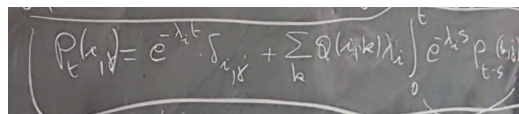


UNIVERSITÉ DE TECHNOLOGIE DE COMPIÈGNE

RO05

MODÉLISATION DES PHÉNOMÈNES ALÉATOIRES

Nikolaos LIMNIOΣ  
(nlimnios@utc.fr)


$$P_t(x) = e^{-\lambda_i t} \delta_{x_i} + \sum_k Q(i,k) \lambda_i \int_0^t e^{-\lambda_i s} p_{t-s}(x) ds$$

Septembre 2020



# Table des matières

<b>1</b>	<b>Rappels de probabilités</b>	<b>1</b>
1.1	Événements et probabilité . . . . .	1
1.2	Indépendance et conditionnement . . . . .	4
1.3	Variables aléatoires . . . . .	8
1.3.1	Définition, Loi et Espérance . . . . .	8
1.3.2	Lois usuelles . . . . .	14
1.4	Les fonctions génératrice et caractéristique . . . . .	18
1.5	Vecteurs aléatoires . . . . .	21
1.6	Probabilité et espérance conditionnelles . . . . .	26
1.7	Suites aléatoires et théorèmes limites . . . . .	29
1.8	Problèmes . . . . .	36
<b>2</b>	<b>Simulation stochastique et méthode de Monte Carlo</b>	<b>38</b>
2.1	Introduction . . . . .	38
2.2	Génération de variables aléatoires . . . . .	39
2.2.1	Suites équiréparties . . . . .	39
2.2.2	Génération d'une variable uniforme . . . . .	39
2.2.3	Génération des variables aléatoires discrètes . . . . .	40
2.2.4	Génération des variables aléatoires réelles . . . . .	41
2.3	Méthode de Monte Carlo . . . . .	42
2.3.1	Mise en oeuvre . . . . .	42
2.3.2	Précision . . . . .	43
2.3.3	Accélération . . . . .	44
2.4	Événements rares . . . . .	45
2.5	Problèmes . . . . .	45
<b>3</b>	<b>Chaînes de Markov</b>	<b>48</b>
3.1	Fonction de transition et exemples . . . . .	48
3.2	Décomposition de l'espace d'état . . . . .	54
3.3	Loi stationnaire et comportement asymptotique . . . . .	59
3.4	Chaînes de Markov périodiques . . . . .	64
3.5	Chaînes de Markov finies . . . . .	67
3.6	Problèmes . . . . .	72
<b>4</b>	<b>Processus de Poisson</b>	<b>74</b>
4.1	Introduction . . . . .	74
4.2	Processus de Poisson homogène . . . . .	74
4.3	Processus de Poisson non homogène . . . . .	75
4.4	Processus de Poisson composé . . . . .	77
4.5	Processus de Poisson marqué . . . . .	78

4.6	Problèmes . . . . .	80
<b>5</b>	<b>Processus de Markov</b>	<b>82</b>
5.1	Fonction de transition et propriété de Markov . . . . .	82
5.2	Générateur et équations de Kolmogorov . . . . .	83
5.3	Loi stationnaire et théorèmes ergodiques . . . . .	85
5.4	Quelques classes remarquables de processus de Markov . . . . .	85
5.4.1	Processus de Poisson . . . . .	85
5.4.2	Processus de naissance . . . . .	86
5.4.3	Processus de mort . . . . .	88
5.4.4	Processus de naissance et de mort . . . . .	89
5.5	Simulation de Monte Carlo . . . . .	92
5.6	Problèmes . . . . .	93
<b>6</b>	<b>Processus de renouvellement</b>	<b>94</b>
6.1	Introduction . . . . .	94
6.2	Equation de renouvellement . . . . .	95
6.3	Théorèmes limites . . . . .	96
6.4	Processus de renouvellement modifié . . . . .	97
6.5	Temps d'attente . . . . .	97
6.6	Processus de renouvellement arrêté . . . . .	98
6.7	Processus de renouvellement alterné . . . . .	99
6.8	Problèmes . . . . .	100
<b>7</b>	<b>Applications</b>	<b>101</b>
7.1	Files d'attente . . . . .	101
7.1.1	Introduction . . . . .	101
7.1.2	Formule de Little . . . . .	102
7.1.3	File M/D/1 : Temps discret . . . . .	103
7.1.4	File M/M/1 . . . . .	104
7.1.5	File M/M/ $\infty$ . . . . .	105
7.1.6	File M/M/m/m . . . . .	106
7.1.7	File M/M/m/ $\infty$ . . . . .	106
7.1.8	File M/M/m/K . . . . .	107
7.2	Fiabilité . . . . .	107
7.2.1	Introduction . . . . .	107
7.2.2	Disponibilité . . . . .	108
7.2.3	Fiabilité . . . . .	108
7.2.4	Fiabilité en temps discret . . . . .	109
7.3	Evolution de populations et modèles de branchement . . . . .	111
7.3.1	Introduction . . . . .	111
7.3.2	Moments et probabilité d'extinction . . . . .	112
7.3.3	Modèles de branchement à temps continu . . . . .	113
7.4	Problèmes . . . . .	114
	<b>Bibliographie</b>	<b>118</b>

# Chapitre 1

## Rappels de probabilités

### 1.1 Événements et probabilité

Il y a beaucoup d'expériences (*lancement d'une pièce de monnaie, jeu de dé, jeu de roulette etc.*) dont le résultat ne peut pas être prévu avec certitude et pour lesquelles on peut seulement parler de probabilité de telle ou telle issue. On dit que le résultat d'une telle expérience est aléatoire. L'événement élémentaire se produit un et un seul chaque fois et qui peut être aussi bien le résultat d'une expérience que l'on a réalisée ou l'observation d'un phénomène dont on ne maîtrise pas la production, que, avec la formulation de type statistique, le tirage d'un individu dans une population. Un événement est un ensemble d'événements élémentaires. Il est bien sûr permis qu'un événement soit constitué par un événement élémentaire tout seul.

Cas remarquables : l'événement impossible (représenté par l'ensemble vide) "rien ne s'est passé" ou "aucun résultat" qui est donc impossible (et sa probabilité sera 0), et l'événement certain "n'importe quoi s'est passé" ou "tout résultat possible" qui est donc certain (et sa probabilité sera 1).

L'ensemble de tous les événements élémentaires possibles est appelé *l'espace de résultats* ou *l'espace fondamental*.

Les notations les plus fréquentes désignent l'événement certain ou *l'espace fondamental* par  $\Omega$ , l'événement élémentaire générique par  $\omega$  (on écrira  $\omega \in \Omega$ ), les événements (élémentaires ou non) par des lettres majuscules  $A, B, C$  etc. Pour exprimer qu'un événement, par exemple  $B$ , s'est ou ne s'est pas réalisé, on écrira  $\omega \in B$  ou  $\omega \notin B$ .

Aussi bien pour travailler sur la situation concrète que l'on modélise, que pour calculer des probabilités, on doit en permanence combiner des événements. Nous rappelons : a) le complémentaire  $\complement A$  ou  $A^c$  correspond à la *négation* logique "non- $A$ "; b) la réunion  $A \cup B$  correspond à la *disjonction* logique " $A$  ou  $B$ "; c) l'intersection  $A \cap B$  correspond à la *conjonction* logique " $A$  et  $B$ ". Si  $A \cap B = \emptyset$  on dit que les événements  $A$  et  $B$  sont *disjoints* ou *incompatibles*; d) on notera enfin que *l'implication* logique " $A \Rightarrow B$ " se traduit par l'inclusion  $A \subset B$ .

▷ **Exemple 1.1** Pour le jeu de pile ou face, on pourra considérer les espaces fondamentaux suivants :

- pour un lancer,  $\Omega = \{P, F\}$ ;
- pour deux lancers,  $\Omega = \{(P, P), (P, F), (F, P), (F, F)\}$ ;
- pour  $n$  lancers,  $\Omega = \{(u_1, \dots, u_n) : u_i \in \{P, F\}\}$ ;
- pour un nombre infini de lancers,  $\Omega = \{P, F\}^{\mathbb{N}}$ .

◁

▷ **Exemple 1.2** L'espace fondamental dans l'étude de la durée de vie d'un appareil est  $\Omega = \mathbb{N}$  ou bien  $\Omega = \mathbb{R}_+$ , selon que le temps est compté de façon discrète ou continue.

◁

Tout événement  $A$  défini sur une expérience aléatoire d'espace fondamental  $\Omega$  est un sous-ensemble de  $\Omega$ , mais tout sous-ensemble de  $\Omega$  n'est pas a priori un événement. Il faut en outre qu'il fasse partie d'un "catalogue" d'événements, appelé *tribu* ou  *$\sigma$ -algèbre d'événements*.

**Définition 1.1.** Soit  $\Omega$  un ensemble.

1. Une famille  $\mathcal{A}$  de parties de  $\Omega$  est appelée *algèbre (de Boole)* si elle contient  $\Omega$  et elle est stable par passage au complémentaire et par la réunion finie, soit telle que
  - a.  $\Omega \in \mathcal{A}$ ;
  - b. si  $A \in \mathcal{A}$ , alors  $A^c \in \mathcal{A}$ ;
  - c. si  $A \in \mathcal{A}$  et si  $B \in \mathcal{A}$ , alors  $A \cup B \in \mathcal{A}$ .
2. Une famille  $\mathcal{C}$  de parties de  $\Omega$  est appelée *classe monotone* si elle contient  $\Omega$  et est stable par réunion (resp. intersection) monotone, soit si pour toute suite croissante (resp. décroissante)  $(A_n)$  d'éléments de  $\mathcal{C}$ , on a  $\bigcup_{n \geq 1} A_n \in \mathcal{C}$  (resp.  $\bigcap_{n \geq 1} A_n \in \mathcal{C}$ ).
3. Une famille  $\mathcal{F}$  de parties de  $\Omega$  est appelée *tribu* ou  *$\sigma$ -algèbre* sur  $\Omega$  si elle contient  $\Omega$  et est stable par passage au complémentaire et par réunion dénombrable, soit telle que
  - i.  $\Omega \in \mathcal{F}$ ;
  - ii. si  $A \in \mathcal{F}$ , alors  $A^c \in \mathcal{F}$ ;
  - iii. si  $A_n \in \mathcal{F}$  pour  $n \geq 1$ , alors  $\bigcup_{n \geq 1} A_n \in \mathcal{F}$ .

Le couple  $(\Omega, \mathcal{F})$  est appelé *espace mesurable* et les éléments de  $\mathcal{F}$  sont les ensembles mesurables. On peut établir une correspondance entre l'espace fondamental et les événements, d'une part, et un espace mesurable et les ensembles mesurables, d'autre part.

- ▷ **Exemple 1.3** 1. L'ensemble des parties de  $\Omega$ , soit  $\mathcal{P}(\Omega)$ , est la plus grande tribu possible sur  $\Omega$ , au sens de l'inclusion. De même, la tribu grossière  $\{\emptyset, \Omega\}$  est la plus petite.
2. L'ensemble  $\mathcal{F}_A = \{A \cap B : B \in \mathcal{F}\}$  est une tribu pour tout  $A \in \mathcal{P}(\Omega)$ .

◁

Toute intersection de tribus est une tribu. Comme  $\mathcal{P}(\Omega)$  est une tribu, on peut aussi définir les tribus engendrées.

**Définition 1.2.** Soit  $\mathcal{H} \subset \mathcal{P}(\Omega)$ . La tribu engendrée par  $\mathcal{H}$  sur  $\Omega$  (désignée  $\sigma(\mathcal{H})$ ) est l'intersection de toutes les tribus sur  $\Omega$  qui contiennent  $\mathcal{H}$ , soit

- $\sigma(\mathcal{H})$  est une tribu sur  $\Omega$ ;
- $\mathcal{H} \subset \sigma(\mathcal{H})$ ;
- si  $\mathcal{E}$  est une tribu quelconque telle que  $\mathcal{E} \supset \mathcal{H}$ , alors  $\mathcal{E} \supset \sigma(\mathcal{H})$ .

- ▷ **Exemple 1.4** Soit  $A \in \mathcal{P}(\Omega)$  et  $\mathcal{H} = \{A\}$ . Alors  $\sigma(\mathcal{H}) = \{\emptyset, A, A^c, \Omega\}$ .

◁

La tribu engendrée par la famille d'ouvert  $\mathcal{O}$  sur  $\mathbb{R}$  est appelée *tribu borélienne* de  $\mathbb{R}$  et désignée par  $\mathcal{B}(\mathbb{R})$ . La tribu borélienne de  $\mathbb{R}$  est engendrée par chacune des familles d'ensembles de  $\mathbb{R}$  suivantes :

- |  |  |
|--|--|
| 1. $\{(a, b) : a, b \in \mathbb{R}, a < b\}$ ; | 2. $\{(a, b] : a, b \in \mathbb{R}, a < b\}$ ; |
| 3. $\{[a, b) : a, b \in \mathbb{R}, a < b\}$ ; | 4. $\{[a, b] : a, b \in \mathbb{R}, a < b\}$ ; |
| 5. $\{(-\infty, b) : b \in \mathbb{R}\}$ ;     | 6. $\{(-\infty, b] : b \in \mathbb{R}\}$ ;     |
| 7. $\{(a, \infty) : a \in \mathbb{R}\}$ ;      | 8. $\{[a, \infty) : a \in \mathbb{R}\}$ .      |

De même, la tribu borélienne sur  $\mathbb{R}^d$ , désignée  $\mathcal{B}(\mathbb{R}^d)$ , est engendrée par les pavés, les boules, les pavés à sommets rationnels etc.

**Théorème 1.1.** (Théorème de la classe monotone) Soient  $\mathcal{C}$  une classe monotone et  $\mathcal{A}$  une algèbre. Si  $\mathcal{A} \subset \mathcal{C}$ , alors  $\sigma(\mathcal{A}) \subset \mathcal{C}$ .

**Définition 1.3.** Soit  $(\Omega, \mathcal{F})$  un espace mesurable. Une fonction  $\mu: \mathcal{F} \rightarrow \overline{\mathbb{R}}_+$  est une mesure positive sur  $(\Omega, \mathcal{F})$  si elle vérifie l'axiome de  $\sigma$ -additivité,

$$\mu\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mu(A_n), \quad A_n \in \mathcal{F}, \quad A_i \cap A_j = \emptyset, i \neq j,$$

et si  $\mu(\emptyset) = 0$ .

La mesure  $\mu$  sur un espace mesurable  $(\Omega, \mathcal{F})$  est dite :

- a) *continue* ou *diffuse*, si pour tout  $\omega \in \Omega$ , et  $\{\omega\} \in \mathcal{F}$  on a  $\mu(\{\omega\}) = 0$ ;
- b) *discrète*, si  $\mu(\Omega \setminus A) = 0$  pour un événement  $A \in \mathcal{F}$  fini ou dénombrable;
- c) *finie* si  $\mu(\Omega) < \infty$ . En particulier, si  $\mu(\Omega) = 1$  la mesure  $\mu$  est une probabilité sur  $(\Omega, \mathcal{F})$  et sera désignée par  $\mathbb{P}$ . Le triplet  $(\Omega, \mathcal{F}, \mathbb{P})$  est dit *espace de probabilité*.
- d)  $\sigma$ -*finie* s'il y a une suite  $\Omega_n$  d'ensembles de  $\mathcal{F}$  tel que

$$\bigcup_{n \geq 1} \Omega_n = \Omega, \quad \mu(\Omega_n) < \infty, n = 1, 2, \dots$$

**Proposition 1.2.** Soit  $\mathcal{A}$  une algèbre. Si deux mesures  $\sigma$ -finies sont égales sur  $\mathcal{A}$ , alors elles sont égales sur la tribu engendrée par  $\mathcal{A}$  (c'est-à-dire  $\sigma(\mathcal{A})$ ).

*Démonstration.* (Esquisse) Soient  $\mu_1$  et  $\mu_2$  les deux mesures  $\sigma$ -finies. La famille d'ensemble

$$\mathcal{C} = \{A \in \mathcal{P}(\Omega) : \mu_1(A) = \mu_2(A)\}$$

est une classe monotone qui contient  $\mathcal{A}$ . Du théorème 1.1 il en résulte  $\mathcal{C} \supset \sigma(\mathcal{A})$ . □

**Remarque 1.1.** Soit  $\mathcal{L}$  une famille stable par intersections finies (i.e.  $A, B \in \mathcal{L} \Rightarrow A \cap B \in \mathcal{L}$ ). On peut démontrer que, si deux probabilités sont égales sur  $\mathcal{L}$ , alors elles sont égales sur la tribu engendrée par  $\mathcal{L}$ .

▷ **Exemple 1.5** (*mesure de Dirac*) Pour tout  $\omega \in \Omega$ , c'est la mesure discrète, finie sur  $(\Omega, \mathcal{P}(\Omega))$ , notée  $\delta_\omega$  et définie par  $\delta_\omega(A) = \begin{cases} 1, & \text{si } \omega \in A \\ 0, & \text{sinon} \end{cases}$ . ◁

▷ **Exemple 1.6** (*mesure de comptage*) Soit  $\Omega$  un ensemble finie ou dénombrable et  $\mathcal{F} = \mathcal{P}(\Omega)$ . La mesure de comptage  $\mu$  est définie sur  $(\Omega, \mathcal{F})$ , à valeurs dans  $\overline{\mathbb{N}}$  par  $\mu(A) = \sum_{\omega \in A} \delta_\omega(A)$ , pour tout  $A \in \mathcal{F}$ . Elle est finie si le cardinal de  $\Omega$  est fini. Dans ce cas on peut définir une probabilité  $\mathbb{P}$  sur  $\mathcal{F}$  par

$$\mathbb{P}(A) = \frac{\mu(A)}{\mu(\Omega)} = \frac{\text{nombre de cas favorables au événement } A}{\text{nombre de cas possibles}} \quad (1.1)$$

appelée *probabilité uniforme* sur  $\Omega$ . ◁

▷ **Exemple 1.7** (*mesure de Lebesgue*) Grâce au théorème de la classe monotone, on peut montrer l'existence et l'unicité d'une mesure  $\lambda$  sur  $(\mathbb{R}, \mathcal{B})$  telle que

$$\lambda([a, b]) = \lambda([a, b)) = \lambda((a, b]) = \lambda((a, b)) = b - a, \quad \text{pour tous } a < b.$$

Elle généralise la notion de la longueur. Elle est continue et  $\sigma$ -finie; de plus elle est invariante par translation.

◁

▷ **Exemple 1.8** (*mesure de Lebesgue-Stieltjes*) Soit  $F: \mathbb{R} \rightarrow \mathbb{R}$  une fonction croissante et continue à droite ( $F$  est dite *fonction de masse*). En utilisant le théorème de la classe monotone, on peut montrer l'existence et l'unicité d'une mesure  $\lambda_F$  sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  telle que

$$\lambda_F((a, b]) = F(b) - F(a), \text{ pour tous } a < b.$$

Cette mesure est appelée mesure Lebesgue-Stieltjes correspondant à  $F$ .

◁

**Remarque 1.2.** 1. Si  $F(x) = x$ , pour tout  $x \in \mathbb{R}$ , alors  $\lambda_F$  est la mesure de Lebesgue.  
2. Pour tout  $x \in \mathbb{R}$  on a

$$\lambda_F(\{x\}) = \lim_{n \rightarrow \infty} \lambda_F((x - 1/n, x]) = \lim_{n \rightarrow \infty} [F(x) - F(x - 1/n)] = F(x) - F(x-).$$

Donc  $\lambda_F(\{x\}) = 0$  si et seulement si  $F$  est continue en  $x$ .

3. On peut définir la mesure de Lebesgue (ou des mesures Lebesgue-Stieltjes sur  $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  de manière similaire.

Si  $\mathbb{P}$  est une probabilité sur l'espace mesurable  $(\Omega, \mathcal{F})$ , alors le triplet  $(\Omega, \mathcal{F}, \mathbb{P})$  est dit *espace de probabilité* ou *espace probabilisé*.

## 1.2 Indépendance et conditionnement

Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $A$  un événement de  $\mathcal{F}$  non négligeable (c'est-à-dire  $\mathbb{P}(A) \neq 0$ ). L'application  $\mathbb{P}(\cdot | A)$  définie sur  $\mathcal{F}$  par

$$\mathbb{P}(B | A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}, \quad B \in \mathcal{F}$$

est une probabilité sur  $\mathcal{F}$  appelée probabilité conditionnelle à  $A$  (ou sachant  $A$ ). La probabilité de  $B$  sachant  $A$ , soit  $\mathbb{P}(B | A)$ , est la probabilité que  $B$  se réalise lorsque l'on sait que  $A$  est réalisé.

Deux événements  $A, B \in \mathcal{F}$  sont dits indépendants si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B) \quad (1.2)$$

Si par exemple  $\mathbb{P}(A) \neq 0$ , cette définition équivaut à affirmer que la probabilité conditionnelle  $\mathbb{P}(B | A) = \mathbb{P}(B)$ . On voit bien ici que  $A$  et  $B$  sont indépendants si et seulement si la probabilité de réalisation de  $A$  ne dépend pas de la réalisation ou non-réalisation de  $B$ .

**Remarque 1.3.** On notera que deux événements non négligeables incompatibles ne sont pas indépendants. De plus, la notion d'indépendance dépend de  $\mathbb{P}$ , contrairement à la notion d'incompatibilité : deux événements peuvent être indépendants pour une probabilité  $\mathbb{P}$  et ne pas l'être pour une probabilité  $\mathbb{P}'$ .

La notion d'indépendance peut être élargie comme suit.

**Définition 1.4.** 1. Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $\mathcal{L}_1, \dots, \mathcal{L}_n \subset \mathcal{F}$  de familles d'événements. Elles sont dites indépendantes (dans leur ensemble) si

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \cdots \mathbb{P}(A_n), \quad A_1 \in \mathcal{L}_1, \dots, A_n \in \mathcal{L}_n. \quad (1.3)$$

2. Les événements de la famille  $(A_i, i \in I) \subset \mathcal{F}$  sont dites indépendants (dans leur ensemble ou mutuellement) si, pour toute sous-famille finie  $J \subset I$  on a

$$\mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \mathbb{P}(A_i) \quad (1.4)$$



3. Soient  $B$  un événement non négligeable de  $\mathcal{F}$  et  $\mathcal{L}_1, \dots, \mathcal{L}_n \subset \mathcal{F}$  de familles d'événements. Elles sont dites indépendantes conditionnellement à  $B$  si elles sont indépendantes pour la probabilité  $\mathbb{P}(\cdot \mid B)$ , soit si

$$\mathbb{P}(A_1 \cap \dots \cap A_n \mid B) = \mathbb{P}(A_1 \mid B) \cdots \mathbb{P}(A_n \mid B), \quad A_i \in \mathcal{L}_i, i = 1, \dots, n. \quad (1.5)$$

Si les événements  $(A_i, i \in I) \subset \mathcal{F}$  sont mutuellement indépendants, alors toutes sous-familles constituées de certains de ces événements sont indépendants aussi. D'autre part, l'indépendance deux à deux n'entraîne pas l'indépendance dans leur ensemble, comme le confirme le contre-exemple suivant.

▷ **Exemple 1.9** (*Bernstein*) Soit  $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$  muni de la tribu  $\mathcal{F} = \mathcal{P}(\Omega)$  et de la probabilité uniforme  $\mathbb{P}$ , soit  $\mathbb{P}(\{\omega_i\}) = \frac{1}{4}$ ,  $i = 1, 2, 3, 4$ . Considérons les trois événements  $A_1 = \{\omega_1, \omega_2\}$ ,  $A_2 = \{\omega_1, \omega_3\}$ ,  $A_3 = \{\omega_1, \omega_4\}$ . On a  $\mathbb{P}(A_i) = \frac{1}{2}$ ,  $i = 1, 2, 3$ . Ces événements sont indépendants deux à deux parce que

$$\begin{aligned} \mathbb{P}(A_1 \cap A_2) &= \mathbb{P}(\{\omega_1\}) = \mathbb{P}(A_1) \mathbb{P}(A_2) = \frac{1}{4}, \\ \mathbb{P}(A_1 \cap A_3) &= \mathbb{P}(\{\omega_1\}) = \mathbb{P}(A_1) \mathbb{P}(A_3) = \frac{1}{4}, \\ \mathbb{P}(A_2 \cap A_3) &= \mathbb{P}(\{\omega_1\}) = \mathbb{P}(A_2) \mathbb{P}(A_3) = \frac{1}{4}. \end{aligned}$$

D'autre côté,

$$\mathbb{P}(A_1 \cap A_2 \cap A_3) = \mathbb{P}(\{\omega_1\}) = \frac{1}{4} \neq \mathbb{P}(A_1) \mathbb{P}(A_2) \mathbb{P}(A_3) = \frac{1}{8}$$

donc les trois événements ne sont pas mutuellement indépendants.

◁

Souvent nous sommes intéressé de conditions sur lesquelles l'indépendance de familles d'événements entraîne l'indépendance de tribus engendrées par eux.

**Proposition 1.3.** *Soient deux familles d'événements indépendantes  $\mathcal{L}_1$  et  $\mathcal{L}_2$  inclus dans  $\mathcal{F}$  et stables par intersections finies. Alors les tribus engendrées  $\sigma(\mathcal{L}_1)$  et  $\sigma(\mathcal{L}_2)$  sont indépendantes.*

*Démonstration.* (esquisse) Soit  $A \in \mathcal{L}_1$  quelconque non négligeable. La probabilité  $\mathbb{P}_1(\cdot) = \mathbb{P}(\cdot \mid A)$  est égale avec  $\mathbb{P}(\cdot)$  sur  $\mathcal{L}_2$  parce que

$$\mathbb{P}_1(B) = \mathbb{P}(B \mid A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B)\mathbb{P}(A)}{\mathbb{P}(A)} = \mathbb{P}(B), \quad B \in \mathcal{L}_2.$$

De la remarque 1.1 il en résulte que  $\mathbb{P}_1(B) = \mathbb{P}(B)$  pour tout  $B \in \sigma(\mathcal{L}_2)$ ; donc

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}(B), \quad A \in \mathcal{L}_1, \quad B \in \mathcal{L}_2.$$

D'une manière similaire on déduit que la relation ci-dessus est vérifiée pour tout  $A \in \sigma(\mathcal{L}_1)$ .  $\square$

**Proposition 1.4** (Cinq formules remarquables).

1. (*Formule de probabilités totales*). Soit  $(B_n, n \in I)$  un système exhaustif d'événements, avec  $\mathbb{P}(B_n) > 0$ , et  $A$  un événement. Nous avons

$$\mathbb{P}(A) = \sum_{n \in I} \mathbb{P}(A \cap B_n) = \sum_{n \in I} \mathbb{P}(A \mid B_n) \mathbb{P}(B_n)$$

2. (Formule de Poincaré ou d'inclusion-exclusion). Pour une suite finie d'événements :  $A_1, A_2, \dots, A_n$ , nous avons

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n \mathbb{P}(A_i) - \sum_{1 \leq i < j \leq n} \mathbb{P}(A_i \cap A_j) + \dots + (-1)^{n-1} \mathbb{P}(\cap_{i=1}^n A_i)$$

3. (Formule de probabilités produit). Pour une suite finie d'événements :  $A_1, A_2, \dots, A_n$ , nous avons

$$\mathbb{P}\left(\bigcap_{i=1}^n A_i\right) = \mathbb{P}(A_1) \mathbb{P}(A_2 | A_1) \mathbb{P}(A_3 | A_1 \cap A_2) \times \dots \times \mathbb{P}(A_n | A_1 \cap \dots \cap A_{n-1}).$$

4. (Formule de produits disjoints). Pour une suite finie d'événements :  $A_1, A_2, \dots, A_n$ , nous avons

$$\mathbb{P}(\cup_{i=1}^n A_i) = \mathbb{P}(A_1) + \mathbb{P}(A_1^c A_2) + \mathbb{P}(A_1^c A_2^c A_3) + \dots + \mathbb{P}(A_1^c A_2^c \dots A_{n-1}^c A_n)$$

5. (Formule de Bayes). Soient  $(B_n, n \in I)$  un système d'événements exhaustif tel que  $\mathbb{P}(B_k) > 0$ , pour tout  $k \in I$  et un événement  $A$  tel que  $\mathbb{P}(A) > 0$ . Alors

$$\mathbb{P}(B_j | A) = \frac{\mathbb{P}(A | B_j) \mathbb{P}(B_j)}{\sum_{k \in I} \mathbb{P}(A | B_k) \mathbb{P}(B_k)}$$

▷ **Exemple 1.10** Soient  $A, B \in \mathcal{F}$  deux événements indépendants. D'après la proposition 1.3 les tribus engendrées par les famille  $\mathcal{L}_1 = \{A\}$  et  $\mathcal{L}_2 = \{B\}$  sont indépendantes. En conséquence, dans chacune des paires  $(A, B^c)$ ,  $(A^c, B)$ ,  $(A^c, B^c)$  les événements sont indépendants.

◁

▷ **Exemple 1.11** (Date d'anniversaire) Combien de personnes choisies au hasard doivent être interrogées sur leur date d'anniversaire pour qu'on trouve au moins deux personnes nées au même jour de l'année. Plus précisément, si on interroge 367 personnes, alors on trouve à coup sûr (avec la probabilité 1) au moins deux personnes nées dans le même jour. Dans ce problème on demande le nombre de personnes pour que cette probabilités soit justement  $> \frac{1}{2}$ .

Ce problème est un cas particulier du problème plus générale connu sous le nom *la distribution de  $r$  boules dans  $n$  urnes*. Ici les jour de l'année sont des urnes  $n = 366$  et les  $r$  personnes interrogées sont les boules. Le nombre de cas possible est  $n^r$ . Si toutes les distributions des boules dans les urnes ont la même probabilité on obtient la dite *statistique de Maxwell-Boltzmann*. Dans cette hypothèse la probabilité d'avoir  $r_1$  boules dans la première urne,  $r_2$  boule dans la deuxième etc. où  $r_j \geq 0$ ,  $j = 1, \dots, n$ ,  $r_1 + \dots + r_n = r$ , est

$$p(r_1, \dots, r_n) = \frac{r!}{r_1! \dots r_n!} n^{-r} \quad (1.6)$$

Pour notre problème, après quelques calculs (pas trop simples !) on trouve le nombre de personnes  $r \cong 21$ . Ce résultat est surprenant car nous nous attendions à une valeur  $n \cong 180$ . Ici l'intuition ne s'accorde pas à la réalité !

◁

▷ **Exemple 1.12** (Problème des rencontres) On demande à quatre personnes d'écrire leur nom sur un papier, puis on mélange les papiers et on les redistribue au hasard. Quelle est la probabilité qu'au moins une personne ait le papier avec son nom ?

Soient les événements  $A_i$  – “la personne  $i$  a reçu le papier avec son nom”,  $i = 1, 2, 3, 4$ . La probabilité cherchée est (on utilise la formule de Poincaré)

$$\begin{aligned}\mathbb{P}\left(\bigcup_{i=1}^4 A_i\right) &= \sum_{i=1}^4 \mathbb{P}(A_i) - \sum_{i < j} \mathbb{P}(A_i \cap A_j) + \sum_{i < j < k} \mathbb{P}(A_i \cap A_j \cap A_k) - \mathbb{P}\left(\bigcap_{i=1}^4 A_i\right) \\ &= 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} = \frac{5}{8}\end{aligned}$$

Ce problème peut être généralisé pour  $n$  personnes d’une manière évidente.

◁

▷ **Exemple 1.13** On constitue une file d’attente en attribuant au hasard des numéros d’ordre à  $n$  personnes. Quelle est la probabilité que deux amis soient distants de  $r$  places (i.e. séparés par  $r - 1$  personnes)? Quelle est la distance la plus probable entre ces deux amis?

L’espace fondamental est l’ensemble des permutations de  $n$  éléments, dont le nombre est  $N = n!$ . L’ensemble des cas favorables est constitué en composant les positions des deux amis à  $r$  places de distance

$$\{(1, r+1), (2, r+2), \dots, (n-r, n)\} \cup \{(r+1, 1), (r+2, 2), \dots, (n, n-r)\}$$

qui possède  $2(n-r)$  éléments, avec les positions des  $n-2$  autres personnes qui possède  $(n-2)!$  éléments. Donc on a  $k = 2(n-r)(n-2)!$  cas favorables, d’où la probabilité cherchée est (voir (1.1))

$$p = \frac{k}{N} = \frac{2(n-r)(n-2)!}{n!} = \frac{2(n-r)}{n(n-1)}.$$

Le maximum de la probabilité se produit pour  $r = 1$  (avec  $p = \frac{2}{n}$ ), c’est-à-dire lorsque les deux amis sont l’un derrière l’autre. Ce résultat est tout à fait contraire à l’intuition la plus courante, qui imagine une distance moyenne voisine de  $\frac{n}{2}$ , et qui conclut imprudemment que cette valeur est aussi la plus probable.

◁

▷ **Exemple 1.14** (*Problème de la ruine du joueur*) Soit le jeu de pile ou face. Le joueur gagne si sort le côté qu’il avait choisi et perd dans le cas contraire avec la probabilité  $\frac{1}{2}$  pour chaque cas. Dans un coup il peut gagner ou perdre 1 € par exemple. Supposons que le joueur entre en jeu avec le capital initial de  $x$  € et qu’il propose de le porter à  $a > x$  €. Le jeu se termine lorsque le joueur ou bien gagne cette somme  $a$  ou bien perd tout son capital. La probabilité que le jeu se termine par une perte totale du capital dépend du capital initial  $x$  et de la somme finale  $a$ . Désignons par  $p(x)$  la probabilité pour que, ayant  $x$  €, le joueur soit ruiné. La formule des probabilités totales donne

$$p(x) = \frac{1}{2}[p(x+1) + p(x-1)] \quad (1.7)$$

pour tous les  $x = 1, \dots, a-1$  (il est évident qu’il faut poser  $p(0) = 1$ , et  $p(a) = 0$ ). L’équation (1.7) ne peut y avoir qu’une seule solution pour des conditions initiales  $p_0 = p(0)$ ,  $p_1 = p(1)$ . Il est facile de vérifier que cette solution est de la forme

$$p(x) = p_0 + (p_1 - p_0)x$$

et pour  $p_0 = 1$  on obtient

$$p(x) = 1 - (1 - p(1))x,$$

d’où, comme  $p(a) = 1 - (1 - p(1))a = 0$ , on trouve  $p(1) = 1 - \frac{1}{a}$  et finalement

$$p(x) = 1 - \frac{x}{a}, \quad x = 0, 1, \dots, a.$$

◁

### 1.3 Variables aléatoires

#### 1.3.1 Définition, Loi et Espérance

Dans la plupart des phénomènes aléatoires, le résultat d'une épreuve peut se traduire par une "grandeur" mathématique, très souvent représentée par un nombre réel, par exemple :

- le nombre total de piles pour  $n$  jeux de pile ou face ;
- le nombre d'appels téléphoniques parvenant à un standard durant un intervalle de temps fixé ;
- le temps de désintégration d'un atome radioactif ;
- la proportion de pièces défectueuses fabriquées par une machine ;
- etc.

Dans le cadre de tel ou tel modèle de la théorie des probabilités décrit par l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ , lorsque l'on envisage une variable aléatoire réelle  $X$ , on suppose qu'elle dépend du hasard (plus exactement de l'issue élémentaire  $\omega \in \Omega$ ) de telle sorte que les probabilités de tous les événements du type  $\{\omega \in \Omega : a \leq X(\omega) \leq b\}$  soient déterminées. Dans les problèmes de la théorie des probabilités il est peu important en général de connaître la dépendance explicite de  $X = X(\omega)$  de  $\omega \in \Omega$ . Ce qui est important, ce sont les probabilités du type  $\mathbb{P}(a \leq X \leq b)$  qui caractérisent la dépendance de la variable  $X$  du hasard ; dans leur ensemble elles montrent la répartition de la probabilité pour le point aléatoire  $X$  de tomber dans tel ou tel intervalle  $[a, b]$ , autrement dit, elles donnent la loi de probabilité de la variable aléatoire  $X$ .

**Définition 1.5.** 1. Soient  $\Omega$  un ensemble,  $(E, \mathcal{E})$  un espace mesurable et  $X: \Omega \rightarrow E$  une application. L'ensemble  $X^{-1}(\mathcal{E})$  est un tribu sur  $\Omega$ , appelée *tribu engendrée par  $X$*  et notée  $\sigma(X)$ , soit

$$\sigma(X) = \{A \in \mathcal{P}(\Omega) : \exists A' \in \mathcal{E} \text{ tel que } A = X^{-1}(A')\}.$$

2. Soient  $(\Omega, \mathcal{F})$  et  $(E, \mathcal{E})$  deux espaces mesurables. Une application  $X: \Omega \rightarrow E$  est dite  $(\mathcal{F}, \mathcal{E})$ -mesurable (ou  $\mathcal{F}$ -mesurable) si  $\sigma(X) \subset \mathcal{F}$ . Si  $(E, \mathcal{E}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ , nous dirons que  $X$  est (une fonction) borélienne.

En probabilités, une application mesurable définie sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  est appelée variable aléatoire, réelle si elle est à valeurs réelles. Souvent on considère de v.a. à valeurs dans  $\overline{\mathbb{R}}$ .

Plus généralement, toute fonction borélienne

$$X = (X_1, \dots, X_d): (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$$

est appelée *d-vecteur aléatoire réel* (ou *variable aléatoire d-dimensionnelle réelle*) si  $d > 1$ .

▷ **Exemple 1.15** La fonction indicatrice d'une partie  $A$  de  $\Omega$  est définie par

$$\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \Omega \setminus A. \end{cases}$$

Une fonction  $X$  est dite *étagée* si

$$X = \sum_{i=1}^n \alpha_i \mathbb{1}_{A_i}, \quad \alpha_i \in \mathbb{R}^*, \quad A_i \subset \Omega.$$

Cette écriture est unique si l'on impose (pour une fonction non identiquement nulle) que  $A_i \cap A_j = \emptyset$  et que  $\alpha_i \neq \alpha_j$  pour  $i \neq j$ .

**Remarque 1.4.** Une fonction étagée est borélienne si et seulement si  $A_i \in \mathcal{F}$  pour tout  $i = 1, \dots, n$ .

Une v.a. réelle sur l'espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  définit une probabilité  $P_X = \mathbb{P} \circ X^{-1}$  sur  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  qui est appelée *loi de probabilité* de  $X$ , c'est-à-dire que

$$P_X(B) = \mathbb{P}[X^{-1}(B)] = \mathbb{P}(X \in B), \quad B \in \mathcal{B}(\mathbb{R}) \quad (1.8)$$

La fonction  $F_X: \mathbb{R} \rightarrow [0, 1]$ , définie par

$$F_X(x) = P_X((-\infty, x]) = \mathbb{P}(X \leq x), \quad x \in \mathbb{R} \quad (1.9)$$

est appelée *fonction de répartition* de  $X$ . On constate immédiatement que  $F_X$  est croissante et continue à droite. De plus,

$$\lim_{x \rightarrow -\infty} F_X(x) = \mathbb{P}(\emptyset) = 0 \text{ et } \lim_{x \rightarrow \infty} F_X(x) = \mathbb{P}(\Omega) = 1.$$

Soient deux réels  $a < b$  et  $F$  la fonction de répartition de la v.a.  $X$ . On a

$$\begin{aligned} P_X(]a, b]) &= \mathbb{P}(a < X \leq b) = F(b) - F(a); & F(a+) &= F(a); \\ P_X([a, b]) &= \mathbb{P}(a \leq X < b) = F(b-) - F(a-); & \mathbb{P}(X = a) &= F(a) - F(a-); \\ P_X(]a, b[) &= \mathbb{P}(a < X < b) = F(b-) - F(a); & P_X(]-\infty, b]) &= \mathbb{P}(X \leq b) = F(b); \\ P_X([a, b]) &= \mathbb{P}(a \leq X \leq b) = F(b) - F(a-); & P_X([a, \infty[) &= \mathbb{P}(X \geq a) = 1 - F(a-). \end{aligned} \quad (1.10)$$

Comme  $]a, a[ = \emptyset$ , on a  $P_X(]a, a]) = 0$ . Même chose également pour  $[a, a[ = \emptyset$  et  $]a, a[ = \emptyset$ .

On appelle v.a. discrète une variable qui ne prend qu'un nombre fini ou dénombrable de valeurs. Une v.a. discrète est caractérisé par l'ensemble des valeurs qu'elle peut prendre, soit  $x_1, x_2, \dots$ , et par l'expression mathématique de la probabilité de ces valeurs. Cette expression mathématique s'appelle la *loi de probabilité* (ou *distribution de probabilité*) de la v.a., et elle est donnée, soit par une formule générale, soit par la liste de ces probabilités  $p_1, p_2, \dots$ . D'habitude la loi d'une v.a. discrète  $X$  est écrit sous la forme

$$X : \begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix} \quad (1.11)$$

où on suppose  $x_i \neq x_j$  pour  $i \neq j$ . Donc

$$p_i = \mathbb{P}(X = x_i), \quad i = 1, 2, \dots \text{ et } \sum_i p_i = 1. \quad (1.12)$$

▷ **Exemple 1.16** On considère le lancer de trois pièces : l'espace fondamental  $\Omega$  comprend 8 événements élémentaires, de probabilité chacun  $\frac{1}{8}$ . On introduit une v.a.  $X: \Omega \rightarrow \mathbb{R}$ , définie par  $X(\omega)$  = nombre de piles de l'événement  $\omega$ . On a

$$X : \begin{pmatrix} 0 & 1 & 2 & 3 \\ \frac{1}{8} & \frac{3}{8} & \frac{3}{8} & \frac{1}{8} \end{pmatrix}$$

La fonction de répartition est

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{1}{2}, & 1 \leq x < 2 \\ \frac{7}{8}, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

&lt;

On appelle v.a. *continue* une variable qui peut prendre toutes les valeurs d'un intervalle (qui peut être un intervalle borné, ou une demi-droite, ou encore  $\mathbb{R}$  tout entier), comme par exemple le temps de fonctionnement d'un composant. La mesure de probabilité est alors "étalée" sur l'intervalle de définition, avec conséquence paradoxale que tous les points ont une probabilité nulle ; en effet, il y a "trop" de points (un ensemble non dénombrable !) dans un intervalle pour que chacun d'eux puisse avoir une probabilité non nulle. Plus précisément

**Définition 1.6.** Une v.a. réelle est appelée continue si  $\mathbb{P}(X = x) = 0$  pour tout  $x \in \mathbb{R}$ .

**Proposition 1.5.** Une v.a. est continue si, et seulement si, sa fonction de répartition est continue.

*Démonstration.* C'est une conséquence de (1.10).  $\square$

**Remarque 1.5.** 1. La fonction de répartition d'une v.a. discrète est en escalier.

2. Une v.a. peut être ni discrète, ni continue (cf. exemple 1.17).

**Définition 1.7.** Soit  $X$  une v.a. réelle avec la fonction de répartition  $F$ . La v.a. est dite *absolument continue* s'il existe une fonction  $f: \mathbb{R} \rightarrow \mathbb{R}_+$  borélienne positive telle que

$$F(x) = \int_{-\infty}^x f(t) dt, \text{ pour tout } x \in \mathbb{R}. \quad (1.13)$$

La fonction  $f$  est appelée *densité (de probabilité)* de  $X$ .

On a

$$\int_{-\infty}^{\infty} f(t) dt = \lim_{x \rightarrow \infty} F(x) = 1. \quad (1.14)$$

Il est évident qu'une v.a. absolument continue est continue et la densité de probabilité coïncide en ses points de continuité avec la dérivée de la fonction de répartition, i.e.  $f(x) = F'(x)$ .

La densité de probabilité a une signification intuitive simple : si  $\delta$  est un "infinitement petit", et si la densité est continue au point  $x$ , alors la probabilité  $\mathbb{P}(x < X \leq x + \delta)$  est infinitement petit proche de  $f(x)\delta$ . En outre, la densité peut être utilisée pour donner par intégration les probabilités d'intervalles (cf. les relations (1.10))

$$\begin{aligned} \mathbb{P}(a < X \leq b) &= \mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) \\ &= \mathbb{P}(a \leq X < b) = \int_a^b f(t) dt. \end{aligned} \quad (1.15)$$

En général, on a

$$\mathbb{P}(X \in B) = \int_B f(t) dt, \quad B \in \mathcal{B}(\mathbb{R}) \quad (1.16)$$

La notion d'indépendance peut être définie pour des v.a. Ainsi les v.a.  $X$  et  $Y$  sur le même espace de probabilité sont dites indépendantes si

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \text{ pour tous } A, B \in \mathcal{B}(\mathbb{R}) \quad (1.17)$$

Plus général, les v.a.  $X_1, \dots, X_n$  sont dites indépendants (dans leur ensemble ou mutuellement) si

$$\mathbb{P}(X_1 \in B_1, \dots, X_n \in B_n) = \mathbb{P}(X_1 \in B_1) \cdots \mathbb{P}(X_n \in B_n) \quad (1.18)$$

pour tous  $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$ .

▷ **Exemple 1.17** Soient les v.a. discrètes  $X_1$  et  $\eta$  avec les lois

$$X_1 : \begin{pmatrix} x_1 & x_2 & \cdots \\ p_1 & p_2 & \cdots \end{pmatrix}; \quad \eta : \begin{pmatrix} 1 & 0 \\ q_1 & q_2 \end{pmatrix}$$

et  $X_2$  une v.a. avec la densité  $f$ . On suppose que  $\eta$  est indépendante de v.a.  $X_1$  et  $X_2$  et on considère la v.a.  $Y = \eta X_1 + (1 - \eta)X_2$ . Il est facile de voir que sous la condition  $\eta = 1$ ,  $Y$  coïncide avec  $X_1$  et sous la condition  $\eta = 0$  avec  $X_2$ . Par conséquent, on a

$$\begin{aligned}\mathbb{P}(a \leq Y \leq b \mid \eta = 1) &= \mathbb{P}(a \leq X_1 \leq b) \\ \mathbb{P}(a \leq Y \leq b \mid \eta = 0) &= \mathbb{P}(a \leq X_2 \leq b)\end{aligned}$$

d'où, en utilisant la formule de la probabilité totale, on obtient

$$\begin{aligned}\mathbb{P}(a \leq Y \leq b) &= q_1 \mathbb{P}(a \leq X_1 \leq b) + q_2 \mathbb{P}(a \leq X_2 \leq b) \\ &= q_1 \sum_{a \leq x_i \leq b} p_i + q_2 \int_a^b f(t) dt,\end{aligned}$$

et on voit que  $Y$  n'est ni discrète, ni continue. En effet, pour  $a \leq x_i \leq b$  on a  $\mathbb{P}(Y = x_i) = p_i \neq 0$  et d'autre côté l'ensemble des valeurs de  $Y$  n'est fini ou dénombrable.

◁

▷ **Exemple 1.18** (*Partie entière*) Soit  $X$  une v.a. positive. La fonction de répartition de la v.a.  $[X]$  (partie entière de  $X$ ) est

$$F_{[X]}(x) = \mathbb{P}([X] \leq x) = \mathbb{P}(X < [x] + 1) = F_X([x] + 1-), \quad x \in \mathbb{R}$$

◁

**Espérance mathématique d'une variable aléatoire.** L'espérance (mathématique), et de manière plus générale les moments, d'une variable aléatoire, jouent un rôle très important dans la théorie et la modélisation stochastique. L'intégrale de Stieltjes est l'outil principale dans le calcul de l'espérance et des moments.

La proposition suivante résume les propriétés de l'intégrale de Stieltjes utiles pour le calcul de l'espérance des v.a. mixtes.

**Proposition 1.6.** *Soit  $F$  une fonction de répartition d'une v.a. réelle  $X$  de loi  $P_X$ . Soit aussi une fonction  $h: \mathbb{R} \rightarrow \mathbb{R}$  dont l'espérance de la v.a.  $h(X)$  existe. Alors :*

1. *Pour tout intervalle  $I$  de  $\mathbb{R}$ , on a  $\int_I 1 dx = P_X(I)$ .*
2. *Pour toute constante réelle  $c$ , on a  $\int_I h dF = \int_I h d(F + c)$ .*
3. *Si  $F$  est constante sur un intervalle ouvert  $I$ , alors  $\int_I h dF = 0$ .*
4. *Pour toute fonction  $h$  définie en  $\xi$ , on a*

$$\int_{[\xi, \xi]} h dF = h(\xi)[F(\xi) - F(\xi-)].$$

5. *Si  $F$  est dérivable sur un intervalle ouvert  $I$  de  $\mathbb{R}$ , alors*

$$\int_I h dF = \int_I h F' dx,$$

*dans le sens que, si l'une de deux intégrales existe, l'autre existe également et sont égales.*

6. *Si  $I = I_1 \cup \dots \cup I_r$ , deux à deux disjoints, alors*

$$\int_I h dF = \sum_{i=1}^r \int_{I_i} h dF.$$

7. Si  $F = \sum_{i=1}^r c_i H_i$ , où les fonctions  $H_i$  sont monotones croissantes et les  $c_i$  des constantes, alors

$$\int_I h dF = \sum_{i=1}^r c_i \int_I h dH_i.$$

Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité,  $X$  une v.a. sur cet espace et  $h: \mathbb{R} \rightarrow \mathbb{R}$  une fonction borélienne.

**Définition 1.8.** On dit que  $h(X)$  possède une *espérance mathématique* ou *valeur moyenne* finie

$$\mathbb{E}[h(X)] = \int_{\Omega} h(X) d\mathbb{P} = \int_{\mathbb{R}} h(x) dP_X = \int_{\mathbb{R}} h(x) F_X(dx) = \int_{\mathbb{R}} h dF_X \quad (1.19)$$

si les intégrales dans (1.19) sont absolument convergentes, i.e.

$$\int_{\mathbb{R}} |h(x)| dP_X < \infty$$

Dans (1.19)  $P_X$  et  $F_X$  sont respectivement la loi et la fonction de répartition de  $X$ .

Il y a deux cas particuliers importants :

1. Si la v.a.  $X$  est discrète avec la loi (1.11), alors

$$\mathbb{E}[h(X)] = \sum_k h(x_k) p_k \text{ si } \sum_k |h(x_k)| p_k < \infty \quad (1.20)$$

2. Si la v.a.  $X$  possède une densité  $f$ , alors

$$\mathbb{E}[h(X)] = \int_{-\infty}^{\infty} h(x) f(x) dx \text{ si } \int_{-\infty}^{\infty} |h(x)| f(x) dx < \infty \quad (1.21)$$

Aux divers choix de la fonction  $h$  correspondent diverses caractéristiques de la v.a.  $X$ . Nous écrirons les formules en général suivant que le lecteur adapte les cas particuliers (1.20) et (1.21)

1. Pour  $h(x) = x$  on a l'*espérance mathématique* de la v.a.  $X$

$$\mathbb{E}(X) = \int_{\mathbb{R}} x F_X(dx) \quad (1.22)$$

2. Pour  $h(x) = x^k$ ,  $k \geq 1$ , on a *moment d'ordre  $k$*  de la v.a.  $X$

$$\mathbb{E}(X^k) = \int_{\mathbb{R}} x^k F_X(dx) \quad (1.23)$$

3. Pour  $h(x) = (x - \mathbb{E}(X))^2$  on a la *variance* de la v.a.  $X$

$$\text{Var}(X) = \int_{\mathbb{R}} (x - \mathbb{E}(X))^2 F_X(dx) \quad (1.24)$$

La variance peut être calculer en utilisant la formule

$$\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2 \quad (1.25)$$

La variance est un paramètre de dispersion : plus les valeurs prises par  $X$  sont groupées, plus sa variance est petite.



Si les v.a.  $X$  et  $Y$  sont indépendantes on a

$$\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y), \quad \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) \quad (1.26)$$

▷ **Exemple 1.19** (*Lancer de deux dés*) Ici  $\Omega = \{1, 2, 3, 4, 5, 6\}^2$  et  $\mathcal{F} = \mathcal{P}(\Omega)$ . Soit  $Y$  la somme des points obtenus. La loi de v.a.  $Y$  est

$$Y : \left( \begin{array}{cccccccccccc} 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ \frac{1}{36} & \frac{2}{36} & \frac{3}{36} & \frac{4}{36} & \frac{5}{36} & \frac{6}{36} & \frac{5}{36} & \frac{4}{36} & \frac{3}{36} & \frac{2}{36} & \frac{1}{36} \end{array} \right)$$

On a

$$\mathbb{E}Y = \frac{2 + 6 + 12 + 20 + 30 + 42 + 40 + 36 + 30 + 22 + 12}{36} = 7$$

◁

▷ **Exemple 1.20** Soit  $X$  une v.a. à valeurs dans  $\mathbb{N}$ . On a  $X = \sum_{n \geq 0} \mathbb{1}_{(X > n)}$ , donc

$$\mathbb{E}(X) = \mathbb{E} \left[ \sum_{n \geq 0} \mathbb{1}_{(X > n)} \right] = \sum_{n \geq 0} \mathbb{P}(X > n). \quad (1.27)$$

◁

▷ **Exemple 1.21** La fonction de répartition,  $F$ , de la durée de vie  $T$  d'un appareil est donnée par la relation :

$$F(x) = \begin{cases} 1 - \exp\{-\lambda x\} & \text{si } x \geq a, \\ 0 & \text{si } x < a, \end{cases}$$

où  $a > 0$ ,  $\lambda > 0$ .

L'espérance de la durée de vie de cet appareil se calcule comme suit.

$$\begin{aligned} \mathbb{E}(T) &= \int_{\mathbb{R}_+} x dF(x) = \int_{[0, a[} x dF + \int_{[a, a]} x dF + \int_{]a, \infty[} x dF \\ &= 0 + a[F(a) - F(a-)] + \int_{]a, \infty[} x F'(x) dx \\ &= a(1 - e^{-\lambda a}) + a e^{-\lambda a} + \frac{1}{\lambda} e^{-\lambda a} \\ &= a + \frac{1}{\lambda} e^{-\lambda a}. \end{aligned}$$

◁

**Inégalités de Markov et de Bienaymé-Tchebichev.** Ces inégalités sont d'une grande importance en probabilités tant au point de vue théorique que pratique.

**Proposition 1.7.** 1. *Inégalité de Markov. Soit une v.a.r.  $X$  possédant le moment d'ordre  $p \geq 1$ . Alors pour tout  $\varepsilon > 0$ , nous avons :*

$$P(|X| \geq \varepsilon) \leq \frac{\mathbb{E}(|X|^p)}{\varepsilon^p}. \quad (1.28)$$

2. *Inégalité de Bienaymé-Tchebichev. Soit une v.a.r.  $X$  de carré intégrable. Alors pour tout  $\varepsilon > 0$ , nous avons :*

$$P(|X - \mathbb{E}X| \geq \varepsilon) \leq \frac{\text{Var}(X)}{\varepsilon^2}. \quad (1.29)$$

### 1.3.2 Lois usuelles

#### Lois discrètes.

1. **Loi de Dirac**-Une v.a. suit une loi de Dirac si  $\mathbb{P}(X = x_0) = 1$  avec  $x_0 \in \mathbb{R}$  et nous noterons  $X \sim \delta_{x_0}$ . Si  $x_0 = 0$  on note  $X \sim \delta$ . On a

$$\mathbb{E}(X) = x_0, \text{ Var}(X) = 0.$$

2. **Loi de Bernoulli**- Une v.a.  $X$  suit une loi de Bernoulli de paramètre  $p \in [0, 1]$ , et nous noterons  $X \sim B(p)$ , si elle a la loi

$$X : \begin{pmatrix} 1 & 0 \\ p & 1-p \end{pmatrix} \quad (1.30)$$

On a

$$\mathbb{E}(X) = p, \text{ Var}(X) = p(1-p). \quad (1.31)$$

L'événement  $(X = 1)$  est dit "succès" et l'événement  $(X = 0)$  "échec" de l'épreuve.

3. **Loi binomiale**- Une v.a.  $X$  suit une loi binomiale de paramètres  $n \in \mathbb{N}^*$  et  $p \in [0, 1]$ , et nous noterons  $X \sim b(n, p)$ , si elle prend les valeurs entières  $k = 0, 1, \dots, n$  avec

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}. \quad (1.32)$$

On a

$$\mathbb{E}(X) = np, \text{ Var}(X) = np(1-p). \quad (1.33)$$

Beaucoup de situations se présentent comme le résultat global d'une succession d'épreuves partielles réitérées dans des conditions similaires. La loi binomiale correspond à la situation appelée *tirage avec remise* ou *schéma de Bernoulli* où on répète exactement à l'identique et de façon indépendante  $n$  épreuves de Bernoulli. Par exemple, on peut considérer une urne contenant des boules blanches et noires. Nous supposons que on tire une boule blanche avec la probabilité  $p$ . Si on tire  $n$  boules avec remise, alors  $X$  est le nombre de boules blanches obtenues.

4. **Loi hypergéométrique**- Une v.a.  $X$  suit une loi hypergéométrique de paramètres entiers non nuls  $N$ ,  $M$  et  $n$  (nous la noterons  $H(N, M, n)$ ) si elle prend les valeurs entières  $\max(0, n + M - N) \leq k \leq \min(M, n)$  avec

$$\mathbb{P}(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}. \quad (1.34)$$

On calcule

$$\mathbb{E}(X) = \frac{nM}{N}, \text{ Var}(X) = n \frac{M}{N} \left(1 - \frac{M}{N}\right) \frac{N-n}{N-1}. \quad (1.35)$$

La loi hypergéométrique peut être présenter comme une succession d'épreuves partielles réitérées où chaque épreuve partielle modifie le contexte et donc les probabilités relatives aux épreuves partielles suivantes. Cette expérimentation, appelée *tirage sans remise*, peut être représentée par une urne qui contient  $N$  boules,  $M$  d'une couleur et  $N - M$  d'une autre couleur. On tire  $n$  boules sans remise. La v.a.  $X$  est égale au nombre de boules de première couleur obtenues.

5. **Loi binomiale négative**- Une v.a.  $X_r$  suit une loi binomiale négative de paramètres  $r \in \mathbb{N}^*$  et  $p \in [0, 1]$  et nous la noterons  $b_-(r, p)$ , si elle prend toutes les valeurs entières  $\{r + k : k \geq 0\}$ , avec

$$\mathbb{P}(X = r + k) = \binom{r + k - 1}{r - 1} p^r (1 - p)^k = \binom{r + k - 1}{k} p^r (1 - p)^k. \quad (1.36)$$

On calcule

$$\mathbb{E}(X_r) = \frac{r}{p}, \text{ Var}(X_r) = \frac{r(1-p)}{p^2} \quad (1.37)$$

La binomiale négative peut être représentée par une urne qui contient  $N$  boules dont  $M$  blanches. On tire  $n$  boules avec remise. La v.a.  $X_r$  est égale au nombre de tirage nécessaires pour obtenir  $r$  boules blanches.

6. **Loi de Poisson**- Une v.a.  $X$  suit une loi de Poisson de paramètre  $\lambda > 0$  et nous la noterons  $P(\lambda)$ , si elle prend toutes les valeurs dans les entiers naturels  $\mathbb{N}$ , avec

$$\mathbb{P}(X = n) = \frac{\lambda^n}{n!} e^{-\lambda} \quad (1.38)$$

C'est la loi des v.a. décrivant des événements rares, nombre d'accidents, d'erreurs de fabrication, d'individus atteints d'une maladie etc., mais pas de tremblements de terre, de disparitions de dinosaures, etc.

On a

$$\mathbb{E}(X) = \text{Var}(X) = \lambda \quad (1.39)$$

▷ **Exemple 1.22** Dans un lac on pêche 1000 poissons qui sont marqués et puis sont affranchis dans le lac. Après une courte période du temps, on pêche 1000 poissons de nouveau. On constate que parmi eux 100 poissons sont marqués. Combien des poissons sont dans le lac ?

C'est un problème de statistique mathématique, un problème d'estimation paramétrique. Il s'agit en fait d'estimer le nombre des poissons du lac. Ça veut dire qu'on ne peut pas préciser exactement ce nombre (soit 8542 poissons!), mais on peut indiquer une valeur autour de laquelle se trouve le nombre exact des poissons. Évidemment ces estimations sont très utiles non seulement pour évaluer le fonds cynégétique, mais en sociologie, biologie, technique etc. La méthode que nous utiliserons est bien connue en statistique et s'appelle *la méthode de maximum de vraisemblance*.

Soient  $N$  (inconnu) le nombre des poissons du lac,  $M = 1000$  le nombre des poissons pêchés la première fois (donc marqués),  $n = 1000$  le nombre des poissons pêchés la deuxième fois,  $k = 100$  le nombre des poissons marqués qui ont été pêchés la deuxième fois.

En utilisant la loi hypergéométrique (cf. la formule (1.34)), la probabilité de l'événement réalisé est

$$p_k(N) = \frac{\binom{n}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$$

ou

$$p_{100}(N) = \frac{\binom{1000}{100} \binom{N-1000}{900}}{\binom{N}{1000}}.$$

Évidemment,  $N \geq 1900$  et, si on suppose  $n = 1900$ , alors on trouve  $p_{100}(1900) \cong 10^{-430}$  ce qui n'est pas vraisemblable! par contre, si on suppose  $N = 10^6$  grand, on trouve la même valeur trop petite pour l'événement qui s'est passé. En somme, on doit choisir pour  $N$  la valeur  $\hat{N}$  pour laquelle la probabilité  $p_{100}(N)$  soit maximale, i.e.

$$p_{100}(\hat{N}) = \max_N p_{100}(N).$$

On trouve  $\hat{N} = 10000$ .

&lt;

▷ **Exemple 1.23** On considère une centrale téléphonique avec un grand nombre des abonnés. Le nombre moyen des appels (flux poissonien) par jour est 50000. Le niveau de surcharge est défini comme le plus petit nombre  $N$ , tel que la probabilité de recevoir au moins  $N$  appels par seconde est moins que 0,001. On demande le niveau de surcharge.

Le nombre moyen d'appels par seconde est  $\frac{50000}{24 \cdot 3600} = 0,58$ . Donc les appels forment un flux poissonien  $P(\lambda)$  avec  $\lambda = 0,58$ . Si  $p(k)$  est la probabilité de  $k$  appels dans une intervalle d'une seconde fixée, alors le niveau de surcharge vérifie

$$p(N) + p(N+1) + \dots \leq 0,001.$$

Comme  $p(N) \gg p(N+1) \gg \dots$  on peut considérer tout simplement

$$p(N) \leq 0,001 \Leftrightarrow \frac{(0,58)^N}{N!} e^{-0,58} \leq 0,001$$

et on trouve  $N = 5$ .

&lt;

### Lois avec densité.

1. **Loi uniforme**- Une v.a. suit une loi uniforme sur l'intervalle  $[a, b]$ , et nous noterons  $X \sim U(a, b)$ , si elle admet la densité

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & x \notin [a, b] \end{cases} \quad (1.40)$$

On a

$$\mathbb{E}(X) = \frac{a+b}{2}, \text{ Var}(X) = \frac{(b-a)^2}{12} \quad (1.41)$$

2. **Loi normale ou de Gauss**- Une v.a. suit une loi gaussienne de paramètres  $m \in \mathbb{R}$  et  $\sigma > 0$ , et nous la noterons  $N(m, \sigma^2)$ , si elle admet la densité

$$\phi(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}} \quad (1.42)$$

On a  $\mathbb{E}(X) = m$  et  $\text{Var}(X) = \sigma^2$ .

*Utilisation pratique* : approximation de la somme (résultante additive) de phénomènes aléatoires petits et nombreux, et pas trop mutuellement dépendantes; utilisée parfois comme approximation de la résultante multiplicative de phénomènes aléatoires petits et nombreux, lorsque l'écart-type est relativement petit devant l'espérance mathématique (positive). De nombreuses lois "naturelles" sont ainsi approchées par une loi normale. La loi  $N(0, 1)$  est appelée *normale centrée réduite* et sa fonction de répartition, i.e.

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt, \quad x \in \mathbb{R},$$

est la *fonction de Laplace*; cette fonction a la propriété  $\Phi(-x) = 1 - \Phi(x)$ ,  $x \in \mathbb{R}$ . Si  $X \sim N(m, \sigma^2)$ , alors

$$\mathbb{P}(a \leq X \leq b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right). \quad (1.43)$$

En particulier, pour  $\alpha > 0$ ,

$$\mathbb{P}(m - \alpha \leq X \leq m + \alpha) = \Phi\left(\frac{\alpha}{\sigma}\right) - \Phi\left(-\frac{\alpha}{\sigma}\right) = 2\Phi\left(\frac{\alpha}{\sigma}\right) - 1. \quad (1.44)$$

3. **Loi gamma**-La densité de probabilité de loi gamma de paramètres  $a > 0$  et  $b > 0$ , notée  $\gamma(a, b)$  est

$$f(x) = \begin{cases} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1.45)$$

où

$$\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx.$$

On a  $\mathbb{E}(X) = \frac{a}{b}$ ,  $\text{Var}(X) = \frac{a}{b^2}$ . La loi  $\gamma(n/2, 1/2)$ ,  $n \in \mathbb{N}^*$ , est appelée loi du  $\chi^2$  à  $n$  degrés de liberté et nous la noterons  $\chi^2(n)$ .

La loi  $\gamma(n, b)$  pour  $n \in \mathbb{N}^*$  et  $b > 0$ , est appelée *loi d'Erlang*, dont la f.r. est

$$F(x) = 1 - \sum_{k=0}^{n-1} e^{-bx} \frac{(bx)^k}{k!}.$$

4. **Loi exponentielle**- La loi  $\gamma(1, \lambda)$  est appelée *loi exponentielle* de paramètre  $\lambda > 0$  et nous la noterons  $\mathcal{E}(\lambda)$ . La loi exponentielle est de loin la loi la plus utilisée dans les applications d'ingénieur.
5. **Loi de Weibull**- La densité de probabilité de loi de Weibull de paramètres  $\alpha > 0$  et  $\beta > 0$ , notée  $W(\alpha, \beta)$ , est

$$f(x) = \begin{cases} \frac{\beta}{\alpha} \left(\frac{x}{\alpha}\right)^{\beta-1} e^{-(x/\alpha)^\beta}, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (1.46)$$

On a

$$\mathbb{E}(X) = \alpha \Gamma\left(1 + \frac{1}{\beta}\right), \quad \text{et} \quad \text{Var}(X) = \alpha^2 \left[ \Gamma\left(1 + \frac{2}{\beta}\right) - \Gamma\left(1 + \frac{1}{\beta}\right)^2 \right] \quad (1.47)$$

La loi de Weibull est utilisée en fiabilité pour modéliser l'usure d'un système mécanique. La loi  $W(\sigma\sqrt{2}, 2)$  est appelée *loi de Rayleigh* de paramètre  $\sigma > 0$ ; elle est utilisée pour modéliser la propagation des ondes radio-électriques.

6. **Loi beta**-Une v.a. suit une loi bêta de paramètres  $p > 0$  et  $q > 0$ , notée  $B(p, q)$  est

$$f(x) = \begin{cases} \frac{1}{B(p, q)} x^{p-1} (1-x)^{q-1}, & x \in [0, 1] \\ 0, & x \notin [0, 1] \end{cases} \quad (1.48)$$

où

$$B(p, q) = \int_0^1 x^{p-1} (1-x)^{q-1} dx.$$

On a

$$\mathbb{E}(X) = \frac{p}{p+q}, \quad \text{Var}(X) = \frac{pq}{(p+q)^2(1+p+q)}. \quad (1.49)$$

7. **Loi de Fisher-Snedecor**- La densité de probabilité de la loi de Fisher-Snedecor de paramètres  $n, m \in \mathbb{N}^*$ , notée  $F(n, m)$ , est

$$f(x) = \frac{n^{\frac{n}{2}} m^{\frac{m}{2}}}{B\left(\frac{n}{2}, \frac{m}{2}\right)} \frac{x^{\frac{n}{2}-1}}{(m+nx)^{\frac{m+n}{2}}}. \quad (1.50)$$

Si  $m > 4$ , alors

$$\mathbb{E}(X) = \frac{m}{m-2}, \quad \text{Var}(X) = \frac{2m^2(m+n-2)}{n(m-4)(m-2)^2} \quad (1.51)$$

8. **Loi de Student**-Une v.a. suit une loi de Student de paramètre  $n \in \mathbb{N}^*$ , et nous noterons  $X \sim t(n)$ , si elle admet la densité

$$f(x) = \frac{n^{-\frac{1}{2}}}{B\left(\frac{1}{2}, \frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}. \quad (1.52)$$

▷ **Exemple 1.24** Soit une v.a.  $X \sim N(m, \sigma^2)$ . On demande la probabilité que  $X$  s'écarte de son espérance mathématique avec plus de  $3\sigma$ .

On a

$$\mathbb{P}(|X - m| > 3\sigma) = 1 - \mathbb{P}(|X - m| \leq 3\sigma) = 1 - \Phi(3) + \Phi(-3) = 2(1 - \Phi(3))$$

Comme  $\Phi(3) \cong 0,99865$  on obtient  $\mathbb{P}(|X - m| > 3\sigma) \cong 0,0027$ .

◁

▷ **Exemple 1.25** Une entreprise produit des billes de roulements avec le diamètre 10 mm. En fait le diamètre est une v.a.  $X \sim N(10; 0,4^2)$ . Au contrôle on accepte seulement les billes avec le diamètre dans l'intervalle  $[9,3; 10,7]$ . On demande le pourcentage des billes défectueuses.

La probabilité qu'une bille soit défectueuse est

$$\begin{aligned} \mathbb{P}(|X - m| > 0,7) &= 1 - \mathbb{P}(|X - m| \leq 0,7) = \left(1 - \Phi\left(\frac{0,7}{0,4}\right)\right) \\ &= 2(1 - \Phi(0,75)) \cong 0,08. \end{aligned}$$

Donc il y a 8% de billes défectueuses.

◁

## 1.4 Les fonctions génératrice et caractéristique

**Définition 1.9.** Soit  $X$  une v.a. à valeurs entières positives. La fonction génératrice de  $X$  est définie par

$$G_X(t) = \mathbb{E}(t^X) = \sum_{n \geq 0} \mathbb{P}(X = n)t^n$$

qui est finie au moins pour  $t \in [-1, 1]$ .

Une fonction génératrice a les propriétés suivantes :

1.  $G_X(0) = \mathbb{P}(X = 0)$  et  $G_X(1) = 1$ . De plus

$$0 \leq G_X(t) \leq 1 \text{ pour } t \in [0, 1]$$

2.  $G_X$  est de classe  $C^\infty$  sur  $(-1, 1)$  (indéfiniment dérivable) et

$$G_X^{(n)}(0) = n! \mathbb{P}(X = n) \quad (1.53)$$

3.  $\mathbb{E}(X) = \lim_{t \searrow 1} G'_X(t)$  ;

4. Si  $X$  est d'espérance finie, on a

$$\text{Var}(X) = G'_X(1)[1 - G'_X(1)] + \lim_{t \searrow 1} G''_X(t)$$

5. Si les v.a.  $X$  et  $Y$  sont indépendantes, alors

$$G_{X+Y} = G_X G_Y$$

6. Soient les v.a.  $(X_n, n \geq 1)$  indépendantes et identiquement distribuées (i.i.d.) à valeurs entières de loi

$$X_1 : \begin{pmatrix} 0 & 1 & 2 & \cdots \\ p_0 & p_1 & p_2 & \cdots \end{pmatrix}.$$

Soit  $N$  une v.a. indépendante de  $(X_n, n \geq 1)$  à valeurs dans  $\mathbb{N}$  et on définit la suite de v.a.

$$\begin{aligned} S_0 &= 0 \\ S_n &= X_1 + \cdots + X_n, \quad n \geq 1. \end{aligned}$$

Alors  $S_N$  est une somme aléatoire avec la fonction génératrice

$$G_{S_N}(t) = G_N(G_{X_1}(t)), \quad 0 \leq t \leq 1 \quad (1.54)$$

Remarquons que si  $G_X$  est dérivable à gauche une fois en 1, alors

$$\mathbb{E}(X) = G'_X(1) \quad (1.55)$$

et si elle l'est deux fois,

$$\text{Var}(X) = G''_X(1) + G'_X(1) - [G'_X(1)]^2 \quad (1.56)$$

▷ **Exemple 1.26** 1. Soit  $X \sim P(\lambda)$ . On a

$$G_X(t) = e^{-\lambda} \sum_{n \geq 1} \frac{(t\lambda)^n}{n!} = e^{\lambda(t-1)}. \quad (1.57)$$

2. Soit  $X \sim B_-(r, p)$ . La fonction génératrice de  $X$  est définie pour  $|t| < \frac{1}{1-p}$ , et

$$\begin{aligned} G_X(t) &= p^r t^r \sum_{k=0}^{\infty} \binom{r+k-1}{k} t^k (1-p)^k \\ &= p^r t^r \sum_{k=0}^{\infty} (-1)^k \frac{(-r)(-r-1) \cdots (-r-k+1)}{k!} t^k (1-p)^k \\ &= \frac{p^r t^r}{(1 - (1-p)t)^r}. \end{aligned} \quad (1.58)$$

3. Si  $X \sim b(n, p)$ , alors  $X = X_1 + \cdots + X_n$ , où les v.a.  $X_1, \dots, X_n$  sont indépendantes et identiquement distribuées  $B(p)$ . Comme  $G_{X_1}(t) = (1-p) + pt$  on a

$$G_X(t) = (1-p + pt)^n. \quad (1.59)$$

◁

▷ **Exemple 1.27** Si  $X \sim P(\lambda)$ ,  $Y \sim P(\mu)$  et  $X, Y$  sont indépendantes, alors

$$G_{X+Y}(t) = G_X(t)G_Y(t) = e^{\lambda(t-1)}e^{\mu(t-1)} = e^{(\lambda+\mu)(t-1)},$$

donc  $X + Y \sim P(\lambda + \mu)$ .

◁

Lois	$\mathbb{P}(X = k)$	$\mathbb{E}[X]$	$\text{Var}[X]$	$G(t)$
$B(p)$	$\mathbb{P}(X = 1) = p$	$p$	$p(1 - p)$	$1 - p + pt$
$b(n, p)$	$C_n^k p^k (1 - p)^{n-k}$	$np$	$np(1 - p)$	$(1 - p + pt)^n$
$\mathcal{P}(\lambda)$	$e^{-\lambda} \frac{\lambda^k}{k!}$	$\lambda$	$\lambda$	$\exp(-\lambda + \lambda t)$
$G(p)$	$p(1 - p)^{k-1}$	$1/p$	$(1 - p)/p^2$	$pt/(1 - t + pt)$

FIGURE 1.1 – Fonctions génératrices des lois discrètes usuelles

▷ **Exemple 1.28** Un jour le propriétaire d'un restaurant où on peut commander par téléphone des mets chinois est indisposé car il a fait la noce la veille. Le nombre de commandes est une v.a.  $N$  qui suit une loi  $P(\lambda)$  et la probabilité que le propriétaire note l'adresse correctement est  $p$ . Soit  $X_i, i = 1, 2, \dots$ , des v.a. indépendantes et identiquement distribuées  $B(p)$  qui marquent si les adresses sont notées correctement ou non. Le nombre de commandes livrées avec succès est  $S_N = X_1 + \dots + X_N$  et, d'après (1.54) et (1.57), la fonction génératrice est

$$G_{S_N}(t) = G_N(1 - p + pt) = e^{\lambda p(t-1)},$$

donc  $S_N \sim P(\lambda p)$ .

Suite à la composition de deux lois, le paramètre  $\lambda$  de la loi de Poisson de la v.a.  $N$  a été diminué à  $\lambda p$ ; ce phénomène est appelé *l'aminçissement* dans la théorie de processus de Poisson. Autres exemples : les clients arrivent (flux poissonien de paramètre  $\lambda$ ) à une station de service pour être servis à la file. La probabilité qu'un client soit content de la prestation est  $p$  ainsi que  $S_N$  est le nombre de clients contents. En outre, nous pouvons imaginer que les appels d'un trafic téléphonique arrivants à une centrale sont de deux types, I et II, et ils sont dirigés par deux lignes différentes. Si le nombre total d'appels arrivés dans une heure est poissonien de paramètre  $\lambda$  et 100% d'entre eux sont de type II, alors le nombre total d'appels dirigés par la ligne de type II suit une loi Poisson de paramètre  $p\lambda$ .

◁

**Définition 1.10.** La fonction  $\varphi: \mathbb{R} \rightarrow \mathbb{C}$ ,

$$\varphi(t) = \mathbb{E}[e^{itX}], \quad t \in \mathbb{R} \quad (1.60)$$

dont la valeur au point  $t$  coïncide avec l'espérance mathématique  $\mathbb{E}[e^{itX}]$  d'une v.a.  $X$  est appelée *fonction caractéristique* de cette v.a.  $X$  (ou de la fonction de répartition correspondante  $F_X$ ).

La fonction  $\varphi(t)$  est continue et bornée,  $|\varphi(t)| \leq 1$  et  $\lim_{|t| \rightarrow \infty} |\varphi(t)| = 0$ .

Supposons que  $X$  soit discrète avec la loi concentrée en des points entiers  $k = 0, \pm 1, \pm 2, \dots$ . Sa fonction caractéristique

$$\varphi(t) = \sum_{k=-\infty}^{\infty} e^{itx_k} p_k, \quad \text{où } p_k = \mathbb{P}(X = k), \quad k \in \mathbb{Z}, \quad (1.61)$$

est alors une fonction périodique (de période  $2\pi$ ) et la formule (1.61) donne son développement en série Fourier de coefficients

$$p_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ix_k t} \varphi(t) dt, \quad k \in \mathbb{Z}. \quad (1.62)$$

Si la loi admet une densité  $p(x)$ , alors la fonction caractéristique s'écrit

$$\varphi(t) = \hat{p}(t) = \int_{-\infty}^{\infty} e^{itx} p(x) dx, \quad t \in \mathbb{R}, \quad (1.63)$$



c'est la transformée de Fourier de la fonction  $p(x)$ , et pour une fonction absolument intégrable  $\varphi$ , la densité  $p(x)$  peut être obtenue par la transformation de Fourier inverse

$$p(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ixt} \varphi(t) dt. \quad (1.64)$$

Si  $\mathbb{E}|X|^n < \infty$ , alors

$$\mathbb{E}X^k = i^k \varphi^{(k)}(0), \quad k = 1, 2, \dots, n. \quad (1.65)$$

Lois	Fonctions caractéristiques
Bernoulli $B(p)$	$pe^{it} + (1-p)$
Binomiale $b(n, p)$	$(pe^{it} + (1-p))^n$
Poisson $\mathcal{P}(\lambda)$	$\exp(-\lambda(1 - \exp it))$
Uniforme sur $[a, b]$	$\frac{e^{ibt} - e^{iat}}{it(b-a)}$
Normale $N(\mu, \sigma^2)$	$\exp(it\mu - \sigma^2 t^2/2)$
Exponentielle $\mathcal{E}(\lambda)$	$\lambda/(\lambda - it)$
Chi-deux $\chi^2(n)$	$(1 - 2it)^{-\frac{n}{2}}$

Une autre propriété intéressante de  $\varphi$  est que si  $\hat{p}(t)$  est la fonction caractéristique de  $p(x)$ , alors  $\widehat{\hat{p}}(x) = 2\pi p(-x)$ . En fait,  $\widehat{\hat{p}}(x)$  est une fonction caractéristique de  $\hat{p}$ , c'est une combinaison des formules (1.63) et (1.64).

▷ **Exemple 1.29** La fonction caractéristique de la fonction  $e^{-a|x|}$ , ( $a > 0$ ) est  $\frac{2a}{a^2+t^2}$ . Maintenant de la formule de la double transformée ci-dessus, nous obtenons  $\widehat{\widehat{e^{-a|x|}}} = \widehat{\frac{2a}{a^2+t^2}} = 2\pi e^{-a|x|}$ . D'où la fonction caractéristique de la loi de Cauchy,  $f_a(x) = \frac{1}{\pi} \frac{a}{a^2+x^2}$ , est donnée par  $e^{-a|t|}$ . Nous en déduisons  $\hat{f}_a(t)\hat{f}_b(t) = \hat{f}_{a+b}(t)$ , ce qui entraîne  $f_a * f_b = f_{a+b}$ .

◁

## 1.5 Vecteurs aléatoires

Nous avons déjà vu qu'un vecteur aléatoire (ou une v.a.  $d$ -dimensionnelle réelle) est une application mesurable  $\mathbf{X}$  définie sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  à valeurs dans  $\mathbb{R}^d$ . Si on note  $X_j = \text{pr}_j \circ \mathbf{X}$ ,  $j = 1, \dots, d$ , alors

$$\mathbf{X} = (X_1, \dots, X_d): (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d)).$$

La fonction  $\mathbf{X}$  est un vecteur aléatoire si et seulement si toutes les applications coordonnées,  $X_j$ ,  $j = 1, \dots, d$ , sont des v.a. réelles.

La loi de  $\mathbf{X}$  (ou loi jointe de  $X_1, \dots, X_d$ ) et une probabilité sur  $\mathcal{B}(\mathbb{R}^d)$ , notée  $P_{\mathbf{X}}$ , définie par

$$\mathbb{P}_{\mathbf{X}}(B) = \mathbb{P}(\mathbf{X}^{-1}(B)) = \mathbb{P}(\{\omega : (X_1(\omega), \dots, X_d(\omega)) \in B\}), \quad B \in \mathcal{B}(\mathbb{R}^d) \quad (1.66)$$

La fonction de répartition  $F_{\mathbf{X}}: \mathbb{R}^d \rightarrow [0, 1]$  du vecteur aléatoire  $\mathbf{X}$  est définie par

$$F_{\mathbf{X}}(x_1, \dots, x_d) = \mathbb{P}_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_d]) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$$

et a des propriétés similaires aux propriétés d'une v.a.<sup>1</sup>

1. On peut discuter sur les conditions nécessaires et suffisantes pour qu'une fonction  $F: \mathbb{R}^d \rightarrow [0, 1]$  soit la fonction de répartition d'un certain vecteur aléatoire. Ainsi, pour  $d > 1$  la condition de croissance dans chaque variable n'est plus suffisante. Par exemple, pour  $d = 2$  on doit ajouter la condition

$$F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1) \geq 0, \quad a_1 < a_2, b_1 < b_2.$$

La densité de probabilité (si elle existe!) de  $\mathbf{X}$  (ou de la loi jointe de  $X_1, \dots, X_d$ ) est une fonction mesurable  $f_{\mathbf{X}}: \mathbb{R} \rightarrow \mathbb{R}_+$  telle que

$$F_{\mathbf{X}}(x_1, \dots, x_d) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} f_{\mathbf{X}}(t_1, \dots, t_d) dt_1 \dots dt_d, \quad x_1, \dots, x_d \in \mathbb{R}.$$

Évidemment,

$$\int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_{\mathbf{X}}(t_1, \dots, t_d) dt_1 \dots dt_d = 1.$$

En général, la densité vérifie la relation

$$\mathbb{P}(\mathbf{X} \in B) = \int \dots \int_B f_{\mathbf{X}}(t_1, \dots, t_d) dt, \quad B \in \mathcal{B}(\mathbb{R}^d), \quad dt = (dt_1 \dots dt_d). \quad (1.67)$$

**Proposition 1.8.** *Les v.a.  $X_1, \dots, X_n$  sont mutuellement indépendantes si et seulement si la fonction de répartition  $F$  du vecteur aléatoire  $(X_1, \dots, X_n)$  vérifie*

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \dots F_{X_n}(x_n), \quad \text{pour tous } x_1, \dots, x_n \in \mathbb{R} \quad (1.68)$$

*Démonstration.* On utilise (1.18) et la proposition 1.3. □

Si le vecteur  $(X_1, \dots, X_n)$  admet une densité de probabilité  $p$ , alors (1.68) est équivalente à

$$p(x_1, \dots, x_n) = p_{X_1}(x_1) \dots p_{X_n}(x_n), \quad \text{pour tous } x_1, \dots, x_n \in \mathbb{R}. \quad (1.69)$$

Les lois *marginale*s de la loi de  $\mathbf{X}$  sont les lois des  $r$ -uplets  $(X_{i_1}, \dots, X_{i_d})$  pour  $1 \leq i_1 < \dots < i_r \leq d$  et  $1 \leq r < d$ .

▷ **Exemple 1.30** (*Loi multinomiale*) Une urne contient des boules et chaque boule porte un numéro  $i$  où  $1 \leq i \leq r$ . La probabilité de tirer une boule portant le numéro  $i$  est  $p_i$ ,  $i = 1, \dots, r$ , et  $p_1 + \dots + p_r = 1$ . On tire  $n$  boules au hasard et avec remise. Soit  $X_i$ ,  $i = 1, \dots, r$  le nombre de boules tirées portant le numéro  $i$ . Le vecteur aléatoire  $\mathbf{X} = (X_1, \dots, X_r)$  a la loi

$$\mathbb{P}(\mathbf{X} = (k_1, \dots, k_r)) = \frac{n!}{k_1! \dots k_r!} p_1^{k_1} \dots p_r^{k_r}, \quad k_1, \dots, k_r \in \mathbb{N}, \quad k_1 + \dots + k_r = n \quad (1.70)$$

Ceci définit la loi multinomiale  $M(n; p_1, \dots, p_r)$ . ◁

▷ **Exemple 1.31** (*Loi normale ou gaussienne multidimensionnelle*) Considérons  $d$  variables aléatoires indépendantes  $Z_1, \dots, Z_d$  de même lois  $N(0, 1)$ . Leur densité de probabilité jointe est

$$\phi(x_1, \dots, x_d) = \frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2} \sum_{k=1}^d x_k^2}, \quad x_1, \dots, x_d \in \mathbb{R}.$$

Soit

$$X_i = \sum_{k=1}^d \sigma_{ik} Z_k + m_i, \quad i = 1, \dots, d \quad (1.71)$$

une transformation linéaire non-dégénérée des variables  $Z_k$ ,  $k = 1, \dots, d$ . On a évidemment

$$\mathbb{E}(X_i) = m_i, \quad \mathbb{E}[(X_i - m_i)(X_j - m_j)] = \sum_{k=1}^d \sigma_{ik} \sigma_{jk}, \quad i, j = 1, \dots, d \quad (1.72)$$

La matrice  $\mathbf{R} = (R_{kj})$  est la *matrice de covariance* des v.a.  $X_1, \dots, X_d$  et  $r_{ij} = \frac{R_{ij}}{\sqrt{R_{ii}R_{jj}}}$  est le *coefficient de corrélation* des v.a.  $X_i$  et  $X_j$ . Il découle des formules (1.72) que la matrice de covariance  $\mathbf{R}$  dans le cas envisagé est  $\mathbf{R} = \boldsymbol{\sigma}\boldsymbol{\sigma}^t$  où  $\boldsymbol{\sigma} = (\sigma_{ij})$  est la matrice de la transformation linéaire (1.71) et  $\boldsymbol{\sigma}^t$  la matrice transposée de  $\boldsymbol{\sigma}$ . On a  $\det \mathbf{R} = (\det \boldsymbol{\sigma})^2$  et  $\det \boldsymbol{\sigma}$  est le jacobien de la transformation (1.71). Pour la densité jointe des v.a.  $(X_1, \dots, X_d)$  on a la formule

$$\begin{aligned} f(\mathbf{x}) &= \frac{1}{(2\pi)^{d/2}(\det \mathbf{R})^{1/2}} \exp \left[ -\frac{1}{2} \sum_{i,j=1}^d b_{ij}(x_i - m_i)(x_j - m_j) \right] \\ &= (2\pi)^{-d/2} (\det \mathbf{R})^{-1/2} \exp \left[ -\frac{1}{2} (\mathbf{x} - \mathbf{m})^\top \mathbf{R}^{-1} (\mathbf{x} - \mathbf{m}) \right], \end{aligned} \quad (1.73)$$

et on note  $\mathbf{X} \sim N_d(\mathbf{m}, \mathbf{R})$ , où  $\mathbf{x} = (x_1, \dots, x_d)$  et  $b(\mathbf{x}) = \sum_{i,j=1}^d b_{ij}(x_i - m_i)(x_j - m_j)$  est la forme

quadratique que prend la somme des carrés  $\sum_{k=1}^d x_k^2$  lors de la transformation linéaire inverse de (1.71). Il est facile de voir que la matrice  $\mathbf{B} = (b_{ij})$  de cette forme quadratique est l'inverse de la matrice de covariance  $\mathbf{R}$ , car la somme des carrés  $\sum_{k=1}^d x_k^2$  devient la forme quadratique

$$\sum_{k=1}^d \left[ \sum_{j=1}^d c_{kj}(x_j - m_j) \right]^2 = \sum_{i,j=1}^d \left( \sum_{k=1}^d c_{ki}c_{kj} \right) (x_i - m_i)(x_j - m_j)$$

où  $\mathbf{C} = (c_{kj})$  est la matrice inverse de  $\boldsymbol{\sigma}$ , de sorte que

$$\mathbf{B} = \mathbf{C}^t \mathbf{C} = [\boldsymbol{\sigma} \boldsymbol{\sigma}^t]^{-1} = \mathbf{R}^{-1}.$$

Une loi dont la densité de probabilité est donnée par (1.73) ainsi que le vecteur aléatoire  $(X_1, \dots, X_d)$  obéissant à cette loi sont appelées *normales* (ou *gaussiennes*). Il est évident que toute transformation linéaire non-dégénérée transforme des variables gaussiennes  $(X_1, \dots, X_d)$  en des variables également gaussiennes.

**Définition 1.11.** Une v.a.v.  $\mathbf{X} = (X_1, \dots, X_d)^\top$  à valeurs dans  $\mathbb{R}^d$  est dite gaussienne, si pour tout vecteur fixé  $\mathbf{a} = (a_1, \dots, a_d)^\top \in \mathbb{R}^d$ , la v.a. réelle  $a_1 X_1 + \dots + a_d X_d$  est gaussienne.

**Proposition 1.9.** Soit  $\mathbf{X}$  une v.a.v. gaussienne à valeurs dans  $\mathbb{R}^d$ . Alors

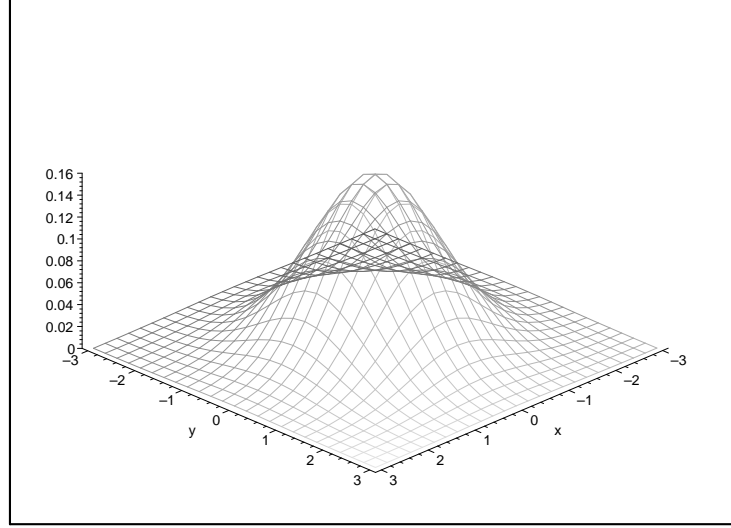
1. chacune de ses coordonnées  $X_i$  est une v.a. gaussienne ;
2. si  $\mathbf{u} : \mathbb{R}^d \rightarrow \mathbb{R}^m$  est une application affine, alors  $\mathbf{u} \circ \mathbf{X}$  est gaussienne. Si, de plus,  $\mathbf{Y} = \mathbf{B}\mathbf{X} + \mathbf{b}$ , alors  $K_{\mathbf{Y}} = \mathbf{B}K_{\mathbf{X}}\mathbf{B}^\top$ .

**Remarque 1.2.** La réciproque de 1. est fausse.

◁

▷ **Exemple 1.32** Vecteur gaussien bidimensionnelle. Soit  $(X_1, X_2)' \sim N_2(\mu, K)$ , avec  $\mu = (\mu_1, \mu_2)'$  et

$$K = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}.$$

FIGURE 1.2 – Densité de la loi normale  $\mathcal{N}_2(0, I)$ .-

Alors, nous obtenons  $\det(K) = \sigma_1^2 \sigma_2^2 (1 - \rho^2)$ , et

$$K^{-1} = \frac{1}{\sigma_1^2 \sigma_2^2 (1 - \rho^2)} \begin{bmatrix} \sigma_2^2 & -\rho \sigma_1 \sigma_2 \\ -\rho \sigma_1 \sigma_2 & \sigma_1^2 \end{bmatrix}.$$

Et donc la densité de  $(X_1, X_2)'$  s'écrit

$$f(x_1, x_2) = \frac{1}{2\pi \sigma_1 \sigma_2 (1 - \rho^2)^{1/2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1 \sigma_2} \right] \right\}.$$

◁

Soit  $\mathbf{X} = (X_1, \dots, X_d)$  un vecteur aléatoire réel défini sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  qui admet une densité de probabilité  $p(x_1, \dots, x_d)$ . Si  $h: \mathbb{R}^d \rightarrow \mathbb{R}$  est une fonction mesurable, alors la fonction  $Y = h \circ \mathbf{X}$  est une v.a. réelle et sa fonction de répartition est donnée par la formule

$$F(y) = \mathbb{P}(Y \leq y) = \int_M f(x_1, \dots, x_d) d\mathbf{x} \quad (1.74)$$

où  $M = \{(x_1, \dots, x_d) \in \mathbb{R}^d : h(x_1, \dots, x_d) \leq y\}$  et  $d\mathbf{x} = dx_1 \cdots dx_d$ .

▷ **Exemple 1.33** Soit le vecteur aléatoire  $(X_1, X_2)$  avec densité de probabilité  $f(x_1, x_2)$ . De (1.74) il en résulte que la fonction de répartition de v.a.  $Y = X_1 + X_2$  est

$$F(y) = \iint_{\{x_1 + x_2 \leq y\}} f(x_1, x_2) dx_1 dx_2$$

et après le changement des variables  $x_1 + x_2 = u$ ,  $x_2 = v$ , on obtient

$$F(y) = \int_{-\infty}^y du \int_{-\infty}^{\infty} f(u - v, v) dv. \quad (1.75)$$

De (1.75) il découle que la densité de  $Y$  est

$$q(y) = \int_{-\infty}^{\infty} f(y - v, v) dv. \quad (1.76)$$

En particulier, si les v.a.  $X_1$  et  $X_2$  sont indépendantes avec les densités  $f_1(x_1)$  et respectivement  $f_2(x_2)$ , alors  $f(x_1, x_2) = f_1(x_1)f_2(x_2)$  (cf. (1.69)) et (1.76) devient

$$q(y) = \int_{-\infty}^{\infty} f_1(y-v)f_2(v)dv. \quad (1.77)$$

L'intégrale de (1.77) est le *produit de convolution* de  $f_1$  et  $f_2$  et on écrit  $q = f_1 * f_2$ . Si, de plus, les v.a.  $X_1$  et  $X_2$  sont positives, alors on a

$$q(y) = \int_0^y f_1(y-v)f_2(v)dv. \quad (1.78)$$

◁

*Espaces  $\mathcal{L}^p$ .* Soit  $\mathcal{L}^p := \mathcal{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ , avec  $1 \leq p < \infty$ , l'espace de v.a. telles que  $\mathbb{E}(|X|^p) < \infty$ , muni de la norme

$$\|X\|_p := [\mathbb{E}(|X|^p)]^{1/p}.$$

Les espaces  $\mathcal{L}^p$  sont des espaces vectoriels normés (et complets).<sup>2</sup>

Soit maintenant  $p = 2$ , et considérons l'espace  $\mathcal{L}^2$  muni du produit scalaire

$$(X | Y) := \mathbb{E}(XY),$$

pour  $X, Y \in \mathcal{L}^2$ , et de la norme associée à ce produit scalaire<sup>3</sup>

$$\|X\|_2 := (X | X)^{1/2} = [\mathbb{E}(|X|^2)]^{1/2}.$$

Nous avons les inégalités suivantes :

- Cauchy-Schwarz : pour  $X, Y \in \mathcal{L}^2$ ,

$$|(X | Y)| \leq \|X\|_2 \|Y\|_2.$$

- Minkowski : pour  $X, Y \in \mathcal{L}^2$ ,

$$\|X + Y\|_2 \leq \|X\|_2 + \|Y\|_2.$$

- Hölder : si  $\frac{1}{p} + \frac{1}{q} = 1$ , pour  $p \geq 1, q \geq 1$ , alors

$$\|XY\|_1 \leq \|X\|_p \|Y\|_q.$$

- Lyapunov : si  $0 < p < q < \infty$ , alors

$$\|X\|_p \leq \|Y\|_q.$$

On peut observer que l'inégalité de Cauchy-Schwarz peut être obtenue par l'inégalité de Hölder pour  $p = q = 1/2$ .

Orthogonalité :

- Si  $(X | Y) = 0$ , les v.a.  $X$  et  $Y$  sont dites orthogonales et nous avons

$$\|X + Y\|_2^2 = \|X\|_2^2 + \|Y\|_2^2,$$

le théorème de Pythagore.

---

2. Un espace vectoriel normé et complet est appelé espace de Banach.

3. Un tel espace vectoriel est appelé espace de Hilbert.

Corrélation  $\rho_{XY}$  :

$$\rho_{XY} = \frac{(X - \mathbb{E}X | Y - \mathbb{E}Y)}{\|X - \mathbb{E}X\|_2 \|Y - \mathbb{E}Y\|_2}.$$

On notera également  $L^p$  l'espace quotient de  $\mathcal{L}^p$  pour l'égalité p.s.

Nous pouvons écrire également que :

$$(X | \mathbb{1}_\Omega) = (\mathbb{E}X)\mathbb{1}_\Omega$$

et

$$\|X - \mathbb{E}X\|_2^2 = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{V}\text{ar}(X).$$

*Variable aléatoire complexe.* Une v.a. complexe, i.e.,  $X : \Omega \rightarrow \mathbb{C}$ , est une v.a.  $X = X_1 + iX_2$ , où  $X_1$  et  $X_2$  sont des v.a. réelles et  $i = \sqrt{-1}$ .

Nous avons :  $\overline{X} = X_1 - iX_2$ ,  $|X|^2 = X_1^2 + X_2^2$ ,  $\mathbb{E}(X) = \mathbb{E}X_1 + i\mathbb{E}X_2$  et  $\mathbb{E}(\overline{X}) = \overline{\mathbb{E}(X)}$ .

Nous pouvons montrer également pour deux v.a. complexes  $X$  et  $Y$  :

$$(X | Y) := \mathbb{E}(X\overline{Y})$$

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}(X))(\overline{Y - \mathbb{E}(Y)})] = \mathbb{E}(X\overline{Y}) - \mathbb{E}(X)\mathbb{E}(\overline{Y}).$$

Avec la définition précédente, pour  $X$  v.a. complexe, nous avons

$$\text{Cov}(X, X) = \mathbb{E}|X - \mathbb{E}(X)|^2 \geq 0$$

ce n'est plus la variance de  $X$ . Néanmoins, elle peut être utilisée pour décrire sa dispersion de  $X$ .

## 1.6 Probabilité et espérance conditionnelles

Considérons d'abord une partition  $\mathcal{B} = \{B_1, \dots, B_n\}$  de  $\Omega$ . Si  $A$  est un événement quelconque, alors la *probabilité conditionnelle de  $A$  sachant  $\mathcal{B}$*  est une v.a. définie par

$$\mathbb{P}(A | \mathcal{B})(\omega) = \sum_{i=1}^n \mathbb{P}(A | B_i) \mathbb{1}_{B_i}(\omega), \quad \omega \in \Omega. \quad (1.79)$$

On doit remarquer que la v.a.  $\mathbb{P}(A | \mathcal{B})$  est  $\sigma(\mathcal{B})$ -mesurable et qu'elle vérifie

$$\int_B \mathbb{P}(A | \mathcal{B}) d\mathbb{P} = \mathbb{P}(A \cap B), \quad B \in \mathcal{B}. \quad (1.80)$$

De plus,  $\mathbb{E}[\mathbb{P}(A | \mathcal{B})] = \mathbb{P}(A)$ .

Ceci suggère la possibilité de définir la probabilité conditionnelle  $\mathbb{P}(A | \mathcal{G})$  par rapport à une sous-tribu  $\mathcal{G}$  quelconque de  $\mathcal{F}$ . Elle sera une v.a.  $\mathcal{G}$ -mesurable vérifiant

$$\int_B \mathbb{P}(A | \mathcal{G}) d\mathbb{P} = \mathbb{P}(A \cap B), \quad B \in \mathcal{G}. \quad (1.81)$$

▷ **Exemple 1.34** Pour  $\mathcal{G} = \{\emptyset, \Omega\}$  on a  $\mathbb{P}(A | \mathcal{G}) = \mathbb{P}(A)$ .

◁

**Remarque 1.6.** La probabilité conditionnelle dépend évidemment de la probabilité  $\mathbb{P}$  et elle est définie par la relation (1.81). Par conséquent elle est définie  $\mathbb{P}$ -p.s.

Nous rappelons quelques propriétés significatives de la probabilité conditionnelle

1.  $\mathbb{P}(\Omega \mid \mathcal{G}) = 1$ ,  $\mathbb{P}$ -p.s.
2. Pour  $A \in \mathcal{F}$ ,  $B \in \mathcal{G}$  on a

$$\mathbb{P}(A \cap B \mid \mathcal{G}) = \mathbb{P}(A \mid \mathcal{G}) \mathbb{1}_B, \mathbb{P} - \text{p.s.} \quad (1.82)$$

En particulier, pour  $A = \Omega$  et  $B \in \mathcal{G}$ , on obtient

$$\mathbb{P}(B \mid \mathcal{G}) = \mathbb{1}_B, \mathbb{P} - \text{p.s.}$$

3. Soit  $\mathbf{X}: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$  un vecteur aléatoire et  $\mathcal{G} = \sigma(\mathbf{X})$ . Alors, pour  $A \in \mathcal{F}$ , il existe une fonction borélienne  $\phi(\cdot, A): \mathbb{R}^d \rightarrow [0, 1]$  telle que

$$\mathbb{P}(A \mid \mathcal{G}) = \mathbb{P}(A \mid \mathbf{X}) = \phi(\mathbf{X}, A), \mathbb{P} - \text{p.s.} \quad (1.83)$$

On utilise la notation  $\phi(\mathbf{x}, A) = \mathbb{P}(A \mid \mathbf{X} = \mathbf{x})$ .

4. De (1.80) et (1.83) on déduit que, pour tout  $C \in \mathcal{B}(\mathbb{R}^d)$ , on a

$$\begin{aligned} \mathbb{P}(A \cap (X \in C)) &= \int_C \phi(\mathbf{x}, A) (\mathbb{P} \circ \mathbf{X}^{-1})(d\mathbf{x}) = \int_C \phi(\mathbf{x}, A) F_{\mathbf{X}}(d\mathbf{x}) \\ &= \int_C \mathbb{P}(A \mid \mathbf{X} = \mathbf{x}) F_{\mathbf{X}}(d\mathbf{x}) \end{aligned} \quad (1.84)$$

où  $F_{\mathbf{X}}$  est la fonction de répartition du vecteur aléatoire  $\mathbf{X}$ .

Pour la simplicité de l'écriture nous allons considérer le cas  $d = 1$ .

**Remarque 1.7.** Pour chaque  $A \in \mathcal{F}$  il peut exister plusieurs fonctions  $\phi$  qui vérifient (1.83). Évidemment, si  $\phi$  et  $\phi'$  sont deux telles fonctions, alors

$$\phi(\cdot, A) = \phi'(\cdot, A), \mathbb{P} \circ X^{-1} - \text{p.s.}$$

La fonction d'ensemble  $\mathbb{P}(\cdot \mid X)$  est  $\sigma$ -additive, mais seulement  $\mathbb{P}$ -p.s. par rapport à  $\omega \in \Omega$ ; par conséquent  $\phi(x, \cdot)$  n'est pas une probabilité sur  $\mathcal{F}$ . Toutefois, la probabilité conditionnelle  $\mathbb{P}(\cdot \mid X)$  est régulière, i.e. on peut sélectionner  $\mathbb{P}(A \mid X)$  dans sa classe d'équivalence telle que  $\mathbb{P}(\cdot \mid X)(\omega)$  soient des probabilités sur  $\mathcal{F}$ , sauf les éléments  $\omega \in \Omega$  appartenant à un événement  $\mathbb{P}$ -négligeable. Donc on peut sélectionner une fonction  $\phi(\cdot, A)$  telle que  $\phi(x, \cdot)$  sont probabilités sur  $\mathcal{F}$  sauf les éléments  $x \in \mathbb{R}$  appartenant à un ensemble  $\mathbb{P} \circ X^{-1}$ -négligeable.

Soit  $Y$  une v.a. réelle et deux événements  $A, B \in \mathcal{F}$  tels que  $\mathbb{P}(B) \neq 0$ . On définit alors l'espérance conditionnelle de  $Y$  sachant  $B$  comme suit :

$$\mathbb{E}[Y \mid B] = \frac{\mathbb{E}[Y \mathbb{1}_B]}{\mathbb{P}(B)} = \frac{1}{\mathbb{P}(B)} \int_B Y(\omega) d\mathbb{P}(\omega). \quad (1.85)$$

De la formule précédente, on peut montrer que  $\mathbb{E}[\mathbb{1}_A \mid B] = \mathbb{P}(A \mid B)$ .

De manière générale, on peut définir l'espérance conditionnelle de  $Y$ , sachant  $\mathcal{G}$ , comme une v.a.  $\mathcal{G}$ -mesurable, à valeurs dans  $\overline{\mathbb{R}}_+$ , notée  $\mathbb{E}(Y \mid \mathcal{G})$ , et telle que

$$\int_B Y d\mathbb{P} = \int_B \mathbb{E}(Y \mid \mathcal{G}) d\mathbb{P}, \text{ pour tout } B \in \mathcal{G}.$$

Évidemment

$$\mathbb{P}(A \mid \mathcal{G}) = \mathbb{E}(\mathbb{1}_A \mid \mathcal{G}), \mathbb{P} - \text{p.s.}$$

**Proposition 1.10** (Propriétés).

Soient  $X$  et  $Y$  deux v.a. intégrables et  $\mathcal{G}$  une  $\sigma$ -sous-tribu de  $\mathcal{F}$  et  $a, b, c \in \mathbb{R}$ . Alors (toutes les égalités et inégalités sont au sens p.s.) :

1.  $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\mathbb{E}[X \mid \mathcal{G}] + b\mathbb{E}[Y \mid \mathcal{G}]$  ;
2.  $\mathbb{E}[\mathbb{E}(X \mid \mathcal{G})] = \mathbb{E}(X)$  ;
3. si  $X \in \mathcal{G}$ , alors  $\mathbb{E}(X \mid \mathcal{G}) = X$ , et si  $X = c$  on a  $\mathbb{E}(c \mid \mathcal{G}) = c$  ;
4.  $\mathbb{E}(X \mid \{\emptyset, \Omega\}) = \mathbb{E}(X)$  ;
5. si  $X \geq 0$  alors  $\mathbb{E}(X \mid \mathcal{G}) \geq 0$  ; et si  $X \geq Y$  p.s. alors  $\mathbb{E}(X \mid \mathcal{G}) \geq \mathbb{E}(Y \mid \mathcal{G})$  ;
6. si  $X$  et  $\mathcal{G}$  sont indépendantes, alors  $\mathbb{E}(X \mid \mathcal{G}) = \mathbb{E}(X)$  ;
7. si  $Y \in \mathcal{G}$  alors  $\mathbb{E}(h(Y)X \mid \mathcal{G}) = h(Y)\mathbb{E}(X \mid \mathcal{G})$  ;
8. si  $\mathcal{H}$  est indépendante de  $\sigma(\sigma(X), \mathcal{G})$ , alors

$$\mathbb{E}(X \mid \sigma(\mathcal{G}, \mathcal{H})) = \mathbb{E}(X \mid \mathcal{G}) ;$$

9.  $|\mathbb{E}(X \mid \mathcal{G})| \leq \mathbb{E}(|X| \mid \mathcal{G})$ .

Soient  $(E, \mathcal{E})$  un espace mesurable,  $Y: (\Omega, \mathcal{F}) \rightarrow (E, \mathcal{E})$  et  $X: \Omega \rightarrow \mathbb{R}$  un v.a. Une fonction  $Q: \mathbb{R} \times \mathcal{E} \rightarrow [0, 1]$  est appelée *loi conditionnelle de Y sachant X* si

1.  $Q(\cdot, B)$  est fonction borélienne pour tout  $B \in \mathcal{E}$  ;
2.  $Q(x, \cdot)$  est une probabilité sur  $\mathcal{E}$  pour tout  $x \in \mathbb{R}$  ;
3. Pour tout  $A \in \mathcal{B}(\mathbb{R})$  et tout  $B \in \mathcal{E}$  on a

$$\mathbb{P}(Y \in B, X \in A) = \int_A Q(x, B)(\mathbb{P} \circ X^{-1})(dx) = \int_A Q(x, B)F_X(dx). \quad (1.86)$$

Pour énumérer les propriétés principales de la loi conditionnelle on considère la probabilité  $\mathbf{Q} = \mathbb{P} \circ (X, Y)^{-1}$  sur  $\mathcal{B}(\mathbb{R}) \times \mathcal{E}$  et une fonction  $\mathbf{Q}$ -intégrable  $h: (\mathbb{R} \times E, \mathcal{B}(\mathbb{R}) \times \mathcal{E}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ . Alors on a

$$\mathbb{P}(Y^{-1}(B) \mid X = x) = Q(x, B), \quad \mathbb{P} \circ X^{-1} - \text{p.s. pour tout } B \in \mathcal{E}; \quad (1.87)$$

$$\mathbb{E}_{\mathbf{P}}[h(X, Y) \mid X] = \left( \int_E h(\cdot, y) Q(\cdot, dy) \right) \circ X, \quad \mathbb{P} - \text{p.s.} \quad (1.88)$$

$$\mathbb{E}_{\mathbf{P}}[h(X, Y) \mid X = x] = \int_E h(x, y) Q(x, dy), \quad \mathbb{P} \circ X^{-1} - \text{p.s.} \quad (1.89)$$

$$\mathbb{E}_{\mathbf{P}}[h(X, Y)] = \int_{\mathbb{R}} \mathbb{E}_{\mathbf{P}}[h(x, Y) \mid X = x](\mathbb{P} \circ X^{-1})(dx) \quad (1.90)$$

Si le vecteur  $(X, Y)$  dans  $\mathbb{R}^2$  a pour densité  $f(x, y)$  alors

$$\mathbb{E}(X \mid Y) = \varphi(Y),$$

avec

$$\varphi(y) = \frac{\int_{\mathbb{R}} xf(x, y)dx}{\int_{\mathbb{R}} f(x, y)dx}.$$

► **Exemple 1.35** Soient  $X$  et  $Y$  deux v.a. réelles avec les fonction de répartitions et les densités de probabilité  $F_1(x)$  et  $p_1(x)$ , respectivement,  $F_2(y)$  et  $p_2(y)$ . On considère les répartitions conditionnelles

$$Q_1(x, A) = \mathbb{P}(Y \in A \mid X = x), \quad x \in \mathbb{R}, A \in \mathcal{B}(\mathbb{R}),$$

$$Q_2(y, B) = \mathbb{P}(X \in B \mid Y = y), \quad y \in \mathbb{R}, B \in \mathcal{B}(\mathbb{R})$$



De (1.86) on a

$$\mathbb{P}(X \in B, Y \in A) = \int_B Q_1(x, A) F_1(dx) = \int_A Q_2(y, B) F_2(dy), \quad A, B \in \mathcal{B}(\mathbb{R})$$

Donc, si  $p(x, y)$  est la densité jointe du vecteur  $(X, Y)$ , on voit que

$$\begin{aligned} Q_1(x, A) p_1(x) &= \int_A p(x, y) dy, \quad \mathbb{P} \circ X^{-1} - \text{p.s.} \\ Q_2(y, B) p_2(y) &= \int_B p(x, y) dx, \quad \mathbb{P} \circ Y^{-1} - \text{p.s.} \end{aligned}$$

C'est pourquoi les fonctions

$$\begin{aligned} p_1(y \mid X = x) &= \begin{cases} \frac{p(x, y)}{p_1(x)} & \text{pour } p_1(x) \neq 0 \\ p_0(y) & \text{pour } p_1(x) = 0 \end{cases} \\ p_2(x \mid Y = y) &= \begin{cases} \frac{p(x, y)}{p_2(y)} & \text{pour } p_2(y) \neq 0 \\ p'_0(x) & \text{pour } p_2(y) = 0 \end{cases} \end{aligned}$$

où  $p_0$  et  $p'_0$  sont densités de probabilités arbitraires nulles sur les ensembles  $\{x \in \mathbb{R} : p_1(x) = 0\}$  respectivement  $\{y \in \mathbb{R} : p_2(y) = 0\}$ , sont appelées *densités conditionnelles*.

◁

## 1.7 Suites aléatoires et théorèmes limites

Soient  $(\Omega, \mathcal{F}, \mathbb{P})$  un espace de probabilité et  $(A_n)$  une suite d'événements de  $\mathcal{F}$ .

*Conjonction et disjonction infinie.*

$$\begin{aligned} \bigcup_{n \geq 1} A_n &= \{\omega \in \Omega : \omega \in A_n \text{ pour au moins un certain } n\}, \\ \bigcap_{n \geq 1} A_n &= \{\omega \in \Omega : \omega \in A_n \text{ pour tout } n\}. \end{aligned}$$

*Suites monotones d'événements.*

**Théorème 1.11.** 1. Si  $(A_n, n \in \mathbb{N})$  est une suite croissante d'événements, i.e.,  $A_1 \subset A_2 \subset \dots \subset A_n \subset \dots$ , alors  $A = \lim_{n \rightarrow \infty} A_n = \bigcup_{n=1}^{\infty} A_n$  et :  $\mathbb{P}(A) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$ .

2. Si  $(A_n, n \in \mathbb{N})$  est une suite décroissante, i.e.,  $A_1 \supset A_2 \supset \dots \supset A_n \supset \dots$ , alors  $A = \lim_{n \rightarrow \infty} A_n = \bigcap_{n=1}^{\infty} A_n$  et :  $\mathbb{P}(A) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$ .

*Limite sup et limite inf.* Soit une suite d'événements,  $(A_n, n \in \mathbb{N})$ , non nécessairement monotone.

— L'ensemble des  $\omega \in \Omega$  appartenant à une infinité de  $A_n$ , soit

$$\limsup_{n \rightarrow \infty} A_n = \overline{\lim}_{n \rightarrow \infty} A_n = \bigcap_{n \geq 0} \bigcup_{k \geq n} A_k = \{\omega \in \Omega : \forall n, \exists k \geq n, \omega \in A_k\}.$$

est appelé *limite supérieure* des  $(A_n)$ .

— L'ensemble des  $\omega \in \Omega$  appartenant à tous  $A_n$  sauf peut-être un nombre fini, soit

$$\liminf_{n \rightarrow \infty} A_n = \underline{\lim}_{n \rightarrow \infty} A_n = \bigcup_{n \geq 0} \bigcap_{k \geq n} A_k = \{\omega \in \Omega : \exists n, \forall k \geq n, \omega \in A_k\}.$$

est appelé *limite inférieure* des  $(A_n)$ .

**Théorème 1.12.** (Lemme de Borel-Cantelli) *Soit  $(A_n)$  une suite d'événements de  $\mathcal{F}$ .*

1. *Si  $\sum_{n \geq 0} \mathbb{P}(A_n) < \infty$  alors  $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0$ .*
2. *Si  $\sum_{n \geq 0} \mathbb{P}(A_n) = \infty$  et la suite  $(A_n)$  est indépendante alors  $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 1$ .*

Soient  $(X_n)$  une suite de v.a. et  $X$  une v.a., toutes définies sur le même espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ .

**Définition 1.12.** La convergence de la suite  $(X_n)$  vers  $X$  est définie comme suit :

1. *Presque sûrement  $(X_n \xrightarrow{\text{p.s.}} X)$  si  $\mathbb{P}(\{\omega \in \Omega : X_n(\omega) \rightarrow X(\omega)\}) = 1$ .*
2. *En probabilité  $(X_n \xrightarrow{\mathbb{P}} X)$  si pour tout  $\varepsilon > 0$ ,*

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

3. *En loi ou faible  $(X_n \xrightarrow{\mathcal{L}} X)$  si  $F_{X_n}(x) \rightarrow F_X(x)$  ponctuellement en tout point de continuité de  $F_X$ .*

4. *En moyenne d'ordre  $p \in \mathbb{N}^*$   $(X_n \xrightarrow{L^p} X)$  si  $X_n \in L^p$  pour tout  $n \in \mathbb{N}^*$ , si  $X \in L^p$  et si  $\lim_{n \rightarrow \infty} \mathbb{E}|X_n - X|^p = 0$ . Les plus utilisées sont les cas  $p = 1$  (convergence en moyenne) et  $p = 2$  (convergence en moyenne quadratique).*

Les liens entre ces types de convergences s'expriment sous forme

$$\text{p.s.} \Rightarrow \mathbb{P} \Rightarrow \mathcal{L}, \quad L^p \Rightarrow \mathbb{P}, \quad L^{p'} \Rightarrow L^p \text{ pour } p' \geq p \quad (1.91)$$

La convergence en loi des v.a. est une propriété de convergence de leurs lois et elle est bien souvent utilisée dans la théorie des probabilités. Cette convergence peut être exprimée avec les fonctions caractéristiques.

**Théorème 1.13.** (le théorème de continuité de P. Lévy) *La suite  $(X_n)$  converge en loi vers  $X$  si et seulement si la suite des fonctions caractéristiques de  $(X_n)$  converge ponctuellement vers la fonction caractéristique de  $X$ .*

▷ **Exemple 1.36** Soit  $X_1, X_2, \dots$  v.a. indépendantes tel que  $\mathbb{P}(X_n = 1) = 1 - \frac{1}{n}$  et  $\mathbb{P}(X_n = n) = \frac{1}{n}$ ,  $n > 1$ . On a  $X_n \xrightarrow{\mathbb{P}} 1$ ,  $n \rightarrow \infty$ , mais  $X_n$  ne converge pas presque sûrement vers 1 pour  $n \rightarrow \infty$ .

En effet, pour tout  $\varepsilon > 0$

$$\mathbb{P}(|X_n - 1| > \varepsilon) = \mathbb{P}(X_n = n) = \frac{1}{n} \rightarrow 0, \quad n \rightarrow \infty.$$

Donc  $X_n \xrightarrow{\mathbb{P}} 1$ ,  $n \rightarrow \infty$ . Pour étudier la convergence p.s. nous rappelons la condition suivante :  $X_n \xrightarrow{\text{a.s.}} X$ ,  $n \rightarrow \infty$ , si et seulement si, pour tout  $\varepsilon > 0$  et  $0 < \delta < 1$ , il existe  $n_0$  tel que pour tout  $n > n_0$

$$\mathbb{P}\left(\bigcap_{m > n} \{|X_m - X| < \varepsilon\}\right) > 1 - \delta. \quad (1.92)$$

Comme pour tous  $\varepsilon > 0$ ,  $\delta \in (0, 1)$  et  $N > \frac{n}{1-\delta} > n$ , on a

$$\begin{aligned} \mathbb{P} \left( \bigcap_{m>n} \{|X_m - 1| < \varepsilon\} \right) &\leq \mathbb{P} \left( \bigcap_{m=n+1}^N \{|X_m - 1| < \varepsilon\} \right) \\ &= \prod_{m=n+1}^N \mathbb{P}(|X_m - 1| < \varepsilon) = \prod_{m=n+1}^N \mathbb{P}(X_m = 1) \\ &= \prod_{m=n+1}^N \left(1 - \frac{1}{m}\right) = \frac{n}{N} < 1 - \delta. \end{aligned}$$

Il n'existe pas  $n_0$  pour lequel la relation (1.92) soit vérifiée, d'où il en résulte que  $X_n$  ne converge pas p.s.

◁

▷ **Exemple 1.37** Soit  $(\Omega, \mathcal{F}, \mathbb{P})$  où  $\Omega = [0, 1]$ ,  $\mathcal{F} = \mathcal{B}([0, 1])$  et  $\mathbb{P}$  la mesure de Lebesgue sur  $\mathbb{R}$ . Soit la suite de v.a.  $(X_n, n \in \mathbb{N}^*)$  définie par

$$X_n(\omega) = \begin{cases} 0, & 0 \leq \omega < \frac{1}{2} + \frac{1}{n} \\ 1, & \frac{1}{2} + \frac{1}{n} \leq \omega \leq 1 \end{cases}$$

et la v.a.

$$X(\omega) = \begin{cases} 1, & 0 \leq \omega < \frac{1}{2} \\ 0, & \frac{1}{2} \leq \omega \leq 1. \end{cases}$$

On a

$$|X_n(\omega) - X(\omega)| = \begin{cases} 1, & 0 \leq \omega < \frac{1}{2} \\ 0, & \frac{1}{2} \leq \omega < \frac{1}{2} + \frac{1}{n} \\ 1, & \frac{1}{2} + \frac{1}{n} \leq \omega \leq 1 \end{cases}$$

d'où il en résulte  $\mathbb{P}(|X_n - X| = 1) = 1 - \frac{1}{n} \rightarrow 1$  pour  $n \rightarrow \infty$  et  $(X_n)$  ne converge pas en probabilité. D'autre part, les fonctions de répartition des v.a.  $X_n, n \in \mathbb{N}^*$ , et  $X$  sont

$$F_{X_n}(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2} + \frac{1}{n}, & 0 \leq x < 1 \\ 1, & x \geq 1 \end{cases}$$

et respectivement

$$F_X(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & 0 \leq x < 1 \\ 1, & x \geq 1, \end{cases}$$

donc  $X_n \xrightarrow{\mathcal{L}} X$  pour  $n \rightarrow \infty$ .

◁

**Théorème 1.14.** (Slutsky) Soient deux suites aléatoires  $(X_n, n \geq 1)$  et  $(Y_n, n \geq 1)$  telles que, lorsque  $n \rightarrow \infty$ ,

$$X_n \xrightarrow{\mathcal{L}} X, \quad \text{et} \quad Y_n \xrightarrow{P} C,$$

où  $X$  est une v.a. et  $C$  une constante. Alors

1.  $X_n \pm Y_n \xrightarrow{\mathcal{L}} X \pm C$  ;
2.  $X_n \cdot Y_n \xrightarrow{\mathcal{L}} X \cdot C$  ;

3.  $\frac{X_n}{Y_n} \xrightarrow{\mathcal{L}} \frac{X}{C}$ , sous réserve que  $C \neq 0$ .

Considérons une suite de v.a.  $(X_n, n \geq 1)$  et supposons que les espérances mathématiques  $a_n = \mathbb{E}(X_n)$ ,  $n = 1, 2, \dots$  existent. Soit

$$Y_n = \frac{(X_1 - a_1) + \dots + (X_n - a_n)}{n}, \quad n = 1, 2, \dots$$

**Définition 1.13.** On dit que la suite  $(X_n, n \geq 1)$  vérifie la loi faible ou forte des grands nombres si  $Y_n \xrightarrow{\mathbf{P}} 0$  ou  $Y_n \xrightarrow{\text{a.s.}} 0$ , respectivement.

Évidemment, si une suite vérifie la loi forte des grands nombres, alors elle vérifie la loi faible également. L'appellation traditionnelle "loi des grands nombres" qui règne encore aujourd'hui ne doit pas tromper : il s'agit d'un "vrai" théorème de mathématiques !

Nous allons donner quelques "lois classiques des grands nombres".

**Théorème 1.15.** (Markov) Soit  $(X_n, n \geq 1)$  une suite de v.a. tel que  $\mathbb{E}(X_i) = a_i$ ,  $\text{Var}(X_i) = \sigma_i^2$ ,  $i = 1, 2, \dots$ . Si

$$\lim_{n \rightarrow \infty} \frac{1}{n^2} \text{Var} \left( \sum_{k=1}^n X_k \right) = 0, \quad (1.93)$$

alors  $(X_n, n \geq 1)$  vérifie la loi faible des grands nombres.

*Démonstration.* Si  $Y_n$  est comme dans la définition 1.13, alors de l'inégalité de Tchebychev on a

$$\mathbb{P}(|Y_n| > \varepsilon) \leq \frac{1}{\varepsilon^2 n^2} \mathbb{E} \left( \sum_{k=1}^n X_k - \sum_{k=1}^n a_k \right)^2 = \frac{1}{\varepsilon^2 n^2} \text{Var} \left( \sum_{k=1}^n X_k \right)$$

d'où il en résulte

$$\lim_{n \rightarrow \infty} \mathbb{P}(|Y_n| > \varepsilon) = 0.$$

□

**Corollaire 1.16.** (Tchebychev) Soit  $(X_n, n \geq 1)$  une suite de v.a. indépendantes telle que  $\mathbb{E}(X_i) = a_i$ ,  $\text{Var}(X_i) = \sigma_i^2$ ,  $i = 1, 2, \dots$ . S'il existe  $0 < M < \infty$  tel que  $\sigma_i^2 \leq M$  pour tous  $i = 1, 2, \dots$ , alors  $(X_n, n \geq 1)$  vérifie la loi faible des grands nombres.

*Démonstration.* Dans ce cas

$$\text{Var} \left( \sum_{k=1}^n X_k \right) = \sum_{k=1}^n \text{Var}(X_k) \leq M n$$

et la condition (1.93) est vérifiée.

□

**Théorème 1.17.** (Kolmogorov) Soit  $(X_n, n \geq 1)$  une suite de v.a. indépendantes, telles que  $\mathbb{E}(X_i) = a_i$ ,  $\text{Var}(X_i) = \sigma_i^2$ ,  $i = 1, 2, \dots$ . Si  $\sum_{n \in \mathbf{N}^*} \frac{\sigma_n^2}{n^2} < \infty$ , alors  $(X_n, n \geq 1)$  vérifie la loi forte des grands nombres.

**Corollaire 1.18.** Dans les conditions du corollaire 1.16, la suite de v.a.  $(X_n, n \geq 1)$  vérifie la loi forte des grands nombres.

▷ **Exemple 1.38** Soit  $(X_n)$  une suite aléatoire i.i.d. de loi  $P$ . Pour tout borélien  $B$ , les v.a.  $\mathbb{1}_{(X_n \in B)}$  sont indépendantes et de même loi ayant l'espérance  $P(B)$ . Par la loi des grands nombres, corollaire 1.18, on a

$$\mathbb{P}(B) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n \mathbb{1}_{(X_k \in B)} = \lim_{n \rightarrow \infty} \frac{\text{nombre de } k \leq n \text{ tels que } X_k \in B}{n}.$$

Ceci justifie l'estimation des probabilités par les fréquences.

◁

▷ **Exemple 1.39** Soit  $(X_n)$  une suite de v.a. à valeurs  $\pm\sqrt{k}$ ,  $k \in \mathbb{N}$  avec les probabilités  $\mathbb{P}(X_1 = 0) = 1$  et

$$\mathbb{P}(X_k = \sqrt{k}) = \mathbb{P}(X_k = -\sqrt{k}) = \frac{1}{k}, \quad \mathbb{P}(X_k = 0) = 1 - \frac{2}{k}, \quad k = 2, 3, \dots$$

Comme  $\mathbb{E}(X_k) = 0$ ,  $k = 1, 2, \dots$  et  $\text{Var}(X_k) = 2$ ,  $k = 2, 3, \dots$  la suite  $(X_n)$  vérifie la loi forte des grands nombres, conformément au corollaire 1.18.

◁

▷ **Exemple 1.40** Soit  $(X_n)$  une suite de v.a. à valeurs  $\pm\sqrt{\ln k}$ ,  $k \in \mathbb{N}^*$ , avec les probabilités  $\mathbb{P}(X_1 = 0) = 1$  et

$$\mathbb{P}(X_k = \sqrt{\ln k}) = \mathbb{P}(X_k = -\sqrt{\ln k}) = \frac{1}{2}, \quad k = 2, 3, \dots$$

Comme  $\mathbb{E}(X_k) = 0$ , et  $\text{Var}(X_k) = \ln k$ ,  $k = 2, 3, \dots$  il en résulte

$$\begin{aligned} \text{Var} \left( \sum_{k=1}^n X_k \right) &= \sum_{k=1}^n \ln k < \sum_{k=1}^n \int_k^{k+1} \ln x dx = \int_1^{n+1} \ln x dx \\ &= (x \ln x - x) \Big|_1^{n+1} = (n+1) \ln(n+1) - n \end{aligned}$$

et

$$\lim_{n \rightarrow \infty} \frac{(n+1) \ln(n+1) - n}{n^2} = 0.$$

Par conséquent les conditions de théorème 1.15 sont satisfaites et la suite  $(X_n)$  vérifie la loi faible des grands nombres.

◁

Dans les applications on a souvent affaire à des v.a. composées d'un grand nombre d'éléments indépendants. Il se trouve que pour des conditions très générales de telles variables sommes sont normalement réparties. Ce fait fondamental s'expriment mathématiquement dans l'une des formes (suivant les conditions) du *théorème de la limite centrale*. Pour énoncer un théorème de la limite centrale nous allons considérer une suite de v.a.  $(X_n, n \geq 1)$  et nous adaptons les notations suivantes

$$a_k = \mathbb{E}(X_k), \quad \sigma_k^2 = \text{Var}(X_k), \quad d_n^2 = \sum_{k=1}^n \sigma_k^2.$$

Un théorème de la limite centrale donne des conditions suffisantes pour que la v.a.

$$Z_n = \frac{1}{d_n} \sum_{k=1}^n (X_k - a_k)$$

converge en loi vers la loi normale  $N(0, 1)$ .

On voit que, si les v.a.  $(X_n, n \geq 1)$  sont indépendantes, alors  $d_n^2 = \text{Var} \left( \sum_{k=1}^n X_k \right)$  et les v.a. sont *centrées réduites* parce que  $\mathbb{E}(Z_n) = 0$ ,  $\text{Var}(Z_n) = 1$ .

Voici un théorème de la limite centrale important en probabilités et statistique mathématique.

**Théorème 1.19.** *Soit une suite  $(X_n, n \geq 1)$  de v.a. indépendantes, possédant toutes la même loi d'espérance mathématique  $a$  et de variance  $\sigma^2 < \infty$ . Alors la suite de v.a.*

$$Z_n = \frac{\sqrt{n}}{\sigma} \left( \frac{X_1 + \cdots + X_n}{n} - a \right), \quad n = 1, 2, \dots$$

*converge en loi vers la loi normale  $N(0, 1)$ . On dit que  $Z_n$  est asymptotiquement normale  $N(0, 1)$  pour  $n \rightarrow \infty$ .*

*Démonstration.* Si  $\varphi(t)$  est la fonction caractéristique de  $(X_1 - a)$  (donc de  $(X_n - a)$  pour tout  $n \geq 1$ ), alors la fonction caractéristique de  $Z_n$  sera

$$\begin{aligned} \varphi_n(t) &= \left[ \varphi \left( \frac{t}{\sigma\sqrt{n}} \right) \right]^n = \left( 1 + \frac{t}{\sigma\sqrt{n}} \varphi'(0) + \frac{t^2}{2n\sigma^2} \varphi''(0) + o(1/n) \right)^n \\ &= \left( 1 - \frac{1}{2n} t^2 + o(1/n) \right)^n. \end{aligned}$$

Donc  $\lim_{n \rightarrow \infty} \varphi_n(t) = e^{-\frac{t^2}{2}}$  et  $Z_n \xrightarrow{\mathcal{L}} N(0, 1)$  conformément au théorème 1.13.  $\square$

**Corollaire 1.20.** (Moivre-Laplace) *Si  $X \sim b(n, p)$  et  $q = 1 - p$ , alors  $Z_n = \frac{X - np}{\sqrt{npq}}$  est asymptotiquement normale  $N(0, 1)$  pour  $n \rightarrow \infty$ .*

*Démonstration.* C'est une conséquence du théorème 1.19 parce que  $X = X_1 + \cdots + X_n$  où  $(X_n, n \geq 1)$  sont i.i.d.,  $X_n \sim B(p)$ .  $\square$

▷ **Exemple 1.41** On lance un dé  $n$  fois et on considère la v.a.  $N$  = nombre de six. À partir de quelle valeur de  $n$  aura-t-on 9 chances sur 10 d'avoir  $|\frac{N}{n} - \frac{1}{6}| < 0,01$  ?

La variable  $N$  est une v.a. binomiale  $b(n, 1/6)$ . Pour  $n$  assez grand on peut utiliser le corollaire 1.20 pour approcher la v.a.  $Z_n = \frac{N - n/6}{\sqrt{5n/6}}$  par une v.a. normale  $N(0, 1)$ . Alors la contrainte

$|\frac{N}{n} - \frac{1}{6}| < 0,01$  peut se réécrire  $|Z_n| < \frac{0,06\sqrt{n}}{\sqrt{5}}$ . Par ailleurs de  $\Phi(x) - \Phi(-x) = 0,9$  ou  $2\Phi(x) - 1 = 0,9$  on obtient  $x = 1,645$  (tables de la loi normale). On aura donc 9 chances sur 10 d'avoir  $|\frac{N}{n} - \frac{1}{6}| < 0,01$  si l'inégalité  $\frac{0,06\sqrt{n}}{\sqrt{5}} \geq 1,645$  sera vérifiée, i.e. si

$$n \geq \left( \frac{1,645\sqrt{5}}{0,06} \right)^2 = 3758,4.$$

◁

▷ **Exemple 1.42** La durée de vie d'un composant est une v.a. qui suit une loi exponentielle de paramètre  $\lambda = 0,2 \times 10^{-3} h^{-1}$ . Si l'on remplace un composant par un composant neuf dès qu'il tombe en panne, quelle est la probabilité qu'au bout de 50 000 heures le composant en fonctionnement soit au moins la dixième ?

Désignons par  $S_n$  la somme des  $n$  variables aléatoires exponentielles (indépendantes) qui représentent les durées de vie des  $n$  premiers composants. D'après les caractéristiques de la loi exponentielle on a  $\mathbb{E}(S_n) = \frac{n}{\lambda}$  et  $\text{Var}(S_n) = \frac{n}{\lambda^2}$  et on s'autorise du théorème 1.19 d'approcher  $\frac{\lambda S_n - n}{\sqrt{n}}$  par une variable normale centrée réduite, ce qui est exactement équivalent à dire que l'on peut approcher  $S_n$  par une variable normale  $N(n/\lambda, n/\lambda^2) = N(5000n, 5000^2n)$ . Énoncer qu'au bout de 50 000 heures le composant en fonctionnement est au moins la dixième revient à dire que l'on a  $S_9 < 50000$ , et il suffit maintenant de calculer  $\mathbb{P}(S_9 < 50000^2)$ , avec  $S_9 \sim N(45000, 15000^2)$ . On a donc la probabilité demandée

$$\mathbb{P}(S_9 < 50000) = \Phi\left(\frac{50000 - 45000}{15000}\right) = \Phi(0,333) = 0,6306.$$

◁

▷ **Exemple 1.43** Un restaurant peut servir 75 repas. La pratique montre que 20% des clients ayant réservé ne viennent pas.

a) Le restaurateur accepte 90 réservations. Quelle est la probabilité qu'il se présente plus de 50 clients ?

b) Combien le restaurateur doit-il accepter de réservations pour avoir une probabilité supérieure ou égale à 0,9 de pouvoir servir tous les clients qui se présenteront ?

Si  $n$  est le nombre des clients ayant réservé, la loi du nombre  $N$  de clients qui se présentent suit une loi binomiale de paramètres  $n$  et  $p = 0,8$ . Conformément au corollaire 1.20 on approche cette loi binomiale par une loi normale de même espérance et de même variance.

a) Si  $n = 90$ ,  $\mathbb{E}(N) = np = 72$  et  $\sigma = \sqrt{npq} = 3,795$ . Alors

$$\mathbb{P}(N > 50) = \mathbb{P}\left(\frac{N - 72}{3,795} > \frac{50 - 72}{3,795}\right) = 1 - \Phi(-5,8) = \Phi(5,8) \cong 1.$$

b) Il s'agit maintenant de résoudre en  $n$  l'inéquation  $\mathbb{P}(N \leq 75) \geq 0,9$ . On a, d'une part :

$$\mathbb{P}(N \leq 75) = \mathbb{P}\left(\frac{N - 0,8n}{\sqrt{0,16n}} \leq \frac{75 - 0,8n}{\sqrt{0,16n}}\right)$$

et, d'autre part  $\Phi(1,281) = 0,9$ . On doit donc avoir  $\frac{75 - 0,8n}{\sqrt{0,16n}} \geq 1,281$ , soit  $75 - 0,8n \geq 1,281\sqrt{0,16n}$ , ce qui conduit, en posant  $\sqrt{n} = x$ , à l'inéquation  $0,8x^2 + 0,5124x - 75 \leq 0$ , d'où  $x \leq 9.367503097$ , i.e.,  $n \leq 87$ .

◁

**Théorème 1.21** (Inégalité de Berry-Essen).

Soient  $X_1, X_2, \dots$  i.i.d. telles que  $\mathbb{E}X_1 = 0$  et  $\mathbb{E}X_1^2 =: \sigma^2$ ,  $0 < \sigma^2 < \infty$ , et  $\mathbb{E}|X_1|^3 =: \rho < \infty$ . Posons  $Y_n = (X_1 + \dots + X_n)/n$ ,  $n \geq 1$ , et soit  $F_n$  la f.r. de la v.a.  $Y_n\sqrt{n}/\sigma$ ,  $n \geq 1$ , et  $\Phi$  la f.r. de  $N(0,1)$ . Alors, pour tous  $x \in \mathbb{R}$  et  $n$ , nous avons

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sigma^3\sqrt{n}},$$

avec  $C < 0,4748$ .

**Théorème 1.22** (Loi du logarithme itéré (LLI)).

Soit une suite de v.a.r  $(X_n, n \geq 1)$ , i.i.d. avec  $\mathbb{E}X_1 = 0$  et  $\text{Var}(X_1) = 1$ . Alors

$$\limsup_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = 1, \quad p.s.$$

$$\liminf_{n \rightarrow \infty} \frac{S_n}{\sqrt{2n \log \log n}} = -1, \quad p.s.$$

## 1.8 Problèmes

**1.1.** On lance deux dés. Soient  $A_0$  l'événement "la somme des points est paire",  $A_1$  l'événement "la somme des points est impaire",  $B$  l'événement "la valeur de la différence des points est égale à 4". Combien y a-t-il d'événements élémentaires dans  $A_0 \cap B$ , dans  $A_1 \cap B$ ?

**1.2.** On a toujours  $\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B \cap C)$ ?

**1.3.** Démontrer l'inégalité

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \leq \sum_{i=1}^n \mathbb{P}(A_i).$$

**1.4.** Montrer que si  $F$  et  $G$  sont fonctions de répartition et  $0 \leq \lambda \leq 1$ , alors  $\lambda F + (1 - \lambda)G$  est une fonction de répartition.

**1.5.** On considère une population composée de 48% d'hommes et de 52% de femmes. La probabilité qu'un homme soit daltonien est 0,05 ; qu'une femme soit daltonienne est 0,0025. Quelle proportion de la population est-elle daltonienne ?

**1.6.** Soit  $N$  une v.a. de Poisson de paramètre  $\lambda$ . Montrer que

$$\mathbb{P}(N \leq n) = \frac{1}{n!} \int_{\lambda}^{\infty} e^{-x} x^n dx$$

et donc

$$\mathbb{P}(N > n) = \frac{1}{n!} \int_0^{\lambda} e^{-x} x^n dx.$$

**1.7.** Soit  $X$  une v.a. réelle. A-t-on  $\text{Var}(10 - X) = \text{Var}(X)$ ,  $-\text{Var}(X)$ ,  $10\text{Var}(X)$  ?

**1.8.** On considère deux v.a. indépendantes  $X$ ,  $Y$ , et des lois :

$$\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \frac{1}{3}, \quad \mathbb{P}(Y = 0) = \mathbb{P}(Y = 1) = \frac{1}{2}.$$

On pose  $Z = |X - Y|$ . Déterminer la loi de  $Z$  ; calculer  $\mathbb{E}(Z)$  et  $\text{Var}(Z)$ .

**1.9.** Une urne contient  $a$  boules rouge et  $b$  boules vertes. On tire une boule au hasard puis on la remet dans l'urne en ajoutant  $c$  boules de la même couleur ; on tire enfin une boule de l'urne dans sa nouvelle composition. On appelle  $X_i$ ,  $i = 1, 2$ , la variable de Bernoulli égale à 1 si la boule tirée au  $i$ -ème tirage est rouge, 0 si elle est verte.

Dresser le tableau de la loi de  $X_1$  et  $X_2$ , et des lois marginales. Calculer les espérance et les variances. Calculer la covariance et le coefficient de corrélation de  $X_1$  et  $X_2$ .

**1.10.** Une banque accepte de ses clients des rouleaux de pièces de 2 € sans contrôler le nombre de pièces (en principe 25). On suppose que 3% des rouleaux contiennent 24 pièces, 96% contiennent 25 pièces et 1% contiennent 26 pièces.

a) Soit  $X$  la v.a. égale au nombre de pièces d'un rouleau. Calculer  $\mathbb{E}(X)$  et  $\text{Var}(X)$ .

b) En utilisant le théorème 1.19 calculer les probabilités pour que 400 rouleaux contiennent :

- moins de 10000 pièces,
- moins de 9990 pièces,
- moins de 9980 pièces.

**1.11.** Soit  $p_n(x) = \frac{n}{\pi(1 + n^2 x^2)}$  la densité de probabilité de  $X_n$ ,  $n = 1, 2, \dots$ . Démontrer que  $X_n \xrightarrow{\text{P}} 0$ .

**1.12.** En utilisant le théorème 1.13 montrer que si  $X_n \xrightarrow{\mathcal{L}} X$ , alors  $aX_n + b \xrightarrow{\mathcal{L}} aX + b$  pour tous  $a, b \in \mathbb{R}$ .

**1.13.** Montrer que si  $X_n \xrightarrow{\text{a.s.}} X$  et  $Y_n \xrightarrow{\text{a.s.}} Y$ , alors  $X_n + Y_n \xrightarrow{\text{a.s.}} X + Y$ . Montrer que le résultat similaire est valable pour la convergence en probabilité.



**1.14.** Soit  $(\lambda_n, n \geq 1)$  une suite strictement croissante de nombres strictement positifs qui tend vers l'infini et  $(X_n, n \geq 1)$  une suite de v.a. telle que, pour chaque  $n \geq 1$ , la v.a.  $X_n$  suit une loi de Poisson de paramètre  $\lambda_n$ . Montrer que la suite de v.a.  $Y_n = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}}$  converge en loi vers une loi normale  $N(0, 1)$ .

**1.15.** On lance un dé  $n$  fois et soit  $X_k$  la v.a. qui représente le nombre des points obtenus au  $k$ -ème lancer.

a) Calculer  $\mathbb{E}(X_k)$  et  $\text{Var}(X_k)$ .

b) Soit  $S_n$  la somme des points obtenues. En utilisant le théorème 1.19, calculer  $\mathbb{P}(3450 < S_n < 3550)$  pour  $n = 1000$ .

**1.16.** Démontrer que la densité jointe des statistiques d'ordre de  $n$  v.a. uniformes et indépendantes sur  $[0, a]$  est

$$f(u_1, \dots, u_n) = \frac{n!}{a^n} \mathbb{1}_{\{0 \leq u_1 \leq \dots \leq u_n\}}.$$

**1.17.** (a) Calculer les fonctions caractéristiques de lois Exponentielle,  $\mathcal{E}(1)$ , et de Laplace,  $L(1)$ . (b) Démontrer qu'une v.a. de Laplace peut se représenter par la différence de deux v.a. exponentielles, de paramètre 1, indépendantes. (c) Calculer la fonction caractéristique de la loi de Cauchy,  $C(0, 1)$ , en se basant sur celle de Laplace.

**1.18.** Montrer que la loi de la v.a.  $N_d(0, I)$ , sur  $\mathbb{R}^d$ , est invariante par rapport à toute transformation orthogonale  $\varphi$ . Une telle loi est dite radiale.

## Chapitre 2

# Simulation stochastique et méthode de Monte Carlo

### 2.1 Introduction

Soit une ellipse incluse dans un rectangle illustrés dans la fig. 2.1. Le problème ici consiste à évaluer l'aire de l'ellipse connaissant l'aire du rectangle. Pour cela on peut faire appel à la loi empirique des grands nombres. On suppose que nous soyons capables d'inscrire des points dans le rectangle de manière complètement aléatoire (uniformément réparties). Par la loi empirique des grands nombres, nous avons :

$$\frac{n_E}{n} \rightarrow \frac{\text{aire de l'ellipse}}{\text{aire du rectangle}}, \quad n \rightarrow \infty, \quad (2.1)$$

où  $n$  est le nombre des points total dans le rectangle, et  $n_E$  le nombre de points dans l'ellipse.

Ainsi nous pouvons donner une bonne approximation de l'aire de l'ellipse dès que le nombre de points  $n$  devient grand, par la formule suivante

$$\text{aire de l'ellipse} \cong \frac{n_E}{n} \times \text{aire du rectangle}. \quad (2.2)$$

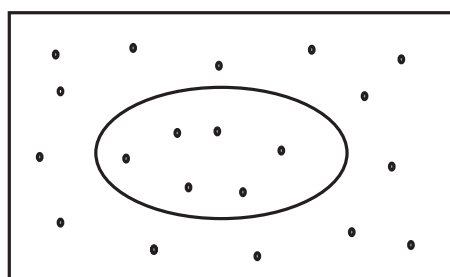


FIGURE 2.1 – Aire d'ellipse par Monte Carlo

Cette méthode de calcul de l'aire de l'ellipse, placée dans un contexte mathématique plus précis, constitue méthode de *Monte Carlo*.

La question qui se pose maintenant concerne la précision de l'approximation (2.2). Par exemple, quelle doit être la valeur minimale de  $n$  pour une précision donnée ? Nous allons voir que pour une bonne approximation, nous avons besoin d'un très grand nombre de points  $n$ ,

et c'est justement le point faible de cette méthode. Néanmoins, nous allons voir qu'il existe de méthodes pour diminuer le nombre de points  $n$  pour une précision donnée.

Un point remarquable qu'on doit souligner dès maintenant est que le calcul de l'aire d'une ellipse, problème complètement déterministe, se transforme à un calcul stochastique (tirage des points aléatoires dans le rectangle).

Afin de pouvoir réaliser la méthode précédente, il faut être capable de générer des points uniformément distribués sur le rectangle. Donc il faut avoir une méthode de génération des variables aléatoires uniformes, et de manière plus générale des v.a. quelconques. C'est ce que nous allons entreprendre dans la section suivante.

## 2.2 Génération de variables aléatoires

Nous allons exposer la génération d'une v.a. uniforme sur l'intervalle  $[0, 1]$ , puis la génération d'une v.a. discrète et enfin d'une v.a. continue quelconques.

### 2.2.1 Suites équiréparties

**Définition 2.1.** Une suite  $(x_n; n \in \mathbb{N}^*)$  à valeurs dans  $[0, 1]$  est dite équirépartie, ou uniformément répartie ou uniformément distribuée, si, pour tous réels  $a, b \in [0, 1]$ , avec  $a < b$ , la suite des entiers  $u_n := \text{Card}\{p : p \leq n, x_p \in [a, b]\}$  est asymptotiquement équivalente à  $n(b - a)$ , lorsque  $n \rightarrow \infty$ .

**Théorème 2.1.** Une suite  $(x_n; n \in \mathbb{N}^*)$  à valeurs dans  $[0, 1]$  est équirépartie ssi pour toute fonction  $f$  Riemann intégrable, on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f(x_k) = \int_0^1 f(x) dx.$$

**Définition 2.2.** Une suite  $(x_n; n \in \mathbb{N}^*)$  à valeurs dans  $\mathbb{R}$  est dite équirépartie modulo 1, ssi la suite de terme général  $x_n - [x_n]$ , à valeurs dans  $[0, 1[$ , est équirépartie.

**Théorème 2.2.** (Critère de Weyl)

Une suite  $(x_n; n \in \mathbb{N}^*)$  est équirépartie modulo 1, ssi pour tout entier  $r \neq 0$ , on a

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n e^{2i\pi r x_k} = 0.$$

### 2.2.2 Génération d'une variable uniforme

La génération ou réalisation des v.a. par des méthodes telles que le jeu des dés, le tirage des cartes ou l'arrêt d'une roue qui tourne n'est plus d'actualité. La génération d'une v.a. se fait actuellement sur ordinateur par une procédure déterministe. Pour cette raison nous appelons ces v.a. *pseudo-aléatoires*.

La méthode la plus utilisée pour générer des v.a. uniforme sur  $[0, 1]$  est basée sur les suites congruentes. Par exemple, partant d'une valeur initiale  $x_0$ , appelée racine, on construit les valeurs  $x_1, x_2, \dots$  par la formule

$$x_n \equiv ax_{n-1} + b \pmod{m}, \quad (2.3)$$

où  $a, b$  et  $m$  sont des constantes entières. Par exemple, nous avons les valeurs  $a = 7^5$ ,  $b = 0$ , et  $m = 2^{31} - 1$  pour un calculateur de 32 bits. Ainsi les valeurs obtenues sont des entiers de 0 à  $m - 1$ . En divisant maintenant ces valeurs par  $m$ , on obtient de valeurs réelles dans  $[0, 1[$ .

Cet algorithme d'obtention de v.a. est appelé *générateur des nombres pseudo-aléatoires*. Il faut noter que la qualité d'un générateur est conditionnée par le choix des nombres  $a$ ,  $b$ ,  $m$ .

*Test d'un générateur.* Nous allons présenter l'utilisation d'un test de  $\chi^2$  pour tester si un générateur est bon ou non, i.e., s'il génère bien des v.a.  $U(0, 1)$ .

On considère une suite de v.a.  $U_k, 1 \leq k \leq N$ , dont les réalisations sont produites par le générateur en question. Cela se fait par l'instruction RAND ou RAN#, etc. (en *Excel* c'est l'instruction ALEA). En suite, on subdivise l'intervalle  $[0, 1]$  en  $r$  (par exemple  $r = 10, 20$ , etc.) sous intervalles réguliers, i.e.,  $I_i = [\frac{i-1}{r}, \frac{i}{r}]$ ,  $1 \leq i \leq r$ . Le nombre de classes doit être en général  $r \cong N^{2/5}$ .

On définit les v.a.

$$N_i := \sum_{k=1}^N \mathbb{1}_{\{U_k \in I_i\}}, \quad 1 \leq i \leq r, \quad (2.4)$$

et le vecteur  $X_r = (\frac{N_1 - N/r}{\sqrt{N/r}}, \dots, \frac{N_r - N/r}{\sqrt{N/r}})$ .

Soit la v.a.  $T_N$  définie comme suit

$$T_N := \|X_r\|^2 = \sum_{i=1}^r \frac{(N_i - N/r)^2}{N/r}, \quad (2.5)$$

où  $1/r$  est la probabilité (uniforme) qu'une réalisation  $U_k(\omega)$  tombe dans un intervalle donné  $I_i$ .

Nous avons le théorème suivant.

**Proposition 2.3.** *La convergence en loi suivante a lieu*

$$T_N \xrightarrow{\mathcal{L}} \chi^2(r-1), \quad N \rightarrow \infty.$$

Donc pour  $N$  assez grand on peut admettre qu'approximativement  $T_N$  suit une loi de chi-deux à  $r-1$  degrés de liberté, notée  $\chi^2(r-1)$ . Pour une réalisation  $t_N$  de  $T_N$ , et pour un niveau  $\alpha$  de risque (en général  $\alpha = 0,05$  ou  $0,01$ ), si  $t_N > \chi_{1-\alpha}^2(r-1)$ , alors on rejette le générateur, et on l'accepte dans le cas contraire.

En fait, la méthode de simulation des v.a. quelconques que nous exposons par la suite, se base sur le théorème fondamental suivant.

**Proposition 2.4.** *Pour toute v.a.  $X$  à valeurs dans  $\mathbb{R}^d$ ,  $d \geq 1$ , il existe une application borélienne  $g : \mathbb{R}^n \rightarrow \mathbb{R}^d$ , et un vecteur  $(U_1, \dots, U_n)$  uniformément repartie sur le cube  $[0, 1]^n$ , tel que la v.a.  $Y := g(U_1, \dots, U_n)$  suit la même loi que  $X$ . En plus ceci peut être réalisé en imposant  $n = 1$  ou  $n = d$ , ou encore  $n \geq 1$  quelconque donné.*

### 2.2.3 Génération des variables aléatoires discrètes

Le problème ici consiste à générer une v.a.d. à partir des réalisations des v.a. uniformes sur  $[0, 1]$ , i.e., mettre en application la proposition 2.4.

*Variable aléatoire de Bernoulli.* Soit  $X$  une v.a. de Bernoulli,  $X \sim B(p)$ ,  $p \in [0, 1]$ , i.e.,

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

Soit une v.a.  $\xi \sim U(0, 1)$ . Alors la v.a.  $Y = \mathbb{1}_{(\xi \leq p)}$  suit une loi  $B(p)$ . En effet,

$$\mathbb{P}(Y = 1) = \mathbb{P}(\mathbb{1}_{(\xi \leq p)} = 1) = \mathbb{P}(\xi \leq p) = p.$$

*Variable aléatoire géométrique* sur  $\mathbb{N}^*$ . Soit  $X$  une v.a. de loi géométrique,  $X \sim G(p)$ ,  $p \in [0, 1]$ , i.e.,

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

Soit  $\xi_1, \xi_2, \dots$  des v.a. i.i.d de loi  $U(0, 1)$ . Si  $\xi_1 > 1 - p, \dots, \xi_{k-1} > 1 - p$  et  $\xi_k \leq 1 - p$ , alors la réalisation de  $X$  sera égale à  $k$ .

*Variable aléatoire discrète quelconque.* Soit une v.a.d.  $X$  à valeurs dans  $E = \{x_1, x_2, \dots, x_n, \dots\}$  et de loi  $p = (p_1, p_2, \dots, p_n, \dots)$ , i.e.,  $p_i = \mathbb{P}(X = x_i)$ . Soit maintenant une v.a.  $\xi \sim \mathcal{U}(0, 1)$ . Alors la v.a.  $Y$  définie par

$$Y = \sum_{i=1}^{\infty} x_i \mathbb{1}_{(p_0 + p_1 + \dots + p_{i-1} \leq \xi < p_1 + \dots + p_{i-1} + p_i)}, \quad (2.6)$$

suit la loi  $p$ , où  $p_0 = 0$ .

### 2.2.4 Génération des variables aléatoires réelles

Nous allons présenter deux méthodes générales pour la génération des v.a.r. appelées *méthode de la fonction inverse* et *méthode de rejet de von Neumann*.

#### Méthode de la fonction inverse

Elle se base sur les résultats suivants.

**Proposition 2.5.** *Soit une v.a.r.  $X$  de f.r.  $F$  continue. Alors la v.a.*

$$Y := F(X) \sim \mathcal{U}(0, 1).$$

*Démonstration.* Comme

$$\{X \leq x\} \subset \{F(X) \leq F(x)\},$$

et

$$\{X > x\} \cap \{F(X) < F(x)\} = \emptyset,$$

on obtient :

$$\begin{aligned} \mathbb{P}(F(X) \leq F(x)) &= \mathbb{P}(F(X) \leq F(x), X \leq x) + \mathbb{P}(F(X) \leq F(x), X > x) \\ &= \mathbb{P}(X \leq x) + \mathbb{P}(F(X) = F(x), X > x) \\ &= \mathbb{P}(X \leq x). \end{aligned}$$

Comme  $F$  est continue, pour tout  $y \in ]0, 1[$ , il existe un  $x \in \mathbb{R}$  tel que  $F(x) = y$ . Alors, par la relation précédente, nous avons :

$$\mathbb{P}(Y \leq y) = \mathbb{P}(F(X) \leq F(x)) = F(x) = y,$$

d'où la conclusion. □

**Proposition 2.6.** *Soit une f.r.  $F$ . Définissons l'inverse généralisée  $F^{-1} : ]0, 1[ \rightarrow \mathbb{R}$ , par*

$$F^{-1}(y) := \inf\{x : F(x) \geq y\}, \quad 0 < y < 1.$$

*Si  $U \sim \mathcal{U}(0, 1)$ , alors la v.a.  $X := F^{-1}(U)$  suit une loi de f.r.  $F$ .*

*Démonstration.* Montrons que, pour tout  $x \in \mathbb{R}$ , tel que  $F(x) \in ]0, 1[$ , et pour tout  $y \in ]0, 1[$ ,  $F(x) \geq y$  si et seulement si  $x \geq F^{-1}(y)$ .

Supposons  $x \geq F^{-1}(y) = \inf\{x : F(x) \geq y\}$ ,  $0 < y < 1$ . Alors, comme  $F$  est croissante et continue à droite,  $\{x : F(x) \geq y\}$  est un intervalle contenant le point extrême gauche. Et donc  $F(x) \geq y$ . Maintenant, supposons  $F(x) < y$ . Alors  $x < \inf\{x : F(x) \geq y\} = F^{-1}(y)$ .

En conséquence

$$\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \leq F(x)) = F(x),$$

ce qui achève la démonstration. □

### Méthode de rejet de von Neumann

Le problème se pose comme suit. On cherche à réaliser une v.a.  $X$  de densité  $f$ , connaissant des réalisations d'une autre v.a.  $Y$  de densité  $g$ .

On suppose que  $\text{supp}(f) \subset \text{supp}(g)$  et que  $f(x)/g(x) \leq C$  pour tout  $x$ ;  $C$  étant une constante. L'algorithme proposé est le suivant :

1. Simuler  $Y$  de  $g$  et  $U$  uniforme sur  $(0, 1)$ ; soient  $Y(\omega)$  et  $U(\omega)$ .
2. Si  $U(\omega) \leq f(Y(\omega))/Cg(Y(\omega))$ , alors  $X(\omega) = Y(\omega)$ ; sinon continue à l'étape 1.

*Démonstration.* Pour  $h$  mesurable bornée on a :

$$\begin{aligned} \mathbb{E}[h(Y) \mid U \leq f(Y)/Cg(Y)] &= \frac{\mathbb{E}[h(Y)\mathbb{1}_{\{U \leq f(Y)/Cg(Y)\}}]}{\mathbb{P}(U \leq f(Y)/Cg(Y))} \\ &= K^{-1} \int \int_{\mathbb{R}^2} h(y)\mathbb{1}_{\{u \leq f(y)/Cg(y)\}}g(y)\mathbb{1}_{[0,1]}(u)dudy \\ &= (CK)^{-1} \int_{\mathbb{R}} h(y)f(y)dy. \end{aligned}$$

Pour  $h = 1$  on obtient  $CK = 1$ , et la démonstration est achevée.  $\square$

### Réalisation des v.a. normales

La méthode de Box-Muller nous fournit des v.a.  $Z_1$  et  $Z_2$  iid de loi  $N(0, 1)$ , partant des v.a.  $U_1$  et  $U_2$  uniformes iid. Les formules sont les suivantes

$$\begin{aligned} Z_1 &= \sqrt{-2 \ln U_1} \cos(2\pi U_2) \\ Z_2 &= \sqrt{-2 \ln U_1} \sin(2\pi U_2). \end{aligned}$$

Ces formules peuvent par répétition nous fournir la réalisation des vecteurs aléatoires normaux ou gaussiens de dimension  $d > 2$ , i.e.,  $Z \sim N_d(0, I)$ , où  $I$  la matrice identité. Pour un vecteur aléatoire gaussien quelconque  $X \sim N_d(m, K)$ , où  $m$  est un vecteur de  $\mathbb{R}^d$  et  $K$  est la matrice de variances-covariances (symétrique et définie positive) de dimension  $d \times d$ , il suffit de trouver une matrice  $C$  telle que  $CC^\top = K$  et prendre  $X = CZ + m$ . Pour une matrice de variances-covariances  $K$  la méthode de Cholesky fournit une matrice  $C$  triangulaire inférieure.

## 2.3 Méthode de Monte Carlo

### 2.3.1 Mise en oeuvre

La méthode de Monte Carlo consiste à calculer des espérances mathématiques des fonctions de v.a., i.e.,

$$\mathbb{E}h(X),$$

avec  $X \in \mathbb{R}^d$ , et  $h : \mathbb{R}^d \rightarrow \mathbb{R}$ ,  $d \geq 1$ . Nous supposons ici que  $\mathbb{E}|h(X)| < \infty$ . En effet, le calcul de l'aire de l'ellipse (section 2.1), était donnée par l'espérance de la fonction indicatrice  $\mathbb{1}_E$ .

Soit  $X_i$ ,  $i \geq 1$  des copies de  $X$  indépendantes. Posons  $Y_i = h(X_i)$ ,  $i \geq 1$ . Par la loi forte des grands nombres, nous avons

$$\bar{Y}_n := \frac{1}{n}(Y_1 + \dots + Y_n) \xrightarrow{p.s.} \mathbb{E}h(X), \quad n \rightarrow \infty. \quad (2.7)$$

**Proposition 2.7.** *La méthode de Monte Carlo converge :*

- en moyenne  $O(n^{-1/2})$ ;
- au pire  $O\left(\left(\frac{\log \log n}{n}\right)^{1/2}\right)$ ,

lorsque  $n \rightarrow \infty$ .

### 2.3.2 Précision

Nous allons supposer ici que  $\mathbb{E}[h^2(X)] < \infty$ . La précision est décrite ici par deux paramètres :  $\varepsilon > 0$  et  $0 < \theta < 1$ , i.e., si

$$R_n := \bar{Y}_n - \mathbb{E}h(X), \quad (2.8)$$

elle s'exprime comme suit

$$\mathbb{P}(|R_n| \geq \varepsilon) \leq \theta. \quad (2.9)$$

La question que l'on se pose est la suivante. Pour une précision donnée  $(\varepsilon, \theta)$ , quelle est la valeur minimale de  $n$  pour attendre cette précision ?

Une première approche pour répondre à cette question consiste à utiliser l'inégalité de Tchebichev. En effet, nous pouvons écrire

$$P(|\bar{Y}_n - \mathbb{E}h(X)| \geq \varepsilon) \leq \frac{\text{Var}(h(X))}{n\varepsilon^2} \quad (2.10)$$

En conséquence, si

$$n \geq n_0 := \frac{\text{Var}(h(X))}{\theta\varepsilon^2}, \quad (2.11)$$

la précision exprimée par (2.9) est satisfaite, et  $n_0$  définie dans (2.11) est la valeur minimale recherchée.

Une approche différente pour estimer la valeur minimale de  $n$  est basée sur l'estimation par intervalles de confiance.

Posons :  $\gamma := \mathbb{E}[Y]$  et  $\sigma^2 := \text{Var}(Y)$ , et

$$S_n^2 := \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2, \quad (2.12)$$

l'estimateur corrigée de la variance  $\sigma^2$ . Par la LFGN nous obtenons

$$S_n^2 \xrightarrow{p.s.} \sigma^2. \quad (2.13)$$

Par le théorème de la limite centrale, nous avons :

$$\frac{\sqrt{n}}{\sigma}(\bar{Y}_n - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (2.14)$$

De (2.13) et du théorème de Slutsky, on obtient

$$\frac{\sqrt{n}}{S_n}(\bar{Y}_n - \gamma) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1). \quad (2.15)$$

En admettant que  $n$  est assez grand, nous pouvons considérer qu'approximativement

$$\frac{\sqrt{n}}{S_n}(\bar{Y}_n - \gamma) \sim N(0, 1),$$

et donc

$$\mathbb{P}\left(-z_{1-\alpha/2} < \frac{\sqrt{n}}{S_n}(\bar{Y}_n - \gamma) < z_{1-\alpha/2}\right) = 1 - \alpha$$

ou

$$\mathbb{P}\left(\bar{Y}_n - z_{1-\alpha/2} \frac{S_n}{\sqrt{n}} < \gamma < \bar{Y}_n + z_{1-\alpha/2} \frac{S_n}{\sqrt{n}}\right) = 1 - \alpha,$$

où  $z_\alpha$  est la  $\alpha$ -quantile de  $\mathcal{N}(0, 1)$ , et nous avons :  $z_\alpha = -z_{1-\alpha}$ .

La largeur de l'intervalle de confiance ci-dessus, i.e.,  $l := 2z_{1-\alpha/2}S_n/\sqrt{n}$  représente la précision de l'approximation.

Lorsque la précision est définie par  $\varepsilon$  et  $\theta$ , on doit avoir  $l \leq \varepsilon$ ,  $\alpha = \theta$ , et donc  $n$  doit vérifier l'inégalité suivante

$$n \geq (2z_{1-\theta/2} \frac{S_n}{\varepsilon})^2. \quad (2.16)$$

**Remarque 2.2.** Il est bien connu que lorsque les v.a.  $Y_i$  suivent la loi normale, alors la v.a.  $\frac{\sqrt{n}}{S_n}(\bar{Y}_n - \gamma)$  suit la loi de Student avec  $n - 1$  degrés de libertés. Et dans le cas où  $Y_i$  ne suivent pas la loi normale, de la limite (2.15), lorsque  $n$  est grand alors la v.a.  $\frac{\sqrt{n}}{S_n}(\bar{Y}_n - \gamma)$  suit approximativement la loi de normale. Néanmoins, la loi de Student approxime la loi normale lorsque  $n$  est assez grand.

### 2.3.3 Accélération

Plus la variance de  $X$  est petite plus la vitesse de convergence est grande. Cette assertion est intuitivement vraie. Nous allons essayer au lieu de la v.a.  $X$ , de travailler avec une autre v.a.  $Y$  telle que  $\text{Var}(Y) < \text{Var}(X)$ .

Afin de simplifier l'exposé, supposons que  $h \geq 0$ . Soit  $f$  la densité de  $X$  et  $g$  celle de  $Y$ . Nous supposons que  $g(x) > 0$  si  $f(x) > 0$ . Alors on peut écrire

$$\mathbb{E}[h(X)] = \mathbb{E}[h(Y)L(Y)] =: \gamma, \quad (2.17)$$

avec

$$L(Y) = \frac{f(Y)}{g(Y)}.$$

Se basant sur cette égalité, nous proposons l'estimateur suivant pour  $\gamma$ .

$$\gamma_n := \frac{1}{n} \sum_{i=1}^n h(Y_i)L(Y_i). \quad (2.18)$$

La variance de cet estimateur est :

$$\begin{aligned} \text{Var}(\gamma_n) &= \frac{1}{n} \text{Var}[h(Y)L(Y)] \\ &= \frac{1}{n} \left\{ \mathbb{E}[h(Y)L(Y)]^2 - [\mathbb{E}(h(Y)L(Y))]^2 \right\} \\ &= \frac{1}{n} \left\{ \mathbb{E}[h(Y)L(Y)]^2 - \gamma^2 \right\}. \end{aligned}$$

Le problème qui se pose maintenant est : quelle est la densité (ou la variable aléatoire  $Y$ ) qui minimise la variance de  $\gamma_n$  ?

Il est clair que pour minimiser la  $\text{Var}(\gamma_n)$  on doit minimiser  $\mathbb{E}[h(Y)L(Y)]^2$ . Ce qui revient à choisir pour densité

$$g(x) = \frac{1}{\gamma} h(x)f(x). \quad (2.19)$$

Dans ce cas là, nous avons :

$$\text{Var}(\gamma_n) = 0 \quad !$$

et donc une seule réalisation suffirait pour obtenir la valeur exacte de  $\gamma$ . Évidemment, on se rend vite compte que cela pose un problème...



## 2.4 Événements rares

Nous allons calculer la probabilité d'un événement à faible probabilité d'occurrence (crash d'avion, explosion d'une centrale nucléaire, panne générale du système informatique d'une banque, etc.). La probabilité d'un tel événement est en générale de l'ordre de  $10^{-10}$ .

Soit un tel événement exprimé par  $\{X \in A\}$  et posons  $h = \mathbb{1}_A$ . Nous avons

$$\gamma = \mathbb{E}h(X) = \mathbb{P}(X \in A).$$

Nous allons distinguer ici deux familles de problèmes : l'une avec erreur relative bornée et l'autre avec erreur relative non bornée.

*Erreur relative.* L'erreur relative  $RE(X)$  d'une v.a.  $X$  se définit par le rapport suivant

$$RE(X) := \frac{\text{écart-type}}{\text{espérance}} \quad (2.20)$$

Pour une v.a.  $X \sim B(\gamma)$ , on a  $RE(X) = 1/\sqrt{\gamma}$ . Pour le cas d'un estimateur  $\gamma_n$  d'une v.a. indicatrice, voir plus haut, nous avons  $\mathbb{E}\gamma_n = \gamma$  et  $Var(\gamma_n) = \gamma(1 - \gamma)/n$ . Donc

$$RE(\gamma_n) \cong \frac{1}{\sqrt{n\gamma}} \rightarrow \infty, \quad \gamma \rightarrow 0.$$

Dans ce cas là l'erreur relative n'est pas bornée. Ces problèmes sont difficiles à traiter.

En revanche, le lecteur pourra vérifier que la loi exponentielle possède une erreur relative bornée.

## 2.5 Problèmes

### 2.1. Une suite congruente.

Pour la suite congruente  $x_n = 5x_{n-1} \bmod 17$  continuer les réalisations  $x_0 = 1, x_1 = 5, x_2 = 8, x_3 = 6, x_4 = 13, x_5 = 14, x_6 = 2$  jusqu'au terme  $x_{10}$ .

### 2.2. (Méthode de rejet pour une v.a. discrète).

Supposons que nous disposons d'une méthode efficace pour générer des réalisations d'une v.a.  $Y$  de loi  $q = \{q_i, i \in \mathbb{N}\}$ . Nous allons utiliser cette v.a. comme base pour simuler une v.a. discrète  $X$  de loi  $p = \{p_i, i \in \mathbb{N}\}$ . Pour cela nous allons simuler d'abord la v.a.  $Y$ , soit  $y = Y(\omega)$ , et accepter cette valeur avec une probabilité  $(p_y/q_y) \wedge 1$ .

#### Algorithme :

1. Soit  $y = Y(\omega)$  une réalisation de  $Y$ .
  2. Soit  $U(\omega)$  une réalisation d'une v.a. uniforme sur  $[0, 1]$ .
  3. Si  $U(\omega) < p_y/q_y$ , alors  $X(\omega) = Y(\omega)$ . Fin. Sinon continuer en 1.
1. Montrer que la v.a. obtenue par cet algorithme a pour loi  $p$ .
  2. Utiliser l'algorithme dans le cas où  $X$  est à valeurs dans  $S = \{1, \dots, 10\}$  et de loi  $p = (0.11; 0.12; 0.090; 0.08; 0.10; 0.09; 0.09; 0.09; 0.10; 0.10)$  et  $Y$  est à valeurs dans  $S$  de loi uniforme.

### 2.3. V.a. géométrique.

Soit une suite de v.a.  $U_n, n \geq 1$  i.i.d. de loi  $\mathcal{U}(0, 1)$  et  $0 < p < 1$  une constante. Définissons la v.a.

$$N := \min\{n : U_n \leq p\}.$$

Montrer que  $N \sim \mathcal{G}(p)$ . Proposer un algorithme de génération d'une v.a. géométrique de paramètre  $p$ .

**2.4. V.a. binomiale.**

Étant données  $n \geq 1$  v.a.  $\mathcal{U}(0, 1)$ , proposer un algorithme pour la réalisation d'une v.a. binomiale  $b(n, p)$ .

**2.5.** Soit  $U \sim \mathcal{U}(0, 1)$ , alors la v.a.

$$X := -\frac{1}{\lambda} \log U.$$

Proposer un algorithme de génération d'une v.a. exponentielle de paramètre  $\lambda$ .

**2.6. V.a. de Poisson.**

Soit une constante  $\lambda > 0$  et une suite de v.a.  $U_n, n \geq 1$ , i.i.d. uniformes sur  $[0, 1]$ . Soit  $N + 1 = \min\{n : \prod_{i=1}^n U_i < e^{-\lambda}\}$ , alors montrer que la v.a.  $N$  suit une loi de Poisson de paramètre  $\lambda$ .

**2.7.** Soit une v.a.  $X$  de f.r.  $F$ , définie par  $F(x) \equiv \sum_{k=1}^n p_k F_k(x)$ , où  $p_k \geq 0$  et  $\sum_{k=1}^n p_k = 1$ .

Partant des réalisations de  $F_i, i = 1, \dots, n$ , donner une réalisation de  $X$ .

**2.8. Loi normale : Méthode de Box-Muller.**

Soient  $U_1$  et  $U_2$  deux v.a. i.i.d. de loi commune  $\mathcal{U}(0, 1)$ . Montrer que les v.a.

$$\begin{aligned} Z_1 &= \sqrt{-2 \ln U_1} \cos(2\pi U_2) \\ Z_2 &= \sqrt{-2 \ln U_1} \sin(2\pi U_2). \end{aligned}$$

sont i.i.d. de loi  $\mathcal{N}(0, 1)$ .

**2.9.** Utiliser la proposition 1.9 pour donner la réalisation d'un vecteur aléatoire gaussien  $N_d(m, \Lambda)$ , où  $m \in \mathbb{R}^d$  et  $\Lambda \in \mathcal{M}_{d,d}(\mathbb{R})$  est une matrice symétrique définie positive.

**2.10.** Soit un vecteur gaussien  $\mathbf{Y} \sim N_3(\mathbf{m}, \Lambda)$ , où  $\mathbf{m} = (1, 2, 3)^\top$  est le vecteur moyen et  $\Lambda$  est la matrices de variances-covariances de  $\mathbf{Y}$ , donnée par

$$\Lambda = \begin{pmatrix} 7/2 & 1/2 & -1 \\ 1/2 & 1/2 & 0 \\ -1 & 0 & 1/2 \end{pmatrix} = \mathbf{A}\mathbf{A}^\top, \quad \text{avec} \quad \mathbf{A} = \begin{pmatrix} \sqrt{7/2} & 0 & 0 \\ 1/\sqrt{14} & \sqrt{6/14} & 0 \\ -\sqrt{2/7} & 1/\sqrt{21} & 1/\sqrt{6} \end{pmatrix}$$

1. Vérifier que  $\Lambda$  est bien une matrice de variances-covariances et que la loi de  $\mathbf{Y}$  n'est pas dégénérée, i.e.,  $\det \Lambda \neq 0$ .
2. Utiliser la méthode de Box-Muller pour réaliser un vecteur gaussien  $\mathbf{Z} \sim N_3(\mathbf{0}, \mathbf{I})$ . On donne quatre réalisations des v.a. uniforme sur  $[0, 1]$  et indépendantes :  $U_1(\omega) = 0.91$ ,  $U_2(\omega) = 0.39$ ,  $U_3(\omega) = 0.13$  et  $U_4(\omega) = 0.51$ .
3. Utiliser la factorisation  $\Lambda = \mathbf{A}\mathbf{A}^\top$  pour obtenir une réalisation du vecteur  $\mathbf{Y}$ .

**2.11.** Supposons qu'il est facile d'obtenir des réalisations de  $F_i, i = 1, \dots, n$ . Générer une v.a. de f.r.  $F$  :

1.  $F(x) \equiv \prod_{k=1}^n F_k(x)$ .
2.  $F(x) \equiv 1 - \prod_{k=1}^n (1 - F_k(x))$ .

**2.12.** Soit un vecteur aléatoire  $(X, Y)$  à valeurs dans  $\mathbb{R}^2$  de densité  $f(x, y)$  et  $U_1, U_2$  deux v.a. i.i.d. de loi uniforme sur  $(0, 1)$ . Notons également  $F(x, y)$  la f.r. de  $(X, Y)$ ,  $F_X(x)$ ,  $F_Y(y)$  les f.r. marginales des  $X$  et  $Y$  respectivement et  $F_{Y|X}(y | x)$  la f.r. conditionnelle de  $Y$  en  $\{X = x\}$ . Nous supposons que les fonctions inverses existent.

1. Montrer que le vecteur aléatoire  $(F_X^{-1}(U_1), F_{Y|X}^{-1}(U_2 | F_X^{-1}(U_1)))$  suit la même loi que  $(X, Y)$ .
2. Soit la densité  $f(x, y) = \frac{2}{3}(x + 2y)\mathbf{1}_{[0,1]^2}(x, y)$ , alors donnez une réalisation du vecteur  $(X, Y)$ .
3. Généraliser cette méthode pour un vecteur aléatoire dans  $\mathbb{R}^d$ ,  $d \geq 3$ .

**2.13.** Soient une v.a.  $X : \Omega \rightarrow E = \{-2, -1, 0, 2.5\}$ , de loi de probabilité  $p = (0.2; 0.2; 0.3; 0.3)$  et une fonction  $h : \mathbb{R} \rightarrow \mathbb{R}$  telle que  $h(x) = \sqrt{|\pi \cos(\pi x)|}$ .

1. Etant donné une réalisation d'une v.a.  $U$  uniform sur  $[0, 1]$ , telle que  $U(\omega) = 0.32$ , donner une réalisation pour chacune de v.a.  $X$  et  $h(X)$ .
2. Monte Carlo. Proposer un schéma de calcul d'une approximation de  $\mathbb{E}[h(X)]$  avec précision  $\varepsilon = \theta = 10^{-2}$ .
3. Définir le nombre de réalisations minimum  $n$  des v.a.  $X$  pour obtenir cette précision.

[Indication : une v.a.  $Z$  avec valeurs dans l'intervalle réel  $[a, b]$  satisfait les inégalités  $a \leq \mathbb{E}(Z) \leq b$  et  $\text{Var}(Z) \leq (b - a)^2/4$ ].

**2.14.** *Réduction de la variance - variable de contrôle.*

Supposons qu'on souhaite estimer  $\gamma = \mathbb{E}[X]$  et soit une autre v.a.  $Y$  de moyenne  $\mu_Y$ . Alors

1. Montrer que la v.a.  $Z = X + C(Y - \mu_Y)$  est un estimateur sans biais de  $\gamma$  pour toute constante  $C \in \mathbb{R}$ .
2. Trouver la valeur de  $C$  qui minimise la variance de l'estimateur  $Z$ . Donner la variance de  $Z$  pour cette valeur.
3. On souhaite estimer  $\gamma := \mathbb{E}(e^U)$ . Trouver une v.a. de contrôle  $Y$  et calculer la variance de  $Z_{c^*}$ . Comparer avec la variance de la v.a.  $X$ .

**2.15.** *Réduction de la variance en fiabilité.*

On souhaite estimer  $r := \mathbb{E}[\phi(X_1, \dots, X_n)]$ , où  $\phi : \{0, 1\}^n \rightarrow \{0, 1\}$  est la fonction de structure d'un système de fiabilité. Soit  $(p_1, \dots, p_n)$  les fiabilités de composants. Les v.a.  $X_i$  sont indépendantes.

Montrer que l'estimation est meilleure en se basant sur la fonction  $\phi(1_i, \mathbf{x})p_i + \phi(0_i, \mathbf{x})(1 - p_i)$  que sur la fonction initiale  $\phi$ .

**2.16.** *Approximation Bootstrap de l'EQM*

Soit  $n$  v.a.r.,  $X_1, \dots, X_n$ , iid de f.r.  $F$ . On cherche à estimer un paramètre  $\gamma = \gamma(F)$  par un estimateur  $T_n = T_n(X_1, \dots, X_n)$ . L'erreur quadratique moyenne (EQM) de cet estimateur est définie par  $EQM(F) := \mathbb{E}_F[(T_n - \gamma)^2]$ . On définit l'approximation bootstrap de  $EQM(F)$  par  $EQM(F_n)$ , i.e.,

$$EQM(F_n) := \mathbb{E}_{F_n}[(T_n(V_1, \dots, V_n) - \gamma(F_n))^2]$$

où  $F_n(x)$  est la f.r. empirique de l'échantillon, i.e.,  $F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}}$  et  $V_i \sim F_n$ .

Si  $T_n(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n V_i^3$  et  $X_1(\omega) = x_1, \dots, X_n(\omega) = x_n$  évaluer par Monte Carlo l'EQM( $F_n$ ).

## Chapitre 3

# Chaînes de Markov

Nous présentons<sup>1</sup> ici les chaînes de Markov homogènes à temps discret et espace d'état discret. Elles constituent une généralisation des suites de variables indépendantes à des suites de variables liées par des relations de dépendance simples : une suite aléatoire est une chaîne de Markov si sa valeur future connaissant le passé et le présent ne dépend que du présent. Le terme temps est relatif car souvent il s'agit plutôt du numéro d'une épreuve que du temps calendaire, par exemple, transitions d'une substance entre les états solide, gazeux et liquide, passages d'un système de l'état de marche à l'état de panne...

### 3.1 Fonction de transition et exemples

Une chaîne de Markov est caractérisée par sa fonction de transition. Nous en présentons les propriétés de base, avec de nombreux exemples de chaînes de Markov usuelles. L'étude de la classification des états de la chaîne et des lois stationnaires conduit ensuite à l'étude asymptotique des chaînes.

Dans la suite,  $(X_n)$  sera une suite de variables aléatoires définies sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$  et à valeurs dans un ensemble fini ou dénombrable  $E$ , dit espace d'état ou de phases.

**Définition 3.1.** [Propriété de Markov] Une suite aléatoire  $(X_n)$  est une chaîne de Markov si pour tous états  $i, j, i_0, i_1, \dots, i_{n-1}$  de  $E$ , et tout entier positif  $n$ , elle vérifie

$$\begin{aligned}\mathbb{P}(X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i) = \\ = \mathbb{P}(X_{n+1} = j \mid X_n = i),\end{aligned}$$

sous réserve que  $\mathbb{P}(X_0 = i_0, X_1 = i_1, \dots, X_{n-1} = i_{n-1}, X_n = i)$  soit non nulle.

Si  $\mathbb{P}(X_{n+1} = j \mid X_n = i) = P(i, j)$  ne dépend pas du temps  $n$ , la chaîne de Markov est dite homogène (par rapport au temps).

Nous étudierons ici seulement les chaînes homogènes.

La fonction  $(i, j) \rightarrow P(i, j)$ , définie sur  $E \times E$ , est appelée fonction de transition de la chaîne (ou fonction de passage en une étape). En effet,

$$\mathbb{P}(X_{n+1} \in A \mid X_0 = i_0, \dots, X_n = i) = \sum_{j \in A} P(i, j)$$

et la fonction  $P: (E, \mathcal{E}) \rightarrow [0, 1]$  définie par  $P(i, A) = \sum_{j \in A} P(i, j)$  est une probabilité de transition de  $E$  dans lui-même. La quantité  $P(i, j)$  est dite probabilité de transition de  $i$  à  $j$  et  $P = (P(i, j))_{(i, j) \in E^2}$  est la matrice de transition (éventuellement infinie). La loi  $\mu$  de  $X_0$  est dite loi initiale de la chaîne. Nous noterons  $\mu(i) = \mu(\{i\})$ ; nous définissons  $\mathbb{P}_\mu(A) = \sum_{i \in E} \mu(i) \mathbb{P}(A \mid$

---

1. Ce chapitre est un extrait du livre : V. Girardin, N. Limnios, *Probabilités*, Vuibert, Paris 2014.

$X_0 = i$ ) et  $\mathbb{E}_\mu$  sera l'espérance correspondant à la probabilité  $\mathbb{P}_\mu$ . Si  $\mu = \delta_i$ , avec  $i \in E$ , nous noterons  $\mathbb{P}_i$  au lieu de  $\mathbb{P}_{\delta_i}$ , autrement dit  $\mathbb{P}_i(A) = \mathbb{P}(A \mid X_0 = i)$ .

**Définition 3.2.** La fonction de transition en  $n$  étapes est définie par la probabilité de passage d'un état à un autre en  $n$  étapes, soit

$$P^n(i, j) = \mathbb{P}(X_{m+n} = j \mid X_m = i) = \mathbb{P}(X_n = j \mid X_0 = i), \quad m \geq 0, n \geq 0,$$

avec  $P^1(i, j) = P(i, j)$  et, par convention,  $\mathbb{P}^0(i, i) = 1$  et  $\mathbb{P}^0(i, j) = 0$  si  $i \neq j$ .

Pour toute fonction  $h$  sur  $E$ , nous noterons  $Ph$  la fonction définie sur  $E$  par  $Ph(i) = \sum_{j \in E} P(i, j)h(j)$ . Si  $Ph = h$ , la fonction  $h$  est dite harmonique ; dans ce cas,  $P^n h = h$  pour tout entier  $n$ . La propriété de Markov peut aussi s'exprimer de manière équivalente par

$$\mathbb{E}[h(X_{n+1}) \mid \mathcal{F}_n] = \mathbb{E}[h(X_{n+1}) \mid X_n] = Ph(X_n), \quad \text{p.s.}$$

pour toute fonction  $h : E \rightarrow \mathbb{R}$  bornée.

### Propriétés d'une fonction de transition

1. La fonction  $j \rightarrow P(i, j)$  définit une probabilité sur  $(E, \mathcal{P}(E))$  pour tout  $i \in E$  ; en effet,  $0 \leq P(i, j) \leq 1$  et  $\sum_{j \in E} P(i, j) = 1$ , puisque la chaîne reste dans  $E$ . Rester en  $i$  est considéré comme une transition virtuelle. Précisément, un temps de saut est un instant où un changement d'état a lieu.
2. (**Identité de Chapman-Kolmogorov**) On a pour tous états  $i$  et  $j$  de  $E$  et tous entiers  $m$  et  $n$ ,

$$\sum_{k \in E} P^n(i, k)P^m(k, j) = P^{n+m}(i, j).$$

En effet,

$$\begin{aligned} P^{n+m}(i, j) &= \mathbb{P}(X_{n+m} = j \mid X_0 = i) \\ &\stackrel{(1)}{=} \sum_{k \in E} \mathbb{P}(X_{n+m} = j, X_n = k \mid X_0 = i) \\ &\stackrel{(2)}{=} \sum_{k \in E} \mathbb{P}(X_{n+m} = j \mid X_n = k, X_0 = i) \mathbb{P}(X_n = k \mid X_0 = i) \\ &\stackrel{(3)}{=} \sum_{k \in E} P^m(k, j)P^n(i, k). \end{aligned}$$

(1) par la formule des probabilités totales, (2) par la formule des probabilités composées et (3) par la propriété de Markov.

3. Si  $E$  est de cardinal fini, il découle du point 2. que la fonction de transition en  $n$  étapes  $P^n$  est la puissance  $n$ -ième de  $P$  pour le produit matriciel ordinaire, et plus généralement,  $P^n P^m = P^{n+m} = P^m P^n$ .

Notons qu'une suite aléatoire qui vérifie l'identité de Chapman-Kolmogorov n'est pas nécessairement une chaîne de Markov.

▷ **Exemple 3.1** *Transmission d'un message binaire dans une population.*— Le message reçu par un individu est bien retransmis avec probabilité  $1 - p$ , où  $p \in ]0, 1[$ . L'espace d'état est  $E = \{0, 1\}$  (soit {non, oui}). On suppose qu'à l'origine le message est "non", soit  $\mathbb{P}(X_0 = 0) = 1$ . On a

$$\begin{aligned} \mathbb{P}(X_n = 1 \mid X_{n-1} = 1) &= \mathbb{P}(X_n = 0 \mid X_{n-1} = 0) = 1 - p, \\ \mathbb{P}(X_n = 0 \mid X_{n-1} = 1) &= \mathbb{P}(X_n = 1 \mid X_{n-1} = 0) = p, \end{aligned}$$

soit la matrice de transition

$$P = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}.$$

La probabilité que le message se transmette sans erreur sur 2 étapes s'écrit

$$\mathbb{P}(X_0 = 0, X_1 = 0, X_2 = 0) = \mu(0)P(0,0)P(0,0) = (1-p)^2.$$

De même, sur 5 étapes,  $\mathbb{P}(X_i = 0, i = 0, \dots, 5) = (1-p)^5$ .

Quelle est la probabilité que le message arrive sans erreur au 2-ième individu, au 5-ième ?

$$\begin{aligned} \mathbb{P}(X_2 = 0) &= \mathbb{P}(X_0 = 0, X_1 = 0, X_2 = 0) + \mathbb{P}(X_0 = 0, X_1 = 1, X_2 = 0) \\ &= (1-p)^2 + p^2. \end{aligned}$$

Cette décomposition est possible pour 2 étapes, plus dure pour 5. Par contre  $\mathbb{P}(X_n = 0) = P^n(0,0)$ . Or on peut montrer par récurrence que

$$P^n = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{(1-2p)^n}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix},$$

d'où

$$\mathbb{P}(X_n = 0) = \frac{1}{2} + \frac{(1-2p)^n}{2}.$$

Remarquons que

$$P^n \rightarrow \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad n \rightarrow +\infty,$$

donc, quels que soient  $p$  et le message de départ, on a autant de chances de recevoir celui-ci que son contraire, voir la figure ???. Ceci illustre un problème important pour les chaînes de Markov : convergence, loi limite (i.e., loi de  $X_n$  lorsque  $n$  tend vers l'infini), oubli du passé, et notamment de la loi initiale.

◁

▷ **Exemple 3.2** *Réservoir de Moran.*- Un réservoir de capacité  $c$  unités de volume ( $c \in \mathbb{N}^*$ ) est observé à des instants  $n$  entiers. Durant l'intervalle de temps  $[n, n+1[$ , une quantité d'eau aléatoire égale à  $Z_n$  unités de volume entre dans le réservoir. Lorsque la capacité  $c$  du réservoir est dépassée, la quantité en excédent est perdue (par débordement). A la fin de l'intervalle  $[n, n+1[$  une unité d'eau, si elle est disponible, quitte le réservoir. Si  $X_n$  est le niveau d'eau du réservoir au temps  $n$ , il est clair que le niveau au temps  $n+1$  est

$$X_{n+1} = (X_n + Z_n - 1)^+ \wedge (c-1).$$

Si les variables  $Z_n$  sont supposées positives et i.i.d. de loi  $\mathbb{P}(Z_n = k) = p_k$ , alors  $(X_n)$  est une chaîne de Markov d'espace d'état  $E = \{0, 1, \dots, c-1\}$  et de matrice de transition

$$\begin{matrix} & \begin{matrix} 0 & 1 & 2 & \cdots & c-2 & c-1 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ c-2 \\ c-1 \end{matrix} & \begin{pmatrix} p_0 + p_1 & p_2 & p_3 & \cdots & p_{c-1} & h_c \\ p_0 & p_1 & p_2 & \cdots & p_{c-2} & h_{c-1} \\ 0 & p_0 & p_1 & \cdots & p_{c-3} & h_{c-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & p_1 & h_2 \\ 0 & 0 & 0 & \cdots & p_0 & h_1 \end{pmatrix} \end{matrix},$$

où  $h_k = \sum_{i \geq k} p_i$  pour  $1 \leq k \leq c$ . Remarquons que  $c$  pourrait également être considéré comme un état de la chaîne, mais seule  $X_0$  pourrait prendre cette valeur.

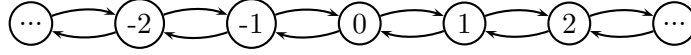


FIGURE 3.1 – Un graphe de marche aléatoire simple (exemple 3.3).

&lt;

▷ **Exemple 3.3** *Marche aléatoire.*– Une marche aléatoire (voir la définition ??) est une chaîne de Markov d'espace d'état  $E = \mathbb{Z}$  et de fonction de transition  $P$  donnée par

$$P(i, j) = p(j - i), \quad (i, j) \in \mathbb{Z}^2,$$

où  $p$  est une fonction de  $\mathbb{Z}$  dans  $[0, 1]$ .

Lorsque  $p(1) = p = 1 - p(-1)$ , avec  $p \in [0, 1]$ , la marche aléatoire est dite simple et l'on a

$$P(i, j) = \begin{cases} p & \text{si } j = i + 1, \\ 1 - p & \text{si } j = i - 1, \\ 0 & \text{sinon.} \end{cases}$$

Son graphe est donné dans la figure 3.1. Une trajectoire de marche aléatoire simple est représentée dans la figure ??.

La probabilité de retour en  $i$  en  $n$  étapes s'écrit

$$P^{2n+1}(i, i) = 0 \quad \text{et} \quad P^{2n}(i, i) = \binom{2n}{n} p^n (1 - p)^n, \quad n \geq 1. \quad (3.1)$$

On montre par récurrence et en utilisant la formule du triangle de Pascal que la probabilité de passer de  $i$  à  $j$  en  $n$  étapes est

$$P^n(i, j) = \binom{n}{(n+j-i)/2} p^{(n+j-i)/2} (1-p)^{(n-j+i)/2} \quad \text{si } n+j-i = 2m > 0,$$

et est nulle sinon.

&lt;

▷ **Exemple 3.4** *Chaîne de naissance et de mort.*– Toute chaîne de Markov d'espace d'état  $E = \mathbb{N}$  et de fonction de transition  $P$  tridiagonale, soit

$$P(i, j) = \begin{cases} p_i & \text{si } j = i + 1, \\ q_i & \text{si } j = i - 1, \\ r_i & \text{si } j = i, \\ 0 & \text{sinon,} \end{cases}$$

pour  $i \geq 0$  et  $j \geq 1$ , est appelée chaîne de naissance et de mort. Si  $q_i = 0$  (respectivement  $p_i = 0$ ) pour tout  $i$ , la chaîne est dite de naissance (respectivement de mort).

Cette chaîne sert de modèle pour beaucoup de phénomènes physiques et biologiques : croissance d'une population, évolution d'un parc de matériels identiques subissant des défaillances et des réparations, évolution d'une substance radioactive, etc. Souvent, elle est considérée avec espace d'état fini : lorsque  $E = \{1, \dots, N\}$ , avec  $p_i = 1 - i/N$ ,  $q_i = i/N$  et  $r_i = 0$ , le modèle est dit d'Ehrenfest et est utilisé dans la théorie des gaz en thermodynamique ; lorsque  $E = \{0, \dots, N\}$ , avec  $r_i = 0$ ,  $p_i = p$ ,  $q_i = q$ , pour tout  $i$  sauf  $p_0 = 0$  et  $q_N = 0$ , c'est la chaîne de la ruine du joueur. Lorsque  $E = \mathbb{Z}$ , avec  $r_i = 0$ ,  $p_i = p$  et  $q_i = 1 - p$  pour tout  $i$ , c'est une marche aléatoire simple.

◁

Les lois finidimensionnelles d'une chaîne de Markov sont déterminées par sa loi initiale et sa matrice de transition, d'après la formule suivante.

**Proposition 3.1.** *Pour tout  $n \geq 0$  et pour toutes parties  $A_0, A_1, \dots, A_n$  de  $E$ , on a*

$$\begin{aligned} \mathbb{P}(X_0 \in A_0, X_1 \in A_1, \dots, X_n \in A_n) &= \\ &= \sum_{i_0 \in A_0} \sum_{i_1 \in A_1} \cdots \sum_{i_n \in A_n} \mu(i_0) P(i_0, i_1) \cdots P(i_{n-1}, i_n). \end{aligned} \quad (3.2)$$

Si dans la relation (3.2), on pose  $A_0 = A_1 = \cdots = A_{n-1} = E$  et  $A_n = A$ , on obtient la loi de  $X_n$ , soit

$$\mathbb{P}(X_n \in A) = \mu P^n(A) = \sum_{i \in E} \sum_{j \in A} \mu(i) P^n(i, j). \quad (3.3)$$

*Démonstration.* Il suffit d'appliquer la formule des probabilités composées et la propriété de Markov. Par exemple,

$$\begin{aligned} \mathbb{P}(X_0 = i_0, X_1 = i_1, X_2 = i_2) &= \\ &= \mathbb{P}(X_2 = i_2 \mid X_0 = i_0, X_1 = i_1) \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) P(X_0 = i_0) \\ &= \mathbb{P}(X_2 = i_2 \mid X_1 = i_1) \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) P(X_0 = i_0) \\ &= \mu(i_0) P(i_0, i_1) P(i_1, i_2), \end{aligned}$$

d'où la loi jointe de  $(X_0, X_1, X_2)$ . □

Réciproquement, si une suite de variables aléatoires  $(X_n)$  à valeurs dans un ensemble fini ou dénombrable  $E$  vérifie la relation (3.2) pour une loi  $\mu$  et une fonction  $P$ , alors c'est une chaîne de Markov de loi initiale  $\mu$  et de fonction de transition  $P$ . Le résultat suivant est même plus général.

**Théorème 3.2** (Construction canonique). *Si  $E$  est un ensemble au plus dénombrable,  $\mu$  une probabilité et  $P$  une fonction de transition sur  $E$ , alors il existe un espace de probabilité sur lequel est définie une chaîne de Markov à valeurs dans  $E$  de loi initiale  $\mu$  et de fonction de transition  $P$ .*

La chaîne ainsi construite est dite canonique et l'espace  $(E^{\mathbb{N}}, \mathcal{P}(E)^{\otimes \mathbb{N}}, \mathbb{P})$  est dit espace canonique.

*Démonstration.* Soit  $(\Omega, \mathcal{F}) = (E^{\mathbb{N}}, \mathcal{P}(E)^{\otimes \mathbb{N}})$ . La formule (3.2) permet de définir une probabilité  $\mathbb{P}$  sur les cylindres infinis  $A_0 \times A_1 \times \cdots \times A_n \times E \times \cdots$  de  $(\Omega, \mathcal{F})$ . Le projecteur sur la  $n$ -ième coordonnée de  $\Omega$  est défini par

$$X_n(\omega) = X_n(\omega_0, \omega_1, \dots, \omega_n, \dots) = \omega_n,$$

et  $(X_n)$  est une chaîne de Markov sur  $(\Omega, \mathcal{F}, \mathbb{P})$  avec les propriétés requises. □

Une caractéristique importante des états d'une chaîne de Markov est leur périodicité.

**Définition 3.3.** Soit  $d_i = \text{p.g.c.d.}\{n \in \mathbb{N}^* : P^n(i, i) > 0\}$  pour  $i \in E$ . Si  $P^n(i, i) = 0$  pour tout  $n \geq 1$ , on pose  $d_i = +\infty$ . Si  $1 < d_i < +\infty$ , l'état  $i$  est dit périodique de période  $d_i$ . Si  $d_i = 1$ , l'état  $i$  est dit apériodique.

Une chaîne de Markov dont tous les états sont périodiques est dite périodique. Un état  $i \in E$  tel que  $P(i, i) = 1$  est dit absorbant. Lorsque le processus entre dans un état absorbant, il y reste indéfiniment.



**Proposition 3.3.** *Si la loi initiale de la chaîne est  $\delta_i$ , avec  $i \in E$  non absorbant, alors :*

1. *le temps de séjour du système dans l'état  $i$  (c'est-à-dire le temps passé par la chaîne en  $i$  avant de le quitter) suit une loi géométrique de paramètre  $1 - P(i, i)$  ;*
2. *la probabilité pour que le système entre dans l'état  $j$  lorsqu'il quitte l'état  $i$  est  $P(i, j)/[1 - P(i, i)]$ .*

*Démonstration.* 1. Soit  $\eta_i$  le temps de séjour dans l'état  $i$ . Pour  $k \geq 1$ , nous avons

$$\begin{aligned} \mathbb{P}_i(\eta_i = k) &= \\ &= P_i(X_{m+1} = i, \dots, X_{m+k-1} = i, X_{m+k} \neq i \mid X_{m-1} \neq i, X_m = i) \\ &= \sum_{j \in E \setminus \{i\}} P(i, i) \dots P(i, i) P(i, j) = P(i, i)^{k-1} [1 - P(i, i)]. \end{aligned}$$

2. Pour  $i \neq j$ , nous avons

$$\begin{aligned} \mathbb{P}_i(X_{n+1} = j \mid X_n = i, X_{n+1} \neq i) &= \\ &= \frac{\mathbb{P}_i(X_{n+1} = j, X_n = i)}{\mathbb{P}_i(X_n = i, X_{n+1} \neq i)} = \frac{\mathbb{P}_i(X_{n+1} = j \mid X_n = i)}{\mathbb{P}_i(X_{n+1} \neq i \mid X_n = i)} = \frac{P(i, j)}{1 - P(i, i)}, \end{aligned}$$

par définition des probabilités conditionnelles.  $\square$

La matrice de coefficients  $Q(i, j) = P(i, j)/[1 - P(i, i)]$ , pour  $i \neq j$  est la matrice de transition de la chaîne de Markov immergée définie aux instants de changement d'état de  $(X_n)$  ; elle sera étudiée dans l'exercice ??.

▷ **Exemple 3.5** Le temps de séjour dans l'état  $i$  d'une chaîne de naissance et de mort suit une loi géométrique  $\mathcal{G}(1 - r_i)$  si  $r_i > 0$ . Il est constant et égal à une unité de temps si  $r_i = 0$ . Si  $r_i = 1$ , le temps de séjour dans  $i$  est infini et  $i$  est absorbant.

◁

On peut ainsi générer une trajectoire d'une chaîne de Markov dans un intervalle de temps donné  $[0, t]$ , par ses temps de saut et les états visités, grâce à l'algorithme suivant :

1. Soit  $i$  l'état initial de la chaîne. Poser  $n := 0$ .  $j := i$ .  $S_0(\omega) := W_0(\omega) = 0$ .
2. Poser  $n := n + 1$ .
  - (a) Soit  $W_n(\omega)$  la réalisation d'une variable de loi géométrique  $\mathcal{G}(1 - P(j, j))$ .
  - (b) Poser  $S_n(\omega) := S_{n-1}(\omega) + W_n(\omega)$ .
  - (c) Soit  $K(\omega)$  la réalisation d'une variable  $K$  à valeurs dans  $E \setminus \{j\}$  de loi donnée par  $\mathbb{P}(K = k) = P(j, k)/[1 - P(j, j)]$  ; poser  $j := K(\omega)$ .
3. Si  $S_n \geq t$ , alors fin.
4. Continuer à l'étape 2.

Notons que si la loi initiale de la chaîne n'est pas  $\delta_i$  mais une loi quelconque sur  $E$ , l'étape 1. de l'algorithme consiste en sa réalisation, suivant l'une des méthodes décrites dans le paragraphe ??.

Les chaînes de Markov multidimensionnelles se définissent et s'étudient comme les chaînes unidimensionnelles. Voyons deux exemples typiques de chaînes bidimensionnelles.

▷ **Exemple 3.6** *Chaîne de Markov produit.* - Considérons deux chaînes de Markov  $(X_n)$  et  $(Y_n)$  indépendantes entre elles, d'espaces d'état  $E_X$  et  $E_Y$ , de lois initiales  $\mu^X$  et  $\mu^Y$ , et de fonctions de transitions  $P^X$  et  $P^Y$  respectivement.

La suite  $(Z_n) = ((X_n, Y_n))$  est une chaîne de Markov d'espace d'état  $E = E_X \times E_Y$ , de loi initiale  $\mu = \mu^X \otimes \mu^Y$  et de fonction de transition définie par

$$P((i, j), (k, l)) = P^X(i, k)P^Y(j, l), \quad (i, j) \in E_X \times E_Y, (k, l) \in E_X \times E_Y.$$

En effet,  $\mathbb{P}[Z_0 = (i, j)] = \mathbb{P}(X_0 = i)\mathbb{P}(Y_0 = j)$ , la propriété de Markov est vérifiée, et nous avons

$$\begin{aligned} P((i, j), (k, l)) &= \mathbb{P}[Z_1 = (k, l) \mid Z_0 = (i, j)] \\ &= \mathbb{P}(X_1 = k, Y_1 = l \mid X_0 = i, Y_0 = j) \\ &= \mathbb{P}(X_1 = k \mid X_0 = i)\mathbb{P}(Y_1 = l \mid Y_0 = j). \end{aligned}$$

Le produit est évidemment généralisable à  $n$  chaînes de Markov indépendantes.

◁

▷ **Exemple 3.7** Soit  $(X_n)$  une chaîne de Markov. Soit  $Z_n = (X_n, X_{n-1})$ . Montrons que  $(Z_n)$  vérifie la propriété de Markov. Posons  $j = (j^2, j^1)$  et  $i_k = (i_k^2, i_k^1)$  pour  $k = 0, \dots, n$ , avec évidemment  $i_k^1 = i_{k-1}^2$  pour  $k = 1, \dots, n$ . Nous avons

$$\begin{aligned} \mathbb{P}(Z_{n+1} = j \mid Z_0 = i_0, \dots, Z_{n-1} = i_{n-1}, Z_n = i_n) &= \\ &= \mathbb{P}(X_{n+1} = j^2 \mid X_0 = i_0^1, X_1 = i_0^2, \dots, X_{n-1} = i_{n-1}^1, X_n = i_n^2) \delta_{i_n^2 j^1} \\ &\stackrel{(1)}{=} \mathbb{P}(X_{n+1} = j^2 \mid X_{n-1} = i_{n-1}^1, X_n = i_n^2) \delta_{i_n^2 j^1} \\ &= \mathbb{P}(X_{n+1} = j^2, X_n = j^1 \mid X_{n-1} = i_{n-1}^1, X_n = i_n^2) = \mathbb{P}(Z_{n+1} = j \mid Z_n = i_n). \end{aligned}$$

(1) par la propriété de Markov appliquée à  $(X_n)$ , où nous conservons  $X_{n-1} = i_{n-1}^1$  pour retrouver  $Z_n$ .

La fonction de transition  $\tilde{P}$  de cette chaîne bidimensionnelle  $(X_n, X_{n-1})$  est donnée par  $\tilde{P}((k, l), (i, j)) = P(j, i) \delta_{kj}$  pour  $i, j, k, l \in E$ . Notons que la fonction de transition  $\tilde{P}$  de la chaîne bidimensionnelle  $(X_{n-1}, X_n)$  est donnée par  $\tilde{P}((k, l), (i, j)) = P(i, j) \delta_{li}$  pour tous  $i, j, k$  et  $l$  de  $E$ .

Ceci se généralise facilement à la chaîne  $k$ -dimensionnelle (appelée chaîne serpent)  $Z_n = (X_n, \dots, X_{n-k})$  pour tout  $k \geq 1$  fixé.

◁

## 3.2 Décomposition de l'espace d'état

Avant d'étudier spécifiquement le comportement asymptotique des chaînes de Markov, il est nécessaire de donner un certain nombre de notions sur la classification des états.

Un état d'une chaîne de Markov peut être caractérisé soit comme récurrent soit comme transient (ou transitoire). Cette distinction est fondamentale pour l'étude des chaînes de Markov.

Soient  $i \in E$  et  $j \neq i$  dans  $E$ . Le temps passé par la chaîne dans l'état  $i$  dans l'intervalle de temps  $[1, n]$  est  $N_i^n = \sum_{k=1}^n \mathbf{1}_{(X_k=i)}$ , et  $N_i = \sum_{k \geq 1} \mathbf{1}_{(X_k=i)}$  est le temps total passé par la chaîne en  $i$ . Le nombre de passages en une étape de  $i$  à  $j$  dans l'intervalle de temps  $[1, n]$  est  $N_{ij}^n = \sum_{k=1}^n \mathbf{1}_{(X_{k-1}=i, X_k=j)}$  et  $N_{ij} = \sum_{k \geq 1} \mathbf{1}_{(X_{k-1}=i, X_k=j)}$  est le nombre de transitions directes de  $i$  à  $j$ . Les variables aléatoires  $N_i$  et  $N_{ij}$  sont à valeurs entières, éventuellement infinies.

Soit  $\rho_{ij} = \mathbb{P}_i(T_j < +\infty) = \mathbb{P}_i(N_j \geq 1)$  la probabilité de passer dans l'état  $j$  en partant de  $i$ , c'est-à-dire que le temps de retour en  $j$  soit fini.

**Définition 3.4.** Si  $\rho_{ii} = 1$ , l'état  $i$  est dit récurrent. Sinon, c'est à dire si  $\rho_{ii} < 1$ , il est dit transient.

Si  $i$  est récurrent, alors, soit  $m_i = \mathbb{E}_i T_i < +\infty$  et  $i$  est dit récurrent positif, soit  $m_i = +\infty$  et  $i$  est dit récurrent nul.

Un état  $i$  est récurrent si, lorsque la chaîne part de  $i$ , elle revient en  $i$  avec probabilité un. En particulier, un état absorbant est évidemment récurrent positif.

▷ **Exemple 3.8** *Ruine du joueur.* - Supposons que la fortune initiale d'un joueur est égale à  $k$  et celle du casino à  $N - k$ . Si le joueur joue un euro à chaque partie avec une probabilité  $p \in ]0, 1[$  de gagner un euro, le jeu constitue une chaîne de naissance et de mort, à valeurs dans  $\{0, \dots, N\}$ , avec  $P(i, i + 1) = p = 1 - P(i, i - 1) = 1 - q$  et  $P(0, 0) = P(N, N) = 1$ .

Les états 0 et  $N$  sont absorbants. Les autres états sont transients puisque la probabilité de passer de  $i$  à 0 ou  $N$  n'est pas nulle. Le temps d'absorption de la chaîne est  $\tau = \inf\{n \in \mathbb{N}^* : X_n = 0 \text{ ou } N\}$ . L'événement "Le joueur est ruiné" correspond à l'absorption de la chaîne en 0. La quantité  $u_k = \mathbb{P}(X_\tau = 0 \mid X_0 = k)$  est la probabilité de ruine du joueur, et celle de ruine du casino est  $1 - u_k$ . On a  $u_0 = 1$ ,  $u_N = 0$  et  $u_k = pu_{k+1} + qu_{k-1}$  pour tout  $1 \leq k \leq N - 1$ . La suite  $(u_k)$  est donc une suite de Fibonacci. En posant  $r = q/p$ , on en déduit que

$$u_k = \begin{cases} \frac{r^k - r^N}{1 - r^N} & \text{si } p \neq q, \\ \frac{N - k}{N} & \text{si } p = 1/2. \end{cases}$$

On pourrait calculer de même la durée moyenne du jeu en déterminant  $\mathbb{E}(\tau \mid X_0 = k) = v_k$ , qui vérifie la relation  $1 + pv_{k+1} + qv_{k-1} = v_k$ .

◁

**Lemme 3.4.** *Pour tous  $(i, j) \in E \times E$  et  $m \in \mathbb{N}^*$ , on a*

$$\mathbb{P}_i(N_j \geq m) = \rho_{ij} \rho_{jj}^{m-1}.$$

*Démonstration.* Par récurrence sur  $m$ .

Pour  $m = 1$ , par définition,  $\mathbb{P}_i(N_j \geq 1) = \rho_{ij}$ . Supposons que le résultat soit vrai pour  $m > 1$ . Soient  $n_1 < n_2 < \dots < n_m < n_{m+1}$  les  $m + 1$  premiers instants de retour de la chaîne dans l'état  $j$  et  $F_m = \{(n_1, \dots, n_m) : 1 \leq n_1 < \dots < n_m < +\infty\}$ . Nous avons

$$\mathbb{P}_i(N_j \geq m + 1) = \sum_{(n_1, \dots, n_{m+1}) \in F_{m+1}} \mathbb{P}_i(X_{n_1} = j, \dots, X_{n_{m+1}} = j).$$

Or, par la formule des probabilités composées et la propriété de Markov,

$$\mathbb{P}_i(X_{n_1} = j, \dots, X_{n_{m+1}} = j) = \mathbb{P}_i(X_{n_1} = j, \dots, X_{n_m} = j) \mathbb{P}_j(X_{n_{m+1} - n_m} = j).$$

Par conséquent,

$$\begin{aligned} \mathbb{P}_i(N_j \geq m + 1) &= \sum_{(n_1, \dots, n_m) \in F_m} \mathbb{P}_i(X_{n_1} = j, \dots, X_{n_m} = j) \sum_{k \geq 1} \mathbb{P}_j(X_k = j) \\ &= \mathbb{P}_i(N_j \geq m) \mathbb{P}_j(T_j < +\infty) \\ &\stackrel{(1)}{=} \rho_{ij} \rho_{jj}^{m-1} \rho_{jj}. \end{aligned}$$

(1) par hypothèse de récurrence. □

**Définition 3.5.** Soient  $(X_n)$  une chaîne de Markov d'espace d'état  $E$  et de matrice de transition  $P$  et  $\alpha \in ]0, 1]$ . La fonction

$$U^\alpha = I + \alpha P + \alpha^2 P^2 + \dots + \alpha^n P^n + \dots$$

définie sur  $E \times E$  et à valeurs dans  $\overline{\mathbb{R}}$  est appelée  $\alpha$ -potentiel de  $(X_n)$  ou de  $P$ .

La fonction  $U^1$  est notée  $U$  et appelée potentiel de la chaîne.

**Proposition 3.5.** *On a*

$$U(i, j) = \begin{cases} \mathbb{E}_i N_j & \text{si } i \neq j, \\ 1 + \mathbb{E}_i N_i & \text{si } i = j, \end{cases} \quad (3.4)$$

et pour  $i \neq j$ , on a aussi  $U(i, j) = \rho_{ij} U(j, j)$ .

*Démonstration.* Par définition,

$$U(i, j) = \sum_{n \geq 0} P^n(i, j) = \sum_{n \geq 0} \mathbb{P}_i(X_n = j) = \sum_{n \geq 0} \mathbb{E}_i[\mathbf{1}_{(X_n=j)}],$$

d'où (3.4). Et donc pour  $i \neq j$ ,

$$\begin{aligned} U(i, j) &= \mathbb{E}_i \left[ \sum_{n \geq 1} \mathbf{1}_{(X_n=j)} \right] = \mathbb{E}_i \left[ \sum_{n \geq T_j} \mathbf{1}_{(X_n=j)} \right] \\ &\stackrel{(1)}{=} \mathbb{E}_i[\mathbf{1}_{(T_j < +\infty)}] \mathbb{E}_j \left[ \sum_{n \geq T_j} \mathbf{1}_{(X_n=j)} \right] = \rho_{ij} U(j, j). \end{aligned}$$

(1) par la propriété de Markov forte. □

Réurrence et transience sont données par les conditions simples suivantes.

**Théorème 3.6.**

*A. Les propositions suivantes sont équivalentes :*

1. *L'état  $i \in E$  est récurrent.*
2.  $\mathbb{P}_i(N_i = +\infty) = 1$ .
3.  $\mathbb{E}_i N_i = +\infty$ , soit  $U(i, i) = +\infty$ .

*B. Les propositions suivantes sont équivalentes :*

1. *L'état  $i \in E$  est transient.*
2.  $\mathbb{P}_i(N_i = +\infty) = 0$ .
3.  $\mathbb{E}_i N_i < +\infty$ , soit  $U(i, i) < +\infty$ .

*Démonstration.* La suite des événements  $(N_i \geq m)$  étant décroissante et grâce au lemme 3.4, nous avons

$$\begin{aligned} \mathbb{P}_i(N_i = +\infty) &= \lim_{m \rightarrow +\infty} \mathbb{P}_i(N_i \geq m) = \lim_{m \rightarrow +\infty} \rho_{ii}^m \\ &= \begin{cases} 0 & \text{si } \rho_{ii} < 1, \text{ état transient,} \\ 1 & \text{si } \rho_{ii} = 1, \text{ état récurrent.} \end{cases} \end{aligned}$$

De plus,

$$\begin{aligned} \mathbb{E}_i N_i &= \sum_{n \geq 0} \mathbb{P}_i(N_i \geq n) = \sum_{n \geq 0} \rho_{ii}^n \\ &= \begin{cases} \rho_{ii}(1 - \rho_{ii})^{-1} < +\infty & \text{si } \rho_{ii} < 1, \text{ état transient} \\ +\infty & \text{si } \rho_{ii} = 1, \text{ état récurrent.} \end{cases} \end{aligned}$$

Ainsi, A et B sont démontrés.

Si  $i$  est transient, on déduit du lemme 3.4 que

$$\mathbb{P}_i(N_i = k) = \mathbb{P}_i(N_i \geq k-1) - \mathbb{P}_i(N_i \geq k) = \rho_{ii}^{k-1} - \rho_{ii}^k = \rho_{ii}^{k-1}(1 - \rho_{ii}),$$

c'est-à-dire que  $N_i$  suit une loi géométrique de paramètre  $1 - \rho_{ii}$  par rapport à  $\mathbb{P}_i$ . □

**Théorème 3.7.** Soient  $j \in E$  un état récurrent et  $i \neq j$  élément de  $E$ .

Alors, soit  $\rho_{ji} = 0$  et dans ce cas  $\mathbb{P}_j(N_i = 0) = 1$ , soit  $\rho_{ji} = 1$  et dans ce cas  $\mathbb{P}_j(N_i = +\infty) = 1$  et  $i$  est aussi un état récurrent.

*Démonstration.* Supposons que  $\rho_{ji} = \mathbb{P}_j(T_i < +\infty) > 0$ . L'état  $j$  étant récurrent, la chaîne y passe infiniment souvent et nous pouvons définir  $N_i^n(j)$ , nombre de passages en  $i$  entre le  $(n-1)$ -ième et le  $n$ -ième passage en  $j$ , pour  $n \geq 1$ . Nous avons  $N_i = \sum_{n \geq 0} N_i^n(j)$ , avec  $N_i^0(j) = 0$ , et, par la propriété de Markov forte, la suite  $(N_i^n(j))$  est i.i.d. pour la probabilité  $P_j$ . Or  $\mathbb{P}_j(T_i < +\infty) > 0$  donc  $\mathbb{P}_j(N_i^1 \geq 1) > 0$  et par conséquent  $\mathbb{P}_j[N_i^n(j) \geq 1] > 0$ . Par le lemme de Borel-Cantelli, nous en déduisons que  $\mathbb{P}_j[\limsup (N_i^n(j) \geq 1)] = 1$ , donc la chaîne passe une infinité de fois en  $i$  en partant de  $j$ , soit  $\mathbb{P}_j(N_i = +\infty) = 1$ , d'où  $\mathbb{P}_j(T_i < +\infty) = 1$ .

Autrement dit,  $\mathbb{P}_j(T_i < +\infty)$  est bien égal à 0 ou 1. Si  $\mathbb{P}_j(T_i < +\infty) = 0$ , nous avons évidemment  $\mathbb{P}_j(N_i = 0) = 1$ . Il reste à montrer que  $i$  est récurrent si  $\mathbb{P}_j(T_i < +\infty) = 1$ . Or

$$\mathbb{P}_j(N_i = +\infty) = \mathbb{P}_j(T_i < +\infty, N_{Y,i} = +\infty),$$

où  $Y_n = X_{T_i+n}$ . Par la propriété de Markov forte, ceci devient

$$1 = \mathbb{P}_j(T_i < +\infty) \mathbb{P}_i(N_i = +\infty),$$

donc  $\mathbb{P}_i(N_i = +\infty) = 1$ . □

En plus de la nature des états (récence ou transience), la possibilité (ou non) de passage entre eux en un nombre fini d'étapes est une propriété fondamentale dans l'étude des chaînes de Markov.

**Définition 3.6.** Soient  $i$  et  $j$  deux états de  $E$ . Si  $P^n(i, j) > 0$  pour un entier  $n$ , on dit que  $i$  conduit à  $j$  et l'on note  $i \rightarrow j$ . Si  $i \rightarrow j$  et  $j \rightarrow i$ , on dit que les états  $i$  et  $j$  communiquent et l'on note  $i \leftrightarrow j$ .

Remarquons que  $n$  ne peut pas être nul si  $i$  et  $j$  sont distincts.

**Proposition 3.8.** Deux états  $i$  et  $j$  distincts communiquent si et seulement si  $\rho_{ij}\rho_{ji} > 0$ .

*Démonstration.* Nous avons

$$\begin{aligned} \mathbb{P}_i(T_j < +\infty) &= \mathbb{P}_i\left[\bigcup_{n \geq 0} (T_j = n)\right] \\ &= \sum_{n \geq 0} \mathbb{P}_i(T_j = n) \end{aligned} \tag{3.5}$$

$$= \sum_{n \geq 0} \mathbb{P}_i(X_k \neq j, 1 \leq k \leq n-1, X_n = j). \tag{3.6}$$

D'après (3.5), si  $\rho_{ij} > 0$ , il existe un entier  $n$  tel que  $\mathbb{P}_i(T_j = n) > 0$ . Or  $\mathbb{P}_i(T_j = n) \leq \mathbb{P}_i(X_n = j)$  et  $\mathbb{P}_i(X_n = j) = P^n(i, j)$  donc  $P^n(i, j) > 0$ .

Réciproquement, si  $P^n(i, j) = 0$  pour tout entier  $n$ , alors  $\mathbb{P}_i(X_k \neq j, 1 \leq k \leq n-1, X_n = j) = 0$ . Nous en déduisons par (3.6) que  $\mathbb{P}_i(T_j < +\infty) = 0$ , c'est-à-dire que  $i$  et  $j$  ne communiquent pas. □

La relation de communication  $\leftrightarrow$  est une relation d'équivalence sur  $E$  :

elle est réflexive puisque  $P^0(i, i) = 1$  ;

symétrique par symétrie de la définition de  $\leftrightarrow$

; transitive puisque  $P^{n+m}(i, k) \geq P^n(i, j)P^m(j, k) > 0$  et  $P^{n'+m'}(k, i) \geq P^{n'}(k, j)P^{m'}(j, i) > 0$ .

On peut donc écrire  $E$  comme réunion disjointe des classes d'équivalence de cette relation.

Remarquons que tous les éléments d'une classe donnée communiquent entre eux et qu'ils ne communiquent avec aucun état extérieur à la classe. Récurrence et transience sont des propriétés de classe.

**Proposition 3.9.** *Tous les états d'une classe de la relation  $\leftrightarrow$  sont de même nature.*

Les états d'une même classe sont donc tous transients, tous récurrents positifs ou tous récurrents nuls. Selon le cas, la classe sera dite transiente ou récurrente positive ou nulle.

*Démonstration.* Soient  $i$  et  $j$  deux éléments d'une même classe. Il existe deux entiers positifs  $n$  et  $m$  tels que  $P^n(i, j) > 0$  et  $P^m(j, i) > 0$ . Nous avons pour tout entier  $r$ ,

$$P^{n+r+m}(i, i) \geq P^n(i, j)P^r(j, j)P^m(j, i),$$

d'où

$$\sum_{r \geq 0} P^{n+r+m}(i, i) \geq P^n(i, j)P^m(j, i) \sum_{r \geq 0} P^r(j, j). \quad (3.7)$$

Or  $\mathbb{E}_i N_i \geq \sum_{r \geq 0} P^{n+r+m}(i, i)$  et  $\mathbb{E}_j N_j = \sum_{r \geq 0} P^r(j, j)$ . Si  $j$  est récurrent,  $\mathbb{E}_j N_j = +\infty$ , ce qui entraîne  $\mathbb{E}_i N_i = +\infty$ , c'est-à-dire que  $i$  est également récurrent. Et si  $i$  est transient, alors  $\mathbb{E}_i N_i < +\infty$ , et (3.7) donne  $\mathbb{E}_j N_j < +\infty$ , donc  $j$  est également transient.  $\square$

Une chaîne de Markov (ou sa matrice de transition) dont tous les états sont récurrents (respectivement transients) est dite récurrente (respectivement transiente). Une chaîne de Markov dont tous les états communiquent est dite irréductible ; elle comporte exactement une classe de communication et sera dite transiente ou récurrente selon le cas.

▷ **Exemple 3.9** *Marche aléatoire simple, suite de l'exemple 3.3.* Par la relation (3.1) page 51, en utilisant la formule de Stirling, nous obtenons

$$P^{2n}(0, 0) \simeq \frac{[4p(1-p)]^n}{\sqrt{\pi n}}.$$

Remarquons que  $4p(1-p) \leq 1$  pour tout  $p \in [0, 1]$ . L'égalité est vérifiée uniquement pour  $p = 1/2$  et dans ce cas  $\sum_{n \geq 0} P^n(0, 0)$  est divergente et l'état 0 est récurrent nul. Elle est convergente pour  $p \neq 1/2$  et, d'après le théorème 3.6, l'état 0 est alors transient.

La chaîne étant irréductible, elle sera récurrente ou transiente suivant la nature de l'état 0.

◁

**Définition 3.7.** Un sous-ensemble  $F$  de  $E$  est dit fermé (ou clos) si pour tout  $i \in F$ , on a  $\mathbb{P}_i(X_1 \in F) = 1$ .

Un ensemble fermé fini est dit final, car si la chaîne rentre dans un tel ensemble, elle y reste indéfiniment. Notons que si un ensemble fermé contient un seul état, alors cet état est absorbant. Une chaîne est irréductible si  $E$  ne contient pas deux sous-ensembles fermés disjoints.

**Proposition 3.10.** *Toute classe de communication récurrente est fermée.*

*Démonstration.* Si la classe  $C$  n'est pas fermée, il existe un état  $i \in C$  qui conduit à un état  $j \notin C$ , donc il existe un  $m \geq 1$ , tel que  $\mathbb{P}_i(X_m = j) > 0$ , mais  $j$  ne conduit pas à  $i$  donc  $\mathbb{P}_i(T_i < +\infty) < 1$ . Par conséquent, la probabilité de repasser infiniment souvent en  $i$  n'est pas égale à 1 et  $i$  est nécessairement transient, et  $C$  aussi d'après la proposition 3.9.  $\square$

Nous admettrons le résultat suivant.

**Théorème 3.11.** *Une chaîne de Markov irréductible est récurrente positive si et seulement si le système d'équations*

$$\sum_{j \in E} P(i, j)x_j = x_i, \quad i \in E, \quad (3.8)$$

*possède une solution non identiquement nulle  $x = (x_i)_{i \in E}$  absolument sommable (soit  $\sum_{i \in E} |x_i| < +\infty$ ). De plus la solution est unique à une constante multiplicative près et les  $x_i$  sont tous de même signe.*

Ce système s'écrit de façon matricielle  $Px = x$ . Si  $E$  est fini, le vecteur colonne  $x = \mathbf{1}$  est donc un vecteur propre de la matrice  $P$ , il est absolument sommable et la chaîne est alors récurrente positive.

### 3.3 Loi stationnaire et comportement asymptotique

Les lois stationnaires sont étroitement liées au comportement asymptotique des chaînes de Markov.

**Définition 3.8.** Soit  $\lambda$  une mesure sur  $E$ . Elle est appelée mesure invariante ou stationnaire de la chaîne de Markov  $(X_n)$  (ou de sa fonction de transition  $P$ ), si

$$\sum_{i \in E} \lambda(i)P(i, j) = \lambda(j), \quad j \in E.$$

Cette relation peut s'écrire sous forme matricielle  $\lambda P = \lambda$ , où  $\lambda = (\lambda(i))_{i \in E}$  est un vecteur ligne invariant de  $P$ .

Si de plus  $\lambda$  est une probabilité, elle est appelée loi invariante ou stationnaire de la chaîne et est notée  $\pi$  en général.

Il est facile de vérifier qu'une mesure invariante de  $P$  est une mesure invariante de  $P^k$ , pour tout entier  $k$ . Si la loi initiale de la chaîne est  $\pi$ , probabilité stationnaire de  $P$ , la loi de  $X_n$  est  $\pi$  aussi pour tout  $n$ , et  $(X_n)$  est dite stationnaire.

▷ **Exemple 3.10** *Une chaîne de Markov binaire.* Soit  $(X_n)$  une chaîne de Markov à deux états, de matrice de transition

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

avec  $0 < p < 1$  et  $0 < q < 1$ . Sa loi stationnaire vérifie  $\pi P = \pi$ , donc est déterminée par les équations

$$\begin{cases} (1-p)\pi(0) + q\pi(1) = \pi(0) \\ \pi(0) + \pi(1) = 1. \end{cases}$$

Par conséquent,  $\pi(0) = q/(p+q)$  et  $\pi(1) = p/(p+q)$ .

◁

▷ **Exemple 3.11** *Loi stationnaire de la chaîne bidimensionnelle de l'exemple 3.7.* Si  $(X_n)$  admet pour loi stationnaire  $\pi = (\pi(i))$ , alors  $(X_{n-1}, X_n)$  a pour loi stationnaire  $\tilde{\pi}$  définie par  $\tilde{\pi}(i, j) = \pi(i)P(i, j)$ , pour  $(i, j) \in E^2$ . En effet, comme  $\pi$  est la loi stationnaire de  $(X_n)$ , elle vérifie

$$\sum_{k \in E} \pi(k)P(k, i)P(i, j) = \pi(i)P(i, j), \quad i, j \in E,$$

soit

$$\sum_{(k, l) \in E^2} \tilde{\pi}(k, l)\delta_{li}P(i, j) = \tilde{\pi}(i, j),$$

c'est-à-dire précisément  $\tilde{\pi}\tilde{P} = \tilde{\pi}$ . De plus,  $\sum_{(i, j) \in E^2} \tilde{\pi}(i, j) = 1$ .

&lt;

**Définition 3.9.** Soit  $\lambda$  une mesure sur  $E$ . Elle est dite réversible pour la chaîne de Markov  $(X_n)$  (ou pour sa fonction de transition  $P$ ), si

$$\lambda(i)P(i, j) = \lambda(j)P(j, i), \quad (i, j) \in E \times E.$$

Une mesure réversible est invariante, puisque l'on a alors

$$\sum_{i \in E} \lambda(i)P(i, j) = \sum_{i \in E} \lambda(j)P(j, i) = \lambda(j) \sum_{i \in E} P(j, i) = \lambda(j), \quad j \in E.$$

De plus, elle est plus facile à déterminer lorsqu'elle existe, comme le confirme l'exemple suivant.

▷ **Exemple 3.12** Chaîne de naissance et de mort, suite de l'exemple 3.4.- Si  $(X_n)$  admet une loi réversible  $\pi$ , alors

$$\pi(i)p_i = \pi(i+1)q_{i+1}, \quad i \geq 0. \quad (3.9)$$

En posant

$$\gamma_i = \frac{p_0 \cdots p_{i-1}}{q_1 \cdots q_i}, \quad i \geq 1, \quad \text{et} \quad \gamma_0 = 1,$$

la relation (3.9) implique que  $\pi(i) = \pi(0)\gamma_i$ . En supposant que  $\sum_{i \geq 0} \gamma_i < +\infty$ , et en sommant les deux membres de l'égalité ci-dessus sur  $i$ , on obtient  $1/\pi(0) = \sum_{i \geq 0} \gamma_i$ , et donc

$$\pi(i) = \frac{\gamma_i}{\sum_{i \geq 0} \gamma_i}, \quad i \geq 0,$$

et  $\pi$  est une loi stationnaire de cette chaîne.

&lt;

Lorsque  $n$  tend vers l'infini, la chaîne ne repasse que dans les états récurrents, c'est-à-dire que  $P^n(i, j)$  tend vers 0 pour tout  $j$  transient. Il reste à étudier les états récurrents, lorsqu'il en existe.

**Théorème 3.12.** Si  $i$  et  $j$  sont des états récurrents, alors,

$$N_i^n/n \xrightarrow{\text{p.s.}} 1/m_i \quad \text{et} \quad N_{ij}^n/n \xrightarrow{\text{p.s.}} P(i, j)/m_i, \quad n \rightarrow +\infty,$$

où  $m_i = \mathbb{E}_i T_i$  est le temps de récurrence moyen en  $i$ , et avec la convention  $1/+\infty = 0$ .

*Démonstration.* Soient  $T_i^n$  le temps du  $n$ -ième passage de la chaîne dans l'état  $i$  et  $u_n = T_i^n - T_i^{n-1}$  le temps entre deux passages successifs dans ce même état, pour  $n \geq 1$ , avec la convention  $u_0 = 0$ . Nous avons  $T_i^{N_i^n} \leq n < T_i^{N_i^n+1}$ , soit

$$\frac{T_i^{N_i^n}}{N_i^n} \leq \frac{n}{N_i^n} < \frac{T_i^{N_i^n+1}}{N_i^n}.$$

Par la propriété de Markov forte,  $(u_n)$  est une suite i.i.d.. Comme  $(N_i^n)$  tend vers l'infini p.s. lorsque  $n$  tend vers l'infini, le théorème ?? infra donne

$$\frac{T_i^{N_i^n}}{N_i^n} = \frac{u_1 + \cdots + u_{N_i^n}}{N_i^n} \xrightarrow{\text{p.s.}} \mathbb{E}_i u_i = m_i.$$

De même,

$$\frac{T_i^{N_i^n+1}}{N_i^n} = \frac{u_1 + \cdots + u_{N_i^n+1}}{N_i^n} = \frac{u_1 + \cdots + u_{N_i^n+1}}{N_i^n + 1} \cdot \frac{N_i^n + 1}{N_i^n} \xrightarrow{\text{p.s.}} \mathbb{E}_i T_i = m_i,$$

d'où le premier résultat.

En suivant la même démarche appliquée à la chaîne de Markov bidimensionnelle  $(X_{n-1}, X_n)$  (voir l'exemple 3.7), nous obtenons le deuxième résultat.  $\square$



**Théorème 3.13.** *Si  $\pi$  est une loi stationnaire de  $(X_n)$ , alors,  $\pi(i) = 1/m_i$  pour tout état  $i$  récurrent.*

*Démonstration.* Par le théorème 3.12 et le théorème de convergence dominée (puisque  $N_i^n \leq n$  pour tout  $i$  et tout  $n$ ), nous obtenons pour tout  $j$  récurrent,

$$\frac{\mathbb{E}_j N_i^n}{n} \rightarrow \frac{1}{m_i}, \quad n \rightarrow +\infty.$$

D'autre part, puisque  $\pi$  est stationnaire, nous avons, pour tout  $k \geq 1$ ,

$$\pi(i) = \sum_{j \in E} \pi(j) P^k(j, i),$$

donc

$$n\pi(i) = \sum_{j \in E} \pi(j) \sum_{k=1}^n P^k(j, i) = \sum_{j \in E} \pi(j) \mathbb{E}_j N_i^n,$$

d'où

$$\pi(i) = \sum_{j \in E} \pi(j) \frac{\mathbb{E}_j N_i^n}{n} \rightarrow \sum_{j \in E} \pi(j) \frac{1}{m_i} = \frac{1}{m_i},$$

le résultat cherché.  $\square$

L'espérance du nombre de passages dans un certain état avant retour dans un état fixé fournit une mesure invariante à toute chaîne récurrente et une probabilité stationnaire à toute chaîne irréductible récurrente positive.

**Théorème 3.14.** *Soit  $(X_n)$  une chaîne de Markov récurrente d'espace d'état  $E$ . Pour tout  $j \in E$ , la mesure  $\lambda^j$  sur  $E$  définie par  $\lambda^j(i) = \mathbb{E}_j N_i^{T_j}$  est une mesure invariante de  $(X_n)$ .*

*De plus, si  $(X_n)$  est irréductible, alors  $0 < \lambda^j(i) < +\infty$  pour tout  $i \in E$ .*

*Démonstration.* Soient  $j \in E$  fixé et  $i \in E$ . Nous avons

$$\lambda^j(i) = \mathbb{E}_j N_i^{T_j} = \sum_{n \geq 1} \mathbb{P}_j(X_n = i, n \leq T_j)$$

avec

$$\begin{aligned} \mathbb{P}_j(X_n = i, n \leq T_j) &= \sum_{k \in E} \mathbb{P}_j(X_n = i, X_{n-1} = k, n \leq T_j) \\ &= \sum_{k \in E} \mathbb{P}_j(X_{n-1} = k, n \leq T_j) P(k, i). \end{aligned}$$

Remarquons que  $\lambda^j(j) = 1$  et que  $\lambda^j(i) = \mathbb{E}_j N_i^{T_j} = \mathbb{E}_j N_i^{T_j-1}$  pour  $i \neq j$ . Par conséquent,

$$\begin{aligned} \lambda^j(i) &= \sum_{k \in E} P(k, i) \sum_{n \geq 1} \mathbb{P}_j(X_{n-1} = k, n \leq T_j) \\ &= \sum_{k \in E} P(k, i) \sum_{n \geq 0} \mathbb{P}_j(X_n = k, n \leq T_j - 1) \\ &= \sum_{k \in E} P(k, i) \mathbb{E}_j N_k^{T_j-1} = \sum_{k \in E} P(k, i) \lambda^j(k), \end{aligned}$$

et  $\lambda^j$  est une mesure invariante de  $P$ .

Si la chaîne est irréductible, il existe pour tout  $i$  des entiers  $n > 0$  et  $m > 0$  tels que  $P^n(i, j) > 0$  et  $P^m(j, i) > 0$ . La mesure  $\lambda^j$  est une mesure invariante de  $P^n$  et de  $P^m$  donc, d'une part,  $\lambda^j(j) \geq \lambda^j(i) P^n(i, j)$  donc  $\lambda^j(i)$  est fini et, d'autre part,  $\lambda^j(i) \geq \lambda^j(j) P^m(j, i) > 0$ .  $\square$

Nous admettrons les résultats suivants.

**Proposition 3.15.** *Soient  $(X_n)$  une chaîne de Markov irréductible récurrente et  $\lambda$  une mesure invariante de  $(X_n)$ . On a*

$$\frac{\lambda(i)}{\lambda(j)} = \lambda^j(i), \quad (i, j) \in E \times E.$$

La mesure  $\lambda^j$  est donc, à une constante multiplicative près, la seule mesure invariante de la chaîne.

**Corollaire 3.16.** *Si  $(X_n)$  est une chaîne de Markov récurrente et  $\lambda$  une mesure invariante de  $(X_n)$ , alors, soit  $\lambda(E) = +\infty$  et  $m_i = +\infty$ , pour tout  $i \in E$ , soit  $\lambda(E) < +\infty$  et il existe une unique probabilité  $\pi$  invariante de  $(X_n)$  et  $m_i = 1/\pi(i)$ , pour tout  $i \in E$ .*

Il existe donc trois types d'espaces d'états pour les chaînes irréductibles :

1. tous les états sont transients ;
2. tous les états sont récurrents nuls ;
3. tous les états sont récurrents positifs et il existe une probabilité stationnaire unique.

Si  $E$  n'est pas irréductible, il peut être décomposé en deux sous-ensembles disjoints,  $E_t$  contenant les états transients et  $E_r$  contenant les états récurrents. De plus, les classes récurrentes  $(C_k)$  de la chaîne constituent une partition de  $E_r$ , voir la figure ??.

Si les états de  $C_k$  sont numérotés de  $|C_1| + \dots + |C_{k-1}| + 1$  à  $|C_1| + \dots + |C_k|$ , la matrice de transition  $P$  de  $(X_n)$  peut s'écrire

$$P = \begin{pmatrix} B & Q_1 & Q_2 & Q_3 & \dots \\ 0 & P_1 & 0 & 0 & \dots \\ 0 & 0 & P_2 & 0 & \dots \\ 0 & 0 & 0 & P_3 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

où  $P_k$  est la matrice de transition de la restriction de  $(X_n)$  à  $C_k$ ,  $B$  est la matrice des transitions de  $E_t$  à  $E_t$  et  $Q_k$  est la matrice des transitions de  $E_t$  à  $C_k$ . De plus, si  $\pi_k$  est la loi stationnaire de la restriction de  $(X_n)$  à  $C_k$ , alors

$$\Pi = ( 0 \mid \alpha_1 \pi_1 \mid \alpha_2 \pi_2 \mid \dots )$$

(où la partition est associée à  $E_t, C_1, C_2, \dots$ ) est une probabilité stationnaire de  $(X_n)$  pour tout  $(\alpha_k)$  tel que  $\alpha_k \geq 0$  et  $\sum_{k \geq 1} \alpha_k = 1$ , comme nous allons le voir sur l'exemple suivant.

▷ **Exemple 3.13** Soit  $(X_n)$  une chaîne de Markov de matrice de transition

$$P = \begin{pmatrix} 0.5 & 0.4 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.3 & 0.6 & 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.6 & 0.1 & 0.1 & 0.1 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.6 & 0.4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.4 & 0.6 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.6 & 0.2 & 0 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & 0.6 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0.3 & 0 & 0 & 0.7 \end{pmatrix},$$

voir la figure 3.2.

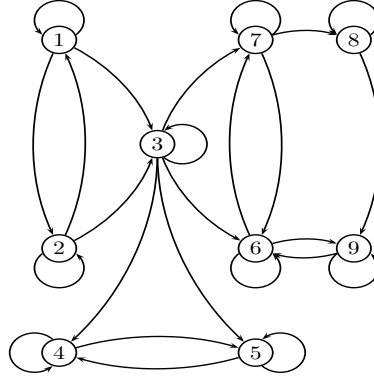


FIGURE 3.2 – Graphe de la chaîne de l'exemple 3.13.

On a  $E = \{1, \dots, 9\}$ , avec  $E_t = \{1, 2, 3\}$ ,  $C_1 = \{4, 5\}$  et  $C_2 = \{6, 7, 8, 9\}$ , dont on déduit que  $\pi(1) = (1/2, 1/2)$  et que  $\pi(2) = (30/73, 15/73, 3/73, 25/73)$ . On vérifie que

$$\Pi = \left( 0, 0, 0, \frac{\alpha}{2}, \frac{\alpha}{2}, \frac{30}{73}(1-\alpha), \frac{15}{73}(1-\alpha), \frac{3}{73}(1-\alpha), \frac{25}{73}(1-\alpha) \right)$$

est bien une loi stationnaire pour la chaîne pour tout  $\alpha \in [0, 1]$ .

◁

**Définition 3.10.** Un état récurrent positif et apériodique est dit ergodique.

Une chaîne de Markov irréductible dont un état est ergodique (et donc dont tous les états sont ergodiques) est dite ergodique.

**Définition 3.11.** Soit  $(X_n)$  une chaîne de Markov de matrice de transition  $P$ . Soit  $i \in E$ . Si

$$P^n(i, j) \rightarrow \pi^i(j), \quad j \in E, \quad n \rightarrow +\infty,$$

pour une probabilité  $\pi^i$  sur  $E$ , cette probabilité est appelée loi limite de la chaîne.

Notons qu'une chaîne de Markov peut admettre plusieurs lois stationnaires et plusieurs lois limites, voir l'exemple 3.14 infra de chaîne périodique. Par contre, toute chaîne ergodique admet une loi limite unique et cette loi est la loi stationnaire de la chaîne d'après le résultat suivant.

**Théorème 3.17** (ergodique). *Si  $(X_n)$  est une chaîne de Markov ergodique de fonction de transition  $P$  et d'espace d'état  $E$ , alors sa loi stationnaire est loi limite unique  $\pi$  de la chaîne, soit*

$$P^n(i, j) \rightarrow \pi(j), \quad i, j \in E, \quad n \rightarrow +\infty.$$

*Démonstration.* Nous allons démontrer ce théorème par la technique de couplage des processus.

Soient  $\mu$  la loi initiale de  $(X_n)$ . Considérons  $(Y_n)$  une chaîne de Markov ergodique indépendante de  $(X_n)$ , ayant le même espace d'état  $E$  et la même fonction de transition  $P$ , et de loi initiale  $\nu$ . La loi stationnaire de  $(Y_n)$  est donc  $\pi$  aussi.

La chaîne bidimensionnelle  $(Z_n) = (X_n, Y_n)$  est aussi ergodique. Clairement, sa loi initiale est  $\mu \otimes \nu$ , sa fonction de transition  $Q$  est définie par  $Q((i, j), (i', j')) = P(i, i')P(j, j')$  et sa loi stationnaire est  $\pi \otimes \pi$ .

Soit  $T = \inf\{n \geq 0 : X_n = Y_n\} = \inf\{n \geq 0 : Z_n = (j, j), j \in E\}$ . Notons  $\bar{\mathbb{P}} = \mathbb{P} \otimes \mathbb{P}$ . Nous avons  $\mathbb{P}_{\mu \otimes \nu}(X_n = j, T \leq n) = \mathbb{P}_{\mu \otimes \nu}(Y_n = j, T \leq n)$ , donc

$$\begin{aligned} \mathbb{P}_{\mu}(X_n = j) &= \mathbb{P}_{\mu \otimes \nu}(X_n = j, T \leq n) + \mathbb{P}_{\mu \otimes \nu}(X_n = j, T > n) \\ &\leq \mathbb{P}_{\nu}(Y_n = j) + \mathbb{P}_{\mu \otimes \nu}(T > n). \end{aligned}$$

De même,

$$\mathbb{P}_\nu(Y_n = j) \leq \mathbb{P}_\mu(X_n = j) + \mathbb{P}_{\mu \otimes \nu}(T > n).$$

Nous en déduisons que

$$|\mathbb{P}_\mu(X_n = j) - \mathbb{P}_\nu(Y_n = j)| \leq \mathbb{P}_{\mu \otimes \nu}(T > n).$$

Or  $\mathbb{P}_{\mu \otimes \nu}(T > n)$  tend vers 0 lorsque  $n$  tend vers l'infini. Pour  $\mu = \delta_i$  et  $\nu = \pi$ , nous obtenons

$$|\mathbb{P}_i(X_n = j) - \pi(j)| \rightarrow 0, \quad n \rightarrow +\infty,$$

d'où le résultat.  $\square$

Le théorème ergodique s'exprime aussi sous la forme suivante.

**Théorème 3.18.** *Si  $(X_n)$  est une chaîne de Markov ergodique de loi stationnaire  $\pi$  et si  $g : E \rightarrow \mathbb{R}$  est telle que  $\sum_{i \in E} \pi(i) |g(i)|$  soit finie, alors*

$$\frac{1}{n} \sum_{k=1}^n g(X_k) \xrightarrow{p.s.} \sum_{i \in E} \pi(i) g(i), \quad n \rightarrow +\infty,$$

### 3.4 Chaînes de Markov périodiques

Nous supposons dans ce paragraphe que la chaîne de Markov est irréductible et nous noterons  $d_i$  la période de l'état  $i \in E$  (voir la définition 3.3).

**Théorème 3.19.** *Si  $i$  et  $j$  communiquent, alors  $d_i = d_j$ .*

*Démonstration.* Nous avons  $P^{d_i}(i, i) > 0$ . Comme  $i \leftrightarrow j$ , il existe deux entiers  $n > 0$  et  $m > 0$  tels que  $P^n(i, j) > 0$  et  $P^m(j, i) > 0$ . Donc

$$P^{m+d_i+n}(j, j) \geq P^m(j, i) P^{d_i}(i, i) P^n(i, j) > 0.$$

De même,  $P^{n+m+2d_i}(j, j) > 0$  et donc  $d_j$  divise  $n + m + 2d_i - (n + m + d_i) = d_i$ . Ainsi  $d_j \leq d_i$ . Par un raisonnement symétrique nous obtenons  $d_i \leq d_j$ . D'où la conclusion.  $\square$

Ainsi la périodicité est une propriété de classe. Autrement dit, si un état  $i$  d'une classe  $C$  est périodique de période  $d_i$ , alors tout autre état  $j \in C$  sera périodique et de même période, soit  $d_j = d_i = d$ . Par conséquent, si une chaîne de Markov irréductible comporte un état périodique de période  $d$ , tous ses états seront  $d$ -périodiques, et la chaîne sera dite  $d$ -périodique.

L'espace d'état  $E$  d'une chaîne  $d$ -périodique peut être décomposé en  $d$  sous-ensembles disjoints, soit  $E_0, E_1, \dots, E_{d-1}$ , de sorte que chaque transition conduit d'un état de  $E_p$  à un état de  $E_{p+1}$ , pour  $p = 0, 1, \dots, d-1$  (avec  $E_d \equiv E_0$ ). Ces ensembles sont appelés classes cycliques de la chaîne. En utilisant ces classes, l'étude de toute chaîne de Markov  $d$ -périodique peut se ramener à celle d'une chaîne apériodique, comme le montre la proposition suivante. Cela revient à observer la chaîne aux instants  $md$ .

**Proposition 3.20.** *Soit  $(X_n)$  une chaîne de Markov  $d$ -périodique de fonction de transition  $P$  et de classes cycliques  $E_0, \dots, E_{d-1}$ .*

1. *La suite  $(Y_m)$  de variables aléatoires définies par*

$$Y_m = X_{md}, \quad m \in \mathbb{N},$$

*est une chaîne de Markov apériodique de fonction de transition  $Q = P^d$ .*

2. *Les classes  $E_0, \dots, E_{d-1}$  sont des ensembles fermés pour  $Q$ , c'est-à-dire que  $Q(i, j) = 0$  pour tous  $i \in E_p$  et  $j \notin E_p$ . De plus, elles sont irréductibles pour  $Q$  (ou pour  $(Y_m)$ ).*

*Démonstration.* 1. Nous avons

$$\begin{aligned} \mathbb{P}(Y_{m+1} = j \mid Y_0 = i_0, Y_1 = i_1, \dots, Y_{m-1} = i_{m-1}, Y_m = i) &= \\ &= \mathbb{P}(X_{(m+1)d} = j \mid X_0 = i_0, X_d = i_1, \dots, X_{(m-1)d} = i_{m-1}, X_{md} = i) \\ &= \mathbb{P}(X_{(m+1)d} = j \mid X_{md} = i) = \mathbb{P}(Y_{m+1} = j \mid Y_m = i), \end{aligned}$$

donc  $(Y_m)$  est une chaîne de Markov. Et comme

$$\mathbb{P}(X_{(m+1)d} = j \mid X_{md} = i) = P^d(i, j),$$

sa fonction de transition est bien  $Q$ .

Nous avons pour tout état  $i$ ,

$$\begin{aligned} p.g.c.d.\{m \in \mathbb{N} : Q^m(i, i) > 0\} &= p.g.c.d.\{m \in \mathbb{N} : P^{md}(i, i) > 0\} \\ &= \frac{1}{d} p.g.c.d.\{n \in \mathbb{N} : P^n(i, i) > 0\} = 1, \end{aligned}$$

donc  $(Y_m)$  est apériodique.

2. Si  $i \in E_p$  et si  $Q(i, j) \neq 0$ , alors nécessairement, par définition des classes cycliques,  $j \in E_{p+md} \equiv E_p$ .

Soient  $i$  et  $j$  dans  $E_p$ . Comme  $(X_n)$  est irréductible, il existe  $n \in \mathbb{N}^*$  tel que  $P^n(i, j) > 0$ , et comme  $i$  et  $j$  sont dans  $E_p$ , nécessairement  $n = md$ , avec  $m > 0$ . Ainsi,  $P^{md}(i, j) > 0$  ou encore  $Q^m(i, j) > 0$  et par conséquent  $i \rightarrow j$ . De même,  $j \rightarrow i$ . Donc  $E_p$  est irréductible.  $\square$

**Théorème 3.21.** Soit  $(X_n)$  une chaîne de Markov irréductible, récurrente positive et  $d$ -périodique, de loi stationnaire  $\pi$ . Soient  $E_0, E_1, \dots, E_{d-1}$ , ses classes cycliques. Posons pour  $i \in E_p$  et  $j \in E_q$  avec  $0 \leq p, q \leq d-1$ ,

$$r = r_{ij} = \begin{cases} q - p & \text{si } q \geq p, \\ d - p + q & \text{si } q < p. \end{cases}$$

Alors

$$P^{nd+r}(i, j) \rightarrow d\pi(j), \quad n \rightarrow +\infty.$$

*Démonstration.* Ecrivons

$$P^{nd+r}(i, j) = \sum_{k \in E_q} P^r(i, k) P^{nd}(k, j), \quad i \in E_p, j \in E_q.$$

Comme  $\sum_{k \in E_q} P^r(i, k) = 1$  pour tout  $i \in E_p$  par définition de  $r$ , par le théorème de convergence dominée, il suffit de montrer que

$$P^{nd}(k, j) = Q^n(k, j) \rightarrow d\pi(j), \quad k, j \in E_q, \quad n \rightarrow +\infty \quad (3.10)$$

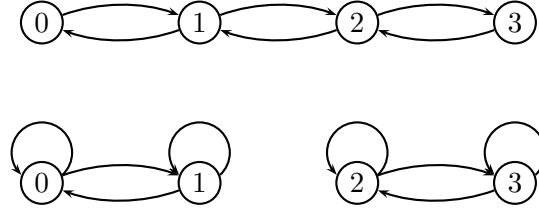
pour obtenir le résultat.

Toute transition se passe uniquement d'un état de  $E_q$  vers un état de  $E_{q+1}$ . Ceci implique que  $\sum_{l \in E_{q+1}} P(k, l) = 1$  pour tout  $k \in E_q$  et donc par le théorème de Fubini que

$$\sum_{l \in E_{q+1}} \sum_{k \in E_q} \pi(k) P(k, l) = \sum_{k \in E_q} \pi(k).$$

Comme  $\pi$  est une loi stationnaire de  $(X_n)$ , nous avons

$$\sum_{l \in E_{q+1}} \sum_{k \in E_q} \pi(k) P(k, l) = \sum_{l \in E_{q+1}} \pi(l).$$

FIGURE 3.3 – Graphes des chaînes de l'exemple 3.14,  $(X_n)$  en haut et  $(Y_n)$  en bas.

Nous déduisons de ces égalités que

$$\sum_{k \in E_q} \pi(k) = c, \quad 0 \leq q \leq d-1,$$

et nécessairement  $c = 1/d$ , puisque la chaîne admet  $d$  classes cycliques.

Comme  $\pi$  est une loi stationnaire de  $(X_n)$  donc une mesure stationnaire de  $Q$  et que  $Q(k, h) = 0$  si  $k \in E_q$  et  $h \notin E_q$ , nous avons

$$\sum_{k \in E_q} \pi(k) Q(k, h) = \pi(h), \quad h \in E_q.$$

Par conséquent,  $(d\pi(k))_{k \in E_q}$  est une loi stationnaire de  $Q_p$ , restriction de  $Q$  sur  $E_p \times E_p$ . Cette matrice de transition est apériodique et, d'après la proposition 3.20, elle est irréductible. Par le théorème ergodique, nous en déduisons (3.10) d'où le résultat.  $\square$

▷ **Exemple 3.14** Considérons la chaîne d'Ehrenfest à quatre états, soit  $E = \{0, 1, 2, 3\}$ , de matrice de transition

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

Cette chaîne est périodique de période  $d = 2$ , puisque

$$P^2 = \begin{pmatrix} 1/3 & 0 & 2/3 & 0 \\ 0 & 7/9 & 0 & 2/9 \\ 2/9 & 0 & 7/9 & 0 \\ 0 & 2/3 & 0 & 1/3 \end{pmatrix}.$$

Les classes cycliques sont ici  $E_0 = \{0, 2\}$  et  $E_1 = \{1, 3\}$ . Par l'équation  $\pi P = \pi$  nous déterminons sa loi stationnaire  $\pi = (1/8, 3/8, 3/8, 1/8)$ . La matrice de transition de la chaîne  $(Y_m)$ , définie par  $Y_m = X_{2m}$  pour  $m \geq 0$ , est  $Q = P^2$  et  $(Y_m)$  est une chaîne apériodique non irréductible.

Les restrictions de la chaîne  $(Y_n)$  à  $E_0$  et  $E_1$  sont ergodiques, de matrices de transitions

$$Q_0 = \begin{pmatrix} 1/3 & 2/3 \\ 2/9 & 7/9 \end{pmatrix} \quad \text{et} \quad Q_1 = \begin{pmatrix} 7/9 & 2/9 \\ 2/3 & 1/3 \end{pmatrix},$$

et de lois stationnaires  $(\pi'_0, \pi'_2) = (1/4, 3/4)$  et  $(\pi'_1, \pi'_3) = (3/4, 1/4)$ , respectivement. En normalisant le vecteur  $(1/4, 3/4, 3/4, 1/4)$ , nous retrouvons la loi stationnaire de la chaîne  $(X_n)$ , soit  $(1/8, 3/8, 3/8, 1/8)$ .

Enfin, par le théorème 3.21, nous obtenons, pour  $r = 0$ ,

$$P^{2n} \rightarrow \begin{pmatrix} 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \end{pmatrix}, \quad n \rightarrow +\infty,$$

et, pour  $r = 1$ ,

$$P^{2n+1} \rightarrow \begin{pmatrix} 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \\ 0 & 3/4 & 0 & 1/4 \\ 1/4 & 0 & 3/4 & 0 \end{pmatrix}, \quad n \rightarrow +\infty.$$

Les graphes de  $(X_n)$  et de  $(Y_m)$  sont donnés dans la figure 3.3.

◁

### 3.5 Chaînes de Markov finies

Savoir qu'une chaîne de Markov ne prend qu'un nombre fini d'états facilite son étude, grâce aux propriétés particulières qui en découlent. Notons que les chaînes de Markov utilisées en fiabilité sont en général finies, le cas des chaînes à deux états étant spécialement intéressant. Une chaîne de Markov  $(X_n)$  dont l'espace d'état  $E$  est de cardinal fini est dite finie. Sa fonction de transition est représentée par une matrice  $P$ . D'après les propriétés de la fonction de transition, les éléments de cette matrice sont positifs et la somme de chaque ligne est égale à 1. Une telle matrice est dite stochastique.

Si une matrice stochastique possède une puissance ne contenant aucun coefficient nul, elle est dite régulière. On peut montrer qu'une chaîne de Markov finie et apériodique est irréductible si et seulement si sa matrice de transition est régulière.

**Théorème 3.22.** *Si  $E$  est fini, alors il contient au moins un état récurrent.*

*Démonstration.* Soient  $j \in E_t$  et  $i \in E$ . Comme  $\mathbb{E}_i N_j$  est finie,  $P^n(i, j)$  tend vers 0 lorsque  $n$  tend vers l'infini. Si  $E$  ne contient que des états transients, alors  $\sum_{j \in E} P^n(i, j)$  tend vers 0 aussi puisque cette somme ne contient qu'un nombre fini de termes. Mais  $P$  est une matrice stochastique, donc

$$\sum_{j \in E} P^n(i, j) = \sum_{j \in E} \mathbb{P}_i(X_n = j) = 1.$$

D'où une contradiction. □

Une chaîne finie irréductible est donc toujours récurrente. La matrice de transition de la chaîne étant stochastique, elle admet le vecteur propre  $\mathbf{1} = (1, \dots, 1)'$  associé à la valeur propre 1. Donc le système (3.8) page 58 admet pour solution le vecteur  $\mathbf{1}$  qui est absolument sommable. On en conclut par le théorème 3.11 que toute chaîne de Markov finie irréductible est récurrente positive, donc a une unique probabilité invariante (par le corollaire 3.16).

**Proposition 3.23.** *Si  $(X_n)$  est une chaîne de Markov finie, alors  $\mathbb{P}_i(X_n \in E_r)$  tend vers 1 lorsque  $n$  tend vers l'infini, pour tout  $i \in E$ .*

Autrement dit, la chaîne entrera p.s. dans l'ensemble des états récurrents et y restera ensuite.

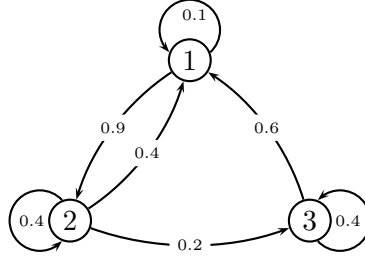


FIGURE 3.4 – Graphe de la chaîne de l'exemple 3.15.

*Démonstration.* Comme  $E$  est fini, si  $\mathbb{P}(X_0 = i) = 1$  avec  $i$  transient, il existe un  $n_i \geq 1$  tel que  $\mathbb{P}_i(X_{n_i} \in E_r) > 0$ .

Par conséquent,  $p = \mathbb{P}(X_M \in E_r \mid X_0 \in E_t) > 0$ , où  $M = \sup_{i \in E_t} n_i$ .

Ainsi,  $\mathbb{P}(X_M \in E_t \mid X_0 \in E_t) = 1 - p$ , et par la propriété de Markov,

$$\mathbb{P}(X_{kM} \in E_t \mid X_0 \in E_t) = (1 - p)^k \rightarrow 0, \quad k \rightarrow +\infty,$$

d'où la conclusion.  $\square$

Notons que la démonstration montre que le résultat est vrai pour toute chaîne du moment que  $E_t$  est fini.

▷ **Exemple 3.15** Soit  $(X_n)$  une chaîne de Markov d'espace d'état  $\{1, 2, 3\}$ , de matrice de transition

$$P = \begin{pmatrix} 0.1 & 0.9 & 0 \\ 0.4 & 0.4 & 0.2 \\ 0.6 & 0 & 0.4 \end{pmatrix}.$$

Son graphe est donné dans la figure 3.4. Tous les états communiquent, la chaîne est irréductible et finie, donc récurrente (d'après le théorème 3.22). Sa loi limite est solution de  $\pi P = \pi$ , soit

$$\begin{cases} 0.1\pi(1) + 0.4\pi(2) + 0.6\pi(3) = \pi(1) \\ 0.9\pi(1) + 0.4\pi(2) = \pi(2) \\ 0.2\pi(2) + 0.4\pi(3) = \pi(3), \end{cases}$$

avec  $\pi(1) + \pi(2) + \pi(3) = 1$ . On calcule  $(\pi(1), \pi(2), \pi(3)) = (1/6, 1/2, 1/3)$ .  $\triangleleft$

**Proposition 3.24.** Une loi  $\pi$  est stationnaire pour une chaîne de Markov  $(X_n)$  finie si et seulement si  $\pi$  est une loi limite de  $(X_n)$ .

*Démonstration.* Le sens direct est vrai pour toute chaîne ergodique à espace d'état fini ou dénombrable d'après le théorème ergodique.

Réciproquement, supposons que  $\pi$  est une loi limite de la chaîne. Nous avons

$$\begin{aligned} \pi(j) &= \lim_{n \rightarrow +\infty} P^n(i, j) = \lim_{n \rightarrow +\infty} \sum_{k \in E} P^{n-1}(i, k) P(k, j) \\ &= \sum_{k \in E} \lim_{n \rightarrow +\infty} P^{n-1}(i, k) P(k, j) = \sum_{k \in E} \pi(k) P(k, j), \end{aligned}$$

car  $E$  étant fini, l'interversion de la limite et de la somme est justifiée.  $\square$

▷ **Exemple 3.16** Soit  $(X_n)$  une chaîne de Markov finie d'espace d'état  $E$  et de matrice de transition  $P$ . Si  $P$  est doublement stochastique, c'est-à-dire si la somme de ses colonnes est égale à 1 comme la somme de ses lignes, alors la probabilité uniforme sur  $E$  est clairement une loi stationnaire (et limite) de  $P$ .



&lt;

La convergence d'une matrice de transition vers sa limite est exponentielle.

**Proposition 3.25.** *Si  $(X_n)$  est une chaîne de Markov finie ergodique, de matrice de transition  $P$  et de loi stationnaire  $\pi$ , alors  $P^n$  converge avec une vitesse exponentielle vers la matrice  $\Pi$  telle que  $\Pi(i, j) = \pi(j)$ , pour  $(i, j) \in E \times E$ .*

*Démonstration.* La chaîne  $(X_n)$  étant irréductible et apériodique,  $P$  est régulière et il existe un entier  $r$  tel que  $\rho = \min_{(i,j) \in E \times E} P^r(i, j) > 0$ .

Soit  $(Y_n)$  une chaîne de Markov ergodique, indépendante de  $(X_n)$ , ayant le même espace d'état  $E = \{1, \dots, N\}$  et la même fonction de transition  $P$ , et de loi initiale  $\pi$ . Par conséquent  $\pi$  est la loi de  $Y_n$  pour tout  $n$ . Posons  $s = \lfloor n/r \rfloor$  pour  $n \geq r$ ; nous avons

$$\begin{aligned} \mathbb{P}_{(k,j)}(X_i \neq Y_i, i \leq n) &\leq \mathbb{P}_{(k,j)}(X_i \neq Y_i, i = r, 2r, \dots, sr) \\ &\leq \mathbb{P}_{(k,j)}(X_r \neq Y_r) \mathbb{P}_{(k,j)}(X_{2r} \neq Y_{2r} \mid X_r \neq Y_r) \cdots \\ &\quad \cdots \mathbb{P}_{(k,j)}(X_{sr} \neq Y_{sr} \mid X_{tr} \neq Y_{tr}, t = 1, \dots, s-1). \end{aligned} \quad (3.11)$$

Or,

$$\begin{aligned} \mathbb{P}(X_r = Y_r \mid X_0 = k, Y_0 = j) &= \sum_{i=1}^N \mathbb{P}_{(k,j)}[(X_r, Y_r) = (i, i)] \\ &= \sum_{i=1}^N P^r(k, j) P^r(i, i) \geq N \rho^2 > 0. \end{aligned}$$

Donc  $\mathbb{P}_{(k,j)}(X_r \neq Y_r) = 1 - \mathbb{P}_{(k,j)}(X_r = Y_r) \leq 1 - N \rho^2$ , et de même pour les autres termes de (3.11), d'où

$$\mathbb{P}_{(k,j)}(X_i \neq Y_i, i \leq n) \leq (1 - N \rho^2)^s.$$

De plus,  $\mathbb{P}_{(k,j)}(X_i \neq Y_i, i \leq n) = \mathbb{P}_{(k,j)}(T_{(i,i)} > n)$ , où  $T_{(i,i)}$  est le temps de retour dans l'état  $(i, i)$  de la chaîne bidimensionnelle  $(X_n, Y_n)$ . Comme  $\pi$  est la loi de  $Y_n$  pour tout  $n$ , la conclusion s'en déduit.  $\square$

▷ **Exemple 3.17** Soit  $(X_n)$  une chaîne de Markov finie d'espace d'état  $E = \{1, \dots, N\}$  et de matrice de transition  $P = (1 - \theta)I + \theta\Pi$ , où  $\theta \in [0, 1]$ ,  $I$  est la matrice identité  $N \times N$  et  $\Pi(i, j) = \pi(j)$  pour tout  $i$ , pour une probabilité  $\pi$  sur  $E$ .

Montrons que  $\pi$  est la loi stationnaire de  $(X_n)$ . Si  $\pi$  est la loi limite de la chaîne, on sait qu'elle est stationnaire. Comme  $\Pi^2 = \Pi$ , on obtient

$$P^n = (1 - \theta)^n I + [1 - (1 - \theta)^n] \Pi.$$

Donc  $P^n$  tend vers  $\Pi$  et la vitesse de convergence est géométrique, égale à  $(1 - \theta)^n$ .

&lt;

Si la chaîne  $(X_n)$  comporte à la fois des états transients et des états récurrents, on peut écrire sa matrice de transition  $P$  sous la forme

$$P = \begin{pmatrix} E_t & E_r \\ Q & B \\ 0 & A \end{pmatrix} \begin{matrix} E_t \\ E_r \end{matrix}.$$

On appelle alors matrice fondamentale de  $(X_n)$  la matrice potentiel  $U$  de  $Q$ , soit

$$U = I + Q + Q^2 + \cdots + Q^n + \cdots$$

à coefficients dans  $\overline{\mathbb{R}}_+$ .

**Théorème 3.26.** Soit  $(X_n)$  une chaîne de Markov finie non irréductible. Soient  $T = \inf\{n \in \mathbb{N} : X_n \in E_r\}$  le temps d'entrée dans  $E_r$  et  $Q$  la restriction de  $P$  à  $E_t \times E_t$ . Alors :

1.  $Q^n$  tend vers la matrice nulle lorsque  $n$  tend vers l'infini ;
2.  $I - Q$  est non singulière et  $U = (I - Q)^{-1}$  ;
3.  $\mathbb{E}_i T = e_i(I - Q)^{-1} \mathbf{1}_{|E_t|}$ ,  $i \in E_t$ , où  $e_i = (0, \dots, 0, 1, 0, \dots, 0)$  est le  $i$ -ème vecteur ligne de la base canonique de  $\mathbb{R}^{|E_t|}$ .

*Démonstration.* 1. Posons  $h_n^i = \sum_{j \in E_t} Q^n(i, j)$  pour  $n \in \mathbb{N}$  et  $i \in E$ . Alors, pour tout  $i \in E_t$ , il existe un  $n \in \mathbb{N}^*$  tel que  $h_n^i > 0$ .

La suite  $(h_n^i)$  est décroissante en  $n$ , puisque

$$h_{n+1}^i = \sum_{j \in E_t} \sum_{\ell \in E_t} Q^n(i, \ell) Q(\ell, j) \leq \sum_{\ell \in E_t} Q^n(i, \ell) = h_n^i.$$

Par conséquent, il existe  $n_i$  et  $C_i$  tels que  $h_n^i < C_i < 1$  pour tout  $n > n_i$ , donc,  $E_t$  étant fini, il existe un  $C$  indépendant de  $i$ , tel que  $h_n^i < C < 1$ .

Nous en déduisons pour un  $n$  fixé que

$$h_{mn+n}^i = \sum_{j \in E_t} \sum_{\ell \in E_t} Q^{mn}(i, \ell) Q^n(\ell, j) \leq C h_{mn}^i \leq C^{m+1}.$$

Or  $C^{m+1}$  tend vers 0 lorsque  $m$  tend vers l'infini. Comme la suite  $(h_n^i)$  est décroissante et admet une sous-suite  $(h_{mn}^i)_{m \in \mathbb{N}}$  convergeant vers zéro, elle converge vers 0 aussi. Comme

$$Q^n(i, k) = \sum_{j \in E_t} Q^{n-1}(i, j) Q(j, k) \leq h_{n-1}^i,$$

$Q^n$  tend vers la matrice nulle.

2. Comme  $Q^n$  tend vers la matrice nulle, pour  $n$  assez grand, les valeurs propres de  $I - Q^n$  sont toutes non nulles et la matrice est non singulière. Nous déduisons de la relation

$$I - Q^n = (I - Q)(I + Q + \dots + Q^{n-1}),$$

que

$$0 \neq \det(I - Q^n) = \det(I - Q) \det(I + Q + \dots + Q^{n-1}),$$

d'où

$$I + Q + \dots + Q^{n-1} = (I - Q)^{-1}(I - Q^n).$$

En faisant tendre  $n$  vers l'infini dans cette égalité, nous obtenons bien par 1. que  $U = (I - Q)^{-1}$ .

3. Pour tout  $i \in E_t$ , nous avons

$$\mathbb{E}_i T = \sum_{j \in E} \mathbb{E}_i [T \mathbf{1}_{(X_1=j)}] = \sum_{j \in E} \mathbb{E}_i (T \mid X_1 = j) P(i, j).$$

Pour les états récurrents,

$$\sum_{j \in E_r} \mathbb{E}_i (T \mid X_1 = j) P(i, j) = \sum_{j \in E_r} P(i, j),$$

et pour les états transients,

$$\begin{aligned} \sum_{j \in E_t} \mathbb{E}_i (T \mid X_1 = j) P(i, j) &= \sum_{j \in E_t} \mathbb{E}_j (T + 1) P(i, j) \\ &= \sum_{j \in E_t} \mathbb{E}_j T P(i, j) + \sum_{j \in E_t} P(i, j), \end{aligned}$$

nous en déduisons que  $\mathbb{E}_i T = 1 + \sum_{j \in E_t} \mathbb{E}_j T P(i, j)$ .

En posant  $L = (\mathbb{E}_1 T, \dots, \mathbb{E}_{|E_t|} T)'$ , ceci s'écrit sous forme matricielle  $L = \mathbf{1}_{|E_t|} + QL$ , ou encore  $(I_{|E_t|} - Q)L = \mathbf{1}_{|E_t|}$ . Et comme la matrice  $I_{|E_t|} - Q$  est non singulière, le résultat s'en déduit.  $\square$

Si la chaîne  $(X_n)$  de matrice de transition  $P$  est ergodique, elle admet une loi stationnaire  $\pi$ . On appelle alors matrice fondamentale de  $(X_n)$  la matrice

$$Z = [I - (P - \Pi)]^{-1},$$

où  $\Pi$  est la matrice définie par  $\Pi(i, j) = \pi(j)$ .

**Théorème 3.27.** *La matrice fondamentale  $Z$  d'une chaîne de Markov ergodique vérifie*

$$Z = I + \sum_{n \geq 0} (P^n - \Pi).$$

*Démonstration.* Comme  $P^n$  tend vers  $\Pi$ , par les égalités  $P^{2n}P = PP^{2n} = P^{n+1}P^n = P^{2n+1}$ , en faisant tendre  $n$  vers l'infini, nous obtenons  $\Pi P = P\Pi = \Pi^2 = \Pi$ , ce qui donne

$$(P - \Pi)^n = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} P^k \Pi^{n-k} = P^n + \sum_{k=0}^{n-1} \binom{n}{k} (-1)^{n-k} \Pi = P^n - \Pi.$$

Par conséquent  $(P - \Pi)^n$  tend vers 0 lorsque  $n$  tend vers l'infini, ce qui conclut la démonstration.  $\square$

Notons que de la relation  $P\mathbf{1}_N = \mathbf{1}_N$ , on déduit que 1 est une valeur propre de  $P$ . On peut montrer que pour une chaîne de Markov finie régulière, c'est la seule valeur propre de module égal à un.

**Proposition 3.28.** *Les valeurs propres de la matrice de transition d'une chaîne de Markov finie sont de module inférieur ou égal à 1.*

*Démonstration.* Posons  $E = \{1, \dots, N\}$  et considérons la norme infinie  $\|a\| = \max_{1 \leq i \leq N} |a_i|$  sur  $\mathbb{C}^N$ . Nous avons pour tout  $a = (a_i) \in \mathbb{C}^N$ ,

$$\|Pa\| = \max_{1 \leq i \leq N} \left| \sum_{j=1}^N P(i, j) a_j \right| \leq \left[ \max_{1 \leq i \leq N} \sum_{j=1}^N P(i, j) \right] \left[ \max_{1 \leq i \leq N} |a_i| \right].$$

Or  $\sum_{j=1}^N P(i, j) = 1$ , donc  $\|Pa\| \leq \|a\|$ .

Pour  $Pa = \lambda a$ , nous en déduisons que  $|\lambda| \leq 1$ .  $\square$

Supposons que  $P$  ait  $N$  valeurs propres distinctes  $1 = \lambda_1 > \lambda_2 > \dots > \lambda_N$ . Soient  $x_1, \dots, x_N$  des vecteurs propres à droite de  $P$  et  $y_1, \dots, y_N$  des vecteurs propres à gauche de  $P$  correspondant aux valeurs propres  $\lambda_1, \dots, \lambda_N$ . Notons  $Q = (x_1, \dots, x_N)$  et  $D = \text{diag}(\lambda_i)$ . On a  $P = QDQ^{-1}$  et par conséquent  $P^n = QD^nQ^{-1}$ . On en déduit que

$$P^n = \sum_{k=1}^N \lambda_k^n A_k = \sum_{k=1}^N \frac{\lambda_k^n}{\langle x_k, y_k \rangle} x_k \cdot y'_k, \quad (3.12)$$

où  $\langle x_k, y_k \rangle$  est le produit scalaire de  $x_k$  et  $y_k$ , et  $x_k \cdot y'_k$  est leur produit matriciel.

Si  $|\lambda_k| < 1$ , alors  $\lambda_k^n$  tend vers 0, lorsque  $n$  tend vers l'infini. La représentation spectrale (3.12) montre que, ceci étant le cas pour tout  $k \geq 2$  lorsque la chaîne est ergodique, alors  $P^n$  tend vers  $A_1$  (à vitesse géométrique  $\lambda_2^n$ ). Et donc  $A_1 = \Pi$ .

### 3.6 Problèmes

**3.1.** Cinq points sont fixés sur un cercle. A chaque temps  $n = 1, 2, \dots$ , une particule saute à un des deux points voisins avec la probabilité  $1/2$  pour chaque voisin.

- Ecrire la matrice de transition ;
- représenter le graphe associé ;
- donner les classes d'états de la chaîne ;
- déterminer une probabilité stationnaire s'il existe.

**3.2.** (Le pays d'Oz). Dans le pays d'Oz il ne fait pas beau deux jours de suite. S'il fait beau (B) un jour, alors la journée suivante il neigera (N) ou il pleuve (P) avec la même probabilité. Si un jour il neige (ou il pleut), le lendemain il fera le même temps avec la probabilité  $1/2$  et il fera un changement avec la probabilité  $1/4$  pour chaque possibilité.

- Donner la matrice de transition pour la chaîne de Markov avec espace d'états  $E = \{P, B, N\}$ .
- Si aujourd'hui il fait beau, quel temps a la plus grande probabilité pour après demain ?
- Donner les classes d'états de la chaîne et déterminer une probabilité stationnaire s'il en existe une.

**3.3.** Donner les classes d'états des chaînes de Markov dont les matrices de transitions sont les suivantes : a)

$$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

b)

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \\ 0 & 1/3 & 1/3 & 1/3 \end{pmatrix}$$

c)

$$P = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \end{pmatrix}$$

d)

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \end{pmatrix}$$

**3.4.** On lance une monnaie  $n$  fois,  $n = 1, 2, \dots$  Pour chaque  $n$ , soit  $X_n$  la différence entre le nombre d'apparitions de la face et le nombre d'apparition de la pile.

- Montrer que  $(X_n, n \in \mathbb{N}^*)$  est une chaîne de Markov et préciser son espace d'états ;
- donner les classes d'états de la chaîne ;
- déterminer la probabilité stationnaire si elle existe.

**3.5.** Deux urnes contiennent  $N$  boules blanches et  $N$  boules noires de sorte que chaque urne contient  $N$  boules. A chaque instant  $n = 1, 2, \dots$  une boule de la première urne et une de la deuxième sont changées entre elles. Soit  $X_n$ ,  $n = 1, 2, \dots$  le nombre de boules blanches de la première urne après le  $n$ -ième changement.

- Montrer que  $(X_n, n \in \mathbb{N})$  est une chaîne de Markov et préciser son espace d'états et sa matrice de transition ;
- donner les classes d'états de la chaîne ;
- déterminer la probabilité stationnaire s'il en existe.

**3.6.** (Médian 2009). Soit une chaîne de Markov  $X = (X_n, n \geq 0)$  d'espace d'état  $E = \{1, 2, 3, 4, 5, 6\}$ , de matrice de transition

$$P = \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 & 0 \\ 1/2 & 1/6 & 1/6 & 1/6 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 1/4 & 1/4 \\ 0 & 0 & 0 & 0 & 1/2 & 1/2 \\ 1/4 & 1/4 & 0 & 0 & 1/4 & 1/4 \end{pmatrix}$$

et de loi initiale  $\alpha = (1, 0, 0, 0, 0, 0)$ .

- Tracer le graphe d'état de cette CM.
- Donner les classes de cette CM en les caractérisant.
- Décrire les évolutions possibles de cette CM et calculer leurs probabilités.
- Nous disposons des réalisations de v.a. uniformes sur  $[0, 1]$  et indépendantes, soient  $U_1(\omega) = 0.55$  et  $U_2(\omega) = 0.70$ . Donner la réalisation d'une trajectoire de la CM jusqu'au premier changement d'état.

**3.7.** (Médian 2010). Soit une chaîne de Markov  $X = (X_n, n \geq 0)$  d'espace d'état  $E = \{1, 2, 3, 4, 5\}$ , de matrice de transition

$$P = \begin{pmatrix} 1/2 & 1/3 & 1/6 & 0 & 0 \\ 3/8 & 1/4 & 1/4 & 0 & 1/8 \\ 0 & 0 & 1/2 & 1/3 & 1/6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

et de loi initiale  $\alpha = (1, 0, 0, 0, 0)$ .

- Tracer le graphe d'état de cette CM, donner les classes de communication en les caractérisant.
- Calculer la probabilité d'absorption de la CM dans l'état 4. Quelle est la probabilité d'absorption par l'état 5 ?
- Nous disposons de la réalisations d'une v.a. uniformes sur  $[0, 1]$ , soit  $U_1(\omega) = 0.55$ . Donner la réalisation du premier état visité après l'état 1.

**3.8.** Soit une chaîne de Markov irréductible récurrent positive et apériodique de loi stationnaire  $\pi$ . Montrer que si sa loi initiale est  $\pi$ , alors elle est strictement stationnaire, c'est-à-dire que la loi de la v.a.  $X_n$  est  $\pi$ .

**3.9.** Soit une chaîne de Markov  $X_n, n \geq 0$ , d'espace d'état  $E$  et de matrice de transition  $P$ . Soit  $0 = S_0 < S_1 < \dots < S_n < \dots$  les instants de changement d'états de la chaîne. Soit encore le processus  $Y_n, n \geq 0$ , défini par la formule

$$Y_n = X_{S_n}, \quad n \geq 0.$$

Montrer que  $Y_n, n \geq 0$ , est une chaîne de Markov et donner son espace d'états et sa matrice de transition en fonction de  $P$ .

## Chapitre 4

# Processus de Poisson

### 4.1 Introduction

Soient les instants d'arrivées de pannes d'un système ou d'un composant,  $0 < T_1 < T_2 < \dots$  sur  $\mathbb{R}_+$ . Considérons un processus de comptage  $N = (N_t; t \geq 0)$ , correspondant à ce phénomène, i.e.,  $N_t$  est le nombre d'arrivées dans l'intervalle de temps  $(0, t]$  avec  $N_0 = 0$ .

Nous définissons :

$$m(t) = \mathbb{E}[N_t] = \int_0^t \lambda(u) du \quad (4.1)$$

appelée fonction espérance, ou fonction moyenne, et

$$\lambda(t) = m'(t) \quad (4.2)$$

intensité du processus  $N$ .

### 4.2 Processus de Poisson homogène

**Définition 4.1.** [Processus de Poisson homogène (PPH)]

Un processus de comptage  $(N_t)$  est appelé processus de Poisson homogène (PPH), s'il vérifie les trois propriétés suivantes :

1. l'application  $t \mapsto N_t(\omega)$  est pour presque tout  $\omega$  une fonction en escaliers croissante et de sauts d'amplitude unité, i.e.,  $N_t - N_{t-} \in \{0, 1\}$  pour tout  $t \geq 0$ , avec  $N_0 = 0$ .
2.  $N$  est à accroissements indépendants, i.e., pour tout  $s < t$  dans  $\mathbb{R}_+$ , la v.a.  $N_t - N_s$  est indépendante de  $(N_u; u \leq s)$ .
3.  $N$  est à accroissements stationnaires, i.e., la loi de la v.a.  $N_t - N_s$  ne dépend que de  $t - s$ .

**Proposition 4.1.**

Pour un PPH, nous avons :

1. pour tout  $t \geq 0$  :  $\mathbb{P}(N_t = 0) = e^{-\lambda t}$  où  $\lambda \geq 0$  est une constante ;
2. pour  $t > 0$  et  $t \rightarrow 0$  :  $\mathbb{P}(N_t \geq 2) = o(t)$  ;
3. pour  $t > 0$  et  $t \rightarrow 0$  :  $\mathbb{P}(N_t = 1) = \lambda t + o(t)$ .

**Proposition 4.2.**

Les temps d'inter-arrivées  $X_n, n \geq 1$ , d'un PPH  $(\lambda)$ , sont des v.a. i.i.d. de loi exponentielle de paramètre  $\lambda$ .

**Proposition 4.3** (Loi d'un PPH).

Si  $N = (N_t; t \geq 0)$  est un PPH, alors pour tout  $t \geq 0$ , nous avons

$$\mathbb{P}(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, 2, \dots \quad (4.3)$$

Nous avons  $m(t) = \mathbb{E}(N_t) = \lambda t$  et  $\text{Var}(N_t) = \lambda t$ . L'intensité d'un PPH est  $m'(t) = \lambda$ . On note alors PPP( $\lambda$ ).

**Proposition 4.4.** Un processus de comptage  $N = (N_t, t \in \mathbb{R}_+)$ , est un processus de Poisson homogène, d'intensité  $\lambda$ , si et seulement si :

1. les sauts de  $N$  sont d'amplitude unité pour tout  $\omega$  ;
2. pour tous  $t, s \geq 0$ , on a

$$\mathbb{E}[N_{t+s} - N_s \mid N_u; u \leq s] = \lambda t.$$

Soit  $N = (N_t; t \geq 0)$  un PPH, alors pour un  $t \geq 0$ , nous définissons :

- age au temps  $t$  (temps de récurrence arrière) :  $U_t := t - S_{N(t)}$  ;
- temps résiduel (temps de récurrence après) :  $V_t := S_{N(t)+1} - t$  ;
- temps total :  $W_t := U_t + V_t$ .

**Proposition 4.5** (Temps de récurrence).

Les fonctions de répartition, marginales et jointes, des  $V_t$  et  $U_t$ , sont :

1.

$$\mathbb{P}(V_t \leq x) = 1 - e^{-\lambda x}, \quad x \geq 0.$$

2.

$$\mathbb{P}(U_t \leq x) = \begin{cases} 1 - e^{-\lambda x}, & 0 \leq x < t \\ 1, & t \leq x. \end{cases}$$

3.

$$\mathbb{P}(U_t \leq x, V_t \leq y) = \begin{cases} 1 - e^{-\lambda(x+y)}, & x > 0, 0 < y < t \\ 0, & t \leq y. \end{cases}$$

**Remarque 4.2.** Observer un PPH à partir de l'instant  $t$ , arbitraire mais fixé, revient à observer un PPH de même paramètre.

**Proposition 4.6** (Comportement asymptotique).

Soit  $N = (N_t; t \geq 0)$  un PPH, nous avons :

1.  $\frac{N_t}{t} \xrightarrow{p.s.} \lambda$ , lorsque  $t \rightarrow \infty$  ;
2.  $\sqrt{t}(\frac{N_t}{t} - \lambda) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \lambda)$ .

**Proposition 4.7** (Loi jointe conditionnelle de temps d'arrivées).

Soit  $N = (N_t; t \geq 0)$  un PPH et  $T_1, \dots, T_n$  les instants d'arrivées sachant que  $\{N_t = n\}$ . Alors la loi de  $(T_1, \dots, T_n)$  est la même que la loi des statistiques d'ordre croissante de  $n$  v.a. uniformes sur  $[0, t]$ .

### 4.3 Processus de Poisson non homogène

**Définition 4.2.** Un processus de comptage  $N = (N_t; t \geq 0)$  est appelé processus de Poisson non homogène (PPNH), s'il il vérifie les propriétés (1) et (2) de la définition 1.

C'est-à-dire qu'il n'est pas à accroissement stationnaires.

Soit  $m^{-1}(t)$  l'inverse généralisée de  $m(t)$  par rapport au temps définit par

$$m^{-1}(t) = \inf\{s : m(s) > t\} \quad (4.4)$$

**Proposition 4.8.** Soit  $N = (N_t; t \geq 0)$  un PPNH et supposons que  $m(t)$  soit continue, alors le processus  $M = (M_t; t \geq 0)$  défini par

$$M_t(\omega) = N_{m^{-1}(t)}(\omega) \quad (4.5)$$

est un PPH avec intensité  $\lambda = 1$ .

**Proposition 4.9** (Loi d'un PPNH).

Soit  $N = (N_t; t \geq 0)$  un PPNH avec  $m(t)$  continue, alors :

1.  $\mathbb{P}(N_{t+h} - N_t = n) = e^{-[m(t+h)-m(t)]} \frac{[m(t+h)-m(t)]^n}{n!}$  pour tout  $t, h \geq 0$  ;
2.  $\mathbb{P}(T_{n+1} - T_n > x \mid T_1, \dots, T_n) = e^{-[m(T_n+x)-m(T_n)]}$  pour tout  $x \geq 0$ .

**Proposition 4.10** (Comportement asymptotique).

Soit  $N = (N_t; t \geq 0)$  un PPNH. Si  $\int_0^\infty \lambda(u) du = \infty$ , nous avons :

1.  $\frac{N_t}{m(t)} \xrightarrow{p.s.} 1$ , lorsque  $t \rightarrow \infty$  ;
2.  $\frac{1}{\sqrt{m(t)}}(N_t - m(t)) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$ .

Ainsi, dans le cadre d'un PPNH où les arrivées,  $T_1, T_2, \dots$ , représenteraient les instant de pannes, la fiabilité, au temps  $t$ , s'écrit

$$R(t) = e^{-m(t)}.$$

▷ **Exemple 4.1** *Fiabilité d'un logiciel.* Une des applications importantes des processus de Poisson est celle de modélisation de la fiabilité des logiciels. Lorsqu'un nouveau logiciel est développé, alors un test est souvent mis en oeuvre pour éliminer les fautes dans le logiciel. Une manière de faire consiste à exécuter le logiciel sur des problèmes connus. Cela se fait sur un temps donné où on note toutes les erreurs manifestées. Lorsque le test s'achève, on essaye d'identifier les fautes spécifiques qui ont causé les erreurs. Le logiciel est alors modifié pour éliminer les fautes. Comme nous ne sommes pas sûrs d'avoir éliminé toutes les fautes du logiciel, un problème important se pose qui concerne l'estimation de la fiabilité du logiciel modifié.

Notons qu'une *faute* réside dans le logiciel et elle peut causer des *erreurs* lors de l'exécution du logiciel. Par exemple, une faute de programmation dans une instruction provoque une erreur à chaque fois que celle ci est exécutée. Un des nombreux modèles utilisés pour estimer la fiabilité d'un logiciel et celui de Goël et Okumoto (1979) basé sur le taux de défaillance.

*Modèle de Goël-Okumoto.*— Supposons maintenant que le logiciel initial contient un nombre inconnu de fautes, désigné par la v.a.  $N$ .

Notons  $N_t$  le processus de comptage des erreurs manifestées dans le temps  $[0, t]$ . Le modèle de Goël-Okumoto est basé sur l'hypothèse suivante :

(H) le nombre moyen d'erreurs détectées dans  $[t, t + \Delta]$  est "essentiellement proportionnel" au nombre moyen d'erreurs non détectées par la longueur  $\Delta$  de l'intervalle.

Nous pouvons écrire aussi que  $m(0) = 0$  et que  $\lim_{t \rightarrow \infty} m(t) = a$  où  $a > 0$  est le nombre moyen d'erreurs dans le logiciel à l'instant  $t = 0$ .

Ainsi, suivant l'hypothèse (H), nous pouvons écrire :

$$m(t + \Delta) - m(t) = b(a - m(t))\Delta + o(\Delta),$$

où  $b$  est une constante  $> 0$ , appelée *taux de détection d'erreur*. De cette égalité on obtient l'équation différentielle suivante

$$m'(t) = b(a - m(t)),$$



d'où

$$m(t) = a(1 - e^{-bt}). \quad (4.6)$$

L'intensité du processus s'écrit

$$\lambda(t) = abe^{-bt}.$$

Ainsi nous avons, sous l'hypothèse d'un PPNH, que la fiabilité au temps  $t \geq 0$ , s'écrit

$$R(t) = \mathbb{P}(N_t = 0) = \exp\{-a(1 - e^{-bt})\}. \quad (4.7)$$

◁

**Remarque 4.4.** Ce résultat est à discuter car ne satisfait pas au comportement générale de la fonction de fiabilité et on a également  $\mathbb{E}(T) = \int_0^\infty R(t)dt = +\infty$ . Cf. problème 4.4.

## 4.4 Processus de Poisson composé

Soient  $Y_1, Y_2, \dots$  de v.a. positives désignant les chocs qu'un composant subit au cours du temps. Les instants d'arrivées de ces chocs forment un processus de Poisson homogène  $N_t, t \geq 0$ , d'intensité  $\lambda$ . Nous supposons que  $N$  et  $Y$  sont indépendantes. Le niveau cumulatif des chocs au temps  $t > 0$  est

$$X_t = \sum_{i=1}^{N_t} Y_i. \quad (4.8)$$

Les effets des chocs sont cumulatifs et le composant peut résister jusqu'au niveau  $z > 0$ , après cette valeur de choc c'est la rupture du composant. Le point  $z$  est appelé *seuil de rupture*.

**Proposition 4.11.** *Le processus de Poisson composé est un processus à accroissements indépendants et stationnaires.*

**Proposition 4.12.**

1.  $\mathbb{E}X_t = \lambda t \mathbb{E}[Y_1]$
2.  $\text{Var } X_t = \lambda t \mathbb{E}[Y_1^2]$
3.  $\text{Cov}(X_s, X_t) = \lambda \mathbb{E}[Y_1^2] \min\{s, t\}$ .

**Proposition 4.13** (Fonction caractéristique).

Nous avons :

$$\varphi_t(u) := \mathbb{E}[e^{iuX_t}] = \exp \left[ -t\lambda \int_{\mathbf{R}} (1 - e^{iuy}) dF(y) \right].$$

▷ **Exemple 4.2** (Processus de risque de Cramér-Lundberg).

Soit le processus  $U$  défini par

$$U_t = u + ct - X_t, \quad (4.9)$$

où  $X_t$  est un processus de Poisson composé défini par la relation (4.8). Où  $u > 0$  représente les fonds propres,  $c > 0$  avec  $ct$  les primes,  $Y_i > 0$  le  $i$ -ème sinistre.

Probabilité de ruine en horizon infini :

$$\zeta(u) := \mathbb{P}(\inf_{t \geq 0} U_t < 0 \mid U_0 = u).$$

Probabilité de ruine en horizon fini (avant  $T$ ) :

$$\zeta(u, T) := \mathbb{P}(\inf_{0 \leq t \leq T} U_t < 0 \mid U_0 = u).$$

◁

## 4.5 Processus de Poisson marqué

Soit un PPH( $\lambda$ )  $N_t$  dont les temps d'arrivées sont  $T_n$ ,  $n \geq 1$ , et  $Y_n$ ,  $n \geq 1$ , une suite de v.a. réelles i.i.d. de fonction de répartition  $G$  et de densité  $g$ , indépendantes du processus de Poisson.

Le processus bidimensionnel  $(T_n, Y_n)$  est appelé *Processus de Poisson marqué*.

**Théorème 4.14.** *Un processus de Poisson marqué  $(T_n, Y_n)$ ,  $n \geq 1$  est un PPH deux-dimensionnel dans le plan  $(t, y)$  où le nombre moyen des points dans une region  $A$  de ce plan est donné par*

$$\mu(A) = \iint_A \lambda g(y) dy dt.$$

▷ **Exemple 4.3** Des impulsions électriques, d'amplitudes aléatoires  $X_i$  au moments aléatoires  $\tau_i$ , arrivent à un oscillographe, en engendrant un flux poissonien ; donc le nombre des impulsions  $N_t$  dans un intervalle  $[0, t]$  vérifie la relation

$$P(N_t = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N},$$

où  $\lambda > 0$  est l'intensité du processus de Poisson homogène.

Après la réception d'impulsion d'amplitude  $X_i$ , sur l'écran de l'oscillographe on peut visualiser la fonction

$$X_i \exp[-\alpha(t - \tau_i)_+] = \begin{cases} 0, & t < \tau_i \\ X_i \exp[-\alpha(t - \tau_i)], & t \geq \tau_i \end{cases}$$

où  $\alpha > 0$  est une constante donnée. Par conséquent, la valeur  $X_i$  suit une décroissance exponentielle jusqu'au moment de l'impulsion suivante (cf. figure 4.1).

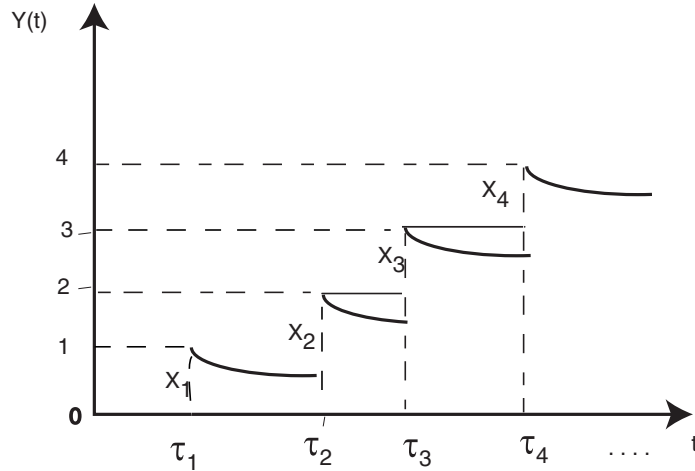


FIGURE 4.1 – Impulsions électriques

Si  $Y_t$  est la valeur enregistrée au moment  $t$ , alors

$$Y_t = \sum_{i=1}^{N_t} X_i \exp[-\alpha(t - \tau_i)_+].$$

En physique il est important de connaître la fonction de répartition de la v.a.  $Y_t$  ou, ce qui est équivalent, sa fonction caractéristique  $\varphi_t(s)$ .

On suppose que les v.a.  $X_i$  sont i.i.d. avec la densité de probabilité  $h(x)$ , donc avec la fonction caractéristique

$$\psi(s) = \int_0^\infty e^{isx} h(x) dx.$$

On peut démontrer que

$$\varphi_t(s) = \exp \left[ -\lambda \int_0^t (1 - \psi(se^{-av})) dv \right].$$

Nous pouvons calculer l'espérance et la variance de  $Y_t$  via sa fonction caractéristique :

$$\mathbb{E}[Y_t] = \lambda \mathbb{E}(X_1) \frac{1 - e^{-\alpha t}}{\alpha}$$

et

$$\text{Var}(Y_t) = \lambda \mathbb{E}(X_1^2) \frac{1 - e^{-2\alpha t}}{2\alpha}.$$

◁

▷ **Exemple 4.4** Soient  $(X_1(t), t \in \mathbb{R}_+)$  et  $(X_2(t), t \in \mathbb{R}_+)$  deux processus de Poisson indépendants de paramètres respectivement  $\lambda_1$  et  $\lambda_2$ . Soit  $X_t = X_1(t) - X_2(t)$ ,  $t \in \mathbb{R}_+$ .

- Le processus  $(X_t, t \in \mathbb{R}_+)$  est-il à accroissements indépendants ?
- Le processus  $(X_t, t \in \mathbb{R}_+)$  est-il un processus de Poisson ?
- Les valeurs  $\mathbb{P}(X_t - X_s = k)$  pour  $0 \leq s < t$  et  $k = 0, \pm 1, \pm 2$ .
- De la définition du processus  $X_t$  il en résulte

$$\begin{aligned} X(t+s) - X_s &= [X_1(t+s) - X_2(t+s)] - [X_1(s) - X_2(s)] = \\ &= [X_1(t+s) - X_1(s)] - [X_2(t+s) - X_2(s)], \end{aligned}$$

$$\begin{aligned} X_s - X_0 &= [X_1(s) - X_2(s)] - [X_1(0) - X_2(0)] = \\ &= [X_1(s) - X_1(0)] - [X_2(s) - X_2(0)]. \end{aligned}$$

Par suite de l'indépendance des quatre v.a. ci-dessus, il en résulte que  $X_{t+s} - X_s$  et  $X_s - X_0$  sont indépendantes également. Donc  $(X_t, t \in \mathbb{R}_+)$  est à accroissements indépendants.

b) La v.a.  $X_t$  peut prendre des valeurs entières, positives, négatives ou nulles ; donc  $(X_t, t \in \mathbb{R}_+)$  n'est pas un processus de Poisson.

c) On a

$$\begin{aligned} \mathbb{P}(X_t - X_s = k) &= \mathbb{P}([X_1(t) - X_1(s)] - [X_2(t) - X_2(s)] = k) = \\ &= \sum_{\nu=0 \vee k}^{\infty} \mathbb{P}(X_1(t) - X_1(s) = \nu) \mathbb{P}(X_2(t) - X_2(s) = \nu - k). \end{aligned}$$

Du fait que les processus  $X_1(t)$  et  $X_2(t)$  soient stationnaires on obtient

$$\begin{aligned} \mathbb{P}(X_i(t) - X_i(s) = k) &= \mathbb{P}(X_i(t-s) - X_i(0) = k) \\ &= \mathbb{P}(X_i(t-s) = k), \quad i = 1, 2. \end{aligned}$$

En posant  $t - s = u$ , il en résulte

$$\mathbb{P}(X_t - X_s = k) = \sum_{\nu=0 \vee k}^{\infty} \mathbb{P}(X_1(u) = \nu) \mathbb{P}(X_2(u) = \nu - k) = P(X(u) = k).$$

Donc  $(X_t, t \in \mathbb{R}_+)$  est un processus stationnaire et

$$\begin{aligned} \mathbb{P}(X_u = k) &= \sum_{\nu=0 \vee k}^{\infty} e^{-\lambda_1 u} \frac{(\lambda_1 u)^\nu}{\nu!} e^{-\lambda_2 u} \frac{(\lambda_2 u)^{\nu-k}}{(\nu-k)!} = \\ &= e^{-(\lambda_1 + \lambda_2)u} \sum_{\nu=0 \vee k}^{\infty} \frac{(\lambda_1 u)^\nu (\lambda_2 u)^{\nu-k}}{\nu! (\nu-k)!}, \end{aligned}$$

d'où

$$P(X_t - X_s = k) = e^{-(\lambda_1 + \lambda_2)(t-s)} \sum_{\nu=0 \vee k}^{\infty} \frac{(\lambda_1(t-s))^\nu (\lambda_2(t-s))^{\nu-k}}{\nu! (\nu-k)!}.$$

◁

## 4.6 Problèmes

### 4.1. Coût d'exploitation d'un composant

Un composant essentiel pour le fonctionnement en continu d'une machine est remplacé à chaque panne par un composant neuf et identique au précédent. Le temps de remplacement est considéré négligeable et la durée de vie du composant suit une loi exponentielle de paramètre  $\lambda$ .

1. Si le coût de chaque remplacement est égal à  $b$  Euros et nous admettons un taux de dévaluation de l'Euro constant et égal à  $a > 0$ . Calculer le coût moyen total des pannes.
2. Si le coût actualisé de chaque panne au temps  $t \geq 0$  et donné par la fonction  $t \mapsto g(t)$  positive sur  $\mathbb{R}_+$ , montrer que le coût moyen total des pannes est  $\lambda \int_0^\infty g(t) dt$ .
3. Retrouver le résultat 1. en utilisant 2.

### 4.2. Deux types de panne

Les instants de panne d'un système forment un processus de Poisson homogène de paramètre  $\lambda$  dont on note  $N(t)$  la fonction de comptage. Une panne donnée peut être de deux types : grave ou bénigne. Les types des différentes pannes sont indépendants les uns des autres, indépendants du processus  $N(t)$  et la probabilité pour qu'une panne donnée soit grave est  $p$  ( $0 < p < 1$ ). On note  $N_1(t)$  la fonction de comptage du processus ponctuel formé par les instants de panne grave.

1. Calculer la loi de  $N_1(t)$ .
2. Expliquer pourquoi  $N_1(t)$  est un processus de Poisson et donner son paramètre.

### 4.3. Processus de risque de Cramér-Lundberg

Le capital d'une compagnie d'assurance est décrit au cours du temps par un processus de risque de Cramér-Lundberg,  $U = (U_t, t \geq 0)$  :

$$U_t = u + ct - \sum_{k=1}^{N_t} Y_k, \quad t \geq 0,$$

où  $u > 0$ ,  $c > 0$ ,  $(N_t, t \geq 0)$  un processus de Poisson homogène de paramètre  $\lambda > 0$  et  $(Y_k, k \geq 1)$  une suite de v.a. positives i.i.d. de fonction de répartition commune  $F$ , indépendante du processus  $(N_t)$ .

1. Exprimer l'événement  $A_1$  = "la ruine de la compagnie ait lieu avant  $t$  lors du premier sinistre", avec  $t > 0$  fixé, en fonction du processus  $U$  et du temps  $T_1$  du premier sinistre.
2. Démontrer que

$$\mathbb{P}(A_1) = \lambda \int_0^t e^{-\lambda s} \bar{F}(u + cs) ds,$$

où  $\bar{F} = 1 - F$ .

3. Soit  $A_2$  l'événement : "la ruine de la compagnie ait lieu avant  $t$  lors du second sinistre". Calculer la probabilité  $\mathbb{P}(A_2)$ .

### 4.4. PPNH et le modèle de Goel-Okumoto

Dans le cadre du modèle de Goel-Okumoto calculer :

1. La fiabilité lorsque  $t \rightarrow \infty$  ? Discuter ce résultat.
2. La fiabilité sur intervalle, définie par  $R_t(x) := \mathbb{P}(N_{t+x} - N_t = 0)$ .
3. Le MTTF (temps moyen jusqu'à la défaillance). Discuter le résultat.
4. Le MTTF conditionnel au temps  $t \geq 0$ , noté  $MTTF_t^c$ , est défini par

$$MTTF_t^c := \mathbb{E}[T_{N_t+1} - t \mid T_{N_t+1} < \infty, N_s, s \leq t]$$

démontrer que

$$MTTF_t^c := \int_0^\infty \frac{e^{ae^{-b(t+x)}} - 1}{e^{ae^{-bt}} - 1} dx.$$

5. Que donnerait cette formule dans le cadre d'un Processus de Poisson non fini, y compris le PPH ? Discuter.

#### 4.5. Processus gamma

Un processus gamma,  $X$ , est un processus croissant, qui croit seulement par des sauts à des instant aléatoires dont le nombre dans des intervalles de temps ouverts est infini. Le nom de "processus gamma" est dû au fait que la densité de probabilité de  $X_t$  est une densité gamma,  $\gamma(bt, c)$ , i.e.,

$$ce^{-cx} \frac{(cx)^{bt-1}}{\Gamma(bt)}, \quad x > 0$$

où  $b$  est le paramètre de forme (shape) et  $c$  le paramètre d'échelle (scale).

1. Montrer que la transformée de Laplace de la loi  $\gamma(t, \alpha)$  est  $[\alpha/(\lambda + \alpha)]^t$ .
2. Montrer que la transformée de Laplace de  $X_t$  a la représentation suivante

$$\mathbb{E}[e^{-\lambda X_t}] := \exp\left[-t \int_{(0, \infty)} (1 - e^{-\lambda x}) \frac{be^{-cx}}{x} dx\right].$$

3. Calculer la moyenne et la variance de  $X_t$ .

#### 4.6. Oscillographe

Dans l'exemple de l'oscillographe, démontrer les formules de l'espérance et de la variance de  $Y_t$ .

# Chapitre 5

## Processus de Markov

### 5.1 Fonction de transition et propriété de Markov

Soit  $E$  un ensemble fini ou dénombrable, et  $P(t) = (p_{ij}(t); i, j \in E)$ ,  $t \in \mathbb{R}_+$ , une matrice de fonctions avec les propriétés

1.  $p_{ij}(t) \geq 0$ ,  $i, j \in E$ ,  $t \in \mathbb{R}_+$
2.  $\sum_{j \in E} p_{ij}(t) = 1$ ,  $i \in E$ ,  $t \in \mathbb{R}_+$
3.  $\lim_{t \rightarrow 0+} p_{ij}(t) = p_{ij}(0) = \delta_{ij}$ ,  $i, j \in E$

où  $\delta_{ij}$  est le symbole de Kronecker.

Une famille de v.a.  $(X_t, t \in \mathbb{R}_+)$  à valeurs dans  $E$  est appelée *processus de Markov homogène par rapport au temps* avec l'espace d'états  $E$ , si

$$\begin{aligned} \mathbb{P}(X_{t_n+t} = j \mid X_{t_n} = i, X_{t_k} = i_k, 0 \leq k \leq n-1) \\ = \mathbb{P}(X_{t_n+t} = j \mid X_{t_n} = i) = p_{ij}(t) \end{aligned} \quad (5.1)$$

pour tous  $0 \leq t_0 < t_1 < \dots < t_n$ ,  $t \geq 0$ ,  $i_0, i_1, \dots, i_{n-1}$ ,  $i, j \in E$ ,  $n \geq 1$ . L'égalité (5.1), appelée *la propriété de Markov*, peut être écrite symboliquement

$$\mathbb{P}(\text{FUTUR} \mid \text{PRÉSENT}, \text{PASSÉ}) = \mathbb{P}(\text{FUTUR} \mid \text{PRÉSENT}),$$

si l'on convient que l'instant  $t_n$  soit le présent (cf. (5.1)).

Nous allons supposer que les trajectoires  $t \mapsto X_t(\omega)$  sont continues à droite et limitées à gauche p.s. La matrice  $P(t)$ , appelée *la fonction (matricielle) de transition* du processus, vérifie *l'identité de Chapman-Kolmogorov*

$$4) \quad p_{ij}(s+t) = \sum_{k \in E} p_{ik}(s)p_{kj}(t), \quad i, j \in E, \quad s, t \geq 0 \quad (5.2)$$

qui dans une forme matricielle s'écrit

$$P(s+t) = P(s)P(t), \quad s, t \geq 0.$$

Ainsi pour qu'une fonction matricielle  $P(t)$  soit la fonction de transition d'un processus de Markov, il faut et il suffit qu'elle vérifie les propriétés 1)-4) ci-dessus.

Les fonctions  $p_{ij}(t)$  ont quelques propriétés remarquables :

**Proposition 5.1.** 1) Pour tous  $i, j \in E$ , la fonction  $p_{ij}(t)$  est uniformément continue sur  $[0, \infty)$ .

2) Pour tous  $i, j \in E$ , la fonction  $p_{ij}(t)$  est soit identiquement nulle, soit strictement positive, sur  $(0, \infty)$ .

## 5.2 Générateur et équations de Kolmogorov

Le générateur d'un processus de Markov,  $Q = (q_{ij}; i, j \in E)$  est défini par

$$q_{ij} := \lim_{t \rightarrow 0+} \frac{p_{ij}(t) - p_{ij}(0)}{t}. \quad (5.3)$$

Pour  $i \neq j \in E$ ,  $q_{ij} = p'_{ij}(0) = \lim_{t \rightarrow 0+} \frac{p_{ij}(t)}{t}$  existe et est fini.

Pour  $i \in E$ ,  $q_i = -p'_{ii}(0) = \lim_{t \rightarrow 0+} \frac{1 - p_{ii}(t)}{t}$  existe et  $\frac{1 - p_{ii}(t)}{t} \leq q_i$ . Un état  $i$  est dit *stable* si  $q_i < \infty$ , *instantané* si  $q_i = \infty$ , et *absorbant* si  $q_i = 0$ .

Si  $q_i < \infty$ , pour tout  $i \in E$ , alors la matrice  $Q = (q_{ij}, i, j \in E)$ , où  $q_{ii} = -q_i$ , est appelée *le générateur (infinitésimal) du processus* ou *la matrice d'intensités de transitions*. En général, on a  $q_{ij} \geq 0$  pour  $i \neq j$  et  $q_{ii} \leq 0$ . De plus

$$\sum_{j \neq i} q_{ij} \leq -q_{ii}. \quad (5.4)$$

Si  $E$  est fini, alors (5.4) devient

$$\sum_{j \neq i} q_{ij} = -q_{ii} < \infty, \quad i \in E. \quad (5.5)$$

**Proposition 5.2.** *Si  $E$  est fini, alors il n'y a pas d'état instantané.*

**Proposition 5.3.** *Les trajectoires du processus sont continues à droite p.s., si, et seulement si, il n'y a pas d'état instantané.*

La relation (5.5) est une condition nécessaire et suffisante pour que  $p_{ij}(t)$  vérifie les équations différentielle

$$p'_{ij}(t) = \sum_{k \in E} q_{ik} p_{kj}(t), \quad i, j \in E, \quad t \geq 0 \quad (5.6)$$

ou en forme matricielle

$$P'(t) = QP(t), \quad t \geq 0.$$

Les équations (5.6) sont appelées *les équations de Kolmogorov arrière (ou rétrograde)*.

En outre, si  $q_j < \infty$ ,  $j \in E$  et la limite

$$\lim_{h \rightarrow 0+} \frac{p_{ij}(h)}{h} = q_{ij}, \quad j \in E$$

est uniforme par rapport à  $i \neq j$ , alors on a

$$p'_{ij}(t) = \sum_{k \in E} p_{ik}(t) q_{kj}, \quad i, j \in E, \quad t \geq 0. \quad (5.7)$$

Les équations (5.7) sont appelées *les équations de Kolmogorov avant (ou progressive)*.

On note le fait important que, si  $E$  est fini, alors les deux équations (5.6) et (5.7) sont vérifiées. Dans ce cas, de ces deux systèmes d'équations, avec les conditions initiales  $p_{ij}(0) = \delta_{ij}$ ,  $i, j \in E$  (ou en forme matricielle  $P(0) = I$ ) on obtient

$$P(t) = \exp(Qt) = I + \sum_{n \in \mathbb{N}^*} \frac{Q^n t^n}{n!}, \quad t \in \mathbb{R}_+. \quad (5.8)$$

Donc, la matrice d'intensités de transitions  $Q$  détermine de manière unique la fonction matricielle de transition  $P(t)$ .

Le premier temps de saut ou le temps de séjour dans un état est défini par

$$T_1 = \sup(t \geq 0 : X_u = X_0 \ \forall u \in [0, t)).$$

Si on pose  $\mathbb{P}_i(\cdot) = \mathbb{P}(\cdot \mid X(0) = i)$ , alors  $\mathbb{P}_i(T_1 \leq t)$  est la fonction de répartition du premier temps de séjour dans l'état  $i$ .

**Proposition 5.4.** Pour  $i, j \in E$ ,  $T_1$  et  $X_{T_1}$  sont indépendantes et on a

$$\mathbb{P}_i(T_1 > t) = e^{-q_i t}, \ t > 0; \quad \mathbb{P}_i(X_{T_1} = j) = \frac{q_{ij}}{q_i} \quad (5.9)$$

On peut définir les temps de sauts du processus comme suit

$$\begin{aligned} T_2 &= \sup(t \geq T_1 : X_u = X_{T_1} \ \forall u \in [T_1, t)) \\ T_3 &= \sup(t \geq T_2 : X_u = X_{T_2} \ \forall u \in [T_2, t)) \\ &\dots \end{aligned}$$

La durée de vie du processus est la v.a.  $\zeta = \sup_n T_n$ . Si  $\mathbb{P}(\zeta = \infty) = 1$ , alors le processus est régulier (non explosif) ; dans le cas contraire, si  $\mathbb{P}(\zeta < \infty) > 0$ , il est non régulier (explosif).

Si  $T_n > 0$  p.s. pour tout  $n \geq 1$ , alors  $(X_t, t \geq 0)$  est un processus de sauts.

**Proposition 5.5.** Il existe une matrice stochastique  $(p_{ij}, i, j \in E)$  avec  $p_{ii} = 0$ ,  $i \in E$ , telle que

$$\mathbb{P}(X_{T_{n+1}} = j, T_{n+1} - T_n \leq t \mid X_{T_n} = i) = p_{ij}(1 - e^{-q_i t}), \ i, j \in E \quad (5.10)$$

C'est-à-dire que les états successifs visités forment une chaîne de Markov  $Y_n = X_{T_n}$ , appelée chaîne immergée (ou incluse), avec la fonction de transition  $P$ . Les temps de séjour dans les différents états, connaissant les états successifs visités par le processus, sont mutuellement indépendants.

**Remarque 5.1.** Nous avons

$$q_{ij} = -q_i \delta_{ij} + q_i a_{ij}; \quad Q = \text{diag}(q_i)(P - I). \quad (5.11)$$

**Proposition 5.6.** (Condition d'explosion de Reuter) Un processus de sauts Markovien est régulier ssi la seule solution non négative bornée de l'équation  $Qy = y$  est  $y = 0$ .

**Proposition 5.7.** Une condition nécessaire et suffisante pour qu'un processus de Markov  $(X_t, t \geq 0)$  soit régulier est  $\sum_{k=1}^{\infty} a_{Y_k}^{-1} = \infty$  p.s.

**Proposition 5.8.** (Équation intégrale de Kolmogorov) Pour  $i, j \in E$ , on a

$$p_{ij}(t) = \delta_{ij} e^{-q_i t} + \sum_{k \in E} p_{ik} \int_0^t q_i e^{-q_i s} p_{kj}(t-s) ds, \ t \geq 0 \quad (5.12)$$

pour  $i$  non absorbant, et  $p_{ij}(t) = \delta_{ij}$  pour  $i$  absorbant.

On note  $\eta_i$  le temps du premier passage dans l'état  $i$ , i.e.  $\eta_i = \inf(t \geq T_1 : X_t = i)$ . Un état  $i \in E$  est dit :

- récurrent si  $\mathbb{P}_i(\sup(t \geq 0 : X_t = i) = \infty) = 1$  ;
- transient si  $\mathbb{P}_i(\sup(t \geq 0 : X_t = i) < \infty) = 1$  ;



- *récurrent positif* si il est récurrent et  $\mu_{ii} = \mathbb{E}_i(\eta_i) < \infty$  ;
- *récurrent nul* si il est récurrent et  $\mu_{ii} = \infty$ .

Si  $\alpha = (\alpha_i, i \in E)$  est une probabilité (initiale) sur  $E$ , alors on pose

$$\mathbb{P}_\alpha(\cdot) = \sum_{i \in E} \alpha_i \mathbb{P}_i(\cdot).$$

Les probabilités  $p_i(t) = \mathbb{P}_\alpha(X_t = i)$ ,  $i \in E$ ,  $t \geq 0$ , sont appelées *probabilités d'états*. Si  $p_j = p_j(0) = \mathbb{P}_\alpha(X_0 = j)$ ,  $j \in E$ , alors les probabilités d'états vérifient les équations

$$p_j(t) = p_j + \sum_{k \in E} \int_0^t q_{kj} p_k(u) du, \quad j \in E. \quad (5.13)$$

### 5.3 Loi stationnaire et théorèmes ergodiques

Une loi  $\pi = (\pi_j, j \in E)$  sur  $E$  est appelée *stationnaire* ou *invariante* si  $\mathbb{P}_\pi(X_t = j) = \pi_j$ , pour tout  $j \in E$ ,  $t \geq 0$ . Cette condition peut être écrite dans une forme matricielle  $\pi P(t) = \pi$  pour tout  $t \geq 0$ . De (5.13) on peut en déduire que  $(\pi_j, j \in E)$  est stationnaire si, et seulement si,  $\sum_{k \in E} q_{kj} \pi_k = 0$ ,  $j \in E$ .

Un processus de Markov  $(X_t, t \geq 0)$  est dit *ergodique* s'il existe une loi  $\pi$  sur  $E$  telle que, pour tout  $i, j \in E$ ,

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_j$$

où de manière équivalente

$$\lim_{t \rightarrow \infty} \sum_{j \in E} |p_{ij}(t) - \pi_j| = 0$$

pour tout  $i \in E$ .

**Proposition 5.9.** *Pour tout état  $i$  dans  $E$ , nous avons*

$$\pi_i = \lim_{t \rightarrow \infty} p_{ii}(t) = \frac{1}{q_i \mu_{ii}},$$

où  $\mu_{ii} = \mathbb{E}_i[\eta_i]$  est le temps de récurrence moyen de l'état  $i$ .

**Proposition 5.10.** *Si le processus de Markov  $(X_t, t \geq 0)$  est irréductible (i.e. la chaîne  $(Y_n, n \in \mathbb{N})$  en est) et possède une probabilité invariante  $\pi$ , alors pour toute fonction  $g$  sur  $E$ , positive et bornée nous avons*

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t g(X(u)) du = \sum_{i \in E} \pi(i) g(i).$$

### 5.4 Quelques classes remarquables de processus de Markov

#### 5.4.1 Processus de Poisson

Nous allons présenter ici le processus de Poisson (vu dans le chapitre précédent) comme un cas particulier d'un processus de Markov. On considère un processus de Markov homogène  $(X_t, t \in \mathbb{R}_+)$ , avec espace d'états  $E = \mathbb{N}$  qui vérifie les conditions :

(1) Si, au moment  $t$ , le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité qu'au moment  $t + \Delta t$  soit dans l'état  $i + 1$  est  $\lambda \Delta t + o(\Delta t)$ , où  $\lambda > 0$  est une constante; donc  $p_{i,i+1}(\Delta t) = \lambda \Delta t + o(\Delta t)$ .

(2) Si, au moment  $t$ , le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité que dans l'intervalle du temps  $(t, t+\Delta t)$  il y ait plus qu'une transition ou qu'au moment  $t+\Delta t$  le processus soit dans un état  $j \neq i+1$ ,  $j \neq i$ , est  $o(\Delta t)$ ; donc  $p_{ij}(\Delta t) = o(\Delta t)$ ,  $j \neq i$ ,  $j \neq i+1$ .

(3) Si, au moment  $t$ , le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité qu'au moment  $t+\Delta t$  soit dans le même état  $i$  est  $1 - \lambda\Delta t + o(\Delta t)$ ; donc  $p_{ii}(\Delta t) = 1 - \lambda\Delta t + o(\Delta t)$ .

Il est facile à voir que le générateur  $Q$  est donné par

$$q_{ij} = \begin{cases} -\lambda, & j = i \\ \lambda, & j = i+1 \\ 0, & \text{ailleurs} \end{cases}$$

et de (5.6) et (5.7) il en résulte que :

$$\begin{aligned} p'_{ij}(t) &= -\lambda p_{ij}(t) + \lambda p_{i+1j}(t), \quad i, j \in \mathbb{N}, \\ p'_{ij}(t) &= -\lambda p_{ij}(t) + \lambda p_{ij-1}(t), \quad i \in \mathbb{N}, \quad j \in \mathbb{N}^*, \\ p'_{i0}(t) &= -\lambda p_{i0}(t), \quad i \in \mathbb{N}. \end{aligned}$$

La solution de ce système d'équations avec les conditions initiales  $p_{ij}(0) = \delta_{ij}$  est

$$p_{ij}(t) = \begin{cases} \frac{(\lambda t)^{j-i}}{(j-i)!} e^{-\lambda t}, & j \geq i \\ 0, & j < i \end{cases}$$

La relation (5.13) pour le processus de Poisson devient

$$\begin{aligned} p_j(t) &= p_j - \lambda \int_0^t [p_j(u) + p_{j-1}(u)] du, \quad j \geq 1, \\ p_0(t) &= -\lambda \int_0^t p_0(u) du \end{aligned}$$

et en dérivant

$$\begin{aligned} p'_j(t) &= -\lambda p_j(t) + \lambda p_{j-1}(t), \quad j \geq 1, \\ p'_0(t) &= -\lambda p_0(t). \end{aligned}$$

Etant donné

$$p_j(0) = \begin{cases} 1, & j = 0 \\ 0, & j = 1, 2, \dots \end{cases}$$

il en résulte  $p_0(t) = e^{-\lambda t}$ , puis  $p_1(t) = \lambda t e^{-\lambda t}$  et en général

$$p_j(t) = \frac{(\lambda t)^j}{j!} e^{-\lambda t}, \quad t \geq 0,$$

ce qui montre que  $X_t$  suit une loi de Poisson de paramètre  $\lambda t$ . On dit que  $(X_t, t \in \mathbb{R}_+)$  engendre un *flux poissonien*.

### 5.4.2 Processus de naissance

Soit  $(X_t, t \in \mathbb{R}_+)$  un processus de Markov avec l'espace d'états  $E = \mathbb{N}$ . Les conditions (1)-(3) du processus de Poisson sont vérifiées de processus de naissance avec cette différence que les probabilités de (1) et (3) seront  $\lambda_i \Delta t + o(\Delta t)$  et respectivement  $1 - \lambda_i \Delta t + o(\Delta t)$ , où  $\lambda_i$ ,  $i \in \mathbb{N}$ , est une suite de constantes positives.

Une transition de l'état  $i$  vers l'état  $i + 1$  correspond à une "naissance".  
Il est facile à voir que le générateur est

$$q_{ij} = \begin{cases} -\lambda_i, & j = i \\ \lambda_i, & j = i + 1 \\ 0, & \text{ailleurs.} \end{cases}$$

Par conséquent les équations (5.6) et (5.7) deviennent

$$\begin{aligned} p'_{ij}(t) &= -\lambda_i p_{ij}(t) + \lambda_i p_{i+1,j}(t), \quad i, j \in \mathbb{N}, \\ p'_{ij}(t) &= -\lambda_j p_{ij}(t) + \lambda_{j-1} p_{i,j-1}(t), \quad i \in \mathbb{N}, j \in \mathbb{N}^*, \\ p'_{i0}(t) &= -\lambda_0 p_{i0}(t), \quad i \in \mathbb{N}. \end{aligned} \tag{5.14}$$

Pour des constantes  $\lambda_i > 0$ ,  $i \in \mathbb{N}$ , distinctes, la solution de ce système avec les conditions initiales  $p_{ij}(0) = \delta_{ij}$ ,  $i, j \in \mathbb{N}$ , est

$$p_{ij}(t) = \begin{cases} \left( \prod_{k=i}^{j-1} \lambda_k \right) \sum_{k=i}^j \frac{e^{-\lambda_k t}}{\prod_{r=i, r \neq k} (\lambda_r - \lambda_k)}, & j > i \\ e^{-\lambda_i t}, & j = i \\ 0, & j < i. \end{cases}$$

Pour les probabilités d'états, de (5.13), on obtient les équations

$$\begin{aligned} p'_i(t) &= -\lambda_i p_i(t) + \lambda_{i-1} p_{i-1}(t), \quad i \in \mathbb{N}^*, \\ p'_0(t) &= -\lambda_0 p_0(t). \end{aligned} \tag{5.15}$$

La solution, avec les conditions initiales  $p_0(0) = 1$ ,  $p_i(0) = 0$ ,  $i \in \mathbb{N}^*$ , est

$$\begin{aligned} p_0(t) &= e^{-\lambda_0 t}, \\ p_i(t) &= \lambda_{i-1} e^{-\lambda_i t} \int_0^t e^{\lambda_i x} p_{i-1}(x) dx, \quad i \in \mathbb{N}^*. \end{aligned}$$

Si  $\sum_{k=0}^{\infty} p_k(t) = 1$  pour tout  $t \in \mathbb{R}_+$ , alors le processus de naissance est appelé *régulier*. Une condition nécessaire et suffisante pour que le processus soit régulier est

$$\sum_{k=0}^{\infty} \frac{1}{\lambda_k} = +\infty. \tag{5.16}$$

Un cas particulier important est quand  $\lambda_i = i\lambda$ ,  $i \in \mathbb{N}$ ,  $\lambda > 0$ . Dans ce cas le processus est appelé *processus linéaire de naissance* (ou *processus de Yule-Furry*). De (5.16) il en résulte que ce processus est régulier. Les relations (5.14) deviennent

$$\begin{aligned} p'_{ij}(t) &= -i\lambda p_{ij}(t) + i\lambda p_{i+1,j}(t), & i, j \in \mathbb{N}, \\ p'_{ij}(t) &= -j\lambda p_{ij}(t) + (j-1)\lambda p_{i,j-1}(t), & i \in \mathbb{N}, j \in \mathbb{N}^*, \\ p'_{i0}(t) &= 0, & i \in \mathbb{N}, \end{aligned}$$

avec la solution

$$p_{ij}(t) = \begin{cases} \binom{j-1}{j-i} e^{-i\lambda t} (1 - e^{-\lambda t})^{j-i}, & j \geq i \\ 0, & j < i. \end{cases}$$

De plus, les relations (5.15) s'écrivent

$$\begin{aligned} p'_i(t) &= -i\lambda p_i(t) + (i-1)\lambda p_{i-1}(t), \quad i \in \mathbb{N}^*, \\ p'_0(t) &= 0. \end{aligned}$$

Comme  $p_0(t) = \text{const.}$ , il y a deux possibilités :  $p_0(t) = 1$ ,  $t \in \mathbb{R}_+$ , donc le processus ne quitte jamais l'état 0, ou  $p_0(t) = 0$ ,  $t \in \mathbb{R}_+$ , donc l'espace d'états est au fond  $\mathbb{N}^*$ . Dans la deuxième hypothèse<sup>1</sup> on considère  $p_1(0) = 1$  et la solution du système est

$$p_j(t) = e^{-\lambda t}(1 - e^{-\lambda t})^{j-1}, \quad j \in \mathbb{N}^*.$$

**Remarque 5.2.** Il est à noter que le processus de Yule suit une loi géométrique de paramètre  $e^{-\lambda t}$ .

### 5.4.3 Processus de mort

Soient  $n_0 \in \mathbb{N}^*$  et  $E = \{0, 1, 2, \dots, n_0\}$ . On considère un processus de Markov homogène  $(X_t, t \in \mathbb{R}_+)$ , avec l'espace d'états  $E$  qui vérifie les conditions :

- (1)  $X(0) = n_0$ .
- (2) Si au moment  $t$  le processus se trouve dans l'état  $i$ ,  $i = 1, 2, \dots, n_0$ , alors la probabilité qu'au moment  $t + \Delta t$  il soit dans l'état  $i - 1$  est  $\mu_i \Delta t + o(\Delta t)$ , où  $\mu_1, \mu_2, \dots, \mu_{n_0}$  sont des constantes données ; donc  $p_{i,i-1}(\Delta t) = \mu_i \Delta t + o(\Delta t)$ .
- (3) Si au moment  $t$  le processus se trouve dans l'état  $i$ ,  $i = 1, 2, \dots, n_0$ , alors la probabilité qu'au moment  $t + \Delta t$  il soit dans le même état  $i$  est  $1 - \mu_i \Delta t + o(\Delta t)$  ; donc  $p_{ii}(\Delta t) = 1 - \mu_i \Delta t + o(\Delta t)$ .
- (4) La probabilité que dans l'intervalle  $(t, t + \Delta t)$  il y ait plus qu'une transition ou une transition différente des ceux décrites au (2) ou (3) est  $o(\Delta t)$  ; donc  $p_{ij}(\Delta t) = o(\Delta t)$ ,  $j \neq i - 1$ ,  $j \neq i$ .

De conditions (2) - (4) on déduit

$$q_{ij} = \begin{cases} -\mu_i, & j = i; \quad i = 1, 2, \dots, n_0 \\ \mu_i, & j = i - 1; \quad i = 1, 2, \dots, n_0 \\ 0, & \text{ailleurs.} \end{cases}$$

De (5.6) et (5.7) il en résulte

$$\begin{aligned} p'_{ij}(t) &= -\mu_i p_{i-1,j}(t) - \mu_i p_{ij}(t), \quad i = 1, 2, \dots, n_0, \quad j \in E, \\ p'_{0j}(t) &= 0, \quad j \in E \end{aligned} \tag{5.17}$$

et

$$\begin{aligned} p'_{ij}(t) &= -\mu_{j+1} p_{ij+1}(t) - \mu_j p_{ij}(t), \quad i, j \in E, \\ p'_{i0}(t) &= \mu_1 p_{i1}(t), \quad j \in E. \end{aligned} \tag{5.18}$$

On peut résoudre (5.17) et (5.18), avec les conditions initiales  $p_{ij}(0) = \delta_{ij}$ , en utilisant la transformation de Laplace.

De (5.13) on déduit que les probabilités d'états vérifient les équations

$$\begin{aligned} p'_i(t) &= -\mu_i p_i(t) + \mu_{i+1} p_{i+1}(t), \quad i = 0, 1, \dots, n_0 - 1, \\ p'_{n_0}(t) &= -\mu_{n_0} p_{n_0}(t), \end{aligned}$$

---

1. On note que (5.16) est vérifiée car  $\sum_{k=1}^{\infty} \frac{1}{k} = +\infty$ .

avec la condition initiale  $p_{n_0}(0) = 1$ . On obtient  $p_{n_0}(t) = e^{-\mu n_0 t}$  et puis les autres probabilités d'états par récurrence.

Pour le cas particulier  $\mu_i = i\mu$ ,  $\mu > 0$ ,  $i \in S$ , les probabilités d'états sont

$$p_i(t) = \binom{n_0}{i} e^{-n_0 \mu t} (e^{\mu t} - 1)^{n_0 - i}, \quad i \in S.$$

#### 5.4.4 Processus de naissance et de mort

Soit le processus de Markov homogène  $(X_t, t \in \mathbb{R}_+)$ , avec l'espace d'états  $E = \mathbb{N}$ , qui vérifie les conditions :

(1) Si au moment  $t$  le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité qu'au moment  $t + \Delta t$  il soit dans  $i + 1$  est  $\lambda_i \Delta t + o(\Delta t)$ , où  $\lambda_i$ ,  $i \in \mathbb{N}$ , sont des constantes strictement positives données ; donc  $p_{i,i+1}(\Delta t) = \lambda_i \Delta t + o(\Delta t)$ .

(2) Si au moment  $t$  le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité qu'au moment  $t + \Delta t$  il soit dans  $i - 1$  est  $\mu_i \Delta t + o(\Delta t)$ , où  $\mu_i$ ,  $i \in \mathbb{N}^*$ , sont des constantes strictement positives données ; donc  $p_{i,i-1}(\Delta t) = \mu_i \Delta t + o(\Delta t)$ .

(3) Si au moment  $t$  le processus se trouve dans l'état  $i \in \mathbb{N}$ , alors la probabilité qu'au moment  $t + \Delta t$  il soit dans le même état  $i$  est

$$1 - (\lambda_i + \mu_i) \Delta t + o(\Delta t), \quad i \in \mathbb{N}^*; \quad 1 - \lambda_0 \Delta t + o(\Delta t), \quad i = 0;$$

donc

$$p_{ii}(\Delta t) = 1 - (\lambda_i + \mu_i) \Delta t + o(\Delta t), \quad i \in \mathbb{N}^*; \quad p_{00}(\Delta t) = 1 - \lambda_0 \Delta t + o(\Delta t).$$

(4) La probabilité que dans l'intervalle de temps  $(t, t + \Delta t)$  il y ait plus qu'une transition ou des transitions différentes des ceux décrites au (1), (2) ou (3) est  $o(\Delta t)$  ; donc  $p_{ij}(\Delta t) = o(\Delta t)$  ;  $j \neq i - 1, j \neq i, j \neq i + 1$ .

Un tel processus peut décrire :

- a) L'évolution d'une population biologique : une transition de l'état  $i$  vers l'état  $i + 1$  correspond à une naissance et une transition de  $i$  à  $i - 1$  à une mort, d'où l'appellation du processus.
- b) L'évolution d'un ensemble de composants identiques dans un système en fonctionnement ou dans un stock. Les termes de naissance et de mort correspondent alors à une défaillance et à une réparation respectivement.
- c) Un système de file d'attente où les termes de naissance et de mort correspondent à l'arrivée et à la fin de service d'un client.

Le générateur est donné par

$$q_{ij} = \begin{cases} -(\lambda_i + \mu_i), & j = i \neq 0 \\ -\lambda_0, & j = i = 0 \\ \lambda_i, & j = i + 1, i \in \mathbb{N} \\ \mu_i, & j = i - 1, i \in \mathbb{N}^* \\ 0, & \text{ailleurs.} \end{cases}$$

De (5.6) et (5.7) il en résulte

$$\begin{aligned} p'_{ij}(t) &= \mu_i p_{i-1j}(t) - (\lambda_i + \mu_i) p_{ij}(t) + \lambda_i p_{i+1j}(t), \quad i \in \mathbb{N}^*, j \in \mathbb{N}, \\ p'_{0j}(t) &= -\lambda_0 p_{0j}(t) + \lambda_0 p_{1j}(t), \quad j \in \mathbb{N} \end{aligned}$$

et

$$\begin{aligned} p'_{ij}(t) &= \lambda_{j-1} p_{ij-1}(t) - (\lambda_j + \mu_j) p_{ij}(t) + \mu_{j+1} p_{ij+1}(t), \quad i \in \mathbb{N}, j \in \mathbb{N}^*, \\ p'_{i0}(t) &= -\lambda_0 p_{i0}(t) + \mu_1 p_{i1}(t), \quad i \in \mathbb{N}. \end{aligned}$$

On peut obtenir la solution en utilisant la transformation de Laplace. Conformément aux relations (5.13) les probabilités d'états vérifient

$$\begin{aligned} p'_i(t) &= \lambda_{i-1}p_{i-1}(t) - (\lambda_i + \mu_i)p_i(t) + \mu_{i+1}p_{i+1}(t), \quad i \in \mathbb{N}^*, \\ p'_0(t) &= \mu_1p_1(t). \end{aligned} \quad (5.19)$$

Si au moment  $t = 0$  le processus se trouve dans l'état  $k_0 \in \mathbb{N}$ , alors on doit considérer les conditions initiales  $p_i(0) = \delta_{ik_0}$ .

Si on pose

$$u_0 = 1, \quad u_k = \frac{\lambda_0\lambda_1 \cdots \lambda_{k-1}}{\mu_1\mu_2 \cdots \mu_k}, \quad k \in \mathbb{N}^*,$$

alors  $(u_i, i \in \mathbb{N})$  est une répartition stationnaire. De plus, si  $\sum_{i \in \mathbb{N}} u_i < \infty$  on a

$$\lim_{t \rightarrow \infty} p_{ij}(t) = \pi_i = \frac{u_j}{\sum_{k \in \mathbb{N}} u_k}.$$

Un cas particulier important de ce processus est le *processus linéaire de naissance et de mort* (le *processus de Feller-Arley*) pour lequel  $\lambda_i = i\lambda$ ,  $\mu_i = i\mu$ ,  $i \in \mathbb{N}$ ,  $\lambda, \mu > 0$ .

Si on considère la fonction génératrice (cf. §1.4)

$$g(s, t) = \sum_{k=0}^{\infty} p_k(t) s^k,$$

alors de (5.19) on déduit

$$\frac{\partial g(s, t)}{\partial t} = [\lambda s^2 - (\lambda + \mu)s + \mu] \frac{\partial g(s, t)}{\partial s}.$$

La solution générale de ce système est

$$g(s, t) = f\left(\frac{\mu - \lambda s}{1 - s} e^{-(\lambda - \mu)t}\right),$$

où  $f$  est une fonction quelconque de classe  $C^1$ . Si  $X(0) = 1$ , alors  $g(s, 0) = s$ , c'est-à-dire  $s = f\left(\frac{\mu - \lambda s}{1 - s}\right)$  et par conséquent  $f(x) = \frac{\mu - x}{\lambda - x}$ .

Donc

$$g(s, t) = \frac{\mu(1 - e^{(\lambda - \mu)t}) - (\lambda - \mu)e^{(\lambda - \mu)t}s}{\mu - \lambda e^{(\lambda - \mu)t} - \lambda(1 - e^{(\lambda - \mu)t})s}. \quad (5.20)$$

En développant la fonction (5.20) en série de puissance de  $s$ , on obtient, pour  $\lambda \neq \mu$ ,

$$\begin{aligned} p_k(t) &= [1 - \alpha(t)][1 - \beta(t)][\beta(t)]^{k-1}, \quad k \in \mathbb{N}^*, \\ p_0(t) &= \alpha(t), \end{aligned} \quad (5.21)$$

où

$$\alpha(t) = \frac{\mu(e^{(\lambda - \mu)t} - 1)}{\lambda e^{(\lambda - \mu)t} - \mu}, \quad \beta(t) = \frac{\lambda(e^{(\lambda - \mu)t} - 1)}{\lambda e^{(\lambda - \mu)t} - \mu}.$$

En utilisant aussi la fonction (5.20) on obtient

$$\begin{aligned} m(t) &= \mathbb{E}[X_t] = e^{(\lambda - \mu)t}, \\ \text{Var}[X_t] &= \frac{\lambda + \mu}{\lambda - \mu} e^{(\lambda - \mu)t} (e^{(\lambda - \mu)t} - 1), \quad \lambda \neq \mu. \end{aligned} \quad (5.22)$$

Un processus linéaire de naissance et de mort est appelé *sous critique, critique ou sur critique* selon que  $\lambda < \mu$ ,  $\lambda = \mu$  ou  $\lambda > \mu$ .

Pour les processus sous critiques ou sur critiques la variance est donnée par (5.22) ; dans le cas critique  $\text{Var}[X_t] = 2\lambda t$ . Un outre,

$$\lim_{t \rightarrow \infty} m(t) = \begin{cases} 0, & \lambda < \mu \\ 1, & \lambda = \mu \\ \infty, & \lambda > \mu. \end{cases}$$

Pour déterminer la probabilité que, pour  $t \rightarrow \infty$ , le processus entre dans  $i = 0$  (la population disparaisse), on utilise (5.21)

$$p_0(t) = \frac{\mu(e^{(\lambda-\mu)t} - 1)}{\lambda e^{(\lambda-\mu)t} - \mu},$$

d'où

$$\lim_{t \rightarrow \infty} p_0(t) = \begin{cases} 1, & \lambda < \mu \\ \frac{\mu}{\lambda}, & \lambda > \mu. \end{cases}$$

▷ **Exemple 5.1** Bien souvent, l'utilisation des processus de Markov homogène est le plus adéquat pour décrire le fonctionnement des systèmes complexes. Au cours du temps, le système passe d'un état à l'autre, à mesure que quelques composants tombent en panne et autres sont réparés. Donc, les défaillances et les réparations des composants sont des événements qui déterminent le passage du système d'un état à l'autre. Par exemple, si les composants d'un système fonctionnent en série, alors le système est défaillant si au moins un de composants tombe en panne. On suppose que seulement un composant peut tomber en panne à un moment  $t$  parce que les autres composants ne se détériorent pendant que le système ne fonctionne pas (les défaillances simultanées sont exclues). Si le nombre de composants est  $n$ , alors il existe  $n + 1$  états possibles dans l'évolution du système. Ainsi, ce phénomène physique peut être décrit par un processus de Markov  $(X_t, t \in \mathbb{R}_+)$  avec l'ensemble d'états  $E = \{0, 1, \dots, n\}$ . L'état 0 signifie que tous les composants fonctionnent ; l'état  $i$ ,  $1 \leq i \leq n$ , signifie la défaillance du  $i$ -ème composant. On suppose que la probabilité de défaillance de ce composant pendant un intervalle de longueur  $\Delta t$  est  $\lambda_i \Delta t + o(\Delta t)$ , et la probabilité de réparation pendant un intervalle de la même longueur est  $\mu_i \Delta t + o(\Delta t)$ , où  $\lambda_i, \mu_i > 0$ ,  $i = 1, 2, \dots, n$  ; la probabilité d'avoir deux ou plus défaillances pendant un intervalle de longueur  $\Delta t$  est  $o(\Delta t)$ .

Nous allons déterminer la fonction de répartition d'une période du fonctionnement continu ; ainsi nous allons considérer seulement l'évolution du système de l'état 0 vers les états  $i, i = 1, 2, \dots, n$ . Pour le modèle décrit ci-dessus le générateur est

$$\begin{aligned} q_{0i} &= \lambda_i, \quad q_{00} = -\lambda_0, \quad \lambda_0 = \sum_{i=1}^n \lambda_i, \\ q_{ij} &= 0, \quad i = 1, 2, \dots, n, \quad j = 0, 1, 2, \dots, n. \end{aligned}$$

En ce qui concerne les probabilités d'états, de (5.13), on a

$$\begin{aligned} p'_0(t) &= -\lambda_0 p_0(t), \quad p_0(0) = 1, \\ p'_i(t) &= \lambda_i p_0(t), \quad p_i(0) = 0, \quad i = 1, 2, \dots, n, \end{aligned}$$

d'où

$$p_0(t) = e^{-\lambda_0 t}, \quad p_i(t) = \frac{\lambda_i}{\lambda_0} (1 - e^{-\lambda_0 t}), \quad i = 1, 2, \dots, n.$$

Donc la fonction de répartition d'une période du fonctionnement continu est

$$F(t) = 1 - p_0(t) = 1 - e^{-\lambda_0 t}.$$

En outre, on en déduit que  $p_i(t) = \frac{\lambda_i}{\lambda_0} F(t)$ ; la valeur  $\frac{\lambda_i}{\lambda_0}$  signifie que la probabilité qu'une période du fonctionnement continu soit achevée par la défaillance du  $i$ -ème composant.

Pour obtenir la fonction de répartition d'une période de réparation, on doit considérer seulement l'évolution du système des états  $i$ ,  $i = 1, 2, \dots, n$ , vers l'état 0. Le générateur est  $q_{0j} = 0$ ,  $j = 0, 1, \dots, n$ ,

$$q_{i0} = \mu_i, \quad q_{ij} = -\mu_i \delta_{ij}, \quad i, j = 1, 2, \dots, n$$

et il conduit au système d'équations différentielles

$$p'_0(t) = \sum_{i=1}^n \mu_i p_i(t), \quad p_0(0) = 0,$$

$$p'_i(t) = -\mu_i p_i(t), \quad p_i(0) = \frac{\lambda_i}{\lambda_0}, \quad i = 1, 2, \dots, n,$$

avec la solution

$$p_0(t) = \sum_{i=1}^n \frac{\lambda_i}{\lambda_0} (1 - e^{-\mu_i t}), \quad p_i(t) = \frac{\lambda_i}{\lambda_0} e^{-\mu_i t}, \quad i = 1, 2, \dots, n.$$

Donc, la fonction de répartition d'une période de réparation est

$$G(t) = p_0(t) = \sum_{i=1}^n \frac{\lambda_i}{\lambda_0} (1 - e^{-\mu_i t}).$$

◁

▷ **Exemple 5.2** On considère une population de  $n+1$  individus, homogène, soumise au risque de contracter une maladie (épidémie). Soit  $X_t$  le nombre d'individus qui ne sont pas contaminés au moment  $t$ . On suppose  $X(0) = n$ , donc au moment initiale  $t = 0$  un seul individu est contaminé. Le processus  $(X_t, t \in \mathbb{R}_+)$  est un processus de mort avec

$$\mu_k = k\mu(n - k + 1), \quad k = 1, 2, \dots, n, \quad \mu > 0.$$

Dans ce cas, les probabilités d'états, pour  $\mu = 1$ , sont

$$p'_k(t) = (k+1)(n-k)p_{k+1}(t) - k(n-k+1)p_k(t), \quad k = 0, 1, \dots, n-1,$$

$$p'_n(t) = -np_n(t).$$

Le modèle ci-dessus peut être amélioré en considérant des individus qui ne sont soumis plus au risque d'être contaminés à cause d'isolation, de vaccination ou de décès.

◁

## 5.5 Simulation de Monte Carlo

Se basant sur les relations (5.9) on peut écrire un algorithme de simulation d'une trajectoire d'un processus de Markov dans  $[0, T]$ .

ALGORITHME :

0. Soit  $x_0$  une réalisation de la v.a.  $X_0 \sim \alpha$ . Posons  $i = 0$  et  $T_0 = 0$ .
1. Posons  $i = i + 1$ . Soit  $w_i$  une réalisation de  $W_i \sim \text{Exp}(q_i)$ ;  $T_i = T_{i-1} + w_i$ .
2. Si  $T_i \geq T$  alors fin.
3. Soit  $l$  une réalisation de la v.a.  $X_i \sim \frac{q_{x_{i-1},j}}{q_{x_{i-1}}}, j \neq x_{i-1}$ . Posons  $x_i = l$ .
4. Continue en 1.



## 5.6 Problèmes

**5.1.** Considérons la fonction matricielle suivante :

$$\mathbf{p}(t) = \begin{pmatrix} 0,8 + 0,2e^{-5t} & 0,2 - 0,2e^{-5t} \\ 0,6 - 0,6e^{-5t} & 0,4 + 0,6e^{-5t} \end{pmatrix}$$

- a) Montrer que  $\mathbf{p}(t)$  est une fonction de transition.
- b) Si  $(X_t, t \geq 0)$  est un processus de Markov avec l'espace d'états  $E = \{1, 2\}$  et fonctions de transition  $\mathbf{p}(t)$ , calculer pour  $t_1 = 1,5$ ;  $t_2 = 3$ ;  $t_3 = 10$  la probabilité  $\mathbb{P}_1(X_{t_1} = 2, X_{t_2} = 1, X_{t_3} = 1)$ .

**5.2.** Considérons un système dont les états successifs visités forment une chaîne de Markov  $Y$  avec espace d'états  $E = \{1, 2, 3\}$  et la matrice de transition

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 0,4 & 0 & 0,6 \\ 0 & 1 & 0 \end{pmatrix}$$

Les temps de séjour dans les différents états sont indépendants et suivent des lois exponentielles de paramètre  $\lambda_1, \lambda_2$  et  $\lambda_3$ . Si  $\lambda_1 = \lambda_2 = \lambda_3 = 1$ , alors

- a) Calculer le générateur du processus de Markov  $X$  décrivant le comportement stochastique du système sur la demi-droite réelle  $x \geq 0$ .
- b) Calculer par un développement limité  $\mathbf{p}(t)$  pour  $t = 1$ .

**5.3.** Supposons qu'un circuit électrique alimente  $m$  machines qui utilisent le courant de façon intermittente. Admettons que les machines fonctionnent indépendamment les unes des autres et que si une machine ne fonctionne pas à l'instant  $t$ , il y a une probabilité  $\lambda\Delta t + o(\Delta t)$ , ( $\lambda > 0$ ), qu'elle nécessite du courant dans l'intervalle  $(t, t + \Delta t)$ . Supposons que les périodes de travail soient des v.a. i.i.d. de fonction de répartition  $F(x) = 1 - e^{-\mu x}$ ,  $x \geq 0$ , ( $\mu \geq 0$ ). Soit  $X_t$  le nombre de machines en action à l'instant  $t$ . Montrer que  $(X_t, t \geq 0)$  est un processus de Markov et déterminer sa fonction de transition.

**5.4.** Un système comporte  $n$  composants dont les durées de vies sont i.i.d. de loi exponentielle de paramètre  $\lambda > 0$ . Lorsque un composant tombe en panne il est remplacé par deux composants identique montés en parallèle. Soit  $X_t$  le nombre de composants formant le système au temps  $t$ . Montrer que

- a)  $(X_t, t \geq 0)$  est un processus de Markov et donner son générateur.
- b)  $(X_t, t \geq 0)$  est un processus explosif.
- c) Calculer la loi de  $\zeta = \sup_n T_n$ .

**5.5.** Soit une file d'attente  $M/M/1/N$ , avec  $\lambda = 0.01$ ,  $\mu = 0.015$ , et  $N = 21$ .

On suppose que le système est stationnaire.

1. Simuler une trajectoire du processus  $X_t, t \geq 0$ , ( $X_t$  = nombre de clients dans le système au temps  $t$ ), sur  $[0, T]$ , avec  $T = 100$ .
2. Estimer, par Monte Carlo, la probabilité de perte d'un client au temps  $t$ , pour une précision souhaitée :  $\theta = 0.01$  et  $\varepsilon = 0.01$ .

## Chapitre 6

# Processus de renouvellement

### 6.1 Introduction

Le processus de renouvellement figure parmi les premiers processus utilisés pour des problèmes d'ingénieur. Ce processus n'est pas markovien.

Soit  $X_1, X_2, \dots$  une suite de v.a. réelles positives i.i.d., avec la fonction de répartition  $F$ ,  $F(0-) = 0$ ,  $F(0) < 1$ , définies sur un espace de probabilité  $(\Omega, \mathcal{F}, \mathbb{P})$ . La suite  $(S_n, n \in \mathbb{N})$ , où

$$S_0 = 0, S_n = X_1 + X_2 + \dots + X_n, n \geq 1 \quad (6.1)$$

est appelée *suite (processus) de renouvellement*.

Dans les applications qui ont conduit à la construction de la théorie de renouvellement, les v.a.  $S_1, S_2, \dots$  représentent les temps de remplacement (renouvellement) des composants détériorés et  $X_1, X_2, \dots$  sont les temps du fonctionnement (les temps entre les arrivées successives). L'espérance mathématique  $\mu = \mathbb{E}(X_k)$ , qui existent pouvant être éventuellement  $+\infty$ , est appelée *la durée moyenne de vie* des composants.

Le nombre de renouvellements dans un intervalle  $[0, t]$  est une v.a. définie par

$$N_t(\omega) = \max\{n : S_n(\omega) \leq t\} = \sum_{n=0}^{\infty} \mathbb{1}_{[0, t]}(S_n(\omega)), t \geq 0, \omega \in \Omega \quad (6.2)$$

et, pour  $n \geq 1$ . Le processus  $N_t, t \geq 0$ , est appelé *processus de comptage*.

Nous avons :

$$\begin{aligned} \mathbb{P}(N_t = n) &= \mathbb{P}(S_{n-1} \leq t, S_n > t) = \int_0^t \mathbb{P}(S_n > t \mid S_{n-1} = x) \mathbb{P}(S_{n-1} \in dx) \\ &= \int_0^t \mathbb{P}(X_n > t - x) \mathbb{P}(S_{n-1} \in dx) \\ &= \int_0^t [1 - F(t - x)] F^{(n-1)}(dx) = F^{(n-1)}(t) - F^{(n)}(t), \end{aligned}$$

où  $F^{(n)}$  (la  $n$ -ième convolution de Stieltjes de  $F$ ) est la fonction de répartition de la v.a.  $S_n$ . Plus précisément,

$$\begin{aligned} F^{(0)}(t) &= \mathbb{1}_{[0, \infty)}(t) = \begin{cases} 1, & t \geq 0 \\ 0, & t < 0 \end{cases}, \\ F^{(1)}(t) &= F(t), \\ F^{(n+1)}(t) &= (F^{(n)} * F)(t) = \int_0^t F^{(n)}(t - x) F(dx), n \geq 1. \end{aligned}$$

**Remarque 6.1.** 1. Dans le cas particulier où  $F(t) = 1 - e^{-\lambda t}$ ,  $t > 0$ , alors le processus de comptage  $(N_t - 1, t \geq 0)$  n'est rien d'autre qu'un processus de Poisson.

2. Nous considérons l'instant  $t = 0$  comme un instant de renouvellement, i.e.  $N_0 = 1$ .

L'espérance mathématique de la fonction de comptage au temps  $t$  est appelée *fonction de renouvellement* notée  $U(t)$ . Nous avons

$$\begin{aligned} U(t) &= \sum_{n=0}^{\infty} \mathbb{E} \mathbb{1}_{[0,t]}(S_n) = \sum_{n=1}^{\infty} \mathbb{P}(N_t \geq n) \\ &= \sum_{n=1}^{\infty} \mathbb{P}(S_{n-1} \leq t) = \sum_{n=0}^{\infty} F^{(n)}(t). \end{aligned} \quad (6.3)$$

Évidemment, le nombre moyen de renouvellement dans l'intervalle  $(s, t]$  est  $U(t) - U(s)$ .

## 6.2 Equation de renouvellement

On observe qu'il existe au moins un renouvellement dans l'intervalle  $(0, t]$  si, et seulement si,  $X_1 \leq t$ . Par conséquent le nombre moyen de renouvellements dans l'intervalle  $(0, t]$ , sachant que  $X_1 = y \leq t$ , est 1 plus le nombre moyen de renouvellements dans l'intervalle  $(y, t]$ . De (6.3) on voit que  $U(t)$  vérifie l'équation

$$U(t) = \mathbb{1}_{[0,\infty)}(t) + \int_0^{\infty} U(t-y)F(dy) \quad (6.4)$$

Cette équation est un cas particulier de ce que nous appelons une *équation de renouvellement* qui est de la forme

$$h = g + F * h \quad (6.5)$$

où  $F$  est une fonction de répartition sur  $\mathbb{R}_+$  et  $h$  et  $g$  sont des fonctions bornées sur des intervalles finies.

**Proposition 6.1.** (a)  $U(\cdot)$  est croissante et continue à droite,  $U(0) = 0$ ,  $U(t) < \infty$  pour tout  $t \in \mathbb{R}_+$  ;

(b) Soit  $h: \mathbb{R} \rightarrow \mathbb{R}$ ,  $h(t) = 0$  pour tout  $t < 0$ , tel que  $\int_0^t g(t-y)U(dy)$  existe et est finie pour tout  $t \in \mathbb{R}_+$ . Alors

$$h(t) = (U * g)(t) = \int_0^t g(t-y)U(dy)$$

est la solution unique de l'équation de renouvellement (6.5) dans la classe des fonctions  $h: \mathbb{R} \rightarrow \mathbb{R}$ ,  $h(t) = 0$  pour tout  $t < 0$  et bornées sur  $[0, t]$ ,  $t > 0$ .

▷ **Exemple 6.1** Soit  $F$  une fonction de répartition exponentielle  $\mathcal{E}(\lambda)$ . Alors,  $F(dy) = \lambda e^{-\lambda y} dy$ ,  $y \geq 0$ , et, pour  $n \geq 1$ , on peut démontrer par récurrence ou en utilisant la transformation de Laplace que

$$F^{(n)}(dy) = \lambda(\lambda y)^{n-1} \frac{e^{-\lambda y}}{(n-1)!} dy$$

d'où

$$\begin{aligned} \sum_{n=1}^{\infty} F^{(n)}(t) &= \sum_{n=1}^{\infty} \int_0^t \lambda(\lambda y)^{n-1} \frac{e^{-\lambda y}}{(n-1)!} dy \\ &= \int_0^t \sum_{n=1}^{\infty} \lambda(\lambda y)^{n-1} \frac{e^{-\lambda y}}{(n-1)!} dy \\ &= \int_0^t \lambda dy = \lambda t. \end{aligned}$$

Il en résulte

$$U(t) = \sum_{n=0}^{\infty} F^{(n)}(t) = 1 + \lambda t.$$

◁

### 6.3 Théorèmes limites

Tout au long de ce paragraphe, nous allons supposer que  $\mu = \mathbb{E}(X_1) < \infty$ .

**Proposition 6.2.** (Loi forte des grands nombres) *On a*

$$\frac{1}{t} N_t \xrightarrow{p.s.} \frac{1}{\mu}, \quad t \rightarrow \infty \quad (6.6)$$

**Proposition 6.3.** (Théorème de la limite centrale) *Si  $0 < \sigma^2 = \text{Var}(X_1) < \infty$ , alors :*

$$\frac{N_t - t/\sigma}{\sigma \sqrt{t/\mu^3}} \xrightarrow{\mathcal{L}} N(0, 1), \quad t \rightarrow \infty \quad (6.7)$$

**Proposition 6.4.** (Théorème élémentaire de renouvellement) *On a*

$$\lim_{t \rightarrow \infty} \frac{U(t)}{t} = \frac{1}{\mu} \quad (6.8)$$

**Proposition 6.5.** (Théorème de renouvellement de Blackwell) *Si  $F$  n'est pas en treillis,<sup>1</sup> alors pour tout  $h > 0$  on a*

$$\lim_{t \rightarrow \infty} [U(t+h) - U(t)] = \frac{h}{\mu} \quad (6.9)$$

*Si  $F$  est en treillis, alors (6.9) est valable lorsque  $h$  est un multiple de la période  $d$ .*

Afin de pouvoir énoncer le résultat suivant dans sa généralité, nous allons introduire les fonctions *direct Riemann intégrables* (DRI). Soit  $g$  une fonction définie sur  $\mathbb{R}_+$  et pour tout  $a > 0$  notons  $\overline{m}_n(a)$  le supremum et  $\underline{m}_n(a)$  l'infimum de  $g$  sur  $[(n-1)a, na]$ . Nous disons que  $g$  est DRI, si les sommes  $\sum_{n \geq 1} \overline{m}_n(a)$  et  $\sum_{n \geq 1} \underline{m}_n(a)$  sont finies pour tout  $a > 0$  et

$$\lim_{a \rightarrow 0} \underline{m}_n(a) = \lim_{a \rightarrow 0} \overline{m}_n(a).$$

**Proposition 6.6.** *Une condition suffisante pour qu'une fonction  $g$  soit DRI est que :*

- $g(t) \geq 0$  pour tout  $t \geq 0$  ;
- elle soit décroissante ;
- $\int_0^{\infty} g(t) dt < \infty$ .

**Proposition 6.7.** (Théorème clé de renouvellement) *Si  $F$  n'est pas en treillis, et  $g: \mathbb{R}_+ \rightarrow \mathbb{R}_+$  est direct Riemann intégrable, nous avons*

$$\lim_{t \rightarrow \infty} \int_0^t g(t-y) U(dy) = \frac{1}{\mu} \int_0^{\infty} g(y) dy. \quad (6.10)$$

*Dans le cas où  $F$  est en treillis avec période  $d$ , et  $\sum_{n \geq 0} g(x+nd) < \infty$ , alors*

$$\lim_{n \rightarrow \infty} \int_0^{x+nd} g(x+nd-y) U(dy) = \frac{d}{\mu} \sum_{n \geq 0} g(x+nd). \quad (6.11)$$

---

1. Une fonction  $F: \mathbb{R} \rightarrow \mathbb{R}$ ,  $F(t) = 0$  pour tout  $t < 0$ , est dite en treillis s'il existe une constante  $d > 0$  telle que  $F$  est constante sur toute l'intervalle ouvert de la forme  $(nd, (n+1)d)$  (treillis de période  $d$ ).

## 6.4 Processus de renouvellement modifié

Supposons  $S_0 > 0$ , indépendante des  $X_n$ ,  $\geq 1$ , avec la fonction de répartition  $F_0$  différente de  $F$ . Le processus  $(S_n, n \in \mathbb{N})$  sera appelé de *renouvellement modifié* ou à *délai*. Si  $U_0$  est la nouvelle fonction de renouvellement, nous pouvons écrire :

$$\begin{aligned} U_0(t) &= \mathbb{E}[N_t] = \sum_{n \geq 0} \mathbb{P}(S_n \leq t) = F_0(t) + \sum_{n=1}^{\infty} \mathbb{P}(S_n \leq t) \\ &= F_0(t) + \sum_{n=0}^{\infty} \mathbb{P}(S_{n+1} \leq t) = F_0(t) + \sum_{n=0}^{\infty} F^{(n+1)}(t) \\ &= F_0(t) + (U_0 * F)(t). \end{aligned}$$

Donc  $U_0$  vérifie l'équation de renouvellement

$$U_0(t) = F_0(t) + \int_0^t U_0(t-y)F(dy). \quad (6.12)$$

Pour  $h > 0$ , nous avons

$$\begin{aligned} U_0(t+h) - U_0(t) &= (F_0 * U)(t+h) - (F_0 * U)(t) \\ &= F_0 * [U(t+h) - U(t)] \\ &= \int_0^{t+h} [U(t+h-y) - U(t-y)]F_0(dy). \end{aligned} \quad (6.13)$$

Si  $F$  n'est pas en treillis, alors, d'après le théorème de renouvellement de Blackwell, nous avons

$$\lim_{t \rightarrow \infty} [U(t+h-y) - U(t-y)] = \frac{h}{\mu} \quad (6.14)$$

et par conséquent le nombre de renouvellement dans  $(t, t+h]$  tend vers la constante  $h/\mu$  lorsque  $t \rightarrow \infty$  et le taux de renouvellement est indépendant de la loi initiale  $F_0$ . On pose le problème de choisir la répartition  $F_0$  tel que  $U_0(t) = \frac{t}{\mu}$ , i.e. l'intensité de renouvellement soit constant. De (6.12) on voit que le seul choix possible est

$$F_0(t) = \frac{t}{\mu} - \frac{1}{\mu} \int_0^t (t-y)F(dy) = \frac{1}{\mu} \int_0^t [1 - F(y)]dy \quad (6.15)$$

Dans ce cas le processus est appelé *processus de renouvellement modifié stationnaire*

## 6.5 Temps d'attente

Considérons un processus de renouvellement simple  $(S_n, n \in \mathbb{N})$  et pour un temps fixe mais arbitraire  $t$ , soient les trois v.a. :  $X_t$ ,  $Y(t)$  et  $L(t)$  définies comme suit

1.  $X_t = t - S_{N_t-1}$  appelée *temps d'attente écoulé* (TAE) ou *age* ;
2.  $Y_t = S_{N_t} - t$  appelée *temps d'attente résiduel* (TAR) ou *durée de vie résiduelle* ;
3.  $L_t = S_{N_t} - S_{N_t-1}$  qui est la durée entre les deux renouvellements consécutifs contenant  $t$ .

**Proposition 6.8.** *Les lois des trois temps d'attente sont :*

$$1. \mathbb{P}(X_t \leq x) = \begin{cases} \int_{t-x}^t [1 - F(t-s)] U(ds), & 0 \leq x < t \\ 1, & x \geq t \end{cases};$$

$$\begin{aligned}
2. \quad \mathbb{P}(Y_t \leq y) &= \int_t^{t+y} [1 - F(t+y-s)] U(ds), y \geq 0; \\
3. \quad \mathbb{P}(L_t \leq x) &= \begin{cases} \int_{t-x}^t [F(x) - F(t-s)] U(ds), & 0 < x \leq t \\ F(x) - F(t) + \int_0^t [F(x) - F(t-s)] U(ds), & x > 0. \end{cases}
\end{aligned}$$

**Corollaire 6.9.** (Comportement asymptotique) On a :

$$\begin{aligned}
1. \quad \lim_{t \rightarrow \infty} \mathbb{P}(X_t \leq x) &= \frac{1}{\mu} \int_0^x [1 - F(s)] ds; \\
2. \quad \lim_{t \rightarrow \infty} \mathbb{P}(Y_t \leq y) &= \frac{1}{\mu} \int_0^y [1 - F(s)] ds; \\
3. \quad \lim_{t \rightarrow \infty} \mathbb{P}(L_t \leq x) &= \frac{1}{\mu} \int_0^x s F(ds).
\end{aligned}$$

▷ **Exemple 6.2** Si  $F$  est une loi exponentielle  $\mathcal{E}(\lambda)$ , on a :

$$\begin{aligned}
1. \quad \mathbb{P}(X_t \leq x) &= \begin{cases} 1 - e^{-\lambda x}, & 0 \leq x < t \\ 1, & x \geq t \end{cases}; \\
2. \quad \mathbb{P}(Y_t \leq y) &= 1 - e^{-\lambda y}, y \geq 0; \\
3. \quad \mathbb{P}(L_t \leq x) &= \begin{cases} 1 - (1 - \lambda x)e^{-\lambda x}, & 0 < x \leq t \\ 1 - (1 + \lambda x)e^{-\lambda x}, & x > t. \end{cases}
\end{aligned}$$

&lt;

▷ **Exemple 6.3** On suppose que les moments d'arrivées des trains de métro dans une station (pour une direction donnée) engendrent un flux poissonien de paramètre  $\lambda$ . Une personne qui arrive à l'arrêt au moment  $t$  peut faire les raisonnements suivants, par rapport au temps  $Y_t$ , jusqu'à l'arrivée du premier train :

a) Par le fait que la loi exponentielle "oublie le passé", le temps moyen d'attente, i.e.  $\mathbb{E}[Y_t]$ , doit être  $1/\lambda$ .

b) le moment où le passager arrive à l'arrêt est choisi au hasard entre deux arrivées consécutives et, pour raisons de symétrie, le temps moyen d'attente doit être  $1/2\lambda$ .

L'explication de ce paradoxe est que la v.a.  $L_t$  qui est l'intervalle entre deux arrivées successives où se trouve le moment  $t$  ne suit pas une loi exponentielle mais la loi donnée dans l'exemple 6.2. Pour cette répartition la valeur moyenne est

$$\mathbb{E}[L_t] = \frac{2}{\lambda} - \frac{1}{\lambda} e^{-\lambda t}$$

et non  $1/\lambda$  comme il a été supposé en a). D'un autre côté, la v.a.  $Y_t$  suit une loi exponentielle (cf. exemple 6.2) et nous avons  $\mathbb{E}[Y_t] = 1/\lambda \cong \mathbb{E}[L_t]/2$  pour  $t \rightarrow \infty$ .

&lt;

## 6.6 Processus de renouvellement arrêté

On suppose qu'à chaque instant de renouvellement il existe la possibilité que le processus s'arrête avec la probabilité  $q \in (0, 1)$  ou il continue avec la probabilité  $p = 1 - q$  indépendamment de l'évolution du processus jusqu'à cet instant-là. Soit  $(T_n, n \geq 1)$  une suite de v.a. i.i.d. avec

$$\mathbb{P}(T_n = 0) = q; \quad \mathbb{P}(T_n = 1) = 1 - q, \quad n \geq 1.$$

On suppose que les suites  $(T_n, n \geq 1)$  et  $(X_n, n \geq 1)$  sont indépendantes. La v.a.  $T_n$ ,  $n \geq 1$ , a la valeur 0 ou 1 selon qu'au  $n$ -ème renouvellement on décide l'arrêt ou la continuation du processus. Soit

$$G(t) = \mathbb{P}(X_n \leq t, T_n = 1) = \mathbb{P}(X_n \leq t)p = pF(t)$$

d'où  $G(\infty) = p$ .

Par conséquent, l'évolution du processus qui peut être arrêté est décrite par une suite de v.a. i.i.d. qui peuvent prendre la valeur  $\infty$  et avec la fonction de répartition  $G(\cdot)$ . Si  $X_n = \infty$ , alors le processus est arrêté au  $n$ -ème renouvellement. La fonction de répartition  $G$  est appelée *impropre* (ou *défectueuse* ou *sous-distribution*), i.e., elle possède un atome à l'infini.

La probabilité que le processus ne soit pas arrêté à l'un de premiers  $n$  renouvellements est  $p^n \rightarrow 0$  lorsque  $n \rightarrow \infty$ . Donc, le processus sera arrêté avec la probabilité 1, après un nombre fini de pas. Le nombre moyen des pas jusqu'à l'arrêt est

$$\sum_{n=1}^{\infty} np^{n-1} = \frac{1}{1-p}.$$

D'autre part, le nombre moyen de renouvellements sur  $[0, \infty)$  est  $U(\infty)$ , d'où il vient

$$U(\infty) = \frac{1}{1-p}.$$

La même chose peut être obtenu de l'équation (6.4) si on remplace  $F$  avec  $G$ .

De plus, la probabilité que  $S_n \leq t$  et que le processus soit arrêté au  $n$ -ème renouvellement est  $qG^{(n)}(t)$  d'où il en résulte que la probabilité que le processus soit arrêté avant  $t$  est

$$q \sum_{n=0}^{\infty} G^{(n)}(t) = qU(t).$$

Donc  $H(t) = qU(t)$  est la répartition du temps de vie du processus. Cette fonction vérifie l'équation de renouvellement

$$H(t) = 1 - G(\infty) + \int_0^t H(t-y)G(dy).$$

## 6.7 Processus de renouvellement alterné

Soient deux suites de v.a. positives et mutuellement indépendantes entre elles  $(X_n, n \in \mathbb{N})$  et  $(Y_n, n \in \mathbb{N})$  et identiquement distribuées des fonctions de répartitions  $F$  et  $G$  respectivement. On considère un système qui peut évoluer parmi deux états 0 et 1. Le système est au moment  $t = 0$  dans l'état 0 et y reste pour un temps aléatoire  $X_1$  au bout duquel il passe dans l'état 1 où y reste un temps aléatoire  $Y_1$ , puis passe de nouveau dans 0 pour un temps aléatoire  $X_2$  et ainsi de suite.

Soit maintenant le processus de renouvellement  $S$  avec

$$S_n = \sum_{i=1}^n (X_i + Y_i), \quad n \geq 1, \quad S_0 = 0 \quad (6.16)$$

avec les temps entre deux arrivées successives  $Z_n = X_n + Y_n$  de fonction de répartition  $H(t) = F * G(t)$ . Si  $H$  est direct Riemann intégrable, alors

$$\lim_{t \rightarrow \infty} P(t) = \frac{\mathbb{E}(X_1)}{\mathbb{E}(X_1) + \mathbb{E}(Y_1)} \quad (6.17)$$

où  $P(t)$  est la probabilité que à l'instant  $t$  le système soit dans l'état 0. En effet,

$$P(t) = \int_0^\infty \mathbb{P}(\text{système est dans 0 à l'instant } t \mid X_1 + Y_1 = x) H(dx).$$

Comme

$$\begin{aligned} & \mathbb{P}(\text{système est dans 0 à l'instant } t \mid X_1 + Y_1 = x) \\ &= \begin{cases} P(t-x), & \text{si } x \leq t \\ \mathbb{P}(X_1 > t \mid X_1 + Y_1 = x), & \text{si } x > t \end{cases} \end{aligned}$$

on a

$$\begin{aligned} P(t) &= \int_0^\infty P(t-x) H(dx) + \int_t^\infty \mathbb{P}(X_1 > t \mid X_1 + Y_1 = x) H(dx) \\ &= \int_0^\infty P(t-x) H(dx) + \int_0^\infty \mathbb{P}(X_1 > t \mid X_1 + Y_1 = x) H(dx), \end{aligned}$$

vue que  $Y_1 \geq 0$  ce qui entraîne  $\mathbb{P}(X_1 > t \mid X_1 + Y_1 < t) = 0$ . Ainsi, la fonction  $P(t)$  vérifie l'équation de renouvellement

$$P(t) = 1 - F(t) + \int_0^\infty P(t-x) H(dx)$$

avec la solution

$$P(t) = 1 - F(t) + \int_0^\infty [1 - F(t-x)] U(dx) \quad (6.18)$$

On applique la proposition 6.7 dans (6.18) et on obtient (6.17).

## 6.8 Problèmes

**6.1.** Les temps où les avions atterrissent sur un aéroport forment un processus de renouvellement avec les temps entre les arrivées successives suivant dont la f.r. est  $F$ . Chaque avion contient un nombre aléatoire de personnes avec une f.r. commune donnée de l'espérance mathématique finie. Trouver une expression pour le taux de personnes arrivées sur une longue période de temps. (*Indication* : Soit  $Z_i$  le nombre de personnes dans  $i$ -ème avion. On suppose que  $(Z_i)$  forment une suite de v.a. mutuellement indépendantes et indépendante des temps d'atterrissages. Le nombre

de personnes arrivées jusqu'au moment  $t$  est  $S(t) = \sum_{i=1}^{N_t} Z_i$  et on doit calculer  $\lim_{t \rightarrow \infty} \frac{S(t)}{t}$ .)

**6.2.** Supposons que le temps d'attente résiduel  $Y_t$  ne dépend pas de  $t$ . Montrer que  $(N_t, t \geq 0)$  est un processus de Poisson. (*Indication* : Soit  $G(y) = \mathbb{P}(Y_t > y)$ . Nous avons

$$G(y) = 1 - F(t+y) + g(y) \int_0^t F(dx)$$

On pose  $H(x) = 1 - F(x)$  et on montre que  $G(y) = H(t)H(t+y)$ ,  $y, t \geq 0$ . Pour  $t = 0$  on a  $G(y) = H(y)/H(0)$  d'où  $G(t+y) = G(t)G(y)$ ,  $y, t \geq 0$  et enfin,  $G(t) = e^{-\lambda t}$ , i.e.  $F(t) = 1 - e^{-\lambda t}$ .)

**6.3.** Montrer que la fonction de renouvellement d'un processus de renouvellement modifié est  $U_0(t) = \lambda t$ .



# Chapitre 7

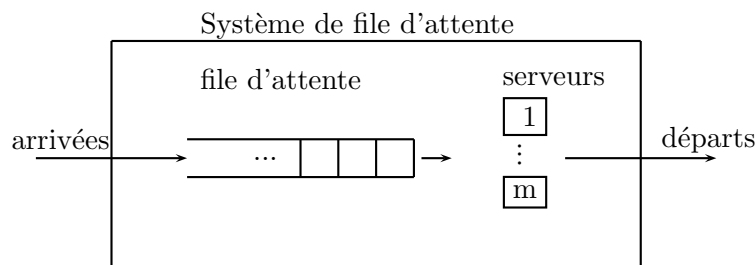
## Applications

### 7.1 Files d'attente

#### 7.1.1 Introduction

Les files d'attente ou systèmes d'attente ou théorie de queue est un domaine très important par ses applications dans des domaines très diverses. Nous pouvons observer des files d'attente dans notre vie quotidienne, par exemple, aux postes, aux bureaux de votes, aux restaurants universitaires, etc. Aujourd'hui on ne peut pas espérer gérer des réseaux téléphoniques sans recours à la théorie de files d'attente.

Le premier papier de recherche sur la modélisation des files d'attente a été publié en 1909, et concernait la gestion des réseaux téléphoniques, par le mathématicien Danois A. K. Erlang (1878-1929).



*Fig. 6.1. Système de file d'attente*

#### Paramètres de description :

1. Le processus d'arrivées
2. Le temps de service par serveur
3. Le nombre de serveurs
4. Le nombre des places disponibles dans le système
5. Le nombre d'individus dans la population étudiée
6. La discipline de service

#### Notation de Kendall

Pour désigner une file d'attente Kendall a proposé une notation compacte en regroupant les paramètres ci-dessus comme suit

$$1./2./3./4./5./$$

et en désignant les processus d'arrivées et de service par les lettres : M pour markovien, D pour déterministe, G pour général. Par exemple, le système de la figure 6.1, dans le cas d'un processus d'arrivé de Poisson et de temps de service exponentielles, est noté :  $M/M/m/\infty/\infty/$  ou simplement  $M/M/m$ .

**Notations :**

$c_j$  le  $j$ -ème client (par ordre d'arrivée),  $j = 1, 2, \dots$

$t_j$  le temps d'arrivée du  $j$ -ème client

$T_j$  le temps entre deux arrivées successives,  $T_j = t_j - t_{j-1}$

$S_j$  le temps de service du client  $c_j$

$D_j$  le temps d'attente dans la file du client  $c_j$

$W_j$  le temps d'attente dans le système du client  $c_j$ ;  $W_j = D_j + S_j$

$X_t$  le nombre de clients dans le système au temps  $t$

$N_t$  nombre d'arrivées dans l'intervalle de temps  $]0, t]$ ;  $N_t := \max\{j : t_j \leq t\}$

$M_t$  nombre de départs dans l'intervalle de temps  $]0, t]$

$I_j(t)$  fonction indicatrice de la présence du client  $c_j$  au temps  $t$  dans le système

On a :

$$I_j(t) := \begin{cases} 1 & \text{si } t_j \leq t \leq t_j + W_j \\ 0 & \text{sinon} \end{cases}$$

on peut écrire :

$$\int_0^\infty I_j(t) dt = W_j$$

et

$$X_t = \sum_{j=1}^{\infty} I_j(t).$$

**Les mesures de performance d'un système d'attente généralement répertoriées sont :**

- le nombre moyen  $L$  de clients dans le système ;
- le nombre moyen  $L_q$  de clients dans la file d'attente ;
- la durée d'attente moyenne  $W_q$  d'un client ;
- la durée de séjour moyenne  $W$  dans le système (attente + service) ;
- le taux d'occupation des postes de service ;
- le pourcentage de clients n'ayant pu être servis ;
- la durée d'une période d'activité, i.e. de l'intervalle de temps durant lequel il y a au moins un client en continue dans le système ;
- etc.

### 7.1.2 Formule de Little

La formule de Little est très intéressante et concerne tout système de file d'attente ergodique, y compris des systèmes déterministe. Considérons un système de file d'attente où les clients arrivent aux temps  $t_1, t_2, \dots, t_n, \dots$  et le nombre d'arrivées dans l'intervalle de temps  $[0, t]$  est  $N(t)$  (les fonctions ici ne sont pas nécessairement aléatoires). Notons  $W_j$  le temps total que le client numéro  $j$  passe dans le système. Nous allons supposer que les fonctions ci-dessus sont ergodiques, i.e., les limites suivantes existent :

(i)

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n W_i = \bar{W}$$

(ii)

$$\lim_{n \rightarrow \infty} \frac{1}{t} N(t) = \lambda$$

(iii)

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t X_u du = L.$$

Alors

$$L = \lambda \bar{W}$$

appelée *formule de Little*.

*Démonstration.* La démonstration de cette formule est obtenue directement de la représentation suivante

$$\frac{1}{t} \int_0^t X_u du = \frac{1}{t} \sum_{j=1}^{N(t)} W_j \wedge (t - t_j) = \frac{N(t)}{t} \frac{1}{N(t)} \sum_{j=1}^{N(t)} W_j \wedge (t - t_j), \quad (7.1)$$

et passage à la limite  $t \rightarrow \infty$ . □

Nous y reviendrons à cette formule pour le cas particulier de files d'attente dans ce qui suit.

### 7.1.3 File M/D/1 : Temps discret

Il s'agit d'une file comprenant un seul serveur dont le temps de service est déterministe. Les arrivées sont indépendantes et identiquement distribuées suivant une loi  $a = (a_i, i = 0, 1, 2, \dots)$  sur  $\mathbb{N}$ . Le nombre des places disponible dans la file est infini.

Notons  $\xi_n$  la v. a. désignant le nombre d'arrivées au temps  $n$ , nous avons :

$$P(\xi_n = k) = a_k \quad n \geq 0, k \geq 0.$$

L'état du système au début de chaque intervalle de temps est défini par le nombre de clients attendant dans le système pour être servis. Soit  $X_n$  le nombre de clients dans la file au début du temps  $n$ .

Nous avons :

$$X_{n+1} = \begin{cases} X_n - 1 + \xi_n & \text{si } X_n \geq 1 \\ \xi_n & \text{si } X_n = 0. \end{cases}$$

Par conséquent,  $X_n$  est une chaîne de Markov avec espace d'états  $E = \mathbb{N}$ . La fonction de transition est :

$$P = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Il est clair que si  $\mathbb{E}[\xi_n] > 1$ , alors la longueur de la file augmente indéfiniment. En revanche, si  $\mathbb{E}[\xi_n] < 1$ , la longueur de la file tend vers un état d'équilibre. Si  $\mathbb{E}[\xi_n] = 1$ , cela entraîne une situation de grande instabilité. Pour étudier la stabilité de ce système, nous avons besoin de deux théorèmes suivants .

**Théorème 7.1.** Soit  $X$  une chaîne de Markov irréductible dont l'espace d'états est  $E = \mathbb{N}$  et  $P$  sa fonction de transition. Une condition nécessaire et suffisante pour que  $X$  soit transitoire est alors que le système d'équations

$$\sum_{j \geq 0} P(i, j) y_j = y_i, \quad i \geq 1, \quad (7.2)$$

ait une solution bornée non réduite à une constante.

**Théorème 7.2.** Une condition nécessaire et suffisante pour qu'une chaîne de Markov irréductible soit récurrente est qu'il existe une suite  $(y_i, i = 1, 2, \dots)$  telle que :

$$\sum_{j \geq 0} P(i, j) y_j \leq y_i, \quad i \geq 1, \quad \text{et } y_i \rightarrow \infty.$$

Le système d'équations (7.2) a une solution bornée non réduite à une constante si

$$\sum_{k \geq 0} k a_k > 1.$$

Posons  $y_k = \xi^k$  et le système d'équations ci-dessus devient :

$$\sum_{j \geq 0} P(i, j) \xi^j = \sum_{j \geq i-1} a_{j-i+1} \xi^j = \xi^i,$$

ou

$$\sum_{j \geq i-1} a_{j-i+1} \xi^{j-i+1} = \xi = \sum_{k \geq 0} a_k \xi^k = \varphi(\xi), \quad i \geq 1.$$

Nous pouvons constater que  $\varphi(0) = a_0 > 0$ ,  $\varphi(1) = 1$  et  $\varphi'(1) = \sum_{k \geq 0} k a_k > 1$ , il existe  $\xi_0$ , un  $0 < \xi_0 < 1$ , tel que  $\varphi(\xi_0) = \xi_0$ .

Le vecteur  $\xi_0^j$ ,  $j \geq 0$ , est la solution bornée non constante cherchée, et donc, d'après le théorème 1, le processus est transitoire.

Lorsque  $\sum_{k \geq 0} k a_k > 1$ , la chaîne de Markov ci-dessus est récurrente. Posons  $y_j = j$ , nous avons

$$\sum_{j \geq 0} P(i, j) j = \sum_{j \geq i-1} j a_{j-i+1} = \sum_{j \geq i-1} j a_{j-i+1} (j - i + 1) - i + 1 = \sum_{k \geq 0} k a_k - 1 + i \leq i$$

et donc, d'après le théorème 2, la chaîne est récurrente.

#### 7.1.4 File M/M/1

C'est un système de file d'attente dont l'entrée est un processus de Poisson homogène d'intensité  $\lambda$ , les temps de service sont i.i.d. de loi exponentielle de paramètre  $\mu$  et il y a un seul réparateur. Dans ce cas, le processus  $X_t$ ,  $t \geq 0$ , désignant le nombre de clients dans le système, est un processus de Markov de naissance et mort, avec  $E = \mathbb{N}$ , dont le générateur est décrit par

$$\lambda_n = \lambda, \quad n \geq 0, \quad \text{et}$$

$$\mu_n = \mu, \quad n \geq 1.$$

Calcul de  $\gamma_j$ ,  $j \geq 0$  :

$$\gamma_j = \frac{\lambda_0 \cdot \lambda_1 \cdots \lambda_{j-1}}{\mu_1 \cdot \mu_2 \cdots \mu_j}, \quad j \geq 1,$$

et  $\gamma_0 = 0$ .

Donc, nous obtenons,

$$\gamma_j = \left(\frac{\lambda}{\mu}\right)^j, \quad j \geq 0.$$

Par suite, si  $\sum_{j \geq 0} \gamma_j < \infty$ , ce qui revient à supposer  $\lambda/\mu < 1$ , alors la probabilité stationnaire du processus  $X$  est  $\pi_j = \gamma_j / \sum_{k \geq 0} \gamma_k$ , ce qui donne

$$\pi_j = a^j (1 - a), \quad j \geq 0,$$

avec  $a = \lambda/\mu$  appelé *trafic offert*. Cette loi est la loi géométrique, sur  $\mathbb{N}$ , de paramètre  $1 - a$ . Il s'interprète comme le nombre moyen d'arrivées dans le système pendant la durée moyenne de service d'un client. La condition  $a < 1$  est donc naturelle pour que le système ne se sature pas.

De la loi géométrique nous obtenons directement :  $\mathbb{E}X_t = a/(1 - a)$  et  $\text{Var}(X_t) = a/(1 - a)^2$ .

**Proposition 7.3.** *Le temps  $W = W_j$ , que le client  $c_j$  passe dans le système est une v.a. exponentielle de paramètre  $(1 - a)\mu$ .*

*Démonstration.* Nous avons

$$\begin{aligned} \mathbb{P}(W > t) &= \sum_{n \geq 0} \mathbb{P}(W > t \mid X_{t_j-} = n) \mathbb{P}(X_{t_j-} = n) \\ &= \sum_{n \geq 0} \mathbb{P}(M_t \leq n) \pi_n \\ &= \sum_{n \geq 0} \left( \sum_{k=0}^n e^{-\mu t} \frac{(\mu t)^k}{k!} \right) a^n (1 - a) \\ &= e^{-\mu t} (1 - a) \sum_{k \geq 0} \left( \sum_{n \geq k} a^n \right) \frac{(\mu t)^k}{k!} \\ &= e^{-(1-a)\mu t}. \end{aligned}$$

De cette proposition, nous obtenons que

$$\mathbb{E}W = \frac{1}{(1 - a)\mu} = \frac{1}{\lambda} \frac{a}{(1 - a)} = \frac{\mathbb{E}X_t}{\lambda}$$

□

C'est-à-dire

$$\mathbb{E}X_t = \lambda \mathbb{E}W,$$

la formule de Little.

### 7.1.5 File M/M/ $\infty$

Le nombre des serveurs dans le système est infini. Par conséquent il n'y a pas d'attente, un client qui arrive est immédiatement pris en service.

Nous avons alors :

$$\lambda_n = \lambda, \quad n \geq 0,$$

$$\mu_n = n\mu, \quad n \geq 1.$$

Nous avons :

$$\gamma_j = \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j = \frac{a^j}{j!}, \quad j \geq 0,$$

et donc  $\sum_{j \geq 0} \gamma_j = e^a < \infty$ .

La probabilité stationnaire est :

$$\pi_n = e^{-a} \frac{a^n}{n!}$$

c'est la loi de Poisson de paramètre  $a$ .

En conséquence  $\mathbb{E}X_t = \text{Var}(X_t) = a$ .

### 7.1.6 File M/M/m/m

Il y a  $m$  place et  $m$  serveurs dans le système. Donc il n'y a pas de possibilité d'attente. Un client qui arrive, soit il trouve un serveur libre et il commence à être servi immédiatement, soit tous les serveurs sont occupés et le client en question quitte le système. L'espace d'états du système est donc fini :  $E = \{0, 1, \dots, m\}$

Nous avons donc :

$$\lambda_n = \begin{cases} \lambda & \text{si } n < m \\ 0 & \text{si } n \geq m, \end{cases}$$

et  $\mu_n = n\mu$ ,  $n = 1, \dots, m$

Nous obtenons :

$$\gamma_j = \frac{1}{j!} \left( \frac{\lambda}{\mu} \right)^j = \frac{a^j}{j!}, \quad j \leq m,$$

Et la loi stationnaire

$$\pi_n = \frac{a^n/n!}{\sum_{j=0}^m a^j/j!}, \quad 0 \leq n \leq m.$$

### 7.1.7 File M/M/m/ $\infty$

La file est infinie et le nombre des serveurs est  $m$ . Nous avons déjà étudié cette file pour  $m = 1$ .

Nous avons donc :  $\lambda_n = \lambda$ ,  $n \geq 0$ , et

$$\mu_n = \begin{cases} n\mu & \text{si } 0 \leq n \leq m \\ m\mu & \text{si } n \geq m. \end{cases}$$

Nous avons :

$$\gamma_j = \begin{cases} \frac{1}{j!} \left( \frac{\lambda}{\mu} \right)^j & \text{si } 0 \leq j \leq m \\ \frac{1}{m!} \left( \frac{\lambda}{\mu} \right)^m \left( \frac{\lambda}{m\mu} \right)^{j-m} & \text{si } j \geq m. \end{cases}$$

D'où

$$\sum_{j \geq 0} \gamma_j = \sum_{j=0}^{m-1} \frac{a^j}{j!} + \frac{a^m}{m!(1-\rho)},$$

où  $\rho = a/m$  et  $a = \lambda/\mu$ . La condition de stabilité est :  $\rho < 1$ .

Et donc, la loi stationnaire est :

$$\pi_0 = \left\{ \sum_{j=0}^{m-1} \frac{a^j}{j!} + \frac{a^m}{m!(1-\rho)} \right\}^{-1},$$

et

$$\pi_n = \begin{cases} \frac{a^n \pi_0}{n!} & \text{si } 1 \leq n \leq m \\ \frac{a^n \pi_0}{m! m^{n-m}} & \text{si } n \geq m. \end{cases}$$

**Proposition 7.4.** La loi du temps d'attente dans la file  $D = D_j$  a la f.r. suivante :

$$\mathbb{P}(D \leq t) = 1 - \frac{a^m \pi_0}{m!(1-\rho)} e^{-m\mu(1-\rho)t}.$$

Et finalement,

$$\mathbb{E}[D] = \frac{\pi_m}{m\mu(1-\rho)^2}.$$

### 7.1.8 File M/M/m/K

Ce système est fini,  $K$  places dans le système, avec  $m$  serveurs ( $m \leq K$ ). Il est donc stable. Nous avons  $\lambda_n = \lambda$  pour  $0 \leq n \leq K-1$  et

$$\mu_n = \begin{cases} n\mu & \text{si } 0 \leq n \leq m \\ m\mu & \text{si } m \leq n \leq K. \end{cases}$$

Nous avons :

$$\gamma_j = \begin{cases} \frac{1}{j!} \left(\frac{\lambda}{\mu}\right)^j & \text{si } 0 \leq j \leq m \\ \frac{1}{m!} \left(\frac{\lambda}{\mu}\right)^m \left(\frac{\lambda}{m\mu}\right)^{j-m} & \text{si } m \leq j \leq K. \end{cases}$$

D'où

$$\sum_{j=0}^K \gamma_j = \sum_{j=0}^{m-1} \frac{a^j}{j!} + \frac{a^m}{m!} \sum_{j=m}^K \rho^{j-m},$$

où  $\rho = a/m$  et  $a = \lambda/\mu$ .

Et la loi stationnaire est :

$$\pi_0 = \left\{ \sum_{j=0}^{m-1} \frac{a^j}{j!} + \frac{a^m}{m!} \sum_{j=0}^{K-m} \rho^j \right\}^{-1},$$

et

$$\pi_n = \begin{cases} \frac{a^n \pi_0}{n!} & \text{si } 1 \leq n \leq m \\ \frac{\pi_0 a^m \rho^{n-m}}{m! m^{n-m}} & \text{si } m \leq n \leq K. \end{cases}$$

Temps d'attente dans la file :  $D$  d'un client quelconque. Notons  $F_D$  sa f.r. Alors nous avons :

**Proposition 7.5.** *La loi du temps d'attente dans la file  $D = D_j$  a la f.r. suivante :*

$$F_D(t) := \mathbb{P}(D \leq t) = 1 - \frac{1}{1 - \pi_K} \sum_{n=m}^{K-1} \pi_n \sum_{i=0}^{n-m} \frac{(m\mu t)^i}{i!} e^{-m\mu t}.$$

et

$$\mathbb{E}[D] = \frac{1}{m\mu(1 - \pi_K)} \sum_{n=m}^{K-1} (n - m + 1) \pi_n.$$

## 7.2 Fiabilité

### 7.2.1 Introduction

Nous allons formuler la fiabilité et la disponibilité d'un système markovien. Si un système (ou un composant) est mis en fonctionnement à l'instant  $t = 0$ , sa fiabilité au temps  $t \geq 0$ , notée  $R(t)$ , est la probabilité que le système ne tombe pas en panne avant  $t$ . La disponibilité (instantanée), notée  $A(t)$ , est la probabilité que le système soit en bon état de marche au temps  $t$ , indépendamment s'il est tombé ou non en panne avant  $t$ .

Les définitions classiques, ci-dessus, de mesures de fiabilité et de disponibilité s'expriment dans le cadre d'un modèle markovien comme suit :

1. Fiabilité :  $R(t) = \mathbb{P}(\forall u \in [0, t], X_u \in U), t \geq 0$

2. Disponibilité instantanée :  $A(t) = \mathbb{P}(X_t \in U), t \geq 0$
3. Maintenabilité :  $M(t) = 1 - \mathbb{P}(\forall u \in [0, t], X_u \in D), t \geq 0$

Soit maintenant un processus de Markov,  $X_t$ , avec espace d'états  $E$ , générateur  $A$ , fonction de transition  $p_t$ , et loi initiale  $\alpha$ ; nous allons voir dans les paragraphes suivants comment les mesures ci-dessus peuvent être calculées. Nous avons besoin de considérer des partitions de matrices et des vecteurs suivant  $U$  et  $D$ . A ce fin, nous rangeons d'abord les états de marche,  $U = \{1, \dots, r\}$ , et puis les états de panne,  $D = \{r+1, \dots, d\}$ .

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

$\alpha = (\alpha_1, \alpha_2)$ . Notons également  $\mathbf{1}_{d,r} = (1, \dots, 1, 0, \dots, 0)^\top$  un vecteur  $d$ -dimensionnel dont les  $r$  premiers éléments sont des 1 et les autres des 0.

### 7.2.2 Disponibilité

Il y a plusieurs disponibilités : disponibilité instantanée, disponibilité limite, disponibilité moyenne sur un intervalle de temps, disponibilité moyenne limite, etc. Nous donnons ici les deux premières qui sont les plus usitées.

**Proposition 7.6.** *La disponibilité instantanée dans un système Markovien décrit ci-dessus est donnée par*

$$A(t) = \alpha e^{tA} \mathbf{1}_{d,r} \quad (7.3)$$

*Démonstration.*

$$\begin{aligned} A(t) &= \mathbb{P}(X_t \in U) = \sum_{j \in U} \mathbb{P}(X_t = j) = \sum_{j \in U} \sum_{i \in E} \mathbb{P}(X_t = j, X_0 = i) \\ &= \sum_{j \in U} \sum_{i \in E} \mathbb{P}(X_t = j \mid X_0 = i) \mathbb{P}(X_0 = i) = \sum_{j \in U} \sum_{i \in E} \alpha(i) P_t(i, j) \\ &= \alpha P_t \mathbf{1}_{d,r}. \end{aligned}$$

La disponibilité stationnaire, notée  $A_\infty$ , est donnée par

$$A_\infty = \sum_{k \in U} \pi(k) = \pi \cdot \mathbf{1}_{d,r} \quad (7.4)$$

□

### 7.2.3 Fiabilité

On considère le processus  $Y$  avec espace d'états  $U \cup \{\Delta\}$ , où  $\Delta$  est un état absorbant. Définissons  $T_D$ , le temps d'entrée dans les états de panne  $D$  :

$$T_D = \inf\{t \geq 0 : X_t \in D\}$$

avec la convention  $\inf \emptyset = +\infty$ , et  $Y$  par

$$Y_t(\omega) = \begin{cases} X_t(\omega) & \text{si } t < T_D \\ \Delta & \text{si } t \geq T_D \end{cases}$$

**Proposition 7.7.** *La fiabilité est donnée par  $R(t) = \alpha_1 e^{tA_{11}} \mathbf{1}_r$*



*Démonstration.*

$$\begin{aligned}
 R(t) &= \mathbb{P}(\forall u \in [0, t], X_u \in U) = \mathbb{P}(Y_t \in U) \\
 &= \sum_{j \in U} \mathbb{P}(Y_t = j) = \sum_{j \in U} \sum_{i \in U} \mathbb{P}(Y_t = j, Y_0 = i) \\
 &= \sum_{j \in U} \sum_{i \in U} \mathbb{P}(Y_t = j \mid Y_0 = i) \mathbb{P}(Y_0 = i) = \sum_{j \in U} \sum_{i \in U} \alpha(i) P_t(i, j) \\
 &= (\alpha_1, 0) P_t \mathbf{1}_{d,r}.
 \end{aligned}$$

□

**Proposition 7.8.** *Temps moyen jusqu'à la défaillance (MTTF : Mean Time To Failure)*

$$MTTF = \mathbb{E}T = -\alpha_1 A_{11}^{-1} \mathbf{1}_r.$$

### 7.2.4 Fiabilité en temps discret

Nous avons défini dans le chapitre 2 les grandeurs de la fiabilité d'un système. La fiabilité d'un système Markovien observé à temps discret  $n \in \mathbb{N}$  peut également être étudiée. L'étude est alors liée de manière naturelle à celle des chaînes de Markov. Le cas d'un système observé à temps continu nécessite l'étude des processus de Markov (indexés par  $\mathbb{R}_+$ ) que nous exposerons dans le chapitre 8.

Soit  $T$  la variable aléatoire à valeurs dans  $\mathbb{N}$  égale à la durée de vie du système. Le taux de défaillance  $h$  est défini par

$$h(n) = \mathbb{P}(T = n \mid T \geq n).$$

Pour que  $h$  soit un taux de défaillance, il faut que  $0 \leq h(n) < 1$  pour tout entier  $n$  et que la série de terme général  $h(n)$  soit divergente.

La loi de  $T$  s'écrit

$$f(n) = \mathbb{P}(T = n) = [1 - h(0)][1 - h(1)] \dots [1 - h(n-1)]h(n),$$

et la fiabilité est

$$R(n) = \mathbb{P}(T > n) = [1 - h(0)][1 - h(1)] \dots [1 - h(n)],$$

et l'on a

$$h(n) = \frac{f(n)}{R(n-1)} = \lambda, \quad n > 0.$$

▷ **Exemple 7.1** *Stock non contrôlé.* - Un composant est pris au hasard dans un stock non contrôlé. La probabilité qu'il soit en bon état à cet instant 0 est  $\alpha$ , et sa durée de vie est alors géométrique. Alors  $T \sim (1 - \alpha)\delta_0 + \alpha\mathcal{G}(p)$ , soit

$$f(n) = \begin{cases} \alpha(1-p)^{n-1}p, & n \geq 1, \\ 1 - \alpha, & n = 0, \end{cases}$$

donc  $R(n) = \alpha(1-p)^n$  pour tout  $n \in \mathbb{N}$ , et enfin

$$h(n) = \begin{cases} p, & n \geq 1, \\ 1 - \alpha, & n = 0. \end{cases}$$

◁

◁

▷ **Exemple 7.2** *Un composant binaire.-* Considérons un composant binaire mis en fonctionnement à l'instant  $n = 0$  dont la durée de vie suit une loi géométrique de paramètre  $p \in ]0, 1[$ . A chaque panne, il est remplacé par un composant neuf identique. Le temps de remplacement est une variable aléatoire de loi géométrique de paramètre  $q \in ]0, 1[$ . Soit  $X_n$  la variable aléatoire égale à 0 si le composant est en bon état à l'instant  $n$ , et à 1 si il est en panne. La suite  $(X_n)$  est une chaîne de Markov d'espace d'état  $E = \{0, 1\}$  et de matrice de transition

$$P = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}$$

Nous avons

$$P^n = \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix}^n = \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{(1-p-q)^n}{p+q} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix},$$

Soit  $\mu$  sa loi initiale. La loi de  $X_n$  est donnée par (3.3) page 52

$$\mathbb{P}_\mu(X_n = 0) = \mu(0)P^n(0, 0) + \mu(1)P^n(1, 0) \quad (7.5)$$

$$= \frac{q}{p+q} + \frac{(1-p-q)^n}{p+q} [p\mu(0) - q\mu(1)] \quad (7.6)$$

et  $\mathbb{P}_\mu(X_n = 1) = 1 - \mathbb{P}_\mu(X_n = 0)$ .

◁

Même pour étudier la fiabilité d'un système complexe (comportant plusieurs composants et plusieurs états), il est souvent suffisant de savoir si l'état dans lequel se trouve le système est un état de bon fonctionnement ou un état défaillant. Pour cela, nous décomposons  $E$  en deux sous-ensembles  $U$  et  $D$ . L'ensemble  $U$  contient les états de bon fonctionnement (ou états de marche, en anglais up-states) et l'ensemble  $D$  les états défaillants (ou états de panne, en anglais down-states).

Dans le cadre de ce modèle, la fiabilité s'exprime comme suit.

$$R(n) = \mathbb{P}(X_k \in U, k \in \{0, \dots, n\}), \quad n \geq 0.$$

Supposons maintenant que la suite  $(X_n)$  des états visités par le système soit une chaîne de Markov, d'espace d'état  $E$  fini, de matrice de transition  $P$ , et de loi initiale  $\mu$ . Dans ce cas, la durée de vie du système est le temps d'entrée dans  $D$ , défini par

$$T = \min\{n \geq 0 : X_n \in D\},$$

avec  $\min \emptyset = +\infty$ . Il est nécessaire de considérer des partitions de matrices et de vecteurs suivant  $U$  et  $D$ . Nous pouvons toujours supposer que  $U = \{1, \dots, r\}$  et  $D = \{r+1, \dots, d\}$ , et écrire  $\mu = [\mu_1, \mu_2]$  et

$$P = \begin{pmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{pmatrix}.$$

La fiabilité s'écrit alors

$$R(n) = \mu_1 P_{11}^n \mathbf{1}_r. \quad (7.7)$$

D'après le théorème 3.26, on a

$$\mathbb{E}T = \sum_{i \in U} \mathbb{E}[T \mathbf{1}_{(X_0=i)}] = \sum_{i \in U} \mathbb{P}(X_0 = i) \mathbb{E}_i T = \mu_1 L,$$

où  $L = (\mathbb{E}_1 T, \dots, \mathbb{E}_m T)' = (I - P_{11})^{-1} \mathbf{1}_r$ . On en déduit que

$$\mathbb{E}T = \mu_1 (I - P_{11})^{-1} \mathbf{1}_r. \quad (7.8)$$

La variance de la durée de vie  $T$  est

$$\text{Var}_i T = V(i) - (L(i))^2,$$

où

$$V = [\mathbb{E}_1(T^2), \dots, \mathbb{E}_m(T^2)]' = (I - P_{11})^{-1} [I + 2P_{11}(I - P_{11})^{-1}] \mathbf{1}_r. \quad (7.9)$$

▷ **Exemple 7.3** Soit  $(X_n)$  la chaîne de Markov de l'exemple 3.15, qui décrit le comportement aléatoire d'un système à 3 états, d'espace d'état  $E = \{1, 2, 3\}$ , de loi initiale  $\mu = (\mu(1), \mu(2), \mu(3)) = (1, 0, 0)$ . Nous supposons que les états 1 et 2 sont les états de bon fonctionnement et que l'état 3 est l'état de panne, soit  $U = \{1, 2\}$  et  $D = \{3\}$ . La fiabilité du système est donnée sous forme matricielle par

$$R(n) = \mu_1 P_{11}^n \mathbf{1}_2 = (1, 0) \begin{pmatrix} 0.1 & 0.9 \\ 0.4 & 0.4 \end{pmatrix}^n \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

avec  $\mu_1 = (\mu(1), \mu(2))$ . Le temps moyen de bon fonctionnement est

$$\mathbb{E}T = \mu_1 (I - P_{11})^{-1} \mathbf{1}_2 = (1, 0) \begin{pmatrix} 0.9 & -0.9 \\ -0.4 & 0.6 \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 25/3.$$

Remarquons que la fiabilité se dégrade (c'est-à-dire tend vers 0) à vitesse géométrique  $((5 + 3\sqrt{17})/20)^n$ , donnée par la plus grande valeur propre de  $P_{11}$ , d'après la représentation spectrale (3.12) page 71.

◁

## 7.3 Evolution de populations et modèles de branchement

### 7.3.1 Introduction

F. Galton (1822-1911) a publié, le 1 avril 1873, dans la revue londonienne *The Educational Times* le problème numéro 4001 qui décrivait l'évolution des noms de familles anglaises et leurs chances d'extinction. Le révérend H. W. Watson (1827-1903), ami de F. Galton, donne une solution à ce problème la même année, le 1 août 1873. Mais Watson a incorrectement conclu sur la probabilité d'extinction des noms. La solution correcte du problème a été obtenue, 60 ans plus tard, en 1930 par G. F. Steffensen.

Plus récemment, des mathématiciens australiens ont découvert un article de I. J. Bienaymé (1796-1878), publié le 29 mars 1845 dans la revue de la *Société Philomatique de Paris*, où l'auteur présente le même problème et il dit que la moyenne des descendants doit être strictement supérieure à 1 pour la survie des noms, mais il ne donne aucune démonstration.

Dans un contexte biologique, nous pouvons énoncer ce problème comme suit. Dans une population biologique chaque individu peut engendrer des descendants. Les individus présents à l'instant initial sont dits appartenir à la génération 0; leurs descendants directs sont dits appartenir à la génération 1, leurs petits-enfants à la génération 2, etc.

Soient  $Z_n$  une v.a. représentant le nombre d'individus de la  $n$ -ième génération,  $n \in \mathbb{N}$ . Notons  $\xi_{n,j}$  le nombre des descendants directs du  $j$ -ième individu de la  $n$ -ème génération. Supposons que les v.a.  $\xi_{n,j}$ ,  $n \in \mathbb{N}$ ,  $j \in \mathbb{N}^*$  sont i.i.d. de loi commune  $p = (p_k, k \geq 0)$  sur  $\mathbb{N}$ .



3. Posons  $u_n := \mathbb{P}(Z_n = 0)$ . Alors la suite  $u_n$  converge vers un nombre  $u$  qui vérifie l'équation

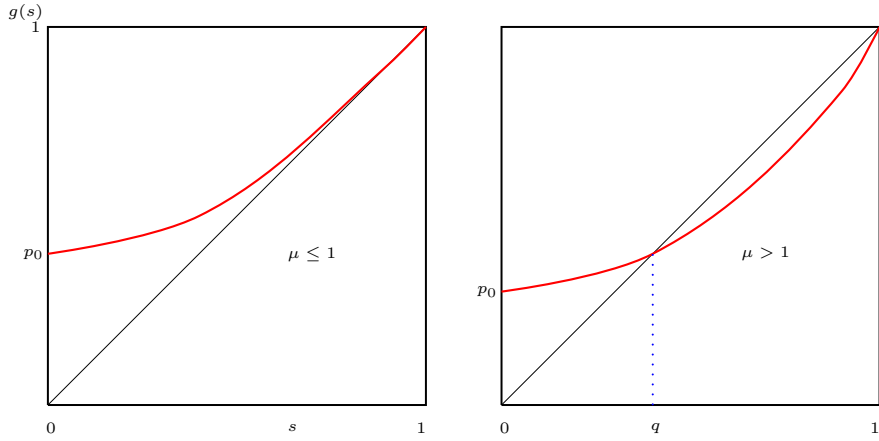
$$u = g(u). \quad (E)$$

4. Si  $w$  est une solution de (E), alors  $u \leq w$ .

5.  $\mathbb{P}(\nu < n) = g_n(0)$  et  $\mathbb{P}(\nu = n) = g_n(0) - g_{n-1}(0)$ .

**Proposition 7.10.** On suppose que  $0 < p_0 < 1$ . Alors

- si  $\mu < 1$ , la population s'éteint p.s., i.e.,  $\lim_{n \rightarrow \infty} \mathbb{P}(Z_n = 0) = 1$  ;
- si  $\mu = 1$ , la population s'éteint p.s., sauf si  $p_1 = 1$  où la probabilité que la population s'éteint est égale 0 ;
- si  $\mu > 1$  alors  $\lim_{n \rightarrow \infty} \mathbb{P}(Z_n = 0) = q$ , où  $q$  est l'unique solution de l'équation (E) dans  $[0, 1[$ .



Le processus est dit critique si  $\mu = 1$ , sous-critique si  $\mu < 1$  et sur-critique si  $\mu > 1$ .

### 7.3.3 Modèles de branchement à temps continu

Soit un processus de branchement, et soit  $T$  le temps aléatoire de durée de vie d'un individu dont la fonction de répartition est  $F$ . Un individu donne naissance à  $j$  individus avec probabilité  $p_j$ ,  $j \geq 1$ . Soit  $M(t)$  le nombre moyen d'individu en vie au temps  $t$ . Alors  $M(t)$  vérifie l'équation suivante

$$M(t) = \bar{F}(t) + \mu \int_0^t M(t-s) dF(s). \quad (7.11)$$

Dans le cas où  $\mu > 1$ , nous obtenons :

$$\lim_{t \rightarrow \infty} e^{-\beta t} M(t) = \frac{\mu - 1}{\beta \mu^2 \int_0^\infty s e^{-\beta s} dF(s)}, \quad (7.12)$$

où  $\beta > 0$  est tel que  $\int_0^\infty e^{-\beta s} dF(s) = 1/\mu$ .

## 7.4 Problèmes

### 7.1. Une banque à deux guichets.

Une banque possède deux guichets ; un pour les commerçants et un pour les clients ordinaires. Les taux d'arrivée et de service pour le guichet de commerçants sont :  $\lambda_1$  et  $\mu_1$  respectivement. De même, les taux d'arrivée et de service pour le guichet des clients ordinaires sont :  $\lambda_2$  et  $\mu_2$  respectivement. Les deux files sont indépendantes et les arrivées forment deux processus de Poisson indépendants et les temps de services suivent des lois exponentielles. Le nombre des places dans la banque peut être considéré infini. Nous supposons que le système est stabilisé. Soient  $\lambda_1 = 6$ ,  $\mu_1 = 12$ ,  $\lambda_2 = 12$  et  $\mu_2 = 24$  (clients / heure).

1. Caractériser les deux files par la notation de Kendall.
2. Donner le nombre moyen de clients dans la banque au temps  $t$  fixé mais quelconque.
3. Donner les conditions de stabilité du système.
4. Donner les temps d'attente moyens passés dans chacune de deux files séparément pour un client.
5. Si on réunit les deux files en une seule, i.e., les deux guichets traitent indifféremment les deux types de clients, avec taux d'arrivée  $\lambda_1 + \lambda_2$  et taux de service commun pour les deux guichets  $(\mu_1 + \mu_2)/2$ , alors donnez :
  - (a) la notation de Kendall pour ce système ;
  - (b) la condition de stabilité ;
  - (c) le nombre moyen de clients dans la banque au temps  $t \geq 0$  (arbitraire mais fixé) ; et
  - (d) le temps d'attente moyen passé dans la file pour un client.
6. Laquelle de ces deux solutions est la plus performante ? Expliquer.

### 7.2. Un système subit des chocs.

Un système subit des chocs à des instants aléatoires suivant un processus de Poisson homogène d'intensité  $\lambda$ . A chaque choc, le système peut tomber en panne avec une probabilité constante et égale à  $p$ . Soient  $T$  le temps de la défaillance du système et  $N$  le nombre de chocs subis jusqu'à la défaillance.

Calculer :

1. la fonction de répartition de  $T$  conditionnelle à  $\{N = n\}$ , c'est-à-dire  $\mathbb{P}(T \leq t \mid N = n)$  ;
2.  $\mathbb{P}(N = n)$ ,  $n \geq 1$ .
3.  $\mathbb{P}(T \leq t)$ ,  $t \geq 0$ .
4. la durée de vie moyenne du système.
5. A.N. Si  $\lambda = 0,001$  et  $p = 0,1$ , calculer  $\mathbb{E}N$  et  $\mathbb{E}T$ .

**7.3.** Un ordinateur tolérant aux fautes contient  $N$  processeurs du même type et il est embarqué dans une navette spatiale (ce problème concerne la période où les astronautes n'effectuaient pas de réparations). Nous allons supposer que chaque processeur peut tomber en panne avec la même probabilité  $\lambda h + o(h)$  dans  $[t, t + h]$ ,  $h > 0$  et  $h \downarrow 0$ , à condition qu'il ne soit pas tombé en panne avant  $t$ . Soit  $X_t$  le nombre des processeurs en vie au temps  $t$ .

1. Montrer que  $X_t, t \geq 0$ , est un processus de Markov et donner sa loi marginale au temps  $t$ .
2. Si  $\tau_j$  est le temps nécessaire pour que le nombre de processeurs en vie soit  $j$ , ( $0 \leq j \leq N$ ), alors calculer son espérance mathématique.
3. Si les processeurs forment un système  $k$ -sur- $N$  :G (c'est un système dont le bon fonctionnement est assuré si au moins  $k$  de ses composants parmi  $N$  sont en bon état), alors calculer sa fiabilité et sa disponibilité au temps  $t$ .

**7.4.** Un système est formé d'un grand nombre de composants en fonctionnement simultané. Chaque composant en fonctionnement au temps  $t$ , a une probabilité  $\lambda h + o(h)$  de tomber en panne dans l'intervalle de temps  $]t, t+h]$ , lorsque  $h \downarrow 0$ . Si le nombre de composants en fonctionnement à l'instant  $t$  est  $N \geq 2$ , montrer que la probabilité que deux composants ou plus tombent en panne dans l'intervalle  $]t, t+h]$ , est en  $o(h)$ , lorsque  $h \downarrow 0$ . Les durées de vies des composants sont indépendantes les unes des autres et les composants sont non réparables.

**7.5.** La consommation en débit d'eau est assurée par deux pompes hydrauliques en fonctionnement simultané. Chaque pompe fournit un débit des  $m$  ( $m > 0$ ) unités de volume par unité de temps lorsque elle est en marche et 0 lorsqu'elle est en panne. Le comportement du système est modélisé par une chaîne de Markov avec espace d'états  $E = \{1, 2, 3\}$ , (1="les deux pompes sont en fonctionnement", 2="une pompe en fonctionnement", 3="zéro pompes en fonctionnement"), et matrice de transition :

$$P = \begin{pmatrix} 0.8 & 0.2 & 0 \\ 0.1 & 0.8 & 0.1 \\ 0 & 0.2 & 0.8 \end{pmatrix}$$

et loi initiale  $\alpha = (1, 0, 0)$ .

1. Donner l'expression du débit moyen au temps  $n > 0$  et calculer sa valeur pour  $n = 2$  et  $m = 5$ .
2. Lorsqu'une pompe tombe en panne son coût de remise en service est égal à  $r \text{ €}$  ( $r > 0$ ). Donner l'expression du coût moyen des réparations sur l'intervalle de temps  $[0, n]$ , ( $n > 0$ ), et calculer sa valeur pour  $n = 3$  et  $r = 100$ .

**7.6.** Un réacteur nucléaire a deux modes de fonctionnement, notés 1 (mode nominal) et 2 (mode dégradé) et deux modes de défaillance, notés 3 (pour fuite de liquide radioactif) et 4 (pour explosion).

Le comportement temporel de ce réacteur est décrit par une CM,  $X = (X_n, n \in \mathbb{N})$  avec espace d'état  $E = \{1, 2, 3, 4\}$  (avec la signification donnée plus haut). Sa matrice de transition est

$$P = \begin{pmatrix} 0,6 & 0,3 & 0,1 & 0 \\ 0,4 & 0,5 & 0 & 0,1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

et de loi initiale  $\alpha = (1/2, 1/2, 0, 0)$ .

1. Caractériser les états de cette CM.
2. Calculer la probabilité que le réacteur explose.
3. Donner l'expression de fiabilité du réacteur au temps  $n$  ( $n \in \mathbb{N}$ ) et la calculer aux temps  $n = 0, 1, 2$ .
4. Calculer le temps moyen de défaillance du réacteur (fuite de liquide radioactif ou explosion).

**7.7.** Considérons deux systèmes de files d'attente. Le système A : M/M/1 avec taux d'arrivées  $\lambda$  et taux de service  $2\mu$  ; et le système B : M/M/2, avec taux d'arrivées  $\lambda$  et taux de service pour chaque serveur  $\mu$ .

1. Calculer les lois stationnaires  $\pi^A$  et  $\pi^B$  de deux systèmes. Préciser les conditions de stabilité pour les deux systèmes.
2. Calculer les temps moyen qu'un client passe dans chacun des deux systèmes, soient  $\mathbb{E}[W_A]$  et  $\mathbb{E}[W_B]$ .
3. Quel système est le plus performant dans les deux cas : (a)  $\lambda/\mu = 0.8$ , et (b)  $\lambda/\mu = 0.9$  ?

*Attention.* Il n'est pas demandé de dériver la formule du temps moyen.

**7.8.** Une banque souhaite évaluer le temps qu'un client passe dans un de ses distributeurs de billets (ATM). Les trois opérations qu'un client peut effectuer, après avoir inséré sa carte bancaire et introduit son code dans le distributeur, sont : 'retirer de l'argent', 'déposer un chèque', 'information sur son compte'. Les durées de ces trois opérations sont aléatoires de lois exponentielles de paramètres :  $\lambda_1$ ,  $\lambda_2$  et  $\lambda_3$  respectivement. Le temps aussi d'introduction de la carte et du code, sont d'une durée aléatoire de loi exponentielle de paramètre  $\lambda$ .

Lorsque le client entre au système, il a une probabilité  $p$  de commencer par retirer de l'argent, une probabilité  $q$  de commencer par déposer un chèque et une probabilité  $r$  de commencer par demander des informations sur son compte (on a  $p+q+r=1$ ). Lorsque il termine une opération, il a une probabilité égale à  $1-\beta$  pour terminer et sortir de l'ATM, et une probabilité  $\beta$  de se diriger vers les deux autres opérations avec probabilités égales chacune à  $\beta/2$ .

Nous allons modéliser ce problème par un processus de Markov  $X = (X_t, t \in \mathbb{R}_+)$ , avec espace d'état  $E = \{1, 2, 3, 4, 5\}$ .

Les significations des cinq états sont :

- 1 : introduction de la carte et du code ;
- 2 : retirer de l'argent ;
- 3 : déposer un chèque ;
- 4 : demander de l'information sur le compte ;
- 5 : sortie de l'ATM (état absorbant).

1. Donner le graphe de transition du processus  $X$ .
2. Donner la matrice de transition  $Q$  de la chaîne de Markov immergée  $(Y_n, n \in \mathbb{N})$ , i.e.,  $Y_n = X_{T_n}$ ,  $n \geq 0$ , avec  $T_n$  le temps du  $n$ -ième saut du processus  $X$ .
3. Calculer le générateur du processus  $X$ .
4. Calculer le temps moyen qu'un client passe dans l'ATM.
5. Donner l'expression de la probabilité qu'un client passe plus de  $x$  unité de temps dans l'ATM ( $x > 0$ ).

**7.9.** L'évolution d'une population biologique est décrite par un processus de branchement  $Z_n, n \geq 0$ , avec  $Z_0 = 2$ . Le nombre des descendants d'un individu quelconque est une variable aléatoire indépendante des autres naissances et suit une loi binomiale  $b(3, p)$ ,  $0 < p < 1$ .

1. Donner la condition, notée  $C$ , sur  $p$ , d'extinction de la population, avec probabilité 1.
2. Sous la condition  $C$ , calculer le nombre moyen d'individus ayant vécu avant l'extinction.
3. Lorsque la condition  $C$  sur  $p$ , n'est pas vérifiée, alors calculer la probabilité d'extinction de la population. Donner sa valeur lorsque  $p = 1/2$ .

**7.10.** L'évolution d'une population biologique est décrite par un modèle de branchement  $Z = (Z_n, n \geq 0)$ , avec  $Z_0 = 1$ . Le nombre des descendants directs d'un individu quelconque est une variable aléatoire,  $\xi$ , de loi  $P$ , et de fonction génératrice  $g$ . Notons également par  $g_n$  la fonction génératrice de  $Z_n$ . On a  $g_1 \equiv g$ .

Soit  $\nu$  le numéro de la première génération lors de l'extinction de la population, i.e.,  $\nu = \min\{n : Z_n = 0\}$  et posons  $u_n := \mathbb{P}(\nu \leq n)$ .

1. Cas général.
  - (a) Montrer que  $u_n = \mathbb{P}(Z_n = 0) = g_n(0)$ .
  - (b) Montrer que  $\mathbb{P}(\nu = n) = g_n(0) - g_{n-1}(0)$ .
2. Cas particulier. Supposons que  $g(s) = q + ps$ , avec  $0 < p < 1$  et  $p + q = 1$ .
  - (a) Donner la loi de probabilité  $P$ .



- (b) Montrer par récurrence que  $g_n(s) = 1 - p^n + p^n s$ , pour  $n \geq 1$ .
  - (c) Calculer  $\mathbb{P}(\nu = n)$  pour  $n \geq 1$ .
  - (d) Calculer  $\mathbb{P}(\nu > n)$  pour  $n \geq 0$ .
  - (e) Calculer le numéro de la génération moyen d'extinction de cette population, i.e.,  $\mathbb{E}\nu$ .
3. Application numérique. Soient  $p = 0.9$  et  $q = 0.1$ .
- (a) Calculer la loi de la v.a.  $\nu$ .
  - (b) Calculer le numéro de la génération moyen d'extinction de la population.

# Bibliographie

- [1] Anderson, W.J. (1991). *Continuous-time Markov chains*, Springer Verlag.
- [2] Asmussen S. (1987). *Applied Probability and Queues*, Wiley, Chichester.
- [3] Barbu V. and Limnios N. (2008). *Semi-Markov Chains and Hidden Semi-Markov Models. Toward Applications. Their use in Reliability and DNA Analysis*, Lecture Notes in Statistics, vol. 191, Springer, New York.
- [4] Billingsley, P. (1986). *Probability and measures*, J. Wiley, N.Y..
- [5] Bouleau, N.(1988). *Processus stochastiques et applications*, Hermann, Paris.
- [6] Çinlar, E. (1975) *Introduction to stochastic processes*, Prentice-Hall, Englewood Cliffs, N.J..
- [7] Ching W.-K., Ng M.K. (2006). *Markov chains - Models, Algorithms and Applications*, Springer's International series In Operations Research, USA.
- [8] Coccozza C. (1997). *Processus stochastiques et fiabilité des systèmes*, SMAI, Springer.
- [9] Cox, D.R. (1962) *Renewal Theory*, Metuhen, London.
- [10] Cox, D.R., Miller, H.D. (1965). *The Theory of Stochastic Processes*, Methuen, London.
- [11] Dacunha-Castelle, D., Duflo, M. (1983). *Probabilités et statistiques : 2- Problèmes à temps mobile*, Masson, Paris.
- [12] Doob, J.L. (1953). *Stochastic processes*, J.Wiley, Inc.
- [13] Feller, W. (1971). *An introduction to probability theory and its applications*, Vol. I & II, J. Wiley, Inc., N.Y..
- [14] Hoel P.G., Port S.C., Stone C.J. (1972). *Introduction to stochastic processes*, Waveland Press, Boston.
- [15] Gihman, I.I., Skorohod, A.V. (1974). *The theory of random processes*, Vol. I, II & III, Springer-Verlag, N.Y..
- [16] Girardin, V., Limnios, N. (2001, 2008, 2014). *Probabilités en vue des applications*, Vol. 1 et Vol. 2, Vuibert, Paris.
- [17] Girardin, V., Limnios, N. (2018). *Applied Probability - From Random Sequences to Stochastic Processes*, Springer, Switzerland.
- [18] Iosifescu, M. (1980). *Finite processes and their applications*, Dover Pub.
- [19] Iosifescu, M., Limnios, N., Oprişan, G. (2007). *Modèles stochastiques*, Hermes/Lavoisier, Paris.
- [20] Karatzas I., Shreve S.E. (1988). *Brownian Motion and Stochastic Calculus*, Springer-Verlag.
- [21] Karlin S., Taylor H.M. (1981). *A Second Course in Stochastic Processes*, Academic Press, San Diego.
- [22] Kemeny, J.G., Snell, J.L. (1976). *Finite Markov chains*, Springer-Verlag, N.Y..
- [23] Korolyuk, V.S., Limnios, N. (2005). *Stochastic systems in merging phase space*, World Scientific.

- [24] Limnios, N., Oprisan, G. (2001). *Semi-Markov processes and reliability*, Birkhäuser, Boston.
- [25] Mackevicius V. (2011). *Introduction to stochastic analysis*, Iste, J. Wiley, London.
- [26] Meyn, S.P., Tweedie, R.L. (1994). *Markov chains and stochastic stability*, Springer, London.
- [27] Neuts, M.F. (1981). *Matrix-Geometric solutions in stochastic models*, The Johns Hopkins University Press.
- [28] Norris J.R. (1997). *Markov chains*, Cambridge University Press, N.Y.
- [29] Pardoux E. (2007). *Processus de Markov et applications*, Dunod, Paris.
- [30] Revuz D., Yor M. (1999). *Continuous Martingales and Brownian Motion*, Springer, 3rd Edition, Berlin.
- [31] Rogers L.C.G., Williams D. (1994). *Diffusions, Markov Processes, and Martingales*, vol. 1 & 2, J. Wiley & Sons, Chichester, U.K.
- [32] Rozanov Y.A. (1969). *Probability Theory : A Concise Course*, Dover Publ., N.Y.
- [33] Ross, S.M. (1969). *Applied probability models with optimization applications*, Dover Publ., Inc., N.Y..
- [34] Seneta E. (1981). *Non-negative Matrices and Markov Chains*, Springer-Verlag.
- [35] Shiryaev, A.N. (1996). *Probability*, 2nd Edition, Springer-Verlag, N.Y..
- [36] Skorokhod A.V. (1991). *Random Processes with Independent Increments*, Kluwer, Dordrecht.
- [37] Stroock D.W. (1993). *Probability Theory : An Analytic View*, Cambridge University Press, Cambridge.
- [38] Tuffin B. (2010). *La simulation de Monte Carlo*, Hermes/Lavoisier.
- [39] Wang,Z., Yang, X. (1992). *Birth and death processes and Markov chains*, Springer.

# TABLE

Fonction de répartition de la loi normale centrée réduite :  $\mathcal{N}(0, 1)$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$$

x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.500 00	0.503 99	0.507 98	0.511 97	0.515 95	0.519 94	0.523 92	0.527 90	0.531 88	0.535 86
0.1	0.539 83	0.543 80	0.547 76	0.551 72	0.555 67	0.559 62	0.563 56	0.567 50	0.571 42	0.575 35
0.2	0.579 26	0.583 17	0.587 06	0.590 95	0.594 84	0.598 71	0.602 57	0.606 42	0.610 26	0.614 09
0.3	0.617 91	0.621 72	0.625 52	0.629 30	0.633 07	0.636 83	0.640 58	0.644 31	0.648 03	0.651 73
0.4	0.655 42	0.659 10	0.662 76	0.666 40	0.670 03	0.673 65	0.677 24	0.680 82	0.684 39	0.687 93
0.5	0.691 46	0.694 97	0.698 47	0.701 94	0.705 40	0.708 84	0.712 26	0.715 66	0.719 04	0.722 40
0.6	0.725 75	0.729 07	0.732 37	0.735 65	0.738 91	0.742 15	0.745 37	0.748 57	0.751 75	0.754 90
0.7	0.758 04	0.761 15	0.764 24	0.767 31	0.770 35	0.773 37	0.776 37	0.779 35	0.782 30	0.785 24
0.8	0.788 14	0.791 03	0.793 89	0.796 73	0.799 55	0.802 34	0.805 11	0.807 85	0.810 57	0.813 27
0.9	0.815 94	0.818 59	0.821 21	0.823 81	0.826 39	0.828 94	0.831 47	0.833 98	0.836 46	0.838 91
1.0	0.841 34	0.843 75	0.846 14	0.848 50	0.850 83	0.853 14	0.855 43	0.857 69	0.859 93	0.862 14
1.1	0.864 33	0.866 50	0.868 64	0.870 76	0.872 86	0.874 93	0.876 98	0.879 00	0.881 00	0.882 98
1.2	0.884 93	0.886 86	0.888 77	0.890 65	0.892 51	0.894 35	0.896 17	0.897 96	0.899 73	0.901 47
1.3	0.903 20	0.904 90	0.906 58	0.908 24	0.909 88	0.911 49	0.913 09	0.914 66	0.916 21	0.917 74
1.4	0.919 24	0.920 73	0.922 20	0.923 64	0.925 07	0.926 47	0.927 86	0.929 22	0.930 56	0.931 89
1.5	0.933 19	0.934 48	0.935 74	0.936 99	0.938 22	0.939 43	0.940 62	0.941 79	0.942 95	0.944 08
1.6	0.945 20	0.946 30	0.947 38	0.948 45	0.949 50	0.950 53	0.951 54	0.952 54	0.953 52	0.954 49
1.7	0.955 43	0.956 37	0.957 28	0.958 19	0.959 07	0.959 94	0.960 80	0.961 64	0.962 46	0.963 27
1.8	0.964 07	0.964 85	0.965 62	0.966 38	0.967 12	0.967 84	0.968 56	0.969 26	0.969 95	0.970 62
1.9	0.971 28	0.971 93	0.972 57	0.973 20	0.973 81	0.974 41	0.975 00	0.975 58	0.976 15	0.976 70
2.0	0.977 25	0.977 78	0.978 31	0.978 82	0.979 32	0.979 82	0.980 30	0.980 77	0.981 24	0.981 69
2.1	0.982 14	0.982 57	0.983 00	0.983 41	0.983 82	0.984 22	0.984 61	0.985 00	0.985 37	0.985 74
2.2	0.986 11	0.986 45	0.986 79	0.987 13	0.987 45	0.987 78	0.988 09	0.988 40	0.988 70	0.988 99
2.3	0.989 28	0.989 56	0.989 83	0.990 10	0.990 36	0.990 61	0.990 86	0.991 11	0.991 34	0.991 58
2.4	0.991 80	0.992 02	0.992 24	0.992 45	0.992 66	0.992 86	0.993 05	0.993 24	0.993 43	0.993 61
2.5	0.993 79	0.993 96	0.994 13	0.994 30	0.994 46	0.994 61	0.994 77	0.994 92	0.995 06	0.995 20
2.6	0.995 34	0.995 47	0.995 60	0.995 73	0.995 85	0.995 98	0.996 09	0.996 21	0.996 32	0.996 43
2.7	0.996 53	0.996 64	0.996 74	0.996 83	0.996 93	0.997 02	0.997 11	0.997 20	0.997 28	0.997 36
2.8	0.997 44	0.997 52	0.997 60	0.997 67	0.997 74	0.997 81	0.997 88	0.997 95	0.998 01	0.998 07
2.9	0.998 13	0.998 19	0.998 25	0.998 31	0.998 36	0.998 41	0.998 46	0.998 51	0.998 56	0.998 61

Queue de la loi :  $1 - \Phi(x)$

x	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
3.	135 10e-5	968 10e-6	687 10e-6	483 10e-6	337 10e-6	233 10e-6	159 10e-6	108 10e-6	723 10e-7	481 10e-7
4.	317 10e-7	207 10e-7	133 10e-7	85 10e-7	54 10e-7	34 10e-7	21 10e-7	13 10e-7	79 10e-8	48 10e-8
5.	29 10e-8	17 10e-8	10 10e-8	58 10e-8	33 10e-8	19 10e-8	11 10e-8	60 10e-10	33 10e-10	18 10e-10

# TABLE 2

**Loi de Chi-2 à  $k$  degrés  
de libertés : valeurs critiques  $\chi_k^2$**

k	5%	1%	0.1 %
1	3.841	6.635	10.828
2	5.991	9.210	13.816
3	7.815	11.345	16.266
4	9.488	13.277	18.467
5	11.070	15.086	20.515
6	12.592	16.812	22.458
7	14.067	18.475	24.322
8	15.507	20.090	26.124
9	16.919	21.666	27.877
10	18.307	23.209	29.588

## Alphabet grec

A	$\alpha$	Alpha
B	$\beta$	Beta
$\Gamma$	$\gamma$	Gamma
$\Delta$	$\delta$	Delta
E	$\varepsilon$ ( $\epsilon$ )	Epsilon
Z	$\zeta$	Zeta
H	$\eta$	Eta
$\Theta$	$\vartheta$ ( $\theta$ )	Theta
I	$\iota$	Iota
K	$\kappa$	Kappa
$\Lambda$	$\lambda$	Lambda
M	$\mu$	Mu
N	$\nu$	Nu
$\Xi$	$\xi$	Ksi
O	$\omicron$	Omicron
$\Pi$	$\pi$	Pi
P	$\rho$ ( $\varrho$ )	Rho
$\Sigma$	$\sigma$ ( $\varsigma$ )	Sigma
T	$\tau$	Tau
$\Upsilon$	$\upsilon$	Upsilon
$\Phi$	$\phi$ ( $\varphi$ )	Phi
X	$\chi$	Chi
$\Psi$	$\psi$	Psi
$\Omega$	$\omega$	Omega