

# Moments d'une variable aléatoire discrète

## Espérance d'une variable aléatoire réelle ou complexe

1. Espérance pour  $X$  variable aléatoire à valeurs dans  $\mathbb{R}^+ \cup \{+\infty\}$ .
2. Pour une variable aléatoire à valeurs dans  $\mathbb{N} \cup \{+\infty\}$ , formule de l'espérance par antirépartition.
3. Existence, et calcul le cas échéant, de l'espérance d'une variable aléatoire complexe  $X$ . Ensemble  $L^1$ .
4. Variables centrées.
5. Espérance d'une variable géométrique, d'une variable de Poisson.
6. Formule de transfert.
7. Linéarité, positivité, croissance, inégalité triangulaire. Caractérisation des variables aléatoires à valeurs dans  $\mathbb{R}^+$  d'espérance nulle.
8. Théorème d'existence d'une espérance par domination.
9. Si  $X$  et  $Y$  sont dans  $L^1$  et indépendantes, alors  $XY$  est dans  $L^1$  et :  $E(XY) = E(X) E(Y)$ . Extension au cas de  $n$  variables aléatoires.

## Variance d'une variable aléatoire réelle, écart type et covariance

10. Si  $E(X^2) < +\infty$ , on note  $X \in L^2$ . Si  $E(X^2) < +\infty$ ,  $X$  est d'espérance finie.
11. Inégalité de Cauchy-Schwarz : si  $X$  et  $Y$  sont dans  $L^2$ ,  $XY$  est dans  $L^1$  et  $|E(XY)| \leq \sqrt{E(X^2)} \sqrt{E(Y^2)}$ . Cas d'égalité.
12. Pour  $X \in L^2$ , variance et écart type de  $X$ . Variables réduites. Caractérisation des variables aléatoires de variance nulle.
13. Formule de Koenig-Huygens.
14. Relation  $V(aX + b) = a^2 V(X)$ . Si  $\sigma(X) > 0$ , la variable aléatoire  $\frac{X - E(X)}{\sigma(X)}$  est centrée réduite.
15. Variance d'une variable géométrique, d'une variable de Poisson.
16. Covariance de deux variables aléatoires de  $L^2$ . Relation  $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$ . Cas de variables indépendantes.
17. Variance d'une somme de  $n$  variables aléatoires.

## Inégalités probabilistes et loi faible des grands nombres

18. Inégalité de Markov.
19. Inégalité de Bienaymé-Tchebychev.
20. Loi faible des grands nombres.

## Fonctions génératrices

21. Fonction génératrice  $G_X$  de la variable aléatoire  $X$  à valeurs dans  $\mathbb{N}$ .
22. Continuité de  $G_X$ .
23. Détermination de la loi de  $X$  par  $G_X$ .
24. La variable aléatoire  $X$  est d'espérance finie si et seulement si  $G_X$  est dérivable en 1 ; dans ce cas  $E(X) = G_X'(1)$ . La démonstration de la réciproque n'est pas exigible.
25. Utilisation de  $G_X$  pour le calcul de  $E(X)$  et  $V(X)$ .
26. Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.
27. Fonction génératrice d'une somme finie de variables aléatoires indépendantes à valeurs dans  $\mathbb{N}$ .

---

Dans ce chapitre, **toutes les variables aléatoires sont discrètes** et définies sur un espace probabilisé  $(\Omega, \mathcal{A}, P)$ . Elles sont à valeurs dans  $\mathbb{C}$ ,  $\mathbb{R}$  ou  $\mathbb{N}$  selon les paragraphes.

# 1 Espérance d'une variable aléatoire réelle ou complexe

## 1.1 existence et calcul

En MPSI, on définit l'espérance d'une variable aléatoire réelle  $X$  à support fini comme la moyenne des valeurs prises par  $X$ , pondérées par leur probabilité d'apparition :

$$E(X) = \sum_{k=1}^n x_k P(X = x_k) = \sum_{x \in X(\Omega)} x P(X = x)$$

L'espérance s'interprète naturellement dans les jeux de hasard comme le gain espéré, d'où l'appellation.

Par exemple, pour  $X$  de loi présentée dans le tableau ci-contre,

$$E(X) =$$

valeur $k$ du gain $X$	-1	1000
$P(X = k)$	$\frac{999}{1000}$	$\frac{1}{1000}$

Exercice 1 : Une urne contient  $n$  boules numérotées de 1 à  $n$ . On tire une boule dans cette urne et on relève son numéro. Soit  $X$  la variable aléatoire égale au numéro relevé. Donner l'espérance de  $X$ .

### Définition 1 – Espérance – cas positif

Quand  $X$  est une variable aléatoire réelle positive, on définit son espérance dans  $[0, +\infty]$  par :

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

### Définition 2 – Espérance – cas sommable

Soit  $X$  est une variable aléatoire à valeurs dans  $\mathbb{C}$ . On dit que  $X$  est *d'espérance finie*, ou encore que  $X$  *admet une espérance*, lorsque la famille  $(xP(X = x))_{x \in X(\Omega)}$  est sommable. Dans ce cas, l'espérance de  $X$  est :

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

et on note  $X \in L^1$ .

Remarques :

- L'espérance de  $X$  ne dépend que de la loi de  $X$ . On pourrait parler de l'espérance d'une loi...
- Toute variable aléatoire finie est d'espérance finie.
- Lorsque  $X(\Omega)$  est dénombrable, on peut écrire  $X(\Omega) = \{x_n, n \in \mathbb{N}\}$ .  $X$  est d'espérance finie si, et seulement si, la série  $\sum x_n P(X = x_n)$  converge absolument. Dans ce cas,

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

et cette quantité est (heureusement !) indépendante de la numérotation choisie pour les valeurs prises par  $X$ .

- Lorsque  $X \geq 0$ , pour montrer que  $X$  est d'espérance finie, on se contente de montrer que  $E(X) < +\infty$ .
- On remarque que  $X$  est d'espérance finie si et seulement si  $|X|$  est d'espérance finie.
- Une variable aléatoire d'espérance nulle est dite *centrée*.

### Propriété 1

Toute variable aléatoire bornée est d'espérance finie.

Nous généraliserons ce résultat avec le « théorème d'existence d'une espérance par domination ».

### Propriété 2 – lois usuelles ♡

- On rappelle que l'espérance de la loi binomiale  $\mathcal{B}(n, p)$  est  $np$ , et l'espérance de  $\mathbf{1}_A$  est  $P(A)$ .
- $X$  de loi géométrique de paramètre  $p$  est d'espérance finie égale à  $\frac{1}{p}$ .
- $X$  de loi de Poisson de paramètre  $\lambda$  est d'espérance finie égale à  $\lambda$ .

Nous considérons souvent des transformées de la ou des variables initiales  $(X^2, e^{tX}, X^k, XY^2, \min(X, Y), \dots)$ . Le théorème qui suit est fondamental : il permet d'étudier les espérances de ces transformées sans chercher les lois de ces transformées. Il agit comme un *transfert* de la loi de  $X$  vers les espérances de ces transformées.

### Théorème 1 – formule de transfert – admis

Soit  $f : X(\Omega) \rightarrow \mathbb{C}$ . La variable aléatoire  $f(X)$  est d'espérance finie si, et seulement si, la famille  $(f(x)P(X = x))_{x \in X(\Omega)}$  est sommable. On a alors :

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

Le plus souvent,  $X(\Omega) \subset \mathbb{C}$ . Mais on peut aussi choisir de considérer deux variables discrètes complexes  $(X_1, X_2)$  et  $X = (X_1, X_2)$ . À ce moment-là,  $X(\Omega) \subset \mathbb{C}^2$ , et la formule de transfert devient :

$$E(f(X_1, X_2)) = \sum_{(x_1, x_2) \in (X_1, X_2)(\Omega)} f(x_1, x_2)P(X_1 = x_1, X_2 = x_2) \text{ (sous réserve de sommabilité)}$$

Exercice 2 : Soit  $X$  de loi de Poisson de paramètre  $\lambda$  et  $t$  réel. Montrer que  $Y = e^{itX}$  est d'espérance finie et calculer son espérance.

Exercice 3 : Calculer  $E(\frac{1}{X})$  pour  $X$  suivant une loi géométrique de paramètre  $p$ .

Exercice 4 : On donne, par l'intermédiaire d'un tableau, la loi du couple  $(X, Y)$ .

$X \setminus Y$	0	1	2	loi de $X$
-1	0	3/16	0	3/16
0	9/16	0	1/16	10/16
1	0	3/16	0	3/16
loi de $Y$	9/16	6/16	1/16	

Calculer  $E(XY)$ .

### Propriété 3 – calcul de l'espérance par antirépartition

Ici, on suppose que  $X$  est à valeurs dans  $\mathbb{N} \cup \{+\infty\}$ .

$$E(X) = \sum_{n=0}^{+\infty} P(X > n) = \sum_{n=1}^{+\infty} P(X \geq n)$$

Exercice 5 : Retrouver, avec cette formule d'antirépartition, l'espérance de  $X$ , variable aléatoire suivant une loi géométrique de paramètre  $p$ .

## 1.2 propriétés de l'espérance

### Théorème 2 – théorème de domination

Si  $X$  et  $Y$  sont deux variables aléatoires vérifiant  $0 \leq |X| \leq Y$ . Si  $Y$  est d'espérance finie, alors  $X$  est aussi d'espérance finie.

### Propriété 4

Dans cette propriété,  $X$  et  $Y$  sont deux variables aléatoires discrètes à valeurs dans  $\mathbb{R}$  ou  $\mathbb{C}$  (en fonction de ce qui a du sens!).

**Linéarité de l'espérance :** Soient  $X$  et  $Y$  deux variables aléatoires d'espérance finie.

Pour  $\lambda, \mu \in \mathbb{C}$ , la variable aléatoire  $\lambda X + \mu Y$  est d'espérance finie et on a :

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y)$$

**Action « centrer » :** si  $X$  est d'espérance finie,  $X - E(X)$  est centrée.

**Positivité de l'espérance :** si  $X \geq 0$ , alors  $E(X) \geq 0$ . De plus,

$$\begin{cases} X & \geq 0 \\ E(X) & = 0 \end{cases} \Rightarrow (X = 0 \text{ presque sûrement})$$

**Croissance de l'espérance :** si  $X$  et  $Y$  sont d'espérance finie et  $X \leq Y$ , on a  $E(X) \leq E(Y)$ .

**Inégalité triangulaire :**  $X$  est d'espérance finie si, et seulement si,  $|X|$  est d'espérance finie et dans ce cas, on a

$$|E(X)| \leq E(|X|)$$

La propriété de linéarité nous dit que l'espérance est une forme linéaire sur  $L^1$ . La propriété de linéarité est très utile. On peut la généraliser à  $n$  variables aléatoires.

### Propriété 5 – linéarité de l'espérance

Soit  $(X_1, \dots, X_n)$  une famille de variables aléatoires de  $L^1$ . On a :

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

Supposons que l'expérience consiste à tirer avec remise des boules dans une urne qui contient des boules vertes, blanches et rouges.  $X$  est la variable aléatoire égale au nombre de tirages à effectuer pour obtenir au moins une fois chaque couleur.

Après un calcul d'espérance, vous trouvez  $E(X) = \frac{4}{3}$ . Qu'écrivez-vous sur votre copie ?

Exercice 6 : On suppose que  $X_1, X_2, \dots, X_n$  sont des variables aléatoires positives d'espérance finie. Montrer, en utilisant le théorème de domination, que  $Z = \inf(X_1, X_2, \dots, X_n)$  et  $T = \sup(X_1, X_2, \dots, X_n)$  sont aussi d'espérance finie.

Exercice 7 :  $n$  personnes se rendent dans un cinéma et chacune choisit l'une des trois salles disponibles, de manière équiprobable. Pour  $i \in \llbracket 1, 3 \rrbracket$ , on note :

$X_i$  la variable aléatoire égale au nombre de personnes choisissant la salle numérotée  $i$

$T_i = \mathbb{1}_{(X_i=0)}$

et on considère la variable aléatoire  $T$  égale au nombre de salles vides.

Écrire  $T$  à l'aide de données de l'énoncé et donner alors  $E(T)$ . On remarquera qu'on n'a pas du tout cherché la loi de  $T$  pour cela !

### Propriété 6

Soient  $X$  et  $Y$  deux variables aléatoires d'espérance finie et **indépendantes**. On a

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

Plus généralement, si  $X_1, X_2, \dots, X_n$  sont d'espérance finie et indépendantes,  $X_1 X_2 \dots X_n$  est d'espérance finie et on a :

$$E(X_1 X_2 \dots X_n) = E(X_1) E(X_2) \dots E(X_n)$$

Exercice 8 : On considère une suite de variables aléatoires  $(X_n)_{n \in \mathbb{N}^*}$ , toutes définies sur le même espace probabilisé, mutuellement indépendantes et suivant la loi, dite de Rademacher de paramètre  $p$  (avec  $0 < p < 1$ ), définie par :

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

On considère la variable aléatoire  $T_n = \prod_{k=1}^n X_k$ .

1. Déterminer l'ensemble des valeurs prises par  $T_n$  puis calculer  $E(T_n)$  et en déduire une relation entre  $P(T_n = 1)$  et  $P(T_n = -1)$ .
2. En déduire la loi de  $T_n$ .

## 2 Variance d'une variable aléatoire réelle

### 2.1 variables aléatoires dont le carré est d'espérance finie

La notation  $X \in L^2$  signifie que  $X^2$  est d'espérance finie. Conformément au programme, on ne soulèvera aucune difficulté quant à la définition précise de  $L^2$ .

### Propriété 7

Si  $E(X^2) < +\infty$ , alors  $X$  est d'espérance finie. Dit autrement,

$$X \in L^2 \Rightarrow X \in L^1$$

### Propriété 8 – inégalité de Cauchy-Schwarz

Si  $X$  et  $Y$  sont dans  $L^2$ ,  $XY$  est dans  $L^1$  et on a :

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

Il y a égalité si et seulement si  $X$  et  $Y$  sont proportionnelles presque sûrement.

## 2.2 définition et propriétés de la variance

### Définition 3

Pour  $X \in L^2$ , on appelle *variance* de  $X$  le réel positif donné par

$$V(X) = E\left((X - E(X))^2\right)$$

et *écart-type* de  $X$  le réel  $\sigma(X) = \sqrt{V(X)}$ .

La variance est la « moyenne » du carré de la distance de  $X$  à  $E(X)$ , et mesure donc la dispersion de  $X$  à sa moyenne.

### Propriété 9

Soit  $X \in L^2$ .

- Pour  $a$  et  $b$  réels,  $aX + b \in L^2$  et on a  $V(aX + b) = a^2V(X)$ .
- $V(X) = 0$  si et seulement si  $X$  est une variable aléatoire presque-sûrement constante.

### Propriété 10 – formule de Kœnig-Huygens

$$V(X) = E(X^2) - E(X)^2$$

Une variable aléatoire d'espérance valant 0 et de variance égale à 1 est dite *centrée réduite*.

### Propriété 11

Pour  $X \in L^2$  de variance non nulle, la variable aléatoire  $\frac{X - E(X)}{\sqrt{V(X)}}$  est une variable aléatoire centrée et réduite.

### Propriété 12 – lois usuelles ♡

- On rappelle que la variance de la loi binomiale  $\mathcal{B}(n, p)$  est  $np(1 - p)$ .
- $X$  de loi géométrique de paramètre  $p$  est de variance finie égale à  $\frac{1-p}{p^2}$ .
- $X$  de loi de Poisson de paramètre  $\lambda$  est de variance finie égale à  $\lambda$ .

## 3 Covariance de deux variables aléatoires de $L^2$

### Définition - propriété 1

Pour  $X$  et  $Y$  dans  $L^2$ , la variable aléatoire  $(X - E(X))(Y - E(Y))$  est d'espérance finie et cette espérance est la *covariance* de  $X$  et  $Y$  :

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

Dans la pratique, on peut calculer la covariance de  $X$  et  $Y$  par la formule de Kœnig-Huygens suivante :

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

Remarques :

- Pour  $X \in L^2$ , on a  $\text{Cov}(X, X) = V(X)$ .
- L'application  $(X, Y) \mapsto \text{Cov}(X, Y)$  est une forme bilinéaire symétrique positive sur  $L^2$  (mais pas définie positive). Par exemple,

$$\text{Cov}(5X + 3Y, Z) =$$

$$\text{Cov}(X, X + 1) =$$

- Lorsque  $\text{Cov}(X, Y) = 0$ , on dit que  $X$  et  $Y$  sont *décorrélées*.
- Si  $X$  et  $Y$  sont indépendantes, on a vu que  $E(XY) = E(X)E(Y)$ , donc  $\text{Cov}(X, Y) = 0$ . Vous avez vu en MPSI que la réciproque est fausse.

$X$  et  $Y$  sont indépendantes



$X$  et  $Y$  sont décorréliées,  $\text{Cov}(X, Y) = 0$

Pour un contre-exemple, on peut reprendre le tableau de loi conjointe en page 3.

— **Une co-variation**

Il s'avère que : «  $\text{Cov}(X, Y) > 0$  signifie que  $X$  et  $Y$  ont tendance à être de même variation », tandis que «  $\text{Cov}(X, Y) < 0$  signifie que  $X$  et  $Y$  ont tendance à être de variation opposée ».

Exercice 9 : On effectue une suite infinie de lancers d'une pièce qui amène Pile avec la probabilité  $p$  ( $p \in ]0, 1[$ ) et Face avec la probabilité  $q = 1 - p$ . On dit que la première série est de longueur  $n$  ( $n \geq 1$ ) si les  $n$  premiers lancers ont amené le même côté et le  $(n + 1)^{\text{e}}$  a amené l'autre côté. On note  $L_1$  la longueur de cette série. On définit de même la longueur  $L_2$  de la deuxième série.

1. Déterminer la loi de  $L_1$  et son espérance.
2. Donner la loi du couple  $(L_1, L_2)$ .
3. Donner la loi de  $L_2$ . Calculer  $E(L_2)$ .
4. On admet que  $\text{Cov}(L_1, L_2)$  existe. Calculer  $\text{Cov}(L_1, L_2)$ .
5. Étudier alors l'indépendance de  $L_1$  et  $L_2$ .

Propriété 13

Soient  $X$  et  $Y$  dans  $L^2$ . Alors  $X + Y$  est dans  $L^2$  et on a :

$$V(X + Y) = V(X) + V(Y) + 2 \text{Cov}(X, Y)$$

Par exemple,

$$V(X - Y) =$$

$$V(aX + bY) =$$

### Propriété 14

Pour toute famille  $(X_1, \dots, X_n)$  de variables aléatoires de  $L^2$ , la variable aléatoire  $X_1 + X_2 + \dots + X_n$  est dans  $L^2$  et on a :

$$V(X_1 + X_2 + \dots + X_n) = \sum_{(i,j) \in \llbracket 1, n \rrbracket^2} \text{Cov}(X_i, X_j)$$

$$V(X_1 + X_2 + \dots + X_n) = \sum_{k=1}^n V(X_k) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

Plutôt que de s'emmêler dans la formule, on peut placer les covariances dans la matrice (dite *matrice de variance-covariance*) suivante

$$C = \begin{pmatrix} V(X_1) & \text{Cov}(X_1, X_2) & \dots & \text{Cov}(X_1, X_n) \\ \text{Cov}(X_1, X_2) & V(X_2) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \text{Cov}(X_{n-1}, X_n) \\ \text{Cov}(X_1, X_n) & \dots & \text{Cov}(X_{n-1}, X_n) & V(X_n) \end{pmatrix}$$

Dans le cas où les  $X_k$  sont deux à deux non corrélées, pour  $i \neq j$ , on a  $\text{Cov}(X_i, X_j) = 0$  et la matrice de variance-covariance de  $X$  est diagonale.

La somme de tous les coefficients de cette matrice est la variance de  $X_1 + X_2 + \dots + X_n$ .

Exercice 10 : Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes suivant toutes la même loi de Bernoulli de paramètre  $p \in ]0, 1[$ . On pose  $Y_n = X_n X_{n+1}$ .

1. Déterminer la loi de  $Y_n$ .
2. Remplir la matrice de variance-covariance (définie ci-dessus) du vecteur  $(Y_1, Y_2, \dots, Y_n)$ .
3. En déduire la variance de  $Z = Y_1 + Y_2 + \dots + Y_n$ .

## 4 Inégalités probabilistes et loi faible des grands nombres

### 4.1 inégalités de concentration

Les *inégalités de concentration* sont des inégalités qui majorent la probabilité qu'une variable aléatoire dévie, s'écarte, d'une certaine valeur (en général son espérance).

#### Théorème 3 – Inégalité de Markov

Soit  $X$  une variable aléatoire réelle, à valeurs positives. On a

$$\forall a > 0, \quad P(X \geq a) \leq \frac{E(X)}{a} \quad (\text{inégalité dans } \mathbb{R}^+ \cup \{+\infty\})$$

Sans l'hypothèse de positivité de  $X$ , on pourra écrire :

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

#### Théorème 4 – Inégalité de Bienaymé-Tchebychev

Soit  $X$  une variable aléatoire réelle de  $L^2$ . On a

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



Donnons des exemples d'inégalités de concentration qu'on peut obtenir avec l'inégalité de Bienaymé-Tchebychev. Soit  $X$  d'espérance  $m$  et de variance  $\sigma^2$ . On applique l'inégalité de Bienaymé-Tchebychev avec  $\varepsilon = a\sigma$  où  $a > 0$ . On obtient  $P(|X - m| \geq a\sigma) \leq \frac{1}{a^2}$ . En particulier :

- $P(|X - m| \geq 2\sigma) \leq \frac{1}{4}$  (il y a moins de 25% de chance que  $X$  s'écarte de sa moyenne de plus de  $2\sigma$ )
- $P(|X - m| \geq 10\sigma) \leq \frac{1}{100}$  (il y a moins de 1% de chance que  $X$  s'écarte de sa moyenne de plus de  $10\sigma$ ).

## 4.2 loi faible des grands nombres

### Théorème 5 – Loi faible des grands nombres

Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes de même loi et de variance finie. Alors, pour tout  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) = 0$$

où  $S_n = \sum_{k=1}^n X_k$  et  $m = E(X_1)$ .

Effectuons une suite d'expériences de Bernoulli indépendantes et notons  $X_i$  la variable aléatoire égale à 1 si la  $i^{\text{ième}}$  expérience donne un succès et valant 0 si c'est un échec.

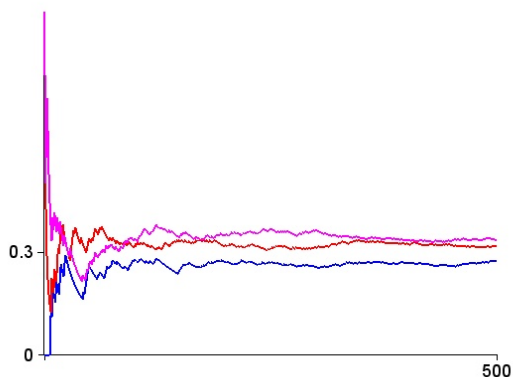
$X_1 + X_2 + \dots + X_n$  est égale au nombre de succès dans les  $n$  premières expériences.

La probabilité empirique de succès (ou *fréquence* de succès) au bout de  $n$  expériences est une réalisation de la variable aléatoire  $\frac{X_1 + X_2 + \dots + X_n}{n}$  avec les  $X_i$  indépendantes d'espérance  $p$  et de variance  $p(1-p)$ . La loi faible des grands nombres traduit l'intuition :

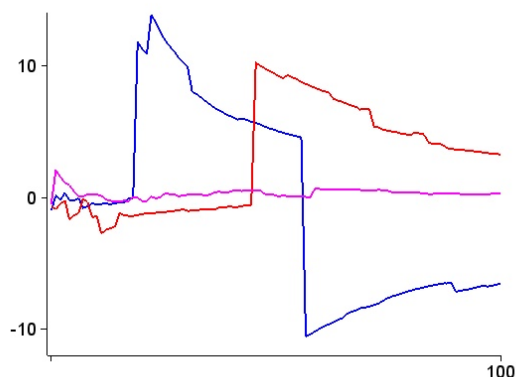
la probabilité empirique de succès tend vers la probabilité théorique de succès ( $p$ ) quand le nombre d'expériences tend vers l'infini.

En 1906, le statisticien anglais Francis Galton (voulant en fait démontrer l'intelligence des experts) demanda à 787 personnes de deviner le poids d'un boeuf exposé dans une foire. Alors que la plupart des réponses prises individuellement étaient largement inexactes, la moyenne des estimations fut exacte à 1% près. Cette idée qu'un grand nombre de personnes peuvent (parfois) donner ensemble une réponse d'expert est le concept de *sagesse des foules*.

Visualisons les évolutions de  $\left(\frac{S_n(\omega)}{n}\right)_{1 \leq n \leq 500}$  pour 3 résultats  $\omega$ , où  $(X_i)_{i \in \mathbb{N}^*}$  est une suite i.i.d. de même loi que  $X$ .



$X$  suit la loi de Bernoulli  $b(0.3)$



$X$  n'a pas d'espérance

## 5 Fonctions génératrices

Dans toute cette partie, on ne considère que des variables aléatoires à valeurs dans  $\mathbb{N}$ .

### Définition 4

La fonction génératrice de la variable aléatoire  $X$  est définie par :

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

On appelle série génératrice la série entière associée. On note  $R$  son rayon de convergence.

- Cette série converge pour  $t = 1$  et  $G_X(1) = \sum_{n=0}^{+\infty} P(X = n) = 1$ .  
Donc  $R \geq 1$ .
- Lorsque  $X$  est une variable aléatoire finie,  $G_X$  est une fonction polynomiale et  $R = +\infty$ .
- Pour tout  $t \in [-1, 1]$ ,  $|t^n P(X = n)| \leq P(X = n)$  et  $\sum P(X = n)$  converge, donc la série entière converge normalement sur  $[-1, 1]$ .  $G_X$  est donc définie et continue sur  $[-1, 1]$ .
- $G_X$  est de classe  $\mathcal{C}^\infty$  sur (au moins) l'intervalle ouvert  $] - 1, 1[$  et on peut dériver terme à terme la somme de la série.

### Propriété 15

La loi d'une variable aléatoire  $X$  à valeurs dans  $\mathbb{N}$  est déterminée par  $G_X$ . Plus précisément, on a :

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

Deux variables aléatoires à valeurs dans  $\mathbb{N}$  ont même loi si, et seulement si, elles ont même fonction génératrice.

### SAVOIR-FAIRE

Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

### Propriété 16

Soit une variable  $X$  à valeurs dans  $\mathbb{N}$ .

- $X \in L^1$  si, et seulement si,  $G_X$  est dérivable en 1. On a alors :

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

- $X \in L^2$  si, et seulement si,  $G_X$  est deux fois dérivable en 1. On a alors :

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

et il faut savoir en déduire  $V(X)$  par la formule de Huygens.

Propriété 17

- Si  $X$  et  $Y$  sont indépendantes et à valeurs dans  $\mathbb{N}$ , alors

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

- Si  $X_1, \dots, X_n$  sont indépendantes et à valeurs dans  $\mathbb{N}$ , alors

$$G_{X_1+X_2+\dots+X_n} = G_{X_1} \cdot G_{X_2} \dots G_{X_n}$$

Exercice 11 : Étudier grâce à cette propriété :

1. la loi de la somme de variables aléatoires indépendantes de lois de Poisson  $\mathcal{P}(\lambda_1), \dots, \mathcal{P}(\lambda_r)$ ,
2. la loi de la somme de variables aléatoires indépendantes de lois binomiales  $\mathcal{B}(n_1, p), \dots, \mathcal{B}(n_r, p)$ .

Un tableau récapitulatif des lois est donné en annexe.

## 6 Annexe : quelques éléments de démonstrations

### Propriété 1

Supposons que  $X$  est bornée. Il existe  $k$  réel tel que  $|X| \leq k$ . On a

$$0 \leq |xP(x=x)| \leq MP(X=x)$$

(On rappelle si besoin que  $X(\Omega)$  est au plus dénombrable).

La famille  $(MP(X=x))_{x \in X(\Omega)}$  est sommable (de somme  $M$ ) donc par comparaison, la famille  $(xP(X=x))_{x \in X(\Omega)}$  l'est aussi, et  $X \in L^1$ .

### Propriété 2 (loi géométrique, loi de Poisson)

Dans les deux cas,  $X$  est à valeurs positives, donc  $X$  admet une espérance finie si et seulement si  $E(X) < +\infty$ .

Commençons par  $\mathcal{G}(p)$ .

$$E(X) = \sum_{k=1}^{+\infty} k(1-p)^{k-1}p$$

La série entière  $\sum x^k$  est de rayon de convergence 1. La fonction somme est dérivable (et même  $\mathcal{C}^\infty$ ) sur l'intervalle ouvert de convergence et ses dérivées s'obtiennent par dérivation terme à terme.

$$\forall x \in ]-1, 1[, \quad f(x) = \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$$

$$\forall x \in ]-1, 1[, \quad f'(x) = \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

et on trouve  $E(X) = \frac{p}{(1-(1-p))^2} = \frac{1}{p} < +\infty$  donc  $X \in L^1$ .

Poursuivons avec la loi de Poisson  $\mathcal{P}(\lambda)$ .

$$\begin{aligned} E(X) &= \sum_{k=0}^{+\infty} k \frac{e^{-\lambda} \lambda^k}{k!} = \sum_{k=1}^{+\infty} \frac{e^{-\lambda} \lambda^k}{(k-1)!} \\ &= \sum_{j=0}^{+\infty} \frac{e^{-\lambda} \lambda \cdot \lambda^j}{j!} = \lambda < +\infty \end{aligned}$$

donc  $X \in L^1$ .

**Propriété 3** Nous travaillons ici dans  $\mathbb{R}^+ \cup \{+\infty\}$ .

$$\begin{aligned} \sum_{n=0}^{+\infty} P(X > n) &= \sum_{n=0}^{+\infty} \sum_{k=n+1}^{+\infty} P(X = k) \\ &= \sum_{k=1}^{+\infty} \sum_{n=0}^{k-1} P(X = k) \text{ par le théorème de Fubini, cas positif} \\ &= \sum_{k=1}^{+\infty} kP(X = k) = E(X) \end{aligned}$$

### Théorème 2 (admis)

(Repose sur le théorème de transfert qui a été admis...). On admettra.

On pose  $Z = (X, Y)$  et on note  $\pi_1$  et  $\pi_2$  les applications projections, ce qui donne  $\pi_2(Z) = Y$ . Par hypothèse,  $Y$  est d'espérance finie, donc par la formule de transfert, la somme en jeu dans ce qui suit est sommable, et

$$E(Y) = \sum_{(x,y) \in Z(\Omega)} \pi_2(x,y)P(X=x, Y=y) = \sum_{(x,y) \in Z(\Omega)} yP(X=x, Y=y)$$

Pour  $(x, y) \in Z(\Omega)$ , il existe  $\omega \in \Omega$  tel que  $(x, y) = (X(\omega), Y(\omega))$ , ce qui fait que :

$$0 \leq |x|P(X=x, Y=y) \leq yP(X=x, Y=y)$$

Donc  $\sum_{(x,y) \in Z(\Omega)} |x|P(X=x, Y=y)$  est sommable. Donc par le théorème de transfert,  $\pi_1(X, Y)$  est d'espérance finie. Autrement dit,  $X$  est d'espérance finie.

### Propriété 4

**linéarité de l'espérance :** On l'admet (trop de questions de sommabilité).

La propriété sur la variable centrée s'en déduit.

**positivité de l'espérance :** si  $X \geq 0$ , alors  $E(X)$  est la somme d'une famille sommable de termes positifs, donc positive. De plus,  $E(0) = 0$  (somme de termes nuls).

Et si  $\begin{cases} X & \geq 0 \\ E(X) & = 0 \end{cases}$  alors pour tout  $x \in X(\Omega)$ , on a  $0 \leq xP(X = x) \leq E(X) = 0$  donc  $P(X = x) = 0$  pour tout  $x$  autre que 0. Enfin,

$$P(X = 0) = 1 - \sum_{x \in X(\Omega) \setminus \{0\}} P(X = x) = 1 - 0 = 1$$

**croissance de l'espérance :** si  $X$  et  $Y$  sont d'espérance finie et  $X \leq Y$ , alors  $0 \leq Y - X$ . Par positivité,  $0 \leq E(Y - X)$ . Et par linéarité,  $0 \leq E(Y) - E(X)$ . Ainsi  $E(X) \leq E(Y)$ .

**inégalité triangulaire :** Admis. Facile pour l'équivalence des espérances finies, puis croissance de l'espérance.

### Propriété 6

Par indépendance de  $X$  et  $Y$ ,  $P(X = x, Y = y) = P(X = x)P(Y = y)$ .

Par la formule de transfert,  $XY$  est d'espérance finie si et seulement si la famille  $(xyP(X = x, Y = y))_{(x,y) \in X(\Omega) \times Y(\Omega)}$  est sommable. Or

$$\begin{aligned} \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |xy|P(X = x, Y = y) &= \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} |xy|P(X = x)P(Y = y) \\ &= E(|X|)E(|Y|) < +\infty \end{aligned}$$

donc  $XY$  est d'espérance finie. De plus, en reprenant le même type de calcul,

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

### Propriété 7 et propriété 8 (Cauchy-Schwarz)

Démonstration intéressante.

Soient  $X$  et  $Y$  dans  $L^1$ .

On connaît l'inégalité :  $|ab| \leq \frac{1}{2}(a^2 + b^2)$ . Avec le théorème de domination, on en déduit le lemme suivant.

#### Lemme 1

Si  $X$  et  $Y$  sont dans  $L^2$ , la variable  $XY$  est dans  $L^1$ .

Remarque : on peut facilement en déduire que

$$L^2 = \{\text{variables aléatoires discrètes réelles sur } (\Omega, \mathcal{A}, P) \text{ telles que } X^2 \text{ soit d'espérance finie}\}$$

est un  $\mathbb{R}$ -espace vectoriel.

On déduit du lemme, avec  $X$  et  $Y = 1$ , que  $X \in L^2 \Rightarrow X \in L^1$ .

• Pour  $t$  réel, considérons  $P(t) = E((tX + Y)^2)$ . Puisque  $(tX + Y)^2 = t^2X^2 + 2tXY + Y^2$ , il s'agit d'une combinaison linéaire de variables d'espérance finie (par le lemme), donc bien d'une variable d'espérance finie (par linéarité). Et

$$P(t) = t^2E(X^2) + 2tE(XY) + E(Y^2)$$

Dans le cas où  $E(X^2) = 0$ , comme  $X^2 \geq 0$ ,  $X^2$  est presque sûrement nulle, et donc  $XY$  et  $X$  sont presque sûrement nulles. L'inégalité attendue est satisfaite et est même une égalité. Dans le cas où  $E(X^2) \neq 0$ ,  $P$  est un polynôme du second degré et de signe constant. Son discriminant  $\Delta$  est donc négatif ou nul. Or  $\Delta = 4((E(XY))^2 - E(X^2)E(Y^2))$ , donc on a l'inégalité cherchée :

$$(E(XY))^2 \leq E(X^2)E(Y^2)$$

• Si  $X = 0$  presque sûrement, on a, on l'a vu, égalité dans l'inégalité précédente.

• Et si  $E(X^2) \neq 0$ , on a égalité dans l'inégalité si, et seulement si,  $\Delta = 0$  si, et seulement si,  $\exists t_0 \in \mathbb{R}$  tel que  $P(t_0) = 0$  si, et seulement si,  $\exists t_0 \in \mathbb{R}$  tel que  $E((t_0X + Y)^2) = 0$ , si, et seulement si,  $Y = -t_0X$  presque sûrement.

En définitive, il y a égalité si, et seulement si, il existe  $a$  et  $b$  réels tels que  $P(aX + bY) = 0$ .

### Propriété 9

• Par linéarité de l'espérance,  $(aX + b - E(aX + b))^2 = a^2(X - E(X))^2$  et on applique à nouveau la linéarité de la variance.

- Si  $X$  est presque sûrement constante, alors il existe  $m$  tel que  $P(X = m) = 1$ . On a  $E(X) = m$ , et donc  $P((X - E(X))^2 = 0) = 1$  et  $E(X - E(X))^2 = 0$ .
  - Si  $V(X) = 0$ , alors  $\begin{cases} E(X - E(X))^2 = 0 \\ (X - E(X))^2 \geq 0 \end{cases}$  donc  $(X - E(X))^2 = 0$  presque sûrement.
- En notant  $m = E(X)$ ,  $P(X = m) = 1$ .

### Propriété 12

- Soit  $X$  de loi géométrique de paramètre  $p$ .  $X(X - 1)$  est à valeurs positives. Par le théorème de transfert, si la quantité suivante est finie,  $X(X - 1)$  sera d'espérance finie.

$$E(X(X - 1)) = \sum_{n=1}^{+\infty} n(n - 1)P(X = n) = \sum_{n=2}^{+\infty} p(1 - p)n(n - 1)(1 - p)^{n-2}$$

La série entière  $\sum x^k$  est de rayon de convergence 1. La fonction somme est dérivable (et même  $C^\infty$ ) sur l'intervalle ouvert de convergence et ses dérivées s'obtiennent par dérivation terme à terme.

$$\forall x \in ] - 1, 1[, \quad f(x) = \sum_{k=0}^{+\infty} x^k = \frac{1}{1 - x}$$

$$\forall x \in ] - 1, 1[, \quad f'(x) = \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1 - x)^2}$$

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

On trouve alors  $E(X(X - 1)) = p(1 - p) \frac{2}{(1 - (1 - p))^3} = \frac{2(1 - p)}{p^2}$ . On remarque que

$$X^2 = X(X - 1) + X \text{ donc par linéarité de l'espérance, } E(X^2) = \frac{2(1 - p)}{p^2} + \frac{1}{p}$$

On termine avec la formule de Huygens.  $X$  est de variance finie égale à  $\frac{1 - p}{p^2}$ .

- Soit  $X$  de loi de Poisson de paramètre  $\lambda$ .  $X(X - 1)$  est à valeurs positives. Par le théorème de transfert, si la quantité suivante est finie,  $X(X - 1)$  sera d'espérance finie.

$$E(X(X - 1)) = \sum_{n=0}^{+\infty} n(n - 1)P(X = n) = \sum_{n=2}^{+\infty} \frac{e^{-\lambda} \lambda^n \cdot \lambda^{n-2}}{(n - 2)!} = e^{-\lambda} \lambda^2 e^\lambda$$

On remarque que

$$X^2 = X(X - 1) + X \text{ donc par linéarité de l'espérance, } E(X^2) = \lambda^2 + \lambda$$

On termine avec la formule de Huygens.  $X$  est de variance finie égale à  $\lambda$ .

### Propriété 13

$$\begin{aligned} V(X + Y) &= E[(X + Y - E(X + Y))^2] = E[(X - E(X) + Y - E(Y))^2] \\ &= E[(X - E(X))^2 + (Y - E(Y))^2 + 2(X - E(X))(Y - E(Y))] \\ &= V(X) + V(Y) + 2E[(X - E(X))(Y - E(Y))] \end{aligned}$$

### Théorème 3, inégalité de Markov

On démontre l'inégalité de Markov dans le cadre suivant : variable aléatoire discrète à valeurs dans  $\mathbb{N}$ . Puis nous donnons une autre façon de faire pour travailler les variables aléatoires indicatrices d'un événement.

Si  $E(X) = +\infty$ , on a bien  $P(X > a) \leq 1 \leq \frac{E(X)}{a}$  pour  $a > 0$ .

Dans ce qui suit,  $X$  est une variable aléatoire à valeurs positives d'espérance finie ;  $a$  est un réel strictement positif.

CAS OÙ  $X(\Omega) = \{i, i \in I\}$  AVEC  $I \subset \mathbb{N}$

On note  $J$  l'ensemble des indices  $i$  de  $I$  pour lesquels  $i \geq a$ .

$$E(X) = \sum_{i \in I} iP(X = i) = \sum_{i \in J} iP(X = i) + \sum_{i \in \bar{J}} iP(X = i)$$

Comme  $X$  est à valeurs positives, pour tout  $i \in \bar{J}$ , on a  $iP(X = i) \geq 0$ . Par positivité pour la somme de séries convergentes,  $\sum_{i \in \bar{J}} iP(X = i) \geq 0$ .

D'autre part, pour  $i \in J$ , on a  $iP(X = i) \geq aP(X = i)$ . Par croissance de la somme de séries convergentes,

$$\begin{aligned} \sum_{i \in J} iP(X = i) &\geq \sum_{i \in J} aP(X = i) \\ &\geq a \sum_{i \in J} P(X = i) \\ &\geq aP(X \geq a) \end{aligned}$$

DÉMONSTRATION AVEC DES VARIABLES ALÉATOIRES INDICATRICES

On introduit  $Y = X \mathbb{1}_{(X \geq a)}$ . On a :

$$\begin{aligned} 0 \leq \mathbb{1}_{(X \geq a)} &\leq 1 \\ 0 \leq X \mathbb{1}_{(X \geq a)} &\leq X \quad \text{car } X \text{ est à valeurs positives} \end{aligned}$$

D'autre part,  $a \mathbb{1}_{(X \geq a)}$  vaut 0 si  $(X < a)$  est réalisé, et vaut  $a$  si  $(X \geq a)$  est réalisé. On a donc :

$$0 \leq a \mathbb{1}_{(X \geq a)} \leq X$$

Par croissance de l'espérance :

$$\begin{aligned} E(a \mathbb{1}_{(X \geq a)}) &\leq E(X) \\ aE(\mathbb{1}_{(X \geq a)}) &\leq E(X) \\ aP(X \geq a) &\leq E(X) \end{aligned}$$

#### **Théorème 4, inégalité de Bienaymé-Tchebychev**

Soit  $X \in L^2$ . Soit  $\varepsilon > 0$ .

On pose  $Y = (X - E(X))^2$ .  $Y$  est à valeurs positives et admet une espérance. On applique à  $Y$  l'inégalité de Markov avec  $a = \varepsilon^2$  :

$$P(Y \geq \varepsilon^2) \leq \frac{V(X)}{\varepsilon^2}$$

Enfin, on remarque que  $(Y \geq \varepsilon^2) = (\sqrt{Y} \geq |\varepsilon|) = (|X - E(X)| \geq \varepsilon)$ . On a bien :

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{V(X)}{\varepsilon^2}$$

#### **Théorème 5, loi faible des grands nombres**

Soit  $(X_n)$  une suite de variables aléatoires indépendantes admettant une même espérance finie  $m$  et une même variance finie  $\sigma^2$ . Par linéarité de l'espérance :

$$\begin{aligned} E\left(\frac{S_n}{n}\right) &= \frac{1}{n} (E(X_1) + E(X_2) + \dots + E(X_n)) \\ &= \frac{1}{n} (nm) = m \end{aligned}$$

Par propriété «  $V(aZ) = a^2V(Z)$  », puis par propriété pour la variance d'une somme de variables aléatoires indépendantes,  $\frac{S_n}{n}$  admet une variance et :

$$\begin{aligned} V\left(\frac{S_n}{n}\right) &= \frac{1}{n^2} (V(X_1) + \dots + V(X_n)) \\ &= \frac{1}{n^2} (V(X_1) + V(X_2) + \dots + V(X_n)) \\ &= \frac{1}{n^2} (n\sigma^2) = \frac{\sigma^2}{n} \end{aligned}$$

Soit  $\varepsilon \geq 0$ . D'après l'inégalité de Bienaymé-Tchebychev :

$$0 \leq P\left(\left|\frac{S_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2}$$

On a  $\lim_{n \rightarrow +\infty} \frac{\sigma^2}{n\varepsilon^2} = 0$ . Par le théorème d'encadrement,

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

#### **Fonctions génératrices pour les lois usuelles**

Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.

- Pour  $X \sim b(p)$ ,  $G_X(t) = P(X = 0) + P(X = 1)t = q + pt$  pour tout  $t$  réel.
- Pour  $X \sim \mathcal{B}(n, p)$ ,

$$G_X(t) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} t^k = (pt + q)^n$$

- Pour  $X$  suivant la loi géométrique de paramètre  $p$ ,

$$G_X(t) = \sum_{k=1}^{+\infty} pq^{k-1} t^k = \sum_{j=0}^{+\infty} pq^j t^{j+1} = tp \frac{1}{1-tq} = \frac{pt}{1-qt}$$

- Pour  $X$  suivant la loi de Poisson de paramètre  $\lambda$ ,

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

### Propriété 16

Conformément au programme, on ne montre que le premier sens direct. Soit  $X$  d'espérance finie, montrons que  $G_X$  est dérivable en 1 et  $G'_X(1) = E(X)$ .

Soit  $R$  le rayon de convergence de  $\sum P(X = n)t^n$ . On a  $R \geq 1$  et  $G_X$  est dérivable sur  $] -R, R[$ , donc sur  $] -1, 1[$ , et

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

- Si  $R > 1$ ,  $G_X$  est dérivable en 1 et  $G'_X(1) = E(X)$ .
- Et si  $R = 1$ , comme  $\sum nP(X = n)$  converge, on peut appliquer le théorème d'Abel radial pour obtenir

$$\lim_{t \rightarrow 1} G'_X(t) = \sum_{n=1}^{+\infty} nP(X = n) = E(X).$$

Par ailleurs, on a vu que  $G_X$  était continue sur  $[0, 1]$ . Par le théorème de la limite de la dérivée,  $G_X$  est dérivable en 1.

La réciproque est admise, conformément au programme, de même que la réciproque pour le 2<sup>e</sup> point (que je passe entièrement).

### Propriété 17

Soient  $X_1, \dots, X_n$  indépendantes. Par la propriété de transfert d'indépendance,  $t^{X_1}, \dots, t^{X_n}$  sont indépendantes. Par propriété pour l'espérance d'un produit de variables aléatoires indépendantes,

$$E(t^{X_1} t^{X_2} \dots t^{X_n}) = E(t^{X_1}) E(t^{X_2}) \dots E(t^{X_n})$$

puis le résultat.