Probabilités 2

Chapitre 12

1 Espérance d'une variable aléatoire discrète

Dans ce paragraphe, on se fixe un espace probabilisé $(\Omega, \mathscr{A}, P)$. On ne considère que des variables aléatoires discrètes sur $(\Omega, \mathscr{A}, P)$ à valeurs dans une partie de $\mathbf{R}$.

1.1 Cas des variables positives

Définition 12.1 (Espérance)

Soit X une variable aléatoire discrète à valeurs dans $[0, +\infty]$.
On appelle espérance de X et on note $E(X)$ la somme de la famille $(xP(X = x))_{x\in X(\Omega)}$. On a donc $E(X) \in [0, +\infty]$.
Si cette somme est finie (c'est-à-dire si la famille $(xP(X = x))_{x\in X(\Omega)}$ est sommable), on dit que X est une variable aléatoire d'espérance finie.
On note $X \in L^1$ ou $X \in L^1(\Omega, [0, +\infty])$ pour dire que X est une variable aléatoire d'espérance finie.

Remarque : Cette année on oubliera la formule

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

En effet dans la majorité des cas, Ω ne sera pas dénombrable et $P(\{\omega\})$ sera toujours nul. Pour utiliser une définition de ce genre, il faudrait développer une notion d'intégrale.

Exemple : On considère une variable aléatoire X qui suit la loi $\mathscr{G}(\frac{1}{2})$. On regarde $Y = 2^X$. C'est -à-dire que l'on lance une pièce équilibré jusqu'à obtenir pile. Si on obtient pile au lancer n (pour la première fois) on gagne 2^n euros. Si on cherche l'espérance de cette variable on trouve

Proposition 12.1

Soit X une variable aléatoire discrète à valeurs dans $[\![0, +\infty]\!] = \mathbf{N} \cup \{+\infty\}$. On a dans $[\![0, +\infty]\!]$,

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

Remarque : Soit X à valeurs entières. Pour tout $n \in \mathbf{N}^*$, $P(X \geqslant n) = P(X > n-1)$. Dans le cas où la variable ne prend que des valeurs finies, on a donc

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

Démonstration :

Théorème 12.2 (Espérance des lois usuelles)

Soit X une variable aléatoire discrète à valeurs réelles.

1. (Loi de Bernoulli) : Soit $p \in [0, 1]$, si $X \sim \mathscr{B}(p)$ alors $E(X) = p$
2. (Loi binomiale) : Soit $p \in [0, 1]$ et $n \in \mathbf{N}^*$, si $X \sim \mathscr{B}(n, p)$ alors $E(X) = np$
3. (Loi géométrique) : Soit $p \in]0, 1]$, si $X \sim \mathscr{G}(p)$ alors $E(X) = \dfrac{1}{p}$.
4. (Loi de Poisson) : Soit $\lambda \in \mathbf{R}_+^*$, si $X \sim \mathscr{P}(\lambda)$ alors $E(X) = \lambda$

Remarques :

1. Dans le cas où $p = 0$, l'espérance de $\mathscr{G}(p)$ est $+\infty$.
2. Il est facile de retenir l'espérance de la loi géométrique. En effet, si on lance un dé équilibré à 10 faces une infinité de fois. Il parait clair que le temps moyen avant le premier 10 est $10 = \dfrac{1}{\frac{1}{10}}$.

Démonstration :

Théorème 12.3 (Formule de transfert pour les variables positives)

Soit X une variable aléatoire discrète et f une fonction définie sur $X(\Omega)$ et à valeurs dans $\mathbf{R}_+$. On a dans $[0, +\infty]$

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

En particulier, $f(X)$ est d'espérance finie si et seulement si la famille $(f(x)P(X = x))_{x \in X(\Omega)}$ est sommable.

Démonstration :

Exemple : Soit $X \sim \mathscr{G}(p)$ avec $p \in]0, 1[$. On cherche à calculer $E(X^2)$. D'après le théorème

$$E(X^2) = \sum_{k \in \mathbf{N}^*} k^2 (1-p)^{k-1} p$$

Proposition 12.4 (Propriétés de l'espérance)

1. (Linéarité) : Soit X et Y deux variables aléatoires discrètes réelles positives. Dans $[0,+\infty]$:
$$E(X+Y)=E(X)+E(Y).$$
2. Soit X et Y deux variables aléatoires discrètes positives telles que $X \leqslant Y$. Dans $[0,+\infty]$, $E(X) \leqslant E(Y)$. C'est-à-dire que si on suppose que $X \leqslant Y$ et Y d'espérance finie alors X est d'espérance finie et $E(X) \leqslant E(Y)$.
3. Soit X une variable aléatoire réelle positive. Si $E(X)=0$ alors X est nulle presque surement c'est-à-dire que $P(X>0)=0$.

Démonstration :

ATTENTION

Dans la démonstration de la linéarité, on décompose $P(X+Y=z)$ en utilisant des termes de la loi conjointe du couple (X,Y) : $P((X,Y)=(x,y))$ mais on ne les décompose pas en $P(X=x)P(Y=y)$. De ce fait, le résultat ne nécessite pas que les variables X et Y soient indépendantes.

1.2 Cas des variables réelles ou complexes

On va généraliser la notion d'espérance au cas des variables aléatoires à valeurs dans $\mathbf{R}$ (mais pas nécessairement positives) et dans $\mathbf{C}$. On note $\mathbf{K}=\mathbf{R}$ ou $\mathbf{C}$.

Définition 12.2 (Espérance)

Soit X une variable aléatoire discrète à valeurs réelles ou complexes.
Elle est dite d'espérance finie si la famille $(xP(X=x))_{x\in X(\Omega)}$ est sommable (ce qui signifie que $E(|X|)<+\infty$).
Dans ce cas on appelle espérance de X la somme de cette famille :
$$E(X)=\sum_{x\in X(\Omega)} xP(X=x).$$
On note $X\in L^1$ ou $X\in L^1(\Omega,\mathbf{K})$

ATTENTION

Dans le cas où X est une variable aléatoire et que l'on veut étudier son espérance.

— Si X est à valeurs réelles positives, on peut toujours définir son espérance dans $[0,+\infty]$.

— Si X est à valeurs réelles non nécessairement positives ou complexes, il faut toujours justifier que la variable aléatoire est d'espérance finie avant de considérer son espérance. Pour cela on étudie dans $[0,+\infty]$ l'espérance de la variable aléatoire $|X|$.

Remarque : On veut éviter les séries semi-convergentes car il n'y a pas lieu d'imposer un ordre de sommation parmi les éléments de $X(\Omega)$.
Exemple : On considère une variable aléatoire X qui suit la loi $\mathscr{G}(\frac{1}{2})$. On regarde $Y = (-2)^X$. C'est -à-dire que l'on lance une pièce équilibré jusqu'à obtenir pile. Si on obtient pile au lancer n (pour la première fois) on « gagne » 2^n euros si n est pair et on « perd » 2^n si n est impair.

Définition 12.3

Soit X une variable aléatoire discrète à valeurs réelles ou complexes. Elle est dite centrée si elle est d'espérance finie et que $E(X) = 0$.

Remarque : Si $X \in L^1$, la variable $X - E(X)$ est centrée.

1.3 Propriétés de l'espérance

La plupart des propriétés vues sur l'espérance des variables aléatoires positives sont encore vraies pour les variables aléatoires réelles ou complexes d'espérance finie.

Théorème 12.5 (Formule de transfert)

Soit X une variable aléatoire discrète et f une fonction définie sur $X(\Omega)$ et à valeurs dans $\mathbf{K}$.

$$f(X) \in L^1(\Omega, \mathbf{K}) \iff (f(x)P(X = x))_{x \in X(\Omega)} \text{ est sommable}$$

Dans ce cas

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

Démonstration :

Proposition 12.6 (Propriétés de l'espérance)

1. (Linéarité) : Soit X et Y deux variables aléatoires discrètes réelles ou complexes d'espérance finie. Soit λ un scalaire, les variables λX et $X+Y$ sont d'espérance finie et
$$E(\lambda X) = \lambda E(X)\,;\, E(X+Y) = E(X) + E(Y).$$
2. Soit X une variable aléatoire discrète réelle ou complexe et Y une variable aléatoire discrète à valeurs positives. On suppose que $|X| \leqslant Y$ et Y d'espérance finie alors X est d'espérance finie et $|E(X)| \leqslant E(|X|) \leqslant E(Y)$.
3. (Positivité) : Soit X une variable aléatoire positive alors $E(X)$ est positive.
4. (Croissance) : Soit X et Y deux deux variables aléatoires discrètes réelles d'espérance finie si $X \leqslant Y$ alors $E(X) \leqslant E(Y)$.

Démonstration :

Proposition 12.7

Soit X et Y deux variables aléatoires discrètes réelles ou complexes. On suppose que X est presque surement égale à Y.

1. La variable $X-Y$ est d'espérance finie et $E(X-Y) = 0$
2. La variable aléatoire X est d'espérance finie si et seulement si Y l'est. Dans ce cas, $E(X) = E(Y)$.

Remarque : Cet énoncé ne figure pas dans le programme ; il faudrait le redémontrer le cas échéant.

Proposition 12.8

Soit X et Y deux variables aléatoires discrètes **indépendantes** d'espérances finies. La variable XY est d'espérance finie et
$$E(XY) = E(X)E(Y).$$

Démonstration :

2 Variance et écart type d'une variable aléatoire réelle

Dans ce paragraphe on se fixe un espace probabilisé $(\Omega, \mathscr{A}, P)$.

2.1 Variables aléatoires réelle de carré sommable

Définition 12.4

Soit X une variable aléatoire discrète réelle. On dit qu'elle est de carré sommable ou qu'elle admet un moment d'ordre 2 si $E(X^2) < +\infty$.
Dans ce cas, on note $X \in L^2$ ou $X \in L^2(\Omega, \mathbf{R})$.

Proposition 12.9

Soit X une variable aléatoire discrète réelle. Si $X \in L^2$ alors $X \in L^1$.

Démonstration :

ATTENTION

C'est une méthode classique qu'il faut connaitre

Exercice : Soit X une variable aléatoire positive. Soit $0 < \alpha \leqslant \beta$. Montrer que X^β est d'espérance finie alors X^α aussi.

Théorème 12.10 (Inégalité de Cauchy-Schwarz)

Soit Y et Y deux variables aléatoires discrètes réelles. On suppose $X \in L^2$ et $Y \in L^2$.

1. La variable aléatoire XY est d'espérance finie et
$$E(XY)^2 \leqslant E(X^2)E(Y^2)$$
2. Il y a égalité si et seulement si X et Y sont proportionnelles presque sûrement c'est-à-dire s'il existe $\lambda \in \mathbf{R}$ tel que $P(X = \lambda Y) = 1$ ou $P(Y = \lambda X) = 1$

Démonstration :
On utilise la preuve « classique » de l'inégalité de Cauchy-Schwarz.

Corollaire 12.11

L'ensemble des variables aléatoires discrètes à valeurs réelles ayant un moment d'ordre 2 est un sous-espace vectoriel de l'ensemble des variables aléatoires discrètes à valeurs réelles

Démonstration : L'ensemble des variables aléatoires discrètes à valeurs réelles ayant un moment d'ordre 2 n'est pas vide car il contient la variable nulle.

Définition 12.5

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$.
On appelle variance de X et on note $V(X)$,
$$V(X) = E((X - E(X))^2).$$
On appelle écart-type et on note $\sigma(X)$,
$$\sigma(X) = \sqrt{V(X)}.$$

Remarque : Comme $X \in L^2$, $X - E(X) \in L^2$.

Proposition 12.12 (Formule de König-Huygens)

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$. On a
$$V(X) = E(X^2) - E(X)^2$$

Remarque : On utilise souvent cette formule pour calculer, avec la formule de transfert la variance :
$$V(X) = \sum_{x \in X(\Omega)} x^2 P(X = x) - \left(\sum_{x \in X(\Omega)} x P(X = x) \right)^2.$$

Démonstration : En utilisant la linéarité de l'espérance :

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

□

Proposition 12.13

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$. Soit a, b deux réels,

$$V(aX + b) = a^2 V(X)\,;\, \sigma(aX + b) = |a|\sigma(X).$$

Démonstration : Il suffit de faire le calcul. □

Proposition 12.14

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$. On suppose que $V(X) = 0$, il existe alors $\lambda \in \mathbf{R}$ telle que X est presque surement égale à λ.

Démonstration : La variable aléatoire $(X - E(X))^2$ est une variable aléatoire positive d'espérance nulle. Elle est nulle presque surement et donc $X = E(X)$ presque surement. □

Proposition-Définition 12.6

Soit X une variable aléatoire discrète réelle admettant un moment d'ordre 2.

1. *Elle est dite réduite si $V(X) = 1$.*
2. *Si $\sigma(X) > 0$, la variable aléatoire $\dfrac{X - E(X)}{\sigma(X)}$ est centrée est réduite.*

2.2 Lois usuelles

Commençons par rappeler les variances des lois usuelles vues en première année

Proposition 12.15

Soit X une variable aléatoire discrète à valeurs réelles.

1. (Loi de Bernoulli) : Soit $p \in [0, 1]$, si $X \sim \mathscr{B}(p)$ alors $V(X) = pq = p(1 - p)$.
2. (Loi binomiale) : Soit $p \in [0, 1]$ et $n \in \mathbf{N}^*$, si $X \sim \mathscr{B}(n, p)$ alors $V(X) = npq$.

Passons aux lois vues cette année,

Théorème 12.16

Soit X une variable aléatoire discrète réelle.

1. (Loi géométrique) : Soit $p \in]0,1]$, si $X \sim \mathcal{G}(p)$ alors $V(X) = \dfrac{q}{p^2}$.
2. (Loi de Poisson) : Soit $\lambda \in \mathbf{R}_+^*$, si $X \sim \mathcal{P}(\lambda)$ alors $V(X) = \lambda$

Démonstration :

2.3 Covariance

On a vu que $E(XY)$ n'était pas, en général égal à $E(X)E(Y)$.

Proposition-Définition 12.7

Soit X et Y deux variables aléatoires discrètes réelles telles que $X \in L^2$ et $Y \in L^2$.
On appelle covariance de X et Y et on note $Cov(X,Y)$:

$$Cov(X,Y) = E((X - E(X))(Y - E(Y))) = E(XY) - E(X)E(Y)$$

Démonstration : Comme $X - E(X) \in L^2$ et $Y - E(Y) \in L^2$ leur produit a une espérance finie et

$$\begin{aligned} E((X - E(X))(Y - E(Y))) &= E(XY - E(X)Y - E(Y)X + E(X)E(Y)) \\ &= E(XY) - 2E(X)E(Y) + E(X)E(Y) \\ &= E(XY) - E(X)E(Y) \end{aligned}$$

□

Proposition 12.17

Soit X et Y deux variables aléatoires discrètes réelles telles que $X \in L^2$ et $Y \in L^2$. Si X et Y sont indépendantes alors $\mathrm{Cov}(X,Y) = 0$.

Démonstration : On a déjà vu que si X et Y était indépendantes alors $E(XY) = E(X)E(Y)$. On en déduit alors que $\mathrm{Cov}(X,Y) = 0$. □

ATTENTION

Ce n'est pas une équivalence, on peut avoir deux variables non indépendantes telles que $\mathrm{Cov}(X,Y) = 0$.

Exemple : On considère une variable aléatoire discrète X dont la loi est définie par $X(\Omega) = \{-1, 0, 1\}$ et

$$P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}$$

On pose alors $Y = X^2$.

Proposition 12.18

Soit E l'espace vectoriel des variables aléatoires discrètes réelles admettant un moment d'ordre 2. La covariance

$$\begin{array}{rccc} \mathrm{Cov}: & E \times E & \longrightarrow & \mathbf{R} \\ & (X, Y) & \longmapsto & \mathrm{Cov}(X, Y) \end{array}$$

est une forme bilinéaire positive. De plus pour tout X dans E,

$$\mathrm{Cov}(X, X) = V(X) = \sigma(X)^2$$

Démonstration : Il suffit d'utiliser la linéarité de l'espérance. □

Remarque : La covariance se comporte presque comme un produit scalaire. Le presque venant du fait qu'elle n'est pas définie, on peut avoir $\mathrm{Cov}(X, X) = V(X) = 0$ sans que X ne soit nul. En effet $V(X) = 0$ implique juste $X = E(X)$ presque surement et pas que $X = 0$.

Proposition 12.19

Soit X et Y deux variables aléatoires discrètes réelles telles que $X \in L^2$ et $Y \in L^2$.

$$|\mathrm{Cov}(X, Y)| \leqslant \sigma(X)\sigma(Y)$$

Démonstration : Il suffit d'utiliser l'inégalité de Cauchy-Schwarz en utilisant pour définition de la covariance $\mathrm{Cov}(X, Y) = E((X - E(X))(Y - E(Y))$ □

Corollaire 12.20

Soit $X_1, \ldots, X_n$ des variables aléatoires discrètes réelles telles que $X_i \in L^2$ pour tout $i \in [\![1 ; n]\!]$.

1. On a
$$V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i) + 2 \sum_{1 \leqslant i < j \leqslant n} \mathrm{Cov}(X_i, X_j)$$
2. Si on suppose que pour tout $(i, j) \in [\![1 ; n]\!]^2$, $i \neq j \Rightarrow \mathrm{Cov}(X_i, X_j) = 0$ alors
$$V(X_1 + \cdots + X_n) = \sum_{i=1}^{n} V(X_i)$$

Démonstration : Ce sont juste les formules classiques de développement par bilinéarité. □

Remarque : L'hypothèse dans la deuxième partie de l'énoncé est en particulier vérifiée si les variables sont deux à deux indépendantes, ce qui est le cas si la famille $(X_1, \ldots, X_n)$ est indépendante.

Exemple : On considère une suite infinie $(X_i)_{i \in \mathbf{N}^*}$ de variables aléatoires indépendantes qui suivent toutes la loi de Bernoulli $\mathscr{B}(p)$ où $p \in]0, 1[$.

Pour tout entier $n \geqslant 1$, on note Z_n le rang du n-ième succès :
$$Z_n : \omega \mapsto \min\{i \in \mathbf{N}^* , \#\{k \leqslant i \, , \, X_k(\omega) = 1\} = n\}$$

D'autre part, on considère $(Y_i)_{i \geqslant 1}$ une suite de variables aléatoires indépendantes suivant toutes la loi géométrique de paramètre p. On note pour $n \geqslant 1$, $T_n = Y_1 + \cdots + Y_n$. On va vérifier que T_n suit la même loi que Z_n. Commençons par déterminer la loi de Z_n.

3 Inégalités probabilistes et loi des grands nombres

La plupart du temps, il n'est pas possible de calculer les probabilités. On peut cependant obtenir des inégalités.

3.1 Inégalités

Théorème 12.21 (Inégalité de Markov)

Soit X une variable aléatoire discrète réelle **positive** et $a > 0$. On a

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

c'est-à-dire, $aP(X \geqslant a) \leqslant E(X)$.

Remarque : L'inégalité est évidente si $E(X) = +\infty$.
Démonstration : Donnons deux démonstrations.

Matheux (Andrei Markov : 1856 - 1922)

Sous la tutelle de Tchebychev, il devint membre de l'Académie des sciences de Saint-Pétersbourg en 1886.
Ses premiers travaux portent sur la théorie des nombres, les formes quadratiques, les fractions continues, les limites d'intégrales et les convergences de séries.
Ses travaux sur la théorie des probabilités l'ont amené à mettre au point les chaînes de Markov qui peuvent représenter les prémices de la théorie du calcul stochastique.
Il a officiellement pris sa retraite en 1905, mais a continué à enseigner jusqu'à la fin de sa vie.
Dès 1906, il commence ses recherches sur le calcul de probabilités. Il introduit de façon précise les processus aléatoires et démontre rigoureusement le théorème de la limite centrale.

Théorème 12.22 (Inégalité de Bienaymé-Tchebychev)

Soit X une variable aléatoire discrète réelle telle que $X \in L^2$.
Soit $a > 0$.

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

Démonstration :

ATTENTION

L'idée derrière la démonstration de l'inégalité de Bienaymé-Tchebycheff est une idée féconde. Il y a d'autres inégalités en probabilité qui se démontrent de la même manière à savoir.

— On étudie une variable aléatoire X

— On considère une fonction f et on pose $Y = f(X)$ où Y est une variable aléatoire positive.

— On applique l'inégalité de Markov de Y en prenant soin de relier les événements construits avec X et ceux construits avec Y.

Matheux (Pafnouti Tchebychev : 1821-1894)

Il est connu pour ses travaux dans les domaines des probabilités, des statistiques, et de la théorie des nombres.
Tchebychev reprend le vaste programme lancé par Jacques Bernoulli, Abraham de Moivre et Siméon Denis Poisson pour énoncer et démontrer de façon rigoureuse des théorèmes limites, c'est-à-dire pour établir les tendances asymptotiques des phénomènes naturels. Il établit une loi des grands nombres très générale et donne une nouvelle et brillante méthode de démonstration basée sur l'inégalité énoncée par Bienaymé et démontrée par lui-même.
En théorie des nombres, Tchebychev obtint en 1848-1852 des résultats corroborant une conjecture de Gauss et Legendre relative à la raréfaction des nombres premiers. Si ces résultats ne lui permirent pas de démontrer la conjecture (le théorème des nombres premiers), ils lui permirent néanmoins de s'en approcher considérablement, et par ailleurs de démontrer une conjecture énoncée par Bertrand : « Pour tout entier n au moins égal à 2, il existe un nombre premier entre n et $2n$ ».

Matheux (Irénée-Jules Bienaymé : 1796-1878)

Irénée-Jules Bienaymé, né à Paris le 28 août 1796 et mort à Paris le 19 octobre 1878, est un probabiliste et statisticien français. Continuateur de l'œuvre de Laplace dont il généralise la méthode des moindres carrés, il contribue à la théorie des probabilités, au développement de la statistique et à leurs applications aux calculs financiers, à la démographie et aux statistiques sociales.
Il a énoncé en particulier l'inégalité de Bienaymé-Tchebychev concernant la loi des grands nombres (1869). Il est aussi à l'origine de ce que l'on appelle aujourd'hui les arbres de Bienaymé-Galton-Watson

Exemple : On considère une expérience aléatoire décrite par l'espace probabilisé $(\Omega, \mathscr{A}), P)$ et A un événement. On note $p = P(A)$. Si on effectue de nombreuses fois (N) l'expérience de manière indépendante et que l'on note X_N le nombre de fois que l'événement A se réalise et si on note $F_N = \dfrac{X_N}{N}$ la fréquence à laquelle l'événement A se réalise alors, il semble logique que F_N va être proche de p.

Remarque : C'est un cas particulier de la loi faible des grands nombres que nous verrons plus loin.

3.2 Loi faible des grands nombres

On peut généraliser l'exemple vu au moment de l'inégalité de Bienaymé-Tchebychev.

Théorème 12.23 (Loi faible des grands nombres)

Soit $(X_n)_{n\geqslant 1}$ une suite de variables aléatoires discrètes réelles deux à deux indépendantes, suivant une même loi et telles que pour tout $n \in \mathbf{N}^*$, $X_n \in L^2$.
On pose $S_n = \sum_{i=1}^{n} X_i$ et $m = E(X_1)$.
Dans ce cas, pour tout $\varepsilon > 0$,

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

Remarques :

1. En statistiques, on s'intéresse souvent à reproduire une même expérience de nombreuses fois. Cela donne naissance pour chaque expérience à une variable aléatoire. De ce fait on récupère une suite de variables aléatoires indépendantes et qui suivent la même loi. On dit que c'est une suite de variables indépendantes et identiquement distribuées (iid).
2. Cela justifie le fait que l'espérance d'une variable aléatoire est bien ce que l'on obtiendra (presque surement) en réalisant **un très grand nombre de fois** une expérience.

Démonstration :

Remarque : On dit que la variable $\frac{S_n}{n}$ converge *en probabilité* vers $E(X)$.

De manière générale quand on se donne une suite de variables aléatoires, (X_k) et une variable aléatoire Y on peut définir plusieurs types de convergence :

— On dit que (X_k) converge *en probabilité* vers Y si, pour tout $\varepsilon > 0$,

$$\lim_{k\to\infty} P(|X_k - Y| \geqslant \varepsilon) = 0$$

— On dit que (X_k) converge *presque surement* vers Y si,

$$P(\lim_{k\to\infty} X_k = Y) = 1$$

Il y a une version « plus forte » qui s'appelle la loi *forte* des grands nombres qui dit que si les X_i ont une espérance finie alors

$$P\left(\lim_{n\to+\infty} \frac{S_n}{n} = m\right) = 1$$

C'est-à-dire que $\frac{S_n}{n}$ converge *presque sûrement* vers $E(X)$.

4 Fonctions génératrices

4.1 Généralités

Définition 12.8 (Fonctions génératrices)

Soit X une variable aléatoire à valeur dans $\mathbf{N}$, on appelle fonction génératrice de X la fonction

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

Proposition 12.24

La série entière $\sum_{n\geqslant 0} P(X = n)t^n$ converge normalement sur $[-1, 1]$.

Démonstration :

Corollaire 12.25

Avec les notations précédentes

1. Le rayon de convergence de la série entière $\sum_{n\geqslant 0} P(X = n)t^n$ est au moins 1.
2. La fonction G_X est continue sur $[-1; 1]$
3. La fonction G_X est de classe $\mathscr{C}^\infty$ sur $]-R, R[$ donc sur $]-1, 1[$

Démonstration :

Remarque : Le rayon de convergence peut être plus grand que 1, par exemple pour une loi de Poisson il est égal à $+\infty$. Il peut de même être juste égal à 1 on verra que c'est en particulier le cas si X n'a pas une espérance finie.

La connaissance de G_X permet d'obtenir de nombreuses informations sur X.

Proposition 12.26

Soit X une variable aléatoire à valeur dans $\mathbf{N}$.

1. $\forall n \in \mathbf{N}, P(X = n) = \dfrac{G_X^{(n)}(0)}{n!}$
2. $G_X(1) = 1$.

Démonstration : Évident □

Théorème 12.27

Soit X une variable aléatoire à valeur dans $\mathbf{N}$.
La variable aléatoire admet une espérance finie si et seulement si G_X est dérivable en 1. Dans ce cas $G'_X(1) = E(X)$.

Démonstration :
Remarques :

1. De même si G_X est de classe $\mathscr{C}^2$ sur $[0, 1]$, par le théorème de transfert on obtient que

$$G''_X(1) = E(X(X-1)) = E(X^2) - E(X)$$

   On peut en déduire que

$$V(X) = G''_X(1) + G'_X(1) - (G'_X(1))^2$$

2. On en déduit que si X n'est pas d'espérance finie alors le rayon de convergence de la série entière qui définie G_X est 1.

Proposition 12.28 (Fonction génératrice d'une somme)

1. Soit X et Y deux variables aléatoires **indépendantes** à valeurs dans $\mathbf{N}$.

$$\forall t \in [-1; 1], G_{X+Y}(t) = G_X(t)G_Y(t)$$

2. Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes à valeurs dans $\mathbf{N}$,

$$\forall t \in [-1, 1], G_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} G_{X_i}(t)$$

Démonstration :

4.2 Exemples

Nous allons établir les fonctions génératrices des lois usuelles.

ATTENTION

Il est bon de connaitre les formules mais il faut surtout savoir les retrouver.