

Variables Aléatoires Réelles Discrètes

Khalid El Amine I.
Department of Mathematics



Dans tout ce chapitre :

- Ω est un ensemble non vide.
- (Ω, \mathcal{A}) est un espace probablisable.
- (Ω, \mathcal{A}, P) est un espace probablisé.
- Toutes les variables aléatoires sont définies sur un espace probablisé (Ω, \mathcal{A}, P) et à valeurs dans l'espace probablisable $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

On considère dans ce chapitre les v.a.r. X telles que $X(\Omega)$ est dénombrable.

1 Variable aléatoire réelle discrète

1.1 Définition et généralités

Définition 1.1 (v.a.r. discrète)

Une variable aléatoire réelle X sur (Ω, \mathcal{A}) est dite *discrète* si $X(\Omega)$ est dénombrable.

- Si $X(\Omega)$ est fini, X est appelée *v.a.r. discrète finie*.
- Si $X(\Omega)$ est infini dénombrable, X est appelée *v.a.r. discrète infinie*.

Théorème 1.2 (Caractérisation importante)

Soit $X : \Omega \rightarrow \mathbb{R}$ une application. On a :

$$X \text{ est une v.a.r.d. sur } (\Omega, \mathcal{A}) \iff \begin{cases} 1) X(\Omega) \text{ est dénombrable} \\ 2) \forall x \in X(\Omega), X^{-1}(\{x\}) \in \mathcal{A} \end{cases}$$

Notation : Pour simplifier, on notera dans la suite :

- $\{X = x\} = X^{-1}(\{x\}) = \{\omega \in \Omega : X(\omega) = x\}$
- $P(X = x) = P(\{X = x\})$
- $P(X = x, Y = y) = P(\{X = x, Y = y\}) = P(\{X = x\} \cap \{Y = y\})$
- $P(X_1 = x_1, \dots, X_n = x_n) = P(\{X_1 = x_1, \dots, X_n = x_n\}) = P(\{X_1 = x_1\} \cap \dots \cap \{X_n = x_n\})$

Théorème 1.3 (Système complet d'événements associé à une v.a.r. discrète)

Si X est une v.a.r. discrète sur (Ω, \mathcal{A}) , alors la famille

$$(\{X = x\})_{x \in X(\Omega)}$$

est un système complet d'événements ; appelé système complet d'événements associé à X .

1.2 Loi de probabilité d'une v.a.r. discrète

Théorème 1.4 Soit X une v.a.r. discrète sur (Ω, \mathcal{A}, P) . La loi de probabilité de X , notée P_X , est une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ et elle est donnée par :

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad P_X(B) = \sum_{x \in B \cap X(\Omega)} P(X = x)$$

- En particulier

$$\sum_{x \in X(\Omega)} P(X = x) = 1$$

En pratique : Pour déterminer la loi de probabilité d'une v.a.r.d. X sur (Ω, \mathcal{A}, P) , il faut déterminer :

$$\begin{cases} 1) X(\Omega) \\ 2) P(X = x) \text{ pour tout } x \in X(\Omega) \end{cases}$$

- La loi de probabilité d'une v.a.r.d. X est parfois donnée par l'ensemble :

$$\{(x, P(X = x)) : x \in X(\Omega)\}$$

ou par la suite réelle :

$$(P(X = x))_{x \in X(\Omega)}$$

Exemple : Jets de deux dés

On jette simultanément deux dés (à six faces) équilibrés discernables. On s'intéresse à la somme des résultats des deux dés.

Modélisation :

- $\Omega = \{1, 2, \dots, 6\}^2 = \{(i, j) \in \mathbb{N}^2 : 1 \leq i, j \leq 6\}$
- $\mathcal{A} = \mathcal{P}(\Omega)$
- P est la probabilité uniforme sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$, i.e. :

$$P(A) = \frac{\text{card}(A)}{36}, \quad \forall A \in \mathcal{A}$$

- $X : \Omega \rightarrow \mathbb{R}$ est définie par $X(i, j) = i + j$.

- X est une v.a.r. car $\mathcal{A} = \mathcal{P}(\Omega)$;
- X est discrète finie car $X(\Omega)$ est fini ;

$$X(\Omega) = \{2, 3, \dots, 12\}$$

- Loi de X : Pour déterminer complètement P_X , la loi de X , on calcule $P(X = k)$. Notons $A_k = \{X = k\} = X^{-1}(\{k\}) = \{(i, j) \in \Omega : i + j = k\}$. On a

$$P(X = k) = P(A_k) = \frac{\text{card}(A_k)}{36}$$

Après calcul, on trouve :

k	2	3	4	5	6	7	8	9	10	11	12
$P(X = k)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

1.3 Fonction de masse

Définition 1.5 (Fonction de masse)

Soit X une v.a.r.d. sur (Ω, \mathcal{A}, P) . On appelle *fonction de masse de X* , la fonction :

$$\begin{aligned} p : \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto p(x) = P(X = x) \end{aligned}$$

Remarques :

- On identifie parfois une fonction de masse de X à sa loi de probabilité.
- p ; la fonction de masse de la v.a.r.d. X vérifie les propriétés suivantes :

1. $\forall x \in \mathbb{R}, \quad p(x) \geq 0$
2. $S_X = \{x \in \mathbb{R} : p(x) > 0\}$ est dénombrable
3. $\sum_{x \in S_X} p(x) = 1$

1.4 Fonction de répartition d'une v.a.r. discrète

Théorème 1.6 Soit X une v.a.r. discrète sur (Ω, \mathcal{A}, P) . La fonction de répartition de X , notée F_X est définie pour tout $t \in \mathbb{R}$ par :

$$\begin{aligned} F_X(t) &= P(X \leq t) \\ &= \sum_{x \in]-\infty, t] \cap X(\Omega)} P(X = x) \\ &= \sum_{\substack{x \in X(\Omega) \\ x \leq t}} P(X = x) \end{aligned}$$

$F_X(t)$ représente la probabilité de toutes les réalisations inférieures ou égales au réel t .

Proposition 1.7 Soit X une v.a.r.d. sur (Ω, \mathcal{A}, P) et F_X sa fonction de répartition. Alors :

1. F_X est croissante sur \mathbb{R} .
2. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$
3. F_X est en escalier sur \mathbb{R} .
4. L'ensemble des points de discontinuité de F_X est $X(\Omega)$ et on a :

$$\forall x \in X(\Omega) : \quad F_X(x^+) = F_X(x) \quad \text{et} \quad F_X(x^+) - F_X(x^-) = P(X = x)$$

Remarques :

- $\forall x \in \mathbb{R}, \quad F_X(x) \in [0, 1]$.
- F_X est continue sur $\mathbb{R} \setminus X(\Omega)$.
- F_X est continue à droite et admet une limite à gauche en tout point de $X(\Omega)$.

Proposition 1.8 Soit X et Y deux v.a.r.d. sur le même (Ω, \mathcal{A}, P) . On a :

$$F_X = F_Y \text{ si et seulement si } P_X = P_Y$$

Remarque : La fonction de répartition F_X de la v.a.r.d. X caractérise complètement sa loi.

1.5 Opérations sur les v.a.r. discrètes

Théorème 1.9 (Opérations algébriques)

Soit X et Y deux v.a.r. discrètes sur (Ω, \mathcal{A}) et $\alpha \in \mathbb{R}$.

- Les applications

$$\alpha X \quad ; \quad X + Y \quad ; \quad XY \quad ; \quad \min(X, Y) \quad ; \quad \max(X, Y)$$

définies, pour tout $\omega \in \Omega$, par :

$$(\alpha X)(\omega) = \alpha X(\omega) \quad ; \quad (X + Y)(\omega) = X(\omega) + Y(\omega) \quad ; \quad (XY)(\omega) = X(\omega)Y(\omega) \quad ;$$

$$(\min(X, Y))(\omega) = \min(X(\omega), Y(\omega)) \quad ; \quad (\max(X, Y))(\omega) = \max(X(\omega), Y(\omega))$$

sont des v.a.r. discrètes sur (Ω, \mathcal{A}) .

- Si $\forall \omega \in \Omega$, $Y(\omega) \neq 0$, alors $\frac{X}{Y}$, définie par $\left(\frac{X}{Y}\right)(\omega) = \frac{X(\omega)}{Y(\omega)}$, est une v.a.r. discrète sur (Ω, \mathcal{A}) .

Théorème 1.10 (Composition)

Soit X une v.a.r. discrète sur (Ω, \mathcal{A}) et $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une application quelconque.

- L'application composée :

$$\varphi(X) = \varphi \circ X$$

définie pour tout $\omega \in \Omega$, par $\varphi(X)(\omega) = \varphi(X(\omega))$ est une v.a.r. discrète sur (Ω, \mathcal{A}) .

- Sa loi est définie par :

$$\left\{ \begin{array}{l} 1) \varphi(X(\Omega)) = \{\varphi(x) : x \in X(\Omega)\} \\ 2) P(\varphi(X) = y) = \sum_{\substack{x \in X(\Omega) \\ \varphi(x) = y}} P(X = x), \quad \forall y \in \varphi(X(\Omega)) \end{array} \right.$$

Exemple : Une urne contient une infinité de boules numérotées sur \mathbb{Z} . Un joueur prélève une boule, de façon que la probabilité que la boule numéro n soit prélevée est $p_n = \frac{1}{2 \times 3^{|n|}}$. La suite $(p_n)_{n \in \mathbb{Z}}$ définit une probabilité sur $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$. En effet :

- $p_n \geq 0$, $\forall n \in \mathbb{Z}$
- Comme $p_{-n} = \frac{1}{2 \times 3^{|-n|}} = \frac{1}{2 \times 3^{|n|}} = p_n$, on a :

$$\sum_{n \in \mathbb{Z}} p_n = 2 \sum_{n \in \mathbb{N}} p_n - p_0 = 2 \sum_{n \in \mathbb{N}} \frac{1}{2 \times 3^n} - \frac{1}{2 \times 3^0} = \sum_{n=0}^{+\infty} \frac{1}{3^n} - \frac{1}{2} = \frac{1}{1 - \frac{1}{3}} - \frac{1}{2} = 1$$

Notons $X = \ll \text{numéro de la boule prélevée} \gg$. X est une v.a.r.d sur $(\mathbb{Z}, \mathcal{P}(\mathbb{Z}))$ dont la loi est donnée par

$$\left\{ \begin{array}{l} 1) X(\Omega) = \mathbb{Z} \\ 2) P(X = k) = \frac{1}{2 \times 3^{|k|}}, \quad \forall k \in \mathbb{Z} \end{array} \right.$$

En prenant $Y = X^2$ (ici $Y = \varphi(X)$ avec $\varphi : \mathbb{R} \rightarrow \mathbb{R}$; $x \mapsto x^2$).

Y est une v.a.r.d. dont l'univers image est

$$Y(\Omega) = \{n^2 : n \in \mathbb{Z}\} = \{n^2 : n \in \mathbb{N}\}$$

On a aussi

$$\forall n \in \mathbb{N}^*, P(Y = n^2) = P(X = n) + P(X = -n) = \frac{1}{3^n}$$

et

$$P(Y = 0) = P(X = 0) = \frac{1}{2}$$

2 Moments des variables aléatoires réelles discrètes

2.1 Espérance

Définition 2.1 (Espérance)

Soit X une v.a.r. discrète sur (Ω, \mathcal{A}, P) . On dit que X admet une espérance si :

$$\sum_{x \in X(\Omega)} |x|P(X = x) < +\infty$$

Le cas échéant, l'espérance de X , notée $E(X)$, est le réel :

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

- Si $X(\Omega) = \{x_1, \dots, x_n\}$, alors l'espérance de X existe toujours :

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

- Si $X(\Omega) = \{x_k ; k \in \mathbb{N}^*\}$, alors sous réserve de convergence absolue de la série :

$$E(X) = \sum_{k=1}^{+\infty} x_k P(X = x_k)$$

Remarque :

- $E(X)$ est la moyenne des valeurs x_k prises par X , pondérées par les probabilités correspondantes $P(X = x_k)$.
- $E(X)$ est un paramètre de position, c'est un nombre autour duquel se répartissent les valeurs possibles prises par X .

Exemple 1 : v.a.r.d. d'espérance nulle

Considérons la v.a.r.d. X sur (Ω, \mathcal{A}, P) dont la loi (déjà vu) est définie par :

$$\begin{cases} 1) X(\Omega) = \mathbb{Z} \\ 2) P(X = n) = \frac{1}{2 \times 3^{|n|}}, \quad \forall n \in \mathbb{Z} \end{cases}$$

On a :

$$\sum_{x \in X(\Omega)} |x|P(X = x) = \sum_{n \in \mathbb{Z}} |n| \frac{1}{2 \times 3^{|n|}} = 2 \sum_{n \in \mathbb{N}^*} \frac{n}{2 \times 3^n} + 0 = \sum_{n \in \mathbb{N}^*} \frac{n}{3^n}$$

Comme la série de terme générale $\frac{n}{3^n}$ converge (critère de D'ALEMBERT), l'espérance de X existe. On a :

$$E(X) = \sum_{x \in X(\Omega)} xP(X = x) = \sum_{n \in \mathbb{Z}} n \frac{1}{2 \times 3^{|n|}} = \sum_{n=1}^{+\infty} \frac{-n}{2 \times 3^n} + \sum_{n=1}^{+\infty} \frac{n}{2 \times 3^n} + 0 = 0$$

Exemple 2 : v.a.r.d. n'admettant pas d'espérance

Considérons la v.a.r.d. X sur (Ω, \mathcal{A}, P) définie par :

$$\begin{cases} 1) X(\Omega) = \mathbb{N}^* \\ 2) P(X = n) = \frac{1}{n(n+1)}, \quad \forall n \in \mathbb{N}^* \end{cases}$$

C'est bien une loi de probabilité car :

- pour tout $n \in \mathbb{N}^*$, $P(X = n) = \frac{1}{n(n+1)} \geq 0$
- $\sum_{x \in X(\Omega)} P(X = x) = \sum_{n \in \mathbb{N}^*} P(X = n) = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1} \right) = 1$, par télescopage.

Cependant

$$\sum_{x \in X(\Omega)} |x|P(X = x) = \sum_{n \in \mathbb{N}^*} |n|P(X = n) = \sum_{n=1}^{+\infty} \frac{1}{(n+1)} = +\infty$$

X n'admet donc pas d'espérance.

2.2 Propriétés de l'espérance

Théorème 2.2 (Théorème de transfert)

Soit X une v.a.r. discrète. Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. Sous réserve de convergence absolue de la série, l'espérance de la v.a.r.d. $\varphi(X)$ vaut :

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

Théorème 2.3 (Linéarité de l'espérance)

Soit X et Y deux v.a.r.d. sur le même (Ω, \mathcal{A}, P) , admettant des espérances. Soit $a, b \in \mathbb{R}$. Alors, la v.a.r.d. $aX + bY$ admet une espérance et on a :

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

En particulier

$$E(X + Y) = E(X) + E(Y) \quad ; \quad E(aX + b) = aE(X) + b$$

et

$$E(aX) = aE(X) \quad ; \quad E(a) = a$$

Remarque : Ce théorème signifie que l'ensemble des v.a.r.d. sur (Ω, \mathcal{A}, P) admettant une espérance est un espace vectoriel sur \mathbb{R} stable par multiplication (c'est une \mathbb{R} -algèbre) et que l'application $X \mapsto E(X)$ est une forme linéaire sur cet espace.

Proposition 2.4 (Croissance de l'espérance)

Soit X et Y deux v.a.r.d. sur le même (Ω, \mathcal{A}, P) , admettant une espérance. On a :

$$Y \leq X \quad \implies \quad E(Y) \leq E(X)$$

En particulier :

$$0 \leq X \quad \implies \quad 0 \leq E(X)$$

et pour $a, b \in \mathbb{R}$:

$$a \leq X \leq b \quad \implies \quad a \leq E(X) \leq b$$

2.3 Variance & Ecart-type

Définition 2.5 (Variance & Ecart-type)

Soit X une v.a.r. discrète admettant une espérance $E(X)$. Sous réserve de convergence de la série, on appelle :

- variance de X :

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

- écart-type de X :

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

Remarque :

- $V(X)$ est l'écart quadratique moyen de la variable aléatoire X par rapport à sa valeur moyenne.
- $V(X)$ est donc un indicateur mesurant la dispersion des valeurs x de X autour de leur moyenne pondérée $E(X)$. Plus précisément, c'est $\sigma(X) = \sqrt{V(X)}$ qui mesure l'étendue de cette dispersion.

Définition 2.6 (v.a.r. centrée (resp. centrée réduite))

Soit X une v.a.r. discrète. On dit que :

- X est centrée, si elle admet une espérance nulle.
- X est centrée réduite, si elle admet une espérance nulle et une variance égale à 1.

Théorème 2.7 Soit X une v.a.r. discrète.

- Si X admet une espérance, alors $X - E(X)$ est centrée.
- Si X admet une variance non nulle, alors $X^* = \frac{X - E(X)}{\sqrt{V(X)}} = \frac{X - E(X)}{\sigma(X)}$ est centrée réduite.

Théorème 2.8 (Inégalité de TCHEBYCHEV)

Soit X une v.a.r. discrète. Si X admet une variance, alors :

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

2.4 Moments d'ordre supérieur

Définition 2.9 (Moments)

Soit X une v.a.r. discrète. Sous réserve de convergence absolue de la série, on appelle :

- moment d'ordre m de X , la quantité

$$E(X^m) = \sum_{x \in X(\Omega)} x^m P(X = x)$$

- moment centré d'ordre m de X , la quantité

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

Proposition 2.10 Soit X une v.a.r.d.. Si X admet un moment d'ordre $m \in \mathbb{N}^*$ alors X admet des moments de tout ordre $j \in \llbracket 1, m \rrbracket$.

Théorème 2.11 (Inégalité de MARKOV)

Soit X une v.a.r.d.. Si X admet un moment d'ordre $m \in \mathbb{N}^*$ alors :

$$\forall x > 0, \quad P(|X| \geq x) \leq \frac{E(|X|^m)}{x^m}$$

Proposition 2.12 Si X est une v.a.r. discrète admettant un moment d'ordre 2 alors :

- $V(X) = E(X^2) - E^2(X)$ (Formule de KOENIG-HUYGHENS)
- Si $a, b \in \mathbb{R}$, alors $V(aX + b) = a^2V(X)$.

Exemple : Moments de l'indicatrice d'un événement

Soit A un événement d'un espace probabilisé (Ω, \mathcal{A}, P) . Nous rappelons que $\mathbf{1}_A$, l'indicatrice de A est une v.a.r.d. sur (Ω, \mathcal{A}) . On a :

$$E(\mathbf{1}_A) = P(A) \quad \text{et} \quad V(\mathbf{1}_A) = P(A)P(A^c)$$

2.5 Variables aléatoires réelles discrètes indépendantes

Théorème 2.13 (Indépendance de v.a.r.d.)

- Deux v.a.r.d. X et Y sont indépendantes si et seulement si : $\forall x \in X(\Omega)$ et $\forall y \in Y(\Omega)$

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

- Une suite de n v.a.r.d. X_1, \dots, X_n sont indépendantes si et seulement si : $\forall x_1 \in X_1(\Omega), \dots, \forall x_n \in X_n(\Omega)$

$$P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \dots P(X_n = x_n)$$

Exemple : Jets de deux dés

On jette simultanément deux dés équilibrés discernables. On prend :

$$\Omega = \{1, 2, \dots, 6\}^2 = \{(i, j) : 1 \leq i, j \leq 6\}$$

$(\Omega, \mathcal{P}(\Omega))$ est muni de la probabilité uniforme.

Si on note $X_1 = \ll \text{valeur du premier dé} \gg$ et $X_2 = \ll \text{valeur du second dé} \gg$, on a

$$X_1(i, j) = i \quad \text{et} \quad X_2(i, j) = j$$

et donc $X(\Omega) = Y(\Omega) = \llbracket 1, 6 \rrbracket$

On vérifie que

$$P(X_1 = i) = P(X_2 = i) = \frac{1}{6}, \quad \forall i \in \llbracket 1, 6 \rrbracket$$

et pour tout $i, j \in \llbracket 1, 6 \rrbracket$

$$P(X_1 = i, X_2 = j) = \frac{1}{36} = \frac{1}{6} \times \frac{1}{6} = P(X_1 = i)P(X_2 = j)$$

Les variables X_1 et X_2 sont indépendantes.

Proposition 2.14 Soit X et Y deux v.a.r. discrètes admettant chacune un moment d'ordre 2. Si X et Y sont indépendantes, alors :

$$E(XY) = E(X)E(Y)$$

et

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