Chapitre 3 : Variables aléatoires discrètes

Introduction : le nombre de piles obtenus au cours d'une série de n lancers d'une pièce est une grandeur numérique X pouvant prendre les valeurs $0, 1, 2, \dots, n$ mais dont la valeur dépend du hasard i.e. des éventualités $\omega \in \Omega$ où $\Omega = \{P, F\}^n$ est l'espace probabilisé décrivant l'expérience aléatoire "lancer une pièce n fois de suite (avec indépendance des lancers) qui se modélise par un espace produit comme vu dans le chapitre précédant. Cette grandeur X s'appelle une variable aléatoire. Mathématiquement c'est une application

$$X : \Omega \to \{0, 1, 2, \dots, n\}$$

(I) Définition d'une variable aléatoire

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé (modélisant une certaine expérience aléatoire) et

$$X : \Omega \longrightarrow \mathbb{R}$$

une application à valeurs discrètes i.e.

$X(\Omega) = \{x_i \,;\, i \in D\}$ est fini ou dénombrable

Cet ensemble est constitué par les valeurs prises par X.

Définition : On dit que X est une variable aléatoire discrète si

$$\forall x_i \in X(\Omega), \ \{\omega \in \Omega \mid X(\omega) = x_i\} \in \mathcal{F} \quad (*)$$

L'ensemble $\{\omega \in \Omega \mid X(\omega) = x_i\}$ des éventualités pour lesquelles X vaut x_i est noté

$[X = x_i]$

On dit que c'est l'événement X prend la valeur x_i.

Remarque 1 : la condition $(*)$ sert à assurer que les ensembles $[X = x_i]$ $(x_i \in X(\Omega))$ sont des événements. Si Ω est fini ou si $\mathcal{F} = \mathcal{P}(\Omega)$ cette condition est automatiquement vérifiée.

Remarque 2 : Si X est une variable aléatoire discrète, les événements

$[X = x_i]$, $i \in D$

forment un système complet d'événements.

Définition : La suite des nombres

$$p_i = \mathbb{P}(X = x_i) \quad i \in D$$

s'appelle la distribution de probabilité de la variable aléatoire X (on dit aussi loi de probabilité de la v.a. X).

(Noter que pour abréger on notera v.a. pour variable aléatoire)

Supposons que les valeurs prises par la v.a. discrète

X puissent s'écrire dans l'ordre croissant

$x_1 < x_2 < \cdots < x_n < \cdots$

(ce qui n'est pas toujours le cas)

On appelle alors <u>probabilités cumulées</u> la suite $(a_k)_{k \in D}$ des nombres

$$a_k = \mathbb{P}(X \leq x_k) = \sum_{i=1}^{k} p_i .$$

Les probabilités cumulées ont un intérêt pratique car pour des valeurs $x_l < x_m$ prises par X, on a :

$$\mathbb{P}(x_l \leq X \leq x_m) = a_m - a_{l-1} .$$

II Les lois de probabilité classiques.

A Loi de Bernoulli :

Une v.a. discrète X telle que $X(\Omega) = \{0, 1\}$ (i.e. prenant deux valeurs 0 ou 1) s'appelle <u>variable aléatoire de Bernoulli</u>. Sa loi de probabilité est déterminée par la valeur

$$p = \mathbb{P}(X = 1)$$

(puisque forcément $\mathbb{P}(X=0) = 1-p$) qui s'appelle <u>le paramètre de la v.a. de Bernoulli</u>.

<u>Exemple typique</u> : Soit $A \in \mathcal{F}$ un événement. L'indicatrice de A est la fonction

$$\mathbb{1}_A : \omega \longmapsto \mathbb{1}_A(\omega) = \begin{cases} 1 \text{ si } \omega \in A \\ 0 \text{ si } \omega \notin A \end{cases}$$

$\mathbb{1}_A$ est une v.a. de Bernoulli de paramètre $p = \mathbb{P}(A)$

B <u>Loi binomiale</u> :

On considère une expérience aléatoire à deux issues :

S (= succès)

E (= échec)

avec $P(S) = p$ et $P(E) = 1-p$ $(0 \leq p \leq 1)$.

On fait n répétitions indépendantes de cette expérience - ce qu'on modélise par l'espace produit $\Omega = \{S, E\}^n$ muni de la probabilité produit $\mathbb{P} = P^{\otimes n}$

<u>Définition</u> : La variable aléatoire X = nombre total de succès (au cours des n répétitions) est appelée variable aléatoire binomiale de paramètres (n, p).

<u>Proposition</u> : La loi d'une v.a. binomiale de paramètres (n,p) est donnée par :

$$\begin{cases} p_k = \mathbb{P}(X=k) = C_n^k p^k (1-p)^{n-k} \\ k = 0, 1, \ldots, n. \end{cases}$$

<u>dém^n</u> : voir le chapitre 2 où l'on a déjà expliqué la façon de procéder :

$[X=k]$ = l'ensemble des n-uplets composés de k lettres S (et $n-k$ lettres E) qui sont au nombre de C_n^k et tous ces n-uplets ont la même probabilité $p^k(1-p)^{n-k}$

C) Loi géométrique ou loi de Pascal

Etant donné une suite illimitée d'épreuves à 2 issues S(succès) ou E(échec) modélisée par l'espace produit infini comme précédemment, on considère l'instant X du premier succès.

Définition la v.a. X instant du premier succès s'appelle v.a. géométrique (ou de Pascal) de paramètre $p = P(S)$.

Proposition: la loi de probabilité d'une v.a. géométrique est donnée par:

$$p_k = \mathbb{P}(X=k) = (1-p)^{k-1} \cdot p \quad (k \in \mathbb{N}^*)$$

D) Loi de Poisson

Définition: On dit qu'une v.a. X à valeurs dans $\mathbb{N}$ suit la loi de Poisson de paramètre $\lambda > 0$ si

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

Remarque: la loi de Poisson n'apparait pas de façon explicite comme les lois précédentes mais comme une loi approchée comme le montre le résultat fondamental suivant:

Proposition: Soit $n \in \mathbb{N}$ et $0 < p < 1$. Supposons que $n \to +\infty$ et $p \to 0$ de telle sorte que $np \to \lambda > 0$

Alors pour tout $k \in \mathbb{N}$ fixé,

$$C_n^k p^k (1-p)^{n-k} \longrightarrow e^{-\lambda} \frac{\lambda^k}{k!}$$

dém^{n}: $C_n^k p^k (1-p)^{n-k} = \frac{1}{k!} (n-k+1)(n-k+2) \dots n p^k (1-p)^{n-k}$

$$= \frac{1}{k!} (np)^k (1-p)^n \left(1 - \frac{k-1}{n}\right) \dots \left(1 - \frac{2}{n}\right)(1-p)^{-k}$$

si on pose $np = \lambda + \varepsilon$ où $\varepsilon \to 0$ alors.

$$(1-p)^n = \left(1 - \frac{\lambda}{n} - \frac{\varepsilon}{n}\right)^n \to e^{-\lambda} \text{ si } \left(\substack{n \to +\infty \\ \text{et } \varepsilon \to 0}\right)$$

cqfd.

Exemple d'utilisation de ce résultat:

si n est grand, p est petit et $pn \sim \lambda > 0$ alors loi $\mathcal{B}(n,p)$ est approximativement égale à la loi de Poisson de paramètre λ.

Par exemple si $n = 100$ et $p = 0,1$ alors $\mathcal{B}(100, 0.1) \simeq \mathcal{P}(\lambda = 10)$.

III) Espérance et variance et moments

Soit X une v.a. discrète prenant les valeurs x_i avec probabilité p_i, $i \in D$ (fini ou dénombrable)

Définition: si $\sum_{i \in D} p_i |x_i| < +\infty$, on dit que X admet un moment d'ordre 1. Alors le nombre

$$\mathbb{E}(X) = \sum_{i \in D} p_i x_i$$ s'appelle espérance de la v.a. X.

Notez que si D est fini, le moment d'ordre r et l'espérance existent toujours.

- si $\mathbb{E}(X)=0$, on dit que la v.a. X est centrée
- si pour $r>0$, $\mathbb{E}(X^r)$ existe, on dit que c'est le moment d'ordre r de X

Définition : si X a un moment d'ordre 2, le nombre $\operatorname{Var} X = \mathbb{E}((X-\mathbb{E}(X))^2)$ existe et s'appelle variance de la v.a. X.

le nombre $\sigma_X = \sqrt{\operatorname{Var} X}$ s'appelle alors écart type de la v.a. X.

Exercice : si Ω est fini, montrer que

$$\mathbb{E}(X) = \sum_{\omega\in\Omega} X(\omega)\,\mathbb{P}(\omega)$$

(on admettra que cette formule est encore vraie si Ω est dénombrable et si $\sum_{\omega\in\Omega} |X(\omega)|\,\mathbb{P}(\omega) < +\infty$)

Propriétés de l'espérance

Théorème (linéarité) : Soient X et Y des v.a. discrètes ayant une espérance et $a,b\in\mathbb{R}$. Alors la v.a. $aX+bY$ a une espérance et

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

dém$^{\text{n}}$: il suffit de considérer le cas $a=b=1$. Supposons pour simplifier que Ω est discret ; on peut alors appliquer l'exercice et

$$\mathbb{E}(X+Y) = \sum_{\omega\in\Omega} (X(\omega)+Y(\omega))\,\mathbb{P}(\omega) =$$

$$\sum_{\omega\in\Omega} \big(X(\omega)\mathbb{P}(\omega) + Y(\omega)\mathbb{P}(\omega)\big) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\omega) + \sum_{\omega\in\Omega} Y(\omega)\mathbb{P}(\omega)$$

$$= \mathbb{E}(X)+\mathbb{E}(Y) \quad \text{cqfd.}$$

Exercice : faire la démonstration dans le cas général

solution : pour $x\in X(\Omega)$ et $y\in Y(\Omega)$ posons

$$A_{xy} = [X=x]\cap[Y=y].$$

Les événements A_{xy} ($(x,y)\in X(\Omega)\times Y(\Omega)$) forment un système complet d'événements et

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

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

Posons $Z=X+Y$ et supposons démontrée la condition d'absolue convergence :

$$\sum_{z\in Z(\Omega)} |z|\,\mathbb{P}(Z=z) < +\infty.$$

Calculons alors $\sum_{z\in Z(\Omega)} z\,\mathbb{P}(Z=z) =$ ⊕

Pour tout $z \in Z(\Omega)$, posons

$$B_z = \{(x,y) \in X(\Omega) \times Y(\Omega) \mid x+y = z\}$$

Alors

$$(*) = \sum_{z \in Z(\Omega)} z \sum_{(x,y) \in B_z} \mathbb{P}([X=x] \cap [Y=y])$$

$$= \sum_{z \in Z(\Omega)} \sum_{(x,y) \in B_z} (x+y)\, \mathbb{P}([X=x] \cap [Y=y])$$

$$= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x+y)\, \mathbb{P}([X=x] \cap [Y=y])$$

car les ensembles B_z, $z \in Z(\Omega)$ forment une partition de $X(\Omega) \times Y(\Omega)$. D'où

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

$$= \underbrace{\sum_{(x,y)} x\, \mathbb{P}(A_{xy})}_{= \mathbb{E}(X)} + \underbrace{\sum_{(x,y)} y\, \mathbb{P}(A_{xy})}_{\mathbb{E}(Y)}$$

(car $\displaystyle\sum_{(x,y)} x\, \mathbb{P}(A_{xy}) = \sum_{x \in X(\Omega)} x \underbrace{\sum_{y \in Y(\Omega)} \mathbb{P}(A_{xy})}$

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

de même pour la 2e somme).

Remarque : la condition d'absolue convergence se démontre de la même façon mais en majorant $|x+y|$ par $|x| + |y|$.

On notera que cette condition d'absolue convergence sert à justifier la sommation par paquets

$$\sum_{z \in Z(\Omega)} \sum_{(x,y) \in B_z} = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)}.$$

Fonctions déterministes et opérations sur les v.a. discrètes.

Soient X et Y des v.a. discrètes définies sur le même espace probabilisé $(\Omega; \mathcal{F}, \mathbb{P})$ et supposons données des applications

$$f : \mathbb{R} \to \mathbb{R} \quad \text{et} \quad g : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$$

Considérons les applications de Ω dans $\mathbb{R}$:

$f(X) : \omega \mapsto f(X(\omega))$

$g(X,Y) : \omega \mapsto g(X(\omega), Y(\omega))$

Théorème : $f(X)$ et $g(X,Y)$ sont des v.a. discrètes

Démonstration : $f(X)(\Omega)$ et $g(X,Y)(\Omega)$ sont des ensembles finis ou dénombrables (clair) il reste à vérifier la condition de mesurabilité. Par exemple pour $f(X)$:

Soit $y \in f(X)(\Omega)$ une valeur prise par $f(X)$ et soit

$$A_y = \{x \in X(\Omega) \mid f(x) = y\}$$

Alors

$$[f(X)=y] = \{\omega\in\Omega ;\ f(X(\omega))=y\} = \bigcup_{x\in A_y}[X=x]$$

$\in \mathcal{F}$ comme réunion finie ou dénombrable des événements $[X=x]\in\mathcal{F}$.

Même démonstration pour montrer que $g(X,Y)$ vérifie la condition de mesurabilité.

Exemples:

1) $\forall n\in\mathbb{N}$, X^n est une v.a.

2) $\forall \lambda\in\mathbb{R}$, λX est une v.a.

3) $X+Y$, $X-Y$ et XY sont des v.a.

(appliquer le théorème avec la fonction g respectivt $(x,y)\mapsto x+y$ (resp. $x-y$, resp. xy).

Le théorème se généralise à un n-uplet $(X_1,\dots,X_n)$ de v.a. et une fonction $g:\mathbb{R}^n\to\mathbb{R}$. Dans ce cas $g(X_1,\dots,X_n)$ est aussi une v.a. discrète.

Théorème: Soit X une v.a. discrète, $g:\mathbb{R}\to\mathbb{R}$ une application et soit $Y=g(X)$. La v.a. Y a une espérance si et seulement si

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

Dans ce cas, on a

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

Démo: $\forall y\in Y(\Omega)$, posons

$$B_y=\{x\in X(\Omega) ;\ g(x)=y\}.$$

Les ensembles B_y $(y\in Y(\Omega))$ forment une partition de $X(\Omega)$ et on a

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

a) ($\Rightarrow$) supposons $\sum_{x\in X(\Omega)}|g(x)|\,\mathbb{P}(X=x)<+\infty$. Alors

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

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

$$=\sum_{y\in Y(\Omega)}|y|\,\mathbb{P}(Y=y)\ (<+\infty)$$ donc Y a un moment d'ordre 1. En reprenant le calcul précédant en changeant $|y|$ en y et $|g(x)|$ en $g(x)$ on trouve

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

Cqfd.

b) ($\Leftarrow$) la réciproque s'obtient en remontant les calculs précédents.

Applications:

1) Le moment d'ordre $r>0$ de la v.a. discrète X existe si et seulement si $\sum_{x\in X(\Omega)} |x|^r \mathbb{P}(X=x)$

Dans ce cas on a

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

2) Si X a un moment d'ordre 2, $\operatorname{Var} X$ existe et on a

$$\operatorname{Var} X = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \sum_{i\in D} p_i x_i^2 - \left(\sum_{i\in D} p_i x_i\right)^2$$

<u>Remarque</u>: Si X a un moment d'ordre 2, X a aussi un moment d'ordre 1. Plus généralement si X a un moment d'ordre p entier >1, X a aussi des moments d'ordre $p-1, p-2, \dots, 1$ (<u>exercice</u>)

<u>Exercices</u>: 1) Si X est de loi binomiale $\mathcal{B}(n,p)$, X a des moments de tous les ordres. En particulier on a

$\mathbb{E}(X) = np$ et $\operatorname{Var} X = np(1-p)$

2) Si X est de loi de Poisson de paramètre $\lambda>0$, X a des moments de tous les ordres. En particulier on a $\mathbb{E}(X) = \operatorname{Var} X = \lambda$.

3) Jeux équitables : Deux joueurs A et B jouent à un jeu où la probabilité de gagner est égale à p pour A et $1-p$ pour B $(0<p<1)$. Les mises de A et B sont respectivement de s et s' euros. Le vainqueur empoche le total des enjeux. Si X et Y sont les gains de A et B respectivement, le jeu est équitable si $\mathbb{E}(X) = \mathbb{E}(Y)$.
Montrer que c'est le cas si $\frac{s}{p} = \frac{s'}{1-p}$.

4) Le jeu de St Petersbourg : on joue à pile ou face jusqu'à l'apparition du 1er pile. Si cela se produit au n-ième coup, le joueur gagne 2^n euros. Quelle serait être sa mise initiale pour pouvoir jouer à un tel jeu ?

IV La loi des grands nombres

A Variables aléatoires indépendantes.

Soient $X_1, X_2, \dots, X_n$ des v.a. définies sur $(\Omega, \mathcal{F}, \mathbb{P})$.

<u>Définition</u>: les v.a. sont dites <u>indépendantes</u> (dans leur ensemble) si pour tout choix de $a^{(i)} \in X_i(\Omega)$ $(i=1,\dots,n)$, les événements $[X_i = a^{(i)}]$ sont indépendants i.e :

$$\mathbb{P}\left(\bigcap_{i=1}^{n} [X_i = a^{(i)}]\right) = \prod_{i=1}^{n} \mathbb{P}(X_i = a^{(i)})$$

Plus généralement <u>une suite</u> $(X_n)_{n\in\mathbb{N}}$ <u>de v.a. indépendantes</u> est telle que pour tout entier N (fixé), les v.a. $X_1, X_2, \dots, X_N$ sont indépendantes

<u>Exemple typique</u>: des v.a. $X_1, \dots, X_n$ sont indépendantes si elles se rapportent à des expériences aléatoires indépendantes. Ainsi si on fait n répétitions indépendantes d'une même expérience et si X_1 = le résultat du 1er essai, …, X_n = le résultat du $n$ième essai alors $X_1, \dots, X_n$ sont des v.a. indépendantes.

Théorème (espérance d'un produit de v.a. indépendantes)
Si les v.a. $X_1, \dots, X_n$ sont indépendantes et ont un moment d'ordre 1, la v.a. produit
$$Z = X_1 \cdot X_2 \cdots X_n$$
a un moment d'ordre 1 et
$$\mathbb{E}(Z) = \prod_{i=1}^{n} \mathbb{E}(X_i)$$

dém$^{\text{n}}$: (cas $n=2$) on pose $X = X_1$ et $Y = X_2$, $Z = XY$.
Supposons X et Y à valeurs positives.
$$Z(\Omega) = \{ z = xy \mid x \in X(\Omega) \text{ et } y \in Y(\Omega) \}.$$
Pour $z \in Z(\Omega)$, soit
$$A_z = \{ (x,y) \in X(\Omega) \times Y(\Omega) \mid xy = z \}$$
Les ensembles A_z $(z \in Z(\Omega))$ forment une partition de $X(\Omega) \times Y(\Omega)$ et
$$[Z=z] = \bigcup_{(x,y) \in A_z} [X=x] \cap [Y=y]$$
(réunion d'événements 2 à 2 incompatibles). Alors
$$\sum_{z \in Z(\Omega)} z\, \mathbb{P}(Z=z) = \sum_{z \in Z(\Omega)} \left(z \sum_{(x,y) \in A_z} \mathbb{P}(X=x)\mathbb{P}(Y=y) \right)$$
(car les événements $[X=x]$ et $[Y=y]$ sont indépendants.
$$= \sum_{z \in Z(\Omega)} \sum_{(x,y) \in A_z} xy\, \mathbb{P}(X=x)\mathbb{P}(Y=y)$$

$$= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} xy\, \mathbb{P}(X=x)\mathbb{P}(Y=y) \quad \text{(sommation par paquets)}$$
$$= \sum_{x \in X(\Omega)} x\, \mathbb{P}(X=x) \sum_{y \in Y(\Omega)} y\, \mathbb{P}(Y=y)$$
$$= \mathbb{E}(X)\mathbb{E}(Y).$$
Si X et Y de signe quelconque, ce qui précède montre que Z a un moment d'ordre 1 ; on peut reprendre le même calcul pour démontrer la formule. cqfd

Corollaire (Variance d'une somme de v.a. indépendantes)
Soient $X_1, \dots, X_n$ des v.a. indépendantes ayant un moment d'ordre 2. Alors $X_1 + \dots + X_n$ a aussi un moment d'ordre 2 et
$$\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \operatorname{Var} X_i$$

dém$^{\text{n}}$: il suffit de supposer $n=2$. Si $X_1 = X$ et $X_2 = Y$, $X+Y$ a un moment d'ordre 2 car XY a un moment d'ordre 1 d'après le Théorème. Alors
$$\operatorname{Var}(X+Y) = \mathbb{E}\left[\left((X - \mathbb{E}(X)) + (Y - \mathbb{E}(Y)) \right)^2 \right]$$
$$= \operatorname{Var} X + \operatorname{Var} Y + 2\big(\underbrace{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}_{=0 \text{ (par le Thm)}}\big) \quad \text{cqfd}$$

B) La loi des grands nombres

Théorème (inégalité de Bienaymé-Tchebychev)

Soit X une v.a. avec moment d'ordre 2. On note $m = \mathbb{E}(X)$ et $\sigma = \sqrt{\operatorname{Var} X}$. Alors pour tout $a > 0$, on a

$$\mathbb{P}\left(\left|\frac{X-m}{\sigma}\right| \geq a\right) \leq \frac{1}{a^2} \quad (*)$$

Démⁿ: Posons $Y = \frac{X-m}{\sigma}$ cette v.a. est centrée et réduite ($\mathbb{E}(Y)=0$ et $\mathbb{E}(Y^2)=1$).

$$1 = \mathbb{E}(Y^2) = \sum_{y \in Y(\Omega)} y^2 \mathbb{P}(Y=y)$$

$$\geq \sum_{\substack{|y| \geq a \\ y \in Y(\Omega)}} y^2 \mathbb{P}(Y=y) \geq a^2 \underbrace{\sum_{\substack{|y|\geq a \\ y \in Y(\Omega)}} \mathbb{P}(Y=y)}_{a^2 \mathbb{P}(|Y| \geq a)}$$

d'où l'inégalité.

Remarque: Si a est grand, $(*)$ nous dit qu'il très improbable que X s'écarte de m de plus de $a\sigma$.

Théorème (loi faible des grands nombres)

Soit $X_1, \dots, X_n, \dots$ une suite de v.a. indépendantes, de même loi d'espérance m et de variance σ^2. Alors

$$\forall \varepsilon > 0, \lim_{n \to +\infty} \mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \varepsilon\right) = 0$$

Démⁿ: Posons $Y_n = \frac{X_1 + \dots + X_n}{n}$. On a:

$$\mathbb{E}(Y_n) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(X_i) = m$$

$$\operatorname{Var}(Y_n) = \frac{1}{n^2} \sum_{i=1}^{n} \operatorname{Var}(X_i) = \frac{\sigma^2}{n}.$$

L'inégalité de Bienaymé-Tchebychev implique:

$$\mathbb{P}\left(|Y_n - m| \geq \frac{a\sigma}{\sqrt{n}}\right) \leq \frac{1}{a^2};$$

En particulier si $a = \frac{\sqrt{n}\,\varepsilon}{\sigma}$, on obtient

$$\mathbb{P}(|Y_n - m| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0 \text{ si } n \to +\infty.$$

Corollaire: Soit $\mathcal{E}$ une expérience aléatoire et A un événement de probabilité p. Soit $\frac{N_A}{N}$ la fréquence relative de A au cours de N répétitions indépendantes de $\mathcal{E}$. Alors

$$\mathbb{P}\left(\left|\frac{N_A}{N} - p\right| \geq \varepsilon\right) \to 0 \text{ quand } N \to +\infty$$

Démⁿ: Posons

$$X_i = \begin{cases} 1 & \text{si } A \text{ se réalise au } i^e \text{ coup} \\ 0 & \text{sinon} \end{cases}$$

les v.a. $(X_i)_{1 \leq i \leq n}$ sont indépendantes et de même loi de Bernoulli avec $\mathbb{E}(X_i) = p$ et $\operatorname{Var} X_i = p(1-p)$ et on a $X_1 + \dots + X_N = N_A$. Le résultat découle alors du théorème.

Remarque: Une suite (X_n) de v.a. converge en probabilité vers une v.a. X si $\forall \varepsilon > 0$,

$$\mathbb{P}(|X_n - X| \geq \varepsilon) \xrightarrow[n \to +\infty]{} 0$$ (à voir dans un autre chapitre).

© Couples et vecteurs aléatoires discrets.

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et soit $V : \Omega \to \mathbb{R}^p$ une application de Ω dans $\mathbb{R}^p$ telle que $V(\Omega)$ soit un sous-ensemble fini ou dénombrable de $\mathbb{R}^p$.

Définition : on dit que V est un vecteur aléatoire discret s'il vérifie la condition de mesurabilité :

$$\forall v \in V(\Omega),\ [V=v] = \{\omega \in \Omega ;\ V(\omega) = v\} \in \mathcal{F}$$

La famille des nombres $(p_v)_{v \in V(\Omega)}$ où $p_v = \mathbb{P}(V=v)$ s'appelle loi de probabilité du vecteur aléatoire V.

Dans la base canonique de $\mathbb{R}^p$ écrivons $V = (X_1, X_2, \dots, X_p)$ i.e. $V(\omega) = (X_1(\omega), \dots, X_p(\omega))$ au lieu du vecteur aléatoire V, on parle aussi de p-uplet $(X_1, X_2, \dots, X_p)$ de variables aléatoires.

Si $p = 2$, on dit que $V = (X_1, X_2)$ est un couple de variables aléatoires.

Les lois de probabilité des v.a. $X_1, \dots, X_p$ s'appellent les lois marginales de V. On les obtient comme suit :

$$\mathbb{P}(X_i = x_i) = \sum_{\substack{v \in V(\Omega) \\ \text{tq. } v_i = x_i}} p_v \quad \text{(exercice facile)}$$

Par exemple si $V = (X, Y)$ est un couple de v.a. de loi $p_v = p_{(x,y)}$, $(x,y) \in V(\Omega)$; Alors

$$\begin{cases} p_x = \mathbb{P}(X=x) = \displaystyle\sum_{y \in Y(\Omega)} p_{(x,y)} \\ \forall x \in X(\Omega) \end{cases}$$

En effet pour $x \in X(\Omega)$,

$$[X=x] = [X=x] \cap \Big(\underbrace{\bigcup_{y \in Y(\Omega)} [Y=y]}_{\Omega} \Big)$$

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

$$= \bigcup_{y \in Y(\Omega)} [(X,Y) = (x,y)]$$

(réunion d'événements 2 à 2 incompatibles)

D'où

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

(V) Notions sur les chaînes de Markov

(A) Introduction:

On utilise généralement une suite $(X_n)_{n\in D}$ $(D=\{1,2,\dots,N\}$ ou $D=\mathbb{N}^*)$ de variables aléatoires pour modéliser un processus aléatoire. L'ensemble D figure le temps et la v.a. X_n est l'état du processus à l'instant n. Par exemple pour le processus décrivant la fortune d'un joueur au cours d'un jeu, X_n est la fortune à l'instant n. Si le processus décrit le cours d'une action en bourse, X_n est la valeur de l'action à l'instant n.

Les processus auxquels on s'intéresse sont ceux tels que la v.a. X_n "dépend" des v.a. du passé $X_0, X_1, \dots, X_{n-1}$. Dans ce paragraphe on va présenter le cas le plus simple :

X_n dépend seulement de X_{n-1}

et $X_n(\Omega)=E$ ensemble fini indépendant de n qu'on appelera <u>espace des états du processus</u>.

Considérons donc un ensemble fini $E=\{x_1,\dots,x_N\}$ et $(X_n)_{n\in\mathbb{N}}$ une suite de v.a. définies sur un espace probabilisé $(\Omega,\mathcal{F},\mathbb{P})$ est à valeurs dans E. En renommant les éléments de E, on peut supposer que $E=\{1,2,\dots,N\}$ ($N\in\mathbb{N}^*$ fixé)

Notation: Pour alléger les notations on remplacera parfois les signes intersection $\cap$ par une virgule par exemple :

$$[X_1=x_1, X_2=x_2, \dots, X_p=x_p] := [X_1=x_1]\cap[X_2=x_2]\cap\dots\cap[X_p=x_p]$$

Définition : Le processus $(X_n)_{n\in\mathbb{N}}$ à valeurs dans l'espace des états E, est une chaîne de Markov si :

$\forall n\in\mathbb{N}$, $\forall i_0, i_1, \dots, i_{n-2}, i, j \in E$,

$$\mathbb{P}(X_n=j \mid X_{n-1}=i, X_{n-2}=i_{n-2}, \dots, X_0=i_0)$$
$$=\mathbb{P}(X_n=j \mid X_{n-1}=i) = p(i,j)$$

(indépendant de n) cette condition s'appelle <u>propriété de Markov</u>.

<u>Remarque</u> : c'est cet ensemble de conditions qui dit que X_n ne dépend que X_{n-1}.

De plus nous avons une condition supplémentaire :

$$\mathbb{P}(X_n=j \mid X_{n-1}=i) = p(i,j)$$

ne dépend plus de n. Ce nombre $p(i,j)$ s'appelle la <u>probabilité de transition de i vers j</u>.

C'est une probabilité conditionnelle : sachant qu'on est dans l'état i à un certain instant, $p(i,j)$ = la probabilité d'aller en j à l'instant suivant.

Rappel d'une formule importante :

Formule de "l'intersection"

Si $A_1, A_2, \dots, A_n$ sont des événements, alors tels $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) =$$

$$\mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1\cap A_2) \cdots \mathbb{P}(A_n|A_1 \cap \dots \cap A_{n-1})$$

$$\left(= \mathbb{P}(A_1)\prod_{i=2}^{n} \mathbb{P}(A_i|A_1\cap\dots\cap A_{i-1})\right)$$

dém^n : on calcule le 2^e membre :

$$\cancel{\mathbb{P}(A_1)}\frac{\cancel{\mathbb{P}(A_1\cap A_2)}}{\cancel{\mathbb{P}(A_1)}}\frac{\mathbb{P}(\cancel{A_1\cap A_2}\cap A_3)}{\mathbb{P}(\cancel{A_1\cap A_2})} \cdots \frac{\mathbb{P}(A_1\cap\dots\cap A_{n-1}\cap A_n)}{\mathbb{P}(\cancel{A_1\cap\dots\cap A_{n-1}})}$$

$$= \mathbb{P}(A_1\cap A_2 \cap \dots \cap A_n) \quad \underline{\text{cqfd}}.$$

Application : Lois de dimensions finie d'une chaîne de Markov.

Soit $(X_n)_{n\in\mathbb{N}}$ une chaîne de Markov dont on connait la <u>loi initiale</u> i.e la loi de la v.a. X_0 :

$$c_i = \mathbb{P}(X_0 = i), \; i \in E$$

et les probabilités de transition

$$p(i,j) \quad i,j \in E.$$

Alors on peut calculer la loi du vecteur $(X_0, X_1, \dots, X_n)$

Théorème : Sous les hypothèses précédentes $\forall i_0, i_1, \dots, i_n \in E$,

$$\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = c_{i_0}\, p(i_0,i_1)p(i_1,i_2)\cdots p(i_{n-1},i_n)$$

dém^n : Appliquer la formule de l'intersection aux événements

$$A_0 = [X_0 = i_0],$$
$$A_1 = [X_1 = i_1]$$
$$\vdots$$
$$A_n = [X_n = i_n]$$

et en tenant compte du fait que

$$\mathbb{P}(A_i | A_0 \cap \dots \cap A_{i-1}) = \mathbb{P}(A_i | A_{i-1})$$

(propriété de Markov).

B Graphe d'une chaîne de Markov

Soit une chaîne de Markov d'espace des états E et de probabilités de transitions $p(i,j)$, $(i,j)\in E^2$.

Propriétés des nombres $p(i,j)$:

a) Pour tout $(i,j)\in E\times E$, $0 \le p(i,j) \le 1$

b) pour tout $i\in E$, $\sum_{j\in E} p(i,j) = 1$

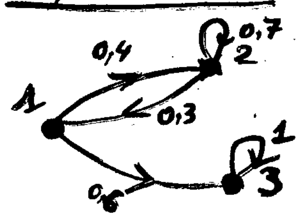
démⁿ : a) évident

b) $\sum_{j\in E} p(i,j) = \sum_{j\in E} \mathbb{P}(X_n=j \mid X_{n-1}=i)$

$= \mathbb{P}\left(\bigcup_{j\in E}[X_n=j] \mid X_{n-1}=i\right) = \mathbb{P}(\Omega \mid X_{n-1}=i) = 1.$

> **Définition** : L'ensemble $\{(i,j)\in E\times E \mid p(i,j)>0\}$ s'appelle graphe de la chaîne de Markov de probabilités de transition $p(i,j)$

Représentation :



Quand $p(i,j)>0$ la flèche entre i et j indique qu'une transition est possible entre i et j

Sur la flèche, on indique la valeur de $p(i,j)$

Ce graphe est un <u>graphe valué</u>

Dans l'exemple ci-dessus l'état 3 est <u>absorbant</u> (un état i est absorbant si $p(i,i) = 1$).

<u>Utilité du graphe</u> : La plupart des questions élémentaires concernent le calcul de la probabilité d'un <u>cheminement</u> donné du processus (X_n). Quand le graphe est simple ces questions sont faciles à résoudre à l'aide du résultat du théorème p.22

Exemples

1) Une personne joue à la roulette avec la stratégie suivante : elle possède 15 euros et veut gagner 1 euro.

Elle mise 1 euro sur le rouge. Si elle gagne elle s'arrête de jouer.

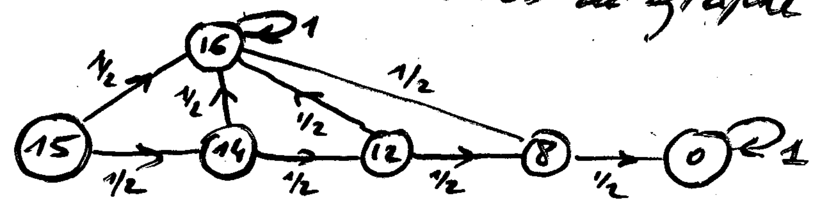
Sinon elle continue à parier sur le rouge en doublant la mise à chaque fois.

Quelle est sa probabilité de gagner ?

Quelle est l'espérance de son gain ?

Quelle est la durée moyenne du jeu ?

La stratégie du joueur peut se modéliser par une chaîne de Markov de graphe :



Les sommets du graphe représentent les états possibles de la fortune du joueur. Les états (16) et (0) sont absorbants (par définition).

a) L'événement "le joueur gagne" se

décompose en les éventualités suivantes (chemins joignant 15 à 16) :

$15 \to 16 \longrightarrow$ probabilité $1/2$

$15 \to 14 \to 16 \longrightarrow$ " $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

$15 \to 14 \to 12 \to 16 \longrightarrow$ " $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}$

$15 \to 14 \to 12 \to 8 \to 16 \to$ " $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{16}$

D'où la probabilité de gagner =

$$\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} = \frac{15}{16}$$

b) Espérance du gain G :

G	1	−15
P	$\frac{15}{16}$	$\frac{1}{16}$

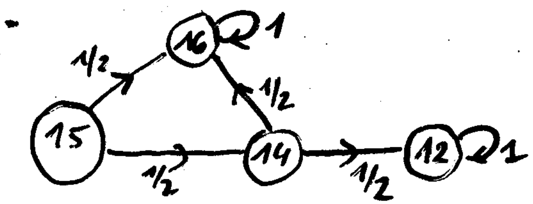
$$E(G) = 1 \cdot \frac{15}{16} - 15 \cdot \frac{1}{16} = 0$$

c) durée moyenne du jeu : c'est la moyenne pondérée des durées de chaque chemin joignant 15 à l'un des états 16 et 0 :

$$E(D) = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 4 \cdot \frac{1}{16} = \frac{15}{8} \simeq 1{,}8$$

2) Même jeu mais le joueur s'arrête s'il n'a pas gagné 1 € au bout de 2 coups :

La chaîne de Markov devient



$\Longrightarrow$ probabilité de gagner $= \frac{1}{2} + \frac{1}{4} = \frac{3}{4}$

espérance de gain $= 1 \cdot \frac{3}{4} - 3 \cdot \frac{1}{4} = 0$

B Matrice d'une chaîne de Markov

On suppose que l'espace des états est $E = \{1, 2, \dots, N\}$

Définition : La matrice de terme général $p(i,j)$ (= la probabilité partant de i d'aller en j à l'instant suivant) est appelée matrice des transitions de la chaîne de Markov

Exemple : pour l'exemple du joueur de roulette si on numérote les états comme suit :

15 ⟶ N°1   12 ⟶ N°3   16 ⟶ N°5

14 ⟶ N°2   8 ⟶ N°4   0 ⟶ N°6

La matrice de la chaîne est donnée par :

$$P = \begin{array}{c} \\ 15 \\ 14 \\ 12 \\ 8 \\ 16 \\ 0 \end{array} \begin{array}{c} \begin{array}{cccccc} 15 & 14 & 12 & 8 & 16 & 0 \end{array} \\ \begin{pmatrix} 0 & 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

Remarques : la somme des coefficients d'une même ligne est toujours égale à 1 car $\sum_{j \in E} p(i,j) = 1$ (propriété déjà vue)

Probabilité d'aller d'un état à un autre en n étapes :

Pour $i, j \in E$, $p(i,j)$ = la probabilité d'aller de i en j à l'instant suivant i.e. en une étape

Soit n un entier fixé et soit

$p^{(n)}(i,j)$ = la probabilité d'aller de i à j en n étapes.

Par exemple sur le graphe de la chaîne de Markov du joueur de roulette, on voit que

$$p^{(2)}(15,12) = 1/4 \qquad p^{(3)}(15,12)=0$$
$$p^{(3)}(15,0) = 0 \qquad p^{(4)}(15,0) = \frac{1}{16}$$

Théorème : La matrice de terme général $p^{(n)}(i,j)$ est la matrice P^n puissance $n$ième de la matrice P.
De plus si $\vec{C} = (c_1, c_2, \dots, c_N)$ est le vecteur ligne loi initiale (i.e. la loi de X_0) alors la loi de X_n est donné par le vecteur ligne $\vec{C}.P^n$

Dém^n : pour $n=2$. On cherche

$$p^{(2)}(i,j) = \mathbb{P}(X_2=j \mid X_0=i).$$

Soit le système complet d'événements $[X_1=\ell]$ $\ell \in \{1,\dots,N\}$. On a

$$p^{(2)}(i,j) = \sum_{\ell=1}^{N} \mathbb{P}(X_2=j, X_1=\ell \mid X_0=i)$$
$$= \sum_{\ell=1}^{N} \frac{\mathbb{P}(X_2=j, X_1=\ell, X_0=i)}{\mathbb{P}(X_0=i)}$$
$$= \sum_{\ell=1}^{N} \frac{\mathbb{P}(X_2=j \mid X_1=\ell, X_0=i)\,\mathbb{P}(X_1=\ell, X_0=i)}{\mathbb{P}(X_0=i)}$$

$$= \sum_{\ell=1}^{N} \mathbb{P}(X_2=j \mid X_1=\ell)\,\mathbb{P}(X_1=\ell \mid X_0=i)$$
$$= \sum_{\ell=1}^{N} p(i,\ell)p(\ell,j) = \text{le coefficient } (i,j) \text{ de la matrice } P^2.$$

Cas général : par récurrence sur n

Loi de X_n :

$$\forall j \in E,\ \mathbb{P}(X_n=j) = \sum_{i=1}^{N} \underbrace{\mathbb{P}(X_n=j \mid X_0=i)}_{p^{(n)}(i,j)}\,\mathbb{P}(X_0=i)$$
$$= \sum_{i=1}^{N} c_i\, p^{(n)}(i,j) = \text{le coefficient d'ordre } j \text{ de}$$

la matrice ligne $\vec{C}P^n$. cqfd.

Cas particulier d'une chaîne de Markov ergodique régulière

Définition : si la matrice P des transitions d'une chaîne de Markov est telle que

$\exists k \in \mathbb{N}^*$ t.q. P^k est à coefficients tous >0. (*)

On dit que la chaîne de Markov est ergodique régulière

La condition (*) signifie qu'en k étapes on peut aller de tout état i en n'importe quel autre état j avec probabilité >0. Il en est de même en nk étapes ($n \in \mathbb{N}^*$). Autrement dit, la chaîne "tourne" sur les différents états "éternellement". Pourtant il y a un phénomène de stabilisation important qui se produit :

la loi de probabilité de X_n tend vers une loi limite qu'on appelle <u>loi stationnaire</u> de la chaîne de Markov

---

<u>Théorème</u> : Pour une chaîne de Markov régulière ergodique, de matrice P, on a :

1) $\lim_{n \to \infty} P^n = A$ existe

2) Toutes les lignes de A sont identiques

$$A = \begin{pmatrix} a_1 \cdots a_N \\ a_1 \cdots a_N \\ \vdots \\ a_1 \cdots a_N \end{pmatrix}$$

3) Les composantes du vecteur $\vec{\alpha} = (a_1, \dots, a_N)$ sont toutes > 0.

4) Quel que soit la loi initiale $\vec{c}$, la loi de X_n $\vec{c}P^n$ tend vers $\vec{\alpha}$ quand $n \to +\infty$

5) Le vecteur $\vec{\alpha}$ est l'unique solution de $\vec{\alpha}P = \vec{\alpha}$ qui soit une loi de probabilité.

<u>dém</u> : plus tard dans le cours. Résultat admis pour le Master MME.

VI Fonctions génératrices

Dans le cas des v.a. discrètes à valeurs entières i.e $X(\Omega) \subset \mathbb{N}$ il existe un outil puissant que nous allons présenter dans ce paragraphe.

Soit $X : \Omega \to \mathbb{N}$ une v.a. à valeurs entières

Définition: On appelle <u>fonction génératrice</u> de la v.a. X, la série entière

$$G_X(t) = \sum_{n=0}^{+\infty} p_n t^n \qquad (t \in \mathbb{R})$$

où $p_n = \mathbb{P}(X = n)$ si $n \in X(\Omega)$ et $p_n = 0$ sinon.

<u>Remarque 1</u> le rayon de convergence de G_X est au moins égal à 1 car la série est convergente pour $t = 1$ donc la série G_X converge absolument pour $|t| < 1$.

<u>Remarque 2</u>: on peut écrire $G_X(t) = \mathbb{E}(t^X)$ car pour tout $t \in]0,1[$ la v.a. t^X existe.

L'utilité de la notion de fonction génératrice réside dans le théorème suivant :

Théorème: Soient G_X et G_Y les fonctions génératrices de deux v.a. <u>indépendantes</u> X et Y à valeurs entières. Alors la v.a. $X+Y$ a une fonction génératrice donnée par :

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

<u>démo</u>: Posons $\sum_i a_i t^i = G_X(t)$ et $\sum_j b_j t^j = G_Y(t)$

On a

$$\mathbb{P}(X+Y=k) = \sum_{j=0}^{k} \mathbb{P}(X+Y=k \mid Y=j)\,\mathbb{P}(Y=j)$$

$$= \sum_{j=0}^{k} \mathbb{P}(X=k-j \mid Y=j)\,\mathbb{P}(Y=j) \quad \text{(form. prob. totale)}$$

$$= \sum_{j=0}^{k} \mathbb{P}(X=k-j)\,\mathbb{P}(Y=j) \quad \text{(indépendance de } X \text{ et } Y)$$

$$= \sum_{j=0}^{k} a_{k-j}\, b_j$$

= le coefficient du terme t^k dans le développement du produit $G_X(t)\,G_Y(t)$. <u>CQFD</u>

<u>autre démo</u> (plus synthétique) pour tout t fixé :

$$G_{X+Y}(t) = \mathbb{E}(t^{X+Y}) = \mathbb{E}(t^X t^Y) =$$

$$= \mathbb{E}(t^X)\,\mathbb{E}(t^Y) \quad \text{(indépendance des v.a. } t^X \text{ et } t^Y)$$

$$= G_X(t)\,G_Y(t).$$

<u>Exemple</u>: Une urne contient 4 boules numérotées 0, 1, 1, 2. On effectue <u>n tirages avec remise</u> dans cette urne. Soit S_n la somme des numérotés tirés. Trouver la loi de probabilité de la v.a. S_n.

<u>Solution</u> : $G_{S_1}(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2 = \left(\frac{1+t}{2}\right)^2$

Les tirages étant avec remise, S_n est la somme de n v.a. indépendantes de même loi que S_1, donc

$$G_{S_n}(t) = \left(G_{S_1}(t)\right)^n \text{ (d'après le Théorème)}$$
$$= \left(\frac{1+t}{2}\right)^{2n}$$

et en développant suivant les puissances de t :

$$\mathbb{P}(S_n = k) = \frac{1}{2^{2n}} C_{2n}^{k} \quad (k = 0, 1, \dots, 2n).$$

Théorème : Soit X une v.a. à valeurs entières et G_X sa fonction génératrice. Si X a un moment d'ordre 2, alors les dérivées à gauche $G_X'(1)$ et $G_X''(1)$ existent en $t=1$ et

$$\mathbb{E}(X) = G_X'(1)$$
$$\mathrm{Var}(X) = G_X''(1) + G_X'(1) - \left(G_X'(1)\right)^2$$

Réciproquement si G_X est 2 fois dérivable en $t=1$, X a un moment d'ordre 2 et les formules précédentes sont valables (avec des vrais dérivées en $t=1$).

dém^{n} : on peut dériver formellement terme à terme la série entière $G_X(t)$ ce qui donne :

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

$$G_X''(t) = \sum_{n=2}^{+\infty} n(n-1) p_n t^{n-2}$$

Si X a un moment d'ordre 2, les dérivées à gauche $G_X'(1)$ et $G_X''(1)$ existent d'après le lemme du prolongement d'Abel. D'où les formules annoncées :

$$\begin{cases} G_X'(1) = \sum_{n=1}^{+\infty} n p_n \\ \qquad = \mathbb{E}(X) \end{cases} \qquad \begin{cases} G_X''(1) = \sum_{n=2}^{+\infty} n(n-1) p_n \\ \qquad = \mathbb{E}(X(X-1)) \\ \qquad = \mathbb{E}(X^2) - \mathbb{E}(X) \end{cases}$$

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

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

La réciproque est facile.

Application : Somme d'un nombre aléatoire de variables aléatoires.

Soit $X_1, X_2, \dots, X_i, \dots$ une suite de v.a. à valeurs entières indépendantes et de même loi :

$$G(t) = \sum_{n=0}^{+\infty} p_n t^n$$

Soit N une v.a. à valeurs entières et indépendante de la suite $(X_i)_{i \in \mathbb{N}}$.

On définit la "somme"

$$S_N = \sum_{i=1}^{N} X_i$$

comme suit :

$$\forall k \in \mathbb{N},\ \forall \omega \in [N=k],\ S_N(\omega) = S_k(\omega)$$

(avec la convention $S_0 = 0$).

On admettra que S_N est une v.a.

Proposition : la v.a. S_N a une fonction génératrice donnée par la formule :

$$G_{S_N}(t) = H(G(t)) = (H \circ G)(t) \quad \boxed{\text{où } H \text{ est la } f^{n} \text{ génératrice de } N}$$

et on peut calculer $\mathbb{E}(S_N)$ et $Var(S_N)$ par les formules

$$\mathbb{E}(S_N) = \mathbb{E}(N)\mathbb{E}(X_1)$$

$$Var(S_N) = \mathbb{E}(N)Var(X_1) + (\mathbb{E}(X_1))^2 Var(N)$$

(si évidemment les X_i et N ont un moment d'ordre 2)

<u>Dém^n</u> : Notons p_j^{*n} le coefficient de t^j dans $G_{S_n}(t)$ (n entier fixé). On a :

$$\mathbb{P}(S_N = j) = \sum_{n=0}^{+\infty} \mathbb{P}(X_1 + \cdots + X_n = j)\,\mathbb{P}(N = n)$$

$$= \sum_{n=0}^{+\infty} q_n\, p_j^{*n}$$

D'où
$$G_{S_N}(t) = \sum_{j=0}^{+\infty} \left(\sum_{n=0}^{+\infty} q_n p_j^{*n} \right) t^j$$

$$= \sum_{n=0}^{+\infty} \left(\sum_{j=0}^{+\infty} p_j^{*n} t^j \right) q_n$$

$$= \sum_{n=0}^{+\infty} q_n\, G_{S_n}(t)$$

$$= \sum_{n=0}^{+\infty} q_n\, (G(t))^n$$

$$= H(G(t))$$

Alors on a $\mathbb{E}(S_N) = H'(1)G'(1) = \mathbb{E}(N)\mathbb{E}(X_1)$

pour $Var(S_N)$ <u>exercice</u>.