Table des matières

Chapitre 26

Variables aléatoires

Dans tout ce qui suit, on se place sur un espace probabilisé fini que l'on note (Ω, P).

1 Variables aléatoires, loi d'une variable aléatoire

1.1 Loi de probabilité

Définition 1.1 (Variable aléatoire)**.**
Soit E un ensemble.
*On appelle **variable aléatoire** sur Ω toute application :*

$$X : \begin{array}{ccc} \Omega & \longrightarrow & E \\ \omega & \longmapsto & X(\omega) \end{array}.$$

L'image de Ω par l'application X est noté $X(\Omega)$.

Remarque. Lorsque $X(\Omega) \subset \mathbf{R}$, on dit que X est une variable aléatoire réelle (notée en abrégé : v.a.r.).

Exemple 1.1.
On lance deux dés à 6 faces parfaitement équilibrés et on regarde les nombres des faces du dessus.
$\Omega = \{(1;1), (1;2), \dots, (6;6)\}$ *ensemble des couples possibles.*

— *On pose X_1 la variable aléatoire réelle qui associe à tout couple le nombre du premier dé.*
$X_1 : \Omega \longrightarrow \{1;2;3;4;5;6\}$
— *On pose X_2 la variable aléatoire réelle qui associe à tout couple le nombre du second dé.*
$X_2 : \Omega \longrightarrow \{1;2;3;4;5;6\}$
— *On pose S la variable aléatoire réelle qui associe à tout couple la somme des nombres, $S = X_1 + X_2$.*
$S : \Omega \longrightarrow \{2;3;4;5;6;7;8;9;10;11;12\}$

Définition 1.2. *(Notations.)*
Avec les notations précédentes :

— *Pour tout $A \in X(\Omega)$, l'ensemble $X^{-1}(A) = \{\omega \in \Omega / X(\omega) \in A\}$ est un évènement de $\mathscr{P}(\Omega)$. On le notera $\{X \in A\}$ au lieu de $X^{-1}(A)$.*
— *Lorsque $A = \{x\}$ est un singleton, on notera $\{X = x\}$ au lieu de $X^{-1}(\{x\})$.*

Remarque.
Les évènements $\{X = x\}$, lorsque x décrit $X(\Omega)$, forment un système complet d'évènements.

Exemple 1.2.
On lance deux dés à 6 faces parfaitement équilibrés et on regarde les nombres des faces du dessus.
$\Omega = \{(1;1),(1;2),\ldots,(6;6)\}$ *ensembles des couples possibles.*

— *On pose X_1 la variable aléatoire réelle qui associe à tout couple le nombre du premier dé.*
$X_1 : \Omega \longrightarrow \{1;2;3;4;5;6\}$

— *On pose X_2 la variable aléatoire réelle qui associe à tout couple le nombre du second dé.*
$X_2 : \Omega \longrightarrow \{1;2;3;4;5;6\}$

— *On pose S la variable aléatoire réelle qui associe à tout couple la somme des nombres, $S = X_1 + X_2$.*
$S : \Omega \longrightarrow \{2;3;4;5;6;7;8;9;10;11;12\}$

— $\{X_1 \in \{2;4;6\}\} = \{(2;1),(2;2),\ldots,(6;6)\}$, *on aurait aussi pu noter cet évènement : « X_1 est pair ».*

— $\{S = 3\} = \{(2;1),(1;2)\}$

Définition 1.3 (Loi de probabilité d'une variable aléatoire)**.**
*On appelle **loi de probabilité** de la variable aléatoire X, la donnée d'un espace probabilisé (Ω, P) et d'une probabilité P_X appelée probabilité image de P par l'application X, définie par :*

$$P_X : \begin{array}{ccl} X(\Omega) & \longrightarrow & [0;1] \\ x & \longmapsto & P_X(x) = P(X = x) \end{array}.$$

Exemple 1.3.
Déterminer les lois de X_1 et de S.

Proposition 1.1.
L'application P_X est une probabilité sur l'univers fini $X(\Omega)$.

Preuve.
- On a pour tout $x \in X(\Omega)$, $P_X(x) = P(X = x) \in [0,1]$.
- $P_X(X(\Omega)) = P(X \in \Omega) = 1$ car X est à valeurs dans Ω.
- Soient A et B deux événements incompatibles, donc les événements $\{X \in A\}$ et $\{X \in B\}$ sont incompatibles et $\{X \in A \cup B\} = \{X \in A\} \cup \{X \in B\}$. On obtient alors :

$$\begin{array}{rcl} P_X(A \cup B) & = & P(\{X \in A \cup B\}) = P(\{X \in A\} \cup \{X \in B\}) \\ & = & P(\{X \in A\}) + P(\{X \in B\}) = P_X(A) + P_X(B). \end{array}$$

Ce qui établit que P_X est bien une probabilité. □

Remarque. On a en particulier $\displaystyle\sum_{x \in X(\Omega)} P(X = x) = \sum_{x \in X(\Omega)} P_X(x) = 1.$

1.2 Composée

Puisqu'une variable aléatoire est une application, nous pouvons la composer avec d'autres applications.

Proposition 1.2 (Loi de l'image d'une v.a.r.).
Soit X une variable aléatoire réelle et $f : X(\Omega) \longrightarrow \mathbf{R}$. On note $Y = f(X)$ l'application composée définie par :

$$\begin{array}{rccc} Y : & \Omega & \longrightarrow & \mathbf{R} \\ & \omega & \longmapsto & f(X(\omega)) \end{array}.$$

Pour tout $y \in Y(\Omega) = f(X(\Omega))$:

$$P_Y(y) = P(Y = y) = \sum_{x \in X(\Omega)/f(x)=y} P(X = x) = \sum_{x \in X(\Omega)/f(x)=y} P_X(x).$$

Remarque. La composée d'une application à valeurs réelles et d'une variable aléatoire est donc aussi une variable aléatoire réelle.

1.3 Fonction de répartition

Définition 1.4 (Fonction de répartition).
Soit X une variable aléatoire réelle (v.a.r.). On appelle ***fonction de répartition*** *de X l'application $F_X : \mathbf{R} \longrightarrow [0, 1]$ définie par :*

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

Proposition 1.3.
On a les propriétés suivantes :

(*i*) *F_X est croissante.*
(*ii*) $\lim_{x \to -\infty} F_X(x) = 0$.
(*iii*) $\lim_{x \to +\infty} F_X(x) = 1$.

Remarques.

— *La donnée d'une loi de probabilité d'une variable aléatoire définit entièrement la fonction de répartition.*

— *Réciproquement, la donnée d'une fonction de répartition d'une variable aléatoire définit entièrement la loi de probabilité. En effet, supposons que $X(\Omega) = \{x_1, \ldots, x_n\}$ et que les réels x_i sont classés par ordre croissant. Alors, pour tout $x_k \in X(\Omega)$:*

$$P(X = x_k) = P(X \leqslant x_k) - P(X < x_k) = P(X \leqslant x_k) - P(X \leqslant x_{k-1}) = F(x_k) - F(x_{k-1}).$$

2 Lois usuelles

2.1 Loi uniforme

Définition 2.1 (Loi uniforme).
Soit $n \in \mathbf{N}^*$. On dit qu'une variable aléatoire X suit la ***loi uniforme*** *sur $\{x_1, \ldots, x_n\}$ si :*

- $X(\Omega) = \{x_1, \ldots, x_n\}$;
- $\forall x_k \in X(\Omega), \quad P_X(x_k) = P(X = x_k) = \dfrac{1}{n}$.

On note $X \hookrightarrow \mathscr{U}(\{x_1, \ldots, x_n\})$.

Remarque. La loi uniforme intervient donc dans les cas d'équiprobabilité.

Exemple 2.1.
X_1 vu précédemment suit la loi uniforme sur $[\![1, 6]\!]$.

Exemple 2.2.
Soient X une variable aléatoire suivant la loi uniforme sur $[\![-1, 3]\!]$ et f la fonction carré. Déterminer la loi de $Y = f(X)$.

2.2 Loi de Bernoulli

Définition 2.2 (Loi de Bernoulli).
Soit $p \in [0, 1]$. On dit qu'une variable aléatoire X suit une ***loi de Bernoulli de paramètre*** *p si :*

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

On note $X \hookrightarrow \mathscr{B}(p)$.

Remarque. La loi de Bernoulli modélise les expériences avec deux issues possibles. Par exemple lancer une pièce. Il y deux issues possibles « PILE » ou « FACE ». Si l'on pose $X(\text{PILE}) = 1$ et $X(\text{FACE}) = 0$ alors X suit une loi de Bernoulli. Le paramètre est alors égal à la probabilité d'obtenir « PILE » .

2.3 Loi binomiale

Définition 2.3 (Loi binomiale).
Soient $n \in \mathbf{N}^*$ et $p \in [0, 1]$. On dit qu'une variable aléatoire X suit une ***loi binomiale de paramètres*** *n* ***et*** *p si :*

- $X(\Omega) = [\![0, n]\!]$;
- $\forall k \in X(\Omega)$, $P(X = k) = \dbinom{n}{k} p^k (1 - p)^{n-k}$.

On note $X \hookrightarrow \mathscr{B}(n, p)$.

Remarques.

- *Les lois de Bernoulli sont des cas particuliers des lois binomiales avec $n = 1$.*

— *La loi binomiale* $\mathscr{B}(n,p)$ *modélise le nombre de succès lors de la répétition de* n *expériences de Bernoulli indépendantes de paramètre* p *(par exemple le nombre de « Face » obtenues lors de* n *lancés de pièces ou nombre de tirage d'une boule blanche lors de* n *tirages d'une boule avec remise dans une urne.)*

3 Couples de variables aléatoires

On notera X, Y deux variables aléatoires sur Ω.

3.1 Loi conjointe

Définition 3.1 (Loi conjointe)**.**
Soient X *et* Y *deux variables aléatoires définies sur un même espace probabilisé* (Ω, P)*.*
On définit la ***loi conjointe*** *du couple* (X, Y) *par la donnée de :*

(i) $X(\Omega) \times Y(\Omega) = \{(x,y) \,/\, x \in X(\Omega);\, y \in Y(\Omega)\}$;

(ii) pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$ *:*

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

Remarque. Un *vecteur aléatoire discret* est un couple de variables aléatoires.

Proposition 3.1 (Formules pour un couple)**.**

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

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

$$P(X + Y = z) = \sum_{x \in X(\Omega)} P(X = x, Y = z - x) = \sum_{y \in Y(\Omega)} P(X = z - y, Y = y).$$

Ces propriétés nous permettent de définir la loi de X et celle de Y.

Définition 3.2 (Lois marginales)**.**
On appelle ***lois marginales*** *du couple* (X, Y)*, les applications définies par :*

$$P_X : X(\Omega) \to [0,1],\ x \mapsto \sum_{y \in Y(\Omega)} P(X = x, Y = y)$$

et

$$P_Y : Y(\Omega) \to [0,1],\ y \mapsto \sum_{x \in X(\Omega)} P(X = x, Y = y).$$

Remarque.
Lorsque l'on connait la loi du couple, on peut donc la loi de chacune des deux variables aléatoires. Par contre, la connaissance des lois de X et de Y ne permet pas toujours de déterminer la loi du couple (X, Y).

Exemple 3.1.

On tire simultanément 2 boules dans une urne contenant 4 boules numérotées de 1 à 4 indiscernables au toucher.

Soit Max la variable aléatoire donnant le plus grand nombre tiré, et Min celle donnant le plus petit nombre tiré.

Déterminer la loi conjointe de (Max, Min) et les lois marginales de Max et Min.

3.2 Loi conditionnelle

Définition 3.3 (Loi conditionnelle)**.**
Soient X et Y deux variables aléatoires définies sur un même espace probabilisé (Ω, P). Soit $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$. On appelle ***loi conditionnelle*** *de X sachant $(Y = y)$, l'application :*

$$\begin{array}{rccl} P(.|(Y=y)) : & X(\Omega) & \longrightarrow & [0;1] \\ & x & \longmapsto & P(X=x|(Y=y)) = P_{Y=y}(X=x). \end{array}$$

Définition 3.4 (Couple de v.a. indépendantes)**.**
Les variables aléatoires X et Y sont dites ***indépendantes*** *si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$:*

$$P(X = x, Y = y) = P(X = x)P(Y = y).$$

Proposition 3.2.
Si X et Y sont indépendantes, alors, pour toute partie $A \subset X(\Omega)$ et toute partie $B \subset Y(\Omega)$, on a :

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B).$$

Remarque. Toutes les définitions et propriétés de ce paragraphe se généralisent à n variables aléatoires sur un même espace probabilisé. On dit que n variables aléatoires réelles discrètes $X_1, \cdots, X_n$ sont **mutuellement indépendantes** si, pour tout $(x_1, \cdots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega)$, les événements $((X_1 = x_1), \cdots, (X_n = x_n))$ sont mutuellement indépendants.

Proposition 3.3 (Somme de v.a. indépendante suivant la loi de Bernoulli)**.**
Soient $n \in \mathbf{N}^*$, $p \in [0;1]$ et $X_1, \ldots, X_n$ des variables aléatoires mutuellement indépendantes suivant la loi $\mathscr{B}(p)$.
Alors la variable aléatoire $X_1 + X_2 + \cdots + X_n$ suit la loi $\mathscr{B}(n, p)$.

Exemple 3.2. *On lance n fois deux pièces A et B. La première est équilibrée ; la deuxième amène FACE avec une probabilité $p \in]0, 1[$ et PILE avec une probabilité $q = 1 - p$.*
Notons X le rang de la première apparition de Face pour la pièce A et Y celui de la 2ème pièce. Les variables X et Y sont indépendantes. Donner la loi conjointe de (X, Y).

Exemple 3.3.
On lance 18 fois de suite un dé à 6 faces parfaitement équilibré.
Quelle est la probabilité d'obtenir exactement 4 fois une face multiple de 3 ?

4 Espérance, variance et écart type

Notons X et Y deux variables aléatoires **réelles** sur Ω (univers fini).

4.1 Espérance

Définition 4.1 (Espérance)**.**
On appelle ***espérance*** *de la variable aléatoire réelle X à valeurs dans un ensemble fini $X(\Omega) = \{x_k \in \mathbf{R} \,/\, k \in [\![1, n]\!]\}$, le réel noté $\mathrm{E}(X)$, et définie par :*

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

Si $\mathrm{E}(X) = 0$, on dit que X est ***centrée****.*

Remarques.

1. *L'espérance est une moyenne des valeurs prises par X, chaque valeur étant pondérée par sa probabilité d'apparition (plus une valeur a de chance d'apparaître, plus elle compte).*
   L'espérance est un barycentre, c'est une valeur de position.
2. *On peut aussi écrire l'espérance comme :*

$$\mathrm{E}(X) = \sum_{x \in X(\Omega)} x \left(\sum_{w \in \Omega / X(w) = x} P(\{\omega\}) \right) = \sum_{w \in \Omega} X(\omega) P(\{\omega\}).$$

Proposition 4.1 (Théorème de transfert)**.**
Soit f une application à valeurs réelles définie sur $X(\Omega) = \{x_k \in \mathbf{R} \,/\, k \in [\![1, n]\!]\}$, alors $f(X)$ est d'espérance donnée par :

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

Remarques.

— *On note aussi* $\mathrm{E}(f(X)) = \displaystyle\sum_{x \in X(\Omega)} f(x) P(X = x)$.

— *L'espérance de $f(X)$ est entièrement déterminée par la loi de X.*

— *On a, en particulier, pour tout $(a, b) \in \mathbf{R}^2$,* $\mathrm{E}(aX + b) = a\,\mathrm{E}(X) + b$.

Théorème 4.1 (Espérance du produit de 2 v.a. indépendantes)**.**
Si X et Y sont ***indépendantes****, alors* $\mathrm{E}(XY) = \mathrm{E}(X) \times \mathrm{E}(Y)$

Attention

Cette formule est fausse si X et Y sont quelconques.
On n'a pas non plus de réciproque ! Deux v.a. peuvent être **corrélées** (c.à.d. $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$), sans qu'elle soit indépendantes.

Exemples 4.1.

1. *Soient $X = Y$ suivant une loi uniforme sur $\{-1; 1\}$, calculer $\mathrm{E}(XY)$ et $\mathrm{E}(X)\,\mathrm{E}(Y)$. Que peut-on en déduire ?*

2. *Soit* $\Omega = \{(1;0),(-1;0),(0;1),(0;-1)\}$ *et* X, Y *v.a.r. donnant respectivement les abscisse et ordonnée d'un point de* Ω.
   Déterminer la loi conjointe de X *et de* Y. *Calculer* $\mathrm{E}(XY)$ *et* $\mathrm{E}(X)\,\mathrm{E}(Y)$. *Que peut-on en déduire ?* X *et* Y *sont-elles indépendantes ?*

Proposition 4.2.
Pour deux v.a.r. X *et* Y, *on a :*

(*i*) *l'espérance est* ***linéaire***, *c'est-à-dire :*

$$\forall a \in \mathbf{R},\ \mathrm{E}(aX+Y) = a\,\mathrm{E}(X) + \mathrm{E}(Y);$$

(*ii*) *l'espérance est* ***positive***, *c'est-à-dire :*

$$X \geqslant 0 \implies \mathrm{E}(X) \geqslant 0;$$

(*iii*) *l'espérance est* ***croissante***, *c'est-à-dire :*

$$X \leqslant Y \implies \mathrm{E}(X) \leqslant \mathrm{E}(Y).$$

Proposition 4.3 (Inégalité de Markov)**.**
Soit X *une v.a.r positive ou nulle , alors :*

$$\forall a > 0, \qquad P(X \geqslant a) \leqslant \frac{\mathrm{E}(X)}{a}.$$

4.2 Variance et écart type

Définition 4.2 (Variance)**.**
On appelle ***variance*** *de* X *le réel :*

$$\mathrm{V}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right) = \mathrm{E}(X^2) - \mathrm{E}(X)^2.$$

Si $\mathrm{V}(X) = 1$, *on dit que* X *est* ***réduite***.

Remarque. La variance mesure la dispersion des valeurs de X autour de l'espérance. On a :

$$\mathrm{V}(X) = 0 \iff P\left(X = \mathrm{E}(X)\right) = 1.$$

Proposition 4.4.

(*i*) *Pour tout* $(a,b) \in \mathbf{R}^2$, *on a :*

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

(*ii*) $\mathrm{V}(X) = \mathrm{E}(X^2) - \mathrm{E}(X)^2$ *(Formule de Huygens).*

Remarque. On a :

$$X \text{ et } Y \text{ indépendantes} \implies \mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y).$$

Attention, la réciproque est fausse.

Théorème 4.2 (Variable aléatoire centrée réduite).
Soit X une variable aléatoire telle que $\mathrm{V}(X) \neq 0$. Alors la variable aléatoire $\dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est une centrée réduite. Elle est appelée ***variable aléatoire centrée réduite associée*** *à X.*

Théorème 4.3 (Inégalité de Bienaymé-Tchebychev).
Pour tout réel $\varepsilon > 0$, on a :

$$P\left(|X - \mathrm{E}(X)| \geqslant \varepsilon\right) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}.$$

Remarque. Cette inégalité nous indique que la probabilité que la variable aléatoire X s'éloigne de la moyenne diminue au fur et à mesure que l'on s'éloigne de cette moyenne. Cette probabilité est d'autant plus petite que ε est grand. Cette éloignement est d'un ordre inversement proportionnel au carré de la distance à la moyenne.

Exemple 4.2.

On lance un dé à 6 faces et on souhaite déterminer s'il est bien équilibré.

On commence par la face numérotée 1, de probabilité d'apparition p_1.

On lance N fois le dé et on note X_1 le nombre de 1 obtenus, $f_1 = \dfrac{X_1}{N}$ la fréquence d'apparition.

1. *Déterminer la loi de X_1. Donner ensuite son espérance et sa variance.*

2. *Déterminer N tel que f_1 soit une approximation de p_1 à 10^{-2} avec une probabilité de 95%.*

Définition 4.3 (Écart-type).
On appelle ***écart-type*** *d'une variable aléatoire réelle X le réel noté $\sigma(X)$, et définie par :*

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

Remarque.

L'intérêt de l'écart type par rapport à la variance c'est son « homogénéité » avec X, si X représente un temps, une longueur, alors l'écart type est du même type que X.

Un écart-type faible correspond à une variable aléatoire resserrée autour de sa moyenne.

4.3 Espérance et variances des lois usuelles

Théorème 4.4 (Espérance et variance des lois usuelles)**.**

(*i*) *Si X est constante, $X = m$ avec $m \in \mathbf{R}$, on dit que X suit une* ***loi certaine****. On a alors :*

$$\mathrm{E}(X) = m \quad \textit{et} \quad \mathrm{V}(X) = 0.$$

(*ii*) *Si $X \hookrightarrow \mathscr{U}(\llbracket 1, n \rrbracket)$, alors :*

$$\mathrm{E}(X) = \frac{n-1}{2} \qquad \textit{et} \qquad \mathrm{V}(X) = \frac{n^2-1}{12}.$$

(*iii*) *Si $X \hookrightarrow \mathscr{B}(p)$ avec $p \in [0,1]$, alors :*

$$\mathrm{E}(X) = p \qquad \textit{et} \qquad \mathrm{V}(X) = p(1-p).$$

(*iv*) *Si $X \hookrightarrow \mathscr{B}(n,p)$ avec $(n,p) \in \mathbf{N}^* \times [0;1]$, alors :*

$$\mathrm{E}(X) = np \qquad \textit{et} \qquad \mathrm{V}(X) = np(1-p).$$

4.4 Tableau récapitulatif

Variable	$X(\Omega)$	$P(X=k)$	$\mathrm{E}(X)$	$\mathrm{V}(X)$
X suit une loi certaine	$\{m\}$	$P(X=m)=1$	m	0
$X \hookrightarrow \mathscr{U}(\llbracket 1, n \rrbracket)$	$\llbracket 1, n \rrbracket$	$P(X=k) = \frac{1}{n}$	$\frac{n-1}{2}$	$\frac{n^2-1}{12}$
$X \hookrightarrow \mathscr{B}(p)$	$\{0,1\}$	$\begin{cases} p & \text{si} \quad k=1 \\ 1-p & \text{si} \quad k=0 \end{cases}$	p	$p(1-p)$
$X \hookrightarrow \mathscr{B}(n,p)$	$\llbracket 0, n \rrbracket$	$\binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$