Université de Gabès

Faculté des Sciences de Gabès

Probabilités et Statistique

LFMa3

HDHIRI Ibtissem

A.U. : 2015-2016

Introduction aux probabilités

0.1 Introduction

La théorie des probabilités est la science qui vise à modéliser les phénomènes aléatoires qui sont des phénomènes dont on ne peut prévoir le résultat à l'avance.

La modélisation d'une experience aléatoire à travers le calcul des probabilités a été introduite par Kolmogorov en 1933. Elle est essentiellement basée sur quatres objets fondamentaux:

a- L'espace d'états ou univers: C'est l'ensemble de tous les résultats possibles de l'experience aléatoire étudiée. Il est habituellement noté Ω et ses éléments, dits événements élémentaires sont notés par ω.

Exemples 0.1.1. • *Pour un tirage de pile ou face, l'univers est donné par:* $\Omega = \{P, F\}$

- *Pour le lancer de deux dés,* $\Omega = \{(i,j);\ 1 \leq i \leq 6;\ 1 \leq j \leq 6\}$

b- Les événements: Ce sont des parties de l'univers Ω. On notera $\mathcal{A}$ l'ensemble des tous les événements. Cet ensemble coincide souvent mais pas toujours avec l'ensemble des parties de Ω. D'une manière générale, $\mathcal{A}$ doit avoir la structure d'une tribu.

Il est utile de savoir que les termes utilisés pour décrire les événement admettent des traductions en language ensembliste:

- Ω : événement certain
- $\emptyset$: événement impossible
- $A \cup B$ Au moins A ou B est réalisé
- $A \cap B$: A et B sont réalisés
- $A^c = \Omega \backslash A$: événement contraire de A.

c- La probabilité : A tout événement A de $\mathcal{A}$; on associe un nombre $\mathbb{P}(A)$ compris entre 0 et 1. Moralement, ce nombre mésure la vraissemblance de la réalisation de l'événement A.

Définition 0.1.1. *Étant donnés un espace d'états* Ω *et une tribu*

d'événements $\mathcal{A}$; une probabilité est une mesure $\mathbb{P}$ sur $(\Omega, \mathcal{A})$ vérifiant $\mathbb{P}(\Omega) = 1$.

Le triplet $(\Omega; \mathcal{A}, \mathbb{P})$ est dit espace probabilisé ou espace de probabilité.

Exemple:

(i) La mesure de Lebesgue sur $[0, 1]$

(ii) Probabilité de Bernouilli $\mathbb{P} = p\delta_1 + (1-p)\delta_0$.

Voici quelques propriétés utiles dont certaines ne sont que des reformulations de la définition. Soient $A, B \in \mathcal{A}$

- Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$.
- $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$
- $\mathbb{P}(A - B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.

Lemma 0.1.1.

Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante au sens de l'inclusion (i.e. $A_n \subset A_{n+1}$ pour tout $n \in \mathbb{N}$) alors la suite $(\mathbb{P}(A_n))_{n\in\mathbb{N}}$ est croissante et $\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\bigcup_{n\in\mathbb{N}} A_n)$

Si $(A_n)_{n\in\mathbb{N}}$ *est une suite décroissante au sens de l'inclusion alors la suite* $(\mathbb{P}(A_n))_{n\in\mathbb{N}}$ *est décroissante et* $\lim_{n\to\infty}\mathbb{P}(A_n) = \mathbb{P}(\bigcap_{n\in\mathbb{N}} A_n)$

d- Variable aléatoire : (v.a. en abrégé) C'est une grandeur qui dépend du résultat de l'experience aléatoire en question. En termes mathématiques, une v.a. X est une application mesurable de $(\Omega;\mathcal{A})$ dans (E,τ) $(E=\mathbb{R}^d,' d\geq 1)$.

Maintenant, on peut définir une nouvelle probabilité $\mathbb{P}_X$ en posant:

$$\mathbb{P}_X(B) = \mathbb{P}[X^{-1}(B)],\ B\in\mathcal{B}(E).$$

Cette probabilité est dite loi de la v.a. X.

0.2 Indépendance d'événements

Intuitivement, dire que deux événements sont indépendants revient à dire que la réalisation de l'un ne donne aucune information sur celle de l'autre.

Exemple : On lance un dé deux fois successives; on note A: 'Avoir un nombre pair au premier lancer' et B:'Avoir 2 au deuxième lancer'.

Définition 0.2.1. *1. Deux événements* A *et* B *sont dits indépendants si* $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B)$

2. *Soit $(A_i)_{i\in I}$ une famille d'événements. Ces événements sont dits indépendants ou mutuellement indépendants si pour toute famille finie J de I, on a :*

$$\mathbb{P}(\bigcap_{j\in J} A_j) = \prod_{j\in J} \mathbb{P}(A_j).$$

Remarque 0.2.1. *Des événements indépendants sont deux à deux indépendants mais la réciproque n'est pas nécesserement vraie.*

Proposition 0.2.1. *Si A et B sont deux événements indépendants alors A^c et B sont également indépendants. En effet; $B = (A \cap B) \cup (A^c \cap B)$; donc $\mathbb{P}(A^c \cap B) = \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B)\mathbb{P}(A^c)$.*

0.3 conditionnement

Définition 0.3.1. *Soit $(\Omega;\ \mathcal{A},\ \mathbb{P})$ un espace de probabilité et B un événement vérifiant $\mathbb{P}(B) > 0$. On définit une nouvelle probabilité $\mathbb{P}_B$ pour tout événement A par:*

$$\mathbb{P}_B(A) = \mathbb{P}[A|B] = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Cette quantité est dite probabilité conditionnelle de A sachant B.

Définition 0.3.2. *Une famille $(B_i)_{i\in\mathbb{N}}$ est dite partition de Ω; ou système complet d'événements si $\bigcup_{i\in\mathbb{N}} B_i = \Omega$ et $B_i \cap B_j = \emptyset,\ \forall i \neq j$.*

Proposition 0.3.1. *Soient I un ensemble au plus dénombrable de $\mathbb{N}$ et $(B_i)_{i\in I}$ une partition de Ω telle que $\mathbb{P}(B_i) > 0$; $\forall i \in I$*

(i) *Pour tout $A \in \mathcal{A}$; on a $\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)$*

(ii) *Pour tout $A \in \mathcal{A}$ tel que $\mathbb{P}(A) > 0$, on a*

$$\mathbb{P}_A(B_j) = \frac{\mathbb{P}_{B_j}(A)\mathbb{P}(B_j)}{\sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)},\ j \in I.$$

La dernière formule est dite formule de Bayes.

Preuve :

(i) On a $A = \dot{\bigcup}_{i\in I} A \cap B_i$ donc

$$\mathbb{P}(A) = \sum_{i\in I} \mathbb{P}(A \cap B_i) = \sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)$$

(ii) On a pour tout $j \in I$;

$$\mathbb{P}_A(B_j) = \frac{\mathbb{P}(A \cap B_j)}{\mathbb{P}(A)} = \frac{\mathbb{P}_{B_j}(A)\mathbb{P}(B_j)}{\sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}.$$

Chapter I

Probabilités sur les ensembles au plus dénombrables

Dans tout ce chapitre, Ω désignera un espace de probabilité au plus dénombrable et on utilisera la convention que si on somme sur un ensemble vide on obtient zéro.

1.1 Construction de probabilité

Une mesure de probabilité $\mathbb{P}$ définie sur Ω est entièrement caractérisée par ses valeurs sur les singletons. En effet, on : $\forall A \in \mathcal{A}$;

$$\mathbb{P}(A) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}).$$

Ceci découle du fait que tout événement A est une réunion disjointe de ses singletons, *i.e.*; $A = \dot{\bigcup}_{\omega \in A} \{\omega\}$.

Inversement, si $\omega \mapsto p(\omega)$ est une application de Ω dans $\mathbb{R}_+$ vérifiant $\sum_{\omega \in \Omega} p(\omega) = 1$, alors on peut facilement montrer que l'application

définie sur $\mathcal{P}(\Omega)$ par

$$P(A) := \sum_{\omega \in A} p(\omega);\ A \in \mathcal{P}(\Omega)$$

est une mesure de probabilité.

1.2 Variables aléatoires discrètes

Les v.a. seront représentées par des lettres majuscules telles que $X;\ Y;\ Z...$

1.2.1 Loi d'une variable aléatoire discrète

Une v.a. X est une application mesurable de Ω dans un espace mesurable (E, τ). Étant donné que l'univers Ω est fini ou dénombrable; il en est de même pour $X(\Omega)$. Dans ce cas, on dit que X est une v.a.discrète (v.a.d.).

La loi de X est la probabilité définie sur $X(\Omega)$ dont la valeur sur les singletons est donnée par:

$$\mathbb{P}_X(\{j\}) = \mathbb{P}[X = j] = \sum_{\omega,\, X(\omega)=j} \mathbb{P}(\omega).$$

On supposera dans la suite que X est une v.a. réelle, *i.e.*, $E = \mathbb{R}$.

Définition 1.2.1. *On appelle fonction de répartition (fdr en abrégé) de la v.a. X et on note F_X la fonction allant de $\mathbb{R}$ dans $[0,1]$ définie par:*

$$F_X(x) = \mathbb{P}_X(]-\infty, x]) = \mathbb{P}(X \leq x);\ x \in \mathbb{R}.$$

Cette fonction vérifie les propriétés suivantes:

Proposition 1.2.1. (i) *La fonction F_X est croissante et vérifie:*

$$\lim_{n\to-\infty} F_X(x) = 0 \ e\ t \lim_{n\to+\infty} F_X(x) = 1.$$

(ii) *C'est une fonction continue à droite limitée à gauche (càdlàg).*

Preuve :

(i) On a pour $x \leq x'$; $]-\infty, x] \subset]-\infty, x']$. La monotonie est donc obtenue en appliquant la probabilité $\mathbb{P}_X$.

Maintenant, on a

$$1 = \mathbb{P}_X(\mathbb{R}) = \mathbb{P}_X(\bigcup_{n\in\mathbb{N}}]-\infty, n]) = \lim_{n\to\infty} \mathbb{P}_X(]-\infty, n]) = \lim_{n\to+\infty} F_X(n).$$

De même; on a :

$$\begin{aligned} 1 = \mathbb{P}_X(\mathbb{R}) &= \mathbb{P}_X(\bigcup_{n\in\mathbb{N}}]-n, +\infty[) \\ &= \lim_{n\to\infty} \mathbb{P}_X(]-n, +\infty[) \\ &= \lim_{n\to+\infty} (1 - F_X(-n)). \end{aligned}$$

D'où le résultat.

(ii) Soit $x \in \mathbb{R}$ et $(h_n)_{n\in\mathbb{N}}$ une suite réelle décroissante telle que $\lim_{n\to\infty} h_n = 0$. On a :

$$F_X(x+h_n) - F_X(x) = \mathbb{P}_X(]x; x+h_n]).$$

Comme la suite $(]x; x+h_n])_n$ est décroissante au sens de l'inclusion, alors en vertu du Lemme 0.1.1; on obtient :

$$\lim_{n\to\infty} F_X(x+h_n) - F_X(x) = \mathbb{P}_X(\bigcap_{n\geq 0}]x; x+h_n]) = \mathbb{P}_X(\emptyset) = 0.$$

On en déduite que la fonction F_X est càd

1.3 Moments d'une v.a.d

Soient $\mathbb{P}$ une probabilité caractérisée par $p_\omega = \mathbb{P}(\{\omega\})$ et X une v.a.d. définie sur $(\Omega, \mathcal{A})$.

Définition 1.3.1. *Si $\sum_{\omega\in\Omega} |X(\omega)|p_\omega < \infty$, on dit que X admet un moment d'ordre 1 et on pose*

$$\mathbb{E}[X] = \sum_{\omega\in\Omega} X(\omega)p_\omega.$$

Cette quantité est dite espérance mathématique ou moyenne de X.

D'une manière générale; pour $k > 0$; on dit que X admet un moment d'ordre k si la v.a. $|X|^k$ admet un moment d'ordre 1. On note $\mathcal{L}^k$ l'ensembles des v.a. admettant des moments d'ordre k.

Les propriétés suivantes sont immédiates:

$(\mathcal{P}_1)$ $\mathcal{L}^1$ est un espace vectoriel sur lequel l'esperance est linéaire

$(\mathcal{P}_2)$ Si X est bornée alors $X \in \mathcal{L}^k$; $\forall k > 0$

$(\mathcal{P}_3)$ Si $X(\omega) = \alpha$; $\forall \omega \in \Omega$; alors $\mathbb{E}[X] = \alpha$.

$(\mathcal{P}_4)$ Si X est une v.a. positive de $\mathcal{L}^1$; alors $\mathbb{E}[X] \geq 0$.

$(\mathcal{P}_5)$ Si l'univers Ω est fini alors toutes les v.a. sont dans $\mathcal{L}^1$.

$(\mathcal{P}_6)$ Pour tout $A \in \mathcal{A}$; $\mathbb{E}[1_A] = \mathbb{P}(A)$.

Proposition 1.3.1. *$\mathcal{L}^2$ est un sous-espace vectoriel de $\mathcal{L}^1$. De plus; pour toute $X \in \mathcal{L}^2$; on a*

$$\mathbb{E}[|X|] \leq \mathbb{E}[X^2]^{\frac{1}{2}}.$$

Preuve : Soient X, Y deux v.a. de $\mathcal{L}^2$ et $\alpha \in \mathbb{R}$. On a $(\alpha X + Y)^2 \leq 2(\alpha^2 X^2 + Y^2)$, ce qui ipmlique que $\alpha X + Y \in \mathcal{L}^2$.

Par ailleurs, le fait que $|X| \leq 1 + |X|^2$ conduit à dire que $\mathcal{L}^2 \subset \mathcal{L}^1$.

Maintenant, $\forall X \in \mathcal{L}^2$, on a :

$$0 \leq \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - 2\mathbb{E}[X\mathbb{E}[X]] + \mathbb{E}[X]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

D'où le résultat.

Définition 1.3.2. *Soit $X \in \mathcal{L}^2$, La variance de X; notée $Var(X)$ est définie par*

$$Var(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

D'après Proposition 1.3.1, cette quantité est positive. On note σ_X sa racine carrée dite écart type de la v.a. X.

Lemma 1.3.1. *(Sommation par paquets) Soient $(u_n)_{n\in\mathbb{N}} \subset \overline{\mathbb{R}_+}$ et $(A_i)_{i\in\mathbb{N}}$ une partition de $\mathbb{N}^*$; on a :*

$$\sum_n u_n = \sum_{k\in\mathbb{N}^*} \sum_{n\in A_k} u_n.$$

Ce résultat reste vrai si le série de terme général u_n est absolument convergente.

Proposition 1.3.2. *Soient X une v.a.d. définie sur Ω et f une fonction allant de $X(\Omega)$ dans $\mathbb{R}$. La v.a. Y admet un moment d'ordre 1 sur $(\Omega;\mathbb{P})$ si et seulement si f admet un moment d'ordre 1 sur*

$(X(\Omega); \mathbb{P}_X)$ *et on a :*

$$\mathbb{E}[Y] = \sum_{\omega \in \Omega} f[X(\omega)]\mathbb{P}(\{\omega\}) = \sum_{k \in X(\Omega)} f(k)\mathbb{P}_X(\{k\}). \tag{1.1}$$

Preuve :On a

$$\begin{aligned} \sum_{\omega \in \Omega} |f[X(\omega)]| p_\omega &= \sum_{k \in X(\Omega)} |f(k)| \sum_{\omega \in X^{(-1)}(\{k\})} \mathbb{P}(\{\omega\}) \\ &= \sum_{k \in X(\Omega)} |f(k)| \mathbb{P}_X(\{k\}). \end{aligned}$$

Ainsi, Y est intégrable sur $(\Omega; \mathbb{P})$ si et seulement si f est integrable sur $(X(\Omega); \mathbb{P}_X)$ et on a:

$$\mathbb{E}[Y] = \sum_{\omega \in \Omega} f[X(\omega)]\mathbb{P}(\{\omega\}) = \sum_{k \in X(\Omega)} f(k)\mathbb{P}_X(\{k\})$$

1.4 Fonction génératrice des probabilités des v.a. entières

Soit X une v.a. à valeurs dans $\mathbb{N}$ (v.a. entière). On pose

$$p_n = \mathbb{P}[X = n];\ n \in \mathbb{N}.$$

Définition 1.4.1. *La fonction génératrice des probabilités de X est la fonction définie sur $[0, 1]$ par:*

$$g_X(x) := \mathbb{E}[x^X] = \sum_{n \geq 0} x^n p_n.$$

Elle sera notée g s'il n'ya pas risque d'ambiguité.

Remarque 1.4.1. *Comme* $\sum_{n\geq 0} p_n = 1$ *alors le rayon de convergence* R_g *de la série entière* $\sum_{n\geq 0} x^n p_n$ *est supérieur ou égal à 1.*

La fonction g est continue et indéfiniment dérivable pour $s < 1$. Pour $s = 1$; on utilisera le lemme suivant:

Lemma 1.4.1. *(Abel) Soit* $\phi(x) := \sum_{n\in\mathbb{N}} a_n x^n$ *une série entière qui converge pour tout* $|x| < 1$. *Si* $\sum_{n\in\mathbb{N}} a_n < \infty$*; alors* $\lim_{s\to 1^-} \phi(s) = \sum_{n\in\mathbb{N}} a_n = \phi(1)$.

D'autre part, on a

$$g^{(n)}(0) = n!\, p_n;\ n \geq 1.$$

Ainsi la fonction génératrice caractérise la loi d'une v.a. entière X.

Proposition 1.4.1. *Soit* X *une v.a. entière et* g *sa fonction génératrice alors on a équivalence entre:*

(i) $X \in \mathcal{L}^1$

(ii) *La v.a.* X *est dérivable à gauche en 1.*

De plus on a $\mathbb{E}[X] = g'(1)$*; avec* $g'(1) = \lim_{s\to 1^-} \frac{g(s)-g(1)}{s-1}$.

Preuve : on a

$$\begin{aligned}\lim_{s\to 1^-}\frac{g(s)-g(1)}{s-1} &= \lim_{s\to 1^-}\sum_n \frac{s^n-1}{s-1}p_n \\ &= \lim_{s\to 1^-}\sum_n (1+s+...s^{n-1})p_n \\ &= \sum_n np_n = \mathbb{E}[X].\end{aligned}$$

On peut également montrer que si la v.a. X admet un moment d'ordre r, alors on a :

$$g^{(r)}(1) = \mathbb{E}[X(X-1)...(X-r+1)].$$

1.5 Quelques lois usuelles

1.5.1 Loi de Bernouilli $\mathcal{B}(p)$; $0 < p < 1$

Elle est généralement associée aux expériences aléatoires ayant seulement deux issues (succès-echec). On dit que X suit la loi de Bernouilli de paramètre p et on note $X \rightsquigarrow \mathcal{B}(p)$ si l'on a

$$\mathbb{P}[X=1] = p = 1 - \mathbb{P}[X=0].$$

On a

$$\mathbb{E}[X] = \sum_{k\in\{0,1\}} k\mathbb{P}[X=k] = 0 + p = p,$$

$$\mathbb{E}[X^2] = \sum_{k\in\{0,1\}} k^2\mathbb{P}[X=k] = p$$

Ainsi

$$Var(X) = p - p^2 = p(1-p).$$

On peut retrouver le même résultat en utilisant la fonction génératrice; on a $\forall x \in [0;1]$

$$g(x) := \mathbb{E}[x^X] = \sum_{n\in\{0;1\}} x^n p_n = p_0 + xp_1 = px + 1 - p$$

D'où; $\mathbb{E}[X] = g'(1) = p$ et

$$Var(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = g''(1) + p - p^2 = p - p^2$$

1.5.2 Loi Binomiale $\mathcal{B}(n,p)$; $n \geq 1$, $0 < p < 1$

On dit que X suit la loi Binomiale de paramètres n et p et on note $X \rightsquigarrow \mathcal{B}(n,p)$ si l'on a

$$\mathbb{P}[X=k] = C_n^k p^k (1-p)^{n-k}.;\ k \in \{0,1,...;n\}.$$

C'est la loi de la somme de n v.a de Bernoulli indépendantes.

On a

$$g(x) = \sum_{k\in\{0;1;...;n\}} x^k p_k = C_n^k (px)^k (1-p)^{n-k} = (px+1-p)^n.$$

Ainsi,

$$\mathbb{E}[X] = g'(1) = np$$

et

$$\begin{aligned} Var(X) &= \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 \\ &= g''(1) + np - (np)^2 \\ &= n(n-1)p + np - n^2 p^2 \\ &= np(1-p). \end{aligned}$$

1.5.3 Loi de Poisson $\mathcal{P}(\lambda)$; $(\lambda > 0)$

On dit que X suit la loi de Poisson de paramètre λ et on note $X \rightsquigarrow \mathcal{P}(\lambda)$ si l'on a

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda k};\ k = 0; 1; 2; ....$$

On a pour $x \in [0; 1]$

$$g(x) = \sum_{k \geq 0} x^k p_k = \sum_{k \geq 0} \frac{(\lambda x)^k}{k!} e^{-\lambda k} = e^{\lambda(x-1)}.$$

D'où

$$\mathbb{E}[X] = g'(1) = \lambda$$

et

$$Var(X) = g''(1) + g'(1) - (g'(1))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

1.5.4 Loi géométrique $\mathcal{G}(p)$; $(0 < p < 1)$

On dit que X suit la loi géométrique de paramètre p et on note $X \rightsquigarrow \mathcal{G}(p)$ si l'on a

$$\mathbb{P}[X = k] = p(1-p)^{k-1};\ k = 1; 2; ....$$

On a

$$\begin{aligned} g(x) &= \sum_{k\geq 1} x^k p_k \\ &= p\sum_{k\geq 1} x^k (1-p)^{k-1} \\ &= px\sum_{k\geq 0} x^k (1-p)^k \\ &= \frac{px}{1-x(1-p)}. \end{aligned}$$

Ce qui implique que

$$\mathbb{E}[X] = g'(1) = \frac{1}{p}$$

et

$$Var(X) = g''(1) + g'(1) - (g'(1))^2 = \frac{1-p}{p^2}.$$

1.6 Variables aléatoires indépendantes

Soient X et Y deux v.a. définies sur $(\Omega; \mathcal{A})$ à valeurs respectives dans des espace au plus dénombrables E et F. On note pour $(i, j) \in E \times F$;

$$\begin{cases} p_i^X = \mathbb{P}[X = i] \\ p_j^Y = \mathbb{P}[Y = j]. \end{cases}$$

Le couple $Z = (X, Y)$ est une v.a. à valeurs dans $E \times F$ et sa loi de probabilité est donnée par

$$\mathbb{P}_Z\{(i, j)\} = \mathbb{P}[Z = (i; j)] = \mathbb{P}[X = i;\ Y = j];\ (i, j) \in E \times F.$$

La loi conditionnelle de Y si $X = i$ et celle de X sachant que $Y = j$ sont respectivement définies par

$$p_j^{Y|X=i} = \mathbb{P}[Y = j|X = i];\ \text{ si } \mathbb{P}[X = i] > 0$$

et

$$p_i^{X|Y=j} = \mathbb{P}[X = i|Y = j];\ \text{ si } \mathbb{P}[Y = j] > 0$$

Proposition 1.6.1. *Pour tout $(i, j) \in E \times F$; on a*

(i) $\mathbb{P}[X = i] = \sum_{j \in F} \mathbb{P}[Z = (i, j)]$

(ii) $\mathbb{P}[Z = (i, j)] = \begin{cases} p_i^X\ p_j^{Y|X=i}, & si\ p_i^X > 0 \\ 0, & sinon. \end{cases}$

Preuve : La preuve est basée sur le fait que l'événement $\{X = i\}$ s'écrit comme étant la réunion disjointe $\bigcup_{j\in F}\{Z = (i, j)\}$.

Définition 1.6.1. *1. Deux v.a.d. X et Y sont dites indépendantes $(X \perp Y)$ si pour tous $A \in \mathcal{P}(E)$ et $B \in \mathcal{P}(F)$; on a*

$$\mathbb{P}[X \in A,\ Y \in B] = \mathbb{P}[X \in A]\ \mathbb{P}[Y \in B] \tag{1.2}$$

2. Des v.a.d. $X_1, ..., X_n$ sont dites indépendantes ou mutuellement indépendantes si $\forall A_i \subset E_i,; i = 1, ..., n$, on a

$$\mathbb{P}[X_1 \in A_1, ...,\ X_n \in A_n] = \prod_{1\leq i\leq n} \mathbb{P}[X_i \in A_i].$$

3. Une famille de v.a.d. $(X_n)_{n\geq 1}$ est dite indépendante si pour tout $n \in \mathbb{N}^*$; les v.a. $X_1, ..., X_n$ sont indépendantes.

Proposition 1.6.2. *Les assertions suivantes sont équivalentes:*

(i) *X et Y sont indépendantes*

(ii) $\mathbb{P}[Z = (i, j)] = \mathbb{P}[X = i]\ \mathbb{P}[Y = j];\ \forall (i, j) \in E \times F$

(iii) $\mathbb{P}[Y = j|X = i] = \mathbb{P}[Y = j],\ \forall (i, j) \in E \times F$ *tels que* $\mathbb{P}[X = i] > 0$.

Preuve :

- (i) $\Rightarrow$ (2i) Il suffit de prendre $A = \{i\}$ et $B = \{j\}$ dans (1.2).
- (ii) $\Rightarrow$ (i) En Vertue du Lemme 1.3.1; on a pour tous $A \in \mathcal{P}(E)$ et $B \in \mathcal{P}(F)$:

$$\begin{aligned}\mathbb{P}[X \in A,\ Y \in B] &= \sum_{(i,j)\in A\times B} \mathbb{P}[X = i]\mathbb{P}[Y = j] \\ &= \sum_{i\in A}\sum_{j\in B} \mathbb{P}[X = i]\mathbb{P}[Y = j] \\ &= \mathbb{P}[X \in A]\ \mathbb{P}[Y \in B]\end{aligned}$$

- (iii) $\Leftrightarrow$ (ii) évidente.

Proposition 1.6.3. *Soient X et Y deux v.a.d. indépendantes et f et g deux fonctions à valeurs réelles définies respectivement sur E et F telles que $f(X) \in \mathcal{L}^1$ et $g(Y) \in \mathcal{L}^1$, alors $f(X)\ g(Y) \in \mathcal{L}^1$ et on a :*

$$\mathbb{E}[f(X)\, g(Y)] = \mathbb{E}[f(X)]\ \mathbb{E}[g(Y)].$$

Preuve :Soient $Z = (X, Y)$ et ϕ la fonction définie sur $E \times F$ par $\phi(x, y) = f(x)g(y);\ \forall (x, y) \in E \times F$. En vertu de (1.2) et

Lemme1.3.1, On a

$$\begin{aligned}\sum_{z\in E\times F}|\phi(z)|\mathbb{P}[Z=z] &= \sum_{(i,j)\in E\times F}|f(i)g(j)|\mathbb{P}[X=i,\,Y=j]\\ &= \sum_{(i,j)\in E\times F}|f(i)|\mathbb{P}[X=i]|g(j)|\mathbb{P}[Y=j]\\ &= \sum_{i\in E}|f(i)|\mathbb{P}[X=i]\sum_{j\in F}|g(j)|\mathbb{P}[Y=j].\end{aligned}$$

Ce qui implique que si $f(X)\in\mathcal{L}^1$ et $g(Y)\in\mathcal{L}^1$, alors $f(X)\,g(Y)\in\mathcal{L}^1$.

D'une manière analogue, en utilisant Lemme1.3.1 pour les séries absolument convergentes on montre que

$$\begin{aligned}\mathbb{E}[f(X)\,g(Y)] &= \sum_{(i,j)\in E\times F}f(i)g(j)\mathbb{P}[X=i,\,Y=j]\\ &= \sum_{i\in E}f(i)\mathbb{P}[X=i]\sum_{j\in F}g(j)\mathbb{P}[Y=j]\\ &= \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)].\end{aligned}$$

Corollaire 1.6.1. *Si X et Y sont deux v.a. indep de carrés intégrables, on a*

$$Var(X+Y)=Var(X)+Var(Y).$$

Preuve :C'est une conséquence directe de la proposition précédente; il suffit de prendre $\phi(x,y)=xy$.

On peut également vérifier le résultat suivant lié aux fonctions génératrices des v.a.d. indep:

Proposition 1.6.4. *Soient X et Y deux v.a. indep définies sur $\mathbb{N}$, on a $g_{X+Y} = g_X.g_Y$*

Exemples 1.6.1. *Soient X et Y deux v.a.d. indépendantes*

- *Si $X \rightsquigarrow \mathcal{B}(n,p)$ et $Y \rightsquigarrow \mathcal{B}(m;p)$ ($n \in \mathbb{N}$ et $0 < p < 1$), alors $g_{X+Y}(s) = g_X(s)\ g_Y(s) = (ps+1-p)^{n+m}$. Comme la fonction génératrice caractérise la loi, ceci implique que $X+Y \rightsquigarrow \mathcal{B}(n+m;p)$*
- *Si $X \rightsquigarrow \mathcal{P}(\lambda)$ et $Y \rightsquigarrow \mathcal{P}(\eta)$ ($\lambda > 0$; $\eta > 0$), alors $g_{X+Y}(s) = g_X(s)\ g_Y(s) = e^{(\lambda+\eta)(x-1)}$, ce qui implique que $X+Y \rightsquigarrow \mathcal{P}(\lambda+\eta)$*