuves de Bernoulli identiques et indépendantes de même paramètre de succès p . On note X le rang de l'épreuve qui a amené le premier succès ou encore le « temps d'attente » du premier succès. Alors X suit la loi géométrique de paramètre p .

Il est essentiel de reconnaître des situations dont la modélisation relève d'une loi géométrique.

★ 1^{er} exemple : On lance une pièce équilibrée et on note X le nombre de lancers nécessaires pour obtenir pile. Alors X suit la loi géométrique de paramètre $1/2$.

★ 2^{ième} exemple : On lance un dé équilibré et on note X le nombre de lancers nécessaires pour obtenir un 6. Alors X suit la loi géométrique de paramètre $1/6$.

Proposition 42 (Espérance et variance d'une variable aléatoire suivant une loi géométrique)

Soient $p \in]0, 1]$ puis X une variable aléatoire suivant la loi géométrique de paramètre p .

Alors, X admet une espérance et une variance données par : $E(X) = \frac{1}{p}$ et $V(X) = \frac{1-p}{p^2}$.

Définition 43 (Variable aléatoire suivant une loi de Poisson)

Soient λ un réel strictement positif et X une variable aléatoire.

On dit que X suit la loi de Poisson de paramètre λ et on note $X \sim \mathcal{P}(\lambda)$ si et seulement si :

- $X(\Omega) = \mathbb{N}$,
- $\forall k \in \mathbb{N}, P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$.

\implies Comment reconnaître une variable aléatoire suivant la loi de Poisson ?

On cherche à compter le nombre X d'événements identiques se produisant dans un intervalle de temps donné. Si l'on sait qu'il y a en moyenne λ événements se produisant de manière indépendante dans un tel intervalle de temps, alors X suit la loi de Poisson de paramètre λ .

Proposition 44 (Espérance et variance d'une variable aléatoire suivant une loi de Poisson)

Soient λ un réel strictement positif puis X une variable aléatoire suivant la loi de Poisson de paramètre λ .

Alors, X admet une espérance et une variance données par : $E(X) = \lambda$ et $V(X) = \lambda$.