



DÉPARTEMENT STPI

TROISIÈME ANNÉE IC

# Probabilités et Statistiques

C. Boyer, S. Grusea et C. Marteau

Année universitaire 2012-2013

---

Simona Grusea  
Bureau 11, département GMM  
simona.grusea@insa-toulouse.fr

Clément Marteau  
Bureau 123, département GMM  
clement.marteau@insa-toulouse.fr

# Table des matières

<b>I</b>	<b>Introduction à la théorie des probabilités</b>	<b>5</b>
I.1	Pourquoi les probabilités ?	5
I.2	Espace probabilisé	6
I.2.1	Définitions	6
I.2.2	Opérations sur les événements	7
I.2.3	Probabilité	8
I.2.4	Probabilité uniforme	8
<b>II</b>	<b>Probabilités conditionnelles et indépendance</b>	<b>11</b>
II.1	Probabilités conditionnelles	11
II.2	Indépendance	12
<b>III</b>	<b>Variables aléatoires réelles</b>	<b>15</b>
III.1	Introduction	15
III.2	Définitions	15
III.2.1	Variable aléatoire	15
III.2.2	Loi de probabilité	15
III.2.3	Fonction de répartition	16
III.3	Différents types de variables aléatoires réelles	17
III.3.1	Variables aléatoires discrètes	17
III.3.2	Variables aléatoires continues	18
III.4	Exemples fondamentaux	19
III.4.1	Variables discrètes	19
a)	Loi de Bernoulli	19
b)	Loi binomiale	19
c)	Loi géométrique	19
d)	Loi de Poisson	19
III.4.2	Variables continues	19
a)	Loi exponentielle	20
b)	Loi Gamma	20
c)	Loi normale de paramètres $(m, \sigma^2)$	20
<b>IV</b>	<b>Manipulation de variables aléatoires</b>	<b>23</b>
IV.1	Caractéristiques des variables aléatoires	23
IV.1.1	Espérance	23
IV.1.2	Propriétés de l'espérance	24

IV.1.3	Variance et écart-type . . . . .	24
IV.2	Indépendance de variables aléatoires . . . . .	25
<b>V</b>	<b>Théorèmes limites</b>	<b>27</b>
V.1	Inégalité de Chebychev . . . . .	27
V.2	Loi (faible) des grands nombres . . . . .	28
V.3	Théorème central limite . . . . .	28
V.4	Approximation d'une loi binomiale . . . . .	29
V.4.1	Approximation par une loi normale . . . . .	29
V.4.2	Approximation par une loi de Poisson . . . . .	30
<b>VI</b>	<b>Estimation</b>	<b>31</b>
VI.1	Introduction . . . . .	31
VI.2	Estimation ponctuelle . . . . .	32
VI.2.1	Propriétés d'un (bon) estimateur . . . . .	32
a)	Consistance . . . . .	32
b)	Biais . . . . .	33
c)	Écart quadratique moyen . . . . .	33
VI.2.2	Estimateur d'une moyenne $m$ ou d'une proportion . . . . .	33
VI.2.3	Estimateur de la variance . . . . .	34
VI.3	Estimation par intervalle de confiance . . . . .	34
VI.3.1	Intervalle de confiance pour la moyenne d'un échantillon gaussien (variance connue) . . . . .	35
VI.3.2	Intervalle de confiance pour la moyenne d'un échantillon gaussien (variance inconnue) . . . . .	35
VI.3.3	Intervalle de confiance pour une proportion . . . . .	36
VI.3.4	Intervalle de confiance pour la variance d'une population gaussienne	37
<b>VII</b>	<b>Tests d'hypothèses</b>	<b>39</b>
VII.1	Exemple introductif . . . . .	39
VII.2	Démarche générale pour la construction d'un test . . . . .	40
VII.3	Test sur la moyenne d'un échantillon gaussien (variance connue) . . . . .	41
VII.3.1	Test unilatéral à droite . . . . .	42
VII.3.2	Test unilatéral à gauche . . . . .	42
VII.3.3	Test bilatéral . . . . .	42
VII.4	Test sur la moyenne d'un échantillon gaussien (variance inconnue) . . . . .	42
VII.5	Cas d'une population non gaussienne . . . . .	42
<b>VIII</b>	<b>Annexe 1: Tables statistiques</b>	<b>43</b>
<b>IX</b>	<b>Annexe 2: Sujets des années précédentes</b>	<b>46</b>

# Chapitre I

## Introduction à la théorie des probabilités

### I.1 Pourquoi les probabilités ?

Dans de nombreux domaines (science, sociologie, médecine, etc...), on s'intéresse à des phénomènes dans lesquels apparaît l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre.

Une expérience est appelée "aléatoire" s'il est impossible de prévoir à l'avance son résultat et si, répétée dans des conditions identiques, elle peut donner des résultats différents:

- succession d'appels à un standard téléphonique non surchargé;
- observation de la durée de vie d'un individu anonyme dans une population;
- observation de la durée de fonctionnement sans panne d'appareil;
- jeu de pile ou face.

Voici d'autres exemples de domaines d'applications des probabilités.

**Fiabilité.** On considère un système formé par plusieurs composants. On s'intéresse à la fiabilité du système: on va chercher à calculer la probabilité que le système fonctionne encore à un instant donné. Il faut pour cela connaître la probabilité que chacun des composants fonctionne à cet instant et tenir compte du fait que les composants ne fonctionnent peut-être pas indépendamment les uns des autres.

**Fatigue des matériaux.** Les données de fatigue des matériaux sont très dispersées. On fait alors appel à des modélisations probabilistes et à des méthodes statistiques afin, par exemple, de construire des intervalles de confiance pour le nombre moyen de cycles jusqu'à la rupture.

**Télécommunications.** En télécommunications, on doit souvent tenir compte du “bruit” dans les systèmes. Par exemple, supposons qu’un système émet soit un 0, soit un 1, et qu’il y a un risque  $p$  que le chiffre émis soit mal reçu. Il est alors intéressant de calculer la probabilité qu’un 0 ait été émis, sachant qu’un 0 a été reçu, ou encore la probabilité qu’il y ait une erreur de transmission.

## I.2 Espace probabilisé

### I.2.1 Définitions

**Définition 1** On appelle *univers* associé à une expérience aléatoire l’ensemble  $\Omega$  de tous les résultats possibles de cette expérience.

Le choix de l’ensemble  $\Omega$  comporte une part d’arbitraire. Il dépend de l’idée que l’on a, a priori, sur les résultats de l’expérience aléatoire. Donnons quelques exemples:

1. On lance une pièce de monnaie. Pour l’ensemble  $\Omega$ , on peut choisir soit  $\Omega = \{ \text{pile, face} \}$ , soit  $\Omega = \{ \text{pile, face, tranche} \}$ .
2. On s’intéresse à l’état de fonctionnement d’un système. Dans ce cas  $\Omega = \{0, 1\}$  avec la convention 0 si le système est en panne et 1 s’il fonctionne.
3. Le résultat de l’expérience aléatoire est le nombre de tirages nécessaires dans un jeu de pile ou face jusqu’à l’obtention du premier “pile”. Dans ce cas,  $\Omega = \{1, 2, 3, \dots\} = \mathbb{N}^*$ .
4. On considère la succession des appels à un standard téléphonique non surchargé et l’on étudie la répartition des instants où le standard reçoit un appel, à partir d’un instant choisi comme origine (on admet que deux appels ne peuvent se produire rigoureusement au même instant et que le phénomène est limité dans le temps). Une réalisation de cet événement est une suite croissante de nombres réels positifs  $t_i$  où  $t_i$  désigne l’instant d’enregistrement du  $i$ ème appel:  $\Omega = \{0 < t_1 < t_2 < \dots < t_n < t_{n+1} < \dots\}$ . L’univers  $\Omega$  est donc une partie de  $(\mathbb{R}^+)^{\mathbb{N}}$ .
5. On considère l’expérience aléatoire “durée de vie d’un individu”. L’ensemble  $\Omega$  est soit l’ensemble  $\mathbb{N}$ , soit  $\mathbb{R}^+$  selon le procédé discontinu ou continu de cette mesure.

Nous constatons que  $\Omega$  peut être fini (exemples 1 et 2), dénombrable (exemples 3 et 5) ou non dénombrable (exemples 4 et 5). Lorsque  $\Omega$  est fini ou dénombrable, on parle d’univers *discret*. Sinon on parle d’univers *continu*.

**Définition 2** Etant donnée une expérience aléatoire, un *événement aléatoire* est une partie de l’ensemble des résultats possibles de l’expérience, c’est donc un sous-ensemble  $A$  de l’univers  $\Omega$ . On dit que l’événement  $A$  est réalisé si le résultat  $\omega$  de l’expérience appartient à  $A$ .

On sait que l'événement  $A$  est réalisé seulement une fois l'expérience aléatoire réalisée.

**Exemples:**

- Si l'on s'intéresse à l'événement suivant: "on a obtenu un chiffre pair lors d'un lancer d'un dé à 6 faces", on introduit  $A = \{2, 4, 6\}$ , qui est un sous-ensemble de  $\Omega = \{1, 2, 3, 4, 5, 6\}$ .
- Si l'on s'intéresse à l'événement suivant: "la durée de vie du composant est supérieure ou égale à 1000 heures",  $A = [1000, +\infty[$  est un sous-ensemble de  $\Omega = \mathbb{R}^+$ .

L'ensemble  $\emptyset$  est appelé l'événement impossible et  $\Omega$  est appelé l'événement certain.

## I.2.2 Opérations sur les événements

Les événements aléatoires étant des ensembles, introduisons les opérations ensemblistes classiques de la théorie des ensembles.

**Définition 3** On appelle événement contraire de  $A$ , noté  $A^C$ , le complémentaire de  $A$  dans  $\Omega$ :

$$A^C = \{\omega \in \Omega : \omega \notin A\}.$$

L'événement contraire  $A^C$  est réalisé si et seulement si  $A$  n'est pas réalisé.

**Exemple :** Si  $A$  est l'événement "la durée de vie du composant est supérieure ou égale à 1000 heures":  $A = [1000, +\infty[$ , l'événement contraire est l'événement "la durée de vie du composant est strictement inférieure à 1000 heures":  $A^C = [0, 1000[$ .

**Définition 4** Soient  $A$  et  $B$  deux événements d'un univers  $\Omega$ .

- L'événement "A et B" est celui qui est réalisé si  $A$  et  $B$  sont réalisés. C'est l'intersection

$$A \cap B = \{\omega \in \Omega : \omega \in A \text{ et } \omega \in B\}.$$

- L'événement "A ou B" est celui qui est réalisé si l'un des deux est réalisé ou si les deux sont réalisés. C'est l'union

$$A \cup B = \{\omega \in \Omega : \omega \in A \text{ ou } \omega \in B\}.$$

- L'inclusion  $A \subset B$  signifie que l'événement  $A$  ne peut être réalisé sans que  $B$  le soit.

**Définition 5** Deux événements  $A$  et  $B$  sont dits **incompatibles** si la réalisation de l'un implique la non-réalisation de l'autre.

Dans l'espace  $\Omega$ , deux événements incompatibles sont représentés par deux parties disjointes. Si  $A \cap B = \emptyset$ , alors  $A$  et  $B$  sont incompatibles. Il est clair, par exemple que  $A$  et  $A^C$  sont incompatibles.

### I.2.3 Probabilité

**Définition 6** Soit  $\Omega$  un univers associé à une expérience aléatoire et soit  $\mathcal{A}$  l'ensemble des parties de  $\Omega$ . Une **probabilité**  $\mathbb{P}$  sur l'espace  $(\Omega, \mathcal{A})$  est une application de  $\mathcal{A}$  dans  $[0, 1]$  telle que

1.  $\mathbb{P}(\Omega) = 1$ .
2. Si  $(A_n)_{n \geq 1}$  est une famille d'événements de  $\mathcal{A}$  2 à 2 incompatibles,

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

Le triplet  $(\Omega, \mathcal{A}, \mathbb{P})$  est appelé espace de probabilité.

On peut déduire de la définition précédente un certain nombre de propriétés.

**Proposition 1** Soient  $A$  et  $B$  deux événements aléatoires.

1.  $\mathbb{P}(\emptyset) = 0$ .
2.  $\mathbb{P} \left( \bigcup_{n=1}^N A_n \right) \leq \sum_{n=1}^N \mathbb{P}(A_n)$ .
3. Si  $A_1, \dots, A_N$  sont deux-à-deux incompatibles,

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

4.  $\mathbb{P}(A^C) = 1 - \mathbb{P}(A)$ .
5. Si  $A \subset B$ ,  $\mathbb{P}(A) \leq \mathbb{P}(B)$ .
6.  $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$ .
7. Si  $\Omega$  est fini ou dénombrable, alors pour tout événement  $A$ ,

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

### I.2.4 Probabilité uniforme

Soit  $\Omega$  un ensemble fini:  $\Omega = \{\omega_1, \dots, \omega_N\}$ . Pour tout  $i \in \{1, 2, \dots, N\}$ , on pose  $\mathbb{P}(\{\omega_i\}) = \frac{1}{N}$ . Alors, pour toute partie  $A$  de  $\Omega$ , on a

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \frac{\text{Card}(A)}{N} = \frac{\text{Card}(A)}{\text{Card}(\Omega)}.$$



Dans le cas du lancer de dé à 6 faces, pour tout  $\omega \in \{1, 2, \dots, 6\}$ ,  $\mathbb{P}(\{\omega\}) = 1/6$ . Si on note l'événement "on a obtenu un chiffre pair" par  $A = \{2, 4, 6\}$ , alors

$$\mathbb{P}(A) = 3/6 = 1/2.$$

**Remarque:** Pour un problème donné, il y a souvent plusieurs modélisations possibles, c'est-à-dire que le choix de l'espace de probabilité n'est pas unique.

**Remarque:** Choisir un élément au hasard signifie que les divers choix possibles sont équiprobables, donc que l'ensemble  $\Omega$  est muni de la probabilité uniforme. Dans ce cas, tous les calculs sont simples et se ramènent souvent à des calculs d'analyse combinatoire.

**Exercice 1.** Une urne contient 12 boules numérotées. Les boules numérotées de 1 à 8 sont de couleur blanche, celles portant les numéros de 9 à 12 sont de couleur noire. On tire une boule au hasard dans cette urne. On s'intéresse à la fois au numéro de cette dernière ainsi qu'à sa couleur.

1. Décrire l'espace des issues associé à cette expérience.
2. Quelle est la probabilité de tirer une boule portant un nombre pair?
3. Calculer la probabilité de l'événement "la boule tirée est blanche".
4. Quelle est la probabilité de tirer une boule noire portant un nombre impair?

**Exercice 2. Le paradoxe du chevalier de Méré**

Voici deux jeux qui étaient fort populaires au XVII<sup>ème</sup> siècle:

- Dans le premier jeu, les joueurs pariaient sur le fait d'obtenir au moins un as en lançant quatre fois un dé à six faces.
- Dans le deuxième jeu, ils pariaient sur l'apparition d'au moins un double as en lançant 24 fois une paire de dés.

Le Chevalier de Méré, un noble de l'époque, adorait ces jeux et, comme beaucoup d'autres joueurs de son époque, il pensait que la probabilité de gagner était la même pour les deux jeux.

Par contre, son expérience lui avait montré que le premier jeu était plus avantageux que le deuxième. Le Chevalier de Méré demanda alors au fameux mathématicien Blaise Pascal de l'éclaircir sur ce problème. Celui-ci, dans un échange de lettres avec son ami Pierre Fermat, réussit à résoudre le fameux "paradoxe", que vous êtes invités à votre tour à résoudre.



# Chapitre II

## Probabilités conditionnelles et indépendance

### II.1 Probabilités conditionnelles

Dans le chapitre précédent, on a parlé de la probabilité d'un événement sans tenir compte de la réalisation d'autres événements. En pratique, on peut considérer plusieurs événements, certains pouvant avoir une influence sur la réalisation d'autres événements.

**Exemple:** On lance deux dés. Soient les événements  $A = \{ \text{la somme est } \geq 11 \}$  et  $B = \{ \text{le lancer du 1er dé donne 6} \}$ . Il est clair que la réalisation de  $B$  influe sur la réalisation de  $A$ .

Supposons que l'on s'intéresse à la réalisation d'un événement  $A$ , tout en sachant qu'un événement  $B$  est réalisé. Si  $A$  et  $B$  sont incompatibles, alors la question est réglée:  $A$  ne se réalise pas. Mais si  $A \cap B \neq \emptyset$ , il est possible que  $A$  se réalise. Cependant, l'espace des événements possibles n'est plus  $\Omega$  tout entier, mais il est restreint à  $B$ . En fait, seule nous intéresse la réalisation de  $A$  à l'intérieur de  $B$ , c'est-à-dire  $A \cap B$  par rapport à  $B$ . Ceci justifie la définition suivante.

**Définition 7** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité. Soient  $A$  et  $B$  deux événements aléatoires tels que  $\mathbb{P}(B) \neq 0$ . On appelle **probabilité conditionnelle** de  $A$  sachant  $B$  la quantité

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

**Remarque:** On a les égalités suivantes:

$$\begin{aligned} \text{Si } \mathbb{P}(B) > 0, \quad \mathbb{P}(A \cap B) &= \mathbb{P}(A|B) \times \mathbb{P}(B). \\ \text{Si } \mathbb{P}(A) > 0, \quad \mathbb{P}(A \cap B) &= \mathbb{P}(B|A) \times \mathbb{P}(A). \end{aligned}$$

**Proposition 2 (formule des probabilités totales)** Soit  $(A_i)_{i \in I}$  une famille d'événements aléatoires formant une partition de  $\Omega$ , c'est-à-dire tels que:

- $\cup_{i \in I} A_i = \Omega$ ;
- $A_i \cap A_j = \emptyset$  pour tout  $i \neq j$ .

On suppose de plus que  $\mathbb{P}(A_i) \neq 0$  pour tout  $i \in I$ . Alors

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|A_i)\mathbb{P}(A_i).$$

**Proposition 3 (formule de Bayes)** Sous les mêmes hypothèses que la proposition précédente, on a:

$$\mathbb{P}(A_i|A) = \frac{\mathbb{P}(A|A_i)\mathbb{P}(A_i)}{\sum_{j \in I} \mathbb{P}(A|A_j)\mathbb{P}(A_j)}.$$

La formule de Bayes (publiée après sa mort en 1763) présente un grand intérêt car elle permet de modifier notre connaissance des probabilités en fonction d'informations nouvelles. Cette formule joue donc un rôle très important dans la statistique bayésienne.

## II.2 Indépendance

**Définition 8** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité, et soient  $A$  et  $B$  deux événements aléatoires. On dit que  $A$  et  $B$  sont **indépendants** si

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

**Remarque:**  $A$  et  $B$  sont indépendants si et seulement si  $\mathbb{P}(A|B) = \mathbb{P}(A)$ . Cette condition signifie que la probabilité de réalisation de l'événement  $A$  n'est pas modifiée par une information concernant la réalisation de l'événement  $B$ .

**Proposition 4** Si  $A$  et  $B$  sont deux événements indépendants alors:

- $A^C$  et  $B$  sont également indépendants;
- $A$  et  $B^C$  sont également indépendants;
- $A^C$  et  $B^C$  sont également indépendants.

Nous allons maintenant définir l'indépendance de plus de 2 événements aléatoires.

**Définition 9** Soit  $(\Omega, \mathcal{A}, \mathbb{P})$  un espace de probabilité. Pour  $n \geq 2$ , soient  $A_1, A_2, \dots, A_n$ , des événements aléatoires.

- Ces événements sont deux à deux indépendants si pour tout couple  $(i, j)$  avec  $i \neq j$  on a

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j).$$

- Ces événements sont indépendants (dans leur ensemble) si pour tout  $k \in \{2, 3, \dots, n\}$  et tout choix d'indices distincts  $i_1, \dots, i_k$ , on a

$$\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \dots \mathbb{P}(A_{i_k}).$$

**Exercice 1.** On s'intéresse à la durée de vie d'un téléphone mobile. Pour tout  $t \geq 0$ , on note  $D_t$  l'évènement  $D_t := \{\text{le téléphone fonctionne encore à la date } t\}$ , et l'on suppose que  $\mathbb{P}(D_t) = \exp(-ct)/(1+t)$ ,  $c > 0$ . Calculer la probabilité que le téléphone fonctionne encore au bout de 2 ans sachant qu'il a déjà fonctionné 1 an.

**Exercice 2.** On considère un système de communication qui émet soit un 0 soit un 1 avec équiprobabilité. À cause du bruit, le signal émis est parfois mal reçu. Soit  $E_i$  l'évènement " $i$  est émis" et soit  $R_i$  l'évènement " $i$  est reçu", pour  $i = 0$  et 1. On suppose que  $\mathbb{P}(R_0/E_0) = 0.7$  et  $\mathbb{P}(R_1/E_1) = 0.8$ .

- Calculer  $\mathbb{P}(E_0/R_1)$ .
- Calculer la probabilité d'une erreur de transmission.

**Exercice 3.** On considère un système formé de 5 constituants indépendants. La probabilité qu'un constituant soit défectueux est de 0,2. Quelle est la probabilité que ce système fonctionne si les constituants sont branchés:

- en série ?
- en parallèle ?

**Exercice 4.** On lance deux dés et on considère les événements  $A$ : "le premier dé donne un nombre pair",  $B$ : "le deuxième dé donne un nombre pair" et  $C$ : "la somme des dés est un nombre pair". Montrer que les événements  $A, B$  et  $C$  sont deux à deux indépendants. Sont-ils indépendants dans leur ensemble?



# Chapitre III

## Variables aléatoires réelles

### III.1 Introduction

Dans de nombreuses expériences aléatoires, on n'est pas intéressé directement par le résultat de l'expérience, mais par une certaine fonction de ce résultat. Considérons par exemple l'expérience qui consiste à observer, pour chacune des  $n$  pièces produites par une machine, si la pièce est défectueuse ou non. Nous attribuerons la valeur 1 à une pièce défectueuse et la valeur 0 à une pièce en bon état. L'univers associé à cette expérience est  $\Omega = \{0, 1\}^n$ . Ce qui intéresse le fabricant est la proportion de pièces défectueuses produites par la machine. Introduisons donc une fonction de  $\Omega$  dans  $\mathbb{R}$  qui à tout  $\omega = (\omega_1, \omega_2, \dots, \omega_n)$  de  $\Omega$  associe le nombre

$$X(\omega) = \sum_{i=1}^n \frac{\omega_i}{n},$$

qui correspond à la proportion de pièces défectueuses associée à l'observation de  $\omega$ . Une telle fonction  $X$  définie sur  $\Omega$  et à valeurs dans  $\mathbb{R}$  s'appelle une variable aléatoire réelle.

### III.2 Définitions

#### III.2.1 Variable aléatoire

**Définition 10** *Étant donné un univers  $\Omega$ , une **variable aléatoire réelle** (v.a.r.) est une application de  $\Omega$  dans  $\mathbb{R}$ :*

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

#### III.2.2 Loi de probabilité

**Définition 11** *Soit  $\Omega$  un univers muni d'une probabilité  $\mathbb{P}$ , et soit  $X$  une v.a.r. On appelle **loi de probabilité** de  $X$ , notée  $\mathbb{P}_X$ , l'application qui à toute partie  $A$  de  $\mathbb{R}$  associe*

$$\mathbb{P}_X(A) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}).$$

**Remarque :** Dans la suite du cours, on utilisera la notation abrégée:  $\mathbb{P}(\{\omega \in \Omega : X(\omega) \in A\}) = \mathbb{P}(X \in A)$ . De même, on notera  $\mathbb{P}(X = x)$  la probabilité  $\mathbb{P}(\{\omega \in \Omega : X(\omega) = x\})$ .

**Proposition 5** *L'application  $\mathbb{P}_X$  définit une probabilité sur  $\mathbb{R}$ .*

### III.2.3 Fonction de répartition

**Définition 12** *La fonction de répartition de la v.a.r.  $X$  est définie par*

$$F_X(x) = \mathbb{P}(X \leq x), \quad \forall x \in \mathbb{R}.$$

**Propriétés de la fonction de répartition:**

1.  $0 \leq F_X(x) \leq 1, \forall x \in \mathbb{R}$ .
2.  $F_X$  tend vers 0 en  $-\infty$  et vers 1 en  $+\infty$ .
3.  $F_X$  est croissante.
4.  $F_X$  est continue à droite.

**Proposition 6** *On a l'identité*

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a), \quad \forall (a, b) \in \mathbb{R}^2, a < b.$$

**Remarque:** On montre facilement que  $F_X$  est continue si et seulement si  $\mathbb{P}(X = x) = 0$  pour tout  $x \in \mathbb{R}$ . On parle alors de loi diffuse ou de v.a.r. continue (voir définition 15).

**Définition 13** *Soit  $X$  une v.a.r. de fonction de répartition  $F_X$  supposée strictement croissante de  $I \subset \mathbb{R}$  dans  $]0, 1[$ . Le **quantile d'ordre**  $\alpha \in ]0, 1[$  de  $X$  est le nombre  $x_\alpha \in I$  tel que  $F_X(x_\alpha) = \alpha$ , ce qui signifie que*

$$\mathbb{P}(X \leq x_\alpha) = \alpha.$$

**Remarques:**

- $x_{1/2}$  est appelé **médiane** de  $X$ . La médiane vérifie les deux égalités

$$P(X \leq x_{1/2}) = 1/2 = P(X > x_{1/2}).$$

- Dans le cas où  $F_X$  n'est pas strictement croissante mais simplement croissante, on définit le quantile d'ordre  $\alpha$  par

$$x_\alpha = \inf \{x \in \mathbb{R} : F_X(x) \geq \alpha\}.$$



## III.3 Différents types de variables aléatoires réelles

### III.3.1 Variables aléatoires discrètes

**Définition 14** Une v.a.r.  $X$  à valeurs dans un ensemble  $\mathcal{X}$  fini ou dénombrable est appelée v.a.r. **discrète**. Dans ce cas, la loi de  $X$  est déterminée par l'ensemble des probabilités:

$$\mathbb{P}_X(x) = \mathbb{P}(X = x), \quad \forall x \in \mathcal{X}.$$

Ainsi, pour toute partie  $A$  de  $\mathcal{X}$ , on a alors:

$$\mathbb{P}_X(A) = \mathbb{P}(X \in A) = \sum_{x \in A} \mathbb{P}(X = x) \quad \text{et} \quad \mathbb{P}_X(\mathcal{X}) = \sum_{x \in \mathcal{X}} \mathbb{P}(X = x) = 1.$$

**Exemple:** Supposons que l'on observe la durée de vie  $T$  d'une ampoule électrique et que cette durée de vie  $T$ , exprimée en heures, satisfait pour tout  $0 < a < b$ ,

$$\mathbb{P}(a < T \leq b) = \exp(-a/100) - \exp(-b/100).$$

On note  $X$  le nombre de périodes complètes de 100 heures que dure l'ampoule. Les valeurs possibles de  $X$  étant entières, la v.a.r.  $X$  est donc discrète. Calculons la fonction de répartition de  $X$ . Comme  $X$  est positive, on a

$$F_X(x) = \mathbb{P}(X \leq x) = 0, \quad \forall x < 0.$$

De plus, pour tout  $n \in \mathbb{N}$ ,

$$\mathbb{P}(X = n) = \mathbb{P}(100n \leq T < 100(n+1)) = \exp(-n) - \exp(-(n+1)).$$

Ainsi, on a donc pour tout  $x \geq 0$ :

$$\begin{aligned} \mathbb{P}(X \leq x) &= \sum_{n=0}^{[x]} \mathbb{P}(X = n) \\ &= 1 - \exp(-([x] + 1)). \end{aligned}$$

On notera que la fonction  $F_X$  est une fonction en escalier.

**Exercice 1.** On lance une paire de dés et on appelle  $X$  le minimum et  $Y$  le maximum des deux nombres obtenus. Déterminer les lois des variables aléatoires  $X$  et  $Y$ .

### III.3.2 Variables aléatoires continues

**Définition 15** Soit  $X$  une v.a.r. qui prend un nombre infini non dénombrable de valeurs. Si  $F_X$  est une fonction continue, on dit que  $X$  est une v.a.r. **continue**. Dans ce cas, la loi de  $X$  est déterminée par l'ensemble des probabilités  $\mathbb{P}(a < X < b)$ , pour tout  $a < b$ .

**Remarque:** Notons que l'on peut mettre  $<$  ou  $\leq$  dans ce qui précède car la variable étant continue, on a  $\mathbb{P}(X = x) = 0$  pour tout  $x \in \mathbb{R}$ .

**Exemple :** Soit  $\lambda > 0$ . Une v.a.r.  $X$  de fonction de répartition

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

est continue.

**Définition 16** Si l'on peut écrire la fonction de répartition d'une variable continue sous la forme

$$F_X(t) = \int_{-\infty}^t f_X(x) dx,$$

où  $f_X$  est une fonction de  $\mathbb{R}$  dans  $\mathbb{R}$ , alors on dit que  $f_X$  est la densité de probabilité de la v.a.r.  $X$ .

Ceci implique que l'on a pour tout  $a < b$ :

$$\mathbb{P}(a < X < b) = F_X(b) - F_X(a) = \int_a^b f_X(x) dx.$$

Cette intégrale étant positive pour tout  $a < b$ , il en résulte que  $f_X \geq 0$ . De plus, puisque  $\lim_{t \rightarrow +\infty} F_X(t) = 1$ , on a

$$\int_{-\infty}^{+\infty} f_X(x) dx = 1.$$

Une densité de probabilité est donc une fonction positive ou nulle, d'intégrale 1, et qui caractérise la loi d'une v.a.r. continue. De plus, en tout point  $x_0 \in \mathbb{R}$  où  $F_X$  est dérivable, on a  $f_X(x_0) = F_X'(x_0)$ .

**Exemple:** Dans l'exemple de la durée de vie  $T$  d'une ampoule électrique,  $T$  a pour densité de probabilité

$$f(x) = \begin{cases} \exp(-x/100)/100 & \text{pour tout } x \geq 0 \\ 0 & \text{pour tout } x < 0. \end{cases}$$

Enfin, établir que deux v.a.r. (discrètes ou continues)  $X$  et  $Y$  ont même loi, c'est démontrer que l'on a l'égalité suivante:

$$\mathbb{P}(a < X \leq b) = \mathbb{P}(a < Y \leq b), \quad a, b \in \mathbb{R}.$$

Ainsi, en faisant tendre  $a$  vers  $-\infty$ , on obtient le résultat suivant:

**Théorème 1** Deux v.a.r. à valeurs dans le même ensemble d'arrivée ont la même loi si et seulement si leurs fonctions de répartition sont égales.

## III.4 Exemples fondamentaux

### III.4.1 Variables discrètes

Soit  $X$  une v.a.r. discrète prenant ses valeurs dans un ensemble  $\{x_1, x_2, \dots, x_n\}$ , éventuellement infini. Alors la loi de  $X$  est caractérisée par l'ensemble des probabilités  $\mathbb{P}(X = x_i)$ , c'est-à-dire les nombres réels positifs  $p_i$  tels que

$$\mathbb{P}(X = x_i) = p_i \quad \text{avec} \quad 0 \leq p_i \leq 1 \quad \text{et} \quad \sum_{i=1}^n p_i = 1.$$

#### a) Loi de Bernoulli

On dit qu'une v.a.r.  $X$  à valeurs dans  $\{0, 1\}$  suit une loi de Bernoulli de paramètre  $p \in ]0, 1[$ , notée  $\mathcal{B}(p)$ , si

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

Par exemple, cette loi intervient lorsque l'on modélise l'état de fonctionnement d'un système. La probabilité que le système fonctionne vaut  $p$  et la probabilité que le système ne fonctionne pas vaut  $1 - p$ . Cette loi s'applique aussi aux jeux de hasard de type binaire comme pile ou face ...

#### b) Loi binomiale

On dit qu'une v.a.r.  $X$  à valeurs dans  $\{0, 1, \dots, n\}$  suit une loi binomiale de paramètres  $(n, p)$ , notée  $\mathcal{B}(n, p)$ , si

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad \forall k \in \{0, \dots, n\}.$$

Cette loi intervient par exemple pour modéliser le nombre de pièces défectueuses dans un lot de  $n$  pièces, qui ont chacune une probabilité  $p$  d'être défectueuse, indépendamment les unes des autres.

#### c) Loi géométrique

On dit qu'une v.a.r.  $X$  à valeurs dans  $\mathbb{N}^*$  suit une loi géométrique de paramètre  $p \in ]0, 1[$ , notée  $\mathcal{G}(p)$ , si

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

Cette loi permet de modéliser le nombre de réalisations indépendantes d'une expérience à 2 issues (succès-échec), jusqu'à l'obtention du premier succès, si à chaque réalisation la probabilité de succès est  $p$ .

#### d) Loi de Poisson

On dit qu'une v.a.r.  $X$  à valeurs dans  $\mathbb{N}$  suit une loi de Poisson de paramètre  $\lambda > 0$ , notée  $\mathcal{P}(\lambda)$ , si

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad k \in \mathbb{N}.$$

Cette loi intervient comme comportement limite de la loi binomiale lorsque  $n \rightarrow +\infty$  et  $np \rightarrow \lambda$ .

### III.4.2 Variables continues

Soit  $X$  une v.a.r. continue. Alors la loi de  $X$  est caractérisée par l'ensemble des probabilités

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x) dx,$$

où  $f_X$  est la densité de probabilité de  $X$  et  $a$  et  $b$  sont deux nombres réels, éventuellement infinis. Comme nous l'avons vu plus haut, il suffit de connaître cette densité pour connaître la loi de  $X$ .

**a) Loi exponentielle**

On dit que  $X$  suit une loi exponentielle de paramètre  $\lambda > 0$ , notée  $\mathcal{E}(\lambda)$ , si la loi de  $X$  a pour densité

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

La loi exponentielle est utilisée en fiabilité. Le paramètre  $\lambda$  représente le taux moyen de défaillance alors que son inverse  $\theta = 1/\lambda$  est “le temps moyen de bon fonctionnement”.

La loi exponentielle s'applique bien aux matériels électroniques ou aux matériels subissant des défaillances brutales.

**b) Loi Gamma**

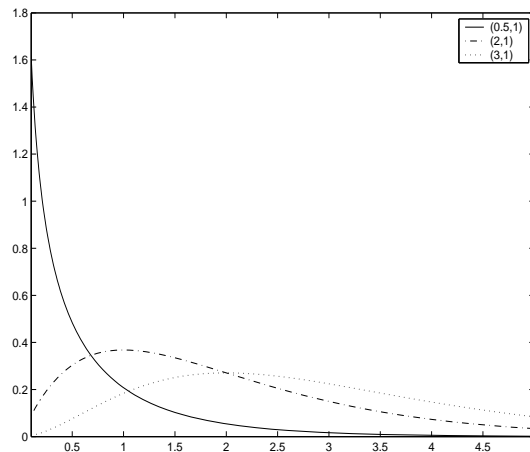
La loi exponentielle est un cas particulier de la famille des lois Gamma.

Soient  $a > 0$  et  $\lambda > 0$ . On dit que  $X$  suit une loi Gamma de paramètres  $(a, \lambda)$ , notée  $\gamma(a, \lambda)$ , si la loi de  $X$  a pour densité

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

où pour tout  $a > 0$ , la célèbre fonction gamma est donnée par  $\Gamma(a) = \int_0^{+\infty} x^{a-1} \exp(-x) dx$ . Pour  $a = 1$ , on retrouve la loi exponentielle de paramètre  $\lambda$ . Le paramètre  $a$  est un paramètre de forme alors que le paramètre  $\lambda$  est un paramètre d'échelle.

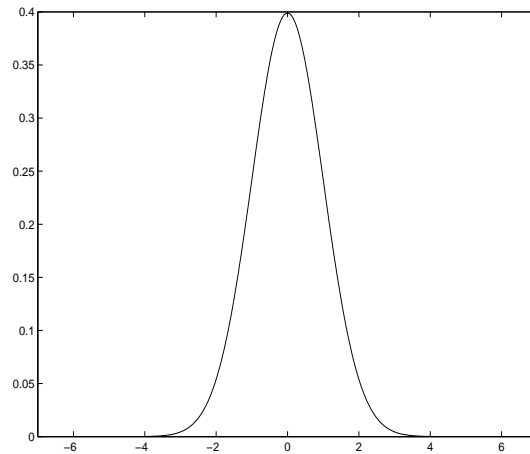
Tracé de la densité de la loi Gamma:

**c) Loi normale de paramètres  $(m, \sigma^2)$** 

Soient  $m \in \mathbb{R}$  et  $\sigma > 0$ . On dit que  $X$  suit une loi normale de paramètres  $(m, \sigma^2)$ , notée  $\mathcal{N}(m, \sigma^2)$ , si la loi de  $X$  a pour densité

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \quad x \in \mathbb{R}.$$

Tracé de la densité de la loi normale de paramètres  $(0, 1)$ :



À cause de sa forme, cette courbe est souvent appelée “courbe en cloche”. Elle présente un axe de symétrie vertical pour  $x = m$ . De plus, comme il n'existe pas d'expression analytique de la fonction de répartition de  $X$ , on utilise alors des tables obtenues par des calculs approchés d'intégrales.

**Exemple:** La loi de la v.a.r.  $X$  représentant le diamètre d'un roulement à bille peut être modélisée par une loi  $\mathcal{N}(m, \sigma^2)$ . Dans ce cas,  $m$  représente le diamètre moyen d'un roulement à bille, et  $\sigma^2$  désigne la précision de cette mesure.

La loi normale s'applique à de nombreux phénomènes, en physique, en économie (erreurs de mesure). Comme on le verra au chapitre V, elle est la forme limite de nombreuses lois discrètes. Elle peut aussi représenter la fin de vie des dispositifs subissant un phénomène de vieillissement: usure, corrosion, etc ... Il faut cependant remarquer que les variables utilisées dans les domaines technologique ou économique sont bien souvent positives. Pour que la loi normale puisse être représentative d'un tel phénomène, il faut que la probabilité d'obtenir des valeurs négatives de la variable soit très faible. Il faut en particulier éviter d'utiliser cette modélisation pour les queues des distributions.



# Chapitre IV

## Manipulation de variables aléatoires

### IV.1 Caractéristiques des variables aléatoires

#### IV.1.1 Espérance

**Définition 17** Soit  $X$  une v.a.r. et  $h$  une application de  $\mathbb{R}$  dans  $\mathbb{R}$ . Donc  $h(X)$  est elle aussi une v.a.r.

- Si  $X$  est discrète à valeurs dans un ensemble  $\mathcal{X}$ , l'espérance de  $h(X)$  est la quantité

$$\mathbb{E}(h(X)) = \sum_{x \in \mathcal{X}} h(x) \mathbb{P}(X = x),$$

pourvu que cette série converge (dans le cas où  $\mathcal{X}$  est infini).

- Si  $X$  est continue et admettant une densité  $f_X$ , l'espérance de  $h(X)$  est la quantité

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) f_X(x) dx,$$

à condition que cette intégrale soit convergente.

Notons que si  $h(x) = x$ , on obtient  $\mathbb{E}(X)$  appelée **espérance mathématique** (ou **moyenne**) de la v.a.r.  $X$ . Par ailleurs, pour  $A \subset \mathbb{R}$ , si l'on définit la v.a.r. suivante:

$$1_{\{X \in A\}} = \begin{cases} 1 & \text{si } X \in A \\ 0 & \text{sinon} \end{cases}$$

qui est appelée fonction caractéristique de l'événement  $\{X \in A\}$ , alors l'espérance de cette v.a.r. est:

$$\mathbb{E}(1_{\{X \in A\}}) = \mathbb{P}(X \in A) = \mathbb{P}_X(A),$$

d'où le lien étroit entre probabilité et espérance.

**Exemple :** Une roulette contient 18 cases rouges, 18 cases noires et une case verte. Si on mise sur une couleur et cette couleur sort, on gagne encore une fois la mise, sinon on perd la mise. Soit  $G$  le gain obtenu quand on joue une fois en misant 1 euro sur le rouge. La loi de la variable aléatoire  $G$  est donnée par  $\mathbb{P}(G = -1) = 19/37$ ,  $\mathbb{P}(G = 1) = 18/37$ . Le gain moyen est  $\mathbb{E}(G) = (-1) \times \frac{19}{37} + 1 \times \frac{18}{37} = -\frac{1}{37}$ . Le jeu n'est donc pas équitable et il est en faveur de la banque.

### IV.1.2 Propriétés de l'espérance

1. L'espérance est linéaire: pour tout  $\alpha, \beta \in \mathbb{R}$ , et pour toutes v.a.r.  $X$  et  $Y$

$$\mathbb{E}(\alpha X + \beta Y) = \alpha \mathbb{E}(X) + \beta \mathbb{E}(Y).$$

2. Si  $X$  est une v.a.r. constante égale à  $a \in \mathbb{R}$ , c'est-à-dire pour tout  $\omega \in \Omega$ ,  $X(\omega) = a$ , alors  $\mathbb{P}(X = a) = 1$  et  $\mathbb{E}(X) = a$ .
3. L'espérance d'une v.a.r. positive est positive. En particulier, si  $X \geq Y$  (ce qui signifie que pour tout  $\omega \in \Omega$ ,  $X(\omega) \geq Y(\omega)$ ), alors  $\mathbb{E}(X - Y) \geq 0$  donc  $\mathbb{E}(X) \geq \mathbb{E}(Y)$ .

L'espérance d'une v.a.r.  $X$  est un indicateur de "localisation" de sa loi:

$$\mathbb{E}(X) \simeq \text{"valeur moyenne de } X\text{"}.$$

Néanmoins, la connaissance de l'espérance mathématique donne peu de renseignements sur cette v.a.r. Ainsi, il faut étudier "l'étalement" de sa loi, c'est-à-dire la dispersion de la v.a.r.  $X$  autour de sa moyenne  $\mathbb{E}(X)$ .

### IV.1.3 Variance et écart-type

Pour rendre positifs les écarts entre  $X$  et son espérance  $\mathbb{E}(X)$ , un autre outil plus facile à manipuler que la valeur absolue, est à notre disposition: la mise au carré. On ne va donc pas calculer la moyenne des écarts mais la moyenne des écarts au carré. C'est ce qu'on appelle la variance.

**Définition 18** La variance de la v.a.r.  $X$  est la quantité:

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

**Propriétés :**

- $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$ .
- $\text{Var}(aX + b) = a^2 \text{Var}(X)$  pour tout  $a, b \in \mathbb{R}$ . En particulier,  $\text{Var}(X + b) = \text{Var}(X)$ .

Afin d'être en mesure de comparer, en termes d'ordre de grandeur, variance et espérance, il faut prendre la racine carrée de la variance. C'est ce qu'on appelle l'écart-type.

**Définition 19** La racine carrée de  $\text{Var}(X)$ , notée  $\sigma_X$ , est appelée écart-type de  $X$ .

**Remarque :** Si  $X$  est une v.a.r. telle que  $\mathbb{E}(X) = m$  et  $\text{Var}(X) = \sigma^2$ , alors la variable  $Y = (X - m)/\sigma$  est d'espérance nulle et de variance 1. On dit que  $Y$  est centrée (d'espérance nulle) et réduite (de variance 1).

**Exemples :**

- Loi de Bernoulli  $\mathcal{B}(p)$ :  $\mathbb{E}(X) = p$  et  $\text{Var}(X) = p(1 - p)$ .
- Loi binomiale  $\mathcal{B}(n, p)$ :  $\mathbb{E}(X) = np$  et  $\text{Var}(X) = np(1 - p)$ .
- Loi géométrique  $\mathcal{G}(p)$ :  $\mathbb{E}(X) = \frac{1}{p}$  et  $\text{Var}(X) = \frac{1-p}{p^2}$ .
- Loi de Poisson  $\mathcal{P}(\lambda)$ :  $\mathbb{E}(X) = \text{Var}(X) = \lambda$ .
- Loi normale  $\mathcal{N}(m, \sigma^2)$ :  $\mathbb{E}(X) = m$  et  $\text{Var}(X) = \sigma^2$ .
- Loi exponentielle  $\mathcal{E}(\lambda)$ :  $\mathbb{E}(X) = \frac{1}{\lambda}$  et  $\text{Var}(X) = \frac{1}{\lambda^2}$ .



**Exercice 1.** Soit  $X$  une variable aléatoire de densité de probabilité

$$f(x) = \begin{cases} (1 + \theta x)/2 & \text{si } -1 \leq x \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

1. Donner les valeurs du paramètre  $\theta$  pour que  $f$  soit bien une densité de probabilité.
2. Calculer l'espérance et la variance de  $X$ .

## IV.2 Indépendance de variables aléatoires

**Définition 20** Deux v.a.r.  $X$  et  $Y$  sont dites **indépendantes** si et seulement si

$$\mathbb{P}(\{X \in A\} \cap \{Y \in B\}) = \mathbb{P}(X \in A) \times \mathbb{P}(Y \in B), \quad \forall A, B \subset \mathbb{R}.$$

On peut montrer que l'indépendance est équivalente à

$$\mathbb{P}(\{X \leq a\} \cap \{Y \leq b\}) = \mathbb{P}(X \leq a) \times \mathbb{P}(Y \leq b), \quad \forall (a, b) \in \mathbb{R}^2,$$

ou encore en termes de fonctions de répartition:

$$F_{X,Y}(a, b) = F_X(a)F_Y(b), \quad \forall (a, b) \in \mathbb{R}^2.$$

**Théorème 2** Deux v.a.r. discrètes  $X$  et  $Y$  sont indépendantes si et seulement si, pour tout couple  $(x, y) \in \mathcal{X} \times \mathcal{Y}$ , on a  $\mathbb{P}(X = x \cap Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y)$ .

**Exemple :** On lance une paire de dés et on note  $X$  (respectivement  $Y$ ) le minimum (respectivement maximum) des deux nombres sortis. On remarque que  $X$  et  $Y$  peuvent prendre toutes deux des valeurs de 1 à 6, mais  $\mathbb{P}(X = i \cap Y = j) = 0$ , pour tout  $i > j$ . Les variables  $X$  et  $Y$  ne sont donc pas indépendantes.



# Chapitre V

## Théorèmes limites

Deux théorèmes mathématiques ont une place particulière en théorie des probabilités et en statistiques: la loi des grands nombres et le théorème central limite. Ils interviennent dans l'étude de phénomènes aléatoires comportant un grand nombre de v.a.r. indépendantes de même loi. Par exemple, pour le premier cité, il apparaît lorsque l'on étudie la proportion de "pile" dans un jeu de pile ou face, ou encore la moyenne de lancers de dé successifs. Quant au second, il nous donne de façon informelle une estimation précise de l'erreur que l'on commet en approchant l'espérance mathématique par la moyenne arithmétique.

### V.1 Inégalité de Chebychev

**Théorème 3** (*inégalité de Chebychev*)

Soit  $\epsilon > 0$  et soit  $X$  une v.a.r. admettant une variance. Alors on a:

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

Cette inégalité permet de comprendre la signification de l'écart-type  $\sigma_X = \sqrt{\text{Var}(X)}$ , au sens où il caractérise la dispersion de la v.a.r. autour de son espérance mathématique:

$$\begin{aligned}\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon\sigma_X) &\leq \frac{1}{\epsilon^2} \\ \mathbb{P}(|X - \mathbb{E}(X)| < \epsilon\sigma_X) &\geq 1 - \frac{1}{\epsilon^2}.\end{aligned}$$

Supposons que  $\epsilon = 10$ . Alors l'événement  $\{|X - \mathbb{E}(X)| \geq 10\sigma_X\}$  a peu de chances de se réaliser car

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq 10\sigma_X) \leq \frac{1}{100}.$$

Supposons maintenant que  $\sigma_X = 0$ . Alors nous obtenons pour tout  $\epsilon > 0$

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \epsilon) \leq 0.$$

Par conséquent

$$\mathbb{P}(|X - \mathbb{E}(X)| > 0) = 0,$$

et donc  $X$  est presque sûrement égale à  $\mathbb{E}(X)$ .

On doit cependant remarquer que, malgré son intérêt théorique certain, l'inégalité de Chebychev présente peu d'intérêt en pratique, car ne faisant pas intervenir la loi de probabilité suivie par la v.a.r. considérée, elle donne une majoration de la probabilité beaucoup trop grande.

## V.2 Loi (faible) des grands nombres

**Théorème 4 (LGN)** Soient  $X_1, \dots, X_n$  des v.a.r. indépendantes, de même loi, et admettant une variance. On note  $m = \mathbb{E}(X_1)$ . Alors, pour tout  $\epsilon > 0$ ,

$$\mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| > \epsilon\right) \xrightarrow{n \rightarrow \infty} 0.$$

Dans ce cas, on dit que la moyenne arithmétique  $\frac{X_1 + \dots + X_n}{n}$  **converge en probabilité** vers l'espérance mathématique  $m$  lorsque  $n$  tend vers  $+\infty$ .

## V.3 Théorème central limite

On a vu que deux v.a.r. ont la même loi si et seulement si leur fonctions de répartition sont égales. Ainsi, la fonction de répartition est souvent utilisée en pratique afin de démontrer l'égalité en loi. On est donc amené à définir la convergence en loi comme la convergence des fonctions de répartition associées.

**Définition 21** Soit  $(Y_n)_{n \in \mathbb{N}}$  une suite de v.a.r. et soit  $Y$  une v.a.r. On dit que  $(Y_n)_{n \in \mathbb{N}}$  **converge en loi** vers  $Y$  si pour tout  $x_0$  point de continuité de la fonction de répartition  $F_Y$  de  $Y$ ,

$$F_{Y_n}(x_0) = \mathbb{P}(Y_n \leq x_0) \xrightarrow{n \rightarrow +\infty} F_Y(x_0) = \mathbb{P}(Y \leq x_0).$$

On note la convergence en loi  $Y_n \xrightarrow{\mathcal{L}} Y$ .

La convergence en loi est réalisée aux points de continuité de  $F_Y$ . C'est la convergence simple de la suite de fonctions de répartition  $F_{Y_n}$ .

Propriété d'additivité de la loi normale: si  $X_1, \dots, X_n$  sont des v.a.r. indépendantes et de même loi  $\mathcal{N}(m, \sigma^2)$ , alors la v.a.r.  $X_1 + \dots + X_n$  suit la loi  $\mathcal{N}(nm, n\sigma^2)$ . Ce résultat implique que la v.a.r. centrée réduite  $\sum_{i=1}^n \frac{X_i - m}{\sigma\sqrt{n}}$  suit la loi normale  $\mathcal{N}(0, 1)$ .

Que se passe-t-il dans le cas général où les v.a.r.  $X_i$  ne sont pas nécessairement normales? Le résultat ci-dessus se transforme alors en un résultat de convergence en loi.

**Théorème 5 (TCL)** Soient  $X_1, \dots, X_n$  des v.a.r. indépendantes, de même loi, et admettant une variance. On note  $m = \mathbb{E}(X_1)$  et  $\sigma^2 = \text{Var}(X_1)$ . Alors

$$\frac{X_1 + \dots + X_n - nm}{\sigma\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1) \quad \text{lorsque } n \rightarrow +\infty.$$

## V.4 Approximation d'une loi binomiale

On contrôle  $n$  pièces et on introduit les v.a.r.  $X_1, \dots, X_n$  définies par  $X_i = 1$  si la  $i$ ème pièce contrôlée est défectueuse, et 0 sinon. On note  $Y = X_1 + \dots + X_n$  le nombre total de pièces défectueuses dans le lot. Alors la v.a.r.  $Y$  suit une loi binomiale de paramètres  $(n, p)$  où  $p$  est la probabilité qu'une pièce soit défectueuse.

### V.4.1 Approximation par une loi normale

Puisque les v.a.r.  $X_i$  sont indépendantes, de même loi, et de variance finie:

$$\mathbb{P}(X_i = 1) = p, \quad \mathbb{P}(X_i = 0) = 1 - p, \quad \mathbb{E}(X_i) = p \quad \text{et} \quad \text{Var}(X_i) = p(1 - p),$$

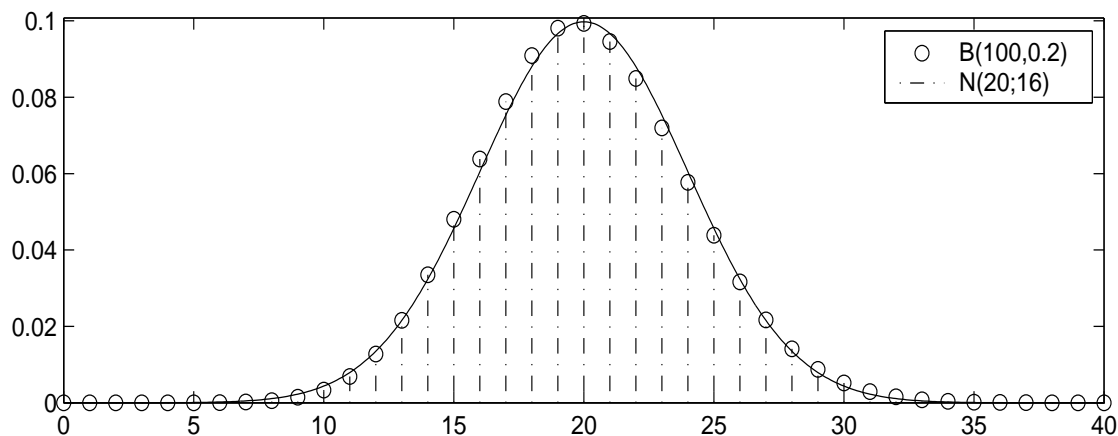
on peut appliquer le TCL:

$$\frac{X_1 + \dots + X_n - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, 1) \quad \text{lorsque} \quad n \rightarrow +\infty.$$

On peut donc approcher la loi de la v.a.r.  $(Y - np)/\sqrt{np(1-p)}$  par une loi normale  $\mathcal{N}(0, 1)$ . Ceci revient à approcher la loi de  $Y$  par une loi  $\mathcal{N}(np, np(1-p))$ . En pratique, on utilise cette approximation si

$$\min(np, n(1-p)) \geq 5.$$

Ainsi, si  $p = 1/2$ , l'approximation est correcte pour  $n \geq 10$ , par contre si  $p = 1/100$ , il faut  $n \geq 500$  pour pouvoir l'utiliser.



Deux remarques importantes doivent être faites. Tout d'abord, une v.a.r. binomiale est une variable discrète à valeurs dans  $\{0, \dots, n\}$ , alors qu'une v.a.r. normale est continue et à valeurs dans  $\mathbb{R}$ . Par ailleurs, dans le cas d'une loi binomiale, un point a une probabilité non nulle alors que dans le cas d'une loi normale, un point est un ensemble de probabilité nulle. Pour ces deux raisons, il faut faire une "correction de continuité" quand on utilise l'approximation d'une loi binomiale par une loi normale. La représentation graphique de la densité d'une variable aléatoire de loi normale est une courbe en "cloche" alors que celle d'une variable aléatoire binomiale est un diagramme en "bâtons". Une valeur approchée de  $\mathbb{P}(X = k)$  est donnée par l'aire comprise entre la courbe en cloche et les droites d'abscisses  $k - 0,5$  et  $k + 0,5$ . On obtient alors

$$\mathbb{P}(X = k) \simeq \mathbb{P}\left(\frac{k - 0,5 - np}{\sqrt{np(1-p)}} < Z < \frac{k + 0,5 - np}{\sqrt{np(1-p)}}\right),$$

ou encore

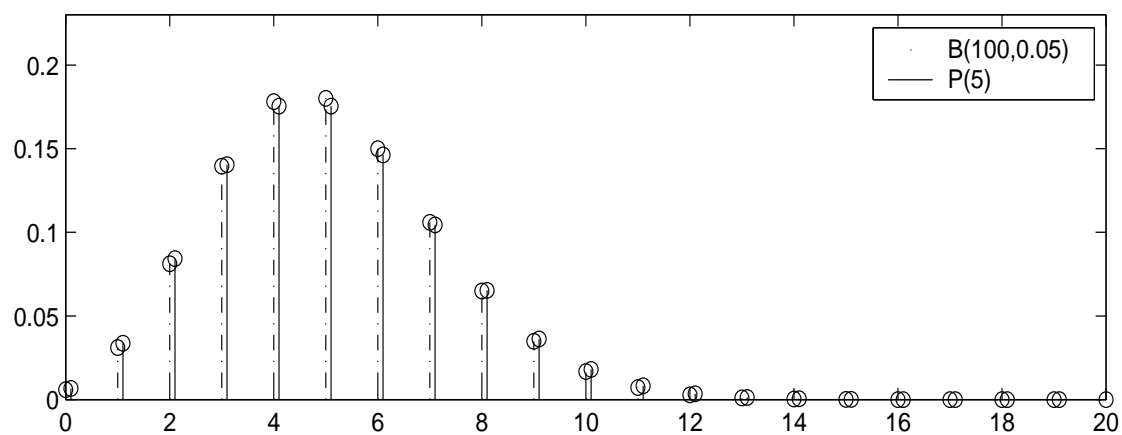
$$\mathbb{P}(X \leq k) \simeq \mathbb{P}\left(Z < \frac{k + 0,5 - np}{\sqrt{np(1-p)}}\right).$$

### V.4.2 Approximation par une loi de Poisson

Lorsque  $p$  est très petit, on utilise plutôt l'approximation de la loi binomiale par une loi de Poisson, qui est satisfaisante pour  $p < 0,1$  et  $n > 50$ .

**Proposition 7** Soit  $Y_n$  une v.a.r. binomiale de paramètres  $(n, p)$ . On suppose que  $n \rightarrow +\infty$  et  $p = \lambda/n$ , où  $\lambda > 0$ . Alors, pour tout  $k \in \mathbb{N}$ ,

$$\lim_{n \rightarrow +\infty} P(Y_n = k) = \exp(-\lambda) \frac{\lambda^k}{k!}.$$



# Chapitre VI

## Estimation

### VI.1 Introduction

Le cadre est le suivant: on dispose d'un nombre fini d'observations sur une population donnée. L'objectif est de pouvoir tirer des conclusions sur certaines caractéristiques de cette population à partir de ces observations. On fait alors une hypothèse raisonnable: il existe une loi de probabilité sous-jacente telle que les "valeurs observables" des différents éléments de la population étudiée puissent être considérées comme des variables aléatoires indépendantes ayant cette loi.

Un aspect important de l'inférence statistique consiste à obtenir des "estimations fiables" des caractéristiques d'une population (de grande taille) à partir d'un échantillon (de petite taille) extrait de cette population. C'est un problème de décision concernant des paramètres qui le plus souvent sont:

- une espérance mathématique  $m$ ;
- une proportion  $p$ ;
- une variance  $\sigma^2$ .

Ces paramètres sont a priori inconnus car la taille réelle de la population étant très grande, il serait trop coûteux de tester tous les éléments de la population. Ainsi, comme un échantillon ne peut donner qu'une information partielle sur la population, les estimations que l'on obtiendra seront inévitablement entachées d'erreurs qu'il s'agit d'évaluer et de minimiser autant que possible.

En résumé, estimer un paramètre inconnu, c'est en donner une valeur approchée à partir des résultats obtenus sur un échantillon aléatoire extrait de la population sous-jacente.

**Exemple:** Une usine produit des composants électroniques. Le fabricant désire connaître la proportion  $p$  de composants défectueux à l'issue du processus de fabrication. Pour obtenir la valeur exacte de  $p$ , il suffirait de tester chacun des composants, opération qui peut se révéler extrêmement coûteuse en temps ou en argent (voire de facto impossible). Une alternative intéressante consiste à ne tester qu'un nombre fini  $n$  de composants prélevés au hasard sur la chaîne de montage. On obtient un ensemble de  $n$  réalisations  $x_i$  de v.a.r. indépendantes de Bernoulli  $X_i$  définies par:

$$X_i = \begin{cases} 1 & \text{si le composant } i \text{ est défectueux} \\ 0 & \text{sinon.} \end{cases}$$

Il est naturel d'estimer  $p$ , la proportion d'objets defectueux dans **la population**, par  $\bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$ , correspondant à la proportion de composants defectueux dans **l'échantillon**. En effet, la LGN nous assure de la convergence en probabilité de la v.a.r.  $\frac{1}{n} \sum_{i=1}^n X_i$  vers l'espérance de  $X_1$ , c'est-à-dire  $p$ . On estimera donc  $p$  par  $\bar{x}_n$ . C'est une estimation dite ponctuelle.

Le but des statistiques n'est pas simplement de donner une estimation ponctuelle d'un paramètre d'intérêt (l'estimateur étant aléatoire... et donc bien souvent faux par nature), mais surtout de donner des garanties sur la précision de l'estimation finale. Dans ce sens, le paragraphe VI.2 liste les propriétés que l'on peut attendre a minima d'un estimateur. La section VI.3 détaille quand à elle la construction d'intervalles de confiance.

## VI.2 Estimation ponctuelle

### VI.2.1 Propriétés d'un (bon) estimateur

#### a) Consistance

**Définition 22** Un  $n$ -échantillon aléatoire issu d'une v.a.r.  $X$  est un ensemble  $(X_1, \dots, X_n)$  de  $n$  v.a.r. indépendantes et de même loi que  $X$ .

Soit  $\theta$  un paramètre associé à la loi de  $X$ , par exemple  $\theta = \mathbb{E}(X)$  ou  $\theta = \text{Var}(X)$ . À partir de l'observation d'un échantillon aléatoire  $(X_1, \dots, X_n)$ , on souhaite estimer le paramètre  $\theta$ .

**Définition 23** Un estimateur  $\hat{\theta}_n$  de  $\theta$  est une fonction qui dépend uniquement du  $n$ -échantillon  $(X_1, \dots, X_n)$ . Il est dit consistant s'il est "proche" de  $\theta$  au sens de la convergence en probabilité, i.e. pour tout  $\epsilon > 0$ ,

$$\mathbb{P}\left(|\hat{\theta}_n - \theta| > \epsilon\right) \xrightarrow[n \rightarrow +\infty]{} 0.$$

Dans l'exemple de l'introduction, la quantité  $\frac{1}{n} \sum_{i=1}^n X_i$  est un estimateur consistant de  $p$  et si, par exemple, on a observé 21 pièces défectueuses sur un lot de 1500 pièces prélevées, l'estimation ponctuelle de  $p$  obtenue est  $\bar{x}_n = 21/1500 = 1,4\%$ .

De manière plus générale, pour estimer l'espérance  $m$  des variables aléatoires  $X_i$ , on utilise souvent la moyenne empirique

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

La LGN garantit en effet la convergence en probabilité vers l'espérance  $m = \mathbb{E}(X_1)$ .

**Exemple:** Considérons une v.a.r.  $X$  représentant le nombre de gripes attrapées par une personne en un an. On peut supposer que  $X$  suit une loi de Poisson de paramètre  $\lambda > 0$ . Chercher la loi de  $X$ , c'est chercher  $\lambda$ , qui n'est autre que l'espérance mathématique de  $X$ . Par conséquent, la LGN nous indique que  $\bar{X}_n$  est un estimateur consistant de  $\lambda$ : pour tout  $\epsilon > 0$ ,

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \lambda\right| \geq \epsilon\right) \xrightarrow[n \rightarrow +\infty]{} 0.$$

L'inégalité de Chebychev permet quant à elle de démontrer le théorème suivant:

**Théorème 6** Soit  $\hat{\theta}_n$  un estimateur de  $\theta$ . Si l'on a:

$$\lim_{n \rightarrow +\infty} \mathbb{E}(\hat{\theta}_n) = \theta \quad \text{et} \quad \lim_{n \rightarrow +\infty} \text{Var}(\hat{\theta}_n) = 0,$$

alors  $\hat{\theta}_n$  est un estimateur consistant de  $\theta$ .



b) **Biais**

**Définition 24** Soit  $\hat{\theta}_n$  un estimateur consistant d'un paramètre  $\theta$ . On appelle **biais** la quantité  $\mathbb{E}(\hat{\theta}_n) - \theta$ . L'estimateur  $\hat{\theta}_n$  est dit **sans biais** si  $\mathbb{E}(\hat{\theta}_n) = \theta$ , et **biaisé** sinon.

**Exemple:** La moyenne empirique  $\bar{X}_n$  est un estimateur consistant et sans biais de l'espérance mathématique  $m$ .

c) **Écart quadratique moyen**

Notons que l'on a

$$\begin{aligned} \mathbb{E} \left\{ (\hat{\theta}_n - \theta)^2 \right\} &= \mathbb{E} \left\{ (\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n) + \mathbb{E}(\hat{\theta}_n) - \theta)^2 \right\} \\ &= \mathbb{E} \left\{ (\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n))^2 + (E(\hat{\theta}_n) - \theta)^2 + 2(\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n))(E(\hat{\theta}_n) - \theta) \right\} \\ &= \text{Var}(\hat{\theta}_n) + (\text{biais})^2, \end{aligned}$$

car le terme  $\mathbb{E} \left\{ (\hat{\theta}_n - \mathbb{E}(\hat{\theta}_n))(\mathbb{E}(\hat{\theta}_n) - \theta) \right\}$  est nul. Ainsi, pour rendre l'écart quadratique moyen  $\mathbb{E} \left\{ (\hat{\theta}_n - \theta)^2 \right\}$  le plus petit possible, il faut que

- $\mathbb{E}(\hat{\theta}_n) = \theta$ , donc choisir un estimateur sans biais,
- la variance  $\text{Var}(\hat{\theta}_n)$  soit faible.

On choisira donc, parmi les estimateurs consistants et sans biais, celui qui a la variance la plus petite. En d'autres termes, si  $\hat{\theta}_n$  est un estimateur consistant et sans biais de  $\theta$ , on a tout intérêt à ce que  $\hat{\theta}_n$  ne varie pas trop autour de sa moyenne. Cette propriété traduit ce que l'on appelle l'efficacité de l'estimateur.

La théorie de l'estimation, qui a émergé au cours du XXème siècle permet, dans des cadres assez généraux, de choisir parmi toutes les statistiques possibles le 'meilleur' estimateur consistant, c'est-à-dire celui qui donnera une estimation ponctuelle la plus proche possible du paramètre d'intérêt et ceci quel que soit l'échantillon.

## VI.2.2 Estimateur d'une moyenne $m$ ou d'une proportion

On considère un  $n$ -échantillon  $(X_1, \dots, X_n)$  issu d'une loi de moyenne  $m$  et de variance  $\sigma^2$ , toutes deux inconnues.

1. d'après la LGN, la moyenne empirique  $\bar{X}_n$  est un estimateur consistant de  $m$ .
2. l'estimateur  $\bar{X}_n$  est sans biais.
3. par indépendance:  $\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$ .
4. loi de  $\bar{X}_n$ :
  - si  $X \sim \mathcal{N}(m, \sigma^2)$ , alors  $\bar{X}_n \sim \mathcal{N}(m, \sigma^2/n)$ .
  - lorsque  $n$  est grand, d'après le TCL, la loi de  $\bar{X}_n$  est approximée par une loi normale  $\mathcal{N}(m, \sigma^2/n)$ .

L'estimation d'une proportion  $p$  est un cas particulier du précédent, au sens où les v.a.r.  $X_i$  considérées sont de Bernoulli de paramètre  $p$ .

### VI.2.3 Estimateur de la variance

**Définition 25** La variance empirique associée à un  $n$ -échantillon  $(X_1, \dots, X_n)$  est définie par

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

**Définition 26** Soit  $(Y_1, \dots, Y_n)$  un  $n$ -échantillon de v.a.r. de loi  $\mathcal{N}(0, 1)$ . On appelle **loi du chi-deux** à  $n$  degrés de liberté la loi de la v.a.r.  $\sum_{i=1}^n Y_i^2$ , et on la note  $\chi_{(n)}^2$ .

**Propriétés de la variance empirique:**

1.  $S_n^2$  est un estimateur consistant de la variance  $\sigma^2$ .
2.  $S_n^2$  est sans biais.
3. loi de  $S_n^2$ : pas de résultat général. Cependant, si  $X \sim \mathcal{N}(m, \sigma^2)$ , alors la v.a.r.  $\frac{n-1}{\sigma^2} S_n^2$  suit une loi du chi-deux à  $n-1$  degrés de liberté  $\chi_{(n-1)}^2$ .

**Remarque:** Puisque  $\mathbb{E}(Y_i) = 0$ , on a  $\mathbb{E}(Y_i^2) = \text{Var}(Y_i) = 1$ . Si  $V$  suit une loi  $\chi_{(n)}^2$ , alors

$$\mathbb{E}(V) = \mathbb{E}(Y_1^2 + \dots + Y_n^2) = n.$$

Ainsi on retrouve le fait que  $S_n^2$  est un estimateur consistant et sans biais de  $\sigma^2$ :

$$\mathbb{E}(S_n^2) = \frac{\sigma^2}{n-1} \mathbb{E}(\chi_{(n-1)}^2) = \sigma^2.$$

## VI.3 Estimation par intervalle de confiance

Les estimations ponctuelles n'apportent pas d'information sur la précision des résultats, c'est-à-dire qu'elles ne tiennent pas compte des fluctuations d'échantillonnage. Pour évaluer la confiance que l'on peut avoir en une valeur, il est nécessaire de déterminer un intervalle contenant, avec une certaine probabilité fixée au préalable, la vraie valeur du paramètre: c'est l'estimation par intervalle de confiance.

**Exemple.** On considère une échantillon d'observations  $X_1, \dots, X_n$  tel que  $X_i \sim \mathcal{N}(m, \sigma^2)$  pour tout  $i \in \{1, \dots, n\}$ , le paramètre  $m$  étant inconnu. On a vu que  $\bar{X}_n \sim \mathcal{N}(m, \sigma^2/n)$ . En particulier, on a donc  $P(\bar{X}_n = m) = 0$ : la probabilité que notre estimation soit juste... est égale à 0!! Pourtant, on sent bien que  $\bar{X}_n$  se rapproche de  $m$  quand  $n$  croît: il s'agit donc d'essayer de quantifier le crédit que l'on peut accorder à notre estimateur.

Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon aléatoire et  $\theta$  un paramètre inconnu de la loi des  $X_i$ .

**Définition 27** Soit  $\alpha \in ]0, 1[$ . S'il existe des v.a.r.  $\theta_{\min}(X_1, \dots, X_n)$  et  $\theta_{\max}(X_1, \dots, X_n)$  telles que

$$\mathbb{P}(\theta \in [\theta_{\min}(X_1, \dots, X_n), \theta_{\max}(X_1, \dots, X_n)]) = 1 - \alpha,$$

on dit alors que  $[\theta_{\min}(X_1, \dots, X_n), \theta_{\max}(X_1, \dots, X_n)]$  est un intervalle de confiance pour  $\theta$ , avec coefficient de sécurité (ou de niveau de confiance)  $1 - \alpha$ . On le note  $IC_{1-\alpha}(\theta)$ .

Dans la pratique, on peut prendre par exemple  $\alpha = 5\%$ , ce qui nous donne un IC à 95%. Cela signifie qu'il y a 95% de chance que la valeur inconnue  $\theta$  soit comprise entre  $\theta_{\min}(x_1, \dots, x_n)$  et  $\theta_{\max}(x_1, \dots, x_n)$ .

### VI.3.1 Intervalle de confiance pour la moyenne d'un échantillon gaussien (variance connue)

Soit  $(X_1, \dots, X_n)$  un  $n$ -échantillon de v.a.r. de loi  $\mathcal{N}(m, \sigma^2)$ . Afin d'estimer  $m$ , on utilise la moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  qui a pour loi  $\mathcal{N}(m, \sigma^2/n)$ . Il en résulte que

$$\sqrt{n} \left( \frac{\bar{X}_n - m}{\sigma} \right) \sim \mathcal{N}(0, 1),$$

et que

$$\mathbb{P} \left( -z_{1-\alpha/2} \leq \sqrt{n} \left( \frac{\bar{X}_n - m}{\sigma} \right) \leq z_{1-\alpha/2} \right) = 1 - \alpha,$$

où  $z_{1-\alpha/2}$  désigne le  $1 - \alpha/2$ -quantile d'une variable gaussienne centrée réduite. Ceci équivaut à

$$\mathbb{P} \left( \bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \leq m \leq \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right) = 1 - \alpha.$$

On obtient donc un IC pour l'espérance  $m$  avec coefficient de sécurité  $1 - \alpha$  dans le cas où  $\sigma$  est connu: il s'agit de l'intervalle aléatoire

$$\left[ \bar{X}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{X}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right].$$

Ainsi, dans les calculs, l'IC est donné par

$$IC_{1-\alpha}(m) = \left[ \bar{x}_n - z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x}_n + z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}} \right],$$

où  $\bar{x}_n$  est l'estimation ponctuelle de  $m$  associée à la réalisation du  $n$ -échantillon  $(X_1, \dots, X_n)$ .

### VI.3.2 Intervalle de confiance pour la moyenne d'un échantillon gaussien (variance inconnue)

On estime la variance  $\sigma^2$ , supposée inconnue, par la variance empirique

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2.$$

Pour la construction d'un intervalle de confiance dans ce cadre, il est alors nécessaire de remplacer dans les formules précédentes la variance par sa version empirique  $S_n^2$ . Il faut donc considérer non plus la quantité  $\sqrt{n} \left( \frac{\bar{X}_n - m}{\sigma} \right)$  mais plutôt

$$\sqrt{n} \left( \frac{\bar{X}_n - m}{S_n} \right),$$

qui ne suit plus une loi normale mais une loi dite **de Student** à  $n - 1$  degrés de liberté, que l'on note  $\mathcal{T}_{n-1}$ . La densité de la loi de Student est une fonction paire, comme la loi normale  $\mathcal{N}(0, 1)$ . On dispose de tables pour obtenir les quantiles de cette loi. On en déduit donc que

$$\mathbb{P} \left( -t_{1-\alpha/2} \leq \sqrt{n} \left( \frac{\bar{X}_n - m}{S_n} \right) \leq t_{1-\alpha/2} \right) = 1 - \alpha,$$

ce qui équivaut à

$$\mathbb{P} \left( \bar{X}_n - t_{1-\alpha/2} \frac{S_n}{\sqrt{n}} \leq m \leq \bar{X}_n + t_{1-\alpha/2} \frac{S_n}{\sqrt{n}} \right) = 1 - \alpha.$$

On obtient donc un IC pour  $m$  avec coefficient de sécurité  $1 - \alpha$ , dans le cas où la variance  $\sigma^2$  est inconnue: il s'agit de l'intervalle aléatoire

$$\left[ \bar{X}_n - t_{1-\alpha/2} \frac{S_n}{\sqrt{n}}, \bar{X}_n + t_{1-\alpha/2} \frac{S_n}{\sqrt{n}} \right].$$

Ainsi, dans les calculs, l'IC est donné par

$$IC_{1-\alpha}(m) = \left[ \bar{x}_n - t_{1-\alpha/2} \frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{1-\alpha/2} \frac{s_n}{\sqrt{n}} \right],$$

où  $\bar{x}_n$  et  $s_n^2$  sont les estimations ponctuelles respectives de la moyenne  $m$  et de la variance  $\sigma^2$ .

### VI.3.3 Intervalle de confiance pour une proportion

Revenons à l'exemple de l'introduction de ce chapitre: on cherche à estimer la proportion  $p$  de composants défectueux produits par une usine. On prélève un lot de  $n$  composants et on note  $X_i$  la v.a.r. qui vaut 1 si la pièce  $i$  est défectueuse, et 0 sinon. On estime  $p$  par la moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ . Les v.a.r.  $X_i$  étant de Bernoulli, on peut alors utiliser l'approximation donnée par le TCL. Soit  $Z \sim \mathcal{N}(0, 1)$  et  $z_{1-\alpha/2}$  le quantile d'ordre  $1 - \alpha/2$  de la loi  $\mathcal{N}(0, 1)$ . Par le TCL,

$$\frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \xrightarrow{\mathcal{L}} Z \sim \mathcal{N}(0, 1), \text{ lorsque } n \rightarrow +\infty.$$

Ceci implique que

$$\mathbb{P} \left( -z_{1-\alpha/2} \leq \frac{\sum_{i=1}^n X_i - np}{\sqrt{np(1-p)}} \leq z_{1-\alpha/2} \right) \xrightarrow{n \rightarrow +\infty} \mathbb{P}(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}) = 1 - \alpha,$$

c'est-à-dire

$$\mathbb{P} \left( \bar{X}_n - z_{1-\alpha/2} \frac{\sqrt{p(1-p)}}{\sqrt{n}} \leq p \leq \bar{X}_n + z_{1-\alpha/2} \frac{\sqrt{p(1-p)}}{\sqrt{n}} \right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha.$$

Ceci ne fournit pas un IC pour  $p$  car les bornes de l'intervalle dépendent de  $p$ . Mais on peut montrer que l'on a le même résultat de convergence, en remplaçant  $p$  dans les bornes de l'intervalle par son estimateur consistant  $\bar{X}_n$ . On obtient alors

$$\mathbb{P} \left( \bar{X}_n - z_{1-\alpha/2} \frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}} \leq p \leq \bar{X}_n + z_{1-\alpha/2} \frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}} \right) \xrightarrow{n \rightarrow +\infty} 1 - \alpha.$$

On dit que l'intervalle

$$\left[ \bar{X}_n - z_{1-\alpha/2} \frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}}, \bar{X}_n + z_{1-\alpha/2} \frac{\sqrt{\bar{X}_n(1-\bar{X}_n)}}{\sqrt{n}} \right]$$

est un IC asymptotique pour le paramètre  $p$ , de coefficient de sécurité  $1 - \alpha$ .

Pour  $\alpha = 5\%$ , on lit dans les tables  $z_{1-\alpha/2} = z_{97,5\%} = 1,96$ . Ainsi, on en déduit que si l'on a observé en pratique 21 pièces défectueuses sur 1500 (c'est-à-dire que l'on remplace dans les calculs la moyenne empirique aléatoire  $\bar{X}_n$  par l'estimation ponctuelle  $\bar{x}_n = 21/1500$ ), l'intervalle de confiance asymptotique pour  $p$  est  $[0.0081, 0.0199]$ .

**Remarque:** Si les v.a.r.  $X_1, \dots, X_n$  ne sont pas gaussiennes mais que  $n$  est assez grand (en pratique supérieur à 30), alors le TCL nous garantit que la moyenne empirique suit approximativement une loi gaussienne. Ceci permet de considérer un très large éventail de situations. Dans le cas particulier où l'intervalle de confiance dépendrait d'une quantité inconnue, il est toujours possible de remplacer cette dernière par un estimateur consistant: la convergence en loi est conservée (lemme de Slutsky).

### VI.3.4 Intervalle de confiance pour la variance d'une population gaussienne

On sait que

$$S_n^2 \sim \frac{\sigma^2}{n-1} \chi_{(n-1)}^2 \text{ et que } \mathbb{E}(S_n^2) = \frac{\sigma^2}{n-1} \mathbb{E}(\chi_{(n-1)}^2) = \sigma^2,$$

c'est-à-dire que  $S_n^2$  est un estimateur sans biais de  $\sigma^2$ . Qui plus est, la LGN nous donne la consistance de  $S_n^2$ . De plus, on lit dans des tables les quantiles d'ordre  $\alpha/2$  et  $1 - \alpha/2$  de la loi du  $\chi_{(n-1)}^2$ , respectivement notés  $v_{\alpha/2}$  et  $v_{1-\alpha/2}$  (il est normal que les quantiles qui nous intéressent ne soient pas opposés car la densité de cette loi n'est pas paire, à l'inverse de la loi normale centrée réduite). On obtient alors

$$\mathbb{P}\left(v_{\alpha/2} \leq \frac{n-1}{\sigma^2} S_n^2 \leq v_{1-\alpha/2}\right) = 1 - \alpha.$$

Ceci équivaut à

$$\mathbb{P}\left(\frac{(n-1)S_n^2}{v_{1-\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)S_n^2}{v_{\alpha/2}}\right) = 1 - \alpha.$$

On obtient donc un IC pour la variance  $\sigma^2$  avec coefficient de sécurité  $1 - \alpha$ : il s'agit de l'intervalle aléatoire

$$\left[ \frac{(n-1)S_n^2}{v_{1-\alpha/2}}, \frac{(n-1)S_n^2}{v_{\alpha/2}} \right].$$

Ainsi, dans les calculs, l'IC est donné par

$$IC_{1-\alpha}(\sigma^2) = \left[ \frac{(n-1)s_n^2}{v_{1-\alpha/2}}, \frac{(n-1)s_n^2}{v_{\alpha/2}} \right],$$

où  $s_n^2$  est l'estimation ponctuelle de  $\sigma^2$  associée à la réalisation du  $n$ -échantillon  $(X_1, \dots, X_n)$ :

$$s_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2.$$



# Chapitre VII

## Tests d'hypothèses

La théorie des tests propose une démarche générale permettant:

- de confronter une hypothèse avec la réalité, ou plus exactement, avec ce que l'on perçoit de la réalité à travers les observations à disposition;
- de prendre une décision à la suite de cette confrontation.

Si les problèmes traités par l'estimation (ponctuelle ou par intervalle de confiance) sont de type quantitatif, i.e. conduisent à un résultat numérique, ceux traités par les tests d'hypothèses sont d'ordre qualitatif, i.e. conduisent à une réponse du type rejet/acceptation de l'hypothèse statistique effectuée.

### VII.1 Exemple introductif

Des relevés effectués pendant de nombreuses années ont permis d'établir que le niveau naturel des pluies dans la Beauce  $X$  en millimètres par an suit une loi normale  $\mathcal{N}(600, 100^2)$ .

Des entrepreneurs, surnommés faiseurs de pluie, prétendaient pouvoir augmenter de  $50mm$  le niveau moyen de pluie, ceci par insémination des nuages au moyen d'iodure d'argent. Leur procédé fut mis à l'essai entre 1951 et 1959 et on releva les hauteurs de pluies suivantes:

Année	1951	1952	1953	1954	1955	1956	1957	1958	1959
mm	510	614	780	512	501	534	603	788	650

Que pouvait-on en conclure? Deux hypothèses s'affrontaient: ou bien l'insémination était sans effet, ou bien elle augmentait réellement le niveau moyen de pluie de  $50mm$ . Ces hypothèses pouvaient se formaliser comme suit, si  $m$  désigne l'espérance mathématique de  $X$ , v.a.r. égale au niveau annuel de pluie.

$$\begin{cases} H_0 : m = 600mm \\ H_1 : m = 650mm. \end{cases}$$

Les agriculteurs hésitaient à opter pour le procédé forcément onéreux des faiseurs de pluie. Ainsi, il fallait donc que l'expérience puisse les convaincre, c'est-à-dire que les faits observés contredisent nettement l'hypothèse  $H_0$ , dite "hypothèse nulle" ( $H_1$  s'appelle "hypothèse alternative"). Les agriculteurs n'étaient donc décidés à abandonner  $H_0$  qu'en présence de faits expérimentaux traduisant une éventualité improbable compte-tenu de  $H_0$ .

Par essence, l'objectif d'un test n'est pas de déterminer si  $H_0$  est fondamentalement vraie ou non, mais plutôt de voir si  $H_0$  est une hypothèse cohérente avec les données observées. On souhaite donc voir  $H_0$  rejetée uniquement dans le cas où les observations la rendent invraisemblable. Ce qui amène généralement à la pratique suivante: on fixe une valeur  $\alpha \in ]0; 1[$ , appelée risque de première espèce, et l'on impose au test d'être tel que si  $H_0$  est vraie, la probabilité de rejeter à tort  $H_0$  soit inférieure à  $\alpha$ .

Dans notre exemple, les agriculteurs choisirent  $\alpha = 0,05$ , c'est-à-dire qu'ils assumaient le risque de se tromper dans 5 cas sur 100, en croyant les promesses des faiseurs de pluie, alors que ces faiseurs de pluie étaient en réalité des charlatans.

Comment décider? Puisqu'il s'agit de "tester" la moyenne  $m$ , il est naturel de s'intéresser à la moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$  que l'on sait consistante vers  $m$  par la LGN. Pour ce test,  $\bar{X}_n$  est appelée *la statistique de test*. Alors, si  $H_0$  est vraie, comme l'expérience porte sur  $n = 9$  années, on a  $\bar{X}_n \sim \mathcal{N}(600, 100^2/9)$ .

Les données relevées indiquent que  $\bar{x}_n = 610,2mm$  (c'est la moyenne obtenue sur toutes les années). L'hypothèse  $H_1$  proposant une moyenne  $m$  alternative supérieure à 600, il est naturel de prendre la règle de décision suivante:

- Si  $\bar{X}_n$  est plus grande qu'un seuil  $t$  (à déterminer), on rejette  $H_0$ ,
- Si  $\bar{X}_n < t$ , on accepte  $H_0$ .

Reste maintenant à choisir le paramètre  $t$  de telle sorte que l'erreur de première espèce soit inférieure ou égale à 5%, autrement dit

$$P_{H_0}(\text{rejeter } H_0) = P_{H_0}(\bar{X}_n > t) \leq 5\%.$$

Il est alors facile de calculer cette dernière probabilité car sous  $H_0$ , la v.a.r.  $Z = \frac{\bar{X}_n - 600}{100/3}$  suit une loi normale  $\mathcal{N}(0, 1)$ :

$$\begin{aligned} P_{H_0}(\bar{X}_n > t) &= \mathbb{P}\left(Z > \frac{t - 600}{100/3}\right) = 5\%, \\ &\Leftrightarrow \frac{t - 600}{100/3} = 1,64, \\ &\Leftrightarrow t = 600 + 1,64 \times \frac{100}{3} \simeq 655. \end{aligned}$$

**Conclusion:** Si la valeur observée de  $\bar{X}_n$  est supérieure à 655, on pourra en conclure que l'amélioration est significative d'un point de vue statistique (et on a au maximum 5% de chances de se tromper en prenant cette décision). Si par contre,  $\bar{x}_n < 655$ , on pourra émettre des doutes sur l'évolution de la pluviométrie. Sur les données de l'énoncé, on trouve  $\bar{x}_n = 610,2$ : on ne rejette donc pas  $H_0$ .

## VII.2 Démarche générale pour la construction d'un test

Dans un test d'hypothèse statistique, il y a deux manières de se tromper:

1 - la possibilité de rejeter l'hypothèse  $H_0$  alors qu'elle est vraie. Le  $\alpha$  donné précédemment borne la probabilité de se tromper dans ce sens, c'est le risque de première espèce.

2 - la possibilité d'accepter  $H_0$  alors qu'elle est fautive. La probabilité de se tromper dans ce sens est le risque de deuxième espèce, et est notée  $\beta$ .

Ainsi, dans notre exemple, la première façon de se tromper est de croire les faiseurs de pluie, alors qu'ils ne sont pour rien dans le résultat obtenu (rejeter  $H_0$  alors qu'elle est vraie), tandis que la seconde manière de se tromper est de ne pas croire les faiseurs de pluie, alors que leur méthode est bonne et que seul le hasard (malencontreux pour eux), dû au faible nombre d'observations, a donné des résultats insuffisants pour convaincre les agriculteurs (accepter  $H_0$  alors qu'elle est fautive).

Supposons maintenant que les faiseurs de pluie ont raison. Alors  $\bar{X}_n$  suit une loi normale  $\mathcal{N}(650, 100^2/9)$ ,



c'est-à-dire que la moyenne en millimètres n'est plus 600 mais 650, sous l'hypothèse  $H_1$ . Ainsi, la probabilité de se tromper est:

$$\beta = \mathbb{P}_{H_1}(\bar{X}_n < 655) = \mathbb{P}\left(Z < \frac{655 - 650}{100/3}\right) = \mathbb{P}(Z < 0,15) = 0,56,$$

ce qui est considérable.

Il faut remarquer au cours de cet exemple le rôle particulier joué par  $H_0$ . Si la forme de la région de rejet est indiquée par la nature de  $H_1$  (650 plus grand que 600), la valeur de  $t$  ne dépend que de  $H_0$ . De plus, les deux hypothèses ne jouent pas des rôles symétriques,  $t$  étant déterminé par  $H_0$  et  $\alpha$ , tandis que  $\beta$  est déterminé par la considération supplémentaire de  $H_1$ .

Enfin il faut remarquer que les deux types d'erreur possibles ne sont en général pas du tout de la même importance. En effet, rejeter l'hypothèse  $H_0$  alors qu'elle est vraie (risque de première espèce, qui est maîtrisé) est beaucoup plus coûteux que de la conserver à tort (risque de deuxième espèce, non maîtrisé).

**Exemple:** Soit  $m$  la moyenne du niveau de radioactivité en picocuries par litre. La valeur  $m = 5$  est considérée comme la valeur critique entre eau potable et non potable. On peut donc souhaiter tester  $H_0 : "m \geq 5"$  contre  $H_1 : "m < 5"$ . L'erreur de première espèce a alors pour conséquence de laisser boire de l'eau toxique, alors que l'erreur de deuxième espèce conduit seulement à jeter de l'eau potable ... Dans tous les cas, il faut trouver une procédure de test qui minimise ces deux types d'erreur.

De manière très générale, on retiendra le schéma suivant lors de l'élaboration d'un test statistique:

1. Choix de  $H_0$  et de  $H_1$ . Fixer  $\alpha$ .
2. Détermination de la statistique de test.
3. Construction d'une règle de décision **cohérente avec les hypothèses**.
4. Calibrage de cette règle en fonction de  $\alpha$  et  $H_0$ .
5. Calcul de la valeur observée de la statistique de test.
6. Conclusion: rejet ou acceptation de  $H_0$  au risque  $\alpha$ .
7. Calcul de la puissance:  $1 - \beta$ .

## VII.3 Test sur la moyenne d'un échantillon gaussien (variance connue)

On suppose que l'on observe les réalisations d'un  $n$ -échantillon  $(X_1, \dots, X_n)$  issu d'une loi  $\mathcal{N}(m, \sigma^2)$  avec  $m$  inconnue et  $\sigma^2$  connue. On se donne donc l'hypothèse nulle  $H_0 : "m = m_0"$ ,  $m_0$  étant une valeur donnée par l'énoncé ou provenant de connaissances théoriques. Ensuite, il existe trois types d'hypothèses alternatives  $H_1$ :

- $H_1 : "m > m_0"$ , test unilatéral à droite;
- $H_1 : "m < m_0"$ , test unilatéral à gauche;
- $H_1 : "m \neq m_0"$ , test bilatéral.

### VII.3.1 Test unilatéral à droite

$H_1$  : “ $m > m_0$ ” (cas de l'exemple introductif).

- Statistique de test: la moyenne empirique  $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ .
- Règle de décision

$$\text{Rejet de } H_0 \iff \bar{X}_n > t,$$

cette dernière étant donnée par la forme de  $H_1$ .

- Calcul de  $\bar{x}_n$ , la valeur observée de  $\bar{X}_n$ .
- Détermination de  $t$ . Sous  $H_0$ , on a que  $\bar{X}_n \sim \mathcal{N}(m_0, \frac{\sigma^2}{n})$ . On définit alors  $Z = \frac{\bar{X}_n - m_0}{\sigma/\sqrt{n}}$ , qui suit  $\mathcal{N}(0, 1)$ . Ainsi on a:

$$\mathbb{P}_{H_0}(\bar{X}_n > t) = \mathbb{P}\left(Z > \frac{\sqrt{n}(t - m_0)}{\sigma}\right).$$

On trouve la valeur de  $t$  à partir de  $\alpha$  et de la table de la loi  $\mathcal{N}(0, 1)$ .

- Conclusion par comparaison de  $t$  et de  $\bar{x}_n$ .

### VII.3.2 Test unilatéral à gauche

$H_1$  : “ $m < m_0$ ”. Même méthode que précédemment, sauf que l'on remplace l'événement  $\{\bar{X}_n > t\}$  par  $\{\bar{X}_n < t\}$ .

### VII.3.3 Test bilatéral

$H_1$  : “ $m \neq m_0$ ”. Même méthode que précédemment, sauf que l'on remplace l'événement  $\{\bar{X}_n > t\}$  par  $\{|\bar{X}_n - m_0| > t\}$ .

## VII.4 Test sur la moyenne d'un échantillon gaussien (variance inconnue)

C'est le test de Student. On suppose que l'on observe les réalisations d'un  $n$ -échantillon  $(X_1, \dots, X_n)$  issu d'une loi  $\mathcal{N}(m, \sigma^2)$  avec  $m$  et  $\sigma^2$  inconnues. Pour cette section, on considère uniquement la construction d'un test bilatéral, les autres cadres suivant exactement le même principe.

On teste  $H_0$  : “ $m = m_0$ ” contre  $H_1$  : “ $m \neq m_0$ ”. Comme  $\sigma^2$  est inconnue, on l'estime par la variance empirique  $S_n^2$ . Par conséquent, la règle de décision est la suivante:

$$\text{Rejet de } H_0 \iff |\bar{X}_n - m_0| > t \iff \frac{|\bar{X}_n - m_0|}{S_n/\sqrt{n}} > \frac{t}{S_n/\sqrt{n}}.$$

Sous  $H_0$ , on sait que  $T_n = \frac{\bar{X}_n - m_0}{S_n/\sqrt{n}}$  suit la loi de Student à  $n - 1$  degrés de liberté. Ainsi, on peut trouver la valeur de  $t$  d'après les tables de la loi de Student et la valeur de  $\alpha$ .

Les tests unilatéraux se font de la même manière.

## VII.5 Cas d'une population non gaussienne

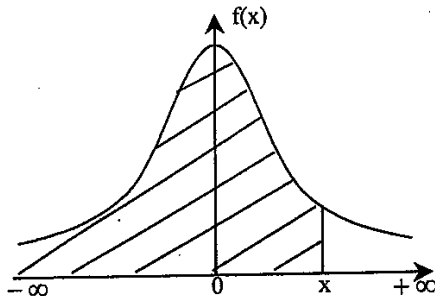
Si le nombre  $n$  d'observations est suffisamment grand, on se ramène à une loi normale grâce au TLC. La loi de  $\bar{X}_n$  est approximée par  $\mathcal{N}\left(m, \frac{\sigma^2}{n}\right)$ . Les procédures de test sont alors les mêmes.

# Chapitre VIII

## Annexe 1: Tables statistiques

### Loi Normale centrée réduite

Probabilité de trouver une valeur inférieure à x.



$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

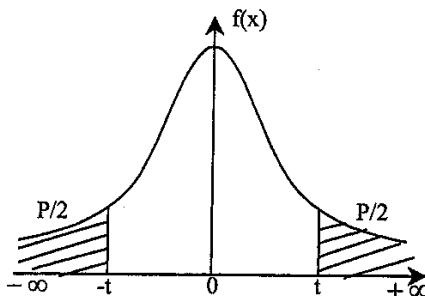
X	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998

Table pour les grandes valeurs de x :

x	3	3,2	3,4	3,6	3,8	4	4,2	4,4	4,6	4,8
F(x)	0,99865003	0,99931280	0,99966302	0,99984085	0,99992763	0,99996831	0,99998665	0,99999458	0,99999789	0,99999921

## Loi de Student

Valeurs de  $t$  ayant la probabilité  $P$  d'être dépassées en valeur absolue.



$\nu$   $P$	90%	80%	70%	60%	50%	40%	30%	20%	10%	5%	1%
1	0,1584	0,3249	0,5095	0,7265	1,0000	1,3764	1,9626	3,0777	6,3137	12,7062	63,6559
2	0,1421	0,2887	0,4447	0,6172	0,8165	1,0607	1,3862	1,8856	2,9200	4,3027	9,9250
3	0,1366	0,2767	0,4242	0,5844	0,7649	0,9785	1,2498	1,6377	2,3534	3,1824	5,8408
4	0,1338	0,2707	0,4142	0,5686	0,7407	0,9410	1,1896	1,5332	2,1318	2,7765	4,6041
5	0,1322	0,2672	0,4082	0,5594	0,7267	0,9195	1,1558	1,4759	2,0150	2,5706	4,0321
6	0,1311	0,2648	0,4043	0,5534	0,7176	0,9057	1,1342	1,4398	1,9432	2,4469	3,7074
7	0,1303	0,2632	0,4015	0,5491	0,7111	0,8960	1,1192	1,4149	1,8946	2,3646	3,4995
8	0,1297	0,2619	0,3995	0,5459	0,7064	0,8889	1,1081	1,3968	1,8595	2,3060	3,3554
9	0,1293	0,2610	0,3979	0,5435	0,7027	0,8834	1,0997	1,3830	1,8331	2,2622	3,2498
10	0,1289	0,2602	0,3966	0,5415	0,6998	0,8791	1,0931	1,3722	1,8125	2,2281	3,1693
11	0,1286	0,2596	0,3956	0,5399	0,6974	0,8755	1,0877	1,3634	1,7959	2,2010	3,1058
12	0,1283	0,2590	0,3947	0,5386	0,6955	0,8726	1,0832	1,3562	1,7823	2,1788	3,0545
13	0,1281	0,2586	0,3940	0,5375	0,6938	0,8702	1,0795	1,3502	1,7709	2,1604	3,0123
14	0,1280	0,2582	0,3933	0,5366	0,6924	0,8681	1,0763	1,3450	1,7613	2,1448	2,9768
15	0,1278	0,2579	0,3928	0,5357	0,6912	0,8662	1,0735	1,3406	1,7531	2,1315	2,9467
16	0,1277	0,2576	0,3923	0,5350	0,6901	0,8647	1,0711	1,3368	1,7459	2,1199	2,9208
17	0,1276	0,2573	0,3919	0,5344	0,6892	0,8633	1,0690	1,3334	1,7396	2,1098	2,8982
18	0,1274	0,2571	0,3915	0,5338	0,6884	0,8620	1,0672	1,3304	1,7341	2,1009	2,8784
19	0,1274	0,2569	0,3912	0,5333	0,6876	0,8610	1,0655	1,3277	1,7291	2,0930	2,8609
20	0,1273	0,2567	0,3909	0,5329	0,6870	0,8600	1,0640	1,3253	1,7247	2,0860	2,8453
21	0,1272	0,2566	0,3906	0,5325	0,6864	0,8591	1,0627	1,3232	1,7207	2,0796	2,8314
22	0,1271	0,2564	0,3904	0,5321	0,6858	0,8583	1,0614	1,3212	1,7171	2,0739	2,8188
23	0,1271	0,2563	0,3902	0,5317	0,6853	0,8575	1,0603	1,3195	1,7139	2,0687	2,8073
24	0,1270	0,2562	0,3900	0,5314	0,6848	0,8569	1,0593	1,3178	1,7109	2,0639	2,7970
25	0,1269	0,2561	0,3898	0,5312	0,6844	0,8562	1,0584	1,3163	1,7081	2,0595	2,7874
26	0,1269	0,2560	0,3896	0,5309	0,6840	0,8557	1,0575	1,3150	1,7056	2,0555	2,7787
27	0,1268	0,2559	0,3894	0,5306	0,6837	0,8551	1,0567	1,3137	1,7033	2,0518	2,7707
28	0,1268	0,2558	0,3893	0,5304	0,6834	0,8546	1,0560	1,3125	1,7011	2,0484	2,7633
29	0,1268	0,2557	0,3892	0,5302	0,6830	0,8542	1,0553	1,3114	1,6991	2,0452	2,7564
30	0,1267	0,2556	0,3890	0,5300	0,6828	0,8538	1,0547	1,3104	1,6973	2,0423	2,7500
40	0,1265	0,2550	0,3881	0,5286	0,6807	0,8507	1,0500	1,3031	1,6839	2,0211	2,7045
50	0,1263	0,2547	0,3875	0,5278	0,6794	0,8489	1,0473	1,2987	1,6759	2,0086	2,6778
60	0,1262	0,2545	0,3872	0,5272	0,6786	0,8477	1,0455	1,2958	1,6706	2,0003	2,6603
80	0,1261	0,2542	0,3867	0,5265	0,6776	0,8461	1,0432	1,2922	1,6641	1,9901	2,6387
100	0,1260	0,2540	0,3864	0,5261	0,6770	0,8452	1,0418	1,2901	1,6602	1,9840	2,6259
120	0,1259	0,2539	0,3862	0,5258	0,6765	0,8446	1,0409	1,2886	1,6576	1,9799	2,6174
200	0,1258	0,2537	0,3859	0,5252	0,6757	0,8434	1,0391	1,2858	1,6525	1,9719	2,6006
$\infty$	0,1257	0,2533	0,3853	0,5244	0,6745	0,8416	1,0364	1,2816	1,6449	1,9600	2,5758

## Chapitre IX

### Annexe 2: Sujets des années précédentes

## Evaluation 1

Durée 45 minutes

### Exercice 1

On interroge des étudiants sur la signification du sigle ADN. A chacun, on propose trois réponses différentes:  $A$ ,  $B$  ou  $C$ , la réponse correcte étant la réponse  $A$ . Tout étudiant connaissant la réponse correcte la donne, sinon il choisit au hasard une des trois réponses proposées. On note  $\theta$  la probabilité qu'un étudiant connaisse la bonne réponse et  $p_A$  la probabilité qu'un étudiant interrogé donne la réponse  $A$ .

1. Exprimer  $\theta$  en fonction de  $p_A$ .
2. Quelle est la probabilité qu'une personne ayant choisi la réponse  $A$  connaisse réellement la signification du sigle ADN?
3. Sur un groupe de 24 étudiants, quelle est l'espérance du nombre de bonnes réponses (justifier).

### Exercice 2

Etant donnée une variable aléatoire  $X$ , on définit la fonction génératrice  $G_X$  de  $X$  comme

$$G_X : [0, 1] \rightarrow \mathbb{R}, \\ t \mapsto G_X(t) = \mathbb{E}[t^X].$$

1. Calculer la fonction génératrice d'une variable binomiale de paramètres  $n$  et  $p$ .
2. Même question pour une variable de Poisson de paramètre  $\lambda$ .
3. Soient  $X$  et  $Y$  deux variables de Poisson indépendantes de paramètres respectifs  $\lambda$  et  $\mu$ . On pose  $Z = X + Y$ . Exprimer  $G_{X+Y}$  en fonction de  $G_X$  et  $G_Y$ .
4. On admet que si deux variables aléatoires ont même fonction génératrice, alors elles ont même loi. Dédurre de la question précédente que  $X + Y$  suit une loi de Poisson de paramètre  $\lambda + \mu$ .

### Exercice 3

Soit  $f$  la fonction définie par

$$f(x) = \begin{cases} 1/2 & \text{si } x \in [-1; 0], \\ a & \text{si } x \in ]0, 1], \\ 0 & \text{sinon.} \end{cases}$$

1. Déterminer  $a$  pour que  $f$  définisse bien une densité de probabilité.
2. Soit  $X$  une variable aléatoire ayant pour densité  $f$ . On pose  $Z = X^2$ . Quelles sont les valeurs possibles de  $Z$ .
3. Calculer la fonction de répartition de  $Z$  et en déduire la densité de  $Z$ .

## Evaluation 2

Durée 1h30

*Tous les documents, calculatrices et téléphones portables sont interdits.*

*Le barème sur 20 est approximatif.*

### Exercice 1 (6 points)

Des électriciens interviennent dans un immeuble pour poser des tableaux électriques. Le fournisseur des ces tableaux a été choisi au hasard (probabilité 0.5) par le commanditaire des travaux parmi deux fabricants  $U$  et  $V$ . Une fois le fournisseur choisi, **tous les tableaux proviennent du même fabricant.**

Les tableaux du fournisseur  $U$  sont défectueux avec probabilité 0.1, ceux provenant du fournisseur  $V$  avec probabilité 0.3. Pour chaque fournisseur, ce mode de fonctionnement est indépendant d'un tableau à l'autre. On considère les événements suivants:  $U$  = "tous les tableaux proviennent du fournisseur  $U$ ",  $V$  = "tous les tableaux proviennent du fournisseur  $V$ " et  $A_i$  = "le  $i$ -ème tableau est défectueux".

1. Donner les valeurs de  $\mathbb{P}(A_1|U)$ ,  $\mathbb{P}(A_1 \cap V)$  et montrer que  $\mathbb{P}(A_1 \cap A_2|U) = (0.1)^2$ .
2. En déduire  $\mathbb{P}(A_1)$  et  $\mathbb{P}(A_1 \cap A_2)$ .
3. Les événements  $A_1$  et  $A_2$  sont-ils indépendants? (on pourra utiliser le fait que  $\mathbb{P}(A_1) = \mathbb{P}(A_2)$ ).
4. Pour tout entier  $n \in \mathbb{N}$ , calculer  $\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \frac{3^n + 1}{2 \times 10^n}$ . (Rappel: tous les tableaux proviennent du même fournisseur).
5. Sachant que les  $n$  premiers tableaux sont défectueux, quelle est la probabilité que le  $(n+1)$ -ème le soit également?
6. Sachant que les  $n$  premiers tableaux sont défectueux, quelle est la probabilité qu'ils proviennent du fournisseur  $U$ ?

### Exercice 2 (7 points)

Soit  $X$  une variable de loi exponentielle de paramètre 1. On pose  $Y = \exp(X/2)$ . On suppose que  $Y$  modélise le temps d'attente d'un bus (en minutes).

1. Montrer que la variable  $Y$  est toujours supérieure ou égale à 1.
2. Quelle est la probabilité d'attendre le bus plus de 20 minutes?
3. Déterminer la fonction de répartition de  $Y$ .
4. Déduire que la variable aléatoire  $Y$  admet comme densité de probabilité la fonction

$$f_Y(y) = \begin{cases} 2y^{-3} & \text{si } y \geq 1, \\ 0 & \text{si } y < 1. \end{cases}$$

5. Calculer le temps moyen d'attente du bus.
6. Sachant que l'on a déjà attendu le bus plus de 10 minutes, quelle est la probabilité que le bus n'arrive pas dans les 10 prochaines minutes?



**Exercice 3** (7 points)

On observe une suite de variables aléatoires  $X_1, \dots, X_n$  indépendantes et de même loi telle que pour tout  $i \in \{1, \dots, n\}$

$$P(X_i = 1) = p, \text{ et } P(X_i = -1) = 1 - p,$$

avec  $p \in ]0, 1[$  un paramètre inconnu. L'objectif de cet exercice est d'estimer  $p$ .

1. Calculer l'espérance et la variance de  $X_1$ .
2. On considère l'estimateur de  $p$  suivant:

$$\hat{p}_n = \frac{\bar{X}_n + 1}{2}, \text{ avec } \bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i.$$

Calculer l'espérance et la variance de l'estimateur  $\hat{p}_n$ .

3. Montrer que pour tout nombre réel  $c > 0$  on a

$$\mathbb{P}(|\hat{p}_n - p| > c) \leq \frac{p(1-p)}{nc^2}.$$

4. Montrer que  $\hat{p}_n$  est un estimateur consistant et sans biais de  $p$ .
5. Énoncer le Théorème de la limite centrale. Donner la loi asymptotique de  $\bar{X}_n$  et en déduire celle de  $\hat{p}_n$ .
6. Montrer que la variable aléatoire

$$D_n = (\hat{p}_n - p) \sqrt{\frac{n}{p(1-p)}}$$

suit approximativement la loi  $N(0, 1)$  pour  $n$  grand.

7. Construire un intervalle de confiance de niveau 95% pour  $p$ .

Indication: On pourra utiliser le fait que  $p(1-p) \leq \frac{1}{4}$ , pour tout  $p \in ]0, 1[$ . On rappelle également que  $F_Z(1, 96) = 0,975$  où  $F_Z$  désigne la fonction de répartition d'une loi  $N(0, 1)$ .

## Evaluation 1

Durée 45 minutes

### Exercice 1 (4,5 points)

Un étudiant de la PO IC3 va en cours avec une probabilité 0.5 quand il fait beau et 0.8 quand il pleut. Sachant qu'à Toulouse, il pleut 3 jours sur 5 en novembre, et que l'étudiant était en classe le 19, quelle est la probabilité ce jour là ait été un jour de pluie? (On considèrera qu'il fait beau lorsqu'il ne pleut pas!)

### Exercice 2 (4,5 points)

Soient  $X$  et  $Y$  deux variables aléatoires indépendantes telles que:

$$P(X = 1) = P(X = -1) = \frac{1}{2} \text{ et } P(Y = 1) = P(Y = 0) = \frac{1}{2}.$$

1. On pose  $Z = XY$ . Quelle est la loi de la variable aléatoire  $Z$ ?
2. Calculer  $\mathbb{E}[Z]$  et  $\text{Var}(Z)$ .
3. Montrer que  $P(\{Y = 0\} \cap \{Z = 1\}) \neq P(Y = 0) \times P(Z = 1)$ . Que peut-on en conclure sur l'indépendance de  $Y$  et  $Z$ ?

### Exercice 3 (5 points)

Soit  $X$  une variable aléatoire de loi géométrique de paramètre  $p$  définie par

$$\mathbb{P}(X = k) = p(1 - p)^{k-1}, \text{ pour tout } k \in \mathbb{N}^*$$

Soit  $X'$  une autre variable aléatoire, indépendante de  $X$  et de loi géométrique de paramètre  $p'$ . On note par  $M$  le minimum de  $X$  et  $X'$ . On cherche à déterminer la loi de  $M$ .

1. Pour  $k \in \mathbb{N}^*$ , calculer  $\mathbb{P}(X > k)$  et  $\mathbb{P}(X' > k)$ .
2. Pour  $k \in \mathbb{N}^*$ , exprimez l'événement  $\{M > k\}$  en fonction des variables aléatoires  $X$  et  $X'$ . Montrer ensuite que

$$\mathbb{P}(M > k) = (1 - (p + p' - pp'))^k.$$

3. Quelle est la loi de  $M$ ? (Indication: on pourra montrer que si  $Y_1$  et  $Y_2$  sont deux variables aléatoires à valeurs dans  $\mathbb{N}$  telles que  $P(Y_1 > k) = P(Y_2 > k)$  pour tout  $k \in \mathbb{N}$ , alors  $Y_1$  et  $Y_2$  ont même loi).

*Tous les documents, calculatrices et téléphones portables sont interdits. Le barème sur 14 est approximatif.*

## Evaluation 2

Durée 1h30

*Tous les documents, calculatrices et téléphones portables sont interdits.  
Le barème est approximatif.*

### Quelques rappels utiles :

- Densité de la loi uniforme  $\mathcal{U}([0, 1])$ :

$$f_U(x) = \begin{cases} 1 & \text{si } x \in [0, 1], \\ 0 & \text{si } x \notin [0, 1]. \end{cases}$$

- Fonction de répartition d'une variable aléatoire  $X$  :  $F_X(x) = \mathbb{P}(X \leq x)$ .
- Quantile d'une variable aléatoire  $Z$  suivant la loi  $\mathcal{N}(0, 1)$  :  $F_Z(1, 96) = 0, 975$ .

### Exercice 1 (7,5 points)

Soit  $U$  une variable aléatoire de loi uniforme  $\mathcal{U}([0, 1])$  et soit  $p \in ]0, 1[$ .

1. Déterminer la loi de la variable aléatoire

$$X = \begin{cases} 1 & \text{si } U \leq p, \\ 0 & \text{si } U > p. \end{cases}$$

2. Soit  $U_1, \dots, U_n$  des variables indépendantes de même loi que  $U$ .

- (a) Soit  $Y$  le nombre de variables  $U_i, 1 \leq i \leq n$  qui sont inférieures à  $p$ , i.e.

$$Y = |\{i : U_i \leq p\}|,$$

où  $|\cdot|$  dénote le cardinal. Quelle est la loi de  $Y$ ? Justifier.

- (b) On considère les variables aléatoires  $V = \max(U_1, \dots, U_n)$  et  $W = \min(U_1, \dots, U_n)$ .  
Calculer  $\mathbb{P}(V \leq p)$  et  $\mathbb{P}(W \leq p)$ .

3. Soit  $\lambda > 0$  et la variable aléatoire  $T = -\frac{\log(U)}{\lambda}$ .

Calculer la fonction de répartition  $F_T$  et la densité  $f_T$  de  $T$ . En déduire la loi de  $T$ .

**Exercice 2** (7 points)

Soit le couple de variables aléatoires  $(X, Y)$  de densité jointe

$$f(x, y) = \begin{cases} \frac{1}{2\sqrt{xy}} & \text{si } (x, y) \in D, \\ 0 & \text{si } (x, y) \notin D, \end{cases}$$

où  $D = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq y < 1\}$ .

1. Montrer que  $f$  est bien une densité de probabilité sur  $\mathbb{R}^2$ .
2. Montrer que la densité marginale de  $X$  est donnée par la formule

$$f_X(x) = \begin{cases} \frac{1}{\sqrt{x}} - 1 & \text{si } x \in ]0, 1[, \\ 0 & \text{si } x \notin ]0, 1[. \end{cases}$$

Calculer la densité marginale  $f_Y$  de  $Y$  et montrer que  $Y$  suit la loi uniforme  $\mathcal{U}([0, 1])$ .

3. Les variables  $X$  et  $Y$  sont-elles indépendantes?
4. Calculer l'espérance et la variance de  $X$  et de  $Y$ .

**Exercice 3** (7 points)

On s'intéresse à la concentration  $\mu$  d'un adjuvant dans un certain type de béton. On réalise pour cela  $n$  mesures  $Y_1, \dots, Y_n$ , ces mesures étant supposées indépendantes et de même loi. Par ailleurs, pour tout  $i \in \{1, \dots, n\}$ , on suppose que  $Y_i = X_i + \epsilon_i$ , où

$$X_i \sim \mathcal{N}(\mu, \sigma_1^2) \text{ et } \epsilon_i \sim \mathcal{N}(0, \sigma_2^2),$$

$X_i$  et  $\epsilon_i$  étant elle-mêmes indépendantes. Les variables  $X_i$  modélisent les fluctuations d'adjuvant d'une mesure à l'autre, les variables  $\epsilon_i$  représentant quand à elles des erreurs liées à l'instabilité du produit chimique permettant la mise en évidence de l'adjuvant. L'objectif est d'estimer  $\mu$ , les termes  $\sigma_1^2$  et  $\sigma_2^2$  étant supposés connus.

1. Quelle est la loi de  $Y_1$ ?
2. Soit  $\bar{Y}_n = n^{-1} \sum_{i=1}^n Y_i$ . Calculer  $\mathbb{E}[\bar{Y}_n]$  et  $\text{Var}(\bar{Y}_n)$ . Donner la loi de  $\bar{Y}_n$ .
3. Construire un intervalle de confiance à 95% pour la moyenne  $\mu$ . Commenter la contribution des erreurs de mesure  $(\epsilon_i)_{i=1, \dots, n}$  dans cet intervalle.
4. On donne  $\sigma_1^2 = \sigma_2^2 = 1/2$ . A l'issue de 100 mesures, on obtient l'estimation  $\bar{y}_n = 0.3g/m^3$ . Calculer numériquement l'intervalle de confiance précédent.
5. Enoncer le théorème centrale limite et donner la loi asymptotique de  $\bar{Y}_n$  quand les  $X_i$  et  $\epsilon_i$  ne sont pas de loi normale.

Probabilités et Statistiques  
Préorientation IC 3ème année

Vendredi 23 octobre 2009

**Contrôle continu 1**  
**Durée : 45 minutes**

Tous les documents, calculatrices et téléphones portables sont interdits.  
Le barème sur 15 est approximatif.

**Exercice 1** (3,5 points)

Soient  $A$  et  $B$  deux événements d'un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$ .  
Démontrez que  $A$  et  $B$  sont indépendants si et seulement si  $A^c$  et  $B$  le sont aussi.

**Exercice 2** (3,5 points)

Soient  $A$  et  $B$  des événements d'un espace de probabilité  $(\Omega, \mathcal{A}, \mathbb{P})$  tels que  $\mathbb{P}(A) = 1/4$   
et  $\mathbb{P}(A \cup B) = 1/3$ . Calculez  $\mathbb{P}(B)$  dans les cas suivants :

- 1)  $A$  et  $B$  sont incompatibles (ou disjoints).
- 2)  $A$  et  $B$  sont indépendants.

**Exercice 3** (4 points)

La production de téléphones portables d'une marque bien connue est assurée par 3 usines  $U_1$ ,  $U_2$  et  $U_3$  qui fabriquent respectivement 30%, 30% et 40% du total. Les proportions des téléphones défectueux sont respectivement de 4%, 3% et 2%.

- 1) Quelle est la probabilité qu'un téléphone choisi au hasard soit défectueux ?
- 2) Quelle est la probabilité qu'un téléphone choisi au hasard et qui est constaté défectueux provienne de l'usine  $U_1$  ? de l'usine  $U_2$  ? de l'usine  $U_3$  ?

**Exercice 4** (4 points)

On considère une variable aléatoire  $X$  suivant la loi géométrique sur  $\mathbb{N}_* = \{1, 2, 3, \dots\}$  de paramètre  $p \in ]0, 1[$  :

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

- 1) On note  $p_k = \mathbb{P}(X = k)$ . Vérifiez que la suite  $(p_k)_{k \in \mathbb{N}_*}$  définit bien une loi de probabilité sur  $\mathbb{N}_*$ .
- 2) Calculez la probabilité  $\mathbb{P}(X > n)$  pour tout  $n \in \mathbb{N}$ .

Probabilités et Statistiques  
Préorientation IC 3ème année

Mardi 19 janvier 2010

**Examen final**  
**Durée: 1h30**

Documents, calculatrices et téléphones portables sont interdits.

Chacune de vos réponses devra être soigneusement justifiée.

Le barème sur 20 est approximatif.

**Quelques rappels utiles :**

- Fonction de répartition d'une variable aléatoire  $X$  :  $F_X(x) = \mathbb{P}(X \leq x)$ .
- Quantiles d'une variable aléatoire  $Z$  suivant la loi  $\mathcal{N}(0, 1)$  :  $F_Z(1, 645) = 0,95$  et  $F_Z(1, 96) = 0,975$ .
- Moyenne empirique d'un  $N$ -échantillon  $(X_1, X_2, \dots, X_N)$  :  $\bar{X}_N = \frac{1}{N} \sum_{i=1}^N X_i$ .
- Somme des  $N$  premiers carrés :  $\sum_{k=1}^N k^2 = \frac{N(N+1)(2N+1)}{6}$ .
- Quelques valeurs numériques :  $\frac{\sqrt{10}}{1,645} = 1,922$  et  $\sqrt{2} = 1,414$ .

**Exercice 1 Tirage de boules (5 pts)**

Une urne contient  $N$  boules numérotées de 1 à  $N$ . On tire deux fois de suite une boule dans l'urne avec remise après le premier tirage. Les deux tirages sont supposés indépendants. On note  $X_1$  la variable aléatoire correspondant au résultat du premier tirage et  $X_2$  celle associée au second tirage. On note enfin  $Y = \max\{X_1, X_2\}$  la valeur maximale obtenue après les deux tirages.

- 1 - Calculez la fonction de répartition de  $Y$ .
- 2 - Déduisez-en que pour tout  $k \in \{1, \dots, N\}$ , on a  $\mathbb{P}(Y = k) = \frac{2k-1}{N^2}$ .
- 3 - Calculez l'espérance de  $Y$ .

**Exercice 2 Chez Arthus-Bertrand (6 pts)**

Un joaillier vend des bagues contenant en moyenne 20 grammes d'or. Un riche entrepreneur décide d'en acheter 10. Si l'on note  $X_i$  la variable aléatoire correspondant à la masse d'or contenue dans la  $i$ -ème bague, alors les variables aléatoires  $X_1, X_2, \dots, X_{10}$  sont indépendantes et de même loi (quelconque), et d'espérance  $m = 20$  grammes. Enfin, on note  $\sigma^2 = \text{Var}(X_i)$  la variance des  $X_i$  et  $U = \sum_{i=1}^{10} X_i$  la variable aléatoire représentant la masse totale d'or contenue dans les 10 bagues.

- 1 - Calculez l'espérance ainsi que la variance de  $U$ .
- 2 - Rappelez l'inégalité de Chebyshev.
- 3 - En appliquant cette inégalité à la variable aléatoire  $U$ , déterminez numériquement  $\sigma_{\max}$ , la valeur maximale de  $\sigma$  pour laquelle l'inégalité suivante est satisfaite :

$$\mathbb{P}(190 \leq U \leq 210) \geq 0,95.$$

Quelle est votre interprétation de cette inégalité ?

4 - On suppose de plus que les  $X_i$  suivent la loi  $\mathcal{N}(20, \sigma^2)$ . Quelle est la loi de  $U$  ? Déterminez alors la valeur de  $\sigma$  pour que l'acheteur ait exactement une probabilité de 95% de posséder une masse totale d'or inférieure à 210 grammes.

### Exercice 3 Loi Kappa (9 pts)

Soit  $X$  la variable aléatoire réelle de densité  $f_X$  définie par:

$$f_X(x) = \begin{cases} C x e^{-\kappa x} & \text{si } x \geq 0; \\ 0 & \text{si } x < 0; \end{cases}$$

où  $\kappa$  est un paramètre inconnu strictement positif.

1 - Déterminez en fonction de  $\kappa$  la valeur de la constante  $C$  pour que  $f_X$  soit effectivement une densité de probabilité sur  $\mathbb{R}$ .

2 - Montrez que l'espérance de la variable aléatoire  $X$  est égale à  $\frac{2}{\kappa}$ .

3 - Montrez que sa variance est égale à  $\frac{2}{\kappa^2}$ .

4 - Calculez la probabilité que  $X$  soit supérieure à son espérance.

5 - On souhaite à présent estimer le paramètre inconnu  $\kappa$ . Pour cela, on note  $(X_1, X_2, \dots, X_N)$  un  $N$ -échantillon issu de la variable aléatoire  $X$ , et

$$D_N = \sqrt{\frac{N}{2}} (\kappa \bar{X}_N - 2).$$

Montrez que lorsque  $N$  est suffisamment grand, la variable aléatoire  $D_N$  suit approximativement la loi  $\mathcal{N}(0, 1)$ .

6 - Déduisez-en un intervalle de confiance à 95% pour le paramètre inconnu  $\kappa$ .

7 - Application numérique. Après une étude sur cet échantillon, on obtient l'estimation ponctuelle suivante:  $\bar{x}_N = \sqrt{2}$ .

a - Pour  $N = 100$ , calculez numériquement l'intervalle de confiance précédent.

b - Même question avec  $N = 10\,000$ . Quel phénomène observez-vous et pourquoi ?