P2 - Variables aléatoires discrètes

Dans tout le chapitre $(\Omega, \mathscr{T}, \mathbb{P})$ est un espace probabilisé.

1 Généralités

1.1 Variables aléatoires discrètes

Définition 1
Soit E un ensemble.
- On appelle variable aléatoire discrète sur Ω une application $X : \Omega \to E$, telle que :
  1. $X(\Omega)$ est au plus dénombrable.
  2. Pour tout $x \in X(\Omega)$, $X^{-1}(\{x\}) = \{\omega \in \Omega / X(\omega) = x\} \in \mathscr{T}$.
- Si $E = \mathbb{R}$, X est appelée *variable aléatoire réelle.*
- Si $E = \mathbb{R}^2$, X est appelé *couple de variables aléatoires réelles.*

Remarque 1
- L'événement $X^{-1}(\{x\})$ se note $(X = x)$.
- Dans le cas d'un univers fini, toutes les applications $X : \Omega \to \mathbb{R}$ sont des variables aléatoires discrètes. Ce n'est pas le cas lorsque l'univers est infini (lorsque $\mathscr{T}$ n'est pas $\mathscr{P}(\Omega)$).

Dans la suite du chapitre, X désignera une variable aléatoire réelle discrète à valeurs dans $E = X(\Omega)$, et on notera :

$$X(\Omega) = \{x_n, n \in \mathbb{N}\}$$

(certains x_n pouvant être égaux si l'ensemble est fini !)

Proposition 1
- Pour tout $A \subset E$, $X^{-1}(A) \in \mathscr{T}$.
  L'événement $X^{-1}(A)$ se note également $\{X \in A\}$ ou $(X \in A)$.

- Pour tout $A \subset E$, et toute suite $(A_i)_{i\in I}$ $(I \subset \mathbb{N})$ de sous-ensembles de E, on a :

$$\left(X \in \overline{A}\right) = \overline{(X \in A)}, \qquad \left(X \in \bigcup_{i\in I} A_i\right) = \bigcup_{i\in I}(X \in A_i), \qquad \left(X \in \bigcap_{i\in I} A_i\right) = \bigcap_{i\in I}(X \in A_i)$$

- Les événements $(X = x_n)_{n\in\mathbb{N}}$ forment un système complet d'événements.

Théorème-Définition 1
Soit U un événement. L'application $\mathbb{1}_U$ définie sur Ω par :

$$\mathbb{1}_U : \left| \begin{array}{l} \Omega \to [0,1] \\ \omega \mapsto \begin{cases} 1 \text{ si } \omega \in U \\ 0 \text{ sinon} \end{cases} \end{array} \right.$$

est une variable aléatoire réelle discrète. Elle est appelée *fonction indicatrice* de U.

1.2 Loi de probabilité

Théorème-Définition 2
On définit l'application :

$$\mathbb{P}_X : \left| \begin{array}{ccc} \mathscr{P}(X(\Omega)) & \to & [0,1] \\ A & \mapsto & \mathbb{P}_X(A) = \mathbb{P}(X \in A) \end{array} \right.$$

$\mathbb{P}_X$ est une probabilité sur $(X(\Omega), \mathscr{P}(X(\Omega)))$, appelée *loi de probabilité* de X.

Définition 2
Deux variables aléatoires discrètes X et Y définies sur Ω sont dites *équiréparties* lorsqu'elles ont la même loi. On note $X \sim Y$.

Remarque 2
- La série de terme général $\mathbb{P}(X = x_n)$ converge, et sa somme vaut 1.

Proposition 2
La loi $\mathbb{P}_X$ de la variable aléatoire discrète X est entièrement déterminée par la donnée de :

1. $X(\Omega) = \{x_n, n \in \mathbb{N}\}$;
2. $\forall n \in \mathbb{N}, \mathbb{P}_X(\{x_n\}) = \mathbb{P}(X = x_n)$ également notée $\mathbb{P}_X(x_n)$.

Définition 3
La *fonction de répartition* d'une variable aléatoire réelle discrète X est la fonction :

$$F_X : \left| \begin{array}{l} \mathbb{R} \to [0,1] \\ x \mapsto F_X(x) = \mathbb{P}(X \leq x) = \displaystyle\sum_{k \in M} \mathbb{P}(X = x_k) \end{array} \right.$$

où $M = \{k \in I / x_k \leq x\}$.

Proposition 3
- F_X est croissante sur $\mathbb{R}$.
- $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.
- Soient a et b deux réels tels que $a \leq b$; alors $\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a)$.

1.3 Couple de variables aléatoires réelles

Proposition 4
Soit f une fonction définie sur E à valeurs dans un ensemble F.
L'application $f \circ X : \omega \in \Omega \to f(X(\omega)) \in F$ est une variable aléatoire réelle discrète, notée $f(X)$.

Proposition 5
Soit $X : \omega \in \Omega \to (X_1(\omega), X_2(\omega)) \in \mathbb{R}^2$ une application définie sur Ω.
X est un couple aléatoire discret si, et seulement si X_1 et X_2 sont deux variables aléatoires discrètes.

Proposition 6
Soient X et Y deux variables aléatoires réelles discrètes, et $(a,b) \in \mathbb{R}^2$.
L'application $aX + bY : \omega \in \Omega \mapsto aX(\omega) + bY(\omega)$ est une variable aléatoire réelle discrète.

Définition 4
Soient X_1 et X_2 des variables aléatoires réelles discrètes.

- On appelle *loi conjointe* de X_1 et X_2, la loi du couple $X = (X_1, X_2)$ entièrement déterminée par la donnée de :
  1. $X_1(\Omega) \times X_2(\Omega)$.
  2. $\forall (x_1, x_2) \in X(\Omega), \mathbb{P}(X = (x_1, x_2)) = \mathbb{P}\left((X_1 = x_1) \cap (X_2 = x_2)\right)$, que l'on note $\mathbb{P}_X(x_1, x_2)$.
- Les lois de X_1 et X_2 sont appelées *lois marginales* du couple X.
- Soit $x \in \mathbb{R}$ tel que $\mathbb{P}(X_1 = x) \neq 0$. L'application :

$$\begin{array}{ll} \mathscr{P}(X_2(\Omega)) & \to [0,1] \\ A & \mapsto \mathbb{P}_{(X_1 = x)}(X_2 \in A) \end{array}$$

  est une probabilité sur $(X_2(\Omega), \mathscr{P}(X_2(\Omega)))$, appelée *loi conditionnelle* sachant $(X_1 = x)$ de X_2.

Remarque 3

- On note $X_1(\Omega) = \{x_i, i \in I\}$ et $X_2(\Omega) = \{y_j, j \in J\}$ où I et J sont des parties de $\mathbb{N}$.
  Pour déterminer les lois marginales à partir de la loi conjointe on utilise le fait que l'événement $\{X_1 = x_i\}$ est la réunion disjointe des événements $\{(X_1 = x_i) \cap (X_2 = y_j)\}$, pour tous $j \in J$:

$$\forall i \in I, \mathbb{P}(X_1 = x_i) = \sum_{j \in J} \mathbb{P}\left((X_1 = x_i) \cap (X_2 = y_j)\right)$$

  de même :

$$\forall j \in J, \mathbb{P}(X_2 = y_j) = \sum_{i \in I} \mathbb{P}\left((X_1 = x_i) \cap (X_2 = y_j)\right)$$

Définition 5
On dit que deux variables aléatoires réelles discrètes X_1 et X_2 sont *indépendantes* si pour tous $x \in X_1(\Omega)$, et $y \in X_2(\Omega)$ les événement $\{X_1 = x\}$ et $\{X_2 = y\}$ sont indépendants.

Remarque 4

- Dans ce cas on a : $\mathbb{P}_X(x, y) = \mathbb{P}_{X_1}(x)\, \mathbb{P}_{X_2}(y)$.

Proposition 7
Soit $X = (X_1, X_2)$ un couple de variables aléatoires indépendantes.
Alors pour tous $A \subset X_1(\Omega)$, et $B \subset X_2(\Omega)$:

$$\mathbb{P}\left((X_1 \in A) \cap (X_2 \in B)\right) = \mathbb{P}(X_1 \in A)\, \mathbb{P}(X_2 \in B)$$

Proposition 8
Si X_1 et X_2 sont deux variables aléatoires réelles indépendantes, alors pour toutes fonctions f et g de $\mathbb{R}$ dans $\mathbb{R}$, $f(X_1)$ et $g(X_2)$ sont indépendantes.

Définition 6
Soient $X_1, ..., X_n$ des variables aléatoires réelles discrètes. On dit qu'elles sont *mutuellement indépendantes* si pour tout $(x_1, ..., x_n) \in X_1(\Omega) \times ... \times X_n(\Omega)$, les événements $(X_1 = x_1), ..., (X_n = x_n)$ sont indépendants.

2 Caractéristiques des variables aléatoires

2.1 Espérance

Définition 7
La variable aléatoire réelle discrète X à valeurs dans $\{x_n, n \in \mathbb{N}\}$ est dite *d'espérance finie* si la série $\sum x_n \mathbb{P}(X = x_n)$ est absolument convergente.

Si tel est le cas, on appelle sa somme *espérance* de X, notée :

$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n \, \mathbb{P}(X = x_n)$$

Une variable aléatoire réelle discrète admettant une espérance nulle est dite *centrée*.

Exemple 1
- Si X est une application constante sur Ω, telle que $\forall \omega \in \Omega, X(\omega) = a$, alors X admet une espérance finie qui vaut a.
- Si U est un événement, la fonction indicatrice $\mathbb{1}_U$ admet une espérance finie qui vaut $\mathbb{P}(U)$.

Théorème 1 Théorème de transfert
Si X est une variable aléatoire réelle discrète, et f une application à valeurs réelles définie sur $X(\Omega) = \{x_n, n \in \mathbb{N}\}$, alors $f(X)$ est d'espérance finie si, et seulement si la série de terme général $f(x_n) \, \mathbb{P}(X = x_n)$ converge absolument. Dans ce cas, on a :

$$\mathbb{E}(f(X)) = \sum_{n=0}^{+\infty} f(x_n) \, \mathbb{P}(X = x_n)$$

Proposition 9
Soient X et Y des variables aléatoires réelles discrètes ayant une espérance finie, alors :

- $\forall \lambda \in \mathbb{R}$, $\mathbb{E}(\lambda X) = \lambda \mathbb{E}(X)$.
- $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.
- $X \geq 0 \Rightarrow \mathbb{E}(X) \geq 0$.
- $X \leq Y \Rightarrow \mathbb{E}(X) \leq \mathbb{E}(Y)$.
- Si X et Y sont indépendantes, $\mathbb{E}(XY) = \mathbb{E}(X) \, \mathbb{E}(Y)$.
  Attention ! La réciproque est fausse.
- Si X est à valeurs dans $\mathbb{N}$, alors $\mathrm{E}(X) = \displaystyle\sum_{n=1}^{+\infty} \mathbb{P}(X \geq n)$.

Proposition 10 Inégalité de Markov
Si X est une variable aléatoire réelle discrète **positive** admettant une espérance finie, alors pour tout réel $a > 0$, on a :

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$$

2.2 Variance

Définition 8
Soient X une variable aléatoire réelle discrète, et $m \in \mathbb{N}^*$. Si X^m admet une espérance finie, on dit que X admet un *moment d'ordre m*, et ce moment d'ordre m est $\mathbb{E}(X^m)$.

Proposition 11
Si X admet un moment d'ordre 2, alors X admet une espérance finie .

Définition 9
Soit X admettant un moment d'ordre 2. On définit la *variance* de X, notée $\mathrm{V}(X)$ par :

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

et *l'écart type* de X, noté $\sigma(X)$ par :

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

Proposition 12
Soit X une variable aléatoire réelle discrète admettant un moment d'ordre 2. Alors :
- $\mathrm{V}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ (Formule de **König-Huygens**).
- $\forall (a, b) \in \mathbb{R}^2, \mathrm{V}(aX + b) = a^2\, \mathrm{V}(X)$.

Proposition 13 Inégalité de Bienaymé-Tchebychev
Soit X une variable aléatoire réelle discrète, admettant un moment d'ordre 2.
Alors pour tout réel $a > 0$:

$$\mathbb{P}\left(|X - \mathbb{E}(X)| \geq a\right) \leq \frac{\mathrm{V}(X)}{a^2}$$

2.3 Covariance

Définition 10
Soient X et Y deux variables aléatoires réelles discrètes admettant une espérance finie.
Si elle existe, on appelle *covariance* du couple (X, Y) le nombre :

$$\mathrm{cov}(X, Y) = \mathbb{E}\left((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right),$$

Remarque 5
- Pour toute variable aléatoire réelle X admettant un moment d'ordre 2, on a : $\mathrm{cov}(X, X) = \mathrm{V}(X)$.
- Sous réserve d'existence, on a : $\mathrm{cov}(X, Y) = \mathrm{cov}(Y, X)$.
- Si X est une variable aléatoire constante ($\forall \omega \in \Omega, X(\omega) = a$), alors pour toute variable aléatoire Y : $\mathrm{cov}(X, Y) = 0$.
- Sous réserve d'existence, $\forall (a, b) \in \mathbb{R}^2, \mathrm{cov}(aX + bY) = a\,\mathrm{cov}(X, Y)$

Proposition 14
Soient X et Y deux variables aléatoires réelles discrètes admettant des moments d'ordre 2.
Alors XY a une espérance finie, et le couple (X, Y) admet une covariance, qui vaut :

$$\mathrm{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\,\mathbb{E}(Y)$$

Proposition 15
Soient X et Y deux variables aléatoires réelles discrètes admettant des moments d'ordre 2.
Si X et Y sont indépendantes, alors $\mathrm{cov}(X, Y) = 0$.

Remarque 6
- On utilise plus fréquemment la contraposée de ce résultat pour montrer que des variables aléatoires admettant des moments d'ordre 2 ne sont pas indépendantes.

Proposition 16 Inégalité de Cauchy-Schwarz
Soient X et Y deux variables aléatoires réelles discrètes admettant des moments d'ordre 2. Alors :

$$(\mathbb{E}(XY))^2 \leq \mathbb{E}(X^2)\,\mathbb{E}(Y^2)$$

Proposition 17
Soient X et Y deux variables aléatoires réelles discrètes admettant des moments d'ordre 2. Alors :
- $X+Y$ admet un moment d'ordre 2, et $\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y) + 2\,\mathrm{cov}(X,Y)$.
- X et Y sont indépendantes $\Rightarrow \mathrm{cov}(X,Y) = 0 \Rightarrow \mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y)$

Définition 11
Soient X et Y deux variables aléatoires réelles discrètes admettant des écarts-types non nuls.
On appelle *coefficient de corrélation linéaire* de X et Y le réel :

$$\rho(X,Y) = \frac{\mathrm{cov}(X,Y)}{\sigma(X)\,\sigma(Y)}$$

Proposition 18
Sous les hypothèses de la définition, on a : $-1 \leq \rho(X,Y) \leq 1$.

2.4 Loi faible des grands nombres

Théorème 2 Loi faible des grands nombres
Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes équiréparties, d'espérance μ.
On note $S_n = \sum_{k=1}^{n} X_k$. Alors pour tout réel $\varepsilon > 0$:

$$\lim_{n\to+\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0$$

Remarque 7
- Autrement dit, la suite des moyennes empiriques $\left(\frac{1}{n}\sum_{k=1}^{n} X_k\right)$ converge *en probabilité* vers $\mathbb{E}(X)$.
- On a plus précisément : $\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\mathrm{V}(X_1)}{n\varepsilon^2}$.

Application :
On admet que pour toute suite dénombrable d'expériences aléatoires associées aux espaces probabilisés $(\Omega_n, \mathscr{P}(\Omega_n), \mathbb{P}_n)$ il existe une tribu $\mathscr{T}$ et une probabilité $\mathbb{P} : \mathscr{T} \to [0,1]$ telles que :
pour toute suite $(A_n)_{n\in\mathbb{N}^*}$ telle qu'il existe $n_0 \geq 1$ tel que $\forall i \in [\![1, n_0]\!], A_i \in \mathscr{P}(\Omega_i)$ et $\forall j > n_0, A_j = \Omega_j$
on a : $\bigcap_{n\in\mathbb{N}^*} A_n \in \mathscr{T}$ et $\mathbb{P}\left(\bigcap_{n\in\mathbb{N}^*} A_n\right) = \prod_{n=1}^{n_0} \mathbb{P}_n(A_n)$.
On dit alors que les expériences aléatoires sont *indépendantes*.
On considère une suite d'expériences aléatoires identiques (c'est-à-dire que les espaces probabilisés sont identiques) et indépendantes, et un événement A de probabilité p dans chaque espace probabilisé.
On note X_n la variable de Bernoulli égale à 1 si A se réalise à la n-ème expérience.
Les variables aléatoires $(X_n)_{n\in\mathbb{N}^*}$ sont indépendantes, équiréparties, d'espérance p.
La fréquence, notée F_n, de réalisation de l'événement A au cours des n premières expériences est égale à la variable aléatoire $\frac{S_n}{n}$.
D'après la loi faible des grands nombres, $\forall \varepsilon > 0$, $\lim_{n\to+\infty} \mathbb{P}\left(|F_n - p| \geq \varepsilon\right) = 0$.
Ainsi, la fréquence tend en probabilité vers p.

3 Fonctions génératrices

Définition 12
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. On définit la *fonction génératrice* G_X de X par :

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

pour les valeurs de t assurant la convergence de la série entière.

Proposition 19
Le rayon de convergence de la série entière est supérieur ou égal à 1.

Proposition 20
Soit X une variable aléatoire à valeurs dans $\mathbb{N}$.
- X admet une espérance finie si, et seulement si G_X est dérivable en 1 et dans ce cas :

$$\mathbb{E}(X) = G'_X(1)$$

- X admet un moment d'ordre 2 si, et seulement si G_X est deux fois dérivable en 1 et dans ce cas :

$$\mathrm{V}(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$$

Proposition 21
Soient X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$. Si X et Y sont indépendantes, alors

$$G_{X+Y} = G_X\, G_Y$$

4 Lois usuelles

4.1 Loi géométrique

On répète des épreuves de Bernoulli identiques et mutuellement indépendantes, jusqu'à l'apparition du premier succès. On étudie la variable aléatoire X qui donne le rang du premier succès.

Définition 13
On dit qu'une variable aléatoire X suit une *loi géométrique* de paramètre $p \in]0, 1[$ si :

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

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

Proposition 22
La loi géométrique est la seule loi discrète à valeurs dans $\mathbb{N}^*$ vérifiant :

$$\forall (n, k) \in \mathbb{N}^2, \mathbb{P}(X > n + k | X > k) = \mathbb{P}(X > n)$$

On dit qu'une variable aléatoire suivant une loi géométrique est *sans mémoire*.

Proposition 23
Soit $X \hookrightarrow \mathscr{G}(p)$. Alors :
- La fonction génératrice G_X a un rayon de convergence R égal à $\dfrac{1}{1-p}$, et

$$\forall t \in] - R, R[, \qquad G_X(t) = \frac{pt}{1 - (1 - p)t}$$

- $\mathbb{E}(X) = \dfrac{1}{p}$ et $\mathrm{V}(X) = \dfrac{1 - p}{p^2}$.

4.2 Loi de Poisson

Définition 14
On dit qu'une variable aléatoire X suit une *loi de Poisson* de paramètre $\lambda > 0$ si :

1. $X(\Omega) = \mathbb{N}$

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

On note : $X \hookrightarrow \mathscr{P}(\lambda)$.

Proposition 24
Soit $X \hookrightarrow \mathscr{P}(\lambda)$. Alors :

- La fonction génératrice G_X a un rayon de convergence égal à $+\infty$, et

$$\forall t \in \mathbb{R}, \qquad G_X(t) = \mathrm{e}^{\lambda(t-1)}$$

- $\mathbb{E}(X) = \lambda$ et $\mathrm{V}(X) = \lambda$

Proposition 25
Soient $X \hookrightarrow \mathscr{P}(\lambda)$ et $Y \hookrightarrow \mathscr{P}(\mu)$. Si X et Y sont indépendantes, alors $X + Y \hookrightarrow \mathscr{P}(\lambda + \mu)$.

Proposition 26 Approximation de la loi binomiale par une loi de Poisson
Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telle que pour tout entier n, $X_n \hookrightarrow \mathscr{B}(n, p_n)$.
Si $\lim\limits_{n\to+\infty} np_n = \lambda > 0$, alors pour tout $k \in \mathbb{N}$ on a :

$$\lim_{n\to+\infty} \mathbb{P}(X_n = k) = \mathrm{e}^{-\lambda} \frac{\lambda^k}{k!}$$

.

Remarque 8

- Cela signifie que la loi de Poisson modélise la loi du nombre de succès se réalisant lors d'un très grand nombre d'expériences, la probabilité d'un succès étant très faible.
  La loi de Poisson est souvent appelée *loi des événements rares.*