
Les variables aléatoires réelles discrètes

EL AMDAOUI Mustapha,

Lycée IBN TIMIYA,
site web: www.elamdaoui.com,
email: elamdaoui@gmail.com

Niveau: MP

Table des matières

I Variables aléatoires discrètes	2
I.1 Variable aléatoire discrète	2
I.2 Lois usuelles	4
I.3 Image d'une variable aléatoire discrète	7
I.4 Vecteurs de variables aléatoires discrètes	8
I.5 Couples de variables indépendantes	10
II Moments d'une variable aléatoire discrète	13
II.1 Espérance	13
II.2 Moments	18
II.3 Variance et écart-type	18
II.4 Covariance et coefficient de corrélation linéaire	20
II.5 Loi faible des grands nombres	23
III Fonctions génératrices	23

I Variables aléatoires discrètes

(Ω, T, P) est un espace probabilisé.

I.1 Variable aléatoire discrète

Définition 1

Soit E un ensemble. On appelle **variable aléatoire discrète** définie sur (Ω, T) à valeurs dans E , toute application X de Ω dans E telle que :

1. $X(\Omega)$ est au plus dénombrable ;
2. Pour tout $x \in E$, $X^{-1}(\{x\}) \in T$

Si de plus $E \subset \mathbb{R}$, la variable est dite réelle

Remarque :

Si $T = \mathcal{P}(\Omega)$ (en particulier si Ω est au plus dénombrable), toute application de Ω dans $\mathbb{R}$ est une variable aléatoire

Exemple

Soit $A \in T$, alors l'application

$$1_A : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$$

est une variable aléatoire réelle

Définition 2

Soit $X : \Omega \rightarrow E$ une variable aléatoire discrète. Pour tout $x \in E$, on note $(X = x)$ ou $[X = x]$ l'événement

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

Pour tout $A \subset E$, on note $(X \in A)$ ou $[X \in A]$ l'événement

$$[X \in A] = \{\omega \in \Omega, X(\omega) \in A\}$$

Remarque :

La notation $(X \in A)$ est compatible avec les opérations ensemblistes

$$\begin{aligned} (X \in A) \cap (X \in B) &= (X \in A \cap B) \\ (X \in A) \cup (X \in B) &= (X \in A \cup B) \\ \overline{(X \in A)} &= \overline{(X \in A)} \\ (X \in A) \setminus (X \in B) &= (X \in A \setminus B) \end{aligned}$$

Définition 3

Si X est de plus est réelle. Pour $x, a, b \in \mathbb{R}$, on introduit les événements

- $[X \geq x] = X^{-1}([x, +\infty[) = \{\omega \in \Omega, X(\omega) \geq x\}$
- $[X \leq x] = X^{-1}(]-\infty, x]) = \{\omega \in \Omega, X(\omega) \leq x\}$
- $[X > x] = X^{-1}([x, +\infty[) = \{\omega \in \Omega, X(\omega) > x\}$
- $[X < x] = X^{-1}([x, +\infty[) = \{\omega \in \Omega, X(\omega) < x\}$
- $[a \leq X \leq b] = X^{-1}([a, b]) = \{\omega \in \Omega, a \leq X(\omega) \leq b\}$

Propriété 1

[Loi d'une variable aléatoire discrète] Soit X une variable aléatoire discrète définie sur (Ω, T) à valeurs dans un ensemble E . Alors

1. Pour tout A sous-ensemble de E , $[X \in A] \in T$ et

$$P_X(A) = P(X \in A) = \sum_{x \in X(\Omega) \cap A} P(X = x)$$

2. L'application

$$\mathbb{P}_X : \begin{cases} \mathcal{P}(X(\Omega)) & \longrightarrow [0, 1] \\ A & \longmapsto \mathbb{P}(X \in A) \end{cases}$$

est une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$. On l'appelle la loi de probabilité de la variable aléatoire X ou probabilité image de $\mathbb{P}$ par X

Preuve:

1. $X(\Omega)$ étant au plus dénombrable,

$$[X \in A] = \bigcup_{x \in X(\Omega) \cap A} [X = x]$$

est une réunion au plus dénombrable d'événements.

2. $\mathbb{P}_X$ est bien définie

► $\mathbb{P}_X(E) = \mathbb{P}(X \in E) = \mathbb{P}(\Omega) = 1$.

► Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements deux à deux incompatibles de $X(\Omega)$. Les événements $(X \in A_n)$ sont deux à deux disjoints et

$$\bigcup_{n \geq 0} (X \in A_n) = \left[X \in \bigcup_{n \geq 0} A_n \right]$$

On en déduit

$$P_X \left[\bigcup_{n \geq 0} A_n \right] = \sum_{n=0}^{+\infty} P(X \in A_n) = \sum_{n=0}^{+\infty} P_X(A_n)$$

Remarque :

[Importante] Si X est discrète, alors pour tout $A \subset \mathbb{R}$, on a :

$$P_X(A) = P(X \in A) = \sum_{x \in X(\Omega) \cap A} P(X = x)$$

Propriété 2

Soit X une variable aléatoire discrète. Alors

1. La loi $\mathbb{P}_X$ est déterminée par les couples $(x, \mathbb{P}(X = x))_{x \in X(\Omega)}$
2. La famille $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements, appelé le système complet d'événements associé X
3. La famille $(\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable de somme 1

Preuve:

1. Pour tout $A \subset \mathbb{R}$, on a :

$$P_X(A) = P(X \in A) = \sum_{x \in X(\Omega) \cap A} P(X = x)$$

2. $\Omega = \bigcup_{x \in X(\Omega)} [X = x]$ union au plus dénombrable d'événements deux à deux incompatibles

3. Par σ -additivité

Définition 4

Soit X et Y deux variables aléatoires discrètes sur Ω prenant les mêmes valeurs.

- Si $P_X = P_Y$, on dit que X et Y suivent la même loi et l'on note $X \sim Y$
- Si la variable X suit une loi usuellement notée L , on écrit $X \rightsquigarrow L$

Théorème 1: Fonction de répartition

Soit X une variable aléatoire discrète réelle. On appelle **fonction de répartition de X** l'application

$$\mathbb{F}_X : \begin{cases} \mathbb{R} & \longrightarrow \mathbb{R} \\ x & \longmapsto \mathbb{P}_X([-\infty, x]) = \mathbb{P}(X \leq x) \end{cases}$$

On a :

1. $\forall x \in \mathbb{R}, F_X(x) \in [0, 1]$
2. F_X est croissante.
3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.
4. Pour tous $a, b \in \mathbb{R}$ tels que $a < b$, $P(a < X \leq b) = F_X(b) - F_X(a)$;
5. F_X est continue à droite en tout point de $\mathbb{R}$
6. F_X est continue à gauche en $x \in \mathbb{R}$ si, et seulement, si $P(X = x) = 0$

Preuve:

1. Proviens de la définition d'une probabilité.
2. Soit $x \leq y$. Alors on a $[X \leq x] \subset [X \leq y]$ et donc $P(X \leq x) \leq P(X \leq y)$, c'est-à-dire $F_X(x) \leq F_X(y)$.
3. Comme F_X est croissante, elle admet des limites en $-\infty$ et $+\infty$. Alors

$$\lim_{x \rightarrow -\infty} F_X(x) = \lim_{n \rightarrow +\infty} F_X(-n) = \lim_{n \rightarrow +\infty} P(X \leq -n)$$

Posons $B_n = [X \leq -n]$, alors $\bigcap_{n \in \mathbb{N}^*} B_n = \emptyset$, car si $\omega \in \bigcap_{n \in \mathbb{N}^*} B_n$, alors pour tout $n \in \mathbb{N}$ on a $X(\omega) \leq -n$, donc $X(\omega) = -\infty$, ce qui est impossible puisque X ne prend que des valeurs finies. D'après la limite monotone

$$\begin{aligned} \lim_{n \rightarrow +\infty} P(X \leq -n) &= \lim_{n \rightarrow +\infty} P(B_n) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{k=1}^n B_k\right) \\ &= P\left(\bigcap_{n \in \mathbb{N}^*} B_n\right) = P(\emptyset) = 0 \end{aligned}$$

4. $\forall (a, b) \in \mathbb{R}^2$, $[a < X \leq b] = [X \leq b] \setminus [X \leq a]$ donc grâce aux propriétés des probabilités,

$$P([a < X \leq b]) = F_X(b) - F_X(a)$$

5. Pour démontrer la continuité à droite, il suffit de poser

$$A_n = \left[x < X \leq x + \frac{1}{n+1} \right]$$

de probabilité $P(A_n) = F\left(x + \frac{1}{n+1}\right) - F_X(x)$. La suite $(A_n)_{n \in \mathbb{N}}$ est croissante telle que $\bigcap_{n=0}^{+\infty} A_n = \emptyset$ et d'après la continuité monotone $\lim_{n \rightarrow +\infty} P(A_n) = 0$. Il en résulte que la suite de terme $F_X\left(x + \frac{1}{n+1}\right)$ converge vers $F_X(x)$. Comme l'on sait que $F_X(x^+) = \lim_{h \rightarrow 0^+} F_X(x+h)$ existe par suite de la croissance de F , il en résulte que $F_X(x^+) = F_X(x)$.

6. On veut montrer que $P(X = x) = F_X(x) - \lim_{x \rightarrow 0^-} F_X(x)$, on pose alors

$$A_n = \left[x - \frac{1}{n+1} < X < x \right]$$

de probabilité $F(x) - P(X = x) = F\left(x - \frac{1}{n+1}\right)$. La suite $(A_n)_{n \in \mathbb{N}}$ est croissante telle que $\bigcap_{n=0}^{+\infty} A_n = \emptyset$ et d'après la continuité monotone $\lim_{n \rightarrow +\infty} P(A_n) = 0$. Il en résulte que $F(x) - P(X = x) = F\left(x - \frac{1}{n+1}\right) \xrightarrow{n \rightarrow +\infty} 0$. Comme l'on sait que $F(x^-) = \lim_{h \rightarrow 0^+} F(x-h)$ existe par suite de la croissance de F , il en résulte que $P(X = x) = F(x) - F(x^-)$

I.2 Lois usuelles**Définition 5: Loi uniforme**

Soit $n \in \mathbb{N}^*$. On dit que X suit la **loi uniforme** $\mathcal{U}(\llbracket 1; n \rrbracket)$ si :

$$X(\Omega) = \llbracket 1; n \rrbracket \text{ et } \forall k \in \llbracket 1; n \rrbracket, \quad P(X = k) = \frac{1}{n}$$

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

Exemple

1. Un dé parfaitement équilibré est lancé. La variable aléatoire X égale au numéro de la face supérieure suit la loi uniforme sur $\llbracket 1, 6 \rrbracket$.

2. Une urne contient n boules numérotée. On en tire une au hasard. La variable X égale au numéro de la boule tirée suit une loi uniforme sur $\llbracket 1, n \rrbracket$.

Définition 6: Loi de Bernoulli

Soit $p \in [0; 1]$. On dit qu'une variable aléatoire X suit la **loi de Bernoulli de paramètre p** si :

$$X(\Omega) = \{0; 1\}$$

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

On note $X \hookrightarrow \mathcal{B}(p)$ ou $X \hookrightarrow \mathcal{B}(1, p)$

Exemple: Modèle probabiliste

On considère un espace probabilisé $(\Omega, T, \mathbb{P})$ et l'on s'intéresse à un événement S appelé succès. On définit la variable

$$X : \omega \longmapsto \begin{cases} 1 & \text{si } \omega \in S \\ 0 & \text{sinon} \end{cases}$$

X est l'indicatrice du succès. Si l'on suppose $p = \mathbb{P}(S)$, alors X suit la loi de Bernoulli du paramètre p

Exemple

La variable indicatrice d'un événement A tel que $0 < P(A) < 1$ est une variable de Bernoulli. Réciproquement, toute variable de Bernoulli est la variable indicatrice de l'événement $\mathbb{P}(X = 1)$.

Définition 7: Loi binomiale

Soit $p \in [0; 1]$ et $n \in \mathbb{N}$. On dit que la variable aléatoire X suit la **loi binomiale de taille n et de paramètre p (notée $\mathcal{B}(n, p)$)** si :

$$X(\Omega) = \llbracket 0; n \rrbracket \quad \text{et} \quad \forall k \in \llbracket 0; n \rrbracket \quad P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

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

Exemple

Une urne contient des boules blanches en proportion p et des boules noires en proportion $1 - p$. On fait n tirages avec remise dans l'urne. La variable égale au nombre de boules blanches obtenues suit une loi binomiale de paramètre (n, p) .

Définition 8: Loi géométrique

Soit $p \in]0; 1[$. On dit qu'une variable aléatoire réelle discrète X suit la **loi géométrique de paramètre p (notée $\mathcal{G}(p)$)** si :

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

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

Exemple

On lance successivement un dé équilibré jusqu'à obtention d'un six. On pose X le nombre de lancers nécessaires. On a

$$\mathbb{P}(X = n) = \left(\frac{5}{6}\right)^{n-1} \frac{1}{6}$$

et donc $X \hookrightarrow \mathcal{G}\left(\frac{1}{6}\right)$

Propriété 3

Si la variable aléatoire X suit la loi géométrique de paramètre $p \in]0, 1[$ à valeurs dans $\mathbb{N}^*$, on a, pour tout entier naturel $k \in \mathbb{N}$.

$$\mathbb{P}(X > k) = q^k$$

Preuve:

Soit $k \in \mathbb{N}$,

$$\mathbb{P}(X > k) = \sum_{n=k+1}^{+\infty} \mathbb{P}(X = n) = \sum_{n=k+1}^{+\infty} p q^{n-1} = p \frac{q^k}{1-q} = q^k$$

Définition 9

Une variable X discrète à valeurs dans $\mathbb{N}^*$ est dite sans mémoire si

1. $\forall n \in \mathbb{N}, \mathbb{P}(X > n) > 0$
2. $\forall n, k \in \mathbb{N}, \mathbb{P}(X > n+k | X > k) = \mathbb{P}(X > n)$

Théorème 2: Loi sans mémoire

Une variable X discrète à valeurs dans $\mathbb{N}^*$. Alors X est sans mémoire si, et seulement si, elle suit une loi géométrique

Preuve:

Soit $q = \mathbb{P}(X > 1)$. La condition imposée donne

$$\mathbb{P}(X > n+1 | X > n) = \mathbb{P}(X > 1) = q$$

Or

$$\mathbb{P}(X > n+1) = \mathbb{P}(X > n+1, X > n) = \mathbb{P}(X > n+1 | X > n) \mathbb{P}(X > n) = q \mathbb{P}(X > n)$$

Par une récurrence immédiate et sachant $\mathbb{P}(X > 0) = 1$, on obtient

$$\forall n \in \mathbb{N}, \mathbb{P}(X > n) = q^n$$

Puis pour tout $n \in \mathbb{N}^*$, on a

$$\mathbb{P}(X = n) = \mathbb{P}(X > n-1) - \mathbb{P}(X > n) = q^{n-1} - q^n = (1-q)q^{n-1}$$

Ainsi, la variable X suit une loi géométrique de paramètre $p = 1 - q$.

Définition 10: Loi de Poisson

Soit $\lambda > 0$. On dit qu'une variable aléatoire réelle discrète X suit une **loi de Poisson** (notée $\mathcal{P}(\lambda)$) si :

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

On note $X \hookrightarrow \mathcal{P}(\lambda)$

Propriété 4: Approximation de la loi de Poisson par la loi binomiale

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de variables aléatoires. On suppose que pour tout $n \in \mathbb{N}$ la variable $X_n \hookrightarrow \mathcal{B}(n, p_n)$ avec $\lim_{n \rightarrow \infty} n p_n = \lambda > 0$. Alors

$$\forall k \in \mathbb{N}, \quad \mathbb{P}(X_n = k) \xrightarrow{n \rightarrow +\infty} \frac{\lambda^k}{k!} e^{-\lambda}$$

Preuve:

Soit $k \in \mathbb{N}$ donc $\forall n \geq 1$:

$$\begin{aligned} \mathbb{P}(X_n = k) &= C_n^k p_n^k (1-p_n)^{n-k} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} \\ &= \frac{1}{k!} (n p_n)^k (1-p_n)^{n-k} \frac{n!}{(n-k)! n^k} \end{aligned}$$

On a :

$$\bullet \quad n p_n \xrightarrow{n \rightarrow +\infty} \lambda \quad \text{donc} \quad (n p_n)^k \xrightarrow{n \rightarrow +\infty} \lambda^k.$$

$$\bullet \quad np_n \xrightarrow{n \rightarrow +\infty} \lambda \text{ donc } p_n \xrightarrow{n \rightarrow +\infty} 0 \text{ donc}$$

$$\begin{aligned} \ln(1-p_n)^{n-k} &= (n-k)\ln(1-p_n) \\ &\sim -(n-k)p_n \\ &\sim -\lambda \end{aligned}$$

$$\text{donc } (1-p_n)^{n-k} \xrightarrow{n \rightarrow +\infty} e^{-\lambda}.$$

$$\bullet \quad \frac{n!}{(n-k)!n^k} = \frac{n(n-1)\cdots(n-(k-1))}{n^k} \sim 1.$$

$$\text{Donc } \lim_{n \rightarrow +\infty} p(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}.$$

I.3 Image d'une variable aléatoire discrète

Soit X une variable aléatoire sur l'espace probabilisé $(\Omega, T, \mathbb{P})$ à valeurs dans un ensemble E .

Définition 11

Si f est une application définie au moins sur $X(\Omega) \subset E$ à valeurs dans un ensemble E' , on note $f(X)$ la variable aléatoire $Y = f \circ X$

$$Y : \Omega \longrightarrow E' \quad \text{avec } Y(\omega) = f(X(\omega))$$

Remarque :

On vérifie qu'il s'agit bien d'une variable aléatoire car

$$\forall y \in Y(\Omega), \quad [Y = y] = \bigcup_{\substack{x \in X(\Omega) \\ f(x) = y}} [X = x] \in T$$

Remarque :

Si la fonction f est une fonction présentant une notation usuelle particulière, on adapte celle-ci à la description de la variable aléatoire $f(X)$. C'est ainsi qu'on pourra écrire $\sqrt{X}$, $|X|$, $\sin(X)$, ...

Théorème 3

La loi de $Y = f(X)$ est entièrement déterminée par celle de X :

$$\forall B \subset Y(\Omega), \quad \mathbb{P}_Y(B) = \mathbb{P}_X(f^{-1}(B))$$

Preuve:

Par définition $\mathbb{P}_Y(B) = \mathbb{P}(Y \in B) = \mathbb{P}(f(X) \in B)$. Or $(f(X) \in B) = (X \in f^{-1}(B))$

Exemple

Si $X \hookrightarrow \mathcal{B}(n, p)$, alors $Y = n - X \hookrightarrow \mathcal{B}(n, 1 - p)$

En effet $Y(\Omega) = \llbracket 0, n \rrbracket$ et pour tout $k \in \llbracket 0, n \rrbracket$, on a :

$$\mathbb{P}(Y = k) = \mathbb{P}(X = n - k) = C_n^{n-k} p^{n-k} (1-p)^k$$

Propriété 5

Soit $X_1, \dots, X_k$ des variables aléatoires réelles définies sur le même espace probabilisé (Ω, T, P) et $f : \mathbb{R}^k \longrightarrow \mathbb{R}$ une fonction continue. Alors

$$f(X_1, \dots, X_k) : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto f(X_1(\omega), \dots, X_k(\omega)) \end{cases}$$

est une variable aléatoire réelle (Ω, T, P)

Preuve:

Conformément au programme ce résultat est admis

Remarque :

Si les variables $X_1, \dots, X_k$ sont toutes discrètes, $f(X_1, \dots, X_k)$ est une variable réelle discrète sans l'hypothèse de la continuité de f

Conséquence 1

Si $X_1, \dots, X_n$ une famille finie de variables aléatoires réelles définies sur le même espace probabilisé (Ω, T, P) , alors $\sum_{i=1}^n X_i$, $\prod_{i=1}^n X_i$, $\min_{1 \leq i \leq n} X_i$ et $\max_{1 \leq i \leq n} X_i$ sont des variables aléatoires réelles

I.4 Vecteurs de variables aléatoires discrètes

Soit X et Y deux variables aléatoires réelles discrètes définies sur un même espace probabilisé (Ω, T, P) à valeurs dans des ensembles E et F respectivement.

Définition 12

On appelle couple défini par les variables aléatoires X et Y la variable $Z = (X, Y) : \Omega \rightarrow E \times F$ déterminée par :

$$\forall \omega \in \Omega, \quad Z(\omega) = (X(\omega), Y(\omega))$$

Remarque :

Il s'agit bien d'une variable aléatoire discrète. En effet $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$ et

$$\forall (x, y) \in Z(\Omega), \quad [Z = (x, y)] = [X = x] \cap [Y = y] \in T$$

Définition 13: Loi conjointe et lois marginales

1. On appelle **loi conjointe du couple** $Z = (X, Y)$, l'application

$$\mathbb{P}_Z : \begin{cases} X(\Omega) \times Y(\Omega) & \longrightarrow [0, 1] \\ (x, y) & \longmapsto \mathbb{P}([X = x] \cap [Y = y]) \end{cases}$$

2. La loi de X est appelé la **première loi marginale de Z** .

3. La loi de Y est appelée la **deuxième loi marginale de Z** .

Notation :

$\mathbb{P}([X = x] \cap [Y = y])$ se note $\mathbb{P}(X = x, Y = y)$

Propriété 6

$(\mathbb{P}(X = x, Y = y))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est une famille de réels positifs sommable de somme 1

Preuve:

Y est une variable aléatoire discrète, donc $([Y = y])_{y \in Y(\Omega)}$ est un système complet d'événements. Par la formule des probabilités totales, pour tout $x \in X(\Omega)$, la famille $(\mathbb{P}(X = x, Y = y))_{y \in Y(\Omega)}$ est sommable et

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

Comme X est une variable aléatoire discrète, donc $([X = x])_{x \in X(\Omega)}$ est un système complet d'événements, alors par σ -additivité la famille $(\mathbb{P}(X = x))_{x \in X(\Omega)}$ est sommable de somme 1. On en déduit, par le critère suffisant de sommabilité, que la famille $(\mathbb{P}(X = x, Y = y))_{(x, y) \in X(\Omega) \times Y(\Omega)}$ est sommable de somme 1.

Corollaire 1

Soit (X, Y) un couple de variables aléatoires discrètes. Alors on a :

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

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

La loi conjointe détermine entièrement ses lois marginales

Preuve:

Soit $x \in X(\Omega)$. La famille $([Y = y])_{y \in Y(\Omega)}$ est un système complet d'événements, alors par la formule des probabilités totales

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

Exemple

Dans une succession de pile ou face pour laquelle la probabilité d'obtenir pile est $p \in]0; 1[$ et la probabilité d'obtenir face est $q = 1 - p$, on note X le rang d'apparition du premier pile et Y le rang d'apparition du deuxième pile.

1. Déterminer la loi du couple (X, Y) ;
2. Déterminer la deuxième loi marginale du couple (X, Y)

1. On a $X(\Omega) = \mathbb{N}^*$ et $Y(\Omega) = \mathbb{N} \setminus \{0, 1\}$. Soit $i \in X(\Omega)$ et $j \in Y(\Omega)$.

- Il est impossible que le deuxième pile arrive avant le premier donc si $i \geq j$ on a

$$\mathbb{P}(X = i, Y = j) = 0$$

- Si $i < j$, on a $[X = i] \cap [Y = j] = F_1 \cap \dots \cap F_{i-1} \cap P_i \cap F_{i+1} \cap \dots \cap F_{j-1} \cap P_j$ et les lancers sont indépendants donc on a

$$P([X = i] \cap [Y = j]) = q^{i-1} \times p \times q^{j-i-1} \times p = q^{j-2} p^2$$

2. On rappelle tout d'abord que $Y(\Omega) = \mathbb{N} \setminus \{0, 1\}$.

Pour tout $j \in Y(\Omega)$, d'après la formule des probabilités totales utilisée avec le système complet d'événements $([X = i])_{i \in X(\Omega)}$, on a :

$$\begin{aligned} P(Y = j) &= \sum_{i=1}^{+\infty} P([X = i] \cap [Y = j]) \\ &= \sum_{i=1}^{j-1} P([X = i] \cap [Y = j]) + \sum_{i=j}^{+\infty} P([X = i] \cap [Y = j]) \\ &= \sum_{i=1}^{j-1} p^2 \times q^{j-2} + \sum_{i=j}^{+\infty} 0 \\ &= p^2 \times q^{j-2} \sum_{i=1}^{j-1} 1 \\ &= (j-1)p^2 \times q^{j-2} \end{aligned}$$

Définition 14: Lois conditionnelles

Soit $y \in Y(\Omega)$ tel que $\mathbb{P}(Y = y) \neq 0$.

On appelle la loi conditionnelle de X sachant $[Y = y]$, l'application

$$\mathbb{P}_{[Y=y]} : \begin{cases} X(\Omega) & \longrightarrow [0, 1] \\ x & \longmapsto \mathbb{P}_{[Y=y]}(X = x) \end{cases}$$

Remarque :

De même pour $x \in X(\Omega)$ tel que $\mathbb{P}(X = x) \neq 0$, on définit loi conditionnelle de Y sachant $[X = x]$

Exemple

Soient X et Y deux variables aléatoires réelles discrètes telles que Y suive la loi de Poisson de paramètre λ , où λ est un réel strictement positif, et que la loi conditionnelle de X sachant $[Y = m]$ soit la loi binomiale de paramètre (m, p) pour tout entier m , où p est un réel de $]0, 1[$.

Donnons la loi conjointe de (X, Y) et la loi de X

Pour tout entier m , par définition de la loi conditionnelle et de la loi binomiale de paramètre (m, p) . Par définition

$$\forall k \in \llbracket 0, m \rrbracket, \quad P_{[Y=m]}(X = k) = \frac{P(X = k, Y = m)}{P(Y = m)}$$

Donc, pour tout entier m et pour tout entier k de $\llbracket 0, m \rrbracket$

$$P(X = k, Y = m) = P_{[Y=m]}(X = k) \times P(Y = m) = C_m^k p^k (1-p)^{m-k} \frac{\lambda^m}{m!} e^{-\lambda}$$

On obtient donc ainsi la loi conjointe

$$P(X = k, Y = m) = \begin{cases} C_m^k p^k (1-p)^{m-k} \frac{\lambda^m}{m!} e^{-\lambda} & \text{si } k \in \llbracket 0, m \rrbracket \\ 0 & \text{sinon} \end{cases}$$

Comme $([Y = m])_{m \in \mathbb{N}}$ est un système complet d'événements, pour tout entier k , on a

$$\begin{aligned} P(X = k) &= \sum_{m=0}^{+\infty} P_{[Y=m]}(X = k) P(Y = m) \\ &= \sum_{m=k}^{+\infty} C_m^k p^k (1-p)^{m-k} \frac{\lambda^m}{m!} e^{-\lambda} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} \sum_{m=k}^{+\infty} \frac{(1-p)^{m-k} \lambda^{m-k}}{(m-k)!} \\ &= e^{-\lambda} \frac{p^k \lambda^k}{k!} e^{(1-p)\lambda} = \frac{(p\lambda)^k}{k!} e^{-p\lambda} \end{aligned}$$

On en déduit que X suit la loi de poisson de paramètre λp

Définition 15: Lois conditionnelles

Soit $B \subset Y(\Omega)$ tel que $\mathbb{P}(Y \in B) \neq 0$.

On appelle **la loi conditionnelle de X sachant $[Y \in B]$** , l'application

$$\mathbb{P}_{[Y \in B]} : \begin{cases} X(\Omega) & \longrightarrow [0, 1] \\ x & \longmapsto \mathbb{P}_{[Y \in B]}(X = x) \end{cases}$$

Remarque :

De même pour $A \subset X(\Omega)$ tel que $\mathbb{P}(X \in A) \neq 0$, on définit **loi conditionnelle de Y sachant $[X \in A]$**

Définition 16: Vecteurs de variables aléatoires

Soit $X_1, \dots, X_n$ des variables aléatoires discrètes sur l'espace probabilisé (Ω, T, P) .

On appelle **vecteur aléatoire discret** défini à partir des variables aléatoires $X_1, \dots, X_n$ la variable aléatoire discrète Z donnée par

$$\forall \omega \in \Omega, \quad Z(\omega) = (X_1(\omega), \dots, X_n(\omega))$$

La loi de la variable Z est appelée **loi conjointe** des variables $X_1, \dots, X_n$ tandis que les lois de $X_1, \dots, X_n$ sont les **lois marginales** de Z .

I.5 Couples de variables indépendantes

Définition 17

On dit que deux variables aléatoires discrètes X et Y sont **indépendantes** si :

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x) \mathbb{P}(Y = y)$$

Théorème 4

Soit X et Y deux variables aléatoires discrètes. Les assertions suivantes sont équivalentes

1. X et Y sont indépendantes
2. Pour tous $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, les événements $[X \in A]$ et $[Y \in B]$ sont indépendants. Autrement-dit pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, on a

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A) \mathbb{P}(Y \in B)$$

Preuve:

$2 \Rightarrow 1$) Soit $(x, y) \in X(\Omega) \times Y(\Omega)$, on prend $A = \{x\}$ et $B = \{y\}$. Les événements $(X = x)$ et $(Y = y)$ sont indépendants donc

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

$2 \Leftarrow 1$) Soit $A \subset X(\Omega)$ et $B \subset Y(\Omega)$, alors $A \times B$ est au plus dénombrable. Par la formule des probabilités totales

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{(x,y) \in A \times B} P(X = x, Y = y) \\ &= \sum_{(x,y) \in A \times B} P(X = x) \times P(Y = y) \end{aligned}$$

En sommant par paquets

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A} \sum_{y \in B} P(X = x) \times P(Y = y) \\ &= \sum_{x \in A} P(X = x) \times \sum_{y \in B} P(Y = y) \\ &= P(X \in A) \times P(Y \in B) \end{aligned}$$

Corollaire 2

Soit X et Y deux variables aléatoires réelles discrètes. Les assertions suivantes sont équivalentes

1. X et Y sont indépendantes
2. Pour tous $x, y \in \mathbb{R}$, on a :

$$\mathbb{P}([X \leq x] \cap [Y \leq y]) = \mathbb{P}(X \leq x) \times \mathbb{P}(Y \leq y)$$

Exemple

- Une urne contient n jetons numérotés de 1 à n . On en tire deux avec remise. Soit X et Y les variables aléatoires égales au premier et au second numéro tiré. On a pour tout $1 \leq i, j \leq n$,

$$\mathbb{P}(X = i, Y = j) = \frac{1}{n^2} = \mathbb{P}(X = i)\mathbb{P}(Y = j)$$

donc les variables X et Y sont indépendantes.

- On effectue la même expérience mais sans remise. On a alors $\mathbb{P}(X = i, Y = i) = 0$ mais $P(X = i)P(Y = i) = \frac{1}{n^2}$ donc les variables ne sont pas indépendantes.

Théorème 5: de Coalitions

Si X et Y sont deux variables aléatoires discrètes indépendantes et si f et g sont deux fonctions numériques définies respectivement sur $X(\Omega)$ et $Y(\Omega)$ alors $f(X)$ et $g(Y)$ sont indépendantes.

Preuve:

Soit $x \in f(X(\Omega))$ et $y \in g(Y(\Omega))$. On a

$$P(f(X) = x, g(Y) = y) = P\left(\left[X \in f^{-1}(\{x\})\right] \cap \left[Y \in g^{-1}(\{y\})\right]\right)$$

Les variables X et Y étant indépendantes, donc

$$\begin{aligned} P(f(X) = x, g(Y) = y) &= P\left(\left[X \in f^{-1}(\{x\})\right] \cap \left[Y \in g^{-1}(\{y\})\right]\right) \\ &= P\left(X \in f^{-1}(\{x\})\right) \times P\left(Y \in g^{-1}(\{y\})\right) \\ &= P(f(X) = x) \times P(g(Y) = y) \end{aligned}$$

Définition 18

Soit $(X_i)_{1 \leq i \leq n}$ une famille de n variables aléatoires discrètes sur l'espace probabilisé $(\Omega, T, \mathbb{P})$.

1. On dit que celles-ci sont deux à deux indépendantes si pour tous entiers $i, j \in \llbracket 1, n \rrbracket$ tels que $i \neq j$ les variables X_i et X_j sont indépendantes
2. On dit que celles-ci sont mutuellement indépendantes, si

$$\forall (x_1, \dots, x_n) \in \prod_{i=1}^n X_i(\Omega), \mathbb{P}\left(\bigcap_{i=1}^n [X_i = x_i]\right) = \prod_{i=1}^n \mathbb{P}(X_i = x_i)$$

Propriété 7

Soit $X_1, \dots, X_n$ des variables aléatoires réelles discrètes. Les assertions suivantes sont équivalentes

1. Les variables aléatoires $X_1, \dots, X_n$ sont mutuellement indépendantes
2. Pour toute famille $(A_i)_{i \in \llbracket 1, n \rrbracket}$ avec $A_i \subset X_i(\Omega)$, les événements $([X_i \in A_i])$ sont mutuellement indépendants

Lemme 1 (des coalitions ou indépendance héritée)

Si $X_1, \dots, X_n$ sont des variables mutuellement indépendantes, alors pour tout m compris entre 1 et $n-1$, et toutes fonction f et g , les variables $X = f_1(X_1, \dots, X_m)$ et $Y = g(X_{m+1}, \dots, X_n)$ sont indépendantes

Preuve:

Non exigible.

Remarque :

Cela signifie que pour tous $k \in \mathbb{N}^*$ et $j_1, \dots, j_k$ distincts dans J

$$P(X_{j_1} \in I_{j_1}, \dots, X_{j_k} \in I_{j_k}) = \prod_{i=1}^k P(X_{j_i} \in I_{j_i})$$

Exemple

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables indépendantes, alors pour tout $n \in \mathbb{N}^*$, les deux variables $S_n = \sum_{k=1}^n X_k$ et X_{n+1} sont indépendantes

Corollaire 3

Si $X_1, \dots, X_r$ sont mutuellement indépendantes chacune de loi $\mathcal{B}(n_i, p)$, alors

$$\sum_{i=1}^r X_i \sim \mathcal{B}\left(\sum_{i=1}^r n_i, p\right)$$

Corollaire 4

Si $X_1, \dots, X_r$ sont mutuellement indépendantes qui suivent la loi de Bernoulli $\mathcal{B}(p)$, alors

$$\sum_{i=1}^r X_i \sim \mathcal{B}(r, p)$$

Propriété 8

Si $X_1, \dots, X_r$ sont mutuellement indépendantes chacune de loi $\mathcal{P}(\lambda_i)$, alors

$$\sum_{i=1}^r X_i \sim \mathcal{P}\left(\sum_{i=1}^r \lambda_i\right)$$

Définition 19

On dit que les variables aléatoires de la famille $(X_i)_{i \in I}$ sont mutuellement indépendantes si toutes ses sous-familles finies sont mutuellement indépendantes.

Théorème 6

Soit $\mathcal{L}$ la loi d'une certaine variable aléatoire discrète.

Il existe un espace probabilisé $(\Omega, T, \mathbb{P})$ sur lequel existe une suite $(X_n)_{n \in \mathbb{N}}$ de variables aléatoires discrètes mutuellement indépendantes et qui sont toutes de loi $\mathcal{L}$

II Moments d'une variable aléatoire discrète

Les variables aléatoires discrètes sont réelles

II.1 Espérance

Définition 20: Espérance d'une variable aléatoire discrète

On dit que la variable discrète X **admet une espérance**, ou que **l'espérance de X existe**, lorsque la famille $(xP(X=x))_{x \in X(\Omega)}$ est sommable. On appelle alors **espérance de X** , le réel

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

Remarque :

1. Toute variable bornée admet une espérance
2. Si X est finie, alors X admet une espérance
3. Si l'ensemble $X(\Omega)$ est infini (nécessairement dénombrable), alors en introduisant $(x_n)_{n \in \mathbb{N}}$ une énumération de celui-ci, la variable X admet une espérance si, et seulement si, il y a convergence absolue de la série $\sum_{n \geq 0} x_n P(X=x_n)$. Auquel cas, on a alors

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

La valeur obtenue ne dépend pas de l'énumération choisie.

Propriété 9: Indicatrice

Soit $A \in \mathcal{T}$, alors 1_A admet une espérance et $\mathbb{E}(1_A) = \mathbb{P}(A)$

Preuve:

Rappelons que la variable aléatoire indicatrice de l'événement A est la variable aléatoire

$$1_A : \begin{cases} \Omega & \longrightarrow \mathbb{R} \\ \omega & \longmapsto \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{sinon} \end{cases} \end{cases}$$

1_A est finie, donc elle admet une espérance et

$$\mathbb{E}(1_A) = 0\mathbb{P}(\overline{A}) + 1\mathbb{P}(A) = \mathbb{P}(A)$$

Propriété 10

Pour tout $a \in \mathbb{R}$, la variable aléatoire réelle discrète certaine égale à a admet a comme espérance.

Preuve:

Soit X la variable aléatoire réelle discrète certaine égale à a . On a $X(\Omega) = \{a\}$ et $\mathbb{P}(X=a) = 1$, donc A admet une espérance et

$$\mathbb{E}(X) = a\mathbb{P}(X=a) = a\mathbb{P}(\Omega) = a$$

Propriété 11

Soit X une variable aléatoire discrète

1. Si $X \hookrightarrow \mathcal{B}(p)$, alors $\mathbb{E}(X) = p$
2. Si $X \hookrightarrow \mathcal{B}(n, p)$, alors $\mathbb{E}(X) = np$
3. Si $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$, alors $\mathbb{E}(X) = \frac{n+1}{2}$
4. Si $X \hookrightarrow \mathcal{G}(p)$, alors $\mathbb{E}(X) = \frac{1}{p}$
5. Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors $\mathbb{E}(X) = \lambda$

Propriété 12: Critère de comparaison

Soit X et Y deux variables aléatoires discrètes telles que Y admet une espérance et $|X| \leq Y$. Alors X admet une espérance et

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|) \leq \mathbb{E}(Y)$$

Preuve:

La famille $(yP(X = x, Y = y))_{(x,y) \in |X|(\Omega) \times Y(\Omega)}$ est sommable, car il s'agit d'une famille de réels positifs

- Pour $y \in Y(\Omega)$ et la famille $(|X| = x)_{x \in X(\Omega)}$ est un système complet d'événements, alors par la formules des probabilités totales la famille $(yP(X = x, Y = y))_{x \in X(\Omega)}$ est sommable de somme $yP(Y = y)$
- La famille $(yP(Y = y))_{y \in Y(\Omega)}$ est sommable car Y admet une espérance

Or $|X| \leq Y$ donc

$$|x|P(X = x, Y = y) \leq yP(X = x, Y = y)$$

En effet, si le terme de probabilité est nul, l'inégalité est vraie, sinon il existe $\omega \in \Omega$ tel que $|X(\omega)| = x$ et $Y(\omega) = y$ donc $|x| \leq y$ et l'inégalité est encore vraie. Par le critère de comparaison la famille $(xP(X = x, Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable. En particulier

- Pour tout $x \in X(\Omega)$ la famille $(xP(X = x, Y = y))_{y \in Y(\Omega)}$ est sommable
- la famille $\left(\sum_{y \in Y(\Omega)} |x|P(X = x, Y = y) \right)_{x \in X(\Omega)} = (|x|P(X = x))_{x \in X(\Omega)}$ est sommable

Donc X admet une espérance et par le théorème de la sommation par paquets

$$|\mathbb{E}(X)| \leq \mathbb{E}(|X|) = \sum_{y \in Y(\Omega)} \sum_{x \in |X|(\Omega)} yP(|X| = x, Y = y)$$

et par probabilité totales

$$\mathbb{E}(|X|) \leq \sum_{y \in Y(\Omega)} yP(Y = y) = \mathbb{E}(Y) < +\infty$$

Théorème 7: Positivité de l'espérance

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, T, P) admettant une espérance

1. Si $X(\Omega) \subset \mathbb{R}^+$, alors $\mathbb{E}(X) \geq 0$
2. Si de plus $\mathbb{E}(X) = 0$, alors $\mathbb{P}(X = 0) = 1$

Preuve:

Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{R}^+$.

1. On a

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \geq 0$$

Comme somme des termes positifs

2. Si $\mathbb{E}(X) = 0$, alors pour tout $x \in X(\Omega) \setminus \{0\}$, on a $\mathbb{P}(X = x) = 0$, donc $\mathbb{P}(X = 0) = 1$

Théorème 8: Inégalité de Markov

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, T, P) telle que $X(\Omega) \subset \mathbb{R}^+$ et possédant une espérance, alors

$$\forall \lambda > 0 \quad \mathbb{P}(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}$$

Preuve:

Soit X une variable aléatoire telle que $X(\Omega) \subset \mathbb{R}^+$.

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x\mathbb{P}(X = x) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} x\mathbb{P}(X = x) \geq \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} \lambda\mathbb{P}(X = x).$$

Or $\mathbb{P}(X \geq \lambda) = \sum_{\substack{x \in X(\Omega) \\ x \geq \lambda}} \mathbb{P}(X = x)$, d'où le résultat.

Corollaire 5 (Inégalité de Markov)

Soit X une variable aléatoire réelle discrète définie sur un espace probabilisé (Ω, T, P) possédant une

espérance, alors

$$\forall \lambda > 0 \quad \mathbb{P}(|X| \geq \lambda) \leq \frac{\mathbb{E}(|X|)}{\lambda}$$

Théorème 9: du transfert

Soit X une variable aléatoire discrète et g une fonction définie au moins sur $X(\Omega)$ et à valeurs dans $\mathbb{R}$. On a équivalence entre :

1. La variable $g(X)$ admet une espérance ;
2. la famille $(g(x)P(X=x))_{x \in X(\Omega)}$ est sommable.

Auquel cas

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

Preuve:

Posons $Y = g(X)$.

1) $\Rightarrow$ 2) Supposons que Y admet une espérance.

- Soit $y \in Y(\Omega)$, la famille $(g(x)P([Y=y] \cap [X=x]))_{x \in X(\Omega)}$ est sommable. On a l'égalité

$$yP([Y=y] \cap [X=x]) = g(x)P([Y=y] \cap [X=x])$$

car l'égalité est vraie quand la probabilité est nulle, mais aussi quand elle est non nulle car il existe un événement ω vérifiant $g(X(\omega)) = y$ et $X(\omega) = x$ donc $g(x) = y$. Or la famille $(yP([Y=y] \cap [X=x]))_{x \in X(\Omega)}$ est sommable, par la formule des probabilités totales, et

$$\begin{aligned} \sum_{x \in X(\Omega)} |g(x)|P([Y=y] \cap [X=x]) &= \sum_{x \in X(\Omega)} |y|P([Y=y] \cap [X=x]) \\ &= |y|P([Y=y]) \end{aligned}$$

- La famille $(|y|P([Y=y]))_{y \in Y(\Omega)}$ est sommable car Y admet une espérance

Donc $(g(x)P([Y=y] \cap [X=x]))_{(x,y) \in X(\Omega) \times Y(\Omega)}$ est sommable. En particulier

- Pour tout $x \in X(\Omega)$, la famille $(|g(x)|P([Y=y] \cap [X=x]))_{y \in Y(\Omega)}$ est sommable de somme, par la formule des probabilités totales, $|g(x)|P([X=x])$
- La famille $(g(x)P([X=x]))_{x \in X(\Omega)}$ est sommable

2) $\Rightarrow$ 1) Même argumentation.

Supposons que Y admet une espérance et montrons l'égalité :

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

La variable $g(X)$ admet une espérance, alors la famille $(yP(g(X)=y))_{y \in g(X)(\Omega)}$ est sommable et

$$\mathbb{E}(g(X)) = \sum_{y \in g(X)(\Omega)} yP(g(X)=y)$$

Par probabilités totales

$$\mathbb{E}(g(X)) = \sum_{y \in g(X)(\Omega)} \sum_{x \in X(\Omega)} yP([g(X)=y] \cap [X=x])$$

Or

$$yP([g(X)=y] \cap [X=x]) = g(x)P([g(X)=y] \cap [X=x])$$

En permutant les sommes

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} \sum_{y \in g(X)(\Omega)} g(x)P([g(X)=y] \cap [X=x])$$

Par probabilités totales

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

Corollaire 6 (Cas de $X(\Omega) = \mathbb{N}$)

Soit X une variable aléatoire discrète avec $X(\Omega) = \mathbb{N}$ et g une fonction définie au moins sur $\mathbb{N}$ et à valeurs dans $\mathbb{R}$. On a équivalence entre :

1. La variable $g(X)$ admet une espérance ;
2. La série $\sum_{n \geq 0} g(n)\mathbb{P}(X=n)$ est absolument convergente

Auquel cas

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

Exemple

Soit $\lambda > 0$ et $X \hookrightarrow \mathcal{P}(\lambda)$. Posons $Y = \frac{1}{X+1}$.
Montrer que Y admet une espérance et la calculer

On a $X(\Omega) = \mathbb{N}$. La série de terme général positif $\frac{1}{n+1} \frac{\lambda}{n!} e^{-\lambda} = \frac{\lambda}{(n+1)!} e^{-\lambda}$ converge. Donc Y admet une espérance

$$\begin{aligned} \mathbb{E}(Y) &= \sum_{n=0}^{+\infty} \frac{\lambda^n}{(n+1)!} e^{-\lambda} \\ &= \frac{e^{-\lambda}}{\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^n}{n!} \\ &= \frac{e^{-\lambda}}{\lambda} \left(\sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} - 1 \right) \\ &= \frac{e^{-\lambda}}{\lambda} (e^{\lambda} - 1) \\ &= \frac{1 - e^{-\lambda}}{\lambda} \end{aligned}$$

Propriété 13

Soit X et Y deux variables aléatoires admettant des espérances. Alors

1. Pour tous réels a et b , la variable $aX + bY$ admet une espérance et

$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$$

L'ensemble des variables aléatoires réelles discrètes admettant des espérances est un $\mathbb{R}$ -espace vectoriel

2. Si $X \geq Y$ presque sûrement, alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$
3. Si $X = Y$ presque sûrement, alors $\mathbb{E}X = \mathbb{E}(Y)$
4. Si de plus X et Y sont indépendantes, alors $Z = XY$ admet une espérance mathématique vérifiant

$$\mathbb{E}(Z) = \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).$$

Preuve:

3 Avec sommabilité

$$\mathbb{E}(X)\mathbb{E}(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy P(X=x)P(Y=y)$$

Par indépendance

$$\mathbb{E}(X)\mathbb{E}(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy P(X=x, Y=y)$$

Par le théorème de transfert

$$\mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(XY)$$

Exemple: Inégalité de Jensen

Si X est une variable aléatoire discrète réelle admettant une espérance, si $f: \mathbb{R} \rightarrow \mathbb{R}$ est une application convexe sur $\mathbb{R}$ et si $Y = f(X)$ admet une espérance alors

$$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$$

f est convexe donc f admet une dérivée à droite en tout point de $\mathbb{R}$ et on a

$$\forall x, y \in \mathbb{R}, f(x) + (y-x)f'_d(x) \leq f(y)$$

Posons $m = E(X)$, alors pour tout $k \in X(\Omega)$ on a :

$$f(m) + (k - m)f'_d(m) \leq f(k)$$

donc

$$f(m)p(X = k) + (k - m)f'_d(m)p(X = k) \leq f(k)p(X = k)$$

Par sommabilité, il vient

$$\sum_{k \in X(\Omega)} f(m)p(X = k) + \sum_{k \in X(\Omega)} (k - m)p(X = k)f'_d(m) \leq \sum_{k \in X(\Omega)} f(k)p(X = k)$$

d'où

$$f(m) \sum_{k \in X(\Omega)} p(X = k) + f'_d(m) \sum_{k \in X(\Omega)} (k - m)p(X = k) \leq \sum_{k \in X(\Omega)} f(k)p(X = k)$$

Or

$$\begin{aligned} \sum_{k \in X(\Omega)} p(X = k) &= 1 \\ \sum_{k \in X(\Omega)} (k - E(X))p(X = k) &= E(X - E(X)) = 0 \\ \sum_{k \in X(\Omega)} f(k)p(X = k) &= E(f(X)) \end{aligned}$$

donc $f(m) \leq E(f(X))$.

Corollaire 7

Si $f(X)$ et $g(Y)$ admettent des espérances avec X et Y variables indépendantes alors

$$\mathbb{E}(f(X)g(Y)) = \mathbb{E}(f(X))\mathbb{E}(g(Y))$$

Propriété 14

Soient $X_1, \dots, X_n$ des variables aléatoires discrètes admettant toutes des espérances. Alors

1. La variable $\sum_{i=1}^n X_i$ admet une espérance et $\mathbb{E}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{E}(X_i)$
2. Si les variables sont indépendantes, alors $\prod_{i=1}^n X_i$ admet une espérance et $\mathbb{E}\left(\prod_{i=1}^n X_i\right) = \prod_{i=1}^n \mathbb{E}(X_i)$

Définition 21: Variable centrée

Si X est une variable aléatoire telle que $\mathbb{E}(X) = 0$, on dit que **X est une variable centrée**.

Propriété 15

Soit X une variable aléatoire discrète admettant une espérance.

La variable aléatoire $\tilde{X} = X - \mathbb{E}(X)$ est une variable aléatoire discrète centrée appelée **la variable aléatoire centrée associée à X** .

Définition 22: Espérance d'un n -upplet

Soit $(X_1, \dots, X_n)$ un n -upplet de variables aléatoires réelles discrètes sur l'espace probabilisé (Ω, T, P) admettant une espérance, on définit le vecteur espérance $\mathbb{E}((X_1, \dots, X_n))$ du vecteur aléatoire $(X_1, \dots, X_n)$ par l'égalité

$$\mathbb{E}((X_1, \dots, X_n)) = (\mathbb{E}(X_1), \dots, \mathbb{E}(X_n))$$

II.2 Moments

Définition 23

Soit X une variable aléatoire réelle discrète et $r \in \mathbb{N}^*$. Si X^r admet une espérance alors on dit que X admet un **moment d'ordre r** qui est le réel

$$m_r(X) = \mathbb{E}(X^r)$$

Remarque :

- Le moment d'ordre 1 est l'espérance
- Toutes les variables aléatoires discrètes finies admettent des moments d'ordre r pour tout $r \in \mathbb{N}^*$.

Remarque :

Notons qu'il découle du théorème du transfert

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

Propriété 16

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

1. X admet des moments d'ordre s pour tout $s \in \llbracket 1; r \rrbracket$;
2. Pour tout $\alpha \in \mathbb{R}$, la variable $X + \alpha$ admet un moment d'ordre r

Preuve:

Si X admet un moment d'ordre r

1. On a, selon $|x| > 1$ ou non,

$$|x|^s \leq 1 + |x|^r$$

ce qui permet d'établir par comparaison la sommabilité de $(x^r \mathbb{P}(X = x))_{x \in X(\Omega)}$

2. On fait appel à la formule du binôme de Newton $(X + \alpha)^r = \sum_{k=0}^r C_r^k \alpha^k X^{r-k}$

Corollaire 8

Si $\mathbb{E}(X^2)$ existe alors $\mathbb{E}(X)$ existe

II.3 Variance et écart-type

Définition 24

Soit X une variable aléatoire admettant un moment d'ordre 2. On appelle **variance de X** le réel :

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

De plus lorsque $\mathbb{V}(X)$ existe, on appelle **écart-type de X** le réel

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

Remarque :

- Si X n'admet pas d'espérance, X ne peut pas admettre de variance.
- La variance est la moyenne du carré de la distance entre les valeurs de X et la moyenne de X . La variance est donc une mesure de dispersion de X par rapport à $\mathbb{E}(X)$.

Théorème 10

Soit X une variable aléatoire admettant un moment d'ordre deux, alors

1. $\mathbb{V}(X) \geq 0$
2. Formule de Kœnig-Huygens : $\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$
3. $\mathbb{V}(X) = 0 \iff \mathbb{P}(X = \mathbb{E}(X)) = 1$

4. Pour tout réels a et b , $aX + b$ admet une variance et $\mathbb{V}(aX + b) = a^2 \mathbb{V}(X)$

Preuve:

Soit X une variable aléatoire admettant un moment d'ordre 2

1. $(X - \mathbb{E}(X))^2$ est une variable positive admettant une espérance, donc $\mathbb{V}(X) = \mathbb{E}((X - \mathbb{E}(X))^2) \geq 0$

2. Posons $m = \mathbb{E}(X)$. En développant

$$(X - m)^2 = X^2 - 2mX + m^2$$

Par le théorème de transfert

$$\begin{aligned} \mathbb{V}(X) &= \sum_{x \in X(\Omega)} (x^2 - 2mx + m^2) P(X = x) \\ &= \sum_{x \in X(\Omega)} x^2 P(X = x) - 2m \sum_{x \in X(\Omega)} x P(X = x) + m^2 \sum_{x \in X(\Omega)} P(X = x) \\ &= \mathbb{E}(X^2) - 2m^2 + m^2 = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \end{aligned}$$

3. $\mathbb{V}(X) = 0$, si et seulement si, $\mathbb{E}((X - \mathbb{E}(X))^2) = 0$, si et seulement si $X - \mathbb{E}(X) = 0$ est presque partout

4. Soit $a, b \in \mathbb{R}$, alors

$$\begin{aligned} \mathbb{V}(aX + b) &= \mathbb{E}((aX + b - \mathbb{E}(aX + b))^2) \\ &= \mathbb{E}((aX - a\mathbb{E}(X))^2) = \mathbb{E}(a^2 (X - \mathbb{E}(X))^2) \\ &= a^2 \mathbb{V}(X) \end{aligned}$$

Exemple

Soit X une variable aléatoire discrète avec $X(\Omega) = \mathbb{N}$ et $P(X = n) = \frac{1}{2^{n+1}}$.

Justifier l'existence de la variance de X et la calculer ?

La variable X est à valeurs positives et la série $\sum_{n \geq 0} \frac{n^2}{2^{n+1}}$ est convergente.

Pour calculer cette somme, exploitons la série entière

$$\forall x \in]-1, 1[, \quad \sum_{n=1}^{+\infty} nx^n = x \left(\sum_{n=0}^{+\infty} x^n \right)' = \frac{x}{(1-x)^2}$$

Redérivons une autre fois

$$\forall x \in]-1, 1[, \quad \sum_{n=1}^{+\infty} n^2 x^{n-1} = \left(\sum_{n=1}^{+\infty} nx^n \right)' = \frac{1+x}{(1-x)^3}$$

donnant : $\mathbb{E}(X^2) = \sum_{n=1}^{+\infty} \frac{n^2}{2^{n+1}} = 3$, puis

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

Exemple

[Loi de Zipf] Soit $s \in]1, +\infty[$. On se donne une variable aléatoire X_s à valeurs dans $\mathbb{N}^*$ de loi

$$\forall n \in \mathbb{N}^*, \quad P(X_s = n) = \frac{n^{-s}}{\zeta(s)}$$

1. Pour quelles valeurs de s , la variable X_s admet une espérance et la calculer

2. Pour quelles valeurs de s , la variable X_s admet une variance et la calculer

1. X_s admet une espérance si, et seulement, si la SATP $\sum_{n \geq 1} nP(X_s = n)$ converge, c'est-à-dire si $s > 2$.

Auquel cas

$$\mathbb{E}(X_s) = \sum_{n=1}^{+\infty} \frac{1}{\zeta(s)n^{s-1}} = \frac{\zeta(s-1)}{\zeta(s)}$$

2. X_s admet une variance si, et seulement, si la SATP $\sum_{n \geq 1} n^2 P(X_s = n)$ converge, c'est-à-dire si $s > 3$.

Auquel cas

$$\mathbb{E}(X_s^2) = \sum_{n=1}^{+\infty} \frac{1}{\zeta(s)n^{s-2}} = \frac{\zeta(s-2)}{\zeta(s)}$$

Par la formule de Huygens kœnig

$$\mathbb{V}(X) = \mathbb{E}(X_s^2) - \mathbb{E}(X_s)^2 = \frac{\zeta(s-2)\zeta(s) - \zeta(s-1)^2}{\zeta(s)^2}$$

Propriété 17

[Variance des lois usuelles] Soit X une variable aléatoire discrète

1. Si $X \hookrightarrow \mathcal{B}(p)$, alors $\mathbb{V}(X) = pq$
2. Si $X \hookrightarrow \mathcal{B}(n, p)$, alors $\mathbb{V}(X) = npq$
3. Si $X \hookrightarrow \mathcal{U}(\llbracket 1; n \rrbracket)$, alors $\mathbb{V}(X) = \frac{n^2 - 1}{12}$
4. Si $X \hookrightarrow \mathcal{G}(p)$, alors $\mathbb{E}(X) = \frac{q}{p^2}$
5. Si $X \hookrightarrow \mathcal{P}(\lambda)$, alors $\mathbb{V}(X) = \lambda$

Définition 25

Si X admet un écart-type non nul, la variable $X^* = \frac{X - \mathbb{E}(X)}{\sigma(X)}$ est appelée **la variable centrée réduite associée à X** .

Théorème 11: Inégalité de Bienaymé-Tchebychev

Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, T, P) possédant un moment d'ordre 2, alors on a

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

Preuve:

On applique l'inégalité de Markov pour la variable $Y = (X - m)^2$ et $\lambda = \varepsilon^2$

II.4 Covariance et coefficient de corrélation linéaire

Propriété 18: Inégalité de Cauchy-Schwarz

Si les variables X et Y admettent chacune un moment d'ordre 2. Alors XY admet une espérance et

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

Il y a égalité si et seulement si $P(X = 0) = 1$ ou il existe $\lambda \in \mathbb{R}$ tel que $P(Y = \lambda X) = 1$

Preuve:

Pour tout $x, y \in \mathbb{R}$, on a $2|xy| = x^2 + y^2$ donc

$$|XY| = \frac{1}{2}(X^2 + Y^2)$$

Puisque les variables X^2 et Y^2 admettent une espérance, la variable XY aussi.

Soit $\lambda \in \mathbb{R}$. Introduisons la variable $Z = (\lambda X + Y)^2 = \lambda^2 X^2 + 2\lambda XY + Y^2$. Par combinaison linéaire, Z admet une espérance et puisque Z est positive

$$\lambda^2 \mathbb{E}(X^2) + 2\lambda \mathbb{E}(XY) + \mathbb{E}(Y^2) \geq 0$$

Cette identité est vraie pour tout $\lambda \in \mathbb{R}$.

- Cas $\mathbb{E}(X^2) > 0$: le trinôme associé au premier membre ne peut posséder deux racines réelles et donc

$$\Delta = 4\mathbb{E}(XY)^2 - 4\mathbb{E}(X^2)\mathbb{E}(Y^2) \leq 0$$

Soit

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

- Cas $\mathbb{E}(X^2) = 0$: On a nécessairement $\mathbb{E}(XY) = 0$ car sinon $2\lambda \mathbb{E}(XY) + \mathbb{E}(Y^2)$ change de signe en $-\frac{\mathbb{E}(Y^2)}{2\mathbb{E}(XY)}$.

| En cas d'égalité, soit X est nulle p.p où il existe $\lambda_0 \in \mathbb{R}$ tel que $\mathbb{E}((\lambda_0 X + Y)^2) = 0$, donc $\lambda_0 X + Y = 0$ p.p

Définition 26

Soient X et Y deux v.a.r admettant des moments d'ordre 2.

1. On appelle **covariance de X et de Y** le nombre réel :

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

2. On appelle coefficient de corrélation de X et de Y si $\sigma(X)\sigma(Y) \neq 0$:

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

3. On dit que X et Y sont non corrélées si $\text{cov}(X, Y) = 0$.

Propriété 19

Soit X, Y, Z des v.a.r admettant des moments d'ordre 2 et soit a et b deux réels

1. $\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
2. $\text{cov}(X, X) = \mathbb{V}(X)$,
3. $\text{cov}(X, Y) = \text{cov}(Y, X)$,
4. $\text{cov}(aX + bZ, Y) = a.\text{cov}(X, Y) + b.\text{cov}(Z, Y)$
5. $\text{cov}(X, Y)^2 \leq \mathbb{V}(X)\mathbb{V}(Y)$
6. Si $\sigma(X)\sigma(Y) \neq 0$, alors
 - a) $|\rho(X, Y)| \leq 1$
 - b) $\rho(X, Y) = 1 \iff \exists \alpha \in \mathbb{R}_+^* \text{ tel que } \tilde{Y} = \alpha \tilde{X} \text{ est presque partout.}$
 - c) $\rho(X, Y) = -1 \iff \exists \alpha \in \mathbb{R}_-^* \text{ tel que } \tilde{Y} = \alpha \tilde{X} \text{ est presque partout.}$

Preuve:

Soit X, Y, Z des VARD admettant des moments d'ordre 2 et soit a et b deux réels

1. Calcul
2. Formule de Huygens ou de Kœnig
3. Facile
4. Conséquence de la linéarité de l'espérance
5. L'inégalité de Cauchy-Schwarz appliquée aux variables $\tilde{X} = X - \mathbb{E}(X)$ et $\tilde{Y} = Y - \mathbb{E}(Y)$
6. a) De 4.) de la propriété précédente, on obtient $|\rho(X, Y)| \leq 1$
 b) $\rho(X, Y) = 1$ donne $\mathbb{E}(\tilde{X}\tilde{Y}) = \sqrt{\mathbb{E}(\tilde{X}^2)}\sqrt{\mathbb{E}(\tilde{Y}^2)}$. Or $\tilde{X}$ et $\tilde{Y}$ ne sont pas nulles presque sûr car $\mathbb{V}(X).\mathbb{V}(Y) \neq 0$, alors il existe $\alpha \in \mathbb{R}^*$ tel que $\tilde{Y} = \alpha \tilde{X}$.
 D'autre part en remplaçant $\tilde{Y}$ par $\alpha \tilde{X}$ dans $\mathbb{E}(\tilde{X}\tilde{Y}) = \sqrt{\mathbb{E}(\tilde{X}^2)}\sqrt{\mathbb{E}(\tilde{Y}^2)}$, on obtient $\alpha = |\alpha|$, donc $\alpha \in \mathbb{R}_+^*$.
 Inversement $\tilde{Y} = \alpha \tilde{X}$, avec $\alpha \in \mathbb{R}_+^*$, donne

$$\rho(X, Y) = \frac{\mathbb{E}(\tilde{X}\tilde{Y})}{\sqrt{\mathbb{E}(\tilde{X}^2)}\sqrt{\mathbb{E}(\tilde{Y}^2)}} = \frac{\alpha \mathbb{E}(\tilde{X}^2)}{|\alpha| \sqrt{\mathbb{E}(\tilde{X}^2)}\sqrt{\mathbb{E}(\tilde{X}^2)}} = 1$$

Finalement $\tilde{Y} = \alpha \tilde{X}$ p.p équivaut à $Y = \alpha X + \beta$ presque sûrement, avec $\beta = \mathbb{E}(Y) - \alpha \mathbb{E}(X)$

Remarque :

1. La première assertion de la propriété est utilisée le plus souvent pour calculer la covariance.
2. Si X et Y sont indépendantes alors elles sont non corrélées $\text{cov}(X, Y) = 0$.
3. Si $\text{cov}(X, Y) \neq 0$, alors X et Y ne sont pas indépendantes.

Attention

Deux v.a.r indépendantes sont non corrélées mais la réciproque est en générale fausse.

Exemple

Soit X suivant une loi uniforme sur $\{-1, 0, 1\}$, on pose $Y = X^2$.
Montrer que X et Y sont non corrélées, et l'indépendance ?

On a $Y(\Omega) = \{0, 1\}$ avec $[Y = 1] = [X = -1] \cup [X = 1]$, ainsi, par le théorème du transfert

$$\mathbb{E}(XY) = P(X = 1, Y = 1) - P(X = -1, Y = 1) = P(X = 1) - P(X = -1) = 0$$

et $\mathbb{E}(X) = 0$, alors $\text{cov}(X, Y) = 0$. Donc les variables ne sont pas corrélées.

D'autre part $P(X = 1, Y = 1) = P(X = 1) = \frac{1}{3}$ et $P(Y = 1) = \frac{2}{3}$, donc $P(X = 1, Y = 1) \neq P(X = 1)P(Y = 1)$

Propriété 20

[Variance de la somme de deux variables] Soit (X, Y) un couple aléatoire discret tel que X et Y ont une variance,

$$\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y) + 2\text{cov}(X, Y).$$

Si de plus X et Y sont indépendantes, alors $\mathbb{V}(X + Y) = \mathbb{V}(X) + \mathbb{V}(Y)$

Preuve:

L'existence de $\mathbb{V}(X + Y)$ est justifiée car $\mathbb{E}(XY)$ existe avec Cauchy-Schwarz et donc le moment d'ordre deux existe pour $X + Y$, donc la variance.

$$\mathbb{E}((X + Y)^2) = \mathbb{E}(X^2) + \mathbb{E}(Y^2) + 2\mathbb{E}(XY)$$

et

$$(\mathbb{E}(X + Y))^2 = \mathbb{E}(X)^2 + \mathbb{E}(Y)^2 + 2\mathbb{E}(X)\mathbb{E}(Y),$$

Par différence et formule de Koenig-Huygens, on obtient le résultat

Exemple

Soit $(X_n)_{n \geq 1}$ une suite de variables aléatoires deux à deux indépendantes et de même loi admettant une variance σ^2 . Calculer l'espérance de la variable aléatoire définie par

$$Y = \frac{1}{n-1} \sum_{k=1}^{n-1} (X_k - X_n)^2$$

Les variables $X_k - X_n$ sont centrées puisque X_k et X_n ont la même espérance. Par suite l'espérance de $(X_k - X_n)^2$ est égale à la variance $\mathbb{V}(X_k - X_n)$, soit à $\mathbb{V}(X_k) + \mathbb{V}(X_n) = 2\sigma^2$ par indépendance de X_k et de X_n . Par suite l'espérance demandée de $\mathbb{V}(Y_n) = 2\sigma^2$

Propriété 21: Variance d'une somme

Soit $(X_1, \dots, X_n)$ un vecteur de n variables aléatoires discrètes admettant de moments d'ordre 2, alors $\sum_{i=1}^n X_i$ admet une variance et on a

$$\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{cov}(X_i, X_j).$$

Si de plus les variables sont indépendantes, alors $\mathbb{V}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \mathbb{V}(X_i)$.

Exemple

Soit X une variable aléatoire discrète suivant une loi binomiale $\mathcal{B}(n, p)$. On interprète X comme la somme de n variables aléatoires indépendantes X_i suivant une loi de Bernoulli de paramètre p . Donc

$$X = \sum_{i=1}^n X_i \quad \text{avec } \forall i \in \llbracket 1, n \rrbracket \quad X_i \sim \mathcal{B}(1, p)$$

d'où $\mathbb{E}(X) = np$ et $\mathbb{V}(X) = npq$ par indépendance des X_i .

Remarque :

Terminologie probabiliste	Terminologie euclidienne
Covariance cov	Produit scalaire $\langle . . \rangle$
Variance $\mathbb{V}$	Norme au carré $\ .\ _2^2$
Écart type σ	Norme $\ .\ $
Non corrélation	Orthogonalité
VAR nulle presque sûr	Vecteur nul

II.5 Loi faible des grands nombres**Théorème 12: Loi faible des grands nombres**

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variables aléatoires réelles **indépendantes** définies sur un espace probabilisé (Ω, T, P) , de même loi, possédant une espérance m et une variance σ^2 . Alors pour tout $\varepsilon > 0$

$$\mathbb{P}_X \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \xrightarrow{n \rightarrow +\infty} 0$$

$$\text{Avec } S_n = \sum_{i=1}^n X_i$$

Preuve:

On a $\mathbb{E} \left(\frac{S_n}{n} \right) = m$ et puisque les variables sont indépendantes, alors $\mathbb{V} \left(\frac{S_n}{n} \right) = \frac{\sigma^2}{n}$. On applique l'inégalité de Bienaymé-Tchebychev et on obtient

$$\forall \varepsilon > 0 \quad \mathbb{P} \left(\left| \frac{S_n}{n} - m \right| \geq \varepsilon \right) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$$

III Fonctions génératrices

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La série entière $\sum_{n \geq 0} P(X = n)z^n$ converge pour $z_0 = 1$, donc, d'après le lemme d'Abel, elle est de rayon de convergence $R_X \geq 1$

Définition 27

On définit sa fonction génératrice G_X par

$$\forall s \in [-1, 1], \quad G_X(s) = \mathbb{E}(s^X) = \sum_{k=0}^{+\infty} \mathbb{P}(X = k)s^k$$

Remarque :

Si X ne prend qu'un nombre fini de valeurs, G_X est un polynôme en s , elle est donc définie sur $\mathbb{R}$ tout entier.

Exemple: X suit la loi de Bernoulli

Si X la loi de Bernoulli $\mathcal{B}(1, p)$, alors

$$G_X(t) = 1 - p + pt$$

Exemple: X suit la loi binomiale

Si X la loi de Bernoulli $\mathcal{B}(n, p)$, alors

$$\begin{aligned} G_X(t) &= \sum_{k=0}^n C_n^k t^k p^k (1-p)^{n-k} \\ &= (1-p + pt)^n \end{aligned}$$

Exemple: X suit la loi de poisson

Si X la loi de poisson $\mathcal{P}(\lambda)$ avec $\lambda > 0$, alors

$$\begin{aligned} G_X(t) &= \sum_{k=0}^{+\infty} \frac{t^k \lambda^k}{k!} e^{-\lambda} \\ &= e^{\lambda(t-1)} \end{aligned}$$

Exemple: X suit la loi géométrique

Si X la loi de poisson $\mathcal{G}(p)$ avec $p \in]0, 1[$, alors $\forall t \in \left] -\frac{1}{q}, \frac{1}{q} \right[:$

$$\begin{aligned} G_X(t) &= \sum_{n=1}^{+\infty} t^n q^{n-1} p \\ &= \frac{tp}{1-qt} \end{aligned}$$

Propriété 22: Caractérisation d'une loi par sa fonction génératrice

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ et G_X sa fonction génératrice, alors

$$\forall n \in \mathbb{N}, \quad P(X = n) = \frac{G_X^{(n)}(0)}{n!}$$

Preuve:

G_X est la somme d'une série entière de rayon $R \geq 1$

Propriété 23

Les assertions suivantes sont équivalentes

1. X admet une espérance
2. G_X est dérivable en 1.

Dans un tel cas $E(X) = G'_X(1)$

Preuve:

$1 \Rightarrow 2)$ Si X admet une espérance, alors la série $\sum_{n \geq 0} nP(X=n)$ converge. Les deux séries $\sum_{n \geq 0} P(X=n)t^n$ et $\sum_{n \geq 1} nP(X=n)t^{n-1}$ sont alors normalement convergentes sur $[-1, 1]$, donc G_X est dérivable en 1 de dérivé $G'_X(1) = \mathbb{E}(X)$

$1 \Leftarrow 2)$ Supposons que G_X est dérivable en 1. Le taux d'accroissement

$$\frac{G_X(t) - G_X(1)}{t - 1} = \sum_{n=1}^{+\infty} \frac{t^n - 1}{t - 1} P(X=n) = \sum_{n=1}^{+\infty} \sum_{k=0}^{n-1} t^k P(X=n)$$

admet une limite finie quand $t \rightarrow 1^-$.

Soit $N \in \mathbb{N}$, on a :

$$\begin{aligned} \sum_{n=1}^N nP(X=n) &= \lim_{t \rightarrow 1^-} \sum_{n=1}^N \sum_{k=0}^{n-1} t^k P(X=n) \\ &\leq \lim_{t \rightarrow 1^-} \sum_{n=1}^{+\infty} \sum_{k=0}^{n-1} t^k P(X=n) = G'_X(1) \end{aligned}$$

La série $\sum_{n \geq 1} nP(X=n)$ est donc convergente car c'est une série à termes positifs aux sommes partielles majorées.

Exemple

Retrouver les espérances des lois usuelles

Propriété 24

Les assertions suivantes sont équivalentes

1. X admet un moment d'ordre 2
2. G_X est deux fois dérivable en 1.

Auquel cas $\mathbb{V}(X) = G_X''(1) + G_X'(1) - (G_X'(1))^2$

Preuve:

$1 \Rightarrow 2$) Si X admet un moment d'ordre 2, il y a convergence de la série $\sum_{n \geq 0} n^2 P(X=n)$ mais aussi de la série $\sum_{n \geq 0} n(n-1)P(X=n)$.

Les trois séries $\sum_{n \geq 0} P(X=n)t^n$, $\sum_{n \geq 1} nP(X=n)t^{n-1}$ et $\sum_{n \geq 0} n(n-1)P(X=n)t^n$ sont alors normalement convergentes sur $[-1, 1]$, donc G_X est deux fois dérivable en 1 de dérivé premier $G_X'(1) = \mathbb{E}(X)$ et de dérivé second

$$G_X''(1) = \mathbb{E}(X(X-1)) = \mathbb{E}(X^2) - \mathbb{E}(X)$$

Soit $\mathbb{E}(X^2) = G_X''(1) + G_X'(1)$. Par la formule de Huygens, on obtient

$$\mathbb{V}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = G_X''(1) + G_X'(1) - (G_X'(1))^2$$

$1 \Leftarrow 2$) Supposons que G_X deux fois dérivable en 1. La fonction G_X est dérivable en 1, donc X admet une espérance. On sait de plus l'expression de $G_X'(t)$ sur $[-1, 1]$.

$$G_X'(t) = \sum_{n=1}^{+\infty} nP(X=n)t^{n-1}$$

$$\frac{G_X'(t) - G_X'(1)}{t-1} = \sum_{n=2}^{+\infty} n \frac{t^{n-1} - 1}{t-1} P(X=n) = \sum_{n=2}^{+\infty} n \sum_{k=0}^{n-2} t^k P(X=n)$$

admet une limite finie quand $t \rightarrow 1^-$.

Soit $N \in \mathbb{N}$, on a :

$$\begin{aligned} \sum_{n=2}^N n(n-1)P(X=n) &= \lim_{t \rightarrow 1^-} \sum_{n=2}^N n \sum_{k=0}^{n-2} t^k P(X=n) \\ &\leq \lim_{t \rightarrow 1^-} \sum_{n=2}^{+\infty} n \sum_{k=0}^{n-2} t^k P(X=n) = G_X''(1) \end{aligned}$$

La série $\sum_{n \geq 2} n(n-1)P(X=n)$ est donc convergente car c'est une série à termes positifs aux sommes partielles majorées.

Or la série $\sum_{n \geq 2} n(n-1)P(X=n)$ converge, alors $\sum_{n \geq 0} n^2 P(X=n)$ converge, c'est-à-dire, la variable X admet un moment d'ordre 2

Exemple

Retrouver les variances des lois usuelles

Propriété 25

Soit X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$ de fonctions génératrices respectives G_X et G_Y . Alors

$$G_{X+Y} = G_X \times G_Y.$$

Preuve:

Comme X et Y sont indépendantes,

$$G_{X+Y}(s) = \mathbb{E}(s^{X+Y}) = \mathbb{E}(s^X s^Y) = \mathbb{E}(s^X) \mathbb{E}(s^Y) = G_X(s) \times G_Y(s)$$

Exemple

Soit X et Y deux variables aléatoires indépendantes de lois de Poisson de paramètres respectifs $\lambda > 0$ et $\mu > 0$. Déterminer, en calculant sa fonction génératrice, la loi de $X + Y$.

On sait que $\forall s \in \mathbb{R}$, on a $G_X(s) = e^{\lambda(t-1)}$ et $G_Y(s) = e^{\mu(t-1)}$. L'indépendance donne

$$\forall s \in \mathbb{R}, \quad G_{X+Y}(s) = G_X(s)G_Y(s) = e^{(\lambda+\mu)(t-1)}$$

On conclut ainsi la loi de la somme :

$$\forall n \in \mathbb{N}, \quad P(X+Y=n) = \frac{G_{X+Y}^{(n)}(0)}{n!} = \frac{(\lambda+\mu)^n}{n!} e^{-\lambda}$$

Propriété 26

Si $X_1, \dots, X_n$ sont des variables aléatoires discrètes indépendantes à valeurs dans $\mathbb{N}$. On note $S_n = \sum_{k=1}^n X_k$, alors

$$G_{S_n} = \prod_{k=1}^n G_{X_k}$$