Université Rennes 2
Licence MASS 2

Introduction aux Probabilités

Arnaud Guyader

Table des matières

1 Espaces probabilisés **1**
- 1.1 Qu'est-ce qu'une probabilité ? . . . 1
  - 1.1.1 Tribu . . . 1
  - 1.1.2 Probabilité . . . 2
- 1.2 Conditionnement . . . 7
- 1.3 Indépendance . . . 10
- 1.4 Exercices . . . 13
- 1.5 Corrigés . . . 25

2 Variables aléatoires discrètes **57**
- 2.1 Loi d'une variable discrète . . . 57
- 2.2 Fonction de répartition . . . 59
- 2.3 Moments d'une variable discrète . . . 61
  - 2.3.1 Espérance . . . 61
  - 2.3.2 Variance . . . 66
  - 2.3.3 Autres moments . . . 68
- 2.4 Corrélation et indépendance . . . 70
- 2.5 Lois usuelles . . . 73
  - 2.5.1 Loi uniforme . . . 74
  - 2.5.2 Loi de Bernoulli . . . 75
  - 2.5.3 Loi binomiale . . . 76
  - 2.5.4 Loi géométrique . . . 77
  - 2.5.5 Loi de Poisson . . . 80
- 2.6 Exercices . . . 82
- 2.7 Corrigés . . . 93

3 Variables aléatoires à densité **115**
- 3.1 Densité d'une variable aléatoire . . . 115
- 3.2 Fonction de répartition . . . 117
- 3.3 Moments d'une variable à densité . . . 121
- 3.4 Lois usuelles . . . 125
  - 3.4.1 Loi uniforme . . . 125
  - 3.4.2 Loi exponentielle . . . 126
  - 3.4.3 Loi normale . . . 128
- 3.5 Exercices . . . 133
- 3.6 Corrigés . . . 145

A Annexes **177**
- A.1 Annales . . . 177
- A.2 Table de la loi normale $X \sim \mathcal{N}(0,1)$ . . . 218

Chapitre 1

Espaces probabilisés

Introduction

Dans ce premier chapitre, on commence par définir axiomatiquement la notion de probabilité sur un ensemble cohérent d'événements (ou tribu). L'idée de probabilité conditionnelle en découle alors très simplement. Elle est entre autres liée à la notion d'indépendance, fondamentale en probabilités comme en statistiques.

1.1 Qu'est-ce qu'une probabilité ?

Avant de définir ce qu'est une probabilité sur un ensemble d'événements, il faut commencer par préciser les propriétés souhaitables pour cet ensemble d'événements.

1.1.1 Tribu

On s'intéresse à une expérience aléatoire dont le résultat est appelé événement élémentaire ω. L'ensemble des résultats possibles, c'est-à-dire l'union des événements élémentaires, est noté Ω et appelé univers ou ensemble fondamental.

Exemples :

1. Lancer d'un dé : on s'intéresse au résultat ω du lancer d'un dé à 6 faces. On a donc $\omega = 1$ ou $\omega = 2$, etc. L'espace fondamental est donc $\Omega = \{1, 2, 3, 4, 5, 6\}$. Cet univers Ω est fini.
2. Infinité de lancers d'une pièce : on lance une infinité de fois une pièce dont l'une des faces est numérotée 0 et l'autre 1. Un événement élémentaire est donc cette fois une suite de 0 et de 1 : $\omega = \{u_1, u_2, \ldots\}$ avec u_n =0 ou 1 pour tout n de $\mathbb{N}^*$. L'espace fondamental est cette fois l'ensemble de toutes les suites possibles formées de 0 et de 1. Cet univers Ω est clairement infini.

Dans la suite, on va vouloir calculer la probabilité de certaines parties de l'espace fondamental Ω. Malheureusement, sauf lorsque Ω sera fini ou dénombrable, on ne pourra pas s'intéresser à l'ensemble $\mathcal{P}(\Omega)$ de toutes les parties de Ω, celui-ci étant en quelque sorte “trop gros”. On se restreindra donc à un sous-ensemble $\mathcal{F}$ de $\mathcal{P}(\Omega)$, qui constituera l'ensemble des parties dont on peut calculer la probabilité. Afin d'obtenir un modèle aussi cohérent que possible, il importe néanmoins d'imposer certaines conditions de stabilité à $\mathcal{F}$: par union, intersection, passage au complémentaire, etc. C'est en ce sens qu'intervient la notion de tribu.

Définition 1.1 (Tribu)
Soit Ω un univers et $\mathcal{F}$ un sous-ensemble de parties de Ω, i.e. $\mathcal{F} \subseteq \mathcal{P}(\Omega)$. On dit que $\mathcal{F}$ est une tribu, ou une σ-algèbre, si elle vérifie les 3 conditions suivantes :
(i) $\Omega \in \mathcal{F}$;
(ii) si A appartient à $\mathcal{F}$, alors son complémentaire $\overline{A}$ (encore noté A^c) appartient aussi à $\mathcal{F}$;
(iii) si $(A_n)_{n\in\mathbb{N}}$ est une suite de $\mathcal{F}$, alors $\bigcup_{n=0}^{+\infty} A_n$ appartient à $\mathcal{F}$.

On appelle dès lors événements les éléments de la tribu $\mathcal{F}$. Rappelons que si A est un événement, alors $\overline{A} = \Omega \setminus A$ est l'événement contraire de A. Par ailleurs, dire que l'événement $\bigcup_{n=0}^{+\infty} A_n$ se réalise signifie que l'un au moins des événements A_n se réalise :

$$\omega \in \bigcup_{n=0}^{+\infty} A_n \;\Leftrightarrow\; \exists n \in \mathbb{N} \;:\; \omega \in A_n.$$

On vérifie sans problème à partir des trois axiomes ci-dessus que toute tribu $\mathcal{F}$ contient l'ensemble vide $\emptyset$, est stable par union finie, intersection finie ou dénombrable. Ainsi, on retiendra qu'une tribu est stable par combinaisons au plus dénombrables d'opérations usuelles sur les ensembles, bref par toutes les manipulations classiques.

Exemples. Voici trois exemples classiques de tribus :
– La tribu triviale : $\mathcal{F} = \{\emptyset, \Omega\}$.
– La tribu engendrée par une partie A de Ω : $\mathcal{F} = \{\emptyset, A, \overline{A}, \Omega\}$.
– La tribu pleine : $\mathcal{F} = \mathcal{P}(\Omega)$.

En pratique, lorsque Ω est fini ou dénombrable, on considère en général la tribu pleine $\mathcal{P}(\Omega)$. C'est le cas par exemple si $\Omega = \{1, 2, 3, 4, 5, 6\}$, ensemble des résultats possibles du lancer d'un dé, ou si $\Omega = \mathbb{N}^*$, date d'apparition du premier Pile dans une succession de lancers d'une pièce (lorsqu'on exclut le cas improbable où Pile n'apparaît jamais). Si Ω n'est pas dénombrable, comme c'est le cas dans l'exemple d'une suite infinie de lancers ($\Omega = \{0, 1\}^{\mathbb{N}^*}$), on ne considérera pas la tribu $\mathcal{F} = \mathcal{P}(\Omega)$, mais une tribu plus petite.

1.1.2 Probabilité

Une fois fixés un univers Ω et une tribu $\mathcal{F}$ de Ω, on peut définir proprement ce qu'est une probabilité sur $(\Omega, \mathcal{F})$. Un point de vocabulaire auparavant : on dit que deux événements A et B sont incompatibles (ou disjoints) si $A \cap B = \emptyset$, et on dit que $(A_n)_{n\geq 0}$ est une suite d'événements deux à deux incompatibles si pour tout couple d'indices distincts (i, j), on a $A_i \cap A_j = \emptyset$.

Définition 1.2 (Probabilité)
On appelle probabilité sur la tribu $\mathcal{F}$ de Ω toute application $\mathbb{P} : \mathcal{F} \to [0, 1]$ telle que
(i) $\mathbb{P}(\Omega) = 1$;
(ii) σ-additivité : si $(A_n)_{n\geq 0}$ est une suite d'événements deux à deux incompatibles de $\mathcal{F}$, alors :

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

On dit alors que $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé.

Exemple. Reprenons l'exemple du lancer de dé. On a vu que l'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}$ et qu'on le munit de la tribu $\mathcal{F} = \mathcal{P}(\Omega)$. On vérifie alors que l'application $\mathbb{P} : \mathcal{F} \to [0, 1]$ qui à $A \in \mathcal{F}$

associe $\mathbb{P}(A) = \#A/6$ est une probabilité sur $\mathcal{F}$, où la notation $\#A$ signifie "cardinal de l'ensemble A".

Généralisation : équiprobabilité sur un univers fini. Dès qu'on considère un univers Ω de cardinal fini sur lequel tout événement élémentaire ω a la même chance d'apparition, on le munira généralement de la même probabilité $\mathbb{P}$ que pour le lancer de dé, appelée équiprobabilité. C'est-à-dire que pour tout événement A, on aura :

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

Nous allons maintenant énoncer diverses propriétés d'une probabilité qui nous seront utiles dans la suite du cours. Rappelons au passage la définition de la soustraction ensembliste "$\setminus$" (figure 1.1) : $B \setminus A = B \cap \overline{A}$.

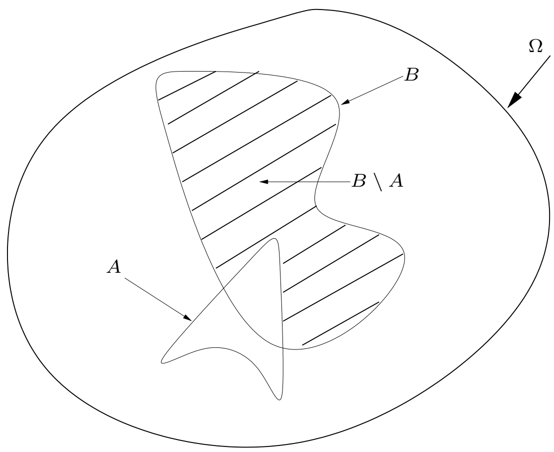


FIGURE 1.1 – Soustraction ensembliste : $B \setminus A = B \cap \overline{A}$.

Propriétés 1.1 (Propriétés d'une probabilité)
Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Tous les ensembles considérés sont supposés appartenir à $\mathcal{F}$.

- *Monotonie : si $A \subseteq B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$. Plus précisément :*

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

- *Additivité forte :*

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

- *Sous$-\sigma-$additivité :*

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n).$$

- *Continuité monotone croissante : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements croissante pour l'inclusion (figure 1.2), alors :*

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

– *Continuité monotone décroissante : si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements décroissante pour l'inclusion (figure 1.3), alors :*

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Preuve.

– Monotonie : il suffit d'appliquer la $\sigma-$additivité avec $A_0 = A$, $A_1 = B \setminus A$ et $A_n = \emptyset$ pour tout $n \geq 2$. Ceci donne :

$$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A),$$

et puisque $\mathbb{P}(B \setminus A) \geq 0$, on a bien $\mathbb{P}(A) \leq \mathbb{P}(B)$.

– Additivité forte : on décompose de façon disjointe

$$A \cup B = (A \setminus (A \cap B)) \cup (A \cap B) \cup (B \setminus (A \cap B)),$$

d'où il vient par $\sigma-$additivité :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \setminus (A \cap B)) + \mathbb{P}(A \cap B) + \mathbb{P}(B \setminus (A \cap B)),$$

et on peut utiliser la propriété précédente :

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

qui aboutit bien à :

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

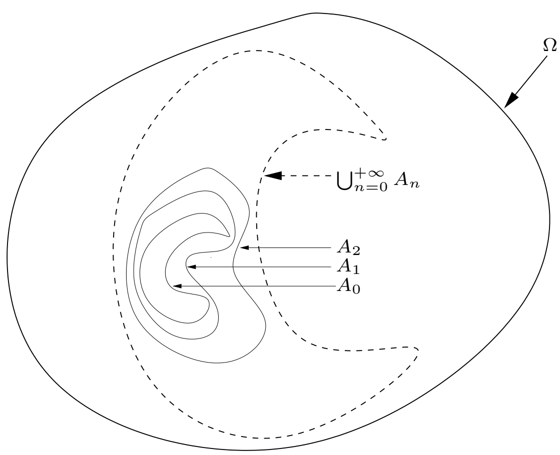


Figure 1.2 – Suite d'ensembles croissante pour l'inclusion.

– Sous-additivité dénombrable : on construit la suite d'ensembles (B_n) comme suit : $B_0 = A_0$ et pour tout $n \geq 1$:

$$B_n = A_n \setminus \left(\bigcup_{k=0}^{n-1} A_k\right).$$

Il est clair que les B_n sont deux à deux disjoints, que $B_n \subseteq A_n$ pour tout n, et que :

$$\bigcup_{n=0}^{+\infty} A_n = \bigcup_{n=0}^{+\infty} B_n.$$

On peut alors appliquer la $\sigma-$additivité :

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(B_n) \leq \sum_{n=0}^{+\infty} \mathbb{P}(A_n),$$

la dernière inégalité provenant de la propriété de monotonie vue ci-dessus.

– Continuité monotone croissante : on reprend la suite d'ensembles (B_n) comme ci-dessus en remarquant que pour tout n :

$$A_n = B_0 \cup B_1 \cup \cdots \cup B_n.$$

Il s'ensuit que :

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} \mathbb{P}(B_n) = \lim_{N\to+\infty} \sum_{n=0}^{N} \mathbb{P}(B_n) = \lim_{N\to+\infty} \mathbb{P}(A_N).$$

– Continuité monotone décroissante : on considère cette fois la suite d'ensembles $(C_n)_{n\geq 0}$ définie par : $C_n = A_0 \setminus A_n$. Par la propriété de monotonie on a donc :

$$\forall n \geq 0 \qquad \mathbb{P}(C_n) = \mathbb{P}(A_0) - \mathbb{P}(A_n).$$

La suite $(C_n)_{n\geq 0}$ est croissante et :

$$\bigcup_{n=0}^{+\infty} C_n = A_0 \setminus \left(\bigcap_{n=0}^{+\infty} A_n\right).$$

Puisque l'intersection des A_n est contenue dans A_0, la monotonie ci-dessus assure que :

$$\mathbb{P}\left(A_0 \setminus \left(\bigcap_{n=0}^{+\infty} A_n\right)\right) = \mathbb{P}(A_0) - \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right).$$

On peut alors appliquer la continuité monotone croissante :

$$\mathbb{P}(A_0) - \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(C_n) = \mathbb{P}(A_0) - \lim_{n\to+\infty} \mathbb{P}(A_n),$$

ce qui donne le résultat voulu, à savoir :

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

■

Remarque. La propriété d'additivité forte se généralise à un nombre quelconque n d'ensembles et a déjà été rencontrée dans des problèmes de dénombrement : c'est la formule de Poincaré (ou d'inclusion-exclusion, ou du crible). Rappelons-la pour $n = 3$:

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

et de façon générale :

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1\leq i_1 < \cdots < i_k \leq n} \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k})\right).$$

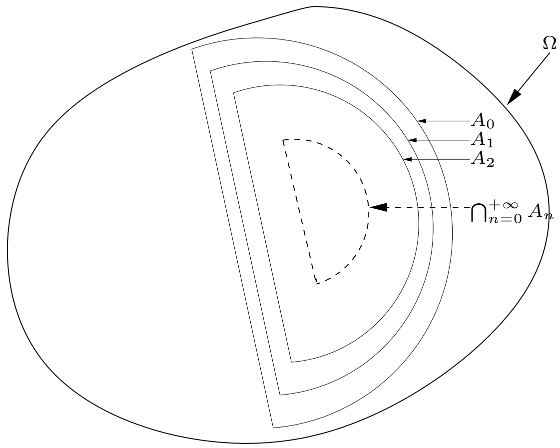


Figure 1.3 – Suite d'ensembles décroissante pour l'inclusion.

Une application est donnée dans l'exercice 1.8.

On a vu que lorsqu'on a équiprobabilité sur un univers fini Ω, la mesure de probabilité $\mathbb{P}$ est celle qui à tout événement A associe le rapport de son cardinal au cardinal de Ω. En d'autres termes $\Omega = \{\omega_1, \ldots, \omega_n\}$ et pour tout $i = 1, \ldots, n$: $p_i = \mathbb{P}(\{\omega_i\}) = 1/n$. Supposer qu'on n'a pas équiprobabilité des événements élémentaires ω_i revient à considérer une séquence $(p_1, \ldots, p_n)$ de nombres positifs et sommant à 1, mais dont tous les coefficients p_i ne sont pas égaux. On définit alors encore une mesure de probabilité sur $\mathcal{P}(\Omega)$ en considérant pour tout événement $A \in \mathcal{P}(\Omega)$:

$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i,$$

où la notation "$i : \omega_i \in A$" signifie que la somme est effectuée sur l'ensemble des indices i pour lesquels $\omega_i \in A$.

Exemple : On lance 3 fois de suite une pièce équilibrée et on compte le nombre de fois où Pile est apparu. On a donc $\Omega = \{0, 1, 2, 3\}$, mais il n'y a pas équiprobabilité puisque les probabilités élémentaires sont $(1/8, 3/8, 3/8, 1/8)$.

Si on veut construire une probabilité $\mathbb{P}$ sur un ensemble infini dénombrable, typiquement sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, on ne peut plus avoir équiprobabilité des événements élémentaires $\{n\}$. Supposons en effet que pour tout $n \in \mathbb{N}$ on ait $\mathbb{P}(\{n\}) = p > 0$, alors la sigma-additivité de $\mathbb{P}$ imposerait que :

$$\mathbb{P}(\mathbb{N}) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} \{n\}\right) = \sum_{n=0}^{+\infty} \mathbb{P}(\{n\}) = \sum_{n=0}^{+\infty} p = +\infty,$$

ce qui est en contradiction avec la condition $\mathbb{P}(\mathbb{N}) = 1$. Une façon de construire une probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ est de généraliser le procédé que l'on vient de voir pour les ensembles finis : considérer une suite $(p_n)_{n \geq 0}$ de nombres positifs telle que la série $\sum_{n \geq 0} p_n$ soit convergente et de somme 1. Comme précédemment, on définit alors pour tout événement $A \in \mathcal{P}(\Omega)$ sa probabilité par :

$$\mathbb{P}(A) = \sum_{i:\omega_i \in A} p_i,$$

la seule différence avec le cas précédent étant que cette fois la somme considérée peut être la somme d'une série (dès lors que le sous-ensemble A est infini).

Exemple : On lance une pièce équilibrée jusqu'à ce que Pile apparaisse (toujours en excluant le cas improbable où Pile n'apparaît jamais). On a donc $\Omega = \{1, 2, \ldots\} = \mathbb{N}^*$. On a clairement $p_1 = \mathbb{P}(\{1\}) = 1/2$, $p_2 = 1/4$ et de façon générale $p_n = 1/2^n$. On reconnaît dans les p_n les termes d'une suite géométrique dont la somme vaut bien 1.

1.2 Conditionnement

La notion de conditionnement sera d'usage constant dans la suite puisqu'elle permet par exemple de tenir compte de l'information dont on dispose déjà pour évaluer la probabilité d'un nouvel événement. Même en l'absence de toute chronologie sur les événements, un détour par un conditionnement astucieux nous permettra souvent d'arriver à nos fins.

Dans tout ce qui suit, $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé arbitraire et tous les ensembles considérés sont des événements de $\mathcal{F}$. Nous commençons par définir la probabilité conditionnelle sachant un événement.

Définition 1.3 (Probabilité conditionnelle)
Soit A et B deux événements, avec $\mathbb{P}(A) > 0$. La probabilité de B sachant A est définie par :

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

Remarque. On peut en fait généraliser la définition de $\mathbb{P}(B|A)$ au cas où A est de probabilité nulle : il suffit de poser $\mathbb{P}(B|A) = 0$.

Concrètement, l'expression "probabilité de B sachant A" signifie "probabilité que B se réalise sachant que A s'est réalisé". La probabilité de B sachant A est donc encore une probabilité au sens usuel du terme (i.e. en particulier un nombre compris entre 0 et 1). Par contre, la probabilité de B peut être faible alors que la probabilité de B sachant A est grande (et réciproquement).

Exemple. Une urne contient 90 boules noires, 9 boules blanches et 1 boule rouge. On tire une boule au hasard : quelle est la probabilité qu'elle soit blanche ? La réponse est bien sûr $\mathbb{P}(B) = 9/100$, donc une probabilité faible. On tire une boule au hasard : quelle est la probabilité qu'elle soit blanche, sachant que la boule tirée n'est pas noire ? Si on note A l'événement "La boule tirée n'est pas noire", on a donc $\mathbb{P}(A) = 1/10$ et la réponse à la question est $\mathbb{P}(B|A) = \mathbb{P}(B \cap A)/\mathbb{P}(A) = \mathbb{P}(B)/\mathbb{P}(A) = 9/10$, donc une grande probabilité.

Puisqu'on peut calculer la probabilité "sachant A" de n'importe quel événement B de la tribu $\mathcal{F}$, une question naturelle est de se demander si $\mathbb{P}(.|A)$ est une probabilité sur $(\Omega, \mathcal{F})$: la réponse est oui. On vérifie en effet facilement les deux conditions sine qua non :

(i) $\mathbb{P}(\Omega|A) = \mathbb{P}(\Omega \cap A)/\mathbb{P}(A) = \mathbb{P}(A)/\mathbb{P}(A) = 1$;
(ii) σ-additivité : si $(B_n)_{n\geq 0}$ est une suite d'événements deux à deux incompatibles de $\mathcal{F}$, alors $(B_n \cap A)_{n\geq 0}$ est aussi une suite d'événements deux à deux incompatibles de $\mathcal{F}$, donc par σ-additivité de $\mathbb{P}$ on a :

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} B_n \middle| A\right) = \frac{\mathbb{P}\left(\bigcup_{n=0}^{+\infty}(B_n \cap A)\right)}{\mathbb{P}(A)} = \frac{\sum_{n=0}^{+\infty} \mathbb{P}(B_n \cap A)}{\mathbb{P}(A)} = \sum_{n=0}^{+\infty} \frac{\mathbb{P}(B_n \cap A)}{\mathbb{P}(A)} = \sum_{n=0}^{+\infty} \mathbb{P}(B_n|A).$$

Ainsi $\mathbb{P}(.|A)$ est une probabilité sur $(\Omega, \mathcal{F})$ et vérifie de fait toutes les propriétés vues précédemment (monotonie, additivité forte, sous-σ-additivité, continuités monotones croissante et décroissante).

Nous allons maintenant énoncer un résultat aussi simple qu'utile, mettant en jeu des conditionnements emboîtés.

Proposition 1.1 (Formule des probabilités composées)
Soit n événements $A_1, \dots, A_n$ tels que $\mathbb{P}(A_1 \cap \dots \cap A_{n-1}) > 0$, alors on a :

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

Preuve. On commence par noter que tous les conditionnements sont justifiés puisque par monotonie :

$$0 < \mathbb{P}(A_1 \cap \dots \cap A_{n-1}) \leq \mathbb{P}(A_1 \cap \dots \cap A_{n-2}) \leq \dots \leq \mathbb{P}(A_1 \cap A_2) \leq \mathbb{P}(A_1).$$

Il reste à remarquer qu'en développant les termes du produit via $\mathbb{P}(B|A) = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$, tous se télescopent sauf le dernier.

■

Remarque. On peut se servir de ce résultat comme d'une poupée russe : soit à calculer $\mathbb{P}(A_n)$, on introduit une séquence croissante d'événements $A_n \subset A_{n-1} \subset \dots \subset A_2 \subset A_1$ et la formule devient tout simplement :

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

Nous passons maintenant à la deuxième formule importante de cette section, dite des probabilités totales. Elle fait intervenir la notion de partition d'un ensemble, encore appelée système complet d'événements.

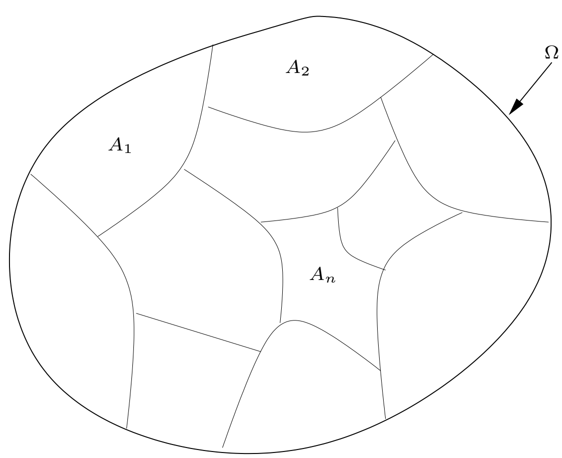


Figure 1.4 – Partition $(A_1, \dots, A_n)$ de Ω.

Définition 1.4 (Partition)
Soit Ω un ensemble et $(A_1, \dots, A_n)$ n sous-ensembles de Ω. On dit que $(A_1, \dots, A_n)$ forme une partition de Ω s'ils sont deux à deux disjoints et si on a :

$$A_1 \cup \dots \cup A_n = \Omega.$$

Bref il suffit de penser aux A_i comme aux pièces d'un puzzle Ω (voir figure 1.4). On va supposer dans la suite tous les $\mathbb{P}(A_i)$ strictement positifs, ce qui légitimera les conditionnements par les A_i. Disposant d'une partition de Ω, l'idée de la formule des probabilités totales est la suivante : si pour tout i on connaît $\mathbb{P}(B|A_i)$ et $\mathbb{P}(A_i)$, alors on peut en déduire $\mathbb{P}(B)$.

Proposition 1.2 (Formule des probabilités totales)
Soit $(\Omega, \mathcal{F}, \mathbb{P})$ *muni d'un système complet d'événements* $(A_1, \dots, A_n)$*, alors pour tout événement* B *on a la décomposition :*

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Preuve. On a tout d'abord d'un point de vue ensembliste (cf. figure 1.5) :

$$B = B \cap \Omega = B \cap (A_1 \cup \dots \cup A_n) = (B \cap A_1) \cup \dots \cup (B \cap A_n),$$

la dernière égalité venant de la distributivité de l'intersection par rapport à l'union (tout comme la multiplication par rapport à l'addition pour les nombres). Il suffit alors de remarquer que la dernière décomposition est une union d'événements deux à deux disjoints (car les A_i le sont), donc on peut appliquer la σ-additivité de $\mathbb{P}$:

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(B \cap A_i) = \sum_{i=1}^{n} \mathbb{P}(B|A_i)\mathbb{P}(A_i),$$

le dernier point venant de l'écriture : $\mathbb{P}(B \cap A_i) = \mathbb{P}(B|A_i)\mathbb{P}(A_i)$.

■

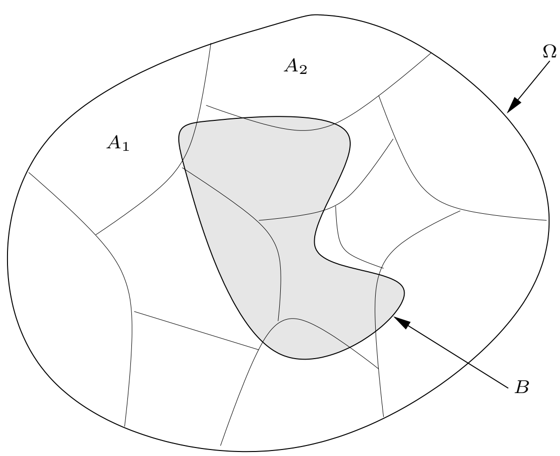


FIGURE 1.5 – Illustration de $B = \bigcup_{i=1}^{n}(B \cap A_i)$.

En pratique, on utilise très souvent cette formule des probabilités totales en conditionnant successivement par un événement et son contraire, c'est-à-dire en prenant tout simplement une partition de type $(A, \overline{A})$, ce qui donne :

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

Remarque. On peut élargir la définition d'une partition à une famille dénombrable $(A_n)_{n\geq 0}$ d'événements deux à deux incompatibles et dont l'union fait Ω (c'est-à-dire qu'il y a toujours exactement l'un des A_n qui se réalise). Dans ce cas la formule des probabilités totales fait intervenir une série :

$$\mathbb{P}(B) = \sum_{n=0}^{+\infty} \mathbb{P}(B|A_n)\mathbb{P}(A_n).$$

Tout est maintenant prêt pour la fameuse formule de Bayes, ou formule de probabilité des causes.

Proposition 1.3 (Formule de Bayes)
Soit $(\Omega, \mathcal{F}, \mathbb{P})$ muni d'une partition $(A_1, \ldots, A_n)$, alors pour tout événement B et pour tout indice j on a :

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{i=1}^n \mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Preuve. C'est l'âne qui trotte. Il suffit en effet d'écrire :

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B \cap A_j)}{\mathbb{P}(B)},$$

puis d'utiliser la décomposition $\mathbb{P}(B \cap A_j) = \mathbb{P}(B|A_j)\mathbb{P}(A_j)$ pour le numérateur et la formule des probabilités totales pour le dénominateur.

■

En pratique, lorsqu'on considère une partition de type $(A, \overline{A})$, cette formule devient :

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B|A)\mathbb{P}(A) + \mathbb{P}(B|\overline{A})\mathbb{P}(\overline{A})}.$$

Une application typique au problème de dépistage d'une maladie est donnée en exercice 1.22.

1.3 Indépendance

La notion d'indépendance intervient de façon constante en probabilités. Intuitivement, deux événements sont indépendants si la réalisation de l'un "n'a aucune influence" sur la réalisation ou non de l'autre. Le but de cette section est de préciser ceci mathématiquement et de l'étendre à plus de deux événements. Dans toute la suite, $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé fixé.

Définition 1.5 (Indépendance de 2 événements)
On dit que deux événements A et B sont indépendants si

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

Si A est tel que $\mathbb{P}(A) > 0$, l'indépendance de A et B s'écrit encore $\mathbb{P}(B|A) = \mathbb{P}(B)$ et on retrouve la notion intuitive d'indépendance : le fait que A se soit réalisé ne change rien quant à la probabilité que B se réalise.

Exemples :

1. On lance un dé deux fois de suite. Soit A l'événement : "Le premier lancer donne un nombre pair" et B l'événement : "Le second lancer donne un nombre pair". L'univers naturel est $\Omega = \{(i, j), 1 \leq i, j \leq 6\}$, ensemble à 36 éléments muni de l'équiprobabilité. Il est clair que $\mathbb{P}(A) = \mathbb{P}(B) = 18/36 = 1/2$ et que :
$$\mathbb{P}(A \cap B) = \frac{9}{36} = \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B),$$
donc A et B sont indépendants.
2. On tire une carte au hasard d'un jeu de 32 cartes. Soit A l'événement : "La carte tirée est un 7" et B l'événement : "La carte tirée est un pique". On a $\mathbb{P}(A) = 1/8$ et $\mathbb{P}(B) = 1/4$. $\mathbb{P}(A \cap B)$ correspond à la probabilité de tirer le 7 de pique donc $\mathbb{P}(A \cap B) = 1/32$. Ainsi on a $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, les événements A et B sont donc indépendants.

Achtung ! Ne pas confondre indépendants et incompatibles ! Deux événements peuvent être indépendants sans être incompatibles (cf. le 7 de pique ci-dessus) et incompatibles sans être indépendants (cf. A et $\overline{A}$ avec $0 < \mathbb{P}(A) < 1$).

Propriétés 1.2 (Indépendance et passage au complémentaire)
Si A et B sont indépendants, alors il en va de même pour :
- *les événements A et $\overline{B}$;*
- *les événements $\overline{A}$ et B ;*
- *les événements $\overline{A}$ et $\overline{B}$.*

Preuve. On montre uniquement le premier point, les autres se prouvant mutatis mutandis de la même façon :
$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A \setminus (A \cap B)) = \mathbb{P}(A) - \mathbb{P}(A \cap B),$$
et on applique maintenant l'indépendance de A et B :
$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)(1 - \mathbb{P}(B)) = \mathbb{P}(A)\mathbb{P}(\overline{B}),$$
ce qui prouve bien l'indépendance de A et $\overline{B}$.

$\blacksquare$

Lorsqu'on considère plus de deux événements simultanément, les choses se compliquent...

Définition 1.6 (Indépendance 2 à 2 & Indépendance mutuelle)
Soit $(A_n)_{n \geq 1}$ une suite d'événements. On dit qu'ils sont :
- *2 à 2 indépendants si pour tout couple (i, j) d'indices distincts, A_i et A_j sont indépendants ;*
- *mutuellement indépendants si pour tout ensemble fini d'indices distincts $(i_1, \ldots, i_k)$, on a*
$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1}) \times \cdots \times \mathbb{P}(A_{i_k}).$$

Exemple. Pour que 3 événements (A, B, C) soient :
- 2 à 2 indépendants, il faut que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$ et $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$;
- mutuellement indépendants, il faut que les 3 relations précédents soient vérifiées et de plus que $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

Il est clair que l'indépendance mutuelle implique l'indépendance 2 à 2 : il suffit de prendre $k = 2$, $i_1 = i$ et $i_2 = j$ pour s'en assurer. La réciproque est cependant fausse, comme le montre l'exemple suivant.

Exemple. On reprend l'exemple des deux lancers successifs d'un dé et on note C l'événement : "La somme des deux lancers est paire". On a donc $\mathbb{P}(C) = 1/2$. On vérifie que les événements (A, B, C) sont 2 à 2 indépendants, mais que :

$$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap B) = \frac{1}{4} \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \frac{1}{8}.$$

En pratique, ce sera l'indépendance mutuelle qui nous intéressera et c'est aussi celle que l'on rencontrera le plus souvent. Ainsi, quand on parlera d'une famille d'événements indépendants (sans plus de précisions), il faudra désormais comprendre **mutuellement indépendants**.

Remarques :

1. Soit une famille $(A_1, \ldots, A_n)$ de n événements, décrits d'une façon ou d'une autre. Supposons qu'on nous demande de prouver l'indépendance (mutuelle) de cette famille. Quel est le nombre N de relations que nous aurions à vérifier ? La réponse est vertigineuse :
$$N = \binom{n}{2} + \binom{n}{3} + \cdots + \binom{n}{n-1} + \binom{n}{n} = \left(\sum_{k=0}^{n} \binom{n}{k}\right) - \binom{n}{1} - \binom{n}{0} = 2^n - n - 1.$$
Rien que pour 10 événements, il y aurait déjà plus de 1000 relations à vérifier ! Ceci n'est bien sûr pas raisonnable. En fait, c'est le contexte qui dicte si l'on a affaire à une famille d'événements indépendants : c'est typiquement le cas lorsqu'on a une répétition d'épreuves (lancers successifs d'une pièce, etc.), le résultat de chacune d'entre elles n'ayant aucune espèce d'influence sur le résultat des autres.
2. La formule de Poincaré se simplifie grandement en cas d'événements indépendants. En effet, la probabilité qu'au moins l'un d'entre eux se réalise est toujours égale à
$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = 1 - \mathbb{P}(\overline{A_1 \cup \cdots \cup A_n}) = 1 - \mathbb{P}(\overline{A_1} \cap \cdots \cap \overline{A_1}),$$
et grâce à l'indépendance :
$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = 1 - \mathbb{P}(\overline{A_1}) \ldots \mathbb{P}(\overline{A_n}) = 1 - (1 - \mathbb{P}(A_1)) \ldots (1 - \mathbb{P}(A_n)) = 1 - (1 - p_1) \ldots (1 - p_n)$$
où $p_1, \ldots, p_n$ représentent les probabilités respectives de $A_1, \ldots, A_n$.

Exercice. On peut montrer que si (A, B, C) sont (mutuellement) indépendants, alors A est indépendant de tout événement formé à partir de B et de C. Prouvons par exemple que A est indépendant de $B \cup C$, c'est-à-dire que $\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}(A)\mathbb{P}(B \cup C)$. On a tout d'abord par distributivité de l'intersection par rapport à l'union :

$$\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}((A \cap B) \cup (A \cap C)),$$

suite à quoi on applique l'additivité forte :

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

et l'indépendance donne :

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

Il suffit alors de noter que par indépendance de B et C, on a $\mathbb{P}(B)\mathbb{P}(C) = \mathbb{P}(B \cap C)$, et d'appliquer la relation d'additivité forte pour obtenir :

$$\mathbb{P}(A \cap (B \cup C)) = \mathbb{P}(A)\mathbb{P}(B \cup C),$$

et la messe est dite.

1.4 Exercices

Exercice 1.1 (Welcome in Rennes 2)

1. Donner le nombre d'anagrammes du mot "laïus". Même question avec "lisier" et "charivari".
2. Généralisation : quel est le nombre de permutations possibles d'un ensemble à n éléments parmi lesquels il y a r paquets $(n_1, \ldots, n_r)$ d'éléments indistinguables entre eux ?
3. Parmi les 10 participants à un tournoi d'échecs, on compte 4 joueurs russes, 3 joueurs indiens, 2 joueurs israéliens et un joueur franco-lusitanien (José de Sousa). Dans le classement final du tournoi apparaît la nationalité du joueur, mais pas son nom. Combien de classements sont possibles ? Combien de classements sont possibles sachant que José est le vainqueur ?
4. Il y a 20 étudiants en Licence MASS 2. En fin de semestre, la moyenne générale de chacun est calculée : combien y a-t-il de classements possibles, en supposant que toutes les notes sont distinctes ?
5. On suppose qu'il y a 10 garçons et 10 filles dans cette classe et on décide de classer les garçons entre eux et les filles entre elles. Combien de classements globaux peut-on avoir ?

Exercice 1.2 (Autour des sommes géométriques)

1. Soit x un nombre réel ou complexe. Rappeler ce que vaut la somme $\sum_{j=0}^{n} x^j$.
2. On organise un tournoi de tennis, pour lequel 32 joueurs sont inscrits. Le tournoi s'effectue en seizièmes, huitièmes, quarts, demis et finale. Combien de matchs sont nécessaires pour désigner le vainqueur ?
3. Imaginons maintenant qu'on ait 32 sprinteurs dont on veut trouver le meilleur. On propose la procédure suivante : ils effectuent une première course et le dernier est éliminé du reste de la compétition, ils effectuent une deuxième course et à nouveau le dernier est éliminé, etc. Le vainqueur de la dernière course (à 2 coureurs, donc) est déclaré meilleur sprinteur. Combien de courses sont nécessaires pour désigner ce vainqueur ? Comparer au résultat de la question précédente.
4. On reprend le tournoi de tennis à 32 joueurs de la question initiale. Combien y a-t-il de déroulements possibles du tournoi, sachant que la place des joueurs sur la feuille de match est fixée ?

Exercice 1.3 (Le podium des MASS 2)

Dans ce qui suit, pour simplifier, on exclut les cas d'égalité de notes de deux étudiants en fin de semestre. On suppose de plus qu'il y a 20 étudiants en Licence MASS 2.

1. En fin de semestre, on récompense le major de chacune des 3 matières importantes du premier semestre (respectivement probabilités, analyse, algèbre) par un prix spécifique à chaque matière (respectivement une médaille d'or, un morceau de craie blanche, un morceau de craie jaune). Combien y a-t-il de triplets possibles (M_p, M_{an}, M_{al}) ?
2. On s'intéresse uniquement à l'épreuve reine du premier semestre (les probabilités) où seront décernées médailles d'or, d'argent et de bronze. Combien y a-t-il de podiums possibles ?
3. L'enseignant n'étant pas suffisamment rétribué, il ne peut offrir de médailles et décide donc de récompenser de la même façon les 3 premiers par un polycopié dédicacé. Combien y a-t-il de dédicaces possibles ?

Exercice 1.4 (Anniversaires)

1. Parmi les 20 étudiants en Licence MASS 2, quelle est la probabilité qu'au moins deux aient leur anniversaire le même jour (ignorer les années bissextiles) ? Quel effectif minimal faudrait-il dans la promotion pour que cette probabilité soit supérieure à 0.5 ? Que vaut cette probabilité pour $n = 50$?

2. Combien devrait-il y avoir d'étudiants en Licence MASS 2 pour qu'avec plus d'une chance sur deux, au moins un autre étudiant ait son anniversaire le même jour que vous ?

Exercice 1.5 (Las Vegas 21)
Un jeu de poker compte 52 cartes et on considère qu'une main est constituée de 5 cartes (poker fermé).

1. Combien y a-t-il de mains possibles ?
2. Quelle est la probabilité d'avoir une quinte flush ?
3. Quelle est la probabilité d'avoir une couleur (mais pas une quinte flush !) ?
4. Quelle est la probabilité d'avoir un carré ?
5. Que deviennent ces probabilités au poker ouvert (ou Texas Hold'em), c'est-à-dire lorsqu'il s'agit de former la meilleur main de 5 cartes parmi 7 ?

Exercice 1.6 (L'art de combiner les combinaisons)

1. Rappeler la formule du binôme de Newton pour $(x+y)^n$, où n est un entier naturel.
2. Dessiner le triangle de Pascal, qui permet de retrouver les valeurs des coefficients binomiaux pour les petites valeurs de n. Pour tout $0 \leq k < n$, simplifier l'expression $\binom{n}{k} + \binom{n}{k+1}$.
3. Calculer $\sum_{k=0}^{n} \binom{n}{k}$, $\sum_{k=0}^{n} (-1)^k \binom{n}{k}$, $\sum_{k=0}^{n} k\binom{n}{k}$, $\sum_{k=0}^{n} \binom{n}{k}/(k+1)$.
4. Calculer $\sum_{k=0}^{n} \binom{n}{k}^2$ en obtenant de deux façons le coefficient de X^n dans le polynôme :
$$P(X) = (1+X)^n(1+X)^n.$$

Exercice 1.7 (Formule de Poincaré)
Dans la suite, tous les ensembles sont finis et on note $\#A$ le cardinal d'un ensemble A.

1. Exprimer $\#(A \cup B)$ en fonction de $\#A$, $\#B$ et $\#(A \cap B)$. Application : dans une classe de lycée, 20 élèves ont pour langues (anglais,espagnol), 15 ont pour langues (anglais,allemand) et 5 étudient les 3 langues. Combien cette classe a-t-elle d'élèves ?
2. Exprimer $\#(A \cup B \cup C)$ en fonction de $\#A$, $\#B$, $\#C$, $\#(A \cap B)$, $\#(A \cap C)$, $\#(B \cap C)$ et $\#(A \cap B \cap C)$.
3. Généralisation : on considère n ensembles $A_1, \ldots, A_n$, on connaît les cardinaux de toutes les intersections possibles de ces ensembles, c'est-à-dire toutes les quantités de la forme
$$\forall k \in \{1, \ldots, n\}, \forall 1 \leq i_1 < \cdots < i_k \leq n, \qquad \#(A_{i_1} \cap \cdots \cap A_{i_k}).$$
Exprimer en fonction de ces quantités le cardinal $\#(A_1 \cup \cdots \cup A_n)$. Cette formule est connue sous le nom de formule de Poincaré, ou formule d'inclusion-exclusion ou encore formule du crible.

Exercice 1.8 (Dérangements)
Les n étudiants de MASS 2 font un repas de classe dans un restaurant et laissent leur manteau au vestiaire en arrivant. Au moment de partir, une panne d'électricité fait que l'employé rend à chacun l'un des manteaux au hasard. Le but de l'exercice est de déterminer la probabilité qu'aucun des étudiants ne récupère le sien. Les étudiants sont numérotés de 1 à n.

1. Combien y a-t-il de répartitions possibles des manteaux parmi les n étudiants ?
2. L'événement A_i signifie : "l'étudiant i a récupéré son manteau". Exprimer grâce aux A_i l'événement A : "aucun des étudiants ne récupère son manteau".
3. Soit $k \in \{1, \ldots, n\}$. Combien y a-t-il de séquences d'indices $(i_1, \ldots, i_k)$ telles que $1 \leq i_1 < \cdots < i_k \leq n$?

4. Que vaut le cardinal $\#(A_{i_1} \cap \cdots \cap A_{i_k})$?
5. Déduire de la formule de Poincaré que $\mathbb{P}(A) = 1 - \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!}$.
6. On peut montrer (cf. cours d'analyse) que pour tout réel x, $e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$. Montrer qu'il y a environ 37% de chances que ce soit le mardi gras absolu en fin de soirée.
7. On appelle dérangement d'un ensemble à n éléments une permutation de cet ensemble qui ne laisse aucun point fixe. Exprimer le nombre d_n de dérangements d'un tel ensemble.

Exercice 1.9 (Traductions ensemblistes d'événements)
Soit Ω un univers muni d'une tribu $\mathcal{F}$ et trois événements A, B et C de $\mathcal{F}$. On sait qu'on peut traduire les événements par des opérations sur les ensembles, par exemple l'événement "A et B se réalisent" s'écrit tout simplement "$A \cap B$". Grâce aux symboles d'union, d'intersection et de passage au complémentaire, déterminer des expressions pour les événements suivants :
- A seul se réalise ;
- A et C se réalisent mais pas B ;
- au moins l'un des trois événements se réalise ;
- au moins deux des trois événements se réalisent ;
- les trois événements se réalisent ;
- aucun ne se réalise ;
- au plus l'un des trois se réalise ;
- au plus deux des trois se réalisent ;
- exactement deux des trois se réalisent ;
- au plus trois se réalisent.

Exercice 1.10 (Exemple de tribu engendrée)
On se place dans l'ensemble $\mathbb{N}$. On considère la tribu $\mathcal{F}$ engendrée par les ensembles

$$S_n = \{n, n+1, n+2\} \text{ avec } n \in \{0, 2, 3, \ldots\}.$$

1. Montrer que pour tout $n \geq 2$, le singleton $\{n\}$ appartient à $\mathcal{F}$.
2. En déduire que toute partie de $\mathbb{N}^{**} = \{2, 3, \ldots\}$ est dans $\mathcal{F}$, autrement dit que $\mathcal{P}(\mathbb{N}^{**}) \subset \mathcal{F}$.
3. Caractériser alors simplement les éléments de $\mathcal{F}$.

Exercice 1.11 (Lancer infini d'une pièce)
On lance une pièce une infinité de fois. Pour tout $i \in \mathbb{N}^*$, on note :

$$A_i = \{\text{le } i\text{-ème lancer donne Pile}\}.$$

1. Décrire par une phrase chacun des événements suivants :

$$E_1 = \bigcap_{i=5}^{+\infty} A_i, \ E_2 = \left(\bigcap_{i=1}^{4} \overline{A_i}\right) \cap \left(\bigcap_{i=5}^{+\infty} A_i\right), \ E_3 = \bigcup_{i=5}^{+\infty} A_i$$

2. Ecrire à l'aide des A_i l'événement : "On obtient au moins une fois Pile après le n-ème lancer".
3. Ecrire à l'aide des A_i les événements
   (a) B_n : "On n'obtient plus que des Pile à partir du n-ème lancer."
   (b) B : "On n'obtient plus que des Pile à partir d'un certain lancer."

Exercice 1.12 (Inégalité de Bonferroni)
Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé.

1. Soit A et B deux événements. Montrer que la probabilité qu'un seul des deux événements se réalise est $\mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B)$.
2. Soit A et B deux événements tels que $\mathbb{P}(A) = 0,9$ et $\mathbb{P}(B) = 0,8$.
   (a) Grâce (par exemple) à l'additivité forte, montrer que $\mathbb{P}(A \cap B) \geq 0,7$.
   (b) Supposons qu'on tire un nombre entier au hasard dans l'ensemble $\Omega = \{1, \dots, 10\}$. Donner un exemple d'événements A et B tels que $\mathbb{P}(A) = 0,9$, $\mathbb{P}(B) = 0,8$ et $\mathbb{P}(A \cap B) = 0,7$.
   (c) Que vaut au maximum $\mathbb{P}(A \cap B)$? De façon générale, quand a-t-on égalité ? En reprenant l'exemple de tirage équiprobable entre 1 et 10, donner un exemple où il y a égalité.
3. Généralisation : soit $A_1, \dots, A_n$ des événements, utiliser la sous-σ-additivité et le passage au complémentaire pour prouver l'inégalité suivante :

$$\mathbb{P}(A_1 \cap \dots \cap A_n) \geq \sum_{i=1}^{n} \mathbb{P}(A_i) - (n-1).$$

   Que vaut au maximum $\mathbb{P}(A_1 \cap \dots \cap A_n)$? Dans quel(s) cas ce maximum est-il atteint ?

Exercice 1.13 (Alea jacta est)

1. On jette 2 dés équilibrés simultanément. Donner, pour tout $i \in \{2, \dots, 12\}$, la probabilité que la somme des résultats fasse i.
2. On répète maintenant l'expérience précédente jusqu'à ce qu'une somme de 5 ou 7 apparaisse. On désigne par E_n l'événement : "Une somme de 5 apparaît au n-ème double jet et sur les $(n-1)$ premiers coups ni la somme de 5 ni celle de 7 n'est apparue."
   (a) Calculer $\mathbb{P}(E_n)$.
   (b) Soit E : "Une somme de 5 apparaît au bout d'un certain nombre de lancers et sur les lancers précédents ni la somme de 5 ni celle de 7 n'est apparue." Décrire E en fonction des E_n et en déduire $\mathbb{P}(E)$.

Exercice 1.14 (Application de la sous-σ-additivité)

Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit $A_1, \dots, A_n$ des événements de $\mathcal{F}$ tels que :

$$\bigcup_{i=1}^{n} A_i = \Omega.$$

Grâce à la sous-σ-additivité, montrer que l'un au moins des événements A_i est de probabilité supérieure ou égale à $\frac{1}{n}$.

Exercice 1.15 (Limites supérieures et inférieures d'ensembles)

Soit $(A_n)_{n \geq 0}$ une suite de parties d'un ensemble Ω. On appelle limite supérieure des A_n et on note $\overline{\lim} A_n$, ou $\limsup_n A_n$, l'ensemble des éléments de Ω qui appartiennent à une infinité de A_n. On appelle limite inférieure des A_n et on note $\underline{\lim} A_n$, ou $\liminf_n A_n$ l'ensemble des éléments de Ω qui appartiennent à tous les A_n sauf à un nombre fini d'entre eux.

1. Soit A et B deux parties de Ω et la suite (A_n) définie par $A_0 = A_2 = \dots = A$ et $A_1 = A_3 = \dots = B$. Déterminer les limites sup et inf des A_n.
2. Ecrire les définitions de $\overline{\lim} A_n$ et $\underline{\lim} A_n$ à l'aide des quantificateurs logiques $\exists$ et $\forall$. Les traduire en termes ensemblistes à l'aide des symboles $\cup$ et $\cap$.
3. Déterminer $\overline{\lim} A_n$ et $\underline{\lim} A_n$ dans les situations suivantes :
   (a) $A_n =]-\infty, n]$ avec $n \geq 0$;

(b) $A_n =]-\infty, -n]$ avec $n \geq 0$;

(c) $A_n =]-1/n, 1/n[$ avec $n > 0$;

(d) $A_n =]-\infty, a_n]$, pour $n \geq 1$, avec :

$$\begin{cases} a_{2p+1} &= -1 - 1/(2p+1) \quad \forall p \geq 0 \\ a_{2p} &= 1 + 1/(2p) \quad \forall p > 0 \end{cases}$$

Exercice 1.16 (Lemme de Borel-Cantelli)
Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé. Soit $(A_n)_{n\geq 0}$ une suite d'éléments de $\mathcal{F}$ et $A = \overline{\lim} A_n$.

1. Par la caractérisation ensembliste de la limite sup, dire pourquoi A appartient à $\mathcal{F}$.
2. Considérons la suite d'ensembles $D_n = \bigcup_{k=n}^{+\infty} A_k$. Montrer qu'elle est décroissante.
3. On suppose que $\sum_{n=0}^{+\infty} \mathbb{P}(A_n) < +\infty$. Via la sous-$\sigma$-additivité, montrer que $\lim_{n\to+\infty} \mathbb{P}(D_n) = 0$.
4. Grâce à la continuité monotone décroissante, en déduire que $\mathbb{P}(A) = 0$. Traduire ce résultat concrètement.
   Remarque : Réciproquement, on montre que si les A_n sont des événements deux à deux indépendants et si $\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = +\infty$, alors $\mathbb{P}(\overline{\lim} A_n) = 1$.

Exercice 1.17 (Ensembles dénombrables)
On dit que E est dénombrable s'il est en bijection avec $\mathbb{N}$. Concrètement, E est dénombrable si on peut numéroter tous ses éléments, i.e. écrire $E = (u_0, u_1, \dots, u_n, \dots)$. Pour montrer qu'un ensemble est dénombrable, il suffit de pouvoir indiquer un procédé de numérotage qui n'oublie aucun élément de E. On parle de "au plus dénombrable" pour dire "fini ou dénombrable".

1. Montrer que l'ensemble $\mathbb{Z}$ des entiers relatifs est dénombrable.
2. Montrer que l'ensemble $\mathbb{Q}$ des nombres rationnels est dénombrable.
3. Montrer que $\mathbb{R}$ n'est pas dénombrable (procédé diagonal de Cantor).

Exercice 1.18 (L'oracle d'Oberhausen)
Lors de la Coupe du Monde de football 2010, avant chacune des 7 rencontres de l'équipe allemande (3 matchs de poule, huitième, quart, demi et "petite finale") ainsi qu'avant la finale (Espagne contre Pays-Bas), Paul le Poulpe avait le choix entre 2 récipients contenant sa nourriture préférée, chacun à l'effigie de l'un des deux adversaires. Le pronostic correspondait au choix du récipient où l'animal allait se nourrir. Il se trouve que les 8 pronostics se sont avérés exacts.

1. Quelle est la probabilité d'un pronostic correct pour un match de poule ? Et pour un match avec élimination directe ?
2. En déduire la probabilité qu'avait Paul le Poulpe de "tomber juste" sur l'ensemble des rencontres ?

Exercice 1.19 (Le poulpe démasqué)
La probabilité de gagner le gros lot au Loto est notée p (environ une chance sur 19 millions).

1. Quelle est la probabilité qu'aucune des N personnes jouant au Loto pour un tirage donné ne remporte le gros lot ?
2. En déduire le nombre de joueurs nécessaires pour qu'il y ait au moins une chance sur deux que le gros lot soit remporté.
3. Combien de "poulpes" (ou autres pronostiqueurs farfelus) étaient nécessaires pour qu'avec une probabilité supérieure à 90%, l'un au moins pronostique les 8 bons résultats ?

Exercice 1.20 (L'art de se tirer une balle dans le pied)
Cet exercice est tiré d'un article de Benjamin Dessus et Bernard Laponche, paru le 3 juin 2011 dans le quotidien *Libération* et intitulé “Accident nucléaire : une certitude statistique”. Au vu des données historiques, la probabilité d'un accident majeur par an pour un réacteur nucléaire est estimée à 3×10^{-4}, obtenue en considérant les 4 accidents majeurs (1 à Tchernobyl, 3 à Fukushima) survenus sur 450 réacteurs en 31 ans. Cette estimation est sujette à débat, mais passons.

1. Il y a 58 réacteurs en France (resp. 143 en Europe). En supposant l'indépendance entre ceux-ci, en déduire la probabilité d'au moins un accident majeur dans les 30 ans à venir en France (resp. en Europe).
2. Donner un équivalent de $1-(1-p)^{nt}$ lorsque p tend vers 0 et nt est fixé.
3. En déduire comment les auteurs en arrivent à écrire une phrase telle que : “Sur la base du constat des accidents majeurs survenus ces trente dernières années, la probabilité d'occurrence d'un accident majeur sur ces parcs serait donc de 50% pour la France et de plus de 100% pour l'Union européenne.”
4. Estimez la note que vous auriez à un contrôle de Probabilités en écrivant une telle phrase.

Exercice 1.21 (Probabilités composées)

1. On considère une urne contenant 4 boules blanches et 3 boules noires. On tire une à une et sans remise 3 boules de l'urne. Quelle est la probabilité que la première boule tirée soit blanche, la deuxième blanche et la troisième noire ?
2. On vous donne 5 cartes au hasard d'un jeu de 52. Quelle est la probabilité que vous ayez une couleur à Pique (i.e. 5 cartes de Pique) ? Quelle est la probabilité que vous ayez une couleur ?

Exercice 1.22 (Le problème du dépistage)

1. Soit $(\Omega,\mathcal{F},\mathbb{P})$ un espace probabilisé. Soit $(H_1,\dots,H_n)$ une partition de Ω en n événements de probabilités non nulles. Soit $A\in\mathcal{F}$ tel que $\mathbb{P}(A)>0$. Rappeler la formule de Bayes (encore appelée formule de probabilité des causes, les H_i étant les causes possibles et A la conséquence).
2. Application : Test de dépistage
   Une maladie est présente dans la population, dans la proportion d'une personne malade sur 1000. Un responsable d'un grand laboratoire pharmaceutique vient vous vanter son nouveau test de dépistage : si une personne est malade, le test est positif à 99%. Néanmoins, sur une personne non malade, le test est positif à 0.2%. Calculer la probabilité qu'une personne soit réellement malade lorsque son test est positif. Qu'en pensez-vous ?

Exercice 1.23 (Composition de familles)
Une population est composée de familles de 0, 1, 2 ou 3 enfants. Il y a une famille sans enfant pour 3 de 1 enfant, 4 de 2 enfants et 2 de 3 enfants. On suppose que les deux sexes sont équiprobables et qu'ils sont indépendants pour deux enfants différents.

1. Donner les probabilités de nombres d'enfants par famille p_0,p_1,p_2,p_3 .
2. On choisit une famille au hasard : quelle est la probabilité qu'il n'y ait aucun garçon ?
3. Toujours pour une famille choisie au hasard, quelle est la probabilité qu'elle ait 2 enfants sachant qu'elle n'a aucun garçon ?

Exercice 1.24 (L'ivresse du gardien de nuit)
Un gardien de nuit a 10 clés, dont une seule marche, pour ouvrir une porte. Il emploie deux méthodes. Méthode A : à jeun, il retire du trousseau les clés déjà essayées ; méthode B : ivre, il remet la clé dans le trousseau après chaque essai.

1. Méthode A : on appelle p_n la probabilité qu'il faille n essais pour ouvrir la porte. Déterminer p_n.

2. Méthode B : on appelle q_n la probabilité qu'il faille n essais pour ouvrir la porte. Déterminer q_n.
3. Le gardien est ivre un jour sur trois. Un jour, après avoir essayé 8 clés, le gardien n'a toujours pas ouvert la porte. Quelle est la probabilité qu'il soit ivre ?

Exercice 1.25 (Urne de Polya)
Une urne contient 4 boules blanches et 6 boules noires. Une boule est tirée au hasard puis on la replace dans l'urne ainsi que 3 autres boules de la même couleur que celle-ci (de sorte qu'il y a alors 13 boules dans l'urne). On tire alors une nouvelle boule au hasard dans l'urne.

1. Calculer la probabilité que la seconde boule tirée soit blanche.
2. Etant donné que la seconde boule tirée est blanche, quelle est la probabilité que la première soit noire ?
3. Généralisation : on considère le même procédé avec initialement B boules blanches, N noires et un ajout de x boules supplémentaires (ainsi précédemment on avait $B = 4$, $N = 6$ et $x = 3$). Montrer que la probabilité que la seconde boule tirée soit blanche est $\frac{B}{B+N}$.

Exercice 1.26 (Transmission bruitée)
Un message doit être transmis d'un point à un autre à travers N canaux successifs. Ce message peut prendre deux valeurs, 0 ou 1. Durant le passage par un canal, le message a la probabilité $p \in]0,1[$ d'être bruité, c'est-à-dire d'être transformé en son contraire, et $(1-p)$ d'être transmis fidèlement. Les canaux se comportent indépendamment les uns des autres.

1. Notons I_n l'événement : "en sortie de n-ème canal, le message est le même que celui transmis initialement." Exprimer $\mathbb{P}(I_{n+1})$ en fonction de $\mathbb{P}(I_n)$ et de p.
2. En notant $p_n = \mathbb{P}(I_n)$, donner une relation de récurrence entre p_{n+1} et p_n. Que vaut p_1 ?
3. On considère une suite $(u_n)_{n\geq 1}$ vérifiant la relation de récurrence :
$$u_{n+1} = (1-2p)u_n + p.$$
Une telle suite est dite arithmético-géométrique. Vérifier que la suite $(v_n)_{n\geq 1}$, définie par $v_n = u_n - \frac{1}{2}$, est géométrique. En déduire v_n en fonction de p et v_1.
4. En déduire p_n en fonction de p pour tout $n \in \{1,\dots,N\}$.
5. Que vaut $\lim_{N\to+\infty} p_N$? Qu'est-ce que ce résultat a d'étonnant à première vue ?

Exercice 1.27 (La roulette de la lose)
Deux joueurs A et B jouent une succession de parties de pile ou face. A chaque coup, A a la probabilité $p \in]0,1[$ de gagner, auquel cas B lui donne 1€, sinon le contraire. Les joueurs A et B disposent en début de partie de 50€ chacun. La partie s'arrête lorsque l'un des deux est ruiné. On cherche la probabilité que A finisse ruiné. Pour tout $n \in \{0,\dots,100\}$, on note p_n la probabilité que A finisse ruiné s'il commence avec n€ et B avec $(100-n)$€.

1. Que valent p_0 et p_{100} ?
2. Notons R_n l'événement : "A finit ruiné en commençant avec n€", c'est-à-dire que $p_n = \mathbb{P}(R_n)$. Décomposer $\mathbb{P}(R_n)$ en conditionnant par le résultat de la première partie, de façon à obtenir une relation de récurrence entre p_{n+1}, p_n et p_{n-1}.
3. On admet que la solution de cette équation est de la forme :
$$p_n = \alpha + \beta \left(\frac{1-p}{p}\right)^n.$$
Déterminer α et β.

4. En déduire la probabilité que A finisse ruiné.
5. De passage à Dinard, vous rentrez au casino et jouez à la roulette : il y a 18 numéros rouges, 18 numéros noirs et 1 numéro vert, le zéro. Vous jouez rouge pour 1€ à chaque fois. Vous commencez avec 50€ et vous arrêtez si vous avez 100€ ou si vous êtes ruiné. Pourquoi valait-il mieux aller baguenauder sur les sentiers côtiers ce jour-là ?
6. Sachant que vous commencez avec 50€ et que vous ne partirez que ruiné ou avec 100€ en poche, quelle tactique vaut-il mieux adapter pour maximiser vos chances de succès ?

Exercice 1.28 (Loi de succession de Laplace)
On dispose de $(N+1)$ urnes, numérotées de 0 à N. L'urne k contient k boules rouges et $(N-k)$ boules blanches. On choisit une urne au hasard. Sans connaître son numéro, on en tire n fois de suite une boule, avec remise après chaque tirage.

1. Quelle est la probabilité que le tirage suivant donne encore une boule rouge sachant que, au cours des n premiers tirages, seules des boules rouges ont été tirées ? Indication : on pourra noter E_n (respectivement E_{n+1}) le fait de tirer n (respectivement $(n+1)$) boules rouges à la suite et décomposer ces deux événements sur la partition $(U_0,\dots,U_N)$ formée par les urnes.
2. Calculer la limite de cette probabilité lorsque N tend vers l'infini. (Rappel sur les sommes de Riemann : si f est continue sur $[0,1]$, alors $\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n} f(k/n)=\int_0^1 f(x)dx$.)

Exercice 1.29 (Il Padrino)

1. On considère deux événements A et B tels que $\mathbb{P}(A)=0.1$, $\mathbb{P}(B)=0.9$ et $\mathbb{P}(A\cup B)=0.91$. A et B sont-ils indépendants ?
2. La Mafia subtilise 10% des colis expédiés de New York par avion. Alice veut envoyer deux cadeaux de Noël à son ami Bob. Elle peut faire soit deux paquets séparés indépendants, soit un paquet groupé. Calculer dans les deux cas les probabilités des événements suivants :
   (a) Un cadeau au moins est bien arrivé.
   (b) Les deux cadeaux sont bien arrivés.
3. On considère trois événements (mutuellement) indépendants A, B et C tels que $\mathbb{P}(A)=0.8$, $\mathbb{P}(B)=0.5$ et $\mathbb{P}(C)=0.2$. Que vaut $\mathbb{P}(A\cup B\cup C)$?

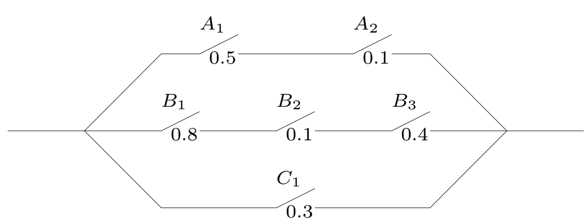


FIGURE 1.6 – Un circuit électrique aléatoire.

Exercice 1.30 (Circuit électrique)
On considère le circuit électrique de la figure 1.6. Chaque relais est en position ouverte ou fermée, la probabilité qu'il soit ouvert étant indiquée sur la figure et les relais se comportant de façon totalement indépendante. Quelle est la probabilité que le courant passe, c'est-à-dire qu'il existe au moins une branche sur laquelle tous les relais sont fermés ?

Exercice 1.31 (Le bandit manchot)
Une machine à sous a trois roues indépendantes, chacune ayant 20 symboles apparaissant de façon équiprobable lorsqu'elle s'arrête de tourner. Les roues de droite et de gauche sont identiques, avec seulement une cloche sur les 20 symboles. La roue du centre est différente et compte 9 cloches.

1. Quelle est la probabilité de remporter le jackpot (3 cloches) ?
2. Calculer la probabilité d'obtenir 2 cloches, mais pas le jackpot.
3. Si au lieu d'une répartition 1-9-1 des cloches, il y a une répartition 3-1-3, que deviennent les résultats des questions précédentes ? Expliquer pourquoi le propriétaire du casino optera plutôt pour la répartition 1-9-1 que 3-1-3.

Exercice 1.32 (Les affres des escales)
Vous voyagez en avion de Los Angeles à Paris avec deux escales, à New York puis à Londres. La probabilité p que votre bagage ne soit pas mis en soute est la même à Los Angeles, New York et Londres. Arrivé à Paris, vous constatez l'absence de votre valise. Calculez les probabilités que celle-ci soit restée à Los Angeles, New York et Londres respectivement.

Exercice 1.33 (Une histoire de montres)
Un lot de montres identiques est reçu par un détaillant parisien. Celui-ci provient de façon équiprobable soit de Hong-Kong, soit de Singapour. L'usine de Hong-Kong produit un article défectueux sur 1000 en moyenne, celle de Singapour un sur 200. Le détaillant inspecte une première montre : elle marche. Sachant ceci, quelle est la probabilité que la deuxième montre inspectée marche elle aussi ?

Exercice 1.34 (Un éléphant ça trompe énormément)
Trois touristes tirent en même temps sur un éléphant au cours d'un safari. On estime la valeur d'un chasseur par sa probabilité d'atteindre la cible en un coup. Ces probabilités sont respectivement 1/4, 1/2 et 3/4. La bête meurt frappée par deux balles. Trouvez pour chacun des chasseurs la probabilité d'avoir raté l'éléphant.

Exercice 1.35 (Une urne à composition variable)
Une urne contient n boules blanches ($n \geq 5$) et 10 boules noires. On tire au hasard et simultanément 10 boules de l'urne.

1. Calculer la probabilité p_n que l'on ait tiré exactement 5 boules noires.
2. Montrer que pour tout $n \geq 5$, on a :
$$\frac{p_{n+1}}{p_n} = \frac{n^2+2n+1}{n^2+7n-44}.$$
3. En déduire les variations de la suite $(p_n)_{n\geq 5}$ et la valeur de n pour laquelle p_n est maximale.

Exercice 1.36 (Les paris plus ou moins vaseux du Chevalier de Méré)
Le Chevalier de Méré était, à la cour de Louis XIV, un joueur impénitent. Il pensait en particulier avoir trouvé deux règles pour gagner de l'argent.

1. Première règle : "Il est avantageux de parier sur l'apparition d'au moins un 6 en lançant un dé quatre fois de suite". Démontrer que c'est vrai.
2. Seconde règle : "Il est avantageux de parier sur l'apparition d'au moins un double 6 en lançant deux dés vingt-quatre fois de suite". Démontrer que c'est faux. Remarque : c'est Blaise Pascal qui lui a prouvé son erreur, les probabilités étaient nées...

Exercice 1.37 (Tirages uniformes sur un segment)
On tire un point au hasard sur le segment $[0, 1]$.

1. Quelle est la probabilité qu'il soit supérieur à $3/4$?
2. Quelle est la probabilité qu'il soit supérieur à $3/4$, sachant qu'il est supérieur à $1/3$?
3. On tire deux points au hasard sur le segment $[0, 1]$, indépendamment l'un de l'autre.

(a) Quelle est la probabilité que le plus petit des deux nombres soit supérieur à $1/3$?

(b) Quelle est la probabilité que le plus grand des deux nombres soit supérieur à $3/4$, sachant que le plus petit des deux est supérieur à $1/3$?

Exercice 1.38 (La loi du minimum)
On considère une urne contenant n jetons numérotés de 1 à n. On tire successivement N fois un jeton, avec remise entre les tirages, et on note le numéro à chaque fois. Soit k un entier naturel fixé entre 1 et n.

1. Quelle est la probabilité P_k que le plus petit des numéros obtenus soit supérieur ou égal à k ?
2. En déduire la probabilité p_k que le plus petit des numéros obtenus soit égal à k.
3. On suppose maintenant $N \leq n$. Que deviennent ces résultats si on ne fait pas de remise entre les N tirages ?

Exercice 1.39 (Fratrie)
Dans cet exercice, on considère qu'à la naissance un enfant a autant de chances d'être une fille qu'un garçon, et ce indépendamment de ses éventuels frères et sœurs.

1. Raoul vient d'une famille de deux enfants. Quelle est la probabilité que l'autre soit une sœur ?
2. Un couple a deux enfants. Quelle est la probabilité que les deux soient des filles sachant que l'aînée en est une ?

Exercice 1.40 (Liouville et les probabilités)
Une urne contient 2 boules rouges et 3 boules noires. Le joueur A commence et gagne s'il tire une boule rouge, sinon c'est à B de tirer (A n'a pas remis la boule rouge dans l'urne). B gagne s'il tire une boule noire, sinon c'est à A de tirer, et ainsi de suite. Quelle est la probabilité que A gagne ? Ce jeu est-il équitable ?

Exercice 1.41 (Pierre-feuille-ciseaux)
On considère ici trois dés à 6 faces un peu particuliers. Le dé A a pour faces $(3, 3, 3, 3, 3, 6)$, le dé B $(2, 2, 2, 5, 5, 5)$, et le dé C $(1, 4, 4, 4, 4, 4)$.

1. Vous lancez simultanément les dés A et B. Quelle est la probabilité que A batte B ?
2. Quelle est la probabilité que B batte C ?
3. Sachant ces résultats, on vous propose de choisir entre le dé A et le dé C pour un nouveau duel. Lequel choisiriez-vous intuitivement ? Que donne le calcul des questions précédentes dans ce cas ?

Exercice 1.42 (Match de tennis)
Dans un match donné, sur son service, un joueur a deux chances sur trois de gagner le point.

1. Calculer la probabilité qu'il a de gagner le jeu sachant qu'il est à 40-40 sur son service (Indication : noter P cette probabilité, P^+ celle de gagner le jeu s'il a l'avantage, P_- celle de gagner le jeu si son adversaire a l'avantage, écrire un système de 3 équations pour les 3 inconnues P_-, P, P^+ et résoudre ce système).
2. Quelle est la probabilité d'arriver à 40-40 ?
3. Quelle est la probabilité que le joueur gagne le jeu en arrivant à 40-30 et en concluant ? en arrivant à 40-15 et en concluant ? en arrivant à 40-0 et en concluant ?
4. Déduire des questions précédentes la probabilité que le joueur gagne le jeu ?
5. Généraliser le résultat précédent en considérant qu'il a une probabilité p de gagner le point sur son service.

Exercice 1.43 (Let's make a deal)
Vous participez à un jeu où l'on vous propose trois portes au choix. L'une des portes cache une voiture à gagner, et chacune des deux autres une chèvre. Vous choisissez une porte, mais sans l'ouvrir ! L'animateur, qui sait où est la voiture, ouvre une autre porte, derrière laquelle se trouve une chèvre. Il vous donne maintenant le choix entre : vous en tenir à votre choix initial, ou changer de porte. Qu'avez-vous intérêt à faire ? Remarque : C'est un problème auquel étaient confrontés les invités du jeu télévisé *Let's make a deal* de Monty Hall (animateur et producteur américain). Il a par ailleurs fait l'objet d'un débat houleux aux Etats-Unis.

Exercice 1.44 (Newton & Galilée)

1. Samuel Pepys écrivit un jour à Isaac Newton : "Qu'est-ce qui est le plus probable : au moins un 6 lorsqu'on lance 6 fois un dé, ou au moins deux 6 lorsqu'on lance 12 fois un dé ?" Calculer les probabilités de ces deux événements.
2. À l'époque de Galilée, on croyait que lorsque 3 dés équilibrés étaient lancés et leurs résultats ajoutés, une somme de 9 avait la même probabilité d'apparaître qu'une somme de 10, puisqu'elles pouvaient chacune être obtenues de 6 façons :
   – pour 9 : 1+2+6, 1+3+5, 1+4+4, 2+2+5, 2+3+4, 3+3+3 ;
   – pour 10 : 1+3+6, 1+4+5, 2+2+6, 2+3+5, 2+4+4, 3+3+4.
   Calculer les probabilités de chacun de ces deux événements pour montrer qu'une somme de 10 est plus probable qu'une somme de 9.

Exercice 1.45 (Peer-to-Peer)
Un logiciel Peer-to-Peer utilise 4 serveurs S_1, S_2, S_3, S_4 de listes de fichiers partagés. S_4 est le plus gros des serveurs et recense 40% des données disponibles. Les données restantes sont distribuées équitablement entre les 3 autres serveurs. Sur la masse des fichiers disponibles, un certain nombre d'entres eux sont défectueux, soit que leur contenu n'est pas conforme à la description qui en est donnée, soit qu'ils contiennent des virus. Les pourcentages de fichiers défectueux sont : 8% pour S_4, 6% pour S_3, 2% pour S_2 et 2% pour S_1.

1. On télécharge un fichier. Quelle est la probabilité que ce fichier soit défectueux ?
2. Sachant que le fichier est défectueux, quelle est la probabilité qu'il provienne du serveur S_4 ?

Exercice 1.46 (Hémophilie)
La reine porte le gène de l'hémophilie avec une probabilité de 0,5. Si elle est porteuse, chaque prince aura une chance sur deux de souffrir de cette maladie, indépendamment l'un de l'autre. Si elle ne l'est pas, aucun prince ne souffrira.

1. Supposons que la reine ait un seul fils. Quelle est la probabilité qu'il soit hémophile ?
2. Supposons maintenant que la reine a eu un seul fils et que celui-ci n'est pas hémophile. Quelle est la probabilité qu'elle soit porteuse du gène ?
3. Toujours en supposant que la reine a eu un fils non hémophile, s'il naît un deuxième prince, avec quelle probabilité sera-t-il hémophile ?

Exercice 1.47 (Dénombrements en vrac)

1. Les initiales de Andrei Kolmogorov sont A.K. Combien y a-t-il d'initiales possibles en tout (on exclut les prénoms et noms composés) ? Combien au minimum un village doit-il avoir d'habitants pour qu'on soit sûr que deux personnes au moins aient les mêmes initiales ?
2. Lors d'une course hippique, 12 chevaux prennent le départ. Donner le nombre de tiercés dans l'ordre (un tiercé dans l'ordre est la donnée du premier, du deuxième et du troisième cheval arrivés, dans cet ordre).
3. Dans un jeu de 32 cartes, on a remplacé une carte autre que la dame de cœur par une seconde dame de cœur. Une personne tire au hasard 3 cartes simultanément. Quelle est la probabilité qu'elle s'aperçoive de la supercherie ?

Exercice 1.48 (Urnes, cartes et dés)

1. Deux urnes contiennent chacune initialement 2 boules noires et 3 boules blanches. On tire au hasard une boule de la première urne, on note sa couleur et on la remet dans la seconde urne. On tire alors au hasard une boule de la seconde urne. Quelle est la probabilité d'obtenir deux fois une boule noire ?
2. Une population possède une proportion $p \in]0,1[$ de tricheurs. Lorsqu'on fait tirer une carte d'un jeu de 52 cartes à un tricheur, il est sûr de retourner un as. Exprimer en fonction de p la probabilité qu'un individu choisi au hasard dans la population retourne un as.
3. On prend un dé au hasard parmi un lot de 100 dés dont 25 sont pipés. Pour un dé pipé, la probabilité d'obtenir 6 est 1/2. On lance le dé choisi et on obtient 6.
   (a) Quelle est la probabilité que ce dé soit pipé ?
   (b) On relance alors ce dé et on obtient à nouveau 6. Quelle est la probabilité que ce dé soit pipé ?
   (c) Généralisation : on lance n fois le dé et à chaque fois on obtient 6. Quelle est la probabilité p_n que ce dé soit pipé ? Que vaut $\lim_{n\to\infty} p_n$? Commenter ce résultat.

Exercice 1.49 (Evénements indépendants)
On considère deux événements indépendants A et B de probabilités respectives 1/4 et 1/3. Calculer :

1. la probabilité que les deux événements aient lieu.
2. la probabilité que l'un au moins des deux événements ait lieu.
3. la probabilité qu'exactement l'un des deux événements ait lieu.

Exercice 1.50 (Un tirage en deux temps)
Une boîte contient une balle noire et une balle blanche. Une balle est tirée au hasard dans la boîte : on remet celle-ci ainsi qu'une nouvelle balle de la même couleur. On tire alors une des trois balles au hasard dans la boîte.

1. Quelle est la probabilité que la seconde balle tirée soit blanche ?
2. Quelle est la probabilité que l'une au moins des deux balles tirées soit blanche ?
3. Quelle est la probabilité que la première balle tirée soit blanche, sachant que l'une au moins des deux balles tirées est blanche ?

Exercice 1.51 (Pièces défectueuses)
Une usine produit des objets par boîtes de deux. Sur le long terme, on a constaté que : 92% des boîtes ne contiennent aucun objet défectueux ; 5% des boîtes contiennent exactement 1 objet défectueux ; 3% des boîtes contiennent 2 objets défectueux. Une boîte est choisie au hasard sur la chaîne de production et on tire au hasard un des deux objets de cette boîte.

1. Quelle est la probabilité que cet objet soit défectueux ?
2. Sachant que cet objet est effectivement défectueux, quelle est la probabilité que l'autre objet de la boîte le soit aussi ?

Exercice 1.52 (Circuits intégrés)
Un atelier reçoit 5000 circuits intégrés : 1000 en provenance de l'usine A et 4000 en provenance de l'usine B. 10% des circuits fabriqués par l'usine A et 5% de ceux fabriqués par l'usine B sont défectueux.

1. On choisit au hasard un circuit intégré à l'atelier. Quelle est la probabilité qu'il soit défectueux ?
2. Sachant qu'un circuit choisi est défectueux, quelle est la probabilité qu'il vienne de l'usine A ?

Exercice 1.53 (Utilité d'un testeur)
Une chaîne de montage d'ordinateurs utilise un lot de processeurs contenant 2% d'éléments défectueux. En début de chaîne, chaque processeur est vérifié par un testeur dont la fiabilité n'est pas parfaite, de telle sorte que la probabilité que le testeur déclare le processeur bon (resp. mauvais) sachant que le processeur est réellement bon (resp. mauvais) vaut 0.95 (resp. 0.94).

1. Calculer la probabilité qu'un processeur soit déclaré bon.
2. Calculer la probabilité qu'un processeur déclaré bon soit réellement bon.
3. Calculer la probabilité qu'un processeur déclaré mauvais soit réellement mauvais.
4. Le testeur est-il utile ?

1.5 Corrigés

Exercice 1.1 (Welcome in Rennes 2)

1. Nombre d'anagrammes du mot "laïus" : $5\,!$=120. Nombre d'anagrammes du mot "lisier" : $6\,!/2\,!$=360. Nombre d'anagrammes du mot "charivari" : $9\,!/(2\,!2\,!2\,!)$=45360.
2. Nombre de permutations possibles d'un ensemble à n éléments parmi lesquels il y a r paquets $(n_1, \ldots, n_r)$ d'éléments indistinguables entre eux :
$$\frac{n!}{n_1! \ldots n_r!}.$$
3. Nombre de classements possibles :
$$\frac{10!}{4!3!2!} = 12600.$$
Nombre de classements possibles sachant que José est le vainqueur :
$$\frac{9!}{4!3!2!} = 1260.$$
4. Nombre de classements possibles : $20! \approx 2,433 \cdot 10^{18}$. Rappelons la formule de Stirling :
$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$
Pour $n = 20$, elle donne $2,423 \cdot 10^{18}$, soit une erreur relative de l'ordre de $0,4\%$.
5. Nombre de classements globaux : $(10!)^2 \approx 1,317 \cdot 10^{13}$.

Exercice 1.2 (Autour des sommes géométriques)

1. Pour l'expression de la somme $S_n = \sum_{j=0}^{n} x^j$, il faut différencier deux cas :
   – si $x = 1$, c'est une somme de 1 et elle vaut tout simplement : $S_n = n + 1$.
   – si $x \neq 1$, c'est la somme des termes d'une suite géométrique de raison x et elle vaut de façon générale :
   $$S_n = \frac{\text{premier terme écrit - premier terme non écrit}}{\text{1-raison}},$$
   ce qui donne ici : $S_n = \frac{1-x^{n+1}}{1-x}$.
2. Il y a 16 seizièmes, 8 huitièmes, 4 quarts, 2 demis et 1 finale, donc le nombre de matchs nécessaires pour désigner le vainqueur est leur somme, soit $16 + 8 + 4 + 2 + 1 = 31$.

3. On élimine 1 sprinteur à chaque course et il faut tous les éliminer sauf un, donc il est clair qu'il faut faire 31 courses. Dans la question précédente, puisque chaque match de tennis éliminait exactement un joueur et qu'on voulait tous les éliminer sauf un, c'était exactement la même chose.
4. On reprend le tournoi de tennis à 32 joueurs de la question initiale. Le nombre S de déroulements possibles du tournoi est

$$S = 2^{16}2^{8}2^{4}2^{2}2^{1} = 2^{31}.$$

Exercice 1.3 (Le podium des MASS 2)

1. Le nombre de triplets possibles (M_p, M_{an}, M_{al}) est $20^3 = 8000$.
2. Le nombre de podiums possibles est le nombre d'arrangements de 3 éléments dans un ensemble à 20 éléments, soit $A_{20}^3 = 6840$.
3. Le nombre de dédicaces possibles est le nombre de combinaisons de 3 éléments dans un ensemble à 20 éléments, soit $\binom{20}{3} = 1140$.

Exercice 1.4 (Anniversaires)

1. Pour simplifier, on considère des années à 365 jours. Les élèves étant considérés comme distinguables, le nombre de 20-uplets d'anniversaires est 365^{20}. Pour calculer la probabilité cherchée, on utilise la ruse classique du passage à l'événement complémentaire, c'est-à-dire qu'on cherche le nombre de 20-uplets d'anniversaires tels que toutes les dates soient distinctes. Il y en a A_{365}^{20}. La probabilité cherchée vaut donc

$$p_{20} = 1 - \frac{A_{365}^{20}}{365^{20}} = 1 - \left(1 - \frac{1}{365}\right)\dots\left(1 - \frac{19}{365}\right) \approx 0,411.$$

Pour que cette probabilité soit supérieure à 0.5, il suffit d'avoir au moins 23 étudiants, puisque $p_{23} \approx 0,507$ tandis que $p_{22} \approx 0,476$. Contrairement à ce qu'une première intuition pourrait laisser croire, dans une assemblée de 50 personnes il y a de très fortes chances que deux personnes aient le même jour d'anniversaire puisque $p_{50} \approx 97\%$. En fait la probabilité est même encore plus grande car la répartition des naissances n'est pas uniforme sur l'année.

La suite (p_n) est représentée par les symboles '+' sur la figure 1.7. Remarquons au passage qu'on peut obtenir facilement une approximation de p_n pour n petit devant 365 puisque :

$$1 - p_n = \frac{A_{365}^{n}}{365^{n}} = \left(1 - \frac{1}{365}\right)\dots\left(1 - \frac{n-1}{365}\right),$$

donc en passant aux logarithmes :

$$\ln(1 - p_n) = \ln\left(1 - \frac{1}{365}\right) + \dots + \ln\left(1 - \frac{n-1}{365}\right),$$

et via l'approximation $\ln(1-u) \approx -u$ au voisinage de 0, on arrive à :

$$\ln(1 - p_n) \approx -\frac{1}{365} - \dots - \frac{n-1}{365} = -\frac{1}{365}\sum_{k=1}^{n-1} k,$$

où l'on reconnaît la somme des termes d'une suite arithmétique :

$$\ln(1 - p_n) \approx -\frac{n(n-1)}{730},$$

c'est-à-dire :

$$p_n \approx 1 - e^{-\frac{n(n-1)}{730}}.$$

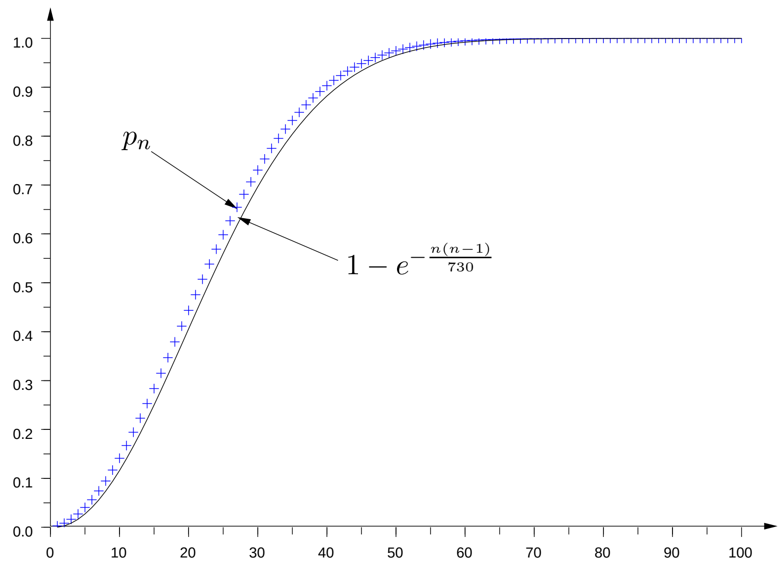


FIGURE 1.7 – Probabilités $p_n \approx 1 - e^{-\frac{n(n-1)}{730}}$ que deux personnes parmi n soient nées le même jour.

En d'autres termes, la suite représentée figure 1.7 est quasiment la version discrétisée de la fonction

$$f : \begin{cases} [0,100] & \to [0,1] \\ x & \mapsto 1 - e^{-\frac{x(x-1)}{730}} \end{cases}$$

2. Soit n le nombre d'étudiants qu'il faudrait pour qu'avec plus d'une chance sur deux, au moins un autre étudiant ait son anniversaire le même jour que vous. Notons A cet événement et calculons la probabilité de $\overline{A}$, c'est-à-dire la probabilité qu'aucun des $(n-1)$ autres étudiants n'ait la même date d'anniversaire que vous. Chaque autre étudiant a donc le choix entre 364 jours dans l'année, ce qui donne :

$$\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - \left(\frac{364}{365}\right)^{n-1}.$$

Nous cherchons donc n tel que cette probabilité soit supérieure à $1/2$, ce qui donne :

$$1 - \left(\frac{364}{365}\right)^{n-1} \geq \frac{1}{2} \Leftrightarrow \left(\frac{364}{365}\right)^{n-1} \leq \frac{1}{2}$$

et en passant aux logarithmes :

$$(n-1)\ln\frac{364}{365} \leq -l\ln 2 \Leftrightarrow n \geq 1 - \frac{\ln 2}{\ln\frac{364}{365}} \approx 253.65$$

Donc il faut au moins 254 étudiants.

Exercice 1.5 (Las Vegas 21)

1. Nombre de mains possibles : $N = \binom{52}{5}$.
2. Soit p_1 la probabilité d'avoir une quinte flush. Pour former une quinte flush, il suffit de choisir la couleur (4 choix) et la valeur de la carte basse (10 choix), donc au total 40 possibilités, soit $p_1 = 40/N \approx 0,00154\%$ de chances.

3. Soit p_2 la probabilité d'avoir une couleur (mais pas une quinte flush). On a 4 choix pour la couleur, puis $\binom{13}{5}$ possibilités pour choisir les 5 cartes à l'intérieur de cette couleur. Pour calculer p_2 il suffit alors d'enlever la probabilité d'avoir une quinte flush :

$$p_2 = \frac{4\binom{13}{5}}{N} - p_1 \approx 0,196\%$$

4. Soit p_3 la probabilité d'avoir un carré. Pour former un carré, on a 13 choix pour la hauteur de la carte et 48 choix pour la carte restante, soit :

$$p_3 = \frac{13 \cdot 48}{N} \approx 0,024\%$$

5. Si on s'intéresse au poker ouvert, ces probabilités changent.
   (a) Nombre de mains possibles : puisqu'il y a 7 cartes et non 5 pour former une combinaison, le nombre de mains possibles est maintenant $N' = \binom{52}{7}$.
   (b) Nombre de quintes flush : soit p'_1 la probabilité d'avoir une quinte flush. Pour former une quinte flush, il suffit de choisir la couleur (4 choix) et la valeur de la carte basse (10 choix). Il reste alors $\binom{47}{2}$ choix possibles pour les 2 cartes restantes, donc a priori on a $4 \times 10 \times \binom{47}{2}$, mais ce faisant on compte certaines quintes flush plusieurs fois, à savoir toutes les quintes flush non royales pour lesquelles l'une des 2 cartes loisibles est celle immédiatement supérieure à la plus haute carte de la quinte. Il faut donc enlever toutes ces quintes flush à 6 cartes, lesquelles sont au nombre de $4 \times 9 \times 46$: 4 choix pour la couleur, 9 choix pour la carte basse de la quinte flush et 46 choix pour la carte restante. Au total on arrive à :

$$p'_1 = \frac{4 \times 10\binom{47}{2} - 4 \times 9 \times 46}{\binom{52}{7}} \approx 0.031\%$$

   (c) Soit p'_2 la probabilité d'avoir une couleur (mais pas une quinte flush). Il s'agit de bien différencier les cas, puisqu'il y a 3 façons d'obtenir une couleur :
   – 5 cartes de même couleur, 2 autres de couleur différente : $4\binom{13}{5}\binom{39}{2}$ possibilités ;
   – 6 cartes de même couleur, 1 autre de couleur différente : $4\binom{13}{6} \times 39$ possibilités ;
   – 7 cartes de même couleur : $4\binom{13}{7}$ possibilités.
   Il suffit d'ajouter tout ça, de diviser par N' puis d'enlever la probabilité p'_1 d'avoir une quinte flush pour obtenir la probabilité d'avoir une couleur :

$$p'_2 = \frac{4\binom{13}{5}\binom{39}{2} + 4\binom{13}{6} \times 39 + 4\binom{13}{7}}{\binom{52}{7}} - p'_1 \approx 3.025\%$$

   (d) Soit p'_3 la probabilité d'avoir un carré. Cette fois il n'y a pas d'embrouille, tout se passe tranquillement. Pour former un carré, on a 13 choix pour la hauteur de la carte et $\binom{48}{3}$ choix pour les 3 cartes restantes, soit :

$$p'_3 = \frac{13\binom{48}{3}}{N} \approx 0,168\%$$

Exercice 1.6 (L'art de combiner les combinaisons)

1. Formule du binôme de Newton :

$$(x+y)^n = x^n + \binom{n}{1}x^{n-1}y + \cdots + \binom{n}{n-1}xy^{n-1} + y^n.$$

2. Le triangle de Pascal consiste à écrire les coefficients intervenant dans la formule du binôme pour des valeurs croissantes de la puissance n. Ainsi, sur la première ligne, puisque $(x+y)^0 = 1$, on écrit simplement 1. Sur la deuxième ligne, puisque $(x+y)^1 = 1 \times x + 1 \times y$, on écrit 1 et 1. Sur la troisième ligne, puisque $(x+y)^2 = 1 \times x^2 + 2 \times xy + 1 \times y^2$, on écrit 1, 2 et 1. Et ainsi de suite, ce qui donne pour les six premières lignes :

$$\begin{array}{cccccc}
1 & & & & & \\
1 & 1 & & & & \\
1 & 2 & 1 & & & \\
1 & 3 & 3 & 1 & & \\
1 & 4 & 6 & 4 & 1 & \\
1 & 5 & 10 & 10 & 5 & 1
\end{array}$$

On peut remarquer que si l'on interprète un blanc comme un zéro, tout coefficient du triangle s'obtient en ajoutant le coefficient au-dessus et le coefficient au-dessus à gauche. A l'intérieur strict du triangle, ceci se traduit mathématiquement comme suit :

$$\forall 0 \leq k < n \qquad \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}.$$

Cette formule peut se prouver en développant les expressions des deux coefficients binomiaux du membre de gauche et en mettant au même dénominateur, ou par un simple raisonnement combinatoire : pour choisir $(k+1)$ objets parmi $(n+1)$, on peut ou bien prendre le dernier, auquel cas il reste ensuite à choisir k objets parmi n, ou bien ne pas prendre le dernier, auquel cas il faut choisir $(k+1)$ objets parmi n.

3. La première somme s'obtient en prenant $x = y = 1$:

$$S_1 = \sum_{k=0}^{n} \binom{n}{k} = (1+1)^n = 2^n.$$

La deuxième somme s'obtient en prenant $x = -1$ et $y = 1$:

$$S_2 = \sum_{k=0}^{n} (-1)^k \binom{n}{k} = (-1+1)^n = 0.$$

La troisième somme se calcule en bidouillant un peu :

$$S_3 = \sum_{k=0}^{n} k\binom{n}{k} = \sum_{k=1}^{n} k\binom{n}{k} = n\sum_{k=1}^{n} \binom{n-1}{k-1},$$

et on effectue le changement d'indice $j = k-1$ pour obtenir S_1 à peu de choses près :

$$S_3 = n\sum_{j=0}^{n-1} \binom{n-1}{j} = n2^{n-1}.$$

La quatrième somme s'obtient aussi en bricolant le bouzin :

$$S_4 = \sum_{k=0}^{n} \frac{\binom{n}{k}}{k+1} = \frac{1}{n+1}\sum_{k=0}^{n} \binom{n+1}{k+1},$$

et on effectue le changement d'indice $j = k+1$ pour obtenir :

$$S_4 = \frac{1}{n+1}\sum_{j=1}^{n+1} \binom{n+1}{j} = \frac{1}{n+1}\left(-1 + \sum_{j=0}^{n+1} \binom{n+1}{j}\right) = \frac{2^{n+1}-1}{n+1}.$$

4. On a d'une part :

$$P(X) = (1+X)^n(1+X)^n = (1+X)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} X^k,$$

donc le coefficient de X^n est $\binom{2n}{n}$. D'autre part, on peut écrire :

$$P(X) = (1+X)^n(1+X)^n = \left(\sum_{k=0}^{n} \binom{n}{k} X^k\right)\left(\sum_{k=0}^{n} \binom{n}{k} X^k\right).$$

Le coefficient de X^n dans ce produit de polynômes s'obtient en faisant la somme de $(n+1)$ coefficients : le coefficient de X^0 dans le premier polynôme par le coefficient de X^n dans le second polynôme, le coefficient de X^1 dans le premier polynôme par le coefficient de X^{n-1} dans le second polynôme, etc. Finalement, en tenant compte du fait que $\binom{n}{k} = \binom{n}{n-k}$, on voit que ce coefficient vaut exactement la somme voulue. On en déduit que :

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

Exercice 1.7 (Formule de Poincaré)

1. On a

$$\#(A \cup B) = \#A + \#B - \#(A \cap B).$$

Application : si on appelle A l'ensemble des élèves ayant pour langues (anglais,espagnol), B l'ensemble des élèves ayant pour langues (anglais,allemand), l'effectif de la classe est donc

$$\#(A \cup B) = \#A + \#B - \#(A \cap B) = 20 + 15 - 5 = 30.$$

2. On a cette fois

$$\#(A \cup B \cup C) = \#A + \#B + \#C - (\#(A \cap B) + \#(A \cap C) + \#(B \cap C)) + \#(A \cap B \cap C)$$

3. Généralisation : la formule de Poincaré s'écrit

$$\#(A_1 \cup \cdots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n} \#(A_{i_1} \cap \cdots \cap A_{i_k}) \right).$$

Elle peut se prouver par récurrence. Elle est vraie aux ordres $n = 2$ et $n = 3$ d'après les questions précédentes. Supposons-la vérifiée à l'ordre $n-1$, c'est-à-dire que

$$\#(A_1 \cup \cdots \cup A_{n-1}) = \sum_{k=1}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k}) \right).$$

Le but est donc de montrer qu'elle est encore satisfaite à l'ordre n. L'associativité de l'union donne tout d'abord

$$\#(A_1 \cup \cdots \cup A_n) = \#((A_1 \cup \cdots \cup A_{n-1}) \cup A_n)$$

et la formule vue pour $n = 2$ impose

$$l\#(A_1 \cup \cdots \cup A_n) = \#(A_1 \cup \cdots \cup A_{n-1}) + \#A_n - \#((A_1 \cup \cdots \cup A_{n-1}) \cap A_n). \qquad (1.1)$$

L'hypothèse de récurrence à l'ordre $n-1$ donne pour la somme des deux premiers termes

$$\begin{aligned}
&\#A_n + \#(A_1 \cup \cdots \cup A_{n-1}) \\
&\qquad = \#A_n + \sum_{k=1}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k}) \right) \\
&\qquad = \#A_1 + \cdots + \#A_n + \sum_{k=2}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k}) \right).
\end{aligned}$$

Par ailleurs, la distributivité de l'intersection par rapport à l'union donne pour le dernier terme de l'équation (1.1)

$$\#((A_1 \cup \cdots \cup A_{n-1}) \cap A_n) = \#((A_1 \cap A_n) \cup \cdots \cup (A_{n-1} \cap A_n)),$$

expression à laquelle nous pouvons appliquer l'hypothèse de récurrence à l'ordre $n-1$ pour obtenir

$$\begin{aligned}
&\#((A_1 \cap A_n) \cup \cdots \cup (A_{n-1} \cap A_n)) \\
&\qquad = \sum_{k=1}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#((A_{i_1} \cap A_n) \cap \cdots \cap (A_{i_k} \cap A_n)) \right) \\
&\qquad = \sum_{k=1}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n) \right)
\end{aligned}$$

ce qui s'écrit encore

$$\begin{aligned}
&\sum_{k=1}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n) \right) \\
&\qquad = \sum_{k=1}^{n-2} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n) \right) \\
&\qquad\qquad + (-1)^{n-2} \#(A_1 \cap \cdots \cap A_{n-1} \cap A_n)
\end{aligned}$$

Au total, l'équation (1.1) devient

$$\begin{aligned}
&\#(A_1 \cup \cdots \cup A_n) \\
&\quad = \#A_1 + \cdots + \#A_n \\
&\qquad + \sum_{k=2}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k}) \right) \\
&\qquad\quad - \sum_{k=1}^{n-2} (-1)^{k-1} \left(\sum_{1 \le i_1 < \cdots < i_k \le n-1} \#(A_{i_1} \cap \cdots \cap A_{i_k} \cap A_n) \right) \\
&\qquad\quad - (-1)^{n-2} \#(A_1 \cap \cdots \cap A_n).
\end{aligned}$$

ou de façon équivalente

$$\begin{aligned}
&\#(A_1 \cup \dots \cup A_n) \\
&\quad = \#A_1 + \dots + \#A_n \\
&\qquad + \sum_{k=2}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \dots < i_k \le n-1} \#(A_{i_1} \cap \dots \cap A_{i_k}) \right) \\
&\qquad\quad + \sum_{k=2}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \dots < i_k = n} \#(A_{i_1} \cap \dots \cap A_{i_k}) \right) \\
&\qquad\qquad + (-1)^{n-1} \#(A_1 \cap \dots \cap A_n).
\end{aligned}$$

Les deux termes intermédiaires se regroupent maintenant de façon naturelle

$$\begin{aligned}
&\#(A_1 \cup \dots \cup A_n) \\
&\quad = \#A_1 + \dots + \#A_n \\
&\qquad + \sum_{k=2}^{n-1} (-1)^{k-1} \left(\sum_{1 \le i_1 < \dots < i_k \le n} \#(A_{i_1} \cap \dots \cap A_{i_k}) \right) \\
&\qquad\quad + (-1)^{n-1} \#(A_1 \cap \dots \cap A_n).
\end{aligned}$$

pour donner finalement

$$\#(A_1 \cup \dots \cup A_n) = \sum_{k=1}^{n} (-1)^{k-1} \left(\sum_{1 \le i_1 < \dots < i_k \le n} \#(A_{i_1} \cap \dots \cap A_{i_k}) \right),$$

qui est la formule de Poincaré à l'ordre n. La récurrence est établie.

Exercice 1.8 (Dérangements)

1. Les étudiants étant numérotés de 1 à n, il y a $n!$ répartitions possibles des manteaux parmi les n étudiants.
2. Pour qu'aucun des étudiants ne récupère son manteau, il faut que le premier étudiant ne récupère pas le sien, le deuxième non plus, ..., le n-ème non plus. On peut donc décrire l'événement A comme suit :

$$A = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n} = \bigcap_{i=1}^{n} \overline{A_i}.$$

3. Soit $k \in \{1, \dots, n\}$. Il y a $\binom{n}{k}$ combinaisons de k indices parmi n. Pour chacune, il y en a une seule telle que $1 \le i_1 < \dots < i_k \le n$. Il y a donc exactement $\binom{n}{k}$ séquences d'indices $(i_1, \dots, i_k)$ telles que $1 \le i_1 < \dots < i_k \le n$.
4. La séquence $(i_1, \dots, i_k)$ étant fixée et l'événement $(A_{i_1} \cap \dots \cap A_{i_k})$ réalisé, il y a k indices parmi n qui ne bougent pas et $(n-k)$ qui permutent. Le cardinal de $(A_{i_1} \cap \dots \cap A_{i_k})$ vaut donc :

$$\#(A_{i_1} \cap \dots \cap A_{i_k}) = (n-k)!$$

5. Puisque $A = \overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n} = \bigcap_{i=1}^{n} \overline{A_i}$ et que $\overline{E_1 \cap E_2} = \overline{E_1} \cup \overline{E_2}$, on en déduit :

$$\overline{A} = \overline{\overline{A_1} \cap \overline{A_2} \cap \dots \cap \overline{A_n}} = A_1 \cup \dots \cup A_n = \bigcup_{i=1}^{n} A_i,$$

et on peut donc utiliser la formule de Poincaré pour calculer $\#\overline{A}$:

$$\#\overline{A} = \sum_{k=1}^{n}(-1)^{k-1}\left(\sum_{1\leq i_1<\cdots<i_k\leq n}\#(A_{i_1}\cap\cdots\cap A_{i_k})\right) = \sum_{k=1}^{n}(-1)^{k-1}\sum_{1\leq i_1<\cdots<i_k\leq n}(n-k)!$$

et comme, pour k fixé, il y a $\binom{n}{k}$ séquences d'indices $(i_1,\dots,i_k)$ telles que $1\leq i_1<\cdots< i_k\leq n$, on arrive à :

$$\#\overline{A} = \sum_{k=1}^{n}(-1)^{k-1}\binom{n}{k}(n-k)! = \sum_{k=1}^{n}(-1)^{k-1}\frac{n!}{k!}.$$

Puisqu'il y a en tout $n!$ répartitions possibles des manteaux parmi les n étudiants, on en déduit bien que :

$$\mathbb{P}(A) = 1-\mathbb{P}(\overline{A}) = 1-\sum_{k=1}^{n}\frac{(-1)^{k-1}}{k!}.$$

6. On peut réécrire la probabilité précédente comme suit :

$$\mathbb{P}(A) = 1+\sum_{k=1}^{n}\frac{(-1)^{k}}{k!} = \sum_{k=0}^{n}\frac{(-1)^{k}}{k!},$$

d'où l'approximation :

$$\mathbb{P}(A) \approx \sum_{k=0}^{+\infty}\frac{(-1)^{k}}{k!} = \frac{1}{e}\approx 0,37.$$

Remarque : puisque la série $\sum\frac{(-1)^k}{k!}$ est alternée, on peut même donner une qualité de l'approximation :

$$\left|\mathbb{P}(A)-\frac{1}{e}\right| \leq \frac{1}{(n+1)!}$$

L'approximation est donc excellente, par exemple pour un effectif de $n=20$ étudiants, le majorant de l'erreur est de l'ordre de 2.10^{-20}.

7. Le nombre d_n de dérangements d'un ensemble à n éléments est donc :

$$d_n = n!\sum_{k=0}^{n}\frac{(-1)^{k}}{k!}.$$

Exercice 1.9 (Traductions ensemblistes d'événements)

On a les décompositions suivantes (cf. figure 1.8) :

- A seul se réalise : $E_1 = A\setminus(B\cup C) = A\cap\overline{B}\cap\overline{C}$;
- A et C se réalisent mais pas B : $E_2 = (A\cap C)\setminus B = A\cap C\cap\overline{B}$;
- au moins l'un des trois événements se réalise : $E_3 = A\cup B\cup C$;
- au moins deux des trois événements se réalisent : $E_4 = (A\cap B)\cup(A\cap C)\cup(B\cap C)$;
- les trois événements se réalisent : $E_5 = A\cap B\cap C$;
- aucun ne se réalise : $E_6 = \overline{A\cup B\cup C} = \overline{A}\cap\overline{B}\cap\overline{C}$;
- au plus l'un des trois se réalise : $E_7 = \overline{E_4}$;
- au plus deux des trois se réalisent : $E_8 = \overline{E_5} = \overline{A}\cup\overline{B}\cup\overline{C}$;
- exactement deux des trois se réalisent : $E_9 = E_4\setminus(A\cap B\cap C)$;
- au plus trois se réalisent : $E_{10} = \Omega$.

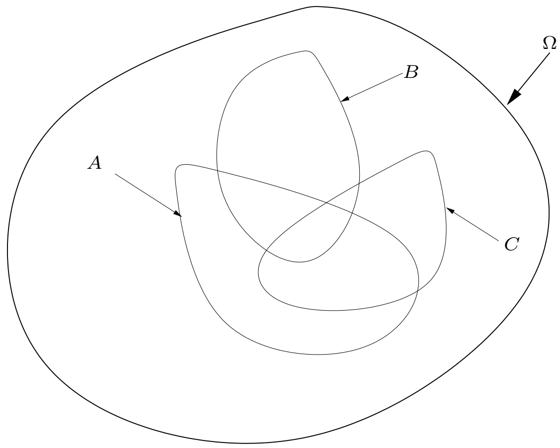


FIGURE 1.8 – L'abécédaire des sous-ensembles.

Exercice 1.10 (Exemple de tribu engendrée)

1. On a $S_0 = \{0, 1, 2\}$, $S_2 = \{2, 3, 4\}$, $S_3 = \{3, 4, 5\}$, $S_4 = \{4, 5, 6\}$, etc. Par stabilité d'une tribu par intersection, on voit donc que pour tout $n \geq 4$, le singleton $\{n\} = S_n \cap S_{n-1} \cap S_{n-2}$ appartient à $\mathcal{F}$. Pour les mêmes raisons, le singleton $\{2\} = S_0 \cap S_2$ appartient à $\mathcal{F}$. Enfin $\{3\} = S_2 \cap \overline{\{2\}} \cap \overline{\{4\}}$ appartient lui aussi à $\mathcal{F}$.
2. Une partie de $\mathbb{N}^{**} = \{2, 3, \dots\}$ est l'union au plus dénombrable de singletons piochés parmi les entiers supérieurs ou égaux à 2. Comme on vient de voir que chacun de ces singletons est dans $\mathcal{F}$ et que $\mathcal{F}$ est stable par union au plus dénombrable, on en déduit que tout sous-ensemble de $\mathbb{N}^{**}$ est dans $\mathcal{F}$. Autrement dit : $\mathcal{P}(\mathbb{N}^{**}) \subset \mathcal{F}$.
3. On voit que $S_0 \setminus \{2\} = \{0, 1\} \in \mathcal{F}$, mais aucun des singletons $\{0\}$ et $\{1\}$ n'appartient à $\mathcal{F}$, autrement dit on ne peut pas séparer 0 et 1 dans $\mathcal{F}$. De fait $A \in \mathcal{F}$ si et seulement si il existe $B \in \mathcal{P}(\mathbb{N}^{**})$ tel que $A = B$ ou $A = \{0, 1\} \cup B$.

Exercice 1.11 (Lancer infini d'une pièce)

1. On peut décrire les événements de la façon suivante :
   - $E_1 = \bigcap_{i=5}^{+\infty} A_i$: à partir du cinquième, tous les lancers donnent Pile ;
   - $E_2 = \left(\bigcap_{i=1}^{4} \overline{A_i}\right) \cap \left(\bigcap_{i=5}^{+\infty} A_i\right)$: les 4 premiers lancers donnent Face et tous les suivants Pile ;
   - $E_3 = \bigcup_{i=5}^{+\infty} A_i$: au moins l'un des lancers à partir du cinquième donne Pile.
2. L'événement E : "On obtient au moins une fois Pile après le n-ème lancer" s'écrit encore : $E = \bigcup_{i=n+1}^{+\infty} A_i$.
3. A l'aide des A_i, on peut écrire :
   (a) $B_n = \bigcap_{i=n}^{+\infty} A_i$.
   (b) $B = \bigcup_{n=1}^{+\infty} \left(\bigcap_{i=n}^{+\infty} A_i\right)$.

Exercice 1.12 (Inégalité de Bonferroni)

1. Soit E l'événement : "un seul des deux événements se réalise". On peut écrire $E = (A \cup B) \setminus (A \cap B)$, ce qui est souvent noté $E = A \Delta B$, différence symétrique de A et de B. Puisque $A \cap B \subset A \cup B$, il est clair que $\mathbb{P}(E) = \mathbb{P}(A \cup B) - \mathbb{P}(A \cap B)$. Il reste à utiliser l'additivité forte :

$$\mathbb{P}(E) = (\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)) - \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - 2\mathbb{P}(A \cap B).$$

2. Soit A et B deux événements tels que $\mathbb{P}(A) = 0,9$ et $\mathbb{P}(B) = 0,8$.
   (a) On utilise l'additivité forte :
   $$\mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = 1,7 - \mathbb{P}(A \cup B),$$
   et puisque $\mathbb{P}(A \cup B) \leq 1$, on en déduit bien que $\mathbb{P}(A \cap B) \geq 0,7$.
   (b) Sur l'espace $\Omega = \{1, \ldots, 10\}$ muni de l'équiprobabilité, considérons $A = \{1, \ldots, 9\}$, $B = \{3, \ldots, 10\}$, ce qui donne $A \cap B = \{3, \ldots, 9\}$. Dans ce cas on a bien $\mathbb{P}(A) = 0,9$, $\mathbb{P}(B) = 0,8$ et $\mathbb{P}(A \cap B) = 0,7$.
   (c) Puisque $A \cap B \subset A$ et $A \cap B \subset B$, on a $\mathbb{P}(A \cap B) \leq \mathbb{P}(A) = 0,9$ et $\mathbb{P}(A \cap B) \leq \mathbb{P}(B) = 0,8$, donc $\mathbb{P}(A \cap B)$ vaut au maximum $\min(\mathbb{P}(A), \mathbb{P}(B)) = 0,8$, avec égalité lorsque $A \cap B = B$, c'est-à-dire lorsque B est contenu dans A. Sur notre exemple, il suffit de prendre $A = \{1, \ldots, 9\}$ et $B = \{1, \ldots, 8\}$. De façon générale, on a $\mathbb{P}(A \cap B) \leq \min(\mathbb{P}(A), \mathbb{P}(B))$, avec égalité lorsque l'un des événements est contenu dans l'autre.
3. Généralisation : en remarquant que
   $$A_1 \cap \cdots \cap A_n = \overline{\overline{A_1} \cup \cdots \cup \overline{A_n}},$$
   et la relation $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$, il vient
   $$\mathbb{P}(A_1 \cap \cdots \cap A_n) = 1 - \mathbb{P}(\overline{A_1} \cup \cdots \cup \overline{A_n}),$$
   or par sous-σ-additivité :
   $$\mathbb{P}(\overline{A_1} \cup \cdots \cup \overline{A_n}) \leq \sum_{i=1}^{n} \mathbb{P}(\overline{A_i}) = \sum_{i=1}^{n} (1 - \mathbb{P}(A_i)).$$
   Au total on obtient bien :
   $$\mathbb{P}(A_1 \cap \cdots \cap A_n) \geq 1 - \sum_{i=1}^{n} (1 - \mathbb{P}(A_i)) = \sum_{i=1}^{n} \mathbb{P}(A_i) - (n-1).$$
   Par ailleurs, le fait que chaque A_i contienne l'intersection des A_i implique que
   $$\mathbb{P}(A_1 \cap \cdots \cap A_n) \leq \min_{1 \leq i \leq n} \mathbb{P}(A_i),$$
   avec égalité si et seulement si l'un des A_i est contenu dans tous les autres.
   Remarque : on peut aussi montrer la minoration de la probabilité d'intersection par récurrence.
   – Elle est évidente pour $n = 1$ et vraie pour $n = 2$ par le raisonnement de la question précédente.
   – Supposons-la vraie à l'indice $n \geq 2$ et considérons les événements $A_1, \ldots, A_{n+1}$. Alors par la formule d'additivité forte et par associativité de l'intersection :
   $$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) = \mathbb{P}((A_1 \cap \cdots \cap A_n) \cap A_{n+1}),$$
   d'où l'on déduit en appliquant l'inégalité avec $n = 2$:
   $$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) \geq \mathbb{P}(A_1 \cap \cdots \cap A_n) + \mathbb{P}(A_{n+1}) - 1.$$
   Il reste à appliquer l'hypothèse de récurrence :
   $$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) \geq \sum_{i=1}^{n} \mathbb{P}(A_i) - (n-1) + \mathbb{P}(A_{n+1}) - 1,$$

ce qui est exactement la formule de récurrence à l'ordre $(n+1)$:

$$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) \geq \sum_{i=1}^{n+1} \mathbb{P}(A_i) - n.$$

Exercice 1.13 (Alea jacta est)

1. Pour tout $i \in \{2, \ldots, 12\}$, on note p_i la probabilité que la somme fasse i. Si on regroupe ces i dans le vecteur ligne $p = [p_2, \ldots, p_{12}]$, on obtient :

$$p = \frac{1}{36}[1, 2, 3, 4, 5, 6, 5, 4, 3, 2, 1].$$

2. On répète maintenant l'expérience précédente jusqu'à ce qu'une somme de 5 ou 7 apparaisse. On désigne par E_n l'événement : "Une somme de 5 apparaît au n-ème double jet et sur les $(n-1)$ premiers coups ni la somme de 5 ni celle de 7 n'est apparue."

   (a) A chacun des $(n-1)$ premiers coups, la probabilité pour que ni une somme de 5 ni une somme de 7 n'apparaisse est $1 - (p_5 + p_7) = 13/18$. Au n-ème coup, la probabilité qu'une somme de 5 apparaisse est $p_5 = 4/36 = 1/9$. On en déduit :

$$\mathbb{P}(E_n) = (1 - (p_5 + p_7))^{n-1} p_5 = \frac{1}{9} \times \left(\frac{13}{18}\right)^{n-1}$$

   (b) Pour que l'événement E ("on s'arrête sur une somme de 5") se réalise, il faut et il suffit que l'un des E_n se réalise, c'est-à-dire en termes ensemblistes : $E = \cup_{n=1}^{\infty} E_n$. Puisque les E_n sont deux à deux incompatibles, la sigma-additivité de $\mathbb{P}$ donne :

$$\mathbb{P}(E) = \sum_{n=1}^{\infty} \mathbb{P}(E_n) = \frac{1}{9} \sum_{n=1}^{\infty} \left(\frac{13}{18}\right)^{n-1},$$

   où on reconnaît la somme des termes d'une suite géométrique de raison 13/18 :

$$\mathbb{P}(E) = \frac{1}{9} \frac{1}{1 - \frac{13}{18}} = \frac{2}{5}.$$

   Ce résultat pouvait se trouver sans ces calculs : puisqu'on va nécessairement s'arrêter sur une somme de 5 ou de 7, la probabilité que l'on s'arrête sur une somme de 5 est tout simplement $\mathbb{P}(E) = p_5/(p_5 + p_7) = 2/5$, et celle qu'on s'arrête sur une somme de 7 est $\mathbb{P}(\overline{E}) = 1 - \mathbb{P}(E) = 3/5$.

Exercice 1.14 (Application de la sous-σ-additivité)

La sous-σ-additivité permet d'écrire :

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

Si tous les $\mathbb{P}(A_i)$ étaient de probabilité strictement inférieure à $1/n$, ceci serait clairement impossible puisque la somme des probabilités serait alors strictement inférieure à 1. On en déduit que l'un au moins des événements A_i est bien de probabilité supérieure ou égale à $\frac{1}{n}$.

Exercice 1.15 (Limites supérieures et inférieures d'ensembles)

1. Soit A et B deux parties de Ω et la suite (A_n) définie par $A_{2n} = A$ et $A_{2n+1} = B$ pour tout $n \in \mathbb{N}$.

- Concernant la limite sup, un élément ω lui appartient s'il est dans une infinité de A_n : soit tous les indices n sont pairs, auquel cas $\omega \in A$, soit tous les indices n sont impairs, auquel cas $\omega \in B$, soit il y en a des pairs et des impairs, auquel cas $\omega \in A \cap B$. Quoi qu'il en soit, il est clair que si ω est dans la limite sup des A_n, on a nécessairement $\omega \in A \cup B$. La réciproque marche aussi : si $\omega \in A \cup B$, le raisonnement précédent permet d'exhiber une infinité de A_n auxquels ω appartient. Ainsi on a $\overline{\lim} A_n = A \cup B$.
- Concernant la limite inf, un élément ω lui appartient s'il est dans tous les A_n sauf un nombre fini. Il existe donc un indice $n_0 \in \mathbb{N}$ tel que :
$$\forall n \geq n_0 \qquad \omega \in A_n.$$
En particulier $\omega \in A_{n_0}$ et $\omega \in A_{n_0+1}$, ainsi $\omega \in A$ et $\omega \in B$, donc $\omega \in A \cap B$. Réciproquement, soit $\omega \in A \cap B$, alors il est clair que $\omega \in A_n$ pour tout $n \in \mathbb{N}$. Ainsi on a $\underline{\lim} A_n = A \cap B$.

2. On peut réécrire automatiquement les définitions de $\overline{\lim} A_n$ et $\underline{\lim} A_n$ à l'aide des quantificateurs logiques $\exists$ et $\forall$. Ceci donne pour la limite sup :
$$\overline{\lim} A_n = \{\omega \in \Omega \ : \ \forall n \in \mathbb{N}, \ \exists k \geq n, \ \omega \in A_k\},$$
et pour la limite inf :
$$\underline{\lim} A_n = \{\omega \in \Omega \ : \ \exists n \in \mathbb{N}, \ \forall k \geq n, \ \omega \in A_k\}.$$
On peut aussi les traduire en termes ensemblistes, en remplaçant $\exists$ par $\cup$ et $\forall$ par $\cap$. Ceci donne pour la limite sup :
$$\overline{\lim} A_n = \bigcap_{n \in \mathbb{N}} \bigcup_{k \geq n} A_k,$$
et pour la limite inf :
$$\underline{\lim} A_n = \bigcup_{n \in \mathbb{N}} \bigcap_{k \geq n} A_k.$$
3. On donne ici les résultats sans les justifier.
   (a) Si $A_n =]-\infty, n]$ pour tout $n \geq 0$, alors $\overline{\lim} A_n = \underline{\lim} A_n = \mathbb{R}$. Lorsque la suite $(A_n)_{n \geq 0}$ est croissante pour l'inclusion, comme c'est le cas ici, on a en fait le résultat général :
$$\overline{\lim} A_n = \underline{\lim} A_n = \bigcup_{n=0}^{\infty} A_n.$$
   (b) Si $A_n =]-\infty, -n]$ pour tout $n \geq 0$, alors $\overline{\lim} A_n = \underline{\lim} A_n = \emptyset$. Lorsque la suite $(A_n)_{n \geq 0}$ est décroissante pour l'inclusion, comme c'est le cas ici, on a en fait le résultat général :
$$\overline{\lim} A_n = \underline{\lim} A_n = \bigcap_{n=0}^{\infty} A_n.$$
   (c) Si $A_n =]-1/n, 1/n[$ pour tout $n > 0$, alors on est à nouveau dans le cas d'une suite décroissante pour l'inclusion et :
$$\overline{\lim} A_n = \underline{\lim} A_n = \bigcap_{n=0}^{\infty} A_n = \{0\}.$$
   (d) Dans ce dernier cas, on a :
$$\underline{\lim} A_n =]-\infty, -1[\subsetneq \overline{\lim} A_n =]-\infty, 1].$$

Exercice 1.16 (Lemme de Borel-Cantelli)

1. On a vu dans l'exercice 1.15 que :

$$A = \bigcap_{n\in\mathbb{N}} \bigcup_{k\geq n} A_k = \bigcap_{n\in\mathbb{N}} D_n.$$

Or pour tout $n \in \mathbb{N}$, l'ensemble $D_n = \bigcup_{k\geq n} A_k$ appartient à $\mathcal{F}$ puisque la tribu $\mathcal{F}$ est stable par union dénombrable. Puisqu'elle est également stable par intersection dénombrable, $A = \bigcap_{n\in\mathbb{N}} D_n$ appartient encore à $\mathcal{F}$.

2. Si on note $D_n = \bigcup_{k=n}^{+\infty} A_k$, on a :

$$D_{n+1} = \bigcup_{k=n+1}^{+\infty} A_k \subset A_n \cup \left(\bigcup_{k=n+1}^{+\infty} A_k\right) = \bigcup_{k=n}^{+\infty} A_k = D_n,$$

et la suite $(D_n)_{n\geq 0}$ est bien décroissante pour l'inclusion.

3. On suppose que $\sum_{n=0}^{+\infty} \mathbb{P}(A_n) < +\infty$. La sous-$\sigma$-additivité permet d'écrire pour tout $k \in \mathbb{N}$:

$$\mathbb{P}(D_n) = \mathbb{P}\left(\bigcup_{k=n}^{+\infty} A_k\right) \leq \sum_{k=n}^{+\infty} \mathbb{P}(A_n),$$

et le terme de droite est le reste d'une série convergente, qui tend donc nécessairement vers 0, d'où a fortiori $\lim_{n\to+\infty} \mathbb{P}(D_n) = 0$.

4. La suite $(D_n)_{n\geq 0}$ étant décroissante pour l'inclusion, on peut appliquer la continuité monotone décroissante :

$$\mathbb{P}(A) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty} D_n\right) = \lim_{n\to+\infty} \mathbb{P}(D_n) = 0.$$

Ainsi, lorsque la série $\sum_{n=\in\mathbb{N}} \mathbb{P}(A_n)$ est convergente, il est improbable qu'une infinité d'événements A_n se produisent simultanément.

Exercice 1.17 (Ensembles dénombrables)

Pour montrer qu'un ensemble est dénombrable, il suffit de pouvoir indiquer un procédé de numérotage qui n'oublie aucun élément de E. C'est ce que nous allons utiliser dans la suite.

1. Pour voir que $\mathbb{Z}$ est dénombrable, il suffit d'écrire :

$$\mathbb{Z} = (0, -1, +1, -2, +2, ...),$$

c'est-à-dire $\mathbb{Z} = (u_n)_{n\geq 0}$, avec :

$$\forall n \in \mathbb{N} \qquad \begin{cases} u_{2n} & = n \\ u_{2n+1} & = -(n+1) \end{cases}$$

2. Pour l'ensemble $\mathbb{Q}$ des rationnels, on exhibe en figure 1.9 un moyen de parcourir l'ensemble des couples (p, q) avec $p \in \mathbb{N}$ et $q \in \mathbb{N}^*$. Puisqu'on ne suppose pas p et q premiers entre eux, l'application $(p, q) \mapsto \frac{p}{q}$ n'est pas bijective, mais peu importe puisqu'elle est surjective donc on n'oublie aucun rationnel positif et c'est bien là l'essentiel : dans la suite $(u_n)_{n\geq 0}$ ainsi obtenue, il suffira ensuite d'éliminer les u_n redondants. Appelons $(v_n)_{n\geq 0}$ cette suite épurée, elle est donc en bijection avec $\mathbb{Q}^+$. Pour obtenir $\mathbb{Q}$ tout entier, on peut alors procéder comme pour $\mathbb{Z}$, en alternant un élément de $\mathbb{Q}^+$ et son opposé dans $\mathbb{Q}^-$. On obtient alors une suite $(q_n)_{n\geq 0}$ décrivant l'ensemble des rationnels.

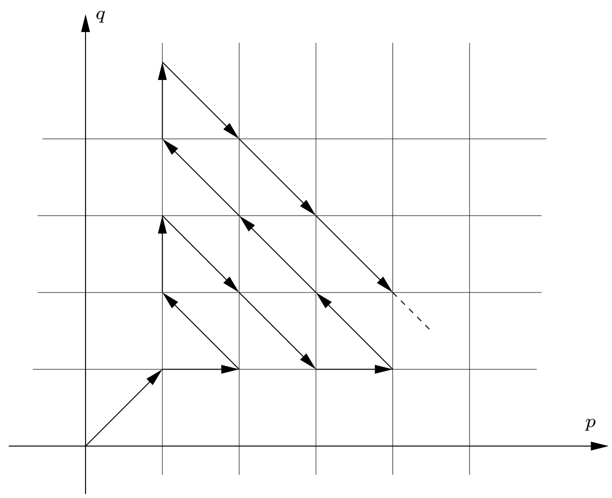


Figure 1.9 – Une façon de parcourir l'ensemble des couples $(p, q) \in \{(0,0)\} \cup \mathbb{N} \times \mathbb{N}^*$.

3. Pour montrer que $\mathbb{R}$ n'est pas dénombrable, il suffit de prouver que l'ensemble $[0, 1[$ ne l'est pas. Pour cela, commençons par rappeler que tout nombre réel x de l'intervalle $[0, 1[$ s'écrit de façon unique sous la forme :

$$x = \sum_{n=1}^{+\infty} x_n 10^{-n} = \frac{x_1}{10} + \frac{x_2}{100} + \cdots + \frac{x_n}{10^n} + \ldots$$

avec $x_n \in \{0, 1, \ldots, 9\}$ pour tout n. C'est le développement décimal de x et on écrit encore $x = 0, \overline{x_1 x_2 \ldots x_n \ldots}$. On convient en général que ce développement décimal ne finit pas par une infinité de 9, c'est-à-dire qu'on écrit $x = 0.3780000$ ou plus succinctement $x = 0.378$, plutôt que $x = 0.37799999\ldots$. On raisonne alors pas l'absurde. Si on suppose $[0, 1[$ dénombrable, il existe une suite $(u_n)_{n \geq 1}$ telle que $[0, 1[= (u_n)_{n \geq 1}$. Chaque terme u_n admet un développement décimal, que l'on convient d'écrire comme suit :

$$u_n = 0, \overline{u_n^1 u_n^2 \ldots u_n^n \ldots}.$$

Vient alors la ruse diabolique de Cantor, connue sous le nom de procédé diagonal : en considérant le nombre $x = 0, \overline{x_1 x_2 \ldots x_n \ldots}$ pour lequel :

$$\forall n \in \mathbb{N} \qquad x_n = \begin{cases} 0 & \text{si } u_n^n \neq 0 \\ 1 & \text{si } u_n^n = 0 \end{cases}$$

Le réel x est encore clairement dans $[0, 1[$, or il est différent de chaque u_n puisque par construction il en diffère au moins par une décimale ($x_n \neq u_n^n$ pour tout n). On a donc une contradiction, ce qui signifie que l'hypothèse de départ était absurde : $[0, 1[$ n'est pas dénombrable.

Exercice 1.18 (L'oracle d'Oberhausen)

1. Que ce soit pour un match de poule ou avec élimination directe, le poulpe se pose sur l'un des deux récipients et ne peut donc pronostiquer un match nul. Pour un match de poule, en notant S (comme Succès), A, B et N les événements correspondant respectivement à un pronostic correct, la victoire de l'équipe A, la victoire de l'équipe B et un match nul, alors la formule des probabilités totales donne :

$$\mathbb{P}(S) = \mathbb{P}(S|A)\mathbb{P}(A) + \mathbb{P}(S|B)\mathbb{P}(B) + \mathbb{P}(S|N)\mathbb{P}(N).$$

On considère a priori que les 3 issues possibles d'une rencontre sont équiprobables : $\mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(N) = 1/3$. Puisque le pouple ne peut prédire de match nul, il vient par contre : $\mathbb{P}(S|A) = \mathbb{P}(S|B) = 1/2$ et $\mathbb{P}(S|N) = 0$. Ainsi, pour un match de poule, la probabilité d'un pronostic juste est $\mathbb{P}(S) = 1/3$. Le même raisonnement montre que pour un match avec élimination directe, la probabilité d'un pronostic juste est cette fois $\mathbb{P}(S) = 1/2$.

2. Pour chacun des 3 matchs de poule, la probabilité de succès était de $1/3$, puis $1/2$ pour chacun des 5 autres matchs. La probabilité de 8 bons pronostics était donc :
$$p = \left(\frac{1}{3}\right)^3 \left(\frac{1}{2}\right)^5 = \frac{1}{864}.$$

Exercice 1.19 (Le poulpe démasqué)

1. La probabilité qu'aucune des N personnes jouant au Loto pour un tirage donné ne remporte le gros lot s'écrit donc $P = (1-p)^N$.
2. Le nombre N de joueurs nécessaires pour qu'il y ait au moins une chance sur deux que le gros lot soit remporté doit donc vérifier :
$$1 - P \geq \frac{1}{2} \Leftrightarrow (1-p)^N \leq \frac{1}{2}$$
En passant aux logarithmes et en utilisant l'approximation $\ln(1-p) \approx -p$ pour p proche de 0, ceci donne :
$$N \geq \frac{-\ln 2}{\ln(1-p)} \approx \frac{\ln 2}{p}.$$
Si $p = 1/n$, ceci nous dit qu'il faut $N \geq n\ln(2) \approx 0.69n$. Pour le Loto, en prenant p égal à une chance sur 19 millions, il faut donc plus de 13 millions de joueurs. Ceci paraît énorme, mais d'une part le nombre de joueurs réguliers au Loto se compte effectivement en millions, d'autre part la plupart d'entre eux jouent plusieurs grilles. De fait il n'est pas rare qu'il y ait un gagnant au Loto : nonobstant une probabilité de gain très faible pour une personne donnée, la probabilité que le gros lot soit remporté par quelqu'un ne l'est pas du tout (l'union fait la force, en quelque sorte).
3. Le même raisonnement peut être appliqué pour l'histoire du poulpe. Le nombre N de "poulpes" nécessaires pour qu'avec une probabilité supérieure à 90%, l'un au moins pronostique les 8 bons résultats se calcule comme suit (en notant $p = 1/864$) :
$$1 - (1-p)^N \geq \frac{9}{10} \Leftrightarrow N \geq \frac{-\ln 10}{\ln(1-p)} \approx \frac{\ln 10}{p} \approx 1990$$
Il suffit donc que 2 000 personnes à travers le monde aient utilisé un moyen ou un autre pour faire des pronostics (par exemple via Mani le perroquet à Singapour) pour qu'il se trouve à coup sûr un oracle de pacotille dans toute la ménagerie...

Exercice 1.20 (L'art de se tirer une balle dans le pied)

1. Puisqu'on suppose l'indépendance entre réacteurs et d'une année à l'autre, le calcul est ici le même que dans l'exercice 1.19. Si p est la probabilité d'un accident majeur par an et par réacteur, si n est le nombre de réacteurs, alors la probabilité d'au moins un accident durant les t prochaines années est :
$$P = 1 - (1-p)^{nt},$$
ce qui donne pour la France :
$$P_F = 1 - (1 - 0.0003)^{58\times 30} \approx 41\%$$

et pour l'Europe :

$$P_E = 1 - (1 - 0.0003)^{143\times 30} \approx 72\%$$

Bien entendu ces résultats n'ont pas grand sens : estimer une probabilité aussi faible que 3×10^{-4} à partir de 4 événements survenus sur une trentaine d'années n'est pas sérieux, d'autant que ceci suppose l'indépendance entre accidents, ce qui n'est clairement pas le cas au Japon.

2. Lorsque p tend vers 0 et nt est fixé, on a $1 - (1-p)^{nt} \approx ntp$. Noter qu'en allant un terme plus loin dans le développement, on obtient

$$1 - (1-p)^{nt} \approx ntp - \frac{nt(nt-1)}{2}p^2$$

L'approximation est donc acceptable si $\frac{nt-1}{2}p \ll 1$.

3. L'idée fabuleuse des auteurs de l'article a consisté à appliquer l'approximation lorsque celle-ci n'est plus valable du tout. En l'occurrence, pour la France, ceci donne

$$ntp = 58 \times 30 \times 0.0003 \approx 0.52$$

c'est-à-dire tout de même une erreur de 11% par rapport à P_F. Mais ce n'est rien par rapport à l'Europe :

$$ntp = 143 \times 30 \times 0.0003 \approx 1.29$$

Encore bravo !

4. En écrivant une telle phrase dans un contrôle, vous auriez 0 avec probabilité 1.

Exercice 1.21 (Probabilités composées)

1. On note B_1 (resp. B_2, N_3) l'événement : "La première (resp. la deuxième, la troisième) boule tirée est blanche (resp. blanche, noire)." On veut donc calculer $\mathbb{P}(B_1 \cap B_2 \cap N_3)$, ce pour quoi on utilise la formule des probabilités composées :

$$\mathbb{P}(B_1 \cap B_2 \cap N_3) = \mathbb{P}(B_1)\mathbb{P}(B_2|B_1)\mathbb{P}(N_3|B_1 \cap B_2).$$

On a initialement 4 boules blanches et 3 boules noires dans l'urne donc $\mathbb{P}(B_1) = 4/7$. Une fois la boule blanche tirée, il reste 3 boules blanches et 3 boules noires donc $\mathbb{P}(B_2|B_1) = 1/2$. Après ces deux premiers tirages, il reste 2 boules blanches et 3 boules noires, d'où $\mathbb{P}(N_3|B_1 \cap B_2) = 3/5$. Au total, on obtient $\mathbb{P}(B_1 \cap B_2 \cap N_3) = \frac{6}{35}$.

2. Supposons qu'on nous donne les 5 cartes les unes après les autres et notons P_i l'événement : "La i-ème carte est un Pique." Dans ce cas, la probabilité p d'avoir une couleur à Pique s'écrit $p = \mathbb{P}(P_1 \cap \cdots \cap P_5)$, d'où par la formule des probabilités composées :

$$p = \mathbb{P}(P_1) \times \mathbb{P}(P_2|P_1) \times \cdots \times \mathbb{P}(P_5|P_1 \cap \cdots \cap P_4).$$

Il reste à voir qu'il y a au premier tirage 13 cartes de Pique sur un total de 52, au second 12 cartes de Pique sur un total de 51, etc., ce qui donne in fine

$$p = \frac{13 \times 12 \times 11 \times 10 \times 9}{52 \times 51 \times 50 \times 49 \times 48} \approx 5 \times 10^{-4}.$$

Puisqu'il y a 4 couleurs possibles, la probabilité d'avoir une couleur est alors tout simplement $4p \approx 0.2\%$ On retrouve le résultat de l'exercice 1.5 (Las Vegas 21).

Exercice 1.22 (Le problème du dépistage)

1. Si $(\Omega, \mathcal{F}, \mathbb{P})$ est un espace probabilisé, $(H_1, \dots, H_n)$ une partition de Ω en n événements de probabilités non nulles et $A \in \mathcal{F}$ tel que $\mathbb{P}(A) > 0$, la formule de Bayes (dite de probabilité des causes) dit que pout tout j entre 1 et n :
$$\mathbb{P}(H_j|A) = \frac{\mathbb{P}(A|H_j)\mathbb{P}(H_j)}{\sum_{i=1}^n \mathbb{P}(A|H_i)\mathbb{P}(H_i)}.$$
2. Si on note T l'événement : "Le test est positif", et M l'événement : "La personne est malade", on cherche donc la probabilité $\mathbb{P}(M|T)$ et la formule de Bayes donne :
$$\mathbb{P}(M|T) = \frac{\mathbb{P}(T|M)\mathbb{P}(M)}{\mathbb{P}(T|M)\mathbb{P}(M) + \mathbb{P}(T|\overline{M})\mathbb{P}(\overline{M})}.$$
D'après l'énoncé, on a $\mathbb{P}(M) = 1/1000$, $\mathbb{P}(T|M) = 0.99$, $\mathbb{P}(T|\overline{M}) = 0.002$, les autres probabilités intervenant dans la formule de Bayes s'en déduisant facilement. Ceci donne $\mathbb{P}(M|T) \approx 1/3$. Le test n'est donc pas si fiable que ça, puisque parmi les positifs, il y a une majorité de faux positifs (voir figure 1.10). Il n'empêche qu'il peut être très utile, en pratique, pour faire une première sélection avant d'effectuer un second test plus fiable (mais plus coûteux) sur les "quelques" patients pour lesquels ce premier test est positif. Le point crucial pour ce genre de test filtre est de ne pas manquer un patient malade, c'est-à-dire d'éviter les faux négatifs. Cette proportion $\mathbb{P}(\overline{T}|M)$ est donnée par l'énoncé et vaut 1%.

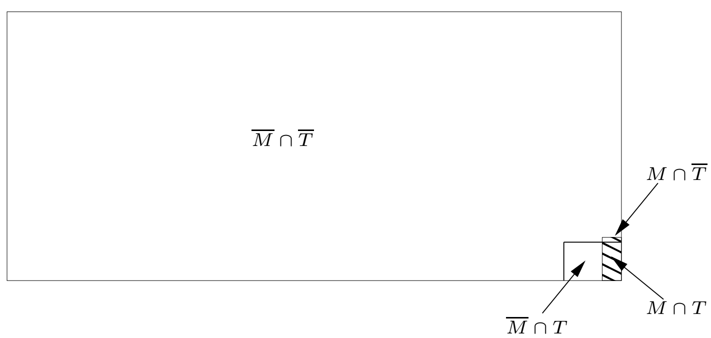


Figure 1.10 – Illustration du test de dépistage.

Exercice 1.23 (Composition de familles)

1. On a d'après le texte : $p_1 = 3p_0$, $p_2 = 4p_0$ et $p_3 = 2p_0$. Puisque la somme des p_i fait 1, on en déduit que :
$$p = [p_0, p_1, p_2, p_3] = \left[\frac{1}{10}, \frac{3}{10}, \frac{4}{10}, \frac{2}{10}\right].$$
2. Notons G l'événement : "Il y a au moins un garçon dans la famille." On cherche donc $\mathbb{P}(\overline{G})$. Nous allons utiliser la formule des probabilités totales via la partition $\Omega = \{E_0, E_1, E_2, E_3\}$ suivant le nombre d'enfants par famille :
$$\mathbb{P}(\overline{G}) = \sum_{i=0}^{3} \mathbb{P}(\overline{G}|E_i)\mathbb{P}(E_i) = \sum_{i=0}^{3} \mathbb{P}(\overline{G}|E_i)p_i,$$
où il reste à voir que pour tout i on a $\mathbb{P}(\overline{G}|E_i) = (1/2)^i$. Finalement on obtient $\mathbb{P}(\overline{G}) = 3/8$.

3. On cherche cette fois la probabilité $\mathbb{P}(E_2|\overline{G})$, il suffit d'inverser le conditionnement :

$$\mathbb{P}(E_2|\overline{G}) = \frac{\mathbb{P}(\overline{G}|E_2)\mathbb{P}(E_2)}{\mathbb{P}(\overline{G})}.$$

D'après la question précédente, on sait que $\mathbb{P}(\overline{G}) = 3/8$, et d'après la première question $\mathbb{P}(E_2) = p_2 = 4/10$. On arrive donc à $\mathbb{P}(E_2|\overline{G}) = 4/15$.

Exercice 1.24 (L'ivresse du gardien de nuit)

1. Méthode A : on appelle p_n la probabilité qu'il faille n essais pour ouvrir la porte. Puisqu'il retire chaque clé après un essai infructueux, il est clair que n peut prendre les valeurs de 1 à 10. On peut calculer les probabilités de proche en proche : la probabilité p_1 est clairement $p_1 = 1/10$. Pour qu'il ouvre la porte au deuxième essai, il faut qu'il se soit trompé au premier, ce qui arrive avec probabilité 9/10 et qu'il ait réussi au second, ce qui arrive avec probabilité 1/9, donc à nouveau $p_2 = 1/10$. En itérant ce raisonnement, on voit sans peine que pour tout n entre 1 et 10, $p_n = 1/10$. Nous parlerons dans ce cas de loi uniforme sur l'ensemble $\{1, \ldots, 10\}$.
Remarque : on pouvait obtenir ce résultat par un autre raisonnement : les 10 clés du trousseau arrivent dans un certain ordre et il n'y aucune raison que la clé qui ouvre la porte soit à une position plutôt qu'à une autre, donc le nombre d'essais nécessaires pour ouvrir la porte est équiréparti entre 1 et 10.
2. Méthode B : cette fois, le nombre n d'essais nécessaire peut prendre toute valeur de $\mathbb{N}^*$. La probabilité q_1 est à nouveau $q_1 = 1/10$. Pour qu'il ouvre la porte au deuxième essai, il faut qu'il se soit trompé au premier, ce qui arrive avec probabilité 9/10, et qu'il ait réussi au second, ce qui arrive avec probabilité 1/10, donc $q_2 = 1/10 \times 9/10$. En itérant ce raisonnement, on voit que :

$$\forall n \in \mathbb{N}^* \qquad q_n = \frac{1}{10}\left(\frac{9}{10}\right)^{n-1}.$$

On dit dans ce cas que le nombre d'essais suit une loi géométrique de paramètre 1/10.
3. Notons $\{N > 8\}$ l'événement : "Après 8 essais, la porte n'est toujours pas ouverte" et, conformément à ce qui précède, A (resp. B) l'événement : "Le gardien est à jeun (resp. ivre)." Notons au passage que $A = \overline{B}$. On cherche donc $\mathbb{P}(B|\{N > 8\})$. On utilise la formule de Bayes :

$$\mathbb{P}(B|\{N > 8\}) = \frac{\mathbb{P}(\{N > 8\}|B)\mathbb{P}(B)}{\mathbb{P}(\{N > 8\}|A)\mathbb{P}(A) + \mathbb{P}(\{N > 8\}|B)\mathbb{P}(B)}.$$

Le texte nous apprend que $\mathbb{P}(B) = 1/3$, donc $\mathbb{P}(A) = 2/3$. Avec des notations naturelles, on obtient d'une part :

$$\mathbb{P}(\{N > 8\}|B) = \sum_{n=9}^{+\infty} q_n = \frac{1}{10}\sum_{n=9}^{+\infty}\left(\frac{9}{10}\right)^{n-1} = \left(\frac{9}{10}\right)^8,$$

puisqu'on a reconnu une série géométrique de raison 9/10. Plus simple encore :

$$\mathbb{P}(\{N > 8\}|A) = p_9 + p_{10} = \frac{2}{10}.$$

Il vient donc $\mathbb{P}(B|E_8) \approx 0,518$.

Exercice 1.25 (Urne de Polya)

1. Notons B_i (respectivement N_i) l'événement : "La boule tirée au i-ème tirage est blanche (respectivement noire)." Le but est donc de calculer la probabilité $\mathbb{P}(B_2)$. Or la probabilité

de tirer une boule blanche au second tirage est facile si on connaît la composition de l'urne avant ce tirage, d'où l'idée naturelle de conditionner par ce qui s'est passé au premier tirage. C'est le principe de la formule des probabilités totales :

$$\mathbb{P}(B_2)=\mathbb{P}(B_2|B_1)\mathbb{P}(B_1)+\mathbb{P}(B_2|N_1)\mathbb{P}(N_1)=\frac{7}{13}\times\frac{4}{10}+\frac{4}{13}\times\frac{6}{10}=\frac{2}{5}.$$

2. On veut cette fois calculer $\mathbb{P}(B_2|N_1)$, ce qui peut se faire comme suit :

$$\mathbb{P}(N_1|B_2)=\frac{\mathbb{P}(B_2\cap N_1)}{\mathbb{P}(B_2)}=\frac{\mathbb{P}(B_2|N_1)\mathbb{P}(N_1)}{\mathbb{P}(B_2)}=\frac{6}{13}.$$

3. Généralisation : l'équation obtenue à la première question devient dans le cas général

$$\mathbb{P}(B_2)=\frac{B+x}{B+N+x}\times\frac{B}{B+N}+\frac{B}{B+N+x}\times\frac{N}{B+N}=\frac{B}{B+N}.$$

Exercice 1.26 (Transmission bruitée)

1. Pour que l'événement I_{n+1} ait lieu, de deux choses l'une : ou bien I_n était réalisé et le message a été bien transmis dans le $(n+1)$-ème canal, ou bien $\overline{I_n}$ était réalisé et le message a été mal transmis dans le $(n+1)$-ème canal. C'est en fait la formule des probabilités totales qui s'applique ici :
$$\mathbb{P}(I_{n+1})=\mathbb{P}(I_{n+1}|I_n)\mathbb{P}(I_n)+\mathbb{P}(I_{n+1}|\overline{I_n})\mathbb{P}(\overline{I_n}),$$
c'est-à-dire :
$$\mathbb{P}(I_{n+1})=(1-p)\mathbb{P}(I_n)+p(1-\mathbb{P}(I_n)).$$

2. On a donc la relation de récurrence :
$$p_{n+1}=(1-p)p_n+p(1-p_n)=(1-2p)p_n+p.$$
La condition initiale est $p_1=1-p$, probabilité que le message n'ait pas été bruité dans le premier canal.

3. On écrit :
$$v_{n+1}=u_{n+1}-\frac{1}{2}=(1-2p)u_n+p-\frac{1}{2},$$
et en remplaçant u_n par $v_n+\frac{1}{2}$, il vient $v_{n+1}=(1-2p)v_n$, donc la suite $(v_n)_{n\geq 1}$ est géométrique de raison $(1-2p)$. On en déduit :
$$\forall n\in\{1,\dots,N\}\qquad v_n=(1-2p)^{n-1}v_1.$$

4. On a la même relation pour p_n que pour $u_n=v_n+\frac{1}{2}$ et puisque $p_1=(1-p)$, on en déduit que :
$$\forall n\in\{1,\dots,N\}\qquad p_n=\frac{1}{2}+\left(\frac{1}{2}-p\right)(1-2p)^{n-1}.$$

5. Pour déterminer $\lim_{N\to+\infty}p_N$, on peut distinguer 3 cas :
   (a) $p=0$: la transmission est fiable et on retrouve bien sûr $p_N=1$ pour tout N.
   (b) $p=1$: chaque passage dans un canal change de façon certaine le message, donc p_N dépend de la parité du nombre de canaux : $p_{2N}=1$ et $p_{2N+1}=0$.
   (c) $0<p<1$: contrairement aux deux situations précédentes, on est dans le cas d'un bruitage aléatoire. On remarque que $\lim_{N\to+\infty}(1-2p)^{N-1}=0$ et $\lim_{N\to+\infty}p_N=\frac{1}{2}$. Ceci signifie que dès que le nombre de canaux devient grand, on est incapable de retrouver le message initial de façon fiable : autant tirer à pile ou face ! C'est un peu le principe du jeu connu sous le nom du "téléphone arabe".

Exercice 1.27 (La roulette de la lose)

1. On a bien sûr $p_0 = 1$ et $p_{100} = 0$.
2. Supposons que A commence avec n€ avec $0 < n < 100$: à la première partie, ou bien il gagne (ce qui arrive avec probabilité p) et la probabilité qu'il se ruine ensuite devient p_{n+1}, ou bien il perd (ce qui arrive avec probabilité $(1-p)$) et la probabilité qu'il se ruine ensuite devient p_{n-1}. La formule des probabilités totales s'écrit donc :
$$p_n = p \times p_{n+1} + (1-p) \times p_{n-1}.$$
3. Si pour tout $n \in \{0, \ldots, 100\}$, on admet que :
$$p_n = \alpha + \beta \left(\frac{1-p}{p}\right)^n,$$
il nous reste simplement à déterminer α et β grâce aux conditions aux bords $p_0 = 1$ et $p_{100} = 0$. Notons $\theta = \frac{1-p}{p}$ afin d'alléger les notations. On a donc à résoudre le système linéaire de deux équations à deux inconnues :
$$\left\{\begin{array}{r} \alpha + \beta = 1 \\ \alpha + \beta\theta^{100} = 0 \end{array}\right. \Longleftrightarrow \left\{\begin{array}{l} \alpha = \frac{\theta^{100}}{\theta^{100}-1} \\ \beta = \frac{-1}{\theta^{100}-1} \end{array}\right.$$
Ceci donne finalement :
$$\forall n \in \{0, \ldots, 100\} \qquad p_n = \frac{\theta^{100} - \theta^n}{\theta^{100} - 1}.$$
4. La probabilité que A finisse ruiné en commençant avec 50€ est donc $p_{50} = \frac{\theta^{100}-\theta^{50}}{\theta^{100}-1}$.
5. A la roulette, la probabilité de gain à chaque partie est $p = 18/37$, donc $\theta = 19/18$, et la probabilité de finir ruiné est : $p_{50} \approx 94\%$. Il valait mieux en effet aller se promener ce jour-là...
6. Tant qu'à être prêt à perdre 50€, le mieux (ou plutôt : le moins pire) est de les miser en une seule fois. La probabilité de finir ruiné est alors simplement $p = 18/37$.

Exercice 1.28 (Loi de succession de Laplace)

1. La probabilité cherchée s'écrit, en suivant l'indication de l'énoncé :
$$p_N = \mathbb{P}(E_{n+1}|E_n) = \frac{\mathbb{P}(E_{n+1} \cap E_n)}{\mathbb{P}(E_n)} = \frac{\mathbb{P}(E_{n+1})}{\mathbb{P}(E_n)},$$
la dernière égalité venant de ce que $E_{n+1} \subseteq E_n$. Les deux termes se traitent alors de la même façon, en décomposant sur la partition $\{U_0, \ldots, U_N\}$:
$$\mathbb{P}(E_n) = \sum_{k=0}^{N} \mathbb{P}(E_n|U_k)\mathbb{P}(U_k) = \frac{1}{N+1} \sum_{k=0}^{N} \mathbb{P}(E_n|U_k),$$
le terme $\frac{1}{N+1}$ venant de l'équiprobabilité pour le choix de l'urne dans laquelle on pioche. Il reste à voir que si on pioche dans l'urne U_k, la probabilité de tirer 1 boule rouge est k/N donc la probabilité de tirer n boules rouges à la suite est $(k/N)^n$. On a donc :
$$p_N = \frac{\frac{1}{N+1}\sum_{k=0}^{N}(k/N)^{n+1}}{\frac{1}{N+1}\sum_{k=0}^{N}(k/N)^n}.$$

2. Pour trouver la limite de (p_N) lorsque le nombre N d'urnes tend vers l'infini, il suffit d'appliquer le résultat sur les sommes de Riemann :
$$\frac{1}{N+1}\sum_{k=0}^{N}(k/N)^n = \frac{N}{N+1}\left(\frac{1}{N}\sum_{k=1}^{N}(k/N)^n\right) \xrightarrow[N\to\infty]{} \int_0^1 x^n dx = \frac{1}{n+1}.$$
On en déduit :
$$\lim_{N\to\infty} p_N = \frac{n+1}{n+2}.$$

Exercice 1.29 (Il Padrino)

1. A et B sont indépendants si et seulement si $\mathbb{P}(A\cap B) = \mathbb{P}(A)\mathbb{P}(B) = 0.09$. Or dans le cas général on sait par la formule d'additivité forte que :
$$\mathbb{P}(A\cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A\cup B) = 0.1 + 0.9 - 0.91 = 0.09,$$
donc A et B sont bien indépendants.
2. (a) Dans le cas d'un paquet groupé, dire qu'un cadeau au moins est bien arrivé signifie que le colis est bien arrivé, donc une probabilité $p = 0.9$. Dans le cas de deux paquets séparés indépendants, on cherche la probabilité complémentaire, c'est-à-dire la probabilité qu'aucun paquet n'arrive : $1 - p' = 0.1 \times 0.1 = 0.01$, donc la probabilité qu'au moins un des deux arrive est $p' = 0.99$.
   (b) Dans le cas d'un paquet groupé, la probabilité est la même qu'en question précédente, c'est-à-dire $p = 0.9$. Dans le cas de deux paquets séparés indépendants, la probabilité est cette fois $p'' = 0.9 \times 0.9 = 0.81$.
3. On considère trois événements (mutuellement) indépendants A, B et C tels que $\mathbb{P}(A) = 0.8$, $\mathbb{P}(B) = 0.5$ et $\mathbb{P}(C) = 0.2$. Pour calculer $\mathbb{P}(A\cup B\cup C)$, deux solutions : ou bien on passe par la formule d'inclusion-exclusion (Poincaré) :
$$\mathbb{P}(A\cup B\cup C) = \mathbb{P}(A)+\mathbb{P}(B)+\mathbb{P}(C)-(\mathbb{P}(A\cap B)+\mathbb{P}(A\cap C)+\mathbb{P}(B\cap C))+\mathbb{P}(A\cap B\cap C)$$
et on utilise l'indépendance pour évaluer les probabilités des intersections. Ou bien on passe par le complémentaire :
$$\mathbb{P}(A\cup B\cup C) = 1 - \mathbb{P}(\overline{A\cup B\cup C}) = 1 - \mathbb{P}(\overline{A}\cap\overline{B}\cap\overline{C}),$$
en se souvenant que l'indépendance de A, B et C équivaut à l'indépendance de leurs complémentaires, d'où :
$$\mathbb{P}(A\cup B\cup C) = 1 - \mathbb{P}(\overline{A})\mathbb{P}(\overline{B})\mathbb{P}(\overline{C}),$$
c'est-à-dire :
$$\mathbb{P}(A\cup B\cup C) = 1 - (1-\mathbb{P}(A))(1-\mathbb{P}(B))(1-\mathbb{P}(C)) = 0.92.$$

Exercice 1.30 (Circuit électrique)

Avec des notations évidentes, on cherche en fait $\mathbb{P}(A\cup B\cup C)$, avec A, B et C indépendants. On applique donc la méthode de l'exercice précédent :
$$\mathbb{P}(A\cup B\cup C) = 1 - (1-\mathbb{P}(A))(1-\mathbb{P}(B))(1-\mathbb{P}(C)).$$
Il reste à voir que par indépendance sur chaque branche on a $\mathbb{P}(A) = \mathbb{P}(A_1)\mathbb{P}(A_2) = 0.45$, $\mathbb{P}(B) = \mathbb{P}(B_1)\mathbb{P}(B_2)\mathbb{P}(B_3) = 0.108$ et $\mathbb{P}(C) = \mathbb{P}(C_1) = 0.7$, ce qui donne au total $\mathbb{P}(A\cup B\cup C) \approx 0.85$.

Exercice 1.31 (Le bandit manchot)

1. Notons p_3 la probabilité de remporter le jackpot, alors :
$$p_3 = \frac{1}{20} \times \frac{9}{20} \times \frac{1}{20} = \frac{9}{8000}.$$
2. Notons p_2 la probabilité d'avoir 2 cloches mais pas le jackpot, alors par le même raisonnement :
$$p_2 = \frac{1 \times 9 \times 19}{8000} + \frac{1 \times 11 \times 1}{8000} + \frac{19 \times 9 \times 1}{8000} = \frac{353}{8000}.$$
3. Si au lieu d'une répartition 1-9-1 des cloches, il y a une répartition 3-1-3, notons p'_3 et p'_2 les nouvelles probabilités. On trouve cette fois :
$$p'_3 = \frac{3}{20} \times \frac{1}{20} \times \frac{3}{20} = \frac{9}{8000},$$
et
$$p'_2 = \frac{3 \times 1 \times 17}{8000} + \frac{3 \times 19 \times 3}{8000} + \frac{17 \times 1 \times 3}{8000} = \frac{273}{8000}.$$
Ainsi la probabilité de remporter le jackpot est inchangée (en moyenne le propriétaire du casino gagne autant de ce point de vue), mais la configuration avec seulement 2 cloches apparaît plus souvent avec le premier système, ce qui peut encourager un client naïf à continuer de jouer. Le propriétaire du casino optera donc plutôt pour la répartition 1-9-1 que 3-1-3.

Exercice 1.32 (Les affres des escales)

Notons tout d'abord A l'événement : “Absence de la valise à Paris”. Notons encore LA (resp. NY et L) l'événement : “La valise est restée à Los Angeles (resp. New York et Londres.)” On cherche donc $\mathbb{P}(LA|A)$ (resp. $\mathbb{P}(NY|A)$ et $\mathbb{P}(L|A)$). Pour la probabilité que la valise soit restée à Los Angeles, il suffit d'inverser le conditionnement :
$$\mathbb{P}(LA|A) = \frac{\mathbb{P}(A|LA)\mathbb{P}(LA)}{\mathbb{P}(A)}.$$
On a $\mathbb{P}(A|LA) = 1$, $\mathbb{P}(LA) = p$ et la probabilité que le bagage n'arrive pas à Paris se calcule facilement par passage au complémentaire : $\mathbb{P}(A) = 1 - \mathbb{P}(\overline{A}) = 1 - (1-p)^3$. Tout ceci donne au total :
$$\mathbb{P}(LA|A) = \frac{p}{1-(1-p)^3}.$$
On raisonne de même pour calculer la probabilité que le bagage soit resté à New York :
$$\mathbb{P}(NY|A) = \frac{\mathbb{P}(A|NY)\mathbb{P}(NY)}{\mathbb{P}(A)},$$
avec à nouveau $\mathbb{P}(A|NY) = 1$ et $\mathbb{P}(A) = 1 - (1-p)^3$. Il reste à voir que pour que le bagage soit resté à New York, il a dû quitter Los Angeles, ce qui arrive avec probabilité $(1-p)$ puis être resté à New York, ce qui arrive avec probabilité p. Ainsi $\mathbb{P}(NY) = p(1-p)$ et :
$$\mathbb{P}(NY|A) = \frac{p(1-p)}{1-(1-p)^3}.$$
De la même façon, la probabilité qu'il soit resté à Londres est :
$$\mathbb{P}(L|A) = \frac{p(1-p)^2}{1-(1-p)^3}.$$
Remarquons qu'au total on a bien :
$$\mathbb{P}(LA|A) + \mathbb{P}(NY|A) + \mathbb{P}(L|A) = 1.$$

Exercice 1.33 (Une histoire de montres)
Notons simplement "1" (resp. "2") l'événement : "La première (resp. la deuxième) montre marche", ainsi que "HK" (resp. "S") l'événement : "Le lot vient de Hong-Kong (resp. de Singapour)." On cherche donc $\mathbb{P}(2|1)$ et on commence par revenir à la définition de la probabilité conditionnelle :

$$\mathbb{P}(2|1) = \frac{\mathbb{P}(2 \cap 1)}{\mathbb{P}(1)}.$$

Pour calculer $\mathbb{P}(1)$ on applique alors la formule des probabilités totales avec la partition (HK, S) :

$$\mathbb{P}(1) = \mathbb{P}(1|HK)\mathbb{P}(HK) + \mathbb{P}(1|S)\mathbb{P}(S).$$

En l'absence d'information, on suppose que le lot de montres a autant de chances de provenir de Hong-Kong que de Singapour, donc $\mathbb{P}(HK) = \mathbb{P}(S) = 1/2$. Ainsi on a :

$$\mathbb{P}(1) = \frac{999}{1000} \times \frac{1}{2} + \frac{199}{200} \times \frac{1}{2} = \frac{997}{1000}.$$

On procède de la même façon pour calculer $\mathbb{P}(2 \cap 1)$:

$$\mathbb{P}(2 \cap 1) = \mathbb{P}(2 \cap 1|HK)\mathbb{P}(HK) + \mathbb{P}(2 \cap 1|S)\mathbb{P}(S).$$

Il reste à voir qu'une fois connue la provenance du lot, les événements "1" et "2" sont indépendants, ce qui donne par exemple :

$$\mathbb{P}(2 \cap 1|HK) = \mathbb{P}(2|HK)\mathbb{P}(1|HK) = \left(\frac{999}{1000}\right)^2.$$

Ainsi on arrive à :

$$\mathbb{P}(2 \cap 1) = \frac{1}{2}\left(\left(\frac{999}{1000}\right)^2 + \left(\frac{199}{200}\right)^2\right).$$

Finalement la probabilité cherchée est :

$$P(2|1) = \frac{\frac{1}{2}\left(\left(\frac{999}{1000}\right)^2 + \left(\frac{199}{200}\right)^2\right)}{\frac{997}{1000}} = \frac{994013}{997000} \approx 0.997.$$

Exercice 1.34 (Un éléphant ça trompe énormément)
Notons C_i l'événement : "Le chasseur i a atteint l'éléphant", et "2" l'événement : "L'éléphant a reçu deux balles." On commence donc par chercher $\mathbb{P}(\overline{C_1}|2)$, qui s'écrit encore :

$$\mathbb{P}(\overline{C_1}|2) = \frac{\mathbb{P}(\overline{C_1} \cap 2)}{\mathbb{P}(2)}.$$

Le numérateur $\mathbb{P}(\overline{C_1} \cap 2)$ est la probabilité que l'éléphant ait été tué de deux balles et que le premier chasseur l'ait manqué, c'est-à-dire :

$$\mathbb{P}(\overline{C_1} \cap 2) = \mathbb{P}(\overline{C_1} \cap C_2 \cap C_3) = \mathbb{P}(\overline{C_1})\mathbb{P}(C_2)\mathbb{P}(C_3) = \frac{9}{32},$$

le passage de l'intersection au produit venant du fait que les tirs sont indépendants. Il reste le dénominateur, pour lequel on décompose l'événement "2" en disant que pour que l'éléphant ait été atteint exactement deux fois, il faut que deux chasseurs l'aient atteint et que le troisième l'ait manqué, ce qui s'écrit :

$$\mathbb{P}(2) = \mathbb{P}(\overline{C_1} \cap C_2 \cap C_3) + \mathbb{P}(C_1 \cap \overline{C_2} \cap C_3) + \mathbb{P}(C_1 \cap C_2 \cap \overline{C_3}),$$

qui se calcule à nouveau via l'indépendance des trois tirs :

$$\mathbb{P}(2) = \frac{3}{4} \times \frac{1}{2} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{2} \times \frac{3}{4} + \frac{1}{4} \times \frac{1}{2} \times \frac{1}{4} = \frac{13}{32}.$$

Finalement on a $\mathbb{P}(\overline{C_1}|2) = 9/13$. Le même raisonnement donne $\mathbb{P}(\overline{C_2}|2) = 3/13$ et $\mathbb{P}(\overline{C_3}|2) = 1/13$. On constate qu'on a bien :

$$\mathbb{P}(\overline{C_1}|2) + \mathbb{P}(\overline{C_2}|2) + \mathbb{P}(\overline{C_3}|2) = 1.$$

Exercice 1.35 (Une urne à composition variable)

1. La probabilité p_n est la probabilité d'avoir tiré 5 boules noires parmi les 10 et 5 boules blanches parmi les n, et ce parmi $\binom{n+10}{10}$ tirages possibles, donc :

$$p_n = \frac{\binom{10}{5}\binom{n}{5}}{\binom{n+10}{10}}.$$

2. Pour tout $n \geq 5$, on obtient bien via la formule précédente :

$$\frac{p_{n+1}}{p_n} = \frac{n^2 + 2n + 1}{n^2 + 7n - 44}.$$

3. Pour connaître les variations de la suite de termes positifs $(p_n)_{n\geq 5}$, il suffit de comparer le rapport p_{n+1}/p_n à 1. On a :

$$\frac{p_{n+1}}{p_n} \leq 1 \;\Leftrightarrow\; n^2 + 2n + 1 \leq n^2 + 7n - 44 \;\Leftrightarrow\; n \geq 9.$$

Remarquons qu'on a plus précisément $p_9 = p_{10}$, ainsi $(p_n)_{n\geq 5}$ est strictement croissante jusqu'à $n = 9$ et strictement décroissante après $n = 10$:

$$p_5 < p_6 < \cdots < p_9 = p_{10} > p_{11} > p_{12} > \ldots$$

Le maximum est atteint pour $n = 9$ et $n = 10$: $p_9 = p_{10} \approx 0,34$.

Exercice 1.36 (Les paris plus ou moins vaseux du Chevalier de Méré)

1. Première règle : "Il est avantageux de parier sur l'apparition d'au moins un 6 en lançant un dé quatre fois de suite". Il suffit de montrer que la probabilité p d'obtenir au moins un 6 sur quatre lancers est supérieure à 1/2. Or la probabilité de n'en obtenir aucun est :

$$1 - p = (5/6)^4 \approx 0.482 \implies p \approx 0.518$$

Cette règle est donc bien avantageuse en moyenne.

2. Seconde règle : "Il est avantageux de parier sur l'apparition d'au moins un double 6 en lançant deux dés vingt-quatre fois de suite". Notons p la probabilité d'apparition au moins un double 6 en 24 lancers. Par le même raisonnement que ci-dessus, on a :

$$1 - p = (35/36)^{24} \approx 0.509 \implies p \approx 0.491$$

Cette règle n'est donc pas avantageuse en moyenne.

Exercice 1.37 (Tirages uniformes sur un segment)

1. La probabilité que le nombre tiré soit supérieur à 3/4 est égale à la longueur du segment $[3/4, 1]$, c'est-à-dire à 1/4.

2. La probabilité qu'il soit supérieur à $3/4$ sachant qu'il est supérieur à $1/3$ est égale au rapport entre la longueur du segment $[3/4, 1]$ et celle du segment $[1/3, 1]$, c'est-à-dire à $3/8$.
3. On tire deux points au hasard sur le segment $[0, 1]$, indépendamment l'un de l'autre.
   (a) La probabilité que le plus petit des deux nombres soit supérieur à $1/3$ est égale à la probabilité que les deux nombres soient supérieurs à $1/3$, donc $p = 2/3 \times 2/3 = 4/9$.
   (b) La probabilité p_0 que le plus grand des deux nombres soit supérieur à $3/4$, sachant que le plus petit des deux est supérieur à $1/3$ s'écrit $p_0 = p_1/p$, où p_1 est la probabilité que le plus petit des deux est supérieur à $1/3$ et le plus grand des deux est supérieur à $3/4$. p_1 est donc encore la probabilité que les deux nombres soient supérieurs à $1/3$ moins la probabilité qu'ils soient tous les deux entre $1/3$ et $3/4$:
$$p_1 = \mathbb{P}([1/3, 1], [1/3, 1]) - \mathbb{P}([1/3, 3/4], [1/3, 3/4]) = \frac{2}{3} \times \frac{2}{3} - \frac{5}{12} \times \frac{5}{12} = \frac{13}{48},$$
   d'où finalement : $p_0 = \frac{39}{64}$.

Exercice 1.38 (La loi du minimum)

1. La probabilité P_k que le plus petit des numéros obtenus soit supérieur ou égal à k est la probabilité que tous les numéros obtenus soient supérieurs ou égaux à k, c'est-à-dire $P_k = \left(\frac{n-k+1}{n}\right)^N$.
2. La probabilité p_k que le plus petit des numéros obtenus soit égal à k est la probabilité que le plus petit des numéros obtenus soit supérieur ou égal à k moins la probabilité qu'il soit supérieur ou égal à $(k + 1)$, soit :
$$\forall k \in \{1, \dots, n\} \qquad p_k = P_k - P_{k+1} = \left(\frac{n-k+1}{n}\right)^N - \left(\frac{n-k}{n}\right)^N.$$
3. Si on ne fait pas de remise entre les N tirages, alors avec les mêmes notations que ci-dessus, on obtient tout d'abord
$$P_k = \begin{cases} \frac{\binom{n-k+1}{N}}{\binom{n}{N}} & \text{si } N \leq n-k+1 \\ 0 & \text{si } N > n-k+1 \end{cases}$$
   D'où l'on déduit :
$$p_k = \begin{cases} \frac{\binom{n-k+1}{N} - \binom{n-k}{N}}{\binom{n}{N}} = \frac{\binom{n-k}{N-1}}{\binom{n}{N}} & \text{si } N \leq n-k+1 \\ 0 & \text{si } N > n-k+1 \end{cases}$$

Exercice 1.39 (Fratrie)

Cet exercice peut se traiter de façon intuitive à fond de cinquième, mais puisqu'il induit parfois en erreur, allons-y piano (accentuer la première syllabe). Notons donc FG l'événement : "Le premier enfant est une fille, le second un garçon." Vu les hypothèses, dans les deux questions, nous avons donc une partition de l'espace fondamental Ω en quatre événements équiprobables : $\Omega = FF \cup FG \cup GF \cup GG$.

1. La probabilité cherchée s'écrit ici :
$$p_1 = \mathbb{P}(FG \cup GF | FG \cup GF \cup GG) = \frac{\mathbb{P}(FG \cup GF)}{\mathbb{P}(FG \cup GF \cup GG)},$$
   et toutes les unions étant disjointes, il vient :
$$p_1 = \frac{\mathbb{P}(FG)}{\mathbb{P}(FG) + \mathbb{P}(GF) + \mathbb{P}(GG)} + \frac{\mathbb{P}(GF)}{\mathbb{P}(FG) + \mathbb{P}(GF) + \mathbb{P}(GG)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}} + \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}} = \frac{2}{3}.$$

2. Cette fois la probabilité cherchée s'écrit :

$$p_2 = \mathbb{P}(FF|FG \cup FF) = \frac{\mathbb{P}(FF)}{\mathbb{P}(FG \cup FF)} = \frac{\mathbb{P}(FF)}{\mathbb{P}(FG) + \mathbb{P}(FF)} = \frac{\frac{1}{4}}{\frac{1}{4} + \frac{1}{4}} = \frac{1}{2}.$$

Exercice 1.40 (Liouville et les probabilités)
Adoptons la notation R_i (respectivement N_i) : "Au i-ème tirage, la boule obtenue est rouge (resp. noire)." En analysant les parties possibles, on voit que la probabilité que A gagne s'écrit :

$$p_a = \mathbb{P}(R_1 \cup (N_1 \cap R_2 \cap R_3)) = \mathbb{P}(R_1) + \mathbb{P}(N_1 \cap R_2 \cap R_3).$$

Le premier terme vaut clairement $2/5$. Quant au second, il peut se calculer par conditionnements successifs :

$$\mathbb{P}(N_1 \cap R_2 \cap R_3) = \mathbb{P}(N_1)\mathbb{P}(R_2|N_1)\mathbb{P}(R_3|N_1 \cap R_2) = \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} = \frac{1}{10}.$$

Finalement, $p_a = 1/2$. Pour autant, le jeu n'est pas équitable puisqu'il y a possibilité de match nul :

$$p_n = \mathbb{P}(N_1 \cap R_2 \cap N_3 \cap R_4) = \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} \times \frac{1}{2} = \frac{1}{10}.$$

Ainsi B est désavantagé, n'ayant qu'une probabilité $2/5$ de gagner la partie.

Exercice 1.41 (Pierre-feuille-ciseaux)

1. Avec des notations claires, la probabilité que A batte B est :

$$\mathbb{P}(A \succ B) = \mathbb{P}(A_6 \cup (A_3 \cap B_2)) = \mathbb{P}(A_6) + \mathbb{P}(A_3)\mathbb{P}(B_2) = \frac{1}{6} + \frac{5}{6} \times \frac{3}{6} = \frac{7}{12}.$$

2. De même, la probabilité que B batte C est :

$$\mathbb{P}(B \succ C) = \mathbb{P}(B_5 \cup (B_2 \cap C_1)) = \mathbb{P}(B_5) + \mathbb{P}(B_2)\mathbb{P}(C_1) = \frac{3}{6} + \frac{3}{6} \times \frac{1}{6} = \frac{7}{12}.$$

3. Si la relation "$\succ$" entre ces dés était transitive, on choisirait le dé A. Mais il n'en est rien, comme le montre le calcul :

$$\mathbb{P}(C \succ A) = \mathbb{P}(C_4 \cap A_3) = \frac{5}{6} \times \frac{5}{6} = \frac{25}{36} > \frac{1}{2}.$$

Ceci explique le titre de l'exercice. Mieux, on peut vérifier que si chaque joueur lance deux dés identiques et effectue la somme, alors on obtient à nouveau une relation non transitive, mais tout est inversé ! A titre d'exemple, nous avons :

$$\mathbb{P}(CC \succ AA) = \mathbb{P}(C_4 \cap C_4 \cap A_3 \cap A_3) = \left(\frac{5}{6}\right)^4 \approx 0.48.$$

Exercice 1.42 (Match de tennis)

1. Partant de 40-40 (ou de l'égalité, ce qui revient au même), que peut-il se passer ? Ou bien le joueur gagne l'échange, ce qui arrive avec probabilité $2/3$, et la probabilité qu'il gagne ensuite le jeu est P^+ ; ou bien le joueur perd l'échange, ce qui arrive avec probabilité $1/3$, et la probabilité qu'il gagne ensuite le jeu est P_-. Pour résumer, nous avons obtenu l'équation :

$$P = \frac{2}{3}P^+ + \frac{1}{3}P_-$$

En raisonnant de même en partant respectivement de “avantage pour le joueur” et “avantage pour son adversaire”, on aboutit finalement au système d'équations :

$$\begin{cases} P^+ = \frac{2}{3} + \frac{1}{3}P \\ P = \frac{2}{3}P^+ + \frac{1}{3}P_- \\ P_- = \frac{2}{3}P \end{cases}$$

lequel se résout sans difficulté :

$$P = \frac{2}{3}\left(\frac{2}{3} + \frac{1}{3}P\right) + \frac{1}{3} \times \frac{2}{3}P \ \Rightarrow P = \frac{4}{5}.$$

Remarque : il est possible d'arriver brutalement au même résultat en décomposant toutes les possibilités de gain du jeu à partir de 40-40 :

$$P = \mathbb{P}(GG \cup GPGG \cup PGGG \cup GPGPGG \cup GPPGGG \cup PGGPGG \cup PGPGGG \cup \dots)$$

Le motif est limpide : le gain du jeu se décompose en une séquence de n couples GP ou PG, conclu par le couple GG. Puisqu'il y a deux choix pour chaque couple, le nombre de séquences possibles de longueur n est 2^n. Il reste à voir que $\mathbb{P}(PG) = \mathbb{P}(GP) = 2/9$ et $\mathbb{P}(GG) = 4/9$ pour arriver à une brave série géométrique :

$$P = \sum_{n=0}^{+\infty} 2^n \times \frac{4}{9}\left(\frac{2}{9}\right)^n = \frac{4}{9}\sum_{n=0}^{+\infty}\left(\frac{4}{9}\right)^n = \frac{4}{5}.$$

2. La probabilité p_3 d'arriver à 40-40 correspond à la probabilité de 6 échanges parmi lesquels 3 ont été remportés par le joueur, les 3 autres par son adversaire. Puisqu'il y a $\binom{6}{3}$ combinaisons de la sorte, on en déduit :

$$p_3 = \binom{6}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^3 = \frac{160}{729}.$$

3. La probabilité p_2 que le joueur gagne le jeu en arrivant à 40-30 et en concluant s'obtient par le même raisonnement :

$$p_2 = \frac{2}{3} \times \binom{5}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^2 = \frac{160}{729}.$$

La probabilité p_1 que le joueur gagne le jeu en arrivant à 40-15 et en concluant :

$$p_1 = \frac{2}{3} \times \binom{4}{3}\left(\frac{2}{3}\right)^3\left(\frac{1}{3}\right)^1 = \frac{64}{243}.$$

La probabilité p_0 que le joueur gagne un jeu blanc :

$$p_0 = \left(\frac{2}{3}\right)^4 = \frac{16}{81}.$$

4. La probabilité P_G que le joueur gagne le set se déduit des calculs précédents :

$$P_G = P \times p_3 + p_2 + p_1 + p_0 = \frac{208}{243} \approx 0.856.$$

5. Tout ce qui précède se généralise en remplaçant 2/3 par p et 1/3 par $q = 1 - p$:

$$P_G = \varphi(p) = \frac{p^2}{1-2pq} \times 20p^3q^3 + 10p^4q^2 + 4p^4q + p^4 = p^4\left(\frac{20pq^3}{1-2pq} + 10q^2 + 4q + 1\right).$$

Remarque : on vérifie bien que $\varphi(0) = 0$, $\varphi(1) = 1$, $\varphi(1/2) = 1/2$, et de façon générale $\varphi(1-p) = 1 - \varphi(p)$. Cette dernière propriété signifie simplement que le graphe de la fonction φ admet $(1/2, 1/2)$ comme centre de symétrie (voir figure 1.11).

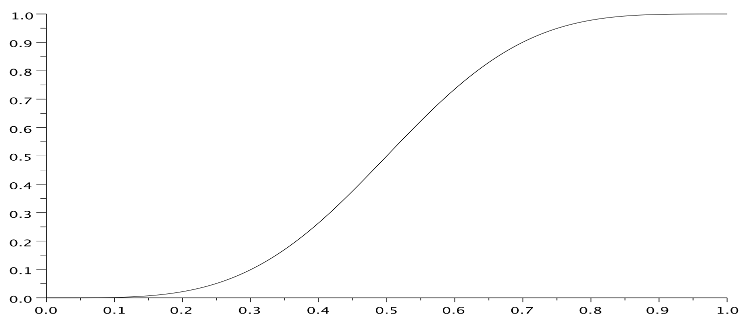


FIGURE 1.11 – Probabilité de gagner le jeu en fonction de la probabilité de gagner le point.

Exercice 1.43 (Let's make a deal)
Supposons, sans perte de généralité, la configuration suivante : (V,C,C), c'est-à-dire que la voiture est derrière la porte 1, les chèvres derrière les portes 2 et 3. Le jeu se déroule alors comme suit :

1. Sans changement de porte :
   (a) le spectateur choisit la porte 1, donc l'animateur ouvre indifféremment l'une des deux autres portes, et le spectateur gagne.
   (b) le spectateur choisit la porte 2, donc l'animateur ouvre la porte 3, et le spectateur perd.
   (c) le spectateur choisit la porte 3, donc l'animateur ouvre la porte 2, et le spectateur perd.
2. Avec changement de porte :
   (a) le spectateur choisit la porte 1, l'animateur ouvre indifféremment l'une des deux autres portes, le spectateur ouvre l'autre et perd.
   (b) le spectateur choisit la porte 2, donc l'animateur ouvre la porte 3, le spectateur ouvre la porte 1 et gagne.
   (c) le spectateur choisit la porte 3, donc l'animateur ouvre la porte 2, le spectateur ouvre la porte 1 et gagne.

Bilan des courses : s'il change de porte, il gagne 2 fois sur 3, sinon seulement 1 fois sur 3. Il vaut donc mieux changer de porte !

Exercice 1.44 (Newton & Galilée)

1. Soit p la probabilité d'obtenir au moins un 6 lorsqu'on lance 6 fois un dé, alors en passant à l'événement complémentaire, on obtient par indépendance des lancers

$$p = 1 - \mathbb{P}(\text{aucun } 6) = 1 - \left(\frac{5}{6}\right)^6 \approx 0.665$$

   Soit q la probabilité d'obtenir au moins un 6 lorsqu'on lance 6 fois un dé, alors en passant à l'événement complémentaire, on obtient cette fois

$$q = 1-\mathbb{P}(\{\text{aucun } 6\}\cup\{\text{un } 6\}) = 1-(\mathbb{P}(\text{aucun } 6)+\mathbb{P}(\text{un } 6)) = 1-\left(\frac{5}{6}\right)^{12}-12\left(\frac{5}{6}\right)^{11}\left(\frac{1}{6}\right)^{1}$$

   soit $q \approx 0.619$. On a donc une probabilité plus grande d'obtenir au moins un 6 en 6 lancers qu'au moins deux 6 en 12 lancers.

2. Sans perte de généralité, supposons les 3 dés discernables et notons $\Omega = \{(i,j,k), 1 \leq i,j,k \leq 6\}$ l'ensemble fondamental. Son cardinal est donc égal à $6^3 = 216$ et, les dés étant équilibrés, chaque triplet (i,j,k) a la même probabilité $\frac{1}{216}$ d'apparaître. Il suffit alors de faire attention aux répétitions éventuelles d'un même chiffre dans le triplet pour pouvoir conclure. En effet, on a par exemple :
$$\mathbb{P}(1+2+6) = \frac{3!}{216} = \frac{6}{216},$$
tandis que
$$\mathbb{P}(1+4+4) = \frac{3}{216}$$
et
$$\mathbb{P}(3+3+3) = \frac{1}{216}.$$
Une fois ceci remarqué, le résultat s'en déduit sans problème :
$$\mathbb{P}(9) = \frac{6+6+3+3+6+1}{216} = \frac{25}{216} \approx 0.116,$$
qui est bien inférieur à
$$\mathbb{P}(10) = \frac{6+6+3+6+3+3}{216} = \frac{27}{216} = \frac{1}{8} = 0.125.$$

Exercice 1.45 (Peer-to-Peer)

1. Avec des notations évidentes, la probabilité que ce fichier soit défectueux est
$$\mathbb{P}(D) = \mathbb{P}(D|S_1)\mathbb{P}(S_1) + \mathbb{P}(D|S_2)\mathbb{P}(S_2) + \mathbb{P}(D|S_3)\mathbb{P}(S_3) + \mathbb{P}(D|S_4)\mathbb{P}(S_4),$$
c'est-à-dire
$$\mathbb{P}(D) = 0.02 \times 0.2 + 0.02 \times 0.2 + 0.06 \times 0.2 + 0.08 \times 0.4 = 0.052$$
2. Sachant que le fichier est défectueux, la probabilité qu'il provienne du serveur S_4 est donc par la formule de Bayes
$$\mathbb{P}(S_4|D) = \frac{\mathbb{P}(D|S_4)\mathbb{P}(S_4)}{\mathbb{P}(D)} = \frac{0.08 \times 0.4}{0.052} = 0.615$$

Exercice 1.46 (Hémophilie)

On note H (respectivement $\overline{H}$) le fait que la reine soit hémophile (respectivement qu'elle ne le soit pas). De même on note H_i ou $\overline{H_i}$ selon que le i-ème fils est hémophile ou non.

1. La probabilité que la reine ait un fils hémophile vaut
$$\mathbb{P}(H_1) = P(H_1|H)\mathbb{P}(H) + \mathbb{P}(H_1|\overline{H})\mathbb{P}(\overline{H}) = \frac{1}{2} \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{1}{4}$$
car d'après le texte on sait que $\mathbb{P}(H_1|H) = \mathbb{P}(H) = \mathbb{P}(\overline{H}) = 1/2$ et $\mathbb{P}(H_1|\overline{H}) = 0$.
2. La probabilité que la reine soit porteuse du gène sachant qu'elle a eu un fils non hémophile s'écrit
$$\mathbb{P}(H|\overline{H_1}) = \frac{\mathbb{P}(\overline{H_1}|H)\mathbb{P}(H)}{\mathbb{P}(\overline{H_1})},$$
or $\mathbb{P}(\overline{H_1}|H) = 1/2$ et par la question précédente
$$\mathbb{P}(\overline{H_1}) = 1 - \mathbb{P}(H_1) = \frac{3}{4}.$$
Au total, la probabilité cherchée vaut donc $\mathbb{P}(H|\overline{H_1}) = 1/3$, quantité logiquement inférieure à $1/2$: le fait de savoir que la reine a eu un fils non hémophile diminue le risque qu'elle soit elle-même hémophile.

3. Nous voulons cette fois calculer

$$\mathbb{P}(H_2|\overline{H_1}) = \frac{\mathbb{P}(H_2 \cap \overline{H_1})}{\mathbb{P}(\overline{H_1})}$$

Le dénominateur a été calculé en question précédente. Pour le numérateur, on utilise à nouveau la formule des probabilités totales :

$$\mathbb{P}(H_2 \cap \overline{H_1}) = \mathbb{P}(H_2 \cap \overline{H_1}|H)\mathbb{P}(H) + \mathbb{P}(H_2 \cap \overline{H_1}|\overline{H})\mathbb{P}(\overline{H})$$

Lorsque la reine ne porte pas le gène, aucun enfant n'est hémophile, donc $\mathbb{P}(H_2 \cap \overline{H_1}|\overline{H}) = 0$. Si par contre elle porte le gène, alors on tient compte de l'indépendance entre enfants vis-à-vis de la maladie, ce qui donne

$$\mathbb{P}(H_2 \cap \overline{H_1}|H) = \mathbb{P}(H_2|H)\mathbb{P}(\overline{H_1}|H) = \frac{1}{4}.$$

Finalement on arrive à : $\mathbb{P}(H_2 \cap \overline{H_1}) = 1/6$. À nouveau, il est cohérent d'obtenir une probabilité inférieure à $\mathbb{P}(H_1)$, probabilité que le premier enfant soit hémophile en l'absence de toute autre information.

Exercice 1.47 (Dénombrements en vrac)
Cet exercice est corrigé en annexe (sujet d'octobre 2009).

Exercice 1.48 (Urnes, cartes et dés)
Cet exercice est corrigé en annexe (sujet d'octobre 2009).

Exercice 1.49 (Evénements indépendants)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 1.50 (Un tirage en deux temps)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 1.51 (Pièces défectueuses)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 1.52 (Circuits intégrés)
Cet exercice est corrigé en annexe (sujet de novembre 2011).

Exercice 1.53 (Utilité d'un testeur)
Cet exercice est corrigé en annexe (sujet de novembre 2011).

Chapitre 2

Variables aléatoires discrètes

Introduction

Lorsque le résultat d'une expérience où intervient le hasard est à valeurs dans un ensemble au plus dénombrable, on parle de variable aléatoire discrète. Celles-ci sont complètement caractérisées par les valeurs qu'elles peuvent prendre et les probabilités avec lesquelles elles les prennent. On définit alors facilement diverses notions utiles en calcul des probabilités : fonction de répartition, espérance, variance, indépendance, etc.

2.1 Loi d'une variable discrète

Dans toute la suite, $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé. On rappelle qu'un ensemble $\mathcal{X} = (x_i)_{i \in I}$ est au plus dénombrable s'il est fini ou dénombrable, c'est-à-dire si on peut énumérer tous ses éléments sous la forme d'une séquence finie ou infinie. Dans toute la suite, $\mathcal{X}$ sera typiquement un sous-ensemble fini de $\mathbb{N}$ ou $\mathbb{N}$ tout entier. La définition qui suit peut sembler un peu abstruse, mais donne le bon cadre pour manipuler des quantités aléatoires.

Définition 2.1 (Variable aléatoire discrète)
Soit $\mathcal{X} = (x_i)_{i \in I}$ un ensemble au plus dénombrable contenu dans $\mathbb{R}$. Une application

$$X : \left\{ \begin{array}{rl} (\Omega, \mathcal{F}, \mathbb{P}) & \to \mathcal{X} = (x_i)_{i \in I} \\ \omega & \mapsto X(\omega) \end{array} \right.$$

est une variable aléatoire discrète si

$$\forall i \in I \qquad \{X = x_i\} := X^{-1}(\{x_i\}) = \{\omega \in \Omega \ : \ X(\omega) = x_i\} \in \mathcal{F}.$$

Justifions brièvement le pourquoi de cette affaire : on aura besoin des probabilités du style $\mathbb{P}(X = x_i)$, ou de probabilités faisant intervenir des unions d'événements $\{X = x_i\}$. Un pré-requis naturel est donc de s'assurer que ces probabilités sont bien définies, autrement dit que ces événements sont bien dans la tribu $\mathcal{F}$.

Remarque. La notion de variable aléatoire discrète est stable par toutes les opérations classiques sur les fonctions : la combinaison linéaire, le produit, le minimum, le maximum de deux variables aléatoires discrètes X et Y sont des variables aléatoires discrètes. Ces propriétés élémentaires étant plus fastidieuses à démontrer que difficiles à concevoir, on n'insistera pas plus.

Exemples :

1. On lance un dé équilibré : on gagne 1€ si le numéro est pair, on perd 1€ sinon. Ceci peut être modélisé par l'application $X : (\Omega, \mathcal{F}, \mathbb{P}) \to \mathcal{X} = \{-1, +1\}$ avec $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $\mathbb{P}$ l'équiprobabilité sur Ω. Il est bien clair que les deux événements $\{X = +1\} = \{2, 4, 6\}$ et $\{X = -1\} = \{1, 3, 5\}$ appartiennent à $\mathcal{F}$, c'est-à-dire que X est bien une variable aléatoire discrète.
2. On lance deux dés équilibrés et on s'intéresse à leur somme. Celle-ci est une variable aléatoire discrète X à valeurs dans $\mathcal{X} = \{2, \dots, 12\}$, $\Omega = \{(i, j), 1 \leq i, j \leq 6\}$ étant muni de la tribu $\mathcal{F} = \mathcal{P}(\Omega)$ et de l'équiprobabilité $\mathbb{P}$.

On voit que la probabilité qui nous intéresse n'est pas tant $\mathbb{P}$ sur l'espace $(\Omega, \mathcal{F})$ que la probabilité de prendre chacune des valeurs x_i. C'est ce qui définit la loi de la variable aléatoire X.

Définition 2.2 (Loi d'une variable aléatoire discrète)
Soit X une variable aléatoire discrète à valeurs dans $\mathcal{X} = (x_i)_{i \in I}$. La loi, ou distribution, de X est la famille des $(p_i)_{i \in I}$, avec :

$$\forall i \in I \qquad p_i = \mathbb{P}(X = x_i).$$

Exemples :

1. Lancer d'un dé : la loi de X est donnée par $p_{-1} = p_{+1} = 1/2$. On dit que X suit une loi uniforme sur l'ensemble $\{-1, +1\}$.
2. Lancer de deux dés : la loi de X est donnée par le vecteur ligne $p = [p_2, \dots, p_{12}]$ suivant :
$$p = \left[\frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}, \frac{5}{36}, \frac{4}{36}, \frac{3}{36}, \frac{2}{36}, \frac{1}{36}\right].$$
Cette loi est illustrée figure 2.1.

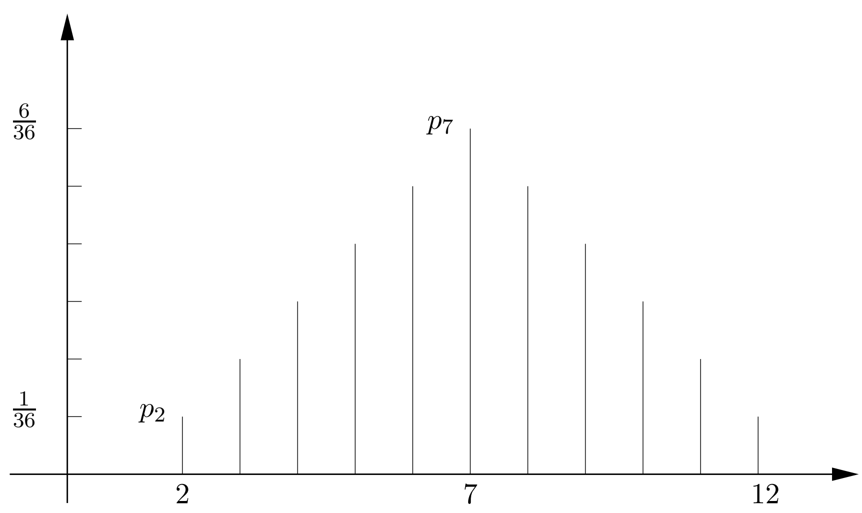


FIGURE 2.1 – Loi de la somme de deux dés.

Propriétés 2.1 (Propriétés de la loi d'une variable discrète)
Soit X une variable aléatoire discrète à valeurs dans $\mathcal{X} = (x_i)_{i \in I}$ et de loi $(p_i)_{i \in I}$, alors on a :

(i) $\forall i \in I$, $0 \leq p_i \leq 1$;

(ii) $\sum_{i \in I} p_i = 1$.

Nous allons maintenant nous intéresser à une fonction de $\mathbb{R}$ dans $\mathbb{R}$ qui permet de caractériser aussi bien les valeurs prises par une variable aléatoire discrète que sa loi : la fonction de répartition.

2.2 Fonction de répartition

Définition 2.3 (Fonction de répartition)
Soit X une variable aléatoire discrète. La fonction de répartition de X est la fonction F définie par :

$$F : \left\{ \begin{array}{ll} \mathbb{R} & \to \mathbb{R} \\ x & \mapsto F(x) = \mathbb{P}(X \leq x) = \sum_{i : x_i \leq x} p_i \end{array} \right. \tag{2.1}$$

Remarque. On trouve encore dans certains ouvrages la définition : $F(x) = \mathbb{P}(X < x)$. Cela ne change pas grand-chose (cf. infra), mais l'usage tend plutôt à imposer la définition que nous venons de donner.

Exemples :

1. Lancer d'un dé : F ne prend que 3 valeurs, à savoir 0 sur $]-\infty, -1[$, $1/2$ sur $[-1, +1[$ et 1 sur $[1, +\infty[$.
2. Lancer de deux dés : F est à nouveau une fonction en escalier (voir figure 2.2). Elle vaut 0 sur $]-\infty, 2[$, $1/36$ sur $[2, 3[$, $3/36$ sur $[3, 4[$, ..., $35/36$ sur $[11, 12[$ et 1 sur $[12, +\infty[$.

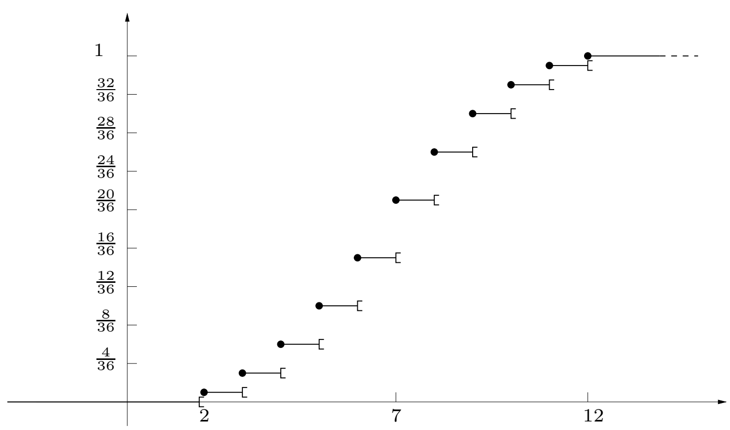


FIGURE 2.2 – Fonction de répartition pour la somme de deux dés.

Sur ces exemples, on peut déjà constater certaines propriétés communes aux deux fonctions de répartition : monotonie, limites en $\pm\infty$, continuité à droite, présence de sauts. Le résultat suivant assure leur généralité.

Propriétés 2.2 (Propriétés d'une fonction de répartition)
Soit X une variable aléatoire discrète à valeurs dans $\mathcal{X} = (x_i)_{i \in I}$. Sa fonction de répartition F a les propriétés suivantes :

1. *F est croissante ;*
2. $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to +\infty} F(x) = 1$ *;*
3. *F est continue à droite ;*
4. $\forall x \in \mathbb{R}$, $\mathbb{P}(X = x) = F(x) - F(x^-)$, *où* $F(x^-) = \lim_{\delta \to 0^+} F(x - \delta)$.

Remarque. Si nous avions pris pour définition $F(x) = \mathbb{P}(X < x)$, les deux dernières propriétés auraient été : F est continue à gauche et $\forall x \in \mathbb{R}$, $\mathbb{P}(X = x) = F(x^+) - F(x)$, où $F(x^+) = \lim_{\delta \to 0^+} F(x + \delta)$.

Preuve.

1. Cette première propriété découle de la monotonie de $\mathbb{P}$. Si $x \leq x'$, alors on a l'inclusion d'événements :
$$\{X \leq x\} = \{\omega \in \Omega \ : \ X(\omega) \leq x\} \subseteq \{\omega \in \Omega \ : \ X(\omega) \leq x'\} = \{X \leq x'\},$$
d'où :
$$F(x) = \mathbb{P}(X \leq x) \leq \mathbb{P}(X \leq x') = F(x').$$
2. Commençons par noter que, F étant croissante sur $\mathbb{R}$, elle admet des limites en $-\infty$ et $+\infty$. En particulier on a $\lim_{x \to -\infty} F(x) = \lim_{n \to +\infty} F(-n)$. Il suffit alors d'appliquer la continuité monotone décroissante de $\mathbb{P}$ en considérant la suite décroissante d'événements $(A_n)_{n \geq 0}$ définie par $A_n = \{X \leq -n\}$:
$$\lim_{n \to +\infty} F(-n) = \lim_{n \to +\infty} \mathbb{P}(A_n) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty} A_n\right) = \mathbb{P}(\emptyset) = 0.$$
De même, puisque $\lim_{x \to +\infty} F(x) = \lim_{n \to +\infty} F(n)$, la limite en $+\infty$ s'obtient via la continuité monotone croissante de $\mathbb{P}$ appliquée à la suite d'ensembles $B_n = \{X \leq n\}$:
$$\lim_{n \to +\infty} F(n) = \lim_{n \to +\infty} \mathbb{P}(B_n) = \mathbb{P}\left(\bigcup_{n=0}^{+\infty} B_n\right) = \mathbb{P}(\Omega) = 1.$$
3. F est continue à droite signifie que F est continue à droite en tout point. Soit donc x_0 un réel quelconque. Puisque F est croissante, elle admet une limite à droite (ainsi qu'à gauche, d'ailleurs) en x_0, que l'on note $F(x_0^+) = \lim_{n \to +\infty} F(x_0 + 1/n)$. Pour montrer qu'elle est continue à droite en ce point, il suffit de montrer que $\lim_{n \to +\infty} F(x_0 + 1/n) = F(x_0)$. On utilise à nouveau la continuité monotone décroissante de $\mathbb{P}$, avec cette fois la suite décroissante d'ensembles $(C_n)_{n \geq 1}$ définie par $C_n = \{X \leq x_0 + 1/n\}$, ce qui donne :
$$\lim_{n \to +\infty} F(x_0 + 1/n) = \lim_{n \to +\infty} \mathbb{P}(C_n) = \mathbb{P}\left(\bigcap_{n=1}^{+\infty} C_n\right) = \mathbb{P}(X \leq x_0) = F(x_0).$$
4. On se sert à nouveau de la continuité monotone décroissante en remarquant que :
$$\{X = x\} = \bigcap_{n=1}^{+\infty} \{x - 1/n < X \leq x\} = \bigcap_{n=0}^{+\infty} D_n.$$
Or, par définition d'une fonction de répartition, on a $\mathbb{P}(D_n) = F(x) - F(x - 1/n)$, donc :
$$F(x) - F(x^-) = \lim_{n \to +\infty} (F(x) - F(x - 1/n)) = \lim_{n \to +\infty} \mathbb{P}(D_n) = \mathbb{P}\left(\bigcap_{n=1}^{+\infty} D_n\right) = \mathbb{P}(X = x).$$

■

La dernière propriété montre que les seuls endroits où F présente des sauts sont ceux où X a des chances de tomber, la hauteur de chaque saut étant égale à la probabilité de tomber en ce point. Ceci était clair sur nos deux exemples précédents (voir en particulier la figure 2.2).

2.3 Moments d'une variable discrète

Dans toute cette section, comme précédemment, X est une variable aléatoire discrète à valeurs dans $\mathcal{X} = (x_i)_{i \in I}$ et de loi $(p_i)_{i \in I}$.

2.3.1 Espérance

Commençons par un rappel sur les séries : on dit que la série numérique $\sum u_n$ est absolument convergente si la série $\sum |u_n|$ est convergente. On sait que l'absolue convergence d'une série implique sa convergence. Si la série $\sum u_n$ est convergente, mais pas absolument convergente, on dit qu'elle est semi-convergente. C'est le cas de la série harmonique alternée $\sum \frac{(-1)^n}{n}$ (dont la somme vaut $-\ln 2$). Etant donné leurs problèmes de commutativité, on fuira comme la peste les séries semi-convergentes.

Définition 2.4 (Espérance)
On dit que la variable X admet une espérance si la série $\sum_{i \in I} x_i p_i$ est absolument convergente, c'est-à-dire si :

$$\sum_{i \in I} |x_i| p_i < +\infty.$$

Si tel est le cas, on appelle espérance de X et on note $\mathbb{E}[X]$ la quantité :

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

Cette définition semble un peu tordue : on commence par vérifier une certaine condition et, si elle est vérifiée, on définit l'objet qui nous intéresse d'une autre façon. Nous reviendrons plus loin sur la raison de tout ceci. Remarquons néanmoins dès à présent que si l'ensemble $\mathcal{X}$ des valeurs possibles de X est fini, il n'y a rien à vérifier et l'espérance de X est toujours définie. Si $\mathcal{X}$ est infini mais contenu dans $\mathbb{R}^+$ ou dans $\mathbb{R}^-$, l'absolue convergence équivaut à la convergence, donc la vérification se fait en même temps que le calcul de l'espérance. Le cas critique est celui de $\mathcal{X}$ infini avec une infinité de valeurs positives et une infinité de valeurs négatives (cf. le second contre-exemple ci-après).

Terminologie. Le terme "espérance" est historique et dû au fait que les probabilités sont nées des jeux d'argent. On peut penser en particulier aux paris du Chevalier de Méré (cf. exercice 1.36) : celui-ci considérait par exemple, à raison, que miser sur l'apparition d'au moins un 6 sur 4 lancers successifs d'un dé était avantageux. Il avait donc **l'espoir** d'un gain positif en moyenne.

Interprétation. L'espérance de X peut être vue comme la moyenne des valeurs x_i pondérées par les probabilités p_i, c'est pourquoi on dit aussi **moyenne** de X pour parler de son espérance (cette moyenne pondérée correspondant bien sûr au barycentre vu dans les petites classes). En particulier, si X prend ses valeurs entre a et b (i.e. $\mathcal{X} \subset [a, b]$), on aura nécessairement $a \leq \mathbb{E}[X] \leq b$.

Exemples :

1. Lancer d'un dé : votre gain moyen est

$$\mathbb{E}[X] = (-1) \times \mathbb{P}(X = -1) + 1 \times \mathbb{P}(X = 1) = 0.$$

   En moyenne, vous ne perdrez ni ne gagnerez donc rien à ce jeu (rien d'étonnant).

2. Lancer de deux dés : l'espérance de la somme X des deux dés est

$$\mathbb{E}[X] = 2\mathbb{P}(X = 2) + \cdots + 12\mathbb{P}(X = 12) = \ldots (\text{petit calcul}) \cdots = 7.$$

Ce résultat est bien naturel, la loi de X étant symétrique par rapport à 7 (cf. figure 2.1). L'espérance apparaît donc comme une mesure de tendance centrale.

3. Loi de Poisson $\mathcal{P}(1)$: on considère une variable aléatoire X à valeurs dans $\mathbb{N}$ et dont la loi[1] est donnée par $p_n = \mathbb{P}(X = n) = e^{-1}/n!$ pour tout $n \geq 0$. Son espérance est bien définie et vaut :
$$\mathbb{E}[X] = \sum_{n=0}^{+\infty} x_n p_n = \sum_{n=0}^{+\infty} n \frac{e^{-1}}{n!} = e^{-1} \sum_{n=0}^{+\infty} \frac{n}{n!} = e^{-1} \sum_{n=1}^{+\infty} \frac{n}{n!},$$
ce qui s'écrit encore :
$$\mathbb{E}[X] = e^{-1} \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} = e^{-1} \sum_{n=0}^{+\infty} \frac{1}{n!} = e^{-1} e = 1.$$
Ainsi une loi de Poisson de paramètre 1 a pour espérance 1. Nous verrons plus loin que ce résultat se généralise : si X suit une loi de Poisson de paramètre λ, alors $\mathbb{E}[X] = \lambda$.

Voyons maintenant deux situations où les choses ne se passent pas bien.

Contre-exemples :

1. Considérons la variable aléatoire X pouvant prendre les valeurs $x_n = n$ pour tout $n \geq 1$, c'est-à-dire que $\mathcal{X} = \{1, 2, \dots\}$ et dont la loi est donnée par $p_n = \mathbb{P}(X = n) = 1/(n(n+1))$. Commençons par remarquer que $(p_n)_{n\geq 1}$ est bien une loi de probabilité puisque :
$$\sum_{n=1}^{+\infty} p_n = \sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = \sum_{n=1}^{+\infty} \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1,$$
puisqu'on reconnaît dans la dernière expression une somme télescopique. L'espérance de X n'est pas définie puisque la série harmonique est divergente :
$$\sum_{n=1}^{+\infty} |x_n| p_n = \sum_{n=1}^{+\infty} x_n p_n = \sum_{n=1}^{+\infty} \frac{1}{n+1} = +\infty.$$

2. Plus retors : considérons la variable aléatoire X pouvant prendre les valeurs $x_n = (-1)^{n+1}(n+1)$ pour tout $n \geq 1$, c'est-à-dire que $\mathcal{X} = \{2, -3, 4, -5, 6, -7, \dots\}$ et dont la loi est donnée par $p_n = \mathbb{P}(X = (-1)^{n+1}(n+1)) = 1/(n(n+1))$. La série $\sum_{n\geq 1} x_n p_n$, bien que convergente :
$$\sum_{n=1}^{+\infty} x_n p_n = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} = \ln 2,$$
n'est pas absolument convergente puisque $\sum_{n\geq 1} |x_n| p_n = \sum_{n\geq 1} \frac{1}{n}$ est la série harmonique. La variable X n'admet donc pas d'espérance.

Problème de commutativité. Supposons une seconde que dans l'exemple précédent on admette que X a une espérance et que celle-ci vaut $\ln 2$, puisqu'après tout la série est bien convergente (cf. figure 2.3 à gauche) :
$$\sum_{n=1}^{+\infty} x_n p_n = 2p_2 - 3p_3 + 4p_4 - 5p_5 + 6p_6 + \cdots = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots = \ln 2.$$

1. On rappelle que pour tout $x \in \mathbb{R}$, on a $e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}$.

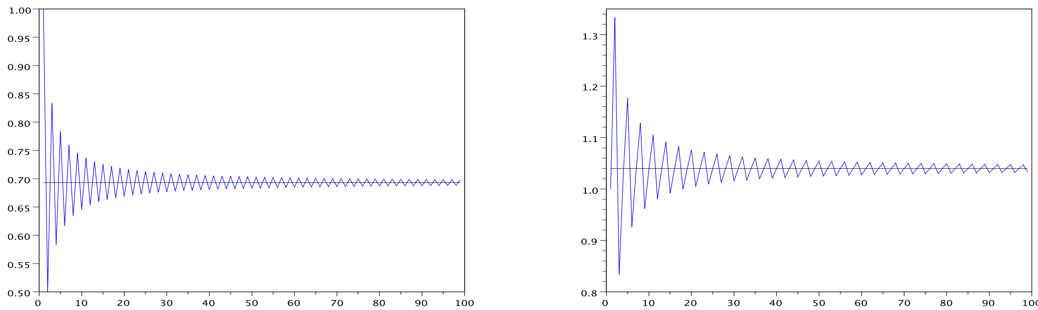


FIGURE 2.3 – Illustration de la non-commutativité de la série harmonique alternée. A gauche : $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots = \ln 2 \approx 0.69$. A droite : $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \cdots = \frac{3}{2} \ln 2 \approx 1.04$.

Considérons alors une variable Y à valeurs dans $\mathcal{Y} = \{2, 4, -3, 6, 8, -5, \dots\}$ et dont la loi est donnée par $\mathbb{P}(Y = (-1)^{n+1}(n+1)) = 1/(n(n+1))$. Ainsi Y prend les mêmes valeurs que X et avec les mêmes probabilités : si l'espérance de X était définie, on s'attendrait donc logiquement à ce qu'il en aille de même pour Y, avec la même valeur moyenne, c'est-à-dire $\ln 2$. Or on peut montrer que le fait de modifier l'ordre de sommation change tout (cf. figure 2.3 à droite) :

$$2p_2 + 4p_4 - 3p_3 + 6p_6 + \cdots = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \cdots = \frac{3}{2} \ln 2.$$

Ce coup de théâtre est dû au fait qu'une série semi-convergente n'est pas commutative. Or, comme on vient de le voir, ce phénomène n'est pas du tout souhaitable si on veut définir proprement la valeur moyenne d'une variable aléatoire, c'est pourquoi on l'avait exclu d'emblée dans la définition de l'espérance.

Revenons à des choses moins pathologiques. Etant donné une variable X dont on connaît la loi, il arrive souvent qu'on veuille calculer non pas l'espérance de X, mais l'espérance d'une fonction de X. Le résultat suivant donne une façon très simple de le faire.

Théorème 2.1 (Théorème de transfert)
Soit X une variable aléatoire discrète et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction, alors $Y = \varphi(X)$ est encore une variable aléatoire discrète et son espérance vaut :

$$\mathbb{E}[Y] = \mathbb{E}[\varphi(X)] = \sum_{i \in I} \varphi(x_i) p_i,$$

sous réserve d'absolue convergence de cette série.

Preuve. Puisque X prend ses valeurs dans un ensemble au plus dénombrable $\mathcal{X} = (x_i)_{i \in I}$, Y prend elle aussi ses valeurs dans un ensemble au plus dénombrable $\mathcal{Y} = (y_j)_{j \in J} = \varphi(\mathcal{X})$. Pour tout indice j de J, notons :

$$E_j = \{i \in I \ : \ \varphi(x_i) = y_j\}$$

et $q_j = \mathbb{P}(Y = y_j) = \sum_{i \in E_j} p_i$. Puisque les $(y_j)_{j \in J}$ et les $(q_j)_{j \in J}$ définissent la loi de Y, son espérance vaut tout bonnement :

$$\mathbb{E}[Y] = \sum_{j \in J} y_j q_j = \sum_{j \in J} y_j \left(\sum_{i \in E_j} p_i \right) = \sum_{j \in J} \left(\sum_{i \in E_j} \varphi(x_i) p_i \right),$$

et puisque les $(E_j)_{j\in J}$ forment une partition de I, on obtient en regroupant les deux symboles de sommation :

$$\mathbb{E}[Y] = \sum_{i\in I} \varphi(x_i)p_i,$$

Dans ce qui précède, toutes les manipulations de sommes sont justifiées par l'absolue convergence de la série $\sum_{i\in I} \varphi(x_i)p_i$.

■

Moyen mnémotechnique. Pour calculer $\mathbb{E}[\varphi(X)]$ et non $\mathbb{E}[X]$, on a juste à remplacer x_i par $\varphi(x_i)$ dans la formule de $\mathbb{E}[X]$.

L'intérêt pratique de ce résultat est le suivant : on n'a pas besoin de commencer par déterminer la loi de Y pour calculer son espérance, il suffit tout simplement de **transférer** la loi de X.

Exemple. Soit X une variable aléatoire discrète de loi uniforme (i.e. équiprobable) sur l'ensemble $\mathcal{X} = \{-2,-1,0,1,2\}$. Considérons maintenant la variable $Y = X^2$, dont on veut calculer l'espérance. Si on ignore le résultat précédent, on doit commencer par trouver la loi de Y, à savoir : Y prend ses valeurs dans l'ensemble $\mathcal{Y} = \{0,1,4\}$ avec les probabilités respectives $1/5$, $2/5$ et $2/5$. Ainsi son espérance vaut :

$$\mathbb{E}[Y] = 0\times\frac{1}{5} + 1\times\frac{2}{5} + 4\times\frac{2}{5} = 2.$$

Si on applique le résultat précédent, on calcule directement :

$$\mathbb{E}[Y] = \mathbb{E}[X^2] = (-2)^2\times\frac{1}{5} + (-1)^2\times\frac{1}{5} + 0^2\times\frac{1}{5} + 1^2\times\frac{1}{5} + 2^2\times\frac{1}{5} = 2,$$

et les Athéniens s'atteignirent.

Proposition 2.1 (Linéarité de l'espérance)
Soit X et Y deux variables aléatoires admettant une espérance, a et b deux réels, alors la variable aléatoire $aX + bY$ admet aussi une espérance et celle-ci vaut :

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y].$$

En particulier, on a :

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b.$$

Preuve. Pour montrer que $\mathbb{E}[aX] = a\mathbb{E}[X]$, il suffit d'appliquer le théorème de transfert :

$$E[aX] = \sum_{i\in I} ax_ip_i = a\sum_{i\in I} x_ip_i = a\mathbb{E}[X].$$

Il reste donc uniquement à prouver que $\mathbb{E}[X+Y] = \mathbb{E}[X]+\mathbb{E}[Y]$. On adopte les notations suivantes : la variable X prend les valeurs $\mathcal{X} = (x_i)_{i\in I}$ avec les probabilités $(p_i)_{i\in I}$, la variable Y les valeurs $\mathcal{Y} = (y_j)_{j\in J}$ avec les probabilités $(q_j)_{j\in J}$, la variable $Z = (X + Y)$ les valeurs $\mathcal{Z} = (z_k)_{k\in K}$ avec les probabilités $(r_k)_{k\in K}$. Enfin, pour tout couple d'indices $(i, j) \in I \times J$, on note $p_{ij} = \mathbb{P}(X = x_i, Y = y_j)$ la probabilité jointe. On a alors :

$$\mathbb{E}[Z] = \sum_{k\in K} z_kr_k = \sum_{k\in K} z_k\left(\sum_{(i,j):x_i+y_j=z_k} p_{ij}\right) = \sum_{k\in K}\left(\sum_{(i,j):x_i+y_j=z_k} z_kp_{ij}\right),$$

que l'on peut casser en deux morceaux puisque $z_k = x_i + y_j$:

$$\mathbb{E}[Z] = \sum_{(i,j)\in I\times J} (x_i + y_j)p_{ij} = \sum_{(i,j)\in I\times J} x_i p_{ij} + \sum_{(i,j)\in I\times J} y_j p_{ij},$$

et l'idée est de regrouper différemment les indices de sommation dans chacune des deux sommes :

$$\mathbb{E}[Z] = \sum_{i\in I} x_i \left(\sum_{j\in J} p_{ij}\right) + \sum_{j\in J} y_j \left(\sum_{i\in I} p_{ij}\right),$$

or on reconnaît les p_i dans la première somme et les q_j dans la seconde :

$$\mathbb{E}[Z] = \sum_{i\in I} x_i p_i + \sum_{j\in J} y_j q_j = \mathbb{E}[X] + \mathbb{E}[Y].$$

Précisons que toutes les opérations effectuées sur les sommes sont légitimes en raison de l'absolue convergence des deux séries $\sum x_i p_i$ et $\sum y_j q_j$. Pour achever la preuve de la proposition, il reste à vérifier que, pour tout réel b, on a $\mathbb{E}[b] = b$. Or b est la variable aléatoire qui prend la seule valeur b avec probabilité 1, donc $\mathbb{E}[b] = b \times 1 = b$.

■

En termes d'algèbre linéaire, ce résultat dit la chose suivante : l'ensemble des variables aléatoires discrètes admettant une espérance est un sous-espace de l'espace des variables aléatoires discrètes et l'espérance est une forme linéaire sur ce sous-espace. Le résultat suivant montre que c'est même une forme linéaire positive.

Proposition 2.2 (Positivité de l'espérance)
Soit X et Y deux variables aléatoires admettant une espérance et telles que $X \leq Y$, alors :

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

En particulier, si $X \geq 0$, i.e. si X ne prend que des valeurs positives, on a $\mathbb{E}[X] \geq 0$.

Dire que $X \leq Y$ signifie que pour tout $\omega \in \Omega$, on a $X(\omega) \leq Y(\omega)$, autrement dit : quel que soit ce qui se passe, on aura toujours X plus petit que Y.

Preuve. Dire que $X \geq 0$ est équivalent à dire que $x_i \geq 0$ pour tout $i \in I$. Ainsi l'espérance $\mathbb{E}[X] = \sum_{i\in I} x_i p_i$ est la somme de termes positifs, elle est donc positive. Par ailleurs, dire que $X \leq Y$ est équivalent à dire que la variable $Z = (Y - X)$ est positive. Le point précédent et la linéarité de l'espérance vue ci-dessus permettent d'en déduire que $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

■

Exemple : min & max. On jette simultanément deux dés et on appelle X (resp. Y) le minimum (resp. le maximum) des deux numéros obtenus. Il est bien clair que $X \leq Y$ et le résultat ci-dessus nous assure, sans aucun calcul, que $\mathbb{E}[X] \leq \mathbb{E}[Y]$.

Corollaire 2.1 (Espérance de la valeur absolue de X)
La variable aléatoire X admet une espérance si et seulement si la variable aléatoire $|X|$ en admet une, auquel cas :

$$\mathbb{E}[X] \leq \mathbb{E}[|X|].$$

Preuve. Par définition, la variable aléatoire X admet une espérance si et seulement si la série $\sum_{i\in I} x_i p_i$ est absolument convergente, c'est-à-dire si et seulement si la série $\sum_{i\in I} |x_i|p_i$ est convergente. Or $|X|$ est la variable aléatoire prenant les valeurs $(|x_i|_{i\in I})$ avec les probabilités $(p_i)_{i\in I}$. Donc elle admet une espérance si et seulement si la série $\sum_{i\in I} |x_i|p_i$ est (absolument) convergente. L'équivalence entre l'existence de $\mathbb{E}[X]$ et celle de $\mathbb{E}[|X|]$ est donc claire. Si cette existence est assurée, il suffit alors de remarquer que :

$$\forall \omega \in \Omega \qquad X(\omega) \leq |X(\omega)|$$

et d'appliquer la propriété de positivité de l'espérance pour en déduire que $\mathbb{E}[X] \leq \mathbb{E}[|X|]$.

■

Remarque. Le raisonnement ci-dessus montre de façon plus générale que si $|X| \leq Y$ et si Y admet une espérance, alors X aussi et $\mathbb{E}[X] \leq \mathbb{E}[Y]$. Ce passage par un majorant est aussi d'usage constant en analyse, pour justifier la convergence de séries numériques et celle d'intégrales généralisées.

2.3.2 Variance

Nous avons dit que l'espérance est une mesure de tendance centrale. Nous allons définir maintenant une mesure de dispersion autour de cette valeur centrale : la variance.

Définition 2.5 (Variance & Ecart-type)
Soit X une variable aléatoire discrète admettant une espérance $\mathbb{E}[X]$. La variance de X est définie par :

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \sum_{i\in I}(x_i - \mathbb{E}[X])^2 p_i,$$

sous réserve de convergence de cette série. On appelle alors écart-type, noté $\sigma(X)$, la racine de la variance : $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Puisque la série $\sum_{i\in I}(x_i - \mathbb{E}[X])^2 p_i$ est à termes positifs, absolue convergence équivaut à convergence. A nouveau, lorsque X ne prend qu'un nombre fini de valeurs, la variance est toujours définie.

Interprétation. De façon générale, la variance d'une variable mesure **la moyenne des carrés des écarts à sa moyenne**. Ainsi, plus la loi d'une variable est étalée autour de sa moyenne, plus sa variance est grande. D'autre part, si X représente une grandeur physique (donc ayant une dimension, par exemple une durée), alors l'écart-type a la même dimension que X, tandis que la variance a cette dimension au carré, ce qui la rend moins parlante en pratique. Le terme écart-type est d'ailleurs à comprendre au sens "écart typique" d'une variable à sa moyenne. Nous y reviendrons plus loin.

Exemples :

1. Lancer d'un dé : nous avons vu que $\mathbb{E}[X] = 0$, sa variance vaut :

$$\mathrm{Var}(X) = (-1-0)^2 \times \frac{1}{2} + (1-0)^2 \times \frac{1}{2} = 1.$$

2. Lancer de deux dés : l'espérance de la somme X des deux dés est $\mathbb{E}[X] = 7$ et sa variance est :

$$\mathrm{Var}[X] = (2-7)^2\mathbb{P}(X=2) + \cdots + (12-7)^2\mathbb{P}(X=12) = \ldots \text{(petit calcul)} \cdots = \frac{35}{6},$$

et son écart-type vaut donc $\sigma(X) = \sqrt{\frac{35}{6}} \approx 2,4$.

Il existe une autre formulation de la variance, qui permet éventuellement d'alléger les calculs : elle est connue sous le nom de formule de König, ou de Huygens-König, par analogie avec la formule de l'énergie cinétique d'un système en mécanique. C'est le premier point des propriétés suivantes.

Propriétés 2.3 (Propriétés de la variance)
Soit X une variable aléatoire discrète, alors sous réserve d'existence de sa variance on a :
(i) $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
(ii) Si a et b sont deux réels, $\text{Var}(aX+b) = a^2\text{Var}(X)$.
(iii) $\text{Var}(X) = 0$ *si et seulement si X est constante.*

Preuve.
(i) Il suffit d'utiliser la linéarité de l'espérance :

$$\text{Var}(X) = \mathbb{E}[(X-\mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2\mathbb{E}[X]X + \mathbb{E}[X]^2] = \mathbb{E}[X^2] - 2\mathbb{E}[\mathbb{E}[X]X] + \mathbb{E}[\mathbb{E}[X]^2],$$

or $\mathbb{E}[X]$ est une quantité déterministe (i.e. non aléatoire) donc il suffit d'appliquer les propriétés vues en Proposition 2.1 :

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

(ii) On applique à nouveau la linéarité de l'espérance :

$$\text{Var}(aX+b) = \mathbb{E}[(aX+b-\mathbb{E}[aX+b])^2] = \mathbb{E}[(aX+b-(a\mathbb{E}[X]+b))^2] = \mathbb{E}[(aX-a\mathbb{E}[X])^2],$$

d'où :

$$\text{Var}(aX+b) = \mathbb{E}[(a(X-\mathbb{E}[X]))^2] = \mathbb{E}[a^2(X-\mathbb{E}[X])^2] = a^2\mathbb{E}[(X-\mathbb{E}[X])^2] = a^2\text{Var}(X).$$

(iii) On rappelle que la variable discrète X prend les valeurs x_i avec les probabilités strictement positives p_i. Supposons X de variance nulle :

$$\sum_{i\in I}(x_i - \mathbb{E}[X])^2 p_i = 0,$$

ainsi on a une série de termes positifs dont la somme est nulle, ce qui n'est possible que si tous les termes sont nuls, c'est-à-dire si :

$$\forall i \in I \qquad (x_i - \mathbb{E}[X])^2 p_i = 0 \;\Leftrightarrow\; x_i = \mathbb{E}[X],$$

la dernière équivalence venant de ce que les p_i sont tous supposés strictement positifs. Ainsi la seule valeur que prend X est sa moyenne, autrement dit cette variable aléatoire est constante (autant dire qu'elle n'a pas grand-chose d'aléatoire...). La réciproque est clairement vraie : si $X(\omega) = x_0$ pour tout $\omega \in \Omega$, alors on vérifie aisément que X admet une espérance et que celle-ci vaut bien sûr $\mathbb{E}[X] = x_0$, et par suite que la variance est nulle.

■

Si on connaît $\mathbb{E}[X]$, il suffit d'après le point (i) de calculer $\mathbb{E}[X^2]$ pour obtenir la variance de X. Ceci peut parfois se faire simplement en remarquant que $X^2 = X(X-1) + X$, d'où il découle que $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X]$. Illustrons-le sur un exemple.

Exemple : Loi de Poisson. Revenons sur l'exemple où $X \sim \mathcal{P}(1)$, loi de Poisson de paramètre 1. Nous avons montré que $\mathbb{E}[X] = 1$, que vaut sa variance ? Puisqu'on connaît $\mathbb{E}[X]$, on se contente de calculer $\mathbb{E}[X(X-1)]$, ce qui donne grâce au théorème de transfert :

$$\mathbb{E}[X(X-1)] = \sum_{n=0}^{+\infty} n(n-1)\frac{e^{-1}}{n!} = e^{-1}\sum_{n=0}^{+\infty}\frac{n(n-1)}{n!} = e^{-1}\sum_{n=2}^{+\infty}\frac{n(n-1)}{n!},$$

qui s'écrit encore :

$$\mathbb{E}[X(X-1)] = e^{-1}\sum_{n=2}^{+\infty}\frac{1}{(n-2)!} = e^{-1}\sum_{n=0}^{+\infty}\frac{1}{n!} = e^{-1}e = 1,$$

d'où l'on déduit que $\mathbb{E}[X^2] = \mathbb{E}[X(X-1)] + \mathbb{E}[X] = 1 + 1 = 2$ et $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1$. Ainsi une variable qui suit une loi de Poisson de paramètre 1 a pour variance 1. On montrera plus généralement que si $X \sim \mathcal{P}(\lambda)$, alors $\mathbb{E}[X] = \mathrm{Var}(X) = \lambda$.

Nous avons dit plus haut que l'écart-type permet d'avoir une idée de l'écart typique entre une variable aléatoire et sa moyenne. Cette idée est précisée par la célèbre inégalité de Tchebychev, également appelée inégalité de Bienaymé-Tchebychev.

Théorème 2.2 (Inégalité de Tchebychev)
Soit X une variable aléatoire discrète admettant une variance, alors :

$$\forall t > 0 \qquad \mathbb{P}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Preuve. Puisqu'on peut voir cette inégalité comme un cas particulier de l'inégalité de Markov, nous donnerons sa preuve en section suivante.

■

Interprétation. Si on pose $t = s\sigma(X)$, l'inégalité de Tchebychev se réécrit pour tout $s > 0$:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq s\sigma(X)) \leq \frac{1}{s^2}.$$

Si on voit l'écart-type $\sigma(X)$ comme une unité d'écart, ceci dit que la probabilité qu'une variable s'éloigne de plus de s unités d'écart de sa moyenne est inférieure à $\frac{1}{s^2}$.

L'aspect remarquable de l'inégalité de Tchebychev est son universalité, puisqu'elle est vérifiée quelle que soit la loi de la variable X (si tant est bien sûr que celle-ci admette une variance). Le prix à payer est qu'elle ne donne souvent en pratique que de médiocres majorations de la queue de distribution. Elle prend néanmoins toute sa puissance pour prouver des résultats généraux, c'est par exemple la méthode typique de démonstration de la loi faible des grands nombres.

Exemple. Une application de l'inégalité de Tchebychev est donnée en exercice 2.21.

2.3.3 Autres moments

On va maintenant généraliser les notions d'espérance et de variance.

Définition 2.6
Soit X une variable aléatoire discrète et $m \in \mathbb{N}^*$. Sous réserve d'existence, on appelle :
(i) moment d'ordre m de X la quantité

$$\mathbb{E}[X^m] = \sum_{i \in I} x_i^m p_i \ ;$$

(ii) moment centré d'ordre m de X la quantité

$$\mathbb{E}[(X - \mathbb{E}[X])^m] = \sum_{i \in I} (x_i - \mathbb{E}[X])^m p_i.$$

Ainsi l'espérance de X est le moment d'ordre 1 et sa variance le moment centré d'ordre 2. Précisons au passage un point de vocabulaire : on dit que X est une variable centrée si $\mathbb{E}[X] = 0$ et qu'elle est réduite si $\text{Var}[X] = 1$. Si X admet une variance non nulle, on dit qu'on centre et réduit X en considérant la variable $Y = (X - \mathbb{E}[X])/\sigma(X)$.

Proposition 2.3
Soit X une variable aléatoire discrète, alors si X admet un moment d'ordre $m \in \mathbb{N}^*$, X admet des moments de tout ordre $n \in \{1, \ldots, m\}$.

Preuve. Effectuons une partition de l'ensemble I d'indices en deux sous-ensembles E_0 et E_1 :

$$\left\{ \begin{array}{l} E_0 = \{i \in I \ : \ |x_i| \leq 1\} \\ E_1 = \{i \in I \ : \ |x_i| > 1\} \end{array} \right.$$

Soit maintenant $n \in \{1, \ldots, m\}$, il nous faut montrer la convergence de la série $\sum_{i \in I} |x_i|^n p_i$, or :

$$\forall i \in I \qquad |x_i|^n \leq \left\{ \begin{array}{l} 1 \text{ si } i \in E_0 \\ |x_i|^m \text{ si } i \in E_1 \end{array} \right.$$

D'où il sort que :

$$\sum_{i \in I} |x_i|^n p_i = \sum_{i \in E_0} |x_i|^n p_i + \sum_{i \in E_1} |x_i|^n p_i \leq \sum_{i \in E_0} p_i + \sum_{i \in E_1} |x_i|^m p_i,$$

et il suit :

$$\sum_{i \in I} |x_i|^n p_i \leq \sum_{i \in I} p_i + \sum_{i \in I} |x_i|^m p_i = 1 + \mathbb{E}[|X|^m] < +\infty,$$

donc l'affaire est entendue.

■

L'existence d'un moment d'ordre élevé assure une décroissance d'autant plus rapide de la queue de la distribution de X à l'infini, comme le montre l'inégalité de Markov. On peut la voir comme une généralisation de l'inégalité de Tchebychev.

Théorème 2.3 (Inégalité de Markov)
Soit X une variable aléatoire discrète, alors si X admet un moment d'ordre $m \in \mathbb{N}^*$, on a :

$$\forall t > 0 \qquad \mathbb{P}(|X| \geq t) \leq \frac{\mathbb{E}[|X|^m]}{t^m}.$$

Preuve. Reprenons l'idée de la preuve ci-dessus, avec cette fois :

$$\left\{ \begin{array}{l} E_0 = \{i \in I \ : \ |x_i| < t\} \\ E_1 = \{i \in I \ : \ |x_i| \geq t\} \end{array} \right.$$

Remarquons d'emblée que $\mathbb{P}(|X| \geq t) = \sum_{i \in E_1} p_i$, d'où l'idée de la décomposition :

$$\mathbb{E}[|X|^m] = \sum_{i \in E_0} |x_i|^m p_i + \sum_{i \in E_1} |x_i|^m p_i \geq \sum_{i \in E_1} |x_i|^m p_i,$$

or pour tout $i \in E_1$, $|x_i|^m \geq t^m$, donc :

$$\mathbb{E}[|X|^m] \geq t^m \sum_{i \in E_1} p_i = t^m \mathbb{P}(|X| \geq t).$$

■

Remarques :

1. L'énoncé ci-dessus est en fait un peu plus général que ce qu'on appelle usuellement l'inégalité de Markov, à savoir que pour toute variable X positive admettant une espérance, on a :

$$\forall t > 0 \qquad \mathbb{P}(X \geq t) \leq \frac{\mathbb{E}[X]}{t}.$$

2. Soit Y une variable admettant une variance : l'inégalité de Tchebychev pour Y se retrouve en considérant $X = (Y - \mathbb{E}[Y])$ et $m = 2$ dans le théorème ci-dessus.

2.4 Corrélation et indépendance

Nous avons vu que l'espérance est linéaire, c'est-à-dire que $\mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y]$. On peut se demander si cette propriété est encore vraie pour la variance. La réponse est non en général et fait intervenir la notion de covariance entre variables aléatoires.

Définition 2.7 (Covariance)
Soit X et Y variables aléatoires discrètes admettant des moments d'ordre 2. La covariance entre X et Y est définie par :

$$\mathrm{Cov}(X,Y) = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y].$$

Cette définition nécessite quelques justifications :

1. Il faut commencer par vérifier que si $\mathbb{E}[X^2]$ et $\mathbb{E}[Y^2]$ existent, alors $\mathbb{E}[XY]$ est bien définie. Il suffit pour ça de partir de l'identité remarquable $(|x|-|y|)^2 = x^2 - 2|xy| + y^2 \geq 0$ pour en déduire que :

$$\forall \omega \in \Omega \qquad 2|X(\omega)Y(\omega)| \leq X^2(\omega)+Y^2(\omega) \;\Rightarrow\; |X(\omega)Y(\omega)| \leq \frac{1}{2}\left(X^2(\omega)+Y^2(\omega)\right) \leq X^2(\omega)+Y^2(\omega),$$

c'est-à-dire succinctement : $|XY| \leq X^2 + Y^2$. Ainsi la variable aléatoire positive $|XY|$ est majorée par une somme de variables admettant chacune une espérance, donc $|XY|$ admet une espérance et idem pour XY. Le premier point est plié.

2. Le second point concerne l'égalité entre les deux formulations de la covariance. Il réside tout simplement sur la linéarité de l'espérance et le fait que la moyenne d'une variable constante est égale à cette constante :

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY - \mathbb{E}[X]Y - \mathbb{E}[Y]X + \mathbb{E}[X]\mathbb{E}[Y]],$$

d'où :

$$\mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] - \mathbb{E}[Y]\mathbb{E}[X] + \mathbb{E}[X]\mathbb{E}[Y],$$

ce qui donne bien au final :

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

L'expression de droite est parfois appelée formule de König, conformément à la formule de la variance.

On peut alors donner plusieurs propriétés de la covariance.

Propriétés 2.4 (Quelques formules sur la covariance)
Soit X et Y variables aléatoires discrètes admettant des moments d'ordre 2. Alors :

1. $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$.
2. $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$.
3. *pour tous réels a, b, c, d :* $\mathrm{Cov}(aX + b, cY + d) = ac\ \mathrm{Cov}(XY)$.
4. $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + 2\mathrm{Cov}(X, Y) + \mathrm{Var}(Y)$.

Preuve. Les deux premiers points sont évidents. Le troisième s'obtient en appliquant la définition de la covariance et en utilisant la linéarité de l'espérance. Détaillons uniquement le dernier :

$$\mathrm{Var}(X+Y) = \mathbb{E}[(X+Y)^2]-(\mathbb{E}[X+Y])^2 = \mathbb{E}[X^2]+2\mathbb{E}[XY]+\mathbb{E}[Y^2]-(\mathbb{E}[X]^2+2\mathbb{E}[X]\mathbb{E}[Y]+\mathbb{E}[Y]^2),$$

et il suffit de bien regrouper les termes :

$$\mathrm{Var}(X + Y) = (\mathbb{E}[X^2] - \mathbb{E}[X]^2) + 2(\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]) + (\mathbb{E}[Y^2] - \mathbb{E}[Y]^2)$$

pour arriver à la formule voulue.

■

Cette démonstration montre que la dernière formule est bien sûr liée à l'identité remarquable vue dans les petites classes : $(x + y)^2 = x^2 + 2xy + y^2$. Elle souligne en particulier que, dans le cas général, la variance n'est pas linéaire puisqu'on n'a pas $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$. Nous allons maintenant préciser ce point.

Définition 2.8 (Coefficient de corrélation)
Soit X et Y variables aléatoires discrètes admettant des variances non nulles. Le coefficient de corrélation entre X et Y est défini par :

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

Si $\rho(X, Y) = \mathrm{Cov}(X, Y) = 0$, X et Y sont dites décorrélées, ce qui est équivalent à dire que :

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Exemple. Soit X qui suit une loi uniforme sur $\{-1, 0, +1\}$ et Y définie par $Y = X^2$. Alors par le théorème de transfert :

$$\mathbb{E}[XY] = \mathbb{E}[X^3] = (-1)^3 \times \frac{1}{3} + 0^3 \times \frac{1}{3} + 1^3 \times \frac{1}{3} = 0.$$

Un calcul similaire montre que $\mathbb{E}[X] = 0$, donc sans même calculer $\mathbb{E}[Y]$, on a aussi $\mathbb{E}[X]\mathbb{E}[Y] = 0$. Il s'ensuit que :

$$\mathrm{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0,$$

c'est-à-dire que X et Y sont décorrélées.

Le coefficient de corrélation est aussi appelé coefficient de corrélation **linéaire**, car il mesure en fait la linéarité entre les deux variables X et Y. C'est ce qu'explique le résultat suivant.

Proposition 2.4 (Inégalité de Cauchy-Schwarz)
Soit X et Y variables aléatoires discrètes admettant des moments d'ordre 2, alors :

$$-1 \leq \rho(X,Y) \leq +1,$$

avec plus précisément :

1. $\rho(X,Y) = -1$ *ssi* $\exists (a,b) \in \mathbb{R}_-^* \times \mathbb{R}$ *tels que* $Y = aX + b$ *;*
2. $\rho(X,Y) = +1$ *ssi* $\exists (a,b) \in \mathbb{R}_+^* \times \mathbb{R}$ *tels que* $Y = aX + b$*.*

Preuve. La démonstration la plus expéditive de ce résultat est basée sur une ruse de sioux. Comme toute variable aléatoire, la variable $(tX+Y)$ est de variance positive, et ce quel que soit le réel t, ce qui s'écrit encore :

$$0 \leq \mathrm{Var}(tX+Y) = \mathrm{Var}(tX) + 2\mathrm{Cov}(tX,Y) + \mathrm{Var}(Y) = t^2\mathrm{Var}(X) + 2\mathrm{Cov}(X,Y)t + \mathrm{Var}(Y),$$

que l'on peut voir comme un trinôme en t. Or un trinôme n'est de signe constant que si son discriminant est inférieur ou égal à 0, c'est-à-dire :

$$\mathrm{Cov}(X,Y)^2 - \mathrm{Var}(X)\mathrm{Var}(Y) \leq 0 \iff |\rho(X,Y)| \leq 1.$$

Supposons $\rho(X,Y) = +1$, alors en remontant les équations ceci implique qu'il existe un réel t_0 tel que $\mathrm{Var}(t_0X+Y) = 0$, donc il existe un réel b tel que $t_0X + Y = b$, c'est-à-dire $Y = -t_0X + b$. Dans ce cas

$$\rho(X,Y) = \frac{\mathrm{Cov}(X, -t_0X+b)}{\sigma(X)\sigma(-t_0X+b)} = \frac{-t_0}{|t_0|},$$

qui vaut 1 si et seulement si t_0 est négatif. Le même raisonnement permet de conclure lorsque $\rho(X,Y) = -1$.

■

Remarque. L'inégalité $|\mathrm{Cov}(X,Y)| \leq \sigma(X)\sigma(Y)$ n'est rien de plus que l'inégalité de Cauchy-Schwarz adaptée au cadre des variables aléatoires. Sa version géométrique dans $\mathbb{R}^n$ muni du produit scalaire usuel et de la norme euclidienne est : pour tous vecteurs u et v de $\mathbb{R}^n$, $\langle u,v\rangle \leq \|u\|\|v\|$. Le coefficient de corrélation de deux variables aléatoires est donc équivalent au cosinus de l'angle entre deux vecteurs.

Interprétation. De façon générale, plus le coefficient de corrélation est proche de 1 en valeur absolue, plus les variables X et Y sont linéairement liées. Un coefficient de corrélation nul signifie donc que les deux variables ne sont pas **linéairement** liées. Il n'empêche qu'elle peuvent être liées par un autre type de relation : c'est ce qui apparaît clairement sur l'exemple ci-dessus où $Y = X^2$, puisqu'une fois X connue, il n'existe plus aucune incertitude sur Y.

Face à ce constat, on aimerait définir le fait qu'il n'existe aucune sorte de relation entre X et Y. La notion pertinente est alors celle d'indépendance, déjà rencontrée dans le premier chapitre et adaptée ici au cas des variables aléatoires.

Définition 2.9 (Indépendance de deux variables)
Soit X et Y variables aléatoires discrètes à valeurs respectives dans $\mathcal{X} = (x_i)_{i\in I}$ et $\mathcal{Y} = (y_j)_{j\in J}$. On dit que X et Y sont indépendantes, noté $X \perp\!\!\!\perp Y$, si :

$$\forall (i,j) \in I \times J \qquad \mathbb{P}(X = x_i, Y = y_j) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j).$$

Concrètement, deux variables sont indépendantes si la valeur prise par l'une n'a aucune espèce d'influence sur la valeur prise par l'autre. La notion d'indépendance est omniprésente en probabilités. Un exemple parmi tant d'autres est celui du lancer simultané de deux dés : il est clair que le résultat X donné par l'un est indépendant du résultat Y donné par l'autre. Voyons maintenant en quoi la notion d'indépendance est plus forte que celle de décorrélation.

Proposition 2.5 (Indépendance ⇒ Décorrélation)
Soit X et Y variables aléatoires discrètes admettant des moments d'ordre 2. Si X et Y sont indépendantes, alors elles sont décorrélées. En particulier, on a alors :

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y).$$

Preuve. L'équivalent du théorème de transfert pour les couples de variables aléatoires permet d'écrire l'espérance de XY de la façon suivante :

$$\mathbb{E}[XY] = \sum_{(i,j)\in I\times J} x_i y_j p_{ij},$$

où $p_{ij} = \mathbb{P}(X = x_i, Y = y_j)$ pour tout couple $(i,j) \in I \times J$. En notant $p_i = \mathbb{P}(X = x_i)$ et $q_j = \mathbb{P}(Y = y_j)$, l'indépendance des deux variables donne alors $p_{ij} = p_i q_j$ et :

$$\mathbb{E}[XY] = \sum_{(i,j)\in I\times J} x_i y_j p_{ij} = \left(\sum_{i\in I} x_i p_i\right)\left(\sum_{j\in J} y_j q_j\right),$$

autrement dit $\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$, ou encore $\mathrm{Cov}(X,Y) = 0$.

■

Remarque. La réciproque est fausse en général. Pour s'en assurer il suffit de reprendre l'exemple où X est uniforme sur $\{-1, 0, +1\}$ et $Y = X^2$. On a vu que $\mathrm{Cov}(X,Y) = 0$, c'est-à-dire que X et Y sont décorrélées. Mais elles ne sont pas indépendantes, puisqu'il suffit par exemple de remarquer que :

$$\mathbb{P}(X=0, Y=0) = \mathbb{P}(X=0) = \frac{1}{3} \neq \frac{1}{9} = \mathbb{P}(X=0)\mathbb{P}(Y=0).$$

Généralisation. L'indépendance entre variables aléatoires se généralise de façon naturelle à plus de deux variables : n variables $X_1, \dots, X_n$ sont (mutuellement) indépendantes si la valeur prise par l'une n'a aucune influence sur les valeurs prises par les autres :

$$\forall (x_1, \dots, x_n) \in \mathcal{X}_1 \times \cdots \times \mathcal{X}_n \qquad \mathbb{P}(X_1 = x_1, \dots, X_n = x_n) = \mathbb{P}(X_1 = x_1) \dots \mathbb{P}(X_n = x_n).$$

La variance de la somme sera alors encore la somme des variances :

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n),$$

tandis que dans le cas général on a :

$$\mathrm{Var}(X_1 + \cdots + X_n) = \sum_{i=1}^n \mathrm{Var}(X_i) + 2 \sum_{1\leq j<k\leq n} \mathrm{Cov}(X_j, X_k).$$

2.5 Lois usuelles

Nous recensons dans cette ultime section quelques lois discrètes classiques. Pour chacune sont précisées espérance, variance ainsi que certaines propriétés particulièrement saillantes.

2.5.1 Loi uniforme

On parle de loi uniforme dès lors qu'il y a équiprobabilité pour les valeurs prises par la variable aléatoire.

Définition 2.10 (Loi uniforme)
On dit que X suit une loi uniforme sur l'ensemble $\{1, \ldots, n\}$, noté $X \sim \mathcal{U}_{\{1,\ldots,n\}}$, si :

$$\forall k \in \{1, \ldots, n\} \qquad \mathbb{P}(X = k) = \frac{1}{n}.$$

Exemple. On lance un dé équilibré et on note X le résultat obtenu. On a alors $X \sim \mathcal{U}_{\{1,\ldots,6\}}$ (voir figure 2.4 à gauche).

Remarque. On parle plus généralement de loi uniforme sur l'ensemble $\{a_1, \ldots, a_n\}$ si :

$$\forall k \in \{1, \ldots, n\} \qquad \mathbb{P}(X = a_k) = \frac{1}{n}.$$

Par exemple, la variable X valant $+1$ si le résultat du lancer d'un dé équilibré est pair, -1 sinon, suit une loi uniforme sur $\{-1, +1\}$ (voir figure 2.4 à droite).

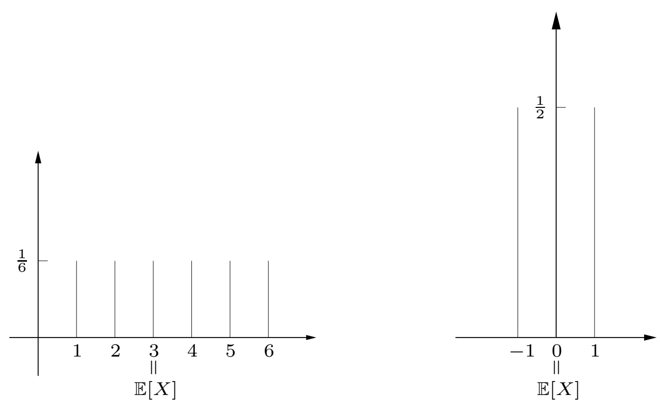


FIGURE 2.4 – A gauche : loi uniforme $\mathcal{U}_{\{1,\ldots,6\}}$. A droite : loi uniforme $\mathcal{U}_{\{-1,+1\}}$.

Proposition 2.6 (Moments d'une loi uniforme)
Si X suit une loi uniforme sur l'ensemble $\{1, \ldots, n\}$, alors :

$$\mathbb{E}[X] = \frac{n+1}{2} \qquad \& \qquad \mathrm{Var}(X) = \frac{n^2-1}{12}.$$

Preuve. Pour l'espérance, elle est basée sur la formule de la somme des termes d'une suite arithmétique :

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

Pour plus de détails, se reporter au corrigé de l'exercice 2.7. Pour la variance, on se sert de la formule de la somme des carrés :

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

laquelle se prouve en développant successivement $(1+1)^3, \dots, (n+1)^3$, en sommant le tout et en simplifiant. Pour ce qui nous concerne, elle s'applique comme suit :

$$\mathbb{E}[X^2] = \sum_{k=1}^{n} \frac{k^2}{n} = \frac{1}{n} \sum_{k=1}^{n} k^2 = \frac{(n+1)(2n+1)}{6},$$

d'où :

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(n+1)(2n+1)}{6} - \left(\frac{n+1}{2}\right)^2 = \frac{n^2-1}{12}.$$

■

Exemple. Pour le lancer d'un dé équilibré où $X \sim \mathcal{U}_{\{1,\dots,6\}}$, on a ainsi $\mathbb{E}[X] = 3,5$ (milieu de l'intervalle $[1,6]$) et $\mathrm{Var}(X) = \frac{35}{12}$ donc $\sigma(X) = \sqrt{\frac{35}{12}} \approx 1,7$. La fonction de répartition de X est représentée figure 2.5.

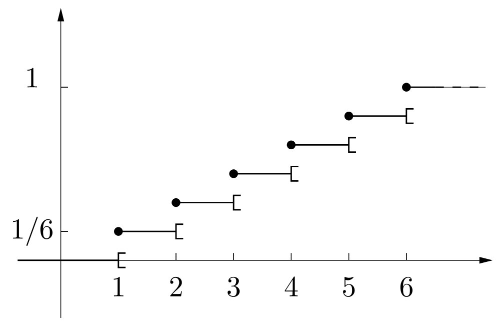


FIGURE 2.5 – Fonction de répartition d'une loi uniforme $\mathcal{U}_{\{1,\dots,6\}}$.

Remarque. Il va de soi que ces formules ne sont plus valables lorsque X suit une loi uniforme sur l'ensemble $\{a_1, \dots, a_n\}$. Dans ce cas général, on n'a rien de mieux que les définitions de l'espérance et de la variance :

$$\mathbb{E}[X] = \frac{1}{n} \sum_{k=1}^{n} a_k \qquad \& \qquad \mathrm{Var}(X) = \frac{1}{n} \sum_{k=1}^{n} a_k^2 - \left(\frac{1}{n} \sum_{k=1}^{n} a_k\right)^2.$$

Par exemple, la variable X valant $+1$ si le résultat du lancer d'un dé équilibré est pair, -1 sinon, a pour moyenne 0 et pour variance 1.

2.5.2 Loi de Bernoulli

On parle de loi de Bernoulli lorsque la variable d'intérêt est binaire.

Définition 2.11 (Loi de Bernoulli)
On dit que X suit une loi de Bernoulli de paramètre $p \in]0,1[$, noté $X \sim \mathcal{B}(p)$, si X ne peut prendre que les valeurs 0 et 1, avec $\mathbb{P}(X=1) = p$ et $\mathbb{P}(X=0) = 1-p = q$.

Exemple. On lance une pièce déséquilibrée dont la probabilité d'apparition de Pile est 3/4. En notant X la variable aléatoire valant 0 pour Face et 1 pour Pile, on a donc $X \sim \mathcal{B}(3/4)$.

Remarque. Plus généralement, la variable X suit une loi de Bernoulli de paramètre p sur $\{a, b\}$ si $P(X = a) = q = 1-p$ et $P(X = b) = p$. Lorsque $a = -1$ et $b = +1$, on parle de loi de Rademacher de paramètre p, noté $X \sim \mathcal{R}(p)$. Par exemple, la variable X valant $+1$ si le résultat du lancer d'un dé équilibré est pair, -1 sinon, suit une loi de Rademacher $\mathcal{R}(1/2)$.

Proposition 2.7 (Moments d'une loi de Bernoulli)
Si X suit une loi de Bernoulli de paramètre p, alors :

$$\mathbb{E}[X] = p \qquad \& \qquad \mathrm{Var}(X) = p(1-p).$$

Preuve. Laissée au lecteur.

■

Remarques :

1. Dans le cas général où X suit une loi de Bernoulli de paramètre p sur $\{a, b\}$, nous avons $\mathbb{E}[X] = a + p(b-a)$ et $\mathrm{Var}(X) = p(1-p)(b-a)^2$.
2. L'étude de la fonction $p \mapsto p(1-p)$ sur $]0,1[$ montre que, parmi les lois de Bernoulli, celle ayant le plus de variance est celle de paramètre $p = \frac{1}{2}$.

Les variables de Bernoulli sont souvent à la base de constructions plus sophistiquées : lois binomiales, lois géométriques, marche aléatoire sur $\mathbb{Z}$, etc. Ce sont d'ailleurs pour des variables de Bernoulli indépendantes et identiquement distribuées (en abrégé i.i.d.) qu'ont été établies les premières versions des deux grands théorèmes des probabilités : la Loi des Grands Nombres et le Théorème Central Limite.

2.5.3 Loi binomiale

La loi binomiale doit son nom aux coefficients binomiaux intervenant dans sa définition.

Définition 2.12 (Loi binomiale)
On dit que X suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0,1[$, noté $X \sim \mathcal{B}(n,p)$, si X est à valeurs dans $\{0, 1, \ldots, n\}$ avec :

$$\forall k \in \{0, 1, \ldots, n\} \qquad \mathbb{P}(X = k) = \binom{n}{k} p^k q^{n-k}.$$

où $q = (1-p)$ est introduit afin d'alléger (un peu) les notations.

La formule du binôme de Newton permet de vérifier qu'on définit bien ainsi une loi de probabilité :

$$\sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1.$$

Commençons par donner une façon naturelle de construire une loi binomiale $\mathcal{B}(n,p)$ à partir d'une loi de Bernoulli $\mathcal{B}(p)$.

Proposition 2.8 (Lien Bernoulli-Binomiale)
Soit $X_1, \ldots, X_n$ n variables indépendantes suivant la même loi de Bernoulli $\mathcal{B}(p)$, alors la variable $X = X_1 + \cdots + X_n$ suit une loi binomiale $\mathcal{B}(n,p)$.

Preuve. Voir le corrigé de l'exercice 2.9.

■

Exemple. On lance n fois de suite une pièce déséquilibrée dont la probabilité d'apparition de Pile à chaque lancer est p. En notant X la somme des résultats X_i obtenus (Face valant 0 et Pile valant 1 comme ci-dessus), X représente donc simplement le nombre de Pile sur les n lancers et on a $X \sim \mathcal{B}(n,p)$.

Proposition 2.9 (Moments d'une loi binomiale)
Si X suit une loi binomiale de paramètres (n,p), alors :

$$\mathbb{E}[X] = np \qquad \& \qquad \mathrm{Var}(X) = np(1-p) = npq.$$

Le lien Bernoulli-Binomiale rend ces formules élémentaires : l'espérance est linéaire dans tous les cas et la variance l'est ici puisque les variables $X_1, \dots, X_n$ sont indépendantes. La propriété suivante découle d'ailleurs du même raisonnement :

$$\left.\begin{array}{c} X \sim \mathcal{B}(n,p) \\ Y \sim \mathcal{B}(m,p) \\ X \perp\!\!\!\perp Y \end{array}\right\} \Rightarrow \ X+Y \sim \mathcal{B}(n+m,p).$$

Autrement dit, la somme de 2 variables binomiales indépendantes et de même paramètre p suit elle aussi une loi binomiale de paramètre p (voir le corrigé de l'exercice 2.9).

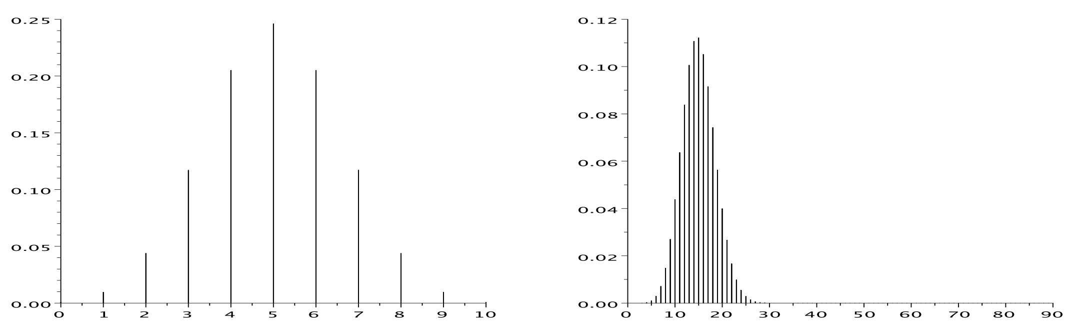


FIGURE 2.6 – Exemples de lois binomiales. A gauche : $X \sim \mathcal{B}(10, 1/2)$. A droite : $Y \sim \mathcal{B}(90, 1/6)$.

Exemples :

1. On lance 10 fois de suite une pièce équilibrée et on note X le nombre de Pile obtenus. On a vu que $X \sim \mathcal{B}(10, 1/2)$. Le nombre moyen de Pile est donc naturellement $\mathbb{E}[X] = 5$ et l'écart-type vaut $\sigma(X) = \sqrt{2.5}$. La loi de X est illustrée figure 2.6 à gauche.
2. On lance 90 fois de suite un dé non pipé et on note Y le nombre de 4 obtenus. Dans ce cas $Y \sim \mathcal{B}(90, 1/6)$. Le nombre moyen de 4 est donc $\mathbb{E}[Y] = 15$ et l'écart-type vaut $\sigma(Y) = \sqrt{12.5}$. La loi de Y est illustrée figure 2.6 à droite. Le Théorème Central Limite explique pourquoi on obtient une forme de courbe "en cloche" typique de la loi normale (cf. fin de Chapitre 3).

2.5.4 Loi géométrique

La loi géométrique est la loi typique du temps d'attente avant apparition d'un certain événement.

Définition 2.13 (Loi géométrique)
On dit que X suit une loi géométrique de paramètre $p \in]0,1[$, noté $X \sim \mathcal{G}(p)$, si X est à valeurs dans $\mathbb{N}^*$ avec :

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

Le nom de cette loi vient bien sûr du fait que la suite $(\mathbb{P}(X=n))_{n\geq 1}$ est géométrique de raison $(1-p)$. Il ne faut donc pas confondre **paramètre** p de la loi $\mathcal{G}(p)$ et **raison** $(1-p)$ de la suite $(\mathbb{P}(X=n))_{n\geq 1}$.

Exemple. On lance un dé équilibré et on appelle X l'indice de la première apparition du numéro 5. La propriété de continuité monotone décroissante permet de montrer que la probabilité que 5 n'apparaisse jamais est nulle, donc on exclut sans vergogne cette éventualité. Ceci fait, la variable aléatoire X prend ses valeurs dans $\mathbb{N}^*$, avec :

$$\forall n\in\mathbb{N}^* \qquad \mathbb{P}(X=n)=\frac{1}{6}\left(\frac{5}{6}\right)^{n-1},$$

c'est-à-dire que $X\sim\mathcal{G}(1/6)$. Cet exemple est typique de la loi géométrique : on répète une expérience jusqu'à la réalisation d'un événement dont la probabilité de réalisation à chaque coup est fixée et égale à p.

Proposition 2.10 (Moments d'une loi géométrique)
Si X suit une loi géométrique de paramètre p, alors :

$$\mathbb{E}[X]=\frac{1}{p} \qquad \& \qquad \mathrm{Var}(X)=\frac{1-p}{p^2}=\frac{q}{p^2}.$$

Preuve. Pour l'espérance, voir le corrigé de l'exercice 2.11. Disons simplement qu'elle est basée sur le développement en série entière suivant :

$$\forall x\in]-1,+1[\qquad \sum_{n=1}^{+\infty}nx^{n-1}=\frac{1}{(1-x)^2}.$$

Le calcul de la variance est lui-même basé sur la dérivation terme à terme de ce développement :

$$\forall x\in]-1,+1[\qquad \sum_{n=1}^{+\infty}n(n-1)x^{n-2}=\frac{2}{(1-x)^3},$$

ainsi que sur l'astuce déjà vue consistant à écrire $\mathbb{E}[X^2]=\mathbb{E}[X(X-1)]+\mathbb{E}[X]$, ce qui donne ici :

$$\mathbb{E}[X(X-1)]=\sum_{n=1}^{+\infty}n(n-1)pq^{n-1}=pq\sum_{n=1}^{+\infty}n(n-1)q^{n-2},$$

d'où :

$$\mathbb{E}[X(X-1)]=\frac{2pq}{(1-q)^3}=\frac{2q}{p^2},$$

et par suite : $\mathrm{Var}(X)=\mathbb{E}[X(X-1)]+\mathbb{E}[X]-\mathbb{E}[X]^2=\frac{q}{p^2}$.

■

Exemples :

1. Dans l'expérience du lancer de dé, le nombre moyen de lancers nécessaires pour voir apparaître le numéro 5 est $\mathbb{E}[X]=6$. L'écart-type est $\sigma(X)=\sqrt{30}\approx 5,5$. La loi de X est illustrée figure 2.7 à gauche.
2. On lance une pièce équilibrée et on note Y l'indice de la première apparition de Pile. Dans ce cas $Y\sim\mathcal{G}(1/2)$. Le temps moyen d'attente est donc $\mathbb{E}[Y]=2$ et l'écart-type vaut $\sigma(Y)=\sqrt{2}$. La loi de Y est illustrée figure 2.7 à droite.

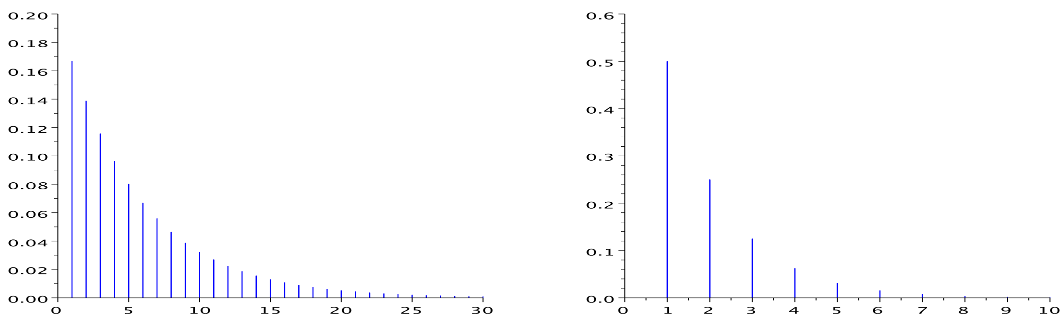


FIGURE 2.7 – Exemples de lois géometriques. A gauche : $X \sim \mathcal{G}(1/6)$. A droite : $Y \sim \mathcal{G}(1/2)$.

Supposons toujours que vous commenciez à lancer votre dé jusqu'à apparition du numéro 5 et qu'après 3 lancers le 5 ne soit toujours pas apparu. Question : quelle est la loi du nouveau temps d'attente jusqu'à apparition du 5 ? Réponse : la même qu'initialement, i.e. une loi géométrique de paramètre $1/6$. Cette propriété, dite d'absence de mémoire, est typique de la loi géométrique (parmi les lois discrètes).

Proposition 2.11 (Absence de mémoire)
Si X suit une loi géométrique de paramètre p, alors :

$$\forall (m,n) \in \mathbb{N} \times \mathbb{N} \qquad \mathbb{P}(X > m+n | X > m) = \mathbb{P}(X > n).$$

Preuve. Voir le corrigé de l'exercice 2.16.

■

Remarque. Parmi les lois à densité que nous verrons au chapitre suivant, la loi exponentielle est la seule à posséder cette propriété. Rien d'étonnant dans cette histoire : de même qu'une suite géométrique peut être considérée comme la version discrète d'une fonction exponentielle, la loi géométrique peut être vue comme la discrétisation en temps de la loi exponentielle (voir exercice 3.16).

Les lois géométriques possèdent une autre propriété remarquable, à savoir leur stabilité par minimisation. La preuve du résultat suivant est donnée dans le corrigé de l'exercice 2.17.

Proposition 2.12 (Minimum de lois géométriques)
Soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois géométriques de paramètres respectifs $p_1, \dots, p_n$, avec $p_i \in]0,1[$ pour tout $i \in \{1, \dots, n\}$. Alors la variable aléatoire $X = \min(X_1, \dots, X_n)$ suit elle-même une loi géométrique, plus précisément

$$X = \min(X_1, \dots, X_n) \sim \mathcal{G}(1 - (1-p_1) \dots (1-p_n)).$$

Le paramètre de cette loi géométrique est clair : notons $E_1, \dots, E_n$ des événements indépendants de probabilités $p_1, \dots, p_n$. La probabilité qu'au moins l'un d'entre eux se réalise est égale à

$$p = \mathbb{P}(E_1 \cup \dots \cup E_n) = 1 - \mathbb{P}(\overline{E_1 \cup \dots \cup E_n}) = 1 - \mathbb{P}(\overline{E_1} \cap \dots \cap \overline{E_1}),$$

et grâce à l'indépendance :

$$p = 1 - \mathbb{P}(\overline{E_1}) \dots \mathbb{P}(\overline{E_n}) = 1 - (1 - \mathbb{P}(E_1)) \dots (1 - \mathbb{P}(E_n)) = 1 - (1-p_1) \dots (1-p_n).$$

2.5.5 Loi de Poisson

Par rapport aux sections précédentes, on peine à donner une modélisation vraiment élémentaire de la loi de Poisson. On verra son interprétation comme loi d'événements rares en fin de section, mais pour l'instant on se contentera de la définir brutalement.

Définition 2.14 (Loi de Poisson)
On dit que X suit une loi de Poisson de paramètre $\lambda > 0$, noté $X \sim \mathcal{P}(\lambda)$, si X est à valeurs dans $\mathbb{N}$ avec :

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

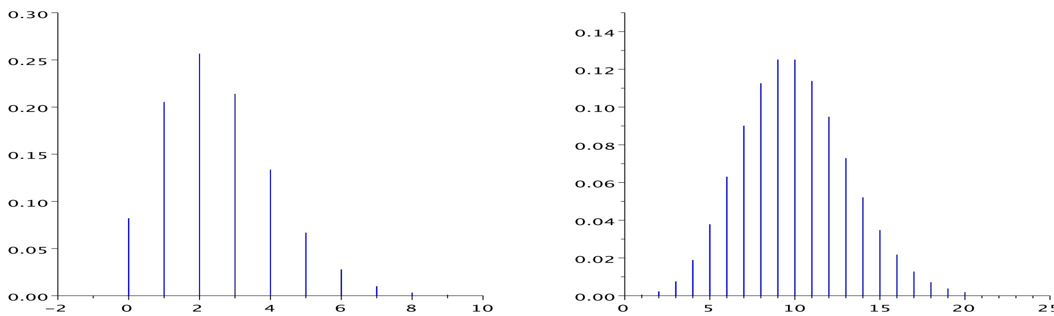


FIGURE 2.8 – Lois de Poisson $\mathcal{P}(2.5)$ et $\mathcal{P}(10)$.

Concernant le mode d'une loi de Poisson, il convient de distinguer deux cas (cf. corrigé de l'exercice 2.23) :

1. si $\lambda \notin \mathbb{N}^*$, on a un seul mode, atteint pour n égal à la partie entière de λ (voir figure 2.8 à gauche) :
$$\max_{n \in \mathbb{N}} p_n = p_{\lfloor \lambda \rfloor} = e^{-\lambda}\frac{\lambda^{\lfloor \lambda \rfloor}}{\lfloor \lambda \rfloor !}.$$
2. si $\lambda \in \mathbb{N}^*$, il y a cette fois deux modes (voir figure 2.8 à droite) :
$$\max_{n \in \mathbb{N}} p_n = p_{\lambda - 1} = p_\lambda = e^{-\lambda}\frac{\lambda^{\lambda}}{\lambda !}.$$

Rappel. Les calculs sur la loi exponentielle font intervenir le développement en série entière de l'exponentielle, qu'il convient donc d'avoir en tête :

$$\forall x \in \mathbb{R} \qquad e^x = \sum_{n=0}^{+\infty} \frac{x^n}{n!}.$$

Ceci est en particulier utile pour le calcul des moments.

Proposition 2.13 (Moments d'une loi de Poisson)
Si X suit une loi de Poisson de paramètre $\lambda > 0$, alors :

$$\mathbb{E}[X] = \lambda \qquad \& \qquad \mathrm{Var}(X) = \lambda.$$

Preuve. Les calculs ont été faits en Section 2.3 dans le cas particulier où $X \sim \mathcal{P}(1)$. La généralisation ne pose aucun problème.

■

Nous avons vu en section précédente que les lois géométriques sont stables par minimisation. Les lois de Poisson, elles, le sont par sommation. La preuve du résultat suivant est donnée dans le corrigé de l'exercice 2.26.

Proposition 2.14 (Somme de variables de Poisson)
Soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois de Poisson de paramètres respectifs $\lambda_1, \dots, \lambda_n$, avec $\lambda_i > 0$ pour tout $i \in \{1, \dots, n\}$, alors la variable $X = X_1 + \cdots + X_n$ suit elle aussi une loi de Poisson, et plus précisément

$$X = X_1 + \cdots + X_n \sim \mathcal{P}(\lambda_1 + \cdots + \lambda_n).$$

Sous certaines conditions, la loi de Poisson peut être vue comme une approximation de la loi binomiale.

Proposition 2.15 (Lien Binomiale-Poisson)
Soit $(p_n)_{n \geq 0}$ une suite de réels compris entre 0 et 1 telle que $\lim_{n\to\infty} np_n = \lambda > 0$. Pour tout $n \geq 0$, soit X_n une variable aléatoire de loi binomiale $\mathcal{B}(n, p_n)$, alors :

$$\forall k \in \mathbb{N} \qquad \mathbb{P}(X_n = k) \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Preuve. Voir le corrigé de l'exercice 2.6.

■

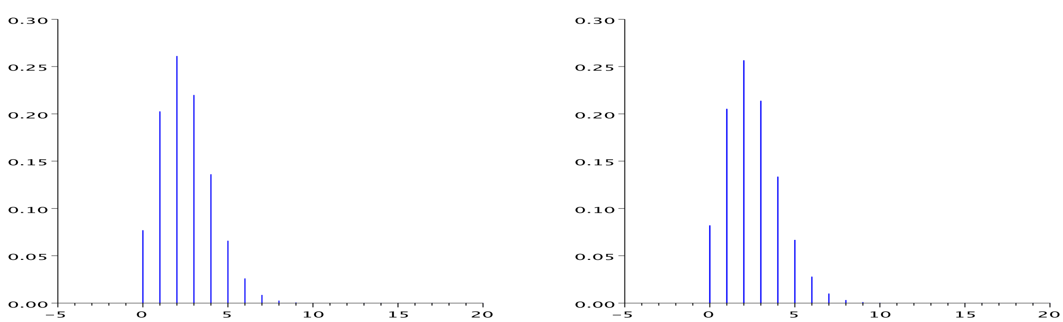


FIGURE 2.9 – Loi binomiale $\mathcal{B}(50, 1/20)$ et loi de Poisson $\mathcal{P}(2.5)$.

Ainsi, pour aller vite, lorsque n est grand, on a $\mathcal{B}(n, p_n) \approx \mathcal{P}(np_n)$. Plus précisément, on dit que $(X_n)_{n\geq 0}$ converge en loi vers une variable aléatoire de loi de Poisson de paramètre λ. C'est parce qu'on peut la voir comme l'approximation d'une loi binomiale lorsque le paramètre p est petit qu'on parle de **loi des événements rares** pour la loi de Poisson. A la louche, on considère souvent que l'approximation est acceptable pour $n > 50$ et $np < 5$. Pour $n = 50$ et $p = 0.05$, la figure 2.9 donne l'allure des deux lois en question.

Exemple. Cette idée est illustrée dans l'exercice 2.22, traitant du surbooking dans les avions : 94 places sont vendues au lieu de 90, car chaque passager ayant réservé a une probabilité (faible, mais réelle, estimée ici à 5%) de ne pas se présenter pour l'embarquement. Le nombre de passagers effectivement absents à l'enregistrement suit donc une loi binomiale $\mathcal{B}(94; 0,05)$, qui peut être approchée par une loi de Poisson $\mathcal{P}(4.7)$.

Remarque. Si n est grand mais si p n'est pas très petit (de sorte que typiquement on n'ait pas $np < 5$), l'approximation de la loi binomiale par une loi de Poisson n'est plus valide. Néanmoins, on verra en fin de Chapitre 3 que le Théorème Central Limite permet alors d'approcher la loi binomiale par une loi normale. Ce phénomène a d'ailleurs été illustré figure 2.6 à droite pour $n = 90$ et $p = 1/6$.

2.6 Exercices

Exercice 2.1 (Loi uniforme)

1. On jette une pièce équilibrée. On appelle X la variable aléatoire qui vaut 0 si on obtient Face et 1 si on obtient Pile. Représenter la fonction de répartition de X.
2. On jette un dé équilibré et on appelle X le résultat du lancer. Donner la loi de X et représenter sa fonction de répartition.
3. De façon générale, on dit que U suit une loi uniforme sur l'ensemble $\{1, \ldots, n\}$ et on note $U \sim \mathcal{U}_{\{1,\ldots,n\}}$ si pour tout i entre 1 et n : $\mathbb{P}(U = i) = 1/n$. Représenter la fonction de répartition de U.
4. Lors d'une visite médicale, n patients de tailles différentes se présentent devant le médecin, et ce dans un ordre aléatoire. On note X le rang de présentation du plus grand d'entre eux. Donner la loi de la variable X.

Exercice 2.2 (Loi de Bernoulli)

1. On jette une pièce dont la probabilité d'apparition de Pile est 2/3. On appelle X la variable aléatoire qui vaut 0 si on obtient Face et 1 si on obtient Pile. Représenter la fonction de répartition de X.
2. De façon générale, on dit que X suit une loi de Bernoulli de paramètre p et on note $X \sim \mathcal{B}(p)$ si X est à valeurs dans $\{0, 1\}$ avec : $\mathbb{P}(X = 1) = p$. Représenter la fonction de répartition de X.
3. D'une urne contenant B boules blanches, N boules noires et 1 boule rouge, on tire simultanément n boules (avec bien sûr $n \leq (B+N+1)$). On appelle X la variable aléatoire égale à 1 si la boule rouge est tirée, 0 sinon : elle suit donc une loi de Bernoulli. Donner son paramètre.
4. Un étudiant aviné sort de chez un ami un jeudi soir comme les autres à Rennes. A l'instant $n = 0$ il est en O et se déplace à chaque instant entier de +1 ou de -1, et ce de façon équiprobable. Soit Y_n la variable aléatoire égale à 1 si à l'instant $2n$ l'étudiant se retrouve à nouveau en son point de départ, et $Y_n = 0$ sinon. Donner le paramètre de cette loi de Bernoulli.

Exercice 2.3 (Loi binomiale)

1. On lance n fois de suite une pièce équilibrée et on note X la somme des résultats obtenus, avec la convention 0 pour Face et 1 pour Pile. Donner la loi de X.
2. On dit que X suit une loi binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in]0, 1[$ et on note $X \sim \mathcal{B}(n, p)$ si X est à valeurs dans $\{0, 1, \ldots, n\}$ avec pour tout $k \in \{0, 1, \ldots, n\}$: $p_k = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. Vérifier que c'est bien une loi de probabilité, c'est-à-dire que la somme des p_k vaut bien 1. Quel lien peut-on faire entre une loi binomiale $\mathcal{B}(n, p)$ et la loi de Bernoulli $\mathcal{B}(p)$?
3. Dans un magasin il y a n clients et m caisses. Chaque client choisit une caisse au hasard et on appelle X le nombre de clients choisissant la caisse numéro 1. Donner la loi de X.

Exercice 2.4 (Loi hypergéométrique)

1. Dans une population de taille N donnée, la proportion favorable à un candidat donné est p. Afin d'estimer p, on interroge un échantillon de n personnes et on appelle X le nombre d'électeurs favorables à ce candidat dans cet échantillon. Donner la loi de X.
2. On dit que X suit une loi hypergéométrique de paramètres $N \in \mathbb{N}^*$, $n \in \{1, \dots, N\}$ et $p = 1 - q \in]0, 1[$ et on note $X \sim \mathcal{H}(N, n, p)$ si X est à valeurs dans $\{0, 1, \dots, n\}$ avec
$$\forall k \in \{0, 1, \dots, n\} \qquad p_k = \mathbb{P}(X = k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$
Vérifier que c'est bien une loi de probabilité, c'est-à-dire que la somme des p_k vaut bien 1.
3. Montrer que lorsque n est très petit devant N, la loi hypergéométrique $\mathcal{H}(N, n, p)$ et la loi binomiale $\mathcal{B}(n, p)$ se ressemblent comme deux gouttes d'eau.

Exercice 2.5 (Loi géométrique)

1. On lance une pièce équilibrée jusqu'à la première apparition de Pile. Quelle est la probabilité que Pile n'apparaisse jamais (on pourra appeler E_n l'événement : "Pile n'apparaît pas durant les n premiers lancers" et appliquer la continuité monotone décroissante) ? On exclut ce cas et on appelle X la variable correspondant à la première apparition de Pile. Donner la loi de X et représenter sa fonction de répartition.
2. On lance un dé équilibré et on appelle X la variable correspondant à la première apparition du numéro 1 (avec la même hypothèse que dans la question précédente). Donner la loi de X et représenter sa fonction de répartition.
3. On dit que X suit une loi géométrique de paramètre $p \in]0, 1[$ et on note $X \sim \mathcal{G}(p)$ si X est à valeurs dans $\mathbb{N}^* = \{1, 2, \dots\}$ avec pour tout $n \in \mathbb{N}^*$: $p_n = \mathbb{P}(X = n) = p(1-p)^{n-1}$. Vérifier que c'est bien une loi de probabilité, c'est-à-dire que la somme des p_n vaut bien 1. Quel lien peut-on faire entre la loi $\mathcal{G}(p)$ et la loi de Bernoulli $\mathcal{B}(p)$?

Exercice 2.6 (Loi de Poisson)

1. On dit que X suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$ et on note $X \sim \mathcal{P}(\lambda)$ si X est à valeurs dans $\mathbb{N}$ avec :
$$\forall n \in \mathbb{N} \qquad p_n = \mathbb{P}(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}.$$
Vérifier que c'est bien une loi de probabilité, c'est-à-dire que la somme des p_n vaut bien 1.
2. Représenter la suite des $(p_n)_{n\geq 0}$ pour $\lambda = 2$, puis pour $\lambda = 20$.
3. Soit $\lambda > 0$ fixé. On lance n fois une pièce amenant Pile avec la probabilité $p_n = \lambda/n$. Soit X_n le nombre de fois où Pile apparaît durant ces n lancers.
   (a) Quelle est la loi de X_n ? Rappeler en particulier $\mathbb{P}(X_n = k)$.
   (b) Pour k fixé entre 0 et n, montrer que $\lim_{n\to\infty} \mathbb{P}(X_n = k) = e^{-\lambda}\frac{\lambda^k}{k!}$. Autrement dit, lorsque n devient grand, la loi binomiale "ressemble" à une loi de Poisson.
   (c) Montrer que le résultat précédent est encore vrai si on suppose juste que :
$$\lim_{n\to+\infty} np_n = \lambda.$$
4. Lors du championnat de France de Ligue 1 de football (saison 2004-2005), on a relevé le nombre de buts marqués par équipe et par match lors de l'ensemble des 38 journées de championnat. En tout, 824 buts ont été marqués lors des 380 matchs disputés (cf. tableau ci-dessous). Soit X la variable correspondant au nombre de buts marqués par une équipe et par match. Déduire la loi de X du tableau. Via $\mathbb{P}(X = 0)$, trouver un paramètre λ tel que la loi de X ressemble à une loi de Poisson $\mathcal{P}(\lambda)$.

Buts marqués par équipe et par match	0	1	2	3	4	5	6	7	8	Total
Équipes ayant marqué ce nombre de buts	268	266	152	53	13	7	0	0	1	760

Exercice 2.7 (Espérance d'une loi uniforme)

1. On jette une pièce équilibrée. On appelle X la variable aléatoire qui vaut 0 si on obtient Face et 1 si on obtient Pile. Donner l'espérance de X.
2. On jette un dé équilibré et on appelle X le résultat du lancer. Que vaut $\mathbb{E}[X]$?
3. Rappeler ce que vaut la somme $1 + 2 + \cdots + n$. En déduire l'espérance de U lorsque $U \sim \mathcal{U}_{\{1,\ldots,n\}}$.
4. On reprend l'exercice sur le gardien de nuit du chapitre 1, lequel a 10 clés pour ouvrir une porte. Dans le cas où il est à jeun et élimine chaque clé après un essai infructueux, quel est le nombre moyen d'essais nécessaires pour ouvrir la porte ?
5. Soit m et n dans $\mathbb{N}^*$. Soit X une variable aléatoire à valeurs dans $\{1, 2, \ldots, mn\}$ et telle que $\forall i \in \{1, 2, \ldots, mn\}$, $\mathbb{P}(X = i) = 1/m - 1/n$.
   (a) Déterminer m en fonction de n pour qu'on ait bien une loi de probabilité.
   (b) Déterminer $\mathbb{E}[X]$ et trouver n tel que $\mathbb{E}[X] = 7/2$.

Exercice 2.8 (Espérance d'une loi de Bernoulli)

1. On jette une pièce dont la probabilité d'apparition de Pile est $2/3$. On appelle X la variable aléatoire qui vaut 0 si on obtient Face et 1 si on obtient Pile. Quelle est l'espérance de X ?
2. De façon générale, que vaut $\mathbb{E}[X]$ lorsque $X \sim \mathcal{B}(p)$?
3. Une roulette a 37 numéros : 18 rouges, 18 noirs et 1 vert (le zéro).
   (a) Si vous misez sur une couleur et que cette couleur sort, vous récupérez 2 fois votre mise, sinon vous perdez votre mise. Supposons que vous misiez 1€ sur rouge, quel est votre bénéfice moyen ?
   (b) Si vous misez sur un numéro (le zéro étant exclu) et que ce numéro sort, vous récupérez 36 fois votre mise, sinon vous la perdez. Supposons que vous misiez 1€ sur le 13, quel est votre bénéfice moyen ?

Exercice 2.9 (Espérance d'une loi binomiale)

1. Calculer $\mathbb{E}[X]$ lorsque X suit une loi binomiale $\mathcal{B}(n, p)$ (on pourra s'inspirer de l'exercice V du TD1).
2. Soit $X_1, \ldots, X_n$ n variables aléatoires indépendantes et identiquement distribuées (en abrégé i.i.d.) suivant la loi de Bernoulli $\mathcal{B}(p)$. Rappeler la loi de $X = X_1 + \cdots + X_n$. Retrouver alors le résultat de la question précédente.
3. Soit $X_1, \ldots, X_{n+m}$ variables aléatoires i.i.d. suivant la loi de Bernoulli $\mathcal{B}(p)$. Loi de $X = X_1 + \cdots + X_n$? Loi de $Y = X_{n+1} + \cdots + X_{n+m}$? Loi de $Z = X + Y$?
4. Soit X une variable aléatoire qui suit une loi binomiale $\mathcal{B}(n, 1/2)$. Chaque réalisation de X est affichée sur un compteur qui est détraqué comme suit : si X n'est pas nul, le compteur affiche la vraie valeur de X ; si X est nul, le compteur affiche un nombre au hasard entre 1 et n (i.e. tiré suivant une loi $\mathcal{U}_{\{1,\ldots,n\}}$). Soit Y la variable aléatoire égale au nombre affiché. Déterminer la loi de Y et son espérance.

Exercice 2.10 (Espérance d'une loi hypergéométrique)

1. On considère n et N dans $\mathbb{N}^*$ avec $n \leq N$ et $p \in]0, 1[$ tel que $Np \in \mathbb{N}^*$. Pour fixer les idées, notre modèle est celui d'une population de taille N dans laquelle Np votants sont favorables au candidat A. Afin d'estimer p on tire au hasard sans remise n individus dans la population et on appelle X le nombre de votants pour A dans cet échantillon. Rappeler la loi suivie par X.

2. On numérote de 1 à Np les Np votants pour le candidat A. Pour tout $k \in \{1, \ldots, Np\}$, on appelle X_k la variable aléatoire qui vaut 1 si l'individu k fait partie de l'échantillon, 0 sinon.
   (a) Donner la relation entre X et $X_1, \ldots, X_{Np}$.
   (b) Montrer que $\mathbb{E}[X_1] = n/N$. En déduire que $\mathbb{E}[X] = np$. Comparer à la moyenne d'une variable aléatoire distribuée suivant une loi binomiale $\mathcal{B}(n, p)$.
3. On tire au hasard et sans remise 5 cartes d'un jeu de 32 cartes. Soit X la variable aléatoire égale au nombre de rois obtenus. Donner sans calculs la loi et l'espérance de X.

Exercice 2.11 (Espérance d'une loi géométrique)
On rappelle que pour tout $x \in]-1, +1[$, on a $\sum_{n=1}^{+\infty} nx^{n-1} = \frac{1}{(1-x)^2}$ et $\sum_{n=1}^{+\infty} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$.

1. On lance un dé équilibré et on appelle X la variable correspondant à la première apparition du numéro 1. Rappeler la loi de X et calculer $\mathbb{E}[X]$.
2. Généralisation : soit $X \sim \mathcal{G}(p)$, que vaut $\mathbb{E}[X]$?
3. On reprend l'exercice sur le gardien de nuit du TD1, lequel a 10 clés pour ouvrir une porte. Dans le cas où il est ivre et remet chaque clé dans le trousseau après un essai infructueux, quel est le nombre moyen d'essais nécessaires pour ouvrir la porte ?
4. On considère des polygones convexes dont le nombre N de côtés est une variable aléatoire ayant pour loi $\mathbb{P}(N = n) = 2^{2-n}$ pour tout $n \geq 3$. Quel est l'espérance du nombre de côtés du polygone ? Quel est l'espérance du nombre de diagonales du polygone ?

Exercice 2.12 (Espérance d'une loi de Poisson)

1. Soit X une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$. Montrer que $\mathbb{E}[X] = \lambda$.
2. Toujours pour $X \sim \mathcal{P}(\lambda)$, on considère alors la variable aléatoire $Y = e^{-X}$. Calculer $\mathbb{E}[Y]$.
3. Un athlète tente de franchir des hauteurs successives numérotées 1,2,$\ldots, n, \ldots$ Il a le droit à un seul essai par hauteur, s'il échoue il est éliminé. On suppose que les sauts sont indépendants les uns des autres et que la probabilité de succès au n-ème saut est $r_n = 1/n$ pour tout $n \in \mathbb{N}^*$. On note X la variable aléatoire égale au numéro du dernier saut réussi.
   (a) Montrer que $\forall n \in \mathbb{N}^* : p_n = \mathbb{P}(X = n) = n/(n+1)!$. Via l'écriture $n = (n+1) - 1$, vérifier qu'on a bien $\sum_{n=1}^{+\infty} p_n = 1$.
   (b) Montrer que $\mathbb{E}[X + 1] = e$. En déduire $\mathbb{E}[X]$.

Exercice 2.13 (Espérance d'une loi arithmétique)
Soit X une variable aléatoire à valeurs dans $\{0, 1, \ldots, n\}$ et telle que $\forall k \in \{0, 1, \ldots, n\}$, $p_k = \mathbb{P}(X = k) = \alpha k$.

1. Déterminer α pour que X soit effectivement une variable aléatoire.
2. On rappelle que $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$. En déduire $\mathbb{E}[X]$.

Exercice 2.14 (Deux dés : somme et différence)
On lance deux dés équilibrés. On note U_1 et U_2 les variables aléatoires correspondant aux résultats obtenus.

1. Rappeler la loi de U_1, son espérance et sa variance.
2. On appelle $X = (U_1 + U_2)$ la somme et $Y = (U_1 - U_2)$ la différence des deux résultats. Que valent $\mathbb{E}[X]$ et $\mathbb{E}[Y]$? Montrer que $\mathbb{E}[XY] = 0$.
3. En déduire que X et Y sont décorrélées. Sont-elles indépendantes ?

Exercice 2.15 (Deux dés : min et max)
On lance deux dés équilibrés. On note U_1 et U_2 les variables aléatoires correspondant aux résultats obtenus. On appelle $X = \min(U_1, U_2)$ le minimum et $Y = \max(U_1, U_2)$ le maximum des deux dés.

1. Donner la loi de X. En déduire $\mathbb{E}[X]$.
2. Exprimer $X+Y$ en fonction de U_1 et U_2. En déduire $\mathbb{E}[Y]$.
3. Exprimer XY en fonction de U_1 et U_2. En déduire $\mathbb{E}[XY]$, puis $\text{Cov}(X, Y)$.

Exercice 2.16 (Memento (absence de mémoire))
Soit $X \sim \mathcal{G}(p)$ loi géométrique de paramètre $p \in]0, 1[$.

1. Soit $n \in \mathbb{N}$. Exprimer $\mathbb{P}(X > n)$ en fonction de p et n. Quel est le lien avec $F(n)$, où F est la fonction de répartition de X ?
2. En déduire la propriété dite “ d'absence de mémoire ” de la loi géométrique, à savoir que :

$$\forall (m, n) \in \mathbb{N} \times \mathbb{N} \qquad \mathbb{P}(X > n + m | X > m) = \mathbb{P}(X > n).$$

Exercice 2.17 (Minimum de lois géométriques)
Soit $X_1 \sim \mathcal{G}(p_1)$ où $p_1 \in]0, 1[$, $X_2 \sim \mathcal{G}(p_2)$ où $p_2 \in]0, 1[$, avec X_1 et X_2 indépendantes. Notons $X = \min(X_1, X_2)$ le minimum de ces deux variables.

1. Quelles sont les valeurs que peut prendre la variable aléatoire X ?
2. Soit $n \in \mathbb{N}$. Exprimer $\mathbb{P}(X > n)$ en fonction de p_1, p_2 et n. En déduire la loi de X.
3. Application : on a en main deux dés qu'on lance en même temps jusqu'à ce qu'apparaisse le numéro 2 sur au moins l'un des deux dés. Quel est le nombre moyen de “ doubles lancers ” nécessaires ?
4. Généralisation : soit n variables indépendantes $X_1, \ldots, X_n$ suivant des lois géométriques de paramètres respectifs $p_1, \ldots, p_n$, avec $p_i \in]0, 1[$ pour tout $i \in \{1, \ldots, n\}$. Donner la loi de la variable aléatoire $X = \min(X_1, \ldots, X_n)$.

Exercice 2.18 (Un problème de natalité)
Supposons qu'à la naissance, la probabilité qu'un nouveau-né soit un garçon est de 1/2. Supposons encore que tout couple engendre jusqu'à obtention d'un garçon. Le but est de trouver la proportion de garçons dans ce modèle théorique.

1. Notons X le nombre d'enfants d'un couple. Donner la loi de la variable aléatoire X.
2. Soit P la proportion de garçons parmi les enfants d'un couple. Exprimer P en fonction de X.
3. En déduire que $\mathbb{E}[P] = \ln 2 \approx 0.69$ (on rappelle que pour tout $x \in [-1, 1[$, $\ln(1 - x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n}$).

Exercice 2.19 (Tirages de cartes)
On note X la variable aléatoire égale au nombre de rois obtenus lorsqu'on tire successivement 5 cartes avec remise dans un jeu de 32 cartes. Préciser la loi de X, son espérance et sa variance.

Exercice 2.20 (Codage redondant)
Un canal de transmission ne peut traiter que des 0 et des 1. En raison des perturbations sur ce canal, un 0 peut être transformé en 1 et un 1 en 0 lors d'une transmission, et ce avec la même probabilité $p = 0,2$ indépendamment à chaque instant. Pour diminuer la probabilité d'erreur, on décide de transmettre 00000 à la place de 0 et 11111 à la place de 1 (codage dit redondant). Si le récepteur décode suivant la règle de la majorité, quelle est la probabilité que le message soit mal interprété ?

Exercice 2.21 (Inégalité de Tchebychev)
On jette 3600 fois un dé et on appelle S le nombre de fois où apparaît le numéro 1.

1. Quelle est la loi de S ? Donner sa moyenne et sa variance.

2. Exprimer sous forme d'une somme la probabilité que ce nombre soit compris strictement entre 480 et 720. Grâce à l'inégalité de Tchebychev, minorer cette probabilité.

Exercice 2.22 (Surbooking)
Des études effectuées par une compagnie aérienne montrent qu'il y a une probabilité 0,05 qu'un passager ayant fait une réservation n'effectue pas le vol. Dès lors, elle vend toujours 94 billets pour ses avions à 90 places. Quelle est la probabilité pour qu'il y ait un problème à l'embarquement ? Indication : pour l'application numérique, on pourra effectuer l'approximation d'une loi binomiale par une loi de Poisson.

Exercice 2.23 (Mode(s) d'une loi de Poisson)
Soit X une variable aléatoire qui suit une loi de Poisson de paramètre $\lambda \in \mathbb{R}_+^*$. Pour tout $n \in \mathbb{N}$, on note $p_n = \mathbb{P}(X = n)$.

1. Former et simplifier le rapport $r_n = p_{n+1}/p_n$.
2. En déduire le (ou les) mode(s) de la loi de Poisson $\mathcal{P}(\lambda)$, c'est-à-dire la (ou les) valeur(s) de n telle(s) que la probabilité p_n soit maximale. Indication : on distinguera les deux cas $\lambda \in \mathbb{N}^*$ et $\lambda \notin \mathbb{N}^*$.

Exercice 2.24 (Parité d'une loi de Poisson)
Soit $\lambda \in \mathbb{R}_+^*$ fixé, on note :

$$S_1 = \sum_{k=0}^{+\infty} \frac{\lambda^{2k}}{(2k)!} \qquad \& \qquad S_2 = \sum_{k=0}^{+\infty} \frac{\lambda^{2k+1}}{(2k+1)!}.$$

1. Calculer $(S_1 + S_2)$ et $(S_1 - S_2)$. En déduire S_1 et S_2.
2. Application : le nombre N de clients entrant dans un magasin en une journée suit une loi de Poisson de paramètre λ. Pierre et Paul distribuent des prospectus aux clients, à raison de 1 prospectus par client. Pierre parie qu'à la fin de la journée, ils auront distribué un nombre pair de prospectus, tandis que Paul soutient qu'ils en auront distribué un nombre impair. Qui gagne en moyenne ?
3. Soit $X \sim \mathcal{P}(\lambda)$. On définit la variable Y de la façon suivante :
   – si X prend une valeur paire, alors $Y = X/2$;
   – si X prend une valeur impaire, alors $Y = 0$.
   Déterminer la loi de Y, son espérance et sa variance.

Exercice 2.25 (Espérance d'une variable à valeurs entières)
Soit $n \in \mathbb{N}^*$ fixé et X une variable aléatoire à valeurs dans $\{1, \dots, n\}$.

1. Rappeler la définition de $\mathbb{E}[X]$.
2. Justifier la formule
$$\mathbb{E}[X] = \mathbb{P}(X \geq 1) + \mathbb{P}(X \geq 2) + \cdots + \mathbb{P}(X \geq n) = \sum_{i=1}^{n} \mathbb{P}(X \geq i).$$
3. On jette 4 dés équilibrés simultanément et on appelle X le minimum obtenu sur ces 4 lancers.
   (a) Quelles valeurs peut prendre la variable aléatoire X ?
   (b) Calculer $\mathbb{P}(X \geq i)$ pour chaque valeur i que peut prendre X.
   (c) En déduire $\mathbb{E}[X]$.
   (d) Soit S la somme des 3 plus gros scores. Déterminer $\mathbb{E}[S]$ (on pourra remarquer que $T = S + X$, où T est le total des quatre dés).

   (e) Déduire des $\mathbb{P}(X \geq i)$ la loi de X, c'est-à-dire $\mathbb{P}(X = i)$ pour chaque valeur i.

4. Généralisation : soit X variable aléatoire à valeurs dans $\mathbb{N}^*$ admettant une espérance.
   (a) En vous inspirant de ce qui précède, donner une nouvelle formulation de $\mathbb{E}[X]$ (on ne demande pas de justifier la convergence de la série).
   (b) On a quatre dés équilibrés en main qu'on lance en même temps jusqu'à ce qu'apparaisse le numéro 2 sur au moins l'un des quatre dés. Quel est le nombre moyen de "quadruples lancers" nécessaires ?

Exercice 2.26 (Somme de variables poissoniennes)
Soit $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$, avec X_1 et X_2 indépendantes. On note $X = (X_1 + X_2)$ la somme de ces deux variables.

1. Soit $n \in \mathbb{N}$ fixé. Justifier la décomposition : $\{X = n\} = \bigcup_{k=0}^{n} \{X_1 = k, X_2 = n - k\}$.
2. En déduire que $X \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
3. Généralisation : soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois de Poisson de paramètres respectifs $\lambda_1, \dots, \lambda_n$, avec $\lambda_i > 0$ pour tout $i \in \{1, \dots, n\}$. Quelle est la loi de la variable $X = X_1 + \cdots + X_n$?

Exercice 2.27 (Un calendrier stochastique)
Supposons que chacune des 12 faces d'un dodécaèdre équilibré corresponde à un mois de l'année. On lance ce dé et on appelle X la variable aléatoire correspondant au nombre de jours du mois obtenu (on considère une année non bissextile).

1. Quelles valeurs peut prendre la variable X ? Avec quelles probabilités ?
2. Représenter la fonction de répartition de X.
3. Calculer l'espérance de X, ainsi que son écart-type.
4. Supposons maintenant qu'on tire au hasard un jour de l'année (toujours supposée non bissextile) et qu'on appelle Y le nombre de jours du mois correspondant à ce tirage. Quelles valeurs peut prendre la variable Y ? Avec quelles probabilités ? Calculer la moyenne de Y.

Exercice 2.28 (Loi à partir des moments)
On considère une variable aléatoire X à valeurs dans $\{0, 1, 2\}$. On sait que $\mathbb{E}[X] = 1$ et $\mathrm{Var}(X) = \frac{1}{2}$. En déduire la loi de X.

Exercice 2.29 (Dés et accidents)

1. On dispose de deux dés qu'on lance simultanément 12 fois de rang et on appelle X le nombre de double six obtenus sur les 12 lancers.
   (a) Quelle est la loi de X ? Donner sa moyenne et sa variance.
   (b) Calculer $\mathbb{P}(X \leq 2)$.
   (c) Que vaut cette quantité si on effectue l'approximation par une loi de Poisson ?
2. Sur la voie express Rennes-Nantes, il y a en moyenne 2 accidents par semaine, cette variable suivant approximativement une loi de Poisson.
   (a) Une semaine, il y a eu 4 accidents. Quelle était la probabilité d'un tel événement ?
   (b) Lorsque X_1 et X_2 sont indépendantes et suivent des lois de Poisson de paramètres respectifs λ_1 et λ_2, quelle est la loi suivie par la variable $X_1 + X_2$?
   (c) En déduire la probabilité qu'il se passe 2 semaines sans accident.

Exercice 2.30 (Test sanguin)
Chacun des soldats d'une troupe de 500 hommes est porteur d'une certaine maladie avec une probabilité 1/1000, indépendamment les uns des autres.

1. Soit X le nombre de soldats porteurs de cette maladie. Quelle est la loi de X ? Rappeler sa moyenne.
2. Par quelle loi peut-on approcher celle de X ? Dans la suite, on pourra faire les calculs avec la loi exacte ou utiliser cette approximation.
3. Cette maladie est détectable à l'aide d'un test sanguin et, pour faciliter les choses, on ne teste qu'un mélange du sang de chacun des 500 soldats. Quelle est la probabilité que le test soit positif, c'est-à-dire qu'au moins une des personnes soit malade ?
4. On suppose que le test a été positif. Dans ce cas, quelle est la probabilité qu'au moins deux personnes soient malades ?
5. L'un des soldats s'appelle Jean, et Jean sait qu'il est porteur de la maladie. Quelle doit être, de son point de vue, la probabilité qu'une autre personne au moins soit porteuse de la maladie ?
6. Le test étant positif, il est décidé que des tests individuels sont menés. Les $(n-1)$ premiers tests sont négatifs, le n-ème est positif : c'est celui de Jean. Quelle est la probabilité, en fonction de n, qu'une des personnes restantes au moins soit malade ?

Exercice 2.31 (Boules blanches et noires)
Un sac contient 8 boules blanches et 2 boules noires. On tire les boules les unes après les autres, sans remise, jusqu'à obtenir une boule blanche. On appelle X le nombre de tirages nécessaires pour obtenir cette boule blanche.

1. Quelles valeurs peut prendre la variable aléatoire X ?
2. Donner la loi de X.
3. Représenter sa fonction de répartition F.
4. Calculer $\mathbb{E}[X]$ et $\text{Var}(X)$.

Exercice 2.32 (Défaut de fabrication)
On admet que la probabilité de défaut pour un objet fabriqué à la machine est égale à 0,1. On considère un lot de 10 objets fabriqués par cette machine. Soit X le nombre d'objets défectueux parmi ceux-ci.

1. Comment s'appelle la loi suivie par X ?
2. Que valent $\mathbb{E}[X]$ et $\text{Var}(X)$?
3. Quelle est la probabilité que le lot comprenne au plus 1 objet défectueux ?
4. Retrouver ce résultat grâce à l'approximation par une loi de Poisson.

Exercice 2.33 (Recrutement)
Une entreprise veut recruter un cadre. Il y a en tout 10 candidats à se présenter pour ce poste. L'entreprise fait passer un test au premier candidat, qui est recruté s'il le réussit. Sinon, elle fait passer le même test au second candidat et ainsi de suite. On suppose que la probabilité qu'un candidat réussisse le test est égale à p, réel fixé compris entre 0 et 1. On appelle alors X la variable aléatoire à valeurs dans $\{1, \dots, 11\}$ qui vaut k si c'est le candidat numéro k qui est recruté, et 11 si aucun candidat n'est recruté.

1. Calculer en fonction de p les probabilités $\mathbb{P}(X=1), \mathbb{P}(X=2), \dots, P(X=10)$. Déterminer aussi $\mathbb{P}(X=11)$.
2. Comment doit-on choisir p pour que la probabilité de ne recruter personne soit inférieure à 1% ?

3. Pour $n \in \mathbb{N}$ fixé, on considère la fonction P définie par :

$$P(x) = 1 + x + \cdots + x^n = \sum_{j=0}^{n} x^j.$$

Exprimer sa dérivée $P'(x)$ sous la forme d'une somme de n termes.

4. Pour $x \neq 1$, écrire plus simplement $P(x)$ (penser à la somme des termes d'une suite géométrique). En déduire une autre expression de $P'(x)$, à savoir :

$$P'(x) = \frac{nx^{n+1} - (n+1)x^n + 1}{(1-x)^2}.$$

5. Déduire des questions précédentes que X a pour moyenne :

$$\mathbb{E}[X] = \frac{1 - (1-p)^{11}}{p}.$$

6. Supposons maintenant qu'il n'y ait pas seulement 10 candidats, mais un nombre infini, et que l'on procède de la même façon. Appelons Y le numéro du candidat retenu. Quelle est la loi classique suivie par Y ? Rappeler son espérance. La comparer à $\mathbb{E}[X]$ lorque $p = 1/2$.

Exercice 2.34 (Lancer de dé)
Un dé équilibré est lancé 10 fois de suite. Déterminer :

1. La probabilité d'au moins un 6 sur les 10 lancers.
2. Le nombre moyen de 6 sur les 10 lancers.
3. La moyenne de la somme des résultats obtenus lors des 10 lancers.
4. La probabilité d'obtenir exactement deux 6 lors des 5 premiers lancers sachant qu'il y en a eu 4 sur les 10 lancers.

Exercice 2.35 (Le dé dyadique)
On appelle "dé dyadique" un dé dont les faces sont numérotées respectivement 2, 4, 8, 16, 32, 64 (au lieu de 1, 2, 3, 4, 5, 6). On jette un dé dyadique équilibré et on appelle X le résultat obtenu.

1. Déterminer l'espérance de X.
2. Calculer l'écart-type de X.
3. Lorsque X_1 et X_2 sont deux variables indépendantes, que vaut $\text{Cov}(X_1, X_2)$?
4. On jette maintenant deux dés dyadiques équilibrés et on appelle Y le produit des résultats obtenus. Calculer l'espérance de Y.
5. (Bonus) Calculer $\mathbb{P}(Y < 20)$.

Exercice 2.36 (Répartition des tailles)
On suppose que dans une population, 1% des gens mesurent plus de 1m92. Supposons que vous tiriez au hasard (avec remise) 200 personnes dans cette population. Appelons X le nombre de personnes de plus de 1m92 dans votre échantillon.

1. Quelle est la loi de X ?
2. Par quelle loi peut-on l'approcher ?
3. Quelle est la probabilité que dans votre échantillon, au moins 3 personnes mesurent plus de 1m92 ?

Exercice 2.37 (Poisson en vrac)
On considère une variable X distribuée selon une loi de Poisson de paramètre $\lambda > 0$. Exprimer en fonction de λ :

1. $\mathbb{E}[3X+5]$.
2. $\mathrm{Var}(2X+1)$.
3. $\mathbb{E}\left[\frac{1}{X+1}\right]$.

Exercice 2.38 (Jeu d'argent)
Un jeu consiste à tirer, indépendamment et avec remise, des tickets d'une boîte. Il y a en tout 4 tickets, numérotés respectivement -2, -1, 0, 3. Votre "gain" X lors d'une partie correspond à la somme indiquée sur le ticket. Par exemple, si vous tirez le ticket numéroté -2, alors $X=-2$ et vous devez donner 2 €, tandis que si vous tirez le ticket 3, alors $X=3$ et vous gagnez 3 €.

1. Donner la loi de X. Calculer son espérance et sa variance.
2. Vous jouez 100 fois de suite à ce jeu et on note S votre gain après 100 parties. En notant X_1 le gain à la première partie, X_2 le gain à la deuxième partie, ..., X_{100} le gain à la centième partie, exprimer S en fonction des X_i.
3. En déduire l'espérance de S et sa variance.
4. Par quelle loi normale peut-on approcher S ? En déduire la probabilité que votre gain sur 100 parties dépasse 25 €.

Exercice 2.39 (Rubrique à brac)

1. Soit T une variable aléatoire suivant une loi géométrique de paramètre p, $0<p<1$. Rappeler la loi de T, son espérance et sa variance.
2. Vous demandez à des personnes choisies au hasard dans la rue leur mois de naissance jusqu'à en trouver une née en décembre. Quel est (approximativement) le nombre moyen de personnes que vous allez devoir interroger ?
3. On jette une pièce équilibrée et on appelle X le nombre de lancers nécessaires pour que Pile apparaisse. Quelle est la loi de X ?
4. Grâce aux moments de X, montrer que $\sum_{n=1}^{+\infty}\frac{n^2}{2^n}=6$.
5. Alice et Bob jouent au jeu suivant : Alice lance une pièce équilibrée jusqu'à ce que Pile apparaisse. Si Pile apparaît dès le premier lancer, Bob lui donne 4 € ; si Pile n'apparaît qu'au deuxième lancer, Bob lui donne 1 € ; si Pile n'apparaît qu'au troisième lancer, elle donne 4 € à Bob ; si Pile n'apparaît qu'au quatrième lancer, elle donne 11 € à Bob, etc. De façon générale, le "gain" d'Alice si Pile n'apparaît qu'au n-ème lancer est $5-n^2$. Notons G la variable aléatoire correspondant à ce gain.
   (a) Calculer la probabilité qu'Alice perde de l'argent lors d'une partie.
   (b) Calculer l'espérance de G.
   (c) Si vous deviez jouer une seule partie, préféreriez-vous être à la place d'Alice ou à la place de Bob ? Et si vous deviez en jouer 100 ?

Exercice 2.40 (Ascenseur pour l'échafaud)
Un ascenseur dessert les 10 étages d'un immeuble, 12 personnes le prennent au rez-de-chaussée et chacune choisit un des 10 étages au hasard.

1. Soit X_1 la variable aléatoire valant 1 si au moins une personne choisit le 1er étage, 0 sinon. Calculer $\mathbb{P}(X_1=1)$ et en déduire la moyenne de X_1.
2. De façon générale, soit X_i la variable aléatoire valant 1 si au moins une personne choisit l'étage i, 0 sinon. Exprimer le nombre d'étages auxquels l'ascenseur s'arrête en fonction des X_i. En déduire le nombre moyen d'étages auxquels l'ascenseur s'arrête.
3. (Bonus) Généralisation : montrer que pour t étages et n personnes, le nombre moyen d'étages desservis est $t(1-(1-\frac{1}{t})^n)$. Que devient cette quantité :

(a) lorsque t tend vers l'infini avec n fixé ? Interpréter.

(b) lorsque n tend vers l'infini avec t fixé ? Interpréter.

Exercice 2.41 (Systèmes de contrôle)
Deux systèmes de contrôle électrique opèrent indépendamment et sont sujets à un certain nombre de pannes par jour. Les probabilités p_n (respectivement q_n) régissant le nombre n de pannes par jour pour le système 1 (resp. 2) sont données dans les tableaux suivants :

Système 1

n	p_n
0	0.07
1	0.35
2	0.34
3	0.18
4	0.06

Système 2

n	q_n
0	0.10
1	0.20
2	0.50
3	0.17
4	0.03

1. Calculer les probabilités des événements suivants :
   (a) Le système 2 a au moins 2 pannes dans la journée.
   (b) Il se produit une seule panne dans la journée.
   (c) Le système 1 a le même nombre de pannes que le système 2.
2. Quel est le nombre moyen de pannes du système 1 par jour ? Comparer à celui du système 2.
3. Supposons que l'équipe de mécaniciens ne puisse réparer qu'un maximum de 6 pannes par jour. Dans quelle proportion du temps ne pourra-t-elle pas suffire à la tâche ?

Exercice 2.42 (Kramer contre Kramer)
On effectue des tirages sans remise dans une urne contenant initialement 3 boules rouges et 3 boules noires jusqu'à obtenir une boule noire. On appelle X le numéro du tirage de cette boule noire (ainsi $X = 1$ si la première boule tirée est noire).

1. Quelles valeurs peut prendre la variable aléatoire X ? Avec quelles probabilités ?
2. Représenter sa fonction de répartition F.
3. Calculer l'espérance et la variance de X.
4. On classe 3 hommes et 3 femmes selon leur note à un examen. On suppose toutes les notes différentes et tous les classements équiprobables. On appelle R le rang de la meilleure femme (par exemple $R = 2$ si le meilleur résultat a été obtenu par un homme et le suivant par une femme). Donner la loi de R.

Exercice 2.43 (Loterie)
Dans une loterie, un billet coûte 1 euro. Le nombre de billets émis est 90000, numérotés de 10000 à 99999, chaque billet comportant donc 5 chiffres. Un numéro gagnant est lui-même un nombre entre 10000 et 99999. Lorsque vous achetez un billet, vos gains possibles sont les suivants :

votre billet correspond au numéro gagnant	10000 euros
vos 4 derniers chiffres sont ceux du numéro gagnant	1000 euros
vos 3 derniers chiffres sont ceux du numéro gagnant	100 euros

1. Quelle est la probabilité d'avoir le numéro gagnant ?
2. Quelle est la probabilité de gagner 1000 euros ?
3. Quelle est la probabilité de gagner 100 euros ?

4. Déterminer votre bénéfice moyen lorsque vous achetez un billet.

Exercice 2.44 (Dé coloré)
Un joueur dispose d'un dé équilibré à six faces avec trois faces blanches, deux vertes et une rouge. Le joueur lance le dé et observe la couleur de la face supérieure :
– s'il observe une face rouge, il gagne 2 euros ;
– s'il observe une face verte, il perd 1 euro ;
– s'il observe une face blanche, il relance le dé et : pour une face rouge, il gagne 3 euros ; pour une face verte, il perd 1 euro ; pour une face blanche, le jeu est arrêté sans gain ni perte.
Soit X la variable aléatoire égale au gain (positif ou négatif) de ce joueur.

1. Quelles sont les valeurs prises par X ? Déterminer la loi de X.
2. Calculer l'espérance de X.
3. Calculer la variance et l'écart-type de X.
4. Le joueur effectue 144 parties successives de ce jeu. Donner une valeur approchée de la probabilité que son gain sur les 144 parties soit positif.

Exercice 2.45 (Beaujolais nouveau)
Le beaujolais nouveau est arrivé.

1. Un amateur éclairé, mais excessif, se déplace de réverbère en réverbère. Quand il se lance pour attraper le suivant, il a 80% de chances de ne pas tomber. Pour gagner le bistrot convoité, il faut en franchir 7. On notera X le nombre de réverbères atteints sans chute.
   (a) Quelles valeurs peut prendre la variable aléatoire X ?
   (b) Préciser sa loi.
2. Quand il sort du café, son étape suivante est l'arrêt de bus. Le nombre de chutes pour y parvenir, noté Y, suit une loi de Poisson $\mathcal{P}(4)$. Calculer la probabilité de faire au plus deux chutes.
3. Arrivé dans l'ascenseur, il appuie au hasard sur un des huits boutons. S'il atteint son étage ou s'il déclenche l'alarme, il sort de l'ascenceur, sinon il réappuie au hasard sur un des huits boutons. Soit Z le nombre de boutons pressés avant d'atteindre son étage ou de déclencher l'alarme.
   (a) Quelle est la loi de Z ?
   (b) Donner son espérance et sa variance.

2.7 Corrigés

Exercice 2.1 (Loi uniforme)

1. Notons F la fonction de répartition de X. On a :

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{2} & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

   Cette fonction de répartition est représentée sur la figure 2.10 à gauche.
2. La fonction de répartition est représentée sur la figure 2.10 au centre.
3. La fonction de répartition est représentée sur la figure 2.10 à droite.

4. La variable X est à valeurs dans $\{1,\dots,n\}$. Calculons la probabilité $\mathbb{P}(X=1)$: il y a a priori $n!$ ordres d'arrivées possibles pour les n patients. Si le plus grand arrive en premier, il reste $(n-1)$ patients qui arrivent ensuite dans un ordre aléatoire, donc $(n-1)!$ possibilités, donc au final $\mathbb{P}(X=1)=(n-1)!/n!=1/n$. On voit que le raisonnement que l'on vient de faire pour calculer $\mathbb{P}(X=1)$ est tout aussi valable pour $\mathbb{P}(X=2)$, $\mathbb{P}(X=3)$, etc. Ainsi X suit une loi uniforme sur $\{1,\dots,n\}$.

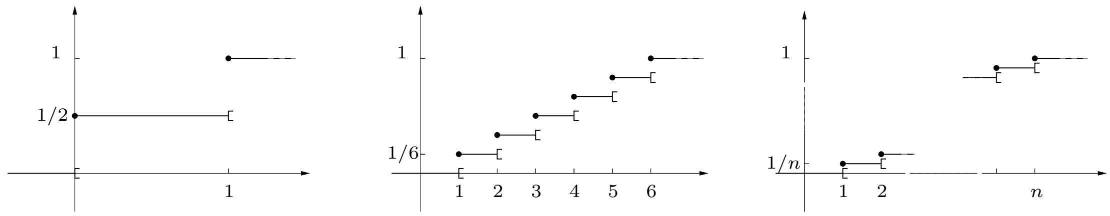


FIGURE 2.10 – Fonctions de répartition des lois uniformes $\mathcal{U}_{\{0,1\}}$, $\mathcal{U}_{\{1,\dots,6\}}$ et $\mathcal{U}_{\{1,\dots,n\}}$.

Exercice 2.2 (Loi de Bernoulli)

1. Notons F la fonction de répartition de X. On a :

$$F(x)=\begin{cases} 0 & \text{si } x<0 \\ \frac{1}{3} & \text{si } 0\le x<1 \\ 1 & \text{si } x\ge 1 \end{cases}$$

   Cette fonction de répartition est représentée sur la figure 2.11 à gauche.
2. La fonction de répartition est représentée sur la figure 2.11 à droite.
3. Si on ne tirait qu'une boule de l'urne, on aurait 1 chance sur $(N+B+1)$ d'obtenir la boule rouge. Si on en tirait 2, on multiplie nos chances par 2, etc. Si on en tire n, on multiplie nos chances par n, il y a donc $n/(N+B+1)$ chances que la boule rouge soit parmi les boules piochées. Ainsi $X\sim\mathcal{B}(n/(N+B+1))$.
4. Il faut calculer la probabilité qu'à l'instant $2n$, l'étudiant soit à nouveau en son point de départ. Pour ça, il faut qu'il ait fait autant de déplacements vers la droite que vers la gauche. Il y a en tout 2^{2n} suites possibles de $+1$ et -1 de longueur $2n$. Parmi celles-ci, seules nous intéressent celles où il y a exactement n fois $+1$. Or il y a $\binom{2n}{n}$ façons de placer ces $+1$ dans la suite de longueur $2n$ correspondant aux déplacements successifs de l'étudiant. La probabilité qu'à l'instant $2n$ l'étudiant soit à nouveau à son point de départ est donc $\binom{2n}{n}/2^{2n}$, donc $X\sim\mathcal{B}(\binom{2n}{n}/2^{2n})$.

Exercice 2.3 (Loi binomiale)

1. La variable X est à valeurs dans $\{0,1,\dots,n\}$. Fixons donc k entre 0 et n et cherchons $\mathbb{P}(X=k)$. Chaque n-uplet ayant la même probabilité 2^{-n} d'apparaître, il suffit de compter combien de ces n-uplets comptent exactement k "1" et $(n-k)$ "0" : ceci revient à choisir une combinaison de k indices parmi n, il y en a donc $\binom{n}{k}$. On en déduit la loi de X :

$$\forall k\in\{0,1,\dots,n\} \qquad \mathbb{P}(X=k)=\frac{\binom{n}{k}}{2^n}.$$

2. Pour vérifier que c'est bien une loi de probabilité, il suffit d'utiliser la formule du binôme (d'où le nom de cette loi) :

$$\sum_{k=0}^{n}p_k=\sum_{k=0}^{n}\binom{n}{k}p^k(1-p)^{n-k}=(p+(1-p))^n=1.$$

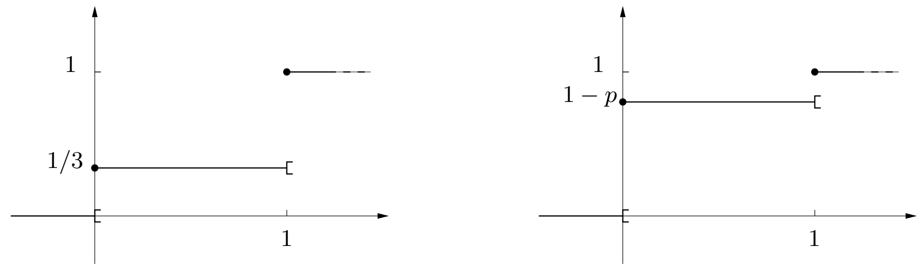


FIGURE 2.11 – Fonctions de répartition des lois de Bernoulli $\mathcal{B}(2/3)$ et $\mathcal{B}(p)$.

Considérons n variables $X_1, \dots, X_n$ qui suivent toutes la même loi de Bernoulli $\mathcal{B}(p)$ et indépendantes c'est-à-dire que :

$$\forall (i_1, \dots, i_n) \in \{0,1\}^n \qquad \mathbb{P}(X_1 = i_1, \dots, X_n = i_n) = \mathbb{P}(X_1 = i_1) \dots \mathbb{P}(X_n = i_n).$$

Considérons alors la variable aléatoire $X = X_1 + \cdots + X_n$. Elle est à valeurs dans $\{0, 1, \dots, n\}$. Fixons k entre 0 et n et cherchons $\mathbb{P}(X = k)$. Pour que $X = k$, il faut que exactement k des variables X_i prennent la valeur 1 et $(n-k)$ la valeur 0, ce qui s'écrit :

$$\mathbb{P}(X = k) = \sum_{(i_1,\dots,i_n):i_1+\cdots+i_n=k} P(X_1 = i_1, \dots, X_n = i_n),$$

donc par l'indépendance des X_i :

$$\mathbb{P}(X = k) = \sum_{(i_1,\dots,i_n):i_1+\cdots+i_n=k} \mathbb{P}(X_1 = i_1) \dots \mathbb{P}(X_n = i_n).$$

Il suffit alors de voir que lorsque $i_1 + \cdots + i_n = k$, la quantité $\mathbb{P}(X_1 = i_1) \dots \mathbb{P}(X_n = i_n)$ est toujours la même, égale à $p^k(1-p)^{n-k}$, donc :

$$\mathbb{P}(X = k) = \sum_{(i_1,\dots,i_n):i_1+\cdots+i_n=k} p^k(1-p)^{n-k} = p^k(1-p)^{n-k} \#\{(i_1, \dots, i_n) : i_1 + \cdots + i_n = k\}.$$

Et comme on l'a vu plus haut, le nombre de n-uplets comptant exactement k “1” et $(n-k)$ “0” est $\binom{n}{k}$, donc :

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \ \Rightarrow \ X \sim \mathcal{B}(n, p).$$

Dit brièvement, une loi binomiale $\mathcal{B}(n, p)$ peut être vue comme la somme de n lois de Bernoulli $\mathcal{B}(p)$ indépendantes.

3. Associons à chacun des n clients une variable X_i valant 1 s'il choisit la caisse numéro 1 et 0 sinon. Pour tout i entre 1 et n, X_i suit donc une loi de Bernoulli. Son paramètre est la probabilité que le client i choisisse la caisse numéro 1, c'est-à-dire $1/m$, donc $X_i \sim \mathcal{B}(1/m)$. Puisque les clients prennent leur décision indépendamment les uns des autres, les variables X_i sont indépendantes. La variable X qui nous intéresse, nombre total de clients à opter pour la caisse numéro 1, est alors tout simplement la somme des X_i. Par la question précédente on en conclut que $X \sim \mathcal{B}(n, 1/m)$.

Exercice 2.4 (Loi hypergéométrique)

1. Supposons d'entrée que $n \leq Np$ et $n \leq Nq$, ce qui est le cas courant lorsque N est bien plus grand que n, et p pas trop petit. La variable X est alors à valeurs dans $\{0, \dots, n\}$. Fixons

maintenant k entre 0 et n. Puisqu'un échantillon de taille $n < N$ est un tirage sans remise, il y a en tout $\binom{N}{n}$ échantillons possibles. Parmi ceux-ci, seuls ceux ayant k votants pour le candidat donné et $(n-k)$ pour l'autre nous intéressent. Or dans la population totale il y a Np votants pour le candidat donné et $N(1-p) = Nq$ votants pour l'autre. Il y a donc $\binom{Np}{k}$ choix possibles d'un côté et $\binom{Nq}{n-k}$ de l'autre. Tout ceci mis ensemble donne :

$$\forall k \in \{0, \ldots, n\} \qquad \mathbb{P}(X = k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}.$$

Remarque : Dans le cas général où on ne suppose plus $n \leq Np$ et $n \leq Nq$, X est à valeurs dans $\{\max(0, n - Nq), \min(n, Np)\}$ et la loi ci-dessus est encore valable.

2. Supposons à nouveau $n \leq Np$ et $n \leq Nq$. Pour vérifier que c'est bien une loi de probabilité, il suffit de montrer que :

$$\sum_{k=0}^{n} \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} = \frac{1}{\binom{N}{n}} \sum_{k=0}^{n} \binom{Np}{k}\binom{Nq}{n-k} = 1.$$

Or $\binom{N}{n}$ est le coefficient de X^n dans le polynôme $P(X) = (1+X)^N$, polynôme que l'on peut encore écrire :

$$P(X) = (1+X)^{Np}(1+X)^{Nq} = \left(\sum_{i=0}^{Np} \binom{Np}{i} X^i\right)\left(\sum_{j=0}^{Nq} \binom{Nq}{n-k} X^j\right).$$

Le coefficient de X^n issu de ce produit s'obtient en sommant les coefficients binomiaux sur tous les couples d'indices (i, j) tels que $i + j = n$, ce qui donne l'égalité :

$$\sum_{k=0}^{n} \binom{Np}{k}\binom{Nq}{n-k} = \binom{N}{n}.$$

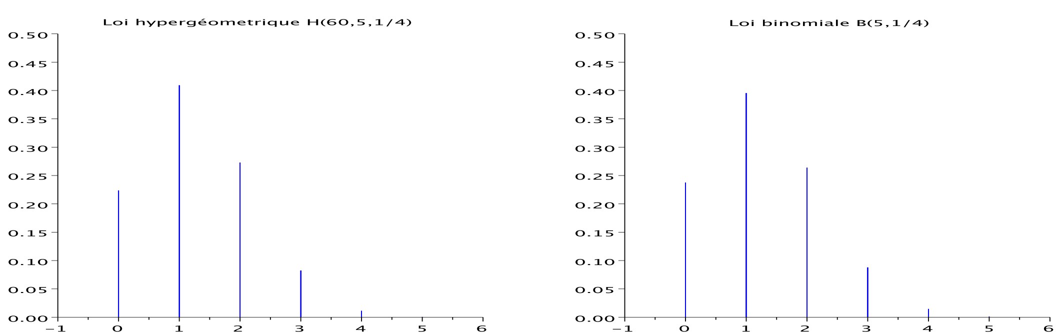


FIGURE 2.12 – Lois hypergéométrique $\mathcal{H}(60, 5, 1/4)$ et binomiale $\mathcal{B}(5, 1/4)$.

3. On suppose n très petit devant N, avec p ni tout proche de 0 ni tout proche de 1, ce qui est typiquement le cas des sondages politiques où N vaut plusieurs dizaines de millions, n environ un millier et p se situe entre 40% et 60%. Supposons que $X \sim \mathcal{H}(N, n, p)$. On a alors pour tout $k \in \{0, \ldots, n\}$:

$$\mathbb{P}(X = k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}} = \frac{\frac{(Np)!}{k!(Np-k)!} \times \frac{(Nq)!}{(n-k)!(Nq-(n-k))!}}{\frac{N!}{n!(N-n)!}},$$

ce qui s'écrit encore :

$$\mathbb{P}(X=k)=\binom{n}{k}\frac{(Np)(Np-1)\dots(Np-k+1)\times(Nq)(Nq-1)\dots(Nq-(n-k)+1)}{N(N-1)\dots(N-n+1)}.$$

Puisque $n \ll N$, on a aussi $k \ll Np$ et $(n-k) \ll Nq$, d'où les approximations suivantes :

$$\mathbb{P}(X=k)=\binom{n}{k}p^kq^{n-k}\frac{(1-\frac{1}{Np})\dots(1-\frac{k-1}{Np})\times(1-\frac{1}{Nq})\dots(1-\frac{n-k-1}{Nq})}{(1-\frac{1}{N})\dots(1-\frac{n-1}{N})}\approx\binom{n}{k}p^kq^{n-k},$$

et on arrive bien à l'approximation d'une loi hypergéométrique par une loi binomiale. Ceci est bien moral : on aurait exactement une loi binomiale si on faisait des tirages avec remise dans la population, or un sondage correspond à un tirage sans remise (on ne sonde pas deux fois la même personne). Cependant, lorsque l'échantillon est de taille négligeable par rapport à la population totale, un tirage avec remise se comporte comme un tirage sans remise puisqu'il y a très peu de chances qu'on pioche deux fois la même personne. En pratique, on effectue l'approximation $\mathcal{H}(N,n,p)\approx\mathcal{B}(n,p)$ dès que $n<N/10$. Ceci est illustré figure 2.12 lorsque $N=60$, $n=5$ et $p=1/4$.

Exercice 2.5 (Loi géométrique)

1. Pour tout $n\in\mathbb{N}$, notons E_n l'événement : "Pile apparaît après le n-ème lancer" et A l'événement : "Pile n'apparaît jamais". On a clairement :

$$A=\bigcap_{n=0}^{+\infty}E_n \ \Rightarrow\ \mathbb{P}(A)=\mathbb{P}\left(\bigcap_{n=0}^{+\infty}E_n\right),$$

et puisque $(E_n)_{n\geq 0}$ est une suite décroissante pour l'inclusion, on peut utiliser la continuité monotone décroissante : $\mathbb{P}(A)=\lim_{n\to+\infty}\mathbb{P}(E_n)$. Or pour que Pile apparaisse après le n-ème lancer, il faut n'obtenir que des Face lors des n premiers jets ce qui arrive avec probabilité 2^{-n}. Ainsi

$$\mathbb{P}(A)=\lim_{n\to+\infty}\frac{1}{2^n}=0,$$

et on dit que A est un événement négligeable, c'est pourquoi on l'exclut dans la suite. La variable aléatoire X est donc à valeurs dans $\mathbb{N}^*$. Pour tout $n\in\mathbb{N}^*$, la probabilité que X soit égale à n est la probabilité qu'on obtienne Face durant les $(n-1)$ premiers lancers et Pile au n-ème, ce qui arrive avec probabilité $\mathbb{P}(X=n)=2^{-n}$.

2. Par le même raisonnement que ci-dessus, la variable X est à valeurs dans $\mathbb{N}^*$, avec cette fois :

$$\forall n\in\mathbb{N}^* \qquad \mathbb{P}(X=n)=\frac{1}{6}\left(\frac{5}{6}\right)^{n-1}.$$

3. Pour vérifier que c'est bien une loi de probabilité, on remarque que (p_n) forme une suite géométrique de raison $(1-p)$ donc on utilise la formule "couteau suisse" des sommes géométriques, à savoir : "Somme = (1er terme écrit - 1er terme non écrit)/(1-la raison)", ce qui donne dans notre cas :

$$\sum_{n=1}^{+\infty}p_n=\frac{p_1-0}{1-(1-p)}=\frac{p}{p}=1.$$

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes et identiquement distribuées suivant la loi de Bernoulli $\mathcal{B}(p)$. Définissons à partir de celles-ci une nouvelle variable X comme le plus petit indice pour lequel X_n vaut 1 :

$$X=\min\{n\geq 1\ :\ X_n=1\}.$$

Comme on l'a vu ci-dessus, la probabilité qu'aucun des X_n ne soit égal à 1 est nulle donc on exclut ce cas et X est alors une variable aléatoire à valeurs dans $\mathbb{N}^*$. Il nous reste à trouver sa loi, or pour tout $n \in \mathbb{N}^*$:

$$\mathbb{P}(X = n) = \mathbb{P}(X_1 = 0, \dots, X_{n-1} = 0, X_n = 1),$$

probabilité qu'on évalue facilement via l'indépendance des X_i :

$$\mathbb{P}(X = n) = \mathbb{P}(X_1 = 0) \dots \mathbb{P}(X_{n-1} = 0)\mathbb{P}(X_n = 1) = (1-p)^{n-1}p,$$

et X suit donc une loi géométrique de paramètre p. La loi géométrique se rencontre typiquement dans les phénomènes d'attente jusqu'à l'apparition d'un événement.

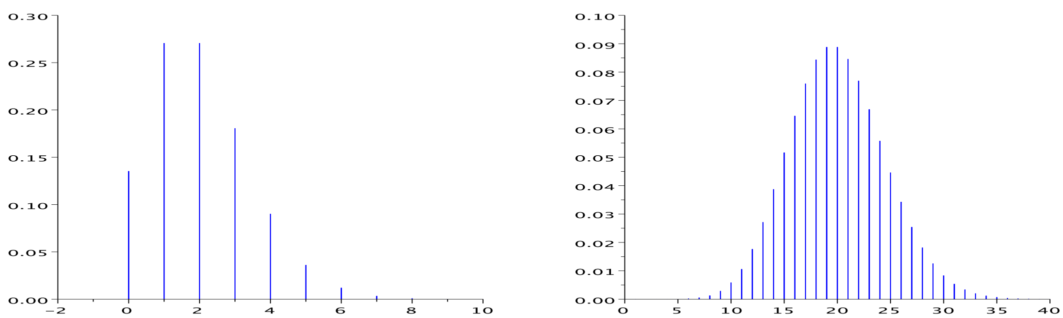


FIGURE 2.13 – Lois de Poisson $\mathcal{P}(2)$ et $\mathcal{P}(20)$.

Exercice 2.6 (Loi de Poisson)

1. Rappelons que pour tout réel x on a $\sum_{n=0}^{+\infty} \frac{x^n}{n!} = e^x$. C'est exactement ce qui s'applique ici :

$$\sum_{n=0}^{+\infty} e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

2. Les lois de Poisson pour $\lambda = 2$ et $\lambda = 20$ sont représentées figure 2.13.
3. Soit $\lambda > 0$ fixé. On lance n fois une pièce amenant Pile avec la probabilité $p_n = \lambda/n$. Soit X_n le nombre de fois où Pile apparaît durant ces n lancers.
   (a) La loi de X_n est la loi binomiale $\mathcal{B}(n, p_n) = \mathcal{B}(n, \lambda/n)$. En particulier :

$$\forall k \in \{0, \dots, n\} \qquad \mathbb{P}(X_n = k) = \binom{n}{k} p_n^k (1-p_n)^{n-k} = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}.$$

   (b) Pour k fixé entre 0 et n, on a :

$$\binom{n}{k} \left(\frac{\lambda}{n}\right)^k = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k = \frac{\lambda^k}{k!} \left(1 - \frac{1}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \xrightarrow[n \to \infty]{} \frac{\lambda^k}{k!}.$$

   D'autre part :

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} = \left(1 - \frac{\lambda}{n}\right)^{-k} e^{n \ln(1-\lambda/n)}.$$

Puisque $\ln(1-x) = -x + o(x)$, on en déduit que le second terme tend vers $e^{-\lambda}$, tandis que le premier tend clairement vers 1, donc :

$$\left(1 - \frac{\lambda}{n}\right)^{n-k} \xrightarrow[n\to\infty]{} e^{-\lambda}.$$

Le tout mis bout à bout donne :

$$\binom{n}{k} p_n^k (1-p_n)^{n-k} \xrightarrow[n\to\infty]{} e^{-\lambda}\frac{\lambda^k}{k!}.$$

Lorsque n devient grand, la loi binomiale "ressemble" à une loi de Poisson.

(c) Le résultat précédent est encore vrai si on suppose juste que $\lim_{n\to+\infty} np_n = \lambda$, puisque d'une part :

$$\binom{n}{k} p_n^k = \frac{(np_n)^k}{k!}\left(1 - \frac{1}{n}\right)\dots\left(1 - \frac{k-1}{n}\right) \xrightarrow[n\to\infty]{} \frac{\lambda^k}{k!},$$

et d'autre part :

$$(1-p_n)^{n-k} = (1-p_n)^{-k}\, e^{n\ln(1-p_n)} \xrightarrow[n\to\infty]{} e^{-\lambda}.$$

4. On considère donc une variable discrète X à valeurs dans $\{0, 1, \dots, 8\}$ et dont la loi $p = [p_0, p_1, \dots, p_8]$ s'obtient en divisant les données du tableau par 760 :

$$p = \left[\frac{268}{760}, \frac{266}{760}, \frac{152}{760}, \frac{53}{760}, \frac{13}{760}, \frac{7}{760}, 0, 0, \frac{1}{760},\right] \approx [0.353, 0.35, 0.2, 0.07, 0.017, 0.009, 0, 0, 0.001].$$

Bien sûr, supposer que X suit une loi de Poisson semble a priori un peu farfelu puisqu'une loi de Poisson prend ses valeurs dans $\mathbb{N}$ tout entier, tandis que X les prend dans $\{0, 1, \dots, 8\}$. Néanmoins on voit en figure 2.13 que les probabilités p_n d'une loi de Poisson décroissent extrêmement vite vers 0, donc l'approximation de X par une variable Y suivant une loi de Poisson $\mathcal{P}(\lambda)$, pour peu qu'elle colle bien sur les premiers termes, n'est pas déraisonnable. Il nous faut trouver une valeur pour le paramètre λ, or on sait que $\mathbb{P}(Y=0) = e^{-\lambda}$, donc on peut proposer par exemple :

$$\lambda = -\ln(\mathbb{P}(Y=0)) \approx -\ln p_0 = -\ln\left(\frac{268}{760}\right) \approx 1.04.$$

Voyons ce que donnent les 8 premiers termes d'une loi de Poisson $\mathcal{P}(1.04)$:

$$[\mathbb{P}(Y=0), \dots, \mathbb{P}(Y=8)] \approx [0.353, 0.368, 0.191, 0.066, 0.017, 0.004, 0.001, 0.000, 0.000].$$

On constate donc que l'approximation par une loi de Poisson est excellente, alors même que nous avons pris pour λ un estimateur on ne peut plus rudimentaire ! Les deux lois sont représentées figure 2.14.

Un estimateur plus sophistiqué est celui du maximum de vraisemblance. Pour une équipe et un match donnés, considérons le nombre de buts marqués comme la réalisation y d'une variable aléatoire Y distribuée suivant une loi de Poisson de paramètre λ. Puisqu'il y a 380 matchs, nous disposons de 760 réalisations $(y_1, \dots, y_{760})$ de 760 variables aléatoires $(Y_1, \dots, Y_{760})$ suivant la même loi de Poisson de paramètre λ. Le principe de l'estimation au maximum de vraisemblance (likelihood en anglais) est de chercher le paramètre $\lambda_{\max}$ qui

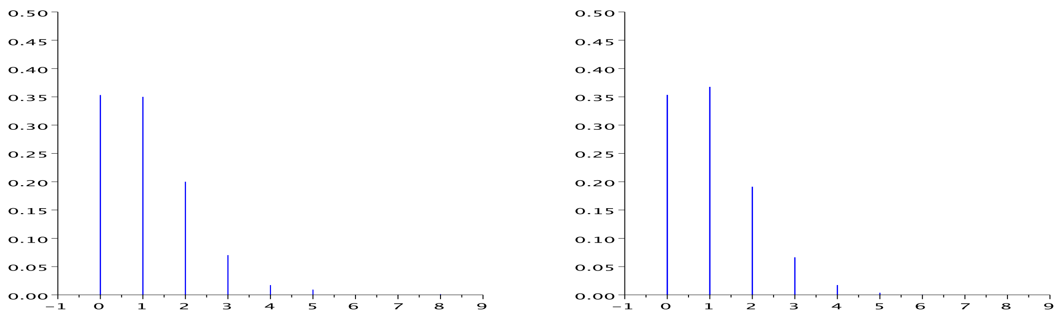


FIGURE 2.14 – Loi empirique du nombre de buts par équipe et par match (à gauche) et son approximation par une loi de Poisson $\mathcal{P}(1.04)$ (à droite).

rend cet ensemble d'observations le plus vraisemblable, c'est-à-dire tel que la fonction du paramètre λ définie par

$$L(\lambda) = \mathbb{P}(Y_1 = y_1, \ldots, Y_{760} = y_{760})$$

soit maximale pour $\lambda = \lambda_{\text{max}}$. Ce calcul se fait facilement si on suppose que les Y_i sont des variables indépendantes. En effet, cette vraisemblance devient alors

$$L(\lambda) = \mathbb{P}(Y_1 = y_1) \ldots \mathbb{P}(Y_{760} = y_{760}) = \prod_{i=1}^{760} e^{-\lambda} \frac{\lambda^{y_i}}{y_i!} = e^{-760\lambda} \frac{\lambda^{\sum_{i=1}^{760} y_i}}{\prod_{i=1}^{760} y_i!}$$

Puisque la fonction logarithmique est croissante, il est équivalent de chercher la valeur de λ pour laquelle le logarithme de $L(\lambda)$ est maximal. Nous passons donc à la log-vraisemblance

$$\ln L(\lambda) = -760\lambda + \ln(\lambda) \sum_{i=1}^{760} y_i - \sum_{i=1}^{760} \ln(y_i!)$$

Pour trouver en quel point cette log-vraisemblance atteint son maximum, il suffit de la dériver

$$(\ln L(\lambda))' = -760 + \frac{\sum_{i=1}^{760} y_i}{\lambda}$$

d'où l'on déduit que l'estimateur au maximum de vraisemblance de λ est

$$\lambda_{\text{max}} = \frac{\sum_{i=1}^{760} y_i}{760} = \frac{824}{760} \approx 1.08.$$

L'estimateur au maximum de vraisemblance est donc tout simplement la moyenne empirique des y_i, souvent notée $\bar{y}$. Ceci n'a rien de choquant intuitivement : le paramètre λ correspondant à la moyenne d'une loi de Poisson $\mathcal{P}(\lambda)$, il est naturel de l'estimer par la moyenne empirique de l'échantillon. Cette fois, les 8 premiers termes d'une loi de Poisson $\mathcal{P}(1.08)$ sont :

$$[\mathbb{P}(Y = 0), \ldots, \mathbb{P}(Y = 8)] \approx [0.340, 0.367, 0.198, 0.071, 0.019, 0.004, 0.001, 0.000, 0.000],$$

qui constitue également une très bonne approximation des données réelles.
Remarque. La principale critique vis-à-vis de l'estimation au maximum de vraisemblance est qu'elle suppose l'indépendance des variables Y_i, laquelle semble peu réaliste dans une compétition sportive.

Exercice 2.7 (Espérance d'une loi uniforme)

1. On a $X \in \{0,1\}$ avec $\mathbb{P}(X=0)=\mathbb{P}(X=1)=1/2$, donc
$$\mathbb{E}[X]=0\times\mathbb{P}(X=0)+1\times\mathbb{P}(X=1)=\frac{1}{2}.$$
2. Pour un dé équilibré, on obtient cette fois :
$$\mathbb{E}[X]=\frac{1}{6}(1+2+3+4+5+6)=\frac{7}{2}=3,5.$$
3. La somme des termes d'une suite arithmétique vaut de façon générale "((1er terme + dernier terme)$\times$ nb de termes)/2", ce qui donne ici :
$$1+2+\cdots+n=\frac{n(n+1)}{2}.$$
On en déduit que lorsque $U\sim\mathcal{U}_{\{1,\dots,n\}}$, son espérance vaut :
$$\mathbb{E}[U]=\frac{1}{n}(1+2+\cdots+n)=\frac{n+1}{2}.$$
4. Pour l'exercice sur le gardien de nuit à jeun, on a vu que le nombre N d'essais nécessaires pour ouvrir la porte suit une loi uniforme sur $\{1,\dots,10\}$, donc le nombre moyen d'essais nécessaires pour ouvrir la porte est $\mathbb{E}[N]=11/2=5,5$.
5. Soit m et n dans $\mathbb{N}^*$. Soit X une variable aléatoire à valeurs dans $\{1,2,\dots,mn\}$ et telle que $\forall i\in\{1,2,\dots,mn\}$, $\mathbb{P}(X=i)=1/m-1/n$.
   (a) Pour qu'on ait bien une loi de probabilité, il faut déjà que $1/m-1/n>0$, donc $m<n$. Par ailleurs la somme des probabilités doit valoir 1 :
$$\sum_{i=1}^{mn}\mathbb{P}(X=i)=1 \iff mn\left(\frac{1}{m}-\frac{1}{n}\right)=1 \iff m=n-1.$$
   Ainsi X suit une loi uniforme sur $\{1,2,\dots,n(n-1)\}$.
   (b) On a donc $\mathbb{E}[X]=\frac{n(n-1)+1}{2}$ et
$$\mathbb{E}[X]=\frac{7}{2} \iff \frac{n(n-1)+1}{2}=\frac{7}{2} \iff n=3.$$

Exercice 2.8 (Espérance d'une loi de Bernoulli)

1. La moyenne de X est $\mathbb{E}[X]=2/3$.
2. De façon générale, lorsque $X\sim\mathcal{B}(p)$, on a :
$$\mathbb{E}[X]=0\times(1-p)+1\times p=p.$$
3. Une roulette a 37 numéros : 18 rouges, 18 noirs et 1 vert (le zéro).
   (a) Si on mise 1€ sur rouge, on gagne 1€ avec probabilité 18/37 et on perd 1€ avec probabilité 19/37 donc notre bénéfice B a pour moyenne :
$$\mathbb{E}[B]=-1\times\frac{19}{37}+1\times\frac{18}{37}=-\frac{1}{37}.$$

(b) Si on mise 1€ sur le 13, notre gain moyen vaut :

$$\mathbb{E}[B] = -1 \times \frac{36}{37} + 35 \times \frac{1}{37} = -\frac{1}{37},$$

donc en moyenne cela revient au même que de miser sur une couleur (en moyenne, mais pas en variance...).

Exercice 2.9 (Espérance d'une loi binomiale)

1. Lorsque X suit une loi binomiale $\mathcal{B}(n,p)$, on peut calculer directement sa moyenne comme suit (en notant $q=(1-p)$ pour alléger les notations) :

$$\mathbb{E}[X] = \sum_{k=0}^{n} k\binom{n}{k}p^k q^{n-k} = \sum_{k=1}^{n} k\frac{n!}{k!(n-k)!}p^k q^{n-k} = np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}q^{(n-1)-(k-1)},$$

et le changement d'indice $j=(k-1)$ donne :

$$\mathbb{E}[X] = np\sum_{j=0}^{n-1}\binom{n-1}{j}p^j q^{(n-1)-j} = np(p+q)^{n-1} = np.$$

Ainsi lorsque $X \sim \mathcal{B}(n,p)$, son espérance vaut $\mathbb{E}[X]=np$.

2. Soit $X_1,\dots,X_n$ n variables aléatoires i.i.d. suivant la loi de Bernoulli $\mathcal{B}(p)$. Rappelons rapidement la loi de X. La variable aléatoire $X = X_1+\cdots+X_n$ peut prendre les valeurs $\{0,\dots,n\}$. Soit donc $k \in \{0,\dots,n\}$, on cherche la probabilité que X soit égale à k. Pour ce faire, il faut que k des n variables X_i prennent la valeur 1 et les $(n-k)$ autres la valeur 0 : il y a $\binom{n}{k}$ combinaisons de cette forme. Par indépendance des X_i, chaque événement de cette forme a alors la même probabilité d'apparition, à savoir $p^k q^{n-k}$. On en déduit que $\mathbb{P}(X=k)=\binom{n}{k}p^k q^{n-k}$, c'est-à-dire que X suit une loi binomiale $\mathcal{B}(n,p)$.
Par linéarité de l'espérance et sachant que $\mathbb{E}[X_1]=\cdots=\mathbb{E}[X_n]=p$, il vient donc :

$$\mathbb{E}[X] = \mathbb{E}[X_1+\cdots+X_n] = \mathbb{E}[X_1]+\cdots+\mathbb{E}[X_n] = np,$$

ce qui est une façon élémentaire de retrouver l'espérance d'une loi binomiale.

3. Soit $X_1,\dots,X_{n+m}$ variables aléatoires i.i.d. suivant la loi de Bernoulli $\mathcal{B}(p)$, alors $X = X_1+\cdots+X_n$, suit une loi binomiale $\mathcal{B}(n,p)$, $Y = X_{n+1}+\cdots+X_{n+m}$ suit une loi binomiale $\mathcal{B}(m,p)$ et $Z = X_1+\cdots+X_{n+m}$ suit une loi binomiale $\mathcal{B}(n+m,p)$. On en déduit que la somme de deux binomiales indépendantes de même paramètre p est encore une binomiale de paramètre p.

4. Y est à valeurs dans $\{1,\dots,n\}$. Soit donc $k\in\{1,\dots,n\}$, on cherche la probabilité que Y soit égale à k. De deux choses l'une : ou bien on a d'entrée $X=k$, ou bien on a $X=0$ puis $U=k$, ce qui traduit par :

$$\mathbb{P}(Y=k) = \mathbb{P}(\{X=k\}\cup(\{X=0\}\cap\{U=k\})) = \mathbb{P}(X=k)+\mathbb{P}(\{X=0\}\cap\{U=k\}),$$

et on utilise maintenant l'indépendance des variables X et U :

$$\mathbb{P}(Y=k) = \mathbb{P}(X=k)+\mathbb{P}(X=0)\mathbb{P}(U=k) = \frac{\binom{n}{k}}{2^n}+\frac{1}{2^n}\times\frac{1}{n} = \frac{1}{2^n}\left(\binom{n}{k}+\frac{1}{n}\right).$$

On en déduit l'espérance de Y :

$$\mathbb{E}[Y] = \sum_{k=1}^{n} k\mathbb{P}(Y=k) = \sum_{k=1}^{n}\frac{k\binom{n}{k}}{2^n}+\frac{1}{2^n}\sum_{k=1}^{n}\frac{k}{n}.$$

A peu de choses près, on reconnaît dans le premier terme l'espérance d'une binomiale et dans le second l'espérance d'une uniforme, d'où :

$$\mathbb{E}[Y] = \frac{n}{2} + \frac{1}{2^n} \times \frac{n+1}{2} = \frac{n}{2} + \frac{n+1}{2^{n+1}}.$$

Exercice 2.10 (Espérance d'une loi hypergéométrique)

1. D'après l'exercice 2.4, X suit une loi hypergéométrique $\mathcal{H}(N, n, p)$.
2. On numérote de 1 à Np les Np votants pour le candidat A. Pour tout $k \in \{1, \dots, Np\}$, on appelle X_k la variable aléatoire qui vaut 1 si l'individu k fait partie de l'échantillon, 0 sinon.
   (a) On a donc $X = X_1 + \cdots + X_{Np}$.
   (b) La variable X_1, ne pouvant prendre que les valeurs 0 et 1, est une variable de Bernoulli. La probabilité qu'elle vaille 1 est la probabilité que le votant 1 fasse partie de l'échantillon formé par un tirage sans remise de taille n parmi N, c'est donc n/N. Ainsi $\mathbb{E}[X_1] = n/N$ et idem pour les autres X_i puisqu'ils ont tous la même loi, d'où par linéarité de l'espérance :
   $$\mathbb{E}[X] = \mathbb{E}[X_1 + \cdots + X_{Np}] = Np\mathbb{E}[X_1] = np.$$
   On retrouve exactement la moyenne d'une variable aléatoire distribuée suivant une loi binomiale $\mathcal{B}(n, p)$.
3. D'après ce qui précède, $X \sim \mathcal{H}(32, 5, 1/8)$ puisque la proportion de rois est égale à $4/32 = 1/8$. Le nombre moyen de rois obtenus est donc $\mathbb{E}[X] = 5/8$.

Exercice 2.11 (Espérance d'une loi géométrique)

1. D'après l'exercice 2.5, X suit une loi géométrique $\mathcal{G}(1/6)$. Son espérance vaut :
   $$\mathbb{E}[X] = \sum_{n=1}^{+\infty} n\frac{1}{6}\left(\frac{5}{6}\right)^{n-1} = \frac{1}{6}\sum_{n=1}^{+\infty} n\left(\frac{5}{6}\right)^{n-1} = \frac{1}{6} \times \frac{1}{(1-5/6)^2} = 6.$$
2. Généralisation : si $X \sim \mathcal{G}(p)$, le même calcul donne $\mathbb{E}[X] = 1/p$.
3. Soit X le nombre d'essais nécessaires pour ouvrir la porte. On a vu dans l'exercice 2.5 que $X \sim \mathcal{G}(1/10)$, donc le nombre moyen d'essais est $\mathbb{E}[X] = 10$.
4. En notant X la variable aléatoire égale à $N - 2$, il apparaît que X est à valeurs dans $\mathbb{N}^*$, avec :
   $$\forall n \in \mathbb{N}^* \qquad \mathbb{P}(X = n) = \mathbb{P}(N = n + 2) = 2^{-n} = \frac{1}{2}\left(\frac{1}{2}\right)^{n-1},$$
   ce qui est exactement dire que $X \sim \mathcal{G}(1/2)$. Ainsi d'une part $\mathbb{E}[X] = 2$, et d'autre part $\mathbb{E}[X] = \mathbb{E}[N - 2] = \mathbb{E}[N] - 2$, d'où $\mathbb{E}[N] = 4$.
   Un polygone à N côtés compte $D = N(N - 3)/2$ diagonales : N choix pour une extrémité, $(N - 3)$ pour l'autre, et on divise par 2 afin de ne pas compter chaque diagonale deux fois. Ainsi :
   $$\mathbb{E}[D] = \frac{1}{2}\mathbb{E}[N(N-3)] = \frac{1}{2}\mathbb{E}[(X+2)(X-1)] = \frac{1}{2}(\mathbb{E}[X(X-1)] + 2\mathbb{E}[X] - 2).$$
   Un petit calcul s'impose :
   $$\mathbb{E}[X(X-1)] = \sum_{n=1}^{+\infty} \frac{n(n-1)}{2}\left(\frac{1}{2}\right)^{n-1} = \frac{1}{4}\sum_{n=1}^{+\infty} n(n-1)\left(\frac{1}{2}\right)^{n-2},$$
   et le rappel en début d'exercice permet d'aboutir à $\mathbb{E}[X(X-1)] = 4$. Au total, le nombre moyen de diagonales est donc $\mathbb{E}[D] = 3$.

Exercice 2.12 (Espérance d'une loi de Poisson)

1. Le calcul détaillé vu en cours pour montrer que $\mathbb{E}[X] = 1$ lorsque $X \sim \mathcal{P}(1)$ se généralise au cas où $X \sim \mathcal{P}(\lambda)$.
2. La moyenne de $Y = e^{-X}$ peut se calculer par le théorème de transfert :
$$\mathbb{E}[Y] = \sum_{n=0}^{+\infty} e^{-n} e^{-\lambda} \frac{\lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{(\lambda e^{-1})^n}{n!},$$
où l'on retrouve le développement en série de l'exponentielle :
$$\mathbb{E}[Y] = e^{-\lambda} e^{\lambda e^{-1}} = e^{\lambda(e^{-1}-1)}.$$
3. (a) La variable X est à valeurs dans $\mathbb{N}^*$. Pour que $X = 1$, il faut que le premier soit réussi, ce qui est certain, et que le second soit raté, ce qui arrive avec probabilité $1/2$, donc $p_1 = \mathbb{P}(X = 1) = 1/2 = 1/(1+1)!$. De façon générale, pour que $X = n$, il faut que les n premiers sauts soient réussis, ce qui arrive avec probabilité $1 \times 1/2 \times \cdots \times 1/n$, et que le $(n+1)$-ème soit un échec, ce qui arrive avec probabilité $1 - r_{n+1} = n/(n+1)$. Ainsi :
$$p_n = \mathbb{P}(X = n) = 1 \times \frac{1}{2} \times \cdots \times \frac{1}{n} \times \frac{n}{n+1} = \frac{n}{(n+1)!}.$$
On vérifie (pour la forme) que c'est bien une loi de probabilité grâce à l'astuce $n = (n+1) - 1$:
$$\sum_{n=1}^{+\infty} p_n = \sum_{n=1}^{+\infty} \left(\frac{1}{n!} - \frac{1}{(n+1)!} \right) = 1,$$
puisque les termes se télescopent.

   (b) Le théorème de transfert donne :
$$\mathbb{E}[X+1] = \sum_{n=1}^{+\infty} (n+1) \frac{n}{(n+1)!} = \sum_{n=1}^{+\infty} \frac{1}{(n-1)!} = e.$$
Par suite $\mathbb{E}[X] + 1 = e$, donc $\mathbb{E}[X] = e - 1$.

Exercice 2.13 (Espérance d'une loi arithmétique)

1. Pour que X soit effectivement une variable aléatoire, il faut que les p_k somment à 1, or :
$$1 = \sum_{k=0}^{n} p_k = \alpha \sum_{k=0}^{n} k = \alpha \frac{n(n+1)}{2} \implies \alpha = \frac{2}{n(n+1)}.$$
2. Avec en tête la relation $\sum_{k=0}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$, on peut y aller à fond de cinquième :
$$\mathbb{E}[X] = \sum_{k=0}^{n} k p_k = \frac{2}{n(n+1)} \sum_{k=0}^{n} k^2 = \frac{2n+1}{3}.$$

Exercice 2.14 (Deux dés : somme et différence)

1. U_1 suit une loi uniforme sur $\{1, \ldots, 6\}$, son espérance vaut $7/2$ et sa variance $35/12$.
2. On appelle $X = (U_1 + U_2)$ la somme et $Y = (U_1 - U_2)$ la différence des deux résultats. U_2 a la même loi que U_1 donc : $\mathbb{E}[X] = \mathbb{E}[U_1] + \mathbb{E}[U_2] = 7$ et $\mathbb{E}[Y] = \mathbb{E}[U_1] - \mathbb{E}[U_2] = 0$. Pour le produit, il suffit d'écrire :
$$\mathbb{E}[XY] = \mathbb{E}[(U_1 + U_2)(U_1 - U_2)] = \mathbb{E}[U_1^2] - \mathbb{E}[U_2^2] = 0,$$
toujours en raison du fait que U_1 et U_2 ont même loi.

3. On a ainsi $\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = 0$, donc X et Y sont décorrélées. Par contre on vérifie aisément que :

$$\mathbb{P}(X = 2, Y = 5) = 0 \neq \mathbb{P}(X = 2)\mathbb{P}(Y = 5) = \frac{1}{36} \times \frac{1}{36},$$

ce qui prouve que X et Y ne sont pas indépendantes. Cette non-indépendance était intuitivement clair : une fois connue X, la variable Y ne peut plus varier qu'entre $2 - X$ et $X - 2$, ce qui montre bien que la valeur prise par X a une incidence sur la valeur prise par Y (et vice versa).

Exercice 2.15 (Deux dés : min et max)

1. La variable X est à valeurs dans $\{1, \dots, 6\}$ et en notant $p_i = \mathbb{P}(X = i)$, quelques calculs donnent
$$[p_1, p_2, p_3, p_4, p_5, p_6] = \frac{1}{36}[11, 9, 7, 5, 3, 1].$$
Il s'ensuit pour l'espérance : $\mathbb{E}[X] = 91/36 \approx 2,5$.
2. Puisque $X + Y = U_1 + U_2$, il vient par linéarité de l'e :
$$\mathbb{E}[Y] = \mathbb{E}[U_1] + \mathbb{E}[U_2] - \mathbb{E}[X] = \frac{161}{36} \approx 4,5.$$
3. De même $XY = U_1U_2$, avec U_1 et U_2 indépendantes, donc :
$$\mathbb{E}[XY] = \mathbb{E}[U_1U_2] = \mathbb{E}[U_1]\mathbb{E}[U_2] = \frac{49}{4}.$$
Ceci donne pour la covariance : $\text{Cov}(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = \frac{1225}{1296} \approx 0.94$.

Exercice 2.16 (Memento (absence de mémoire))

1. On a :
$$\mathbb{P}(X > n) = \sum_{k=n+1}^{+\infty} \mathbb{P}(X = k) = \sum_{k=n+1}^{+\infty} p(1-p)^{k-1} = p \sum_{k=n+1}^{+\infty} (1-p)^{k-1},$$
où l'on reconnaît une somme géométrique, donc :
$$\mathbb{P}(X > n) = p\frac{(1-p)^n}{1-(1-p)} = (1-p)^n.$$
On en déduit la fonction de répartition au point n :
$$F(n) = \mathbb{P}(X \leq n) = 1 - \mathbb{P}(X > n) = 1 - (1-p)^n.$$
2. Par définition de la probabilité conditionnelle, on a alors $\forall(m,n) \in \mathbb{N} \times \mathbb{N}$:
$$\mathbb{P}(X > n + m | X > m) = \frac{\mathbb{P}(\{X > n + m\} \cap \{X > m\})}{\mathbb{P}(X > m)} = \frac{\mathbb{P}(X > n + m)}{\mathbb{P}(X > m)},$$
puisque l'événement $\{X > n + m\}$ implique l'événement $\{X > m\}$. Grâce à la question précédente, on a donc :
$$\mathbb{P}(X > n + m | X > m) = \frac{(1-p)^{n+m}}{(1-p)^m} = (1-p)^n = \mathbb{P}(X > n).$$

Exercice 2.17 (Minimum de lois géométriques)

1. Tout comme les variables X_1 et X_2, la variable aléatoire X est à valeurs dans $\mathbb{N}^*$.
2. Soit $n \in \mathbb{N}$ fixé, on peut écrire grâce à l'indépendance de X_1 et X_2 :
$$\mathbb{P}(X > n) = \mathbb{P}(X_1 > n, X_2 > n) = \mathbb{P}(X_1 > n)\mathbb{P}(X_2 > n),$$
or ces quantités ont été vues dans l'exercice précédent :
$$\mathbb{P}(X > n) = (1-p_1)^n(1-p_2)^n = ((1-p_1)(1-p_2))^n = (1-(1-(1-p_1)(1-p_2)))^n.$$
Puisqu'il est clair que $0 < 1-(1-p_1)(1-p_2) < 1$, on en déduit que $X = \min(X_1, X_2)$ suit elle-même une loi géométrique, et plus précisément $X \sim \mathcal{G}(1-(1-p_1)(1-p_2))$.
3. Application : On s'intéresse donc à la variable $X = \min(X_1, X_2)$, où $X_1 \sim \mathcal{G}(1/6)$, $X_2 \sim \mathcal{G}(1/6)$, avec X_1 et X_2 indépendantes. On vient de voir que $X \sim \mathcal{G}(11/36)$, donc le nombre moyen de “ doubles lancers ” nécessaires est $\mathbb{E}[X] = 36/11$.
4. Généralisation : soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois géométriques de paramètres respectifs $p_1, \dots, p_n$, avec $p_i \in]0,1[$ pour tout $i \in \{1, \dots, n\}$. Le même calcul que ci-dessus permet de montrer que :
$$X = \min(X_1, \dots, X_n) \sim \mathcal{G}(1-(1-p_1)\dots(1-p_n)).$$

Exercice 2.18 (Un problème de natalité)

1. X est à valeurs dans $\mathbb{N}^*$ et on vérifie facilement que X suit une loi géométrique $\mathcal{G}(1/2)$.
2. Puisqu'il y a exactement un garçon parmi les X enfants, $P = 1/X$.
3. Le calcul d'espérance s'écrit par le théorème de transfert :
$$\mathbb{E}[P] = \mathbb{E}[1/X] = \sum_{n=1}^{+\infty} \frac{1}{n} \times \frac{1}{2^n}.$$
Rappelons le développement en série entière :
$$\forall x \in [-1, +1[\qquad \ln(1-x) = -\sum_{n=1}^{+\infty} \frac{x^n}{n},$$
qu'il suffit d'appliquer ici en $x = 1/2$ pour obtenir $\mathbb{E}[P] = \ln 2 \approx 0.69$. Ainsi, à la génération suivante, il y a bien plus de garçons que de filles, ce qui n'est pas étonnant (cf. figure 2.15) mais risque de poser très vite des problèmes de renouvellement de population.

Exercice 2.19 (Tirages de cartes)

Nous avons affaire à un cas d'école : $X \sim \mathcal{B}(5, 1/8)$, d'espérance $5/8$ et de variance $35/64$.

Exercice 2.20 (Codage redondant)

Le nombre N d'erreurs lors d'une transmission suit une loi binomiale $\mathcal{B}(5; 0,2)$. Pour que le message soit mal interprété après un décodage à la majorité, il faut et il suffit qu'il y ait eu au moins 3 erreurs lors de la transmission. La probabilité p de mauvaise interprétation vaut donc :
$$p = \mathbb{P}(N=3) + \mathbb{P}(N=4) + \mathbb{P}(N=5) = \sum_{k=3}^{5} \binom{5}{k} (0,2)^k (0,8)^{5-k} \approx 0,058.$$

Le codage redondant permet donc de passer de 20% d'erreurs de transmission à seulement 5,8%.

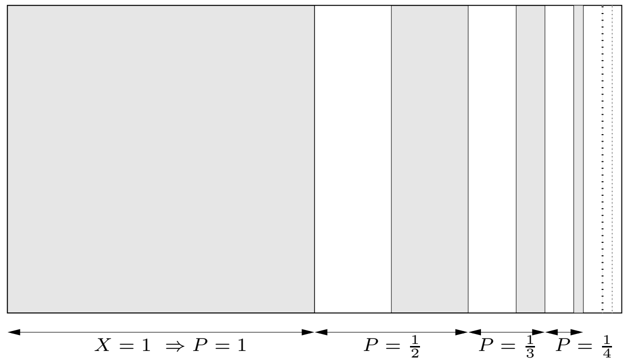


FIGURE 2.15 – Proportion de garçons (surface grisée).

Exercice 2.21 (Inégalité de Tchebychev)

1. La variable S est la somme de 3600 variables indépendantes de Bernoulli de paramètre $1/6$. On en déduit que S suit une loi binomiale $\mathcal{B}(3600, 1/6)$, de moyenne $\mathbb{E}[S] = 3600 \times 1/6 = 600$ et de variance $\mathrm{Var}(S) = 3600 \times 1/6 \times 5/6 = 500$.
2. La probabilité que S soit compris strictement entre 480 et 720 est donc :

$$\mathbb{P}(480 < S < 720) = \sum_{n=481}^{719} \mathbb{P}(S = n) = \sum_{n=481}^{719} \binom{3600}{n} \left(\frac{1}{6}\right)^n \left(\frac{5}{6}\right)^{3600-n} \tag{2.2}$$

L'inégalité de Tchebychev permet de minorer cette probabilité comme suit :

$$\mathbb{P}(480 < S < 720) = \mathbb{P}(-120 < S - 600 < 120) = \mathbb{P}(-120 < S - \mathbb{E}[S] < 120),$$

ce qui s'écrit encore :

$$\mathbb{P}(480 < S < 720) = 1 - \mathbb{P}(|S - \mathbb{E}[S]| \geq 120) \geq 1 - \frac{\mathrm{Var}(S)}{120^2},$$

ce qui donne :

$$\mathbb{P}(480 < S < 720) \geq 0.965.$$

Il y a donc au moins 96,5% de chances que le nombre de 1 obtenus sur les 3600 lancers soit entre 480 et 720. Le calcul sur machine de la somme (2.2) donne en fait $\mathbb{P}(480 < S < 720) \approx 1 - 10^{-7}$. Ainsi la borne donnée par l'inégalité de Tchebychev est très pessimiste : lorsqu'on lance 3600 fois un dé, il y a environ une chance sur 10 millions que le nombre de 1 ne soit pas compris entre 480 et 720.

Exercice 2.22 (Surbooking)

Puisqu'il y a 94 places vendues et que pour chacune d'entre elles, il y a 5% de chances que le passager ne soit pas là pour l'embarquement, le nombre S de personnes absentes à l'embarquement suit une loi binomiale $\mathcal{B}(94, 0.05)$. La probabilité qu'il y ait trop de monde à l'embarquement est donc :

$$p = \mathbb{P}(S \leq 3) = \mathbb{P}(S = 0) + \cdots + \mathbb{P}(S = 3) = \binom{94}{0} \left(\frac{5}{100}\right)^0 \left(\frac{95}{100}\right)^{94} + \cdots + \binom{94}{3} \left(\frac{5}{100}\right)^3 \left(\frac{95}{100}\right)^{91}.$$

Via l'approximation d'une loi binomiale $\mathcal{B}(94, 0.05)$ par une loi de Poisson $\mathcal{P}(94 \times 0.05) = \mathcal{P}(4, 7)$, ceci est à peu près égal à :

$$p \approx e^{-4,7} \frac{4,7^0}{0!} + \cdots + e^{-4,7} \frac{4,7^3}{3!} \approx 0,310.$$

Il y a donc 31,0% de risques de problème avec cette technique de surbooking.
Remarque : si on n'effectue pas l'approximation poissonienne, on obtient en fait $p = 30,3\%$.

Exercice 2.23 (Mode(s) d'une loi de Poisson)

1. Pour tout $n \in \mathbb{N}$, on obtient $r_n = \frac{p_{n+1}}{p_n} = \frac{\lambda}{n+1}$.
2. Pour connaître les variations de la suite (p_n), il suffit donc de comparer r_n à 1 (on rappelle que tous les p_n sont strictement positifs). On distingue donc deux cas :
   (a) $\lambda \in \mathbb{N}^*$: pour $n = \lambda - 1$, on a $r_n = 1$, donc il y a deux modes :
   $$\max_{n\in\mathbb{N}} p_n = p_{\lambda-1} = p_\lambda = e^{-\lambda}\frac{\lambda^\lambda}{\lambda!}.$$
   (b) $\lambda \notin \mathbb{N}^*$: on a cette fois un seul mode, atteint pour n égal à la partie entière de λ :
   $$\max_{n\in\mathbb{N}} p_n = p_{\lfloor\lambda\rfloor} = e^{-\lambda}\frac{\lambda^{\lfloor\lambda\rfloor}}{\lfloor\lambda\rfloor!}.$$

Exercice 2.24 (Parité d'une loi de Poisson)

1. On a d'une part :
   $$S_1 + S_2 = \sum_{k=0}^{+\infty}\frac{\lambda^{2k}}{(2k)!} + \sum_{k=0}^{+\infty}\frac{\lambda^{2k+1}}{(2k+1)!} = \sum_{n=0}^{+\infty}\frac{\lambda^n}{n!} = e^\lambda,$$
   et d'autre part :
   $$S_1 - S_2 = \sum_{k=0}^{+\infty}\frac{\lambda^{2k}}{(2k)!} - \sum_{k=0}^{+\infty}\frac{\lambda^{2k+1}}{(2k+1)!} = \sum_{n=0}^{+\infty}\frac{(-\lambda)^n}{n!} = e^{-\lambda}.$$
   On en déduit que
   $$S_1 = \frac{e^\lambda + e^{-\lambda}}{2} \qquad \& \qquad S_2 = \frac{e^\lambda - e^{-\lambda}}{2}.$$
2. Application : Notons S le nombre de prospectus distribués dans la journée. Par hypothèse $S \sim \mathcal{P}(\lambda)$. La probabilité que ce nombre soit pair vaut :
   $$p = \sum_{k=0}^{+\infty}\mathbb{P}(S = 2k) = \sum_{k=0}^{+\infty}e^{-\lambda}\frac{\lambda^{2k}}{(2k)!} = e^{-\lambda}S_1 = \frac{1+e^{-2\lambda}}{2}.$$
   Par conséquent la probabilité que ce nombre soit impair vaut :
   $$q = 1 - p = \frac{1-e^{-2\lambda}}{2}.$$
   Puisque $e^{-2\lambda} > 0$, on a $p > q$. En moyenne, c'est donc Pierre qui gagne.
3. Lorsque X décrit l'ensemble des entiers naturels pairs, $X/2$ décrit l'ensemble de tous les entiers naturels, donc Y est à valeurs dans $\mathbb{N}$. Soit $n \in \mathbb{N}^*$, la probabilité que Y soit égale à n correspond à la probabilité que X soit égale à $2n$:
   $$\forall n \in \mathbb{N}^*, \qquad \mathbb{P}(Y = n) = \mathbb{P}(X = 2n) = e^{-\lambda}\frac{\lambda^{2n}}{(2n)!}.$$
   On en déduit que :
   $$\mathbb{P}(Y = 0) = 1 - \sum_{n=1}^{+\infty}e^{-\lambda}\frac{\lambda^{2n}}{(2n)!} = 1 - e^{-\lambda}(S_1 - 1) = 1 - \frac{(1+e^{-\lambda})^2}{2}.$$

L'espérance de Y vaut :

$$\mathbb{E}[Y]=\sum_{n=0}^{+\infty}n\mathbb{P}(Y=n)=\sum_{n=1}^{+\infty}n\mathbb{P}(Y=n)=\frac{e^{-\lambda}}{2}\sum_{n=1}^{+\infty}(2n)\frac{\lambda^{2n}}{(2n)!},$$

qui s'écrit aussi :

$$\mathbb{E}[Y]=\frac{\lambda e^{-\lambda}}{2}\sum_{n=1}^{+\infty}\frac{\lambda^{2n-1}}{(2n-1)!}=\frac{\lambda e^{-\lambda}}{2}S_2=\lambda\frac{1-e^{-2\lambda}}{4}.$$

Pour calculer $\mathrm{Var}(Y)$, on utilise la décomposition :

$$\mathrm{Var}(Y)=\mathbb{E}[Y^2]-\mathbb{E}[Y]^2=\mathbb{E}[Y(Y-1/2)]+\frac{\mathbb{E}[Y]}{2}-\mathbb{E}[Y]^2,$$

où le seul terme à calculer est le premier :

$$\mathbb{E}[Y(Y-1/2)]=\sum_{n=0}^{+\infty}n(n-1/2)\mathbb{P}(Y=n)=\frac{1}{4}\sum_{n=1}^{+\infty}2n(2n-1)\mathbb{P}(Y=n),$$

ce qui donne après simplifications :

$$\mathbb{E}[Y(Y-1/2)]=\frac{\lambda^2 e^{-\lambda}}{4}\sum_{n=1}^{+\infty}\frac{\lambda^{2n-2}}{(2n-2)!}=\frac{\lambda^2 e^{-\lambda}}{4}S_1=\lambda^2\frac{1+e^{-2\lambda}}{8}.$$

Au final on obtient donc :

$$\mathrm{Var}(Y)=\lambda^2\frac{1+e^{-2\lambda}}{8}+\lambda\frac{1-e^{-2\lambda}}{8}-\lambda^2\frac{(1-e^{-2\lambda})^2}{16}.$$

Exercice 2.25 (Espérance d'une variable à valeurs entières)

1. En notant p_i la probabilité que X prenne la valeur i, l'espérance de X vaut

$$\mathbb{E}[X]=1\times p_1+2\times p_2+\cdots+n\times p_n=\sum_{i=1}^{n}i\times p_i$$

2. La formule précédente peut se réécrire sous la forme

$$\mathbb{E}[X]=(p_1+p_2+\cdots+p_n)+(p_2+\cdots+p_n)+(p_3+\cdots+p_n)+\cdots+(p_{n-1}+p_n)+p_n$$

Et en remarquant que de façon générale, pour tout i entre 1 et n :

$$p_i+\cdots+p_n=\mathbb{P}(X=i)+\cdots+\mathbb{P}(X=n)=\mathbb{P}(X\geq i)$$

on arrive bien à la formule dite de sommation des queues (*tail sum formula*) :

$$\mathbb{E}[X]=\mathbb{P}(X\geq 1)+\mathbb{P}(X\geq 2)+\cdots+\mathbb{P}(X\geq n)=\sum_{i=1}^{n}\mathbb{P}(X\geq i).$$

3. On jette 4 dés équilibrés simultanément et on appelle X le minimum obtenu sur ces 4 lancers.
   (a) La variable aléatoire X peut prendre les valeurs 1, 2, 3, 4, 5, 6.

(b) Pour que X soit supérieure ou égale à i, il faut que les 4 dés prennent une valeur supérieure ou égale à i. Notons $U_1, \dots, U_4$ les 4 variables correspondant aux valeurs prises par ces dés. Ce sont des variables indépendantes donc

$$\mathbb{P}(X \geq i) = \mathbb{P}(U_1 \geq i, U_2 \geq i, U_3 \geq i, U_4 \geq i) = \mathbb{P}(U_1 \geq i)\mathbb{P}(U_2 \geq i)\mathbb{P}(U_3 \geq i)\mathbb{P}(U_4 \geq i)$$

et puisqu'elles sont uniformes sur $\{1, \dots, 6\}$, on a

$$\mathbb{P}(U_1 \geq i) = \mathbb{P}(U_1 = i) + \cdots + \mathbb{P}(U_1 = 6) = \frac{1}{6} + \cdots + \frac{1}{6} = (7-i)/6,$$

d'où l'on déduit

$$\mathbb{P}(X \geq i) = \left(\frac{7-i}{6}\right)^4.$$

(c) De la formule de sommation des queues, on déduit alors

$$\mathbb{E}[X] = \sum_{i=1}^{6} \left(\frac{7-i}{6}\right)^4 = \left(\frac{6}{6}\right)^4 + \left(\frac{5}{6}\right)^4 + \cdots + \left(\frac{1}{6}\right)^4 \approx 1.755$$

(d) Soit S la somme des 3 plus gros scores. En notant $T = (U_1 + U_2 + U_3 + U_4)$ le total des quatre dés, il vient $T = X + S$, d'où $S = T - X$, et puisque $\mathbb{E}[U_1] = 7/2$ on en déduit

$$\mathbb{E}[S] = \mathbb{E}[T] - \mathbb{E}[X] = 4\mathbb{E}[U_1] - \mathbb{E}[X] = 4 \times \frac{7}{2} - \mathbb{E}[X] \approx 12.24$$

(e) Pour chaque valeur i entre 1 et 5, on peut écrire

$$\mathbb{P}(X = i) = \mathbb{P}(X \geq i) - \mathbb{P}(X \geq i+1) = \left(\frac{7-i}{6}\right)^4 - \left(\frac{7-(i+1)}{6}\right)^4 = \left(\frac{7-i}{6}\right)^4 - \left(\frac{6-i}{6}\right)^4.$$

4. Généralisation : soit X variable aléatoire à valeurs dans $\mathbb{N}^*$ admettant une espérance.

(a) Pour une variable à valeurs dans $\mathbb{N}^*$, la formule de sommation des queues s'écrit

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

(b) Soit T_1, T_2, T_3, T_4 les temps aléatoires nécessaires pour faire apparaître le 2 sur les 4 dés respectivement. Ces variables sont indépendantes et de même loi géométrique de paramètre $1/6$. Notons T la variable aléatoire correspondant au minimum de ces temps, $T = \min(T_1, T_2, T_3, T_4)$. Pour tout n de $\mathbb{N}^*$, on a par le même raisonnement que ci-dessus

$$\mathbb{P}(T \geq n) = \mathbb{P}(T_1 \geq n)^4,$$

or

$$\mathbb{P}(T_1 \geq n) = \sum_{k=n}^{+\infty} \mathbb{P}(T_1 = k) = \frac{1}{6} \sum_{k=n}^{+\infty} \left(\frac{5}{6}\right)^{k-1} = \frac{1}{6} \frac{(5/6)^{n-1}}{1 - 5/6} = \left(\frac{5}{6}\right)^{n-1}.$$

On en déduit que $\mathbb{P}(T \geq n) = (5/6)^{4(n-1)}$ et de la formule de sommation des queues :

$$\mathbb{E}[T] = \sum_{n=1}^{+\infty} \left(\frac{5}{6}\right)^{4(n-1)} = \frac{1}{1 - (5/6)^4}.$$

Remarque : ce résultat peut bien sûr se retrouver en disant que le minimum de 4 variables géométriques indépendantes de paramètre $1/6$ est géométrique de paramètre $1 - (5/6)^4$.

Exercice 2.26 (Somme de variables poissoniennes)

1. Soit $n \in \mathbb{N}$ fixé. Puisque X_1 et X_2 sont à valeurs dans $\mathbb{N}$, leur somme est égale à n si et seulement si il existe $k \in \{0, \dots, n\}$ tel que $X_1 = k$ et $X_2 = (n-k)$. C'est exactement ce que traduit l'égalité : $\{X = n\} = \bigcup_{k=0}^n \{X_1 = k, X_2 = n-k\}$.
2. X est à valeurs dans $\mathbb{N}$, donc pour connaître sa loi il suffit de calculer $\mathbb{P}(X = n)$ pour tout $n \in \mathbb{N}$, ce qui donne :
$$\mathbb{P}(X = n) = \mathbb{P}\left(\bigcup_{k=0}^n \{X_1 = k, X_2 = n-k\}\right) = \sum_{k=0}^n \mathbb{P}(X_1 = k, X_2 = n-k),$$
et on applique l'indépendance de X_1 et X_2 pur poursuivre le calcul :
$$\mathbb{P}(X = n) = \sum_{k=0}^n \mathbb{P}(X_1 = k)\mathbb{P}(X_2 = n-k) = \sum_{k=0}^n e^{-\lambda_1}\frac{\lambda_1^k}{k!}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!},$$
et la fin du calcul glisse comme un pet sur une toile cirée grâce à la formule du binôme :
$$\mathbb{P}(X = n) = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^n \binom{n}{k}\lambda_1^k\lambda_2^{n-k} = e^{-(\lambda_1+\lambda_2)}\frac{(\lambda_1+\lambda_2)^n}{n!}.$$
On en déduit que $X \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
3. Généralisation : soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois de Poisson de paramètres respectifs $\lambda_1, \dots, \lambda_n$, avec $\lambda_i > 0$ pour tout $i \in \{1, \dots, n\}$, alors la variable $X = X_1 + \cdots + X_n$ suit elle aussi une loi de Poisson, de paramètre $\lambda = \lambda_1 + \cdots + \lambda_n$.

Exercice 2.27 (Un calendrier stochastique)

1. La variable X peut prendre les valeurs $\{28, 30, 31\}$ avec les probabilités $1/12$, $4/12$ et $7/12$.
2. La fonction de répartition de X est représentée figure 2.16.

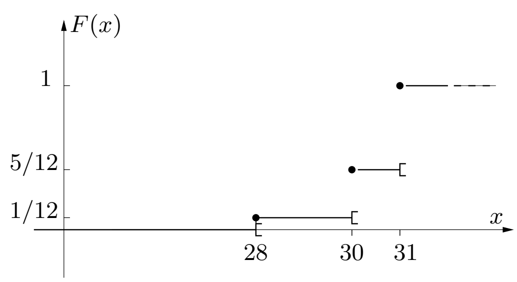


Figure 2.16 – Fonction de répartition de X.

3. L'espérance de X est donc
$$\mathbb{E}[X] = 28 \times \frac{1}{12} + 30 \times \frac{4}{12} + 31 \times \frac{7}{12} = \frac{365}{12} \approx 30.42$$
On a par ailleurs
$$\sigma(X) = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2} = \sqrt{\left(28^2 \times \frac{1}{12} + 30^2 \times \frac{4}{12} + 31^2 \times \frac{7}{12}\right) - \left(\frac{365}{12}\right)^2} \approx 0.86$$
4. La variable Y peut elle aussi prendre les valeurs $\{28, 30, 31\}$. Mais à la différence de X, la probabilité qu'elle prenne :

– la valeur 28 est $28/365$;
– la valeur 30 est $(4 \times 30)/365 = 120/365$;
– la valeur 31 est $(7 \times 31)/365 = 217/365$;

La moyenne de Y est donc

$$\mathbb{E}[Y] = 28 \times \frac{28}{365} + 30 \times \frac{120}{365} + 31 \times \frac{217}{365} = \frac{11111}{365} \approx 30.44$$

Exercice 2.28 (Loi à partir des moments)
On considère une variable aléatoire X à valeurs dans $\{0, 1, 2\}$. On sait que $\mathbb{E}[X] = 1$ et $\mathrm{Var}(X) = \frac{1}{2}$. Notons p_0, p_1 et p_2 les probabilités que X prenne respectivement les valeurs 0, 1 et 2. Le fait que les p_i doivent sommer à 1 et les connaissances de l'espérance et de la variance permettent d'écrire un système linéaire de 3 équations à 3 inconnues, lequel se résout sans difficulté :

$$\left\{ \begin{array}{l} p_0 + p_1 + p_2 = 1 \\ 0 \times p_0 + 1 \times p_1 + 2 \times p_2 = 1 \\ 0^2 \times p_0 + 1^2 \times p_1 + 2^2 \times p_2 - 1^2 = 1/2 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} p_0 = 1/4 \\ p_1 = 1/2 \\ p_2 = 1/4 \end{array} \right.$$

On reconnaît en fait une loi binomiale : $X \sim \mathcal{B}(2, 1/2)$.

Exercice 2.29 (Dés et accidents)

1. On dispose de deux dés qu'on lance simultanément 12 fois de rang et on appelle X le nombre de double six obtenus sur les 12 lancers.
   (a) La variable X suit une loi binomiale : $X \sim \mathcal{B}(12, 1/36)$. Moyenne et variance valent donc $\mathbb{E}[X] = 1/3$ et $\mathrm{Var}(X) = 35/108$.
   (b) On a
   $$\mathbb{P}(X \leq 2) = \binom{12}{0}\left(\frac{1}{36}\right)^0\left(\frac{35}{36}\right)^{12} + \binom{12}{1}\left(\frac{1}{36}\right)^1\left(\frac{35}{36}\right)^{11} + \binom{12}{2}\left(\frac{1}{36}\right)^2\left(\frac{35}{36}\right)^{10}$$
   d'où $\mathbb{P}(X \leq 2) \approx 0.996$.
   (c) On approche la loi de X par une loi de Poisson de même moyenne, donc $X \approx \mathcal{P}(1/3)$, d'où
   $$\mathbb{P}(X \leq 2) \approx e^{-\frac{1}{3}}\frac{\left(\frac{1}{3}\right)^0}{0!} + e^{-\frac{1}{3}}\frac{\left(\frac{1}{3}\right)^1}{1!} + e^{-\frac{1}{3}}\frac{\left(\frac{1}{3}\right)^2}{2!} \approx 0.995$$
   Ainsi l'approximation fonctionne très bien, bien que le premier paramètre de la loi binomiale n'est pas très grand (ici $n = 12$).
2. (a) Puisque $X \sim \mathcal{P}(2)$, on a
   $$\mathbb{P}(X = 4) = e^{-2} \times \frac{2^4}{4!} \approx 0.09$$
   (b) Les lois de Poisson sont stables par sommation indépendante : $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$.
   (c) La question précédente permet d'affirmer que $Y \sim \mathcal{P}(4)$, d'où
   $$\mathbb{P}(Y = 0) = e^{-4} \times \frac{4^0}{0!} \approx 0.018.$$
   On pouvait également arriver à ce résultat directement en utilisant l'indépendance des accidents d'une semaine à l'autre :
   $$\mathbb{P}(Y = 0) = \mathbb{P}(\{X_1 = 0\} \cap \{X_2 = 0\}) = \mathbb{P}(X_1 = 0)\mathbb{P}(X_2 = 0) = e^{-2} \times \frac{2^0}{0!} \times e^{-2} \times \frac{2^0}{0!}$$
   d'où l'on retrouve bien : $\mathbb{P}(Y = 0) = e^{-4}$.

Exercice 2.30 (Test sanguin)

1. Soit X le nombre de soldats porteurs de cette maladie. Puisque les soldats sont porteurs ou non de la maladie indépendamment les uns des autres, X suit une loi binomiale $\mathcal{B}(500, 1/1000)$. Ainsi sa moyenne vaut-elle $\mathbb{E}[X] = 500 \times 1/1000 = 1/2$.
2. Dans ces conditions, on peut approcher la loi de X par une loi de Poisson de même moyenne, c'est-à-dire $X \approx \mathcal{P}(1/2)$.
3. La probabilité que le test soit positif est, en utilisant la loi binomiale :
$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - \left(1 - \frac{1}{1000}\right)^{500} \approx 0.3936$$
Si on utilise l'approximation par la loi de Poisson, on trouve :
$$\mathbb{P}(X \geq 1) = 1 - \mathbb{P}(X = 0) = 1 - e^{-\frac{1}{2}} \approx 0.3935$$
Donc, pour cette probabilité, l'erreur d'approximation est très faible, de l'ordre de 10^{-4}.
4. On cherche cette fois une probabilité conditionnelle, à savoir :
$$\mathbb{P}(X \geq 2 | X \geq 1) = \frac{\mathbb{P}(\{X \geq 2\} \cap \{X \geq 1\})}{\mathbb{P}(X \geq 1)} = \frac{\mathbb{P}(X \geq 2)}{\mathbb{P}(X \geq 1)},$$
avec, toujours par l'approximation de Poisson :
$$\mathbb{P}(X \geq 2) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1)) = 1 - \left(e^{-\frac{1}{2}} + \frac{1}{2}e^{-\frac{1}{2}}\right),$$
d'où
$$\mathbb{P}(X \geq 2 | X \geq 1) = \frac{1 - e^{-\frac{1}{2}}}{1 - \frac{3}{2}e^{-\frac{1}{2}}} \approx 0.229$$
5. Si Jean est malade, la probabilité qu'il y ait au moins une autre personne malade est par contre la probabilité que parmi 499 personnes, au moins une soit atteinte. En d'autres termes, on est ramené au calcul de la question 3 en remplaçant 500 par 499 :
$$p = 1 - \left(1 - \frac{1}{1000}\right)^{499} \approx 0.3930$$
6. Le raisonnement de la question précédente s'applique *mutatis mutandis*. En notant p_n la probabilité, en fonction de n, qu'une des personnes restantes au moins soit malade, on obtient
$$p_n = 1 - \left(1 - \frac{1}{1000}\right)^{500-n} \approx 1 - e^{-\frac{500-n}{1000}}$$

Exercice 2.31 (Boules blanches et noires)
Cet exercice est corrigé en annexe (sujet de décembre 2009).

Exercice 2.32 (Défaut de fabrication)
Cet exercice est corrigé en annexe (sujet de décembre 2009).

Exercice 2.33 (Recrutement)
Cet exercice est corrigé en annexe (sujet de décembre 2009).

Exercice 2.34 (Lancer de dé)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 2.35 (Le dé dyadique)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 2.36 (Répartition des tailles)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 2.37 (Poisson en vrac)
Cet exercice est corrigé en annexe (sujet de novembre 2010).

Exercice 2.38 (Jeu d'argent)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 2.39 (Rubrique à brac)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 2.40 (Ascenseur pour l'échafaud)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 2.41 (Systèmes de contrôle)
Cet exercice est corrigé en annexe (sujet de novembre 2011).

Exercice 2.42 (Kramer contre Kramer)
Cet exercice est corrigé en annexe (sujet de novembre 2011).

Exercice 2.43 (Loterie)
Cet exercice est corrigé en annexe (sujet de novembre 2011).

Exercice 2.44 (Dé coloré)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Exercice 2.45 (Beaujolais nouveau)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Chapitre 3

Variables aléatoires à densité

Introduction

Contrairement au chapitre précédent, on considère ici des variables aléatoires à valeurs dans $\mathbb{R}$ ou un intervalle de $\mathbb{R}$. Parmi celles-ci, seules celles dont la loi admet une représentation intégrale nous intéresseront : on parle alors de variables à densité, ou absolument continues. On pourra ainsi retrouver l'ensemble des notions vues pour les variables discrètes.

3.1 Densité d'une variable aléatoire

Dans toute la suite, $(\Omega, \mathcal{F}, \mathbb{P})$ désigne un espace probabilisé. On commence par définir la notion de variable aléatoire de façon très générale, valable en particulier pour les variables discrètes. Rappelons qu'un intervalle I de $\mathbb{R}$ est un ensemble d'un seul tenant, c'est-à-dire de la forme $I = [a, b]$ ou $I =]a, b]$ ou $I = [a, b[$ ou $I =]a, b[$, avec a et b deux réels, voire $-\infty$ et/ou $+\infty$ à condition d'ouvrir les crochets.

Définition 3.1 (Variable aléatoire)
Une application

$$X : \left\{ \begin{array}{rl} (\Omega, \mathcal{F}, \mathbb{P}) & \to \mathbb{R} \\ \omega & \mapsto X(\omega) \end{array} \right.$$

est une variable aléatoire si pour tout intervalle I de $\mathbb{R}$:

$$\{X \in I\} := X^{-1}(I) = \{\omega \in \Omega \ : \ X(\omega) \in I\} \in \mathcal{F}.$$

Commençons par remarquer que si X prend ses valeurs dans un sous-ensemble au plus dénombrable $\mathcal{X}$ de $\mathbb{R}$, on retrouve la définition 2.1 d'une variable discrète puisqu'alors :

$$\{X \in I\} = \bigcup_{i \ : \ x_i \in I} \{X = x_i\} \in \mathcal{F},$$

car la tribu $\mathcal{F}$ est stable par union au plus dénombrable. Réciproquement, si $I = \{x_i\} = [x_i, x_i]$, alors suivant la définition ci-dessus :

$$\{X \in I\} = \{X = x_i\} \in \mathcal{F},$$

et la boucle est bouclée.

Expliquons en deux mots la définition générale ci-dessus. L'intérêt de supposer qu'un événement appartient à $\mathcal{F}$ est bien sûr d'être assuré qu'on pourra en calculer la probabilité. Mais dans le cas où X prend ses valeurs dans $\mathbb{R}$ tout entier (ou dans un intervalle de $\mathbb{R}$), on ne veut plus se restreindre à des événements de la forme "X prend la valeur x" comme dans le cas discret, car ceux-ci seront généralement de probabilité nulle, donc sans grand intérêt. Le fait de supposer qu'on va pouvoir calculer la probabilité de n'importe quel intervalle est bien plus pertinent, car à eux seuls les intervalles engendrent une famille très riche de sous-ensembles de $\mathbb{R}$, connue sous le nom de tribu borélienne.

Comme dans le cas discret, la notion de variable aléatoire est stable par toutes les opérations classiques sur les fonctions : la combinaison linéaire, le produit, le minimum, le maximum de deux variables aléatoires X et Y sont encore des variables aléatoires.

Outre le cas discret, un cas confortable de variable aléatoire à valeurs dans $\mathbb{R}$ est celui de variable admettant une densité. Dans toute la suite, on restera volontairement évasif quant aux hypothèses précises sur cette densité, le bon cadre pour toutes ces notions étant celui de l'intégrale de Lebesgue, qui ne sera vu qu'en troisième année. Pour fixer les idées, on pourra par exemple se dire que f est continue par morceaux sur $\mathbb{R}$, ce qui sera le cas dans la plupart des exemples.

Définition 3.2 (Variable à densité)
Soit X une variable aléatoire à valeurs dans $\mathbb{R}$. On dit que X est à densité, ou absolument continue, s'il existe une fonction $f : \mathbb{R} \to \mathbb{R}$ vérifiant :

1. $f \geq 0$,
2. $\int_{-\infty}^{+\infty} f(x)dx = 1$,

et telle que pour tout intervalle I de $\mathbb{R}$:

$$\mathbb{P}(X \in I) = \int_I f(x)dx.$$

Dans ce cas, f est appelée densité de X.

Remarque. Notons d'emblée que si X admet une densité, alors la probabilité qu'elle prenne une valeur donnée x_0 est nulle puisque :

$$\mathbb{P}(X = x_0) = \mathbb{P}(X \in [x_0, x_0]) = \int_{x_0}^{x_0} f(x)dx = 0.$$

Interprétation. Bien que $f(x_0)$ soit positif pour tout x_0, $f(x_0)$ n'est pas une probabilité : on peut en particulier avoir $f(x_0) > 1$! Il n'empêche que si $[x_0 - \delta/2, x_0 + \delta/2]$ est un "petit" intervalle centré en x_0, la probabilité de tomber dans cet intervalle est :

$$\mathbb{P}(X \in [x_0 - \delta/2, x_0 + \delta/2]) = \int_{x_0-\delta/2}^{x_0+\delta/2} f(x)dx \approx \delta f(x_0).$$

Autrement dit, à largeur d'intervalle δ fixée, X a d'autant plus de chances de tomber dans un intervalle centré en x_0 que $f(x_0)$ est grand. Pratiquement, imaginons qu'on représente un très grand nombre de réalisations $X_1 = x_1, \dots, X_n = x_n$ tirées indépendamment suivant la même loi que X : la **densité** des X_i sera alors la plus élevée là où f prend ses plus grandes valeurs.

Exemples :

1. Loi uniforme : conformément à ce qui vient d'être dit, une variable X se répartissant de façon uniforme dans le segment $[0, 1]$ est définie par la densité $f = \mathbb{1}_{[0,1]}$, c'est-à-dire $f(x) = 1$ si $x \in [0, 1]$ et $f(x) = 0$ sinon. On a bien d'une part $f \geq 0$ et d'autre part :

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^1 1dx = 1.$$

Par ailleurs, pour tous points $a \leq b$ de $[0,1]$, la probabilité de tomber entre a et b est :

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx = (b-a),$$

donc ne dépend que de la longueur de l'intervalle, ce qui est bien conforme à l'intuition d'une variable uniforme : X a autant de chances de tomber dans $[0,1/3]$ que dans $[2/3,1]$. On dit que X suit une loi uniforme sur le segment $[0,1]$ et on note $X \sim \mathcal{U}_{[0,1]}$.

2. Loi exponentielle : prenons maintenant $f(x) = e^{-x}$ si $x \geq 0$ et $f(x) = 0$ si $x < 0$. Alors à nouveau $f \geq 0$ et :

$$\int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} e^{-x}dx = \left[-e^{-x}\right]_0^{+\infty} = 1,$$

donc f définit bien une densité. Puisque f décroît à vitesse exponentielle vers 0 lorsque x tend vers l'infini (cf. figure 3.1 à gauche), on voit cette fois que X a plutôt tendance à prendre des valeurs petites. On dit qu'elle suit une loi exponentielle de paramètre 1 et on note $X \sim \mathcal{E}(1)$.

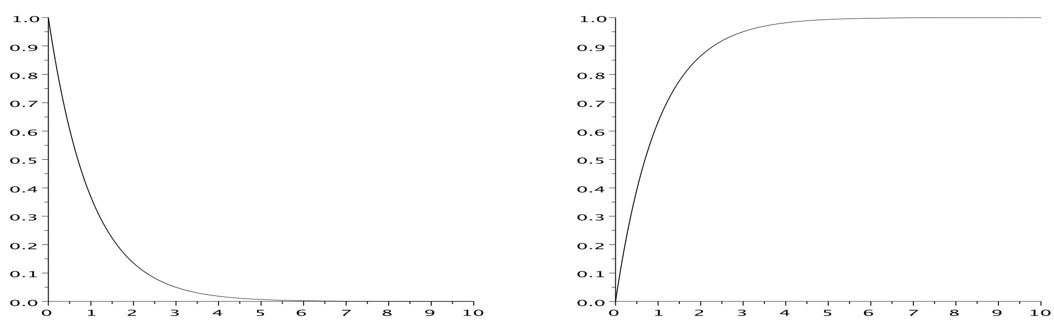


Figure 3.1 – Densité et fonction de répartition de la loi exponentielle de paramètre 1.

3.2 Fonction de répartition

Comme dans le chapitre précédent, on peut définir très facilement la fonction de répartition d'une variable aléatoire absolument continue. Le lien entre intégrale et primitive en fait d'ailleurs un outil bien plus puissant que dans le cas discret.

Définition 3.3 (Fonction de répartition)
Soit X une variable aléatoire absolument continue, de densité f. La fonction de répartition de X est la fonction F définie par :

$$F : \left\{ \begin{array}{ll} \mathbb{R} & \to \mathbb{R} \\ x & \mapsto F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^x f(t)dt. \end{array} \right.$$

La fonction de répartition permet de calculer la probabilité de tomber dans n'importe quel intervalle, par exemple :

$$\mathbb{P}(0 < X \leq 1) = \mathbb{P}(0 < X < 1) = \mathbb{P}(0 \leq X \leq 1) = \mathbb{P}(0 \leq X < 1) = \int_0^1 f(t)dt = F(1) - F(0).$$

Exemples :

1. Loi uniforme : dans le cas où $f = \mathbb{1}_{[0,1]}$, un petit calcul donne
$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$
2. Loi exponentielle : si $f(x) = e^{-x}\mathbb{1}_{\{x\geq 0\}}$, on a (voir aussi figure 3.1 à droite)
$$F(x) = \begin{cases} 0 & \text{si } x \leq 0 \\ 1 - e^{-x} & \text{si } x \geq 0 \end{cases}$$

On retrouve sur ces exemples les propriétés vues dans le cas discret, et même mieux : monotonie, limites en $\pm\infty$, continuité (et non plus simplement continuité à droite). Elles sont en fait toujours vraies.

Propriétés 3.1 (Propriétés d'une fonction de répartition)
Soit X une variable aléatoire absolument continue, de densité f. Sa fonction de répartition F a les propriétés suivantes :

1. *F est croissante ;*
2. *$\lim_{x\to-\infty} F(x) = 0$, $\lim_{x\to+\infty} F(x) = 1$;*
3. *F est continue sur $\mathbb{R}$.*

Preuve. Les deux premiers points se prouvent comme dans le cas discret. La continuité à droite se montre aussi comme dans le cas discret, il reste donc simplement à montrer la continuité à gauche en tout point. Soit donc x_0 réel fixé. Puisque F est croissante sur $\mathbb{R}$, elle admet une limite à gauche en x_0, notée $F(x_0^-)$ et qui vérifie donc $F(x_0^-) = \lim_{n\to+\infty} F(x_0 - 1/n)$. Nous voulons maintenant prouver que $F(x_0^-) = F(x_0)$, ou de façon équivalente que $\lim_{n\to+\infty} F(x_0 - 1/n) = F(x_0)$. Supposons pour simplifier que f est continue par morceaux, alors f est bornée au voisinage de x_0, disons par M, d'où :
$$|F(x_0) - F(x_0 - 1/n)| = \left| \int_{x_0-1/n}^{x_0} f(x)dx \right| \leq \int_{x_0-1/n}^{x_0} |f(x)|dx,$$
ce qui donne :
$$|F(x_0) - F(x_0 - 1/n)| \leq \int_{x_0-1/n}^{x_0} M dx = \frac{M}{n} \xrightarrow[n\to+\infty]{} 0,$$
et la continuité à gauche est prouvée. Ainsi F est bien continue sur $\mathbb{R}$.

■

Quid des sauts de F ? Nous avions vu dans le cas discret que :
$$\forall x \in \mathbb{R} \qquad \mathbb{P}(X = x) = F(x) - F(x^-).$$
Cette relation est encore vraie, mais devient une tautologie, puisqu'elle ne dit rien de plus que 0=0. L'intérêt était pourtant de voir le lien entre la loi de X et sa fonction de répartition. Le résultat suivant rétablit ce lien.

Proposition 3.1 (Lien entre fonction de répartition et densité)
Soit X une variable aléatoire absolument continue, de densité f et de fonction de répartition F. Alors en tout point où f est continue, F est dérivable et on a $F'(x) = f(x)$.

Preuve. Soit x_0 un point de continuité de f. Il s'agit ici de montrer que la limite du taux de variation de F en x_0 existe et vaut $f(x_0)$, c'est-à-dire :

$$\frac{F(x_0+\delta)-F(x_0)}{\delta} \xrightarrow[\delta\to 0]{} f(x_0),$$

ce qui s'écrit encore :

$$\forall \varepsilon > 0, \exists \delta > 0, |x - x_0| \leq \delta \implies \left|\frac{F(x_0+\delta)-F(x_0)}{\delta} - f(x_0)\right| \leq \varepsilon.$$

Soit donc $\varepsilon > 0$ fixé. Puisque f est continue en x_0, il existe $\delta > 0$ tel que $|x - x_0| \leq \delta$ implique $|f(x) - f(x_0)| \leq \varepsilon$. Il s'ensuit donc de façon générale :

$$\left|\frac{F(x_0+\delta)-F(x_0)}{\delta} - f(x_0)\right| = \left|\frac{1}{\delta}\int_{x_0}^{x_0+\delta}(f(x)-f(x_0))dx\right| \leq \frac{1}{\delta}\int_{x_0}^{x_0+\delta}|f(x)-f(x_0)|dx,$$

et pour $|x - x_0| \leq \delta$, il vient :

$$\left|\frac{F(x_0+\delta)-F(x_0)}{\delta} - f(x_0)\right| \leq \varepsilon,$$

ce qui achève la preuve.

■

Exemples :

1. Loi uniforme : hormis en 0 et en 1, F est dérivable partout, avec $F'(x) = f(x)$.
2. Loi exponentielle : hormis en 0, F est dérivable partout, avec $F'(x) = f(x)$.

Remarque. Réciproquement, on peut montrer que si une variable aléatoire X admet une fonction de répartition F, définie par $F(x) = \mathbb{P}(X \leq x)$, qui est continue partout, dérivable sauf éventuellement en un nombre fini de points et de dérivée notée f aux points de dérivabilité, alors X est absolument continue et de densité f.

Exemple. Supposons que X suive une loi uniforme sur $[0,1]$ et définissons la variable aléatoire Y par $Y = X^2$. Y admet-elle une densité ? Si oui, quelle est-elle ? Pour répondre à cette question, on passe par la fonction de répartition F de Y, qui est donc définie pour tout réel y par $F(y) = \mathbb{P}(Y \leq y)$. Puisque X ne prend ses valeurs qu'entre 0 et 1, il est clair qu'il en va de même pour Y, d'où l'on déduit que $F(y) = 0$ pour $y \leq 0$ et $F(y) = 1$ pour $y \geq 1$. Considérant maintenant $y \in]0,1[$, on peut écrire :

$$F(y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \mathbb{P}(X \leq \sqrt{y}) = \sqrt{y},$$

la dernière égalité étant issue du calcul de la fonction de répartition de X vu précédemment. Puisque $y \mapsto \sqrt{y}$ est dérivable sur $]0,1[$, il en va de même pour F, avec $F'(y) = \frac{1}{2\sqrt{y}}$. On en déduit que Y est absolument continue, de densité f définie par :

$$f(y) = \begin{cases} 0 & \text{si } y \leq 0 \\ \frac{1}{2\sqrt{y}} & \text{si } 0 < y < 1 \\ 0 & \text{si } y \geq 1 \end{cases}$$

Densité et fontion de répartition de Y sont données figure 3.2.
La technique vue sur cet exemple pour trouver la densité de Y se généralise en fait à tout "bon" changement de variable $Y = \varphi(X)$.

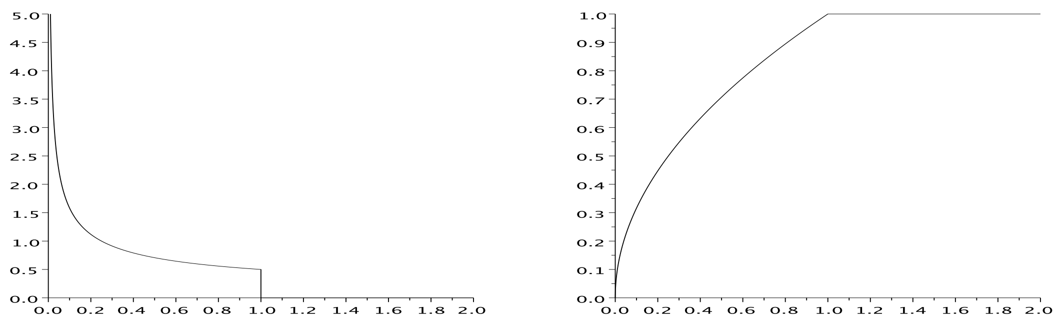


FIGURE 3.2 – Densité et fonction de répartition de $Y = X^2$, où $X \sim \mathcal{U}_{[0,1]}$.

Proposition 3.2 (Changement de variable)
Soit X une variable aléatoire à valeurs dans l'intervalle I et admettant une densité f_X. Soit alors $Y = \varphi(X)$, avec φ dérivable et bijective, variable aléatoire à valeurs dans l'intervalle J. Alors Y admet pour densité f_Y définie par :

$$f_Y(y) = \frac{f_X(\varphi^{-1}(y))}{|\varphi'(\varphi^{-1}(y))|},$$

en tout point y de J tel que $\varphi'(\varphi^{-1}(y)) \neq 0$.

Preuve. Il suffit de généraliser le raisonnement de l'exemple précédent. Notons F_Y la fonction de répartition de Y, définie pour tout réel y par $F_Y(y) = \mathbb{P}(Y \leq y)$, alors la relation entre X et Y permet d'écrire, en supposant par exemple φ croissante :

$$F_Y(y) = \mathbb{P}(\varphi(X) \leq y) = \mathbb{P}(X \leq \varphi^{-1}(y)) = F_X(\varphi^{-1}(y)),$$

en notant F_X la fonction de répartition de X. Ainsi, en tout point y où la fonction $y \mapsto F_X(\varphi^{-1}(y))$ est dérivable, F_Y l'est aussi et :

$$F_Y'(y) = F_X'(\varphi^{-1}(y))(\varphi^{-1}(y))' = \frac{f_X(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))},$$

avec $\varphi'(\varphi^{-1}(y)) > 0$ puisque φ est croissante. Si φ est décroissante, les calculs précédents deviennent :

$$F_Y(y) = \mathbb{P}(\varphi(X) \leq y) = \mathbb{P}(X \geq \varphi^{-1}(y)) = 1 - F_X(\varphi^{-1}(y)),$$

d'où :

$$F_Y'(y) = -F_X'(\varphi^{-1}(y))(\varphi^{-1}(y))' = -\frac{f_X(\varphi^{-1}(y))}{\varphi'(\varphi^{-1}(y))} = \frac{f_X(\varphi^{-1}(y))}{|\varphi'(\varphi^{-1}(y))|},$$

puisque cette fois $\varphi'(\varphi^{-1}(y)) < 0$.

■

Remarque. Cette formule est bien sûr liée à la formule de changement de variable connue pour les intégrales. Il n'est pas utile de la retenir, à condition de savoir la retrouver via le passage par la fonction de répartition vu sur l'exemple ci-dessus.

3.3 Moments d'une variable à densité

Cette section est calquée sur celle du chapitre 2, en remplaçant les sommes par des intégrales, les x_i par x et les p_i par $f(x)$. On considère dans toute la suite une variable aléatoire X de densité f sur $\mathbb{R}$.

Définition 3.4 (Espérance)
On appelle espérance de la variable X la quantité :

$$\mathbb{E}[X] = \int_{\mathbb{R}} xf(x)dx,$$

sous réserve de convergence de cette intégrale.

Remarque. Par analogie avec le cas discret, on aurait pu s'attendre à ce que soit requise l'absolue convergence de l'intégrale dans la définition de l'espérance. C'est en fait inutile : si X est à valeurs dans l'intervalle I, alors l'intégrale définissant l'espérance peut être généralisée pour 2 grandes raisons : borne(s) de l'intervalle infinie(s) et/ou fonction infinie en l'une ou les deux bornes. Quoi qu'il en soit, l'étude de la convergence équivaut à celle de l'absolue convergence puisque $xf(x)$ est de signe constant au voisinage de la borne étudiée.

Interprétation. Comme dans le cas discret, l'espérance de X peut être vue comme la moyenne des valeurs x pondérées par les probabilités infinitésimales $f(x)dx$, c'est pourquoi on dit aussi **moyenne** de X pour parler de son espérance. En particulier, si X prend ses valeurs entre a et b (i.e. $\mathcal{X} \subset [a, b]$), on aura nécessairement $a \leq \mathbb{E}[X] \leq b$.

Exemples :

1. Loi uniforme : si $X \sim \mathcal{U}_{[0,1]}$, alors son espérance vaut

$$\mathbb{E}[X] = \int_0^1 xdx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2},$$

ce qui est bien la moyenne attendue, par symétrie de la loi de X autour de $1/2$.

2. Loi exponentielle : si $X \sim \mathcal{E}(1)$, alors une intégration par parties donne

$$\mathbb{E}[X] = \int_0^{+\infty} xe^{-x}dx = \left[-xe^{-x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-x}dx = \left[-e^{-x}\right]_0^{+\infty} = 1.$$

Voyons maintenant une loi classique n'ayant pas d'espérance.

Contre-exemple : la loi de Cauchy. On considère une variable aléatoire X de densité $f(x) = \frac{1}{\pi(1+x^2)}$ (cf. figure 3.3). La fonction f est bien une densité puisqu'elle est positive et intègre à 1. On dit que X suit une loi de Cauchy de paramètre 1. Néanmoins, si on veut calculer son espérance, soit formellement $\int_{-\infty}^{+\infty} xf(x)dx$, il faut la convergence de l'intégrale en $+\infty$ et en $-\infty$. Or en $+\infty$, on a $f(x) \sim \frac{1}{x}$, donc l'intégrale est divergente et X n'admet pas d'espérance.

Revenons à des choses moins pathologiques. Etant donné une variable X dont on connaît la loi, il arrive souvent qu'on veuille calculer non pas l'espérance de X mais l'espérance d'une fonction de X. Le résultat suivant donne une façon très simple de le faire. Sa preuve est admise.

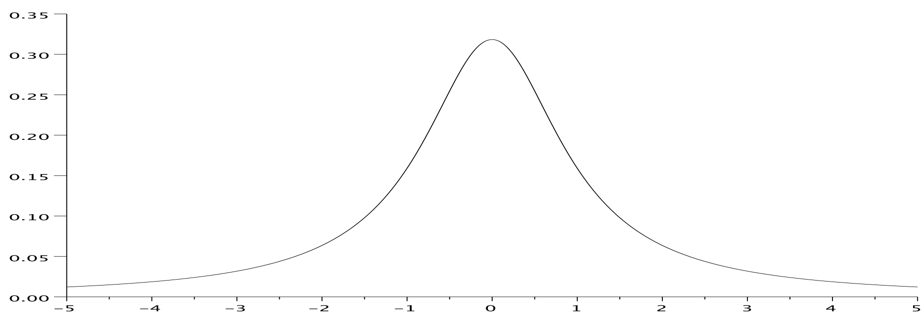


FIGURE 3.3 – Densité d'une variable aléatoire suivant une loi de Cauchy.

Théorème 3.1 (Théorème de Transfert)
Soit X une variable aléatoire à densité et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction, alors l'espérance de $\varphi(X)$ vaut :

$$\mathbb{E}[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) f(x) dx,$$

sous réserve d'absolue convergence de cette intégrale.

Remarques :

1. Etant donné que $\varphi(x)$ peut osciller entre des valeurs positives et négatives aux bords de l'intervalle d'intégration, on est cette fois obligé d'imposer l'absolue convergence.
2. Si φ est dérivable et bijective, on peut appliquer le résultat de changement de variable vu plus haut : $Y = \varphi(X)$ admet une densité f_Y, donc son espérance vaut :

$$\mathbb{E}[Y] = \int_J y f_Y(y) dy,$$

et le changement de variable $y = \varphi(x)$ dans l'intégrale donne :

$$\mathbb{E}[Y] = \int_I \varphi(x) f_Y(\varphi(x)) |\varphi'(x)| dx.$$

La valeur absolue vient de ce qu'on a mis l'intervalle I dans le "bon sens", ce qui n'est automatique que lorsque φ est croissante. Il reste alors à faire le lien avec la Proposition 3.2 pour retrouver :

$$\mathbb{E}[Y] = \int_I \varphi(x) f(x) dx.$$

Moyen mnémotechnique. Pour calculer $\mathbb{E}[\varphi(X)]$ et non $\mathbb{E}[X]$, on a juste à remplacer x par $\varphi(x)$ dans la formule de $\mathbb{E}[X]$.

L'intérêt pratique de ce résultat est le même que dans le cas discret : si on appelle $Y = \varphi(X)$ la variable aléatoire qui nous intéresse, on n'a pas besoin de commencer par déterminer sa densité pour calculer son espérance, il suffit tout simplement de se servir de celle de X.

Exemple. On reprend l'exemple où $X \sim \mathcal{U}_{[0,1]}$ et $Y = X^2$. Puisqu'on a vu que Y admet pour densité $f_Y(y) = \frac{1}{2\sqrt{y}} \mathbb{1}_{\{0<y<1\}}$, on a :

$$\mathbb{E}[Y] = \int_{\mathbb{R}} y f(y) dy = \int_0^1 \frac{\sqrt{y}}{2} dy = \left[\frac{y\sqrt{y}}{3}\right]_0^1 = \frac{1}{3}.$$

Mais on peut tout aussi bien appliquer directement le théorème de transfert :

$$\mathbb{E}[Y] = \mathbb{E}[X^2] = \int_{\mathbb{R}} x^2 f(x)dx = \int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}.$$

On énumère maintenant un ensemble de propriétés concernant l'espérance.

Propriétés 3.2 (Propriétés de l'espérance)

1. *Pour tous réels a et b : $\mathbb{E}[aX+b] = a\mathbb{E}[X] + b$.*
2. *Si $X \geq 0$, i.e. si X ne prend que des valeurs positives, on a $\mathbb{E}[X] \geq 0$.*
3. *Si $X \leq Y$, alors $\mathbb{E}[X] \leq \mathbb{E}[Y]$.*

Preuve. Ce sont les mêmes que dans le cas discret.

■

Remarque. Contrairement au cas discret, la notion de variable absolument continue n'est pas stable par combinaison linéaire. Autrement dit, deux variables aléatoires X et Y peuvent avoir chacune une densité sans que la variable $Z = (X + Y)$ en ait une. Il suffit pour s'en convaincre de prendre $X \sim \mathcal{U}_{[0,1]}$ et $Y = 1 - X$.

Nous avons dit que l'espérance est une mesure de tendance centrale, nous allons voir maintenant une mesure de dispersion autour de cette valeur centrale : la variance.

Définition 3.5 (Variance & Ecart-type)
Soit X une variable aléatoire de densité f et admettant une espérance $\mathbb{E}[X]$. La variance de X est définie par :

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \int_{\mathbb{R}} (x - \mathbb{E}[X])^2 f(x)dx,$$

sous réserve de convergence de cette intégrale. On appelle alors écart-type, noté $\sigma(X)$, la racine de la variance : $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.

Interprétation. C'est la même que dans le cas discret : de façon générale, la variance d'une variable mesure la moyenne des carrés des écarts à sa moyenne. Si X représente une grandeur physique, alors l'écart-type a la même dimension que X, tandis que la variance a cette dimension au carré, ce qui la rend moins parlante en pratique. Le terme écart-type est encore à comprendre au sens "écart typique" d'une variable à sa moyenne.

Exemples :

1. Loi uniforme : si $X \sim \mathcal{U}_{[0,1]}$, nous avons vu que $\mathbb{E}[X] = 1/2$, sa variance vaut donc :

$$\mathrm{Var}(X) = \int_0^1 \left(x - \frac{1}{2}\right)^2 dx = \left[\frac{(x-\frac{1}{2})^3}{3}\right]_0^1 = \frac{1}{12}.$$

2. Loi exponentielle : si $X \sim \mathcal{E}(1)$, nous avons vu que $\mathbb{E}[X] = 1$, sa variance vaut donc :

$$\mathrm{Var}(X) = \int_0^{+\infty} (x-1)^2 e^{-x} dx,$$

et après deux intégrations par parties successives, on trouve $\mathrm{Var}(X) = 1$.

Les propriétés suivantes, ainsi que leurs preuves, sont les mêmes que dans le cas discret.

Propriétés 3.3
Soit X une variable aléatoire à densité, alors sous réserve d'existence de sa variance :
(i) $\mathrm{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$.
(ii) Si a et b sont deux réels, $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$.

On va maintenant généraliser les notions d'espérance et de variance.

Définition 3.6
Soit X une variable aléatoire de densité f et $m \in \mathbb{N}^*$. Sous réserve de convergence de l'intégrale, on appelle :
(i) moment d'ordre m de X la quantité

$$\mathbb{E}[X^m] = \int_{\mathbb{R}} x^m f(x)dx.$$

(ii) moment centré d'ordre m de X la quantité

$$\mathbb{E}[(X-\mathbb{E}[X])^m] = \int_{\mathbb{R}} (x-\mathbb{E}[X])^m f(x)dx.$$

Ainsi l'espérance de X est le moment d'ordre 1 et sa variance le moment centré d'ordre 2. Rappelons aussi que X est dite centrée si $\mathbb{E}[X]=0$ et réduite si $\mathrm{Var}[X]=1$. On dit qu'on centre et réduit X en considérant la variable $Y = \frac{X-\mathbb{E}[X]}{\sigma(X)}$. Le moment d'ordre 3 de Y est alors appelé coefficient d'asymétrie (skewness) de X et le moment d'ordre 4 de Y est appelé kurtosis, ou coefficient d'aplatissement, de X.

Proposition 3.3
Soit X une variable aléatoire à densité, alors si X admet un moment d'ordre $m \in \mathbb{N}^*$, X admet des moments de tout ordre $j \in \{1, \dots, m\}$.

L'existence d'un moment d'ordre élevé assure une décroissance d'autant plus rapide de la queue de la distribution de X à l'infini, comme le montre l'inégalite de Markov. Tout comme le résultat précédent, ce théorème est admis.

Théorème 3.2 (Inégalité de Markov)
Soit X une variable aléatoire à densité, alors si X admet un moment d'ordre $m \in \mathbb{N}^*$, on a :

$$\forall t > 0 \qquad \mathbb{P}(|X| \geq t) \leq \frac{\mathbb{E}[|X|^m]}{t^m}.$$

Une première conséquence de l'inégalité de Markov : une variable ayant des moments de tout ordre a une queue de distribution à décroissance plus rapide que n'importe quelle fraction rationnelle (par exemple exponentielle). Une seconde conséquence : Tchebychev.

Théorème 3.3 (Inégalité de Tchebychev)
Soit X une variable aléatoire admettant une variance, alors :

$$\forall t > 0 \qquad \mathbb{P}(|X-\mathbb{E}[X]| \geq t) \leq \frac{\mathrm{Var}(X)}{t^2}.$$

Interprétation. C'est la même que dans le cas discret. Si on pose $t = s\sigma(X)$, l'inégalité de Tchebychev se réécrit pour tout $s > 0$:

$$\mathbb{P}(|X - \mathbb{E}[X]| \geq s\sigma(X)) \leq \frac{1}{s^2}.$$

Si on voit l'écart-type $\sigma(X)$ comme une unité d'écart, ceci dit que la probabilité qu'une variable s'éloigne de plus de s unités d'écart de sa moyenne est inférieure à $\frac{1}{s^2}$.

Remarque. Dans le cas discret, tout a été démontré facilement grâce à la seule théorie des séries numériques. Dès qu'on passe au cas absolument continu, ça se corse. En particulier, la plupart des résultats énoncés dans ce chapitre ont été prouvés sous des hypothèses inutilement restrictives sur la densité f. Ceci vient du fait que le bon cadre théorique pour traiter des probabilités en toute généralité est celui de l'intégrale de Lebesgue, laquelle permet d'unifier et d'englober les deux situations, mais dépasse largement le cadre de ce cours introductif.

3.4 Lois usuelles

Comme dans le cas discret, il existe un certain nombre de lois classiques pour les variables à densité. Nous en détaillons ici quelques-unes.

3.4.1 Loi uniforme

La loi uniforme sert à préciser ce qu'on entend par un énoncé du type : "On tire un point au hasard entre 0 et 1". C'est l'équivalent continu de la loi uniforme vue dans le cas discret.

Définition 3.7 (Loi uniforme)
Soit a et b deux réels, avec $a < b$. On dit que X suit une loi uniforme sur l'intervalle $[a, b]$, noté $X \sim \mathcal{U}_{[a,b]}$, si X admet pour densité $f(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x)$.

Remarque. On peut tout aussi bien ouvrir ou fermer les crochets au bord de l'intervalle.

Espérance et variance se calculent alors sans problème, ceci a d'ailleurs déjà été fait en Section 3.3 dans le cas d'une loi uniforme sur $[0, 1]$.

Proposition 3.4 (Moments d'une loi uniforme)
Si X suit une loi uniforme sur $[a, b]$, alors :

$$\mathbb{E}[X] = \frac{a+b}{2} \qquad \& \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}.$$

L'espérance de X correspond donc au milieu du segment $[a, b]$, ce qui n'a rien d'étonnant si l'on pense à l'espérance comme à une valeur moyenne. La fonction de répartition ne pose pas non plus de difficultés.

Proposition 3.5 (Fonction de répartition d'une loi uniforme)
Si X suit une loi uniforme sur $[a, b]$, alors sa fonction de répartition F est :

$$F(x) = \begin{cases} 0 & si\ x \leq a \\ \frac{x-a}{b-a} & si\ a \leq x \leq b \\ 1 & si\ x \geq b \end{cases}$$

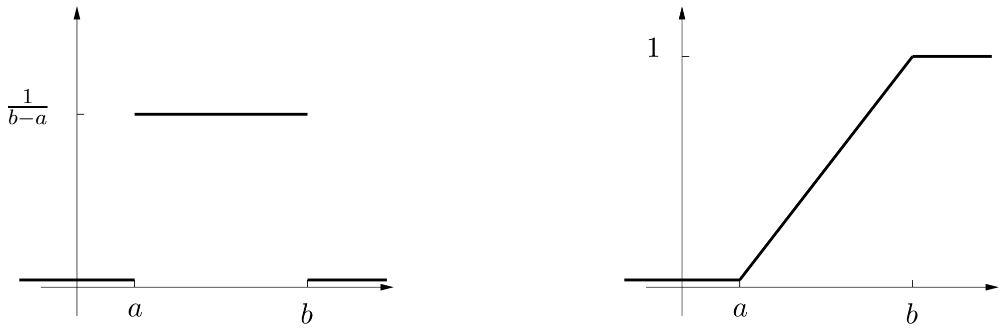


FIGURE 3.4 – Densité et fonction de répartition de la loi uniforme sur $[a, b]$.

Remarque. Densité et fonction de répartition de la loi uniforme sur $[a, b]$ sont représentées figure 3.4.

Universalité de la loi uniforme. Soit X une variable aléatoire de fonction de répartition F : $\mathbb{R} \to]0,1[$ bijective (donc strictement croissante) et notons $F^{-1} :]0,1[\to \mathbb{R}$ son inverse (strictement croissante elle aussi). Tirons maintenant une variable uniforme U sur $]0,1[$ à l'aide d'une calculatrice ou d'un logiciel (fonction usuellement appelée `rand`), et calculons $X = F^{-1}(U)$. Question : quelle est la loi de X ? Pour y répondre, il suffit de calculer sa fonction de répartition F_X. Or pour tout réel x, on a :

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(F(X) \leq F(x)) = \mathbb{P}(U \leq F(x)) = F_U(F(x)) = F(x),$$

puisque la fonction de répartition d'une loi uniforme est l'identité entre 0 et 1. Bilan des courses : X a pour fonction de répartition F. Ainsi, partant de la simulation d'une loi uniforme, on peut simuler une variable aléatoire ayant une loi arbitraire, si tant est que sa fonction de répartition soit facilement inversible. Cette astuce est abondamment utilisée dans les logiciels de calcul scientifique.

Exemple : Simulation d'une variable de Cauchy. On veut simuler une variable X distribuée selon une loi de Cauchy, c'est-à-dire de densité $f(x) = \frac{1}{\pi(1+x^2)}$. Sa fonction de répartition vaut donc pour tout réel x :

$$F(x) = \int_{-\infty}^{x} \frac{1}{\pi(1+t^2)} dt = \frac{1}{\pi}[\arctan t]_{-\infty}^{x} = \frac{1}{2} + \frac{1}{\pi} \arctan x.$$

Donc pour tout $u \in]0,1[$, on a :

$$u = F(x) \iff u = \frac{1}{2} + \frac{1}{\pi} \arctan x \iff x = \tan(\pi(u - 1/2)),$$

ce qui est exactement dire que pour tout $u \in]0,1[$, $F^{-1}(u) = \tan(\pi(u - 1/2))$. Il suffit donc de simuler $U \sim \mathcal{U}_{]0,1[}$ et de calculer $X = \tan(\pi(U - 1/2))$ pour obtenir une variable X ayant une loi de Cauchy.

3.4.2 Loi exponentielle

La loi exponentielle est la version en temps continu de la loi géométrique. De fait, elle intervient souvent dans la modélisation des phénomènes d'attente.

Définition 3.8 (Loi exponentielle)
Soit λ un réel strictement positif. On dit que X suit une loi exponentielle de paramètre λ, noté $X \sim \mathcal{E}(\lambda)$, si X admet pour densité :

$$f(x) = \lambda e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}}.$$

Une variable exponentielle est donc à valeurs dans $[0, +\infty[$. Plus λ est grand et plus la variable a des chances de prendre des valeurs proches de 0. Ceci se reflète dans l'expression de son espérance.

Proposition 3.6 (Moments d'une loi exponentielle)
Si X suit une loi exponentielle de paramètre λ, alors :

$$\mathbb{E}[X] = \frac{1}{\lambda} \qquad \& \qquad \mathrm{Var}(X) = \frac{1}{\lambda^2}.$$

Preuve. Voir le corrigé de l'exercice 3.4.

■

Généralisation. Grâce à une récurrence elle aussi basée sur une intégration par parties, il est facile d'exprimer le moment d'ordre n d'une loi exponentielle :

$$\forall n \in \mathbb{N}^* \qquad \mathbb{E}[X^n] = \int_0^{+\infty} x^n \lambda e^{-\lambda x} dx = \frac{n!}{\lambda^n}.$$

Proposition 3.7 (Fonction de répartition d'une loi exponentielle)
Si X suit une loi exponentielle de paramètre λ, alors sa fonction de répartition F est :

$$F(x) = \begin{cases} 0 & si\ x \leq 0 \\ 1 - e^{-\lambda x} & si\ x \geq 0 \end{cases}$$

Densité et fonction de répartition de la loi exponentielle de paramètre $\lambda = 4$ sont représentées figure 3.5.

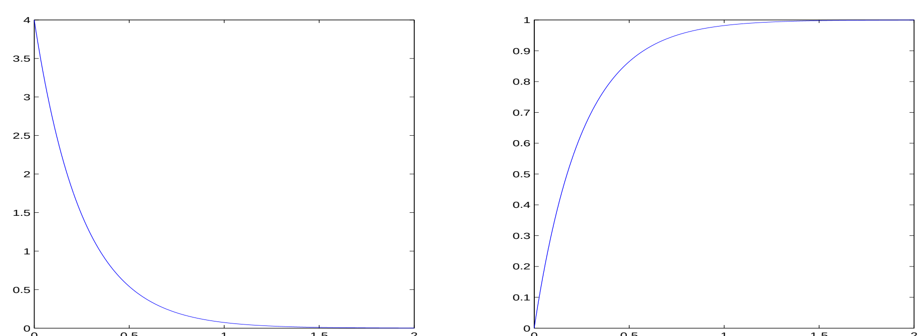


FIGURE 3.5 – Densité et fonction de répartition de la loi exponentielle de paramètre $\lambda = 4$.

Remarque. Cette expression permet de voir que la médiane d'une loi exponentielle n'est pas égale à sa moyenne, puisque :

$$\mathbb{P}(X > \mathbb{E}[X]) = \mathbb{P}(X > 1/\lambda) = 1 - \mathbb{P}(X \leq 1/\lambda) = 1 - F(1/\lambda) = \frac{1}{e} \approx 0,37 < 0,5.$$

Ainsi une variable exponentielle a plus de chances de tomber en dessous de sa moyenne qu'au-dessus. La médiane m est quant à elle la solution de :

$$F(m) = \frac{1}{2} \Leftrightarrow 1 - e^{-\lambda m} = \frac{1}{2} \Leftrightarrow m = \frac{\ln 2}{\lambda}$$

On parle parfois de demi-vie plutôt que de médiane (cf. exercice III, sujet de décembre 2010).

Proposition 3.8 (Minimum de lois exponentielles)
Soit n variables indépendantes $X_1, \dots, X_n$ suivant des lois exponentielles de paramètres respectifs $\lambda_1, \dots, \lambda_n$. Alors la variable aléatoire $X = \min(X_1, \dots, X_n)$ suit elle-même une loi exponentielle, plus précisément

$$X = \min(X_1, \dots, X_n) \sim \mathcal{E}(\lambda_1 + \dots \lambda_n).$$

Preuve. Voir le corrigé de l'exercice 3.10.

■

Proposition 3.9 (Absence de mémoire)
Soit X une variable aléatoire suivant une loi exponentielle de paramètre λ. Alors pour tout couple (x, t) de réels positifs, nous avons $\mathbb{P}(X > x + t | X > x) = \mathbb{P}(X > t)$. Autrement dit, la loi exponentielle n'a pas de mémoire.

Preuve. Voir le corrigé de l'exercice 3.5.

■

Ces deux dernières propriétés on un goût de déjà-vu : ce sont en effet les mêmes que pour la loi géométrique. Le lien entre lois géométrique et exponentielle est formalisé en exercice 3.16 : on montre que si $X \sim \mathcal{E}(\lambda)$ alors en notant $Y = \lceil X \rceil$ la variable égale à sa partie entière supérieure, on a $Y \sim \mathcal{G}(1 - e^{-\lambda})$.

3.4.3 Loi normale

La loi normale est sans conteste la loi la plus importante des probabilités et des statistiques. Ce rôle prépondérant est dû au Théorème Central Limite, dont nous dirons un mot en fin de section, mais dont l'exposé détaillé dépasse le cadre de ce cours.

Définition 3.9 (Loi normale)
Soit m et σ deux réels, avec $\sigma > 0$. On dit que X suit une loi normale, ou gaussienne, de paramètres m et σ^2, noté $X \sim \mathcal{N}(m, \sigma^2)$, si X admet pour densité :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

En particulier, si $m = 0$ et $\sigma = 1$, on dit que X suit une loi normale centrée réduite.

Des exemples de densités de lois normales, appelées aussi courbes en cloches, sont donnés figure 3.6.

Remarques :

1. Il n'existe pas d'expression analytique simple pour une primitive de la fonction e^{-x^2}. Nous admettrons le résultat suivant :

$$\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}, \tag{3.1}$$

ce qui assure au moins que la densité de la loi normale centrée réduite est bien une densité !

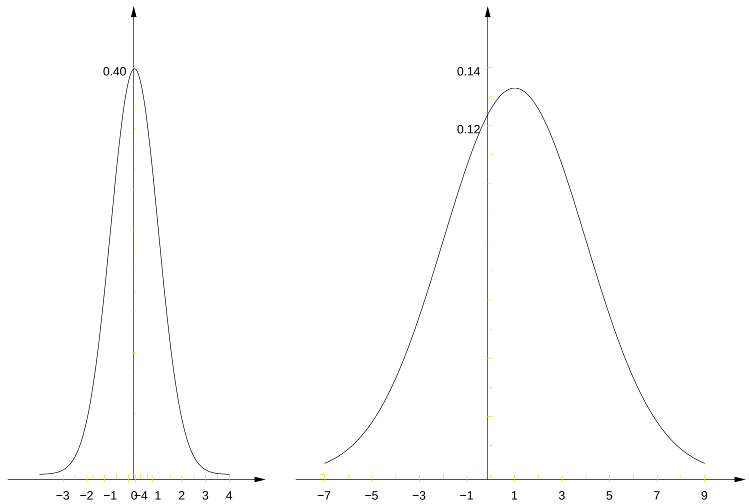


FIGURE 3.6 – Densités des lois normales $\mathcal{N}(0,1)$ (à gauche) et $\mathcal{N}(2,9)$ (à droite).

2. On peut étendre la définition ci-dessus au cas où $\sigma = 0$ en considérant qu'alors la variable aléatoire X n'est plus aléatoire, mais déterministe, et ne prend que la valeur m.

Les lois normales sont stables par transformation affine, comme nous allons maintenant le montrer.

Propriétés 3.4 (Loi normale & transformation affine)

1. *Soit $X \sim \mathcal{N}(m,\sigma^2)$ une variable gaussienne, a et b deux réels, a étant non nul, alors la variable $Y = aX+b$ suit elle aussi une loi normale. Plus précisément $Y \sim \mathcal{N}(am+b, a^2\sigma^2)$.*
2. *En particulier, si $X \sim \mathcal{N}(m,\sigma^2)$, la variable $Y = \frac{X-m}{\sigma}$ suit une loi normale centrée réduite.*
3. *Réciproquement, si $X \sim \mathcal{N}(0,1)$, alors $Y = \sigma X + m \sim \mathcal{N}(m,\sigma^2)$.*

Preuve. Il suffit d'appliquer la formule de changement de variable vue en Proposition 3.2 à la fonction $\varphi(x) = ax+b$. Sa dérivée est constante égale à a et son inverse est $\varphi^{-1}(y) = (y-b)/a$, donc la formule

$$f_Y(y) = \frac{f_X(\varphi^{-1}(y))}{|\varphi'(\varphi^{-1}(y))|}$$

donne dans notre cas de figure :

$$f_Y(y) = \frac{1}{|a|\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{\left(\frac{y-b}{a} - m\right)^2}{2\sigma^2}\right\},$$

qui s'écrit encore :

$$f_Y(y) = \frac{1}{\sqrt{2\pi(a\sigma)^2}} e^{-\frac{(y-(am+b))^2}{2(a\sigma)^2}},$$

où l'on reconnaît la densité d'une loi normale $\mathcal{N}(am+b, a^2\sigma^2)$. Les deux derniers points en découlent directement.

■

L'intérêt de ce résultat est de ramener tous les calculs sur les lois normales à des calculs sur la loi normale centrée réduite.

Proposition 3.10 (Moments d'une loi normale)
Si X suit une loi normale de paramètres (m, σ^2), alors :

$$\mathbb{E}[X] = m \qquad \& \qquad \mathrm{Var}(X) = \sigma^2.$$

Preuve. Soit $X \sim \mathcal{N}(0,1)$, alors par définition de l'espérance :

$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} \frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Cette intégrale est doublement généralisée mais converge en $-\infty$ et $+\infty$, puisqu'en ces deux bornes il suffit par exemple de constater que :

$$xe^{-\frac{x^2}{2}} = o(x^{-2}).$$

Ainsi $\mathbb{E}[X]$ est l'intégrale d'une fonction impaire sur un domaine d'intégration symétrique par rapport à 0, donc $\mathbb{E}[X] = 0$. Voyons maintenant sa variance. Par définition :

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X^2] = \int_{-\infty}^{+\infty} \frac{x^2}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx.$$

Cette fois encore, il n'y a aucun problème de convergence de l'intégrale. Une intégration par parties donne :

$$\mathrm{Var}(X) = \left[-\frac{x}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx,$$

qui, grâce à la formule (3.1), donne bien $\mathrm{Var}(X) = 1$. Passons maintenant au cas général : si $X \sim \mathcal{N}(0,1)$ et si $Y = \sigma X + m$, alors d'après ce qui vient d'être vu $Y \sim \mathcal{N}(m, \sigma^2)$, et d'après les propriétés classiques sur espérance et variance, on a d'une part :

$$\mathbb{E}[Y] = \mathbb{E}[\sigma X + m] = \sigma \mathbb{E}[X] + m = m,$$

et d'autre part :

$$\mathrm{Var}(Y) = \mathrm{Var}(\sigma X + m) = \sigma^2 \mathrm{Var}(X) = \sigma^2.$$

■

Généralisation. Il est assez facile d'obtenir tous les moments d'une loi normale $\mathcal{N}(0,1)$. Les moments impairs ne nécessitent aucun calcul puisque par intégration d'une fonction impaire sur $]-\infty, +\infty[$, il est clair que $\forall n \in \mathbb{N}$, $\mathbb{E}[X^{2n+1}] = 0$. Pour les moments d'ordres pairs, une récurrence basée sur une intégration par parties permet d'aboutir à la formule suivante :

$$\forall n \in \mathbb{N} \qquad \mathbb{E}[X^{2n}] = \frac{(2n)!}{2^n \ n!}$$

En raison de l'absence de primitive simple de e^{-x^2}, la fonction de répartition n'a pas d'expression plus synthétique que sa formulation intégrale. Si $Y \sim \mathcal{N}(m, \sigma^2)$, alors sa fonction de répartition est :

$$F(y) = \mathbb{P}(Y \leq y) = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-m)^2}{2\sigma^2}} dt.$$

On peut toujours par contre se ramener à une loi normale centrée réduite puisque :

$$F(y) = \mathbb{P}(Y \leq y) = \mathbb{P}\left(\frac{Y-m}{\sigma} \leq \frac{y-m}{\sigma} \right) = \Phi\left(\frac{y-m}{\sigma} \right),$$

si l'on convient de noter Φ la fonction de répartition de la loi normale $\mathcal{N}(0,1)$, c'est-à-dire que si $X \sim \mathcal{N}(0,1)$:

$$\Phi(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Puisqu'elle n'admet pas d'expression analytique élémentaire, la fonction Φ est tabulée : dans le tableau en Annexe A.2 sont rassemblées les valeurs de $\Phi(x)$ pour $0 \leq x \leq 3,29$. Notons que par symétrie de la densité de la loi normale centrée réduite par rapport à 0, les valeurs de $\Phi(-x)$ s'en déduisent :

$$\Phi(-x) = \mathbb{P}(X \leq -x) = \mathbb{P}(X \geq x) = 1 - \mathbb{P}(X < x) = 1 - \Phi(x),$$

autrement dit la courbe de Φ admet le point $(0, 1/2)$ comme centre de symétrie. Ceci est illustré figure 3.7.

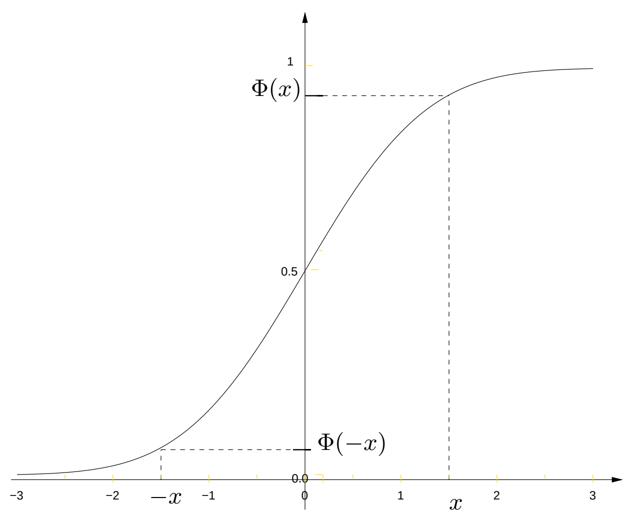


FIGURE 3.7 – Fonction de répartition Φ d'une loi normale $\mathcal{N}(0,1)$ et relation : $\Phi(-x) = 1 - \Phi(x)$.

Concentration. Supposons qu'on tire des nombres selon une loi normale $\mathcal{N}(m, \sigma^2)$, par exemple avec un ordinateur. Alors plus l'écart-type σ est faible et plus on a des chances d'obtenir des résultats autour de la moyenne m : 68% de tomber à distance inférieure ou égale à σ, 95% de tomber à distance inférieure ou égale[1] à 2σ, 99,7% de tomber à distance inférieure ou égale à 3σ. Ceci est illustré figure 3.8.

Exemple : le test du Q.I. La distribution des résultats au test de la WAIS (Weschler Adult Intelligence Scale), ou test du Q.I. pour adultes, est gaussienne et celui-ci a été calibré pour que sa moyenne soit égale à 100 et son écart-type égal à 15. Il y a donc 68% de la population adulte dont le quotient intellectuel est compris entre 85 et 115.

Terminologie et Théorème Central Limite. La loi normale tire son nom de ce qu'elle apparaît de façon "naturelle" ou "normale" dans de très nombreux phénomènes. Ceci est dû au Théorème Central Limite. En voici la version la plus simple : si $(X_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes et identiquement distribuées (en abrégé i.i.d.) admettant une variance, alors en

1. Un encadrement plus précis pour l'intervalle de confiance à 95% est $[m - 1.96\sigma; m + 1.96\sigma]$.

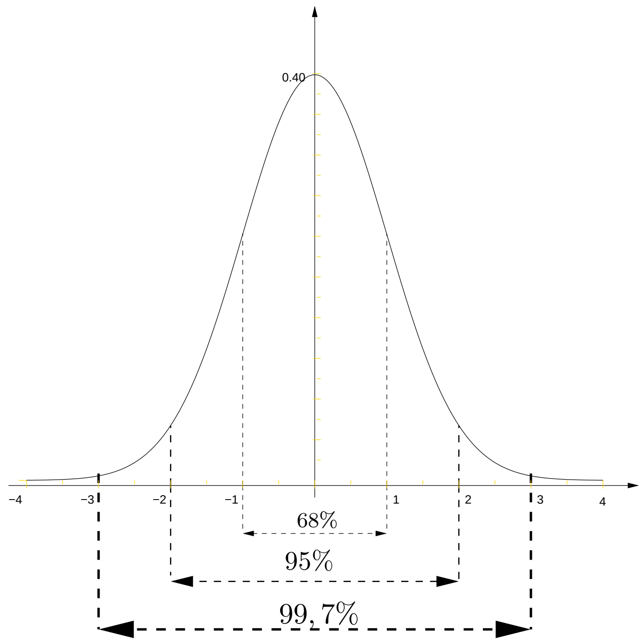


FIGURE 3.8 – Concentration autour de la moyenne d'une loi $\mathcal{N}(0,1)$.

notant $S_n = X_1 + \cdots + X_n$ leurs sommes partielles, on a la convergence en loi suivante :

$$\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\ \mathrm{Var}(X_1)}} \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1),$$

c'est-à-dire que pour tout intervalle (a,b) de $\overline{\mathbb{R}}$, on a :

$$\mathbb{P}\left(a \leq \frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\ \mathrm{Var}(X_1)}} \leq b\right) \xrightarrow[n\to+\infty]{} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}\, dx.$$

Autrement dit, la somme d'un grand nombre de variables aléatoires i.i.d. se comporte comme une loi normale. Dit grossièrement et quitte à choquer les puristes, on peut considérer que si n est "assez grand", S_n suit quasiment une loi normale de moyenne $n\mathbb{E}[X_1]$ et d'écart-type $\sqrt{n\ \mathrm{Var}(X_1)}$: $S_n \approx \mathcal{N}(n\mathbb{E}[X_1], n\ \mathrm{Var}(X_1))$.

Universalité. L'aspect remarquable de ce résultat tient bien sûr au fait que la loi commune des X_n peut être n'importe quoi ! Celle-ci peut aussi bien être discrète qu'absolument continue, mixte ou singulière. La seule chose requise est l'existence de la variance. Avec la Loi des Grands Nombres, ce résultat peut être considéré comme le plus important en probabilités et statistiques.

Exemple. Revenons à l'exercice 2.21. Rappel des épisodes précédents : après 3600 jets d'un dé équilibré, la question était d'évaluer la probabilité p que le nombre S de 1 apparus soit compris entre 480 et 720. L'expression exacte était donnée par la somme de termes binomiaux et une minoration avait été fournie par l'inégalité de Tchebychev : $p \geq 0,965$. Or nous sommes dans le cadre typique d'application du théorème central limite, avec $n = 3600$ et les X_i ayant pour loi commune la distribution de Bernoulli $\mathcal{B}(1/6)$. Ainsi :

$$\frac{S_n - n\mathbb{E}[X_1]}{\sqrt{n\ \mathrm{Var}(X_1)}} = \frac{S-600}{\sqrt{500}} \approx \mathcal{N}(0,1).$$

Et on cherche :

$$p = \mathbb{P}(480 < S < 720) = \mathbb{P}\left(\frac{-120}{\sqrt{500}} < \frac{S-600}{\sqrt{500}} < \frac{120}{\sqrt{500}}\right) = 2\Phi(5.36) - 1,$$

ce qui donne $p \approx 1 - 8.10^{-8}$. Le calcul sur machine de la probabilité p via son expression exacte donne en fait :

$$\mathbb{P}(480 < S < 720) = \sum_{n=481}^{719} \mathbb{P}(S = n) = \sum_{n=481}^{719} \binom{3600}{n} \left(\frac{1}{6}\right)^n \left(\frac{5}{6}\right)^{3600-n} \approx 1 - 11.10^{-8}.$$

Ceci montre que l'approximation gaussienne donnée par le théorème central limite est excellente. La minoration par l'inégalité de Tchebychev était par contre très pessimiste : il n'y a concrètement à peu près aucune chance que le nombre de 1 ne soit pas compris entre 480 et 720.

Remarque. Nous avons ainsi obtenu une approximation de la loi binomiale $\mathcal{B}(n,p)$ par une loi normale $\mathcal{N}(np, np(1-p))$ lorsque n est grand. Celle-ci ne fonctionne cependant que si p n'est pas trop petit, plus précisément il ne faut pas que p soit de l'ordre de $1/n$. Dans cette situation, comme expliqué en Section 2.5.5, c'est l'approximation par une loi de Poisson $\mathcal{P}(np)$ qui est pertinente (cf. exercice 3.26).

3.5 Exercices

Exercice 3.1 (Espérance et variance d'une loi uniforme)
Soit X une variable aléatoire qui suit une loi uniforme sur le segment $[0,1]$.

1. Calculer sa moyenne $\mathbb{E}[X]$ et sa variance $\mathrm{Var}(X)$.
2. De façon générale, calculer $\mathbb{E}[X^n]$, moment d'ordre n de X.
3. Soit a et b deux réels tels que $a < b$. Comment définiriez-vous la densité d'une variable aléatoire X uniforme sur le segment $[a,b]$? Donner alors $\mathbb{E}[X]$ et $\mathrm{Var}(X)$.

Exercice 3.2 (Loi de Cauchy)
On dit que X suit une loi de Cauchy de paramètre 1 si X admet pour densité f avec :

$$\forall x \in \mathbb{R} \qquad f(x) = \frac{c}{1+x^2}.$$

1. Déterminer c pour que f soit bien une densité.
2. Calculer et représenter la fonction de répartition F de X.
3. Montrer que X n'a pas d'espérance.
4. Soit Y une variable aléatoire uniforme sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$. Déterminer la loi de $X = \tan Y$ (on pourra passer par sa fonction de répartition).

Exercice 3.3 (Densités parabolique et circulaire)
Soit X une variable aléatoire de densité $f(x) = c(1-x^2)\mathbb{1}_{\{-1<x<1\}}$.

1. Déterminer c pour que f soit bien une densité de probabilité.
2. Quelle est la fonction de répartition de X ?
3. Calculer $\mathbb{E}[X]$ et $\mathrm{Var}(X)$.
4. Mêmes questions avec la densité $f(x) = c\sqrt{1-x^2}\mathbb{1}_{\{-1<x<1\}}$. Indication : on pourra penser au changement de variable $x = \cos t$.

Exercice 3.4 (Loi exponentielle)
Soit $\lambda > 0$ fixé. On dit que X suit une loi exponentielle de paramètre λ si X admet pour densité $f(x) = \lambda e^{-\lambda x}\mathbb{1}_{\{x\geq 0\}}$. On note alors $X \sim \mathcal{E}(\lambda)$.

1. Représenter f. Vérifier que f est bien une densité.
2. Calculer et représenter la fonction de répartition F.
3. Calculer espérance et variance de X.
4. La durée de vie T en années d'une télévision suit une loi de densité $f(t) = \frac{1}{8}e^{-\frac{t}{8}}\mathbb{1}_{\{t\geq 0\}}$.
   (a) Quelle est la durée de vie moyenne d'une telle télévision ? Et l'écart-type de cette durée de vie ?
   (b) Calculer la probabilité que votre télévision ait une durée de vie supérieure à 8 ans.

Exercice 3.5 (Absence de mémoire)
Soit X une variable qui suit une loi exponentielle de paramètre λ.

1. Calculer $\mathbb{P}(X > t)$ pour tout $t \geq 0$.
2. En déduire que la loi exponentielle a la propriété d'absence de mémoire, c'est-à-dire que :
$$\forall (x,t) \in \mathbb{R}^+ \times \mathbb{R}^+ \qquad \mathbb{P}(X > x+t | X > x) = \mathbb{P}(X > t).$$
3. Application : la durée de vie T en années d'une télévision suit une loi exponentielle de moyenne 8 ans. Vous possédez une telle télévision depuis 2 ans, quelle est la probabilité que sa durée de vie soit encore d'au moins 8 ans à partir de maintenant ?

Exercice 3.6 (Durée de vie)
Un appareil comporte six composants de même modèle, tous nécessaires à son fonctionnement. La densité de la durée de vie T d'un composant est donnée par $f(t) = \frac{t}{16}e^{-\frac{t}{4}}\mathbb{1}_{\{t\geq 0\}}$, l'unité de temps étant l'année.

1. Vérifier que f est bien une densité de probabilité.
2. Calculer $\mathbb{E}[T]$ et $\mathrm{Var}(T)$.
3. Quelle est la probabilité qu'un composant fonctionne durant au moins six ans à partir de sa mise en marche ? En déduire la probabilité que l'appareil fonctionne durant au moins six ans à partir de sa mise en marche.

Exercice 3.7 (Loi de Pareto)
La variable aléatoire T, représentant la durée de vie en heures d'un composant électronique, a pour densité $f(t) = \frac{10}{t^2}\mathbb{1}_{\{t>10\}}$.

1. Calculer $\mathbb{P}(T > 20)$.
2. Quelle est la fonction de répartition de T ? La représenter.
3. Quelle est la probabilité que parmi 6 composants indépendants, au moins 3 d'entre eux fonctionnent durant au moins 15 heures.

Exercice 3.8 (Tirages uniformes sur un segment)
Soit X un point au hasard sur le segment $[0,1]$, c'est-à-dire que $X \sim \mathcal{U}_{[0,1]}$.

1. Quelle est la probabilité que X soit supérieur à $3/4$?
2. Quelle est la probabilité que X soit supérieur à $3/4$, sachant qu'il est supérieur à $1/3$?
3. Le point X définit les deux segments $[0,X]$ et $[X,1]$. Quelle est la probabilité pour que le rapport entre le plus grand et le plus petit des deux segments soit supérieur à 4 ?
4. On tire deux points X et Y au hasard sur le segment $[0,1]$, indépendamment l'un de l'autre.

(a) Quelle est la probabilité que le plus petit des deux nombres soit supérieur à $1/3$?

(b) Quelle est la probabilité que le plus grand des deux nombres soit supérieur à $3/4$, sachant que le plus petit des deux est supérieur à $1/3$?

Exercice 3.9 (Problèmes de densité)

1. Soit X une variable aléatoire qui suit une loi uniforme sur le segment $[0,1]$ et soit $Y = 1-X$. Donner la fonction de répartition de Y. En déduire la densité de Y. Est-ce que la variable aléatoire $Z = (X+Y)$ admet une densité ?
2. On construit une variable aléatoire X en commençant par lancer une pièce équilibrée : si on obtient Pile, alors $X = 1$; si on obtient Face, X est le résultat d'un tirage uniforme dans le segment $[0,1]$. Donner la fonction de répartition de X.

Exercice 3.10 (Minimum d'exponentielles)

1. On considère deux variables aléatoires indépendantes X_1 et X_2 exponentielles de paramètres respectifs λ_1 et λ_2. Soit $Y = \min(X_1, X_2)$ le minimum de ces deux variables.
   (a) Pour tout réel y, calculer $\mathbb{P}(X_1 > y)$.
   (b) En déduire $\mathbb{P}(Y > y)$, puis la fonction de répartition F de la variable Y.
   (c) En déduire que Y suit une loi exponentielle de paramètre $\lambda_1 + \lambda_2$.
2. Deux guichets sont ouverts à une banque : le temps de service au premier (respectivement second) guichet suit une loi exponentielle de moyenne 20 (respectivement 30) minutes. Alice et Bob arrivent ensemble à la banque : Alice choisit le guichet 1, Bob le 2. En moyenne, au bout de combien de temps sort le premier ?
3. En moyenne, combien de temps faut-il pour que les deux soient sortis ? (Indication : le max de deux nombres, c'est la somme moins le min.)

Exercice 3.11 (Think Tank)

Dans une station-service, la demande hebdomadaire en essence, en milliers de litres, est une variable aléatoire X de densité $f(x) = c(1-x)^4 \mathbb{1}_{\{0<x<1\}}$. Ce modèle suppose en particulier que la demande ne dépasse jamais 1000 litres.

1. Déterminer c pour que f soit une densité. Représenter f.
2. Calculer la fonction de répartition F.
3. La station est approvisionnée une fois par semaine. Quelle capacité doit avoir le réservoir de cette station pour que la probabilité d'épuiser l'approvisionnement soit inférieure à 10^{-5} ?

Exercice 3.12 (Loi polynomiale)

Soit X une variable aléatoire de densité $f(x) = c(x+x^2)$ sur l'intervalle $[0,1]$.

1. Déterminer c pour que f soit effectivement une densité.
2. Calculer la fonction de répartition F de X. Donner l'allure de F.
3. Calculer l'espérance et l'écart-type de X.

Exercice 3.13 (Ambulance et accidents)

Une station d'ambulances se situe au kilomètre 30 d'une route de 100 kms de long. Les accidents sont supposés arriver uniformément sur cette route. L'ambulance roule à 100 km/h pour intervenir sur le lieu d'un accident. Notons T le temps écoulé (en minutes) entre l'appel à la station et l'arrivée de l'ambulance sur le lieu de l'accident.

1. Quelles valeurs peut prendre la variable aléatoire T ?
2. Que vaut $\mathbb{P}(T > 30)$?
3. Plus généralement, calculer $\mathbb{P}(T > t)$ en fonction de t.

4. Déterminer la densité de T, sa moyenne et sa variance.

Exercice 3.14 (Minimum d'uniformes)
Soit n variables aléatoires $U_1, \ldots, U_n$ indépendantes et de même loi uniforme sur $[0,1]$. On considère la variable $X = \min(U_1, \ldots, U_n)$.

1. Que vaut $\mathbb{P}(U > t)$ lorsque $U \sim \mathcal{U}_{[0,1]}$ et $t \in [0,1]$?
2. Calculer la fonction de répartition F de la variable X.
3. En déduire la densité et l'espérance de X.

Exercice 3.15 (Racines d'un trinôme aléatoire)
La variable U suit une loi uniforme sur $[0,5]$. On considère le trinôme aléatoire $P(x) = 4x^2 + 4Ux + U + 2$.

1. Donner l'expression de son discriminant en fonction de U.
2. Etudier le signe de la fonction $D(u) = u^2 - u - 2$ sur $\mathbb{R}$.
3. En déduire la probabilité pour que P ait deux racines réelles distinctes.

Exercice 3.16 (Lien entre lois exponentielle et géométrique)
Soit X une variable aléatoire suivant une loi exponentielle $\mathcal{E}(1)$, et $Y = \lceil X \rceil$ la variable égale à sa partie entière supérieure (c'est-à-dire que $\lceil 2.8 \rceil = 3$ et $\lceil 4 \rceil = 4$).

1. Quelles valeurs peut prendre Y ? Avec quelles probabilités ? Quelle loi reconnaissez-vous ? En déduire $\mathbb{E}[Y]$ et $\mathrm{Var}(Y)$.
2. Soit alors $Z = Y - X$. Dans quel intervalle Z prend-elle ses valeurs ? Déterminer sa fonction de répartition F (elle fait intervenir une série).
3. En déduire que sa densité vaut
$$f(z) = \frac{e^z}{e-1} \mathbb{1}_{[0,1]}(z).$$
4. Préciser $\mathbb{E}[Z]$.

Exercice 3.17 (Moments d'une loi normale)
Pour tout $n \in \mathbb{N}$, on note :
$$I_n = \int_{-\infty}^{+\infty} x^n e^{-\frac{x^2}{2}}\, dx.$$

1. Déterminer I_0 et I_1.
2. Montrer que, pour tout $n \in \mathbb{N}$, on a : $I_{n+2} = (n+1)I_n$.
3. Donner alors I_{2n+1} pour tout $n \in \mathbb{N}$. Pouvait-on prévoir ce résultat sans calculs ?
4. Déterminer I_{2n} pour tout $n \in \mathbb{N}$.
5. Soit X une variable aléatoire gaussienne de moyenne 1 et de variance unité. Déterminer $\mathbb{E}[X^4]$.

Exercice 3.18 (Vitesse d'une molécule)
La vitesse d'une molécule au sein d'un gaz homogène en état d'équilibre est une variable aléatoire de densité :
$$f(x) = ax^2 e^{-\frac{mx^2}{2kT}} \mathbb{1}_{\{x \geq 0\}},$$
où k est la constante de Boltzmann, T la température absolue et m la masse de la molécule. Déterminer a en fonction de ces paramètres.

Exercice 3.19 (Loi log-normale)
Soit m et σ deux réels, avec $\sigma > 0$. On dit que X suit une loi log-normale, ou de Galton, de paramètres (m, σ^2), noté $X \sim \mathcal{LN}(m, \sigma^2)$, si $Y = \ln X$ suit une loi normale $\mathcal{N}(m, \sigma^2)$. Cette loi intervient lors de la multiplication d'un grand nombre de variables indépendantes et positives. En linguistique, elle sert à modéliser le nombre de mots dans une phrase.

1. Supposons que $X \sim \mathcal{LN}(0, 1)$. Exprimer sa fonction de répartition F à l'aide de la fonction de répartition Φ de la loi normale centrée réduite.
2. En déduire que sa densité est :
$$f(x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{\ln^2 x}{2}} \mathbb{1}_{\{x>0\}}.$$
Représenter f.
3. Montrer que son espérance vaut $\mathbb{E}[X] = \sqrt{e}$ et sa variance $\mathrm{Var}(X) = e(e-1)$.
4. Un tas de sable est composé de grains homogènes sphériques. La diamètre X d'un grain suit la loi $\mathcal{LN}(-0,5; 0,09)$, l'unité étant le millimètre. On passe le tas au crible d'un tamis dont les trous sont circulaires, de diamètre 0,5 mm. Quelle est la proportion de grains de sable passant à travers le tamis ?

Exercice 3.20 (La Belle de Fontenay)
On suppose que la masse X d'une pomme de terre Belle de Fontenay suit une loi normale de moyenne $m = 200$ g et d'écart-type $\sigma = 70$ g. Quelle est la probabilité qu'une pomme de terre :

1. pèse plus de 250 grammes ?
2. pèse moins de 180 grammes ?
3. ait une masse comprise entre 190 et 210 grammes ?

Exercice 3.21 (Quantile et variance)

1. Supposons que X suive une loi normale de moyenne 12 et de variance 4. Trouver la valeur q telle que $\mathbb{P}(X > q) = 0,1$.
2. Soit $X \sim \mathcal{N}(5, \sigma^2)$. Déterminer la variance σ^2 telle que $\mathbb{P}(X > 9) = 0,2$.

Exercice 3.22 (Répartition des tailles)
La taille d'un homme âgé de 25 ans suit une loi normale de moyenne 175 et d'écart-type 6.

1. Quel est le pourcentage d'hommes ayant une taille supérieure à 1m85 ?
2. Parmi les hommes mesurant plus de 1m80, quelle proportion mesure plus de 1m92 ?

Exercice 3.23 (Choix de machine)
La longueur des pièces (en mm) produites par une machine A (resp. B) suit une loi normale $\mathcal{N}(8; 4)$ (resp. $\mathcal{N}(7,5; 1)$). Si vous voulez produire des pièces de longueurs 8 ± 1 mm, quelle machine vaut-il mieux choisir ?

Exercice 3.24 (Approximation gaussienne)
Soit X le nombre de Pile obtenus en 400 lancers d'une pièce équilibrée.

1. Grâce à l'approximation normale, estimer $\mathbb{P}(190 \leq X \leq 210)$.
2. Idem pour $\mathbb{P}(210 \leq X \leq 220)$.
3. Reprendre les questions précédentes pour une pièce biaisée où $\mathbb{P}(\text{Pile}) = 0.51$.

Exercice 3.25 (Sondage)
Deux candidats, Alice et Bob, sont en lice lors d'une élection. On note p la proportion d'électeurs pour Alice dans la population totale. Afin d'estimer p, on effectue un sondage (avec remise) auprès de n personnes. Notons X le nombre d'électeurs favorables à Alice dans cet échantillon.

1. Quelle est la loi suivie par X ?
2. Grâce à l'approximation normale, donner en fonction de n et p un intervalle où X a 95% de chances de se situer.
3. Donner un estimateur naturel $\hat{p}$ de p. Quelle est sa moyenne ?
4. Donner en fonction de n et p un intervalle où $\hat{p}$ a 95% de chances de se situer.
5. Donner un majorant de $x(1-x)$ lorsque $x \in [0,1]$. En déduire un intervalle de confiance à 95% pour p.
6. Quelle est la taille de cet intervalle lorsqu'on interroge 1000 personnes ?
7. Combien de personnes faut-il interroger pour obtenir une estimation à $\pm 2\%$?

Exercice 3.26 (Surbooking (bis))
Reprenons le contexte de l'exercice 2.22 : des études effectuées par une compagnie aérienne montrent qu'il y a une probabilité 0,05 qu'un passager ayant fait une réservation n'effectue pas le vol. Dès lors, elle vend toujours 94 billets pour ses avions à 90 places. On veut évaluer la probabilité qu'il y ait un problème à l'embarquement, c'est-à-dire qu'il y ait au plus 3 absents.

1. Estimer cette probabilité en utilisant l'approximation d'une loi binomiale par une loi normale.
2. Comparer à la vraie valeur d'une part et à la valeur obtenue par l'approximation de Poisson d'autre part. Comment expliquez-vous que l'approximation gaussienne ne marche pas ici ?

Exercice 3.27 (Queue de la gaussienne)
On appelle fonction de Marcum, ou queue de la gaussienne, la fonction définie pour tout réel x par :

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}}\, dt.$$

1. Soit X une variable aléatoire qui suit une loi normale centrée réduite $\mathcal{N}(0,1)$. Représenter la densité de X, puis $Q(x)$ sur ce même dessin. Soit F la fonction de répartition de X : donner la relation entre $F(x)$ et $Q(x)$.
2. Soit $x > 0$ fixé. Dans l'intégrale définissant $Q(x)$, effectuer le changement de variable $t = x+u$ et, tenant compte de $e^{-ux} \leq 1$, montrer qu'on a :
$$Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}}.$$
3. Pour $t \geq x > 0$, montrer que :
$$\frac{1+\frac{1}{t^2}}{1+\frac{1}{x^2}} \leq 1 \leq \frac{t}{x}.$$
4. En déduire que :
$$\frac{1}{(1+\frac{1}{x^2})\sqrt{2\pi}} \int_x^{+\infty} \left(1+\frac{1}{t^2}\right) e^{-\frac{t^2}{2}}\, dt \leq Q(x) \leq \frac{1}{x\sqrt{2\pi}} \int_x^{+\infty} t e^{-\frac{t^2}{2}}\, dt.$$
5. Calculer la dérivée de $\frac{1}{t} e^{-\frac{t^2}{2}}$. En déduire que, pour tout $x > 0$, on a :
$$\frac{1}{(1+\frac{1}{x^2})x\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq Q(x) \leq \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$
6. En déduire un équivalent de $Q(x)$ en $+\infty$.

7. Application : en communications numériques, pour une modulation binaire, les symboles transmis valent $\pm\sqrt{E_b}$, où E_b est appelée énergie moyenne par bit. Quand il transite par un canal à bruit gaussien, le signal reçu en sortie Y est égal à la somme du symbole d'entrée et d'une variable aléatoire indépendante $B \sim \mathcal{N}(0, \frac{N_0}{2})$, où N_0 est appelé puissance moyenne du bruit.
   (a) Supposons que le symbole d'entrée soit $+\sqrt{E_b}$. Donner la loi de Y en fonction de E_b et N_0. Même question si le symbole d'entrée est $-\sqrt{E_b}$.
   (b) On reçoit $y \in \mathbb{R}$ en sortie de canal, mais on ignore ce qu'était le symbole d'entrée : quelle règle simple proposez-vous pour décider si en entrée le symbole émis était a priori équiprobablement $+\sqrt{E_b}$ ou $-\sqrt{E_b}$?
   (c) Montrer que la probabilité d'erreur P_e faite avec cette règle de décision est :
$$P_e = Q\left(\sqrt{\frac{2E_b}{N_0}}\right).$$
   La quantité $\frac{E_b}{N_0}$ est appelée rapport signal à bruit et intervient très souvent en communications numériques (on l'exprime usuellement en décibels).

Exercice 3.28 (Entropie d'une variable aléatoire)
Si X est une variable aléatoire réelle admettant une densité f, on appelle entropie de X la quantité (si elle est définie) :
$$h(X) = \mathbb{E}[-\ln f(X)] = -\int_{-\infty}^{+\infty} f(x)\ln f(x)\,dx.$$
Grosso modo, l'entropie d'une variable aléatoire mesure le degré d'incertitude qu'on a sur l'issue d'un tirage de cette variable aléatoire.

1. Supposons que $X \sim \mathcal{N}(0,1)$, loi normale centrée réduite. Montrer qu'elle a pour entropie :
$$h(X) = \frac{1}{2}(1+\ln(2\pi)).$$
2. Supposons que $X \sim \mathcal{N}(m,\sigma^2)$. Montrer qu'elle a pour entropie : $h(X) = \frac{1}{2}(1+\ln(2\pi\sigma^2))$. Ainsi, au moins pour les lois normales, l'entropie est d'autant plus grande que la variance est grande. On va montrer dans la suite que, parmi les variables aléatoires de variance donnée, celles qui ont la plus grande entropie sont celles qui suivent une loi normale.
3. Soit donc $X_1 \sim \mathcal{N}(0,\sigma^2)$, dont la densité est notée φ, et X_2 une variable aléatoire centrée de densité f et de variance σ^2, c'est-à-dire que :
$$\int_{-\infty}^{+\infty} x^2 f(x)\,dx = \sigma^2.$$
   On suppose pour simplifier que f est strictement positive sur $\mathbb{R}$.
   (a) Vérifier que (sous réserve d'existence des intégrales) :
$$h(X_2) = \int_{-\infty}^{+\infty} f(x)\ln\frac{\varphi(x)}{f(x)}\,dx - \int_{-\infty}^{+\infty} f(x)\ln\varphi(x)\,dx.$$
   (b) Montrer que pour tout $x>0$, $\log x \leq x-1$. En déduire que :
$$\int_{-\infty}^{+\infty} f(x)\ln\frac{\varphi(x)}{f(x)}\,dx \leq 0.$$

(c) Montrer que :

$$-\int_{-\infty}^{+\infty} f(x)\ln\varphi(x)\,dx = \frac{1}{2}(1+\ln(2\pi\sigma^2)).$$

(d) En déduire que $h(X_2) \leq h(X_1)$.

Exercice 3.29 (Nul n'est censé ignorer la loi normale)

1. On appelle premier quartile q_1 (respectivement troisième quartile q_3) d'une variable aléatoire X à densité le réel tel que $\mathbb{P}(X \leq q_1) = 1/4$ (respectivement $\mathbb{P}(X \leq q_3) = 3/4$). Déterminer le premier et le troisième quartile d'une loi normale de moyenne 20 et d'écart-type 5.
2. Un groupe de 200 étudiants passe en début d'année un examen et les notes sont approximativement distribuées suivant une loi normale de moyenne 9 et d'écart-type 2. L'enseignant décide de faire des séances de rattrapage pour les étudiants dont les notes sont les plus faibles mais il ne peut encadrer que 30 étudiants. Quelle est la note limite permettant à un étudiant de bénéficier du rattrapage ?
3. La durée de la grossesse, en jours, est modélisée par une loi normale de moyenne 270 et de variance 100. Lors d'un procès en attribution de paternité, l'un des pères putatifs peut prouver son absence du pays sur une période allant du 290e au 240e jour avant la naissance. Quelle est la probabilité qu'il puisse être le père malgré cet alibi ?
4. Dans une université, une promotion de première année ne doit pas dépasser 200 étudiants. En se basant sur le constat que seulement un candidat accepté sur trois viendra effectivement à la rentrée, la politique de l'université est d'accepter systématiquement 500 étudiants.
   (a) Sur 500 candidats acceptés, quelle est la loi de la variable X correspondant au nombre d'étudiants effectivement présents à la rentrée ?
   (b) En utilisant l'approximation normale, estimer la probabilité qu'il y ait plus de 200 étudiants présents à la rentrée.

Exercice 3.30 (Loi bêta)

On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} c\,x(1-x) & \text{si } 0 \leq x \leq 1 \\ 0 & \text{ailleurs} \end{cases} \tag{3.1}$$

1. Evaluer la constante c pour que f soit une densité de probabilité. Représenter f.
2. Déterminer la fonction de répartition F de X. La représenter.
3. Calculer $\mathbb{P}(1/4 < X < 3/4)$.
4. Déterminer espérance et variance de X.
5. Minorer $\mathbb{P}(1/4 < X < 3/4)$ grâce à l'inégalité de Tchebychev.
6. Pour tout $n \in \mathbb{N}^*$, déterminer le moment d'ordre n de X.

Exercice 3.31 (Loi de Rayleigh)

On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} x\,e^{-\frac{x^2}{2}} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{3.1}$$

1. Vérifier que f est bien une densité de probabilité. Donner l'allure de f. On dit que X suit une loi de Rayleigh de paramètre 1.
2. Déterminer la fonction de répartition F de X. Donner son allure.

3. Déterminer la médiane de X, c'est-à-dire la valeur m telle que $\mathbb{P}(X > m) = 1/2$.
4. Rappeler ce que vaut la quantité

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx$$

En déduire la valeur de $\int_0^{+\infty} e^{-\frac{x^2}{2}} dx$.

5. Grâce (par exemple) à une intégration par parties, montrer que $\mathbb{E}[X] = \sqrt{\frac{\pi}{2}}$
6. Soit U une variable aléatoire distribuée suivant une loi uniforme sur $]0, 1]$.
   (a) Rappeler ce que vaut la fonction de répartition F_U de U.
   (b) On considère maintenant la variable aléatoire $X = \sqrt{-2 \ln U}$. Dans quel intervalle X prend-elle ses valeurs ?
   (c) En passant par sa fonction de répartition F_X, montrer que la variable aléatoire X suit une loi de Rayleigh de paramètre 1.

Exercice 3.32 (Loi de Rademacher et marche aléatoire)
Soit X une variable suivant une loi de Bernoulli de paramètre $1/2$.

1. Rappeler la moyenne et la variance de X.
2. On considère maintenant la variable aléatoire $Y = 2X - 1$. Quelles valeurs peut prendre Y ? Avec quelles probabilités ? On dit que Y suit une loi de Rademacher.
3. Calculer la moyenne et la variance de Y.
4. On considère maintenant 100 variables aléatoires indépendantes $Y_1, \ldots, Y_{100}$, chacune suivant la loi de Rademacher. On note $S_{100} = Y_1 + \cdots + Y_{100}$ la somme de ces variables.
   (a) Quelles valeurs peut prendre la variable S_{100} ? Préciser sa moyenne et sa variance.
   (b) Un homme ivre quitte un troquet : il fait des pas d'un mètre, un coup à droite, un coup à gauche, et ce de façon équiprobable et indépendante. Au bout de 100 pas, dans un rayon de combien de mètres autour de son point de départ va-t-il se trouver avec 95% de chances ?

Exercice 3.33 (Précipitation vs. précision)

1. La quantité annuelle de précipitations (en cm) dans une certaine région est distribuée selon une loi normale de moyenne 140 et de variance 16.
   (a) Quelle est la probabilité qu'en une année il pleuve plus de 150 cm ?
   (b) Quelle est la probabilité qu'à partir d'aujourd'hui, il faille attendre au moins 10 ans avant d'obtenir une année avec une quantité annuelle de pluie supérieure à 150 cm ?
2. La largeur (en cm) d'une fente entaillée dans une pièce suit une loi normale de moyenne $m = 2$ et d'écart-type σ. Les limites de tolérance sont données comme étant 2 ± 0.012.
   (a) Si $\sigma = 0.007$, quel sera le pourcentage de pièces défectueuses ?
   (b) Quelle est la valeur maximale que peut prendre σ de sorte que le pourcentage de pièces défectueuses ne dépasse pas 1% ?

Exercice 3.34 (Loi de Weibull)
On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} 3x^2 \, e^{-x^3} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{3.1}$$

1. Vérifier que f est bien une densité de probabilité. On dit que X suit une loi de Weibull.

2. Calculer la dérivée de f. En déduire le mode de X, c'est-à-dire l'abscisse du point où f est maximale.
3. Représenter f.
4. Déterminer la fonction de répartition F de X. Donner son allure.
5. Supposons que la durée de vie (en années) d'un élément soit distribuée selon la loi de Weibull ci-dessus.
   (a) Quelle est la probabilité que cet élément dure plus de 2 ans ?
   (b) Quelle est la probabilité que sa durée de vie soit comprise entre un an et deux ans ?
   (c) Quelle est la probabilité que sa durée de vie soit supérieure à deux ans sachant qu'il fonctionne encore au bout d'un an ?

Exercice 3.35 (Loi du khi-deux)
Soit X une variable distribuée selon une loi normale centrée réduite $\mathcal{N}(0,1)$.

1. Rappeler la moyenne et la variance de X. En déduire $\mathbb{E}[X^2]$.
2. Rappeler la densité de X. Grâce à une intégration par parties et en utilisant la question précédente, montrer que $\mathbb{E}[X^4] = 3$.
3. Soit $Y = X^2$. Exprimer la variance de Y en fonction des moments de X. Déduire des questions précédentes que $\text{Var}(Y) = 2$.
4. Soit $n \in \mathbb{N}^*$ un entier naturel non nul. Si $X_1, \dots, X_n$ sont des variables indépendantes et identiquement distribuées suivant la loi normale centrée réduite, on dit que la variable $S_n = X_1^2 + \cdots + X_n^2$ suit une loi du khi-deux à n degrés de liberté, noté $S_n \sim \chi_n^2$.
   (a) Calculer la moyenne et la variance de S_n.
   (b) On tire 200 variables gaussiennes centrées réduites, on les élève au carré et on les ajoute pour obtenir un nombre S. Donner un intervalle dans lequel se situe S avec environ 95% de chances.

Exercice 3.36 (Loi de Laplace)
On considère une variable aléatoire X dont la densité f est donnée par :

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{1}{2}e^{-|x|},$$

où $|x|$ représente la valeur absolue de x, c'est-à-dire $|x| = x$ si $x \geq 0$ et $|x| = -x$ si $x \leq 0$.

1. Vérifier que f est bien une densité sur $\mathbb{R}$. Représenter f.
2. On note F la fonction de répartition de X. Calculer $F(x)$ (on distinguera les cas $x \leq 0$ et $x \geq 0$). Représenter F.
3. Montrer que $\mathbb{E}[X] = 0$.
4. Pour tout $n \in \mathbb{N}$, on appelle I_n l'intégrale définie par :
   $$I_n = \int_0^{+\infty} x^n e^{-x} dx.$$
   (a) Combien vaut I_0 ?
   (b) Montrer que pour tout $n \in \mathbb{N}^*$, $I_n = nI_{n-1}$. En déduire que $I_n = n!$ pour tout $n \in \mathbb{N}$.
5. Pour tout $n \in \mathbb{N}$, calculer $\mathbb{E}[X^{2n}]$. Que vaut $\text{Var}(X)$?
6. Pour tout $n \in \mathbb{N}$, que vaut $\mathbb{E}[X^{2n+1}]$?

Exercice 3.37 (Autour de la loi normale)
On considère une variable aléatoire X de loi normale $\mathcal{N}(0,1)$.

1. Montrer que, pour tout $n \in \mathbb{N}$, on a : $\mathbb{E}[X^{n+2}] = (n+1)\mathbb{E}[X^n]$ (intégrer par parties).
2. Que vaut $\mathbb{E}[X^2]$? Déduire de ce résultat et de la question précédente la valeur de $\mathbb{E}[X^4]$.
3. Que vaut $\mathbb{E}[X^3]$?
4. Soit Y la variable aléatoire définie par $Y = 2X + 1$.
   (a) Quelle est la loi de Y ?
   (b) Déterminer $\mathbb{E}[Y^4]$ (on pourra utiliser la formule du binôme et les moments de X trouvés précédemment).
5. A l'aide de la table de la loi normale, déterminer $\mathbb{P}(|X| \geq 2)$. Que donne l'inégalité de Tchebychev dans ce cas ? Comparer et commenter.
6. On considère maintenant que X suit une loi normale de moyenne 7 et d'écart-type 4.
   (a) Déterminer $\mathbb{P}(X \leq 8)$ et $\mathbb{P}(5 \leq X \leq 9)$.
   (b) Déterminer q tel que $\mathbb{P}(X > q) = 0,9$.
7. La taille des enfants d'un collège est distribuée selon une loi normale de moyenne m et d'écart-type σ. On sait qu'un cinquième des élèves mesurent moins de 1m50 et que 10% des élèves mesurent plus de 1m80. Déterminer m et σ.

Exercice 3.38 (Variable à densité)
Soit X une variable aléatoire de densité $f(x) = \frac{c}{x^4}\mathbb{1}_{\{x\geq 1\}}$.

1. Déterminer c pour que f soit bien une densité. Représenter f.
2. Calculer la fonction de répartition F et la représenter.
3. Déterminer la médiane de X, c'est-à-dire la valeur m telle que $\mathbb{P}(X > m) = 1/2$.
4. Calculer l'espérance de X et sa variance.
5. Déterminer le moment d'ordre 3 de X.

Exercice 3.39 (Diamètre d'une bille)
Le diamètre d'une bille est distribué suivant une loi normale de moyenne 1 cm. On sait de plus qu'une bille a une chance sur trois d'avoir un diamètre supérieur à 1.1 cm.

1. Déterminer l'écart-type de cette distribution.
2. Quelle est la probabilité qu'une bille ait un diamètre compris entre 0.2 et 1 cm ?
3. Quelle est la valeur telle que 3/4 des billes aient un diamètre supérieur à cette valeur ?

Exercice 3.40 (Tchernobyl for ever)
Soit T une variable aléatoire distribuée suivant une loi exponentielle de paramètre $\lambda > 0$.

1. Rappeler ce que valent densité, fonction de répartition, espérance et variance de T (on ne demande pas les calculs).
2. Pour tout $t > 0$, que vaut $\mathbb{P}(T > t)$?
3. On appelle demi-vie la durée h telle que $\mathbb{P}(T > h) = 1/2$. Déterminer h en fonction de λ.
4. Le strontium 90 est un composé radioactif très dangereux que l'on retrouve après une explosion nucléaire. Un atome de strontium 90 reste radioactif pendant une durée aléatoire T qui suit une loi exponentielle, durée au bout de laquelle il se désintègre. Sa demi-vie est d'environ 28 ans.
   (a) Déterminer le paramètre λ de la loi de T.
   (b) Calculer la probabilité qu'un atome reste radioactif durant au moins 50 ans.

(c) Calculer le nombre d'années nécessaires pour que 99% du strontium 90 produit par une réaction nucléaire se soit désintégré.

Exercice 3.41 (Durée de vie d'un processeur)
On modélise la durée de vie d'un processeur (en années) par une loi exponentielle de paramètre 1/2.

1. Que vaut la durée de vie moyenne d'un tel processeur ?
2. Avec quelle probabilité le processeur fonctionne-t-il plus de six mois ?
3. Chaque vente de processeur rapporte 100 euros à son fabriquant, sauf s'il doit être échangé pendant les six mois de garantie, auquel cas il ne rapporte plus que 30 euros. Combien rapporte en moyenne un processeur ?

Exercice 3.42 (Densité quadratique)
On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} c\,x^2 & 0 \leq x \leq 3 \\ 0 & \text{ailleurs} \end{cases}$$

1. Evaluer la constante c pour que f soit une densité de probabilité. Donner l'allure de f.
2. Déterminer la fonction de répartition F de X. Donner son allure.
3. Calculer $\mathbb{P}(1 < X < 2)$.
4. Déterminer espérance et variance de X.
5. Pour tout $n \in \mathbb{N}^*$, déterminer le moment d'ordre n de X.

Exercice 3.43 (Accidents et fréquence cardiaque)

1. On considère que, pour un conducteur, le nombre de kilomètres avant le premier accident suit une loi normale d'espérance 35000 km avec un écart-type de 5000 km. Pour un conducteur choisi au hasard, déterminer la probabilité :
   (a) qu'il ait eu son premier accident avant d'avoir parcouru 25000 km.
   (b) qu'il ait eu son premier accident après avoir parcouru 25000 km et avant 40000 km.
   (c) qu'il n'ait pas eu d'accident avant d'avoir parcouru 45000 km.
   (d) Au bout de combien de kilomètres peut-on dire que 80% des conducteurs ont eu leur premier accident ?
2. La fréquence cardiaque chez un adulte en bonne santé est en moyenne de 70 pulsations par minute, avec un écart-type de 10 pulsations. Soit X la variable aléatoire représentant la fréquence cardiaque chez un adulte.
   (a) A l'aide de l'inégalité de Tchebychev, minorer $\mathbb{P}(50 < X < 90)$.
   (b) Si on suppose maintenant que X suit une loi normale, que vaut $\mathbb{P}(50 < X < 90)$?

Exercice 3.44 (Loi de Gumbel)

1. On considère la fonction g définie pour tout réel x par $g(x) = e^{-e^{-x}}$. Calculer ses limites en $-\infty$ et $+\infty$, sa dérivée, et donner l'allure de g.
2. Vérifier que la fonction f définie pour tout réel x par $f(x) = e^{-x-e^{-x}}$ est une densité.
3. Soit X une variable aléatoire de loi exponentielle de paramètre 1. Rappeler ce que vaut la fonction de répartition F de X. Donner son allure.
4. Soit X_1 et X_2 des variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre 1, et soit $M = \max(X_1, X_2)$ la variable aléatoire correspondant au maximum de ces deux variables. Pour tout réel x, calculer $\mathbb{P}(M \leq x)$. En déduire la densité de M.

5. On note maintenant $M_n = \max(X_1, \ldots, X_n)$, où $X_1, \ldots, X_n$ sont variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre 1. Pour tout réel x, calculer $F_n(x) = \mathbb{P}(M_n \leq x)$.
6. Soit u un réel fixé, que vaut $\lim_{n\to+\infty}(1 - \frac{u}{n})^n$? En déduire que pour tout réel x

$$\lim_{n\to+\infty} F_n(x + \ln n) = g(x).$$

3.6 Corrigés

Exercice 3.1 (Espérance et variance d'une loi uniforme)
Soit X une variable aléatoire qui suit une loi uniforme sur le segment $[0, 1]$, c'est-à-dire $f(x) = \mathbb{1}_{[0,1]}(x)$.

1. Par définition de l'espérance, on a

$$\mathbb{E}[X] = \int_{\mathbb{R}} xf(x)dx = \int_0^1 xdx = \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}.$$

Ceci pouvait se voir sans calculs : la moyenne d'une variable uniforme est le milieu des extrémités du segment où elle tombe. Pour la variance, on obtient :

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \int_0^1 x^2dx - \frac{1}{4} = \left[\frac{x^3}{3}\right]_0^1 - \frac{1}{4} = \frac{1}{12}.$$

2. De façon générale, le moment d'ordre n de X vaut

$$\mathbb{E}[X^n] = \int_0^1 x^n dx = \left[\frac{x^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1}.$$

3. Soit a et b deux réels tels que $a < b$. Comme son nom l'indique, la densité uniforme sur le segment $[a, b]$ doit être constante sur ce segment, donc de la forme $f(x) = c\ \mathbb{1}_{[a,b]}(x)$. Il reste à déterminer c pour que f soit effectivement une densité, or

$$\int_{\mathbb{R}} xf(x)dx = \int_a^b cdx = [cx]_a^b = c(b - a),$$

quantité qui doit valoir 1 par définition d'une densité, d'où $c = (b - a)$, et

$$f(x) = \frac{1}{b-a}\mathbb{1}_{[a,b]}(x).$$

Le même type de calculs qu'en première question donne alors $\mathbb{E}[X] = (a + b)/2$ (l'interprétation étant la même que ci-dessus : en moyenne, on tombe au milieu de l'intervalle) et $\mathrm{Var}(X) = (b - a)^2/12$, formule à mettre en parallèle avec $\mathrm{Var}(X) = (n^2 - 1)/12$ d'une loi uniforme discrète sur $\{1, \ldots, n\}$.

Exercice 3.2 (Loi de Cauchy)

1. Rappelons qu'une primitive de $1/(1 + x^2)$ est la fonction arctan, fonction croissante de $\mathbb{R}$ dans $] - \frac{\pi}{2}, \frac{\pi}{2}[$ avec $\lim_{x\to-\infty} \arctan x = -\frac{\pi}{2}$ et $\lim_{x\to+\infty} \arctan x = \frac{\pi}{2}$. Ainsi

$$\int_{\mathbb{R}} f(x)dx = \int_{-\infty}^{+\infty} \frac{c}{1+x^2}dx = c\,[\arctan x]_{-\infty}^{+\infty} = c\pi,$$

donc $c = 1/\pi$ et la densité d'une variable de Cauchy de paramètre 1 (cf. figure 3.9 à gauche) est
$$\forall x \in \mathbb{R} \qquad f(x) = \frac{1}{\pi(1+x^2)}.$$
Remarque : De façon plus générale, une loi de Cauchy de paramètre $a > 0$ a pour densité $f(x) = \frac{a}{\pi(a^2+x^2)}$.

2. La fonction de répartition F de X est définie pour tout réel x par :
$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} \frac{dt}{\pi(1+t^2)} = \frac{1}{\pi}\left[\arctan t\right]_{-\infty}^{x} = \frac{1}{2} + \frac{1}{\pi}\arctan x.$$
Cette fonction de répartition est représentée figure 3.9 à droite.

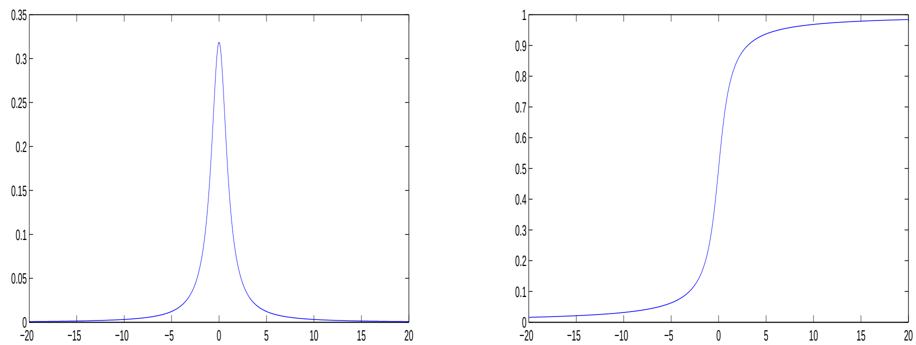


Figure 3.9 – Densité et fonction de répartition d'une loi de Cauchy.

3. Par définition, la variable X de densité f admet une espérance si l'intégrale $\int_{\mathbb{R}} xf(x)dx$ est convergente. Or dans notre cas
$$\int_{\mathbb{R}} xf(x)dx = \int_{-\infty}^{+\infty} \frac{x}{\pi(1+x^2)}dx = \int_{-\infty}^{0} \frac{x}{\pi(1+x^2)}dx + \int_{0}^{+\infty} \frac{x}{\pi(1+x^2)}dx,$$
intégrale doublement généralisée, qui converge si et seulement si les deux intégrales sont convergentes. Or
$$\int_{0}^{+\infty} \frac{x}{\pi(1+x^2)}dx = \frac{1}{2\pi}\left[\ln(1+x^2)\right]_{0}^{+\infty} = +\infty,$$
donc l'intégrale définissant l'espérance est divergente. Par conséquent la variable X n'admet pas d'espérance. Ceci est dû au fait que les queues de la densité de X ne décroissent pas assez vite vers zéro lorsque x tend vers $\pm\infty$. La loi de Cauchy est un exemple typique de loi à queue lourde (*heavy-tailed distribution*).

4. Si Y est une variable aléatoire uniforme sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, sa densité est
$$f(y) = \frac{1}{\pi}\mathbb{1}_{\left]-\frac{\pi}{2},\frac{\pi}{2}\right[}(x).$$
La fonction tangente établit une bijection de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ vers $]-\infty, +\infty[$, donc X est à valeurs dans $\mathbb{R}$ tout entier. Calculons sa fonction de répartition en utilisant le fait que la fonction arctan est la réciproque de la fonction tan et qu'elle est croissante : pour tout réel x, on a
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\tan Y \leq x) = \mathbb{P}(Y \leq \arctan x).$$

Le nombre $\arctan x$ est entre $-\pi/2$ et $+\pi/2$: la variable Y étant uniforme sur cet intervalle, sa probabilité de tomber à gauche de $\arctan x$ est le rapport entre la longueur du segment $]-\pi/2, \arctan x[$ et celle de l'intervalle $]-\pi/2, +\pi/2[$, c'est-à-dire

$$F(x) = \frac{\arctan x - (-\pi/2)}{\pi} = \frac{1}{2} + \frac{1}{\pi}\arctan x,$$

que l'on reconnaît être la fonction de répartition d'une loi de Cauchy. Autrement dit, X suit une loi de Cauchy de paramètre 1.
Remarque : ceci donne un moyen très simple de simuler une variable de Cauchy à partir d'une variable uniforme.

Exercice 3.3 (Densités parabolique et circulaire)

1. Pour que f soit bien une densité de probabilité, il faut que son intégrale soit égale à 1 :

$$1 = \int_{\mathbb{R}} f(x)dx = c\int_{-1}^{1}(1-x^2)dx = c\left[x - \frac{x^3}{3}\right]_{-1}^{1} = \frac{4c}{3},$$

donc $c = 3/4$ et $f(x) = \frac{3}{4}(1-x^2)\mathbb{1}_{\{-1<x<1\}}$. Cette densité est représentée sur le graphique de gauche de la figure 3.10.

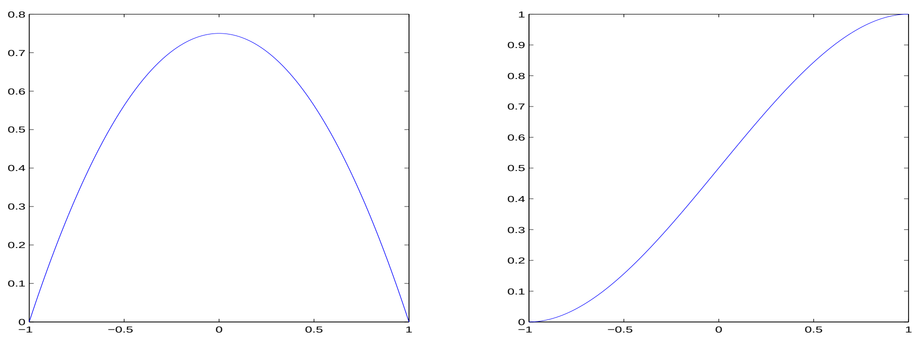


FIGURE 3.10 – Densité et fonction de répartition de la loi parabolique.

2. La variable X est à valeurs dans $]-1,+1[$ donc F est nulle à gauche de -1 et vaut 1 à droite de 1. Pour $x \in]-1,1[$, on a

$$F(x) = \int_{-1}^{x}\frac{3}{4}(1-t^2)dt = \cdots = \frac{1}{2} + \frac{3}{4}x - \frac{x^3}{4}.$$

Cette fonction de répartition est symétrique par rapport au point $(0, 1/2)$, ce qui est dû au fait que la variable X est symétrique par rapport à 0 (cf. graphique de droite de la figure 3.10).

3. La densité de X étant symétrique par rapport à 0, son espérance vaut 0. Quant à sa variance, on obtient :

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X^2] = \frac{3}{4}\int_{-1}^{1}x^2(1-x^2)dx = \frac{3}{4}\int_{-1}^{1}(x^2-x^4)dx,$$

ce qui s'intègre sans problème :

$$\mathrm{Var}(X) = \frac{3}{4}\left[\frac{x^3}{3} - \frac{x^5}{5}\right]_{-1}^{1} = \frac{1}{5}.$$

4. (a) Pour la densité $f(x) = c\sqrt{1-x^2}\mathbb{1}_{\{-1<x<1\}}$, nous allons utiliser le changement de variable $x = \cos t$. Pour cela, rappelons que la fonction cosinus établit une bijection de $[0, \pi/2]$ vers $[0, 1]$, d'où :

$$1 = c\int_{-1}^{1}\sqrt{1-x^2}dx = 2c\int_0^1\sqrt{1-x^2}dx = 2c\int_0^{\frac{\pi}{2}}\sqrt{1-\cos^2 t}\sin t dt = 2c\int_0^{\frac{\pi}{2}}\sin^2 t dt,$$

et on applique la formule de linéarisation $\sin^2 t = (1-\cos 2t)/2$ pour finir :

$$1 = c\int_0^{\frac{\pi}{2}}(1-\cos 2t)dt = c\left[t - \frac{\sin 2t}{2}\right]_0^{\frac{\pi}{2}} = c\frac{\pi}{2},$$

d'où $c = 2/\pi$. Ce résultat pouvait se voir sans aucun calcul : l'aire sous la courbe définie par $y = \sqrt{1-x^2}$ est tout simplement la surface d'un demi-cercle de rayon 1, donc elle vaut $\pi/2$.

(b) Le raisonnement de la question précédente va néanmoins nous servir pour déterminer la fonction de répartition F de X. Comme dans l'exemple précédent, commençons par noter que F est nulle à gauche de -1 et vaut 1 à droite de 1. De même, puisque la densité de X est paire, sa fonction de répartition est symétrique par rapport au point $(0, 1/2)$, c'est-à-dire que $F(-x) = 1 - F(x)$. En particulier on a $F(0) = 1/2$. On peut donc se contenter de calculer $F(x)$ pour $x \in]0, 1[$, ce qui donne

$$F(x) = \int_{-1}^{x}\frac{2}{\pi}\sqrt{1-t^2}dt = \frac{1}{2} + \frac{2}{\pi}\int_{\arccos x}^{\frac{\pi}{2}}\sqrt{1-\cos^2 u}\sin u du,$$

et il suffit alors de bidouiller tout ça comme précédemment pour arriver à :

$$F(x) = 1 - \frac{1}{\pi}\arccos x + \frac{1}{2\pi}\sin(2\arccos x),$$

et via la relation $\sin 2t = 2\cos t\sin t = 2\cos t\sqrt{1-\cos^2 t}$, il vient finalement pour tout $x \in [0, 1]$:

$$F(x) = 1 + \frac{x\sqrt{1-x^2} - \arccos x}{\pi},$$

et cette formule est encore valide pour $x \in [-1, 0]$. La densité et la fonction de répartition de X sont représentées figure 3.11.

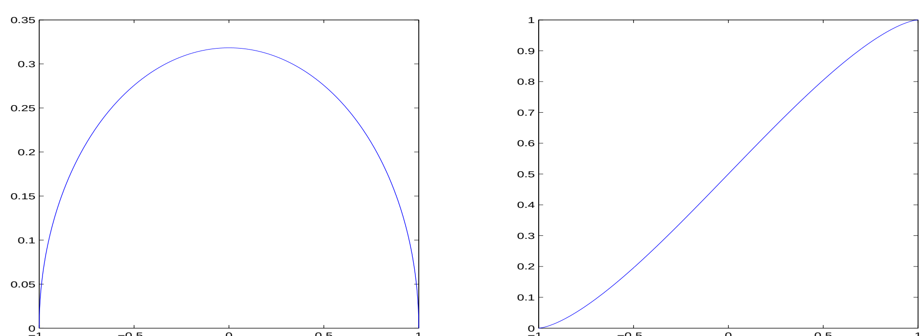


FIGURE 3.11 – Densité et fonction de répartition de la loi circulaire.

(c) Par symétrie de X, il est clair que $\mathbb{E}[X] = 0$. Ainsi sa variance est égale à son moment d'ordre 2 :

$$\text{Var}(X) = \mathbb{E}[X^2] = \frac{2}{\pi}\int_{-1}^{1} x^2\sqrt{1-x^2}dx = \frac{4}{\pi}\int_{0}^{1} x^2\sqrt{1-x^2}dx = \frac{4}{\pi}\int_{0}^{\frac{\pi}{2}} \cos^2 t \sin^2 t dt$$

et via les relations $\sin 2t = 2\sin t\cos t$ et $\sin^2(2t) = (1-\cos 4t)/2$, ceci donne

$$\text{Var}(X) = \frac{1}{2\pi}\left[t - \frac{\sin 4t}{4}\right]_0^{\frac{\pi}{2}} = \frac{1}{4}.$$

Exercice 3.4 (Loi exponentielle)

1. Puisque λ est positif, vérifier que f est une densité revient à vérifier que son intégrale vaut 1 :
$$\int_{\mathbb{R}} f(x)dx = \int_0^{+\infty} \lambda e^{-\lambda x}dx = \left[-e^{-\lambda x}\right]_0^{+\infty} = 1,$$
donc f est bien une densité, représentée (pour $\lambda = 1/8$) sur le graphique de gauche figure 3.12.
2. Puisque X est à valeurs dans $[0, +\infty[$, il est clair que $F(x) = 0$ pour $x \leq 0$. Pour $x \geq 0$, il vient :
$$F(x) = \int_0^x f(t)dt = \int_0^x \lambda e^{-\lambda t}dt = \left[-e^{-\lambda t}\right]_0^x = 1 - e^{-\lambda x}.$$
Cette fonction de répartition est représentée (pour $\lambda = 1/8$) sur le graphique de droite figure 3.12.

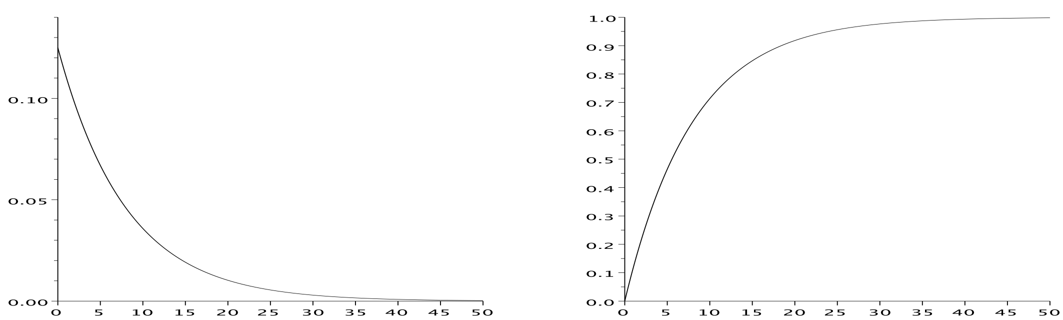


FIGURE 3.12 – Densité et fonction de répartition de la loi exponentielle de paramètre $\lambda = \frac{1}{8}$.

3. Le calcul de l'espérance de X se fait par une intégration par parties :
$$\mathbb{E}[X] = \int_0^{+\infty} x \times \lambda e^{-\lambda x}dx = \left[-xe^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x}dx.$$
Or $\lambda > 0$ donc $\lim_{x\to+\infty} xe^{-\lambda x} = 0$, l'exponentielle imposant sa limite à x. Ainsi :
$$\mathbb{E}[X] = \left[-\frac{1}{\lambda}e^{-\lambda x}\right]_0^{+\infty} = \frac{1}{\lambda}.$$
Ce résultat est intuitivement cohérent : plus λ est grand, plus X prend souvent des valeurs proches de 0, donc plus sa moyenne est faible. Pour la variance, on commence par calculer $\mathbb{E}[X^2]$, ce qui se fait derechef via une intégration par parties :
$$\mathbb{E}[X^2] = \int_0^{+\infty} x^2 \times \lambda e^{-\lambda x}dx = \left[-x^2e^{-\lambda x}\right]_0^{+\infty} + \int_0^{+\infty} 2xe^{-\lambda x}dx = 2\int_0^{+\infty} xe^{-\lambda x}dx.$$

Or cette intégrale a quasiment été calculée ci-dessus :

$$2\int_0^{+\infty} xe^{-\lambda x}dx = \frac{2}{\lambda}\int_0^{+\infty} x \times \lambda e^{-\lambda x}dx = \frac{2}{\lambda}\mathbb{E}[X] = \frac{2}{\lambda^2}.$$

Ainsi $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 1/\lambda^2$.

4. La durée de vie T en années d'une télévision suit une loi de densité $f(t) = \frac{1}{8}e^{-\frac{t}{8}}\mathbb{1}_{\{t\geq 0\}}$.
   (a) On reconnaît une loi exponentielle : $T \sim \mathcal{E}(1/8)$. La durée de vie moyenne de cette télévision est donc $\mathbb{E}[T] = 8$ ans. L'écart-type de cette durée de vie est $\sigma(T) = 8$ ans.
   (b) La probabilité que cette télévision ait une durée de vie supérieure à 8 ans est :

$$\mathbb{P}(T \geq 8) = 1 - \mathbb{P}(T \leq 8) = 1 - F(8) = 1 - \left(1 - e^{-8/8}\right) = e^{-1} \approx 0.37.$$

   Il y a donc environ 37% de chances que cette télévision dure plus de 8 ans.

Exercice 3.5 (Absence de mémoire)

1. D'après l'exercice précédent, on a pour tout $t \geq 0$:

$$\mathbb{P}(X > t) = 1 - \mathbb{P}(X \leq t) = 1 - F(t) = 1 - \left(1 - e^{-\lambda t}\right) = e^{-\lambda t}.$$

2. Pour tout couple $(x, t) \in \mathbb{R}^+ \times \mathbb{R}^+$, nous avons donc :

$$\mathbb{P}(X > x + t|X > x) = \frac{\mathbb{P}(\{X > x + t\} \cap \{X > x\})}{\mathbb{P}(X > x)} = \frac{\mathbb{P}(X > x + t)}{\mathbb{P}(X > x)} = \frac{e^{-\lambda(x+t)}}{e^{-\lambda x}} = e^{-\lambda t}$$

   donc d'après la question précédente $\mathbb{P}(X > x + t|X > x) = \mathbb{P}(X > t)$. La loi exponentielle n'a pas de mémoire.
3. Application : la probabilité cherchée s'écrit

$$\mathbb{P}(X > 2 + 8|X > 2) = \mathbb{P}(X > 8) = e^{-1} \approx 0.37.$$

Exercice 3.6 (Durée de vie)

Un appareil comporte six composants de même modèle, tous nécessaires à son fonctionnement. La densité de la durée de vie T d'un composant est donnée par $f(t) = \frac{t}{16}e^{-\frac{t}{4}}\mathbb{1}_{\{t\geq 0\}}$, l'unité de temps étant l'année.

1. La fonction f est positive donc pour être une densité de probabilité, il suffit qu'elle somme à 1, ce qui se vérifie par une intégration par parties :

$$\int_0^{+\infty} \frac{t}{16}e^{-\frac{t}{4}}dt = \left[-\frac{t}{4}e^{-\frac{t}{4}}\right]_0^{+\infty} + \int_0^{+\infty} \frac{1}{4}e^{-\frac{t}{4}}dt = \left[-e^{-\frac{t}{4}}\right]_0^{+\infty} = 1.$$

2. Considérons une variable aléatoire $X \sim \mathcal{E}(\lambda)$, alors on montre facilement par récurrence que son moment d'ordre n est $n!/\lambda^n$, c'est-à-dire :

$$\mathbb{E}[X^n] = \int_0^{+\infty} x^n \times \lambda e^{-\lambda x}dx = \frac{n!}{\lambda^n}.$$

   Cette relation générale permet de calculer espérance et variance de T en considérant le cas particulier où $X \sim \mathcal{E}(1/4)$:

$$\mathbb{E}[T] = \int_0^{+\infty} \frac{t^2}{16}e^{-\frac{t}{4}}dt = \frac{1}{4}\int_0^{+\infty} \frac{t^2}{4}e^{-\frac{t}{4}}dt = \frac{\mathbb{E}[X^2]}{4} = 8.$$

Pour le calcul de la variance, on commence par le moment d'ordre 2 :

$$\mathbb{E}[T^2] = \int_0^{+\infty} \frac{t^3}{16} e^{-\frac{t}{4}} dt = \frac{1}{4} \int_0^{+\infty} \frac{t^3}{4} e^{-\frac{t}{4}} dt = \frac{\mathbb{E}[X^3]}{4} = 96.$$

Il s'ensuit que :

$$\mathrm{Var}(T) = \mathbb{E}[T^2] - (\mathbb{E}[T])^2 = 32.$$

Remarque : Lois Gamma. Soit $n \in \mathbb{N}^*$ et $\lambda > 0$. On dit que T suit une loi Gamma de paramètres n et λ, noté $T \sim \Gamma(n, \lambda)$, si T admet pour densité :

$$f(t) = \frac{(\lambda t)^{n-1}}{(n-1)!} \times \lambda e^{-\lambda t} \mathbb{1}_{\{t \geq 0\}}.$$

On voit ainsi que la variable étudiée dans cet exercice suit une loi $\Gamma(2, 1/4)$. L'intérêt des lois Gamma vient de la remarque suivante : si $T_1, \ldots, T_n$ sont des variables indépendantes et de même loi $\mathcal{E}(\lambda)$, alors leur somme T suit une loi Gamma : $T = T_1 + \cdots + T_n \sim \Gamma(n, \lambda)$. De la connaissance des lois exponentielles, on déduit alors sans calculs lourdingues que $\mathbb{E}[T] = n\mathbb{E}[T_1] = n/\lambda$ et $\mathrm{Var}(T) = n\mathrm{Var}(T_1) = n/\lambda^2$.

3. La probabilité qu'un composant fonctionne durant au moins six ans à partir de sa mise en marche est :

$$\mathbb{P}(T \geq 6) = \int_6^{+\infty} \frac{t}{16} e^{-\frac{t}{4}} dt = \left[-\frac{t}{4} e^{-\frac{t}{4}} \right]_6^{+\infty} + \int_6^{+\infty} \frac{1}{4} e^{-\frac{t}{4}} dt,$$

c'est-à-dire :

$$\mathbb{P}(T \geq 6) = \frac{3}{2} e^{-\frac{3}{2}} + \left[-e^{-\frac{t}{4}} \right]_6^{+\infty} = \frac{5}{2} e^{-\frac{3}{2}} \approx 0.56.$$

Puisque les six composants, supposés indépendants, sont tous nécessaires à son fonctionnement, la probabilité que l'appareil fonctionne durant au moins six ans à partir de sa mise en marche est donc : $p = \mathbb{P}(T \geq 6)^6 = \frac{15625}{64} e^{-9} \approx 0.03$

Exercice 3.7 (Loi de Pareto)

1. Le calcul se fait sans problème à partir de la densité :

$$\mathbb{P}(T > 20) = \int_{20}^{+\infty} \frac{10}{t^2} dt = \left[-\frac{10}{t} \right]_{20}^{+\infty} = \frac{1}{2}.$$

Remarque : la loi de Pareto est un autre exemple de distribution à queue lourde. En particulier, on voit ici qu'elle n'admet pas d'espérance puisque

$$\int_{\mathbb{R}} t f(t) dt = \int_{10}^{+\infty} \frac{10}{t} dt = [-10 \ln t]_{10}^{+\infty} = +\infty.$$

2. Notons F la fonction de répartition de T. Puisque T est à valeurs dans l'intervalle $[10, +\infty[$, F est nulle sur $]-\infty, 10[$. Pour $t \geq 10$, on a :

$$F(t) = \int_{10}^{t} \frac{10}{u^2} du = \left[-\frac{10}{u} \right]_{10}^{t} = 1 - \frac{10}{t}.$$

La densité et la fonction de répartition sont représentée sur la figure 3.13.

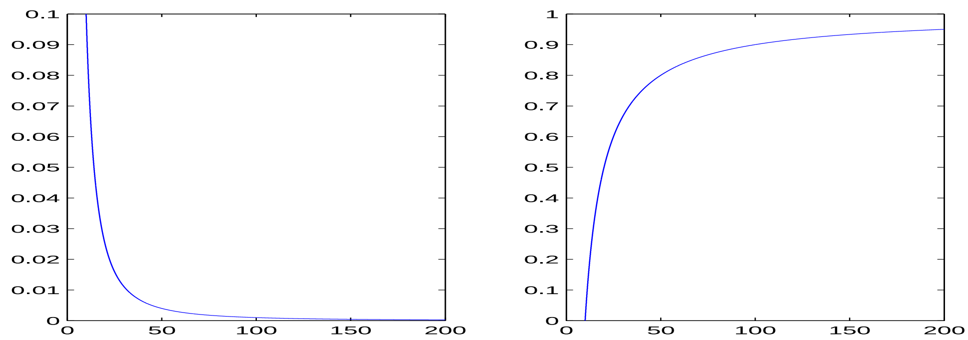


FIGURE 3.13 – Densité et fonction de répartition de la loi de Pareto.

3. Commençons par calculer la probabilité p qu'un composant fonctionne durant au moins 15 heures. Celle-ci s'écrit :
$$p = \mathbb{P}(T > 15) = \int_{15}^{+\infty} \frac{10}{t^2} dt = \left[-\frac{10}{t} \right]_{15}^{+\infty} = \frac{2}{3}.$$
Puisque les composants sont indépendants, le nombre de composants parmi les 6 à être encore en fonctionnement après 15 heures est une variable X distribuée suivant une loi binomiale $\mathcal{B}(6, 2/3)$. Ainsi la probabilité P que parmi 6 composants indépendants, au moins 3 d'entre eux fonctionnent durant au moins 15 heures s'écrit :
$$P = \mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X < 3) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)),$$
qui se développe en :
$$P = 1 - \left(\binom{6}{0} \left(\frac{2}{3}\right)^0 \left(\frac{1}{3}\right)^6 + \binom{6}{1} \left(\frac{2}{3}\right)^1 \left(\frac{1}{3}\right)^5 + \binom{6}{2} \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^4 \right) = \frac{656}{729} \approx 0.9$$
Il y a donc environ 90% de chances qu'au moins 3 des composants fonctionnent durant au moins 15 heures.

Exercice 3.8 (Tirages uniformes sur un segment)
Cet exercice a été traité dans le chapitre 1 (exercice 1.37). La seule différence ici se trouve dans la formulation en termes de variables aléatoires.

Exercice 3.9 (Problèmes de densité)

1. Puisque X prend ses valeurs dans $[0, 1]$, $Y = 1 - X$ aussi, donc si on note F la fonction de répartition de Y, il en découle que F est nulle à gauche de 0 et vaut 1 à droite de 1. Pour $y \in [0, 1]$, il suffit de se ramener à X :
$$F(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(1 - X \leq y) = \mathbb{P}(X \geq 1 - y) = 1 - (1 - y) = y,$$
où l'on retrouve la fonction de répartition de la loi uniforme sur $[0, 1]$, ainsi $Y \sim \mathcal{U}_{[0,1]}$. ce résultat était intuitivement clair : puisque X se distribue de façon uniforme sur $[0, 1]$, il en va de même pour $1-X$. Concernant la variable aléatoire $Z = (X+Y)$, on a $Z = X+(1-X) = 1$, c'est-à-dire que Z ne prend que la valeur 1 et n'a rien d'aléatoire. On peut aussi voir Z comme une variable discrète prenant la valeur 1 avec probabilité 1. Ce petit exemple montre que la somme de deux variables à densité n'a pas forcément de densité (tandis que la somme de deux variables discrètes reste toujours une variable discrète).

2. La variable X est à valeurs dans $[0,1]$, donc sa fonction de répartition est nulle sur $\mathbb{R}^-$ et vaut 1 sur $[1,+\infty[$. Notons P (respectivement F) l'événement "la pièce donne Pile (respectivement Face)" et U le résultat du tirage uniforme. Considérons $x \in [0,1[$, alors par indépendance du lancer de dé et du tirage uniforme il vient :

$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\{U \leq x\} \cap \{F\}) = \mathbb{P}(U \leq x)\mathbb{P}(F) = \frac{x}{2}.$$

Ainsi F est une fonction continue par morceaux, avec un saut d'amplitude $1/2$ en 1 (cf. figure 3.14). Autrement dit, la variable X n'est ni discrète ni absolument continue, c'est une sorte de cocktail.

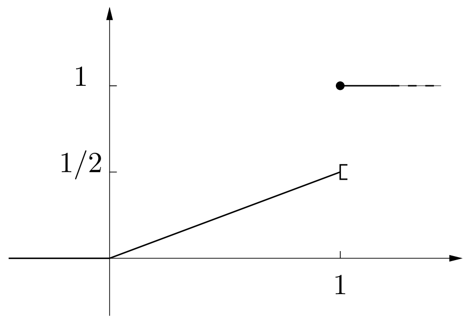


Figure 3.14 – Fonction de répartition de la variable "cocktail".

Exercice 3.10 (Minimum de variables exponentielles)

1. On considère deux variables aléatoires indépendantes X_1 et X_2 exponentielles de paramètres respectifs λ_1 et λ_2. Soit $Y = \min(X_1, X_2)$ le minimum de ces deux variables.
   (a) Pour tout réel y :
   $$P(X_1 > y) = 1 - P(X_1 \leq y) = 1 - F_1(y),$$
   où F_1 est la fonction de répartition de la loi exponentielle $\mathcal{E}(\lambda_1)$. Ainsi, $P(X_1 > y) = 1$ si $y \leq 0$, et si $y \geq 0$ nous avons
   $$P(X_1 > y) = 1 - \left(1 - e^{-\lambda_1 y}\right) = e^{-\lambda_1 y}.$$
   (b) La variable Y, en tant que minimum de deux variables positives, est elle-même positive donc $P(Y > y) = 1$ pour tout $y \leq 0$. Pour $y > 0$, utilisons l'indépendance des deux variables pour transformer une probabilité d'intersection en produit de probabilités :
   $$P(Y > y) = \mathbb{P}(\min(X_1, X_2) > y) = \mathbb{P}(X_1 > y, X_2 > y) = \mathbb{P}(X_1 > y)\mathbb{P}(X_2 > y)$$
   c'est-à-dire :
   $$P(Y > y) = e^{-\lambda_1 y}e^{-\lambda_2 y} = e^{-(\lambda_1+\lambda_2)y}.$$
   En notant F la fonction de répartition de la variable Y, il en découle que pour tout $y > 0$:
   $$F(y) = 1 - P(Y > y) = 1 - e^{-(\lambda_1+\lambda_2)y},$$
   et bien sûr $F(y) = 0$ si $Y \leq 0$.
   (c) La fonction de répartition caractérisant complètement la loi d'une variable aléatoire, on en déduit que Y suit une loi exponentielle de paramètre $(\lambda_1 + \lambda_2)$.

2. Pour répondre à cette question, il suffit d'appliquer le résultat précédent en considérant X_1 (respectivement X_2) comme le temps mis par Alice (respectivement Bob) pour pouvoir sortir. L'énoncé implique que $X_1 \sim \mathcal{E}(1/20)$ et $X_2 \sim \mathcal{E}(1/30)$. Le temps mis par le premier pour sortir est la variable aléatoire $Y = \min(X_1, X_2)$, laquelle suit donc une loi exponentielle de paramètre $1/20 + 1/30$: $Y \sim \mathcal{E}(5/60)$. En moyenne, le premier sort donc au bout de $\mathbb{E}[Y] = 60/5 = 12$ minutes.
3. Le temps nécessaire pour que les deux soient sortis correspond à la variable $X = \max(X_1, X_2)$. La loi de X n'est plus une bête exponentielle : on pourrait facilement la déterminer via sa fonction de répartition, mais nul besoin ici puisqu'on ne s'intéresse qu'à sa moyenne. Il suffit en effet de remarquer que $X + Y = X_1 + X_2$ pour en déduire que
$$\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] - \mathbb{E}[Y] = 20 + 30 - 12 = 38 \text{ minutes.}$$
Remarque : Notons que par définition du minimum, nous avons à la fois $Y \leq X_1$ et $Y \leq X_2$, donc il n'est pas étonnant de voir que $\mathbb{E}[Y] \leq \mathbb{E}[X_1]$ et $\mathbb{E}[Y] \leq \mathbb{E}[X_2]$ (propriété de positivité de l'espérance). Idem pour le maximum.

Exercice 3.11 (Think Tank)

1. Comme d'habitude, il faut que c vérifie :
$$1 = \int_{\mathbb{R}} f(x)dx = c\int_0^1 (1-x)^4 dx = c\left[-\frac{(1-x)^5}{5}\right]_0^1 = \frac{c}{5},$$
donc il faut $c = 5$, ce qui donne
$$f(x) = 5(1-x)^4 \mathbb{1}_{\{0<x<1\}}.$$
2. La fonction de répartition F est nulle à gauche de 0 et vaut 1 à droite de 1. Pour $0 \leq x \leq 1$, on a :
$$F(x) = \int_0^x f(t)dt = \left[-(1-t)^5\right]_0^x = 1 - (1-x)^5.$$
Densité et fonction de répartition sont représentées figure 3.15.

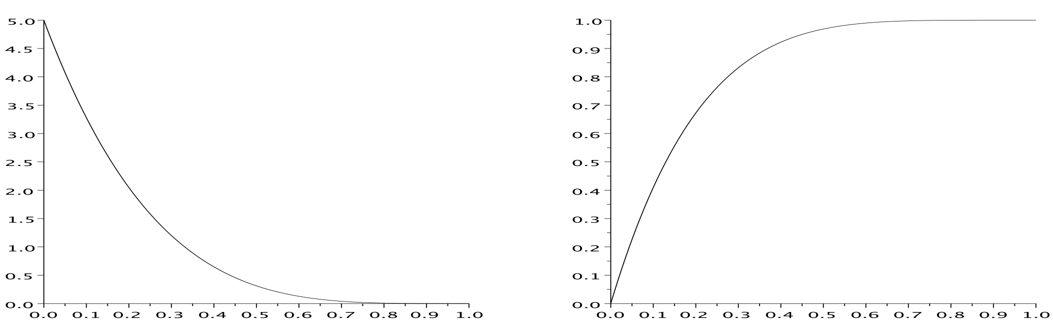


FIGURE 3.15 – Densité et fonction de répartition de la variable X (demande hebdomadaire).

3. Puisque X représente la demande hebdomadaire, nous cherchons la valeur x telle que $\mathbb{P}(X > x) \leq 10^{-5}$, c'est-à-dire tel que $1 - F(x) \leq 10^{-5}$, or :
$$1 - F(x) \leq 10^{-5} \;\Leftrightarrow\; (1-x)^5 \leq 10^{-5} \;\Leftrightarrow\; 1 - x \leq 10^{-1} \;\Leftrightarrow\; x \geq 0.9.$$
Il faut donc que la capacité du réservoir soit d'au moins 900 litres.

Exercice 3.12 (Loi polynomiale)

1. Pour que f soit effectivement une densité, il faut que
$$1 = \int_{\mathbb{R}} f(x)dx = c\int_0^1 (x+x^2)dx = c\left[\frac{x^2}{2}+\frac{x^3}{3}\right]_0^1 = \frac{5c}{6},$$
donc il faut $c = 6/5$, ce qui donne $f(x) = \frac{6}{5}(x+x^2)\mathbb{1}_{\{0<x<1\}}$.
2. La fonction de répartition F est nulle à gauche de 0 et vaut 1 à droite de 1. Pour $0 \leq x \leq 1$, on a :
$$F(x) = \int_0^x f(t)dt = \frac{6}{5}\left[\frac{t^2}{2}+\frac{t^3}{3}\right]_0^x = \frac{6}{5}\left(\frac{x^2}{2}+\frac{x^3}{3}\right).$$
Densité et fonction de répartition sont représentées figure 3.16.

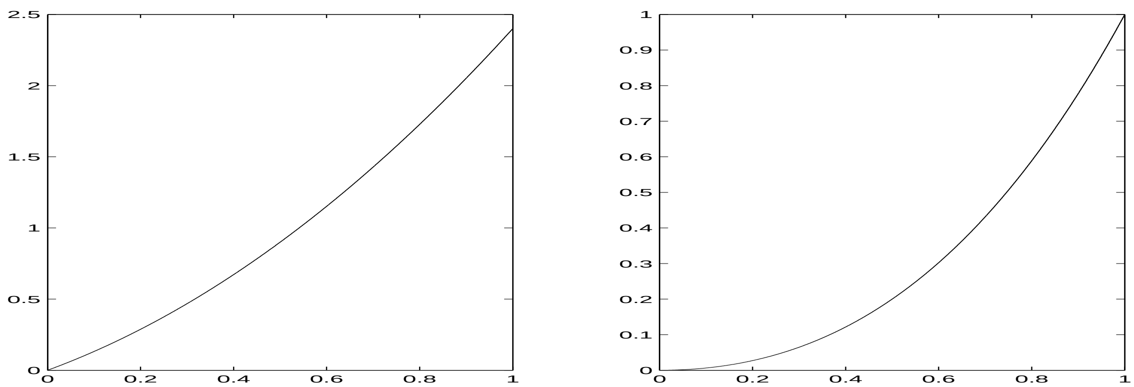


FIGURE 3.16 – Densité et fonction de répartition de la variable X (volume de vente hebdomadaire).

3. L'espérance de X vaut :
$$\mathbb{E}[X] = \frac{6}{5}\int_0^1 x(x+x^2)dx = c\left[\frac{x^3}{3}+\frac{x^4}{4}\right]_0^1 = \frac{7}{10}.$$
Pour l'écart-type, on commence par calculer le moment d'ordre 2 de X :
$$\mathbb{E}[X^2] = \frac{6}{5}\int_0^1 x^2(x+x^2)dx = \frac{6}{5}\left[\frac{x^4}{4}+\frac{x^5}{5}\right]_0^1 = \frac{27}{50}.$$
L'écart-type de X s'en déduit :
$$\sigma(X) = \sqrt{\mathbb{E}[X^2]-(\mathbb{E}[X])^2} = \frac{\sqrt{5}}{10} \approx 0.224.$$

Exercice 3.13 (Ambulance et accidents)

1. Au mieux l'accident a lieu devant la station, auquel cas $T = 0$. Au pire, l'accident se produit à 70 kms de la station et il faut alors $70 \times 60/100 = 42$ minutes à l'ambulance pour arriver. La variable T prend donc ses valeurs dans l'intervalle $[0, 42]$.
2. Puisque l'ambulance roule à 100 km/h pour intervenir sur le lieu d'un accident, la quantité $\mathbb{P}(T > 30)$ est la probabilité qu'un accident ait lieu à une distance supérieure à 50 kms de la station, c'est-à-dire entre le kilomètre 80 et le kilomètre 100 de la route. Puisque les accidents sont supposés arriver uniformément sur les 100 kms, on en déduit que la quantité cherchée vaut :
$$\mathbb{P}(T > 30) = \frac{100-80}{100} = \frac{1}{5}.$$
Il y a donc une chance sur 5 que l'ambulance mette plus d'une demi-heure à intervenir.

3. Pour le calcul plus général, commençons par voir qu'un accident situé à plus de trente kilomètres de la station ne peut se situer qu'en un endroit (disons à l'ouest de la station), tandis qu'un accident situé à moins de trente kilomètres de la station peut se situer en deux endroits (à l'est ou à l'ouest). Il faut $30 \times 60/100 = 18$ minutes pour faire 30 kms. Ainsi, pour $t \in [18, 42]$, le raisonnement de la question précédente s'applique :
$$\mathbb{P}(T > t) = \frac{100 - (30 + t \times 100/60)}{100} = \frac{7}{10} - \frac{t}{60}.$$
Pour $t \in [0, 18]$, on passe à l'événement complémentaire :
$$\mathbb{P}(T > t) = 1 - \mathbb{P}(T \leq t) = 1 - \frac{(30 + t \times 100/60) - (30 - t \times 100/60)}{100} = 1 - \frac{t}{30}.$$
Par ailleurs, il est clair que $\mathbb{P}(T > t) = 1$ pour tout $t \leq 0$ et $\mathbb{P}(T > t) = 0$ pour tout $t \geq 42$.
4. La fonction de répartition F de la variable T est nulle à gauche de 0 et vaut 1 à droite de 42, donc la densité f est nulle à gauche de 0 et à droite de 1. Pour $t \in [0, 18]$, nous pouvons écrire :
$$F(t) = 1 - \mathbb{P}(T > t) = \frac{t}{30} \;\Rightarrow\; f(t) = \frac{1}{30},$$
tandis que pour $t \in [18, 42]$:
$$F(t) = 1 - \mathbb{P}(T > t) = \frac{3}{10} + \frac{t}{60} \;\Rightarrow\; f(t) = \frac{1}{60}.$$
La densité de T est donc une fonction constante par morceaux. Dans ces conditions, la moyenne de T vaut :
$$\mathbb{E}[T] = \int_{\mathbb{R}} t f(t) dt = \frac{1}{30}\int_0^{18} t dt + \frac{1}{60}\int_{18}^{42} t dt = \frac{1}{30}\left[\frac{t^2}{2}\right]_0^{18} + \frac{1}{60}\left[\frac{t^2}{2}\right]_{18}^{42} = 17.4$$
Le temps d'intervention moyen est donc de 17 minutes et 24 secondes.
Pour la variance, commençons par calculer le moment d'ordre 2 de T :
$$\mathbb{E}[T^2] = \int_{\mathbb{R}} t^2 f(t) dt = \frac{1}{30}\int_0^{18} t^2 dt + \frac{1}{60}\int_{18}^{42} t^2 dt = \frac{1}{30}\left[\frac{t^3}{3}\right]_0^{18} + \frac{1}{60}\left[\frac{t^3}{3}\right]_{18}^{42} = 444.$$
On en déduit $\mathrm{Var}(X) = 444 - 17.4^2 = 141.24$, donc $\sigma(X) \approx 11.88$. L'écart-type du temps d'intervention est donc un peu inférieur à 12 minutes.

Exercice 3.14 (Minimum d'uniformes)

1. On connaît la fonction de répartition d'une loi uniforme donc pour $t \in [0, 1]$:
$$\mathbb{P}(U > t) = 1 - \mathbb{P}(U \leq t) = 1 - t.$$
2. La variable X est à valeurs dans l'intervalle $[0, 1]$ donc sa fonction de répartition vaut 0 sur $\mathbb{R}^-$ et 1 sur $[1, +\infty[$. Pour $t \in [0, 1]$, passons à l'événement complémentaire :
$$F(x) = \mathbb{P}(X \leq x) = 1 - \mathbb{P}(X > x) = 1 - \mathbb{P}(U_1 > x, \dots, U_n > x),$$
et appliquons l'indépendance des variables U_i :
$$F(x) = 1 - \mathbb{P}(U_1 > x) \dots \mathbb{P}(U_n > x) = 1 - (1 - x)^n.$$

3. Notons f la densité de X. Il est clair que $f = 0$ sur $\mathbb{R}^-$ et sur $[1, +\infty[$. Pour $x \in [0, 1]$, il suffit de dériver la fonction de répartition F, ce qui donne au final :
$$f(x) = n(1-x)^{n-1}\mathbb{1}_{\{0 \leq x \leq 1\}}.$$
Le calcul de l'espérance de X est alors automatique :
$$\mathbb{E}[X] = \int_{\mathbb{R}} xf(x)dx = \int_0^1 xn(1-x)^{n-1}dt,$$
où l'on applique naturellement une intégration par parties :
$$\mathbb{E}[X] = [-x(1-x)^n]_0^1 + \int_0^1 (1-x)^n dx = \left[-\frac{(1-x)^{n+1}}{n+1}\right]_0^1 = \frac{1}{n+1}.$$
Dans le cas particulier où $n = 1$, on a $X = U_1$ et on retrouve $\mathbb{E}[X] = 1/2$, moyenne d'une loi uniforme sur $[0, 1]$.

Exercice 3.15 (Racines d'un trinôme aléatoire)

1. Le discriminant de P vaut :
$$\Delta = 16(U^2 - U - 2) = 16(U+1)(U-2).$$
2. On peut écrire $D(u) = (u+1)(u-2)$, d'où l'on déduit que D admet les 2 racines -1 et 2, est strictement négative sur $]-1, +2[$ et strictement positive sur $]-\infty, -1[$ et sur $]+2, +\infty[$.
3. Pour que P ait deux racines réelles distinctes, il faut et il suffit que son discriminant soit strictement positif. Au vu des deux questions précédentes et puisque U prend ses valeurs sur $[0, 5]$, on en déduit que la probabilité p que P ait deux racines réelles distinctes vaut :
$$p = \mathbb{P}(U \in]-\infty, -1[\cup]+2, +\infty[) = \mathbb{P}(2 < U \leq 5) = \frac{3}{5}.$$

Exercice 3.16 (Lien entre lois exponentielle et géométrique)

1. Puisque X est à valeurs dans $]0, +\infty[$, Y est à valeurs dans $\mathbb{N}^*$. Pour tout $n \in \mathbb{N}^*$, on a :
$$\mathbb{P}(Y = n) = \mathbb{P}(n-1 < X \leq n) = F(n) - F(n-1) = \left(1 - e^{-n}\right) - \left(1 - e^{-(n-1)}\right) = e^{-(n-1)} - e^{-n}$$
qui s'écrit encore :
$$\mathbb{P}(Y = n) = \left(1 - e^{-1}\right)e^{-(n-1)}.$$
On voit que Y suit donc une loi géométrique de paramètre $1 - 1/e$, noté $Y \sim \mathcal{G}(1 - 1/e)$. Sa moyenne vaut donc
$$\mathbb{E}[Y] = \frac{1}{1 - 1/e} = \frac{e}{e-1} \approx 1.58.$$
Sa variance vaut quant à elle :
$$\mathrm{Var}(Y) = \frac{1 - (1 - 1/e)}{(1 - 1/e)^2} = \frac{e}{(e-1)^2} \approx 0.92.$$
2. Soit alors $Z = Y - X$. la variable Z est à valeurs dans $[0, 1[$ puisque pour tout réel x, $x \leq \lceil x \rceil < x + 1$. Soit donc $z \in [0, 1[$: dire que $Z = Y - X = \lceil X \rceil - X$ est inférieure à z, c'est dire que X est à distance inférieure à z de l'entier supérieur le plus proche, c'est-à-dire qu'il existe $n \in \mathbb{N}^*$ tel que $n - z \leq X \leq n$. Formalisons ceci :
$$\mathbb{P}(Z \leq z) = \mathbb{P}\left(\bigcup_{n=1}^{+\infty}\{n - z \leq X \leq n\}\right).$$

Puisqu'on a affaire à une union d'événements deux à deux disjoints, la sigma-additivité s'applique :

$$\mathbb{P}(Z \leq z) = \sum_{n=1}^{+\infty} \mathbb{P}(n - z \leq X \leq n) = \sum_{n=1}^{+\infty} (F(n) - F(n-z)),$$

où F est comme précédemment la fonction de répartition d'une loi exponentielle de paramètre 1 :

$$\mathbb{P}(Z \leq z) = \sum_{n=1}^{+\infty} \left(\left(1 - e^{-n}\right) - \left(1 - e^{-(n-z)}\right) \right) = \sum_{n=1}^{+\infty} (e^z - 1)\, e^{-n},$$

et on reconnaît une série géométrique de raison $1/e$:

$$\mathbb{P}(Z \leq z) = (e^z - 1) \sum_{n=1}^{+\infty} e^{-n} = (e^z - 1)\, \frac{e^{-1}}{1 - e^{-1}} = \frac{e^z - 1}{e - 1}.$$

On a ainsi déterminé la fonction de répartition de la variable aléatoire Z.

3. Sa densité s'en déduit par dérivation :

$$f(z) = \frac{e^z}{e-1} \mathbb{1}_{[0,1]}(z).$$

4. Pour trouver $\mathbb{E}[Z]$, inutile de passer par la densité, il suffit d'utiliser les moyennes des lois géométrique et exponentielle : $\mathbb{E}[Z] = \mathbb{E}[Y - X] = \mathbb{E}[Y] - \mathbb{E}[X] = \frac{e}{e-1} - 1 = \frac{1}{e-1} \approx 0.58$.

Exercice 3.17 (Moments d'une loi normale)

1. $I_0 = \sqrt{2\pi}$ puisqu'on reconnaît la densité d'une loi normale centrée réduite. Pour I_1, on a :

$$I_1 = \int_{-\infty}^{+\infty} x e^{-\frac{x^2}{2}}\, dx = \left[-e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} = 0.$$

2. Pour tout $n \in \mathbb{N}$, on peut écrire :

$$I_{n+2} = \int_{-\infty}^{+\infty} x^{n+2} e^{-\frac{x^2}{2}}\, dx = \int_{-\infty}^{+\infty} (x^{n+1})(x e^{-\frac{x^2}{2}})\, dx,$$

et on effectue une intégration par parties :

$$I_{n+2} = \left[-x^{n+1} e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} (n+1) x^n e^{-\frac{x^2}{2}}\, dx = (n+1) I_n,$$

la dernière égalité venant du fait que l'exponentielle l'emporte sur la puissance :

$$\lim_{x \to +\infty} x^{n+1} e^{-\frac{x^2}{2}} = \lim_{x \to -\infty} x^{n+1} e^{-\frac{x^2}{2}} = 0.$$

3. Puisque $I_1 = 0$, on en déduit que $I_3 = 0$, puis que $I_5 = 0$, et de proche en proche il est clair que $I_{2n+1} = 0$ pour tout $n \in \mathbb{N}$. Ce résultat était d'ailleurs clair sans calculs puisqu'on intègre une fonction impaire sur un domaine symétrique par rapport à 0.

4. Pour les indices pairs, on a $I_2 = 1 \times I_0 = \sqrt{2\pi}$, puis $I_4 = 3 \times I_2 = 3 \times 1 \times I_0 = 3\sqrt{2\pi}$, et de proche en proche :

$$I_{2n} = (2n-1) \times (2n-3) \times \cdots \times 3 \times 1 \times I_0 = \frac{(2n)!}{2^n n!} \sqrt{2\pi}.$$

5. Pour déterminer $\mathbb{E}[X^4]$, il y a deux méthodes équivalentes.
   – Méthode analytique : on écrit l'espérance sous forme d'intégrale :
$$\mathbb{E}[X^4] = \int_{-\infty}^{+\infty} \frac{x^4}{\sqrt{2\pi}} e^{-\frac{(x-1)^2}{2}}\, dx,$$
   et on effectue le changement de variable $u = x - 1$, ce qui donne :
$$\mathbb{E}[X^4] = \int_{-\infty}^{+\infty} \frac{(u+1)^4}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}\, du.$$
   On utilise la formule du binôme : $(u+1)^4 = u^4 + 4u^3 + 6u^2 + 4u + 1$, et on peut alors tout exprimer en fonction des I_n :
$$\mathbb{E}[X^4] = \frac{1}{\sqrt{2\pi}}(I_4 + 4I_3 + 6I_2 + 4I_1 + I_0) = 10.$$
   – Méthode probabiliste : l'idée est la même, puisqu'on sait que si $X \sim \mathcal{N}(1,1)$, alors $Y = X - 1 \sim \mathcal{N}(0,1)$. Donc, par les calculs faits avant, on sait que $\mathbb{E}[Y] = \mathbb{E}[Y^3] = 0$, $\mathbb{E}[Y^2] = 1$ et $E[Y^4] = 3$. Or on a : $\mathbb{E}[X^4] = \mathbb{E}[(Y+1)^4] = \mathbb{E}[Y^4] + 4\mathbb{E}[Y^3] + 6\mathbb{E}[Y^2] + 4\mathbb{E}[Y] + 1 = 3 + 6 + 1 = 10$.

Exercice 3.18 (Vitesse d'une molécule)
Afin de se ramener à la loi normale, posons $\sigma^2 = (kT)/m$, de sorte que la densité en question s'écrit
$$f(x) = ax^2 e^{-\frac{x^2}{2\sigma^2}} \mathbb{1}_{\{x \geq 0\}},$$
et par un argument de parité évident, il vient
$$\int_0^{+\infty} f(x)dx = \int_0^{+\infty} ax^2 e^{-\frac{x^2}{2\sigma^2}} dx = \frac{1}{2}\int_{-\infty}^{+\infty} ax^2 e^{-\frac{x^2}{2\sigma^2}} dx = \frac{a\sqrt{2\pi\sigma^2}}{2}\int_{-\infty}^{+\infty} x^2 \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx,$$
écriture qui fait apparaître le moment d'ordre 2 d'une variable X suivant une loi normale centrée et d'écart-type σ, lequel ne pose pas problème puisqu'il est égal à sa variance :
$$\sigma^2 = \mathrm{Var}(X) = \mathbb{E}[X^2] = \int_{-\infty}^{+\infty} x^2 \frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} dx.$$
Ainsi, puisque f est une densité :
$$1 = \int_0^{+\infty} f(x)dx = \frac{a\sqrt{2\pi\sigma^2}}{2} \times \sigma^2 \;\Leftrightarrow\; a = \sqrt{\frac{2}{\pi}}\sigma^{-3} = \sqrt{\frac{2}{\pi}}\left(\frac{m}{kT}\right)^{\frac{3}{2}}.$$

Exercice 3.19 (Loi log-normale)
1. Puisque $X = e^Y$, on commence par remarquer que X ne prend que des valeurs positives, si bien que si l'on convient de noter F la fonction de répartition de X, il s'ensuit que $F(x) = 0$ pour tout $x \leq 0$. Prenons maintenant $x > 0$, alors :
$$F(x) = \mathbb{P}(X \leq x) = \mathbb{P}(e^Y \leq x) = \mathbb{P}(Y \leq \ln x) = \Phi(\ln x),$$
   où Φ désigne comme d'habitude la fonction de répartition de la loi normale centrée réduite.
2. La densité f de la variable X s'obtient alors en dérivant cette fonction de répartition et en utilisant le fait que Φ' correspond à la densité de la gaussienne standard. Pour tout $x > 0$:
$$f(x) = F'(x) = \frac{1}{x}\Phi'(\ln x) = \frac{1}{x\sqrt{2\pi}} e^{-\frac{\ln^2 x}{2}}$$
   et bien entendu $f(x) = 0$ pour tout $x \leq 0$. Cette densité est représentée figure 3.17.

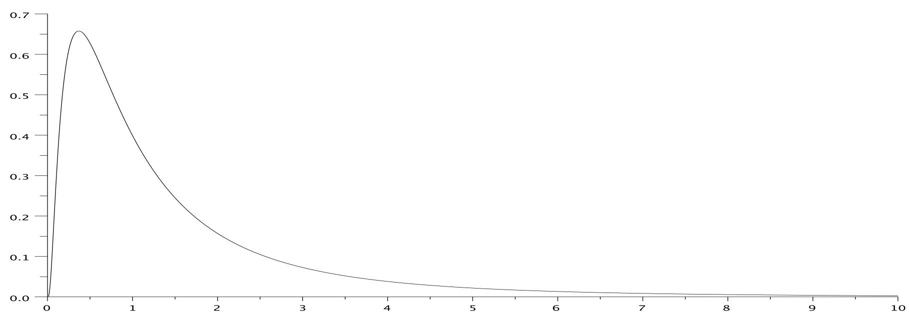


FIGURE 3.17 – Densité d'une loi log-normale.

3. Pour calculer la moyenne de X, on applique le théorème de transfert et on bricole un peu pour se ramener à la densité d'une gaussienne :

$$\mathbb{E}[X] = \mathbb{E}[e^Y] = \int_{-\infty}^{+\infty} e^y \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = \sqrt{e} \int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-1)^2}{2}}}{\sqrt{2\pi}} dy$$

et on reconnaît à l'intérieur de l'intégrale la densité d'une loi normale réduite et de moyenne 1, donc cette intégrale vaut 1 et $\mathbb{E}[X] = \sqrt{e}$. Le calcul du moment d'ordre 2 se fait suivant le même principe :

$$\mathbb{E}[X^2] = \mathbb{E}[e^{2Y}] = \int_{-\infty}^{+\infty} e^{2y} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = e^2 \int_{-\infty}^{+\infty} \frac{e^{-\frac{(y-2)^2}{2}}}{\sqrt{2\pi}} dy$$

et on reconnaît à l'intérieur de l'intégrale la densité d'une loi normale réduite et de moyenne 2, donc cette intégrale vaut 1 et $\mathbb{E}[X^2] = e^2$, ce qui donne bien $\mathrm{Var}(X) = e(e-1)$.

4. Il s'agit ici de calculer la probabilité que X soit inférieure à $0,5$. Le plan de vol est limpide : on commence par se ramener à une loi normale $\mathcal{N}(-0,5; 0,09)$, laquelle est ensuite centrée et réduite :

$$\mathbb{P}(X < 0,5) = \mathbb{P}(\ln X < \ln 0,5) = \mathbb{P}(Y < -\ln 2) = \mathbb{P}\left(\frac{Y+0,5}{0,3} < \frac{0,5-\ln 2}{0,3}\right),$$

c'est-à-dire, puisque $0,5 - \ln 2 \approx -0,19 < 0$:

$$\mathbb{P}(X < 0,5) = \Phi\left(\frac{0,5-\ln 2}{0,3}\right) = 1 - \Phi\left(\frac{\ln 2 - 0,5}{0,3}\right) \approx 1 - \Phi(0,64) \approx 0,26.$$

En moyenne, avec ce modèle, 26% des grains de sable passent à travers le tamis.

Exercice 3.20 (La Belle de Fontenay)
Dans tout l'exercice, Φ est la fonction de répartition de la gaussienne standard.

1. La probabilité qu'une pomme de terre pèse plus de 250 grammes s'écrit

$$\mathbb{P}(X > 250) = \mathbb{P}\left(\frac{X-200}{70} > \frac{250-200}{70}\right) = 1 - \Phi(5/7) \approx 0.24$$

2. De même, la probabilité qu'une pomme de terre pèse moins de 180 grammes est

$$\mathbb{P}(X < 180) = \mathbb{P}\left(\frac{X-200}{70} < \frac{180-200}{70}\right) = \Phi(-2/7) = 1 - \Phi(2/7) \approx 0.39$$

3. Enfin, la probabilité qu'elle ait une masse comprise entre 190 et 210 grammes vaut

$$\mathbb{P}(190 \leq X \leq 210) = \mathbb{P}\left(\frac{190-200}{70} \leq \frac{X-200}{70} \leq \frac{210-200}{70}\right)$$

d'où

$$\mathbb{P}(190 \leq X \leq 210) = \Phi(1/7) - \Phi(-1/7) = 2\Phi(1/7) - 1 \approx 0.11$$

Exercice 3.21 (Quantile et variance)

1. Avant tout calcul, notons qu'il est clair que q est supérieure à la moyenne 12 de cette loi normale. Plus précisément

$$\mathbb{P}(X > q) = 0.1 \Leftrightarrow \mathbb{P}(X \leq q) = 0.9 \Leftrightarrow \mathbb{P}\left(\frac{X-12}{2} \leq \frac{q-12}{2}\right) = 0.9$$

ce qui est encore dire que

$$\mathbb{P}(X > q) = 0.1 \Leftrightarrow \Phi\left(\frac{q-12}{2}\right) = 0.9 \Leftrightarrow \frac{q-12}{2} \approx 1.28 \Leftrightarrow q \approx 14.56$$

2. On peut y aller à fond de cinquième :

$$\mathbb{P}(X > 9) = 0.2 \Leftrightarrow \mathbb{P}(X \leq 9) = 0.8 \Leftrightarrow \mathbb{P}\left(\frac{X-5}{\sigma} \leq \frac{4}{\sigma}\right) = 0.8 \Leftrightarrow \frac{4}{\sigma} \approx 0.84$$

donc $\sigma \approx 4.76$.

Exercice 3.22 (Répartition des tailles)

Notons X la variable aléatoire correspondant à la taille (en centimètres) d'un homme choisi au hasard. D'après ce modèle, on a donc $X \sim \mathcal{N}(175, 6^2)$.

1. Le pourcentage d'hommes ayant une taille supérieure à 1m85 est donc

$$\mathbb{P}(X > 185) = \mathbb{P}(\frac{X-185}{6} > \frac{175-185}{6}) = \Phi(-5/3) = 1 - \Phi(5/3) \approx 0.05$$

2. Parmi les hommes mesurant plus de 1m80, la proportion mesurant plus de 1m92 s'écrit

$$\mathbb{P}(X > 192 | X > 180) = \frac{\mathbb{P}(\{X > 192\} \cap \{X > 180\})}{\mathbb{P}(X > 180)} = \frac{\mathbb{P}(X > 192)}{\mathbb{P}(X > 180)}$$

et il suffit alors de mener les calculs comme dans la question précédente pour voir que

$$\mathbb{P}(X > 192 | X > 180) = \frac{1 - \Phi(17/6)}{1 - \Phi(5/6)} \approx \frac{1 - 0.9977}{1 - 0.7967} \approx 0.01.$$

Exercice 3.23 (Choix de machine)

Notons X (respectivement Y) la variable aléatoire correspondant à la longueur d'une pièce (en mm) produite par la machine A (respectivement B). Le texte spécifie que $X \sim \mathcal{N}(8; 4)$ et $Y \sim \mathcal{N}(7.5; 1)$. Pour savoir quelle machine il vaut mieux choisir, il suffit de comparer $\mathbb{P}(7 \leq X \leq 9)$ à $\mathbb{P}(7 \leq Y \leq 9)$, ce qui donne respectivement

$$\mathbb{P}(7 \leq X \leq 9) = \mathbb{P}\left(\frac{7-8}{2} \leq \frac{X-8}{2} \leq \frac{9-8}{2}\right) = \Phi(1/2) - \Phi(-1/2) = 2\Phi(1/2) - 1 \approx 0.38$$

et

$$\mathbb{P}(7 \leq Y \leq 9) = \mathbb{P}\left(-\frac{1}{2} \leq \frac{Y-7.5}{1} \leq \frac{3}{2}\right) = \Phi(3/2) - \Phi(-1/2) = \Phi(3/2) + \Phi(1/2) - 1 \approx 0.62$$

Il est donc évident qu'il faut choisir la machine B, même si en moyenne elle ne fait pas des pièces de 8 mm.

Exercice 3.24 (Approximation gaussienne)
Soit X le nombre de Pile obtenus en 400 lancers d'une pièce équilibrée.

1. Pour tout i entre 1 et 400, notons X_i la variable valant 1 si le i-ème lancer donne Pile, 0 s'il donne Face. Les variables $X_1, \dots, X_{400}$ sont donc indépendantes et identiquement distribuées suivant la loi de Bernoulli de paramètre 1/2. Avec ces notations, il est clair que $X = X_1 + \cdots + X_{400}$ correspond au nombre de Pile obtenus sur les 400 lancers et nous sommes typiquement dans le cadre d'application du Théorème Central Limite : la variable $(X - 400\mathbb{E}[X_1])/\sqrt{400\mathrm{Var}(X_1)}$ suit approximativement une loi normale centrée réduite. Puisque $\mathbb{E}[X_1] = 1/2$ et $\mathrm{Var}(X_1) = 1/4$, ceci signifie que
$$\frac{X-200}{10} \approx \mathcal{N}(0,1) \;\Leftrightarrow\; X \approx \mathcal{N}(200, 10^2)$$
Avec cette approximation, la quantité $\mathbb{P}(190 \leq X \leq 210)$ est la probabilité qu'une loi normale ne s'éloigne pas de plus d'un écart-type de sa moyenne, c'est-à-dire environ 68%.
2. Le calcul de $\mathbb{P}(210 \leq X \leq 220)$ se fait comme d'habitude par centrage et réduction :
$$\mathbb{P}(210 \leq X \leq 220) = \mathbb{P}\left(1 \leq \frac{X-200}{10} \leq 2\right) = \Phi(2) - \Phi(1) \approx 0.14$$
3. Notons Y le nombre de Pile obtenus en 400 lancers pour une pièce biaisée où $\mathbb{P}(\text{Pile}) = 0.51$. L'approximation normale donne cette fois
$$\frac{Y-204}{9.9998} \approx \mathcal{N}(0,1) \;\Leftrightarrow\; Y \approx \mathcal{N}(200, 9.998^2)$$
ou en arrondissant
$$\frac{Y-204}{10} \approx \mathcal{N}(0,1) \;\Leftrightarrow\; Y \approx \mathcal{N}(200, 10^2)$$
c'est-à-dire que l'écart-type est quasiment le même que pour X, mais la moyenne diffère un peu. Autrement dit, si l'on trace les densités de X et Y, celle de Y est la même que celle de

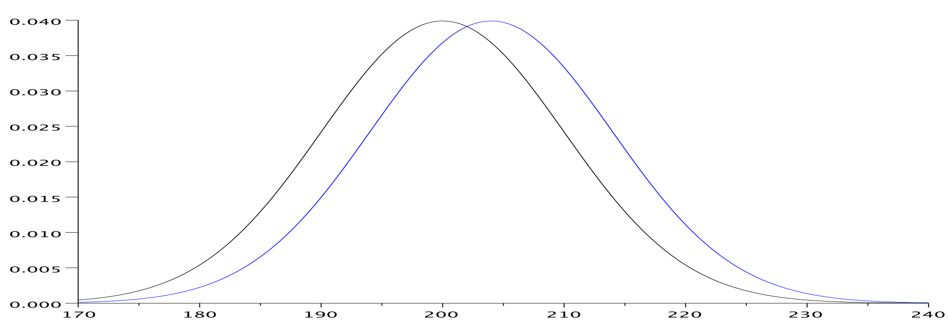


FIGURE 3.18 – Densités des lois $\mathcal{N}(200, 10^2)$ et $\mathcal{N}(204, 10^2)$.

X mais translatée de +4 selon l'axe des abscisses (cf. figure 3.18), ce qui est cohérent avec les résultats suivants :
$$\mathbb{P}(190 \leq Y \leq 210) = \mathbb{P}\left(-1.4 \leq \frac{Y-204}{10} \leq 0.6\right) = \Phi(0.6) - \Phi(-1.4) \approx 0.645$$
donc $\mathbb{P}(190 \leq Y \leq 210) < \mathbb{P}(190 \leq X \leq 210)$. A contrario
$$\mathbb{P}(210 \leq Y \leq 220) = \mathbb{P}\left(0.6 \leq \frac{Y-204}{10} \leq 1.6\right) = \Phi(1.6) - \Phi(0.6) \approx 0.22$$
donc $\mathbb{P}(210 \leq Y \leq 220) > \mathbb{P}(210 \leq X \leq 220)$.

Exercice 3.25 (Sondage)

1. Puisque le sondage est fait avec remise, la loi suivie par X est binomiale de paramètres n et p. Notons que si le sondage était fait sans remise dans une population totale de taille N (comme c'est le cas en pratique, puisqu'on n'interroge pas deux fois la même personne), ce serait une loi hypergéométrique $\mathcal{H}(N, n, p)$ (voir les exercices 2.4 et 2.10). Cependant, et comme déjà mentionné, dès que n est négligeable devant N, ces deux lois sont très proches l'une de l'autre : en d'autres termes, si le nombre de sondés est très faible par rapport à la population totale, il y a très peu de chances d'interroger deux fois la même personne lorsqu'on effectue un tirage avec remise.
2. Il est implicite ici que n est grand et p pas trop proche de 0, de sorte que nous pouvons approcher la loi binomiale de paramètres n et p par une loi normale de mêmes espérance et variance, à savoir np et $np(1-p)$:
$$\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p))$$
Or une loi normale se concentre à 95% dans un intervalle centré en sa moyenne et de rayon égal à 1,96 fois l'écart-type, soit :
$$\mathbb{P}(X \in [np - 1.96\sqrt{np(1-p)}, np + 1.96\sqrt{np(1-p)}]) = 0.95$$
3. Un estimateur naturel $\hat{p}$ de p est tout bonnement la proportion empirique d'électeurs favorables au candidat, c'est-à-dire $\hat{p} = X/n$. Puisque $\mathbb{E}[X] = np$, on a $\mathbb{E}[\hat{p}] = p$. On dit que $\hat{p}$ est un estimateur non biaisé de p, c'est-à-dire qu'en moyenne, avec cet estimateur, on ne se trompe pas.
4. On a vu qu'avec 95% de chances
$$np - 1.96\sqrt{np(1-p)} \leq X \leq np + 1.96\sqrt{np(1-p)}$$
d'où l'on déduit aussitôt qu'avec 95% de chances
$$p - 1.96\sqrt{\frac{p(1-p)}{n}} \leq \hat{p} = \frac{X}{n} \leq p + 1.96\sqrt{\frac{p(1-p)}{n}}$$
5. L'étude de la fonction $x \mapsto x(1-x)$ permet de voir que pour tout $x \in [0,1]$, $0 \leq x(1-x) \leq 1/4$, maximum atteint pour $x = 1/2$. De cette majoration et de la question précédente, on déduit qu'avec 95% de chances
$$p - \frac{0.98}{\sqrt{n}} \leq p - 1.96\sqrt{\frac{p(1-p)}{n}} \leq \hat{p} \leq p + 1.96\sqrt{\frac{p(1-p)}{n}} \leq p + \frac{0.98}{\sqrt{n}}$$
d'où un intervalle de confiance à 95% pour p
$$\hat{p} - \frac{0.98}{\sqrt{n}} \leq p \leq \hat{p} + \frac{0.98}{\sqrt{n}}$$
6. La taille de cet intervalle de confiance à 95% lorsqu'on interroge n personnes est $1.96/\sqrt{n}$ donc pour 1000 personnes, il est de diamètre 0.062, soit une marge d'erreur de $\pm 3.1\%$.
7. Pour savoir combien de personnes interroger pour obtenir un intervalle de confiance à 95% de rayon 2%, il suffit de résoudre
$$\frac{0.98}{\sqrt{n}} \leq 0.02 \;\Leftrightarrow\; n \geq 2401$$
Il faut sonder environ 2400 personnes pour atteindre cette précision. Notons qu'en pratique, ce n'est pas ce type de sondage qui est utilisé, mais plutôt des sondages stratifiés ou par quotas.

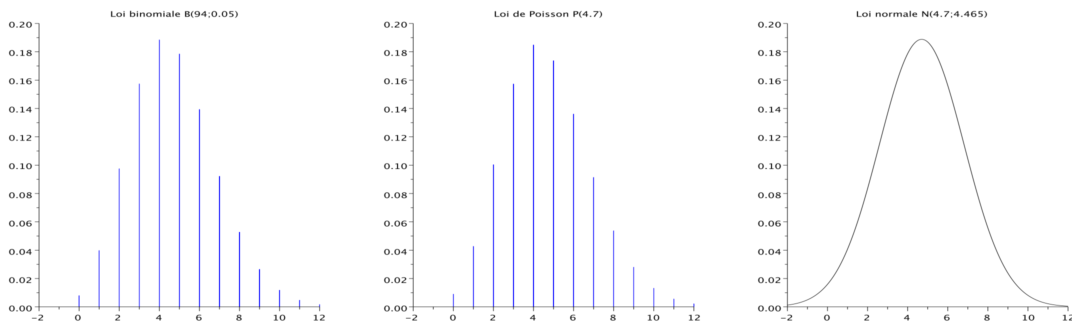


FIGURE 3.19 – Lois binomiale, de Poisson et normale pour le surbooking.

1. Le nombre d'absents à l'embarquement est une variable aléatoire S qui suit une loi binomiale $\mathcal{B}(94, 0.05)$. Puisque $\mathbb{E}[S] = 4.7$ et $\mathrm{Var}(S) = 4.465$, on peut être tenté d'approcher cette loi binomiale par une loi normale $\mathcal{N}(4.7; 4.465)$ pour estimer le nombre d'absents et la probabilité qu'il y ait trop de monde à l'embarquement. La figure 3.19 représente les lois binomiale, de Poisson et normale. Puisqu'une loi normale prend ses valeurs dans $\mathbb{R}$ et non dans $\mathbb{N}$, on coupe la poire en deux et on estime cette probabilité par

$$\mathbb{P}(\mathcal{N}(4.7; 4.465) \leq 3.5) = 1 - \Phi\left(\frac{4.7 - 3.5}{\sqrt{4.465}}\right) \approx 1 - 0.7157 = 0.284$$

2. La probabilité réelle qu'il y ait trop de monde à l'embarquement est en fait :

$$\mathbb{P}(S \leq 3) = \mathbb{P}(S = 0) + \cdots + \mathbb{P}(S = 3) = \binom{94}{0}\left(\frac{5}{100}\right)^0\left(\frac{95}{100}\right)^{94} + \cdots + \binom{94}{3}\left(\frac{5}{100}\right)^3\left(\frac{95}{100}\right)^{91}$$

donc $\mathbb{P}(S \leq 3) \approx 0.303$. L'approximation par une loi normale aboutit donc à une erreur relative de $(0.303 - 0.284)/0.303 \approx 6.3\%$. La figure 3.20 illustre ce phénomène. Par ailleurs, nous avons vu que l'approximation de la loi binomiale $\mathcal{B}(94, 0.05)$ par une loi de Poisson $\mathcal{P}(94 \times 0.05) = \mathcal{P}(4, 7)$ donne comme estimation :

$$\mathbb{P}(\mathcal{P}(4, 7) \leq 3) = e^{-4,7}\frac{4, 7^0}{0!} + \cdots + e^{-4,7}\frac{4, 7^3}{3!} \approx 0, 310$$

soit une erreur relative de 2.3%.

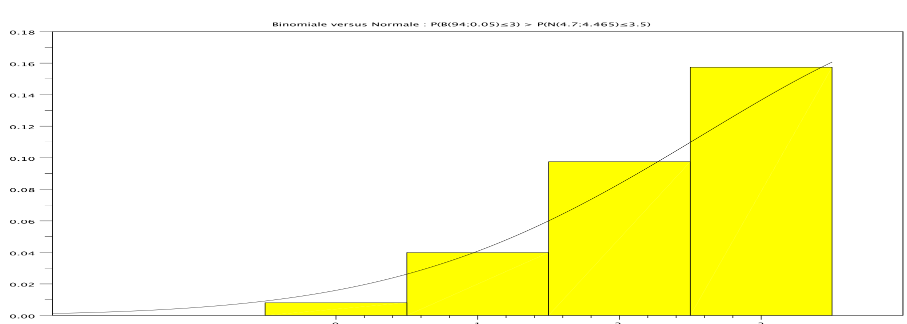


FIGURE 3.20 – Comparaison des probabilités de problème à l'embarquement.

Exercice 3.27 (Queue de la gaussienne)
On appelle fonction de Marcum, ou queue de la gaussienne, la fonction définie pour tout réel x par :

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{t^2}{2}}\, dt.$$

1. Pour tout réel x, on a $F(x) = 1 - Q(x)$.
2. Soit $x > 0$ fixé. Le changement de variable $t = x + u$ et le fait que $e^{-ux} \leq 1$ pour x et u positifs donne
$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{(x+u)^2}{2}}\, du = e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-ux} e^{-\frac{u^2}{2}}\, du \leq e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{u^2}{2}}\, du$$
et on aura reconnu la densité de la gaussienne standard
$$\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{u^2}{2}}\, du = \frac{1}{2}\left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{u^2}{2}}\, du\right) = \frac{1}{2}$$
ce qui donne bien pour tout x positif
$$Q(x) \leq \frac{1}{2} e^{-\frac{x^2}{2}}.$$
3. Pour $t \geq x > 0$, on a
$$1 + \frac{1}{t^2} \leq 1 + \frac{1}{x^2} \;\Rightarrow\; \frac{1 + \frac{1}{t^2}}{1 + \frac{1}{x^2}} \leq 1$$
L'inégalité de droite est encore plus évidente.
4. On en déduit alors
$$\frac{1}{(1 + \frac{1}{x^2})\sqrt{2\pi}} \int_x^{+\infty} \left(1 + \frac{1}{t^2}\right) e^{-\frac{t^2}{2}}\, dt \leq \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} 1 \times e^{-\frac{t^2}{2}}\, dt \leq \frac{1}{x\sqrt{2\pi}} \int_x^{+\infty} t e^{-\frac{t^2}{2}}\, dt$$
5. Pour tout réel non nul t
$$\left(\frac{1}{t} e^{-\frac{t^2}{2}}\right)' = -\left(1 + \frac{1}{t^2}\right) e^{-\frac{t^2}{2}}$$
Ainsi
$$\int_x^{+\infty} \left(1 + \frac{1}{t^2}\right) e^{-\frac{t^2}{2}}\, dt = \left[-\frac{1}{t} e^{-\frac{t^2}{2}}\right]_x^{+\infty} = \frac{1}{x} e^{-\frac{x^2}{2}}$$
et l'inégalité de gauche est acquise. Celle de droite est encore plus simple puisque
$$\int_x^{+\infty} t e^{-\frac{t^2}{2}}\, dt = \left[-e^{-\frac{t^2}{2}}\right]_x^{+\infty} = e^{-\frac{x^2}{2}}$$
Au total, on a bien montré que pour tout $x > 0$
$$\frac{1}{(1 + \frac{1}{x^2}) x\sqrt{2\pi}} e^{-\frac{x^2}{2}} \leq Q(x) \leq \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$
6. Cet encadrement permet de voir que
$$\frac{Q(x)}{\frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}} \xrightarrow[x \to +\infty]{} 1$$
d'où un équivalent très simple de $Q(x)$ lorsque x tend vers $+\infty$:
$$Q(x) \sim \frac{1}{x\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

7. Application

   (a) Si le symbole d'entrée est $+\sqrt{E_b}$ (respectivement $-\sqrt{E_b}$), alors $Y \sim \mathcal{N}(+\sqrt{E_b}, \frac{N_0}{2})$ (respectivement $Y \sim \mathcal{N}(-\sqrt{E_b}, \frac{N_0}{2})$). De façon générale, $Y = X + B$ où B est le bruit additif, supposé gaussien centré de variance $N_0/2$ et indépendant de X, variable aléatoire binaire correspondant au symbole d'entrée.

   (b) Intuitivement, on se dit que le symbole d'entrée était plus vraisemblablement $+\sqrt{E_b}$ (respectivement $-\sqrt{E_b}$) si la sortie y est positive (respectivement négative). Cette règle est en effet la bonne si les symboles d'entrée sont équiprobables, c'est-à-dire si $\mathbb{P}(X = +\sqrt{E_b}) = \mathbb{P}(X = +\sqrt{E_b}) = 1/2$. Il suffit de comparer les probabilités conditionnelles pour s'en convaincre. Il convient juste d'adapter la formule de Bayes et celle des probabilités totales au cas d'un cocktail entre loi discrète et loi à densité, ce qui donne ici :

$$\mathbb{P}(X = +\sqrt{E_b}|y) = \frac{f(y|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b})}{f(y)}$$

d'où

$$\mathbb{P}(X = +\sqrt{E_b}|y) = \frac{f(y|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b})}{f(y|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b}) + f(y|X = -\sqrt{E_b})\mathbb{P}(X = -\sqrt{E_b})}$$

Il reste à tenir compte du fait que les symboles d'entrée sont équiprobables et des densités respectives de la réponse Y connaissant X pour obtenir

$$\mathbb{P}(X = +\sqrt{E_b}|y) = \frac{e^{-\frac{(y-\sqrt{E_b})^2}{N_0}}}{e^{-\frac{(y-\sqrt{E_b})^2}{N_0}} + e^{-\frac{(y+\sqrt{E_b})^2}{N_0}}} = \frac{1}{1 + e^{-4\frac{\sqrt{E_b}}{N_0}y}} \tag{3.2}$$

On en déduit automatiquement :

$$\mathbb{P}(X = -\sqrt{E_b}|y) = 1 - \mathbb{P}(X = +\sqrt{E_b}|y) = \frac{e^{-4\frac{\sqrt{E_b}}{N_0}y}}{1 + e^{-4\frac{\sqrt{E_b}}{N_0}y}}$$

et par suite

$$\frac{\mathbb{P}(X = +\sqrt{E_b}|y)}{\mathbb{P}(X = -\sqrt{E_b}|y)} = e^{4\frac{\sqrt{E_b}}{N_0}y}$$

de sorte que ce rapport est supérieur à 1 si et seulement si y est positif, et la règle de décision au maximum de vraisemblance correspond bien à la règle intuititive donnée ci-dessus.

Remarque : si les symboles d'entrée ne sont pas équiprobables, il faut en tenir compte dans la règle de décision. Supposons par exemple que $\mathbb{P}(X = +\sqrt{E_b}) = 3/4$, alors l'équation (3.2) devient

$$\mathbb{P}(X = +\sqrt{E_b}|y) = \frac{3}{3 + e^{-4\frac{\sqrt{E_b}}{N_0}y}}$$

et

$$\frac{\mathbb{P}(X = +\sqrt{E_b}|y)}{\mathbb{P}(X = -\sqrt{E_b}|y)} = 3\ e^{4\frac{\sqrt{E_b}}{N_0}y}$$

Ainsi on décide que le symbole d'entrée était $X = +\sqrt{E_b}$ si

$$3\ e^{4\frac{\sqrt{E_b}}{N_0}y} > 1 \iff y > \tau = \frac{-\ln 3}{4} \times \frac{N_0}{\sqrt{E_b}}$$

Ces résultats admettent une interprétation graphique très simple : les points d'abscisses 0 et τ sont les points d'intersection des fonctions $y \mapsto f(y|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b})$ et $y \mapsto f(y|X = -\sqrt{E_b})\mathbb{P}(X = -\sqrt{E_b})$ respectivement lorsque $\mathbb{P}(X = +\sqrt{E_b}) = 1/2$ et $\mathbb{P}(X = +\sqrt{E_b}) = 3/4$ (voir figure 3.21 dans le cas où $N_0 = \sqrt{E_b} = 1$, d'où en particulier $\tau = -\ln 3/4 \approx -0.27$).

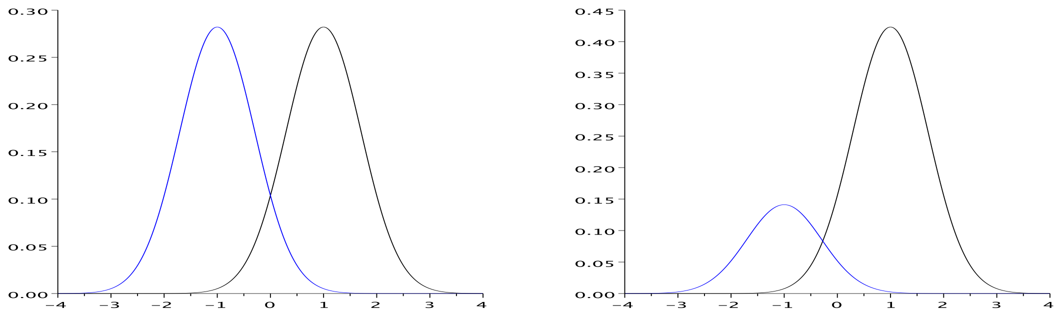


FIGURE 3.21 – Fonctions $y \mapsto f(y|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b})$ et $y \mapsto f(y|X = -\sqrt{E_b})\mathbb{P}(X = -\sqrt{E_b})$ lorsque $\mathbb{P}(X = +\sqrt{E_b}) = 1/2$ (à gauche) et $\mathbb{P}(X = +\sqrt{E_b}) = 3/4$ (à droite).

(c) Dans le cas où les symboles d'entrée sont équiprobables, la probabilité d'erreur P_e est égale à la somme de la probabilité de décider $+\sqrt{E_b}$ alors que le symbole d'entrée était $-\sqrt{E_b}$ et vice-versa :

$$P_e = \mathbb{P}(Y > 0|X = -\sqrt{E_b})\mathbb{P}(X = -\sqrt{E_b}) + \mathbb{P}(Y < 0|X = +\sqrt{E_b})\mathbb{P}(X = +\sqrt{E_b})$$

et par symétrie des rôles, en notant toujours B le bruit additif :

$$P_e = \mathbb{P}(B > \sqrt{E_b}) = \mathbb{P}\left(\frac{B}{\sqrt{N_0/2}} > \frac{\sqrt{E_b}}{\sqrt{N_0/2}}\right) = Q\left(\sqrt{\frac{2E_b}{N_0}}\right).$$

Exercice 3.28 (Entropie d'une variable aléatoire)
Si X est une variable aléatoire réelle admettant une densité f, on appelle entropie de X la quantité (si elle est définie) :

$$h(X) = \mathbb{E}[-\ln f(X)] = -\int_{-\infty}^{+\infty} f(x) \ln f(x)\, dx.$$

1. Si $X \sim \mathcal{N}(0, 1)$, alors son entropie s'écrit

$$h(X) = -\int_{-\infty}^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \ln\left(\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right) dx = \frac{\ln(2\pi)}{2}\int_{-\infty}^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx + \frac{1}{2}\int_{-\infty}^{+\infty} x^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx$$

or

$$\int_{-\infty}^{+\infty} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = 1 \qquad \text{et} \qquad \int_{-\infty}^{+\infty} x^2 \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\, dx = \mathbb{E}[X^2] = \mathrm{Var}(X) = 1$$

d'où en effet

$$h(X) = \frac{1}{2}(1 + \ln(2\pi)).$$

2. Le même calcul que ci-dessus montre que si $X \sim \mathcal{N}(m, \sigma^2)$, alors elle a pour entropie : $h(X) = \frac{1}{2}(1 + \log(2\pi\sigma^2))$.

3. Soit donc $X_1 \sim \mathcal{N}(0, \sigma^2)$, dont la densité est notée φ, et X_2 une variable aléatoire centrée de densité f et de variance σ^2, c'est-à-dire que :
$$\int_{-\infty}^{+\infty} x^2 f(x)\,dx = \sigma^2.$$
On suppose pour simplifier que f est strictement positive sur $\mathbb{R}$.
   (a) Sous réserve d'existence des intégrales, par définition de l'entropie
$$h(X_2) = -\int_{-\infty}^{+\infty} f(x)\ln f(x)\,dx = \int_{-\infty}^{+\infty} f(x)\left(\ln\frac{\varphi(x)}{f(x)} - \ln\varphi(x)\right)\,dx$$
   ce qui donne bien
$$h(X_2) = \int_{-\infty}^{+\infty} f(x)\ln\frac{\varphi(x)}{f(x)}\,dx - \int_{-\infty}^{+\infty} f(x)\ln\varphi(x)\,dx.$$
   (b) Pour montrer que pour tout $x > 0$, $\ln x \leq x - 1$, il suffit par exemple d'étudier la fonction $g : x \mapsto x - 1 - \ln x$ sur $]0, +\infty[$. Sa dérivée est $g'(x) = 1 - 1/x$, qui est négative sur $]0, 1]$ et positive sur $[1, +\infty[$. Son minimum est donc $g(1) = 0$, autrement dit g est bien positive sur son domaine de définition. On en déduit que :
$$\int_{-\infty}^{+\infty} f(x)\log\frac{\varphi(x)}{f(x)}\,dx \leq \int_{-\infty}^{+\infty} f(x)\left(\frac{\varphi(x)}{f(x)} - 1\right)\,dx = \int_{-\infty}^{+\infty} \varphi(x)\,dx - \int_{-\infty}^{+\infty} f(x)\,dx$$
   or f et φ étant toutes deux des densités, elles intègrent à 1 et le majorant vaut bien 0.
   (c) On a alors
$$-\int_{-\infty}^{+\infty} f(x)\ln\varphi(x)\,dx = -\int_{-\infty}^{+\infty} f(x)\ln\left(\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}\right)\,dx$$
   qui se calcule sans difficultés
$$-\int_{-\infty}^{+\infty} f(x)\ln\varphi(x)\,dx = \frac{\ln(2\pi)}{2}\int_{-\infty}^{+\infty} f(x)\,dx + \frac{1}{2}\int_{-\infty}^{+\infty} x^2 f(x)\,dx$$
   en ayant en tête que
$$\int_{-\infty}^{+\infty} f(x)\,dx = 1 \qquad \text{et} \qquad \int_{-\infty}^{+\infty} x^2 f(x)\,dx = \mathbb{E}[X_2^2] = \mathrm{Var}(X_2) = \sigma^2$$
   Au total on a bien
$$-\int_{-\infty}^{+\infty} f(x)\log\varphi(x)\,dx = \frac{1}{2}(1 + \ln(2\pi\sigma^2)).$$
   (d) Des trois questions précédentes et du calcul de l'entropie pour une variable gaussienne $X_1 \sim \mathcal{N}(0, \sigma^2)$, on déduit que
$$h(X_2) \leq \frac{1}{2}(1 + \ln(2\pi\sigma^2)) = h(X_1),$$
   c'est-à-dire que, à variance donnée, c'est la loi normale qui réalise le maximum de l'entropie.

Exercice 3.29 (Nul n'est censé ignorer la loi normale)
On note comme d'habitude Φ la fonction de répartition de la loi normale centrée réduite.

1. Commençons par le troisième quartile. Par définition de celui-ci, puis centrage-réduction et lecture dans la table de la loi normale, on a
$$\mathbb{P}(X \leq q_3) = 0.75 \;\Leftrightarrow\; \mathbb{P}\left(\frac{X-20}{5} \leq \frac{q_3-20}{5}\right) = 0.75 \;\Leftrightarrow\; \frac{q_3-20}{5} \approx 0.67 \;\Leftrightarrow\; q_3 \approx 23.35$$
et par symétrie d'une loi normale par rapport à sa moyenne :
$$\frac{q_1+q_3}{2} = 20 \;\Rightarrow\; q_1 \approx 16.65$$

2. Seuls $20/300 = 15\%$ des étudiants seront concernés par le rattrapage, on cherche donc la note x telle que
$$\mathbb{P}(X \leq x) = 0.15 \;\Leftrightarrow\; \mathbb{P}\left(\frac{X-9}{2} \leq \frac{x-9}{2}\right) = 0.15 \;\Leftrightarrow\; \Phi\left(\frac{9-x}{2}\right) = 0.85$$
d'où il sort
$$\mathbb{P}(X \leq x) = 0.15 \;\Leftrightarrow\; \frac{9-x}{2} \approx 1.04 \;\Leftrightarrow\; x \approx 6.92$$
Ainsi, en gros, les étudiants ayant une note inférieure à 7 pourront suivre les cours de rattrapage.

3. On cherche cette fois $\mathbb{P}(X \notin [240, 290]) = 1 - \mathbb{P}(240 \leq X \leq 290)$, où $X \sim \mathcal{N}(270, 10^2)$, or
$$\mathbb{P}(240 \leq X \leq 290) = \mathbb{P}\left(\frac{240-270}{10} \leq \frac{X-270}{10} \leq \frac{290-270}{10}\right) = \Phi(2) - \Phi(-3)$$
c'est-à-dire
$$\mathbb{P}(240 \leq X \leq 290) = \Phi(2) + \Phi(3) - 1 \approx 0.9759$$
Au vu de sa période d'absence à l'étranger, il y a donc environ seulement 2.4% de chances qu'il puisse être le père.

4. (a) Chaque étudiant accepté a une probabilité 1/3 d'être effectivement présent à la rentrée. Les décisions des étudiants étant supposées indépendantes les unes des autres, la loi de X est binomiale : $X \sim \mathcal{B}(500, 1/3)$.

   (b) Nous sommes dans le cadre d'application du théorème central limite : puisque $\mathbb{E}[X] = 500/3$ et $\text{Var}(X) = 1000/9$, nous faisons l'approximation : $X \approx \mathcal{N}(500/3, 1000/9)$, donc
$$\mathbb{P}(X > 200) \approx \mathbb{P}\left(\frac{X-500/3}{\sqrt{1000/9}} > \frac{200-500/3}{\sqrt{1000/9}}\right) \approx 1 - \Phi(3.16) \approx 8 \times 10^{-4}$$
   Si le modèle est bon, il y a donc très peu de risques d'être en sureffectif à la rentrée.

Exercice 3.30 (Loi bêta)
On considère une variable aléatoire X de densité
$$f(x) = \begin{cases} c\,x(1-x) & \text{si } 0 \leq x \leq 1 \\ 0 & \text{ailleurs} \end{cases} \tag{3.2}$$

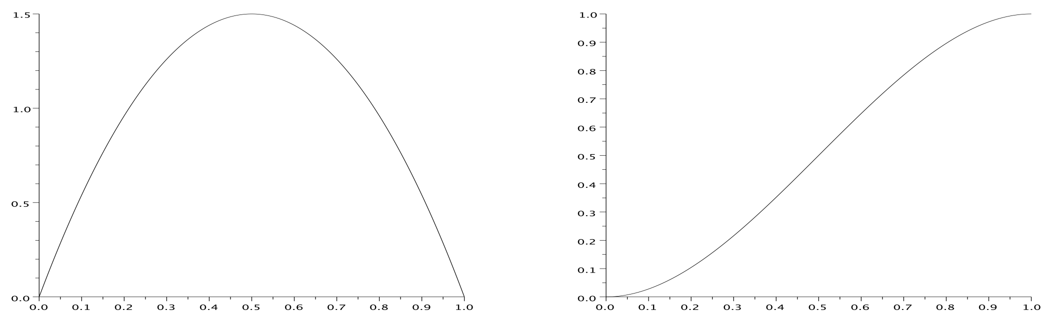


FIGURE 3.22 – Densité $f(x) = 6x(1-x)\mathbb{1}_{[0,1]}(x)$ et fonction de répartition F associée.

1. La constante c doit être positive pour que f le soit et f doit sommer à 1, or :
$$c\int_0^1 x(1-x)dx = c\int_0^1 (x-x^2)dx = c\left[\frac{x^2}{2}-\frac{x^3}{3}\right]_0^1 = \frac{c}{6} \ \Rightarrow \ c = 6$$
et $f(x) = 6x(1-x)\mathbb{1}_{[0,1]}(x)$. Cette densité est représentée figure 3.22 à gauche.
2. La fonction de répartition F est nulle à gauche de 0, vaut 1 à droite de 1, et pour $0 \leq x \leq 1$ un petit calcul s'impose
$$F(x) = \int_{-\infty}^x f(t)dt = \int_0^x 6t(1-t)dt = \left[3t^2 - 2t^3\right]_0^x = 3x^2 - 2x^3.$$
Cette fonction de répartition est représentée figure 3.22 à droite.
3. On en déduit immédiatement que $\mathbb{P}(1/4 < X < 3/4) = F(3/4) - F(1/4) = 11/16$.
4. L'espérance de X vaut
$$\mathbb{E}[X] = \int_{-\infty}^{+\infty} xf(x)dx = \int_0^1 6x^2(1-x)dx = \left[2x^3 - \frac{3}{2}x^4\right]_0^1 = \frac{1}{2},$$
ce qui était évident par symétrie de la densité autour de 1/2. Pour le calcul de la variance, on commence par le moment d'ordre 2 :
$$\mathbb{E}[X^2] = \int_0^1 6x^3(1-x)dx = \left[\frac{3}{2}x^4 - \frac{6}{5}x^5\right]_0^1 = \frac{3}{10},$$
d'où
$$\operatorname{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{3}{10} - \frac{1}{4} = \frac{1}{20}.$$
5. L'inégalité de Tchebychev permet de majorer la probabilité qu'a une variable aléatoire de s'éloigner de sa moyenne en fonction de sa variance. Précisément, elle nous assure que pour tout $t > 0$
$$\mathbb{P}\left(|X - \mathbb{E}[X]| \geq t\right) \leq \frac{\operatorname{Var}(X)}{t^2} \ \Leftrightarrow \ \mathbb{P}\left(-t < X - \mathbb{E}[X] < t\right) \geq 1 - \frac{\operatorname{Var}(X)}{t^2}$$
Puisque $\mathbb{E}[X] = 1/2$, on l'applique ici avec $t = 1/4$, ce qui donne
$$\mathbb{P}\left(\frac{1}{4} < X < \frac{3}{4}\right) = \mathbb{P}\left(-\frac{1}{4} < X - \frac{1}{2} < \frac{1}{4}\right) \geq 1 - \frac{\operatorname{Var}(X)}{(1/4)^2} = \frac{1}{5}.$$
Bien entendu, cette borne est inférieure à la vraie valeur 11/16 trouvée précédemment.

6. Pour tout $n \in \mathbb{N}^*$, le moment d'ordre n de X vaut
$$\mathbb{E}[X^n] = \int_0^1 6x^{n+1}(1-x)dx = 6\left[\frac{x^{n+2}}{n+2} - \frac{x^{n+3}}{n+3}\right]_0^1 = \frac{6}{(n+2)(n+3)}.$$
On vérifie bien sûr que ceci coïncide avec les résultats trouvés pour $\mathbb{E}[X]$ et $\mathbb{E}[X^2]$.

Exercice 3.31 (Loi de Rayleigh)
On considère une variable aléatoire X de densité
$$f(x) = \begin{cases} c\ x\ e^{-\frac{x^2}{2}} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{3.2}$$

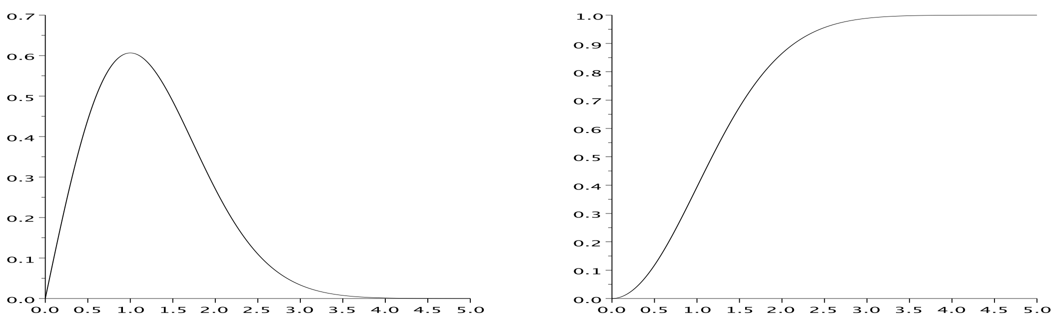


FIGURE 3.23 – Densité $f(x) = x\ e^{-\frac{x^2}{2}}\mathbb{1}_{[0,+\infty[}(x)$ et fonction de répartition F associée.

1. Il est clair d'une part que f est bien positive, et d'autre part que son intégrale sur $[0, +\infty[$ est égale à 1, puisque
$$\int_0^{+\infty} xe^{-\frac{x^2}{2}}dx = \left[-e^{-\frac{x^2}{2}}\right]_0^{+\infty} = 1.$$
Cette densité est représentée figure 3.23 à gauche.
2. La fonction de répartition F de X vaut 0 sur $]-\infty, 0]$ et pour $x \geq 0$, on a
$$F(x) = \int_0^x te^{-\frac{t^2}{2}}dt = \left[-e^{-\frac{t^2}{2}}\right]_0^x = 1 - e^{-\frac{x^2}{2}}.$$
Cette fonction de répartition est représentée figure 3.23 à droite.
3. La médiane de X vérifie $\mathbb{P}(X > m) = 1/2$, ce qui équivaut à dire que $F(m) = 1/2$, donc
$$1 - e^{-\frac{m^2}{2}} = \frac{1}{2} \Leftrightarrow m = \sqrt{2\ln 2}.$$
4. On reconnaît la densité d'une loi normale centrée réduite donc
$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = 1.$$
Par parité de la fonction $x \mapsto e^{-\frac{x^2}{2}}$ sur $\mathbb{R}$, il est clair que
$$\int_0^{+\infty} e^{-\frac{x^2}{2}}dx = \frac{1}{2}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx = \frac{\sqrt{2\pi}}{2}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}}dx\right) = \sqrt{\frac{\pi}{2}}.$$

5. On a donc via une intégration par parties

$$\mathbb{E}[X] = \int_0^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = \int_0^{+\infty} x \times \left(xe^{-\frac{x^2}{2}}\right) dx = \left[-xe^{-\frac{x^2}{2}}\right]_0^{+\infty} + \int_0^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{\frac{\pi}{2}}.$$

6. Soit U une variable aléatoire distribuée suivant une loi uniforme sur $]0,1]$.
   (a) La fonction de répartition F_U vaut 0 à gauche de 0, 1 à droite de 1, et $F_U(u) = u$ pour $0 \leq u \leq 1$.
   (b) Puisque U prend ses valeurs entre 0 et 1, la variable aléatoire $X = \sqrt{-2\ln U}$ prend ses valeurs dans l'intervalle $[0, +\infty[$.
   (c) De fait sa fonction de répartition F_X vaut 0 sur $\mathbb{R}^-$, et pour $x \geq 0$, il suffit d'écrire que

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\sqrt{-2\ln U} \leq x) = \mathbb{P}(0 \leq -2\ln U \leq x^2) = \mathbb{P}(U \geq e^{-\frac{x^2}{2}})$$

   c'est-à-dire

$$F_X(x) = 1 - F_U(e^{-\frac{x^2}{2}}) = 1 - e^{-\frac{x^2}{2}}.$$

   Ainsi X suit bien une loi de Rayleigh de paramètre 1 et la messe est dite.

Exercice 3.32 (Loi de Rademacher et marche aléatoire)
Soit X une variable suivant une loi de Bernoulli de paramètre $1/2$.

1. Moyenne et variance valent respectivement $\mathbb{E}[X] = 1/2$ et $\mathrm{Var}(X) = 1/4$.
2. La variable aléatoire $Y = 2X - 1$ peut prendre les valeurs -1 et 1, et ce de façon équiprobable.
3. Sa moyenne et sa variance découlent directement de celles de X : $\mathbb{E}[Y] = 2\mathbb{E}[X] - 1 = 0$ et $\mathrm{Var}(Y) = 2^2\mathrm{Var}(X) = 1$.
4. (a) Une seconde de réflexion permet de voir que la variable S_{100} peut prendre toutes les valeurs paires de -100 à 100, c'est-à-dire que $S_{100} \in \{-100, -98, \dots, 98, 100\}$. Par linéarité de l'espérance, on a tout d'abord $\mathbb{E}[S_{100}] = 100\mathbb{E}[Y_1] = 0$, et par indépendance des variables intervenant dans la somme, on a aussi $\mathrm{Var}[S_{100}] = 100\mathrm{Var}[Y_1] = 100$.
   (b) Nous sommes exactement dans le cadre de la question précédente : l'ivrogne part du point d'abscisse 0, puis $Y_1 = 1$ s'il fait son premier pas à droite (respectivement $Y_1 = -1$ s'il fait son premier pas à gauche), et ainsi de suite jusqu'au centième pas. La variable S_{100} correspond donc à l'abscisse de l'aviné au bout de 100 pas. Puisque les pas sont indépendants et identiquement distribués, on peut appliquer le théorème central limite pour approcher la loi de cette variable : $S_{100} \approx \mathcal{N}(0, 10^2)$. De fait, il y a 95% de chances de trouver le poivrot à une distance inférieure à deux fois l'écart-type de sa moyenne, c'est-à-dire à une distance inférieure à 20 mètres de son point de départ.

Exercice 3.33 (Précipitation vs. précision)
On note Φ la fonction de répartition de la loi normale centrée réduite.

1. La quantité annuelle de précipitations (en cm) dans une certaine région est une variable X distribuée selon une loi normale de moyenne 140 et de variance 16.
   (a) La probabilité qu'en une année il pleuve plus de 150 cm est d'environ 0.6% puisque

$$\mathbb{P}(X \geq 150) = 1 - \Phi\left(\frac{5}{2}\right) \approx 0.0062$$

   (b) La probabilité qu'il faille attendre au moins 10 ans est la probabilité qu'un événement de probabilité 0.9938 se répète 10 années de suite, c'est-à-dire

$$p = 0.9938^{10} \approx 0.9397 \approx 94\%$$

2. La largeur (en cm) d'une fente entaillée dans une pièce est une variable X qui suit une loi normale de moyenne $m = 2$ et d'écart-type σ. Les limites de tolérance sont données comme étant 2 ± 0.012.
   (a) Si $\sigma = 0.007$, le pourcentage de pièces défectueuses est
   $$1 - \mathbb{P}(2 - 0.012 \leq X \leq 2 + 0.012) = 2 - 2\Phi(1.71) \approx 0.0892$$
   soit environ 9% de pièces défectueuses.
   (b) Avec le même raisonnement, la valeur maximale que peut prendre σ de sorte que le pourcentage de pièces défectueuses ne dépasse pas 1% est telle que
   $$2 - 2\Phi\left(\frac{0.012}{\sigma}\right) \leq 0.01 \;\Leftrightarrow\; \Phi\left(\frac{0.012}{\sigma}\right) \geq 0.995 \;\Leftrightarrow\; \frac{0.012}{\sigma} \geq 2.58$$
   c'est-à-dire $\sigma \leq 0.0047 \approx 0.005$.

Exercice 3.34 (Loi de Weibull)
On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} 3x^2\, e^{-x^3} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \tag{3.2}$$

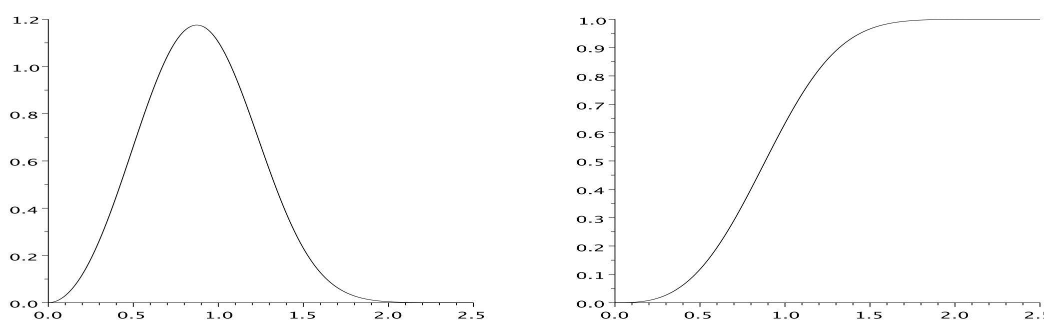


Figure 3.24 – Densité et fonction de répartition de la loi de Weibull.

1. f est positive et elle intègre à 1 puisque
   $$\int_{-\infty}^{+\infty} f(x)dx = \int_0^{+\infty} 3x^2\, e^{-x^3} dx = \left[-e^{-x^3}\right]_0^{+\infty} = 1.$$
2. La dérivée de f est bien sûr nulle à gauche de 0, et pour tout $x \geq 0$:
   $$f'(x) = 3x(2 - 3x^3)e^{-x^3}$$
   Le mode de f se situe donc au point $x_0 = (2/3)^{1/3} \approx 0.87$.
3. La représentation de f est fournie figure 3.24 à gauche.
4. La fonction de répartition F de X est nulle sur $]-\infty, 0]$ et pour tout $x \geq 0$ on trouve
   $$F(x) = \int_0^x f(t)dt = \left[-e^{-t^3}\right]_0^x = 1 - e^{-x^3}.$$
   Sa représentation de f est fournie figure 3.24 à droite.

5. Supposons que la durée de vie (en années) d'un élément soit distribuée selon la loi de Weibull ci-dessus.
   (a) La probabilité que cet élément dure plus de 2 ans est
   $$\mathbb{P}(X \geq 2) = 1 - \mathbb{P}(X \leq 2) = 1 - F(2) = e^{-8} \approx 0.03\%$$
   (b) La probabilité que sa durée de vie soit comprise entre un an et deux ans est
   $$\mathbb{P}(1 \leq X \leq 2) = F(2) - F(1) = e^{-1} - e^{-8} \approx 36.8\%$$
   (c) La probabilité que sa durée de vie soit supérieure à deux ans sachant qu'il fonctionne encore au bout d'un an vaut
   $$\mathbb{P}(X \geq 2 | X \geq 1) = \frac{\mathbb{P}(X \geq 2)}{\mathbb{P}(X \geq 1)} = \frac{1 - F(2)}{1 - F(1)} = e^{-7} \approx 0.09\%$$

Exercice 3.35 (Loi du khi-deux)
Soit X une variable distribuée selon une loi normale centrée réduite $\mathcal{N}(0,1)$.

1. $\mathbb{E}[X] = 0$ et $\text{Var}(X) = 1$, donc $\mathbb{E}[X^2] = \text{Var}(X) + (\mathbb{E}[X])^2 = 1$.
2. X a pour densité
   $$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$
   Le théorème de transfert assure alors que
   $$\mathbb{E}[X^4] = \int_{-\infty}^{+\infty} x^4 f(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^3 \times x e^{-\frac{x^2}{2}} dx$$
   d'où en intégrant par parties
   $$\mathbb{E}[X^4] = \frac{1}{\sqrt{2\pi}} \left[-x^3 e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + 3 \int_{-\infty}^{+\infty} x^2 \times \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = 3\mathbb{E}[X^2] = 3$$
3. Si $Y = X^2$, alors
   $$\text{Var}(Y) = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = \mathbb{E}[X^4] - (\mathbb{E}[X^2])^2 = 3 - 1 = 2.$$
4. Soit $n \in \mathbb{N}^*$ un entier naturel non nul, $X_1, \dots, X_n$ des variables iid de loi normale centrée réduite, et $S_n = X_1^2 + \cdots + X_n^2$.
   (a) Par linéarité de l'espérance et équidistribution des variables X_i, il est clair que $\mathbb{E}[S_n] = n\mathbb{E}[X_1^2] = n$. Par indépendance des X_i^2, on a de plus $\text{Var}(S_n) = n\text{Var}(X_1^2) = 2n$ d'après la question précédente.
   (b) Dans cette situation, le théorème central limite nous dit que S est approximativement distribuée comme une gaussienne de moyenne $\mathbb{E}[S] = 200$ et de variance $\text{Var}(S) = 400$, soit $S \approx \mathcal{N}(200, 20^2)$. Avec environ 95% de chances, S sera donc à distance inférieure à deux fois l'écart-type de sa moyenne, i.e. entre 160 et 240.

Exercice 3.36 (Loi de Laplace)
Cet exercice est corrigé en annexe (sujet de décembre 2009).

Exercice 3.37 (Autour de la loi normale)
Cet exercice est corrigé en annexe (sujet de décembre 2009).

Exercice 3.38 (Variable à densité)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 3.39 (Diamètre d'une bille)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 3.40 (Tchernobyl for ever)
Cet exercice est corrigé en annexe (sujet de décembre 2010).

Exercice 3.41 (Durée de vie d'un processeur)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Exercice 3.42 (Densité quadratique)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Exercice 3.43 (Accidents et fréquence cardiaque)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Exercice 3.44 (Loi de Gumbel)
Cet exercice est corrigé en annexe (sujet de décembre 2011).

Annexe A

Annexes

A.1 Annales

Université de Rennes 2
Licence MASS 2
Arnaud Guyader

Mercredi 21 Octobre 2009
Durée : 1 heure

Contrôle de Probabilités

I. QCM
Chaque réponse correcte rapporte 0.5 point, chaque réponse incorrecte enlève 0.25 point.

1. Soit Ω un ensemble muni d'une tribu $\mathcal{F}$. Est-ce que $\emptyset \in \mathcal{F}$?
   □ Oui.
   □ Non.
2. Soit Ω un ensemble. Est-ce que $\mathcal{P}(\Omega)$, ensemble des parties de Ω, est une tribu ?
   □ Oui.
   □ Non.
3. Soit Ω un ensemble et A un sous-ensemble de Ω. Est-ce que $\{\emptyset, A, \Omega\}$ est une tribu ?
   □ Oui.
   □ Non.
4. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $A_1, \ldots, A_n$ éléments de $\mathcal{F}$. Cochez la (ou les) affirmation(s) vraie(s) parmi les suivantes :
   □ $\mathbb{P}(A_1 \cup \cdots \cup A_n) \leq \sum_{i=1}^n \mathbb{P}(A_i)$.
   □ $\mathbb{P}(A_1 \cup \cdots \cup A_n) = \sum_{i=1}^n \mathbb{P}(A_i)$.
   □ $\mathbb{P}(A_1 \cup \cdots \cup A_n) \geq \min_{1\leq i\leq n} \mathbb{P}(A_i)$.
   □ $\mathbb{P}(A_1 \cup \cdots \cup A_n) \leq \max_{1\leq i\leq n} \mathbb{P}(A_i)$.
5. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n\geq 0}$ une suite d'éléments de $\mathcal{F}$ croissante pour l'inclusion. A-t-on $\mathbb{P}(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} \mathbb{P}(A_n)$?
   □ Oui.
   □ Non.
6. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $(A_n)_{1\leq i\leq n}$ éléments de $\mathcal{F}$. A-t-on $\mathbb{P}(A_n) = \mathbb{P}(A_n|A_{n-1})\mathbb{P}(A_{n-1}|A_{n-2})\ldots\mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$?
   □ Oui.
   □ Non.
7. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $(A_n)_{1\leq i\leq n}$ une partition de $\mathcal{F}$ et A un élément quelconque de $\mathcal{F}$. A-t-on $\mathbb{P}(A) =$

$\sum_{i=1}^{n} \mathbb{P}(A|A_i)$?
- ☐ Oui.
- ☐ Non.

8. Que vaut la somme $S_N = \sum_{n=1}^{N} (\frac{3}{4})^n$?
   - ☐ $S_N = 3(1-(\frac{3}{4})^{N+1})$.
   - ☐ $S_N = 4(1-(\frac{3}{4})^{N+1})$.
   - ☐ $S_N = 3(1-(\frac{3}{4})^{N})$.
   - ☐ $S_N = 4(1-(\frac{3}{4})^{N})$.
9. Que vaut la somme $\sum_{k=0}^{n} \frac{\binom{n}{k}}{2^k}$?
   - ☐ 2^n.
   - ☐ 2^{-n}.
   - ☐ $(3/2)^n$.
   - ☐ $(1/2)^n$.

II. Dénombrements

1. Les initiales de Andréï Kolmogorov sont A.K. Combien y a-t-il d'initiales possibles en tout ? Combien au minimum un village doit-il avoir d'habitants pour qu'on soit sûr que deux personnes au moins aient les mêmes initiales ?
2. Lors d'une course hippique, 12 chevaux prennent le départ. Donner le nombre de tiercés dans l'ordre (un tiercé dans l'ordre est la donnée du premier, du deuxième et du troisième cheval arrivés, dans cet ordre).
3. Dans un jeu de 32 cartes, on a remplacé une carte autre que la dame de cœur par une seconde dame de cœur. Une personne tire au hasard 3 cartes simultanément. Quelle est la probabilité qu'elle s'aperçoive de la supercherie ?

III. Probabilités

1. Deux urnes contiennent chacune initialement 2 boules noires et 3 boules blanches. On tire au hasard une boule de la première urne, on note sa couleur et on la remet dans la seconde urne. On tire alors au hasard une boule de la seconde urne. Quelle est la probabilité d'obtenir deux fois une boule noire ?
2. Une population possède une proportion $p \in]0,1[$ de tricheurs. Lorsqu'on fait tirer une carte d'un jeu de 52 cartes à un tricheur, il est sûr de retourner un as. Exprimer en fonction de p la probabilité qu'un individu choisi au hasard dans la population retourne un as.
3. On prend un dé au hasard parmi un lot de 100 dés dont 25 sont pipés. Pour un dé pipé, la probabilité d'obtenir 6 est 1/2. On lance le dé choisi et on obtient 6.
   (a) Quelle est la probabilité que ce dé soit pipé ?
   (b) On relance alors ce dé et on obtient à nouveau 6. Quelle est la probabilité que ce dé soit pipé ?
   (c) (Bonus) Généralisation : on lance n fois le dé et à chaque fois on obtient 6. Quelle est la probabilité p_n que ce dé soit pipé ? Que vaut $\lim_{n\to\infty} p_n$? Commenter ce résultat.

Université de Rennes 2
Licence MASS 2
Arnaud Guyader

Mercredi 21 Octobre 2009
Durée : 1 heure

Corrigé du Contrôle

I. QCM

1. Soit Ω un ensemble muni d'une tribu $\mathcal{F}$. Est-ce que $\emptyset \in \mathcal{F}$?
   ☒ Oui.
   ☐ Non.
2. Soit Ω un ensemble. Est-ce que $\mathcal{P}(\Omega)$, ensemble des parties de Ω, est une tribu ?
   ☒ Oui.
   ☐ Non.
3. Soit Ω un ensemble et A un sous-ensemble de Ω. Est-ce que $\{\emptyset, A, \Omega\}$ est une tribu ?
   ☐ Oui.
   ☒ Non.
4. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $A_1, \dots, A_n$ éléments de $\mathcal{F}$. Cochez la (ou les) affirmation(s) vraie(s) parmi les suivantes :
   ☒ $\mathbb{P}(A_1 \cup \dots \cup A_n) \leq \sum_{i=1}^n \mathbb{P}(A_i)$.
   ☐ $\mathbb{P}(A_1 \cup \dots \cup A_n) = \sum_{i=1}^n \mathbb{P}(A_i)$.
   ☒ $\mathbb{P}(A_1 \cup \dots \cup A_n) \geq \min_{1\leq i\leq n} \mathbb{P}(A_i)$.
   ☐ $\mathbb{P}(A_1 \cup \dots \cup A_n) \leq \max_{1\leq i\leq n} \mathbb{P}(A_i)$.
5. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé et $(A_n)_{n\geq 0}$ une suite d'éléments de $\mathcal{F}$ croissante pour l'inclusion. A-t-on $\mathbb{P}(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n\to+\infty} \mathbb{P}(A_n)$?
   ☐ Oui.
   ☒ Non.
6. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $(A_n)_{1\leq i\leq n}$ éléments de $\mathcal{F}$. A-t-on $\mathbb{P}(A_n) = \mathbb{P}(A_n|A_{n-1})\mathbb{P}(A_{n-1}|A_{n-2}) \dots \mathbb{P}(A_2|A_1)\mathbb{P}(A_1)$?
   ☐ Oui.
   ☒ Non.
7. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace probabilisé, $(A_n)_{1\leq i\leq n}$ une partition de $\mathcal{F}$ et A un élément quelconque de $\mathcal{F}$. A-t-on $\mathbb{P}(A) = \sum_{i=1}^n \mathbb{P}(A|A_i)$?
   ☐ Oui.
   ☒ Non.
8. Que vaut la somme $S_N = \sum_{n=1}^N (\frac{3}{4})^n$?
   ☐ $S_N = 3(1 - (\frac{3}{4})^{N+1})$.
   ☐ $S_N = 4(1 - (\frac{3}{4})^{N+1})$.
   ☒ $S_N = 3(1 - (\frac{3}{4})^{N})$.
   ☐ $S_N = 4(1 - (\frac{3}{4})^{N})$.
9. Que vaut la somme $\sum_{k=0}^n \frac{\binom{n}{k}}{2^k}$?
   ☐ 2^n.
   ☐ 2^{-n}.
   ☒ $(3/2)^n$.
   ☐ $(1/2)^n$.

II. Dénombrements

1. Puisqu'il y a 26 lettres dans l'alphabet, il y a en tout $26^2 = 676$ initiales possibles (on exclut ici les prénoms composés). Pour que deux personnes au moins aient les mêmes initiales, un village doit donc compter au moins 677 habitants.
2. Il y a 12 possibilités pour le premier cheval arrivé, 11 pour le deuxième et 10 pour le troisième, donc $N = 12 \times 11 \times 10 = A_{12}^3 = 1320$ tiercés possibles.

3. Pour que la personne se rende compte de la supercherie, il faut que parmi les 3 cartes tirées il y ait les 2 dames de cœur. Le nombre de tirages possibles de 3 cartes est le nombre de combinaisons de 3 éléments parmi 32, c'est-à-dire $\binom{32}{3} = 4960$. Le nombre de tirages contenant les 2 dames de cœur est 30 puisque seule la dernière carte est au choix. La probabilité qu'on s'aperçoive du problème est donc $p = \frac{30}{4960} \approx 0,006$.

III. Probabilités

1. Notons N_1 (resp. N_2) l'événement : "Tirage d'une boule noire au premier (resp. second) tirage". La probabilité cherchée est donc :
$$p = \mathbb{P}(N_1 \cap N_2) = \mathbb{P}(N_2|N_1)\mathbb{P}(N_1).$$
Dans l'urne initiale il y a 2 noires et 3 blanches donc $\mathbb{P}(N_1) = 2/5$. Sachant qu'on a pioché une noire au premier tirage, on connaît la composition de la seconde urne avant le second tirage : il y a 3 noires et 3 blanches, donc $\mathbb{P}(N_2|N_1) = 1/2$. Ainsi $p = 1/5$.
2. Notons A l'événement : "L'individu tire un as" et T l'événement : "L'individu est un tricheur". On cherche donc $\mathbb{P}(A)$ que l'on décompose sur la partition $(T, \overline{T})$ via la formule des probabilités totales :
$$\mathbb{P}(A) = \mathbb{P}(A|T)\mathbb{P}(T) + \mathbb{P}(A|\overline{T})\mathbb{P}(\overline{T}).$$
On connaît $\mathbb{P}(T) = p$ et $\mathbb{P}(A|T) = 1$. Il reste à voir que $\mathbb{P}(A|\overline{T}) = 1/13$ puisqu'il y a 4 as pour 52 cartes. On arrive finalement à :
$$\mathbb{P}(A) = \frac{1 + 12p}{13}.$$
3. On note simplement 6 l'événement : "Le lancer donne un 6", et $\mathcal{P}$ l'événement : "Le dé est pipé".
   (a) On cherche ici $\mathbb{P}(\mathcal{P}|6)$ et on utilise la formule de Bayes :
   $$\mathbb{P}(\mathcal{P}|6) = \frac{\mathbb{P}(6|\mathcal{P})\mathbb{P}(\mathcal{P})}{\mathbb{P}(6|\mathcal{P})\mathbb{P}(\mathcal{P}) + \mathbb{P}(6|\overline{\mathcal{P}})\mathbb{P}(\overline{\mathcal{P}})}.$$
   Le texte nous dit que $\mathbb{P}(\mathcal{P}) = 1/4$ et $\mathbb{P}(6|\mathcal{P}) = 1/2$, par ailleurs on sait que $\mathbb{P}(6|\overline{\mathcal{P}}) = 1/6$, ce qui donne au total :
   $$\mathbb{P}(\mathcal{P}|6) = \frac{\frac{1}{2} \times \frac{1}{4}}{\frac{1}{2} \times \frac{1}{4} + \frac{1}{6} \times \frac{3}{4}} = \frac{1}{2}.$$
   (b) On note 66 l'événement : "Les 2 lancers donnent 6", donc on cherche cette fois $\mathbb{P}(\mathcal{P}|66)$ et on utilise à nouveau la formule de Bayes :
   $$\mathbb{P}(\mathcal{P}|66) = \frac{\mathbb{P}(66|\mathcal{P})\mathbb{P}(\mathcal{P})}{\mathbb{P}(66|\mathcal{P})\mathbb{P}(\mathcal{P}) + \mathbb{P}(66|\overline{\mathcal{P}})\mathbb{P}(\overline{\mathcal{P}})}.$$
   Il suffit alors de voir que par indépendance des lancers on a $\mathbb{P}(66|\mathcal{P}) = \mathbb{P}(6|\mathcal{P})\mathbb{P}(6|\mathcal{P}) = 1/4$ et $\mathbb{P}(66|\overline{\mathcal{P}}) = \mathbb{P}(6|\overline{\mathcal{P}})\mathbb{P}(6|\overline{\mathcal{P}}) = 1/36$. Ceci donne $\mathbb{P}(\mathcal{P}|66) = 3/4$.
   (c) On note n l'événement : "Les n lancers donnent 6", et on raisonne comme ci-dessus :
   $$p_n = \mathbb{P}(\mathcal{P}|n) = \frac{\mathbb{P}(n|\mathcal{P})\mathbb{P}(\mathcal{P})}{\mathbb{P}(n|\mathcal{P})\mathbb{P}(\mathcal{P}) + \mathbb{P}(n|\overline{\mathcal{P}})\mathbb{P}(\overline{\mathcal{P}})} = \frac{\left(\frac{1}{2}\right)^n \times \frac{1}{4}}{\left(\frac{1}{2}\right)^n \times \frac{1}{4} + \left(\frac{1}{6}\right)^n \times \frac{3}{4}} = \frac{1}{1 + \left(\frac{1}{3}\right)^{n-1}}.$$
   On a donc $\lim_{n\to\infty} p_n = 1$. Que le dé soit pipé ou non, la probabilité d'obtenir n fois de suite un 6 tend vers zéro. Néanmoins, on a bien plus de chances que ce phénomène arrive avec un dé pipé qu'avec un dé non pipé : c'est ce que dit $\lim_{n\to\infty} p_n = 1$.

Université de Rennes 2
Licence MASS 2
Durée : 2 heures

Lundi 14 Décembre 2009
Calculatrice autorisée
Aucun document

Contrôle de Probabilités

I. Boules blanches et noires
Un sac contient 8 boules blanches et 2 boules noires. On tire les boules les unes après les autres, sans remise, jusqu'à obtenir une boule blanche. On appelle X le nombre de tirages nécessaires pour obtenir cette boule blanche.

1. Quelles valeurs peut prendre la variable aléatoire X ?
2. Donner la loi de X.
3. Représenter sa fonction de répartition F.
4. Calculer $\mathbb{E}[X]$ et $\mathrm{Var}(X)$.

II. Défaut de fabrication
On admet que la probabilité de défaut pour un objet fabriqué à la machine est égale à 0,1. On considère un lot de 10 objets fabriqués par cette machine. Soit X le nombre d'objets défectueux parmi ceux-ci.

1. Comment s'appelle la loi suivie par X ?
2. Que valent $\mathbb{E}[X]$ et $\mathrm{Var}(X)$?
3. Quelle est la probabilité que le lot comprenne au plus 1 objet défectueux ?
4. Retrouver ce résultat grâce à l'approximation par une loi de Poisson.

III. Recrutement
Une entreprise veut recruter un cadre. Il y a en tout 10 candidats à se présenter pour ce poste. L'entreprise fait passer un test au premier candidat, qui est recruté s'il le réussit. Sinon, elle fait passer le même test au second candidat et ainsi de suite. On suppose que la probabilité qu'un candidat réussisse le test est égale à p, réel fixé compris entre 0 et 1. On appelle alors X la variable aléatoire à valeurs dans $\{1, \dots, 11\}$ qui vaut k si c'est le candidat numéro k qui est recruté, et 11 si aucun candidat n'est recruté.

1. Calculer en fonction de p les probabilités $\mathbb{P}(X = 1), \mathbb{P}(X = 2), \dots, P(X = 10)$. Déterminer aussi $\mathbb{P}(X = 11)$.
2. Comment doit-on choisir p pour que la probabilité de ne recruter personne soit inférieure à 1% ?

3. Pour $n \in \mathbb{N}$ fixé, on considère la fonction P définie par :

$$P(x) = 1 + x + \cdots + x^n = \sum_{j=0}^{n} x^j.$$

Exprimer sa dérivée $P'(x)$ sous la forme d'une somme de n termes.

4. Pour $x \neq 1$, écrire plus simplement $P(x)$ (penser à la somme des termes d'une suite géométrique). En déduire une autre expression de $P'(x)$, à savoir :

$$P'(x) = \frac{nx^{n+1} - (n+1)x^n + 1}{(1-x)^2}.$$

5. Déduire des questions précédentes que X a pour moyenne :

$$\mathbb{E}[X] = \frac{1 - (1-p)^{11}}{p}.$$

6. Supposons maintenant qu'il n'y ait pas seulement 10 candidats, mais un nombre infini, et que l'on procède de la même façon. Appelons Y le numéro du candidat retenu. Quelle est la loi classique suivie par Y ? Rappeler son espérance. La comparer à $\mathbb{E}[X]$ lorque $p = 1/2$.

IV. Autour de la loi normale

On considère une variable aléatoire X de loi normale $\mathcal{N}(0,1)$.

1. Montrer que, pour tout $n \in \mathbb{N}$, on a : $\mathbb{E}[X^{n+2}] = (n+1)\mathbb{E}[X^n]$ (intégrer par parties).
2. Que vaut $\mathbb{E}[X^2]$? Déduire de ce résultat et de la question précédente la valeur de $\mathbb{E}[X^4]$.
3. Que vaut $\mathbb{E}[X^3]$?
4. Soit Y la variable aléatoire définie par $Y = 2X + 1$.
   (a) Quelle est la loi de Y ?
   (b) Déterminer $\mathbb{E}[Y^4]$ (on pourra utiliser la formule du binôme et les moments de X trouvés précédemment).
5. A l'aide de la table de la loi normale, déterminer $\mathbb{P}(|X| \geq 2)$. Que donne l'inégalité de Tchebychev dans ce cas ? Comparer et commenter.
6. On considère maintenant que X suit une loi normale de moyenne 7 et d'écart-type 4.
   (a) Déterminer $\mathbb{P}(X \leq 8)$ et $\mathbb{P}(5 \leq X \leq 9)$.
   (b) Déterminer q tel que $\mathbb{P}(X > q) = 0,9$.
7. (Bonus) La taille des enfants d'un collège est distribuée selon une loi normale de moyenne m et d'écart-type σ. On sait qu'un cinquième des élèves mesurent moins de 1m50 et que 10% des élèves mesurent plus de 1m80. Déterminer m et σ.

V. Loi de Laplace

On considère une variable aléatoire X dont la densité f est donnée par :

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{1}{2}e^{-|x|},$$

où $|x|$ représente la valeur absolue de x, c'est-à-dire $|x| = x$ si $x \geq 0$ et $|x| = -x$ si $x \leq 0$.

1. Vérifier que f est bien une densité sur $\mathbb{R}$. Représenter f.
2. On note F la fonction de répartition de X. Calculer $F(x)$ (on distinguera les cas $x \leq 0$ et $x \geq 0$). Représenter F.

3. Montrer que $\mathbb{E}[X] = 0$.
4. Pour tout $n \in \mathbb{N}$, on appelle I_n l'intégrale définie par :

$$I_n = \int_0^{+\infty} x^n e^{-x} dx.$$

   (a) Combien vaut I_0 ?
   (b) Montrer que pour tout $n \in \mathbb{N}^*$, $I_n = nI_{n-1}$. En déduire que $I_n = n!$ pour tout $n \in \mathbb{N}$.
5. Pour tout $n \in \mathbb{N}$, calculer $\mathbb{E}[X^{2n}]$. Que vaut $\mathrm{Var}(X)$?
6. Pour tout $n \in \mathbb{N}$, que vaut $\mathbb{E}[X^{2n+1}]$?

Université de Rennes 2
Licence MASS 2
Durée : 2 heures

Lundi 14 Décembre 2009
Calculatrice autorisée
Aucun document

Corrigé du Contrôle

I. Boules blanches et noires
Un sac contient 8 boules blanches et 2 boules noires. On tire les boules les unes après les autres, sans remise, jusqu'à obtenir une boule blanche. On appelle X le nombre de tirages nécessaires pour obtenir cette boule blanche.

1. Puisqu'il n'y a que deux boules noires, la variable aléatoire X ne peut prendre que les valeurs 1, 2 ou 3.
2. Notons B_i (resp. N_i) le fait de tirer une boule blanche (resp. noire) au i-ème tirage. On peut alors écrire :
$$\mathbb{P}(X=1)=\mathbb{P}(B_1)=\frac{8}{10}=\frac{4}{5}.$$
Pour le calcul suivant, on procède par conditionnement :
$$\mathbb{P}(X=2)=\mathbb{P}(N_1\cap B_2)=\mathbb{P}(N_1)\mathbb{P}(B_2|N_1)=\frac{2}{10}\times\frac{8}{9}=\frac{8}{45}.$$
Enfin, de la même façon :
$$\mathbb{P}(X=3)=\mathbb{P}(N_1\cap N_2\cap B_3)=\mathbb{P}(N_1)\mathbb{P}(N_2|N_1)\mathbb{P}(B_3|N_1\cap N_2)=\frac{2}{10}\times\frac{1}{9}\times\frac{8}{8}=\frac{1}{45},$$
et la loi de X est ainsi complètement déterminée. Remarque : au passage, on vérifie bien que $4/5+8/45+1/45=1$.

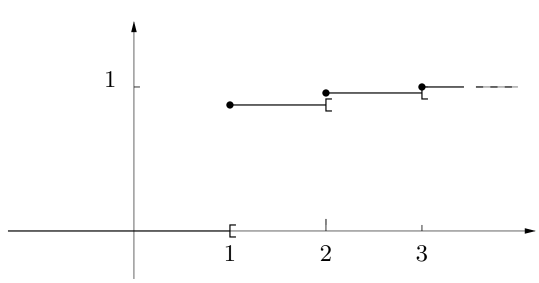


FIGURE A.1 – Fonction de répartition de la variable X.

3. Sa fonction de répartition F s'en déduit facilement et est représentée figure A.1.

4. Par définition de l'espérance d'une variable discrète, la moyenne de X vaut donc :

$$\mathbb{E}[X] = 1 \times \frac{4}{5} + 2 \times \frac{8}{45} + 3 \times \frac{1}{45} = \frac{11}{9} \approx 1,22.$$

Quant à sa variance, on commence par calculer $\mathbb{E}[X^2]$:

$$\mathbb{E}[X^2] = 1^2 \times \frac{4}{5} + 2^2 \times \frac{8}{45} + 3^2 \times \frac{1}{45} = \frac{77}{45} \approx 1,71.$$

D'où l'on déduit : $\text{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{88}{405} \approx 0,22$.

II. Défaut de fabrication

On admet que la probabilité de défaut pour un objet fabriqué à la machine est égale à 0,1. On considère un lot de 10 objets fabriqués par cette machine. Soit X le nombre d'objets défectueux parmi ceux-ci.

1. La variable aléatoire X suit une loi binomiale $\mathcal{B}(10; 0,1)$.
2. Il s'ensuit que $\mathbb{E}[X] = 10 \times 0,1 = 1$ et $\text{Var}(X) = 10 \times 0,1 \times 0,9 = 0,9$.
3. La probabilité p que le lot comprenne au plus 1 objet défectueux s'écrit :

$$p = \mathbb{P}(X=0) + \mathbb{P}(X=1) = \binom{10}{0}(0,1)^0(0,9)^{10} + \binom{10}{1}(0,1)^1(0,9)^9 \approx 0,7361.$$

4. Puisque le paramètre 0,1 de la binomiale est petit, l'approximation poissonienne consiste à dire que X suit à peu près une loi de Poisson de paramètre le produit des deux paramètres de la binomiale, soit $X \approx \mathcal{P}(1)$. Ainsi la probabilité p que le lot comprenne au plus 1 objet défectueux vaut approximativement :

$$p = \mathbb{P}(X=0) + \mathbb{P}(X=1) \approx e^{-1}\frac{1^0}{0!} + e^{-1}\frac{1^1}{1!} = \frac{2}{e} \approx 0,7358.$$

L'approximation par une loi de Poisson est donc très bonne dans ce cas.

III. Recrutement

1. Pour alléger les notations, nous noterons $q = (1-p)$. Jusque $k = 10$, tout se passe comme pour une loi géométrique de paramètre p. En effet, pour que $X = k$, il faut que les $(k-1)$ premiers candidats aient échoué au test, ce qui arrive avec probabilité q^{k-1}, et que le k-ème candidat ait réussi, ce qui arrive avec probabilité p. Puisque les résultats des candidats au test sont indépendants, il vient :

$$\forall k \in \{1, \dots, 10\} \qquad \mathbb{P}(X = k) = pq^{k-1}.$$

Enfin, dans l'éventualité où aucun candidat ne réussit le test, on obtient $\mathbb{P}(X = 11) = q^{10}$.

2. Pour que la probabilité de ne recruter personne soit inférieure à 1%, il faut et il suffit que :

$$\mathbb{P}(X = 11) \leq 0,01 \iff (1-p)^{10} \leq 0,01 \iff p \geq 1 - 0,01^{1/10} \approx 0,369.$$

3. La dérivée de P s'écrit :

$$P'(x) = 1 + 2x + \cdots + nx^{n-1} = \sum_{j=1}^{n} jx^{j-1}.$$

4. Pour $x \neq 1$, la somme des termes d'une suite géométrique de raison x donne :
$$P(x) = \frac{1 - x^{n+1}}{1 - x}.$$
Par dérivation de cette formule, on obtient :
$$P'(x) = \frac{nx^{n+1} - (n+1)x^n + 1}{(1-x)^2}.$$
5. La moyenne de X s'écrit :
$$\mathbb{E}[X] = 1 \times p + 2 \times pq + \cdots + 10 \times pq^9 + 11 \times q^{10} = p\left(1 + 2q + \cdots + 10q^9\right) + 11q^{10},$$
où l'on reconnaît la fonction P' lorsque $n = 10$:
$$\mathbb{E}[X] = p \times P'(q) + 11q^{10}.$$
Il reste à appliquer la formule obtenue précédemment pour $P'(x)$, à tout mettre au même dénominateur et à simplifier pour obtenir :
$$\mathbb{E}[X] = \frac{1 - (1-p)^{11}}{p}.$$
6. S'il y a une "infinité" de candidats, alors Y suit une loi géométrique $\mathcal{G}(p)$. Son espérance vaut dans ce cas : $\mathbb{E}[Y] = 1/p$. On voit qu'elle est supérieure à $\mathbb{E}[X]$, mais de très peu, puisque pour $p = 1/2$ on a $\mathbb{E}[Y] = 2$ et :
$$\mathbb{E}[X] = \frac{1 - \left(1 - \frac{1}{2}\right)^{11}}{\frac{1}{2}} \approx 1,999.$$

IV. Autour de la loi normale

On considère une variable aléatoire X de loi normale $\mathcal{N}(0,1)$.

1. Soit $n \in \mathbb{N}$ fixé, alors par le théorème de transfert le moment d'ordre $(n+2)$ de X s'écrit :
$$\mathbb{E}[X^{n+2}] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{n+2} e^{-\frac{x^2}{2}}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^{n+1} \left(xe^{-\frac{x^2}{2}}\right) dx,$$
que l'on peut intégrer par parties :
$$\mathbb{E}[X^{n+2}] = \frac{1}{\sqrt{2\pi}} \left[-x^{n+1} e^{-\frac{x^2}{2}}\right]_{-\infty}^{+\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (n+1)x^n e^{-\frac{x^2}{2}}\, dx,$$
et puisque la quantité entre crochets est nulle, ceci donne bien : $\mathbb{E}[X^{n+2}] = (n+1)\mathbb{E}[X^n]$.
2. Il en découle par exemple que :
$$\mathbb{E}[X^2] = 1 \times \mathbb{E}[X^0] = \mathbb{E}[1] = 1.$$
Il vient alors :
$$\mathbb{E}[X^4] = 3\mathbb{E}[X^2] = 3.$$
3. On obtient de même :
$$\mathbb{E}[X^3] = 2\mathbb{E}[X] = 0,$$
puisque par hypothèse la variable X est centrée. De façon générale, il est clair que tous les moments d'ordres impairs d'une loi normale centrée sont nuls.

4. Soit Y la variable aléatoire définie par $Y = 2X + 1$.
   (a) Par stabilité de la loi normale par transformation affine, Y suit elle aussi une loi normale et plus précisément : $Y \sim \mathcal{N}(1, 4)$.
   (b) Pour déterminer le moment d'ordre 4 de Y, on utilise la formule du binôme :
   $$Y^4 = (2X + 1)^4 = 1 + 8X + 24X^2 + 32X^3 + 16X^4,$$
   d'où par linéarité de l'espérance :
   $$\mathbb{E}[Y^4] = 1 + 8\mathbb{E}[X] + 24\mathbb{E}[X^2] + 32\mathbb{E}[X^3] + 16\mathbb{E}[X^4] = 73.$$

5. A l'aide de la table et en notant comme d'habitude Φ la fonction de répartition de la loi normale centrée réduite, on a :
   $$\mathbb{P}(|X| \geq 2) = \mathbb{P}(X \leq -2) + \mathbb{P}(X \geq 2) = \Phi(-2) + (1 - \Phi(2)) = 2(1 - \Phi(2)) \approx 0,0456.$$
   L'inégalité de Tchebychev donne dans ce cas :
   $$\mathbb{P}(|X| \geq 2) = \mathbb{P}(|X - \mathbb{E}[X]| \geq 2) \leq \frac{\mathrm{Var}(X)}{2^2} = \frac{1}{4}.$$
   On voit donc que cette majoration d'une quantité d'environ 5% par 25% est très grossière. Mais n'oublions pas que l'inégalité de Tchebychev est universelle, en ce sens qu'elle s'applique à toute variable aléatoire admettant une variance. Elle ne peut donc être précise dans toutes les situations.

6. On considère maintenant que X suit une loi normale de moyenne 7 et d'écart-type 4.
   (a) Pour pouvoir utiliser la table, l'idée est de centrer et réduire X :
   $$\mathbb{P}(X \leq 8) = \mathbb{P}\left(\frac{X - 7}{4} \leq \frac{8 - 7}{4}\right) = \Phi(0,25) \approx 0,5987.$$
   Sur le même principe :
   $$\mathbb{P}(5 \leq X \leq 9) = \mathbb{P}\left(-\frac{1}{2} \leq \frac{X - 7}{4} \leq \frac{1}{2}\right) = \Phi(0,5) - \Phi(-0,5) = 2\Phi(0,5) - 1 \approx 0,383.$$
   (b) Pour déterminer q, il suffit d'écrire :
   $$\mathbb{P}(X > q) = 0,9 \;\Leftrightarrow\; \mathbb{P}\left(\frac{X - 7}{4} > \frac{q - 7}{4}\right) = 0,9$$
   qui vaut encore :
   $$\mathbb{P}\left(\frac{X - 7}{4} \leq \frac{7 - q}{4}\right) = 0,9 \;\Leftrightarrow\; \Phi\left(\frac{7 - q}{4}\right) = 0,9.$$
   La lecture de la table permet d'en déduire que :
   $$\frac{7 - q}{4} \approx 1,29 \;\Leftrightarrow\; q \approx 1,88.$$

7. D'après l'énoncé, il est clair que la moyenne de taille m des élèves se situe entre 1m50 et 1m80. Traduisons plus précisément les informations de l'énoncé :

$$\mathbb{P}(X \leq 1,5) = 0,2 \;\Leftrightarrow\; \mathbb{P}\left(\frac{X-m}{\sigma} \leq \frac{1,5-m}{\sigma}\right) = 0,2 \;\Leftrightarrow\; \Phi\left(\frac{1,5-m}{\sigma}\right) = 0,2$$

c'est-à-dire :

$$\Phi\left(\frac{m-1,5}{\sigma}\right) = 0,8 \;\Leftrightarrow\; \frac{m-1,5}{\sigma} = 0,84 \;\Leftrightarrow\; m - 0,84\sigma = 1,5.$$

Le même raisonnement fournit l'équation :

$$\mathbb{P}(X \geq 1,8) = 0,1 \;\Leftrightarrow\; \frac{1,8-m}{\sigma} = 1,28 \;\Leftrightarrow\; m + 1,28\sigma = 1,8.$$

Il suffit alors de résoudre le système linéaire de deux équations à deux inconnues :

$$\left\{\begin{array}{l} m - 0,84\sigma = 1,5 \\ m + 1,28\sigma = 1,8 \end{array}\right. \Longleftrightarrow \left\{\begin{array}{l} m \approx 1,62 \\ \sigma \approx 0,14 \end{array}\right.$$

Ainsi la taille des élèves de ce collège est distribuée selon une loi normale de moyenne 1m62 et d'écart-type 14cm.

V. Loi de Laplace

On considère une variable aléatoire X dont la densité f est donnée par :

$$\forall x \in \mathbb{R}, \qquad f(x) = \frac{1}{2}e^{-|x|}.$$

On dit dans ce cas que X suit une loi de Laplace de paramètre 1.

1. La fonction f est positive donc il suffit de vérifier qu'elle intègre à 1. Puisqu'elle est paire, le calcul de l'intégrale se ramène à l'intervalle $[0, +\infty[$:

$$\int_{-\infty}^{+\infty} f(x)dx = 2\int_0^{+\infty} \frac{1}{2}e^{-|x|}dx = \int_0^{+\infty} e^{-x}dx = \left[-e^{-x}\right]_0^{+\infty} = 1,$$

et qui fait bien de f une densité sur $\mathbb{R}$. Sa représentation est donnée figure A.2, à gauche.

2. Pour $x \leq 0$, la fonction de répartition s'écrit :

$$F(x) = \int_{-\infty}^{x} f(u)du = \frac{1}{2}\int_{-\infty}^{x} e^u du = \frac{1}{2}\left[e^u\right]_{-\infty}^{x} = \frac{e^x}{2}.$$

En particulier, on a $F(0) = 1/2$ (ce qui est logique puisque f est paire). Maintenant, pour $x \geq 0$:

$$F(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} f(u)du + \int_0^x f(u)du = F(0) + \frac{1}{2}\int_0^x e^{-u}du,$$

c'est-à-dire :

$$F(x) = \frac{1}{2}(1 + e^{-x}).$$

La représentation de F est donnée figure A.2, à droite.

3. La densité f est impaire sur le domaine $]-\infty, +\infty[$ symétrique par rapport à 0 et l'intégrale généralisée définissant $\mathbb{E}[X]$ est convergente, donc $\mathbb{E}[X] = 0$.

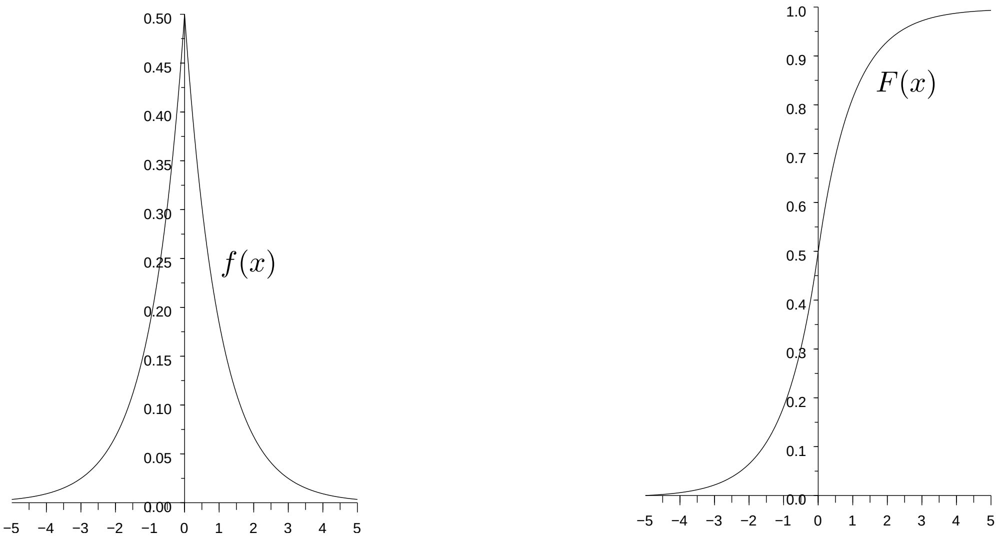


Figure A.2 – Densité (à gauche) et fonction de répartition (à droite) pour une loi de Laplace.

4. Pour tout $n \in \mathbb{N}$, on appelle I_n l'intégrale définie par :

$$I_n = \int_0^{+\infty} x^n e^{-x} dx.$$

   (a) Le calcul de I_0 a déjà été fait pour vérifier que f est une densité. On a obtenu $I_0 = 1$.

   (b) Pour tout $n \in \mathbb{N}^*$, une intégration par parties donne :

$$I_n = \left[-x^n e^{-x}\right]_0^{+\infty} + \int_0^{+\infty} n x^{n-1} e^{-x} dx = n I_{n-1}.$$

   On a donc pour tout $n \in \mathbb{N}$:

$$I_n = n I_{n-1} = n(n-1) I_{n-2} = \cdots = n! I_0 = n!$$

5. Puisque la fonction $x \mapsto \frac{1}{2} x^{2n} e^{-|x|}$ est paire et intégrable sur $\mathbb{R}$, on a :

$$\mathbb{E}[X^{2n}] = \int_{-\infty}^{+\infty} x^{2n} \frac{e^{-|x|}}{2} dx = \int_0^{+\infty} x^{2n} e^{-x} dx = I_{2n} = (2n)!$$

   En particulier : $\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X^2] = 2! = 2$.

6. De l'imparité de la fonction $x \mapsto \frac{1}{2} x^{2n+1} e^{-|x|}$, on déduit que $\forall n \in \mathbb{N} : \mathbb{E}[X^{2n+1}] = 0$.

Université de Rennes 2
Licence MASS 2
Durée : 1 heure 45

Lundi 22 Novembre 2010
Calculatrice autorisée
Aucun document

Contrôle de Probabilités

I. Evénements indépendants
On considère deux événements indépendants A et B de probabilités respectives $1/4$ et $1/3$. Calculer :

1. la probabilité que les deux événements aient lieu.
2. la probabilité que l'un au moins des deux événements ait lieu.
3. la probabilité qu'exactement l'un des deux événements ait lieu.

II. Un tirage en deux temps
Une boîte contient une balle noire et une balle blanche. Une balle est tirée au hasard dans la boîte : on remet celle-ci ainsi qu'une nouvelle balle de la même couleur. On tire alors une des trois balles au hasard dans la boîte.

1. Quelle est la probabilité que la seconde balle tirée soit blanche ?
2. Quelle est la probabilité que l'une au moins des deux balles tirées soit blanche ?
3. Quelle est la probabilité que la première balle tirée soit blanche, sachant que l'une au moins des deux balles tirées est blanche ?

III. Pièces défectueuses
Une usine produit des objets par boîtes de deux. Sur le long terme, on a constaté que : 92% des boîtes ne contiennent aucun objet défectueux ; 5% des boîtes contiennent exactement 1 objet défectueux ; 3% des boîtes contiennent 2 objets défectueux. Une boîte est choisie au hasard sur la chaîne de production et on tire au hasard un des deux objets de cette boîte.

1. Quelle est la probabilité que cet objet soit défectueux ?
2. Sachant que cet objet est effectivement défectueux, quelle est la probabilité que l'autre objet de la boîte le soit aussi ?

IV. Lancer de dé
Un dé équilibré est lancé 10 fois de suite. Déterminer :

1. La probabilité d'au moins un 6 sur les 10 lancers.
2. Le nombre moyen de 6 sur les 10 lancers.
3. La moyenne de la somme des résultats obtenus lors des 10 lancers.

4. (Bonus) La probabilité d'obtenir exactement deux 6 lors des 5 premiers lancers sachant qu'il y en a eu 4 sur les 10 lancers.

V. Le dé dyadique

On appelle "dé dyadique" un dé dont les faces sont numérotées respectivement 2, 4, 8, 16, 32, 64 (au lieu de 1, 2, 3, 4, 5, 6). On jette un dé dyadique équilibré et on appelle X le résultat obtenu.

1. Déterminer l'espérance de X.
2. Calculer l'écart-type de X.
3. Lorsque X_1 et X_2 sont deux variables indépendantes, que vaut $\mathrm{Cov}(X_1, X_2)$?
4. On jette maintenant deux dés dyadiques équilibrés et on appelle Y le produit des résultats obtenus. Calculer l'espérance de Y.
5. (Bonus) Calculer $\mathbb{P}(Y < 20)$.

VI. Répartition des tailles

On suppose que dans une population, 1% des gens mesurent plus de 1m92. Supposons que vous tiriez au hasard (avec remise) 200 personnes dans cette population. Appelons X le nombre de personnes de plus de 1m92 dans votre échantillon.

1. Quelle est la loi de X ?
2. Par quelle loi peut-on l'approcher ?
3. Quelle est la probabilité que dans votre échantillon, au moins 3 personnes mesurent plus de 1m92 ?

VII. Poisson en vrac

On considère une variable X distribuée selon une loi de Poisson de paramètre $\lambda > 0$. Exprimer en fonction de λ :

1. $\mathbb{E}[3X + 5]$.
2. $\mathrm{Var}(2X + 1)$.
3. $\mathbb{E}\left[\frac{1}{X+1}\right]$.

Université de Rennes 2
Licence MASS 2
Durée : 2 heures

Lundi 22 Novembre 2010
Calculatrice autorisée
Aucun document

Corrigé du Contrôle

I. Evénements indépendants

On considère deux événements indépendants A et B de probabilités respectives 1/4 et 1/3.

1. La probabilité que les deux événements aient lieu vaut $\mathbb{P}(A \cap B)$, et en tenant compte de l'indépendance il vient
$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) = \frac{1}{12}.$$
2. La probabilité que l'un au moins des deux événements ait lieu est $\mathbb{P}(A \cup B)$, ce qui donne
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) = \frac{1}{3} + \frac{1}{4} - \frac{1}{12} = \frac{1}{2}.$$
3. La probabilité qu'exactement l'un des deux événements ait lieu s'écrit par exemple
$$\mathbb{P}((A \cup B) \setminus (A \cap B)) = \mathbb{P}(A \cup B) - \mathbb{P}(A \cap B) = \frac{1}{2} - \frac{1}{12} = \frac{5}{12}.$$

II. Un tirage en deux temps

Une boîte contient une balle noire et une balle blanche. Une balle est tirée au hasard dans la boîte : on remet celle-ci ainsi qu'une nouvelle balle de la même couleur. On tire alors une des trois balles au hasard dans la boîte.

1. Notons B_2 l'événement "la seconde balle tirée est blanche". Avec des notations naturelles, la probabilité de cet événement s'écrit :
$$\mathbb{P}(B_2) = \mathbb{P}(B_2|B_1)\mathbb{P}(B_1) + \mathbb{P}(B_2|N_1)\mathbb{P}(N_1) = \frac{2}{3} \times \frac{1}{2} + \frac{1}{3} \times \frac{1}{2} = \frac{1}{2}.$$
Ce résultat est logique puisque les 2 couleurs jouent des rôles complètement symétriques.
2. La probabilité que l'une au moins des deux balles tirées soit blanche vaut :
$$\mathbb{P}(B_2 \cup B_1) = 1 - \mathbb{P}(N_2 \cap N_1) = 1 - \mathbb{P}(N_2|N_1)\mathbb{P}(N_1) = 1 - \frac{2}{3} \times \frac{1}{2} = \frac{2}{3}.$$
3. La probabilité que la première balle tirée soit blanche, sachant que l'une au moins des deux balles tirée est blanche est
$$\mathbb{P}(B_1|B_2 \cup B_1) = \frac{\mathbb{P}(B_1 \cap (B_2 \cup B_1))}{\mathbb{P}(B_2 \cup B_1)} = \frac{\mathbb{P}(B_1)}{\mathbb{P}(B_2 \cup B_1)} = \frac{\frac{1}{2}}{\frac{2}{3}} = \frac{3}{4}.$$

III. Pièces défectueuses

Une usine produit des objets par boîtes de deux. Sur le long terme, on a constaté que : 92% des boîtes ne contiennent aucun objet défectueux ; 5% des boîtes contiennent exactement 1 objet défectueux ; 3% des boîtes contiennent 2 objets défectueux. Une boîte est choisie au hasard sur la chaîne de production et on tire au hasard un des deux objets de cette boîte.

1. Notons D l'événement "l'objet est défectueux", B_0 (respectivement B_1, B_2) l'événement "l'objet vient d'une boîte contenant 0 (respectivement 1, 2) objet défectueux". La probabilité que l'objet tiré soit défectueux se décompose alors via la formule des probabilités totales :

$$\mathbb{P}(D) = \mathbb{P}(D|B_0)\mathbb{P}(B_0) + \mathbb{P}(D|B_1)\mathbb{P}(B_1) + \mathbb{P}(D|B_2)\mathbb{P}(B_2).$$

Or il est clair que $\mathbb{P}(D|B_0) = 0$, $\mathbb{P}(D|B_1) = 0.5$ et $\mathbb{P}(D|B_2) = 1$, donc

$$\mathbb{P}(D) = 0 \times 0.92 + 0.5 \times 0.05 + 1 \times 0.03 = 0.055.$$

Il y a donc 5.5% de chances que cet objet soit défectueux.

2. Notons D' l'événement "l'autre objet est également défectueux". La probabilité cherchée s'écrit donc $\mathbb{P}(D'|D)$ et peut se calculer comme suit :

$$\mathbb{P}(D'|D) = \frac{\mathbb{P}(D' \cap D)}{\mathbb{P}(D)} = \frac{\mathbb{P}(B_2)}{\mathbb{P}(D)} = \frac{0.03}{0.055} = \frac{6}{11}.$$

IV. Lancer de dé

Un dé équilibré est lancé 10 fois de suite.

1. Notons p la probabilité d'au moins un 6 sur les 10 lancers. $1-p$ est donc la probabilité de n'avoir aucun 6 sur les 10 lancers, d'où

$$1 - p = \left(\frac{5}{6}\right)^{10} \Rightarrow p = 1 - \left(\frac{5}{6}\right)^{10} \approx 0.84.$$

2. Le nombre de 6 sur les 10 lancers est une variable aléatoire X qui suit une loi binomiale $\mathcal{B}(10, 1/6)$. On en déduit que le nombre moyen de 6 sur les 10 lancers vaut

$$\mathbb{E}[X] = 10 \times \frac{1}{6} = \frac{5}{3}.$$

3. Soit $U_1, \dots, U_{10}$ les nombres obtenus aux lancers successifs. Les variables U_i suivent toutes une loi uniforme sur $\{1, \dots, 6\}$, de moyenne 3.5. La somme des nombres obtenus est alors la variable $S = U_1 + \cdots + U_{10}$, dont la moyenne vaut :

$$\mathbb{E}[S] = \mathbb{E}[U_1 + \cdots + U_{10}] = \mathbb{E}[U_1] + \cdots + \cdots + \mathbb{E}[U_{10}] = 10 \times \mathbb{E}[U_1] = 35.$$

4. Notons X_5 (respectivement X_5') le nombre de 6 obtenus sur les 5 premiers (respectivement derniers) lancers et $X_{10} = X_5 + X_5'$ le nombre de 6 obtenus sur les 10 lancers. La probabilité d'obtenir exactement deux 6 lors des 5 premiers lancers sachant qu'il y en a eu 4 sur les 10 lancers s'écrit alors

$$\mathbb{P}(X_5 = 2|X_{10} = 4) = \frac{\mathbb{P}(\{X_5 = 2\} \cap \{X_{10} = 4\})}{\mathbb{P}(X_{10} = 4)} = \frac{\mathbb{P}(\{X_5 = 2\} \cap \{X_5' = 2\})}{\mathbb{P}(X_{10} = 4)}.$$

Les 5 premiers lancers sont bien entendu indépendants des 5 derniers, donc ceci s'écrit encore

$$\mathbb{P}(X_5 = 2|X_{10} = 4) = \frac{\mathbb{P}(X_5 = 2)\mathbb{P}(X_5' = 2)}{\mathbb{P}(X_{10} = 4)}.$$

Pour pouvoir plier l'affaire, il reste à noter que les variables aléatoires X_5, X_5' et X_{10} suivent toutes des lois binomiales, et plus précisément $X_5 \sim \mathcal{B}(5, 1/6)$, $X_5' \sim \mathcal{B}(5, 1/6)$, et $X_{10} \sim \mathcal{B}(10, 1/6)$. Ceci donne :

$$\mathbb{P}(X_5 = 2|X_{10} = 4) = \frac{\left(\binom{5}{2}(1/6)^2(5/6)^3\right)^2}{\binom{10}{4}(1/6)^4(5/6)^6} = \frac{10}{21}.$$

V. Le dé dyadique

On appelle "dé dyadique" un dé dont les faces sont numérotées respectivement 2, 4, 8, 16, 32 et 64. On jette un dé dyadique équilibré et on appelle X le résultat obtenu.

1. Puisque le dé est équilibré, la probabilité de chaque occurence est égale à 1/6, d'où :

$$\mathbb{E}[X] = \frac{1}{6}(2 + 4 + 8 + 16 + 32 + 64) = 21.$$

2. L'écart-type de X vaut quant à lui

$$\sigma(X) = \sqrt{\mathrm{Var}(X)} = \sqrt{\mathbb{E}[X^2] - (\mathbb{E}[X])^2}.$$

Il suffit donc de calculer

$$\mathbb{E}[X^2] = \frac{1}{6}(2^2 + 4^2 + 8^2 + 16^2 + 32^2 + 64^2) = 910$$

pour en déduire

$$\sigma(X) = \sqrt{910 - 21^2} \approx 21.66$$

3. Lorsque X_1 et X_2 sont deux variables indépendantes, elle sont a fortiori décorrélées et leur covariance est nulle. On en déduit en particulier que $\mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2]$.
4. On jette maintenant deux dés dyadiques équilibrés, appelons X_1 le résultat du premier, X_2 celui du second et $Y = X_1X_2$ le produit des deux. L'espérance de Y vaut donc

$$\mathbb{E}[Y] = \mathbb{E}[X_1X_2] = \mathbb{E}[X_1]\mathbb{E}[X_2] = (\mathbb{E}[X])^2 = 441.$$

5. Pour que le produit des deux dés fasse moins de 20, il faut avoir l'une des six combinaisons suivantes : (2,2), (2,4), (2,8), (4,2), (4,4), (8,2). Sur un total de 36 combinaisons équiprobables, ceci fait donc $\mathbb{P}(Y < 20) = 6/36 = 1/6$.

VI. Répartition des tailles

On suppose que dans une population, 1% des gens mesurent plus de 1m92. Supposons que vous tiriez au hasard (avec remise) 200 personnes dans cette population. Appelons X le nombre de personnes de plus de 1m92 dans votre échantillon.

1. La loi de X est binomiale $\mathcal{B}(200, 0.01)$.
2. On peut l'approcher par une loi de Poisson $\mathcal{P}(200 \times 0.01) = \mathcal{P}(2)$.
3. La probabilité que l'échantillon compte au moins 3 personnes de plus de 1m92 vaut donc

$$p = \mathbb{P}(X \geq 3) = 1 - \mathbb{P}(X < 3) = 1 - (\mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2)).$$

On peut calculer cette quantité directement avec la loi binomiale ou via l'approximation poissonienne. Dans le premier cas, ceci donne

$$p = 1 - \left(\binom{200}{0}0.01^0 0.99^{200} + \binom{200}{1}0.01^1 0.99^{199} + \binom{200}{2}0.01^2 0.99^{198}\right) \approx 0.323321$$

Tandis que dans le second, il vient

$$p = 1 - \left(e^{-2}\frac{2^0}{0!} + e^{-2}\frac{2^1}{1!} + e^{-2}\frac{2^2}{2!} \right) \approx 0.323324$$

L'approximation est donc excellente.

VII. Poisson en vrac

On considère une variable X distribuée selon une loi de Poisson de paramètre $\lambda > 0$.

1. $\mathbb{E}[3X+5] = 3\mathbb{E}[X] + 5 = 3\lambda + 5$.
2. $\mathrm{Var}(2X+1) = 4\mathrm{Var}(X) = 4\lambda$.
3. Ce dernier calcul mérite quelques détails. Le théorème de transfert donne

$$\mathbb{E}\left[\frac{1}{X+1}\right] = \sum_{n=0}^{+\infty} \frac{1}{n+1} e^{-\lambda} \frac{\lambda^n}{n!},$$

où l'on voit apparaître $(n+1)!$ au dénominateur, d'où l'idée de forcer un peu les choses au numérateur :

$$\mathbb{E}\left[\frac{1}{X+1}\right] = e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^n}{(n+1)!} = \frac{e^{-\lambda}}{\lambda} \sum_{n=0}^{+\infty} \frac{\lambda^{n+1}}{(n+1)!} = \frac{e^{-\lambda}}{\lambda} \sum_{n=1}^{+\infty} \frac{\lambda^n}{n!}$$

pour faire apparaître la série de l'exponentielle :

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

Université de Rennes 2
Licence MASS 2
Durée : 1 heure 45

Mercredi 15 Décembre 2010
Calculatrice autorisée
Aucun document

Contrôle de Probabilités

I. Variable à densité
Soit X une variable aléatoire de densité $f(x) = \frac{c}{x^4}\mathbb{1}_{\{x\geq 1\}}$.

1. Déterminer c pour que f soit bien une densité. Représenter f.
2. Calculer la fonction de répartition F et la représenter.
3. Déterminer la médiane de X, c'est-à-dire la valeur m telle que $\mathbb{P}(X > m) = 1/2$.
4. Calculer l'espérance de X et sa variance.
5. Déterminer le moment d'ordre 3 de X.

II. Diamètre d'une bille
Le diamètre d'une bille est distribué suivant une loi normale de moyenne 1 cm. On sait de plus qu'une bille a une chance sur trois d'avoir un diamètre supérieur à 1.1 cm.

1. Déterminer l'écart-type de cette distribution.
2. Quelle est la probabilité qu'une bille ait un diamètre compris entre 0.2 et 1 cm ?
3. Quelle est la valeur telle que 3/4 des billes aient un diamètre supérieur à cette valeur ?

III. Tchernobyl for ever
Soit T une variable aléatoire distribuée suivant une loi exponentielle de paramètre $\lambda > 0$.

1. Rappeler ce que valent densité, fonction de répartition, espérance et variance de T (on ne demande pas les calculs).
2. Pour tout $t > 0$, que vaut $\mathbb{P}(T > t)$?
3. On appelle demi-vie la durée h telle que $\mathbb{P}(T > h) = 1/2$. Déterminer h en fonction de λ.
4. Le strontium 90 est un composé radioactif très dangereux que l'on retrouve après une explosion nucléaire. Un atome de strontium 90 reste radioactif pendant une durée aléatoire T qui suit une loi exponentielle, durée au bout de laquelle il se désintègre. Sa demi-vie est d'environ 28 ans.
   (a) Déterminer le paramètre λ de la loi de T.
   (b) Calculer la probabilité qu'un atome reste radioactif durant au moins 50 ans.
   (c) Calculer le nombre d'années nécessaires pour que 99% du strontium 90 produit par une réaction nucléaire se soit désintégré.

IV. Jeu d'argent

Un jeu consiste à tirer, indépendamment et avec remise, des tickets d'une boîte. Il y a en tout 4 tickets, numérotés respectivement -2, -1, 0, 3. Votre "gain" X lors d'une partie correspond à la somme indiquée sur le ticket. Par exemple, si vous tirez le ticket numéroté -2, alors $X = -2$ et vous devez donner 2 €, tandis que si vous tirez le ticket 3, alors $X = 3$ et vous gagnez 3 €.

1. Donner la loi de X. Calculer son espérance et sa variance.
2. Vous jouez 100 fois de suite à ce jeu et on note S votre gain après 100 parties. En notant X_1 le gain à la première partie, X_2 le gain à la deuxième partie, ..., X_{100} le gain à la centième partie, exprimer S en fonction des X_i.
3. En déduire l'espérance de S et sa variance.
4. Par quelle loi normale peut-on approcher S ? En déduire la probabilité que votre gain sur 100 parties dépasse 25 €.

V. Rubrique à brac

1. Soit T une variable aléatoire suivant une loi géométrique de paramètre p, $0 < p < 1$. Rappeler la loi de T, son espérance et sa variance.
2. Vous demandez à des personnes choisies au hasard dans la rue leur mois de naissance jusqu'à en trouver une née en décembre. Quel est (approximativement) le nombre moyen de personnes que vous allez devoir interroger ?
3. On jette une pièce équilibrée et on appelle X le nombre de lancers nécessaires pour que Pile apparaisse. Quelle est la loi de X ?
4. Grâce aux moments de X, montrer que $\sum_{n=1}^{+\infty} \frac{n^2}{2^n} = 6$.
5. Alice et Bob jouent au jeu suivant : Alice lance une pièce équilibrée jusqu'à ce que Pile apparaisse. Si Pile apparaît dès le premier lancer, Bob lui donne 4 € ; si Pile n'apparaît qu'au deuxième lancer, Bob lui donne 1 € ; si Pile n'apparaît qu'au troisième lancer, elle donne 4 € à Bob ; si Pile n'apparaît qu'au quatrième lancer, elle donne 11 € à Bob, etc. De façon générale, le "gain" d'Alice si Pile n'apparaît qu'au n-ème lancer est $5 - n^2$. Notons G la variable aléatoire correspondant à ce gain.
   (a) Calculer la probabilité qu'Alice perde de l'argent lors d'une partie.
   (b) Calculer l'espérance de G.
   (c) Si vous deviez jouer une seule partie, préféreriez-vous être à la place d'Alice ou à la place de Bob ? Et si vous deviez en jouer 100 ?

VI. Ascenseur pour l'échafaud

Un ascenseur dessert les 10 étages d'un immeuble, 12 personnes le prennent au rez-de-chaussée et chacune choisit un des 10 étages au hasard.

1. Soit X_1 la variable aléatoire valant 1 si au moins une personne choisit le 1er étage, 0 sinon. Calculer $\mathbb{P}(X_1 = 1)$ et en déduire la moyenne de X_1.
2. De façon générale, soit X_i la variable aléatoire valant 1 si au moins une personne choisit l'étage i, 0 sinon. Exprimer le nombre d'étages auxquels l'ascenseur s'arrête en fonction des X_i. En déduire le nombre moyen d'étages auxquels l'ascenseur s'arrête.
3. (Bonus) Généralisation : montrer que pour t étages et n personnes, le nombre moyen d'étages desservis est $t(1 - (1 - \frac{1}{t})^n)$. Que devient cette quantité :
   (a) lorsque t tend vers l'infini avec n fixé ? Interpréter.
   (b) lorsque n tend vers l'infini avec t fixé ? Interpréter.

Université de Rennes 2
Licence MASS 2
Durée : 2 heures

Mercredi 15 Décembre 2010
Calculatrice autorisée
Aucun document

Corrigé du Contrôle

I. Variable à densité
Soit X une variable aléatoire de densité $f(x) = \frac{c}{x^4}\mathbb{1}_{\{x\geq 1\}}$.

1. Pour que f soit bien une densité, il faut que la constante c soit positive et telle que l'intégrale de f soit égale à 1, c'est-à-dire :

$$1 = \int_1^{+\infty} \frac{c}{x^4}dx = c\left[\frac{-1}{3x^3}\right]_1^{+\infty} = \frac{c}{3},$$

donc $c = 3$ et $f(x) = \frac{3}{x^4}\mathbb{1}_{\{x\geq 1\}}$. Cette densité est représentée figure A.3 à gauche.

2. La fonction de répartition F est nulle à gauche de 1, et pour $x \geq 1$ on a :

$$F(x) = \int_1^x \frac{3}{t^4}dt = 1 - \frac{1}{x^3}.$$

Cette fonction de répartition est représentée figure A.3 à droite.

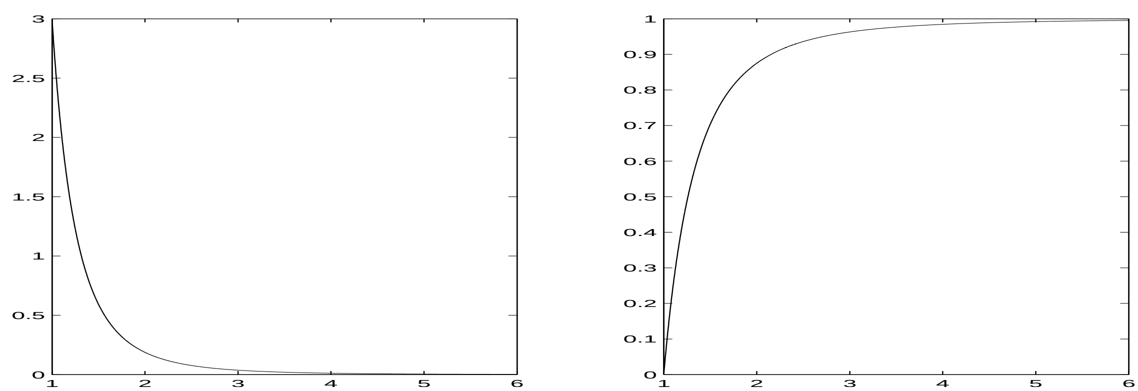


FIGURE A.3 – Densité f et fonction de répartition F.

3. Puisque $\mathbb{P}(X > m) = 1 - \mathbb{P}(X \leq m) = 1 - F(m)$, il nous suffit de résoudre :

$$1 - F(m) = \frac{1}{2} \Leftrightarrow \frac{1}{m^3} = \frac{1}{2} \Leftrightarrow m = 2^{1/3} \approx 1.26.$$

4. L'espérance de X se calcule comme suit :

$$\mathbb{E}[X] = \int_1^{+\infty} \frac{3x}{x^4} dx = \left[\frac{-3}{2x^2}\right]_1^{+\infty} = \frac{3}{2}.$$

Pour la variance, on commence par calculer le moment d'ordre 2 :

$$\mathbb{E}[X^2] = \int_1^{+\infty} \frac{3x^2}{x^4} dx = \left[\frac{-3}{x}\right]_1^{+\infty} = 3,$$

d'où l'on déduit :

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{3}{4}.$$

5. Le moment d'ordre 3 de X se calcule a priori de la même façon :

$$\mathbb{E}[X^3] = \int_1^{+\infty} \frac{3x^3}{x^4} dx = [3 \ln x]_1^{+\infty} = +\infty...$$

... Blood on the wall ! En fait X n'admet pas de moment d'ordre 3.

II. Diamètre d'une bille

Le diamètre d'une bille est distribué suivant une loi normale de moyenne 1 cm. On sait de plus qu'une bille a une chance sur trois d'avoir un diamètre supérieur à 1.1 cm.

1. Notons σ l'écart-type de cette distribution et X la variable correspondant au diamètre d'une bille, alors $Y = (X - 1)/\sigma$ suit une loi normale centrée réduite, de fonction de répartition notée comme d'habitude Φ. Le texte nous indique que $\mathbb{P}(X > 1.1) = 1/3$. Procédons par centrage et réduction de X :

$$\mathbb{P}(X > 1.1) = \frac{1}{3} \Leftrightarrow \mathbb{P}\left(\frac{X-1}{\sigma} > \frac{1.1-1}{\sigma}\right) = \frac{1}{3} \Leftrightarrow \mathbb{P}\left(Y > \frac{0.1}{\sigma}\right) = \frac{1}{3} \Leftrightarrow 1-\Phi(0.1/\sigma) = \frac{1}{3}.$$

Il nous reste donc à trouver dans la table la valeur $q = 0.1/\sigma$ telle que $\Phi(q) = 2/3$, ce qui donne $0.1/\sigma \approx 0.43$, donc $\sigma \approx 0.23$ cm.

2. Avec la même technique et les mêmes notations, la probabilité qu'une bille ait un diamètre compris entre 0.2 et 1 cm vaut :

$$\mathbb{P}(0.2 < X < 1) = \mathbb{P}\left(\frac{0.2-1}{0.23} < \frac{X-1}{0.23} < \frac{1-1}{0.23}\right) = \mathbb{P}(-3.48 < Y < 0) = \Phi(0) - \Phi(-3.48),$$

d'où : $\mathbb{P}(0.2 < X < 1) = \Phi(0) + \Phi(3.48) - 1 \approx \frac{1}{2}$, puisque d'après la table $0.995 \leq \Phi(3.48) < 1$, donc $\Phi(3.48) \approx 1$.

3. Soit q la valeur telle que 3/4 des billes aient un diamètre supérieur à q. Il nous faut donc résoudre :

$$\mathbb{P}(X > q) = \frac{3}{4} \Leftrightarrow \mathbb{P}\left(Y > \frac{q-1}{0.23}\right) = \frac{3}{4} \Leftrightarrow \Phi((1-q)/0.23) = \frac{3}{4},$$

ce qui donne via la table $(1-q)/0.23 \approx 0.67$ donc $q \approx 0.85$.

III. Tchernobyl for ever

Soit T une variable aléatoire distribuée suivant une loi exponentielle de paramètre $\lambda > 0$.

1. Si on note f la densité, F la fonction de répartition, alors

$$f(t) = \lambda e^{-\lambda t} \mathbb{1}_{\{t \geq 0\}} \qquad \text{et} \qquad F(t) = (1 - e^{-\lambda t}) \mathbb{1}_{\{t \geq 0\}}.$$

   On montre que $\mathbb{E}[T] = \frac{1}{\lambda}$ et $\text{Var}(T) = \frac{1}{\lambda^2}$.
2. Pour tout $t > 0$, on a donc $\mathbb{P}(T > t) = 1 - F(t) = e^{-\lambda t}$.
3. La demi-vie vérifie donc :

$$e^{-\lambda h} = \frac{1}{2} \Leftrightarrow -\lambda h = \ln(1/2) \Leftrightarrow h = \frac{\ln 2}{\lambda}.$$

4. Le strontium 90 est un composé radioactif très dangereux que l'on retrouve après une explosion nucléaire. Un atome de strontium 90 reste radioactif pendant une durée aléatoire T qui suit une loi exponentielle, durée au bout de laquelle il se désintègre. Sa demi-vie est d'environ 28 ans.
   (a) D'après la question précédente, le taux de désintégration vaut donc $\lambda = \ln 2/h \approx 0.0248$.
   (b) La probabilité qu'un atome reste radioactif durant au moins 50 ans est :

$$\mathbb{P}(T > 50) = e^{-50\lambda} \approx e^{-50 \times 0.0248} \approx 0.29$$

   (c) Calculer le nombre d'années nécessaires pour que 99% du strontium 90 produit par une réaction nucléaire se soit désintégré revient à trouver la durée t telle que $\mathbb{P}(T > t) = 0.01$, c'est-à-dire :

$$e^{-0.0248t} = 0.01 \Leftrightarrow -0.0248t = \ln(0.01) \Leftrightarrow t = \frac{\ln(0.01)}{-0.0248} \approx 185.7$$

   Il faut donc près de 186 ans pour qu'il ne reste plus que 1% de strontium 90.

IV. Jeu d'argent

Un jeu consiste à tirer, indépendamment et avec remise, des tickets d'une boîte. Il y a en tout 4 tickets, numérotés respectivement -2, -1, 0, 3. Votre "gain" X lors d'une partie correspond à la somme indiquée sur le ticket. Par exemple, si vous tirez le ticket numéroté -2, alors $X = -2$ et vous devez donner 2 €, tandis que si vous tirez le ticket 3, alors $X = 3$ et vous gagnez 3 €.

1. X prend les quatres valeurs -2, -1, 0, 3 avec la même probabilité, c'est-à-dire $1/4$. On vérifie sans difficultés que sa moyenne est nulle (c'est donc un jeu équitable) et que sa variance vaut $7/2$.
2. On a clairement $S = X_1 + \cdots + X_{100}$.
3. Par linéarité de l'espérance et du fait que les X_i ont la même loi, on a donc $\mathbb{E}[S] = 100\mathbb{E}[X_1] = 0$. Du fait que les X_i ont la même loi et sont indépendantes, on déduit $\text{Var}(S) = 100\text{Var}(X_1) = 350$.
4. Puisque S est la somme d'un grand nombre de variables indépendantes et de même loi, le théorème central limite permet d'approcher la loi de S par une loi normale centrée et de variance 350. La probabilité que notre gain sur 100 parties dépasse 25 € est donc :

$$\mathbb{P}(S > 25) = \mathbb{P}\left(\frac{S}{\sqrt{350}} > \frac{25}{\sqrt{350}}\right) = 1 - \Phi(25/\sqrt{350}) \approx 0.09.$$

   Sur 100 parties, on a donc environ 9% de chances de gagner plus de 25 €.

V. Rubrique à brac

1. Dire que T est une variable aléatoire suivant une loi géométrique de paramètre p signifie que T est à valeurs dans $\mathbb{N}^*$, avec
$$\forall n \in \mathbb{N}^* \qquad \mathbb{P}(T=n) = p(1-p)^{n-1}$$
Son espérance vaut $\mathbb{E}[T] = 1/p$ et sa variance $\mathrm{Var}(T) = q/p^2$.
2. Si on suppose que le nombre de naissance est équiréparti sur les 12 mois de l'année (ceci signifie en particulier qu'on ne tient pas compte du fait que certains mois ont plus de jours que d'autres et que les naissances ne sont en fait pas équiréparties sur l'année) et que les individus interrogés sont indépendants du point de vue du mois de naissance, alors le nombre de personnes que l'on doit interroger suit une loi géométrique de paramètre $1/12$. Ainsi le nombre moyen de personnes à interroger est $\mathbb{E}[T] = 12$.
3. La loi de X est gémétrique de paramètre $1/2$.
4. On reconnaît dans la série concernée le moment d'ordre 2 d'une variable X distribuée suivant une loi géométrique de paramètre $1/2$:
$$X \sim \mathcal{G}(1/2) \;\Rightarrow\; \mathbb{E}[X^2] = \sum_{n=1}^{+\infty} \frac{n^2}{2^n}.$$
Or $\mathbb{E}[X^2] = \mathrm{Var}(X) + (\mathbb{E}[X])^2 = 2 + 4 = 6$.
5. (a) Pour calculer la probabilité p qu'Alice perde de l'argent lors d'une partie, on passe par l'événement complémentaire, à savoir le fait qu'Alice gagne de l'argent, ce qui arrive si et seulement si Pile apparaît au premier ou au deuxième lancer. En notant toujours X la variable géométrique introduite précédemment, on a donc :
$$p = 1 - (\mathbb{P}(G=4) + \mathbb{P}(G=1)) = 1 - (\mathbb{P}(X=1) + \mathbb{P}(X=2)) = 1 - \left(\frac{1}{2} + \frac{1}{4}\right) = \frac{1}{4}.$$
   (b) La variable G prend les valeurs $5 - n^2$ avec les probabilités $1/2^n$ pour tout $n \in \mathbb{N}^*$, donc son espérance vaut :
$$\mathbb{E}[G] = \sum_{n=1}^{+\infty} \frac{5-n^2}{2^n} = 5\sum_{n=1}^{+\infty} \frac{1}{2^n} - \sum_{n=1}^{+\infty} \frac{n^2}{2^n} = 5 - 6 = -1.$$
   On pouvait aussi trouver ce résultat en appliquant le théorème de transfert puisque la variable G est tout simplement égale à $5 - X^2$, d'où :
$$\mathbb{E}[G] = \mathbb{E}[5 - X^2] = 5 - \mathbb{E}[X^2] = 5 - 6 = -1.$$
   (c) Sur une partie, Alice a 3 chances sur 4 de gagner de l'argent, donc sur ce principe on pourrait préférer être du côté d'Alice. Néanmoins, on a vu aussi qu'en moyenne, par partie, elle perd 1 €. Ceci signifie en gros que lorsqu'elle gagne (ce qui arrive 3 fois sur 4) elle gagne peu, tandis que lorsqu'elle perd (ce qui arrive 1 fois sur 4), elle peut perdre beaucoup. Si on se met à la place de Bob, tout se passe un peu comme s'il achetait un ticket de Loto à prix variable (1 ou 4 €) : il ne va probablement rien récupérer, mais s'il gagne il peut éventuellement gagner beaucoup. Tout dépend donc si on est joueur ou non...
   Il n'y a par contre plus aucune ambiguïté lorsqu'Alice et Bob jouent un grand nombre de parties. Notons σ l'écart-type de G et $S = G_1 + \cdots + G_{100}$ le gain d'Alice sur 100 parties, alors le théorème central limite dit que S est approximativement distribuée selon une loi normale de moyenne $100 \times \mathbb{E}[G] = -100$ et d'écart-type 10σ. Puisque la moyenne de cette loi normale est négative, S a plus de chances d'être négatif que positif, donc on préférera être à la place de Bob.

VI. Ascenseur pour l'échafaud

Un ascenseur dessert les 10 étages d'un immeuble, 12 personnes le prennent au rez-de-chaussée et chacune choisit un des 10 étages au hasard.

1. X_1 est nulle si aucune des 12 personnes ne choisit l'étage 1, ce qui arrive avec probabilité :

$$\mathbb{P}(X_1 = 0) = \left(\frac{9}{10}\right)^{12} \Rightarrow \mathbb{P}(X_1 = 1) = 1 - \left(\frac{9}{10}\right)^{12} \Rightarrow \mathbb{E}[X_1] = 1 - \left(\frac{9}{10}\right)^{12}.$$

2. Soit N le nombre (aléatoire) d'étages auxquels l'ascenseur s'arrête. Par définition, $N = X_1 + \cdots + X_{10}$. Les X_i ont toutes la même loi, donc le nombre moyen d'étages desservis est :

$$\mathbb{E}[N] = \mathbb{E}[X_1 + \cdots + X_{10}] = 10\,\mathbb{E}[X_1] = 10\left(1 - \left(\frac{9}{10}\right)^{12}\right).$$

Remarque : Les variables X_i sont de même loi mais pas indépendantes puisqu'il est par exemple clair qu'elles ne peuvent pas être toutes nulles simultanément :

$$\mathbb{P}(X_1 = 0, \ldots, X_{10} = 0) = 0 \neq \left(\frac{9}{10}\right)^{120} = \mathbb{P}(X_1 = 0) \ldots \mathbb{P}(X_{10} = 0).$$

En particulier le calcul de la variance de N n'est pas aussi simple que celui de l'espérance.

3. La généralisation du raisonnement précédent pour t étages et n personnes est immédiate :

$$\mathbb{E}[N] = t\left(1 - \left(\frac{t-1}{t}\right)^n\right).$$

(a) Lorsque t tend vers l'infini avec n fixé, on a

$$\left(\frac{t-1}{t}\right)^n = \left(1 - \frac{1}{t}\right)^n \sim 1 - \frac{n}{t}$$

donc

$$\mathbb{E}[N] = t\left(1 - \left(\frac{t-1}{t}\right)^n\right) \sim t\left(1 - \left(1 - \frac{n}{t}\right)\right) = n,$$

autrement dit $\lim_{t\to\infty} \mathbb{E}[N] = n$. Ceci est logique : lorsqu'il y a un très grand nombre d'étages, il y a très peu de chances que plusieurs personnes choisissent le même, donc le nombre d'étages desservis correspond approximativement au nombre de personnes.

(b) A contrario, lorsque n tend vers l'infini avec t fixé, on a

$$\left|\frac{t-1}{t}\right| < 1 \Rightarrow \left(\frac{t-1}{t}\right)^n \xrightarrow[n\to\infty]{} 0$$

donc

$$\mathbb{E}[N] = t\left(1 - \left(\frac{t-1}{t}\right)^n\right) \xrightarrow[n\to\infty]{} t.$$

Ceci est logique aussi : lorsque le nombre de personnes est beaucoup plus grand que le nombre d'étages, l'ascenseur s'arrête en général à tous les étages.

Université Rennes 2
Licence MASS 2

Lundi 7 Novembre 2011
Durée : 1h30
Calculatrice autorisée
Aucun document

Contrôle de Probabilités

I. Circuits intégrés

Un atelier reçoit 5000 circuits intégrés : 1000 en provenance de l'usine A et 4000 en provenance de l'usine B. 10% des circuits fabriqués par l'usine A et 5% de ceux fabriqués par l'usine B sont défectueux.

1. On choisit au hasard un circuit intégré à l'atelier. Quelle est la probabilité qu'il soit défectueux ?
2. Sachant qu'un circuit choisi est défectueux, quelle est la probabilité qu'il vienne de l'usine A ?

II. Systèmes de contrôle

Deux systèmes de contrôle électrique opèrent indépendamment et sont sujets à un certain nombre de pannes par jour. Les probabilités p_n (respectivement q_n) régissant le nombre n de pannes par jour pour le système 1 (resp. 2) sont données dans les tableaux suivants :

Système 1

n	p_n
0	0.07
1	0.35
2	0.34
3	0.18
4	0.06

Système 2

n	q_n
0	0.10
1	0.20
2	0.50
3	0.17
4	0.03

1. Calculer les probabilités des événements suivants :
   (a) Le système 2 a au moins 2 pannes dans la journée.
   (b) Il se produit une seule panne dans la journée.
   (c) Le système 1 a le même nombre de pannes que le système 2.
2. Quel est le nombre moyen de pannes du système 1 par jour ? Comparer à celui du système 2.
3. Supposons que l'équipe de mécaniciens ne puisse réparer qu'un maximum de 6 pannes par jour. Dans quelle proportion du temps ne pourra-t-elle pas suffire à la tâche ?

III. Utilité d'un testeur

Une chaîne de montage d'ordinateurs utilise un lot de processeurs contenant 2% d'éléments défectueux. En début de chaîne, chaque processeur est vérifié par un testeur dont la fiabilité n'est pas parfaite, de telle sorte que la probabilité que le testeur déclare le processeur bon (resp. mauvais) sachant que le processeur est réellement bon (resp. mauvais) vaut 0.95 (resp. 0.94).

1. Calculer la probabilité qu'un processeur soit déclaré bon.
2. Calculer la probabilité qu'un processeur déclaré bon soit réellement bon.
3. Calculer la probabilité qu'un processeur déclaré mauvais soit réellement mauvais.
4. (Bonus) Le testeur est-il utile ?

IV. Kramer contre Kramer

On effectue des tirages sans remise dans une urne contenant initialement 3 boules rouges et 3 boules noires jusqu'à obtenir une boule noire. On appelle X le numéro du tirage de cette boule noire (ainsi $X = 1$ si la première boule tirée est noire).

1. Quelles valeurs peut prendre la variable aléatoire X ? Avec quelles probabilités ?
2. Représenter sa fonction de répartition F.
3. Calculer l'espérance et la variance de X.
4. On classe 3 hommes et 3 femmes selon leur note à un examen. On suppose toutes les notes différentes et tous les classements équiprobables. On appelle R le rang de la meilleure femme (par exemple $R = 2$ si le meilleur résultat a été obtenu par un homme et le suivant par une femme). Donner la loi de R.

V. Loterie

Dans une loterie, un billet coûte 1 euro. Le nombre de billets émis est 90000, numérotés de 10000 à 99999, chaque billet comportant donc 5 chiffres. Un numéro gagnant est lui-même un nombre entre 10000 et 99999. Lorsque vous achetez un billet, vos gains possibles sont les suivants :

votre billet correspond au numéro gagnant	10000 euros
vos 4 derniers chiffres sont ceux du numéro gagnant	1000 euros
vos 3 derniers chiffres sont ceux du numéro gagnant	100 euros

1. Quelle est la probabilité d'avoir le numéro gagnant ?
2. Quelle est la probabilité de gagner 1000 euros ?
3. Quelle est la probabilité de gagner 100 euros ?
4. Déterminer votre bénéfice moyen lorsque vous achetez un billet.

Université Rennes 2
Licence MASS 2

Lundi 7 Novembre 2011
Durée : 1h30
Calculatrice autorisée
Aucun document

Corrigé du Contrôle

I. Circuit intégré

Un atelier reçoit 5000 circuits intégrés : 1000 en provenance de l'usine A et 4000 en provenance de l'usine B. 10% des circuits fabriqués par l'usine A et 5% de ceux fabriqués par l'usine B sont défectueux.

1. Par la formule des probabilités totales, la probabilité qu'un circuit pris au hasard soit défectueux est (avec des notations évidentes) :

$$\mathbb{P}(D) = \mathbb{P}(D|A)\mathbb{P}(A) + \mathbb{P}(D|B)\mathbb{P}(B) = \frac{10}{100} \times \frac{1000}{5000} + \frac{5}{100} \times \frac{4000}{5000} = 0.06$$

2. Sachant qu'un circuit choisi est défectueux, la probabilité qu'il vienne de l'usine A se déduit alors de la formule de Bayes :

$$\mathbb{P}(A|D) = \frac{\mathbb{P}(D|A)\mathbb{P}(A)}{\mathbb{P}(D)} = \frac{\frac{10}{100} \times \frac{1000}{5000}}{0.06} = \frac{1}{3}.$$

II. Systèmes de contrôle

Deux systèmes de contrôle électrique opèrent indépendamment et sont sujets à un certain nombre de pannes par jour. Les probabilités p_n (respectivement q_n) régissant le nombre n de pannes par jour pour le système 1 (resp. 2) sont données dans les tableaux suivants :

Système 1

n	p_n
0	0.07
1	0.35
2	0.34
3	0.18
4	0.06

Système 2

n	q_n
0	0.10
1	0.20
2	0.50
3	0.17
4	0.03

1. (a) La probabilité P que le système 2 ait au moins 2 pannes dans la journée s'écrit :

$$P = q_2 + q_3 + q_4 = 0.50 + 0.17 + 0.03 = 0.70$$

   (b) Puisque les deux systèmes sont indépendants, la probabilité P' qu'il se produise une seule panne dans la journée est :

$$P' = p_0q_1 + p_1q_0 = 0.049$$

(c) A nouveau par indépendance des deux systèmes, la probabilité P'' que le système 1 ait le même nombre de pannes que le système 2 vaut :

$$P'' = p_0q_0 + p_1q_1 + p_2q_2 + p_3q_3 + p_4q_4 \approx 0.28$$

2. Soit X_1 la variable aléatoire correspondant au nombre de pannes du système 1 en une journée. Le nombre moyen de pannes par jour du système 1 est donc :

$$\mathbb{E}[X_1] = 0 \times p_0 + 1 \times p_1 + 2 \times p_2 + 3 \times p_3 + 4 \times p_4 = 1.81$$

De même, le nombre moyen de pannes par jour du système 2 est égale à :

$$\mathbb{E}[X_2] = 0 \times q_0 + 1 \times q_1 + 2 \times q_2 + 3 \times q_3 + 4 \times q_4 = 1.83$$

En moyenne, le système 2 a donc un peu plus de pannes par jour.

3. La proportion du temps Q durant laquelle l'équipe de réparation ne pourra pas suffire à la tâche correspond à la probabilité qu'il y ait plus de 6 pannes dans la même journée, c'est-à-dire à la probabilité qu'il y en ait 7 ou 8, soit :

$$Q = p_4q_3 + p_3q_4 + p_4q_4 = 0.0174 \approx 0.017$$

III. Utilité d'un testeur

Une chaîne de montage d'ordinateurs utilise un lot de processeurs contenant 2% d'éléments défectueux. En début de chaîne, chaque processeur est vérifié par un testeur dont la fiabilité n'est pas parfaite, de telle sorte que la probabilité que le testeur déclare le processeur bon (resp. mauvais) sachant que le processeur est réellement bon (resp. mauvais) vaut 0.95 (resp. 0.94).

1. Par la formule des probabilités totales, la probabilité qu'un processeur soit déclaré bon est :

$$\mathbb{P}(DB) = \mathbb{P}(DB|B)\mathbb{P}(B) + \mathbb{P}(DB|M)\mathbb{P}(M)$$

Or $\mathbb{P}(DB|M) = 1 - \mathbb{P}(DM|M) = 1 - 0.94 = 0.06$, d'où :

$$\mathbb{P}(DB) = 0.95 \times 0.98 + 0.06 \times 0.02 = 0.9322 \approx 0.932$$

2. La probabilité qu'un processeur déclaré bon soit réellement bon s'en déduit :

$$\mathbb{P}(B|DB) = \frac{\mathbb{P}(DB|B)\mathbb{P}(B)}{\mathbb{P}(DB)} \approx 0.999$$

3. Par le même raisonnement, la probabilité qu'un processeur déclaré mauvais soit réellement mauvais est :

$$\mathbb{P}(M|DM) = \frac{\mathbb{P}(DM|M)\mathbb{P}(M)}{\mathbb{P}(DM)},$$

avec $\mathbb{P}(DM) = 1 - \mathbb{P}(DB) \approx 0.068$, donc :

$$\mathbb{P}(M|DM) = \frac{0.94 \times 0.02}{0.068} \approx 0.28$$

4. Il y a plusieurs réponses possibles à cette question. La première revient à comparer le pourcentage d'ordinateurs défectueux sans et avec testeur. De ce point de vue la réponse est claire : sans testeur, il y en avait 2% ; avec testeur, il n'y en a plus qu'environ 0,1% puisque les mauvais processeurs déclarés bons (faux négatifs) sont en proportion

$$\mathbb{P}(DB \cap M) = \mathbb{P}(DB|M)\mathbb{P}(M) = 0.06 \times 0.02 \approx 0.001$$

Ainsi, le taux de processeurs défectueux effectivement utilisés pour le montage d'ordinateurs a été divisé par 20, ce qui peut sembler tout à fait satisfaisant. Néanmoins, cette réponse n'est pas complètement convaincante car elle ne tient pas compte du fait que ceci s'est fait au détriment de processeurs bons qui ont été déclarés mauvais (faux positifs). Une façon de préciser ce point est la suivante : d'un point de vue purement comptable, le testeur est utile s'il fait gagner de l'argent à l'entreprise. Grosso modo, la réponse dépend donc de ce qui coûte le plus entre :
– vendre un ordinateur défectueux et devoir le changer ensuite,
– ne pas vendre un ordinateur qui fonctionnerait.
Supposons que le bénéfice retiré de la vente d'un ordinateur opérationnel est b euros et que le déficit engendré par la vente d'un ordinateur défectueux est d euros. Dans le premier modèle, sans testeur, le bénéfice moyen par ordinateur est donc :

$$B_1 = \mathbb{P}(B) \times b - \mathbb{P}(M) \times d = 0.98b - 0.02d$$

En effet, en moyenne sur 1000 processeurs, 980 ont rapporté b euros et 20 ont fait perdre d euros.
Dans le second modèle, le bénéfice moyen par ordinateur est par contre :

$$B_2 = \mathbb{P}(DB \cap B) \times b - \mathbb{P}(DB \cap M) \times d - \mathbb{P}(DM \cap B) \times b + \mathbb{P}(DM \cap M) \times 0,$$

avec :
– $\mathbb{P}(DB \cap B) = \mathbb{P}(DB|B)\mathbb{P}(B) = 0.95 \times 0.98 = 0.931$;
– $\mathbb{P}(DB \cap M) = \mathbb{P}(DB|M)\mathbb{P}(M) = 0.06 \times 0.02 \approx 0.001$;
– $\mathbb{P}(DM \cap B) = \mathbb{P}(DM|B)\mathbb{P}(B) = 0.05 \times 0.98 \approx 0.049$;
– $\mathbb{P}(DM \cap M) = \mathbb{P}(DM|M)\mathbb{P}(M) = 0.94 \times 0.02 \approx 0.019$
En effet, en moyenne sur 1000 processeurs, 931 sont déclarés bons et le sont, donc rapportent chacun b euros, 1 est déclaré bon mais est mauvais donc fait perdre d euros, et 49 ont été déclarés mauvais alors qu'ils étaient bons, donc il n'ont pas été vendus alors qu'ils auraient dû rapporter chacun b euros. Les 19 processeurs restants étant mauvais et déclarés comme tels, ils n'ont engendré ni perte ni profit. Bref, le bénéfice moyen par ordinateur est cette fois :

$$B_2 = 0.882b - 0.001d.$$

Avec ce point de vue, le testeur est utile si $B_2 > B_1$, c'est-à-dire si :

$$0.882b - 0.001d > 0.98b - 0.02d \;\Leftrightarrow\; d > \frac{0.098}{0.019} \times b \approx 5.2 \times b$$

Ainsi, si le déficit engendré par la vente d'un ordinateur défectueux est environ 5 fois plus élevé que le bénéfice engendré par la vente d'un ordinateur qui marche, alors le testeur est utile. Sinon on peut s'en passer.

IV. Kramer contre Kramer

On effectue des tirages sans remise dans une urne contenant initialement 3 boules rouges et 3 boules noires jusqu'à obtenir une boule noire. On appelle X le numéro du tirage de cette boule noire.

1. La variable aléatoire X peut prendre les valeurs 1,2,3,4. En notant N_i (respectivement R_i) l'événement : "Le tirage i est une boule noire (resp. rouge)", on obtient pour la loi de X les probabilités suivantes :
   – $\mathbb{P}(X = 1) = \mathbb{P}(N_1) = 3/6 = 10/20$;
   – $\mathbb{P}(X = 2) = \mathbb{P}(N_2 \cap R_1) = \mathbb{P}(N_2|R_1)\mathbb{P}(R_1) = 3/5 \times 3/6 = 3/10 = 6/20$;

– $\mathbb{P}(X=3)=\mathbb{P}(N_3\cap R_2\cap R_1)=\mathbb{P}(N_3|R_1\cap R_2)\mathbb{P}(R_1\cap R_2)=\mathbb{P}(N_3|R_1R_2)\mathbb{P}(R_2|R_1)\mathbb{P}(R_1)$, c'est-à-dire $\mathbb{P}(X=3)=3/4\times 2/5\times 3/6=3/20$;
– $\mathbb{P}(X=4)=1-(\mathbb{P}(X=1)+\mathbb{P}(X=2)+\mathbb{P}(X=3))=1/20$.

2. La fonction de répartition F de la variable aléatoire X est représentée figure A.4.

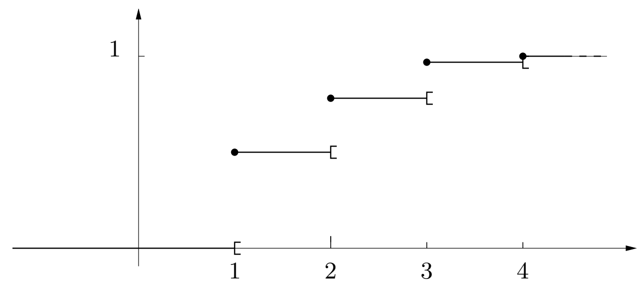


Figure A.4 – Fonction de répartition F de la variable X.

3. L'espérance de X vaut :
$$\mathbb{E}[X]=1\times\frac{10}{20}+2\times\frac{6}{20}+3\times\frac{3}{20}+4\times\frac{1}{20}=\frac{7}{4}=1.75$$
Sa variance vaut $\mathrm{Var}(X)=\mathbb{E}[X^2]-(\mathbb{E}[X])^2$, avec :
$$\mathbb{E}[X^2]=1^2\times\frac{10}{20}+2^2\times\frac{6}{20}+3^2\times\frac{3}{20}+4^2\times\frac{1}{20}=\frac{77}{20}=3.85$$
d'où $\mathrm{Var}(X)=63/80=0.7875$.
4. Puisque toutes les notes sont différentes et tous les classements équiprobables, la loi de R est la même que celle de X vue précédemment.

V. Loterie

Dans une loterie, un billet coûte 1 euro. Le nombre de billets émis est 90000, numérotés de 10000 à 99999, chaque billet comportant donc 5 chiffres. Un numéro gagnant est lui-même un nombre entre 10000 et 99999. Lorsque vous achetez un billet, vos gains possibles sont les suivants :

votre billet correspond au numéro gagnant	10000 euros
vos 4 derniers chiffres sont ceux du numéro gagnant	1000 euros
vos 3 derniers chiffres sont ceux du numéro gagnant	100 euros

1. La probabilité d'avoir le numéro gagnant est égale à 1/90000.
2. Le numéro gagnant étant fixé (par exemple 23456), vous gagnez 1000 euros si vous avez l'un des numéros 13456, 33456, 43456, ..., 93456. Autrement dit, 8 possiblités qui correspondent aux 8 choix possibles pour la première décimale (ne pas oublier que 23456 ne convient pas puisque dans ce cas vous gagnez 10000 euros). Ainsi, la probabilité de gagner 1000 euros est 8/90000.
3. Même raisonnement : a priori il y a 90 choix possibles pour les deux premières décimales, mais 9 d'entre eux ne conviennent pas (gains de 1000 ou 10000 euros). Ainsi la probabilité de gagner 100 euros est 81/90000.
4. Le gain moyen par billet est donc
$$\mathbb{E}[G]=10000\times\frac{1}{90000}+1000\times\frac{8}{90000}+100\times\frac{81}{90000}=\frac{29}{100}=0.29$$
Puisqu'un billet coûte 1 euro, votre bénéfice moyen est de -0.71 euro par billet. En d'autres termes, si vous jouez 100 fois de suite à ce jeu, vous perdrez en moyenne 71 euros.

Université Rennes 2
Licence MASS 2

Lundi 12 Décembre 2011
Durée : 2 heures
Calculatrice autorisée
Aucun document

Contrôle de Probabilités

I. Durée de vie d'un processeur
On modélise la durée de vie d'un processeur (en années) par une loi exponentielle de paramètre 1/2.

1. Que vaut la durée de vie moyenne d'un tel processeur ?
2. Avec quelle probabilité le processeur fonctionne-t-il plus de six mois ?
3. Chaque vente de processeur rapporte 100 euros à son fabriquant, sauf s'il doit être échangé pendant les six mois de garantie, auquel cas il ne rapporte plus que 30 euros. Combien rapporte en moyenne un processeur ?

II. Densité quadratique
On considère une variable aléatoire X de densité

$$f(x) = \begin{cases} c\,x^2 & 0 \leq x \leq 3 \\ 0 & \text{ailleurs} \end{cases}$$

1. Evaluer la constante c pour que f soit une densité de probabilité. Donner l'allure de f.
2. Déterminer la fonction de répartition F de X. Donner son allure.
3. Calculer $\mathbb{P}(1 < X < 2)$.
4. Déterminer espérance et variance de X.
5. Pour tout $n \in \mathbb{N}^*$, déterminer le moment d'ordre n de X.

III. Accidents et fréquence cardiaque

1. On considère que, pour un conducteur, le nombre de kilomètres avant le premier accident suit une loi normale d'espérance 35000 km avec un écart-type de 5000 km. Pour un conducteur choisi au hasard, déterminer la probabilité :
   (a) qu'il ait eu son premier accident avant d'avoir parcouru 25000 km.
   (b) qu'il ait eu son premier accident après avoir parcouru 25000 km et avant 40000 km.
   (c) qu'il n'ait pas eu d'accident avant d'avoir parcouru 45000 km.
   (d) Au bout de combien de kilomètres peut-on dire que 80% des conducteurs ont eu leur premier accident ?

2. La fréquence cardiaque chez un adulte en bonne santé est en moyenne de 70 pulsations par minute, avec un écart-type de 10 pulsations. Soit X la variable aléatoire représentant la fréquence cardiaque chez un adulte.
   (a) A l'aide de l'inégalité de Tchebychev, minorer $\mathbb{P}(50 < X < 90)$.
   (b) Si on suppose maintenant que X suit une loi normale, que vaut $\mathbb{P}(50 < X < 90)$?

IV. Dé coloré

Un joueur dispose d'un dé équilibré à six faces avec trois faces blanches, deux vertes et une rouge. Le joueur lance le dé et observe la couleur de la face supérieure :
– s'il observe une face rouge, il gagne 2 euros ;
– s'il observe une face verte, il perd 1 euro ;
– s'il observe une face blanche, il relance le dé et : pour une face rouge, il gagne 3 euros ; pour une face verte, il perd 1 euro ; pour une face blanche, le jeu est arrêté sans gain ni perte.
Soit X la variable aléatoire égale au gain (positif ou négatif) de ce joueur.

1. Quelles sont les valeurs prises par X ? Déterminer la loi de X.
2. Calculer l'espérance de X.
3. Calculer la variance et l'écart-type de X.
4. Le joueur effectue 144 parties successives de ce jeu. Donner une valeur approchée de la probabilité que son gain sur les 144 parties soit positif.

V. Beaujolais nouveau

Le beaujolais nouveau est arrivé.

1. Un amateur éclairé, mais excessif, se déplace de réverbère en réverbère. Quand il se lance pour attraper le suivant, il a 80% de chances de ne pas tomber. Pour gagner le bistrot convoité, il faut en franchir 7. On notera X le nombre de réverbères atteints sans chute.
   (a) Quelles valeurs peut prendre la variable aléatoire X ?
   (b) Préciser sa loi.
2. Quand il sort du café, son étape suivante est l'arrêt de bus. Le nombre de chutes pour y parvenir, noté Y, suit une loi de Poisson $\mathcal{P}(4)$. Calculer la probabilité de faire au plus deux chutes.
3. Arrivé dans l'ascenseur, il appuie au hasard sur un des huits boutons. S'il atteint son étage ou s'il déclenche l'alarme, il sort de l'ascenceur, sinon il réappuie au hasard sur un des huits boutons. Soit Z le nombre de boutons pressés avant d'atteindre son étage ou de déclencher l'alarme.
   (a) Quelle est la loi de Z ?
   (b) Donner son espérance et sa variance.

VI. Loi de Gumbel

1. On considère la fonction g définie pour tout réel x par $g(x) = e^{-e^{-x}}$. Calculer ses limites en $-\infty$ et $+\infty$, sa dérivée, et donner l'allure de g.
2. Vérifier que la fonction f définie pour tout réel x par $f(x) = e^{-x-e^{-x}}$ est une densité.
3. Soit X une variable aléatoire de loi exponentielle de paramètre 1. Rappeler ce que vaut la fonction de répartition F de X. Donner son allure.

4. Soit X_1 et X_2 des variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre 1, et soit $M = \max(X_1, X_2)$ la variable aléatoire correspondant au maximum de ces deux variables. Pour tout réel x, calculer $\mathbb{P}(M \leq x)$. En déduire la densité de M.
5. On note maintenant $M_n = \max(X_1, \ldots, X_n)$, où $X_1, \ldots, X_n$ sont variables aléatoires indépendantes et identiquement distribuées de loi exponentielle de paramètre 1. Pour tout réel x, calculer $F_n(x) = \mathbb{P}(M_n \leq x)$.
6. Soit u un réel fixé, que vaut $\lim_{n\to+\infty}(1 - \frac{u}{n})^n$? En déduire que pour tout réel x

$$\lim_{n\to+\infty} F_n(x + \ln n) = g(x).$$

Université Rennes 2
Licence MASS 2

Lundi 12 Décembre 2011
Durée : 2 heures
Calculatrice autorisée
Aucun document

Corrigé du Contrôle

I. Durée de vie d'un processeur
On modélise la durée de vie d'un processeur (en années) par une loi exponentielle de paramètre 1/2.

1. Notons T la variable aléatoire modélisant cette durée de vie. Ainsi $T \sim \mathcal{E}(1/2)$, d'où $\mathbb{E}[T] = 2$ ans.
2. Pour tout $t \geq 0$, la fonction de répartition est $F(t) = \mathbb{P}(T \leq t) = 1 - e^{-t/2}$, d'où sa fonction de survie $\mathbb{P}(T > t) = 1 - F(t) = e^{-t/2}$. Ainsi la probabilité que le processeur fonctionne plus de six mois, i.e. une demi-année :
$$\mathbb{P}(T > 1/2) = e^{-1/4} \approx 0.78$$
3. Notons G la variable correspondant à ce que rapporte un processeur. Elle prend donc 2 valeurs, 100 et 30, avec les probabilités 0.78 et 0.22, d'où en moyenne :
$$\mathbb{E}[G] = 100 \times 0.78 + 30 \times 0.22 = 84.6 \text{ €}.$$

II. Densité quadratique
On considère une variable aléatoire X de densité
$$f(x) = \begin{cases} c\,x^2 & 0 \leq x \leq 3 \\ 0 & \text{ailleurs} \end{cases}$$

1. Pour que f soit une densité de probabilité, il faut qu'elle soit positive et intègre à 1 :
$$1 = \int_0^3 f(x)dx = c\int_0^3 x^2dx = c\left[\frac{x^3}{3}\right]_0^3 = 9c \;\Rightarrow\; c = \frac{1}{9}.$$
La fonction $f : x \mapsto \frac{x^2}{9}\mathbb{1}_{\{0\leq x\leq 3\}}$ est représentée sur la figure A.5 à gauche.
2. Puisque X ne tombe qu'entre 0 et 3, sa fonction de répartition F est nulle à gauche de 0 et vaut 1 à droite de 3. Pour $0 \leq x \leq 3$, il vient
$$F(x) = \int_0^x f(u)du = \frac{1}{9}\int_0^x u^2du = \frac{1}{9}\left[\frac{u^3}{3}\right]_0^x = \frac{x^3}{27}.$$
La fonction F est représentée sur la figure A.5 à droite.

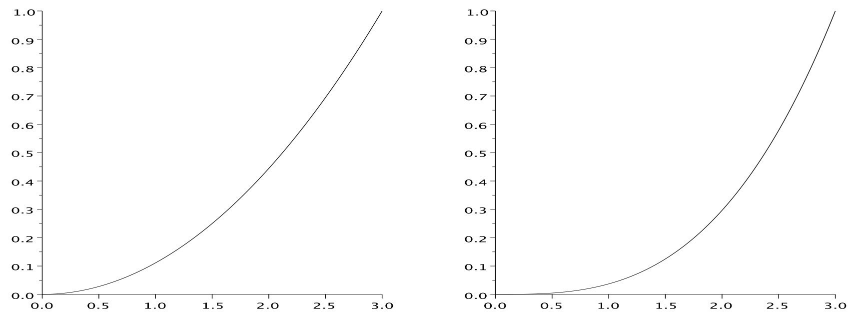


FIGURE A.5 – Densité f et fonction de répartition F.

3. Il va de soi que la quantité cherchée se déduit de la fonction de répartition :

$$\mathbb{P}(1 < X < 2) = F(2) - F(1) = \frac{7}{27}.$$

4. L'espérance de X vaut

$$\mathbb{E}[X] = \int_0^3 xf(x)dx = \frac{1}{9}\int_0^3 x^3 dx = \frac{1}{9}\left[\frac{x^4}{4}\right]_0^3 = \frac{9}{4}$$

La variance se déduit alors du moment d'ordre 2, lequel se calcule de façon comparable :

$$\mathbb{E}[X^2] = \int_0^3 x^2 f(x)dx = \frac{1}{9}\int_0^3 x^4 dx = \frac{1}{9}\left[\frac{x^5}{5}\right]_0^3 = \frac{27}{5}$$

Par conséquent : $\text{Var}(X) = \frac{27}{5} - \left(\frac{9}{4}\right)^2 = \frac{27}{80}$.

5. Pour tout $n \in \mathbb{N}^*$, nous avons

$$\mathbb{E}[X^n] = \int_0^3 x^n f(x)dx = \frac{1}{9}\int_0^3 x^{n+2} dx = \frac{1}{9}\left[\frac{x^{n+3}}{n+3}\right]_0^3 = \frac{3^{n+1}}{n+3}.$$

III. Accidents et fréquence cardiaque

1. Si on note Y la variable aléatoire correspondant au kilométrage lors du premier accident, l'énoncé implique que la variable X définie par $X = \frac{Y-35000}{5000}$ suit une loi normale centrée réduite, dont la fonction de répartition est notée Φ conformément à l'usage.
   (a) La probabilité cherchée s'écrit :

   $$p_1 = \mathbb{P}(Y \leq 25000) = \mathbb{P}\left(\frac{Y-35000}{5000} \leq \frac{25000-35000}{5000}\right) = \mathbb{P}(X \leq -2) = \Phi(-2)$$

   ce qui donne $p_1 = 1 - \Phi(2) \approx 0.0228$. Il y a donc environ 2.3% de chances qu'un conducteur ait son premier accident avant d'avoir parcouru 25000 km.
   (b) La probabilité cherchée s'écrit cette fois :

   $$p_2 = \mathbb{P}(25000 \leq Y \leq 40000) = \mathbb{P}\left(\frac{25000-35000}{5000} \leq \frac{Y-35000}{5000} \leq \frac{40000-35000}{5000}\right)$$

   soit

   $$p_2 = \mathbb{P}(-2 \leq X \leq 1) = \Phi(1) - \Phi(-2) \approx 0.8413 - 0.0228 = 0.8185$$

(c) La variable Y étant centrée en 35000, qui est le milieu du segment $[25000, 45000]$, il est clair que $p_3 = p_1$ (au besoin, faire un dessin pour s'en convaincre). Si on aime faire des calculs inutiles, ceci s'écrit :

$$p_3 = \mathbb{P}(Y \leq 45000) = \mathbb{P}\left(\frac{Y - 35000}{5000} \leq \frac{45000 - 35000}{5000}\right) = \mathbb{P}(X \leq 2) = \Phi(2) \approx 0.0228$$

(d) On cherche le quantile q tel que $\mathbb{P}(Y \leq q) = 0.8$, or

$$\mathbb{P}(Y \leq q) = 0.8 \;\Leftrightarrow\; \mathbb{P}\left(X \leq \frac{q - 35000}{5000}\right) = 0.8 \;\Leftrightarrow\; \Phi\left(\frac{q - 35000}{5000}\right) = 0.8$$

ce qui donne

$$\frac{q - 35000}{5000} \approx 0.84 \;\Leftrightarrow\; q \approx 39200 \text{ km}$$

2. La fréquence cardiaque chez un adulte en bonne santé est en moyenne de 70 pulsations par minute, avec un écart-type de 10 pulsations. Soit X la variable aléatoire représentant la fréquence cardiaque chez un adulte.

(a) A l'aide de l'inégalité de Tchebychev, nous obtenons

$$\mathbb{P}(50 < X < 90) = \mathbb{P}(|X - \mathbb{E}[X]| < 20) = 1 - \mathbb{P}(|X - \mathbb{E}[X]| > 20)$$

et puisque $\mathrm{Var}(X) = 100$, on obtient au final :

$$\mathbb{P}(50 < X < 90) \geq 1 - \frac{\mathrm{Var}(X)}{20^2} = \frac{3}{4}$$

(b) Si on suppose que X suit en fait une loi normale, ce qui est tout à fait raisonnable, on se rend compte que la probabilité est en fait bien plus grande :

$$\mathbb{P}(50 < X < 90) = \mathbb{P}\left(-2 < \frac{X - 20}{10} < 2\right) = \Phi(2) - \Phi(-2) = 2 \times \Phi(2) - 1 \approx 0.9544$$

IV. Dé coloré

Un joueur dispose d'un dé équilibré à six faces avec trois faces blanches, deux vertes et une rouge. Le joueur lance le dé et observe la couleur de la face supérieure :
– s'il observe une face rouge, il gagne 2 euros ;
– s'il observe une face verte, il perd 1 euro ;
– s'il observe une face blanche, il relance le dé et : pour une face rouge, il gagne 3 euros ; pour une face verte, il perd 1 euro ; pour une face blanche, le jeu est arrêté sans gain ni perte.

Soit X la variable aléatoire égale au gain (positif ou négatif) de ce joueur.

1. Les valeurs prises par X sont $\{-1, 0, 2, 3\}$. Avec des notations allant de soi, la loi de X est alors donnée par :
   – $\mathbb{P}(X = -1) = \mathbb{P}(V_1 \cup (B_1 \cap V_2)) = \mathbb{P}(V_1) + \mathbb{P}(B_1)\mathbb{P}(V_2) = \frac{2}{6} + \frac{3}{6} \times \frac{2}{6} = \frac{1}{2}$;
   – $\mathbb{P}(X = 0) = \mathbb{P}(B_1 \cap B_2) = \mathbb{P}(B_1)\mathbb{P}(B_2) = \frac{3}{6} \times \frac{3}{6} = \frac{1}{4}$;
   – $\mathbb{P}(X = 2) = \mathbb{P}(R_1) = \frac{1}{6}$;
   – $\mathbb{P}(X = 3) = \mathbb{P}(B_1 \cap R_2) = \mathbb{P}(B_1) \times \mathbb{P}(R_2) = \frac{3}{6} \times \frac{1}{6} = \frac{1}{12}$.

2. L'espérance de X est donc :

$$\mathbb{E}[X] = -1 \times \frac{1}{2} + 0 \times \frac{1}{4} + 2 \times \frac{1}{6} + 3 \times \frac{1}{12} = \frac{1}{12}.$$

3. Pour la variance de X, commençons par calculer son moment d'ordre 2 :

$$\mathbb{E}[X^2] = (-1)^2 \times \frac{1}{2} + 0^2 \times \frac{1}{4} + 2^2 \times \frac{1}{6} + 3^2 \times \frac{1}{12} = \frac{23}{12},$$

d'où

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{23}{12} - \left(\frac{1}{12}\right)^2 = \frac{275}{144} \approx 1.91$$

et l'écart-type vaut : $\sigma(X) = \sqrt{\frac{275}{144}} \approx 1.38$.

4. Notons S la variable correspondant au gain du joueur sur 144 parties successives de ce jeu. Il est clair que $S = X_1 + \cdots + X_{144}$, où X_i représente le gain à la partie i. Puisque les 144 variables X_i sont indépendantes et identiquement distribuées, le Théorème Central Limite s'applique, à savoir que S suit approximativement une loi normale $\mathcal{N}(144 \times \mathbb{E}[X], 144 \times \mathrm{Var}(X)) = \mathcal{N}(12, 275)$. Dès lors, une valeur approchée de la probabilité que le gain sur les 144 parties soit positif est :

$$\mathbb{P}(S > 0) = \mathbb{P}\left(\frac{S-12}{\sqrt{275}} > \frac{-12}{\sqrt{275}}\right) \approx 1 - \Phi\left(\frac{-12}{\sqrt{275}}\right) = \Phi\left(\frac{12}{\sqrt{275}}\right) \approx 0.7642.$$

Le joueur a donc environ 76% de chances d'avoir un gain global positif sur 144 parties successives.

V. Beaujolais nouveau

Le beaujolais nouveau est arrivé.

1. Un amateur éclairé, mais excessif, se déplace de réverbère en réverbère. Quand il se lance pour attraper le suivant, il a 80% de chances de ne pas tomber. Pour gagner le bistrot convoité, il faut en franchir 7. On notera X le nombre de réverbères atteints sans chute.
   (a) La variable aléatoire X peut prendre les valeurs $\{0, 1, \ldots, 6, 7\}$.
   (b) La loi de X est en gros celle d'une variable géométrique commençant à 0 et "tronquée" à droite puisque

   $$\forall k \in \{0, 1, \ldots, 6\} \qquad \mathbb{P}(X = k) = 0.2 \times 0.8^k$$

   et $\mathbb{P}(X = 7) = 0.8^7$.
2. La variable Y suivant une loi de Poisson $\mathcal{P}(4)$, la probabilité de faire au plus deux chutes est tout simplement

$$\mathbb{P}(Y \leq 2) = \mathbb{P}(Y = 0) + \mathbb{P}(Y = 1) + \mathbb{P}(Y = 2) = e^{-4}\left(\frac{4^0}{0!} + \frac{4^1}{1!} + \frac{4^2}{2!}\right) = 13e^{-4} \approx 0.238$$

3. Avant toute chose, notons qu'il est clair d'après l'énoncé que l'alarme fait partie des 8 boutons.
   (a) Sur les 8 boutons, 2 conduisent à l'arrêt du "jeu" (le bon étage ou l'alarme), pour les 6 autres on rejoue. En ce sens, la variable Z suit une loi géométrique de paramètre $1/4$.
   (b) On en déduit que $\mathbb{E}[Z] = 4$ et $\mathrm{Var}(Z) = 12$.

VI. Loi de Gumbel

1. On considère la fonction g définie pour tout réel x par $g(x) = e^{-e^{-x}}$. Puisque $\lim_{x \to -\infty} e^{-x} = +\infty$ et $\lim_{x \to +\infty} e^{-x} = 0$, il s'ensuit que $\lim_{x \to -\infty} g(x) = 0$ et $\lim_{x \to +\infty} g(x) = 1$. Par ailleurs, la fonction g est dérivable en tant que composée de fonctions dérivables et sa dérivée vaut $g'(x) = e^{-x-e^{-x}}$. L'allure de g est donnée en figure A.6 à droite.

2. Puisque l'exponentielle est partout positive, il en va de même pour f. Il reste à vérifier que son intégrale somme à 1, or il suffit pour cela de remarquer que f n'est rien d'autre que la dérivée de g :

$$\int_{-\infty}^{+\infty} f(x)dx = [g(x)]_{-\infty}^{+\infty} = \lim_{x\to+\infty} g(x) - \lim_{x\to-\infty} g(x) = 1 - 0 = 1.$$

L'allure de f est donnée en figure A.6 à gauche. Lorsqu'une variable X a pour densité f, on dit qu'elle suit une loi de Gumbel.

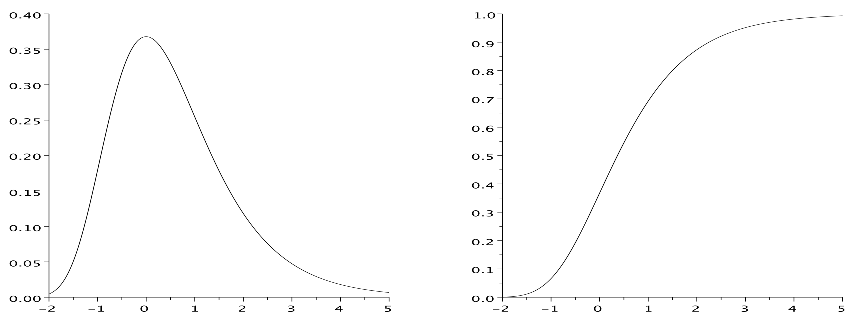


Figure A.6 – Fonctions f et g, densité et fonction de répartition d'une loi de Gumbel.

3. Si X suit une loi exponentielle de paramètre 1, sa fonction de répartition F vaut 0 pour $x \leq 0$ et $F(x) = 1 - e^{-x}$ pour $x \geq 0$.
4. Les variables X_1 et X_2 ne prenant que des valeurs postives, c'est a fortiori le cas pour la variable M, donc $\mathbb{P}(M \leq x) = 0$ si $x \leq 0$. Pour $x \geq 0$, nous avons par indépendance de X_1 et X_2 :

$$\mathbb{P}(M \leq x) = \mathbb{P}(\max(X_1, X_2) \leq x) = \mathbb{P}(\{X_1 \leq x\} \cap \{X_2 \leq x\}) = \mathbb{P}(X_1 \leq x)\mathbb{P}(X_2 \leq x)$$

et via la question précédente

$$\mathbb{P}(M \leq x) = \left(1 - e^{-x}\right)\left(1 - e^{-x}\right) = \left(1 - e^{-x}\right)^2.$$

Nous avons donc calculé la fonction de répartition F_2 de la variable M. Sa dérivée f_2 est la densité de M. Celle-ci vaut bien entendu 0 pour $x \leq 0$, tandis que pour $x \geq 0$

$$f_2(x) = F_2'(x) = 2e^{-x}\left(1 - e^{-x}\right).$$

5. Mutatis mutandis, les arguments précédents s'appliquent ici et aboutissent à

$$F_n(x) = \mathbb{P}(M_n \leq x) = \left(1 - e^{-x}\right)^n \mathbb{1}_{\{x\geq 0\}}.$$

6. u étant un réel fixé, il est clair que pour n suffisamment grand, nous avons $|u/n| < 1$, de sorte que nous pouvons sans vergogne passer à la forme exponentielle-logarithmique de la quantité en question et utiliser le développement limité $\ln(1 - x) = -x + o(x)$:

$$\left(1 - \frac{u}{n}\right)^n = e^{n\ln\left(1-\frac{u}{n}\right)} = e^{n\left(-\frac{u}{n}+o\left(\frac{1}{n}\right)\right)} = e^{-u+o(1)} \xrightarrow[n\to+\infty]{} e^{-u}$$

Pour tout réel x, pour n suffisamment grand, nous avons $x + \ln n > 0$ et la formule obtenue pour F_n donne donc :

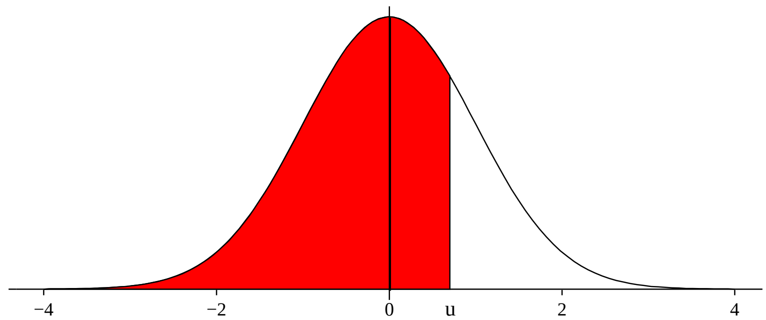
$$F_n(x + \ln n) = \left(1 - e^{-(x+\ln n)}\right)^n = \left(1 - e^{-x}e^{-\ln n}\right)^n = \left(1 - \frac{e^{-x}}{n}\right)^n$$

à la suite de quoi nous pouvons appliquer le résultat précédent avec $u = e^{-x}$ pour obtenir

$$\lim_{n \to +\infty} F_n(x + \ln n) = e^{-e^{-x}} = g(x).$$

Dans le jargon, on dit que la suite de variables aléatoires $(M_n - \ln n)_{n \geq 0}$ converge en loi vers une variable aléatoire qui suit une loi de Gumbel. Dit autrement, le maximum d'un grand nombre de variables i.i.d. exponentielles tend vers l'infini à vitesse $\ln n$, et après translation de ce maximum par $-\ln n$, l'aléa qui reste suit en gros une loi de Gumbel. C'est pourquoi on dit que la loi de Gumbel est une des lois des extrêmes. En hydrologie, par exemple, elle peut servir à modéliser les crues d'un fleuve.

A.2 Table de la loi normale $X \sim \mathcal{N}(0,1)$



Valeurs de $\Pr(X \leq u)$ en fonction de u.

u	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995

Bibliographie

[1] Nicolas Bouleau. *Probabilités de l'ingénieur*. Hermann, 2002.

[2] Pierre Brémaud. *Introduction aux probabilités : modélisation des phénomènes aléatoires*. Springer, 1984.

[3] Rick Durrett. *Elementary Probability for Applications*. Cambridge University Press, 2009.

[4] Dominique Foata et Aimé Fuchs. *Calcul des probabilités*. Dunod, 1998.

[5] Alain Combrouze et Alexandre Dédé. *Probabilités et Statistiques / 1*. PUF, 1996.

[6] Alain Combrouze et Alexandre Dédé. *Probabilités et Statistiques / 2*. PUF, 1998.

[7] Jean Guégand, Jean-Louis Roque et Christian Lebœuf. *Cours de probabilités et de statistiques*. Ellipses, 1998.

[8] Gilles Pagès et Claude Bouzitat. *En passant par hasard... Les probabilités de tous les jours*. Vuibert, 2000.

[9] Eva Cantoni, Philippe Huber et Elvezio Ronchetti. *Maîtriser l'aléatoire (Exercices résolus de probabilités et statistique)*. Springer, 2006.

[10] François Husson et Jérôme Pagès. *Statistiques générales pour utilisateurs (2. Exercices et corrigés)*. Presses Universitaires de Rennes, 2005.

[11] Philippe Barbe et Michel Ledoux. *Probabilités*. Belin, 1998.

[12] Guy Auliac, Christiane Cocozza-Thivent, Sophie Mercier et Michel Roussignol. *Exercices de probabilités*. Cassini, 1999.

[13] Valérie Girardin et Nikolaos Limnios. *Probabilités*. Vuibert, 2001.

[14] Jean Jacod et Philip Protter. *L'essentiel en théorie des probabilités*. Cassini, 2003.

[15] Marie Cottrell, Valentine Genon-Catalot, Christian Duhamel et Thierry Meyre. *Exercices de probabilités*. Cassini, 1999.

[16] Gérard Frugier. *Exercices ordinaires de probabilités*. Ellipses, 1992.

[17] Geoffrey R. Grimmett and David R. Stirzaker. *One Thousand Exercises in Probability*. Oxford University Press, New York, 2001.

[18] Geoffrey R. Grimmett and David R. Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, 2001.

[19] Jacques Harthong. *Calcul des probabilités*. Format électronique, 2001.

[20] Michel Métivier. *Probabilités : dix leçons d'introduction*. Ellipses, 1987.

[21] Jean-Yves Ouvrard. *Probabilités 1*. Cassini, 1998.

[22] Jim Pitman. *Probability*. Springer, 1999.

[23] Sheldon M. Ross. *Initiation aux probabilités*. Presses polytechniques et universitaires romandes, 1987.

[24] Charles Suquet. *Introduction au Calcul des Probabilités*. Format électronique, 2005.