

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEURE ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE 20 AOUT 1955 SKIKDA

FACULTE DE SCIENCES
DEPARTEMENT DE MATHEMATIQUES

Probabilités

Cours et exercices

Karima Kimouche

Année universitaire : 2019/2020

Avant-propos

Ce polycopié est le support du cours de « Probabilités » destiné principalement aux étudiants de la licence (2^{ème} année) de mathématiques de l'université 20 août 1955 Skikda. Cette première version du polycopié a été élaborée avec beaucoup de détails et clarté. Il est constitué de trois chapitres qui vont permettre aux étudiants de mieux comprendre, d'assimiler, d'approfondir, et de visualiser ou mettre en lumière la raison principale de ce manuscrit. A la fin de chaque chapitre on pourra trouver une série d'exercices.

Le contenu de ce polycopié est structuré en trois chapitres :

Le premier chapitre est consacré aux notions de base sur les probabilités, on y définira l'expérience aléatoire et événements, la probabilité sur un espace fondamental fini : la probabilité conditionnelle, la formule des probabilités totales.

Le deuxième chapitre concerne les variables aléatoires. Nous avons donné les définitions et les principales propriétés de la loi de probabilité et la fonction de répartition dans les deux cas discret et continu, nous avons traité ensuite l'espérance et la variance, enfin nous avons étudié les inégalités en probabilités.

Compte tenu l'importance des lois de probabilités le troisième chapitre a illustré les définitions et les principales propriétés de ses lois usuelles : Lois de probabilités discrètes et loi de probabilités absolument continues. La convergence en loi et l'approximation entre les principales lois usuelles ont été considéré.

Pour élaborer ce support, je me suis appuyé sur différentes références, des ouvrages reconnus dans la discipline. Il est certain que la première version de cet ouvrage n'est pas perfectible, et qu'elle contient certaines erreurs. Alors, j'invite tous les lecteurs, étudiants ou enseignants à me faire parvenir leurs remarques et commentaires pour que je puisse modifier et corriger ce polycopié.

Table des matières

1	Rappels sur les probabilités	3
1.1	Introduction	3
1.2	Expérience aléatoire et événements	3
1.2.1	Langage et notation ensembliste	4
1.3	Probabilité sur un espace fondamental fini	4
1.3.1	Construction d'une probabilité	5
1.3.2	Probabilité uniforme (ou équiprobabilité)	6
1.3.3	Probabilité conditionnelle	7
1.4	Généralisation des probabilités composés	8
1.4.1	Événements indépendants	9
1.4.2	Formule des probabilités totales	10
1.5	Exercices	11
2	Variables aléatoires	13
2.1	Définition	13
2.2	Loi de probabilité, Fonction de répartition	14
2.2.1	Loi d'une variable discrète	14
2.2.2	Loi d'une variable aléatoire à densité (ou continues)	15
2.3	Espérance et variance d'une variable aléatoire	16
2.3.1	Mesurer la diffusion d'une variable : variance et écart type	18
2.3.2	Corrélation de deux variables aléatoires	18
2.4	Inégalités en probabilités	19
2.4.1	Inégalité de Markov et de Bienaymé-Tchebychev	19
2.4.2	Inégalité de Cauchy-Schwarz et Minkowski	21
2.4.3	Inégalité de Jensen	22

2.5	Exercices	22
3	Lois de probabilités	25
3.1	Lois de probabilités discrètes usuelles	25
3.2	Lois de probabilités absolument continues usuelles	34
3.2.1	La convergence en loi	52
3.2.2	Approximation d'une loi hypergéométrique par une loi binomiale	52
3.2.3	Approximation d'une loi binomiale par une loi de Poisson	53
3.2.4	Approximation d'une loi binomiale par une loi normale	55
3.2.5	Approximation d'une loi de Poisson par une loi normale	55
3.3	Exercices	57

Chapitre 1

Rappels sur les probabilités

1.1 Introduction

Des actions comme lancer un dé, tirer une carte d'un jeu, observer la durée de vie d'une ampoule électrique, etc sont des expériences aléatoires. Leur résultat n'est connu que lorsque l'expérience aléatoire a pris fin. On peut cependant envisager a priori des éventualités (des résultats possibles) mais seule la fin de l'expérience nous permet de savoir si telle éventualité est réalisée ou non. Un événement est un ensemble d'éventualités qui est réalisé si l'une de ces éventualités est réalisée. Dans une expérience aléatoire certains événements sont plus probables que d'autres ; d'où l'idée de chiffrer les chances qu'a un événement de se produire.

1.2 Expérience aléatoire et événements

Définition 1 *Une expérience ayant un nombre fini d'issues possibles est appelé expérience aléatoire (ou épreuve) s'il est impossible de savoir à l'avance quelle en sera le résultat.*

L'ensemble de tous les résultats possibles est appelé univers ou "ensemble des éventualités" ou "ensemble fondamental"; généralement noté Ω .

Chaque sous ensemble de Ω contenant un seul élément, c'est à dire chaque issue possible est appelé événement élémentaire. Chaque partie de Ω est un événement.

Ω est l'événement certain. \emptyset est l'événement impossible.

Remarque 2 *Si on note chaque issues possibles e_1, \dots, e_n alors $\Omega = \{e_1, \dots, e_n\}$ et chaque $\{e_i\}$ est alors un événement élémentaire. Une expérience aléatoire est déterminée par l'expérience que l'on*

effectue et donc l'univers aussi, c'est à dire que si on change d'expérience aléatoire, on change aussi d'univers!

Exemple 3 Lancer un dé peut se représenter par l'ensemble $\Omega = \{1, 2, 3, 4, 5, 6\}$.

"Obtenir un nombre pair" est représenté par l'événement $A = \{2, 4, 6\}$,

"Obtenir le nombre 3" est représenté par l'événement $B = \{3\}$,

"Obtenir un dix" est l'événement impossible,

"Obtenir un nombre entier compris entre 1 et 6" est un événement certain.

1.2.1 Langage et notation ensembliste

écriture ensembliste	formulation probabiliste
Ω	univers, "événement certains"
$\{\omega\} \subset \Omega$	événement élémentaire
$A \subset \Omega$	événement A
$A \subset B$	$A \implies B$
$\bar{A} = C_{\Omega}^A$	événement contraire
$A \cap B$	est l'événement qui se produit si A et B sont réalisées simultanément
$A \cup B$	A ou B réalisé
$A \cap B = \emptyset$	A et B sont incompatibles
$(\Omega = A \cup B)$ et $(A \cap B = \emptyset)$	(A ou B est certain) et (A et B sont incompatibles) A et B forment un système complet

1.3 Probabilité sur un espace fondamental fini

Définition 4 L'univers fini est probabilisé si à chaque événement $A \in \mathcal{P}(\Omega)$ est associé un nombre $P(A) \geq 0$, appelé probabilité de A de telle sorte que pour $A, B \in \mathcal{P}(\Omega)$ les deux propriétés suivantes soient satisfaites :

$$P(\Omega) = 1$$

$$A \cap B = \emptyset \implies P(A \cup B) = P(A) + P(B)$$

Le couple (Ω, P) où l'application P définie sur $\mathcal{P}(\Omega)$ vérifie les conditions ci-dessus est appelé espace probabilisé. Le couple $(\Omega, \mathcal{P}(\Omega))$ est appelé espace probabilisable.

Les propriétés immédiates de la probabilité sont les suivantes :

Proposition 5 Si (Ω, P) est un espace probabilisé fini, pour tous $A, B \in \mathcal{P}(\Omega)$, on a

$$0 \leq P(A) \leq 1$$

$$P(\overline{A}) = 1 - P(A)$$

$$P(\emptyset) = 0$$

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$. De plus si A_1, \dots, A_n sont des événements deux à deux incompatibles

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

$$A \subset B \implies P(A) \leq P(B)$$

1.3.1 Construction d'une probabilité

Proposition 6 Soit $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ un espace fondamental fini. La probabilité P définie sur $\mathcal{P}(\Omega)$ est déterminé par la probabilité des événements élémentaires $P(\{\omega_i\}) = p_i, i = \overline{1, n}$ tel que : $0 \leq p_i \leq 1$ et $\sum_{i=1}^n p_i = 1$.

Exemple 7 Soit $\Omega = \{a, b, c\}$ et P probabilité sur $\mathcal{P}(\Omega)$ tel que :

$$P(\{b\}) = 2P(\{a\}), P(\{c\}) = 2P(\{b\})$$

Calculer $P(\{a\}), P(\{b\})$ et $P(\{c\})$.

Ecrire tout les événements de l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ et calculer la probabilité de chaque événement.

On sait que :

$$\begin{aligned} P(\Omega) &= 1 \Leftrightarrow P(\{a\} \cup \{b\} \cup \{c\}) = 1 \\ &\Leftrightarrow P(\{a\}) + P(\{b\}) + P(\{c\}) = 1 \end{aligned}$$

On pose $P(\{a\}) = p_1, P(\{b\}) = p_2$ et $P(\{c\}) = p_3$.

alors

$$\begin{cases} p_1 + p_2 + p_3 = 1 \\ p_2 = 2p_1 \\ p_3 = 2p_2 = 4p_1 \end{cases} \Leftrightarrow \begin{cases} p_1 = \frac{1}{7} \\ p_2 = \frac{2}{7} \\ p_3 = \frac{4}{7} \end{cases}$$

$$\mathcal{P}(\Omega) = \{\emptyset, \{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\}, \Omega\};$$

$$P(\emptyset) = 0, P(\{a\}) = \frac{1}{7}, P(\{b\}) = \frac{2}{7}, P(\{c\}) = \frac{4}{7}, P(\Omega) = 1,$$

$$P(\{a, b\}) = P(\{a\}) + P(\{b\}) = \frac{3}{7}, P(\{a, c\}) = \frac{5}{7}, P(\{b, c\}) = \frac{6}{7}$$

1.3.2 Probabilité uniforme (ou équiprobabilité)

Définition 8 P est appelée probabilité uniforme si

$$\forall i = \overline{1, n}; P(\{\omega_i\}) = p_i = \frac{1}{\text{card}\Omega} = \frac{1}{n}$$

Si la probabilité est uniforme sur Ω , pour tout événement A , on a

$$P(A) = \frac{\text{card}A}{\text{card}\Omega},$$

En d'autres termes plutôt anciens

$$P(A) = \frac{\text{card}A}{\text{card}\Omega} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possible}}$$

Exemple 9 On lance un dé et soit l'événement A "avoir un nombre impair"

$$\Omega = \{1, 2, 3, 4, 5, 6\}, A = \{1, 3, 5\}, P(A) = \frac{\text{card}A}{\text{card}\Omega} = \frac{3}{6} = \frac{1}{2}.$$

Pour calculer en détails ces probabilités on utilise des résultats de combinatoire, dont on rappelle les plus classiques.

- Nombre de permutations d'un ensemble à n éléments : $n!$.
- Nombre de p -uples dans un ensemble à n éléments : n^p .
- Nombre de p -uples d'éléments distincts dans un ensemble à n éléments : $A_n^p = n(n-1)\dots(n-p+1)$.
- Nombre de parties d'un ensemble à n éléments : 2^n .
- Nombre de parties à p éléments dans un ensemble à n éléments : $C_n^p = \binom{n}{p}$.

1.3.3 Probabilité conditionnelle

Définition 10 Soient $(\Omega, \mathcal{P}(\Omega), P)$ espace probabilité et A un événement de $\mathcal{P}(\Omega)$ tel que $P(A) \neq 0$, l'application de $\mathcal{P}(\Omega)$ dans $[0, 1]$ définie par :

$$\forall B \in \mathcal{P}(\Omega), P(B/A) = \frac{P(B \cap A)}{P(A)}$$

est une probabilité sur $\mathcal{P}(\Omega)$ appelée probabilité conditionnelle B sachant A

On note $P(B/A)$ par $P_A(B)$

Propriétés

$$P(\Omega/A) = 1$$

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

$$P(A \cap B) = P(B/A) \cdot P(A) = P(A/B) \cdot P(B) \text{ (formule des probabilités composées)}$$

Exemple 11 Dans une ville on a 40% des citoyens ont les cheveux châtain et 25% les yeux marrons et 15% cheveux châtain avec yeux marrons. On a choisit un citoyen de aléatoire de cette ville.

Si ses cheveux sont châtain quelle est la probabilité que ses yeux soient marron ? (on va calculer $P(B/A)$)

On note : A "citoyen au cheveux châtain" et B "citoyen au yeux marrons"

$$P(A) = 0,4; P(B) = 0,25; P(A \cap B) = 0,15.$$

Alors

$$P(B/A) = \frac{P(B \cap A)}{P(A)} = \frac{0,15}{0,4} = 0,37$$

Si ses yeux sont marrons quelle est la probabilité que ses cheveux ne soient pas châtain ?

On va chercher

$$P(\bar{A}/B) = \frac{P(\bar{A} \cap B)}{P(B)}$$

$$\text{On a } B = (A \cap B) \cup (\bar{A} \cap B) \implies P(B) = P[(A \cap B) \cup (\bar{A} \cap B)]$$

$$\implies P(B) = P(A \cap B) + P(\bar{A} \cap B) \text{ (car } (A \cap B) \text{ et } (\bar{A} \cap B) \text{ sont incompatibles)}$$

$$\implies P(\bar{A} \cap B) = P(B) - P(A \cap B)$$

d'ou

$$P(\bar{A}/B) = \frac{P(B) - P(A \cap B)}{P(B)} = \frac{0,25 - 0,15}{0,25} = 0,4.$$

Quelle est la probabilité que ses cheveux ne soient pas châains et que ses yeux ne soient pas marrons ?

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A \cup B)$$

$$= 1 - [P(A) + P(B) - P(A \cap B)]$$

$$= 1 - [0,4 + 0,25 - 0,15]$$

$$= 0,5$$

1.4 Généralisation des probabilités composés

Soient A_1, \dots, A_n une famille de n événement tel que : $P(A_1 \cap A_2 \cap \dots \cap A_n) \neq 0$, alors

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2/A_1) \cdot P(A_3/A_1 \cap A_2) \dots P(A_n/A_1 \cap A_2 \cap \dots \cap A_{n-1})$$

Exemple 12 Soit 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 1^{ère} boule tirée soit blanche, la 2^{ème} blanche et la 3^{ème} noire ?

Soit : A_i "la $i^{\text{ème}}$ boule tirée est blanche", B_i "la $i^{\text{ème}}$ boule tirée est noire"

On cherche $P(A_1 \cap A_2 \cap B_3)$

$$\begin{aligned} P(A_1 \cap A_2 \cap B_3) &= P(A_1) \cdot P(A_2/A_1) \cdot P(B_3/A_1 \cap A_2) \\ &= \frac{4}{7} \cdot \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P(B_3 \cap (A_1 \cap A_2))}{P(A_1 \cap A_2)} \\ &= \frac{4}{7} \cdot \frac{3}{6} \cdot \frac{3}{5} = \frac{6}{35} \end{aligned}$$

1.4.1 Événements indépendants

Définition 13 On dit que A et B sont indépendants par rapport à la probabilité P si $P(A \cap B) = P(A) \cdot P(B)$.

Remarque 14 Si deux événement A et B sont indépendants alors $P(A \setminus B) = P(A)$ et $P(B \setminus A) = P(B)$.

$$\begin{aligned} A \text{ et } B \text{ indépendants} &\iff A \text{ et } \overline{B} \text{ sont indépendants} \\ &\iff \overline{A} \text{ et } B \text{ sont indépendants} \\ &\iff \overline{A} \text{ et } \overline{B} \text{ sont indépendants} \end{aligned}$$

Exemple 15 On lance une pièce de monnaie et un dé de faces numérotés de 1 à 6.

Calculer la probabilité d'avoir face et le numéro 2.

Soient les événements : A "avoir face", B "avoir le numéro 2"

$$P(A) = \frac{1}{2}, P(B) = \frac{1}{6}$$

$A \cap B$ "avoir face et le numéro 2", on cherche $P(A \cap B)$?

Il est clair que A et B sont indépendants alors $P(A \cap B) = P(A) \cdot P(B) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$

Vérifions que A et B sont indépendants. on détermine $\Omega = \left\{ \begin{array}{l} (F, 1), (F, 2), (F, 3), (F, 4), (F, 5), (F, 6), \\ (P, 1), (P, 2), (P, 3), (P, 4), (P, 5), (P, 6) \end{array} \right\}$

$$A \cap B = \{(F, 2)\} \implies P(A \cap B) = \frac{\text{nombre de cas favorable}}{\text{nombre de cas possible}} = \frac{1}{12} = P(A) \cdot P(B)$$

$\implies A$ et B sont indépendants.

Définition 16 Une famille $(A_n)_{n \in \mathbb{N}}$ d'événements est dite système complet d'événements si les A_n sont deux à deux incompatibles et leurs réunion est l'événement certain Ω .

Théorème 17 (De Bayes) Soient $(\Omega, \mathcal{P}(\Omega), P)$ espace probabilisé fini et $A, B_1, B_2, \dots, B_n, (n+1)$ événements tel que $P(A) > 0$ et $\{B_1, B_2, \dots, B_n\}$ forme un système complet d'événements. Alors on a pour tout $k = \overline{1, n}$

$$P(B_k \setminus A) = \frac{P(B_k) \cdot P(A \setminus B_k)}{\sum_{k=1}^n P(B_k) P(A \setminus B_k)}$$

1.4.2 Formule des probabilités totales

Soit (B_k) un système complet d'événements de probabilités non nulles pour tout événement A on a

$$P(A) = \sum_{k=1}^n P(A \cap B_k) = \sum_{k=1}^n P(A \setminus B_k) P(B_k).$$

Exemple 18 Un Pharmacien importe des flacons de médicaments de trois laboratoires différents : 50% du laboratoire A, 30% du laboratoire B, 20% du laboratoire C. 4% des flacons provenant de A sont avariés, 5% de ceux qui provient de B sont le sont aussi, ainsi que 7% de ceux qui provient de C.

Le pharmacien, choisit un flacon d'une manière aléatoire ;

Quelle est la probabilité qu'il provient de A et qu'il soit avarié.

Quelle est la probabilité qu'il provient de C quant en s'aperçoit qu'il avarié et qu'il soit avarié.

Solution : Soient les événements

A " le flacon provient du laboratoire A "

B " le flacon provient du laboratoire B "

C " le flacon provient du laboratoire C "

D " le flacon est avarié "

On a $P(A) = 0,5$; $P(B) = 0,3$; $P(C) = 0,2$; $P(D \setminus A) = 0,04$; $P(D \setminus B) = 0,05$; $P(D \setminus C) = 0,07$

On calcul la probabilité de réalisation simultanée des deux événements A et D c'est à dire $P(A \cap D)$.

$P(A \cap D) = P(D \setminus A) \cdot P(A) = 0,04 \times 0,5 = 0,02$.

$P(C \setminus D)$?

L'application du théorème de Bayes donne

$$\begin{aligned} P(C \setminus D) &= \frac{P(D \setminus C) \cdot P(C)}{P(D \setminus A) P(A) + P(D \setminus B) P(B) + P(D \setminus C) P(C)} \\ &= \frac{0,07 \times 0,2}{0,04 \times 0,5 + 0,05 \times 0,3 + 0,07 \times 0,2} \\ &\simeq 0,2857 \end{aligned}$$

1.5 Exercices

Exercice 19 On suppose l'ensemble fondamental $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$. Quelle est parmi ces fonction définie une probabilité :

1. (a) $P(\{\omega_1\}) = \frac{1}{2}, P(\{\omega_2\}) = \frac{1}{3}, P(\{\omega_3\}) = \frac{1}{4}, P(\{\omega_4\}) = \frac{1}{5}$.

(b) $P(\{\omega_1\}) = \frac{1}{2}, P(\{\omega_2\}) = \frac{1}{4}, P(\{\omega_3\}) = -\frac{1}{4}, P(\{\omega_4\}) = \frac{1}{2}$.

(c) $P(\{\omega_1\}) = \frac{1}{2}, P(\{\omega_2\}) = \frac{1}{4}, P(\{\omega_3\}) = \frac{1}{4}, P(\{\omega_4\}) = 0$.

2. Est-il possible de définir une probabilité P sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ telle que : $P(A) = \frac{1}{3}, P(B) = \frac{1}{3}, P(C) = \frac{2}{3}$; avec $\Omega = \{\omega_1, \omega_2, \omega_3\}, A = \{\omega_2, \omega_3\}, B = \{\omega_1, \omega_3\}, C = \{\omega_1, \omega_2\}$.

Exercice 20 Soient A et B deux événements tels que : $P(A) = \frac{3}{8}, P(B) = \frac{1}{2}, P(A \cap B) = \frac{1}{4}$.

- Calculer les probabilités suivantes :

$$P(A \cup B), P(\bar{A}), P(\bar{B}), P(\bar{B} \cap \bar{A}), P(\bar{B} \cup \bar{A}), P(A - B), P(B - A)$$

Exercice 21 1. Combien peut-on définir sur $\Omega = \{a, b, c\}$ de probabilités P telles que

a) $P(\{a, b\}) = \frac{1}{4}$,

b) $p(\{a, b\}) = P(\{b, c\}) = \frac{1}{4}$,

c) $p(\{a, b\}) = P(\{b, c\}) = \frac{3}{4}$.

2. Soit A et B deux événements aléatoires associés à un espace probabilisé (Ω, P) .

- Montrer que $P(A \cap B) - P(A)P(B) = P(B)P(A^c) - P(B \cap A^c)$.

Exercice 22 Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé

1. Soient A, B et C trois événements, déterminer $P(A \cup B \cup C)$.

2. Soient A et B deux événements indépendants, montrer que :

a) A et B sont aussi indépendants.

b) $P(B \setminus A) = P(B \setminus \bar{A})$.

3. Soient A et B deux événements. Si $P(A) = 0,5, P(B) = 0,3$ et $P(A \cup B) = 0,65$. Les événements A et B sont-ils indépendants en probabilité ?

Exercice 23 Dans la classe de 6^{ème}, 25% des élèves n'ont pas réussi l'épreuve de Mathématiques, 15% celle de chimie et 10% des élèves n'ont pas réussi en chimie ni en maths.

1. a) S'il n'a pas réussi chimie, quelle est la probabilité qu'il n'ait pas réussi en maths ?
- b) S'il n'a pas réussi les maths, quelle est la probabilité qu'il n'ait pas réussi en chimie ?
- c) Quelle est la probabilité qu'il ait au moins un échec dans les deux branches ?

Exercice 24 En europe occidentale, 5% des garçons et 0,25% des filles naissent daltoniens. 51% des naissances (bébés) concernent des garçons.

1. a) Quelle est la proportion d'enfants daltoniens dans le total ?
- b) Quelle est la probabilité de garçons dans la population des bébés daltoniens ?

Exercice 25 Dans une population Ω , deux maladies M_1 et M_2 sont présentes respectivement chez 10% et 20%. On suppose que le nombre de ceux qui souffrent des deux maladies est négligeable. On entreprend un dépistage systématique des maladies M_1 et M_2 . Pour cela, on applique un test qui réagit sur 90% des malades de M_1 , sur 70% des malades M_2 et sur 10% des individus qui n'ont aucune de ces deux affections.

1. Quand on choisit au hasard un individu ω dans Ω , quelle est la probabilité pour que le test réagisse ?
2. Sachant que pour un individu ω , le test a réagi, donner les probabilités :
 - Pour que le test ait réagi à cause de la maladie M_1 .
 - Pour que le test ait réagi à cause de la maladie M_2 .
 - Pour que le test ait réagi alors que l'individu n'est infecté par qu'aucune des deux maladies M_1 et M_2 .

Chapitre 2

Variables aléatoires

2.1 Définition

On considère un ensemble Ω muni d'une probabilité P .

Définition 26 Une variable aléatoire X est une fonction de l'ensemble fondamental Ω à valeurs dans \mathbb{R} , $X : \Omega \rightarrow \mathbb{R}$.

Lorsque la variable X ne prend que des valeurs discrètes, on parle de variable aléatoire discrète.

Exemple 27 On jette deux dés distincts et on s'intéresse à la somme des points. On note X cette variable aléatoire, elle est définie par

$$\begin{aligned} X & : \quad \Omega \rightarrow \mathbb{R} \text{ avec } \Omega = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\} \\ (\omega_1, \omega_2) & \mapsto \omega_1 + \omega_2 \end{aligned}$$

L'ensemble des valeurs possibles de X est $\{2, 3, \dots, 12\}$.

Exemple 28 On lance toujours deux dés, mais cette fois on s'intéresse au plus grand chiffre Y obtenu. On a alors

$$\begin{aligned} Y & : \quad \Omega \rightarrow \mathbb{R} \text{ avec } \Omega = \{(1, 1), (1, 2), \dots, (6, 5), (6, 6)\} \\ (\omega_1, \omega_2) & \mapsto \max(\omega_1, \omega_2). \end{aligned}$$

La variable Y est à valeurs dans $\{1, 2, \dots, 6\}$.

2.2 Loi de probabilité, Fonction de répartition

La loi de probabilité d'une variable aléatoire permet de connaître les chances d'apparition des différentes valeurs de cette variable.

On se place sur l'espace de probabilité (Ω, P) .

Définition 29 Soit X une variable aléatoire. La loi de probabilité de X est définie par la fonction F_X , appelée fonction de répartition de la variable X , définie par

$$\begin{aligned} F_X &: \mathbb{R} \rightarrow [0, 1] \\ x &\mapsto P(X \leq x). \end{aligned}$$

On dit que deux variables aléatoires X et Y ont la même loi si elles ont la même fonction de répartition $F_X = F_Y$.

Remarque 30 Soit I un intervalle de \mathbb{R} . L'événement $\{X \leq x\}$ représente l'ensemble des valeurs $\omega \in \Omega$ telles que $X(\omega)$ soit inférieur à x , i.e. $\{X \leq x\} = \{\omega \in \Omega : X(\omega) \leq x\}$. La loi de X est en générale notée $\mathcal{L}(X)$ ou $\mathcal{L}oi(X)$.

Remarque 31 On a $P(X \in \mathbb{R}) = 1$, car $P(X \in \mathbb{R}) = P(\{\omega \in \Omega : X(\omega) \in \mathbb{R}\}) = P(\Omega) = 1$.

Propriétés La fonction de répartition est une fonction croissante à valeur dans $[0, 1]$ telle que $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$, mais elle n'est pas forcément continue.

Remarque 32 Soit $a \leq b$, on a $P(X \in [a, b]) = P(X < b) - P(X < a)$.

$$P(a \leq X \leq b) = F_X(b) - F_X(a).$$

2.2.1 Loi d'une variable discrète

La fonction de répartition d'une variable discrète est constante par morceaux. Si X est une variable discrète à valeurs dans $\{x_1, \dots, x_n\}$ avec $x_1 < \dots < x_n$ alors pour $x \in \mathbb{R}$

$$F_X(x) = \sum_{i=1}^k P(X = x_i) \text{ avec } k \text{ tel que } x_k \leq x < x_{k+1}.$$

De même, si X prend une infinité de valeurs $\{x_1, \dots, x_n, \dots\}$ avec $x_1 < \dots < x_n \dots$, on a pour $x \in \mathbb{R}$

$$F_X(x) = \sum_{i=1}^k P(X = x_i) \text{ avec } k \text{ tel que } x_k \leq x < x_{k+1}.$$

Les sauts de la fonction de répartition F_X ont lieu en les points x_i et la hauteur du saut au point x_i est égale à $P(X = x_i)$. Il suffit donc de calculer la fonction de répartition aux points x_i .

Proposition 33 Si X est à valeurs discrètes dans $\{x_1, \dots, x_n\}$ (ou $\{x_1, \dots, x_n, \dots\}$), la loi de X est entièrement caractérisée par $\{P(X = x_i) : i \geq 1\}$.

On remarque que

pour tout $i \geq 1$, $P(X = x_i) \in [0, 1]$,

$\sum_{i \geq 1} P(X = x_i) = 1$ (En effet, $1 = P(x \in \mathbb{R}) = P(X = x_i)$.)

Exemple 34

Reprenons les deux premiers exemples précédents, qui décrivent effectivement des variables discrètes. Pour trouver la loi de ces variables on utilise la proposition précédente.

X est à valeurs dans $\{2, 3, \dots, 12\}$, donc $P(X = k) = 0$ pour $k \notin \{2, 3, \dots, 12\}$

k	2	3	4	5	6	7	8
$P(X = k)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36
$F_X(k)$	1/36	3/36	6/36	10/36	15/36	21/36	26/36

9	10	11	12
4/36	3/36	2/36	1/36
30/36	33/36	35/36	1

Exemple 35

Y est à valeurs dans $\{1, 2, \dots, 6\}$, donc $P(Y = k) = 0$ pour $k \notin \{1, 2, \dots, 6\}$.

k	1	2	3	4	5	6
$P(Y = k)$	1/36	3/36	5/36	7/36	9/36	11/36
$F_Y(k)$	1/36	4/36	9/36	16/36	25/36	1

2.2.2 Loi d'une variable aléatoire à densité (ou continues)

Considérons la durée de vie d'une bactérie. On conçoit facilement que la probabilité que cette durée de vie vaille exactement une certaine valeur est nulle. Par exemple, il est quasiment impossible qu'une bactérie vive exactement 1 an 0 mois, 0 heure, 0 minute

La fonction de répartition d'une telle variable est par conséquent continue. On peut par contre s'intéresser à la probabilité que la bactérie vive moins d'un an.

Définition 36 Une variable aléatoire X est à densité, ou continue, s'il existe une fonction f définie sur \mathbb{R} telle que la fonction de répartition de X s'écrit

$$\forall x \in \mathbb{R}, F_X(x) = \int_{-\infty}^x f(t) dt$$

où f est une fonction intégrable sur \mathbb{R} satisfaisant les conditions suivantes :

1. $f(t) \geq 0$ pour tout $t \in \mathbb{R}$,
2. $\int_{-\infty}^{+\infty} f(t) dt = 1$.

Une fonction qui vérifie les conditions 1 et 2 est appelée densité de probabilité.

Propriétés Soit X une variable aléatoire à densité. Alors pour tout $x \in \mathbb{R}$, $P(X = x) = 0$.

Preuve. Soit $x \in \mathbb{R}$. On considère l'intervalle réduit à un point $I = \{x\}$. On a

$$P(X = x) = P(X \in I) = \int_x^x f(t) dt = 0.$$

Remarque 37 1. La probabilité $P(X \in [a, b]) = \int_a^b f(t) dt$ correspond à l'aire de la surface comprise entre la courbe de f et l'axe des abscisses sur l'intervalle $[a, b]$.

2. La fonction de répartition d'une variable à densité est continue.

Proposition 38 Soit X une variable aléatoire de fonction de répartition F_X . Si F_X est continue sur \mathbb{R} et dérivable sur \mathbb{R} (sauf peut-être en un nombre fini de points), alors X est une variable à densité f donnée par $f(x) = F_X'(x)$.

2.3 Espérance et variance d'une variable aléatoire

L'idée intuitive de l'espérance puise son origine dans les jeux de hasard. Considérons le jeu suivant : on lance un dé plusieurs fois de suite. Supposons que pour une mise de 1 DA, on gagne 1 DA si le résultat obtenu est pair, 2 DA si le résultat est 1 ou 3, et on perd 3 DA si le résultat est 5. Est-il intéressant de jouer à ce jeu ? Quel peut-être le gain moyen ?

Soit X la variable aléatoire correspondant au nombre dinars gagnés ou perdus. La loi de X est

k	-3	1	2
$P(X = k)$	1/6	1/2	1/3

L'espérance de gain, noté $E[X]$, est alors

$E[X] = -3 \times 1/6 + 1 \times 1/2 + 2 \times 1/3 = 2/3$. Le joueur gagne donc en moyenne 2/3 de DA pour une mise de 1 DA. . .

Définition 39 L'espérance d'une variable aléatoire X est notée $E[X]$. Elle représente la valeur moyenne prise par la variable X .

Si X est une variable discrète à valeurs dans $D = \{x_1, \dots, x_n\}$, son espérance est

$$E[X] = x_1P(X = x_1) + \dots + x_nP(X = x_n) = \sum_{i=1}^n x_iP(X = x_i).$$

Si X est une variable discrète à valeurs dans l'ensemble infini $D = \{x_i : i \geq 1\}$, lorsque la somme est bien définie, son espérance est

$$E[X] = \sum_{i=1}^{+\infty} x_iP(X = x_i).$$

Définition 40

Si X est une variable à densité f , lorsque l'intégrale est bien définie, son espérance est

$$E[X] = \int_{-\infty}^{+\infty} x_i f(x) dx.$$

Lorsqu'une variable X vérifie $E[X] = 0$, on dit que la variable est centrée.

Propriétés

L'espérance est linéaire : soient a et $b \in \mathbb{R}$, deux variables aléatoires X et Y d'espérance finie alors

$$E[aX + bY] = aE[X] + bE[Y].$$

Si $X > 0$, alors $E[X] \geq 0$.

Si $X \leq Y$, alors $E[X] \leq E[Y]$.

Exemple 41 Supposons que la durée de vie T d'une bactérie est modélisée par la loi exponentielle de densité $f(t) = \lambda \exp(-\lambda t)$ pour $t \geq 0$ pour une certaine valeur de λ . Alors sa durée de vie moyenne est $E(T) = 1/\lambda$.

2.3.1 Mesurer la diffusion d'une variable : variance et écart type

On a vu que l'espérance correspondant à la valeur moyenne d'une variable aléatoire. L'écart type représente l'écart moyen (la distance moyenne) entre la variable et sa moyenne. Elle mesure la dispersion d'une variable, plus l'écart-type est grand plus la variable prend des valeurs qui peuvent être éloignées les unes des autres, plus l'écart-type est petit plus la variable prend des valeurs proches de sa moyenne.

Définition 42 La variance d'une variable aléatoire X , notée $Var(X)$, est définie par

$$Var(X) = E [(X - E(X))^2]$$

L'écart type est la racine carrée de la variance :

$$\sigma(X) = \sqrt{Var(X)}.$$

Lorsqu'une variable X vérifie $Var(X) = 1$, on dit que la variable est réduite.

La variance s'écrit aussi $Var(X) = E[X^2] - E[X]^2$

Propriétés :

$Var(X) = 0$ ssi X est constante.

Soient a et $b \in \mathbb{R}$, alors $Var(aX + b) = a^2 Var(X)$.

Exemple 43 Supposons que la durée de vie T d'une bactérie est modélisée par la loi exponentielle de densité $f(t) = \lambda \exp(-\lambda t)$ pour $t \geq 0$ pour une certaine valeur de λ . La variance de la durée de vie de la bactérie étudiée est $Var(T) = 1/\lambda^2$.

2.3.2 Corrélation de deux variables aléatoires

Définition 44 Soient X et Y deux v.a. réelles intégrables. La covariance de X et Y , ou coefficient de corrélation, est définie par

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

Lorsque $Cov(X, Y) = 0$, on dit que X et Y sont non corrélées.

Remarque 45 Si $\sigma(X) > 0$ et $\sigma(Y) > 0$, on définit le coefficient de corrélation de X et Y par

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

Lemme 46 Si X et Y sont indépendantes, alors elles sont non corrélées.

Remarque 47 La réciproque est fausse.

2.4 Inégalités en probabilités

2.4.1 Inégalité de Markov et de Bienaymé-Tchebychev

En théorie des probabilités, l'inégalité de Markov donne une majoration de la probabilité qu'une variable aléatoire réelle à valeurs positives soit supérieure ou égale à une constante positive. Cette inégalité a été nommée ainsi en l'honneur d'Andreï Markov.

Proposition 48 *Si X est une v.a dans $\mathcal{L}^1(\Omega, P)$ on a pour tout $\lambda > 0$*

$$P(|X| > \lambda) \leq \frac{E(|X|)}{\lambda}.$$

Preuve. La v.a $|X|$ peut s'écrire $|X| = |X| \cdot \mathbf{1}_{\{|X| > \lambda\}} + |X| \cdot \mathbf{1}_{\{|X| \leq \lambda\}}$, et par additivité et positivité de l'espérance on a

$$E(|X|) \geq E(|X| \cdot \mathbf{1}_{\{|X| > \lambda\}})$$

Or, $|X| \cdot \mathbf{1}_{\{|X| > \lambda\}} \geq \lambda \cdot \mathbf{1}_{\{|X| > \lambda\}}$ et par conséquent

$$E(|X|) \geq \lambda E(\mathbf{1}_{\{|X| > \lambda\}})$$

c'est à dire

$$E(|X|) \geq \lambda P(\{|X| > \lambda\}).$$

La proposition précédente est une version quantitative du fait que la probabilité que X prenne de grandes valeurs a tendance à être petite.

Définition 49 *On dit qu'une v.a X est dans $\mathcal{L}^2(\Omega, P)$ si son carré est P -intégrable,*

$$E(|X|^2) < \infty.$$

Proposition 50 *Si X est une v.a dans $\mathcal{L}^2(\Omega, P)$ on a pour tout $\lambda > 0$*

$$P(|X| > \lambda) \leq \frac{E(|X|^2)}{\lambda^2}.$$

Preuve. Il suffit de remarquer que $\{|X| > \lambda\} = \{X^2 > \lambda^2\}$ et d'appliquer la proposition précédente à la v.a $Y = X^2$.

Il existe une version plus générale de cette inégalité.

Théorème 51 Soit X une variable aléatoire de X est une v.a dans $\mathcal{L}^p(\Omega, P)$ on a pour tout $\lambda > 0$

$$P(|X| \geq \lambda) \leq \frac{E(|X|^p)}{\lambda^p}.$$

Preuve. La démonstration tient entièrement au fait que pour tout λ strictement positif, $\lambda^p \mathbf{1}_{\{|X| \geq \lambda\}} \leq |X|^p$. Ici $\mathbf{1}_{\{|X| \geq \lambda\}}$ désigne l'indicatrice de l'événement $\{|X| \geq \lambda\}$. Par croissance de l'espérance, on obtient :

$$E(|X|^p) \geq E(\lambda^p \mathbf{1}_{\{|X| \geq \lambda\}}) = \lambda^p E(\mathbf{1}_{\{|X| \geq \lambda\}}) = \lambda^p P(\{|X| \geq \lambda\}).$$

En divisant de part et d'autre de l'inégalité par λ^p on trouve le résultat recherché.

Corollaire 52 Soit ϕ une fonction croissante positive ou nulle sur un intervalle I . Soit X une variable aléatoire réelle définie sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$ telle que $P(X \in I) = 1$.

Alors :

$$\forall b \in I | \phi(b) > 0, P(X \geq b) \leq \frac{E(\phi(X))}{\phi(b)}.$$

Preuve. On applique l'inégalité de Markov à $Y = \phi(X)$ et $a = \phi(b)$ pour obtenir :

$$\forall b > 0, P(\phi(X) \geq \phi(b)) \leq \frac{E(\phi(X))}{\phi(b)}$$

La croissance de ϕ entraîne : $\{X \geq b\} \implies \{\phi(X) \geq \phi(b)\}$. Par conséquent : $P(X \geq b) \leq \frac{E(\phi(X))}{\phi(b)}$.

L'inégalité de Bienaymé-Tchebychev est une inégalité de concentration permettant de montrer qu'une variable aléatoire prendra avec une grande probabilité une valeur relativement proche de son espérance. Ce résultat s'applique dans des cas très divers, nécessitant la connaissance de peu de propriétés (seules l'espérance et la variance doivent être connues), et permet de démontrer la loi faible des grands nombres.

Ce théorème doit son nom aux mathématiciens Irénée-Jules Bienaymé, qui fut le premier à le formuler, et Pafnouti Tchebychev qui le démontra. Appliquée à la v.a $Y = X - E(X)$ la proposition précédente donne le théorème de Bienaymé-Tchebychev :

Théorème 53 Si X est une v.a dans $\mathcal{L}^2(\Omega, P)$ on a pour tout $\lambda > 0$

$$P(|X - E(X)| > \lambda) \leq \frac{Var(|X|)}{\lambda^2}.$$

Si on note $\sigma = \sqrt{Var(X)}$ l'écart type on a donc

$$P(|X - E(X)| > \lambda\sigma) \leq \frac{1}{\lambda^2}.$$

Le théorème de Bienaymé-Tchebychev permet d'obtenir les probabilités des déviations importantes de la v.a X par rapport à sa moyenne.

2.4.2 Inégalité de Cauchy-Schwarz et Minkowski

Théorème 54 (Cauchy-Schwarz) Si X et Y sont dans $\mathcal{L}^2(\Omega, P)$ alors le produit $X.Y$ est dans $\mathcal{L}^1(\Omega, P)$ et on a

$$E(|X.Y|) \leq E(|X|^2)^{1/2} . E(|Y|^2)^{1/2} .$$

En outre, l'inégalité précédente est une égalité si et seulement si il existe un réel tel que P -presque sûrement $Y = \lambda X$ ou $X = \lambda Y$ (X et Y sont colinéaires).

Preuve. Rappelons que pour tous réels a, b on a $|a.b| \leq \frac{1}{2}(a^2 + b^2)$. Par conséquent, $|X.Y| \leq \frac{1}{2}(X^2 + Y^2)$ et en prenant l'espérance de chaque membre de l'inégalité on obtient

$$E(|X.Y|) \leq \frac{1}{2} (E(|X|^2) + E(|Y|^2)) ,$$

ce qui démontre que $X.Y$ est intégrable.

Supposons $E(X^2) \neq 0$ et considérons pour t réel la quantité suivante qui est clairement toujours positive ou nulle (espérance d'un carré) :

$$E((tX + Y)^2) = E(X^2)t^2 + 2E(X.Y)t + E(Y^2) .$$

(si $E(X^2) = 0$ et $E(Y^2) \neq 0$ on échange X et Y , si $E(X^2) = E(Y^2) = 0$ il n'y a rien à démontrer car X et Y sont nulles P -ps). Vue comme fonction de t c'est un polynôme quadratique si $E(X^2) \neq 0$. Comme il ne prend que des valeurs positives ou nulles son discriminant $\Delta = 4E(X.Y)^2 - E(X^2)E(Y^2)$ doit être négatif ou nul (sinon ce polynôme admettrait deux racines réelles distinctes et serait strictement négatif entre les racines). Ceci n'est rien d'autre que l'inégalité annoncée.

L'égalité a lieu si et seulement si le discriminant s'annule. Or, dans ce cas le polynôme quadratique $E((tX + Y)^2)$ admet une racine réelle t_0 (et une seule). On a donc $E(t_0X + Y) = 0$ ce qui signifie que $t_0X + Y = 0$ P -presque sûrement.

Théorème 55 (Minkowski) L'espace $\mathcal{L}^2(\Omega, P)$ est un R -espace vectoriel et on a

$$E(|X + Y|^2)^{\frac{1}{2}} \leq E(|X|^2)^{1/2} + E(|Y|^2)^{1/2} .$$

Preuve. Il suffit de démontrer que $E((X + Y)^2)$ est finie si $E(X^2)$ et $E(Y^2)$ le sont. Or,

$$E((X + Y)^2) = E(X^2) + 2E(X.Y) + E(Y^2) ,$$

et comme

$$E((X.Y)) \leq E(X^2)^{1/2} . E(Y^2)^{1/2}$$

on a

$$\begin{aligned} E((X + Y)^2) &\leq E(X^2) + E(X^2)^{1/2} . E(Y^2)^{1/2} + E(Y^2) \\ &= \left(E(|X|^2)^{1/2} + E(|Y|^2)^{1/2} \right)^2 . \end{aligned}$$

2.4.3 Inégalité de Jensen

En mathématiques, l'inégalité de Jensen est une relation utile et très générale concernant les fonctions convexes, due au mathématicien danois Johan Jensen et dont il donna la preuve en 1906. On peut l'écrire de deux manières : discrète ou intégrale. Elle apparaît notamment en analyse, en théorie de la mesure et en probabilités (théorème de Rao-Blackwell), mais également en physique statistique, en mécanique quantique et en théorie de l'information (sous le nom d'inégalité de Gibbs).

On rappelle qu'une fonction $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ est convexe si

$$\varphi(tx + (1 - t)y) \leq t\varphi(x) + (1 - t)\varphi(y)$$

pour tout choix de $x, y \in \mathbb{R}$ et de $t \in [0, 1]$. La fonction φ est concave si l'inégalité inverse a lieu, c'est-à-dire si $-\varphi$ est convexe.

Théorème 56 *Soit φ une fonction convexe sur un intervalle réel I et X une variable aléatoire à valeurs dans I , dont l'espérance $E(\varphi(X))$ existe. Alors,*

$$\varphi(E(X)) \leq E(\varphi(X)).$$

2.5 Exercices

Exercice 57 *Soit X une v.a.r discrète prenant les valeurs 3, 4, 5 et 6. Déterminer la loi de probabilité de X sachant que :*

$$P(X < 5) = \frac{1}{3}, P(X > 5) = \frac{1}{2}, P(X = 3) = P(X = 4).$$

Exercice 58 On considère un dé cubique truqué dont les faces sont numérotés de 1 à 6 et on note X la variable aléatoire donnée par le numéro de la face du dessus. On suppose que le dé est truqué de sorte que la probabilité d'obtenir une face est proportionnelle au numéro inscrit sur cette face.

1. Déterminer la loi de X , calculer son espérance.
2. On pose $Y = 1/X$. Déterminer la loi de Y , et son espérance.

Exercice 59 On lance 2 dés et on appelle Z la v.a.r égale à la valeur absolue de la différence des numéros obtenus. Déterminer la loi de Z , sa fonction de répartition, son espérance et sa variance.

Exercice 60 Parmi les fonctions suivantes définies sur \mathbb{R} , déterminer lesquelles sont la densité d'une variable aléatoire à densité. Calculer le cas échéant leur fonction de répartition et préciser si elles admettent une espérance.

$$f_1(x) = \begin{cases} \cos x & \text{si } x \in [0, \frac{\pi}{2}] \\ 0 & \text{sinon.} \end{cases}, \quad f_2(x) = \frac{1}{1+x^2}, x \in \mathbb{R},$$

$$f_3(x) = \frac{e^x}{(e^x+1)^2}, x \in \mathbb{R}, \quad f_4(x) = \begin{cases} 1+x & \text{si } x \in [-1, 0] \\ 1-x & \text{si } x \in [0, 1] \\ 0 & \text{sinon} \end{cases},$$

$$f_5(x) = \begin{cases} \frac{1}{|x|^3} & \text{si } |x| > 1 \\ 0 & \text{sinon} \end{cases}, \quad f_6(x) = \sin x + 1, x \in \mathbb{R}.$$

Exercice 61 Déterminer si les fonctions suivantes sont des densités de probabilité et si oui déterminer la fonction de répartition de la VAR associée à cette densité.

$$f(t) = \begin{cases} 0 & \text{si } t < 0 \\ 4te^{-2t} & \text{si } t \geq 0 \end{cases}, h(t) = \begin{cases} 0 & \text{si } t < 0 \\ \frac{1}{2 \ln 2} e^{-t} \ln(1+e^t) & \text{si } t \geq 0 \end{cases},$$

On pensera au changement de variable $u = e^{-t}$.

Exercice 62 Calculer, si elles existent, l'espérance et la variance de la variable X dont une densité est :

$$g(t) = \begin{cases} 0 & \text{si } t < 0 \\ 4te^{-2t} & \text{si } t \geq 0 \end{cases}, h(t) = \begin{cases} 0 & \text{si } t < 1 \\ \frac{4 \ln t}{t^3} & \text{si } t \geq 1 \end{cases}$$

Exercice 63 Soit f la fonction définie sur \mathbb{R} par :

$$f(x) = \frac{1}{2(1+|x|)^2}.$$

- Démontrer que f est la densité de probabilité d'une variable aléatoire X . On note F sa fonction de répartition (qu'on ne demande pas de calculer).
- On considère la variable aléatoire $Y = \ln(1 + |X|)$ et on note G sa fonction de répartition G . Exprimer G en fonction de F .
- En déduire que Y admet une densité que l'on calculera.
- Reconnaître la loi de Y .

Exercice 64 Soit X une variable aléatoire dont une densité est la fonction f définie sur \mathbb{R} par :

$$f(x) = \begin{cases} kx(x-2) & \text{si } 0 \leq x \leq 2 \\ 0 & \text{sinon} \end{cases}$$

- Déterminer la valeur de k pour que la fonction f définie une densité de probabilité de X .
- Déterminer la fonction de répartition F de X puis calculer l'espérance et la variance.

Chapitre 3

Lois de probabilités

Dans ce chapitre nous présentons les lois de probabilité les plus utiles et les plus utilisées dans les diverses applications de la Statistique. Nous en donnons également les caractéristiques principales, ainsi que leurs propriétés essentielles.

3.1 Lois de probabilités discrètes usuelles

Parmi les lois des variables discrètes, nous avons la grande famille des lois des Indicatrices et associées : lois de Bernoulli, lois Binomiales, lorsqu'il s'agit de tirages avec remise, et les lois Hypergéométriques pour des tirages sans remise. Nous présentons également les versions multivariées de ces lois : les lois Multinomiales et les lois Polyhypergéométriques ; de plus nous abordons les lois Géométriques. Nous avons ensuite les lois de Poisson qui sont très utiles pour les phénomènes observés dans le temps. Nous terminerons avec les lois Uniformes discrètes qui sont plutôt d'une utilisation formelle.

Loi de Bernoulli, $B(p)$, avec $p \in]0, 1[$.

Une variable aléatoire X de Bernoulli est une variable qui ne prend que deux valeurs : l'échec (auquel on associe la valeur 0) et le succès (auquel on associe la valeur 1) d'une expérience.

Cette expérience est appelée épreuve de Bernoulli. Par exemple, on souhaite savoir si une cellule est atteinte par un virus. On associe la valeur 1 si elle est atteinte (succès) et la valeur 0 si elle est saine (échec).

La loi est donnée par : $P(X = 1) = p$, $P(X = 0) = 1 - p$,

(i.e. $P(X = k) = p^k q^{1-k}$, $k \in \{0, 1\}$).

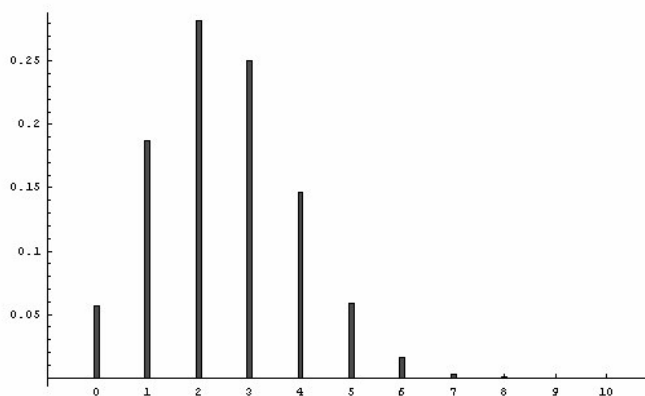


FIG. 3.1 – Loi de Probabilités $B\left(10, \frac{1}{4}\right)$.

$$E(X) = p, \text{var}(X) = pq.$$

Loi Binomiale, $B(n, p)$, avec $p \in]0, 1[$, $n \in \mathbb{N}$, $n \geq 1$.

La loi Binomiale est utilisée pour modéliser un "sondage avec remise". C'est la loi du nombre de succès lorsqu'on renouvelle n fois de manière indépendante une épreuve de Bernoulli de paramètre p . On note X le nombre de succès obtenus à l'issue des n épreuves. Sa loi s'appelle loi Binomiale de paramètres n et p .

La loi est donnée par : $P(X = k) = C_n^k p^k q^{n-k}$, $k \in \{0, \dots, n\}$.

$$E(X) = np, \text{var}(X) = npq.$$

On peut écrire le nombre de succès X à l'aide des résultats de chaque épreuve de Bernoulli. On note X_i le résultat de la i ème expérience :

$$X_i = \begin{cases} 1 & \text{si la } i^{\text{ème}} \text{ expérience est réussie,} \\ 0 & \text{si la } i^{\text{ème}} \text{ expérience est échec.} \end{cases}$$

On a alors $X = X_1 + \dots + X_n$. Toute variable de Bernoulli peut s'écrire de cette façon.

Exemple 65 Dans une entreprise, il y a dix imprimantes identiques fonctionnant de façon indépendante tous les jours. Chaque jour, la probabilité qu'une imprimante tombe en panne est égale à 0,002. Le risque de panne un jour donné est indépendant des pannes survenues les jours précédents.

- Déterminer la probabilité qu'une imprimante tombe en panne au moins une fois pendant un mois (de 30 jours). Donner la valeur arrondie aux millièmes.

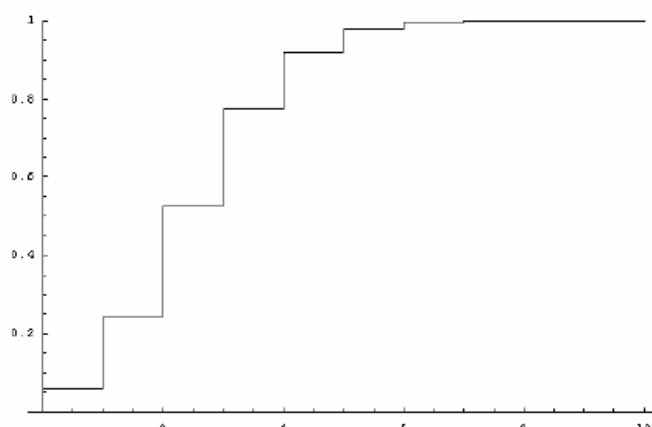


FIG. 3.2 – Fonction de répartition $B\left(10, \frac{1}{4}\right)$.

On note X la variable aléatoire correspondant au nombre de jours de panne d'une imprimante pendant les 30 jours.

On a alors $X = \{0, 1, 2, \dots, 29, 30\}$. On considère l'épreuve de Bernoulli qui consiste à prendre un jour au hasard parmi les 30 jours du mois et ayant les issues possibles S : « l'imprimante est en panne » et $E = \bar{S}$: « l'imprimante n'est pas en panne ».

On a $P(S) = 0,002$ et $P(E) = 1 - 0,002 = 0,998$

Le risque de panne un jour donné est indépendant des pannes survenues les jours précédents, chaque épreuve de Bernoulli est indépendante des autres, la loi de probabilité de X suit donc la loi binomiale de paramètres 30 et 0,002 notée aussi $B(30; 0,002)$.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - 0,998^{30} \simeq 0,058 \text{ (i.e. } P(X = 0) = C_{30}^0 \times 0,002^0 \times 0,998^{30})$$

La probabilité que l'imprimante tombe en panne au moins une fois en 30 jours est d'environ 0,058.

Loi Multinomiale $\mathcal{M}(n, p_1, \dots, p_k)$

Soit une expérience possédant k résultats possibles A_1, \dots, A_k de probabilités respectives p_1, \dots, p_k . Soit X_i la v.a qui est égale au nombre x_i de réalisations de A_i au cours de N épreuves indépendantes (problème de tirage successifs avec remise).

La probabilité que A_i soit réalisé x_i fois, est égale à $p_i^{x_i}$ (événement indépendant). La probabilité que les kA_i soient réalisés, est égale à $\prod_{i=1}^k p_i^{x_i}$ fois le nombre de permutations avec répétitions de x_i

parmi $n = \sum_{i=1}^k x_i$, c'est-à-dire fois $\frac{n!}{\prod_{i=1}^k x_i!}$; $\sum_{i=1}^k p_i = 1$.

loi de probabilités :

$$P(X = x) = P(X_1 = x_1, \dots, X_n = x_n) = \frac{n!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k p_i^{x_i},$$

$$\text{avec } \sum_{i=1}^k x_i = n \text{ et } \sum_{i=1}^k p_i = 1$$

$E(X)$ a pour composantes les espérances marginales $E(X_i)$.

$Var(X)$ a pour composantes les variances marginales $Var(X_i)$

Loi Hypergéométrique $\mathcal{H}(N, m, n)$, avec $N \geq 1, (m, n) \in \{1, \dots, N\}^2$.

La loi hypergéométrique est utilisée pour modéliser un "sondage sans remise". C'est le cas de pratiquement tous les sondages (notamment lorsqu'on veut étudier la conformité d'un lot de médicaments, étudier le nombre de cellules atteintes par un virus, etc. . .).

Le paramètre N est l'effectif de la population totale, m celui de la sous-population à laquelle on s'intéresse et n la taille de l'échantillon observé.

Exemple 66 *Un lac contient N poissons. On en pêche m qu'on marque et qu'on remet à l'eau. On pêche à nouveau n poissons. Quelle est la probabilité de pêcher k poissons marqués ?*

La loi est donnée par :

$$P(X = k) = \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \text{ si } k \in \{0, \dots, \min(m, n)\}.$$

$$E(X) = np, \text{ var}(X) = np(1-p) \frac{N-n}{N-1} \text{ tel que } p = \frac{m}{N}.$$

Loi Poly-Hypergéométrique $\mathcal{PH}(N_1, \dots, N_r; n), n \in \mathbb{N}$

Nous généralisons les lois Hypergéométriques au cas multivarié.

Soit un ensemble contenant N unités et partitionné en r sous-ensembles, chacun contenant N_j unités, $j = 1, \dots, r$. Les unités du sous-ensemble d'ordre j possèdent exclusivement, par exemple, une caractéristique \mathcal{C}_j , pour tous les $j = 1, \dots, r$. Nous avons $N = N_1 + \dots + N_r$. Nous procédons à n tirages sans remise. Nous considérons le v.a. $X = (X_1, \dots, X_r)^t$, de dimension r , dont la composante

d'ordre j , notée X_j , admet comme réalisation le nombre d'unités possédant la caractéristique \mathcal{C}_j parmi les unités extraites, et ceci pour tout les $j = 1, \dots, r$. Alors la loi de ce v.a. est donnée par :

$$P(X = (k_1, \dots, k_r)^t) = \frac{C_{N_1}^{k_1} C_{N_2}^{k_2} \dots C_{N_r}^{k_r}}{C_N^n},$$

pour $k_j = 0, 1, \dots, n$ et $j = 1, \dots, r$, de plus nous avons $k_1 + \dots + k_r = n$. Nous appelons la loi de loi X Poly-hypergéométrique de dimension r et de paramètres $N_1, \dots, N_r, n \in \mathbb{N}$. Nous notons $\mathcal{L}(X) = \mathcal{PH}(N_1, \dots, N_r; n)$.

Modélisation : Le v.a. X peut s'écrire comme somme des indicatrices vectorielles Y_1, \dots, Y_n . L'indicatrice Y_i a toutes ses composantes nulles à l'exception de celle dont l'indice correspond à l'ordre de la caractéristique possédée par l'unité extraite lors de ce tirage d'ordre i , pour tous les $i = 1, \dots, n$.

Remarque 67 Pour montrer que la somme des probabilités vaut bien 1, nous utilisons l'identité généralisée de Vandermonde :

$$C_N^n = \sum_{\substack{0 \leq k_1, \dots, k_r \\ k_1 + \dots + k_r = n}} C_{N_1}^{k_1} C_{N_2}^{k_2} \dots C_{N_r}^{k_r}, \text{ pour tous } n, N_1, \dots, N_r \in \mathbb{N}.$$

Celle-ci s'obtient en identifiant les coefficients de même degré dans les deux polynômes en a :

$$(1 + a)^N = (1 + a)^{N_1} (1 + a)^{N_2} \dots (1 + a)^{N_r}.$$

Approximation : Si l'ensemble est assez grand et si nous extrayons assez peu d'unités, en pratique $N \geq 50$ et $n \leq 0.1N$, alors nous pouvons assimiler des tirages sans remise à des tirages avec remise. C'est-à-dire que dans ces conditions nous avons :

$$\mathcal{L}(X) = \mathcal{PH}(N_1, \dots, N_r; n) \approx \mathcal{M}\left(n; \frac{N_1}{N}, \dots, \frac{N_r}{N}\right).$$

Remarque 68 Si $\mathcal{L}(X) = \mathcal{PH}(N_1, \dots, N_r; n)$ alors pour tout $j = 1, \dots, r$, la loi marginale de X_j est une loi Hypergéométrique $\mathcal{H}(N_j, N - N_j; n)$. La loi marginale du v.a. $(X_{j_1}, X_{j_2}, n - X_{j_1} - X_{j_2})$ est une loi Poly-hypergéométrique $\mathcal{PH}(N_{j_1}, N_{j_2}, N - N_{j_1} - N_{j_2}; n)$. La loi conditionnelle de X_{j_1} sachant que $X_{j_2} = k_{j_2}$ est une loi Hypergéométrique $\mathcal{H}(N_1, N - N_2; n - k_{j_2})$.

Proposition 69 Si $\mathcal{L}(X) = \mathcal{PH}(N_1, \dots, N_r; n)$ alors la moyenne théorique de ce vecteur est le vecteur défini par :

$$E(X) = \begin{pmatrix} E(X_1) \\ \vdots \\ E(X_r) \end{pmatrix} = \begin{pmatrix} n \frac{N_1}{N} \\ \vdots \\ n \frac{N_r}{N} \end{pmatrix}.$$

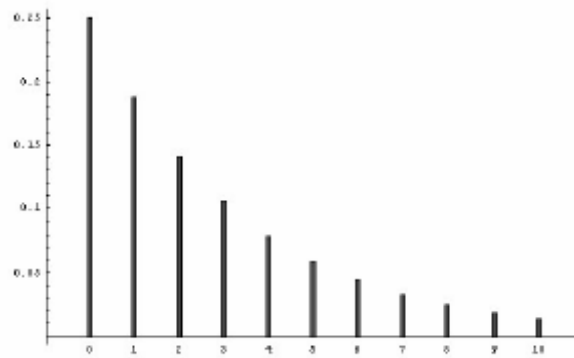


FIG. 3.3 – Loi de probabilités $G\left(\frac{1}{4}\right)$.

Proposition 70 Si $\mathcal{L}(X) = \mathcal{PH}(N_1, \dots, N_r; n)$ alors la matrice des variances-covariances théoriques de ce vecteur est la matrice définie par :

$$\begin{aligned} \sum(X) &= E[(X - E(X))^t (X - E(X))] \\ &= \frac{N-n}{N-1} \begin{pmatrix} n \frac{N_1}{N} \left(1 - \frac{N_1}{N}\right) & -n \frac{N_1}{N} \frac{N_2}{N} & \dots & -n \frac{N_1}{N} \frac{N_r}{N} \\ -n \frac{N_2}{N} \frac{N_1}{N} & n \frac{N_2}{N} \left(1 - \frac{N_2}{N}\right) & \dots & -n \frac{N_2}{N} \frac{N_r}{N} \\ \vdots & \vdots & \ddots & \vdots \\ -n \frac{N_r}{N} \frac{N_1}{N} & -n \frac{N_r}{N} \frac{N_2}{N} & \dots & n \frac{N_r}{N} \left(1 - \frac{N_r}{N}\right) \end{pmatrix}. \end{aligned}$$

Remarque 71 Nous constatons que le vecteur espérance est le même que celui de la loi Multinomiale. Dans la matrice des variances-covariances nous voyons apparaître les variances et les covariances de la loi Multinomiale, avant tirages, mais corrigée par l'indice d'exhaustivité $\frac{N-n}{N-1}$. Nous notons que, lorsque $n = N$, la matrice des variances-covariances est nulle, nous connaissons exactement le nombre d'unités de l'ensemble; si N est très grand par rapport à n , alors l'indice d'exhaustivité est proche de 1 et nous avons presque la même matrice des variances-covariances que celle de la loi Multinomiale.

Remarque 72 Pour voir ces propriétés il suffit d'utiliser l'identité généralisée de Vandermonde et la propriété $kC_l^k = lC_{l-1}^{k-1}$.

Loi Géométrique, $G(p)$, avec $p \in]0, 1[$.

La loi géométrique est la loi du premier succès.

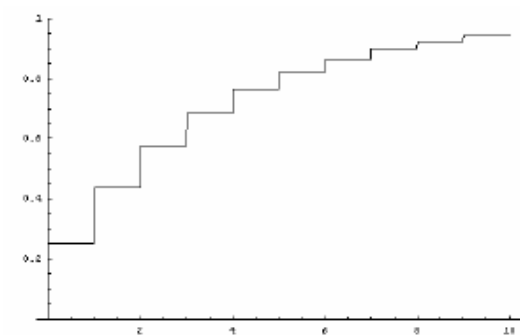


FIG. 3.4 – Foction de répartition $G(\frac{1}{4})$.

Exemple 73 On lance une pièce truquée jusqu'à ce qu'on obtienne une fois "Pile". On note p la probabilité de tomber sur "Pile". On veut connaître la probabilité d'avoir "Pile" au premier lancer, au deuxième, . . . , au k ème lancer, . . . On note X le nombre de lancers nécessaires pour avoir un succès.

La loi est donnée par : $P(X = k) = p(1 - p)^{k-1}$ avec $k \in \mathbb{N}, k \geq 1$.

$$E(X) = \frac{1}{p}, \text{var}(X) = \frac{1-p}{p^2}.$$

Une formule utile quand on veut faire des calculs avec la loi géométrique :

$$1 + x^2 + \dots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Loi de Poisson, $P(\lambda)$, avec $\lambda > 0$ un réel.

La loi de Poisson est utilisée pour modéliser le comptage d'événements rares, c'est à dire des événements ayant une faible probabilité de réalisation : maladies rares, accidents mortels rares, le titrage d'une solution virale, mutations ou recombinaisons dans une séquence génétique, pannes, radioactivité...

Exemple 74

La loi est donnée par $P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$ avec $k \in \mathbb{N}$.

$$E(X) = \text{var}(X) = \lambda.$$



FIG. 3.5 – Loi de probabilités $P(10)$.

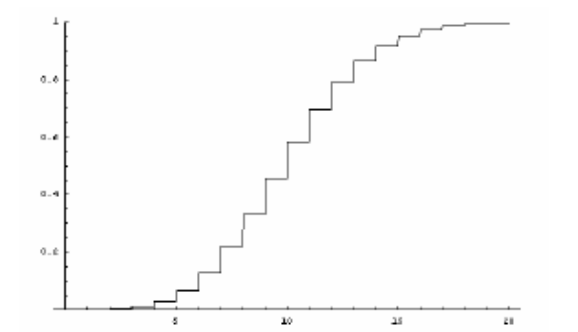


FIG. 3.6 – Fonction de répartition $P(10)$.

Une formule utile :

$$e^\lambda = 1 + \lambda + \frac{\lambda^2}{2} + \frac{\lambda^3}{3!} + \dots = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$$

Exemple 75 une suspension bactérienne contient 5000 bactéries/litre. On ensemence à partir de cette suspension, 50 boîtes de Pétri à raison d'1cm³ par boîte. Si X représente le nombre de colonies par boîte, alors la loi de probabilité de X est $P(\lambda = 5)$.

La probabilité qu'il n'y ait aucune colonie sur la boîte de Pétri est : $P(X = 0) = \frac{5^0 e^{-5}}{0!} = 0,0067$ (0,67% de chance)

La probabilité qu'il n'y ait au moins une colonie sur la boîte de Pétri est : $P(X > 0) = 1 - P(X = 0) = 1 - \frac{5^0 e^{-5}}{0!} = 0,9933$ (99,33% de chance)

Le nombre moyen de colonies attendu sur la boîte de Pétri est : $E(X) = \lambda = 5$ colonies.

$var(X) = \lambda = 5$ colonies.

Loi Uniforme, $\mathcal{U}(\{1, \dots, N\})$, avec $N \in \mathbb{N}^*$

Définition 76 Une v.a. X suit une loi uniforme discrète à valeur dans l'ensemble $\{1, \dots, N\}$ avec $N \in \mathbb{N}^*$ si :

$$P(X = k) = \frac{1}{N}, \forall k \in \{1, \dots, N\}.$$

Nous notons $\mathcal{L}(X) = \mathcal{U}(\{1, \dots, N\})$.

Modélisation : Cette loi décrit l'expérience qui consiste à extraire une unité au hasard dans un ensemble contenant N unités. L'expression «au hasard» veut dire que le dispositif d'extraction est tel que chaque unité a les mêmes chances d'être choisie. L'exemple le plus classique est celui du lancer d'un dé équilibré à N faces.

Propriété 1 :

Si $\mathcal{L}(X) = \mathcal{U}(\{1, \dots, N\})$, nous avons les résultats suivants :

$$E(X) = \frac{N+1}{2}, Var(X) = \frac{N^2-1}{12}.$$

Remarque 77 La loi Uniforme a une distribution symétrique autour de son espérance mais sa forme est platykurtique ou sous-gaussienne ou encore sous-normale, c'est-à-dire que la distribution est plus aplatie que la distribution gaussienne.

3.2 Lois de probabilités absolument continues usuelles

Parmi les lois des variables continues, une grande partie est réservée aux lois de Gauss ou Normales, les lois Log-Normales, ainsi qu'aux lois associées : lois de Student, lois du Khi-deux et lois de Fisher-Snedecor. Puis nous présentons les lois Gamma, avec comme cas particulier les lois Exponentielles, les lois Bêta, les lois de Pareto, utiles en Économie, les lois de Weibull, très utiles en Fiabilité et les lois de Gumbel utiles en assurances. Ensuite nous avons les lois de Cauchy et les lois Uniformes continues qui sont plutôt d'une utilisation formelle.

Loi Normale (ou loi Gaussienne), $\mathcal{N}(m, \sigma^2)$, avec m, σ deux réels tel $\sigma > 0$.

Cette loi introduit indépendamment par Gauss et par Laplace est la plus fréquemment rencontrée dans la nature. De ce fait leur importance en Statistique est primordiale. Nous pouvons affirmer que sans elle la Statistique n'existerait pas.

La loi Normale est une loi centrale dans la théorie des probabilités. Elle est notamment très utilisée en statistique.

Définition 78

On dit que la v.a.r continue suit la loi normale de paramètre m et σ lorsqu'elle admet pour densité de probabilité la fonction

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}} \text{ avec } x \in \mathbb{R}$$

On dit que la v.a.r continue suit la loi normale centrée réduite si $m = 0$ et $\sigma = 1$.

Notons que l'on peut calculer formellement la probabilité d'un intervalle. En effet, la fonction de répartition $F_X(x)$

$$F_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^x \exp \left\{ -\frac{1}{2} \frac{(t-m)^2}{\sigma^2} \right\} dt \text{ avec } x \in \mathbb{R}$$

$$E(X) = m, \text{ var}(X) = \sigma^2.$$

Pour passer d'une loi normale quelconque $\mathcal{N}(m, \sigma)$ à la loi normale centrée réduite $\mathcal{N}(0, 1)$ on a

Modélisation : Lorsqu'une mesure ou observation est le cumul d'un très grand nombre d'autres variables, indépendantes entre elles et individuellement négligeables, alors cette mesure peut être la réalisation d'une v.a. X qui suit une loi \mathcal{N} . Cette propriété, connue sous le nom de Théorème de la Limite Centrale, est celle qui donne toute son importance à la loi Normale.

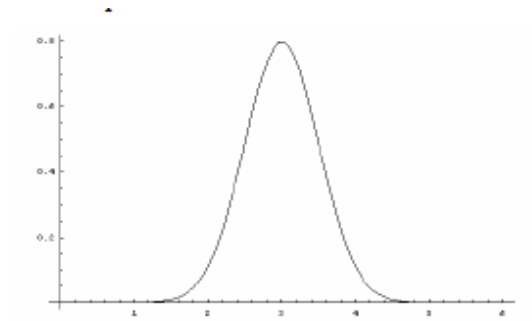


FIG. 3.7 – Loi de probabilités $\mathcal{N}(3, 0.5)$.

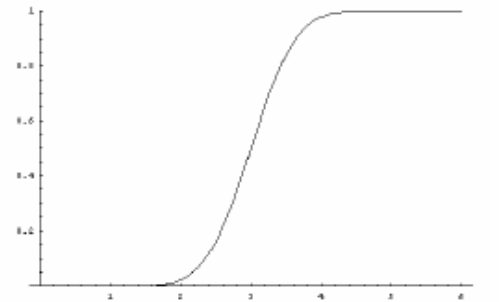
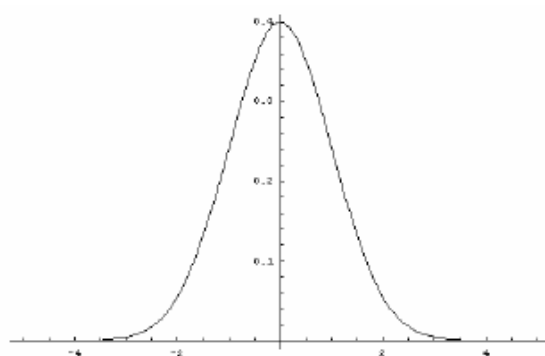


FIG. 3.8 – Fonction de répartition $\mathcal{N}(3, 0.5)$.

FIG. 3.9 – Loi de probabilités $\mathcal{N}(0, 1)$.

Corollaire 79 Soit X une variable aléatoire réelle continue de densité de probabilité $\mathcal{N}(m, \sigma^2)$ avec $\sigma > 0$ et $m \in \mathbb{R}$. On pose $Y := \frac{X-m}{\sqrt{\sigma^2}}$. Alors, la loi Y suit la loi de probabilité $\mathcal{N}(0, 1)$.

On a

$$\begin{aligned} P(a \leq X \leq b) &= P\left(\frac{a-m}{\sigma} < Y < \frac{b-m}{\sigma}\right) \\ &= \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right) \end{aligned}$$

où $\Phi(\cdot)$ la fonction de répartition de la loi centrée réduite.

Conséquemment, la connaissance de la fonction Φ est suffisante. Etudions ses propriétés : $\Phi(0) = \frac{1}{2}$, et pour tout $t \in \mathbb{R}$, l'égalité suivante est vérifiée : $\Phi(-t) = 1 - \Phi(t)$.

Exemple 80 Soit une variable aléatoire réelle X suivant la loi $\mathcal{N}(5, 4)$.

Calculons $P(1 < X < 7)$. On considère d'abord $Y := \frac{X-5}{\sqrt{4}}$. Alors $Y \sim \mathcal{N}(0, 1)$ puis :

$$\begin{aligned} P(1 < X < 7) &= P\left(\frac{1-5}{2} < \frac{X-5}{2} < \frac{7-5}{2}\right) \\ &= P(-2 < Y < 1) \\ &= \Phi(1) - \Phi(-2) \\ &= \Phi(1) + \Phi(2) - 1. \end{aligned}$$

Or, d'après la table, $\Phi(1) \approx 0.8413$ et $\Phi(2) \approx 0.9772$ d'où $P(1 < X < 7) \approx 0.8185$.

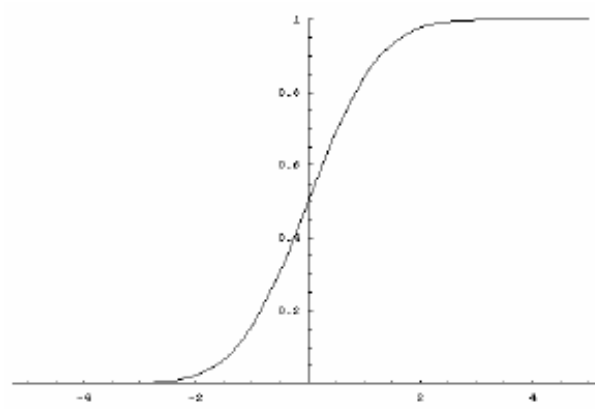


FIG. 3.10 – Fonction de répartition $\mathcal{N}(0, 1)$.

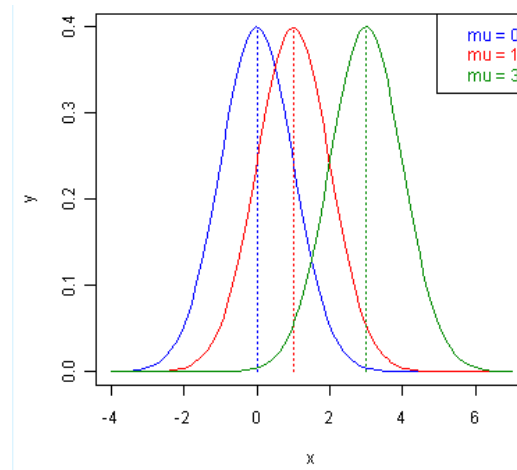
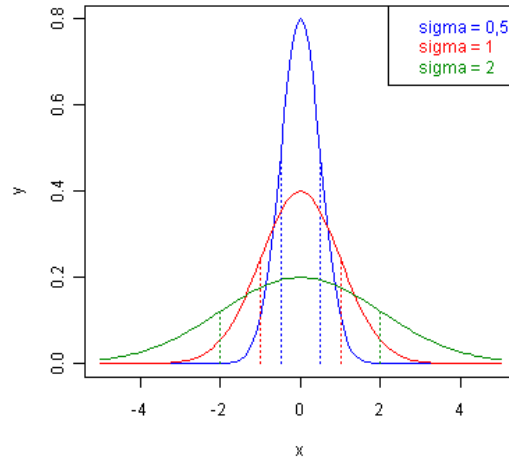


FIG. 3.11 – Densités normales avec $\sigma = 1$.

FIG. 3.12 – Densités normales avec $m = 0$.

Loi Log-normale $LN(m, \sigma^2, x)$ avec $m \in \mathbb{R}, \sigma > 0$ réel

Une variable X suit une loi Log-Normale $LN(m, \sigma^2, X_0)$, si $\ln(X)$ suit une loi normale $\mathcal{N}(m, \sigma^2)$. Cette loi permet de modéliser par exemple le temps de survie des bactéries en présence de désinfectant, le dosage de certains médicaments ...

Densité :

$$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} e^{-\frac{(\ln(x)-m)^2}{2\sigma^2}} \text{ avec } x > 0.$$

Loi de Khi-deux χ_n^2

Les premières lois du Khi-deux ont été mentionnées par Helmer puis développées par Pearson.

Définition 81 Soit X_1, \dots, X_n des v.a. indépendantes de loi Normale $\mathcal{N}(\mu_i; 1)$, $i = 1, \dots, n$, respectivement. Nous posons :

$$X = \sum_{i=1}^n X_i^2.$$

Si $\mu_1 = \dots = \mu_n = 0$, alors la loi de X est une loi de Khi-deux à n degré de liberté. Ceci noté $\mathcal{L}(X) = \chi_n^2$.

Si au moins un des $\mu_i, i = 1, \dots, n$ n'est pas nul, alors la loi de X est une loi du Khi-deux excentrée à n degré de liberté et avec un paramètre d'excentricité $\delta = \sum_{i=1}^n \mu_i^2$. Ceci est noté $\chi_n^2(\delta)$.

Modélisation : Ces lois sont utilisées dans l'étude des écarts quadratiques, en particulier pour l'estimation de variances. Elles interviennent également dans les tests sur les variances, les lois du Khi-deux sous les hypothèses nulles et les lois du Khi-deux excentrées sous les contre-hypothèses.

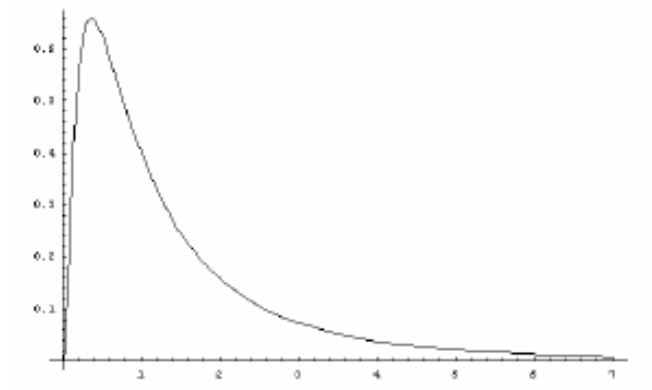


FIG. 3.13 – Loi de probabilité $\mathcal{LN}(0, 1, 0)$.

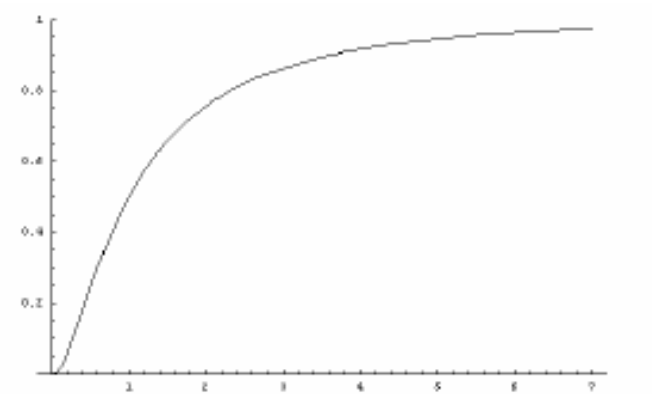


FIG. 3.14 – Fonction de répartition $\mathcal{LN}(0, 1, 0)$.

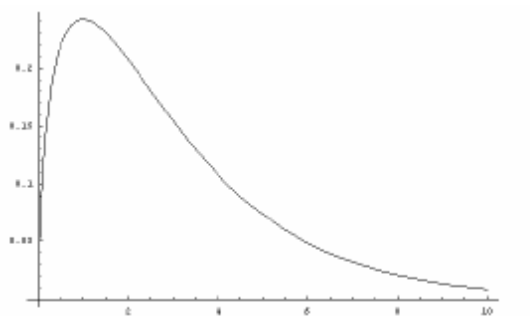


FIG. 3.15 – Loi de probabilités χ_3^2 à 3 degré de liberté.

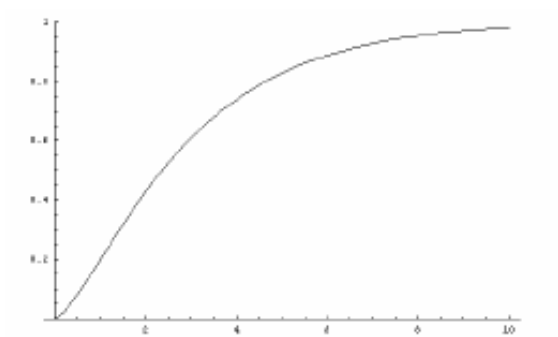


FIG. 3.16 – Fonction de répartition de χ_3^2 à 3 degré de liberté.

Propriété : Si $\mathcal{L}(X) = \chi_n^2(\delta)$ alors $E(X) = n + \delta$, $Var(X) = 2n + 4\delta$ et sa densité s'écrit :

$$f_X(t) = e^{-\frac{t+\delta}{2}} \sum_{j=0}^{+\infty} \left(\frac{\delta}{2}\right)^j \frac{t^{\frac{n}{2}+j-1}}{j! 2^{\frac{n}{2}+j} \Gamma\left(\frac{n}{2} + j\right)} I_{]0, +\infty[}(t).$$

Remarque 82 1. Une loi du Khi-deux est une loi Gamma. De même, une loi du Khi-deux excentrée est une loi Gamma excentrée.

2. Lorsque $\delta = 0$, il n'y a plus que le premier terme dans la somme et nous reconnaissons la densité de la loi de Gamma $\mathcal{GA}\left(\frac{n}{2}; \frac{1}{2}\right)$.

Propriété :

1. Si deux v.a. X, Y sont indépendantes et de loi $\mathcal{L}(X) = \mathcal{N}(\mu; 1)$ et $\mathcal{L}(X) = \chi_n^2$ alors :

$$\mathcal{L}(X^2 + Y) = \chi_{n+1}^2(\mu^2).$$

2. Si deux v.a. X_1, X_2 sont indépendantes et de loi $\mathcal{L}(X_1) = \chi_{n_1}^2(\delta_1)$ et $\mathcal{L}(X_2) = \chi_{n_2}^2(\delta_2)$ alors :

$$\mathcal{L}(X_1 + X_2) = \chi_{n_1+n_2}^2(\delta_1 + \delta_2).$$

Loi de student t_n

Le premier calcul de ce type de loi a été présenté par Helmert, ensuite Gosset a publié, en 1908, une forme générale sous le pseudonyme Student. Enfin Fisher les a utilisées dans des tests d'hypothèses sous le nom «lois de Student».

Définition 83 Soient Y, Z deux v.a. indépendantes, respectivement de loi Normale $\mathcal{L}(Y) = \mathcal{N}(\delta; 1)$ et de loi Khi-deux $\mathcal{L}(Z) = \chi_n^2$. Nous posons :

$$X = \frac{Y}{\sqrt{\frac{Z}{n}}}.$$

Si $\delta = 0$, alors la loi de X est une loi de Student à n degré de liberté. Ceci est noté $\mathcal{L}(X) = \mathcal{T}_n$. Si $\delta \neq 0$, alors la loi de X est une loi du Student non centrée à n degré de liberté et avec un paramètre de non centralité δ . Ceci est noté $\mathcal{L}(X) = \mathcal{T}_n(\delta)$.

Modélisation : Cette loi est utilisée dans l'étude des moyennes observées. En particulier pour l'estimation de la moyenne théorique d'une loi Normale. Elle intervient également dans les tests sur les moyennes de lois Normales, les lois de Student sous les hypothèses nulles et les lois de Student non centrées sous les contre-hypothèses.

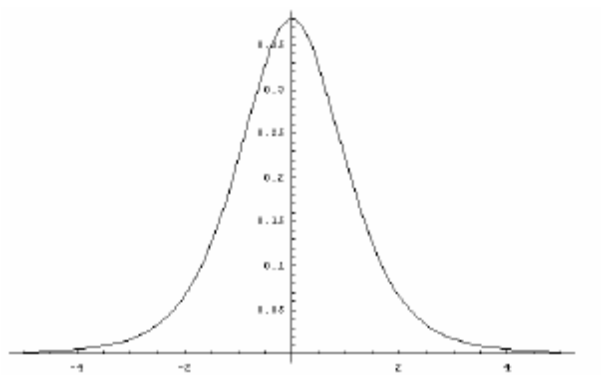


FIG. 3.17 – Loi de probabilités t_5 à 5 degrés de liberté.

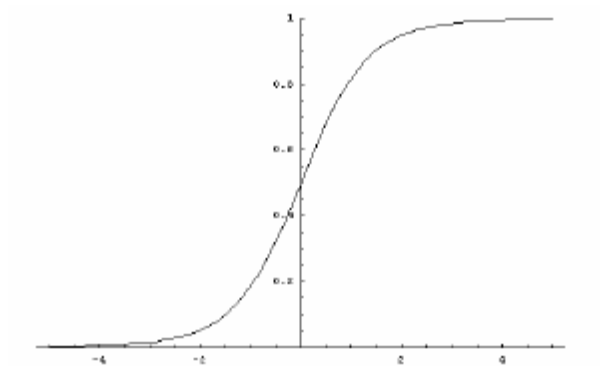


FIG. 3.18 – Fonction de répartition t_5 à 5 degré de liberté.

Propriété : Si $\mathcal{L}(X) = \mathcal{T}_n(\delta)$ alors : $E(X) = \delta \sqrt{\frac{n}{2}} \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n}{2})} \forall 1 < n, \text{Var}(X) = (1 + \delta^2) \frac{n}{n-2}, \forall 2 < n$. et sa densité s'écrit :

$$f_X(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}} \forall t \in \mathbb{R}.$$

Remarque 84 1. Lorsque $\delta = 0$, nous retrouvons le fait que la loi \mathcal{T}_n , pour $n > 1$, est centrée. De plus, toujours pour \mathcal{T}_n , avec cette fois-ci $n > 2$, nous retrouvons le fait que la variance est supérieure à 1 mais s'en rapproche lorsque $n \rightarrow +\infty$.

2. Lorsque $n = 1$, alors $\mathcal{T}_1(\delta) = \mathcal{CA}(\delta; 1)$, loi de Cauchy. (C'est cette propriété qui explique la condition d'existence de la moyenne théorique d'une loi de Student)

3. Si $\mathcal{L}(X) = \mathcal{T}_n(\delta)$ alors $\mathcal{L}(X^2) = \mathcal{F}_{1,n}(\delta)$, loi de Fisher.

Loi de Fisher-Snedecor $\mathcal{F}_{n_1, n_2}(\delta)$

La première loi dite de Fisher-Snedecor a été proposée par Fisher. Leurs applications ont été développées par Snedecor.

Définition 85 Soient X_1, X_2 deux v.a. indépendantes de loi $\mathcal{L}(X_1) = \chi_{n_1}^2(\delta)$ et $\mathcal{L}(X_2) = \chi_{n_2}^2$ respectivement. Nous posons :

$$X = \frac{\frac{X_1}{n_1}}{\frac{X_2}{n_2}}.$$

Si $\delta = 0$, alors la loi de X est une loi de Fisher-Snedecor à n_1, n_2 degrés de liberté. Nous pouvons également dire plus simplement loi de Fisher ou encore loi de Fisher et éventuellement préciser n_1 degré de liberté au numérateur et n_2 degrés de liberté au dénominateur. Ceci est noté $\mathcal{L}(X) = \mathcal{F}_{n_1, n_2}$. Si $\delta \neq 0$, alors la loi de X est une loi de Fisher - Snedecor excentrée à n_1, n_2 degrés de liberté et avec un paramètre d'excentricité. Ceci est noté $\mathcal{L}(X) = \mathcal{F}_{n_1, n_2}(\delta)$.

Propriété :

1. Si $\mathcal{L}(X) = \mathcal{F}_{n_1, n_2}(\delta)$ alors : $E(X) = \frac{n_2(n_1 + \delta)}{n_1(n_2 - 2)}, \forall n_2 > 2, \text{Var}(X) = 2 \left(\frac{n_2}{n_1}\right)^2 \frac{(n_1 + \delta)^2 + (n_1 + 2\delta)(n_2 - 2)}{(n_2 - 2)^2(n_2 - 4)}, \forall n_2 > 4$ et sa densité s'écrit :

$$f_X(t) = e^{-\frac{\delta}{2}} \sum_{j=0}^{+\infty} \frac{\delta^j}{j! 2^j} \frac{n_1^{j + \frac{n_1}{2}} t^{j-1 - \frac{n_1}{2}}}{B(j + \frac{n_1}{2}; \frac{n_2}{2}) (n_1 + tn_2)^{j + \frac{n_1 + n_2}{2}}} I_{]0, +\infty[}(t),$$

où $B(n_1, n_2)$ est la fonction Bêta.

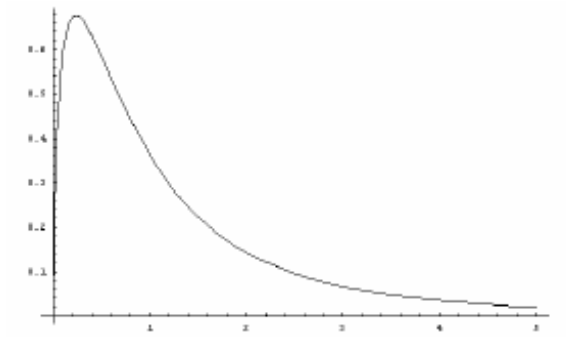


FIG. 3.19 – Loi de probabilités $F_{3,5}$ à 3 et 5 degré de liberté.

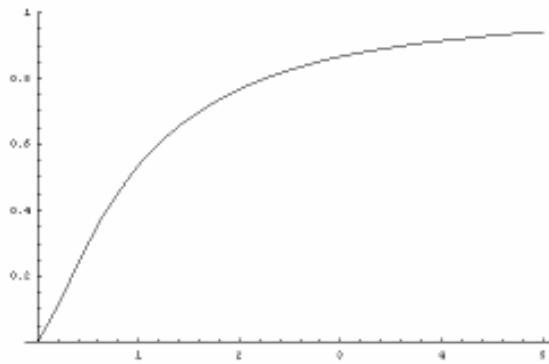


FIG. 3.20 – Fonction de répartition $F_{3,5}$ à 3 et 5 degré de liberté.

Modélisation : Ces lois sont utilisées dans la comparaison de variances. En particulier pour l'estimation du rapport de deux variances théoriques de deux lois Normales. Elles interviennent également dans les tests sur les variances de lois Normales : les lois de Fisher - Snedecor sous les hypothèses nulles et les lois de Fisher - Snedecor excentrées sous les contre-hypothèses (cf. la partie du site concernant ces tests). Mais la plus importante utilisation des ces lois concerne les modèles linéaires normaux et les analyses de variance en particulier.

Remarque 86 1. Si $\mathcal{L}(X) = \mathcal{F}_{n_1, n_2}$ alors $\mathcal{L}\left(\frac{1}{X}\right) = \mathcal{F}_{n_2, n_1}(\delta)$.

2. Lorsque $\delta = 0$, nous avons une moyenne théorique qui ne dépend pas du degré de liberté du numérateur.

3. Lorsque $\delta = 0$, il n'y a plus que le premier terme dans la somme et nous avons la densité de la loi \mathcal{F}_{n_1, n_2} .

Loi de Gamma, $\mathcal{GA}(\alpha; \beta)$, $\alpha, \beta \in \mathbb{R}_+^*$

Nous présentons la famille de lois Gamma ou d'Euler très utiles pour les propriétés de décroissance rapide de leur fonction de survie. Pour cela nous introduisons une fonction spéciale.

Définition 87 Nous appelons fonction Gamma la fonction définie par :

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1} e^{-t} dt, \forall \alpha \in \mathbb{R}_+^*.$$

Proposition 88 La fonction Gamma satisfait aux relations suivantes :

$$\Gamma(\alpha + 1) = \alpha \Gamma(\alpha), \forall \alpha \in \mathbb{R}_+^*,$$

$$\Gamma(n + 1) = n!, \forall n \in \mathbb{N},$$

$$\Gamma(0.5) = \pi.$$

Définition 89 Une v.a. X suit une loi Gamma ou eulerienne de paramètres $\mathcal{GA}(\alpha; \beta)$, $\alpha, \beta \in \mathbb{R}_+^*$ si elle admet pour densité de probabilité la fonction :

$$f_X(t) = \begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} & \text{si } t > 0, \\ 0 & \text{sinon.} \end{cases}$$

Ceci est noté $\mathcal{L}(X) = \mathcal{GA}(\alpha; \beta)$. La loi $\mathcal{GA}(1; \beta) = \mathcal{E}(\beta)$ est appelée loi exponentielle de paramètre β .

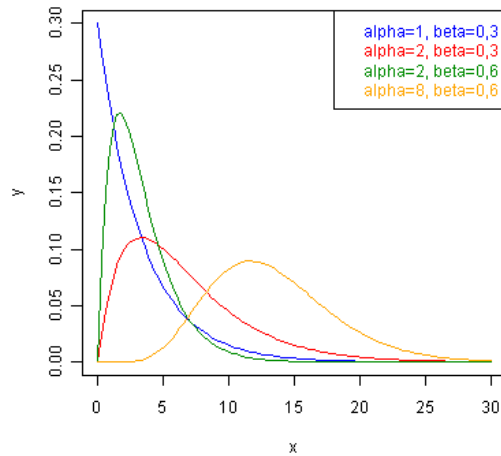


FIG. 3.21 – Densités Gamma.

Propriété : Si $\mathcal{L}(X) = \mathcal{GA}(\alpha; \beta)$ alors : $E(X) = \frac{\alpha}{\beta}$, $Var(X) = \frac{\alpha}{\beta^2}$.

Modélisation : La loi Gamma peut décrire des phénomènes de durée de vie, en assurance pour l'étude du temps écoulé entre deux sinistres dans des portefeuilles à risques hétérogènes ou encore pour prendre en compte cette hétérogénéité. En général pour des distributions fortement asymétriques avec une décroissance rapide en queue de distribution, une loi Gamma peut être un bon modèle.

Remarque 90 1. A l'exception des lois exponentielles, il n'existe pas d'expression explicite pour la fonction de répartition $F_X(t) = P(X \leq t)$.

2. La loi $\mathcal{GA}(\frac{n}{2}; \frac{1}{2}) = \chi_n^2$ est appelée loi du Khi-deux.

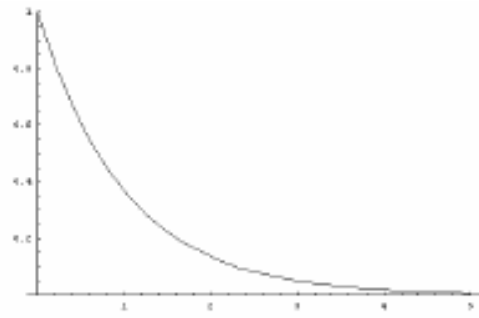
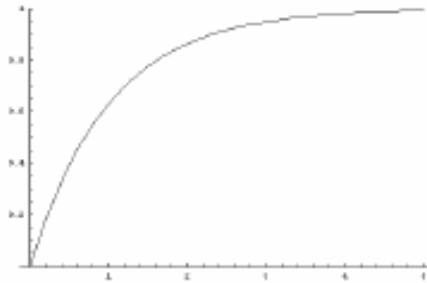
3. Soit $X_i, i = 1, \dots, n$, des variables aléatoires indépendantes de loi $\mathcal{GA}(\alpha_i; \beta)$ respectivement.

Alors pour tout nombre $c \in \mathbb{R}_+^*$ nous avons :

$$\mathcal{L}(c(X_1 + \dots + X_n)) = \mathcal{GA}\left(\alpha_1 + \dots + \alpha_n; \frac{\beta}{c}\right).$$

Loi Exponentielle, $\mathcal{E}(\lambda)$, avec $\lambda > 0$ un réel.

Cette densité de probabilité permet en général de modéliser des durées de vie d'êtres non soumis au vieillissement (par exemple, la durée de vie d'une bactérie) ou des temps d'attente (par exemple, le temps d'attente entre deux signaux synaptiques).

FIG. 3.22 – Loi de probabilités $\mathcal{E}(1)$.FIG. 3.23 – Fonction de répartition $\mathcal{E}(1)$.

On dit que la v.a.r suit la loi exponentielle de paramètre λ lorsque X admet pour densité de probabilité la fonction

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

On peut calculer la fonction de répartition $F_X(x)$

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}, \text{var}(X) = \frac{1}{\lambda^2}.$$

La loi exponentielle est la seule loi continue qui vérifie la propriété d'absence de mémoire :

Si $X \sim \mathcal{E}(\lambda)$, alors pour tout $s, t > 0$ $P(X > t + s | X > t) = P(X > s)$.

La loi exponentielle est celle de la mortalité des êtres qui ne seraient pas soumis au vieillissement : à chaque moment ils ont la même probabilité de mourir dans l'unité de temps qu'il leur reste,

quelque soit leur âge.

Loi de Bêta $\mathcal{BE}(\alpha; \beta)$, $\alpha, \beta \in \mathbb{R}_+^*$

Nous présentons la famille de lois Bêta liées aux lois Gamma. Pour cela nous introduisons une fonction spéciale.

Définition 91 Nous appelons fonction Bêta incomplète la fonction définie par :

$$B(t; \alpha; \beta) = \int_0^t u^{\alpha-1} (1-u)^{\beta-1} du, \forall t \in]0, 1[, \alpha, \beta \in \mathbb{R}_+^*.$$

Nous notons $B(\alpha; \beta) = B(1; \alpha; \beta)$ et nous l'appelons fonction Bêta.

Propriété : La fonction Bêta est liée à la fonction Gamma de la manière suivante :

$$B(\alpha; \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)}.$$

Définition 92 Une v.a. X suit une loi Bêta de paramètre $\alpha, \beta \in \mathbb{R}_+^*$, si elle admet pour densité de probabilité la fonction :

$$f_x(t) = \begin{cases} \frac{1}{B(\alpha; \beta)} t^{\alpha-1} (1-t)^{\beta-1} & \text{si } t \in]0, 1[, \\ 0 & \text{sinon.} \end{cases}$$

Ceci est noté $\mathcal{L}(X) = \mathcal{BE}(\alpha; \beta)$. La loi $\mathcal{BE}(\alpha; \beta) = \mathcal{U}(]0, 1[)$ est la loi Uniforme sur l'intervalle $]0, 1[$.

Propriété : Si $\mathcal{L}(X) = \mathcal{BE}(\alpha; \beta)$ alors : $E(X) = \frac{\alpha}{\alpha + \beta}$, $Var(X) = \frac{\alpha\beta}{(\alpha + \beta)^2(1 + \alpha + \beta)}$.

Modélisation : La loi Bêta, outre ses aspects théoriques, présente un intérêt certain. Comme le montre le graphique ci-dessous, elle admet différentes asymétries sur l'intervalle $]0, 1[$, selon les valeurs des paramètres. Elle apparaît également lors de l'observation de lois de Bernoulli pour modéliser la variation de leurs paramètres lorsque ces derniers sont aléatoires. Enfin les lois Bêta apparaissent dans l'étude des statistiques d'ordre.

Remarque 93 1. Si $\mathcal{L}(X) = \mathcal{BE}(\alpha; \beta)$ alors la fonction de répartition $F_X(t) = P(X \leq t)$ est la fonction Bêta incomplète normalisée sur l'intervalle $]0, 1[$, de plus :

$$F_X(t) = \begin{cases} 0 & \text{si } t \leq 0, \\ \frac{B(t; \alpha; \beta)}{B(\alpha; \beta)} & \text{si } t \in]0, 1[, \\ 1 & \text{si } t \geq 1. \end{cases}$$

Il est possible, lorsque sont des nombres entiers, d'écrire la fonction de répartition sous forme de somme.

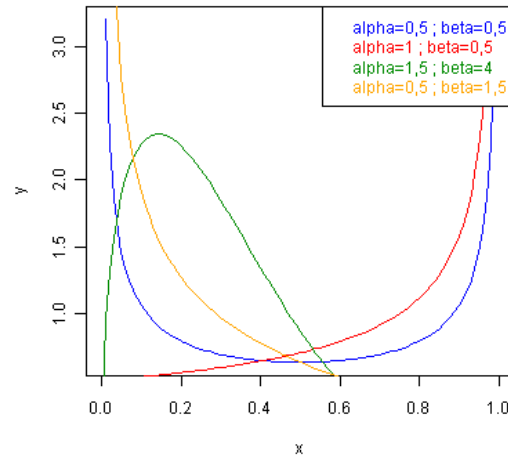


FIG. 3.24 – Densités Bêta.

2. Si $\mathcal{L}(X) = \mathcal{BE}(\alpha; \beta)$ alors $\mathcal{L}(1 - X) = \mathcal{BE}(\beta; \alpha)$.

3. Soit deux v.a. X, Y indépendantes et telles que $\mathcal{L}(X) = \mathcal{GA}(\alpha_1; \beta)$ et $\mathcal{L}(Y) = \mathcal{GA}(\alpha_2; \beta)$, alors

$$\mathcal{L}\left(\frac{X}{X+Y}\right) = \mathcal{BE}(\alpha_1; \alpha_2).$$

Loi de Weibull, $W(a, b)$, avec $a > 0, b > 0$ réels.

La loi de Weibull est utilisée en démographie pour modéliser le vieillissement et en épidémiologie pour modéliser la distribution de probabilité de la durée d'incubation d'une maladie infectieuse.

Densité :

$$f(x) = \frac{a}{b} \left(\frac{x}{b}\right)^{a-1} \exp\left(-\left(\frac{x}{b}\right)^a\right) \text{ avec } x > 0.$$

Fonction de répartition :

$$F(x) = 1 - \exp\left(-\left(\frac{x}{b}\right)^a\right) \text{ avec } x > 0.$$

On constate que pour $a = 1$, on retrouve la loi exponentielle. La loi exponentielle est celle de la mortalité des être vivants qui ne seraient pas soumis au vieillissement : à chaque moment ils ont la même probabilité de mourir dans l'unité de temps qu'il leur reste, quelque soit leur âge. Plus a est grand plus le vieillissement se fait pesant (la mortalité augmente avec l'âge). Le cas $a < 1$ correspond à un monde dans lequel plus on vieillirait, moins forte serait la probabilité de mourir dans l'unité de temps qui vient.

Loi de Cauchy $\mathcal{CA}(x_0; \alpha)$, $\alpha \in \mathbb{R}_+^*$ et $x_0 \in \mathbb{R}$

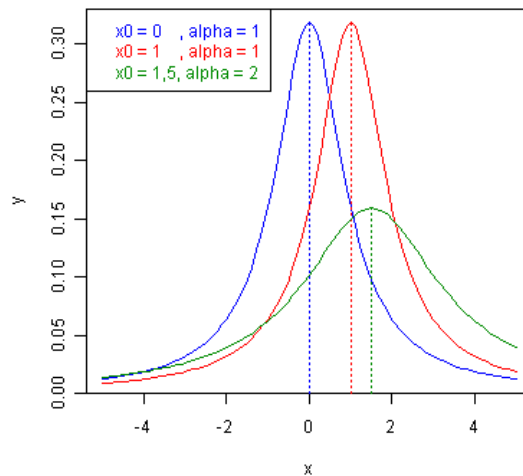
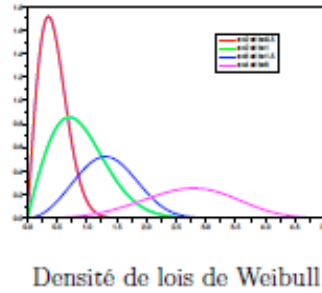


FIG. 3.25 – Densité de Cauchy.

Les lois de Cauchy sont intéressantes par les aspects théoriques particuliers qu'elles présentent.

Définition 94 Une v.a. continue X suit une loi de Cauchy, de paramètres $\alpha \in \mathbb{R}_+^*$ (paramètre d'échelle) et $x_0 \in \mathbb{R}$ (paramètre de position) si elle admet pour densité de probabilité la fonction :

$$f_X(t) = \frac{1}{\pi} \frac{\alpha}{\alpha^2 + (t - x_0)^2}, t \in \mathbb{R}.$$

Ceci est noté $\mathcal{L}(X) = \mathcal{CA}(x_0; \alpha)$.

Propriété : Si $\mathcal{L}(X) = \mathcal{CA}(x_0; \alpha)$ alors la v.a. X ne possède ni moyenne théorique ni a fortiori variance théorique. Sa fonction de répartition est donnée par :

$$F_X(t) = \frac{1}{2} + \frac{1}{\pi} \operatorname{arctg} \left(\frac{t - x_0}{\alpha} \right).$$

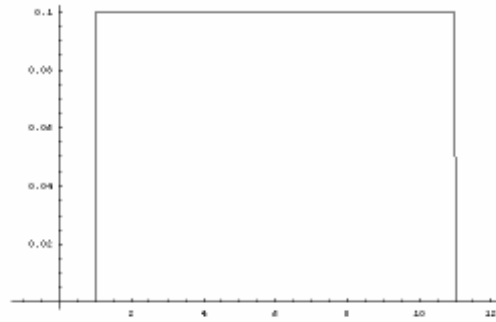


FIG. 3.26 – Loi de probabilités $\mathcal{U}([1, 11])$.

Remarque 95 Si deux v.a. X et Y sont indépendantes et de loi $\mathcal{N}(0; 1)$ alors : $\mathcal{L}\left(\frac{X}{Y}\right) = \mathcal{CA}(0; 1)$.

Loi Uniforme, $\mathcal{U}([a, b])$, avec $a, b \in \mathbb{R}$, $a < b$.

Cette loi est l'analogie continu de l'équiprobabilité dans le cas discret. Elle permet de modéliser le tirage d'un nombre aléatoire dans l'intervalle $[a, b]$.

On dit que la v.a.r suit la loi uniforme sur $[a, b]$ lorsque X admet pour densité de probabilité la fonction

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b[\\ 0 & \text{sinon} \end{cases}$$

Fonction de répartition :

$$F_X(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a} & \text{si } x \in [a, b] \\ 1 & \text{si } x > b \end{cases}$$

$$E(X) = \frac{a+b}{2}, \text{var}(X) = \frac{(b-a)^2}{12}.$$

Exemple 96 Lors d'une étude du comportement animal, on a relâché des oiseaux dont l'orientation a été rendue très difficile. On s'attend alors à ce que les oiseaux choisissent au hasard leur direction. On peut modéliser la direction prise par un oiseau de la façon suivante. On considère X l'angle entre le nord et la direction prise par l'oiseau (selon le sens des aiguilles d'une montre). La variable X suit une loi uniforme entre 0 et 360 degrés ($X \sim \mathcal{U}([0, 360])$) et

$$f(x) = \begin{cases} \frac{1}{360} & \text{si } x \in [0, 360[\\ 0 & \text{sinon} \end{cases}, F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x}{360} & \text{si } x \in [0, 360] \\ 1 & \text{si } x > 360 \end{cases}.$$

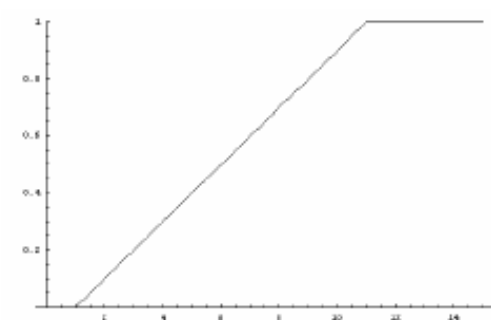


FIG. 3.27 – Fonction de répartition $\mathcal{U}([1, 11])$.

3.2.1 La convergence en loi

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a.r définies sur un espace probabilisé $(\Omega, \mathcal{P}(\Omega), p)$, F_n leurs fonctions de répartition, et X une v.a.r définie sur ce même espace, de fonction de répartition F . On dit que la suite (X_n) converge en loi vers X si et seulement si : en tout point x où F est continue on a $\lim_{n \rightarrow \infty} F_n(x) = F(x)$.

3.2.2 Approximation d'une loi hypergéométrique par une loi binomiale

Reprenons, l'exemple des poissons. Lorsque la taille N de la population de poissons dans le lac est très grande, le fait d'enlever 1 poisson ne changera quasiment pas la proportion $p = \frac{m}{N}$ de poissons marqués.

De même, si on en enlève un nombre fini n . Par conséquent, chaque poisson pêché a peu près la même probabilité p d'être marqué.

On peut donc approcher la loi Hypergéométrique par la loi binomiale $B(n, p)$ où p est la proportion de "poissons marqués".

De manière générale, lorsque la taille de la population est grande, on utilise souvent la binomiale même pour un "sondage sans remise".

Proposition :

La loi binomiale approche la loi hypergéométrique lorsque N tend vers l'infini.

Preuve. On a

$$\begin{aligned}
 P(X = k) &= \frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \\
 &= \frac{m!}{k! (m-k)!} \frac{(N-m)!}{(n-k)! (N-m-n+k)!} \frac{n! (N-n)!}{N!} \\
 &= C_n^k \left(\frac{m}{N} \frac{m-1}{N-1} \cdots \frac{m-k+1}{N-k+1} \right) \left(\frac{N-m}{N-k} \frac{N-m-1}{N-k-1} \right. \\
 &\quad \left. \cdots \frac{N-m-n+k+1}{N-n-1} \right)
 \end{aligned}$$

si on suppose maintenant que $N \rightarrow \infty$ et que $\frac{m}{N} \rightarrow p$, on a

$$\begin{aligned}
 \frac{m}{N} &\rightarrow p, \quad \frac{m-1}{N-1} = \frac{\frac{m}{N} - \frac{1}{N}}{1 - \frac{1}{N}} \rightarrow p, \dots, \\
 \frac{m-k+1}{N-k+1} &= \frac{\frac{m}{N} - \frac{k-1}{N}}{1 - \frac{k-1}{N}} \rightarrow p; \\
 \frac{N-m}{N-k} &= \frac{1 - \frac{m}{N}}{1 - \frac{k}{N}} \rightarrow 1-p, \quad \frac{N-m-1}{N-k-1} \rightarrow 1-p, \dots, \\
 \frac{N-m-n+k+1}{N-n-1} &\rightarrow 1-p
 \end{aligned}$$

En conclusion : si $N \rightarrow \infty$ et que $\frac{m}{N} \rightarrow p$, on a : $\frac{C_m^k C_{N-m}^{n-k}}{C_N^n} \rightarrow C_n^k p^k (1-p)^{n-k}$.

Dès que $N \geq 10n$, on pourra dire que l'on peut approcher la loi hypergéométrique par la loi binomiale.

Soit X une v.a.r suivant la loi hypergéométrique $H(100, 5, 4)$. Nous allons calculer $P(X \geq 1)$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \frac{C_5^0 C_{95}^4}{C_{100}^4} \simeq 0,188$$

Calcul approché : on approche la loi $H(100, 5, 4)$ par la loi $\mathcal{B}(4, 0.05)$

$$P(X \geq 1) = 1 - P(X = 0) = 1 - C_4^0 (0.05)^0 (0.95)^4 \simeq 0,185.$$

3.2.3 Approximation d'une loi binomiale par une loi de Poisson

Proposition : Une loi de Poisson de paramètre $\lambda > 0$ peut être obtenue comme une limite de loi binomiale $\mathcal{B}(n; p_n)$, où la suite (p_n) est supposée positive, décroissante, tend vers zéro et telle que : $\lim_{n \rightarrow \infty} np_n = \lambda$.

Preuve. On a

$$P(X_n = k) = C_n^k p_n^k (1 - p_n)^{n-k} = \frac{n!}{(n-k)!n^k} \frac{(np_n)^k}{k!} (1 - p_n)^{n-k}$$

comme

$$\begin{aligned} \frac{n!}{(n-k)!n^k} &= \frac{n}{n} \frac{n-1}{n} \frac{n-2}{n} \dots \frac{n-k+1}{n} \\ &= 1 \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \end{aligned}$$

Donc $\lim \frac{n!}{(n-k)!n^k} = 1$. On écrira d'abord : $(1 - p_n)^{n-k} = \exp\{(n-k) \log(1 - p_n)\}$. Quand n tend vers l'infini, p_n tend vers 0 et donc $\log(1 - p_n) \sim -p_n$.

Par conséquent : $(n-k) \log(1 - p_n) \sim -(n-k)p_n \sim -\lambda$.

On a donc : $\lim_n (n-k) \log(1 - p_n) = -\lambda$.

Et enfin $\lim_n (1 - p_n)^{n-k} = e^{-\lambda}$.

En conclusion, lorsque n tend vers l'infini, la loi binomiale $\mathcal{B}(n, p)$ converge vers une loi de poisson $\mathcal{P}(\lambda)$ puisque :

$$P(X_n = k) \xrightarrow[n \rightarrow \infty]{} \frac{\lambda^k}{k!} e^{-\lambda}.$$

Si X suit une Binomiale $\mathcal{B}(n, p)$, avec n grand ($n \geq 30$) et p petit ($p \leq 0,1$). Alors pour tout $k \geq 0$; $P(X = k) \simeq P(Y = k)$; où Y suit une loi de Poisson de paramètre $\lambda = np$.

Soit X une v.a.r suivant la loi binomiale $\mathcal{B}(100, 0.05)$. Nous allons calculer $P(X = 2)$

$$P(X = 2) = C_{100}^2 (0.05)^2 (0.95)^{98} \simeq 0,0812$$

Calcul approché : on approche la loi $\mathcal{B}(100, 0.05)$ par la loi $\mathcal{P}(5)$

$$P(X = 2) = e^{-5} \times \frac{5^2}{2!} \simeq 0,0843.$$

3.2.4 Approximation d'une loi binomiale par une loi normale

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de v.a.r telle que X_n suit la loi binomiale $\mathcal{B}(n, p)$, alors la suite aléatoire $\frac{X_n - np}{\sqrt{npq}}$ converge en loi vers la loi normale centrée réduite $\mathcal{N}(0, 1)$. (Théorème de Moivre-Laplace)

En pratique lorsque n est grand ($n \geq 30$), si $np \geq 5$ et $n(1 - p) \geq 5$, alors on constate que la distribution binomiale tend vers la distribution normale de moyenne np et de variance $np(1 - p)$ (i.e. $\mathcal{B}(np, npq)$).

Soit X une variable qui suit la loi $\mathcal{B}(900, 0.5)$. On cherche à calculer $P(405 \leq X \leq 495)$.

Pour le calcul exact, il nous faudrait calculer des combinaisons avec de très grand nombres, ce qui nécessite un ordinateur et ne donne parfois qu'une valeur approchée.

On remarque ici que l'on est dans les conditions où l'on peut approximer notre loi binomiale par une loi normale $\mathcal{N}(450, 225)$ car $900 > 30$ et $900 \times 0,5 = 450 > 5$.

Cependant pour la loi normale $\mathcal{N}(450, 225)$ il n'est toujours pas facile de calculer $P(405 \leq X \leq 495)$ car on ne connaît pas la fonction de répartition d'une loi normale quelconque. Nous sommes donc obligés de nous ramener à la loi normale centrée réduite en nous intéressant à la variable centrée réduite associée à X .

On a

$$\begin{aligned} P(405 \leq X \leq 495) &= P\left(\frac{405 - 450}{\sqrt{225}} \leq \frac{X - 450}{\sqrt{225}} \leq \frac{495 - 450}{\sqrt{225}}\right) \\ &= P\left(-3 \leq \frac{X - 450}{15} \leq 3\right) \\ &\simeq \Phi(3) - \Phi(-3) \\ &\simeq \Phi(3) - (1 - \Phi(3)) \\ &\simeq 2\Phi(3) - 1 \\ &\simeq 0.9974 \end{aligned}$$

3.2.5 Approximation d'une loi de Poisson par une loi normale

Lorsque son paramètre λ est grand (en pratique $\lambda \geq 18$), une loi de Poisson peut être approchée par une loi normale d'espérance λ et de variance λ .

Le principe est analogue à celui utilisé pour l'approximation de la loi binomiale par la loi normale.

Si on considère X qui suit une loi de Poisson de paramètre 64 que l'on cherche à calculer $P(X \leq 74)$, on a intérêt à approcher la loi de X par la loi $\mathcal{N}(64, 64)$ et donc $\frac{X-64}{8}$ suit la loi normale centrée réduite. On a donc

$$P(X \leq 74) = P\left(\frac{X - 64}{8} \leq 1.25\right) \simeq \Phi(1.25) \simeq 0.8944.$$

Exemple 97 *On effectue un contrôle de fabrication sur des pièces dont une proportion $p = 0.02$ est défectueuse.*

On contrôle un lot de 1000 pièces : Soit X la variable aléatoire : «nombre de pièces défectueuses parmi 1000». Quelle est la vraie loi de X ? (on ne donnera que la forme générale); quel est son espérance, son écart-type ?

En approchant cette loi par celle d'une loi normale adaptée, calculez la probabilité pour que X soit compris entre 18 et 22 ($P(18 \leq X \leq 22)$); on fera les calculs avec et sans correction de continuité. On fera également les calculs avec la vraie loi pour comparer.

Exemple 98

La loi de X est la loi binomiale $B(1000; 0.02)$, d'espérance $E(X) = n.p = 1000 \times 0.02 = 20$, d'écart-type $\sqrt{19.6}$ $\text{var}(X) = n.p.q = 1000 \times 0.02 \times 0.98 = 19.6$.

En approchant cette loi par celle d'une loi normale de paramètre $m = 20$, écart-type $\sqrt{19.6}$.

$$P(18 \leq X \leq 22) = P\left(\frac{17.5 - 20}{\sqrt{19.6}} \leq \frac{X - 20}{\sqrt{19.6}} \leq \frac{22.5 - 20}{\sqrt{19.6}}\right) \simeq 0.424.$$

Sans correction de continuité on trouve

$$P(18 \leq X \leq 22) = P\left(\frac{18 - 20}{\sqrt{19.6}} \leq \frac{X - 20}{\sqrt{19.6}} \leq \frac{22 - 20}{\sqrt{19.6}}\right) \simeq 0.346.$$

Approchée par la loi de Poisson de paramètres : espérance 20 et variance 20, on trouve $P(18 \leq X \leq 22) \simeq 0.423$.

Enfin par la vraie loi binomiale : on trouve $P(18 \leq X \leq 22) \simeq 0.427$.

3.3 Exercices

Exercice 99 Un QCM comporte 5 affirmations. Pour chaque affirmation, on doit répondre par vrai (V) si l'affirmation est toujours vraie, faux (F) si elle est toujours fausse ou par (P) si on ne peut pas conclure. Une réponse au QCM est une suite de 5 lettres parmi V , F ou P .

1. Quel est le nombre de réponses possibles pour le QCM ?
2. Un candidat répond au hasard et on appelle X la variable aléatoire donnant le nombre de bonnes réponses. Indiquer la loi suivie par X et préciser ses paramètres et calculer $E(X)$.

Exercice 100 Un jeu de hasard est formé d'un dispositif lançant de façon aléatoire une fléchette dans une cible ayant la forme suivante

B	B	B	B	B	B	B	B	B	J	J	J	V	V	R
R	V	V	J	J	J	B	B	B	B	B	B	B	B	B

La fléchette atteint toujours une case et une seule. Les 30 cases blanches (B), jaunes (J), vertes (V) ou rouges (R) ont toutes la même probabilité d'être atteintes. Si la fléchette atteint une case rouge, le joueur gagne 8 DA, si la fléchette atteint une case verte, le joueur gagne 5 DA, si la fléchette atteint une case jaune, le joueur gagne rien et ne perd rien, si la fléchette atteint une case blanche, le joueur perd a DA, la lettre a désignant un réel positif.

1. On note X la variable aléatoire représentant le gain algébrique du joueur (compté négativement quand il perd).
 - a) Donner la loi de probabilité de X .
 - b) Calculer a pour que le jeu soit équitable (c'est-à-dire pour que $E(X)$ soit nulle).
2. Un joueur est considéré comme gagnant s'il a obtenu un gain strictement positif.
 - a) Quelle est la probabilité p qu'un joueur gagne ?
 - b) Un joueur joue 5 parties consécutives indépendantes. Quelle est la probabilité qu'il gagne exactement 2 fois ? exactement 5 fois ? Quel est le nombre moyen de parties gagnantes ?

Exercice 101 On prend au hasard, en même temps, trois ampoules dans un lot de 15 dont 5 sont défectueuses. Calculer la probabilité des événements :

- a) Au moins une ampoule est défectueuse ;
- b) Les 3 ampoules sont défectueuses ;

c) *Exactement une ampoule est défectueuse.*

Exercice 102 *L'oral d'un concours comporte au total 100 sujets ; les candidats tirent au sort trois sujets et choisissent alors le sujet traité parmi ces trois sujets. Un candidat se présente en ayant révisé 60 sujets sur les 100.*

1. *Quelle est la probabilité pour que le candidat ait révisé :*

- a) *Les trois sujets tirés ;*
- b) *Exactement deux sujets sur les trois sujets ;*
- c) *Aucun des trois sujets.*

2. *Définir une variable aléatoire associée à ce problème et donner sa loi de probabilité, son espérance.*

Exercice 103 *Une population comporte en moyenne une personne mesurant plus de 1.90 m sur 80 personnes. Sur 100 personnes, calculer la probabilité qu'il y ait au moins une personne mesurant plus de 1.90 m. Sur 300 personnes, calculer la probabilité qu'il y ait au moins une personne mesurant plus de 1.90 m.*

Exercice 104 *Soit X une variable aléatoire qui suit la loi uniforme sur $[-5, 15]$. Calculer : $P(X \leq 2)$, $P(-1 \leq X \leq 1)$, $P_{(X>0)}(-1 \leq X \leq 2)$. Soit Y la variable aléatoire égale à $\frac{X+5}{10}$. Calculer $P_{(X \leq 10)}(Y \geq 1)$.*

Exercice 105 *Soit X une variable aléatoire suivant la loi normale $\mathcal{N}(200; 15^2)$.*

- *Déterminer le réel $u > 0$ tel que $P(200 - 2u \leq X \leq 200 + 2u) = 0,9$.*

Exercice 106 *Une enquête est effectuée auprès de familles de 4 personnes afin de connaître leur achat de lait en 1 mois. Sur l'ensemble des personnes interrogées, la consommation de ce produit forme une population gaussienne avec une moyenne de 20L et un écart-type de 6.*

En vue d'une conception d'une campagne de pub, on souhaite connaître le pourcentage des faibles consommateurs (c'est-à-dire moins de 10 L/mois) et le pourcentage des grands consommateurs (c'est-à-dire plus de 30 L/mois).

1. *Calculer ces 2 pourcentages.*
2. *Au-dessous de quel nombre de litres achetés se trouvent 75% des consommateurs ?*
3. *Combien de litres au maximum consomme la moitié des consommateurs ?*
4. *Au-dessus de quelle consommation se trouve 1/3 de la population ? et 2/3 de la population ?*