XIV. Espérance, variance, covariance etc

1er janvier 2025

Table des matières

C - Espérance et variance

Contenus	Capacités & commentaires
a) Espérance d'une variable aléatoire discrète réelle ou complexe	
Espérance d'une variable aléatoire à valeurs dans $[0,+\infty]$, définie par $$\mathrm{E}(X)=\sum_{x\in X(\Omega)} xP(X=x).$$	On adopte la convention $xP(X=x)=0$ lorsque $x=+\infty$ et $P(X=+\infty)=0$.
Variable aléatoire X à valeurs réelles ou complexes d'espérance finie, espérance de X.	X est d'espérance finie si la famille $\big(xP(X=x)\big)_{x\in X(\Omega)}$ est sommable. Dans ce cas, la somme de cette famille est l'espérance de X. Variable centrée.
Pour X variable aléatoire à valeurs dans $\mathbb{N}\cup\{+\infty\}$, relation : $$\mathrm{E}(X)=\sum_{n=1}^{+\infty} P(X\geqslant n).$$	
Espérance d'une variable géométrique, de Poisson.	
Formule de transfert : $f(X)$ est d'espérance finie si et seulement si la famille $\big(f(x)P(X=x)\big)_{x\in X(\Omega)}$ est sommable. Dans ce cas : $$\mathrm{E}\Big(f(X)\Big)=\sum_{x\in X(\Omega)} f(x)P(X=x).$$	On remarque que la formule s'applique aux couples, aux n-uplets de variables aléatoires.
Linéarité de l'espérance. Si $\lvert X\rvert\leqslant Y$ et $\mathrm{E}(Y)<+\infty$, alors X est d'espérance finie. Positivité, croissance de l'espérance. Si X est positive et d'espérance nulle, alors $(X=0)$ est presque sûr.	
Pour X et Y deux variables aléatoires indépendantes d'espérance finie, alors XY est d'espérance finie et : $$\mathrm{E}(XY)=\mathrm{E}(X)\,\mathrm{E}(Y).$$	Extension au cas de n variables aléatoires.

Contenus	Capacités & commentaires
b) Variance d'une variable aléatoire discrète réelle, écart type et covariance	
Si X^2 est d'espérance finie, X est d'espérance finie. Inégalité de Cauchy-Schwarz : si X^2 et Y^2 sont d'espérance finie, alors XY l'est aussi et : $$\mathrm{E}(XY)^2\leqslant \mathrm{E}(X^2)\,\mathrm{E}(Y^2)$$	Cas d'égalité.
Variance, écart type.	Notations $\mathrm{V}(X)$, $\sigma(X)$. Variable réduite.
Relation $\mathrm{V}(X)=\mathrm{E}(X^2)-\mathrm{E}(X)^2$.	
Relation $\mathrm{V}(aX+b)=a^2\,\mathrm{V}(X)$.	Si $\sigma(X)>0$, la variable $\dfrac{X-\mathrm{E}(X)}{\sigma(X)}$ est centrée réduite.
Variance d'une variable géométrique, de Poisson.	
Covariance de deux variables aléatoires. Relation $\mathrm{Cov}(X,Y)=\mathrm{E}(XY)-\mathrm{E}(X)\,\mathrm{E}(Y)$, cas de deux variables indépendantes.	

c) Fonctions génératrices

Fonction génératrice de la variable aléatoire X à valeurs dans $\mathbb{N}$: $$G_X(t) = \mathrm{E}\left(t^X\right) = \sum_{n=0}^{+\infty} P(X=n)\,t^n.$$	La série entière définissant G_X est de rayon $\geqslant 1$ et converge normalement sur $[-1,1]$. Continuité de G_X. Les étudiants doivent savoir calculer rapidement la fonction génératrice d'une variable aléatoire de Bernoulli, binomiale, géométrique, de Poisson.
La loi d'une variable aléatoire X à valeurs dans $\mathbb{N}$ est caractérisée par sa fonction génératrice G_X.	
La variable aléatoire X est d'espérance finie si et seulement si G_X est dérivable en 1 ; dans ce cas $\mathrm{E}(X) = {G_X}'(1)$.	La démonstration de la réciproque n'est pas exigible. Utilisation de G_X pour calculer $\mathrm{E}(X)$ et $\mathrm{V}(X)$.
Fonction génératrice d'une somme de deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.	Extension au cas d'une somme finie de variables aléatoires indépendantes.

d) Inégalités probabilistes

Inégalité de Markov.	
Inégalité de Bienaymé-Tchebychev.	
Loi faible des grands nombres : si $(X_n)_{n\geqslant 1}$ est une suite i.i.d. de variables aléatoires de variance finie, alors en notant $S_n = \sum_{k=1}^{n} X_k$ et $m = \mathrm{E}(X_1)$, pour tout $\varepsilon > 0$: $$P\left(\left\|\frac{S_n}{n} - m\right\| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.$$	Loi faible des grands nombres : si $(X_n)_{n\geqslant 1}$ est une suite i.i.d. de variables aléatoires de variance finie, alors, pour tout $\varepsilon > 0$, $$P\left(\left\|\frac{S_n}{n} - m\right\| \geqslant \varepsilon\right) \xrightarrow[n\to\infty]{} 0,$$ où $S_n = \sum_{k=1}^{n} X_k$ et $m = \mathrm{E}(X_1)$. Les étudiants doivent savoir retrouver, avec $\sigma = \sigma(X_1)$: $$P\left(\left\|\frac{S_n}{n} - m\right\| \geqslant \varepsilon\right) \leqslant \frac{\sigma^2}{n\varepsilon^2}.$$

Dans tout ce chapitre, $(\Omega, \mathscr{A}, P)$ est un espace probabilisé, et X et Y deux variables aléatoires discrètes réelles ou complexes définies sur Ω.

1 Espérance

1.1 Définition

Rappel 1.1.1.
Si X est **finie**, et si l'on note $X(\Omega) = \{x_1, \dots, x_n\}$, alors on appelle ***espérance*** de X le réel $\mathrm{E}(X) = \sum_{k=1}^{n} x_k P(X = x_k)$.

Étendons cette définition dans le cas d'une variable aléatoire discrète infinie :

Définition 1.1.2 (Espérance)**.**

1. Si X est à valeurs dans $[0, +\infty]$, on appelle ***espérance de*** X le nombre $\mathrm{E}(X) = \sum_{x \in X(\Omega)} xP(X = x)$. Ce nombre est soit un réel (fini), et dans ce cas on dit que X ***admet une espérance finie***, soit $+\infty$. On adpotera la convention $xP(X = x) = 0$ si $x = +\infty$ et $P(X = +\infty) = 0$.
2. Si X est à valeurs réelles ou complexes, on dit que X ***admet une espérance finie*** si la famille $(xP(X = x))_{x \in X(\Omega)}$ est sommable. Dans ce cas ***l'espérance de*** X est le réel ou complexe $\mathrm{E}(X) = \sum_{x \in X(\Omega)} xP(X = x)$.

Remarque 1.1.3. 1. Même si la situation $\mathrm{E}(X) = +\infty$ est possible dans le cas d'une variable aléatoire réelle positive telle que $P(X = +\infty) > 0$, on dira en général que X admet une espérance uniquement dans le cas où cette espérance est finie.

2. Si $X(\Omega)$ est indicé par $\mathbb{N}$, c'est-à-dire si $X(\Omega) = (x_k)_{k \in \mathbb{N}}$, X a une espérance finie si et seulement si la série $\sum_{n \in \mathbb{N}} x_k P(X = x_k)$ est absolument convergente.
3. La sommabilité est exigée afin que la valeur de la somme ne dépende pas de l'ordre de sommation des termes. La simple convergence de la série (dans le cas où $X(\Omega)$ est indicé par $\mathbb{N}$) ne suffit pas.
4. Pour montrer que X admet une espérance et la calculer, dans le cas où X n'est pas réelle positive, il faudra donc d'abord montrer que $\sum_{x \in X(\Omega)} |x| P(X = x)$ converge, puis calculer $\sum_{x \in X(\Omega)} xP(X = x)$.
5. Dans le cas d'une variable aléatoire réelle positive, la convergence de $\sum_{x \in X(\Omega)} xP(X = x)$ permet de conclure à l'existence d'une espérance (finie).

Exemple 1.1.4 (important)**.**
Soit $c \in \mathbb{C}$ et X une variable aléatoire valant c presque sûrement. Alors $\sum_{x \in X(\Omega)} xP(X = x) = cP(X = c) + \sum_{x \in X(\Omega),\ x \neq c} xP(X = x) = c + 0 = c$.

Exemple 1.1.5.
Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $P(X = n) = \frac{1}{2^{n+1}}$. Cette définition est licite puisque $\sum_{n=0}^{+\infty} \frac{1}{2^{n+1}} = 1$.
Alors X est à valeurs positives donc

$$\begin{aligned} \mathrm{E}(X) &= \sum_{n=0}^{+\infty} \frac{n}{2^{n+1}} = \sum_{n=1}^{+\infty} \frac{n}{2^{n+1}} \\ &= \frac{1}{4} \sum_{n=1}^{+\infty} \frac{n}{2^{n-1}} \end{aligned}$$

On reconnaît là la dérivée de la série entière géométrique, évaluée en $\frac{1}{2}$, donc $\mathrm{E}(X) = 1$.

Exercice 1.1.6.
Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$ et pour tout $n \in \mathbb{N}^*$, $P(X = n) = \dfrac{1}{n(n+1)}$. Montrer que la définition de cette loi est licite et calculer l'espérance de X.

Dans le cas d'une variable aléatoire à valeurs entières naturelles, nous avons également la propriété suivante :

Proposition 1.1.7.
Soit X à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Alors $\mathrm{E}(X) = \displaystyle\sum_{n=1}^{+\infty} P(X \geqslant n) = \sum_{n=0}^{+\infty} P(X > n)$.

Démonstration.
Soit $n \in \mathbb{N}$. Remarquons que $(X \geqslant n+1) \subset (X = n)$, donc $(X = n) = (X \geqslant n) \backslash (X \geqslant n+1)$. Il en vient $P(X = n) = P(X \geqslant n) - P(X \geqslant (n+1))$. Si l'on fixe $N \in \mathbb{N}$,

$$\begin{aligned}
\sum_{n=0}^{N} nP(X = n) = \sum_{n=1}^{N} nP(X = n) &= \sum_{n=1}^{N} n\left[P(X \geqslant n) - P(X \geqslant n+1)\right] \\
&= \sum_{n=1}^{N} nP(X \geqslant n) - \sum_{n=1}^{N} nP(X \geqslant n+1) \\
&= \sum_{n=1}^{N} nP(X \geqslant n) - \sum_{n=0}^{N} nP(X \geqslant n+1) \\
&= \sum_{n=1}^{N} nP(X \geqslant n) - \sum_{n=1}^{N+1} (n-1)P(X \geqslant n) \\
&= \sum_{n=1}^{N+1} P(X \geqslant n) - (N+1)P(X \geqslant (N+1)) \qquad (*).
\end{aligned}$$

Cette égalité permet d'observer que

$$\sum_{n=0}^{N} nP(X = n) \leqslant \sum_{n=1}^{N+1} P(X \geqslant n)$$

donc la convergence de $\displaystyle\sum_{n \geqslant 1} P(X \geqslant n)$ implique que X est d'espérance finie.
Réciproquement, si X est d'espérance finie alors, en remarquant que $(X \geqslant N+1) = \displaystyle\bigsqcup_{n=N+1}^{+\infty} (X = n)$,

$$(N+1)P(X \geqslant N+1) = (N+1) \sum_{n=N+1}^{+\infty} P(X = n) \leqslant \sum_{n=N+1}^{+\infty} nP(X = n)$$

donc $(N+1)P(X \geqslant N+1) \xrightarrow[N \to +\infty]{} 0$ et avec $(*)$ on en tire la convergence de $\displaystyle\sum_{n \geqslant 1} P(X \geqslant n)$, et également en passant à la limite quand $N \to +\infty$,

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

La convergence de l'une de ces deux séries étant équivalente à celle de la seconde, si les deux divergent l'égalité précédente est toujours valable, et les deux membres valent $+\infty$. □

Exemple 1.1.8.
On effectue des tirages d'une boule avec remise dans une urne de n boules numérotées. On arrête les tirages lorsque le numéro de la boule tirée est supérieur ou égal au numéro de la boule obtenue au précédent tirage. On note X le nombre de tirages effectués. Calculons $\mathbb{E}(X)$.
X est à valeurs dans $\mathbb{N}\backslash \{0,1\} \cup \{+\infty\}$, nous allons donc utiliser la formule $\mathrm{E}(X) = \displaystyle\sum_{k=2}^{+\infty} P(X \geqslant k)$.
Soit $k \in \mathbb{N}^*$. L'évènement $(X \geqslant k)$ correspond à l'évènement « les numéros des k premières boules tirées forment une suite strictement décroissante ». Or il y autant de suites strictement décroissantes de k éléments de $[\![1, n]\!]$ qu'il y a de manières de choisir k éléments parmi n : en effet, étant donnée une suite strictement décroissante de k éléments, on peut lui associer un unique ensemble de k éléments, formé par ses termes. Et réciproquement, à chaque ensemble de k éléments on peut associer une unique suite strictement décroissante, puisqu'il y a une seule manière d'ordonner ces éléments dans l'ordre strictement décroissant.

Ainsi il y a $\binom{n}{k}$ suites strictement décroissantes de k éléments de $[\![1, n]\!]$.

Donc $P(X \geqslant k) = \frac{\binom{n}{k}}{n^k}$, qui est nulle dès que $k > n$. D'où

$$\begin{aligned} E(X) &= \sum_{k=2}^{n} \frac{\binom{n}{k}}{n^k} = \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{n}\right)^k - 1 - \binom{n}{1} \frac{1}{n} \\ &= \left(1 + \frac{1}{n}\right)^n - 2. \end{aligned}$$

1.2 Propriétés

Proposition 1.2.1 (Linéarité de l'espérance)**.**
Soit $\lambda \in \mathbb{K}$. Si X et Y admettent toutes deux une espérance finie, alors $X + \lambda Y$ admet une espérance finie et $E(X + \lambda Y) = E(X) + \lambda E(Y)$.

Démonstration.
Nous allons montrer séparément que $E(\lambda Y) = \lambda E(Y)$ et $E(X+Y) = E(X) + E(Y)$, ce qui est assurera le résultat voulu.
Le cas $\lambda = 0$ est évident, traitons le cas $\lambda \neq 0$. Commençons par remarquer que $(\lambda Y)(\Omega) = \{\lambda y,\ y \in Y(\Omega)\}$. Alors

$$\begin{aligned} E(\lambda Y) &= \sum_{z \in (\lambda Y)(\Omega)} z P(\lambda Y = z) = \sum_{y \in Y(\Omega)} \lambda y P(\lambda Y = \lambda y) \\ &= \lambda \sum_{y \in Y(\Omega)} y P(Y = y) = \lambda E(Y). \end{aligned}$$

Montrons ensuite que $E(X + Y) = E(X) + E(Y)$. Soit $x \in X(\Omega)$. Par la formule des probabilités totales, $P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y)$, donc en sommant par paquets, $E(X) = \sum_{x \in X(\Omega)} x P(X = x) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} x P(X = x, Y = y)$.
De même, $E(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} y P(X = x, Y = y)$.
Par sommabilité, $E(X) + E(Y) = \sum_{(x,y) \in X(\Omega) \times Y(\Omega)} (x + y) P(X = x, Y = y)$.

En sommant par paquets selon la valeur de $z = x + y$, il vient finalement

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

soit $E(X) + E(Y) = \sum_{z \in (X+Y)(\Omega)} z P(X + Y = z) = E(X + Y)$. □

Corollaire 1.2.2.
Si X admet une espérance finie, et soit $a, b \in \mathbb{C}$. Alors $aX + b$ admet une espérance finie et $E(aX + b) = aE(X) + b$.

Démonstration.
Il suffit d'utiliser la linéarité de l'espérance et le fait que $E(b) = b$. □

Proposition 1.2.3 (Positivité de l'espérance)**.**
Soit X une variable aléatoire réelle positive.

1. $E(X) \geqslant 0$;
2. Réciproquement, si $E(X) \geqslant 0$, alors $X = 0$ presque sûrement.

Démonstration.
1. $E(X) = \sum_{x \in X(\Omega)} x P(X = x)$, et tous les termes de cette somme sont positifs.
2. Soit $x \in X(\Omega)$ tel que $x > 0$. Alors, si $P(X = x) \neq 0$, $xP(X = x) > 0$ et donc $E(X) \geqslant xP(X = x) > 0$ ce qui est absurde. Donc $P(X = x) = 0$. Finalement, $P(X \neq 0) = P(X > 0) = \sum_{x \in X(\Omega),\ x > 0} P(X = x) = 0$.

□

Corollaire 1.2.4 (Croissance de l'espérance)**.**
Soit X et Y deux variables aléatoires réelles admettant une espérance finie. Si $X \leqslant Y$, alors $E(X) \leqslant E(Y)$.

Démonstration.
Si $X \leqslant Y$, alors $Y - X \geqslant 0$ donc $E(Y - X) \geqslant 0$. Par linéarité nous avons alors $E(Y) - E(X) \geqslant 0$. □

Remarque 1.2.5.
Si $X \geqslant 0$ presque sûrement, alors on peut montrer que $\mathrm{E}(X) \geqslant 0$. De même, si $X \leqslant Y$ presque sûrement, alors $\mathrm{E}(X) \leqslant \mathrm{E}(Y)$.

Définition 1.2.6.
Une variable aléatoire d'espérance nulle est dite ***centrée***.

Corollaire 1.2.7.
$X - \mathrm{E}(X)$ est centrée.

Théorème 1.2.8.
Soit X une variable aléatoire réelle ou complexe et Y une variable aléatoire réelle positive, telles que $|X| \leqslant Y$. Si Y admet une espérance finie, alors X aussi.

Démonstration.
$|X|$ est à valeurs réelles positives, donc elle admet une espérance, éventuellement infinie.
Par probabilités totales, $\mathrm{E}(|X|) = \sum_{x\in|X|(\Omega)} \sum_{y\in Y(\Omega)} xP(|X| = x, Y = y)$.
Soit $(x, y) \in |X|(\Omega) \times Y(\Omega)$. S'il n'existe aucun $\omega \in \Omega$ tel que $|X|(\omega) = x$ et $Y(\omega) = y$, alors $P(|X| = x, Y = y) = 0$, donc $xP(|X| = x, Y = y) = yP(|X| = x, Y = y)$.
Sinon, il existe $\omega \in \Omega$ tel que $|X|(\omega) = x$ et $Y(\omega) = y$, et alors $x \leqslant y$ car $|X| \leqslant Y$, donc $xP(|X| = x, Y = y) \leqslant yP(|X| = x, Y = y)$.
Dans tous les cas, $\mathrm{E}(|X|) \leqslant \sum_{x\in|X|(\Omega)} \sum_{y\in Y(\Omega)} yP(|X| = x, Y = y)$.
Toujours par probabilités totales, $\mathrm{E}(Y) = \sum_{y\in Y(\Omega)} \sum_{x\in|X|(\Omega)} yP(|X| = x, Y = y) =$.
En réordonnant la somme, ce qui est possible car Y admet une espérance finie, $\mathrm{E}(Y) = \sum_{x\in|X|(\Omega)} \sum_{y\in Y(\Omega)} yP(|X| = x, Y = y) \geqslant \mathrm{E}(|X|)$, d'où le résultat. □

1.3 Formule de transfert

Théorème 1.3.1 (Formule de transfert)**.**
Soit X une variable aléatoire discrète à valeurs dans un ensemble E, et f une fonction définie sur $X(\Omega)$ et à valeurs réelles ou complexes.
Alors $f(X)$ admet une espérance finie si et seulement si la famille $(f(x)P(X = x))_{x\in X(\Omega)}$ est sommable.
Dans ce cas, $\mathrm{E}(f(X)) = \sum_{x\in X(\Omega)} f(x)P(X = x)$.

Démonstration.
Elle est théoriquement au programme mais repose essentiellement sur des arguments de sommabilité et n'a pas d'intérêt particulier. Elle figure en annexe à la fin de ce chapitre. □

Exercice 1.3.2.
Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}$ et pour tout $n \in \mathbb{N}$, $P(X = n) = \dfrac{1}{2^{n+1}}$. Calculer $\mathrm{E}(X^2)$.

Remarque 1.3.3.
Puisqu'un couple ou un n-uplet de variables aléatoires discrètes est une variable aléatoire discrète, la formule de transfert s'applique aux couples ou n-uplets.
Par exemple, pour un couple (X, Y) de variables aléatoires réelles admettant une espérance finie et $f : \mathbb{R}^2 \to \mathbb{R}$, nous avons

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

En particulier, $\mathrm{E}\left[XY\right] = \sum_{x\in X(\Omega),\ y\in Y(\Omega)} xyP(X = x, Y = y)$.

1.4 Variables indépendantes

Théorème 1.4.1.
Si X et Y sont indépendantes et admettent une espérance finie, alors XY aussi et $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$.

Démonstration.

$$\begin{aligned}
\mathrm{E}(X)\mathrm{E}(Y) &= \left(\sum_{x\in X(\Omega)} xP(X=x)\right)\left(\sum_{y\in Y(\Omega)} xP(Y=y)\right)\\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X=x)P(Y=y)\\
&\qquad \text{en réarrangeant les termes par sommabilité}\\
&= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xyP(X=x, Y=y) \text{ par indépendance}\\
&= \sum_{z\in(XY)(\Omega)} z \times \left(\sum_{(x,y)\in X(\Omega)\times Y(\Omega),\ xy=z} P(X=x,Y=y)\right)\\
&\qquad \text{en regroupant les termes par paquets}\\
&= \sum_{z\in(XY)(\Omega)} zP(XY=z) \text{ par la formule des probabilités totales}\\
&= \mathrm{E}(XY)
\end{aligned}$$

□

1.5 Lois usuelles

Proposition 1.5.1 (Espérance des lois usuelles, rappels de première année)**.**
Soit $p \in [0,1]$, $n \in \mathbb{N}$, $a, b \in \mathbb{Z}$.

1. L'espérance d'une variable aléatoire suivant la loi de Bernoulli de paramètre p est p.
2. L'espérance d'une variable aléatoire suivant la loi binomiale de paramètres n et p vaut np.
3. L'espérance d'une variable aléatoire suivant la loi uniforme sur $[\![a, b]\!]$, avec $a < b$, vaut $\dfrac{a+b}{2}$.

Théorème 1.5.2 (Espérance des lois géométrique et de Poisson)**.** 1. Si $X \sim \mathscr{G}(p)$, alors $\mathrm{E}(X) = \dfrac{1}{p}$;
2. Si $X \sim \mathscr{P}(\lambda)$, alors $\mathrm{E}(X) = \lambda$.

Démonstration. 1. On rappelle le résultat sur la dérivée de la série géométrique : pour tout $x \in]-1,1[$, $\displaystyle\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$.
On en tire $\displaystyle\mathrm{E}(X) = \sum_{n=1}^{+\infty} n(1-p)^{n-1}p = p.\frac{1}{(1-(1-p))^2} = \frac{1}{p}$.
2. $\displaystyle\mathrm{E}(X) = \sum_{n=0}^{+\infty} n\mathrm{e}^{-\lambda}\frac{\lambda^n}{n!} = \sum_{n=1}^{+\infty} \mathrm{e}^{-\lambda}\frac{\lambda^n}{(n-1)!} = \lambda\mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = \lambda$.

□

Exercice 1.5.3.
Si $X \sim \mathscr{G}(p)$, calculer $\mathrm{E}(\frac{1}{X})$ et $\mathrm{E}(X^2)$.

2 Variance

Dans cette section, X et Y sont des variables aléatoires **réelles**.

2.1 Définition

Proposition 2.1.1.
Si X^2 est d'espérance finie, alors X aussi.

Démonstration.
Il suffit de remarquer le fait suivant : pour tout $\omega \in \Omega$, soit $|X(\omega)| > 1$ auquel cas

$|X(\omega)| \leqslant X^2(\omega)$, ou alors $|X(\omega)| \leqslant 1$. Dans tous les cas, $|X(\omega)| \leqslant 1 + X^2(\omega)$. Or 1 et X^2 admettent une espérance finie, donc par somme $1 + X^2$ aussi, et par majoration X aussi.
On peut aussi utiliser, avec $y = 1$, l'inégalité importante « pour tout $x, y \in \mathbb{R}$, $|xy| \leqslant \frac{1}{2}(x^2 + y^2)$ », qui se démontre en observant que $(x \pm y)^2 \geqslant 0$. □

Remarque 2.1.2.
1. Ce résultat peut aussi se démontrer à partir de l'inégalité de Cauchy-Schwarz, qui sera vue juste après – la démonstration utilise d'ailleurs la même idée.
2. Si $p \in \mathbb{N}^*$ et $|X|^p$ admet une espérance, on dit que X ***admet un moment d'ordre*** p, qui vaut alors $\mathrm{E}(X^p)$. L'espérance d'une variable aléatoire est donc son moment d'ordre 1 (s'il existe). Ce vocabulaire est hors-programme en PSI.

Proposition 2.1.3 (Inégalité de Cauchy-Schwarz)**.**
Si X^2 et Y^2 sont toutes deux d'espérance finie, alors XY aussi et

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

Démonstration.
Reprenons la seconde idée de la démonstration de **2.1.1** : pour tout $x, y \in \mathbb{R}$, $(x \pm y)^2 \geqslant 0$, donc après développement $-x^2 - y^2 \leqslant 2xy \leqslant x^2 + y^2$. Alors $|xy| \leqslant \frac{1}{2}(x^2 + y^2)$.
Ceci implique que $|XY| \leqslant \frac{1}{2}(X^2 + Y^2)$. Puisque X^2 et Y^2 sont d'espérance finie, par somme et majoration, XY aussi.
La fin de la démonstration est analogue à la démonstration de l'inégalité de Cauchy-Schwarz pour un produit scalaire, vue en première année.
Soit $\lambda \in \mathbb{R}$. Considérons $\varphi : \lambda \mapsto \mathrm{E}\left((\lambda X + Y)^2\right)$. Cette application est bien définie car par combinaison linéaire, $(\lambda X + Y)^2 = \lambda^2 X^2 + 2XY + Y^2$ admet bien une espérance finie.
Par positivité de $(\lambda X + Y)^2$, $\varphi \geqslant 0$. Mais en développant, $\varphi(\lambda) = \lambda^2 \mathrm{E}(X^2) + 2\lambda \mathrm{E}(XY) + \mathrm{E}(Y^2)$, qui est donc une fonction polynomiale réelle de degré au plus 2 en λ. Cette fonction étant de signe constant, son discriminant est négatif est nul, ce qui donne exactement $\mathrm{E}^2(XY) \leqslant \mathrm{E}\left(X^2\right)\mathrm{E}\left(Y^2\right)$. □

La proposition **2.1.1** permet de définir la ***variance*** et l'***écart type*** d'une variable aléatoire :

Définition 2.1.4 (Variance et écart type)**.**
Si X^2 admet une espérance finie, on apelle ***variance de*** X le réel

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

Et on appelle ***écart type de*** X le réel $\sigma(X) = \sqrt{\mathrm{V}(X)}$.

Remarque 2.1.5.
La variance est bien définie dans ce cas car $(X - \mathrm{E}(X)^2 = X^2 - 2\mathrm{E}(X)X + \mathrm{E}^2(X)$, qui admettent toutes des espérances finies par hypothèse et grâce à la proposition **2.1.1**.
Et l'écart type est bien défini également, car en tant qu'espérance d'une variable aléatoire mise au carré et donc positive, $\mathrm{V}(X)$ est positive.

Remarque 2.1.6.

1. La variance et l'écart-type ne dépendent que de la loi de X.
2. Intuitivement la variance mesure la « moyenne » du carré des écarts à la « moyenne ». C'est donc une mesure de la dispersion des valeurs autour de la valeur moyenne prise par la variable aléatoire X. Si la variable X est exprimée dans une unité u, la variance sera exprimée dans l'unité u^2 (et l'écart-type dans l'unité u).

Exercice 2.1.7.
Quelle est la variance d'une variable aléatoire constante ?

2.2 Propriétés

Proposition 2.2.1 (Formule de König-Huygens)**.**
Si $\mathrm{V}(X)$ existe, $\mathrm{V}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X)$.

Démonstration.
Simple développement :

$$\begin{aligned}\mathrm{V}(X) &= \mathrm{E}\big((X-\mathrm{E}(X))^2\big) = \mathrm{E}\big(X^2 - 2\mathrm{E}(X)X + \mathrm{E}^2(X)\big)\\ &= \mathrm{E}(X^2) - 2\mathrm{E}(X)\mathrm{E}(X) + \mathrm{E}(\mathrm{E}^2(X)) = \mathrm{E}(X^2) - 2\mathrm{E}^2(X) + \mathrm{E}^2(X)\\ &= \mathrm{E}(X^2) - \mathrm{E}^2(X).\end{aligned}$$

□

Remarque 2.2.2.
En pratique on utilise presque tout le temps cette formule plutôt que la définition pour calculer une variance.

Proposition 2.2.3.
Soit $a, b \in \mathbb{R}$. Si X admet une variance, $aX + b$ aussi et $\mathrm{V}(aX+b) = a^2\mathrm{V}(X)$.

Démonstration.
Là encore il suffit de développer :
$(aX+b)^2 = a^2X^2 + 2abX + b^2$, qui est bien d'espérance finie par combinaison linéaire, donc $aX + b$ admet une variance. Puis,
$\mathrm{V}(aX+b) = \mathrm{E}\left(a^2X^2 + 2abX + b^2\right) - (a\mathrm{E}(X)+b)^2 = a^2\left(\mathrm{E}\left(X^2\right) - \mathrm{E}(X)^2\right) = a^2\mathrm{V}(X)$.
□

Définition 2.2.4.
On dit que X est ***réduite*** si elle admet une variance valant 1.

Proposition 2.2.5.
Si X admet une variance, $\dfrac{X - \mathrm{E}(X)}{\sigma(X)}$ est centrée réduite.

2.3 Lois usuelles

Proposition 2.3.1 (Variance des lois usuelles, rappels de première année)**.**
Soit $p \in [0,1]$ et $n \in \mathbb{N}^*$.

1. Toute variable aléatoire constante p.s. est de variance nulle.
2. Toute variable aléatoire réelle suivant la loi de Bernoulli de paramètre p a pour variance $p(1-p)$.
3. Toute variable aléatoire réelle suivant la loi binomiale de paramètres p et n a pour variance $np(1-p)$.
4. Toute variable aléatoire rélle suivant la loi uniforme sur $[\![1, n]\!]$, a pour variance $\dfrac{n^2-1}{12}$.

Théorème 2.3.2 (Variance des lois géométrique et de Poisson)**.**
1. Si $X \sim \mathscr{G}(p)$, alors $\mathrm{V}(X) = \dfrac{1-p}{p^2}$;
2. Si $X \sim \mathscr{P}(\lambda)$, alors $\mathrm{V}(X) = \lambda$.

Démonstration.
1. Dans le calcul de l'espérance de la loi géométrique nous avons utilisé que pour tout $x \in]-1,1[$, $\displaystyle\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$, en dérivant une fois la série entière géométrique. En la dérivant une seconde fois nous obtenons $\displaystyle\sum_{n=2}^{+\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$.
   Nous allons alors écrire $\mathrm{V}(X) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) - \mathrm{E}^2(X)$.
   Or, par la formule de transfert, $\displaystyle\mathrm{E}(X(X-1)) = \sum_{n=1}^{+\infty} n(n-1)p(1-p)^{n-1} = p(1-p)\sum_{n=2}^{+\infty} n(n-1)(1-p)^{n-2} = \frac{2p(1-p)}{p^3}$.
   Finalement, $\mathrm{V}(X) = \dfrac{2(1-p)}{p^2} + \dfrac{1}{p} - \dfrac{1}{p^2} = \dfrac{2-2p+p-1}{p^2} = \dfrac{1-p}{p^2}$.
2. Écrivons là encore $\mathrm{V}(X) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) - \mathrm{E}^2(X)$.

$$\begin{aligned}
\mathrm{E}(X(X-1)) &= \sum_{n=0}^{+\infty} n(n-1)\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} = \sum_{n=2}^{+\infty} n(n-1)\mathrm{e}^{-\lambda}\frac{\lambda^k}{k!} \\
&= \sum_{n=2}^{+\infty} \mathrm{e}^{-\lambda}\frac{\lambda^k}{(k-2)!} = \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda}\frac{\lambda^{k+2}}{k!} \\
&= \lambda^2 \mathrm{e}^{-\lambda}\sum_{n=0}^{+\infty}\frac{\lambda^k}{k!} = \lambda^2
\end{aligned}$$

donc $\mathrm{V}(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

□

3 Covariance

Dans cette section encore, X et Y sont des variables aléatoires **réelles**.

Définition 3.0.1 (Covariance).
Si X^2 et Y^2 ont une espérance finie, on définit la ***covariance de X et Y*** comme le réel $\mathrm{Cov}(X,Y) = \mathrm{E}\left[(X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right]$.

Remarque 3.0.2.
$\mathrm{Cov}(X,X)$ n'est autre que $V(X)$ et $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$.

Proposition 3.0.3.
Dans le cadre de la définition précédente,

$$\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$$
$$\mathrm{V}(X+Y) = \mathrm{V}(X) + \mathrm{V}(Y) + 2\mathrm{Cov}(X,Y).$$

Démonstration.
Simples développements. □

Corollaire 3.0.4.
Si X et Y sont indépendantes, $\mathrm{Cov}(X,Y) = 0$ et $\mathrm{V}(X+Y) = \mathrm{V}(X)+\mathrm{V}(Y)$.

Démonstration.
Si X et Y sont indépendantes, $\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$. □

Remarque 3.0.5.
La réciproque est fausse ! Deux variables peuvent être non corrélées (c'est-à-dire de covariance nulle) sans être indépendantes. Considérer par exemple $X \sim \mathscr{U}(-1,0,1)$ et $Y = X^2$.

Les résultats précédents se généralisent à une somme finie de variables aléatoires :

Proposition 3.0.6.
Soit $X_1,\dots,X_n$ des variables aléatoires admettant une variance.

1. $\mathrm{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathrm{V}(X_k) + 2\sum_{1\leqslant i<j\leqslant n} \mathrm{Cov}(X_i,X_j)$.
2. Si les X_k sont deux à deux indépendantes, $\mathrm{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathrm{V}(X_k)$.

Démonstration.
1. Par récurrence à partir du cas $n=2$.
2. Direct. On remarquera l'indépendance mutuelle ne permet pas de conclure, c'est bien l'indépendance deux à deux qui est supposée.

□

Exemple 3.0.7.
Ce dernier résultat est utilisé pour calculer la variance d'une variable aléatoire X telle que $X \sim \mathscr{B}(n,p)$. En effet, si $X_1,\dots,X_n$ sont n variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre p, alors $\sum_{k=1}^{n} X_k \sim \mathscr{B}(n,p)$. Or $\mathrm{V}\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} \mathrm{V}(X_k) = np(1-p)$, par indépen-

dance. Comme $X \sim \sum_{k=1}^{n} X_k$ et comme la variance d'une variable aléatoire ne dépend que de sa loi, alors $\mathrm{V}(X) = \mathrm{V}\left(\sum_{k=1}^{n} X_k\right) = np(1-p)$.

4 Inégalités probabilistes

Dans cette section encore, X et Y sont des variables aléatoires **réelles**.

4.1 Inégalité de Markov

Théorème 4.1.1 (Inégalité de Markov)**.**
Si X est une variable aléatoire réelle **positive** d'espérance finie, alors pour tout $a \in \mathbb{R}_+$,

$$aP(X \geqslant a) \leqslant \mathrm{E}(X).$$

En particulier, si $a > 0$,

$$P(X \geqslant a) \leqslant \frac{\mathrm{E}(X)}{a}.$$

Démonstration.

$$\begin{aligned}
\mathrm{E}(X) &= \sum_{x \in X(\Omega)} xP(X = x) \\
&= \sum_{x \in X(\Omega),\ x<a} xP(X = x) + \sum_{x \in X(\Omega),\ x \geqslant a} xP(X = x) \\
&\geqslant 0 + \sum_{x \in X(\Omega),\ x \geqslant a} aP(X = x) \\
&\geqslant a \sum_{x \in X(\Omega),\ x \geqslant a} P(X = x) \\
&\geqslant aP(X \geqslant a).
\end{aligned}$$

□

Corollaire 4.1.2.
Si X^2 est d'espérance finie et $\varepsilon \in \mathbb{R}_+^*$, $P(|X| \geqslant \varepsilon) \leqslant \dfrac{\mathrm{E}(X^2)}{\varepsilon^2}$.

4.2 Inégalité de Bienaymé-Tchebychev

Elle découle de l'inégalité de Markov.

Théorème 4.2.1 (Inégalité de Bienaymé-Tchebychev)**.**
Si X admet une variance, alors pour tout $\varepsilon \in \mathbb{R}_+^*$,

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

Démonstration.
On observe que $(|X - \mathrm{E}(X)| \geqslant \varepsilon) = ((X - \mathrm{E}(X))^2 \geqslant \varepsilon^2)$, et on applique l'inégalité de Markov :
$\mathrm{V}(X) = \mathrm{E}((X - \mathrm{E}(X))^2) \geqslant \varepsilon^2 P(|X - \mathrm{E}(X)| \geqslant \varepsilon)$. □

Remarque 4.2.2.
Cette inégalité contrôle la déviation d'une variable aléatoire par rapport à sa moyenne.

Exemple 4.2.3.
On jette 3600 fois un dé équilibré à 6 faces. Nous voulons minorer la probabilité que le nombre d'apparitions du numéro 1 soit compris entre 480 et 720 .
Soit S la variable aléatoire comptant le nombre d'apparitions du chiffre 1 au cours de ces lancers.
S suit une loi binomiale de paramètres 3600 et $1/6$. On sait donc que $\mathrm{E}(S) = 600$ et $\mathrm{V}(S) = 500$. Remarquons en outre que

$$480 < S < 720 \iff -120 < S - 600 < 120 \iff |S - 600| < 120$$

Par l'inégalité de Bienaymé-Tchebychev,

$$P(|S-600| \geqslant 120) \leqslant \frac{500}{120^2} = \frac{5}{144}.$$

On en déduit que

$$P(480 < S < 720) \geqslant 1 - \frac{5}{144} = \frac{139}{144}.$$

En particulier, la probabilité que le numéro 1 apparaisse entre 480 et 720 fois au cours de ces 3600 lancers est supérieur à $0,96$.

4.3 Loi faible des grands nombres

Théorème 4.3.1 (Loi faible des grands nombres)**.**
Soit $(X_n)_{n\geqslant 0}$ une suite de variables aléatoires deux à deux indépendantes et suivant une même loi (on dit qu'elles sont ***indépendantes et identiquement distribuées***, souvent abrégé en i.i.d).
Si celles-ci admettent une variance, alors en introduisant m leur espérance commune et

$$S_n = \sum_{k=1}^{n} X_k$$

on a

$$P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \xrightarrow[n\to+\infty]{} 0.$$

Démonstration.
Les X_k ont tous le même écart type σ, et la même variance σ^2. Par indépendance, $\mathrm{E}(S_n) = nm$ et $\mathrm{V}(S_n) = n\sigma^2$.
Grâce à l'inégalité de Bienaymé-Tchebychev, $P(|S_n - nm| \geqslant n\varepsilon) \leqslant \dfrac{n\sigma^2}{n^2\varepsilon^2}$.
Or $(|S_n - nm| \geqslant n\varepsilon) = \left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right)$, donc $P\left(\left|\frac{S_n}{n} - m\right| \geqslant \varepsilon\right) \leqslant \dfrac{\sigma^2}{n\varepsilon^2} \xrightarrow[n\to+\infty]{} 0.$ □

Remarque 4.3.2.
Cette loi exprime que la moyenne des valeurs d'un échantillon modélisant empiriquement une variable aléatoire tend vers l'espérance de cette variable aléatoire quand la taille de l'échantillon tend vers $+\infty$.
Ainsi, si vous lancez un dé à 6 faces un très grand nombre de fois, la fréquence d'apparition du nombre 1 tendra vers $\frac{1}{6}$.

5 Fonctions génératrices

Dans cette section, X et Y sont des variables aléatoires **à valeurs dans** $\mathbb{N}$.

5.1 Définition

Définition 5.1.1 (Fonction génératrice)**.**
On appelle ***fonction génératrice de*** X la série entière $\displaystyle\sum_{n\geqslant 0} P(X=n)t^n$ en la variable t.
Là où elle est définie, on note sa somme $G_X(t) = \displaystyle\sum_{n=0}^{+\infty} P(X=n)t^n$.

Remarque 5.1.2.

1. Par la formule de transfert, si $G_x(t)$ existe, $G_X(t) = \mathrm{E}(t^X)$.
2. Par probabilités totales G_X est toujours définie en 1 et $G_X(1) = 1$.

Théorème 5.1.3.
G_X est de rayon de convergence supérieur ou égal à 1, et converge normalement sur $[-1,1]$.

Démonstration.
Puisque G_X converge en 1, son rayon de convergence est nécessairement supérieur ou égal à 1.
Notons bien que cela n'implique pas la convergence normale sur $[-1,1]$, ni même sur $]-1,1[$, mais seulement sur tout segment inclus dans $]-1,1[$.
Cela dit, G_X est définie sur $[-1,1]$ au moins, et comme pour tout $n\in\mathbb{N}$, $P(X=n)\geqslant 0$,

$\sup_{t\in[-1,1]} |P(X=n)t^n| = P(X=n)$.

Comme $\sum_{n\geqslant 0} P(X=n)$ converge, il y a bien convergence normale sur $[-1,1]$. □

Corollaire 5.1.4.
G_X est définie et continue au moins sur $[-1,1]$, et elle est de classe $\mathscr{C}^\infty$ sur $]-1,1[$.

Proposition 5.1.5.
La connaissance de la loi de X équivaut à la connaissance de sa fonction génératrice G_X.

Démonstration.
($\Rightarrow$) Par définition même, connaître la loi de X permet de connaître toutes les $P(X=n)$ et donc G_X.
($\Leftarrow$) Réciproquement, si l'on connaît G_X, alors par la formule de Taylor, pour tout $n\in\mathbb{N}$, $P(X=n) = \dfrac{G_X^{(n)}(0)}{n!}$ et donc l'on connaît la loi de X. □

Remarque 5.1.6.
Ce dernier résultat implique donc que si $G_X = G_Y$ si et seulement si $X \sim Y$.
On peut donc montrer que deux variables aléatoires suivent la même loi en montrant qu'elles ont la même fonction génératrice.

5.2 Fonctions génératrices des lois usuelles

Proposition 5.2.1.

1. Si $X \sim \mathscr{B}(p)$, alors G_X est de rayon de convergence infini et pour tout $t\in\mathbb{R}$, $G_X(t) = (1-p)+pt$;
2. Si $X \sim \mathscr{B}(n,p)$, alors G_X est de rayon de convergence infini et pour tout $t\in\mathbb{R}$, $G_X(t) = \sum_{k=0}^{n} \binom{n}{k} (pt)^k(1-p)^{n-k} = (1-p+pt)^n$;
3. Si $X \sim \mathscr{G}(p)$, alors G_X est de rayon de convergence $\dfrac{1}{1-p}$ et pour tout $t \in \left]-\frac{1}{1-p}, \frac{1}{1-p}\right[$, $G_X(t) = \sum_{n=0}^{+\infty} pt^n(1-p)^{n-1} = \dfrac{pt}{1-(1-p)t}$;
4. Si $X \sim \mathscr{P}(\lambda)$, alors G_X est de rayon de convergence infini et pour tout $t\in\mathbb{R}$, $G_X(t) = \sum_{n=0}^{+\infty} \mathrm{e}^{-\lambda}\dfrac{(\lambda t)^n}{n!} = \mathrm{e}^{\lambda(t-1)}$.

5.3 Fonction génératrice, espérance et variance

Théorème 5.3.1.
X a une espérance finie si et seulement si G_X est dérivable en 1, et dans ce cas $\mathrm{E}(X) = G_X'(1)$.

Démonstration.
Commençons par remarquer que si le rayon de convergence de G_X est strictement supérieur à 1, G_X est de toute façon de classe $\mathscr{C}^\infty$ sur un intervalle ouvert contenant 1, donc elle est dérivable en 1. Sa série dérivée converge alors en 1, et il se trouve qu'en 1 elle vaut $\sum_{n=0}^{+\infty} nP(X=n)$: X est donc d'espérance finie.
Le cas où le rayon de convergence de G_X vaut 1 est plus compliqué. Dans ce cas G_X n'est pas définie au-delà de $[-1,1]$, et donc sa dérivabilité en 1 équivaut à sa dérivabilité à gauche en 1.
($\Rightarrow$) Si $\mathrm{E}(X)$ existe et est finie, alors $\sum_{n=0}^{+\infty} nP(X=n)$ converge, et même absolument car elle est à termes positifs. Or les fonctions $u_n : t\mapsto P(X=n)t^n$ sont de classe $\mathscr{C}^1$ et $\sum_{n\geqslant 0} u_n'(t) = \sum_{n=1}^{+\infty} nP(X=n)t^{n-1}$.
Alors la série de fonctions $\sum_{n\geqslant 0} u_n'(t)$ converge normalement sur $[-1,1]$.
Le théorème de dérivation d'une série de fonctions assure alors que G_X est dérivable (à gauche) en 1, et que $G_X'(1) = \sum_{n=0}^{+\infty} nP(X=n) = \mathrm{E}(X)$.

($\Leftarrow$) La preuve de cette réciproque n'est pas exigible.
Réciproquement, supposons G_X dérivable à gauche en 1 , et soit $t \in [0,1[$. Le théorème des accroissements finis montre que :

$$\exists\, c_t \in]t,1[\text{ tq } \frac{G_X(t)-G_X(1)}{t-1} = G'_X(c_t) = \sum_{n=1}^{+\infty} n\mathbb{P}(X=n)c_t^{n-1}.$$

Tous les termes de la série étant positifs on a donc

$$\forall\, N \in \mathbb{N},\ \sum_{n=1}^{N} nP(X=n){c_t}^{n-1} \leqslant \frac{G_X(t)-G_X(1)}{t-1}.$$

Donc en faisant tendre t vers 1 on a $\forall\, N \in \mathbb{N},\ \sum_{n=1}^{N} nP(X=n) \leqslant G'_X(1)$.

Cela implique que la série à termes positifs $\sum_{n\in\mathbb{N}} n\mathbb{P}(X=n)$ converge, donc $\mathrm{E}(X)$ existe et $\mathrm{E}(X) \leqslant G'_X(1)$.
Mais en reprenant la première relation on a alors

$$\frac{G_X(t)-G_X(1)}{t-1} = \sum_{n=1}^{+\infty} nP(X=n){c_t}^{n-1} \leqslant \sum_{n=1}^{+\infty} nP(X=n),$$

et en faisant tendre à nouveau t vers 1, on obtient l'autre inégalité. □

Théorème 5.3.2.
X a une variance si et seulement si G_X est dérivable deux fois en 1, et dans ce cas

$$\begin{aligned}\mathrm{E}(X(X-1)) &= G''_X(1)\\ \mathrm{E}(X^2) &= G'_X(1)+G''_X(1)\\ \mathrm{V}(X) &= G'_X(1)+G''_X(1)-(G'_X)^2(1).\end{aligned}$$

Démonstration.
Le programme n'est pas très clair quant à l'exigibilité de cette démonstration.
Le premier point se démontre comme pour le théorème précédent, nous n'en dirons pas plus.
Par contre il faut savoir calculer une espérance et une variance à partir de G_X.
Pour les dernières égalités, remarquons que $\mathrm{V}(X) = \mathrm{E}(X(X-1)) + \mathrm{E}(X) - \mathrm{E}^2(X)$. □

Exercice 5.3.3.
Retrouver les espérances et les variances des lois présentées dans la proposition **5.2.1** en utlisant G_X.

5.4 Fonction génératrice d'une somme de variables aléatoires

Théorème 5.4.1.
Si X et Y sont indépendantes, alors $G_{X+Y} = G_X + G_Y$.

Démonstration.
Soit $t \in [-1,1]$.
Tout part de $G_X = \mathrm{E}(t^X)$.
En effet, si X et Y sont indépendantes, t^X et t^Y aussi.
Alors $G_{X+Y}(t) = \mathrm{E}\big(t^{X+Y}\big) = \mathrm{E}\big(t^X t^Y\big) = \mathrm{E}(t^X)\mathrm{E}(t^Y) = G_X(t)G_Y(t)$. □

Corollaire 5.4.2.
Si $X_1,\dots,X_n$ sont des variables aléatoires mutuellement indépendantes, $G_{X_1+\ldots+X_n})G_{X_1} \times \ldots \times G_{X_n}$.

Exemple 5.4.3.
1. Sachant qu'une loi $\mathscr{B}(n,p)$ peut être simulée par la somme de n loi $\mathscr{B}(p)$ indépendantes, on retrouve que si $X \sim \mathscr{B}(n,p)$ alors $G_X(t) = (1-p+pt)^n$.
2. Soit X et Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs λ et μ. Alors $X+Y \sim \mathscr{P}(\lambda+\mu)$.

6 Annexe : démonstration de la formule de transfert

Démonstration.
(merci à B. Arsac).
Notons $Y = f(X) = f \circ X$. Et commençons par nous occuper de la première proposition : « $f(X)$ est d'espérance finie si et seulement si la famille $(P(X=x)\ f(x))_{x\in X(\Omega)}$ est sommable ».
Par définition, Y est d'espérance finie si et seulement si la famille $(P(Y=y)y)_{y\in Y(\Omega)}$ est sommable. Par définition encore, Y est d'espérance finie si et seulement si la famille $(P(Y=y)|y|)_{y\in Y(\Omega)}$ est sommable.

Rappelons que $X(\Omega)$ est fini ou dénombrable, et que par conséquent $Y(\Omega) = f(X(\Omega))$ l'est.

On a $Y(\Omega) = f(X(\Omega))$. Il est donc naturel de partitionner $X(\Omega)$ en classant ses éléments suivant leur image par f. On définit donc, pour tout $y \in Y(\Omega)$,

$$I_y = f^{-1}(\{y\}) = \{x \in X(\Omega)\ ;\ f(x) = y\}$$

L'évènement $(Y = y)$ est l'évènement $(f(X) = y)$, qui peut donc s'écrire $\left(X \in f^{-1}(\{y\})\right)$, ou encore $(X \in I_y)$. Par additivité ou σ-additivité (suivant si I_k est fini ou non),

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

(la sommabilité de la famille au second membre est automatique, par σ-additivité de la probabilité P). C'est-à-dire

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

Multiplions par $|y|$, en remarquant que $f(x) = y$ pour tout $x \in I_y$:

$$|y|\ P(Y = y) = \sum_{x \in I_y} |f(x)| P(X = x)$$

Comme $X(\Omega)$ est la réunion disjointe des I_y, c'est-à-dire

$$X(\Omega) = \bigcup_{y \in Y(\Omega)} I_y \qquad \text{et} \qquad y \neq y' \Rightarrow I_y \cap I_{y'} = \varnothing$$

le théorème de sommabilité par paquets dit que la famille $(|y|\ P(Y = y))_{y \in Y(\Omega)}$ est sommable si et seulement si la famille $(|f(x)| P(X = x))_{x \in X(\Omega)}$ l'est. On obtient donc la première proposition.

Pour la deuxième proposition, le travail est presque fait : la sommabilité de la famille $(P(X = x)\ f(x))_{x \in X(\Omega)}$ permet de faire une sommation par paquets en utilisant la partition $(I_y)_{y \in Y(\Omega)}$
(on note toujours $Y = f(X) = f \circ X$) :

$$\begin{aligned}\sum_{x \in X(\Omega)} P(X = x)\ f(x) &= \sum_{y \in Y(\Omega)} \left(\sum_{x \in I_y} P(X = x) f(x) \right) = \sum_{y \in Y(\Omega} \left(y \sum_{x \in I_y} P(X = x) \right) \\ &= \sum_{y \in Y(\Omega)} y P(Y = y) = E(Y) = E(f(X))\end{aligned}$$

□

7 Exercices classiques

7.1 Calculs d'espérance et de variance

Une secrétaire effectue, une première fois, un appel téléphonique vers n correspondants distincts.
On admet que les n appels constituent n expériences indépendantes et que, pour chaque appel, la probabilité d'obtenir le correspondant demandé est p ($p \in]0, 1[$).
Soit X la variable aléatoire représentant le nombre de correspondants obtenus.

1. Donner la loi de X. Justifier.
2. La secrétaire rappelle une seconde fois, dans les mêmes conditions, chacun des $n - X$ correspondants qu'elle n'a pas pu joindre au cours de la première série d'appels. On note Y la variable aléatoire représentant le nombre de personnes jointes au cours de la seconde série d'appels.
   a) Soit $i \in [\![0, n]\!]$. Déterminer, pour $k \in \mathbb{N}$, $P_{(X=i)}(Y = k)$.
   b) Prouver que $Z = X + Y$ suit une loi binomiale dont on déterminera le paramètre.
   Indication : on pourra utiliser, sans la prouver, l'égalité suivante : $\binom{n-i}{k-i}\binom{n}{i} = \binom{k}{i}\binom{n}{k}$.
   c) Déterminer l'espérance et la variance de Z.

7.2 Un couple de variables aléatoires

Soient X et Y deux variables aléatoires définies sur un même espace probabilisé $(\Omega, \mathscr{A}, P)$ et à valeurs dans $\mathbb{N}$.
On suppose que la loi du couple (X, Y) est donnée par :
$\forall (i, j) \in \mathbb{N}^2$, $P((X = i) \cap (Y = j)) = \dfrac{1}{\mathrm{e}\ 2^{i+1} j!}$

1. Déterminer les lois de X et de Y.
2. a) Prouver que $1 + X$ suit une loi géométrique et en déduire l'espérance et la variance de X.

b) Déterminer l'espérance et la variance de Y.

3. Les variables X et Y sont-elles indépendantes ?
4. Calculer $P(X = Y)$.

7.3 Inégalité de Bienaymé-Tchebychev

1. Rappeler l'inégalité de Bienaymé-Tchebychev.
2. Soit (Y_n) une suite de variables aléatoires indépendantes, de même loi et et telle que $\forall n \in \mathbb{N}$, $Y_n \in L^2$.
   On pose $S_n = \sum_{k=1}^{n} Y_k$.
   Prouver que : $\forall\, a \in]0, +\infty[$, $P\left(\left|\dfrac{S_n}{n} - E(Y_1)\right| \geqslant a\right) \leqslant \dfrac{V(Y_1)}{na^2}$.
3. **Application** : On effectue des tirages successifs, avec remise, d'une boule dans une urne contenant 2 boules rouges et 3 boules noires.
   À partir de quel nombre de tirages peut-on garantir à plus de 95% que la proportion de boules rouges obtenues restera comprise entre 0, 35 et 0, 45 ?
   Indication : considérer la suite (Y_i) de variables aléatoires de Bernoulli où Y_i mesure l'issue du $i^{\text{ème}}$ tirage.

7.4 Calculs d'espérance et de variance grâce à la fonction génératrice

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ dont la fonction génératrice est

$$G_X(t) = \frac{t}{2 - t^2} \qquad \text{pour tout } t \in]-\sqrt{2}, \sqrt{2}[$$

1. Calculer la loi de X.
2. Reconnaître la loi de $Y = \dfrac{1}{2}(X + 1)$. En déduire l'espérance et la variance de X.

7.5 Détermination d'une fonction génératrice

Soit X une variable aléatoire à valeurs dans $\mathbb{N}$, de loi de probabilité donnée par : $\forall n \in \mathbb{N}$, $P(X = n) = p_n$.
La fonction génératrice de X est notée G_X et elle est définie par $G_X(t) = E[t^X] = \sum_{n=0}^{+\infty} p_n t^n$.

1. Prouver que l'intervalle $]-1, 1[$ est inclus dans l'ensemble de définition de G_X.
2. Soit X_1 et X_2 deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.
   On pose $S = X_1 + X_2$.
   Démontrer que $\forall t \in]-1, 1[$, $G_S(t) = G_{X_1}(t) G_{X_2}(t)$:

   a) en utilisant le produit de Cauchy de deux séries entières.

   b) en utilisant uniquement la définition de la fonction génératrice par $G_X(t) = E[t^X]$.

   Remarque : on admettra, pour la question suivante, que ce résultat est généralisable à n variables aléatoires indépendantes à valeurs dans $\mathbb{N}$.
3. Un sac contient quatre boules : une boule numérotée 0, deux boules numérotées 1 et une boule numérotée 2.
   Soit $n \in \mathbb{N}^*$. On effectue n tirages successifs, avec remise, d'une boule dans ce sac.
   On note S_n la somme des numéros tirés.
   Soit $t \in]-1, 1[$.
   Déterminer $G_{S_n}(t)$ puis en déduire la loi de S_n.