

## TD1 — Théorèmes-limites et transformées de lois

**Exercice 1.1.** *Marche aléatoire.* Soit  $(X_k)_{k \geq 1}$  une suite de variables indépendantes et de même loi  $\frac{1}{2}(\delta_{-1} + \delta_1)$ . Pour tout  $n \geq 0$ , soit  $S_n = X_1 + \dots + X_n$ . Montrer que, pour tout  $\varepsilon > 0$ ,  $\mathbb{P}(S_n > \varepsilon n) \rightarrow 0$  quand  $n \rightarrow \infty$ . Qu'en est-il de  $\mathbb{P}(S_n > 0)$  ?

**Exercice 1.2.** Soit  $(U_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes et de même loi uniforme sur la boule euclidienne unité de  $\mathbb{R}^2$ , i.e. sur  $B(0, 1) = \{(x, y) \in \mathbb{R}^2 ; x^2 + y^2 \leq 1\}$ . On définit par récurrence  $X_0 = (1, 0)$  et  $X_{n+1} = \|X_n\|_2 U_{n+1}$ . Montrer que  $(\frac{1}{n} \log \|X_n\|_2)_{n \geq 1}$  converge presque sûrement et préciser sa limite.

**Exercice 1.3.** Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires réelles indépendantes et de même loi  $\mu$ . Pour tout  $n \geq 1$ , on définit la mesure empirique  $L_n = \frac{1}{n} \sum_{k=1}^n \delta_{X_k}$ . Montrer que, presque sûrement,  $(L_n)_{n \geq 1}$  converge étroitement vers  $\mu$ .

**Exercice 1.4.** Soit  $N$  une variable aléatoire à valeurs dans  $\mathbb{N}$  et  $(X_n)_{n \geq 1}$  une suite de variables aléatoires i.i.d. à valeurs dans  $\mathbb{N}$  et indépendante de  $N$ . Soit  $Z = \sum_{n=1}^N X_n$ . Calculer la fonction génératrice des probabilités de  $Z$  en fonction de celles de  $N$  et de  $X_1$ . Déterminer l'espérance de  $Z$ . **Application** — Dans le cas où  $N$  suit la loi de Poisson de paramètre  $\lambda \in ]0, +\infty[$  et  $X_1$  suit la loi de Bernoulli de paramètre  $p \in ]0, 1[$ , préciser la loi de  $Z$ .

**Exercice 1.5.** *Inégalité de Hoeffding.*

1) Soit  $(X_n)_{n \geq 1}$  une suite de variable indépendantes et de même loi donnée par  $\mathbb{P}(X_1 = 1) = 1/2 = \mathbb{P}(X_1 = -1)$ . Pour tout  $n \geq 0$ , soit  $S_n = X_1 + X_2 + \dots + X_n$ . Montrer que, pour tout  $x \geq 0$  et pour tout  $n \geq 0$ ,  $\mathbb{P}(S_n \geq x\sqrt{n}) \leq e^{-x^2/2}$ . [Indication : on pourra appliquer l'inégalité de Markov pour majorer la probabilité  $\mathbb{P}(e^{\lambda S_n} \geq e^{\lambda x \sqrt{n}})$  avec  $\lambda > 0$ .]

2) Comparer le résultat précédent avec celui donné par le théorème-limite central.

3) Soit  $X$  une variable aléatoire réelle centrée et à valeurs dans  $[a, b]$ . Montrer que, pour tout  $\lambda \in \mathbb{R}$ ,

$$\log \mathbb{E}[e^{\lambda X}] \leq \frac{(b-a)^2 \lambda^2}{8}.$$

4) Soit  $(X_1, \dots, X_n)$  une famille de variables aléatoires réelles indépendantes. Pour tout  $k \in \llbracket 1, n \rrbracket$ , on suppose que  $X_k$  est à valeurs dans  $[a_k, b_k]$ . Soit aussi  $S_n = X_1 + \dots + X_n$ . Montrer que, pour tout  $\varepsilon > 0$ ,

$$\mathbb{P}(S_n - \mathbb{E}[S_n] \geq \varepsilon) \leq \exp\left(-\frac{2\varepsilon^2}{\sum_{k=1}^n (b_k - a_k)^2}\right).$$

*Remarque.* La fonction  $\Lambda_X : \mathbb{R} \rightarrow ]-\infty, +\infty]$  donnée par  $\Lambda_X(\lambda) = \log \mathbb{E}[e^{\lambda X}]$  est appelée *fonction génératrice des cumulants* de  $X$ .

## Pour s'entraîner

**Exercice 1.6.** Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes et de même loi  $\mathcal{E}(1)$ .

1) Pour tout  $n \geq 1$ , expliciter la loi de  $S_n = X_1 + \dots + X_n$ .

2) En utilisant la méthode de Laplace, déterminer un équivalent de  $\mathbb{P}(S_n \geq n)$ .

3) En utilisant le théorème-limite central, retrouver la formule de Stirling.

**Exercice 1.7.** Pour tout  $n \geq 1$ , soit  $(m_n, \sigma_n) \in \mathbb{R} \times \mathbb{R}_+$  et  $X_n \sim \mathcal{N}(m_n, \sigma_n^2)$ . Montrer que  $(X_n)_{n \geq 1}$  converge en loi si et seulement si  $((m_n, \sigma_n))_{n \geq 1}$  converge et identifier la limite.

## Pour aller plus loin

**Exercice 1.8.** Soit  $p \in [0, 1]$ . Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires indépendantes et de même loi  $(1-p)\delta_{-1} + p\delta_1$ . Pour tout  $n \geq 0$ , soit  $S_n = X_1 + X_2 + \dots + X_n$ . On note  $T_0 = \inf \{n \geq 1 ; S_n = 0\} \in \mathbb{N} \cup \{+\infty\}$  le temps de retour en 0. Pour tout  $n \geq 0$ , on note  $u(n) = \mathbb{P}(S_n = 0)$  et  $v(n) = \mathbb{P}(T_0 = n)$  et on définit, pour tout  $s \in [0, 1[$ ,

$$U(s) = \sum_{n=0}^{\infty} u(n)s^n \quad \text{et} \quad V(s) = \sum_{n=1}^{\infty} v(n)s^n.$$

- 1) Montrer que, pour tout  $1 \leq m \leq n$ ,  $\mathbb{P}(S_n = 0 \mid T_0 = m) = u(n-m)$ .
- 2) Montrer que, pour tout  $n \geq 1$ ,

$$u(n) = \sum_{m=0}^n u(n-m)v(m),$$

puis que  $U = 1 + UV$ .

- 3) Montrer que, pour tout  $k \geq 0$ ,  $u(2k) = \binom{2k}{k}(p(1-p))^k$  et  $u(2k+1) = 0$ . En déduire que, pour tout  $s \in [0, 1[$ ,

$$U(s) = \frac{1}{\sqrt{1-4p(1-p)s^2}} \quad \text{et} \quad V(s) = 1 - \sqrt{1-4p(1-p)s^2}.$$

- 4) En déduire que  $\mathbb{P}(T_0 < \infty) = 1 - |2p-1|$ .
- 5) Montrer que, si  $p = 1/2$ ,  $\mathbb{E}[T_0] = +\infty$ .

## TD2 — Chaînes de Markov — Définition

**Exercice 2.1.** *Tribu cylindrique.* Soit  $E$  un ensemble,  $\mathcal{B}$  une tribu sur  $E$  contenant les singletons,  $y \in E$ ,  $f: E \rightarrow \mathbb{R}$  ( $\mathcal{B}, \mathcal{B}(\mathbb{R})$ )-mesurable, et  $a \in \mathbb{R}$ .

- 1) Montrer que les parties de  $E^{\mathbb{N}}$  suivantes sont dans la tribu cylindrique  $\mathcal{B}^{\otimes \mathbb{N}}$  :
  - 1a)  $A = \{(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}} ; \exists k \in \mathbb{N} \ x_k = y\}$  ;
  - 1b)  $B = \{(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}} ; \text{i.s. } x_k = y\}$  ;
  - 1c)  $C = \{(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}} ; \sup_{k \in \mathbb{N}} f(x_k) = a\}$  ;
  - 1d)  $D = \{(x_k)_{k \in \mathbb{N}} \in E^{\mathbb{N}} ; \lim_{k \rightarrow \infty} f(x_k) = a\}$ .
- 2) Montrer que l'application  $H^y: E^{\mathbb{N}} \rightarrow \overline{\mathbb{N}}$  définie pour  $x = (x_k)_{k \in \mathbb{N}}$  par  $H^y(x) = \inf \{k \in \mathbb{N} ; x_k = y\}$  est  $(\mathcal{B}^{\otimes \mathbb{N}}, \mathfrak{P}(\overline{\mathbb{N}}))$ -mesurable.
- 3) Les résultats subsistent-ils si l'on remplace  $\mathbb{N}$  par  $\mathbb{R}$  (i.e.  $E^{\mathbb{N}}$  par  $E^{\mathbb{R}}$ ) ?

**Exercice 2.2.** On dispose de deux pièces : l'une est normale et équilibrée, l'autre présente "face" des deux côtés. On choisit l'une des deux pièces au hasard, puis on la lance une infinité de fois. On modélise la suite des résultats des lancers par un processus  $(X_n)_{n \geq 1}$  à valeurs dans  $\{P, F\}$ .

- 1) On a obtenu pile au  $n^{\text{e}}$  lancer. Quelle est la probabilité d'obtenir face au  $(n+1)^{\text{e}}$  lancer ?
- 2) On a obtenu face au  $n^{\text{e}}$  lancer. Quelle est la probabilité d'obtenir face au  $(n+1)^{\text{e}}$  lancer ?
- 3) On a obtenu pile au  $(n-1)^{\text{e}}$  lancer et face au  $n^{\text{e}}$  lancer. Quelle est la probabilité d'obtenir face au  $(n+1)^{\text{e}}$  lancer ?
- 4) Le processus  $(X_n)_{n \geq 1}$  est-il une chaîne de Markov ?

**Exercice 2.3.** Soit  $(X_n)_{n \geq 0}$  la marche simple sur  $\mathbb{Z}$ , avec probabilité  $p \in [0, 1]$  de se déplacer à droite. Parmi les fonctions  $f$  suivantes, pour lesquelles  $(f(X_n))_{n \geq 0}$  est-elle une chaîne de Markov ? Le cas échéant, on en précisera le noyau de transition.

- 1)  $f: \mathbb{Z} \rightarrow \mathbb{Z}$ ,  $f(x) = 2x - 1$ .
- 2)  $f: \mathbb{Z} \rightarrow \mathbb{N}$ ,  $f(x) = |x|$ .
- 3)  $f: \mathbb{Z} \rightarrow \mathbb{Z}/N\mathbb{Z}$ ,  $f(x) = x \bmod N$ .

### Pour s'entraîner

**Exercice 2.4.** Soit  $p \in ]0, 1[$  et  $(X_n)_{n \geq 0} \sim \mathcal{B}(p)^{\otimes \mathbb{N}}$ . Dans chacun des cas suivants, dire si la suite est une chaîne de Markov et, le cas échéant, préciser son espace d'états, sa loi initiale et son noyau de transition :

- 1)  $(X_n)_{n \geq 0}$  ;
- 2)  $(S_n)_{n \geq 0}$ , où  $S_n = X_1 + \dots + X_n$  ;
- 3)  $(T_n)_{n \geq 0}$ , où  $T_n = S_1 + \dots + S_n$  ;
- 4)  $((T_n, S_n))_{n \geq 0}$ .

**Exercice 2.5.** Montrer que la distance à la racine de la marche sur l'arbre  $r$ -aire enraciné est une chaîne de Markov dont on précisera le noyau de transition.

### Pour aller plus loin

**Exercice 2.6.** Soit  $E$  un ensemble dénombrable, muni de la tribu  $\mathfrak{P}(E)$  de toutes les parties de  $E$ . On définit, pour  $x = (x_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$  et  $y = (y_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ ,  $d(x, y) = \exp(-\inf \{n \in \mathbb{N} ; x_n \neq y_n\})$ .

- 1) Montrer que  $d$  est une distance sur  $E^{\mathbb{N}}$ .
- 2) Montrer que l'ensemble  $D \subset E^{\mathbb{N}}$  des suites stationnaires (i.e. constantes à partir d'un certain rang) est dénombrable et dense dans  $(E^{\mathbb{N}}, d)$ .
- 3) Montrer que la tribu cylindrique  $\mathfrak{P}(E)^{\otimes \mathbb{N}}$  est égale à la tribu borélienne  $\mathcal{B}(E^{\mathbb{N}}, d)$ .

## TD3 — Chaînes de Markov — Loi au temps $n$

**Exercice 3.1.** *Modèle d'Ehrenfest (élémentaire).* On considère deux urnes et on place deux boules dans la première urne. À chaque instant  $n \geq 0$ , on tire une boule au hasard et on la change d'urne.

- 1) Modéliser le nombre de boules dans la première urne par une chaîne de Markov  $(X_n)_{n \geq 0}$ , dont on précisera l'espace d'états, le graphe, la matrice de transition et l'état initial.
- 2) Pour tout  $n \geq 0$ , déterminer la loi  $\mu_n$  de  $X_n$ .
- 3) Calculer l'espérance du temps de premier passage (ou retour) en 2.
- 4) Montrer que  $(\mu_n)_{n \geq 0}$  ne converge pas, mais que  $(\frac{1}{n+1}(\mu_0 + \mu_1 + \dots + \mu_n))_{n \geq 0}$  converge et préciser sa limite.

**Exercice 3.2.** *La météo en Syldavie.* Le temps évolue de façon assez simple en Syldavie : s'il fait beau, alors il fera encore beau le lendemain avec une probabilité de  $5/6$  ; s'il pleut, alors il pleuvra le lendemain avec une probabilité de  $2/3$ . C'est Noël et il fait beau. Quelle est la probabilité qu'il fasse beau à Noël prochain ?

**Exercice 3.3.** *Marche sur l'hypercube.* L'objectif de l'exercice est de calculer explicitement la loi au temps  $n$  de la marche sur l'hypercube  $E = (\mathbb{Z}/2\mathbb{Z})^d$ , dont le noyau de transition est donné par  $Q(x, y) = \frac{1}{d} \mathbb{1}_{\exists i \ y=x+e^i}$ , où  $e^i = (0, \dots, 0, 1, 0, \dots, 0)$ , le 1 étant en  $i^e$  position. On considère l'espace vectoriel  $V$  des fonctions de  $E$  dans  $\mathbb{R}$ , que l'on munit du produit scalaire

$$\langle f | g \rangle := \frac{1}{2^d} \sum_{x \in E} f(x)g(x).$$

Pour tout  $u \in E$ , on définit  $\varepsilon_u(x) = (-1)^{\sum_{i=1}^d u_i x_i}$ .

- 1) Montrer que  $(\varepsilon_u)_{u \in E}$  est une base orthonormée de  $V$ .
- 2) Pour  $p = \frac{1}{d} \sum_{i=1}^d \delta_{e^i}$  et  $u \in E$ , calculer  $\langle p | \varepsilon_u \rangle$ .
- 3) En remarquant que  $Q(x, y) = p(y - x)$ , montrer que

$$Q(x, y) = \frac{1}{2^d} \sum_{u \in E} \left(1 - \frac{2|u|}{d}\right) \varepsilon_u(y - x),$$

où  $|u| = |\{i \in \llbracket 1, d \rrbracket ; u_i = 1\}|$ .

- 4) Montrer que, pour tout  $(x, u, v) \in E^3$ ,

$$\frac{1}{2^d} \sum_{y+z=x} \varepsilon_u(y) \varepsilon_v(z) = \mathbb{1}_{u=v} \varepsilon_u(x).$$

- 5) En déduire que, pour tout  $n \geq 1$ ,

$$Q^n(x, y) = \frac{1}{2^d} \sum_{u \in E} \left(1 - \frac{2|u|}{d}\right)^n \varepsilon_u(y - x).$$

- 6) En déduire que, pour toute distribution initiale, la loi au temps  $2n$  de la marche sur l'hypercube converge vers une limite que l'on précisera.

## Pour s'entraîner

**Exercice 3.4.** *Automate.* Un automate lit une suite de 0 et de 1, par exemple

100110111101...

de la façon suivante : partant de la gauche, il compte le nombre de 1 successifs obtenus, en saturant à 2. Il obtient ici

100120121201...

On suppose désormais que la suite donnée à l'automate est une suite de variables aléatoires i.i.d. de loi  $\mathcal{B}(1/2)$ .

- 1) Modéliser la suite renvoyée par l'automate par une chaîne de Markov  $(X_n)_{n \geq 0}$  dont on précisera l'espace d'états, le graphe, la matrice de transition et la loi initiale  $\mu$ .
- 2) Montrer que  $(\mathcal{L}_\mu(X_n))_{n \geq 0}$  converge et préciser sa limite  $\pi$ .
- 3) Retrouver  $\pi$  par un argument élémentaire.

## Pour aller plus loin

**Exercice 3.5.** *Marche sur le graphe complet.* Soit  $k \geq 2$  et  $G = (V, E)$  le graphe complet à  $k$  sommets, autrement dit le graphe à  $k$  sommets dans lequel toute paire de sommets est reliée par une arête. La marche aléatoire simple sur  $G$  est la chaîne de Markov d'espace d'états  $V$  dont la transition est définie par : si l'on est en un sommet  $v \in V$ , on choisit uniformément au hasard un autre sommet  $v' \in V \setminus \{v\}$  et on s'y déplace. On notera  $Q$  la matrice de transition et  $(X_n)_{n \geq 0}$  le processus.

- 1) Donner l'expression de  $Q$ .
- 2) Pour tout  $n \geq 0$ , calculer  $Q^n$ . [Indication : on pourra écrire  $Q = \frac{1}{k-1}(J_k - I_k)$ .]
- 3) Pour toute loi  $\mu$  sur  $V$ , étudier la convergence de  $(\mathcal{L}_\mu(X_n))_{n \geq 0}$ .
- 4) Si  $v \in V$ , déterminer la loi, sous  $\mathbb{P}_v$ , du temps de premier passage (ou retour) en  $v$ .

## TD4 — Chaînes de Markov — Classes de communication et analyse à un pas

**Exercice 4.1.** *Variables aléatoires i.i.d.* Soit  $E$  un ensemble dénombrable. Soit  $(X_n)_{n \geq 0}$  une suite de variables aléatoires i.i.d. sur  $E$ . Donner les classes de communication de la chaîne de Markov  $(X_n)_{n \geq 0}$  et préciser leurs natures (fermée ou non).

**Exercice 4.2.** Soit  $E = \llbracket 1, 3 \rrbracket$  et  $\mu_0$  une probabilité sur  $E$ . On considère une chaîne de Markov  $(X_n)_{n \geq 0}$  sur  $E$  de loi initiale  $\mu_0$  et de transition

$$Q = \begin{pmatrix} 3/4 & 0 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/2 & 0 & 1/2 \end{pmatrix}.$$

- 1) Donner les classes de communication de la transition  $Q$  et préciser leurs natures (fermée ou non).
- 2) Montrer que  $(\mathcal{L}(X_n))_{n \geq 0}$  converge. On notera  $\pi$  sa limite.
- 3) Vérifier que  $\pi$  est l'unique probabilité telle que  $\pi Q = \pi$ .

**Exercice 4.3.** Abel dit à Caïn : « Nous allons lancer une pièce jusqu'à ce que les mots PPP ou PFP apparaissent. Si c'est PPP, c'est toi qui gagne, si c'est PFP, c'est moi. Ce jeu est honnête, puisque chaque mot a la probabilité  $1/8$  d'apparaître. » Qu'en pensez-vous ?

- 1) Modéliser le jeu par une chaîne de Markov dont on précisera l'espace d'états, le graphe, la matrice de transition et la distribution initiale.
- 2) Calculer la probabilité qu'Abel gagne.
- 3) Quelle est la durée moyenne du jeu ?

**Exercice 4.4.** *Loi du temps de premier passage.* Sur  $E = \llbracket 1, 3 \rrbracket$ , on considère la transition donnée par

$$Q = \begin{pmatrix} 1 & 0 & 0 \\ 1/2 & 1/6 & 1/3 \\ 1/3 & 3/5 & 1/15 \end{pmatrix}.$$

Soit  $(X_n)_{n \geq 0}$  une suite de variables aléatoires à valeurs dans  $E$  telle que, sous  $\mathbb{P}_\mu$ ,  $(X_n)_{n \geq 0}$  soit une chaîne de Markov de transition  $Q$  et de loi initiale  $\mu$ . Soit  $T_3 = \inf \{n \geq 1 ; X_n = 3\} \in \mathbb{N}^* \cup \{\infty\}$ .

- 1) Pour tout  $n \geq 1$ , on note

$$F_n = \begin{pmatrix} \mathbb{P}_1(T_3 = n) \\ \mathbb{P}_2(T_3 = n) \\ \mathbb{P}_3(T_3 = n) \end{pmatrix}.$$

Montrer que  $(F_n)_{n \geq 1}$  satisfait une relation de récurrence de la forme  $F_{n+1} = AF_n$  où  $A$  est une matrice à préciser.

- 2) Pour tout  $n \geq 1$ , calculer  $F_n$ .
- 3) En déduire, pour tout  $k \in E$ ,  $\mathcal{L}_k(T_3)$ .

### Pour s'entraîner

**Exercice 4.5.** *Ruine du joueur.* Pour  $a \in \mathbb{Z}$ , on considère la marche aléatoire simple symétrique sur  $\mathbb{Z}$  issue de  $a$  : pour tout  $n \geq 0$ ,  $S_n = a + X_1 + \dots + X_n$  où les  $X_i$  sont i.i.d. de loi  $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ .

- 1) Montrer que  $(S_n)_{n \geq 0}$  est une chaîne de Markov, dont on précisera l'espace d'états, le graphe, la matrice de transition et la distribution initiale.
- 2) Soit  $N \in \mathbb{N}$ . Pour  $0 \leq a \leq N$ , calculer la probabilité d'atteindre d'abord  $N$  avant 0.
- 3) Quelle est le temps moyen d'atteinte de  $\{0, N\}$  ?

### Pour aller plus loin

**Exercice 4.6.** Montrer qu'une transition sur un ensemble fini admet toujours une classe de communication fermée.

## TD5 — Chaînes de Markov — Récurrence et transience

**Exercice 5.1.** *Château de cartes.* Soit  $(p_k)_{k \geq 0}$  une suite de réels de  $[0, 1]$ . Soit  $(X_n)_{n \geq 0}$  une chaîne de Markov sur  $\mathbb{N}$  de transition donnée par : pour tout  $k \geq 0$ ,  $Q(k, k+1) = p_k$  et  $Q(k, 0) = 1 - p_k$ .

- 1) Dessiner le graphe de la transition.
- 2) Donner une condition nécessaire et suffisante pour que la transition soit irréductible.
- 3) En notant  $T_0 = \inf \{n \geq 1 ; X_n = 0\}$ , calculer  $\mathbb{P}_0(T_0 = \infty)$ .
- 4) Soit  $p \in ]0, 1[$ . On suppose que, pour tout  $k \in \mathbb{N}$ ,  $p_k = p$ . L'état 0 est-il récurrent ou transient ?
- 5) On suppose que, pour tout  $k \in \mathbb{N}$ ,  $p_k = 1 - 2^{-k-1}$ . L'état 0 est-il récurrent ou transient ?

**Exercice 5.2.** *Marche simple symétrique sur  $\mathbb{Z}^d$ .* Soit  $d \geq 1$ . On considère, sur le groupe  $\mathbb{Z}^d$ , la marche aléatoire  $(X_n)_{n \geq 0}$  engendrée par  $\pm e_1, \dots, \pm e_d$  avec probabilité uniforme. Après avoir vérifié que  $(X_n)_{n \geq 0}$  est une chaîne de Markov dont la transition est irréductible, étudier la récurrence de la transition. [Indication — Calculer  $\mathbb{E}_0[V_0]$ .]

**Exercice 5.3.** *Marche aléatoire sur  $\mathbb{N}$  réfléchi au bord.* Une particule dans  $\mathbb{N}$  se déplace chaque seconde : avec probabilité  $p \in ]0, 1[$ , d'un pas vers la droite et, avec probabilité  $q = 1 - p$ , d'un pas vers la gauche. Sauf quand elle est en 0, auquel cas elle se déplace, avec probabilité 1, en 1.

- 1) Modéliser la trajectoire de la particule par une chaîne de Markov dont on précisera la transition.
- 2) Montrer que la transition est irréductible.
- 3) Étudier la récurrence de la transition (en fonction de  $p$ ). [Indication —  $\mathbb{P}_0(T_0 < \infty) = \mathbb{P}_1(H_0 < \infty)$ .]

**Exercice 5.4.** *Chaîne aux temps de saut.* Soit  $E$  un ensemble dénombrable et  $Q$  un noyau de transition sur  $E$  sans état absorbant. Soit  $(X_n)_{n \geq 0}$  une chaîne de Markov de transition  $Q$ . On définit la suite  $(J_k)_{k \geq 0}$  des temps de saut de  $(X_n)_{n \geq 0}$  par  $J_0 = 0$  et, pour tout  $k \geq 0$ ,  $J_{k+1} = \inf \{n > J_k ; X_n \neq X_{J_k}\}$ .

- 1) Montrer que, pour tout  $k \geq 0$ ,  $J_k$  est un temps d'arrêt fini presque sûrement.
- 2) On définit, pour tout  $k \geq 0$ ,  $Y_k = X_{J_k}$  et  $S_k = J_k - J_{k-1}$  (on pose arbitrairement  $S_0 = 1$ ). Montrer que  $((S_k, Y_k))_{k \geq 0}$  est une chaîne de Markov, dont on précisera la transition.
- 3) En déduire que  $(Y_k)_{k \geq 0}$  est une chaîne de Markov de transition donnée par  $\check{Q}(x, y) = \mathbb{1}_{y \neq x} Q(x, y) (1 - Q(x, x))^{-1}$ .
- 4) Montrer que  $Q$  et  $\check{Q}$  ont les mêmes classes de communication, et que les classes correspondantes sont de même nature (fermée ou non, récurrente ou transiente).

### Pour s'entraîner

**Exercice 5.5.** *Marche sur un arbre.* Soit  $k \geq 1$ . On note  $T_k$  l'arbre enraciné dont chaque sommet a exactement  $k$  enfants. On note  $(X_n)_{n \geq 0}$  la marche aléatoire sur le graphe  $T_k$ .

- 1) Vérifier que la transition est irréductible.
- 2) Étudier la récurrence de la transition.

**Exercice 5.6.** *Chaîne ralentie.* Soit  $E$  un ensemble dénombrable,  $Q$  un noyau de transition sur  $E$  et  $\alpha \in ]0, 1[$ . On définit  $\hat{Q} = \alpha I + (1 - \alpha)Q$ . Montrer que  $Q$  et  $\hat{Q}$  ont les mêmes classes de communication, et que les classes correspondantes sont de même nature (fermée ou non, récurrente ou transiente).

### Pour aller plus loin

**Exercice 5.7.** *Musculature markovienne.* Soit  $(X_n)_{n \geq 0}$  une chaîne de Markov sur un ensemble dénombrable  $E$ . Soit  $C$  une classe récurrente et  $y \in C$ .

- 1) Montrer que, pour tout  $x \in C$ ,  $\mathbb{P}_x(H_y < \infty) = 1$ .
- 2) Montrer que, pour tout  $x \in C$ ,  $\mathbb{P}_x(V_y = \infty) = 1$ .
- 3) Montrer que, pour tout  $x \in E$ ,  $\mathbb{P}_x(H_y < \infty) = \mathbb{P}_x(H^C < \infty)$ .

## TD6 — Chaînes de Markov — Mesures invariantes

**Exercice 6.1.** *Modèle d'Ehrenfest.* On considère deux urnes et on y répartit  $N$  boules. À chaque instant  $n \geq 0$ , on tire une boule au hasard et on la change d'urne.

- 1) Modéliser le nombre de boules dans la première urne par une chaîne de Markov  $(X_n)_{n \geq 0}$  sur  $E = \llbracket 0, N \rrbracket$  dont on précisera le graphe et la matrice de transition  $Q$ .
- 2) Montrer que  $Q$  admet une unique probabilité invariante que l'on déterminera.
- 3) Pour tout  $x \in E$ , on note  $T_x = \inf \{n \geq 1 ; X_n = x\}$ . Comparer  $\mathbb{E}_N[T_N]$  et  $\mathbb{E}_{\lfloor N/2 \rfloor}[T_{\lfloor N/2 \rfloor}]$ .
- 4) Pour tout  $x \in E$  et  $n \geq 0$ , calculer  $\mathbb{E}_x[X_n]$ . [Indication — Pour tout  $y \in E$ , calculer  $\mathbb{E}_x[X_{n+1} | X_n = y]$ .]

**Exercice 6.2.** *Processus de naissance et de mort.* On considère la transition  $Q$  sur  $\mathbb{N}$  donnée par  $Q(0, 0) = 1/2$ ,  $Q(0, 1) = 1/2$  et, pour tout  $x \geq 1$ ,

$$Q(x, y) = \begin{cases} 1/3 & \text{si } y = x - 1 \\ 1/2 & \text{si } y = x \\ 1/6 & \text{si } y = x + 1. \end{cases}$$

- 1) Montrer que la transition est irréductible.
- 2) Montrer que la transition admet une probabilité réversible que l'on calculera.
- 3) Calculer l'espérance du temps de premier passage en 0 pour la chaîne de Markov de transition  $Q$  partant de 0.

**Exercice 6.3.** *Marche du cavalier.* Sur un échiquier, un cavalier part de la case A1 et se déplace au hasard. Au bout de combien de temps en moyenne le cavalier revient-il sur la case A1 ?

### Pour s'entraîner

**Exercice 6.4.** *Nombre de particules.* Pour tout  $n \geq 0$ ,  $X_n$  désigne le nombre de particules présentes à l'instant  $n$  dans un volume  $V$ . On suppose que, pendant l'intervalle de temps  $[n, n + 1[$ , chacune des  $X_n$  particules présentes a une probabilité  $p \in ]0, 1[$  de quitter  $V$  et que, pendant ce même temps, un nombre aléatoire de nouvelles particules, suivant une loi de Poisson de paramètre  $a$ , entre dans  $V$ . On suppose que les différents phénomènes aléatoires ainsi considérés sont indépendants les uns des autres.

- 1) Montrer que  $(X_n)_{n \geq 0}$  est une chaîne de Markov sur  $\mathbb{N}$ , dont on précisera la matrice de transition  $Q$ .
- 2) Pour tout  $t \in \mathbb{R}$ , calculer directement  $\mathbb{E}_x[e^{itX_1}]$ .
- 3) En déduire la fonction caractéristique de  $X_1$  lorsque  $X_0$  suit une loi de Poisson.
- 4) Montrer que, pour une valeur convenable de  $\theta$ , la loi de Poisson  $\mathcal{P}(\theta)$  est invariante pour  $Q$ .
- 5) Montrer que  $Q$  est irréductible et récurrente positive.
- 6) Pour  $x \in E$ , la fréquence des passages en  $x$  entre l'instant 0 et l'instant  $n$  converge-t-elle quand  $n \rightarrow \infty$  ?

**Exercice 6.5.** Un étudiant range ses trois livres préférés ensemble sur une étagère. Chaque matin, il prend l'un des trois, la probabilité qu'il prenne le  $i^e$  (en partant de la gauche) étant  $p_i$ . Le soir, il range le livre à la gauche des deux autres. Sur une année, combien de fois les livres ont-ils été rangés dans l'ordre 1, 2, 3 ?

### Pour aller plus loin

**Exercice 6.6.** Pour chacun des modèles suivants, déterminer le caractère récurrent positif par l'identification des mesures invariantes, et retrouver le caractère récurrent ou transient par l'identification des mesures excessives :

- 1) chaîne aux temps de saut (voir exercice 5.4) ;
- 2) château de cartes (voir exercice 5.1) ;
- 3) marche sur  $\mathbb{N}$  réfléchi en 0 (voir exercice 5.3) ;
- 4) marche simple sur  $\mathbb{Z}$  (avec  $Q(x, x + 1) = p$  et  $Q(x, x - 1) = 1 - p$  pour  $p \in ]0, 1[$ ).

## TD7 — Chaînes de Markov — Convergence

**Exercice 7.1.** On considère une chaîne de Markov  $(X_n)_{n \geq 0}$  sur  $E = \llbracket 1, 4 \rrbracket$  de matrice de transition

$$Q = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1/2 & 0 & 1/4 & 1/4 \\ 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

- 1) Montrer qu'il existe une unique probabilité  $\pi$  invariante pour  $Q$ . Calculer  $\pi$ .  $\pi$  est-elle réversible ?
- 2) Montrer que la proportion de temps passé par la chaîne dans l'état 3 entre les instants 0 et  $n$  converge  $\mathbb{P}_1$ -p.s. quand  $n \rightarrow \infty$  et calculer sa limite.
- 3) Montrer que  $\frac{1}{n} \sum_{k=0}^{n-1} X_k(X_k + 1)$  converge  $\mathbb{P}_2$ -p.s. quand  $n \rightarrow \infty$  et calculer sa limite.
- 4) Montrer que  $\frac{1}{n} \sum_{k=0}^{n-1} X_k X_{k+1}$  converge  $\mathbb{P}_3$ -p.s. quand  $n \rightarrow \infty$  et calculer sa limite.
- 5) Montrer que, pour tout  $t \in \mathbb{R}$ ,  $(\mathbb{E}_4[e^{tX_n}])_{n \geq 0}$  converge et calculer sa limite.

**Exercice 7.2.** *Marche aléatoire sur  $\mathbb{N}$  réfléchie au bord (épisode 2).* On reprend le contexte de l'exercice 5.3. Une particule dans  $\mathbb{N}$  se déplace chaque seconde : avec probabilité  $p \in ]0, 1/2[$ , d'un pas vers la droite et, avec probabilité  $q = 1 - p$ , d'un pas vers la gauche. Sauf quand elle est en 0, auquel cas elle se déplace, avec probabilité 1, en 1.

- 1) Montrer que  $\pi = (1 - \frac{1}{2q})\delta_0 + \frac{1}{2q}\mathcal{G}(2 - \frac{1}{q})$  est une probabilité invariante pour  $Q$ .
- 2) La transition est-elle récurrente ou transiente ?
- 3) Montrer que  $\frac{1}{n} \sum_{k=0}^{n-1} X_k$  converge  $\mathbb{P}_0$ -p.s. lorsque  $n \rightarrow \infty$  et préciser sa limite.
- 4) Calculer l'espérance du temps de retour en 2 pour la chaîne partant de 2.
- 5) Calculer l'espérance du temps d'atteinte de 4 pour la chaîne partant de 0.

**Exercice 7.3.** *Renouvellement.* Soit  $(S_k)_{k \geq 1}$  une suite de variables aléatoires indépendantes et de même loi, à valeurs dans  $\mathbb{N}^*$  et intégrables.

- 1) Montrer que le processus  $(X_n)_{n \geq 0}$  défini par  $X_n = \inf \{m \geq n ; \exists k \geq 0 S_1 + \dots + S_k = m\} - n$  est une chaîne de Markov, dont on précisera la transition  $Q$ .
- 2) Étudier la convergence de  $(\mathbb{P}(X_n = 0))_{n \geq 0}$ .
- 3) Interprétation — Antoine pratique le badminton en club. Quelle est la probabilité que le cordage de sa raquette porte-bonheur soit cassé le jour du championnat départemental de mai prochain ?

## Pour s'entraîner

**Exercice 7.4.** *Chaîne non irréductible.* On admet pour le moment le théorème suivant.

**Théorème.** Soit  $E$  un ensemble dénombrable et  $Q$  une matrice de transition sur  $E$ . Soit  $(X_n)_{n \geq 0}$  une chaîne de Markov sur  $E$  de transition  $Q$ . Soit  $C$  une classe fermée récurrente positive apériodique,  $\pi_C$  l'unique distribution invariante supportée par  $C$  et  $H^C$  le temps d'atteinte de  $C$ .

$$\forall z \in C \quad \mathbb{P}(X_n = z) \xrightarrow[n \rightarrow \infty]{} \mathbb{P}(H^C < \infty) \pi_C(z).$$

Considérons l'exemple  $E = \llbracket 1, 7 \rrbracket$  et

$$Q = \begin{pmatrix} 0 & 2/3 & 1/3 & 0 & 0 & 0 & 0 \\ 1/3 & 0 & 0 & 1/6 & 1/2 & 0 & 0 \\ 0 & 0 & 1/5 & 4/5 & 0 & 0 & 0 \\ 0 & 0 & 2/3 & 1/3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1/3 & 0 & 2/3 \\ 0 & 0 & 0 & 0 & 1/6 & 1/2 & 1/3 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 & 0 \end{pmatrix}.$$

- 1) Déterminer les classes de communication de la transition, ainsi que leur nature.
- 2) Donner l'ensemble des probabilités invariantes pour  $Q$ .
- 3) En notant  $A$  la réunion des classes fermées, calculer, sous  $\mathbb{P}_1$ , l'espérance du temps d'atteinte de  $A$ .
- 4) Montrer que  $(\mathcal{L}_1(X_n))_{n \geq 0}$  converge et préciser sa limite.
- 5) Prouver le théorème énoncé ci-dessus.

## Pour aller plus loin

**Exercice 7.5.** *Processus de Galton-Watson surcritique.* Soit  $\nu$  une probabilité sur  $\mathbb{N}$  de moyenne  $m > 1$  et de variance  $\sigma^2 < \infty$ . Soit  $(\xi_{n,i})_{(n,i) \in (\mathbb{N}^*)^2}$  une famille de variables aléatoires indépendantes et de même loi  $\nu$ . Soit  $(Z_n)_{n \geq 0}$  la suite de variables aléatoires définie par la récurrence suivante :  $Z_0 = 1$  et, pour tout  $n \geq 0$ ,

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_{n,i}.$$

Pour tout  $n \geq 0$  et  $s \in [0, 1]$ , soit  $f_n(s) = \mathbb{E}[s^{Z_n}]$ .

- 1) Montrer que  $(Z_n)_{n \geq 0}$  est une chaîne de Markov.
- 2) Uniquement dans cette question, on regarde le cas  $\nu = p\delta_0 + (1-p)\delta_2$  (avec  $p < \frac{1}{2}$ ). Donner le noyau de transition de  $(Z_n)_{n \geq 0}$ , ainsi que ses classes de communication et leur nature (fermée ou non).
- 3) Montrer que  $f_1$  admet exactement deux points fixes : 1 et  $q \in [0, 1[$ .
- 4) Montrer que, pour tout  $n \geq 1$ ,  $f_n = f_1 \circ \dots \circ f_1$ , où  $f_1$  apparaît  $n$  fois.
- 5) Pour tout  $n \geq 0$ , calculer  $\mathbb{E}[Z_n]$  et  $\text{Var}(Z_n)$ .
- 6) Montrer que  $\mathbb{P}(\exists n \geq 0 \ Z_n = 0) = \lim_{n \rightarrow \infty} \mathbb{P}(Z_n = 0) = q$ .
- 7) Montrer que  $\mathbb{P}(Z_n \rightarrow \infty \mid \forall n \geq 0 \ Z_n > 0) = 1$ .

## TD8 — Chaînes de Markov — Théorème-limite central

**Exercice 8.1.** *La météo en Syldavie (épisode 2).* On rappelle (voir exercice 3.2) que le temps évolue de façon assez simple en Syldavie : s'il fait beau, alors il fera encore beau le lendemain avec une probabilité de  $5/6$  ; s'il pleut, alors il pleuvra le lendemain avec une probabilité de  $2/3$ . Donner une valeur approchée de la probabilité qu'il y ait plus de 275 jours de beau temps en Syldavie dans l'année 2026.

**Exercice 8.2.** *Variante pour la variance.* Soit  $E$  un ensemble fini,  $Q$  une transition sur  $E$  irréductible,  $\pi$  l'unique probabilité invariante pour  $Q$ ,  $f: E \rightarrow \mathbb{R}$  avec  $\pi(f) = 0$ .

1) Montrer que

$$\lim_{n \rightarrow \infty} \text{Var}_{\pi} \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(X_k) \right) = \text{Var}_{\pi}(f) + 2 \lim_{n \rightarrow \infty} \sum_{k=1}^n \left( 1 - \frac{k}{n} \right) \text{Cov}_{\pi}(f(X_0), f(X_k)).$$

2) Montrer que, si  $Q$  est de plus apériodique,

$$\lim_{n \rightarrow \infty} \text{Var}_{\pi} \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} f(X_k) \right) = \text{Var}_{\pi}(f) + 2 \sum_{k=1}^{\infty} \text{Cov}_{\pi}(f(X_0), f(X_k)).$$

Le reste de l'exercice a pour objectif de prouver que la limite des variances est bien la variance-limite apparaissant dans le théorème-limite central markovien.

3) Montrer que l'équation, dite de **Poisson**, d'inconnue  $g: E \rightarrow \mathbb{R}$ ,  $(I - Q)g = f$  admet une unique solution, à une constante additive près.

4) Montrer que la suite  $(\sum_{k=0}^{n-1} (1 - \frac{k}{n}) Q^k f)_{n \geq 1}$  converge simplement et que sa limite  $g$  est solution de l'équation de Poisson.

5) Montrer que

$$\pi(2fg - f^2) = \text{Var}_{\pi}(f) + 2 \lim_{n \rightarrow \infty} \sum_{k=1}^n \left( 1 - \frac{k}{n} \right) \text{Cov}_{\pi}(f(X_0), f(X_k)).$$

6) Soit  $x \in E$ . Montrer que  $\tilde{g}(y) = \mathbb{E}_y [\sum_{k=0}^{H_x} f(X_k)]$  est solution de l'équation de Poisson.

7) Montrer que

$$\pi(2f\tilde{g} - f^2) = \pi(x) \mathbb{E}_x \left[ \left( \sum_{k=1}^{T_x} f(X_k) \right)^2 \right].$$

8) En déduire que, pour toute loi initiale  $\lambda$ , sous  $\mathbb{P}_{\lambda}$ ,  $n^{-1/2} \sum_{k=0}^{n-1} f(X_k) \rightarrow \mathcal{N}(0, \sigma_f^2)$  avec

$$\sigma_f^2 = \text{Var}_{\pi}(f) + 2 \lim_{n \rightarrow \infty} \sum_{k=1}^n \left( 1 - \frac{k}{n} \right) \text{Cov}_{\pi}(f(X_0), f(X_k)).$$

9) **Application** — Montrer le théorème-limite central pour la marche simple sur le graphe complet (voir exercice 3.5). Calculer la variance limite et la comparer à celle du théorème-limite central pour des variables indépendantes et de loi la probabilité invariante. Prolonger l'étude pour la chaîne ralentie (voir exercice 5.6).

### Pour s'entraîner

**Exercice 8.3.** Sur  $E = \mathbb{Z}$ , on considère la transition  $Q$  donnée par :  $Q(0, 0) = 1/2$  et, pour tout  $x \geq 1$ ,  $Q(0, x) = x^{-3}/(2c)$ ,  $Q(x, -x) = Q(-x, 0) = 1$ , où  $c = \sum_{x=1}^{\infty} x^{-3}$ . On se donne une loi  $\lambda$  sur  $E$  et  $(X_n)_{n \geq 0} \sim \text{Markov}(\lambda, Q)$ .

1) Montrer que  $Q$  est irréductible et récurrente positive.

2) Calculer la probabilité  $\pi$  invariante pour  $Q$ .

3) Montrer que  $n^{-1/2} \sum_{k=0}^{n-1} X_k \xrightarrow{(d)} \delta_0 = \mathcal{N}(0, 0)$  quand  $n \rightarrow \infty$  (en particulier  $(X_n)_{n \geq 0}$  vérifie un théorème-limite central).

4) Montrer toutefois que, pour tout  $x \in \mathbb{Z}$ ,

$$\mathbb{E}_x \left[ \left( \sum_{k=0}^{T_x-1} |X_k| \right)^2 \right] = +\infty.$$

5) Montrer aussi que

$$\text{Var}_\lambda \left[ \left( \frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} X_k \right)^2 \right] \xrightarrow{n \rightarrow \infty} +\infty.$$

## Pour aller plus loin

**Exercice 8.4.** Calculer la variance-limite  $\sigma_f^2$  dans les cas suivants et comparer  $\sigma_f^2$  et  $\text{Var}_\pi(f)$  :

- 1) le château de cartes (voir exercice 5.1) et  $f(y) = \mathbb{1}_{y=x} - \pi(x)$  ;
- 2) la marche sur l'hypercube (voir exercice 3.3 et  $f(y) = \mathbb{1}_{y=x} - \pi(x)$  ;
- 3) la marche simple symétrique sur  $\mathbb{Z}/n\mathbb{Z}$ .

## TD9 — Vecteurs gaussiens

**Exercice 9.1.** Soit  $(X, \varepsilon)$  un couple de variables aléatoires indépendantes tel que  $X \sim \mathcal{N}(0, 1)$  et  $\mathbb{P}(\varepsilon = 1) = \frac{1}{2} = \mathbb{P}(\varepsilon = -1)$ .

- 1) Montrer que  $\varepsilon X$  et  $\varepsilon |X|$  ont même loi que  $X$ .
- 2) Le vecteur  $(X, \varepsilon X)$  est-il gaussien ?

**Exercice 9.2.** Soit  $(X_1, X_2)$  un vecteur gaussien centré à valeurs dans  $\mathbb{R}^2$ , de matrice de covariance  $\Gamma = \begin{pmatrix} 3 & 6 \\ 6 & 12 \end{pmatrix}$ . Montrer que les variables aléatoires  $X_1$  et  $X_2$  sont proportionnelles.

**Exercice 9.3.** Soit  $X = (X_1, X_2)^\top \sim \mathcal{N}(0, I_2)$ .

- 1) Montrer que, si  $O \in O_2(\mathbb{R})$ , alors  $OX$  a même loi que  $X$ .
- 2) Montrer que  $X_1 X_2$  a même loi que  $\frac{1}{2}(X_1^2 - X_2^2)$ . [Indication : on pourra utiliser la formule  $X_1 X_2 = \frac{1}{4}((X_1 + X_2)^2 - (X_1 - X_2)^2)$ .]

**Exercice 9.4.** Soit  $(X_n)_{n \geq 1}$  une suite de variables aléatoires réelles i.i.d. et de carré sommable. Pour tout  $n \geq 0$ , soit  $S_n = \sum_{k=1}^n X_k$ . Montrer que  $\mathbb{P}(S_n > 0, S_{2n} < 0)$  converge quand  $n \rightarrow \infty$  et préciser sa limite.

### Pour s'entraîner

**Exercice 9.5.** Dans  $\mathbb{R}^d$ , soit  $X = (X_1, \dots, X_d)^\top$  un vecteur gaussien centré de matrice de covariance  $\Gamma$ .

- 1) Montrer que  $\mathbb{E}[\|X\|_2^2] = \text{tr}(\Gamma)$ .
- 2) Montrer que  $\mathbb{E}[X_1^4] = 3 \text{Var}(X_1)^2$  et  $\mathbb{E}[X_1^2 X_2^2] = \text{Var}(X_1) \text{Var}(X_2) + 2 \text{Cov}(X_1, X_2)^2$ .
- 3) En déduire que  $\mathbb{E}[\|X\|_2^4] = \text{tr}(\Gamma)^2 + 2 \text{tr}(\Gamma^2)$ .
- 4) Soit  $\lambda_1, \dots, \lambda_d$  les valeurs propres de  $\Gamma$  et  $\rho := \max \{\lambda_k ; k \in \llbracket 1, d \rrbracket\}$ . Montrer que, pour tout  $\alpha < 1/\rho$ ,

$$\mathbb{E}[e^{\alpha \|X\|_2^2/2}] = \prod_{k=1}^d \frac{1}{\sqrt{1 - \alpha \lambda_k}}.$$

### Pour aller plus loin

**Exercice 9.6.** *Entropie.* Si  $\mu$  est une mesure de probabilité sur  $\mathbb{R}^d$  de densité  $f$  par rapport à la mesure de Lebesgue, on définit son *entropie* par

$$H(\mu) = - \int f(x) \log(f(x)) dx.$$

- 1) Soit  $m \in \mathbb{R}^d$  et  $\Gamma \in \mathcal{M}_d(\mathbb{R})$  une matrice symétrique positive. Montrer que  $\mathcal{N}(m, \Gamma)$  est à densité par rapport à la mesure de Lebesgue si et seulement si  $\Gamma$  est inversible (i.e. définie positive), et dans ce cas la densité est donnée par

$$\phi_{m, \Gamma}(x) = \frac{1}{\sqrt{(2\pi)^d \det(\Gamma)}} e^{-(x-m)^\top \Gamma^{-1} (x-m)}.$$

- 2) Pour  $m \in \mathbb{R}^d$  et  $\Gamma \in \mathcal{M}_d(\mathbb{R})$  symétrique définie positive, calculer l'entropie de  $\mathcal{N}(m, \Gamma)$ .
- 3) Soit  $\mu$  une mesure de probabilité sur  $\mathbb{R}^d$  de densité  $f$  par rapport à la mesure de Lebesgue et telle que  $\int x d\mu(x) = m$  et  $\int \|x - m\|_2^2 d\mu(x) = d\sigma^2$ . Montrer que

$$H(\mu) \leq H(\mathcal{N}(m, \sigma^2 I_d)),$$

avec égalité si et seulement si  $\mu = \mathcal{N}(m, \sigma^2 I_d)$ .

## TD10 — Statistique — Estimation non asymptotique

**Exercice 10.1.** Soit  $n \in \mathbb{N}^*$ . On considère le modèle statistique  $(\mathbb{N}, \mathcal{P}(\mathbb{N}), (\mathcal{B}(n, p))_{p \in [0,1]})$ . Si  $\alpha \in \mathbb{R}$ , montrer qu'il existe un estimateur sans biais de  $p^\alpha$  si et seulement si  $\alpha \in \llbracket 0, n \rrbracket$ .

**Exercice 10.2.** Soit  $n \in \mathbb{N}^*$ . On considère un modèle statistique  $(\Omega, \mathcal{A}, (\mathbb{P}_\theta)_{\theta \in \Theta}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_\theta)_{\theta \in \Theta})^{\otimes n}$ . On suppose que toutes les lois  $P_\theta$  ont un moment d'ordre 2 fini. On note  $X = (X_1, \dots, X_n)$  l'observation.

- 1) On suppose que  $\theta \in \Theta \mapsto \mathbb{E}_\theta[X_1]$  n'est pas constante. Montrer que le meilleur (au sens du risque quadratique) estimateur  $T(X)$  affine (en  $X$ ) et sans biais de  $\mathbb{E}_\theta[X_1]$  est la moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ .
- 2) On définit la variance empirique de  $(X_1, \dots, X_n)$  par  $\hat{\sigma}_n^2 := \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^2$ . Calculer  $\mathbb{E}_\theta[\hat{\sigma}_n^2]$  et en déduire un estimateur sans biais de  $\sigma^2(\theta) := \text{Var}_\theta(X_1)$ .

**Exercice 10.3.** Soit  $n \in \mathbb{N}^*$ . On considère le modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathcal{U}([0, \mu]))_{\mu \in \mathbb{R}_+^*})^{\otimes n}$ . On note  $X = (X_1, \dots, X_n)$  l'observation et on définit  $\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$ ,  $(X_{(1)}, \dots, X_{(n)})$  le  $n$ -uplet  $X$  ordonné par ordre croissant (statistiques d'ordre),  $M_n := X_{(n)} = \max(X_1, \dots, X_n)$ , et  $L_n := \frac{1}{2}(X_{(\lfloor (n+1)/2 \rfloor)} + X_{(\lceil (n+1)/2 \rceil)})$  (médiane de  $(X_1, \dots, X_n)$ ).

- 1) Montrer que  $2\bar{X}_n$  est un estimateur sans biais de  $\mu$ . Construire avec  $2\bar{X}_n$  un intervalle de confiance pour  $\mu$ .
- 2) Montrer que  $M_n$  est un estimateur biaisé de  $\mu$ . Construire avec  $M_n$  un intervalle de confiance pour  $\mu$ .
- 3) Montrer que  $2L_n$  est un estimateur sans biais de  $\mu$ . Construire avec  $2L_n$  un intervalle de confiance pour  $\mu$ .

*Remarque.* L'estimation donnée par  $M_n$  a été utilisée par les Alliés pendant la Seconde Guerre Mondiale pour évaluer le nombre de chars allemands. Ils sont arrivés au sixième des évaluations données par leurs services de renseignements. Et cette évaluation statistique était bien meilleure<sup>1</sup> !

### Pour s'entraîner

**Exercice 10.4.** *Répartition allélique.* Lorsqu'un gène a deux allèles  $A$  et  $a$ , les génotypes  $AA$ ,  $Aa$  et  $aa$  se présentent dans la population avec les fréquences respectives  $p^2$ ,  $2p(1-p)$  et  $(1-p)^2$ , où  $p \in [0, 1]$ . Proposer un modèle statistique et un estimateur sans biais de  $p$ .

**Exercice 10.5.** *Réduction de la variance.* On veut estimer  $p = \int_5^\infty \phi(x)dx$ , où  $\phi(x) = \exp(-x^2/2)/\sqrt{2\pi}$ . On dispose d'un échantillon  $(X_k)_{k \in \llbracket 1, n \rrbracket}$  de loi  $\mathcal{N}(0, 1)$ . On considère les estimateurs

$$\hat{p}_1 = \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{X_k > 5} \quad \text{et} \quad \hat{p}_2 = \frac{1}{n} \sum_{k=1}^n \frac{\phi(X_k + 5)}{\phi(X_k)} \mathbb{1}_{X_k > 0}.$$

- 1) Montrer que les deux estimateurs sont sans biais.
- 2) Montrer que  $p \leq \phi(5)/5$ .
- 3) Calculer les variances des deux estimateurs.
- 4) Quelle valeur de  $n$  pour estimer  $p$  à  $10^{-7}$  ?

### Pour aller plus loin

**Exercice 10.6.** *Estimateur de Haldane.* On cherche à estimer la probabilité qu'un dé tombe sur 6 en le lançant jusqu'à obtenir  $r$  fois le 6. En notant  $N_r$  le nombre de lancers, on estime  $p$  par  $\hat{p}_r := r/N_r$ .

- 1) Calculer le biais de  $\hat{p}_1$  en fonction de  $p$ .
- 2) Montrer que, pour tout  $r \geq 1$ , pour tout  $n \geq r$ ,  $\mathbb{P}(N_r = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}$ .
- 3) Le généticien John Burdon Sanderson Haldane propose, pour  $r \geq 2$ , d'utiliser plutôt  $\tilde{p}_r := (r-1)/(N_r-1)$ . Qu'en pensez-vous ?

1. Voir [https://fr.wikipedia.org/wiki/Problème\\_du\\_char\\_d'assaut\\_allemand](https://fr.wikipedia.org/wiki/Problème_du_char_d'assaut_allemand) et les références citées.

## TD11 — Statistique — Estimation asymptotique

**Exercice 11.1.** On considère un modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_\theta)_{\theta \in \Theta})^{\otimes \infty}$ . On note  $X = (X_n)_{n \geq 1}$  l'observation et, pour tout  $n \in \mathbb{N}^*$ , on définit la fonction de répartition empirique des  $n$  premières observations par  $F_n(x) := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{X_i \leq x}$ . Montrer que, pour tout  $x \in \mathbb{R}$ ,  $(F_n(x))_{n \geq 1}$  est un estimateur de  $F_\theta(x) := P_\theta([-\infty, x])$  sans biais, fortement consistant et asymptotiquement normal.

**Exercice 11.2.** *Stabilisation de la variance.* On considère le modèle statistique  $([0, 1], \mathcal{B}([0, 1]), (\mathcal{B}(p))_{p \in ]0, 1[})^{\otimes \infty}$ . Chercher une fonction  $\psi: ]0, 1[ \rightarrow \mathbb{R}$  telle que  $\sqrt{n}(\psi(\bar{X}_n) - \psi(p)) \rightarrow \mathcal{N}(0, 1)$  et en déduire un intervalle de confiance asymptotique pour  $p$ .

**Exercice 11.3.** À la suite de l'exercice 10.3, dans le modèle avec une infinité d'observations, montrer que :

- 1)  $(2\bar{X}_n)_{n \geq 1}$  est un estimateur de  $\mu$  fortement consistant et asymptotiquement normal ;
- 2)  $(M_n)_{n \geq 1}$  est un estimateur de  $\mu$  fortement consistant et tel que  $(n(\mu - M_n))_{n \geq 1}$  converge en loi vers une limite que l'on déterminera ;
- 3)  $(2L_n)_{n \geq 1}$  est un estimateur de  $\mu$  fortement consistant et asymptotiquement normal.

Dans chacun des cas, proposez un intervalle de confiance asymptotique pour  $\mu$ .

### Pour s'entraîner

**Exercice 11.4.** On considère un modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_\theta)_{\theta \in \Theta})^{\otimes \infty}$ . On suppose que toutes les lois  $P_\theta$  ont un moment d'ordre 4 fini. Montrer que la variance empirique  $(\hat{\sigma}_n^2)_{n \geq 1}$  est un estimateur de la variance  $\int x^2 dP_\theta(x) - (\int x dP_\theta(x))^2$  asymptotiquement normal et préciser sa variance asymptotique.

**Exercice 11.5.** On considère le modèle statistique  $(\Omega, \mathcal{A}, (\mathbb{P}_\lambda)_{\lambda \in \mathbb{R}_+^*}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathcal{E}(\lambda))_{\lambda \in \mathbb{R}_+^*})^{\otimes \infty}$ . On note  $X = (X_n)_{n \geq 1}$  l'observation. Pour tout  $n \geq 1$ , soit  $L_n = \min(X_1, \dots, X_n)$  et  $U_n = \max(X_1, \dots, X_n)$ .

- 1) Pour  $n \geq 1$  et  $\lambda > 0$ , soit  $Y_{\lambda, n} := n\lambda L_n$ . Montrer que, sous  $\mathbb{P}_\lambda$ ,  $(Y_{\lambda, n})_{n \geq 1}$  converge en loi et préciser sa limite.
- 2) Pour  $n \geq 1$  et  $\lambda > 0$ , soit  $Z_{\lambda, n} := \lambda U_n - \log(n)$ . Montrer que, sous  $\mathbb{P}_\lambda$ ,  $(Z_{\lambda, n})_{n \geq 1}$  converge en loi et préciser sa limite.

Application — On dispose de  $n$  ampoules de même facture, dont on modélise les durées de vie par des variables de lois exponentielles indépendantes.

- 3) On observe les durées de vie des  $n$  ampoules. Proposer un estimateur de la durée de vie moyenne. Est-il sans biais ? Est-il consistant ? Est-il asymptotiquement normal ? Proposer aussi un intervalle de confiance (asymptotique).
- 4) On sait seulement combien de temps a fonctionné l'ampoule qui a grillé le plus tôt. Proposer un estimateur de la durée de vie moyenne. Est-il sans biais ? Est-il consistant ? Est-il asymptotiquement normal ? Proposer aussi un intervalle de confiance (asymptotique).
- 5) On sait seulement combien de temps a fonctionné l'ampoule qui a grillé le plus tard. Proposer un estimateur de la durée de vie moyenne. Est-il sans biais ? Est-il consistant ? Est-il asymptotiquement normal ? Proposer aussi un intervalle de confiance (asymptotique).

### Pour aller plus loin

**Exercice 11.6.** On considère le modèle statistique  $(\{0, 1\}^{\mathbb{N}}, \mathfrak{P}(\{0, 1\})^{\otimes \mathbb{N}}, (\mathbb{P}_p)_{p \in ]0, 1[})$  où, pour tout  $p \in ]0, 1[$ ,  $\mathbb{P}_p = \text{Markov}(\delta_0, Q_p)$  avec  $Q_p = \begin{pmatrix} 1-p & p \\ 1 & 0 \end{pmatrix}$ . Proposer un intervalle de confiance asymptotique pour  $p$ .

## TD12 — Statistique — Méthode des moments

**Exercice 12.1.** On considère le modèle statistique  $(\Omega, \mathcal{A}, (\mathbb{P}_\lambda)_{\lambda \in \mathbb{R}_+^*}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathcal{E}(\lambda))_{\lambda \in \mathbb{R}_+^*})^{\otimes \infty}$ . On note  $X = (X_n)_{n \geq 1}$  l'observation.

- 1) Proposer un estimateur des moments de  $\lambda$  à l'aide de la fonction  $m: \lambda \mapsto \mathbb{E}_\lambda[X_1]$ . Montrer qu'il est asymptotiquement normal.
- 2) Proposer un estimateur des moments de  $\lambda$  à l'aide de la fonction  $v: \lambda \mapsto \text{Var}_\lambda(X_1)$ . Montrer qu'il est asymptotiquement normal.
- 3) Pour tout  $t > 0$ , proposer un estimateur des moments de  $\lambda$  à l'aide de la fonction  $q_t: \lambda \mapsto \mathbb{P}_\lambda(X_1 > t)$ . Montrer qu'il est asymptotiquement normal.
- 4) Comparer les estimateurs ci-dessus.

### Pour s'entraîner

**Exercice 12.2.** On considère le modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_c)_{c \in \mathbb{R}_+^*})^{\otimes \infty}$  où, pour tout  $c \in \mathbb{R}_+^*$ ,  $P_c$  est la loi de Cauchy de paramètre  $c$ , autrement dit la loi dont la densité par rapport à la mesure de Lebesgue est donnée par

$$f_c(x) = \frac{c}{\pi(c^2 + x^2)}.$$

Proposer un estimateur des moments de  $c$  et étudier ses propriétés de convergence.

**Exercice 12.3.** On considère le modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_{\alpha, \beta})_{(\alpha, \beta) \in (\mathbb{R}_+^*)^2})^{\otimes \infty}$  où, pour tout  $(\alpha, \beta) \in (\mathbb{R}_+^*)^2$ ,  $P_{\alpha, \beta} = \text{Beta}(\alpha, \beta)$  est la loi dont la densité par rapport à la mesure de Lebesgue est donnée par

$$f_{\alpha, \beta}(x) = \mathbb{1}_{0 < x < 1} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{1-\alpha} (1-x)^{1-\beta}.$$

Proposer un estimateur des moments de  $(\alpha, \beta)$  à l'aide de la fonction  $g: (\alpha, \beta) \mapsto (\int x dP_{\alpha, \beta}(x), \int x^2 dP_{\alpha, \beta}(x))$ . Proposer également une région de confiance asymptotique pour  $(\alpha, \beta)$ .

### Pour aller plus loin

**Exercice 12.4.** On considère le modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_{m, \sigma^2})_{(m, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*})^{\otimes \infty}$  où, pour tout  $(m, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+^*$ ,  $P_{m, \sigma^2} = \mathcal{N}(m, \sigma^2)$ . Proposer un estimateur des moments de  $(m, \sigma^2)$  à l'aide de la fonction quantile  $Q_{m, \sigma^2} := F_{m, \sigma^2}^{-1}$ , où, pour tout  $x \in \mathbb{R}$ ,  $F_{m, \sigma^2}(x) = P_{m, \sigma^2}([-\infty, x])$ . Montrer qu'il est asymptotiquement normal.

## TD13 — Statistique — Modèles dominés

**Exercice 13.1.** On considère un modèle statistique dominé  $(\Omega, \mathcal{A}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ . On note  $\mu$  une mesure dominante. Pour tout  $\theta \in \Theta$ , on note  $f_\theta$  la densité de  $\mathbb{P}_\theta$  par rapport à  $\mu$  et on suppose que  $f_\theta > 0$ . Pour  $(\theta_0, \theta_1) \in \Theta^2$ , on définit

$$D(\theta_0, \theta_1) = \int \log \left( \frac{f_{\theta_1}(x)}{f_{\theta_0}(x)} \right) f_{\theta_1}(x) \mu(dx).$$

- 1) Montrer que  $D$  est bien définie (à valeurs dans  $\overline{\mathbb{R}}$ ).
- 2) Montrer que  $D \geq 0$  avec égalité si et seulement si  $\mathbb{P}_{\theta_0} = \mathbb{P}_{\theta_1}$ .
- 3) Dans le modèle  $(\Omega, \mathcal{A}, (\mathbb{P}_\theta)_{\theta \in \Theta}) = (\Omega, \mathcal{A}, (\mathcal{N}(m, 1))_{m \in \mathbb{R}})$ , calculer  $D$ .
- 4) On suppose que  $\Theta = \mathbb{R}^k$  et que le modèle a les mêmes propriétés de régularité que dans le cours. Montrer que, pour tout  $\theta_0 \in \Theta$ ,

$$D(\theta_0, \theta) = \frac{I(\theta_0)}{2} (\theta - \theta_0)^2 + o((\theta - \theta_0)^2) \quad \text{quand } \theta \rightarrow \theta_0.$$

**Exercice 13.2.** Déterminer l'estimateur du maximum de vraisemblance dans les modèles statistiques suivants et en déduire une région de confiance asymptotique :

- 1)  $(\mathbb{N}, \mathfrak{P}(\mathbb{N}), (\mathcal{P}(\lambda))_{\lambda \in \mathbb{R}_+^*})^{\otimes \infty}$  ;
- 2) pour tout  $k \in \mathbb{R}_+^*$ ,  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (\Gamma(k, \lambda))_{\lambda \in \mathbb{R}_+^*})^{\otimes \infty}$  ;
- 3)  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (\Gamma(k, \lambda))_{(k, \lambda) \in (\mathbb{R}_+^*)^2})^{\otimes \infty}$  ;
- 4)  $(\mathbb{N}^*, \mathfrak{P}(\mathbb{N}^*), (\mathcal{G}(p))_{p \in ]0, 1]})^{\otimes \infty}$ .

### Pour s'entraîner

**Exercice 13.3.** Dans le modèle  $(\mathbb{N}, \mathfrak{P}(\mathbb{N}), (\mathcal{B}(k, p))_{(k, p) \in \mathbb{N}^* \times ]0, 1]})^{\otimes \infty}$ , déterminer l'estimateur du maximum de vraisemblance.

**Exercice 13.4.** Soit  $E$  un ensemble fini et  $\lambda$  une mesure de probabilité sur  $E$ . On note  $\text{Trans}(E)$  l'ensemble des noyaux de transition sur  $E$ . On considère le modèle statistique  $(E^{\mathbb{N}}, \mathfrak{P}(E)^{\otimes \mathbb{N}}, (\text{Markov}(\lambda, Q))_{Q \in \text{Trans}(E)})$ . Déterminer l'estimateur du maximum de vraisemblance.

### Pour aller plus loin

**Exercice 13.5. Nombre d'erreurs.** Jeanne et Étienne sont chargés de relire séparément un mémoire, afin d'en corriger les fautes d'orthographe. On suppose qu'il y a  $n$  fautes dans le mémoire, que Jeanne détecte chaque faute (indépendamment des autres) avec probabilité  $p \in ]0, 1]$  et qu'Étienne détecte chaque faute (indépendamment des autres et indépendamment de Jeanne) avec probabilité  $q \in ]0, 1]$ . On note  $A$  le nombre de fautes détectées par Jeanne,  $B$  le nombre de fautes détectées par Étienne et  $C$  le nombre de fautes qu'ils ont tous les deux détectées.

- 1) En supposant  $n$  connu, proposer des estimateurs sans biais pour  $p$  et  $q$ . Montrer que  $AB/n^2$  et  $C/n$  sont des estimateurs sans biais de  $pq$  et qu'ils sont consistants (lorsque  $n \rightarrow \infty$ ).
- 2) En supposant  $n$  inconnu, proposer un estimateur  $\hat{n}$  de  $n$  tel que  $\hat{n}/n$  soit un estimateur consistant de 1 (quand  $n \rightarrow \infty$ ).

## TD14 — Statistique — Tests d'hypothèses

### Pour s'entraîner

**Exercice 14.1.** On considère le modèle statistique  $(\Omega, \mathcal{A}, (\mathbb{P}_m)_{m \in \mathbb{R}}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathcal{N}(m, 1))_{m \in \mathbb{R}})^{\otimes \infty}$ . Soit  $m_0 \in \mathbb{R}$ . On souhaite tester  $H_0: m \leq m_0$  contre  $H_1: m > m_0$ . Si  $\phi$  est un test statistique, on définit sa (fonction) **probabilité de rejet** (de l'hypothèse nulle) par  $\rho_\phi: m \in \mathbb{R} \mapsto \mathbb{E}_m[\phi] (= \mathbb{P}_m(\phi = 1))$ .

1) Pour tout  $n \in \mathbb{N}^*$ , calculer la probabilité de rejet du test  $\phi_n = \mathbb{1}_{\bar{X}_n > m_0}$  et en déduire les erreurs de première et seconde espèce.

2) Soit  $\alpha \in ]0, 1[$ . Pour tout  $n \in \mathbb{N}^*$ , déterminer la valeur minimale de  $x_n \in \mathbb{R}$  de sorte que le test  $\psi_n = \mathbb{1}_{\bar{X}_n > x_n}$  soit de niveau  $\alpha$ . Préciser alors la probabilité de rejet et les erreurs de première et seconde espèce. Montrer aussi que la suite de tests  $(\psi_n)_{n \geq 1}$  est **consistante**, i.e., pour tout  $m \in \Theta_1 := ]m_0, \infty[$ ,  $\rho_{\psi_n}(m) \rightarrow 1$  quand  $n \rightarrow \infty$  (la puissance  $\psi_n$  tend simplement vers 1).

3) Soit  $n \in \mathbb{N}^*$ . Pour tout  $m \in \mathbb{R}$ , on note  $f_m$  la densité de  $\mathcal{N}(m, 1)$  par rapport à la mesure de Lebesgue. Montrer que, pour tout  $m_1 > m_0$ , il existe  $c > 0$  tel que  $\psi_n = \mathbb{1}_{(f_{m_1}/f_{m_0})^{\otimes n}(X_1, \dots, X_n) > c}$ . En déduire que  $\psi_n$  est **uniformément plus puissant** que tout test  $\psi$  fonction de  $(X_1, \dots, X_n)$  et de niveau  $\alpha$ , au sens où, pour tout  $m \in \Theta_1$ ,  $\rho_{\psi_n}(m) \geq \rho_\psi(m)$  (la puissance de  $\psi_n$  est supérieure à celle de  $\psi$  en tout point).

**Exercice 14.2.** *Exponentielles translatées.* On considère le modèle statistique  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (P_\theta)_{\theta \in \mathbb{R}})^{\otimes \infty}$  où, pour tout  $\theta \in \mathbb{R}$ ,  $dP_\theta(x) = \mathbb{1}_{x > \theta} e^{-(x-\theta)} dx$ . Soit  $\theta_0 \in \mathbb{R}$  et  $\alpha \in ]0, 1[$ . Pour  $H_0: \theta \leq \theta_0$  contre  $H_1: \theta > \theta_0$ , déterminer le test du rapport de vraisemblance de niveau  $\alpha$  et d'erreur de seconde espèce minimale. Vérifier que ce test est consistant.

**Exercice 14.3.** Une étude médicale vise à connaître les risques de vertiges liés à la prise d'un certain médicament. Le médicament ne sera retiré de la vente si la probabilité de subir des vertiges après prise du médicament est supérieure à 10 %. Le laboratoire pharmaceutique affirme quant à lui que cette probabilité est inférieure à 8 %. Au total, 1163 patients soignés avec ce médicament ont été suivis et on a relevé les résultats suivants :

| vertiges | pas de vertiges |
|----------|-----------------|
| 108      | 1055            |

Quelle décision prendre au vu de ces résultats ?

### Pour aller plus loin

**Exercice 14.4.** On souhaite savoir si les chimpanzés préfèrent le rouge, le bleu, le vert ou le jaune. Un chercheur réalise donc l'expérience suivante sur 83 chimpanzés : il leur montre 4 cartons (un par couleur, les 4 cartons sont de forme identique et sont présentés dans une disposition aléatoire) ; il attend ensuite que le chimpanzé choisisse l'un des cartons et note la couleur choisie. Les résultats obtenus sont les suivants :

| rouge | bleu | vert | jaune |
|-------|------|------|-------|
| 30    | 22   | 15   | 16    |

Cette expérience est-elle une preuve expérimentale du fait que les chimpanzés ont une préférence de couleur ?

**Exercice 14.5.** *Comparaison de moyennes.* Soit  $\Theta \subset \mathcal{P}(\mathbb{R})^2$  un ensemble de couples de lois sur  $\mathbb{R}$  admettant un moment d'ordre 2 fini non nul. On considère le modèle statistique  $(\mathbb{R}^{\mathbb{N}^*} \times \mathbb{R}^{\mathbb{N}^*}, \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}^*} \otimes \mathcal{B}(\mathbb{R})^{\otimes \mathbb{N}^*}, (\mathbb{P}_{\mu, \nu})_{(\mu, \nu) \in \Theta})$  où  $\mathbb{P}_{\mu, \nu} = \mu^{\otimes \mathbb{N}^*} \otimes \nu^{\otimes \mathbb{N}^*}$ . On note  $(X, Y) = ((X_k)_{k \geq 1}, (Y_l)_{l \geq 1})$  l'observation.

1) On suppose ici que  $\mathbb{E}_{\mu, \nu}[X_1] = \mathbb{E}_{\mu, \nu}[Y_1]$ . En notant, pour tout  $(m, n) \in (\mathbb{N}^*)^2$ ,  $\hat{\sigma}_{X, m}^2 = \overline{(X - \bar{X}_m)^2}$  et  $\hat{\sigma}_{Y, n}^2 = \overline{(Y - \bar{Y}_n)^2}$ , montrer que

$$T_{m, n} := \frac{\bar{X}_m - \bar{Y}_n}{\sqrt{\frac{\hat{\sigma}_{X, m}^2}{m} + \frac{\hat{\sigma}_{Y, n}^2}{n}}} \xrightarrow[m \rightarrow \infty]{(d)} \mathcal{N}(0, 1).$$

2) Construire un test asymptotique pour  $H_0: \mathbb{E}_{\mu, \nu}[X_1] = \mathbb{E}_{\mu, \nu}[Y_1]$  contre l'hypothèse  $H_1: \mathbb{E}_{\mu, \nu}[X_1] \neq \mathbb{E}_{\mu, \nu}[Y_1]$ .