Dans ce chapitre, $(\Omega, \mathcal{A}, P)$ désigne un espace probabilisé.

1 Définitions

1.1 Variables aléatoires discrètes

Définition 1.1.1 (Variable aléatoire discrète) 1. *On appelle variable aléatoire discrète sur Ω à valeurs dans E (un ensemble) toute application X de Ω dans E telle que :*
- *L'image $X(\Omega)$ est au plus dénombrable (fini ou dénombrable).*
- *Pour tout x de $X(\Omega)$, l'image réciproque $X^{-1}(\{x\})$ appartient à la tribu $\mathcal{A}$.*

2. *Pour toute partie A de E, l'événement $X^{-1}(A)$ est noté $\{X \in A\}$ ou $(X \in A)$. Dans le cas où A est un singleton $\{x\}$, on emploie plutôt les notations $\{X = x\}$ ou $(X = x)$.*
3. *Dans la plupart des situations, E est un sous-ensemble de $\mathbb{R}$ et on dit dans ce cas que X est une variable aléatoire réelle. Dans ce cas, si $A =]-\infty, x]$ pour un certain x de $\mathbb{R}$, on emploie plutôt les notations $\{X \leq x\}$ ou $(X \leq x)$. Pour les autres intervalles, nous avons le même type de notations.*
4. *Si $E = \mathbb{R}^n$, alors $X = (X_1, ..., X_n)$ est appelé vecteur aléatoire. Pour i dans $[\![1, n]\!]$, les X_i sont des fonctions définies sur Ω à valeurs réelles*

Remarque 1.1.1 1. *Une variable aléatoire est donc une application (et pas une variable) et elle n'est pas aléatoire ! !*
2. *Ainsi $X(\Omega)$ peut s'écrire $X(\Omega) = \{x_1, ..., x_n\}$ si $X(\Omega)$ est fini ou $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$ si $X(\Omega)$ est dénombrable.*
3. *IMPORTANT $(X = x)_{x \in X(\Omega)}$ est un système complet d'événements de Ω.*

Proposition 1.1.1 (Image réciproque d'un ensemble) *Soit $U \subset X(\Omega)$. Alors $X^{-1}(U)$ est aussi un événement.*

Démonstration : On a : $X^{-1}(U) = \bigcup_{x \in U} X^{-1}(\{x\})$. Comme U est par hypothèse une partie d'un ensemble au plus dénombrable, U est au plus dénombrable.
Il suffit donc de savoir que pour tout $x \in X(\Omega)$, on a : $X^{-1}(\{x\}) \in \mathcal{A}$ pour obtenir par réunion : $X^{-1}(U) \in \mathcal{A}$, car une tribu est stable par réunion finie ou dénombrable.

Proposition 1.1.2 (Variable aléatoire $f(X)$) *Soit $X : \Omega \to E$ une variable aléatoire discrète et $f : E \to E'$ une application. Alors $f \circ X$ est une variable aléatoire que l'on note $f(X)$.*

Démonstration : Nous avons $f(X)(\Omega) = f(X(\Omega)) = \{f(x_n),\ n \in I\}$, si on note $X(\Omega) = \{x_n,\ n \in I\}$, avec I un ensemble fini ou $I = \mathbb{N}$. Ainsi $f(X)(\Omega)$ est au plus dénombrable. Soit $x' \in f(X)(\Omega)$. Nous avons : $f(X)^{-1}(\{x'\}) = \{\omega \in \Omega,\ f(X(\omega)) = x'\} = \{\omega \in \Omega,\ X(\omega) \in f^{-1}(\{x'\})\} = X^{-1}\left(f^{-1}(\{x'\}) \cap X(\Omega)\right)$ qui est dans $\mathcal{A}$ grâce à la proposition précédente, car X est une variable aléatoire discrète.

Exemple 1.1.1 1. *Soit (E_n) une suite d'événements de $(\Omega, \mathcal{A}, P)$ et $Z = \sum_{n=0}^{+\infty} 1_{E_n}$. Montrer que Z est une variable aléatoire discrète.*

2. *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé. Soit $k \in \mathbb{N}^*$. On pose $C_k = \displaystyle\bigcup_{n\in\mathbb{N}} \bigcap_{p\geq n} \left(|X_p - X| \leq \frac{1}{k}\right)$.*

   On pose $B = \{\omega \in \Omega,\ \lim_{n\to+\infty} X_n(\omega) = X(\omega)\}$.

   (a) *Montrer que $P(B) = \lim_{k\to+\infty} P(C_k)$.*

   (b) *Si pour tout $\varepsilon \in \mathbb{R}_+^*$, la série $\sum P(|X_n - X| > \varepsilon)$ converge, alors (X_n) converge presque-sûrement vers X.*

1.2 Loi d'une variable aléatoire

Définition 1.2.1 (Loi de probabilité de X) *Soit $X : \Omega \to E$ une variable aléatoire discrète. On pose :*

$$P_X : \left\{ \begin{array}{ccc} X(\Omega) & \to & [0,1] \\ x & \mapsto & P(X = x) \end{array} \right. .$$

On a aussi : $\forall A \in \mathcal{P}(X(\Omega)),\ P(X \in A) = \displaystyle\sum_{x\in A} P(X = x)$.

On pose : $\forall A \in \mathcal{P}(X(\Omega)), P_X(A) = P(X \in A)$.
P_X est appelée loi de probabilité de X.

Proposition 1.2.1 (Loi de probabilité définit une probabilité) *L'application*
$P_X : \left\{ \begin{array}{ccc} \mathcal{P}(X(\Omega)) & \to & [0,1] \\ A & \mapsto & P(X \in A) \end{array} \right.$ *définit une probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$.*

Démonstration : La tribu que l'on considère sur $X(\Omega)$ est $\mathcal{P}(X(\Omega))$, car $X(\Omega)$ est au plus dénombrable.

- P_X est à valeurs dans $[0,1]$, car P est à valeurs dans $[0,1]$.
- $P_X(X(\Omega)) = P(X \in X(\Omega)) = 1$, car $(X \in X(\Omega)$ est toujours réalisé.)
- Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux disjoints de $X(\Omega)$, on a :
  $P_X(\displaystyle\bigcup_{n\in\mathbb{N}} A_n) = P(\{\omega \in \Omega,\ X(\omega) \in \bigcup_{n\in\mathbb{N}} A_n) = P(\bigcup_{n\in\mathbb{N}} \{\omega \in \Omega,\ X(\omega) \in A_n\})$. Comme les A_n sont

deux à deux disjoints, alors la réunion $\bigcup_{n\in\mathbb{N}}\{\omega\in\Omega,\ X(\omega)\in A_n\}$ l'est aussi et donc :

$$P_X(\bigcup_{n\in\mathbb{N}} A_n)=\sum_{n=0}^{+\infty}P(\{\omega\in\Omega,\ X(\omega)\in A_n\})=\sum_{n=0}^{+\infty}P(X\in A_n)=\sum_{n=0}^{+\infty}P_X(A_n).$$

Remarque 1.2.1 *1. Donner la loi d'une variable aléatoire, c'est donner $X(\Omega)=\{x_0,...,x_n,...\}$ (espace de départ de P_X) puis les $P(X=x_i)$ pour tout i de $\mathbb{N}$ (les valeurs de $P_X(x_i)$).*

2. *On a : $\displaystyle\sum_{x\in X(\Omega)}P(X=x)=1$, car $(X=x)_{x\in X(\Omega)}$ forme un système complet d'événements.*
3. *Une variable aléatoire permet de transporter la probabilité P définie sur Ω vers une probabilité sur les valeurs de X. On s'intéresse plus qu'aux valeurs de X et on raisonne sur un espace probabilisé plus restreint $(X(\Omega),\mathcal{P}(X(\Omega)),P_X)$, mais qui nous donne juste l'information dont on a besoin. Ainsi lorsque l'on étudie une variable aléatoire, on ne s'intéresse pas vraiment à l'espace probabilisé de départ $(\Omega,\mathcal{A},P)$.*
4. *Ainsi lorsqu'une expérience modélise une succession infinie d'expériences succès-échecs indépendantes, nous avons eu des difficultés pour définir un espace probabilisé adéquat. Mais si on s'intésse par exemple à la variable aléatoire X donnant le moment du premier succès, qui est à valeurs dans $\mathbb{N}^*\cup\{+\infty\}$, on déplace la difficulté sur $X(\Omega)$ qui est dénombrable et il est plus simple de définir un espace probabilisé sur $X(\Omega)$.*
   Ainsi on déplace le problème sur un espace de taille plus petite contenant seulement l'information dont on a besoin, ce qui permet de s'affranchir de la difficulté théorique posée par l'espace de toutes les expériences infinies succès-échecs.

   Attention, si on a : $P(X=x)=0$, cela ne veut pas nécessairement dire que la valeur x n'est pas atteinte par X. Cela signifie que l'événement $(X=x)$ est quasi impossible.

Définition 1.2.2 (Variables aléatoires de même loi) *On dit que deux variables aléatoires disctrètes X et Y ont même loi si $P_X=P_Y$. On note alors $X\sim Y$.*

Remarque 1.2.2 *1. ATTENTION : deux variables aléatoires peuvent avoir la même loi sans être égales. Par exemple on lance une pièce équilibrée. Soit X la variable aléatoire qui vaut 1 si on tombe sur pile et 0 si on tombe sur face. On pose $Y=1-X$. Alors X et Y ont la même loi sans être égales.*
En effet Y vaut 1 si on tombe sur face et 0 si on tombe sur pile. Ainsi
$P(Y=0)=P(X=0)=1/2$ (autant de chances d'avoir pile ou face) et
$P(Y=1)=P(X=1)=1/2$.

2. *$X_1,...,X_n,...$ sont dites identiquement distribuées si elles ont toutes la même loi.*
3. *Si $X\sim Y$, alors X et Y ne sont pas forcément définies sur le même espace probabilisé.*

Exemple 1.2.1 *1. $X:\Omega\to\{c\}$ est une variable aléatoire constante. Une variable aléatoire X est quasi constante ou certaine s'il existe $x_0\in X(\Omega)$ tel que $P(X=x_0)=1$, ce qui signifie que : $\forall x\in X(\Omega)\setminus\{x_0\},\ P(X=x)=0$.*

2. *Soit X une variable aléatoire réelle. Déterminer $\lim\limits_{x\to+\infty}P(X\le x)$ et $\lim\limits_{x\to-\infty}P(X\le x)$.*

3. *Une urne contient n boules blanches numérotées de 1 à n et m boules noires numérotées de 1 à m. On tire une à une et sans remise toutes les boules de l'urne. Soit X la variable aléatoire indiquant le nombre de tirages effectués jusqu'au retrait de la première boule blanche et Y la variable aléatoire indiquant le nombre de tirages effectués jusqu'au retrait de la première boule numérotée 1.*

   (a) *Déterminer la loi de X.*

   (b) *Déterminer la loi de Y.*

   (c) *En déduire que pour p, q dans $\mathbb{N}^*$, avec : $p \leq q$, on a : $\displaystyle\sum_{k=p}^{q} \binom{k}{p} = \binom{q+1}{p+1}$.*

4. *On effectue une série de lancers d'une pièce, avec une probabilité p d'avoir Pile. Soit X la variable aléatoire comptant le nombre de lancers nécessaires afin d'obtenir exactement r Pile, avec $r \in \mathbb{N}^*$. Déterminer $P(X = n)$, pour n dans $\mathbb{N}$.*

5. *Un joueur prélève n boules successivement avec remise dans une urne contenant N boules numérotées de 1 à N. On considère les variables aléatoires réelles Y et Z égales respectivement au plus grand et au plus petit numéro des n boules prélevées. Donner les lois de Y et Z.*

Proposition 1.2.2 (Loi de $f(X)$) *Soit X une variable aléatoire discrète sur $(\Omega, \mathcal{A}, P)$ et f une application définie sur $X(\Omega)$. La loi de $f(X)$ est donnée par :*

$$\forall y \in f(X)(\Omega),\ P(f(X) = y) =$$

Démonstration : Soit $y \in f(X)(\Omega) = f(X(\Omega))$. On a : $(f(X) = y) = \{\omega \in \Omega,\ f(X(\omega)) = y\} = \{\omega \in \Omega,\ X(\omega) \in f^{-1}(\{y\})\} = (X \in f^{-1}(\{y\})) = \bigcup_{x \in f^{-1}(\{y\})} (X = x)$. L'ensemble $f^{-1}(\{y\})$ est au plus dénombrable, car c'est un partie de $X(\Omega)$ qui est dénombrable. Ainsi $\bigcup_{x \in f^{-1}(\{y\})} (X = x)$ est une union au plus dénombrable et disjointe.

Proposition 1.2.3 ($X \sim Y \Rightarrow f(X) \sim f(Y)$) *Soient X et Y deux variables aléatoires discrètes de même loi, définies sur respectivement $(\Omega_1, \mathcal{A}_1, P_1)$ et $(\Omega_2, \mathcal{A}_2, P_2)$.*
Soit f une application définie sur $X(\Omega_1) = Y(\Omega_2)$. Alors on a $f(X) \sim f(Y)$.

Démonstration : Tout d'adord, on a bien : $f(X)(\Omega_1) = f(Y)(\Omega_2)$, car $X(\Omega_1) = Y(\Omega_2) = Z$.
Soit $z \in Z$. Grâce à démonstration de la proposition précédente,
$(f(X) = z) = \bigcup_{x \in f^{-1}(\{z\})} (X = x)$ et $(f(Y) = z) = \bigcup_{x \in f^{-1}(\{z\})} (Y = x)$. Les unions précédentes étant disjointes, alors $P(f(X) = z) = \sum_{x \in f^{-1}(\{z\})} P(X = x) = \sum_{x \in f^{-1}(\{z\})} P(Y = x) = P(f(Y) = z)$, car X et Y ont la même loi.

2 Lois usuelles

2.1 Loi uniforme

Définition 2.1.1 (Loi uniforme) *Soient X une variable aléatoire et E un ensemble fini. On dit que X suit la loi uniforme sur E, si :*

$$\forall e \in E,\ P(X = e) = \frac{1}{|E|}.$$

Dans ce cas pour tout A de $\mathcal{P}(E)$, on a : $P(X \in A) = \dfrac{|A|}{|E|}$.
On note $\mathcal{U}(E)$ cette loi et on écrit $X \sim \mathcal{U}(E)$.

Cette situation se rencontre lorsque l'on effectue des tirages équiprobables sur un espace fini, ce qui est souvent le cas lorsque l'on effectue un tirage au hasard.

Exemple 2.1.1 *On répète le lancer d'un dé (que l'on suppose équilibré) à 12 faces jusqu'à ce que le dé produise un résultat pair que l'on divise finalement par 2. La variable aléatoire X prend la valeur obtenue. Quelle est la loi de X ?*

X est à valeurs dans $[\![1,6]\!]$. Soit $k \in [\![1,6]\!]$.

Pour avoir $(X = k)$, le premier nombre pair obtenu doit être $2k$.

Soit A_l : « obtenir $2k$ au lancer l et avant que des impairs ».
On a : $P(A_l) = \left(\frac{1}{2}\right)^{l-1} \frac{1}{12}$, par indépendance des lancers.

De plus, on a : $(X = k) = \bigcup_{l=1}^{+\infty} A_l$ et cette union est disjointe, donc :

$P(X = k) = \sum_{l=1}^{+\infty} \left(\frac{1}{2}\right)^{l-1} \frac{1}{12} = \frac{1}{12} \sum_{p=0}^{+\infty} \left(\frac{1}{2}\right)^{p} = \frac{1}{12} \times \frac{1}{1-\frac{1}{2}} = \frac{1}{6}$. Ainsi $X \sim \mathcal{U}([\![1,6]\!])$.

2.2 Loi de Bernoulli

Définition 2.2.1 (Loi de Bernoulli) *Soit $p \in [0,1]$. On dit qu'une variable aléatoire X suit la loi de Bernoulli de paramètre p si*

$$X(\Omega) = \{0,1\} \quad \text{et} \quad P(X=1) = p\ ,\ P(X=0) = 1-p.$$

On note $\mathcal{B}(p)$ cette loi et on écrit $X \sim \mathcal{B}(p)$.

La loi de Bernoulli modélise toute expérience à deux issues que l'on appelle succès ou échec. X vaut 1 en cas de succès et 0 en cas d'échec.

Exemple 2.2.1

1. *Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $A \in \mathcal{A}$ un événement de probabilité p.*
   Alors $1_A : \left\{ \begin{array}{ccl} \Omega & \to & \{0;1\} \\ x & \mapsto & \left\{ \begin{array}{lll} 1 & si & x \in A \\ 0 & si & x \notin A \end{array} \right. \end{array} \right.$ est une variable aléatoire discrète,
   car $1_A(\Omega) = \{0,1\}$ est fini et $1_A^{-1}(\{1\}) = A$ et $1_A^{-1}(\{0\}) = \overline{A}$ sont dans $\mathcal{A}$.

   De plus $1_A \sim \mathcal{B}(P(A))$, car $P(\{\omega \in \Omega,\ 1_A(\omega) = 1\} = P(A)$.
   Ainsi une variable aléatoire suivant une loi de Bernoulli modélise toute expérience où un événement A se produit avec une probabilité p et ne se produit pas avec une probabilité $1-p$.

2. *Soient $A_1, ..., A_n$ des événements de Ω. Que représente la variable aléatoire $\sum_{k=1}^{n} 1_{A_k}$?*

2.3 Loi Binomiale

Définition 2.3.1 (Loi Binomiale) *Soient $n \in \mathbb{N}^*$ et $p \in [0,1]$. On dit qu'une variable aléatoire X suit la loi binomiale de paramètres (n,p) si*

$$X(\Omega) = [\![0,n]\!] \quad \text{et} \quad \forall k \in [\![0,n]\!],\ \ P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

On note $\mathcal{B}(n,p)$ cette loi et on écrit $X \hookrightarrow \mathcal{B}(n,p)$.

Une expérience peut se modéliser avec une loi binomiale $\mathcal{B}(n,p)$ lorsque l'on effectue n expériences aléatoires indépendantes à deux issues : succès et échec avec une probabilité p d'avoir un succès et donc une probabilité $1-p$ d'avoir un échec. Dans ce cas la variable aléatoire X compte le nombre de succès lors de ces n expériences.

Remarque 2.3.1 *Si on a $X \hookrightarrow \mathcal{B}(n,p)$, alors $n-X \hookrightarrow \mathcal{B}(n,1-p)$, car ici on compte le nombre d'échecs.*

Exemple 2.3.1 *Soit $n \in \mathbb{N}^*$ et $p \in]0,1[$. Soit S_n une variable aléatoire suivant une loi $\mathcal{B}(n,p)$. Montrer que $p = 1/2$ si et seulement si : $\forall n \in \mathbb{N}^*, \ \max_{k \in [\![0,2n]\!]} P(S_{2n} = k) \leq P(S_{2n} = 0) \quad (*)$.*

2.4 Loi géométrique

Soit $p \in]0;1]$. Dans ce paragraphe, on posera $q = 1 - p$.

Définition 2.4.1 (Loi géométrique) *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $p \in]0;1]$. On appelle variable aléatoire géométrique de paramètre p une variable aléatoire X telle que $X(\Omega) = \mathbb{N}^*$ et*

$$\forall k \in \mathbb{N}^*, \quad P(X = k) = p(1-p)^{k-1} = pq^{k-1}.$$

On note $X \sim \mathcal{G}(p)$, et l'on dit que X suit une loi géométrique de paramètre p.

Remarque 2.4.1 *On a bien : $\displaystyle\sum_{k=1}^{+\infty} p(1-p)^{k-1} = p \sum_{l=0}^{+\infty} (1-p)^l = p \times \frac{1}{1-(1-p)} = 1.$*

Exemple 2.4.1 *Si $X \sim \mathcal{G}(p)$, alors :*
$\forall n \in \mathbb{N}, \ P(X > n) =$

Proposition 2.4.1 (Interprétation d'une loi géométrique) *On considère une succession infinie d'expériences mutuellement indépendantes succès-échec où la probabilité d'un succès pour chacune des expériences est p (épreuves de Bernoulli de paramètre p). Soit X donnant le rang du premier succès. Alors X suit une loi $\mathcal{G}(p)$.*

Démonstration : Il s'agit d'une loi binomiale négative pour $m = 1$ (revoir l'exemple 1.2.1).

Remarque 2.4.2

1. *En reprenant les notations précédentes, on a $P(X = +\infty) = 0$.*
   En effet :

2. *Ceci permet de simuler par exemple un jeu de pile ou face où l'on se demande au bout de combien de lancers on va obtenir pile pour la première fois.*

Exemple 2.4.2

1. *On dispose initialement d'une urne constituée d'une boule blanche, et d'une pièce de monnaie équilibrée. On effectue des lancers de la pièce selon la règle suivante :*
   - *si on obtient « Face » on ajoute une boule noire dans l'urne et on relance la pièce ;*
   - *si on obtient « Pile » on tire une boule dans l'urne et on arrête l'expérience.*

Quelle est la probabilité d'obtenir une boule blanche à la fin de l'expérience ?

2. *Soit X une variable aléatoire à valeurs dans $\mathbb{N}^*$ telle que : $\forall n \in \mathbb{N},\ P(X > n) > 0$. Montrer que X suit une loi géométrique si et seulement si : $\forall n, k \in \mathbb{N},\ P(X > n+k | X > n) = P(X > k)$.*
   - *On suppose que X suit une loi $\mathcal{G}(p)$.*

   Grâce à l'exemple 2.4.1, on a : $P(X > n+k|X > n) = \dfrac{P\left((X > n+k) \cap (X > n)\right)}{P(X > n)} = \dfrac{P(X > n+k)}{P(X > n)} = \dfrac{(1-p)^{n+k}}{(1-p)^n} = (1-p)^k = P(X > k).$

2.5 Loi de Poisson

Définition 2.5.1 (Loi de Poisson) *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisable et $\lambda > 0$.*
On appelle variable aléatoire de Poisson de paramètre λ une variable aléatoire X telle que $X(\Omega) = \mathbb{N}$ et

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

On note $X \sim \mathcal{P}(\lambda)$.

Remarque 2.5.1 *On a bien :* $\displaystyle\sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{+\infty}\frac{\lambda^k}{k!} = e^{-\lambda}.e^{\lambda} = 1.$

Exemple 2.5.1 *Soit $(X_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires telles que pour tout n de $\mathbb{N}$, on ait : $X_n \sim \mathcal{B}(n, p_n)$, avec : $\displaystyle\lim_{n\to+\infty} np_n = \lambda > 0$. Montrer que : $\forall k \in \mathbb{N},\ \displaystyle\lim_{n\to+\infty} P(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$.*

Remarque 2.5.2 *Nous remarquons que lorsque n tend vers $+\infty$, alors p_n tend vers 0.*
Ainsi une loi de Poisson modélise un nombre d'apparitions d'un événement rare (p_n est petit) dans une suite très grande (n grande) d'événements. Dans les occurences d'un événement rare, on peut remplacer la variable aléatoire de loi $\mathcal{B}(n, p_n)$ qui modélise le nombre d'événements rares réalisés au cours de n expériences par une variable aléatoire de loi $\mathcal{P}(np_n)$.
Ceci a l'avantage de simplifier les calculs et permet de travailler que sur un paramètre au lieu de deux pour la loi binomiale. Par ailleurs comme nous le verrons plus tard, ces deux variables aléatoires ont la même espérance. Ainsi, par ce remplacement, on ne change pas le nombre moyen d'apparition de ces événements rares.
Les lois de Poisson se rencontrent concrètement pour modéliser le nombre d'événements d'un certain type qui se produisent dans une période de temps donnée. Ce nombre peut être le nombre de clients se présentant dans un magasin, le nombre d'objets défectueux d'une chaîne de production, le nombre d'atomes radioactifs désintégrés dans un laps de temps donné,...

3 Couples de variables aléatoires

3.1 Définitions et loi d'un couple de variables aléatoires

Définition 3.1.1 (Couple de variables aléatoires) *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soient deux variables aléatoires discrètes $X : \Omega \to E$ et $Y : \Omega \to F$. On définit le couple de variables aléatoires (X, Y) comme la variable aléatoire*

$$(X, Y) : \left\{ \begin{array}{ccc} \Omega & \to & E \times F \\ \omega & \mapsto & (X(\omega), Y(\omega)) \end{array} \right. .$$

(X, Y) est UNE variable aléatoire.

Proposition 3.1.1 (La variable aléatoire (X, Y))

1. *(X, Y) est encore une variable aléatoire discrète.*
2. *Toute variable aléatoire discrète définie sur $(\Omega, \mathcal{A}, P)$ et à valeur dans $E \times F$ peut s'écrire sous la forme (X, Y), où X et Y sont des variables aléatoires discrètes à valeurs respectivement dans E et F.*

Démonstration :

1. Si on note $Z = (X, Y)$, nous avons : $Z(\Omega) \subset X(\Omega) \times Y(\Omega)$. Or $X(\Omega)$ et $Y(\Omega)$ sont dénombrables, alors $X(\Omega) \times Y(\Omega)$ l'est aussi. Ainsi $Z(\Omega)$ est dénombrable en tant que sous-ensemble d'un ensemble dénombrable.
   Par ailleurs, soit $z = (x, y) \in Z(\Omega)$. Ainsi $(Z = z) = (X = x) \cap (Y = y)$, qui est encore dans $\mathcal{A}$, car $(X = x)$ et $(Y = y)$ le sont aussi, car X et Y sont des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, P)$.
2.

Définition 3.1.2 (Loi d'un couple de variables aléatoires et lois marginales)

1. *La loi conjointe de* X *et* Y *est la loi de* (X, Y)*, noté* $P_{(X,Y)}$.
   Ainsi la loi du couple est la donnée, pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$*, des probabilités* $P((X, Y) = (x, y)) = P(X = x, Y = y) = P((X = x) \cap (Y = y))$.
2. *Les lois marginales de* (X, Y) *sont les lois de* X *et de* Y. P_X *est appelée première loi marginale et* P_Y *est appelée deuxième loi marginale.*

Proposition 3.1.2 (Lois marginales) *Soit* (X, Y) *un couple de variables aléatoires. Nous disposons de la loi conjointe de* X *et* Y $(P(\{(X = x) \cap (Y = y)\})$ *pour* (x, y) *dans* $X(\omega) \times Y(\Omega))$*. Alors pour tout* x *de* $X(\Omega)$ *et* y *de* $Y(\Omega)$*, on a les lois marginales :*

$$P(X = x) = P_X(x) =$$

$$P(Y = y) = P_Y(y) =$$

Démonstration : Identique à la démonstration de sup.

Exemple 3.1.1 1. *Soient* X *et* Y *deux variables aléatoires à valeurs dans* $\mathbb{N}$ *telles que :*

$$\forall (k, n) \in \mathbb{N}^2,\ P((X = k) \cap (Y = n)) = \begin{cases} \binom{n}{k} \left(\frac{1}{2}\right)^n p(1-p)^n & si\ k \leq n \\ 0\ sinon \end{cases},\ avec\ p \in]0, 1[.$$

(a) *Soient* $k \in \mathbb{N}$ *et* $f : x \mapsto \sum_{n=k}^{+\infty} \binom{n}{k} x^n$*. Donner le rayon de convergence de* f *et calculer* $f(x)$ *lorsque* x *est dans l'intervalle ouvert de convergence.*

(b) *Vérifier que cela définit bien une loi de probabilité pour le couple* (X, Y).

(c) *Déterminer la loi de* Y.

(d) *Déterminer la loi de* X.

2. *Soit (X,Y) un couple de variables aléatoires à valeurs dans $\mathbb{N}^2$, dont la loi est donnée par :*
$$\forall (j,k)\in\mathbb{N}^2,\ P((X,Y)=(j,k))=\frac{(j+k)\left(\frac{1}{2}\right)^{j+k}}{ej!k!}.$$
   (a) *éterminer les lois marginales de X et Y.*
   (b) *Quelle est la loi de $Z=X+Y$?*

   (a) *On a :* $Y(\Omega)=\mathbb{N}$. *Soit* $k\in\mathbb{N}$.
   $P(Y=k)=\sum_{j=0}^{+\infty}P((X,Y)=(j,k))=\sum_{j=0}^{+\infty}\frac{(j+k)\left(\frac{1}{2}\right)^{j+k}}{ej!k!}$.
   Or : $\underbrace{\sum_{j=0}^{+\infty}\frac{j\left(\frac{1}{2}\right)^{j+k}}{ej!k!}}_{=0 \text{ si } j=0}=\frac{\left(\frac{1}{2}\right)^{k+1}}{ek!}\sum_{j=1}^{+\infty}\frac{\left(\frac{1}{2}\right)^{j-1}}{(j-1)!}=\frac{\left(\frac{1}{2}\right)^{k+1}}{ek!}\sum_{l=0}^{+\infty}\frac{\left(\frac{1}{2}\right)^{l}}{l!}=\frac{\left(\frac{1}{2}\right)^{k+1}}{ek!}e^{\frac{1}{2}}=\frac{\left(\frac{1}{2}\right)^{k+1}}{\sqrt{e}k!}$. *De plus :* $\sum_{j=0}^{+\infty}\frac{k\left(\frac{1}{2}\right)^{j+k}}{ej!k!}=\frac{k\left(\frac{1}{2}\right)^{k}}{ek!}\sum_{j=0}^{+\infty}\frac{\left(\frac{1}{2}\right)^{j}}{j!}=$
   $\frac{k\left(\frac{1}{2}\right)^{k}}{ek!}e^{\frac{1}{2}}=\frac{k\left(\frac{1}{2}\right)^{k}}{\sqrt{e}k!}$.
   Ainsi $P(Y=k)=\frac{\left(\frac{1}{2}\right)^{k+1}}{\sqrt{e}k!}+\frac{k\left(\frac{1}{2}\right)^{k}}{\sqrt{e}k!}$.
   Pour des raisons de symétrie, X a la même loi que Y.

   (b) *On a :* $Z(\Omega)=\mathbb{N}$. *Soit* $k\in\mathbb{N}$.
   Comme on a l'union disjointe $(Z=k)=\bigcup_{l=0}^{n}((X=l)\cap(Y=k-l))$, *alors :* $P(Z=k)=$
   $\sum_{l=0}^{k}P((X=l)\cap(Y=k-l))=\sum_{l=0}^{k}\frac{k\left(\frac{1}{2}\right)^{k}}{el!(k-l)!}=\frac{k\left(\frac{1}{2}\right)^{k}}{ek!}\sum_{l=0}^{k}\frac{k!}{l!(k-l)!}=\frac{k\left(\frac{1}{2}\right)^{k}}{ek!}\sum_{l=0}^{k}\binom{k}{l}=\frac{k\left(\frac{1}{2}\right)^{k}}{ek!}2^k=\frac{k}{ek!}$.

Remarque 3.1.1 *ATTENTION, les lois marginales ne permettent pas d'obtenir la loi conjointe. Par exemple, on se donne une urne contenant 3 boules blanches et 4 boules noires. On tire successivement deux boules.*
On envisage deux modes de tirages : avec et sans remise.
Dans ces deux cas on note X la variable aléatoire qui vaut 1 si la première boule tirée est blanche et 0 si elle est noire et on note Y la variable aléatoire qui vaut 1 si la deuxième boule tirée est blanche et 0 si elle est noire. Donner dans un tableau les lois conjointes et marginales correspondant aux deux situations. Constater ensuite que les lois marginales dans les deux cas sont les mêmes mais pas les lois conjointes.

Avec remise :

$X \backslash Y$	1	0	
1	$\frac{3}{7}\times\frac{3}{7}=\frac{9}{49}$	$\frac{3}{7}\times\frac{4}{7}=\frac{12}{49}$	$P(X=1)=\frac{21}{49}=\frac{3}{7}$
0	$\frac{4}{7}\times\frac{3}{7}=\frac{12}{49}$	$\frac{4}{7}\times\frac{4}{7}=\frac{16}{49}$	$P(X=0)=\frac{28}{49}=\frac{4}{7}$
	$P(Y=1)=\frac{21}{49}=\frac{3}{7}$	$P(Y=0)=\frac{28}{49}=\frac{4}{7}$	

Sans remise :

$X \backslash Y$	1	0	
1	$\frac{3}{7}\times\frac{2}{6}=\frac{1}{7}$	$\frac{3}{7}\times\frac{4}{6}=\frac{2}{7}$	$P(X=1)=\frac{3}{7}$
0	$\frac{4}{7}\times\frac{3}{6}=\frac{2}{7}$	$\frac{4}{7}\times\frac{3}{6}=\frac{2}{7}$	$P(X=0)=\frac{4}{7}$
	$P(Y=1)=\frac{3}{7}$	$P(Y=0)=\frac{4}{7}$	

Pour calculer par exemple $P((X=1)\cap(Y=0))$, dans le cas sans remise, on a deux méthodes :
- *Le nombre de tirages possibles est 7×6 (7 choix pour la première boule et 6 pour la deuxième). Le nombre de tirage avec une première boule blanche et une deuxième noire est 3×4.*
- $P((X=1)\cap(Y=0))=P(X=1)P(Y=0|X=1)=\frac{3}{7}\times\frac{4}{6}$.

Proposition 3.1.3 (n-uplet de variables aléatoires) *1. Soient $n \in \mathbb{N}^*$ et $X_1, ..., X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A})$, à valeurs dans des ensembles $E_1, ..., E_n$. L'application*

$$\left\{ \begin{array}{lll} \Omega & \to & E_1 \times ... \times E_n \\ \omega & \mapsto & (X_1(\omega), ..., X_n(\omega)) \end{array} \right. .$$

définit une variable aléatoire discrète, notée $(X_1, ..., X_n)$. C'est la loi conjointe de $X_1, ..., X_n$.

2. *Réciproquement toute variable aléatoire discrète à valeurs dans $E_1 \times ... \times E_n$ est de la forme $(X_1, ..., X_n)$, avec $X_1, ..., X_n$ des variables aléatoires discrète définies sur le même espace probabilisable. Elles sont appeléss loi marginales de $(X_1, ..., X_n)$.*

Démonstration : Généraliser la preuve de la proposition 3.1.1.

Exemple 3.1.2 *On considère un dé truqué à six faces. La probabilité d'obtenir la face k est notée p_k. On lance ce dé n fois et on note N_k le nombre d'apparition de la face k lors de ces n lancers pour k dans $[\![1, 6]\!]$.*

1. *Donner la loi de $(N_1, ..., N_6)$.*
2. *Donner le loi de (N_1, N_2).*

Définition 3.1.3 (Loi conditionnelle) *Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé, $A \in \mathcal{A}$ tel que $P(A) \neq 0$ et X une variable aléatoire discrète définie sur Ω. La loi conditionnelle de X sachant A est la probabilité sur $X(\Omega)$ définie par*

$$P_{X|A} : \left\{ \begin{array}{lll} X(\Omega) & \to & [0,1] \\ x & \mapsto & P_A(X = x) = \dfrac{P((X = x) \cap A)}{P(A)} \end{array} \right. .$$

C'est la loi de X si on munit Ω de la probabilité P_A.

Exemple 3.1.3 *Un joueur dispose de N dés équilibrés. Il lance une première fois ceux-ci et on note X_1 le nombre de six obtenus. Il met de côté les dés correspondants et relance les autres dés (s'il en reste). On note X_2 le nombre de six obtenus et on répète l'expérience définissant ainsi une suite de variables aléatoires $X_1, X_2, \cdots$.*
La variable $S_n = X_1 + \cdots X_n$ correspondont alors au nombre de six obtenu après n lancers.

1. *Vérifier que S_n suit une loi binomiale dont on précisera les paramètres.*
2. *Montrer qu'il est presque sûr qu'il existe un rang n pour lequel $S_n = N$.*
3. *Soit $T = \min\{n \geq 1,\ S_n = N\} \cup \{+\infty\}$. Déterminer la loi de T.*

3.2 Variables aléatoires indépendantes

Définition 3.2.1 (Variables aléatoires indépendantes) *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soient X et Y deux variables aléatoires discrètes. On dit que X et Y sont indépendantes si pour tout $(x, y) \in X(\Omega) \times Y(\Omega)$, les événements $(X = x)$ et $(Y = y)$ sont indépendants :*

$$\forall (x, y) \in X(\Omega) \times Y(\Omega), \quad P(X = x, Y = y) = P((X = x) \cap (Y = y)) = P(X = x)P(Y = y)$$

On note cela $X \perp\!\!\!\perp Y$.

Remarque 3.2.1 *IMPORTANTE : soient X et Y deux variables aléatoires indépendantes à valeurs dans $\mathbb{N}$. Alors pour tout n de $\mathbb{N}$, on a : $P(X + Y = n) =$*

Exemple 3.2.1 *1. Dans l'exemple 3.1.1, nous avons considéré X et Y deux variables aléatoires à valeurs dans $\mathbb{N}$ telles que :*

$$\forall (k,n) \in \mathbb{N}^2,\ P((X=k)\cap(Y=n)) = \begin{cases} \binom{n}{k}\left(\frac{1}{2}\right)^n p(1-p)^n & \text{si } k \leq n \\ 0 \text{ sinon} \end{cases},\quad \text{avec } p \in]0,1[.$$

Les variables aléatoires X et Y sont-elles indépendantes ?

2. *Soient X_1 et X_2 deux variables aléatoires indépendantes suivant une loi de Poisson de paramètre respectifs λ_1 et λ_2.*
   (a) *Déterminer la loi de $Z = X_1 + X_2$ (on donnera dans le paragraphe sur les fonctions génératrices une méthode plus rapide).*
   (b) *Déterminer la loi de X_1 sachant $(Z = n)$.*

3. *Soient $p \in]0,1[$, $\lambda \in \mathbb{R}_+^*$ et X, Y deux variables aléatoires définies sur $(\Omega, \mathcal{A}, P)$. On suppose que Y suit une loi de Poisson de paramètre λ et que $X(\Omega) = \mathbb{N}$, avec pour tout m de $\mathbb{N}$, la loi conditionnelle de X sachant $(Y = m)$ est une loi $\mathcal{B}(m,p)$.*
   (a) *Déterminer la loi de X.*

   (b) *Le nombre de clients venant dans un magasin pour acheter des poivrons suivent une loi de Poisson de paramètre λ. Chaque client achète au hasard un poivron rouge ou vert. On note R le nombre de client choisissant un poivron rouge et V le nombre de clients choisissant un poivron vert.*
   i) Déterminer les lois de R et V.
   ii) Sont-elles indépendantes ?

4. *Soient* X *et* Y *des variables aléatoires indépendantes suivant une loi* $\mathcal{G}(p)$*, avec* $p \in]0,1[$*. Quelle est la probabilité pour que* $M = \begin{pmatrix} X & 1 \\ 0 & Y \end{pmatrix}$ *soit diagonalisable ?*

5. *Soient* X *et* Y *i.i.d. suivant une loi* $\mathcal{G}(p)$*. On pose :* $q = 1 - p$ *Soient* $V = \min(X, Y)$ *et* $W = X - Y$*. Montrer que* V *et* W *sont indépendantes. On montrera dans l'exemple 3.2.2 que* V *suit une loi* $\mathcal{G}(1 - p^2)$*.*

6. *Soient* X *une variable aléatoire réelle discrète telle que* $X(\Omega)$ *soit infini et* X_1, X_2 *deux variables*

aléatoires indépendantes suivant la loi de X. On suppose que $X_1 + X_2 \sim 2X$.

(a) *Montrer qu'il existe $x \in \mathbb{R}$ tel que $P(X = x)$ soit maximal.*

(b) *Montrer que X est presque-sûrement constante.*

(a) On pose $X(\Omega) = \{x_n\}_{n\in\mathbb{N}}$.

On a $\sum_{n=0}^{+\infty} P(X = x_n) = 1$. Il existe un $y \in X(\Omega)$ tel que $P(X = y) > 0$ (sinon la somme serait nulle).

Comme la série $\sum P(X = x_n)$ converge alors $\lim_{n\to+\infty} P(X = x_n) = 0$. Il existe donc $N \in \mathbb{N}$ tel que : $\forall n \geq N,\ P(X = x_n) \leq P(X = y)/2 < P(X = y)$. Ainsi dans l'ensemble fini $\{P(X = x_n),\ 0 \leq n \leq N-1\}$, il y a un plus grand élément $P(X = x)$ et :
$\forall n \geq N,\ P(X = x_n) \leq P(X = y) \leq P(X = x)$, car $P(X = y)$ est dans l'ensemble $\{P(X = x_n),\ 0 \leq n \leq N-1\}$.

(b) On a alors $P(X = x) = P(2X = 2x) = P(X_1 + X_2 = 2x)$.

Ainsi, on a : $P(X = x) = \sum_{y\in X(\Omega)} P(\{X_2 = 2x - y\} \cap \{X_1 = y\})$ et par indépendance de X_1 et X_2, on a donc : $P(X = x) = \sum_{y\in X(\Omega)} P(X_2 = 2x - y)\, P(X_1 = y)$,

puis puisque les variables aléatoires suivent la même loi, $P(X = x) = \sum_{y\in X(\Omega)} P(X = 2x - y)\, P(X = y)$.

On a également : $P(X = x) = \sum_{y\in X(\Omega)} P(X = x)\, P(X = y)$, puisque les $\{X = y\}_{y\in X(\Omega)}$ forment un système complet d'événements, d'où :

$\sum_{y\in X(\Omega)} (P(X = x) - P(X = 2x - y))\, P(X = y) = 0$ par hypothèse, il s'agit d'une série à termes positifs de somme nulle, donc :
$\forall y \in X(\Omega),\ 2x - y \in X(\Omega)$ et $P(X = x) = P(X = 2x - y)$ dès que $P(X = y) > 0$.
Mais alors, puisque $P(X = 2x - y) > 0$, on a : $P(X = 2x - y) = P(X = 2x - (2x - y))$, d'où $P(X = x) = P(X = y)$. Ainsi, X suit une loi uniforme sur
$A = \{x \in X(\Omega),\ P(X = x) > 0\}$, qui est donc nécessairement fini, sinon
$\sum_{z\in A} P(X = z)$ diverge. Supposons que cet ensemble ne soit pas un singleton, soit $a = \min(A)$ et pour $b \in A$ avec $b \neq a$, on a : $2a - b = a + (a - b) < a$,
ce qui contredit la minimalité de a. Ainsi A est un singleton.

Proposition 3.2.1 (Evénements et variables aléatoires indépendantes) *Si X et Y sont indépendantes, alors pour tout $A \subset X(\Omega)$ et $B \subset Y(\Omega)$ on a : $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$.*

Démonstration : On a $(X \in A) \cap (Y \in B) = \bigcup_{(x,y)\in A\times B} (X = x) \cap (Y = y)$ et cette réunion est disjointe et elle est dénombrable, car $A \times B$ l'est en tant que sous-ensemble de $X(\Omega) \times Y(\Omega)$. Ainsi, on a : $P((X \in A) \cap (Y \in B)) = \sum_{(x,y)\in A\times B} P((X = x) \cap (Y = y)) = \sum_{(x,y)\in A\times B} P(X = x)P(Y = y)$, car X et Y sont indépendantes. Par ailleurs la famille $\sum_{(x,y)\in A\times B} P(X = x)P(Y = y)$ est sommable, car elle est à termes positifs et sa somme est finie (elle vaut $P((X \in A) \cap (Y \in B))$). Ainsi :
$\sum_{(x,y)\in A\times B} P(X = x)P(Y = y) = \sum_{x\in A} P(X = x) \sum_{y\in B} P(Y = y) = P(X \in A)P(Y \in B)$.

Définition 3.2.2 (n variables aléatoire indépendantes) *Soient $X_1, \ldots, X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, P)$. On dit que les X_i sont indépendantes si pour tout $(x_1, \ldots, x_n)$ dans $X_1(\Omega) \times \ldots \times X_n(\Omega)$, les événements $(X_i = x_i)_{1\leq i\leq n}$ sont indépendants.*

Remarque 3.2.2 *En pratique, il est rare qu'on utilise la définition précédente à cause du nombre de cas à considérer. On préférera utiliser la caractérisation suivante :*

Proposition 3.2.2 (Condition d'indépendance de n variables aléatoires) *Soient $X_1, \ldots, X_n$ des variables aléatoires discrètes définies sur $(\Omega, \mathcal{A}, P)$. Les X_i sont indépendantes si et seulement si*

$$\forall (x_1, \ldots, x_n) \in X_1(\Omega) \times \ldots \times X_n(\Omega), \qquad P(X_1 = x_1, \ldots, X_n = x_n) = \prod_{i=1}^{n} P(X_i = x_i)$$

Démonstration : Voir la démonstration de sup.

Proposition 3.2.3 (Évenements et n variables aléatoires indépendantes) *Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et soient $X_1, \ldots, X_n$ des variables aléatoires discrètes indépendantes. Pour tout $i \in \{1, \ldots, n\}$ soit A_i une partie de $X_i(\Omega)$. Alors les événements $(X_i \in A_i)_{1\leq i\leq n}$ sont indépendants :*

$$\forall J \subset \{1, \ldots, n\}, \qquad P\left(\bigcap_{j\in J} (X_j \in A_j)\right) = \prod_{j\in J} P(X_j \in A_j)$$

Démonstration : Soit $J \subset \{1,\ldots,n\}$. On a : $\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \bigcup_{(x_j)_{j\in J}\in\prod_{j\in J}A_j} \left(\bigcap_{j\in J}(X_j = x_j)\right)$ et cette réunion est disjointe. Comme les A_j sont dénombrable, alors le produit cartésien $\prod_{j\in J} A_j$ l'est aussi, car nous avons un nombre fini d'ensemble qui le constituent. Ainsi $P\left(\bigcap_{j\in J}(X_j \in A_j)\right) = \sum_{(x_j)_{j\in J}\in\prod_{j\in J}A_j} P\left(\bigcap_{j\in J}(X_j = x_j)\right) = \sum_{(x_j)_{j\in J}\in\prod_{j\in J}A_j} \prod_{j\in J} P\left(X_j = x_j\right)$. En utilisant le même argument de sommabilité que la proposition 3.2.1, mais cette fois-ci avec $|J|$ sommes, on a :
$\sum_{(x_j)_{j\in J}\in\prod_{j\in J}A_j} \prod_{j\in J} P\left(X_j = x_j\right) = \prod_{j\in J}\sum_{x_j\in A_j} P\left(X_j = x_j\right) = \prod_{j\in J} P(X_j \in A_j)$.

Exemple 3.2.2 *Soient $N \in \mathbb{N}^*$, $p \in]0,1[$ et $q = 1-p$. On considère $X_1,...,X_N$ des variables aléatoires définies sur $(\Omega,\mathcal{A},P)$, indépendantes et suivant une loi $\mathcal{G}(p)$. On pose $Y = \min(X_1,...,X_N)$. Quelle est la loi de Y ?*

Remarque 3.2.3 *Attention l'indépendance implique l'indépendance deux à deux, mais la réciproque est fausse. Par exemple soient X et Y deux variables aléatoires de même loi et indépendantes avec $P(X=1) = P(X=-1) = 1/2$. On pose $Z = XY$.*
On a : $P(Z=1) = P\left((X=1,Y=1)\cup(X=-1,Y=-1)\right) = P\left(X=-1,Y=-1\right) + P\left(X=1,Y=1\right) = P(X=-1)P(Y=-1) + P(X=1)P(Y=1) = 1/2$, car la réunion
$(X=1,Y=1)\cup(X=-1,Y=-1)$ est disjointe et ensuite X et Y sont indépendantes.
Z étant à valeurs dans $\{1,-1\}$, on a : $P(Z=-1) = 1/2$.
Soit $\epsilon_1,\epsilon_2 \in \{1,-1\}$. On a : $P(Z=\epsilon_1, X=\epsilon_2) = P(Y=\frac{\epsilon_1}{\epsilon_2} = \epsilon_1\epsilon_2, X=\epsilon_2) =$
$P(Y=\epsilon_1\epsilon_2)P(X=\epsilon_2) = 1/4 = P(Z=\epsilon_1)P(X=\epsilon_2)$. Ainsi X et Z sont indépendantes. De même Z et Y sont indépendantes. On a donc l'indépendance deux à deux. Cependant :
$P(Z=1,X=1,Y=-1) = 0 \neq 1/8 = P(Z=1)P(X=1)P(Y=-1)$.

Proposition 3.2.4 (Loi binomiale comme somme de Bernoulli indépendantes)
$X_1,\ldots,X_n$ des variables aléatoires mutuellement indépendantes et suivant la loi $\mathcal{B}(p)$.
Alors $X = X_1+\ldots+X_n$ suit une loi $\mathcal{B}(n,p)$.

Proposition 3.2.5 (Lemme des coalitions) *Soient $X_1,...,X_n$ des variables aléatoires indépendantes et $f:(X_1,...,X_k)(\Omega)\to F$ et $g:(X_{k+1},...,X_n)(\Omega)\to G$ deux applications. Alors les variables aléatoires $f(X_1,...,X_k)$ et $g(X_{k+1},...,X_n)$ sont indépendantes.*

Démonstration : Montrons d'abord le résultat pour $n=2$ et $k=1$.
Notons $Y_1 = f(X_1)$ et $Y_2 = g(X_2)$. Tout d'abord Y_1 et Y_2 sont deux variables aléatoires.
Fixons $y_1 \in Y_1(\Omega)$ et $y_2 \in Y_2(\Omega)$. Remarquons que :

$$[Y_1 = y_1] \;=\; \{\omega\in\Omega \,/\, f(X_1(\omega)) = y_1\} \;=\; \{\omega\in\Omega \,/\, X_1(\omega)\in f^{-1}(\{y_1\})\} \;=\; [X_1 \in f^{-1}(\{y_1\})]$$

et $[Y_2 = y_2] = [X_2 \in g^{-1}(\{y_2\})]$. Or X_1 et X_2 sont indépendantes, donc les événements $[X_1 \in f^{-1}(\{y_1\})]$ et $[X_2 \in g^{-1}(\{y_2\})]$ le sont aussi. Ainsi $[Y_1 = y_1]$ et $[Y_2 = y_2]$ sont indépendants pour tout $y_1 \in Y_1(\Omega)$ et $y_2 \in Y_2(\Omega)$ donc Y_1 et Y_2 sont indépendantes.

On montre ensuite que les deux vecteurs aléatoires $Y = (X_1, \ldots, X_k)$ et $Z = (X_{k+1}, \ldots, X_n)$ sont indépendants.

Soient $(x_1, \ldots, x_k) \in \prod_{i=1}^{k} X_i(\Omega)$ et $(x_{k+1}, \ldots, x_n) \in \prod_{i=k+1}^{n} X_i(\Omega)$,

$$(Y = (x_1, \ldots, x_k)) = \bigcap_{i=1}^{k} (X_i = x_i) \text{ et } (Z = (x_{k+1}, \ldots, x_n)) = \bigcap_{i=k+1}^{n} (X_i = x_i)\,.$$

On en déduit, d'après l'indépendance des variables X_i que

$$P\left((Y = (x_1, \ldots, x_k)) \cap (Z = (x_{k+1}, \ldots, x_n))\right) = P\left(\bigcap_{i=1}^{n} (X_i = x_i)\right) = \prod_{i=1}^{n} P\left(X_i = x_i\right)$$

De même

$$P\left(Y = (x_1, \ldots, x_k)\right) \times P\left(Z = (x_{k+1}, \ldots, x_n)\right) = \prod_{i=1}^{k} P\left(X_i = x_i\right) \times \prod_{i=k+1}^{n} P\left(X_i = x_i\right) = \prod_{i=1}^{n} P\left(X_i = x_i\right).$$

On a bien montré que Y et Z étaient indépendantes.

En rassemblant les deux points $f(Y) = f(X_1, ..., X_m)$ et $g(Z) = g(X_{m+1}, ..., X_n)$ sont indépendantes.

Corollaire 3.2.1 *Soient I un ensemble fini, $(I_1, ..., I_k)$, avec $k \in \mathbb{N}^*$ une partition de I et $(X_i)_{i \in I}$ une famille de variables aléatoires indépendantes. Pour $i \in [\![1, k]\!]$, soit une application $f_i : \prod_{j \in I_i} X_j(\Omega) \to F_i$.*

Alors la famille de variables aléatoires $(f_i((X_j)_{j \in I_i}))_{i \in [\![1,k]\!]}$ est indépendante.

Démonstration : Généraliser la preuve précédente.

Exemple 3.2.3 *On considère une variable aléatoire X à valeurs dans $\mathbb{Z}^d$, $(X_k)_{k \in \mathbb{N}^*}$ une suite de variables aléatoires mutuellement indépendantes suivant chacune la loi de X et définies sur un même espace probabilisé.*

La suite de variables aléatoires $(S_n)_{n \in \mathbb{N}}$ est définie par $S_0 = 0$ et : $\forall n \in \mathbb{N}^*,\ S_n = \sum_{k=1}^{n} X_k$.

On note R la variable aléatoire à valeurs dans $\mathbb{N}^* \cup \{+\infty\}$ définie par

$$R = \begin{cases} \min\{n \in \mathbb{N}^*,\ S_n = 0\} & si\ \ \{n \in \mathbb{N}^*,\ S_n = 0\} \neq \emptyset \\ +\infty & sinon \end{cases}$$

On admet que l'on a défini des variables aléatoires. Soit $n \in \mathbb{N}^*$.

Montrer que : $\mathbb{P}(S_n = 0) = \sum_{k=1}^{n} \mathbb{P}(R = k)\mathbb{P}(S_{n-k} = 0_d)$.

Définition 3.2.3 (Famille de variables aléatoires indépendantes) *Soit* I *un ensemble et* $(X_i)_{i\in I}$ *une famille de variables aléatoires définies sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$. *On dit que c'est une famille de variables aléatoires indépendantes si toute sous-famille finie est constituée de variables aléatoires indéndantes.*

Théorème 3.2.1 (Réalisation d'une suite de variables aléatoires) *Pour toute suite* (P_n) *de lois de probabilité discrètes, il existe un espace probabilisé* $(\Omega, \mathcal{A}, P)$ *et une suite* (X_n) *de variables aléatoires discrètes indépendantes sur* $(\Omega, \mathcal{A}, P)$ *telle que, pour tout* n *de* $\mathbb{N}$, *la loi* P_{X_n} *soit* P_n.

Démonstration : Admis.

Définition 3.2.4 (Suite de variables aléatoires i.i.d.) *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variables aléatoires discrètes définies sur un espace probabilisé* $(\Omega, \mathcal{A}, P)$. *On dit que que ces variables aléatoires sont indépendantes identiquement distribuées, noté i.i.d., si toutes ces variables aléatoires sont de même loi et la famille* $(X_n)_{n\in\mathbb{N}}$ *est indépendante.*

Exemple 3.2.4 1. *On parle de processus de Bernoulli de paramètre p lorsque l'on a une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires indépendantes suivant toutes la loi de Bernoulli $\mathcal{B}(p)$. Cela modélise par exemple un jeu de Pile ou Face infini. Par exemple si on note X la variable aléatoire donnant le premier n tel que $(X_n = 1)$, alors X suit une loi*

2. *Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes définies sur un espace probabilisé $(\Omega, \mathcal{A}, P)$ et suivant toutes une loi de Bernoulli de paramètre p, avec : $0 < p < 1$. Soit $k \in \mathbb{N}^*$ et on note Y_k le temps d'attente du k-ème succès. On pose $Z_1 = Y_1$ et pour $k \geq 2$: $Z_k = Y_k - Y_{k-1}$. Montrer que $(Z_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes identiquement distribuées. On posera $q = 1 - p$.*

4 Moyenne et dispersion

4.1 Espérance

4.1.1 Définitions

Définition 4.1.1 (Espérance) 1. *Soit X une variable aléatoire discrète, à valeur dans $\mathbb{R}_+ \cup \{+\infty\}$. L'espérance de X notée $\mathbb{E}(X)$ est la somme de la famille $(xP(X = x))_{x\in X(\Omega)}$ dans $[0, +\infty]$ soit :*

$$\mathbb{E}(X) = \sum_{x\in X(\Omega)} xP(X = x).$$

2. *Soit X une variable aléatoire discrète complexe. On dit que X est d'espérance fini si l'espérance de $|X|$ est finie, c'est-à -dire que la famille $(xP(X = x))_{x\in X(\Omega)}$ est sommable. Dans ce cas l'espérance de X notée $\mathbb{E}(X)$ est la somme dans $\mathbb{R}$:*

$$\mathbb{E}(X) = \sum_{x\in X(\Omega)} xP(X = x).$$

Dans ce cas, on notera cela : $X \in L^1$.

3. *On dit que X est centré si $\mathbb{E}(X) = 0$.*

Remarque 4.1.1 1. *Si $X(\Omega) = \{x_n,\, n \in \mathbb{N}\}$ et que l'espérance de X est finie, alors $\sum_{n=0}^{+\infty} |x_n| P(X = x_n)$ converge et : $\mathbb{E}(X) = \sum_{n=0}^{+\infty} x_n P(X = x_n)$*

2. *Dans le cas de convergence absolue, la valeur de la somme $\sum_{n=0}^{+\infty} x_n P(X = x_n)$ ne dépend pas de l'ordre d'énumération des valeurs x_n. C'est en effet le cas de n'importe quelle série absolument convergente.*
3. *Si $X(\Omega)$ est fini alors nécessairement $\mathbb{E}(X)$ existe.*
4. *On parle de moyenne pondérée car on peut aussi écrire dans le cas fini : $\mathbb{E}(X) = \dfrac{\sum_{i=1}^{n} x_i P(X = x_i)}{\sum_{i=1}^{n} P(X = x_i)}$.*
5. *$\mathbb{E}(X)$ ne dépend que de la loi de X. On pourra donc parler d'espérance d'une loi.*
6. *Dans le cas d'une variable aléatoire X positive, si $P(X = +\infty) > 0$, alors $(+\infty) \times P(X = +\infty) = +\infty$ et donc $\mathbb{E}(X) = +\infty$.*
   Si $P(X = +\infty) = 0$, on adoptera la convention $(+\infty) \times P(X = +\infty) = 0$.

Proposition 4.1.1 (Espérance nulle d'une variable aléatoire positive) *Soit X une variable aléatoire discrète positive telle que $\mathbb{E}(X) = 0$. Alors $X = 0$ presque-sûrement.*

Démonstration : On pose $X(\Omega) = \{x_n,\ n \in \mathbb{N}\}$ si $X(\Omega)$ est dénombrable, le cas fini se traitant de la même manière. Ainsi on a : $\sum_{n=0}^{+\infty} x_n P(X = x_n) = 0$. C'est une somme de nombres positifs, donc : $\forall n \in \mathbb{N},\ x_n P(X = x_n) = 0$. Pour $n \in \mathbb{N}$, si x_n est non nul, alors $P(X = x_n) = 0$. Comme $\sum_{n=0}^{+\infty} P(X = x_n) = 1$, alors il existe $k \in \mathbb{N}$ tel que $P(X = x_k) \neq 0$ et donc par contraposée $x_k = 0$. Ainsi comme pour $n \neq k$, on a $P(X = x_n) = 0$, car $x_n \neq x_k$, alors $1 = P(X = x_k) = P(X = 0)$, donc $X = 0$ presque-sûrement.

Proposition 4.1.2 (Autre formule de l'espérance) *Soit X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Dans $\mathbb{R}_+ \cup \{+\infty\}$, on a :*

$$\mathbb{E}(X) = \sum_{k=1}^{+\infty} P(X \geq k) = \sum_{l=0}^{+\infty} P(X > l).$$

Démonstration : Si $a = P(X = +\infty) > 0$, alors grâce au dernier point de la remarque 4.1.1, on a : $\mathbb{E}(X) = +\infty$ et de plus : $\forall x \in \mathbb{N}^*,\ P(X \geq k) \geq P(X = +\infty) = a$ et comme $a > 0$, alors $\sum_{k=1}^{+\infty} a = +\infty$ et donc $\sum_{k=1}^{+\infty} P(X \geq k) = +\infty$.

On suppose maintenant que $P(X = +\infty) = 0$ et on peut supposer que X soit à valeurs dans $\mathbb{N}$.

Montrons d'abord la relation : $\sum_{k=0}^{N} kP(X = k) = \sum_{k=1}^{N} P(X \geq k) - NP(X \geq N + 1)$, pour N dans $\mathbb{N}^*$.

On suppose que X est dans L^1. Montrons le résultat dans ce cas :

On suppose que X n'est pas dans L^1.

Remarque 4.1.2 *Ainsi pour X une variable aléatoire à valeurs dans $\mathbb{N} \cup \{+\infty\}$, X est dans L^1 si et seulement si $\sum_{k \geq 1} P(X \geq k)$ converge.*

Exemple 4.1.1 *1. Soit $p \in]0,1[$ et $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. de loi $\mathcal{G}(p)$. On pose $Z_n = \max(X_1, ..., X_n)$ et $\beta_n = E(Z_n)$. On note $q = 1 - p$.*
Montrer que : $-\frac{1}{\ln(q)} \sum_{j=1}^{n} \frac{1}{k} \leq \beta_n \leq -\frac{1}{\ln(q)} \sum_{j=1}^{n} \frac{1}{j} + 1$.

2. *Soit* (E_n) *une suite d'événements de* $(\Omega, \mathcal{A}, P)$ *et* $Z = \sum_{n=0}^{+\infty} 1_{E_n}$. *Montrer que si* $\sum P(E_n)$ *converge, alors* Z *est d'espérance finie et déterminer celle-ci.*

Proposition 4.1.3 (Espérance pour Ω fini) *Si Ω est un ensemble fini, alors*

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

4.1.2 Espérance de lois classiques

Proposition 4.1.4 (Espérance d'une loi quasi constante) *Si X est une variable aléatoire réelle constante ou quasi constante (il existe $x \in \mathbb{R}$ telle que $P(X = x) = 1$) alors*

$$E(X) = E(x) = x.$$

Proposition 4.1.5 (Espérance d'une loi uniforme) *Soit X une variable aléatoire ayant pour valeurs $x_1, ..., x_n$ et suivant une loi uniforme. Alors*

$$E(X) = \frac{\sum_{i=1}^{n} x_i}{n}.$$

Ainsi $E(X)$ est la moyenne naturelle de $x_1, ..., x_n$.

Exemple 4.1.2 *Si $X \hookrightarrow \mathcal{U}([\![a, b]\!])$, alors* $E(X) = \dfrac{1}{b-a+1} \displaystyle\sum_{k=a}^{b} k = \dfrac{1}{b-a+1} \times \dfrac{a+b}{2}(b-a+1) = \dfrac{a+b}{2}$.

Proposition 4.1.6 (Espérance d'une loi de Bernoulli) *Soit X une variable aléatoire de loi $\mathcal{B}(p)$. Alors*

$$E(X) = p.$$

Exemple 4.1.3

Soit $A \in \mathcal{A}$. On a $E(\mathbf{1}_A) =$

Proposition 4.1.7 (Espérance d'une loi binomiale) *Soit X une variable aléatoire de loi $\mathcal{B}(n, p)$. Alors*

$$E(X) = np.$$

Exemple 4.1.4 *On reprend l'exemple 3.1.3. Un joueur dispose de N dés équilibrés. Il lance une première fois ceux-ci et on note X_1 le nombre de six obtenus. Il met de côté les dés correspondants et relance les autres dés (s'il en reste). On note X_2 le nombre de six obtenus et on répète l'expérience définissant ainsi une suite de variables aléatoires $X_1, X_2, \cdots$.*
La variable $S_n = X_1 + \cdots X_n$ correspondont alors au nombre de six obtenu après n lancers. déterminer $E(X_n)$, pour $n \in \mathbb{N}^*$.

Proposition 4.1.8 (Espérance d'une loi géométrique) *Si X suit une loi géométrique de paramètre p, alors :*

$$\mathbb{E}(X) = \frac{1}{p}.$$

Démonstration : On pose $q = 1 - p$ et donc : $q \in]0, 1[$. On a : $\forall k \in \mathbb{N}^*,\ kP(X = k) = pkq^{k-1}$. Or la série entière $\sum_{k \geq 1} kx^{k-1}$ a pour rayon de convergence 1 et $\sum_{k=1}^{+\infty} kx^{k-1} = \frac{d}{dx}\underbrace{\left(\sum_{k=0}^{+\infty} x^k\right)}_{=\frac{1}{1-x}} = \frac{1}{(1-x)^2}$, pour $x \in]-1, 1[$. Ainsi $\mathtt{E}(X)$ existe et $\mathtt{E}(X) = p\sum_{k=1}^{+\infty} kq^{k-1} = p \times \frac{1}{(1-q)^2} = \frac{p}{p^2} = \frac{1}{p}$.

Exemple 4.1.5 *Une urne contient N boules. On effectue des tirages avec remise. Soit X la variable aléatoire désignant le nombre de tirages nécessaires pour tirer une deuxième fois la boule obtenue au premier tirage. Calculer $\mathbb{E}(X)$.*

Proposition 4.1.9 (Espérance d'une loi de Poisson) *Si X suit une loi de Poisson de paramètre λ, alors :*

$$\mathbb{E}(X) = \lambda.$$

Démonstration : Dans $\mathbb{R}_+ \cup \{+\infty\}$, on a : $\mathtt{E}(X) = \sum_{k=0}^{+\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{+\infty} k\frac{\lambda^k}{k!}e^{-\lambda} = \sum_{k=1}^{+\infty} \frac{\lambda^k}{(k-1)!}e^{-\lambda} = \sum_{l=0}^{+\infty} \frac{\lambda^{l+1}}{l!}e^{-\lambda} = \lambda e^{-\lambda}\sum_{l=0}^{+\infty} \frac{\lambda^l}{l!} = \lambda e^{-\lambda}e^{\lambda} = \lambda$ et $\mathtt{E}(X)$ existe.

4.1.3 Propriétés de l'espérance

Proposition 4.1.10 (Formule de transfert) *Soit $X : \Omega \to E$ une variable aléatoire discrète. Soit f une fonction de $X(\Omega)$ dans $\mathbb{R}$.*
La variable $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)P(X=x))_{x\in X(\Omega)}$ est sommable et dans ce cas :

$$\mathtt{E}(f(X)) =$$

L'espérance de $f(X)$ est entièrement déterminée par P_X et f.

Démonstration : Soit Y la variable aléatoire $Y = f(X)$. Pour $y \in Y(\Omega)$, on note $A_y = f^{-1}(\{y\})$, c'est-à-dire l'ensemble des antécédents de y par f. La famille $(A_y)_{y\in Y(\Omega)}$ est une partition au plus dénombrable de $X(\Omega)$ et :

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

d'où comme la réunion est disjointe : $P(Y=y) \;=\; \sum_{x\in A_y} P(X=x)$.

1. Supposons d'abord que la famille $\Big(f(x)P(X=x)\Big)_{x\in X(\Omega)}$ est sommable. Alors, on peut sommer par paquets des termes positifs :

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

   Par conséquent Y a une espérance finie. Les égalités précédentes sont alors encore vraies sans les valeurs absolues et ainsi :

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

2. Réciproquement, supposons que Y est d'espérance finie. Alors $\Big(|y|P([Y=y])\Big)_{y\in Y(\Omega)}$ est sommable, donc on peut sommer par paquets des termes positifs :

$$\sum_{y\in Y(\Omega)} |y|P([Y=y]) = \sum_{y\in Y(\Omega)}\sum_{x\in A_y} |f(x)|P(X=x) = \sum_{x\in X(\Omega)} |f(x)|P(X=x).$$

   Ainsi la famille $\Big(f(x)P(X=x)\Big)_{x\in X(\Omega)}$ est sommable, c'est-à-dire $f(X)$ admet une espérance.

Remarque 4.1.3 *En particulier si la variable aléatoires considérée est un couple de variable aléatoire discrètes (X,Y), la formule de transfert s'écrit :*

$$\mathtt{E}(f(X,Y)) =$$

Exemple 4.1.6 *1. Soit X une variable aléatoire suivant une loi $\mathcal{G}(p)$. Déterminer $\mathtt{E}(1/X)$.*

2. *Soit* (X,Y) *un couple de variables aléatoires à valeurs dans* $\mathbb{N}^2$*, dont la loi est donnée par :* $\forall (j,k)\in\mathbb{N}^2,\ P((X,Y)=(j,k))=\dfrac{(j+k)\left(\frac{1}{2}\right)^{j+k}}{ej!k!}$*. Montrer que* $E(2^{X+Y})$ *existe et la calculer.*

3. *On reprend l'exemple 1.2.1. Une urne contient* n *boules blanches numérotées de* 1 *à* n *et* m *boules noires numérotées de* 1 *à* m*. On tire une à une et sans remise toutes les boules de l'urne. Soit* X *la variable aléatoire indiquant le nombre de tirages effectués jusqu'au retrait de la première boule blanche. On a vu que :* $\forall l\in[\![1,m+1]\!],\ P(X=l)=\dfrac{\binom{m+n-l}{n-1}}{\binom{m+n}{n}}$*. Déterminer* $E(X)$*.*

Proposition 4.1.11 (Linéarité, positivité, croissante de l'espérance) *Soient* $X,\,Y\in L^1$*.*

1. *Pour tout* a,b *de* $\mathbb{K}$*,* $aX+bY$ *admet une espérance et :* $E(aX+bY)=aE(X)+bE(Y)$*.*
2. *Si* $X\geq 0$*, alors* $E(X)\geq 0$*.*
3. *Si* $X\leq Y$*, alors* $E(X)\leq E(Y)$*.*
4. $|E(X)|\leq E(|X|)$*.*

Démonstration :

1. On pose $X(\Omega)\times Y(\Omega)=\{(u_n,v_n),\ n\in\mathbb{N}\}$ la formule de transfert appliquée respectivement aux fonctions : $(x,y)\mapsto x$ et $(x,y)\mapsto y$ permettent d'avoir sous réserve d'existence : $\mathrm{E}(X)=\sum_{n=0}^{+\infty}u_nP(X=u_n,Y=v_n)$ et $\mathrm{E}(Y)=\sum_{n=0}^{+\infty}v_nP(X=u_n,Y=v_n)$. De plus ces deux séries convergent absolument car par exemple $\mathrm{E}(|X|)$ existe et donc en utilisant la formule de transfert $(x,y)\mapsto|x|$, on a : $\mathrm{E}(|X|)=\sum_{n=0}^{+\infty}|u_n|P(X=u_n,Y=v_n)$ qui converge. Ainsi la série $\sum(au_n+bv_n)P(X=u_n,Y=v_n)$ converge absolument et donc : $aX+bY$ admet une espérance.

Sa valeur est : $\mathtt{E}(aX+bY)=\sum_{n=0}^{+\infty}(au_n+bv_n)P(X=u_n,Y=v_n)=$

$a\sum_{n=0}^{+\infty}u_nP(X=u_n,Y=v_n)+b\sum_{n=0}^{+\infty}v_nP(X=u_n,Y=v_n)=a\mathtt{E}(X)+b\mathtt{E}(Y)$.

2. Si $X\geq 0$, alors $X(\Omega)=\{x_n,\ n\in\mathbb{N}\}$, avec tous les x_n positifs et donc :
   $\mathtt{E}(X)=\sum_{n=0}^{+\infty}x_nP(X=x_n)\geq 0$.
3. Si $X\leq Y$, alors $0\leq \mathtt{E}(Y-X)=\mathtt{E}(Y)-\mathtt{E}(X)$, par linéarité.
4. Si $X(\Omega)=\{x_n,\ n\in\mathbb{N}\}$, alors $\sum_{n=0}^{+\infty}|x_n|P(X=x_n)$ converge par hypothèse, donc $\mathtt{E}(|X|)$ existe par la formule de transfert. De plus $|\mathtt{E}(X)|=|\sum_{n=0}^{+\infty}x_nP(X=x_n)|\leq\sum_{n=0}^{+\infty}|x_n|P(X=x_n)=\mathtt{E}(|X|)$.

Exemple 4.1.7

1. *Soit $p\in]0,1[$ et Y à valeurs dans $\mathbb{N}$ telle que : $\forall n\in\mathbb{N},\ P(Y=n)=p(1-p)^n$. Déterminer l'espérance de Y sans calculs.*

2. *Une puce se déplace sur un axe gradué. Elle part de 0. À chaque étape, elle se déplace d'une unité à gauche ou à droite de façon équiprobable. Chaque déplacement est indépendant des autres. Soit $n\in\mathbb{N}^*$ et on pose X_n la variable aléatoire qui vaut 1 si lors de la n-ème étape la puce s'est déplacée à droite et -1 sinon. On pose $S_n=\sum_{k=1}^{n}X_k$. Déterminer $E(S_n)$.*

3. *Un ascenseur amène m personnes à n étages, chaque personne s'arrêtant à un étage de façon équiprobable. Quel est le nombre moyen d'arrêts ?*

4. *Soit $A_1,A_2,\ldots,A_n$ des événements d'un espace probabilisé $(\Omega,\mathcal{A},P)$. Montrer la formule du crible :*
   $$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{k=1}^{n}\left((-1)^{k-1}\sum_{1\leq i_1<i_2<\ldots<i_k\leq n}P\left(A_{i_1}\cap A_{i_2}\cap\ldots\cap A_{i_k}\right)\right).$$

5. *Chez un marchand de journaux, on peut acheter des pochettes contenant chacune une image. La collection complète comporte en tout N images distinctes. On note X_k le nombre d'achats ayant permi l'obtention de k images distinctes. En particulier, $X_1 = 1$ et X_N est le nombre d'achats nécessaires à l'obtention de la collection complète.*
   Déterminer l'espérance de X_N.

   Sachant $X_k = i$ (pour $i \geq k$), la variable aléatoire $X_{k+1} - X_k$ resprésente le temps d'attente pour avoir une nouvelle image alors que l'on a eu k images différentes. Cela peut se modéliser par une succession de succès-échec indépendants, où l'on a un succès lorsque l'on obtient une nouvelle image. La probabilité d'un succès est $\dfrac{N-k}{N}$, car il faut tomber sur une des $N-k$ images que l'on n'a pas encore eues. $X_{k+1} - X_k$ représente le temps d'attente pour avoir un succès.

   Ainsi $X_{k+1} - X_k$ suit une loi $\mathcal{G}\left(\dfrac{N-k}{N}\right)$, puis : $E_{(X_k=i)}(X_{k+1} - X_k) = \dfrac{N}{N-k}$ (espérance de $X_{k+1} - X_k$ avec la probabilité $P(\,.\,|X_k = i)$).

   En utilisant le fait que $(X_k = i)_{i\geq k}$ soit un système complet d'événements et le théorème de Fubini dans $\mathbb{R}_+$, on a :

   $$E(X_{k+1} - X_k) = \sum_{j=1}^{+\infty} jP(X_{k+1} - X_k = j) = \sum_{j=1}^{+\infty} j \sum_{i=k}^{+\infty} P(X_{k+1} - X_k = j|X_k = i)P(X_k = i) = \sum_{i=k}^{+\infty} P(X_k = i) \sum_{j=1}^{+\infty} jP(X_{k+1} - X_k = j|X_k = i) =$$
   $$\sum_{i=k}^{+\infty} P(X_k = i)E_{(X_k=i)}(X_{k+1} - X_k) = \sum_{i=k}^{+\infty} P(X_k = i)\frac{N}{N-k} = \frac{N}{N-k}.$$

   On a : $X_N = \sum_{k=1}^{N-1}(X_{k+1} - X_k) + X_1$, *donc :* $E(X_N) = \sum_{k=1}^{N-1} E(X_{k+1} - X_k) + E(X_1) = \sum_{k=1}^{N-1} \frac{N}{N-k} + 1 = N\sum_{l=1}^{N}\frac{1}{l}$.

6. *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires i.i.d. suivant toutes la loi d'une variable aléatoire X à valeurs dans $\mathbb{N}$. Pour $n \in \mathbb{N}^*$, on pose $R_n = card\{X_1, ..., X_n\}$. Soit $a \in \mathbb{N}$.*
   (a) *Montrer que : $E(R_n) \leq a + nP(X \geq a)$.*
   (b) *En déduire que $E(R_n) = o(n)$.*

Proposition 4.1.12 (Comparaison et espérance) *Soient X et Y deux variables aléatoires discrètes réelles définies sur $(\Omega, \mathcal{A}, P)$, telles que $|X| \leq Y$. Si Y est dans L^1, alors X aussi.*

Démonstration : Notons $X(\Omega) = \{x_i \,/\, i \in I\}$ et $Y(\Omega) = \{y_j \,/\, j \in J\}$, avec I et J dénombrables.

Grâce au système complet d'événements $\Big([Y=y_j]\Big)_{j\in J}$, on a pour tout $i \in I$:

$$|x_i|P(X=x_i) \;=\; |x_i|\sum_{j\in J} P([X=x_i]\cap[Y=y_j]) \;\leq\; \sum_{j\in J} y_j P([X=x_i]\cap[Y=y_j]). \tag{1}$$

Or la famille $\Big(y_jP([X=x_i]\cap[Y=y_j])\Big)_{(i,j)\in I\times J}$ est sommable de somme $E(Y)$. En effet, les égalités suivantes sont licites puisque Y est positive et admet une espérance finie :

$$E(Y) \;=\; \sum_{j\in J} y_j P(Y=y_j) \;=\; \sum_{j\in J} y_j \sum_{i\in I} P([X=x_i]\cap[Y=y_j]) \;=\; \sum_{i\in I}\sum_{j\in J} y_j P([X=x_i]\cap[Y=y_j]).$$

Ainsi l'inégalité (1) permet d'affirmer que X admet une espérance finie.

Exemple 4.1.8 *Si X est une variable aléatoire bornée, alors X est d'espérance finie, car*

En particulier, si Y est une variable aléatoire réelle et discrète , alors $\cos(Y)$ est d'éspérance finie.

Proposition 4.1.13 (Espérance d'un produit de variables aléatoires indépendantes)
$(\Omega,\mathcal{A},P)$ un espace probabilisé et soient X et Y deux variables aléatoires indépendantes discrètes dans L^1. Alors $E(XY)$ existe et

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

Démonstration : On note $X(\Omega)=\{x_k\,/\,k\in K\}$ et $Y(\Omega)=\{y_\ell\,/\,\ell\in L\}$, avec K et L des ensembles au plus dénombrables. On considère le couple (X,Y) de variables aléatoires à valeurs dans $\mathbb{R}^2$ et f l'application de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par $f(x,y)=|xy|$. Alors, d'après le théorème de transfert, la variable aléatoire $|XY|$ admet une espérance finie si et seulement si la famille
$\Big(|x_ky_\ell|P([(X,Y)=(k,\ell)])\Big)_{(k,\ell)\in K\times L}$ est sommable. Or X et Y sont d'espérances finies, d'où :

$$\left(\sum_{k\in K}|x_k|P(X=x_k)\right)\left(\sum_{\ell\in L}|y_\ell|P(X=y_\ell)\right) \;=\; \sum_{(k,\ell)\in K\times L}|x_ky_\ell|P(X=x_k)P(Y=y_\ell) \text{ qui est fini.}$$

De plus X et Y sont indépendantes, ce qui permet d'affirmer que :
$\forall (k,\ell)\in K\times L,\quad P(X=x_k)P(Y=y_\ell) \;=\; P\Big((X,Y)=(x_k,y_\ell)\Big)$. On en déduit ainsi que $|XY|$ admet une espérance finie.
Les égalités ci-dessous sans les valeurs absolues sont donc encore valables et prouvent que
$E(XY) \;=\; E(X)E(Y)$.

Corollaire 4.1.1 *Soient $X_1,...,X_n$ des variables aléatoiresindépendantes définies sur le même espace et dans L^1. Alors on a $E(X_1X_2...X_n)=\prod_{i=1}^{n}E(X_i)$.*

Démonstration : Par récurrence.

Exemple 4.1.9 1. *On reprend l'exemple 3.2.4. Soient $q\in\mathbb{N}$, avec $q\geq 3$ et $(\theta_n)_{n\geq 0}$ une suite de variables aléatoires indépendantes identiquement distribuées suivant une loi uniforme sur $\{0,1/q,2/q,...,(q-1)/q\}$. On pose $T_0=0$ et, pour $n\in\mathbb{N}$, $T_{n+1}=T_n+3+\sin(2\pi(T_n-\theta_n))$. Déterminer $E(T_n)$.*

2. *Soient $n \in \mathbb{N}^*$, A_{ij} pour $1 \leq i, j \leq n$, des variables aléatoires indépendantes, identiquement distribuées, telles que $P(A_{ij} = 1) = P(A_{ij} = -1) = 1/2$. On note A la matrice $(A_{ij})_{1 \leq i,j \leq n}$. Déterminer l'espérance de* $\det(A)$

Remarque 4.1.4 *ATTENTION : on peut avoir $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ sans que X et Y soient indépendantes.*
Par exemple, on considère le couple (X, Y) dont la loi conjointe est donnée par le tableau suivant

Y \ X	$X = 1$	$X = -1$	*Loi de Y*
$Y = -1$	$1/8$	$1/8$	$1/4$
$Y = 0$	$3/8$	$1/8$	$1/2$
$Y = 1$	$1/8$	$1/8$	$1/4$
Loi de X	$5/8$	$3/8$	1

La loi de XY est donc donnée par $P(XY = 1) = P(XY = -1) = 1/4$ et $P(XY = 0) = 1/2$.
On a donc $\mathbb{E}(XY) = 0$. Comme $\mathbb{E}(X) = 1/4$ et $\mathbb{E}(Y) = 0$, on a bien : $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Cependant, $P(X = 1, Y = 1) = 1/8 \neq P(X = 1)P(Y = 1) = 5/32$, donc X et Y ne sont pas indépendantes.

4.2 Variance

La variance et l'écart type sont des mesures de dispersion autour de la moyenne (espérance). L'écart type s'exprime avec la même unité que la variable aléatoire.

Définition 4.2.1 (Espace L^2) *Soit X une variable aléatoire réelle discrète. Si X^2 est d'espérance finie, on dit que X est dans L^2.*

Proposition 4.2.1 ($\mathbb{E}(X^2)$ existe implique que $\mathbb{E}(X)$ existe) *Soit X une variable aléatoire réelle discrète dans L^2. Alors X est dans L^1 ($\mathbb{E}(X)$ existe).*

Démonstration :

Définition 4.2.2 (Variance, écart type) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $X : \Omega \to \mathbb{R}$ *une variable aléatoire réelle discrète dans* L^2 $(\mathbb{E}(X^2)$ *existe). La variance de* X *est notée* $\mathbb{V}(X)$ *et est la moyenne des carrées des écarts des valeurs de* X *par rapport à leur moyenne.*

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

L'écart type est la racine carrée de la variance

$$\sigma(X) = \sqrt{\mathbb{V}(X)}.$$

On dit que X *est réduite si* $\mathbb{V}(X) = 1$.

Remarque 4.2.1

1. *La variance est bien définie, car* $(X - \mathbb{E}(X))^2 = X^2 - 2\mathbb{E}(X)X + (\mathbb{E}(X))^2$. *Comme* $\mathbb{E}(X)$ *est une constante, alors l'espérance de* $(X - \mathbb{E}(X))^2$ *est finie en tant que combinaison linéaire de variables aléatoires admettant une espérance.*
2. *L'écart type est bien défini, car d'après la croissance de l'espérance et le fait que :* $(X - \mathbb{E}(X))^2 \geqslant 0$, *on a :* $\mathbb{V}(X) \geqslant 0$.
3. *La variance mesure la proximité de* X *par rapport à la moyenne, c'est-à-dire la dispersion de la variable aléatoire* X. *On aurait pu s'intéresser à* $\mathbb{E}(|X - \mathbb{E}(X)|)$, *l'écart moyen, mais on a préféré* $\mathbb{E}\left((X - \mathbb{E}(X))^2\right)$, *car on a plus de propriétés calculatoires.*
4. *Il y a une différence entre la variance et l'écart type. On peut voir la variance comme une longueur au carrée, ainsi elle n'a pas la même « unité » que* $\mathbb{E}(X)$. *L'écart type, lui est homogène à une longueur, donc comparable à* $\mathbb{E}(X)$.
5. *Deux variables aléatoires* X *et* Y *ayant la même loi, ont la même variance.*

Proposition 4.2.2 (Variance nulle) *Soit* $X : \Omega \to \mathbb{R}$ *une variable aléatoire réelle discrète dans* L^2. *Si* $V(X) = 0$, *alors* X *est presque-sûrement constante. En particulier* $X = \mathbb{E}(X)$ *presque-sûrement.*

Démonstration : $(X - \mathrm{E}(X))^2$ est une variabale aléatoire discrète positive admettant une espérance nulle, donc $(X - \mathrm{E}(X))^2 = 0$ presque-sûrement, soit $X = \mathrm{E}(X)$ presque-sûrement.

Proposition 4.2.3 (Formule de König-Huygens) *Soit* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et soit* $X : \Omega \to \mathbb{R}$ *une variable aléatoire réelle discrète dans* L^2.

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

Démonstration : Identique à celle de sup.

Remarque 4.2.2 *Pour calculer une variance et en particulier* $\mathbb{E}(X^2)$, *parfois il peut être pratique de calculer* $\mathbb{E}(X(X \pm 1))$, *en effet, :* $\mathbb{E}(X^2) = \mathbb{E}(X(X \pm 1)) \mp \mathbb{E}(X)$.

Exemple 4.2.1

1. *Un sauteur en hauteur doit franchir successivement une série de barres de plus en plus hautes ; toutes les barres doivent être franchies et le parcours s'arrête au premier échec. La probabilité de passer la n-ème hauteur est* $1/n$. *Soit* X *le nombre de sauts réussis. Quelle est la variance de* X *?*

2. *Soient* $n \in \mathbb{N}^*$, A_{ij} *pour* $1 \leq i, j \leq n$, *des variables aléatoires indépendantes, identiquement distribuées, telles que* $P(A_{ij} = 1) = P(A_{ij} = -1) = 1/2$. *On note* A *la matrice* $(A_{ij})_{1 \leq i,j \leq n}$. *Dans l'exemple 4.1.9, nous avons vu que :* $\forall i, j \in [\![1, n]\!]$, $E(A_{i,j}) = 0$ *et* $E(\det(A)) = 0$. *Déterminer la variance de* $\det(A)$.

Proposition 4.2.4 (Opérations sur la variance) *Soient* $(\Omega, \mathcal{A}, P)$ *un espace probabilisé et* X *une variable aléatoire réelle dans* L^2. *Soit* a *et* b *deux réels*

$$\mathtt{V}(aX + b) = a^2 \mathtt{V}(X).$$

$$\sigma\,(aX + b) = |a|\sigma(X).$$

Démonstration : Voir la démonstration de sup.

Proposition 4.2.5 (Construction d'une variable aléatoire centrée réduite) *Soit* X *une variable aléatoire non presque-sûrement constante* ($\sigma(X) > 0$) *dans* L^2. *On pose* $X^* = \dfrac{X - \mathtt{E}(X)}{\sigma(X)}$. *Alors* X^* *est une variable aléatoire centrée et elle est aussi réduite.*

Démonstration : $\mathtt{E}(X^*) = \mathtt{E}\left(\dfrac{X - \mathtt{E}(X)}{\sigma(X)}\right) = \dfrac{1}{\sigma(X)}\mathtt{E}\left(X - \mathtt{E}(X)\right) = \dfrac{1}{\sigma(X)}\left(\mathtt{E}(X) - \mathtt{E}(\mathtt{E}(X))\right) = \dfrac{1}{\sigma(X)}(\mathtt{E}(X) - \mathtt{E}(X)) = 0$.

$$\mathtt{V}(X^*) = \mathtt{V}\left(\frac{1}{\sigma(X)}X - \underbrace{\frac{\mathtt{E}(X)}{\sigma(X)}}_{\text{constante}}\right) = \frac{1}{\sigma(X)^2}\mathtt{V}(X) = \frac{\mathtt{V}(X)}{\mathtt{V}(X)} = 1.$$

Proposition 4.2.6 *Soient* X *une variable aléatoire réelle,* $p \in]0, 1[$ *et* $q = 1 - p$.

1. *Si X suit une loi $\mathcal{U}(\{x_1, x_2, ..., x_n\})$, alors* $V(X) = \sum_{i=1}^{n} \frac{x_i^2}{n} - \left(\sum_{i=1}^{n} \frac{x_i}{n}\right)^2$.
2. *Si X suit une loi $\mathcal{B}(p)$, alors* $V(X) = p(1-p) = pq$.
3. *Si X suit une loi $\mathcal{B}(n,p)$, alors* $V(X) = np(1-p) = npq$.
4. *Si X suit une loi $\mathcal{G}(p)$, alors* $V(X) = \frac{1-p}{p^2} = \frac{q}{p^2}$.
5. *Si X suit une loi $\mathcal{P}(\lambda)$, alors* $V(X) = \lambda$.

Démonstration : Montrons l'avant dernière point (le dernier point a été fait dans le chapitre 9).
Si X suit une loi $\mathcal{G}(p)$, alors : $\mathtt{E}(X^2) = \sum_{k=1}^{+\infty} k^2 P(X=k) = \sum_{k=1}^{+\infty} k^2 q^{k-1} p =$
$\sum_{k=1}^{+\infty} k(k-1)q^{k-1}p + \sum_{k=1}^{+\infty} kq^{k-1}p = pq \sum_{k=2}^{+\infty} k(k-1)q^{k-2} + \mathtt{E}(X)$, car ces deux sommes convergent. De plus :
$\forall x \in]-1,1[,\ \sum_{k=2}^{+\infty} k(k-1)x^{k-2} = \frac{d^2}{dx^2}(\sum_{k=0}^{+\infty} x^k) = \frac{d^2}{dx^2}\left(\frac{1}{1-x}\right) = \frac{2}{(1-x)^3}$.
Ainsi : $\mathtt{E}(X^2) = \frac{2pq}{(1-p)^3} + \frac{1}{p} = \frac{2q}{p^2} + \frac{p}{p^2} = \frac{2-p}{p^2}$, puis $\mathtt{V}(X) = \mathtt{E}(X^2) - (\mathtt{E}(X))^2 = \frac{1-p}{p^2}$.

4.3 Covariance

Proposition 4.3.1 (Inégalité de Cauchy-Schwarz) *Soient X et Y deux variables aléatoires discrètes dans L^2. Alors XY admet une espérance ($XY \in L^1$) et :*

On a égalité si et seulement

Démonstration : L'existence de l'espérance de XY provient de $|XY| \leq \frac{1}{2}(X^2 + Y^2)$.
Comme pour établir l'inégalité de Cauchy-Schwarz dans des espaces préhilbertiens, considérons la fonction polynomiale μ définie sur $\mathbb{R}$ par $\mu : t \mapsto \mathtt{E}\Big((tX+Y)^2\Big) = t^2\mathtt{E}(X^2) + 2t\mathtt{E}(XY) + \mathtt{E}(Y^2)$ qui est une fonction positive, par positivité de l'espérance.
Lorsque $\mathtt{E}(X^2) = 0$, cette fonction de t est alors affine et toujours positive, donc elle est constante, donc son coefficient directeur est nul. Ainsi : $\mathtt{E}(XY) = 0$ et l'inégalité est vérifiée.
On se place alors dans le cas où $E(X^2) > 0$. La fonction μ est un trinôme du second degré ne prenant que des valeurs positives, on en déduit que le discriminant de μ est négatif ou nul, ce qui donne : $\Delta = 4\left(\mathtt{E}(XY)\right)^2 - 4\mathtt{E}(X^2)\mathtt{E}(Y^2) \leq 0$, d'où : $\left(\mathtt{E}(XY)\right)^2 \leq \mathtt{E}(X^2)\mathtt{E}(Y^2)$.
Étudions le cas d'égalité.
Si $\mathtt{E}(X^2) = 0$, donc X^2 est nulle presque-sûrement (car positive) et donc X aussi. Ainsi $1.X + 0.Y = 0$ presque sûrement.
Si $\mathtt{E}(X^2) \neq 0$, si on a égalité, alors $\Delta = 0$ donc μ admet une racine réelle double t_0. Dans ce cas : $\mathtt{E}\Big((t_0X+Y)^2\Big) = 0$ et donc on déduit que $t_0X + Y = 0$ presque sûrement.
Réciproquement, on suppose qu'il existe $(\alpha, \beta) \in \mathbb{R}^2 \setminus \{(0,0\}$ tel que $\alpha X + \beta Y = 0$ presque-sûrement. Il existe donc $t \in \mathbb{R}$ tel que $X = tY$ (si $\alpha \neq 0$) ou $Y = tX$ (si $\beta \neq 0$), presque sûrement. Dans les deux cas, on vérifie l'égalité.

Définition 4.3.1 (Covariance,) *Soient X, Y deux variables aléatoires discrètes admettant une variance. La covariance de X et Y, notée $cov(X,Y)$ est le réel*

$$cov(X,Y) = E\left[(X - E(X)) \cdot (Y - E(Y))\right].$$

Proposition 4.3.2 (Opérations sur la covariance)
1. *$cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y)$.*
2. *$cov(X,X) = V(X)$.*

3. *Si X et Y sont indépendantes, alors $cov(X,Y)=0$.*

Démonstration : Voir démonstration de sup.

Remarque 4.3.1 1. *cov est bilinéaire : si X,Y,Z sont trois variables aléatoires et λ un réel, alors $cov(\lambda X+Y,Z)=\lambda cov(X,Z)+cov(Y,Z)$ et $cov(X,\lambda Y+Z)=\lambda cov(X,Y)+cov(X,Z)$. Cela provient de la linéarité de $X\mapsto E(X\cdot Y)-E(X)\cdot E(Y)$ et $Y\mapsto E(X\cdot Y)-E(X)\cdot E(Y)$.*

2. *La covariance est symétrique : $cov(X,Y)=cov(Y,X)$.*
3. *Si $cov(X,Y)=0$, cela n'implique pas que X et Y sont indépendantes (voir remarque 4.1.4).*
4. *On a : $|cov(X,Y)|=|E[(X-E(X))\cdot(Y-E(Y))]|\leq\sqrt{E[(X-E(X))^2]}\cdot\sqrt{E[(Y-E(Y))^2]}=\sqrt{V(X)}.\sqrt{V(Y)}=\sigma(X).\sigma(Y)$, grâce à l'inégalité de Cauchy-Schwarz.*

Exemple 4.3.1 1. *Soient A et B deux événements d'un espace probabilisé. Montrer que : $|P(A\cap B)-P(A)P(B)|\leq\dfrac{1}{4}$.*

On a : $|P(A\cap B)-P(A)P(B)|=|E(1_{A\cap B})-E(1_A)E(1_B)|=|E(1_A1_B)-E(1_A)E(1_B)|=|cov(1_A,1_B)|=\sqrt{V(1_A)V(1_B)}$, grâce à la remarque précédente.

Si on note $p=P(A)$, alors $V(1_A)=p(1-p)\leq\frac{1}{4}$, car le maximum de la fonction $x\mapsto x(1-x)$ est atteint en $x=1/2$ (moyenne des racines du trinoôme du second degré $X(1-X)$). De même $V(1_B)\leq$

$frac14$ ce qui donne le résultat.

2. ***(Inégalité de Paley-Zigmund)** Soit X une variable aléatoire à valeurs positives admettant un moment d'ordre deux, avec $E(X^2)>0$. Pour α dans $]0,1[$, montrer que : $P(X\geq\alpha E(X))\geq(1-\alpha)^2\dfrac{(E(X))^2}{E(X^2)}$.*

Proposition 4.3.3 (Variance d'une somme) 1. *Soient X,Y deux variables aléatoires discrètes dans L^2. On a : $V(X+Y)=V(X)+V(Y)+2cov(X,Y)$.*

2. *Plus généralement si $X_1,...,X_n$ sont des variables aléatoires aléatoires discrètes dans L^2, on a :*

$$V\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n V(X_i)+\sum_{\substack{1\leq i,j\leq n\\ i\neq j}}cov(X_i,X_j)=\sum_{i=1}^n V(X_i)+2\sum_{1\leq i<j\leq n}cov(X_i,X_j).$$

Démonstration : Voir démonstration de sup.

Corollaire 4.3.1 (Variance et indépendance) 1. *Soient X,Y deux variables aléatoires discrètes indépendantes dans L^2. On a : $V(X+Y)=V(X)+V(Y)$.*

2. *$X_1,...,X_n$, sont des variables aléatoires deux à deux indépendantes dans L^2, alors : $V\left(\sum_{i=1}^n X_i\right)=\sum_{i=1}^n V(X_i)$.*

Démonstration : Voir démonstration de sup.

Exemple 4.3.2

Une urne contient $2n$ boules. Parmi ces boules n portent le numéro 0 et les n autres portent les numéros de 1 à n. On tire n boules de l'urne. Soit S la somme des numéros tirés. Déterminer $E(S)$ et $V(S)$.

4.4 Loi faible des grands nombres

Théorème 4.4.1 (Inégalité de Markov) *Soit X est une variable aléatoire discrète réelle positive admettant une espérance.*

$$\forall a \in \mathbb{R}_+^*,\ P(X \geq a) \leq$$

Démonstration : On a : $aP(X \geq a) = a\mathtt{E}(1_{X\geq a}) = \mathtt{E}(a1_{X\geq a})$.
Soit $\omega \in \Omega$.
Si $X(\omega) \geq a$, alors : $a1_{X\geq a}(\omega) = a \leq X(\omega)$.

Si $X(\omega) < a$, alors : $a1_{X\geq a}(\omega) = 0 \leq X(\omega)$, car X est positive.
Dans tous les cas, on a : $a1_{X\geq a} \leq X$, puis : $aP(X \geq a) = \mathtt{E}(a1_{X\geq a}) \leq \mathtt{E}(X)$.

Remarque 4.4.1 *Si X est de signe quelconque, on a donc :* $\forall a \in \mathbb{R}_+^*,\ P(|X| \geq a) \leq \dfrac{\mathtt{E}(|X|)}{a}$.

Exemple 4.4.1 *Soient $X, Y, (X_i)_{i\geq 1}$ des variables aléatoires réelles, discrètes et bornées, définies sur le même espace probabilisé. On suppose que les X_i sont indépendantes. Nous rappelons que nous avons démontré dans le cours sur les séries entières que :* $\forall \lambda \in \mathbb{R},\ \operatorname{ch} \lambda \leq \exp\left(\dfrac{\lambda^2}{2}\right)$.

1. *Soit $\lambda > 0$. Démontrer que, pour tout $a \in \mathbb{R}$, on a :* $\mathbb{P}(Y \geq a) \leq e^{-\lambda a}\mathbb{E}\big(e^{\lambda Y}\big)$.
2. *En déduire que :* $\forall a \in \mathbb{R}_+^*,\ \mathbb{P}(|Y| \geq a) \leq e^{-\lambda a}\mathbb{E}\big(e^{\lambda Y}\big) + e^{-\lambda a}\mathbb{E}\big(e^{-\lambda Y}\big)$.
3. *Soient $\lambda \geq 0$ et $x \in [-1, 1]$. Montrer que :* $\exp(\lambda x) \leq \operatorname{ch}(\lambda) + x \operatorname{sh}(\lambda) \leq \exp\left(\dfrac{\lambda^2}{2}\right) + x \operatorname{sh}(\lambda)$.
4. *Démontrer que si la variable aléatoire X prend ses valeurs dans $[-1, 1]$ et est centrée, alors on a pour tout $\lambda \geq 0$, on a :* $\mathbb{E}\left(e^{\lambda X}\right) \leq \exp\left(\dfrac{\lambda^2}{2}\right)$ *et* $\mathbb{E}\left(e^{-\lambda X}\right) \leq \exp\left(\dfrac{\lambda^2}{2}\right)$.
5. *Montrer que si les variables aléatoires indépendantes X_i prennent leurs valeurs dans $[-1, 1]$ et sont centrées, alors :* $\forall n \in \mathbb{N}^*, \forall a \in \mathbb{R}_+^*,\ \mathbb{P}\left(\left|n^{-1/2}\sum_{i=1}^{n} X_i\right| \geq a\right) \leq 2\exp\left(-\dfrac{a^2}{2}\right)$.

Théorème 4.4.2 (Inégalité de Bienaymé-Tchebychev) *Si X est une variable aléatoire discrète réelle dans L^2, et si t est un réel strictement positif,*

Démonstration : $P(|X-\mathtt{E}(X)|\geq t)=P((X-\mathtt{E}(X))^2\geq t^2)$, car on a : $|X-\mathtt{E}(X)|\geq t$ si et seulement si : $(X-\mathtt{E}(X))^2\geq t^2$, car t est positif. La variable aléatoire $(X-\mathtt{E}(X))^2$ étant positive et admetteant une espérance, on a grâce à l'inégalité de Markov, $P\left((X-\mathtt{E}(X))^2\geq t^2\right)\leq\dfrac{\mathtt{E}\left((X-\mathtt{E}(X))^2\right)}{t^2}=\dfrac{\mathtt{V}(X)}{t^2}$.

Exemple 4.4.2

1. *Soient $\theta, x\in\mathbb{R}_+^*$ distincts. Déterminer $\displaystyle\lim_{n\to+\infty}e^{-n\theta}\sum_{k\leq nx}\frac{(n\theta)^k}{k!}$. On distinguera les cas $x<\theta$ et $\theta<x$.*

2. *Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes qui suivent une loi de Bernoulli de paramètre p_i telle que $\displaystyle\lim_{n\to+\infty}\frac{1}{n}\sum_{i=1}^{n}p_i=p$.*

   Montrer que : $\forall\varepsilon>0,\ \displaystyle\lim_{n\to+\infty}P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-p\right|\geq\varepsilon\right)=0$.

   On note : $\forall n\in\mathbb{N}^*,\ e_n=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}p_i$ et $Y_n=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i$. Soit $\varepsilon\in\mathbb{R}_+^*$. Il existe $N\in\mathbb{N}$ tel que : $\forall n\geq N,\ |e_n-p|\leq\varepsilon/2$.

   Commes les variables aléatoires X_i sont dans L^2, alors Y_n l'est aussi et $E(Y_n)=e_n$, puis $V(Y_n)=\dfrac{1}{n^2}\displaystyle\sum_{i=1}^{n}p_i(1-p_i)\leq\dfrac{1}{n^2}\sum_{i=1}^{n}1=\dfrac{1}{n}$.

   Grâce à l'inégalité de Bienaymé-Tchebytchev, on a : $P(|Y_n-e_n|\geq\varepsilon/2)\leq\dfrac{4V(Y_n)}{\varepsilon^2}\leq\dfrac{4}{\varepsilon n}\underset{n\to+\infty}{\to}0$.

   Il existe donc $N_1\geq N$ tel que : $\forall n\geq N_1,\ P(|Y_n-e_n|\geq\varepsilon/2)\leq\varepsilon$.
   Soit $n\geq N_1$. Si $|Y_n-p|\geq\varepsilon$, alors $|Y_n-e_n|=|(Y_n-p)-(e_n-p)|\geq||Y_n-p|-|e_n-p||\geq|Y_n-p|-|e_n-p|\geq\varepsilon-\varepsilon/2=\varepsilon/2$.
   Ainsi on a : $(|Y_n-p|\geq\varepsilon)\Rightarrow(|Y_n-e_n|\geq\varepsilon/2)$, puis : $\forall n\geq N_1,\ P\left(|Y_n-p|\geq\varepsilon\right)\leq P\left(|Y_n-e_n|\geq\varepsilon/2\right)\leq\varepsilon$.
   Ainsi on a : $\displaystyle\lim_{n\to+\infty}P\left(|Y_n-p|\geq\varepsilon\right)=0$.

Théorème 4.4.3 (Loi faible des grands nombres) *Soit* $(Y_n)_{n\in\mathbb{N}^*}$ *une suite de variables aléatoires discrètes dans* L^2 *i.i.d.. On pose* $S_n = \sum_{i=1}^{n} Y_i$ *et on note* m *l'espérance commune de* $Y_1, ..., Y_n$ $(m = \mathtt{E}(Y_1))$ *et* σ *leur écart-type commun* $(\sigma = \sigma(Y_1))$*.*
On a : $\forall a \in \mathbb{R}_+^*, \forall n \in \mathbb{N}^*,\ P\left(\left|\frac{S_n}{n} - m\right| \geq a\right) \leq \frac{\sigma^2}{na^2}$.
En particulier : $\forall a \in \mathbb{R}_+^*,\ \lim_{n\to+\infty} P\left(\left|\frac{S_n}{n} - m\right| \geq a\right) = 0$.

Démonstration : S_n/n est dans L^2, en tant que combinaison linéaire de variables aléatoires dans L^2.
On a $\mathtt{E}(S_n/n) = \frac{1}{n}\sum_{i=1}^{n}\mathtt{E}(Y_i) = m$ et par indépendance des Y_n, on a :
$\mathtt{V}(S_n/n) = \frac{1}{n^2}\mathtt{V}(S_n) = \frac{1}{n^2}\sum_{i=1}^{n}\mathtt{V}(Y_i) = \frac{\sigma^2}{n}$. Grâce à l'inégalité de Bienaymé-Tchebychev, on a :
$P\left(\left|\frac{S_n}{n} - m\right| \geq a\right) = P\left(\left|\frac{S_n}{n} - \mathtt{E}\left(\frac{S_n}{n}\right)\right| \geq a\right) \leq \frac{\mathtt{V}(S_n/n)}{a^2} = \frac{\sigma^2}{na^2}$.

Remarque 4.4.2 *On pose* $\overline{Y_n} = \frac{\sum_{i=1}^{n} Y_i}{n}$*, qui est la moyenne empirique des résultats* $Y_1, ..., Y_n$*. Ce théorème montre que plus* n *est grand, plus la probabilité de voir* $\overline{Y_n}$ *s'approcher de son espérance* m *est proche de* 1*. La loi des grands nombres justifie l'interprétation fréquentiste des probabilités. En effet si on regarde la fréquence d'un événement lorsque l'on répète indéfiniment une expérience aléatoire, celle-ci tend vers la probabilité à priori de cet événement.*

Exemple 4.4.3

1. *Soient* $x \in [0,1]$*,* $\alpha \in \mathbb{R}_+^*$ *et* $n \in \mathbb{N}^*$*. On considère une variable aléatoire* S_n *suivant une loi* $\mathcal{B}(n,x)$*. Montrer que* $P\left(\left|\frac{S_n}{n} - x\right| > \alpha\right) \leq \frac{1}{4n\alpha^2}$*.*

2. *Une urne contient* 2 *boules rouges et* 3 *boules noires. On tire successivement avec remise des boules de cette urne. À partir de combien de tirages peut-on garantir à* 95% *que la proportion de boules rouges tirées est dans l'intervalle* $]0,35; 0,45[$*.*

   Soit $n \in \mathbb{N}^*$*. On considère la variable aléatoire* Y_i *valant* 1 *si la i-ème boule tirée est rouge et* 0 *sinon.* Y_i *suit une loi de Bernoulli de paramètre* $p = 2/5 = 0,4$*. Les variables* Y_i *suivent la même loi, sont mutuellement indépendantes et dans* L^2*. On a :* $\forall i \in \mathbb{N}^*,\ m = E(Y_i) = 0,4$ *et* $\sigma^2 = V(Y_i) = 0,4(1-0,4) = 0,24$*.*
   On pose $S_n = \sum_{i=1}^{n} Y_i$ *qui représente le nombre de boules rouges obtenues au cours de* n *tirages et* $T_n = S_n/n$ *qui est la proportion de boules rouges obtenues lors de ces tirages.*

   On cherche n *à partir duquel* $P(0,35 < T_n < 0,45) \geq 0,95$*.*
   Or $P(0,35 < T_n < 0,45) = P(0,4-0,05 < T_n < 0,4+0,05) = P\left(-0,05 < \frac{S_n}{n} - m < 0,05\right) = P\left(\left|\frac{S_n}{n} - m\right| < 0,05\right) = 1 - P\left(\left|\frac{S_n}{n} - m\right| \geq 0,05\right)$*.*
   On veut donc
   $P\left(\left|\frac{S_n}{n} - m\right| \geq 0,05\right) \leq 1 - 0.95 = 0,05$*. Or grâce à la proposition précédente, on a :* $P\left(\left|\frac{S_n}{n} - m\right| \geq 0,05\right) \leq \frac{0,24}{n(0,05)^2}$*. Il suffit donc d'avoir :* $\frac{0,24}{n(0,05)^2} \leq 0,05$*, soit* $\frac{0,24}{(0,05)^3} \leq n$*, soit* $1920 \leq n$*.*

3. ***(Loi forte des grands nombres dans le cas*** L^4***)*** *Soit* (X_n) *une suite de variables aléatoires indépendantes identiquement distribuées réelles telles que :* $\forall n \in \mathbb{N},\ X_n^4 \in L^1$*, on dit que* X_n *est dans* L^4*. On pose* $m = E(X_1)$*,* $V_2 = E\left((X_1 - E(X_1))^2\right)$ *et*
   $V_4 = E\left((X_1 - E(X_1))^4\right)$*.*

   Soit $\varepsilon > 0$*. Pour* $n > 0$*, on note* A_n^ε *l'événement* $\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - m\right| \geq \varepsilon\right)$*.*

(a) *Donner une majoration de* $P(A_n^\varepsilon)$ *à l'aide de* n, V_2 *et* V_4.

(b) *En déduire que presque-sûrement* $\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ *converge vers* m.

Remarque 4.4.3 Nous pouvons même montrer la convergence presque-sûre pour des variables aléatoires dans les cas où les X_i sont dans L^2. On reprend les mêmes notations que dans l'exemple précédent.

Pour montrer que S_n/n converge presque-sûrement vers m, il faut montrer que $S'_n = \frac{1}{n}\sum_{i=1}^{n} Y_i$ converge presque-sûrement vers 0.

Grâce à la proposition précédente, on a : $\forall \varepsilon \in \mathbb{R}_+^*, \forall n \in \mathbb{N}^*,\ P\left(\left|\frac{S'_n}{n}\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon}$, puis : $\forall \varepsilon \in \mathbb{R}_+^*, \forall n \in \mathbb{N}^*,\ P\left(\left|\frac{S'_{n^2}}{n^2} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n^2\varepsilon}$, donc $\sum P\left(\left|\frac{S'_{n^2}}{n^2} - m\right| \geq \varepsilon\right)$ converge et comme dans la conclusion de l'exemple précédent, la suite $\left(\frac{S'_{n^2}}{n^2}\right)_{n\geq 1}$ converge presque-sûrement vers 0. On appelle A le sous-ensemble de Ω pour lequel cela est réalisé.
Soit $n \in \mathbb{N}^*$. On pose $k_n = \lfloor \sqrt{n} \rfloor$.
On a $k_n \leq \sqrt{n} < k+1$, donc : $k_n^2 \leq n \leq k_n^2 + 2k_n + 1$.
On pose $T_n = \frac{1}{n}\sum_{i=k_n^2+1}^{n} Y_i$.
Soit $\varepsilon \in \mathbb{R}_+^*$. On a : $V(T_n) = \frac{n-k_n^2}{n^2}V_2 \leq \frac{2k_n+1}{n^2}V_2 \leq \frac{2\sqrt{n}+1}{n^2}V_2 = O\left(\frac{1}{n^{3/2}}\right)$.
Or grâce à l'inégalité de Bienaymé-Tchebitchev, on a : $P(|T_n| \geq \varepsilon) \leq \frac{V_n}{\varepsilon^2} = O\left(\frac{1}{n^{3/2}}\right)$, donc comme dans l'exemple précédent, (T_n) converge presque-sûrement vers 0. On appelle B le sous-ensemble de Ω pour lequel cela est réalisé.
Comme une intersection au plus dénombrables d'événements presque-sûrs est presque-sûr, alors $A \cap B$ est de probabilité 1 et sur cet ensemble, on a : $S'_n = \frac{k_n^2}{n}\frac{S'_{k_n^2}}{k_n^2} + T_n$. Comme $1 \leq \frac{k_n^2}{n} \leq 1 + 2\frac{1}{\sqrt{n}} + \frac{1}{n}$, alors $\lim_{n\to+\infty} \frac{k_n^2}{n} = 1$, puis $(S'_n)_{n\geq 1}$ converge presque-sûrement vers 0, ce qui donne le résultat.

4.5 Fonctions ou séries génératrices

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

Définition 4.5.1 (Fonctions génératrices) *On appelle fonction ou série génératrice de la variable aléatoire X la fonction*

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

Remarque 4.5.1 *1. La formule avec l'espérance s'obtient grâce à la formule de transfert.*
2. Si X a un nombre fini de valeurs, alors G_X est une fonction polynôme.

Proposition 4.5.1 (La série entière G_X) *1. $G_X(1) =$*
2. La série génératrice G_X d'une variable aléatoire entière est une série entière de rayon de convergence au moins égal à 1 et G_X est au moins définie sur $[-1,1]$.
3. Soit on a : $\forall k \in \mathbb{N},\ P(X=k) =$
4. Deux variables aléatoires X et Y à valeurs dans $\mathbb{N}$ ont la même loi si et seulement si

Démonstration :

1. $G_X(1) = \sum_{n=0}^{+\infty} P(X=n) = 1$, car X est à valeurs dans $\mathbb{N}$.
2. Soit $n \in \mathbb{N}$ et $t \in [-1,1]$. On a : $|t^n P(X=n)| \leq P(X=n)$ et $\sum_{n\geq 0} P(X=n)$ converge (voir le premier point). Ainsi $\sum_{n\geq 0} t^n P(X=n)$ converge absolument pour t dans $[-1,1]$, donc G_X est définie sur $[-1,1]$ et a un rayon de convergence R_X au moins égal à 1.
3. Soit $k \in \mathbb{N}$. G_X est ainsi $\mathcal{C}^\infty$ sur $]-1,1[\subset]-R_X, R_X[$ et on peut la dériver terme à terme : $\forall t \in]-1,1[,\ G_X^{(k)}(t) = \sum_{n=k}^{+\infty} n(n-1)...(n-k+1)t^{n-k}P(X=n) = \sum_{n=k}^{+\infty} \frac{n!}{(n-k)!}t^{n-k}P(X=n)$.
   En particulier $G_X^{(k)}(0) = k!P(X=k)$, puis $P(X=k) = \frac{G_X^{(k)}(0)}{k!}$.
4. Par unicité du DSE, $G_X = G_Y$ si et seulement si : $\forall k \in \mathbb{N},\ P(X=k) = P(Y=k)$ si et seulement si X et Y ont la même loi.

Remarque 4.5.2 *On suppose que le rayon de convergence de G_X est R_X. Pour $k \in \mathbb{N}^*$, on a :*
$G_{kX}(t) =$

Proposition 4.5.2 *1. La série entière* $\sum_{n=0}^{+\infty} P(X=n)z^n$ *converge normalement sur* $\overline{D}(0,1)$.

2. G_X *est continue sur* $[-1,1]$.

Démonstration :

1. Soit $n \in \mathbb{N}$ et on pose $f_n : z \mapsto z^n P(X=n)$. On a : $\forall z \in \overline{D}(0,1),\ |f_n(z)| \leq P(X=n)$, donc $\|f_n\|_\infty \leq P(X=n)$, donc $\sum_{n\geq 0} \|f_n\|_\infty$ converge.
2. • Pour tout n de $\mathbb{N}$, f_n est continue sur $[-1,1]$.
   • $\sum_{n\geq 0} f_n$ converge normalement sur $[-1,1]$.

   Ainsi $\sum_{n=0}^{+\infty} f_n$ est continue sur $[-1,1]$.

Exemple 4.5.1 *1. Une pièce a une probabilité* p *de tomber sur « pile » et* $q = 1-p$ *de tomber sur « face ». On effectue successivement une infinité de lancers. On note* S_n *le numéro du tirage pour lequel on obtient « pile » pour la* $n^{\text{ème}}$ *fois (on convient que* $S_n = +\infty$ *si cet événement n'est pas réalisé). Déterminer* G_{S_n}. *On rappelle que l'on a vu dans l'exemple 1.2.1 que* S_n *est à valeurs dans* $\mathbb{N} \cap [n, +\infty[\cup \{+\infty\}$. *Nous avons vu dans le chapitre 10 que* $P(S_n = +\infty) = 0$, *car le motif avec n Piles consécutifs apparaît au moins une fois et par ailleurs :*
$\forall k \geq n,\ P(S_n = k) = \binom{k-1}{n-1} p^n q^{k-n}.$

2. On considère dans toute cette partie que la pièce est équilibrée, c'est-à-dire que $p = \frac{1}{2}$.
On note N_n *le nombre de séries lors des n premiers lancers :*

- *La première série est donc de longueur* $k < n$ *si les* k *premiers lancers ont amené le même côté de la pièce et le* $(k+1)$*-ième l'autre côté et de longueur* n *si les* n *premiers lancers ont amené le même côté de la pièce ;*
- *La dernière série se termine nécessairement au n-ième lancer ».*

Déterminer la loi de N_n *à l'aide de sa séries génératrice* G_n *en exprimant d'abord* G_n *en fonction de* G_{n-1}, *pour* $n \geq 2$. *On notera* P_n *(resp.* F_n*) l'événement « avoir Pile (resp. Face) au n-ème lancer.*

Proposition 4.5.3 (Fonctions génératrices de lois usuelles)

1. *Si* $X \hookrightarrow \mathcal{B}(p)$ *alors* $G_X(t) = (1-p) + tp$.
2. *Si* $X \hookrightarrow \mathcal{B}(n,p)$ *alors* $G_X(t) = ((1-p)+tp)^n$.
3. *Si* $X \hookrightarrow \mathcal{G}(p)$ *alors* $G_X(t) = \dfrac{pt}{1-(1-p)t} = \dfrac{pt}{1-qt}$ *et le rayon de convergence est* $\dfrac{1}{q}$, *avec* $q = 1-p$.
4. *Si* $X \hookrightarrow \mathcal{P}(\lambda)$ *alors* $G_X(t) = e^{-\lambda+\lambda t} = e^{\lambda(t-1)}$ *et le rayon de convergence est* $+\infty$.

Démonstration :

1. $\forall t \in \mathbb{R},\ G_X(t) = t^0 P(X=0) + t^1 P(X=1) = 1-p+tp$.
2. $\forall t \in \mathbb{R},\ G_X(t) = \displaystyle\sum_{k=0}^{n} t^k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^{n} \binom{n}{k} (tp)^k (1-p)^{n-k} = (tp+1-p)^n$.
3. $G_X(t) = \displaystyle\sum_{k=1}^{+\infty} t^k q^{k-1} p = pt \sum_{k=1}^{+\infty} (tq)^{k-1} = pt \sum_{l=0}^{+\infty} (tq)^l = pt \times \frac{1}{1-qt} = \frac{pt}{1-(1-p)t} = \frac{pt}{1-qt}$, avec $|qt| < 1$.
4. $\forall t \in \mathbb{R},\ G_X(t) = \displaystyle\sum_{k=0}^{+\infty} t^k \frac{\lambda^k}{k!} e^{-\lambda} = e^{-\lambda} \sum_{k=0}^{+\infty} \frac{(t\lambda)^k}{k!} = e^{-\lambda} e^{\lambda t} = e^{\lambda(t-1)}$.

Proposition 4.5.4 (Fonctions génératrices et espérances) *X est d'espérance finie si et seulement si G_X est dérivable en* 1.
Si tel est le cas, alors $\mathbb{E}(X) = G'_X(1)$.

Démonstration : Si $E(X)$ existe.

Réciproquement si G_X est dérivable en 1 :

Remarque 4.5.3 *On retrouve ainsi aisément les espérances des variables aléatoires usuelles. Si* $X \sim \mathcal{G}(p)$. *La fonction* G_X *est dérivable sur* $[0,1]$, *car* $1 < \dfrac{1}{q}$, *donc* $\mathbb{E}(X)$ *existe et :* $\forall t \in [0,1],\ G'_X(t) = \dfrac{p(1-(1-p)t)+p(1-p)t}{(1-(1-p)t)^2} = \dfrac{p}{(1-(1-p)t)^2}$.
Ainsi : $\mathbb{E}(X) = G'_X(1) = \dfrac{p}{p^2} = \dfrac{1}{p}$.
Si $X \sim \mathcal{P}(\lambda)$. *La fonction* G_X *est dérivable sur* $\mathbb{R}$, *donc* $\mathbb{E}(X)$ *existe et :*
$\forall t \in \mathbb{R},\ G'_X(t) = \lambda e^{-\lambda+\lambda t}$, *puis :* $\mathbb{E}(X) = G'_X(1) = \lambda$.

Exemple 4.5.2 *On reprend l'exemple 4.5.1 Une pièce a une probabilité* $p \in]0,1[$ *de tomber sur « pile » et* $q = 1-p$ *de tomber sur « face ». On effectue successivement une infinité de lancers. On note* S_n *le numéro du tirage pour lequel on obtient « pile » pour la* $n^{\text{ème}}$ *fois (on convient que* $S_n = +\infty$ *si cet événement n'est pas réalisé). Quelle est l'espérance de* S_n *?*

Proposition 4.5.5 (Fonctions génératrices et variances) *X est dans* L^2 *si et seulement si* G_X *est deux fois dérivable en 1.*
Si tel est le cas, alors $\mathbb{V}(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.

Démonstration : Comme précédemment, si $\mathrm{E}(X^2)$ existe alors $\mathrm{E}(X)$ aussi, les séries $\sum n^2 P(X=n)$ et $\sum nP(X=n)$ convergent donc et $\sum n(n-1)P(X=n)$ converge aussi et le théorème de dérivation des séries de fonction s'applique deux fois, prouvant que
$G''_X(1) = \displaystyle\sum_{n=2}^{+\infty} n(n-1)P(X=n) = \mathrm{E}(X(X-1)) = \mathrm{E}(X^2) - \mathrm{E}(X)$ (par la formule de transfert) donc $\mathrm{E}(X^2) = G''_X(1) + G'_X(1)$.
Et finalement $V(X) = G''_X(1) + G'_X(1) - G'_X(1)^2$.
Inversement, si G_X est deux fois dérivable en 1, alors comme précédemment on montre que : $\sum n(n-1)P(X=n)$ converge donc $\sum n^2 P(X=n)$ aussi et X^2 est donc d'espérance finie.

Remarque 4.5.4

1. *On retrouve ainsi aisément les variances des variables aléatoires usuelles.*
2. *La formule de la variance est plus à savoir retrouver qu'à apprendre.*
   En effet $G''_X(1) =$

   puis $\mathbb{E}(X^2) =$

Remarque 4.5.5 *1. On peut etrouver les variances de lois usuelles.*

Si $X \sim \mathcal{G}(p)$*. On a :* $\forall t \in [0,1],\ G''_X(t) = \dfrac{2p(1-p)}{(1-(1-p)t)^3}$*, puis* $G''_X(1) = \dfrac{2p(1-p)}{p^3} = \dfrac{2(1-p)}{p^2}$.

Ainsi $V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2 = \dfrac{2(1-p)}{p^2} + \dfrac{p}{p^2} - \dfrac{1}{p^2} = \dfrac{1-p}{p^2}$.

Si $X \sim \mathcal{P}(\lambda)$*, on a :* $\forall t \in \mathbb{R},\ G''_X(t) = \lambda^2 e^{-\lambda+\lambda t}$*, donc :* $V(X) = \lambda^2 + \lambda - \lambda^2 = \lambda$.

2. *Par récurrence sur* $n \geq 1$*, on peut démontrer que si* X *est une variable aléatoire à valeurs dans* $\mathbb{N}$*, alors* G_X *est* r *fois dérivable en 1, si et seulement si* X^r *admet une espérance et dans ce cas :* $G_X^{(r)}(1) = E(X(X-1)...(X-r+1))$*. Ceci nous permet de calculer* $E(X^r)$*, à l'aide de* $E(X)$,...,$E(X^{r-1})$.

Exemple 4.5.3 *On reprend l'exemple 4.5.1 Une pièce a une probabilité* $p \in]0,1[$ *de tomber sur « pile » et* $q = 1-p$ *de tomber sur « face ». On effectue successivement une infinité de lancers. On note* S_n *le numéro du tirage pour lequel on obtient « pile » pour la* $n^{\text{ème}}$ *fois (on convient que* $S_n = +\infty$ *si cet événement n'est pas réalisé). Quelle est l'espérance de* S_n *?*

Proposition 4.5.6 *1. Soient deux variables aléatoires* X *et* Y *à valeurs dans* $\mathbb{N}$ *indépendantes. Alors on a :* $G_X G_Y = G_{X+Y}$.

2. *Plus généralement soient* $X_1, ..., X_n$ *des variables aléatoires à valeurs dans* $\mathbb{N}$ *mutuellement indépendantes. Alors on a :* $G_{X_1+...+X_n} = G_{X_1} \times ... \times G_{X_n}$.

Démonstration :

Le deuxième point se prouve par récurrence.

Remarque 4.5.6 *La réciproque est fausse.*

Exemple 4.5.4 *1. Soient* X *et* Y *deux variables indépendantes suivant respectivement une loi* $\mathcal{P}(\lambda)$ *et* $\mathcal{P}(\mu)$*. Alors* $X+Y$ *suit une loi*

2. *Soient* X *et* Y *deux variables indépendantes suivant respectivement une loi* $\mathcal{B}(n,p)$ *et* $\mathcal{B}(m,p)$*. Alors* $X+Y$ *suit une loi*

$\forall t \in [-1,1],\ G_{X+Y}(t) \underbrace{=}_{X \text{ et } Y \text{ est indépendante}} G_X(t)G_Y(t) = (1-p+tp)^n(1-p+tp)^m = (1-p+tp)^{n+m}$*. Ceci est la fonction génératrice d'une variable aléatoire suivant une loi* $\mathcal{B}(m+n,p)$ *et une fonction génératrice caractérise la loi.*

Réciproquement, soit $m, n \in \mathbb{N}^*$ *et* X, Y *deux variables aléatoires indépendantes avec* $X \sim \mathcal{B}(n,p)$ *et* $Y \sim \mathcal{B}(m,q)$*. Montrer que si* $X+Y$ *suit une loi binomiale, alors* $p = q$.

On a : $\forall t \in [-1,1],\ G_{X+Y}(t) \underbrace{=}_{X \text{ et } Y \text{ est indépendante}} G_X(t)G_Y(t) = (1-p+tp)^n(1-q+tq)^m$*. Or la série génératrice d'une loi binomiale n'a qu'une racine, donc* $p = q$.

3. *Peut-on piper deux dés de sorte que, les lancers étant supposés indépendants, leur somme suive la loi uniforme sur* $\{2, ..., 12\}$ *(on pourra utiliser des séries génératrices) ?*

4. ***(Formule de Wald)***

Soient $(X_k)_{k\in\mathbb{N}^*}$ *une suite variables aléatoires discrètes indépendantes de même loi sur* $(\Omega, \mathcal{A}, P)$, *à valeurs dans* $\mathbb{N}$. *Soit* $p \in \mathbb{N}^*$ *et on pose :* $\forall\omega \in \Omega,\ S_p(\omega) = \displaystyle\sum_{k=1}^{p} X_k(\omega)$ *et* $S_0 = 0$. *On considère une variable aléatoire* N, *à valeurs dans* $\mathbb{N}$, *sur le même espace probabilisé, indépendante des variables* X_k. *On définit une variable aléatoire* Y *par :* $Y = S_N$.

(a) *Déterminer* G_Y *en fonction de* G_N *et* G_X *avec* $X = X_1$.

(b) *Lorsque* $\mathbb{E}(X)$ *et* $\mathbb{E}(N)$ *existent, montrer que* $\mathbb{E}(Y)$ *existe et l'exprimer en fonction de* $\mathbb{E}(X)$ *et* $\mathbb{E}(N)$.

5 Résumé concernant les lois classiques

En notant $q = 1 - p$.

Nom	Paramètres	$X(\Omega)$	$P(X=k)$	$\mathtt{E}(X)$	$\mathtt{V}(X)$	G_X	RCV
Constante	c	$\{c\}$	1	c	0	t^c $(c \in \mathbb{N})$	$+\infty$
Uniforme	$a < b \in \mathbb{N}$	$[\![a,b]\!]$	$\dfrac{1}{b-a+1}$	$\dfrac{a+b}{2}$	$\dfrac{(b-a+1)^2-1}{12}$	$\dfrac{t^a - t^{b+1}}{(b-a+1)(1-t)}$	$+\infty$
Bernoulli	$p \in]0,1[$	$\{0,1\}$	p $(k=1)$	p	pq	$q+pt$	$+\infty$
Binomiale	$(n,p) \in \mathbb{N} \times]0,1[$	$[\![0,n]\!]$	$\dbinom{n}{k} p^k q^{n-k}$	np	npq	$(q+pt)^n$	$+\infty$
Géométrique	$p \in]0,1[$	$\mathbb{N}^*$	pq^{k-1}	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$	$\dfrac{pt}{1-qt}$	$\dfrac{1}{q}$
Poisson	$\lambda \in \mathbb{R}_+^\star$	$\mathbb{N}$	$\dfrac{e^{-\lambda}\lambda^k}{k!}$	λ	λ	$e^{\lambda(t-1)}$	$+\infty$

6 Complément : démonstration du théorème d'approximation de Weierstrass

Soit $f \in \mathcal{C}([0,1],\mathbb{R})$. Soient $n \in \mathbb{N}^*$ et $x \in [0,1]$. On pose : $B_n(f)(x) = \displaystyle\sum_{k=0}^{n} \binom{n}{k} f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$ (polynôme de Bernstein).
On considère S_n qui est une variable aléatoire suivant une loi $\mathcal{B}(n,x)$.
On rappelle que dans un exemple du cours, on a montré grâce à l'inégalité de Bienaymé-Tchebytchev que : $P\left(\left|\dfrac{S_n}{n} - x\right| > \alpha\right) \leq \dfrac{1}{4n\alpha^2}$.

1. Montrer que $\mathtt{E}\left[f\left(\dfrac{S_n}{n}\right)\right] = B_n(f)(x)$.
2. Soit $\varepsilon \in \mathbb{R}_+^*$. Justifier qu'il existe α dans $\mathbb{R}_+^*$ tel que :
   $\forall a,b \in [0,1],\ |a-b| \leq \alpha \Rightarrow |f(a) - f(b)| \leq \varepsilon$.
   Majorer $\left|f\left(\dfrac{k}{n}\right) - f(x)\right|$, pour tout k de $[\![0,n]\!]$ vérifiant $|k/n - x| \leq \alpha$.
3. Montrer que $\left|\displaystyle\sum_{|k/n-x|>\alpha} \left(f\left(\frac{k}{n}\right) - f(x)\right) P(S_n = k)\right| \leq 2\|f\|_\infty P\left(\left|\frac{S_n}{n} - x\right| > \alpha\right)$.

4. Montrer qu'il existe n_0 dans $\mathbb{N}$ tel que : $\forall n \geq n_0,\ \|B_n(f) - f\|_\infty \leq 2\varepsilon$.
5. En déduire que tout fonction f de $\mathcal{C}([0,1], \mathbb{R})$ est limite uniforme sur $[0,1]$ d'une suite de fonctions polynomiales.