VARIABLES ALÉATOIRES DISCRÈTES

Cours

Dans tout ce chapitre, $(\Omega, \mathscr{A}, P)$ désigne un espace probabilisé c'est-à-dire Ω désigne un ensemble, $\mathscr{A}$ une tribu sur Ω et P une probabilité sur $(\Omega, \mathscr{A})$.

I. Généralités sur les variables aléatoires discrètes

A. Définition

Définition 1

On appelle *variable aléatoire discrète sur* $(\Omega, \mathscr{A})$ toute application X définie sur Ω telle que :

- $X(\Omega)$ est un ensemble au plus dénombrable,
- pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) \in \mathscr{A}$.

- L'ensemble image $X(\Omega) = \{X(\omega), \omega \in \Omega\}$ est l'ensemble des valeurs prises par X.

Cet ensemble étant fini ou dénombrable, il peut être décrit en extension.
S'il est fini de cardinal $N \in \mathbb{N}^*$: $X(\Omega) = \{x_1, x_2, \ldots, x_N\}$ et s'il est dénombrable : $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$.
On peut donc écrire dans le cas général :

$$X(\Omega) = \{x_i,\ i \in I\} \text{ où } I \subset \mathbb{N} \text{ et les } x_i \text{ étant tous distincts.}$$

- Soit $x \in X(\Omega)$. L'ensemble $X^{-1}(\{x\})$ est l'image réciproque du singleton $\{x\}$ par X c'est-à-dire l'ensemble des antécédents de x par X :

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

Dire que cet ensemble appartient à la tribu $\mathscr{A}$, c'est dire que cet ensemble est un événement.
L'événement $X^{-1}(\{x\})$ est plutôt noté $(X = x)$ ou $\{X = x\}$.
On pourra donc considérer la probabilité $P((X = x))$ que l'on notera plus simplement $P(X = x)$.

- Soit $A \subset X(\Omega)$. L'ensemble $X^{-1}(A)$ est l'image réciproque de l'ensemble A par X c'est-à-dire :

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

On a $X^{-1}(A) = X^{-1}\left(\bigcup_{x \in A} \{x\}\right) = \bigcup_{x \in A} X^{-1}(\{x\}) \in \mathscr{A}$ par stabilité de $\mathscr{A}$ par union au plus dénombrable.
Ainsi, $X^{-1}(A)$ est un événement qui est plutôt noté $(X \in A)$ ou $\{X \in A\}$.

Dans le cas particulier où X est à valeurs réelles c'est-à-dire $X(\Omega) \subset \mathbb{R}$, on notera pour tout $x \in \mathbb{R}$:

$$(X \leqslant x) = \{X \in]-\infty, x]\},\ (X > x) = \{X \in]x, +\infty[\},\ \textit{etc.}$$

▸ Vous avez étudié en première année les variables aléatoires dans le cas Ω fini et $\mathscr{A} = \mathscr{P}(\Omega)$.
Si $\mathscr{A} = \mathscr{P}(\Omega)$ alors la condition 2 est toujours réalisée.

Exemple 1 : Un joueur lance une pièce jusqu'à obtenir *pile*.
Soit X la variable aléatoire égale au nombre de lancers effectués.
Déterminer $X(\Omega)$.

Définition/Proposition 2

Si X est une variable aléatoire discrète et f une fonction définie sur $X(\Omega)$ alors l'application $f \circ X$ est une variable aléatoire discrète notée $f(X)$.

Exemple 2 : Une urne contient des boules numérotées -2, -1, 0, 1 et 2 (une de chaque).
Un joueur mise 1 euro puis il tire une boule de l'urne.
Il gagne en euros le carré du numéro de la boule tirée.
Soit X la variable aléatoire égale au numéro de la boule tirée et G la variable égale au gain algébrique du joueur.
Exprimer G en fonction de X.

Dans toute la suite, X désigne une variable aléatoire discrète définie sur $(\Omega, \mathscr{A})$.

Proposition 3

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

B. Loi de probabilité

1. Définition

Définition/Proposition 4

Soit X une variable aléatoire discrète définie sur $(\Omega, \mathscr{A})$.

- L'application $P_X : \begin{array}{ccc} \mathscr{P}\big(X(\Omega)\big) & \longrightarrow & [0,1] \\ A & \longmapsto & P(X \in A) \end{array}$ est une probabilité sur $X(\Omega)$.
  Elle est appelée *la loi de X*.
- P_X est entièrement déterminée par la distribution de probabilités $\big(P(X = x)\big)_{x \in X(\Omega)}$.

Dans la pratique, déterminer la loi d'une variable aléatoire X, c'est donner :
1. l'ensemble des valeurs prises par X c'est-à-dire $X(\Omega)$,
2. les probabilités correspondantes c'est-à-dire $P(X = x)$ pour tout $x \in X(\Omega)$.

Proposition 5

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

Cette propriété peut être utile s'il reste une seule valeur x_0 de $X(\Omega)$ pour laquelle on ne connaît pas $P(X=x_0)$ car on a :

$$P(X=x_0) = 1 - \sum_{x \in X(\Omega) \setminus \{x_0\}} P(X=x).$$

Exemple 2 (suite) : Déterminer la loi de X et la loi de G.

Définition/Proposition 6

Soit X et Y deux variables aléatoires discrètes.

- Si X et Y ont la même loi alors on note $X \sim Y$.
- Soit f une fonction définie sur $X(\Omega) \cup Y(\Omega)$.
  Si $X \sim Y$ alors $f(X) \sim f(Y)$.

2. Lois usuelles

a) Loi uniforme

Définition 7

Soit E un ensemble fini non vide.
On dit que X *suit la loi uniforme sur* E et on note $X \sim \mathscr{U}(E)$ lorsque :

$$X(\Omega) = E \quad \text{ et } \quad \forall x \in E,\ P(X=x) = \frac{1}{\text{Card}(E)}.$$

Situation-type : Toutes les valeurs prises par X sont équiprobables.

Cas particuliers : $X \sim \mathscr{U}(\llbracket 1,n \rrbracket)$, $X \sim \mathscr{U}(\llbracket 0,n \rrbracket)$ $(n \in \mathbb{N}^*)$.

Exemple 3 : Une urne contient n boules numérotées de 1 à n.
On tire au hasard successivement et avec remise 3 boules dans cette urne.
Pour tout $i \in \{1,2,3\}$, on note X_i la variable aléatoire égale au numéro de la boule obtenue au tirage numéro i.
On note M la variable aléatoire égale au plus grand numéro des 3 boules tirées.

1. Soit $i \in \{1,2,3\}$. Déterminer la loi de X_i. Pour tout $k \in \llbracket 1,n \rrbracket$, déterminer $P(X_i \leqslant k)$.
2. Écrire M en tant que fonction de X_1, X_2 et X_3.
3. Pour tout $k \in \llbracket 1,n \rrbracket$, déterminer $P(M \leqslant k)$.
4. En déduire la loi de M.

b) Loi de Bernoulli

Définition 8

Soit $p \in [0,1]$.
On dit que X *suit la loi de Bernoulli de paramètre* p et on note $X \sim \mathscr{B}(p)$ lorsque :

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

Situation-type : Si $X(\Omega) = \{0,1\}$ alors $X \sim \mathscr{B}\big(P(X=1)\big)$.

- On appelle *épreuve de Bernoulli* toute expérience aléatoire qui n'admet que deux issues possibles : le succès et l'échec.

La variable aléatoire qui vaut 1 en cas de succès et 0 en cas d'échec suit la loi de Bernoulli de paramètre p, où p est la probabilité de succès.

- Soit A est un événement. La variable aléatoire qui prend la valeur 1 lorsque A est réalisé et 0 sinon, est appelée *variable indicatrice de l'événement* A et elle est notée $\mathbb{1}_A$. On a $\mathbb{1}_A \sim \mathscr{B}(P(A))$.

Exemple 4 : On lance une pièce pour laquelle la probabilité d'obtenir *Face* est $\frac{1}{3}$.
Soit X la variable aléatoire qui prend la valeur 1 lorsque *Face* est obtenue et 0 sinon.
Déterminer la loi de X.

c) Loi binômiale

Définition 9

Soit $n \in \mathbb{N}^*$ et $p \in [0,1]$.
On dit que X *suit la loi binômiale de paramètres* n *et* p et on note $X \sim \mathscr{B}(n,p)$ lorsque :

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

Notons que la loi binômiale de paramètres 1 et p est la loi de Bernoulli de paramètre p : $\mathscr{B}(1,p)$.

Situation-type : Soit $n \in \mathbb{N}^*$. On réalise successivement n épreuves de Bernoulli indépendantes, de même probabilité de succès $p \in [0,1]$.
Soit X la variable aléatoire égale au nombre de succès durant ces n épreuves.
Alors X suit la loi binômiale de paramètres n et p.

Exemple 5 : Nombre de succès lors de tirages successifs avec remise
Une urne contient 2 boules blanches et 5 boules noires.
On effectue 3 tirages successifs avec remise dans cette urne.
Soit X la variable aléatoire égale au nombre de boules blanches obtenues.
Déterminer la loi de X.

d) Loi géométrique

Définition 10

Soit $p \in]0,1[$.
On dit que X *suit la loi géométrique de paramètre* p et on note $X \sim \mathscr{G}(p)$ lorsque :

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

Situation-type : On réalise une suite illimitée d'épreuves de Bernoulli indépendantes, de même probabilité de succès $p \in]0,1[$ (ou on réalise de telles épreuves de Bernoulli jusqu'à obtenir un succès).
Soit X la variable aléatoire égale au rang du premier succès.
Alors X suit la loi géométrique de paramètre p.

Exemple 1 (suite) : Déterminer la loi de X.

Exemple 6 : Rang du premier succès lors de tirages successifs avec remise
Une urne contient 2 boules blanches et 5 boules noires.
On effectue une infinité de tirages successifs dans cette urne avec remise (ou jusqu'à obtenir une boule blanche).
Soit X la variable aléatoire égale au rang d'apparition de la première boule blanche.
Déterminer la loi de X.

Proposition 11

Soit X une variable aléatoire suivant la loi géométrique de paramètre p (où $p \in]0,1[$).
Alors pour tout $k \in \mathbb{N}$, on a $P(X > k) = (1-p)^k$.

e) Loi de Poisson

Définition 12

Soit $\lambda \in]0, +\infty[$.
On dit que X *suit la loi de Poisson de paramètre* λ et on note $X \sim \mathscr{P}(\lambda)$ lorsque :

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

Notons que cette loi ne pourra pas être déduite d'une expérience aléatoire.

Exemple 7 : La loi de Poisson comme loi limite de la loi binômiale
Soit $\lambda \in]0, +\infty[$.
Pour tout $n \in \mathbb{N}$, X_n est une variable aléatoire suivant la loi binômiale de paramètres n et p_n.
On suppose que $\lim\limits_{n \to +\infty} np_n = \lambda$.

Montrer que pour tout $k \in \mathbb{N}$, on a $\lim\limits_{n \to +\infty} P(X_n = k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$.

Interprétation : En particulier, pour n grand, on peut considérer que $\mathscr{B}(n, \frac{\lambda}{n}) \approx \mathscr{P}(\lambda)$.
La loi de Poisson permet de modéliser le nombre de réalisations d'un *événement rare* (événement qui se réalise avec une faible probabilité) dans une période de temps donnée.
Par exemple, si X est le nombre de désintégrations radioactives enregistrées par un compteur Geiger pendant un intervalle de temps donné alors on considère que X suit une loi de Poisson (dont le paramètre est habituellement estimé grâce aux statistiques).

3. Loi conditionnelle

Définition/Proposition 13

Soit X une variable aléatoire discrète définie sur $(\Omega, \mathscr{A})$.
Soit B un événement de probabilité non nulle.
L'application $\begin{array}{ccc} \mathscr{P}\big(X(\Omega)\big) & \longrightarrow & [0,1] \\ A & \longmapsto & P_B(X \in A) \end{array}$ est une probabilité sur $X(\Omega)$.
Elle est appelée *la loi conditionnelle de X sachant B*.

Dans la pratique, déterminer la loi d'une variable aléatoire X sachant un événement B, c'est donner :

1. l'ensemble des valeurs prises par X c'est-à-dire $X(\Omega)$,
2. les probabilités correspondantes c'est-à-dire $P_B(X = x) = \dfrac{P((X = x) \cap B)}{P(B)}$ pour tout $x \in X(\Omega)$.

On remarquera que si $x \in X(\Omega)$ vérifie $(X = x) \cap B = \varnothing$ alors $P_B(X = x) = 0$.
On pourrait se contenter de donner $X(B) = \{X(\omega), \omega \in B\}$ c'est-à-dire l'ensemble des valeurs prises par X lorsque B se réalise.

Exemple 8 : On dispose de deux pièces, la première donnant *pile* avec une probabilité p et la deuxième avec une probabilité q (où $(p, q) \in]0, 1[^2)$.
À chaque tour on lance les deux pièces et on s'arrête lorsque que la deuxième donne *face*.
On note alors T le nombre de tours réalisés et X le nombre de *pile* donnés par la première pièce.

1. Déterminer la loi de T.
2. Pour tout $n \in T(\Omega)$, déterminer la loi de X sachant $(T = n)$.
3. En déduire la loi de X.

C. Couples et n-uplets de variables aléatoires

1. Définition

Définition/Proposition 14

Soit X et Y deux variables discrètes définies sur un même espace probabilisé.
L'application $(X, Y) \ : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & X(\Omega) \times Y(\Omega) \\ \omega & \longmapsto & (X(\omega), Y(\omega)) \end{array} \right.$ est une variable aléatoire discrète, appelée *couple des variables aléatoires discrètes X et Y*.

Étudier le couple (X, Y), c'est étudier <u>simultanément</u> les variables aléatoires X et Y.

Exemple 9 : On lance deux dés.
Soit X la variable aléatoire égale au plus petit résultat des deux dés et Y au plus grand.
Déterminer $(X, Y)(\Omega)$.

Proposition 15

La famille $\big((X = x) \cap (Y = y)\big)_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est un système complet d'événements.

L'événement $(X = x) \cap (Y = y)$ est aussi noté $(X = x, Y = y)$.

On définit plus généralement le n-uplet de variables aléatoires discrètes $(X_1, \dots, X_n)$ comme la variable aléatoire discrète :

$$(X_1, \dots, X_n) : \begin{cases} \Omega & \longrightarrow & X_1(\Omega) \times \dots \times X_n(\Omega) \\ \omega & \longmapsto & (X_1(\omega), \dots, X_n(\omega)). \end{cases}$$

2. Lois de probabilités associées

Définition 16

Soit (X, Y) un couple de variables aléatoires discrètes.

- On appelle *loi conjointe* de X et Y la loi du couple (X, Y).
- On appelle *première loi marginale* du couple (X, Y) la loi de X.
  On appelle *première loi marginale* du couple (X, Y) la loi de Y.

- Dans la pratique, déterminer la loi conjointe de X et Y, c'est donner :
  1. l'ensemble $(X, Y)(\Omega)$ (ens. des couples de valeurs que peuvent prendre X et Y simultanément)
  2. les probabilités $P(X = x, Y = y)$ pour tout $(x, y) \in (X, Y)(\Omega)$.

- Lorsque les variables aléatoires X et Y sont finies, la loi du couple (X, Y) peut être représentée par un tableau à double entrée :

$X \setminus Y$	y_1	y_2	$\dots$	y_m
x_1	$P(X = x_1, Y = y_1)$	$P(X = x_1, Y = y_2)$	$\dots$	$P(X = x_1, Y = y_m)$
x_2	$P(X = x_2, Y = y_1)$	$P(X = x_2, Y = y_2)$	$\dots$	$P(X = x_2, Y = y_m)$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
x_n	$P(X = x_n, Y = y_1)$	$P(X = x_n, Y = y_2)$	$\dots$	$P(X = x_n, Y = y_m)$

- On a :

$$\sum_{(x,y) \in X(\Omega) \times Y(\Omega)} P(X = x, Y = y) = \sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} P(X = x, Y = y) = \sum_{y \in Y(\Omega)} \sum_{x \in X(\Omega)} P(X = x, Y = y) = 1.$$

- Lorsque l'on dispose d'un couple (X, Y) de variables aléatoires, on peut également s'intéresser à :
  - ⋆ la loi conditionnelle de X sachant l'événement $[Y = y]$ pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$,
  - ⋆ la loi conditionnelle de Y sachant l'événement $[X = x]$ pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$.

Exemple 10 : On considère un couple (X, Y) de variables aléatoires.

1. On connaît la loi conjointe.
   Comment obtient-on les lois marginales ? (Illustration sur le tableau ci-dessus.)
   Comment obtient-on les lois conditionnelles ?
2. On connaît la loi de Y et la loi de X sachant $[Y = y]$ pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$.
   Comment obtient-on la loi conjointe ?
   Comment obtient-on les lois marginales ?

Exemple 9 (suite) : Déterminer la loi conjointe de X et Y.
En déduire les lois marginales du couple (X, Y).

On peut de la même façon s'intéresser à diverses lois autour d'un n-uplet $(X_1, \dots, X_n)$: la loi du n-uplet, les lois marginales et diverses lois conditionnelles.

Connaître les lois marginales ne suffit pas pour connaître la loi du couple en général. Par contre, cela suffit si l'on sait de plus que les deux variables aléatoires sont indépendantes.

3. Variables aléatoires indépendantes

Définition 17

Soit X et Y deux variables aléatoires définies sur Ω.
On dit que les variables aléatoires X et Y sont *indépendantes* et on note $X \perp\!\!\!\perp Y$ lorsque :

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

Les assertions suivantes sont équivalentes :

(i) Les variables aléatoires X et Y sont indépendantes.

(ii) Pour tout $(x,y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants.

(iii) Pour tout $x \in X(\Omega)$ tel que $P(X = x) \neq 0$, la loi de Y sachant $(X = x)$ est la loi de Y.

(iv) Pour tout $y \in Y(\Omega)$ tel que $P(Y = y) \neq 0$, la loi de X sachant $(Y = y)$ est la loi de X.

Proposition 18

Si $X \perp\!\!\!\perp Y$ alors pour toutes parties $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, on a :

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

Proposition 19

Si $X \perp\!\!\!\perp Y$ alors pour toutes fonctions f et g, définies respectivement sur $X(\Omega)$ et $Y(\Omega)$, on a $f(X) \perp\!\!\!\perp g(Y)$.

Définition 20

Soit $n \in \mathbb{N}^*$. Soit $X_1, \dots, X_n$ des variables aléatoires définies sur Ω.
Les variables $X_1, \dots, X_n$ sont dites *indépendantes* lorsque :

$$\forall\,(x_1, \dots, x_n) \in X_1(\Omega) \times \cdots \times X_n(\Omega),\ P\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} P(X_i = x_i).$$

L'indépendance implique l'indépendance deux à deux mais la réciproque n'est pas vraie.

Exemple : Soit $n \in \mathbb{N}^*$. Si $X_1, \dots, X_n$ sont n variables aléatoires indépendantes suivant chacune la loi $\mathscr{B}(p)$ alors $X_1 + \cdots + X_n$ suit la loi $\mathscr{B}(n,p)$.

Proposition 21

Si les variables $X_1, \dots, X_n$ sont indépendantes alors pour toutes parties $A_1 \subset X_1(\Omega)$,..., $A_n \subset X_n(\Omega)$, on a $P\left(\bigcap_{i=1}^{n}(X_i \in A_i)\right) = \prod_{i=1}^{n} P(X_i \in A_i)$.

Proposition 22

Si les variables $X_1, \dots, X_n$ sont indépendantes alors pour toutes fonctions $f_1, \dots, f_n$ telles que pour tout $k \in [\![1, n]\!]$, f_k est définie sur $X_k(\Omega)$, les variables $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes.

Théorème 23 (*Lemme des coalitions*)

Si les variables $X_1, \dots, X_n$ sont indépendantes alors pour tout $m \in [\![1, n-1]\!]$, toute variable aléatoire fonction de $X_1, \dots, X_m$ est indépendante de toute variable aléatoire fonction de $X_{m+1}, \dots, X_n$.

Exemple : Si les variables aléatoires $X_1, \dots, X_{n+1}$ sont indépendantes alors les variables aléatoires $\prod_{k=1}^{n} X_k$ et X_{n+1} sont indépendantes, de même que les variables aléatoires $\sum_{k=1}^{n} X_k$ et X_{n+1} (utile dans des récurrences).

Ce résultat est encore vrai pour plus de deux coalitions.

Définition 24

Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires discrètes définies sur Ω.

- On dit que $(X_n)_{n\in\mathbb{N}}$ est une *suite de variables indépendantes* lorsque pour tout $n \in \mathbb{N}$, les variables aléatoires $X_0, \dots, X_n$ sont indépendantes.
- On dit que $(X_n)_{n\in\mathbb{N}}$ est une *suite i.i.d.* lorsque c'est une suite de variables indépendantes et identiquement distribuées c'est-à-dire telles que pour tout $n \in \mathbb{N}$, $X_n \sim X_0$.

Exemple : Jeu de pile ou face infini
On lance une pièce une infinité de fois.
Pour tout $n \in \mathbb{N}^*$, on note X_n la variable aléatoire égale à 1 si on obtient *pile* au $n$ème lancer et 0 sinon.
Alors la suite $(X_n)_{n\in\mathbb{N}^*}$ est une suite i.i.d. de variables aléatoires de Bernoulli.

II. Moments de variables aléatoires discrètes

A. Espérance

1. Définition

Définition 25

Soit X une variable aléatoire discrète.

- $\boxed{\text{Dans le cas } X(\Omega) \subset [0, +\infty]}$
On appelle *espérance de* X l'élément de $[0, +\infty]$ défini par :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$
Convention : $xP(X = x) = 0$ lorsque $x = +\infty$ et $P(X = x) = 0$.
- $\boxed{\text{Dans le cas } X(\Omega) \subset \mathbb{K} \text{ où } \mathbb{K} = \mathbb{R} \text{ ou } \mathbb{K} = \mathbb{C}}$
X est dite *d'espérance finie* lorsque la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable.
Dans ce cas, on appelle *espérance de* X l'élement de $\mathbb{K}$ défini par :
$$E(X) = \sum_{x \in X(\Omega)} xP(X = x).$$

- Notons que $(xP(X = x))_{x \in X(\Omega)}$ étant une famille au plus dénombrable de $[0, +\infty]$ ou de $\mathbb{K}$, les notions de sommabilité et de somme ont été définies en préambule du chapitre « Espaces probabilisés ».

- Au lieu de « X est d'espérance finie » , on rencontre parfois « X a une espérance ».

- L'espérance est la moyenne des issues possibles pondérées par leurs probabilités.

- L'espérance de toutes les lois usuelles est à connaître et à savoir retrouver (*cf* tableau p.11).

Notons en particulier que si A est un événement, $\mathbb{1}_A \sim \mathscr{B}(P(A))$ et donc $E(\mathbb{1}_A) = P(A)$.

Proposition 26

Hyp. On suppose que X est à valeurs dans $\mathbb{N} \cup \{+\infty\}$.
On a :
$$E(X) = \sum_{n=1}^{+\infty} P(X \geqslant n).$$

Exemple 11 : Retrouver l'espérance de la loi géométrique de paramètre $p \in]0, 1[$.

Définition 27

Une variable aléatoire est dite *centrée* lorsque $E(X) = 0$.

Lois discrètes usuelles

	Nom	Notation	Support $X(\Omega)$	Valeur de $P(X=k)$ pour tout $k \in X(\Omega)$	Espérance	Variance	Série génératrice
Lois finies	Loi uniforme sur $[\![1,n]\!]$ $n \in \mathbb{N}^*$	$\mathscr{U}([\![1,n]\!])$	$[\![1,n]\!]$	$\dfrac{1}{n}$	$\dfrac{n+1}{2}$	$\dfrac{n^2-1}{12}$	$\forall t \in \mathbb{R},\ G_X(t) = \begin{cases} \dfrac{t}{n}\dfrac{1-t^n}{1-t} & \text{si } t \neq 1 \\ 1 & \text{si } t = 1 \end{cases}$
	Loi de Bernoulli de paramètre p $p \in [0,1]$	$\mathscr{B}(p) = \mathscr{B}(1,p)$	$\{0,1\}$	$P(X=1) = p$ $P(X=0) = 1-p$	p	$p(1-p)$	$\forall t \in \mathbb{R},\ G_X(t) = 1-p+pt$
	Loi binômiale de paramètres n et p $n \in \mathbb{N}^*,\ p \in [0,1]$	$\mathscr{B}(n,p)$	$[\![0,n]\!]$	$\dbinom{n}{k}p^k(1-p)^{n-k}$	np	$np(1-p)$	$\forall t \in \mathbb{R},\ G_X(t) = (1-p+pt)^n$
Lois infinies	Loi géométrique de paramètre p $p \in]0,1[$	$\mathscr{G}(p)$	$\mathbb{N}^*$	$p(1-p)^{k-1}$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$R = \frac{1}{1-p} > 1$ et $\forall t \in]-R,R[$: $G_X(t) = \dfrac{pt}{pt-t+1}$
	Loi de Poisson de paramètre λ $\lambda \in]0,+\infty[$	$\mathscr{P}(\lambda)$	$\mathbb{N}$	$e^{-\lambda}\dfrac{\lambda^k}{k!}$	λ	λ	$\forall t \in \mathbb{R},\ G_X(t) = e^{\lambda(t-1)}$

2. Théorème du transfert

Théorème 28 (*Formule de transfert*)

Soit X une variable aléatoire discrète. Soit f une fonction définie sur $X(\Omega)$.

- $\boxed{\text{Dans le cas } f(X)(\Omega) \subset [0,+\infty]}$

  On a :
  $$E\big(f(X)\big) = \sum_{x\in X(\Omega)} f(x)P(X=x).$$
- $\boxed{\text{Dans le cas } f(X)(\Omega) \subset \mathbb{K} \text{ où } \mathbb{K}=\mathbb{R} \text{ ou } \mathbb{K}=\mathbb{C}}$

  $f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)P(X=x)\big)_{x\in X(\Omega)}$ est sommable.
  Dans ce cas :
  $$E\big(f(X)\big) = \sum_{x\in X(\Omega)} f(x)P(X=x).$$

Ce théorème permet de déterminer si $f(X)$ est d'espérance finie et de calculer $E(f(X))$ sans avoir à déterminer la loi de $f(X)$.

Exemple 12 : On suppose que X suit une loi géométrique de paramètre p. Calculer $E\left(\dfrac{1}{X}\right)$.

Ce théorème peut s'appliquer avec un couple ou un n-uplet de variables aléatoires.

Exemple 13 : Soit $n \in \mathbb{N}$, $n \geqslant 2$. Une urne contient n boules numérotées de 1 à n. On tire deux boules successivement et sans remise dans cette urne. Soit X la variable aléatoire égale au premier numéro obtenu et Y la variable aléatoire égale au deuxième numéro obtenu.
Déterminer la loi du couple (X,Y) puis calculer $E(XY)$.

3. Propriétés

Proposition 29 (*Linéarité de l'espérance*)

- Si X et Y sont d'espérance finie alors pour tout $\lambda \in \mathbb{K}$, $\lambda X + Y$ est d'espérance finie et on a :
  $$E(\lambda X + Y) = \lambda E(X) + E(Y).$$
- Si $X_1,\dots,X_n$ sont d'espérance finie alors pour tout $(\lambda_1,\dots,\lambda_n) \in \mathbb{K}^n$, $\sum_{i=1}^{n} \lambda_i X_i$ est d'espérance finie et on a $E\left(\sum_{i=1}^{n} \lambda_i X_i\right) = \sum_{i=1}^{n} \lambda_i E(X_i)$.

- Si les variables aléatoires sont toutes à valeurs dans $[0,+\infty]$ et si les scalaires sont aussi des éléments de $[0,+\infty]$ alors les formules restent valables sans l'hypothèse d'espérance finie.
- Si X est d'espérance finie alors la variable $X - E(X)$ est centrée.

Exemple 14 : On suppose que $X \sim \mathscr{B}(n,p)$. En écrivant X comme une somme de n variables suivant une loi de Bernoulli, retrouver la valeur de $E(X)$.

Proposition 30

Si $|X| \leqslant Y$ et si $E(Y) < +\infty$ alors X est d'espérance finie.

Proposition 31

- *Positivité :* Si $X \geqslant 0$ alors $E(X) \geqslant 0$.
- *Croissance :* Si $X \leqslant Y$ et X et Y sont d'espérance finie alors $E(X) \leqslant E(Y)$.
- *Nullité pour une variable positive :*
  Si $X \geqslant 0$ et $E(X) = 0$ alors $(X = 0)$ est presque sûr.

En particulier, si $X(\Omega) \subset [a, b]$ et X est d'espérance finie alors on a $a \leqslant E(X) \leqslant b$.

Théorème 32

Si X et Y sont indépendantes et d'espérance finie alors XY est d'espérance finie et on a :

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

Corollaire 33

Si $X_1, \dots, X_n$ sont n variables indépendantes et d'espérance finie alors $\prod_{k=1}^{n} X_k$ est d'espérance finie et on a :

$$E\left(\prod_{k=1}^{n} X_k\right) = \prod_{k=1}^{n} E(X_k).$$

B. Variance et covariance

Dans ce paragraphe, les variables aléatoires sont à valeurs réelles.

1. Inégalité de Cauchy-Schwarz

Proposition 34

Si X^2 est d'espérance finie alors X est d'espérance finie.

Théorème 35 (*Inégalité de Cauchy-Schwarz*)

Si X^2 et Y^2 sont d'espérance finie alors XY est d'espérance finie et on a :

$$\big(E(XY)\big)^2 \leqslant E(X^2)E(Y^2).$$

Cas d'égalité :
$\big(E(XY)\big)^2 = E(X^2)E(Y^2)$ si et seulement si $X = 0$ ps ou il existe $a \in \mathbb{R}$ tel que $Y = aX$ ps.

2. Variance et écart-type

Définition/Proposition 36

- On dit que X *admet une variance* lorsque X^2 est d'espérance finie.
- Dans ce cas, la variable aléatoire $\big(X - E(X)\big)^2$ est d'espérance finie.
  On appelle alors *variance de* X et on note $V(X)$ le réel :

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

- Au lieu de « X admet une variance », on rencontre parfois « X a une variance finie ».

- Par positivité de l'espérance, on obtient que $V(X)$ est un réel positif.

Par le cas de nullité pour une variable positive, $V(X) = 0$ si et seulement si X est constante ps.

- La variance est l'espérance des carrés des écarts à l'espérance.

Elle est un indicateur de la dispersion de la variable aléatoire.

Proposition 37

Si X^2 est d'espérance finie alors on a :

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

- Par le théorème du transfert :

$$X \text{ admet une variance si et seulement si } \sum_{x \in X(\Omega)} x^2 P(X = x) < +\infty$$

et dans ce cas :

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

- La variance des lois usuelles est à connaître et à savoir retrouver (*cf* tableau).

Définition 38

On suppose que X admet une variance.
On appelle *écart-type de* X le réel $\sigma(X) = \sqrt{V(X)}$.

Définition 39

Une variable aléatoire est dite *réduite* lorsque $V(X) = 1$.

Proposition 40

Soit a et b deux réels.
Si X admet une variance alors $aX + b$ aussi et on a $V(aX + b) = a^2 V(X)$.

Proposition 41

On suppose que X admet une variance non nulle.
Alors la variable $\dfrac{X - E(X)}{\sigma(X)}$ est centrée et réduite.

Exemple 15 : Soit X une variable aléatoire vérifiant $X(\Omega) = \mathbb{N}^*$ et $\forall n \in \mathbb{N}^*,\ P(X = n) = \dfrac{1}{n(n+1)}$.

1. Calculer l'espérance de X. Est-ce que X admet une variance ?
2. La variable $\sqrt{X}$ est-elle d'espérance finie ? admet-elle une variance ?
3. La variable $(-1)^X X$ est-elle d'espérance finie ?
4. Montrer que la variable $(-1)^X\sqrt{X}$ est d'espérance finie et la donner sous forme d'une somme. Cette variable admet-elle une variance ?

3. Covariance

Définition/Proposition 42

On suppose que X et Y admettent une variance.

- La variable aléatoire $\big(X - E(X)\big)\big(Y - E(Y)\big)$ est d'espérance finie.
- On appelle *covariance de X et Y* le réel :

$$\mathrm{Cov}(X,Y) = E\Big(\big(X - E(X)\big)\big(Y - E(Y)\big)\Big).$$

- Si X admet une variance alors $\mathrm{Cov}(X,X) = V(X)$.
- La covariance est symétrique : $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.
- Par l'inégalité de Cauchy-Schwarz, on a $|\mathrm{Cov}(X,Y)| \leqslant \sigma(X)\sigma(Y)$.
- Si X est constante presque sûrement alors $\mathrm{Cov}(X,Y) = 0$.

Proposition 43

On suppose que X et Y admettent une variance.
On a :

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

Proposition 44

On suppose que X et Y admettent une variance.
Si $X \perp\!\!\!\perp Y$ alors $\mathrm{Cov}(X,Y) = 0$.

Proposition 45 (*Bilinéarité*)

- On suppose que X, Y et Z admettent une variance. Soit $\lambda \in \mathbb{R}$.
  On a :
  $$\mathrm{Cov}(\lambda X + Y, Z) = \lambda \mathrm{Cov}(X, Z) + \mathrm{Cov}(Y, Z).$$
- Soit $n \in \mathbb{N}^*$. On suppose que $X_1, \ldots, X_n$ et Y admettent une variance.
  Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$. On a :
  $$\mathrm{Cov}\left(\sum_{i=1}^{n} \lambda_i X_i, Y\right) = \sum_{i=1}^{n} \lambda_i \mathrm{Cov}(X_i, Y).$$

Par symétrie, ces propriétés sont également vraies à droite.

Proposition 46 (*Variance d'une somme de deux variables aléatoires*)

- Si X et Y admettent une variance alors $X + Y$ aussi et on a :
  $$V(X+Y) = V(X) + V(Y) + 2\mathrm{Cov}(X, Y).$$
- Si de plus X et Y sont indépendantes alors :
  $$V(X+Y) = V(X) + V(Y).$$

Proposition 47 (*Généralisation à un nombre fini de variables aléatoires*)

- Si $X_1, \ldots, X_n$ admettent une variance alors $\sum\limits_{i=1}^{n} X_i$ aussi et on a :
  $$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i) + \sum_{\substack{(i,j)\in[\![1,n]\!]^2 \\ i\neq j}} \mathrm{Cov}(X_i, X_j) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i, X_j).$$
- Si de plus les variables $X_1, \ldots, X_n$ sont deux à deux indépendantes, on a :
  $$V\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} V(X_i).$$

Exemple 16 : Retrouver la variance de X lorsque $X \sim \mathscr{B}(n, p)$ en écrivant X comme une somme de n variables aléatoires indépendantes suivant toutes la loi $\mathscr{B}(p)$.

Exemple 17 : On suppose que $X_1, \ldots, X_n$ admettent une variance.
On note $M = (m_{i,j})_{1\leqslant i,j\leqslant n}$ la matrice de variances-covariances associée c'est-à-dire définie par :

$$\forall (i,j) \in [\![1,n]\!]^2,\ m_{i,j} = \mathrm{Cov}(X_i, X_j).$$

On remarquera qu'il s'agit d'une matrice symétrique réelle.
Montrer que pour tout $X \in \mathscr{M}_{n,1}(\mathbb{R})$, $X^{\top} M X \geqslant 0$.

C. Fonction génératrice d'une variable aléatoire à valeurs dans $\mathbb{N}$

Définition 48

Soit X une variable aléatoire discrète à valeurs dans $\mathbb{N}$.
On appelle *fonction génératrice de* X et on note G_X la fonction définie par :

$$G_X(t) = E\left(t^X\right) = \sum_{n=0}^{+\infty} P(X = n)t^n.$$

Par le théorème du transfert, la variable t^X est d'espérance finie si et seulement si la série $\sum P(X = n)t^n$ converge absolument et dans ce cas, $E(t^X) = \sum_{n=0}^{+\infty} P(X = n)t^n$.
Cela revient donc à l'étude d'une série entière.

Proposition 49

- La série entière $\sum P(X = n)t^n$ a un rayon de convergence supérieur ou égal à 1.
- En tant que série de fonctions, elle converge normalement sur $[-1, 1]$.
- La fonction génératrice G_X est définie et continue sur $[-1, 1]$.

Proposition 50

La loi d'une variable aléatoire à valeurs dans $\mathbb{N}$ est caractérisée par sa fonction génératrice.

La fonction génératrice des lois usuelles est à connaître et à savoir retrouver (voir tableau)

Théorème 51

On suppose que X est à valeurs dans $\mathbb{N}$.

- X est d'espérance finie si et seulement si G_X est dérivable en 1.
  Dans ce cas, on a :
  $$E(X) = G'_X(1).$$
- X admet une variance si et seulement si G_X est deux fois dérivable en 1.
  Dans ce cas, on a :
  $$E(X(X-1)) = G''_X(1) \text{ d'où par linéarité } V(X) = G''_X(1) + G'_X(1)(1 - G'_X(1)).$$

Si la série entière $\sum P(X = n)t^n$ a un rayon de convergence strictement supérieur à 1 alors X est d'espérance finie et a une variance (puisque G_X est de classe $\mathscr{C}^\infty$ sur l'intervalle ouvert de convergence).
Si le rayon de convergence est égal à 1 alors « (deux fois) dérivable en 1 » signifie « (deux fois) dérivable à gauche en 1 ».

Exemple 18 : Retrouver l'espérance et la variance des lois $\mathscr{B}(n, p)$ et $\mathscr{P}(\lambda)$ à l'aide de ce théorème.

Théorème 52

Les variables considérées sont à valeurs dans $\mathbb{N}$.

- Si $X \perp\!\!\!\perp Y$ alors $G_{X+Y} = G_X \times G_Y$.
- Si les variables aléatoires $X_1, \ldots, X_n$ sont indépendantes alors $G_{\sum_{i=1}^n X_i} = \prod_{i=1}^{n} G_{X_i}$.

Exemple 19 : Montrer que $X \sim \mathscr{P}(\lambda)$, $Y \sim \mathscr{P}(\mu)$ et $X \perp\!\!\!\perp Y$ alors $X + Y \sim \mathscr{P}(\lambda + \mu)$.

Exemple 20 : Retrouver la fonction génératrice de la loi $\mathscr{B}(n, p)$.

III. Inégalités

A. Inégalité de Markov

Proposition 53 (*Inégalité de Markov*)

On suppose que X est positive.
Pour tout $a > 0$, on a :

$$P(X \geqslant a) \leqslant \frac{E(X)}{a}.$$

En appliquant l'inégalité de Markov à $|X|^r$ pour $r \in \mathbb{N}^*$ (comme $|X| \geqslant 0$), on obtient :

$$\forall a > 0,\ P(|X|^r \geqslant a) \leqslant \frac{E(|X|^r)}{a}.$$

Exemple 21 : Soit X une variable aléatoire discrète. Soit $\alpha \in \mathbb{R}$.
Montrer que $P(X \geqslant \alpha) \leqslant e^{-\alpha} E(e^X)$.

B. Inégalité de Bienaymé-Tchebychev

Proposition 54 (*Inégalité de Bienaymé-Tchebychev*)

On suppose que X^2 est d'espérance finie.
Pour tout $\varepsilon > 0$, on a :

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

- Si l'on connaît la loi de X, on peut calculer les probabilités $P(X \geqslant a)$ et $P(|X - E(X)| \geqslant \varepsilon)$. Grâce aux inégalités de Markov et Bienaymé-Tchebychev, même si l'on ne connaît pas la loi de X, on peut obtenir une majoration de ces probabilités à condition de connaître l'espérance et la variance (mais cette majoration est assez grossière).
- L'inégalité de Bienaymé-Tchebychev rappelle que la variance est un indicateur de dispersion.
- Par passage à l'événement contraire, on a :

$$\forall \varepsilon > 0,\ P(|X - E(X)| < \varepsilon) \geqslant 1 - \frac{V(X)}{\varepsilon^2}.$$

C. Loi faible des grands nombres

Théorème 55

On suppose que $(X_n)_{n\geqslant 1}$ est une suite de variables i.i.d. de variance finie.
On note $m = E(X_1)$ et $\sigma = \sigma(X_1)$. Pour tout $n \in \mathbb{N}^*$, on pose $S_n = \sum_{k=1}^{n} X_k$.
Alors pour tout $\varepsilon > 0$, on a :

$$\forall n \in \mathbb{N}^*,\quad P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2} \text{ et donc } \lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) = 0.$$

- Par passage au complémentaire, on a $\forall \varepsilon > 0,\ \lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| < \varepsilon\right) = 1$.
- On réalise une même expérience aléatoire plusieurs fois, dans les mêmes conditions.
  - Pour tout $n \in \mathbb{N}^*$, on note X_n le résultat obtenu à la n-ème expérience (supposé numérique). Si l'on réalise cette expérience un grand nombre de fois alors d'après la loi des grands nombres, la moyenne des résultats obtenus (moyenne empirique) est très proche de l'espérance du résultat (moyenne théorique).
  - On s'intéresse à un événement A et pour tout $n \in \mathbb{N}$, on note X_n la variable aléatoire qui prend la valeur 1 si A est réalisé à la n-ème expérience et 0 s'il ne l'est pas.
    Si l'on réalise cette expérience un grand nombre de fois alors d'après la loi des grands nombres, la fréquence de réalisation de l'événement A est très proche de la probabilité de l'événement A.
    Ainsi, la probabilité d'un événement coïncide avec la notion historique intuitive de fréquence limite d'apparition de cet événement lors d'une succession infinie de réalisations de l'expérience aléatoire.
- Ce résultat permet d'estimer une espérance ou une probabilité à l'aide de simulations informatiques.