Moments d'une variable aléatoire discrète

Les attentes

1. Pour une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$, formule de l'espérance par antirépartition.
2. Existence, et calcul le cas échéant, de l'espérance d'une variable aléatoire complexe X. On note : $X \in L^1$.
3. Formule de transfert.
4. Théorème d'existence d'une espérance par domination.
5. Si X et Y sont dans L^1 et indépendantes, alors XY est dans L^1 et : $\mathrm{E}(XY) = \mathrm{E}(X)\,\mathrm{E}(Y)$.
6. Si $E(X^2) < +\infty$, on note : $X \in L^2$. Auquel cas, X est d'espérance finie.
7. Savoir définir la variance de X. Formule de Koenig-Huygens.
8. Connaître les espérance et variance des lois de référence.
9. Savoir définir la covariance de X et Y, et donner la relation de Huygens.
10. Formules pour $E(aX+b)$, $V(aX+b)$, $E(\sum\limits_{i=1}^{n} \lambda_i X_i)$, $V(X+Y)$, $V(\sum\limits_{i=1}^{n} X_i)$ avec les X_i indépendantes.
11. Savoir définir la fonction génératrice de X variable aléatoire à valeurs dans $\mathbb{N}$. Savoir calculer G_X pour les lois de référence.
12. Détermination de la loi de X à l'aide de G_X.

1. Inégalité de Cauchy-Schwarz et cas d'égalité.
2. Quelle est la variable aléatoire centrée et réduite déduite de X ?
3. Inégalité de Markov, inégalité de Bienaymé-Tchebichev, loi faible des grands nombres.
4. Savoir utiliser G_X pour calculer $E(X)$ et $V(X)$. À quelle condition nécessaire et suffisante sur G_X, X est-elle d'espérance finie ?
5. Fonction génératrice d'une somme de variables aléatoires indépendantes.

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 Famille sommable

La famille $(u_i)_{i\in I}$ est sommable lorsque

2 Espérance d'une variable aléatoire réelle ou complexe

2.1 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 :

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

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,

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

$\mathrm{E}(X) =$

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 :

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

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 :

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

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

Remarques :

- L'espérance de X ne dépend que de la loi de X. On peut parler de l'espérance d'une loi.
- On remarque que X est d'espérance finie si et seulement si $|X|$ est d'espérance finie.
- Toute variable aléatoire finie, c'est-à-dire telle que $X(\Omega)$ est un ensemble fini, 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,

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

  et cette quantité est (heureusement !) indépendante de la numérotation choisie pour les valeurs prises par X.
- 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 $\mathbb{1}_A$ 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 λ est d'espérance finie égale à λ.

Nous considérons souvent des transformées de la ou des variables initiales (X^2, e^{tX}, X^k, XY^2, $\min(X,Y)$,...). 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 $f : X(\Omega) \to \mathbb{C}$. 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 :

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

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

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

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

Propriété 3 – calcul de l'espérance par antirépartition

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

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

Exercice 4 : Retrouver, avec cette formule d'antirépartition, l'espérance de X, variable aléatoire suivant une loi géométrique de paramètre p.

2.2 propriétés de l'espérance

Théorème 2 – théorème de domination

Si X et Y sont deux variables aléatoires vérifiant $0 \leqslant |X| \leqslant Y$. 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 :

$$\mathrm{E}(\lambda X + \mu Y) = \lambda \mathrm{E}(X) + \mu \mathrm{E}(Y)$$

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

Positivité de l'espérance : si $X \geqslant 0$, alors $\mathrm{E}(X) \geqslant 0$. De plus,

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

Croissance de l'espérance : si X et Y sont d'espérance finie et $X \leqslant Y$, on a $\mathrm{E}(X) \leqslant \mathrm{E}(Y)$.

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

$$|\mathrm{E}(X)| \leqslant \mathrm{E}(|X|)$$

Propriété 5 – linéarité de l'espérance

Soit $(X_1, \ldots, X_n)$ une famille de variables aléatoires de L^1. Toute combinaison linéaire de $X_1, \ldots, X_n$ est dans L^1 et on a, pour $a_1, \ldots, a_n$ scalaires,

$$\mathrm{E}(a_1 X_1 + a_2 X_2 + \cdots + a_n X_n) = a_1 \mathrm{E}(X_1) + a_2 \mathrm{E}(X_2) + \cdots + a_n \mathrm{E}(X_n)$$

$$\mathrm{E}(X_1 + X_2 + \cdots + X_n) = \mathrm{E}(X_1) + \mathrm{E}(X_2) + \cdots + \mathrm{E}(X_n)$$

> 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 $\mathrm{E}(X) = \dfrac{4}{3}$. Qu'écrivez-vous sur votre copie ?

Exercice 5 : On suppose que $X_1, X_2, \ldots, X_n$ sont des variables aléatoires positives d'espérance finie. Montrer, en utilisant le théorème de domination, que $Z = \inf(X_1, X_2, \ldots, X_n)$ et $T = \sup(X_1, X_2, \ldots, X_n)$ sont aussi d'espérance finie.

Exercice 6 : n personnes se rendent dans un cinéma et chacune choisit l'une des trois salles disponibles, de manière équiprobable. Pour $i \in [\![1, 3]\!]$, 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 $\mathrm{E}(T)$. 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

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

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

$$\mathrm{E}\left(X_1 X_2 \ldots X_n\right) = \mathrm{E}\left(X_1\right)\mathrm{E}\left(X_2\right)\ldots\mathrm{E}\left(X_n\right)$$

Exercice 7 : 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 :

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

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 $\mathrm{E}(T_n)$ et en déduire une relation entre $P\left(T_n = 1\right)$ et $P\left(T_n = -1\right)$.
2. En déduire la loi de T_n.

3 Variance d'une variable aléatoire réelle

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

3.1 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 $\mathrm{E}\left(X^2\right) < +\infty$, alors X est d'espérance finie. Dit autrement,

$$X \in L^2 \;\Rightarrow\; X \in L^1 \quad \text{ou encore } L^2 \subset L^1$$

Propriété 8 – inégalité de Cauchy-Schwarz

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

$$|\mathrm{E}(XY)| \leqslant \sqrt{\mathrm{E}\left(X^2\right)}\,\sqrt{\mathrm{E}\left(Y^2\right)}$$

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

3.2 définition et propriétés de la variance

Définition 3

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

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

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

La variance est la « moyenne » du carré de la distance de X à $\mathrm{E}(X)$, 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 $\mathrm{V}(aX + b) = a^2\mathrm{V}(X)$.
- $\mathrm{V}(X) = 0$ si et seulement si X est une variable aléatoire presque-sûrement constante.

Propriété 10 – formule de Kœnig-Huygens

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

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 $\dfrac{X - \mathrm{E}(X)}{\sqrt{\mathrm{V}(X)}}$ 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 λ est de variance finie égale à λ.

4 Covariance de deux variables aléatoires de L^2

Définition - propriété 1

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

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

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

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

Remarques :

— Pour $X \in L^2$, on a $\mathrm{Cov}(X,X) = \mathrm{V}(X)$.

— L'application $(X,Y) \mapsto \mathrm{Cov}(X,Y)$ est une forme bilinéaire symétrique positive sur L^2 (mais pas définie positive). Par exemple,

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

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

— Lorsque $\mathrm{Cov}(X,Y)=0$, on dit que X et Y sont *décorrélées*.

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

$$\boxed{X \text{ et } Y \text{ sont indépendantes}} \quad \begin{array}{c}\Longrightarrow \\ \not\Longleftarrow\end{array} \quad \boxed{X \text{ et } Y \text{ sont décorrélées, } \mathrm{Cov}(X,Y)=0}$$

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

Exercice 8 : 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 ($n \geqslant 1$) si les n premiers lancers ont amené le même côté et le $(n+1)^{\mathrm{e}}$ 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 $\mathrm{Cov}(L_1,L_2)$ existe. Calculer $\mathrm{Cov}(L_1,L_2)$.
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 :

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

Par exemple,

$$\mathrm{V}(X-Y) =$$

$$\mathrm{V}(aX+bY) =$$

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 :

- Si $X_1, \dots, X_n$ sont indépendantes,

$$\mathrm{V}(X_1 + X_2 + \dots + X_n) = \mathrm{V}(X_1) + \mathrm{V}(X_2) + \dots + \mathrm{V}(X_n)$$

- Dans le cas général,

$$\mathrm{V}(X_1 + X_2 + \dots + X_n) = \sum_{(i,j)\in[\![1,n]\!]^2} \mathrm{Cov}(X_i, X_j)$$

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

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

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

Dans le cas où les X_k sont deux à deux non corrélées, pour $i \neq j$, on a $\mathrm{Cov}(X_i, X_j) = 0$ 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 9 (extrait Mines 2024) : Soit $X_1, \dots, X_n$ indépendantes et de même loi, donnée par

$$\mathrm{P}(X_1 = -1) = \frac{1}{2} = \mathrm{P}(X_1 = 1)$$

On pose $S_n = \sum\limits_{k=1}^{n} X_k$. Donner l'espérance et la variance de S_n.

Exercice 10 : Soit $(X_n)_{n\geqslant 1}$ 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 $Z = Y_1 + Y_2 + \dots + Y_n$.

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 la fonction définie par :

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

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 $R \geqslant 1$.

— Lorsque X est une variable aléatoire finie, G_X est une fonction polynomiale et $R = +\infty$.

— Pour tout $t \in [-1, 1]$, $|t^n P(X = n)| \leqslant P(X = n)$ donc en posant $f_n : t \mapsto t^n P(X = n)$, on a

$$0 \leqslant \|f_n\|_{\infty,\, [-1,1]} \leqslant P(X = n)$$

La série $\sum P(X = n)$ converge, donc la série entière converge normalement, donc uniformément, sur $[-1, 1]$. Par transfert de continuité, G_X est 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 :

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

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

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

et il faut savoir en déduire $\mathrm{V}(X)$ 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

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

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 $\mathcal{B}(n_1, p), \dots, \mathcal{B}(n_r, p)$.

Exercice 12 : On donne, pour tout réel t, $G_X(t) = ae^{1+t^2}$.

1. Donner a.
2. Donner la loi de X.
3. Donner, à l'aide de G_X, l'espérance et la variance de X.

6 Inégalités probabilistes et loi faible des grands nombres

6.1 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

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

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

$$\forall a > 0, \quad P(|X| \geqslant a) \leqslant \frac{\mathrm{E}(|X|)}{a}$$

Théorème 4 – Inégalité de Bienaymé-Tchebychev

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

$$\forall \varepsilon > 0, \quad P\Big(|X - \mathrm{E}(X)| \geqslant \varepsilon\Big) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$$

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 σ^2. On applique l'inégalité de Bienaymé-Tchebychev avec $\varepsilon = a\sigma$ où $a > 0$. On obtient $P(|X - m| \geqslant a\sigma) \leqslant \frac{1}{a^2}$. En particulier :

— $P(|X - m| \geqslant 2\sigma) \leqslant \frac{1}{4}$ (il y a moins de 25% de chance que X s'écarte de sa moyenne de plus de 2σ)

— $P(|X - m| \geqslant 10\sigma) \leqslant \frac{1}{100}$ (il y a moins de 1% de chance que X s'écarte de sa moyenne de plus de 10σ).

Exercice 13 (oral Mines-Télécom, extrait) : Soient $X_1, \ldots, X_n$ des variables aléatoires indépendantes qui suivent toutes la loi uniforme sur $\{-1, 1\}$. On pose $S_n = \sum_{i=1}^{n} X_i$.

1. Pour $n \in \mathbb{N}^*$ et $t > 0$, calculer $E(e^{tS_n})$.
2. Montrer que pour tout réel t, $\mathrm{ch}(t) \leqslant \exp(\frac{t^2}{2})$.
3. Montrer que pour $a > 0$, $P(S_n \geqslant a) \leqslant e^{-\frac{a^2}{2n}}$.

6.2 loi faible des grands nombres

Théorème 5 – Loi faible des grands nombres

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

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

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

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.

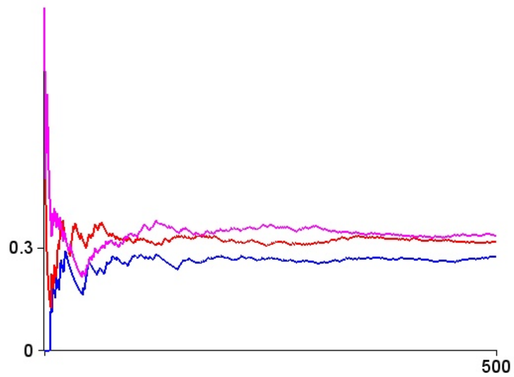
$X_1 + X_2 + \cdots + X_n$ est égale au nombre de succès dans les n premières expériences.

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

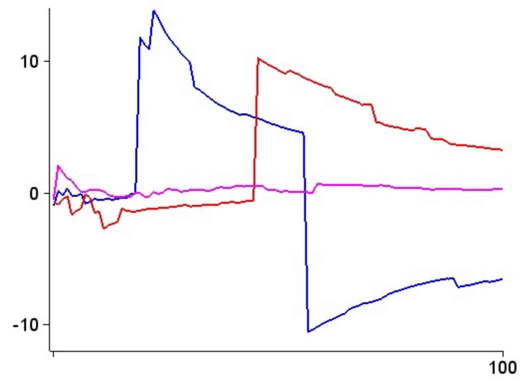
la probabilité empirique de succès tend vers la probabilité théorique de succès (p) quand le nombre d'expériences tend vers l'infini.

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*.

Nous visualisons, en complétant les légendes, les évolutions de $\left(\frac{S_n(\omega)}{n}\right)_{1\leqslant n\leqslant 500}$ pour 3 résultats ω, où $(X_i)_{i\in\mathbb{N}^*}$ est une suite de variables indépendantes et de même loi que X.



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



X n'a pas d'espérance

7 Annexe : quelques éléments de démonstrations

Propriété 1

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

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

(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 $\mathcal{P}(\lambda)$.

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

donc $X \in L^1$.

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

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

Théorème 2 (admis)

(Repose sur le théorème de transfert qui a été admis...). On admettra.
On pose $Z = (X, Y)$ et on note π_1 et π_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

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

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

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

Donc $\sum\limits_{(x,y)\in Z(\Omega)} |x|P(X = x, Y = y)$ 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 $X \geqslant 0$, alors $\mathrm{E}(X)$ est la somme d'une famille sommable de termes positifs, donc positive.
De plus, $\mathrm{E}(0) = 0$ (somme de termes nuls).
Et si $\begin{cases} X & \geqslant 0 \\ E(X) & = 0 \end{cases}$ alors pour tout $x \in X(\Omega)$, on a $0 \leqslant xP(X = x) \leqslant \mathrm{E}(X) = 0$ 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 $X \leqslant Y$, alors $0 \leqslant Y - X$. Par positivité, $0 \leqslant E(Y - X)$. Et par linéarité, $0 \leqslant E(Y) - E(X)$. Ainsi $E(X) \leqslant E(Y)$.

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

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

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

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

Propriété 7 et propriété 8 (Cauchy-Schwarz)

Démonstration intéressante.
Soient X et Y dans L^1.
On connaît l'inégalité : $|ab| \leqslant \frac{1}{2}(a^2 + b^2)$. 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 $P(t) = \mathrm{E}((tX + Y)^2)$. 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

$$P(t) = t^2\mathrm{E}(X^2) + 2t\mathrm{E}(XY) + \mathrm{E}(Y^2)$$

Dans le cas où $\mathrm{E}(X^2) = 0$, comme $X^2 \geqslant 0$, 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 $\mathrm{E}(X^2) \neq 0$, P est un polynôme du second degré et de signe constant. Son discriminant Δ est donc négatif ou nul. Or $\Delta = 4((\mathrm{E}(XY))^2 - \mathrm{E}(X^2)\mathrm{E}(Y^2))$, donc on a l'inégalité cherchée :

$$(\mathrm{E}(XY))^2 \leqslant \mathrm{E}(X^2)\mathrm{E}(Y^2)$$

• Si $X = 0$ presque sûrement, on a, on l'a vu, égalité dans l'inégalité précédente.
• Et si $\mathrm{E}(X^2) \neq 0$, 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 $\mathrm{E}((t_0X + Y)^2) = 0$, si, et seulement si, $Y = -t_0X$ presque sûrement.
En définitive, il y égalité si, et seulement si, il existe a et b réels tels que $P(aX + bY = 0) = 1$.

Propriété 9

- Par linéarité de l'espérance, $(aX+b-\mathrm{E}(aX+b))^2 = a^2(X-\mathrm{E}(X))^2$ 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 $\mathrm{E}(X)=m$, et donc $P((X-\mathrm{E}(X))^2=0)=1$ et $\mathrm{E}(X-\mathrm{E}(X))^2=0$.
  Si $\mathrm{V}(X)=0$, alors $\begin{cases} \mathrm{E}(X-\mathrm{E}(X))^2=0 \\ (X-\mathrm{E}(X))^2 \geqslant 0 \end{cases}$ donc $(X-\mathrm{E}(X))^2=0$ presque sûrement.
  En notant $m=\mathrm{E}(X)$, $P(X=m)=1$.

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.
  $$\mathrm{E}(X(X-1)) = \sum_{n=1}^{+\infty} n(n-1)P(X=n) = \sum_{n=2}^{+\infty} p(1-p)n(n-1)(1-p)^{n-2}$$
  La série entière $\sum x^k$ est de rayon de convergence 1. La fonction somme est dérivable (et même $\mathcal{C}^\infty$) sur l'intervalle ouvert de convergence et ses dérivées s'obtiennent par dérivation terme à terme.
  $$\forall x \in]-1,1[, \quad f(x) = \sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$$
  $$\forall x \in]-1,1[, \quad f'(x) = \sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$
  $$\forall x \in]-1,1[, \quad f''(x) = \sum_{k=2}^{+\infty} k(k-1)x^{k-1} = \frac{2}{(1-x)^3}$$
  On trouve alors $\mathrm{E}(X(X-1)) = p(1-p)\frac{2}{(1-(1-p))^3} = \frac{2(1-p)}{p^2}$. On remarque que
  $$X^2 = X(X-1)+X \text{ donc par linéarité de l'espérance, } \mathrm{E}(X^2) = \frac{2(1-p)}{p^2} + \frac{1}{p}$$
  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 λ. $X(X-1)$ est à valeurs positives. Par le théorème de transfert, si la quantité suivante est finie, $X(X-1)$ sera d'espérance finie.
  $$\mathrm{E}(X(X-1)) = \sum_{n=0}^{+\infty} n(n-1)P(X=n) = \sum_{n=2}^{+\infty} \frac{e^{-\lambda}\lambda^2.\lambda^{n-2}}{(n-2)!} = e^{-\lambda}\lambda^2 e^{\lambda}$$
  On remarque que
  $$X^2 = X(X-1)+X \text{ donc par linéarité de l'espérance, } \mathrm{E}(X^2) = \lambda^2+\lambda$$
  On termine avec la formule de Huygens. X est de variance finie égale à λ.

Propriété 13

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

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 $\mathrm{E}(X)=+\infty$, on a bien $P(X>a) \leqslant 1 \leqslant \frac{\mathrm{E}(X)}{a}$ 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 $i \geqslant a$.

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

Comme X est à valeurs positives, pour tout $i \in \overline{J}$, on a $iP(X=i) \geqslant 0$. Par positivité pour la somme de séries convergentes, $\sum_{i \in \overline{J}} iP(X=i) \geqslant 0$.
D'autre part, pour $i \in J$, on a $iP(X=i) \geqslant aP(X=i)$. Par croissance de la somme de séries convergentes,

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

démonstration avec des variables aléatoires indicatrices
On introduit $Y = X\,\mathbb{1}_{(X \geqslant a)}$.
$a\mathbb{1}_{(X \geqslant a)}$ vaut 0 si $(X < a)$ est réalisé, et vaut a si $(X \geqslant a)$ est réalisé. On a donc :

$$0 \leqslant a\mathbb{1}_{(X \geqslant a)} \leqslant X$$

Par croissance de l'espérance :

$$\begin{aligned} \mathrm{E}\left(a\mathbb{1}_{(X \geqslant a)}\right) &\leqslant \mathrm{E}(X) \\ a\mathrm{E}\left(\mathbb{1}_{(X \geqslant a)}\right) &\leqslant \mathrm{E}(X) \\ aP(X \geqslant a) &\leqslant \mathrm{E}(X) \end{aligned}$$

Théorème 4, inégalité de Bienaymé-Tchebychev

Soit $X \in L^2$. Soit $\varepsilon > 0$.
On pose $Y = (X - \mathrm{E}(X))^2$. Y est à valeurs positives et admet une espérance. On applique à Y l'inégalité de Markov avec $a = \varepsilon^2$:

$$P(Y \geqslant \varepsilon^2) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$$

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

$$P(|X - \mathrm{E}(X)| \geqslant \varepsilon) \leqslant \frac{\mathrm{V}(X)}{\varepsilon^2}$$

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 σ^2. Par linéarité de l'espérance :

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

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

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

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

$$0 \leqslant P(|\frac{S_n}{n} - m| \geqslant \varepsilon) \leqslant \frac{\sigma^2}{n\varepsilon^2}$$

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

$$\lim_{n \to +\infty} P(|\frac{S_n}{n} - m| \geqslant \varepsilon) = 0$$

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 λ,

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

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 $G'_X(1) = \mathrm{E}(X)$.
Soit R le rayon de convergence de $\sum P(X=n)t^n$. On a $R \geqslant 1$ et G_X est dérivable sur $]-R, R[$, donc sur $]-1,1[$, et $G'_X(t) = \sum\limits_{n=1}^{+\infty} nP(X=n)t^{n-1}$.

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

$\lim\limits_{t \to 1} G'_X(t) = \sum\limits_{n=1}^{+\infty} nP(X=n) = \mathrm{E}(X)$.
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 2è 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,

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

puis le résultat.