Université Paris Villetaneuse Année 2013/2014
Clément Foucart

Probabilités Prépa Capes

Probabilités et Statistiques: Eléments de cours et exercices

Les programmes de mathématiques dans l'enseignement secondaire ainsi qu'en classes préparatoires et BTS mettent de plus en plus en avant la théorie des probabilités. S'il est rare que les exercices de probabilités donnés au lycée et dans les premières années universitaires nécessitent des raisonnements difficiles en analyse réelle, plusieurs types de difficultés apparaissent lorsque l'on enseigne ou que l'on étudie les probabilités. D'une part, des énoncés mal construits (trop succincts ou au contraire trop longs) peuvent rapidement bloquer le lecteur [1]. D'autre part, on peut être amené à mobiliser des connaissances en analyse ou en algèbre (calculs de séries, d'intégrales, algèbre linéaire, matrice à inverser).

La théorie de la mesure n'est pas au programme du CAPES. Nous nous bornerons donc à l'usage de notions probabilistes sans donner plus de détails quant à leurs définitions en théorie de la mesure. Il faut néanmoins bien comprendre que probabilités et intégration partagent beaucoup de choses, comme nous le verrons dans les chapitres suivants, probabilités discrètes (respectivement continues) amènent à l'étude de séries (respectivement d'intégrales). Afin de donner un cadre clair, nous introduirons rapidement les tribus (c'est d'ailleurs au programme en classes préparatoires économiques et sociales, voie scientifique). D'un point de vue pédagogique, introduire les tribus permet de s'entraîner à manipuler les sous-ensembles.

Les exercices accompagnés du signe # sont corrigés. Ils ont été dans leur grande majorité rédigés par Alexandre Genadot et Pierre Veuillez (http ://mathsece.free.fr/) [2]. Afin de vous entraîner, vous pouvez consulter les références suivantes :

Ouvrages utilisés pour la rédaction de ce polycopié

1) Intégration et Probabilités, Auliac, Cocozza-Thivent, Mercier, Roussignol, EdiScience.
2) Probabilités, analyse de données et Statistiques, Saporta, Technip.
3) Polycopié du cours de préparation au capes de Paris 6. Frédérique Petit et Béatrice de Tilière (disponible en ligne sur la page de B. de Tilière)
4) Probabilités et statistiques pour le CAPES externe et l'Agrégation interne de Mathématiques (Jérôme Escoffier)
5) Annales des ECRICOMES (pour les problèmes et exercices de probabilités : http ://mathsece.free.fr/sujetsent
6) Analyse pour l'agrégation, Zuily et Queffélec. Dunod,

Références pouvant être consultées

1) Probabilités Tome 1, licence- capes Jean-Yves Ouvrard, chez Cassini
2) Introduction au calcul des probabilités et à la statistique, Jean-François Delmas, disponible en ligne
3) Probability and random processes Grimmett and Stirzaker third edition, Oxford

1. et ceci est vrai à tout niveau.
2. J'en profite pour les remercier ici.

Table des matières

Chapitre 1

Rappels fondamentaux

Ce chapitre est du niveau de première et deuxième année. Il est conseillé d'y revenir si vous avez des lacunes.

Nous commençons par des rappels de base sur les ensembles et les opérations sur les ensembles. Nous terminons par des rappels sur les ensembles dénombrables. Ces ensembles jouent un rôle primordial lors de l'étude des probabilités discrètes. Les notions présentées ici font parties des cours de première et deuxième année de mathématiques. Nous renvoyons le lecteur aux ouvrages consacrés.

1.1 Rudiments sur les ensembles et les fonctions

1.1.1 Notions de base sur les ensembles

Soit Ω un ensemble (typiquement Ω sera inclus dans $\mathbb{N}$, $\mathbb{R}$ ou $\mathbb{R}^d$ pour $d \geq 2$). A est un sous-ensemble (on dit aussi partie) de Ω, si A est une collection d'éléments de Ω : au sens où tous les éléments de A sont éléments de Ω. On note alors

$$A \subset \Omega,$$

la relation $\subset$ est la relation d'*inclusion*.

Rappel : On raisonne souvent par double inclusion pour montrer que deux ensembles sont égaux :

$$A = B \Longleftrightarrow A \subset B \text{ et } B \subset A.$$

On note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω. C'est à dire :

$$\mathcal{P}(\Omega) = \{A; A \text{ sous-ensemble de } \Omega\}.$$

Opérations sur les ensembles :

– Ensemble complémentaire. Soit $A \in \mathcal{P}(\Omega)$. Le complémentaire de A est l'ensemble des éléments de Ω qui n'appartiennent pas à A :

$$A^c = \{x \in \Omega; x \notin A\}.$$

On note $\emptyset$ le complémentaire de Ω. Cet ensemble s'appelle l'ensemble vide.

– Réunion : La réunion de deux ensembles A et B est notée $A \cup B$ et est définie par :

$$x \in A \cup B \Longleftrightarrow x \in A \text{ ou } x \in B$$

Attention, le "ou" n'est pas *exclusif*, c'est à dire que x peut appartenir à la fois à A et à B.

– Intersection : l'intersection de deux ensembles A et B est notée $A \cap B$ et est definie par :

$$x \in A \cap B \Longleftrightarrow x \in A \text{ et } x \in B$$

– L'ensemble des éléments de A qui ne sont pas dans B est noté $A \setminus B$. On dit "A privé de B". On le définit par :

$$A \setminus B = A \cap B^c.$$

Ces opérations se généralisent à des *familles* d'ensembles : soit $I \subset \mathbb{N}$ et $(A_i, i \in I)$. Leur réunion est notée $\bigcup_{i \in I} A_i$, leur intersection $\bigcap_{i \in I} A_i$

$$x \in \bigcup_{i \in I} A_i \Longleftrightarrow \exists i \in I \text{ tel que } x \in A_i, \text{ donc } x \notin \bigcup_{i \in I} A_i \Longleftrightarrow \forall i \in I, x \notin A_i$$
$$x \in \bigcap_{i \in I} A_i \Longleftrightarrow \forall i \in I, x \in A_i, \text{ donc } x \notin \bigcap_{i \in I} A_i \Longleftrightarrow \exists i \in I \text{ tel que } x \notin A_i.$$

Remarque 1.1.1 *Toutes ces opérations sont stables dans $\mathcal{P}(\Omega)$.*

On rappelle les règles de calculs :

- $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
- $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
- $(A \cup B)^c = A^c \cap B^c$
- $(A \cap B)^c = A^c \cup B^c$

Plus généralement :

$$(\cup_{i \in I} A_i)^c = \cap_{i \in I} A_i^c \text{ et } (\cap_{i \in I} A_i)^c = \cup_{i \in I} A_i^c.$$

Définition 1.1.1 (Partition) *Une famille de parties de Ω, $(A_i, i \in I)$ est une partition de Ω si ;*
i) $\cup_{i \in I} A_i = \Omega$
ii) $\forall i \in I, j \in I : i \neq j \Longrightarrow A_i \cap A_j = \emptyset$.

Définition 1.1.2 (Produit cartésien) *Le produit cartésien de deux ensembles E et F est notè, $E \times F$ et est défini par :*

$$E \times F = \{(x, y); x \in E \text{ et } y \in F\}.$$

Plus généralement, le produit cartésien de n ensembles est :

$$E_1 \times E_2 \times E_3 \times ... \times E_n = \prod_{i=1}^{n} E_i = \{(x_1, x_2, ..., x_n); \forall i \in [|1, n|], x_i \in E_i\}.$$

1.1.2 Ensembles et fonctions

On donne ici quelques rudiments sur les applications : les définitions d'injectivité, surjectivité, de l'image et de l'image réciproque d'un ensemble par une application.
Soient E et F deux ensembles, soit f une application qui va de E dans F :

Définition 1.1.3 *La fonction f est dite*

- *injective si :*
$$\forall x, y \in E, f(x) = f(y) \Longrightarrow x = y.$$
*Dans ce cas, un élément de F **a au plus** un antécédent par f dans E.*

- *surjective si :*
$$\forall z \in F, \exists x \in E; f(x) = z.$$
*Dans ce cas, tout élément de F **a au moins** un antécédent par f.*

- *bijective si : injective et surjective. Dans ce cas, pour tout z dans F, il existe un et un seul $x \in E$ tel que $f(x) = z$.*

Définition 1.1.4 *Soit f une application de E dans F, A une partie de E et B une partie de F.*
– *On appelle image de A par f l'ensemble noté $f(A)$ défini par*
$$f(A) = \{y \in F; \exists x \in A \text{ tel que } y = f(x)\}.$$
– *On appelle image réciproque de B par f l'ensemble noté $f^{-1}(B)$ défini par*
$$f^{-1}(B) = \{x \in E; f(x) \in B\}.$$

Remarque 1.1.2 *La notation f^{-1} est abusive car f n'est pas toujours bijective et n'a donc pas toujours de réciproque, l'ensemble $f^{-1}(B)$ existe même si f n'est pas bijective.*

Proposition 1.1.1 (Image réciproque et union) *Soient $A, B \in \mathcal{P}(F)$*

- $f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B)$
- $f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B)$
- $f^{-1}(A^c) = f^{-1}(A)^c$

Preuve laissée en exercice.

1.2 Ensembles dénombrables

Pour clôturer ces généralités, et avant de passer aux probabilités, nous rappelons la notion d'ensemble dénombrable, ainsi que des éléments de dénombrement. Le dénombrement est fondamental dans les probabilités discrétes. Nous traitons ces questions à part afin de se concentrer dans la suite sur les probabilités.

1.2.1 Ensembles au plus dénombrables

Définition 1.2.1 *Un ensemble E est dit au plus dénombrable s'il existe une injection de E dans $\mathbb{N}$*

Deux cas sont possibles :
- l'ensemble E est infini et dans ce cas, on peut démontrer qu'il existe une bijection de E dans $\mathbb{N}$. L'ensemble est dit dénombrable.
- L'ensemble E est fini et dans ce cas, on peut dénombrer combien il possède d'éléments.

Remarque 1.2.1 (Exemples) *Ces propriétés sortent du cadre du cours, mais sont à savoir démontrer.*
- *L'ensemble des rationnels $\mathbb{Q}$ est dénombrable.*
- *L'ensemble des réels $\mathbb{R}$ est non dénombrable.*

1.2.2 Ensembles finis

On dit qu'un ensemble est fini s'il est vide, ou s'il existe un entier naturel n et une bijection de $[|1, n|] := \{1, 2, ..., n\}$ dans E. Si $E \neq \emptyset$, l'entier n est appelé **cardinal** de E. On note Card $E = n$.
Par convention Card $\emptyset = 0$. On a Card $E = n$ si et seulement si les éléments de E peuvent être notés $e_1, ..., e_n$ où les e_k sont distincts. Dénombrer un ensemble fini E correspond à déterminer le nombre déléments de E, c'est à dire son cardinal. Plusieurs méthodes sont possibles. Néanmoins, rappelons les trois conditions qui doivent être remplies pour que le dénombrement soit correct :

1) S'assurer que seuls des éléments de E sont comptés
2) S'assurer que l'on en oublie pas
3) S'assurer que l'on ne compte pas plusieurs fois le même élément.

Propriétes des cardinaux

Proposition 1.2.1 *Soit E un ensemble fini. Si F est un ensemble en bijection avec E, alors F est fini et Card $E =$ Card F.*

Démonstration Si E est un ensemble fini de cardinal n alors il existe ϕ une bijection de $[|1, n|]$ dans E. Soit f une bijection de E dans F, la composée de deux bijections est une bijection, donc l'application $\phi \circ f$ est une bijection. De plus le cardinal de F est n. $\square$

Proposition 1.2.2 *Soit E un ensemble fini. Toute partie A de E est finie et Card $A \leq$ Card E. Si A est une partie de E et Card $A =$ Card$E$ alors $A = E$.*

Proposition 1.2.3 *Soit A et B deux parties de E, ensemble fini.*
1) Si A et B sont disjointes, Card $(A \cup B) =$ Card $(A) +$ Card (B)
2) Card $(A \setminus B) =$ Card $A -$ Card B
3) Card $A^c =$ Card $E -$ Card A
4) Card $(A \cup B) =$ Card $A +$ Card $B -$ Card $(A \cap B)$

Proposition 1.2.4 (Formule du Crible) *Soit* $(A_i, i \in I)$ *une famille de parties de l'ensemble fini* E *alors pour tout entier* $n \geq 1$ *;*

$$Card\ (\cup_{i=1}^{n} A_i) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{1 \leq i_1 < i_2 < i_3 < ... < i_k \leq n} Card\ (\cap_{l=1}^{k} A_{i_l})$$

Si $(A_i, i \in I)$ *est une partition de* E *alors*

$$Card\ (\cup_{i=1}^{n} A_i) = Card\ (E)$$

Preuve. Par récurrence.

Cette formule correspond à la formule du crible dans le cadre des probabilités sur des ensembles finis avec l'hypothèse d'*équiprobabilité*. Nous verrons qu'elle se généralise à toutes les probabilités (pas seulement pour la probabilité uniforme).

Proposition 1.2.5 (Cardinal d'un produit d'ensembles finis) *Soit* $(E_i, i \in [|1, n|])$ *une famille d'ensembles finis. On a :*

$$Card\ (E_1 \times E_2 \times ... \times E_n) = (Card\ E_1)(Card\ E_2)...(Card\ E_n).$$

Remarque 1.2.2 *Cette formule est* ***fondamentale*** *en dénombrement. Une bonne compréhension de sa signification et de sa preuve est le meilleur départ possible en probabilités discrètes.*

Démonstration. On se concentre sur $n = 2$, une fois le résultat établi pour $n = 2$, on peut raisonner par récurrence. Soient A et B, deux ensembles de cardinal respectif n et p. On va procéder **par récurrence** sur p. Initialisons, la récurrence : supposons $p = 1$, par définition puisque Card $A = n$, il existe une bijection $\phi : A \mapsto \{1, ..., n\}$. L'application :

$$\phi : (a, b) \in A \times B \mapsto \phi(a) \in \{1, ..., n\}$$

est une bijection, donc Card $(A \times B) = n$.
Hypothèse de récurrence : Supposons que la propriété est vraie au rang p.
Hérédité : On note A et B deux ensembles avec B de cardinal $p+1$. On note $A = \{a_1, ..., a_n\}$, $B = \{b_1, ..., b_p, b_{p+1}\}$

$$A \times B = \{(a, b); a \in A, b \in B\} = A \times (\{b_1, ..., b_p\}) \cup A \times \{b_{p+1}\}.$$

Par hypothèse de récurrence Card $(A \times \{b_1, ..., b_p\}) = np$, donc

$$\text{Card}\,(A \times B) = np + n = n(p + 1).$$

La propriété étant héréditaire, initialisée, on conclut qu'elle est vraie pour tout entier p par récurrence. □

Arrangements et combinaisons

Définition 1.2.2 *On appelle p-liste d'un ensemble E à n éléments, toute suite de p éléments de E. C'est à dire un élément de E^p.*

Remarque 1.2.3 – *L'ordre des p-éléments est important.*
– *Une p-liste peut contenir plusieurs fois le même élément.*
– *On parle aussi de p-uplet, ou p-uple.*

Théorème 1.2.6 *Il y a n^p p-listes d'un ensemble à n éléments.*

Démonstration. On choisit p éléments dans E : on a donc à compter le nombre d'éléments de $\underbrace{E \times E ... \times E}_{p \text{ fois}}$. D'après la proposition 1.2.5, il y a $(Card\ E)^p$ éléments.

On peut également raisonner de la façon suivante (c'est en fait tout à fait équivalent, mais la rédaction est différente) : On choisit d'abord le premier élément de la liste : on a n possibilités. On choisit ensuite le second élément, on a également n possibilités, et ainsi de suite p fois. Le nombre de p-listes possibles est alors le produit de toutes ces possibilités : c'est à dire n^p. □
En fait, on peut identifier p-liste d'un ensemble à n éléments et application de $\{1, ..., n\}$ dans $\{1, ..., p\}$. On a donc le théorème suivant :

Théorème 1.2.7 *Soient A et B de cardinal respectif n et p. Le nombre d'applications de A dans B est $(Card\ A)^{Card\ B}$.*

Définition 1.2.3 *Un **arrangement** de p éléments parmi n est une liste de p éléments **distincts** d'un ensemble E à n éléments. Si $Card\ E = n$, un arrangement de n éléments est une **permutation** de E.*

Théorème 1.2.8 *On note A_n^p le nombre d'arrangements de p parmi n. On a :*

$$A_n^p = n(n-1)...(n-p+1) = \frac{n!}{(n-p)!}.$$

Remarque 1.2.4 *On note souvent $(n)_p$ (symbole de Pochhammer) A_n^p (néanmoins cette notation n'est pas utilisée en secondaire).*

Démonstration. Notons $\mathcal{A}_n^p$ l'ensemble des arrangements de p parmi n. On souhaite démontrer que

$$\text{Card } \mathcal{A}_n^p = \frac{n!}{(n-p)!}.$$

On procède par récurrence sur p. Si $p = 1$ alors $A_n^1 = n$ (et ce pour tout n). Soit $p \geq 1$, supposons que la propriété est vraie au rang p. Clairement $\mathcal{A}_n^{p+1}$ est en bijection avec $\mathcal{A}_n^p \times A_{n-p}^1$. Donc par le Théorème 1.2.5

$$A_n^{p+1} = A_n^p A_{n-p}^1.$$

Par hypothèse de récurrence, et car $A_{n-p}^1 = n - p$, on a

$$A_n^{p+1} = \frac{n!}{(n-p)!} \times (n-p) = \frac{n!}{(n-p-1)!}.$$

□

Définition 1.2.4 (Combinaison) *Soit E un ensemble fini de cardinal n. On appelle combinaison de k éléments de E toute partie de E de cardinal k. On note $\binom{n}{k}$ le nombre de combinaisons de k éléments parmi n. Autrement dit $\binom{n}{k}$ est le nombre de façons de choisir k éléments parmi n sans tenir compte de leur ordre. Il faut bien faire la différence entre* ***partie et liste*** *!*

Théorème 1.2.9 *Pour tout $k \leq n$*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Démonstration. On va faire une récurrence sur n. On note la propriété au rang n, P_n :

$$« \tbinom{n}{k} = \tfrac{n!}{k!(n-k)!} ».$$

On initialise la propriété : $n = 0$, on compte les parties de l'ensemble vide ; il n'y a qu'une partie l'ensemble vide lui même : $\binom{0}{0} = 1$.
On montre maintenant l'hérédité, c'est à dire : $P_n \Longrightarrow P_{n+1}$. Soit $E = \{a_1, ..., a_{n+1}\}$.
Soit $k = 0$, $\binom{n+1}{k} = 1$: la seule partie de E avec 0 élément est l'ensemble vide.
Soit $k \geq 1$, une partie de E contient l'élément a_{n+1} ou non. On écrit donc l'ensemble des parties à k éléments de E comme l'union disjointe des parties à k éléments sans a_{n+1} et des parties à k éléments avec a_{n+1}. Comptons le nombre de parties. On a $\binom{n}{k}$ parties de E contenant k éléments sans a_{n+1}. Il reste à compter les parties de E à k éléments contenant a_{n+1}. Ces parties s'écrivent de la forme $B \cup \{a_{n+1}\}$ avec B une partie de $\{a_1, ..., a_n\}$ à $k-1$ éléments. On en déduit qu'il y a $\binom{n}{k-1}$ parties de E à k éléments contenant a_{n+1}. Par hypothèse de récurrence $\binom{n}{k-1} = \frac{n!}{(k-1)!(n-k+1)!}$, donc

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} = \frac{(n+1)!}{k!(n+1-k)!}.$$

□

Remarque 1.2.5 *Au long de la preuve, on a démontré un résultat intéressant, à savoir*

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}.$$

Cette relation permet de calculer les premières valeurs de $\binom{n}{k}$. Construire le triangle de Pascal.

Proposition 1.2.10 (Formule de Vandermonde) $\forall 0 \leq p \leq n+m$

$$\binom{n+m}{p} = \sum_{k=0}^{p} \binom{n}{k}\binom{m}{p-k}.$$

Démonstration. Soient E un ensemble de cardinal n, F un ensemble disjoint de cardinal m. Dénombrer les parties à p éléments de $E \cup F$.

1.3 Exercices

Exercice 1.3.1 *Une classe comporte 20 étudiants. Douze filles et huit garçons. Le professeur décide de désigner un groupe de travail de trois élèves chargé de préparer un devoir maison.*
1) Combien de groupes de travail de trois élèves est il possible de former ?
2) Combien y a-t-il de groupes constitués de trois filles ?
3) Combien y a-t-il de groupe avec deux filles et un garçon ?

Exercice 1.3.2 *On organise un championnat de Curling. Sept équipes sont qualifiées.*
Si chaque équipe rencontre une seule fois chacune des autres équipes, quel nombre de matchs doit-on prévoir ?
Si chaque équipe rencontre deux fois chacune des autres équipes (match aller, match retour) quel nombre de matchs doit-on prévoir ?

Exercice 1.3.3 *On considère n points distincts d'un cercle avec $n \geq 3$. En joignant ces points deux à deux, combien détermine-t-on de droites ? Combien existe-t-il de triangles ayant leurs sommets en ces points ?*

Exercice 1.3.4 *Un restaurant propose trois entrées, deux plats et quatre desserts. Un menu se compose d'une entrée, un plat, un dessert. Quel est le nombre de menus ?*

Exercice 1.3.5 *On étudie le lectorat de trois revues (a=Libération, b=Le Monde, c=Le Figaro). Sur 100 personnes interrogées dans la rue, 57 lisent a, 42 lisent b, 38 lisent c, 22 lisent a et b, 14 lisent b et c, 16 lisent a et c, 8 lisent a,b et c.*
Calculer le nombre de personnes
1) qui ne lisent que a et b, que b et c, que a et c
2) qui ne lisent que a, que b, que c
3) qui ne lisent aucune des trois revues.

Exercice 1.3.6 (#) *Soit $n \in \mathbf{N}^*$. On pose :*

$$f(x) = (1+x)^n$$

1. *Développer $(1+x)^n$.*
2. *En déduire la valeur des sommes suivantes :*

$$A = \sum_{k=0}^{n} \binom{n}{k}, \quad B = \sum_{k=0}^{n} \binom{n}{k} 2^k, \quad C = \sum_{k=0}^{n} \binom{n}{k} (-1)^k$$

3. *Calculer la dérivée $f'(x)$ de deux façons.*
4. *En déduire la somme :*

$$D = \binom{n}{1} + 2\binom{n}{2} + 3\binom{n}{3} + \cdots + n\binom{n}{n}$$

Correction.

1. D'après la formule du binôme :

$$f(x) = (x+1)^n = \sum_{p=0}^{n} \binom{n}{p} x^p 1^{n-p} = \sum_{p=0}^{n} \binom{n}{p} x^p = \binom{n}{0} + \binom{n}{1} x + \binom{n}{2} x^2 + \cdots + \binom{n}{n} x^n$$

2. On remarque, grâce à la question précédente, que :

$$A = f(1) = (1+1)^n = 2^n$$

$$B = f(2) = (1+2)^n = 3^n$$

$$C = f(-1) = (1-1)^n = 0 \qquad \text{car } n \geq 1 \text{ par hypothèse (rappelons que } 0^0 = 1)$$

3. Calculons $f'(x)$ de deux façons. D'une part $f(x) = (1+x)^n$ donc :

$$f'(x) = n(1+x)^{n-1}$$

On a utilisé la formule : $(u^n)' = nu'u^{n-1}$. D'autre part :

$$f(x) = \binom{n}{0} + \binom{n}{1} x + \binom{n}{2} x^2 + \cdots + \binom{n}{n} x^n$$

Donc :

$$f'(x) = \binom{n}{1} x + 2\binom{n}{2} x + \cdots + n\binom{n}{n} x^{n-1}$$

4. Avec la question précédente, on voit que :

$$D = f'(1) = n(1+1)^{n-1} = n2^{n-1}$$

Exercice 1.3.7 (#) *Un étudiant va acheter trois livres de math et deux bandes dessinées. Dans le magasin, il y a dix livres de math et vingt bandes dessinées.*

1. *De combien de façons l'étudiant peut-il faire ses achats ?*
2. *De retour chez lui, il forme une pile avec ses nouveaux livres. De combien de façons peut-il le faire ?*
3. *Même question si l'étudiant souhaite que ses livres de math se trouvent en bas de la pile, et ses bandes dessinées en haut.*
4. *Même question si l'étudiant souhaite que ses livres de math se suivent dans la pile (mais pas nécessairement les bandes dessinées).*

Correction.

1. Pour l'achat des livres, l'ordre n'a pas d'importance. L'étudiant choisit 3 livres de math parmi 10 et 2 bandes dessinées parmi 20. Donc le nombre de façons de faire ces achats est : $\binom{10}{3}\binom{20}{2} = \frac{10\times9\times8}{3\times2} \times \frac{20\times19}{2} = 10 \times 3 \times 4 \times 10 \times 19 = 22800$.
2. Ici l'ordre compte. L'étudiant choisit un de ses 5 livres pour être le premier de la pile, puis un des 4 autres pour être le deuxième, etc. Il a donc $5 \times 4 \times 3 \times 2 = 5! = 120$ façons de constituer la pile. (Nombre de permutations d'un ensemble de cardinal 5)

3. Dans cette question, l'étudiant forme une pile avec ses 3 livres de math : il a $3 \times 2 = 6$ façons de le faire. Puis une pile avec ses bandes dessinées : il a 2 façons de le faire. Enfin, il place la pile de livres de math sous la pile de bande dessinées, mais il n'y a qu'une seule façon de faire cela. Il a donc $6 \times 2 = 12$ façons de constituer la pile de 5 livres.
4. Ici, l'étudiant forme une pile avec ses 3 livres de math (donc 6 façons de le faire, on l'a vu), une pile avec ses bandes dessinées (2 façons de le faire), puis il insère la pile de livre de math :
   – sous les bandes dessinées
   – ou bien entre les deux bandes dessinées
   – ou bien au dessus des bandes dessinées
   Finalement, il a donc $6 \times 2 \times 3 = 36$ façons de constituer la pile.

Exercice 1.3.8 10 *livres doivent être rangés dans une bibliothèque, dont* 4 *livres de maths,* 3 *livres de chimie,* 2 *livres de littérature, et* 1 *d'histoire. On souhaite ranger les livres de manière que les livres du même sujet soient regroupés sur l'étagère. Combien d'arrangements sont possibles ?*

Exercice 1.3.9
1. *On appelle partage par paire d'un ensemble toute partition dont les parties contiennent chacune deux éléments. Quel est le nombre de partages par paires d'un ensemble de* $2n$ *éléments ?*
2. *On considère* 32 *joueurs de tennis, de combien de façons peut-on organiser le premier tour d'un tournoi de tennis en simple ? En double ?*

Exercice 1.3.10 Soit $\Omega = \{1, 2, 3, 4\}$. Décrire toutes les parties de Ω, puis vérifier que $\text{card}(\mathcal{P}(\Omega)) = 2^4$. Démontrer que si card $(\Omega) = n$ alors $\text{card}(\mathcal{P}(\Omega)) = 2^n$

Exercice 1.3.11 Montrer que pour tous $n, p \in \mathbb{N}$,

$$\binom{n}{p} = \binom{n}{n-p} \quad \text{et} \quad \binom{n+1}{p+1} = \binom{n}{p} + \binom{n}{p+1}.$$

Exercice 1.3.12 Soient A , B et C trois parties d'un même ensemble E.

1. A l'aide de dessins, justifier les lois de Morgan :

$$(A \cup B)^c = A^c \cap B^c \quad \text{et} \quad (A \cap B)^c = A^c \cup B^c.$$

2. A l'aide de dessins, Justifier les relations ensemblistes suivantes :

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{et} \quad A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

3. Donner une forme simplifiée des expressions :

$$(A \cup B) \cap (A \cup B^c) \quad \text{et} \quad (A \cup B) \cap (A^c \cup B) \cap (A^c \cup B^c).$$

Chapitre 2

Espaces probabilisés

Afin de motiver les définitions qui suivent, voici deux exemples types de problèmes où les probabilités se révèlent nécessaires. Il est courant que des exercices soient énoncés de cette manière en secondaire ou dans les premières années d'université.

Exemple 2.0.1 (Contrôle de production) *Une usine produit à grande échelle une pièce mécanique. Le processus de fabrication est complexe et ne permet pas de détecter à tout moment si une pièce est défectueuse. Néanmoins, le fabricant a des engagements envers ses clients et souhaite vérifier que son stock ne contient pas trop de pièces défectueuses. Pour en estimer le nombre, on peut par exemple prendre 100 pièces et compter combien sont défectueuses. Si on en trouve 3, on a envie de conclure que 3% de la production sera défectueuse. Ce n'est évidemment pas précis, et le calcul des probabilités est nécessaire pour répondre à la question.*

Exemple 2.0.2 (Fiabilité) *Imaginons qu'un système embarque plusieurs composants électroniques qui tombent en panne à des temps aléatoires. Afin d'assurer la fiabilité du système, on a besoin d'informations précises sur ces temps aléatoires. Cela nous conduira à considérer des lois de probabilités.*

Dans cette section, nous présentons les fondements de la théorie. Les exercices associés sont un peu abstraits mais permettent une compréhension plus globale de la théorie.

2.1 Terminologie probabiliste et notion de probabilité

La première étape pour étudier un phénomène aléatoire est d'identifier les résultats possibles d'une expérience impliquant ce phénomène.

Définition 2.1.1 *On note Ω l'ensemble de tous les résultats possibles. On appelle cet ensemble l'* ***univers*** *(on parle également d'ensemble fondamental ou d'espace d'états).*

Reprenons l'exemple 2.0.1. Supposons que sur $N = 10000$ pièces, un nombre m de pièces sont défectueuses. On tire maintenant au sort $n = 100$. Comment décrire Ω ?

- On peut prendre pour univers Ω l'ensemble des tirages de n pièces sans tenir compte de l'ordre : c'est à dire l'ensemble des combinaisons de n éléments parmi N. Le cardinal de Ω est donc le nombre de combinaisons de n éléments parmi N : $\binom{N}{n} = \frac{N!}{n!(N-n)!}$.
- On peut prendre pour univers l'ensemble des tirages tenant compte de l'ordre : dans ce cas, il s'agit de l'ensemble des arrangements de n éléments parmi N. Le cardinal de Ω est donc A_n^N.

- On peut également prendre pour univers l'ensemble décrit par le nombre de pièces défectueuses tirées : $\Omega = \{0, 1, ..., n\}$.

Il n'y a donc pas unicité de l'espace Ω. Plusieurs choix sont possibles, le meilleur est toujours celui qui mène aux calculs les plus simples... Dans l'exemple de durée de fonctionnement, s'il y a un seul composant, l'espace de probabilité naturel pour modéliser le temps de panne est $\Omega = \mathbb{R}_+$. S'il y a N composants, on peut prendre $\Omega = \mathbb{R}_+^N$.

Une fois le choix de l'univers clair, on peut s'intéresser à certains **événements**. Par exemple « le nombre de pièces défectueuses est inférieur à 5 (dans l'exemple 2.0.1) » ou « le composant ne tombe pas en panne dans les 10 prochaines années ». Si l'on peut décrire un événement à l'aide d'une phrase, il faut bien comprendre que c'est un sous-ensemble de l'univers Ω[1] Le tableau suivant donne les correspondances entre ensembles et événements.

Notations	vocabulaire ensembliste	vocabulaire probabiliste
ω	élément de Ω	tirage, résultat possible ou expérience
$\omega \in A$	ω appartient à A	l'expérience ω réalise A
$A \subset B$	A contenu dans B	l'événement A **entraîne** l'événement B
$A \cup B$	réunion de A et B	A **ou** B
$A \cap B$	intersection de A et B	A **et** B
A^c	ensemble complémentaire de A	événement "non A"
$\emptyset$	ensemble vide	événement impossible
Ω	ensemble plein	événement certain
$A \cap B = \emptyset$	A et B disjoints	A et B événements incompatibles.

L'objectif est donc d'étudier les occurrences de ces événements. On entend par occurrence, la probabilité d'apparition de l'événement : c'est à dire un réel appartenant au segment $[0, 1]$, vérifiant certaines propriétés naturelles (voir Définition 2.1.5). Lorsque Ω est non-dénombrable, il est impossible d'attribuer à chaque élément de $\mathcal{P}(\Omega)$, une probabilité vérifiant ces propriétés et on ne peut définir $\mathbb{P}$ que sur un sous-ensemble strict de $\mathcal{P}(\Omega)$. Ces difficultés mènent à la notion de *tribu* qui peut être passée en première lecture.

Définition 2.1.2 *Une tribu (aussi appelée σ-algèbre) sur un ensemble Ω est une famille $\mathcal{F}$ de parties de Ω, vérifiant les trois propriétés suivantes :*
- $\Omega \in \mathcal{F}$
- *Pour tout élément A de $\mathcal{F}$, A^c appartient à $\mathcal{F}$ (on parle de stabilité par passage au complémentaire)*
- *Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{F}$, $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{F}$.*

Proposition 2.1.1 – *L'intersection dénombrable de tribus est une tribu (voir Exercice).*
- *En général, la réunion de deux tribus n'est pas une tribu !*

Définition 2.1.3 *La tribu engendrée par un ensemble de parties $\mathcal{G}$ de Ω est la plus petite tribu,*

1. c'est une difficulté lorsque l'on débute en probabilité.

au sens de l'inclusion, contenant $\mathcal{G}$. On a donc

$$\sigma(\mathcal{G}) := \bigcap_{\mathcal{A} \text{ tribu } ; \mathcal{G} \in \mathcal{A}} \mathcal{A}.$$

Définition 2.1.4 *Une tribu importante, à laquelle nous ferons implicitement référence dans le chapitre 4, est la* ***tribu des boréliens.*** *La tribu des boréliens est la tribu sur $\mathbb{R}^d$ engendrée par l'ensemble $\mathcal{G}$ des pavés ouverts (des intervalles ouverts si $d = 1$). On la notera $\mathcal{B}(\mathbb{R}^d)$.*

Introduisons maintenant la notion de probabilité. La première notion de probabilité étudiée en secondaire est celle d'équiprobabilité dans un univers Ω fini. Supposons que de par l'expérience tous les éléments de Ω ont la même chance d'être obtenu (c'est la notion d'*équiprobabilité*) on définit une notion de probabilité de la façon suivante. Soit A un événement. La probabilité de A est la proportion d'éléments de A dans Ω. On définit

$$\mathbb{P}(A) = \frac{\text{Card } A}{\text{Card } \Omega}.$$

Les propriétés du cardinal d'un ensemble conduisent immédiatement à la célèbre formule

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

En effet Card $(A \cup B) =$ Card $A +$ Card $B -$ Card $(A \cap B)$. Cette notion de probabilité est d'un certain point de vue la plus simple et ramène le calcul à des questions de dénombrement (parfois loin d'être faciles).

Remarque 2.1.1 *Ces dernières années, les programmes ne mettent pas l'accent sur le dénombrement, néanmoins vous devez en connaître les rudiments.*

La situation d'équiprobabilité est un cas particulier (on parlera dans la suite de loi uniforme). Pour avoir une définition tout à fait générale et pouvoir par exemple étudier les probabilités pour des expériences liées à des univers infinis, suivons l'*approche fréquentiste.* Lorsque l'on peut répéter un grand nombre de fois une même expérience aléatoire, on constate que *la fréquence d'apparition* d'un événement A, c'est à dire le quotient

$$\frac{\text{nombre de fois que A se réalise}}{\text{nombre de fois que l'on fait l'expérience}}$$

fluctue de moins en moins et semble converger vers une limite au fur et à mesure que le nombre d'expériences grandit. On note cette quantité

$$\mathbb{P}(A).$$

De façon heuristique, si $\mathbb{P}(A)$ est proche de 1, A s'est produit très souvent et donc lors d'une expérience, l'événement A a beaucoup de chances de se réaliser. Au contraire si $\mathbb{P}(A)$ est proche de 0, alors A s'est peu produit et la chance qu'il se réalise est petite.
Si A et B sont des événements incompatibles, alors le nombre de fois ou l'événement $A \cup B$ se réalise est simplement la somme du nombre de fois où A s'est réalisé et du nombre de fois ou B s'est réalisé. En "passant à la limite", on a si $A \cap B = \emptyset$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Présentée telle quelle la notion de probabilité n'est pas du tout rigoureuse. En effet, nous ne pouvons pas a priori manipuler des limites sans savoir si elles existent. Cependant, cette heuristique nous indique déjà qu'une probabilité est une application qui va de l'ensemble des événements à l'intervalle $[0,1]$ et qui de plus vérifient certaines propriétés d'additivité. Nous allons poser comme axiome qu'une probabilité vérifie la σ-additivité, qui est une généralisation de l'additivité vue ci-dessus.

Définition 2.1.5 *Si $\mathcal{F}$ est une tribu sur l'ensemble Ω, une probabilité sur la tribu $\mathcal{F}$ est une application notée $\mathbb{P}$ de $\mathcal{F}$ dans $[0,1]$ telle que :*
– $\mathbb{P}(\Omega) = 1$
– *Pour toute suite (finie ou infinie) $(A_n, n \geq 1)$ d'événements deux à deux disjoints de $\mathcal{F}$,*

$$\mathbb{P}(\bigcup_{n\geq 1} A_n) = \sum_{n\geq 1} \mathbb{P}(A_n).$$

Le triplet $(\Omega, \mathcal{F}, \mathbb{P})$ est appelé espace probabilisé.

Avec cette définition, c'est un **théorème** qui établira la convergence des fréquences d'apparition de cet événement vers sa probabilité (Loi des grands nombres, voir Chapitre 5).

Remarque 2.1.2 *Il est légitime de se poser la question de l'intérêt de la notion de tribu. Lorsque Ω est au plus dénombrable, cela ne pose en général pas de problème, on prendra $\mathcal{F} = \mathcal{P}(\Omega)$. Lorsque Ω est infini non dénombrable, typiquement $\Omega = \mathbb{R}$, on ne peut pas définir une probabilité sur la tribu $\mathcal{P}(\Omega)$. Il existe des parties de $\mathbb{R}$ qui n'appartiennent pas aux boréliens. Nous n'irons pas plus loin sur ces questions qui sortent du programme. Mais il est bon de savoir qu'il existe des ensembles non boréliens.*

Dans la proposition suivante sont rassemblées des propriétés élémentaires **mais fondamentales** liées à la définition :

Proposition 2.1.2 – *Pour tout événement A, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$*

– *Pour tous événements A et B, $\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c)$. Plus généralement si $(A_i, i \in I)$ est une partition de Ω alors $\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(B \cap A_i)$.*

– *Si $A \subset B$ alors $\mathbb{P}(A) \leq \mathbb{P}(B)$ et $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.*

– *Pour tous événements A, B : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.*

– *Si $A_1, A_2, ..., A_n, ...$ est une suite d'événements : on a*

$$\mathbb{P}\left(\bigcap_n A_n\right) = 1 - \mathbb{P}\left(\bigcup_n A_n^c\right) \text{ et } \mathbb{P}\left(\bigcup_n A_n\right) = 1 - \mathbb{P}\left(\bigcap_n A_n^c\right)$$

– *Soient n événements $A_1, A_2, ..., A_n$ alors $\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1^n} \mathbb{P}(A_i)$*

Proposition 2.1.3 (Propriété de la limite monotone) *1) Si $A_1, A_2, ..., A_n, ...$ est une suite croissante (pour l'inclusion) d'événements et que l'on note $A = \bigcup A_n$ alors*

$$\mathbb{P}(A) = \lim_{n\to\infty} \uparrow \mathbb{P}(A_n).$$

2) Si $A_1, A_2, ..., A_n, ...$ est une suite décroissante (pour l'inclusion) d'événements et que l'on note $A = \bigcap A_n$ alors

$$\mathbb{P}(A) = \lim_{n\to\infty} \downarrow \mathbb{P}(A_n).$$

Démonstration. On démontre seulement 1). La preuve de 2) est similaire, il suffit de passer au complémentaire. Si $(A_n, n \geq 1)$ est une suite croissante au sens de l'inclusion alors $(\mathbb{P}(A_n), n \geq 1)$ est une suite réelle croissante majorée par 1. La suite est donc convergente. On définit les événements suivants $B_1 = A_1, B_2 = A_2 \setminus A_1$,

$$B_i = A_i \setminus A_{i-1}.$$

Clairement $\bigcup_{i=1}^n B_i \subset \bigcup_{i=1}^n A_i$. Vérifions l'inclusion réciproque. Soit $\omega \in \bigcup_{i=1}^n A_i$. On pose $k = \min\{i \in [|1,n|]; \omega \in A_i\}$. De deux choses l'une, si $k = 1, \omega \in A_1 = B_1$, si $k > 1, \omega \in A_k \setminus A_{k-1} = B_k$. Donc $\bigcup_{i=1}^n A_i \subset \bigcup_{i=1}^n B_i$. Par σ-additivité :

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(B_i)$$

Or $\sum_{k=1}^n \mathbb{P}(B_k) = \mathbb{P}\left(\bigcup_{k=1}^n B_k\right) = \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \mathbb{P}(A_n)$, donc

$$(\mathbb{P}(A_n), n \geq 1) \text{ converge.}$$

□

Proposition 2.1.4 (Formule du Crible) *Soit $(A_i, i \in I)$ une famille d'événements alors pour tout entier $n \geq 1$;*

$$\mathbb{P}\left(\cup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1\leq i_1<i_2<i_3<...<i_k\leq n} \mathbb{P}(\cap_{l=1}^k A_{i_l}).$$

Démonstration. A vous !

Trois cas fondamentaux sont à distinguer concernant l'univers Ω.

Univers fini

Supposons que l'expérience donne p résultats possibles $\Omega = \{a_1, ..., a_p\}$. On note la probabilité de l'événement élémentaire a_i, $\mathbb{P}[\{a_i\}] = p_i$. Les événements $\{a_i\}$ sont disjoints et quelque soit l'événement A, on a $A = \bigcup_{i;a_i\in A}\{a_i\}$ et donc

$$\mathbb{P}(A) = \sum_{i\in[|1,p|], a_i\in A} p_i.$$

Lorsque Ω est fini le calcul d'une probabilité revient au calcul d'une *somme finie*.
Lorsque de plus, l'expérience a des réalisations équiprobables, $p_i = p$ élement de $[0, 1]$ indépendant de i et

$$\mathbb{P}(A) = \sum_{i;a_i\in A} p_i = \sum_{i;a_i\in A} p = p\text{Card } A.$$

On obtient ainsi avec $A = \Omega$, $p = \frac{1}{\text{Card } \Omega}$. Dans ce cas, le calcul se ramène à un problème de dénombrement. Nous retiendrons la formule

$$p = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Exemple 2.1.1 (Les anniversaires) *On cherche à calculer la probabilité qu'au moins deux personnes parmi n soient nées le même jour. On formalise le problème en prenant $\Omega = [|1, 365|]^n$, c'est à dire l'ensemble des anniversaires possibles parmi les n. On suppose que tous les jours sont équiprobables. Soit A l'événement « Au moins, deux personnes ont leur anniversaire le même jour » Calculer $\mathbb{P}[A]$ peut sembler difficile, il faut penser à regarder ce que signifie l'événement contraire :*

$$A^c : \text{« chaque personne parmi les } n \text{ est née un jour différent. »}$$

Puisqu'on a supposé que tout était équiprobable,

$$\mathbb{P}[A^c] = \frac{\textit{nombres de cas favorables}}{\textit{nombres de cas possibles}} = \frac{A_{365}^n}{365^n}$$

On a finalement

$$\mathbb{P}[A] = 1 - \mathbb{P}[A^c] = 1 - \frac{A_{365}^n}{365^n} \approx 0.5 \textit{ pour } n = 23.$$

Nous avons vu précédemment que plusieurs choix d'univers sont parfois possibles pour une même expérience aléatoire (néanmoins nous parlerons de **l**'univers). Certains choix aident les calculs. On reprend l'exemple 2.0.1 du contrôle de production où nous allons voir apparaître une première loi usuelle importante : la loi hypergéométrique.

Exercice 2.1.1 (#) *On tire n pièces sans remise dans une production de N pièces parmi lesquelles m sont défectueuses. Quelle est la probabilité de l'événement : « le nombre de pièces tirées défectueuses est égal à k » ?*

Solution : Si on choisit de prendre comme univers Ω, l'ensemble des tirages de n pièces parmi N sans tenir compte de l'ordre. C'est à dire l'ensemble des combinaisons de n parmi N. Chaque élément de Ω a donc pour probabilité $\binom{N}{n}$, et nous sommes dans une situation d'équiprobabilité : (il y a autant de raison de tirer un n-uplet non ordonné qu'un autre). l'événement A_k : « le nombre de pièces tirées qui sont défectueuses est égal à k »correspond à la partie de Ω des n-uplets non ordonnés qui contiennent exactement k objets défectueux. On a donc

$$\mathbb{P}(A_k) = \frac{\text{Card } A_k}{\text{Card } \Omega} = \frac{\text{Card } A_k}{\binom{N}{n}}.$$

On cherche donc à calculer Card A_k. Si $k \geq n + 1$ alors clairement Card $A_k = 0$. Si $k > m$ ou $n - k > N - m$ alors Card $A_k = 0$ (on ne peut pas tirer plus de pièces défectueuses que m, ni avoir plus de pièces non-défectueuses que $N - m$). Dans les autres cas : on compte Card $A_k = \binom{k}{m}\binom{n-k}{N-m}$. Donc finalement :

$$\mathbb{P}(A_k) = \frac{\binom{k}{m}\binom{n-k}{N-m}}{\binom{N}{n}}.$$

C'est la loi hypergéométrique. L'exercice est résolu.

Prenons maintenant l'univers des tirages où l'ordre compte, c'est à dire que Ω est l'ensemble des arrangements de n parmi N. Le fait de prendre l'ordre des pièces que l'on tire en compte, ne change rien quant au fait de tirer une pièce défectueuse ou non. Il n'y a pas plus de chance de tirer tel ou tel arrangement. Nous sommes à nouveau dans un cadre équiprobable. On a cette fois Card $\Omega = A_N^n$. De la même façon que précédemment, on identifie l'événement à $B_k = \{$arrangements de n objets parmi N avec k défectueux$\}$. Il faut calculer Card B_k :
Pour réaliser un n-uplet ordonné avec k éléments défectueux, on commence par choisir la place des éléments défectueux : on a $\binom{n}{k}$ possibilités. On choisit un k-uplet ordonné parmi les m défectueux : on a A_m^k possibilités. Il faut aussi choisir un $(n-k)$-uplet ordonné parmi les $N-m$ objets non défectueux : on en a A_{N-m}^{n-k}. Finalement

$$\begin{aligned}
\text{Card } B_k &= \binom{n}{k} A_m^k A_{N-m}^{n-k} \\
&= \frac{n!}{k!(n-k)!} \frac{m!}{(m-k)!} \frac{(N-m)!}{(N-m-n+k)!} \\
&= n! \frac{m!}{k!(m-k)!} \frac{(N-m)!}{(n-k)!(N-m-n+k)!} \\
&= n! \binom{m}{k} \binom{n-k}{N-m}
\end{aligned}$$

On a donc

$$\mathbb{P}(B_k) = \frac{n!\binom{m}{k}\binom{n-k}{N-m}}{A_N^n} = \frac{\binom{k}{m}\binom{n-k}{N-m}}{\binom{N}{n}}.$$

Ce qui correspond à la probabilité calculée précédemment. Il faut noter que du fait de l'expérience, il n'est pas naturel de prendre en compte l'ordre des pièces que l'on tire et que le calcul est plus dangereux.

Univers infini dénombrable

Si $\Omega = \{a_1, a_2, ..., a_i, ...\}$, on note $p_i = \mathbb{P}(\{a_i\})$. Comme précédemment $\Omega = \bigcup_{i\geq 1}\{a_i\}$, et $\sum_{i\geq 1} p_i = 1$. De plus tout événement A peut s'écrire $A = \bigcup_{i:a_i\in A}\{a_i\}$, d'où

$$\mathbb{P}(A) = \sum_{i;a_i\in A} p_i.$$

Donc lorsque Ω est dénombrable, le calcul d'une probabilité sera lié à celui de la somme d'une certaine série. Il est important de noter que dans le cas d'un univers infini dénombrable il n'y a plus de notion d'équiprobabilité ! En effet supposons que p_i est une constante p, l'égalité $\sum_{i\geq 1} p_i = 1$ n'est plus vérifiée.

Remarque 2.1.3 *A ce propos, tout texte commençant par une phrase du type « On choisit un nombre au hasard »n'a pas de sens mathématique. L'ensemble des entiers est infini dénombrable, et vous ne pouvez pas choisir de façon équiprobable un entier.*

Univers infini non-dénombrable

Le dernier cas à distinguer est celui où Ω est non-dénombrable. Vous pouvez facilement imaginer une expérience où l'univers est $\mathbb{R}$ (exemple ?).

Exemple 2.1.2 (Fléchettes) *On considère l'expèrience suivante : on lance une fléchette sur une cible. On s'intéresse à la position de la fléchette. Un choix d'univers Ω est clairement $\mathcal{P}(D)$ où D est le disque représentant la cible. Soit $A \in \mathcal{P}(\Omega)$, on a envie de définir la probabilité que A se réalise par*

$$\mathbb{P}(A) = \frac{\textit{aire de } A}{\textit{aire de } D}.$$

Cependant la théorie de la mesure nous dit qu'il existe des sous-ensembles A pour lesquels on ne peut pas définir d'aire. Il faut se restreindre aux parties de Ω qui ont une aire bien définie.

- *Un univers associé à l'étude de la distance entre la fléchette et le centre du disque est $\Omega = \mathbb{R}^2$ et on peut prendre la tribu des boréliens.*
- *Pour étudier la trajectoire de la fléchette, on peut prendre $\Omega = \mathcal{C}(\mathbb{R}_+, D)$: l'ensemble des fonctions continues. Le choix de la tribu est plus difficile et sort du cadre du cours.*

2.2 Probabilité conditionnelle

On définit dans ce paragraphe la notion de probabilité conditionnelle.

Définition 2.2.1 *Soient A et B deux événements avec $\mathbb{P}(B) > 0$. On appelle probabilité conditionnelle de A sachant B et on note*

$$\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

La quantité $\mathbb{P}_B(A)$ est la probabilité que l'événement A se réalise sachant que B est arrivé. Si on sait que B est arrivé, on peut restreindre l'univers à $\Omega \cap B$.

Remarque 2.2.1
- *Pour comprendre pourquoi on pose cette définition, on peut reprendre l'approche fréquentiste. De façon heuristique, si on note $N_{A\cap B}$ et N_B respectivement le nombre de fois que A et B se réalisent et le nombre de fois que B se réalise, parmi N expériences. On a*

$$\mathbb{P}_B(A) = \lim_{N\to\infty} \frac{N_{A\cap B}}{N_B} = \lim_{N\to\infty} \frac{N_{A\cap B}}{N} \times \frac{N}{N_B} = \lim_{N\to\infty} \frac{N_{A\cap B}}{N} \times \lim_{N\to\infty} \frac{N}{N_B} = \frac{\mathbb{P}(A\cap B)}{\mathbb{P}(B)}.$$

- *Il existe une autre notation couramment utilisée pour les probabilités conditionnelles : $\mathbb{P}(A|B)$. Cette notation n'est pas celle au programme du secondaire. D'autre part, il faut bien comprendre que $A|B$ n'est pas un événement ! $A|B$ seul n'étant pas défini ne doit pas apparaître dans vos copies !*

Proposition 2.2.1 *Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et B élément de $\mathcal{F}$ tel que $\mathbb{P}(B) > 0$. Alors l'application $\mathbb{P}_B$ de $\mathcal{F}$ dans $[0,1]$ est une probabilité.*

Démonstration. Laissée en exercice.

Exercice 2.2.1 *On reprend l'exemple du contrôle de production : N objets, m défectueux. On suppose que l'on tire les éléments de l'échantillon les uns après les autres. La probabilité que le premier objet tiré soit défectueux est* $\frac{m}{N}$. *Quelle est la probabilité de tirer un objet défectueux au deuxième tirage sachant que l'on a tiré un objet défectueux au premier tirage ?*

Exercice 2.2.2 *[Test de maladie] Un laboratoire pharmaceutique met au point un test pour détecter une maladie. On sait que la maladie touche 1% de la population. On effectue le test sur un échantillon de malades et voit que 95% sont positifs. On effectue ensuite le test sur une population de non malades et 5% sont déclarés positifs par le test. Formaliser l'énoncé (donner un univers). Est-ce un bon test ?*

Les événements à considérer sont : « Malade », « Non malade », « Test positif », « Test négatif ». Les données du problèmes sont alors : $\mathbb{P}(\text{Malade}) = 0.01$, $\mathbb{P}_{\text{Malade}}(\text{Test positif}) = 0.95$ et $\mathbb{P}(\text{Test négatif}) = 0.05$. Une façon de juger la qualité du test est de calculer la probabilité d'être malade sachant que le test est positif. C'est à dire $\mathbb{P}_{\text{Test positif}}(\text{Malade})$. Pour faire ce calcul, on a besoin de la formule de Bayes.

2.2.1 Formule des probabilités totales et formule de Bayes

Proposition 2.2.2 (formule des probabilités composées) *Soient n événements* $A_1, ..., A_n$ *tel que* $\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_{n-1}) > 0$, *on a :*

$$\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1 \cap A_2)...\mathbb{P}(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1}).$$

Démonstration. Si $\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_{n-1}) > 0$ alors pour tout $1 \leq i \leq n-1$ $\mathbb{P}(A_1 \cap A_2 \cap ... \cap A_i) > 0$. Faire une récurrence.

Proposition 2.2.3 *Soit* B *un événement tel que* $\mathbb{P}(B) > 0$ *et* $\mathbb{P}(B^c) > 0$ *alors pour tout événement* A *:*

$$\mathbb{P}(A) = \mathbb{P}_B(A)\mathbb{P}(B) + \mathbb{P}_{B^c}(A)\mathbb{P}(B^c).$$

Démonstration. On a $A = (A \cap B) \cup (A \cap B^c)$ et les événements $A \cap B$ et $A \cap B^c$ sont disjoints donc

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c),$$

on conclut par définition des probabilités conditionnelles.

On généralise cette proposition :

Proposition 2.2.4 (Formule des probabilités totales) *Si* $(B_i, i \geq 1)$ *est une famille au plus dénombrable d'événements formant une partition de* Ω *et* $\mathbb{P}(B_i) > 0$ *pour tout* $i \in I$, *alors pour tout événement* A, *on a :*

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i).$$

Démonstration. Laissée en exercice.

Remarque 2.2.2 (Lien avec les arbres pondérés) *Les formules des probabilités composées et des probabilités totales sont à rapprocher de la représentation en arbre pondéré utilisée au lycée pour représenter une expérience aléatoire. On rappelle qu'un arbre pondéré est un arbre tel que :*

- *La somme des pondérations (ou probabilités) des noeuds issus d'un même sommet donne* 1.
- *La probabilité d'un chemin est le produit des probabilités des branches qui le composent.*
- *La pondération de la branche allant du sommet A vers le sommet B est la probabilité conditionnelle de B sachant A, $\mathbb{P}_A(B)$.*

La construction d'un arbre pondéré obéit au principe suivant : le noeud racine réprésente l'événement certain ω, et si un noeud représente un événement A, on construit autant de fils qu'il y a d'événements $A \cap E_n$, où (E_n) est une partition de Ω. La formule des probabilités composées est la traduction exacte de l'item 2. ci-dessus : la probabilité d'un noeud de l'arbre est la probabilité de l'intersection des événements qui composent le chemin qui le relie au sommet. La formule des probabilités totales dans ce contexte dit que la probabilité d'un événement est la somme des probabilités des feuilles de l'arbre qui concernent cet événement.

Exemple 2.2.1 *On dispose de deux urnes, l'une contenant* 1 *boule blanche et* 3 *noires, l'autre* 2 *blanches et* 2 *noires. On choisit uniformément au hasard parmi les deux urnes, puis on tire au hasard une boule dans l'urne choisie. On considére la probabilité de l'événement $A =$ « on tire une boule noire ». L'univers Ω est $\{(i,j),\ i = 1,2,\ j = 1,2,3,4\}$, où i est le numéro de l'urne et j celui de la boule. On considére la partition $\Omega = E_1 \cup E_2$, où $E_i = \{(i,j),\ j = 1,2,3,4\}$ qui est l'événement consistant à choisir l'urne i. Si on numérote en premier les boules noires, l'événement A s'écrit $\{(1,1),(1,2),(1,3),(2,1),(2,2)\}$. Pour tout i, la probabilité conditionnelle $\mathbb{P}_{E_i}$ est la probabilité uniforme sur E_i. Donc $\mathbb{P}_{E_1}(A) = 3/4$ et $\mathbb{P}_{E_2}(A) = 1/2$. D'après la formule des probabilités totales, $\mathbb{P}(A) = \mathbb{P}_{E_1}(A)\mathbb{P}(E_1) + \mathbb{P}_{E_2}(A)\mathbb{P}(E_2) = \frac{3}{4} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{5}{8}$.*

Proposition 2.2.5 (Formule de Bayes simple) *Soient A et B deux événements tels que $\mathbb{P}(A)$, $\mathbb{P}(B)$ et $\mathbb{P}(B^c)$ sont strictement positifs. On a :*

$$\mathbb{P}_A(B) = \frac{\mathbb{P}_B(A)\mathbb{P}(B)}{\mathbb{P}_B(A)\mathbb{P}(B) + \mathbb{P}_{B^c}(A)\mathbb{P}(B^c)}.$$

Démonstration. On a $\mathbb{P}_A(B) = \frac{\mathbb{P}(B\cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A\cap B)\mathbb{P}(B)}{\mathbb{P}(A)}$. On décompose ensuite le dénominateur par la formule des probabilités totales : $\mathbb{P}(A) = \mathbb{P}_B(A)\mathbb{P}(B) + \mathbb{P}_{B^c}(A)\mathbb{P}(B^c)$. □

A nouveau, cette proposition se généralise, à une famille formant une partition de Ω.

Proposition 2.2.6 (Formule de Bayes) *Si $(B_i, i \in I)$ est une famille qui forme une partition de Ω telle que pour tout $i \in I$, $\mathbb{P}(B_i) > 0$ alors, quelque soit l'événement A, si $\mathbb{P}(A) > 0$*

$$\mathbb{P}_A(B_i) = \frac{\mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}{\sum_{i\in I} \mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}.$$

Démonstration. Même argument que précédemment.

Les deux dernières formules permettent de calculer des probabilités conditionnelles $\mathbb{P}_A(B)$ connaissant les probabilités « inverses »$\mathbb{P}_B(A)$.

Exercice 2.2.3 Calculer la probabilité d'avoir tiré une boule dans la première urne sachant que c'est une noire.

Sous son apparente simplicité, la formule de Bayes est le départ d'une théorie entière appelée la statistique bayésienne.

Nous reprenons maintenant la question de l'exercice 2.2.2 concernant le test de maladie. La formule de Bayes va nous permettre de calculer la probabilité qu'un individu soit malade sachant qu'il a été testé positif. On note M l'événement « malade », T l'événement « test positif ». On a

$$\mathbb{P}_T(M) = \frac{\mathbb{P}_M(T)\mathbb{P}(M)}{\mathbb{P}_M(T)\mathbb{P}(M) + \mathbb{P}_{M^c}(T)\mathbb{P}(M^c)} = 0.16.$$

Résultat, le test n'est pas bon.

2.2.2 Indépendance

Nous passons maintenant à une notion fondamentale en probabilités : la notion d'indépendance. Si la connaissance d'un événement n'influe en rien sur la probabilité d'un autre, on dit que les événements sont indépendants. La notion de probabilité conditionnelle formalise cela. Si B est un événement avec probabilité strictement positive et $\mathbb{P}(A|B) = \mathbb{P}(A)$ alors A et B sont dit *indépendants*. Par définition des probabilités conditionnelles cela équivaut à la définition suivante :

Définition 2.2.2 *Deux événements A et B sont indépendants si*

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B).$$

Pour illustrer cette notion, on reprend l'exemple de contrôle de production.

Exemple 2.2.2 (Contrôle de production non destructif) *On a N objets, m défectueux. On tire au hasard n objets parmi les N en remettant chaque objet après l'avoir tiré. (on parle de tirage avec remise). Soit l' événement A_i :« l'objet tiré au ième tirage est défectueux ». Montrer que A_i et A_j sont indépendants pour tout $i \neq j$.*

Exemple 2.2.3 *Cette fois on ne remet pas l'objet tiré. (tirage sans remise). Calculer $\mathbb{P}(A_1 \cap A_2)$ et conclure que ce ne sont pas des événements indépendants*

Proposition 2.2.7 *Soient A et B deux événements indépendants alors A^c et B sont indépendants, ainsi que A et B^c et A^c et B^c.*

Démonstration. Laissée en exercice.

On généralise maintenant la notion d'indépendance à une famille d'événements :

Définition 2.2.3 (Indépendance mutuelle) *Soit une famille $(A_i, i \in I)$ d'événements. On dit qu'elle est formée d'événements (mutuellement) indépendants si pour toute sous famille finie $i_1, i_2, ..., i_p$ de I on a :*

$$\mathbb{P}\left(\bigcap_{k=1}^{p} A_{i_k}\right) = \prod_{k=1}^{p} \mathbb{P}(A_{i_k}).$$

Attention à bien noter que l'indépendance de A_i et A_j pour tout $i \neq j$ n'implique pas l'indépendance de la famille $(A_i, i \in I)$. Donnons un contre-exemple : $\Omega = \{1, 2, 3, 4\}$, muni de la probabilité uniforme. Considérer $A = \{1, 2\}$, $B = \{1, 3\}$, $C = \{2, 3\}$.

2.3 Variables aléatoires : notion de loi et de moments

On a vu dans la section précédente, les difficultés que le choix de l'univers pouvait impliquer. Dans la plupart des cas, on se concentre sur les événements que l'on souhaite étudier. Plus généralement, on définit la notion de variables aléatoires définies sur un espace Ω abstrait.
Une variable aléatoire est une **application** dont la valeur dépend du résultat de l'expérience aléatoire. On rappelle que $\mathcal{B}(\mathbb{R}^d)$ est la tribu des boréliens, c'est à dire la tribu engendrée par les pavés $\prod_{i=1}^n]a_i, b_i[$.

Définition 2.3.1 *Une variable aléatoire X est une application*

$$X : \omega \in \Omega \mapsto X(\omega) \in \mathbb{R}^d$$

telle que pour tout $A \in \mathcal{B}(\mathbb{R}^d)$

$$X^{-1}(A) = \{\omega \in \Omega; X(\omega) \in A\} \in \mathcal{F}. \tag{2.1}$$

Cela correspond à la notion de mesurabilité : X est une fonction $\mathcal{F}$-mesurable à valeurs dans l'espace euclidien. En probabilité, cela signifie que l'on peut étudier l'événement « X tombe dans $\prod_{i=1}^n [a_i, b_i]$ ». Pour tout $B \in \mathcal{B}(\mathbb{R}^d)$, on identifiera l'événement $X^{-1}(B)$ avec « $X \in B$ ».

On admet la proposition suivante issue de la théorie de la mesure.

Proposition 2.3.1 *$\mathcal{L}^0(\Omega, \mathcal{F})$ l'ensemble des variables aléatoires à valeurs dans l'espace euclidien $\mathbb{R}^d$ est un espace vectoriel.*

Exemple 2.3.1 (Fonction indicatrice) *Soit A un événement. On définit*

$$\mathbf{1}_A : \omega \mapsto \begin{cases} 1 \text{ si } \omega \in A \\ \\ 0 \text{ sinon.} \end{cases}$$

Exemple 2.3.2 *On lance une fléchette sur une cible et cherche à connaître la distance entre la fléchette et le centre de la cible. On pose $\Omega = \mathbb{R}^2$ et*

$$X : w = (x, y) \mapsto \sqrt{x^2 + y^2}.$$

X est une fonction continue (donc mesurable) et vérifie la condition 2.1.

Remarque 2.3.1 *Lorsque l'on peut prendre $\mathcal{F} = \mathcal{P}(\Omega)$ (par exemple, lorsque Ω est fini, et que l'on souhaite que tous les événements élémentaires soient dans la tribu) alors la condition 2.1 est toujours vraie.*

Proposition 2.3.2 *L'ensemble des variables aléatoires définies sur un même espace $(\Omega, \mathcal{F})$ forme une algèbre : en particulier la somme et le produit de deux variables aléatoires sont des variables aléatoires.*

2.3.1 Loi d'une variable aléatoire

Le plus souvent, (typiquement lorsque la variable aléatoire prend un nombre infini de valeurs), lorsque l'on étudie la variable aléatoire X, on n'étudie pas $X(\omega)$ en chaque ω, mais la probabilité avec laquelle X se *répartit* dans les boréliens. La variable aléatoire X est ainsi souvent caractérisée par sa *fonction de répartition* ou par sa *loi.*

Définition 2.3.2 *On appelle loi ou distribution d'une variable aléatoire X définie sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$, l'application*

$$P_X : B \in \mathcal{B}(\mathbb{R}^d) \mapsto \mathbb{P}(X^{-1}(B)).$$

On a le résultat fondamental suivant :

Proposition 2.3.3 *L'application P_X est une probabilité sur $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.*

Démonstration. A faire en cours.

Exemple 2.3.3 (jeu de dé) *On joue au jeu suivant : on lance un dé équilibré. On réalise un gain nul si on obtient* 1, 1 *euro si on obtient* 2, 3 *ou* 4 *et* 2 *euros si le résultat est* 5 *et* 4 *euros si le résultat est* 6. *On définit la variable aléatoire X donnant le montant du gain. Déterminer sa loi.*

Proposition 2.3.4 *Si X est une variable aléatoire sur $(\Omega, \mathcal{F})$. Soit g une fonction continue par morceaux sur $X(\Omega)$ alors l'application $g \circ X$ est une variable aléatoire. On la note souvent $g(X)$.*

2.3.2 Fonctions de répartition et indépendance

On se place dans le cas des variables aléatoires réelles : c'est à dire à valeurs dans $\mathbb{R}$. On note souvent v.a.r pour variable aléatoire réelle.

Définition 2.3.3 *On appelle fonction de répartition de X, l'application :*

$$F : x \in \mathbb{R} \mapsto \mathbb{P}(X \leq x).$$

Comme mentionné précédemment l'événement « $X \leq x$ » s'identifie à $X^{-1}(]-\infty, x])$.

Proposition 2.3.5 *Toute fonction de répartition vérifie les propriétés suivantes :*

1. *F est croissante*
2. $\lim_{x \to -\infty} F(x) = 0$ *et* $\lim_{x \to +\infty} F(x) = 1$
3. *F est continue à droite en tout point de $\mathbb{R}$*
4. *F a une limite à gauche en tout point x de $\mathbb{R}$, et*

$$\forall x \in \mathbb{R}, F(x) - F(x-) = \mathbb{P}[X = x]$$

   où $F(x-) = \lim_{t \to x; t < x} F(t)$.
5. *L'ensemble des points de discontinuité de F est au plus dénombrable.*

Remarque 2.3.2 *La fonction de répartition peut être définie pour des vecteurs aléatoires, en prenant la relation d'ordre* $(x_1, ..., x_p) \leq (x'_1, ..., x'_p)$ *si* $x_i \leq x'_i$ *pour* $i \in [|1, n|]$

Démonstration.
1. Soit $y \geq x$, $F(y) - F(x) = \mathbb{P}[x < X \leq y] \geq 0$

2. $F(-n) = \mathbb{P}(X \in]-\infty, n])$. Soit $A_n = X^{-1}(]-\infty, n])$, $(A_n, n \geq 1)$ est une suite décroissante d'événements. Par limite monotone : $\lim\limits_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}(\bigcap_{k=1}^{\infty} A_k)$, or $\bigcap_{k=1}^{\infty} A_k = \emptyset$ donc

$$\lim \mathbb{P}(A_n) = \lim F(-n) = 0.$$

   F étant une fonction monotone, on a $\lim_{x\to-\infty} F(x) = \lim_{n\to-\infty} F(n)$.

3. Comme F est monotone (croissante), pour montrer que F est continue à droite, il suffit de montrer que $F(x + \frac{1}{n}) \underset{n\to\infty}{\longrightarrow} F(x)$. La suite d'événements $A_n := X^{-1}\left(]-\infty, x + \frac{1}{n}]\right)$ est décroissante, donc la suite $(\mathbb{P}(A_n), n \geq 1)$ converge et

$$\lim_{n\to\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \mathbb{P}(X^{-1}(]-\infty, x]) = F(x).$$

4. A nouveau, on étudie $F\left(x - \frac{1}{n}\right)$, on cherche à montrer que cette suite a une limite quand n tend vers $+\infty$. On note $A_n = X^{-1}\left(]-\infty, x - \frac{1}{n}]\right)$. La suite $(A_n, n \geq 1)$ est croissante donc $\mathbb{P}(A_n, n \geq 1)$ converge et sa limite est

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right).$$

   On a $\bigcup_{n=1}^{\infty} A_n = X^{-1}\left(]-\infty, x[\right)$, par suite

$$\lim_{n\to\infty} F(x - \frac{1}{n}) = \mathbb{P}(X < x).$$

   La fonction F a donc une limite à gauche en x, elle vaut $\mathbb{P}(X < x)$. De plus on a

$$F(x) - \lim_{t\to x^-} F(t) = \mathbb{P}(X \leq x) - \mathbb{P}(X < x) = \mathbb{P}(X = x).$$

5. On note $D_n := \{x_0 \text{ tel que } \lim\limits_{x\to x_0^+} F(x) - F(x_0) \geq \frac{1}{n}\}$. Comme F est croissante et comprise entre 0 et 1, D_n contient au plus n éléments. Soit D l'ensemble des points de discontinuité, $D = \bigcup_{n=1}^{\infty} D_n$, donc D est une union dénombrable d'ensembles finis. □

Le théorème suivant nous dit que si les propriétés 1, 2, 3 sont vérifiées alors la fonction F est la fonction de répartition d'une variable aléatoire réelle.

Théorème 2.3.6 *Toute fonction* $F : \mathbb{R} \mapsto [0, 1]$ *telle que*

1. F est croissante

2. $\lim_{x\to-\infty} F(x) = 0$ *et* $\lim_{x\to+\infty} F(x) = 1$

3. F *est continue à droite en tout point de* $\mathbb{R}$

est la fonction de répartition d'une variable aléatoire réelle.

Proposition 2.3.7 *La fonction de répartition de* X *caractérise sa loi.*

Démonstration. Pour tout $x < y$, $P_X(]x, y]) = F(y) - F(x)$, donc F donne P_X sur les parties du types $]x, y]$, comme ces parties engendrent la tribu borélienne, cela caractérise P_X sur tout $\mathcal{B}(\mathbb{R})$.

Le terme variable aléatoire est particulièrement utilisé pour des variables réelles. Lorsque X est à valeurs dans $\mathbb{R}^d$, on parle plus couramment de vecteurs aléatoires.
Lorsque $d \geq 2$, la définition générale (Définition 2.3.1) correspond à la définition d'un vecteur aléatoire. De façon équivalente, un vecteur aléatoire est une application

$$X : \omega \in \Omega \mapsto \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ . \\ . \\ . \\ X_n(\omega) \end{pmatrix}$$

où chaque X_i est une v.a.r.

On introduit maintenant **l'indépendance des variables aléatoires**.

Définition 2.3.4 *Deux variables aléatoires* X *et* Y *sont indépendantes si* $\mathbb{P}(X \in A$ *et* $Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ *pour tous* A, B *boréliens.*

Définition 2.3.5 (indépendance mutuelle) $(X_i, i \in I)$ *est une famille de variables aléatoires indépendantes si pour toute sous-famille* $\{i_1, ..., i_p\}$ *finie de* I*, on a, pour toute famille de boréliens* $(B_k)_{1\leq k\leq p}$

$$\mathbb{P}\left(\bigcap_{k=1}^{p}(X_{i_k} \in B_k)\right) = \prod_{k=1}^{p}\mathbb{P}(X_{i_k} \in B_k).$$

Exemple fondamental : si A et B sont deux événements indépendants alors $\mathbf{1}_A$ et $\mathbf{1}_B$ sont deux variables aléatoires indépendantes.

L'indépendance correspond à une structure produit « cachée ». Pour construire deux variables indépendantes sur un même espace probabilisé. On peut munir le produit cartésien $X(\Omega)\times Y(\Omega)$ de la probabilité produit $P_X \otimes P_Y$ définie par

$$P_X \otimes P_Y(]a, b]\times]c, d]) = P_X(]a, b])P_Y(]c, d]) \text{ pour tout } a < b, c < d.$$

Critère d'indépendance : Si

$$X : \omega \in \Omega \mapsto \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ . \\ . \\ . \\ X_n(\omega) \end{pmatrix}$$

alors $X_1, ..., X_n$ sont indépendantes si et seulement si $P_{(X_1,...,X_n)} = P_{X_1} \otimes P_{X_2} \otimes ... \otimes P_{X_n}$.

2.4 Exercices

Exercice 2.4.1 (Univers) *Proposer un univers Ω pour chacune des expériences suivantes :*
1) On lance deux dés. Etant donné que le premier dé donne 3, quelle est la probabilité que le total des deux dés dépasse 6 ?
2) On lance un dé rouge et un dé bleu et on observe les points donnés par chacun d'eux
3) On tire une pièce plusieurs fois jusqu'à obtenir face pour la première fois. On observe combien de lancers on doit faire. Soit A :« on obtient face après un nombre pair de lancers ». A est il un événement ?

Exercice 2.4.2 Exprimer chacun des événements suivants à l'aide des opérations de complémentation, union et intersection :
« A et B se réalisent, mais pas C » ;
« au moins un des trois événements A , B, C se réalise » ;
« au plus un des trois événements A , B, C se réalise » ;
« aucun des trois événements A , B, C ne se réalise » ;
« les trois événements A , B, C se réalisent ».

Exercice 2.4.3 (#) *Trois usines pharmaceutiques A, B, et C produisent respectivement 40%, 35% et 25% du nombre total de comprimés achetés par un grossiste. Chacune de ces usines produit respectivement 5, 6, et 3% de comprimés défectueux. Le qualiticien de l'entreprise reçoit une nouvelle livraison.*

1. *Déterminer les probabilités des différentes possibilités suivantes : provenir de A et être défectueux, provenir de A et être conforme, provenir de B et être défectueux, provenir de B et être conforme, provenir de C et être défectueux, provenir de C et être conforme.*
2. *Dans cette livraison, on prend un comprimé au hasard. Quelle est la probabilité p_1 pour qu'il soit défectueux ? Quelle est la probabilité p_2 pour qu'il soit conforme ?*
3. *Dans cette livraison, on prend un comprimé au hasard, on constate qu'il est défectueux. Quelle est la probabilité qu'il ait été fabriqué dans l'usine A ?*

Solution. Reécrivons l'énoncé "mathématiquement". Notons D (resp. D^c) l'événement "le comprimé est défectueux (resp. non défectueux)". On a $\mathbb{P}(A) = 40/100$, $\mathbb{P}(B) = 35/100$ et $\mathbb{P}(C) = 25/100$. De plus, $\mathbb{P}(D|A) = 5/100$, $\mathbb{P}(D|B) = 6/100$ et $\mathbb{P}(D|C) = 3/100$.

1. $\mathbb{P}(A \cap D) = \mathbb{P}(D|A)\mathbb{P}(A) = 2/100$; $\mathbb{P}(A \cap D^c) = \mathbb{P}(A) - \mathbb{P}(A \cap D) = 38/100$.
   $\mathbb{P}(B \cap D) = 2,1\%$; $\mathbb{P}(B \cap D^c) = 32,9\%$.
   $\mathbb{P}(C \cap D) = 0.75\%$; $\mathbb{P}(C \cap D^c) = 24,25\%$.
2. $p_1 = \mathbb{P}(D) = \mathbb{P}(A \cap D) + \mathbb{P}(B \cap D) + \mathbb{P}(C \cap D) = 4,85\%$; $\mathbb{P}(D^c) = 95,15\%$.
3. $p_2 = \mathbb{P}(A|D) = \mathbb{P}(A \cap D)/\mathbb{P}(D) \simeq 41\%$.

Exercice 2.4.4 *Une urne contient cinq boules roules et trois boules noires. On tire au hasard une boule.*
Si cette boule est noire, on arrête le jeu. Si cette boule est rouge, on replace la boule dans le sac, on ajoute deux boules rouges et on procède à un deuxième tirage.

1) Représenter l'expérience à l'aide d'un arbre pondéré.
2) Quelle est la probabilité de tirer deux boules rouges ?

Exercice 2.4.5 (Allèle) *On suppose qu'un gène est formé de deux allèles A et a. Un individu peut avoir l'un des trois génotypes AA, Aa, aa.*
Un enfant hérite d'un allèle de chacun de ses parents, chacun d'eux étant choisi au hasard. Par exemple si le père est de type Aa, et la mère AA, alors les enfants peuvent être du types AA ou Aa.
On considère une population dont les proportions des génotypes pour les hommes comme pour les femmes sont de p_0 pour AA, q_0 pour Aa, r_0 pour aa.
1) On suppose que le génotype de l'enfant est formé au hasard, on le note (x, y) avec x le génotype de son père et y celui de sa mère.
  a) Quelle est la probabilité qu'un enfant ait deux parents de type AA ?
  b) un parent de type AA et un parent de type Aa ?
  c) deux parents de type Aa ?
2) quelle est la probabilité que l'enfant soit de type AA sachant que :
  a) les deux parents sont de type AA ?
  b) un des deux parents est de type AA et l'autre de type Aa ?
  c) les deux parents sont de types Aa ?
3)-a Calculer la probabilité p_1 pour que l'enfant soit de type AA
b) calculer la probabilité r_1 pour que l'enfant soit de type aa
c) calculer la probabilité q_1 que l'enfant soit de type Aa

Exercice 2.4.6 (Rats) *On fait une expérience sur le comportement des rats. Ils doivent choisir entre 4 portes d'apparence identique. En réalité une seule porte est la bonne (et permet au rat de sortir), les deux autres portes ramènent le rat a son point de départ. L'expérience a lieu jusqu'à ce que le rat sorte. On suppose qu'il évite les portes choisies auparavant et qu'il choisit de façon équiprobable entre celles qu'il n'a pas encore ouvertes.*
Soient les événements :

S_1= *« le rat sort la première fois. »*

S_2= *« le rat sort la deuxième fois. »*

S_3= *« le rat sort la troisième fois. »*

S_4= *« le rat sort la quatrième fois. »*

Construire un arbre pondéré et donner la probabilité de chacun des événements.

Exercice 2.4.7 (Kids) *Une famille a deux enfants. On cherche la probabilité que les deux enfants soient des garçons, sachant qu'au moins est un garçon ?*
1) Prenons en compte l'âge des enfants : le plus jeune et le plus âgé peuvent être chacun des garçons. Donner l'univers. On suppose que chaque configuration a la même probabilité.
2) Calculer la probabilité.
3) Sachant maintenant que le plus jeune enfant est un garçon, quelle est la probabilité que les deux soient des garçons ?

Exercice 2.4.8 (Urnes #) *On considère deux urnes. Chaque urne contient des boules colorées. La première urne contient 2 boules blanches et 3 boules noires. La deuxième urne contient 3 blanches et 4 noires. On tire une boule au hasard dans la première urne et on la place dans la deuxième. On suppose que toutes les boules ont la même chance d'être tirée (rappeler le nom de cette hypothèse). On tire alors au hasard une boule dans la deuxième urne et on l'examine. Quelle est la probabilité que la boule soit noire ?*

Correction. On note les urnes I et II. L'événement qui nous intéresse est

$$A = \text{ « la boule tirée dans II est noire » }.$$

On va utiliser la formule des probabilités totales. Soit B l'événement, « on a tiré une boule noire dans l'urne I »Ḋ'après la formule des probabilités totales, on a :

$$\mathbb{P}[A] = \mathbb{P}[A|B]\mathbb{P}[B] + \mathbb{P}[A|B^c]\mathbb{P}[B^c].$$

Sachant qu'on a tiré une boule noire dans la première urne (notée I) ; la probabilité que l'on tire une noire dans la deuxième (notée II) est $\mathbb{P}[A|B] = \frac{\mathbb{P}[A\cap B]}{\mathbb{P}[B]} = \frac{5}{8}$
$\mathbb{P}[A|B^c] = \frac{4}{8} = \frac{1}{2}$ et $\mathbb{P}[B] = \frac{3}{5}$. Donc,

$$\mathbb{P}[A] = \frac{5}{8}.\frac{3}{5} + \frac{1}{2}\frac{2}{5} = \frac{23}{40}.$$

Exercice 2.4.9 Un domino est un rectangle comportant deux parties $[i|j]$. Sur chacune de ces parties figurent entre 0 et 6 points. Les dominos $[i|j]$ et $[j|i]$ ne sont pas distingués.

1. Combien de dominos différents peut-on ainsi obtenir ?
2. On suppose que l'on a un jeu complet de dominos. On prend l'un de ces dominos au hasard.
   (a) Quelle est la probabilité que ce soit un double ?
   (b) Quelle est la probabilité que la somme des points soit 6 ?
   (c) Quelle est la probabilité que ce soit un double si l'on sait que la somme des points vaut 6 ?
3. On retire du jeu les dominos dont la somme des points est 6, et on en tire un parmi les dominos restants. Quelle est la probabilité que ce soit un double ?
4. On dit que des dominos sont amis si l'une de leurs parties a le même nombre de points. A partir du jeu complet, on tire deux dominos. Quelle est la probabilité qu'ils soient amis ?

Exercice 2.4.10 (Poker #) *On joue au poker avec un jeu de 32 cartes. Il y a donc*
– *4 couleurs : carreau, trèfle, pique et coeur.*
– *8 hauteurs : 7, 8, 9, 10, valet, dame, roi, as.*

Chaque joueur reçoit 5 cartes. Quelle est la probabilité qu'un joueur ait :

1) un carré : quatre cartes de même hauteur
2) un brelan : trois cartes de même hauteur
3) une paire : deux cartes de même hauteur
4) deux paires

5) un full (brelan+ paire) ?

Solution. L'univers Ω est l'ensemble des mains possibles, c'est-à-dire l'ensemble des combinaisons de 5 éléments parmi 32. On suppose que toutes les mains sont équiprobables, autrement dit on munit Ω de la probabilité uniforme $\mathbb{P}(A) = \frac{\text{card }(A)}{\text{card }(\Omega)}$.

1) Déterminons le cardinal de l'ensemble correspondant à l'événement $A =$« le joueur possède un carré. »
  - Choix de la hauteur : $\binom{8}{1} = 8$ (on choisit un élément parmi 8)
  - Choix de la couleur : $\binom{4}{4} = 1$ (on choisit une couleur différente pour chaque carte : deux cartes de même hauteur et de même couleur sont identiques !)
  - Choix de la hauteur de la 5ème carte : $\binom{7}{1} = 7$: il n'y a que 4 cartes de chaque hauteur, une fois choisie la hauteur du carré, il reste seulement 7 hauteurs potentielles
  - Choix de la couleur de l a 5ème carte : $\binom{4}{1} = 4$.

$$\text{On a donc } \mathbb{P}(A) = \frac{8\times1\times7\times4}{\binom{32}{5}} = \frac{224}{201376} \approx 0,00111.$$

2) Soit l'événement $A =$« le joueur possède un brelan. »
  - Choix de la hauteur des trois cartes : $\binom{8}{1} = 8$ (on choisit un élément parmi 8)
  - Choix de la couleur : $\binom{4}{3} = 4$ (on choisit trois couleurs différentes pour chaque carte : deux cartes de même hauteur et de même couleur sont identiques !)
  - Choix de la hauteur des cartes restantes : $\binom{7}{2}$: (les deux cartes restantes sont d'une hauteur différente, sinon nous aurions un full).
  - Choix de la couleur des cartes restantes : 4^2.

$$\text{On a donc } \mathbb{P}(A) = \frac{8\times4\times\binom{7}{2}\times4^2}{\binom{32}{5}}.$$

3) Soit l'événement $A =$« le joueur possède une paire. »
  - Choix des hauteurs : $\binom{8}{1}$ (on choisit 1 élément parmi 8)
  - Choix des couleurs : $\binom{4}{2}$ (on choisit deux couleurs différentes pour chaque carte)
  - Choix des hauteurs des cartes restantes : $\binom{7}{3}$: (les cartes restantes sont d'une hauteur différente, sinon nous aurions au moins un brelan).
  - Choix de la couleur des cartes restantes : 4^3.

$$\text{On a donc } \mathbb{P}(A) = \frac{\binom{8}{1}\times\binom{4}{2}\times\binom{7}{3}\times4^3}{\binom{32}{5}}.$$

4) Soit l'événement $A =$« le joueur possède deux paires. »
  - Choix des hauteurs : $\binom{8}{2}$ (on choisit deux éléments parmi 8)
  - Choix des couleurs : $\binom{4}{2}^2$ (on choisit deux couleurs différentes pour chaque paire)
  - Choix de la hauteur de la carte restante : $\binom{6}{1}$: (la carte restante est d'une hauteur différente, sinon nous aurions un full).
  - Choix de la couleur de la carte restante : 4.

$$\text{On a donc } \mathbb{P}(A) = \frac{\binom{8}{2}\times\binom{4}{2}^2\times\binom{6}{1}\times4}{\binom{32}{5}}.$$

5) Soit l'événement $A =$« le joueur possède un full. »
  - Choix des hauteurs : $\binom{8}{1}$ (on choisit un élément parmi 8)
  - Choix des couleurs : $\binom{4}{3}$
  - Choix de la hauteur des cartes restantes : $\binom{7}{1}$: (les cartes restantes sont d'une hauteur différente, sinon nous aurions un carré).
  - Choix des couleurs de la paire $\binom{4}{2}$.

$$\text{On a donc } \mathbb{P}(A) = \frac{\binom{8}{1}\times\binom{4}{3}\times\binom{7}{1}\times\binom{4}{2}}{\binom{32}{5}}.$$

Exercice 2.4.11 *Les coefficients a, b, c de l'équation quadratique $ax^2 + bx + c = 0$ sont détermines en lançant trois fois un de équilibré. Quelle est la probabilité que les racines soient réelles ? Complexes non réelles ?*

Exercice 2.4.12 Un antivirus assure avec une fiabilité de $a = 95\%$ la détection d'un malware M lorsqu'il est effectivement présent. Cependant, le test indique aussi un résultat faussement "positif" pour $b = 1\%$ des systèmes réellement sains à qui on l'applique. On suppose qu 'une proportion $p = 0.5\%$ des systèmes ont le malware M.
1) En considérant que les paramètres a, b et p sont des probabilités, déterminer la probabilité qu'un système soit vraiment atteint sachant qu'il a un test positif ?
2) On suppose maintenant que a, b et p sont définis comme ci dessus mais qu'on ne connait plus leur valeur numérique. A quelles conditions sur ces valeurs le test est-il utile ?

Exercice 2.4.13 Pour lutter contre les spams, les messageries ont mis en place des techniques en repérant certains mots ou expressions[2] pouvant figurer dans les emails : par exemple

$$\{\text{merveilleux ; gratuit ; argent ; vous avez gagné ; cadeau}\}.$$

On dit qu'il y a *détection* si cette famille de mots est présente dans le mail. Le filtre fonctionne alors de la façon suivante :

Si **sachant qu'il y a détection**, la probabilité d'être un spam dépasse le seuil de 90%, l'email est déplacé dans les courriers indésirables. Sinon il n'est pas déplacé.

1) On suppose qu'il n'y a qu'un seul mot dans la famille. Soient les événements D=« détection » et S=« spam ». Redémontrer la formule de Bayes

$$\mathbb{P}_D[S] = \frac{\mathbb{P}_S(D)\mathbb{P}(S)}{\mathbb{P}_S(D)\mathbb{P}(S) + \mathbb{P}_{\bar{S}}(D)\mathbb{P}(\bar{S})}.$$

2) Lors de la phase d'apprentissage (lorsque vous indiquez à votre messagerie qu'un email est indésirable) ; l'anti-spam compte le nombre de spams contenant le mot. On suppose qu'il y a 70% des spams qui contiennent le mot, et 15% des non-spams qui le contiennent. En approchant les probabilités par ces quantités que valent $\mathbb{P}_S(D)$ et $\mathbb{P}_{\bar{S}}(D)$?
On suppose $\mathbb{P}(S) = 0.7$. Dans ces conditions si vous recevez un email contenant ce mot sera-t-il déplacé dans les courriers indésirables ? L'est-il si $\mathbb{P}(S) = 0.5$?

3) On considère maintenant une famille de n mots. Pour tout $i \in \{1, .., n\}$, on appelle M_i l'événement « le i-ème mot est présent ». On suppose que les événements $(S \cap M_i)_{i\geq 1}$ sont indépendants, ainsi que les événements $(\bar{S} \cap M_i)_{i\geq 1}$. L'événement de détection D correspond à avoir tous les mots dans le mail : $M_1 \cap M_2 \cap ... \cap M_n$. On note $p_i = \mathbb{P}_S(M_i)$, $q_i = \mathbb{P}_{\bar{S}}(M_i)$ et $s = \mathbb{P}(S)$, démontrer que :

$$\mathbb{P}_D[S] = \frac{s^n p_1 p_2 ... p_n}{s^n p_1 p_2 ... p_n + (1-s)^n q_1 q_2 ... q_n}.$$

Indication : remarquer que $M_1 \cap M_2 \cap S = M_1 \cap S \cap M_2 \cap S$.

2. A plan for spam http ://www.paulgraham.com/spam.html, voir aussi *filtrage bayésien anti-spam dans wikipedia* (qui contient des erreurs de maths !)

4) On suppose qu'il y a les 5 mots et expressions suivants dans notre filtre.

Mot :	merveilleux	argent	gratuit	cadeau	vous avez gagné
parmi les spams :	70%	75%	40%	60%	85%
parmi les non-spams :	30%	15%	10%	60%	2%

Si $s = 0.5$, un mail contenant ces 5 mots sera-t'il déplacé en indésirable ? Qu'en est-il s'il contient seulement les mots argent et gratuit ?

Exercice 2.4.14 *Montrer que la probabilité qu'exactement un des événements A et B se réalise est*

$$\mathbb{P}[A] + \mathbb{P}[B] - 2\mathbb{P}[A \cap B]$$

Exercice 2.4.15 *Montrer que*

$$\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) \geq \sum_{i=1}^{n} \mathbb{P}(A_i) - \sum_{1 \leq i < k \leq n} \mathbb{P}(A_i \cap A_k)$$

Exercice 2.4.16 *Calculer la probabilité qu'une main de 13 cartes tirée à partir d'un paquet de 52 cartes contienne exactement deux rois et un as. Quelle est la probabilité qu'elle contienne exactement un as sachant qu'elle contient un roi ?*
On rappelle qu'il y a 4 rois et 4 as dans un paquet.

Exercice 2.4.17 *On tire un dé deux fois. Calculer les probabilités des événements suivants*
a) Un 6 est réalisé exactement une fois.
b) Les deux numéros sont impairs
c) la somme des numéros est 4
d) la somme des numéros est divisible par 3 *?*

Exercice 2.4.18 *On dispose de deux pièces A et B.*
– *La pièce A est équilibrée, au sens où elle donne face et pile avec probabilité $\frac{1}{2}$.*
– *La pièce B donne face avec probabilité p et pile avec probabilité $1 - p$*
On effectue une succession de lancers selon le procédé suivant :
– *On choisit une des deux pièces A, B au hasard, on la lance*
– *A chaque lancer, si on obtient face, on garde la pièce pour le lancer suivant, sinon on change de pièce.*
Pour tout entier naturel, on note A_n l'événement

« le n-ième lancer se fait avec la pièce A »

Soit $a_n = \mathbb{P}[A_n]$, étudier la suite $(a_n, n \geq 1)$ et commenter vos résultats.

Exercice 2.4.19 On dispose de deux dés, l'un honnête , et l'autre pipé pour lequel la probabilité d'obtenir 6 est 1/3 tandis que les autres faces ont toutes la même probabilité.
1) On choisit un dé au hasard et on le lance. Quelle est la probabilité de faire 6 ?
2) On choisit un dé au hasard et on le lance 4 fois. Quelle est la probabilité d'obtenir au moins une fois 6 ?
3) On joue 4 fois en choisissant chaque fois au hasard l'un des dés. Quelle est la probabilité d'obtenir au moins une fois 6 ?

Exercice 2.4.20 On note Ω l'ensemble des huit résultats de trois lancers successifs d'une pièce de monnaie et on considère les deux événements suivants : A="le premier jet donne un pile", B="pile est amené au moins deux fois". Si on suppose que tous les éléments de Ω sont équibrobables, A et B sont-ils indépendants ?

Exercice 2.4.21 *Montrer que si F et G sont des fonctions de répartition alors pour tout $\lambda \in (0,1)$, $\lambda F + (1-\lambda)G$ est une fonction de répartition.*

Exercice 2.4.22 *Soit X une variable aléatoire réelle. Donner les fonctions de répartitions des valeurs aléatoires suivantes :*

$$X^+ = \max\{0, X\}, X^- = -\min\{0, X\}, |X| = X^+ + X^-, -X$$

Exercice 2.4.23 *On appelle* ***médiane****, un réel m tel que*

$$\lim_{y \uparrow m} F(y) \leq \frac{1}{2} \leq F(m).$$

Montrer que toute fonction de répartition a au moins une médiane. Montrer que l'ensemble des médianes est un intervalle fermé de $\mathbb{R}$

Exercice 2.4.24 *Soient X et Y deux variables aléatoires réelles de fonctions de répartition F et G définies par :* $F(x) = \begin{cases} 0 \ si \ x < -1 \\ \frac{1}{3} \ si \ -1 \leq x < 2. \\ 1 \ si \ x \geq 2 \end{cases}$ $G(x) = \begin{cases} 0 \ si \ x \leq 0 \\ x^2 \ si \ 0 \leq x \leq \frac{1}{2}. \\ 1 - \frac{3}{4}e^{1/2-x} \ si \ x \geq \frac{1}{2} \end{cases}$

1) Déterminer les réels x et y tels que $\mathbb{P}[X = x] > 0$, et $\mathbb{P}[Y = y] > 0$.
2) Déterminer la fonction de répartition de $aX + b$ où $a > 0$.
3) On suppose que X et Y sont indépendantes. Déterminer les fonctions de répartitions de $max(X, Y)$ et $min(X, Y)$.

Exercice 2.4.25 *Soient $\mathcal{F}$ et $\mathcal{G}$ deux tribus.*
1) Soit $A \in \mathcal{P}(\Omega)$, montrer que $\{\emptyset, \Omega, A, A^c\}$ est une tribu.
2) Soient A et B des éléments de $\mathcal{F}$, montrer que $A \cap B$, $A \setminus B$ appartiennent à $\mathcal{F}$.
3) Montrer que $\mathcal{F} \cap \mathcal{G}$ est une tribu.
4) Montrer que $\mathcal{F} \cup \mathcal{G}$ n'est pas forcément une tribu.

Chapitre 3

Variables aléatoires discrètes

Dans ce chapitre, nous étudions les probabilités *discrètes*. Les exemples étudiés précédemment sont pour la majorité des exemples discrets. Nous avons déjà vu que le dénombrement joue une rôle important.

3.1 Généralités

Définition 3.1.1 *Une v.a.r X est discrète si $X(\Omega)$ est au plus dénombrable.*

Remarque 3.1.1 *Cela ne signifie pas forcément que X est à valeurs dans les entiers. La variable peut prendre ses valeurs dans un sous-ensemble dénombrable de $\mathbb{R}$ tel que $\mathbb{Q}$.*

3.1.1 Loi et fonction de répartition

Soit X une variable discrète, $X(\Omega) = \{x_i, i \in I\}$. La loi de X, P_X est déterminée par

$$P_X(B) = \mathbb{P}(X \in B) = \mathbb{P}\left(\bigcup_{i;x_i \in B} \{X = x_i\}\right) = \sum_{i;x_i \in B} \mathbb{P}(X = x_i) \text{ avec } B \subset X(\Omega)$$

. Finalement la loi d'une variable discrète est déterminée par l'ensemble des valeurs prises $X(\Omega)$ et par les valeurs $\mathbb{P}(X = x_i)$. Il est important de noter que l'on a toujours

$$\sum_{i \in I} \mathbb{P}(X = x_i) = 1.$$

Inversement, on a la propriété suivante :

Proposition 3.1.1 *Soit I une partie non vide de $\mathbb{N}$. Soient $(x_i, i \in I)$ une famille de réels et $(p_i, i \in I)$ une famille de réels positifs vérifiant $\sum_{i \in I} p_i = 1$. On peut définir la variable X avec :*
$X(\Omega) = \{x_i, i \in I\}$
$\forall i \in I$, $\mathbb{P}(X = x_i) = p_i$.

Exemple 3.1.1 *Soit X une variable aléatoire à valeurs dans $\mathbb{N}$ telle que*

$$\forall n \in \mathbb{N}, \mathbb{P}(X = n) = e^{-1}\frac{1}{n!}$$

Montrer que cela définit bien une loi.

Lorsque X est discrète, la fonction de répartition est une fonction en escalier : sur chaque intervalle $]x_i, x_{i+1}[$ la fonction $F : x \mapsto \mathbb{P}(X \leq x)$ est constante.

3.1.2 Espérance

La notion d'espérance correspond à celle de moyenne.

Définition 3.1.2 *Soit X une v.a.r discrète. Si $X(\Omega) = \{x_i, i \in I\}$ avec I dénombrable, et si $\sum_{i\in I} |x_i|\mathbb{P}[X = x_i] < \infty$ alors l'espérance de X est définie par*

$$\mathbb{E}(X) := \sum_{i\in I} x_i\mathbb{P}[X = x_i].$$

Dans le cas où I est fini, l'hypothèse de convergence de la série est toujours vérifiée. Dans le cas infini dénombrable, on suppose que la série de terme général $x_i\mathbb{P}[X = x_i]$ est *absolument convergente.* L'absolue convergence de la série garantie que l'espérance de X ne dépend pas de l'ordre des x_i. On rappelle qu'il existe des séries convergentes, non absolument convergentes, dont la somme dépend de l'ordre pris sur ses termes. Penser à $\sum_{n\in\mathbb{Z}} \frac{1}{n}$ ou $\sum_{n\geq 1} \frac{(-1)^n}{n}$ qui sont non absolument convergentes. Il faut bien comprendre que l'espérance n'existe pas toujours, même dans le cas discret ! Si la variable aléatoire discrète est à valeurs positives alors son espérance existe ou est infinie.

Exemple 3.1.2 (Paradoxe de Saint Petersbourg) *On tire une pièce, si elle tombe sur face le jeu s'arrête, si elle tombe sur pile le joueur gagne deux fois sa mise et rejoue. On continue ainsi de suite en doublant le gain du joueur à chaque fois qu'il rejoue. Imaginons qu'il ait misé 1 euro. Quelle est la loi de son gain ? Quelle mise maximale le joueur doit il accepter pour jouer ?*
On note X son gain. On a $\mathbb{P}(X = 2^n) = \frac{1}{2^n}$ pour $n \in \mathbb{N}^\star$. On a bien

$$\sum_{n=1}^{\infty} \mathbb{P}(X = 2^n) = 1.$$

On remarque tout d'abord que la loi ne dépend pas de la mise de départ. De plus on a $\sum_{n\geq 1} 2^n\mathbb{P}(X = 2^n] = \sum_{n\geq 1} 1 = \infty$. Finalement, la moyenne du gain étant infinie, le joueur devrait a priori toujours jouer.

Proposition 3.1.2 *Si $X(\Omega)$ possède un maximum et un minimum, alors $\mathbb{E}(X)$ existe et*

$$x_{min} \leq \mathbb{E}[X] \leq x_{max}$$

Démonstration. A faire.

Théorème 3.1.3 *Soit Ω fini ou dénombrable, si la série est absolument convergente, on a*

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

Démonstration. On note $A_i = \{\omega \in \Omega; X(\omega)x_i\}$. On a

$$\mathbb{P}(X = x_i) = \mathbb{P}(X^{-1}(\{x_i\}) = \mathbb{P}(A_i) = \sum_{\omega\in A_i} \mathbb{P}(\{\omega\})$$

On a donc

$$\begin{aligned}\sum_{i\in I} x_i \mathbb{P}(X = x_i) &= \sum_i \sum_{\omega\in A_i} \mathbb{P}(\{\omega\}) \\ &= \sum_i \sum_{\omega\in A_i} x_i \mathbb{P}(\{\omega\}) \\ &= \sum_i \sum_{\omega\in A_i} X(\omega)\mathbb{P}(\{\omega\}) \\ &= \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\})\end{aligned}$$

□

Ce théorème est essentiel du point de vue théorique. Jusqu'à maintenant le rôle de l'univers était quelque peu caché puisque nous travaillions directement sur l'espace probabilisé $(X(\Omega), \mathcal{B}(\mathbb{R}), P_X)$. Dans la plupart des cas, les calculs sont plus simples sur ce triplet que sur celui de l'univers. Néanmoins le théorème 3.1.3 permet de travailler avec différentes variables aléatoires simultanément.

Proposition 3.1.4 (linéarité) *Soient X et Y deux variables aléatoires discrètes sur Ω. On suppose qu'elles possèdent une espérance. Soit $\lambda \in \mathbb{R}$. Alors, la variable aléatoire $X + \lambda Y$ est discrète et a pour espérance :*

$$\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda\mathbb{E}[Y].$$

Démonstration. La série $\sum_{\omega\in\Omega}(X(\omega)+\lambda Y(\omega))\mathbb{P}(\{\omega\})$ est absolument convergente car c'est une somme de deux séries absolument convergentes (utiliser l'inégalité triangulaire). On a donc :

$$\sum_{\omega\in\Omega}(X(\omega) + \lambda Y(\omega))\mathbb{P}(\{\omega\}) = \sum_{\omega\in\Omega} X(\omega)\mathbb{P}(\{\omega\}) + \lambda \sum_{\omega\in\Omega} Y(\omega)\mathbb{P}(\{\omega\}) = \mathbb{E}[X] + \lambda\mathbb{E}[Y].$$

□

Le théorème 3.1.3 et la propriété 3.1.4 restent vrais pour toute variable aléatoire, pas forcément discrète, définie sur n'importe quel univers Ω. Du point de vue de *la théorie de la mesure*, l'espérance n'est rien d'autre que l'intégrale de X contre la mesure $\mathbb{P}$:

$$"\mathbb{E}[X] = \int_\Omega X(\omega)\mathbb{P}(\mathrm{d}\omega)."$$

Proposition 3.1.5 *Soient X et Y deux variables aléatoires. Supposons que $0 \le X \le Y$ et que Y admet un moment d'ordre 1. Alors X admet un moment d'ordre 1 et*

$$\mathbb{E}[X] \le \mathbb{E}[Y].$$

Démonstration. Laissée en exercice.

3.1.3 Composition d'une variable aléatoire et d'une fonction et indépendance

Soit X une variable aléatoire discrète. Soit g une fonction continue par morceaux définie sur un ensemble J contenant $X(\Omega)$, clairement la variable $g \circ X$ est une variable discrète : on a $(g \circ X)(\Omega) = \{g(x_i), i \in I\}$. Le théorème suivant est très important.

Théorème 3.1.6 (théorème de transfert)

$$\mathbb{E}[g(X)] = \sum_{i \in I} g(x_i)\mathbb{P}(X = x_i)$$

Démonstration. On note $A_i = \{\omega \in \Omega; X(\omega) = x_i\}$.

$$\begin{aligned}\mathbb{E}[g(X)] &= \sum_{\omega \in \Omega} g(X(\omega))\mathbb{P}(\{\omega\}) = \sum_{i \in I} \sum_{\omega \in A_i} g(x_i)\mathbb{P}(\{\omega\}) \\ &= \sum_{i \in I} g(x_i) \sum_{\omega \in A_i} \mathbb{P}(\{\omega\}) = \sum_{i \in I} g(x_i)\mathbb{P}(A_i).\end{aligned}$$

□

Proposition 3.1.7 *Soit g une fonction positive. Si $\mathbb{E}[g(X)] = 0$ alors*

$$\mathbb{P}[g(X) = 0] = 1.$$

On dit que la variable $g(X)$ est nulle presque-sûrement.

Démonstration. On a $\mathbb{E}[g(X)] = \sum_{i \in I} g(x_i)\mathbb{P}[X = x_i] = 0$. Comme $g(x_i)\mathbb{P}[X = x_i]$ est positif, tous les termes sont nuls et si $\mathbb{P}[X = x_i] > 0$ alors $g(x_i) = 0$. Autrement dit $g(x) = 0$ si $x \in \bigcup_{i \in I, \mathbb{P}[X = x_i] > 0} \{x_i\}$. On a de plus $P_X \left(\bigcup_{i \in I, \mathbb{P}[X = x_i] = 0} \{x_i\} \right) = 0$, donc

$$\mathbb{P}\left(\bigcup_{i \in I, \mathbb{P}[X = x_i] > 0} \{x_i\} \right) = 1$$

□

L'indépendance entre deux variables aléatoires X, Y peut s'exprimer avec l'opérateur d'espérance.

Proposition 3.1.8 *X et Y sont deux variables indépendantes si pour toutes fonctions f et g continues bornées,*

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$$

Démonstration Si X et Y sont indépendantes alors la propriété est vraie par définition pour des fonctions f et g simples. Toute fonction continue bornée est limite de fonctions simples.

Critère d'indépendance. Soient X et Y deux variables discrètes telles que $X(\Omega) = \{x_1, x_2, ...\}$ et $Y(\Omega) = \{y_1, y_2, ...\}$. X et Y sont indépendantes si

$$\mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_i).$$

3.1.4 Moments, variance et écart-type

Définition 3.1.3 *On appelle* **moment d'ordre** k *de* X *l'espérance si elle existe de la variable* X^k *avec* $k \geq 2$
On a d'après le théorème de transfert :

$$\mathbb{E}[X^k] = \sum_{i \in I} x_i^k \mathbb{P}(X = x_i)$$

Définition 3.1.4 *On appelle* **variance** *de* X*, l'espérance de* $(X - \mathbb{E}[X])^2$ *:*

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

On a immédiatement $\mathbb{V}(X) = \sum_i (x_i - \mathbb{E}[X])^2 \mathbb{P}[X = x_i]$. On mesure l'écart de la variable aléatoire par rapport à sa moyenne. C'est une mesure de dispersion autour de $\mathbb{E}(X)$.

Proposition 3.1.9 *Si* X *admet une variance et* $\mathbb{V}(X) = 0$ *alors* $X = 0$ *avec probabilité* 1

Théorème 3.1.10 *Si* X *possède un moment d'ordre 2 alors*

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

3.2 Lois discrètes usuelles

On présente les lois classiques au programme.

Loi de Bernoulli

Définition 3.2.1 *Soit* $p \in [0, 1]$*,* X *suit la loi de Bernoulli de paramètre* p *si*

1. $X(\Omega) = \{0, 1\}$

2. $\mathbb{P}(X = 1) = p$ *et* $\mathbb{P}(X = 0) = 1 - p$

La fonction indicatrice d'un événement est une variable aléatoire qui suit une loi de Bernoulli. Soit $A \in \mathcal{F}$, $X = \mathbf{1}_A$ a pour loi, la loi de Bernoulli avec paramètre $p = \mathbb{P}(A)$.

Proposition 3.2.1 $\mathbb{E}(X) = p$ *et* $\mathbb{V}(X) = p(1 - p)$

Loi uniforme sur $[|1, n|]$

Définition 3.2.2 X *a pour loi, la loi uniforme sur* $[|1, n|]$ *si*
1 $X(\Omega) = [|1, n|]$
2 $\mathbb{P}(X = k) = \frac{1}{n}$.

Proposition 3.2.2 *On a* $\mathbb{E}[X] = \frac{n+1}{2}$ *et* $\mathbb{V}(X) = \frac{n^2-1}{12}$

Démonstration. Utiliser les formules

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

Loi de Poisson

La variable aléatoire X suit une loi de Poisson de paramètre λ (loi sur $\mathbb{N}$) si :

$$\mathbb{P}[X = k] = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Il faut savoir calculer les moments !

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Plus généralement

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

On en déduit que $\mathbb{E}[X^2 - X] = \lambda^2$, d'où

$$\mathbb{E}[X^2] = \lambda^2 + \lambda$$

et

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

Loi binomiale

On considère n tirages équiprobables indépendants dans un ensemble composé de deux types d'éléments, le premier (I) en proportion p, le deuxième en proportion $q = 1 - p$. Soit X le nombre d'éléments du premier type tirés parmi l'échantillon de taille n. La variable aléatoire X suit une loi binomiale. Supposons qu'à chaque élément échantillonné $i \in [|1, n|]$, on associe la variable aléatoire Y_i qui vaut 1 si i est de type I, 0 sinon. Les variables aléatoires Y_i sont des variables de Bernoulli de paramètre p indépendantes. Il est évident que le nombre X d'éléments de type (I) dans l'échantillon vérifie :

$$X = \sum_{i=1}^{n} Y_i$$

Définition 3.2.3 *Soient $n \in \mathbb{N}^\star$ et $p \in [0, 1]$. X suit la loi binomiale de paramètre n et p si*

1. $X(\Omega) = [|0, n|]$
2. $\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$

On a $\mathbb{E}[X] = np$ et $\mathbb{V}(X) = np(1-p)$.

Loi hypergéométrique

On a déjà rencontré cette loi dans le problème de contrôle de production 2.0.1. On redonne ici l'expérience type menant à une loi hypergéométrique. On considère un tirage équiprobable **sans remise** de n éléments dans une population de taille $N \geq n$. On s'intéresse à un type donné (I) (par exemple *défectueux*) d'éléments de la population, que l'on supposera être en proportion p (Np est donc un entier). Soit X le nombre d'éléments de type I présents dans l'échantillon de taille n.

Définition 3.2.4 *X suit une loi hypergéométrique de paramètres N, n, p si*

$$\mathbb{P}[X = k] = \frac{\binom{Np}{k}\binom{N(1-p)}{n-k}}{\binom{N}{n}},$$

avec $\max\{0; n - N(1-p)\} \leq k \leq \min\{n, Np\}$.

Proposition 3.2.3 $\mathbb{E}[X] = np$, $\mathbb{V}(X) = \frac{N-n}{N-1} np(1-p)$.

Démonstration. Pour tout $i \in [|1, N|]$, soit

$$\epsilon_i = \begin{cases} 0 & \text{si } i \text{ n'appartient pas à l'échantillon} \\ 1 & \text{sinon} \end{cases}$$

Il y a $\binom{N}{n}$ échantillons de taille n possibles et $\binom{N-1}{n-1}$ échantillons possibles de taille n contenant i. On a donc

$$\mathbb{P}[\epsilon_i = 1] = \frac{\binom{N-1}{n-1}}{\binom{N}{n}} = \frac{n}{N}.$$

On peut écrire

$$X = \sum_{i=1}^{N} Y_i \epsilon_i$$

où Y_i prend la valeur 1 si i est de type I, 0 sinon. On en déduit :

$$\mathbb{E}[X] = \sum_{i=1}^{N} \mathbb{E}[Y_i \epsilon_i] = \sum_{i=1}^{N} \mathbb{E}[Y_i]\mathbb{E}[\epsilon_i] = \sum_{i=1}^{N} p\frac{n}{N} = np$$

De même, on calcule

$$\mathbb{V}(X) = \frac{N-n}{N-1} np(1-p).$$

(exercice)

Loi géométrique

On considère une population dont une proportion p est composée d'éléments de type I donné. On désire obtenir 1 élément de ce type en procédant à une suite de tirage équiprobables et indépendants : Soit Y le nombre de tirage nécessaires pour obtenir un élément de type I. La loi de Y est la loi géométrique de paramètre p sur $\mathbb{N}^\star$: c'est à dire :

$$\mathbb{P}[Y = k] = (1-p)^{k-1} p$$

On a tiré $k-1$ éléments de type II, le k-ième est de type I. Les moments sont

$$\mathbb{E}[Y] = \tfrac{1}{p} \text{ et } \mathbb{V}(Y) = \tfrac{1-p}{p^2}.$$

Loi binomiale négative

Il s'agit d'une généralisation de la loi géométrique. Cette fois ci on désire obtenir n éléments de type I. On note Y le nombre de tirages nécessaires pour avoir n éléments. L'événement $\{Y = k\}$ signifie que k tirages ont donné un élément de type II, et n tirages ont donné un élément de type I (dont le dernier). Le nombre de choix possibles est $\binom{n-1}{n-1+k}$, on en déduit :

$$\mathbb{P}[Y = k] = \binom{n-1}{n-1+k} p^n (1-p)^k.$$

Moments :

$$\mathbb{E}[X] = n\tfrac{1-p}{p} \text{ et } \mathbb{V}(X) = n\tfrac{1-p}{p^2}.$$

3.3 Vecteurs aléatoires discrets

3.3.1 Vecteurs aléatoires discrets : lois marginales, indépendance

Soit X un vecteur aléatoire discret. Soient $X_1, X_2, ..., X_n$ ses composantes, c'est à dire

$$X : \omega \in \Omega \mapsto \begin{pmatrix} X_1(\omega) \\ X_2(\omega) \\ . \\ . \\ . \\ X_n(\omega) \end{pmatrix}.$$

Exemple 3.3.1 *On lance deux fois un dé équilibré. On pose $\Omega = [|1, 6|]^2$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $\mathbb{P}$ la probabilité uniforme. Soit X le vecteur aléatoire bi-dimensionnel :*

$$X : \omega \in \Omega \mapsto \begin{pmatrix} U(\omega) = \omega_1 + \omega_2 \\ V(\omega) = \max(\omega_1, \omega_2) \end{pmatrix}.$$

Exercice 3.3.1 *Les variables U et V sont elles indépendantes ? Déterminer la loi de X.*

Définition 3.3.1 (Lois marginales) *On appelle lois marginales de X, les lois de $X_1, ..., X_n$.*

La loi de X est aussi appelée loi conjointe des $X_i, 1 \leq i \leq n$. Les propriétés suivantes se généralisent au cas multidimensionnel, pour alléger les notations, on considère un vecteur bidimensionnel ($n = 2$). Soit $X = (U, V)$. La loi de U se déduit de la loi de (U, V) :

$$\mathbb{P}(U = u_i) = \sum_{j \in I} \mathbb{P}(U = u_i, V = v_j).$$

Exercice 3.3.2 *Reprenons le vecteur aléatoire de l'exemple précédent. Déterminer ses lois marginales.*

La loi conditionnelle de V sachant $U = u_i$ est la donnée de

$$\left(v_j ; \frac{\mathbb{P}(V = v_j \text{ et } U = u_i)}{\mathbb{P}(U = u_i)} \right).$$

Exercice 3.3.3 *Déterminer la loi de U sachant $V = 3$.*

Proposition 3.3.1 *Soient X et Y deux v.a.r discrètes indépendantes possédant une espérance*

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y].$$

Démonstration.

Proposition 3.3.2 (Admis) *Si $X_1, ..., X_n$ sont indépendantes, alors toute fonction de $X_1, ..., X_p$ est indépendantes de toute fonction de $X_{p+1}, ..., X_n$.*

3.3.2 Structure algébrique, covariance et corrélation

On note $\ell^0(\Omega, \mathcal{F}, \mathbb{P})$ l'ensemble des variables aléatoires discrètes définies sur un même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. L'espace ℓ^0 est un espace vectoriel pour la loi de composition interne $+$:

$$(X + Y) \text{ est l'élément de } \ell^0 \text{ qui à } \omega \in \Omega \text{ associe } X(\omega) + Y(\omega)$$

et la loi de composition externe scalaire :

$$\forall \lambda \in \mathbb{R},\ (\lambda.X)(\omega) = \lambda X(\omega).$$

Proposition 3.3.3 *Soit ℓ^1 l'ensemble des variables aléatoires discrètes possédant une espérance, i.e*

$$\ell^1 = \{X \in \ell^0; \mathbb{E}[|X|] < \infty\}.$$

1) ℓ^1 est un sous-espace vectoriel de ℓ^0
2) $\mathbb{E}$ est une forme linéaire sur ℓ^1

Proposition 3.3.4 *Soit $\ell^2_{\mathbb{P}}$ l'ensemble des éléments de $\ell^0_{\mathbb{P}}$ possédant un moment d'ordre 2, i.e*

$$\ell^2_{\mathbb{P}} = \{X \in \ell^0; \mathbb{E}[X^2] < \infty\}$$

$\ell^2_{\mathbb{P}}$ est un sous-espace vectoriel de $\ell^0_{\mathbb{P}}$ inclus dans $\ell^1_{\mathbb{P}}$.

Démonstration. Utiliser $ab \leq \frac{1}{2}(a^2 + b^2)$

Définition 3.3.2 (Covariance) *On appelle* ***covariance*** *l'application*

$$Cov : (X, Y) \in \ell^2 \times \ell^2 \mapsto \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Par définition $Cov(X, X) = \mathbb{V}(X)$. On a également la propriété importante suivante :

Proposition 3.3.5

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

Remarque 3.3.1 *Cette formule est importante et se simplifie lorsque $Cov(X_i, X_j) = 0$ pour tout $i \neq j$. On dit que les variables ne sont pas « corrélées ».*

Proposition 3.3.6 $Cov(X, Y) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right].$

On donne ici quelques propriétés d'algèbre bilinéaire associées à la covariance.

Théorème 3.3.7 *L'application Cov est une forme bilinéaire symétrique positive sur* $\ell^2_{\mathbb{P}}$.

1) Symétrie : $Cov(X,Y) = Cov(Y,X)$.
2) Bilinéarité : immédiat par linéarité de $\mathbb{E}$.
3) Positivité : $Cov(X,X) = \mathbb{E}\left((X - \mathbb{E}(X))^2\right) \geq 0$.

La covariance n'est **pas définie-positive**. On a seulement que $Cov(X,X) = 0 \Longrightarrow X$ est constante presque sûrement.

Remarque 3.3.2 *Cov est un produit scalaire sur le sous-espace vectoriel des variables aléatoires centrées (c'est à dire d'espérance nulle), quotienté par la relation d'équivalence notée "=" :* $X = Y \iff X - Y = 0$ *avec probabilté 1. Nous ne rentrons pas dans les détails ici, mais il faut reconnaître dans la notation* $\ell^1_{\mathbb{P}}, \ell^2_{\mathbb{P}}$ *les espaces « ℓ^p » de la théorie des séries. En effet dans le cadre des séries, on a*

$$\ell^p := \{(x_i, i \in \mathbb{N}) \in \mathbb{R}^{\mathbb{N}} : \sum_{i \geq 0} |x_i|^p < \infty\}$$

Dans notre cadre, $\ell^p_{\mathbb{P}}$ *correspond à l'espace* ℓ^p *où* x_i *est muni du poids* $\mathbb{P}[x_i]$. *Autrement dit :* $\ell^p_{\mathbb{P}}$ *est isomorphe à*

$$\{(x_i, i \in \mathbb{N}^\star) : \sum_{i \geq 1} |x_i|^p \mathbb{P}(x_i) < \infty\}.$$

Définition 3.3.3 (Coefficient de corrélation) *Soient X et Y deux éléments de L^2 de variance strictement positives. On appelle coefficient de corrélation linéaire le réel :*

$$\rho(X,Y) := \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)}.$$

On rappelle que $\sigma(X)$ est l'écart-type de la variable aléatoire X, c'est à dire $\sigma(X) = \sqrt{\mathbb{V}(X)} = \sqrt{Cov(X,X)}$.

Remarque 3.3.3 *On met en garde le lecteur quant à la non-transitivité de la corrélation :* $\rho(X,Y) > 0$ *et* $\rho(X,Z) > 0$ ***n'implique pas*** $\rho(Y,Z) > 0$. *C'est assez contre-intuitif et source d'erreurs potentielles lors d'interprétation.*

Proposition 3.3.8 *Quelque soit le couple (X,Y) de variables aléatoires de L^2, on a :*

$$|\rho(X,Y)| \leq 1$$

Démonstration. Il faut démontrer

$$|Cov(X,Y)| \leq \sigma(X)\sigma(Y).$$

Soit $\lambda \in \mathbb{R}$, on a

$$\mathbb{V}(X + \lambda Y) = \mathbb{V}(X) + \lambda^2 \mathbb{V}(Y) + 2\lambda Cov(X,Y).$$

Le trinôme en λ à droite dans l'égalité ci-dessus est toujours positif puisque $\mathbb{V}(X + \lambda Y) \geq 0$. On en déduit que son discriminant est négatif ou nul :

$$\Delta = 4(Cov(X,Y))^2 - 4\mathbb{V}(X)\mathbb{V}(Y) \leq 0.$$

On en déduit donc le résultat. □

Remarque 3.3.4 *C'est la même preuve que celle de l'inégalité de Cauchy-Schwarz. Il faut impérativement la connaître.*

3.3.3 Loi multinomiale

On commence par une généralisation du théorème binomiale : Soient $x_1, ..., x_k$ k réels, soit n un entier. On a

$$(x_1 + ... + x_k)^n = \sum_{n_1,...,n_k; n_1+...n_k=n} \frac{n!}{n_1!n_2!...n_k!} x_1^{n_1}...x_k^{n_k}.$$

On note parfois

$$\binom{n}{n_1, n_2, ..., n_k} = \frac{n!}{n_1!n_2!...n_k!}$$

Ce nombre est le coefficient multinomial. On passe maintenant à la loi multinomiale. Il s'agit d'une généralisation de la loi binomiale. On considère n tirages équiprobables, indépendants dans une population composée d'éléments de m types, chaque type j étant en proportion p_j :

$$\sum_{j=1}^{m} p_j = 1.$$

On définit la variable aléatoire N_j pour $j = 1, .., m$, comme étant le nombre d'éléments de type j figurant dans l'échantillon. Le vecteur $(N_1, ..., N_m)$ suit une loi multinomiale de paramètre $n, p_1, ..., p_m$. On la notera $\mathcal{M}(n, n_1, ..., n_m)$.

Définition 3.3.4 *Le vecteur aléatoire $X = (N_1, ..., N_m)$ suit une loi multinomiale de paramètres $n, p_1, ..., p_m$ si*

$$\mathbb{P}(X = (n_1, ..., n_m)) = \frac{n!}{n_1!...n_m!} p_1^{n_1}...p_m^{n_m}$$

où $\sum_{i=1}^{m} n_i = n$,

$$\frac{n!}{n_1!...n_m!}$$

est le nombre de partitions de $\{1, ..., n\}$ avec des sous-ensembles de cardinaux $n_1, ..., n_m$.

Proposition 3.3.9 *La j-ème loi marginale de X est une loi binomiale de paramètre n, p_j*

Un élément tiré au hasard est soit du type j avec probabilité p_j soit d'un autre type avec probabilité $(1 - p_j)$: le nombre total d'éléments du type j dans l'échantillon suit donc une loi binomiale de paramètre n, p_j. On va retrouver cela par le calcul. Soit $\mathbf{n}_j$ fixé. On définit l'événement A_j défini par

$$A_j := \{(n_1, ..., n_{j-1}, \mathbf{n}_j, n_{j+1}, ..., n_m); \sum_{i \neq j} n_i = n - \mathbf{n}_j\}$$

$$\begin{aligned}
\mathbb{P}[N_j = \mathbf{n}_j] = \mathbb{P}[X \in A_j] &= \sum_{A_j} \frac{n!}{n_1!...n_m!} p_1^{n_1}...p_m^{n_m} \\
&= \frac{n!}{\mathbf{n}_j!(n - \mathbf{n}_j)!} p_j^{\mathbf{n}_j} (1 - p_j)^{n - \mathbf{n}_j} \sum_{A_j} \frac{(n - \mathbf{n}_j)!}{n_1!...n_{j-1}!n_{j+1}!...n_m!} \\
&= \binom{n}{\mathbf{n}_j} p_j^{\mathbf{n}_j} (1 - p_j)^{n - \mathbf{n}_j}
\end{aligned}$$

Il n'y a aucune raison pour que les variables aléatoires N_l, N_j, soient non corrélées, et a fortiori indépendantes. On calcule la covariance de deux marginales N_j, N_l. Le couple (N_j, N_l) suit une loi multinomiale de paramètres : $n, p_j, p_l, 1-(p_j+p_l)$. En effet, lorsque l'on tire un échantillon de taille n, chaque élément tiré est soit de type j avec probabilité p_j, soit de type l avec probabilité p_l, soit d'un autre type avec probabilité $1-(p_j+p_l)$. On en déduit

$$\begin{aligned}
\mathbb{E}[N_j N_l] &= \sum_{(n_j,n_l);n_j+n_l\leq n} n_j n_l \frac{n!}{n_j! n_l! (n-n_j-n_l)!} p_j^{n_j} p_l^{n_l} (1-p_l-p_j)^{n-n_j-n_l} \\
&= \sum_{(n_j,n_l);n_j+n_l\leq n} \frac{(n-2)!}{(n_j-1)!(n_l-1)!(n-2-(n_j-1)-(n_l-1))!} p_j^{n_j-1} p_l^{n_l-1} (1-p_l-p_j)^{n-2-(n_j-1)-(n_l-1)} \\
&\qquad \times n(n-1) p_j p_i \\
&= n(n-1) p_j p_l
\end{aligned}$$

ce qui n'est pas égal à $\mathbb{E}[N_j]\mathbb{E}[N_l] = n^2 p_j p_l$.

3.3.4 Somme de variables aléatoires indépendantes

Somme de variables de Bernoulli

Proposition 3.3.10 *Soient $X_1, ..., X_n$ n variables indépendantes de loi de Bernoulli de paramètre p. La variable aléatoire*

$$Y = \sum_{i=1}^{n} X_i$$

suit une loi binomiale de paramètre (n,p).

Remarque 3.3.5 *C'est de cette façon que nous avons introduit la loi binomiale au Chapitre 3. En fait nous allons voir que nous pouvons toujours construire une variable aléatoire de loi binomiale comme somme de Bernoulli indépendantes.*

Ce qui suit n'est pas clairement au programme, mais permet de comprendre l'indépendance. On considère une variable de Bernoulli de paramètre p sur un espace $(\Omega, \mathcal{F}, \mathbb{P})$. On appelle succès, et note S, l'événement la variable de Bernoulli réalise 1. Afin de répéter n fois l'expérience, de façon indépendante, on va construire n *copies indépendantes* :

On se place sur l'espace $(\Omega^n, \mathcal{F}^{\otimes n}, \mathbb{P}^{\otimes n})$ où $\mathcal{F}^{\otimes n}$ et $\mathbb{P}^{\otimes n}$ sont respectivement la tribu produit et la probabilité produit. On définit

$$X_i : \omega \in \Omega^n \mapsto \omega_i.$$

Par définition de la tribu produit $\mathbb{P}^{\otimes n}$, on a

$$\begin{aligned}
\mathbb{P}^{\otimes n}(X_i = 1) &= \mathbb{P}^{\otimes n}(\Omega \times \Omega \times ..., \times \{S\}, \times ... \times \Omega) \\
&= \mathbb{P}(\Omega) \times \mathbb{P}(\{S\}) ... \times \mathbb{P}(\Omega) \\
&= \mathbb{P}[\{S\}] = p.
\end{aligned}$$

Soit $(x_1, ..., x_n) \in \{0,1\}^n$. On a

$$\mathbb{P}^{\otimes n}[X_1 = x_1, ..., X_n = x_n] = \mathbb{P}^{\otimes n}[X_1 = x_1] ... \mathbb{P}^{\otimes n}[X_n = x_n] = p^k (1-p)^{n-k}$$

où k est le nombre de $x_i, i \in [|1, |n]$ valant 1. On conclut immédiatement que la variable $Y = \sum_{i=1}^n X_i$ suit la loi suivante :

$$\mathbb{P}[Y = k] = (\text{ le nombre de façon de choisir } k \text{ éléments parmi } n) \times p^k(1-p)^{n-k} = \binom{n}{k} p^k(1-p)^{n-k}.$$

Somme de variables binomiales

Proposition 3.3.11 *Soient $X_1, ..., X_k$ des variables aléatoires binomiales de paramètres respectifs $(n_1, p), ..., (n_k, p)$. La variable aléatoire*

$$Y = \sum_{i=1}^k X_i$$

est une de loi binomiale de paramètre $(\sum_{i=1}^k n_i, p)$.

D'après la proposition précédente, chaque variable X_i est somme de n_i variables de Bernoulli indépendantes. On les numérote de la façon suivante :

$$X_1 = \sum_{i=1}^{n_1} B_i, X_2 = \sum_{i=n^1+1}^{n_2} B_i, ..., X_k = \sum_{i=n^1+...+n_{k-1}+1}^{n_k} B_i$$

On a donc

$$Y = \sum_{i=1}^k X_i = \sum_{i=1}^k \sum_{j=n_1+...+n_{i-1}+1}^{n_i} B_j = \sum_{l=1}^{n_1+...+n_k} B_l.$$

D'après la proposition précédente, on a donc que Y suit une loi binomiale de paramètre $(n_1 + ... + n_k, p)$.

Somme de variables de Poisson

Proposition 3.3.12 *Soient $X_1, ..., X_k$ des variables aléatoires définies sur un même espace, indépendantes de loi de Poisson de paramètre respectif λ_i. La variable aléatoire*

$$Y = \sum_{i=1}^k X_i$$

suit une loi de Poisson de paramètre $\sum_{i=1}^k \lambda_i$

On le fait dans le cas $k = 2$, et le soin de faire une récurrence sur k est laissé au lecteur. On se donne donc X_1 et X_2 variables telles que

$$\mathbb{P}[X_1 = j] = \frac{\lambda_1^j}{j!} e^{-\lambda_1}$$

$$\mathbb{P}[X_2 = l] = \frac{\lambda_2^l}{l!} e^{-\lambda_2}$$

On cherche à calculer $\mathbb{P}[X_1 + X_2 = n]$. On a :

$$\begin{aligned}
\mathbb{P}[X_1 + X_2 = n] &= \mathbb{P}\left(\bigcup_{i=0}^{n} \{X_1 = i\} \cap \{X_2 = n-i\}\right) \\
&= \sum_{i=0}^{n} \mathbb{P}(X_1 = i, X_2 = n-i) \\
&= \sum_{i=0}^{n} \mathbb{P}(X_1 = i)\mathbb{P}(X_2 = n-i) \\
&= \sum_{i=0}^{n} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \frac{\lambda_2^{n-i}}{(n-i)!} e^{-\lambda_2} \\
&= e^{-(\lambda_1+\lambda_2)} \sum_{i=0}^{n} \frac{1}{i!} \sum_{i=0}^{n} \binom{n}{i} \lambda_1^i \lambda_2^{n-i} = e^{-(\lambda_1+\lambda_2)} \frac{(\lambda_1+\lambda_2)^i}{i!}.
\end{aligned}$$

Un outil très important pour étudier plus généralement les sommes de variables aléatoires est la notion de fonction génératrice. Cette notion n'est pas exigible mais apparaît implicitement dans de nombreux exercices. Nous soulignons également qu'une fonction génératrice n'est rien d'autre qu'une série entière (totalement au programme.) Soit X une variable aléatoire à valeurs dans $\mathbb{N}$. La fonction génératrice associée à X est définie de la façon suivante :

$$G_X : s \in]-1,1[\mapsto \mathbb{E}[s^X] = \sum_{k=0}^{\infty} s^k \mathbb{P}[X = k].$$

La fonction génératrice caractérise entièrement la loi.

Proposition 3.3.13 *Soient X et Y deux variables aléatoires. X et Y ont même loi si et seulement si $G_X = G_Y$.*

Démonstration. Que peut-on dire des coefficients de deux séries entières égales ?

3.4 Exercices

Exercice 3.4.1 *Soit n un entier ≥ 1. On considère l'expérience aléatoire qui consiste à lancer successivement n fois un dé (équilibré) et à noter les n résultats obtenus (dans l'ordre).*

1. *Décrire l'univers Ω associé à cette expérience aléatoire. Déterminer le cardinal de Ω.*
2. *Calculer la probabilité de l'évènement $A =$ « on obtient au moins un 6 ».*
3. *Démontrer que pour avoir une probabilité supérieure ou égale à 0,9 d'obtenir au moins un 6, il faut et il suffit que :*
$$n \geq \frac{\ln(10)}{\ln(6) - \ln(5)}$$

Exercice 3.4.2 *Une urne contient trois boules blanches et trois boules noires. On tire simultanément quatre boules de l'urne au hasard. Quelle est la probabilité qu'il y ait autant de blanches que de noires ?*

Exercice 3.4.3 *On considère un jeu de fléchettes. La cible est un disque de rayon* 3*. Par simplicité, nous supposons que le joueur atteint toujours la cible. Supposons que le centre de la cible est situé à l'origine du plan* $\mathbb{R}^2$*. Nous souhaitons étudier les emplacements des fléchettes sur la cible. Nous supposons que la cible est divisée en trois cercles centrés à l'origine* C_1, C_2, C_3 *de rayons respectifs* $1, 2, 3$*. Ces cercles divisent la cible en trois anneaux* A_1, A_2, A_3 *avec*

$$A_k = \{(x, y) \in \mathbb{R}^2 ; k - 1 \leq \sqrt{x^2 + y^2} < k\}.$$

1) Donner un univers Ω*.*
2) On suppose que la probabilité de tomber dans un anneau est proportionnelle à son aire. Calculer

$$\mathbb{P}(A_1), \mathbb{P}(A_2), \mathbb{P}(A_3)$$

3) Soit maintenant X *la variable aléatoire qui donne le nombre de points selon l'emplacement de la fléchette. On suppose que la règle du jeu est la suivante : Si la fléchette tombe dans* A_k *alors le joueur remporte* $3 - k + 1$ *points. Calculer l'espérance de* X
4) Donner la fonction de répartition de X *et tracer la.*
5) On suppose maintenant que le joueur n'atteint pas la cible avec une probabilité donnée $p > 0$*. On suppose que si le joueur atteint la cible alors il marque des points selon les mêmes règles que précédemment. Sinon, il marque* 0 *point. Mêmes questions que précédemment.*

Exercice 3.4.4 (Minimum de deux variables aléatoires géométriques) *Dans cet exercice,* p *désigne un réel de* $]0;1[$ *et* $q = 1 - p$*. On considère deux variables aléatoires* X *et* Y*, indépendantes, et suivant toutes deux la même loi géométrique de paramètre* p*.*

1. *Soit* $i \in \mathbb{N}^\star$*. En remarquant que* $(X \leq i) = \bigcup_{k=1}^{i} (X = k)$*, montrer que* $\mathbb{P}(X > i) = q^i$*.*
2. *Vérifier que l'égalité précédente est encore vraie lorsque* $i = 0$*.*
3. *On définit une variable aléatoire* Z *en posant* $Z = \min(X, Y)$*, c'est-à-dire :*

$$Z = \begin{cases} X & \text{si } X \leq Y \\ Y & \text{si } Y < X \end{cases}$$

   Expliquer pourquoi, pour tout $i \in \mathbb{N}$ *:*

$$(Z > i) = (X > i) \cap (Y > i)$$

4. *En déduire* $P(Z > i)$*, pour* $i \in \mathbb{N}$*.*
5. *Soit* $i \in \mathbf{N}^*$*. En utilisant le fait que* $(Z > i - 1) = (Z > i) \cup (Z = i)$*, calculer* $P(Z = i)$*, pour* $i \in \mathbb{N}^*$*.*
6. *Démontrer que* Z *suit la loi géométrique de paramètre* $1 - q^2$*.*

Exercice 3.4.5 Le trousseau de clefs d'un gardien de nuit comporte dix clefs, dont une seule ouvre la porte du poste de garde. Pour qu'il y pénètre, il y a deux scenarios possibles :
• Cas A : il prend une clef au hasard, l'essaie, la met de côté si elle n'ouvre pas, et ainsi de suite.
• Cas B : il prend une clef au hasard, l'essaie, mais la laisse sur le trousseau si elle n'ouvre pas, et ainsi de suite.

On désigne respectivement par X_A et X_B les variables aléatoires égales aux nombres d'essais (y compris le bon) avant succès, dans le premier et le second scenario. Déterminer la loi de probabilité et la fonction de répartition de X_A et de X_B.
Calculer $\mathbb{E}[X_A]$ et $\mathbb{E}[X_B]$.

Le gardien utilise la méthode B un jour sur trois. Un jour, après avoir essayé 8 clefs, il n'a toujours pas ouvert la porte. Quelle est la probabilité pour qu'il ait utilisé la méthode B ?

Exercice 3.4.6 Soit n un entier naturel non nul. Une personne effectue n lancers indépendants d'une pièce parfaitement équilibrée. Soit X_n le nombre de "piles" obtenus.
1) Quelle est la loi de X_n ? quelle est son espérance ? sa variance ?
2) Calculer la probabilité pour qu'à l'issue des n lancers, le nombre de piles soit strictement supérieur au nombre de faces.

Exercice 3.4.7 (#) Une entreprise de construction produit des objets sur deux chaînes de montage A et B qui fonctionnent indépendemment l'une de l'autre. Pour une chaîne donnée, les fabrications des pièces sont indépendantes. On suppose que A produit 60% des objets et B produit 40% des objets. La probabilité qu'un objet construit par la chaine A soit défectueux est 0.1 alors que la probabilité pour qu'un objet construit par la chaine B soit défectueux est 0.2.

1. On choisit au hasard un objet à la sortie de l'entreprise. On constate que cet objet est défectueux. Calculer la probabilité de l'événement “l'objet provient de la chaîne A“ .
2. On suppose de plus que le nombre d'objets produits en une heure par A est une variable aléatoire Y qui suit une loi de Poisson de paramètre $\lambda = 20$.
   On considère la variable aléatoire X représentant le nombre d'objets défectueux produits par la chaîne A en une heure.
   (a) Rappeler la loi de Y ainsi que la valeur de l'espérance et de la variance de Y.
   (b) Soient k et n deux entiers naturels, déterminer la probabilité conditionnelle $\mathbb{P}_{(Y=n)}[X = k]$. (On distinguera les cas $k \leq n$ et $k > n$).
   (c) En déduire, en utilisant le système complet d'événements $(Y = i)_{i\in\mathbb{N}}$, que X suit une loi de Poisson de paramètre 2 .

Solution.

1. Pour un objet pris à la sortie, $P(A) = 0.6$ et $P(B) = 0.4$
   Soit D ="l'objet est défectueux”.
   On a $P(D/A) = 0.1$ et $P(D/B) = 0.2$ et comme (A, B) est un système complet d'événements,
$$\begin{aligned} P(D) &= P(D/A)\,P(A) + P(D/B)\,P(B) \\ &= 0.1 \cdot 0.6 + 0.2 \cdot 0.4 \\ &= 0.14 \end{aligned}$$
   Si l'objet est défectueux, la probabilité de l'événement “l'objet provient de la chaîne A“ est $P(A/D)$ que l'on calcule par la formule de Bayes :
$$\begin{aligned} P(A/D) &= \frac{P(A\cap D)}{P(D)} = \frac{P(D/A)\,P(A)}{P(D)} \\ &= \frac{0.1\cdot 0.6}{0.14} = \frac{0.06}{0.14} = \frac{6}{14} = \frac{3}{7} \end{aligned}$$

2. On suppose de plus que le nombre d'objets produits en une heure par A est une variable aléatoire Y qui suit une loi de Poisson de paramètre $\lambda = 20$.

   On considère la variable aléatoire X représentant le nombre d'objets défectueux produits par la chaîne A en une heure.

   (a) On a $Y(\Omega) = \mathbb{N}$ et pour tout entier n : $P(Y = n) = \frac{\lambda^n e^{-\lambda}}{n!}$. $E(Y) = \lambda = 20$ et $V(Y) = \lambda = 20$

   (b) Quand $Y = n$, X est le **nombre** d'objet défectueux parmi n, qui sont défectueux **indépendemment** les un des autres avec une même probabilité 0.1. Donc sachant $Y = n$, $X \hookrightarrow \mathcal{B}(n, 0.1)$ et

   $P[X = k/Y = n] = 0$ si $k > n$ et $P[X = k/Y = n] = C_n^k 0.1^k 0.9^{n-k}$ si $k \leq n$

   (c) Comme $(Y = n)_{n\in\mathbb{N}}$, est un système complet d'événements on a pour tout entier k :

$$P(X = k) = \sum_{n=0}^{+\infty} P[X = k/Y = n]\, P(Y = n)$$

   série convergente dont on calcule la somme partielle en distinguant suivant que $n \geq k$ ou $n < k$:

$$\begin{aligned}
\sum_{k=0}^{M} P[X = k/Y = n]\, P(Y = n) &= \sum_{n=0}^{k-1} P[X = k/Y = n]\, P(Y = n) \\
&\quad + \sum_{n=k}^{M} P[X = k/Y = n]\, P(Y = n) \\
&= 0 + \sum_{n=k}^{M} C_n^k 0.1^k 0.9^{n-k} \frac{20^n e^{-20}}{n!} \\
&= \left(\frac{0.1}{0.9}\right)^k e^{-20} \sum_{n=k}^{M} \frac{n!}{k!\,(n-k)!n!} (0.9 \cdot 20)^n \\
&= \left(\frac{1}{9}\right)^k e^{-20} \frac{1}{k!} \sum_{n=k}^{M} \frac{1}{(n-k)!} 18^n \\
&= \left(\frac{1}{9}\right)^k e^{-20} \frac{1}{k!} \sum_{m=0}^{M-k} \frac{1}{m!} 18^{m+k} \\
&\to \left(\frac{1}{9}\right)^k e^{-20} \frac{1}{k!} 18^k e^{18} = \frac{2^k e^{-2}}{k!}
\end{aligned}$$

   donc X suit une loi de Poisson de paramètre 2

Exercice 3.4.8 On lance successivement n boules au hasard dans N cases numérotées de 1 à N avec $N \geq 2$. On suppose que les différents lancers de boules sont indépendants et que la probabilité pour qu'une boule quelconque tombe dans une case donnée est $\frac{1}{N}$. Une case peut contenir plusieurs boules. Le gain étant fonction du nombre de cases atteintes, on étudie la variable aléatoire T_n égale au nombre de cases non vides à l'issue des n lancers.

1. Déterminer en fonction de n et N les valeurs prises par T_n (on distinguera deux cas : $n \leq N$ et $n > N$).
2. Donner les lois de T_1 et T_2.
3. Déterminer, lorsque $n \geq 2$, la probabilité des événements $(T_n = 1)$, $(T_n = 2)$, $(T_n = n)$ (pour la dernière probabilité, on distinguera deux cas : $n > N$ et $n \leq N$).
4. À l'aide de la formule des probabilités totales, justifier l'égalité (I) suivante, pour tout entier k tel que $1 \leq k \leq n$.

$$(I) \qquad P(T_{n+1} = k) = \frac{k}{N} P(T_n = k) + \frac{N-k+1}{N} P(T_n = k-1)$$

5. Afin de calculer l'espérance $\mathbb{E}(T_n)$ de la variable T_n, on considère la fonction polynomiale G_n définie par :

$$\forall x \in \mathbb{R},\ G_n(x) = \sum_{k=1}^{n} P(T_n = k) x^k$$

   (a) Quelle est la valeur de $G_n(1)$?
   (b) Exprimer $\mathbb{E}(T_n)$ en fonction de $G_n'(1)$.
   (c) En utilisant la relation (I), montrer que :

$$\forall x \in \mathbb{R},\ G_{n+1}(x) = \frac{1}{N}(x - x^2) G_n'(x) + x G_n(x)$$

   (d) En dérivant l'expression précédente, en déduire que :

$$\mathbb{E}(T_{n+1}) = \left(1 - \frac{1}{N}\right) E(T_n) + 1$$

   (e) Prouver enfin que l'espérance de T_n est donnée par :

$$\mathbb{E}(T_n) = N\left(1 - \left(1 - \frac{1}{N}\right)^n\right)$$

   (f) L'entier N étant fixé, calculer la limite de $\mathbb{E}(T_n)$ lorsque n tend vers $+\infty$. Interpréter le résultat.

Exercice 3.4.9 *On lance n fois ($n \geq 3$) une pièce de monnaie parfaitement équilibrée. Pour tout $i \in [|1, n|]$, on pose $A_i =$« on a obtenu pile au i-ème lancer ».*

1. *Soit X le nombre total de « piles » obtenu. Donner sa loi, son espérance et sa variance.*
2. *Si, à l'issue des n lancers, on obtient par exemple pile-pile-face-face-face-pile-..., on dit que l'on a une première série de longueur 2, parce qu'on a obtenu 2 piles au début (et pas plus), et une deuxième série de longueur 3, parce qu'on a obtenu ensuite 3 faces (et pas plus). Autre exemple : si l'on obtient face-face-pile-face-..., on a une première série de longueur 2 et une deuxième série de longueur 1. On note S_1 la longueur de la première série et S_2 la longueur de la deuxième série si elle existe. On convient de poser $S_2 = 0$ s'il n'y a pas de deuxième série, c'est-à-dire si $S_1 = n$. Déterminer $S_1(\Omega)$.*
3. *Justifier l'égalité $(S_1 = 1) = (A_1 \cap \overline{A_2}) \cup (\overline{A_1} \cap A_2)$. En déduire, en justifiant soigneusement, que $\mathbb{P}(S_1 = 1) = \dfrac{1}{2}$.*

4. *En étudiant l'événement $(S_1 = k)$ à la manière de la question précédente, démontrer que :*

$$\forall k \in [|1, n-1|],\ \mathbb{P}(S_1 = k) = \left(\frac{1}{2}\right)^k$$

5. *Démontrer que* $\mathbb{P}(S_1 = n) = \left(\frac{1}{2}\right)^{n-1}$.
6. *Vérifier, en utilisant les formules obtenues dans les deux questions précédentes, que :*

$$\sum_{k=1}^{n} \mathbb{P}(S_1 = k) = 1$$

7. *Soit* $x \in \mathbb{R} \setminus \{1\}$. *Démontrer que pour tout* $n \in \mathbb{N}^*$ *:*

$$\sum_{i=1}^{n} i x^{i-1} = \frac{n x^{n+1} - (n+1)x^n + 1}{(1-x)^2}.$$

   Calculer l'espérance de S_1.
8. *Justifier l'égalité* $S_2(\Omega) = [|0, n-1|]$.
9. *Calculer* $\mathbb{P}(S_2 = 0)$.
10. *Soit* $k \in [|1, n|]$. *Montrer que :*

$$\mathbb{P}(S_2 = 1 | S_1 = k) = \begin{cases} 0 & si\ k = n \\ 1 & si\ k = n-1 \\ \frac{1}{2} & si\ k \leq n-2 \end{cases}$$

11. *En déduire que :*

$$\mathbb{P}(S_2 = 1) = \frac{1}{2}$$

Exercice 3.4.10 Une urne contient N boules numérotées de 1 à N. On en tire n une à une ($n \leq N$). On étudiera d'abord le cas **avec remise**, puis **sans remise**.
Soit X et Y le plus petit et le plus grand des nombres obtenus
a) Calculer $\mathbb{P}(X \geq x)$ pour tout $x \in [|1, n|]$. En déduire la loi de X.
b) Calculer $\mathbb{P}(Y \leq y)$ pour tout $y \in [|1, n|]$. En déduire la loi de Y.

Exercice 3.4.11 Soit a un nombre réel, et X une variable aléatoire à valeurs dans $\mathbb{N}$, telle que : pour tout $k \in \mathbb{N}$,

$$\mathbb{P}[X = k] = \frac{a}{2^k k!}$$

a) Déterminer a.
b) La variable aléatoire X admet elle une espérance ? Si oui, la déterminer.
c) La variable aléatoire X admet elle une variance ? Si oui, la déterminer.

Exercice 3.4.12 Calculer la fonction de répartition de :
– la loi uniforme sur $\{1, ..., n\}$
– la loi de Bernoulli de paramètre p
– la loi Binomiale de paramètres 4 et $\frac{1}{2}$

– la loi géométrique de paramètre p.

Exercice 3.4.13 On admet que si $0 \leq x < 1$ et $r \in \mathbb{N}^\star$,

$$\frac{1}{(1-x)^r} = \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} x^k$$

Considérons une urne contenant des boules blanches en proportion p, des boules noires en proportion $1-p$.
On effectue des tirages avec remise dans cette urne jusqu'à l'obtention de r boules blanches. Soit X le nombre de boules noires obtenues (avant la $r$ème boule blanche).
1) Calculer $\mathbb{P}[X = k]$ pour tout $k \in \mathbb{N}$.
2) Montrer que l'on obtient presque-sûrement r boules blanches
3) Calculer $\mathbb{E}[X]$.

Exercice 3.4.14 Une urne contient N boules numérotées de 1 à N. On en tire n une à une ($n \leq N$). On étudiera d'abord le cas **avec remise**, puis **sans remise**.
Soit X et Y le plus petit et le plus grand des nombres obtenus
a) Calculer $\mathbb{P}(X \geq x)$ pour tout $x \in [|1, n|]$. En déduire la loi de X.
b) Calculer $\mathbb{P}(Y \leq y)$ pour tout $y \in [|1, n|]$. En déduire la loi de Y.

Exercice 3.4.15 Un garagiste dispose de 2 voitures de location. Chacune est utilisable en moyenne 4 jours sur 5. Il loue les voitures avec une marge de 50 euros par jour et par voiture. On considère X la variable aléatoire égale au nombre de clients se présentant chaque jour pour louer une voiture. On suppose $X(\Omega) = \{0, 1, 2, 3\}$ avec

$$\mathbb{P}[X = 0] = 0.1 \quad \mathbb{P}[X = 1] = 0.3 \quad \mathbb{P}[X = 2] = 0.4 \quad \mathbb{P}[X = 3] = 0.2.$$

a) Déterminer la loi de la variable aléatoire Y représentant le nombre de clients satisfaits par jour
b) Calculer la marge moyenne par jour.

Exercice 3.4.16 Une urne contient initialement une boule blanche et une boule rouge.
On effectue des tirages successifs d'une boule dans l'urne selon le protocole suivant : après chaque tirage, la boule tirée est remise dans l'urne et on rajoute dans l'urne une boule de la même couleur, avant le tirage suivant. Pour tout entier non nul n, on note X_n la variable aléatoire égale au nombre de boules blanches obtenues au cours des n premiers tirages.
– Déterminer la loi de X_1.
– Déterminer la loi de X_2.
– Montrer que la loi de X_n est donnée par $\mathbb{P}(X_n = k) = \frac{1}{n+1}$.

Exercice 3.4.17 (#) Soit N une variable aléatoire de Poisson de paramètre λ et $(X_i)_{i \geq 0}$ une suite de variables de Bernoulli de paramètre p, indépendantes entre elles et de N. Montrer que

$$S = \sum_{i=1}^{N} X_i$$

est une variable de Poisson de paramètre $p\lambda$.

Solution. Pour $n \geq 0$,

$$\mathbb{P}(S=n) = \mathbb{P}\left(\sum_{i=1}^{N} X_i = n\right) = \sum_{k\geq 0} \mathbb{P}\left(\sum_{i=1}^{N} X_i = n, N=k\right) = \sum_{k\geq 0} \mathbb{P}\left(\sum_{i=1}^{k} X_i = n, N=k\right).$$

Comme N est indépendant de $X_1, \ldots, X_k$, on a

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i = n, N=k\right) = \mathbb{P}\left(\sum_{i=1}^{k} X_i = n\right) \cdot \mathbb{P}(N=k).$$

De plus, comme on considère une somme de k variables de Bernoulli de paramètre p et indépendantes, $\sum_{i=1}^{k} X_i$ est une v.a. binomiale de paramètres k et p :

$$\mathbb{P}\left(\sum_{i=1}^{k} X_i = n\right) = \begin{cases} 0 & \text{si } n > k \\ \binom{k}{n} p^n (1-p)^{k-n} & \text{sinon.} \end{cases}$$

Comme N est une v.a. $\mathcal{P}(\lambda)$, $\mathbb{P}(N=k) = e^{-\lambda}\frac{\lambda^k}{k!}$.
Finalement, on a

$$\mathbb{P}(S=n) = \sum_{k\geq n} \binom{k}{n} p^n (1-p)^{k-n} e^{-\lambda}\frac{\lambda^k}{k!} = \frac{p^n e^{-\lambda}}{n!} \sum_{k\geq n} \frac{k!}{k!(k-n)!}(1-p)^{k-n}\lambda^k$$

et par un changement de variables dans la somme :

$$\mathbb{P}(S=n) = \frac{p^n e^{-\lambda}}{n!} \sum_{k\geq 0} \frac{1}{k!}(1-p)^k \lambda^{k+n} = \frac{\lambda^n p^n e^{-\lambda}}{n!} e^{\lambda(1-p)} = \frac{(\lambda p)^n}{n!} e^{-\lambda p}.$$

Exercice 3.4.18 (Loi multinomiale #) Une boîte contient n jetons numérotés de 1 à n. On tire un jeton au hasard, on note son numéro et on le remet dans la boîte. Si le numéro de ce jeton est i, alors on tire au hasard et sans remise i jetons de la boîte que l'on distribue au hasard dans trois urnes U_1, U_2 et U_3 (vides au départ). Pour $k = 1, 2, 3$, on note X_k la variable aléatoire désignant le nombre de jetons de l'urne U_k après cette opération.

1. Soit X la variable aléatoire donnant le numéro du jeton que l'on a tiré au départ dans la boîte. Quelle est la loi de X ? Déterminer la loi conjointe du couple (X_k, X), où $k = 1, 2, 3$.
2. Calculer pour $k = 1, 2, 3$, l'espérance de X_k (On pourra utiliser $X = X_1 + X_2 + X_3$).
3. (a) Trouver la loi du vecteur aléatoire (X_1, X_2, X).
   (b) En déduire la loi du vecteur aléatoire (X_1, X_2, X_3).

solution

1. La loi de X est la loi uniforme sur $\{1, .., n\}$. Pour $1 \leq i \leq n$ et $0 \leq j \leq n$,

$$\mathbb{P}(X_k = j, X = i) = \mathbb{P}(X_k = j | X = i)\mathbb{P}(X = i) = \mathbb{P}(X_k = j | X = i) \cdot \frac{1}{n}$$

et sachant $\{X = i\}$, la v.a. X_k est une binomiale de paramètres $1/3$ et i. On a donc

$$\mathbb{P}(X_k = j, X = i) = \begin{cases} 0 & \text{si } j > i \\ \frac{1}{n}\binom{j}{i}\left(\frac{1}{3}\right)^j \left(\frac{2}{3}\right)^{i-j} & \text{si } j \leq i \leq n. \end{cases}$$

2. Comme $X = X_1 + X_2 + X_3$ et que les trois v.a. X_1, X_2 et X_3 ont même loi,

$$3\mathbb{E}[X_1] = \mathbb{E}[X] = \frac{1}{n}\frac{n(n+1)}{2}$$

et donc $\boxed{\mathbb{E}[X_1] = \frac{n+1}{6}}$.

3. (a) Pour $1 \leq i \leq n$ et $0 \leq j, k \leq n$,

$$\mathbb{P}(X_1 = j, X_2 = k, X = i) = \mathbb{P}_{(X=i)}(X_1 = j, X_2 = k)\mathbb{P}(X = i) = \frac{1}{n}\mathbb{P}_{(X=i)}(X_1 = j, X_2 = k).$$

Conditionnellement à $\{X = i\}$, l'expérience revient à placer i boules dans 3 urnes. L'espace des configurations possibles, qui est de cardinal 3^i, est muni de la probabilité uniforme. De plus, le nombre de configurations avec j boules dans la 1ère urne et k boules dans la seconde est

$$\frac{i!}{j!k!(i-j-k)!}$$

si $j + k \leq i$ et 0 sinon.
On a donc

$$\mathbb{P}(X_1 = j, X_2 = k, X = i) = \frac{1}{n}\frac{1}{3^i}\frac{i!}{j!k!(i-j-k)!} \quad \text{si } j + k \leq i \leq n.$$

(b) Par la question précédente, on a directement

$$\mathbb{P}(X_1 = j, X_2 = k, X_3 = l) = \frac{1}{n}\frac{1}{3^{j+k+l}}\frac{(j+k+l)!}{j!k!l!} \quad \text{si } j + k + l \leq n.$$

Exercice 3.4.19 (#) On joue à pile ou face avec une pièce non équilibrée. A chaque lancer la probabilité d'obtenir pile est 2/3, et face 1/3. Les lancers sont supposés indépendants. On note X la variable aléatoire réelle égale au nombre de lancers nécessaires pour l'obtention, pour la première fois de deux piles consécutifs.
Soit n un entier naturel non nul. On note p_n la probabilité de l'événement $(X = n)$.
1) Expliciter les événements $(X = 2), (X = 3), (X = 4), (X = 5)$. Déterminer la valeur de p_1, p_2, p_3, p_4 et p_5
2) A l'aide de la formule des probabilités totales, en distinguant deux cas selon le résultat du premier lancer, montrer que

$$\forall n \geq 3, p_n = \frac{2}{9}p_{n-2} + \frac{1}{3}p_{n-1}.$$

3) En déduire l'expression de p_n en fonction de n pour $n \geq 1$.
4) Calculer $\mathbb{E}[X]$.

Exercice 3.4.20 1) Déterminer α, β, γ tels que pour tout $k \in \mathbb{N}^\star$,

$$\frac{1}{k(k+1)(k+2)} = \frac{\alpha}{k} + \frac{\beta}{k+1} + \frac{\gamma}{k+2}$$

2) Soit $a \in \mathbb{R}$. Déterminer a pour que

$$p_k = \frac{a}{k(k+1)(k+2)}$$

soit une loi de probabilité.
3) Cette loi de probabilité admet elle une espérance ? une variance ?

Exercice 3.4.21 Soit X une variable aléatoire à valeurs dans $\mathbb{N}^\star$.
1) Montrer que pour tout $n \in \mathbb{N}^\star$,

$$\sum_{k=1}^{n} k\mathbb{P}[X = k] = \left(\sum_{i=1}^{n-1} \mathbb{P}[X > k]\right) - n\mathbb{P}(X > n).$$

2) On suppose que $\lim_{n\to\infty} n\mathbb{P}(X > n) = 0$ et que la série de terme général $u_k = \mathbb{P}(X > k)$ converge. Montrer que X admet une espérance et que

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

3) Réciproquement, on suppose que X admet une espérance. Montrer que

$$\lim_{n\to\infty} n\mathbb{P}(X > n) = 0$$

et en déduire

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}[X > k].$$

Exercice 3.4.22 Soit X une v.a.r discrète telle que $X(\Omega) = \mathbb{N}$. Montrer que
1) Si $\mathbb{E}(X)$ existe, $\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}[X \geq k]$.
2) Si $\mathbb{E}[X(X-1)]$ existe,

$$\mathbb{E}[X(X-1)] = 2\sum_{p=2}^{\infty}\sum_{n=p}^{\infty} \mathbb{P}[X \geq n].$$

Exercice 3.4.23 On considère un entier naturel N supérieur ou égal à 3. Une urne contient N boules numérotées de 1 à N. On y effectue des tirages successifs d'une boule avec remise de la boule tirée après chaque tirage, jusqu'à obtenir pour la première fois un numéro déjà tiré. On note alors T_N le rang aléatoire de ce dernier tirage.
Par exemple, si on a obtenu successivement les numéros 1-5-4-7-3-5, la variable T_N prend la valeur 6. Alors que si on a obtenu 5-4-2-2 la variable T_N prend la valeur 4.
1) Déterminer la loi, l'espérance et la variance de T_3.
2) Soit $N \geq 3$
  a) Déterminer l'ensemble des valeurs que peut prendre T_N.
  b) Calculer $\mathbb{P}[T_N = 2]$, $\mathbb{P}[T_N = 3]$ et $\mathbb{P}[T_N = N+1]$
  c) Prouver pour tout entier k de $\{1, ..., N\}$, les égalités

$$\mathbb{P}[T_N > k] = \frac{N!}{(N-k)!N^k} = \prod_{i=0}^{k-1}\left(1 - \frac{i}{N}\right).$$

En déduire la loi de la variable aléatoire T_N

Exercice 3.4.24 On joue à pile ou face avec une pièce non équilibrée. A chaque lancer la probabilité d'obtenir pile est 2/3, et face 1/3. Les lancers sont supposés indépendants. On note X la variable aléatoire réelle égale au nombre de lancers nécessaires pour l'obtention, pour la première fois de deux piles consécutifs.
Soit n un entier naturel non nul. On note p_n la probabilité de l'événement $(X = n)$.
1) Expliciter les événements $(X = 2), (X = 3), (X = 4), (X = 5)$. Déterminer la valeur de p_1, p_2, p_3, p_4 et p_5
2) A l'aide de la formule des probabilités totales, en distinguant deux cas selon le résultat du premier lancer, montrer que

$$\forall n \geq 3, p_n = \frac{2}{9}p_{n-2} + \frac{1}{3}p_{n-1}.$$

3) En déduire l'expression de p_n en fonction de n pour $n \geq 1$.
4) Calculer $\mathbb{E}[X]$.

Exercice 3.4.25 1) Déterminer α, β, γ tels que pour tout $k \in \mathbb{N}^\star$,

$$\frac{1}{k(k+1)(k+2)} = \frac{\alpha}{k} + \frac{\beta}{k+1} + \frac{\gamma}{k+2}$$

2) Soit $a \in \mathbb{R}$. Déterminer a pour que

$$p_k = \frac{a}{k(k+1)(k+2)}$$

soit une loi de probabilité.
3) Cette loi de probabilité admet elle une espérance ? une variance ?

Exercice 3.4.26 (A faire !) Soit $\mathbb{P}$ une probabilité et X une variable aléatoire à valeurs dans $\mathbb{N}^\star$.
1) Montrer que pour tout $n \in \mathbb{N}^\star$,

$$\sum_{k=1}^{n} k\mathbb{P}[X = k] = \left(\sum_{i=1}^{n-1} \mathbb{P}[X > k]\right) - n\mathbb{P}(X > n).$$

2) On suppose que $\lim_{n\to\infty} n\mathbb{P}(X > n) = 0$ et que la série de terme général $u_k = \mathbb{P}(X > k)$ converge. Montrer que X admet une espérance et que

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

3) Réciproquement, on suppose que X admet une espérance. Montrer que

$$\lim_{n\to\infty} n\mathbb{P}(X > n) = 0$$

et en déduire

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} \mathbb{P}[X > k].$$

Exercice 3.4.27 Soit X une v.a.r discrète telle que $X(\Omega) = \mathbb{N}$. Montrer que
1) Si $\mathbb{E}(X)$ existe, $\mathbb{E}(X) = \sum_{k=1}^{\infty} \mathbb{P}[X \geq k]$.
2) Si $\mathbb{E}[X(X-1)]$ existe,

$$\mathbb{E}[X(X-1)] = 2 \sum_{p=2}^{\infty} \sum_{n=p}^{\infty} \mathbb{P}[X \geq n].$$

Exercice 3.4.28 Soient X et Y deux variables aléatoires indépendantes, à valeurs dans $\mathbb{N}$.
- Déterminer en fonction de la loi de X et de la loi de Y, la loi de $Z = X + Y$.
- Déterminer la loi de $T = \min(X, Y)$.

Exercice 3.4.29 – Montrer que la somme de n variables aléatoires indépendantes de Bernoulli de paramètre p est une loi binomiale de paramètres n, p.
- Montrer que la somme de n variables aléatoires de Poisson de paramètres $\lambda_1, ..., \lambda_n$ respectivement est une loi de Poisson de paramètre $\sum_{i=1}^{n} \lambda_i$

Exercice 3.4.30 Une urne contient une boule blanche et une boule noire, les boules étant indiscernables au toucher. On y prélève une boule, chaque boule ayant la même probabilité d'être tirée, on note sa couleur, et on la remet dans l'urne avec c boules de la couleur de la boule tirée. On répète cette épreuve, on réalise ainsi une succession de tirages ($n \geq 2$).
On considère les variables aléatoires $(X_i)_{1 \leq i \leq n}$ définies par :

$X_i = 1$ si on obtient une boule blanche au i-ème tirage.
$X_i = 0$ sinon.

On définit, pour $2 \leq p \leq n$, la variable aléatoire Z_p par

$$Z_p = \sum_{i=1}^{p} X_i$$

1) Déterminer la loi du couple (X_1, X_2). En déduire la loi de X_2.
2) Déterminer la loi de Z_2
3) a) Déterminer $Z_p(\Omega)$. Soit $p \leq n-1$
b) Déterminer $\mathbb{P}(X_{p+1} = 1 | Z_p = k)$ pour $k \in Z_p(\Omega)$. En utilisant la formule des probabilités totales, montrer que

$$\mathbb{P}[X_{p+1} = 1] = \frac{1 + c\mathbb{E}[Z_p]}{2 + pc}$$

c) Montrer par récurrence que pour tout $p \in [|1, n|]$,

$$\mathbb{P}[X_p = 1] = \mathbb{P}[X_p = 0] = \frac{1}{2}$$

Exercice 3.4.31 *Soit une pièce avec probabilité d'avoir face p.*
- *On lance la pièce n fois, soit X la variable aléatoire représentant le nombre de faces obtenues, et Y le nombre de piles. Quelle est la loi de X, la loi de Y ? montrer que les variables aléatoires X et Y ne sont pas indépendantes.*
- *On lance maintenant la pièce N fois où N est un nombre aléatoire qui suit une loi de Poisson. On note encore X et Y le nombre de faces et le nombre de piles. Déterminer la loi du couple (X, Y), en déduire que X et Y sont indépendantes, et donner leurs lois respectives.*

Exercice 3.4.32 On considère n urnes numérotées de 1 à n. La première urne contient des boules blanches et des boules noires ; la proportion de boules blanches est p_1. Les urnes suivantes contiennent chacune a boules blanches et a boules noires.
On effectue n tirages de la manière suivante : on tire une boule de la premiere urne que l'on place dans la deuxième urne, puis on tire une boule de la deuxième urne que l'on place dans la troisième, et ainsi de suite jusqu'au tirage dans la dernière urne.

Pour $1 \leq k \leq n$, on désigne par X_j la variable aléatoire égale à 1 si la boule tirée dans la kème urne est blanche, égale à 0 si la boule tirée de la $k$ème urne est noire.
1) Déterminer les lois de probabilités de X_1 et X_2, puis leurs espérances et leurs variances en fonction de p_1 et de a.
2) Démontrer qu'il existe une valeur de p_1 pour laquelle X_1 et X_2 suivent la même loi de probabilité.
3) Pour cette valeur de p_1 étudier l'indépendance de X_1 et X_2. Pour $1 \leq k \leq n$, on pose $p_k = \mathbb{P}[X_k = 1]$ et $q_k = \mathbb{P}[X_k = 0]$.
4) Démontrer qu'il existe une matrice M dépendant de a telle que pour $k \in [|1, n-1|]$

$$\begin{pmatrix} p_{k+1} \\ q_{k+1} \end{pmatrix} = M \begin{pmatrix} p_k \\ q_k \end{pmatrix}$$

5) Déterminer M^n pour $n \in \mathbb{N}$, puis déterminer la loi de X_n.
6) Déterminer $\lim\limits_{n\to\infty} p_n$ et $\lim\limits_{n\to\infty} q_n$.

Exercice 3.4.33 Soit (X, Y) un couple de variables aléatoires à valeurs dans $\mathbb{N} \times \mathbb{N}$ tel que : $\forall (j, k) \in \mathbb{N}^2$,

$$\mathbb{P}(\{X = j\} \cap \{Y = k\}) = \frac{(j+k)\lambda^{j+k}}{ej!k!}$$

a) Déterminer λ. (On pourra étudier $f : x \mapsto 2xe^{2x-1} - 1$) sur $\mathbb{R}_+$.
b) Déterminer les lois de X et de Y. Les variables X et Y sont elles indépendantes ?

Exercice 3.4.34 Soit X et Y deux variables indépendantes vérifiant :

$$\mathbb{P}[X = n] = \mathbb{P}[Y = n] = \frac{1}{4}\left(\frac{1 + a^n}{n!}\right)$$

pour tout $n \in \mathbb{N}$.
1) Déterminer a.
2) Calculer $\mathbb{E}[X]$ et $\mathbb{V}(X)$
3) Déterminer la loi de $S = X + Y$.

Exercice 3.4.35 On admet que si $0 \leq x < 1$ et $r \in \mathbb{N}^\star$,

$$\frac{1}{(1-x)^r} = \sum_{k=0}^{\infty} \binom{k+r-1}{r-1} x^k$$

Considérons une urne contenant des boules blanches en proportion p, des boules noires en proportion $1 - p$.
On effectue des tirages avec remise dans cette urne jusqu'à l'obtention de r boules blanches. Soit X le nombre de boules noires obtenues (avant la $r$ème boule blanche).

1) Calculer $\mathbb{P}[X = k]$ pour tout $k \in \mathbb{N}$.
2) Montrer que l'on obtient presque-sûrement r boules blanches
3) Calculer $\mathbb{E}[X]$.

Exercice 3.4.36 (#) Un joueur va au casino avec une fortune $a \in \mathbb{N}$. A chaque partie, il peut gagner 1 euro avec une probabilité p et perdre 1 euro avec une probabilité $q = 1 - p$. Le but du joueur est alors de jouer jusqu'à l'obtention de la fortune $c \geq a$, $c \in \mathbb{N}$ mais il doit s'arrêter s'il est ruiné. On note $s_c(a)$ sa probabilité de succès (atteindre c avant la ruine).

1. Calculer $s_c(0)$ et $s_c(c)$
2. Montrer, pour $a > 0$, en raisonnant sur ce qui s'est passé au premier coup, la relation
$$s_c(a) = ps_c(a+1) + qs_c(a-1)$$
3. Déduire la valeur de $s_c(a)$ suivant que $p = 0,5$ ou $p \neq 0,5$.
4. Application numérique : Calculer la valeur précédente avec $a = 900$; $c = 1000$; $a = 100$; $c = 20000$ dans les cas $p = 0$; 5 et $p = 18/38$.

Solution

1. Si on démarre avec une fortune nulle, on ne peut atteindre la somme c : $\boxed{s_c(0) = 0}$. Si on démarre avec c, on a atteint la somme c et donc $\boxed{s_c(c) = 1}$.
2. Au premier coup, on peut soit gagner 1 euro (avec proba p), soit en perdre 1 (avec proba q). Si on part d'une fortune a et que l'on gagne au premier coup, on se retrouve avec une fortune de $a + 1$ et la probabilité d'atteindre c vaut alors $s_c(a + 1)$. On a la même chose si on perd au premier coup et on a bien la formule de récurrence pour $1 \leq a \leq c - 1$:
$$s_c(a) = ps_c(a+1) + qs_c(a-1)$$
3. Si $\underline{p = 0}$ (et donc $q = 1$), on a $s_c(a) = s_c(a - 1)$ pour $1 \leq a \leq c - 1$. On a donc $s_c(a) = 0$ si $a < c$ et $s_c(c) = 1$.
Le cas le plus intéressant est bien sûr $\underline{p > 0}$. On a alors une suite récurrente d'ordre 2 qui vérifie
$$ps_c(a+1) = s_c(a) - qs_c(a-1)$$
et $s_c(0) = 0$, $s_c(c) = 1$. Pour trouver son expression générale, on doit trouver les racines de son équation caractéristique $pr^2 - r + q = 0$. Après résolution, on trouve que cette équation a deux racines : 1 et q/p.
Si $q = p$, c'est-à-dire si $\underline{\text{p=1/2}}$, 1 est la racine double de l'équation caractéristique et alors on a
$$s_c(a) = (\alpha a + \beta)1^a = \alpha a + \beta$$
avec α et β à déterminer. D'après la question a), $0 = s_c(0) = \beta$ et $1 = s_c(c) = \alpha c + \beta$. On en déduit que $\beta = 0$ et $\alpha = 1/c$. Cela donne au final $\boxed{s_c(a) = a/c}$.
Si $q \neq p$, c'est-à-dire si $\underline{p \neq 1/2}$, l'équation caractéristique a deux racines distinctes (1 et q/p) et alors
$$s_c(a) = \alpha \left(\frac{q}{p}\right)^a + \beta 1^a = \alpha \left(\frac{q}{p}\right)^a + \beta$$

avec α et β à déterminer. D'après la question a), $0 = s_c(0) = \alpha + \beta$ et $1 = s_c(c) = \alpha \left(\frac{q}{p}\right)^c + \beta$ et alors $\alpha = -\beta = (1-(q/p)^c)^{-1}$ ce qui entraîne

$$\boxed{s_c(a) = \frac{1-(q/p)^a}{1-(q/p)^c}}\ .$$

4. Application numérique : si $p = 1/2$, $s_{1000}(900) = 9/10$. si $p = 18/38$, $s_{1000}(900) = 2,66 \cdot 10^{-5}$.

Exercice 3.4.37 (Une marche aléatoire) On étudie le cours en bourse d'une action. On suppose que les variations journalières sont indépendantes les unes des autres. On convient de noter 0 le cours au début de l'observation. On suppose que chaque jour, l'action monte d'une unité $(+1)$ avec probabilité p ou descend d'une unité (-1) avec probabilité $1-p$. On note X_{2n}, le cours constaté le $2n$-ème jour.
Par exemple si n=2, et que le cours à baisser les trois premiers jours et monté le quatrième jour, on a $X_4 = -1-1-1+1 = -2$
1) Quelles sont les valeurs prises par X_{2n} ?
2) On note Y_{2n} le nombre de jours parmi les $2n$ jours d'observation où l'action a monté ; et Z_{2n} le nombre de jours parmi les $2n$ jours, où l'action a baissé. Quelles sont les lois de probabilité de Y_{2n} et Z_{2n} ? Donner leur espérance.
3) Quelles relations lient d'une part n et $Y_{2n}Z_{2n}$, et d'autre part X_{2n}, Y_{2n} et Z_{2n} ?
En déduire une expression de X_{2n}? Montrer que pour tout $k \in [|-n, n-]$

$$\mathbb{P}[X_{2n} = 2k] = \binom{2n}{n+k} p^{n+k} q^{n-k}$$

Exercice 3.4.38 On dispose d'un dé équilibré à 6 faces et d'une pièce truquée telle que la probabilité d'apparition de pile soit égale à p. Soit N un entier naturel non nul. On effectue N lancers du dé ; si n est le nombre de 6 obtenus, on lance alors n fois la pièce.
On définit trois variables aléatoires X, Y, Z de la manière suivante :

Z indique le nombre de 6 tirés parmi les N lancers de dé.

X indique le nombre de piles obtenus parmi les lancers de la pièce.

Y indique le nombre de faces obtenues au lancers de la pièce.

1) Préciser la loi de Z, son espérance, sa variance.
2) Pour $k \in \mathbb{N}, n \in \mathbb{N}$, déterminer la probabilité conditionnelle $\mathbb{P}[X = k | Z = n]$.
3) Montrer que pour tout couple d'entiers naturels (k, n) :

$$\begin{aligned} \mathbb{P}[X = k \text{ et } Z = n] &= \binom{n}{k}\binom{N}{n} p^k (1-p)^{n-k} \left(\frac{5}{6}\right)^{N-n} \left(\frac{1}{6}\right)^n \quad \text{si } 0 \le k \le n \le N \\ &= 0 \text{ sinon.} \end{aligned}$$

4) Calculer la probabilité $\mathbb{P}(X = 0)$.

5) Montrer que pour tout couple d'entiers naturels (k, n) tels que $0 \leq k \leq n \leq N$

$$\binom{n}{k}\binom{N}{n} = \binom{N}{k}\binom{N-k}{n-k}.$$

En déduire la probabilité $\mathbb{P}(X = k)$.
6) Montrer que la variable aléatoire X suit une loi binomiale de paramètre $(N, \frac{p}{6})$. Quelle est la loi de Y ?
7) Calculer la covariance de (X, Y). X et Y sont-elles indépendantes ?
8) Déterminer la loi du couple (X, Y).

Exercice 3.4.39 Soit X et Y deux variables aléatoires dans $\mathbb{N}^\star$, telles que

$$\mathbb{P}[\{X = i\} \cap \{Y = j\}] = \frac{a}{2^{i+j}}, \text{ pour tout } (i, i) \in \mathbb{N}^\star \times \mathbb{N}^\star$$

a) Calculer a
b) Déterminer les lois marginales de X et Y.
c) X et Y sont elles indépendantes ?

Exercice 3.4.40 Soit X une variable aléatoire uniforme sur $[|1, n|]$ et Y une variable aléatoire uniforme sur $[|1, X|]$, déterminer la loi de Y et calculer $\mathbb{E}[Y]$.

Exercice 3.4.41 Soit n un entier naturel non nul, et X et Y deux variables aléatoires indépendantes suivant toutes les deux des lois uniformes sur $\{1, ..., n\}$.

1. Calculer $\mathbb{P}(X = Y)$ et $\mathbb{P}(X \geq Y)$.
2. Déterminer la loi de $D = Y - X$.

Exercice 3.4.42 Soit X_1, X_2, X_3 des variables de Bernoulli, deux à deux indépendantes, prenant les valeurs 0 ou 1 avec probabilité 1/2. On considère les variables

$$Y = X_1X_2 \qquad \text{et} \qquad Z = X_2X_3$$

1. Déterminer la loi jointe de (Y, Z).
2. Les variables aléatoires X et Y sont-elles indépendantes ?
3. Calculer l'esperance et la variance de $Y + Z$ et YZ

Exercice 3.4.43 Soit n un entier naturel non nul, et X et Y deux variables aléatoires indépendantes suivant toutes les deux des lois uniformes sur $\{1, ..., n\}$.

1. Calculer $\mathbb{P}(X = Y)$ et $\mathbb{P}(X \geq Y)$.
2. Déterminer la loi de $D = Y - X$.

Exercice 3.4.44 Soit X_1, X_2, X_3 des variables de Bernoulli, deux à deux indépendantes, prenant les valeurs 0 ou 1 avec probabilité 1/2. On considère les variables

$$Y = X_1X_2 \qquad \text{et} \qquad Z = X_2X_3$$

1. Déterminer la loi jointe de (Y, Z).
2. En déduire les lois marginales de Y et Z.

3. Calculer $\mathrm{Cov}(Y, Z)$ Les variables aléatoires X et Y sont-elles indépendantes ?
4. Calculer l'espérance et la variance de $Y + Z$ et YZ

Exercice 3.4.45 (Processus d'assainissement informatique) *Soit $p \in]0, 1[$, une entreprise dispose de N copies d'un même logiciel. Une proportion p est infectée par un virus, et il est impossible de discerner un logiciel sain d'un contaminé. On suppose que $N = nm$ avec n et m deux entiers strictement plus grand que 1. Un des responsables propose la méthode suivante pour assainir le lot :*

- *Les N copies forment la génération 0.*
- *On prélève n logiciels au hasard et avec remise dans la génération 0. On les copie chacun m fois. Les $N = nm$ logiciels obtenus constituent la génération 1.*
- *On itère le procédé. Durant tout le processus, la copie d'un logiciel sain est saine, d'un logiciel contaminé, contaminée.*

Le statisticien pense que si la proportion p est faible, on a de bonnes chances d'obtenir un lot sain après un assez grand nombre d'opérations.
On souhaite vérifier si effectivement ce procédé fonctionne.

Préliminaire.
On considère un couple de variables aléatoires (X, Y) définies sur $(\Omega, \mathcal{F}, \mathbb{P})$. Soit k et l deux entiers strictement positifs. On suppose que X prend q valeurs réelles : $x_1, ..., x_q$, et Y, r valeurs réelles : $y_1, ..., y_r$. Soit g une fonction réelle. Pour tout $1 \leq j \leq q$, soit

$$E_j = \sum_{i=1}^{r} g(y_i)\mathbb{P}[Y = y_i | X = x_j].$$

Démontrer que

$$\mathbb{E}[g(Y)] = \sum_{j=1}^{q} E_j \mathbb{P}[X = x_j]. \tag{3.1}$$

On passe maintenant au problème en question. Soit $k \in \mathbb{N}$, on note T_k le nombre de copies infectées obtenues parmi les n tirées dans la génération k, pour constituer la génération $k + 1$

1) Quelles sont les valeurs prises par T_k ? Pour j décrivant l'ensemble de ces valeurs, déterminer la loi conditionnelle de T_{k+1} sachant $T_k = j$.

2) En déduire, à l'aide de (3.1), une relation entre $\mathbb{E}[T_{k+1}]$ et $\mathbb{E}[T_k]$ et montrer que pour tout k, $\mathbb{E}[T_k] = np$.

3) On considère maintenant la variable $Z_k = T_k(n - T_k)$.
  - *a) En utilisant le préliminaire avec (T_k, T_{k+1}) et une fonction g convenablement choisie. Montrer que la suite de terme général $\mathbb{E}[Z_k]$ est géométrique.*
  - *b) Donner l'expression de $\mathbb{E}[Z_k]$ en fonction de p, n et k.*
  - *c) Montrer que la suite $(\mathbb{E}[Z_k], k \geq 0)$ tend vers 0.*

4 a) Que signifie concrètement l'événement $Z_k = 0$?
  - *b) Quelle est la plus petite valeur de $j(n - j)$ lorsque j parcourt $\{1, 2, ..., n - 1\}$. ?*
  - *c) En déduire, à l'aide de 3-c), que*

$$\mathbb{P}(0 < T_k < n) \underset{k \to \infty}{\longrightarrow} 0.$$

  - *d) Calculer la limite de $\mathbb{P}[Z_k = 0]$ et interpréter le résultat.*

e) Déterminer en fonction de n uniquement, une valeur de k telle que $\mathbb{P}[Z_k = 0] > 0,99$

5) On cherche maintenant à calculer la probabilité de l'événement F=« A partir d'un certain rang ; les générations ne sont constituées que de logiciels infectées ».

a) Montrer que la suite d'événements $(\{T_k = n\})_{k\geq 1}$ est croissante pour l'inclusion.

b) Que représente alors $\mathbb{P}[F]$ pour la suite, $(\mathbb{P}[T_k = n])_{k\geq 1}$?

6) Soit $j \in [|1, n-1|]$, déterminer la limite de $\mathbb{P}[T_k = j]$ quand k tend vers l'infini.

7) En utilisant le résultat de la question 2), donner la valeur de $\mathbb{P}[F]$.

8) On cherche maintenant à calculer la probabilité de l'événement G=« A partir d'un certain rang, les générations ne sont constituées que de copies saines. » Calculer $\mathbb{P}[G]$, que pensez vous de la méthode du statisticien ?

Séries génératrices

Révision : séries

Exercice 3.4.46 (#) *Déterminer le rayon de convergence des séries entières suivantes :*

1. $\sum_n \frac{n!}{(2n)!} x^n$
2. $\sum_n \frac{\sqrt{n}}{2^n+1} x^{2n}$
3. $\sum_n \frac{(1+i)^n z^{3n}}{n2^n}$
4. $\sum_n n^{\ln n} z^n$

Solution. On utilise le résultat suivant : pour une série $\sum_n a_n X^n$, si la limite

$$\lim_{n\to+\infty} \left|\frac{a_{n+1}}{a_n}\right| = L$$

existe (limite éventuellement infinie), alors son rayon de convergence est égal à $1/L$.

1. Posons $a_n = \frac{n!}{(2n)!}$. Alors

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n+1}{(2n+1)(2n+2)} \xrightarrow[n\to+\infty]{} 0.$$

Le rayon de convergence est donc $\boxed{R = +\infty}$.

2. Posons $a_n = \frac{\sqrt{n}}{2^n+1}$. Alors

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\sqrt{n+1}(2^n+1)}{\sqrt{n}(2^{n+1}+1)} \xrightarrow[n\to+\infty]{} 1/2.$$

Le rayon de convergence de la série $\sum_n a_n X^n$ est donc $R = 2$ et celui de la série considérée ici est $\boxed{R = \sqrt{2}}$ (prendre $X = x^2$).

3. Comme $|1+i| = \sqrt{2}$, on a

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n}{n+1}\frac{\sqrt{2}}{2} \xrightarrow[n\to+\infty]{} 1/\sqrt{2}.$$

Le rayon de convergence de la série $\sum_n a_n X^n$ est donc $R = \sqrt{2}$ et celui de la série considérée ici est $\boxed{R = (\sqrt{2})^{1/3} = 2^{1/6}}$.

4.
$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{(n+1)^{\ln(n+1)}}{n^{\ln n}} = \exp\left((\ln(n+1))^2 - (\ln n)^2\right).$$

Or,

$$(\ln(n+1))^2-(\ln n)^2 = (\ln(n+1)-\ln n)(\ln(n+1)+\ln n) = \ln(1+1/n)\ln(n+n^2) \underset{n\to+\infty}{\sim} \frac{1}{n}\ln(n+n^2)$$

donc

$$\left|\frac{a_{n+1}}{a_n}\right| \underset{n\to+\infty}{\longrightarrow} e^0 = 1$$

et $\boxed{R=1}$.

Exercice 3.4.47 (#) *Pour $n \in \mathbb{N}$ on définit sur $[0;1], u_n = X^n(1-X)$.*

1. *Montrer que $\sum_n u_n$ converge simplement sur $[0;1]$ et calculer $\sum_n u_n(x)$ pour tout $x \in [0;1]$.*
2. *Montrer que $\sum_n u_n$ ne converge pas normalement sur $[0;1]$*
   (a) *en utilisant la définition,*
   (b) *en utilisant le Théorème de continuité des séries.*
3. *Vérifier que $\sum_n u_n$ converge normalement sur $[0;a]$ pour tout $a \in]0;1[$.*

Solution.

1. La série que l'on étudie est à termes positifs ou nuls. Pour montrer que $\sum_n u_n$ converge simplement sur $[0;1]$, montrons que pour $x \in [0,1]$, la limite
$$\lim_{N\to+\infty} \sum_{n=0}^{N} x^n(1-x)$$
existe. Si $x=1$, $\sum_{n=0}^{N} u_n(1) = 0$ et on a donc la convergence voulue. Si $0 \le x < 1$,
$$\sum_{n=0}^{N} u_n(x) = \sum_{n=0}^{N} x^n(1-x) = (1-x)\frac{1-x^{N+1}}{1-x} = 1 - x^{N+1} \underset{N\to+\infty}{\longrightarrow} 1.$$
La série $\sum_n u_n$ converge donc simplement sur $[0;1]$ et sa somme S est définie par $S(1)=0$ et $S(x)=1$ si $x \in [0,1[$.
2. (a) Une série converge uniformément si
$$\sup_{x\in[0,1]} \left|\sum_{n=0}^{N} u_n(x) - S(x)\right| \underset{N\to+\infty}{\longrightarrow} 0.$$
Cherchons la borne supérieure de la fonction $|\sum_{n=0}^{N} u_n - S|$ sur $[0,1]$: on a
$$\left|\sum_{n=0}^{N} u_n(x) - S(x)\right| = \begin{cases} |1-x^{N+1}-1| = x^{N+1} & \text{si } x<1 \\ 0 & \text{si } x=1 \end{cases}.$$
La borne supérieure qui vaut donc 1 ne converge pas vers 0 : la convergence n'est pas uniforme.

(b) On peut aussi le montrer en utilisant la propriété suivante : si une série de fonctions continues converge uniformément, alors sa somme est également continue.

Ici, toutes les fonctions u_n pour $n \geq 0$ sont continues sur $[0,1]$ mais S n'est pas continue en 1 : la convergence ne peut donc pas être uniforme.

3. Soit $0 < a < 1$. Dire que la convergence est normale sur $[0,a]$ revient à dire que la série

$$\sum_n \sup_{x\in[0,a]} |u_n(x)|$$

converge. Pour cela, cherchons un équivalent de $\sup_{x\in[0,a]} |u_n(x)|$ quand $n \to +\infty$. En étudiant les variations de la fonction u_n, on voit qu'elle est croissante sur $[0, \frac{n}{n+1}]$ puis décroissante sur $[\frac{n}{n+1}, 1]$. Comme $n/(n+1)$ tend vers 1 et comme $a < 1$, pour n assez grand on a $n/(n+1) > a$. La fonction u_n est alors croissante sur $[0,a]$ et atteint donc son maximum en a :

$$\sup_{x\in[0,a]} |u_n(x)| = a^n(1-a) \text{ pour } a \text{ assez grand}$$

et la série de terme général $(a^n(1-a))_n$ est convergente puisque $a < 1$ (voir question a)) et la série $\sum_n u_n$ converge normalement sur $[0,a]$.

Séries génératrices

Exercice 3.4.48 (#) *Calculer à l'aide de la fonction génératrice l'espérance et la variance de la loi géométrique de paramètre p et de la loi de Poisson de paramètre λ.*

Solution. On traite d'abord le cas où X est une v.a. de Poisson de paramètre λ :

$$\mathbb{P}(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}, \quad k \geq 0.$$

Pour $s \in]-1,1[$, on a

$$G_X(s) = \mathbb{E}\left[s^X\right] = \sum_{k\geq 0} s^k \mathbb{P}(X=k) = \sum_{k\geq 0} s^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k\geq 0}\frac{(s\lambda^k)}{k!} = e^{-\lambda}e^{s\lambda} = \boxed{e^{\lambda(s-1)}}.$$

Comme l'espérance de X existe (puisque la série de terme général $(k\mathbb{P}(X=k))_k$ converge absolument), on a

$$\mathbb{E}[X] = \lim_{s\to 1,\ s<1} G'_X(s).$$

Or, $G'_X(s) = \lambda e^{\lambda(s-1)}$ et donc $\boxed{\mathbb{E}[X] = \lambda}$. De même, $\mathbb{E}[X^2]$ existe et alors

$$\mathbb{E}[X(X-1)] = \lim_{s\to 1,\ s<1} G''_X(s) = \lim_{s\to 1,\ s<1} \lambda^2 e^{\lambda(s-1)} = \lambda^2.$$

On a donc $\mathbb{V}(X) = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - \mathbb{E}[X]^2 = \lambda^2 + \lambda - \lambda^2 = \boxed{\lambda}$.
On procède de la même manière si X est une v.a. géométrique : si $-1 < s < 1$,

$$G_X(s) = \sum_{k\geq 0} s^k\mathbb{P}(X=k) = \sum_{k\geq 1} s^k p(1-p)^{k-1} = ps\sum_{k\geq 1}(s(1-p))^{k-1} = \boxed{\frac{sp}{1-s(1-p)}}$$

(la dernière égalité est valide car $(1-p)s < 1$).
En dérivant, on obtient comme avant $\boxed{\mathbb{E}[X] = 1/p}$ et $\boxed{\mathbb{V}(X) = (1-p)/p^2}$.

Exercice 3.4.49 (#) *Soit X et Y deux variables aléatoires indépendantes de lois de Poisson de paramètres respectifs $\lambda > 0$ et $\gamma > 0$. Déterminer, en calculant sa fonction génératrice, la loi de $X+Y$.*

Solution. Pour $s \in]-1,1[$, on a

$$G_{X+Y}(s) = \mathbb{E}[s^{X+Y}] = \mathbb{E}[s^X s^Y] = \mathbb{E}[s^X]\mathbb{E}[s^Y] = G_X(s)G_Y(s)$$

car les 2 v.a. sont indépendanntes. Or, $G_X(s) = e^{\lambda(s-1)}$ et $G_Y(s) = e^{\gamma(s-1)}$ donc

$$G_{X+Y}(s) = e^{(\lambda+\gamma)(s-1)}.$$

La série génératrice caractérisant la loi d'une v.a., on en déduit que $X+Y$ est une v.a. de Poisson de paramètre $\lambda+\gamma$.

Exercice 3.4.50 (#) *Soit $(X_k)_{1\le k\le n}$ une suite de v.a. indépendantes telles que pour $1 \le k \le n$, X_k suit une loi uniforme discrète sur $\{1,\dots k\}$, c'est-à-dire que pour $i \in \{1,\dots k\}$,*

$$P(X_k = i) = \frac{1}{k}.$$

Déterminer la fonction génératrice de X_k et celle de $Y_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$.

Solution. Commençons par la série génératrice de X_k qui est une v.a. uniforme sur $\{1,\dots,k\}$:

$$\mathbb{P}(X_k = i) = \frac{1}{k}, \quad 1 \le i \le k.$$

Pour $s \in]-1,1[$, on a

$$G_{X_k}(s) = \sum_{i=1}^{n} s^i \mathbb{P}(X_n = i) = \frac{1}{k}\sum_{i=1}^{n} s^i = \boxed{\frac{s}{k}\frac{1-s^k}{1-s}}.$$

Calculons maintenant la série génératrice de $Y_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$. On a

$$G_{Y_n}(s) = \mathbb{E}\left[s^{Y_n}\right] = \mathbb{E}\left[\left(s^{1/n}\right)^{X_1+\cdots+X_n}\right] = \prod_{k=1}^{n}\mathbb{E}\left[\left(s^{1/n}\right)^{X_k}\right]$$

car les v.a. $(X_k)_k$ sont indépendantes. On a donc

$$G_{Y_n}(s) = \prod_{k=1}^{n} G_{X_k}(s^{1/n}) = \prod_{k=1}^{n}\frac{s^{1/n}}{k}\frac{1-s^{k/n}}{1-s^{1/n}} = \boxed{\frac{s}{\left(1-s^{1/n}\right)^n n!}\prod_{k=1}^{n}\left(1-s^{k/n}\right)}$$

Exercice 3.4.51 *Soit n fixé, et $(X_i)_{i=1,\dots,n}$ une suite de v.a. de Bernoulli indépendantes de paramètre $p_n = \lambda/n$; $X_i = 1$ modélise que le i-ème assuré subit un sinistre. Le nombre d'assurés subissant un sinistre est donc $Y_n = X_1 + \ldots + X_n$. On suppose que les X_i sont indépendants ; le fait que p_n est petit avec n modélise que le risque de sinistre pour chaque assuré est petit devant le nombre d'assurés, λ représentant le "nombre d'assurés sinistrés espéré".*

1. *Soit $Z \sim P(\lambda)$. Calculer la série génératrice G_Z de Z.*

2. *Calculer la série génératrice* G_{X_i} *de* X_i.
3. *Calculer la série génératrice* G_{Y_n} *de* Y_n.
4. *Calculer* $\lim_{n\to+\infty} G_{Yn}$. *On rappelle que* $\ln(1+u) = u(1+\epsilon(u))$, *où* $lim_{u\to 0}\epsilon(u) = 0$. *Conclusion ?*

Exercice 3.4.52 (#) *(Somme aléatoire) Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de v.a. iid à valeurs dans* $\mathbb{N}$ *et* T *une variable aléatoire, à valeurs dans* $\mathbb{N}$, *et indépendantes des* $(X_n)_{n\in\mathbb{N}}$. *On pose* $S_0 = 0$ *et pour tout* $n \geq 1$, $S_n = \sum_{k=1}^n X_k$ *et* $S(w) = S_{T(w)}(w)$.

1. *Montrer que* $G_S = G_T \circ G_{X_1}$.
2. *Montrer que si* X_1 *et* T *ont une espérance, alors l'espérance de* S *existe et* $\mathbb{E}[S] = \mathbb{E}[T]\mathbb{E}[X]$.
3. *On prend pour les* $(X_n)_{n\in\mathbb{N}}$ *des v.a. iid de loi de Bernoulli de paramètre* p, *et pour* T *une loi de Poisson de paramètre* λ. *Quelle est la loi de* S *?*

Solution.

1. On a

$$G_S(s) = \sum_{k\geq 0} s^k \mathbb{P}(S=k) = \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(S=k, T=n) = \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(S_n=k, T=n).$$

De plus, comme $S_n = X_1 + X_2 + \cdots + X_n$ et que T est indépendant des v.a. $X_1, \dots, X_n$, on a

$$G_S(s) = \sum_{k\geq 0} s^k \sum_{n\geq 0} \mathbb{P}(S_n=k)\mathbb{P}(T=n).$$

On veut intervertir les deux signes de sommation. Pour cela, il faut montrer que la série double converge absolument. Puisque $|s| < 1$,

$$\sum_{k\geq 0}\sum_{n\geq 0} \left|s^k \mathbb{P}(S_n=k)\mathbb{P}(T=n)\right| \leq \sum_{k\geq 0}\sum_{n\geq 0} \mathbb{P}(S_n=k)\mathbb{P}(T=n) = \sum_{k\geq 0} \mathbb{P}(S=k) = 1 < +\infty$$

(on a refait les mêmes calculs que précédemment mais dans l'autre sens). On peut intervertir les deux sommes, on a alors

$$G_S(s) = \sum_{n\geq 0} \mathbb{P}(T=n) \sum_{k\geq 0} s^k \mathbb{P}(S_n=k) = \sum_{n\geq 0} \mathbb{P}(T=n) G_{S_n}(s).$$

Or, S_n est la somme de n variables indépendantes et de même loi, on a donc

$$G_{S_n}(s) = \prod_{i=1}^n G_{X_i}(s) = (G_{X_1}(s))^n$$

et

$$G_S(s) = \sum_{n\geq 0} \mathbb{P}(T=n)\left(G_{X_1}(s)\right)^n = G_T(G_{X_1}(s)),$$

ce qui est la relation à trouver.

2. On justifie l'existence de $\mathbb{E}[S]$ en montrant que la suite $\left(\sum_{k=0}^{K} k\mathbb{P}(S=k)\right)_{K\geq 0}$ est bornée et alors la série $\sum_{k\geq 0} k\mathbb{P}(S=k)$ est absolument convergente.
Comme $\mathbb{E}[S]$ existe, on sait que $\mathbb{E}[S] = G'_S(1)$. En utilisant a), on a
$$\mathbb{E}[S] = G'_{X_1}(1)G'_T(G_{X_1}(1)) = \mathbb{E}[X_1]G'_T(1) = \mathbb{E}[X_1]\mathbb{E}[T]$$
On a utilisé que pour toute série génératrice, $G(1) = 1$.
3. On a dans ce cas particulier $G_T(s) = e^{\lambda(s-1)}$ et $G_{X_1}(s) = ps + 1 - p$. En utilisant la relation trouvée à la question a), on a $G_S(s) = e^{\lambda(ps+1-p-1)} = e^{p\lambda(s-1)}$ et donc S est une v.a. de Poisson de paramètre $p\lambda$.

Chapitre 4

Variables aléatoires à densité

Ebauche de motivation : la planche de Galton.

On répète plusieurs fois l'expérience suivante : on laisse tomber une bille sur une planche

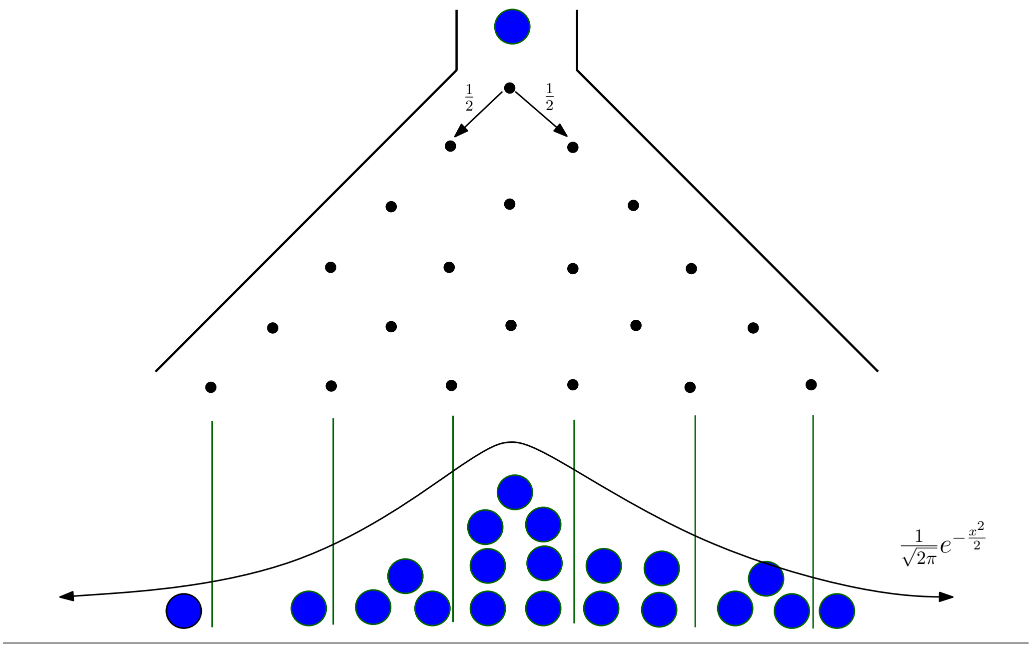


FIGURE 4.1 – Planche de Galton

inclinée munie de clous. Lorsque la bille rebondit sur le clou, elle a une chance sur deux de tomber à droite ou à gauche. On récolte les billes et observe leurs répartitions. Ce que Galton observe avec son expérience est la tendance des billes à se répartir d'une façon particulière. De façon heuristique, lorsque l'on augmente le nombre de billes, l'histogramme se rapproche d'une courbe. Il s'agit de la courbe de Gauss, et cette convergence est appelée le théorème de *de Moivre-Laplace*. Il apparait ainsi qu'à partir d'expériences simples, on passe d'un « monde discret » à « un monde continu ».

Mathématiquement, le phénomène observé par Galton s'exprime ainsi :

$$\mathbb{P}\left[2\sqrt{n}\left(\frac{S_n}{n}-\frac{1}{2}\right)\leq z\right]\longrightarrow\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}e^{-\frac{u^2}{2}}du.$$

Où S_n suit une loi binomiale de paramètres $(n, \frac{1}{2})$. Ces résultats de convergence seront étudiées dans le chapitre 5. Nous établissons dans un premier temps la notion de densité.

4.1 Généralités

Définition 4.1.1 *Soit X une variable aléatoire réelle, on dit que X admet une densité f si sa fonction de répartition F est*
- *continue*
- *de la forme :*

$$F(x) = \int_{-\infty}^{x} f(t)dt, \text{ pour tout } x \in \mathbb{R}$$

avec
1. *f est positive*
2. *f possède un nombre fini de points de discontinuité*
3. *$\int_{-\infty}^{\infty} f(x)dx = 1$.*

Remarque 4.1.1 *La fonction f n'est pas unique, puisque l'on peut changer sa valeur en un point sans changer la valeur de l'intégrale. Néanmoins nous parlerons de la densité.*

Proposition 4.1.1 *En tout point x_0 où f est continue, F est dérivable et*

$$F'(x_0) = f(x_0).$$

Savoir faire la démonstration de cette propriété est très important. La propriété correspond en fait au théorème fondamental de l'analyse reliant intégrale et dérivée.

Démonstration. On écrit le taux d'accroissement en x_0 de la fonction F :

$$\frac{F(x_0 + h) - F(x_0)}{h} = \frac{1}{h}\int_{x_0}^{x_0+h} f(t)dt$$

Soit $\epsilon > 0$, par continuité de f en x_0, il existe h tel que si $|x - x_0| \leq h$ alors $|f(x) - f(x_0)| \leq \epsilon$. On écrit donc

$$\begin{aligned}\left|\frac{F(x_0 + h) - F(x_0)}{h} - f(x_0)\right| &= \left|\frac{1}{h}\int_{x_0}^{x_0+h} (f(x) - f(x_0))\, dx\right| \\ &\leq \frac{1}{h}\int_{x_0}^{x_0+h} |f(x) - f(x_0)|dx \\ &\leq \frac{1}{h} h\epsilon \leq \epsilon\end{aligned}$$

Si h tend vers 0, le taux d'accroissement de F en x_0 admet une limite et sa limite vaut $f(x_0)$. Par définition de la dérivée (en tant que limite du taux d'accroissement) :

$$F'(x_0) = f(x_0).$$

□ La notion d'intégrale généralisée est sous-jacente à l'étude des probabilités continues. Une piqûre de rappel concernant les intégrales généralisées peut être utile...

4.1.1 Loi et fonction de répartition

Proposition 4.1.2 *Soit X une variable aléatoire réelle admettant une densité f. Pour tout $(a,b) \in \mathbb{R}^2$ avec $a \leq b$*

$$\mathbb{P}[a < X \leq b] = \int_a^b f(t)dt$$

Démonstration. On a $F(b) = \mathbb{P}[X \leq b]$. l'événement $\{a < X \leq b\}$ correspond à $\{X \leq b\} \setminus \{X \leq a\}$. Donc

$$\begin{aligned}
\mathbb{P}[a < X \leq b] &= F(b) - F(a) \\
&= \int_{-\infty}^b f(t)dt - \int_{-\infty}^a f(t)dt \\
&= \int_a^b f(t)dt \text{ par la relation de Chasles}
\end{aligned}$$

Proposition 4.1.3 *Pour une variable à densité, $\mathbb{P}[X = a] = 0$ pour tout $a \in \mathbb{R}$.*

Démonstration. La probabilité de l'événement $\{X = a\}$ est la limite à gauche de $F(a) - F(x)$ lorsque x tend vers a. Par continuité de F, $F(a) - F(x) \underset{x \to a; x < a}{\longrightarrow} 0$. □

4.1.2 Espérance

Définition 4.1.2 *Soit X une v.a.r de densité f. On appelle espérance de X, le réel*

$$\mathbb{E}[X] := \int_{-\infty}^{+\infty} tf(t)dt$$

définie uniquement si cette intégrale converge absolument.

Il faut voir l'*analogie* entre cette définition et la définition de l'espérance d'une variable discrète :

$$E[X] = \sum_{i \in I} x_i \mathbb{P}[X = x_i].$$

La série est remplacée par l'intégrale. Dans le cas discret, c'est $\mathbb{P}[X = x_i]$ qui donne le poids de probabilité de l'événement élémentaire $\{X = x_i\}$. Dans le cas continu, ces éléments élémentaires sont donnés par la dérivée de la fonction de répartition $F'(x)$ c'est à dire $f(x)$ (en dehors d'une famille finie de points).

Contrairement au cadre discret, les variables aléatoires à densité ne forment pas un espace vectoriel : Soient X et Y à densité, il n'y a aucune raison a priori que $X + Y$ ait une densité. Un exemple trivial est donnée par X et $Y = -X$: si X a une densité alors il en est de même pour Y (le démontrer). Dans ce cas : $X + Y = 0$ presque sûrement. On dit que la loi de $X + Y$ est *dégénérée*, $\mathbb{P}[X + Y = 0] = 1$.

Remarque 4.1.2 (Où est l'univers ?) *D'aucun aura remarqué, que l'univers Ω n'apparait que rarement dans cette partie où l'on traite des variables à densité. Nous avons déjà vu* *qu'une fois les variables aléatoires discrètes définies sur un espace abstrait Ω, on peut en pratique tout calculer en considérant sa loi : c'est à dire en travaillant sur l'espace probabilisé $(X(\Omega), \mathcal{B}(\mathbb{R}), P_X)$. Le théorème 3.1.3 nous dit que*

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

La théorie de la mesure permet de définir une notion d'intégrale sur un espace Ω général contre toute probabilité $\mathbb{P}$, et montre que pour toute variable aléatoire réelle définie sur un triplet $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\mathbb{E}[X] = \int_\Omega X(\omega)\mathrm{d}\mathbb{P}(\omega) \tag{4.1}$$

On énonce les théorèmes fondamentaux suivants sans donner de démonstration.

Théorème 4.1.4 *Soient X et Y deux variables à densité admettant une espérance, soit λ un réel. La variable aléatoire $X + \lambda Y$ admet une espérance (mais pas forcément une densité !) et*

$$\mathbb{E}[X + \lambda Y] = \mathbb{E}[X] + \lambda\mathbb{E}[Y]$$

Une façon de démontrer ce résultat est de travailler avec un couple de variables aléatoires dont les lois marginales correspondent à la loi de X et la loi de Y. Cela nécessiterait de travailler avec des intégrales multiples, nous ferons ce genre de calculs lorsque l'on étudiera les vecteurs aléatoires.

4.1.3 Composition d'une variable aléatoire et d'une fonction et indépendance

Théorème 4.1.5 (Théorème de transfert) *Soient X une variable aléatoire réelle de densité f et g une fonction continue par morceaux réelle telle que $|g|f$ soit intégrable sur R. Alors $g(X)$ est une variable aléatoire réelle et possède une espérance :*

$$\mathbb{E}[g(X)] = \int_{-\infty}^{+\infty} g(x)f(x)dx.$$

Corollaire 4.1.6 *Pour tout A borélien,*

$$\mathbb{E}[\mathbf{1}_A(X)] = \int_A f(x)dx$$

En fait, la variable aléatoire $\mathbf{1}_A(X)$ est une variable de Bernoulli de paramètre $\int_A f(x)dx$.

A priori, le calcul de la loi de la v.a.r $g(X)$ n'est pas dans le programme ; mais il faut comprendre que cela correspond à un changement de variable "dans f". Il peut arriver lors d'une épreuve mélangeant analyse et probabilité, que vous soyez amenés à faire ce genre de calcul. On mentionne donc que si g est C^1 strictement croissante et d'inverse C^1 (en fait il suffit que g soit un difféomorphisme) alors $g(X)$ est à densité. En effet, soit F_g sa fonction de répartition,

$$F_g(x) := \mathbb{P}[g(X) \leq x] = \mathbb{P}[X \leq g^{-1}(x)] = \int_{-\infty}^{g^{-1}(x)} f(t)dt.$$

Finalement,

$$F_g'(x) = f(g^{-1}(x))(g^{-1})'(x) = f(g^{-1}(x))(g' \circ g^{-1}(x))^{-1}.$$

4.1.4 Moments, variance et écart-type

Définition 4.1.3 (moment) *Soit $k \in \mathbb{N}^\star$. On appelle moment d'ordre k l'espérance si elle existe de la variable X^k. D'après le théorème de transfert, c'est donc*

$$\mathbb{E}[X^k] = \int_{-\infty}^{\infty} x^k f(x)dx.$$

Remarque 4.1.3 *L'ensemble des variables aléatoires à densité ne forme pas un sous-espace vectoriel de $\mathcal{L}^0$, contrairement aux variables discrètes ! L'ensemble des variables aléatoires définies sur un même espace Ω muni d'une tribu $\mathcal{F}$, $\mathcal{L}^0$ contient les deux cas : discret et à densité. Il faut savoir qu'il contient beaucoup plus ; il y a par exemple des variables non discrètes sans densité...*

Définition 4.1.4 (Variance) *On appelle variance de X l'espérance de la variable $(X-\mathbb{E}[X])^2$:*

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

Par le théorème de transfert :

$$\mathbb{V}(X) = \int_{-\infty}^{+\infty} (x - \mathbb{E}[X])^2 f(x)\mathrm{d}x$$

Pour calculer $\mathbb{V}$, on utilise souvent la formule :

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

4.2 Lois à densité usuelles

Loi uniforme

X suit une loi uniforme sur le segment $[a, b]$ si sa densité est donnée par :

$$f(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(x).$$

On a $F(x) = \frac{x-a}{b-a}$ pour $x \in [a, b]$. Les moments sont :

$$\mathbb{E}[X] = \frac{a+b}{2} \text{ et } \mathbb{V}(X) = \frac{(b-a)^2}{12}.$$

Loi exponentielle

Une variable aléatoire X suit une loi exponentielle de paramètre θ si sa densité est

$$f(x) = \theta e^{-\theta x}\mathbf{1}_{\mathbb{R}^+}(x).$$

On a

$$\mathbb{E}[X] = \frac{1}{\theta} \text{ et } \mathbb{V}(X) = \frac{1}{\theta^2}.$$

La loi exponentielle est un cas particulier de la loi suivante.

Loi Gamma

On rappelle la définition de la fonction Γ pour tout réel $p > 0$:

$$\Gamma(p) = \int_0^\infty e^{-x} x^{p-1} dx$$

Une variable aléatoire suit une loi Gamma de paramètre $\theta > 0$ et $p > 0$ (notée $\Gamma(p, \theta)$) si sa densité est :

$$f(x) = \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{p-1} \mathbf{1}_{\mathbb{R}^+}(x).$$

Proposition 4.2.1

$$\mathbb{E}[X^r] = \frac{\Gamma(p+r)}{\theta^r \Gamma(p)}$$

Démonstration.

$$\mathbb{E}[X^r] = \int_0^\infty \frac{\theta^p}{\Gamma(p)} e^{-\theta x} x^{r+p-1} dx = \int_0^\infty \frac{\theta^p}{\Gamma(p)} e^{-u} \left(\frac{u}{\theta}\right)^{r+p-1} \frac{1}{\theta} du$$

On obtient donc

$$\mathbb{E}[X^r] = \frac{1}{\theta^r} \frac{\Gamma(p+r)}{\Gamma(p)}.$$

□

Les propriétés de la fonction Gamma sont très importantes. On en rappelle quelques unes, sans donner de démonstration. **Néanmoins, vous devez savoir étudier la fonction** Γ.

- $\Gamma(p) = (p-1)\Gamma(p-1)$
- $\Gamma(p) = (p-1)!$ si p est entier
- $\Gamma(1/2) = \sqrt{\pi}$
- $\Gamma(p+1) \underset{p \to \infty}{\sim} \left(\frac{p}{e}\right)^p \sqrt{2\pi p}(1 + \frac{1}{12p})$

Loi normale

Une variable aléatoire X suit la loi gaussienne $\mathcal{N}(m, \sigma^2)$, où $m \in \mathbb{R}$, $\sigma \geq 0$, si sa loi a pour densité sur $\mathbb{R}$: $\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$ quand $\sigma > 0$, et si $X = m$ quand $\sigma = 0$.

Remarque 4.2.1 *Il faut impérativement savoir démontrer que*

$$\int_{-\infty}^\infty \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x = \sqrt{2\pi}.$$

Il n'existe pas de forme explicite à l'aide de fonctions usuelles de $x \mapsto \int_0^x e^{-\frac{u^2}{2}} du$.

Calculons les moments :

1) $\mathbb{E}[X] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-m)^2}{2\sigma^2}} \mathrm{d}x$. On pose $u = \frac{x-m}{\sigma}$, on obtient

$$\mathbb{E}[X] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma u + m) e^{-u^2/2} \sigma \mathrm{d}u = m + \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} u e^{-u^2/2} \sigma \mathrm{d}u = m$$

car la fonction à intégrer dans la partie droite est impaire.

2)

$$\begin{aligned}\mathbb{E}[X^2] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-m)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (\sigma u + m)^2 e^{-u^2/2} \sigma du \\ &= m^2 + \frac{2\sigma^2}{\sqrt{\pi}} \Gamma(3/2).\end{aligned}$$

On a $\Gamma(3/2) = \frac{1}{2}\Gamma(1/2) = \frac{\sqrt{\pi}}{2}$, et donc

$$\mathbb{E}[X^2] = m^2 + \sigma^2.$$

Si X est centrée réduite , c'est-à-dire lorsque $m = 0$ et $\sigma^2 = 1$, la loi normale s'appelle **gaussienne standard**.

On a les propriétés intéressantes (mais non-exigibles) suivantes.

Exercice 4.2.1 (#) Montrer les assertions suivantes :

1) $\mathbb{P}(X > t) \leq \frac{1}{2} e^{-t^2/2}$.

2) $\mathbb{P}(X > t) \underset{t \to \infty}{\sim} \frac{1}{\sqrt{2\pi} t} e^{-t^2/2}$.

3) $\mathbb{E}[e^{zX}] = e^{z^2/2}$ pour tout $z \in \mathbb{R}$.

Solution. Preuve de 1)

$$\begin{aligned}\mathbb{P}[X > t] &= (2\pi)^{-1/2} \int_t^{\infty} e^{-x^2/2} dx \\ &= (2\pi)^{-1/2} \int_0^{\infty} e^{-(x+t)^2/2} dx \\ &\leq (2\pi)^{-1/2} e^{-t^2/2} \int_0^{\infty} e^{-x^2/2} dx = \frac{1}{2} e^{-t^2/2}.\end{aligned}$$

Preuve de 2)

$$\mathbb{P}[X > t] = \frac{(2\pi)^{-1/2} e^{-t^2/2}}{t} \int_0^{\infty} e^{-y} e^{-y^2/2t^2} dy$$

et cette intégrale tend vers $\int_0^{\infty} e^{-y} dy = 1$.

Preuve de 3)

$$\begin{aligned}\mathbb{E}[e^{zX}] &= (2\pi)^{-1/2}\int_{-\infty}^{\infty} e^{zx-x^2/2}dx\\ &= (2\pi)^{-1/2}e^{z^2/2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-z)^2}dx\\ &= (2\pi)^{-1/2}e^{z^2/2}\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-z)^2}dx = e^{z^2/2}.\end{aligned}$$

Loi Bêta

Une variable aléatoire suit une loi Bêta de paramètre $\alpha > 0$ et $\beta > 0$ si sa densité est :

$$f(x) = \frac{1}{B(\alpha,\beta)}x^{\alpha-1}(1-x)^{\beta-1}\mathbf{1}_{[0,1]}(x),$$

où

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx.$$

Exercice 4.2.2 Démontrer que

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

Nous terminons ce catalogue avec une dernière loi un peu moins classique.

Loi de Weibull

Une variable aléatoire suit une loi Weibull de paramètre $\lambda > 0$ et $k > 0$ si sa densité est :

$$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(x/\lambda)^k}\mathbf{1}_{\mathbb{R}^+}(x).$$

Cette loi est utilisée dans de nombreux domaines (par exemple en fiabilité, en logistique et en météo).

Proposition 4.2.2 – *La fonction de répartition de la loi de Weibull est*

$$F(x) = 1 - e^{-(x/\lambda)^k} pour x \geq 0 \ \text{et}\ 0 \ \text{si}\ x < 0.$$

- $\mathbb{E}[X] = \lambda\Gamma(1+1/k)$
- $\mathbb{V}[X] = \lambda^2\left[\Gamma\left(1+\frac{2}{k}\right) - \left(\Gamma\left(1+\frac{1}{k}\right)\right)^2\right]$

Démonstration. Exercice.

Remarque 4.2.2 *Il est intéressant de voir le rôle des paramètres dans la forme de la densité.*

4.3 Vecteurs aléatoires à densité

Définition 4.3.1 *Un vecteur aléatoire X de Ω dans $\mathbb{R}^p$ admet une densité s'il existe une fonction $f : \mathbb{R}^p \to \mathbb{R}$ positive, intégrable et d'intégrale valant 1, telle que : $\forall\ (x_1, ..., x_p) \in \mathbb{R}^p$*

$$\mathbb{P}[X_1 \leq x_1, X_2 \leq x_2, ..., X_p \leq x_p] = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}...\int_{-\infty}^{x_p} f(u_1,u_2,...,u_p)du_1...du_p$$

4.3.1 Vecteurs aléatoires à densité : lois marginales, indépendance

Proposition 4.3.1 *Soit X un vecteur aléatoire à densité. Les variables aléatoires $X_1, ..., X_p$ sont alors également à densité. Une densité pour X_i est*

$$f_{X_i} : x \mapsto \int_{-\infty}^{\infty} f(u_1, ..., u_{i-1}, x, u_{i+1}, ..., u_p) du_1 ... du_{i-1} du_{i+1} ... du_p$$

Définition 4.3.2 *Les fonctions f_{X_i}, pour $i = 1, ..., p$ sont appelées densités marginales du vecteur X.*

Théorème 4.3.2 (Théorème de factorisation) *Soit $X = (X_1, ..., X_p)$ un vecteur aléatoire. Les variables aléatoires $X_1, ...X_p$ sont indépendantes si et seulement si le vecteur X a une densité f à variables séparées : c'est à dire :*

$$f(x_1, ..., x_p) = f_1(x_1)...f_p(x_p).$$

Dans ce cas, on a : $f_i = \lambda_i f_{X_i}$ avec $\prod \lambda_i = 1$.

Démonstration. On démontre que si la densité s'écrit sous forme de produit alors les variables sont indépendantes :

$$\begin{aligned}
\mathbb{P}[X_1 \leq x_1, ..., X_p \leq x_p] &= \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_p} f(u_1, ..., u_p) du_1 ... du_p \\
&= \int_{-\infty}^{x_1} ... \int_{-\infty}^{x_p} f_1(u_1)...f_p(u_p) du_1 ... du_p \\
&= \frac{1}{\lambda_1} \int_{-\infty}^{x_1} f_1(u_1) du_1 .. \frac{1}{\lambda_p} \int_{-\infty}^{x_p} f_p(u_p) du_p \\
&= \int_{-\infty}^{x_1} f_{X_1}(u_1) du_1 ... \int_{-\infty}^{x_p} f_{X_p}(u_p) du_p \\
&= \mathbb{P}[X_1 \leq x_1] ... \mathbb{P}[X_p \leq x_p]
\end{aligned}$$

4.3.2 Loi normale multidimensionnelle

Définition 4.3.3 (Vecteur gaussien) *Un vecteur aléatoire $X = (X_1, ..., X_p)$ est un vecteur gaussien si pour tous réels $\lambda_1, ..., \lambda_p$,*

$$\lambda_1 X_1 + ... \lambda_p X_p \text{ suit une loi normale.}$$

Remarque 4.3.1 *On en déduit que si A est une application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^d$, et si X est gaussien dans $\mathbb{R}^p$ alors AX est gaussien dans $\mathbb{R}^d$*

De façon plus explicite, on a la définition suivante en terme de densité :

Définition 4.3.4 *Soit $m \in \mathbb{R}^p$, et Σ une matrice (p, p) réelle symétrique positive, X est un vecteur normal de dimension p si sa densité est donnée par*

$$f(x) = \frac{1}{(2\pi)^{p/2} \sqrt{Det(\Sigma)}} \exp\left(-\frac{1}{2} \langle (x - m), \Sigma^{-1}(x - m) \rangle\right)$$

La matrice Σ est appelée matrice de variance-covariance. En effet, la matrice Σ est donnée par $\Sigma_{i,i} = \mathbb{V}(X_i)$ et $\Sigma_{i,j} = cov(X_i, X_j)$. Si la matrice est diagonale, alors les variables ne sont pas corrélés. En fait, dans le « monde gaussien » une matrice diagonale correspond à un vecteur à composantes indépendantes.

Théorème 4.3.3 *Les composantes d'un vecteur gaussien sont indépendantes si et seulement si elles ne sont pas corrélées.*

Remarque 4.3.2 *Il faut bien faire attention à ce que dit ce théorème. En général l'indépendance n'est pas équivalente à la corrélation nulle. Ce n'est vrai que si on a démontré au préalable que le vecteur est gaussien. En particulier, on peut construire des exemples de variables aléatoires X, Y de loi normale, telles que X et Y ne soient pas indépendantes et $cov(X, Y) = 0$, bien sûr dans ce cas, le vecteur (X, Y) n'est pas gaussien.*

Exercice 4.3.1 *Montrer que si X et Y sont deux variables aléatoires réelles indépendantes de loi normale de paramètre m et σ^2, alors $X + Y$ et $X - Y$ sont indépendantes. Indication : montrer que $(X + Y, X - Y)$ est un vecteur gaussien.*

4.3.3 Sommes de deux variables aléatoires indépendantes : convolution

On ne rentrera pas dans les détails. Le produit de convolution n'est pas au programme (mais est tombé en problème en 2001). On démontre que si X et Y sont deux variables aléatoires à densité **indépendantes**, alors $X + Y$ est également à densité. On donne de plus une forme intégrale de sa densité.

Théorème 4.3.4 (produit de convolution) *Soit X une v.a.r de densité f et Y une v.a.r indépendante de densité g. La variable aléatoire $Z = X + Y$ a pour densité :*

$$h : z \mapsto \int_{-\infty}^{\infty} f(z - y) g(y) \mathrm{d}y.$$

la fonction h est appelée convolée de f et g, et notée $f \star g$.

Remarque 4.3.3 *L'opération $\star$ est l'opération de convolution, c'est une opération commutative, non inversible. Ce dernier point n'est pas trivial et peut faire l'objet d'un problème d'analyse.*

Démonstration. Soit $z \in \mathbb{R}$, on cherche à exprimer la fonction de répartition de $Z = X + Y$. On note f et g les densités de X et Y. Ces variables étant indépendantes, la densité de (X, Y) est $\phi(x, y) = f(x)g(y)$. On intègre cette fonction de $\mathbb{R}^2$ sur le domaine $\{(x, y), x + y \leq z\}$:

$$\begin{aligned}
\mathbb{P}[X + Y \leq z] &= \int_{\mathbb{R}^2} \mathbf{1}_{\{(x,y), x+y \leq z\}} g(x) f(y) dx dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z-y} f(x) dx \right) g(y) dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{z} f(u - y) du \right) g(y) dy \\
&= \int_{-\infty}^{z} \int_{-\infty}^{\infty} f(u - y) g(y) du dy
\end{aligned}$$

avec le changement de variable $u = x + y$.

Théorème 4.3.5 *Soient $X_1, ..., X_n$ des variables aléatoires indépendantes de loi gaussienne de paramètres respectifs $(m_1, \sigma_1^2), ..., (m_n, \sigma_n^2)$. La variable aléatoire $Y = \sum_{i=1}^n X_i$ suit une loi gaussienne de paramètres $\sum_{i=1}^n m_i, \sum_{i=1}^n \sigma_i^2$.*

4.4 Exercices

Exercice 4.4.1 (#) *Déterminer c (si c'est possible) pour que les fonctions suivantes soient des densités de probabilité sur $\mathbb{R}$:*

1. $f(x) = c\mathbf{1}_{[a,b]}(x)$ *où a et b sont deux réels tels que $a < b$*
2. $f(x) = ce^{-\alpha x}1_{[0,+\infty[}(x)$*, où $\alpha \in \mathbb{R}$*
3. $f(x) = \frac{c}{1+x^2}$
4. $f(x) = c\left(\frac{a}{x}\right)^{\alpha+1} 1_{[a,+\infty[}(x)$ *pour $a > 0$ et $\alpha \in \mathbb{R}$.*
5. $f(x) = c\sin(2x + \frac{\pi}{4})1_{[-\frac{\pi}{8},\frac{\pi}{8}]}(x)$

Solution.

1. $c = 1/(b-a)$
2. Si $\alpha \leq 0$, l'intégrale $\int_{-\infty}^{\infty} f(x)dx$ n'est pas convergente : il n'existe donc pas de c répondant à la question. Si $\alpha > 0$, $x \mapsto ce^{-\alpha x}$ est intégrable sur $]0, +\infty[$ et l'intégrale vaut α. On a donc $c = \alpha$.
3. $c = 1/\pi$
4. Si $\alpha \leq 0$, impossible. Si $\alpha > 0, c = \alpha/a$
5. $c = 2$

Exercice 4.4.2 (#) *Soit X une variable aléatoire suivant une loi dont la densité de probabilité est donnée par*

$$\begin{array}{cccc} f_X & : & \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ & & x & \longmapsto & xe^{-x^2/2}1_{\mathbb{R}_+^\star}(x). \end{array}$$

Après avoir démontré leurs existences, déterminer l'espérance et la variance de la variable aléatoire X^2.

Solution. On pose $Y = X^2$. On a alors $\mathbb{E}[Y] = \mathbb{E}[X^2]$ et $\mathbb{V}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = \mathbb{E}[X^4] - \mathbb{E}[X^2]^2$. L'existence de l'espérance et de la variance découle de la comparaison avec les intégrales de Riemann.

Calcul de l'espérance. Par définition,

$$\mathbb{E}[Y] = \int_0^{+\infty} x^2 xe^{-x^2/2}dx.$$

On introduit alors les fonctions suivantes

$$\begin{array}{cccc} u & : & \mathbb{R}_+ & \rightarrow & \mathbb{R}_+ \\ & & x & \longmapsto & -e^{-x^2/2}, \end{array}$$

$$\begin{array}{cccc} v & : & \mathbb{R}_+ & \rightarrow & \mathbb{R}_+ \\ & & x & \longmapsto & x^2, \end{array}$$

On a alors

$$\mathbb{E}[Y] = \int_0^{+\infty} u'(x)v(x)dx.$$

Puisque u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+$, la formule d'intégration par parties donne

$$\mathbb{E}[Y] = [-x^2 e^{-x^2/2}]_0^{+\infty} + 2\int_0^{+\infty} xe^{-x^2/2}dx = 2\int_0^{+\infty} xe^{-x^2/2}dx = 2[-e^{-x^2/2}]_0^{+\infty} = 2.$$

Calcul de la variance. Nous avons ensuite

$$\mathbb{E}[Y^2] = \int_0^{+\infty} x^4 xe^{-x^2/2}dx.$$

On introduit alors les fonctions suivantes

$$\begin{array}{rcccc} u & : & \mathbb{R}_+ & \to & \mathbb{R}_+ \\ & & x & \longmapsto & -e^{-x^2/2}, \end{array}$$

$$\begin{array}{rcccc} v & : & \mathbb{R}_+ & \to & \mathbb{R}_+ \\ & & x & \longmapsto & x^4, \end{array}$$

On a alors

$$\mathbb{E}[Y^2] = \int_0^{+\infty} u'(x)v(x)dx.$$

Puisque u et v sont de classe $\mathcal{C}^1$ sur $\mathbb{R}_+$, la formule d'intégration par parties donne

$$\mathbb{E}[Y^2] = [-x^4 e^{-x^2/2}]_0^{+\infty} + 4\int_0^{+\infty} x^3 e^{-x^2/2}dx = 4\int_0^{+\infty} x^2 xe^{-x^2/2}dx.$$

On effectue alors une nouvelle intégration par parties (ou alors on remarque que l'on a déjà fait le calcul plus haut) pour obtenir

$$\mathbb{E}[Y^2] = 8\int_0^{+\infty} xe^{-x^2/2}dx = 8[-e^{-x^2/2}]_0^{+\infty} = 8.$$

On en déduit que

$$\mathbb{V}[Y] = \mathbb{E}[Y^2] - \mathbb{E}[Y]^2 = 4.$$

Exercice 4.4.3 *Pour quelles valeurs de $\alpha \geq 0$, l'espérance $\mathbb{E}(X^\alpha)$ est-elle finie, si la densité de la loi de X est*

$$\begin{array}{rcccc} f_X & : & \mathbb{R} & \longrightarrow & \mathbb{R}_+ \\ & & x & \longmapsto & C(1+x^2)^{-m} \end{array}$$

où $C \in \mathbb{R}$ et $m \in \mathbb{R}$.

Réponse : $\alpha < 2m - 1$.

Exercice 4.4.4 *Soient $a < b$ deux réels. Calculer $\mathbb{E}\left(e^{-X}\right)$ si X suit une loi uniforme sur l'intervalle $[a, b]$, c'est-à-dire si la densité de X est $f_X = \frac{1}{b-a}\mathbf{1}_{[a,b]}$.*

Réponse. $\mathbb{E}\left(e^{-X}\right) = \frac{e^{-b}-e^{-a}}{b-a}$.

Exercice 4.4.5 (#) *1. Soit F la fonction de répartition d'une variable aléatoire X. On suppose que F est bijective. Soit U une variable aléatoire uniforme sur $[0,1]$. Montrer que $F^{-1}(U)$ a même loi que X.*

2. On suppose que X est une v.a. à valeurs dans $\mathbb{R}$ et ayant pour densité :

$$f : y \longmapsto \frac{1}{\pi(1+y^2)}$$

(on dit que X est une v.a. de Cauchy de paramètre 1).

3. Calculer la fonction de répartition de X.

4. Montrer que si U est une v.a. uniforme sur $[0,1]$, alors

$$Y = \tan\left(\pi\left(U - \frac{1}{2}\right)\right)$$

a même loi que X.

Solution.

1. Déterminons la fonction de répartition de $F^{-1}(U)$:

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x).$$

2.
$$F(x) = \int_{-\infty}^{x} \frac{dy}{\pi(1+y^2)} = \frac{1}{\pi}\left(\operatorname{Arctan}(x) + \frac{\pi}{2}\right).$$

3. F est strictement croissante et continue : F est donc inversible. Calculons son inverse F^{-1}. Soit $y \in [0,1]$,

$$y = \frac{1}{\pi}\left(\operatorname{Arctan}(x) - \frac{\pi}{2}\right) \Longleftrightarrow \pi\left(y - \frac{1}{2}\right) = \operatorname{Arctan}(x) \Longleftrightarrow x = \tan\left(\pi\left(y - \frac{\pi}{2}\right)\right)$$

et donc $F^{-1}(y) = \tan\left(\pi\left(y - \frac{\pi}{2}\right)\right)$. On conclut que $F^{-1}(U)$ a même loi que X par la question 1.

Exercice 4.4.6 (#) *Soit X une variable aléatoire positive ayant une densité <u>continue</u> f et telle que $\mathbb{E}[X^r]$ est finie pour $r \geq 1$.*

1. Montrer que

$$\lim_{x \to +\infty} x^r \mathbb{P}(X > x) = 0.$$

Indication : trouver une expression de $\mathbb{P}(X > x)$ en fonction de f.

2. En déduire que

$$\mathbb{E}[X^r] = \int_0^{\infty} r x^{r-1} \mathbb{P}(X > x) dx.$$

Indication : Que vaut la dérivée de $x \longmapsto \mathbb{P}(X > x)$?

Solution.

1. On a $\mathbb{P}(X > x) = \int_x^{+\infty} f(t)dt$ et donc
$$x^r\mathbb{P}(X > x) = x^r \int_x^{+\infty} f(t)dt \leq \int_x^{+\infty} t^r f(t)dt$$
et ce dernier terme tend vers 0 quand $x \to +\infty$ puisque $\mathbb{E}[X^r] = \int_0^{+\infty} t^r f(t)dt$ est finie.
2. On fait une intégration par parties.
Puisque f est continue, on a pour $x > 0$, $(\mathbb{P}(X > x))' = -f(x)$. On a alors pour $A > 0$,
$$\int_x^A rt^{r-1}\mathbb{P}(X > x)dt = \Big[t^r\mathbb{P}(X > t)\Big]_0^A + \int_0^A t^r f(t)dt$$
Le premier terme tend vers 0 d'après a) et le second vers $\mathbb{E}[X^r]$ qui est finie.

Exercice 4.4.7 (#) *Soit X une variable aléatoire telle que X et $2X$ ont même fonction de répartition. Donner une équation vérifiée par la fonction de répartition de X et en déduire sa loi.*

Solution. On note F la fonction de répartition de X. Soit $x \in \mathbb{R}$. On a
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(2X \leq x) = \mathbb{P}\left(X \leq \frac{x}{2}\right) = F\left(\frac{x}{2}\right).$$
En réitérant, on a alors que pour tous $x \in \mathbb{R}$ et $n \in \mathbb{N}$, $F(x) = F\left(\frac{x}{2^n}\right)$ et en faisant tendre n vers $+\infty$, on a donc $F(x) = F(0^+)$ si $x > 0$ et $F(x) = F(0^-)$ si $x < 0$. Or, F est continue à droite donc $F(0^+) = F(0)$. La fonction F est donc constante sur $]-\infty, 0[$ et sur $[0, +\infty[$. Mais, par propriété des fonctions de répartition, $F(-\infty) = 0$ et $F(+\infty) = 1$.
On en déduit que $F(x) = 0$ si $x < 0$ et $F(x) = 1$ si $x \geq 0$, ce qui signifie que $\mathbb{P}(X = 0) = 1$ et donc X est nulle.

Exercice 4.4.8 *Soit X une variable aléatoire de loi exponentielle de paramètre λ. On désigne par $[.]$ la fonction partie entière.*
1) On pose $Y = [X]$. Donner la loi de Y, son espérance et sa variance.
2) Soit Z l'entier le plus proche de Y, donner la loi de Z.

Exercice 4.4.9 *Soit X une variable aléatoire de loi normale centrée réduite. On note $Y = \frac{X^2}{2}$. Déterminer la fonction de répartition de Y et sa densité.*

Exercice 4.4.10 *Soient X et Y deux variables aléatoires réelles indépendantes de densités respectives f et g et de fonctions de répartition respectives F et G. Donner les densités de $U = \max(X, Y)$ et $V = \min(X, Y)$.*

Exercice 4.4.11 *Soit X une variable aléatoire à densité. Soit g une fonction C^1 croissante. Montrer que*
$$\mathbb{E}[g(X)] = g(0) + \int_0^{\infty} g'(t)\mathbb{P}[X > t]dt$$

Exercice 4.4.12 *Soit $(X_1, ..., X_n)$, n variables indépendantes identiquement distribuées. Déterminer la fonction de répartition de $\min(X_1, ..., X_n)$. En supposant que la loi de X_1 est une densité, exprimer la densité de $\min(X_1, ..., X_n)$.*

Exercice 4.4.13 (Exercice fondamental) *Soient X et Y deux variables aléatoires (à densité ou discrètes) avec un moment d'ordre 2. En pensant à l'inégalité de Cauchy-Schwarz, démontrer que*
$$\mathbb{E}[XY] \leq \sqrt{\mathbb{E}[Y^2]}\sqrt{\mathbb{E}[X^2]}.$$

Exercice 4.4.14 La durée en heures de fonctionnement d'un ordinateur est une variable aléatoire continue de densité :
$$f : x \mapsto \lambda e^{-\frac{x}{36}} \text{ si } x \geq 0; 0 \text{ sinon.}$$

1) Calculer λ.
2) Quelle est la probabilité que l'ordinateur fonctionne entre 10 et 36 heures ?
3) Quelle est la probabilité qu'il fonctionne moins de 36 heures ?

Exercice 4.4.15 *La RATP estime que le retard en minutes du bus 156 est une variable aléatoire X de densité de probabilité f avec*
$$f(x) = ax(10 - x) \textit{ si } x \in [0, 10] \textit{ et } 0 \textit{ sinon.}$$

1) Trouver a tel que f soit bien une densité de probabilité.
2) Calculer le retard moyen d'un bus et sa variance
3) Déterminer la fonction de répartition de X et calculer la probabilité qu'un bus ait un retard supérieur à 8 minutes.
4) Soit $Y = X^2$, déterminer la fonction de répartition de Y puis sa densité de probabilité.

Exercice 4.4.16 (#) *Soit f la fonction définie sur $\mathbb{R}$ par :*
$$\begin{cases} f(t) = \dfrac{2}{(1+t)^3} & si \quad t \geq 0 \\ f(t) = 0 & si \quad t < 0 \end{cases}$$

1. *Montrer que f est une densité d'une variable aléatoire Z*
2. *Déterminer la fonction de répartition F_Z de Z.*
3. *Justifier la convergence de l'intégrale :*
$$\int_0^{+\infty} \frac{2t}{(1+t)^3} dt$$
La calculer en effectuant le changement de variable $u = t + 1$.
4. *Prouver que Z admet une espérance et la déterminer.*
5. *Z admet-elle une variance ?*

Solution.

1. On vérifie les caractéristiques d'une densité :
   - f est positive sur $\mathbb{R}$
   - continue sur $\mathbb{R}^*$

- f est prolongeable à gauche et à droite en 0 donc $\displaystyle\int_{-\infty}^{+\infty} f$ est impropre en $\pm\infty$.

  $\int_{-\infty}^{0} f = \int_{-\infty}^{0} 0 = 0$ (converge)

  $\displaystyle\int_0^K f = \int_0^X \frac{2}{(1+t)^3}dt = \left[-\frac{1}{(1+t)^2}\right]_0^K = 1 - \frac{1}{(1+K)^2} \to 1$

  donc $\displaystyle\int_0^{+\infty} f$ converge et vaut 1 et $\displaystyle\int_{-\infty}^{+\infty} f$ converge et vaut 1.

  Donc f est bien une densité de variable aléatoire.

2. On a $F_Z(x) = \int_{-\infty}^{x} f(t)\,dt$ donc
   - si $x < 0 : F_Z(x) = 0$
   - si $x \geq 0 : F_Z(x) = \displaystyle\int_{-\infty}^{0} 0 + \int_0^x f = 1 - \frac{1}{(1+x)^2}$

3. l'intégrale $\displaystyle\int_0^{+\infty} \frac{2t}{(1+t)^3}dt$ n'est impropre qu'en $+\infty$

   En $+\infty$ on a un équivalent simple : $\displaystyle\frac{2t}{(1+t)^3} = \frac{2t}{t^3(1+1/t)^3} \sim \frac{2}{t^2} > 0$.

   Or l'intégrale de Riemann $\displaystyle\int_1^{+\infty} \frac{1}{t^2}dt$ est convergente car $2 > 1$. Donc par comparaison d'intégrale à termes positifs, $\displaystyle\int_0^{+\infty} \frac{2t}{(1+t)^3}dt$ converge.

   On effectue le changement de variable (u est de classe C^1sur $[1, K+1]$) $u = t+1 : du \leftrightarrow dt : t = 0 \leftrightarrow r = 1$

   $$\begin{aligned}\int_0^K \frac{2t}{(1+t)^3}dt &= \int_1^{K+1} \frac{2(u-1)}{u^3}du = \int_1^{K+1} \frac{2}{u^2} - \frac{2}{u^3}du \\ &= \left[\frac{-2}{u} + \frac{1}{u^2}\right]_1^{K+1} = \frac{-2}{K+1} + \frac{1}{(K+1)^2} + 1 \\ &\underset{K\to\infty}{\longrightarrow} 1\end{aligned}$$

   donc $\displaystyle\int_0^{+\infty} \frac{2t}{(1+t)^3}dt = 1$

4. Espérance ?

   On étudie la convergence $\int_{-\infty}^{+\infty} tf(t)\,dt$ impropre en $\pm\infty$.
   - $\displaystyle\int_{-\infty}^{0} tf(t)\,dt = 0$
   - $\displaystyle\int_0^{+\infty} tf(t)\,dt = \int_0^{+\infty} \frac{2t}{(1+t)^3}dt = 1$ (d'après la question précédente)
   - donc $\int_{-\infty}^{+\infty} tf(t)\,dt$ converge et Z admet une espérance $E(Z) = 1$

5. Comme $\displaystyle\frac{2t^2}{(1+t)^3} \sim \frac{2}{t}$ dont l'intégrale diverge en $+\infty$, alors Z^2 n'a pas d'espérance et Z n'a pas de variance.

Exercice 4.4.17 (#) *On considère trois variables aléatoires X, Y et Z indépendantes et suivant la loi exponentielle de paramètre λ.*

1. *Déterminer la loi de* $Y + Z$ *(on pourra calculer sa densité).*
2. (a) *Déterminer la densité de la variable aléatoire* $D = Y + Z - X$ *sur* $\mathbb{R}^+$.
   (b) *Calculer* $\mathbb{P}(X \leq Y + Z)$.
3. *Déterminer l'événement complémentaire de l'événement*

$$\{X \leq Y + Z\} \cap \{Y \leq Z + X\} \cap \{Z \leq X + Y\}$$

et calculer sa probabilité.

4. *Quelle est la probabilité pour qu'on puisse construire un triangle (éventuellement aplati) dont les côtés aient pour longueurs* X, Y *et* Z *?*

solution

1. $Y + Z$ est encore une variable aléatoire à valeurs dans $\mathbb{R}^+$ et comme Y et Z sont indépendantes, une densité f_{Y+Z} de $Y + Z$ peut être définie par convolution, ce qui donne pour $x > 0$

$$f_{Y+Z}(x) = \int_{-\infty}^{+\infty} f_Y(t) f_Z(x - t) dt = \int_0^x \lambda e^{-\lambda t} \lambda e^{-\lambda(x-t)} = \lambda^2 x e^{-\lambda x}.$$

De plus, $f_{Y+Z}(x) = 0$ pour $x \leq 0$.

2. (a) La variable aléatoire $-X$ est indépendante de $Y + Z$ et, à nouveau par convolution, pour $x > 0$,

$$f_D(x) = \int_{-\infty}^{+\infty} f_{Y+Z}(t) f_{-X}(x - t) dt. \tag{4.2}$$

On a donc besoin de calculer la densité de $-X$. On utilise la méthode vue en TD. Soit Φ une fonction continue bornée de $\mathbb{R}$ dans $\mathbb{R}$.

$$\mathbb{E}[\Phi(-X)] = \int_0^{+\infty} \Phi(-y) \lambda e^{-\lambda y} dy \overset{u=-y}{=} \int_{-\infty}^0 \Phi(u) \lambda e^{\lambda u} du.$$

La densité de $-X$ est donc $u \longmapsto \lambda e^{\lambda u} \mathbf{1}_{]-\infty,0]}(u)$ et en revenant à (4.2), on a pour $x > 0$,

$$f_D(x) = \int_{-\infty}^{+\infty} \lambda^2 t e^{-\lambda t} \mathbf{1}_{[0,+\infty[}(t) \lambda e^{\lambda(x-t)} \mathbf{1}_{]-\infty,0]}(x - t) dx = \lambda^3 e^{\lambda x} \int_x^{+\infty} t e^{-2\lambda t} dt.$$

Le calcul de cette intégrale se fait à l'aide d'une intégration par parties, ce qui donne pour $x > 0$

$$f_D(x) = \frac{\lambda}{4}(2\lambda x + 1) e^{-\lambda x}.$$

(b) On a

$$\mathbb{P}(X \leq Y + Z) = \mathbb{P}(D \geq 0) = \int_0^{+\infty} f_D(x) dx = \frac{\lambda}{4} \int_0^{+\infty} (2\lambda x + 1) e^{-\lambda x} dx.$$

En intégrant à nouveau par parties, on obtient : $\mathbb{P}(D \geq 0) = 3/4$.

3. Posons $C = \{X \leq Y + Z\} \cap \{Y \leq Z + X\} \cap \{Z \leq X + Y\}$. L'événement complémentaire est $C^c = \{X > Y + Z\} \cup \{Y > Z + X\} \cup \{Z > X + Y\}$. Les trois événements constituant la réunion précédente sont deux à deux disjoints (si on avait par exemple $\{X > Y + Z\}$ et $\{Y > Z + X\}$, on aurait $\{X > 2Z + X\}$, ce qui est clairement impossible puisque Z est à valeurs dans $\mathbb{R}^+$). De plus, d'après la question b), chacun de ces événements a pour probabilité $1/4$.
   On en déduit que $\mathbb{P}(C^c) = 3/4$.
4. On peut construire un triangle de côtés donnés si et seulement si chacun de ces nombres est inférieur ou égal à la somme des deux autres, ce qui correspond à la réalisation de l'événement C de probabilité $1/4$.

Exercice 4.4.18 *Dans cet exercice, n désigne un entier naturel non nul.*

1. *Soit f_n la fonction définie par : $f_n(x) = \begin{cases} nx^{n-1} & si\ x \in [0;1] \\ 0 & sinon \end{cases}$*
   Montrer que f_n est une densité de probabilité.
2. *On considère une variable aléatoire X_n réelle dont une densité de probabilité est f_n. On dit alors que X_n suit une loi monôme d'ordre n.*
   (a) *Reconnaître la loi de X_1.*
   (b) *Dans le cas où n est supérieur ou égal à 2, déterminer la fonction de répartition F_n de X_n, ainsi que son espérance $E(X_n)$ et sa variance $V(X_n)$.*
3. *On considère deux variables aléatoires U_n et V_n définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, suivant la loi monôme d'ordre n $(n \geqslant 2)$ et indépendantes, c'est-à-dire qu'elles vérifient en particulier l'égalité suivante :*

$$\forall x \in \mathbb{R}, \quad P(U_n \leqslant x \cap V_n \leqslant x) = P(U_n \leqslant x) P(V_n \leqslant x)$$

   On pose $M_n = \sup(U_n, V_n)$ et on admet que M_n est une variable aléatoire définie, elle aussi, sur $(\Omega, \mathcal{A}, P)$.
   (a) *Pour tout réel x, écrire, en justifiant la réponse, l'événement $(M_n \leqslant x)$ à l'aide des événements $(U_n \leqslant x)$ et $(V_n \leqslant x)$.*
   (b) *En déduire une densité de M_n. Vérifier que M_n suit une loi monôme dont on donnera l'ordre, puis déterminer sans calcul $E(M_n)$.*
   (c) *On pose $T_n = \inf(U_n, V_n)$. Exprimer $M_n + T_n$ en fonction de U_n et V_n, puis en déduire, sans calcul d'intégrale, la valeur de $E(T_n)$.*

Dans cet exercice, n désigne un entier naturel non nul.

1. Soit f_n la fonction définie par : $f_n(x) = \begin{cases} nx^{n-1} & \text{si } x \in [0;1] \\ 0 & \text{sinon} \end{cases}$
   f_n est positive sur $\mathbb{R}$ et continue sur $\mathbb{R} - \{0, 1\}$ (en fait, en 0 elle est continue)
   $\int_{-\infty}^{+\infty} f\ _n$est impropre en $\pm\infty$ car f_n est continue par morceaux sur $\mathbb{R}$
   $\int_{-\infty}^{0} f_n = \int_{-\infty}^{0} 0$ converge et est nulle.
   $\int_{1}^{+\infty} f_n = \int_{1}^{+\infty} 0 = 0$
   $\int_0^1 f_n = \int_0^1 nx^{n-1}dx = [x^n]_{x=0}^1 = 1$ car $0^n = 0$ pour $n \in \mathbb{N}^*$
   Donc $\int_{-\infty}^{+\infty} f_n$ converge et vaut 1.
   Donc f_n est bien une densité de probabilité.

2. On considère une variable aléatoire X_n réelle dont une densité de probabilité est f_n. On dit alors que X_n suit une loi monôme d'ordre n.

   (a) Pour $n = 1$ la densité est : $f_1(x) = \begin{cases} 1 & \text{si } x \in [0;1] \\ 0 & \text{sinon} \end{cases}$ donc X_1 suit une loi uniforme sur $[0,1]$

   (b) La fonction de réparition de X_n est donnée par :
   $F_n(x) = \int_{-\infty}^{x} f$ où il faut distinguer suivant que $x < 0 : 0 \leq x \leq 1$ et $x > 1$:
   – si $x < 1$ alors $F_n(x) = \int_{-\infty}^{x} 0 = 0$
   – si $0 \leq x \leq 1$ alors $F_n(x) = \int_{-\infty}^{0} 0 + \int_0^x nt^{n-1}dt = 0 + [t^n]_0^x = x^n$
   – si $1 < x$ alors $F_n(x) = \int_{-\infty}^{0} 0 + \int_0^1 nt^{n-1}dt + \int_1^x 0 = 0 + [t^n]_0^1 + 0 = 1$
   Pour calculer $E(X_n)$ on étudie la convergence et on calcule $\int_{-\infty}^{+\infty} tf_n(t)\,dt$ qui est impropre en $\pm\infty$
   $\int_{-\infty}^{0} t\,f_n(t)\,dt = \int_{-\infty}^{0} 0 = .0$
   $\int_1^{+\infty} t\,f_n(t)\,dt = \int_1^{+\infty} 0 = 0$
   $$\int_0^1 t\,f_n(t)\,dt = \int_0^1 t\,nt^{n-1}dt = \int_0^1 nt^n dt = \left[\frac{n}{n+1}t^{n+1}\right]_{t=0}^1 = \frac{n}{n+1}$$
   Donc X_n a une espérance et $\boxed{E(X_n) = \frac{n}{n+1}}$
   De même pour $E(X_n^2) = \int_0^1 t^2 f_n(t)\,dt = \int_0^1 nt^{n+1}dt = \left[\frac{n}{n+2}t^{n+2}\right]_0^1 = \frac{n}{n+2}$
   Donc X a une variance et
   $$\begin{aligned} V(X_n) &= E(X_n^2) - E(X_n^2) = \frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2 \\ &= n\frac{n^2+2n+1-n(n+2)}{(n+2)(n+1)^2} \\ &= \frac{n}{(n+2)(n+1)^2} \end{aligned}$$

3. On considère deux variables aléatoires U_n et V_n définies sur le même espace probabilisé $(\Omega, \mathcal{A}, P)$, suivant la loi monôme d'ordre n $(n \geqslant 2)$ et indépendantes, c'est-à-dire qu'elles vérifient en particulier l'égalité suivante :
   $$\forall x \in \mathbb{R}, \quad P(U_n \leqslant x \cap V_n \leqslant x) = P(U_n \leqslant x)\,P(V_n \leqslant x)$$
   On pose $M_n = \sup(U_n, V_n)$ et on admet que M_n est une variable aléatoire définie, elle aussi, sur $(\Omega, \mathcal{A}, P)$.

   (a) $(M_n \leqslant x)$ est l'événement "le plus grand es tplus petit que x" c'est à dire "tous sont plus petit que x"
   et $(M_n \leqslant x) = (U_n \leqslant x) \cap (V_n \leqslant x)$

   (b) Comme U_n et V_n sont indépendantes, on a alors : $p(M_n \leqslant x) = p(U_n \leqslant x) \cdot p(V_n \leqslant x)$
   En notant F la fonction de répartition d'une loi monôme et G celle de M_n on a alors :
   $G(x) = F(x) \cdot F(x) = F(x)^2$ pour tout x réel.
   Comme F est la fonctio de répartition d'une variable à densité, on sait

– que F est continue sur $\mathbb{R}$
– de classe C^1 sur $\mathbb{R}-\{0,1\}$ (là où f_n est continue)
– et que $F' = f_n$

Donc

– que G est continue sur $\mathbb{R}$ comme composée de fonctions continues.
– G de classe C^1 sur $\mathbb{R}-\{0,1\}$
– et que $G' = 2F(x)F'(x) = 2F(x)f_n(x) = \begin{cases} 2x^n nx^{n-1} = 2nx^{2n-1} & \text{si } x \in [0;1] \\ 0 & \text{sinon} \end{cases}$

Donc M_n est une variable à densité et suit une loi monôme d'ordre $2n$

et donc $\boxed{E(M_n) = \dfrac{2n}{2n+1}}$

(c) On pose $T_n = \inf(U_n, V_n)$.
On a $M_n + T_n = U_n + V_n$ (car l'un est le plus petit et l'auter le plus grand)
De plus $E(U_n) = E(V_n) = \dfrac{n}{n+1}$ alors $E(U_n+V_n) = E(U_n) + E(V_n) = \dfrac{2n}{n+1}$
Et comme $T_n = M_n + T_n - M_n = U_n + V_n - M_n$ on a finalement

$$\begin{aligned} E(T_n) &= E(U_n+V_n) - E(M_n) \\ &= \frac{2n}{n+1} - \frac{2n}{2n+1} \\ &= \frac{2n^2}{(n+1)(2n+1)} \end{aligned}$$

Exercice 4.4.19 (Inégalité de Paley-Zigmund) *Soit X une variable positive, non nulle, ayant un moment d'ordre 2. Montrer*

$$\mathbb{P}(X \geq \alpha\mathbb{E}[X]) \geq (1-\alpha)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}$$

pour tout $\alpha \in]0,1[$.

Exercice 4.4.20 (#) *Soit X une variable aleatoire normale de moyenne $\mu = 2$ de variance $\sigma^2 = 4$.*

1. *Soit la variable $T = (X-\mu)/\sigma$. Quelle loi de probabilité suit T ?*
2. *On considère l'événement $\{X < 1,5\}$. Quel est l'événement équivalent pour T ? Quelle est la probabilité correspondante ?*
3. *Même question pour l'évenement $\{X > -2\}$.*
4. *Soit l'évenement $\{-1 \leq T < 1\}$. Quel est l'événement équivalent pour X ? Quelle est sa probabilité ?*
5. *Déterminer les valeurs x telles que $\mathbb{P}(X < x) = 0,76$; $\mathbb{P}(X \geq x) = 0,6$ et $\mathbb{P}(0 \leq X < x) = 0,40$.*

Solution.

2. $\mathbb{P}(X < 1,5) = \mathbb{P}(T < -0,25) = 1 - \mathbb{P}(T \leq 0,25) = \boxed{0,401}$
3. $\mathbb{P}(X > -2) = \mathbb{P}(T > -2)\mathbb{P}(T \leq 2) = \boxed{0,977}$

4. $\mathbb{P}(-1 \leq T < 1) = \mathbb{P}(T < 1) - \mathbb{P}(T < -1) = \mathbb{P}(T < 1) - \mathbb{P}(T > 1) = \mathbb{P}(T < 1) - (1 - \mathbb{P}(T \leq 1))$ et donc $\mathbb{P}(-1 \leq T < 1) = 2\mathbb{P}(T < 1) - 1 = \boxed{0,682}$.
5. – $\mathbb{P}(X < x) = \mathbb{P}(T < \frac{x-2}{2})$. Donc $\frac{x-2}{2} = 0,71$ et $\boxed{x = 3,42}$.

   – $\mathbb{P}(X \geq x) = \mathbb{P}\left(T \geq \frac{x-2}{2}\right) = \mathbb{P}\left(T \leq \frac{2-x}{2}\right)$ et donc $\boxed{x = 1,5}$.

   – $\mathbb{P}(0 \leq X < x) = \mathbb{P}\left(-1 \leq T < \frac{x-2}{2}\right) = \mathbb{P}\left(T < \frac{x-2}{2}\right) - \mathbb{P}\left(T < -1\right)$. On a alors

   $$\mathbb{P}(0 \leq X < x) = \mathbb{P}\left(T < \frac{x-2}{2}\right) - 1 + \mathbb{P}(T \leq 1) = \mathbb{P}\left(T < \frac{x-2}{2}\right) - 0,16$$

   (car d'après la table de la loi normale, $\mathbb{P}(T \leq 1) = 0,84$).
   On cherche donc x tel que $\mathbb{P}\left(T < \frac{x-2}{2}\right) = 0,56$. On a alors-+ $\boxed{x = 2,3}$.

Exercice 4.4.21 (#) *Le taux de glycémie d'une population est répartie suivant une loi normale. Une enquête auprès de l'ensemble de cette population montre que 84,1% des individus ont un taux inférieur à 1,40 g/l et 2,3% ont un taux supérieur à 1,60 g/l.*

1. *A l'aide de la table de la loi normale, déterminer la moyenne et la variance du modèle.*
2. *En admettant qu'un taux de glycémie supérieur à 1,30 g/l nécessite un traitement, quel pourcentage de cette population devra t'on traiter ?*

Solution.

1. Notons respectivement m et σ^2 la moyenne et la variance du taux de glycémie que l'on note X. On a

   $$0,841 = \mathbb{P}(X < 1,4) = \mathbb{P}\left(\frac{X-m}{\sigma} < \frac{1,4-m}{\sigma}\right).$$

   Comme X est une v.a. normale $\mathcal{N}(m, \sigma^2)$, $N = \frac{X-m}{\sigma}$ est une v.a. normale $\mathcal{N}(0,1)$. En utilisant la table de la loi normale, on a donc $\frac{1,4-m}{\sigma} = 1$. De même,

   $$0,977 = \mathbb{P}(X < 1,6) = \mathbb{P}\left(N < \frac{1,6-m}{\sigma}\right)$$

   et $\frac{1,6-m}{\sigma} = 2$.
   Pour trouver m et σ, il suffit donc de résoudre le système

   $$\sigma + m = 1,4 \quad 2\sigma + m = 1,6$$

   qui a pour solution $\boxed{m = 1,2 \ \sigma = 0,2}$. X est donc une v.a. $\mathcal{N}(1,2; 0,04)$.
2. Il suffit de calculer $\mathbb{P}(X > 1,3) = \mathbb{P}(N > 0,5) = 1 - \mathbb{P}(N \leq 0,5) = \boxed{0,31}$.

Exercice 4.4.22 (#) *Soit X une variable aléatoire gaussienne centrée et réduite (on note $X \sim \mathcal{N}(0,1)$). Trouver la densité de la variable aléatoire $Y = X^2$.*

Solution. La densité de X^2 est

$$f(y) = \frac{e^{-y/2}}{\sqrt{2\pi y}} \mathbf{1}_{[0,+\infty[}(y).$$

Exercice 4.4.23 On considère N de loi normale $\mathcal{N}(0,1)$ et ε, indépendante de N, telle que $\mathbb{P}(\varepsilon=-1)=\mathbb{P}(\varepsilon=1)=1/2$.

1. Montrer que $N' := \varepsilon N$ a même loi que N (on pourra d'abord montrer que $-N$ a même loi que N).
2. Est-ce que $N+N'$ est une v.a. normale ?

Solution.

1. Calculons la fonction de répartition de N'. Pour $x\in\mathbb{R}$,
$$\begin{aligned}\mathbb{P}(N'\leq x) &= \mathbb{P}(\varepsilon N\leq x)=\mathbb{P}(\varepsilon=1,\varepsilon N\leq x)+\mathbb{P}(\varepsilon=-1,\varepsilon N\leq x)\\ &= \mathbb{P}(\varepsilon=1,N\leq x)+\mathbb{P}(\varepsilon=-1,N\geq -x)\\ &= \mathbb{P}(\varepsilon=1)\mathbb{P}(N\leq x)+\mathbb{P}(\varepsilon=-1)\mathbb{P}(N\geq -x)\end{aligned}$$
puisque les variables ε et N sont indépendantes. De plus, en utilisant la définition de ε et la propriété de symétrie de la loi normale centrée réduite,
$$\mathbb{P}(N'\leq x)=\frac{1}{2}\mathbb{P}(N\leq x)+\frac{1}{2}\mathbb{P}(N\leq x)=\mathbb{P}(N\leq x).$$
Les 2 v.a. N et N' ont donc <u>même fonction de répartition</u> : elles ont donc <u>même loi</u>.
2. En utilisant l'indépendance de N et ε, on remarque que
$$\mathbb{P}(N+N'\neq 0)=\mathbb{P}((\varepsilon+1)N\neq 0)=\mathbb{P}(\varepsilon+1\neq 0)\mathbb{P}(N\neq 0)=\frac{1}{2}\cdot 1=\frac{1}{2}$$
(on rappelle que comme N est une variable à densité, on a $\mathbb{P}(N\neq 0)=1-\mathbb{P}(N=0)=1$). On a donc montré que $\mathbb{P}(N+N'=0)=1/2$. Cette probabilité étant non nulle, $N+N'$ ne peut pas être une variable à densité et a fortiori pas une v.a. normale.

<u>Rappel de cours :</u> Si N et N' sont deux variables normales $\mathcal{N}(\mu_1,\sigma_1^2)$ et $\mathcal{N}(\mu_2,\sigma_2^2)$ **indépendantes**, alors $N+N'$ est une v.a. normale $\mathcal{N}(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$.
Cet exercice fournit donc un contre-exemple à cette propriété quand les 2 v.a. sont dépendantes.

Exercice 4.4.24 (difficile) *Soient a et b deux réels strictement positifs et $s\in]0,1[$.*
1) Etablir la convergence de l'intégrale $J=\int_0^{+\infty}\frac{1}{e^u+a}du$. Calculer sa valeur (indication calculer aJ).

On considère une suite de variables aléatoires $(Y_k)_{k\geq 1}$ définies sur $(\Omega,\mathcal{F},\mathbb{P})$, indépendantes et suivant la même loi exponentielle de paramètre b. On considère également sur le même espace probabilisé Ω, une variable aléatoire N indépendantes des Y_k, suivant une loi géométrique de paramètre s.

On va étudier $Z:=\max\{Y_1,...,Y_N\}$.

2) Soit $j\in\mathbb{N}^\star$ et $t>0$. Calculer la probabilité conditionnelle $\mathbb{P}[Z\leq t|N=j]$
3) En déduire la fonction de répartition de Z et déterminer une densité.
4) Montrer que Z a une espérance et que l'on a
$$\mathbb{E}[Z]=\frac{-\ln(s)}{b(1-s)}.$$

Seul le résultat de la question 4 est nécessaire pour traiter les questions suivantes :

5) Soit g la fonction définie sur $[0,\infty[$ *par* $g(0)=1$ *et* $g(t)=\frac{te^{-t}}{1-e^{-t}}$ *pour* $t>0$.
– *Montrer que g est continue et bornée sur* $[0,+\infty[$.
– *Etablir que pour tout* $t>0$ *et* $n\in\mathbb{N}$, *on a l'égalité*

$$g(t)=g(t)e^{-(n+1)t}+\sum_{k=0}^{n} te^{-(k+1)t}.$$

6) Justifier la convergence de l'intégrale $\int_0^{+\infty} te^{-(k+1)t}dt$ *et la calculer.*
7) Montrer alors que $\int_0^{\infty} g(t)dt$ *est convergence et égale à* $\sum_{k=1}^{\infty}\frac{1}{k^2}$
8) On rappelle que $\sum_{k=1}^{\infty}\frac{1}{k^2}=\frac{\pi^2}{6}$. *Montrer que la valeur moyenne de la quantité* $\mathbb{E}[Z]$ *sur* $]0,1[$, *c'est à dire* $\int_0^1 \frac{-\ln(s)}{b(1-s)}ds$, *est égale à* $\frac{\pi^2}{6b}$.

Exercice 4.4.25 On considère une fonction F définie par

$$F(x)=a \quad \text{si} \quad x\leq -1\ ,$$

$$F(x)=bx+c \quad \text{si} \quad x\in]-1,1[\ ,$$

$$F(x)=d \quad \text{si} \quad x\geq 1\ .$$

1) Déterminer les valeurs de a, b, c et d pour que F soit une fonction de répartition d'une variable aléatoire à densité.
2) Représenter graphiquement F.
3) Déterminer la densité de probabilité associée à F.

Exercice 4.4.26 Soit f la fonction de $\mathbb{R}$ dans $\mathbb{R}^+$ définie par

$$\forall x\in[2,4[\quad f(x)=\lambda\frac{1}{(1-x)^2}\ ;\quad f(x)=0\quad x\notin[2,4[\ .$$

1) Déterminer $\lambda\in\mathbb{R}$ pour que f soit une densité de probabilité.
2) Soit X une variable aléatoire de densité f. Déterminer sa fonction de répartition F.
3) Calculer l'espérance $E(X)$ et la variance $V(X)$ de la variable aléatoire X. On pourra utiliser les égalités

$$x=(x-1)+1\ ,\quad x^2=(x-1)^2+2(x-1)+1\ .$$

Exercice 4.4.27 Soit X une variable aléatoire réelle de densité p_X. Déterminer la densité de la variable aléatoire Y dans les cas suivants :
– $Y=aX+b$ où a,b sont des nombres réels et $a\neq 0$;
– $Y=X^2$;
– $Y=\exp(X)$.
Résoudre cet exercice en utilisant les deux méthodes suivantes : le calcul de la fonction de répartition, l'utilisation du théorème de transfert.

Exercice 4.4.28 Soit X une variable aléatoire de loi normale centrée réduite.
– Démontrer que pour tout $k\in\mathbb{N}$, l'intégrale $\int_0^{\infty} x^k e^{-\frac{x^2}{2}}dx$ est convergente. En déduire que la variable aléatoire X admet des moments de tout ordre.
– Démontrer que l'application $x\in\mathbb{R}\mapsto p_x(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ est bien une densité de probabilité.
– Montrer que X admet 0 comme espérance et 1 comme variance.

– Montrer que la variable aléatoire $Y = \sigma X + m$ avec σ, m deux réels non-nuls suit une loi normale $\mathcal{N}(m, \sigma)$. En déduire que Y admet des moments de tout ordre et a pour espérance m et variance σ^2.

Exercice 4.4.29 On considère une variable aléatoire réelle X dont la fonction de répartition $F_X(x)$ est donnée par :

$$F_X(x) = 0 \quad \text{si} \quad x < 0$$

$$F_X(x) = 1 - (1 + \frac{x}{2})e^{-\frac{x}{2}} \quad \text{si} \quad x \geq 0$$

1) $F_X(x)$ est-elle continue sur $\mathbb{R}$?
2) Déterminer $\lim_{x \to +\infty} F_X(x)$. Interprétation ?
3) Calculer la densité de probabilité $f_X(x)$. Quel est le mode de X ?
4) Calculer l'espérance et l'écart-type de X.
représentative de $F_X(x)$.
5) Calculer $P(1 \leq X < 2)$.

Exercice 4.4.30 Soit T une variable aléatoire réelle, de densité de probabilité :

$$f_T(t) = 0 \quad \text{si} \quad t \notin [-1, +1]$$

$$f_T(t) = \lambda(1 - t^2) \quad \text{si} \quad t \in [-1, +1]$$

1) Calculer λ et représenter graphiquement $f_T(t)$.
2) Déterminer la fonction de répartition $F_T(t)$ et tracer sa courbe représentative.
3) Calculer la probabilité de l'événement $|T| \geq \frac{1}{2}$. Représenter cette probabilité sur chacun des deux graphiques précédents.
4) Calculer l'espérance et la variance de T. Quelle est sa médiane ?

Exercice 4.4.31 Soit U une variable aléatoire de loi uniforme sur $[0, 1]$ (sa densité de probabilité, f_U est définie par $f_U(x) = \mathbf{1}_{[0,1]}(x)$). On pose

$$X = -\frac{1}{\lambda} \ln(U) \ ;$$

F_X est la fonction de répartition de X et f_X sa densité de probabilité.
1. Rappeler la fonction de répartition $F_U(x) = P(U \leq x)$, $x \in \mathbb{R}$, de la variable aléatoire U.
2. Déterminer la fonction de répartition F_X et la densité de probabilité de la variable aléatoire X.
3. Quelle est la loi de X ? Donner les valeurs de $E(X)$ et $V(X)$. On rappelle que

$$\int_0^{+\infty} x^n e^{-x} dx = n! \ .$$

Exercice 4.4.32 La durée en heures de fonctionnement d'un ordinateur est une variable aléatoire continue de densité :

$$f : x \mapsto \lambda e^{-\frac{x}{100}} \text{ si } x \geq 0; 0 \text{ sinon.}$$

1) Calculer λ.
2) Quelle est la probabilité que l'ordinateur fonctionne entre 50 et 100 heures ?

3) Quelle est la probabilité qu'il fonctionne moins de 100 heures ?

Exercice 4.4.33 (Inégalité de Paley-Zigmund) Soit X une variable positive, non nulle, ayant un moment d'ordre 2. Montrer

$$\mathbb{P}\left(X \geq \alpha \mathbb{E}[X]\right) \geq (1-\alpha)^2 \frac{(\mathbb{E}[X])^2}{\mathbb{E}[X^2]}$$

pour tout $\alpha \in]0,1[$.

Chapitre 5

Théorèmes limites

Dans ce chapitre, on s'intéresse au comportement asymptotique de suites de variables aléatoires. Deux théorèmes majeurs en probabilité sont au programme.
- La loi des grands nombres.
- Le théorème de la limite centrale (ou théorème central limite)

Il y a plusieurs versions du premier théorème et nous nous bornerons à la version *faible*. Le premier théorème est qualifié de "loi". Ce théorème fait la jonction entre la définition axiomatique que l'on a donnée et l'approche *fréquentiste*, qui nous a servi de motivation : nous allons *démontrer* que si A est un événement associé à une expérience, alors :

$$\frac{\text{nombre de fois que A se réalise}}{\text{nombre de fois que l'on fait l'expérience}} \longrightarrow \mathbb{P}(A)$$

Le deuxième théorème permet d'étudier les fluctuations de ces quantités autour de la limite. C'est un théorème central en statistiques. Nous considérons des variables aléatoires soit discrètes, soit à densité.

5.1 Notions de convergence de suites aléatoires

5.1.1 Convergence en probabilité et loi faible des grands nombres.

Définition 5.1.1 *Soit $(X_n, n \geq 1)$ une suite de v.a.r. et X une v.a.r. définies sur le même espace probabilisé. La suite $(X_n, n \geq 1)$ converge en probabilité si :*

$$\forall a > 0, \mathbb{P}[|X_n - X| > a] \underset{n\to\infty}{\longrightarrow} 0.$$

Cela signifie donc que la probabilité que l'écart entre X_n et X dépasse a tend vers 0. Cette convergence correspond à une certaine distance sur l'espace L^0 des variables aléatoires, c'est hors programme (voir par exemple Zuily-Quéffelec : Analyse pour l'agrégation p515)

Inégalités de Markov et de Tchebychev

Proposition 5.1.1 (Inégalité de Markov) *Soit X une variable possédant un moment d'ordre* 1 *alors :*

$$\forall a > 0, \mathbb{P}[|X| > a] \leq \frac{\mathbb{E}[|X|]}{a}$$

Démonstration. Soit $A = \{|X| > a\}$, on a $\mathbb{E}[\mathbf{1}_A] = \mathbb{P}(A)$. De plus

$$|X| = |X|1_{|X|\leq a} + |X|1_{|X|>a} \geq a1_{|X|>a}$$

d'où $\mathbb{P}[|X| > a] \leq \frac{\mathbb{E}[|X|]}{a}$. □

Proposition 5.1.2 (Inégalité de Tchebychev) *Soit X une variable possédant un moment d'ordre 2, alors*

$$\forall a > 0, \mathbb{P}[|X - \mathbb{E}(X)| > a] \leq \frac{\mathbb{V}(X)}{a^2}.$$

Démonstration. C'est immédiat par l'inégalité de Markov appliquée à $Y = |X - \mathbb{E}(X)|^2$. Vous pouvez également suivre le même argument que précédemment (à savoir faire).

Théorème 5.1.3 (Loi faible des grands nombres I) *Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles définies sur un même espace $(\Omega, \mathcal{F}, \mathbb{P})$* ***indépendantes, de même espérance et variance****. On définit la moyenne empirique des variables (X_n) par $Z_n = \frac{1}{n}\sum_{i=1}^n X_i$. Alors,*

$$Z_n \underset{n\to\infty}{\longrightarrow} \mathbb{E}[X_1] \text{ en probabilité.}$$

Démonstration. L'inégalité de Tchebycheff nous donne

$$\mathbb{P}\left(|\frac{S_n}{n} - \mathbb{E}\left(\frac{S_n}{n}\right)| > t\right) = \mathbb{P}\left(|\frac{S_n}{n} - \mathbb{E}[X]| > t\right) \leq \frac{\mathbb{V}\left(\frac{S_n}{n}\right)}{t^2}.$$

D'autre part

$$\frac{\mathbb{V}\left(\frac{S_n}{n}\right)}{t^2} = \frac{1}{n^2t^2}\mathbb{V}(S_n).$$

Par indépendance, cette dernière quantité est égale à

$$\frac{1}{n^2t^2}\sum_{i=1}^n \mathbb{V}(X_i) = \frac{\mathbb{V}(X)}{nt^2}.$$

On conclut que $\frac{S_n}{n}$ converge en probabilité vers la constante $\mathbb{E}[X]$. □

Définition 5.1.2 *Sous réserve d'existence des quantités manipulées, on dit que la suite $(X_n, n \geq 1)$*
- *converge en moyenne vers X si $\lim_{n\to\infty} \mathbb{E}\left(|X_n - X|\right) = 0$*
- *converge en moyenne quadratique vers X si $\lim_{n\to\infty} \mathbb{E}\left(|X_n - X|^2\right) = 0$*

Proposition 5.1.4 *Si $(X_n, n \geq 1)$ converge en moyenne vers X, alors*

$$\mathbb{E}[X_n] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[X]$$

Démonstration. Par la seconde inégalité triangulaire

$$||\mathbb{E}[X_n]| - |\mathbb{E}[X]|| \leq \mathbb{E}[|X_n - X|]$$

La quantité à droite tend vers 0 par définition. □
On énonce alors le théorème suivant :

Théorème 5.1.5 (Convergence en moyenne quadratique) *Soit $(X_n, n \geq 1)$ une suite de variables aléatoires réelles indépendantes de même loi et de carré intégrable (i.e possédant un moment d'ordre 2). La suite des moyennes empiriques $\frac{1}{n}\sum_{i=1}^{n} X_i$ converge en moyenne quadratique vers $\mathbb{E}[X_1]$.*

Démonstration. On a :

$$\mathbb{E}\left(||\frac{1}{n}\sum_{i=1}^{n} X_i - \mathbb{E}[X_1]||^2\right) = V\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right).$$

Or par indépendance, et car les variables on même loi, on a :

$$\mathbb{V}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathbb{V}(X_i) = \frac{1}{n}\mathbb{V}(X_1).$$

On obtient la convergence souhaitée. □

Proposition 5.1.6 *La convergence en moyenne quadratique implique la convergence en moyenne, qui implique la convergence en probabilité.*

Démonstration...

5.1.2 Convergence en loi

Définition 5.1.3 *Soit $(X_n, n \geq 1)$ une suite de variables aléatoires. On note leur fonction de répartition F_n. Soit X une variable aléatoire. de fonction de répartition F. On dit que $(X_n, n \geq 1)$ converge en loi vers X et on note*

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$$

Si pour tout x point de continuité de F on a

$$F_n(x) \underset{n\longrightarrow\infty}{\longrightarrow} F(x).$$

Comme son nom l'indique, cette convergence est relative aux lois des variables et pas aux variables elles-mêmes contrairement à la convergence en probabilité, on ne peut pas munir L^0 d'une distance pour cette convergence (on peut le faire pour les mesures de probabilités associées). On admet la proposition suivante :

Proposition 5.1.7 (Convergence en loi et fonctions tests)

$$X_n \xrightarrow[n\to\infty]{\mathcal{L}} X$$

si et seulement si, pour toute fonction continue bornée ϕ de $\mathbb{R}^d$ dans $\mathbb{R}$, on a

$$\mathbb{E}[\phi(X_n)] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[\phi(X)].$$

La proposition nous permet de prendre des fonctions continues bornées à la place d'indicatrice de fermés. Sa démonstration (que nous ne faisons pas) est basée sur une approximation de l'indicatrice par des fonctions régulières. La démonstration n'est pas triviale et n'est pas au programme.

Proposition 5.1.8 *On a l'implication suivante*

$$X_n \underset{n\to\infty}{\longrightarrow} X \text{ en probabilité} \Longrightarrow X_n \underset{n\to\infty}{\longrightarrow} X \text{ en loi.}$$

La réciproque est fausse.

Démonstration. On va utiliser la proposition précédente. Soit $(X_n, n \geq 1)$ une suite de v.a qui converge en probabilité vers X. Si ϕ est une fonction continue à support compact, alors elle est uniformément continue. Pour tout $\epsilon > 0$, il existe $\eta > 0$ tel que si $|x-y| \leq \eta$ on a $|\phi(x) - \phi(y)| \leq \epsilon$. On a donc

$$\begin{aligned}
|\mathbb{E}[\phi(X_n) - \phi(X)]| &\leq \mathbb{E}[|\phi(X_n) - \phi(X)|] \\
&= \mathbb{E}[|\phi(X_n) - \phi(X)|\mathbf{1}_{\{|X_n - X| \leq \eta\}}] + \mathbb{E}[|\phi(X_n) - \phi(X)|\mathbf{1}_{\{|X_n - X| > \eta\}}] \\
&\leq \epsilon + 2||\phi||_\infty \mathbb{P}[|X_n - X| > \eta].
\end{aligned}$$

On a donc

$$\limsup_{n\to\infty} |\mathbb{E}[\phi(X_n) - \phi(X)]| \leq \epsilon.$$

Ceci est vrai pour tout $\epsilon > 0$ donc

$$\lim_{n\to\infty} |\mathbb{E}[\phi(X_n) - \phi(X)]| = 0,$$

on conclut par la proposition précédente.□

5.1.3 Approximations de certaines lois et théorème central limite

Théorème 5.1.9 (Approximation loi hypergéométrique-loi binomiale) *Soit* $(X_k, k \geq 1)$ *une suite de variables aléatoires de loi hypergéométrique de paramètres* (k, n, p)*. La suite* $(X_k, k \geq 1)$ *converge en loi vers une variable* X *binomiale de paramètre* (n, p)

Remarque 5.1.1 *La loi hypergéométrique résulte d'un tirage sans remise alors que la loi binomiale d'un tirage avec remise. A la limite, que l'on "remette les boules" dans l'urne ou pas, on obtient la loi binomiale.*

Démonstration. Tout d'abord, la variable X_k prend ses valeurs dans $[|\max\{0; n-k(1-p)\}, \min\{n, kp\}|]$. On a de plus

$$\begin{aligned}
\mathbb{P}[X_k = j] &= \frac{\binom{kp}{j}\binom{k(1-p)}{n-j}}{\binom{k}{n}} \\
&= \frac{(kp)!}{j!(kp-j)!} \times \frac{k(1-p)}{(n-j)!(k(1-p)-n+j)!} \times \frac{n!(k-n)!}{n!} \\
&= \binom{n}{j} \frac{(kp)!}{(kp-j)!} \times \frac{k(1-p)}{(k(1-p)-n+j)!} \times \frac{(k-n)!}{n!}.
\end{aligned}$$

En remarquant que $\frac{q!}{(q-l)!} = A_q^l \underset{q\to\infty}{\sim} q^l$, on obtient

$$\mathbb{P}[X_k = j] \underset{k\to\infty}{\sim} \binom{n}{k} \frac{(Np)^k (N(1-p))^{n-k}}{N^n} = \binom{n}{k} p^k (1-p)^{n-k}.$$

A partir d'un certain rang k, $n-k(1-p) \leq 0$ et $kp \geq n$, X_k prend donc ses valeurs dans $[|0,n|]$. D'après l'équivalence obtenue, pour tout $j \in [|0;n|], \mathbb{P}[X_k = j] \underset{k\to\infty}{\longrightarrow} \mathbb{P}[X = j]$, où X suit une loi binomiale. On en déduit la convergence simple des F_k vers F sur l'ensemble des points de continuité de F. En effet, si $x \in [0,n] \setminus [|0,n|]$, alors x est un point de continuité de F et

$$F_k(x) = \mathbb{P}[X_k \leq x] = \sum_{j=0}^{[x]} \mathbb{P}[X_k = j] \underset{k\to\infty}{\longrightarrow} \sum_{j=0}^{[x]} \mathbb{P}[X = j] = F(x).$$

Pour conclure, il reste à remarquer $F_k(x) = F(x) = 0$ si $x < 0$ et $F_k(x) = F(x) = 1$ si $x > n$. □

Théorème 5.1.10 (Approximation de la loi binomiale par la loi de Poisson) *Soit* $(X_k, k \geq 1)$ *une suite de variables aléatoires réelles. Chaque* X_k *suit une loi binomiale de paramètres* $k, \frac{\lambda}{k}$*. Le réel* λ *est fixé. Alors la suite* $(X_k, k \geq 1)$ *converge en loi vers* X *une variable aléatoire de Poisson de paramètre* λ*.*

démonstration. Dans la preuve du théorème précédent, on a vu qu'il était en fait suffisant pour les variables discrètes de montrer les convergences des probabilités $\mathbb{P}[X_k = j]$ vers $\mathbb{P}[X = j]$. On a pour tout $k \geq j$,

$$\begin{aligned}
\mathbb{P}[X_k = j] &= \binom{k}{j} p^j (1-p)^{k-j} \\
&= \binom{k}{j} \left(\frac{\lambda}{k}\right)^j \left(1 - \frac{\lambda}{k}\right)^{k-j} \\
&= \frac{\lambda^j}{j!} \frac{k!}{(k-j)!} \frac{\left(1-\frac{\lambda}{k}\right)^{-j}}{k^j} \left(1 - \frac{\lambda}{k}\right)^k \underset{k\to\infty}{\longrightarrow} \frac{\lambda^j}{j!} e^{-\lambda}.
\end{aligned}$$

La dernière ligne vient du fait que

$$\left(1 - \frac{\lambda}{k}\right)^k \underset{k\to\infty}{\longrightarrow} e^{-\lambda}$$

(à savoir démontrer) et de

$$\frac{k!}{(k-j)!} \frac{\left(1-\frac{\lambda}{k}\right)^{-j}}{k^j} \underset{k\to\infty}{\longrightarrow} 1.$$

□

Théorème 5.1.11 (Théorème central limite) *Soit* $(X_n, n \geq 1)$ *une suite de variables aléatoires réelles indépendantes et de même loi, admettant une espérance (notée* m*) et une variance (notée* σ^2*). Soit*

$$S_n = \sum_{i=1}^n X_i.$$

On a $\mathbb{E}[S_n] = nm$ *et* $V(S_n) = n\sigma^2$*. La suite des variables aléatoires centrées réduites :*

$$\left(\frac{S_n - nm}{\sqrt{n}\sigma}; n \geq 1\right)$$

converge en loi vers une loi gaussienne $\mathcal{N}(0,1)$*. Autrement dit :*

$$\mathbb{P}\left(a \leq \frac{S_n - nm}{\sqrt{n}\sigma} \leq b\right) \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-t^2/2} dt$$

Démonstration. La preuve la plus efficace repose sur les fonctions caractéristiques, outils que nous ne développons pas dans ce cours. Il existe des preuves *élémentaires*. Nous verrons en problème, un cas particulier. Pour le cas général, voir par exemple l'article de Jérôme Depauw [1]
Le théorème central limite met en avant le caractère universel de la loi gaussienne, il justifie la modélisation d'une erreur aléatoire par une loi normale. Souvent, une erreur aléatoire est modélisée par l'addition d'un grands nombres d'événements indépendants.

Proposition 5.1.12 (intervalle de confiance) *Supposons que* $\mathbb{V}(X_1) = 1$*, on déduit du théorème précédent le résultat suivant : la probabilité que* $\mathbb{E}[X_1]$ *se trouve dans l'intervalle aléatoire*

$$\left[\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{b}{\sqrt{n}}, \frac{1}{n}\sum_{i=1}^{n} X_i - \frac{a}{\sqrt{n}}\right]$$

est égale à

$$\mathbb{P}(a \leq Z \leq b) = \frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}\,dx.$$

Cette dernière vaut environ 0.95 *pour* $b = -a = 1.96$*. On dit alors que l'intervalle*

$$\left[\frac{1}{n}\sum_{i=1}^{n} X_i - \frac{1.96}{\sqrt{n}}, \frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1.96}{\sqrt{n}}\right]$$

est un intervalle de probabilité (on dit aussi intervalle de confiance) pour le paramètre de la moyenne $\theta = \mathbb{E}[X_1]$ *à 95%*

On termine ce chapitre en mentionnant la loi forte des grands nombres. Ce théorème dépasse le cadre du cours mais est fondamental en théorie des probabilités.

5.1.4 Convergence presque-sûre et loi forte des grands nombres

Définition 5.1.4 *On dit que la suite* $(X_n, n \geq 1)$ *converge presque-sûrement vers* X*, si*

$$\mathbb{P}\left(\{\omega \in \Omega; X_n(\omega) \underset{n\to\infty}{\longrightarrow} X(\omega)\}\right) = 1.$$

Cette notion de convergence est très forte. Elle implique la convergence en probabilité (mais n'est pas équivalente). Pour établir la convergence presque-sûre d'une suite de variables aléatoires des informations précises sur la « trajectoire »de la suite sont nécessaires (et rarement accessibles). On peut maintenant énoncer la loi forte des grands nombres.

Théorème 5.1.13 (Loi forte des grands nombres) *Soit* $(X_n, n \geq 1)$ *une suite de variables indépendantes avec un moment d'ordre* 1*, alors*

$$\lim \tfrac{1}{n}\textstyle\sum_{i=1}^{n} X_i = \mathbb{E}[X_1] \text{ presque sûrement.}$$

1. disponible ici : **www.lmpt.univ-tours.fr/∼depauw/TCLRMS_WEB.pdf.**

La démonstration de ce théorème requiert au minimum le *lemme de Borel-Cantelli*, directement issu des axiomes posés sur la probabilité $\mathbb{P}$ dans la définition. Si on considère une suite d'événements $(A_i, i \geq 1)$ indépendants de même probabilité p. La variable aléatoire

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}_{A_i}$$

correspond à la fréquence d'apparition des événements A_i pour i entre 1 et n, et $\mathbb{E}[X_1] = p$. La loi *faible* des grands nombres nous dit que cette fréquence converge en probabilité vers p. La loi forte nous dit que cette convergence est presque-sûre et nous permet de revenir à la définition heuristique d'une probabilité en tant que limite d'une fréquence d'apparition...

5.2 Exercices.

Exercice 5.2.1 *Soit $(X_n, n \geq 1)$ une suite de variables aléatoires de loi géométrique sur $\mathbb{N}^\star$ de paramètre $\frac{\theta}{n}$ avec $\theta \in \mathbb{R}_+^\star$. Montrer que la suite $\left(\frac{X_n}{n}\right)_{n\geq 1}$ converge en loi vers une v.a.r suivant la loi exponentielle de paramètre θ.*

Exercice 5.2.2 *Soit $(X_n, n \geq 1)$ une suite de v.a.r indépendantes identiquement distribuées de loi uniforme sur $[0,1]$. On note $m_n = \min\{X_1, ..., X_n\}$ et $M_n = \max\{X_1, ..., X_n\}$, montrer que $(m_n, n \geq 1)$ converge en probabilité vers 0 et $(M_n, n \geq 1)$ vers 1.*

Exercice 5.2.3 *Soit $(X_n, n \geq 1)$ une suite de v.a.r telle que $\mathbb{E}[X_n] \underset{n\to\infty}{\longrightarrow} a$ et $\mathbb{V}(X_n) \longrightarrow 0$, montrer que $X_n \underset{n\to}{\longrightarrow} a$ en probabilité.*

Exercice 5.2.4 *Soit f une fonction continue bornée de $\mathbb{R}$ dans $\mathbb{R}$. Montrer que*

$$\lim_{n\to\infty}\sum_{k=0}^{\infty} e^{-n\lambda}\frac{(n\lambda)^k}{k!} f\left(\frac{k}{n}\right) = f(\lambda).$$

Indication : utiliser une suite de variables de Poisson de paramètre λ, et utiliser la loi faible des grands nombres. La loi forte est hors programme.

Exercice 5.2.5 *Soit $(X_n)_{n\geq 1}$ une suite de v.a. i.i.d. suivant la loi normale $N(1,3)$. Montrer que la suite de terme général*

$$\frac{1}{n}\sum_{i=1}^{n} X_i e^{X_i}$$

converge en probabilités et déterminer sa limite.

Exercice 5.2.6 *1) Soit f une fonction uniformément continue de $\mathbb{R}^d$ dans $\mathbb{R}$ et $(X_n, n \in \mathbb{N})$ une suite qui converge en probabilité vers X, alors $(f(X_n), n \geq 0)$ converge en probabilité vers $f(X)$.*
2) Montrer que si $(X_n, n \geq 0)$ et $(Y_n, n \geq 0)$ convergent respectivement vers X et Y en probabilité alors $(X_n + Y_n, n \geq 1)$ converge vers $X + Y$ en probabilité.

3) Soit $(X_n, n \geq 1)$ une suite de v.a.r. indépendantes de même loi telle que $\mathbb{E}[|X_1|] < \infty$. Vérifier que

$$\frac{X_n}{n} = \frac{\sum_{i=1}^n X_i}{n} - \frac{n-1}{n}\frac{\sum_{i=1}^{n-1} X_i}{n-1},$$

en déduire à l'aide de la question 2), que $\frac{X_n}{n}$ converge en probabilité vers 0.

Exercice 5.2.7 *Soient $\{X_n\}_{n\geq 1}$ une suite de variables aléatoires i.i.d telles que $X_1 \sim \mathcal{P}(1)$. Pour tout entier naturel non nul n on pose $S_n = \sum_{k=1}^n X_k$.*

1. *Déterminer la limite en loi de $\left\{\frac{S_n - n}{\sqrt{n}}\right\}_{n\geq 1}$.*
2. *Montrer que*

$$e^{-n}\sum_{k=0}^n \frac{n^k}{k!} \underset{n\to+\infty}{\longrightarrow} \frac{1}{2}.$$

Exercice 5.2.8 (Une loi faible des grands nombres pour des variables non indépendantes) *Soit $(X_n, n \geq 1)$ une suite de v.a.r. indépendantes de même loi admettant un moment d'ordre 2. On pose $Y_n = X_n X_{n+1}$ et $S_n = \sum_{k=1}^n Y_k$.*
1) Les variables Y_n sont-elles indépendantes ?
2) Montrer que pour tout $\epsilon > 0$,

$$\mathbb{P}\left(|\frac{S_n}{n} - \mathbb{E}[Y_1]| > \epsilon\right) \leq \frac{1}{\epsilon^2 n^2} V(S_n)$$

3) En déduire que $\frac{S_n}{n}$ converge en probabilité vers $\mathbb{E}[X_1]^2$.

Exercice 5.2.9 (Inégalité de Hölder) Soient p et q deux réels positifs tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soient X et Y deux variables aléatoires possédant respectivement un moment d'ordre p et un moment d'ordre q. L'inégalité d'Hölder que l'on va démontrer est

$$\mathbb{E}[|XY|] \leq (\mathbb{E}[|X|^p])^{\frac{1}{p}} (\mathbb{E}[|Y|^q])^{\frac{1}{q}}$$

1) Soient x, y deux réels positifs. En utilisant la concavité de ln, c'est-à-dire pour tout $\lambda \in]0, 1[$

$$\ln(\lambda x + (1-\lambda)y) \geq \lambda \ln(x) + (1-\lambda)\ln(y),$$

montrer l'inégalité (appelée inégalité de Young)

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}.$$

2) En posant $\tilde{X} = \frac{X}{\mathbb{E}[|X|^p]^{1/p}}$ et $\tilde{Y} = \frac{Y}{\mathbb{E}[|Y|^q]^{1/q}}$, montrer l'inégalité d'Hölder.
3) On suppose que X a un moment d'ordre s. Soit $0 \leq r \leq s$. En appliquant l'inégalité d'Hölder aux variables $|X|^r$ et 1 avec $p = \frac{s}{r}$, et $q = \frac{p}{p-1}$ son conjugué, montrer

$$\mathbb{E}[|X|^r] \leq \mathbb{E}[|X|^s]^{1/p}.$$

En déduire que X a un moment d'ordre r pour tout $r \leq s$.

Exercice 5.2.10 [Loi forte des grands nombres] Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes telles que pour tout $k \geq 1$, $\mathbb{E}[X_k] = 0$ et $\mathbb{E}[X_k^4] \leq K$ où K est un réel positif indépendant de k. On pose $S_n = \sum_{k=1}^n X_k$, on va montrer que

$$\frac{S_n}{n} \underset{n\to\infty}{\longrightarrow} 0 \text{ p.s.}$$

et

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] \underset{n\to\infty}{\longrightarrow} 0.$$

1) En utilisant le résultat de la dernière question de l'exercice précédent. Justifier que $\mathbb{E}[X_i^2]$, $\mathbb{E}[X_i^3]$ ont un sens. Montrer que

$$\mathbb{E}[S_n^4] = \mathbb{E}\left[\sum_{k\geq 1} X_k^4 + 6\sum_{i<j} X_i^2 X_j^2\right].$$

2) Montrer que $\mathbb{E}[X_i^2]^2 \leq \mathbb{E}[X_i^4]$, en déduire

$$\mathbb{E}[S_n^4] \leq nK + 3n(n-1)K \leq 3Kn^2.$$

3) Montrer que

$$\mathbb{E}\left[\sum_{n\geq 1}(S_n/n)^4\right] \leq 3K\sum_{n\geq 1}\frac{1}{n^2}.$$

En déduire que

$$\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] \underset{n\to\infty}{\longrightarrow} 0.$$

4) Montrer que $\mathbb{E}\left[\sum_{n\geq 1}(S_n/n)^4\right] < \infty$ implique que

$$\mathbb{P}[S_n/n \underset{n\to\infty}{\longrightarrow} 0] = 1.$$

Exercice 5.2.11 Soient (X_i) une suite de variables aléatoires indépendantes de même loi de Poisson de paramètre 1.
1) Quelle est la loi de $S_n = \sum_{i=1}^n X_i$?
2) On pose $S_n^\star = \frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var(S_n)}}$. En utilisant le théorème central limite, déterminer

$$\lim \mathbb{P}(S_n^\star \leq 0).$$

3) En déduire un équivalent de $\sum_{k=0}^n \frac{n^k}{k!}$.

Exercice 5.2.12 On dispose d'un dé. On cherche à savoir s'il est truqué ou non. Pour cela on va étudier la différence entre la fréquence d'apparition de 6 lors de n lancers de dé et 1/6. Soit N_n le nombre de 6 obtenus parmi n lancers. Déterminer la loi de N_n. On définit f_n par :

$$f_n = \frac{N_n}{n}.$$

Déterminer l'espérance et la variance de f_n

1 En utilisant l'inégalité de Bienaymé-Tchebychev, montrer que

$$\mathbb{P}\left(|F_n - \frac{1}{6}| \geq 0.01\right) \leq \frac{5}{36n \times 0.01^2}$$

2) Estimer avec l'inégalité précédente, le nombre de lancers que l'on doit effectuer pour qu'au cours de ces lancers la fréquence d'apparition de 6 diffère de 1/6 d'au plus 1/100 avec un risque d'erreur inférieur à 5/100 ?

3) Estimer ce nombre avec le théorème central limite.

Chapitre 6

Eléments de statistiques

Nous nous concentrons sur la statistique inférentielle. La statistique descriptive est également au programme et fait régulièrement l'objet de questions à l'oral. Le lecteur pourra par exemple consulter le chapitre 5 du livre de Saporta Probabilités, analyse des données et Statistique.

6.1 Estimation

L'objectif de la statistique inférentielle est d'estimer les paramètres d'une loi de probabilité régissant une expérience aléatoire. Il s'agit de mettre en place les outils pour estimer des paramètres à partir des réalisations d'une expérience. Les théorèmes limites vus précédemment sont fondamentaux dans la théorie de l'estimation. Il y a deux types d'estimation, l'estimation ponctuelle : on approche directement le paramètre en donnant une valeur numérique et l'estimation par intervalle : la valeur recherchée est encadrée par des bornes aléatoires avec une probabilité donnée (par exemple 95%). Soit X une variable aléatoire et $X_1, X_2, ..., X_n$ n variables aléatoires de même loi que X. On cherche à partir des valeurs prises par ces n variables aléatoires à avoir des renseignements sur la loi de X. Si par exemple :

1) la variable aléatoire X suit une loi de Poisson, comment peut on estimer son paramètre λ à partir des n observations ?
2) la variable aléatoire X suit une loi normale, comment peut on estimer ses paramètres m et σ à partir des n observations ?
3) la variable aléatoire est inconnue mais possède une espérance, peut on estimer $\mathbb{E}[X]$?

Soit Θ l'ensemble des paramètres d'une loi. Par exemple, $\Theta = \mathbb{R}_+$ pour la loi exponentielle, ou $\Theta = [0, 1]$ pour la loi de Bernoulli.

Définition 6.1.1 *Soit X une variable aléatoire de loi à paramètres dans Θ. Un échantillon de taille n de X est un vecteur aléatoire $(X_1, ..., X_n)$ dont les composantes ont même loi que X. Etant donné un échantillon de taille n, on appelle estimateur de θ, une variable aléatoire de la forme*

$$T_n = f(X_1, ..., X_n)$$

La variable aléatoire T_n est un estimateur ***convergent*** *de θ si*

$$\forall \theta \in \Theta, \forall \alpha > 0, \mathbb{P}\left(|T_n - \theta| > \alpha\right) \underset{n\to\infty}{\longrightarrow} 0.$$

On définit maintenant la notion de biais d'un estimateur (au programme en BTS).

Définition 6.1.2 (Biais d'un estimateur) *Soit T_n un estimateur du paramètre θ,*

- *Si pour tout* $\theta \in \Theta, \mathbb{E}[T_n] = \theta$, *l'estimateur est dit sans biais.*
- *Si pour tout* $\theta \in \Theta, \mathbb{E}[T_n] \underset{n\to\infty}{\longrightarrow} \theta$, *l'estimateur est dit asymptotiquement sans biais.*

Remarque 6.1.1 *Il faut noter que la propriété d'être non biaisé n'est pas indispensable. Il arrive même que certains estimateurs biaisés donnent de meilleurs résultats !*

Proposition 6.1.1 *Tout estimateur asymptotiquement sans biais dont la variance tend vers* 0 *est convergent.*

Démonstration. Utiliser l'inégalité de Markov.

Définition 6.1.3 *Supposons que* T_n *a un moment d'ordre 2. Le risque quadratique de l'estimateur* T_n *est la quantité*

$$R_{T_n}(\theta) := \mathbb{E}[(T_n - \theta)^2]$$

Il mesure l'écart en moyenne quadratique entre θ et T_n. La notion de risque quadratique permet de comparer deux estimateurs : l'estimateur S_n est meilleur que T_n si

$$\forall \theta \in \Theta, R_{S_n}(\theta) \leq R_{T_n}(\theta).$$

Définition 6.1.4 – *Soit* $(X_1, ..., X_n)$ *un échantillon de taille* n *de* X.

- *La moyenne* ***empirique*** *est*

$$\overline{X}_n := \frac{1}{n}\sum_{i=1}^{n} X_i.$$

- *La variance* ***empirique*** *est*

$$S_n^2 := \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

- *La variance* ***empirique corrigée*** *est*

$$\Sigma_n^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X}_n)^2.$$

Proposition 6.1.2 *Soit* X *une variable aléatoire de moyenne* m *et d'écart-type* σ, *alors*

$$\mathbb{E}[\overline{X}_n] = m, \quad \mathbb{V}(\overline{X}_n) = \frac{\sigma^2}{n}$$
$$\mathbb{E}[S_n^2] = \frac{n-1}{n}\sigma^2, \quad \mathbb{E}[\Sigma_n^2] = \sigma^2.$$

Démonstration. Exercice.

Proposition 6.1.3 *Soit* X *une variable aléatoire de moyenne* m *et de variance* σ^2.

1) La moyenne empirique $\overline{X}_n$ est un estimateur sans biais et convergent de m. Si on se donne une réalisation ω, une estimation ponctuelle de la moyenne est donnée par

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

où $x_i = X_i(\omega)$ sont les données.

2) La variance empirique σ_n est un estimateur sans biais de la variance. Si on se donne une réalisation ω, une estimation ponctuelle de la variance est donnée par

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

où $x_i = X_i(\omega)$ sont les données.

3) La variance empirique corrigée Σ_n est un estimateur sans biais de la variance. Si on se donne une réalisation ω, une estimation ponctuelle de la variance est donnée par

$$\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

où $x_i = X_i(\omega)$ sont les données.

4) A ces trois estimateurs classiques, on peut rajouter celui du moment d'ordre r :

$$\hat{m}_r := \frac{1}{n}\sum_{i=1}^{n} x_i^r.$$

Démonstration.

1) Laissée au lecteur (appliquée la loi faible des grands nombres).
2) Laissée au lecteur.
3) Les variables X_i étant indépendantes de même loi, nous avons

$$\mathbb{E}[\Sigma_n] = \mathbb{E}\left[\frac{1}{n-1}\sum_{k=1}^{n}(X_k - \bar{X}_n)^2\right] = \frac{n}{n-1}\mathbb{E}[(X_1 - \bar{X}_n)^2].$$

Nous avons

$$\begin{aligned}\mathbb{E}[(X_1 - \bar{X}_n)^2] &= \mathbb{E}[X_1^2] - \frac{2}{n}\sum_{i=1}^{n}\mathbb{E}[X_1X_i] + \frac{1}{n^2}\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \\ &= \mathbb{E}[X_1^2] - \frac{2}{n}\sum_{i=1}^{n}\mathbb{E}[X_1X_i] + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}[X_iX_j].\end{aligned}$$

Les variables aléatoires (X_i) étant indépendantes et identiquement distribuées nous avons :

$$\mathbb{E}[X_iX_j] = \mathbb{E}[X_1^2] \text{ si } i = j$$
$$\mathbb{E}[X_iX_j] = \mathbb{E}[X_1]^2 \text{ si } i \neq j.$$

On a donc

$$-\frac{2}{n}\sum_{i=1}^{n}\mathbb{E}[X_1X_i] = -\frac{2}{n}\left[\mathbb{E}[X_1^2]+(n-1)\mathbb{E}[X_1]^2\right] \frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbb{E}[X_iX_j]$$
$$= \frac{1}{n^2}\sum_{i=1}^{n}\left[\mathbb{E}[X_1^2]+(n-1)\mathbb{E}[X_1]^2\right]$$
$$= \frac{1}{n}\left[\mathbb{E}[X_1^2]+(n-1)\mathbb{E}[X_1]^2\right].$$

On en déduit :

$$\mathbb{E}[(X_1-\bar{X}_n)^2] = \mathbb{E}[X_1^2] - \frac{1}{n}\left[\mathbb{E}[X_1^2]+(n-1)\mathbb{E}[X_1]^2\right] = \frac{n-1}{n}(\mathbb{E}[X_1^2]-\mathbb{E}[X_1]^2).$$

Finalement $\mathbb{E}[\Sigma_n^2] = \mathbb{E}X_1^2] - \mathbb{E}[X_1]^2 = \sigma^2$. □

Une méthode classique pour construire un estimateur est *la méthode des moments* basée sur une application directe de la loi des grands nombres. La méthode des moments consiste à exprimer θ en termes des moments $\mathbb{E}[X], \mathbb{E}[X^2], ...$ puis de remplacer chaque moment par le moment empirique lui correspondant. Cela fournit un estimateur convergent de θ. L'expression peut aussi faire intervenir σ^2 que l'on remplace alors par S_n^2 ou Σ_n^2.

En pratique, on calcule $\mathbb{E}[X], \mathbb{E}[X^2]$ etc « jusqu'à » obtenir à l'aide des moments une expression inversible de θ. On inverse alors cette expression, ce qui donne θ en fonction de $\mathbb{E}[X], \mathbb{E}[X^2], ...$ et on remplace $\mathbb{E}[X]$ par $\bar{X}_n$, $\mathbb{E}[X^2]$ par $\hat{m}_2$ (ou par $\Sigma_n^2 + \bar{X}_n^2$) etc.

Exemple 6.1.1 *Soit X une variable aléatoire prenant les valeurs $\{0,1,2\}$ avec les probabilités*

$$\mathbb{P}[X=0] = 1-p_1-p_2, \mathbb{P}[X=1] = p_1, \mathbb{P}[X=2] = p_2;$$

alors

$$\hat{p_2} = \frac{1}{2}\left(S_n^2 + \overline{X}_n^2 - \overline{X}_n\right)$$

est un estimateur sans biais du paramètre p_2.

Un estimateur classique est l'**estimateur du maximum de vraisemblance** (E.M.V). On suppose que la loi a pour paramètre $\theta \in \mathbb{R}$. Soit $(X_1, ..., X_n)$ un échantillon. Notons $L(x_1, ..., x_n; \theta)$ la densité de $(X_1, ...X_n)$. Lorsque $(x_1, ..., x_n)$ est fixé, $\theta \mapsto L(x_1, ..., x_n, \theta)$ est appelée vraisemblance de θ.

Définition 6.1.5 *On appelle estimateur du maximum de vraisemblance la valeur de θ qui rend maximale $L(x_1, ..., x_n, \theta)$.*

Puisque le logarithme est une fonction strictement croissante, on peut chercher à maximiser la fonction $\theta \mapsto -\ln L(x_1, ..., x_n, \theta)$. En pratique, un E.M.V est une solution de l'équation de la vraisemblance :

$$\frac{\partial}{\partial\theta}\ln L(X,\theta) = 0.$$

Intervalles de confiance et lois usuelles en statistiques

Nous avons déjà brièvement introduit la notion d'intervalle de confiance dans la proposition 5.1.12. Au lieu de donner une valeur (estimation ponctuelle) du paramètre θ on cherche deux bornes B_1 et B_2 entre lesquelles on espère que se trouve la vraie valeur du paramètre. Les bornes sont des fonctions des variables aléatoires $X_1, ..., X_n$.

Définition 6.1.6 – *Soit $B_1 = f_1(X_1, ..., X_n)$ et $B_2 = f_2(X_1, ..., X_n)$. On appelle intervalle de probabilité pour le paramètre θ le couple (B_1, B_2) et $\mathbb{P}(B_1 \leq \theta \leq B_2)$ est la probabilité de recouvrement (appelée plus souvent niveau de confiance).*

Si cette probabilité est égale à α, $[B_1, B_2]$ est appelé intervalle de probabilité $1-\alpha$. En général on choisit une valeur faible pour α et on construit l'intervalle de sorte que la probabilité de recouvrement soit égale à $1-\alpha$.

– *On appelle intervalle de confiance $1-\alpha$ (noté souvent $IC_{1-\alpha}(\theta)$ pour θ toute réalisation $[b_1, b_2]$ d'un intervalle de probabilité de recouvrement égale à $1-\alpha$ $[B_1, B_2]$.*

Pour comprendre le principe de construction d'un intervalle de confiance nous allons nous concentrer dans un premier temps sur le cas d'un intervalle de confiance pour la moyenne dans le cas gaussien.

Soit $(X_1, ..., X_n)$ un échantillon de loi normale $\mathcal{N}(\mu, \sigma^2)$. On suppose σ connu le paramètre à estimer est μ. Comme l'échantillon est gaussien, $\bar{X}_n$ suit une loi normale de paramètres $(\mu, \sigma^2/n)$. Pour pouvoir utiliser les tables gaussiennes, on définit $T = \frac{\bar{X}_n - \mu}{s/\sqrt{n}}$ qui suit une loi gaussienne standard. La variable T est appelée *statistique pivotale*, elle dépend de $X_1, ..., X_n$ et $\theta = \mu$ mais sa loi est indépendante de μ. Cette statistique est utile car elle va permettre d'encadrer le paramètre μ. Dans un premier temps, on observe que

$$\mathbb{P}[q_{\alpha/2} < T < q_{1-\alpha/2}] = 1-\alpha,$$

où q_p désigne le quantile d'ordre p pour la loi normale $\mathcal{N}(0,1)$. Comme cette loi est symétrique, on a $q_p = -q_{1-p}$ et on peut donc écrire que

$$\mathbb{P}\left(-q_{1-\alpha/2} \leq \sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \leq q_{1-\alpha/2}\right) = 1-\alpha.$$

En isolant μ pour l'encadrer, on obtient

$$\mathbb{P}\left(\bar{X}_n - q_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{X}_n + q_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha.$$

On pose

$$B_1 = \bar{X}_n - q_{1-\alpha/2}\frac{\sigma}{\sqrt{n}} \text{ et } B_2 = \bar{X}_n + q_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$$

l'intervalle $[B_1, B_2]$ est un intervalle de probabilité $1-\alpha$ pour la moyenne μ. Il est important de comprendre que ce sont B_1 et B_2 qui sont aléatoires et non pas μ. En pratique on utilise les réalisations des variables aléatoires, un intervalle de confiance est une réalisation $[b_1, b_2]$.

Remarque 6.1.2 *La largeur de l'intervalle de confiance est une fonction décroissante de α et de n.*

Deux lois utiles : la loi du khi-deux et la loi de Student

Définition 6.1.7 *Soit $d \geq 1$. Par définition, la loi du Khi-deux à d degrés de liberté est la loi de la variable aléatoire $Y_d = \sum_{i=1}^{d} N_i^2$ où $N_1, ..., N_d$ sont d variables gaussiennes centrées réduites.*

Proposition 6.1.4 *Soit Y_d une variable aléatoire qui suit une loi du khi-deux à d degrés de liberté, alors*
- *sa densité est donnée par $f_{Y_d}(x) = k_d x^{d/2-1} e^{-x/2} 1_{[0,\infty[}(x)$ avec $k_d = \frac{1}{2^{d/2}\Gamma(d/2)}$*
- *Y_d est intégrable et $\mathbb{E}[Y_d] = d$*
- *Si $d \geq 30, \mathbb{P}[\sqrt{Y_d} - \sqrt{2d-1} \in I] \approx \mathbb{P}[N \in I]$ où N est une Gaussienne standard.*

Définition 6.1.8 *On dit que la variable aléatoire T_d suit la loi de Student à d degrés de liberté si $T_d = N\sqrt{\frac{d}{Y_d}}$ où N et Y_d sont deux variables indépendantes de loi respective la loi normale centrée réduite, et la loi du khi-deux à d degrés de liberté.*

Proposition 6.1.5 *Soit T_d une variable aléatoire suivant la loi de Student. Sa densité est*

$$f_{T_d}(x) = c_d(1 + x^2/d)^{-\frac{d+1}{2}},$$

où

$$c_d = \frac{\Gamma(\frac{d+1}{2})}{\sqrt{\Pi d}}$$

$$p_{T_d}(x) \underset{d\to\infty}{\longrightarrow} \frac{e^{-x^2/2}}{\sqrt{2\pi}}.$$

Dans les applications, si $d \geq 30$ alors $\mathbb{P}[T_d \in I] \approx \mathbb{P}[N \in I]$.

Intervalle de confiance pour une moyenne et une variance.

Soient $X_1, ..., X_n$ n variables aléatoires gaussiennes, indépendantes et de même loi $\mathcal{N}(\mu, \sigma^2)$. Deux cas sont à envisager pour la détermination d'un intervalle de confiance : selon que σ soit connu ou non. Dans le premier cas, on a montré précédemment que $B_1 = \bar{X}_n - q_{1-\alpha/2}\frac{\sigma}{n}$ et $B_2 = \bar{X}_n + u_{1-\alpha/2}\frac{\sigma}{\sqrt{n}}$ forment les bornes d'un intervalle de probabilité $1-\alpha$. Dans le seoond cas, où σ est inconnu. On a que

$$\frac{(n-1)\Sigma_n^2}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$$

suit une loi du khi-deux à $n-1$ degrés de liberté. On définit $T = \sqrt{n}\frac{(\bar{X}_n - \mu}{\Sigma_n}$. On a

$$T = \frac{\sqrt{n}(\bar{X}_n - \mu)/\sigma}{\sqrt{(n-1)\frac{S^2}{\sigma^2(n-1)}}} = \frac{N}{K}$$

avec N de loi $\mathcal{N}(0,1)$ et K indépendante de loi du khi-deux à $n-1$ degrés de liberté. T est une statistique pivotale puisque sa loi est indépendante des paramètres. On en déduit un intervalle de probabilité $1-\alpha$ pour μ car on a $\mathbb{P}[t_{n-1;\alpha/2} < T < t_{n-1;1-\alpha/2}] = 1-\alpha$, où $t_{n-1;\alpha/2}$ est le quantile d'ordre $\frac{\alpha}{2}$ pour la loi de Student à $n-1$ degrés de liberté. On a donc

$$\mathbb{P}[t_{n-1:\alpha/2} < \sqrt{n}\frac{\bar{X}_n - \mu}{\Sigma_n} < t_{n-1,1-\alpha/2}] = 1-\alpha,$$

en isolant μ et en remarquant que $t_{n-1;\alpha/2} = -t_{n-1;1-\alpha/2}$ car la loi de Student est symétrique par rapport à 0, on obtient

$$\mathbb{P}\left(\bar{X}_n - t_{n-1;1-\alpha/2}\frac{\Sigma_n}{\sqrt{n}} < \mu < \bar{X}_n + t_{n-1;1-\alpha/2}\frac{\Sigma_n}{\sqrt{n}}\right) = 1-\alpha;$$

ce qui permet de conclure.
Nous terminons par la construction d'un intervalle de confiance pour une variance. La statistique pivotale utilisée est

$$T = \frac{(n-1)S^2}{\sigma^2}.$$

Elle suit une loi du Khi-deux à $n-1$ degrés de liberté. Elle est comprise entre les quantiles $r_{n-1,\alpha/2}$ et $r_{n-1,1-\alpha/2}$ avec probabilité $1-\alpha$, ce qui permet d'obtenir un intervalle de confiance $1-\alpha$ pour σ^2 :

$$IC_{1-\alpha}(\sigma^2) = \left[\frac{(n-1)\Sigma_n^2}{r_{n-1,1-\alpha/2}}; \frac{(n-1)\Sigma_n^2}{r_{n-1,\alpha/2}}\right].$$

Nous nous sommes limités aux variables aléatoires gaussiennes. Cependant en utilisant le théorème central limite, si nos variables possèdent une variance, on peut approcher des intervalles de confiance par ceux déterminés précédemment. Evidemment il s'agira d'approximation et le niveau de confiance ne sera pas exactement respecté mais ces approximation sont en général très bonnes pour $n \geq 30$.

6.2 Exercices

Exercice 6.2.1 (#) On a relevé le nombre de dents cariées chez 100 enfants âgés de 7 ans dans une région de France. Les résultats obtenus sont les suivants :

Nb de dents cariées	0	1	2	3	4	5	6	7
Nb d'enfants	30	34	14	10	4	5	1	2

Calculer la moyenne empirique et la variance empirique du nombre de dents cariées.

Solution. La moyenne empirique est

$$\overline{X} = \frac{30 \times 0 + 34 \times 1 + \cdots + 2 \times 7}{100} = 1,53.$$

et la variance empirique est

$$S^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k - \overline{X})^2 = \frac{1}{99}\left[30 \times (0-\overline{X})^2 + 34 \times (1-\overline{X})^2 + \cdots + 2 \times (7-\overline{X})^2\right] = 2,36.$$

Exercice 6.2.2 (#) Un statisticien observe le nombre d'ampoules défaillantes à la sortie d'une chaîne de production. Il veut estimer la probabilité p qu'une ampoule soit défaillante. Pour cela, il compte le nombre de N_n , égaux à zero. Il propose d'estimer par : $p_n = N_n/n$

1. Montrer que p_n est un estimateur sans biais de p.
2. Calculer son risque quadratique.

Solution. On a $N_n = X_1 + \cdots + X_n$ où $X_i = 1$ si l'ampoule i est défaillante et 0 sinon. Les v.a. sont donc toutes de loi de Bernoulli de paramètre p. On suppose de plus qu'elles sont indépendantes.

1. Si les X_i sont i.i.d., N_n est une variable binomiale de paramètres p et n. On a alors
$$\mathbb{E}\left[\frac{N_n}{n}\right] = \frac{1}{n}\mathbb{E}[N_n] = \frac{np}{n} = \boxed{p}$$
et l'estimateur est bien sans biais.
2. Le risque quadratique est par définition
$$\mathbb{E}\left[(p_n - p)^2\right] = \mathbb{E}\left[\left(\frac{N_n}{n} - p\right)^2\right] = \frac{1}{n^2}\mathbb{E}\left[(N_n - np)^2\right] = \frac{1}{n^2}\mathbb{V}(N_n) = \frac{np(1-p)}{n^2} = \boxed{\frac{p(1-p)}{n}}.$$

Exercice 6.2.3 (#) Considérons un échantillon de taille 1 issu d'une population suivant une loi de Poisson de paramètre $\lambda > 0$ (ceci revient donc simplement à considérer une variable aléatoire X avec $X \sim \mathcal{P}(\lambda)$).

1. Montrer que l'estimateur $\delta(X) = 1_{[X=0]}$ est non biaisé pour $e^{-\lambda}$.
2. On cherche à présent à estimer $e^{-3\lambda}$. Que pouvez-vous dire sur la statistique $(-2)^X$?
3. Proposer en toute généralité un estimateur sans biais pour $e^{-n\lambda}$.

Solution.

1. On a $\mathbb{E}[\delta(X)] = \mathbb{E}[\mathbf{1}_{\{X=0\}}] = \mathbb{P}(X = 0) = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-\lambda}$.
2. De même,
$$\mathbb{E}\left[(-2)^X\right] = \sum_{k\geq 0}(-2)^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{-2\lambda} = e^{-3\lambda}.$$
3. On vérifie que $(1-n)^X$ est un estimateur sans biais de $e^{-n\lambda}$.

Exercice 6.2.4 (#) On considère un échantillon $(X_i)_{1\leq i\leq n}$ avec $n \geq 2$ le modèle avec P_K et tel que les v.a. sont i.i.d. de loi uniforme sur $\{1, \ldots, K\}$. Le but de cet exercice est de proposer deux estimateurs du paramètre $K > 0$.

1. A l'aide de l'espérance de $X \sim P_K$, proposer un estimateur $\hat{K}$ de K. Déterminer son risque quadratique $R_K(\hat{K}, K)$.
2. (a) Montrer que l'E.M.V. est $\tilde{K} = \max(X_1, \ldots, X_n)$.
   (b) Montrer que sous P_K, $\tilde{K}$ vérifie
$$P(\tilde{K} \leq k) = \frac{k^n}{K^n}$$

(c) Soit X une variable aléatoire à valeurs dans $\{1,\dots,K\}$. Montrer que

$$E[X]=\sum_{k=0}^{K-1}P(X>k).$$

(d) En deduire le biais de $\tilde{K}$.

Solution.

1. La moyenne empirique est un estimateur sans biais de l'espérance des $(X_i)_{1\leq i\leq n}$ qui sont uniformes sur $\{1,\dots,K\}$. Or, cette espérance vaut $\mathbb{E}[X_1]=(K+1)/2$. On propose donc comme estimateur de K :

$$\boxed{\hat{K}=2\overline{X}-1=\frac{2}{n}\sum_{i=1}^{n}X_i-1}.$$

Cet estimateur est sans biais (puisque la moyenne empirique l'est) donc le risque quadratique de $\hat{K}$ est sa variance :

$$R_K(\hat{K},K)=\mathbb{V}\left(\frac{2}{n}\sum_{i=1}^{n}X_i-1\right)=\frac{4}{n^2}\mathbb{V}\left(\sum_{i=1}^{n}X_i\right)=\frac{4}{n}\mathbb{V}\left(X_1\right)$$

car les X_i sont indépendantes et de même loi. De plus,

$$\mathbb{E}[X^2]=\frac{1}{K}\sum_{k=1}^{K}k^2=\frac{1}{K}\frac{K(K+1)(2K+1)}{6}=\frac{(K+1)(2K+1)}{6}$$

et donc

$$\mathbb{V}(X)=\frac{(K+1)(2K+1)}{6}-\left(\frac{K+1}{2}\right)^2=\frac{K^2-1}{12}$$

et le risque quadratique est

$$\boxed{R_K(\hat{K},K)=\frac{K^2-1}{3n}}.$$

2. (a) On a

$$L(K,k_1,\dots,k_n)=\mathbb{P}(X_1=k_1)\mathbb{P}(X_2=k_2)\cdots\mathbb{P}(X_n=k_n)=\begin{cases}1/K^n & \text{si } \forall i,\ k_i\leq K\\ 0 & \text{sinon.}\end{cases}$$

Or, l'assertion $\big(\forall i,\ k_i\leq K\big)$ est équivalente à $\big(\max(k_1,k_2,\dots,k_n)\leq K\big)$.
La fonction L est donc nulle jusqu'à $\max(k_1,\dots,k_n)$ puis est décroissante et strictement positive.
Son maximum est donc atteint en $\max(k_1,\dots,k_n)$ et l'E.M.V. est $\boxed{\tilde{K}=\max(X_1,\dots,X_n)}$.

(b)

$$\mathbb{P}(\tilde{K}\leq k)=\mathbb{P}(\max(X_1,\dots,X_n)\leq k)=\mathbb{P}(X_1\leq k,\dots,X_n\leq k)=\mathbb{P}(X_1\leq k)^n=\frac{k^n}{K^n}$$

car les variables sont indépendantes et de loi uniforme.

3. On a
$$\sum_{k=0}^{K-1} \mathbb{P}(X > k) = \sum_{k=0}^{K-1} \sum_{j=k+1}^{K} \mathbb{P}(X = j) = \sum_{k=0}^{K-1} \sum_{j=k}^{K-1} \mathbb{P}(X = j + 1)$$
En inversant les 2 signes de sommation, on a
$$\sum_{k=0}^{K-1} \mathbb{P}(X > k) = \sum_{j=0}^{K-1} \sum_{k=0}^{j} \mathbb{P}(X = j{+}1) = \sum_{j=0}^{K-1} \mathbb{P}(X = j{+}1) \sum_{k=0}^{j} 1 = \sum_{j=0}^{K-1} \mathbb{P}(X = j{+}1)(j{+}1) = \mathbb{E}[X].$$

4. En utilisant iii) et ii), le biais de $\tilde{K}$ est
$$\mathbb{E}[\tilde{K}] - K = \sum_{k=0}^{K-1} \mathbb{P}(\tilde{K} > k) - K = \sum_{k=0}^{K-1} (1 - \mathbb{P}(\tilde{K} \leq k)) - K = \boxed{\frac{1}{K^n} \sum_{k=0}^{K-1} k^n}.$$
Le biais est donc strictement positif : $\tilde{K}$ n'est pas sans biais.

Exercice 6.2.5 (#) Pour $n = 1$, on considère le modèle des lois binomiales sur l'espace $\mathcal{X} = \{0, \ldots, m\}, m \in \mathbb{N}$ et
$$f(x, \theta) = \binom{m}{x} \theta^x (1 - \theta)^{m-x}, x \in \mathcal{X},$$
avec le paramètre $\theta \in [0, 1]$.

1. Montrer qu'une fonction $g(\theta)$ admet un estimateur sans biais si et seulement si g est un polynôme de degré $< m$, et dans ce cas un tel estimateur est unique.
2. Soit $g(\theta) = \theta^2$. Montrer que l'estimateur sans biais associé s'annule en $x = 0$ et $x = 1$.

Solution On considère une variable X_1 de loi binomiale de paramètres m et θ avec θ à estimer.

1. Dire que $g(\theta)$ admet un estimateur sans biais signifie qu'il existe une fonction Φ : $\{0, \ldots, m\} \to \mathbb{R}$ telle que
$$\mathbb{E}\left[\Phi(X_1)\right] = g(\theta).$$
Or, comme X_1 est une v.a. binomiale de paramètres m et θ, on a
$$\mathbb{E}\left[\Phi(X_1)\right] = \sum_{k=0}^{m} \Phi(k) \binom{m}{k} \theta^k (1 - \theta)^{m-k}$$
et donc s'il existe Φ telle qu'on ait l'égalité $\sum_{k=0}^{m} \Phi(k) \binom{m}{k} \theta^k (1 - \theta)^{m-k} = g(\theta)$, g est nécessairement un polynôme de degré strictement inférieur à m. La réciproque et l'unicité découlent du fait que la famille de polynômes $(X^k (1 - X)^{m-k})_{0 \leq k \leq m}$ est une base de l'espace vectoriel des polynômes de degrés inférieurs ou égaux à m (c'est une famille libre et elle a $m + 1$ éléments).

2. D'après ce qui précède, on doit avoir $\displaystyle\sum_{k=0}^{m} \Phi(k) \binom{m}{k} \theta^k (1-\theta)^{m-k} = \theta^2$. En prenant $\theta = 0$, on a $\Phi(0) = 0$. On a alors l'égalité :

$$\theta^2 = \sum_{k=1}^{m} \Phi(k) \binom{m}{k} \theta^k (1-\theta)^{m-k} = \theta \sum_{k=1}^{m} \Phi(k) \binom{m}{k} \theta^{k-1} (1-\theta)^{m-k}.$$

Comme 0 est racine double du polynôme $\theta \sum_{k=1}^{m} \Phi(k) \binom{m}{k} \theta^{k-1} (1-\theta)^{m-k}$, on doit avoir

$$0 = \sum_{k=1}^{m} \Phi(k) \binom{m}{k} 0^{k-1} (1-0)^{m-k} = \Phi(1).$$

Chapitre 7

Sujets des examens passés et problèmes

Epreuve de probabilités

Calculs d'intégrales et notion d'entropie

Exercice 7.0.6 Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

1) Supposons $\Omega = \{1, ..., n\}$. On définit l'entropie de la probabilité $\mathbb{P}$ par

$$H(\mathbb{P}) = -\sum_{k=1}^{n} p_k \ln(p_k)$$

où $p_k = \mathbb{P}(\{k\})$. Avec pour convention $0 \ln(0) = 0$.

a) Montrer que H est à valeurs dans $\mathbb{R}_+$ et trouver $\mathbb{P}$ telle que $H(\mathbb{P}) = 0$.
b) Calculer $H(\mathbb{P})$ lorsque $\mathbb{P}$ est la loi uniforme sur $\{1, ..., n\}$.
c) On va montrer que la probabilité uniforme réalise le maximum de l'entropie sur $\Omega = \{1, ..., n\}$.

On rappelle que la fonction ln est concave. On **admet** que pour tous λ_k réels positifs tels que $\sum_{k=1}^{n} \lambda_k = 1$ et tous $x_1, ..., x_k$ positifs :

$$\sum_{k=1}^{n} \lambda_k \ln(x_k) \leq \ln\left(\sum_{k=1}^{n} \lambda_k x_k\right).$$

Soit $(q_k, k \in [|1, n|])$ et $(p_k, k \in [|1, n|])$ deux probabilités sur $\{1, ..., n\}$. Démontrer à l'aide de l'inégalité ci-dessus que

$$\sum_{k=1}^{n} q_k \ln(p_k) \leq \sum_{k=1}^{n} q_k \ln(q_k).$$

Conclure avec $p_k = 1/n$ pour tout $k \in \{1, ...n\}$.

2) a) Rappeler la définition d'une densité de probabilité sur $\mathbb{R}$.

Soit f une densité de probabilité sur $\mathbb{R}$. On note

$$H_f = -\int_{\mathbb{R}} f(x) \ln(f(x)) dx,$$

lorsque cette intégrale a un sens et $H_f = \infty$ sinon. La quantité H_f est appelée entropie de la loi de densité f.

b) Calculer l'entropie de la loi normale $\mathcal{N}(0,1)$. On rappelle que la densité de la loi normale centrée réduite est $x \mapsto \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$, x variant dans $\mathbb{R}$.

c) Soient f et g deux fonctions strictement positives. Ecrire (et justifier) l'égalité

$$\frac{g(x)}{f(x)} = 1 + \frac{g(x) - f(x)}{f(x)}.$$

En utilisant l'inégalité $\ln(1+z) \leq z$ pour tout $z > -1$, établir que si f, g sont des densités

$$\int_{\mathbb{R}} \ln(g(x)/f(x)) f(x) dx \leq 0.$$

d) En prenant dans l'inégalité ci-dessus g la densité de la loi normale, démontrer que la loi normale réalise le maximum d'entropie sur les densités vérifiant $\int_{\mathbb{R}} x f(x) dx = 0$ et $\int_{\mathbb{R}} x^2 f(x) dx = 1$.

En thermodynamique, l'entropie mesure l'état de désorganisation interne d'un système. Boltzmann (1870) a montré qu'un système isolé évolue spontanément vers l'état le plus désordonné. La théorie de l'information a été introduite par Shannon (1948) qui associe entropie et information manquante sur le système : toute évolution spontanée d'un système isolé se fait avec perte d'information sur le système. De façon intuitive, les lois gaussiennes reflètent un manque d'information. Dans une certaine mesure, cela justifie le choix de cette fonction comme "fonction d'erreur".

Calculs binomiaux

On considère un serveur informatique consommant 500 watts. Pour améliorer la sûreté de fonctionnement d'un serveur informatique, on envisage d'introduire de la *redondance* (*), c'est-à-dire d'équiper le serveur de plusieurs exemplaires d'un même composant. On étudie deux configurations :

(I) **Redondance des alimentations :** On met 3 alimentations de 300 watts chacune.
(II) **Redondance des disques durs :** Le serveur est configuré en RAID5 (**), c'est-à-dire qu'il dispose de 4 disques durs et continue à fonctionner avec un disque dur en panne.

On suppose que la probabilité qu'une alimentation tombe en panne est p et celle d'un disque dur est q. On suppose que tous les composants sont indépendants.

1) Soit un serveur avec la configuration (I)

a) Combien d'alimentations peuvent tomber en panne sans que le serveur ne tombe en panne ?
b) Pour tout $i \in \{1, 2, 3\}$, on définit A_i la variable aléatoire qui vaut 1 si la i-ème alimentation tombe en panne, 0 sinon. Quelle est la loi de A_i ?

c) On note N_A le nombre d'alimentations en panne. Exprimer N_A en fonction de A_1, A_2, A_3. Quelle est la loi de N_a ?

d) On suppose qu'aucun composant différent des alimentations ne peut tomber en panne. Justifier que la probabilité de panne (notée $P_a(p)$) est donnée par $\mathbb{P}(N_a \geq 2)$ et montrer que

$$P_a(p) = p^2(3 - 2p).$$

2) Soit un serveur en configuration (II).

a) Soit N_D le nombre de disques durs en panne. Déterminer la loi de N_D (vous pourrez introduire des variables aléatoires de Bernoulli indépendantes de paramètre q).

b) On suppose qu'aucun composant différent des disques durs ne peut tomber en panne. Montrer que la probabilité que le serveur tombe en panne (notée $P_D(q)$) vérifie

$$P_D(q) = q^2(3q^2 - 8q + 6).$$

3) On suppose $p = q$. En étudiant le signe de la fonction $p \mapsto P_A(p) - P_D(p)$, décider quelle configuration parmi (I) et (II) est la plus fiable.

Tranport d'information : loi binomiale/loi de Poisson

Soit $n \in \mathbb{N}^\star$. On considère un système comportant n transmetteurs **indépendants** : $I_1, ..., I_n$. On suppose que I_1 porte l'information 0. Il communique avec I_2 et lui donne l'information qu'il porte (c'est-à-dire 0) avec probabilité $p \in]0, 1[$ ou l'information contraire (c'est-à-dire 1) avec probabilité $1 - p$. Soit $2 \leq k \leq n$, chaque I_k reçoit ainsi une information de I_{k-1} et :

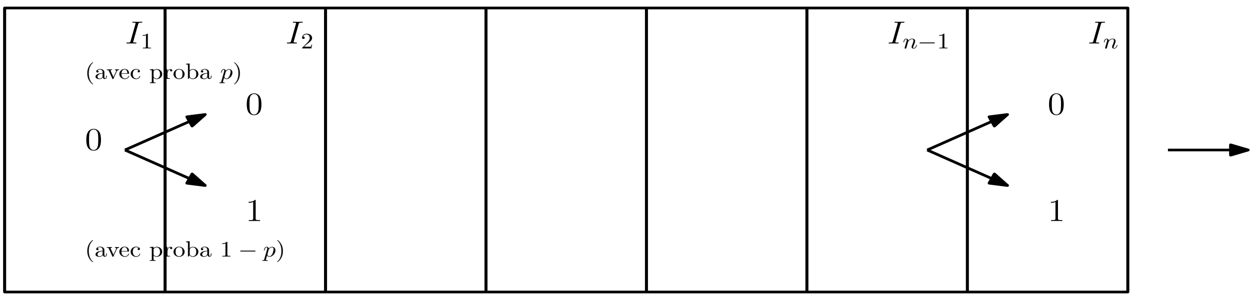


FIGURE 7.1 – Modèle de transmission

- transmet cette information à I_{k+1}, avec probabilité p
  ou
- transmet l'information contraire à I_{k+1}, avec probabilité $1 - p$ (on dira qu'il *ment*).

On récupère l'information à la sortie de I_n.

1) On pose

$$T_n := \min\{k \in \{1, ..., n\}; I_k \text{ ment}\} \text{ et } T_n = 0 \text{ si aucun transmetteur ne ment.}$$

Quelles sont les valeurs prises par T_n ? Déterminer la loi de T_n.

Les questions qui suivent sont indépendantes de la question 1. On s'intéresse maintenant à la probabilité que l'information reçue coïncide avec l'information de départ (c'est-à-dire **0**).

Première méthode

2) Soit u_n la probabilité que I_n délivre la bonne information. Enoncer la formule des probabilités totales et prouver la relation de récurrence suivante

$$u_n = pu_{n-1} + (1-p)(1-u_{n-1}).$$

3) Démontrer par récurrence que pour tout $n \geq 1$, $u_n = \frac{1+(2p-1)^n}{2}$ et étudier la limite de cette suite lorsque $n \to \infty$.

Deuxième méthode

4) Soit $x, y \in \mathbb{R}_+$. Rappeler la formule du binôme de Newton et à l'aide de celle-ci montrer que

$$\frac{1}{2}\left((x+y)^n + (y-x)^n\right) = \sum_{k=0}^{[n/2]} \binom{n}{2k} x^{2k} y^{n-2k}$$

$[n/2]$ désigne la partie entière du réel $n/2$.

5) Pour tout $i \geq 1$, on note X_i la variable aléatoire définie de la façon suivante.

$$X_i = 1 \text{ si } I_i \text{ ment, } X_i = 0 \text{ sinon.}$$

En justifiant soigneusement, donner la loi (avec ses paramètres) de la variable aléatoire $S_n = \sum_{i=1}^n X_i$.

6) Justifier que la variable aléatoire S_n est le nombre de « transmetteurs menteurs » et justifier l'égalité des événements « I_n donne 0 »= « S_n a une valeur paire ».

7) A l'aide de la question 4, calculer la probabilité qu'une variable aléatoire binomiale de paramètres n, p ait une valeur paire. En déduire la probabilité que I_n délivre la bonne information.

Approximation par une loi de Poisson

On suppose maintenant que la probabilité qu'un transmetteur ne mente pas est $p_n = \lambda/n$, avec $\lambda > 0$.

8) Soit $(S_n)_{n\geq 1}$ une suite de variables aléatoires de loi binomiale (n, p_n). Démontrer la propriété du cours suivante : pour tout $k \in \mathbb{N}$, $\mathbb{P}(S_n = k) \underset{n\to\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^k}{k!}$. On rappelle que $\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}$.

9) Démontrer que

$$\frac{e^{-\lambda} + e^{\lambda}}{2} = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!}.$$

10) Soit Z une variable aléatoire de loi de Poisson de paramètre λ. On rappelle que cela signifie $\mathbb{P}(Z = k) = e^{-\lambda}\frac{\lambda^k}{k!}$. A l'aide de la question précédente, montrer que $\mathbb{P}(Z \text{ est paire}) = \frac{e^{-2\lambda}+1}{2}$.

11) Soit $(a_n)_{n\geq 1}$ la suite définie par $a_n = \mathbb{P}(I_n \text{ donne } 0)$. Déterminer sa limite lorsque n tend vers l'infini. Entre le modèle où la probabilité de transmission est p et ne dépend pas de n et ce modèle, quel est celui qui transmet le bon message avec la plus grande probabilité lorsque n tend vers l'infini ?

QCM et loi hypergéométrique

0) Soient X et Y deux variables aléatoires discrètes à valeurs dans $\{0, ..., m\}$. Rappeler la définition de $\mathbb{E}[Y]$. Soit $x \in \{0, ..., m\}$, on note $\mathbb{E}_{(X=x)}[Y]$ l'espérance de Y par rapport à la probabilité conditionnelle $\mathbb{P}_{(X=x)}$: c'est-à-dire

$$\mathbb{E}_{(X=x)}[Y] = \sum_{y=0}^{m} y\mathbb{P}_{(X=x)}(Y = y), \text{ montrer que } \mathbb{E}[Y] = \sum_{x=0}^{m} \mathbb{E}_{(X=x)}[Y]\mathbb{P}(X = x).$$

On fait passer un examen sous forme d'un QCM (questions à choix multiples). L'examen se compose de 20 questions tirées au hasard et *sans remise* parmi 100 questions possibles recouvrant le programme. A chaque question est proposée 4 réponses possibles, *seule une réponse est correcte*. Si le candidat choisit la bonne réponse il a un point, sinon 0 point.

On considère un candidat ayant appris une proportion p du programme ($100p \in \mathbb{N}$).

1) On note X le nombre de questions figurant parmi les 20 que le candidat a révisées. Justifier brièvement que X suit une loi hypergéométrique de paramètres $(100, 20, p)$

$$\mathbb{P}[X = x] = \frac{\binom{100p}{x}\binom{100(1-p)}{20-x}}{\binom{100}{20}}.$$

Rappeler la signification du coefficient binomial $\binom{n}{k}$ et justifier brièvement la formule ci-dessus.

Ces X questions rapportent X points au candidat. On suppose dans la suite que $X(\Omega) = \{0, ..., 20\}$.

2) Etant donnée une question non révisée, le candidat choisit uniformément au hasard une réponse parmi les 4 possibles. Quelle est la probabilité que le candidat choisisse la bonne réponse ?

3) Le nombre de questions non révisées par le candidat est $20 - X$. On les numérote par $\{1, 2, ..., 20 - X\}$. A toute question $i \in \{1, 2, ..., 20 - X\}$, on associe une variable aléatoire X_i suivant une loi de Bernoulli de paramètre $1/4$: $X_i = 1$ si le candidat répond bien à la question i, $X_i = 0$ sinon. Les variables $(X_i, i \geq 1)$ sont indépendantes.

On note Y la somme des points obtenus en répondant aux questions **non révisées**.
a) Exprimer Y, à l'aide d'une somme, en fonction des X_i et de X.
b) Soit $x \in \{0, ..., 20\}$, sachant $X = x$, justifier que Y suit une loi binomiale de paramètre $(20 - x, 1/4)$.
c) Que vaut $\mathbb{E}_{(X=x)}[Y]$? (on rappelle que l'espérance d'une loi binomiale de paramètre (m, q) est mq).
d) En utilisant la formule $\mathbb{E}[Y] = \sum_{x=0}^{20} \mathbb{E}_{(X=x)}[Y]\mathbb{P}(X = x)$, montrer que $\mathbb{E}[Y] = 5 - 5p$.
Indication : On rappelle que $\mathbb{E}[X] = 20p$ *(vu en cours).*

4) On note N la note du candidat. Exprimer N en fonction de X et Y et déterminer l'espérance de N en fonction de p. Pour quelle valeur de p a-t-on $\mathbb{E}[N] = 10$?

5) On détermine maintenant la loi de $N = X + Y$. Montrer les trois égalités

$$\mathbb{P}[N = n] = \sum_{x=0}^{n} \mathbb{P}[X = x, Y = n - x] = \sum_{x=0}^{n} \mathbb{P}[X = x]\mathbb{P}_{(X=x)}[Y = n - x]$$
$$= \sum_{x=0}^{n} \frac{\binom{100p}{x}\binom{100(1-p)}{20-x}}{\binom{100}{20}} \binom{20 - x}{n - x} \left(\frac{1}{4}\right)^{n-x} \left(\frac{3}{4}\right)^{20-n}.$$

6) (Question bonus) On suppose maintenant que l'on enlève 1/3 de point à chaque mauvaise réponse et on note N' la note obtenue. Sans chercher la loi explicite de la note, mais en introduisant à la place des X_i, la variable aléatoire X_i' prenant comme valeur $-1/3$ avec probabilité 3/4 et 1 avec probabilité 1/4 démontrer que

$$\mathbb{E}[N'] = 20p.$$

---

* la redondance consiste à disposer plusieurs exemplaires d'un même équipement ou d'un même processus ou de tout autre élément participant à une solution électronique, mécanique ou industrielle. Selon les circonstances elle est utile :
1. pour augmenter la capacité totale ou les performances d'un système,
2. pour réduire le risque de panne,
3. pour combiner ces deux effets.

** En informatique, le mot RAID désigne les techniques permettant de répartir des données sur plusieurs disques durs afin d'améliorer soit la tolérance aux pannes, soit la sécurité, soit les performances de l'ensemble, ou une répartition de tout cela. La technologie RAID a été élaborée par un groupe de chercheurs de l'université de Californie à Berkeley en 1987. Ils étudièrent la possibilité de faire reconnaître deux disques durs ou plus comme une seule entité par le système. Ils obtinrent pour résultat un système de stockage aux performances bien meilleures que celles des systèmes à disque dur unique, mais doté d'une très mauvaise fiabilité. Les chercheurs s'orientèrent alors vers des architectures redondantes, afin d'améliorer la tolérance aux pannes du système de stockage.
(Source Wikipedia)

7.1 Fonctions additives et loi exponentielle

Préliminaire sur les fonctions additives

Considérons Φ une fonction *monotone* additive de $\mathbb{R}$ i.e. $\Phi(x + y) = \Phi(x) + \Phi(y)$ pour tous $x, y \in \mathbb{R}$.
a) Montrer que $\Phi(q) = q\Phi(1)$ pour tout $q \in \mathbb{Q}$.
b) Montrer que si $\Phi(1) = 0$ alors Φ est identiquement nulle.
c) Montrer que si $\Phi(1) \neq 0$ alors $\Phi(x) = \Phi(1)x$
d) Soit Ψ une fonction monotone vérifiant $\Psi(x+y) = \Psi(x)\Psi(y)$ pour tous $x, y \in \mathbb{R}$. Démontrer qu'il existe a tel que $\Psi(x) = e^{ax}$.

l'objectif de cette partie est de caractériser la loi exponentielle par son *absence de mémoire*. Soit X une variable aléatoire telle que :

$$\mathbb{P}_{X>s}[X > t + s] = \mathbb{P}[X > t] \tag{7.1}$$

e) Montrer que la loi exponentielle vérifie (7.1).
f) Montrer inversement que si X vérifie (7.1) alors X a une loi exponentielle.

7.2 Théorème de Weierstrass

Weierstrass a démontré que toute fonction continue sur le segment $[0, 1]$ est limite uniforme sur ce segment d'une suite de polynômes. Bernstein a découvert une nouvelle démonstration qui utilise les probabilités. On commence par des résultats basiques d'analyse. Pour toute fonction continue f sur $[0, 1]$,

$$||f||_\infty := \sup_{x\in[0,1]} |f(x)|.$$

Soit f une fonction continue sur $[0, 1]$, on définit son module de continuité par

$$m : \delta \mapsto \sup\{|f(u) - f(v)|; |u - v| \leq \delta\}.$$

1) Montrer que $|f(x) - f(y)| \leq 2||f||_\infty$ pour tout $x, y \in [0, 1]$.

2) Montrer que si $|x - y| \leq \delta$ alors $|f(x) - f(y)| \leq m(\delta)$. Montrer que $m(\delta) \underset{\delta\to 0}{\longrightarrow} 0$.

3) Soit $(b_n)_{n\geq 0}$ une suite réelle, rappeler la définition avec les quantificateurs de $b_n \underset{n\to\infty}{\longrightarrow} 0$.
Soit $n \geq 1$, on considère le polynôme

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} x^k (1 - x)^{n-k} f\left(\frac{k}{n}\right).$$

4) Soient $X_1, ..., X_n$ des variables aléatoires indépendantes de carré intégrable et

$$S_n = \sum_{i=1}^{n} X_i.$$

Montrer que

$$\mathbb{P}\left(|\frac{S_n}{n} - \mathbb{E}[X]| > t\right) \leq \frac{\mathbb{V}(X)}{nt^2}.$$

En déduire la loi faible des grands nombres.

5) On prend maintenant des variables X_i de Bernoulli de paramètre x. Montrer que

$$\mathbb{P}\left(|\frac{S_n}{n} - x| > t\right) \leq \frac{1}{4nt^2}.$$

6) Montrer que

$$\mathbb{E}\left[f\left(\frac{S_n}{n}\right)\right] = B_n(x).$$

7) Montrer que

$$f(x) - B_n(x) = \mathbb{E}\left[f(x) - f\left(\frac{S_n}{n}\right)\right],$$

en déduire l'inégalité

$$|f(x) - B_n(x)| \leq \mathbb{E}\left[|f(x) - f\left(\frac{S_n}{n}\right)|\right].$$

8) Soit $\delta > 0$, en considérant l'événement $A = \{x \in [0,1]; |x - \frac{S_n}{n}| \leq \delta\}$, montrer que

$$\mathbb{E}\left[|f(x) - f\left(\frac{S_n}{n}\right)|\right] \leq m(\delta) + 2||f||_\infty \mathbb{E}[1_{|x - \frac{S_n}{n}| > \delta|}].$$

9 En utilisant la question (4), déduire que

$$\mathbb{E}\left[|f(x) - f\left(\frac{S_n}{n}\right)|\right] \leq m(\delta) + \frac{||f||_\infty}{2n\delta^2}.$$

10) Montrer que

$$||f - B_n||_\infty \leq m(\delta) + \frac{||f||_\infty}{2n\delta^2}.$$

Soit $\epsilon > 0$.
11) Montrer qu'il existe δ_ϵ tel que $m(\delta_\epsilon) \leq \epsilon/2$.
12) Déterminer n_0 tel que pour tout $n \geq n_0$, $||f - B_n||_\infty \leq \epsilon$. Conclure.

7.3 Marche aléatoire

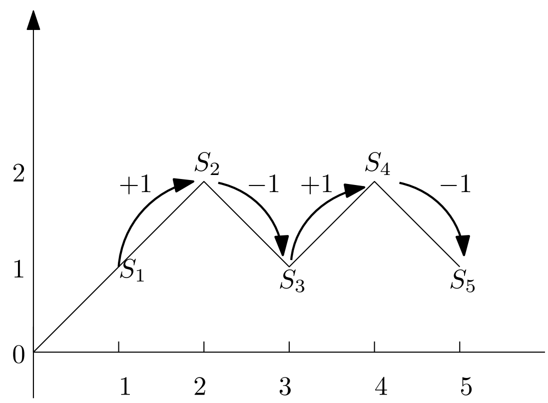
On se donne un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel est définie une suite $(X_n, n \in \mathbb{N}^\star)$ de variables aléatoires indépendantes identiquement distribuées. La loi de X_1 est donnée par

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

On définit

$$S_0 = 0 \text{ et } S_n = \textstyle\sum_{k=1}^n X_k.$$

La suite de variables aléatoires $(S_n, n \geq 1)$ est appelée *marche aléatoire symétrique.*



Plusieurs questions naturelles se posent, la marche aléatoire, retourne-t-elle à l'origine ? Si oui, combien de temps met-on en moyenne pour y revenir ?

L'objectif du problème est de répondre à ces deux questions. On définit

$$T = \begin{cases} \min\{n \geq 1; S_n = 0\} \text{ si } \{n \geq 1; S_n = 0\} \neq \emptyset \\ \\ +\infty \text{ sinon.} \end{cases}$$

La variable aléatoire T est le premier temps de retour à 0 de la marche aléatoire.

Le problème contient trois parties. La troisième partie est purement analytique. N'hésitez pas à admettre des questions d'analyse pour avancer.

Etude du temps de retour à l'origine :

1) On considère les événement E_n=« la marche aléatoire n'a pas touché 0 entre les temps 1 et n » et A_k=« La marche aléatoire visite 0 pour la dernière fois avant l'instant n, à l'instant k ». Justifier brièvement les égalités

$$E_n = (T > n) \cup (T = +\infty)$$

$$A_k = (S_k = 0) \cap \left(\bigcap_{i=k+1}^{n} (S_i \neq 0) \right)$$

2) Montrer que

$$\mathbb{P}(A_k) = \mathbb{P}(S_k = 0)\mathbb{P}_{(S_k=0)} \left(\bigcap_{i=k+1}^{n} (S_i \neq 0) \right).$$

On rappelle que la notation $\mathbb{P}_{(S_k=0)}(.)$ correspond à la probabilité conditionnelle sachant $S_k = 0$.

3) Soit k un entier fixé, on définit $S'_n = S_{n+k}$ pour tout $n \geq 0$. En remarquant que $(S'_n, n \geq 0)$ sachant $S'_0 = 0$ a même loi que $(S_n, n \geq 0)$, montrer que

$$\mathbb{P}(A_k) = \mathbb{P}(S_k = 0)\mathbb{P}\left(E_{n-k}\right).$$

4) Justifier que $\mathbb{P}(\bigcup_{k=0}^{n} A_k) = 1$ et montrer que

$$\sum_{k=0}^{n} \mathbb{P}(S_k = 0)\mathbb{P}\left(E_{n-k}\right) = 1.$$

5) Soit $x \in [0, 1[$. On considère les suites $(a_n, n \geq 0)$ et $(b_n, n \geq 0)$ définies par

$$a_n = \mathbb{P}[S_n = 0]x^n \text{ et } b_n = \mathbb{P}[E_n]x^n.$$

Justifier que les séries de terme général $(a_n, n \geq 0)$ et $(b_n, n \geq 0)$ sont convergentes. En utilisant le produit de Cauchy (voir l'encadré à la fin du sujet) montrer que

$$\frac{1}{1-x} = \left(\sum_{n=0}^{+\infty} \mathbb{P}[S_n = 0]x^n \right) \left(\sum_{n=0}^{\infty} \mathbb{P}[E_n]x^n \right).$$

6) Justifier que $\mathbb{P}[S_{2n+1} = 0] = 0$ pour tout n. On définit une suite de variables aléatoires indépendantes $(Y_i, i \geq 1)$ de la façon suivante : $Y_i = 1$ si $X_i = 1$ et 0 sinon. Justifier l'identité des événements

$$(S_{2n} = 0) = \left(\sum_{i=1}^{2n} Y_i = n \right).$$

Quelles sont les lois respectives de Y_i et $\sum_{i=1}^{2n} Y_i$? En déduire que

$$\mathbb{P}[S_{2n} = 0] = \frac{\binom{2n}{n}}{4^n}.$$

7) En admettant [1]

$$\sum_{n=0}^{\infty} \frac{\binom{2n}{n}}{4^n} x^{2n} = \frac{1}{\sqrt{1-x^2}},$$

montrer que

$$\sum_{n=0}^{\infty} \mathbb{P}[E_n] x^n = \sqrt{\frac{1+x}{1-x}}.$$

8) En remarquant que l'événement $(T = +\infty)$ est inclus dans E_n pour tout n, montrer que pour tout $x \in [0,1[$,

$$\mathbb{P}[T = +\infty] x^n \leq \mathbb{P}[E_n] x^n \text{ et en déduire que } \sum_{n=0}^{\infty} \mathbb{P}[T = +\infty] x^n \leq \sum_{n=0}^{\infty} \mathbb{P}[E_n] x^n.$$

Montrer que pour tout $x \in [0,1[$

$$\mathbb{P}[T = +\infty] \leq \sqrt{1-x^2}.$$

En faisant tendre x vers 1, montrer par encadrement que

$$\mathbb{P}[T = +\infty] = 0.$$

Qu'en déduisez vous pour la marche aléatoire, avec quelle probabilité retourne-t-elle à l'origine ?

Etude de l'espérance du temps de retour à l'origine :

9) Justifier que $\mathbb{P}[T > n] = \mathbb{P}[E_n]$ et en déduire que pour tout $x \in [0,1[$,

$$\sum_{n=0}^{\infty} \mathbb{P}[T > n] x^n = \sqrt{\frac{1+x}{1-x}}.$$

Montrer que

$$\sum_{n \geq 0} \mathbb{P}[T > n] = \infty.$$

On rappelle que si $\sum_{n \geq 1} a_n x^n$ est une série entière de rayon de convergence R alors

$$\lim_{x \to R^-} \sum_{n=1}^{\infty} a_n x^n = \sum_{n=1}^{\infty} a_n R^n \in \mathbb{R}_+ \cup \{+\infty\}.$$

10) Justifier que pour tout $k \in \mathbb{N}$,

$$\mathbb{P}[T = k] = \mathbb{P}[T > k-1] - \mathbb{P}[T > k].$$

En utilisant cette relation, établir que pour tout entier naturel $n \geq 2$,

$$\sum_{k=0}^{n} k\mathbb{P}[T = k] = \sum_{k=0}^{n-1} \mathbb{P}[T > k] - n\mathbb{P}[T > n].$$

1. On le démontre dans la dernière partie du sujet

11) On suppose que $\mathbb{E}[T] < +\infty$. Montrer que

$$n\mathbb{P}[T > n] = n \sum_{k=n+1}^{\infty} \mathbb{P}[T = k] \text{ et } n \sum_{k=n+1}^{\infty} \mathbb{P}[T = k] \leq \sum_{k=n+1}^{\infty} k\mathbb{P}[T = k].$$

En déduire que $n\mathbb{P}[T > n]$ tend vers 0 lorsque n tend vers l'infini, et que

$$\mathbb{E}[T] = \sum_{k=0}^{\infty} \mathbb{P}[T > k].$$

En particulier, si $\mathbb{E}[T] < +\infty$ alors $\sum_{k=0}^{\infty} \mathbb{P}[T > k] < \infty$.

12) Montrer à l'aide de la question précédente et de la question 9 que

$$\mathbb{E}[T] = +\infty.$$

La moyenne du temps de retour à l'origine est donc infinie. Cette propriété remarquable est liée au faite que si la marche aléatoire s'éloigne beaucoup de 0, il lui faut beaucoup de temps pour y revenir.

Développement en série entière de $\phi : x \to \frac{1}{\sqrt{1-x}}$.

La fonction $\phi : x \to \frac{1}{\sqrt{1-x}}$ est de classe C^∞ sur $[0, 1[$. Par la formule de Taylor avec reste intégral à l'ordre n, on a au point $x \in [0, 1[$:

$$\phi(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \phi^{(k)}(0) + \int_0^x \frac{(x-t)^n}{n!} \phi^{(n+1)}(t)dt.$$

13) On note $\phi^{(n)}$ la dérivée n-ième de ϕ, montrer par récurrence que

$$\phi^{(n)}(x) = \frac{(2n)!}{4^n n!}(1-x)^{-\frac{2n+1}{2}}.$$

14) Soit $t \in [0, x]$, démontrer que

$$\frac{(x-t)^n}{n!} \phi^{(n+1)}(t) = (n+1) \frac{\binom{2n+2}{n+1}}{4^{n+1}} \left(\frac{x-t}{1-t} \right)^n \times \frac{1}{(1-t)^{3/2}}$$

15) On a [2] l'inégalité suivante : $\frac{\binom{2n+2}{n+1}}{4^{n+1}} \leq 1$. En étudiant la fonction $t \mapsto x - \frac{x-t}{1-t}$, montrer que pour tout $t \in [0, x]$, $0 \leq \frac{x-t}{1-t} \leq x$. Montrer à partir de ces inégalités que

$$0 \leq \frac{(x-t)^n}{n!} \phi^{(n+1)}(t) \leq \frac{(n+1)x^n}{(1-x)^{3/2}},$$

en déduire

$$\int_0^x \frac{(x-t)^n}{n!} \phi^{(n+1)}(t)dt \underset{n\to+\infty}{\longrightarrow} 0.$$

16) Etablir l'égalité préalablement admise

$$\frac{1}{\sqrt{1-x^2}} = \sum_{n=1}^{\infty} \frac{\binom{2n}{n}}{4^n} x^{2n}.$$

2. $\frac{\binom{2n+2}{n+1}}{4^{n+1}} = \mathbb{P}[S_{2n+2} = 0] \leq 1$

Rappel sur le produit de Cauchy :

si $(a_n, n \geq 1)$ et $(b_n, n \geq 1)$ sont deux suites positives et forment les termes généraux de séries convergentes, alors la série de terme général

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

converge et sa somme vaut

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

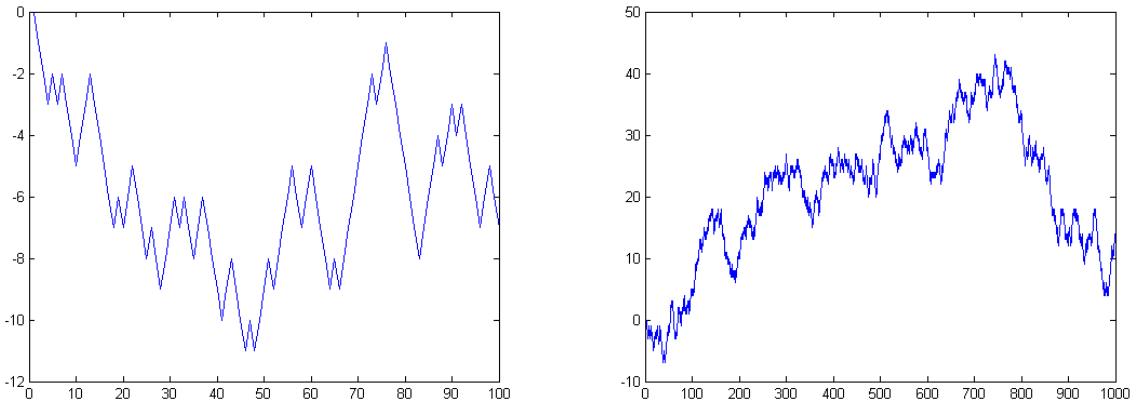


FIGURE 7.2 – Réalisation d'une marche aléatoire symétrique, n=100 et n=1000

Les marches aléatoires jouent un rôle important en théorie des probabilités. L'échelle de temps dans la simulation à droite est dix fois plus petite que celle de gauche. Cette courbe approche une fonction aléatoire célèbre : le mouvement Brownien. Cette fonction aléatoire est avec probabilité 1, continue, non-dérivable, non monotone par morceaux. Ces fonctions très irrégulières sont typiques en théorie des probabilités (plus précisément en théorie des processus stochastiques.) Il n'est pas facile de construire des fonctions déterministes avec de telles propriétés. Le mouvement Brownien est enseigné en M2. Il est impliqué dans de nombreux modèles probabilistes pour la physique, la biologie ou la finance.

Chapitre 8

Problèmes d'analyse pour les probabilités

Intégrale de Gauss

Soit
$$f : x \in \mathbb{R} \mapsto e^{-x^2/2}.$$

1) Montrer que f est intégrable sur $\mathbb{R}$.
2) Soit
$$G := \int_0^\infty f(x)\mathrm{d}x.$$
   Montrer que
$$\int_{-\infty}^\infty f(x)\mathrm{d}x = 2G.$$
3) Soit
$$H := \int_{\mathbb{R}_+ \times \mathbb{R}_+} e^{-(x^2+y^2)}\mathrm{d}x\mathrm{d}y.$$
   Montrer que $H = G^2$.
4) Rappeler brièvement la définition des coordonnées polaires. En passant aux coordonnées polaires, montrer que
$$H = \int_{\mathbb{R}_+ \times [0, \frac{\pi}{2}]} e^{-r^2} r\mathrm{d}r\mathrm{d}\theta.$$
   En déduire la valeur de H puis de G.

Intégrales de Wallis et la formule de Stirling

La formule de Stirling est l'équivalent suivant :
$$n! \underset{n\to\infty}{\sim} \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Afin de l'établir, on introduit les intégrales de Wallis : pour $n \in \mathbb{N}$,
$$I_n = \int_0^{\pi/2} \sin^n(t)\mathrm{d}t.$$

1) Montrer que $(I_n)_{n\in\mathbb{N}}$ est décroissante et convergente.
2) Montrer que pour tout $n \in \mathbb{N}$ tel que $n \geq 2$,

$$I_n = \frac{n-1}{n-2} I_{n-2}.$$

3) Calculer I_{2n} et I_{2n+1}.
4) Démontrer que $\frac{I_{2n}}{I_{2n+1}} \underset{n\to\infty}{\longrightarrow} 1$.
5) En déduire que

$$\left[\frac{1.3.5...(2n-1)}{2.4.6...2n}\right]^2 n \underset{n\to\infty}{\longrightarrow} \frac{1}{\pi}.$$

6) En remarquant que $2.4.6...2n = 2^n n!$, montrer que

$$\frac{\sqrt{n}(2n)!}{2^{2n}(n!)^2} \underset{n\to\infty}{\longrightarrow} \frac{1}{\sqrt{\pi}},$$

7) Soit f une fonction $C^2([0,1],\mathbb{R})$. Montrer qu'il existe $\xi \in [0,1]$ tel que :

$$\int_0^1 f(x)dx = \frac{1}{2}\left(f(0) + f(1)\right) - \frac{1}{12} f''(\xi).$$

Indication : introduire $\phi(x) = \frac{1}{2}x(1-x)$, en particulier remarquer que $\int_0^1 f(x)dx = -\int_0^1 \phi''(x) f(x)$.

8) En appliquant la formule précédente, montrer que pour $1 \leq k \leq n-1$:

$$\int_k^{k+1} \ln(x)\mathrm{d}x = \frac{1}{2}(\ln(k) + \ln(k+1)) + \frac{1}{12}\xi_k^{-2}$$

avec $k \leq \xi_k \leq k+1$.
9) Montrer que

$$\int_1^n \ln(x)\mathrm{d}x = \sum_{k=1}^n \ln(k) - \frac{1}{2}\ln(n) + \frac{1}{12}\sum_{k=1}^n \frac{1}{\xi_k^2}.$$

10) En déduire que

$$\ln(n!) = \left(n + \frac{1}{2}\right)\ln(n) - n + 1 - \frac{1}{12}\sum_{k=1}^n \frac{1}{\xi_k^2}.$$

11) On pose $\gamma_n = 1 - \frac{1}{12}\sum_{k=1}^n \frac{1}{\xi_k^2}$, justifier que γ_n est une suite convergente et que

$$n! = n^{n+\frac{1}{2}} e^{-n} e^{\gamma_n}.$$

12) On pose $c_n = e^{\gamma_n}$, on cherche la limite c de cette suite. Montrer en calculant $\frac{c_n^2}{c_{2n}}$ que

$$c = \lim_{n\to\infty} \sqrt{2}\frac{2^{2n}(n!)^2}{\sqrt{n}(2n)!}$$

13) En utilisant la question 6) montrer que $c = \sqrt{2\pi}$, en déduire la formule de Stirling.

Pour Z de loi $N(0,1)$,

$$\Phi(t) = F_Z(t) = P(Z \leq t) = \int_{-\infty}^{t} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

t	$0,00$	$0,01$	$0,02$	$0,03$	$0,04$	$0,05$	$0,06$	$0,07$	$0,08$	$0,09$
$0,0$	$0,5000$	$0,5040$	$0,5080$	$0,5120$	$0,5160$	$0,5199$	$0,5239$	$0,5279$	$0,5319$	$0,5359$
$0,1$	$0,5398$	$0,5438$	$0,5478$	$0,5517$	$0,5557$	$0,5596$	$0,5636$	$0,5675$	$0,5714$	$0,5753$
$0,2$	$0,5793$	$0,5832$	$0,5871$	$0,5910$	$0,5948$	$0,5987$	$0,6026$	$0,6064$	$0,6103$	$0,6141$
$0,3$	$0,6179$	$0,6217$	$0,6255$	$0,6293$	$0,6331$	$0,6368$	$0,6406$	$0,6443$	$0,6480$	$0,6517$
$0,4$	$0,6554$	$0,6591$	$0,6628$	$0,6664$	$0,6700$	$0,6736$	$0,6772$	$0,6808$	$0,6844$	$0,6879$
$0,5$	$0,6915$	$0,6950$	$0,6985$	$0,7019$	$0,7054$	$0,7088$	$0,7123$	$0,7157$	$0,7190$	$0,7224$
$0,6$	$0,7257$	$0,7291$	$0,7324$	$0,7357$	$0,7389$	$0,7422$	$0,7454$	$0,7486$	$0,7517$	$0,7549$
$0,7$	$0,7580$	$0,7611$	$0,7642$	$0,7673$	$0,7704$	$0,7734$	$0,7764$	$0,7794$	$0,7823$	$0,7852$
$0,8$	$0,7881$	$0,7910$	$0,7939$	$0,7967$	$0,7995$	$0,8023$	$0,8051$	$0,8078$	$0,8106$	$0,8133$
$0,9$	$0,8159$	$0,8186$	$0,8212$	$0,8238$	$0,8264$	$0,8289$	$0,8315$	$0,8340$	$0,8365$	$0,8389$
$1,0$	$0,8413$	$0,8438$	$0,8461$	$0,8485$	$0,8508$	$0,8531$	$0,8554$	$0,8577$	$0,8599$	$0,8621$
$1,1$	$0,8643$	$0,8665$	$0,8686$	$0,8708$	$0,8729$	$0,8749$	$0,8770$	$0,8790$	$0,8810$	$0,8830$
$1,2$	$0,8849$	$0,8869$	$0,8888$	$0,8907$	$0,8925$	$0,8944$	$0,8962$	$0,8980$	$0,8997$	$0,9015$
$1,3$	$0,9032$	$0,9049$	$0,9066$	$0,9082$	$0,9099$	$0,9115$	$0,9131$	$0,9147$	$0,9162$	$0,9177$
$1,4$	$0,9192$	$0,9207$	$0,9222$	$0,9236$	$0,9251$	$0,9265$	$0,9279$	$0,9292$	$0,9306$	$0,9319$
$1,5$	$0,9332$	$0,9345$	$0,9357$	$0,9370$	$0,9382$	$0,9394$	$0,9406$	$0,9418$	$0,9429$	$0,9441$
$1,6$	$0,9452$	$0,9463$	$0,9474$	$0,9484$	$0,9495$	$0,9505$	$0,9515$	$0,9525$	$0,9535$	$0,9545$
$1,7$	$0,9554$	$0,9564$	$0,9573$	$0,9582$	$0,9591$	$0,9599$	$0,9608$	$0,9616$	$0,9625$	$0,9633$
$1,8$	$0,9641$	$0,9649$	$0,9656$	$0,9664$	$0,9671$	$0,9678$	$0,9686$	$0,9693$	$0,9699$	$0,9706$
$1,9$	$0,9713$	$0,9719$	$0,9726$	$0,9732$	$0,9738$	$0,9744$	$0,9750$	$0,9756$	$0,9761$	$0,9767$
$2,0$	$0,9772$	$0,9778$	$0,9783$	$0,9788$	$0,9793$	$0,9798$	$0,9803$	$0,9808$	$0,9812$	$0,9817$
$2,1$	$0,9821$	$0,9826$	$0,9830$	$0,9834$	$0,9838$	$0,9842$	$0,9846$	$0,9850$	$0,9854$	$0,9857$
$2,2$	$0,9861$	$0,9864$	$0,9868$	$0,9871$	$0,9875$	$0,9878$	$0,9881$	$0,9884$	$0,9887$	$0,9890$
$2,3$	$0,9893$	$0,9896$	$0,9898$	$0,9901$	$0,9904$	$0,9906$	$0,9909$	$0,9911$	$0,9913$	$0,9916$
$2,4$	$0,9918$	$0,9920$	$0,9922$	$0,9925$	$0,9927$	$0,9929$	$0,9931$	$0,9932$	$0,9934$	$0,9936$
$2,5$	$0,9938$	$0,9940$	$0,9941$	$0,9943$	$0,9945$	$0,9946$	$0,9948$	$0,9949$	$0,9951$	$0,9952$
$2,6$	$0,9953$	$0,9955$	$0,9956$	$0,9957$	$0,9959$	$0,9960$	$0,9961$	$0,9962$	$0,9963$	$0,9964$
$2,7$	$0,9965$	$0,9966$	$0,9967$	$0,9968$	$0,9969$	$0,9970$	$0,9971$	$0,9972$	$0,9973$	$0,9974$
$2,8$	$0,9974$	$0,9975$	$0,9976$	$0,9977$	$0,9977$	$0,9978$	$0,9979$	$0,9979$	$0,9980$	$0,9981$
$2,9$	$0,9981$	$0,9982$	$0,9982$	$0,9983$	$0,9984$	$0,9984$	$0,9985$	$0,9985$	$0,9986$	$0,9986$
$3,0$	$0,9987$	$0,9987$	$0,9987$	$0,9988$	$0,9988$	$0,9989$	$0,9989$	$0,9989$	$0,9990$	$0,9990$
$3,1$	$0,9990$	$0,9991$	$0,9991$	$0,9991$	$0,9992$	$0,9992$	$0,9992$	$0,9992$	$0,9993$	$0,9993$
$3,2$	$0,9993$	$0,9993$	$0,9994$	$0,9994$	$0,9994$	$0,9994$	$0,9994$	$0,9995$	$0,9995$	$0,9995$
$3,3$	$0,9995$	$0,9995$	$0,9995$	$0,9996$	$0,9996$	$0,9996$	$0,9996$	$0,9996$	$0,9996$	$0,9997$
$3,4$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9997$	$0,9998$
$3,5$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$	$0,9998$
$3,6$	$0,9998$	$0,9998$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$
$3,7$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$
$3,8$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$	$0,9999$
$3,9$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$	$1,0000$

Exemples : $\Phi(0,25) \simeq 0,5987$, $\quad \Phi(-0,32) = 1 - \Phi(0,32) \simeq 1 - 0,6255 = 0,3745$