

Chapitre 21 - Variables aléatoires

1 Variables aléatoires

1.1 Définition

Définition 1.1. Soit (Ω, P) un espace probabilisé fini et E un ensemble.

On appelle variable aléatoire une application définie sur (Ω, P) à valeurs dans E : $X : \begin{matrix} \Omega & \longrightarrow & E \\ \omega & \longmapsto & X(\omega) \end{matrix}$.

Lorsque E est une partie de \mathbb{R} , on parle de variable aléatoire réelle.

L'ensemble des valeurs susceptibles d'être prises par X est $X(\Omega)$ défini par $X(\Omega) = \{x \in E \mid \exists \omega \in \Omega : x = X(\omega)\}$.

Lorsque Ω est fini, $X(\Omega)$ est fini et on notera souvent $X(\Omega) = (x_i)_{i \in \llbracket 1, n \rrbracket}$ avec $n = |\Omega|$.

Remarque 1.1. Par défaut, on suppose que les variables aléatoires sont des variables aléatoires réelles (V.A.R.).

Définition 1.2. Soit une variable aléatoire $X : \Omega \longrightarrow E$. Soit A une partie de E : $A \subset E$.

On définit l'événement $(X \in A) = \{\omega \in \Omega \mid X(\omega) \in A\}$.

On a donc $(X \in A) = X^{-1}(A)$ et on le note aussi $(X \in A) = \{X \in A\}$.

Définition 1.3. Pour une variable aléatoire réelle X et pour a, b réels, on définit les événements suivants :

$(X = a) = \{\omega \in \Omega \mid X(\omega) = a\} = \{\text{tous les résultats } \omega \text{ tels que } X(\omega) = a\}$

$(X \leq a) = \{\omega \in \Omega \mid X(\omega) \leq a\} = \{\text{tous les résultats } \omega \text{ tels que } X(\omega) \leq a\}$

$(a \leq X < b) = \{\omega \in \Omega \mid a \leq X(\omega) < b\} = \{\text{tous les résultats } \omega \text{ tels que } a \leq X(\omega) < b\}$

1.2 Exemples

1.3 Loi de probabilité d'une v.a. réelle finie

Définition 1.4. Soit $X : \Omega \rightarrow E$ une variable aléatoire finie. L'application $P_X : \begin{matrix} \mathcal{P}(X(\Omega)) & \longrightarrow & [0, 1] \\ A & \longmapsto & P(X^{-1}(A)) \end{matrix}$ est une probabilité appelée loi de probabilité de la v.a. X et $(X(\Omega), P_X)$ est un espace probabilisé.

Définition 1.5. Soit X une v.a. finie $X : \Omega \longrightarrow E$. On appelle loi de probabilité de X la donnée de toutes les valeurs prises par X : $X(\Omega) = \{x_i\}_{i \in I}$ et de toutes les probabilités $(P(X = x_i))_{x_i \in X(\Omega)}$.

Remarque 1.2. On peut utiliser un tableau

x	x_1	x_2	x_3	x_4	x_5	\dots
$P(X = x)$	$P(X = x_1)$	$P(X = x_2)$	$P(X = x_3)$	$P(X = x_4)$	$P(X = x_5)$	\dots

Théorème 1.1. Soit $\{(x_i, p_i) \mid i \in I\}$ une partie finie de \mathbb{R}^2 telle que les x_i soient distincts.

$\{(x_i, p_i) \mid i \in I\}$ est la loi de probabilité d'une variable aléatoire finie si et seulement si $\begin{cases} \forall i \in I, & p_i \geq 0 \\ \sum_{i \in I} p_i = 1 \end{cases}$.

1.4 Système complet associé à une v.a. finie

Proposition 1.2. Soit $X : \Omega \longrightarrow E$ une v.a. sur un espace probabilisé fini.

$((X = x_i))_{x_i \in X(\Omega)}$ est un système complet d'événements.

Corollaire 1.3. On en déduit que $\sum_{x_i \in X(\Omega)} P(X = x_i) = 1$.

1.5 Loi d'une fonction d'une variable aléatoire finie

Définition 1.6. Soit X une v.a. finie sur (Ω, P) avec $X(\Omega) = \{x_1, x_2, \dots, x_n\}$. Soit $f : D \subset E \longrightarrow E$ une fonction telle que $X(\Omega) \subset D$.

Alors $Y = f \circ X$ est une v.a. finie sur Ω .

De plus,

$$Y(\Omega) = (f \circ X)(\Omega) = \{f(x_1), f(x_2), \dots, f(x_n)\}$$

et

$$P(Y = y_j) = \sum_{f(x_i)=y_j} P(X = x_i) \text{ (somme sur toutes les valeurs } x_i \text{ telles que } f(x_i) = y_j \text{).}$$

2 Lois usuelles

2.1 Loi uniforme

Définition 2.1. On dit que $X : \Omega \longrightarrow \mathbb{R}$ suit une loi uniforme sur $\llbracket 1, n \rrbracket$ si $\forall k \in \llbracket 1, n \rrbracket, \quad P(X = k) = \frac{1}{n}$.

On note $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$.

Définition 2.2. On dit que $X : \Omega \longrightarrow \mathbb{R}$ suit une loi uniforme sur $\{x_1, x_2, \dots, x_n\}$ si $\forall k \in \llbracket 1, n \rrbracket, \quad P(X = x_k) = \frac{1}{n}$.

On note $X \hookrightarrow \mathcal{U}(\{x_1, x_2, \dots, x_n\})$.

2.2 Loi de Bernoulli

Définition 2.3. On dit que $X : \Omega \longrightarrow \mathbb{R}$ suit une loi de Bernoulli de paramètre p avec $0 < p < 1$ si $X(\Omega) = \{0, 1\}$ et $P(X = 1) = p$ donc $P(X = 0) = 1 - p$.

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

Définition 2.4. Soit A une partie d'un ensemble Ω . On appelle fonction indicatrice de A notée χ_A la fonction $\chi_A : \Omega \longrightarrow \{0, 1\}$ définie pour $\omega \in \Omega$ par

$$\chi_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases}.$$

Remarque 2.1. Lors d'une expérience aléatoire, soit un événement A qui est réalisé (succès) ou qui ne l'est pas (échec), on peut modéliser cette situation avec une variable aléatoire X telle que l'événement $(X = 1) = A$ modélise le succès et l'événement $(X = 0) = \bar{A}$, l'échec de l'expérience. X est alors la fonction indicatrice de A .

2.3 Loi Binomiale

Définition 2.5. On dit que $X : \Omega \longrightarrow \mathbb{R}$ suit une loi binomiale de paramètre n et p avec $0 < p < 1$ et $n \in \mathbb{N}$ si $X(\Omega) = \llbracket 0, n \rrbracket$ et $\forall k \in \llbracket 0, n \rrbracket, \quad P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$

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

Remarque 2.2. On vérifie
$$\sum_{k=0}^n P(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1^n = 1.$$

Remarque 2.3. Une variable aléatoire X de loi binomiale $\mathcal{B}(n, p)$ modélise le nombre de succès lors de la répétition de n expériences aléatoires de Bernoulli indépendantes.

Remarque 2.4. Une variable aléatoire X de loi binomiale $\mathcal{B}(n, p)$ modélise le nombre d'obtention de boules rouges pour n tirages avec remise dans une urne contenant une proportion p de boules rouges.

3 Couple de variables aléatoires

3.1 Loi conjointe

Définition 3.1. Soit (Ω, P) un espace probabilisé. Soit X, Y deux v.a. sur Ω .

On appelle couple de v.a. (X, Y) l'application
$$\begin{array}{ccc} \Omega & \longrightarrow & E^2 \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array}.$$

C'est une variable aléatoire sur E^2 . On notera $(X = x_i, Y = y_j)$ l'événement $\{\omega \in \Omega | X(\omega) = x_i \text{ et } Y(\omega) = y_j\}$.

Définition 3.2. On appelle loi conjointe du couple (X, Y) la loi de la variable aléatoire (X, Y) c'est-à-dire la donnée de
— toutes les valeurs prises par le couple $(X, Y) = X(\Omega) \times Y(\Omega) = (x_i)_{i \in I} \times (y_j)_{j \in J} = \{(x_i, y_j) | x_i \in X(\Omega), y_j \in Y(\Omega)\}$

— et de toutes les probabilités

$$p_{i,j} = (P(X = x_i, Y = y_j))_{x_i \in X(\Omega), y_j \in Y(\Omega)} = (P((X = x_i) \cap (Y = y_j)))_{x_i \in X(\Omega), y_j \in Y(\Omega)}$$

ce que l'on pourra noter $((x_i, y_j), p_{i,j})_{i \in I, j \in J}$.

Théorème 3.1. Soit $\{((x_i, y_j), p_{i,j}) | i \in I, j \in J\}$ une partie finie de $\mathbb{R}^2 \times \mathbb{R}$ telle que les x_i soient distincts et que les y_j soient distincts.

$\{((x_i, y_j), p_{i,j}) | i \in I, j \in J\}$ est la loi de probabilité d'un couple de variables aléatoires réelles finies si et seulement si
$$\begin{cases} \forall i \in I, \forall j \in J, & p_{i,j} \geq 0 \\ \sum_{i \in I, j \in J} p_{i,j} = 1 \end{cases}.$$

Remarque 3.1. Par définition, comme cette somme est finie, on peut sommer d'abord sur les lignes ou d'abord sur les colonnes :

$$\sum_{i \in I, j \in J} p_{i,j} = \sum_{i \in I} \sum_{j \in J} p_{i,j} = \sum_{j \in J} \sum_{i \in I} p_{i,j} = 1$$

3.2 Lois marginales

Définition 3.3. Soit (X, Y) un couple de v.a. Les lois de X et Y s'appellent les lois marginales du couple (X, Y) .

Proposition 3.2. On a

$$\text{pour tout } x_i \in X(\Omega), \quad P(X = x_i) = \sum_{y_j \in Y(\Omega)} P(X = x_i, Y = y_j) \text{ notée } p_{i,\bullet}$$

et

$$\text{pour tout } y_j \in Y(\Omega), \quad P(Y = y_j) = \sum_{x_i \in X(\Omega)} P(X = x_i, Y = y_j) \text{ notée } p_{\bullet,j}.$$

3.3 Loi conditionnelles

Définition 3.4. Soit Y une v.a. sur (Ω, P) telle que $Y(\Omega) = \{y_j\}_{j \in J}$. Soit A un événement de $\mathcal{P}(\Omega)$ de probabilité non nulle.

La loi de Y conditionnée par A ou loi conditionnelle de Y sachant A est l'ensemble des probabilités

$$(P_A(Y = y_j))_{y_j \in Y(\Omega)} = \left(\frac{P((Y = y_j) \cap A)}{P(A)} \right)_{y_j \in Y(\Omega)}$$

Définition 3.5. Soit X et Y deux v.a. sur (Ω, P) telles que $X(\Omega) = (x_i)_{i \in I}$ et $Y(\Omega) = (y_j)_{j \in J}$.

Soit $j \in J$. La loi de X conditionnée par $(Y = y_j)$ est l'ensemble des valeurs

$$(P_{(Y=y_j)}(X = x_i))_{x_i \in X(\Omega)} = \left(\frac{P(Y = y_j \cap X = x_i)}{P(Y = y_j)} \right)_{i \in I}$$

Soit $i \in I$. La loi de Y conditionnée par $(X = x_i)$ est l'ensemble des valeurs

$$(P_{(X=x_i)}(Y = y_j))_{y_j \in Y(\Omega)} = \left(\frac{P(X = x_i \cap Y = y_j)}{P(X = x_i)} \right)_{j \in J}$$

3.4 Fonction de deux variables aléatoires

Soit X et Y deux v.a. réelles sur Ω . Comme le couple (X, Y) est une variable aléatoire réelle sur Ω , on peut définir pour une fonction $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ la variable aléatoire $Z : \Omega \rightarrow \mathbb{R}$ par $Z = g(X, Y)$.

On a $Z(\Omega) = \{g(x_i, y_j) | x_i \in X(\Omega), y_j \in Y(\Omega)\}$: ensemble des valeurs prises par Z

Les $g(x_i, y_j)$ ne sont pas nécessairement distincts.

On a $P(Z = z_k) = \sum_{\substack{(x_i, y_j) \text{ tels que} \\ g(x_i, y_j) = z_k}} P((X = x_i) \cap (Y = y_j))$

4 Variables aléatoires indépendantes

4.1 Indépendance d'un couple de variables aléatoires

Définition 4.1. Soit (X, Y) un couple de v.a. sur (Ω, P) avec $X(\Omega) = (x_i)_{i \in I}$ et $Y(\Omega) = (y_j)_{j \in J}$.

On dit que X et Y sont indépendantes pour la probabilité P si

$$P(X = x_i, Y = y_j) = P(X = x_i) \times P(Y = y_j) \text{ pour tout } x_i \in X(\Omega) \text{ et } y_j \in Y(\Omega).$$

Proposition 4.1. Si X et Y sont 2 v.a. indépendantes sur (Ω, P) alors pour toutes parties $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, on a

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

4.2 Variables aléatoires mutuellement indépendantes

Définition 4.2. Soit X_1, X_2, \dots, X_n des v.a. sur un même espace probabilisé (Ω, P) .

On dit que X_1, X_2, \dots, X_n sont mutuellement indépendantes pour la probabilité P si $\forall x_1 \in X_1(\Omega), \forall x_2 \in X_2(\Omega), \dots, \forall x_n \in X_n(\Omega)$,

$$P(X_1 = x_1 \cap X_2 = x_2 \cap \dots \cap X_n = x_n) = P(X_1 = x_1) \cdot P(X_2 = x_2) \dots P(X_n = x_n)$$

Théorème 4.2. Si X_1, X_2, \dots, X_n sont des v.a. mutuellement indépendantes sur (Ω, P) , alors

quelque soit $(A_1, A_2, \dots, A_n) \in \prod_{i=1}^n X_i(\Omega)$, les événements $(X_i \in A_i)_{i=1, \dots, n}$ sont mutuellement indépendants pour la probabilité P .

Proposition 4.3. Si X_1, X_2, \dots, X_n sont des v.a. mutuellement indépendantes sur (Ω, P) , alors elles sont indépendantes deux à deux.

4.3 Somme de v.a. suivant la loi de Bernoulli

Proposition 4.4. Soit (X_1, X_2, \dots, X_n) des v.a. mutuellement indépendantes suivant la loi de Bernoulli de même paramètre p avec $X_k(\Omega) = \{0, 1\}$. Alors la v.a. $X_1 + X_2 + \dots + X_n$ suit la loi $\mathcal{B}(n, p)$ binomiale de paramètres n et p .

Remarque 4.1. On utilise cette proposition pour modéliser n expériences identiques et indépendantes avec 2 issues (succès et échec). La variable aléatoire somme compte le nombre de succès.

4.4 Indépendance de fonctions de v.a. indépendantes

Théorème 4.5. Soit X, Y deux v.a. sur (Ω, P) fini. Soit f, g deux fonctions définies respectivement sur $X(\Omega)$ et $Y(\Omega)$.

Si X et Y sont indépendantes, alors $f(X)$ et $g(Y)$ sont indépendantes.

5 Moments d'une v.a. réelle finie

5.1 Espérance

Définition 5.1. Soit X une v.a.r. sur Ω . On appelle espérance de X le réel $E(X) = \sum_{x \in X(\Omega)} xP(X = x)$.

Ce qui s'écrit également $E(X) = \sum_{i \in I} x_i P(X = x_i)$ avec $X(\Omega) = \{x_i | i \in I\}$.

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

Proposition 5.1. Si pour tout $x \in X(\Omega)$, on a $a \leq x \leq b$, alors $a \leq E(x) \leq b$.

5.2 Propriétés de l'espérance : linéarité et croissance

Proposition 5.2. Soit X, Y deux v.a.r. sur (Ω, P) et a, b réels. On a

$$E(aX + bY) = aE(X) + bE(Y)$$

Proposition 5.3. Soit X, Y deux v.a.r. sur Ω .

Si on a $X \leq Y$, c'est à dire $\forall \omega \in \Omega, X(\omega) \leq Y(\omega)$, alors $E(X) \leq E(Y)$.

5.3 Théorème de transfert

Théorème 5.4. Soit g une fonction définie sur $X(\Omega)$ et X une v.a.r. sur (Ω, P) finie.

$$\text{On a } E(g(X)) = \sum_{x \in X(\Omega)} g(x)P(X = x) = \sum_{i \in I} g(x_i)P(X = x_i).$$

5.4 Espérance et indépendance

Théorème 5.5. Soit X et Y deux v.a.r. sur Ω .

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

5.5 Variance et écart-type

Définition 5.2. Soit X une v.a.r. finie. On appelle variance de X le réel $V(X) = E((X - E(X))^2)$ et écart-type de X le réel $\sigma(X) = \sqrt{V(X)}$.

$$\text{On a donc } V(X) = \sum_{x_i \in X(\Omega)} (x_i - E(X))^2 P(X = x_i).$$

Remarque 5.1. On a $V(X) \geq 0$ donc $\sigma(X)$ existe.

Théorème 5.6 (Formule de Kœnig-Huygens).

On a pour une v.a.r. finie X :

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

Théorème 5.7. Soit X une v.a. finie. On a pour tous réels a, b

$$V(aX + b) = a^2 V(X).$$

5.6 Espérance et variance des lois usuelles finies

— Si X est une v.a. finie constante, $X = c$ avec $P(X = c) = 1$, alors

$$E(X) = c \text{ et } V(X) = 0.$$

— Si X suit la loi uniforme sur $\llbracket 1, n \rrbracket : \mathcal{U}(\llbracket 1, n \rrbracket)$ alors

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

— Si X suit la loi uniforme sur $\Omega : X(\omega) = \{x_1, x_2, \dots, x_n\}$ et $P(X = x_i) = \frac{1}{n}$, alors

$$E(X) = \frac{1}{n} \sum_{i=1}^n x_i \text{ et } V(X) = \frac{1}{n} \sum_{i=1}^n x_i^2 - \left(\frac{1}{n} \sum_{i=1}^n x_i \right)^2.$$

— Si X suit une loi de Bernoulli de paramètre $p : \mathcal{B}(p)$, alors

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

— Si X suit une loi binomiale de paramètres n et $p : \mathcal{B}(n, p)$, alors

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

Démonstration.

— Soit X une variable aléatoire réelle suivant une loi uniforme : $X \hookrightarrow \mathcal{U}(\llbracket 1, n \rrbracket)$.

$$\text{Alors } E(X) = \sum_{k=1}^n k \times \frac{1}{n} = \frac{n(n+1)}{2} \times \frac{1}{n} \implies \boxed{E(X) = \frac{n+1}{2}}.$$

On calcule également $E(X^2) = \sum_{k=1}^n k^2 \times \frac{1}{n} = \frac{n(n+1)(2n+1)}{6} \times \frac{1}{n} = \frac{(n+1)(2n+1)}{6}$

où on a utilisé $\sum_{k=0}^n k^2 = \frac{n(n+1)(2n+1)}{6}$.

$$\text{Et } V(X) = E(X^2) - E(X)^2 = \frac{(n+1)(2n+1)}{6} - \frac{(n+1)^2}{4} = \frac{2(n+1)(2n+1) - 3(n+1)^2}{12}.$$

On trouve

$$V(X) = \frac{4n^2 + 6n + 2 - 3n^2 - 6n - 3}{12} \Rightarrow \boxed{V(X) = \frac{n^2 - 1}{12}}.$$

— Soit X une variable aléatoire réelle suivant une loi de Bernoulli : $X \hookrightarrow \mathcal{B}(p)$.

X prend les valeurs 0 avec la probabilité $1 - p$ et la valeur 1 avec p .

Alors $E(X) = 0 \times (1 - p) + 1 \times p$ soit $\boxed{E(X) = p}$. On calcule également $E(X^2) = 0^2 \times (1 - p) + 1^2 \times p = p$.

Et $V(X) = E(X^2) - E(X)^2 = p - p^2$ soit $\boxed{V(X) = p(1 - p)}$.

— Soit X une variable aléatoire réelle suivant une loi binomiale : $X \hookrightarrow \mathcal{B}(n, p)$.

Alors $E(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1 - p)^{n-k}$. Mais on sait que $k \binom{n}{k} = n \binom{n-1}{k-1}$ pour $k \geq 1$.

On obtient :

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

On reconnaît une formule du binôme :

$$E(X) = np(p + 1 - p)^{n-1} \text{ soit } \boxed{E(X) = np}. \text{ (résultat facile à obtenir par linéarité)}$$

On calcule maintenant $E(X(X - 1))$ qui donnera $E(X^2) - E(X)$:

$$E(X(X - 1)) = \sum_{k=0}^n k(k - 1) \binom{n}{k} p^k (1 - p)^{n-k}. \text{ Mais on sait que } k(k - 1) \binom{n}{k} = n(n - 1) \binom{n-2}{k-2} \text{ pour } k \geq 2.$$

On obtient :

$$E(X(X - 1)) = 0 + 0 + \sum_{k=2}^n n(n - 1) \binom{n-2}{k-2} p^k (1 - p)^{n-k} = n(n - 1)p^2 \sum_{k=2}^n \binom{n-2}{k-2} p^{k-2} (1 - p)^{(n-2)-(k-2)}.$$

On reconnaît une formule du binôme :

$$E(X(X - 1)) = n(n - 1)p^2(p + 1 - p)^{n-2} = n(n - 1)p^2.$$

Et,

$$V(X) = E(X^2) - E(X)^2 = E(X(X - 1)) + E(X) - E(X)^2 = n(n - 1)p^2 + np - n^2p^2 = np((n - 1)p + 1 - np).$$

D'où $\boxed{V(X) = np(1 - p)}$.

□

5.7 Inégalité de Bienaymé-Tchebychev

Théorème 5.8. Si X est une v.a. définie sur un espace probabilisé (Ω, P) d'espérance $E(X)$ et de variance $V(X)$, alors

$$\forall \varepsilon > 0, \quad P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}.$$